

Berzsényi Zoltán

**Növénytermesztési kísérletek
tervezése és értékelése**

A könyvet a Földművelésügyi Minisztérium Intézményközi Tankönyvkiadási Szakértő Bizottsága az egyetemek és főiskolák agrártudományi karain kiegészítő tankönyvként javasolja.

Berzsenyi Zoltán

Növénytermesztési kísérletek tervezése és értékelése



AGROINFORM
Budapest
2016

A könyvet *Sváb János* (1922–1986) és *Győrffy Béla* (1928–2002) emlékére ajánlom.

Támogatók:



Magyar Tudományos Akadémia
MTA ATK Mezőgazdasági Intézet, Martonvásár
Nyugat-Magyarországi Egyetem, Mosonmagyaróvár
Kaposvári Egyetem Agrár- és Környezettudományi Kar, Kaposvár

Lektorálta:

Dr. Józsa Sándor
a biológiai tudomány kandidátusa

© Berzsényi Zoltán

© AGROINFORM

Kiadta:

AGROINFORM
Budapest

Felelős kiadó:

Bolyki Etelka
ügyvezető igazgató

Szerkesztette:

Sándor Anna
Agroinform Stúdió

ISBN 978-963-502-954-9



AGROINFORM Kiadó és Nyomda Kft.
Felelős vezető: Stekler Mária
1149 Budapest, Angol u. 34.
www.agroinform.hu
2015/37

Tartalomjegyzék

Szerző	XI
Bevezetés	1
1. Hazai kísérletezés és kísérletértékelés rövid története	5
2. Bevezetés a kísérletes kutatásba	8
2.1. A kutatás jellemzői és különböző kutatási megközelítések	8
2.2. A kísérletek különböző típusai	9
2.3. A növény-környezet interakciók	10
2.4. Növénytermesztési kísérletek sajátosságai	11
2.5. A kísérletek tervezésének alapszabályai	12
2.6. A kísérleti adatok statisztikai analízisének célja	14
3. A szántóföldi kísérletezés alapelvei és alapfogalmai	16
4. Kísérletek tervezése	29
4.1. Gondolatok megfogalmazása	29
4.2. Célok meghatározása	30
4.3. A populáció és a minta meghatározása	31
4.4. Hipotézisek megfogalmazása	34
4.5. Hipotézisvizsgálat	34
4.6. Várható kezeléskülönbségek	36
4.7. Változók	36
4.8. A kezelések kiválasztása	37
4.9. Ismétlés	38
4.10. Blokk-képzés	41
4.11. Randomizáció	43
4.12. Kovariánsok	44
5. Szántóföldi kísérleti elrendezések és a kísérletek szerkezete	46
5.1. Egytényezős kísérletek	46
5.1.1. Teljesen véletlen elrendezés (CRD)	47
5.1.2. Véletlen blokkelrendezés (RCB, RBD)	48
5.1.3. Latin négyzet elrendezés (LS)	49
5.1.4. Csoportosított blokkelrendezés	54
5.1.5. Nem-teljes blokk-elrendezés	55
5.2. Faktoriális kísérletek	57
5.2.1. Kéttényezős véletlen blokkelrendezés	57
5.2.2. Osztott parcellás (split-plot) elrendezés	59
5.2.3. Sávós (strip-plot vagy criss-cross) elrendezés	60
5.2.4. Három- és többtényezős kísérletek	61
5.3. Kísérletsorozatok	69
5.4. Tartamkísérletek	72
5.5. Üzemi (demonstrációs) kísérletek	73

6. Adatbevitel és adatfeltárás	75
6.1. Adattípusok	75
6.2. A változók típusai mérési skála szerint	76
6.3. Adatellenőrzés és exploratív adatelemzés	79
6.4. Számítógépek használata a kísérletek tervezésében és értékelésében ..	79
6.5. Kísérletek számítógépes tervezése	82
7. Leíró statisztika	87
7.1. Helyzet-mutatók	87
7.2. Szóródás mutatók	90
7.3. A minta megbízhatóságának mérése: az átlag (középérték) hibája	94
7.4. Ferdeség és csúcsosság	100
7.5. A leíró statisztikai mutatók kiszámítása GenStat programmal	100
7.6. Grafikus ábrázolás	102
7.7. Ábrák szerkesztése GenStat programmal	109
7.8. Táblázatok	109
8. Egymintás és kétmintás paraméteres próbák	111
8.1. Bevezetés	111
8.2. Egymintás t-próba	112
8.3. Kétmintás t-próba	114
8.3.1. Kétmintás t-próba, feltételezve egyenlő populáció varianciát	115
8.3.2. Kétmintás t-próba, feltételezve nem egyenlő populáció varianciát	117
8.4. Páros t-próba	118
8.5. Két minta varianciája közötti különbség analízise: F-próba	120
8.6. Két-mintás kísérletek tervezése és analízise	122
8.7. A Student t-próba kiszámítása a GenStat felhasználásával	123
9. A varianciaanalízis (ANOVA) alapelvei	126
9.1. A varianciaanalízis történeti háttere	126
9.2. A varianciatáblázat szerkezete	129
9.3. Az ANOVA feltételei	131
9.4. A varianciaanalízis feltételeinek ellenőrzése GenStat programmal	133
9.5. Összehasonlítás a kezelésközégek között	135
9.5.1. Többszörös terjedelem vizsgálatok	135
9.5.2. A különbség standard hibája	135
9.5.3. Legkisebb szignifikáns különbség (SzD, LSD)	137
9.5.4. Tukey-féle szignifikáns különbség (HSD)	138
9.5.5. Student –Newman – Keuls többszörös terjedelem próba	139
9.5.6. Dunnnett próba	141
9.5.7. Számítógépes értékelés GenStat programmal	142
9.5.8. Duncan többszörös terjedelem próba (DMRT)	145
9.5.9. Számítógépes értékelés GenStat programmal	151
10. Egytényezős kísérletek varianciaanalízise	152
10.1. Teljesen véletlen elrendezés (CRD)	152
10.1.1. Randomizáció	152
10.1.2. Egyenlő ismétlésszám	157
10.1.3. Kísérlet leírása	157
10.1.4. Varianciaanalízis lépései	157

10.1.5. Varianciaanalízis GenStat programmal	163
10.1.6. Nem egyenlő ismétlésszám	165
10.1.7. Kísérlet leírása	166
10.1.8. Varianciaanalízis lépései	167
10.1.9. Varianciaanalízis GenStat programmal	170
10.2. Véletlen blokkalrendezés (RCB)	172
10.2.1. Kísérlet leírása	172
10.2.2. Varianciaanalízis lépései	173
10.2.3. Varianciaanalízis GenStat programmal	177
10.3. Latin négyzet (LS) elrendezés	180
10.3.1. Randomizáció	180
10.3.2. Kísérlet leírása	181
10.3.3. Varianciaanalízis lépései	182
10.3.4. Varianciaanalízis GenStat programmal	186
10.4. Csoportosított blokk-elrendezés	190
10.4.1. Randomizáció és elrendezés	191
10.4.2. A kísérlet leírása	192
10.4.3. Varianciaanalízis lépései	193
10.4.4. Varianciaanalízis GenStat programmal	199
10.5. Latintégla elrendezés	205
10.5.1. Kiegyensúlyozott latintégla elrendezés	206
10.5.2. Részben kiegyensúlyozott latintégla elrendezés	217
10.6. Kezelés struktúra	222
10.6.1. Kísérlet leírása	224
10.6.2. A varianciaanalízis lépései	224
10.6.3. Varianciaanalízis GenStat programmal	227
10.6.4. Faktoriális kezelés struktúra	228
11. Kéttényezős kísérletek varianciaanalízise	231
11.1 Kéttényezős teljesen véletlen elrendezés (CRD)	238
11.1.1. Kísérlet leírása	238
11.1.2. Varianciaanalízis lépései	238
11.1.3. Varianciaanalízis GenStat programmal	243
11.2. Fő hatások és interakciók	246
11.2.1. Kéttényezős faktoriális kísérlet varianciaanalízise	248
11.3. Kéttényezős véletlen blokkalrendezés	252
11.3.1. Kísérlet leírása	252
11.3.2. Varianciaanalízis lépései	254
11.3.3. Az eredmények értelmezése	258
11.3.4. Varianciaanalízis GenStat programmal	259
11.4. Osztott parcellás (split-plot) elrendezés	262
11.4.1. Randomizálás	263
11.4.2. Az osztott parcellás elrendezés varianciaanalízise	265
11.4.3. Kísérlet leírása	268
11.4.4. Varianciaanalízis lépései	269
11.4.5. Az eredmények értelmezése	275
11.4.6. Varianciaanalízis GenStat programmal	276
11.5. Sávós (strip-plot vagy criss-cross) elrendezés	280
11.5.1. Randomizáció és elrendezés	280
11.5.2. Kísérlet leírása	281

11.5.3. Varianciaanalízis lépései	282
11.5.4. Az eredmények értelmezése	289
11.5.5. Varianciaanalízise GenStat programmal	290
11.6. Hierarchikus vagy beágyazott (nested) kísérletek.	293
11.7. Kéttényezős kísérlet ismétlés nélkül	298
12. Háromtényezős kísérletek varianciaanalízise	301
12.1. Háromtényezős teljesen véletlen (CRD) elrendezés	303
12.1.1. Kísérlet leírása	303
12.1.2. Varianciaanalízis GenStat programmal	304
12.2. Kétszeresen osztott parcellás (split-split-plot) elrendezés	307
12.2.1. Randomizáció és elrendezés	307
12.2.2. Kísérlet leírása	307
12.2.3. Varianciaanalízis lépései	309
12.2.4. Varianciaanalízis GenStat programmal	322
12.3. 2 ⁿ Faktoriális kísérletek	326
12.3.1. Kísérlet leírása	328
12.3.2. Varianciaanalízis GenStat programmal	330
12.3.3. A kísérlet eredményeinek értékelése	334
13. Időbeni mérések analízise	335
13.1. Kísérlet leírása	336
13.2. Varianciaanalízis lépései	338
13.3. Varianciaanalízis GenStat programmal	341
13.4. Kísérlet leírása	345
13.5. Az ANOVA és az REML eredményének összehasonlítása	345
14. Kísérletsorozatok kombinált varianciaanalízise	347
14.1. Kísérletsorozat: évek analízise	350
14.1.1. Kísérlet leírása	351
14.1.2. Varianciaanalízis lépései	352
14.1.3. Varianciaanalízis GenStat programmal	355
14.2. Technológiai adaptációs kísérletsorozat: különböző helyek analízise	358
14.2.1. Kísérlet leírása	359
14.2.2. Varianciaanalízis lépései	359
14.3. Technológiai adaptációs kísérletsorozat: Műtrágyázási kísérlet	363
14.3.1. Kísérlet leírása	363
14.3.2. Varianciaanalízis lépései	364
14.3.3. Kísérletsorozat számítógépes értékelése GenStat programmal	371
14.4. 2 ⁿ Faktoriális kísérletsorozat kombinált varianciaanalízise	381
14.4.1. Kísérlet leírása	381
14.4.2. A variancia homogenitásának vizsgálata	382
14.4.3. Előzetes analízis	383
14.4.4. Részletes analízis	384
14.4.5. A statisztikai analízis eredményének értelmezése	386
14.5. Ültetvény kísérletsorozatok varianciaanalízise	387
14.5.1. Komplet kétirányú táblázatok	388
14.5.2. Kísérlet leírása	389
14.5.3. Inkomplet kétirányú táblázatok	391
14.5.4. Kísérlet leírása	392

14.5.5. Regresszió analízis a komplett kétirányú táblázatokkal	395
14.5.6. Regresszió analízis inkomplett kétirányú táblázatokkal.	398
15. Üzemi kísérletek	401
15.1. Technológia-fejlesztési kísérletek	402
15.1.1. A tesztelő hely kiválasztása	402
15.1.2. A kísérletek elrendezése	402
15.1.3. Adatgyűjtés	403
15.1.4. Adatelemzés	403
15.1.5. Varianciaanalízis a GenStat programmal	408
15.2. Technológia összehasonlítási kísérletek rendszere	412
15.2.1. Kísérleti elrendezés.	412
15.2.2. Adatgyűjtés	414
15.2.3. Terméshiány elemzés	414
15.2.4. Varianciaanalízis GenStat programmal	421
15.3. Termesztési rendszerek vizsgálata (FSR)	427
15.3.1. Elemi üzemi FSR kísérletek	427
15.3.2. Üzemi FSR kísérlet kiegészítő kezelésekkel	430
15.3.3. Üzemi FSR kísérlet két komponensű javított gyakorlattal	434
15.3.4. Számítógépes értékelés GenStat programmal	440
16. Tartamkísérletek	442
16.1. A tartamkísérletek jelentősége és sajátosságai	442
16.2. A tartamkísérletek értékelésének módszerei	444
16.2.1. A tartamkísérletek standard és javított tervezése.	444
16.2.2. Összegező statisztikai mutatók.	446
16.2.3. Évenkénti és kombinált varianciaanalízis	447
16.2.4. Kumulatív terméselemzés	448
16.2.5. Stabilitásanalízis	450
16.2.6. A tartamkísérletek elemzése vegyes (mixed) modellekkel.	452
16.3. A martonvásári tartamkísérletek értékelésének eredményei	453
16.3.1. Kukorica és búza vetésforgó kísérlet eredményei	453
16.3.2. Istállótrágya és műtrágya hatása a kukorica termésére tartamkísérletekben	457
16.3.3. Kukorica hibridek N-műtrágya reakciójának vizsgálata tartam- kísérletekben	460
16.3.4. Növénytermesztési tényezők hatása a kukorica termésére és termésstabilitására	462
17. Nem-paraméteres próbák	468
17.1. Bevezetés	468
17.2. Két független minta közötti különbség analízise	468
17.3. Páros, nem-független minták közötti különbségek analízise	472
17.3.1. Az előjel teszt.	472
17.3.2. Wilcoxon előjelteszt	474
17.4. Nem-paraméteres többmintás analízis	476
17.4.1. Kruskal-Wallis teszt	476
17.4.2. Nem-paraméteres többszörös összehasonlítási teszt	479
17.5. Randomizált blokk elrendezésű kísérlet nem-paraméteres analízise	480
17.6. Paraméteres és nem-paraméteres módszerek összehasonlítása.	483

18. Korrelációs számítás és regresszió-analízis	485
18.1. Korrelációs számítás	487
18.1.1. Bevezetés a korreláció különböző típusaihoz	487
18.1.2. Pearson-féle korrelációs koefficiens (r)	488
18.1.3. Nem-paraméteres korrelációs számítás: Spearman rangkorreláció ..	495
18.2. Lineáris regresszió-analízis	498
18.2.1. A kétváltozós lineáris modell	498
18.2.2. A regressziós egyenlet meghatározása: a legkisebb négyzetek módszere	498
18.2.3. A regressziós egyenes szignifikanciájának tesztelése	502
18.2.4. A regressziós egyenes illeszkedésének tesztelése: determinációs koefficiens (r^2)	505
18.2.5. A regressziós egyenes megbízhatóságának tesztelése	506
18.2.6. Az előrejelzések (predikciók) megbízhatóságának tesztelése ..	506
18.2.7. A gyomboritottság és a kukorica szemtermése közötti lineáris összefüggés	509
18.2.8. A növényszám és a levélterület index (LAI) közötti lineáris összefüggés	510
18.2.9. Lineáris regresszióanalízis GenStat programmal	511
18.3. Nem-lineáris modellek	514
18.3.1. Lineáris és nem lineáris modellek	514
18.3.2. Kukorica hibridek műtrágya-reakciójának vizsgálata különböző függvényekkel	518
18.3.3. Aszimptotikus növekedési függvények	523
18.3.4. A vetésidő hatása a kukorica hibridek növekedésére	527
18.3.5. Nem-lineáris regresszióanalízis GenStat programmal	529
18.4. Többszörös regresszió-analízis	530
18.4.1. Többszörös lineáris regressziós modell	530
18.4.2. A többszörös lineáris regressziós modell szignifikanciájának tesztelése	534
18.4.3. A többszörös regressziós modell illeszkedésének tesztelése ..	535
18.4.4. A többszörös lineáris regressziós modell feltételeinek tesztelése	536
18.4.5. Interakció a prediktor változók között a többszörös regresszió-analízisben	537
18.4.6. Többszörös lineáris regresszió a GenStat programmal	538
18.4.7. Az eredmények értelmezése	539
18.4.8. Változók kiválasztása a regresszió-analízisben	540
18.4.9. Kukorica hibridek növekedésének vizsgálata többszörös regresszióanalízissel	541
18.4.10. A búzatermés és a növekedési mutatók regresszióanalízise ..	543
18.4.11. Path-analízis	544
18.4.12. A növekedési mutatók közötti összefüggések path-diagramja ..	548
18.4.13. Többszörös regresszió analízis GenStat programmal	550
Felhasznált irodalom	551
Melléklet	557
Statisztikai táblázatok	558
Tárgymutató	584

Szerző



Berzsenyi Zoltán, növénytermesztő kutató, az MTA doktora, 1982-től dolgozott az MTA Mezőgazdasági Kutatóintézetében, Martonvásáron. 1984 és 2000 között a Kukorica Agrotechnikai Osztály, 2000 és 2010 között a Növénytermesztési Osztály tudományos osztályvezetője, ezt követően kutató professor emeritus. 2002 és 2013 között egyetemi tanár, 2013-tól professor emeritus a Kaposvári Egyetemen. Szakterületének kiemelkedő, nemzetközileg is elismert kutatója. Agrármérnöki oklevelet a Keszthelyi Agrártudományi Főiskolán szerzett 1967-ben. 1972-ben doktorált a Gödöllői Agrártudományi Egyetemen. Kandidátusi értekezését Sváb János módszertani irányításával készítette el a „Kukoricatermesztés biometriai elemzése” címmel és védte meg az akadémián 1980-ban. Az MTA

doktora címet 1994-ben kapta meg. Akadémiai doktori értekezésének címe: „Növekedésanalízis a kukoricatermesztési kutatásokban”.

Kutató munkája az intézetben a szabadföldi kukoricatermesztési kísérletekhez, különösképpen Győrffy Béla növénytermesztési tartamkísérleteihez és szakmai munkásságához kapcsolódott. A növénytermesztési kutatásokba egy teljesen új agrotechnikai irányzatot vezetett be, a szabadföldi kísérletek több paraméteres növekedésanalízisét. Jelentősek eredményei a kutatásmódszertan területén, a biometriában, mind a többváltozós módszerek alkalmazásában, mind a tartamkísérletek értékelési módszereinek fejlesztésében.

A növénytermesztési kutatásban és oktatásban az új növénytermesztés-fiziológiai (crop physiology) irányzatot képviseli. Polifaktoriális kísérletekben kutatásainak célja a genotípus, a termesztéstechnológia és a környezeti tényezők közötti pozitív interakciók kvantifikálása és szabályozása révén a termésmenvelés további útjainak feltárása. A növekedésanalízis kutatások bevezetésével és a kísérletértékelés terén végzett módszertani kutatásaival iskolateremtő a növénytermesztésben.

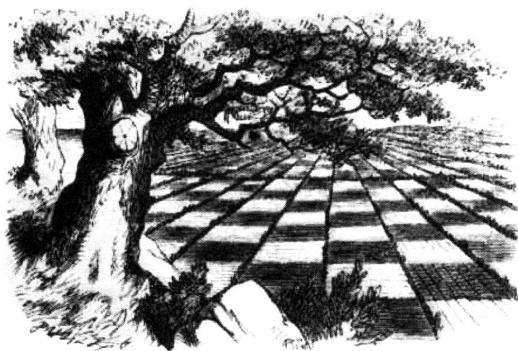
Részt vesz a PhD képzésben. 2013-ig törzstagja a Szent István Egyetem Növénytudományok Doktori Iskolájának. A Nyugat-Magyarországi Egyetemen, a „Precíziós növénytermesztés módszertana” c. doktori iskolában 15 éven keresztül oktatta a PhD hallgatóknak a „Szabadföldi kísérletek beállításának és biometriai értékelésének módszerei” c. tantárgyat. Oktatója a Pannon Egyetem Fesztetics Doktori Iskolájának.

Címzetes egyetemi tanár a Pannon Agrártudományi Egyetemen 1991-től; egyetemi magántanár a Debreceni Egyetemen 1998-tól. Az MTA Növénytermesztési Bizottság tagja 1987-től, az MTA Biomatematikai és Biometriai Interdiszciplináris Bizottság tagja 1997 és 2011 között. 1998 és 2001 között Széchenyi professzori ösztöndíjban részesült. 2012-ben az MTA Talajtani, Vizgazdálkodási és Növénytermesztési Bizottsága kiemelkedő szakmai munkásságának elismeréseként a Surányi János Emlékérmet adományozta részére. A szakterület magyar és idegen nyelvű tudományos folyóirataiban 106, szerkesztett konferencia kiadványokban 45 publikációja jelent meg. Szerzője 3 könyvnek, 21 könyvfejezetnek. Több mint 300 SCI független hivatkozást tart számon.

Bevezetés

Milyen lenne a világ mezőgazdasági kutatás nélkül? Kétségtelen, hogy a mezőgazdasági kutatás mélységesen hatott az élet minőségére és Bolygónk fenntarthatóságára. Az elmúlt évszázadban, a varianciaanalízis (ANOVA) eljárásai alapvetők voltak a mezőgazdasági kutatásban, amely a következő kérdéshez vezet: Milyen lenne a világ ANOVA nélkül? Valószínűleg a világ hasonló felfedezésekhez vezetett volna, azonban sokkal több évet és több erőfeszítést vett volna igénybe, és több tévedést hordozott volna magában.

Egy olyan világban, amely imperfekt, imprecíz és bizonytalan, a biometriai analízis és tervezés módszereket nyújt, hogy elősegítse a kísérletek eredményeinek interpretálását. A biometria fő célja, hogy előírjon tudományos standardokat a kísérleti megfigyelésekre, amelyek ki vannak téve örökletes biológiai, továbbá fizikai, időbeni és térbeni variációnak.



1. kép. „... és milyen különös ország volt ... A föld ... fel volt osztva négyzetekre ...”
Lawis Carroll: 1871. *The Adventures of Alice Through the Looking Glass* (Alice Csodaországban).
2. Chapter. Macmillan, U.K. p. 30.

Lawis Carroll (1871) könyvének és a fenti idézetnek (1. kép) megjelenése után közel másfél évszázad elteltével, bárki, aki ellátogat egy növénytermesztési kutatóintézetbe – és még inkább, aki elrepül felette – látni fogja, mint Alice, hogy a föld fel van osztva kis négyzetekre, vagy inkább négyszögekre (2. kép), melyeknek mindegyike egy-egy kezelés.

A különböző kezelések (akár műtrágya dózisosk, akár különböző fajták) összehasonlítása szükségessé teszi, hogy a kezelések a lehető legközelebb kerüljenek egymáshoz. Az is igaz, hogy a kísérletező beállíthatna minden kezelést különböző táblákon. Ekkor jogosan állíthatná, hogy úgy alkalmazta a kezeléseket, ahogy egy gazda tenné, azonban ha egy tábla volna minden kezelés, komoly nehézség lenne az interpretációban. A táblák ugyanis különböznek számos okból. Nem tudhatnánk, hogy a megfi-

gyelt különbségek egy kísérlet táblái között mindenképpen meglennének-e vagy azokat részben az alkalmazott kezeléseknél lehetne tulajdonítani.

Ahhoz, hogy elkerüljük ezt a hibát, a kísérletező a másik extrém irányba mehetne. Kiválaszthatna egy kis területet, amely egységesnek (uniform) tűnik és ezután feloszthatná azt kisebb részekre, ún. parcellákra. Ezután mindegyik kezelést a kiválasztott parcellákon állítaná be. Ekkor ki lenne téve egy másik kritikának, nevezetesen, hogy a parcellái túl kicsik ahhoz, hogy azokon a kezeléseket megfelelően alkalmazzuk. Következésképpen, bizonyos kompromisszum szükséges a két extrém parcellaméret között.



2. kép. 1843-ban beállított és azóta folyamatosan fenntartott Broadbalk őszi búza trágyázási kísérlet a Rothamsted Kutató Intézetben (UK)

Számos specifikusság van, amely a kísérletezésben alapvetőnek tekinthető. Egyet már említettünk, nevezetesen a megfelelő parcellaméret használatát. Egy másik specifikusság az „ismétlés”, azaz a több parcella használata minden egyes kezelésre. Ha csak egy parcellát használnánk minden kezelésre, nem volna módszer annak eldöntésére, mennyire pontosan állítottuk be a kísérletet. Ha több parcella van, akkor a köztük levő különbség megmutatja, mennyi nem szabályozott természetes variáció van a mérésekben. Növénytermesztési vonatkozásban ez különösen fontos, gondoljunk pl. az évjárat hatásra.

Egy kísérlet további fontos jellemzője a randomizálás, amely minden esetben szükséges objektivitási okból. Ha a kísérletező kijelentheti, hogy a parcellákat a kezelésekre véletlen számok vagy kártyakeverés, stb. alapján jelölte ki, biztosak lehetünk abban, hogy a legjobb parcellák nem szándékosan lettek kiválasztva bizonyos kedvezményezett kezelésekre. Van azonban egy fontosabb ok a randomizálásra, nevezetesen a pontosság meghatározása. Ahol ismétlés van, azaz több parcella minden kezelésre, ezeknek a parcelláknak az összehasonlítása megmutatja, mennyi variációt vezettünk be a talajban, a növényben levő különbségekkel, a mérések pontatlanságával stb.

Egy utolsó szempont a helyi variáció szabályozására vonatkozik. Ha a kísérleti terület észlelhető szisztematikus különbségeket (pl. lejtés, talajtermékenység) mutat, akkor blokkok kialakítása javasolt. A blokkok nem szükségszerűen azonos méretűek, fontos, hogy minden blokk, amennyire lehetséges egységes legyen. A fő különbségek

a blokkok között legyenek. Ha a szisztematikus különbségek nem nyilvánvalóak és nem ismertek a terület múltjából, a parcellákat elhelyezhetjük random az egész területen belül. Ekkor csak egy blokk van, és az elrendezés teljesen randomizált.

Általánosan elfogadott, hogy a biometria a kutatás minden típusának nélkülözhetetlen módszere. A kutatás-specializáció állandóan sokszorozódó területei vezettek a rendelkezésre álló biometriai eljárások számának és diverzitásának hasonló mértékű növekedéséhez. A mezőgazdasági kutatásban különböző biometriai módszerek vannak a növényekre és állatokra, a laboratóriumi és szántóföldi kísérletekre, a genetikai és fiziológiai vagy molekuláris biológiai kutatásra. Ez a diverzitás jelzi a megfelelő biometriai eljárások rendelkezésre állását a legtöbb kutatási problémára és ugyancsak jelzi annak nehézségét, hogy megtaláljuk a legmegfelelőbb eljárást a specifikus kísérletekre. Nyilvánvalóan ez a bonyolultság fokozódik, ahogy az eljárások száma nő.

Egy korrekt biometriai eljárás kiválasztása egy adott kísérletre feltételezi a jártasságot a biometriában és a kísérlet tárgykörében egyaránt. A kettő közül az egyik alapos ismerete önmagában nem elegendő. Úgy véljük, a legjobb alternatíva, hogy kapjon a mezőgazdasági kutató (szakember) egy biometriai háttérrel ahhoz, hogy korrekten kiválaszthassa azt a biometriai eljárást, amely a legmegfelelőbb a kísérleteihez. Egy olyan szakkönyv (tankönyv) összeállítására törekedtünk, melyet a növénytermesztés (és tudományterületei) bármely területén kísérleteket végző szakember használni tud.

Először bemutatjuk az egyszerűbb és általánosabban használt biometriai eljárásokat a mezőgazdasági kutatásban, speciális hangsúllyal a szántóföldi kísérletekre. A bevezető fejezetek magukba foglalják a kísérletezés alapelveit és alapfogalmait; a szántóföldi kísérleti elrendezéseket és a kísérletek szerkezetét; a leíró statisztikai mutatókat (helyzet és szóródás mutatók); az egy- és kétmintás statisztikai mutatókat (t-próbák).

A kísérlet-értékelési módszerek gerincét a varianciaanalízis (ANOVA) képezi. Az egytényezős, kéttényezős és háromtényezős kísérletek különböző típusainak (elrendezéseinek) varianciaanalízise után a kísérletsorozatok, tartamkísérletek és üzemi kísérletek varianciaanalízise következik. A kísérletek varianciaanalízisét kétféle módon, először a hagyományos kézi számolással (zsebszámológép), lépésenként mutatjuk be, ezt követően a számítógépes értékelést vázoljuk fel a GenStat programcsomag felhasználásával. A varianciaanalízis hagyományos, kézi számoláson alapuló módszere elősegíti a speciális módszerek megértését és számítógépes program hiányában a varianciaanalízis korrekt elvégzését. A GenStat programmal történő értékelés bemutatja a kísérletek értékelésének korszerű módszerét, beleértve az adatok bevitelének és az értékelési lehetőségeknek kiválasztását.

A szakkönyv a nem-paraméteres próbák áttekintése után röviden foglalkozik a korrelációs számítással, a kétváltozós lineáris és nem-lineáris regresszióanalízissel és a többszörös regresszióanalízissel. A példák jelentős részében a világviszonylatban is legfontosabb három növényvel, a kukoricával, búzával és rizzsel végzett kísérletek adatait értékeljük. A példák magukba foglalnak ültetvény és állat kísérleteket is. Ily módon a mezőgazdasági kísérletek széles körében alkalmazhatók. Ezen kívül, minimális matematikai és biometriai elméletet és szakkifejezést használunk és a példák valós kísérletek adataiból származnak, felhasználva a martonvásári növénytermesztési kísérletek eredményeit, melyeket képekkel illusztrálunk.

A kézi számolás lépéseinél alapvetően *Gomez és Gomez (1984)*, illetve *Ireland (2010)*, továbbá *Sváb (1981)* és *Petersen (1994)* módszerét követtük. Következésképpen nem

törekedhettünk egységes jelölés- és lépésrendszer kialakítására. A könyvben tárgyalt kísérlettypusok felölelik *Sváb* (1981) által ismertetett kísérleteken túlmenően más, a gyakorlatban fontos kísérleteket is (pl. 2ⁿ faktoriális, latintégla, hierarchikus kísérletek, tartamkísérletek). Olyan kísérleti adatbázis megalkotására törekedtünk, amelyek hiteles kísérleteken, illetve felvetett kérdéseken alapulnak és vezérfonalként szolgálhatnak a kutató (kísérletező szakember, PhD hallgató) saját kísérleteinek értékeléséhez. A GenStat programot a Rothamstedi Kutató Intézet fejlesztette ki a mezőgazdasági kísérletek értékelésére (*Payne et al.*, 2003). Könnyen kezelhető statisztikai programcsomag, melyet világszerte használnak és a könyv Szerzője – több, mint 10 éves tapasztalata alapján – különösen alkalmasnak találta a növénytermesztési kísérletek értékelésére (info@vsn-intl.com; <http://www.vsn-intl.com>).

A könyv szerzője növénytermesztő kutató, akit az a nemes cél vezérelt hogy – a kísérletezésben eltöltött több, mint három évtized tapasztalatainak birtokában – elősegítse a növénytermesztési kísérletek korrekt tervezését, beállítását és értékelését. A könyv döntően a klasszikus biometriai módszereket alkalmazza (*Sváb*, 1981; *Gomez és Gomez*, 1984; *Petersen*, 1994), bár tartalmaz példákat az ún. vegyes (mixed) modell (fix és random hatások egyaránt tartalmaz) használatára is. A növénytermesztési kísérletekben elsősorban a beállított fix kezeléshatások érdeklik a szakembereket, ugyanakkor bizonyos kísérlettypusokban (kísérletsorozatok, tartamkísérletek, nem kiegyensúlyozott kísérletek) egyre nagyobb a jelentősége a vegyes modellek alkalmazásának (*Gbur et al.*, 2012).

A szakkönyv javasolható a kísérletekkel foglalkozó, illetve a kísérletek iránt érdeklődő szakembereknek, kutatóknak, szaktanácsadóknak, oktatóknak, hallgatóknak (MSC, PhD hallgatóknak), akik a kísérlet tervezés és értékelés gyakorlati módszerit kívánják elsajátítani és kevésbé kívánnak elmélyedni a kísérletértékelés elméleti (biometriai) kérdéseiben. Hiánypótló mű, amennyiben a növénytermesztési kísérletek tervezésével és értékelésével foglalkozó szakkönyv hosszabb idő óta nem hozzáférhető az oktatásban, illetve az érdeklődő szakemberek számára.

Köszönettel tartozom Győrffy Béla (1928–2002) és Balla László (1933–2014) egykori igazgatóimnak, Bedő Zoltán főigazgatónak, hogy az MTA ATK Mezőgazdasági Intézetében, Martonvásáron, biztosították az alkotó feltételeket tudományos kutató munkámhoz. Köszönöm a Növénytermesztési Osztály valamennyi munkatársának az egzakt szabadföldi kísérleti tevékenységet és kutatómunkám kitarató támogatását. Köszönettel tartozom Prof. Kuroli Gézának és Prof. Reisinger Péternek a Nyugat-Magyarországi Egyetemen (Mosonmagyaróvár), akik meghívtak a „Szántóföldi kísérletek beállításának és biometriai értékelésének módszerei” c. PhD tantárgy oktatására. Köszönöm Józsa Sándor (Pannon Egyetem) professzornak a könyv biometriai szempontból történő áttekintését, értékes észrevételeit és javaslatait. Külön köszönettel tartozom Kizmus Lajos egykori kollégámnak, a könyv szerkesztésében nyújtott értékes segítségéért. Végül köszönöm Bolyki Istvánnak, az Agroinform igazgatójának, hogy vállalta a szakkönyv megjelentetését.

2015. június

Dr. Berzsenyi Zoltán

1. Hazai kísérletezés és kísérletértékelés rövid története

A kísérletezés és a kísérletértékelés különböző szintjei és tudományos módszerei a különböző időszakokban – a kor tudományos és technológiai fejlettségével összhangban – mindenkor hatékonyan hozzájárultak az alkalmazott kutatási eredmények gyakorlati hasznosításához a növénytermesztésben.

Cserhádi Sándor (1852–1909) egész életén keresztül fő feladatául tekintette olyan híd létesítését, amely segíti az alapozó, természettudományi felfedezések gyakorlatba való átvitelét. Ez volt munkásságának vezérlő alapelve. 1896-ban Cserhádi erről így írt: „A mezőgazdasági kísérletezés éppen olyan fontos, mint a szakoktatás, mindkettővel mezőgazdasági ismeretek fejlesztésére törekszünk...”, ezért kiválóan szükséges, hogy a kísérleti ügyet mihamarabb olyan módon fejlesszük, hogy annak a gyakorlati gazda hasznát lássa.” Cserhádi Sándor korszerű és napjainkban is érvényes alapelveket fogalmazott meg a trágyázási kísérletekről az „Általános és különleges növénytermelés” (1905) c. könyvében és más tudományos dolgozataiban. A hazai növénytermesztési kísérletezés úttörőjének tekinthetjük.

Várallyay György (1900–1954) alapelvei a kísérleti eredmények gyakorlatba történő átvételére napjainkban éppúgy érvényesek, mint ahogy az Ő korában is azok voltak. Téziseit a „Növénytermelés” c. tudományos folyóirat első évfolyamának első számában, 1952-ben fejtette ki. „Természetesen a kísérleti eredmények gyakorlatba való átvételének megvannak a maga módjai és eszközei. Egyik eddig eléggé ki nem



Sváb János
(1922–1986)

használt eszköz a gyakorlati klíma- és talajvizsgálat: a gyakorlati termőhely ismeret.” „Tehát a kísérleti hely és a mezőgazdasági nagyüzem termőhely viszonyainak összevetése a legjobb eszköz és mód arra, hogy a kísérleti eredmények nagy és megfelelő területre érvényesíttessenek... A nagyüzemi termesztést üzemi feltételek között végzett, több éven át tartó, ismétléses kisparcellás kísérletekkel irányítsuk olyanformán, hogy ezek eredményeit tömeges, gyakorlati talajvizsgálatokkal, üzemi talajterképekkel érvényesítsük a mezőgazdasági nagyüzem egyes tábláira.” A kisparcellás kísérletekből adódó eredmények széles körű üzemi kísérletekkel való ellenőrzésének szükségességét *Sarkadi János* (1920–2006) már az 1950-es években hangsúlyozta.

Sváb Jánost (1922–1986) tekinthetjük a hazai biometria egyik megalapítójának és legnagyobb egyéniségének. Sváb János széles látókörű, modern szemléletű, nagy formátumú tudós volt. Évtizedek múltán világosan látjuk szakmai és emberi nagyságát. Nem mások által kitaposott utat járt, hanem úttörő volt, iskolaalapító és példakép. Tudományos meggyőződését követte, bátran és elszántan, nem karriert akart építeni, hanem

a tudományt akarta gazdagítani hazája és embertársai javára (Prékopa, 2011). Az Országos Fajtakísérleti Intézet munkatársaként került kapcsolatba a biometriával, ami egész életének kutató munkáját meghatározta. Jelentős szerepe volt abban, hogy az intézet a fajtakísérleteket új alapokra helyezte.

Az intézet országosan központi szerepe lehetővé tette, hogy a Sváb János által kidolgozott módszerek és feldolgozott anyagok sokfelé eljussanak, és közkinccsé váljanak. Az ezt követő években, 1971–75 között, az országos műtrágyázási tartamkísérletek tervezésével és értékelésével foglalkozott, Láng Géza akadémikus vezetésével, majd az Agrobotanikai Intézet biometrikusa lett. A Gödöllői Agrártudományi Egyetemen egy kutatócsoport vezetőjeként nagyüzemi táblasoros adatok feldolgozásával foglalkozott, matematikai szempontból pedig a többváltozós statisztikai analízis módszereit kezdte alkalmazni. Bebizonyította, hogy a nagyüzemi adatok is alkalmasak a tudományos eredmények elérésére, hipotézisek igazolására. Sváb János 1979-ben hirdette meg azt a tézisét, mely szerint a kutatás két iránya, a kísérletes kutatás és a táblasoros elemzés a kutatási rendszerben egységet alkotva, együtt szolgálják a növénytermesztés fejlődését. Az üzemi termesztés táblasoros elemzésével (exploratív adatelemzés) e kutatási irányzat arra keresett választ, hogy mely termesztési tényezők befolyásolják legerősebben a táblánkénti átlagterméseket, egyben a meglévő termesztéstechnológiai elemek és a termés olyan összefüggéseinek feltárására is mód nyílt, amelyre a lehetséges kombinációk sokrétűsége miatt a kísérletes kutatás nem alkalmas. Ez már a többváltozós biometriai módszerek bevezetésének időszaka volt.

Akadémiai doktori disszertációjában a többváltozós statisztikai analízis mezőgazdasági alkalmazásaival foglalkozott. Utolsó éveiben aktívan dolgozott az Agrártudományi Egyetemen, Gödöllőn és az MTA Mezőgazdasági Kutatóintézetében, Martonvásáron. Több könyve jelent meg. A kísérletek értékelésében fontos mérföldkő Sváb János: „Biometriai módszerek a kutatásban” c. könyve, melyet többször kiadták (1967, 1973, 1979). A populációgenetika alapjai (1971) könyve külföldön is megjelent. „Többváltozós módszerek a biometriában” című könyve (1979) megismerteti az olvasók tág körével ezt a gyakorlati szempontból nagyfontosságú módszertant.

Sváb János által képviselt tudományos iskolában és általa szervezett posztgraduális tudományos továbbképző kurzusokon több évtizeden keresztül kutatók generációi sajátították el a szántóföldi kísérletezésnek és a kísérletek biometriai értékelésének módszereit. Sváb János a „Biometriai módszerek a kutatásban” c. könyvében az elődökről így emlékezik meg: „Erkölcsei kötelezettséget jelent annak az útnak a folytatása, amelyet a mezőgazdaság területén a korszerű kísérleti módszerek hazai ismertetésével néhány lelkes úttörő, így dr. Telegdy-Kováts László, Horváth Alajos, dr. Berzsényi-Janosits László és dr. Terts István elindított.”

Sváb János után a biometria oktatása és a kísérletértékelés kutatása az agráregyetemenek és -főiskolákon, illetve kutatóintézetekben folytatódott. Györffy Béla (1928–2002) és munkatársai az MTA Mezőgazdasági Kutatóintézetében, Martonvásáron már az 1960-as évektől kezdődően arra törekedtek, hogy országos viszonylatban megismertessék a szakemberekkel (a termesztési rendszereken keresztül) a növénytermesztési kísérletek beállításának és értékelésének korszerű módszereit (1.1. kép). Az 1970-es évektől kezdődően a kutatóintézetekben, agráregyetemenek, termesztési rendszerekben és nemesítő intézetekben országsszerte beállított termesztési és fajtakísérletek

jelentősen hozzájárultak a növénytermesztési technológiák fejlesztéséhez és a fontosabb kultúrnövények termésének növekedéséhez.

Szóab közeli munkatársaként *Baráth Csabáné* több éven keresztül részt vett a nemzetközi biometriai társaság munkájában és első szerzője volt az 1996-ban megjelent „Biometria” c. könyvnek. 1995-ben jelent meg *Précsényi I.* ökológus előadásai alapján a KLTE Ökológiai Tanszékén *Précsényi et al.*: Alapvető kutatásszervezési, statisztikai és projektértékelési módszerek a szupraindividuális biológiában c. könyv. Az elméleti biometria kiemelkedő hazai művelője volt *Juhász-Nagy Pál* ökológus (1935–1993). A biomatematika elméleti kérdéseit tárgyalja *Izsák, J.–Juhász-Nagy, P.–Varga, Z.*: 1981. Bevezetés a biomatematikába. c. könyve. *Podani J.*: 1997. Bevezetés a többváltozós biológiai adatfeltárás rejtelmeibe c. könyve a növénytermesztők számára is javasolható. *Tomcsányi P.* akadémikus a kutatómódszertant oktatta több mint egy évtizeden keresztül a doktorandusz hallgatóknak. „Általános kutatómódszertan” c. könyve 2000-ben jelent meg.

Harnos Zsolt (1941–2009) akadémikus több évtizeden keresztül a hazai biometriai kutatásokat irányította és szervezte a 3 évenként megrendezésre kerülő biometriai konferenciákat. Az akadémián Biomatematikai és Biometriai Interdiszciplináris Bizottságot hozott létre és azt irányította. 2005-ben jelent *Harnos, Zs.–Ladányi, M.*: „Biometria agrártudományi alkalmazásokkal” c. könyv, amely agráregyetemen és -főiskolákon javasolt tankönyv. *Hajtman B.*: 2012. Bevezetés a biomatematikába c. könyve közérthető módon rövid áttekintést ad az adatok gyűjtéséről, rendszerezéséről és a biostatistikai módszerekről.



1. 1. kép. Györfly Béla (középen) a martonvásári tenyészkertben növénytermesztőknek ismerteti kukoricatermesztési kísérleteinek eredményeit

2. Bevezetés a kísérletes kutatásba

2.1. A kutatás jellemzői és különböző kutatási megközelítések

A kutatás tágabb értelemben úgy definiálható, mint a kérdések feltevésének és megválaszolásának folyamata. A kutatás modernebb definíciója, hogy egy tárgynak vagy problémának szisztematikus és objektív vizsgálata azzal a céllal, hogy új információt és alapelveket fedezzünk fel és ezáltal növeljük ismereteinket. A kutatási folyamat négy stádiumot foglal magában (Ireland, 2010):

(1) *Kutatási kérdés azonosítása.* A kutatási kérdés megadja a kutatás célját és tárgyát. Kétségtelenül igaz, hogy jó kutatás először is jó kérdésekből származik. Egy kutatási kérdés megválaszolása azonban állandóan új kutatási kérdésekhez vezet.

(2) *Adatok gyűjtése.* Az adatok egyszerű feljegyzett tények, pl. mérések, számolások, leírások, melyeknek önmagukban kevés a jelentésük, amíg fel nem dolgozzuk azokat és fel nem tárjuk összefüggéseiket. A kutatási folyamatban az adatgyűjtés magában foglalhatja egy kísérlet végzését, egy felmérés végzését vagy megfigyeléseket és rögzítéseket.

(3) *Információ nyújtása.* Az információ a begyűjtött adatok feldolgozásából származik és megállapításokat tartalmaz az észlelt igazságról vagy választ ad sajátos kérdésekre. Az adatok feldolgozását objektív adatelemzési és interpretációs módszerekkel érjük el.

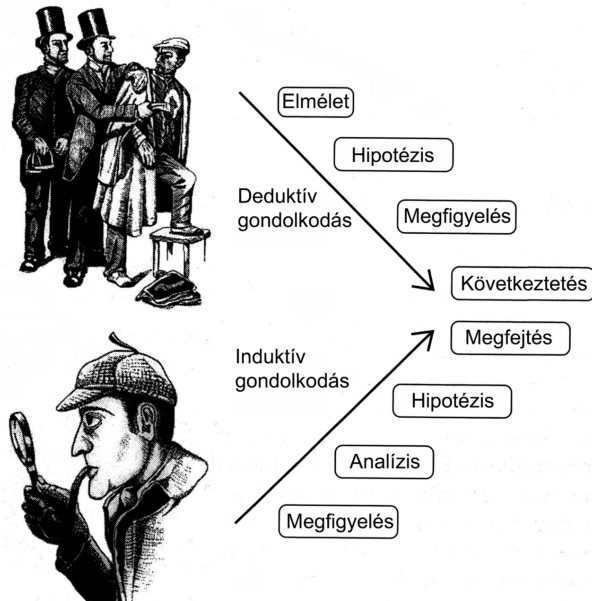
(4) *Ismeretek bővítése.* Ismeret a különböző információk egybevetéséből és integrálásából származik, ténylegesen lehetővé teszi egy sajátos folyamat vagy tárgy (szubjektum) általános megértését.

A tudományos megközelítés tradicionálisan empirikus természetű, *kísérletezésen és megfigyelésen* alapul. A tudományos kutatást elsődlegesen a hipotézis tesztelés, értékelés, a hipotézis módosítása vagy szükség szerint új hipotézis felállítása és ezután újbóli tesztelés folyamatán keresztül végezzük.

Gyakran megkülönböztetünk *kvantitatív kutatást*, amely magában foglalja a változók közvetlen numerikus mérését és a *kvalitatív kutatást*, amely az adatok és folyamatok leírását tartalmazza. Míg a kvantitatív adatok analízise numerikusan pontos statisztikai méréseken alapul, a kvalitatív adatok analízise problematikus, azonban magában foglalhat „pszeudo-kvantitatív” megközelítést, melynek során a kvalitatív változót átalakítjuk kvantitatív mérőszámmá.

A tradicionális tudományos gondolkodásmód, amelyet *deduktív*nak hívnak, egy gondolattal kezdődik, amely néhány sajátos jelenségre vonatkozik, melyből tesztelhető hipotézis formálható. A tesztelést követően a hipotézis megerősíthető vagy elutasítható és ez egy módosított vagy új hipotézis formálásához vezet. A hipotézis vizsgálatnak ez a folytonos folyamata a jelenség valódi alapját eredményezheti, amelyet így megfejtethünk. Másrésztől, az *induktív gondolkodás* magában foglal egy logikus láncolatot, amely a deduktív gondolkodással ellentétes irányba halad. Azaz, a folya-

mat elkezdődik a kérdéses jelenség semleges, független megfigyelésének sorozatával. A megfigyelések integrációján keresztül, mintázatok és trendek válnak nyilvánvalóvá, amely viszont hipotézisek felállításához vezethet, amelyek ezután a jelenségek magyarázatát adják. Tehát a deduktív konklúzió a jelenségről az előre felállított hipotézis tesztelésén alapul, körültekintően kontrollált körülmények között, míg az induktív gondolkodás során megmagyarázó hipotéziseket indukálunk a jelenség többszöri megfigyelése alapján természetes környezetben. A deduktív és induktív gondolkodás megkülönböztetését jól illusztrálja két szórakoztató nyomozó, *Lestrade* (Scotland Yard) és *Sherlock Holmes* (detektív) karaktere (2.1. ábra).



2.1. ábra. Deduktív és induktív gondolkodás illusztrálása (Ireland, 2010 alapján)

Természetesen, a modern tudományos kutatás mindkét megközelítés aspektusait magában foglalja a problémák megoldásához. Ezáltal a tudományos kutatóknak, akik nagyon specifikus kutatási problémával foglalkoznak, ezeknek a különböző kutatási megközelítéseknek az ismerete szükségessé válhat, hogy elősegítsék új innovatív módszerek alkalmazhatóságát és lehetővé tegyék számos kutatási probléma lehetséges értelmezését és megoldását.

2.2. A kísérletek különböző típusai

Azokat a kísérleteket, amelyekben egy vagy több körülményt szándékosan változtatunk meg, míg az összes többi körülményt, amennyire lehetséges, konstansnak tartjuk fenn, *manipulációs kísérleteknek* hívjuk. Nyilvánvalóan, ilyen kísérleteket legkönnyebben egy kutató laboratórium körültekintően kontrollált (szabályozott) körülményei között lehet végezni. A mezőgazdaságban és kertészetben azonban a növények mesterséges konstans körülmények közötti nevelésének csak korlátozott gyakorlati

értéke van. Miközben megállapítható egy manipulációs laboratóriumi kísérletben, hogy egy bizonyos faktornak van egy sajátos hatása a növényre, rendszerint szükséges lesz szántóföldi kísérletet beállítani, annak megállapítására, hogy a vizsgált faktornak ugyanaz-e a hatása a növényre szántóföldi körülmények között. *Szántóföldi kísérletre* van szükség ennek megállapítására. A biológusok és ökológusok természetesen még jobban függnék a szabadföldi kísérletektől, a valóságban a laboratóriumi kísérletek kevésbé fontosak számukra.

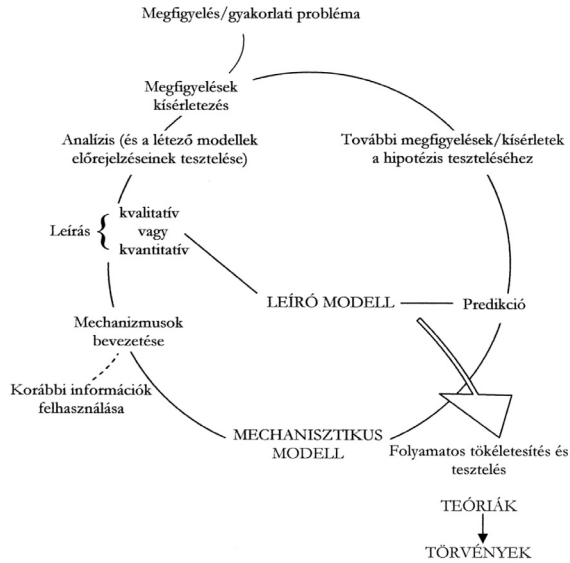
A szabadföldi kísérletben valószínűnek kell tekinteni, hogy a külső körülmények nem lesznek konstansok, hanem változhatnak térben és időben egyaránt. A szabadföldi kísérletek tervezése mögötti cél ezáltal az, hogy biztosítsuk a változó körülmények hatásának egyenletes eloszlását a kísérleti parcellákon, ezáltal lehetővé téve a kezelés faktor hatásának vizsgálatát. Ezt a kísérleti tervezés folyamatán keresztül érjük el, és míg a laboratóriumi kísérletek tervezése szabályozott körülmények között nagyon egyszerű lehet, a szabadföldi kísérletek tervezése ritkán egyszerű és a valóságban nagyon bonyolult lehet.

2.3. A növény-környezet interakciók

A növénytermesztési kutatás körütekintően megtervezett és több ismétlésben beállított szántóföldi kísérleteken alapul, melyekből részletes adatokat kapunk, rendszerint több éves időszakban, mielőtt bármely dedukciót vagy konklúziót levonhatnánk. A növénytermesztési kutatás módszerét négy általános lépés jellemzi: a probléma meghatározása, hipotézis felállítása, bizonyítékok kutatása (elsősorban kísérleti úton) és következtetések levonása. Ezt a folyamatot illusztrálja nagyon leegyszerűsített formában az 2.2. *ábra*. A kutatás megfigyelési és kísérleti fázisai egyformán fontosak, éppen úgy, mint a modellezési fázis. A tudományos ismeret nem csupán megfigyelési adatokat jelent, hanem magában foglalja a hipotézist, amely összhangban van az adatokkal, és a folyamatos interakció a hipotézis (ahogyan gondoljuk, hogy a rendszer működik) és a megfigyelési adatok (ahogy a rendszer ténylegesen működik) között vezet az előrehaladáshoz.

A vizsgálat kezdeti stádiumai a növény-környezet rendszer viselkedésének többnyire kvalitatív leírását adják. További előrehaladás bármely rendszer megértésében szükségessé teszi a kvantitatív megközelítést. A matematikai modellek felhasználhatók a hipotézisek felállítására, tesztelésére, továbbá predikciók generálására. Lehetnek empirikusak, amikor nem teszünk kísérletet a szóban forgó mechanizmus leírására. A mechanisztikus modell rendszerint megkísérel megmagyarázni a jelenséget az organizáció mélyebb szintjén és valószínűleg a növény-környezet rendszer pontosabb predikcióját adja eltérő környezetben (2.2. *ábra*).

A kutatás bármely ága, ahogy a kvalitatívtól a kvantitatív irányba halad, egyszer eléri azt a pontot, ahol a biometria használata a teória és a kísérlet összekapcsolására egyre fontosabbá és hasznosabbá válik. Nem lehet eléggé hangsúlyozni, hogy a biometria csupán egy módszer, amely lehetővé teszi a kutatónak, hogy biológiai hipotézisét kifejezze olyan módon, hogy a kvantitatív predikció lehetővé váljék, amely azután összehasonlítható a valós világgal.



2.2. ábra. A modellek szerepe a tudományos kutatásban (Jones, 1992)

2.4. Növénytermesztési kísérletek sajátosságai

A növénytermesztés, mint a biológiának bármely más területe, a komplex folyamatok megértésére törekszik. Számos forrásból és diszciplínából az ismeretek szintézise szükséges a sürgős és gyakorlati problémák megoldásához. A növénytermesztés ritkán tud elérni teljes megértést a tábla szinten és a gazdaság szintjén, mivel számos összefüggés és visszacsatolás van az időjárás, a talaj, a növény és az agrotechnika között. A növénytermesztési kutatás ezért szigorú kísérletezési módszertant fejlesztett ki. Valóban, a növénytermesztés kialakította a kísérleti és elméleti módszertan széles skáláját, többek között a szabadföldi kísérleteket, a fiziológiai kutatásokat (szabályozott környezetben) és az ökofiziológiai megközelítéseket (Berzsenyi, 2009a).

A szabadföldi kísérletek az agronómia repertoárját képezik évtizedek óta és a mezőgazdasági kutatás fontos részei maradnak még hosszú ideig. Ritkán identifikálnak kapcsolatot a fizikai környezet és a termés között, így hely-specifikusak, az egyik helyen kapott eredmények ritkán alkalmazhatók máshol. Az agronómia különböző területeinek, amelyek kutatják azokat a mechanizmusokat, amelyek összekapcsolják a vizet, a talajt és a növényt, fontos funkciója a szántóföldi kísérleti eredmények kiterjesztése. A szántóföldi kísérletek hozzájárulnak az ok és okozat megértéséhez, lehetővé teszik továbbá a növények és termesztési szisztémák jövőbeni teljesítményének előrejelzését. Ebben az összefüggésben, a modern eljárások, mint amilyenek a számítógépes termés és növekedési modellek, fontos szerepet játszanak a kísérleti eredmények kiterjesztésében (Berzsenyi, 2009a).

A növény és a környezet összefüggésének tanulmányozására fontosak a szabályozott környezetben (pl. üvegházban, fitotronban) végzett növénytermesztési kísérletek, ahol a szabadföldi kísérletek heterogenitása és a tényezők közötti kapcsolatok kiküszöbölhetők. Klímakamrákban vizsgáljuk a fényt, a vizet, a tápanyagokat, a

hőmérséklet, a fotoperiódus és más változók independens hatásait. E kísérletek lényeges sajátossága, hogy megtöriünk néhány kapcsolatot a környezeti változók között. Experimentális korlátozása a kapcsolatoknak a folyamatok és a környezet között, gyakran kivált nem tipikus reakciókat a növényben. A növénytermesztőknek nem szabad feltételezni, hogy ilyen esetekben a vizsgált faktor kis mértékű hatást fejt ki, vagy egyáltalán nincs hatással szántóföldön. A végső cél az, hogy visszatérjünk az eredeti problémára szántóföldön, és ehhez a feltört kapcsolatok igazi hatására a megismert folyamatokból kell következtetni.

A valóságban nagyon nehéz megtenni ezt a visszatérést a szántóföldre. A problémákkal különböző skálákon kell megbirkózni a különböző tudomány-területeket képviselő szakembereknek, mint amilyen a talajtan, a fiziológia és a genetika. Amikor a diszciplínák izoláltak maradnak, a szántóföldi skálájú problémákat ritkán oldják meg. Még akkor is, ha a diszciplínák összekapcsolódnak, a skálázásnak alapvető tudományos kihívása van, amennyiben valamilyen elméleti vagy matematikai eljárást használunk arra, hogy előre jelezzük a kísérleti mérésekből azt, ami valószínűleg bekövetkezik a növényállomány szintjén.

A tudomány néhány területén olyan keveset tudunk a faktorok közötti kapcsolatokról a populáció, a tábla vagy a felvételezési terület skáláján, hogy a következtetések csupán becslésnek tekinthetők. A globális környezeti változás vegetációra gyakorolt hatásának előrejelzése nagyon bizonytalan, mivel a kísérleteket nem lehet végezni olyan skálán, mint amelyre a kérdés vonatkozik. A bevezetett gének és fajták sorsa ugyancsak sajátos skálázási problémákat képvisel. A növénytermesztés szerencsés abban az értelemben, hogy a kísérletezés a termesztési gyakorlathoz közeli, reális skálán lehetséges.

2.5. A kísérletek tervezésének alapszabályai

A kísérletek tervezése alatt az egész folyamatot értjük, beleértve a kezdeti kutatói kérdések feltevését, a vizsgálni kívánt populáció meghatározását, mérésekhez a minták kiválasztását, az alkalmazni kívánt kezelések meghatározását és elosztását, a mérni kívánt változók azonosítását és az alkalmazandó adatelemzési módszerek meghatározását. Bármely kísérlet tervezésének három fő célja van (*Ireland, 2010*):

(1) Biztosítani, hogy a valódi kezelés-hatások elkülöníthetők legyenek mind a random (véletlen) variációtól a szubjektumok (kezelések) között és mind a külső kísérleti körülmények nem egyenletes hatásaitól.

(2) Biztosítani, hogy a megfigyelések/mérések torzítás nélkül végezhetőek legyenek.

(3) Biztosítani, hogy a begyűjtött adatok analízise objektív és megalapozott legyen.

Nyilvánvalónak tűnik, azonban hangsúlyozni szükséges, hogy a korrekt és pontos kísérleti tervezés alapvető ahhoz, hogy a kísérleti eredményekből valós következtetéseket lehessen levonni. Számos kísérlet sikertelen lesz a nem megfelelő kísérleti tervezés miatt.

Bármely kísérlet-tervezési folyamat fő szakaszai az alábbiak:

A kísérlet céljainak megállapítása. Világosan meg kell határozni a specifikus kérdést, amelyre a kísérlet irányul vagy a hipotézist, melyet tesztelni kívánunk.

A kísérlethez rendelkezésre álló források azonosítása. A kísérlethez rendelkezésre álló forrásokat meg kell határozni. Számításba kell venni nemcsak a rendelkezésre álló kísérleti anyag mennyiségét, hanem a rendelkezésre álló időt és munkaerőt a kísérlet végzéséhez. Számos kísérlet tönkrement a túlméretezett tervezés miatt, amely a kísérletet teljesen végrehajthatatlanná tette.

A kísérlet helyének és beállítási körülményeinek becslése. A külső körülmények egyöntetűségének mértékét becsülni kell. A kísérleti tervet lényegesen módosítani kell, ha a hely és/vagy a külső körülmények nem egyformák (uniform).

A vizsgálni kívánt populációk azonosítása. Fontos, hogy világosan meghatározzuk a populációt, amely a kísérlet tárgyát képezi. Ha mintákat használunk, fontos, hogy a populációt, amelyre a következtetések vonatkoznak, meghatározzuk a kísérlet előtt, és a minták a megfelelő populációból származzanak. Fontos szabály, hogy a statisztikai analízis nem teszi lehetővé a következtetések extrapolálását (kiterjesztését) egyik populációból a másikra.

A stabilitás nagyságának számbavétele, amely valószínűleg bekövetkezik a mintákon belül. Ahol lehetséges, a variáció és az egyenetlenség okait és mértékét a vizsgált anyagban identifikálni szükséges előzetesen és számba kell venni, hogyan minimalizálható a variabilitás.

Az elvégzendő megfigyelések/mérések típusának és ezáltal a kísérletekből származó adatok típusának azonosítása. A kutatónak előre, nagyon világosan kell tudnia, milyen változók lesznek mérve, pl. növénymagasság, gyökér/hajtás arány, a sztomák száma egységnyi levélterületen, és ezáltal tudnia kell, hogy milyen adattípusok keletkeznek, pl. folytonos mérési adatok, nem folytonos számolt adatok, stb. Ez fontos, mivel az adatok típusa meghatározza az alkalmazható adatelemzés módszerét.

Kezelés-csoportok azonosítása és a kezelések kijelölése. Előre pontosan specifikálni szükséges a kezelések számát és minden kezelés szükséges ismétléseinek számát. A kezeléseket ezután elosztjuk a parcellákra vagy parcella-csoportokra egy azonosított eljárással. Egy világos, torzításmentes eljárást a minták kiválasztására a soron következő mérésekhez ki kell alakítani.

Mivel gyakorlatilag lehetetlen eltávolítani minden variációt a kísérleti anyagból, és ezáltal nem lehetséges eltávolítani minden kísérleti hibát, fontos, hogy a hibát a minimális elfogadható szinten tartsuk és egyenletes eloszlása legyen az egész kísérletben. Ez általában elérhető a kísérleti tervezés két szabályának betartásával: (i) randomizáljunk teljes mértékben, ahogy lehetséges, (ii) annyi ismétlés legyen, amennyi lehetséges.

A legmegfelelőbb módszer azonosítása a kísérleti adatok elemzésére. Az alkalmazni kívánt statisztikai analízis módszerét azonosítani kell a kísérlet-tervezési folyamat részeként. A módszer kiválasztása függ az elvégezni tervezett mérések típusától, a keletkezett adatok típusától. A kísérlet-tervezésnek biztosítani kell, hogy érvényes statisztikai analízis módszer legyen alkalmazható az adatok leírására és az adatokból következtetések levonására. Értéktelen az a kísérlet, amely objektíven nem analizálható adatokat képez.

2.6. A kísérleti adatok statisztikai analízisének célja

Számos vizsgálat az alap- és alkalmazott kutatásban szükségessé teszi numerikus adatok begyűjtését, melyekből következtetést kell levonni. A statisztikai analízis alapvetően a matematikai módszereknek egy sorozata, amely lehetővé teszi a felhasználónak, hogy leírja a számszerű adatokat és objektíven értékelje az adatokból levonható következtetések megbízhatóságát. A matematikai analízis fő szabályait a következőképpen lehet összefoglalni (*Ireland, 2010*):

- Leírni és összegezni az adatsorokat.
- Értékelni annak érvényességét, hogy a felhasznált minták megfelelően reprezentálják-e a populációt.
- Jelentős trendek identifikálása az adatokban, a természetes háttér-variációval szemben.
- Elősegíteni objektív következtetések levonását az adatsorozatból.

Amennyiben objektív statisztikai analízis nem alkalmazható a kvantitatív kísérleti adatok értékelésére, akkor az ilyen adatokból levont kísérleti következtetéseknek nagyon limitált az értéke. Annak biztosítására, hogy a kísérlet képes legyen analízálható adatokat szolgáltatni, nagyon fontos, hogy az alkalmazni kívánt statisztikai analízis módszereit identifikáljuk a kísérleti tervezés során, az adatok begyűjtését megelőzően.

(1) Adatsorok leírása, összegezése és összehasonlítása

A statisztikai analízis használatának egyik fő célja az adatsorok összegezése, hogy érzékelhessük az adatokban levő trendeket. Az átlagok közötti különbség nagyon könnyen inkább a növények természetes variációjának tulajdonítható, mint pl. a műtrágya kezelésnek. Minthogy az adatok variabilitása alapvetően hatással van arra, hogy képesek legyünk interpretálni az adatokat, ezt a variabilitást is le kell írunk. A variabilitást számos statisztikai mutatóval mérhetjük, melyeket együttesen a szóródás mutatóinak hívunk. Amikor statisztikai mutatókkal írunk le egy adatsort, akkor a centrális tendencia (közéérték) és a szóródás mérőszámai egyaránt fontosak, mivel meghatározzák a levonható következtetések érvényességét. A szóródás mutatószámai magukban foglalják az adatok terjedelmét, varianciáját és szórását.

(2) Az adatok variabilitásának problémája a kísérletek analízisében

Az adatelemzés sokkal könnyebb lenne, ha az adatoknak nem lenne természetes variabilitása. A világban minden ki van téve természetes és random variabilitásnak. A kérdés az, hogyan tudjuk az adatokban megfigyelt trendeket, amelyek valamilyen felismert oknak tulajdoníthatók, elkülöníteni a természetes random variációtól?

A természetes variabilitás egy alapvető probléma, amely minden adattípus interpretációját megnehezíti. A statisztikai analízis egyik fő célja, hogy adjon egy módszert ennek a problémának a megoldására és tegye lehetővé reális trendek és hatások identifikálását az adatokban lévő természetes variáció háttérével szemben. Ezt főként úgy érjük el, hogy meghatározzuk egy sajátos oknak tulajdonítható sajátos hatás valószínűségét, miközben elfogadjuk, hogy a természetes variabilitás miatt sohasem lehetünk 100%-osan biztosak bármely általunk levont következtetés korrektségében (helyességében). A valóságban a statisztika nem tud bizonyítani egy dolgot. Ehe-

lyett a statisztikát arra használjuk, hogy kimutassuk annak valószínűségét, hogy egy hatásnak van bizonyos oka. Érthetően, ha a valószínűség nagyon nagy, és ugyanazt az eredményt kapjuk újból és újból az ismétléses kísérletekben, akkor a következtetések széleskörűen elfogadhatók és az eredmény hozzájárul ismereteink bővítéséhez. Az értelmezési bizonytalanságok hátterét a többváltozós elemző módszerek alkalmazásánál Józsa (1988) tanulmányozta.

(3) *Statisztikai adatelemzés az adattrendek értékelésére és következtetések levonására*

A kutatóknak el kell fogadni azt a tényt, hogy sohasem lehetnek 100%-osan biztosak abban, hogy egy kísérletben megfigyelt sajátos hatásnak van egy sajátos oka. Amit azonban tehetünk, hogy meghatározzuk annak *valószínűségét*, hogy az eset bekövetkezik. Lehetséges meghatározni annak valószínűségét, hogy bizonyos trendek vagy megfigyelt különbségek egy sajátos kezelésnek vagy oknak tulajdoníthatók-e. Ezt egy statisztikai eljárással érhetjük el, melyet *szignifikancia vizsgálatnak* hívunk. A szignifikancia vizsgálatok becsülik annak valószínűségét, hogy a megfigyelt adatok összeegyeztethetők-e a negatív hipotézissel, hogy nincs kezelés-hatás. Ha ezt követően kiderül, hogy megfelelően kicsi annak valószínűsége, hogy az adatok támogatják a negatív hipotézist, például kevesebb, mint 5%, akkor levonható az a következtetés, hogy valóban van kezelés-hatás és ez a kezelés-hatás elsődlegesen értelmezi a megfigyelt adatokat. A szignifikancia vizsgálatok képezik az alapját a kísérleti adatok analízisének.

3. A szántóföldi kísérletezés alapelvei és alapfogalmai

A mezőgazdasági kutatás számos területén, a növénytermesztésben különösen, a kisparcellás szabadföldi kísérletek az alapvető forrásai új információ szerzésének. A kutatási programoknak számos célja lehet: (a) Új fajták előállítása, amelyek jobbak a meglévő fajtáknál termőképességben, minőségben, betegségekkel szembeni rezisztenciában stb. (b) A termelési tényezőkben bekövetkezett változások hatásának tanulmányozása és meghatározása azoknak a faktoroknak, illetve a szintjeiknek, amelyek optimalizálják a termést, a minőséget és más tulajdonságokat. (c) Új technológiák, új kultúrnövényfajok és a potenciális mezőgazdasági termelés új területeinek vizsgálata. (d) Új ismeretek bemutatása a termelőknek, akik végső soron határoznak arról, hogy használják-e azokat vagy nem. (e) Alapvető megértése azoknak a faktoroknak, amelyek szabályozzák a kultúrnövények produkcióját és azt, ahogyan ezek a faktorok működnek. E célok megvalósítása érdekében a növénytermesztő kutatók valamilyen típusú kísérletet állítanak be. Nagyon sok esetben ezek kisparcellás szabadföldi kísérletek.

Kísérlet (experimentum). Olyan mesterségesen előidézett és tervszerűen ellenőrzött természeti folyamat vagy folyamatok sorozata, melynek célja megfigyelések gyűjtése a folyamat okairól, okozatairól és kapcsolataikról. A kísérlet kiindulási alapja rendszerint egy hipotézis, melynek alapján szabatosan meghatározható a kísérlet célja. A kísérletnek kritériuma, hogy a hatótényezőket mesterségesen működtetjük. Ha ez nem így történik, megfigyelésről beszélünk (*Jánossy et al.*: 1966. Biometriai értelmező szótár).

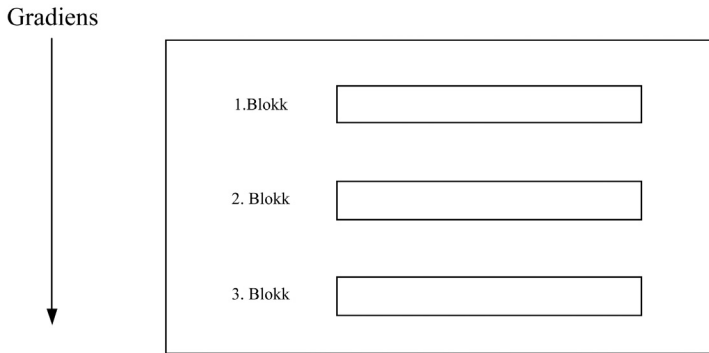
A kísérlet célja meghatározni a kísérletező által szabályozott változók (kísérleti változók vagy kezelés faktorok) hatásait a kísérleti egység (pl. parcella) egy vagy több jellemzőjére (reakció változó). Fő meghatározó sajátosságai: (1) a kísérletező random (véletlenszerűen) jelöli ki a kezeléseket vagy a kísérleti változó szintjeit a kísérleti egységre és (2) mindegyik kísérleti változónak két vagy több szintje van (*Hurlbert*, 2013).

Kísérleti hiba és csökkentése. A kísérlet eredményeit terhelő hiba, az egyformán kezelt kísérleti egységek (parcellák) közötti variációt jelenti. Számos forrása van a szabadföldi kísérletek variációjának. Oka lehet a kísérlet nem megfelelő tervezése és végrehajtása (parcella mérete és alakja, a kísérlet elhelyezése), de okozhatják véletlen tényezők (mérési hibák, egyedi variáció stb.). A kísérletek gondos technikai végrehajtásán kívül megfelelő kísérleti elrendezéssel, módszerekkel csökkenthetjük a kísérleti hibát. A modern kísérleti tervezés célja egyrészt, hogy mérjük a kísérleti hibát, másrészt csökkentjük a lehetséges mértékig. Általánosságban, a természetes variabilitás fő forrásai szabadföldi kísérletekben a következők: (1) a növény variabilitása, (2) szezonális variabilitás és (3) a talaj variabilitása. A modern kísérleti módszertan számos olyan eljárást foglal magában, amely lehetővé teszi a kutatónak, hogy mérje és szabályozza a kísérleti hibát. Ezek a következők: (1) ismétlés, (2) randomizálás és (3) blokkképzés.

Ismétlés. Kísérletekben ismétlésről beszélünk, ha ugyanazt a kezelést több kísérleti egységen vizsgáljuk. Az ismétlés elsősorban a következő célokat szolgálja: 1. A kísérletekből kapott eredmények pontosságának fokozása. A kezeléseket ugyanis annál nagyobb pontossággal hasonlítjuk össze, minél több ismétlésből számítjuk a kezeléslagokat. Az ismétlések számának növelésével ugyanis csökkenthető a középérték szórása (a középérték hibája). 2. Általában csak az ismétlés teszi lehetővé, hogy a hiba variancia becslésével a kísérleti hiba nagyságát mérjük. 3. Megnöveli a következtetések megbízhatóságát. Az ismétlések számát az alábbi tényezők határozzák meg: 1. A kísérleti terület talajában megfigyelt variabilitás mintázata és nagysága. 2. Az eltérés nagysága (a kezelések között), amelyet ki akarunk mutatni. 3. Az indokolt szignifikancia szint. 4. A kezelések száma. 5. A források nagysága, amelyet a kísérletre fordíthatunk. 6. Korlátok a költségekben, munkaerőben, időben stb.

Randomizálás. Ha valamilyen tetszőleges eljárás során, mesterségesen szerephez juttatjuk a véletlent, randomizálásról beszélünk. Kiemelkedően nagy a jelentősége a kísérletek elrendezésében. Itt a randomizálás a kezelések véletlenszerű hozzárendelését jelenti a kísérleti egységekhez (parcellákhoz). Megfelelő randomizációval bármely parcellának azonos a valószínűsége, hogy bármely kezelésre kijelöljük. Néhány kísérleti elrendezésnél (pl. blokkok kialakításával) a randomizáció korlátozott bizonyos módon, de nincs olyan kísérleti terv, amelyből teljesen kimaradna. Legalább két oka van a randomizációnak: 1. Kizárjuk az elfogultságot. A randomizáció biztosítja, hogy egyetlen kezelés sem kerül előnybe vagy hátrányba azért, mert azokra a parcellákra került. 2. Biztosítjuk a függetlenséget a megfigyelések között. Ez szükséges a megalapozott szignifikancia vizsgálatokhoz és intervallum becslésekhez. Általában csak a randomizálás teszi lehetővé azt, hogy megfigyelési, illetve kísérleti adatainkat a biometria módszereivel értékelhessük. Ezért a randomizálás a kísérletek elrendezésének elengedhetetlen alapkövetelménye. A gyakorlatban a randomizálást legelőnyösebben véletlen számok táblázatainak felhasználásával hajthatjuk végre. A számítógépes kísérlettervezés magában foglalja a randomizációt is.

Blokkok kialakítása. A kísérleti egységekből (parcellákból) a kísérlet végrehajtását megelőzően kialakított csoportokat nevezzük blokkoknak. Az eljárás itt az, hogy csoportosítjuk a parcellákat (kísérleti egységeket), homogén egységek blokkjaiba vagy csoportjaiba. A kezeléseket ezt követően véletlenszerűen jelöljük ki a parcellákba, a blokkokon belül. Az alapvető cél az, hogy a kísérleti hibából eltávolítsuk a blokkok közötti variációt. A kezelés összehasonlítások ezt követően a blokkokon belül történnek, olyan parcellákon, amelyek amennyire csak lehet, hasonlóak. A döntő a sikeres blokk képzésnél az, hogy maximalizáljuk a variációt a blokkok között és minimalizáljuk a variációt a parcellák között a blokkon belül. A blokkok alakjának és irányának megválasztásához ismernünk kell a variáció forrását (pl. talajtermékenységi gradiens, tábla lejtése, rovar invázió stb.). Ha a gradiens egyirányú (azaz csak egy gradiens van), használjunk hosszú, keskeny blokkot és úgy helyezzük el, hogy a blokk hosszabbik oldala merőleges legyen a gradiens (pl. lejtés) irányára (3.1. ábra). Ha két gradiens van egymásra merőlegesen, akkor használjunk négyzet alakú blokkot és kétirányú blokk-képzést (pl. latin négyzet elrendezés).



3.1. ábra. A blokkok elhelyezése lejtős területen a gradiensre merőlegesen

Számos oka van a blokk-képzésnek: 1. Növelheti a precizitást azáltal, hogy a blokkok közötti különbségeket eltávolítjuk a kísérleti hibából a kísérleti adatok analízise során. 2. Mivel az összehasonlításokat az egyforma parcellák blokkjain belül tesszük, a megfigyelt különbségek még jobban a kezelések közötti valódi különbségeket tükrözik vissza. 3. Néha növelheti az információt a kísérletről. A legegyszerűbb esetekben minden blokkon belül valamennyi vizsgált kezelést beállítjuk egy-egy kísérleti egységen (pl. véletlen blokkelrendezés). A blokkok méretének csökkentése érdekében gyakran előnyösebb egy blokkba csak a kezelések egy részét beállítani (nem teljes blokk elrendezés).



3.1. kép. Tenyészedény kísérlet inhomogén fitotron növénynevelő kamrában Martonvásáron, a hőmérsékleti gradiens hatásának tanulmányozására herbicid-érzékenységi kísérletben

Kísérleti egység. A legkisebb rendszer vagy kísérleti anyag egysége, amely egy kezelést (vagy kezeléskombinációt) kap és amelyet függetlenül kezelünk a többi ilyen rendszertől a kísérlet minden szakaszában. A függetlenség azt is jelenti, hogy egyik kísérleti egységnek nem lehet hatása a másik kísérleti egységre. A földnek egy parcellája, egy tehen vagy a tehenek egy csoportja vagy egy fészekaljja baromfi, mind kísérleti egységek. A csoportos takarmányozási kísérletben a tehenek csoportja a kísér-

leti egység, míg az egyedi takarmányozási kísérletekben mindenegyes állat a kísérleti egység. Látható, hogy egy kísérleti egység állhat egy egyedből vagy az egyedeknek egy csoportjából, és ennek világosnak kell lenni a kísérleti elrendezésből.



3.2. kép. Tenyészláda kísérlet (balra) és petricsésze kísérlet csíráztató kamrában (jobbra)

Parcella. A kísérletek egyöntetűen kezelt legkisebb egysége, függetlenül attól, hogy növény, állat, mikroorganizmus a kísérlet tárgya. A szántóföldi kísérletekben, a növény termesztését szolgáló földterületet jelenti, amely a kísérleti kezelést kapta, amelyre megfigyelésünk vonatkozik és alapul szolgál a kezelésre adott reakció értékelésére. Más esetekben a parcella lehet egy vagy több bokor vagy gyümölcsfa. Nem parcellás kísérletekben külön megnevezzük a megfigyelési egységet, amely lehet pl. tenyészedeny, petricsésze, mikroszkóp látótere, egy növény, egy levél stb. (3.1–3.2. kép).

Parcella mérete és alakja. Számos faktor van hatással a parcellaméretre. Ezek közül a legfontosabbak: 1. A kultúrnövény típusa. 2. A kísérlet típusa. 3. A kutatási program fázisa 4. A kísérleti hely variabilitása. 5. A szegélyhatások jelenléte és természete. 6. A kísérletben használt gép típusa. 7. A kezelések száma és típusa. 8. A rendelkezésre álló földterület. 9. Költségek. A különböző faktorok hatását a parcella méretére a 3.2. ábra szemlélteti.

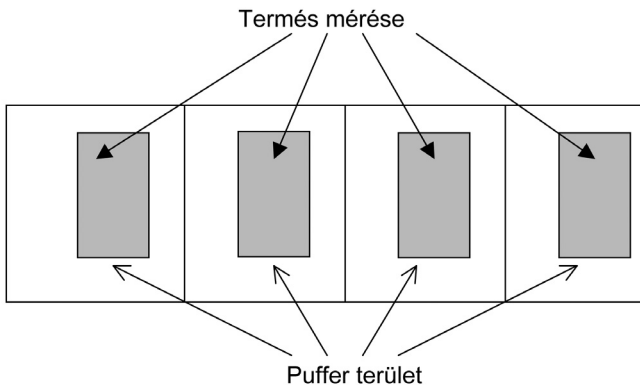
Faktor	Növekvő parcellaméret	
	Kis parcellák	Nagy parcellák
Talaj variabilitás	Egységes	Heterogén
Kultúrnövény	Gyep → Gabonafélék → Kapás növények	Fák → Legelő
Kutatás fázisa	Korai	Késői
Kísérlet típusa	Nemesítés → Trágyázás	Talajművelés → Öntözés
Gépesítés	Nincs	Kutatás → Üzemi méret

3.2. ábra. A parcella méretére ható faktorok

A parcelláknak elég nagyoknak kell lenni ahhoz, hogy géppel végezhessük el a talajművelést, a vetést, a műtrágyázást, a növényvédelmet, a betakarítást és egyéb kísérleti munkákat. A szegélyek eltávolítása után a nettó parcella megfelelő nagyságú legyen a termésméréshez. Ezen túlmenően nincs célja a parcellaméret további növelésének, ugyanis a kísérlet precizitása gyorsabban nő az ismétlések számának, mint a parcella méretének növelésével. Másrészt, a nagyon kis parcelláknak tendenciában nagyobb a variabilitása, mivel: (i) A betakarítási veszteségeknek vagy mérési hibáknak relatíve nagyobb a hatása. (ii) A variabilitás sorba vetett (kapás) növényeknél kis parcellákon nagy lehet a kevesebb növényszám miatt. (iii) A kompetíció és a szegélyhatás valószínűleg viszonylag nagyobb kis parcellákon.

Nagyság szerint megkülönböztetünk mikro-, kis- és nagyparcellás kísérleteket. Általában a talajművelési, trágyázási és növényápolási kísérletekben a parcella nagysága 70–100 m², öntözési kísérletekben >100 m². Fajta összehasonlító, gyomirtási és egyéb növényvédelmi, továbbá vetésidő kísérletek parcellamérete többnyire 15–20 m². Üzemi kísérletekben a parcella területe kb. 0.1 ha, a mikroparcellás kísérletekben viszont csak pár m². Szántóföldi kísérletekben a parcella alakját a kísérlet célja és a technikai adottságok, illetve lehetőségek szerint választjuk meg; leggyakrabban téglalap, négyzet, de lehet kör vagy legyező alakú is (pl. sűrítési kísérletben).

Bruttó parcella és nettó parcella. A bruttó parcella a teljes parcella elnevezése. Nettó parcellának nevezzük a homlok- és szegély-, illetve puffer sorok nélküli parcellát. A szegély- és puffer sorokat, illetve növényeket eltávolítjuk a termésmérés előtt. A kísérleti kezeléseket bruttó parcellákon állítjuk be, a termést viszont nettó parcellákon mérjük (3.3. ábra).



3.3. ábra. Bruttó parcella a teljes parcella (puffer területekkel együtt).
A termés mérése a nettó parcellákon történik

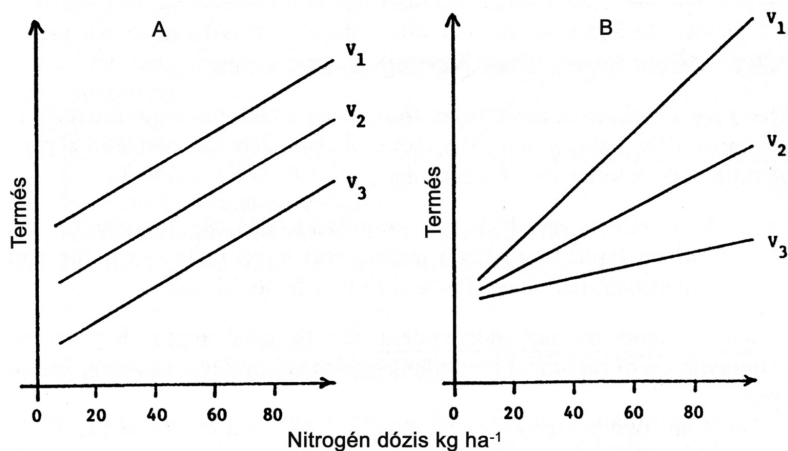
Szegélyhatás. A szomszédos terület hatása a kísérleti parcella növényeinek szélső soraira. E pozitív vagy negatív hatás kiküszöbölésére szoktunk kiegyenlítő sorokat (puffer sorok), esetleg egész parcellákat alkalmazni (puffer parcellák). A szegélyhatások olyan általánosak, hogy majdnem minden kísérlet beállításánál törekedni kell minimalizálásukra. A szokásos gyakorlat az, hogy a kísérleti parcella értékelése előtt eltávolítjuk és a termés értékelésekor nem vesszük figyelembe a szélső sorokat (lásd nettó parcella).

Kezelés. Egytényezős kísérletben a szóban forgó tényező vizsgált szintjeit nevezük kezeléseknek. Így pl. trágyázási kísérletben a különböző összetételű vagy adagú műtrágyák; fajtakísérletben az összehasonlított fajták; talajművelési kísérletben a talajművelés különböző módozatai (pl. konvencionális, minimális talajművelés) jelentik a kezeléseket. Többtényezős kísérletben a vizsgált tényezők (faktorok) különböző szintjeinek kombinációit nevezük kezeléseknek, bár itt inkább kezeléskombinációkról beszélünk.

Kontroll (standard kezelés). A kísérletnek az a kezelése, melyhez a kísérletben vizsgált többi kezelést hasonlítjuk. Trágyázási kísérletben például a trágyázatlan vagy a szokásosan trágyázott kezelés lehet a kontroll. Gyomirtási kísérletben általában több kontroll kezelést állítunk be, így gyomos és kapás kontrollt, standard herbicidet.

Kezeléshatás. A kezelésátlag és a standard kezelés (pl. a műtrágyázási kísérletben a kezeletlen kontroll) átlagának különbsége. A faktoriális hatás a faktoriális kísérletekben jelentkező főhatások és kölcsönhatások közös megjelölése. Tartamkísérletekben megkülönböztetünk közvetlen hatásokat, utóhatásokat és kumulatív hatásokat.

Kölcsönhatás (interakció). Ha különböző tényezők egymás hatását befolyásolják, a tényezők között kölcsönhatás van. A kölcsönhatás pozitív, ha a tényezők összehatása az egyenkénti hatásuk összegénél nagyobb, negatív, ha kisebb. A kölcsönhatás számszerű értékét az összehatás és a tényezőkénti hatások összege közötti különbség adja. Ha az összehatás az egyenkénti hatások összegével azonos, a tényezők hatása egymástól független, nincs kölcsönhatás. A kölcsönhatást szorzójellel jelöljük (pl. $A \times B$). Két tényező kölcsönhatását elsőrendű kölcsönhatásnak, három tényező kölcsönhatását másodrendű kölcsönhatásnak nevezük. A kölcsönhatásokat leggyakrabban a varianciaanalízisben, a faktoriális kísérletek és kísérletsorozatok értékelésekor vizsgáljuk. A 3.4. ábra illusztrálja a lehetséges termésreakciókat, amikor három fajtát vizsgáltak négy N-műtrágyaszinten. Az A. ábrán a fajták közötti különbség ugyanaz az N-műtrágyázás minden szintjén. Nincs interakció a N mennyisége és a fajták között. Másrészt a B. ábra azt mutatja, hogy a fajták közötti különbség nő, ahogy a N dózis növekszik. Itt N dózis \times fajta interakció van. Nem mondhatunk semmit a fajtakülönbségekről anélkül, hogy egyidejűleg ne specifikálnánk a N műtrágyázás szintjét.



3.4. ábra. Lehetséges termésreakciók, amikor három fajtát termesztünk négy műtrágyaszinten

Reakció struktúra. A reakció változók listája, melyeket mérni kívánunk és a mintavételi terv, amely specifikálja, mikor, hol és melyik komponensét fogjuk mérni vagy megfigyelni (és feljegyezni) a kísérleti egységnek.

Kísérleti elrendezés. Kísérletekben a parcellák és a parcellákon a vizsgált kezelések tervszerű elhelyezése, illetve csoportosítása valamely előírt rendszer szerint. A kísérleti elrendezés célja, hogy az összehasonlított kezeléseknél lehetőleg azonos körülményeket biztosítsunk, illetve az ellenőrizetlen, tudatosan nem irányított hatások egyenlően érvényesüljenek az összes kezelést. A megfelelő kísérleti elrendezést a kísérleti kérdéstől, a vizsgált tényezőktől, a kezelésektől és ismétlésektől függően választjuk meg. A kísérleti elrendezés és a kísérlet biometriai értékelése egymással szorosan összefügg, az elrendezés módja egyértelműen meghatározza az alkalmas értékelési eljárást.

Variáció. A biológiában a változékonyság (variabilitás) mennyiségi vagy minőségi megnyilvánulása. A biometriában általánosságban egy populáció elemeinek különbözőségét jelenti. Valamely tulajdonság variációja egy populáción belül az eloszlással, illetve eloszlástípussal, valamint különböző mutatókkal (pl. szórás, terjedelem stb.) jellemezhető.

Variancia (szórásnégyzet, s^2) és szórás (s). A megfigyelési adatoknak, valószínűségi változóknak, populációknak, eloszlásoknak – a középérték mellett – a variancia a legfontosabb jellemzője. Független megfigyelésekből, mérésekből stb. származó n elemű x_1, x_2, \dots, x_n adatsor (minta) varianciáját úgy kapjuk meg, hogy az átlagtól való eltérések négyzetösszegét elosztjuk $n-1$ -gyel. Tehát a számláló az eltérésnégyzet-összeg, a nevező pedig a szabadságfok. A variancia négyzetgyöke a szórás (s).

Szabadságfok (df, DF). A szabadságfok egyszerűen a megfigyelések száma a mintánkban, amely „szabadon változik”, amikor a varianciát becsüljük. Mielőtt már meghatároztuk az átlagot, akkor csak $n - 1$ megfigyelés változik szabadon, mivel ismerve az átlagot és $n - 1$ megfigyelést, az utolsó megfigyelés fix.

Varianciaanalízis. A varianciaanalízis egy megalapozott statisztikai eljárás, melyet hipotézisek tesztelésére használunk, a variáció forrásainak felosztásával, a variancia komponensek becsülésével, a maradék variancia csökkentésével és értelmezésével és a hatások szignifikanciájának meghatározásával (McIntosh, 2015). A biometriában az egyik legáltalánosabban alkalmazott kísérlet értékelési módszer. Az adatok teljes varianciáját (pontosabban a teljes eltérésnégyzet-összeget) bontja fel különböző komponensekre. A felbontás célja kettős: egyrészt a szisztematikus (csoportok közötti) hatások kiszűrésével a véletlen hatások okozta ún. hiba-variancia becsülése, másrészt egyes szisztematikus hatások (csoportok közötti különbségek) szignifikanciájának vizsgálata. A felbontás módja igen különböző lehet aszerint, hogy hány tényező kísérletet értékelünk és milyen a kísérlet elrendezése. Az eredményeket ún. varianciatáblázatban szokás feltüntetni. Egy véletlen blokkelrendezésű kísérlet varianciatáblázatának szerkezete a következő:

Tényező	df	SS	MS	F-érték
Összes	$vr-1$			
Blokk	$r-1$			
Kezelés	$v-1$			Kezelés MS/Hiba MS
Hiba	$(r-1)(v-1)$			

Itt a teljes (összes) variancia a blokkok (r) közötti, a kezelések (v) közötti és a hiba-varianciára bontható. Az első kettő a blokkok, illetve a kezelések közötti szisztematikus különbségeket fejezi ki, az utolsó tényező pedig az összes egyéb (többnyire ismeretlen) tényezőknek az eredményekre gyakorolt hatását méri. E hatásokat a kísérletben véletlen hatásoknak tekintjük és összegüket kísérleti hibának nevezzük. A varianciaanalízissel ennek a véletlen hibának a varianciáját becsüljük.

A varianciatáblázat első oszlopa (Tényező) a variancia forrásait, a második oszlop a szabadságfokokat (df), a harmadik oszlop az eltérésnégyzet-összegeket (SS , SQ) tartalmazza. A következő oszlop (SS/df), a közepes négyzetes eltérés (MS , MQ). Az utolsó oszlopban van a számított F -érték. Az eltérésnégyzet-összeg (SS) minden egyes adat és számtani középértéke közötti különbség négyzetének összege.

$$\text{Képlete: } SS = \sum_{i=1}^n (x_i - \bar{x})^2, \text{ ahol, } n \text{ az adatok száma.}$$

A varianciaanalízisben az adatok főátlagtól való eltérésének négyzetösszegét, az ún. összes SS -t az egyes tényezők hatásának és kölcsönhatásának, valamint a maradék eltéréseknek megfelelő részek összegére bontjuk fel. Ha ezeket a megfelelő szabadságfokokkal elosztjuk, a tényezőkénti közepes négyzetes eltéréseket (MS) kapjuk.

A kísérleti hiba-varianciát a hiba MS értékével becsüljük. A kezelés MS értéke a kezelések közötti különbségeket méri, és annál nagyobb, minél nagyobbak ezek a különbségek. Ha a kezelések középértékei között valójában nincsenek különbségek, hanem a mért eltéréseket csak a véletlen hibák okozzák, akkor a kezelés MS is csak a hiba-varianciát fejezi ki. Értéke ilyenkor közel áll a hiba MS értékéhez. Ha azonban a kezelések középértékei valóban különbözöek, akkor a kezelés MS általában nagyobb, mint a hiba MS .

A két variancia hányadosa alapján, F -próbával (l. később) vizsgálhatjuk, hogy az eltérés csak a véletlen következménye-e. Az F -próba jelen esetben a két variancia homogenitásának vizsgálatán keresztül a kezelések középértékeinek homogenitását vizsgálja. Ha tehát a próba szignifikáns eredményt ad, akkor ez azt mutatja, hogy a kezelések középértékei között különbségek vannak. A továbbiakban a hiba-variancia (Hiba MS) ismeretében egyszerű t -próbákkal vizsgálhatjuk, hogy mely kezelések különböznek szignifikánsan, vagy valamelyik többszörös terjedelemprobát (pl. Duncan próba) ajánlatos alkalmazni.

Bonyolultabb kísérletek varianciaanalízisében tovább bontjuk a varianciát. Faktoriális kísérletekben (l. később) az összes tényező, valamint a kölcsönhatások is különválaszthatók. A varianciaanalízis és a vele összekapcsolt F -próbák és t -próbák alkalmazása bizonyos feltételekhez van kötve (normális eloszlás, hibák függetlensége, homogén szórás, hatások additivitása). Amennyiben a feltételek még megközelítőleg sem teljesülnek, adat-transzformációt (pl. logaritmus-transzformáció) célszerű végezni.

Varianciaanalízis modelljei. A varianciaanalízisben rendszerint lineáris additív modellt használunk. A megfigyeléseket (pl. termés) összegnek tekintjük, amely a μ (konstans) vagy főátlag, egy e hibakomponens, valamint a vizsgált tényezők főhatásaiból és kölcsönhatásaiból tevődik össze. A hatások szempontjából három esetet különböztetjük meg: 1. *Fix modell*: egy lineáris modell, amelyben csak fix hatások és maradék (reziduális) tagok vannak, 2. *Random (vagy véletlen) modell*: az összes hatás

valószínűségi változó és 3. *Vegyes (mixed) modell*: a hatások közül egyesek valószínűségi változók, mások fix hatások. A kísérletek elrendezéseinek ismertetésekor (5. fejezet) a varianciaanalízis fix modelljeit adjuk meg.

Kezelés-struktúra. A varianciaanalízisben specifikál egy modell formulát (a kezeléssel modell) és definiálja az összehasonlítani kívánt kezelés (fix) tagokat.

Blokk-struktúra. Specifikálja a modell formulát (a blokk modellt), levezetve a kísérleti elrendezésből; definiálja a blokk tényezőket és a hiba tényezőket a varianciaanalízishez.

Kontraszt (lineáris kontraszt, lineáris összehasonlítás). Tetszőleges x_1, x_2, \dots, x_n mennyiségek (változók, paraméterek, megfigyelési adatok, stb.) egy kontrasztjának nevezzük ezek bármely olyan $k = l_1x_1 + l_2x_2 + \dots + l_nx_n$ alakú lineáris függvényét, amelyben az együtthatók összege zéró, vagyis $l_1 + l_2 + \dots + l_n = 0$ (Jánossy *et al.*, 1966).

Szignifikanciaszint. Valamely statisztikai próbához előre megadott, rögzített hibahatár. Általában α -val jelölik, de a gyakorlatban a P jelölés is szokásos. A szokásos értékek: $\alpha = 0.05; 0.01; 0.001$. A szignifikanciaszintet igen gyakran százalékos formában adjuk meg: $\alpha\% = 5\%, 1\%, 0.1\%$. Minél kisebb szignifikanciaszintet választunk, annál ritkábban fordulhat elő, hogy tévesen elvetjük a nullhipotézist (másodfajú hiba). A legfontosabb próbák különböző szignifikancia szintekhez tartozó kritikus értékeire táblázatok állnak rendelkezésre (pl. F-táblázat).

Próba ereje. Valamely statisztikai próbával kapcsolatban a próba erejének (vagy röviden erőnek) nevezzük azt a valószínűséget, amellyel a H_0 hipotézist elutasítjuk, amikor az alternatív hipotézis (H_1) teljesül. A próba erejét $1 - \beta$ -val jelöljük, ahol β a másodfajú hiba. Mivel egy próbára vonatkoztatva az α (a maximális elsőfajú hiba) tetszőlegesen kicsi lehet, a próba jóságát a próba erejével jellemezhetjük. Minél kisebb a másodfajú hiba, annál nagyobb lesz a próba ereje, annál inkább képes két közeli populáció közötti különbséget detektálni. A próbák ereje a mintaszám emelésével növelhető.

Szignifikáns differencia (SzD, LSD). Általában így nevezzük valamely statisztikai próbában a megfigyelt különbség kritikus értékét. Ha a megfigyelt különbség a szignifikáns differenciát meghaladja, a nullhipotézist (azaz, hogy nincs különbség a kezeléseik között) (1. később) elutasítjuk, és a különbséget szignifikánsnak tekintjük. Az SzD-t akkor számítjuk ki, ha a varianciaanalízisben az F-próba szignifikáns kezeléshatást mutatott ki.

Kísérletek típusai. Megkülönböztetünk egytényezős és faktoriális kísérleteket, kísérletsorozatokat, tartamkísérleteket és üzemi (demonstáriós) kísérleteket.

Egytényezős kísérlet. Azokat a kísérleteket, amelyekben csak egyetlen tényező (faktor) változik, míg az összes többi tényező konstans (nem változik), egytényezős kísérleteknek hívjuk. Ilyen kísérletekben, a kezeléseik csupán az egyetlen változó faktor különböző szintjeit tartalmazzák. Minden többi faktort egyformán alkalmazunk minden parcellára egyetlen előírt szinten. Egytényezős kísérletek pl. a fajta, a trágyázási, a növényvédelmi (herbicidek, inszekticidek, fungicidek), a vetésidő és a növényszám kísérletek, ahol különböző fajtákat, a tápanyagok, növényvédő szerek különböző dózisait, eltérő vetésidők és különböző növényszámok hatását vizsgáljuk.

Faktoriális kísérlet (két- és többtényezős kísérlet, polifaktoriális kísérlet). Két- és több tényező hatásának egyidejű vizsgálatára beállított kísérlet. Az egyes tényezőket különböző szinteken (fokozatokon) vizsgálja. A faktoriális kísérletbe a tényezők szint-

jeinek kombinációit (az ún. kezeléskombinációkat) állítjuk be, ezek alkotják a kísérlet kezeléseit. A leggyakoribb és legegyszerűbb esetekben, a vizsgált szintek összes lehetséges kombinációját beállítjuk. A faktoriális kísérletek eredményei alapján az egyes tényezők hatását egy vagy több másik tényező különböző szintjein is vizsgálhatjuk, így a tényezők átlagos hatása (főhatása) mellett a kölcsönhatásokra vonatkozó következtetéseket is levonhatunk. A faktoriális kísérletek reprezentációs szélessége, valamint az eredményeikből nyerhető információ mennyisége általában lényegesen nagyobb, mint az egytényezős kísérleteké.

Kísérletsorozat. Kísérletek olyan sorozata, amelyet ugyanannak a kérdésnek a vizsgálatára azonos kezelésekkel folytatnak le térben és/vagy időben megismételve. A szántóföldi kísérletek pl. több helyen és több évben megismételhetők. A legtöbb biológiai kísérleti kérdésre (melyet nagyobb számú nem ellenőrizhető környezeti hatás befolyásol) csak kísérletsorozat eredményeiből kaphatunk kielégítő választ.

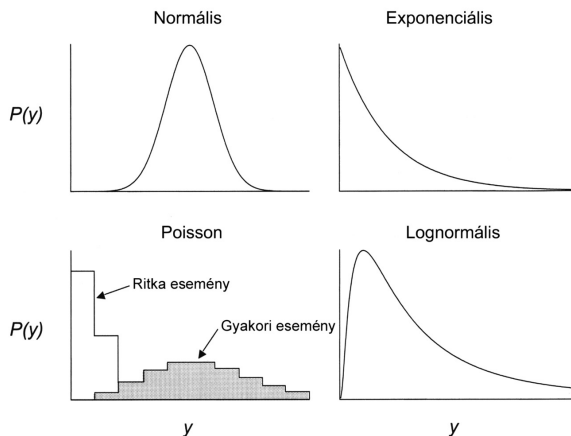
Tartamkísérlet. Olyan kísérlet, amelynek az idő a legjelentősebb tényezője, vagyis huzamos ideig van beállítva az időtényező hatásának vizsgálata céljából. A tartamkísérletben ugyanazokat a kezeléseket ugyanazon a parcellákon állítjuk be minden évben. Általában 10–15 év után beszélhetünk tartamkísérletről. E kísérletekben tanulmányozzuk a természet fenntarthatóságát, a különböző kísérleti kezeléseknél a tartamhatását a lassan változó folyamatokra (pl. talaj szerves anyag tartalma stb.). A tartamkísérletek ideális lehetőséget adnak a vetésciklusok, tápelem ciklusok, továbbá a termés és az időjárás közötti összefüggés tanulmányozásához. A hosszabb időszak alatt különösen a halmozódó (kumulatív) hatások és utóhatások mérésének a lehetősége nagy. A legrégebbi tartamkísérleteket több mint 160 évvel ezelőtt Rothamstedben (Anglia) állították be. A hazai tartamkísérletek (pl. Martonvásáron, Keszthelyen) több, mint 50 évesek (3.3. kép).



3.3. kép. Martonvásáron 1961-ben Györfly Béla által beállított vetésciklus tartamkísérlet

Üzemi (demonstrációs) kísérlet. Az üzemi (farm, gazdaság) termelés keretében beállított és annak körülményeit reprezentáló kísérlet. Különösen fontos szerepük van a tájtermesztés kialakításában. Interdiszciplináris kísérletek, és miközben kiegészítik a kisparcellás kísérleteket, nem helyettesítik azokat. Célja: új fajták, termelési eljárások stb. adaptálhatóságának, illetve egzakt kisparcellás kísérletekben megelőzően már vizsgált kezeléshatásoknak az ellenőrzése üzemi körülmények között. A kezeléseket általában ismétlés nélkül állítják be, ezért az egyes gazdaságok képezik a kísérlet egy-egy ismétlését, blokkját. Mivel az üzemi kísérletek precizitása eltér a kisparcellás kísérletekétől, nagyobb ismétlésszám (általában 5–8) szükséges. Az üzemi kísérletek értékelésére a variancianalízist a közelmúltban dolgozták ki.

A változók eloszlása. A legtöbb statisztikai eljárás épít a változó (vagy a hiba tényező a statisztikai modellből) valószínűségi eloszlásának ismeretére. A **normális (vagy Gauss) eloszlás** egy szimmetrikus valószínűségi eloszlás, jellegzetes harang alakú görbe (3.5. ábra). Meghatározható az átlaggal (μ) és a varianciával (σ^2). A standard normális eloszlás (z eloszlás) egy normális eloszlás zéró átlaggal és 1 varianciával. A normális eloszlású változók értékeinek zöme az átlag (μ) körül tömörül. Bármely normális eloszlásra könnyen meghatározhatjuk, hogy a megfigyelések mekkora hányada fordul elő a populációban bizonyos távolságra az átlagtól: a populáció 50%-a esik $\mu \pm 0.674\sigma$ közé, a populáció 95%-a esik $\mu \pm 1.960\sigma$ közé, a populáció 99%-a esik $\mu \pm 2.576\sigma$ közé. A normális eloszlás a legfontosabb valószínűségi eloszlás az adatelemzésben, ugyanis feltételezzük, hogy az analizálni kívánt változók (vagy az eltérések az illesztett modelltől) normális eloszlást követnek.



3.5. ábra. A változók normális, exponenciális, Poisson és lognormál valószínűségi eloszlása (Quinn és Keough, 2002 alapján)

Biológiai változóknak gyakran kifejezetten ferde az eloszlásuk. Egyik ferde eloszlás a **lognormál eloszlás**, amely azt jelenti, hogy a változó logaritmusa normális eloszlású. Azok a mért változók a biológiában, melyek nem lehetnek zérónál kisebbek (pl. hosszúság, tömeg stb.) gyakran lognormál eloszlást követnek. Az **exponenciális eloszlás** egy másik ferde eloszlás, melyet olyan esetekben használunk, amikor a vizs-

gált mennyiségnek egy intervallumon bekövetkező változása arányos a mennyiségnek az intervallum elején felvett értékével.

Fontos valószínűségi eloszlás a diszkrét változókra a **Poisson eloszlás**, amely olyan dolgokra (eseményekre) vonatkozik, amelyek egy meghatározott mintaegységben viszonylag ritkán fordulnak elő. Például, a gyomnövények száma, amelyek egy specifikált mintaterületen fordulnak elő egy kultúrnövény állományban; a gomba spórák száma egy levélen; a gólok száma 90 percben egy futballmérkőzésen. Mivel egy esemény bekövetkezésének valószínűsége alacsony és a zéró érték csonkítja az eloszlást, a Poisson eloszlás tipikusan aszimmetrikus, egy hosszabb elnyúló jobboldali végponttal és egy csúccsal az eloszlás baloldal irányába. A μ kis értékeire a Poisson eloszlás kifejezetten ferde, azonban mihelyt μ nagyobb, mint öt, az eloszlás szimmetrikus (3.5. ábra).

A **binominális eloszlás** a diszkrét eloszlásokhoz tartozik, azaz a változó csupán bizonyos értékeket vehet fel. Például, a pénzfeldobás kimenetele csak fej vagy írás lehet. Folyamatok csupán két lehetséges kimenettel gyakoriak a biológiában, pl. az állatok egy kísérletben vagy életben maradnak vagy elpusztulnak, egy adott növényfaj egy cönológiai felvételezési mintában jelen lehet vagy hiányzik. Ezt az eloszlást először Bernoulli fedezte fel és posztumusz publikálták 1713-ban. A binomiális eloszlást használhatjuk arra, hogy előrejelezzük egy sajátos esemény kimenetelét. Például, egy vetőmagtermelő előrejelzheti a vetőmag tételek számát, amelyek tartalmazzák az élő magok adott arányát.

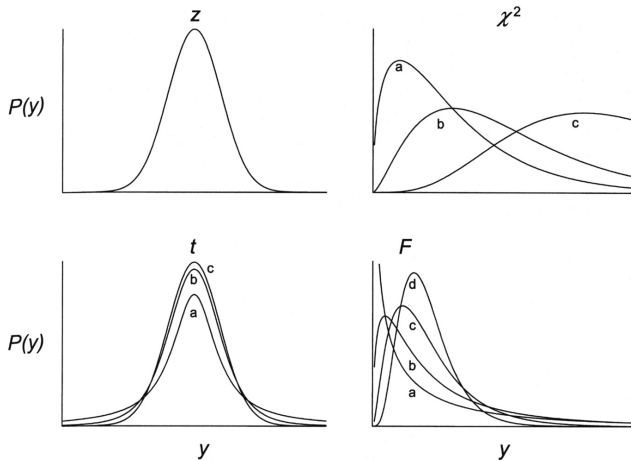
A binomiális eloszlás alakját mindegyik megfigyelés előfordulásának valószínűsége irányítja. Ahol mindegyik megfigyelés előfordulásának valószínűsége egyenlő, azaz $P = 0.5$, az eloszlás csúcsa a tartomány közepén helyezkedik el. Ha azonban két, egymást kölcsönösen kizáró megfigyelés valószínűsége nem egyenlő, akkor a csúcs eltolódik a tartomány egyik vagy másik végébe, ferde eloszlást eredményezve.

A statisztikai jellemzők eloszlása. Gyakran használjuk ezeket az eloszlásokat becslésre és hipotézis tesztelésére. Négy különösen fontos közülük (Quinn és Keough, 2002).

1. A **z vagy normális eloszlás** reprezentálja egy random változó valószínűségi eloszlását, azaz a minta statisztikai jellemző és populáció értéke közötti különbség arányát a populáció statisztikai jellemző szórásához (3.6. ábra).

2. **Student t-eloszlás** reprezentálja egy random változó valószínűségi eloszlását, azaz a minta statisztikai jellemző és populáció értéke közötti különbség arányát a minta statisztikai jellemző eloszlásának szórásához. A t-eloszlás egy szimmetrikus eloszlás, amely nagyon hasonló a normális eloszláshoz, végtelenbe tart mindkét irányban. Alakja még hasonlóbba válik növekvő mintaméretnél (3.6. ábra). A Student t-eloszlás része az F-eloszlásnak, de egyszerűen megkaphatjuk direkt úton is, az átlag standardizálása vezet a t-eloszláshoz. Azok az F-eloszlások, amelyekben a számláló szabadságfoka 1, pontosan megegyeznek a t-eloszlások négyzetével. Egy egyszerű minta statisztikai mutatót átalakíthatunk egy t értékre és felhasználjuk a t eloszlást, hogy meghatározzuk annak valószínűségét, hogy megkapjuk ezt a t értéket (kisebbit vagy nagyobbat) a populáció paraméter specifikált értékére.

3. A χ^2 (khi-négyzet) eloszlás (3.6. ábra) reprezentálja egy változó valószínűségi eloszlását, amely a standard normális eloszlás értékének négyzete. A χ^2 eloszlás értékeit zéró és végtelen határolja. A varianciának χ^2 eloszlása van, így ezt az eloszlást használjuk a populáció varianciák intervallum becslésére. Arra is használhatjuk a χ^2 eloszlást, hogy meghatározzuk annak valószínűségét, hogy megkapjunk egy minta különbséget (kisebbit vagy nagyobbat) a megfigyelt érték és a modell által előrejelzett érték között.



3.6. ábra. Négy gyakori statisztikai mutató valószínűségi eloszlása. A t, χ^2 és F eloszlásokra mutatják az eloszlásokat és az eloszlások alakjának változását három vagy négy szabadságfokra (a-tól d-ig, növekvő sorrendben) (Quinn és Keough, 2002 alapján)

4. Az **F-eloszlás** (3.6. ábra) reprezentálja egy változó valószínűségi eloszlását, azaz két független χ^2 változó arányát, mindegyik osztva a d.f. (szabadságfok) értékével. Mivel a varianciák χ^2 eloszlásúak, az F-eloszlást használjuk a variancia arányokra vonatkozó hipotézisek tesztelésére. Az F-eloszlás értékét zéró és a végtelen határolja. Felhasználhatjuk az F-eloszlást, hogy meghatározzuk annak valószínűségét, hogy kapjunk egy minta variancia arányt (kisebbit vagy nagyobbat) a varianciák közötti valódi arány specifikált értékére.

Ezek az eloszlások táblázatokba vannak foglalva a szakkönyvekben (l. melléklet) és programozva vannak a legtöbb statisztikai szoftverben. Ezeket az eloszlásokat használjuk, hogy reprezentáljuk a minta statisztikai jellemzők (z, t, χ^2 vagy F) eloszlását, mely várható egy populáció vagy populációk ismételt random mintavételéből.

Ha számítógépes programmal végezzük a számolást, akkor nemcsak a statisztikai jellemző kiszámított értékét, hanem a hozzátartozó valószínűséget is megkapjuk. Egyszerűbb programok vagy zsebszámológép használata esetén a kapott statisztikai jellemzőhöz nekünk kell a valószínűséget meghatároznunk. Ehhez táblázatok állnak rendelkezésre a statisztikai szakkönyvekben. Ezek használatának ismeretére mindenkinek szüksége van. Az eloszlások táblázataiban egy érték (egy statisztikai jellemző értékének) valószínűsége rendszerint azt jelenti, hogy annak és a nála nagyobb összes értéknek az előfordulása mennyire valószínű.

4. Kísérletek tervezése

A tervezés minden kísérlet egyik legfontosabb szakasza. E nélkül a kísérletnek nincs iránya, nincs értelme, nincs célja és ezáltal nem fogunk elérni vele semmit. Egy kísérlet mindig egy *gondolattal* kezdődik, amelyet *célokban* kell megfogalmazni. Ezeknek a céloknak világosnak, tömörnek, mérhetőnek és elérhetőnek kell lenniük. A célokat *hipotézisekbe* kell formálni, hogy lehetőség legyen a statisztikai analízisre, amely támogatja bármely konklúzió levonását. A *populációt*, amelyre a kísérlet vonatkozik, ebben a stádiumban kell figyelembe venni. Számba kell venni, hogyan vehetők reprezentatív minták ebből a populációból. A hipotézisvizsgálat szignifikanciáját most kell meghatározni.

A kísérlet tervezési stádiumában el kell dönteni, milyen változókat fogunk mérni, hogyan történik a mérés és milyen módszert fogunk használni az adatok feljegyzésére. A kezeléseket ki kell választani, hogy megfeleljenek a kísérleti céloknak, gondolva a források bármilyen korlátjára.

A kísérletezés egyik fontos alapelve az *ismétlés*. Az ismétlések száma, amely szükséges bármely kísérletben, befolyásolja a kísérletnek azt a képességét, hogy feltárjon kezelésátlag különbségeket, ha azok fennállnak. Az ismétlések számát akkor kell meghatározni, amikor tervezzük a kísérletet.

A *blokkok* kialakítását arra használjuk, hogy csoportosítsuk a kísérleti egységeket együtt homogén sorozatokba, miközben megengedjük, hogy a kísérlet heterogén anyagot (pl. földterületet) használjon. Ez lehetővé teszi, hogy a kísérleti következtetések több helyzetben alkalmazhatók legyenek. Egy kísérletben használt blokkok számát a homogén csoportok lehetséges száma határozza meg, amelyeket a kísérleti anyag képez és nem a szükséges ismétlések száma.

A kezelések elosztása a kísérleti egységekhez befolyásolhatja az eredményt, előnyt adva az egyik kezelésnek a másik hátrányára, ha nem objektíven történik. Ezt elkerülhetjük a kezelések kísérleti egységekre történő elosztásának *randomizálásával*. A randomizáció érvényesíti a kísérlet statisztikai analízisét is.

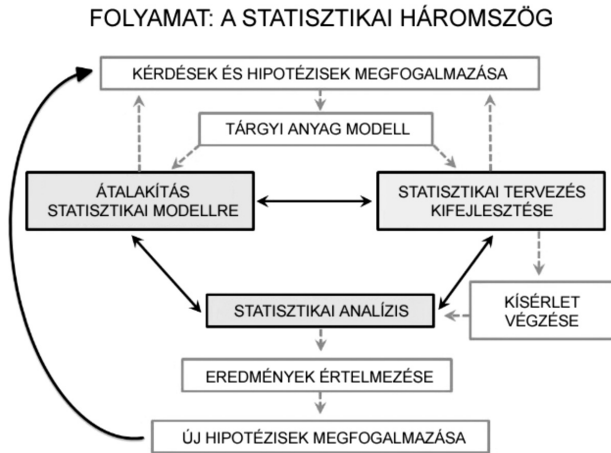
A tudományos kísérletezés logikus lépéseinek folyamatábráját a 4.1. ábra szemlélteti.

Az egész tervezési folyamat során figyelembe kell venni a rendelkezésre álló forrásokat. Néha kompromisszumra van szükség ahhoz, hogy beállítsuk a kísérletet limitált forrásokkal. Figyelembe kell venni az emberi és pénzügyi forrásokat, a rendelkezésre álló időt, a szükséges eszközöket és anyagokat.

4.1. Gondolatok megfogalmazása

A kutatók kérdéseket tesznek fel, amelyek új gondolatokhoz vezethetnek, vagy régi gondolatok átalakíthatók, amelyek viszont megadják az alapvető gondolkodást, amely a *kísérlet* mögött van. A bizonytalan gondolatokat hamarosan átfogalmazzák. Megvitatás a kollégákkal gyorsan kiválasztja, hogy a kérdés marad megválaszolatlan vagy érdemes további vizsgálatokra. Valakinek már lehet egy válasza, amely megfe-

lel a kutatónak, egészen addig, amíg a következő gondolat megérlelődik. Ha nem, a kutató elkezdheti kialakítani a kérdéseket és megfogalmaz specifikus célt vagy specifikus célok sorozatát.



4.1. ábra. A tudományos kísérletezés logikus lépéseinek folyamatábrája a visszacsatolási hurokkal, amely lehetővé tesz új tudományos hipotézis és kísérleti tervezés módosítást a jövőbeni kísérletezéshez. A három központi doboz képezi a statisztikai háromszöget (Casler, 2015 alapján)

4.2. Célok meghatározása

Elsődlegesen, a céloknak reálisnak, mérhetőnek és elérhetőnek kell lenniük. A célokat világosan és egyértelműen kell kifejezni. Sok ember lesz bevonva a javasolt kísérlet tervezésébe és beállításába és rendkívül fontos, hogy a bevont team teljesen megértse a munka célját mindegyik szakaszban. A bizonytalan célok keveredést és félreértést okozhatnak. Végül biztosítsuk, hogy a célok nem tértek el az eredeti gondolattól. Jó megoldás, hogy egy független szakértő ellenőrizze az értelmezhetőséget és a világosságot. A célok valóban azt fejezik-e ki, ami a szándékunk volt?

Fő célok

Néhány kísérletnek csak egy célja van, amely nyilvánvalóan a fő cél. A legtöbb kísérletnek azonban számos specifikus célja van. El kellene döntenie az elején, melyik lesz a fő cél, mivel ennek hatása lesz a jövőbeni döntésekre. Ahhoz, hogy a kísérlet sikeres legyen, semmit sem kellene levonni a fő célból, tekintet nélkül bármely akadályra, amely felmerülhet. Ha egyetlen cél sem tekinthető a legfontosabbnak, akkor meg kell fontolni kisebb kísérletekre bontani vagy újból meghatározni a célokat. Kiseb kísérletekkel a végrehajtandó műveletek kevésbé komplexek lesznek, csökkentve annak esélyét, hogy valamit rosszul végzünk.

Másodlagos célok

A kísérletek költségesek és időigényesek, ezért érdemes másodlagos célokat megfogalmazni. A túl kevés cél elpazarolja a forrásokat. A másodlagos célok értékesebb információt adhatnak, azonban nem szabad megengedni, hogy csökkentsék a fő célokat. A túl sok másodlagos cél azonban elhomályosíthatja a végeredményt, így ezeket óvatosan kell kiválasztani. Ha egy súlyos csapás következik be és a fő célokat fel kell adni, akkor az egyik másodlagos cél előléphet fő célnak és nem veszett el így minden. A kísérlet folytatható, azonban gondolni kell arra, hogy az eredeti célok nem valósíthatók meg. Szükségessé válhat az eredeti cél újrafogalmazása.

Példa. Búzatermés. Fő cél: Meghatározni a nitrogén műtrágyázás szintjét, amely maximális búzatermést ad három fajta természetesekeor. Másodlagos cél: A búza terméskísérlet másodlagos célja lehet meghatározni, hogy a négy fungicid kezelés (C: kontroll, T_1 és T_2 : egyszeri permetezés két időpontban, S_2 : kétszeri permetezés) közül melyik korlátozza legjobban a betegségi szintet betakarításkor az A, B és C búzafajtákban. Ez a cél kiegészítő információt nyújthat a négy fungicid vizsgálatával a három fajtán, a nitrogén műtrágya minden szintjére. Ha valamilyen probléma merült volna fel a nitrogén műtrágya kezelésekre beállításakor, ez a másodlagos cél (fungicidek alkalmazása) mégis információt nyújthat zéró nitrogén műtrágya szinten.

4.3. A populáció és a minta meghatározása

Teljes megbízhatóságot egy kísérletben csak akkor kapunk, amikor minden lehetséges egyed a vizsgált populációban mérni tudunk. A gyakorlatban azonban a teljes populációk többnyire túl nagyok ahhoz, hogy ezt megtehessek és ehelyett a populációnak egy mintáját választjuk ki, hogy képviselje a populációt és a mérések a kiválasztott mintára korlátozódnak. Mintákra van akkor is szükség, amikor a mérések destruktívak, például, a változást a növekvő növény biomasszájában csak a növényállományból egy más után (szukcesszív) vett mintákból lehet meghatározni (növekedésanalízis).

A kutató végső célja következtetések levonása, amelyek jóllehet a minta adatokon alapulnak, kiterjeszthetők (extrapolálhatók) az egész populációra. Például, ha arra következtetünk, hogy egy adott műtrágya növeli a búza mintanövények száraztömegét, jó lenne, ha ezt a következtetést ki lehetne terjeszteni egy sokkal nagyobb búzanövény populációra, nem csupán a kiválasztott minta néhány növényére. Ahhoz, hogy ezt megtehessek, szükséges lesz megállapítani, mennyire pontosan reprezentálja a minta a populációt. A minta adatokat ezáltal arra használjuk, hogy reprezentálják a populáció adatokat és a kiszámított minta mutatókat (statisztikákat) használjuk fel a populáció paraméterek becslésére. További statisztikai módszereket alkalmazunk annak értékelésére, mennyire reprezentálja a minta az egész populációt, amelyből származik és ezáltal mekkora bizalom helyezhető bármely következtetésre, melyet a minta adatokból vonunk le, azonban ezt követően a populációra alkalmazunk.

Egy kísérletből, amely mintákat használ, a levont következtetések érvényessége teljes mértékben attól függ, mennyire jól reprezentálja a kiválasztott minta a populációt. A kérdés, amely ezáltal felmerül az, hogyan tudjuk valóban megítélni, mennyire jól reprezentálja a minta a populációt, amikor nem ismerjük a populáció paramétereit.

A minta-kiválasztás folyamata annak biztosítására, hogy a lehető legpontosabban reprezentálja azt a populációt, amelyből származik, kritikus része a kísérlet tervezési folyamatnak. Két fő akadálya van az érvényes minta-kiválasztásnak, melyen úrrá kell lenni: a vizsgálatot végző torzított minta-kiválasztása és a szisztematikus variáció jelenléte a populációban, melyből a mintát kiválasztottuk. Torzított minta kiválasztás könnyen bekövetkezik, ha a vizsgálatot végző személy nem eléggé körültekintő a mintavételi eljárás során. A szisztematikus variáció kifejezés utal a variációnak a trendjére, amely áthalad a populáción.

A szabadföldi kísérletek területének többségében ritka, hogy egy kísérletet abszolút egységes körülmények között állítunk be. A növénykísérlet példánál maradva, legtöbb szántóföldön növekvő növény ki van téve a klimatikus és környezeti feltételek egész sorának, amely szisztematikus variációt okoz a szántóföldön termesztett növény növekedésében és termelésében. Mind a kísérleti kezelések megosztását a parcellákra és a minták kiválasztását a mérésekhez oly módon kell végezni, hogy az alapul szolgáló variációt figyelembe vegyük és lehetőség szerint különítsük el a specifikus kezeléshatásoktól.

Sor	Kontroll parcellák																			
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	0.50	0.52	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51
2	0.52	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51
3	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51
4	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51
5	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51
6	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51
7	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51
8	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51
9	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51
10	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51
11	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51
12	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51
13	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51
14	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51
15	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51
16	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51
17	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51
18	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51
19	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51
20	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51
Sor	Kezelt parcellák																			
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
21	0.67	0.67	0.67	0.67	0.67	0.67	0.67	0.67	0.67	0.67	0.67	0.67	0.67	0.67	0.67	0.67	0.67	0.67	0.67	0.67
22	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60
23	0.55	0.55	0.55	0.55	0.55	0.55	0.55	0.55	0.55	0.55	0.55	0.55	0.55	0.55	0.55	0.55	0.55	0.55	0.55	0.55
24	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65
25	0.72	0.72	0.72	0.72	0.72	0.72	0.72	0.72	0.72	0.72	0.72	0.72	0.72	0.72	0.72	0.72	0.72	0.72	0.72	0.72
26	0.52	0.52	0.52	0.52	0.52	0.52	0.52	0.52	0.52	0.52	0.52	0.52	0.52	0.52	0.52	0.52	0.52	0.52	0.52	0.52
27	0.48	0.48	0.48	0.48	0.48	0.48	0.48	0.48	0.48	0.48	0.48	0.48	0.48	0.48	0.48	0.48	0.48	0.48	0.48	0.48
28	0.67	0.67	0.67	0.67	0.67	0.67	0.67	0.67	0.67	0.67	0.67	0.67	0.67	0.67	0.67	0.67	0.67	0.67	0.67	0.67
29	0.64	0.64	0.64	0.64	0.64	0.64	0.64	0.64	0.64	0.64	0.64	0.64	0.64	0.64	0.64	0.64	0.64	0.64	0.64	0.64
30	0.59	0.59	0.59	0.59	0.59	0.59	0.59	0.59	0.59	0.59	0.59	0.59	0.59	0.59	0.59	0.59	0.59	0.59	0.59	0.59
31	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60
32	0.68	0.68	0.68	0.68	0.68	0.68	0.68	0.68	0.68	0.68	0.68	0.68	0.68	0.68	0.68	0.68	0.68	0.68	0.68	0.68
33	0.62	0.62	0.62	0.62	0.62	0.62	0.62	0.62	0.62	0.62	0.62	0.62	0.62	0.62	0.62	0.62	0.62	0.62	0.62	0.62
34	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50
35	0.55	0.55	0.55	0.55	0.55	0.55	0.55	0.55	0.55	0.55	0.55	0.55	0.55	0.55	0.55	0.55	0.55	0.55	0.55	0.55
36	0.61	0.61	0.61	0.61	0.61	0.61	0.61	0.61	0.61	0.61	0.61	0.61	0.61	0.61	0.61	0.61	0.61	0.61	0.61	0.61
37	0.63	0.63	0.63	0.63	0.63	0.63	0.63	0.63	0.63	0.63	0.63	0.63	0.63	0.63	0.63	0.63	0.63	0.63	0.63	0.63
38	0.61	0.61	0.61	0.61	0.61	0.61	0.61	0.61	0.61	0.61	0.61	0.61	0.61	0.61	0.61	0.61	0.61	0.61	0.61	0.61
39	0.67	0.67	0.67	0.67	0.67	0.67	0.67	0.67	0.67	0.67	0.67	0.67	0.67	0.67	0.67	0.67	0.67	0.67	0.67	0.67
40	0.63	0.63	0.63	0.63	0.63	0.63	0.63	0.63	0.63	0.63	0.63	0.63	0.63	0.63	0.63	0.63	0.63	0.63	0.63	0.63

Populáció átlag: Kontroll átlag: 0.558 kg; Új műtrágya kezelésátalaga: 0.641 kg

Feltételezve 15 növényből álló mintaméretet, három mintavételi módszer eredménye a következő:

(a) A növények kiválasztása mindegyik parcella közepéből.

Kontroll minta átlag: 0.524 kg

Új műtrágya minta átlag: 0.603 kg

(b) Egyforma helyzetű növények kiválasztása mindkét kezelésből.

Kontroll minta átlag: 0.607 kg

Új műtrágya minta átlag: 0.643 kg

(c) A növények teljesen random kiválasztása, azonban kihagyva a szegélysorokat.

Kontroll minta átlag: 0.553

Új műtrágya minta átlag: 0.605 kg

4.2. ábra. Példa a mintavétel különböző módszereinek hatására (Ireland, 2010)

A minta kiválasztásának problémáját és fontosságát illusztrálja a 4.2. ábra. A nagyon egyszerű szántóföldi kísérlet 40 burgonyával tartalmaz, 20 növényvel soronként. A sorok felét standard műtrágyával kezelték (kontroll) és a másik felét egy új műtrágyával. Mindenegyes növény természetét (kg) mutatja a 4.2. ábra. Tétélezzük fel,

hogy nem lehet mérni minden növény termését, ezáltal mintavétel szükséges. Három lehetséges mintavételi módszert mutatunk be, a populáció átlagértékével együtt. Nem meglepő, hogy mindegyik mintavétel a populációátlag eltérő becslését adja; a mintaátlagok jelentősen változnak és egyik sincs különösen közel a tényleges populációátlaghoz.

A minta kiválasztás első aranyszabálya, hogy a torzított mintavétel elkerülése céljából a mintáknak amennyire lehetséges, randomnak kell lenni. A statisztikai eljárások alkalmazása a populáció paraméterek becslésére a minta adatokból csak akkor érvényes, amikor randomizált mintákat használunk. Ha nem azonosítható szisztematikus variáció van a populációban, akkor a mintákat teljesen véletlenszerűen (random) kell venni. Ezen az elven alapuló kísérleti elrendezést *teljesen véletlen elrendezésnek* (CRD) hívjuk. Számos útja van a random kiválasztásnak. Meg kell azonban jegyezni, hogy még a teljesen randomizált elrendezésnél sem vonjuk be a kísérleti parcella szélén levő növényeket, hogy elkerüljük a lehetséges „szegélyhatást”, amely torzítja az adatokat.

Ahol egy szisztematikus variáció jelenlétét azonosítjuk egy populációban, a teljesen randomizált eljárás a minták kiválasztására nem javasolt. A megoldás a randomizált blokk elrendezés (RBD) használata. A randomizált blokk elrendezésű kísérletben a populáció fel van osztva egyenlő méretű blokkokra. A kezeléseknél egy teljes sorozatát random helyezzük el minden blokkban, azonban eltérő randomizációt alkalmazunk minden blokkban. Ez az eljárás biztosítja, hogy a kezelésminták magukba foglaljanak olyan növényeket, melyek elterjedtek az egész populációban, azonban egyidejűleg magában foglal elegendő méretű randomizációt, hogy eltávolítsa a torzított kiválasztás lehetőségét.

Az *ismétlés szükségessége* a minta kiválasztásában egészen nyilvánvaló. Ha bizonyos lehetnének abban, hogy minden egyed egy populációban ugyanolyan méretű, és a kísérleti kezelést követően minden egyed ugyanolyan új méretet ér el, akkor nem volna szükség az ismétlésre és egyetlen egyed a mérések előtt és után elegendő lenne. Minden populációban azonban van egy jellemző természetes variáció az egyedek között és ezt a jellemző variációt súlyosbítja bármely nem egyöntetűség jelenléte a kísérleti körülményekben. Ahhoz, hogy egy jó becslést kapjunk a tényleges méretnek, számos egyedet kell megmérnünk a populációból.

Van egy másik okunk arra, miért van szükségünk ismételt mérésekre. Tudni akarjuk, hogy a kísérletben van-e valódi szignifikáns különbség két vagy több eltérően kezelt populáció között, vagyis egy olyan különbség, amely a kísérleti kezeléseknél tulajdonítható, sokkal inkább, mint a random variációnak. Ahhoz, hogy választ kapjunk erre a kérdésre, elvégzünk egy statisztikai analízist mindegyik populációból vett minta méréseken. Az ilyen analízis, melyet statisztikai szignifikancia vizsgálatnak hívunk, függ a variancia nagyságának kvantitatív ismeretétől, amely jelen van minden mintában. A variáció nagyságát mintán belül természetesen csak akkor határozhatjuk meg, ha a minta bizonyos számú ismételt megfigyelést tartalmaz. Továbbá, a statisztikai eljárás a minta megbízhatóságának meghatározására, azaz mennyire közel várható a minta átlag a populáció átlaghoz, ugyancsak függ a mintától, amely számos ismétlést tartalmaz.

A statisztikai analízis nyújthat bizonyos segítséget a megfelelő mintaméret kiválasztásában. A minta megbízhatóságát a *standard hibával* (standard error) határozzuk meg. Bármely kísérletben a standard hibát, amely természetesen csökken a növekvő mintamérettel, a lehető legalacsonyabban kell tartani.

4.4. Hipotézisek megfogalmazása

A populációkra (sokaságokra) vonatkozó különféle feltevéseket *hipotéziseknek*, az azok helyességének mintavételi eredményekre alapozott vizsgálatát pedig *hipotézisvizsgálatnak* nevezzük. A hipotézisek a vizsgált sokaság(ok) eloszlására vagy az adott eloszlás(ok) egy vagy több paraméterére vonatkozhatnak. A különféle hipotézisek vizsgálatára szolgáló eljárásokat *próbáknak* (más néven teszteknek) hívjuk.

A hipotézisvizsgálat lényegében mindig annak mérlegelése, hogy a vizsgálat tárgyának megfelelő sokaságra vonatkozó valamely állítás mennyire hihető a mintavétel eredményeinek tükrében. Ha a mintavétel eredménye összhangban áll a sokaságra vonatkozó feltevessel, akkor nincs okunk arra, hogy kétségbe vonjuk a feltevés helyességét. Ellenkező esetben azonban minden okunk megvan arra. Így az első esetben az állítást célszerű fenntartani, mert az valószínűleg (de nem biztosan!) igaz, a második esetben pedig célszerű azt elvetni, mert az valószínűleg nem igaz.

Egy *hipotézis kifejezi* a változót, a statisztikát és az összehasonlítást. Minden célnak van egy *nullhipotézise* és egy *alternatív hipotézise*. A hipotézis a cél újbóli kifejezése statisztikai módon. Ennek ki kell fejeznie a (minket érdeklő) változót, a vizsgált statisztikai mutatót és a (tervezett) összehasonlítást. Vegyünk figyelembe egy búzatermesztési kísérletet és azt a hipotézist, hogy „nincs különbség az A, B és C búzafajták termésében”. Ez egy valós hipotézis; a bennünket érdeklő változó a termést, a tervezett összehasonlítás a három búzafajta és a vizsgált statisztikai mutató az átlag. A statisztikai mutató egy összefoglaló érték, amelyet egy mintában végzett megfigyelésekből számítunk ki. Variancia, szórás és átlag példák a statisztikai mutatóra.

A cél kifejezése statisztikai hipotézisként megengedi a formális statisztikai vizsgálatot. Minden hipotézisvizsgálat céljára egyszerre két hipotézis, egy ún. nullhipotézis és egy azzal szembenálló másik állítás, egy ún. alternatív hipotézis fogalmazandó meg. Erre azért van szükség, mert a hipotézisvizsgálat során a kettő közül azt fogjuk igaznak tekinteni, elfogadni, amelyik a mintavétel eredménye alapján hihetőbbnek tűnik a másiknál. A nullhipotézis jelölésére a H_0 , az alternatív hipotézis jelölésére pedig H_1 szimbólumot használjuk. A két hipotézist oly módon kell megfogalmazni, hogy azok (a) akármelyikét is tekinthetjük majd hihetőbbnek a másiknál, megválaszolható legyen a bennünket érdeklő kérdés, (b) egyszerre ne lehessenek igazak.

Az a hipotézis, amelynek helyességéről közvetlenül döntünk, mindig a nullhipotézis. Ha azonban a nullhipotézis és az alternatív hipotézis kölcsönösen kizárják egymást, akkor a H_0 hipotézisre vonatkozó döntés közvetetten mindig döntést jelent a H_1 -re vonatkozóan is: a nullhipotézis elfogadása egyúttal H_1 elvetését, a H_0 elvetése egyben H_1 elfogadását is jelenti.

A *nullhipotézis* annak állítása, hogy a szóban forgó statisztikai mutatónak van egy sajátos értéke. Az *alternatív hipotézis* kifejezi azt, hogy a statisztikai mutató értéke eltér a nullhipotézisben specifikált értéktől.

4.5. Hipotézisvizsgálat

A statisztikai következtetés egyik leggyakrabban alkalmazott fajtája a hipotézisvizsgálat, amely előre felállított, kérdés formájában megfogalmazott hipotézisre ad igen vagy nem választ. Ez a válasz természetesen tartalmaz bizonyos valószínűséggel

jellemzett bizonytalanságot (Hajtman, 2012). A hipotézisvizsgálat leggyakoribb kérdéstípusai a következők: történt-e valamilyen változás az általunk vizsgált beavatkozás (kezelés) hatására; van-e a vizsgálatban szereplő két vagy több csoport között különbség a vizsgált változó szempontjából; van-e kapcsolat a két változó között.

A hipotézisvizsgálat módszere a statisztikai próba. A statisztikai próba úgy működik, hogy kiszámítunk egy statisztikai jellemzőt, majd ezt az értéket egybevetjük az illető statisztikai jellemzőnek a nullhipotézis igaz volta esetén érvényes eloszlásával. Megnézzük, mennyire valószínű, hogy ezt az értéket kapjuk, ha a hipotézis-párból a nullhipotézis, a hatás tagadása igaz (Hajtman, 2012). Szükségünk van tehát a hatást mérő statisztikai jellemző nulla körüli eloszlására. A statisztikai próbák többségében a négy eloszlás: a normális eloszlás, a χ^2 -eloszlás, az F-eloszlás és a t-eloszlás.

A hipotézis helyességének ellenőrzése céljából a próbafüggvény lehetséges értékeinek tartományát alkalmas osztópontok segítségével két egymást át nem fedő részre bontjuk: egy *elfogadási tartományra* és egy *visszautasítási* – más néven *kritikus* – *tartományra*. E két tartomány határait úgy választjuk meg, hogy a próbafüggvény a nullhipotézis fennállása esetén előre megadott nagy valószínűséggel az elfogadási tartományba essen. A kritikus tartományba esés α valószínűségét (vagy ha azt százalékos formában fejezzük ki, akkor annak százszorosát) *szignifikancia szintnek* nevezzük. A szignifikancia-szint megválasztásával kapcsolatban megjegyezzük, hogy azt általában kicsinek – 0.05-nek vagy ennél is kisebb értéknek – szokás választani a gyakorlatban.

A valóságnak a nullhipotézisben rögzített állapottól való meghatározott irányú eltérései egyoldali alternatív hipotézisként írhatók fel. Ha az *egyoldali alternatív hipotézis* fennállása esetén a próbafüggvény alacsonyabb értéket vesz fel, mint H_0 fennállásakor, *baloldali*, ellenkező esetben *jobboldali alternatív hipotézisről* beszélünk. A valóságnak a nullhipotézisben rögzített állapottól való tetszőleges irányú eltérései kétoldali alternatív hipotézisként fogalmazhatók meg. A *kétoldali alternatív hipotézis* fennállása esetén a próbafüggvény értéke akár kisebb, akár nagyobb lehet, mint H_0 fennállásakor (Hunyadi et al., 2000).

A hipotézisvizsgálat során kétféle hiba követhető el: I. típusú (elsőfajú) és II. típusú (másodfajú) hiba. Elsőfajú hibát akkor követünk el, ha elvetjük a H_0 nullhipotézist, pedig az a valóságban igaz. Másodfajú hiba elkövetésére ezzel szemben akkor kerül sor, ha elfogadjuk a nullhipotézist, pedig az a valóságban nem igaz. Másképpen fogalmazva: az elsőfajú hiba H_0 téves elvetése, a másodfajú hiba pedig H_0 téves elfogadása (4.1. táblázat).

4.1. táblázat. Az adatokból levont következtetések, kifejezve a próba erejében, szignifikanciában és I. típusú és II. típusú hibában (Steel és Torrie, 1980)

A populáció valós állapota	A minta adatokból levont következtetések	
	Elfogadjuk a nullhipotézist és H_0 -t helyesnek mondjuk	Visszautasítjuk a nullhipotézist és H_1 -et mondjuk korrektnek
A nullhipotézis, H_0 korrekt	A következtetést $1-\alpha$ valószínűséggel tesszük	I. típusú (elsőfajú) hibát követünk el A következtetést α valószínűséggel végezzük, amely a szignifikancia szint
A nullhipotézis téves, H_1 korrekt	II. típusú (másodfajú) hibát követünk el A következtetést β valószínűséggel végezzük	A következtetést $1-\beta$ valószínűséggel végezzük, amely a próba ereje

Az elsőfajú hiba elkövetésének valószínűsége megegyezik az α szignifikancia-szinttel, a másodfajú hiba elkövetésének β valószínűsége azonban csak akkor határozható meg számszerűen, ha pontosan tudjuk azt, hogy H_0 helyett a valóságban milyen egyszerű hipotézis áll fenn.

4.6. Várható kezeléskülönbségek

Fogalmazzunk meg további hipotéziseket, hogy teszteljük a várható kezeléskülönbségeket, amikor tervezzük a kísérletet. Különbségek a specifikus kezelések vagy kezeléscsoportok között különösen érdekesek lehetnek és ezek a fontos összehasonlítások származtathatók a kísérlet kezeléseinek ismeretéből. Például, a növényi betegség szintjei a kezeletlen kontroll parcellákon várhatóan különböznek a fungiciddal kezelt parcellák betegség szintjeitől. Ilyen esetekben fontos lehet további nullhipotéziseket megfogalmazni a tervezési stádiumban. Ez lehetővé teszi a kezelés átlagokban a várható különbségek részletesebb vizsgálatát.

További hipotézisek használhatók, hogy becsüljük a kezeléskülönbségeket görbevonalú hatások (reakciók) esetén, ha a kezelések összefüggő kezelések sorozatát képezik, mint amilyen egy kemikália növekvő dózisa vagy takarmány adalékok különböző mennyiségei a szarvasmarhák takarmányozásánál. Például, a növekvő nitrogén dózisek hatása a termésszintekre egyenes vonalú hatás (reakció) vagy görbe vonalú, vagyis a termésnövekedés fokozatosan csökken ahogy a dózis növekszik? Ha görbe vonalú, milyen mértékben értelmezhető egy másodfokú görbével vagy más magasabb fokú polinom függvényvel, mint amilyen a harmadfokú görbe?

Ha egy hipotézist egy specifikus kezelés összehasonlításra megfogalmazzunk a tervezési stádiumban, a hipotézist nem befolyásolják az adatok, mivel azok még nem ismertek. A hipotézis megfogalmazás teljesen a kezelések ismeretén alapul és nem a megfigyelt kezeléskülönbségek nagyságán. Ez azt jelenti, hogy annak az esélye, hogy visszautasítjuk a nullhipotézist és megállapítjuk, hogy kezeléskülönbség fennáll, egyenlő az I. típusú hiba valószínűségével, amely egyenlő a teszt szignifikancia szintjével.

A *tervezett összehasonlítások* ellentétként (kontraszt) ismertek és a görbevonalú hatások vizsgálatára tervezettek polinomiális ellentétként (kontraszt) ismertek. Ha egy-nél több tervezett összehasonlítást használunk, akkor mindegyiknek *kölcsonösen függetlennek*, vagy más szavakkal kölcsonösen *ontogonálisnak* kell lenni. A tervezett összehasonlítások számának nem szabad többnek lenni, mint a rendelkezésre álló szabadságfokok száma. A *szabadságfokok* nagysága visszatükrözi a „megfigyelések” számát, amelyek szabadon változnak, miután bizonyos korlátokat vezetünk be az adatokra. Például, ha n megfigyelésünk van és az átlag érték ismert, $n-1$ megfigyelés bármilyen értéket vehet, azonban az n -edik megfigyelés értékét meghatározza az $n-1$ megfigyelés értéke és átlaga.

4.7. Változók

Bármely kísérletben mérni fogjuk a változókat egyik vagy másik formában. A legérdekesebb változók azok, amelyekre a kísérletben alkalmazott kezelések hatással vannak. Ezeket *reakció változóknak* (response variable) hívjuk és értékük változhat egyik kísérleti egységtől a másikig. Például reakció változó a súlygyarapodás a takar-

mányozási kísérletben és termés a fajtakísérletben. Más változókat is mérhetünk, ezek lehetnek háttér változók vagy kovariánsok.

Szükséges dönteni kezdetben arról, hogy mi lesz a reakció változó. Ezt a fő célkitűzés határozza meg. A fő változónak képesnek kell lenni megválaszolni az eredeti kérdést és lehetséges, hogy a mért változók sorozatából vezetjük le. További reakció változókat kell mérni, hogy eleget tegyünk a másodlagos céloknak. Ezek ugyancsak lehetnek származtatott változók.

Miután kiválasztottuk a mérni kívánt változókat, óvatos megfontolás szükséges a mérések gyakoriságának és időpontjának meghatározásakor. A méréseket úgy kell végezni, hogy az eredmények a lehető legszélesebb körben legyenek felhasználhatók. A célok gyakran meghatározzák, hogy a méréseket bizonyos intervallumokban végezzük az egész kísérlet folyamán vagy csak a kísérlet végén. Bizonyos esetekben a mérések idejét és gyakoriságát meghatározza maga a változó. Bizonyos kísérletekben lehet, hogy nem világos, mikor érjük el a befejezés pontját. A befejezési pontot világosan definiálni szükséges a kísérlet tervezésekor.

Néhány változót azért mérünk, hogy további háttér információt adjon, amely segít az interpretációban. Ezekre a változókra lehetséges, hogy nincs közvetlen hatása a vizsgált változónak. A kísérletek tervezésekor ezeket a változókat óvatosan kell kiválasztani, nehogy csökkentsék a reakció változókat vagy elpazarolják a forrásokat. Számba kell venni, hogy információ nyerhető-e más forrásból; csapadék adatok rendelkezésre állnak-e a Meteorológiai Szolgálatól. További információval rendelkezhetünk az egyes kísérleti egységekről, kiegészítő változók értékeinek formájában, melyekről feltételezzük, hogy valamilyen módon kapcsolatban vannak a reakció változóval. Az ilyen változókat kovariánsoknak hívjuk. Egy vagy több kovariáns használata az analízisben segített eltávolítani a variáció egy részét a kísérleti egységek között.

A mérési technikáknak megegyezőnek kellene lenni a kísérlet minden helyén és minden évében. A potenciális megfigyelő (felvételező) hatást figyelembe kell venni a kísérlet tervezése során. A kísérlet tervezési stádiumában dönteni kell az alkalmazni kívánt mérési módszerekről és bármely hatást, amelyet a reakció változóra kifejtethet, figyelembe kell venni.

A mintavételi eljárás változik a mért változónak megfelelően és a méréshez használt műszerek korlátaitól függően. Ne túlozzuk el az eredmények pontosságát. A számítások után kerekítjük az eredményeket, ne pedig azt megelőzően. Jegyezzünk fel minden hiányzó értéket egy megfelelő kóddal, mint amilyen a csillag (*).

A becsléseket olyan fokú precizitással (tizedes helyek, vagy jelentős számjegyek) kell feljegyezni, amely megfelel a használt műszernek. Bármely számítás eredményét megfelelő fokú precizitással kell közölni. Minden decimális (tizedes) helyet fel kellene használni a végső eredményig, hogy minimalizáljuk a kerekítési hibákat. Az eredmények rögzíthetők elektronikus formában vagy füzetekbe. Az elektronikus adatrögzítés megfelelőbb módszer az adatok bevitelére, jóllehet megfelelő biztonsági másolat (back-up) is szükséges.

4.8. A kezelések kiválasztása

Számos kísérlet célja, hogy összehasonlítsuk, hogyan különbözik a reakció változó a különböző kezelésekből. Ezek lehetnek kapcsolatban nem levő kezelések, ahol nincs

természetes sorrendje a kezeléseknek, mint a különböző fungicidek listája, melyeket használni fogunk. Alternatív módon, a kezelések kapcsolatban lehetnek, a kezelések meghatározott sorrendjével, mint egyetlen fungicid alkalmazása különböző dózisokban. E két helyzet kombinációja ugyancsak használható. Ilyen esetben a kezelés a fungicid és a dózis kombinációja lehet. Egy kezelés bármi lehet, amely alkalmas szabályozott applikációra a kísérlet követelményeinek megfelelően.

A kezelés faktorok és faktorszintek kiválasztása függ a kísérleti céloktól. Minden kezelés faktor kombináció nem mindig lehet praktikus. Ahol két vagy több kezelés faktort használunk, az interakció a szintjeik között, érdekes lehet. A kezelés faktorok és faktor szintek végső kiválasztásánál figyelembe kell venni a legmegfelelőbb statisztikai tervezést, a mezőgazdasági praktikumot és az interpretációt. A kiválasztott faktor szintek függenek a céloktól. Ha a kísérlet célja a kezelések összehasonlítása, a használt faktor szintek különböznek attól, amelyet akkor használunk, ha a kísérlet célja megvizsgálni az összefüggést a reakció változó és a kezelés faktor szintjei között.

Amikor a kezelés szintek egy kezelés faktorban növekvő dózisokat képviselnek, lehetséges megvizsgálni egy trendnek a meglétét a növekvő kezelésszintekre adott reakcióban. Ha csak két kezelésszintet használunk, akkor csak egy egyenes vonalú reakciót vizsgálhatunk. Minimálisan három szinttel a görbe vonalú hatások vagy trendek fennállása vizsgálható, jöllehet a kezelésszintek száma korlátozza bármely trendnek a komplexitását, amely vizsgálható. Másodfokú trendek azonosításához minimálisan három kezelésszint szükséges. Egy standard vagy null kezelést (kontroll kezelés) magában foglalhat a kísérlet, ha az összehasonlítások a kontroll és a kísérleti kezelések között érdekesek, vagy a kontroll kezelések hasznos háttér információt adnak, mint amilyen a betegség szintje kezelés nélkül. Bizonyos esetekben speciális gondoskodást igényel a tervezéskor, jöllehet a legtöbb esetben extra (külön) kezelésként vagy kezelésként vannak számba véve. Ha a kontroll kezelést háttér információként használjuk, akkor nem szükséges, hogy a többi kezeléssel együtt analizáljuk; csupán arra használjuk, hogy segítsen interpretálni az eredményeket. A kontroll kezelések előzetes ismeretének mértéke, a kísérlet célja és a gyakorlati szempontok fogják meghatározni a bevonható kontroll kezelések számát.

4.9. Ismétlés

A kísérletek egyik alapelve az ismétlés. Statisztikai értelemben az ismétlést rendszerint úgy definiáljuk, mint mindegyik kezelés alkalmazását egy vagy több kísérleti egységre. Rendszerint két kísérleti egység nem adja ugyanazt az eredményt, még akkor sem, ha ugyanazt a kísérleti kezelést kapta mindegyik. A biológiai adatok változóak, így ahhoz, hogy a kísérlet tudományosan és statisztikailag érvényes legyen, ismétléseket kell alkalmazni.

Fontos megjegyezni, hogy mindenegyik kezelés ismétlése szükséges a kísérleti egységeken, sokkal inkább, mint ismételt mérések egyetlen kísérleti egységen. Néhány esetben a kezeléseket nem egyenlő számban ismétljük meg, jöllehet rendszerint minden kezelést egyenlően ismétlünk meg. A nem egyenlő ismétlés magában foglalhat komplex statisztikai tervezést, amely bonyolultabb analízist és interpretációt eredményez. Ezek a tervezések néha szükségesek, különösen amikor a források korlátozottak.

Ha körültekintően tervezünk és végezzük az analízist, amely komplexebb lehet, azonban az interpretáció lehet egyszerű.

Esetenként egy ismétléses kísérleteket állítunk be. Ezekben az esetekben a kezelés hatását össze kell hasonlítani a háttér variabilitás bizonyos becslésével. A variabilitásnak ezt a mérését kezdetben kell meghatározni és rendszerint néhány kezelés-kombinációból áll, amely nem várható, hogy bármilyen hatást mutasson, pl. magasabb rendű interakciók. Ezek a kísérletek rendszerint nagy több tényezős faktoriális kísérletek, amelyekben a magasabb rendű interakciókat összevonjuk, hogy a kísérleti hibának egy becslését adják. A gyakorlatban a magasabb rendű interakciók rendszerint nem szignifikánsak és nincs gyakorlati értelmezésük, így igazolt az interakciók összevonása és hiba tagként való használatuk. Az interakció tagok összevonásának módszerét, hogy hiba tagot képezzenek, akkor is használhatjuk, amikor kevés maradék (reziduális) szabadságfok van, jóllehet bármilyen analízis megkezdése előtt dönteni kell arról, hogy melyik interakciókat vonjuk össze. Egy probléma ezzel a módszerrel az, hogy ha egy interakció tag szignifikáns és összevonjuk a hiba tagba, akkor a szignifikáns interakció eredményt nem találjuk meg. A módszer egyik előnye, hogy több kezelés faktort használhatunk anélkül, hogy sokkal több kísérleti egységre lenne szükségünk; egyetlen ismétléses kísérlet végezhető, felhasználva a nem fontos interakció tagokat a hiba tagként, hogy a bennünket érdeklő kezeléshatásokat megvizsgáljuk.

Az ismétlésnek két alapvető funkciója van a kísérletekben. Mivel a kísérletek többsége nem tudja mérni az egész populációt, mintát veszünk és a mintát használjuk fel arra, hogy becsüljük a populáció különböző jellemzőit, pl. a kezelés populáció átlagot a minta átlaggal becsüljük. A becslés javul, ahogy több megfigyelést végzünk, azaz a populáció átlag pontosabb becslését érjük el a nagyobb minták vételével. Más szavakkal, a kísérletnek nő az ereje (a hatásfoka) a kezelés-különbségek feltárásában, ha az ismétlések száma nő, miközben minden más körülmény azonos marad.

A természetes variáció a kísérleti egységek között azt jelenti, hogy két hasonló kísérleti egység, amely ugyanazt a kezelést kapja, nem adja ugyanazt a reakciót. Ez a természetes variáció *kísérleti hibaként* vagy háttér variációként ismert. Ennek a természetes variációnak a mérését megkaphatjuk, mint egy varianciát, amelyet hiba varianciának vagy maradék (reziduális) varianciának hívunk. A kísérleti hiba vagy háttér variabilitást összehasonlítjuk a kezelés variabilitással, hogy meghatározzuk, van-e kezeléshatás. Nem lehetséges mérni a kísérleti hibát ismétlés használata nélkül és így ismétlés szükséges a kezeléshatások teszteléséhez.

Egy kezelés ismétléseinek száma az egyik faktor, amely meg fogja határozni, hogy vajon egy kezelés-különbség kimutatható lesz-e, ha fennáll. Túl kevés ismétlés a szükséges különbségek kimutatásának képtelenségéhez vezet, feltéve, hogy a különbségek ténylegesen fennállnak. Túl sok ismétlés használata ugyancsak a források pazarlását eredményezi. A szükséges ismétlések számát meghatározza a teszt indokolt ereje, a kezelésátlagok közötti különbség nagysága, amelyet kimutatni akarunk, az alkalmazott teszt szignifikancia szintje és az adatok háttér variabilitása.

Adott szignifikancia szintre, hiba variációra és kimutatandó kezelés-különbség nagyságra a teszt ereje nő az ismétlések számának növekedésével. Egy specifikált szignifikancia szintre, teszt erejére, és a reakció változó kezelés átlagai közötti különbségre, a különbségek kimutatásához szükséges ismétlések száma függ a *reakció vál-*

tozó variabilitásától. Minél nagyobb a variabilitás, annál nagyobb számú ismétlésre van szükség. A variancia becslés konstans lesz egy adott reakció-változóra. Az ismétlések szükséges száma egy specifikált különbség kiszámításához egy adott szignifikancia szintre és variabilitás szintre nő, ahogy a kimutatandó különbség nagysága csökken. Az összefüggés azonban a különbség nagysága és az ismétlések száma között nem lineáris. Ha a szükséges ismétlések száma megkétszereződik, a különbség, amelyet kimutathatunk, nem feleződik.

Egy kompromisszumot azonban elérhetünk a kimutatandó különbség nagysága, a statisztika ereje, a teszt szignifikanciája és a vizsgálni kívánt kezelések száma között abból a célból, hogy megfelelő ismétlést kapjunk. Például, ha a kezelések költségesek, vagy csak kevés állat áll rendelkezésre vagy korlátozott földterület van, akkor kompromisszumot kell elérni.

Fontos, hogy elérjük a lehető legjobb *becslését a hiba varianciának*. Adott körülmények között (szignifikancia szint, statisztikai erő, háttér variabilitás és a kezelés különbség nagysága) lesz egy minimális számú ismétlés, amely szükséges ahhoz, hogy különbséget mutassunk ki a kezelés átlagértékekben. Ugyancsak lesz egy ismétlésszám, amely felett nincs gyakorlati vagy statisztikai előny. Az ismétlések kiválasztott számának e két extrém érték között kell elhelyezkedni, hogy az adott kísérlet hatékony legyen statisztikai és ökonómiai oldalról egyaránt.

Jóllehet számos változót mérhetünk egy kísérletben, mégis rendszerint a fő reakcióváltozót használjuk fel a kísérletben szükséges ismétlések számának meghatározásához. Abban az esetben, ha több változó egyformán fontos, a legvariábilisabb változót használjuk fel az ismétlések szükséges számának meghatározásához. A reakcióváltozó hiba varianciáját, vagy kísérleti hibáját használjuk referenciaként, amellyel szemben teszteljük a minta átlagok közötti különbségeket és ez az a variancia, amelyet felhasználunk a szükséges ismétlések számának kiszámításához.

Mennyi ismétlésre van szükség? Soha sincs csak egy válasz erre a kérdésre. A szóban forgó kísérlet számos sajátosságától függ. Gyakran az információ-szükséglet az ismétlések számának kiszámításához több forrásból áll rendelkezésre. Mielőtt bármely próbálkozást tennénk, a kérdés megválaszolására egy döntést kell hozni arról, hogy a kezelések között mekkora reakciókülönbség érdekes (jelentős). A különbség nagyságát meghatározzák gyakorlati szempontok, és gyakorlati jelentőségűnek kell lenni. Például, nincs gyakorlati értéke találni 0.01 t ha^{-1} különbséget a búza termésében. Nincs szabály ennek meghatározására, csupán a józan ész, azonban akadályozhatja a használt felvételező (mérő) eszközök pontossága. Háttér variáció, mint aminek a búzatermés különbsége tulajdonítható két egyforma méretű földterületen (parcellán), még akkor is, amikor nem alkalmaztunk kezelést, egy másik tényező és számos mód van ennek kifejezésére.

A variabilitás legáltalánosabb mérése a minta *variancia*, amely információt ad arról, hogyan csoportosulnak az egyedi minta megfigyelések mintaátlaguk körül. A hiba variancia megbízható becslésének a reakcióváltozóra vonatkozóan rendelkezésre kellene állni. A hiba variancia „méri a variabilitást” a megfigyelések között, nem értelmezett okoknak vagy kísérleti hibának tulajdoníthatóan és nevezhetjük reziduális varianciának.

A hiba variabilitás becslését megkaphatjuk számos forrásból, beleértve az irodalmat és más történelmi forrásokat. Gyakran egy előző kísérletet tekintünk reak-

cióváltozónak hasonló körülmények között, és a hiba variabilitás becslése ebből a kísérletből felhasználható. Abban a kevés esetben, ahol nincs irodalom és nincs adat előző kísérletről, egy *pilot (próba)* tanulmány végezhető, hogy megadja ezt az információt. Egy *pilot (próba)* tanulmány a kísérlet kis skálájú lefuttatása, hogy információt kapjunk a variabilitásról, és teszteljünk bármely új applikációs módszert a fő kísérlet beállítása előtt. Még további két információ szükséges, mielőtt válaszolni tudnánk arra a kérdésre, hogy mennyi ismétlés szükséges. Ezek a teszt szignifikanciája (annak az állításnak a valószínűsége, hogy különbség áll fenn, amikor nincs különbség) és a teszt ereje (annak az állításnak a korrektsége, hogy bizonyíték van a különbség létezésére). Mihelyt ezekről döntöttünk kiválaszthatunk egy megfelelő statisztikai vizsgálatot, amelyet azután felhasználhatunk a minta méretének meghatározására.

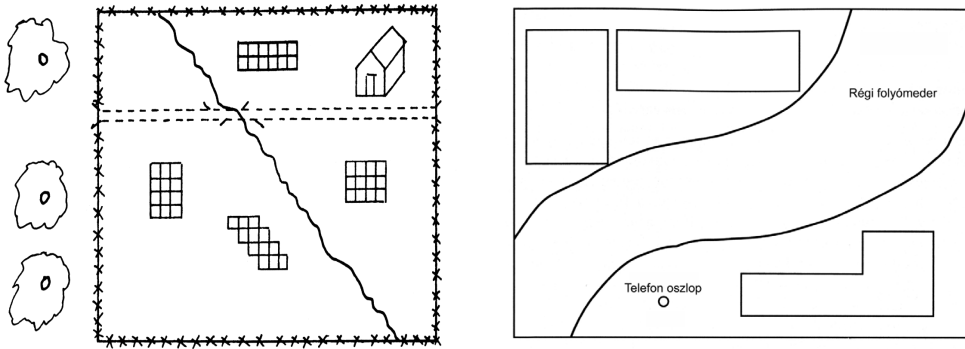
A szignifikancia szintek és a teszt erejének megválasztása kapcsolatban van a kockázattal, visszautasítjuk a nullhipotézist, amikor az a valóságban korrekt és a tesztnek azzal a képességével, hogy korrekten felismerjük, amikor a nullhipotézis hibás. A szignifikancia szintet gyakran 5%-nak választjuk, mivel ez az a kockázat, amelyre az emberek felkészülnek a legtöbb helyzetben, azonban nincs statisztikai oka annak, amiért más értéket nem tudnánk választani. A statisztikai erőt rendszerint 80%-nak vesszük, jóllehet nincs statisztikai oka, amiért ezt az értéket választjuk.

Számos módszer van, ami felhasználható a kísérletben szükséges ismétlések számának kiszámítására. Mindegyik módszer ugyanazon az alapelven alapul és hasonló eredményeket ad. Mindegyik különböző formulákat ad, példáulkal illusztrálva a szükséges ismétlések számának kalkulációját. A részleteket *Stael és Torrie (1980)* és *Cochran és Cox (1957)* ismerteti.

4.10. Blokk-képzés

A kísérleti egységeket, amelyek hasonló reakciókat adnak egy kezelésre, homogénnek írjuk le. Ha az egységek homogének, akkor a variáció az egységek között kicsi lesz. A blokk a homogén kísérleti egységeknek egy csoportja. A blokk-képzés mögötti gondolat az, hogy képezzük a kísérleti egységeknek egy csoportját, amelyek hasonlóak, amennyire lehetséges, azonban maguk a blokkok lehetnek heterogének. Számos kísérletben a blokk minden kezelés egyetlen ismétlését tartalmazza, ez a legegyszerűbb eset. Az is lehetséges, hogy lefuttatunk egy kísérletet több mint egy teljes ismétléssel egy blokkban vagy csak részleges ismétléssel egy blokkban. A parcellák egy blokkon belül homogének, azonban maguk a blokkok heterogének. A kezeléseket random (véletlenszerűen) kell elhelyezni a kísérleti egységeken, minden blokknál külön-külön.

Mielőtt döntünk a kísérleti egységek (parcellák) blokk-képzéséről, a variáció minden ismert forrását, amely hatással lehet a mért reakcióra, identifikálni szükséges, pl. születési tömeg, kondíció pontszáma, elővetemény, talajtípus. Minden blokkon belül a variáció ismert forrásait amennyire lehet, konstansan kell tartani. Például, ha a víznek a lefolyása a táblán változik a tábla lejtése miatt, akkor a kísérleti egységeket úgy kellene csoportosítani, hogy a blokkon belül minden parcella hasonló víz lefolyási sajátosságokkal rendelkezzen. Ezt rendszerint úgy érjük el, hogy a blokkokat a lejtésre merőlegesen és nem a lejtés irányába helyezük el (4.3. ábra).



4.3. ábra. A blokkok elhelyezése a kísérleti terület fizikai sajátosságainak figyelembevételével

Ha a kísérleti egységek szántóföldi parcellák, akkor a blokkoknak nem szükséges, hogy szabályos alakjuk legyen, feltéve, hogy homogének. Ha egy állategyed a kísérleti egység, akkor a blokknak tartalmaznia kell hasonló tömegű állatokat vagy a blokk-képzés elvégezhető úgy, hogy egyetlen blokk-faktort használunk vagy két vagy több kombinált faktort.

A blokk-képzés elődleges célja a *kísérleti hiba csökkentése* azáltal, hogy kiküszöböljük a variáció ismert forrásának hozzájárulását a kísérleti egységek közötti variációból.. Ezt azáltal érjük el, hogy a kísérleti egységeket (parcellákat) blokkokba csoportosítjuk úgy, hogy a *variabilitást* minden *blokkon belül minimalizáljuk* és a *blokkok között* a variabilitást *maximalizáljuk*. Mivel csak a blokkon belüli variáció válik a kísérleti hiba részévé, a blokk-képzés akkor a leghatékonyabb, ha a kísérleti terület rendelkezik a variabilitás előre jelezhető mintázatával. Előre jelezhető mintázat esetén, a parcella alakja és a blokk iránya úgy választható, hogy a variáció többségét megmagyarázza a blokkok közötti különbség és a kísérleti parcellákat ugyanazon a blokkon belül a lehető leg-egységesebben tartjuk.

Két fontos döntés van, amelyet meg kell tennünk, hogy eljussunk egy megfelelő és hatékony blokk-képzéshez. Ezek:

- A variabilitás forrásának feltárása, hogy felhasználhassuk a blokk-képzés alapjául.
- A blokk alakjának és irányának kiválasztása.

A variáció ideális forrása, amelyet felhasználhatunk a blokk-képzésre, nagy hatású és jól előre jelezhető. Például:

- Talaj heterogenitás egy műtrágyázási vagy fajtakísérletben, ahol a termésadatok azok, amelyek elsődlegesen érdekelnek bennünket.
- Rovar migráció iránya, egy inszekticid kísérletben, ahol a rovarok betelepülése az, amely elsődlegesen érdekel bennünket.
- Egy tábla lejtése, a növény vízstressz-reakciójának tanulmányozásakor.

A blokk-képzéshez a variabilitás specifikus forrásainak identifikálása után, a blokkoknak a méretét és alakját kell kiválasztani, hogy maximáljuk a variabilitást a blokkok között. Irányelvek ebben a vonatkozásban:

1. Ha a gradiens egyirányú (azaz csak egy gradiens van), használjunk hosszú és keskeny blokkokat. Továbbá, úgy orientáljuk ezeket a blokkokat, hogy a hosszabbik oldaluk merőleges legyen a gradiens irányára.

2. Ha talajtermékenységi gradiens fordul elő két irányban és az egyik gradiens sokkal erősebb, mint a másik, hagyjuk figyelmen kívül a gyengébb gradienst és az egyirányú gradiensre vonatkozó irányelveket kövessük.
3. Ha két talajtermékenységi gradiens fordul elő és mindkét gradiens egyformán erős és egymásra merőleges, válasszuk ki az egyik alábbi alternatívát:
 - Használjunk olyan blokkokat, amelyek lehetőség szerint négyzet alakúak.
 - Használjunk hosszú és keskeny blokkokat, melyek hosszabbik oldala merőleges az egyik gradiens irányára és használjuk a kovariancia eljárást, hogy vigyázzunk a másik gradiensre.
 - Használjunk latin négyzet elrendezést kétirányú blokk képzéssel (egyirányú blokk képzés mindegyik gradiensre).
4. Ha a variabilitás mintázata nem előre jelezhető, a blokkok lehetőség szerint négyzet alakúak legyenek.

Bármelyik blokk kialakítást használjuk, a blokkok azonossága és használatuk célja következetes legyen a kísérletezés teljes időszakában. Vagyis, ha bármikor olyan variáció forrás létezik, amelyet a kutató nem képes szabályozni, biztosítanunk kell azt, hogy az ilyen variáció a blokkok között forduljon elő és ne a blokkokon belül. Például, ha bizonyos műveletek, mint az inszekticidek kijuttatása vagy az adatgyűjtés nem fejezhető be az egész kísérletben egyetlen nap alatt, a feladat az, hogy fejezzük be ugyanazon blokk minden parcelláján a műveletet ugyanazon a napon. Ily módon a variáció a napok között (amelyet fokozhat az időjárási tényező) a blokk-variáció részévé válik és ily módon kizárjuk a kísérleti hibából. Ha egynél több személy végez méréseket a kísérletben, ugyanannak a felvételezőnek kellene végezni a méréseket egy blokk minden parcellájában. Ily módon a variáció a felvételezők között, ha van, a blokk variáció része és nem a kísérleti hibáé.

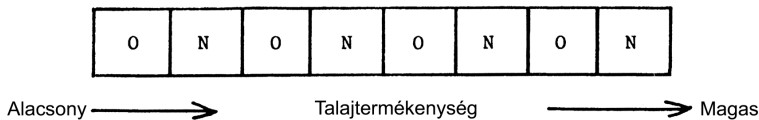
4.11. Randomizáció

A kezeléseket el kell osztani a kísérleti egységekre torzítatlan módon, hogy biztosítsuk azt, hogy egy sajátos kezelés nincs folytonosan hátrányban vagy előnyben az egymást követő ismétlésekben. A randomizáció a kezelések elosztása a kísérleti egységekre torzítatlan módon.

A randomizáció védelmet ajánl a véletlenszerű torzításokkal szemben. Ha egy kísérletet korekten randomizálunk, akkor a kísérlet nem torzíthatja az eredményeket semmilyen módon. A randomizációt, talán legfontosabb okként azért használjuk, hogy biztosítsuk az adatok statisztikai analízisének érvényességét. Legtöbb statisztikai analízis különböző feltételeket tesz az adatokra. Az egyik ilyen feltétel csaknem minden statisztikai analízisnél az, hogy a kísérleti hibák függetlenek. Az egyetlen mód, hogy ennek a feltételnek megfeleljünk, a kezelések random kijelölése a kísérleti egységekre. Ha a korrekt randomizációs eljárást nem használjuk, akkor a random hiba feltételt nem tudjuk teljesíteni és a statisztikai analízis érvénytelen lesz.

Tételezzük fel, hogy egy búzanemesítő elhatározta, hogy öt parcellán állítja be az új fajtát és ugyancsak öt parcellán a régi fajtát, hogy megvizsgálja a különbséget a két fajta termése között. Kijelölt egy sávot a 10 parcellának és a fajtákat a 4.4. ábrának megfelelően helyezte el a parcellákon. Tételezzük fel továbbá, hogy a nemesítő nem tudta, hogy egy talajtermékenységi gradiens alacsony értékű a sáv egyik végén

és fokozatosan növekedve, magasa értékű a sáv másik végén. Mindegyik parcella párban a régi (O) fajta után következett az új (N) fajta, ezáltal az új fajta mindig magasabb talajtermékenységu parcellába került. Egy megfigyelt különbség a fajták között ezáltal visszatükrözi a talajtermékenység különbséget is az új fajta előnyeként, valamint a fajták közötti különbséget, ha egyáltalán fennáll. E probléma elkerüléséhez a kezeléseket véletlenszerűen kell kijelölni a parcellákra, valamelyik randomizálási eljárással.



4.4. ábra. A régi (O) és új (N) fajták szisztematikus kijelölése a parcellák sávjában, talajtermékenységi gradienssel

Különböző pseudo-random eszközök, mint az érme feldobás, kockadobás, telefonkönyv használata, és így tovább, néha javasoltak, ezek azonban hajlamosak bevezetni nem-random elemeket a szelekciós eljárásba. Bármely megfelelő módszer felhasználja a véletlen számok statisztikai táblázatát vagy egy számológépet, és minden előállított számot fel kell használni. Annyi random számot kell generálni, amennyi kísérleti egység (parcella) van. Egy véletlen számok táblázata olyan táblázat, amelyben bármely lehetséges belépő egyenlő valószínűséggel fordul elő bármely adott pozícióban. Megtalálhatók minden statisztikai táblázatokat tartalmazó könyvben. Ahhoz, hogy generáljunk a véletlen számoknak egy sorozatát, a kezdő pontot a táblázatban véletlenszerűen választjuk ki és a négy lehetséges irány (balra, jobbra, fel vagy le) egyikét ugyancsak random választjuk ki. A számok sorozatát úgy állítjuk elő, hogy vesszük az egymás után következő számokat a táblázatból a kiválasztott irányban. Ha a táblázat szélét elérjük, egy új irányt választunk ki és az egymás utáni számokat használjuk egészen addig, amíg nincs szükség több random számra.

Az előállított random számok sorrendjét használjuk, hogy elosszuk a kezeléseket a kísérleti egységeken. Egy random számot osztunk ki minden kísérleti egységre. Az alapelv az, hogy ezeket a random számokat rangsoroljuk és a kezelés kódokat kijelöljük a sorrendnek. Ez a módszer garantálja, hogy a kezelések teljesen véletlenül kerüljenek a kísérleti egységekre és ez megengedi, hogy a legtöbb statisztikai analízis egyik feltételének megfeleljünk.

4.12. Kovariánsok

Az alap gondolat a kovariáns mögött az, hogy bizonyos mértékű variáció a reakció változóban nem a kezeléseknak vagy a háttér hiba variabilitásnak tulajdonítható, hanem a variációnak a kovariánsban. Ha a kovariánsnak tulajdonítható variációt eltávolítjuk a reakció változó kezelés hatásának becslését megelőzően, akkor a becslés hiba variancia, amellyel a kezelés-hatást teszteljük, csökkenni fog és így a kísérlet pontosabb lesz.

Előfordul, hogy regresszió-analízissel nem tudjuk elég pontosan meghatározni a függő és független változók közötti kapcsolatot, mivel a kvantitatív változókon kívül kvalitatív jellegű változók is jelen vannak. Így először ezek hatását kell kiküsz-

öböltni, majd ezután határozható meg az X és Y változók közötti regressziós kapcsolat. Erre szolgál a kovariancia-analízis (ANCOVA), amely a varianciaanalízis és a regresszióanalízis kombinációja (Sváb, 1981) és egy standard módszer a kovariánsok számbavételére az analízisben. A kovariancia analízisben a kovariáns hatását a reakció változóra regresszió analízissel távolítjuk el a reziduális (maradék) variabilitásból.

A kovariánsok használhatók bármelyik kísérleti elrendezésnél, azonban az analízis eredményének értelmezése bonyolult lehet a többfaktoros kísérletekben. Az is feltételezhető, hogy az összefüggés a kovariáns és a reakció változó között görbe vonalú, pl. kvadratikus. Az analízist ugyanolyan módon végezzük, mintha az egyenes vonalú összefüggést használnánk, azonban eltérő regressziós modellel. Több kovariáns is használható. Fontos, hogy a kovariánsra ne legyen hatása a kezeléseknek. Következésképpen, a kovariánsokat mérhetjük a kezelések beállítása előtt. Például minden parcella előző termése felhasználható, hogy visszatükörözze a parcellánkénti különbségeket a parcella termékenységben, ha ugyanazokat a parcellákat használtuk újból a búzatermés kísérlethez.

Példa. Egy búzatablában május, június és július hónapokban havonta 8 db 400 cm²-es parcellákban levágták a búzát és feljegyezték a tőszámát. A parcellánkénti tömeget szárazanyag g-ban adták meg. A kutatók vizsgálták, vajon különbözik-e egységnyi területen a búza tömege a különböző hónapokban. A tömeg varianciaanalízise előtt megvizsgálták, van-e összefüggés a középértékek és a hozzájuk tartozó szórás között. Mivel találtak közöttük összefüggést, a szórás stabilizálására a tömeg adatok log transzformáltját alkalmazták a további analízisben.

A tömegadatokra elvégzett ANOVA nem adott szignifikáns eredményt, vagyis egyszerű varianciaanalízissel nem tudtak a hónapok között különbséget kimutatni. A tőszámnak azonban szignifikáns hatása volt a tömegre, így kovariancia analízissel folytatták. A tőszám hatásának eltávolítása után a hónapok között szignifikáns különbséget mutattak ki. A tőszámra korrigált átlagok esetében egy növekvő trendet találtak, amelyet a tőszám zavaró hatása elfedett az eredeti adatok esetében. A példa részletes analízisét *Précsényi et al.* (1995) közli.

Példa. Egy biológiai kísérletben vizsgálták a nagyobb levélkárosodás (két szint: természetes növényevő kártétele, mesterségesen megnövelt kártétel a növényevő fokozott kártételének utánzására) hatását a ligetszépe (*Oenothera macrocarpa*) évelő növény virágzásának virág tulajdonságaira Missouri államban (*Quinn és Keough*, 2002). A reakció változó a párta átmérő és a virágtömlő hossza volt, melyeknek változásai különbségeket idézhetnek elő a beporzási preferenciában és hatásfokban. A virágzási sorrendet (szukcesszív szezonális virágzás) használták kovariánsként, hogy elősegítsék értelmezni a virágzási tulajdonságok variációját és a kártétel hatások erőteljesebb tesztjét kapják.

A kovariancia analízissel részletesen foglalkozik pl. *Sváb* (1981) és *Snedecor és Cochran* (1989). A cukorrépa hozamkísérleteket kovariancia-analízissel értékelte *Soós és E. Majzik* (1980).

5. Szántóföldi kísérleti elrendezések és a kísérletek szerkezete

A kísérleti elrendezés jelenti a kísérletekben a parcellák és a parcellákon a vizsgált kezelések tervszerű elhelyezését, illetve csoportosítását valamely előírt rendszer szerint. A kísérleti elrendezés célja, hogy az összehasonlított kezeléseknél lehetőleg azonos körülményeket biztosítsunk, illetve az ellenőrizetlen, tudatosan nem irányított hatások egyenlően érvényesüljenek az összes kezelésen. A megfelelő kísérleti elrendezést a kísérleti kérdéstől, a vizsgált tényezőktől, a kezelések és ismétlések számától függően választjuk meg. A kísérleti elrendezés és a kísérlet biometriai értékelése egymással szorosan összefügg; az elrendezés módja egyértelműen meghatározza az alkalmas értékelési eljárást. A kísérletek elrendezésének számos oka van: (i) A kísérleti hiba becslését adja. (ii) Növeli a precizitást a parcellák tervezett csoportosításával és a csoport különbségek eltávolításával a kísérleti hibából. (iii) Információt ad, amely a szignifikancia vizsgálatokhoz szükséges. (iv) Elősegíti a kezelések alkalmazását, az agrotechnikai eljárásokat és a kísérleti parcellák betakarítását. A kísérleti elrendezést befolyásoló tényezők: 1. A kísérleti hely fizikai és topográfiai sajátosságai. 2. A talaj variabilitás nagysága és mintázata. 3. A kezelések száma és természete. 4. Vizsgálatba bevont kultúrnövény. 5. A kísérlet időtartama. 6. A használni kívánt gépek sajátosságai. 7. A kezelések között kimutatni kívánt különbség nagysága. 8. A tervezett szignifikancia szint. 9. A kísérleti források (vetőmag, műtrágya, növényvédőszer) mennyisége, amelyet a vizsgálatra szánunk. 10. A kísérlet végrehajtásának költsége (pénz, idő, személyzet). A legfőbb szabály: válasszuk ki a legegyszerűbb kísérleti elrendezést, amely megadja a megkívánt precizitást, a rendelkezésre álló források korlátainak figyelembe vételével. A kísérleti elrendezések típusait a következő sorrendben tárgyaljuk: egytényezős, kéttényezős, három- és többtényezős kísérletek, kísérletsorozatok, tartamkísérletek és üzemi kísérletek (Berzsenyi, 2010).

5.1. Egytényezős kísérletek

A kísérleti tervezésnek két csoportja alkalmazható az egytényezős kísérletekre. Az egyik csoport a teljes (komplett) blokk elrendezés, amely kevés kezelést tartalmazó kísérletekre alkalmas és olyan blokkokkal jellemezhető, amelyek mindegyike a kezelések legalább egy teljes sorozatát tartalmazza. A másik a nem-teljes (inkomplett) blokk elrendezés, amely a nagyszámú kezelést tartalmazó kísérletekre alkalmas és olyan blokkokkal jellemezhető, amelyek a vizsgálni kívánt kezeléseknél csak egy részét (frakcióját) tartalmazzák. A komplett blokk-elrendezéshez tartozik a teljesen randomizált, randomizált komplett blokk, latin négyzet, a csoportosított blokk-elrendezés, míg a nem-komplett blokk-elrendezéshez a latin téglalap elrendezés.

5.1.1. Teljesen véletlen elrendezés (CRD)

A kezeléseket teljesen véletlenül jelöljük ki, úgy hogy minden kísérleti egységnek (parcella) azonos az esélye, hogy bármely kezelést kapja. A variációnak két forrása van: az egyik a kezelés variáció, a másik a kísérleti hiba. A kettő relatív nagyságát használjuk arra, hogy jelezzük, hogy a megfigyelt különbség a kezelések között valós vagy a véletlennek tulajdonítható. A kezelések közötti különbségről akkor mondjuk, hogy valós, ha a kezelés variáció megfelelően nagyobb, mint a kísérleti hiba. A CRD-nél, bármely különbséget az azonos kezelést kapott kísérleti egységek között kísérleti hibának tekintünk. Ezáltal a CRD csak homogén kísérleti egységeket tartalmazó kísérletekben megfelelő, mint amilyenek az üvegházi, fitotron és labor kísérletek, ahol környezeti hatásokat viszonylag könnyű szabályozni. Szántóföldi kísérleteknél, ahol általában nagy a variáció a kísérleti parcellák között, olyan környezeti faktorokban, mint a talaj, a CRD-t ritkábban használjuk. Egy teljesen véletlen elrendezésű kísérletet (hat kezelés és négy ismétlés) az 5.1. ábra szemléltet.

1 B	4 F	7 C	10 C	13 E	16 E	19 A	22 F
2 E	5 A	8 A	11 A	14 B	17 D	20 F	23 B
3 C	6 B	9 D	12 C	15 F	18 E	21 D	24 D

5.1. ábra. Teljesen véletlen elrendezés (CRD) hat kezeléssel (A,B,C,D,E és F) és négy ismétléssel (n = 24)

Előnyök, hátrányok és felhasználás

A CRD-nek számos előnye van: 1. *Flexibilitás*. Bármilyen számú kezelés és ismétlés használható. Továbbá, a kezelésenkénti ismétlések számának nem szükséges azonosnak lenni minden kezelésnél, jóllehet az elrendezés akkor a leghatékonyabb, ha a kezelések ismétlése egyenlő. 2. *Egyszerű statisztikai analízis*. Az analízist nem komplikálja a kezelések nem egyenlő ismétlése. Ez nem áll fenn a bonyolultabb kísérleti elrendezéseknél. 3. *Hiányzó parcellák*. Az analízist nem komplikálják hiányzó adatok. 4. *Maximális hiba szabadságfok (df vagy FG)*. Azonos összes parcellaszámnál és kezelésnél egyetlen más kísérleti elrendezés sem ad több hiba szabadságfokot, mint a CRD. A CRD fő hátránya az alacsony precizitás, ha a parcellák nem homogének. Jóllehet, más kísérleti elrendezések precízebbek lehetnek, a CRD-nek számos felhasználása lehet: 1. Ha a kísérleti terület viszonylag egységes vagy a csoportosításnak nincs elfogadható alapja. 2. Ahol a parcellák (kezelések) nagy hányada várhatóan nem ad reakciót vagy tönkremehet a kísérlet során. 3. Ahol a parcellák száma korlátozott.

A teljesen véletlen elrendezés statisztikai modellje:

A modellt a következő formában írhatjuk fel:

$$(\text{termés}) = (\text{főátlag}) + (\text{kezelés-hatás}) + (\text{maradék})$$

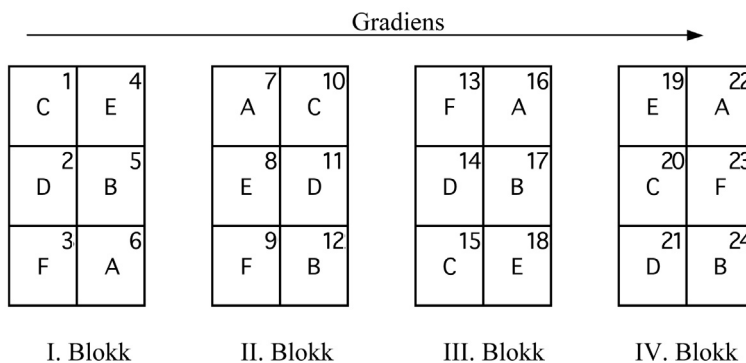
$$y_{ij} = \mu + T_i + e_{ij} \quad i = 1 \dots t; \quad j = 1 \dots r_i$$

ahol: y_{ij} : az i -edik kezelésben részesült j -edik parcella adata (pl. termése), μ : a populáció átlagos reakciója, T_i : az i -edik kezelés átlagos hatása, r_i : az i -edik kezelés ismétléseinek száma, e_{ij} : random kísérleti hiba. A teljesen véletlen elrendezés (CRD) varianciaanalízis táblázata:

Variáció forrása	df	SS	MS	F-érték
Kezelés (T_i)	df_T	SS_T	$MS_T = SS_T/df_T$	$F_T = MS_T/MS_E$
Maradék (e_{ij})	df_E	$SS_E = SS_{TOT} - SS_T$	$MS_E = SS_E/df_E$	
Összes	df_{TOT}	SS_{TOT}		

5.1.2. Véletlen blokkelrendezés (RCB, RBD)

A véletlen blokkelrendezés az egyik leggyakrabban használt kísérleti elrendezés a mezőgazdasági kutatásban. Különösen alkalmas szántóföldi kísérletekre, ahol a kezelések száma nem nagy és a kísérleti terület előre jelezhető talajtermékenységi vagy más gradienssel rendelkezik. Az RCB fő megkülönböztető tulajdonsága az egyenlő méretű blokkok jelenléte, amelyeknek mindegyike tartalmazza az összes kezelést. A blokk-képzés elsődleges célja a kísérleti hiba csökkentése azáltal, hogy kiküszöböljük a variáció ismert forrásának hozzájárulását a kísérleti egységek között. Ezt azáltal érjük el, hogy a kísérleti egységeket (parcellákat) blokkokba csoportosítjuk úgy, hogy a variabilitást minden blokkon belül minimalizáljuk és a blokkok között a variabilitást maximalizáljuk. A variabilitás három forrása van az RCB elrendezésben: kezelés, ismétlés (vagy blokk), és kísérleti hiba. Figyeljük meg, hogy ez eggyel több, mint a CRD-nél. Egy véletlen blokkelrendezésű kísérletet hat kezeléssel és négy ismétléssel a 5.2. ábra szemléltet.



5.2. ábra. Véletlen blokkelrendezésű kísérlet (RCB) hat kezeléssel (A,B,C,D,E és F) és négy ismétléssel (I-IV) ($n = 24$)

Előnyök, hátrányok és felhasználás

A véletlen blokkelrendezésnek számos előnye van: 1. A kísérleti hibából eltávolíthatja a variáció egyik forrását és így növeli a precizitást. 2. Blokkok képzése különböző környezeti körülmények között bővíti a kísérleti eredmények érvényességét. 3. Akárhány kezelést és bármennyi blokkot képezhetünk. Az egyetlen korlátozás, hogy minden kezelést azonos számban kell megismételni minden blokkban. 4. Az eredmények biometriai analízise viszonylag egyszerű. *Van néhány hátránya is ennek a kísérlet elrendezésnek:* 1. A hiányzó adatok nehézséget okoznak az analízisben. 2. A kezelés jelölése hibásan a parcellára a nem megfelelő blokkba, problémákhoz vezet az analízisben. 3. Ez az elrendezés kevésbé hatékony, mint mások, ha egynél több nem kívánatos variáció forrás (gradiens) van. *A véletlen blokkelrendezésnek több felhasználási területe van:* 1. Felhasználható, hogy kiszűrjük a nem kívánatos variáció forrást. Gyakran megfelelő precizitást ad, és így nincs szükség bonyolultabb elrendezésre. 2. A blokk-képző faktor átlagainak korrekt becslését adja. Ezáltal ezek az átlagok a véletlen blokkelrendezéssel becsülhetők.

A véletlen blokkelrendezés statisztikai modellje:

(termés) = (főátlag) + (ismétlés vagy blokk-hatás) + (kezelés-hatás) + (maradék)

$$y_{hi} = \mu + T_i + Bl_h + e_{hi} \quad i = 1 \dots t; \quad h = 1 \dots b$$

ahol: y_{hi} : a h-adik blokkban i-edik kezelésben részesült parcella adata (pl. termése), μ : a populáció átlagos reakciója, T_i : az i kezelés átlagos hatása, Bl_h : a h blokk hatása, e_{hi} : random kísérleti hiba.

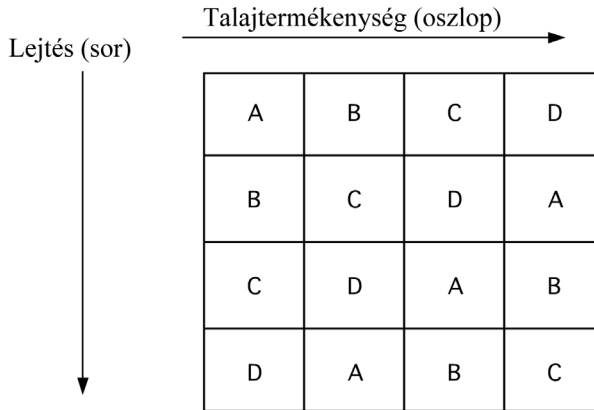
Véletlen blokk elrendezés (RCB) varianciaanalízis táblázata:

Variáció forrása	df	SS	MS	F-érték
Blokk (Bl_h)	df_{Bl}	SS_{Bl}	$MS_{Bl} = SS_{Bl} / df_{Bl}$	
Kezelés (T_i)	df_T	SS_T	$MS_T = SS_T / df_T$	$F_T = MS_T / MS_E$
Maradék (e_{hi})	df_E	$SS_E = SS_{TOT} - SS_T - SS_{Bl}$	$s^2 = SS_E / df_E$	
Összes	df_{TOT}	SS_{TOT}		

5.1.3. Latin négyzet elrendezés (LS)

A latin négyzet (LS) elrendezés legfontosabb sajátossága, hogy egyidejűleg képes a kísérleti egységek közötti variáció két ismert forrását figyelembe venni. Ezeket a variáció forrásokat független blokk-képzési kritériumként kezeli, szemben az RCB-vel, ahol egy variáció forrás volt. A kétirányú blokk-képzés (sorirányú és oszlopírányú blokkok) azt jelenti, hogy minden kezelés csak egyszer fordul elő minden sor-blokkban és egyszer minden oszlop-blokkban. Ez az eljárás teszi lehetővé, hogy becsüljük a variációt a sor-blokkokban és az oszlop-blokkokban és elkülönítsük ezeket a kísérleti hibából. Az LS elrendezés néhány felhasználási területe: (i) szabadföldi kísérletekben,

ahol két gradiens (pl. talajtermékenységi és a terület lejtése) egymásra merőleges, (ii) inszekticid szabadföldi kísérletekben, ahol a rovarok migrációjának van egy előrejelezhető iránya, amely merőleges a kísérleti terület fő talajtermékenységi gradiensére. A latin négyzet elrendezésben a kezelések száma azonos az ismétlések számával. Az 5.3. ábra egy latin négyzet kísérleti elrendezést mutat be.



5.3. ábra. Latin négyzet kísérlet elrendezési terve két egymásra merőleges gradiens hatásának figyelembevételével

Előnyök, hátrányok és felhasználás

A latin négyzet elrendezésnek egyetlen fő előnye van: 1. Lehetővé teszi a kísérletezőnek, hogy a variáció két forrását szabályozza. *Számos hátránya van:* 1. A hiba FG kicsi, ha csak néhány kezelést tartalmaz. 2. A kísérlet nagyon nagy lesz, ha a kezelések száma nagy, mivel a parcellák száma a kezelések számának négyzete. 3. A statisztikai analízist bonyolítják a hiányzó parcellák és a kezelések téves beállítása. *A latin négyzet elrendezés korlátozott számú alkalmazása:* 1. Hasznos, ha a variáció két forrását kell szabályozni. 2. Olyan kísérletekben alkalmazzuk, ahol a kezelések száma nem kisebb négy-nél és nem nagyobb nyolcnál. A variációnak itt négy forrása van: sor, oszlop, kezelés és kísérleti hiba.

A latin négyzet elrendezés statisztikai modellje:

(termés) = (főátlag) + (sor-hatás) + (oszlop-hatás) + (kezelés-hatás) + (maradék)

$$y_{ijk} = \mu + R_i + C_j + T_k + e_{ijk} \quad i = 1 \dots t; j = 1 \dots t; k = 1 \dots t$$

ahol: y_{ijk} : az i -edik sorban és a j -edik oszlopban k -adik kezelésben részesült parcella adata (pl. termése), μ : a populáció átlagos reakciója, R_i : az i -edik sor átlagos hatása, C_j : a j -edik oszlop átlagos hatása, T_k : a k -adik kezelés átlagos hatása, e_{ijk} : random kísérleti hiba. Latinnégyzet elrendezés (LS) varianciaanalízis táblázata:

Variáció forrása	df	SS	MS	F-érték
Sor (R_i)	df_R	SS_R	$MS_R=SS_R/df_R$	
Oszlop (C_j)	df_C	SS_C	$MS_C=SS_C/df_T$	
Kezelés (T_k)	df_T	SS_T	$MS_T=SS_T/df_T$	$F_T=MS_T/MS_E$
Maradék (e_{ijk})	df_E	$SS_E=SS_{TOT}-SS_R-SS_C-SS_T$	$MS_E=SS_E/df_E$	
Összes	df_{TOT}	SS_{TOT}		

Megemlítjük még a *Graeco-latin négyzet elrendezést*. Ebben kezelések az ismétlésekben három különböző módon vannak csoportosítva, hogy a variáció három különböző forrásának hatását kiegyenlítsék minden kezelésre. Nagyon ritkán használják, mivel a kísérleti egységek (parcellák) ritkán kiegyenlítettek mindhárom csoportosításnak megfelelően.

Többszörös latin négyzet elrendezés

A latin négyzet elrendezés egyik hátránya, hogy kiiktatva a variáció két forrását az eredeti egységekből (parcellákból), elkerülhetetlenül csökkentjük a rendelkezésre álló szabadságfokok számát a fennmaradó random variáció becslésére és ez a kezelek összehasonlítását kevésbé pontosá teszi. Például, három kezeléssel csak 2 hiba d.f. van, összehasonlítva a véletlen blokkelrendezés 4 d.f.-ével, azonos számú kísérleti egységnél. Négy kezelésnél 6 hiba d.f. van, összehasonlítva a véletlen blokkelrendezésnél 9 d.f. van. A t-eloszlás értékei 5%-os szignifikancia szinten 4.30 a 2 d.f.-nél és 2.78 a 4 d.f.-nél. Ahhoz, hogy egy adott nagyságú különbséget szignifikánsnak észleljünk ugyanannál a szintnél egy 3×3 -as latin négyzetnél, mint egy azonos méretű randomizált blokk elrendezésnél, a két blokk-képzési kritérium használatának a latin négyzetben a hiba MS-t a randomizált blokk hiba MS értékének 65%-ára (azaz kb. 40%-kal) kell csökkenteni (Mead *et al.*, 1993).

A	B	C
C	B	A
A	C	B

(a)

C	B	A
B	A	C
A	C	B

(b)

B	A	A	C	B	C
C	B	C	B	A	A
A	C	B	A	C	B

5.4. ábra. A többszörös latinnégyzet elrendezés két formája: (a) két független latinnégyzet és (b) két latinnégyzet egyesítve, hogy 3×6 -os téglalapot képezzen

E nehézség leküzdésének egyik útja, miközben megtartjuk a két blokk-képzési faktor előnyét, hogy egynél több latin négyzetet használunk, amely több mint megduplázza a hibára rendelkezésre álló szabadságfokot. Két úton tehetjük ezt meg, mint az 5.4. ábra mutatja.

Az 5.4. (a) ábrán a két (vagy több) négyzet teljesen függetlenként van kezelve, a randomizáció egymástól független, mint ahogy a kezelések randomizációja történik a véletlen blokk elrendezés különböző blokkjaiban.

Az 5.4. (b) ábrán a két (vagy több) négyzet egyesítve van és egy téglalapot formál, amelyben minden kezelés egyszer jelenik meg minden oszlopban és kétszer (vagy többször) mindegyik sorban. Ehhez az elrendezéshez, két különböző négyzet van feltüntetve kezdetben, és a sorok, oszlopok randomizációja és a kezelések elosztása úgy történik, mint korábban, kivéve, hogy most van hat oszlopunk, nem két elkülönített sorozata három-három oszlopnak.

A lényeges különbség a két elrendezés között az, hogy az (a) esetben, ahol két független négyzetünk van, a sor hatásról az 1. négyzetben nem feltételezzük, hogy bármilyen hasonlósága van a sor hatásokhoz a 2. négyzetben. Ugyanez igaz az oszlop hatásokra. A „téglalapban” a sor hatásokról feltételezzük, hogy ugyanaz mindkét eredeti négyzögre és ezáltal mindenoszlopra. Az oszlop hatásokról a két eredeti négyzetre feltételezzük, hogy nincsenek kapcsolatban. Ez a különbség a feltételezésben elkerülhetetlenül az analízis különböző formáihoz vezet. Az (a) esetben, ahol két független négyzetünk van, az eltérésnégyzet összegeket a sorokra és oszlopokra mindegyik négyzetre külön-külön számítjuk ki. Kiszámítunk egy eltérésnégyzet összeget a négyzetek közötti általános különbségre és kiszámítjuk a kezelés eltérésnégyzet összeget, felhasználva a kezelés összegeket mindkét négyzet számbavételével. A variancia táblázat szerkezete a két független latinnégyzet elrendezésnél:

Variáció forrása	SS	d.f.
Négyzetek között	$\frac{G_1^2 + G_2^2}{9} - \frac{[\sum(x)]^2}{18}$	1
Sorok	(sor SS) ₁ + (sor SS) ₂	4
Oszlopok	(oszlop SS) ₁ + (oszlop SS) ₂	4
Kezelések	$\frac{T_1^2 + T_2^2 + T_3^2}{6} - \frac{[\sum(x)]^2}{18}$	2
Hiba	kivonással	6
Összes	$\sum(x^2) - \frac{[\sum(x)]^2}{18}$	17

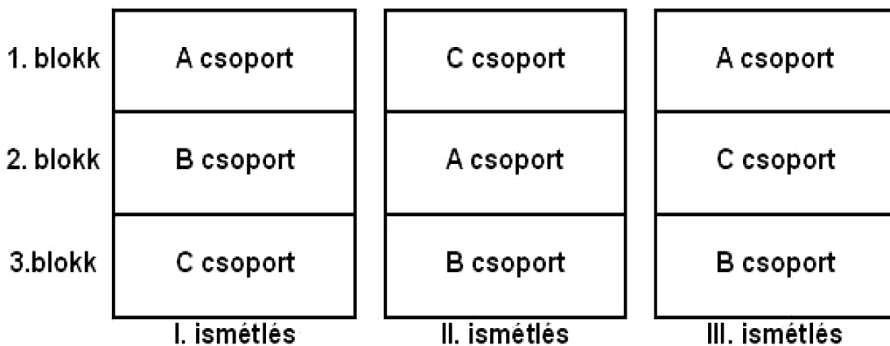
Variancia táblázat szerkezete a két latinnégyzet egyesítésekor:

Variáció forrása	SS	d.f.
Sorok	$\frac{R_1^2 + R_2^2 + R_3^2}{6} - \frac{[\sum(x)]^2}{18}$	2
Oszlopok	$\frac{C_1^2 + C_2^2 + C_3^2 + C_4^2 + C_5^2 + C_6^2}{3} - \frac{[\sum(x)]^2}{18}$	5
Kezelések	$\frac{T_1^2 + T_2^2 + T_3^2}{6} - \frac{[\sum(x)]^2}{18}$	2
Hiba	kivonással	8
Összes	$\sum(x)^2 - \frac{[\sum(x)]^2}{18}$	17

G_1 és G_2 főösszegek az első és második négyzetben; T_1 , T_2 és T_3 kezelésszegek a hat egységre, három-három mindegyik négyzetben.

A (b) esetre, a latin téglára, az eltérésnégyzet összegek ki vannak számítva a sor összegek, az oszlop összegek és a kezelés összegek közötti variációra, amely a következő varianciaanalízist eredményezi:

A legfontosabb különbség a két analízis között a több hiba d.f. a (b) esetben. A (b) esetnek ez a fölénye igaz a latin négyzetek minden méretére és akkor is, ha kettőnél több négyzetet használunk. Visszatükrözi a (b) esetben használt komolyabb feltételezéseket, hogy a különbségek a sorok között hasonlóak mind a hat oszlopra, míg az (a) esetben a sor hatásoknak megengedett, hogy különbözőek legyenek a két négyzetben.



5.5.a. ábra. Random elhelyezése a három fajtacsoportnak (A, B és C) három blokkba mindhárom ismétlésben, amely az első lépést jelenti a csoportosított blokk elrendezés randomizációs folyamatában

7	8	1	4	14	38	39	43	44	33	2	1	13	5	15	1. blokk
12	3	13	10	9	31	42	36	32	45	11	4	3	14	7	
6	11	2	5	15	40	34	35	41	37	9	12	10	6	8	
21	20	27	18	22	1	11	9	4	6	31	33	35	38	44	2. blokk
16	25	17	24	30	12	10	7	15	13	36	34	37	42	41	
26	19	28	29	23	3	2	8	5	14	39	43	32	45	40	
41	31	32	44	43	17	27	22	25	28	16	25	23	20	17	3. blokk
36	37	40	35	38	26	18	21	19	20	22	19	30	28	21	
33	39	42	45	34	16	23	24	30	29	18	24	26	29	27	
I. ismétlés					II. ismétlés					III. ismétlés					

5.5.b. ábra. A csoportosított blokk kísérlet elrendezése, amely magában foglal 45 fajtát, felosztva három csoportra, mindegyik csoport tartalmaz 15 fajtát. A kísérletet három ismétlésben állították be

5.1.4. Csoportosított blokk-elrendezés

A *csoportosított blokk-elrendezés* (group balanced block design) elsődleges sajátossága a kezelések csoportosítása homogén blokkokba, a kezdések kiválasztott tulajdonságai alapján. Míg a latintégla elrendezésnél a homogenitást blokkon belül a *kísérleti parcellák* olyan csoportosításával éri el, amely a kísérleti terület valamilyen ismert heterogenitási mintázatának ismeretén alapul, addig a csoportosított elrendezés ugyanezt a *kezelések* olyan csoportosításával éri el, amely a kezelések valamilyen ismert jellemzőjén alapul.

A csoportosított elrendezésben, ugyanahhoz a csoporthoz tartozó kezeléseket mindig ugyanabban a blokkban teszteljük, azonban azokat, amelyek eltérő csoportokhoz tartoznak, sohasem teszteljük ugyanabban a blokkban. Ezáltal, a precizitás, amellyel a különböző kezeléseket összehasonlítjuk, nem ugyanaz minden összehasonlításnál. Ugyanis ahhoz a csoporthoz tartozó kezeléseket egy magasabb fokú precizitással vannak összehasonlítva, mint azok, amelyek különböző csoportokhoz tartoznak. A csoportosított elrendezést általánosan használjuk a fajta kísérletekben, ahol a hasonló morfológiai tulajdonságokkal rendelkező fajtákat együtt helyezük el ugyanabban a csoportban.

Két leggyakrabban használt kritérium a fajták csoportosítására: (i) Növénymagasság, abból a célból, hogy elkerüljük a jelentősebb kompetíciós hatást, amikor jelentősen eltérő magasságú növények növekednek a szomszédos parcellákban. (ii) Növekedés időtartama, abból a célból, hogy minimalizáljuk a kompetíciós hatást és elősegítsük a betakarítási műveletet. A kísérletek egy másik típusa, ahol a csoportosított elrendezést használjuk az, amelyik vegyszeres inszekticid, vagy herbicid kezeléseket foglal magában és a kezeléseket fel lehet osztani hasonló permetezési eljárásokra, hogy elősegítsük a vegyszerek szabadföldi kijuttatását.

Az 5.5.a-5.5.b. ábra 45 rizsfajta csoportosított elrendezését mutatja be három ismétlésben. A fajtákat a növekedés időtartama alapján sorolták három csoportba: A csoportban a 105 napnál rövidebb növekedési időtartamú fajták, B csoportban 105 és 115 nap közötti növekedési időtartamú fajták, C csoportban 115 napnál hosszabb növekedési időtartamú fajták vannak. Mindegyik csoport 15 fajtát tartalmaz.

5.1.5. Nem-teljes blokk-elrendezés

Elméletileg a komplett blokkelrendezések, mint amilyen a véletlen blokk (RCB) és a latin négyzet (LS) elrendezés, bármilyen számú kezeléshez alkalmazhatók. A teljes (komplett) blokkelrendezések azonban kevésbé hatékonyá válnak, ahogy a kezelések száma nő, elsődlegesen azért, mert a blokk mérete arányosan nő a kezelések számával, és a kísérleti kezelések homogenitását nagy blokkon belül bonyolult fenntartani. Azaz, a komplett blokkelrendezés kísérleti hibája általában várhatóan nő a kezelések számával.

A nagy számú kezelést tartalmazó egytényezős kísérletek alternatív elrendezési módja az inkomplett (nem teljes) blokkelrendezés, amelynek egyike a rácselrendezés. Mint a neve mutatja, minden blokk az inkomplett blokk elrendezésben nem tartalmaz minden kezelést és meglehetősen kis blokkméret fenntartható még akkor is, ha a kezelések száma nagy. Kisebb blokkokkal a kísérleti parcellák homogenitása ugyanabban a blokkban könnyebben fenntartható és általában nagyobb pontosság várható. Megvan az ára a nagyobb pontosságnak az inkomplett blokkelrendezés használatakor. Ezek a következők: (i) A kezeléseknél vagy az ismétléseknél vagy mindkettőnél a száma nem rugalmas (nem választható meg önkényesen). (ii) Nem egyenlő mértékű a pontosság a kezeléscsoportok összehasonlításakor. (iii) Az adatelemzés bonyolultabb.

Jóllehet nincs külön szabály arra, hogy mekkora lehet a kezelések száma az inkomplett blokkelrendezés használatakor, az alábbi irányelvek segítségével szolgálhatnak: (i) Az inkomplett blokkelrendezés előnye a komplett blokkelrendezéssel szemben fokozódik a kísérleti anyagban levő variabilitás növekedésével. Általánosságban, bármikor, ha a blokkméret az RCB elrendezésnél túl nagy ahhoz, hogy fenntartsion egy megfelelő szintű homogenitást a kísérleti parcellák között ugyanazon a blokkon belül, az inkomplett blokkelrendezés használatát komolyan fontolóra kell venni. (ii) Az inkomplett blokkelrendezés adatelemzése komplexebb, mint a komplett blokkelrendezésé. Ezért olyan esetekben, amikor a számítógépes lehetőségek és szolgáltatások kevésbé hozzáférhetők, az inkomplett blokkelrendezést utolsó megoldásként vegyük számításba.

Általánosságban, az inkomplett blokkelrendezés a kisebb blokkmérettel várhatóan nagyobb pontosságot ad, mint a komplett blokkelrendezés (5.6. ábra). Ezáltal az inkomplett blokkelrendezést általában előnyben kell részesíteni, ha a használatához szükséges forrásoknak (pl. több ismétlés, nem flexibilis számú kezelések, és komplexebb adatelemzés) eleget tudunk tenni. A bevezetett extra blokk-struktúra a szabadföldi variáció szabályozására extra tagként jelenik meg a modellben, amely most a következő:

$$(\text{termés}) = (\text{főátlag}) + (\text{ismétlés-hatás}) + (\text{blokk-hatás}) + (\text{kezelés-hatás}) + (\text{maradék})$$

		Blokkok							
		1	2	3	4	5	6	7	
Alacsony termé- kenység		42	12	15	37	17	22	4	
		21	26	25	18	28	36	16	
		10	35	6	29	11	8	2	
		23	38	41	3	31	24	30	Magas
		39	1	27	40	14	34	7	termé-
		19	21	33	5	9	13	32	kenység
		→							
Trend iránya									

5.6. ábra. Egy 42 kezelést tartalmazó kísérlet I. ismétlésének felosztása blokkokra az inkomplett blokk-elrendezés illusztrálására, hogy a kísérleti területen a talajtermékenységi trendet figyelembe vegyünk

A *latintégla (lattice design)* a legáltalánosabban használt inkomplett blokk-elrendezés a mezőgazdasági kutatásban. Megfelelő flexibilitás van a tervezésében ahhoz, hogy alkalmazása egyszerűbb legyen, mint más inkomplett blokk-elrendezésé. Leggyakrabban a kiegyensúlyozott (balanced) és a részben kiegyensúlyozott (partially balanced) latintégla elrendezést használják. Mindkettőnél a kezelések számának tökéletes négyzetszámnak kell lenni. A *kiegyensúlyozott latintégla elrendezést a következő alaptulajdonságok jellemzik*: 1. a kezelések száma (t) négyzet szám (azaz $t = k^2$, úgymint 25, 36, 49, 64, 81, 100 stb.). Jóllehet ez a követelmény szigorúnak tűnhet először, rendszerint könnyű betartani a gyakorlatban. Ahogy a kezelések száma nagy lesz, néhányal több vagy néhányal kevesebb kezelést rendszerint könnyű teljesíteni. Például, ha egy növénynevelő 80 fajta teljesítményét kívánja vizsgálni egy kiegyensúlyozott rácselrendezésű kísérletben, mindössze eggyel több fajtát kell beállítania a kísérletbe. Vagy ha 82 vagy 83 fajtája van, könnyen kihagy egyet vagy kettőt. 2. A blokk mérete (k) egyenlő a kezelések számának négyzetgyökével (azaz, $k = t^{1/2}$). 3. Az ismétlések száma (R) eggyel több, mint a blokk nagysága, azaz $r = (k + 1)$. Azaz, a szükséges ismétlések száma 6, ha a kezelések száma 25, 7 ha a kezelések száma 36, 8 ha a kezelések száma 49, és így tovább. Egy latin téglakísérlet elrendezését a 5.7. ábra szemlélteti.

1. Blokk	T_9	T_3	T_6	T_5	T_4	T_6	T_7	T_5	T_3	T_1	T_9	T_5
2. Blokk	T_8	T_5	T_2	T_3	T_2	T_1	T_6	T_8	T_1	T_4	T_8	T_3
3. Blokk	T_7	T_1	T_4	T_9	T_8	T_7	T_2	T_4	T_9	T_7	T_2	T_6
	I. ismétlés			II. ismétlés			III. ismétlés			IV. ismétlés		

5.7. ábra. Egy 3×3 -as kiegyensúlyozott latin téglakísérlet elrendezése, amely kilenc kezelést (T_1, T_2, \dots, T_9) foglal magában

A leggyakoribb a *négyzetes latintégla elrendezés* (square lattice) ($t = k$) és a *derékszögű latintégla elrendezés* (rectangular lattice) ($t = k - 1$). Hatékony elrendezések állnak rendelkezésre az r , t és k más értékeire is. Az általánosított latintégla-elrendezésnek ezt az osztályát *alfa elrendezésnek* hívják és alkalmazható a paraméter kombinációk tág körére. A CycDesignN software csomagot javasolják a latintégla-elrendezés tervezéséhez (Williams et al., 2002).

5.2. Faktoriális kísérletek

Azt a kísérletet, amelyben a kezelések tartalmazzák két vagy több faktor kiválasztott szintjeinek összes lehetséges kombinációját, faktoriális kísérletnek hívjuk. Például, egy kísérletet, amely két faktort foglal magában, mindegyiket két szinten, úgy mint két fajta és két nitrogén dózis, 2×2 -es vagy 2^2 faktoriális kísérletnek nevezzük. Kezeléseik tartalmazzák a két faktor két szintjének négy lehetséges kombinációját. Ha a 2^2 faktoriális kísérletet kiterjesztjük, hogy foglaljon magában egy harmadik faktort, mondjuk gyomirtást két szinten, a kísérlet $2 \times 2 \times 2$ vagy 2^3 faktoriális kísérlet lesz, a következő nyolc kezelés-kombinációval.

Jegyezzük meg, hogy a faktoriális kifejezés leír egy specifikus módot, amellyel a kezeléseket képezzük, és semmiképpen sem utal az alkalmazott kísérleti elrendezésre. Például, ha a szóban forgó 2^3 faktoriális kísérlet randomizált blokk elrendezésű akkor a kísérlet korrekt megnevezése 2^3 faktoriális kísérlet randomizált komplett blokk elrendezésben. A kezelések száma gyorsan nő a faktorok számának vagy minden faktorban a szintek számának növekedésével. Egy faktoriális kísérletben, amely magában foglal öt fajtát, négy nitrogén szintet és három gyomirtási módot, a kezelések összes száma $5 \times 4 \times 3 = 60$. Tehát kerüljük el az indokolatlan használatát a faktoriális kísérleteknek a nagy méretük, komplexitásuk és költségük miatt. Továbbá, nem bölcs dolog elkötelezni magunkat egy nagy kísérletre a vizsgálat kezdetén, amikor több kisebb előzetes kísérlet ugyanazt az eredményt adhatja. Tehát, jóllehet egy faktoriális kísérlet értékes információt szolgáltat az interakcióról, és kétségt kívül több információt nyújt, mint az egytényezős kísérlet, praktikus szempontok limitálhatják használatát.

A faktoriális kísérletek elrendezésére, az egytényezős kísérletekhez hasonlóan, alkalmazhatók a legegyszerűbb kísérleti elrendezések (pl. véletlen blokkelrendezés). Ilyen elrendezések alkalmazása esetén mindegyik tényező hatását, valamint az összes kölcsönhatásokat ugyanolyan pontossággal vizsgáljuk. Ha valamelyik tényező szintjei csak nagyobb egységeken alkalmazhatók, vagy ha egyes tényezők hatását, illetve kölcsönhatásait pontosabban kívánjuk vizsgálni, míg mások vizsgálatában kisebb pontossággal is megelégszünk, akkor az osztott-parcellás elrendezéseket, illetve ezek különböző változatait célszerű alkalmazni (split-plot, sávós, kétszeresen osztott parcellás elrendezés). A faktoriális kísérletek a belső ismétlések következtében, általában kevesebb ismétléssel állíthatók be, mint az egytényezős kísérletek (Yates, 1937).

5.2.1. Kéttényezős véletlen blokkelrendezés

Előnyök, hátrányok és felhasználás

Egyik legegyszerűbb kéttényezős kísérleti elrendezés. Az egytényezős véletlen blokkelrendezéstől abban különbözik, hogy itt az egyes kezelések két tényező vál-

tozatainak összes lehetséges kombinációi. Akkor alkalmazzuk, ha minden kombináció közötti különbséget azonos pontossággal kívánjuk elbírálni és ha mindkét tényező változatai közötti különbség elbírálására súlyt helyezünk. Az elrendezés technikai hátránya, hogy az egyik tényező kezelése kis parcellákon gyakran nem kivitelezhető. Egy 4×2 -es faktoriális kísérlet véletlen blokkelrendezésben az 5.8. ábrán látható. Két faktor van az ábrán: tenyészterület faktor négy szinten (S_1 - S_4) és nitrogén faktor két szinten (N_1 - N_2), a kezeléskombinációk száma 8.

I.		II.		III.	
S_4N_2	S_4N_1	S_3N_1	S_1N_1	S_3N_1	S_4N_2
S_3N_2	S_1N_1	S_1N_2	S_2N_1	S_4N_1	S_3N_2
S_2N_1	S_1N_2	S_2N_2	S_4N_2	S_2N_2	S_1N_1
S_2N_2	S_3N_1	S_3N_2	S_4N_1	S_1N_2	S_2N_1

5.8. ábra. 4×2 -es faktoriális kísérlet randomizált blokk elrendezésben, 3 ismétlésben. A parcellák kezelése a tenyészterület (S) és nitrogén műtrágya (N) faktorok kombinációi

A kétféteyezős véletlen blokkelrendezésű kísérlet statisztikai modellje:

A lineáris additív modell a következő formában írható fel:

$$(\text{termés}) = (\text{főátlag}) + (\text{ismétlés-hatás}) + (A \text{ tényező hatása}) + (B \text{ tényező hatása}) + (A,B \text{ kölcsönhatás}) + (\text{maradék})$$

$$y_{hij} = \mu + Bl_h + A_i + B_j + (AB)_{ij} + e_{hij} \quad h = 1 \dots r; i = 1 \dots a; j = 1 \dots b$$

ahol y_{hij} : a h-adik blokk adata (pl. termés), amely az A kezelés i-edik szintjét és a B kezelés j-edik szintjét kapja, μ : minden kezelés átlagos reakciója, Bl_h : a h-adik blokk hatása, A_i : az A kezelés átlagos hatása i-edik szinten, B_j : a B kezelés átlagos hatása j-edik szinten, $(AB)_{ij}$: az A (i-edik szinten) és B (j-edik szinten) kezelés közötti, interakció átlagos hatása, e_{hij} : random kísérleti hiba. Kétféteyezős randomizált blokk elrendezésű kísérlet varianciaanalízis táblázata:

Variáció forrása	df	SS	MS	F-érték
Blokk (Bl_h)	df_{Bl}	SS_{Bl}	$MS_{Bl} = SS_{Bl}/df_{Bl}$	
A faktor (A_i)	df_A	SS_A	$MS_A = SS_A/df_A$	$F_A = MS_A/MS_E$
B faktor (B_j)	df_B	SS_B	$MS_B = SS_B/df_B$	$F_B = MS_B/MS_E$
AB kölcsönhatás (AB_{ij})	df_{AB}	SS_{AB}	$MS_{AB} = SS_{AB}/df_{AB}$	$F_{AB} = MS_{AB}/MS_E$
Maradék (e_{hij})	df_E	$SS_E = SS_{TOT} - SS_{Bl} - SS_{AB} - SS_A - SS_B$	$MS_E = SS_E/df_E$	
Összes	df_{TOT}	SS_{TOT}		

5.2.2. Osztott parcellás (split-plot) elrendezés

Az osztott parcellás elrendezés különösen alkalmas a kéttényezős kísérletekre, melyek több kezelést tartalmaznak, mint amennyit a véletlen blokkelrendezésbe el lehetne helyezni. Az osztott parcellás elrendezésben az egyik faktort a *főparcellákba* helyezzük el. Ezt a faktort a *főparcella faktornak* hívjuk. A főparcellát *alparcellákra* osztjuk, amelybe a második faktort, az *alparcella faktort* helyezzük el. Tehát minden főparcella az alparcella kezelése (azaz az alparcella faktor szintjei) blokkjává válik. Az osztott parcellás elrendezésnél a főparcella faktor-hatás mérésének precizitását feláldozzuk azért, hogy javítsuk az alparcella faktor pontosságát. Az alparcella-faktor fő hatásának és interakciójának mérése a főparcella-faktorról sokkal pontosabb, mint egy randomizált blokkelrendezésnél. Másrészt, a főfaktor-kezelések (pl. a főparcella-faktor szintjei) hatásának mérése kevésbé pontos, mint a megfelelő véletlen blokkelrendezésnél. Mivel az osztott-parcellás elrendezésnél a parcella mérete és a hatások mérésének precizitása nem azonos mind a két faktorra, a megfelelő faktor megalapozott kijelölése akár főparcellának, akár alparcellának, rendkívül fontos. Egy osztott parcellás kísérlet elrendezése az 5.9. ábrán látható. A hat nitrogénszint (N_0-N_5) jelöli a főparcellákat és a négy fajta (V_1-V_4) az alparcellákat a főparcellákon belül.

N_4 N_3 N_1 N_0 N_5 N_2						N_1 N_0 N_5 N_2 N_4 N_3						N_0 N_1 N_4 N_5 N_3 N_2					
V_2	V_1	V_1	V_2	V_4	V_3	V_1	V_4	V_3	V_1	V_1	V_3	V_4	V_3	V_3	V_1	V_2	V_1
V_1	V_4	V_2	V_3	V_3	V_2	V_3	V_1	V_4	V_2	V_4	V_2	V_2	V_4	V_2	V_3	V_3	V_4
V_3	V_2	V_4	V_1	V_2	V_1	V_2	V_2	V_1	V_4	V_2	V_4	V_1	V_1	V_4	V_2	V_4	V_2
V_4	V_3	V_3	V_4	V_1	V_4	V_4	V_3	V_2	V_3	V_3	V_1	V_3	V_2	V_1	V_4	V_1	V_3
I. ismétlés						II. ismétlés						III. ismétlés					

5.9. ábra. A négy fajta (V_1 , V_2 , V_3 és V_4) random elhelyezése a hat nitrogén szint, mint főparcella alparcelláiba, mindhárom ismétlésben

Előnyök, hátrányok és felhasználás

Számos előnye van az osztott-parcellás kísérletnek: 1. Lehetővé teszi olyan faktorok hatékony használatát, amelyek különböző parcellaméreteket kívánnak az alkalmazásukhoz. 2. Lehetővé teszi új kezelések bevezetését a már korábban beállított kísérletekbe. 3. Nagyobb precizitást ad néhány faktoriális hatás becslésekor. *Ennek a tervezésnek természetesen vannak hátrányai is:* 1. A főparcella faktort kisebb precizitással becsüljük, mint az alparcella faktort. Ezáltal gyakran nagy különbségek szükségesek ahhoz, hogy a főparcella-faktor hatása szignifikáns legyen, miközben az alparcella faktornak szignifikáns lehet a hatása, amelynek viszont túl kicsi a gyakorlati jelentősége. 2. A statisztikai analízis komplexebb az RBD-nél, mivel különböző hiba-szórások (standard errors) szükségesek a különböző összehasonlításokhoz. *Számos felhasználási területe van az osztott parcellás elrendezésnek:* 1. Olyan kísérletekben, amelyekben egy faktort nagyobb parcellákon kell beállítani, mint más faktorokat. 2. Egy új faktort állítunk be a már meglévő kísérletbe.

Az osztott parcellás elrendezés statisztikai modellje:

(termés) = (főátlag) + (ismétlés-hatás) + (A tényező hatása) + (főparcella maradék) + (B tényező hatása) + (A.B kölcsönhatás) + (alparcella maradék)

$$y_{hij} = \mu + Bl_h + A_i + em_{hi} + B_j + (AB)_{ij} + e_{hij} \quad h = 1 \dots r; i = 1 \dots a; j = 1 \dots b$$

ahol: μ : az összes kezelés átlagos reakciója, Bl_h : a h-adik blokk hatása, A_i : az A kezelés i-edik szintjének átlagos hatása, B_j : a B kezelés j-edik szintjének átlagos hatása, $(AB)_{ij}$: az A (i-edik szinten) és B (j-edik szinten) kezelés közötti interakció átlagos hatása, em_{hi} : a fő-parcella hiba, e_{hij} : alparcella hiba.

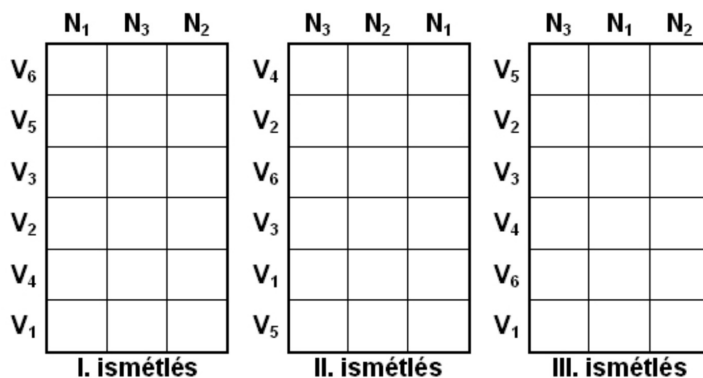
Az osztott parcellás kísérlet varianciatáblázata:

Variáció forrása	df	SS	MS	F-érték
Blokk (Bl_h)	df_{Bl}	SS_{Bl}	$MS_{Bl} = SS_{Bl} / df_{Bl}$	
A faktor (A_i)	df_A	SS_A	$MS_A = SS_A / df_A$	$F_A = MS_A / MS_{EM}$
Főparcella maradék (em_{hi})	df_{EM}	$SS_{EM} = SS_{MTOT} - SS_A - SS_{Bl}$	$MS_{EM} = SS_{EM} / df_{EM}$	
Főparcella összes	df_{MTOT}	SS_{MTOT}		
B faktor (B_j)	df_B	SS_B	$MS_B = SS_B / df_B$	$F_B = MS_B / MS_E$
AB interakció (AB_{ij})	df_{AB}	SS_{AB}	$MS_{AB} = SS_{AB} / df_{AB}$	$F_{AB} = MS_{AB} / MS_E$
Alparcella maradék (e_{hij})	df_E	$SS_E = SS_{TOT} - SS_B - SS_{AB} - SS_{MTOT}$	$MS_E = SS_E / df_E$	
Összes	df_{TOT}	SS_{TOT}		

5.2.3. Sávós (strip-plot vagy criss-cross) elrendezés

A sávós kísérleti elrendezés különösen alkalmas a kéttényezős kísérletre, amelyben a megkívánt precizitás a két faktor közötti *interakció hatás mérésére* nagyobb, mint a két faktor közül bármelyik fő hatásának mérésére. Ezt *három parcellaméret* használatával valósítjuk meg: 1. Vertikális-sávós parcella az első faktorra (*vertikális faktor*). 2. Horizontális-sávós parcella a második faktorra (*horizontális faktor*). 3. Kölcsönhatás parcella (két faktor közötti kölcsönhatásra). A vertikális-sávós parcella és a horizontális-sávós parcella mindig merőleges egymásra. Nincs azonban összefüggés a méreteik között, nem úgy mint az osztott parcellás elrendezés főparcellája és alparcellája esetében. Tehát, a sávós elrendezésnél, a precizitás mértékét mindkét faktor főhatásával összefüggésben feláldozzuk azért, hogy javítsuk az interakció hatás precizitását.

A sávós elrendezést az 5.10. ábra szemlélteti. Az egyik kezelés faktor (fajta, V) randomizálva van a sorokra mindegyik ismétlésen belül. A másik kezelés faktor (nitrogén, N) randomizálva van az oszlopokra mindegyik ismétlésen belül. Az ilyen blokk-struktúra sorai és oszlopai egyidejűleg és egymástól függetlenül randomizáltak az ismétlésen belül.



5.10. ábra. A sávos kísérletben a három nitrogén dózis (N_1 , N_2 és N_3) random elhelyezése a vertikális sávokba és hat fajta (V_1 , V_2 , V_3 , V_4 , V_5 és V_6) random elhelyezése a horizontális sávokba, mindhárom ismétlésben

A sávos elrendezésű kísérlet variancia táblázatának szerkezete az alábbi:

Variáció forrása	Szabadságfok	Eltérés négyzet összeg	Közepes négyzetes eltérés	Számított F-érték	Táblázati F-érték	
					5%	1%
Ismétlés	$r-1 = 2$					
Horizontális faktor (A)	$a-1=5$					
Hiba (a)	$(r-1)(a-1)=10$					
Vertikális faktor (B)	$b-1=2$					
Hiba (b)	$(r-1)(b-1)=4$					
A × B	$(a-1)(b-1)=10$					
Hiba (c)	$(r-1)(a-1)(b-1)=20$					
Összes	$rab-1=53$					

5.2.4. Három- és többtényezős kísérletek

A kéttényezős kísérlet kiterjeszhető, hogy magában foglaljon egy harmadik faktort, a háromtényezős kísérlet kiterjeszhető, hogy magában foglaljon egy negyedik faktort, és így tovább. Két fontos következménye van azonban annak, amikor további faktorokat adunk egy kísérlethez:

1. Gyors növekedés van a kezelések számában, amelyeket tesztelnünk kell.
2. Növekedés következik be az interakciók számában és típusában. Például, egy háromtényezős kísérletben négy interakció hatást vizsgálhatunk. Egy négytényezős kísérletben tíz interakció hatás van.

Jóllehet nem kívánatos egy nagy kísérlet magas költsége és komplexitása miatt, ugyanakkor nagyon értékes lehet az információ-növekedés, melyet a faktorok közötti interakció hatásából kapunk. Következésképpen, a kutató döntése a kísérletbe bevonni kívánt faktorok számáról egy kompromisszumon alapul, amely figyelembe veszi a lehető legtöbb interakció hatását, másrészt a kísérlet méretét a rendelkezésre álló források korlátain belül tartja.

A fő hatások és az alacsonyabb rendű interakciók kevésbé értelmezhetővé válnak, ha szignifikáns interakcióba vannak bevonva. Az interakciók rendje eggyel kevesebb, mint a kezelés faktorok száma, amelyek be vannak vonva az interakcióba. Így két kezelés faktor közötti interakció elsőrendű interakció, míg a négy kezelés faktor közötti harmadrendű interakció.

Előnyök

Számos főhatás vizsgálható egy kísérletben, feltéve, hogy az interakció tagok nem szignifikánsak. Többszörös interakciók számos kezelésfaktor között vizsgálhatók.

Hátrányok

A kísérletek nagyon nagygyá és komplexekké válhatnak. A magasabb rendű interakciókat bonyolult értelmezni. Minden kezelésfaktor kombináció egyenlő ismétlése szükséges, jöllehet becslési eljárások használhatók, ha csak néhány hiányzó érték van.

Randomizáció

Hasonlóan a kéttényezős faktoriális kísérlethez, a kezelés faktor kombinációk randomizáltak a kísérleti egységekre a kísérleti tervnek megfelelően (teljesen randomizált, randomizált blokk vagy latin négyzet elrendezés).

A varianciaanalízist ugyanúgy végezzük, mint a kéttényezős faktoriális kísérleteknél és ezáltal ugyanaz a formája is. A kezelés hatások értelmezése hasonló a kéttényezős faktoriális kísérlethez. A nem szignifikáns eredményeket ugyanúgy értelmezzük. Ha azonban az interakció hatások statisztikailag szignifikánsak, akkor bármely főhatás vagy alacsonyabb rangú interakció hatás (a szignifikáns interakcióval összefüggésben) értelmezése nem vagy csak korlátozott mértékben lehetséges. Ezáltal a többtényezős faktoriális elrendezésű kísérlet értelmezése mindig a legmagasabb interakció taggal kezdődik. Az interpretáció azután folytatódik a következő rendű interakció taggal és így tovább, egészen a főhatások értelmezéséig. Mindegyik szakaszban a szignifikáns interakciók jelenléte megakadályozhatja bármely hatás értelmezését, amely hozzájárult ehhez az interakcióhoz.

Számos kísérleti elrendezés van, amely használható a három- vagy többtényezős kísérletekben. Ezek az elrendezések négy kategóriába sorolhatók, nevezetesen, egyfaktoros, kétfaktoros, háromfaktoros kísérleti elrendezések és frakcionális faktoriális elrendezések.

Egytényezős kísérleti elrendezések alkalmazása

Minden egytényezős kísérleti elrendezés alkalmazható a három- vagy többfaktoros kísérletekhez. Minden faktoriális kezelés-kombinációt úgy tekintünk, mintha egyetlen faktor szintjei lennének.

Illusztrációként vegyünk egy háromtényezős kísérletet, amely magában foglal két fajtát, négy nitrogén szintet és három gyomirtási módszert, három ismétlésben. Ha randomizált komplett blokkelrendezést (RCB) használunk, a $2 \times 4 \times 3 = 24$ faktoriális

kezeléskombinációt teljesen véletlenül jelöljük ki a 24 kísérleti parcellára, mindhárom ismétlésben. Egy szántóföldi kísérleti elrendezést a 5.11. ábra mutatja.

$V_2N_0W_3$	$V_2N_3W_1$	$V_1N_3W_2$	$V_1N_2W_3$	$V_1N_3W_3$	$V_2N_2W_3$	$V_1N_3W_1$	$V_1N_2W_1$	$V_2N_0W_2$	$V_2N_2W_3$	$V_1N_0W_2$	$V_1N_3W_2$	$V_1N_2W_1$	$V_2N_3W_1$	$V_1N_1W_1$	$V_2N_3W_3$	$V_1N_1W_2$	$V_2N_2W_3$
$V_1N_1W_1$	$V_2N_1W_1$	$V_1N_0W_3$	$V_2N_0W_2$	$V_1N_0W_1$	$V_2N_2W_1$	$V_2N_1W_1$	$V_2N_3W_3$	$V_2N_1W_2$	$V_1N_0W_3$	$V_1N_2W_3$	$V_2N_2W_1$	$V_2N_3W_2$	$V_1N_0W_2$	$V_2N_2W_2$	$V_2N_1W_3$	$V_2N_0W_2$	$V_2N_0W_3$
$V_1N_1W_2$	$V_1N_0W_2$	$V_1N_1W_3$	$V_1N_2W_1$	$V_2N_3W_3$	$V_2N_1W_3$	$V_2N_3W_2$	$V_2N_1W_3$	$V_2N_0W_3$	$V_1N_1W_1$	$V_1N_1W_2$	$V_1N_1W_3$	$V_2N_0W_1$	$V_1N_1W_3$	$V_2N_2W_1$	$V_1N_2W_2$	$V_1N_0W_3$	$V_1N_3W_3$
$V_1N_2W_2$	$V_1N_3W_1$	$V_2N_3W_2$	$V_2N_2W_2$	$V_2N_0W_1$	$V_2N_1W_2$	$V_1N_0W_1$	$V_2N_0W_1$	$V_1N_3W_3$	$V_2N_3W_1$	$V_1N_2W_2$	$V_2N_2W_2$	$V_2N_1W_1$	$V_1N_3W_2$	$V_1N_3W_1$	$V_1N_2W_3$	$V_2N_1W_2$	$V_1N_0W_1$
I. ismétlés						II. ismétlés						III. ismétlés					

5.11. ábra. A $2 \times 4 \times 3$ -as faktoriális kísérlet, amely magában foglal két fajtát (V_1, V_2), négy nitrogén szintet (N_0, N_1, N_2, N_3 és N_4) és három gyomirtási módszert (W_1, W_2 és W_3) randomizált blokk elrendezésben, három ismétlésben

A variancia-táblázat szerkezete a $2 \times 4 \times 3$ -as faktoriális kísérletre véletlen blokk (RCB) elrendezésben:

Variáció forrása	Szabadságfok	Eltérés négyzet összeg	Közepes négyzetes eltérés	Számított F-érték	Táblázati F-érték	
					5%	1%
Ismétlés	$r-1 = 2$					
Kezelés	$vnw-1 = 23$					
Fajta (V)	$v-1 = 1$					
Nitrogén (N)	$n-1 = 3$					
Gyomirtás (W)	$w-1 = 2$					
$V \times N$	$(v-1)(n-1)=3$					
$V \times W$	$(v-1)(w-1)=2$					
$N \times W$	$(n-1)(w-1)=6$					
$V \times N \times W$	$(v-1)(n-1)(w-1)=6$					
Hiba	$(r-1)(vnw-1)=46$					
Összes	$rvnw-1=71$					

r = ismétlések száma; v , n és w a három faktor (V, N és W) szintjei

Jegyezzük meg, hogy az RCB elrendezésnél csak egy parcellaméret van és csupán egy hiba variancia minden hatás (azaz három főhatás, három kétfaktoros interakció hatás és egy háromfaktoros interakció hatás) szignifikanciájának tesztelésére, úgy-hogy minden hatást azonos pontossággal mérünk. Tehát egy komplett blokk elrendezést, mint az RCB, akkor indokolt használni, ha:

- Minden hatás (azaz, főhatás és kölcsönhatás) egyformán fontos és ezáltal ugyanolyan pontossággal mérhető.
- A kísérleti egységek elég homogének ahhoz, hogy a homogenitás magas szintjét érjük el blokkon belül.

Mivel a három- vagy többfaktoros kísérlet magában foglal nagyszámú kezelést, a homogenitást a kísérleti egységekben ugyanazon a blokkon belül bonyolult elérni, és ezáltal a komplett blokk elrendezést nem túl gyakran használják.

Kétfényező kísérleti elrendezések alkalmazása

Mindegyik kétfaktoros kísérleti elrendezés alkalmazható a három- vagy többfaktoros kísérletekhez. Bemutatjuk egy háromtényező kísérlet tervezésének eljárásait a már leírt $2 \times 4 \times 3$ -as faktoriális kísérlettel.

1. Lépés. Osszuk fel a tesztelni kívánt k faktort két csoportra, k_1 faktorok az egyik csoportba és k_2 faktorok egy másik csoportba (ahol $k_1 + k_2 = k$). Helyezzük azonos csoportba azokat a faktorokat, amelyeket azonos pontossággal akarunk mérni. Mindegyik csoport bármilyen számú faktort tartalmazhat.

Példánkban $k = 3$ faktor, fajta és gyomirtás kombinálható egy csoportba és a megmaradó faktor, a nitrogén, egy másik csoportba. Tehát az I. csoport tartalmaz két faktort ($k_1 = 2$) és a II. csoport tartalmaz egy faktort ($k_2 = 1$).

2. Lépés. Tekintsük a k_1 faktorok faktoriális kezelés kombinációját az I. csoportban egyetlen faktor szintjeinek, amelyet A-nak hívunk és a k_2 faktorok faktoriális kombinációját a II. csoportban egyetlen faktor szintjeinek, melyet B faktornak hívunk. Tehát, a k -faktoros kísérletet átalakítottuk kétfaktoros kísérletté, amely magában foglal két újonnan kialakított faktort, A- és B-t.

$V_1 W_2$	$V_2 W_1$	$V_1 W_3$	$V_2 W_3$	$V_1 W_1$	$V_2 W_2$	$V_1 W_2$	$V_2 W_2$	$V_2 W_3$	$V_1 W_1$	$V_1 W_3$	$V_2 W_1$	$V_2 W_1$	$V_2 W_2$	$V_1 W_2$	$V_1 W_1$	$V_2 W_3$	$V_1 W_3$
N_2	N_3	N_1	N_2	N_3	N_1	N_2	N_3	N_0	N_0	N_2	N_3	N_2	N_3	N_0	N_1	N_3	N_2
N_0	N_2	N_0	N_3	N_0	N_2	N_3	N_1	N_1	N_2	N_1	N_1	N_3	N_1	N_3	N_2	N_1	N_0
N_3	N_1	N_2	N_0	N_2	N_0	N_1	N_0	N_3	N_1	N_3	N_0	N_0	N_2	N_1	N_3	N_0	N_1
N_1	N_0	N_3	N_1	N_1	N_3	N_0	N_2	N_2	N_3	N_0	N_2	N_1	N_0	N_2	N_0	N_2	N_3
I. ismétlés						II. ismétlés						III. ismétlés					

5.12. ábra. A $2 \times 4 \times 3$ -as faktoriális kísérlet, amely magában foglal két fajtát (V_1, V_2), négy nitrogén szintet (N_0, N_1, N_2, N_3 és N_4) és három gyomirtási módszert (W_1, W_2 és W_3), osztott parcellás elrendezésben, nitrogén szintekkel az alparcellában, három ismétlésben

Példánkra, a $3 \times 2 = 6$ faktoriális kezeléskombinációt a fajta és a gyomirtás között az I. csoportban úgy tekintjük, mint az újonnan képzett A faktor hat szintjét, és a nitrogén négy szintjét, mint az újonnan képzett B faktor szintjeit. Tehát a $2 \times 4 \times 3$ -as faktoriális kísérlet most úgy tekinthető, mint egy 6×4 -es faktoriális kísérlet, amely A és B faktorokat foglal magában.

3. Lépés. Válasszunk ki egy megfelelő kétfényező kísérleti elrendezést és alkalmazzuk a szimulált kétfaktoros kísérlethez, melyet a 2. lépésben alakítottunk ki.

Példánkra, ha az osztott parcellás elrendezést a B (nitrogén) faktornal, mint alparcella faktornal használjuk, a kísérleti elrendezést az 5.12. ábra mutatja és a megfelelő varianciaanalízis szerkezete az alábbi:

Variáció forrása	Szabadságfok	Eltérés négyzet összeg	Közepes négyzetes eltérés	Számított F-érték	Táblázati F-érték	
					5%	1%
Ismétlés	$r-1 = 2$					
Főparcella faktor	$vw - 1 = 5$					
Fajta (V)	$v-1 = 1$					
Gyomirtás (W)	$w-1 = 2$					
V x W	$(v-1)(w-1) = 2$					
Hiba (a)	$(r-1)(vw-1) = 10$					
Alparcella faktor (N)	$n-1 = 3$					
N x V	$(n-1)(v-1) = 3$					
N x W	$(n-1)(w-1) = 6$					
N x V x W	$(n-1)(v-1)(w-1) = 6$					
Hiba (b)	$vw(r-1)(n-1) = 36$					
Összes	$rvnw-1 = 71$					

Szimulált kétfaktoros kísérlet, melyben a főparcella faktor két eredeti faktor (V és W) kombinációja, és az alparcella faktor reprezentálja a harmadik faktort (N). r = ismétlések száma; v , n és w a három faktor (V, N és W) szintjei

Jegyezzük meg, hogy az osztott parcellás elrendezésben két parcellaméret és két hiba közepes négyzetes eltérés (MS) van a különböző hatások szignifikanciájának tesztelésére:

- Hiba (a) MS a fajta főhatásának, a gyomirtási eljárás főhatásának és az interakció hatásuknak a tesztelésére.
- Hiba (b) MS a nitrogén műtrágya főhatásának és interakciójának tesztelésére a másik két faktoralal.

Mivel a hiba (b) MS várhatóan kisebb, mint a hiba (a) MS, a pontosság mértéke a nitrogén hatásának mérésére várhatóan nagyobb, mint a fajtáé vagy a gyomirtásé. Tehát, az osztott parcellás elrendezés (vagy a sávos elrendezés) megfelelő a három- vagy többfaktoros kísérletre, ha az alábbi két feltétel teljesül:

- A faktoriális kezeléskombinációk összes száma túl nagy egy véletlen blokk elrendezéshez.
- A k faktor könnyen felosztható két csoportra, azonosítható különbségekkel a pontosság kívánt szintjével mindegyik csoportra.

Három- vagy többtényezős kísérleti elrendezések alkalmazása

A három- vagy többfaktoros kísérletekre kifejlesztett specifikus kísérleti elrendezések a mezőgazdasági kutatásban az *osztott parcellás* vagy a *sávos elrendezés* kiterjesztései.

Példánkra, az osztott parcellás elrendezés kiterjeszthető, hogy behelyezzük a harmadik faktort mindegyik alparcella további felosztásán keresztül al-alparcellába és kiterjeszthető tovább, hogy beépítsük a negyedik faktort mindegyik al-alparcella további felosztásán keresztül al-al-alparcellába, és így tovább. Az eredményül kapott elrendezést kétszeresen osztott parcellásnak (split-split-plot), háromszorosan osztott

parcellásnak (split-split-split-plot) hívjuk. Egy kétszeresen osztott parcellás elrendezést háromfaktoros kísérlethez használunk, az első faktort elhelyezzük a főparcellában, a második faktort az alparcellában és a harmadik faktort az al-alparcellában. Ily módon nincs szükség a három faktor kombinálására két csoportba, hogy szimuláljunk egy kétfaktoros kísérletet, ahogy ez szükséges, ha az osztott parcellás elrendezést egy háromfaktoros kísérlethez használjuk. Hasonlóan, a sávos (strip-plot) elrendezés kiterjeszhető, hogy beépítsük a harmadik faktort, a negyedik faktort, és így tovább, minden interakció parcella felosztásán keresztül alparcellára és mindegyik alparcella további felosztásával al-alparcellára. A kapott elrendezés sávos-osztott parcellás (strip-split-plot) elrendezés, sávos kétszeresen osztott-parcellás (strip-split-split-plot) elrendezés stb.

$N_2M_1V_1$	$N_2M_3V_3$	$N_2M_2V_3$	$N_1M_2V_1$	$N_1M_3V_1$	$N_1M_1V_2$	$N_3M_1V_2$	$N_3M_2V_2$	$N_3M_3V_1$
$N_2M_1V_3$	$N_2M_3V_1$	$N_2M_2V_2$	$N_1M_2V_2$	$N_1M_3V_2$	$N_1M_1V_1$	$N_3M_1V_3$	$N_3M_2V_3$	$N_3M_3V_3$
$N_2M_1V_2$	$N_2M_3V_2$	$N_2M_2V_1$	$N_1M_2V_3$	$N_1M_3V_3$	$N_1M_1V_3$	$N_3M_1V_1$	$N_3M_2V_1$	$N_3M_3V_2$
$N_1M_3V_2$	$N_1M_1V_1$	$N_1M_2V_1$	$N_5M_2V_1$	$N_5M_1V_2$	$N_5M_3V_1$	$N_5M_2V_1$	$N_5M_1V_1$	$N_5M_3V_3$
$N_1M_3V_1$	$N_1M_1V_3$	$N_1M_2V_2$	$N_5M_2V_3$	$N_5M_1V_3$	$N_5M_3V_2$	$N_5M_2V_3$	$N_5M_1V_3$	$N_5M_3V_2$
$N_1M_3V_3$	$N_1M_1V_2$	$N_1M_2V_3$	$N_5M_2V_2$	$N_5M_1V_1$	$N_5M_3V_3$	$N_5M_2V_2$	$N_5M_1V_2$	$N_5M_3V_1$
$N_5M_3V_1$	$N_5M_2V_1$	$N_5M_1V_3$	$N_4M_2V_2$	$N_4M_3V_3$	$N_4M_1V_3$	$N_1M_2V_2$	$N_1M_3V_1$	$N_1M_1V_3$
$N_5M_3V_3$	$N_5M_2V_2$	$N_5M_1V_2$	$N_4M_2V_1$	$N_4M_3V_1$	$N_4M_1V_1$	$N_1M_2V_3$	$N_1M_3V_2$	$N_1M_1V_2$
$N_5M_3V_2$	$N_5M_2V_3$	$N_5M_1V_1$	$N_4M_2V_3$	$N_4M_3V_2$	$N_4M_1V_2$	$N_1M_2V_1$	$N_1M_3V_3$	$N_1M_1V_1$
$N_4M_2V_1$	$N_4M_1V_1$	$N_4M_3V_3$	$N_2M_3V_2$	$N_2M_2V_2$	$N_2M_1V_1$	$N_4M_2V_2$	$N_4M_3V_3$	$N_4M_1V_2$
$N_4M_2V_2$	$N_4M_1V_3$	$N_4M_3V_1$	$N_2M_3V_3$	$N_2M_2V_3$	$N_2M_1V_2$	$N_4M_2V_1$	$N_4M_3V_2$	$N_4M_1V_3$
$N_4M_2V_3$	$N_4M_1V_2$	$N_4M_3V_2$	$N_2M_3V_1$	$N_2M_2V_1$	$N_2M_1V_3$	$N_4M_2V_3$	$N_4M_3V_1$	$N_4M_1V_1$
$N_3M_1V_3$	$N_3M_2V_1$	$N_3M_3V_3$	$N_3M_2V_3$	$N_3M_3V_3$	$N_3M_1V_1$	$N_2M_3V_3$	$N_2M_1V_3$	$N_2M_2V_2$
$N_3M_1V_1$	$N_3M_2V_3$	$N_3M_3V_1$	$N_3M_2V_2$	$N_3M_3V_1$	$N_3M_1V_2$	$N_2M_3V_2$	$N_2M_1V_1$	$N_2M_2V_3$
$N_3M_1V_2$	$N_3M_2V_2$	$N_3M_3V_2$	$N_3M_2V_1$	$N_3M_3V_2$	$N_3M_1V_3$	$N_2M_3V_1$	$N_2M_1V_2$	$N_2M_2V_1$

I. ismétlés

II. ismétlés

III. ismétlés

5.13. ábra. Az $5 \times 3 \times 3$ -as faktoriális kísérlet split-split-plot elrendezésben, öt nitrogén szinttel (N_1 , N_2 , N_3 , N_4 és N_5), mint főparcella kezelés, három agrotechnikai eljárással (M_1 , M_2 és M_3), mint alparcella kezelés és három fajtaival (V_1 , V_2 és V_3), mint al-alparcella kezelés, három ismétlésben

Kétszeresen osztott parcellás (split-split-plot) elrendezés

A kétszeresen osztott parcellás elrendezés az osztott parcellás (split-plot) elrendezés kiterjesztése, hogy beillesszük a harmadik faktort. Különösen megfelelő az olyan háromfaktoros kísérletekre, ahol három különböző szintű pontosság kívánatos a különböző hatásokra. A kísérlet elrendezését bemutatjuk egy $5 \times 3 \times 3$ -as faktoriális kísérlettel, három ismétlésben (5.13. ábra). A kezelések magukba foglalják az öt nitro-

gén szintet, mint főparcellát, a három agrotechnikai eljárást, mint alparcellát és három búzafajtát, mint al-alparcellát. A kétszeresen osztott parcellás elrendezés variancia táblázatának szerkezete:

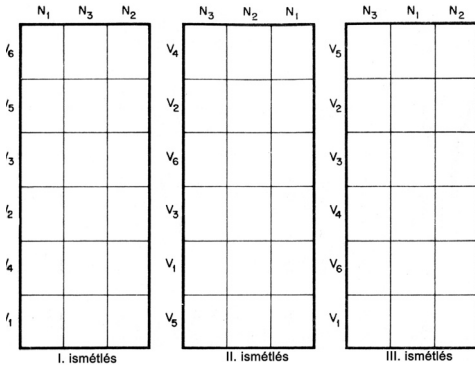
Variáció forrása	Szabadságfok	Eltérés négyzet összeg	Közepes négyzetes eltérés	Számított F-érték	Táblázati F-érték	
					5%	1%
Főparcella analízis:						
Ismétlés	$r-1 = 2$					
Főparcella faktor (A)	$a-1=4$					
Hiba (a)	$(r-1)(a-1)=8$					
<i>Alparcella analízis:</i>						
Alparcella faktor (B)	$b-1=2$					
A × B	$(a-1)(b-1)=8$					
Hiba (b)	$a(r-1)(b-1)=20$					
<i>Al-alparcella analízis:</i>						
Al-alparcella faktor(C)	$c-1=2$					
A × C	$(a-1)(c-1)=8$					
B × C	$(b-1)(c-1)=4$					
A × B × C	$(a-1)(b-1)(c-1)=16$					
Hiba (c)	$ab(r-1)(c-1)=60$					
Összes	$rabc-1=134$					

Sávós-osztott parcellás (strip-split-plot) elrendezés

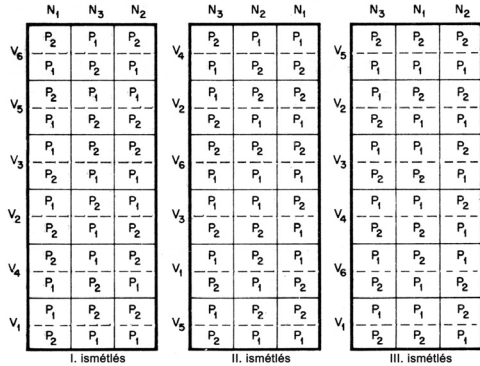
A sávós-osztott parcellás elrendezés a sávós elrendezés kiterjesztése, amelyben a kereszteződési (interakció) parcellát felosztjuk alparcellákra, hogy beillesszünk egy harmadik faktort. A sávós-parcellás elrendezést két fő tulajdonság jellemzi:

1. Négy parcellaméret van – a horizontális sáv, a vertikális sáv, a kereszteződési parcella és az alparcella.
2. Négy szintje van a pontosságoknak, amellyel a különböző faktorok hatásait mérhetjük, a legmagasabb szint megfelel az alparcella faktornak és a többi faktornal alkotott interakcióinak.

Példa. Egy háromfaktoros kísérletben vizsgálni kívánták két vetési mód (P_1 és P_2), három nitrogén dózis (N_1 , N_2 és N_3) hatását hat búzafajta (V_1 , V_2 , V_3 , V_4 , V_5 és V_6) termesére. A kísérletet három ismétlésben állították be, nitrogén volt a vertikális faktor, a fajta a horizontális faktor és a vetési mód az alparcella. A sávós-osztott parcellás kísérlet elrendezésének lépéseit az 5.14.–5.15. ábra mutatja.



5.14. ábra. A hat fajta (V₁, V₂, V₃, V₄, V₅ és V₆) random elhelyezése a horizontális sávba és a három nitrogén dózis (N₁, N₂ és N₃) random elhelyezése a vertikális sávba mindhárom ismétlésben, mint a sávos-osztott parcellás kísérlet első lépése



5.15. ábra. A 3 × 6 × 2-es faktoriális kísérlet sávos-osztott parcellás elrendezésben, hat fajttal (V₁, V₂, V₃, V₄, V₅ és V₆), mint horizontális kezelés, három nitrogén dózissal (N₁, N₂ és N₃), mint vertikális kezelés és két vetési móddal (P₁ és P₂), mint alparcella kezelés, három ismétlésbe

A sávos-osztott parcellás kísérlet variancia táblázatának szerkezete:

Variáció forrása	Szabadságfok	Eltérés négyzet összeg	Közepes négyzetes eltérés	Számított F-érték	Táblázati F-érték	
					5%	1%
Ismétlés	$r-1 = 2$					
Vertikális faktor (A)	$a-1=2$					
Hiba (a)	$(r-1)(a-1)=4$					
Horizontális faktor (B)	$b-1=5$					
Hiba (b)	$(r-1)(b-1)=10$					
A × B	$(a-1)(b-1)=10$					
Hiba (c)	$(r-1)(a-1)(b-1)=20$					
Al-alparcella faktor(C)	$c-1=1$					
A × C	$(a-1)(c-1)=2$					
B × C	$(b-1)(c-1)=5$					
A × B × C	$(a-1)(b-1)(c-1)=10$					
Hiba (d)	$ab(r-1)(c-1)=36$					
Összes	$rbc-1=107$					

Frakcionális faktoriális elrendezés

Eltérően az 1–3. pontban bemutatott elrendezésektől, ahol a faktoriális kezeléskombinációk teljes sorozatát bevontuk a vizsgálatba, a frakcionális faktoriális elrendezés (FFD), mint a neve is utal rá, a faktoriális kezeléskombinációk teljes sorozatának csupán egy frakcióját foglalja magában. Az FFD nyilvánvaló előnye a kísérlet

méretének csökkentése, amely kívánatos lehet, amikor a faktoriális kezeléskombinációk komplett (teljes) sorozata túl nagy a gyakorlati megvalósításra. Ezt a fontos előnyt azonban a becsülhető hatások számának csökkentésével érjük el. A kísérlet elrendezését részletesen ismerteti *Gomez és Gomez* (1984).

5.3. Kísérletsorozatok

Kísérletek olyan sorozata, amelyet ugyanannak a kérdésnek a vizsgálatára azonos kezelésekkel folytatnak le térben és/vagy időben megismételve. A szántóföldi kísérletek pl. több helyen és több évben megismételhetők. A legtöbb biológiai kísérleti kérdésre (melyet nagyobb számú nem ellenőrizhető környezeti hatás befolyásol) csak kísérletsorozat eredményeiből kaphatunk kielégítő választ.

Fontos teljesen meghatározni a statisztikai modellt, még akkor is, ha a kutató nem érdekelt az évek vagy helyek főhatásának tesztelésében. Ahhoz, hogy korrekten analízünk egy kísérletet, minden tényezőt számba kell venni, mivel a legtöbb számítógépes program automatikusan összevon valamennyi figyelembe nem vett variáció forrást a hiba tényezővel. A kísérletsorozatok értékelését részletesen ismerteti *Sváb* (1981). *McIntosh* (1982) bemutat komplett varianciaanalízis táblázatokat az egytényezős és osztott parcellás kísérletsorozatokra, amelyek felhasználhatók a variancia források és a megfelelő F-arányok azonosításában.

Több évben és több helyen megismételt kísérlet egy lehetséges additív lineáris modellje a következő:

$$Y_{ijkl} = \mu + Y_i + L_j + YL_{ij} + B_{(ij)k} + T_l + YT_{il} + LT_{jl} + YLT_{ijl} + BT_{(ij)kl}$$

ahol: Y_{ijkl} : az $ijkl$ -edik kísérleti egységen (parcella) mért reakció, μ : főátlag, Y_i : az i -edik év hatása, L_j : a j -edik hely hatása, YL_{ij} : az Y i -edik szintjének és L_j -edik szintjének interakciója, $B_{(ij)k}$: k -edik blokk hatása az i -edik éven és j -edik helyen belül, T_l : az l -edik kezelés hatása, YT_{il} : az Y i -edik szintjének és T l -edik szintjének interakciója, LT_{jl} : az L j -edik szintjének és T l -edik szintjének interakciója, YLT_{ijl} : Y i -edik szintjének, L j -edik szintjének és T l -edik szintjének interakciója, $BT_{(ij)kl}$: kísérleti hiba, $NID(0, \sigma_{BT}^2)$.

Az egy- és kéttényezős kísérletsorozatok variancia táblázatának szerkezetét az alábbi példák szemléltetik (MSTAT-C, 1991):

1. Egytényezős, véletlen blokkelrendezésű kísérletsorozat variancia táblázata a helyek (vagy évek) figyelembevételével

Variancia forrása	df
Hely (L)	$l-1$
Hiba	$l(r-1)$
A tényező	$a-1$
LA	$(l-1)(a-1)$
Hiba	$l(r-1)(a-1)$

2. Egytényezős, véletlen blokkrendezésű (RCBD) kísérletsorozat variancia táblázata, kombinálva helyek és évek szerint, új hely minden évben

Variancia forrása	df
Év (Y)	$y-1$
L(Y)	$y(l-1)$
R (LY)	$y/(r-1)$
A tényező	$a-1$
YA	$(y-1)(a-1)$
LA(Y)	$y(l-1)(a-1)$
Hiba	$y(r-1)(a-1)$

3. Egytényezős, véletlen blokkrendezésű (RCBD) kísérletsorozat variancia táblázata, kombinálva helyek és évek szerint, ugyanaz a hely minden évben

Variancia forrása	df
Év (Y)	$y-1$
Hely (L)	$l-1$
YL	$(y-1)(l-1)$
R(LY)	$y/(r-1)$
A tényező	$a-1$
YA	$(y-1)(a-1)$
LA	$(l-1)(a-1)$
YLA	$(y-1)(l-1)(a-1)$
Hiba	$y/(r-1)(a-1)$

4. Kéttényezős, véletlen blokkrendezésű (RCBD) kísérletsorozat variancia táblázata, a helyek (vagy évek) figyelembevételével:

Variancia forrása	df
Hely (L)	$l-1$
R(L)	$l/(r-1)$
A tényező	$a-1$
LA	$(l-1)(a-1)$
B tényező	$b-1$
LB	$(l-1)(b-1)$
AB	$(a-1)(b-1)$
LAB	$(l-1)(a-1)(b-1)$
Hiba	$l/(r-1)(a-1)(b-1)$

5. Kéttényezős, véletlen blokkelrendezésű (RCBD) kísérletsorozat variancia táblázata, a helyek és évek figyelembevételével, új hely minden évben

Variancia forrása	df
Év (Y)	$y-1$
L(Y)	$y(l-1)$
R(LY)	$y(l(r-1))$
A tényező	$a-1$
YA	$(y-1)(a-1)$
LA(Y)	$y(l-1)(a-1)$
B tényező	$b-1$
YB	$(y-1)(b-1)$
LB(Y)	$y(l-1)(b-1)$
AB	$(a-1)(b-1)$
YAB	$(y-1)(a-1)(b-1)$
LAB(Y)	$y(l-1)(a-1)(b-1)$
Hiba	$y(l(r-1))(ab-1)$

6. Kéttényezős, véletlen blokkelrendezésű (RCBD) kísérletsorozat variancia táblázata, a helyek és évek figyelembevételével, ugyanaz a hely, azonban randomizálva minden évben

Variancia forrása	df
Év (Y)	$y-1$
Hely (L)	$l-1$
YL	$(y-1)(l-1)$
R(LY)	$y(l(r-1))$
A tényező	$a-1$
YA	$(y-1)(a-1)$
LA	$(l-1)(a-1)$
YLA	$(y-1)(l-1)(a-1)$
B tényező	$b-1$
YB	$(y-1)(b-1)$
LB	$(l-1)(b-1)$
YLB	$(y-1)(l-1)(b-1)$
AB	$(a-1)(b-1)$
YAB	$(y-1)(a-1)(b-1)$
LAB	$(l-1)(a-1)(b-1)$
YLAB	$(y-1)(l-1)(a-1)(b-1)$
Hiba	$y(l(r-1))(ab-1)$

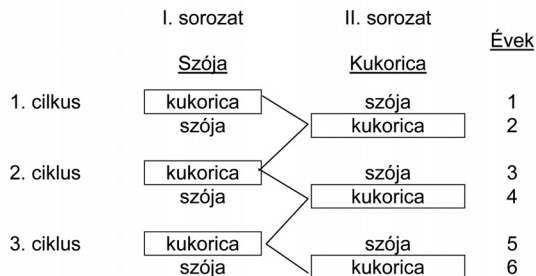
5.4. Tartamkísérletek

A tartamkísérletben a kezeléseket minden parcellán beállítjuk a kísérlet kezdetekor, jöllehet egyes vetésforgó kísérletekben a parcellák szakaszolva lehetnek, hogy lehetővé tegyék a rotáció minden növényének termesztését minden évben. A parcellán végzett mérések általában minden évben megtörténnek a termés és más növényi tulajdonságok esetében. A tartamkísérletekben a kumulatív hatásokat becsüljük, amelyek a direkt (a kezelés tárgyevi hatása) és a reziduális (korábbi évek utóhatása) hatások kombinációi (Cady, 1991). Ezek az időben ismételt mérési adatokat valamilyen statisztikai analízisnek vetjük alá a kezelések időbeni kumulatív hatásának vizsgálata céljából. Ennek eredményeként, az idő \times kezelés interakció gyakran az analízis fókuszába kerül, amely vetésforgó kísérletben kiterjeszhető a ciklusok és sorozatok vizsgálatára.

A tartamkísérletek elrendezése követhet bármilyen standard kísérleti elrendezést, beleértve a randomizált blokk és az osztott parcellás elrendezéseket (Yates, 1954; Patterson, 1964; Sváb, 1981). Vetésváltás kísérletek tervezését és elrendezését, valamint az ún. lépcsős kiterítés szerkezetét Sváb (1981) ismerteti. Lépcsős kiterítés esetén – technikai nehézségek miatt – nem javasolja az osztott parcellás elrendezést, bár elméletileg ez a helyes megoldás. Teljesen kielégítő, ha az újonnan indított kísérletet a már futó kísérlet mellé, hasonló körülmények közé helyezzük. A különböző megfigyelési szakaszokban újonnan indított tartamkísérleteket időben végzett kísérletsorozatként értékelhetjük.

Példa: Trágyázási kezelések összehasonlítása vetésforgó kísérletben (Cady és Mason, 1964)

A kukorica termésadatokat analizálták kukorica és szója két éves vetésforgó kísérletben. A kísérletet hat éven keresztül folytatták egy bevezető év után. A rotáció mindkét fázisát beállították minden évben, amely kukoricatermést eredményezett mind egyik évben. A trágyázási kezelések nitrogén, foszfor és kálium kombinációk dózisait és alkalmazási idejét jelentették, randomizált blokkelrendezésben és három ismétlésben. A kísérlet szerkezetét az 5.16. ábra szemlélteti. Az első évben kukorica volt termesztve a kísérlet I. sorozatában (szója volt termesztve egy évvel korábban, az előzetes évben). A parcelláknak mindkét sorozata 16 kezelést tartalmazott három ismétlésben, randomizált blokk elrendezésben. A sorozatok száma egyenlő a rotáció hosszával, példánkban, kettővel. Hat év során a rotáció háromszor ismétlődött, vagyis a rotációnak három ciklusa volt. A modellben az év komponens felbontható sorozatokra, ciklusokra és sorozat \times ciklus interakcióra. Feltételezzük, hogy a sorozatok és a ciklusok random tényezők.



5.16. ábra. Sematikus diagram a kísérleti struktúra illusztrálására

A rotációs kísérlet struktúráját beépítve, az alábbi modellt kapjuk:

$$Y_{ijkl} = M + S_i + R_{ij} + C_k + (SC)_{ik} + (RC)_{ijk} + T_l + (ST)_{il} + (RT)_{ijl} + (CT)_{kl} + (SCT)_{ikl} + (RCT)_{ijkl}$$

$$i = 1, 2$$

$$j = 1, 2, 3$$

$$k = 1, 2, 3$$

$$l = 1, 2, \dots, 16$$

ahol: M = főátlag, S_i = az i -edik sorozat additív hatása, R_{ij} = a j -edik ismétlés additív hatása az i -edik sorozatban, C_k = a k -adik ciklus additív hatása, $(SC)_{ik}$ = az i -edik sorozat és a k -adik ciklus kombinációjának additív interakció hatása, $(RC)_{ijk}$ = a j -edik ismétlésben az i -edik sorozatban a k -adik ciklussal alkotott additív interakció hatása, T_l = az l -edik kezelés additív hatása, $(ST)_{il}$ = i -edik sorozat és az l -edik kezelés additív interakció hatása, $(RT)_{ijl}$ = j -edik ismétlésnek az i -edik sorozatban és az l -edik kezelésnek az additív interakció hatása, $(CT)_{kl}$ = a k -adik ciklus és az l -edik kezelés additív interakció hatása, $(SCT)_{ikl}$ = az i -edik sorozat és a k -adik ciklus és az l -edik kezelés additív három faktoros interakció hatása, és $(RCT)_{ijkl}$ = a j -edik ismétlésnek az i -edik sorozatban és a k -adik ciklusnak és az l -edik kezelésnek az additív három faktoros interakciója.

A kísérlet részletes értékelése *Cady és Mason (1964)* közleménye tartalmazza. A varianciaanalízis vázlatát az alábbi táblázat mutatja:

Variáció forrása	d.f.
Sorozatok (S)	$s-1$
Ciklusok (C)	$c-1$
SC	$(s-1)(c-1)$
Ismétlés S(R/S)-en belül	$s(r-1)$
CR/S	$s(r-1)(c-1)$
Kezelés (T)	$t-1$
TS	$(t-1)(s-1)$
TR/S	$s(t-1)(r-1)$
TC	$(t-1)(c-1)$
TSC	$(t-1)(s-1)(c-1)$
TCR/S	$s(t-1)(c-1)(r-1)$
Összes	$scrt-1$

5.5. Üzemi (demonstrációs) kísérletek

A kísérleteket a gazdaság tábláin (üzemi kísérletek) osztályozhatjuk, mint technológia-fejlesztési és technológia-összehasonlítási kísérletek. A kisparcellás kísérletekben alkalmazott biometriai eljárások kismértékű módosítással felhasználhatók a technológia-fejlesztési kísérletekben (15.1. fejezet). A technológia-összehasonlítási kísérlet

k teszt faktoral úgy tekinthető, mint egy 2^k faktoriális kísérlet, amelyben minden faktor két szintje reprezentálja az előírt szintet az új technológiában és a termelői gyakorlat (technológia) szintjét. A kísérlet méretének csökkentésére általában a kezelések három különböző sorozatát tesztelik (15.2. fejezet). A termesztési rendszerek vizsgálatkor (FSR) a javított technológiának (kiegészítő kezelésekkel) tulajdonítható minimális terménynövekedés (LSI) szignifikanciájának tesztje a kísérleti adatok varianciaanalízisén alapul (15.3. fejezet).

6. Adatbevétel és adatfeltárás

A statisztika foglalkozik az *adatok gyűjtésével, elemzésével és értelmezésével* (interpretálás) a lehetséges legmegfelelőbb módon. Mindhárom aspektus egyformán fontos. Az adatok minőségének fontosságát nem lehet eléggé hangsúlyozni.

6.1. Adattípusok

A tudományos kutatásban legtöbbször az ún. **mérhető adatokkal** találkozunk. Elnevezésük utal rá, hogy az ilyen adatot általában valamilyen mérés eredményeképp kapjuk. Gondoljunk itt elsősorban a távolság-, idő-, tömeg- és anyagmennyiség adatokra, de szóba jönnek a származtatott mennyiségek is, mint a sebesség, gyorsulás, erő, koncentráció stb. Ezen adatok legfőbb jellemzője a folytonosság. Ez lényegében annyit jelent, hogy (bizonyos határok közt) minden érték előfordulhat, más szóval az adatok között akármilyen különbségek is lehetnek. Ennek következménye, hogy az adatok gyakorlatilag mind különbözőek, csak mérési eljárásunk érzékenységén múlik, hogy megtaláljuk-e ezeket a különbségeket.

A számadatok egy másik csoportját nem folytonos, hanem egymástól elszigetelt, **diszkrét számadatok** alkotják. Ezek sohasem mérés, hanem számlálás eredményei, tehát mindig pozitív egész számok. Van úgy, hogy egyet sem találunk valamiből, ezért a nullát is fel kell vennünk a lehetséges adatok közé.

A következő típus különleges helyet foglal el az adatok közt: nagyon ritka, hogy a vizsgálat (megfigyelés, felmérés, kísérlet) során ilyen adatokat rögzítsenünk, annál gyakoribb azonban, hogy a legtöbbször méréssel meghatározott adatokat ilyenekké minősítik át, és ennek megfelelően elemzik. Ezek az ún. **rangsorolható adatok**. Mindössze annyit kell tudnunk, melyik nagyobb, melyik kisebb a másiknál, de hogy mennyivel kisebbek vagy nagyobbak, az már lényegtelen, és legtöbbször meg sem állapítható. Kiértékelésük során úgy járunk el, hogy nagyság szerinti sorba rakva megszámozzuk azokat (1-es, 2-es, ...) és eredeti értékük helyett ezekkel a rangszámokkal dolgozunk a továbbiakban.

Az adatok következő típusa nem számokból, hanem különféle megállapításokból áll. Ilyen lehet valakinek a foglalkozása, a szeme színe, betegsége és számtalan más, egyáltalán nem számszerű és legtöbbször nem is számszerűsíthető adat. Gyakran nevezik az ilyen adatokat **megállapítható adatoknak**, hiszen nem mérés vagy számlálás, hanem egy – szóval vagy mondattal történő – megállapítás útján jutunk hozzájuk. Az ilyen megállapítás a jellemzett személy vagy tárgy valamilyen kategóriába sorolását jelenti, ezért kategóriás adatoknak is hívják. Ha pedig azt akarjuk hangsúlyozni, hogy ezek nem számadatok, kvalitatív adatoknak is hívhatjuk ezeket. Különösen gyakran fordul elő ezen adatok legegyszerűbb változata, amikor mindössze két érték lehetséges: egészséges vagy beteg, élő vagy meghalt, kezelték vagy nem kezelték. Ezeket bináris adatoknak is hívjuk.

Adatbevitel

- Készítsünk tartalék másolatot az adatokról és tartsuk biztonságos helyen.
- Gépi másolatoknak (hard copy) tartósnak és jól áttekinthetőnek kell lenni.

Az adatok gyűjtése költséges. A legtöbb esetben az adattárolás elektronikus formában történik.

Adatok

- A kvalitatív adatoknak nincs numerikus sorrendje.
- A kvantitatív adatok numerikus sorrendbe rendezhetők.
- Kvalitatív adatok: pl. állatok neme, 1: hím-, 2: nőnemű. Alma osztályozása szín szerint. 1, 2, 3 stb. jelenthetnek különböző színeket, vagy szavakat fejezhetnek ki, amelyek leírják a színeket, így zöld, sárga, piros.
- Kvantitatív adatok: melyek értékei numerikus sorrendbe helyezhetők. Születési súly, termésátlag stb. Matematikai műveletek végezhetők ilyen adatokkal, jöhetnek az eredmények nem mindig értelmezhetők.

Adattípusok

- Diszkrét adatoknak véges számú lehetséges értéke van az értékeknek egy tartományában.
- Folytonos adatoknak végtelen számú lehetséges értéke van.

A kvantitatív adatok lehetnek diszkrét vagy folytonos adatok, a kvalitatív adatok csak diszkrét adatok lehetnek.

Diszkrét adatok

A zéró feletti pozitív egész számok sorozata és egy anyajuh bárányainak száma példa a diszkrét adatokra.

A diszkrét adatoknak speciális formáját képezik a *bináris adatok*. Ezeknek csak két értéke lehet. Pl. egy változó, amely azt mutatja, hogy gyomok jelen vannak-e egy parcellán vagy hiányoznak, lehet 1-es értéke (jelen vannak a gyomok) vagy 2-es értéke (hiányoznak).

Folytonos adatok

Minden értéket felvehetnek az adott tartományon belül. Egy állat tömege, egy növény termése, peszticid mennyisége víz mintában.

6.2. A változók típusai mérési skála szerint

A *változó* szót a statisztikában arra használjuk, hogy jelezzen egy jellemzőt vagy tulajdonságot, melyet lehetséges mérni (Afifi et al., 2004). Amikor mérünk valamit, egy numerikus modellt készítünk arról a dologról, amit mérünk. Követünk bizonyos szabályt, hogy kijelöljünk egy számot a sajátos jellemző mindegyik szintjére, melyet mérünk.

Fizikai méretek, mint a magasság és tömeg, közvetlenül mérhetők, felhasználva fizikai műszereket. Másrészt, az olyan tulajdonságokat, mint a gondolkodó képes-

ség vagy egy személy depressziójának állapota, közvetetten kell mérni. Választhatunk egy sajátos intelligencia tesztet és meghatározzuk az „intelligencia” változó legyen egy pont eredmény, amelyet el kell érni ezen a teszten. Hasonlóan meghatározhatjuk a „depresszió” változót úgy, mint a pozitív válaszok számát a kérdések sorozatára. Jóllehet, amit mi akarunk mérni, az a depresszió foka, befejezzük az igen válaszok megszámlálásával bizonyos kérdésekre. Ezek a példák rámutatnak az alapvető különbségre a közvetlen fizikai mérések és az absztrakt változók között.

Bármely adott egyedén vagy dolgon számos különböző jellemzőt mérhetünk. Számos változóval foglalkozhatunk, mint amilyen egy bizonyos egyed kora, magassága, éves bevétele, neme, vallása és depresszió szintje. Hasonlóan mérhetjük a vállalatok jellemzőit, mint a különböző pénzügyi jellemzőket.

A változók Stevens-féle osztályozása

Ebben a rendszerben megkülönböztetünk nominális, ordinális, intervallum vagy arány változókat (Stevens, 1951). A változók típusait figyelembe kell venni a statisztikai próba kiválasztásánál.

Nominális változók

Nominális változóknál minden megfigyelés számos eltérő kategória (osztály) egyikéhez tartozik. A kategóriák (osztályok) nem szükségszerűen numerikusak, jóllehet számok használhatók, hogy reprezentálják azokat. Például a „nem” egy nominális változó. Egy egyed neme vagy hím vagy nő. Használhatunk bármely két szimbólumot, mint amilyen M és F, a két kategória reprezentálására. Az adatelemzésben számokat használunk szimbólumként, mivel sok számítógépes programot úgy terveztek, hogy csak numerikus szimbólumot kezeljen. Mivel a kategóriák rendezhetők bármely kívánatos sorrendbe, bármely adatsor használható, hogy reprezentálja azokat. Például, használhatunk 0 és 1 számokat a hím és nő kategóriák reprezentálására. Használhatunk 1 és 2 számokat, hogy elkerüljük a zéró üres helyet. Bármely más két szám is használható, amíg következetesen használjuk azokat. A 6.1. táblázat további példákat mutat be.

6.1. táblázat. A változók típusai a mérési skála szerint (Stevens, 1951)

Mérési skála	Alap empirikus művelet	Példák
Nominális	Meghatározza a kategóriák egyenlőségét	Vállalat nevek Fajok Vallások Futball játékosok száma
Ordinális	Meghatározza a nagyobb, mint vagy kisebb, mint (sorrend)	Ásványok keménysége Borok rangsora
Intervallum	Meghatározza a szintek közötti különbség egyenlőségét	Hőmérséklet Fahrenheit fokban Naptári napok
Arány	Meghatározza szintek arányának egyenlőségét	Magasság Súly Sűrűség

Ordinális változók

Kategóriákat használunk az *ordinális változóknál* is, azonban ismert sorrend is van. Például az ásványok és sziklák osztályozhatók a keménység szintje szerint (*Mohs* keménységi skála). A legkeményebb ásvány a gyémánt és a legpuhább a zsírkö. Bár mely tíz szám használható, hogy reprezentálja a kategóriákat, amíg nagyság szerint rendezve vannak. Például, 0–10 egész szám természetesen használható. Tehát az alapvető tapasztalati érvényesség, amely meghatározza az ordinális változót, az hogy vajon egyik megfigyelés nagyobb-e mint a másik.

Gyakran a vizsgálatot végzők osztályozzák az embereket vagy megkérlik őket, hogy osztályozzák magukat bizonyos folytonosság (kontinuum) mentén. Például, egy orvos osztályozhatja egy páciens betegség állapotát úgy, mint nem = 1, gyenge = 2, közepes = 3, és súlyos = 4. Érthetően, a növekvő számok jelölnek növekvő súlyosságot, azonban az nem biztos, hogy a különbség a nem beteg és a gyenge betegség között ugyanaz, mint a gyenge eset és a közepes eset között. Ezáltal, Steven osztályozási rendszerének megfelelően, ez egy ordinális (sorrendi) változó.

Intervallum változók

Egy *intervallum változó* olyan változó, amelyben a különbségek két egymást követő érték között mindig ugyanazok. Például, a „hőmérséklet” változót Fahrenheit fokban intervallum skálán mérjük, mivel a különbség 12° és 13° között ugyanaz, mint a különbség 13° és 14° között vagy a különbség bármely két egymást követő hőmérséklet között. Ezzel szemben, *Mohs* keménységi skálája nem felel meg ennek a feltételnek, mivel az intervallumok az egymást követő kategóriák (osztályok) között nem szükségszerűen ugyanazok. A skálának meg kell felelni annak az alapvető tapasztalati érvényességnek, hogy megőrzi az intervallumok egyenlőségét.

Arány változók

Az *arány változók* intervallum változók egy természetes ponttal, amely reprezentálja a mérés eredetét (kiindulását), azaz a természetes zéró pontot. Például, a magasság egy arány változó, mivel zéró magasság egy természetesen definiált pont a skálán. Megváltoztathatjuk a mérés mértékegységét (pl. cm-ről inch-re), azonban mégis megőrizzük a zéró pontot és a magasság bármely két értéke közötti arányt is. A leíró statisztikai mutatókat a mérési skálától függően a 62. táblázat mutatja.

6.2. táblázat. Leíró statisztikai mutatók a mérési skálától függően (Stevens, 1951)

Osztályozás	Grafikus ábrázolás	Eloszlás középpontjának mérése	Eloszlás variabilitásának mérése
Nominális	Oszlopdigram Kördiagram	Módusz	Binomiális vagy multinominális variancia
Ordinális	Hisztogram	Medián	Tartomány $P_{75} - P_{25}$
Intervallum	Hisztogram területek mérhető	Átlag = \bar{x}	Szórás = s
Arány	Hisztogram területek mérhető	Geometriai átlag Harmonikus átlag	Variációs koefficiens = s/\bar{x}

Egy adott osztályozási skálánál a felette levő mutatók mindegyike alkalmazható.

6.3. Adatellenőrzés és exploratív adatelemzés

Ha összegyűjtöttük a mintaszámnak megfelelő adatokat és még egyszer végiggondoltuk azokat a lépéseket, ahol valamilyen hibát (mintavételi, mérési, preconcepciós hibák) elkövethettünk, akkor nekifoghatunk adataink analízisének. A legtöbb esetben az adattárolás elektronikus formában történik. Készítsünk másolatot az adatokról és tartsuk biztonságos helyen.

- Az adatokat ellenőrizni kell, hogy helyesen lettek-e feljegyezve (rögzítve).
- Az adatokat jóvá kell hagyni, azaz ellenőrizzük hihetőségre és következeteségre.
- Extrém értékeket ellenőrizni kell, ezek lehetnek hibásak vagy kívülállók.

Célszerű exploratív (felderítő) adatanalízissel kezdeni, ami rámutat néhány nyilvánvaló összefüggésre és ötleteket adhat a tesztelésekhez vagy a további vizsgálatokhoz. Az exploratív adatanalízis egyszerű számításokat és ábrázolásokat foglal magában. A számítások közül legfontosabbak: (i) *maximum és minimum*, (ii) *helyzet-mutatók* (módusz, medián, kvartilis, számtani átlag), (iii) *szóródási mutatók* (terjedelem, interkvartilis terjedeleme, szórás, standard hiba, konfidencia határok és intervallumok, variációs koefficiens), (iv) *megfigyelések, illetve a hiányzó adatok száma*. A grafikus ábrázoláshoz leggyakrabban oszlop diagramot, hisztogramot, kvartilis ábrát és szórás diagramot használunk.

Konfirmatív adatanalízis. Miután megismerkedtünk adataink struktúrájával, rátérhetünk a konfirmációs (megerősítő) vagy döntéshozó adatanalízisre, azaz predikcióink tesztelésére. A biometriai módszerek közül a növénytermesztési kutatásokban leggyakrabban az alábbiakat használjuk: (i) *Paraméteres próbák* (átlagok és varianciák összehasonlítása). (ii) *Varianciaanalízis* (kettőnél több átlag összehasonlítása). (iii) *Regressióanalízis* (azonos megfigyelési egységen mért két vagy több kvantitatív változó közötti összefüggés jellegének vizsgálata) és *korrelációs számítás* (a változók közötti lineáris kapcsolat szorosságának vizsgálatára szolgáló módszer). (iv) *Többváltozós módszerek* (főkomponens analízis, faktoranalízis, többszörös varianciaanalízis, diszkriminancia analízis, klaszteranalízis és a korrespondencia-analízis) (Précsényi et al., 1995).

6.4. Számítógépek használata a kísérletek tervezésében és értékelésében

A legtöbb statisztikai fogalmat, amelyet ebben a könyvben használunk, a számítógépek felfedezése előtt kifejlesztettek. A statisztikai módszerek nagyon valószínűleg különböző utakon fejlődtek volna, ha a komputerek már 100 évvel ezelőtt rendelkezésre álltak volna, sokkal inkább mint 40 évvel ezelőtt. Azonban azok az alapelvek, amelyeken a kísérleti adatok analízise vagy az adatok összefüggés-vizsgálata alapul, nem lett volna különböző. Napjainkban közel minden statisztikai adatelemzés számítógéppel történik, amely magában foglalja egy statisztikai csomag használatát. Számos statisztikai csomag áll rendelkezésre manapság, hogy segítse a kísérletezőt az adatelemzésben.

A statisztikai szoftverek nemcsak számításokat hajtanak végre, hanem prezentálják az eredményeket grafikusán, hogy elősegítsék az interpretációt és az eredmények köz-

analízisét megelőzően, jöllehet, amikor ilyen derivált változókat számítunk ki számítógépes programmal, nagyon jó általános szabály megnézni az új változó értékeket, úgy hogy vizuálisan ellenőrizni lehessen bármely értéket, amelyek gyanúsak tűnik.

Az értékeknek ez az ellenőrzése minden oszlopon belül egy lehetőség, amely rendszerint rendelkezésre áll. Néhány program egyszerűen megadja a maximum és minimum értékeket egy oszlopban. Mások azonosítani tudják azokat az értékeket és a megfelelő egységeket, amelyek a specifikált tartományon kívül helyezkednek el. Az ellenőrzés megadásán kívül, egyszerű összesítések állhatnak rendelkezésre minden oszlopra. Ezek rendszerint magukba foglalják az átlagot, a szórást, a mediánt és más százalékos mutatókat. Grafikus output is rendelkezésre áll a legtöbb csomagban és rendszeren magában foglalja a gyakorisági eloszlásokat az egyes változókra és a szóródás diagramokat a változó párokra.

Programcsomagok nagy mennyiségű adatok analízisére magukban foglalják az adatkezelési módszerek sokkal szélesebb körét. A módszerek egyik szabálya az adat alsorozatok tabulálását adja, kiválasztva azokat a megfigyeléseket (sorokat), amelyek megfelelnek azoknak a feltételeknek, amelyeket egy vagy több változóval (oszlop) kifejezve definiáltunk. Például, ha a sorok megfelelnek különböző gazdaságoknak, akkor egy oszlop meghatározhatja a gazdaság típusát (növénytermesztés, állattenyésztés, vegyes, stb.), és előállít egy altáblázatot, amely magában foglalja az eredményeket csak a növénytermesztő gazdaságokra. Alternatívaként, egy altáblázat magában foglalhat csak bizonyos méretű gazdaságokat vagy a sajátos geográfiai területen elhelyezkedő gazdaságokat.

A táblázat manipulálási lehetőségen kívül, a csomagok a nagy adatmennyiségekre tipikusan magukba foglalják az adatok grafikus ábrázolásának tág lehetőségeit. Ezek a csomagok ugyancsak magukba foglalják egyszerű összegező statisztikai mutatók számításának lehetőségét és vagy eljárásokat a komplexebb analízishez vagy legalább a sor vagy derivált adat alsorozatok tárolásának módszerét, amelyeket komplexebb analízisnek vetünk alá, felhasználva más csomagokat. Folyamatosan új programcsomagok válnak elérhetővé és javított eljárásokat fejlesztenek ki. Ezáltal az alapvető statisztikai módszerek megfelelő megértése döntő fontosságú.

A tananyagban jöllehet hangsúlyozzuk a statisztikai analízisnek és az adatok interpretációjának alapelveit, és a módszereket egyszerű példákkal illusztráljuk, egyúttal demonstrálni kívánjuk, hogy a számítógépek elvégeznek minden szükséges számítást. A számítások alapvetően nem kapcsolódnak egy sajátos számítógépes csomaghoz és nem kísérünk meg összehasonlító becslést adni a rendelkezésre álló programcsomagokról. Ennek az elvnek nem mond ellent az a gyakorlat, hogy a kísérleti példák számítógépes lefuttatásához a Genstat programot használjuk.

A véletlen blokkrendezésű kísérlet példáján keresztül bemutatjuk a kísérletek számítógépes tervezését. A kísérletek számítógépes értékelését (varianciaanalízisét) a kapcsolódó fejezetekben ismertetjük.

6.5. Kísérletek számítógépes tervezése

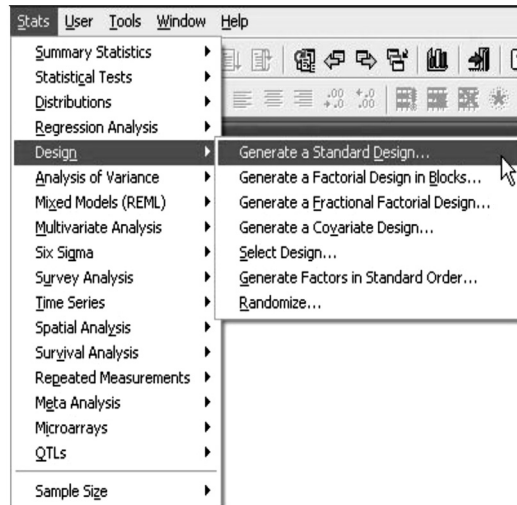
RCBD előállítás

Az elrendezés típusát kiválaszthatjuk, használva a Design lista dobozt (6.2. ábra). A kategóriák megegyeznek az Analysis of Variance menüvel. A teljesen randomizált blokk-elrendezés egyike azoknak az elrendezéseknek (design), melyeket a *Generate a Standard Design* menü kínál. A lenti ábrán az egytényezős, a kéttényezős és az általános kezelésstruktúra elrendezés (mindegyik véletlen blokk) látható (6.3. ábra). A jobb oldali részábrán a *Design* 1–5 kezelés faktor beállítását teszi lehetővé.

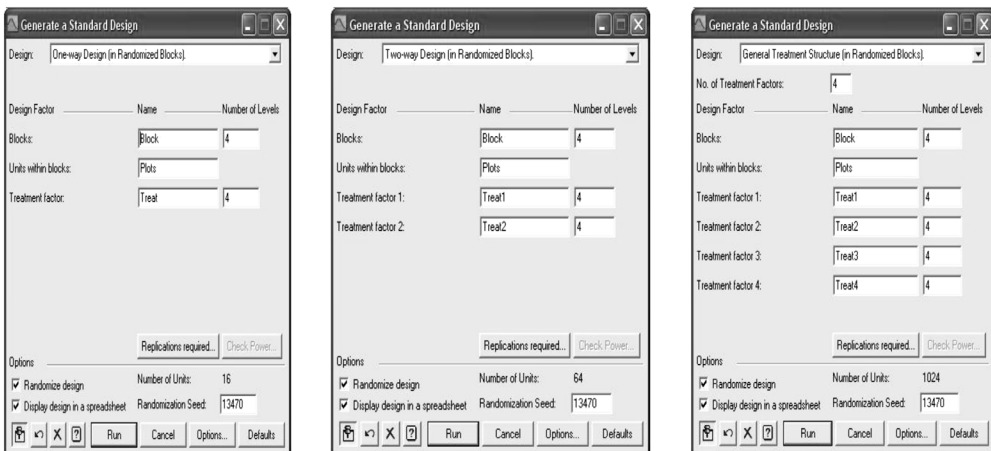
A GenStat a tervezési menük sorozatával rendelkezik, amely lehetővé teszi, hogy

1. kiválasszuk a kísérlet elrendezését
2. megállapítsuk a kezelések szintjeit
3. megnevezzük a faktorokat
4. megformázzunk egy nem randomizált tervet
5. randomizáljuk a kezeléseket
6. kinyomtassuk a tervet
7. meghatározzuk a blokk-struktúrát és a kezelés struktúrát (ha kiegyensúlyozott = balanced) az ANOVA-hoz
8. ellenőrizzük az elrendezést (ha kiegyensúlyozott), felhasználva az ANOVA-t
9. *Standard Design* menü elmentheti a faktorokat a file-ban. Utasításokat is használhatunk.

A *Generate a Standard Design* menü lehetővé teszi, hogy előállítsunk számos standard kísérleti elrendezést. Megkaphatjuk, ha rákattintunk a Stats-ra a főmenüben és kiválasztjuk a Design-t, ezután *Generate a Standard Design*-t.



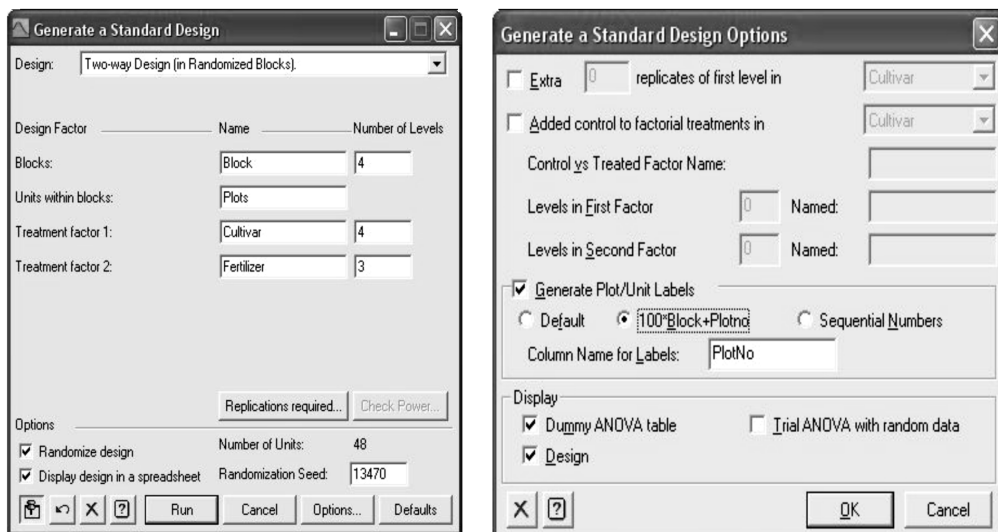
6.2. ábra. A kísérlet elrendezés típusának kiválasztása



6.3. ábra. Egytényezős, kéttényezős és általános kísérlettervezés

RCBD elrendezés 2 kezelés faktorial

Használjuk a fő menüt, hogy specifikáljuk a faktor neveket, meghatározzuk a szintek számát, a randomizációt, elmentsük az elrendezést egy munkalapon, és döntsünk az ismétlésről. A menü a baloldali ábrán előállít egy véletlen blokk elrendezést 4 blokkal, a kísérleti egység neve Plots, az 1. kezelés faktor neve Cultivar 4 szinttel, a 2. kezelés faktor neve Fertilizer 3 szinttel.



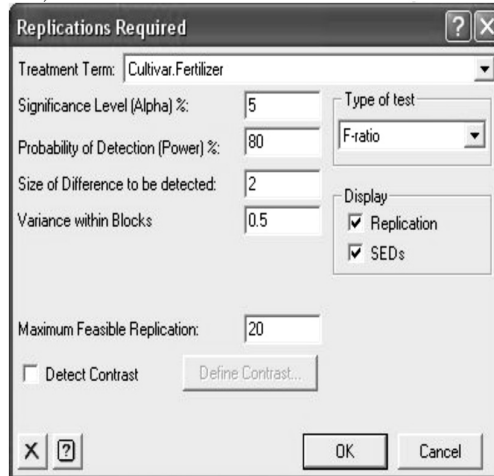
6.4. ábra. Kéttényezős RCBD kísérlet tervezése

A Generate a Standard Design Options menü (jobb oldali részábra) további szabályozást tesz lehetővé (6.4. ábra). Az Options menüt használjuk, hogy meghatározzunk extra ismétléseket/kontrollokat vagy előállítsunk kísérleti egység megnevezéseket. Az alábbi ábrán a Generate Plot/Unit Labels dobozt bejelöltük, hogy címkéket formáljunk a kísérleti egységek formálására. Gyakran megfelelőbb egyszerű numerikus kódokat használni. A címkék egész számok lesznek, 1, 2 stb. Ezeket a PlotNo néven mentjük el, specifikálva a Column Name for Labels ablakban. A menü kinyomtatja az elrendezést és mivel a Dummy ANOVA table is be van jelölve, előállít egy varianciaanalízis táblázatot is. Rákattintunk az OK-ra, hogy visszatérjünk a főmenübe.

Szükséges ismétlések az RCBD-ben

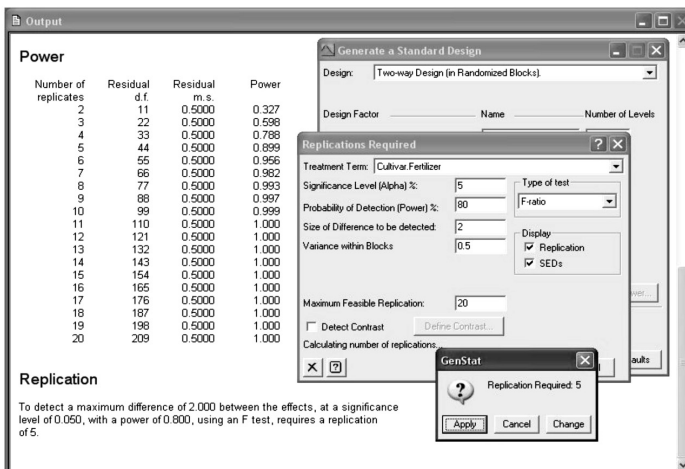
Visszatérve a Generate a Standard Design menühöz, rákattintunk a Replications Required gombra, megjelenít egy menüt, amely lehetővé teszi az ismétlés meghatározását (6.5. ábra). Az RCBD-ben az ismétlések száma a blokkok számát jelenti. A számításokhoz a GenStat-ban meg kell adni, hogy melyik kezelésre vonatkozik (itt az interakcióra). A Type of Test specifikálja, hogyan becsüljük azt, itt az F-arányt választottuk. Szükséges dönteni az alábbiakról: szignifikancia szint, amelyet használni tervezünk a különbség észlelésére; annak valószínűsége, amellyel kimutatni akarjuk a

szignifikanciát (teszt ereje); a legkisebb különbség nagysága, amelyet kimutatni akarunk (itt 2); mekkora blokkon belüli variancia várható (azaz maradék MS, itt 0.5). A maximális lehetséges ismétlés itt 20.



6.5. ábra. Az ismétlések tervezése az RCBD elrendezésben

Rákattintva az OK-ra, megjelenik az alábbi menü, amely jelzi az ismétlések szükséges számát (6.6. ábra). A Replication és SEDs dobozokat már bejelöltük, így a GenStat kinyomtat egy táblázatot, amely megadja az észlelés valószínűségét (Power), a maradék MS-t (közepes négyzetes eltérés) egészen 20 ismétlésig és egy közleményt a szükséges ismétlésekről.



6.6. ábra. Az ismétlések tervezésének outputja

A GenStat menü mutatja, hogy a szükséges ismétlések száma 5. Ahhoz, hogy a hatások között 2.0 maximális különbséget kimutassunk, 0.05 szignifikancia szinten, 80%-os észlelési valószínűséggel, felhasználva az F-próbát, 5 ismétlést tesz szükségessé. Rákattintunk az Apply gombra.

Állítsunk elő egy RCBD kísérletet

A Replications Required gomb rendelkezésre áll minden tervezéshez, ahol az ismétlés egyszerűen módosítható az egyik faktor szintjének megváltoztatásával. A Generate Standard Design menü jelenik meg újból, mint aktív ablak. Rákattintunk a Run-ra, hogy előállítsuk az elrendezést és az alábbi output-ot.

Treatment combinations on each unit of the design

Block	1	2	3	4	5
Plots					
1	1 3	4 3	3 3	4 2	4 3
2	3 3	4 1	2 2	1 3	2 1
3	1 2	4 2	1 3	2 2	4 1
4	4 1	1 2	4 2	4 1	4 2
5	2 3	2 3	1 1	1 1	3 1
6	4 2	1 3	1 2	2 1	2 3
7	2 1	3 2	2 1	4 3	3 3
8	2 2	3 1	4 1	1 2	1 3
9	4 3	2 1	2 3	3 1	1 2
10	1 1	3 3	3 2	3 3	2 2
11	3 2	2 2	4 3	3 2	1 1
12	3 1	1 1	3 1	2 3	3 2

Treatment factors are listed in the order: Cultivar, Fertilizer.

Analysis of variance

Source of variation	d.f.
Block stratum	4
Block.Plots stratum	
Cultivar	3
Fertilizer	2
Cultivar.Fertilizer	6
Residual	44
Total	59

Állítsunk elő egy RCBD-t egy számolótáblán

Mivel Display design in spreadsheet doboz be volt jelölve, a faktorok egy munkalapon jelennek meg a kísérlet elrendezésének megfelelően (6.7. ábra). Ezután használhatjuk a standard spreadsheet menüket, hogy meghatározzuk pl. a szinteket, illetve megnevezéseket.

The image shows a spreadsheet application window with the 'Factor' menu open. The menu options include: New, Column, Factor, Calculate, Delete, Insert, Select, Restrict/Filter, Sort..., Manipulate, Sheet, Book, Add, Export, Update, and Set as Active Sheet. The 'Factor' menu is expanded, showing options like 'Change Levels...', 'Reorder Levels...', 'Calculate...', 'Product/Combine...', 'Divide...', 'Recode...', 'Standardize Levels...', 'Convert to', 'Edit Levels and Labels...', 'Edit Levels...', 'Edit Labels...', and 'Reference Level...'. A mouse cursor is pointing at 'Edit Levels and Labels...'. To the right, a spreadsheet window titled 'Spreadsheet [Book:1]' displays a table with the following data:

Row	PlotNo	Blocks	Plots	Cultivar	Fertilizer
1	101	1	1	2	3
2	102	1	2	1	2
3	103	1	3	4	1
4	104	1	4	4	2
5	105	1	5	2	1
6	106	1	6	3	1
7	107	1	7	4	3
8	108	1	8	3	3
9	109	1	9	1	1
10	110	1	10	1	3
11	111	1	11	2	2
12	112	1	12	3	2
13	201	2	1	2	3
14	202	2	2	2	1
15	203	2	3	3	1
16	204	2	4	2	2
17	205	2	5	4	3
18	206	2	6	1	2

6.7. ábra. Az RCBD kísérlet számolótáblája

7. Leíró statisztika

A leíró statisztika hasznos az analízis mindegyik szakaszában, az adatok ellenőrzésétől és összegezésétől a kezdeti értelmező (exploratív) analízisig, az eredmények végső prezentálásán és a következtetéseken keresztül. A leíró statisztika használata magában foglalja az összegező statisztikák kiszámítását, mint amilyen az átlag, minimális és maximális értékek, szórás, megfigyelések száma és hiányzó értékek. Ezeket a statisztikai mutatókat arra használjuk, hogy leírjuk a populációk *helyzetét*, *szóródását* (diszperzió/variabilitás) és *alakját*, amelyből az adatok származnak. Egyetlen mutató sem írja le tökéletesen az adatok sorozatát.

Az adatokban levő információk betekintésére az alábbi mutatók hasznosak:

- *A megfigyelések száma*: minél nagyobb a minta mérete, annál jobban reprezentálja a populációt.
- *Minimum és maximum*: segít feltárni a kiugró értékeket és hibás adatokat.
- *A helyzet mérőszámai*: jelzést adnak az adatok helyzetéről.
- *A szóródás mérőszámai*: az adatok variabilitásáról adnak jelzést.
- *Hiányzó adatok száma*: az adatok teljességéről ad felvilágosítást.

7.1. Helyzet-mutatók

Számtani átlag

A számtani átlag a legfontosabb és leggyakrabban használt mutatója a centrális tendenciának (a helyzetnek). A populáció átlagát μ , a minta átlagát jelöli. Kiterjedten használják a statisztikai analízisben és felhasznál minden tényleges megfigyelést az adatsorban. Túlságosan befolyásolhatja azonban egy vagy több nagyon extrém érték. Éppen ezért nem mindig a legmegfelelőbb középérték a ferde eloszlású adatokra. Legmegfelelőbb a szimmetrikus eloszlású adatok helyzet mutatójának kiszámítására. Ha a minta megfigyelések szimmetrikusak voltak a számtani átlagértékre, akkor a medián értéke és a számtani átlag azonos.

Az átlag értékek kiszámítása nem megfelelő kvalitatív adatokra. Ha az adatok nyilvánvaló csoportokat tartalmaznak, az összes adat számtani átlag értéke kevésbé értékes, jöllehet megfelelő lehet mindegyik csoportra külön-külön.

A számtani átlag a megfigyelések összege osztva a megfigyelések számával. A következő képlettel számítjuk ki:

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{\sum x}{n}$$

ahol x_i az i -edik egyedi megfigyelés értéke; minta számtani átlaga, és n a megfigyelések (adatok) száma a mintában.

Medián

A mediánt gyakran használjuk olyan adatsorok centrális tendenciájának leírására. Olyan adatsorokban, melyek vagy kevés extrém magas vagy kevés extrém alacsony értékeket tartalmaznak, összehasonlítva az értékek többségével, a medián még egy jó jelzést adhat az adatok centrális tendenciájáról, míg az átlagok hajlamosak eltolódní a középsőtől az extrém értékek irányába.

Ezt a helyzet-mutatót diszkrét adatokra használjuk, de felhasználható folytonos adatokra is, különösen, ha ferde eloszlásúak. A medián érték csak a középén elhelyezkedő megfigyelési értékeket használja, ha az értékek sorrendbe vannak állítva, a tehát a medián értékét nem befolyásolja semmilyen extrém érték, legyen az magas vagy alacsony. *A medián az az érték, amelyiknél az összes előforduló érték fele kisebb, fele nagyobb.* Számos fontos adatelemzési módszer alapul az értékek rangsorba állításán és ezt követően a medián értékek összehasonlításán (pl. nem-paraméteres próbák).

A medián a megfigyelések középső értéke, amikor azok nagyság szerinti sorrendben vannak elrendezve. Definíció szerint a medián az az érték, amelynek egyenlő számú megfigyelése van bármelyik oldalán. Ez nyilvánvalóan a középső megfigyelési érték, *ha páratlan számú megfigyelés van.* Ha páros számú megfigyelés van, akkor a medián a két középső megfigyelés összegének a fele. Következőképpen számítjuk ki:

$$M = x_{(n+1)/2} \quad \text{ha } n \text{ páratlan szám}$$

$$M = \frac{x_{n/2} + x_{1+n/2}}{2} \quad \text{ha } n \text{ páros szám}$$

ahol x_n az n -edik minta megfigyelés értéke, M a medián, és n a megfigyelések száma a mintában.

Módusz

A módusz a helyzet mutatója, amelyre nincsenek hatással a kiugró (extrém) értékek. A módusz *a leggyakrabban előforduló megfigyelés értéke* (az az érték, melynél a gyakorisági eloszlás eléri a csúcst). Ezáltal, ez a helyzet-mutató leghasznosabb megszámtolt adatokra. Bizonyos esetekben egynél több módusz lehetséges (több módusú). Ez igen gyakran annak a jele, hogy a vizsgált sokaság (populáció) elemei valamilyen lényeges szempontból különböznek egymástól, s annak több részre bontása kívánatos az elemzés során.

Folytonos adatoknál valószínűtlen, hogy egy módusz érték előfordul, mivel nem valószínű, hogy két megfigyelés pontosan ugyanaz legyen. Ennek ellenére, a módusz felhasználható folytonos adatoknál. Osztály-intervallumokat választunk és az az osztály, amely az adatok legnagyobb gyakoriságát tartalmazza, az a módusz osztály. Ez egy hasznos jelzést ad az adatok eloszlásának általános alakjáról, azonban önmagában nem különösen hasznos összefoglaló mutató, mivel az osztály-intervallumok kiválasztása megfelelően szubjektív módon történik. A módusznak azonban nincs szerepe a komplexebb kvantitatív adatelemzési módszerekben, melyeket a kísérleti adatok elemzésére használunk.

Kvartilisek

A kvartilisek helyzet mutatók, amelyeket diszkrét vagy folytonos adatokhoz használhatunk. Az *első kvartilis* vagy *alsó kvartilis* az az érték, amely alatt a megfigyelések egynegyede elhelyezkedik. A *harmadik kvartilis* vagy *felső kvartilis* az az érték, amely felett a megfigyelések egy-negyede található. Ebből következik, hogy az adatok 50%-a az alsó és a felső kvartilisek között található. A medián érték a második kvartilis, az az érték, amely felett és alatt a megfigyelések fele helyezkedik el. Ezek a kvartilisek mellőzik a tényleges megfigyelési értékek többségét és nem befolyásolják extra értékek.

A kvartilisek kiszámításához az adatokat először növekvő érték szerinti sorrendbe rendezzük.

Az *első kvartilis* a megfigyelés értéke a $(N+1)/4$. pozícióban és a *harmadik kvartilis* a megfigyelés értéke a $3(N+1)/4$. pozícióban, ahol N a megfigyelések száma a mintában. Ha ezek nem eredményeznek egész számokat, interpoláció használható a tényleges kvartilis becslésére.

Az alábbi példa adatsorok forrása: Collins, C. A., Seeney, F. M.: 1999. Statistical experiment design and interpretation. John Wiley & Sons, New York.

Példa

Figyeljük meg a következő három adatsort:

1 adatsor: 3, 5, 6, 8, 1, 6, 6, 4, 1, 2, 5, 2, 6

A megfigyelések száma: 13

Számtani átlag: megfigyelések összege/13 = 55/13 = 4.23

A tényleges értékek előfordulásának gyakorisága a következő:

Érték	1	2	3	4	5	6	7	8
Gyakoriság	2	2	1	1	2	4	0	1

Ezáltal a módusz az 6.

Adatok sorrendje: 1, 1, 2, 2, 3, 4, 5, 5, 6, 6, 6, 6, 8.

Megfigyelések száma: $n = 13$

Medián érték = $(13+1)/2$. megfigyelés értéke = hetedik megfigyelés = 5.

2. adatsor. 81.5, 69.9, 58.9, 74.0, 89.9, 54.8, 72.9, 73.9, 64.7, 64.4, 79.3, 83.1, 89.1, 78.3, 74.4.

A megfigyelések száma: $n = 15$

Számtani átlag: megfigyelések összege/15 = 1109.1/15 = 73.9

A számtani átlag és a medián érték hasonló erre az adatsorra, jelezve a szimmetrikus eloszlást.

A megfigyelések gyakorisága a következő módusz osztályokban:

Érték	50.00–	55.00–	60.00–	65.00–	70.00–	75.00–	80.00–	85.00–
		54.95	59.95	64.95	69.95	74.95	79.95	84.95
Gyakoriság	1	1	2	1	4	2	2	2

Tehát, a módusz osztály 70.00–74.95

Adatok sorrendje: 54.8, 58.9, 64.4, 64.7, 69.9., 72.9, 73.9, 74.0, 74.4, 78.3, 79.3, 81.5, 83.1, 89.1, 89.9.

Megfigyelések száma: $n = 15$

Medián érték = $(15+1)/2$. megfigyelés értéke = 8. megfigyelés értéke = 74.0.

Az első kvartilis: $a(15+1)/4$. megfigyelés értéke = a 4. megfigyelés értéke = 64.7.

A harmadik kvartilis: $3(15+1)/4$. megfigyelés értéke = 12. megfigyelés értéke = 81.5.

3. adatsor. 11, 20, 18, 20, 90, 22, 28, 32, 46, 45, 27, 31.

A megfigyelések száma: $n = 12$

Átlag érték = megfigyelések összege/12 = $390/12 = 32.5$.

Ez az érték teljesen különböző a medián értéktől. A számtani átlagot befolyásolta a 90-es kiugró érték. Mivel az adatok ferde eloszlásúak, a medián a legjobb a helyzet mérésére.

Adatok sorrendje: 11, 18, 20, 20, 22, 27, 28, 31, 32, 45, 46, 90

Megfigyelések száma: $n = 12$

Medián érték = $(12/2)$ és $(12/2+1)$ megfigyelés összegének a fele = a 6. és 7. megfigyelés összegének a fele = $(27+28)/2=27.5$.

Az érték a minta többségét reprezentálja. Nem befolyásolja az extrém 90-es érték. Mivel az adatok ferde eloszlásúak a medián a helyzet legjobb mutatója.

A mediánt ugyanannyi tizedesértékkel közzöljük, mint az eredeti adatokat.

Az első kvartilis: $a(12+1)/4$. megfigyelés értéke = a 3.25. megfigyelés értéke

Ezáltal az érték egy-negyeddel fekszik a harmadik és negyedik rangsorolt megfigyelés között. Az első kvartilis a 3. megfigyelés értéke + $0.25 \cdot (4. \text{ megfigyelés} - 3. \text{ megfigyelés}) = 20 + 0.25 \cdot (20-20)=20$.

A harmadik kvartilis: $3(12+1)/4$. megfigyelés értéke = 9.75. megfigyelés értéke. Ezáltal az érték háromnegyeddel a kilencedik és tizedik rangsorolt megfigyelés között fekszik.

A harmadik kvartilis a 9. megfigyelés értéke + $0.75 \cdot (10 \text{ megfigyelés} - 9 \text{ megfigyelés}) = 32 + 0.75 \cdot (45-32)=42$

7.2. Szóródás mutatók

A korábban bemutatott helyzet-mutatók nem adnak semmilyen információt arról, hogy a helyzet mérése mennyire jól képviseli a megfigyeléseket. A szóródás mutatóit arra használjuk, hogy leírjuk a megfigyelések „alakját”. Ez különösen fontos, amikor megkíséreljük megbecsülni, hogy a megfigyelések a különböző változókból ugyanabból a populációból származnak-e, amikor ismernünk kell az egyedi megfigyelések változásának mértékét a mintán belül. A megfigyelések szorosan csoportosulhatnak együtt, szóródhatnak az értékek széles tartományában vagy klasztereket formálhatnak különböző pontokban. A szóródás mértékegysége azonos az alapadatokéval.

Terjedelem

Ez a szóródás legegyszerűbb mérése és rendszerint diszkrét adatokhoz kapcsolódik. Ha a helyzet megfelelő mérőszáma a medián, akkor a terjedelem közölhető a szóródás mutatójaként. A szóródásnak ez a mérése felhasználható arra, hogy összehasonlítsuk a hasonló méretű adatsorozatok variabilitását, azonban minthogy a terjedelem tendenciájában nő a minta méretének növekedésével, nem közvetlenül alkalmazható különböző méretű mintákra. Főként arra használjuk, hogy azonosítsuk a hibákat és a kívülálló értékeket.

A megfigyelések egy sorozatának terjedelme a sorozatban megtalálható *legnagyobb és legkisebb megfigyelési érték közötti különbség*. Egy extrém érték nagymértékben megnövelheti a tartományt, és minthogy csak a maximum és minimum értékeket használja fel, csak az adatsorozat általános szóródásáról ad képet, tehát tendenciájában nem informatív nagy adatsorozatokra. Például, két adatsorozat terjedelme ugyanaz lehet, azonban az egyikben a megfigyelések szorosan csoportosulhatnak együtt egy extrém értékkel, és a másik sorozatban mindegyik megfigyelés egyenlő távolságra helyezkedik el.

Interkvartilis terjedelem

A szóródásnak ez a mutatója ugyancsak kapcsolatban van a mediánnal, a helyzet mutatójával, mivel ez is a rangsorban levő mutatókat használja. Hasznos mutatója a szóródásnak, ha a minta értékek felső és alsó határai (limitjei) korlátlanok vagy nem meghatározottak.

Az interkvartilis tartományra kevésbé vannak hatással az extrém értékek, mint a tartományra. Következésképpen, hasznosabb mérőszáma a variabilitásnak, amikor a minta extrém értékeket tartalmaz vagy aszimmetrikus. Az interkvantilis tartomány a felső és alsó kvartilisek közötti különbség. Minthogy a kvartilisek a sorrendbe tett adatsorozatot négy egyforma részre osztja fel, az interkvantilis tartomány a megfigyelések középső 50%-át tartalmazza.

Példa

Figyeljük meg az alábbi két adatsorozatot:

2. *adatsor* (sorrendbe rendezve): 54.8, 58.9, 64.4, 64.7, 69.9, 72.9, 73.9, 74.0, 74.4, 78.3, 79.3, 81.5, 83.1, 89.1, 89.9.

Minimális érték = 54.8

Maximális érték = 89.9

Terjedelem = $89.9 - 54.8 = 35.1$

Az első kvartilis: 64.7

A harmadik kvartilis: 81.5

Interkvartilis tartomány: $81.5 - 64.7 = 16.8$

3. *adatsor* (sorrendbe rendezve): 11, 18, 20, 20, 22, 27, 28, 31, 32, 45, 46, 90

Minimális érték = 11

Maximális érték = 90

Terjedelem = $90 - 11 = 79$

Ezt a terjedelmet befolyásolta a 90-es extrém érték; e nélkül a terjedelem 35 lenne. Abból a célból, hogy pontosan leírjuk egy adatsor eloszlását, amely korlátolt számú extrém adatot tartalmaz, gyakran az *interkvartilis terjedelmet* számítjuk ki.

Első kvartilis: 20

Harmadik kvartilis: 42

Interkvartilis tartomány: $42 - 20 = 22$

Az interkvartilis tartományra kevésbé vannak hatással az extrém értékek, mint a tartományra.

Meg kell jegyezni, hogy a kvartilisek kiszámításának több módszere van, és ezek kismértékben eltérő eredményt adnak.

Eltérésnégyzet-összeg (SS)

Egy tökéletesebb megközelítés az adatsor szóródásának leírására az, amelyik számba veszi az adatok gyakorisági eloszlását és meghatározza a különbséget minden egyes érték és az átlagos érték között. Ezeknek a különbségeknek az összegezése ezután megadja az adatok variabilitásának relatív mérőszámát. Az összegezés előtt azonban az átlagtól való eltéréseket négyzetre emeljük, abból a célból, hogy eltávolítsuk a negatív értékeket (egyébként az átlagtól való eltérések összege zéró lenne!).

A négyzetes eltérések összegezése az *eltérésnégyzet-összeg* (SS) statisztikai mutatót eredményezi. Képlete:

$$SS = \sum_{i=1}^n (x_i - \bar{x})^2$$

ahol n az adatok számát jelöli.

Példa. Az átlagtól való eltérések négyzetének hatása a variáció mérésére. Két mintának (A és B) azonos az átlaga és hasonló az eltérés összege, $\sum (x - \bar{x})$, azonban az „A” mintának kétszer akkora az eltérésnégyzet összege, $\sum (x - \bar{x})^2$, mint a B mintának, jelezve, hogy nagyobb a variabilitása.

A minta ($\bar{x} = 10$)			B minta ($\bar{x} = 10$)		
x	$[x - \bar{x}]$	$[x - \bar{x}]^2$	x	$[x - \bar{x}]$	$[x - \bar{x}]^2$
6	4	16	10	0	0
10	0	0	12	2	4
14	4	16	8	2	4
9	1	1	12	2	4
11	1	1	8	2	4
\bar{x}	10		10		
$\Sigma[x - \bar{x}]$	10		8		
$\Sigma[x - \bar{x}]^2 (= SS)$		34			16

Variancia (σ^2 , s^2)

Jóllehet az eltérésnégyzet-összeg ad egy mérőszámot az adatsorozat összes varianciájáról, nagysága nő az adatok számának növekedésével. Ez nem kedvező, amikor különböző méretű adatsorok varianciáját hasonlítjuk össze. Ezáltal, abból a célból, hogy az átlagos variáció megfelelő mérőszámát kapjuk, az eltérésnégyzet-összeget elosztjuk a minta vagy a populáció méretével. Ez egy statisztikai mutatót eredményez, melyet hívhatnánk „átlagos eltérésnégyzet-összegnek”, azonban ténylegesen *varianciának* hívjuk.

Képlete:

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}$$

Tehát a számláló az eltérés-négyzet, a nevező pedig a szabadságfok.

Az adatokban kifejezésre jutó hatásokat (pl. kezelések) a variancia elemzésével tanulmányozzuk és értékeljük. Rendszerint úgy, hogy a variancia felbontásának segítségével különválasztjuk ezeket a hatásokat.

Szórás

Ez a variabilitás mutatója, amelyet legtöbbször ismernek és leggyakrabban használnak. Meg kell jegyezni, hogy a variancia a szórás négyzete és ezáltal nem ugyanaz a mértékegysége, mint az alapadatoké. A variancia ezáltal sohasem használható leíró statisztikai mutatóként.

Ahogy a megfigyelések száma nő, a variabilitás mérőszáma csökkenni fog és a populáció átlag (minta átlag) becslése nagyobb precizitással fog történni. A megfigyelések számának duplájára növelése nem csökkenti felére a szórást. Meg kell jegyezni, hogy a szórás csökkentése növeli a populáció átlag becslésének precizitását, azonban az nem szükségszerűen következik, hogy a becslés olyan pontos, mint ahogy a pontosság vonatkozik az átlagos értékre. Egy kis szórás összehasonlítva az átlagos érték nagyságával, magában foglalja, hogy a minta megfigyelések szorosan csoportosulnak átlagértékeik körül. Egy nagy szórás összehasonlítva az átlagértékkel, magában foglalja, hogy a megfigyelések nagymértékben szóródnak átlagértékük körül.

A megfigyelések sorozatának szórása a megfigyelések számtani átlagától vett eltérések négyzetes átlagának négyzetgyöke. A minta szórását a következőképpen számítjuk ki:

$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}}$$

ahol s a minta szórása; x_i az i -edik egyedi minta megfigyelés; \bar{x} a minta átlag; és az $(n-1)$ -et a számlálóban álló kifejezés szabadságfokának nevezik. A számlálóban az eltérésnégyzet összeg található.

Variációs koefficiens

A szóródás eddig tárgyalt minden mutatójának ugyanaz volt a mértékegysége, mint az egyedi megfigyeléseknek. Bizonyos célból hasznosabb mérni a szóródást relatív mértékegységben. A *variációs koefficiens* a minta variabilitás mérőszáma a helyzetéhez viszonyítva. A szórás és a minta átlag abszolút értékének hányadosa. Kifejezhetjük százalékban is. Ezt a mutatót arra használjuk, hogy közvetlenül összehasonlítsuk a különböző méretű minták variabilitását.

A variációs koefficiens kicsi értéke magában foglalja, hogy a minta megfigyelések szorosan csoportosulnak középértékük körül. Egy nagy variációs koefficiens kifejezi, hogy a megfigyelések tág tartományban szóródnak középértékük körül. A variációs koefficiens egyik előnye, hogy független a mértékegységektől, amelyben a változott mérték, feltéve, hogy a mérési skálák zérónál kezdődnek. A variációs koefficiens lehetővé teszi adatsorok variációjának közvetlen összehasonlítását, amelyek eltérő nagyságrendű értékeket tartalmaznak és ezáltal nagyon különböző az átlaguk.

Kiszámítása a következő formulával történik:

$$CV\% = \frac{s}{x} 100$$

ahol: CV% a variációs koefficiens százalékban; s a minta szórása és a minta átlaga.

Példa

Figyeljük meg a következő adatsort (2. *adatsor*):

81.5, 69.9, 58.9, 74.0, 89.9, 54.8, 72.9, 73.9, 64.7, 64.4, 79.3, 83.1, 89.1, 78.3, 74.4.

A megfigyelések száma, $n=15$

$$\sum_{i=1}^{15} x_i^2 = (81.5^2 + 69.9^2 + 58.9^2 + \dots + 83.1^2 + 89.1^2 + 78.3^2 + 74.4^2) = 83446.75$$

$$\sum_{i=1}^{15} x_i = (81.5 + 69.9 + 58.9 + 74.0 + \dots + 79.3 + 83.1 + 89.1 + 78.3 + 74.4) = 1109.1$$

$$\text{Minta szórása} = \sqrt{\frac{83446.75 - \frac{1109.1^2}{15}}{15 - 1}} = 10.21$$

A minta átlagos értéke: 73.9

A minta szórása: 10.21

CV% = $(10.21/73.9)100 = 13.82$.

7.3. A minta megbízhatóságának mérése: az átlag (középérték) hibája

Amikor egyetlen minta átlagára hivatkozunk, a következtetés az, hogy a minta átlag reprezentálja az egész populáció átlagát, amelyből a minta származik. Ezáltal fontos, hogy képesek legyünk jelezni, mennyire jól reprezentálja a minta a populációt. Ezt a statisztikai mutatót az átlag (középérték) hibájának (szórásának) hívjuk és jelzi a

minta átlag megbízhatóságát, a populáció átlag becsléseként. Mihelyt a minta szórását (s) meghatároztuk, a középérték hibáját (SE) úgy számítjuk ki, hogy elosztjuk a megfigyelések számának négyzetgyökével.

Minden megfigyelt adatot felhasználnak és folytonos adatokat használnak. Nagyon befolyásolják az extrém értékek és a szórás csökken a minta méretének növekedésével egy adott populációban. Ha a hibaszórást minden további magyarázat nélkül használjuk, akkor ez azt jelenti, hogy az átlag szórásáról van szó, bármely más szórás további kvalifikációt indokol. Gyakorlatilag a leggyakrabban használt szórások az átlag szórása (standard error of mean) és két átlag közötti különbség szórása (standard error of the difference between two means).

Egynél több random minta vehető egy populációból, amely a populáció átlagértékének egynél több becsléséhez vezet. Mindegyik mintának megvan a maga átlagértéke (a populáció átlag becslése) és a szórása. A különböző minta átlagok kialakítják a saját eloszlásukat, amely egy normális eloszlás, függetlenül az eredeti populáció eloszlástól. Az átlagértékeknek ez az eloszlása természetesen rendelkezik saját átlagértékkel és szórással. Az átlagértékek eloszlásának szórását egy adott méretű mintára vonatkozóan, hívjuk az *átlag (középérték) szórásának*.

A fenti elveket alkalmazzuk akkor is, ha egynél több populációból veszünk mintát. Ebben az esetben az átlagértékeknek egynél több eloszlása van, egy mindegyik populációra. Összehasonlítások végezhetők két átlagérték között, annak tesztelésére, hogy populációjuk ugyanaz vagy különböző. Más szavakkal, a két átlagérték közötti különbség egyenlő-e zéróval? Minthogy a két populáció mindegyike rendelkezik az átlagok eloszlásával, a két átlag közötti különbségnek is lesz egy eloszlása. A különbségek eloszlásának szórását egy adott mintaméretre, hívjuk a *különbség szórásának*.

Egy minta átlagértékének (középértékének) szórását (a középérték hibáját, SE) a következőképpen számítjuk ki:

$$s_{\bar{x}} = \frac{s}{\sqrt{n}}$$

ahol:

$s_{\bar{x}}$: az átlag szórása

s : a minta szórása

n : a megfigyelések száma a mintában.

Minthogy a minta mérete, n az egyenlet nevezője, eltekintve a négyzetgyöktől, az SE-nek automatikusan nőni kell, ahogy a minta mérete csökken. Ez egy nagyon fontos dolgot jelent, vagyis minél kisebb a minta, annál kevésbé megbízható. Mivel fontos, hogy jelzést adjunk a minta megbízhatóságáról, az átlag hibáját (SE) egy *alapvető leíró statisztikai mutató*nak kell tekinteni, melyet *mindig közölni kell a minta átlag prezentálásakor*. Figyeljük meg, hogy az SE-t általában mint plusz/mínusz értéket fejezzük ki az átlag körül, például, ha egy minta átlag 1.5 és az átlag hibája 0.25, akkor ezt egyszerűen úgy fejezhetjük ki, mint: 1.50 ± 0.25 .

Két minta átlagérték közötti különbség szórását (SED) az alábbi formulával számítjuk ki:

$$s_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}$$

ahol:

$s_{\bar{x}_1 - \bar{x}_2}$: a két minta átlagérték közötti különbség szórása,

s_1 : az első minta szórása,

n_1 : a megfigyelések száma az első mintában,

s_2 : a második minta szórása,

n_2 : a megfigyelések száma a második mintában.

Példa

Tekintsük meg az alábbi két adatsort:

2. adatsor: 81.5, 69.9, 58.9, 74.0, 89.9, 54.8, 72.9, 73.9, 64.7, 64.4, 79.3, 83.1, 89.1, 78.3, 74.4.

A megfigyelések száma, $n=15$

A minta szórása = 10.21

4. adatsor: 14.8, 15.6, 15.2, 14.8, 13.5, 14.2, 15.3, 25.0, 23.3, 26.2, 27.5, 24.9, 24.6, 25.0, 26.1.

A megfigyelések száma: $n = 15$

A minta szórása: 5.54

A minta átlag szórása a 2. adatsorra $= \frac{10.21}{\sqrt{15}} = 2.64$

A minta átlag szórása a 4. adatsorra $= \frac{5.54}{\sqrt{15}} = 1.43$

A két minta átlagérték közötti különbség szórása =

$$= \sqrt{\frac{(15-1)10.21^2 + (15-1)5.54^2}{15+15-2} \left(\frac{1}{15} + \frac{1}{15} \right)} = 3.00$$

Példa. Az adatok szóródását kifejező statisztikai mutatók kiszámítása (Ireland, 2010).

30–30 tenyészedenyben nevelt paradicsom növény növekedési rátája (szárztömeg g^{-1} nap $^{-1}$), melyeket egy új műtrágyával (Extragrow) és a standard műtrágyával (kontroll) kezeltek

Növekedési ráta (szárztömeg g^{-1} nap $^{-1}$)					
Standard műtrágya kezelés (kontroll)			„Extragrow” műtrágya kezelés		
0.8	1.3	0.9	1.9	1.7	2.0
1.1	1.4	0.4	2.3	2.2	1.5
0.7	1.8	0.8	1.0	0.8	1.6
0.8	0.9	1.3	1.0	1.1	0.7
1.2	1.1	1.0	1.8	1.2	0.7

A táblázat folytatása

Növekedési ráta (száraztömeg g ⁻¹ nap ⁻¹)					
Standard műtrágya kezelés (kontroll)			„Extragrow” műtrágya kezelés		
1.7	1.6	1.4	1.1	0.9	0.9
1.3	1.0	0.4	1.8	1.5	1.3
1.0	1.2	1.3	2.2	0.7	1.0
1.5	0.7	1.2	1.2	0.9	0.8
0.5	0.5	1.3	1.7	1.1	1.0

Az adatok egyszerű megfigyelése nem ad jelzést a műtrágya hatékonyságáról. Néhány leíró statisztikai mutató szükséges, hogy összegezzük az adatokat és lehetővé tegyük az adatok összehasonlítását.

	Kontroll	„Extragrow” műtrágya kezelés
Megfigyelések száma(n) =	30	30
Összes (Σx) =	32.10	39.60
Átlag (\bar{x}) =	1.07	1.32
Terjedelem =	1.40	1.60

Eltérésnégyzet-összeg kiszámítása (SS)

Σx^2 =	38.39	59.28
$(\Sigma x)^2$ =	1030.41	1568.16
$SS = \sum x^2 - \frac{(\sum x)^2}{n}$ =	4.04	7.01

Minta variancia (s^2) kiszámítása

$$s^2 = \frac{SS}{n-1} = 0.14 \quad 0.24$$

Figyeljük meg, hogy az adatok mintából származnak és a számláló a variancia kiszámításához $n - 1$, azaz a minta szabadságfoka inkább, mint n .

Szórás (s) kiszámítása

$$s = \sqrt{s^2} = 0.37 \quad 0.49$$

A műtrágyával kezelt mintának nagyobb a variabilitása, mint a kontrollnak.

Az átlag hibájának (szórásának) (SE) kiszámítása

$$SE = \frac{s}{\sqrt{n}} = \pm 0.07 \quad \pm 0.09$$

Figyeljük meg, hogy a fenti számításban, a várható gyakorlattal összhangban, a végső megoldásokat eggyel több tizedes helyre kerekítettük, mint az eredeti adatok.

Konfidencia-határok és -intervallumok

A különböző minta statisztikák, mint az átlagos értékek és szórások a populáció paraméterek (populáció átlag és szórás) becslései. E paraméterek valódi populáció értékeit nem ismerjük, kivéve, ha az adatokat az egész populációból begyűjtöttük. Lehetséges azonban becsülni a számított minta statisztikák megbízhatóságát oly módon, hogy a populáció paramétereit a statisztikai mutatókhoz illesztett konfidencia határokkal becsüljük.

A *konfidencia határok* definiálják egy intervallum felső és alsó határát, amely adott valószínűséggel magában foglalja a valódi populáció átlagértéket. Például, ha az alsó és felső 95%-os konfidencia határokat L_l és L_u reprezentálja, akkor 0.95 annak valószínűsége, hogy a valódi populáció átlag L_l és L_u között fog elhelyezkedni. A *konfidencia intervallum* a távolság a felső és alsó konfidencia határok között. Ezáltal, egy 95%-os konfidencia intervallum kifejezi, hogy 95% a valószínűsége annak, hogy az intervallum magában foglalja a valódi populáció átlagértéket.

Különböző fokú bizonyosság alkalmazható a konfidencia határookra és intervallumokra. Minél nagyobb fokú a bizonyosság, annál szélesebbé válik a konfidencia intervallum. Tehát, egy adott adatsorra és mintaméretre a konfidencia intervallum (amelyben az átlagérték elhelyezkedik) szélesebb lesz, mondjuk 99%-os bizonyosságnál, mint 95%-os bizonyosságnál. Ez nagyobb biztonságot ad arra, hogy a populáció átlag a konfidencia határok között van, azonban kevesebb bizonyosság van az átlag tényleges valódi értékéről, mivel a határok távolabb vannak egymástól.

A konfidencia intervallum szélessége csökkenthető az átlag szórásának csökkentésével. Ezt a tervezési stádiumban kell figyelembe venni. A konfidencia határok normális eloszlásra szimmetrikusak a minta átlag körül. Aszimptotikusak lehetnek azonban a minta átlagra más eloszlásoknál. Óvatosnak kell lenni, ha az adatokat transzformáltuk vagy ha ismert, hogy az eloszlás aszimmetrikus.

A formula 100(1- α)% konfidencia határ kiszámítására a következő:

$$\text{alsó } 100 (1-\alpha)\% \text{ konfidencia határ, } L_l = \bar{x} - t_{n-1, \alpha/2} S_{\bar{x}}$$

$$\text{felső } 100 (1-\alpha)\% \text{ konfidencia határ, } L_u = \bar{x} + t_{n-1, \alpha/2} S_{\bar{x}}$$

ahol: L_l : az alsó konfidencia határ; L_u : a felső konfidencia határ; \bar{x} : minta átlag; $t_{n-1, \alpha/2}$: a t-mutató n-1 szabadságfokkal; α : a valószínűségi szint, 0.05 95%-os valószínűségre; s: az átlag szórása; n: a megfigyelések száma a mintában.

Példa: Tekintsük meg a következő adatsort:

2. adatsor: 81.5, 69.9, 58.9, 74.0, 89.9, 54.8, 72.9, 73.9, 64.7, 64.4, 79.3, 83.1, 89.1, 78.3, 74.4.

A minta átlagos értéke: 73.9

A minta átlag szórása: 2.64

95%-os t-próba értéke 14 szabadságfoknál: 2.145

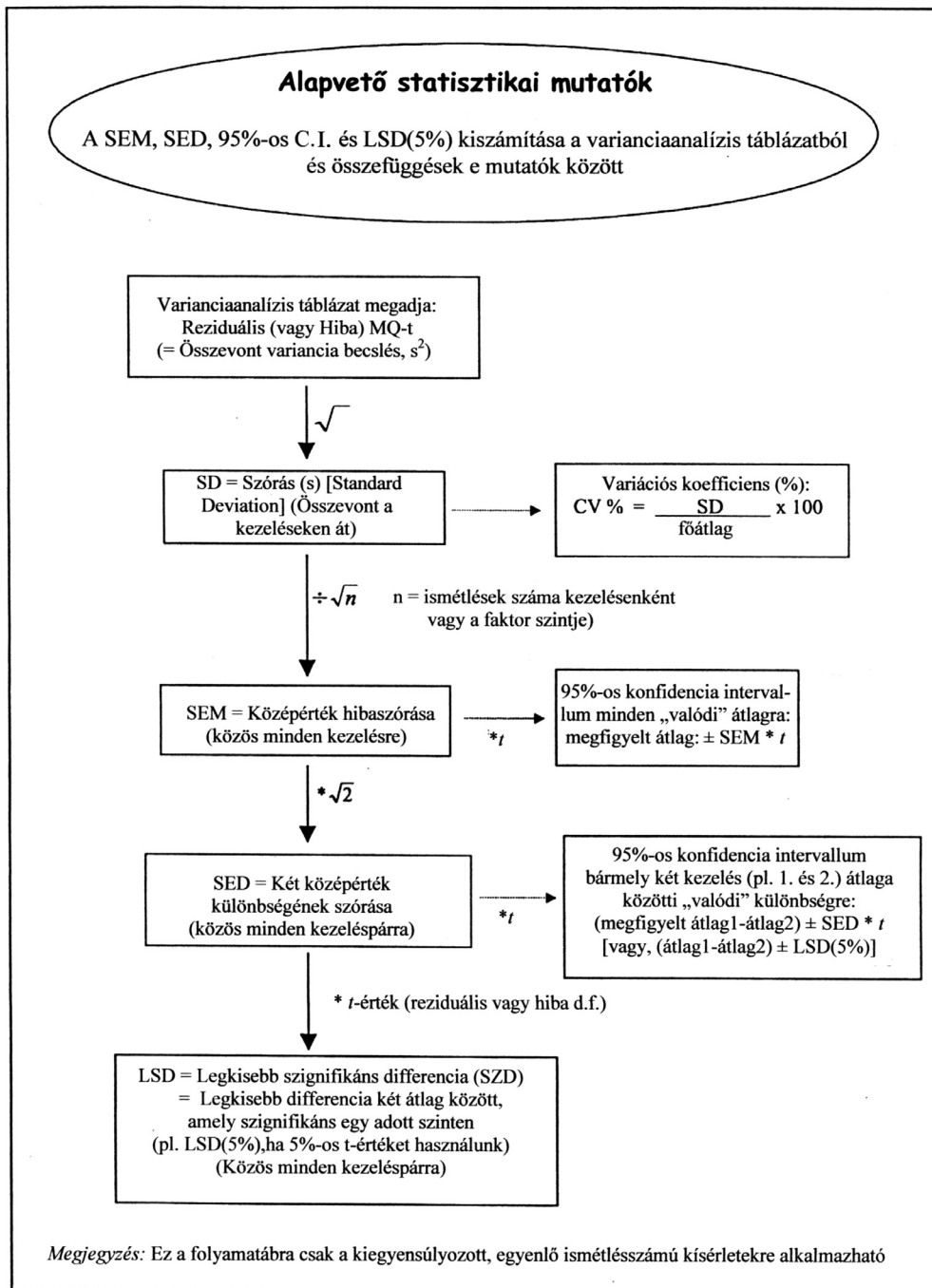
Alsó 95%-os konfidencia határ: $73.9 - 2.145 \cdot 2.64 = 68.2$

Felső 95%-os konfidencia határ: $73.9 + 2.145 \cdot 2.64 = 79.6$

95%-os konfidencia intervallum: $79.6 - 68.2 = 11.4$

Ezáltal 95%-os bizonyossággal állítható, hogy a 68.2–79.6 intervallum a valódi populáció átlagértéke, amelynek becsült értéke 73.9.

A leíró statisztikai mutatók közötti összefüggést a 7.1. ábra szemlélteti (Saville és Rowarth, 2008 alapján)



7.1. ábra. Az alapvető statisztikai mutatók közötti összefüggések folyamatábrája (Saville és Rowarth, 2008 alapján)

7.4. Ferdeség és csúcsosság

Amikor az adatokat analizáljuk mindenegyes megfigyelés helyzetére való utalás nélkül, a leggyakoribb analízis az átlagnak (a valószínűségi eloszlás első momentumának) és a varianciának (második momentum), valamint a valószínűségi eloszlás magasabb momentumainak kiszámítása. A valószínűségi eloszlás harmadik momentuma a ferdeség, melynek mértéke kvantitatív módon leírható a ferdeség koefficiensével, rendszerint g_1 szimbólummal jelölve. Számos különböző módszer van e koefficiens meghatározására.

A ferdeség (aszimmetria) Pearson-féle koefficiense, g_1 a következő:

$$g_1 = \frac{3(\text{átlag} - \text{medián})}{\text{szórás}}$$

A ferdeség jellemzi az eloszlás aszimmetriájának mértékét az átlag körül. Szimmetrikus eloszlások átlaga és mediánja egybeesik, ferdeség koefficiensük zéróval egyenlő. A ferdeség koefficiensnek előjelét és nem az abszolút értékét használjuk információként az eloszlás szimmetriájának vagy aszimmetriájának jellemzésére. Egy pozitív koefficiens az átlagnál kisebb mediánnal együtt egy hosszabb végpontot jelez a hisztogram jobboldalán, és egy negatív előjel egy hosszabb végpontot jelez a baloldalon, az átlagnál nagyobb medián értékkel együtt. Általában, -1 vagy $+1$ határokon kívüli értékek olyan ferde eloszlást jeleznek, amelyek észrevehetően eltérnek a szimmetrikus normális eloszlástól.

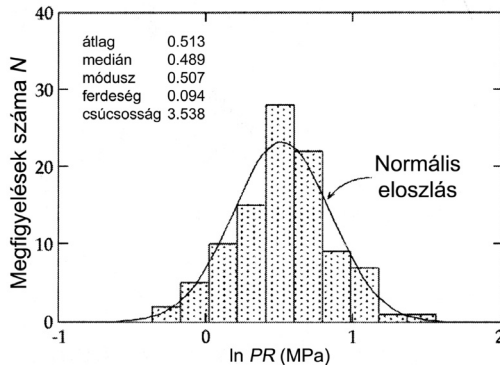
A valószínűségi eloszlás negyedik momentuma a csúcsosság. A csúcsossági mutatók közül az alábbi emeljük ki:

$$K = \frac{Q_3 - Q_1}{2(D_9 - D_1)}$$

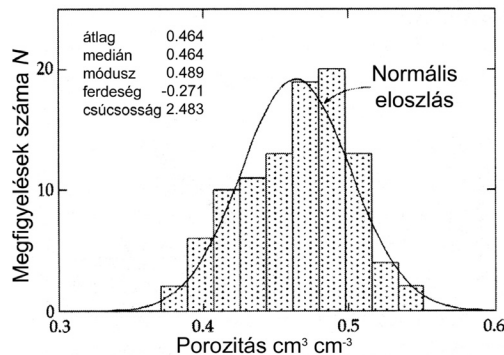
Ez azon a megfigyelésen alapul, hogy minél csúcsosabb egy eloszlás, annál kisebb a felső és alsó kvartilis különbségének fele a két szélső decilis – D_9 és D_1 – különbségéhez viszonyítva. A normális eloszlás esetében $K \approx 0.263$, ami a K mutató értékeléséhez bizonyos támpontot ad. A K mérőszáma valójában nem a csúcsosság, hanem a lapultság mérőszáma, mert minél lapultabb a vizsgált gyakorisági eloszlás, annál nagyobb K -értéket kapunk. Minél csúcsosabb ezzel szemben az eloszlás, annál kisebb lesz K értéke (Hunyadi et al., 2000). A lapos eloszlásúaknak kevés megfigyelésük van az eloszlás közepe közelében, míg a csúcsos eloszlásúaknak sok megfigyelésük van az eloszlás közepe közelében. Egy-egy példát a normális, ferde és csúcsos eloszlásra az 7.2–7.3. ábra mutat.

7.5. A leíró statisztikai mutatók kiszámítása GenStat programmal

Abból a célból, hogy kiszámítsuk a statisztikai mutatókat több mintára a GenStat felhasználásával, az adatsorokat be kell vinni a GenStat számolóablába, mint külön adatszlopok. A GenStat megnyitása és az adatok bevitele után, a *Stats* lehetőséget választjuk ki a GenStat eszköztárban, hogy megjelenítsük a statisztikai legördülő menüt. Az egérrel kijelöljük a *Summary Statistics* (összefoglaló statisztika) lehetőséget, hogy feltárjunk egy másodlagos *summary statistics* legördülő menü oldalát. A GenStat számolóablát és a legördülő statisztikai választási menüt a 4. ábra illusztrálja.



7.2. ábra. Talaj penetrométer ellenállásának mérése 1 m-ként egy 100 m-es szelvény mentén. A ferdeség megközelítőleg zéró, csúcosság > 3 értéke bizonyos csúcosságot jelez. Közel normális eloszlásúnak tekinthető (Nielsen és Wendroth, 2003 alapján)



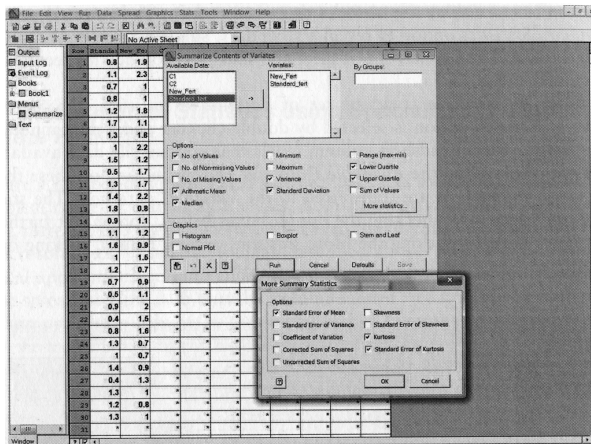
7.3. ábra. Talaj porozitás mérése 1 m-enként egy 100 m-es szelvény mentén. Az átlag és a medián közel azonos, negatív ferdeség enyhe végpontot jelez balra, és a csúcosság < 3 értéke kevesebb csúcosságot jelez, mint a normális eloszlás (Nielsen és Wendroth, 2003 alapján)

Row	Breakfast	lunch	dinner
1	0.8	1.8	
2	1.1	2.2	
3	0.7	1	
4	0.6	1	
5	1.2	1.8	
6	1.7	1.1	
7	1.3	1.8	
8	1	2.2	
9	1.5	1.2	
10	0.5	1.7	
11	1.3	1.7	
12	1.4	2.2	
13	1.6	0.8	
14	0.8	1.1	
15	1.1	1.2	
16	1.6	0.8	
17	1	1.8	
18	1.2	0.7	
19	0.7	0.8	
20	0.6	1.1	
21	0.9	2	
22	0.4	1.2	
23	0.8	1.4	
24	1.3	0.7	
25	1	0.7	
26	1.4	0.8	
27	0.4	1.3	
28	1.3	1	
29	1.4	0.8	
30	0.8	1.9	
31	1.1	2.3	

7.4. ábra. GenStat számolótábla és legördülő statisztikai menük

Abból a célból, hogy a leíró statisztika standard sorozatát megkapjuk, a *Summarize Contents of Variates* lehetőséget választjuk ki dupla kattintással és ekkor adatbeviteli doboz jelenik meg. A számolótábla oszlop feliratok, amelyek tartalmazzák az analízishez rendelkezésre álló adatokat, az *Available Data* dobozban jelennek meg és ezekre kattintva, az analízishez szükséges adatsorokat átvisszük a *Variables* adat dobozba. A szükséges statisztikai mutatókat úgy választjuk ki, hogy megvizsgáljuk az alatta levő *Option* (lehetőség) dobozokat, azzal a megjegyzéssel, hogy további mutatók rendelkezésre állnak, rákattintva a *More statistics* választási lehetőségre. Végül, rákattintva a *Run*-ra eredményezi a kiszámított mutatókat, melyek az *Output window*-ban (kimeneti ablak) jelennek meg. A kimeneti ablak megjelenítéséhez vagy az *Output* választható ki (dupla kattintással) az ablakok megjelenítési fán, a számláló tábla bal oldalán vagy kiválasztható a *Window* lehetőségen keresztül a fő eszköztárban. Az *output* kijelölhető és másolás-beillesztéssel betehető a *Word* dokumentumba megformázás céljából, ha szükséges.

A *GenStat* számolótábla és az *output window* az 5. ábrán látható a kezeletlen és a műtrágyával kezelt paradicsom növények növekedési rátájára vonatkozóan.



7.5. ábra. GenStat számolótábla és beállítási dobozok a leíró statisztikához

7.6. Grafikus ábrázolás

Egy kép ezer szó értékű. Ez igaz a statisztikai analízisben is: sokkal könnyebb megnézni egy ábrát, mint elolvasni egy hosszú leírást. Egy ábra képes az információt világosan és pontosan bemutatni és színek, árnyékolások, skálák és szimbólumok használata hangsúlyoz egy sajátos üzenetet az adatokban. A grafikonok azonban oly módon is prezentálhatók, hogy a tények torzítottak és a hangsúlyozott üzenetek rejtve maradhatnak. A szándék mindig az legyen, hogy az adatokban levő információt világos és tömör formában mutassuk be. Néha előnyösebb az információt táblázat formájában közölni.

A grafikonok használhatók arra, hogy mutassák a nyilvánvaló összefüggést a változók között, vagy feltárják az adatsorozat eloszlását, és nagyon hasznosak a kívülálló adatok vagy az adatokban levő más anomáliák jelzésére. Segítségül szolgálnak

GenStat output: Summary statistics**Summary statistics: 1. adatsor**

Number of observations = 13
 Number of missing values = 0
 Mean = 4.231
 Median = 5
 Minimum = 1
 Maximum = 8
 Range = 7
 Lower quartile = 2
 Upper quartile = 6
 Standard deviation = 2.242
 Standard error of mean = 0.622
 Variance = 5.026
 Coefficient of variation = 52.99
 95% Confidence interval for mean
 Lower bound = 2.876
 Upper bound = 5.586

Summary statistics: 3. adatsor

Number of observations = 12
 Number of missing values = 0
 Mean = 32.5
 Median = 27.5
 Minimum = 11
 Maximum = 90
 Range = 79
 Lower quartile = 20
 Upper quartile = 38.5
 Standard deviation = 20.87
 Standard error of mean = 6.026
 Variance = 435.7
 Coefficient of variation = 64.23
 95% Confidence interval for mean
 Lower bound = 19.24
 Upper bound = 45.76

Summary statistics: 2. adatsor

Number of observations = 15
 Number of missing values = 0
 Mean = 73.94
 Median = 74
 Minimum = 54.8
 Maximum = 89.9
 Range = 35.1
 Lower quartile = 66
 Upper quartile = 80.95
 Standard deviation = 10.21
 Standard error of mean = 2.637
 Variance = 104.3
 Coefficient of variation = 13.81
 95% Confidence interval for mean
 Lower bound = 68.29
 Upper bound = 79.60

Summary statistics: 4. adatsor

Number of observations = 15
 Number of missing values = 0
 Mean = 20.4
 Median = 23.3
 Minimum = 13.5
 Maximum = 27.5
 Range = 14
 Lower quartile = 14.9
 Upper quartile = 25
 Standard deviation = 5.542
 Standard error of mean = 1.431
 Variance = 30.72
 Coefficient of variation = 27.17
 95% Confidence interval for mean
 Lower bound = 17.33
 Upper bound = 23.47

továbbá az objektumok két- vagy több csoportja közötti vizuális összehasonlításra, vagy egy ábrában vagy az ábrák sorozatán keresztül. Óvatosság szükséges a grafikonon belül az információ értelmezésében, mivel a tengelyek nem megfelelő használata eltorzíthatja a valódi képet.

Figyelmet kell fordítani arra, hogy a grafikon világos és könnyen érthető legyen. Minden tengelyen feltüntetjük a változó nevét és mértékegységét. Ha egy skála nem nullánál kezdődik, világosan jelezni kell. A skála megszakítást rendszerint cikcakk formájában jelöljük a tengelyen vagy megszakítjuk a vonalat párhuzamos meredekségű vonalakkal. Megjegyezzük, hogy több számítógépes csomag nem ajánl lehetőséget az ilyen skála megszakításra, azonban ilyen esetben is utalni kell arra, hogy a skála

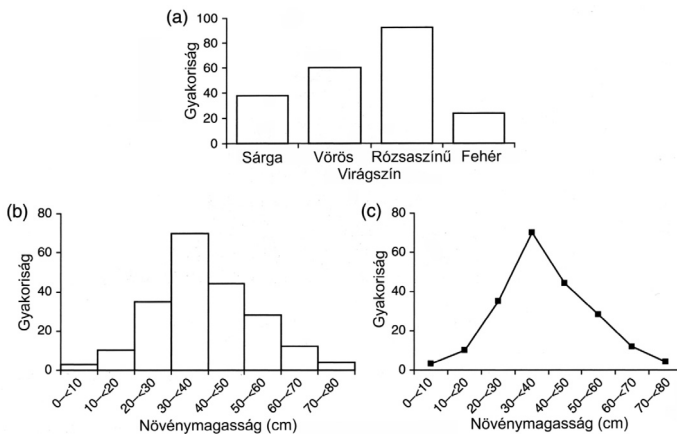
nem nullánál kezdődik. Ha a skála túl kicsi, akkor az információk elfedettek lehetnek, azonban, ha túl nagy, az információk elferdítettek lehetnek, felnagyítva kis különbségeket. Egy hasznos szabálya a lineáris trendek ábrázolásakor az, hogy a skálát úgy válasszuk meg, hogy a vonal meredeksége 30° és 45° között legyen.

A grafikon bemutatására és a grafikon típusának kiválasztására különös figyelmet fordítsunk. A grafikon egyik formája nagyon eltérő képet adhat ugyanarról az adatsorozatról, ha egy másik formában mutatjuk be. Egy grafikon, amelyben a pontokat összekötjük, eltérő értelemmel bír, mint ugyanaz a grafikon, amelyen a pontokat nem kötöttük össze. A szimbólumok használhatók, mint amilyenek a csillagok és pontok, azonban ha egynél több szimbólumot használunk, akkor azok könnyen megkülönböztethetők legyenek.

Oszlopdigramok

Oszlopdigramot használunk, hogy megnézzük egy változó eloszlását, amely diszkrét adatokat tartalmaz. Az oszlopdigram információt ad az eloszlás szimmetriájáról és az értékek tartományáról, és segíthet identifikálni extrém értékeket vagy kivülálló adatokat.

Mindegyik oszlop reprezentál egy eltérő értéket és az oszlop magassága arányos ennek az értéknek a gyakoriságával. Az oszlop bármilyen szélességű lehet, jóllehet félreérthető, ha néhány oszlop szélesebb, mint a többi, ezért rendszerint minden oszlop szélességét azonosnak választjuk. Mindegyik oszlopot elkülönítjük a következő oszloptól egy térközzel, mivel ez jelzi, hogy mindegyik oszlop eltérő és az adatok diszkréték (7.6. ábra).



7.6. ábra. Gyakorisági eloszlás ábrák: (a) oszlopdigram, (b) hisztogram, (c) gyakorisági poligon

A legegyszerűbb módja egy oszlopdigram ábrázolásának eltérő változókkal, hogy ábrázoljuk a tényleges gyakoriságot a változó lehetséges értékeink tartománya szerint, ezután az oszlop magassága reprezentálja a tényleges gyakoriságot. Ha a tényleges gyakoriságot használjuk, közvetlen összehasonlítás az oszlopdigramok között csak akkor végezhető, ha mindegyik ábra azonos összes megfigyeléssel rendelkező változókat használ. Például, az 5 értékű gyakoriságnak más értelmezése van, ha össze-

sen 10 megfigyelésből származik, mint amikor 100 megfigyelésből származik. Relatív gyakoriságok használata lehetővé teszi, hogy összehasonlítást tegyünk eltérő számú összes megfigyelésből származó változók oszlopdiagramjai között. Ebben az esetben a magasság a megfigyelések hányada, amelyeknek ez az értéke.

Az összehasonlítás az oszlopdiagramok között egyszerűbbé válik, ha egynél több változót rajzolunk ábránként. Mindegyik változónak hasonló értéktartománya legyen. Az oszlopok sorozata minden eltérő értékre lehet osztott, csoportosított, átfedésben levő vagy szomszédos, a bemutatni kívánt információtól függően.

Hisztogram

A hisztogram a folytonos adat-változó megjelölője az oszlopdiagramnak. Információt nyújt az eloszlás szimmetriájáról és az értékek tartományáról, továbbá segíthet identifikálni extrém értékeket és kívülálló adatokat (7.6. ábra). A változó értékek tartománya csoportokra van felosztva és mindegyik csoportot egy osztály-intervallumnak hívunk. A hisztogram különbözik az oszlopdiagramtól, amennyiben egy derékszögű négyszög van rajzolva minden osztályintervallum fölé. A legegyszerűbb esetben egy derékszögű négyszög területe, nem pedig a magassága a megfigyelések gyakorisága, amely az osztályintervallumba esik és nincs térköz kihagyva a négyszögek között. Ha az osztályintervallum egységnyi szélességű, akkor az oszlop magassága egyenlő lesz a területtel és ezáltal a tényleges gyakoriságot reprezentálja. Az osztály-intervallumok nem szükségszerűen egyenlő szélességűek és ezekben az esetekben a vertikális tengelyt „gyakoriság denzitással” nevezzük meg.

Hasonlóan az oszlopdiagramokhoz, ha a tényleges gyakoriságokat használunk, akkor közvetlen összehasonlítás a hisztogramok között csak akkor tehető, ha mindegyik hisztogram ugyanolyan összes megfigyelésszámú változót használ. Minden összehasonlítandó hisztogram osztály-intervallumának azonosnak kell lenni. A gyakorlatban, feltételezve az osztály-intervallumok azonosságát, a gyakoriság ábrázolható a vertikális tengelyen, anélkül, hogy hatással lenne az ábra általános alakjára. Ez az eset a legtöbb statisztikai csomagnál. Óvatosság szükséges nyilvánvalóan a vertikális tengely megnevezésénél, amikor hisztogramokkal foglalkozunk, hogy elkerüljük a keveredést az összehasonlítások végzésekor.

A relatív gyakoriságok használata lehetővé teszi, hogy összehasonlítást végezzünk a különböző összes megfigyelésszámú változók hisztogramjai között. Ebben az esetben a négyszög területe a megfigyelések aránya, amely az osztály-intervallumban van.

A hisztogramhoz kiválasztott osztály-intervallumoknak kölcsönösen kizárónak kell lenni, úgyhogy minden adatpontnak egy és csakis egy intervallumba kell esni. Ezt elérhetjük annak biztosításával, hogy az intervallumok lefedik az adatok teljes tartományát és azáltal, hogy az intervallumokat eggyel több tizedeshellyel vagy jelentős értékkel fejezzük ki, mint amit az adatoknál használtunk. Minden négyszög az intervallum középpontja felett helyezkedik el. Vagy a középső értékek vagy az intervallum szélső értékei vannak identifikálva a négyszög alatt. Néha a végső osztály-intervallumok nyitott-végűek, mivel nem kívánatos, hogy magában foglaljon több intervallumot, hogy lefedjen egy vagy két extrém értéket. Ezekben az esetekben az utolsó intervallumok nem ábrázolhatók a hisztogramban. Például, egy állat életkorában az osztály-intervallumok lehetnek <5 nap, 6–10 nap, 11–20 nap, és 21 nap vagy több.

Nincsenek szigorú szabályok a használni kívánt osztály-intervallumok számára, azonban van néhány javaslat. Az egyik leggyakrabban idézett az, hogy használjunk 10 osztály-intervallumot, a másik az, hogy használjunk \sqrt{n} osztály-intervallumot, ahol n a megfigyelések összes száma. A hisztogram osztály-intervallumainak számát illetően megemlítjük a $k=1+\log_2 n$ lehetőséget is. Alternatív módon, az adatok megadják maguknak a nyilvánvaló intervallumot, pl. 1-es intervallum-szélesség hasznos lehet olyan adatokra, melyek tartománya 0 és 15 fok közé esik. Megemlítjük a $k = 1 + \log_2 n$ lehetőséget is.

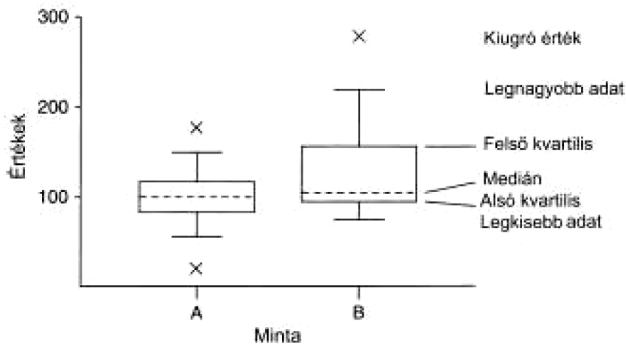
Bármilyen szabályt is használunk, fontos, hogy ne válasszunk olyan intervallumokat, amelyek túl szűkek vagy túl tágak, mivel befolyásolják a hisztogram alakját. Ha az intervallumok túl kicsik, akkor kevés megfigyelést fognak tartalmazni és a hisztogram nagyon szabálytalanként jelenik meg, miközben ha az intervallumok túl szélesek, a minta eloszlásnak csak egy nagyon durva reprezentációját fogjuk elérni.

Kvartilis (vagy doboz) ábra

A kvartilis ábrák használhatók arra, hogy bemutassunk egy kvantitatív változó fő tulajdonságait. Különösen hasznosak, amikor összehasonlítjuk a durván azonos méretű adatok csoportjait, mivel különbségeket a variabilitásban és a helyzetben világosan illusztrálják a kvartilis ábrák. A kvartilis ábra az eloszlás öt tulajdonságát mutatja: legkisebb adat, alsó kvartilis, medián, felső kvartilis, legnagyobb adat.

Leggyakrabban ezek az ábrák tartalmazzak egy dobozt, a doboz hossza egyenlő az interkvartilis tartománnyal és a doboz oldalai jelölik az alsó és a felső kvartilist, úgyhogy a megfigyelések 50%-a dobozon belül helyezkedik el. A doboz mindkét oldalából a szakaszok (whiskers) kiterjednek a minimális és maximális értékig. A medián helyzetét a dobozon belül egy vonal jelzi. Ha a medián a doboz hosszanti oldalának közepén helyezkedik el, ez a szimmetrikus eloszlást jelzi. A leíró statisztika aktuális értékei is fel vannak tüntetve a 7.7. ábrán.

A minta:	átlag	=	100.0	B minta:	átlag	=	129.0
(szimmetrikus	medián	=	100	(ferde	medián	=	105
eloszlás)	szórás	=	35.1	eloszlás)	szórás	=	53.1

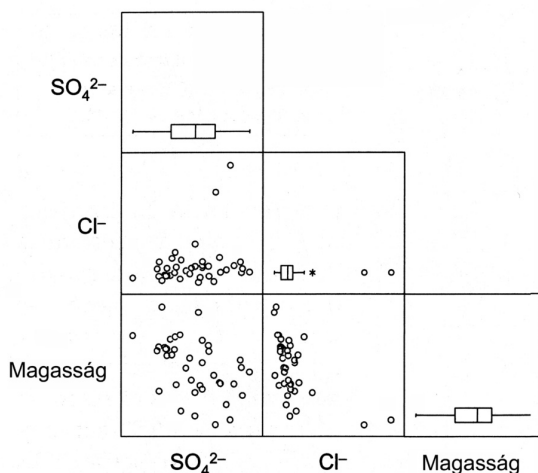


7.7. ábra. A dobozábra szerkezete két mintára. A szakaszok (legnagyobb adat-legkisebb adat, whiskers) magukba foglalnak minden adat értéket, melyek az $1.5 \times$ interkvartilis tartományon belül helyezkednek el a kvartiliseken kívül.

A doboz két oldalán a szakaszok kiterjednek a legnagyobb és a legkisebb értékig a doboz oldalainak 1.5-ös inter-kvartilis tartományán belül. Egyes adatpontok, amelyek a szakaszokon kívül esnek, elkülönítetten azonosítva vannak az ábrán, a háromszoros inter-kvartilis tartományon kívüli értékek ki vannak emelve, tehát jelzést adnak az extrém értékekre és a lehetséges kívülálló értékekre. A változók közötti összehasonlítások könnyen elvégezhetők úgy, hogy egy vagy több kvartilis ábrát helyezünk el ugyanabban a grafikonban.

Szóródás ábra

A szóródás ábrák kétváltozós ábrák, egyik változó ábrázolása a másik változó értékei szerint (7.8. ábra). Ezáltal felhasználhatók arra, hogy identifikáljunk potenciális összefüggéseket a változók között, jöllehet más a felhasználási területük. A változók közötti potenciális összefüggések formáját, mint amilyen a lineáris vagy kvadratikussá, tudni kell a formális analízis előtt, amely csak az adatok ábrázolásával érhető el. Ezek az ábrák arra is felhasználhatók, hogy detektáljuk a megfigyelések csoportosulását és becsljük, mennyire jól írta le egy statisztikai modell az adatokat. A szóródás diagramok számos módon prezentálhatók. Egy hasznos technika a csoportok azonosítására az ábrán az, hogy különböző ábrázolási szimbólumokat használunk.



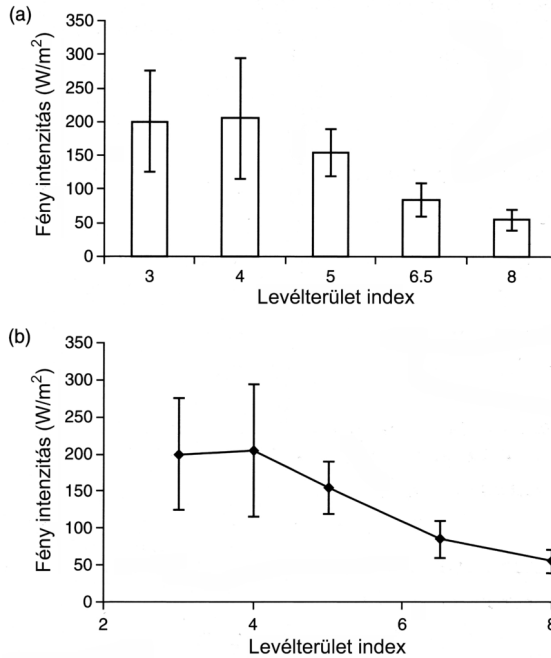
7.8. ábra. Szórásdiagram mátrix három változóra (tengerszint feletti magasság, SO_4^{2-} koncentráció és Cl^- koncentráció) 39 helyre erdős vízgyűjtő medencére Catskill Mountains-ben, New York államban (Quinn és Keough, 2002)

Egy másik hasznos technika egyesíteni az egymást követő pontokat. Ez különösen hasznos, amikor a faktorok közötti interakció hatásokat vizsgáljuk. A pontok egyesítése nem mindig megfelelő (pl. reziduális ábránál) és minden esetet vizsgálni kellene a saját érdeme szerint.

A szóródás ábrák segíthetnek identifikálni szokatlan megfigyeléseket, mint pl. egyetlen pont távol a pontok fő csoportosulásától vagy egy pont, amely távol van azoktól, melyek egy egyenesen helyezkednek el.

Hiba ábra

A hiba ábrák az adatok variabilitásáról adnak jelzést. Feltüntethetők a hisztogramok fölött vagy a pontok körül egy szóródás ábrán. Ezek a hiba szakaszok (bar) lehetnek előrejelzett standard hibák, ha illesztett egyenest ábrázolunk, vagy egyszerűen átlagok szórásai, amikor átlag értékeket ábrázolunk. Az ábrával együtt meg kell nevezni, milyen hiba szakaszcsoportról (bar) van szó. Az átlag értékek hiba szakaszokkal ábrázolva mutatják az átlagot és plusz vagy mínusz egyszeres szórást (7.9. ábra).



7.9. ábra. A hibaszávok használata mutatja az átlag szórását az oszlopdiagramon (a) és vonaldiagramon (b). Az oszlopdiagram (a) mutatja az átlagos fényintenzitást a talaj felszínén búzaállományban, eltérő levélfelület indexeknél (a vertikális hibaszávok jelzik a \pm szórást, $n = 10$). A vonaldiagram (b) mutatja az átlagos fényintenzitást a talaj felszínén különböző levélfelület indexeknél (a vertikális hibaszávok jelzik a \pm szórást, $n = 10$)

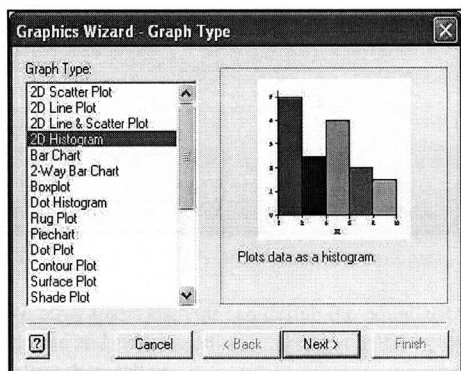
Kördiagram

A kördiagram egy kör, amely jelzi az összes adatot, amelyek körszeletekre vannak felosztva, jelezve a csoportok arányát az adatokon belül. A kördiagram ábráján a körszelet nagysága arányos az előfordulási gyakorisággal. Kvantitatív és kvalitatív változók ábrázolására egyaránt alkalmas.

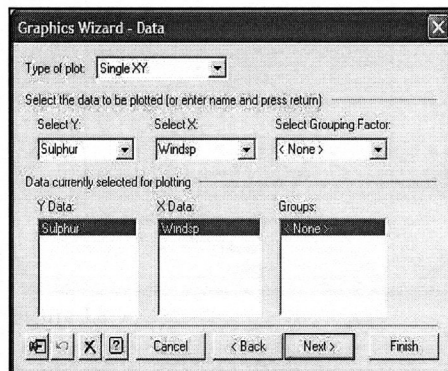
7.7. Ábrák szerkesztése GenStat programmal

Ábrák készítéséhez kiválasztjuk a *Graphics* menüből a *Create Graph* lehetőséget. Ez kinyitja a *Graphics Wizard* (ábra varázsló) első menüjét, mint a 7.10. ábra mutatja. A menüből kiválasztottuk a 2-D Histogram lehetőséget a *Graph Type* listából. A *Next* gombra kattintva, kinyílik a második menü, amely az adatok bevitelére szolgál. Az ábrakészítés további részleteit a GenStat kézikönyv tartalmazza.

Számos statisztikai vizsgálat nem egyetlen változóra, hanem több változó összefüggésére koncentrál. Egy szóródás diagram készítéséhez kiválasztottuk a *2-D Scatter Plot* lehetőséget a *Graphics Wizard* menüben (7.11. ábra).



7.10. ábra. Ábra típusok a Graphic Wizard menüben



7.11. ábra Több változó összefüggésének ábrázolása a Graphic Wizard menüben

7.8. Táblázatok

- A táblázatok elősegítik a statisztikai analízis értelmezését.
- Hasznosak az eredmények bemutatásában.
- A táblázatok legyenek egyszerűek.

A táblázatok hasznosak a numerikus vagy szöveges adatok és statisztikai mutatók közlésére tömör formában. Segíthetnek bemutatni az összefüggéseket a változók között és segítenek az eredmények interpretálásában is. Felhasználhatók prezentálni összefoglaló adatokat, átlagos értékeket, szórásokat és a megfigyelések számát, felhasználhatók továbbá adatsorok bemutatására is. Ha egynél több információt közlünk egyetlen cellában (táblamezőben), akkor a táblázatot világosan el kell látni jegyzetekkel a cella tartalmának sorrendjére vonatkozóan. Nem mindig könnyű összeállítani egy világos táblázatot és figyelmet kell fordítani az alábbi szempontokra. Egy táblázat bonyolultságát nagymértékben meghatározza a felhasznált osztályozási faktorok száma. Egy *osztályozási faktor* a változó hasonló egységeit egybe csoportosítja. Ezek a csoportosítások azután felhasználhatók arra, hogy definiálják a táblázat sorait és oszlopait, amelyek viszont identifikálják a táblázat celláit. Egy táblázat dimenziói közvetlenül megfelelnek a táblázat definiálásához használt csoportosítási faktorok számának. Két csoportosítási faktor szükséges egy kétirányú táblázat meghatározásához, három szükséges a háromirányú táblázat meghatározásához, és így tovább. Amikor háromnál több csoportosítási faktort használunk, egy faktor oszlopait vagy sorait

hierarchikusan kell egymáson belül elhelyezni, hogy prezentálják az információt egy két-dimenziós formátumban, pl. papíron vagy a számítógép képernyőjén. Ahogy egyre több csoportosítási faktort használunk, a táblázat egyre bonyolultabbá válik és nehezebbé válik olvasni és megérteni a táblázat tartalmát. Általánosságban, nehezebb olvasni és megérteni a három vagy négy dimenziós táblázatot. Ha szükséges, egy nagy táblázat felosztható kisebb táblázatokra, hogy világosabban prezentálják az információt. A táblázatokat a lehető legegyszerűbben kell összeállítani.

A táblázat térköze és beosztása (szerkezete) nagyon fontos. Üres sorok és oszlopok megfontolt használata segíthet egy táblázatot könnyebben olvashatóvá tenni. Segít az értelmezésben ahogy a sorok és oszlopok sorrendjét prezentáljuk. Például, ha különböző oszlopok információt tartalmaznak a különböző évekre, akkor az oszlopokat kronologikus sorrendbe célszerű rendezni. Könnyebb megnézni egy táblázat oszlopait, mint keresztülnézni egy soron, így segít ha a táblázatnak több a sora, mint az oszlopa, jöllehet ez nem mindig lehetséges, ha információt számos változóból mutatunk be egy táblázatban. Ha egy táblázatban az oszlopokat akarjuk összehasonlítani, akkor egymáshoz közel kell azokat elhelyezni.

Legyünk óvatosak a számítógépes output-ok táblázatban történő közvetlen felhasználásában, mivel nagyon zavaró és félrevezető lehet közölni túl sok tizedes-jegyvet a táblázatban. A helyzet mérőszámait is közölni kellene a fontos mért adatok mellett. Egy standard gyakorlat, hogy a variabilitás mérőszámokat eggyel több jelentős számjeggyel adjuk meg, mint a helyzet mérőszámait.

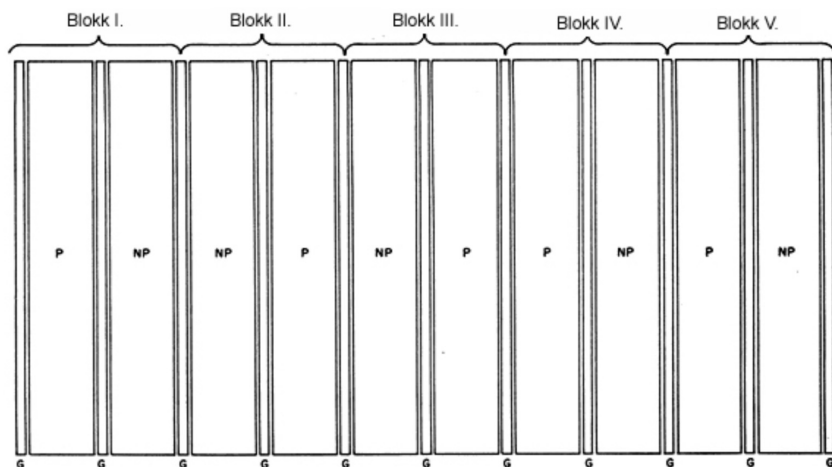
Vegyük figyelembe, hogy hány összefoglaló statisztikai mutatót közlünk a táblázatban. Például átlagolni az oszlopokon keresztül helytelen lenne, ha az oszlopok teljesen eltérő változókat reprezentálnak. Bármikor, ha átlagértékeket közlünk, a variabilitás megfelelő mértékét és a megfigyelések számát ugyancsak közölni kellene ugyanabban a táblázatban. Például, ha a táblázat egyszerűen csupán összegezi az adatokat, akkor a szórást vagy az átlag szórását és a megfigyelések számát közölni kellene minden átlagértékkel. Ha a táblázatot arra használjuk, hogy összehasonlítást végzünk az átlagértékek között, akkor a két átlagérték közötti különbség szórása megfelelőbb lenne.

Tudományos közleményekben a táblázatoknak érthetőnek kell lenni anélkül, hogy hivatkoznánk a szöveges részre. A táblázat címének világosnak és önmagát magyarázóknak kell lenni, és ha egynél több táblázat van, értelemszerűen kell beszámolni azokat. Az olvasót nem kell arra kényszeríteni, hogy a szöveges részhez térjen vissza, hogy megtudja a táblázat tartalmát. A sorokat és oszlopokat világosan és megkülönböztethető módon kell megjelölni, a mértékegység feltüntetésével, ha azok oszloponként vagy soronként különböznek. Ha a mértékegységek ugyanazok, akkor a táblázat kevésbé látszik zsúfoltnak, ha a mértékegységet a címben közöljük. Bármely rövidítést vagy kifejezést, melyet a fejlécben vagy mint cella- tartalmat használtunk, magyarázatokkal kell ellátni a lábjegyzetben. Például „*” használható, hogy jelezzen egy hiányzó értéket vagy jelezze a szignifikancia szintjét és „na” jelzi, hogy a feljegyzett statisztika nem volt alkalmazható erre a sor és oszlop kombinációra. Rendszerint minden cellát kitöltünk, ha azonban a cellákat üresen hagyjuk, azt magyarázattal kell ellátni.

8. Egymintás és kétmintás paraméteres próbák

8.1. Bevezetés

A statisztikai hipotézisek eldöntésére alkalmas próbák két csoportba különíthetők el, parametrikus és nem-parametrikus próbák. A parametrikus próbák közös jellemzője, hogy feltételezik a vizsgált változók normáeloszlását. A másik feltétel, hogy a mérések legalább intervallumskálán történjenek, tehát kategóriába sorolt adatokon nem lehet pl. t-próbát végezni. Az elnevezés onnan ered, hogy az általuk tesztelt nullhipotézisek a populáció eloszlásának valamely – elméletileg is jól leírható – paraméteréről állítanak valamit. Előnyeik közé tartozik, hogy ha a feltételeik teljesülnek, akkor érzékenyebbek a nem-parametrikus megfelelőjüknél. Másik tulajdonságuk, hogy jól kidolgozott elmélet áll a háttérükben, így a próbastatisztikák eloszlása jól ismert (Précsényi et al., 1995).



8.1. ábra. Védelemben részesített (P) és kontroll (NP) parcellák randomizált blokk elrendezésben, öt ismétlésben, elválasztó sorokkal (G) a parcellák között

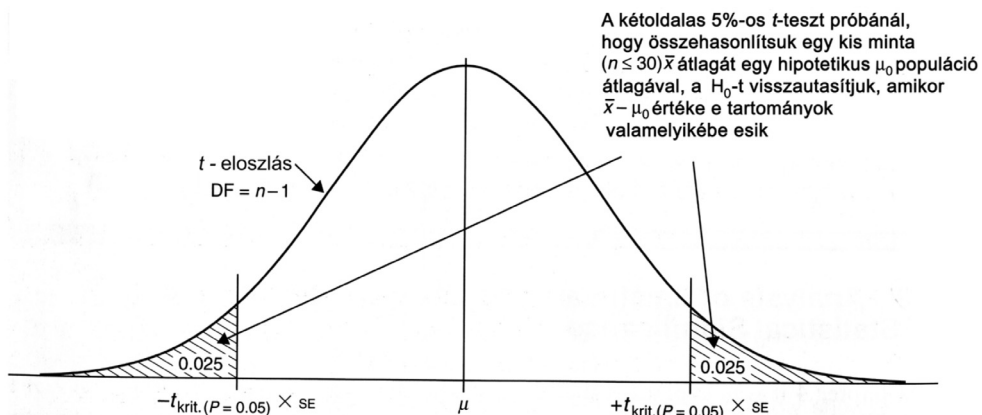
Az egyik leggyakoribb és valóban a legegyszerűbb típusa a kísérlet tervezésnek, amikor két összehasonlítandó kezelést alkalmazunk két független populációban és egyetlen mintát választunk ki ezután a két populáció mindegyikéből a méréshez. Ezután szükséges, hogy teszteljük a szignifikáns különbséget a két populáció között, felhasználva a minta adatokat a populáció paraméterek becslésére. Ez a helyzet különösen gyakran bekövetkezik ott, ahol az egyik populáció reprezentál egy kísérleti kezelést és a másik populáció reprezentál egy kezeletlen kontrollt (8.1. ábra). Például

termésvesztések becslése vagy egy új műtrágyával kezelt búzapopulációból vett minta termését összehasonlítjuk a kezeletlen búzapopulációból vett minta termésével. Kevésbé gyakori, mégis eléggé ismert az a helyzet, amikor egyetlen mintát veszünk egy populációból és meg kell tudnunk, hogy a minta átlag szignifikánsan különbözik-e egy eltérő populáció elméleti vagy hipotetikus átlagától. Például, egy újonnan nemesített tehénfajta átlagos tejhozama szignifikánsan különbözik-e egy meglévő tehénfajta már ismert átlagos tejhozamától. Mindkét esetben, ha a mért minták olyan populációkból származnak, amelyekről biztonsággal feltételezhető, hogy normális eloszlásúak, akkor a paraméteres szignifikancia teszt elvégezhető.

Két paraméteres teszt létezik két minta összehasonlítására, nevezetesen a z-teszt, ahol az analízis nagy mintákon ($n > 30$) alapul, és a Student t-teszt, melyet eredetileg kis minták ($n \leq 30$) analízisére dolgoztak ki, azonban valójában alkalmazható bármely minta méretnél (Ireland, 2010). A következőkben a Student t-tesztet ismertetjük részletesebben. Ha azonban a két összehasonlítandó minta olyan populációkból származik, melyekről nem feltételezhető, hogy normális eloszlásúak, és nem kívánjuk adattranszformációval elérni a normális eloszlást, akkor egy alternatív, nem paraméteres eljárásra van szükség. A t próbák alapelveit és kiszámítási módszereit Ireland (2010) és Précsényi et al. (1995) alapján ismertetjük.

8.2. Egmintás t-próba

A Student t-teszt használatát azzal kezdjük, hogy összehasonlítjuk a populáció átlagát (melyet kis mintából becslünk) egyetlen hipotetikus populáció átlaggal. A teszt célja meghatározni, hogy vajon a μ_0 hipotetikus populáció átlag és az \bar{x} mintaátlag közötti különbség megfigyelt értékének valószínűsége kisebb-e, mint a kiválasztott kritikus valószínűség, a null hipotézis (H_0) visszautasítására. Tehát az 5%-os Student t-tesztnél a H_0 -t visszautasítjuk, ha $\bar{x} - \mu_0$ a t-eloszlás 2.5%-ának bármelyik oldalára esik, mint ahogy a 8.2. ábra mutatja.



8.2. ábra. A Student t-eloszlás tartománya a null hipotézis visszautasítására a kétoldalas 5%-os Student t-próbánál, összehasonlítva egy megfigyelt kis minta átlagot (\bar{x}) egy hipotetikus populáció átlaggal (μ_0), SE = standard hiba (Ireland, 2010)

A probléma az, hogy mivel mintákkal dolgozunk, nem ismerjük a populáció paraméterek értékét és a populáció szórását, σ -t csak a minta s szórásán keresztül becsülhetjük, amelyet a populációból származtattunk. Minthogy a minta átlagok sorozatának szórását megadja az átlag szórása, a $t_{\text{megfigyelt}}$ értékét megkapjuk ha a populáció és a minta átlag közötti különbséget elosztjuk a minta szórásával (SE). Minthogy a szórást s/\sqrt{n} adja meg:

$$t_{\text{megfigyelt}} = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

ahol: \bar{x} = minta átlag, μ = populáció átlag, s = minta szórása, n = megfigyelések száma a mintában.

A $t_{\text{megfigyelt}}$ értékét ezután összehasonlítjuk a Student t-eloszlás kritikus értékeivel (*Melléklet 1. táblázat*). Mivel a Student t-eloszlás függ a minta méretétől, a Student t-táblázathoz kell használni a minta szabadságfokát, amely $n-1$. Ha a $t_{\text{megfigyelt}}$ nagyobb, mint a kritikus t-érték, akkor a null hipotézis elutasítható a valószínűség adott szintjén. Fordítva, ha $t_{\text{megfigyelt}}$ kisebb, mint a t kritikus értéke, akkor a null hipotézist el kell fogadni. A Student t-teszt elvégezhető egyoldali vagy kétoldali próbaként, mint az alábbi példa illusztrálja.

Példa. A szignifikancia Student t-próbája egy kis minta átlaga és egy hipotetikus populáció átlag közötti különbség analizésére

Egy kétoldali próba annak meghatározására, hogy egy újonnan nemesített tehénfajtának szignifikánsan megváltozott-e a fehérje kazein szintje a tejben, összehasonlítva egy meglévő tejelő tehénfajtaival.

Tételezzük fel, hogy a meglévő tejelő fajta átlagos tej kazein tartalma ismert és 3.2 mg g^{-1} . A tenyésztők tudni akarják, hogy történt-e szignifikáns változás az új tehénfajta tejének kazein tartalmában, összehasonlítva a meglévő fajtaival. Az újonnan nemesített 15 tehén tej mintájában az átlagos kazein szint 3.7 mg g^{-1} volt és a szórás (s) 0.55 mg g^{-1} .

Null hipotézis: Nincs különbség az újonnan nemesített fajta populáció átlaga és a hipotetikus populáció átlag között ($H_0: \mu = \mu_0$).

A próba statisztika értékét a következőképpen számítjuk ki:

$$t_{\text{megfigyelt}} = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{3.7 - 3.2}{0.55/\sqrt{15}} = 3.52$$

Mivel a minta mérete, $n = 15$, a szabadságfoka a t-próbához: $15-1 = 14$.

5%-os szignifikancia tesztre a t kritikus értéke egy kétoldali próbára a t-táblázatból ($\alpha = 0.05$, d.f. = 14): 2.145. A $t_{\text{megfigyelt}}$ értéke nagyobb mint 2.145, ezáltal a null hipotézist visszautasítjuk.

Következtetés: az újonnan nemesített tejelő tehénfajta tejének kazein tartalma szignifikánsan különbözik a régebbi tehénfajta tejének kazein tartalmától ($P < 0.05$).

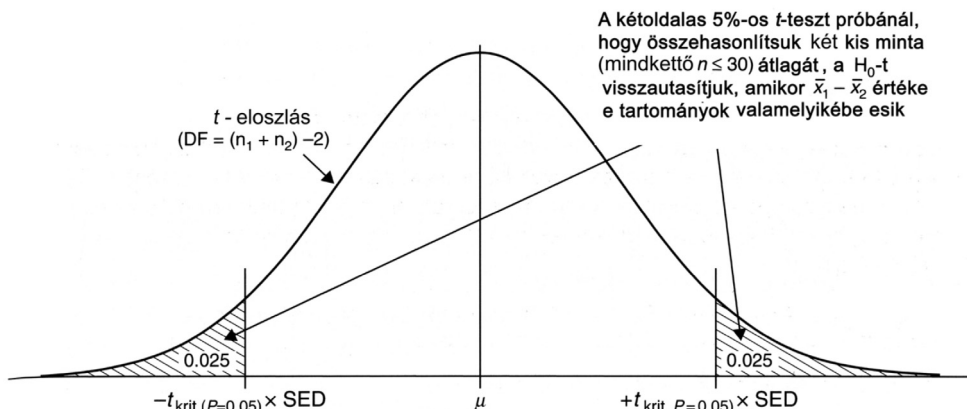
Egyoldali próba: az analízis egyoldali Student t-próbaként is elvégezhető, megvizsgálva, hogy az új nemesített tehénfajta teheneinek nagyobb-e a tej kazein tartalma, mint a régi tehénfajtáké. Ebben az esetben a H_0 -t a következő formában állítjuk fel: az új fajta populáció átlaga nem nagyobb, mint a hipotetikus populáció átlag ($H_0: \mu \leq \mu_0$). A kritikus egyoldali érték $\alpha = 0.05$ -nél 1.761. Minthogy $t_{\text{megfigyelt}} > 1.761$, a H_0 -t vissza-

utasítjuk és levonjuk azt a következtetést, hogy az új tehénfajta tejében a kazein tartalom szignifikánsan nagyobb, mint a meglévő tehénfajta tej kazein tartalma ($P < 0.05$).

8.3. Kétmintás t-próba

Az egyik legáltalánosabban használt statisztikai szignifikancia teszt a Student t-teszt két kis független minta közötti különbség analizésére, ahol mindkét mintáról biztonsággal feltételezhető, hogy normális eloszlású populációkból származnak. Ez a teszt felhasználható bármely mintaméret összehasonlítására, azonban használni kell, ha bármelyik minta mérete ≤ 30 . A $t_{\text{megfigyelt}}$ teszt statisztikát (mutatót) a Student t-tesztre megkapjuk, ha két minta átlag különbségét elosztjuk a különbség standard hibájával (két átlagérték közötti különbség szórásával) (SED):

$$t_{\text{megfigyelt}} = \frac{\text{minta átlagok közötti különbség}}{\text{SED}}$$



8.3. ábra. A Student t-eloszlás tartománya a null hipotézis visszautasítására egy kétoldalu 5%-os t-próbában, vizsgálva a különbséget két mintaátlag (\bar{x}_1 és \bar{x}_2) között, SED = különbség standard hibája (Ireland, 2010)

Miután meghatároztuk a $t_{\text{megfigyelt}}$ értékét, viszonyítjuk a t kritikus értékéhez. Ezt illusztrálja a 8.3. ábra.

A SED meghatározásának módszere függ azonban attól, hogy a két populáció varianciája, melyből a minták származnak, egyenlő vagy nem. Ez valóban megfelelő feltétel a kísérleti helyzetek többségében, mivel különböző kezeléseket alkalmazva az egyedek populációjára megváltoztathatja a kezelt egyedek egy sorozatának átlagát, összehasonlítva egy másikkal, anélkül, hogy hatással lenne az átlag körüli értékek eloszlására. Ha a két populációról feltételezhető, hogy egyforma a variációjuk, akkor a SED számítása során a két minta variancia összevonható, hogy a populáció variancia pontosabb becslését adja. Ha azonban a két populáció varianciáról nem feltételezhető, hogy egyenlő, akkor egy eltérő matematikai megközelítés szükséges.

8.3.1. Kétmintás t-próba, feltételezve egyenlő populáció varianciát

Ahol a két kis mintáról feltételezhető, hogy egyenlő varianciájú két populációból származik, a $t_{\text{megfigyelt}}$ statisztikai mutatót megkapjuk:

$$t_{\text{megfigyelt}} = \frac{\bar{x}_1 - \bar{x}_2}{SED} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s_p^2 \times \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}}$$

ahol: \bar{x}_1, \bar{x}_2 = az 1. és 2. minta átlagai

n_1, n_2 = az 1. és 2. minta mérete

s_p^2 = a két minta összevont varianciája

$$= \frac{s_1^2(n-1) + s_2^2(n_2-1)}{(n_1-1) + (n_2-1)}$$

s_1^2, s_2^2 = az 1. és 2. minta varianciája

Mivel ez a számítás magában foglalja a két minta variáció egyesítését, hogy egy összevont mutatót, s_p^2 , képezzünk, a szabadságfok (d.f.) a teszthez a két minta szabadságfokának az összege, azaz: d.f. = $(n_1 - 1) + (n_2 - 1)$

A $t_{\text{megfigyelt}}$ értékét ezután összehasonlítjuk a Student t-táblázat kritikus értékével a megfelelő valószínűsénél és szabadságfoknál. Ha a megfigyelt érték nagyobb vagy egyenlő a kritikus értékkel, akkor a null hipotézis elutasítható a valószínűség adott szintjén. Jegyezzük meg, ha a tesztet kétoldalú tesztként végezzük, akkor a $t_{\text{megfigyelt}}$ előjele (azaz pozitív vagy negatív) figyelmen kívül hagyható. Ha azonban a tesztet egyoldalú tesztként végezzük, akkor fontos, hogy a hipotetikusán kisebb átlagot vonjunk le a nagyobból, amikor kiszámítjuk a $t_{\text{megfigyelt}}$ értékét és a $t_{\text{megfigyelt}}$ előjelét ezt követően vegyük figyelembe, amikor összehasonlítjuk a kritikus egyoldalú t-értékkel a Student t-táblázatból.

Példa. Student t-teszt két kis minta átlaga közötti különbség analizésére (feltételezve egyenlő varianciát)

(a) Két-oldalú teszt annak meghatározására, hogy szignifikánsan különbözik-e a növények kis mintájának a növekedési rátája, melyeket egy új műtrágyával, illetve egy standard műtrágyával kezeltek.

Tenyészedényben tesztelt paradicsom növényeket random helyeztek el egy üvegházban és kezeltek vagy egy új „Extragrow” műtrágyával vagy egy standard műtrágyával (kontroll). Meghatározták a növekedési rátáját mind az „Extragrow” műtrágyával kezelt és mind a kontroll műtrágyával kezelt növények random kiválasztott kisebb mintáinak ($n = 30$). (Jegyezzük meg, hogy a mintaméreteknek nem szükségszerűen kell egyenlőnek lenni ehhez a teszthez.) A következő adatokat és minta statisztikákat kapták (Ireland, 2010):

30 növminta növekedési rátája (g szárazanyag nap ⁻¹)					
„Extragrow” műtrágya			Standard műtrágya		
1.9	1.7	2.0	0.8	1.3	0.9
2.3	2.2	1.5	1.1	1.4	0.4
1.0	0.8	1.6	0.7	1.8	0.8
1.0	1.1	0.7	0.8	0.9	1.3
1.8	1.2	0.7	1.2	1.1	1.0
1.1	0.9	0.9	1.7	1.6	1.4
1.8	1.5	1.3	1.3	1.0	0.4
2.2	0.7	1.0	1.0	1.2	1.3
1.2	0.9	0.8	1.5	0.7	1.2
1.7	1.1	1.0	0.5	0.5	1.3
Minta A: „Extragrow” műtrágya	$\bar{x} = 1.32$ $s = 0.49$ $n = 30$		Minta B: Standard műtrágya	$\bar{x} = 1.07$ $s = 0.37$ $n = 30$	

Null hipotézis: Az „Extragrow” műtrágyával és a standard műtrágyával kezelt növények átlagos növekedési rátája ugyanaz ($H_0: \mu_1 = \mu_2$).

Először kiszámítjuk mindkét minta összevont varianciáját,

$$s_p^2 = \frac{s_a^2(n-1) + s_b^2(n_2-1)}{(n_a-1) + (n_b-1)} = \frac{0.49^2(30-1) + 0.37^2(30-1)}{(30-1) + (30-1)} = 0.19$$

Ezt követően $t_{\text{megfigyelt}}$ statisztikai mutatót számítjuk ki:

$$t_{\text{megfigyelt}} = \frac{[\bar{x}_a - \bar{x}_b]}{\sqrt{s_p^2 \times \left(\frac{1}{n} + \frac{1}{n_b}\right)}} = \frac{[1.32 - 1.07]}{\sqrt{0.19 \times \left(\frac{1}{30} + \frac{1}{30}\right)}} = 2.22$$

A Student t-táblázatból (Melléklet 1. táblázat) a kétoldalú kritikus t-érték $\alpha = 0.05$ -nél és d.f. = 58-nál 2.002. A t megfigyelt értéke nagyobb, mint 2.002; ezáltal annak valószínűsége, hogy olyan megfigyelt t-értéket kapunk, amikor a null hipotézis igaz, kisebb mint 5%. A H_0 -t ezáltal visszautasítjuk.

Következtetés: szignifikáns különbség van az „Extragrow” műtrágyával és a standard műtrágyával kezelt növények növekedési rátájában ($P < 0.05$).

(b) Egyoldalú teszt annak megállapítására, hogy vajon a növények kis mintájának, melyeket az „Extragrow” új műtrágyával kezelték, szignifikánsan nagyobb-e a növekedési rátájuk, mint a növények kis mintájának, melyeket standard műtrágyával kezelték (ugyanazok az adatok, mint fent az (a)-ban).

Meghatározták az „Extragrow” műtrágyával kezelt és a standard műtrágyával kezelt növények random kiválasztott kis mintájának ($n=30$) növekedési rátáját. Az alábbi adatokat és minta statisztikákat kapták:

Minta A: $\bar{x} = 1.32$
„Extragrow” $s = 0.49$
műtrágya $n = 30$

Minta B: $\bar{x} = 1.07$
Standard $s = 0.37$
műtrágya $n = 30$

Null hipotézis: az „Extragrow” műtrágyával kezelt növények átlagos növekedési rátája nem nagyobb, mint a standard műtrágyával kezelt növények átlagos növekedési rátája ($H_0: \mu_1 \leq \mu_2$).

A $t_{\text{megfigyelt}}$ statisztikai mutatót pontosan úgy számítjuk ki, mint fent (a) kétoldalú tesztnél, jóllehet a mutató előjelét meg kell jegyezni:

$$t_{\text{megfigyelt}} = + 2.22$$

A Student t-táblázatból (Melléklet 1. táblázat) az egyoldalú kritikus t-érték $\alpha = 0.05$ -nél és DF = 58-nál 1.672 (azaz egyenlő a kétoldalú t-teszttel $\alpha = 0.10$ -nél). Figyelembe véve az előjelet, a $t_{\text{megfigyelt}}$ értéke nagyobb, mint 1.672; ezáltal annak valószínűsége, hogy a $t_{\text{megfigyelt}}$ értékét megkapjuk valós null hipotézis esetén, kisebb, mint 5%. A H_0 -t ezáltal visszautasítjuk.

Következtetés: az „Extragrow” műtrágyával kezelt paradicsom növények szignifikánsan nagyobb növekedési rátával rendelkeznek, mint a standard műtrágyával kezelt növények ($P < 0.05$).

8.3.2. Kétmintás t-próba, feltételezve nem egyenlő populáció varianciát

A homogén populáció variancia feltételezése nem mindig érvényesül. Gyakran feltételezzük, hogy amikor egy kezelést alkalmazunk, az hatással lesz az adatok centrális tendenciájára, anélkül, hogy hatással lenne az adatok uniformitására. Természetesen nem mindig ez az eset és a kezelés számos típusa ténylegesen hatással van a kezelt populációk szóródására és ezáltal varianciájára. Továbbá, két populáció variancia ugyancsak különböző lehet egyszerűen a jelen levő random variáció különböző nagyságának tulajdoníthatóan. Ilyen esetekben nem használhatunk összevont varianciát a két mintából egyetlen homogén populáció variancia (σ^2) becsléseként. Ehelyett a két minta variancia mindegyike megadja a két eltérő populáció variancia független becslését és mindkettőt ezáltal közvetlenül bevonjuk az átlagok közötti különbség standard hibájának (SED) meghatározásába. Ezáltal, ahol a varianciák nem egyenlők, a $t_{\text{megfigyelt}}$ értékét a következő módon számítjuk ki:

$$t_{\text{megfigyelt}} = \frac{(\bar{x}_1 - \bar{x}_2)}{\text{SED}} = \frac{(\bar{x}_1 - \bar{x}_2)}{\sqrt{\left(\frac{s_1^2}{n_1}\right) + \left(\frac{s_2^2}{n_2}\right)}}$$

ahol: \bar{x}_1, \bar{x}_2 = az 1. és 2. minta átlaga;

s_1^2, s_2^2 = az 1. és 2. minta varianciája;

n_1, n_2 = az 1. és 2. minta mérete.

Míg ez a $t_{\text{megfigyelt}}$ viszonylag egyszerű kiszámításának tűnik, van egy probléma, amennyiben az ily módon kiszámított $t_{\text{megfigyelt}}$ értéke nem felel meg tökéletesen a Student t-eloszlásnak. Ennek eredményeként növeli az 1. típusú statisztikai hiba elkövetésének valószínűségét, azaz a null hipotézis visszautasítását, amikor az ténylegesen igaz. Nagy mintákra ($n \geq 30$) a hiba elég kicsi és figyelmen kívül hagyható, azonban ahogy a minták kisebbé válnak, a hiba észlelhető lesz. Ennek korrigálásához a szabadságfokok elég bonyolult módosítását kell elvégezni, amely visszaállítja a $t_{\text{megfigyelt}}$ érté-

két a Student t-eloszláshoz. Ennek eléréséhez a szabadságfokokat a következő egyenlettel becsüljük:

$$d.f. = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2 - 1}}$$

A fenti egyenlet (amely napjainkban könnyen meghatározható a legtöbb statisztikai számítógépes programmal) rendszerint nem egész számú szabadságfokot eredményez, amely esetben a legközelebbi legkisebb egész számot használjuk. Mihelyt a $t_{\text{megfigyelt}}$ és a d.f. meghatározása megtörtént, a $t_{\text{megfigyelt}}$ értékét összehasonlítjuk a kritikus Student t-táblázati értékkel, hogy becsüljük a null hipotézis korrektségének valószínűségét és a tesztet ugyanúgy folytatjuk, mint a varianciák egyenlősége esetén már leírtuk.

8.4. Páros t-próba

Eddig két független, véletlenül kiválasztott minta összehasonlítását vitattuk meg. Egészen általános azonban, hogy két olyan mérés-sorozat között végzünk összehasonlítást, amelyek az egyedek azonos mintájából származnak. Ebben az esetben az adatok két sorozata nem lehet független és „párosnak” írjuk le. Ha feltételezhetjük, hogy a minták normális eloszlású populációkból származnak, akkor a szignifikáns különbségeket a páros minták között vizsgálhatjuk a páros mintás t-teszttel.

Páros minták kétféle módon képződnek. Az első esetben, a mérések két elkülönített sorozatát elvégezhetjük ugyanazon a változón, különböző időben. Például, tíz alma respirációs (légzési) rátáját a tárházban mérhetjük az 1. hét végén és ugyanennek a tíz almának a respirációs rátáját mérhetjük a 2. hét végén. Ebben az esetben azok az alma egyedek, amelyeknek legnagyobb volt a respirációs rátája az 1. hét végén, nagy valószínűséggel a legnagyobb respirációs rátát mutatják a 2. hét végén, míg a legalacsonyabb respirációs rátával rendelkező egyedek az 1. hét végén, nagy valószínűséggel a legalacsonyabb respirációs rátát mutatják a 2. héten is. Világos ezáltal, hogy a respirációs adatok két sorozata nem tekinthető függetlennek. A másik mód, ahogy páros minták gyakran keletkeznek, ahol az adatok egyetlen mintája az egyidejű mérés két sorozatát adja. Az ilyen adatokat gyakran kétváltozósoknak hívjuk. Például, tíz levél minta felső részén (színén) a levélszőrök száma és ugyanannak a tíz levélnek a fonákján a levélszőrök száma egyidejűleg feljegyezhető. Újból, a mérések két sorozata nem kezelhető függetlenként, hanem kapcsolatban vannak, ugyanabból a levél-mintából származnak.

A két mintás null hipotézis a páros mintás t-tesztre az, hogy a populációk átlaga, amelyekből a minták származnak, egyenlő. A teszt első lépése, hogy meghatározzuk minden méréspár között a számszerű különbséget, például a respirációs ráta között az 1. héten és a 2. héten minden almára a tároló helyiségben. Ha a mérések két sorozatának főátlaga teljesen azonos, akkor a különbségek átlagának minden adatpár között zérónak kell lenni. Minél jobban eltér a különbségek átlaga zérótól, annál nagyobb a

valószínűsége, hogy szignifikáns különbség van két minta között. Ezáltal, mindaddig, amíg az adatok normális eloszlású populációkból származnak, a Student t-tesztet használhatjuk annak vizsgálatára, hogy a különbségek átlaga a mérés-párok között (\bar{d}) szignifikánsan különbözik-e a zero értéktől. A teszt ezáltal azonossá válik a Student t-teszttel, vagyis a minta átlagnak egy hipotetikus átlaggal (μ_0) történő összehasonlításával. Először kiszámítjuk az adatpárok különbségét, majd a mérés-párok közötti

különbségek átlagát (\bar{d}):
$$\bar{d} = \frac{\sum_{i=1}^n d_i}{n}$$

Mivel ebben az esetben μ_0 zero, a teszt mutató kiszámításának egyenlete a következő:

$$t_{\text{megfigyelt}} = \frac{\bar{d}}{s/\sqrt{n}}$$

ahol: \bar{d} = a páros mérések közötti különbségek átlaga;

μ = elméleti átlag (azaz zero);

s = a különbség értékek szórása;

n = az adatok száma.

A fenti egyenlet átalakítható egy alternatív formává, melynek könnyebb a használata a gyakorlatban:

$$t_{\text{megfigyelt}} = \frac{\sum d}{\sqrt{\frac{n \sum d^2 - (\sum d)^2}{n-1}}}$$

ahol: $\sum d$ = a mérések mindenegybes párja közötti különbségek összege.

A szabadságfokok száma a tesztre $n - 1$, ahol n az adatpárok száma, azaz a különbség értékek száma. A t megfigyelt értékét összehasonlítjuk a kritikus t értékkel, amely várható lenne a null hipotézis esetén. A kritikus t értéket a Student t -táblázatból kapjuk, az α valószínűség kívánt szintjén és megfelelő szabadságfoknál. Ha a megfigyelt érték nagyobb vagy egyenlő a kritikus értékkel, akkor a null hipotézis elutasítható, és szignifikáns különbség állapítható meg a páros minták között az adott valószínűségi szinten. Az alábbi példa bemutatja a páros t -próba kiszámítását.

Példa. Páros t -próba két nem-független minta közötti különbség szignifikanciájának vizsgálatára

Egy teszt annak meghatározására, hogy van-e időbeni változás a tárolt alma minták respirációs rátájában

Mind a tíz címkézett alma respirációs rátáját mérték, amikor szeptemberben tárházba helyezték és ugyanennek a tíz almának újramérték a respirációs rátáját 4 héttel később, októberben. Az eredmények az alábbiak:

Alma sorszáma	Respirációs ráta szeptemberben (mmol O ₂ /g/h)	Respirációs ráta októberben (mmol O ₂ /g/h)	Különbség (d)	d ²
1	10.3	12.2	+1.9	3.61
2	11.4	12.1	+0.7	0.49
3	10.9	13.1	+2.2	4.84
4	12.0	11.9	-0.1	0.1
5	10.0	12.0	+2.0	4.00
6	11.9	12.9	+1.0	1.00
7	12.2	11.4	-0.8	0.64
8	12.3	12.1	-0.2	0.04
9	11.7	13.5	+1.8	3.24
10	12.0	12.3	+0.3	0.09
n = 10	$\bar{x}_1 = 11.47$	$\bar{x}_2 = 12.35$	$\sum d = 8.8$ $(\sum d)^2 = 77.44$	$\sum d^2 = 17.96$

Különbségek átlaga, $\bar{d} = 0.88$

Különbségek szórása, $s = 1.065$

Null hipotézis: az eltárolt almák átlagos respirációs rátája szeptemberben és októberben ugyanaz ($H_0: \mu_1 = \mu_2$).

$$t_{\text{megfigyelt}} = \frac{\bar{d}}{s/\sqrt{n}} = \frac{0.88}{1.065/\sqrt{10}} = 2.613$$

Alternatív számítás:

$$t_{\text{megfigyelt}} = \frac{\sum d}{\sqrt{\frac{n \sum d^2 - (\sum d)^2}{n-1}}} = \frac{8.8}{\sqrt{\frac{10 \times 17.96 - 77.44}{10-1}}} = 2.613$$

A Student t táblázatból a kritikus érték $\alpha = 0.05$ -nél és d.f. = 9 értéknél 2.262. Mivel $t_{\text{megfigyelt}}$ nagyobb, mint 2.262, a null hipotézist visszautasítjuk.

Következtetés: szignifikáns különbség van az almák szeptemberben és októberben mért respirációs rátája között ($P < 0.05$).

8.5. Két minta varianciája közötti különbség analízise: F-próba

A 8.3. fejezetben rámutattunk arra, hogy a Student t-teszt számítása függ attól, hogy a két populáció varianciája, amelyből a minták származnak, statisztikailag egyenlő-e. Nagyon gyakran a kísérleti tervezés megtekintése és korábbi vizsgálatok tapasztalata lehetővé teszi, hogy megfelelő véleményt alkossunk a populáció varianciák valószínű homogenitásáról. Ha azonban nem biztos, hogy a populáció varianciák statisztikai egyenlőségének feltétele fennáll, akkor a feltételt formálisan tesztelni szükséges. Továbbá, bizonyos kísérletekben a kezelés hatása a kísérleti anyag varianciájára esetenként önmagában az elsődleges vizsgálati cél. Például, egy kezelést alkalmaztak egy vetőmag tételnél azzal a szándékkal, hogy javítsák a csírázás egyöntetűségét, inkább,

mint a csírázás mennyiségét. Ebben az esetben a csírázási ráta varianciájában levő különbség vizsgálata a vetőmag tételek között sokkal fontosabb lesz, mint az átlagok közötti különbség vizsgálata.

A legegyszerűbb rendelkezésre álló teszt két variancia közötti különbség vizsgálatára az *F arány teszt (F-próba)*. Újból azonban valószínűtlen, hogy ismernénk a populáció varianciákat, úgy hogy a tesztet a mint varianciákkal végezzük, feltételezve, hogy a minta statisztikák megfelelő becslését adják a populációk paramétereinek, amelyekből a minták származnak. Ha két mintának hasonló a varianciája, akkor varianciájuk hányadosa (aránya) közel lesz az 1-hez; minél nagyobb a különbség a varianciák között, annál távolabbra kerül ez az arány az egységnyitől. Az egyik minta variancia aránya a másikhoz egy statisztikai mutatót ad, amelyet F-arálynak hívunk. Mielőtt meghatároztuk az F-arány értékét, tesztelhetjük a kritikus F-értékkel szemben, amelyet az F-táblázatból keressük ki, annak becslése céljából, hogy a két variancia statisztikailag különbözik-e.

Ebben az esetben az F-arányt úgy határozzuk meg, hogy elosztjuk a nagyobb varianciát a kisebbel, azaz:

$$F\text{-arány} = s_a^2 / s_b^2$$

ahol: s_a^2 = a nagyobb varianciával rendelkező minta varianciája;
 s_b^2 = a kisebb varianciával rendelkező minta varianciája.

A tesztet kétoldalú próbaként végezzük és ezáltal a kétoldalú kritikus F-érték táblázatot kell használnunk (*Melléklet 3. táblázat*) a szükséges valószínűségnél (rendszerint $\alpha=0.05$) és mindkét minta szabadságfokánál, ahol v_1 = a nagyobb varianciával rendelkező minta d.f.-je, v_2 = a kisebb varianciával rendelkező minta d.f.-je. A null hipotézis erre a tesztre az, hogy nincs különbség a populáció varianciák között ($H_0: \sigma_1^2 = \sigma_2^2$). Ha a megfigyelt F-érték $\geq F_{\text{kritikus}}$, akkor a null hipotézist vissza kell utasítani és a varianciák közötti különbséget kell feltételezni. Ha, viszont, az F-arány $< F_{\text{kritikus}}$, akkor a két varianciát statisztikailag egyformának kell tekinteni.

Példa: F-arány teszt annak megállapítására, hogy két minta variancia statisztikailag homogén-e.

Kétoldalú teszt annak meghatározására, hogy az „Extragrow” új műtrágyával és a standard műtrágyával kezelt paradicsom növények növekedési rátájának varianciája szignifikánsan különbözik-e (a példa adatai a 8.3.1. fejezetben).

Minta A:	$\bar{x} = 1.32$	Minta B	$\bar{x} = 1.07$
„Extragrow”	$s = 0.49$	Standard	$s = 0.37$
Műtrágya	$s^2 = 0.242$	Műtrágya	$s^2 = 0.137$
	$n = 30$		$n = 30$

Null hipotézis: az Extragrow és a standard műtrágyával kezelt növények varianciája egyenlő ($H_0: \sigma_1^2 = \sigma_2^2$).

$$F_{\text{arány}} = s_{\text{nagyobb}}^2 / s_{\text{kisebb}}^2 = 0.242 / 0.137 = 1.77$$

A kétoldalú kritikus F-érték (Melléklet 3. táblázat) $\alpha = 0.05$ -nél és $v_1 = 29$, $v_2 = 29$ szabadságfoknál: 2.10. Mivel a megfigyelt F-érték $< F$ kritikus, a null hipotézist el kell fogadni.

Következtetés: nincs szignifikáns különbség az Extragrow és a standard műtrágyával kezelt növények növekedési rátájának varianciája között. A minta átlagok közötti különbség vizsgálható a Student t-teszttel, feltételezve egyenlő populáció varianciát.

8.6. Két-mintás kísérletek tervezése és analízise

Ez a fejezet rámutatott arra, hogy viszonylag egyszerű az olyan kísérletek tervezése, amelyek magukba foglalják két kezelés hatásának összehasonlítását, vagy egyetlen kezelés összehasonlítását egy kontroll kezeléssel. Mégis számos fontos szempont figyelembevétele indokolt. Az adatanalízis kiválasztott módszere függ több tervezési kritériumtól és feltételezéstől a begyűjtött adatokra vonatkozóan. Különösen, tudni kell, hogy a minták normális eloszlású populációkból származnak-e és a populáció varianciákról feltételezhető-e, hogy statisztikailag egyenlők. Ha e feltételek egyikének sem felelnek meg, akkor egy alternatív nem-paraméteres módszerre van szükség. A tervezésben a fő kérdés az, hogy a minták függetlenek vagy nem függetlenek, azaz vajon a két adatminta az egyedek különböző sorozatából vagy az egyedek ugyanazon sorozatából származik-e. Rendkívül fontos, hogy ezeket a kérdéseket megválaszoljuk a kísérlet végzését megelőzően, sokkal inkább, mint a kísérlet befejezésekor, amikor túl késő lehet biztosítani, hogy rendelkezésre álljon az adatanalízishez egy érvényes módszer.

Egy bonyolultabb kérdés a két-mintás kísérletek tervezésére vonatkozóan akkor merül fel, amikor alkalmilag egy valódi választás áll rendelkezésre egy független tervezés vagy egy páros-mintás tervezés használata között. Számos bizonyíték van arra, hogy ahol lehetséges, a páros minta megközelítést kellene előnyben részesíteni, mivel a páros t-próba statisztikailag erősebb és nagyobb esély van kis különbségek szignifikanciájának identifikálására. Ennek oka egészen világos. Amikor két független mintát használunk, mindegyik tartalmaz kis mennyiségű random mintán belüli variációt, amely nem szükségszerűen hasonló, és minthogy mindkét minta variancia a teszt statisztika (mutató) kiszámításának komponense, hatásuk az analízisre összeadódik. A páros minták esetén, mivel a két adatsor az egyedek ugyanazon sorozatából származik, a mintán belüli variáció valószínűleg kevésbé változik a minták között és kioltják egymást, amikor az adatpárok közötti különbségeket számítjuk. Következésképpen, a páros t-teszt érzékenyebb lesz a kis különbségekre és képes lesz jobban feltárni szignifikáns különbségeket, amikor a minta méretek kicsik, mint a független mintás t-teszt. Másrészt, ha hibás mintát választottunk ki véletlenül, akkor ennek hatása összekeveredne a páros minta tervezés esetén, mivel ugyanazt a mintát használjuk mindkét adatsorban, melyeket összehasonlítottunk. Általánosságban, amit megvitattunk itt az, hogy amikor ugyanahhoz a kísérleti tervhez alkalmazzuk, egy független mintás t-tesztnél nagyobb esélye van 1. típusú statisztikai hibát elkövetni (a null hipotézis visszautasítása, amikor az igaz), azonban kisebb az esélye a 2. típusú statisztikai hiba előidézésének (a null hipotézis elfogadása, amikor az hibás), mint a páros t-tesztnél. A vizsgálatot végzőnek kell ezáltal eldöntenie, melyik a legmegfelelőbb megközelí-

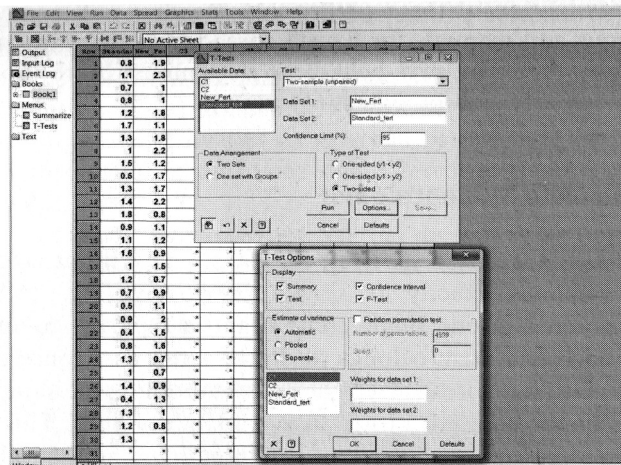
tés, a vizsgált kísérlet jellege és a feltett kutatási kérdés alapján. Rá kell mutatni arra is, hogy a szoftverek ma már feleslegessé teszik a statisztikai táblázatok használatát.

Jóllehet elég általános, a két mintás kísérleti elrendezés viszonylag egyszerű és csak limitált információt ad a kísérletről. A mezőgazdasági és kertészeti kutatásban, különösen a szántóföldi kísérletek esetén sokkal általánosabb kettőnél több kezelés egyidejű alkalmazása, amely elvezet a több kezeléssel kísérlet tervezéshez. A kettőnél több kezelést tartalmazó kísérletek analízise egy eltérő megközelítést, a varianciaanalízist igényli és ennek módszereit a következő fejezetekben mutatjuk be.

8.7. A Student t-próba kiszámítása a GenStat felhasználásával

A Student t-teszt elvégzéséhez, hogy összehasonlítsunk két mintát, az adatokat be kell vinni a GenStat számolótáblába, két oszlopba. Alternatív módon, az adatok bevihetők egyetlen oszlopba és egy külön oszlopot használunk annak azonosítására, hogy az adatok melyik mintához tartoznak. Ezt az oszlopot faktorként azonosítjuk.

A Student t-próba futtatásához a *Statistical Tests* beállítást választjuk a *Stats* legördülő menüből, ezután a *One- and two-sample t-tests* beállítást választjuk, hogy megjelenjen a *T-Tests* beviteli doboz. E dobozon belül a t-próbával összehasonlítandó két változót kiválasztjuk és választunk az egyoldalas (one-tailed test) és kétoldalas (two-tailed test) próba között (ez utóbbi az alap-beállítás). Ha az adatokat egyetlen oszlopba és a csoportképző faktort külön oszlopba vittük be, akkor ezt jelezni kell a *Data Arrangement* dobozban. Egy *Options* doboz további analízist kínál, ha szükséges, beleértve egy *F-próbát*, hogy becsüljük a minta varianciák homogenitását és meghatározzuk a mintaátlagok közötti különbség konfidencia intervallumát. Megjelenít továbbá egy választást a variancia számításának módjára a t-próbán belül, amely lehet összevont vagy meghatározható külön-külön mindegyik mintára. (Ez függ attól, hogy a minta varianciák homogének-e.) Végül, rákattintva a *Run*-ra, a t-próba végeredményét kapjuk, amely megjelenik a kimeneti ablakban. A GenStat számolótáblát és a kimeneti ablakot a Student t-próbára a 8.4. ábra mutatja.



8.4. ábra. GenStat számolótábla és menü dobozok a független minták Student t-próbájához

Figyeljük meg, hogy a 8.4. ábrán bemutatott példában egy variancia összehasonlítási F-próbát futtattunk, amely kimutatta, hogy a minta varianciák szignifikánsan nem különböztek ($P = 0.14$). A GenStat ezt követően meghatározta a Student t-próbát, feltételezve az egyenlő varianciát. A különbség a minták között szignifikáns volt ($P = 0.030$).

Kétoldalú t-teszt annak meghatározására, hogy szignifikánsan különbözik-e a növények növekedési rátája, melyeket az Extragrow vagy a Standard műtrágyával kezeltek

Two-sample t-test

Variates: Extragrow, Standard.

Test for equality of sample variances

Test statistic $F = 1.73$ on 29 and 29 d.f.

Probability (under null hypothesis of equal variances) = 0.14

Summary

Sample	Size	Mean	Variance	Standard deviation	Standard error of mean
Extragrow	30	1.320	0.2417	0.4916	0.08975
Standard	30	1.070	0.1394	0.3734	0.06817

Difference of means: 0.250

Standard error of difference: 0.113

95% confidence interval for difference in means: (0.02440, 0.4756)

Test of null hypothesis that mean of Extragrow is equal to mean of Standard

Test statistic $t = 2.22$ on 58 d.f.

Probability = 0.030

Páros t-próba annak meghatározására, hogy van-e időbeni változás a tárolt alma minták respirációs rátájában

***** Two-sample T-test (paired) *****

Calculated using one-sample t-test with the null hypothesis that the mean of RespirOctob - RespirSept is equal to 0

One-sample t-test

Variate: Y[1].

Summary

Sample	Size	Mean	Variance	Standard deviation	Standard error of mean
Respir Octob–Respir Sept	10	0.8800	1.135	1.065	0.3369

95% confidence interval for mean: (0.1178, 1.642)

Test of null hypothesis that mean of RespirOctob-RespirSept is equal to 0

Test statistic $t = 2.61$ on 9 d.f.

Probability = 0.028

9. A varianciaanalízis (ANOVA) alapelvei

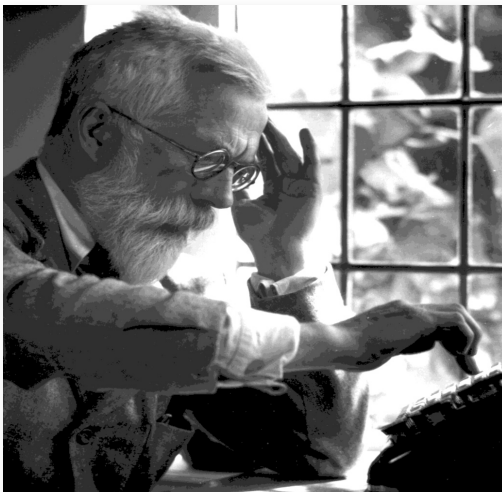
9.1. A varianciaanalízis történeti háttere

A korai szántóföldi növénytermesztési és növénynemesítési kutatás elsődlegesen a termés (hozam) kísérletekre fókuszált. Az általános cél a kezelések (talajkezelések és növényfajták) hatásainak és interakcióinak alapvető tudományos megértése volt. Mérték a növények termését és kezeléssátlagokat számoltak ki, hogy tanácsot adjanak a termelőknek a növényi produkció maximalizálására. A 20. század kezdetén az agronómia tudományos diszciplínává kezdett fejlődni, és az agronómusok fokozott figyelmet fordítottak javaslataik hihetőségére, felismerve, hogy a parcellák átlagtermései változnak olyan faktoroknak tulajdoníthatóan, melyek szabályozási lehetőségeiken kívül vannak.

A kísérlettervezésnek ebben a korai időszakában, a közös kutatási cél volt, hogy alkossanak egy statisztikai mutatót, amely a megbízhatóság mérését adja a kezeléssátlagok megbízhatóságában és megismételhetőségében. Ha az egyik kezelés átlagtermése nagyobb volt, mint a kontroll, a kutatók tudni akarták, hogy ez a különbség „szignifikáns” volt-e. Az ANOVA előtt a kezeléseket a parcellákon megismételték, hogy kiszámítsák a szórást (SD). A kezelt parcellák ismétlései azonban szisztematikusan (nem randomizáltak) voltak elrendezve, mivel tévesen úgy gondolták, hogy összehasonlítva a kezelt parcellákat a szomszédos kontroll parcellákkal, jobban ki tudják mutatni a szignifikáns kezeléskülönbségeket. Az agronómusok ugyancsak aggodtak amiatt, hogy a kezelt és kontroll parcellák közötti szignifikáns különbségek

gyakran nem voltak következetesek az évek és a helyek között. A fajták relatív termésének variációi az évek között továbbra is egy statisztikai dilemma maradt, mivel az évek random hatása nem ismételhető.

Az 1920-as években *Fisher, R.A.* (1890–1962), brit statisztikus fejlesztette ki az ANOVA-t, megvizsgálva a rothamstedi búza kísérleti parcellák termés variációit, amelyek folyamatos talajkezeléseket kaptak, több mint félévszázadon keresztül. *Fisher* az ANOVA-t használta, hogy elkülönítse a nagy változásokat az átlagtermésben más változásoktól és meghatározza a szignifikáns különbségeket a kezelések között.



Fisher, R. A. (1890–1962)

Az ANOVA tehát egy statisztikai módszer, amely felosztja az összes variációt a szántóföldi parcellák között, elkülönült és független „variáció forrásokra”.

Ahhoz, hogy az ANOVA korrekten funkcionáljon, a kezelés parcellákat (kísérleti egységek) meg kell ismételni és randomizálni szükséges. Fisher radikálisan megváltoztatta a kísérlettervezési gyakorlatot, amikor demonstrálta, hogy a kezelések szisztematikus elrendezése a kezelésátlagok torzított becslését, a hiba variáció túlbecslését és alulbecslését, valamint korreláló hibákat eredményezett. Tehát az ANOVA szűkességessé teszi a kezelések ismétlését a kísérleti hiba (a kezeléseken belüli variáció) becsléséhez és a randomizációt ahhoz, hogy a szignifikancia választott szintje vagy a valószínűség korrekt legyen.

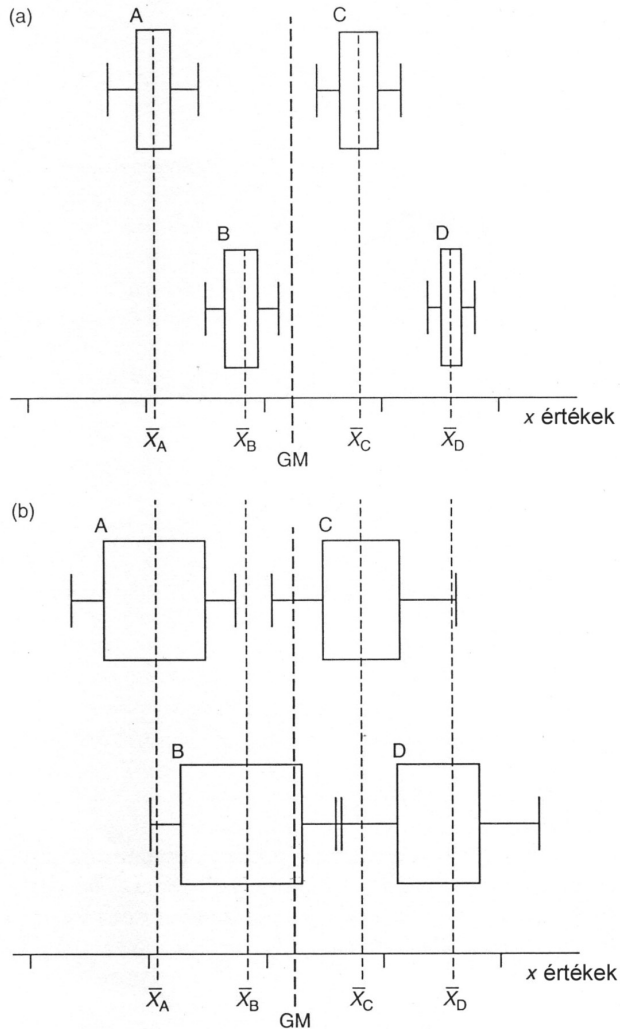
A legegyszerűbb esetben, az ANOVA felosztja a parcellák közötti összes variációt két független forrásra: variáció a kezeléseket között és variáció a kezeléseken belül (hívják kísérleti hibának, maradék hibának, hiba közepes négyzetes eltérésnek és hiba varianciának). Feltételezve, hogy az adatok normális és független eloszlásúak, az F-arányt, amely összehasonlítja a kezeléseket közötti variációt a kezeléseken belüli variációval, használjuk fel arra, hogy teszteljük a null hipotézist, mely szerint a kezelésátlagok egyenlők. Minél nagyobb az F-arány, annál kisebb a valószínűsége annak, hogy a kezelésátlagok egyenlők, és az átlagokat szignifikánsan különbözónak tekintjük, amikor egy nagyobb F-érték véletlen bekövetkezésének valószínűsége azonos vagy kisebb a szignifikancia szintnél.

A szignifikáns kezelés-hatások 1 főlé növelik a minták (kezelések) közötti variancia arányát, a mintán (kezelésen) belüli varianciához viszonyítva. Ezáltal, minél nagyobb lesz a minták közötti variancia aránya a mintán belüli varianciához viszonyítva, annál nagyobb lesz az összes variancia aránya a kísérletben, amely a kezelés-hatásnak tulajdonítható.

Az ANOVA alapelvét mutatja be a 9.1. ábra, ahol a doboz ábrát (kvartilis diagramot) használtuk négy minta reprezentálására, amely az első esetben szignifikáns különbségeket mutat, míg a második esetben nem mutat szignifikáns különbségeket. Mindegyik esetben minden ekvivalens mintának ugyanaz a középértéke; azonban az első esetben, ahol a különbségek szignifikánsak, a mintáknak alacsony a mintán belüli varianciája, amely egy nagy F-értéket eredményez. A második esetben, ahol a különbségek nem szignifikánsak, a mintáknak viszonylag magas a mintán belüli varianciája, amely alacsony F-értéket ad.

A varianciaanalízist először a fix hatású modellekre (Modell I.) vezették be, azonban hamarosan kiterjesztették a random hatású modellekre (Modell II.) (Fisher, 1925). Az ANOVA-t kezdetben a fix hatásokra alkalmazták, hogy teszteljék a szignifikáns különbségeket a kiválasztott kezelésátlagok között, míg a random hatású modellt arra használták, hogy becsüljék a random faktorokkal összefüggő variációt, mint amilyenek az évek. A fix hatású modellek olyan kezeléseket tartalmaznak, amelyeket azért választottunk ki, hogy reprezentáljanak specifikus kezeléseket vagy a kezelések szintjeit. Tehát, a fix hatásokat használjuk, hogy szűk körű következtetéseket vonjunk le és teszteljük a kezelésátlagok közötti különbség hiányának hipotézisét. Ezzel szemben, a random hatásokat arra használjuk, hogy széles körű következtetéseket vonjunk le az egész populációról, beleértve a populáció szintjeit, amelyek nincsenek bevonva a kísérletbe. A random hatások szintjeiről feltételezzük, hogy egy nagy populáció random mintái és reprezentálják a populáción belüli variációt. Például, ha a

helyek képeznek egy random faktort, akkor a kísérleti helyeket random választjuk ki, hogy a helyek populációjának random mintáját reprezentálják. Eisenhart (1947) javasolta, hogy a varianciaanalízist alkalmazni lehetne a vegyes (kevert) hatású modellekre, amelyek fix és random kezelés faktorokat egyaránt tartalmaznak. Napjainkban a számítógépek teljesítményének növekedése felgyorsította az előrehaladást a vegyes modellek alkalmazásában (Gbur et al., 2012).



9.1. ábra. Összefüggés a mintán belüli variabilitás, az F-arány és a szignifikáns különbség megléte között egy egytényezős ANOVA-ban. A doboz ábrák mutatják négy minta (A-D) eloszlását a főátlag (GM) körül két helyzetben, amelyekben (a) van szignifikáns különbség és (b) nincs szignifikáns különbség a minta átlagok között. Az (a) helyzetben alacsony mintán belüli variancia van, összehasonlítva a minták közötti varianciával, amely nagy F-arányt idéz elő és szignifikáns különbség megléte következtetünk. A (b) helyzetben magas mintán belüli variancia van a minták közötti varianciával összehasonlítva, amely alacsony F-arányt idéz elő és nem szignifikáns különbségre lehet következtetni (Ireland, 2010).

A varianciaanalízist leggyakrabban annak meghatározására használjuk, hogy az átlagok közötti különbségek nagyobbak-e, mint ami várható lenne a random variációnak tulajdoníthatóan. Minél kisebb annak valószínűsége, hogy a különbségek kizárólag random variációnak tulajdoníthatók, annál valószínűbb, hogy az átlagok szignifikánsan különböznek. Szántóföldi kispácellás kísérletekre, Fisher személyes preferenciája volt, hogy a kezeléskülönbséget akkor tekintjük szignifikánsnak, amikor $P \leq 0.05$ (Fisher, 1926), és gyakorlati célból, klasszikus referenciája, „*Statistical Methods for Research Workers*” csak egy F-táblázatot foglalt magában, amely kritikus F-értéket adott $P \leq 0.05$ -re (Fisher, 1925). Számítógépes szoftver hiányában az F-táblázatokat használjuk, hogy kikeressük a kritikus F-értéket (minimális F-érték, amely szükséges a szignifikáns különbség megállapításához).

Míg Fisher kifejlesztette a valószínűség használatának kereteit, Nyman és Pearson (1933) egy másik megközelítést dolgozott ki a hipotézis tesztelésére, amely egyidejűleg figyelembe veszi a hibák két típusának arányait. Az I. típusú hiba akkor következik be, amikor visszautasítunk egy valós null hipotézist és a II. típusú hiba bekövetkezik, amikor elmulasztunk visszautasítani egy téves null hipotézist. Javasolták, hogy a kísérlet analízise előtt a kutató válassza meg az I. típusú hiba mértékét (α), melyet szignifikancia szintnek hívunk.

A klasszikus ANOVA feltételek a szignifikancia tesztelésére a megfigyelések normális és független eloszlása 0 átlaggal és közös σ_ε^2 varianciával [NID $(0, \sigma_\varepsilon^2)$]. Eltérések ezektől a feltételektől hatással lehetnek az F-értékre és a szignifikancia szintre. A gyakorlatban, ezek a feltételek gyakran nem pontosan állnak fenn, ezért statisztikai eljárásokat (adat-transzformáció, hiányzó parcellaadat becslése) fejlesztettek ki a feltételek megsértésének enyhítésére. Szerencsére az F-tesztek viszonylag robusztusak és kismértékű eltérések az alapfeltételektől kis hatással vannak az ANOVA eredményére.

9.2. A varianciatáblázat szerkezete

A varianciaanalízis egy keretet ad a hipotézis tesztelésére és a hatások varianciájának becslésére egy lineáris modellben. Az ANOVA minden egyes számítási lépésének eredményét gyakran egy komplett variancia táblázatban mutatjuk be. Ez a táblázat magában foglalja a variancia forrásait, a szabadságfokot (df), az eltérésnégyzet összeget (SS) és a közepes négyzetes eltérést (MS) minden hatásra a lineáris modellben, együtt az F-értékkal és a P értékkel, melyeket a hatások szignifikanciájának tesztelésére használunk.

Ahhoz, hogy hatékony legyen, egy ANOVA táblázat szerkezetének vissza kell tükröznie a kutató célját és elég részletesnek kell lenni ahhoz, hogy megalapozza az eredmények interpretációját. Egy alapvető szempont a kutatási célok tekintetében, hogy a hatások a lineáris modellben fixek vagy randomok. Gyakran egészen nyilvánvaló, így a műtrágya kezelések, fajták, hőmérsékletek és herbicid szintek tipikusan fix hatások, míg a hiba variáció a kísérleti és mintavételi tervvel kapcsolatosan random hatás. Ugyanakkor az évek és környezetek kategorizálása fix vagy random hatásnak különösen problematikus. A döntés, hogy egy hatást fixnek vagy randomnak tekintünk, kihat a következtetésekre és az ANOVA számításra, és ezt a statisztikai módszerek leírásának tartalmaznia kell (Piepho et al., 2003).

A **variancia források** az ANOVA táblázat alapját képezik és tömör számszerűsítését adják a kísérlet lineáris modelljében levő tagoknak. Alapvetően, minden variancia forrás független, együttesen az összes varianciát adják a kísérletben. Egy részletes ANOVA táblázat, amely magában foglal minden variancia forrást a lineáris modellben, egy hatékony módja a lineáris modell leírásának. Statisztikai könyvek (Gomez és Gomez, 1984; Mead *et al.*, 2003; Sváb, 1981) megadnak ANOVA táblázatokat, melyek megegyezést mutatnak a variancia források és a lineáris modell tagok között. A leggyakrabban használt számítógépes programok outputjai azonban nem mindig adnak olyan ANOVA táblázatot, amely megegyezik a szakkönyvek ANOVA táblázataival. Ilyen esetekben a legmegfelelőbb ANOVA táblázat megválasztása és a számítógépes output interpretálása komoly kihívást jelenthet a kutatónak.

Egy ANOVA táblázat magában foglalja minden variancia forrás **szabadságfokát (df)**. A df különösen hasznos, amikor komplex kísérlettervezést és analízist ír le. A df ismeretében könnyen meghatározhatjuk az ismétléseknek, faktoroknak, minden egyes faktor szintjeinek, a kísérleti egységeknek és a kísérlet mintavételi egységeinek számát. Ezen kívül a df lehetővé teszi az analízis és a következtetések helytállóságának ellenőrzését.

Az **eltérésnégyzet összegeket (SS)** arra használjuk, hogy felosszuk az összes variációt a variancia források által meghatározott variációkra. Adott, hogy a variancia források egymástól függetlenek, minden egyes variancia forrás összege az összes SS-t adja. Jóllehet az SS-nek elsődleges szerepe van a modell illesztésében és az illesztés jóságának meghatározásában, szerepük a hipotézis tesztelésben egy közbenső lépés az MS számításakor. Tehát az SS-t tipikusan nem tüntetjük fel a végső ANOVA táblázatokban a szignifikancia tesztelésekor.

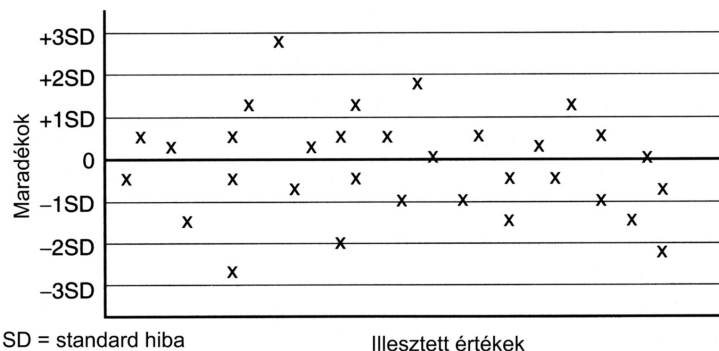
A **közepes négyzetes eltérés (MS)** becsüli az összes variációt, összefüggésben minden variancia forrással és az SS/df hányadosaként számítjuk ki. Az MS becsülésének fő szerepe van a hipotézis vizsgálatban és a variancia komponensek becsülésében. Az MS-t használjuk, hogy összehasonlítsuk a különböző hatásoknak tulajdonítható variációt és magában foglalja az F-érték számlálóját és nevezőjét. Magában foglalja a random hatások MS értékét, amelyek felhasználhatók további F-tesztekhez és az átlag (közép-érték) szórásának (hibájának) (SE) kiszámításához. Az SE-t úgy számítjuk ki, mint $\sqrt{(HibaMS/n)}$, ahol n a megfigyelések száma, melyet az átlag kiszámításához használunk.

Az **F-értékek** (F-mutatók vagy F-arányok) variancia arányok. Az F-arányt annak becsülésére használjuk, hogy a variancia forrás hatásának MS értéke szignifikánsan nagyobb-e a variancia forrás hiba MS értékénél. Elméletileg, ha a null hipotézis valós (a vizsgált hatás zéró), a várható MS az F számlálójában és nevezőjében azonos és $F = 1$. Tehát a null hipotézist visszautasítjuk, ha a számított F-érték felülmúlja a kritikus F-értéket a szignifikancia adott szintjén. A **P-érték** egy összegező statisztikai mutató, egy standard, általános skálát (0.0–1.0) ad, amely alkalmazható az F, t és χ^2 statisztikai mutatók tesztelésére. Az ANOVA táblázatok magukba foglalják a P értékeket vagy *, ** és ***jelöléseket, melyek reprezentálják a szignifikanciát 0.05, 0.01 és 0.001 valószínűségeknél. A P-értékek a végső statisztikai mutatók az ANOVA táblázatban és a szignifikáns P-értéket gyakran az ANOVA végső céljának tekintjük.

9.3. Az ANOVA feltételei

1. A minták legyenek függetlenek.
2. A minták normális eloszlású populációból származzanak.
3. A minta varianciák legyenek homogének.

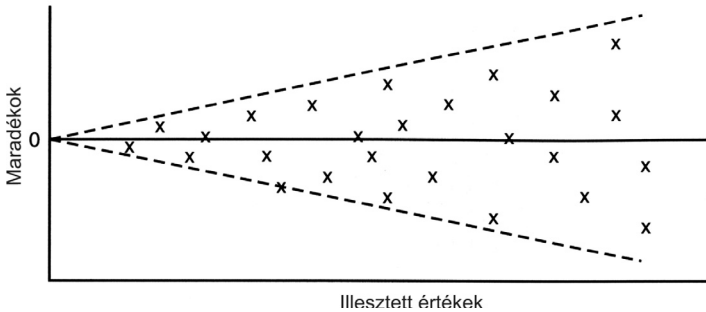
A kölcsönösen független megfigyelések feltétele megköveteli a random mintavételt és hogy az eredmények egy kísérleti egységből nem befolyásolhatják a szomszédos kísérleti egységek (parcellák) eredményeit. Néhány esetben ezt azt jelentheti, hogy a kísérleti egységeket fizikailag szeparálni kell térben vagy időben: például egy szántóföldi kísérletben, ahol a kezelés permetezést jelent, puffer parcellák válhatnak szükségessé a kísérleti egységek között, hogy ezáltal biztosítsuk azt, hogy az egyik kísérleti egységen végzett permetezés ne sodródjon át a következő kísérleti egységre. A feltételnek oly módon is eleget teszünk, hogy biztosítjuk a kezelések randomizáltan kerülnek a kísérleti egységekre, felhasználva egy megfelelő randomizálási eljárást. Ha ezt a feltételt megsértjük, ez nem korrigálható az analízisben; azonban a feltételezéseknek megfelelően a kísérlet, amennyiben követtük a korrekt kísérleti eljárást.



9.2. ábra. Reziduális ábra szórás (SD) határokkal, amely elfogadhatónak tűnik, azaz a pontok véletlen szórása egy horizontális sávban helyezkedik el a zero reziduális érték körül (Collins és Seeney, 1999)

A konstans kezelés variabilitás feltétel annak a biztosítására vonatkozik, hogy a kísérleti egységek a különböző kezelésekre megegyező módon adnak reakciót (reagálnak). Ez lehetséges, hogy nem mindig igaz az adatgyűjtésnél. Lehetséges, hogy a benne rejlő variabilitások a megfigyelések között nőnek, ahogy maguk a megfigyelések is nőnek nagyságukban (terjedelmükben). A nem konstans variancia más okai lehetnek a kiszámíthatatlan (erratikus) reakciók egy kezelésre vagy ferde adat eloszlások. Ennek ellenőrzésére a legegyszerűbb mód, hogy elkészítjük a modell reziduuumok (maradék) ábráját az illesztett értékek szerint. A pontoknak random kell szóródnia az ábrán és semmilyen mintázatnak sem szabad megnyilvánulni. A pontoknak az ábrán egy horizontális sávot kell képezni a zero reziduuum körül, mint ez a 9.2. ábrán látható. Ha a hibák normális eloszlásúak, akkor a legtöbb maradék plusz vagy mínusz egy szórás között helyezkedik el, többségük plusz vagy mínusz két szórás között és nagyon kevés maradék helyezkedik el a plusz vagy mínusz három szórásán kívül (9.2. ábra).

Lehetséges, hogy a pontok csoportosulnak, azonban ilyen esetekben minden csoporton belül a pontok random szóródásúak legyenek és a pontok szóródása a zero reziduum körül konstans legyen minden csoportnál. Például, ha a pontok az ábrán tölcser alakúak, mint a 9.3. ábrán látható, akkor a konstans kezelés variancia feltétel nem teljesül. Ennek a feltételnek a megsértése feloldható az adatok transzformálásával és az újbóli analízisével. Vannak formális statisztikai tesztek, amelyek megvizsgálják a variancia homogenitását, azonban ezek érzékenyek a normálistól való eltérésre.



9.3. ábra. Resziduális ábra olyan mintázattal, amely nem konstans kezelés varianciát mutat (Collins és Seeney, 1999)

A hibák függetlenségének feltétele a modell reziduumok eloszlására vonatkozik. Szabadföldi kísérletekben a szomszédos parcellák hasonló terméseket adnak és ha egy kezelés minden ismétlése (parcellája) szomszédos lenne, a hibákról nem lehetne feltételezni, hogy függetlenek. Egy adott kísérleti elrendezésnél a kezelések korrekt randomizációját követve, biztosíthatjuk a hibák függetlenségét. A hibák nem-függetlensége nem korrigálható a kísérlet beállítása után. Ez a feltétel ellenőrizhető a reziduumok logikus sorrendbe rendezésével a nagyságrendjüktől függetlenül, mint amilyen a parcella sorrend vagy az illesztett értékek növekvő sorrendjében; a reziduális értékekről feltételezzük, hogy random sorrendben következnek. Például pozitív vagy negatív értékek nagy sorrendje nem lenne valószínű és a reziduumok függetlenségének hiányát jeleznék. Az előjel teszt használható a reziduumok előjelére, logikus sorrendbe rendezésük után, és a nem szignifikáns eredmény jelzi azt, hogy a reziduumok függetlenek.

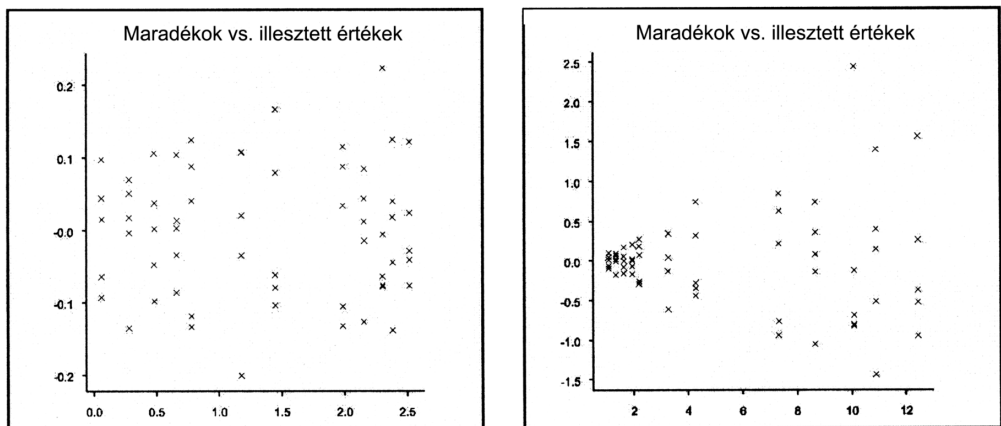
A hibák normális eloszlásának feltétele ugyancsak a modell reziduumok eloszlására vonatkozik. A reziduumok hisztogramja egy vizuális bemutatását adja annak, hogy ez a feltétel teljesült-e. Ha a hibák normális eloszlásúak, akkor a reziduumok nagy része a plusz vagy mínusz egyszeres szórás között helyezkedik el; a reziduumok többsége a plusz vagy mínusz kétszeres szórás között helyezkedik el és csak kevés reziduum található a plusz vagy mínusz háromszoros szóráson kívül. A reziduumok tipikus ábráját az illesztett értékek szerint a 9.4. ábra mutatja. A reziduumok normális eloszlása az adatok sajátossága sokkal inkább, mint az elrendezése, és ennek a feltételnek a megsértése valószínűleg az aszimmetrikus adateloszlásnak tulajdonítható. Ezekben az esetekben az adatok normális eloszlásának a hiánya korrigálható az adatok transzformálásával. Számos statisztikai teszt van, amelyeket felhasználhatunk az adatok normális eloszlásának ellenőrzésére.

9.4. A varianciaanalízis feltételeinek ellenőrzése GenStat programmal

A variancia homogenitása

Feltételezzük, hogy a variancia homogén, azaz a random (véletlen) variáció mértéke hasonló minden kísérleti egységnél. A variancia homogenitása könnyen becsülhető a maradékok (reziduumok, a random hiba becslései) ábrázolásával az illesztett értékek ellenében; ha a variancia homogén, a reziduális értékeknek egy egyenes (uniform) sávon belül kellene elhelyezkedniük (9.4. ábra).

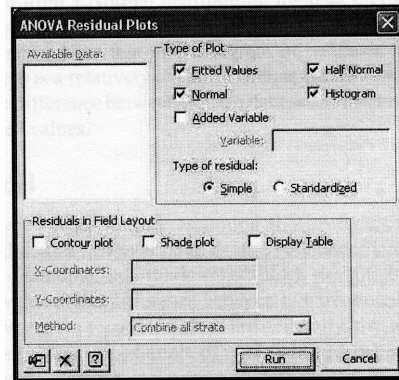
Egészen általános, különösen a megszámlolt adatoknál, hogy a reziduumok variációja nő, ahogy a reakció értéke növekszik (9.4. ábra). Ebben az esetben a kezelések közötti különbségek standard hibáját (szórását) túlbecsüljük a kezelések közötti különbségre alacsony átlagokkal, és alulbecsüljük a nagyobb átlagok közötti különbségre, és ezáltal téves következtetésre juthatunk. Ha a maradékok ábrája az illesztett értékektől függően nem homogén varianciát mutat, az adatok transzformációját kell megfontolni.



9.4. ábra. Reziduális értékek vs. illesztett értékek a varianciák homogén (bal oldali ábra) és nem homogén (jobb oldali ábra) eloszlásakor

A reziduumok szabályszerűsége és függetlensége

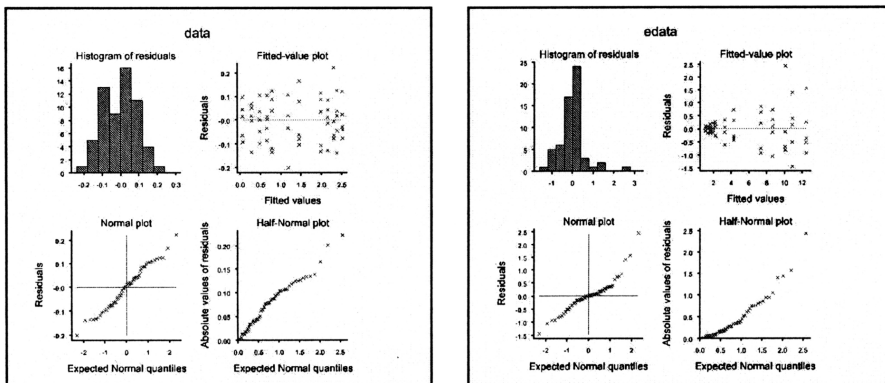
A varianciaanalízis feltételezi, hogy az adatok random hibát tartalmaznak (a maradékokkal becsüljük), amely független és normális eloszlású minden adatértékre. A reziduális értékek nem-szabályszerűsége rendszerint ugyancsak összefügg a varianciák homogenitásának hiányával és grafikusan számos módon vizsgálható. Először, a reziduális értékek ábrázolhatók hisztogram formájában – ennek megközelítőleg normális eloszlásúnak kellene lenni, vagyis nem-ferde harang alakú eloszlásnak. Alternatív módon egy normális ábra vagy fél-normális ábra használható. Ez ábrázolja a rendezett reziduumokat (vagy abszolút értékeiket) a normális eloszlás negyedei (quantiles) ellenében. Ha a reziduumoknak normális az eloszlása, ezeknek az ábráknak egyenes vonalúnak kellene lenniük.



9.5. ábra. A varianciaanalízis feltételeinek grafikus ellenőrzése az ANOVA Residual Plots menüvel

Ezek az ábrák, együtt a reziduumok ábrájával az illesztett értékektől függően, elkészíthetők az *ANOVA Residual Plots* menüvel. Ezt megkapjuk, ha rákattintunk a *Further Output* gombra az *Analysis of Variance* menüben, és azután a *Residual Plots* gombra az *ANOVA Further Output* menüben (9.5. ábra). A menü lehetővé teszi, hogy kiválasszuk azokat az ábrákat, melyeket látni szeretnénk. A 9.6. ábra alapértelmezett beállítással volt elkészítve. A 9.6. ábra baloldali részábrája normális eloszlású adatokból, míg a jobboldali részábrára nem normális eloszlású adatokból származik, ahol a variancia nő a reakció változó nagyságával. Figyeljük meg, hogy a reziduumok hisztogramja a 9.6. ábrán kissé ferde eloszlású, azonban viszonylag kis különbség van a normális és fél-normális ábrák között. Legvilágosabb a különbség a két adatsorozat között a reziduumok ábráján az illesztett értékektől függően.

Egy kiugró adat olyan extrém megfigyelés, amely nagy reziduális értékű egységhez vezet. GenStat ANOVA figyelmeztetéseket ad, ha az egységek nagy reziduális értékekkel rendelkeznek, összehasonlítva az egység standard hibájával (szórásával). A feltételek megsértése gyakran korrigálható az adatok transzformálásával, felhasználva a *Calculate* menüt. Különböző transzformációk megfelelőek különböző adatokra.



9.6. ábra. A varianciaanalízis feltételeinek grafikus ellenőrzése. A bal oldali részábrára normális eloszlású, míg a jobboldali részábrára nem normális eloszlású adatokból származik

9.5. Összehasonlítás a kezelésközvetlenek között

9.5.1. Többszörös terjedelelem vizsgálata

A varianciaanalízis feltárja, hogy, létezik-e szignifikáns különbség a kezelésközvetlenek minden lehetséges kombinációja között a faktoriális kísérletben. A varianciaanalízis azonban nem tárja fel bármely meglévő szignifikáns különbség pontos helyét. A fix hatások ANOVA-t követve, ún. többszörös terjedelelem vizsgálata végezhető, hogy kiderítsük, mely kezelésközvetlenek között van szignifikáns különbség.

Számos többszörös összehasonlítási teszt áll rendelkezésre. Talán a legegyszerűbb a **legkisebb szignifikáns különbség (SzD, LSD)** kiszámítása, amely lényegében a Student-féle t-teszt egyik típusa. Ez a teszt azonban nem veszi figyelembe a kísérletben levő minta átlagok számát, és ezáltal viszonylag nagy a kockázata az I. típusú hiba elkövetésének. Ebből az okból, használatát körültekintően kell tervezni, a kísérlet előtt. Szélesebb körben elfogadott az átlagok közötti különbségek általános érvényű tesztelésére a kísérlet beállítását követően a **Tukey HSD teszt** és a hasonló **Student-Newman-Keuls próba**. Mindkettő a q mutató eloszlásán alapul, melynek kritikus értéke változik az összehasonlítani kívánt minta átlagok összes számával, és kevésbé vannak kitéve az I. típusú hibának, mint az LSD teszt, azonban nagyobb a valószínűsége a II. típusú hiba elkövetésének. Egy másik hasonló teszt a **Dunnett teszt**, amely egyetlen kontroll minta átlagot hasonlít össze az összes többi minta átlaggal a kísérletben.

Azokban a kísérletekben, amelyekben minden lehetséges kezelés-pár értékelése szükséges, az SzD próba rendszerint nem megfelelő. Ez különösen igaz akkor, amikor a kezelésekek száma nagy. Ilyen esetekben hasznos a **Duncan többszörös terjedelelem próba (DMRT)**. A DMRT alkalmazásának eljárása hasonló az SzD próbához, eltérően azonban az SzD próbától, amelyben csak egyetlen érték szükséges bármely kezelés-pár összehasonlításhoz a szignifikancia adott szintjén, a DMRT szükségessé teszi az értékek sorozatának kiszámítását, mindegyik kezelés-pár összehasonlításához.

Fel kell hívni a figyelmet egy fontos szabályra, éspedig arra, hogy a többszörös összehasonlítási próbákat nem szabad elvégezni, ha az ANOVA nem mutatott ki szignifikáns kezeléshatást. Ha az ANOVA nem szignifikáns F-értéket eredményez egy bizonyos kezelés faktorra, akkor e kezelésközvetlenek közötti különbségek további analízise teljesen érvénytelen. A legtöbb többszörös összehasonlítási teszt alapja a két minta átlag közötti különbség standard hibája.

9.5.2. A különbség standard hibája

A különbség standard hibáját (két középérték különbségének szórását) (SED) két átlagra megadja a minta varianciák összege, osztva a megfelelő minta nagysággal, n-nel és négyzetgyökük:

$$SED = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

Mivel az ANOVA eredeti feltétele, hogy minden minta variancia egyenlő, ezáltal a reziduális (maradék) variancia kiszámított értéke ($MS_{\text{maradék}}$) felhasználható minden minta varianciájának mérőszámaként a kísérletben. A fenti egyenlet ezáltal egyszerűsíthető:

$$SED = \sqrt{MS_{\text{maradék}} \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}$$

Továbbá, ha az összehasonlítandó minta méretek egyenlőek, akkor a fenti egyenlet tovább egyszerűsíthető:

$$SED = \sqrt{\frac{2 \times MS_{\text{maradék}}}{n}}$$

A SED a véletlen hiba mérőszámát (mutatóját) adja, mérve a különbséget két átlag között, amellyel a megfigyelt különbség összehasonlítható. Felhasználható a következő többszörös összehasonlítási példákban.

A legkisebb szignifikáns különbség (SzD, LSD), a Tukey-féle szignifikáns különbség (HSD), a Student-Newman-Keuls többszörös terjedelem próba (SNK) és a Dunnett próba kiszámítását az alábbi kísérlet adataival mutatjuk be. A Duncan többszörös terjedelem próbát (DMRT) egy másik kísérlettel illusztráljuk.

Kísérlet leírása. Az MTA Mezőgazdasági Kutatóintézetében, Martonvásáron, a nitrogén műtrágyázás hatását vizsgáltuk a búzafajták klorofill-tartalmát jellemző SPAD értékekre 2007-ben. A N-műtrágya dózisa 0, 80, 160 és 240 kg ha⁻¹ volt. A búzafajták átlagában kifejezett SPAD értékeket az egytényezős véletlen blokk elrendezésnek megfelelő varianciaanalízissel értékeltük. Az ismétlés és a nitrogén szintek a faktorok, a SPAD érték a változó az adatbevitelnél. A kísérlet számítógépes adattáblázata az alábbi:

Ismétlés	Nitrogén szintek	SPAD értékek
1	1	42.57
2	1	41.03
3	1	42.48
4	1	42.57
1	2	53.46
2	2	53.78
3	2	53.99
4	2	54.06
1	3	53.94
2	3	54.53
3	3	54.65
4	3	55.97
1	4	55.71
2	4	56.31
3	4	55.92
4	4	56.13

A varianciaanalízis eredménytáblázata az alábbi:

Analysis of variance					
Variate: SPADertek					
Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Ismetles stratum	3	1.5892	0.5297	1.74	
Ismetles.*Units* stratum					
Nitrogenszint	3	494.1993	164.7331	539.70	<.001
Residual	9	2.7471	0.3052		
Total	15	498.5356			

Megállapítható, hogy a N-műtrágyának $P < 0.01$ szinten szignifikáns hatása volt a búzafajták SPAD értékére. A variancia táblázatból azonban nem tudjuk, hogy a N-műtrágyázás mely szintjei között volt szignifikáns különbség. Először bemutatjuk négy többszörös összehasonlítási teszt (legkisebb szignifikáns különbség, Tukey HSD próba, Student–Newmann–Keuls próba és a Dunnett próba) kézi úton történő kiszámításának módszerét.

9.5.3. Legkisebb szignifikáns különbség (SzD, LSD)

A legkisebb szignifikáns különbség (LSD) a legkisebb numerikus különbség, amelynek meg kell lenni két minta átlag között egy kísérletben, ahhoz, hogy két átlagot szignifikánsan különbözőnek tekinthessünk egy adott valószínűségi szinten. Feltételezve, hogy a mintáknak egyenlő a varianciája és hasonló a mérete, ekkor:

$$LSD = t \times SED$$

ahol: t = Student t érték, melyet a Student t táblázatból (Melléklet, 1. táblázat) kapunk meg, a valószínűség kívánatos szintjén (rendszerint $\alpha = 0.05$) és a maradék szabadságfoknál; SED = a különbség standard hibája.

Kiszámítva az LSD értékét, bármely minta pár (kezelés pár), melynek átlagai között olyan numerikus különbség van, amely egyenlő vagy nagyobb, mint az LSD, szignifikáns különbségnek tekinthető a valószínűségnek azon a szintjén, amelynél az LSD-t meghatároztuk.

A fő probléma az LSD teszttel az, hogy összegezi az I. típusú hiba elkövetésének valószínűségét minden esetben, amikor két átlagot összehasonlítunk egy kísérletben. Ezáltal, abból a célból, hogy csökkentsük az I. típusú hiba bekövetkezését, az LSD-t korlátozott számú pár összehasonlításnál kellene használni, és az összehasonlítandó átlagokat a kísérlet beállítását megelőzően kellene kiválasztani. Az LSD próba használatát a példa mutatja.

Kísérlet. A legkisebb szignifikáns különbség kiszámítása az egytényezős ANOVA-t követően. Adatok az ANOVA-ból. A kísérletben az N-műtrágyázás hatását vizsgáltuk a búzafajták klorofill-tartalmát jellemző SPAD értékekre.

$MS_{\text{maradék}}$ (az ANOVA táblából) = 0.3052 (maradék DF = 9)
 t (a Student's táblázatból; $\alpha = 0.05$, maradék DF = 9) = 2.262
 n (minden mintára) = 4

$$LSD_{(\alpha=0.05)} = t \times \sqrt{\frac{2 \times MS_{\text{maradék}}}{n}} = 2.262 \times \sqrt{\frac{2 \times 0.3052}{4}} = 0.884$$

Következtetés: ha a megfigyelt különbség az előre kiválasztott kezelés párok minta átlagai között ≥ 0.884 , akkor a különbség szignifikáns ($P < 0.05$).

9.5.4. Tukey-féle szignifikáns különbség (HSD)

A Tukey-féle szignifikáns különbség (HSD) annyiban különbözik az LSD próbától, hogy függetlenül attól, hány minta átlagot hasonlítunk össze, az I. típusú hiba elkövetésének valószínűsége konstans szinten marad és nem halmozódik, ahogy több mintát hasonlítunk össze. Ilyen okból a Tukey HSD próba felhasználható bármilyen számú mintaátlag összehasonlítására, miután a kísérlet befejeződött és amikor az összehasonlítások nem voltak előre megtervezve. Másrésztől a Tukey HSD próba kevésbé erős, mint az LSD próba és kockáztatja kis, de valódi különbségek feltárását.

A Tukey-féle HSD próbánál a kritikus q értéket használjuk fel a minimális kritikus különbség (MCD) kiszámítására, amely a legkisebb numerikus különbség, amelynek meg kell lenni egy mintaátlag pár között, hogy a két minta közötti különbséget szignifikánsnak tekintsük.

Az első lépés az eljárásban az, hogy meghatározzuk a minta átlagok közötti különbségek standard hibáját, a Tukey becslés szerint (Tukey SED a jelölése). Ezt megkapjuk, ha a maradék varianciát ($MS_{\text{maradék}}$) elosztjuk 2-vel, majd elosztjuk a minta méret reciprokának összegével, és vesszük ennek a négyzetgyökét. Ahol a mint méretek egyenlők, a számítás leegyszerűsödik. Egyenlő minta méretre:

$$\text{Tukey SED} = \sqrt{\frac{MS_{\text{maradék}}}{n}}$$

Nem egyenlő minta méretre:

$$\text{Tukey SED} = \sqrt{\frac{MS_{\text{maradék}}}{2} \times \left(\frac{1}{n_a} + \frac{1}{n_b} \right)}$$

ahol: n_a, n_b = az összehasonlításra kiválasztott minták mérete.

A próba statisztika kritikus értékét (q_{krit}) megkapjuk a (Melléklet, 9. táblázat) a valószínűség (α) kiválasztott kritikus szintjén, maradék szabadságfoknál (V) és a kísérletben levő minták (kezelések) számára (k). A minimális kritikus különbséget

(MCD) az átlagok között úgy határozzuk meg, hogy megszorozzuk a Tukey SED értékét a q_{kritikus} értékével:

$$\text{MCD} = q_{\text{krit}} \times \text{Tukey SED}$$

Ha a különbség a kiválasztott minta átlag párok között egyenlő vagy nagyobb, mint az MCD, akkor a populációk, amelyekből a minták származnak, szignifikánsan különbözőnek tekinthetők a próba adott szignifikancia szintjén.

A Tukey próba végső eredményének bemutatásakor, hasznos eljárás összeállítani egy táblázatot, amely mutatja a számított különbséget minden lehetséges minta pár között. Ily módon nagyon könnyű összehasonlítani a megfigyelt különbségeket az átlagok között az MCD-vel és így megkülönböztetni azokat a kezeléseket, amelyek szignifikánsan különböznek.

Kísérlet. A Tukey HSD teszt kiszámítása az egytényezős ANOVA-t követően. Adatok az ANOVA-ból. A kísérletben az N-műtrágyázás hatását vizsgáltuk a búzafajták klorofill-tartalmát jellemző SPAD értékekre.

$$\text{MS}_{\text{maradék}} \text{ (az ANOVA táblázatból)} = 0.3052 \text{ (DF = 9)}$$

$$n = 4$$

$$\text{Tukey SED} = \sqrt{\frac{\text{MS}_{\text{maradék}}}{n}} = 0.2762$$

$$\text{Kritikus } q \text{ érték a táblázatból } (\alpha = 0.05, \text{maradék DF} = 9, k = 4) = 4.415$$

$$\text{MCD az átlagok között} = 4.415 \times 0.2762 = 1.219$$

Következtetés: ha a megfigyelt különbség bármely minta pár átlagai között > 1.219 , akkor a különbség szignifikáns ($P < 0.05$).

Mintaátlagok közötti különbségek táblázata:

(A Tukey HSD próbával feltárt szignifikáns különbségeket * jelöli)

	Különbség a mintaátlagok között			
	240	160	80	0
Nitrogén dózis	($\bar{x} = 56.02$)	($\bar{x} = 54.77$)	($\bar{x} = 53.82$)	($\bar{x} = 42.16$)
0 ($\bar{x} = 42.16$)	13.86*	12.61*	11.66*	
80 ($\bar{x} = 53.82$)	2.195*	0.95		
160 ($\bar{x} = 54.77$)	1.25*			
240 ($\bar{x} = 56.02$)				

9.5.5. Student –Newman – Keuls többszörös terjedelem próba

A Student-Newman-Keuls (SNK) próba egy kompromisszumnak tekinthető az LSD és HSD próbák között; kissé erősebb, mint a Tukey HSD próba, jóllehet legtöbb esetben a kétféle próba ugyanazt az eredményt adja. Ha azonban a különbség a két átlag között

közel van a szignifikáns kritikus különbséghez, nagyobb a valószínűsége a szignifikáns hatás feltárásának az SNK használatakor, mint a Tukey próba esetében. Vitatható azonban, hogy a Tukey HSD próbának kisebb az esélye az I. típusú hiba előállításában, és ezáltal biztonságosabb választás, jóllehet nincs általános egyetértés ebben.

Az SNK próba elvégzéséhez a minta átlagokat először nagyság szerint rendezni kell. Az átlagok közötti különbség standard hibáját (szórását) (SED) ugyanazzal a formulával becsüljük, mint a Tukey HSD próbánál. A q próba statisztika kritikus értékét ezután megkapjuk a Studentizált terjedelem táblázatból (Melléklet 9. táblázat), a szükséges kritikus valószínűségnél (α), a maradék szabadságfoknál (V) és a rangsorolt minta átlagok számánál (p), amely abban a tartományban helyezkedik el, amelyet meghatároz két összehasonlítandó minta átlag. (Ezáltal p helyettesíti $k-t$, melyet a Tukey HSD próbában használtunk.) Például, ha a különbséget a második legnagyobb minta átlag (sorrend = 2) és az ötödik legnagyobb minta átlag (sorrend = 5) között kellene tesztelni, akkor p értéke 4 lenne. A SED-et ezután megszorozzuk a q_{krit} értékkel, hogy megkapjunk egy kritikus minimális különbséget, amelynek el kell különíteni a két kiválasztott minta átlagot, hogy a szignifikáns különbséget közöttük igazoljuk. Egy minimális kritikus különbséget (MCD) először úgy kapunk meg, hogy összehasonlítjuk a legkisebb átlagot és a legnagyobb átlagot, és ezután a következő legkisebb és legnagyobb átlagot, és így tovább. Megfigyelhető, hogy két szomszédos átlagra a Student-Newman-Keuls MCD valóban ugyanaz lesz, mint az LSD érték. Mihelyt minden MCD értéket meghatároztunk, a különbség minden átlag pár között megvizsgálható, és szignifikáns különbség igazolható az átlagok között, ahol a számszerű különbség az átlagok között egyenlő vagy nagyobb, mint a megfelelő MCD.

A Student-Newman-Keuls próba kiszámítását a kísérlet mutatja. Ha az SNK próba eredményét a kísérletben összehasonlítjuk a Tukey HSD próbával ugyanazokra az adatokra, megfigyelhető, hogy a 160 és 80 kg ha⁻¹ N dózisok összehasonlításakor, az SNK próba szignifikáns különbséget jelzett, míg a Tukey próba nem. Ez mutatja azt a pontot, hogy az SNK próbának viszonylag nagyobb a statisztikai ereje, mint a Tukey próbának, kis különbségek szignifikánsnak történő feltárásában.

Kísérlet. A Student-Newman-Keuls többszörös terjedelem próba kiszámítása az egytényezős ANOVA-t követően. A kísérletben az N-műtrágyázás hatását vizsgáltuk a búzafajták klorofill-tartalmát jellemző SPAD értékekre.

$$MS_{\text{maradék}} \text{ (az ANOVA táblázatból)} = 0.3052 \text{ (DF = 9)}$$

$$n = 4$$

$$TukeySED = \sqrt{\frac{MS_{\text{maradék}}}{n}} = 0.2762$$

Kezelés minták rangsorolva átlagaik szerint	N-műtrágya dózis	Átlag	Sorrend
	0	42.16	1
	80	53.82	2
	160	54.77	3
	240	56.02	4

A minimális kritikus különbség (MCD) meghatározása minden páros összehasonlításra rangsorban.

Sorrend	N- műtrágya kezelés	Különbség az átlagok között ($\bar{x}_1 - \bar{x}_2$)	p	$q_{(p; DF \text{ maradék}, \alpha=0.05)}$	MCD	Szig-nifikancia
4 vs 3	240 vs 160	1.245	2	3.199	0.884	$P < 0.05$
4 vs 2	240 vs 80	2.195	3	3.949	1.091	$P < 0.05$
4 vs 1	240 vs 0	13.855	4	4.415	1.219	$P < 0.05$
3 vs 2	160 vs 80	0.95	2	3.199	0.884	$P < 0.05$
3 vs 1	160 vs 0	12.61	3	3.949	1.091	$P < 0.05$
2 vs 1	80 vs 0	11.66	2	3.199	0.884	$P < 0.05$

Következtetés: ahol a megfigyelt különbség bármely mintaátlagok között ($\bar{x}_1 - \bar{x}_2$) \geq MCD érték, akkor a különbség azokra az átlagokra szignifikáns ($P < 0.05$).

9.5.6. Dunnett próba

A Dunnett próbát akkor használjuk, amikor a kutató ahelyett, hogy érdekelt lenne minden lehetséges minta átlag közötti összehasonlításban, valóban csupán egy specifikált minta átlag, melyet nevezhetünk „kontrollnak” és a többi minta átlag összehasonlításában érdekelt. Ebben az esetben egy kissé eltérő valószínűségi eloszlást használunk attól, amit a Tukey próbánál használtunk, és ezt q' eloszlásnak jelöljük.

A különbség standard hibáját a következőképpen becsüljük:

$$SED = \sqrt{\frac{2 \times MS_{\text{maradék}}}{n}} \quad SED = \sqrt{MS_{\text{maradék}} \left(\frac{1}{n_a} + \frac{1}{n_{\text{kontroll}}} \right)}$$

Egyenlő minta méretre: Nem egyenlő minta méretre:
 ahol: $MS_{\text{maradék}}$ = maradék MS az ANOVA-ból,
 n = összehasonlítandó minták mérete

A q' próba statisztika kritikus szintjét megkapjuk a q' valószínűségi eloszlásból a szükséges kritikus α valószínűségnél, maradék DF-nél és k mintánál (Melléklet 10. táblázat). A minimális kritikus különbséget (MCD), minden minta átlag és a kontroll átlag közötti különbség szignifikanciájának vizsgálatára a következőképpen határozzuk meg:

$$MCD = q'_{\text{krit}} \times SED$$

Mint ahogy lehetséges specifikálni, hogy vajon a kontroll átlagnak nagyobbnak vagy kisebbnek kellene lenni, mint az összehasonlított minta átlag, a Dunnett próba egyoldali próbaként végezhető, ha szükséges. Az egyoldali kritikus q' értéket külön táblázat tartalmazza (Melléklet 10. táblázat).

Kísérlet. A Dunnett próba kiszámítása az egytényezős ANOVA-t követően. Adatok az ANOVA-ból származnak. A kísérletben az N-műtrágyázás hatását vizsgáltuk a búzafajták klorofill-tartalmát jellemző SPAD értékekre.

Tételezzük fel, hogy az N-műtrágyázás nélküli kezelés lett előzetesen kontrollként azonosítva, amellyel minden más kezelést összehasonlítottunk.

$$MS_{\text{maradék}} \text{ (az ANOVA táblázatból)} = 0.3052 \text{ (DF} = 9)$$

$$n = 4$$

$$SED = \sqrt{\frac{2 \times MS_{\text{maradék}}}{n}} = 0.391$$

A minimális kritikus különbség (MCD) meghatározása minden páros összehasonlításra a kontroll és a kezelések között.

N-műtrágya kezelések összehasonlítása	Különbség az átlagok között ($\bar{x}_1 - \bar{x}_2$)	SED	$q' (k; DF_{\text{maradék}},$ $\alpha=0.05)$	MCD	Szigni- fikancia
2 vs 1	11.66	0.391	2.81	1.099	P < 0.05
3 vs 1	12.61	0.391	2.81	1.099	P < 0.05
4 vs 1	13.86	0.391	2.81	1.099	P < 0.05

Következtetés: ahol a megfigyelt különbség bármely kezelésközé és a kontroll közlé között ≥ 1.099 , ott a különbség szignifikáns (P < 0.05).

9.5.7. Számítógépes értékelés GenStat programmal

A menük kiválasztásának sorrendje: Stats – Analysis of Variance – General ... – One-way ANOVA (in Randomized Blocks) – Options – Multiple Comparisons – a legördülőpróbák közül kiválasztjuk azt, amelyiket használni kívánunk. A GenStat-ban rendelkezésre álló többszörös terjedelem próbák közül a következőket választottuk ki: LSD próba, Tukey's próba, Student-Newman-Keuls próba, Duncan's többszörös terjedelem próba, Scheffe próba. A Scheffe próbával elvégzünk minden összehasonlítást, miközben az általános I. típusú hiba szintet a kívánatos α szinten tartjuk. Ahhoz, hogy ezt elérjük, minden egyedi összehasonlítást α -nál alacsonyabb szinten kell tesztelni. Rendszerint nem tervezett összehasonlítások tesztelésére használjuk.

Least significant differences of means (5% level)

Table	Nitrogenszint
rep.	4
d.f.	9
l.s.d.	0.884

Tukey's 95% confidence intervals**Nitrogenszint**

Comparison	Difference	Lower 95%	Upper 95%	Significant
4 vs 3	1.245	0.025	2.465	yes
4 vs 2	2.195	0.975	3.415	yes
4 vs 1	13.855	12.635	15.075	yes
3 vs 2	0.950	-0.270	2.170	no
3 vs 1	12.610	11.390	13.830	yes
2 vs 1	11.660	10.440	12.880	yes

Mean

4	56.02	
3	54.77	
2	53.82	
1	42.16	

	Mean	
4	56.02	a
3	54.77	b
2	53.82	b
1	42.16	c

Student-Newman-Keuls test

Experiment-wise error rate = 0.05

Nitrogenszint

Comparison-wise error rates

Range probability		c
2	0.95	2.262
3	0.95	2.792
4	0.95	3.122

Comparison	Difference	Lower 95%	Upper 95%	t	Significant
4 vs 3	1.245	0.361	2.129	3.19	yes
4 vs 2	2.195	1.104	3.286	5.62	yes
4 vs 1	13.855	12.635	15.075	35.47	yes
3 vs 2	0.950	0.066	1.834	2.43	yes
3 vs 1	12.610	11.519	13.701	32.28	yes
2 vs 1	11.660	10.776	12.544	29.85	yes

4	56.02
3	54.77
2	53.82
1	42.16

	Mean	
4	56.02	a
3	54.77	b
2	53.82	c
1	42.16	d

Duncan's multiple range test

Experiment-wise error rate = 0.05

Nitrogen szint

Comparison-wise error rates

Range	probability	c
2	0.9500	2.262
3	0.9025	2.361
4	0.8574	2.418

Comparison	Difference	Lower 95%	Upper 95%	t	Significant
4 vs 3	1.245	0.361	2.129	3.19	yes
4 vs 2	2.195	1.273	3.117	5.62	yes
4 vs 1	13.855	12.910	14.800	35.47	yes
3 vs 2	0.950	0.066	1.834	2.43	yes
3 vs 1	12.610	11.688	13.532	32.28	yes
2 vs 1	11.660	10.776	12.544	29.85	yes

	Mean	
4	56.02	
3	54.77	
2	53.82	
1	42.16	

	Mean	
4	56.02	a
3	54.77	b
2	53.82	c
1	42.16	d

Scheffe test

Experiment-wise error rate = 0.05

NitrogenszintTwo means are significantly different if $|t| \geq 3.4041$

Critical values for t

5%	3.404
1%	4.580
0.1%	6.458

Comparison	Difference	Lower 95%	Adjusted Upper 95%	t	probability	Significant
4 vs 3	1.245	-0.085	2.575	3.19	0.0677	no
4 vs 2	2.195	0.865	3.525	5.62	0.0027	yes
4 vs 1	13.855	12.525	15.185	35.47	0.0000	yes
3 vs 2	0.950	-0.380	2.280	2.43	0.1889	no
3 vs 1	12.610	11.280	13.940	32.28	0.0000	yes
2 vs 1	11.660	10.330	12.990	29.85	0.0000	yes

	Mean		
4	56.02		
3	54.77		
2	53.82		
1	42.16		

	Mean	
4	56.02	a
3	54.77	ab
2	53.82	b
1	42.16	c

9.5.8. Duncan többszörös terjedelem próba (DMRT)

A DMRT értékek kiszámítása, hasonlóan az SzD próbához, elsődlegesen függ az összehasonlítandó kezeléscsoportok különbségének szórásától (standard hibájától) (s_d , SED).

Kísérlet leírása: A DMRT kiszámításának lépéseit az összes lehetséges kezeléscsoport összehasonlítására egy olyan CRD elrendezésű kísérlettel mutatjuk be, amely hét inszekticid kezelés hatását, négy ismétlésben vizsgálták a rizs termésére (Gomez és Gomez, 1984). A CRD elrendezésű inszekticid kísérlet ismétlésekenkénti alapadatai:

Kezelés	Szemtermés kg ha ⁻¹				Kezelés Összeg T	Kezelés átlag
	I. ismétlés	II. ismétlés	III. ismétlés	IV. ismétlés		
1	2537	2069	2104	1797	8507	2127
2	3366	2591	2211	2544	10712	2678
3	2536	2459	2827	2385	10207	2552
4	2387	2453	1556	2116	8512	2128
5	1997	1679	1649	1859	7184	1796
6	1796	1704	1904	1320	6724	1681
7 (kontroll)	1401	1516	1270	1077	5264	1316
Főösszeg G					57110	
Főátlag						2040

A kísérlet varianciaanalízisének eredménye a GenStat programmal a következő:

Analysis of variance					
Variate: Termes					
Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Kezeles	6	5587175.	931196.	9.83	<.001
Residual	21	1990238.	94773.		
Total	27	7577412.			

Tables of means							
Variate: Termes							
Grand mean 2040.							
Kezeles	1	2	3	4	5	6	7
	2127.	2678.	2552.	2128.	1796.	1681.	1316.

Standard errors of differences of means	
Table	Kezeles
rep.	4
d.f.	21
s.e.d.	217.7

A DMRT kiszámításának lépései a következők:

1. Lépés. Rangsoroljuk az összes kezelésátlagot csökkenő (vagy növekvő) sorrendben. Szokásos a kezelés-párokat a preferencia sorrendjében rangsorolni. Károsító előfordulási adatoknál, a rangsor rendszerint a legkevésbé fertőzött kezeléstől a legsúlyosabban fertőzött kezelések irányába tart. Példánkban a hét kezelésátlagot csökkenő termés sorrendjében rendeztük:

Kezelés	Átlagtermés kg ha ⁻¹	Rangsor
T ₂	2678	1
T ₃	2552	2
T ₄	2128	3
T ₁	2127	4
T ₅	1796	5
T ₆	1681	6
T ₇	1316	7

2. Lépés. Számítsuk ki az s_d értékét (l. 10.1.4. fejezet), példánkban a következőképpen:

$$s_d = \sqrt{\frac{2s^2}{r}} = \sqrt{\frac{2(94.773)}{4}} = 217.68 \text{ kg ha}^{-1}$$

3. Lépés. Számítsuk ki a legrövidebb szignifikáns tartományok (t-1) értékét:

$$R_p = \frac{(r_p)(s_d)}{\sqrt{2}} \quad p = 2, 3, \dots, t \text{ értékre}$$

ahol t a kezelések összes száma, s_d két kezelésátlag különbségének szórása (standard hibája), r_p a szignifikáns Student-féle tartományok táblázati értékei a melléklet 14. táblázatából és p a távolság a rangsorban az összehasonlítandó kezelés-párok között (azaz, p = 2 két átlagra egymás után következő rangsorban, és p = t a legnagyobb és legkisebb átlagra). A studentizált terjedelem a gyakorlatban sokszor alkalmazott próbastatisztika, különösen a középértékek többszörös összehasonlításakor. Példánkra, az r_p értékek 21 hiba d.f. értékkel és 5%-os szignifikancia szinten a melléklet 14. táblázatából:

p	r_p (.05)
2	2.94
3	3.09
4	3.18
5	3.24
6	3.30
7	3.33

A (t-1) = 6 R_p érték kiszámítása:

p	R_p
2	$\frac{(2.94)(217.68)}{\sqrt{2}} = 453$
3	$\frac{(3.09)(217.68)}{\sqrt{2}} = 476$
4	$\frac{(3.18)(217.68)}{\sqrt{2}} = 489$
5	$\frac{(3.24)(217.68)}{\sqrt{2}} = 499$
6	$\frac{(3.30)(217.68)}{\sqrt{2}} = 508$
7	$\frac{(3.33)(217.68)}{\sqrt{2}} = 513$

4. Lépés. Azonosítsuk és csoportosítsuk együtt minden kezelésátlagot, melyek szignifikánsan nem különböznek egymástól.

A) Számítsuk ki a különbséget a legnagyobb kezelésátlag és a legnagyobb R_p érték között (R_p érték p = t-nél) és nyilvánítsunk minden kezelésátlagot, melyeknek értéke kisebb, mint a számított különbség, szignifikánsan különbözőnek a legnagyobb kezelésátlagtól.

Ezt követően, számítsuk ki a tartományt a megmaradó kezeléscsoportok (azok az átlagok, melyeknek értéke nagyobb, mint vagy egyenlő a legnagyobb átlag és a legnagyobb R_p érték közötti különbséggel) között és hasonlítsuk össze ezt a tartományt az R_p értékkel $p = m$ -nél, ahol m a kezelések száma a csoportban. Ha a számított tartomány kisebb, mint a megfelelő R_p érték, akkor minden m kezeléscsoportot a csoportban egymástól szignifikánsan nem különbözöknek kell tekinteni.

Végül, húzzunk egy vertikális vonalat, amely összeköt minden átlagot, amelyet egymástól szignifikánsan nem különbözöknek nyilvánítottunk.

Példáinkra, a különbség a legnagyobb R_p érték (R_p érték $p = 7$ -nél), vagyis 513 és a legnagyobb kezeléscsoport, azaz 2678 között: $2678 - 513 = 2165 \text{ kg ha}^{-1}$. Minden kezeléscsoport, a T_3 kivételével, kisebb, mint a számított 2165 kg ha^{-1} különbség, ezáltal ezeket a kezeléseket a T_2 -től szignifikánsan különbözöknek nyilvánítjuk.

Az $m = 2$ megmaradó kezeléscsoportból (T_2 és T_3), amelyeknek értékei nagyobbak, mint a számított 2165 kg ha^{-1} különbség, számítsuk ki a tartományt úgy, mint $2678 - 2552 = 126 \text{ kg ha}^{-1}$ és hasonlítsuk össze ezt az R_p értékkel $p = m = 2$ -nél, azaz 453-mal. Mivel a számított különbség kisebb, mint az R_p érték $p = 2$ -nél, a T_2 és T_3 átlagot egymástól szignifikánsan nem különbözöknek tekintjük. Egy vertikális vonalat húzzunk, hogy összekössük ezt a két átlagot:

Kezelés	Termésátlag, kg ha^{-1}
T_2	2678
T_3	2552
T_4	2128
T_1	2127
T_5	1796
T_6	1681
T_7	1316

B) Számítsuk ki a különbséget a második legnagyobb kezeléscsoport és a második legnagyobb R_p érték között (R_p érték $p = t - 1$ -nél) és nyilvánítsunk minden kezeléscsoportot, melynek értéke kevesebb, mint ez a különbség, szignifikánsan különbözöknek a második legnagyobb kezeléscsoporttól. Az m_1 megmaradó kezeléscsoportra, melynek értéke nagyobb, illetve egyenlő a számított különbséggel, számítsuk ki tartományukat és hasonlítsuk össze a megfelelő R_p értékkel (R_p érték $p = m_1$ -nél). Nyilvánítsunk minden kezelést a tartományon belül, szignifikánsan nem különbözöknek egymástól, ha a tartomány kisebb, mint a megfelelő R_p érték.

Példáinkra, a különbséget a második legnagyobb R_p érték (az R_p érték $p = 6$ -nál) és a második legnagyobb kezeléscsoport (T_3 átlag) között úgy számítjuk ki, mint $2552 - 508 = 2044 \text{ kg ha}^{-1}$. Mivel a T_5 , T_6 és T_7 kezelések átlagai kisebbek, mint 2044 kg ha^{-1} , szignifikánsan különbözöknek nyilvánítjuk a T_3 átlagától.

Az $m_1 = 3$ megmaradó kezeléscsoportok, melyeket nem nyilvánítottunk szignifikánsan különbözöknek, a T_3 , T_4 és T_1 kezelések. Tartományukat úgy számítjuk ki, mint $T_3 - T_1 = 2552 - 2127 = 425 \text{ kg ha}^{-1}$, melyet összehasonlítunk a megfelelő R_p értékkel $p = m_1 = 3$ -nál, vagyis 476 kg ha^{-1} -ral. Mivel a tartomány kisebb, mint az R_p érték

$p = 3$ -nál, a három megmaradó átlag szignifikánsan nem különbözik egymástól. A vertikális vonalat ezután megrajzoljuk, hogy összekösse a T_3 , T_4 és T_1 kezeléseket, az alábbiak szerint:

Kezelés	Termésátlag, kg ha ⁻¹
T2	2678
T3	2552
T4	2128
T1	2127
T5	1796
T6	1681
T7	1316

C) Folytassuk az eljárást a harmadik legnagyobb kezelésátlaggal, azután a negyedikkel, és így tovább, amíg minden kezelésátlagot megfelelően összehasonlítunk. Példánkban, a folyamat folytatódik a harmadik legnagyobb kezelésátlaggal. A különbséget a harmadik legnagyobb kezelésátlag (T_4 átlag) és a harmadik legnagyobb R_p érték (R_p érték $p = 5$ -nél) között úgy számítjuk ki, mint $2128 - 499 = 1629$ kg ha⁻¹. Kizárólag T_7 kezelés átlaga kevesebb, mint a számított 1629 kg ha⁻¹ különbség. Tehát T_7 szignifikánsan különbözik T_4 -től. A négy megmaradó kezelés, amelyek szignifikánsan nem különböznek, a T_4 , T_1 , T_5 és T_6 . Tartományát úgy számítjuk ki, mint $T_4 - T_6 = 2128 - 1681 = 447$ kg ha⁻¹. Mivel a számított tartomány kisebb, mint a megfelelő 489 kg ha⁻¹ R_p érték $p = 4$ -nél, mind a négy megmaradó átlagot egymástól szignifikánsan nem különbözőnek nyilvánítjuk. Egy vertikális vonalat húzunk, hogy összekösse a T_4 , T_1 , T_5 és T_6 átlagokat, mint az alábbiak mutatják:

Kezelés	Termésátlag, kg ha ⁻¹
T2	2678
T3	2552
T4	2128
T1	2127
T5	1796
T6	1681
T7	1316

Ezen a ponton, ugyanez a folyamat folytatható a negyedik legnagyobb kezelésátlaggal, és így tovább. Mivel azonban T_7 átlaga az egyedül kívülálló a megtett csoportosításból, egyszerűbb összehasonlítani a T_7 átlagot, felhasználva a megfelelő R_p értékeket, a megmaradó átlagokkal (nevezetesen: T_1 , T_5 és T_6). Ezeket az összehasonlításokat a következőképpen végezzük:

$$T_1 \text{ vs } T_7: 2127 - 1316 = 811 > R_p \text{ (} p = 4\text{-nél): } 489$$

$$T_5 \text{ vs } T_7: 1796 - 1316 = 480 > R_p \text{ (} p = 3\text{-nál): } 476$$

$$T_6 \text{ vs } T_7: 1681 - 1316 = 365 < R_p \text{ (} p = 2\text{-nél): } 453$$

A három összehasonlításból egyedül a T_6 és T_7 közötti különbség kisebb, mint a megfelelő R_p érték. Tehát T_6 és T_7 szignifikánsan nem különbözik egymástól. Egy vertikális vonalat húzhatunk, hogy összekösse a T_6 és T_7 átlagait, az alábbiak szerint:

Kezelés	Termésátlag, kg ha ⁻¹
T_2	2678
T_3	2552
T_4	2128
T_1	2127
T_5	1796
T_6	1681
T_7	1316

Mivel elértük sorrendben az utolsó kezelést (T_7), befejeződött a folyamat, hogy csoportosítsuk egybe mindazokat a kezeléseket, amelyek szignifikánsan nem különböznek egymástól.

5. Lépés. Prezentáljuk a próba eredményét a következő módok egyikével:

1. Használjuk a vonal jelölést, ha a kezeléseket sorrendje az eredmények prezentálásában rangsoraiknak megfelelően rendezhető.

2. Használjuk az alfabetikus jelölést, ha a kezeléseket kívánatos sorrendje az eredmények prezentálásában nem a rangsoraikon alapul. Az alfabetikus jelölés levezethető egyszerűen a vonal jelölésből, jelölve ugyanazzal a betűvel minden kezelésátlagot, melyet ugyanaz a vertikális vonal köt össze. Szokásos használni *a* betűt az első vonalra, *b* betűt a másodikra, *c*-t a harmadikra, és így tovább. Példánkban, *a*, *b*, *c* és *d* jelöli a négy vertikális vonalat az 5.12. táblázat szerint:

Kezelés	Termésátlag, kg ha ⁻¹
T_2	2678 ^a
T_3	2552 ^b
T_4	2128 ^c
T_1	2127
T_5	1796
T_6	1681 ^d
T_7	1316

A végső prezentációt, felhasználva az alfabetikus jelölést, az alábbi táblázat mutatja. Figyeljük meg, hogy egy vagy két betűvel jelölhető minden egyes kezelés. Például, csak egy betűvel, *a*-val jelöljük a T_2 kezelést, míg két betűvel, *ab*-vel jelöljük a T_3 -as kezelést.

DMRT minden lehetséges kezelés-pár összehasonlítására:

Kezelés	Átlagtermés kg ha ⁻¹	DMRT
T ₁	2127	bc
T ₂	2678	a
T ₃	2552	ab
T ₄	2128	bc
T ₅	1796	c
T ₆	1681	cd
T ₇	1316	d

Bármely két átlag, melyet azonos betű követ, szignifikánsan nem különbözik 5%-os szinten

9.5.9. Számítógépes értékelés GenStat programmal

Duncan's multiple range test

Experiment-wise error rate = 0.05

Mean	Mean	
2 2678	2 2678	a
3 2552	3 2552	ab
4 2128	4 2128	bc
1 2127	1 2127	bc
5 1796	5 1796	c
6 1681	6 1681	cd
7 1316	7 1316	d

10. Egytényezős kísérletek varianciaanalízise

Azokat a kísérleteket, amelyekben csak egyetlen tényező (faktor) változik, míg az összes többi tényező konstans (nem változik), *egytényezős kísérleteknek* hívjuk. Ilyen kísérletekben, a kezelések csupán az egyetlen változó faktor különböző szintjeit tartalmazzák. Minden többi faktort egyformán (uniform) alkalmazunk minden parcellára egyetlen előírt szinten. Például, legtöbb *fajtakísérlet* egytényezős kísérlet, amelyben az egyetlen változó faktor a fajta és a faktor szintek (azaz kezelések) a különböző fajták. Kizárólag az elvetett fajták különböznek egyik kísérleti parcellától a másikig és minden agrotechnikai faktort, mint a műtrágya, a károsítók elleni védekezés és a vízgazdálkodás, egyformán alkalmazunk minden parcellára.

További példák az egytényezős kísérletekre:

- *Trágyázási kísérletek*, melyekben pl. egyetlen tápanyag különböző dózisait vizsgáljuk.
- *Növényvédelmi kísérletek*, melyekben különböző herbicidek, inszekticidek és fungicidek hatását vizsgáljuk.
- *Növényszám kísérletek*, melyekben különböző növényszámokat tesztelünk.
- *Vetésidő kísérletek*, melyekben különböző vetésidők hatását vizsgáljuk.

Kétféle kísérleti elrendezés alkalmazható az egytényezős kísérletekre. Az egyik csoport a *komplett blokk elrendezés*, amely kevés kezelést tartalmazó kísérletekre alkalmas és olyan blokkokkal jellemezhető, amelyek mindegyike a kezelések legalább egy teljes sorozatát tartalmazza. A másik a *nem komplett blokk elrendezés*, amely a nagyszámú kezelést tartalmazó kísérletekre alkalmas és olyan blokkokkal jellemezhető, amelyek a vizsgálni kívánt kezeléseknek csak egy részét (frakcióját) tartalmazzák.

Bemutattunk három *komplett blokk-elrendezést* (teljesen randomizált, randomizált komplett blokk, és latin négyzet elrendezés) és két *nem-komplett blokk-elrendezést* (latin-téglá és csoportosított blokk-elrendezés).

10.1. Teljesen véletlen elrendezés (CRD)

A kezeléseket teljesen véletlenül jelöljük ki, úgy hogy minden kísérleti egységnek (parcella) azonos az esélye, hogy bármely kezelést kapja. A variációnak két forrása van: az egyik a kezelés variáció, a másik a kísérleti hiba. A kettő relatív nagyságát használjuk arra, hogy jelezzük, hogy a megfigyelt különbség a kezelések között valós vagy a véletlennek tulajdonítható.

10.1.1. Randomizáció

Egy CRD kísérlet randomizációjának lépésenkénti eljárása egy szántóföldi kísérletben négy kezeléssel (A, B, C és D) és öt ismétléssel a következő:

1. lépés. Határozzuk meg a kísérleti parcellák összes számát (n), mint a kezelések (t) és ismétlések (r) számának szorzatát, azaz $n = (r)(t)$. Példánkban $n = (5)(4) = 20$.

2. lépés. Adjunk egy parcella számot minden kísérleti parcellának egy megfelelő formában, például folyamatosan 1-től n -ig. Példánkban a parcella számokat 1-től 20-ig adtuk meg, mint az 10.1. táblázat mutatja.

3. lépés. A kísérleti parcellákra jelöljük ki a kezeléseket bármelyik alábbi randomizációs eljárással, melyeket Gomez és Gomez (1984) alapján ismertetjük.

A) Véletlen számok táblázatával. A lépések a következők:

A₁ lépés. Jelöljük meg egy kezdőpontot a véletlen számok táblázatban, becsukva szemünket és rámutatva a lap bármelyik pozíciójára. Példánkban, a kezdőpont a hatodik sor és a tizenkettedik (egyedi) oszlop kereszteződése, mint a nyíl mutatja.

Véletlen számok táblázata

14620	95430	12951	81953	17629
09724	85125	48477	42783	70473
56919	17803	95781	85069	61594
97310	78209	51263	52396	82681
07585	28040	2693	64531	70570
		↓		
25950	85189	69374	37904	06759
82937	16405	81497	20863	94072
60819	27364	59081	72635	49180
59041	38475	03615	84093	49731
74208	69516	79530	47649	53046
39412	03642	87497	29735	14308
48480	50075	11804	24956	72182
95318	28749	49512	35408	21814
72094	16385	90185	72635	86259
63158	49753	84279	56496	30618
19082	73645	09182	73649	56823
15232	84146	87729	65584	83641
94252	77489	62434	20965	20247
72020	18895	84948	53072	74573
48392	06359	47040	05695	79799
37950	77387	35495	48192	84518
09394	59842	39573	51630	78548
34800	28055	91570	99154	39603
36435	75946	85712	06293	85621
28187	31824	52265	80494	66428

A₂ lépés. Felhasználva az A₁ lépésben kapott kezdőpontot, jelöljük ki a számokat vertikálisan lefelé, amíg elérjük az $n = 20$ különálló három számjegyű random számot. A három számjegy azért előnyös, mivel kevésbé valószínű, hogy magában foglal azonos értékeket, szemben az egy- és kétjegyű számokkal. Példánkban, a 20 különálló három jegyű random szám a megjelenés sorrendjében a következő:

Véletlen szám	Sorrend	Véletlen szám	Sorrend
937	1	918	11
149	2	772	12
908	3	243	13
361	4	494	14
953	5	704	15
749	6	549	16
180	7	957	17
951	8	157	18
018	9	571	19
427	10	226	20

A₃ lépés. Az A₂ lépésben kapott számokat tegyük növekvő vagy csökkenő rangsorba. Példánkban, a 20 random számot a legkisebbtől a legnagyobbig rangsoroltuk:

Véletlen szám	Sorrend	Rangsor	Véletlen szám	Sorrend	Rangsor
937	1	17	918	11	16
149	2	2	772	12	4
908	3	15	243	13	6
361	4	7	494	14	9
953	5	19	704	15	12
749	6	13	549	16	10
180	7	4	957	17	20
951	8	18	157	18	3
018	9	1	571	19	11
427	10	8	226	20	5

A₄ lépés. Osszuk fel az A₃ lépésben kapott n rangsort t csoportra, mindegyik r számot tartalmaz, abban a sorrendben ahogy a random számok megjelennek. Példánkban, a 20 rangsort négy csoportra osztottuk, mindegyik öt számot tartalmaz:

Csoport szám	Rangsor a csoportban				
1	17	2	15	7	19
2	13	4	18	1	8
3	16	14	6	9	12
4	10	20	3	11	5

A₅ lépés. Jelöljük meg a kezeléseket az n kísérleti parcellán, felhasználva az A₄ lépés csoport számát, mint kezelés számot és a megfelelő rangsort mindegyik csoportban, mint parcella számot, amelyen a megfelelő kezelést bejelöljük. Példánkban, az első csoportot jelöljük meg A kezelésnek és a 17, 2, 15, 7 és 19 parcellaszám kapja ezt

a kezelést; a második csoportot B kezelésnek jelöljük és a 13, 4, 18, 1 és 8 parcellaszám kapja ezt a kezelést; a harmadik csoportot jelöljük C kezelésnek és a 16, 14, 6, 9 és 12 parcellaszám kapja ezt a kezelést; a negyedik csoportot jelöljük D kezelésnek és a 10, 20, 3, 11 és 5 parcellaszám kapja ezt a kezelést. A kísérlet végső elrendezését az 10.1. táblázat mutatja.

10.1. táblázat. A teljesen véletlen elrendezésű (CRD) kísérlet elrendezése
4 kezeléssel és öt ismétléssel

1	2	3	4
B	A	D	B
5	6	7	8
D	C	A	B
9	10	11	12
C	D	D	C
13	14	15	16
B	C	A	C
17	18	19	20
A	B	A	D

1–20: parcella számok, A, B, C és D: a kezelések jelölése

B) Kártya húzásával. A lépések a következők:

B₁ lépés. Egy rendes játékkártya csomagból húzzunk ki n lapot, egyenként, és keverjük meg a megmaradt lapokat minden húzás után. Az eljárás nem használható akkor, amikor a kísérleti egységek összes száma 52-nél több, mivel csak 52 lap van egy csomagban. Példánkban, a 20 kihúzott lap és a lapok kihúzásának sorrendje látható az alábbiakban:

1	2	3	4	5	6	7	8	9	10	Sorszám
J	3	A	Q	K	5	6	9	9	8	
♥	♦	♦	♥	♣	♥	♣	♠	♥	♠	
11	12	13	14	15	16	17	18	19	20	Sorszám
A	4	3	2	7	10	A	4	6	10	
♥	♦	♥	♠	♣	♠	♣	♠	♥	♣	

B₂ lépés. Rangsoroljuk a B₁ lépésben kihúzott 20 lapot a rangnak (♠ ♣ ♥ ♦) megfelelően, és számozzuk be a lapokat (2 a legkisebb, A a legnagyobb). Példánkban a 20 lap rangsorolva van a legkisebttől a legnagyobbig, ahogy következnek:

1	2	3	4	5	6	7	8	9	10	Sorszám
14	7	9	15	5	11	2	19	13	18	Rangsor
J	3	A	Q	K	5	6	9	9	8	
♥	♦	♦	♥	♣	♥	♣	♠	♥	♠	

11	12	13	14	15	16	17	18	19	20	Sorszám
16	8	10	1	3	20	6	17	12	4	Rangsor
A	4	3	2	7	10	A	4	6	10	
♥	♦	♥	♣	♣	♠	♣	♠	♥	♣	

B₃ lépés. Jelöljük meg a t kezelést az n parcellán, felhasználva a B₂ lépésben kapott rangsort, mint parcella számot. Kövessük az A₄ és A₅ lépések eljárását. Példánkban a négy kezelést a következőképpen jelöltük ki a 20 parcellára:

Kezelés	Parcella kijelölés				
A	14	7	9	15	5
B	11	2	19	13	18
C	16	8	10	1	3
D	20	6	17	12	4

C) Sorshúzással. A lépések a következők:

C₁ lépés. Készítsünk elő n azonos nagyságú papírdarabot és osszuk fel t csoportra, mindegyik csoportban legyen r papírdarab. Címkezzük meg az azonos csoport minden papírdarabját ugyanazzal a betűvel (vagy számmal) egy kezelésnek megfelelően. Egységesen hajtsuk össze az n megcímkézett papírdarab mindegyikét, keverjük össze alaposan, tegyük be egy tartóba. Példánkban, 20 papírdarab van, öt-öt darab A, B, C és D megjelöléssel.

C₂ lépés. Húzzunk ki egy-egy papírdarabot egyenként, anélkül, hogy pótolnánk és minden húzás után keverjük össze a papírdarabokat. Példánkban, a kezelés megjelölés és a papírdarabok kihúzásának sorrendje a következő lehet:

Kezelés megjelölés	D	B	A	B	C	A	D	C	B	D
Sorrend	1	2	3	4	5	6	7	8	9	10
Kezelés megjelölés	D	A	A	B	B	C	D	C	C	A
Sorrend	11	12	13	14	15	16	17	18	19	20

C₃ lépés. Jelöljük ki a kezeléseket a parcellákra a megfelelő kezelés megjelölés és sorrend alapján, mint ahogy a C₂ lépésben leírtuk. Példánkban, az A kezelést kapják a 3, 6, 12, 13 és 20 parcella számok; a B kezelést a 2, 4, 9, 14 és 15 számú parcellák; a C kezelést az 5, 8, 16, 18 és 19 számú parcellák; és a D kezelést az 1, 7, 10, 11 és 17 számú parcellák kapják.



10.1. kép. A gyomnövények elleni posztemergens gyomirtási kísérlet ICIA 0051 (SC9925) herbicid-kombinációkkal a martonvásári növénytermesztési kísérleti területen

2. lépés. A teljesen véletlen elrendezés analízisének általános formája t kezeléssel és kezelésként r ismétléssel az alábbi:

Variáció forrása	SS	d.f.	MS
Kezelés	$\frac{[\sum(x_1)]^2 + [\sum(x_2)]^2 + \dots + [\sum(x_t)]^2}{r} - \frac{[\sum(x)]^2}{tr}$	$t - 1$	
Hiba	$\sum(x^2) - \frac{[\sum(x_1)]^2 + [\sum(x_2)]^2 + \dots + [\sum(x_t)]^2}{r}$	$t(r-1)$	s^2
Összes	$\sum(x^2) - \frac{[\sum(x)]^2}{tr}$	$tr-1$	

3. lépés. Meghatározzuk a szabadságfokot (df) minden variáció forrásra. Jelölje t a kezelések számát és r az ismétlések számát:

$$\text{Összes d.f.} = (r)(t) - 1 = (4)(14) - 1 = 55$$

$$\text{Kezelés d.f.} = t - 1 = 14 - 1 = 13$$

$$\text{Hiba d.f.} = t(r-1) = 14(4-1) = 42$$

A hiba d.f. kivonással is meghatározható:

$$\text{Hiba d.f.} = \text{Összes d.f.} - \text{Kezelés d.f.} = 55 - 13 = 42$$

Állítsuk össze a variancia-táblázatot a következőképpen:

Variáció forrása	Szabadságfok (df)	Eltérés négyzet-összeg (SS)	Közepes négyzetes eltérés (MS)	Számított F-érték	Táblázati F-érték	
					5%	1%
Kezelés	13					
Kísérleti hiba	42					
Összes	55					

4. lépés. Jelölje X_i az i -edik parcella adatát (mérési eredményt), T_i az i -edik kezelés összegét és n a kísérleti parcellák számát [(azaz $n = (r)(t)$)], kiszámítjuk a korrekciós tényezőt (C.F.) és a különböző eltérésnégyzet-összegeket (SS):

$$\text{Korrekciós tényező (C.F.)} = \frac{G^2}{n}$$

ahol G = a kísérlet főösszege

$$\text{Összes SS} = \sum_{i=1}^n X_i^2 - \text{C.F.}$$

$$\text{Kezelés SS} = \frac{\sum_{i=1}^t T_i^2}{r} - \text{C.F.}$$

$$\text{Hiba SS} = \text{Összes SS} - \text{Kezelés SS}$$

Példánkban, a T és G értékeket felhasználva a 10.2. táblázatból, az eltérésnégyzet-összegeket (SS) a következőképpen számítjuk ki:

$$\text{C.F.} = \frac{(373.3)^2}{(4)(14)} = 2488.4$$

$$\text{Összes SS} = \left[(7.61)^2 + (6.74)^2 + \dots + (6.44)^2 \right] - 2488.4 = 69.6$$

$$\text{Kezelés SS} = \frac{(29.46)^2 + (14.5)^2 + \dots + (25.38)^2}{4} - 2488.4 = 55.35$$

$$\text{Hiba SS} = 69.6 - 55.35 = 14.25$$

5. lépés. Kiszámítjuk a közepes négyzetes eltéréseket (MS) a variáció minden forrására oly módon, hogy mindegyik SS-t elosztjuk a megfelelő d.f. értékkel:

$$\text{Kezelés MS} = \frac{\text{Kezelés SS}}{t-1} = \frac{55.35}{13} = 4.26$$

$$\text{Hiba MS} = \frac{\text{Hiba SS}}{t(r-1)} = \frac{14.25}{(14)(3)} = 0.34$$

6. lépés. Kiszámítjuk az F-értéket a kezeléskülönbség szignifikanciájának tesztelésére:

$$F = \frac{\text{Kezelés MS}}{\text{Hiba MS}} = \frac{4.26}{0.34} = 12.53$$

Jegyezzük meg, hogy az F-értéket csak akkor érdemes kiszámítani, ha a hiba d.f. elég nagy a hiba variancia megbízható becslésére. Általános irányelvként, az F-értéket csak akkor számítsuk ki, ha a hiba d.f. hat, vagy annál nagyobb.

7. lépés. Kikeressük a táblázati F-értéket: $f_1 = \text{kezelés d.f.} = (t-1)$ és $f_2 = \text{hiba d.f.} = t(r-1)$. Példánkban a táblázati F érték $f_1 = 13$ és $f_2 = 42$ szabadságfokkal 1.99 az 5%-os szignifikancia szinten és 2.64 az 1%-os szignifikancia szinten.

8. lépés. A kiszámított értékeket beírjuk a varianciatáblázatba. Példánk eredményét a 10.3. táblázat tartalmazza.

10.3. táblázat. A kukorica gyomnövényei elleni posztemergens herbicid-kísérlet variancia-táblázata (CRD, egyenlő ismétléssel)

Variáció forrása	Szabadságfok (df)	Eltérés négyzetösszeg (SS)	Közepes négyzetes eltérés (MS)	Számított F-érték	Táblázati F-érték	
					5%	1%
Kezelés	13	55.35	4.26	12.53**	1.99	2.64
Kísérleti hiba	42	14.25	0.35			
Összes	55	69.6				

CV = 8.87% ** = szignifikáns 1%-os szinten

9. lépés. Hasonlítsuk össze a kiszámított F-értéket a táblázati F-értéssel és határozzuk meg a kezelések közötti különbségek szignifikanciáját a következő szabályok szerint.

Ha a kiszámított F-érték nagyobb, mint a táblázati F-érték 1%-os szignifikancia szinten, akkor a kezeléskülönbségről azt mondjuk, hogy *erősen szignifikáns*. Ezt általában két csillaggal jelöljük a számított F-értéken a varianciatáblázatban.

Ha a kiszámított F-érték nagyobb, mint a táblázati F-érték 5%-os szignifikancia szinten, de kisebb vagy egyenlő a táblázati F-értéssel a szignifikancia 1%-os szintjén, akkor a kezeléskülönbségről azt mondjuk, hogy *szignifikáns*. Az ilyen eredményt egy csillaggal jelöljük a számított F-értéken a variancia táblázatban.

Ha a számított F-érték kisebb vagy egyenlő a táblázati F-értéssel a szignifikancia 5%-os szintjén, akkor a kezelés-különbségről azt mondjuk, hogy *nem szignifikáns*. Az ilyen eredményt ^{ns}-sel jelöljük a számított F-értéken a variancia-táblázatban.

Jegyezzük meg, hogy a nem szignifikáns F-teszt a varianciaanalízisben azt jelzi, hogy a kísérlet eredménytelen volt abban, hogy különbséget mutasson ki a kezelések között. Semmi esetre sem bizonyítja azt, hogy minden kezelés (hatása) ugyanaz, mivel az F vizsgálaton alapuló nem szignifikáns kezeléskülönbség oka lehet nagyon kicsi vagy nulla kezeléskülönbség vagy nagyon nagy kísérleti hiba, vagy mindkettő.

Tehát, bármikor, ha az F-teszt nem szignifikáns, a kutatónak meg kell vizsgálni a kísérleti hiba nagyságát és a kezeléscsoportok közötti számszerű különbséget. Ha mindkét érték nagy, a kísérletet meg kell ismételni és erőfeszítéseket szükséges tenni a kísérleti hiba csökkentésére, úgy, hogy a kezelések közötti különbséget, ha van, ki lehessen mutatni. Másfelől, ha mindkét érték kicsi, a kezelések közötti különbség valószínűleg túl alacsony ahhoz, hogy gazdasági jelentősége legyen, így új kísérlet beállítása nem szükséges.

Példánkban a számított F-érték 12.53, nagyobb, mint a táblázati 2.64 F-érték, 1%-os szignifikancia szinten. Ha a kezeléskülönbség erősen szignifikáns, ez azt jelenti, hogy 100 közül 1-nél kisebb a valószínűsége annak, hogy a kezelések között megfigyelt különbségeket a véletlennek lehetne tulajdonítani. Meg kell jegyezni, hogy a szignifikáns F-érték azt igazolja, hogy van különbség bizonyos vizsgált kezelések között, azonban nem specifikálja azt vagy azokat a kezelés-párokat, amelyek szignifikánsan különböznek. Ez utóbbi az SzD (szignifikáns differencia) érték meghatározásán alapszik.

10. lépés. Számítsuk ki a főátlagot és a variációs koefficiens (CV):

$$\text{Főátlag} = \frac{G}{n} \quad \text{és} \quad CV\% = \frac{\sqrt{\text{Hiba MQ}}}{\text{fő átlag}} \times 100$$

$$\text{Példánkban a főátlag} = \frac{373.3}{56} = 6.67 \quad \text{és} \quad CV\% = \frac{\sqrt{0.35}}{6.67} \times 100 = 8.87$$

A CV jelöli a pontosság mértékét, amellyel a kezeléseket összehasonlítjuk és a kísérlet megbízhatóságának az indexe. Kifejezi a kísérleti hibát az átlag százalékában, tehát minél nagyobb a CV érték, annál kisebb a kísérlet megbízhatósága. A CV értéket általában a variancia-táblázat alá írjuk.

A CV jelentősen ingadozik a kísérlet típusától, a termesztett növénytől és a mért tulajdonságtól függően. Egy tapasztalt kutató megfelelő becslését adhatja egy bizonyos CV érték elfogadhatóságának egy adott típusú kísérletre. Tapasztalatok szerint a gabonafélék termésére a CV elfogadható tartománya fajtakísérletekben 6-8%, trágyázási kísérletekben 10-12%, inszekticid és herbicid kísérletekben 13-15%. Más növénytulajdonságokra a CV jelentősen különbözik a termésétől. Például szabadföldi kísérletekben a CV rizstermesre 10% körüli, hajtások számára 20% és növénymagasságra 3%.

Szignifikáns differencia kiszámítása (egyenlő ismétlésszám)

A teljesen randomizált elrendezésű kísérletben 14 kezelést (12 herbicid kezelést és két kontroll kezelést) vizsgáltunk négy ismétlésben. A kísérlet célja volt meghatározni a herbicid kezelések hatását a gyomos kontrollhoz viszonyítva, továbbá azonosítani olyan herbicid-kombinációkat, amelyeknél a kukorica termése szignifikánsan nagyobb volt a 3. kezelésnél (sulcotrione herbicid önmagában). Ebben a példában a megfelelő összehasonlítás a tervezett összehasonlítások sorozata, amelyben mindegyik herbicid kezelés átlagát összehasonlítjuk a gyomos kontrollal, illetve a 3. kezelés átlagával.

Az SzD vizsgálat lépései a következők:

1. lépés. Számítsuk ki az átlagos különbséget a 3. kezelés és a többi herbicid kezelés mindegyike között, mint a 10.4. táblázat mutatja.

2. lépés. Számítsuk ki az SzD értékét α szignifikancia szinten:

$$SzD_{\alpha} = t_{\alpha} \sqrt{\frac{2s^2}{r}}$$

Példánkban a hiba MS, vagyis $s^2 = 0.35$, a hiba szabadságfok 42, és az ismétlések száma 4. A táblázati t érték $n = 42$ szabadságfoknál 2.021 az 5%-os és 2.704 az 1%-os szignifikancia szinten. A számított SzD értékek:

$$SzD_{5\%} = 2.021 \sqrt{\frac{2(0.35)}{4}} = 0.845 \text{ t ha}^{-1}$$

$$SzD_{1\%} = 2.704 \sqrt{\frac{2(0.35)}{4}} = 1.131 \text{ t ha}^{-1}$$

10.4. táblázat. A 3. kezelés és a többi herbicid kezelés átlagtermései közötti összehasonlítás, felhasználva az SzD vizsgálatot

Herbicid kezelés	Átlagtermés t ha ⁻¹	Eltérés a 3. herbicid kezeléstől t ha ⁻¹	Eltérés a gyomos kontrolltól
1. Kapált kontrol	7.37	–	3.74**
2. Gyomos kontrol	3.63	–	–
3. ICIA-0051	6.43	–	2.8**
4. ICIA-0051	6.93	0.5 ^{ns}	3.3**
5. ICIA-0051+Hungazin PK	7.42	0.99*	3.79**
6. ICIA-0051+Hungazin PK	7.30	0.87*	3.67**
7. ICIA-0051+Lentagran EC	7.50	1.07*	3.87**
8. ICIA-0051+Lentagran EC	7.62	1.19*	3.99**
9. ICIA-0051+Basagran	6.7	0.27 ^{ns}	3.07**
10. ICIA-0051+Basagran	7.08	0.65 ^{ns}	3.45**
11. ICIA-0051+Banvel 480	6.38	-0.05 ^{ns}	3.33**
12. ICIA-0051+Banvel 480	6.96	0.53 ^{ns}	3.33**
13. ICIA-0051+CGA136 872	5.66	-0.77 ^{ns}	2.03**
14. ICIA-0051+CGA136 872	6.35	-0.08 ^{ns}	2.72**

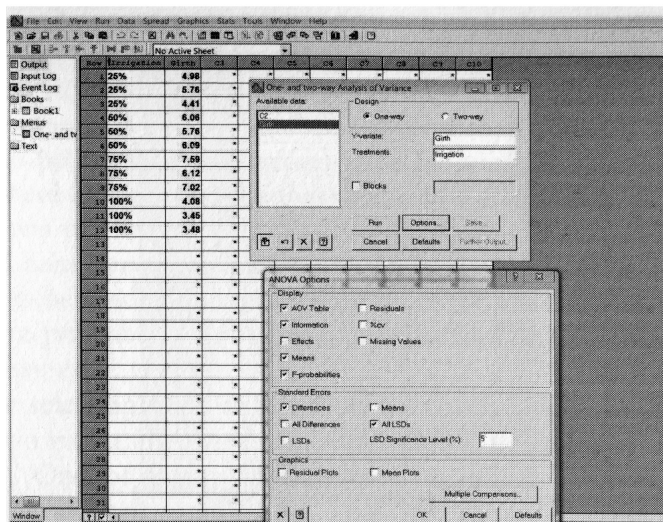
**= szignifikáns 1%-os szinten, *= szignifikáns 5%-os szinten, ^{ns} = nem szignifikáns

3. lépés. Összehasonlítjuk a kezelések kontrolltól való eltérését az SzD értékével és bejelöljük a szignifikáns differenciát megfelelő csillagokkal. Példánkban mindegyik

herbicid kezelés 1%-os szinten szignifikáns terménynövekedést eredményezett a gyomos kontrollhoz viszonyítva. Megjegyezzük, hogy a terménynövekedés 0.1%-os szinten is szignifikáns volt ($SzD_{0,1\%}$ értéke 1.459). A 3. kezeléshez viszonyítva az 5-8. herbicid-kombinációk 5%-os szinten szignifikáns terménynövekedést eredményeztek. A kapált kontrollhoz viszonyítva – a gyomos kontrollon kívül – a 3. és 11. kezelésben volt 5%-os szignifikancia szinten kisebb a termés.

10.1.5. Varianciaanalízis GenStat programmal

Abból a célból, hogy lefuttassuk a varianciaanalízist (ANOVA) a kísérleti adatokat be kell vinni az egyik oszlopba és egy külön oszlopba azt a faktort, melyet az adatok csoportosítására használunk. Az adatokat közvetlenül bevitethetjük a GenStat számoló-táblába vagy az Excel táblázatkezelőbe és ezután fájlként átvisszük a GenStat-ba (vagy másolás-beillesztés úton).



10.1. ábra. GenStat számológéptábla és menü dobozok az egytényezős varianciaanalízishez (ANOVA)

Egy teljesen randomizált egytényezős ANOVA futtatásához a GenStat számlá-lótáblát úgy kell kialakítani, mint a 10.1. ábra mutatja. Figyeljük meg, hogy az első oszlop a csoportosítási faktor, melyet piros felkiáltójel mutat az oszlop címében; a második oszlop tartalmazza az adatokat. Az egy- és kéttényezős ANOVA (One- and Two-way ANOVA) adatbeviteli dobozát megkapjuk az Analysis of Variance választási lehetőségen át a Stats menüben és a One-way ANOVA kiválasztásával. Az adatokat tartalmazó címkét ezután bevisszük az Y-variáns dobozba és a csoportosítási faktort tartalmazó oszlop címkét a Treatments dobozba. Ha szükséges, kiválaszthatunk egy options dobozt, amely további statisztikát kínál, mint pl. a szignifikáns különbség (LSD) és a többszörös összehasonlító tesztek (multiple comparisons). Az analízist a Run gombra kattintással futtatjuk le és az analízis eredménye megjelenik a GenStat

Output Window-ban. Az egytényezős CRD kísérlet varianciaanalízisét lefuttathatjuk az *Analysis of Variance* menüből a *General* kiválasztásával és a legördülő lehetőségek közül *One-way ANOVA (no Blocking)* kiválasztásával.

Az adatok bevitelekor az ismétlést és a kezelést faktornak tekintjük, a termést változónak. Az ismétlés faktornak négy, a kezelés faktornak 14 szintje van.

Kezelés	Kukorica termés
1	7.61
2	2.9
3	6.44
4	7.3
5	7.32
6	7.43
7	6.78
8	7.77
.	.
1	6.74
2	3.68
3	6.14
4	6.88
.	.
2	3.34
3	6.49
4	6.78
.	.
11	7.32
12	7.42
13	4.57
14	6.44

A varianciaanalízis számítógépes outputja a következő:

Analysis of variance

Variate: Termes

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Kezeles	13	55.6308	4.2793	12.58	<.001
Residual	42	14.2888	0.3402		
Total	55	69.9196			

Tables of means

Variate: Termes

Grand mean 6.666

Kezeles	1	2	3	4	5	6	7
	7.365	3.625	6.425	6.925	7.420	7.303	7.502
Kezeles	8	9	10	11	12	13	14
	7.622	6.700	7.082	6.383	6.963	5.660	6.345

Standard errors of differences of means

Table	Kezeles
rep.	4
d.f.	42
s.e.d.	0.4124

Least significant differences of means (5% level)

Table	Kezeles
rep.	4
d.f.	42
l.s.d.	0.8323

Duncan's multiple range test**Kezeles**

	Mean	
2	3.625	a
13	5.660	b
14	6.345	bc
11	6.383	bc
3	6.425	bcd
9	6.700	cde
4	6.925	cde
12	6.963	cde
10	7.082	cde
6	7.303	cde
1	7.365	de
5	7.420	e
7	7.502	e
8	7.622	e

10.1.6. Nem egyenlő ismétlésszám

Mivel a számítási eljárás a CRD-re nem túlságosan komplikált, amikor az ismétlések száma különbözik a kezeléseik között, a CRD-t általában használják olyan tanulmányokban, ahol a kísérleti anyag nehézkessé teszi egyenlő számú ismétlés használatát minden kezelésre. Néhány példa ezekre az esetekre:

- Állattetési kísérletek, ahol az állatok száma az egyes tényezetekben nem azonos.
- Kísérletek a rovarcsapdában elfogott különböző rovarfajok testhosszának összehasonlítására.
- Kísérletek, amelyeket eredetileg egyenlő számú ismétlésben állítottak be, azonban egyes kísérleti egységek (parcellák) megsemmisültek a kísérletezés során.



10.2. kép. Klóracetanilid herbicidek gyomirtó és fitotoxikus hatásának szakértői szemléje a martonvásári tenyészkertben 1987-ben

10.1.7. Kísérlet leírása

Az MTA Mezőgazdasági Kutatóintézetében, Martonvásáron 1987-ben kisparcellás szántóföldi kísérletben vizsgáltuk különböző klóracetanilid herbicidek gyomirtó és fitotoxikus hatását, illetve mértük hatásukat a kukorica szemtermésére (10.2. kép). A preemergens herbicid kísérletben a propaklór, alaklór, metolaklór és acetoklór herbicidek hatását hasonlítottuk össze 1, 2, 4, 8 és 16 l ha⁻¹ dózisban. A kísérletben a kapált és gyomos kontrollal együtt 22 kezelés hatását vizsgáltuk négy ismétlésben és teljesen véletlen (CRD) elrendezésben (Berzsenyi *et al.*, 1994). Korábbi fitotron kísérletekben megállapítottuk, hogy a klóracetanilid herbicidek fitotoxikus hatása nő a következő sorrendben: propaklór, alaklór, metolaklór és acetoklór (Berzsenyi és Gyórfy, 1989). A kísérlet 4. ismétlésében vadkártétel miatt 4 parcellát ki kellett zárunk az értékelésből.

10.5. táblázat. Kukorica preemergens herbicid kísérlet terméseredményei különböző hatóanyagú és dózisú herbicid- kezelésekkel (CRD kísérlet nem egyenlő számú ismétléssel)

Kezelés	Dózis kg ha ⁻¹	Kukorica szemtermése t ha ⁻¹				Kezelés Összeg (T)	Kezelés átlag
		Ismétlés					
		I.	II.	III.	IV.		
1. Kapált kontroll	–	9.612	9.572	9.049	8.863	37.096	9.274
2. Gyomos kontroll	–	1.732	2.965	1.909	1.985	8.591	2.148
3. Propaklór	1	4.742	6.411	3.104	5.584	19.841	4.960
4.	2	5.343	6.187	5.049	*	16.579	5.526
5.	4	8.43	8.534	10.544	6.299	33.807	8.452
6.	8	10.178	7.205	7.143	9.846	34.372	8.593

A 10.5. táblázat folytatása

Kezelés	Dózis kg ha ⁻¹	Kukorica szemtermése t ha ⁻¹				Kezelés Összeg (T)	Kezelés átlag
		Ismétlés					
		I.	II.	III.	IV.		
7.	16	9.911	9.028	9.522	9.185	37.646	9.412
8. Alaklór	1	7.316	5.528	5.992	7.023	25.859	6.465
9.	2	7.963	8.333	5.234	7.157	28.687	7.172
10.	4	7.284	7.735	7.56	*	22.579	7.526
11.	8	9.565	8.855	7.526	5.54	31.486	7.872
12.	16	9.718	7.339	10.784	6.938	34.779	8.695
13. Metolaklór	1	7.673	6.323	5.192	5.989	25.177	6.295
14.	2	8.855	6.226	6.603	6.37	28.054	7.014
15.	4	8.449	7.286	8.245	*	23.98	7.993
16.	8	7.698	7.936	6.227	5.898	27.759	6.940
17.	16	6.79	7.77	7.874	6.195	28.629	7.158
18. Acetoklór	1	8.121	8.117	8.622	8.018	32.878	8.220
19.	2	7.502	8.533	9.253	*	25.288	8.429
20.	4	7.826	7.585	9.492	6.488	31.391	7.848
21.	8	6.703	8.149	9.828	7.385	32.065	8.016
22.	16	7.26	7.063	7.504	7.574	29.401	7.350
Főösszeg (G)						615.9	
Főátlag							7.334

*A hiányzó adatok jelölése

10.1.8. Varianciaanalízis lépései

A varianciaanalízis lépései a CRD elrendezésnél, nem egyenlő számú ismétlésnél az alábbiak.

1. lépés. Csoportosítsuk az adatokat kezelésként és számítsuk ki a kezelés összeget (T) és a főösszeget (G). Az adatokat a 10.5. táblázat tartalmazza.

2. lépés. Állítsuk össze a variancia-táblázatot a következőképpen:

Variáció forrása	Szabadságfok (df)	Eltérés négyzet-összeg (SS)	Közepes négyzetes eltérés (MS)	Számított F-érték	Táblázati F-érték	
					5%	1%
Kezelés						
Kísérleti hiba						
Összes						

3. lépés. Jelölje t a kezeléseket számát és n az összes megfigyelés (parcella) számát, a szabadságfokok számát a variáció minden forrására az alábbiak szerint határozzuk meg:

$$\text{Összes d.f.} = n - 1 = 84 - 1 = 83$$

$$\text{Kezelés d.f.} = t - 1 = 22 - 1 = 21$$

$$\text{Hiba d.f.} = \text{Összes d.f.} - \text{Kezelés d.f.} = 83 - 21 = 62$$

4. lépés. A 4. táblázat kezelésösszegének (T) és főösszegének (G) felhasználásával számítsuk ki a korrekciós tényezőt és a különböző eltérésnégyzet összegeket (SS) a következőképpen:

$$C.F. = \frac{G^2}{n} = \frac{(615.9)^2}{84} = 4515.87$$

$$\text{Összes SS} = \sum_{i=1}^n X_i^2 - C.F. = [(9.612)^2 + (9.572)^2 + \dots + (7.574)^2] - 4515.87 = 291.13$$

Kezelés SS =

$$= \sum_{i=1}^4 \frac{T_i^2}{r_i} - C.F. = \left[\frac{(37.096)^2}{4} + \dots + \frac{(16.579)^2}{3} + \dots + \frac{(29.401)^2}{4} \right] - 4515.87 = 212.13$$

$$\text{Hiba SS} = \text{Összes SS} - \text{Kezelés SS} = 291.13 - 212.13 = 79.0$$

5. lépés. Kövessük „CRD egyenlő ismétléssel” c. fejezet 5–10. lépését. Példánkra a teljes varianciaanalízist a 5. táblázat tartalmazza. Az F-vizsgálat eredménye erősen szignifikáns különbséget mutat a kezelések között.

10.6. táblázat. A herbicid kísérlet (10.5. táblázat) termésadatainak varianciaanalízise (CRD nem egyenlő számú ismétléssel)

Variáció forrása	Szabadságfok (df)	Eltérés négyzet-összeg (SS)	Közepes négyzetes eltérés (MS)	Számított F-érték	Táblázati F-érték	
					5%	1%
Kezelés	21	212.13	10.101	7.93**	3.47	4.53
Kísérleti hiba	62	79	1.274			
Összes	83	291.13				

** = szignifikáns 1%-os szinten

Szignifikáns differencia kiszámítása (nem egyenlő ismétlésszám)

Egy teljesen véletlen elrendezésű kísérletben 22 kezelés volt (20 gyomirtási kezelés és két kontroll) és nem egyenlő ismétléssel. Meghatározni kívántuk, hogy a 20 gyomirtási kezelés bármelyike jobb-e a kapált kontroll kezelésnél. Ebben a példában a megfelelő összehasonlítás a tervezett pár-összehasonlítás, amelyben a 20 gyomirtási kezelés mindegyikét összehasonlítjuk a kapált kontroll kezeléssel. A lépések a következők:

1. lépés. Számítsuk ki az átlagos különbséget a kontroll kezelés és a 20 gyomirtási kezelés mindegyike között, mint ahogy a 10.7. táblázat mutatja.

2. lépés. Számítsuk ki az SzD értéket α szignifikancia szinten:

Mivel bizonyos kezeléseknél négy ismétlése van és más kezeléseknél három, az SzD értékeknek két sorozatát kell kiszámolni. Felhasználjuk a hiba MS, vagyis s^2 érté-

ket (1.274), a hiba szabadságfokot (62) és a táblázati t értéket 62 szabadságfoknál, amely 2.00 az 5%-os és 2.66 az 1%-os szignifikanciaszinten. Ezekből az adatokból a kétféle SzD értéket a következőképpen számítjuk ki:

- A kontroll kezelés (négy ismétlés) és mindegyik négy ismétlésben beállított gyomirtási kezelés összehasonlítására az SzD értékét az alábbi módon számítjuk ki:

$$SzD_{5\%} = 2.00 \sqrt{\frac{2(1.274)}{4}} = 1.596 \text{ t ha}^{-1}$$

$$SzD_{1\%} = 2.66 \sqrt{\frac{2(1.274)}{4}} = 2.123 \text{ t ha}^{-1}$$

- A kontroll kezelés (négy ismétlésben) és minden három ismétlésben beállított gyomirtási kezelés összehasonlítására az SzD értékét a következőképpen számítjuk ki:

$$SzD_{\alpha} = (t_{\alpha})(s_d)$$

$$\text{ahol } s_d = \sqrt{s^2 \left(\frac{1}{r_i} + \frac{1}{r_j} \right)}$$

$$\text{tehát } SzD_{5\%} = 2.00 \sqrt{1.274(1/3 + 1/4)} = 1.719 \text{ t ha}^{-1}$$

$$SzD_{1\%} = 2.66 \sqrt{1.274(1/3 + 1/4)} = 2.287 \text{ t ha}^{-1}$$

3. lépés. Hasonlítsuk össze mindegyik kezelés kapált kontrolltól való eltérését a megfelelő SzD értékével és csillaggal jelöljük a szignifikanciát. A vizsgálat eredményét minden kezelés-párra a 10.7. táblázat tartalmazza, amely jelzi, hogy mely herbicid kezelések termése különbözött szignifikánsan a kontroll kezeléstől. Az eredmények értelmezéséhez a 10.7. táblázat utolsó két oszlopában feltüntettük a kezelésenkénti gyomborítottságot (%) és a fitotoxikus hatást (1–9-es skála alapján). Látható, hogy a propaklór 1 és 2 l ha⁻¹ dózisban szignifikánsan kisebb termést eredményezett a kontrollhoz képest. Az alaklór 1, 2 és 4 l ha⁻¹-es dóziséknél volt szignifikánsan kisebb termés. A metolaklór első két dózisában a nem megfelelő gyomirtó hatás, a magas dózisokban a fitotoxikus hatás miatt csökkent szignifikánsan a termés. Az acetoklór gyomirtó hatása volt a legjobb, 16 l ha⁻¹ dózisban azonban fitotoxikus hatású volt és szignifikáns terméscsökkenést okozott. A gyomos kontrollhoz viszonyítva, minden kezelésben szignifikánsan nagyobb volt a termés.

10.7. táblázat. A klóracetanalid herbicidek hatása a kukorica termésére a gyomirtó és fitotoxikus hatásuktól függően. A táblázatban feltüntettük a kapált kontrolltól való eltérés szignifikanciáját

Kezelés	Dózis kg ha ⁻¹	Ismét- lések száma	Termés- átlag t ha ⁻¹	Eltérés a Kontroll- tól t ha ⁻¹	SzD _{5%} t ha ⁻¹	SzD _{1%} t ha ⁻¹	Gyom- borítás % 60 DAT	Fito- toxikus hatás 30 DAT
1. Kapált kontroll	–	4	9.274	–	–	–	–	–
2. Gyomos kontroll	–	4	2.148	–7.126**	1.596	2.123	84	–
3. Propaklór	1	4	4.960	–4.314**	1.596	2.123	45	2.3
4.	2	3	5.526	–3.748**	1.719	2.287	38	2.5
5.	4	4	8.452	–0.822 ^{ns}	1.596	2.123	24	2.8
6.	8	4	8.593	–0.681 ^{ns}	1.596	2.123	25	2.8
7.	16	4	9.412	0.138 ^{ns}	1.596	2.123	12	3.3
8. Alaklór	1	4	6.465	–2.809**	1.596	2.123	48	2.2
9.	2	4	7.172	–2.102*	1.596	2.123	44	2.8
10.	4	3	7.526	–1.748*	1.719	2.287	18	2.8
11.	8	4	7.872	–1.402 ^{ns}	1.596	2.123	12	3.0
12.	16	4	8.695	–0.579 ^{ns}	1.596	2.123	12	4.0
13. Metolaklór	1	4	6.295	–2.979**	1.596	2.123	44	2.8
14.	2	4	7.014	–2.26**	1.596	2.123	35	2.8
15.	4	3	7.993	–1.281 ^{ns}	1.719	2.287	24	3.8
16.	8	4	6.940	–2.334**	1.596	2.123	23	4.0
17.	16	4	7.158	–2.116*	1.596	2.123	15	6.0
18. Acetoklór	1	4	8.220	–1.054 ^{ns}	1.596	2.123	30	2.8
19.	2	3	8.429	–0.845 ^{ns}	1.719	2.287	13	3.0
20.	4	4	7.848	–1.426 ^{ns}	1.596	2.123	11	4.0
21.	8	4	8.016	–1.258 ^{ns}	1.596	2.123	4	5.3
22.	16	4	7.350	–1.924*	1.596	2.123	3	6.5

DAT: napok száma a kelés után. SzD_{5%}: gyomborítotttság: 14%, fitotoxikus hatás: 0.9. Fitotoxikus hatás: 1–9 skála alapján.

10.1.9. Varianciaanalízis GenStat programmal

A kísérleti adatok bevitele és értékelése azonos a CRD egyenlő ismétléssel fejezet-
részben leírtakkal. A számítógépes értékelés eredménye az alábbi:

Analysis of variance

Variate: Termes

Source of variation	d.f.	(m.v.)	s.s.	m.s.	v.r.	F pr.
Kezeles	21		216.767	10.322	8.10	<.001
Residual	62	(4)	78.987	1.274		
Total	83	(4)	290.814			

Tables of means

Variate: Termes

Grand mean 7.33

Kezel2	1	2	3	4	5	6	7
	9.27	2.15	4.96	5.53	8.45	8.59	9.41
Kezel2	8	9	10	11	12	13	14
	6.46	7.17	7.53	7.87	8.69	6.29	7.01
Kezel2	15	16	17	18	19	20	21
	7.99	6.94	7.16	8.22	8.43	7.85	8.02
Kezel2	22						
	7.35						

Standard errors of differences of means

Table	Kezeles
rep.	4
d.f.	62
s.e.d.	0.798

Least significant differences of means (5% level)

Table	Kezel2
rep.	4
d.f.	62
l.s.d.	1.595

Least significant differences of means (1% level)

Table	Kezel2
rep.	4
d.f.	62
l.s.d.	2.121

(Not adjusted for missing values)

10.2. Véletlen blokkelrendezés (RCB)

A véletlen blokkelrendezés (randomized complete block, RCB) az egyik leggyakrabban használt kísérleti elrendezés a mezőgazdasági kutatásban. Különösen alkalmas szántóföldi kísérletekre, ahol a kezelések száma nem nagy és a kísérleti terület előre jelezhető termékenységi gradienssel rendelkezik. Az RCB fő megkülönböztető tulajdonsága az egyenlő méretű blokkok jelenléte, amelyeknek mindegyike tartalmazza az összes kezelést.

10.2.1. Kísérlet leírása

Az MTA Mezőgazdasági Kutatóintézetében, Martonvásáron, Győrffy Béla által 1960-ban beállított Komplex I. jelzésű véletlen blokkelrendezésű kísérletben, négy ismétlésben, 5 növénytermesztési tényező hatását 7 kezeléskombinációban vizsgáltuk a kukorica termésére (Győrffy, 1969; Berzsényi – Győrffy, 1995; Berzsényi – Dang, 2008a) (10.3. kép). A vizsgált tényezők: talajművelés, trágyázás, növényszám, fajta és ápolás. Minden tényezőnek 2 fokozata van, egy minimális és egy optimális. Az 1. kezelésben minden tényező minimális, a 2. kezelésben minden tényező optimális fokozatban van. A 3–7. kezelésben úgy alakítottuk ki a kezeléskombinációkat, hogy egy növénytermesztési tényezőt mindig minimális fokozatban hagytunk, a többi pedig optimális fokozatban. Így a 3. kezelésben a talajművelés, a 4. kezelésben a trágyázás, az 5. kezelésben a növényszám, a 6. kezelésben a fajta és a 7. kezelésben az ápolás van minimumban. Az 1983. évi termésadatokat ismétlésenként az 10.8. táblázat mutatja.



10.3. kép. Többtényezős tartamkísérlet a martonvásári kísérleti területen

A kísérletben vizsgált tényezők és fokozataik a következők:

Növénytermesztési tényező	Minimális fokozat	Optimális fokozat
Talajművelés	Sekélyszántás 12–14 cm-re	Mélyszántás 24–28 cm-re
Trágyázás	Trágyázás nélkül	60 t ha ⁻¹ istállótrágya 4 évenként, N ₂₈₀ P ₁₂₀ K ₁₄₀ évente
Növényszám	35 000 fő ha ⁻¹	70 000 fő ha ⁻¹
Fajta	szabadelvirágzású: Aranyözön és Mindszentpusztai sárga	adott évben két legjobb hibrid
Ápolás	gyenge, késői egyelés és 2 megkésett kapálás	jó, teljes gyommentesség, időben egyelve

10.2.2. Varianciaanalízis lépései

A variabilitásnak három forrása van az RCB elrendezésben: a kezelés, az ismétlés (vagy blokk), és a kísérleti hiba. Figyeljük meg, hogy ez eggyel több, mint a CRD-nél, az ismétlés hozzáadása miatt, amely megfelel a blokkok közötti variabilitásnak. Először a kézi számolás lépéseit, ezt követően a számítógépes varianciaanalízis eredményét mutatjuk be.

1. lépés. Csoportosítsuk az adatokat kezelésként és ismétlésként és számítsuk ki a kezelés összeget (T), az ismétlés összeget (R) és a főösszeget (G), mint azt az 10.8. táblázat mutatja.

10.8. táblázat. Hét különböző kezeléskombináció hatása a kukorica szemtermésére a véletlen blokkelrendezésű kísérletben

Kezelés	Szemtermés t ha ⁻¹				Kezelés összeg (T)	Kezelés átlag
	I.	II.	III.	IV.		
	ismétlés					
1. Minden tényező minimumban	2.43	2.25	1.94	2.4	9.02	2.255
2. Minden tényező optimumban	9.49	9.78	9.25	10.24	38.76	9.69
3. Talajművelés minimumban	9.14	8.11	9.36	9.17	35.78	8.945
4. Trágyázás minimumban	4.95	7.28	4.35	5.12	21.7	5.425
5. Növényszám minimumban	7.06	6.92	7.15	6.82	27.95	6.988
6. Fajta minimumban	3.69	5.42	5.49	5.59	20.19	5.048
7. Növényápolás minimumban	5.44	6.56	7.22	6.1	25.32	6.33
Ismétlés összeg (R)	42.2	46.32	44.76	45.44		
Főösszeg (G)					178.72	
Főátlag						6.383

2. lépés. Készítsük el a variancia táblázatot.

Variáció forrása	Szabadságfok (df)	Eltérés négyzet összeg (SS)	Közepes négyzetes eltérés (MS)	Számított F-érték	Táblázati F-érték	
					5%	1%
Ismétlés						
Kezelés						
Hiba						
Összes						

3. lépés. Felhasználva az ismétlések számát (r), a kezelések számát (t), meghatározzuk a szabadságfokot minden variáció forrásra:

$$\text{Összes d.f.} = rt - 1 = 28 - 1 = 27$$

$$\text{Ismétlés d.f.} = r - 1 = 4 - 1 = 3$$

$$\text{Kezelés d.f.} = t - 1 = 7 - 1 = 6$$

$$\text{Hiba d.f.} = (r - 1)(t - 1) = (3)(6) = 18$$

Figyeljük meg, hogy hasonlóan, mint a CRD-ben, a hiba d.f. kiszámítható kivonással is:

$$\text{Hiba d.f.} = \text{Összes d.f.} - \text{Ismétlés d.f.} - \text{Kezelés d.f.} = 27 - 3 - 6 = 18$$

4. lépés. Számítsuk ki a korrekciós faktort és a különböző négyzetösszegeket (SS) a következőképpen:

$$C.F. = \frac{G^2}{rt} = \frac{(178.72)^2}{(4)(7)} = 1140.74$$

$$\text{Összes SS} = \sum_{i=1}^t \sum_{j=1}^r x_{ij}^2 - C.F. = [(2.43)^2 + (9.49)^2 + \dots + (6.1)^2] - 1140.74 = 161.23$$

$$\text{Ismétlés MS} = \frac{\sum_{j=1}^r R_j^2}{t} - C.F. = \frac{(42.2)^2 + (46.32)^2 + (44.76)^2 + (45.44)^2}{7} - 1140.74 = 1.346$$

$$\text{Kezelés SS} = \frac{\sum_{i=1}^t T_i^2}{r} - C.F. = \frac{(9.02)^2 + \dots + (25.32)^2}{4} - 1140.74 = 150.44$$

$$\begin{aligned} \text{Hiba SS} &= \text{Összes SS} - \text{Ismétlés SS} - \text{Kezelés SS} \\ &= 161.23 - 1.346 - 150.44 = 9.45 \end{aligned}$$

5. lépés. Számítsuk ki az MS értékeket minden variáció forrásra, úgy, hogy az SS értékeket elosztjuk a megfelelő szabadságfokok számával:

$$\text{Ismétlés MS} = \frac{\text{Ismétlés SS}}{r-1} = \frac{1.346}{3} = 0.4488$$

$$\text{Kezelés MS} = \frac{\text{Kezelés SS}}{t-1} = \frac{150.44}{6} = 25.07$$

$$\text{Hiba MS} = \frac{\text{Hiba SS}}{(r-1)(t-1)} = \frac{9.45}{18} = 0.525$$

6. lépés. Számítsuk ki az F értékeket a kezelés-különbségek vizsgálatára:

$$F = \frac{\text{Kezelés MS}}{\text{Hiba MS}} = \frac{25.07}{0.525} = 47.77$$

7. lépés. Hasonlítsuk össze a számított F-értéket a táblázati F-értékkal $f_1 = \text{kezelés d.f.}$ és $f_2 = \text{hiba d.f.}$ értéknél. Példánkban a táblázati F-érték $f_1 = 6$ és $f_2 = 18$ szabadságfoknál 3.66 az 5%-os szignifikancia szinten, és 4.01 az 1%-os szinten. Mivel a számított F-érték (47.77) nagyobb, mint a táblázati F-érték 1%-os szinten, levonjuk azt a következtetést, hogy a kísérlet $P = 1\%$ -os szignifikáns különbséget mutatott ki a hét kezelés között.

8. lépés. Számítsuk ki a variációs koefficiens:

$$CV = \frac{\sqrt{\text{Hiba MS}}}{\text{fő átlag}} \times 100 = \frac{\sqrt{0.5249}}{6.383} \times 100 = 11.35\%$$

9. lépés. Írjuk be a 3–8. lépésben kiszámított értékeket a varianciaanalízis táblázatba (10.9. táblázat).

10.9. táblázat. A véletlen blokk (RCB) elrendezésben beállított kukorica kísérlet (Komplex I.) termésadatainak varianciaanalízise

Variáció forrása	Szabadságfok (df)	Eltérés négyzet összeg (SS)	Közepes négyzetes eltérés (MS)	Számított F-érték	Táblázati F-érték	
					5%	1%
Ismétlés	3	1.346	0.449			
Kezelés	6	150.44	25.073	47.77**	3.66	4.01
Hiba	18	9.448	0.525			
Összes	27	161.23				

cv = 11.35%, ** = $P=1\%$ -os szinten szignifikáns

10. Lépés. Szignifikáns különbség kiszámítása:

A véletlen blokkelrendezésű kísérletben az átlagos különbség szórását bármely kezelésátlag párra a következőképpen számítjuk ki:

$$s_d = \sqrt{\frac{2s^2}{r}} = \sqrt{\frac{2(0.525)}{4}} = 0.512$$

ahol r az ismétlések száma, s_d a hiba MS a variancia táblázatból.

Számítsuk ki az SzD értéket a szignifikancia α szintjén: $SzD_\alpha = t_\alpha \sqrt{\frac{2s^2}{r}}$

Példánkra a hiba közepes négyzetes eltérés 0.525, a hiba szabadságfoka 18, és az ismétlések száma négy. A táblázati t érték $n = 18$ szabadságfokra 2.101 az 5%-os és 2.878 az 1%-os szignifikancia szinten. Az SzD értékek a következők:

$$SzD_{0.05} = 2.101 \sqrt{\frac{2(0.525)}{4}} = 1.076$$

$$SzD_{0.01} = 2.878 \sqrt{\frac{2(0.525)}{4}} = 1.474$$

Közzöljük a kezelésenkénti átlagterméseket az $SzD_{5\%}$ értékkel:

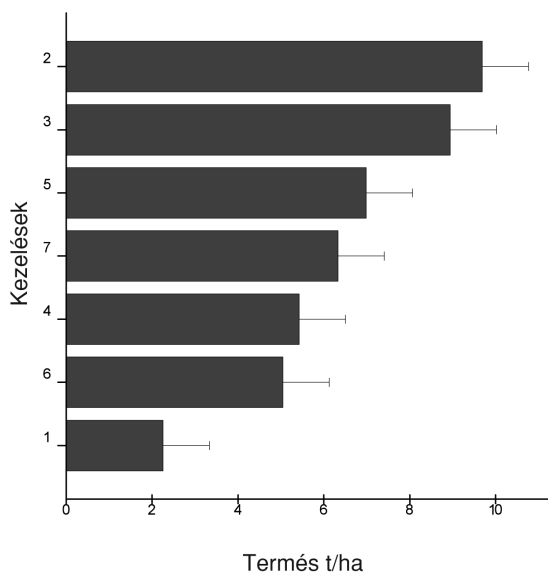
Kezelés	Termésátlag t ha ⁻¹
1. Minden tényező minimumban	2.26
2. Minden tényező optimumban	9.69
3. Talajművelés minimumban	8.95
4. Trágyázás minimumban	5.43
5. Növényszám minimumban	6.99
6. Fajta minimumban	5.05
7. Növényápolás minimumban	6.33
SzD _{5%}	1.076

Az eredmények értelmezése

Az SzD érték, illetve a Duncan többszörös terjedelem vizsgálat (l. számítógépes output) alapján megállapítható, hogy a 4. és a 6. kezelés szignifikánsan nem különbözik egymástól (ezt a Duncan tesztnél azonos betűk előfordulása jelzi), a többi kezelés között szignifikáns különbség van.

A növénytermesztési tényezők hatását a kukorica termésére az 10.2. ábra szemlélteti. Minden tényező minimális, illetve optimális szintjének összehasonlításakor legnagyobb a különbség az átlagos termésreakcióban (7.43 t ha⁻¹). Egy-egy termesztési tényező minimális szintjének hatását vizsgálva (a többi optimuma esetén), megállapítható, hogy a talajművelés mélységének hatása kismértékű (0.74 t ha⁻¹ termésnövekedés az optimumhoz képest). Megállapítható továbbá, hogy amikor a trágyázás, illetve

a genotípus a minimumban levő tényező, a terméscsökkenés – az optimumhoz viszonyítva – jelentős $> 4 \text{ t ha}^{-1}$ (átlagos termésreakció 5.43, illetve 5.05 t ha^{-1}). A növény-szám, illetve a növényápolás minimális szintjénél a terméscsökkenés 2.7, illetve 3.36 t ha^{-1} (átlagos reakció 6.99 és 6.33 t ha^{-1}).



10.2. ábra. A növénytermesztési tényezők hatása a kukorica termésére

Kezelések: 1. Minden tényező minimumban, 2. Minden tényező optimumban, 3. Talajművelés minimumban, 4. Trágyázás minimumban, 5. Növényszám minimumban, 6. Fajta minimumban, 7. Növényápolás minimumban

10.2.3. Varianciaanalízis GenStat programmal

Az egytényezős véletlen blokkelrendezésű kísérlet (RCB) számítógépes értékelése hasonló a teljesen randomizált elrendezésű (CRD) kísérletéhez. Az RCB esetében azonban az ismétlést blokk csoportosítási faktorként külön oszlopba visszük be a GenStat számlálótáblába. Az ANOVA adat input dobozban az oszlop címkét, amely a blokk csoportosítási faktort tartalmazza, bevisszük a Blocks dobozba. Az RCB elrendezésben a blokkok hatása tehát el van különítve a kezelések hatásától. A Run billentyűre kattintással az ANOVA futtatható. Megjegyezzük, hogy a véletlen blokkelrendezésű (RCB) kísérlet lefuttatható az ANOVA General menüjével. A Design legördülő választási lehetőségeiből a One-way ANOVA (in Randomized Blocks)-t választjuk.

A kísérleti adatok bevitele a számítógépbe:

Ismétlést és kezelést faktoroknak, a termést változóknak tekintjük. Az ismétlés szintjeinek száma 4, a kezeléseké 7.

Ismétlés	Kezelés	Termés t ha ⁻¹
1	1	2.43
2	1	2.25
3	1	1.94
4	1	2.4
1	2	9.49
2	2	9.78
3	2	9.25
4	2	10.24
1	3	9.14
2	3	8.11
3	3	9.36
4	3	9.17
1	4	4.95
2	4	7.28
3	4	4.35
4	4	5.12
1	5	7.06
2	5	6.92
3	5	7.15
4	5	6.82
1	6	3.69
2	6	5.42
3	6	5.49
4	6	5.59
1	7	5.44
2	7	6.56
3	7	7.22
4	7	6.1

A kísérlet varianciaanalízisének számítógépes outputja:

Analysis of variance

Variate: Yield

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Replication stratum	3	1.3463	0.4488	0.85	
Replication.*Units* stratum					
Treatment	6	150.4401	25.0734	47.77	<.001
Residual	18	9.4482	0.5249		
Total	27	161.2346			

Tables of means

Variate: Yield

Grand mean 6.38

Treatment	1	2	3	4	5	6	7
	2.25	9.69	8.95	5.42	6.99	5.05	6.33

Standard errors of differences of means

Table	Treatment
rep.	4
d.f.	18
s.e.d.	0.512

Least significant differences of means (5% level)

Table	Treatment
rep.	4
d.f.	18
l.s.d.	1.076

Duncan's multiple range test**Treatment**

	Mean	
1	2.255	a
6	5.047	b
4	5.425	bc
7	6.330	cd
5	6.987	d
3	8.945	e
2	9.690	e

10.3. Latin négyzet (LS) elrendezés

A véletlen blokk elrendezés megkísérelti szabályozni a helyi (lokális) variáció egy formáját. Gyakran lehetséges azonban több mint egy típusú helyi variáció, melyet szabályozni kívánunk és szeretnénk használni egy olyan kísérleti elrendezést, amely magában foglalja a helyi szabályozás két formáját. Példák az ilyen helyzetekre:

(1) szántóföldi kísérletek, ahol vagy ismert, hogy két termékenységi trend keresztezi egymást, vagy ahol semmit sem tudunk a termékenységi trendekről és egy trend figyelembevétele bármely irányban szükséges;

(2) állatkísérletek, ahol különbségek az almok (egyszerre született utódok) között és a kezdeti tömegek variabilitása egyaránt nyilvánvaló variancia-források;

(3) vagy üvegházi kísérletek, ahol variáció lehet az üvegházban a fényviszonyoknak és hőmérsékleti különbségeknek tulajdoníthatóan. A latin négyzet elrendezést arra is használják, hogy úrrá legyenek azon a problémán, hogy nem képesek elvégezni minden mérést ugyanabban az időben. Például, ha arra volt lehetőség, hogy csak egy ismétlésben mérjenek meg minden kezelést naponta, akkor a mérés napja lehet egy faktor a latin négyzet elrendezésben.

Egytényezős kísérletekben 4, 5, 6, 7 és 8 kezelés összehasonlítására alkalmas kísérleti elrendezés, feltéve, hogy az ismétlések száma azonos a kezelések számával. A latin négyzet elrendezésben ugyanis a kezelések és ismétlések számának azonosnak kell lennie. Igen nagy előnye, hogy ha ugyanabban a sorban vagy oszlopban több parcella tönkremegy (pl. víznyomás, rovarkártétel), akkor akár egy sor, akár egy oszlop (teljes ismétlés) kihagyható, és a kísérlet véletlen blokkelrendezésként értékelhető.

A latin négyzet elrendezés abban különbözik a véletlen blokkelrendezéstől, hogy az összes kezelés egy-egy parcellájából két irányban képezünk blokkokat. A latin négyzet vázrajza szerint ezért sorokat és oszlopokat különböztetünk meg. Minden sor és minden oszlop magában foglalja az összes kezelés egy-egy parcelláját. Ezen a megszorításon kívül az elrendezés véletlen.

A szántóföldi, üvegházi és egyéb kísérletekben, ahol egymásra merőleges sávban, úgynevezett sor és oszlop irányú szisztematikus hibahatások várhatók, latin négyzet elrendezéssel az ilyen típusú hibák hatása a kísérlet hibaszórásából kiszűrhető és a kísérlet pontossága ezzel fokozható. Szántóföldi kísérletekben a latin négyzetet felhasználhatjuk arra, hogy eltávolítsuk két, egymásra merőleges gradiens, mondjuk, a lejtés és a talajtermékenység hatását. Itt a legjobb szántóföldi elrendezés a latin négyzet, négyzet alakú parcellákkal.

10.3.1. Randomizáció

A latin négyzet elrendezés randomizálásához kezdjük a kiinduló négyzettel:

1. Első sor alfabetikus sorrendben.

2. A következő sorokban ugyanabban a sorrendben, de egy parcellával balra eltolva írjuk be a betűket.

A B C D

BCDA
 CDAB
 DABC

Ezután randomizáljuk a sorok sorrendjét, mondjuk 2, 1, 3, 4:

BCDA
 ABCD
 CDAB
 DABC

Végül, randomizáljuk az oszlopok sorrendjét, mondjuk, 4, 2, 3, 1:

ACDB
 DBCA
 BDAC
 CABD



10.4. kép. Növénytermesztési klasszikus tartamkísérletek Martonvásáron az intézet kísérleti területen

Az utolsó randomizáció adja a kezelések elhelyezését a parcellákra a szántóföldi kísérleti tervben. Figyeljük meg, hogy a latin négyzet tulajdonságait megőriztük.

10.3.2. Kísérlet leírása

1959-ben az MTA Mezőgazdasági Kutatóintézetének kísérleti területén beállított kukorica monokultúra tartamkísérletben vizsgáltuk különböző trágyázási kezeléseknek a hatását a kukorica termésére (Berzsenyi és Győrffy, 1997; Berzsenyi és Dang, 2004; Berzsenyi et al., 2011) (10.4. kép). A kísérletben eredetileg feltett kérdés az volt, hogyha

a négyévenként kijuttatott 35 t ha⁻¹ és 70 t ha⁻¹ istállótrágya NPK hatóanyagát felerészben vagy egészen szervesetlen NPK műtrágyával pótoljuk, helyettesíthető-e az istállótrágya műtrágyával. A kísérletet latin négyzet elrendezésben állították be, hét kezelést és hét ismétlést foglal magában. A kísérleti parcella mérete 8 m × 10 m = 80 m². A kísérlet kezelése az alábbiak: 1. Kontroll, trágyázás nélkül, 2. 35 t ha⁻¹ istállótrágya négyévenként, 3. 17.5 t ha⁻¹ istállótrágya négyévenként + NPK műtrágya-kiegészítés, 4. 35 t ha⁻¹ istállótrágya hatóanyagának megfelelő mennyiségben NPK műtrágya, 5. 70 t ha⁻¹ istállótrágya négyévenként, 6. 35 t ha⁻¹ istállótrágya négyévenként + NPK műtrágya-kiegészítés, 7. 70 t ha⁻¹ istállótrágya hatóanyagának megfelelő mennyiségben NPK műtrágya. Az 1997. évi termésadatok varianciaanalízisét mutatjuk be, a hagyományos kézi számolással és számítógépes értékeléssel. A kísérlet elrendezési tervét a 10.3. ábra mutatja. A kezeléseket sorszámát (1–7) A–G betűvel helyettesítettük.

VII.	D	A	G	E	C	B	F
VI.	E	D	F	C	G	A	B

Elválasztó út

V.	G	C	A	F	B	E	D
IV.	F	G	D	B	A	C	E
III.	B	F	E	A	D	G	C

Elválasztó út

II.	C	E	B	G	F	D	A
I.	A	B	C	D	E	F	G

10.3. ábra. A 7 × 7-es latin négyzet trágyázási tartamkísérlet elrendezési terve a martonvásári kísérleti területen (I. – VII. jelöli a sorokat)

10.3.3. Varianciaanalízis lépései

Négy forrása van a variációnak az LS elrendezésnél, kettővel több, mint a CRC elrendezésnél és eggyel több, mint az RCD elrendezésnél. Az ANOVA számításakor az összes eltérésnégyzet összeget (SS) felosztjuk a sorhatás, az oszlophatás, a kezelés-hatás és a maradék (reziduális) hatás között. Feltételezzük, hogy a kezelésnek nincs interakciója sem a sor faktoral, sem pedig az oszlop faktoral; valóban, ha interakció lenne jelen, akkor a latin négyzet elrendezés érvénytelené válna.

A fő gyakorlati probléma a latin négyzet elrendezés használatával a megfelelő maradék szabadságfok biztosítása, amely lehetővé teszi az érvényes analízist. Ha a kezeléseket száma három, akkor a DF_{maradék} csupán kettő, míg négy kezelésnél a DF_{maradék} még mindig csupán hat. Csak öt kezelésnél éri el a DF_{maradék} a javasolt minimális 12-es értéket az érvényes ANOVA eljáráshoz, amelynél a parcellák összes száma 25 lesz (vagyis N²). A kísérlet típusától függően a 25 parcella kihívást jelenthet a rendelkezésre álló források tekintetében.

A lépésenkénti eljárás a varianciaanalízisben a következő:

1. lépés. Két alaptáblázatot készítünk. Az első alaptáblázatba az adatokat a kísérleti elrendezésnek (10.3. ábra) megfelelően írjuk be (10.10. táblázat). A táblázat adatait mindkét irányban összegezzük. Az R oszlopban vannak az elrendezés szerinti sorösszegek, a C sorban az elrendezés szerinti oszlopösszegek. Az R oszlop és a C sor adatainak összege egymással egyezik, a kísérlet főösszegét (G) adják.

10.10. táblázat. A 7 × 7-es latin négyzet kísérlet adatai a kísérleti terv szerint rendezve.

Az adatok kukorica szemtermését mutatják t ha⁻¹-ben (A kezeléseket A-G betűvel jelöltük.)

Sorok	Oszlopok							Sor- összeg R
	1	2	3	4	5	6	7	
1	3.38A	6.14B	7.32C	9.79D	6.74E	9.10F	9.08G	51.55
2	6.71C	5.62E	4.91B	9.37G	9.37F	10.2D	3.79A	49.97
3	4.77B	7.79F	5.61E	3.99A	8.79D	9.67G	7.46C	48.08
4	7.64F	7.93G	7.90D	5.20B	3.32A	7.38C	6.00E	45.37
5	7.74G	5.20C	2.84A	8.75F	5.43B	6.37E	8.34D	44.67
6	4.70E	6.86D	8.04F	6.88C	9.65G	3.99A	4.42B	44.54
7	7.89D	2.06A	7.57G	6.78E	6.98C	4.98B	8.08F	44.34
Oszlopösszeg (C) Főösszeg (G)	42.83	41.6	44.19	50.76	50.28	51.69	47.17	328.52

2. lépés. A második alaptáblázatban a 10.10. táblázat adatait az elrendezési terv (10.3. ábra) segítségével sorok és kezelések szerint átrendeztük. Az átrendezett adatokat a 10.11. táblázat tartalmazza. Utóbbi táblázatban az adatokat mindkét irányban összegezzük. Az R sor és a T oszlop adatainak összege egymással egyezik, a kísérlet főösszegét (G) adja.

10.11. táblázat. A 7 × 7-es latin négyzet kísérlet adatai a 10.3. ábra és az 10.10. táblázat adataiból kezelések és sorok szerint rendezve

Kezelések	Sorok							Kezelés- összeg (T)
	1	2	3	4	5	6	7	
A	3.38	3.79	3.99	3.32	2.84	3.99	2.06	23.37
B	6.14	4.91	4.77	5.2	5.43	4.42	4.98	35.85
C	7.32	6.71	7.46	7.38	5.2	6.88	6.98	47.93
D	9.79	10.2	8.79	7.9	8.34	6.86	7.89	59.77
E	6.74	5.62	5.61	6	6.37	4.7	6.78	41.82
F	9.1	9.37	7.79	7.64	8.75	8.04	8.08	58.77
G	9.08	9.37	9.67	7.93	7.74	9.65	7.57	61.01
Sorösszeg (R) Főösszeg (G)	51.55	49.97	48.08	45.37	44.67	44.54	44.34	328.52

3. lépés. Vázoljuk fel a varianciaanalízis táblázatot:

Variáció forrása	SS	d.f.	MS
Sor	$\frac{R_1^2 + R_2^2 + \dots + R_t^2}{t} - \frac{[\sum(x)]^2}{t^2}$	t - 1	
Oszlop	$\frac{C_1^2 + C_2^2 + \dots + C_t^2}{t} - \frac{[\sum(x)]^2}{t^2}$	t - 1	
Kezelés	$\frac{T_1^2 + T_2^2 + \dots + T_t^2}{t} - \frac{[\sum(x)]^2}{t^2}$	t - 1	
Hiba	kivonással	(t - 1) (t - 2)	s ²
Összes	$\sum(x^2) - \frac{[\sum(x)]^2}{t^2}$	t ² - 1	

4. lépés. Jelölje t a kezelések számát, határozzuk meg a szabadságfokot minden variáció forrásra:

$$\text{Összes d.f.} = t^2 - 1 = 49 - 1 = 48$$

$$\text{Sor d.f.} = \text{Oszlop d.f.} = \text{Kezelés d.f.} = t - 1 = 7 - 1 = 6$$

$$\text{Hiba d.f.} = (t - 1) (t - 2) = (7 - 1) (7 - 2) = 12$$

Hiba d.f. megkapható kivonással is:

$$\text{Hiba d.f.} = \text{Összes d.f.} - \text{Sor d.f.} - \text{Oszlop d.f.} - \text{Kezelés d.f.} = 48 - 6 - 6 - 6 = 30$$

5. lépés. Számítsuk ki a korrekciós faktort és a különböző eltérésnégyzet összegeket (SS):

$$C.F. = \frac{G^2}{t^2} = \frac{(328.52)^2}{49} = 2202.56$$

$$\text{Összes SS} = \sum X^2 - C.F. = \left[(3,38)^2 + (6,14)^2 + \dots + (8,08)^2 \right] - 2202.56 = 200.3$$

$$\text{Sor SS} = \frac{\sum R^2}{t} - C.F. = \frac{(51.55)^2 + (49.97)^2 + \dots + (44.34)^2}{7} - 2202.56 = 14.48$$

$$\text{Oszlop SS} = \frac{\sum C^2}{t} - C.F. = \frac{(42.83)^2 + (41.6)^2 + \dots + (47.17)^2}{7} - 2202.56 = 7.41$$

$$\text{Kezelés SS} = \frac{\sum T^2}{t} - C.F. = \frac{(23.37)^2 + (35.85)^2 + \dots + (61.01)^2}{7} - 2202.56 = 172.61$$

$$\text{Hiba SS} = \text{Összes SS} - \text{Sor SS} - \text{Oszlop SS} - \text{Kezelés SS} = 200.4 - 14.48 -$$

$$7.41 - 172.61 = 5.8$$

6. lépés. Számítsuk ki a közepes négyzetes eltérést (MS) mindegyik variáció forrásra úgy, hogy elosztjuk az SQ értékeket a megfelelő szabadságfokok számával.

$$\text{Sor MS} = \frac{\text{Sor SS}}{t-1} = \frac{14.48}{6} = 2.413$$

$$\text{Oszlop MS} = \frac{\text{Oszlop SS}}{t-1} = \frac{7.41}{6} = 1.235$$

$$\text{Kezelés MS} = \frac{\text{Kezelés SS}}{t-1} = \frac{172.61}{6} = 28.77$$

$$\text{Hiba MS} = \frac{\text{Hiba SS}}{(t-1)(t-2)} = \frac{5.8}{(6)(5)} = 0.194$$

7. lépés. Számítsuk ki az F értéket a kezelés-hatás tesztelésére:

$$F = \frac{\text{Kezelés MS}}{\text{Hiba MS}} = \frac{28.77}{0.194} = 148.3$$

8. lépés. Hasonlítsuk össze a számított F-értéket a táblázati F-értéssel, $f_1 =$ kezelés d.f. = $t-1$ és $f_2 =$ hiba d.f. = $(t-1)(t-2)$ értékeknél. Példánkban a táblázati F-érték $f_1 = 6$ és $f_2 = 30$ szabadságfoknál 2.42 az 5%-os, 3.47 az 1%-os és 5.12 a 0.1%-os szignifikancia szinten. Mivel a számított F érték nagyobb, mint a táblázati F érték 0.1%-os szignifikancia szinten, a kezelés különbség 0.1%-os szinten szignifikáns.

9. lépés. Számítsuk ki a variációs koefficiens:

$$CV = \frac{\sqrt{\text{Hiba MS}}}{\text{Fő átlag}} \times 100 = \frac{\sqrt{0.194}}{6.704} \times 100 = 6.57\%$$

10. lépés. Írjuk be a kiszámított SS, MS értékeket, az F értékeket és a szabadságfokokat a 10.12. táblázatba.

10.12. táblázat. A 7 x 7-es latin négyzet kísérlet kukorica termésadatainak varianciatáblázata

Variáció forrása	Szabadság fok (d.f.)	Eltérés négyzet összeg (SS)	Közepes négyzetes eltérés (MS)	Számított F-érték	Táblázati F-érték	
					1%	0.1%
Sor	6	14.48	2.413			
Oszlop	6	7.41	1.235			
Kezelés	6	172.61	28.77	148.3***	3.47	5.12
Hiba	30	5.8	0.194			
Összes	48					

CV = 6.75% *** = szignifikáns 0.1%-os szinten

Jegyezzük meg, hogy jóllehet az F-teszt a varianciaanalízisben mutatja a szignifikáns különbséget a hét vizsgált trágyázási kezelés átlagtermése között, azonban nem identifikálja a kezeléscsoportokat vagy speciális párokat, amelyek különböznek. E kérdés megválaszolására számítjuk ki a szignifikáns differenciát.

11. lépés. A szignifikáns differencia kiszámítása:

$$SzD_{5\%} = t_{5\%} \sqrt{\frac{2Hiba MS}{r}} = 2.04 \sqrt{\frac{(2)(0.194)}{7}} = 0.480 \text{ t ha}^{-1}$$

Eredmények értelmezése (közlése)

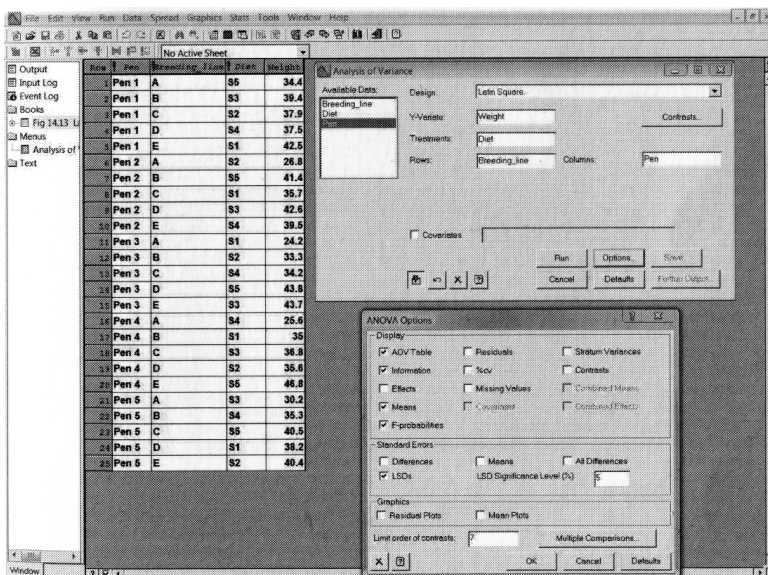
Összefoglalva megállapítható, hogy a kukorica termése legnagyobb volt a magas szintű trágyázásnál, akár NPK műtrágya formájában, akár a hatóanyag-azonosság alapján fele arányban istállótrágya formájában juttattuk ki. Az istállótrágya és az NPK műtrágya megfelelő kombinációja lehetővé teszi a magas kukoricatermés elérését. Közöljük a kezelésenkénti terméseredményeket, az $SzD_{5\%}$ -kal együtt:

Trágyázási kezelések	Átlagtermés t ha ⁻¹
1(A)	3.339e
2(B)	5.121d
3(C)	6.847b
4(D)	8.539a
5(E)	5.974c
6(F)	8.396a
7(G)	8.716a
$SzD_{5\%}$	0.480

Megjegyzés: Az „átlagtermés” oszlopban azonos betűjeleket tartalmazó kezelések nem különböznek szignifikánsan egymástól a Duncan többszörös terjedelem teszt alapján.

10.3.4. Varianciaanalízis GenStat programmal

Az Analysis of Variance menüből a General beállítást választjuk ki, hogy megjelenítsük a varianciaanalízis bemeneti dobozt és ezután kiválasztjuk a Latin Square-t. Az adatokat tartalmazó oszlop címkéket, vagyis a kezelés faktort, a sor faktort és az oszlop faktort bevisszük sorrendben az Y-variate dobozba, a Treatment dobozba, azután a Rows dobozba és a Columns dobozba. Más kimeneti beállítások, beleértve a többszörös összehasonlítási tesztek, megkaphatók az Options dobozon keresztül. A GenStat számolótáblát, az adatbeviteli dobozokat és a kimeneti ablakokat az analízisnek erre a típusára a 10.4. ábra mutatja, felhasználva a kísérleti adatokat.



10.4. ábra. GenStat számolótábla és menü dobozok a latin négyzet varianciaanalízishez

A kísérlet alapadatait az alábbi formában visszük be a GenStat számolótáblába. Az Oszlop, Sor és Kezelés képezik a faktorokat, míg a termés a vizsgált változó.

A trágyázási tartamkísérlet 1997. évi parcellánkénti terméseredményei ($t\ ha^{-1}$) és az adatok számítógépes alapadat táblázata (a felkiáltójel a faktorokat jelzi):

Oszlop!	Sor!	Kezelés!	Termés	Oszlop!	Sor!	Kezelés!	Termés
1	1	1	3.38	5	1	7	7.74
1	2	2	6.14	5	2	3	5.2
1	3	3	7.32	5	3	1	2.84
1	4	4	9.79	5	4	6	8.75
1	5	5	6.74	5	5	2	5.43
1	6	6	9.1	5	6	5	6.37
1	7	7	9.08	5	7	4	8.34
2	1	3	6.71	6	1	5	4.7
2	2	5	5.62	6	2	4	6.86
2	3	2	4.91	6	3	6	8.04
2	4	7	9.37	6	4	3	6.88
2	5	6	9.37	6	5	7	9.65
2	6	4	10.2	6	6	1	3.99
2	7	1	3.79	6	7	2	4.42
3	1	2	4.77	7	1	4	7.89
3	2	6	7.79	7	2	1	2.06
3	3	5	5.61	7	3	7	7.57

A táblázat folytatása

Oszlop!	Sor!	Kezelés!	Termés	Oszlop!	Sor!	Kezelés!	Termés
3	4	1	3.99	7	4	5	6.78
3	5	4	8.79	7	5	3	6.98
3	6	7	9.67	7	6	2	4.98
3	7	3	7.46	7	7	6	8.08
4	1	6	7.64				
4	2	7	7.93				
4	3	4	7.9				
4	4	2	5.2				
4	5	1	3.32				
4	6	3	7.38				
4	7	5	6				

Analysis of variance

Variate: Termés

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Sor stratum	6	14.4762	2.4127	12.45	
Oszlop stratum	6	7.4100	1.2350	6.37	
Sor.Oszlop stratum					
Kezelés	6	172.6071	28.7678	148.40	<.001
Residual	30	5.8158	0.1939		
Total	48	200.3090			

Tables of means

Variate: Termés

Grand mean 6.704

Kezelés	1	2	3	4	5	6	7
	3.339	5.121	6.847	8.539	5.974	8.396	8.716

Standard errors of differences of means

Table	Kezelés
rep.	7
d.f.	30
s.e.d.	0.2353

Least significant differences of means (5% level)

Table	Kezelés
rep.	7
d.f.	30
l.s.d.	0.4806

Duncan's multiple range test**Kezelés**

Mean

1	3.339	a
2	5.121	b
5	5.974	c
3	6.847	d
6	8.396	e
4	8.539	e
7	8.716	e

A latin négyzet elrendezésben beállított kísérlet varianciaanalízisét lefuttattuk véletlen blokk elrendezés szerint is. Figyeljük meg, hogy a Maradék MS, az s_d és az SzD értéke közel kétszerese, összehasonlítva a korrekelt latin négyzet elrendezés szerinti varianciaanalízissel.

A kétféle analízis eredményének összehasonlítása világosan mutatja annak fontosságát, hogy a varianciaanalízis típusának meg kell egyeznie a kísérlet elrendezésével. Tehát súlyos hibát követnénk el, ha a latin négyzet elrendezésben beállított kísérletet a véletlen blokk elrendezésnek megfelelő varianciaanalízissel értékelnénk.

A 7 × 7-es latin négyzet elrendezésben beállított trágyázási tartamkísérlet varianciaanalízise randomizált blokkelrendezés szerint
Analysis of variance

Variate: Termes

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
sor stratum	6	7.4100	1.2350	2.19	
sor.*Units* stratum					
Kezeles	6	172.6071	28.7678	51.04	<.001
Residual	36	20.2919	0.5637		
Total	48	200.3090			

Tables of means

Variate: Termes

Grand mean 6.70

Kezelés	1	2	3	4	5	6	7
	3.34	5.12	6.85	8.54	5.97	8.40	8.72

Standard errors of differences of means

Table	Kezeles
rep.	7
d.f.	36
s.e.d.	0.401

Least significant differences of means (5% level)

Table	Kezeles
rep.	7
d.f.	36
l.s.d.	0.814

10.4. Csoportosított blokk-elrendezés

A csoportosított blokk-elrendezés (Group balanced block design) elsődleges sajátossága a kezelések csoportosítása homogén blokkokba, a kezelések kiválasztott jellemzői alapján. Míg a latintégla (lattice design) a homogenitást blokkon belül a kísérleti parcellák olyan csoportosításával éri el, amely a kísérleti terület valamilyen ismert heterogenitási mintázatának ismeretén alapul, addig a csoportosított elrendezés ugyanezt a kezelések olyan csoportosításával éri el, amely a kezelések valamilyen ismert jellemzőjén alapul.

A csoportosított elrendezésben, ugyanahhoz a csoporthoz tartozó kezeléseket mindig ugyanabban a blokkban teszteljük, azonban azokat, amelyek eltérő csoportokhoz tartoznak, sohasem teszteljük ugyanabban a blokkban. Ezáltal, a precizitás, amellyel a különböző kezeléseket összehasonlítjuk, nem ugyanaz minden összehasonlításnál. Ugyanis az azonos csoporthoz tartozó kezelések egy magasabb fokú precizitással vannak összehasonlítva, mint azok, amelyek különböző csoportokhoz tartoznak.

A csoportosított elrendezést általánosan használjuk a fajta kísérletekben, ahol a hasonló morfológiai tulajdonságokkal rendelkező fajtákat együtt helyezük el ugyanabban a csoportban. Két leggyakrabban használt kritérium a fajták csoportosítására:

- Növénymagasság, abból a célból, hogy elkerüljük a jelentősebb kompetíciós hatást, amikor jelentősen eltérő magasságú növények növekednek a szomszédos parcellákban.
- Növekedés időtartama, abból a célból, hogy minimalizáljuk a kompetíciós hatást és elősegítsük a betakarítási műveletet.
- A kísérletek egy másik típusa, ahol a csoportosított elrendezést használjuk, az inszekticid, vagy herbicid kísérletek, amelyekben a kezeléseket fel lehet osztani hasonló permetezési eljárásokra, hogy elősegítsük a vegyszerek szabadföldi kijuttatását.



10.5. kép. Kukorica teljesítménykísérlet a hibridek termőképességének összehasonlítására

10.4.1. Randomizáció és elrendezés

A randomizáció és az elrendezés lépései a következők:

1. lépés. Az előírt csoportosítási kritérium alapján, csoportosítsuk a kezeléseket s csoportba, mindegyik tartalmaz t/s kezelést, ahol t a kezelések összes száma. Példánkban, a fajtákat három csoportba soroltuk (A, B és C), mindegyik csoport 15 fajtát tartalmaz, a eltérő növekedési időtartamuknak megfelelően.

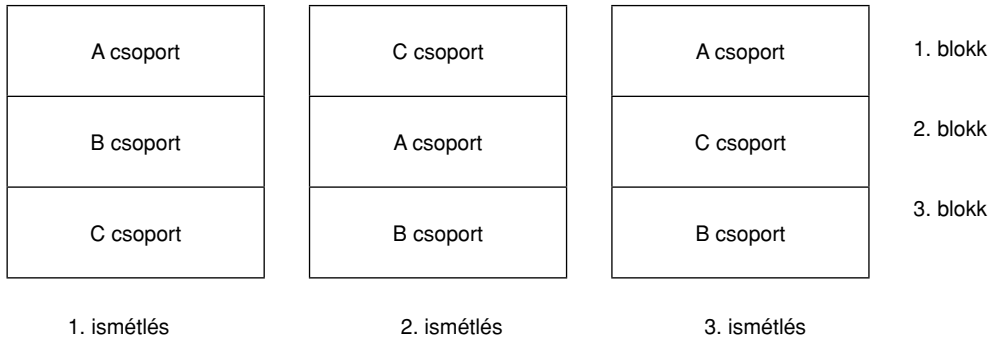
2. lépés. Osszuk fel a kísérleti területet r ismétlésre, mindegyik tartalmazzon t kísérleti parcellát. Példánkban a kísérleti területet három ismétlésre osztjuk fel, mindegyik magában foglal $(3)(15) = 45$ kísérleti parcellát.

3. lépés. Mindegyik ismétlést osszuk fel s blokkra, mindegyik tartalmazza a t/s kísérleti parcellát. Példánkban a három ismétlés mindegyikét felosztjuk három blokkra, mindegyik tartalmaz $45/3=15$ kísérleti parcellát.

4. lépés. Felhasználva a randomizációs sémákat, az s csoportot véletlenszerűen helyezük el az első ismétlés s blokkjában. Ezután függetlenül ismétljük meg az eljárást a megmaradó ismétlésekre.

Példánkban, az A, B és C fajtacsoportokat random helyeztük el az I ismétlés három blokkjára, majd a II. ismétlésbe és végül a III. ismétlésbe. Az eredményt az 10.5. ábra mutatja.

5. lépés. A három blokk mindegyikébe ismétlésenként, random helyezük el a t/s kezelést a 4. lépésben meghatározott blokkokba. Példánkban, az I. ismétlés első blokkjával kezdve, random jelöljük ki az A csoport 15 fajtáját a blokk 15 parcellájára. Az eljárást ismétljük meg a fennmaradt 8 blokkban, egymástól függetlenül. A végső eredményt a 10.6. ábra mutatja.



10.5. ábra. A fajták három csoportjának (A, B és C) random elhelyezése három blokkban a három ismétlés mindegyikébe. Ez képviseli a randomizációs folyamat első lépését a csoportosított blokk elrendezésű kísérletben

7	8	1	4	14	38	39	43	44	33	2	1	13	5	15	
12	3	13	10	9	31	42	36	32	45	11	4	3	14	7	1. blokk
6	11	2	5	15	40	34	35	41	37	9	12	10	6	8	
21	20	27	18	22	1	11	9	4	6	31	33	35	38	44	2. blokk
16	25	17	24	30	12	10	7	15	13	36	34	37	42	41	
26	19	28	29	23	3	2	8	5	14	39	43	32	45	40	3. blokk
41	31	32	44	43	17	27	22	25	28	16	25	23	20	17	
36	37	40	35	38	26	18	21	19	20	22	19	30	28	21	
33	39	42	45	34	16	23	24	30	29	18	24	26	29	27	
1. ismétlés	2. ismétlés	3. ismétlés													

10.6. ábra. A csoportosított blokk kísérlet elrendezése, amely magában foglal 45 fajtát, felosztva három csoportra, mindegyik csoport tartalmaz 15 fajtát. A kísérletet három ismétlésben állították be

10.4.2. A kísérlet leírása

Az MTA Mezőgazdasági Kutatóintézetében, Martonvásáron, az ún. I. fajtakísérletben összehasonlítottuk a legjobb hazai és nemzetközi kukorica hibridek termését és más fontos agronómiai tulajdonságait (10.5. kép). Az évente vizsgált mintegy 100 hibrid különböző érésidő csoportokhoz tartozott, amelyet a különböző FAO értékek jellemeztek. A hibridek öt FAO érésidő csoporthoz tartoztak, az igen koraitól a késői érésidőig: A: FAO 100-199, B: 200-299, C: 300-399, D:400-499, E: 500-599. A hibridek érésidőjének eltéréseit jól jellemezte a vetéstől az 50%-os nővirágzásig eltelt napok száma: A: 86, B: 89, C: 92, D: 95, E: 97. A kísérletet négy ismétlésben állítottuk be csoportosított blokk elrendezésben. A II. és III. ismétlés kiegészítő öntözésben (50 mm) részesült. A varianciaanalízishez 50 kukorica hibrid 1984. évi adatait használjuk fel (csoportonként 10–10 hibrid) (10.13. táblázat).

10.4.3. Varianciaanalízis lépései

A csoportosított blokkelrendezés varianciaanalízisének lépéseit mutatjuk be az alábbiakban, felhasználva az 10.13. táblázat adatait.

1. lépés. A csoportosított blokk elrendezés varianciaanalízisének felvázolása t kezeléssel, s csoporttal és r ismétléssel:

Variáció forrása	Szabadság- fok	SS	MS
Ismétlés	$r-1$		
Kezeléscsoport	$s-1$		
Hiba (a)	$(r-1)(s-1)$		
Kezelések az 1. csoporton belül	$\frac{t}{s} - 1$		
Kezelések a 2. csoporton belül	$\frac{t}{s} - 1$		
	\vdots		
Kezelések s csoporton belül	$\frac{t}{s} - 1$		
Hiba (b)	$s(r-1)\left(\frac{t}{s} - 1\right)$		
Összes	$(r)(t) - 1$		

2. lépés. Számítsuk ki a kezelés összegeket (T), az ismétlés összegeket (R) és a főösszeget (G), ahogy az 10.13. táblázat mutatja.

10.13. táblázat. 50 kukorica hibrid szemtermés adatai, csoportosított blokkelrendezésű kísérletben, 10 hibrid csoportonként

Hibrid sorszáma	Szemtermés t ha ⁻¹				Hibrid összesen (T)
	I. ismétlés	II. ismétlés	III. ismétlés	IV. ismétlés	
1	6.64	8.74	7.07	8.74	31.19
2	8.23	7.65	8.77	9.96	34.61
3	4.85	6.97	6.98	7.44	26.24
4	5.55	8.33	7.87	7.93	29.68
5	4.24	7.2	7.05	6.96	25.45
6	6.83	7.56	10.28	8.59	33.26
7	8.96	8.99	10.09	10.33	38.37
8	9.96	9.01	11.08	9.97	40.02
9	7.8	6.87	7.78	8.07	30.52
10	8.33	9.03	9.82	6.88	34.06
11	8.33	11.13	10.83	8.99	39.28
12	8.88	11.42	9.58	9.89	39.77
13	9.18	10.21	9.85	10.67	39.91
14	9.75	10.61	9.87	8.8	39.03

A táblázat folytatása

Hibrid sorszáma	Szemtermés t ha ⁻¹				Hibrid összesen (T)
	I. ismétlés	II. ismétlés	III. ismétlés	IV. ismétlés	
15	6.65	9.65	9.68	9.09	35.07
16	6.08	9.4	8.44	10.02	33.94
17	8.27	11.82	10.5	8.97	39.56
18	8.91	10.15	11.17	8.62	38.85
19	11.32	12.13	12.05	11.17	46.67
20	9.41	10.34	11.45	8.03	39.23
21	10.16	11.31	12.42	10.9	44.79
22	11.4	11.46	12.24	10.94	46.04
23	9.66	10.03	10.71	10.49	40.89
24	8.82	9.23	10.87	8.33	37.25
25	10.86	11.82	11.55	9.22	43.45
26	7.99	11.39	9.03	8.77	37.18
27	10.54	12.74	10.96	11.5	45.74
28	7.47	10.57	10.35	9.62	38.01
29	10.44	11.79	10.99	11.06	44.28
30	8.88	10.9	9.51	10.1	39.39
31	10.74	13.49	13.03	10.92	48.18
32	10.92	13.15	12.73	10.15	46.95
33	9.71	11.61	12.54	10.36	44.22
34	7.63	11.6	11.53	9.97	40.73
35	8.92	13.16	12.27	10.48	44.83
36	8.42	12.26	12.4	9.66	42.74
37	8.01	10.45	11.23	8.08	37.77
38	8.98	13.36	12.4	11.29	46.03
39	9.74	12.79	12.32	9.82	44.67
40	11.52	13.3	13.23	10.12	48.17
41	7.42	12.59	10.84	8.84	39.69
42	9.59	12.49	12.69	8.17	42.94
43	8.25	11.23	11.13	10.07	40.68
44	9.38	14.47	12.93	10.5	47.28
45	9.75	13.78	13.45	9.87	46.85
46	8.2	11.91	11.8	9.86	41.77
47	10.41	13.05	12.31	11.34	47.11
48	9.7	12.1	11.68	11.32	44.8
49	8.04	11.83	12.51	11.76	44.14
50	11.43	13.94	13.38	10.77	49.52
Ism. összeg (R)	441.2	551.0	545.2	483.4	
Főösszeg (G)					2020.8

A csoport tartalmazza az 1–10, a B csoport a 11–20, a C csoport a 21–30, a D csoport 31–40 és az E csoport a 41–50 sorszámú fajtákat.

10.14. táblázat. Az ismétlés × csoport táblázat az összes termésre, az 10.13. táblázat adataiból számolva

Csoport	Termés összeg (RS)				Csoport összeg (S)
	I. ismétlés	II. ismétlés	III. ismétlés	IV. ismétlés	
A	71.39	80.35	86.79	84.87	323.4
B	86.78	106.9	103.4	94.25	391.3
C	96.22	111.2	108.6	100.9	417.0
D	94.59	125.2	123.7	100.8	444.3
E	92.17	127.4	122.7	102.5	444.8

3. lépés. Készítsük el az ismétlés × csoport kétirányú táblázatot az összes termésre (RS), és számítsuk ki a csoport összegeket (S), ahogy a 10.14. táblázat mutatja. Ezután számítsuk ki a korrekciós faktort (C.F.), az összes SS-t, az ismétlés SS-t, a csoport SS-t és a hiba (a) SS-t:

$$C.F. = \frac{G^2}{rt} = \frac{(2020.8)^2}{(4)(50)} = 20418.2$$

$$\text{Összes SS} = \sum X^2 - C.F. = [(6.64)^2 + \dots + (10.77)^2] - 20418.2 = 704.8$$

$$\text{Ismétlés SS} = \frac{\sum R^2}{t} - C.F. = \frac{(441.2)^2 + \dots + (483.4)^2}{50} - 20418.2 = 165.34$$

$$\text{Csoport SS} = \frac{\sum S^2}{rt/s} - C.F. = \frac{(323.4)^2 + \dots + (444.8)^2}{(4)(50)/(5)} - 20418.2 = 252.84$$

$$\begin{aligned} \text{Hiba(a)SS} &= \frac{\sum (RS)^2}{t/s} - C.F. - \text{Ismétlés SS} - \text{Csoport SS} = \\ &= \frac{[(71.39)^2 + \dots + (102.5)^2]}{(50)/(5)} - 20418.2 - 165.34 - 252.84 = 45.82 \end{aligned}$$

4. lépés. Számítsuk ki az SS értékeket a kezelések között az i -edik csoporton belül:

$$\text{Kezelések SS az } i\text{-edik csoporton belül} = \frac{\sum_{j=1}^{t/s} T_{ij}^2}{r} - \frac{S_i^2}{rt/s}$$

ahol a T_{ij} a j -edik kezelés összeg az i -edik csoportban és S_i az i -edik csoport összeg.

Példánkban, a hibridek közötti SS-t az öt csoport mindegyikében a következőképpen számítjuk ki:

Hibridek SS az A csoporton belül:

$$ASS = \frac{\sum T_A^2}{4} - \frac{S_A^2}{(4)(50)/5} = \frac{(31.19)^2 + \dots + (34.06)^2}{4} - \frac{(323.4)^2}{40} = 50.06$$

Hibridek SS a B csoporton belül:

$$B\ SS = \frac{\Sigma T_B^2}{4} - \frac{S_B^2}{(4)(50)/5} = \frac{(39.28)^2 + \dots + (39.23)^2}{4} - \frac{(391.3)^2}{40} = 25.61$$

Hibridek SS a C, D és E csoporton belül:

$$C\ SS = \frac{\Sigma T_C^2}{4} - \frac{S_C^2}{(4)(50)/5} = \frac{(44.79)^2 + \dots + (39.39)^2}{4} - \frac{(417.0)^2}{40} = 29.03$$

$$D\ SS = \frac{\Sigma T_D^2}{4} - \frac{S_D^2}{(4)(50)/5} = \frac{(48.18)^2 + \dots + (48.17)^2}{4} - \frac{(444.3)^2}{40} = 24.19$$

$$E\ SS = \frac{\Sigma T_E^2}{4} - \frac{S_E^2}{(4)(50)/5} = \frac{(39.69)^2 + \dots + (49.52)^2}{4} - \frac{(444.8)^2}{40} = 22.82$$

Itt, T_A , T_B , T_C , T_D és T_E a kezelés összegekre vonatkozik, és S_A , S_B , S_C , S_D és S_E vonatkozik az A csoport, B csoport, C csoport, D csoport és E csoport összegekre.

5. lépés. Számítsuk ki a hiba (b) SS-t:

Hiba (b) SS = Összes SS – (az összes többi SS összege) = 704.8 – (165.34 + 252.84 + 45.82 + 50.06 + 25.61 + 29.03 + 24.19 + 22.82) = 89.09

6. lépés. Számítsuk ki az MS értékeket mindegyik variáció forrásra, úgy, hogy elosztjuk az SS-t a megfelelő d.f. értékkel:

$$\text{Ismétlés MS} = \frac{\text{Ismétlés SS}}{r-1} = \frac{165.34}{3} = 55.11$$

$$\text{Csoport MS} = \frac{\text{Csoport SS}}{s-1} = \frac{252.84}{4} = 63.21$$

$$\text{Hiba (a) MS} = \frac{\text{Hiba(a) SS}}{(r-1)(s-1)} = \frac{45.82}{(3)(4)} = 3.82$$

$$\text{Fajták A csoporton belül MS} = \frac{\text{Fajták A csoporton belül SS}}{(t/s)-1} = \frac{50.06}{9} = 5.56$$

$$\text{Fajták B csoporton belül MS} = \frac{\text{Fajták B csoporton belül SS}}{(t/s)-1} = \frac{25.61}{9} = 2.85$$

$$\text{Fajták C csoporton belül MS} = \frac{\text{Fajták C csoporton belül SS}}{(t/s)-1} = \frac{29.03}{9} = 3.23$$

$$\text{Fajták D csoporton belül MS} = \frac{\text{Fajták D csoporton belül SS}}{(t/s)-1} = \frac{24.19}{9} = 2.69$$

$$\text{Fajták E csoporton belül MS} = \frac{\text{Fajták E csoporton belül SS}}{(t/s)-1} = \frac{22.82}{9} = 2.54$$

$$\text{Hiba (b) MS} = \frac{\text{Hiba (b) SS}}{s(r-1)[(t/s)-1]} = \frac{89.09}{135} = 0.6599$$

7. lépés. Számítsuk ki a következő F-értékeket:

$$F(\text{csoport}) = \frac{\text{Csoport MS}}{\text{Hiba (a) MS}} = \frac{63.21}{3.82} = 16.55$$

$$F(\text{fajták A csoporton belül}) = \frac{\text{Fajták A csoporton belül MS}}{\text{Hiba(b) MS}} = \frac{5.56}{0.6599} = 8.43$$

$$F(\text{fajták B csoporton belül}) = \frac{\text{Fajták B csoporton belül MS}}{\text{Hiba(b) MS}} = \frac{2.85}{0.6599} = 4.32$$

$$F(\text{fajták C csoporton belül}) = \frac{\text{Fajták C csoporton belül MS}}{\text{Hiba(b) MS}} = \frac{3.23}{0.6599} = 4.89$$

$$F(\text{fajták D csoporton belül}) = \frac{\text{Fajták D csoporton belül MS}}{\text{Hiba(b) MS}} = \frac{2.69}{0.6599} = 4.08$$

$$F(\text{fajták E csoporton belül}) = \frac{\text{Fajták E csoporton belül MS}}{\text{Hiba(b) MS}} = \frac{2.54}{0.6599} = 3.85$$

8. lépés. Mindegyik számított F-értékre keressük ki a megfelelő táblázati F-értéket az adott szignifikancia szinten, f_1 = a számláló MS d.f. értéke és f_2 = a nevező MS d.f. értéke.

Példánkban, a táblázati F-érték, amely megfelel az F (csoport) értéknek, $f_1 = 4$ és $f_2 = 12$ szabadságfoknál, 3.26 az 5%-os és 5.41 az 1%-os szignifikancia szinten. A három számított F (fajták csoporton belül) értéknek megfelelő táblázati F-értékek $f_1 = 9$ és $f_2 = 135$ szabadságfoknál 1.95 az 5%-os és 2.56 az 1%-os szignifikancia szinten.

9. lépés. Számítsuk ki a két variációs koefficiens, amely megfelel a két hiba MS értéknek.

$$CV(a) = \frac{\sqrt{\text{Hiba (a) MS}}}{\text{foatlag}} \times 100 = \frac{\sqrt{3.82}}{10.104} \times 100 = 19.34\%$$

$$CV(b) = \frac{\sqrt{\text{Hiba (b) MS}}}{\text{foatlag}} \times 100 = \frac{\sqrt{0.6599}}{10.104} \times 100 = 8.04\%$$

10. lépés. Írjuk be a varianciaanalízis 3–9. lépésében kapott értékeket a 10.15. táblázatba. Az eredmények szignifikáns különbséget mutatnak a hibridek öt csoportjának átlagértékei között és szignifikáns különbséget a hibridek között az öt csoport mind-egyikében.

10.15. táblázat. A csoportosított blokk elrendezésű kísérlet variancia táblázata

Variáció forrása	Szabad- ságfok	SS	MS	Számított F-érték	Táblázati F-érték	
					5%	1%
Ismétlés	3	165.34	55.11			
Fajta csoport	4	252.84	63.21	16.55**	3.26	5.41
Hiba (a)	12	45.82	3.82			
Fajták az A csoporton belül	9	50.06	5.56	8.43**	1.95	2.56
Fajták a B csoporton belül	9	25.61	2.85	4.32**	1.95	2.56
Fajták a C csoporton belül	9	29.03	3.23	4.89**	1.95	2.56
Fajták a D csoporton belül	9	24.19	2.69	4.08**	1.95	2.56
Fajták az E csoporton belül	9	22.82	2.54	3.85**	1.95	2.56
Hiba (b)	135	89.09	0.6599			
Összes	199	704.8				

cv (a) = 19.34%, cv (b) = 8.04%, ** = szignifikáns 1%-os szinten

Szignifikáns differencia kiszámítása

Azonos csoporton belül két kezelés között:

$$SzD_{5\%} = t_{5\%} \sqrt{\frac{2Hiba(b) MS}{r}} = 1.98 \sqrt{\frac{(2)(0.6599)}{4}} = 1.137$$

Két csoportátlag között, ha a kezelések száma azonos:

$$SzD_{p\%} = t_{5\%} \sqrt{\frac{2Hiba(a) MS}{rs}} = 2.179 \sqrt{\frac{(2)(3.82)}{(4)(5)}} = 1.347$$

Különböző csoportokban levő két kezelés között, ha a kezelések száma azonos:

$$SzD_{5\%} = t_{5\%} \sqrt{\frac{2[(p-1)(Hiba(b) MS + Hiba(a) MS)]}{(r)(p)}}$$

$$= 1.98 \sqrt{\frac{(2)(10-1)(0.6599 + 3.82)}{(4)(10)}} = 1.24$$

ahol p kezelések száma csoportonként

Eredmények közlése

1. Közöljük a varianciatáblázatot.
2. Közöljük a csoportátlagokat az $SzD_{5\%}$ feltüntetésével.
3. Közöljük a kezelésenkénti termésátlagokat, de csoportonként elkülönítve, az $SzD_{5\%}$ -kal. A táblázat alatt feltüntetjük a különböző csoportokban levő kezelések közötti különbségre vonatkozó $SzD_{5\%}$ -ot.

10.4.4. Varianciaanalízis GenStat programmal

A varianciaanalízis lépései:

Analysis of Variance – General – Two-way ANOVA (in Randomized Blocks)

Y-variate: Yield

Treatment 1: Variety

Treatment 2: VarGroup

Blocks: Ismetles/VarGroup (kézzel kell beírni a"/" jeltől)

Először elvégezzük a varianciaanalízist az összes adatra, ezt követően külön-külön az adatsoportokra (érésidőnként). A végső variancia táblázatot a 10.15. táblázat alapján állítjuk össze.

Analysis of variance (az összes adatra)

Variate: Yield

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Ismetles stratum	3	165.5880	55.1960	14.95	
Ismetles.VarGroup stratum					
VarGroup	4	252.8270	63.2068	17.12	<.001
Residual	12	44.3079	3.6923	5.60	
Ismetles.VarGroup.*Units* stratum					
Variety	9	33.1135	3.6793	5.58	<.001
VarGroup.Variety	36	118.8293	3.3008	5.00	<.001
Residual	135	89.0807	0.6599		
Total	199	703.7464			

Tables of means

Variate: Yield

Grand mean 10.104

VarGroup	1	2	3	4	5		
	8.085	9.783	10.425	11.107	11.119		
Variety	1	2	3	4	5	6	7
	10.156	10.515	9.597	9.698	9.782	9.444	10.428
Variety	8	9	10				
	10.385	10.514	10.518				
VarGroup	Variety	1	2	3	4	5	6
1		7.798	8.653	6.560	7.420	6.363	8.315
2		9.820	9.943	9.978	9.758	8.768	8.485
3		11.197	11.510	10.223	9.312	10.862	9.295
4		12.045	11.738	11.055	10.183	11.207	10.685
5		9.922	10.735	10.170	11.820	11.712	10.442
VarGroup	Variety	7	8	9	10		
1		9.593	10.005	7.630	8.515		
2		9.890	9.712	11.668	9.807		
3		11.435	9.502	11.070	9.848		
4		9.443	11.508	11.168	12.043		
5		11.777	11.200	11.035	12.380		

Standard errors of differences of means

Table	VarGroup	Variety	VarGroup Variety
rep.	40	20	4
s.e.d.	0.4297	0.2569	0.6939
d.f.	12	135	66.38

Except when comparing means with the same level(s) of

VarGroup	0.5744
d.f.	135

Least significant differences of means (5% level)

Table	VarGroup	Variety	VarGroup Variety
rep.	40	20	4
l.s.d.	0.9362	0.5080	1.3853
d.f.	12	135	66.38

Except when comparing means with the same level(s) of

VarGroup	1.1360
d.f.	135

Analysis of variance (FAO 100-199 csoportra)

Variate: Yield

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Ismetles stratum	3	14.1186	4.7062	5.47	
Ismetles.VarGroup.*Units* stratum					
Variety	9	50.1734	5.5748	6.48	<.001
Residual	27	23.2144	0.8598		
Total	39	87.5064			

Tables of means

Variate: Yield

Grand mean 8.08

Variety	1	2	3	4	5	6	7
	7.80	8.65	6.56	7.42	6.36	8.31	9.59
Variety	8	9	10				
	10.00	7.63	8.51				

Standard errors of differences of means

Table	Variety
rep.	4
d.f.	27
s.e.d.	0.656

Least significant differences of means (5% level)

Table	Variety
rep.	4
d.f.	27
l.s.d.	1.345

Analysis of variance (FAO 200-299 csoportra)

Variate: Yield

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Ismetles stratum	3	24.7708	8.2569	10.57	
Ismetles.VarGroup.*Units* stratum					
Variety	9	25.3988	2.8221	3.61	0.005
Residual	27	21.0898	0.7811		
Total	39	71.2594			

Tables of means

Variate: Yield

Grand mean 9.78

Variety	1	2	3	4	5	6	7
	9.82	9.94	9.98	9.76	8.77	8.48	9.89
Variety	8	9	10				
	9.71	11.67	9.81				

Standard errors of differences of means

Table	Variety
rep.	4
d.f.	27
s.e.d.	0.625

Least significant differences of means (5% level)

Table	Variety
rep.	4
d.f.	27
l.s.d.	1.282

Analysis of variance (FAO 300-399 csoportra)

Variate: Yield

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Ismetles stratum	3	14.3548	4.7849	8.13	
Ismetles.VarGroup.*Units* stratum					
Variety	9	28.5663	3.1740	5.39	<.001
Residual	27	15.8979	0.5888		
Total	39	58.8190			

Tables of means

Variate: Yield

Grand mean 10.43

Variety	1	2	3	4	5	6	7
	11.20	11.51	10.22	9.31	10.86	9.29	11.43
Variety	8	9	10				
	9.50	11.07	9.85				

Standard errors of differences of means

TableVariety
 rep. 4
 d.f. 27
 s.e.d. 0.543

Least significant differences of means (5% level)

TableVariety
 rep. 4
 d.f. 27
 l.s.d. 1.113

Analysis of variance (FAO 400-499 csoportra)

Variate: Yield

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Ismetles stratum	3	73.3861	24.4620	66.96	
Ismetles.VarGroup.*Units* stratum					
Variety	9	24.5310	2.7257	7.46	<.001
Residual	27	9.8633	0.3653		
Total	39	107.7804			

Tables of means

Variate: Yield

Grand mean 11.107

Variety	1	2	3	4	5	6	7
	12.045	11.737	11.055	10.182	11.207	10.685	9.442
Variety	8	9	10				
	11.507	11.167	12.042				

Standard errors of differences of means

Table	Variety
rep.	4
d.f.	27
s.e.d.	0.4274

Least significant differences of means (5% level)

Table	Variety
rep.	4
d.f.	27
l.s.d.	0.8769

Analysis of variance (FAO 500-599 csoportra)

Variate: Yield

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Ismetles stratum	3	83.2657	27.7552	39.41	
Ismetles.VarGroup.*Units* stratum					
Variety	9	23.2733	2.5859	3.67	0.004
Residual	27	19.0152	0.7043		
Total	39	125.5542			

Tables of means

Variate: Yield

Grand mean 11.12

Variety	1	2	3	4	5	6	7
	9.92	10.73	10.17	11.82	11.71	10.44	11.78
Variety	8	9	10				
	11.20	11.03	12.38				

Standard errors of differences of means

Table	Variety
rep.	4
d.f.	27
s.e.d.	0.593

Least significant differences of means (5% level)

Table	Variety
rep.	4
d.f.	27
l.s.d.	1.218

10.5. Latintégla elrendezés

Elméletileg a komplett blokkelrendezések, mint amilyen a véletlen blokk (RCB) és a latin négyzet (LS) elrendezés, bármilyen számú kezeléshez alkalmazhatók. A teljes (komplett) blokkelrendezések azonban kevésbé hatékonyá válnak, ahogy a kezelések száma nő, elsődlegesen azért, mert a blokk mérete arányosan nő a kezelések számával, és a kísérleti kezelések homogenitását nagy blokkon belül bonyolult fenntartani. Azaz, a komplett blokkelrendezés kísérleti hibája általában várhatóan nő a kezelések számával.

A nagy számú kezelést tartalmazó egytényezős kísérletek alternatív elrendezési módja az inkomplett (nem teljes) blokkelrendezés, amelynek egyike a latintégla elrendezés. Mint a neve mutatja, minden blokk az inkomplett blokk elrendezésben nem tartalmaz minden kezelést és meglehetősen kis blokkméret fenntartható még akkor is, ha a kezelések száma nagy. Kisebb blokkokkal a kísérleti parcellák homogenitása ugyanabban a blokkban könnyebben fenntartható és általában nagyobb pontosság várható.

Megvan az ára a nagyobb pontosságnak az inkomplett blokkelrendezés használatakor. Ezek a következők:

- (a) A kezeléseknek vagy az ismétléseknek vagy mindkettőnek a száma nem rugalmas (nem választható meg önkényesen).
- (b) Nem egyenlő mértékű a pontosság a kezelésátlagok összehasonlításakor.
- (c) Az adatelemzés bonyolultabb.



10.6. kép. Latintégla kísérleti elrendezés a martonvásári kukoricanevelési tenyészkertben (Vécsy Attila felvétele)

Jóllehet nincs külön szabály arra, hogy mekkora lehet a kezelések száma az inkomplett blokkelrendezés használatakor, az alábbi irányelvek segítségével szolgálhatnak (Gomez és Gomez, 1984):

Variabilitás a kísérleti anyagban. Az inkomplett blokkelrendezés előnye a komplett blokkelrendezéssel szemben fokozódik a kísérleti anyagban levő variabilitás növekedésével. Általánosságban, bármikor, ha a blokkméret az RCB elrendezésnél túl nagy ahhoz, hogy fenntartson egy megfelelő szintű homogenitást a kísérleti parcellák között ugyanazon a blokkon belül, az inkomplett blokkelrendezés használatát komolyan fontolóra kell venni. Például, öntözött rizskísérletben, ahol a kísérleti parcellák várhatóan relatíve homogének, az RCB elrendezés valószínűleg megfelelő egy fajtakísérletben 25 fajttal. Másrészt, ugyanennél a kísérletnél száraz körülmények között, ahol a kísérleti parcellák várhatóan kevésbé homogének, a latintégla elrendezés hatékonyabb lehet.

Számítógépes háttér és szolgáltatás. Az inkomplett blokkelrendezés adatelemzése komplexebb, mint a komplett blokkelrendezésé. Ezért olyan esetekben, amikor a számítógépes lehetőségek és szolgáltatások kevésbé hozzáférhetők, az inkomplett blokkelrendezést utolsó megoldásként vegyük számításba.

Általánosságban, az inkomplett blokkelrendezés a kisebb blokkmérettel várhatóan nagyobb pontosságot ad, mint a komplett blokkelrendezés. Ezáltal az inkomplett blokkelrendezést általában előnyben kell részesíteni, ha a használatához szükséges forrásoknak (pl. több ismétlés, nem flexibilis számú kezelések, és komplexebb adatelemzés) eleget tudunk tenni.

A latintégla elrendezés a legáltalánosabban használt inkomplett blokkelrendezés a mezőgazdasági kutatásban (10.6. kép). Megfelelő flexibilitás van a tervezésében ahhoz, hogy alkalmazása egyszerűbb legyen, mint más inkomplett blokkelrendezésé. Leggyakrabban a kiegyensúlyozott (balanced) és a részben kiegyensúlyozott (partially balanced) latintégla elrendezést használják. Mindkettőnél a kezelések számának pontos négyzetszámnak kell lenni.

10.5.1. Kiegyensúlyozott latintégla elrendezés

A kiegyensúlyozott latintégla elrendezést a következők jellemzik:

1. A kezelések száma (t) tökéletes négyzet szám (azaz $t = k^2$, úgymint 25, 49, 64, 81 stb.). Jóllehet ez a követelmény szigorúnak tűnhet először, rendszerint könnyű betartani a gyakorlatban. Ahogy a kezelések száma nagy lesz, néhányval több vagy néhányval kevesebb kezelést rendszerint könnyű teljesíteni. Például, ha egy növénynemesítő 80 fajta teljesítményét kívánja vizsgálni egy kiegyensúlyozott latintégla elrendezésű kísérletben, mindössze eggyel több fajtát kell beállítania a kísérletbe. Vagy ha 82 vagy 83 fajtája van, könnyen kihagy egyet vagy kettőt.

2. A blokk mérete (k) egyenlő a kezelések számának négyzetgyökével (azaz, $k = t^{1/2}$).

3. Az ismétlések száma (R) eggyel több, mint a blokk nagysága, azaz $r = [(k + 1)]$. Azaz, a szükséges ismétlések száma 6, ha a kezelések száma 25, 7 ha a kezelések száma 36, 8 ha a kezelések száma 49, és így tovább.

10.5.1.1. Randomizálás és elrendezés

A latintégla elrendezésű kísérlet randomizálását és a kezelések elrendezését kilenc kezelést magában foglaló szántóföldi kísérlettel mutatjuk be. Négy ismétlés van, mindegyik tartalmaz három inkomplett blokkot és mindegyik inkomplett blokkban három kezelés van.

Az egyes lépések a következők:

1. lépés. A kísérleti területet osszuk fel $r = (k + 1)$ ismétlésre, mindegyik ismétlés tartalmaz $t = k^2$ kísérleti parcellát. Ebben a példában, a kísérleti területet $r = 4$ ismétlésre osztjuk fel, mindegyik tartalmaz $t = 9$ kísérleti parcellát, mint ahogy az 10.7. ábra mutatja.

1. blokk	1 2 3	10 11 12	19 20 21	28 29 30
2. blokk	4 5 6	13 14 15	22 23 24	31 32 33
3. blokk	7 8 9	16 17 18	25 26 27	34 35 36
	I. ismétlés	II. ismétlés	III. ismétlés	IV. ismétlés

10.7. ábra. A kísérleti terület felosztása, amely tartalmaz 36 parcellát (1, 2, ..., 36) négy ismétlésben, mindegyik magában foglal három inkomplett blokkot 3–3 parcellával.

Ez az első lépés a 3×3 -as kiegyensúlyozott latintégla elrendezésében

2. lépés. Mindegyik ismétlést osszuk fel k inkomplett blokkra, mindegyik tartalmaz k kísérleti parcellát. A példánkban, mindegyik ismétlést $k = 3$ inkomplett blokkra osztjuk fel, mindegyik tartalmaz $k = 3$ kísérleti parcellát (10. 7.ábra).

3. lépés. Válasszunk ki a kezelések számának megfelelő latintégla elrendezésű kísérleti tervet. Esetünkben a 3×3 kiegyensúlyozott rácselrendezést az 10.16. táblázat mutatja.

10.16. táblázat. A 3×3 -as kiegyensúlyozott latintégla elrendezés alapterve, amely magában foglal kilenc kezelést (1,2, ..., 9), blokkonként három parcellával és négy ismétlésben

Inkomplett blokkok számozása	Kezelések száma			
	I. ismétlés	II. ismétlés	III. ismétlés	IV. ismétlés
1	1 2 3	1 4 7	1 5 9	1 6 8
2	4 5 6	2 5 8	2 6 7	2 4 9
3	7 8 9	3 6 9	3 4 8	3 5 7

4. lépés. Randomizáljuk a kiválasztott alapterv ismétlés elrendezését, követve egy megfelelő randomizációs sémát. Pédánkban a véletlen számok táblázata módszert alkalmazzuk:

Kiválasztunk négy három számjegyű véletlen számot a véletlen számok táblázatból; például, 372, 217, 963 és 404. Tegyük rangsorba a legkisebttől a legnagyobbig:

Véletlen szám	Sorrend	Rangsor
372	1	2
217	2	1
963	3	4
404	4	3

Használjuk a sorrendet, hogy reprezentálja az alapterv meglévő ismétlés számát és a rangsort, hogy reprezentálja az új terv ismétlés számát. Tehát, az alapterv első ismétlése (sorrend = 1) az új terv második ismétlésévé válik (rangsor = 2), az alapterv második ismétlése az új terv első ismétlésévé válik, és így tovább. Az új terv kimenetele ebben a lépésben:

Nem komplett	Kezelés szám			
	I. ismétlés	II. ismétlés	III. ismétlés	IV. ismétlés
1	1 4 7	1 2 3	1 6 8	1 5 9
2	2 5 8	4 5 6	2 4 9	2 6 7
3	3 6 9	7 8 9	3 5 7	3 4 8

5. lépés. Randomizáljuk az inkomplett blokkokat minden ismétlésen belül, követve egy megfelelő randomizációs sémát. Példánkban, ugyanazt a randomizációs sémát használtuk, mint a 4. lépésben, hogy újra jelöljük a három inkomplett blokkot mind a négy ismétlésben. Négy független randomizációs folyamat után, az újra kijelölt inkomplett blokkok a következőt mutathatják:

Inkomplett blokk szám az alaptervben	Újra kijelölt inkomplett blokk szám az új tervben			
	I. ismétlés	II. ismétlés	III. ismétlés	IV. ismétlés
1	3	2	3	1
2	2	1	1	3
3	1	3	2	2

Mint látható, az I. ismétlésre, az alapterv 1. blokkja az új terv 3. blokkjává válik, a 2. blokk marad ugyanazon a helyen és az alapterv 3. blokkja az új terv első blokkjává válik. A II. ismétlésre, az alapterv 1. blokkja az új terv 2. blokkjává válik, az alapterv 2. blokkja az új terv 1. blokkjává válik, és így tovább. Az új terv kimenetele ennél a lépésnél:

Inkomplett blokk szám	Kezelés szám			
	I. ismétlés	II. ismétlés	III. ismétlés	IV. ismétlés
1	3 6 9	4 5 6	3 5 7	1 5 9
2	2 5 8	1 2 3	1 6 8	3 4 8
3	1 4 7	7 8 9	2 4 9	2 6 7

6. lépés. Randomizáljuk a kezelés elrendezést mindegyik inkomplett blokkban. Példánkban, random újra jelöljük ki a három kezelést a 12 inkomplett blokk mindegyikében, követve ugyanazt a randomizációs sémát, mint a 4. és 5. lépésben. 12 független randomizációs folyamat után, az újra kijelölt kezelés sorrendek a következőt mutathatják:

Kezelés sorrend az alaptervben	Újra kijelölt kezelés sorrend az új tervben					
	I. ismétlés			II. ismétlés		
	Blokk 1	Blokk 2	Blokk 3	Blokk 1	Blokk 2	Blokk 3
1	2	3	2	2	3	3
2	3	2	3	1	2	2
3	1	1	1	3	1	1

Kezelés sorrend az alaptervben	Újra kijelölt kezelés sorrend az új tervben					
	III. ismétlés			IV. ismétlés		
	Blokk 1	Blokk 2	Blokk 3	Blokk 1	Blokk 2	Blokk 3
1	3	3	1	1	3	2
2	2	1	2	3	1	3
3	1	2	3	2	2	1

Az I. ismétlés 1. inkomplett blokk ebben az esetben az alapterv 1. kezelés sorrendje (kezelés 3) az új terv 2. kezelés sorrendje lett, az alapterv 2. kezelés sorrendje (kezelés 6) az új terv 3. kezelés sorrendje lett, az alapterv 3. kezelés sorrendje (kezelés 9) az új terv 1. kezelés sorrendje lett, és így tovább. Az új terv kimenete ebben a lépésben:

Inkomplett blokk szám	Kezelés szám			
	I. ismétlés	II. ismétlés	III. ismétlés	IV. ismétlés
1	9 3 6	5 4 6	7 5 3	1 9 5
2	8 5 2	3 2 1	6 8 1	4 8 3
3	7 1 4	9 8 7	2 4 9	7 2 6

10.17. táblázat. A hajtások száma m²-enként egy 16 műtrágyakezelést magában foglaló 4 × 4-es kiegyensúlyozott latintégla elrendezésű kísérletben^a

Blokk szám	Hajtás db/m ²				Blokk összeg (B)	Blokk szám	Hajtás db/m ²				Blokk összeg (B)
	I. ismétlés						II. ismétlés				
	(1)	(2)	(3)	(4)			(1)	(5)	(9)	(13)	
1	147	152	167	150	616	5	140	165	182	152	639
	(5)	(6)	(7)	(8)			(10)	(2)	(14)	(6)	
2	127	155	162	172	616	6	97	155	192	142	586
	(9)	(10)	(11)	(12)			(7)	(15)	(3)	(11)	
3	147	100	192	177	616	7	155	182	192	192	721
	(13)	(14)	(15)	(16)			(16)	(8)	(12)	(4)	
4	155	195	192	205	747	8	182	207	232	162	783

A 10.17. táblázat folytatása

Blokk szám	Hajtás db/m ²				Blokk összeg (B)	Blokk szám	Hajtás db/m ²				Blokk összeg (B)
Ismétlés összeg R ₁					2595	Ismétlés összeg R ₂					2729
III. ismétlés						IV. ismétlés					
	(1)	(6)	(11)	(16)		(1)	(14)	(7)	(12)		
9	155	162	177	152	646	13	220	202	175	205	802
	(5)	(2)	(15)	(12)			(13)	(2)	(11)	(8)	
10	182	130	177	165	654	14	205	152	180	187	724
	(9)	(14)	(3)	(8)			(5)	(10)	(3)	(16)	
11	137	185	152	152	626	15	165	150	200	160	675
	(13)	(10)	(7)	(4)			(9)	(6)	(15)	(4)	
12	185	122	182	192	681	16	155	177	185	172	689
Ismétlés összeg R ₃					2607	Ismétlés összeg R ₄					2890
V. ismétlés											
	(1)	(10)	(15)	(8)							
17	147	112	177	147	583						
	(9)	(2)	(7)	(16)							
18	180	205	190	167	742						
	(13)	(6)	(3)	(12)							
19	172	212	197	192	773						
	(5)	(14)	(11)	(4)							
20	177	220	205	225	827						
Ismétlés összeg R ₅					2925						

^aA zárójelben levő értékek megfelelnek a kezelésszámoknak.

7. lépés. A 6. lépés randomizációs folyamatának végső eredményét alkalmazzuk az 10.7. ábra szántóföldi elrendezéséhez, amely a 10.8. ábrát, a végső elrendezést adja. Jegyezzük meg, hogy a kiegyensúlyozott latintégla elrendezés fontos sajátossága, hogy minden kezeléspár csak egyszer fordul elő ugyanabban a blokkban. Például, kezelés 1 csak egyszer jelenik meg kezelés 4-gyel és 7-tel az I. ismétlés 3. blokkban; a 2. és 3. kezeléssel a II. ismétlés 2. blokkban, a 6. és 8. kezeléssel a III. ismétlés 2. blokkban; az 5. és 9. kezeléssel a IV. ismétlés 1. blokkban. E sajátosság következményeként, a pontosság mértéke minden kezeléspár összehasonlításához a kiegyensúlyozott rácselrendezésben ugyanaz minden párra.

I. blokk	T ₉	T ₃	T ₆	T ₅	T ₄	T ₆	T ₇	T ₅	T ₃	T ₁	T ₉	T ₅
II. blokk	T ₈	T ₅	T ₂	T ₃	T ₂	T ₁	T ₆	T ₈	T ₁	T ₄	T ₈	T ₃
III. blokk	T ₇	T ₁	T ₄	T ₉	T ₈	T ₇	T ₂	T ₄	T ₉	T ₇	T ₂	T ₆
	I. ismétlés			II. ismétlés			III. ismétlés			IV. ismétlés		

10.8. ábra. A 3 x 3-as kiegyensúlyozott latintégla elrendezésű kísérlet terve, amely magában foglal kilenc kezelést (T₁, T₂, ..., T₉)

10.5.1.2. Kísérlet leírása

A számítási eljárást 16 műtrágya kezelést tartalmazó szántóföldi rizskísérlet hajtás-szám adataival mutatjuk be (Gomez és Gomez, 1984). A kísérlet 4×4 kiegyensúlyozott latintégla elrendezés öt ismétléssel. Az adatokat a 10.17. táblázat mutatja.

10.5.1.3. Varianciaanalízis lépései

A latintégla elrendezés varianciaanalízisét Cochran és Cox (1957) írta le részletesen. Gomez és Gomez (1984) által bemutatott példával mi is ezt követjük. Hasonló módszert mutat be Petersen (1994). A variációnak négy forrása van, amely kiszámítható a kiegyensúlyozott latintégla elrendezésű kísérletben: ismétlés, kezelés, inkomplett blokk és kísérleti hiba. Az RCB-hez viszonyítva az inkomplett blokk egy további variáció forrás és visszatükrözi a különbségeket ugyanazon ismétlés inkomplett blokkjai között. A lépések a következők:

1. lépés. Számítsuk ki a blokk összegeket (B) és az ismétlés összegeket (R), mint ahogy a 10.17. táblázat mutatja.

2. lépés. Számítsuk ki a kezelés összegeket (T) és a főösszeget (G), mint ahogy a 10.18. táblázat 2. oszlopa mutatja.

10.18. táblázat. Korrigált és nem korrigált kezelésösszegek kiszámítása a 4×4 -es kiegyensúlyozott latintégla elrendezésű adatokra a 2. táblázatból

Kezelés-szám	Kezelés összeg (T)	Blokk összeg (B _i)	W= 4T-5B _i +G	T'= T+μW	M' = $\frac{T'}{5}$
1	809	3286	552	829	166
2	794	3322	312	805	161
3	908	3411	323	920	184
4	901	3596	-630	878	176
5	816	3411	-45	814	163
6	848	3310	588	869	174
7	864	3562	-608	842	168
8	865	3332	546	885	177
9	801	3312	390	815	163
10	581	3141	365	594	119
11	946	3534	-140	941	188
12	971	3628	-510	953	191
13	869	3564	-598	848	170
14	994	3588	218	986	197
15	913	3394	428	928	186
16	866	3593	-755	839	168
Összes	13746 (G)	54983	0	-	-

3. lépés. Mindegyik kezelésre számítsuk ki a B_t értéket, mint a szummáját azoknak a blokk összegeknek, amelyben a szóban forgó kezelés előfordul. Például az 5. kezelés a példánkban megtalálható a 2, 5, 10, 15 és 20-as jelű blokkban. Tehát a B_t -t az 5. kezelésre úgy számítjuk ki, mint a 2, 5, 10, 15 és 20-as blokkok összegét összeadjuk, vagyis $B_5 = 616 + 639 + 654 + 675 + 827 = 3411$.

A B_t értékeket mind a 16 kezelésre a 10.18. táblázat 3. oszlopa mutatja. Jegyezzük meg, hogy a B_t értékek összege minden kezelés figyelembevételével egyenlő $(k)(G)$ -vel, ahol k a blokkok száma.

4. lépés. Mindegyik kezelésre számítsuk ki a W értékeket:

$$W = kT - (k + 1) B_t + G$$

Például, W értékét az 5. kezelésre a következőképpen számítjuk ki:

$$W_5 = 4(816) - (5)(3411) + 13746 = -45$$

A W értékeket mind a 16 kezelésre a 10.18. táblázat 4. oszlopa tartalmazza. Jegyezzük meg, hogy a W értékek összege minden kezelés figyelembevételével 0-val egyenlő.

5. lépés. Készítsük el a varianciatáblázatot, feltüntetve a variancia forrását és a szabadságfokokat:

Variáció forrása	Szabadságfokok (d.f.) száma	SS	MS	F-érték
Ismétlés	$k = 4$			
Kezelés (nem korrigált)	$k^2 - 1 = 15$			
Blokk (korrigált)	$k^2 - 1 = 15$			
Blokkon belüli hiba	$(k-1)(k^2-1) = 45$			
Kezelés (korrigált)	$[(k^2 - 1) = 15]$			
Tényleges hiba	$[(k-1)(k^2-1) = 45]$			
Összes	$k^2(k+1) - 1 = 79$			

6. lépés. Eljutottunk most a varianciaanalízisig. Az összes SS, az ismétlés SS és a kezelés SS kiszámítása a szokásos módon történik. Számítsuk ki az összes SS-t, az ismétlés SS-t és a kezelés (nem korrigált) SS-t a következőképpen:

$$C.F. = \frac{G^2}{(k^2)(k+1)} = \frac{(13746)^2}{(16)(5)} = 2361906$$

$$\text{Összes SS} = \sum X^2 - C.F. = [(147)^2 + (152)^2 + \dots + (225)^2] - 2361906 = 58856$$

$$\text{Ismétlés SS} = \frac{\sum R^2}{k} - C.F. = \frac{(2595)^2 + (2729)^2 + \dots + (2925)^2}{16} - 2361906 = 5946$$

$$\begin{aligned} \text{Kezelés (nem korrigált) SS} &= \frac{\sum T^2}{(k+1)} - C.F. = \\ &= \frac{(809)^2 + (794)^2 + \dots + (866)^2}{5} - 2361906 = 26995 \end{aligned}$$

7. lépés. Számítsuk ki a blokk (korrigált) SS-t [azaz, négyzetösszegek a blokkra ismétlésen belül, korrigálva a kezelés hatásokra]:

$$\text{Blok (korr.) SS} = \frac{\sum_{i=1}^t W_i^2}{(k^3)(k+1)} = \frac{(552)^2 + (312)^2 + \dots + (-755)^2}{(64)(5)} = 11382$$

8. lépés. Számítsuk ki a blokkon belüli (intrablokk) hiba SS-t:

Blokkon belüli hiba SS = Összes SS – Ismétlés SS – Kezelés (nem korr.) SS – Blokk (korr.) SS = 58856 – 5946 – 26995 – 11382 = 14533

9. lépés. Számítsuk ki a blokk (korrigált) közepes négyzetes eltérést (MS) és intrablokk (blokkon belüli) hiba MS értékeket:

$$\text{Blok (korr.) MS} = \frac{\text{Blok (korr.) SS}}{k^2 - 1} = \frac{11382}{15} = 759$$

$$\text{Intrablokk hiba MS} = \frac{\text{Intrablokk hiba SS}}{(k-1)(k^2-1)} = \frac{14,533}{(3)(15)} = 323$$

10. lépés. Minden egyes kezelésre számítsuk ki a korrigált kezelés összeg T' értékeket:

$$T' = T + \mu W$$

ahol

$$\mu = \frac{\text{Blok (korr.) MS} - \text{Intrablokk hiba MS}}{k^2 [\text{Blok (korr.) MS}]}$$

Jegyezzük meg, hogy ha az intrablokk hiba MS nagyobb, mint a blokk (korrigált) MS, akkor μ -t zérónak tekintjük és nem szükséges kiigazítás a kezelésre és semmilyen további kiigazítás (korrekció) nem szükséges. Az F-vizsgálatot a kezeléshatás szignifikanciájára ekkor a szokásos módon végezzük, mint a kezelés (nem korr.) MS és az intrablokk hiba MS arányát, és a 10–14. lépés és a 17. lépés elhagyható.

A példánkban, az intrablokk hiba MS kisebb, mint a blokk (korr.) MS. Ennél fogva, a korrekciós faktort (μ) a következőképpen számítjuk ki:

$$\mu = \frac{759 - 323}{16(759)} = 0.0359$$

A T' értéket például az 5. kezelésre úgy számítjuk ki, mint: $T'_5 = 816 + (0.0359)(-45) = 814$. A T' értékek eredményeit mind a 16 kezelésre a 10.18. táblázat 5. oszlopa.

11. lépés. Mindegyik kezelésre számítsuk ki a korrigált átlagos M' értéket:

$$M' = \frac{T'}{k+1}$$

A példánkban, az M' értékét az 5. kezelésre úgy számítjuk ki, mint: $814/5 = 163$. az M' értékek mind a 16 kezelésre a 10.18. táblázat utolsó oszlopában található.

12. lépés. Számítsuk ki a korrigált kezelés MS értékeket a következő formulával:

$$\begin{aligned} \text{Kezelés (korr.) MS} &= \left[\frac{1}{(k+1)(k^2-1)} \left\| \sum T_i^2 - \frac{G^2}{k^2} \right\| \right] \\ &= \left[\frac{1}{(5)(15)} \left\{ [(829)^2 + (805)^2 + \dots + (839)^2] - \frac{(13746)^2}{16} \right\} \right] = 1602 \end{aligned}$$

13. lépés. Számítsuk ki a tényleges hiba MS értékét:

Tényleges hiba MS = (Intrablokk hiba MS)(1+kμ) = 323 [1 + 4 (0.0359)] = 369

Számítsuk ki a megfelelő CV értékeket:

$$\text{CV} = \frac{\sqrt{\text{Tényleges hiba MS}}}{\text{Foatlag}} \times 100 = \frac{\sqrt{369}}{172} \times 100 = 11.2\%$$

14. lépés. Számítsuk ki az F-értéket a kezeléskülönbségek vizsgálatához:

$$F = \frac{\text{Kezelés (korr.) MS}}{\text{Tényleges hiba MS}} = \frac{1602}{369} = 4.34$$

15. lépés. Hasonlítsuk össze a számított F-értéket a táblázati F-értékkal, $f_2 = (k^2-1) = 15$ és $f_2 = (k-1)(k^2-1) = 45$ szabadságfoknál. Mivel a számított F-érték nagyobb, mint a táblázati F-érték, 1%-os szignifikancia szinten, a kezelés különbséget erősen szignifikánsnak tekintjük.

16. lépés. A 6–9. és a 12–15. lépésben kiszámított értékeket írjuk be a varianciatáblázatba. A végső eredményt a 10.19. táblázat tartalmazza.

17. lépés. Becsüljük meg a kiegyensúlyozott latintégla elrendezéssel kapott nagyobb mértékű pontosságot, az RCB elrendezéssel szemben:

$$\text{R.E.} = \frac{100[\text{Blokk(korr.)SS} + \text{Intrablokk hiba SS}]}{k(k^2-1)(\text{Tényleges hiba MS})} = \frac{100(11382 + 14533)}{(4)(16-1)(369)} = 117\%$$

Tehát a 4 × 4-es kiegyensúlyozott latintégla elrendezés használata a kísérleti pontosságot 17%-kal növelte a fölé, amelyet az RCB elrendezésnél kaptunk volna.

10.19. táblázat. Varianciaanalízis eredménytáblázata (4x4-es kiegyensúlyozott latintégla elrendezés)

Variáció forrása	Szabadságfok száma	SS	MS	Számított F-érték	Táblázati F-érték	
					5%	1%
Ismétlés	4	5946				
Kezelés (nem korr.)	15	26995				
Blokk (korr.)	15	11382	759			
Intrablokk hiba	45	14533	323			
Kezelés (korr.)	(15)	–	1.602	4.34**	1.90	2.47
Tényleges hiba	(45)	–	369			
Összes	79	58856				

18. lépés. Szignifikáns differencia (SzD) kiszámítása

Az SzD kiszámítása a kiegyensúlyozott latintégla elrendezésű kísérlet adataiból magában foglal két fontos korrekciót: (a) a korrigált kezelésátlagot használjuk az átlagos különbség kiszámításában, (b) a tényleges hiba MS-t használjuk az átlagos különbség standard hibájának kiszámításában.

A 4×4 kiegyensúlyozott latintégla elrendezésű kísérlet kezeléseit, a kiigazított átlagokat, a kontroll kezeléshez viszonyított különbségeket az 10.20. táblázat tartalmazza. Feltételezzük, hogy a 16 kezelés egyike (a 10. kezelés) nem műtrágyázott kontroll kezelés és a kutató meghatározni kívánja, hogy a 15 műtrágyakezelésre van-e szignifikáns termésreakció. Ebből a célból a megfelelő kezelés-összehasonlítás a tervezett kezeléspár összehasonlítás, melynek során a 15 műtrágya-kezelés mindegyikét összehasonlítjuk a kontroll kezeléssel.

Az SzD kiszámításának lépései:

1. lépés. Számítsuk ki az átlagos differenciát a kontroll (10. kezelés) és a 15 műtrágyakezelés mindegyike között, felhasználva a kiigazított kezelés-átlagokat.

2. lépés. Számítsuk ki az SzD értéket α szignifikancia szinten:

$$SzD_{\alpha} = t_a \sqrt{\frac{2(\text{tényleges hiba MS})}{r}}$$

ahol a tényleges hiba MS-t a 13. lépésben határoztuk meg.

Példánkban minden kezelés ötször van megismételve és a tényleges hiba MS 45 szabadságfokkal 369. A táblázati t-érték $n = 45$ szabadságfoknál 2.016 az 5%-os szignifikancia szinten és 2.693 az 1%-os szinten. Tehát az SzD értékek a következők:

$$SzD_{5\%} = 2.016 \sqrt{\frac{2(369)}{5}} = 24 / m^2$$

$$SzD_{1\%} = 2.693 \sqrt{\frac{2(369)}{5}} = 33 / m^2$$

3. lépés. Összehasonlítjuk minden egyes kezeléspár átlagos különbségét a számított SzD értékekkel és csillaggal jelöljük a szignifikancia szinteket. Például az átlagos különbség az 1. kezelés és a kontroll kezelés között 47 hajtás m^{-2} , amely meghaladja az SzD_{1%}-ot. Mind a 15 kezelés-pár reakciója erősen szignifikáns (10.20. táblázat).

10.20. táblázat. Az átlagos hajtásszám összehasonlítása a 15 műtrágyázott és a kontroll kezelés között, felhasználva az SzD vizsgálatot

Kezelések száma	Átlag (korrigált) db m^{-2}	Eltérés a kontrolltól db m^{-2}	Kezelések száma	Átlag (korrigált) db m^{-2}	Eltérés a kontrolltól db m^{-2}
1	166	47**	9	163	44**
2	161	42**	10	119	–
3	184	65**	11	188	69**
4	176	57**	12	191	72**
5	163	44**	13	170	51**
6	174	55**	14	197	78**
7	168	49**	15	186	67**
8	177	58**	16	168	49**

**1%-os szinten szignifikáns

10.5.1.4. Variacionális GenStat programmal

A GenStat programban a menük kiválasztásának sorrendje a következő: Stats-Analysis of Variance – General ... – Lattice Design. A variancia táblázat kitöltése: Y-variate: termés, Treatments: kezelés, Replication: Ismétlés, Blocks: Blokk, Units within blocks: Parcella, Name of Pseudo-factor pointer: PS.

A GenStat program által számított korrigált kezelések eredménye teljesen meggyezik a kézi számolás eredményével. A variancia táblázat szerkezete azonban eltérő. A „standard error of differences of combined means” felhasználásával kiszámított SzD érték kismértékben eltér a kézi számolás alapján meghatározott értéktől. Az analízis outputja a következő:

Analysis of variance

Variate: Termes

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Replic stratum	4	5946.1	1486.5		
Replic.Blokk stratum					
Treatment	15	19980.0	1332.0		
Replic.Blokk.Plots stratum					
Treatment	15	18396.2	1226.4	3.80	<.001
Residual	45	14533.3	323.0		
Total	79	58855.6			

Tables of combined means

Variate: Termes

Treatment	1	2	3	4	5	6	7
	165.8	161.0	183.9	175.7	162.9	173.8	168.4
Treatment	8	9	10	11	12	13	14
	176.9	163.0	118.8	188.2	190.5	169.5	197.2
Treatment	15	16					
	185.7	167.8					

Standard errors of differences of combined means

Table	Treatment
rep.	5
s.e.d.	13.59
effective d.f.	53.72

Except when comparing means with the same level(s) of

s.e.d.	12.15
effective d.f.	53.72

Az adatokat a következő formában írjuk be a GenStat számolótáblába:

Ismétlés	Blokk	Parcella	Kezelés	Termés
1	1	1	1	147
1	1	2	2	152
1	1	3	3	167
1	1	4	4	150
1	2	5	5	127
1	2	6	6	155
1	2	7	7	162
1	2	8	8	172
1	3	9	9	147
1	3	10	10	100
1	3	11	11	192
1	3	12	12	177
1	4	13	13	155
1	4	14	14	195
1	4	15	15	192
1	4	16	16	205
2	5	17	1	140
2	5	18	5	165
2	5	19	9	182
2	5	20	13	152

10.5.2. Részben kiegyensúlyozott latintégla elrendezés

A részben kiegyensúlyozott latintégla elrendezés hasonló a kiegyensúlyozott latin-tégla elrendezéshez, azonban az ismétlések számának megválasztásában nagyobb rugalmasságot enged meg. Míg a részben kiegyensúlyozott latintégla elrendezés szükségessé teszi, hogy a kezelések számának egy pontos négyzetnek kell lenni és blokk-méret pontosan a négyzetgyöke ennek a kezelésszámnak, addig az ismétlések számára nincs előírás a kezelések számának függvényében. Valóban, az ismétlések bármely száma használható egy részben kiegyensúlyozott latintégla elrendezésben.

A részben kiegyensúlyozott latintégla elrendezést két ismétléssel *egyszerű* (simple) latintégla; három ismétléssel *háromszoros* (triple) latintégla; négy ismétléssel *négyszeres* (quadruple) és így tovább, latintégla elrendezésnek hívjuk. Az ilyen rugalmasság az ismétlések számának megválasztásában azonban a szimmetria elvesztését eredményezi a kezelések elhelyezésében a blokkokban (azaz, bizonyos kezelés-párok sohasem jelennek meg együtt ugyanabban az inkomplett blokkban. Következésképpen, a kezelési-párokat, amelyeket ugyanabban az inkomplett blokkban tesztelünk, hasonlítjuk össze a pontosságnak egy olyan szintjével, amely magasabb, mint azoknál a kezelési-pároknál, melyeket nem ugyanabban az inkomplett blokkban tesztelünk. A ke-

zelésátlagok összehasonlításának egynél több szintje miatt az adatelemzés bonyolultabb. A latintégla elrendezésre a biometriai szakkönyvekben kísérleti tervek állnak rendelkezésre. Az analízis kézi számításra alapuló módszerét *Gomez és Gomez (1984)* ismerteti részletesen.

10.5.2.1. Kísérlet leírása

Egy gabonanemesítő ki akart fejleszteni egy új, nagy termékenyséű árpaajtát. A folyamat részeként kiválasztott 22 ígéretes vonalat, hogy tesztelje három kontroll fajtával, ismétléses fajtakísérletben. A kísérleti terület talaj-variabilitása miatt elhatározta, hogy egy egyszerű (simple), 5×5 -ös latintégla elrendezést használ, két ismétléssel. A kísérlet szántóföldi elrendezését és a termést (kg parcella⁻¹) az alábbi táblázat mutatja. A kézi számolás módszerét *Petersen (1984)* ismerteti.

Az árpa termése (kg parcella⁻¹) az egyszerű, 5×5 -ös latintégla elrendezésű kísérletben

Ismétlés	Blokk	Termés kg parcella ⁻¹				
		(19) [†] 18.2	(16) 13.0	(18) 9.5	(17) 6.7	(20) 10.1
1	1	(12) 13.3	(13) 11.4	(15) 14.2	(14) 11.9	(11) 13.4
	2	(1) 15.0	(2) 12.4	(3) 17.3	(4) 20.5	(5) 13.0
	3	(22) 7.0	(24) 5.9	(21) 14.0	(25) 19.2	(23) 7.8
	4	(9) 11.9	(7) 15.2	(10) 17.2	(8) 16.3	(6) 16.0
	5	(23) 7.7	(18) 15.2	(3) 19.1	(8) 15.5	(13) 14.7
2	1	(5) 15.8	(20) 18.0	(10) 18.8	(15) 14.4	(25) 20.0
	2	(22) 10.2	(12) 11.5	(2) 17.0	(17) 11.0	(7) 15.3
	3	(14) 10.9	(24) 4.7	(9) 10.9	(4) 16.6	(19) 9.8
	4	(6) 20.0	(16) 21.1	(11) 16.9	(21) 10.9	(1) 15.0
	5					

[†]Szelektációs szám van zárójelben.

10.5.2.2. Varianciaanalízis GenStat programmal

A kísérlet termésadait GenStat programmal értékeltük. Az adatok bevitele és a program menük kiválasztásának sorrendje megegyezik a 10.5.1.4. fejezetben leírtakkal. A számítógépes értékelés outputja sorrendben a következőket tartalmazza: variancia táblázat, nem korrigált termésátlagok, átlagok hibája (szórása), átlagok közötti különbségek hibája (s_d , SED), szignifikáns differencia, kombinált (korrigált) átlagok és a hozzájuk tartozó statisztikai mutatók.

Analysis of variance

Variate: Termes

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Ismetles stratum	1	18.727	18.727		
Ismetles.Blokk stratum					
Kezeles	8	252.666	31.583		
Ismetles.Blokk.Parcella stratum					
Kezeles	24	447.013	18.626	3.43	0.007
Residual	16	86.975	5.436		
Total	49	805.381			

Tables of means

Variate: Termes

Grand mean 13.83

Kezeles	1	2	3	4	5	6	7
PS[1]	3	3	3	3	3	5	5
PS[2]	5	3	1	4	2	5	3
	14.28	14.19	17.74	20.63	13.60	17.14	14.60
Kezeles	8	9	10	11	12	13	14
PS[1]	5	5	5	2	2	2	2
PS[2]	1	4	2	5	3	1	4
	15.30	13.34	17.06	14.32	11.78	12.48	13.37
Kezeles	15	16	17	18	19	20	21
PS[1]	2	1	1	1	1	1	4
PS[2]	2	5	3	1	4	2	5
	13.39	17.56	9.57	13.12	17.31	14.48	11.16
Kezeles	22	23	24	25			
PS[1]	4	4	4	4			
PS[2]	3	1	4	2			
	7.52	6.72	6.81	18.23			

Standard errors of means

Table	Kezeles
rep.	2
d.f.	16
e.s.e.	1.951

Except when comparing means with the same level(s) of

PS[1]	1.806
PS[2]	1.806

Standard errors of differences of means

Table	Kezeles
rep.	2
d.f.	16
s.e.d.	2.759

Except when comparing means with the same level(s) of

PS[1]	2.554
PS[2]	2.554

Least significant differences of means (5% level)

Table	Kezeles
rep.	2
d.f.	16
l.s.d.	5.848

Except when comparing means with the same level(s) of

PS[1]	5.414
PS[2]	5.414

Tables of combined means

Variate: Termes

Kezeles	1	2	3	4	5	6	7
PS[1]	3	3	3	3	3	5	5
PS[2]	5	3	1	4	2	5	3
	14.68	14.48	18.00	19.46	14.05	17.62	14.96
Kezeles	8	9	10	11	12	13	14
PS[1]	5	5	5	2	2	2	2
PS[2]	1	4	2	5	3	1	4
	15.64	12.25	17.59	14.79	12.13	12.80	12.27
Kezeles	15	16	17	18	19	20	21
PS[1]	2	1	1	1	1	1	4
PS[2]	2	5	3	1	4	2	5
	13.90	17.27	9.17	12.69	15.45	14.24	11.88
Kezeles	22	23	24	25			
PS[1]	4	4	4	4			
PS[2]	3	1	4	2			
	8.13	7.30	5.96	19.00			

Standard errors of differences of combined means

Table	Kezeles
rep.	2
s.e.d.	2.528
effective d.f.	22.69

Except when comparing means with the same level(s) of

PS[1]	2.432
effective d.f.	20.86
PS[2]	2.432
effective d.f.	20.86

Standard errors of combined means

Table	Kezeles
rep.	2
e.s.e.	1.788
effective d.f.	22.69

Except when comparing means with the same level(s) of

PS[1]	1.720
effective d.f.	20.86
PS[2]	1.720
effective d.f.	20.86

10.5.2.3. Az eredmények értelmezése

Egy gabonanemesítő beállított egy fajtakísérletet, hogy értékelje 22 új árpafajta terméspotenciálját a három kontroll fajtahoz viszonyítva. A kísérleti hely varianciája miatt a kísérletet 5×5 -ös egyszerű latintégla elrendezésben állította be. Az 1, 13 és 25 számú fajta volt a kontroll. A 13-as fajta volt a hosszú távú kontroll, az 1-es fajtát régóta használták a termesztésben a régióban, és a 25-ös volt az utolsó nemesített fajta a kísérleti állomáson. A terméseredményeket a fenti Genstat output tartalmazza (Tables of combined means).

Az adatsorra a szignifikáns különbség (SzD) $P = 5\%$ -os szinten 4.41 volt. (Jegyezzük meg, hogy t_{α} egyoldalú, az 5% -os t érték 16 hiba d.f.-nél 1.746.) Tehát öt új fajta nagyobb termést adott, mint a hosszú távú kontroll, melynek átlaga $12.80 \text{ kg parcella}^{-1}$ volt. Egy új nemesített fajtának (4-es fajta) $19.46 \text{ kg parcella}^{-1}$ termése szignifikánsan nagyobb volt, mint a helyi kontroll (1-es fajta) $14.69 \text{ kg parcella}^{-1}$ termése. Egyetlen újonnan szelektált fajta termése sem haladta meg a legutoljára nemesített fajtáét, melynek termése $19.0 \text{ kg parcella}^{-1}$ volt.

10.6. Kezelés struktúra

Számos kísérletben a kezeléseknak van egy bizonyos struktúrája (szerkezete), úgy-hogy sajátos összehasonlításokat tehetünk a kezelések csoportjai között és csoport-jain belül, minden lehetséges kezelés-pár egyszerű összehasonlítása helyett (Mead *et al.* 2003).

A kezelés struktúrára egy egyszerű példa egy kísérlet, amelyben számos új gabo-natörzset hasonlítottak össze egy standard törzssel. Ebben a helyzetben két alaphipo-tézis lehet, melyet a kísérletező tesztelni kíván: először, vajon az új törzsek, mint egy csoport mutat-e különbséget a termésben a kontrollhoz viszonyítva, és másodsor, vajon jelentős variáció van-e az új törzsek között. Ezeknek a hipotéziseknek a tesztelé-sére feloszthatjuk a kezelés SS-t komponensekre, amelyek reprezentálják (1) a külön-biséget a kontroll és az új törzsek között, és (2) a variációt az új törzsek között. Tételez-zük fel, hogy van öt új törzsünk és egy kontroll egy randomizált blokk kísérletben, négy blokkban (ismétlésben), összesen 24 parcellával. Ha kezelés összegeket T_1, T_2, T_3, T_4, T_5 és T_C a kezelés SS a következő:

$$\frac{T_1^2 + T_2^2 + T_3^2 + T_4^2 + T_5^2 + T_C^2}{4} - \frac{G^2}{24}$$

ahol G a termés főátlaga. Az összehasonlítás az új törzsek és a kontroll között úgy tekinthető, mint egy összehasonlítás a 20 kísérleti egység (négy az öt új törzs mind-egyikére) átlaga és a négy egység átlaga között. Ennek a különbségnek megfelelő SS:

$$\frac{T_N^2}{20} + \frac{T_C^2}{4} - \frac{G^2}{24}$$

ahol T_N a T_1, T_2, T_3, T_4 és T_5 összege, vagyis az öt új törzs termésösszege. Ahhoz, hogy mérjük a variációt az új törzsek között, kiszámítjuk az SS-t az öt kezelés közötti variá-cióra a következőképpen:

$$\frac{T_1^2 + T_2^2 + T_3^2 + T_4^2 + T_5^2}{4} - \frac{T_N^2}{24}$$

Ez természetesen a kezelés SS, melyet akkor számítottunk volna ki, ha a kísérlet csak az öt új törzset tartalmazta volna. Ennek a két SS, (2) és (3) kifejezések, 1 és 4 d.f. szabadságfokkal rendelkezik, míg a (2) és (3) kifejezés összege a teljes kezelés SS (1). A két SS megfelel a variáció különböző okainak és mindegyik MS összehasonlítható a hiba MS-sel, felhasználva a megfelelő F-eloszlást a variáció szignifikanciájának tesz-

telésére. Bármely két kezelésátlag közötti különbség SE-ja $\sqrt{(2s^2/4)}$, és a kontroll átlag, valamint az öt új törzs átlaga közötti különbség SE-ja $\sqrt{(s^2/4 + s^2/20)}$, mivel a két átlagot 4 és 20 egységből számítottuk ki.

Ha a kontroll kezelést többször ismétljük meg minden blokkban, ugyanezt a mód-szert alkalmazzuk. Például, tételezzük fel, hogy van öt blokkunk, mindegyik tartal-

maz hat új törzset és a kontroll négyszeri ismétlését. Tételezzük fel továbbá, hogy a négy kontroll kezelés mindegyik blokkban azonosítható és jelölésük 7, 8, 9 és 10. Kísérletünk úgy tekinthető, hogy tíz kezelést tartalmaz, vagyis 1, 2, 3, 4, 5, 6, C₁, C₂, C₃ és C₄; a teljes kezelés SS kiszámítható a következőképpen:

$$\frac{T_1^2 + T_2^2 + T_3^2 + T_4^2 + T_5^2 + T_6^2 + T_7^2 + T_8^2 + T_9^2 + T_{10}^2}{5} - \frac{G^2}{50}$$

Ez felosztható három komponensre:

1. Variáció mérése a kontroll és az új törzsek között:

$$\frac{T_N^2}{30} + \frac{T_C^2}{20} - \frac{G^2}{50}$$

2. Variáció mérése az új törzsek között:

$$\frac{T_1^2 + T_2^2 + T_3^2 + T_4^2 + T_5^2 + T_6^2}{5} - \frac{T_N^2}{30}$$

3. Variáció mérése a kontrollok között:

$$\frac{T_7^2 + T_8^2 + T_9^2 + T_{10}^2}{5} - \frac{T_C^2}{20}$$

Természetesen az MS ebből az utolsó SS-ből körülbelül ugyanakkora lenne, mint a hiba MS, mivel nincs valós különbség a 7, 8, 9 és 10 kezelés között. A másik két komponens, (1) és (2), melyet ezt megelőzően számítottunk ki, és mindegyik MS összehasonlítható a hiba MS-sel.

A kontrollok és az új kezelések közötti összehasonlítás a kezelés összehasonlítás legegyszerűbb formája a két kezelésátlag alapvető összehasonlítását követően. Amikor a kezelések egy kísérletben komplikáltabb struktúrával rendelkeznek, kívánatos lehet felbontani a kezelés SS-t számos komponensre, mindegyik megfelel a kezelések bizonyos csoportjai közötti összehasonlításnak. Ez hasonló módon végezhető, mint amit használtunk a kontrollokat és új törzseket tartalmazó kísérletben. A kezeléscsoportok közötti minden összehasonlításra kiszámítjuk az összegeket a különböző csoportokra, T₁, T₂ ...; ezeknek az összegeknek a szummáját, $\sum T = T_1 + T_2 + \dots$. Az SS-t ezután így számítjuk ki:

$$\frac{T_1^2}{r_1} + \frac{T_2^2}{r_2} + \dots - \frac{[\sum(T)]^2}{\sum r}$$

ahol r₁ az egységek (parcellák) összes száma, amely hozzájárul az összes T₁-hez, r₂ az összes T₂-höz, ..., és $\sum(r) = r_1 + r_2 + \dots$. Az alapvető szabály az, hogy elosztunk minden négyzetösszeget azoknak az egységeknek (parcella) a számával, melyeket bevontunk ebbe az összeg számításba. A szabadságfok minden SS-re a csoportok száma mínusz 1.

10.6.1. Kísérlet leírása

Kezelés SS felbontása összehasonlításra

Példaként vegyünk egy tavaszi káposzta kísérletet, melyben nyolc kezelést állítottak be minden blokkban, véletlen blokk elrendezésben, négy ismétlésben. A kísérletben három nitrogén forrást hasonlítottak össze, mindegyiket két szinten, egy nitrogén nélküli kontroll kezeléssel, a kontroll két ismétlésben volt beállítva minden blokkban. Az adatok a következők (Mead *et al.*, 2003):

	Kezelés szintje	Blokk				Összesen
		I	II	III	IV	
Nitro-chalk	1	70.3	72.5	79.0	86.2	308.0
	2	61.0	75.1	71.3	65.2	272.6
Ammónia szulfát	1	75.5	63.0	65.4	67.7	271.6
	2	64.7	66.0	64.3	68.8	263.8
Nitrát	1	85.2	80.5	83.6	92.3	341.6
	2	65.9	78.7	71.1	76.8	292.5
Kontroll	1	35.7	39.6	45.5	50.5	171.3
	2	38.5	36.3	42.1	61.4	178.3
		496.8	511.7	522.3	568.9	2099.7

10.6.2. A varianciaanalízis lépései

A kezdeti varianciaanalízis nem veszi figyelembe a kezelések struktúráját és a számítógépes analízishez az információ a következő:

Termés	Blokk	Kezelés
70.3	1	1
72.5	2	1
79.0	3	1
86.2	4	1
61.0	1	2
75.1 ...	2....	2 ...
36.3	2	8
42.1	3	8
61.4	4	8

A modell a következő:

Blokk + Kezelés

Variáciatáblázat:

Variáció forrása	SS	d.f.	MS	F-érték
Blokk	363.33	3		
Kezelés	6205.16	7	886.45	26.36***
Hiba	706.57	21	33.65 = s ²	
Összes	7257.06	31		
c.f.	137773.12			

A kezelés SS-t feloszthatjuk számos hasznos komponensre. Ehhez több összegre (total) van szükség:

Két kontroll kezelés összege: 349.6

Két nitro-chalk szint összege: 580.6

Két ammónia szint összege: 535.4

Két nitrát szint összege: 634.1

Hat nitrogén kezelés összege: 1750.1

Az első komponens összehasonlítás a kontrollok és a nitrogén kezelések között:

$$\frac{(349.6)^2}{8} + \frac{(1750.1)^2}{24} - c.f. = 5123.15 \quad c.f. = \frac{(2099.7)^2}{32} = 137773.12$$

A második komponens összehasonlítás a három nitrogén forrás között:

$$\frac{(580.6)^2 + (535.4)^2 + (634.1)^2}{8} - \frac{(1750.1)^2}{24} = 610.29$$

A harmadik összehasonlítás a két szint között mindegyik nitrogén forrásra:

$$\frac{(308.0)^2 + (272.6)^2}{4} - \frac{(580.6)^2}{8} = 156.65 \quad (\text{nitro-chalk})$$

$$\frac{(271.6)^2 + (263.8)^2}{4} - \frac{(535.4)^2}{8} = 7.61 \quad (\text{ammónia-szulfát})$$

$$\frac{(341.6)^2 + (292.5)^2}{4} - \frac{(634.1)^2}{8} = 301.35 \quad (\text{nitrát})$$

Végül, kiszámíthatunk egy komponenset a két kontroll kezelés közötti variációra (melynek hasonlóknak kellene lenni a hiba MS-hez, mivel a kontroll kezelések azonosak voltak):

$$\frac{(171.3)^2 + (178.3)^2}{4} - \frac{(349.6)^2}{8} = 6.12$$

Az utolsó komponens SS-t rendszerint nem számítják ki, a teljesség kedvéért mutatjuk be.

A kezelés SS felbontását a következőképpen tudjuk összegezni:

Variáció forrása	SS	d.f.	MS	F-érték
Kontroll vs. nitrogén	5123.15	1	5123.15	152.2***
Nitrogén források	610.29	2	305.15	9.1**
Nitro-chalk szintjei	156.65	1	156.65	4.7*
Ammónia szintjei	7.60	1	7.60	0.2
Nitrát szintjei	301.35	1	301.35	9.0**
Kontrollok között	6.12	1	6.12	0.2
Összes kezelés	6205.16	7	886.45	26.4***
Hiba	706.57	21	33.65	

Az eredmények bemutatására kiszámítjuk a kezelésközvetlen átlagokat és az átlagok közötti különbségek különböző SE értékeit.

	Kezelés átlagok		
	1. szint	2. szint	Átlag
Nitro-chalk	77.0	68.2	72.6
Ammónia	67.9	66.0	67.0
Nitrát	85.4	73.1	79.3
Átlag			72.9
Kontroll			43.7

A SE-okat a következőképpen számítjuk ki:

(a) A kontroll kezelésközvetlen átlag és a hat nitrogén kezelés összehasonlítására:

$$\sqrt{\left(\frac{s^2}{8} + \frac{s^2}{24}\right)} = \sqrt{\left[33.65\left(\frac{1}{8} + \frac{1}{24}\right)\right]} = 2.37$$

(b) Kontroll kezelés összehasonlítására egy nitrogén forrás átlaggal vagy két nitrogén forrás összehasonlítására:

$$\sqrt{\left(\frac{s^2}{8} + \frac{s^2}{8}\right)} = \sqrt{(33.65/4)} = 2.90$$

(c) Kontroll átlag összehasonlítására bármely forrás bármely szintjével:

$$\sqrt{\left(\frac{s^2}{8} + \frac{s^2}{4}\right)} = \sqrt{[33.65(3/8)]} = 3.55$$

(d) Bármely két átlag összehasonlítására azonos vagy különböző források sajátos szintjére:

$$\sqrt{(2s^2 / 4)} = \sqrt{(33.65 / 2)} = 4.10$$

Eredmények összefoglalása: Minden nitrogén kezelés jelentősen növelte a termést, a kontroll kezeléssel összehasonlítva. A növekedés 22 és 42 között változott a különböző forrásoknál, a legnagyobb átlagos növekedés (35) a nitrátnak tulajdonítható. Az 1. szint nagyobb termést eredményezett, mint a 2. szint minden forrásnál, jól lehet a különbség a nitro-chalk és a nitrát között volt szignifikáns.

10.6.3. Varianciaanalízis GenStat programmal

A menük kiválasztásának sorrendje: Stats – Analysis of Variance – General ... - One-way ANOVA (in Randomized Blocks). Y-variate: Termés, Treatments: Kezelés, Blocks: Ismétlés. A kontrasztok számításának sorrendje: Contrasts-ra kattintunk a One-way ANOVA fő menüben, Contrast Factor: Kezelés, Contrast Type: Comparison, Number of Contrasts: 5. A megjelenő kontraszt táblázatban a contrast 1–5-öt kitöltjük az alábbi outputnak megfelelően. A Run-ra kattintva megkapjuk az ANOVA eredményét.

Analysis of variance

Contrast 1	1	1	1	1	1	1	-3	3
Contrast 2	1	-1	0	0	0	0	0	0
Contrast 3	0	0	1	-1	0	0	0	0
Contrast 4	0	0	0	0	1	-1	0	0
Contrast 5	0	0	0	0	0	0	1	-1

Variate: Termes

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Ismetles stratum	3	363.33	121.11	3.60	
Ismetles.*Units* stratum					
Kezeles	7	6205.16	886.45	26.35	<.001
Contrast 1	1	5123.14	5123.14	152.26	<.001
Contrast 2	1	156.65	156.65	4.66	0.043
Contrast 3	1	7.61	7.61	0.23	0.639
Contrast 4	1	301.35	301.35	8.96	0.007
Contrast 5	1	6.12	6.12	0.18	0.674
Residual	21	706.58	33.65		
Total	31	7275.06			

Tables of means

Variate: Termes

Grand mean 65.6

Kezeles	1	2	3	4	5	6	7
	77.0	68.1	67.9	65.9	85.4	73.1	42.8
Kezeles	8						
	44.6						

Standard errors of differences of means

Table	Kezeles
rep.	4
d.f.	21
s.e.d.	4.10

Least significant differences of means (5% level)

Table	Kezeles
rep.	4
d.f.	21
l.s.d.	8.53

A „Nitrogén források” SS értékét megkapjuk, ha az 1–5. kontraszt SS összegét kivonjuk a Kezelés SS-ből, azaz $6205.16 - 5594.87 = 610.29$. A Kezelés 7 d.f. értékéből 5 d.f.-t használtunk fel a kontrasztokkal, tehát a „Nitrogén Források” szabadságfoka 2.

10.6.4. Faktoriális kezelés struktúra

Számos kísérletben, a javasolt kezelésstruktúra csak akkor alakul ki, amikor a kísérletező meghatározza a kísérlet célját, és kezdetben gyakran kevésbé nyilvánvaló lesz a mintázat a kezelések és kölcsönös kapcsolataik választásában. Számos előnye van azonban annak, amikor a kezeléseket határozott struktúra mintázattal használjuk, mint azt egyszerűen mutatja a négy tápanyag-ellátottságnál termesztett káposzta hozam (termés) adatsora (Mead *et al.*, 1993).

Műtrágya	Termés (t acre ⁻¹)
(1) 0 P, 0 K	20.5
(2) 0 P, 1 cwt K	19.8
(3) 0.6 cwt P ₂ O ₅ , 0 K	26.7
(4) 0.6 cwt P ₂ O ₅ , 1 cwt K ₂ O	30.5

1 acre = 0.4047 ha és 1 cwt = 50.8 kg

A nyilvánvaló következtetés ezekből az adatokból, feltételezve, hogy a termések SE értékei megfelelően kicsik, hogy a P alkalmazásának hatása a termésre függ a K alkalmazásának szintjétől és vice versa. Tehát, ha nem alkalmazunk K-t, 0.6 cwt P₂O₅ 6.2 t acre⁻¹-rel növeli a termést, míg ha 1 cwt K₂O-t alkalmazunk, 0.6 cwt P₂O₅ a termést 10.7 t acre⁻¹-rel növeli. Fordítva, ha nem alkalmazunk P-t, K alkalmazása nem növeli a termést, azonban, ha P-t használunk, akkor a K 4 tonnával növeli a termést acre-enként.

Ebben a kísérletben a négy kezelésnek egy világosan definiálható struktúrája van, magában foglalja a P (0 vagy 0.6) két szintje egyikének minden lehetséges kombinációját a K (0 vagy 1) két szintjének egyikével. Ezt világosan mutatja a kétirányú táblázat.

	0 P	0.6 cwt P ₂ O ₅
0 K	20.5	26.7
1 cwt K ₂ O	19.8	30.5

A kezelésstruktúrának ezt a formáját hívjuk „faktoriálisnak” és lényeges, amikor számos különböző típusú kezelés direkt hatásait és kombinált kereszt irányú hatásait (interakcióit) vizsgáljuk. Ebben a stádiumban be kell vezetni néhány definíciót:

Faktor: Egy típusa a kezelésnek (pl. fajta, műtrágya, vagy takarmány).

Szint: Egy faktor sorozatnak egy sajátos kezelése (pl. A fajta, vagy 50 kg K₂O, vagy 12% fehérje).

Kísérleti kezelés: Minden faktor egy szintjének kombinációja a kísérletben.

Komplett faktoriális kísérlet: Egy kísérlet, amely tartalmazza az összes lehetséges kísérleti kezelést, lehetőség szerint több ismétlésben.

E kifejezések használatának bemutatására vegyünk egy kísérletet, amelyben összehasonlítjuk a P két szintjének és a K két szintjének a hatását három káposzta-fajtára. A kísérlet három faktort (P, K és fajta) tartalmaz és ezeknek a faktoroknak két-két szintje (P és K: hiány és alkalmazás), illetve három szintje (V₁, V₂ és V₃ fajta) van. Egy tipikus kísérleti kezelés pl. tartalmazza a V₂ növényfajtát, melyet 0 P és 1 cwt K₂O kijuttatásával termesztünk. 12 (2×2×3) különböző kísérleti kezelés van és a teljes kísérlet tartalmaz három randomizált blokkot, mindegyik blokkban 12 parcellával. Egy ilyen kísérletet 2×2×3-as faktoriális (vagy néha 2² × 3 faktoriális) kísérletként jelöljük.

Amikor a kísérleti kezeléseket faktoriális kísérletben tervezzük, a kezeléseket véletlenszerűen helyezzük el a parcellákra (kísérleti egységekre) a blokkban, a kezelésstruktúrától eltekintünk. Ennek megfelelő módja az, hogy beszámozzuk a kísérleti kezeléseket egymás után 1-től kezdődően és ezután a számokat véletlenszerűen helyezzük el a parcellákra. A 2×2×3-as faktoriális kísérlet kezeléseinek számozása a következő lehet:

1. 0 P, 0 K, V₁ fajta
2. 0 P, 0 K, V₂ fajta
3. 0 P, 0 K, V₃ fajta
4. 0 P, 1 cwt K₂O, V₁ fajta
5. 0 P, 1 cwt K₂O, V₂ fajta
6. 0 P, 1 cwt K₂O, V₃ fajta
7. 0.6 cwt P₂O₅, 0 K, V₁ fajta
8. 0.6 cwt P₂O₅, 0 K, V₂ fajta
9. 0.6 cwt P₂O₅, 0 K, V₃ fajta
10. 0.6 cwt P₂O₅, 1 cwt K₂O, V₁ fajta
11. 0.6 cwt P₂O₅, 1 cwt K₂O, V₂ fajta
12. 0.6 cwt P₂O₅, 1 cwt K₂O, V₃ fajta

A kezelések randomizált sorrendje az első blokkban lehet: 1, 7, 4, 10, 12, 5, 8, 3, 6, 11, 9, 2 és a teljes, randomizált kísérletet a 10.21. táblázat mutatja, a kezelések részletes jelölése csak az 1. blokkban van megadva.

10.21. táblázat A 2×2×3-as faktoriális kísérlet kezeléseinek randomizált elrendezése három blokkban

1. Blokk	2. Blokk	3. Blokk
1. 0 P, 0 K, V ₁ fajta	5	6
7. 0.6 cwt P ₂ O ₅ , 0 K, V ₁ fajta	4	2
4. 0 P, 1 cwt K ₂ O, V ₁ fajta	11	11
10. 0.6 cwt P ₂ O ₅ , 1 cwt K ₂ O, V ₁ fajta	1	7
12. 0.6 cwt P ₂ O ₅ , 1 cwt K ₂ O, V ₃ fajta	7	8
5. 0 P, 1 cwt K ₂ O, V ₂ fajta	12	1
8. 0.6 cwt P ₂ O ₅ , 0 K, V ₂ fajta	2	12
3. 0 P, 0 K, V ₃ fajta	3	9
6. 0 P, 1 cwt K ₂ O, V ₃ fajta	9	10
11. 0.6 cwt P ₂ O ₅ , 1 cwt K ₂ O, V ₂ fajta	6	4
9. 0.6 cwt P ₂ O ₅ , 0 K, V ₃ fajta	10	5
2. 0 P, 0 K, V ₂ fajta	8	3

11. Kéttényezős kísérletek varianciaanalízise

Az előző fejezetekben bemutattuk az egytényezős kísérletek analízisét a szignifikáns kezelés-különbségek megállapítására. Gyakran azonban egynél több faktort vizsgálunk egyidejűleg ugyanabban a kísérletben, mindegyik faktort különböző szinteken alkalmazva. Például a műtrágya hatását a növény növekedésére tesztelhetjük számos dózisban, különböző talajtípusokon. Ebben a kísérletben két faktort vizsgálunk egyidejűleg, nevezetesen a műtrágyát és a talajt, és ezt kéttényezős kísérletnek hívjuk. Ily módon egy háromtényezős kísérlet is tervezhető, hogy megvizsgáljuk három tényező hatását. Például a műtrágya hatását vizsgálhatjuk különböző növényfajtákon, melyeket különböző talajtípusokon állítottunk be ugyanabban a kísérletben. A gyakorlatban a háromtényezős kísérletek viszonylag ritkák, azonban a számítógépek korszakában könnyen analízálhatók. Jóllehet elméletileg lehetséges, a háromnál többtényezős kísérlet oly sok forrást és követelményt támaszt, mint amilyen az ismétlések nagy száma a statisztikai analízishez, hogy többé-kevésbé ismeretlen.

A többtényezős kísérleteknek egy másik típusa, amellyel találkozunk a mezőgazdaságban és a kertészetben, különösen szántóföldön, a 2^k kísérlet. Ez olyan kísérlet, amelyben k kezelés faktor van, ahol k ténylegesen egy elég nagy szám lehet, azonban mindegyik faktort csupán két szinten vizsgáljuk. Például, három takarmány-kiegészítés kombinált hatása egy állatetelési kísérletben a növekedés rátájára vizsgálható lenne egyszerűen úgy, hogy hozzáadjuk vagy elvonjuk mindegyik kiegészítőt a takarmányból. Ez összesen 2^3 kezelést ad, mindegyik megfelelő számú ismétlésben. Ilyen kísérlet lehet teljesen randomizált, azonban gyakrabban randomizált blokk elrendezésként állítjuk be.

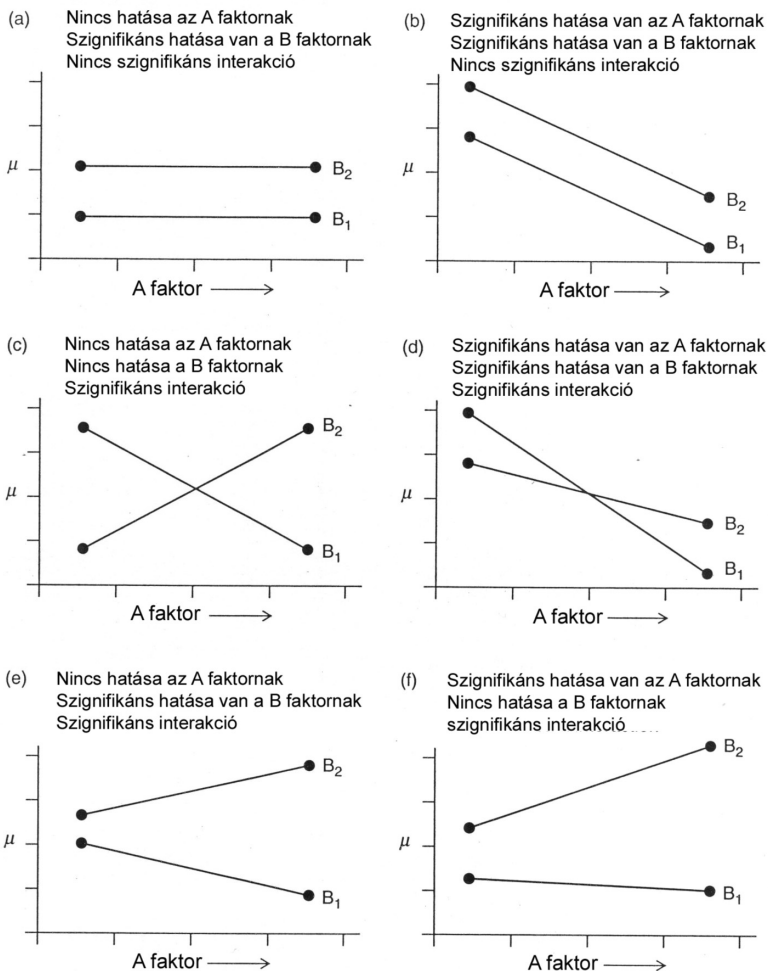
A legfontosabb különbség az egyfaktoros (egytényezős) és a két (vagy több) faktoros kísérletek között az, hogy az utóbbi esetben az egyedi kezelés-hatásokon kívül a kezelés faktorok közötti interakció (kölsönhatás) lehetősége is fennáll, mint egy lehetséges fontos hatás, amely befolyásolja az eredményt és analízist indokol.

Kezelések közötti interakció

Amikor két vagy több kezelés faktort alkalmazunk egyidejűleg egy többtényezős kísérletben, fennáll a kezelés faktorok közötti interakció lehetősége. Az „interakció” kifejezés alatt azt értjük, hogy egy kezelés hatása módosíthatja a szubjektum (növény, állat) reakcióját egy másik egyidejű kezelésre. Például, egy szövetkultúra kísérletben egy bizonyos növény-növekedési hormon szükséges optimális koncentrációját a maximális növekedési rátához vizsgálták különböző cukor koncentrációnál a növekedési közegben. Úgy találták, hogy a zeatin hormon bizonyos szintjének hozzáadása önmagában 10%-os növekedési ráta emelkedést eredményezett és a cukor bizonyos szintje önmagában hasonlóan 10%-os növekedést adott. Ha azonban a zeatint és a cukrot együtt adták, 25%-os növekedési ráta emelkedést kaptak. A hormon és a cukor együtt olyan hatást eredményezett, amely nagyobb volt az egyedi hatások összegénél. A reak-

ciónak ezt a típusát *pozitív szinergista interakciónak* hívjuk. Néha azonban negatív interakció is bekövetkezhet, amikor a reakció egy kombinált kezelésre csökken, összehasonlítva az egyedi kezelés-hatások összegével. Az interakciók néhány típusát, amely bekövetkezhet a kéttényezős kísérletekben, az 11.1. ábra mutatja.

Fontos megérteni, hogy egy szignifikáns interakció megléte független bármely más szignifikáns hatás meglététől. Például, egy üvegházi kísérletben, az uborka termése nem mutatott szignifikáns reakciót a kiegészítő világítás szintjének növelésére és hasonlóan nem mutatott szignifikáns reakciót az üvegház léghőmérsékletének a CO_2 koncentráció növelésére. Amikor azonban mind a fénynek és mind a CO_2 koncentrációnak a szintjét együtt növelték, egy szignifikáns termésmenyeget következett be. Ez két kezelés közötti pozitív interakcióra volt példa, olyan esetre, amikor önmagában egyik kezelés sem volt képes szignifikáns hatást kiváltani (11.1.c. ábra).



11.1. ábra. A faktorok közötti interakció különböző típusainak bemutatása, amely bekövetkezhet egy kéttényezős kísérletben. Az A faktor növekvő szintjeinek hatása a B faktor két szintjén a μ populáció átlagára hat különböző esetre (Ireland, 2010 alapján)

A kezelések közötti interakció nagyon fontos a növénytermesztésben. Nagyon gyakran az egyedi kezelés-hatások nem ismertek előre és kérdéses, hogy vajon kölcsönösen hatnak-e egymásra, adva egy olyan reakciót, amely azután a kutató elsődleges érdeklődésévé válik és a fő oka lesz többtényezős kísérlet beállításának.

Kéttényezős varianciaanalízis (ANOVA) alapelvei és lépései

A kéttényezős kísérletek magukban foglalják két kezelés faktor egyidejű alkalmazását, mindegyiket egynél több szinten. Az ilyen kísérletek szükségessé tesznek egy olyan analízist, amely képes elkülöníteni a kombinált kezeléshatást komponenseire és ezáltal lehetővé teszi, hogy a két kezelés faktor mindegyikének (pl. talaj és műtrágya) és interakciónak a szignifikáns hatását meghatározzuk egymástól függetlenül. Ha a minták normális eloszlású populációkból származnak, akkor a parametrikus kéttényezős varianciaanalízist (kéttényezős ANOVA) alkalmazzuk.

Az egytényezős ANOVA-nál a kísérletben meglevő összes variációt felosztottuk az egyetlen kezelés és a variáció random forrásai között. A kéttényezős ANOVA egy logikus kiterjesztése ennek az eljárásnak; azonban most minden kezelést faktort a variáció elkülönült forrásának tekintjük. Tehát az összes variációt az adatsorozatban, melyet az összes eltérésnégyzet összeg ír le, felosztjuk mindegyik kezelés faktor és a random (vagy reziduális) variáció között. Ezen kívül, amíg ismétlés van a mintákon belül, a kezelések közötti interakció is analizálható, mint a variáció további forrása. Egy kéttényezős faktoriális kísérletben (megfelelő ismétléssel) ezáltal három elkülönült null hipotézist tesztelhetünk:

$H_{0(1)}$ = nincs hatása az A kezelésnek,

$H_{0(2)}$ = nincs hatása a B kezelésnek,

$H_{0(3)}$ = nincs A×B kezelés interakció.

A kéttényezős ANOVA ezután az F-próba három elkülönült értékét adja, vagyis egy F-értéket a két kezelés faktor mindegyikére és egy F-értéket az interakciónak, melyeket felhasználjuk arra, hogy egymástól függetlenül teszteljük a három null hipotézist.

A kéttényezős ANOVA feltételei és korlátai

A kéttényezős ANOVA feltételei ugyanazok, mint az egytényezősé, vagyis:

1. A minták legyenek függetlenek.
2. A minták normális eloszlású populációkból származzanak.
3. A minták azonos varianciájú populációkból származzanak.

A kéttényezős kísérletek esetében el kell különíteni azokat az eseteket, amikor van ismétlés, és azokat, amikor nincs ismétlés a mintákon belül. Nyilvánvalóan, egy egytényezős kísérlet ismétlés nélkül tarthatatlan, mivel a minta átlagok egyetlen megfigyelésen alapulnának és nem lenne mintán belüli variancia. Lehetséges azonban lefolytatni egy kéttényezős kísérletet ismétlés nélkül úgy, hogy az egyik kezelést faktor különböző szintjeinek értékeit használjuk ismétlésként a másik kezeléscsoportra. Amikor azonban nincs ismétlés, akkor interakció nem észlelhető és a variáció források egyszerűen a két kezelést faktor és a reziduális (maradék) variáció. Ennek oka egészen nyilvánvaló. Ismétlés nélkül nem számítható csoporton belüli variancia és ezáltal a reakció egy kezelésre nem hasonlítható össze a második kezelés különböző alkalma-

zási szintjeinél. Tehát kezelés interakció csak akkor válik észlelhető variáció forrásá, amikor a faktorok száma ≥ 3 (Ireland, 2010).

A kéttényezős teljesen randomizált kísérlet varianciaanalízisének (ANOVA) standard számítási eljárása

A számos összegezési eljárás miatt rendkívül fontos, hogy az adatokat a kéttényezős ANOVA-hoz jól áttekinthető formában rendezzük táblázatba. Egy kéttényezős kísérlet tipikus adattáblázatának formáját a teljesen randomizált elrendezéshez és számos szükséges paraméter definícióját az 11.1. táblázat mutatja. Meg kell említeni, hogy van egy nagyon fontos gyakorlati korlát, éspedig az, hogy a mintáknak egyforma méretűnek kell lenni.

Az ANOVA számítása az összes eltérésnégyzet összeg ($SS_{\text{összes}}$) kiszámításával kezdődik, amely méri az összes variációt a kísérletben. Az összes minták közötti eltérésnégyzet összeget ($SS_{\text{minták között}}$), melyet gyakran hívnak „fő hatás” SS-nek, határozzuk meg ezután. Ez az adat tartalmazza az összes variációt a két kezelés faktornak tulajdoníthatóan, plusz a variációt a kezelés interakciónak tulajdoníthatóan. Az eltérésnégyzet összegeket mindegyik kezelés faktornak (SS_A és SS_B) tulajdoníthatóan határozzuk meg ezután egymástól függetlenül. A variáció az interakciónak ($SS_{A \times B}$) tulajdoníthatóan egyszerűen úgy számítható ki, mint a két egyedi kezelés SS összege és a minták közötti SS közötti különbség. Végül, a random variációt a kísérletben, melyet a reziduális (maradék) SS képvisel, a minta varianciák és a megfelelő szabadságfokok alapján számítjuk ki. Ez azonban sokkal könnyebben meghatározható, mint az összes SS és a minták közötti SS közötti különbség.

11.1. táblázat. Kéttényezős, ismétléses faktoriális kísérlet adattáblázatának tipikus formája

B kezelés szintje	A kezelés szintje				B kezelés összeg
	1	2	3	n	
1	X_1	X_1	X_1	X_1	$T_{B,1}$
	X_2	X_2	X_2	X_2	
	X_3	X_3	X_3	X_3	
	X_n	X_n	X_n	X_n	
	$\sum_{1,1}$	$\sum_{1,2}$	$\sum_{1,3}$	$\sum_{1,n}$	
2	X_1	X_1	X_1	X_1	$T_{B,2}$
	X_2	X_2	X_2	X_2	
	X_3	X_3	X_3	X_3	
	X_n	X_n	X_n	X_n	
	$\sum_{2,1}$	$\sum_{2,2}$	$\sum_{2,3}$	$\sum_{2,n}$	
3	X_1	X_1	X_1	X_1	$T_{B,3}$
	X_2	X_2	X_2	X_2	
	X_3	X_3	X_3	X_3	
	X_n	X_n	X_n	X_n	
	$\sum_{3,1}$	$\sum_{3,2}$	$\sum_{3,3}$	$\sum_{3,n}$	
n	X_1	X_1	X_1	X_1	$T_{B,n}$
	X_2	X_2	X_2	X_2	
	X_3	X_3	X_3	X_3	
	X_n	X_n	X_n	X_n	
	$\sum_{n,1}$	$\sum_{n,2}$	$\sum_{n,3}$	$\sum_{n,n}$	
A Kezelés összeg	$T_{A,1}$	$T_{A,2}$	$T_{A,3}$	$T_{A,n}$	$\sum X$ (Főösszeg)

A szabadságfokokat, amelyek kapcsolódnak ezekhez a felosztott SS értékekhez, ugyancsak meg kell határozni. A következő paramétereket definiáljuk:

k = a kezelés minták összes száma,

n = ismétlések (mérések) száma mintánként,

N = a mérések összes száma az egész kísérletben ($= k \times n$),

N_A = az A kezelés szintjeinek száma,

N_B = a B kezelés szintjeinek száma,

n_a = a megfigyelések összes száma az A kezelés minden szintjén ($= n \times N_B$)

n_b = a megfigyelések száma a B kezelés minden szintjén ($= n \times N_A$)

A kéttényezős ANOVA kiszámításának lépései a következők:

- 1. lépés.** A korrekciós tényező (CT) kiszámítása.
- 2. lépés.** Az összes eltérésnégyzet összeg ($SS_{\text{összes}}$) kiszámítása.
- 3. lépés.** A minták közötti eltérésnégyzet összeg ($SS_{\text{minták között}}$) kiszámítása.
- 4. lépés.** Az egyes kezelés eltérésnégyzet összegek ($SS_{\text{kezelés}}, SS_A, SS_B$) kiszámítása.
- 5. lépés.** Az interakció eltérésnégyzet összegek ($SS_{A \times B}$) kiszámítása.
- 6. lépés.** A maradék (reziduális) eltérésnégyzet összeg ($SS_{\text{reziduális}}$) kiszámítása.
- 7. lépés.** A szabadságfokok meghatározása mindegyik variáció forrásra

A szabadságfokokat (DF) mindegyik eltérésnégyzet-összeg adataira úgy számítjuk ki, mint az egytényezős ANOVA-nál, kivéve, hogy a kezelésekek közötti interakció DF-et ugyancsak meg kell határozni. Az interakció DF-et egyszerűen a két kezelés DF szorzataként számítjuk ki.

8. lépés. A közepes négyzetes eltérések (variancia) meghatározása mindegyik variáció forrásra. Mindegyik eltérésnégyzet összeget most átalakítjuk közepes négyzetes eltéréssé (vagy varianciává) azáltal, hogy elosztjuk a hozzájuk tartozó szabadságfokkal.

9. lépés. Az F-érték kiszámítása és az eredmények bemutatása az ANOVA táblázatban

Az F-érték meghatározásának módszere függ az alkalmazott ANOVA modelltől.

1. Fix hatások ANOVA I Modell. A leggyakoribb modellben, amelyben minden kezelés faktor szintje fix és azt a kutató határozza meg (ANOVA I Modell), a nullhipotézis megállapítja, hogy nincs különbség a fix szintek között, amelyeken a kezeléseket alkalmazták, pl. 0, 5, 10 és 15 kg ha⁻¹ műtrágya dózisok hatása között. A nullhipotézis az A kezelésekre, amelyeket fix szinteken alkalmaztak, a következő formában írható fel: $\mu_{A1} = \mu_{A2} = \dots = \mu_{AN}$, ahol N a kezelés szintek száma. Ha null hipotézis lenne igaz, akkor a variancia a kezelésekek tulajdoníthatóan és a random hiba hatásoknak tulajdoníthatóan egyenlő lenne. Ezáltal az F-próbát a variáció minden kezelés forrására és az interakcióra megadja minden kezelés közepes négyzetes eltérésének aránya a reziduális közepes négyzetes eltéréshez. Ez megadja a variáció mérőszámát a kezelésnek tulajdoníthatóan, a random hatásnak tulajdonítható varianciához viszonyítva. Tehát:

F-érték az A kezelésre: $MS_{\text{kezelés A}} / MS_{\text{maradék}}$

F-érték a B kezelésre: $MS_{\text{kezelés B}} / MS_{\text{maradék}}$

F-érték az A × B interakcióra: $MS_{A \times B \text{ interakció}} / MS_{\text{maradék}}$

2. Random hatások ANOVA II Modell. Ahol egy kezelést randomizált szinteken alkalmazunk, a null hipotézis az, hogy nincs kezelés hatás az alkalmazás egyetlen lehetséges szintjén sem, pl. az éves csapadék mennyiségének nincs hatása egy kultúrnövény lisztharmat fertőzöttségére. Statisztikai formában kifejezve, ez egyenlő a null hipotézissel, amely megállapítja, hogy a random kezelés faktor nem vált ki extra varianciát a kísérletben. Azokban a viszonylag nem gyakori esetekben, ahol a kéttényezős kísérletben mindkét kezelés faktort random szinteken alkalmazzuk (azaz ANOVA II Modell), a null hipotézis szerint a variancia a kezeléseknél és az interakciónak tulajdoníthatóan egyenlő (H_0):

Az F-próbát a randomizált kezelésre ezáltal úgy határozzuk meg, mint a kezelés MS arányát az interakció MS-hez, míg az interakció F-érték megmarad az interakció MS-nek a reziduális (maradék) MS-hez való arányaként. Tehát:

$$F\text{-érték (minden kezelés)} = MS_{\text{kezelés}} / MS_{\text{interakció}}$$

$$F\text{-érték (interakció)} = MS_{\text{interakció}} / MS_{\text{maradék}}$$

Vegyes hatások ANOVA III Modell. Amikor egyik faktort fix szinteken, a második faktort random szinteken alkalmazzuk, akkor az ANOVA III Modellt használjuk. Ebben az esetben a null hipotézis kimondja, hogy nincs kezelés-hatás és a fix hatások varianciája egyenlő az interakciónak tulajdonítható varianciával, míg a randomizált kezelés varianciája egyenlő lesz a reziduális varianciával. Az F-arányok meghatározása a következő:

$$F\text{-érték (fix kezelés szintek)} = MS_{\text{kezelés}} / MS_{\text{interakció}}$$

$$F\text{-érték (random kezelés szintek)} = MS_{\text{kezelés}} / MS_{\text{maradék}}$$

$$F\text{-érték (interakció)} = MS_{\text{interakció}} / MS_{\text{maradék}}$$

A kéttényezős ANOVA kiszámításának végső szakaszait a kéttényezős ANOVA táblázat mutatja (11.2. táblázat).

11.2. táblázat. A kéttényezős ANOVA táblázat standard formája (Ireland, 2010)

Variáció forrása	Eltérés négyzet-összeg (SS)	Szabadság-fok (DF)	Közepes négyzetes eltérés (MS)	F-érték		
				Fix hatások I. modell ANOVA	Random hatások II. modell ANOVA	Vegyes hatások III. modell ANOVA ^a
Kezelés faktor A	$SS_{\text{kezelés A}}$	$DF_{\text{faktor A}}$	$\frac{SS_{\text{A kezelés}}}{DF_{\text{A tényező}}}$	$\frac{MS_{\text{A kezelés}}}{MS_{\text{maradék}}}$	$\frac{MS_{\text{A kezelés}}}{MS_{\text{A} \times \text{B}}}$	$\frac{MS_{\text{A kezelés}}}{MS_{\text{A} \times \text{B}}}$
Kezelés faktor B	$SS_{\text{kezelés B}}$	$DF_{\text{faktor B}}$	$\frac{SS_{\text{B kezelés}}}{DF_{\text{B tényező}}}$	$\frac{MS_{\text{B kezelés}}}{MS_{\text{maradék}}}$	$\frac{MS_{\text{B kezelés}}}{MS_{\text{A} \times \text{B}}}$	$\frac{MS_{\text{B kezelés}}}{MS_{\text{maradék}}}$
A×B interakció	$SS_{\text{A} \times \text{B}}$	$DF_{\text{A} \times \text{B}}$	$\frac{SS_{\text{A} \times \text{B}}}{DF_{\text{A} \times \text{B}}}$	$\frac{MS_{\text{A} \times \text{B}}}{MS_{\text{maradék}}}$	$\frac{MS_{\text{A} \times \text{B}}}{MS_{\text{maradék}}}$	$\frac{MS_{\text{A} \times \text{B}}}{MS_{\text{maradék}}}$
Maradék (mintán belül)	$SS_{\text{maradék}}$	$DF_{\text{maradék}}$	$\frac{SS_{\text{maradék}}}{DF_{\text{maradék}}}$			
Összes	$SS_{\text{összes}}$	$DF_{\text{összes}}$				

^aVegyes hatások modell, feltételezve, hogy az A faktor szintjei fixek és a B faktor szintjei random hatásúak.

10. lépés. A szignifikancia valószínűségének meghatározása

Abból a célból, hogy meghatározzuk a szignifikancia valószínűségét, a számított F-értéket mind a kezelésekre és mind az interakcióra, összehasonlítjuk sorjában az egyoldali kritikus F-értékkal (F-érték táblázat). Az F-érték táblázatában számlálóként (V_1) és nevezőként (V_2) egyaránt felhasználjuk a közepes négyzetes eltérések szabadságfokát, amikor meghatározzuk az F-értéket. Ezáltal a fix hatás (I Modell) ANOVA esetében V_1 lesz a kezelés vagy az interakció DF és V_2 lesz a reziduális DF. A random vagy a vegyes hatású ANOVA (II és III Modell) esetében V_2 lesz az interakció DF, a kezelések hatásának tesztelésekor. Ha a megfigyelt F-érték egy sajátos kezelésre vagy az interakcióra egyenlő vagy nagyobb mint a kritikus F-érték, akkor a null hipotézist visszautasítjuk és a kezelésnek vagy az interakciónak a szignifikáns hatását igazolhatjuk a valószínűségnek azon a szintjén, amelyen a statisztikai próbát végeztük.

11.1 Kéttényezős teljesen véletlen elrendezés (CRD)

11.1.1. Kísérlet leírása

Az MTA Mezőgazdasági Kutatóintézetében, Martonvásáron, fitotroni kísérletben vizsgáltuk a kukorica beltenyésztett vonalak acetoklór herbiciddel szembeni érzékenységét és az AD-67 jelű antidótum protektív hatását (Berzsenyi *et al.*, 1994, 1997) (11.1. kép). A kísérletet PGV 3 típusú növénynevelő kamrában állítottuk be 18/13 °C nappali/éjszakai hőmérsékleten és 80%-os relatív páratartalomnál. Az acetoklór herbicid kijuttatása preemergensen (vetés után, kelés előtt) történt speciális permetező fülkében. A kéttényezős kísérlet egyik tényezője a kukorica vonal, a másik tényezője a herbicid kezelés volt. A kísérletben hat beltenyésztett vonal érzékenységét vizsgáltuk. Az acetoklór herbicid dózisa 1, 2 és 4 l ha⁻¹ volt önmagában, illetve AD-67 antidótummal kombinálva, ezáltal a kezeletlen kontrollal együtt hét herbicid kezelés volt. A kísérletet 3 ismétlésben teljesen véletlen elrendezésben (CRD) állítottuk be, és hetente kétszer újra randomizáltuk. A kísérlet időtartama 21 nap volt. A kísérlet lebontásakor növényenként mértük a hajtás hosszát és tömegét, az adatokból tenyészedény átlagokat számoltunk. A varianciaanalízishez a hajtáshossz adatokat minden beltenyésztett vonal kezeletlen kontrolljának %-ában fejeztük ki (11.3. táblázat).



11.1. kép. Két kukorica beltenyésztett vonal (0643, Szuvenir) herbicid érzékenységének vizsgálata fitotron kísérletben, Martonvásáron. Felső tenyészedény sor az antidótum protektív hatását, az alsó tenyészedény sor a herbicid fitotoxikus hatását mutatja antidótum nélkül, eltérő dózisoknál

11.1.2. Varianciaanalízis lépései

1. lépés. Vázzoljuk fel a variancia táblázatot. Jelöljük az A faktor (beltenyésztett vonal) szintjeit a-val és a B faktor (herbicid) szintjeit b-vel.

Variáció forrása	Szabadságfok	SS	MS	Számított F-érték	Táblázati F-érték	
					5%	1%
Kezelés	$ab - 1 = 35$					
A faktor	$a - 1 = 5$					
B faktor	$b - 1 = 5$					
A × B	$(a - 1)(b - 1) = 25$					
Hiba	$ab(r - 1) = 72$					
Összes	$rab - 1 = 107$					

1.1.3. táblázat. Az acetoklór herbicid hatása hat beltenyészett kukoricavonal hajtáshosszára, három ismétlésben, teljesen véletlen elrendezésű kéttényezőes fitotron kísérletben

Herbicid		Hajtáshossz a kontroll %-ában			Összes (T)
Hatóanyaga	Dózis l ha ⁻¹	Ismétlés			
		I.	II.	III.	
<i>W 401</i>					
1. acetoklór	1	69.27	66.27	77.5	213.04
2. acetoklór	2	73.59	57.31	67.17	198.07
3. acetoklór	4	51.81	57.79	56.99	166.59
4. acetoklór + AD-67	1	97.59	87.34	81.84	266.77
5. acetoklór + AD-67	2	89.9	90.21	87.45	267.56
6. acetoklór + AD-67	4	77.65	76.36	79.09	233.1
<i>EPI-21</i>					
1. acetoklór	1	70.0	58.53	63.72	192.25
2. acetoklór	2	46.28	33.08	32.95	112.31
3. acetoklór	4	26.13	25.76	32.09	83.98
4. acetoklór + AD-67	1	85.69	85.78	75.11	246.58
5. acetoklór + AD-67	2	76.79	67.03	58.86	202.68
6. acetoklór + AD-67	4	68.03	59.29	36.66	163.98
<i>761</i>					
1. acetoklór	1	53.29	74.22	82.17	209.68
2. acetoklór	2	46.44	47.41	58.76	152.61
3. acetoklór	4	46.24	38.03	40.58	124.85
4. acetoklór + AD-67	1	81.9	84.04	85.26	251.2
5. acetoklór + AD-67	2	70.99	72.1	85.61	228.7
6. acetoklór + AD-67	4	71.66	70.26	58.83	200.75
<i>To 337</i>					
1. acetoklór	1	42.96	70.0	49.66	162.62
2. acetoklór	2	29.62	66.59	41.67	137.88
3. acetoklór	4	32.37	23.81	30.54	86.72
4. acetoklór + AD-67	1	86.88	95.41	90.14	272.43
5. acetoklór + AD-67	2	78.08	82.52	88.71	249.31
6. acetoklór + AD-67	4	61.14	66.11	53.33	180.58
<i>HMv 565-3</i>					
1. acetoklór	1	76.61	76.0	85.81	238.42
2. acetoklór	2	77.85	77.82	81.73	237.4
3. acetoklór	4	73.79	71.02	76.67	221.48
4. acetoklór + AD-67	1	89.52	85.34	91.53	266.39
5. acetoklór + AD-67	2	82.58	85.85	85.14	253.57
6. acetoklór + AD-67	4	77.4	81.78	77.16	236.34
<i>HMv 124-2</i>					
1. acetoklór	1	67.07	55.18	40.94	163.19
2. acetoklór	2	45.27	40.52	36.96	122.75
3. acetoklór	4	24.32	32.13	17.14	73.59
4. acetoklór + AD-67	1	95.26	90.59	78.8	264.65
5. acetoklór + AD-67	2	79.09	75.92	74.57	229.58
6. acetoklór + AD-67	4	68.19	54.09	55.09	177.37
Főösszeg					7089

2. lépés. Számítsuk ki a kezelés összegeket (T) és a főösszeget (G), mint a 11.3. táblázat mutatja, és számítsuk ki az összes SS-t és a hiba SS-t, követve az egytényezős RCB kísérletnél leírt eljárást:

$$C.F. = \frac{G^2}{rab} = \frac{(7\ 089)^2}{(3)(6)(6)} = 465\ 314$$

$$\text{Összes SS} = \sum X^2 - C.F. = \left[(69.27)^2 + (66.27)^2 + \dots + (55.09)^2 \right] - 465\ 314 = 42\ 399$$

$$\text{Kezelés SS} = \frac{\sum T^2}{r} - C.F. = \frac{(213.04)^2 + \dots + (177.37)^2}{3} - 465\ 314 = 38\ 039.3$$

$$\text{Hiba SS} = \text{Összes SS} - \text{Kezelés SS} = 42\ 399 - 38\ 039.3 = 4\ 359.7$$

Az előzetes varianciaanalízist, a különböző számított SS értékekkel a 11.4. táblázat tartalmazza

11.4. táblázat. A herbicid kísérlet adatainak előzetes varianciaanalízise

Variáció forrása	Szabadságfok	SS	MS	Számított F-érték	Táblázati F-érték	
					5%	1%
Kezelés	35	38039.3	1086.8	17.95**	1.62	1.98
Hiba	72	4359.7	60.55			
Összes	107	42399				

3. lépés. Készítsük el a faktor A × faktor B összegek táblázatát, a számított faktor A és faktor B összegekkel. Példánkban, a beltenyészített vonal × herbicid (AB) összege, a beltenyészített vonal összeggel (A) és a herbicid összeggel (B) együtt a 11.5. táblázatban található.

11.5. táblázat. A beltenyészített vonal × herbicid adatok összesítő táblázata

Herbicid kezelés	Hajtáshossz % összeg (AB)						Herbicid összeg (B)
	L ₁	L ₂	L ₃	L ₄	L ₅	L ₆	
H ₁	213.04	192.25	209.68	162.62	238.42	163.19	1179.2
H ₂	198.07	112.31	152.61	137.88	237.4	122.75	961.02
H ₃	166.59	83.98	124.85	86.72	221.48	73.59	757.21
H ₄	266.77	246.58	251.2	272.43	266.39	264.65	1568.02
H ₅	267.56	202.68	228.7	249.31	253.57	229.58	1431.4
H ₆	233.1	163.98	200.75	180.58	236.34	177.37	1192.12
Beltenyészített vonal összeg (A)	1345.13	1001.78	1167.79	1089.54	1453.6	1031.13	7088.97

L₁–L₆: beltenyészített vonalak, H₁–H₆: herbicid kezelések jelölése.

4. lépés. Számítsuk ki a kezelés SS három komponensét:

$$A \text{ SS} = \frac{\sum A^2}{rb} - \text{C.F.} = \frac{(1345.13)^2 + \dots + (1031.13)^2}{(3)(6)} - 465314 = 9127.7$$

$$B \text{ SS} = \frac{\sum B^2}{ra} - \text{C.F.} = \frac{(1179.2)^2 + \dots + (1192.12)^2}{(3)(6)} - 465314 = 24473.9$$

$$A \times B = \text{Kezelés SS} - A \text{ SS} - B \text{ SS} = 38039.3 - 9127.7 - 24473.9 = 4437.7$$

5. lépés. Számítsuk ki az MS értékeket mindegyik variáció forrásra úgy, hogy az SS értékét elosztjuk a megfelelő d.f.-fel:

$$A \text{ MS} = \frac{A \text{ SS}}{a-1} = \frac{9127.7}{5} = 1825.5$$

$$B \text{ MS} = \frac{B \text{ SS}}{b-1} = \frac{24473.9}{5} = 4894.8$$

$$A \times B \text{ MS} = \frac{A \times B \text{ SS}}{(a-1)(b-1)} = \frac{4437.7}{25} = 177.5$$

$$\text{Hiba MS} = \frac{\text{Hiba SS}}{ab(r-1)} = \frac{4359.7}{72} = 60.55$$

6. lépés. Számítsuk ki az F-értéket a három komponens mindegyikére:

$$F(A) = \frac{A \text{ MS}}{\text{Hiba MS}} = \frac{1825.5}{60.55} = 30.15$$

$$F(B) = \frac{B \text{ MS}}{\text{Hiba MS}} = \frac{4894.8}{60.55} = 80.84$$

$$F(A \times B) = \frac{A \times B \text{ MS}}{\text{Hiba MS}} = \frac{177.5}{60.55} = 2.93$$

7. lépés. Hasonlítsuk össze mindegyik számított F-értéket a táblázati F-értékkal. A táblázati F-érték a beltenyészett vonalra és a herbicidre egyaránt, $f_1 = 5$ és $f_2 = 70$ szabadságfoknál 2.35 az 5%-os szignifikancia szinten és 3.29 az 1%-os szignifikancia szinten. Az eredmény azt mutatja, hogy a beltenyészett vonal és a herbicid fő hatása és a köztük levő interakció egyaránt 5%-os szinten szignifikáns.

8. lépés. Számítsuk ki a variációs koefficiens:

$$CV\% = \frac{\sqrt{\text{Hiba MS}}}{\text{foatlag}} = \frac{\sqrt{60.55}}{84.2} = 9.24\%$$

11.6. táblázat. Az 5 × 5-ös faktoriális kísérlet adatainak variancia táblázata

Variáció forrása	Szabadságfok	SS	MS	Számított F-érték	Táblázati F-érték	
					5%	1%
Kezelés	35	38039.3	1086.8	17.95**	1.62	1.98
Belt. vonal (A)	5	9127.7	1825.5	30.15**	2.35	3.29
Herbicidek (B)	5	24473.9	4894.8	80.84**	2.35	3.29
A × B	25	4437.7	177.5	2.93**	1.67	2.07
Hiba	72	4359.7	60.55			
Összes	107	42399				

* szignifikáns 5%-os szinten, ** szignifikáns 1%-os szinten

9. lépés. A 4–8. lépésben kiszámított értékeket írjuk be a variancia táblázatba (11.6. táblázat). Az eredményekből látható, hogy szignifikáns interakció van a beltenyészett vonal (genotípus) és a herbicidek között, jelezve, hogy a genotípusos különbséget szignifikánsan befolyásolja a herbicidek kezelése és hogy a herbicidek hatása szignifikánsan különbözik a vizsgált beltenyészett vonalaknál. A genotípus és a herbicidek fő hatása egyaránt szignifikáns. Ugyanakkor az MS értékek alapján megállapítható, hogy a genotípus MS 10-szerese, a herbicidek kezelése MS 26-szorosa a kölcsönhatás MS-nek.

Szignifikáns differencia kiszámítása:

(a) Bármely két kombináció között:

$$SzD_{5\%} = t_{5\%} \sqrt{\frac{2 \text{Hiba MS}}{r}} = 2.0 \sqrt{\frac{2(60.55)}{3}} = 12.71$$

(b) Az A tényező változatai között a B tényező változatainak átlagában, ha a kölcsönhatás nem szignifikáns:

$$SzD_{5\%} = t_{5\%} \sqrt{\frac{2 \text{Hiba MS}}{rb}} = 2.00 \sqrt{\frac{2(60.55)}{18}} = 5.19$$

(c) A B tényező változatai között az A tényező változatainak átlagában, ha a kölcsönhatás nem szignifikáns:

$$SzD_{5\%} = t_{5\%} \sqrt{\frac{2 \text{Hiba MS}}{ra}} = 2.00 \sqrt{\frac{2(60.55)}{18}} = 5.19$$

Ha a kölcsönhatás legalább $P = 5\%$ -os szinten szignifikáns, akkor a tényezők változatai közötti különbséget általában nem a másik tényező változatainak átlagában, hanem kombinációnként vizsgáljuk.

Az eredmények értelmezése

Fitotroni kísérletben vizsgáltuk 6 kukorica beltenyészett vonal acetoklór herbicidekkel szembeni érzékenységét és az AD-67 antidótum protektív hatását az acetoklór 1, 2

és 4 l ha^{-1} dózisában. Szignifikáns különbséget lehetett megállapítani a beltenyészett vonalak között az acetoklór érzékenységükben és az antidótum protektív hatásában (11.2. ábra). Az EPI-21, a HMv 124-2 és a To 337 beltenyészett vonalak szignifikánsan érzékenyebbek voltak az acetoklórra. Hajtásmagasságuk a kontroll %-ában 40 és 43% között változott, a herbicid dózisok átlagában. A HMv 565-3, a W 401 és a 761 jelölésű beltenyészett vonalak szignifikánsan toleránsabbak voltak az acetoklór herbiciddel szemben. Hajtásmagasságuk a kontroll %-ában 54 és 78% között változott, amikor a herbicid dózisok átlagában fejeztük ki. Az AD-67 antidótum protektív hatása megfelelő volt, a hajtásmagasság a kontroll 66 és 87%-a között változott a herbicid dózisok és beltenyészett vonalak átlagában (Berzsenyi et al., 1994, 1997). A szignifikáns interakció MS értéke jelentősen kisebb volt, összehasonlítva a fő hatásokéval, amely lehetővé teszi a főhatások korlátozott értelmezését. A beltenyészett vonalak reakcióinak vizsgálata a kontrasztok módszerével kimutatta, hogy szignifikáns különbség van az 1. és 5. vonal vs. többi vonal, az 1 vs. 5. vonal, valamint a 3. vonal vs. 2., 4. és 6. vonal reakciója között (11.3. ábra). A Duncan teszt alapján – a fajták átlagában – az 1. és 6. herbicid kezelés szignifikánsan nem különbözik, azonban a többi herbicid kezelés között szignifikáns különbség van (11.4. ábra).

11.1.3. Varianciaanalízis GenStat programmal

Analysis of variance

Variate: Hajtásmagasság

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Fajta	5	9130.96	1826.19	30.16	<.001
Herbicid	5	24476.64	4895.33	80.86	<.001
Fajta.Herbicid	25	4433.31	177.33	2.93	<.001
Residual	72	4359.03	60.54		
Total	107	42399.95			

Tables of means

Variate: Hajtásmagasság

Grand mean 65.64

Fajta	1	2	3	4	5	6	
	74.73	55.65	64.88	60.53	80.76	57.29	
Herbicid	1	2	3	4	5	6	
	65.51	53.39	42.07	87.11	79.52	66.23	
Fajta	Herbicid	1	2	3	4	5	6
1		71.01	66.02	55.53	88.92	89.19	77.70
2		64.08	37.44	27.99	82.20	67.56	54.66
3		69.89	50.87	41.62	83.73	76.23	66.92
4		54.21	45.96	28.91	90.81	83.10	60.20
5		79.47	79.13	73.83	88.80	84.52	78.78
6		54.40	40.93	24.53	88.21	76.53	59.12

Standard errors of differences of means

Table	Fajta	Herbicide	Fajta Herbicide
rep.	18	18	3
d.f.	72	72	72
s.e.d.	2.594	2.594	6.353

Least significant differences of means (5% level)

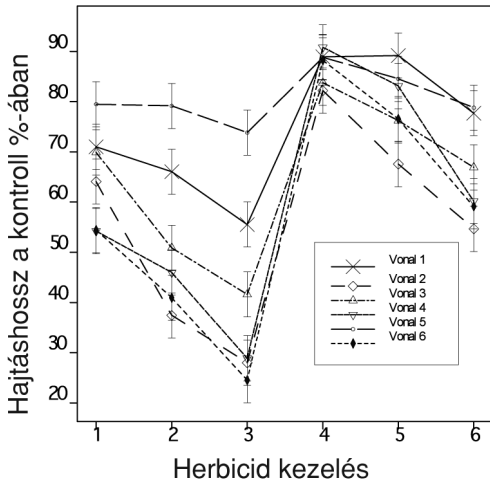
Table	Fajta	Herbicide	Fajta Herbicide
rep.	18	18	3
d.f.	72	72	72
l.s.d.	5.170	5.170	12.665

Duncan's multiple range test**Fajta**

	Mean	
2	55.65	a
6	57.29	a
4	60.53	ab
3	64.88	b
1	74.73	c
5	80.76	d

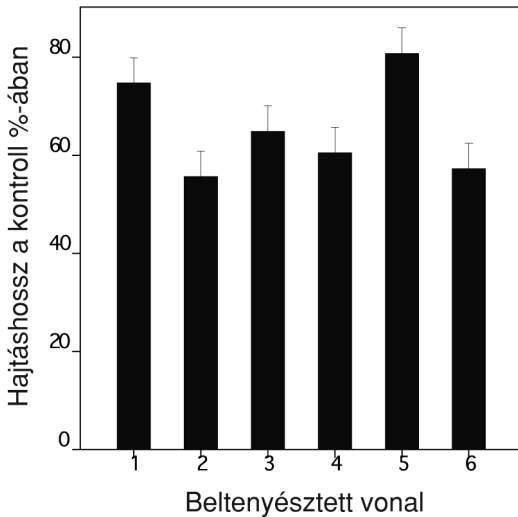
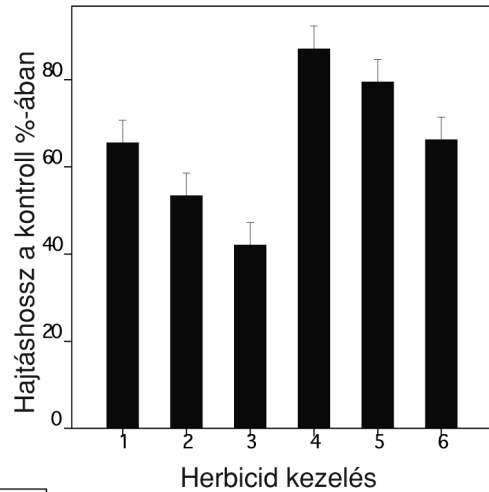
Herbicide

	Mean	
3	42.07	a
2	53.39	b
1	65.51	c
6	66.23	c
5	79.52	d
4	87.11	e



11.2. ábra. A vizsgált 6 kukorica beltenyésztett vonal hajtáshosszának reakciója a különböző herbicid kezelésekre (1–3: növekvő herbicid dózis antidótum nélkül, 4–6: antidótummal)

11.3. ábra. A herbicid kezelések hatása – a beltenyésztett vonalak átlagában – a kukoricánövény hajtáshosszának alakulására



11.4. ábra. A vizsgált 6 beltenyésztett vonal hajtáshosszának alakulása (a herbicid kezelések átlagában)

11.2. Fő hatások és interakciók

Egy faktoriális kísérlet varianciaanalízisének, amely két faktort tartalmaz, dönté-
nünk kell arról, hogyan bontjuk fel a kezelés SS-t. Abból a célból, hogy megválaszol-
junk számos kérdést, amelyet a faktoriális struktúra tartalmaz, az SS felbontásának
egyik nyilvánvaló módszere az, hogy vesszük az első komponens SS-t, hogy megfe-
leljen az egyik faktor szintjei közötti variációnak; a többi komponens ezután reprezen-
tálni fogja a második faktor szintjei közötti variációt az első faktor minden szintjére.
Egy egyszerű példaként vegyünk számba egy kísérletet három takarmányféleséggel
(0,1 és 2), összehasonlítva hímivarú és nőivarú (M, F) sertésekkel. Ha a kezelés össze-
gek T_{0M} , T_{0F} , T_{1M} , T_{1F} , T_{2M} , T_{2F} és a kísérlet r blokkot tartalmaz, mindegyik hat egység-
gel, akkor a teljes kezelés SS a következő:

$$\frac{T_{0M}^2 + T_{0F}^2 + T_{1M}^2 + T_{1F}^2 + T_{2M}^2 + T_{2F}^2}{r} - c.f.$$

Ez felosztható ortogonális komponens SS-re két módon.

(a) Egy komponens reprezentálja az összes különbséget a hím- és nőivarú serté-
sek között:

$$(i) \frac{T_M^2 + T_F^2}{3r} - c.f.$$

ahol $T_M = T_{0M} + T_{1M} + T_{2M}$, és a két komponens reprezentálja a variációt a takarmá-
nyok között, mindegyik nemre,

$$(ii) \frac{T_{0M}^2 + T_{1M}^2 + T_{2M}^2}{r} - \frac{T_M^2}{3r}$$

$$\text{és (iii) } \frac{T_{0F}^2 + T_{1F}^2 + T_{2F}^2}{r} - \frac{T_F^2}{3r}$$

(b) Egy komponens reprezentálja a variációt a három takarmány között:

$$(i) \frac{T_0^2 + T_1^2 + T_2^2}{2r} - c.f.$$

és három komponens reprezentálja a különbséget a hím- és nőivarú sertések között
mindegyik takarmányra:

$$(ii) \frac{T_{0M}^2 + T_{0F}^2}{r} - \frac{T_0^2}{2r}$$

$$(iii) \frac{T_{1M}^2 + T_{1F}^2}{r} - \frac{T_1^2}{2r}$$

$$(iv) \frac{T_{2M}^2 + T_{2F}^2}{r} - \frac{T_2^2}{2r}$$

E felosztások bármelyike megfelelő lehet egy sajátos kísérletre. Egyik sem használja
azonban a faktoriális struktúra szimmetriájának előnyét. Mindkét faktor figyelembe
veszi, először egy faktor szintjei közötti nagy különbségeket, és másodsor, e faktor
minden szintjére a második faktor szintjei közötti különbségeket.

Egy faktoriális kezelés SS felbontásának standard módszere minden faktort egyformán kezel. A tipikus felosztás a takarmány \times nem kísérletre:

$$(i) \frac{T_M^2 + T_F^2}{3r} - c.f.$$

képviseli az általános különbséget a nemek között;

$$(ii) \frac{T_0^2 + T_1^2 + T_2^2}{2r} - c.f.$$

képviseli az általános különbséget a takarmányok között; és a megmaradó SS

$$\frac{T_{0M}^2 + T_{0F}^2 + T_{1M}^2 + T_{1F}^2 + T_{2M}^2 + T_{2F}^2}{r} - c.f. - (i) - (ii)$$

Mivel (i) és (ii) kifejezések képviselnek minden információt a két nem közötti átlagos különbségről, és a három takarmány közötti átlagos különbségről, az utolsó SS-nek meg kell felelni a két nem reakciója közötti különbségnek a különböző takarmányokra. A kezelés SS ily módon történő felosztásának két előnye van, éspedig, felhasználja a faktoriális struktúra szimmetriáját és egyszerűen általánosítható a kettőnél több faktoros kísérletekre.

A három SS komponens:

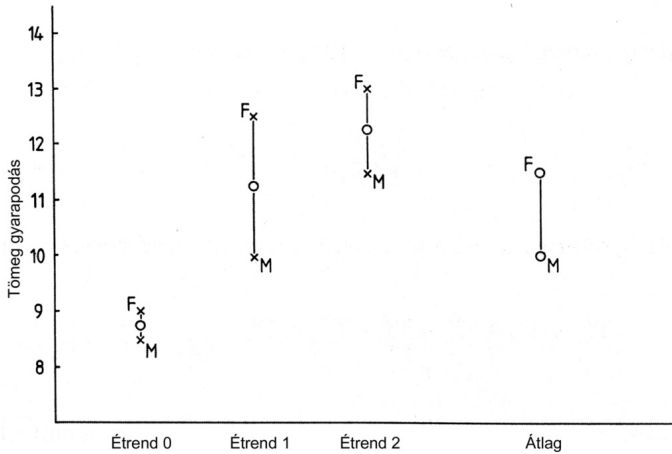
- (i) SS a nemek fő hatására,
- (ii) SS a takarmányok fő hatására és
- (iii) SS a nemek és a takarmányok közötti interakcióra.

Formálisan, meghatározhatjuk egy faktor fő hatását, mint az átlagtermések közötti különbséget e faktor különböző szintjeire, a többi faktor összes szintjének átlagában. Meghatározhatjuk az interakciót két faktor között, mint az átlagtermések közötti különbség variációját az egyik faktor különböző szintjeire a másik faktor különböző szintjein keresztül. E definíciók gyakorlati interpretációja legkönnyebben megmagyarázható diagram formájában.

Tételezzük fel, hogy a sertések heti hipotetikus átlagos tömeggyarapodása a ivar (nem) \times takarmány kísérletben:

Ivar csoport	0 jelű takarmány	1-es jelű takarmány	2-es jelű takarmány
Hímivarú	8.5 lb	10.0 lb	11.5 lb
Nőivarú	9.0 lb	12.5 lb	13.0 lb

Ezek a hozamok reprezentálhatók grafikusán, mint a 11.5. ábra mutatja. Átlagos hozamok a takarmányokra (a nemek átlagában): 8.75, 11.25 és 12.25 a 0, 1 és 2 jelű takarmányra, és a körök mutatják a három takarmányféleség oszlopában. A takarmányok összehasonlításának fő hatása e három érték közötti különbség. Hasonlóan, az átlagos hozamok a nemekre (a takarmányok átlagában) 10.0 és 11.5 az M-re és az F-re, melyet a körök mutatnak az „átlag” oszlopban a 11.5 ábrán. Az F-M különbségek a takarmányokra 0.5, 2.5 és 1.5, és az interakció a takarmányok és a nemek között tartalmazza e három érték közötti variációt. Ábrázolva, a hozam variációt reprezentálja az F magasságának variációja az M felett a három takarmány átlagában.



11.5. ábra. Hat kezelés átlagtermésének grafikus ábrázolása, amely mutatja a főhatásokat és az interakciókat (Mead et al., 2003)

Fontos realizálni, hogy az interakció két faktor között szimmetrikus. Megvitattuk a takarmányok és a nemek közötti interakciót kifejezve a hím- és nőivarú sertések közötti különbség változásának útján, a három takarmány változásának figyelembevételével. Éppen olyan jól megvitathatnánk, hogyan különböznek a három takarmány közötti eltérések a hímivarú sertésekre vonatkozóan, a megfelelő különbségekből a nőivarú sertésekre vonatkozóan. Például, a hímivarú sertésekre az átlagos heti súlygyarapodás az 1-es jelű takarmányra a 0-jelű és a 2-es jelű takarmányra kapott súlygyarapodás között volt, míg a nőivarú sertésekre a gyarapodás az 1-es jelű takarmányra majdnem ugyanaz volt, mint a 2-es jelű takarmányra, és mindkettőre sokkal nagyobb volt, mint a gyarapodás a 0 jelű takarmányra. Még jelentősebben, az 1-es jelű takarmány 1.5 lb-nél nagyobb gyarapodást eredményezett, összehasonlítva a 0 jelű takarmánnyal a hímivarú sertésekre, azonban a nőivarú sertésekre az 1-es jelű takarmányra a gyarapodás 3.5 lb-vel nagyobb volt, mint a 0 jelű takarmányra. A különbségeknek ez az összehasonlítása természetesen pontosan ugyanazt az információt adja, mint az F-M különbség összehasonlítása, 0.5 a 0 jelű takarmányra és 2.5 az 1-es jelű takarmányra.

11.2.1. Kéttényezős faktoriális kísérlet varianciaanalízise

A kéttényezős faktoriális kísérlet varianciaanalízise a következő. Tételezzük fel, hogy r blokkunk van, t kezeléssel. A kezelések A faktor a szintjének és B faktor b szintjének összes kombinációja ($ab = t$). Legyen az A faktor 1. szintjének és a B faktor 1. szintjének összege r ismétlésben T_{11} , az A faktor 1. szintjének és a B faktor 2. szintjének összege T_{12} , és így tovább, az A faktor a szintjének és a B faktor b szintjének összege T_{ab} . Legyen T_{A1} az A faktor 1. szintjének $r \times b$ összege (összegezve a B kezelés minden szintjén, valamint az ismétlésekben) és T_{B1} a B faktor 1. szintjének az $r \times a$ összege.

A varianciaanalízis első részében a kísérleti kezelések struktúráját nem vesszük figyelembe és az analízist a randomizált blokk elrendezésnek megfelelően végezzük. Számítógépes programhoz a modell: Blokk + Kezelés

Variáció forrása	SS	d.f.	MS
Blokk	$\frac{B_1^2 + B_2^2 + \dots + B_r^2}{t} - c.f.$	r-1	
Kezelés	$\frac{T_{11}^2 + T_{12}^2 + T_{13}^2 + \dots + T_{ab}^2}{r} - c.f.$	t-1	
Hiba	Kivonással	(t-1)(r-1)	
Összes	$\sum (y^2) - c.f.$	tr-1	
c.f.	$[\sum (y)]^2 / rt$		

A kezelés SS felbontása:

Variáció forrása	SS	d.f.	MS
A tényező fő hatása	$\frac{T_{A1}^2 + T_{A2}^2 + \dots}{rb} - c.f.$	a-1	
B tényező fő hatása	$\frac{T_{B1}^2 + T_{B2}^2 + \dots}{ra} - c.f.$	b-1	
AB interakció	kivonással	(a-1)(b-1)	
Összes kezelés variáció	$\frac{T_{11}^2 + T_{12}^2 + \dots + T_{ab}^2}{r} - c.f.$	ab-1	

A kezelés SS három komponensének mindegyik MS értéke összehasonlítható a hiba MS-sel, és az így kapott F-értéket összehasonlítjuk a táblázati F-értékkel. Például, az A tényező fő hatásának teszteléséhez az $F = [(A) MS / \text{hiba MS}]$ és összehasonlítjuk az F-eloszlással (a-1) és (t-1)(r-1) szabadságfoknál.

Gyakran egyesítjük az analízis két szakaszát, helyettesítve a kezelés SS-t az első táblázatban a második táblázat három komponensével, amely a varianciaanalízis következő formáját adja:

Variáció forrása	SS	d.f.	MS	F-érték
Blokk		(r-1)		
A faktor fő hatása		(a-1)		
B faktor fő hatása		(b-1)		
AxB interakció		(a-1)(b-1)		
Hiba		(ab-1)(r-1)	s^2	
Összes		(abr-1)		

Számítógépes analízisnél meg kell határoznunk a három faktort – Blokk, faktor A, faktor B – az adatokkal, a következő formában: Termés Blokk Faktor A Faktor B
Az illesztett modell:

$$\text{Blokk} + A + B + A \times B$$

ahol $A \times B$ jelöli az A és B faktor interakcióját.

11.7. táblázat. A 2 × 2-es hipotetikus faktoriális kísérlet eredményei és prezentálásuk formája (Mead et al., 2003)

Kezelésátlagok		F-értékek		Prezentálás formája		
(1)						
A ₁	A ₂	Fő hatás (A)	20.0**	A ₁	19	
B ₁	16 20	Fő hatás (B)	38.2***	A ₂	24	
B ₂	22 28	Interakció	0.8	B ₁	18	
				B ₂	25	
SE az A ₁ összehasonlítására A ₂ -vel vagy B ₁ összehasonlítására B ₂ -vel = 1.12						
(2)						
A ₁	A ₂	Fő hatás (A)	20.0**	A ₁	A ₂	Átlag
B ₁	16 18	Fő hatás (B)	39.2***	B ₁	16 18	17
B ₂	20 28	Interakció	7.2*	B ₂	20 28	24
				Átlag	18 23	
SE a négy kombináció átlag összehasonlítására = 1.58 SE a fő hatások összehasonlítására = 1.12						
(3)						
A ₁	A ₂	Fő hatás (A)	20**	A ₁	A ₂	
B ₁	13 24	Fő hatás (B)	12.8**	B ₁	13 24	
B ₂	23 22	Interakció	28.8***	B ₂	23 22	
SE bármely két átlag összehasonlítására = 1.58						

Amennyiben a lényeges hatások a faktoriális kísérletben csak a fő hatások, akkor az eredmények korlátozhatók sorban mindegyik faktor különböző szintjeinek átlagtermésére, a másik faktor szintjeinek átlagában. Más szavakkal, ha az interakciók a faktorok között nem fontosak, akkor a termésátlagok a különböző faktorok szintjeinek kombinációira nem adnak több információt, mint a termésátlagok mindegyik faktor minden szintjére külön-külön. Másrészt, ha az F-érték az A × B interakcióra szignifikáns, akkor az eredmények közlésekor a termésátlagokat az A és B szintjeinek minden kombinációjára meg kell adni. Amikor egy interakció F-értéke szignifikáns, számszerű értéke sokkal kisebb lehet, mint a megfelelő fő hatás F-értéke, vagy lehet hasonló értékű. Az előző esetben, amely gyakoribb, az értelmezése lehet az, hogy jelentős különbség van, mondjuk az A faktor szintjeinek hatása között, azonban erre a különbségre hatással van a B faktor szintje. Következésképpen, a prezentált eredményeknek magukban kell foglalni az átlagtermést mindegyik faktor (A és B) minden szintjére, továbbá a kezelések kétoldalú táblázatát az A és B szintjeinek összes kombinációjára. Ha azonban az A × B

interakció F-értéke hasonló vagy nagyobb, mint az A és B fő hatások F-értéke, akkor a B hatás variációja az A szintjeinek átlagában legalább olyan nagy, mint az átlagos hatások és mindegyik faktor fő hatása egyértelműen nem értelmezhető és csak a kezelésátlagok kétirányú táblázatát kell közölni (Mead *et al.*, 2003).

Az eredmények prezentálásának és interpretációjának módszereit a 11.7. táblázat mutatja, a kezelésátlagok, a fő hatások és az interakció F-értékek, valamint a prezentáció formájának feltüntetésével. Az adatok három 2×2 -es hipotetikus faktoriális kísérletből származnak, négy blokkal, a hiba MS értékét 9 d.f. alapján 5-nek vesszük.

11.3. Kéttényezős véletlen blokkelrendezés

Az egytényezős kísérletekben használt bármelyik komplett blokk-elrendezés alkalmazható a faktoriális kísérletekben. A randomizációs és az elrendezés eljárásai közvetlenül alkalmazhatók, csupán eltekintünk a faktoriális kezelések faktor összetételétől és minden kezelést úgy tekintünk, mintha nem volnának kapcsolatban. A varianciaanalízisben az egytényezős kísérletnél tárgyaltak közvetlenül alkalmazhatók. További számítási lépések szükségesek azonban, amelyek megfelelnek az egyes faktorok fő hatásainak és interakcióinak. Az alábbiakban bemutatjuk a kéttényezős, RCB elrendezésű kísérlet varianciaanalízisének lépésenkénti folyamatát.

11.3.1. Kísérlet leírása

Az MTA Mezőgazdasági Kutatóintézetében, Martonvásáron kéttényezős véletlen blokkelrendezésű kísérletben vizsgáltuk az őszi búzafajták N-műtrágya reakcióját (11.2. kép). A búzát Norfolk típusú vetésgörögben termesztettük, a növényi sorrend a következő volt: őszi búza, kukorica, tavaszi árpa, borsó. A kísérletet négy ismétlésben állítottuk be, és a búzafajták N-műtrágya reakcióját nyolc N-szinten vizsgáltuk (kg ha^{-1}): 0, 40, 80, 120, 160, 200, 240 és 280, jelölésük: $N_0, N_1, N_2, N_3, N_4, N_5, N_6$ és N_7 . A kísérlet 2004. évi adataiból öt búzafajta (V_1-V_5) termésreakcióját vizsgáljuk. A kísérlet magában foglal nyolc nitrogén dózist, öt búzafajtát és négy ismétlést. A kísérleti adatokat az 11.8. táblázat, a 40 faktoriális kezeléskombináció listáját a 11.9. táblázat mutatja.



11.2. kép. A martonvásári kéttényezős vetésgörög kísérlet búza parcellái (előtérben)

11.8. táblázat. Öt őszi búzafajta szemtermése nyolc nitrogénszinten és négy ismétlésben a kéttényezős véletlen blokkalrendezésű kísérletben

Nitrogén dózisa kg ha ⁻¹	Búza szemtermése t ha ⁻¹				Kezelés összeg (T)
	I. ismétlés	II. Ismétlés	III. ismétlés	IV. ismétlés	
V ₁					
0 (N ₀)	6.82	7.69	6.9	7.14	28.55
40 (N ₁)	7.53	8.0	6.27	7.37	29.17
80 (N ₂)	7.37	7.14	7.22	7.22	28.95
120 (N ₃)	7.53	6.98	7.53	7.53	29.57
160 (N ₄)	6.98	7.92	7.37	6.12	28.39
200 (N ₅)	7.53	7.14	6.82	7.84	29.33
240 (N ₆)	6.51	6.12	7.29	8.39	28.31
280 (N ₇)	7.22	5.96	6.75	8.08	28.01
V ₂					
0 (N ₀)	8.31	8.94	8.31	9.25	34.81
40 (N ₁)	8.86	9.18	10.12	9.02	37.18
80 (N ₂)	9.33	9.1	8.78	9.73	36.94
120 (N ₃)	8.63	8.39	9.65	9.49	36.16
160 (N ₄)	8.94	9.73	7.53	6.59	32.79
200 (N ₅)	8.63	8.55	7.53	8.55	33.26
240 (N ₆)	8.63	8.24	7.53	8.86	33.26
280 (N ₇)	8.63	8.63	6.12	8.55	31.93
V ₃					
0 (N ₀)	7.45	8.24	8.08	8.39	32.16
40 (N ₁)	8.31	9.18	7.22	8.86	33.57
80 (N ₂)	8.31	8.0	8.71	8.94	33.96
120 (N ₃)	7.61	6.98	7.84	8.71	31.14
160 (N ₄)	8.16	8.31	6.27	6.43	29.17
200 (N ₅)	7.53	6.9	4.94	5.33	24.7
240 (N ₆)	5.1	4.71	6.35	6.27	22.43
280 (N ₇)	4.08	5.41	5.49	6.43	21.41
V ₄					
0 (N ₀)	7.84	8.31	7.69	8.47	32.31
40 (N ₁)	8.63	8.71	8.94	8.24	34.52
80 (N ₂)	8.39	8.39	8.63	8.71	34.12
120 (N ₃)	8.16	7.84	7.29	8.0	31.29
160 (N ₄)	7.92	8.86	7.22	5.88	29.88
200 (N ₅)	8.39	7.84	5.96	6.2	28.39
240 (N ₆)	6.75	6.98	7.06	7.84	28.63
280 (N ₇)	5.96	8.47	4.55	6.9	25.88
V ₅					
0 (N ₀)	7.14	6.75	7.22	8.47	29.58
40 (N ₁)	8.31	8.63	8.24	8.71	33.89
80 (N ₂)	8.08	8.39	8.78	8.94	34.19
120 (N ₃)	8.78	9.02	9.49	9.25	36.54
160 (N ₄)	8.71	9.33	8.31	8.47	34.82

A 11.8. táblázat folytatása

Nitrogén dózisa kg ha ⁻¹	Búza szemtermése t ha ⁻¹				Kezelés összeg (T)
	I. ismétlés	II. Ismétlés	III. ismétlés	IV. ismétlés	
200 (N ₅)	9.02	8.63	8.39	8.71	34.75
240 (N ₆)	8.47	8.55	8.16	8.63	33.81
280 (N ₇)	8.0	7.14	8.16	8.47	31.77
Ism. összeg (R)	312.5	317.3	300.7	319	
Főösszeg (G)					1249.5

11.9. táblázat. Öt búzafajta és nyolc nitrogénszint kezeléskombinációi egy 5 × 8-as faktoriális kísérletben

Nitrogén dózisa kg ha ⁻¹	Faktoriális kezeléskombinációk				
	Fajták				
	V ₁	V ₂	V ₃	V ₄	V ₅
0 (N ₀)	N ₀ V ₁	N ₀ V ₂	N ₀ V ₃	N ₀ V ₄	N ₀ V ₅
40 (N ₁)	N ₁ V ₁	N ₁ V ₂	N ₁ V ₃	N ₁ V ₄	N ₁ V ₅
80 (N ₂)	N ₂ V ₁	N ₂ V ₂	N ₂ V ₃	N ₂ V ₄	N ₂ V ₅
120 (N ₃)	N ₃ V ₁	N ₃ V ₂	N ₃ V ₃	N ₃ V ₄	N ₃ V ₅
160 (N ₄)	N ₄ V ₁	N ₄ V ₂	N ₄ V ₃	N ₄ V ₄	N ₄ V ₅
200 (N ₅)	N ₅ V ₁	N ₅ V ₂	N ₅ V ₃	N ₅ V ₄	N ₅ V ₅
240 (N ₆)	N ₆ V ₁	N ₆ V ₂	N ₆ V ₃	N ₆ V ₄	N ₆ V ₅
280 (N ₇)	N ₇ V ₁	N ₇ V ₂	N ₇ V ₃	N ₇ V ₄	N ₇ V ₅

11.3.2. Varianciaanalízis lépései

1. lépés. Jelöljük az ismétlések számát r-rel, az A faktor (fajta) szintjeit a-val és a B faktor (nitrogén) szintjeit b-vel. Vázoljuk fel a variancia táblázatot.

Variáció forrása	Szabadságfok (d.f.)	SS	MS	Számított F-érték	Táblázati F-érték	
					5%	1%
Ismétlés	r-1=3					
Kezelés	ab-1=39					
Fajta (A)	a-1=4					
Nitrogén (B)	b-1=7					
A × B	(a-1)(b-1)=28					
Hiba	(r-1)(ab-1)=117					
Összes	rab-1=159					

2. lépés. Számítsuk ki a kezelés összegeket (T), ismétlés összegeket (R), és a főösszeget (G), mint a 11.10. táblázat mutatja, és számítsuk ki az összes SS-t, az ismétlés SS-t és a hiba SS-t, követve az egytényezős RCB kísérletnél leírt eljárást:

$$C.F. = \frac{G^2}{rab} = \frac{(1249.5)^2}{(4)(5)(8)} = 9757.8$$

$$\text{Összes SS} = \Sigma X^2 - C.F. = \left[(6.82)^2 + (7.69)^2 + \dots + (8.47)^2 \right] - 9757.8 = 204.2$$

$$\text{Ismétlés SS} = \frac{\Sigma R^2}{ab} - C.F. = \frac{(312.5)^2 + \dots + (319.0)^2}{(5)(8)} - 9757.8 = 5.126$$

$$\text{Kezelés SS} = \frac{\Sigma T^2}{r} - C.F. = \frac{(28.55)^2 + \dots + (31.77)^2}{(4)} - 9757.8 = 136.7$$

$$\text{Hiba SS} = \text{Összes SS} - \text{Ismétlés SS} - \text{Kezelés SS} = 204.2 - 5.126 - 136.7 = 62.37$$

Az előzetes varianciaanalízist, a különböző kiszámított SS értékekkel a 11.10. táblázat tartalmazza.

11.10. táblázat. Az adatok előzetes varianciaanalízise

Variáció forrása	Szabadságfok (d.f.)	SS	MS	Számított F-érték	Táblázati F-érték	
					5%	1%
Ismétlés	3	5.126	1.71	3.21*	2.68	3.94
Kezelés	39	136.7	3.51	6.59**	1.49	1.75
Hiba	117	62.37	0.533			
Összes	159	204.2				

*szignifikáns 5%-os szinten, **szignifikáns 1%-os szinten

3. lépés. Készítsük el a faktor A × faktor B kétirányú táblázat összegeket, a számított faktor A és faktor B összegekkel. Példánkban, a fajta × nitrogén (AB) összege, a fajta összeg (A) és a nitrogén összeg (B) együtt a 11.11. táblázatban található.

11.11. táblázat. A fajta × nitrogén termésadatok összesítő táblázata

Nitrogén	Termés összeg (AB)					Nitrogén összeg (B)
	V ₁	V ₂	V ₃	V ₄	V ₅	
N ₀	28.55	34.81	32.16	32.31	29.58	157.41
N ₁	29.17	37.18	33.57	34.52	33.89	168.33
N ₂	28.95	36.94	33.96	34.12	34.19	168.16
N ₃	29.57	36.16	31.14	31.29	36.54	164.7
N ₄	28.39	32.79	29.17	29.88	34.82	155.05
N ₅	29.33	33.26	24.7	28.39	34.75	150.43
N ₆	28.31	33.26	22.43	28.63	33.81	146.44
N ₇	28.01	31.93	21.41	25.88	31.77	139
Fajta összeg (A)	230.28	276.33	228.54	245.02	269.35	1249.52

4. lépés. Számítsuk ki a kezelés SS három frakcionális komponensét:

$$A \text{ SS} = \frac{\sum A^2}{rb} - \text{C.F.} = \frac{(230.28)^2 + \dots + (269.35)^2}{(4)(8)} - 9757.8 = 61.01$$

$$B \text{ SS} = \frac{\sum B^2}{ra} - \text{C.F.} = \frac{(157.41)^2 + \dots + (139)^2}{(4)(5)} - 9757.8 = 39.81$$

$$A \times B \text{ SS} = \text{Kezelés SS} - \text{ASS} - \text{BSS} = 136.7 - 61.01 - 39.81 = 35.88$$

5. lépés. Számítsuk ki az MS értéket mindegyik variáció forrásra, úgy hogy az SS értékét elosztjuk a megfelelő d.f.-fel:

$$A \text{ MS} = \frac{A \text{ SS}}{a-1} = \frac{61.01}{4} = 15.25$$

$$B \text{ MS} = \frac{B \text{ SS}}{b-1} = \frac{39.81}{7} = 5.69$$

$$A \times B \text{ MS} = \frac{A \times B \text{ SS}}{(a-1)(b-1)} = \frac{35.88}{(4)(7)} = 1.28$$

$$\text{Hiba MS} = \frac{\text{Hiba SS}}{(r-1)(ab-1)} = \frac{62.37}{(3)[(5)(8)-1]} = 0.533$$

6. lépés. Számítsuk ki az F-értéket a három faktoriális komponens mindegyikére:

$$F(A) = \frac{A \text{ MS}}{\text{Hiba MS}} = \frac{15.25}{0.533} = 28.61$$

$$F(B) = \frac{B \text{ MS}}{\text{Hiba MS}} = \frac{5.68}{0.533} = 10.66$$

$$F(A \times B) = \frac{A \times B \text{ MS}}{\text{Hiba MS}} = \frac{1.28}{0.533} = 2.40$$

7. lépés. Hasonlítsuk össze mindegyik számított F-értéket a táblázati F-értékkel. A fajtára $f_1 = 4$ és $f_2 = 117$ szabadságfoknál 2.44 az 5%-os szignifikancia szinten és 3.47 az 1%-os szignifikancia szinten. A nitrogénre $f_1 = 7$ és $f_2 = 117$ szabadságfoknál 2.08 az 5%-os szignifikancia szinten és 2.79 az 1%-os szignifikancia szinten. A kölcsönhatásra $f_1 = 28$ és $f_2 = 117$ szabadságfoknál a táblázati F-érték 1.54 az 5%-os és 1.83 az 1%-os szignifikancia szinten. Az eredmény mutatja, hogy a faktor A (fajta) fő hatása és a faktor B (nitrogén) fő hatása, valamint az A \times B interakció egyaránt 1%-os szinten szignifikáns.

8. lépés. Számítsuk ki a variációs koefficiens:

$$CV = \frac{\sqrt{\text{Hiba MS}}}{\text{Foatlag}} 100 = \frac{\sqrt{0.533}}{7.81} \times 100 = 9.35\%$$

9. lépés. A 4–8. lépésben kiszámított értékeket írjuk be a varianciaanalízis táblázatba, melyet az 11.12. táblázat mutat. Az eredményekből látható, hogy szignifikáns interakció van a fajta és a nitrogén között, jelezve, hogy a fajtakülönbséget befolyásolja szignifikánsan az alkalmazott nitrogén szint és hogy a nitrogén hatás szignifikánsan különbözik a vizsgált fajtáknál. A fajta és a nitrogén fő hatása és kölcsönhatásuk egyaránt szignifikáns volt.

11.12. táblázat. A 5 × 8-as faktoriális RCB kísérlet adatainak varianciaanalízise

Variáció forrása	Szabadságfok (d.f.)	SS	MS	Számított F-érték	Táblázati F-érték	
					5%	1%
Ismétlés	3	5.126	1.71	3.21*	2.68	3.94
Kezelés	39	136.7	3.51	6.59**	1.49	1.75
Fajta (A)	(4)	61.01	15.25	28.61**	2.44	3.47
Nitrogén(B)	(7)	39.81	5.69	10.66**	2.08	2.79
A × B	(28)	35.88	1.28	2.40**	1.54	1.83
Hiba	117	62.37	0.533			
Összes	159	204.2				

*szignifikáns 5%-os szinten, **szignifikáns 1%-os szinten

Szignifikáns különbség (SzD) kiszámítása

Egy faktoriális kísérletben több típusú kezelésátlag van. Például, egy 2 × 3-as faktoriális kísérlet magában foglal A faktort két szinttel és B faktort három szinttel, négy típusú átlagot tartalmaz, amelyek összehasonlíthatók:

1. Két A átlag, átlagolva a B faktor mindhárom szintjén
2. Három B átlag, átlagolva az A faktor mindkét szintjén
3. Hat A átlag, két átlag a B faktor három szintjének mindegyikén
4. Hat B átlag, három átlag az A faktor két szintjének mindegyikén.

Az 1. típusú átlag 3 r megfigyelés átlaga; a 2. típusú átlag 2r megfigyelés átlaga; a harmadik és negyedik típus r megfigyelés átlaga. Tehát, az $s_d = (2s^2/r)^{1/2}$ formula csak arra az átlag különbségre megfelelő, amely vagy a 3. típusú vagy a 4. típusú átlagot foglalja magában. Az 1. típusra és a 2. típusra, az r osztót a formulában helyettesíteni kell 3r-rel, illetve 2r-rel. Vagyis, ahhoz, hogy összehasonlítsunk két A átlagot a B faktor minden szintjén átlagolva, az s_d értéket úgy számítjuk ki, mint $(2s^2/3r)^{1/2}$. Hasonlóan ahhoz, hogy összehasonlítsuk bármely B átlag kezeléspárt, átlagolva az A faktor összes szintjén, az s_d értéket úgy számítjuk ki, mint $(2s^2/2r)^{1/2}$ vagy $(s^2/r)^{1/2}$.

(a) Bármely két kombináció között:

$$SzD_{5\%} = t_{5\%} \sqrt{\frac{2\text{Hiba MS}}{r}} = 1.98 \sqrt{\frac{2(0.533)}{4}} = 1.022$$

(b) Az A tényező változatai között a B tényező változatainak átlagában, ha a kölcsönhatás nem szignifikáns:

$$SzD_{5\%} = t_{5\%} \sqrt{\frac{2\text{Hiba MS}}{rb}} = 1.98 \sqrt{\frac{2(0.533)}{(4)(8)}} = 0.361$$

(c) A B tényező változatai között az A tényező változatainak átlagában, ha a kölcsönhatás nem szignifikáns:

$$SzD_{5\%} = t_{5\%} \sqrt{\frac{2\text{Hiba MS}}{ra}} = 1.98 \sqrt{\frac{2(0.533)}{(4)(5)}} = 0.457$$

Ha a kölcsönhatás legalább $P = 5\%$ -os szinten szignifikáns, akkor a tényezők változatai közötti különbségeket általában nem a másik tényező változatainak átlagában, hanem kombinációként vizsgáljuk.

11.3.3. Az eredmények értelmezése

Mivel a fajta és az N-műtrágyázás között a kölcsönhatás szignifikáns volt, a kezelések fő hatásának értelmezése korlátozott. A tényezők változatai közötti különbséget kombinációként vizsgáljuk (11.13. táblázat). Mivel azonban a kezelések fő hatásának MS értéke 4.4-szer, illetve 11.9-szer nagyobb volt a kölcsönhatás MS értékénél, közöljük mindegyik faktor szintjeinek átlagtermését a másik faktor átlagában, az SzD értékekkel együtt.

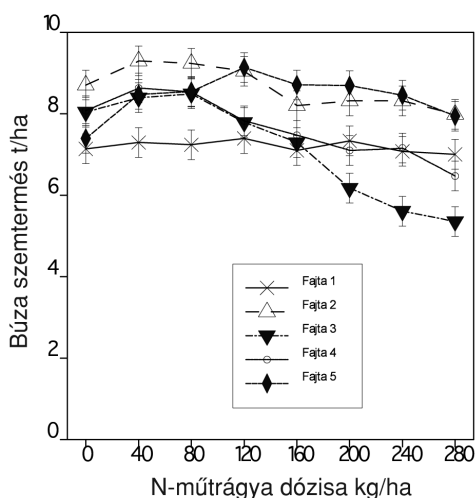
11.13. táblázat. A fajta \times nitrogén kölcsönhatás terméstáblázata a kéttényezős véletlen blokk elrendezésű kísérletben

N-kezelések	Búzafajták					
	V1	V2	V3	V4	V5	Átlag
N0	7.138	8.703	8.04	8.078	7.395	7.871
N1	7.293	9.295	8.393	8.63	8.473	8.417
N2	7.238	9.235	8.49	8.53	8.548	8.408
N3	7.393	9.04	7.785	7.823	9.135	8.235
N4	7.098	8.198	7.293	7.47	8.705	7.753
N5	7.333	8.315	6.175	7.098	8.688	7.522
N6	7.078	8.315	5.608	7.158	8.453	7.322
N7	7.003	7.983	5.353	6.47	7.943	6.950
Átlag	7.196	8.635	7.142	7.657	8.417	

A \times B kölcsönhatás $SzD_{5\%}$: 1.022 t ha⁻¹, A tényező változatai között a B tényező átlagában $SzD_{5\%}$: 0.321 t ha⁻¹, B tényező változatai között a A tényező átlagában $SzD_{5\%}$: 0.457 t ha⁻¹

A kölcsönhatás jól szemléltethető grafikonon (11.6. ábra). A búzafajták N-reakciójának 3 típusa különíthető el az ábra alapján. Az 1. fajta (nagyon régi standard fajta) egyáltalán nem mutat N-műtrágya reakciót, termése 7.003 és 7.393 t ha⁻¹ között kevésbé változik. A 2. és 5. fajta átlagtermése volt a legnagyobb (8.235 és 8.417 t ha⁻¹), N-műtrágya reakciójukra jellemző, hogy 80, illetve 120 kg ha⁻¹ N-dózisnál érték el a maximális termést és ennél nagyobb N-dózisnál termésük kismértékben csökkent. A 3. és 4. fajta termése 80 kg ha⁻¹ N-dózisnál volt a legnagyobb, ennél nagyobb N-dózisnál termésük meredeken csökkent, amely különösen a 3. fajtára volt jellemző.

A termés a fajták átlagában, 40–120 kg ha⁻¹ N-dózis tartományban volt a legnagyobb és ennél nagyobb dózisonál a termés szignifikánsan csökkent. A búzafajták közül – az N-kezelések átlagában – szignifikánsan legnagyobb volt a termése a 2. és 5. fajtának és legkisebb a 3. és 1. fajtának (11.13. táblázat).



11. 6. ábra. A búzafajták N-műtrágya reakciója a kéttényezős véletlen blokk elrendezésű kísérletben

11.3.4. Varianciaanalízis GenStat programmal

Az adatok bevitelekor az ismétlés, fajta és nitrogén képezik a faktorokat, a termés a változó. Az ismétlésnek 4, a fajtának 5, a nitrogénnek 8 szintje van. A varianciaanalízis lépései a következők: *Stats – Analysis of Variance – General – Two-Way ANOVA in Randomized Blocks*. *Y-variate* a termés, *Treatment 1* a fajta, *Treatment 2* a nitrogén, *Blocks* az ismétlés. *Options*-ra kattintás lehetőséget ad az LSD és *Multiple Comparisons* kiválasztására. A *Run* parancsra a program elvégzi a varianciaanalízist. *Further Output – Means plot* kiválasztása lehetővé teszi az adatok ábrázolását. Esetünkben vonalas ábra: *Lines – Factor for X-axis*: nitrogén, *Groups*: fajta – *Difference*: Plot Around Every Mean.

A számítógépes értékelés outputja a következő:

Analysis of variance

Variate: Termes

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Ismetles stratum	3	5.0947	1.6982	3.21	
Ismetles.*Units* stratum					
Nitrogen	7	39.4802	5.6400	10.66	<.001
Fajta	4	60.6830	15.1707	28.67	<.001
Nitrogen.Fajta	28	36.1864	1.2924	2.44	<.001
Residual	117	61.9046	0.5291		
Total	159	203.3490			

Tables of means

Variate: Termes

Grand mean 7.810

Nitrogen	1	2	3	4	5	6	7
	7.871	8.417	8.408	8.235	7.753	7.522	7.322

Nitrogen	8
	6.950

Fajta	1	2	3	4	5
	7.196	8.635	7.142	7.657	8.417

Nitrogen	Fajta	1	2
1		7.138	8.703
2		7.293	9.295
3		7.238	9.235
4		7.393	9.040
5		7.098	8.198
6		7.333	8.315
7		7.078	8.315
8		7.003	7.983

Nitrogen	Fajta 3
1	8.040
2	8.393
3	8.490
4	7.785
5	7.293
6	6.175
7	5.608
8	5.353

Nitrogen	Fajta	4	5
1		8.078	7.395
2		8.630	8.473
3		8.530	8.548
4		7.823	9.135
5		7.470	8.705
6		7.098	8.688
7		7.158	8.453
8		6.470	7.943

Standard errors of differences of means

Table	Nitrogen	Fajta	Nitrogen Fajta
rep.	20	32	4
d.f.	117	117	117
s.e.d.	0.2300	0.1818	0.5143

Least significant differences of means (5% level)

Table	Nitrogen	Fajta	Nitrogen Fajta
rep.	20	32	4
d.f.	117	117	117
l.s.d.	0.4555	0.3601	1.0186

Duncan's multiple range test

Nitrogen

	Mean	
8	6.950	a
7	7.322	ab
6	7.522	bc
5	7.753	bc
1	7.871	cd
4	8.235	de
3	8.408	e
2	8.417	e

Fajta

	Mean	
3	7.142	a
1	7.196	a
4	7.657	b
5	8.417	c
2	8.635	c

11.4. Osztott parcellás (split-plot) elrendezés

Néha a faktoriális kísérletekben a legkisebb egység (parcella), amelyen egy faktor különböző szintjeit állítjuk be, lehet egészen nagy, feltétlenül nagyobb, mint a többi faktor különböző szintjeinek megfelelő parcellaméret. Példaként szolgálnak a talajművelési kísérletek, ahol mindegyik talajművelési módot nagyobb földterületen kell beállítani, míg a többi faktor, mint a fajta, vagy a növény tenyészterülete sokkal kisebb parcellákon állítható be. Egy másik példa lehet egy tejtermelési faktoriális kísérlet különböző legeltetési és fejési módszerekkel, mint faktorokkal. Az ideális méret különböző fejési módszerekre lehet egyetlen tehén, azonban a legkisebb megfelelő parcella különböző legeltetési eljárásokra jelentősen nagyobb annál, mint amelyet egyetlen tehén lelegel.

Az osztott parcellás elrendezésben az egyik faktort a *főparcellákba* helyezük el. Ezt a faktort a *főparcella faktornak* hívjuk. A főparcellát *alparcellákra* osztjuk, amelybe a második faktort, az *alparcella faktort* helyezük el. Tehát minden főparcella az alparcella kezelések (azaz az alparcella faktor szintjei) blokkjává válik. Az osztott parcellás elrendezésnél a főparcella faktor-hatás mérésének precizitását feláldozzuk azért, hogy javítsuk az alparcella faktor pontosságát. Az alparcella-faktor fő hatásának és interakciójának mérése a főparcella-faktorral sokkal pontosabb, mint egy randomizált blokk-elrendezésnél. Másrészt, a főparcella-faktor kezelések (a főparcella-faktor szintjei) hatásának mérése kevésbé pontos, mint a megfelelő véletlen blokk-elrendezésnél. Mivel az osztott-parcellás elrendezésnél a parcella mérete és a hatások mérésének precizitása nem azonos mind a két faktorra, a megfelelő *faktor kijelölése akár főparcellának, akár alparcellának, rendkívül fontos*. Ehhez az alábbi szempontok figyelembevételét javasoljuk:

1. *A pontosság mértéke*. Mivel a B faktor precizitásának mértéke nagyobb, mint az A faktoré, a B faktort jelöljük alparcellának és az A faktort főparcellának. Például a növénynevelő, aki tíz kukoricafajtáját akarja értékelni három műtrágya szinten, egy 10×3 -as faktoriális kísérletben, valószínűleg nagyobb precizitást kíván a fajta összehasonlításához, mint a műtrágya-reakcióra. Ezáltal a fajtát jelöli meg alparcella faktornak és a műtrágyát főparcella faktornak. Másrészt, egy agronómus, aki tanulmányozni kívánja a tíz ígéretes fajta műtrágya-reakcióját, valószínűleg nagyobb pontosságot akar a trágyareakciónak, mint a fajtahatásnak, és a fajtát jelölné meg főparcellának és a műtrágyát alparcellának.
2. *A főhatások relatív nagysága*. Ha egyik faktor (B faktor) főhatása várhatóan sokkal nagyobb és könnyebb kimutatni, mint a másik faktor (A faktor) hatását, a B faktort kellene megjelölni főparcellának és az A faktort alparcellának. Ez növeli annak az esélyét, hogy a kisebb hatású A faktor szintjei között különbséget észleljünk. Például, a trágyázás \times fajta kísérletben a kutató a fajtákat az alparcellában és a trágyázást a főparcellában helyezi el, mivel úgy gondolja, hogy a műtrágya hatása sokkal nagyobb a fajta hatásánál.
3. *Agrotechnikai eljárások*. Az egyik faktorhoz szükséges agrotechnikai eljárások szükségessé tehetik a nagyobb parcellák használatát. Gyakorlati célszerűségből az ilyen faktort helyezük el a főparcellába. Például, egy öntözési kísérletben az öntözési kezeléseket jelöljük ki főparcellaként, hogy minimalizáljuk a víz mozgását a szomszédos parcellák között és ily módon csökkentjük a szegélyhatást.

Vagy, egy kísérletben, ahol több búzafajta teljesítményét értékeljük, különböző műtrágya szinteken, a kutató főparcellának jelölheti meg a trágyázást, hogy minimalizálja annak szükségességét, hogy a különböző trágyázási szintekkel kezelt parcellákat egymástól elkülönítsük puffer sávokkal.

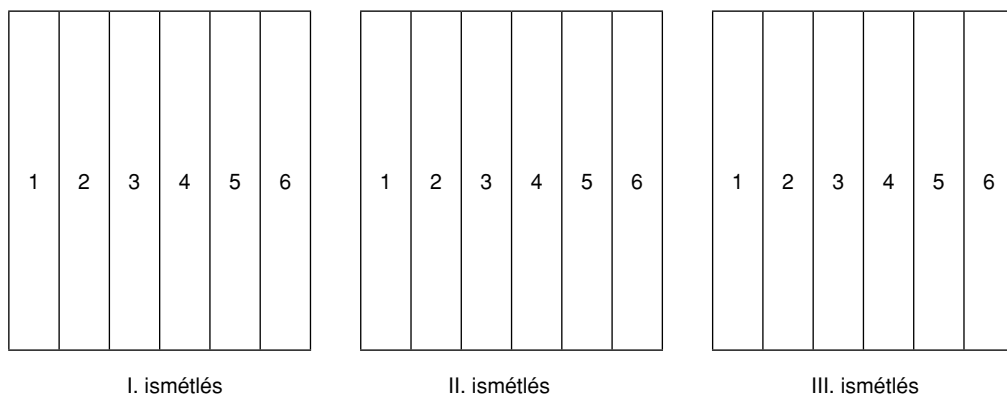
Az osztott parcellás kísérletben a randomizációt és a varianciaanalízist két szinten végezzük – a főparcella szintjén és az alparcella szintjén.

11.4.1. Randomizálás

Két különálló randomizációs folyamat van az osztott parcellás elrendezésben – egy a főparcellára és egy másik az alparcellára. Mindegyik ismétlésben, a főparcella kezeléseket jelöljük ki először random a főparcellákra, ezután random jelöljük ki az alparcella kezeléseket minden főparcellán belül.

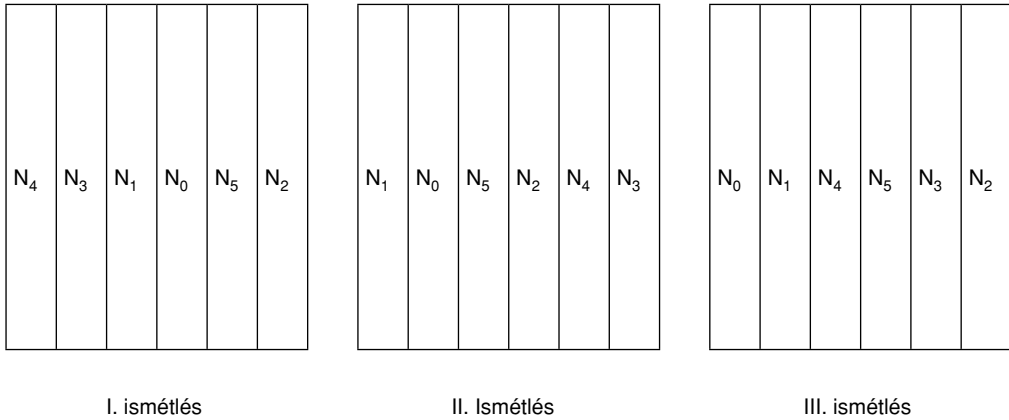
Az osztott parcellás elrendezés randomizációs lépéseiben a jelöli a főparcella kezeléseket számát, b az alparcella kezeléseket számát és r az ismétlések számát. Példaként egy két faktoros kísérletet mutatunk be, amely magában foglalhat nitrogén szintet (főparcella kezeléseket) és négy rizsfajtát (alparcella kezeléseket), három ismétlésben (Gomez és Gomez, 1984 alapján).

1. lépés. A kísérleti területet felosztjuk $r = 3$ blokkba, mindegyiket tovább osztjuk $a = 6$ főparcellára, mint az 11.7. ábra mutatja.



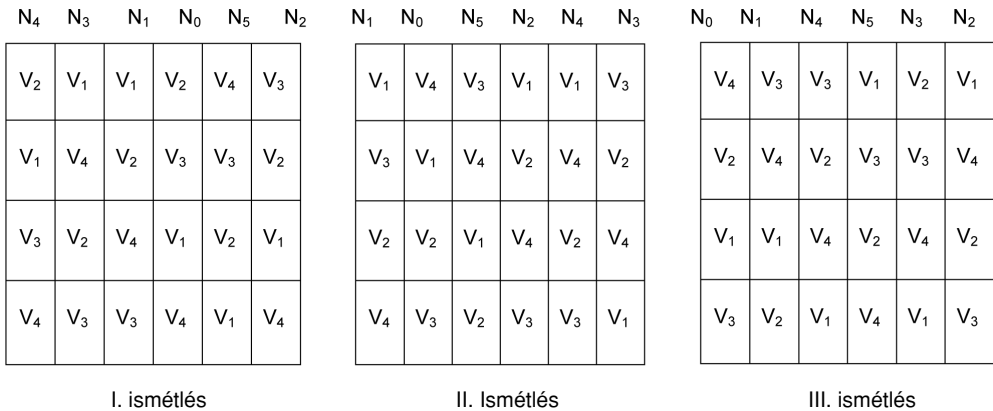
11.7. ábra. A kísérleti terület felosztása három blokkra (ismétlésre), mindegyik magában foglalhat főparcellát, mint az osztott parcellás kísérlet beállításának első lépése, amely tartalmaz három ismétlést és hat főparcella kezelést

2. lépés. Kövessük az RCB randomizációs eljárást a = 6 kezeléssel és r = 3 ismétléssel, random kijelölve a 6 nitrogén kezelést a 6 főparcellára mind a 3 blokkban. (11.8. ábra).



11.8. ábra. A hat nitrogén szint (N₀, N₁, N₂, N₃, N₄ és N₅) random elhelyezése a hat főparcellán

3. lépés. Az $(r)(a) = 18$ főparcella mindegyikét osszuk fel $b = 4$ alparcellára és a randomizációs eljárást követve a $b = 4$ kezelésre és a $(r)(a) = 18$ ismétlésre, random jelöljük ki a 4 fajtát a 4 alparcellába a 18 főparcella mindegyikében. Az eredményt a 11.9. ábra mutatja.



11.9. ábra. A négy rizsfajtát (V₁, V₂, V₃ és V₄), mint alparcellát és a hat nitrogén kezelést (N₀, N₁, N₂, N₃, N₄ és N₅), mint főparcellát tartalmazó kéttényezős osztott parcellás kísérlet elrendezése, három ismétlésben

Figyeljük meg, hogy egy osztott-parcellás kísérlet szántóföldi elrendezése, mint azt a 11.9. ábra mutatja, az alábbi sajátosságokkal rendelkezik:

1. A főparcella mérete b-szer nagyobb az alparcella méreténél. Példánkban, 4 fajtával ($b = 4$) a főparcella mérete 4-szerese az alparcella méretének.

2. Mindegyik főparcella kezelést r -szer teszteljük, míg minden alparcella kezelést $(a)(r)$ -szer tesztelünk. Tehát az alparcella kezelések tesztelésének száma mindig nagyobb, mint a főparcelláké és ez az elsődleges oka az alparcella kezelések nagyobb pontosságának, a főparcella kezelésekhez viszonyítva. Példánkban a nitrogén 6 szintjének mindegyike 3-szor volt tesztelve, míg a 4 fajta mindegyike 18-szor.

Az osztott parcellás kísérletnek a használatát két faktorról tekintettük át, azonban kiterjeszhető *több faktorra*. A faktorokat csoportosíthatjuk a főparcellákra és az alparcellákra. A randomizáció két részben megy végbe, az előzőhöz hasonlóan: azokat a kísérleti kezeléseket, amelyek tartalmazzák a faktorszintek minden kombinációját a főparcella csoportban, beállítjuk véletlenszerűen a főparcellákba minden blokkban és ezt követően az alparcella faktor kombinációkat beállítjuk véletlenszerűen az alparcellákba minden főparcellában.

Főparcellák

P ₃ N ₀ H ₁	P ₃ N ₀ H ₄	P ₃ N ₀ H ₃	P ₃ N ₀ H ₂	P ₂ N ₀ H ₃	P ₂ N ₀ H ₂	P ₂ N ₀ H ₁	P ₂ N ₀ H ₄	(1) P ₁ N ₀
P ₂ N ₁ H ₃	P ₂ N ₁ H ₁	P ₂ N ₁ H ₄	P ₂ N ₁ H ₂	P ₁ N ₀ H ₁	P ₁ N ₀ H ₂	P ₁ N ₀ H ₃	P ₁ N ₀ H ₄	(2) P ₁ N ₁
P ₃ N ₁ H ₃	P ₃ N ₁ H ₂	P ₃ N ₁ H ₄	P ₃ N ₁ H ₁	P ₁ N ₁ H ₂	P ₁ N ₁ H ₄	P ₁ N ₁ H ₃	P ₁ N ₁ H ₁	(3) P ₂ N ₀
								(4) P ₂ N ₁
								(5) P ₃ N ₀
								(6) P ₃ N ₁

11.10. ábra. Az első blokk elrendezése egy osztott parcellás kísérletben két főparcella (P és N) és egy alparcella (H) faktorról

Példaként vegyünk számba egy zellernövény kísérletet, amelyben három vetőmag szaporítási és két tápanyag-ellátottsági szintet hasonlítunk össze négy betakarítási időpontban. A szaporítási és tápanyag faktorokat főparcellákba állítottuk be és a betakarítások képezik az alparcellákat. Mindegyik blokkban van hat főparcella kezelésünk és ezeknek a sorrendjét randomizáltuk (pl. 5, 3, 4, 1, 6, 2). Mindegyik főparcellát felosztottuk négy alparcellára a négy különböző betakarítási időpontnak, H₁, H₂, H₃ és H₄ és az időpontokat random jelöltük ki az alparcellákra minden főparcellában (pl. az első főparcellában 1, 4, 3 és 2; a második főparcellában 3, 2, 1 és 4; és a harmadik főparcellában 3, 1, 4 és 2). A kétszakaszos randomizációs eljárást függetlenül használjuk mindegyik blokkra. A 11.10. ábra mutatja a teljes randomizációt az első blokkra.

11.4.2. Az osztott parcellás elrendezés varianciaanalízise

Az osztott parcellák analízise, jóllehet inkább komplex, egy logikus kiterjesztése a kéttényezős randomizált blokk elrendezésű ANOVA-nak. Az analízist két részben végezzük, először a főparcella kezelést analizáljuk, ezt követi az alparcella kezelés analízise.

A főparcella kezelés analízise

Ahhoz, hogy analizáljuk a főparcellára elhelyezett kezelés hatását (A kezelés), az alparcellákat homogén ismétléseknek tekintjük. A főparcella kezelés ezután úgy analizálható, mintha randomizált blokk elrendezésen belül lenne elhelyezve. Az ANOVA felosztja az összes főparcella eltérésnégyzet összeget (SS_{MP}) a blokk eltérésnégyzet összeg (SS_{blokk}), a főparcella kezelés eltérésnégyzet összeg ($SS_{kezelés A}$) és a főparcella reziduális (maradék) hiba eltérésnégyzet-összeg ($SS_{MP\ maradék}$) között. A közepes négyzetes eltéréseket ezekre az adatokra a szokásos módon határozzuk meg, elosztva az SS értékeket a megfelelő szabadságfokokkal. A főparcella MS aránya a maradék MS-hez ad egy F-értéket, amely lehetővé teszi a főparcella kezelés szignifikanciájának becslését.

Az alparcella kezelés analízise

Az alparcellákra kijelölt kezelések (B kezelés) analíziséhez az alparcellákra vonatkozó összes eltérésnégyzet összeget ($SS_{alparcella}$) felosztjuk az alparcella kezelés ($SS_{kezelésB}$) eltérésnégyzet összeg, az alparcella és a főparcella közötti interakció ($SS_{A \times B}$) eltérésnégyzet összeg és a maradék tag (amely képviseli az alparcella hibát) ($SS_{SP\ maradék}$) között. Ezekből közepes négyzetes eltéréseket (MS) képezünk minden tagra. Az alparcella MS aránya a maradék MS-hez ad egy F-értéket, amely lehetővé teszi az alparcella kezelés-hatás szignifikanciájának becslését. Hasonlóan, az interakció MS aránya az alparcella maradék MS-hez ad egy F-értéket, amely lehetővé teszi az interakció szignifikanciájának becslését. A kéttényezős osztott parcellás elrendezésű kísérlet ANOVA táblázatát az 11.14. táblázat mutatja.

Általában feltételezzük, hogy a nagyobb parcellák változékonnyabbak, mint a kis parcellák (ez igazolása a blokk-képzésnek szántóföldi kísérletben), úgyhogy várhatóan a főparcella hiba MS, s_a^2 , rendszerint nagyobb, mint az alparcella hiba MS, s_b^2 . Ez azt jelenti, hogy az összehasonlítások az alparcella kezelése között és az interakciók a főparcella kezelésekkal rendszerint pontosabbak, mint az összehasonlítások a főparcella kezelése között, és pontosabbak, mintha a szokásos véletlen blokk elrendezést használtuk volna. Egy másik ok az osztott parcellás elrendezés használatára bizonyos vizsgálatoknál, ahol az egyik faktor főhatása kevésbé érdekes, mert már alaposan megvizsgálták, vagy lényegében kevésbé fontos.

Ritka azonban, hogy az osztott parcellás elrendezés nagy javulást eredményezne az alparcella összehasonlítások pontosságában. Az egyetlen meggyőző ok az osztott parcellás elrendezés használatára a nagy parcellák praktikus szükségessége bizonyos kezelés faktorokra.

Nagyon fontos, hogy a varianciaanalízis struktúrája pontosan megfeleljen a kísérleti elrendezés struktúrájának. Amikor egy osztott parcellás elrendezést használunk, a varianciaanalízisnek magában kell foglalni a főparcella variációt, elkülönítve az alparcella variációtól. Néha az osztott parcellás kísérletet randomizált blokk kísérletként analizálják, amely helytelen és hibás következtetéshez vezet.

Az osztott parcellás kísérletek analízise, statisztikai software-t használva, elképzelhetően bonyolultabb, mint egy egyszerű randomizált blokk elrendezése. Specifikusan, a kísérletek értékelésére írt software lehetővé teszi a kísérlet struktúrájának

specifikálását, úgy hogy a főparcellák és az alparcellák hiba tagjai elkülönülnek és a megfelelő hiba tagokat használjuk az F-próbák osztójaként. A legjobb software kiszámítja továbbá a különböző SE értékeket összehasonlításokra a főhatás és interakció átlagok között.

11.14. táblázat. Az osztott parcellás elrendezés ANOVA táblázatainak standard formája

Variáció forrása	SS	DF	MS	F-arány
Főparcella analízis				
Blokkok	SS_{blokk}	$N_{\text{blokk}} - 1$	SS/DF	$\frac{MS_{\text{blokkok}}}{MS_{\text{MPmaradék}}}$
Főparcella (A) kezelés	$SS_{\text{kezelésA}}$	$N_A - 1$	SS/DF	$\frac{MS_{\text{kezelés A}}}{MS_{\text{MPmaradék}}}$
Főparcella maradék	$SS_{\text{MPmaradék}}$	$(N_A - 1) \times (N_{\text{blokk}} - 1)$ (vagy kivonással)	SS/DF	
Főparcella összes	$SS_{\text{főparcella összes}}$	$(N_{\text{blokk}} \times N_A) - 1$	SS/DF	
Alparcella analízis				
Alparcella (B) kezelés	$SS_{\text{kezelésB}}$	$N_B - 1$	SS/DF	$\frac{MS_{\text{kezelés B}}}{MS_{\text{SPmaradék}}}$
A x B interakció	$SS_{\text{AxB interakció}}$	$(N_A - 1) \times (N_B - 1)$	SS/DF	$\frac{MS_{\text{AxB}}}{MS_{\text{SPmaradék}}}$
Alparcella maradék	$SS_{\text{SP maradék}}$	$N_A \times (N_B - 1) \times (N_{\text{blokk}} - 1)$ (vagy kivonással)	SS/DF	
Alparcella összes	$SS_{\text{SP összes}}$	$N_A \times N_{\text{blokk}} \times (N_B - 1)$		
Összes	$SS_{\text{összes}}$	$N - 1$		

Ahol N = a kezelések összes száma az egész kísérletben; N_{blokk} = a blokkok száma; N_A az A kezelés faktor szintjeinek száma; N_B a B kezelés faktor szintjeinek száma.

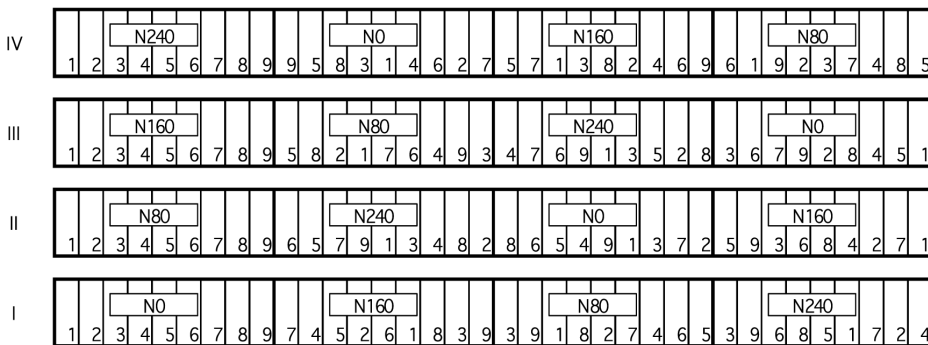
Amikor egyszerűbb statisztikai csomagokat használunk az osztott parcellás vagy még komplikáltabb kísérletek analízisére, a felhasználónak meg kell érteni az illeszteni kívánt modellt és specifikálni kell a hiba tagokat, melyeket osztóként használunk a variancia arány (F-vizsgálat) tesztekben. Sok csomagban a SE-eket a különböző szintű összehasonlításokhoz kézi úton kell kiszámolni.



11.3. kép. A martonvásári N-műtrágyázási tartamkísérlet kukorica monokultúrában (bal oldali főparcella N_0 , a jobb oldali főparcella N_{80} kezelés). Kísérlet kezdete: 1961

11.4.3. Kísérlet leírása

A martonvásári N-műtrágyázási tartamkísérletben osztott parcellás (split-plot) elrendezésben vizsgáljuk a kukorica hibridek N-műtrágya reakcióját (Gyórfy, 1979a; Berzsényi, 1993; Berzsényi és Dang, 2003) (11.3. kép). A négy ismétlésben beállított kísérletben az N-műtrágya kezelések képezik a főparcellákat és a kukorica hibridek az alparcellákat. A négy N-műtrágya kezelésben az N dózisa (kg ha^{-1}) a következő: 0, 80, 160 és 240 (jelölésük: N_1 , N_2 , N_3 és N_4). A kísérlet 1995. évi adataival hat eltérő kukorica hibrid (jelölésük: H_1 - H_6) N-műtrágya reakciójának varianciaanalízisét mutatjuk be. A kísérlet elrendezését a 11.11. ábra, a kísérlet szemtermés adatait a 11.15. táblázat tartalmazza.



11.11. ábra. A martonvásári kéttényezős, osztott parcellás N-műtrágyázási kísérlet elrendezése. I-IV: ismétlés, N_0 , N_{80} , N_{160} és N_{240} : N-műtrágya kezelések, 1-9: kukorica hibrid

11.15. táblázat. Hat kukorica hibrid szemtermés 1995. évi adatai négy eltérő nitrogén szinten, négy ismétlésben a martonvásári tartamkísérletben

Kukorica hibrid	Kukorica szemtermése t ha ⁻¹			
	I. ismétlés	II. ismétlés	III. ismétlés	IV. ismétlés
	N ₁ (0 kg ha ⁻¹ N)			
H ₁	4.09	5.05	5.05	3.64
H ₂	5.11	5.44	5.2	3.41
H ₃	5.03	5.66	4.73	4.48
H ₄	4.71	6	4.81	4.43
H ₅	5.29	6.57	5.98	5.41
H ₆	5.98	5.84	6.39	4.41
	N ₂ (80 kg ha ⁻¹ N)			
H ₁	9.9	8.8	7.97	6.68
H ₂	9.06	8.99	9.17	6.9
H ₃	8.42	9.11	7.05	5.88
H ₄	9.58	9.38	7.05	5.81
H ₅	10.35	10.53	9.79	7.65
H ₆	11.42	12.07	11.26	9.17
	N ₃ (160 kg ha ⁻¹ N)			
H ₁	8.13	10.36	8.46	6.41
H ₂	11.5	9.81	10.33	9.2
H ₃	10.24	10.23	10.54	7.37
H ₄	11.7	11.24	9.06	9.96
H ₅	11.86	11.84	12.7	9.26
H ₆	12.43	14.67	12.48	11.09
	N ₄ (240 kg ha ⁻¹ N)			
H ₁	8.18	10.23	6.79	8.29
H ₂	11.77	10.77	9.34	9.5
H ₃	9.21	11.92	10.07	8.66
H ₄	10.17	10.2	9.48	8.21
H ₅	10.15	13.5	12.56	9.46
H ₆	12.31	14.79	13.81	11.14

11.4.4. Varianciaanalízis lépései

Az osztott parcellás kísérlet varianciaanalízise felosztható főparcella analízisre és alparcella analízisre. Jelöljük A-val a főparcella faktort és B-vel az alparcella faktort. Végezzük el a varianciaanalízist'

1. lépés. Készítsük el az osztott parcellás elrendezés variancia táblázatát.

Variáció forrása	Szabadságfok	SS	MS	Számított F-érték	Táblázati F-érték	
					5%	1%
Ismétlés	r-1=3					
Főparcella faktor (A)	a-1=3					
Hiba (a)	(r-1)(a-1)=9					
Alparcella faktor (B)	b-1=5					
A × B	(a-1)(b-1)=(15)					
Hiba (b)	a(r-1)(b-1)=60					
Összes	rab-1=95					

2. lépés. Készítsünk el két összesítő táblázatot:

A. Az ismétlés \times faktor A kétirányú összegező táblázat, az ismétlés összegekkel, a faktor A összesen és a kiszámított főösszeg. Példánkban, az ismétlés \times nitrogén összegező táblázatot (RA), az ismétlés összegekkel (R), nitrogén összegekkel (A), és a főösszeggel (G) a 11.16. táblázat mutatja.

B. A faktor A \times faktor B összesítő táblázat, faktor B összegekkel. Példánkra, a nitrogén \times fajta összegező táblázatot (AB), a fajta összegekkel (B) a 11.17. táblázat mutatja.

11.16. táblázat. Az ismétlés \times nitrogén termés összesen táblázata

Nitrogén	Termés összesen (RA)				Nitrogén összes (A)
	I. ismétlés	II. ismétlés	III. ismétlés	IV. ismétlés	
N ₁	30.21	34.56	32.16	25.78	122.71
N ₂	58.73	58.88	52.29	42.09	211.99
N ₃	65.86	68.15	63.57	53.29	250.87
N ₄	61.79	71.41	62.05	55.26	250.51
Ismétlés összes (R)	216.59	233.0	210.07	176.42	
Főösszeg (G)					836.08

11.17. táblázat. A nitrogén \times hibrid termés összesen táblázata

Nitrogén	Termés összesen (AB)					
	H ₁	H ₂	H ₃	H ₄	H ₅	H ₆
N ₁	17.83	19.16	19.9	19.95	23.25	22.62
N ₂	33.35	34.12	30.46	31.82	38.32	43.92
N ₃	33.36	40.84	38.38	41.96	45.66	50.67
N ₄	33.49	41.38	39.86	38.06	45.67	52.05
Hibrid összesen (B)	118.03	135.5	128.6	131.79	152.9	169.26

3. lépés. Számítsuk ki a korrekciós faktort és az SS értékeket a főparcella analízishez:

$$C.F. = \frac{G^2}{rab} = \frac{(836.08)^2}{(4)(4)(6)} = 7281.56$$

$$\text{Összes SS} = \Sigma X^2 - C.F. = \left[(4.09)^2 + \dots + (11.14)^2 \right] - 7281.56 = 705.44$$

$$\text{Ismétlés SS} = \frac{\Sigma R^2}{ab} - C.F. = \frac{(216.59)^2 + (233.0)^2 + (210.07)^2 + (176.42)^2}{(4)(6)} - 7281.56 = 70.67$$

$$A \text{ (nitrogén) SS} = \frac{\Sigma A^2}{rb} - C.F. = \frac{(122.71)^2 + \dots + (250.51)^2}{(4)(6)} - 7281.56 = 455.52$$

$$\begin{aligned} \text{Hiba (a) SS} &= \frac{\Sigma(RA)^2}{b} - C.F. - \text{Ismétlés SS} - \text{ASS} = \\ &= \frac{(30.21)^2 + \dots + (55.26)^2}{(6)} - 7281.56 - 70.56 - 455.52 = 10.92 \end{aligned}$$

4. lépés. Számítsuk ki az SS értékeket az alparcella analízishez.

$$B(\text{hibrid}) \text{ SS} = \frac{\Sigma B^2}{ra} - \text{C.F.} = \frac{(118.03)^2 + \dots + (169.26)^2}{(4)(4)} - 7281.56 = 107.5$$

A x B (nitrogén x hibrid) SS =

$$\frac{\Sigma(AB)^2}{r} - \text{C.F.} - \text{BSS} - \text{ASS} = \frac{(17.83)^2 + \dots + (52.05)^2}{(4)} - 7281.56 - 107.5 - 455.52 = 24.92$$

$$\text{Hiba (b) SS} = \text{Összes SS} - (\text{összes többi SS összege}) = 705.44 - (70.67 + 455.52 + 10.92 + 107.5 + 24.92) = 35.91$$

5. lépés. A variáció minden forrására számítsuk ki az MS értékeket úgy, hogy az SS-t osztjuk a megfelelő d.f. értékkel:

$$\text{Ismétlés MS} = \frac{\text{Ismétlés SS}}{r-1} = \frac{70.67}{3} = 23.56$$

$$A \text{ MS} = \frac{A \text{ SS}}{a-1} = \frac{455.52}{3} = 151.84$$

$$\text{Hiba (a) MS} = \frac{\text{Hiba (a) SS}}{(r-1)(a-1)} = \frac{10.92}{9} = 1.213$$

$$B \text{ MS} = \frac{B \text{ SS}}{b-1} = \frac{107.5}{5} = 21.5$$

$$A \times B \text{ MS} = \frac{A \times B \text{ SS}}{(a-1)(b-1)} = \frac{24.92}{15} = 1.66$$

$$\text{Hiba (b) MS} = \frac{\text{Hiba(b)SS}}{a(r-1)(b-1)} = \frac{35.91}{60} = 0.599$$

6. lépés. Számítsuk ki az F-értéket minden hatásra, amelyet tesztelni szükséges úgy, hogy elosztjuk mindegyik MS-t a megfelelő hiba tényezővel:

$$F(A) = \frac{A \text{ MS}}{\text{Hiba(a)MS}} = \frac{151.84}{1.213} = 125.18$$

$$F(B) = \frac{B \text{ MS}}{\text{Hiba(b)MS}} = \frac{21.5}{0.599} = 35.89$$

$$F(A \times B) = \frac{A \times B \text{ MS}}{\text{Hiba(b)MS}} = \frac{1.66}{0.599} = 2.77$$

7. lépés. Mindegyik kiszámított F-értékhez, melynek értéke nem kisebb, mint 1, keressük ki a megfelelő táblázati F-értékét a kívánt szignifikancia szinten, f_1 = a számláló MS d.f. értéke és f_2 = a nevező MS d.f. értéke. Például, a táblázati F-érték az F (A × B)-re 1.86 az 5%-os és 2.40 az 1%-os szignifikancia szintre.

8. lépés. Számítsuk ki a két variációs koefficiens, az egyik megfelel a főparcella analízisnek és a másik az alparcella analízisnek:

$$CV(a) = \frac{\sqrt{\text{Hiba (a) MS}}}{\text{Főátlag}} \times 100 = \frac{\sqrt{1.213}}{8.709} \times 100 = 12.65\%$$

$$CV(b) = \frac{\sqrt{\text{Hiba (b) MS}}}{\text{Főátlag}} \times 100 = \frac{\sqrt{0.599}}{8.709} \times 100 = 8.89\%$$

A CV(a) értéke jelzi a főparcella faktornál elért pontosságot. A CV(b) értéke jelzi az alparcella faktorra és főparcella × alparcella kölcsönhatásra elért pontosságot. A CV(b) értékétől azt várjuk, hogy kisebb legyen, mint a CV(a), mivel a főparcella faktort feltételezhetően kisebb pontossággal mérjük, mint az alparcella faktort. Kísérleti adatainkra a CV (a) értéke nagyobb, mint a CV (b) értéke (12.65% vs. 8.89%). Ez a trend azonban nem mindig érvényesül.

9. lépés. A 3–8. lépésben számított valamennyi értéket írjuk be a variancia táblázatba (11.18. táblázat). Hasonlítsuk össze mindegyik számított F-értéket a megfelelő táblázati F-értékkel és jelöljük meg a megfelelő csillagokkal a szignifikanciát.

Példánkban, mindhárom hatás (a két főhatás és az interakció hatás) erősen szignifikáns. Szignifikáns interakciónál óvatosság szükséges az eredmények értelmezésénél.

11.18. táblázat. A 4 × 6-os osztott parcellás faktoriális kísérlet adatainak varianciaanalízise

Variáció forrása	Szabadságfok	SS	MS	Számított F-érték	Táblázati F-érték	
					5%	1%
Ismétlés	3	70.67	23.56			
Nitrogén (A)	3	455.52	151.84	125.18**	3.86	6.99
Hiba (a)	9	10.92	1.213			
Hibrid (B)	5	107.5	21.5	35.89**	2.37	3.34
A × B	15	24.92	1.66	2.77**	1.86	2.40
Hiba (b)	60	35.91	0.599			
Összes	95	705.44				

Az SzD kiszámítása az osztott parcellás kísérletben

Egy osztott parcellás kísérletben, két változó faktorról és két hiba tényezővel, négy különböző típusú pár-összehasonlítás lehetséges. Ezek saját SzD értékeket igényelnek. Az összehasonlítások az alábbiak (Gomez és Gomez, 1984):

- (i) Összehasonlítás két főparcella kezelés átlaga között az összes alparcella kezelés átlagában.
- (ii) Összehasonlítás két alparcella kezelés átlag között az összes főparcella kezelés átlagában.
- (iii) Összehasonlítás két alparcella kezelés átlag között ugyanabban a főparcella kezelésben.
- (iv) Összehasonlítás két főparcella kezelésátlag között ugyanabban vagy eltérő alparcella kezeléseknél (azaz bármely két kezeléskombináció átlaga).

A 11.19. táblázat megadja a formulákat a kezeléskülönbség megfelelő szórásának (s_d) kiszámítására minden egyes pár összehasonlításra. Amikor az s_d (= s.e.d vagy SE) kiszámítása magában foglal egynél hibátényezőt, mint a 4. típusú összehasonlításnál, a standard táblázati t érték nem használható közvetlenül és a súlyozott táblázati t értéket kell kiszámolni. A formulák a súlyozott t érték kiszámítására a 11.20. táblázatban találhatók.

11.19. táblázat. Két kezelésátlag közötti különbség szórása a négy típusú kezelés-pár összehasonlítására osztott parcellás kísérletnél (Gomez és Gomez, 1984)

Sorszám	Kezelés-pár összehasonlítás típusa	s_d
1.	Két főparcella átlaga között (minden alparcella átlagában)	$\sqrt{\frac{2E_a}{rb}}$
2.	Két alparcella átlaga között (minden főparcella átlagában)	$\sqrt{\frac{2E_b}{ra}}$
3.	Két alparcella átlaga között ugyanabban a főparcella kezelésben	$\sqrt{\frac{2E_b}{r}}$
4.	Két főparcella átlaga között ugyanabban vagy különböző alparcella kezeléseknél	$\sqrt{\frac{2[(b-1)E_b + E_a]}{rb}}$

E_a = hiba (a) MS, E_b = hiba (b) MS, r = az ismétlések száma, a = főparcella kezelések száma, b = alparcella kezelések száma.

Példánkban, a 4×6 -os faktoriális kísérletben a varianciaanalízis (11.18. táblázat) erősen szignifikáns interakciót mutatott a nitrogén és a hibrid között, jelezve, hogy a hibridhatások változnak a kijuttatott műtrágya dózisával. Ezáltal az összehasonlítás a nitrogénátlagok között az összes hibrid átlagában vagy a hibrid átlagok között az összes nitrogén dózis átlagában nem hasznos. Sokkal megfelelőbb átlag összehasonlítások a hibridátlagok között ugyanannál a nitrogén dózisonál vagy ugyanazon hibrid nitrogén dózis átlagai között. Mivel azonban a pár összehasonlítás ugyanazon hibrid nitrogén dózis átlagai között nem megfelelő a nitrogén dózis kezelésekvantitatív természete miatt, csak a hibridátlagok közötti összehasonlítást ugyanannál a nitrogén dózisonál mutatjuk be.

Az SzD kiszámításának lépései két hibrid átlag összehasonlítására ugyanannál a nitrogén dózisonál (azaz, két alparcella átlag között ugyanannál a főparcella kezelésnél) a következők:

1. lépés. Számítsuk ki a kezelésközépek közötti különbség szórását (s_d vagy s.e.d.), követve a táblázat 3. típusú összehasonlítási formuláját:

$$s_d = \sqrt{\frac{2E_b}{r}} = \sqrt{\frac{2(0.599)}{4}} = 0.5473 \text{ t ha}^{-1}$$

ahol E_b értéke a varianciaanalízis Hiba (b) MS értékével azonos (11.18. táblázat).

2. lépés. Keressük ki a táblázati t értéket $n = \text{hiba (b) d.f.} = 60$ -nál, amely 2.00 az 5%-os és 2.66 az 1%-os szignifikanciaszinten.

3. lépés. $SzD_{\alpha} = (t_{\alpha}) (s_d)$ formula alapján számítsuk ki az SzD értékeket 5%-os és 1%-os szignifikanciaszinten:

$$SzD_{5\%} = (2.00)(0.5473) = 1.0946 \text{ t ha}^{-1}$$

$$SzD_{1\%} = (2.66)(0.5473) = 1.4558 \text{ t ha}^{-1}$$

11.20. táblázat. Súlyozott táblázati t-értékek összefüggésben a különböző átlag összehasonlításokkal, melyeknek átlag különbség hibája (s.e.d.) egynél több hiba MS értéket foglal magában (Gomez és Gomez, 1984 alapján)

Kezelés összehasonlítás			Súlyozott táblázati t-érték
Forrás			
Sorszám	Táblázatszám	Összehasonlítás szám	
1		4	$\frac{(b-1)E_b t_b + E_a t_a}{(b-1)E_b + E_a}$
2		3	$\frac{(b-1)E_c t_c + E_a t_a}{(b-1)E_c + E_a}$
3		4	$\frac{(a-1)E_c t_c + E_b t_b}{(a-1)E_c + E_b}$
4		8	$\frac{(b-1)E_b t_b + E_a t_a}{(b-1)E_b + E_a}$
5		9	$\frac{(c-1)E_c t_c + E_b t_b}{(c-1)E_c + E_b}$
6		10	$\frac{(c-1)E_c t_c + E_b t_b}{(c-1)E_c + E_b}$
7		11	$\frac{(c-1)E_c t_c + E_a t_a}{(c-1)E_c + E_a}$
8		12	$\frac{b(c-1)E_c t_c + (b-1)E_b t_b + E_a t_a}{b(c-1)E_c + (b-1)E_b + E_a}$

A, b, c, E_a , E_b és E_c definíciója a 11.19., 11.28. és 12.12. táblázatokban található; t_a , t_b és t_c táblázati t értékek $n = \text{d.f.}$ az E_a , E_b és E_c -nek megfelelően.

11.4.5. Az eredmények értelmezése

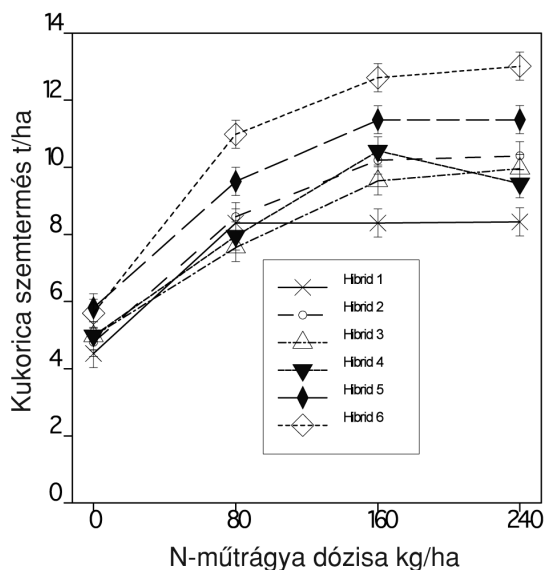
Készítsük el a hibrid \times nitrogén kétirányú táblázatot a kezelésközvetlenekkel és közöljük az SzD értékeket két hibrid átlag összehasonlítására azonos nitrogén dózisonál (11.21. táblázat). Minden hibrid pár (azonos nitrogén szinten) összehasonlítására számítsuk ki az átlagos különbséget és hasonlítsuk össze az SzD értékkel. Például, az egyik bennünket érdeklő átlagos különbség lehet a H₁ és H₅ hibrid között 0 N kg ha⁻¹-nél, amely 4.457 – 5.812 = 1.355 kg ha⁻¹. Mivel az átlagos különbség nagyobb, mint az SzD érték, a különbség szignifikáns P = 5%-os szinten.

11.21. táblázat. A hat kukorica hibrid átlagos termése négy nitrogén szinten vizsgálva osztott parcellás elrendezésnél

Nitrogén dózis kg ha ⁻¹	Termésátlag ^a t ha ⁻¹					
	H ₁	H ₂	H ₃	H ₄	H ₅	H ₆
0	4.457	4.790	4.975	4.987	5.812	5.655
80	8.337	8.530	7.615	7.955	9.580	10.980
160	8.340	10.210	9.595	10.490	11.415	12.667
240	8.372	10.345	9.965	9.515	11.417	13.012

^aNégy ismétlés átlaga. Az SzD értékek két hibrid összehasonlításához azonos nitrogén dózisonál, 1.0946 t ha⁻¹ az 5%-os és 1.4558 t ha⁻¹ az 1%-os szignifikancia szinten.

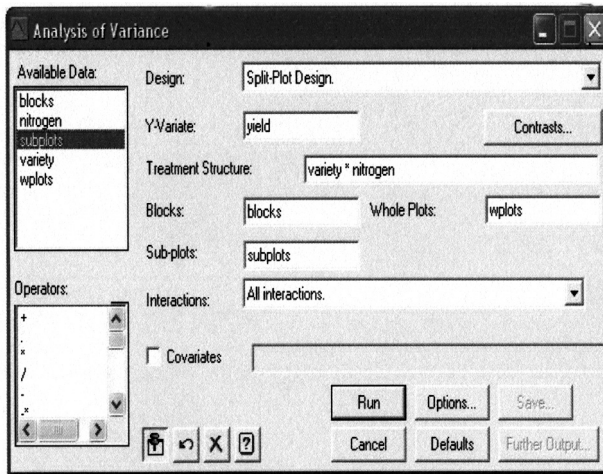
A vizsgált kukorica hibridek N-műtrágya reakcióját a 11.12. ábra szemlélteti. Látható, hogy az N \times hibrid interakcióért elsősorban az 1. és 4. hibrid eltérő N-műtrágya reakciója felelős. A kukorica hibridek N-műtrágya reakciójáról a tartamkísérletben részletes adatokat közöl Berzsényi és Dang (2003).



11.12. ábra. N-műtrágyázási tartamkísérletben 1995. évben vizsgált 6 eltérő kukorica hibrid N-műtrágya reakciója

11.4.6. Varianciaanalízis GenStat programmal

Az Analysis of Variance választási lehetőségei közül a General Analysis of Variance menüből a Split-plot Design-t választjuk. Az adatfájlban külön oszlopok tartalmazzák a termésadatokat, a blokkokat, a főparcellákat és az alparcellákat. Az Available Data doboz tartalmazza a megfelelő input adatokat. A kísérleti elrendezést leíró modellt be kell írni a Treatment Structure input dobozba. Példánkban a kezelés struktúra N-műtrágyázás × Hibrid, amely magában foglalja a faktorok fő hatásának és interakciójuknak a kiszámítását. Az Interactions dobozban rendszerint az alapértelmezés szerinti All interactions-t választjuk (11.13. ábra).



11.13. ábra. GenStat menü doboz az osztott parcellás ANOVA számításához

A varianciaanalízis eredménye sorrendben tartalmazza a variancia táblázatot, az átlagokat, az átlagok különbségének szórását és a szignifikáns differenciákat. A 11.22. táblázat tartalmazza a kísérlet alapadatait.

Analysis of variance

Variate: Termes

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Ismetles stratum	3	70.6753	23.5584	19.33	
Ismetles.Fertilisation stratum					
Fertilisation	3	455.4616	151.8205	124.59	<.001
Residual	9	10.9671	1.2186	2.04	
Ismetles.Fertilisation.Hybrid stratum					
Hybrid	5	107.5182	21.5036	35.97	<.001
Fertilisation.Hybrid	15	25.0422	1.6695	2.79	0.002
Residual	60	35.8701	0.5978		
Total	95	705.5345			

Tables of means

Variate: Termes

Grand mean 8.709

Fertilisation	1	2	3	4				
	5.113	8.833	10.453	10.438				
Hybrid	1	2	3	4	5	6		
	7.377	8.469	8.037	8.237	9.556	10.579		
Fertilisation	Hybrid	1	2	3	4	5	6	
1		4.457	4.790	4.975	4.987	5.812	5.655	
2		8.337	8.530	7.615	7.955	9.580	10.980	
3		8.340	10.210	9.595	10.490	11.415	12.667	
4		8.372	10.345	9.965	9.515	11.417	13.012	

Standard errors of differences of means

Table	Fertilisation	Hybrid	Fertilisation Hybrid
rep.	24	16	4
s.e.d.	0.3187	0.2734	0.5922
d.f.	9	60	56.40

Except when comparing means with the same level(s) of

Fertilisation	0.5467
d.f.	60

Least significant differences of means (5% level)

Table	Fertilisation	Hybrid	Fertilisation Hybrid
rep.	24	16	4
l.s.d.	0.7209	0.5468	1.1860
d.f.	9	60	56.40

Except when comparing means with the same level(s) of

Fertilisation	1.0936
d.f.	60

11.22. táblázat. Nitrogén műtrágyázási tartamkísérlet 1995. évi alapadatai

Ismétlés	N-szint	Hibrid	Termés	Ismétlés	N-szint	Hibrid	Termés
1	1	1	4.09	1	3	1	8.13
2	1	1	5.05	2	3	1	10.36
3	1	1	5.05	3	3	1	8.46
4	1	1	3.64	4	3	1	6.41
1	1	2	5.11	1	3	2	11.5
2	1	2	5.44	2	3	2	9.81
3	1	2	5.2	3	3	2	10.33
4	1	2	3.41	4	3	2	9.2
1	1	3	5.03	1	3	3	10.24
2	1	3	5.66	2	3	3	10.23
3	1	3	4.73	3	3	3	10.54
4	1	3	4.48	4	3	3	7.37
1	1	4	4.71	1	3	4	11.7
2	1	4	6	2	3	4	11.24
3	1	4	4.81	3	3	4	9.06
4	1	4	4.43	4	3	4	9.96
1	1	5	5.29	1	3	5	11.86
2	1	5	6.57	2	3	5	11.84
3	1	5	5.98	3	3	5	12.7
4	1	5	5.41	4	3	5	9.26
1	1	6	5.98	1	3	6	12.43
2	1	6	5.84	2	3	6	14.67
3	1	6	6.39	3	3	6	12.48
4	1	6	4.41	4	3	6	11.09
1	2	1	9.9	1	4	1	8.18
2	2	1	8.8	2	4	1	10.23
3	2	1	7.97	3	4	1	6.79
4	2	1	6.68	4	4	1	8.29
1	2	2	9.06	1	4	2	11.77
2	2	2	8.99	2	4	2	10.77
3	2	2	9.17	3	4	2	9.34

A 11.22. táblázat folytatása

Ismétlés	N-szint	Hibrid	Termés	Ismétlés	N-szint	Hibrid	Termés
4	2	2	6.9	4	4	2	9.5
1	2	3	8.42	1	4	3	9.21
2	2	3	9.11	2	4	3	11.92
3	2	3	7.05	3	4	3	10.07
4	2	3	5.88	4	4	3	8.66
1	2	4	9.58	1	4	4	10.17
2	2	4	9.38	2	4	4	10.2
3	2	4	7.05	3	4	4	9.48
4	2	4	5.81	4	4	4	8.21
1	2	5	10.35	1	4	5	10.15
2	2	5	10.53	2	4	5	13.5
3	2	5	9.79	3	4	5	12.56
4	2	5	7.65	4	4	5	9.46
1	2	6	11.42	1	4	6	12.31
2	2	6	12.07	2	4	6	14.79
3	2	6	11.26	3	4	6	13.81
4	2	6	9.17	4	4	6	11.14

11.5. Sávos (strip-plot vagy criss-cross) elrendezés

A sávos elrendezés különösen megfelel azokra a két-faktoros kísérletekre, melyekben két faktor közötti interakció (kölsönhatás) mérésének kívánt pontossága nagyobb, mint bármelyik két faktor fő hatásának mérése.

Ez három parcellaméret használatával valósítható meg:

- (1) Vertikális sávú parcella az első faktorra – *a vertikális faktor*
- (2) Horizontális sávú parcella a második faktorra – *a horizontális faktor*
- (3) Interakció parcella a két faktor közötti interakcióra

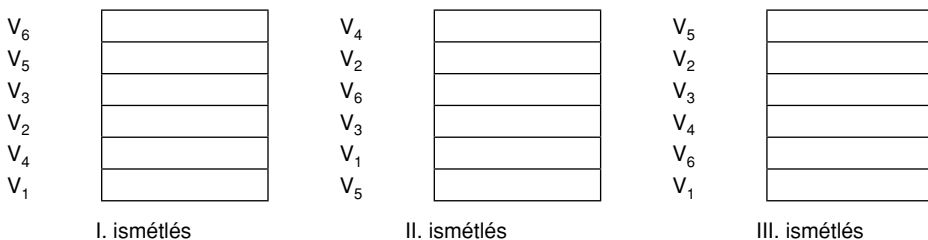
A vertikális sávú parcella és a horizontális sávú parcella mindig merőleges egymásra. Nincs azonban összefüggés méreteik között, eltérően az osztott parcellás elrendezés főparcellájából és alparcellájától. Az interakció parcella természetesen a legkisebb. Tehát, a sávos parcella elrendezésben a pontosság fokát összefüggésben mindkét faktor fő hatásával feláldozzuk azért, hogy javítsuk az interakció pontosságát.

11.5.1. Randomizáció és elrendezés

A sávos parcellás elrendezés randomizációjának folyamata két független randomizációs folyamatot tartalmaz – egyet a horizontális faktorra és egy másikat a vertikális faktorra. A két eljárás végrehajtásának sorrendje lényegtelen.

Jelölje A a horizontális faktort és B a vertikális faktort, továbbá „a” és „b” jelölje szintjeiket. A korábbiakhoz hasonlóan r jelöli az ismétlések számát. A két-faktoros sávos elrendezésű kísérlet lépéseit bemutatjuk egy példán keresztül, amely magában foglal hat fajtát (horizontális kezelések) és három nitrogén dózist (vertikális kezelések), három ismétlésben.

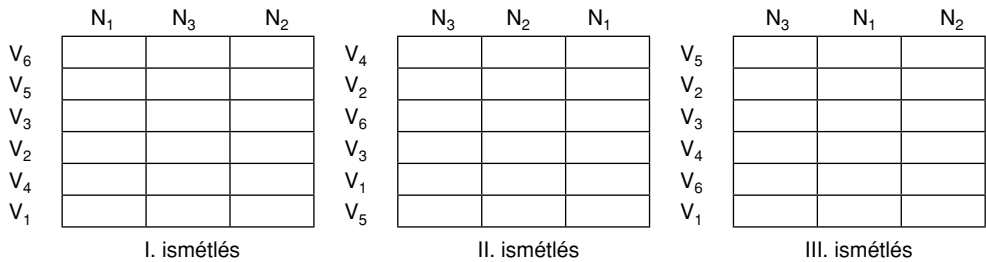
1. lépés. Jelöljük ki a horizontális parcellákat úgy, hogy a kísérleti területet felosztjuk $r = 3$ blokkra és mindegyiket felosztjuk $a = 6$ horizontális sávra. Követjük a randomizációs eljárást a randomizált teljes blokk elrendezésre $a = 6$ kezeléssel és $r = 3$ ismétléssel és random jelöljük ki a hat fajtát a hat horizontális sávra a három blokk mindegyikében, elkülönítetten és függetlenül. Ennek eredményét az 11.14. ábra mutatja.



11.14. ábra. A hat fajta (V_1, V_2, V_3, V_4, V_5 és V_6) random elhelyezése a horizontális sávokban a sávos elrendezésű kísérlet három ismétlésében

2. lépés. Jelöljük ki a vertikális parcellákat oly módon, hogy felosztjuk mindegyik blokkot $b = 3$ vertikális sávra. Kövessük a randomizációs eljárást a randomizált teljes blokk elrendezésre $b = 3$ kezeléssel és $r = 3$ ismétléssel és random jelöljük ki a három

nitrogén dózist a három vertikális sávra mindhárom blokkban, elkülönítetten és függetlenül. A végső elrendezést a 11. 15. ábra szemlélteti.



11. 15. ábra. A hat fajta (V₁, V₂, V₃, V₄, V₅ és V₆) mint horizontális kezelések és a három nitrogén dózis (N₁, N₂, N₃), mint vertikális kezelések elrendezése egy sávos parcellás kísérletben, három ismétlésben

11.5.2. Kísérlet leírása

A sávos elrendezésű kísérlet varianciaanalízisét a martonvásári növényszám kísérlet 1995. évi adataival mutatjuk be (Győrffy, 1979a; Berzsenyi et al., 1994b). A kéttényezős kísérlet magában foglal hat kukorica hibridet (horizontális sáv) és kilenc növényszámot (vertikális sáv), négy ismétlésben. A hibridek eltérő tenyésztési csoportokat és növényszám reakció típusokat képviselnek (jelölésük: H₁ – H₆). A folytonos sűrítést megvalósító növényszámok 20 10³ és 100 10³ t ha⁻¹ között változtak, 10³ t ha⁻¹ kezeléskénti különbséggel (jelölésük: D₁ – D₉).

A kísérletet eredetileg Győrffy Béla tervezte meg, kezdetben mikroparcellákra, melyet később parcella vetőgépre adaptáltunk. Megjegyezzük, hogy a kísérletben 19 kukorica hibrid növényszám reakcióját vizsgáltuk. A kísérlet elrendezését az 11.4. kép mutatja, az adatok az 11.23. táblázatban találhatóak.



11.4. kép. Sávos elrendezésű kukorica tenyésztési kísérlet a martonvásári növénytermesztési kísérleti területen

11.5.3. Varianciaanalízis lépései

A sávos parcellás elrendezés varianciaanalízise három részre osztható: a horizontális-faktor analízis, a vertikális-faktor analízis, és az interakció analízis.

11.23. táblázat. Hat kukorica hibrid termésadatai kilenc növényszámnál a martonvásári sávos elrendezésű növényszám kísérletben, négy ismétlésben

Ismétlés	Hibrid	Növ. szám	Termés	Ismétlés	Termés	Ismétlés	Termés	Ismétlés	Termés
1	1	1	3.88	2	4.88	3	4.56	4	3.32
1	1	2	6.31	2	4.65	3	4.55	4	5.3
1	1	3	6.72	2	5.77	3	4.65	4	5.34
1	1	4	8.65	2	6.12	3	5.52	4	7.01
1	1	5	7.89	2	7.05	3	6.07	4	7.74
1	1	6	8.01	2	6.17	3	5.96	4	8.35
1	1	7	6.81	2	7.89	3	6.72	4	6.56
1	1	8	6.71	2	7.8	3	6.69	4	6.17
1	1	9	6.42	2	5.34	3	5.43	4	7.32
1	2	1	4.87	2	5.54	3	5.54	4	3.45
1	2	2	5.51	2	6.26	3	6.22	4	7.55
1	2	3	7.94	2	7.64	3	6.56	4	7.39
1	2	4	9.74	2	7.69	3	8.1	4	8.16
1	2	5	8.74	2	7.28	3	6.6	4	8.87
1	2	6	6.89	2	7.02	3	5.26	4	8.71
1	2	7	5.61	2	6.54	3	6.13	4	6.62
1	2	8	5.11	2	6.94	3	6.21	4	6.5
1	2	9	3.82	2	4.92	3	4.64	4	6.45
1	3	1	5.7	2	4.91	3	4.82	4	4.34
1	3	2	5.86	2	5.94	3	5.48	4	4.44
1	3	3	7.67	2	6.4	3	5.28	4	5.14
1	3	4	8.23	2	6.52	3	5.29	4	5.5
1	3	5	8.5	2	6.47	3	6.01	4	7.19
1	3	6	8.41	2	6.04	3	5.86	4	7.37
1	3	7	7.61	2	7.62	3	7.48	4	7.39
1	3	8	7.05	2	7.43	3	6.51	4	7.86
1	3	9	7.22	2	5.07	3	6.41	4	7.57
1	4	1	5.71	2	7.53	3	4.09	4	5.66
1	4	2	7.7	2	7.43	3	7.13	4	6.41
1	4	3	8.87	2	7.86	3	6.31	4	8.67
1	4	4	9.35	2	8.07	3	8.02	4	8.09
1	4	5	10.16	2	9.04	3	7.38	4	9.58
1	4	6	9.5	2	7.98	3	8.38	4	8.92
1	4	7	8.57	2	9.07	3	7.13	4	9.54

A 11.23. táblázat folytatása

Ismétlés	Hibrid	Növ. szám	Termés	Ismétlés	Termés	Ismétlés	Termés	Ismétlés	Termés
1	4	8	9.05	2	9.46	3	6.32	4	8.0
1	4	9	8.43	2	8.19	3	6.28	4	9.53
1	5	1	5.19	2	4.4	3	4.71	4	4.0
1	5	2	6.72	2	5.97	3	5.74	4	5.21
1	5	3	7.01	2	7.89	3	7.76	4	4.78
1	5	4	8.97	2	9.34	3	7.56	4	8.33
1	5	5	7.69	2	9.58	3	8.1	4	8.83
1	5	6	8.72	2	9.44	3	7.89	4	9.29
1	5	7	8.87	2	9.56	3	7.55	4	7.14
1	5	8	7.04	2	9.18	3	8.49	4	6.56
1	5	9	5.46	2	6.97	3	8.52	4	7.13
1	6	1	8.42	2	8.51	3	6.81	4	8.13
1	6	2	6.5	2	9.17	3	7.89	4	6.73
1	6	3	10.07	2	8.49	3	7.36	4	9.47
1	6	4	8.19	2	9.47	3	8.66	4	8.24
1	6	5	8.83	2	7.27	3	9.43	4	9.71
1	6	6	9.53	2	8.06	3	7.19	4	10.49
1	6	7	9.4	2	10.1	3	8.67	4	9.5
1	6	8	8.31	2	11.54	3	8.87	4	8.54
1	6	9	7.15	2	8.36	3	9.97	4	10.09

1. lépés. Szerkesszük meg a sávos parcellás elrendezésű kísérlet variancia táblázatát:

Variáció forrása	Szabadságfok (d.f.)	Eltérés-négyzet összeg (SS)	Közepes négyzetes eltérés (MS)	Számított F	Táblázati F	
					5%	1%
Ismétlés	$r-1 = 3$					
Horizontális faktor (A)	$a-1 = 5$					
Hiba (a)	$(r-1)(a-1) = 15$					
Vertikális faktor (B)	$b-1 = 8$					
Hiba (b)	$(r-1)(b-1) = 24$					
A x B	$(a-1)(b-1) = 40$					
Hiba (c)	$(r-1)(a-1)(b-1) = 120$					
Összes	$rab-1 = 215$					

2. lépés. Szerkesszük meg a három összegező táblázatot:

1. Ismétlés \times horizontális faktor összegező táblázat az ismétlés összegekkel, a horizontális faktor összegekkel és főösszeggel. Példánkban, az ismétlés \times fajta táblázat összegeket (RA), az ismétlés összegekkel (R), a fajta összegekkel (A) és a főösszeggel (G) a 11.24. táblázat mutatja.

11.24. táblázat. A termésösszegek ismétlés \times hibrid táblázata az 1. táblázat adataiból számolva

Hibrid	Termésösszeg (RA)				Hibrid összeg (A)
	I. ismétlés	II. ismétlés	III. ismétlés	IV. ismétlés	
H1	61.4	55.67	50.15	57.11	224.33
H2	58.23	59.83	55.26	63.7	237.02
H3	66.25	56.4	53.14	56.8	232.59
H4	77.34	74.63	61.04	74.4	287.41
H5	65.67	72.33	66.32	61.27	265.59
H6	76.4	80.97	74.85	80.9	313.12
Ismétlésösszeg (R)	405.29	399.83	360.76	394.18	
Főösszeg (G)					1560.06

2. Az ismétlés \times vertikális faktor táblázat összegeket kiszámítjuk a vertikális faktor összegekkel. Példánkban, az ismétlés \times növényszám táblázat összegeket (RB) a növényszám összegekkel (B) a 11. 25. táblázat tartalmazza.

11.25. táblázat. A termésösszegek ismétlés \times növényszám táblázata a 11.23. táblázat adataiból számítva

Növényszám	Termés összeg (RB)				Növényszám összeg (B)
	I. ismétlés	II. ismétlés	III. ismétlés	IV. ismétlés	
D ₁	33.77	35.77	30.53	28.9	128.97
D ₂	38.6	39.42	37.01	35.64	150.67
D ₃	48.28	44.05	37.92	40.79	171.04
D ₄	53.13	47.21	43.15	45.33	188.82
D ₅	51.81	46.69	43.59	51.92	194.01
D ₆	51.06	44.71	40.54	53.13	189.44
D ₇	46.87	50.78	43.68	46.75	188.08
D ₈	43.27	52.35	43.09	43.63	182.34
D ₉	38.5	38.85	41.25	48.09	166.69

3. A horizontális faktor \times vertikális faktor táblázat összegek. Példánkban a hibrid \times növényszám táblázat összegeket (AB) a 11.26. táblázat mutatja.

11.26. táblázat. A termésösszegek hibrid \times növényszám táblázata az 11.23. táblázat adataiból számítva

Hibrid	Termés összeg (AB)								
	D ₁	D ₂	D ₃	D ₄	D ₅	D ₆	D ₇	D ₈	D ₉
H ₁	16.4	20.81	22.48	27.33	28.75	28.49	27.98	27.37	24.51
H ₂	19.4	25.54	29.53	33.69	31.49	27.88	24.9	24.76	19.83
H ₃	19.77	21.72	24.49	25.54	28.17	27.68	30.1	28.85	26.27
H ₄	22.99	28.67	31.71	33.53	36.16	34.78	34.31	32.83	32.43
H ₅	18.3	23.64	27.44	34.2	34.2	35.34	33.12	31.27	28.08
H ₆	31.87	30.29	35.39	34.56	35.24	35.27	37.67	37.26	35.57

3. lépés. Számítsuk ki a korrekciós faktort és az összes eltérésnégyzet összegeket (SS):

$$C.F. = \frac{G^2}{rab} = \frac{(1560.06)^2}{(4)(9)(6)} = 11\,267.53$$

$$\text{Összes SS} = \sum X^2 - C.F. = \left[(3.88)^2 + \dots + (10.09)^2 \right] - 11\,267.53 = 11\,818 - 11\,267.53 = 550.47$$

4. lépés. Számítsuk ki az eltérésnégyzet összegeket a horizontális faktorra:

$$\begin{aligned} \text{Ismétlés SS} &= \frac{\sum R^2}{ab} - C.F. = \frac{(405.29)^2 + (399.83)^2 + (360.76)^2 + (394.18)^2}{(9)(6)} - 11\,267.53 = \\ &= 11\,289.81 - 11\,267.53 = 22.284 \end{aligned}$$

$$A \text{ (hibrid) SS} = \frac{\sum A^2}{rb} - C.F. = \frac{(224.33)^2 + \dots + (313.12)^2}{(4)(9)} - 11\,267.53 = 171.0$$

$$\begin{aligned} \text{Hiba (a) SS} &= \frac{\sum (RA)^2}{b} - C.F. - \text{Ismétlés SS} - A \text{ SS} = \frac{(61.4)^2 + \dots + (80.9)^2}{9} - 11\,267.53 - \\ &22.284 - 171.0 = 27.74 \end{aligned}$$

5. lépés. Számítsuk ki az eltérésnégyzet összegeket (SS) a vertikális analízishez:

$$\begin{aligned} B(\text{növényszám}) \text{ SS} &= \frac{\sum B^2}{ra} - C.F. = \frac{(128.97)^2 + \dots + (166.69)^2}{(4)(6)} - 11\,267.53 = 11\,424.041 - \\ &11\,267.53 = 156.51 \end{aligned}$$

$$\begin{aligned} \text{Hiba (b) SS} &= \frac{\sum (RB)^2}{a} - C.F. - \text{Ismétlés SS} - B \text{ SS} = \frac{(33.77)^2 + \dots + (48.09)^2}{6} - 11\,267.53 - \\ &- 22.284 - 156.51 = 52.51 \end{aligned}$$

6. lépés. Számítsuk ki az eltérésnégyzet összegeket (SS) az interakció analízishez:

$$\begin{aligned} A \times B \text{ (fajta} \times \text{nitrogén) SS} &= \frac{\sum (AB)^2}{r} - C.F. - A \text{ SS} - B \text{ SS} = \frac{(10715)^2 + \dots + (7476)^2}{3} - \\ &11\,267.53 - 156.51 - 171.0 = 56.46 \end{aligned}$$

$$\begin{aligned} \text{Hiba (c) SS} &= \text{Összes SS} - (\text{az összes többi SS összege}) = 550.47 - (22.248 + 156.51 + 52.51 + \\ &171.0 + 27.74 + 56.46) = 63.97 \end{aligned}$$

7. lépés. Számítsuk ki a közepes négyzetes eltéréseket (MS) minden variáció forrásra, úgy, hogy elosztjuk az SS értékeket a megfelelő d.f. értékekkel:

$$\text{Ismétlés MS} = \frac{22.28}{3} = 7.43$$

$$A \text{ MS} = \frac{171.0}{5} = 34.2$$

$$\text{Hiba (a) MS} = \frac{27.74}{15} = 1.85$$

$$\text{B MS} = \frac{156.51}{8} = 19.56$$

$$\text{Hiba (b) MS} = \frac{52.51}{24} = 2.19$$

$$\text{A x B MS} = \frac{56.46}{40} = 1.41$$

$$\text{Hiba (c) MS} = \frac{63.97}{120} = 0.53$$

8. lépés. Számítsuk ki az F-értékeket:

$$F(A) = \frac{A \text{ MS}}{\text{Hiba (a) MS}}$$

$$F(B) = \frac{B \text{ MS}}{\text{Hiba (b) MS}}$$

$$F(A \times B) = \frac{A \times B \text{ MS}}{\text{Hiba (c) MS}}$$

$$F(A) = \frac{34.2}{1.85} = 18.49$$

$$F(B) = \frac{19.56}{2.19} = 8.93$$

$$F(A \times B) = \frac{1.41}{0.53} = 2.66$$

9. lépés. Mindegyik számított F-értékre, melynek értéke nem kisebb 1-nél, keressük ki a megfelelő táblázati F-értéket f_1 = a számláló MS d.f. értéke és f_2 = a nevező MS d.f. értéke, a kívánt szignifikancia szinten.

10. lépés. Számítsuk ki a három variációs koefficiens a három hiba MS-nek megfelelően:

$$cv(a) = \frac{\sqrt{\text{Hiba(a)MS}}}{\text{Főátlag}} \times 100$$

$$cv(b) = \frac{\sqrt{\text{Hiba(b)MS}}}{\text{Főátlag}} \times 100$$

$$cv(c) = \frac{\sqrt{\text{Hiba(c)MS}}}{\text{Főátlag}} \times 100$$

A cv (a) jelzi a pontosság mértékét a horizontális faktoral összefüggésben, cv (b) a vertikális faktoral és cv (c) a két faktor közötti interakcióval összefüggésben. A cv (c) értéke várhatóan a legkisebb és ezáltal az interakció-hatás mérésének pontossága a legnagyobb. A cv (a)-ra és a cv (b)-re azonban nincs alapunk azt várni, hogy egyik kisebb vagy nagyobb legyen, mint a másik.

$$cv (a) = \frac{\sqrt{1.85}}{7.22} \times 100 = 18.84\%$$

$$cv (b) = \frac{\sqrt{2.19}}{7.22} \times 100 = 20.5\%$$

$$cv (c) = \frac{\sqrt{0.53}}{7.22} \times 100 = 10.08\%$$

11. lépés. A 3–10. lépésben kiszámított értékeket írjuk be a variancia táblázatba (11.27. táblázat). Hasonlítsuk össze a számított F-értéket a megfelelő táblázati F-értékkel és jelöljük meg a szignifikancia szinteket csillagokkal. Példánkban, mindhárom F-érték, azaz a hibrid fő hatása, a növényszám fő hatása és a hibrid × növényszám interakció egyaránt szignifikáns. Szignifikáns interakció esetén megfelelő óvatosság szükséges az eredmények értelmezésekor.

A varianciaanalízis erősen szignifikáns interakciót mutat a hibrid és a növényszám között, jelezve, hogy a hibrid hatások változnak a növényzámtól függően. Ezáltal az összehasonlítás a növényszám átlagok között minden hibrid átlagán keresztül vagy a hibrid átlagok között minden növényszám átlagán keresztül nem hasznos. A legmegfelelőbb átlag összehasonlítás a hibrid átlagok között, ugyanannál a növényzámnál vagy ugyanazon hibrid növényszám átlagai között.

11.27. táblázat. A 3 × 3-as sávos strip-plot elrendezésű faktoriális kísérlet adatainak varianciaanalízise^a

Variáció forrása	Szabadságfokok (d.f.)	Eltérés négyzet összeg (SS)	Közepes négyzetes eltérés (MS)	Számított F-érték ^b	Táblázati F-érték	
					5%	1%
Ismétlés	3	22.28	7.43			
Hibrid (A)	5	171.0	34.2	18.49***	2.90	4.56
Hiba (a)	15	27.74	1.85			
Növényszám (B)	8	156.51	19.56	8.93***	2.36	3.36
Hiba (b)	24	52.51	2.19			
A × B	40	56.46	1.41	2.66***	1.49	1.75
Hiba (c)	120	63.97	0.53			
Összes	215	550.47				

^acv(a) = 18.84%, cv (b) = 20.5% cv (c) = 10.08%.

^b***= szignifikáns 0.1%-os szinten

A szignifikáns különbség (SzD) kiszámítása

Az osztott parcellás elrendezéshez hasonlóan, a sávok parcellás elrendezésénél a kezelés-pár összehasonlításának négy típusa van, mindegyik szükségessé teszi a saját SzD érték kiszámítását. Ez a négy típus és a megfelelő formulák a vonatkozó s_d értékek kiszámításához, a 11.28. táblázatban található. A fenti példában, mivel az interakció hatás a hibrid és a növényszám között szignifikáns (11.27. táblázat), az egyetlen megfelelő típusú pár összehasonlítás az a hibridek között ugyanannál a növényzámmal.

11. 28. táblázat. Az átlagos különbség szórása (s_d) a pár összehasonlítás négy típusára sávok parcella elrendezésénél (Gomez és Gomez, 1984)

Sorszám	Pár összehasonlítás típusa	s_d^a
1.	Két horizontális átlag között (az összes vertikális kezelés átlagában)	$\sqrt{\frac{2E_a}{rb}}$
2.	Két vertikális átlag között (az összes horizontális kezelés átlagában)	$\sqrt{\frac{2E_b}{ra}}$
3.	Két horizontális átlag között a vertikális faktor azonos szintjén	$\sqrt{\frac{2[(b-1)E_c + E_a]}{rb}}$
4.	Két vertikális átlag között a horizontális faktor azonos szintjén	$\sqrt{\frac{2[(a-1)E_c + E_b]}{ra}}$

^a E_a = hiba (a) MS, E_b = hiba (b) MS, E_c = hiba (c) MS, r = ismétlések száma, a = horizontális sáv faktor szintjei, b = vertikális sáv faktor szintjei.

1. lépés. Számítsuk ki az s_d értéket, követve a 11.28. táblázat formuláját a 3. típusú összehasonlításhoz:

$$s_d = \sqrt{\frac{2[(b-1)E_c + E_a]}{rb}} = \sqrt{\frac{2[(8)(0.53) + 1.85]}{(4)(9)}} = 0.5817 \text{ t ha}^{-1}$$

2. lépés. Mivel két hiba tag (E_a és E_c) van a fenti formulában, kiszámítjuk a súlyozott táblázati t -értéket:

A táblázati t -érték, amely megfelel az E_a -nak $n = 15$ d.f.-nél (azaz t_a) és a táblázati t -érték, amely megfelel az E_c -nek $n = 120$ d.f.-nél (azaz t_c) az 5%-os és 1%-os szignifikanciaszintnél:

$$t_a(0.05) = 2.131 \text{ és } t_a(0.01) = 2.947$$

$$t_c(0.05) = 1.980 \text{ és } t_c(0.01) = 2.617$$

Számítsuk ki a táblázati súlyozott t -értéket, követve a megfelelő formulát (azaz 2. formula) a súlyozott t értékek kiszámítását tartalmazó táblázatban:

$$t' = \frac{(b-1)E_c t_c + E_a t_a}{(b-1)E_c + E_a}$$

$$t'(0.05) = \frac{(8)(0.53)(1.98) + (1.85)(2.131)}{(8)(0.53) + 1.85} = 2.026$$

$$t'(0.01) = \frac{(8)(0.53)(2.617) + (1.85)(2.947)}{(8)(0.53) + 1.85} = 2.717$$

3. lépés. Számítsuk ki az SzD értékeket 5%-os és 1%-os szignifikancia szinten:

$$\text{SzD}_{0.05} = t'(0.05)(s_d)$$

$$\text{SzD}_{0.05} = t'(0.05)(s_d) = (2.026)(0.5817) = 1.1784 \text{ t ha}^{-1}$$

$$\text{SzD}_{0.01} = t'(0.01)(s_d) = (2.717)(0.5817) = 1.5806 \text{ t ha}^{-1}$$

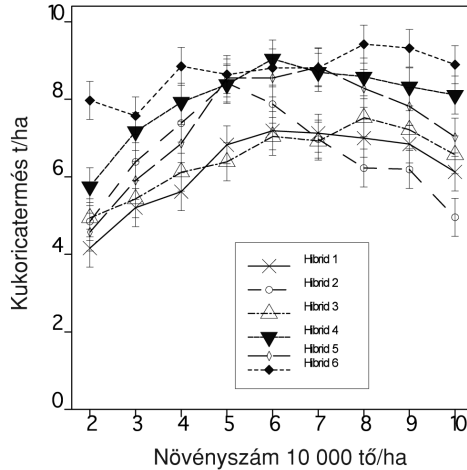
11.5.4. Az eredmények értelmezése

Szerkesszük meg az átlagok hibrid \times növényszám kétirányú táblázatát az SzD értékekkel, melyet a 11.29. táblázat mutat. Például, annak meghatározásához, hogy H_1 átlagos termése szignifikánsan különbözik-e a H_6 termésétől $60 \cdot 10^3$ tó ha^{-1} növény számnál, az átlagos különbségeket ($8.810 - 7.188 = 1.622$) összehasonlítjuk a kiszámított SzD értékekkel, azaz 1.1784 t ha^{-1} -ral és 1.5806 t ha^{-1} -ral. Mivel az átlagos különbség nagyobb, mint az SzD érték 5%-os és 1%-os szignifikancia szinten, a H_6 és H_1 átlagtermése szignifikánsan különbözik a $60 \cdot 10^3$ tó ha^{-1} növény számnál. A növény szám \times hibrid interakció hatását a 11.16. ábra szemlélteti.

11.29. táblázat. A hat kukorica hibrid átlagtermése (az ismétlések átlagában) kilenc eltérő növény számnál vizsgálva, sávos elrendezésű kísérletben

Hibrid	Kukorica szemtermése t ha^{-1}								
	D ₁	D ₂	D ₃	D ₄	D ₅	D ₆	D ₇	D ₈	D ₉
H ₁	4.160	5.202	5.620	6.825	7.188	7.122	6.995	6.842	6.127
H ₂	4.850	6.385	7.383	8.422	7.873	6.970	6.225	6.190	4.957
H ₃	4.942	5.430	6.123	6.385	7.043	6.920	7.525	7.213	6.567
H ₄	5.748	7.168	7.928	8.383	9.040	8.695	8.578	8.320	8.107
H ₅	4.575	5.910	6.860	8.550	8.550	8.835	8.280	7.817	7.020
H ₆	7.967	7.572	8.848	8.640	8.810	8.817	9.417	9.315	8.892

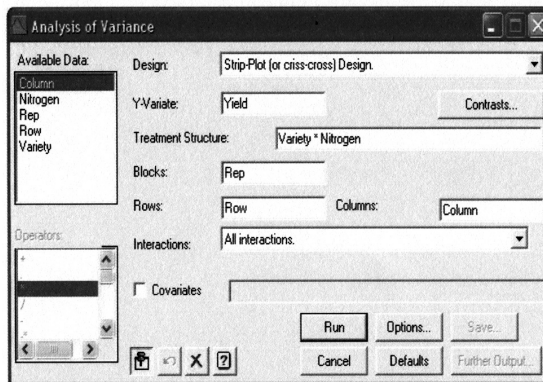
Az SzD értékek két hibrid összehasonlítására azonos növény számnál 1.1784 t ha^{-1} az 5%-os és 1.5806 t ha^{-1} 1%-os szignifikancia szinten.



11.16. ábra. Növényszám × hibrid interakció hatása a kukorica termésére

11.5.5. Varianciaanalízise GenStat programmal

A sávós (strip-plot vagy criss-cross) elrendezés megfelel az ismételt sor és oszlop elrendezésnek. Az egyik kezelés faktor (itt a hibrid) randomizálva van az oszlopokra minden ismétlésen belül, a másik kezelés faktor (itt a növényszám) randomizálva van a sorokra minden ismétlésen belül. Az ilyen blokk struktúra sorai és oszlopai keresztezik egymást (randomizáltak egyidejűleg és egymástól függetlenül) az ismétlésen belül, azaz: ismétlés/(sor × oszlop).



11.17. ábra. A sávós elrendezésű kísérlet varianciaanalízisének menü doboza

Az Analysis of Variance választási lehetőségei közül a General Analysis of Variance menüből a Strip-plot (or criss-cross) Design-t választjuk (11.17. ábra). Az adatfájlban külön oszlopok tartalmazzák a termésadatokat, a blokkokat, a sorokat és oszlopokat. Az Available Data doboz tartalmazza a megfelelő input adatokat. Y-variate: Termés, Treatment Struktüre: Növényszám × Hibrid, Blocks: Ismétlés, Rows: Növényszám, Columns: Hibrid.

Analysis of variance

Variate: Termes

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Ismetles stratum	3	22.5326	7.5109		
Ismetles.Hibrid stratum					
Hibrid	5	171.6904	34.3381	18.37	<.001
Residual	15	28.0373	1.8692	3.47	
Ismetles.Novszam stratum					
Novszam	8	156.8656	19.6082	9.07	<.001
Residual	24	51.8753	2.1615	4.01	
Ismetles.Hibrid.Novszam stratum					
Hibrid.Novszam	40	56.3529	1.4088	2.61	<.001
Residual	120	64.7310	0.5394		
Total	215	552.0852			

Tables of means

Variate: Termes

Grand mean 7.225

Hibrid	1	2	3	4	5	6	
	6.231	6.584	6.461	7.996	7.377	8.698	
Novszam	1	2	3	4	5	6	7
	5.374	6.278	7.127	7.868	8.084	7.893	7.837
Novszam	8	9					
	7.616	6.945					
Hibrid Novszam	1	2	3	4	5	6	
1	4.160	5.202	5.620	6.825	7.188	7.122	
2	4.850	6.385	7.383	8.422	7.873	6.970	
3	4.942	5.430	6.123	6.385	7.043	6.920	
4	5.748	7.168	7.928	8.383	9.040	8.695	
5	4.575	5.910	6.860	8.550	8.550	8.835	
6	7.967	7.572	8.848	8.640	8.810	8.817	
Hibrid Novszam	7	8	9				
1	6.995	6.842	6.127				
2	6.225	6.190	4.957				
3	7.525	7.213	6.567				
4	8.578	8.320	8.107				
5	8.280	7.817	7.020				
6	9.417	9.315	8.892				

Standard errors of differences of means

Table	Hibrid	Novszam	Hibrid Novszam
rep.	36	24	4
s.e.d.	0.3222	0.4244	0.6919
d.f.	15	24	96.03

Except when comparing means with the same level(s) of

Hibrid	0.6363
d.f.	92.47
Novszam	0.5862
d.f.	98.55

Least significant differences of means (5% level)

Table	Hibrid	Novszam	Hibrid Novszam
rep.	36	24	4
l.s.d.	0.6868	0.8759	1.3735
d.f.	15	24	96.03

Except when comparing means with the same level(s) of

Hibrid	1.2637
d.f.	92.47
Novszam	1.1631
d.f.	98.55

Least significant differences of means (1% level)

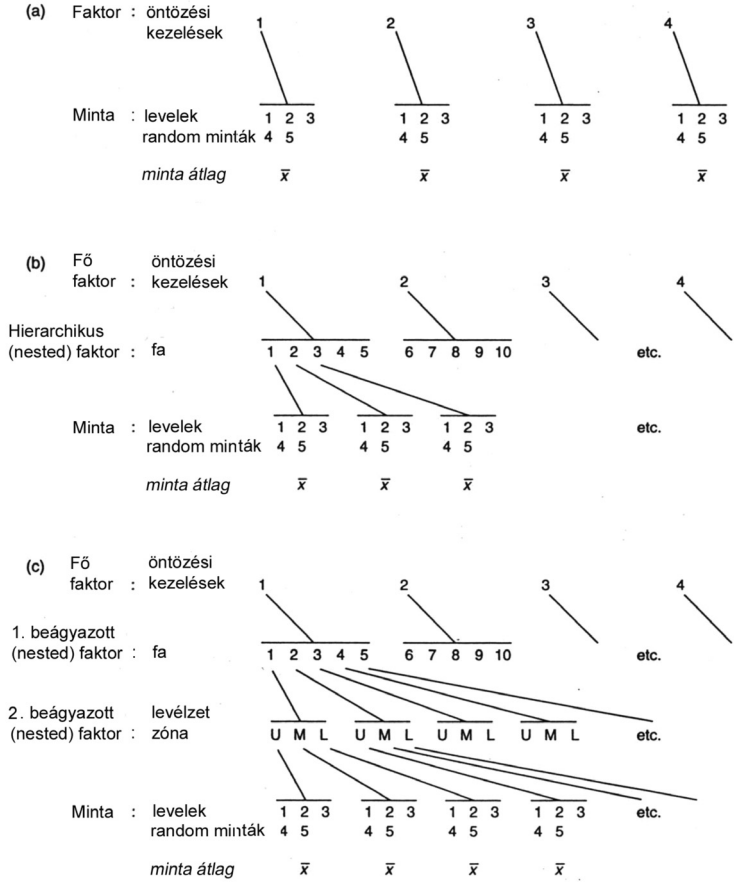
Table	Hibrid	Novszam	Hibrid Novszam
rep.	36	24	4
l.s.d.	0.9496	1.1870	1.8184
d.f.	15	24	96.03

Except when comparing means with the same level(s) of

Hibrid	1.6735
d.f.	92.47
Novszam	1.5396
d.f.	98.55

11.6. Hierarchikus vagy beágyazott (nested) kísérletek

Az eddig tárgyalt összes faktoriális kísérletben mindenegyes ismétlés a kísérletben egyenlő rangú és hozzájárul egyetlen méréshez. Például, egy kísérletben, hogy megvizsgáljuk az öntözés hatását a fák vízellátottságára, egy adott számú levél begyűjthető véletlenszerűen a fák eltérően öntözött parcelláiból és mérjük a vízellátottságukat. Ezután meghatározzuk a minta átlagok sorozatát, és ha a különbségek elég nagyok, akkor levonhatjuk azt a következtetést, hogy az öntözés szignifikáns hatással van a levél vízellátottságára. Nem leszünk képesek azonban semmit sem mondani a fákról, amelyekről a leveleket begyűjtöttük. Alternatív módon, random kiválasztunk egy adott számú fát minden parcellában és adott számú levél mintát veszünk (alminta) a fák mindegyikéről. Most lehetőség nyílik arra, hogy megvizsgáljuk a variációt a fák között a parcellán belül, valamint a különbségeket a kezelt parcellák között. Az elrendezésnek ezt a típusát beágyazott (nested) vagy hierarchikus elrendezésnek hívjuk, mivel egy faktor, ebben az esetben a különböző fák beágyazottak egy másik faktoron belül, ebben az esetben az öntözésen belül.



11.18. ábra. A randomizált faktoriális elrendezés és a hierarchikus (beágyazott) faktoriális kísérleti elrendezés összehasonlítása. A kísérletben az öntözés hatását vizsgáljuk a levél vízellátottságára különböző fákon, felhasználva (a) teljesen randomizált elrendezést; (b) egytényezős hierarchikus elrendezést; (c) két hierarchikus faktor elrendezést (Ireland, 2010)

A hierarchikus tervezés fő sajátossága, amely megkülönbözteti a normális faktoriális tervezéstől, hogy a beágyazott faktor kategóriái a fő faktor minden szintjén különbözőek. Példánkban, különböző fák adják a levélmintákat a fő öntözési kezeléseken mindegyikén belül. Ez lényeges különbség a normál faktoriális tervezés és a hierarchikus tervezés között a 11.18. ábrán látható.

A beágyazott elrendezések alkalmazhatók a bonyolultabb töbfőtenyezős kísérletekre is. Például, a fa öntözési kísérletben érdekes lehet meghatározni, hogy a levélzet pozíciója (elhelyezkedés) milyen mértékben befolyásolja a levél víztartalmát. Ebben az esetben a fa levélzete (lombozata) minden öntözési parcellán belül felosztható felső, középső és alsó zónára. A leveleknek egy adott száma ezután begyűjthető mindegyik levélzónából minden fán belül. Ezt illusztrálja a 11.18. ábra.

A hierarchikus elrendezés varianciaanalízise

A hierarchikus elrendezés statisztikai analízisében szükséges tesztelni a null hipotézist a fő faktor hatására mindegyik beágyazott faktorra vonatkozóan. Az ANOVA-nak ezáltal számba kell venni a variációt a fő csoportok (azaz a fő faktor) között és az alcsoportok (azaz a beágyazott faktorok) között, valamint a variációt a mintákon belül (azaz reziduális hatások). Így, a fa öntözési kísérletben szükséges meghatároznunk az eltérésnégyzet összegeket (SS) és közepes négyzetes eltéréseket (MS) az öntözés kezelésre, a fákra, amelyek beágyazottak az öntözés kezelésekre és a levelekre a mintákon belül. A számítási eljárást az alábbi példa mutatja.

Példa. Az öntözés hatása a fák leveleinek vízellátására.

Négy öntözési kezelést (A,B,C,D) alkalmaztak a fák egy mintájára. A levél víztartalmát határozták meg öt random kiválasztott levél alcsoportban az öt fa mindegyikéről, amelyek beágyazva voltak (nested) mindegyik öntözési kezelésbe. (A kísérleti tervet az 11.18. ábra mutatja.) A táblázat mutatja az összes és átlagos levél víztartalmát a levelek minden almintájára és minden öntözési kezelésre.

Öntözés kezelés	Levél víztartalma (g/g levél zöldtömeg) (alcsoport összegek és átlagok, öt levelet foglalnak magukban minden fáról)					Öntözés összes	Öntözés átlag	
			Fa					
				Fa				
A		1	2	3	4	5		
	Összes	4.04	3.77	4.25	4.13	4.21	20.40	0.816
	Átlag	0.808	0.754	0.85	0.826	0.842		
				Fa				
B		6	7	8	9	10		
	Összes	4.24	3.94	4.37	3.92	4.26	20.73	0.830
	Átlag	0.848	0.788	0.874	0.784	0.852		
				Fa				
C		11	12	13	14	15		

A táblázat folytatása

Öntözés kezelés		Levél víztartalma (g/g levél zöldtömeg) (alcsoport összegek és átlagok, öt levelet foglalnak magukban minden fáról)					Öntözés összes	Öntözés átlag
		Összes	3.54	3.25	3.47	3.54		
	Átlag	0.708	0.65	0.694	0.708	0.736		
				Fa				
D		16	17	18	19	20		
	Összes	3.27	3.25	3.29	3.14	3.39	16.34	0.654
	Átlag	0.654	0.65	0.658	0.628	0.678		

$$\Sigma x = 74.95$$

$$\Sigma x^2 = 56.9513$$

$N = a$ levelek összes száma = 100 $n =$ levelek száma alcsoportonként = 5

$N_{\text{öntözés}} =$ öntözés kezelések száma = 4 $n_{\text{öntözés}} =$ levelek száma öntözési
kezelésenként = 25

$N_{\text{alcsoport}} =$ alcsoportok (fák) száma = 20

Null hipotézis: különböző mennyiségű öntöző vízzel kezelt, különböző fák leveleinek a víztartalma ugyanaz.

Analízis: a korrekciós tagot (CT) és az összes eltérésnégyzet összeget a szokásos módon határozzuk meg:

$$CT = (\Sigma x)^2 / N = 74.95^2 / 100 = 56.1750$$

$$SS_{\text{összes}} = \Sigma x^2 - CT = 56.9513 - 56.1750 = 0.7763$$

$$DF = N - 1 = 99$$

Az eltérésnégyzet összegeket minden alcsoportra, amelyek tartalmazzák a főfaktor (öntözés) hatást és a „nested” faktor (fák) hatást, a következőképpen határozzuk meg:

$$SS_{\text{alcsoport}} = \Sigma T_{\text{alcsoport}}^2 / n - CT = (4.04^2 + 3.77^2 + \dots + 3.39^2) / 5 - 56.1750 = 0.6515$$

$$DF = N_{\text{alcsoport}} - 1 = 19$$

$$SS_{\text{öntözés}} = T_{\text{öntözés}}^2 / n_{\text{öntözés}} - CT = (20.40^2 + 20.73^2 + 17.48^2 + 16.34^2) / 25 - 56.175 = 0.5625$$

$$DF = N_{\text{alcsoport}} - 1 = 3$$

Az eltérésnégyzet összeget a beágyazott (nested) faktorra és a maradék (residuális) hatásokra kivonással számítjuk ki:

$$SS_{\text{fák}} = SS_{\text{alcsoportok}} - SS_{\text{öntözés}} = 0.6510 - 0.5625 = 0.0885$$

$$DF = DF_{\text{alcsoportok}} - DF_{\text{öntözés}} = 16$$

$$SS_{\text{maradék}} = SS_{\text{összes}} - SS_{\text{alcsoportok}} = 0.7763 - 0.6510 = 0.1252$$

$$DF = DF_{\text{összes}} - DF_{\text{alcsoportok}} = 80$$

Az ANOVA-ban a közepes négyzetes eltérés (MS) minden egyes nested (beágyazott) szinten becsüli a variabilitást erre a szintre, plussz minden ez alatti szintre. Az F-értékeket ezáltal úgy határozzuk meg, hogy összehasonlítjuk az MS-t minden faktorra a közvetlenül alatta lévő faktor MS-sel a nested (beágyazott) hierarchiában.

ANOVA táblázat

Variáció forrása	SS	DF	MS	F-érték	F _{kritikus}		Szignifikancia
					α = 0.05	α = 0.01	
Minden alcsoport	0.6510	19					
Fő faktor (öntözés)	0.5625	3	0.1875	$\frac{0.1875}{0.0055} = 34.01$	3.24	5.29	$P < 0.01$
Nested faktor (fák)	0.0885	16	0.0055	$\frac{0.0055}{0.0016} = 3.44$	1.77	2.24	$P < 0.01$
Maradék	0.1252	80	0.0016				
Összes	0.7763	99					

Következtetés: szignifikáns különbség van az öntözés kezelések között ($P < 0.01$) és a fák között a kezeléseken belül ($P < 0.01$). Ezáltal, miközben az öntözés kezelések szignifikáns hatását állapítottuk meg, a vizsgálat ugyancsak jelzi, hogy a fák nem egy-ségesek (uniform) és egymástól szignifikánsan eltérő reakciót adnak.

Az ANOVA-ban a közepes négyzetes eltérés (MS) mindegyik beágyazott szinten belül becsüli a variabilitást erre a szintre, plusz minden szintre ez alatt a hierarchiában. Az F-értékeket ezáltal úgy határozzuk meg, hogy összehasonlítjuk az MS-t minden faktorra a közvetlen alatta levő faktoriall, tehát:

$$\text{F-arány (fő faktor)} = \frac{MS_{\text{főfaktor}}}{MS_{\text{beágyazott faktor}}} \text{ és}$$

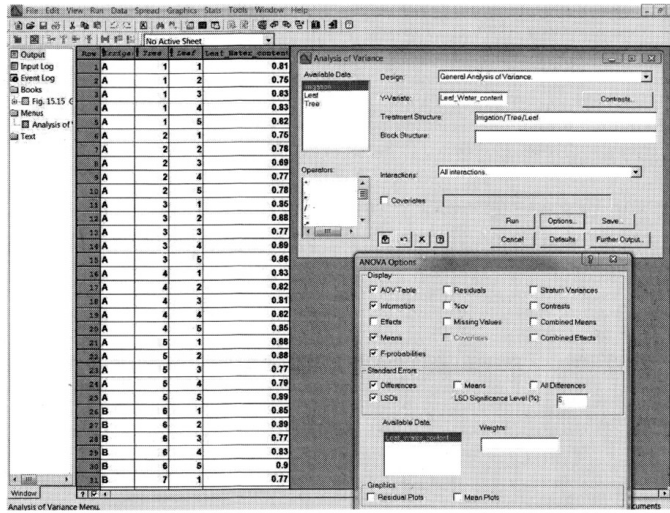
$$\text{F-arány (beágyazott faktor)} = \frac{MS_{\text{beágyazottfaktor}}}{MS_{\text{maradék}}}$$

Ezáltal az öntözési kísérletben, melyet az 11.18. ábra illusztrál, az F-értékeket a következőképpen kapjuk meg:

$$\text{F-arány (fő faktor)} = \frac{MS_{\text{öntözés}}}{MS_{\text{fák}}} \text{ és}$$

$$\text{F-arány (beágyazott faktor)} = \frac{MS_{\text{fák}}}{MS_{\text{maradék}}}$$

AGenStat számolótáblát és a menü dobozokat a hierarchikus ANOVA-hoz a 11.19. ábra mutatja.



(b) GenStat output window

Analysis of variance					
Variate: Leaf_Water_content					
Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Irrigation	3	0.5625310	0.1875103		
Irrigation.Tree	16	0.0885040	0.0055315		
Irrigation.Tree.Leaf	80	0.1252400	0.0015655		
Total	99	0.7762750			

Tables of means

Variate: Leaf_Water_content											
Grand mean 0.75											
Irrigation	A	B	C	D							
Irrigation	0.82	0.83	0.70	0.65							
Irrigation A	1	2	3	4	5	6					
Irrigation A B	0.81	0.75	0.85	0.83	0.83	0.85					
Irrigation B	7	8	9	10	11	12					
Irrigation B C	0.79	0.87	0.78	0.85	0.71	0.65					
Irrigation C	13	14	15	16	17	18					
Irrigation C D	0.69	0.71	0.74	0.65	0.65	0.66					
Irrigation D	19	20									
	0.63	0.68									
Standard errors of differences of means											
Table	Irrigation	Irrigation Tree	Irrigation Tree Leaf								
rep.	25	5	1								
d.f.	.	.	.								
s.e.d.	.	.	.								
Least significant differences of means (5% level)											
Table	Irrigation	Irrigation Tree	Irrigation Tree Leaf								
rep.	25	5	1								
d.f.	.	.	.								
l.s.d.	.	.	.								

11.19. ábra. (a) GenStat számológépi és menü dobozok a nested (hierarchikus ANOVA) elvégzéséhez egy fő faktorról, (b) GenStat kimeneti ablak

11.7. Kéttényezős kísérlet ismétlés nélkül

Természetesen szokatlan beállítani egy kéttényezős kísérletet anélkül, hogy magában foglalna ismétlést, amikor azonban korlátozott kísérleti anyag áll rendelkezésre, ez szükségessé válhat. Ismétlés nélkül kezelés interakció nem mutatható ki, ezáltal az ANOVA csupán a variáció három lehetséges forrására irányul, nevezetesen két kezelés faktorra, melyeket kezelés varianciaként mutatunk ki, és a természetes random variáció, melyet reziduális varianciaként mutatunk ki. Ahhoz, hogy lehetővé váljék mindkét variancia meghatározása, mindkét kezelésre az eltérésnégyzet összegeket elkülönítetten kell kiszámolni és a tipikus adattáblázat hasonló lesz ahhoz, amit a 11.30. táblázatban bemutatunk (Ireland, 2010).

11.30. táblázat. Az adattáblázat tipikus formája a kéttényezős faktoriális kísérlethez, ismétlés nélkül

Kezelés faktor B szint	Kezelés faktor A szint				Kezelés B összesen
	1	2	3	n	
1	$x_{1,1}$	$x_{1,2}$	$x_{1,3}$	$x_{1,n}$	T_{B1}
2	$x_{2,1}$	$x_{2,2}$	$x_{2,3}$	$x_{2,n}$	T_{B2}
3	$x_{3,1}$	$x_{3,2}$	$x_{3,3}$	$x_{3,n}$	T_{B3}
n	$x_{n,1}$	$x_{n,2}$	$x_{n,3}$	$x_{n,n}$	T_{Bn}
Kezelés A összesen	$T_{A,1}$	$T_{A,2}$	$T_{A,3}$	$T_{A,n}$	Σx (Főösszeg)

A következő tagokat határozzuk meg:

N_A = az A kezelés alkalmazásának szintjei

N_B = a B kezelés alkalmazásának szintjei

N = a megfigyelések összes száma (= $N_A \times N_B$)

A kéttényezős, ismétlés nélküli ANOVA kiszámításának folyamata a következő:

1. lépés: A korrekciós tag (CT) kiszámítása:

$$CT = \frac{\left(\sum X\right)^2}{N}$$

2. lépés: Az összes eltérésnégyzet összegek kiszámítása

$$SS_{\text{összes}} = \sum (X - \bar{X})^2 = \sum X^2 - CT$$

3. lépés: A kezelés eltérésnégyzet összeg kiszámítása

Ahol két kezelés faktort, A-t és B-t alkalmazunk, a kezelés eltérésnégyzet összegek kiszámítása:

$$SS_A = \sum (\bar{T}_A - GM)^2 = \left(\frac{\sum T_A^2}{n_A}\right) - CT \quad \text{és} \quad SS_B = \sum (\bar{T}_B - GM)^2 = \left(\frac{\sum T_B^2}{n_B}\right) - CT$$

ahol: $\sum T_A^2 = A$ kezelés eltérésnégyzet összegek

$\sum T_B^2 = B$ kezelés eltérésnégyzet összegek

$n_A = a$ megfigyelések száma az A kezelés minden szintjére ($=N_B$)

$n_B = a$ megfigyelések száma a B kezelés minden szintjére ($=N_A$)

4. lépés: A reziduális eltérésnégyzet összegek kiszámítása (azaz SS a random variációnak tulajdoníthatóan)

Mivel a kezelés SS és a reziduális SS összege az összes SS-t adja, a reziduális SS megkapható egyszerűen kivonással:

$$SS_{\text{reziduális}} = SS_{\text{összes}} - (SS_A + SS_B)$$

5. lépés: A szabadságfokok (DF) meghatározása a variáció minden forrására

Összes DF = $N - 1$ (a megfigyelések összes száma mínusz 1)

Kezelés A DF = $N_A - 1$ (a kezelés A szintjeinek összes száma mínusz 1)

Kezelés B DF = $N_B - 1$ (a kezelés B szintjeinek összes száma mínusz 1)

Reziduális DF = $(n_A - 1) \times (n_B - 1)$ (ez egyenlő az összes DF mínusz a két kezelés DF)

6. lépés: A közepes négyzetes eltérés (variancia) kiszámítása mindegyik variáció forrásra.

A közepes négyzetes eltérést mindegyik variáció forrásra úgy számítjuk ki, hogy elosztjuk mindegyik SS értékét a megfelelő szabadságfokkal:

$$\text{Kezelés A MS} = SS_{\text{kezelés A}} / DF_{\text{kezelés A}}$$

$$\text{Kezelés B MS} = SS_{\text{kezelés B}} / DF_{\text{kezelés B}}$$

$$\text{Reziduális MS} = SS_{\text{reziduális}} / DF_{\text{reziduális}}$$

7. lépés: Az F-mutató kiszámítása és az eredmények prezentálása az ANOVA táblázatban

A fix hatású ANOVA I Modellben az F-arányt megkapjuk minden kezelés faktorra úgy, hogy a kezelés közepes eltérésnégyzetet elosztjuk a reziduális közepes eltérésnégyzettel. Ha a II vagy III Modell ANOVA-t használjuk, akkor az F-értékeket minden kezelés faktorra úgy határozzuk meg, mint azt bemutattuk (standard számítási eljárás). A végső szakaszokat a számításban a kéttényezős ANOVA táblázat mutatja:

11.31. táblázat. Kéttényezős ANOVA (ismétlés nélkül) táblázat standard formája

Variáció forrása	Eltérésnégyzet összegek (SS)	Szabadság fokok (DF)	Közepes négyzet eltérések (MS)	F-arány ^a
Kezelés A faktor	$SS_{\text{kezelésA}}$	$N_A - 1$	$SS_{\text{kezelésA}} / DF_{\text{kezelésA}}$	$\frac{MS_{\text{kezelésA}}}{MS_{\text{maradék}}}$
Kezelés B faktor	$SS_{\text{kezelésB}}$	$N_B - 1$	$SS_{\text{kezelésB}} / DF_{\text{kezelésB}}$	$\frac{MS_{\text{kezelésB}}}{MS_{\text{maradék}}}$
Maradék (mintán belül)	$SS_{\text{maradék}}$	$(n_a - 1) \times (n_b - 1)$	$SS_{\text{maradék}} / DF_{\text{maradék}}$	
Összes	$SS_{\text{összes}}$	$N - 1$		

^a fix hatások ANOVA modell feltételezve

8. Lépés: Összehasonlítás az F kritikus értékével és következtetés levonása

Abból a célból, hogy mindegyik kezelés szignifikanciáját egymástól függetlenül becsüljük, a megfelelő F arány értékeket összehasonlítjuk a táblázati F mutató kritikus értékével a valószínűség kívánatos szintjén és a kezelés szabadságfoknál (V_1) és a maradék szabadságfoknál (V_2). Ha a megfigyelt F-arány bármelyik kezelés faktorra nagyobb vagy egyenlő a kritikus F-értékkel, akkor a null hipotézis visszautasítható és a kezelés faktor szignifikáns hatását igazolhatjuk.

12. Háromtényezős kísérletek varianciaanalízise

A háromtényezős faktoriális kísérletek használata sokkal kevésbé általános, mint az egy- és kéttényezős kísérleteké, mivel sokkal nagyobb a forrásigényük. Ahol háromtényezős faktoriális kísérletre van szükség, az analízis ANOVA-val elvileg a kéttényezős ANOVA logikus kiterjesztése és a számítás könnyen elvégezhető, ha rendelkezésre áll számítógépes program. Mivel azonban a három faktor között több interakció lehetséges, a nehézség általában az eredmények interpretálása és nem maga a számítási eljárás. Az interakciók száma világossá válik, amikor lineáris modellel írjuk le a háromtényezős faktoriális kísérletet. Az általános lineáris modell a következő:

$$y = \mu + \alpha + \beta + \gamma + (\alpha \times \beta) + (\alpha \times \gamma) + (\beta \times \gamma) + (\alpha \times \beta \times \gamma) + \varepsilon$$

ahol: y = a vizsgált változó mért értéke;

μ = a változó populációjának átlaga (főátlag);

α = az α kezelés faktor szintjének kvantitatív hatása a változóra;

β = a β kezelés faktor szintjének kvantitatív hatása a változóra;

γ = a γ kezelés faktor szintjének kvantitatív hatása a változóra;

$\alpha \times \beta \times \gamma$ = az α , β és γ kezelés faktorok interakciójának kvantitatív hatása a változóra;

ε = a random (hiba) hatások kvantitatív hatása a változóra.

12.1. táblázat. A háromtényezős ANOVA táblázat standard formája

Variáció forrása	Eltérésnégyzet összeg (SS)	Szabadságfokok (DF)	Közepes négyzetes eltérések (MS)	F-arány
A kezelés	$SS_{\text{kezelésA}}$	$DF_{\text{kezelésA}}$	$\frac{SS_{\text{kezelésA}}}{DF_{\text{kezelésA}}}$	$\frac{MS_{\text{kezelésA}}}{MS_{\text{maradék}}}$
B kezelés	$SS_{\text{kezelésB}}$	$DF_{\text{kezelésB}}$	$\frac{SS_{\text{kezelésB}}}{DF_{\text{kezelésB}}}$	$\frac{MS_{\text{kezelésB}}}{MS_{\text{maradék}}}$
C kezelés	$SS_{\text{kezelésC}}$	$DF_{\text{kezelésC}}$	$\frac{SS_{\text{kezelésC}}}{DF_{\text{kezelésC}}}$	$\frac{MS_{\text{kezelésC}}}{MS_{\text{maradék}}}$
AxB interakció	SS_{AxB}	DF_{AxB}	$\frac{SS_{\text{AxB}}}{DF_{\text{AxB}}}$	$\frac{MS_{\text{AxB}}}{MS_{\text{maradék}}}$
AxC interakció	SS_{AxC}	DF_{AxC}	$\frac{SS_{\text{AxC}}}{DF_{\text{AxC}}}$	$\frac{MS_{\text{AxC}}}{MS_{\text{maradék}}}$
BxC interakció	SS_{BxC}	DF_{BxC}	$\frac{SS_{\text{BxC}}}{DF_{\text{BxC}}}$	$\frac{MS_{\text{BxC}}}{MS_{\text{maradék}}}$
AxBxC interakció	SS_{AxBxC}	DF_{AxBxC}	$\frac{SS_{\text{AxBxC}}}{DF_{\text{AxBxC}}}$	$\frac{MS_{\text{AxBxC}}}{MS_{\text{maradék}}}$
Maradék	$SS_{\text{maradék}}$	$DF_{\text{maradék}}$	$\frac{SS_{\text{maradék}}}{DF_{\text{maradék}}}$	
Összes	$SS_{\text{összes}}$	$DF_{\text{összes}}$		

Következésképpen megnőtt a tesztelni kívánt null hipotézisek száma:

$H_{0(1)}$ = nincs hatása az A kezelésnek

$H_{0(2)}$ = nincs hatása a B kezelésnek

$H_{0(3)}$ = nincs hatása a C kezelésnek

$H_{0(4)}$ = nincs hatása az A \times B kezelés interakciónak

$H_{0(5)}$ = nincs hatása az A \times C kezelés interakciónak

$H_{0(6)}$ = nincs hatása a B \times C kezelés interakciónak

$H_{0(7)}$ = nincs hatása az A \times B \times C kezelés interakciónak

Az eltérésnégyzet-összegek felbontásának eljárása a háromtényezős ANOVA-nál követi ugyanazokat a lépéseket, mint a kéttényezős ANOVA-nál, jóllehet most négy és nem egy interakciót kell meghatározni. Az ANOVA Modell I esetében az F mutató mindenegyes fő kezelésre és minden interakcióra a kezelés vagy az interakció MS aránya lesz a reziduális (maradék) MS-hez. A végső ANOVA táblázat formáját a 12.1. táblázat mutatja. Az F-arányokat használjuk, hogy teszteljünk minden null hipotézist egymástól függetlenül a szokásos módon.

12.1. Háromtényezős teljesen véletlen (CRD) elrendezés

12.1.1. Kísérlet leírása

Háromtényezős faktoriális kísérletben vizsgálták három különböző mulcs típus és négy különböző öntözési szint hatását a nyírfacsemetek növekedésére három eltérő növény-sűrűségnél. A fák törzsének körméretét ezt követően határozták meg (Ireland, 2010). A kísérlet alapadatait az 12.2. táblázat tartalmazza, a variancia táblázat szerkezetét a 12.3. táblázat mutatja. A következőkben bemutatjuk a háromtényezős kísérlet varianciaanalízisét a GenStat programmal.

12.2. táblázat. A mulcs különböző típusainak hatása, különböző öntözési szintekkel és különböző növényszámokkal kombinálva, a fiatal nyírfák növekedésre

Növény- szám	Fák körmérete (cm)											
	Mulcs nélkül				Fakéreg mulcs				Plasztik mulcs			
	Öntözés (VK %)				Öntözés (VK %)				Öntözés (VK %)			
	0	33	66	100	0	33	66	100	0	33	66	100
1	15	16	21	21	18	18	23	21	20	19	20	21
	13	20	19	19	17	20	26	18	16	23	22	18
	17	19	23	14	18	27	27	16	16	26	25	9
	11	15	22	17	12	15	24	17	14	19	16	17
	12	21	26	18	20	24	28	20	22	20	26	18
\bar{x}	13.6	18.2	22.2	17.8	17.0	20.8	25.6	18.4	17.6	21.4	21.8	16.6
2	16	18	22	24	15	20	25	22	15	23	18	18
	19	20	26	18	22	23	26	15	22	18	21	14
	12	21	20	18	14	27	19	13	18	27	23	13
	14	21	21	22	15	24	27	20	19	26	20	20
	14	21	18	17	13	21	20	16	24	26	19	15
\bar{x}	15.0	20.2	21.4	19.8	15.8	23.0	23.4	17.2	19.6	24.0	20.2	16.0
5	10	13	16	15	10	17	19	14	11	19	16	14
	9	10	19	15	14	17	27	16	20	17	22	18
	13	14	20	15	23	24	24	12	22	20	21	11
	10	11	23	22	18	15	28	12	21	17	16	10
	9	15	16	16	18	16	22	12	23	17	20	12
\bar{x}	10.2	12.6	18.8	16.6	16.6	17.8	24.0	13.2	19.4	18.0	19.0	13.0

12.3. táblázat. A háromtényezős teljesen véletlen (CRD) elrendezésű kísérlet variancia táblázata

Variáció forrása	Szabadság- fokok	Eltérés-négyzet összeg (SS)	Közepes négyzetes eltérés (MS)	Számított F-érték	Táblázati F-érték	
					5%	1%
A faktor	a-1					
B faktor	b-1					
A × B	(a-1)(b-1)					
C faktor	c-1					
A × C	(a-1)(c-1)					
B × C	(b-1)(c-1)					
A × B × C	(a-1)(b-1)(c-1)					
Hiba	abc (r-1)					

12.1.2. Varianciaanalízis GenStat programmal

Az adatok GenStat számolótáblába történő bevitelekor a mulcs, az öntözés és a növénytűrűség képezik a faktorokat és a fák körmérete a változó. A varianciaanalízis műveleteinek sorrendje: Stats, Analysis of Variance, General Analysis of Variance. Y-variate: Stem girth, Treatment Structure: mulch × irrigation × plant density, Block structure: replication.

Az eredmények értelmezése

A GenStat programmal történt értékelés outputja négy részből áll. Először a variancia táblázat látható a szokásos szerkezetben, ezután az átlagértékek táblázatai következnek. A fő hatások szintjei szerinti értékek után a kölcsönhatások szerinti értékek láthatók. Az átlagértékek táblázatai után a különbségek szórásai (SED) és a legkisebb szignifikáns különbségek ($SzD_{5\%}$) következnek.

Megállapítható, hogy 0.1%-os szinten szignifikáns különbség van a mulcs kezelése, a az öntözési kezelése és a növényszám kezelése között. A kettős interakciók közül 0.1%-os szinten szignifikáns az öntözés × mulcs kölcsönhatás, illetve 5%-os szinten szignifikáns az öntözés × növényszám kölcsönhatás. A mulcs × növényszám és a hármas kölcsönhatás nem szignifikáns. Következésképpen, az öntözés, a mulcs és a növényszám közvetlen hatása a szignifikáns kettős interakciók miatt csak korlátozottan értelmezhető.

Analysis of variance

Variate: Stemgirth

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Irrigation	3	986.417	328.806	33.23	<.001
Mulcs	2	158.811	79.406	8.03	<.001
Density	2	327.411	163.706	16.55	<.001
Irrigation.Mulcs	6	464.567	77.428	7.83	<.001
Irrigation.Density	6	153.967	25.661	2.59	0.020
Mulcs.Density	4	41.722	10.431	1.05	0.382
Irrigation.Mulcs.Density					
	12	69.300	5.775	0.58	0.853
Residual	144	1424.800	9.894		
Total	179	3626.994			

Tables of means

Variate: Stemgirth

Grand mean 18.49

	1	2	3	4
Irrigation	16.09	19.56	21.82	16.51
Mulcs	17.20	19.40	18.88	

Density	1	2	3		
	19.25	19.63	16.60		
Irrigation	Mulcs	1	2	3	
1		12.93	16.47	18.87	
2		17.00	20.53	21.13	
3		20.80	24.33	20.33	
4		18.07	16.27	15.20	
Irrigation	Density	1	2	3	
1		16.07	16.80	15.40	
2		20.13	22.40	16.13	
3		23.20	21.67	20.60	
4		17.60	17.67	14.27	
Mulcs	Density	1	2	3	
1		17.95	19.10	14.55	
2		20.45	19.85	17.90	
3		19.35	19.95	17.35	
Irrigation	Mulcs	Density	1	2	3
1	1		13.60	15.00	10.20
	2		17.00	15.80	16.60
	3		17.60	19.60	19.40
2	1		18.20	20.20	12.60
	2		20.80	23.00	17.80
	3		21.40	24.00	18.00
3	1		22.20	21.40	18.80
	2		25.60	23.40	24.00
	3		21.80	20.20	19.00
4	1		17.80	19.80	16.60
	2		18.40	17.20	13.20
	3		16.60	16.00	13.00

Standard errors of differences of means

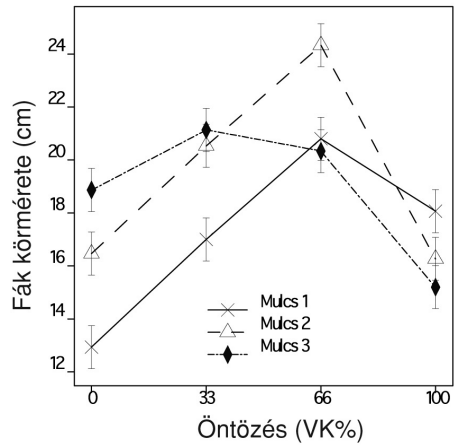
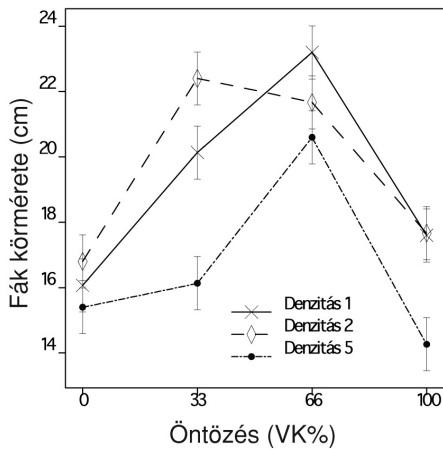
Table	Irrigation	Mulcs	Density	Irrigation Mulcs
rep.	45	60	60	15
d.f.	144	144	144	144
s.e.d.	0.663	0.574	0.574	1.149

Table	Irrigation Density	Mulcs Density	Irrigation Mulcs Density
rep.	15	20	5
d.f.	144	144	144
s.e.d.	1.149	0.995	1.989

Least significant differences of means (5% level)

Table	Irrigation	Mulcs	Density	Irrigation Mulcs
rep.	45	60	60	15
d.f.	144	144	144	144
l.s.d.	1.311	1.135	1.135	2.270

Table	Irrigation Density	Mulcs Density	Irrigation Mulcs Density
rep.	15	20	5
d.f.	144	144	144
l.s.d.	2.270	1.966	3.932



12.1. ábra. A fák körmérete az öntözés × növényszám (bal oldali ábra) és az öntözés × mulcs (jobb oldali ábra) interakciótól függően

12.2. Kétszeresen osztott parcellás (split-split-plot) elrendezés

A kétszeresen osztott parcellás elrendezés az osztott parcellás kísérlet kiterjesztése, hogy befogadjon egy harmadik faktort. Egyedülálló módon alkalmas a háromtényezős kísérletekre, ahol három eltérő szintű pontosság indokolt a különböző hatásokra. A pontosság minden szintjét megjelöljük a hatásokra a három faktor mindegyikével összefüggésben. Ezt a tervet két fontos sajátosság jellemzi:

1. Három parcella méret van a három faktornak megfelelően, nevezetesen, a legnagyobb parcella (főparcella) a főparcella faktorhoz, a közbelső méretű parcella (alparcella) az alparcella faktorra, és a legkisebb parcella (al-alparcella) az al-alparcella faktorra.

2. A pontosságnak három szintje van, a főparcella faktor kapja a legkisebb mértékű pontosságot és az al-alparcella a legnagyobb mértékű pontosságot.

Az eljárást a randomizációra, az elhelyezésre és a varianciaanalízisre egy $5 \times 3 \times 3$ -as faktoriális kísérlettel, három ismétlésben mutatjuk be. A kezeléseket képezi a nitrogén öt szintje mint főparcella, három agrotechnikai eljárás, mint alparcella és három búzafajta, mint al-alparcella. Az ismétlések számát r -rel jelöljük; A, B és C jelöli a főparcella faktort, az alparcella faktort és az al-alparcella faktort; míg a , b és c utal az A, B és C faktorok kezelésszintjeire.

12.2.1. Randomizáció és elrendezés

Három lépés van a kétszeresen osztott parcellás kísérlet randomizálásában és elrendezésében:

1. lépés. Osszuk fel a kísérleti területet r ismétlésre és mindegyik ismétlést a főparcellákra. Ezután random jelöljük ki az a főparcella kezeléseket az a főparcellákra, elkülönítetten és függetlenül az r ismétlés mindegyikére, követve bármelyik randomizációs sémát.

Példánkban, a területet felosztottuk három ismétlésre és mindegyik ismétlést öt főparcellára. Ezután az öt nitrogénszintet (N_1, N_2, N_3, N_4 és N_5) random elhelyeztük az öt főparcellára mindegyik ismétlésben. Az eredményt az 12.2. ábra mutatja.

2. lépés. Osszuk fel mindegyik főparcellát b alparcellára, és b alparcella kezeléseket random jelöljük ki, elkülönítetten és függetlenül az $(r)(a)$ főparcella mindegyikére.

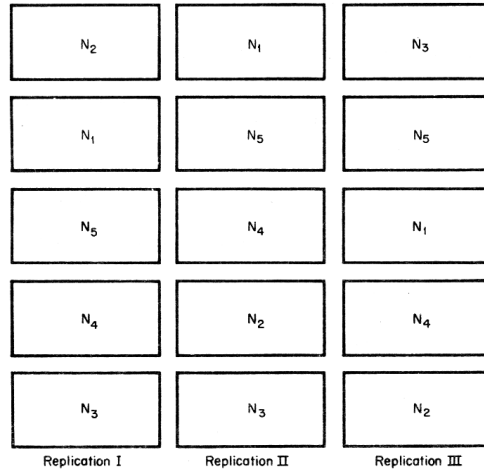
Példánkban, mindegyik főparcellát három alparcellára osztottuk fel, amelyekben a három agrotechnikai eljárást (M_1, M_2 és M_3) random jelöltük ki. A randomizáció folyamatot $(r)(a) = 15$ -ször megismételtük. Az eredmény a 12.3. ábrán látható.

3. lépés. Osszuk fel mindegyik alparcellát c al-alparcellára, amelyekben a c alparcella kezeléseket random jelöljük ki, elkülönítetten és függetlenül, az $(r)(a)(b)$ alparcellára.

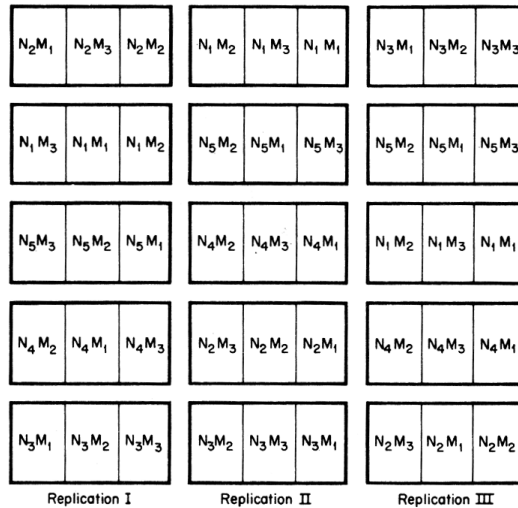
Példánkban, mindegyik alparcellát három al-alparcellára osztottuk fel, amelyekbe a három fajtát (V_1, V_2 és V_3) random helyeztük el. Ezt a randomizációs folyamatot megismételtük $(r)(a)(b) = 45$ -ször. A végső elrendezést a 12.4. ábra mutatja.

12.2.2. Kísérlet leírása

Az MTA Mezőgazdasági Kutatóintézetében, Martonvásáron a vetésidő, a N műtrágyázás és a genotípus hatását vizsgáltuk a kukorica szemtermésére (Berzsenyi

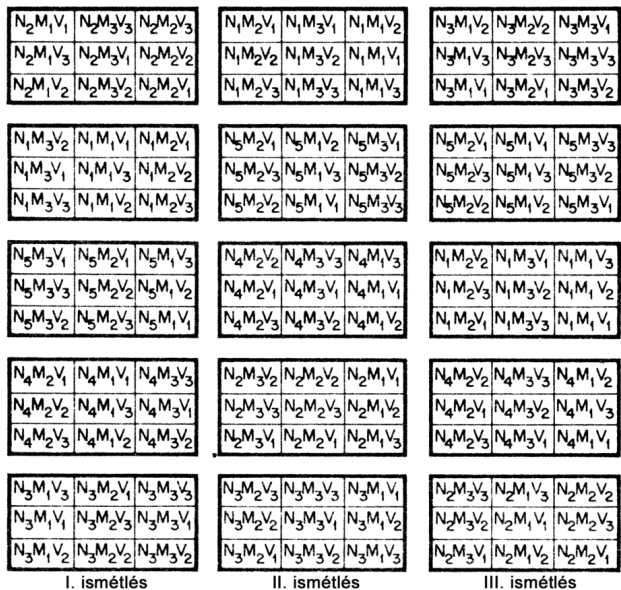


12.2. ábra. Az öt nitrogén szint (N_1 , N_2 , N_3 , N_4 és N_5) random elhelyezése a főparcellákon mindhárom ismétlésben, a split-split-plot elrendezés első lépéseként

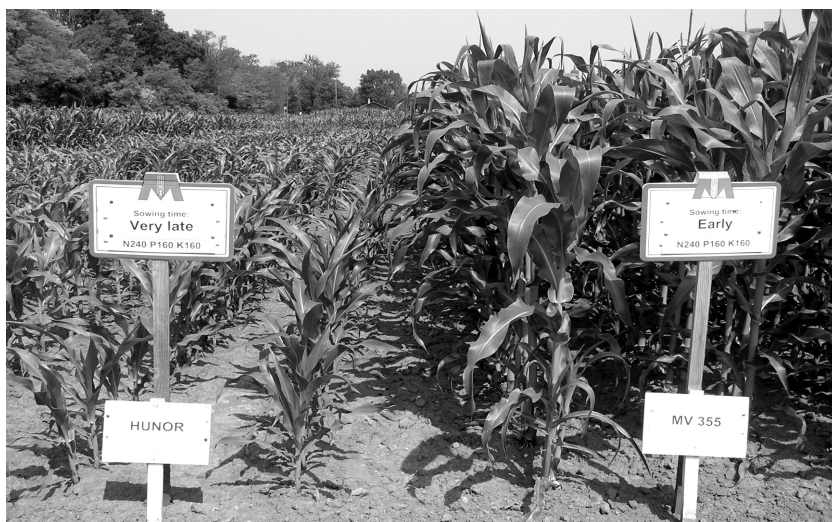


12.3. ábra. A három agrotechnikai eljárás (M_1 , M_2 és M_3) random elhelyezése a három alparcellán mind a 15 főparcellában, a kétszeresen osztott parcellás elrendezés második lépéseként

és Dang, 2008b) (12.1. kép). A három tényezős, split-split-plot elrendezésű kísérletben a N-műtrágya kezelések képezték a főparcellákat, a vetésidők az alparcellákat és a kukorica hibridek az al-alparcellákat. Az N-kezelések a következők voltak: 0, 60, 120, 180 és 240 kg ha⁻¹ (jelölésük: (N_1 , N_2 , N_3 , N_4 és N_5)). A vetés négy időpontban történt: 10 nappal az optimális időpont előtt (korai, S_1), optimális időpontban (április 24 körül, S_2), 10 nappal az optimális időpont után (késői, S_3) és 20 nappal az optimális időpont után (igen késői, S_4). A négy kereskedelmi forgalomban levő hibrid eltérő tenyésztő csoportokat képezett a FAO 200-299, FAO 300-399 és FAO 400-499 érésidő csoportokban (jelölésük: H_1 - H_4). A kísérlet 2006. évi termésszámainak varianciaanalízisét mutatjuk be.



12.4. ábra. Az 5 × 3 × 3-as kétszeresen osztott parcellás kísérlet elrendezése az öt nitrogén szinttel (N₁, N₂, N₃, N₄ és N₅) mint főparcella kezelések, a három agrotechnikai eljárással (M₁, M₂ és M₃) mint alparcella kezelések és a három fajtaival (V₁, V₂ és V₃) mint al-alparcella kezelések, három ismétlésben



12.1. kép. Kétszeresen osztott parcellás (split-split-plot) vetésidő kísérlet a martonvásári tenyészkertben (balra: nagyon késői, jobbra: korai vetésidő)

12.2.3. Varianciaanalízis lépései

A split-split-plot (kétszeresen osztott parcellás) elrendezésben beállított 5×4×4-es faktoriális kísérlet termésadatait az 12.4. táblázat tartalmazza. A varianciaanalízis lépései a következők.

12.4. táblázat. A kukorica hibridek szentermése a négy ismétléses, kétszeresen osztott parcellás 5×4×4-es marionvásári tartamkísérlésben 2006. évben

Ve- tés- idő	Szentermés t ha ⁻¹															
	H ₁				H ₂				H ₃				H ₄			
	I	II	III	IV	I	II	III	IV	I	II	III	IV	I	II	III	IV
	N ₁ (0 kg N ha ⁻¹)															
V ₁	10.427	6.37	9.183	7.402	8.111	4.793	6.387	7.122	8.404	6.726	9.466	7.178	9.45	5.178	5.855	6.783
V ₂	8.369	9.418	7.411	7.385	8.459	6.753	5.493	5.991	8.694	7.308	6.943	6.255	6.226	8.422	6.202	7.408
V ₃	7.363	7.034	7.497	8.475	7.602	5.203	7.321	9.146	6.546	6.639	8.247	8.243	8.1	5.978	6.793	5.617
V ₄	5.868	6.545	5.929	7.302	5.872	3.459	6.01	6.102	4.571	5.702	4.56	6.291	6.605	6.193	5.065	4.99
	N ₂ (60 kg N ha ⁻¹)															
V ₁	12.644	11.263	11.48	11.701	10.192	7.858	9.11	10.508	12.55	10.437	13.129	10.783	10.782	8.253	8.965	9.462
V ₂	11.367	11.924	9.769	11.724	11.532	10.706	8.399	10.867	12.287	13.656	8.838	11.704	10.846	12.286	8.587	12.664
V ₃	9.31	10.259	9.721	12.226	10.574	7.148	9.726	10.858	10.858	9.642	9.27	11.392	11.48	9.624	8.459	9.528
V ₄	10.626	9.915	8.833	11.199	10.079	7.37	8.189	10.609	9.846	8.942	9.812	10.173	9.801	7.037	9.132	8.549
	N ₃ (120 kg N ha ⁻¹)															
V ₁	13.023	13.771	11.864	12.687	11.639	11.372	10.555	9.336	11.748	12.35	13.482	11.886	11.498	11.981	10.877	11.857
V ₂	12.063	11.733	10.754	11.994	11.326	13.201	11.697	12.72	12.828	13.6	11.713	11.156	12.771	14.879	10.576	12.051
V ₃	11.295	10.008	11.363	11.451	10.255	9.668	12.292	11.665	10.317	13.601	12.532	11.388	12.878	12.37	10.752	11.285
V ₄	10.623	12.363	9.502	10.779	9.254	7.227	10.339	9.541	10.462	11.692	10.632	11.899	12.382	14.397	10.293	12.375
	N ₄ (180 kg N ha ⁻¹)															
V ₁	13.063	11.837	11.288	13.741	12.579	10.078	9.593	10.982	11.809	11.426	11.748	13.478	13.335	11.993	11.335	10.904
V ₂	13.729	12.019	11.456	13.239	12.003	9.774	11.285	11.889	14.114	14.796	12.713	12.656	13.889	12.338	10.012	12.054
V ₃	10.805	11.433	11.281	10.748	11.342	8.801	11.026	10.624	12.425	12.206	11.737	12.264	12.138	10.731	10.114	11.657
V ₄	11.238	8.711	10.959	9.497	12.259	8.88	7.996	9.66	11.263	11.433	10.6	10.656	13.247	10.964	9.952	11.388
	N ₅ (240 kg N ha ⁻¹)															
V ₁	11.668	12.133	12.241	11.333	11.815	11.437	10.279	10.625	11.349	10.386	12.396	11.949	12.901	12.597	11.263	12.071
V ₂	12.284	13.811	12.023	13.36	11.484	12.516	11.018	10.441	12.568	14.533	11.602	12.982	12.851	13.353	9.426	10.723
V ₃	11.682	11.206	11.235	11.151	11.804	11.195	10.608	10.114	11.623	13.492	12.954	10.975	13.365	14.236	9.945	11.775
V ₄	10.38	11.841	11.379	12.024	10.995	10.127	7.243	9.244	10.223	13.221	11.999	11.436	11.779	11.799	11.667	8.108

1. lépés. Szerkesszük meg a variancia táblázatot a kétszeresen osztott parcellás elrendezéshez.

Variáció forrása	Szabadság- fokok	Eltérés- négyzet összeg (SS)	Közepes négyzetes eltérés (MS)	Számított F-érték	Táblázati F-érték	
					5%	1%
Főparcella analízis						
Ismétlés	$r-1 = 3$					
Főparcella faktor (A)	$a-1 = 4$					
Hiba (a)	$(r-1)(a-1) = 12$					
Alparcella analízis:						
Alparcella faktor (B)	$b-1 = 3$					
A × B	$(a-1)(b-1) = 12$					
Hiba (b)	$a(r-1)(b-1) = 45$					
Al-alparcella analízis:						
Al-alparcella faktor (C)	$c-1 = 3$					
A × C	$(a-1)(c-1) = 12$					
B × C	$(b-1)(c-1) = 9$					
A × B × C	$(a-1)(b-1)(c-1) = 36$					
Hiba (c)	$ab(r-1)(c-1) = 180$					
Összes	$rabc-1 = 319$					

2. lépés. A. Szerkesszük meg az ismétlés × A faktor kétirányú táblázat összegeket és számítsuk ki az ismétlés összeget, a faktor A összeget és a főátlagot. Példánkban, az ismétlés × nitrogén táblázatát az összegeknek (RA), a nitrogén összeggel (A) és a főösszeggel (G) a 12.5. táblázat mutatja.

B. Számítsuk ki a korrekciós faktort és a különböző SS (eltérésnégyzet) összegeket:

$$C.F. = \frac{G^2}{rabc} = \frac{(3318.2)^2}{(4)(5)(4)(4)} = 34407.7$$

12.5. táblázat. A termésösszegek ismétlés × N-műtrágya táblázata, az 12.4. táblázat adataiból számolva

Nitrogén	Termésösszeg (RA)				Nitrogén összes (A)
	I. ismétlés	II. ismétlés	III. ismétlés	IV. ismétlés	
N ₁	120.7	101.7	108.4	111.7	442.5
N ₂	174.8	156.3	151.4	173.9	656.4
N ₃	184.4	194.2	179.2	184.1	741.9
N ₄	199.2	177.4	173.1	185.4	735.1
N ₅	188.8	197.9	177.3	178.3	742.3
Ismétlés összes (R)	867.9	827.5	789.4	833.4	
Főösszeg (G)					3318.2

$$\text{Összes SS} = \sum x^2 - C.F. = \left[(10.427)^2 + \dots + (8.108)^2 \right] - 34407.7 = 1641.3$$

$$\text{Ismétlés SS} = \frac{\sum R^2}{abc} - C.F. = \frac{(867.9)^2 + \dots + (833.4)^2}{(5)(4)(4)} - 34407.7 = 38.73$$

$$\text{A (nitrogén) SS} = \frac{\sum A^2}{rbc} - C.F. = \frac{(442.5)^2 + \dots + (742.3)^2}{(4)(4)(4)} - 34407.7 = 1037.05$$

$$\begin{aligned} \text{Hiba (a) SS} &= \frac{\sum (RA)^2}{bc} - C.F. - \text{IsmétlésSS} - \text{ASS} = \\ &= \frac{(120.7)^2 + \dots + (178.3)^2}{(4)(4)} - 34407.7 - 38.73 - 1037.05 = 49.83 \end{aligned}$$

3. lépés. Végezzük el az alparcella analízist.

A) Szerkesszünk meg két táblázat összeget:

(i) Számítsuk ki a faktor A × faktor B kétirányú táblázat összegeket, a számított faktor B összeggel. Példánkban, a nitrogén × vetésidő táblázat összegek (AB), a számított vetésidő összegekkel (B) a 12.6. táblázat mutatja.

(ii) Számítsuk ki az ismétlés×faktor A×faktor B háromirányú táblázat összegeket. Példánkban, az ismétlés×nitrogén×vetésidő táblázat összegeket (RAB) a 12.7. táblázat tartalmazza.

B) Számítsuk ki a különböző eltérésnégyzet összegeket:

12.6. táblázat. A nitrogén × vetésidő termésösszegek táblázata, a 12.4. táblázat adataiból számítva

Nitrogén	Termés összeg (AB)			
	S ₁	S ₂	S ₃	S ₄
N1	118.8	116.7	115.8	91.1
N2	169.1	177.2	160.1	150.1
N3	189.9	195.1	183.1	173.8
N4	189.2	198.0	179.3	168.7
N5	186.4	195.0	187.4	173.5
Vetésidő összes (B)	853.4	882.0	825.7	757.2

$$B \text{ (vetésidő) SS} = \frac{\sum B^2}{rac} - C.F. = \frac{(853.4)^2 + \dots + (757.2)^2}{(4)(5)(4)} - 34407.7 = 109.15$$

$$A \times B \text{ (nitrogen} \times \text{vetésidő) SS} = \frac{\sum (AB)^2}{rc} - C.F. - ASS - BSS =$$

$$= \frac{(118.8)^2 + \dots + (173.5)^2}{(4)(4)} - 34407.7 - 1037.05 - 109.15 = 12.1$$

$$\text{Hiba (b) SS} = \frac{\sum (RAB)^2}{c} - C.F. - \text{Ismétlés SS} - A \text{ SS} - \text{Hiba (a) SS} - B \text{ SS} - A \times B \text{ SS} =$$

$$= \frac{(36.39)^2 + \dots + (40.81)^2}{4} - 34407.7 - 38.73 - 1037.05 - 49.83 - 109.15 - 12.1 = 71.94$$

12.7. táblázat. Az ismétlés \times nitrogén \times vetésidő termésösszegek táblázata,
a 12.4. táblázat adataiból számítva

Vetésidő	Termésösszeg (RAB)			
	I. ismétlés	II. ismétlés	III. ismétlés	IV. ismétlés
N ₁ (0 kg N ha ⁻¹)				
S ₁	36.39	23.07	30.89	28.49
S ₂	31.75	31.90	26.05	27.04
S ₃	29.61	24.85	29.86	31.48
S ₄	22.92	21.90	21.56	24.69
N ₂ (60 kg N ha ⁻¹)				
S ₁	46.17	37.81	42.68	42.45
S ₂	46.03	48.57	35.59	46.96
S ₃	42.22	36.67	37.18	44.00
S ₄	40.35	33.26	35.97	40.53
N ₃ (120 kg N ha ⁻¹)				
S ₁	47.91	49.47	46.78	45.77
S ₂	48.99	53.41	44.74	47.92
S ₃	44.75	45.65	46.94	45.79
S ₄	42.72	45.68	40.77	44.59
N ₄ (180 kg N ha ⁻¹)				
S ₁	50.79	45.33	43.96	49.11
S ₂	53.73	48.93	45.47	49.84
S ₃	46.71	43.17	44.16	45.29
S ₄	48.01	39.99	39.51	41.20
N ₅ (240 kg N ha ⁻¹)				
S ₁	47.73	46.55	46.18	45.98
S ₂	49.19	54.21	44.07	47.51
S ₃	48.47	50.13	44.74	44.02
S ₄	43.38	46.99	42.29	40.81

4. lépés. Készítsük el az al-alparcella analízist.

A) Szerkesszünk meg három táblázat összeget.

(i) Számítsuk ki a faktor A × faktor C kétirányú táblázat összeget a faktor C összeggel. Példánkban, a nitrogén × hibrid táblázat összeget (AC), a hibrid összeggel (C) az 12.8. táblázat mutatja.

(ii) Számítsuk ki a faktor B × faktor C kétirányú táblázat összeget.

Példánkban, a vetésidő × fajta táblázat összeget (BC) a 12.9. táblázat mutatja.

(iii) Számítsuk ki a faktor A × faktor B × faktor C háromirányú táblázat összeget. Példánkban, a nitrogén × vetésidő × hibrid táblázat összeget (ABC) a 12.10. táblázat mutatja.

12.8. táblázat. A nitrogén × hibrid termésösszegek táblázata
a 12.4. táblázat adataiból számolva

Nitrogén	Termés összeg (AC)			
	H ₁	H ₂	H ₃	H ₄
N ₁	122.0	103.8	111.8	104.9
N ₂	174.0	153.7	173.7	155.5
N ₃	185.3	172.1	191.3	193.2
N ₄	185.0	168.8	195.3	186.1
N ₅	189.8	170.9	193.7	187.9
Hibrid összes (C)	856.1	769.3	865.4	827.5

12.9. táblázat. A vetésidő × fajta termésösszegek táblázata
a 12.4. táblázat adataiból számolva

Vetésidő	Termés összeg (BC)			
	H ₁	H ₂	H ₃	H ₄
S ₁	229.1	194.4	222.7	207.3
S ₂	225.8	207.6	230.9	217.6
S ₃	205.5	197.0	216.4	206.8
S ₄	195.5	170.5	195.4	195.7

12.10. táblázat. A nitrogén × vetésidő × hibrid termésösszegek táblázata
a 12.4. táblázat adataiból számolva

Vetésidő	Termésösszeg (ABC)			
	H ₁	H ₂	H ₃	H ₄
N ₁ (0 kg N ha ⁻¹)				
S ₁	33.38	26.41	31.77	27.27
S ₂	32.58	26.70	29.20	28.26
S ₃	30.37	29.27	29.68	26.49
S ₄	25.64	21.44	21.12	22.85

A 12.10. táblázat folytatása

Vetésidő	Termésösszeg (ABC)			
	H ₁	H ₂	H ₃	H ₄
N ₂ (60 kg N ha ⁻¹)				
S ₁	47.09	37.67	46.90	37.46
S ₂	44.78	41.50	46.48	44.38
S ₃	41.52	38.31	41.16	39.09
S ₄	40.57	36.25	38.77	34.52
N ₃ (120 kg N ha ⁻¹)				
S ₁	51.34	42.90	49.47	46.21
S ₂	46.54	48.94	49.30	50.28
S ₃	44.12	43.88	47.84	47.28
S ₄	43.27	36.36	44.69	49.45
N ₄ (180 kg N ha ⁻¹)				
S ₁	49.93	43.23	48.46	47.57
S ₂	50.44	44.95	54.28	48.29
S ₃	44.27	41.79	48.63	44.64
S ₄	40.41	38.80	43.95	45.55
N ₅ (240 kg N ha ⁻¹)				
S ₁	47.38	44.16	46.08	48.83
S ₂	51.48	45.16	51.69	46.35
S ₃	45.27	43.72	49.04	49.32
S ₄	45.62	37.61	46.88	43.35

B. Számítsuk ki az eltérésnégyzet (SS) összegeket:

$$C \text{ (hibrid) SS} = \frac{\sum C^2}{rab} - C.F. = \frac{(856.1)^2 + \dots + (827.6)^2}{(4)(5)(4)} - 34407.7 = 74.4$$

$$A \times C \text{ SS} = \frac{\sum (AC)^2}{rb} - C.F. - ASS - CSS =$$

$$= \frac{(122.0)^2 + \dots + (187.9)^2}{(4)(4)} - 34407.7 - 1037.05 - 74.4 = 25.43$$

$$B \times C \text{ SS} = \frac{\sum (BC)^2}{ra} - \text{C.F.} - B \text{ SS} - C \text{ SS} = \frac{(229.1)^2 + \dots + (195.7)^2}{(4)(5)} -$$

$$-34407.7 - 109.15 - 74.4 = 8.7$$

$$A \times B \times C \text{ SS} =$$

$$\frac{\sum (ABC)^2}{r} - \text{C.F.} - A \text{ SS} - B \text{ SS} - C \text{ SS} - A \times B \text{ SS} - A \times C \text{ SS} - B \times C \text{ SS} =$$

$$= \frac{(33.38)^2 + \dots + (43.35)^2}{4} - 34407.7 - 1037.05 - 109.15 - 74.4 - 12.1 - 25.43 - 8.7 = 37.72$$

$$\text{Hiba (c) SS} = \text{Összes SS} - (\text{minden többi SS összege}) =$$

$$= 1641.3 - (38.73 + 1037.05 + 49.83 + 109.15 + 12.1 + 71.94 + 74.4 + 25.43 + 8.7 + 37.72) = 176.25$$

5. lépés. A variáció minden forrására számítsuk ki a közepes négyzetes eltérést (MS), az SS értékeket elosztva a megfelelő d.f. értékekkel.

$$\text{Ismétlés MS} = \frac{\text{Ismétlés SS}}{r-1} = \frac{38.73}{3} = 12.91$$

$$A \text{ MS} = \frac{A \text{ SS}}{a-1} = \frac{1037.05}{4} = 259.26$$

$$\text{Hiba (a) MS} = \frac{\text{Hiba (a) SS}}{(r-1)(a-1)} = \frac{49.83}{(3)(4)} = 4.15$$

$$B \text{ MS} = \frac{B \text{ SS}}{b-1} = \frac{109.15}{3} = 36.38$$

$$A \times B \text{ MS} = \frac{A \times B \text{ SS}}{(a-1)(b-1)} = \frac{12.1}{(4)(3)} = 1.008$$

$$\text{Hiba (b) MS} = \frac{\text{Hiba (b) SS}}{a(r-1)(b-1)} = \frac{71.94}{(5)(3)(3)} = 1.599$$

$$C \text{ MS} = \frac{C \text{ SS}}{c-1} = \frac{74.4}{3} = 24.8$$

$$A \times C \text{ MS} = \frac{A \times C \text{ SS}}{(a-1)(c-1)} = \frac{25.43}{(4)(3)} = 2.12$$

$$B \times C \text{ MS} = \frac{B \times C \text{ SS}}{(b-1)(c-1)} = \frac{8.7}{(3)(3)} = 0.97$$

$$A \times B \times C \text{ MS} = \frac{A \times B \times C \text{ SS}}{(a-1)(b-1)(c-1)} = \frac{37.72}{(4)(3)(3)} = 1.048$$

$$\text{Hiba (c) MS} = \frac{\text{Hiba (c) SS}}{ab(r-1)(c-1)} = \frac{176.25}{(5)(4)(3)(3)} = 0.979$$

6. lépés. Számítsuk ki az F-értékeket minden hatásra úgy, hogy mindegyik közepes négyzetes eltérést (MS) elosztjuk a megfelelő hiba MS értékkel:

$$F(A) = \frac{A \text{ MS}}{\text{Hiba (a) MS}} = \frac{259.26}{4.15} = 62.47$$

$$F(B) = \frac{B \text{ MS}}{\text{Hiba (b) MS}} = \frac{36.38}{1.599} = 22.75$$

$$F(A \times B) = \frac{A \times B \text{ MS}}{\text{Hiba (b) MS}} = \frac{1.008}{1.599} < 1$$

$$F(C) = \frac{C \text{ MS}}{\text{Hiba (c) MS}} = \frac{24.8}{0.979} = 25.33$$

$$F(A \times C) = \frac{A \times C \text{ MS}}{\text{Hiba (c) MS}} = \frac{2.12}{0.979} = 2.165$$

$$F(B \times C) = \frac{B \times C \text{ MS}}{\text{Hiba (c) MS}} = \frac{0.97}{0.979} < 1$$

$$F(A \times B \times C) = \frac{A \times B \times C \text{ MS}}{\text{Hiba (c) MS}} = \frac{1.048}{0.979} = 1.07$$

7. lépés. Minden hatásra, melyre a kiszámított F-érték 1-nél nem kisebb, keressük ki a megfelelő táblázati F-értéket, f_1 = a számláló MS d.f. értéke és f_2 = a nevező MS d.f. értéke, 5%-os és 1%-os szignifikancia szinten.

8. lépés. Számítsuk ki a három variációs koefficienset a három hiba tagnak megfelelően:

$$CV(a) = \frac{\sqrt{\text{Hiba(a) MS}}}{\text{Főátlag}} \times 100 = \frac{\sqrt{4.15}}{10.369} \times 100 = 19.65\%$$

$$CV(b) = \frac{\sqrt{\text{Hiba}(b)MS}}{\text{Főátlag}} \times 100 = \frac{\sqrt{1.599}}{10.369} \times 100 = 12.20\%$$

$$CV(c) = \frac{\sqrt{\text{Hiba}(c)MS}}{\text{Főátlag}} \times 100 = \frac{\sqrt{0.979}}{10.369} \times 100 = 9.54\%$$

A CV (a) értéke mutatja a pontosság mértékét, összefüggésben a főparcella faktor főhatásával, a cv (b) jelzi az alparcella faktor főhatásának és a főparcellával való interakciónak a pontosságát és a cv (c) értéke jelzi az al-alparcella faktor fő hatásának és a többi faktorialkottott interakcióinak pontosságát. Normálisan, a három variációs koefficiens nagysága csökken cv (a)-tól cv (b)-ig és cv (c)-ig. Példánkban is ez a sorrend érvényesült.

9. lépés. Írjuk be a 2–8. lépésben kapott értékeket a varianciatáblázatba, és hasonlítsuk össze a számított F-értéket a megfelelő táblázati F-értékkel és jelöljük meg csillagokkal a szignifikancia szintet.

Példánkban az eredmények (12.11. táblázat) jelzik, hogy a három faktor interakciója (nitrogén × vetésidő × hibrid) nem szignifikáns, és csupán egy két faktoros interakció (nitrogén × hibrid) szignifikáns.

12.11. táblázat. Az 12.4. táblázatban közölt szemtermés adatok varianciaanalízise (Kétszeresen osztott parcellás elrendezés)

Variáció forrása	Szabadságfokok	Eltérés-négyzet összegek (SS)	Közepes négyzetes eltérés (MS)	Számított F-érték	Táblázati F-érték	
					5%	1%
<i>Főparcella analízis</i>						
Ismétlés	3	38.73	12.91			
Nitrogén (A)	4	1037.05	259.26	62.47**	3.26	5.41
Hiba (a)	12	49.83	4.15			
<i>Alparcella analízis</i>						
Vetésidő (B)	3	109.15	36.38	22.25**	2.82	4.26
A × B	12	12.1	1.008	<1	–	–
Hiba (b)	45	71.94	1.599			
<i>Al-alparcella analízis</i>						
Hibrid (C)	3	74.4	24.8	25.33**	2.65	3.88
A × C	12	25.43	2.12	2.165*	1.80	2.28
B × C	9	8.7	0.97	< 1		
A × B × C	36	37.72	1.048	1.07 ^{ns}	1.49	1.74
Hiba (c)	180	176.25	0.979			
Összes	319	1641.3				

CV (a) = 19.65%, CV(b) = 12.2%, CV(c) = 9.54% * = szignifikáns 5%-os szinten, ** = szignifikáns 1%-os szinten, ^{ns} = nem szignifikáns.

Szignifikáns differencia (SzD) kiszámítása (Kétszeresen osztott parcellás elrendezés)

A kétszeresen osztott parcellás elrendezésre a pár összehasonlításnak 12 típusa van, mindegyik szükségessé teszi a saját SzD értékeit. Ezek a pár összehasonlítások, a megfelelő s_d értékek kiszámításának megfelelő formuláival a 12.12. táblázatban láthatók.

12.12. táblázat. Az átlagos különbség szórása (standard hiba) a pár összehasonlítás 12 típusának mindegyikére kétszeresen osztott parcellás (split-split-plot) elrendezésnél (Gomez és Gomez, 1984 alapján)

Sor-szám	Pár összehasonlítás típusa	s_d
1.	Két főparcella átlag között (minden alparcella és al-alparcella kezelés átlagában)	$\sqrt{\frac{2E_a}{rbc}}$
2.	Két alparcella átlag között (minden főparcella és al-alparcella kezelés átlagában)	$\sqrt{\frac{2E_b}{rac}}$
3.	Két alparcella átlag között (minden al-alparcella kezelés átlagában) a főparcella faktor azonos vagy eltérő szintjein	$\sqrt{\frac{2E_b}{rc}}$
4.	Két al-alparcella átlag között (minden főparcella és alparcella kezelés átlagában)	$\sqrt{\frac{2E_c}{rab}}$
5.	Két al-alparcella átlag között a főparcella szintjén (minden alparcella kezelés átlagában)	$\sqrt{\frac{2E_c}{rb}}$
6.	Két al-alparcella átlag között (minden főparcella és alparcella kezelés átlagában)	$\sqrt{\frac{2E_c}{ra}}$
7.	Két al-alparcella faktor átlaga között a főparcella és alparcella kezeléseket ugyanazon kombinációjánál	$\sqrt{\frac{2E_c}{r}}$
8.	Két főparcella átlag között (minden al-alparcella kezelés átlagában) az alparcella faktor azonos vagy eltérő szintjein	$\sqrt{\frac{2[(b-1)E_b + E_a]}{rbc}}$
9.	Két alparcella átlag között (minden főparcella kezelés átlagában) az al-alparcella faktor azonos vagy eltérő szintjein	$\sqrt{\frac{2[(c-1)E_c + E_b]}{rac}}$
10.	Két alparcella átlag között a főparcella és az al-alparcella kezeléseket ugyanazon kombinációjánál	$\sqrt{\frac{2[(c-1)E_c + E_b]}{rc}}$
11.	Két főparcella átlag között (az összes alparcella kezelés átlagában) az al-alparcella és az al-alparcella faktor azonos vagy különböző szintjein	$\sqrt{\frac{2[(c-1)E_c + E_a]}{rbc}}$
12.	Két főparcella átlag között az alparcella és az al-alparcella kezeléseket ugyanazon kombinációjánál	$\sqrt{\frac{2[b(c-1)E_c + (b-1)E_b + E_a]}{rbc}}$

E_a = hiba (a) MS, E_b = hiba (b) MS, E_c = hiba (c) MS, r = ismétlések száma, a = főparcella kezeléseket száma, b = alparcella kezeléseket száma, c = al-alparcella kezeléseket száma.

A varianciatáblázatban (12.11. táblázat) mind a három főhatás és egy interakció hatás a nitrogén és a hibrid között szignifikáns. Következésképpen, csupán a következő típusú pár összehasonlításokat tudjuk megvizsgálni:

1. Összehasonlítás a négy vetésidő között: vetésidők minden hibrid és nitrogén dózis átlagában – mivel a vetésidőt magában foglaló egyik interakció hatás sem szignifikáns.

2. Összehasonlítás a négy hibrid között, mindegyik vetésidő átlagában, azonban ugyanannál a nitrogén dózisonál – mivel a nitrogén \times hibrid interakció szignifikáns.

3. Összehasonlítás az öt nitrogén dózis között, mindegyik vetésidő átlagában, azonban ugyanannál a hibridnél, mivel a nitrogén \times hibrid interakció szignifikáns. Mivel a pár összehasonlítás megfelelő csupán az 1. és 2. összehasonlításra, és nem a 3. összehasonlításra, ahol a kezelés (azaz a nitrogén dózisek) kvantitatív, az SzD kiszámítási eljárást az 1. és 2. pár összehasonlításra adjuk meg.

Az 1. összehasonlításra (összehasonlítás négy vetésidő között a hibridek és a nitrogén dózisek átlagában) az SzD kiszámításának lépései a következők:

1. lépés. Számítsuk ki az s_d értéket, követve a 12.12. táblázat 2. típusú összehasonlításra megadott formuláját.

$$s_d = \sqrt{\frac{2E_b}{rac}} = \sqrt{\frac{2(1.599)}{(4)(5)(4)}} = 0.20$$

2. lépés. Keressük ki a megfelelő táblázati t értéket $n = \text{hiba (b) d.f.} = 45$ d.f.-re, amely 5%-os szignifikancia szintre 2.014 és 1%-os szignifikancia szinten 2.690.

3. lépés. Számítsuk ki az SzD értékeket az 5%-os és 1%-os szignifikancia szintekre:

$$SzD_\alpha = (t_\alpha) (s_d)$$

$$SzD_{5\%} = (2.014) (0.20) = 0.403 \text{ t ha}^{-1}$$

$$SzD_{1\%} = (2.690) (0.20) = 0.538 \text{ t ha}^{-1}$$

4. lépés. Számítsuk ki a négy vetésidő átlagtermését az összes nitrogén dózis és hibrid átlagában:

Vetésidő	Átlagtermés t ha ⁻¹
S1	10.669
S2	11.024
S3	10.321
S4	9.464

5. lépés. A 4. lépésben kiszámított átlagtermések felhasználásával, számítsuk ki az átlagos különbséget a bennünket érdeklő bármelyik vetésidő párra és hasonlítsuk össze a 3. lépésben kiszámított SzD értékkel. Például az S_1 és S_2 összehasonlításakor, kiszámítjuk az átlagos különbséget: $11.024 - 10.669 = 0.355 \text{ t ha}^{-1}$. Mivel a számított átlag-

gos különbség kisebb, mint az SzD érték 5%-os szignifikancia szinten, megállapítjuk, hogy a S_1 és S_2 közötti különbség nem szignifikáns.

A 2. összehasonlításnál (összehasonlítás a négy hibrid között, mindegyik vetésidő átlagában, azonban ugyanannál a nitrogén dózissal) a lépésenkénti eljárás az SzD kiszámítására:

1. lépés. Számítsuk ki az s_d értéket, követve a 12.12. táblázat 5. típusú összehasonlításának formuláját:

$$s_d = \sqrt{\frac{2E_c}{rb}} = \sqrt{\frac{2(0.979)}{(4)(4)}} = 0.350 \text{ t ha}^{-1}$$

2. lépés. Keressük ki a táblázati t értéket $n = \text{hiba (c) d.f.} = 180 \text{ d.f.}$ -re, amely 1.96 az 5%-os és 2.58 az 1%-os szignifikancia szinten.

3. lépés. Számítsuk ki az SzD értéket, 5%-os és 1%-os szignifikancia szinten:

$$\text{SzD}_\alpha = (t_\alpha) (s_d)$$

$$\text{SzD}_{5\%} = (1.96) (0.350) = 0.686 \text{ t ha}^{-1}$$

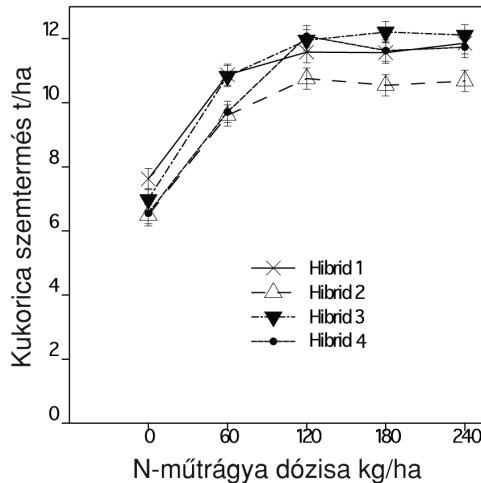
$$\text{SzD}_{1\%} = (2.58) (0.350) = 0.903 \text{ t ha}^{-1}$$

4. lépés. Szerkesszünk egy hibrid \times nitrogén kétirányú táblázatot az átlagos értékekre, a négy vetésidő átlagában, mint a 12.13. táblázat mutatja. Ahhoz, hogy összehasonlítsuk a hibrid átlagok bármely párját, ugyanannál a nitrogén dózissal, számítsuk ki az átlagos különbséget és hasonlítsuk össze a 3. lépésben kiszámított SzD értékkel. Példánkban, a H_1 és H_2 összehasonlításakor 180 kg N ha^{-1} dózissal, az átlagos különbség $11.565 - 10.548 = 1.017 \text{ t ha}^{-1}$. Mivel ez az átlagos különbség nagyobb, mint az SzD értéke 1%-os szignifikancia szinten, levonjuk a következtetést, hogy H_1 és H_2 hibridek átlagtermése 180 kg N ha^{-1} dózissal szignifikánsan különbözik $P = 1\%$ -os szinten. A 12.5. ábra jól szemlélteti, hogy a szignifikáns N-műtrágyázás \times hibrid interakcióért elsősorban a 2. és 4. hibrid eltérő reakciója felelős.

12.13. táblázat. Négy kukorica hibrid átlagtermése öt eltérő nitrogén dózissal, kétszeresen osztott parcellás elrendezésnél (alapadatok a 12.4. táblázatban vannak)

Nitrogén dózis kg ha^{-1}	Átlagtermés t ha^{-1}			
	H_1	H_2	H_3	H_4
0	7.624	6.489	6.986	6.554
60	10.873	9.608	10.832	9.716
120	11.580	10.755	11.955	12.076
180	11.565	10.548	12.208	11.628
240	11.859	10.684	12.105	11.741

Négy vetésidő átlaga, mindegyik négy ismétlésben. Az SzD értékek két hibrid összehasonlításához ugyanannál a nitrogén dózissal 0.686 t ha^{-1} az 5%-os és 0.903 t ha^{-1} az 1%-os szignifikancia szintnél.



12.5. ábra. Szignifikáns N-műtrágyázás × hibrid interakció ábrázolása a vetésidő kísérletben

12.2.4. Varianciaanalízis GenStat programmal

A varianciaanalízis lépéseinek sorrendje a következő: Stats – Analysis of Variance – General Analysis of Variance – Split-Split-Plot- Design. Az adatokat az Available Data tartalmazza. Y-variate: Yield2006, Treatment Structure: Fertilisation × Sowing × Hybrid. Blocks: Replication, Whole Plots: Fertilisation, Sub-plots: Sowing, Sub-Sub-Plots: Hybrid. Options-re kattintás lehetőséget ad az LSD és a Multiple Comparisons kiválasztására. A Run parancsra a program elvégzi a varianciaanalízist. Further Output – Means plot kiválasztása lehetővé teszi az adatok ábrázolását.

Analysis of variance

Variate: Yield2006

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Replication stratum	3	38.7134	12.9045	3.11	
Replication.Fertilisation stratum					
Fertilisation	4	1037.4757	259.3689	62.45	<.001
Residual	12	49.8418	4.1535	2.52	
Replication.Fertilisation.Sowing stratum					
Sowing	3	107.2188	35.7396	21.68	<.001
Fertilisation.Sowing	12	11.0792	0.9233	0.56	0.862
Residual	45	74.1767	1.6484	1.71	

Replication.Fertilisation.Sowing.Hybrid stratum					
Hybrid	3	70.1585	23.3862	24.22	<.001
Fertilisation.Hybrid	12	24.5357	2.0446	2.12	0.018
Sowing.Hybrid	9	15.1902	1.6878	1.75	0.081
Fertilisation.Sowing.Hybrid					
	36	39.3719	1.0937	1.13	0.293
Residual	180	173.7827	0.9655		
Total	319	1641.5446			

Tables of means

Variate: Yield2006

Grand mean 10.369

Fertilisation		1	2	3	4	5	
		6.913	10.257	11.592	11.487	11.598	
Sowing		1	2	3	4		
		10.669	11.024	10.321	9.464		
Hybrid		1	2	3	4		
		10.700	9.617	10.817	10.343		
Fertilisation	Sowing		1	2	3	4	
	1		7.427	7.296	7.238	5.691	
	2		10.570	11.072	10.005	9.382	
	3		11.870	12.191	11.445	10.860	
	4		11.824	12.373	11.208	10.544	
	5		11.653	12.186	11.710	10.842	
Fertilisation	Hybrid		1	2	3	4	
	1		7.624	6.489	6.986	6.554	
	2		10.873	9.608	10.832	9.716	
	3		11.580	10.755	11.955	12.076	
	4		11.565	10.548	12.208	11.628	
	5		11.859	10.684	12.105	11.741	
Sowing	Hybrid		1	2	3	4	
	1		11.456	9.719	11.134	10.367	
	2		11.292	10.378	11.547	10.878	
	3		10.277	9.849	10.818	10.341	
	4		9.776	8.523	9.771	9.786	
Fertilisation	Sowing	Hybrid		1	2	3	4
	1	1		8.345	6.603	7.943	6.816
		2		8.146	6.674	7.300	7.064
		3		7.592	7.318	7.419	6.622
		4		6.411	5.361	5.281	5.713

2	1	11.772	9.417	11.725	9.365
	2	11.196	10.376	11.621	11.096
	3	10.379	9.576	10.290	9.773
	4	10.143	9.062	9.693	8.630
3	1	12.836	10.725	12.366	11.553
	2	11.636	12.236	12.324	12.569
	3	11.029	10.970	11.959	11.821
	4	10.817	9.090	11.171	12.362
4	1	12.482	10.808	12.115	11.892
	2	12.611	11.238	13.570	12.073
	3	11.067	10.448	12.158	11.160
	4	10.101	9.699	10.988	11.388
5	1	11.844	11.039	11.520	12.208
	2	12.869	11.365	12.921	11.588
	3	11.318	10.930	12.261	12.330
	4	11.406	9.402	11.720	10.838

Standard errors of differences of means

Table	Fertilisation	Sowing	Hybrid	Fertilisation Sowing
rep.	64	80	80	16
s.e.d.	0.3603	0.2030	0.1554	0.5332
d.f.	12	45	180	41.79
Except when comparing means with the same level(s) of Fertilisation				0.4539
d.f.				45

Table	Fertilisation Hybrid	Sowing Hybrid	Fertilisation Sowing Hybrid
rep.	16	20	4
s.e.d.	0.4694	0.3371	0.8040
d.f.	33.49	193.06	156.90
Except when comparing means with the same level(s) of Fertilisation			0.7537
d.f.	180	193.06	
Sowing		0.3107	
d.f.	180		
Fertilisation.Sowing			0.6948
d.f.	180		
Fertilisation.Hybrid			0.7537
d.f.	193.06		

Least significant differences of means (5% level)

Table	Fertilisation	Sowing	Hybrid	Fertilisation Sowing
rep.	64	80	80	16
l.s.d.	0.7850	0.4089	0.3066	1.0763
d.f.	12	45	180	41.79

Except when comparing means with the same level(s) of

Fertilisation	0.9142
d.f.	45

Table	Fertilisation Hybrid	Sowing Hybrid	Fertilisation Sowing Hybrid
rep.	16	20	4
l.s.d.	0.9544	0.6648	1.5880
d.f.	33.49	193.06	156.90

Except when comparing means with the same level(s) of

Fertilisation	0.6855	1.4866
d.f.	180	193.06

Sowing	0.6131
d.f.	180

Fertilisation.Sowing	1.3710
d.f.	180

Fertilisation.Hybrid	1.4866
d.f.	193.06

12.3. 2ⁿ FAKTORIÁLIS KÍSÉRLETEK

A faktoriális kísérletek speciális típusa olyan kísérleteket tartalmaz, amelyben minden faktornak két szintje van. Az ilyen kísérleteket általában 2ⁿ faktoriálisnak hívjuk. Számos faktornak természeténél fogva két szintje van, ilyenek a szóban forgó faktor jelenléte vagy hiánya, az állatok nem. Ezen kívül, a faktorok gyakran hasznosan alkalmazhatók két jelentősen eltérő szinten, hogy a kísérlet jó eséllyel feltárjon interakciót más faktorokkal. A 2ⁿ kísérletek specifikus sajátossága, hogy minden főhatás és interakció (és SS értékeik) meghatározható a kezelésátlagok közötti egyszerű különbségek formájában (Mead *et al.*, 2003).

Egy faktor fő hatása a két szint átlaga közötti különbség. Például az A faktorra a fő hatás:

$$A = \frac{T_{A_1} - T_{A_2}}{(N/2)}$$

ahol T_{A_1} és T_{A_2} az összes többi faktor összege az A faktor 1. és 2. szintjén, és N az egységek (parcellák) összes száma a kísérletben ($N/2$ az A 1. szintjére és $N/2$ az A 2. szintjére).

Az interakció két faktor között a különbség az egyik faktor hatása között, mérve a másik faktor két szintjén. Az A és B faktorra az interakció

$$\frac{(T_{A_1B_1} - T_{A_2B_1}) - (T_{A_1B_2} - T_{A_2B_2})}{(N/2)}$$

ahol, például $T_{A_1B_1}$ minden más faktor összege az A faktor 1. szintjének és a B faktor 1. szintjének kombinációjára.

Meghatároztuk az interakciót az A faktor szintjei közötti különbségként kifejezve a B faktor két szintjén. Az interakció szimmetrikus A-ban és B-ben és ez a meghatározás azonos az alábbival:

$$\frac{(T_{A_1B_1} - T_{A_1B_2}) - (T_{A_2B_1} - T_{A_2B_2})}{(N/2)}$$

és átrendezve:

$$AB = \frac{(T_{A_1B_1} + T_{A_2B_2}) - (T_{A_1B_2} + T_{A_2B_1})}{(N/2)}$$

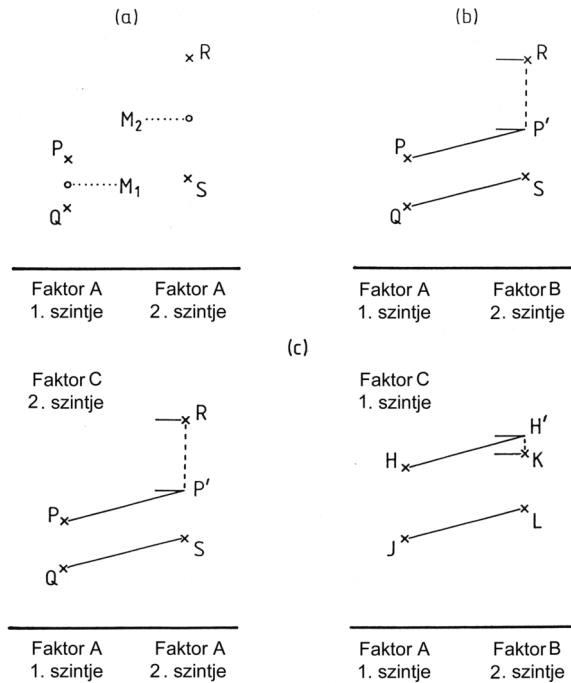
A háromtényezős interakció A, B és C faktorra a következőképpen határozható meg:

$$ABC = \frac{[(T_{A_1B_1C_1} - T_{A_2B_1C_1}) - (T_{A_1B_2C_1} - T_{A_2B_2C_1})] - [(T_{A_1B_1C_2} - T_{A_2B_1C_2}) - (T_{A_1B_2C_2} - T_{A_2B_2C_2})]}{(N/2)}$$

Vagy újból felhasználva a szimmetrikus formát:

$$ABC = \frac{[T_{A_1B_1C_1} - T_{A_2B_1C_1} - T_{A_1B_2C_1} - T_{A_1B_1C_2} + T_{A_2B_1C_2} + T_{A_1B_2C_2} + T_{A_2B_2C_1} - T_{A_2B_2C_2}]}{(N/2)}$$

Segítségül szolgálhat az interakciók megértésében a 2ⁿ típusú faktoriális kísérletek-nél a fő hatások és interakciók grafikonon történő bemutatása (12.6. ábra). Az 12.6(a) és 12.6(b) ábrán a P, Q, R és S jelű pontok képviselik a termést a vertikális skálán a négy kombinációra, azaz A₁B₁, A₁B₂, A₂B₁ és A₂B₂. Az 12.6(a) ábrán az A faktor fő hatása az M₂-M₁ különbsége, ahol M₂ az R és S átlaga, M₁ a P és Q átlaga. Az 12.6 (b) ábrán A és B közötti interakció R-P' különbsége, ahol PP' párhuzamos a QS-sel. Az 12.6(c) ábrán a két diagram mutatja az AB interakció értékeket elkülönítve a C két szintjére és az ABC három faktoros interakció az R-P' és K-H' közötti különbség, amely ebben az esetben egyenlő a két távolság összegével, mivel K-H' negatív (azaz, az AB interakciónak különböző előjele van a C két szintjére; egyik esetben a második szint kombinációja „jobb, mint várható lenne” és a másik esetben rosszabb).



12.6. ábra. A fő hatások, a 2-faktoros interakciók és 3-faktoros interakciók grafikus ábrázolása (Mead et al., 2003 alapján)

A 2ⁿ faktoriális kísérletekben az összes kezelés fő hatás és az interakció komponensei kiszámíthatók a kezelés összegekből. Van egy alternatív egyenrangú számítási módszer, amely a fő hatások és az interakciók algebrai meghatározásán alapul. Minden komponens SS-t úgy számítunk ki, mint az összeg-kombinációk négyzete, osztva az egységek összes számával a kísérletben. Tehát, egy 2³ faktoriális kísérletben, A, B és C faktorokkal és a nyolc kezelés r blokkjával a komponens SS-ek a fő hatásokra, a kettős és hármas interakciókra a következőképpen számítható ki:

$$A \text{ fő hatás SS} = \frac{(T_{A_1} - T_{A_2})^2}{8r} \text{ vagy } \frac{T_{A_1}^2 + T_{A_2}^2}{4r} - c.f.$$

$$\begin{aligned}
 \text{B fő hatás SS} &= \frac{(T_{B_1} - T_{B_2})^2}{8r} \text{ vagy } \frac{T_{B_1}^2 + T_{B_2}^2}{4r} - \text{c.f.} \\
 \text{AB interakció SS} &= \frac{(T_{A_1 B_1} - T_{A_1 B_2} - T_{A_2 B_1} + T_{A_2 B_2})^2}{8r} \text{ vagy} \\
 &\frac{T_{A_1 B_1}^2 + T_{A_1 B_2}^2 + T_{A_2 B_1}^2 + T_{A_2 B_2}^2}{2r} - \text{c.f.} - \text{SS(A)} - \text{SS(B)} \\
 \text{SS ABC} &= \\
 &\frac{(T_{A_1 B_1 C_1} - T_{A_2 B_1 C_1} - T_{A_1 B_2 C_1} - T_{A_1 B_1 C_2} + T_{A_2 B_1 C_2} + T_{A_1 B_2 C_2} + T_{A_2 B_2 C_1} - T_{A_2 B_2 C_2})^2}{(N/2)}
 \end{aligned}$$

vagy kivonva minden fő hatás SS-t és két faktoros interakció SS-t az összes kezelés SS-ből.

12.3.1. Kísérlet leírása

Az MTA Mezőgazdasági Kutatóintézetében Gyórfly Béla 1961 és 1967 között négy növénytermesztési tényező hatását és kölcsönhatását vizsgálta a kukorica szemtermésére. A vizsgált tényezők: A: Trágyázás, B: Növényszám, C: Fajta, D: Ápolás. A négy tényező mindegyikét két szinten vizsgálták kukorica monokultúra kísérletben. A 2^4 (poli)faktoriális kísérlet 16 kezeléskombinációt tartalmazott. A kísérletet négy ismétlésben állították be, 38 m²-es parcellákon. A kísérlet eredményeiről Gyórfly és munkatársai több tudományos publikációban közölnek adatokat (Gyórfly et al., 1969). A kísérlet módszertani szempontból napjainkban is rendkívül értékes és a 2^n faktoriális kísérlet értékelését ennek a kísérletnek az adataival mutatjuk be. A négytényezős, 2^4 faktoriális kísérlet varianciaanalízise kézi számolással rendkívül hosszadalmas, ezért az értékeést GenStat programmal mutatjuk be. Megjegyezzük, hogy az intézetben 1966-ban az értékelést O'sváth János vezetésével kézi számolással végezték, és ennek eredménye (tartamkísérletek dokumentációs füzetei) pontosan megegyezik az alábbiakban bemutatott számítógépes értékeléssel. A 16 kísérleti kezelés parcellánkénti random elrendezését és a kezeléskombinációkat az I–IV. ismétlésben a 12.7. ábra és a 12.14. táblázat mutatja.

IV.	14	16	6	8	12	3	10	4	15	7	13	5	11	1	9	2
-----	----	----	---	---	----	---	----	---	----	---	----	---	----	---	---	---

III.	12	10	15	13	14	9	16	11	6	2	8	1	4	5	3	7
------	----	----	----	----	----	---	----	----	---	---	---	---	---	---	---	---

II.	5	7	9	11	1	15	2	13	3	16	4	14	8	12	6	10
-----	---	---	---	----	---	----	---	----	---	----	---	----	---	----	---	----

I.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
----	---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----

I.–IV.: ismétlések jelölése; 1–16: parcellák jelölése.

12.7. ábra. A 2^4 faktoriális kísérlet elrendezése és kezeléskombinációi

A kísérleti parcellák kezeléskombinációja:

IV. ism.	ABcd	abCd	AbCD	aBCD	aBcd	ABcd	Abcd	ABcD	abcd	aBCd	abCD	AbCd	aBcD	abcd	AbcD	ABCD
III. ism.	aBcd	Abcd	abcd	abCD	ABcd	Abcd	abCd	aBcD	AbCD	ABCD	aBCD	abcd	ABcD	AbCd	ABcd	aBCd
II. ism.	AbCd	aBCd	AbcD	aBcD	abcd	abcd	ABCD	abCD	ABcd	abCd	ABcD	ABcd	aBCD	aBcd	AbCD	Abcd
I. ism.	abcd	ABCD	ABcd	ABcD	AbCd	AbCD	aBCd	aBCD	Abcd	Abcd	aBcD	aBcd	abCD	ABcd	abcD	abCd

Kezeléskombinációk:

1. abcd	5. AbCd	9. AbcD	13. abCD
2. ABCD	6. AbCD	10. Abcd	14. ABcd
3. ABcd	7. aBCd	11. aBcD	15. abcd
4. ABcD	8. aBCD	12. aBcd	16. abCd

Tényezők:

Trágyázás
Növényszám
Fajta
Ápolás

Szint:

a A
b B
c C
d D

12.14. táblázat. A 2⁴ faktoriális kísérlet kezelés szintjei

Vizsgált tényezők	Kezelések	
	0 szint	1 szint
Trágyázás (a, A)	Trágyázás nélkül	60 kg ha ⁻¹ N, 30 kg ha ⁻¹ P ₂ O ₅ , 60 kg ha ⁻¹ K ₂ O évenként
Növényszám (b, B)	20 000 tő ha ⁻¹ , 70 × 70 cm tenyészterület	40 000 tő ha ⁻¹ , 70 × 35 cm tenyészterület
Fajta (c, C)	Szabadelvirágzású fajta: Mindszentpusztai sárga	Hibrid: Martonvásári 1
Ápolás (d, D)	Megkéssett egyelés, 2 megkéssett kapálás	Vegyszeres gyomirtás + kézi kapálás

12.3.2. Varianciaanalízis GenStat programmal

A trágyázás (A), a növényszám (B), a fajta (C) és a növényápolás (D) a faktorok, mindegyik két szinttel (jelölésük 1 és 2). A termésadatok 100 kg ha⁻¹-ban vannak megadva. A számítógépbe a termésadatokot a kísérlet elrendezésének megfelelő sorrendben vittük be.

A varianciaanalízishez a General analysis of variance menüt választjuk. Y-variát a termés 1965, kezelés struktúra: A×B×C×D, Blokk struktúra: Ismétlés

ism.	A	B	C	D	Termés	ism.	A	B	C	D	Termés
1	1	1	1	1	17.45	3	1	1	1	1	18.3
1	2	2	2	2	56.76	3	2	2	2	2	49.14
1	2	2	1	1	24.07	3	2	2	1	1	32.055
1	2	2	1	2	57.27	3	2	2	1	2	51.36
1	2	1	2	1	22.38	3	2	1	2	1	21.125
1	2	1	2	2	51.46	3	2	1	2	2	43.47
1	1	2	2	1	31.195	3	1	2	2	1	25
1	1	2	2	2	49.605	3	1	2	2	2	45.63
1	2	1	1	2	50.41	3	2	1	1	2	39.015
1	2	1	1	1	29.015	3	2	1	1	1	23.535
1	1	2	1	2	39.755	3	1	2	1	2	37.425
1	1	2	1	1	26.36	3	1	2	1	1	26.505
1	1	1	2	2	38.605	3	1	1	2	2	33.775
1	2	2	2	1	39.505	3	2	2	2	1	27.575
1	1	1	1	2	38.845	3	1	1	1	2	33.805
1	1	1	2	1	31.06	3	1	1	2	1	20.14
2	1	1	1	1	18.095	4	1	1	1	1	30.51
2	2	2	2	2	61.925	4	2	2	2	2	61.82
2	2	2	1	1	25.61	4	2	2	1	1	18.66
2	2	2	1	2	53.87	4	2	2	1	2	41.69
2	2	1	2	1	24.455	4	2	1	2	1	23.945
2	2	1	2	2	47.145	4	2	1	2	2	35.41
2	1	2	2	1	37.32	4	1	2	2	1	28.51
2	1	2	2	2	49.065	4	1	2	2	2	34.525
2	2	1	1	2	48.095	4	2	1	1	2	47.925
2	2	1	1	1	18.405	4	2	1	1	1	9.42
2	1	2	1	2	45.235	4	1	2	1	2	34.53
2	1	2	1	1	26.615	4	1	2	1	1	18.265
2	1	1	2	2	32.28	4	1	1	2	2	36.16
2	2	2	2	1	26.83	4	2	2	2	1	26.18
2	1	1	1	2	41.6	4	1	1	1	2	30.205
2	1	1	2	1	26.63	4	1	1	2	1	25.205

Analysis of variance

Variate: Yield1965

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Ismetles stratum	3	413.35	137.78	5.67	
Ismetles.*Units* stratum					
A	1	406.65	406.65	16.74	<.001
B	1	637.47	637.47	26.24	<.001
C	1	188.80	188.80	7.77	0.008
D	1	5965.34	5965.34	245.55	<.001
A.B	1	20.52	20.52	0.84	0.363
A.C	1	2.44	2.44	0.10	0.753
B.C	1	82.57	82.57	3.40	0.072
A.D	1	564.75	564.75	23.25	<.001
B.D	1	26.03	26.03	1.07	0.306
C.D	1	23.09	23.09	0.95	0.335
A.B.C	1	1.71	1.71	0.07	0.792
A.B.D	1	0.33	0.33	0.01	0.908
A.C.D	1	1.89	1.89	0.08	0.782
B.C.D	1	31.87	31.87	1.31	0.258
A.B.C.D	1	0.44	0.44	0.02	0.893
Residual	45	1093.23	24.29		
Total	63	9460.48			

Tables of means

Variate: Yield1965

Grand mean 34.65

A	1	2		
	32.13	37.17		
B	1	2		
	31.50	37.81		
C	1	2		
	32.93	36.37		
D	1	2		
	25.00	44.31		
A	B	1	2	
1		29.54	34.72	
2		33.45	40.90	
A	C	1	2	
1		30.22	34.04	
2		35.65	38.70	

B	C	1	2			
1		30.91	32.08			
2		34.95	40.66			
A	D	1	2			
1		25.45	38.82			
2		24.55	49.80			
B	D	1	2			
1		22.48	40.51			
2		27.52	48.10			
C	D	1	2			
1		22.68	43.19			
2		27.32	45.42			
A	B	1	2			
	C	1	2	1	2	
1		28.60	30.48	31.84	37.61	
2		33.23	33.67	38.07	43.72	
A	B	1	2			
	D	1	2	1	2	
1		23.42	35.66	27.47	41.97	
2		21.54	45.37	27.56	54.23	
A	C	1	2			
	D	1	2	1	2	
1		22.76	37.68	28.13	39.96	
2		22.60	48.70	26.50	50.89	
B	C	1	2			
	D	1	2	1	2	
1		20.59	41.24	24.37	39.79	
2		24.77	45.14	30.26	51.06	
A	B	C	1	2	2	
		D	1	2	1	2
1	1		21.09	36.11	25.76	35.20
	2		24.44	39.24	30.51	44.71
2	1		20.09	46.36	22.98	44.37
	2		25.10	51.05	30.02	57.41

Standard errors of differences of means

Table	A	B	C	D
rep.	32	32	32	32
d.f.	45	45	45	45
s.e.d.	1.232	1.232	1.232	1.232

Table	A	A	B	A
	B	C	C	D
rep.	16	16	16	16
d.f.	45	45	45	45
s.e.d.	1.743	1.743	1.743	1.743

Table	B	C	A	A
	D	D	B	B
			C	D
rep.	16	16	8	8
d.f.	45	45	45	45
s.e.d.	1.743	1.743	2.464	2.464

Table	A	B	A
	C	C	B
	D	D	C
			D
rep.	8	8	4
d.f.	45	45	45
s.e.d.	2.464	2.464	3.485

Least significant differences of means (5% level)

Table	A	B	C	D
rep.	32	32	32	32
d.f.	45	45	45	45
l.s.d.	2.482	2.482	2.482	2.482

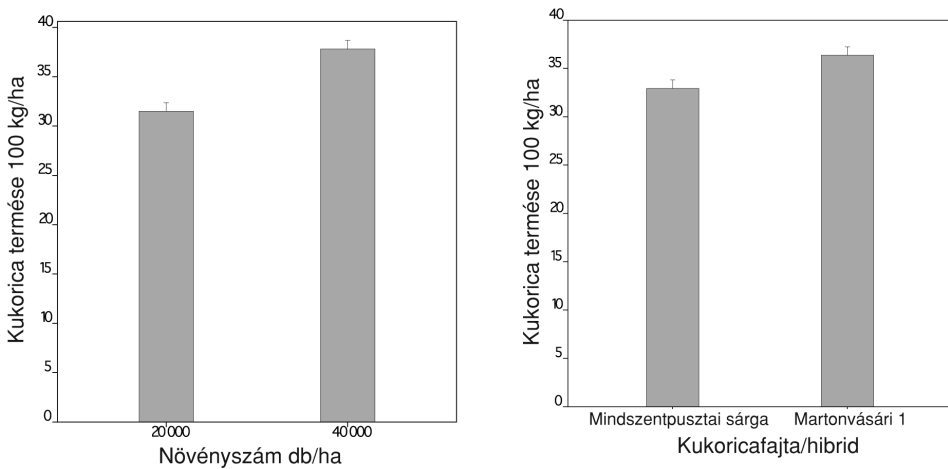
Table	A	A	B	A
	B	C	C	D
rep.	16	16	16	16
d.f.	45	45	45	45
l.s.d.	3.510	3.510	3.510	3.510

Table	B	C	A	A
	D	D	B	B
			C	D
rep.	16	16	8	8
d.f.	45	45	45	45
l.s.d.	3.510	3.510	4.964	4.964

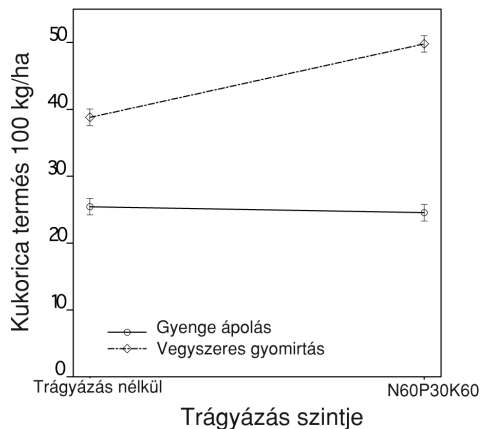
Table	A	B	A
	C	C	B
	D	D	C
			D
rep.	8	8	4
d.f.	45	45	45
l.s.d.	4.964	4.964	7.020

12.3.3. A kísérlet eredményeinek értékelése

A kísérletben mind a négy faktor, azaz a trágyázás (A), növényszám (B), fajta (C) és növényápolás (D) fő hatása szignifikáns volt 0.1%-os, illetve 1%-os (fajta) szinten. A kettős kölcsönhatások közül a trágyázás \times növényápolás (A \times D) kölcsönhatás volt szignifikáns, 0.1%-os szinten. A hármas kölcsönhatások nem voltak szignifikánsak. Az A \times D kölcsönhatás azt jelenti, hogy a trágyázás hatása függ a növényápolás szintjétől. Az optimális vs. alacsony növényszám 631 kg ha⁻¹, a hibrid termesztése a szabadelvirágzási fajtához viszonyítva 344 kg ha⁻¹ termésnövekedést eredményezett. A növényszám és fajta közvetlen fő hatása a kukorica termésére a 12.8. ábra szemlélteti. Mint a 12.9. ábra mutatja, a növényápolás alacsony szintjén nincs trágyahatás, míg jó növényápolásnál a trágyázás 1098 kg ha⁻¹ termésnövekedést eredményezett.



12.8. ábra. A növényszám és a fajta fő hatása a kukorica termésére a 2⁴ faktoriális kísérletben



12.9. ábra. A trágyázás \times növényápolás szignifikáns kölcsönhatása a kukorica termésére a 2⁴ faktoriális kísérletben. Trágyázás és növényápolás egyaránt két szinten

13. Időbeni mérések analízise

Ha egy tulajdonságot egy kísérletben ugyanazokon a kísérleti parcellákon különböző időpontokban mérünk, a kutató rendszerint a változás rátáját vizsgálja az egyik időszaktól a másikig. Fontos tehát meghatározni a kölcsönhatást (interakció) a kezelés és a mérési időpont között. Ehhez az egyes mérési időszakokra elvégzett varianciaanalízis önmagában nem elegendő. Általános megközelítés az ún. kombinált varianciaanalízis, vagyis a mérés idejét egy kiegészítő faktornak tekintjük a kísérletben, mintha alparcella, azaz a legkisebb kísérleti egység lenne (Gomez és Gomez 1984, Petersen 1994). A varianciaanalízis szerkezete az időbeni mérésre a randomizált blokk elrendezésnél megfelel a split-plot elrendezésnek (13.1. táblázat), míg a split-plot kísérletnél megfelel a standard kétszeresen osztott parcellás (split-split-plot) elrendezésnek (13.2. táblázat), amelyben az alparcella, illetve az al-alparcella a mérés (megfigyelés) ideje.

13.1. táblázat. Az időbeni mérések összevont varianciatáblázata RCB kísérleti elrendezésre

Variáció forrása	Szabadságfokok	Eltérésnégyzet összeg	Közepes négyzetes eltérés	Számított F-érték	Táblázati 5%	F-érték 1%
Ismétlés	$r-1$					
Kezelés (T)	$t-1$					
Hiba (a)	$(r-1)(t-1)$					
Megfigyelés ideje (P)	$p-1$					
T x P	$(t-1)(p-1)$					
Hiba (b)	$t(r-1)(p-1)$					
Összes	$rtp-1$					

r = ismétlések száma, t = kezelések száma, p = megfigyelések idejének száma

A módszer népszerűsége annak köszönhető, hogy az analízis viszonylag egyszerű. Ha a megfigyelési idő kvantitatív típusú (mint amilyen a növény kronológiai ideje, kifejezve pl. a vetéstől eltelt napok számában) és legalább három stádiumot felölel, akkor az interakció SS felbontása javasolt az ortogonális polinomok módszerével (Sváb 1981, Gomez és Gomez 1984).

Mead *et al.* (1993) rámutatott arra, hogy az osztott parcellás varianciaanalízis megközelítés az időt úgy kezeli, mint egy faktort, sokkal inkább, mint a mérések időbeni ismétlését. Továbbá az ismételt mérések analízisének számba kell venni a sorozatkorrelációt a mérések között ugyanazon a kísérleti parcellán. Egy újabb megközelítés (Piepho *et al.* 2003) az, hogy az ismételt méréseket vegyes modell (vagy más néven kevert modell, Jánossy *et al.* 1966) (mixed model) keretében analízáljuk, felhasználva a maximum likelihood (ML) vagy a korlátozott maximum likelihood (RML) mód-

szert. A vegyes modell (REML) egyaránt becsüli a fix kezelés-hatásokat és a random variancia komponenseket. Fontos előnye a korrelációs (variancia-kovariancia) mátrix rendelkezésre állása és a nem-komplett adatsorok hatékony analízise. *Pieopho et al.* (2004) javasolta az ismételt mérések vegyes modell (REML) megközelítését összehasonlítni a többváltozós varianciaanalízis (MANOVA) eredményével.

13.2. táblázat. Az időbeni mérések összevont varianciaanalízisének formája egy split-plot elrendezésre

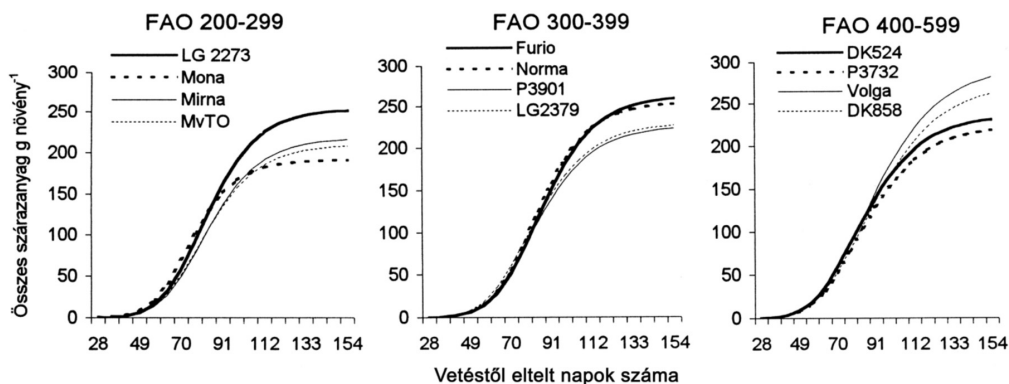
Variáció forrása	Szabadság fokok	Eltérés négyzet összeg	Közepes négyzetes eltérés	Számított F-érték	Táblázati F-érték	
					5%	1%
Ismétlés	$r-1$					
Főparcella kezelés(A)	$a-1$					
Hiba (a)	$(r-1)(a-1)$					
Alparcella kezelés (B)	$b-1$					
A x B	$(a-1)(b-1)$					
Hiba (b)	$a(r-1)(b-1)$					
Időbeni megfigyelések (C)	$p-1$					
A x C	$(a-1)(p-1)$					
B x C	$(b-1)(p-1)$					
A x B x C	$(a-1)(b-1)(p-1)$					
Hiba (c)	$ab(r-1)(p-1)$					
Összes	$rabp-1$					

A többszöri mérések célja gyakran a kezeléshatások időbeni mintázatának (a változás időbeni rátájának) feltárása különböző függvények illesztésével. Az időbeni mérések másik módszerénél, a regresszióanalízisnél a függvénykapcsolatot vizsgáljuk a reakció változó (Y) és a megfigyelési idő (X) között külön-külön minden kezelésre és összehasonlítjuk a regresszió paramétereit a különböző kezelésekből. Bármely modell alkalmas lehet, amennyiben megfelel a biológiai elvárásoknak és az adatok mintázatának. A modell paramétereit, amelyek sajátos derivált változók, ezt követően analizáljuk.

Különböző tenyészidejű kukorica hibridek növekedésének (száranyag termelésének) időbeni dinamikáját, Richards függvény illesztése alapján a 13.1. ábra szemlélteti.

13.1. Kísérlet leírása

A számítások itt egy RCB kísérlet adataira vonatkoznak, melynek célja a nitrogén műtrágyázás hatékonyságának vizsgálata rizs növényen (*Gomez és Gomez, 1984*).



13.1. ábra. Különböző tenyészidejű (FAO szám) kukorica hibridek növekedésének (szárazanyag produkciójának) időbeni dinamikája, Richards függvény illesztése alapján (Berzsenyi és Dang, 2000a)

Nyolc műtrágya-kezelést vizsgáltak négy ismétlésben. A talaj nitrogéntartalmának adatait a rizsnövény három növekedési stádiumában gyűjtötték be (13.3. táblázat).

13.3. táblázat. A talaj nitrogén tartalmának adatai, nyolc műtrágya kezelésre, RCB elrendezésben. A méréseket a növény három növekedési stádiumában (P_1 – P_3) végezték

Kezelések száma	Talaj nitrogén tartalma, %											
	I. ismétlés			II. ismétlés			III. ismétlés			IV. ismétlés		
	P_1	P_2	P_3	P_1	P_2	P_3	P_1	P_2	P_3	P_1	P_2	P_3
1	3.26	1.88	1.40	2.98	1.74	1.24	2.78	1.76	1.44	2.77	2.00	1.25
2	3.84	2.36	1.53	3.74	2.14	1.21	3.09	1.75	1.28	3.36	1.57	1.17
3	3.50	2.20	1.33	3.49	2.28	1.54	3.03	2.48	1.46	3.36	2.47	1.41
4	3.43	2.32	1.61	3.45	2.33	1.33	2.81	2.16	1.40	3.32	1.99	1.12
5	3.43	1.98	1.11	3.24	1.70	1.25	3.45	1.78	1.39	3.09	1.74	1.20
6	3.68	2.01	1.26	3.24	2.33	1.44	2.84	2.22	1.12	2.91	2.00	1.24
7	2.97	2.66	1.87	2.90	2.74	1.81	2.92	2.67	1.31	2.42	2.98	1.56
8	3.11	2.53	1.76	3.04	2.22	1.28	3.20	2.61	1.23	2.81	2.22	1.29

13.2. Varianciaanalízis lépései

Jelölje t a kezelések számát, r az ismétlések számát és p az időbeni mérések számát, melyet minden parcellán elvégeztek. Az adatelemzés lépésenkénti eljárása a következő:

1. lépés. Végezzük el a varianciaanalízist a megfigyelések p stádiumára, követve a standard varianciaanalízist a kísérleti elrendezésnek megfelelően. Példánkra az elrendezés az RCB. A varianciaanalízis eredményét $p = 3$ stádiumra a 13.4. táblázat mutatja.

13.4. táblázat. Három egyenkénti varianciaanalízis (RCB elrendezés) eredménye, mindegyik növekedési stádiumra, a 13.3. táblázat adatai alapján

Variáció forrása	Szabadságfok	Közepes négyzetes eltérés (MS)		
		P_1	P_2	P_3
Ismétlés	3	0.302412	0.019654	0.061458
Kezelés	7	0.211634**	0.415005**	0.063686 ^{ns}
Hiba	21	0.039843	0.039356	0.026011
Összes	31			

2. lépés. Vizsgáljuk meg a p hiba varianciák homogenitását. A varianciák homogenitásának vizsgálatára a χ^2 -tesztet (általánosan Bartlett-tesztnek hívják) alkalmazzuk, amikor kettőnél több varianciát tesztelünk. A lépésenkénti eljárás a következő:

(i) Mindegyik hiba varianciának (s^2) számítsuk ki a 10-es alapú logaritmus értékét. Ezután számítsuk ki a k varianciára az s^2 összeget és a $\log s^2$ összeget. Példánkra:

Növekedési stádium	s^2	$\log s^2$
P_1	0.039843	-1.39965
P_2	0.039356	-1.40499
P_3	0.026011	-1.58484
Összesen	0.105210	-4.38948

(ii) Számítsuk ki az összevont hiba varianciát:

$$s_p^2 = \frac{\sum_{i=1}^k s_i^2}{k} = \frac{0.105210}{3} = 0.03507$$

(iii) Legyen f a szabadságfokok száma mindegyik s_i^2 -re, számítsuk ki a χ^2 -értékeket a következő formulával:

$$\chi^2 = \frac{(2.3026)(f)(p \log s_p^2 - \sum \log s_i^2)}{1 + \frac{(p+1)}{3pf}}$$

Példánkban, a χ^2 -teszt a három hiba MS varianciájának homogenitás vizsgálatára a következő:

$$\chi^2 = \frac{(2.3026)(21)[3 \log 0.03507 - (-4.38948)]}{1 + \frac{(3+1)}{3(3)(21)}}$$

$$= 1.15$$

Összehasonlítjuk a számított χ^2 értéket a táblázati χ^2 értékkel, $(p-1) = 2$ szabadságfoknál. Mivel a számított χ^2 érték kisebb, mint a megfelelő táblázati χ^2 érték 5%-os szinten, amely 5.99, a variancia heterogenitása nem mutatható ki.

3. lépés. A variancia homogenitás tesztjének eredménye alapján alkalmazzuk a megfelelő varianciaanalízist. Ha a variancia heterogenitását mutattuk ki, választunk egy megfelelő adattranszformációt, amely stabilizálja a hiba varianciát és számítsuk ki az összevont varianciaanalízist a transzformált adatok alapján. Példánkban, mivel a hiba variancia heterogenitása nem volt kimutatható, nincs szükség adattranszformációra.

Ha a hiba variancia heterogenitása nem áll fenn, számítsuk ki az összevont varianciaanalízist az eredeti adatok alapján az összes p megfigyelésből. Példánkban, mivel a hiba variancia heterogenitása nem mutatható ki, a számítást az összevont varianciaanalízissel folytatjuk.

Az összevont varianciaanalízis általános formáját az időbeni mérésekre egy RCB elrendezésre az 13.1. táblázat mutatja. A szükséges számítási eljárás követi a split-plot elrendezés standard varianciaanalízisét t kezeléssel, mint főparcella kezelés és p időbeni megfigyeléssel, mint alparcella kezelés.

Az összevont varianciaanalízis eredményét példánkra az 13.5. táblázat mutatja. Az eredmény erősen szignifikáns interakciót mutat a kezelések és a növekedési stádiumok között, jelezve, hogy a kezeléshatások szignifikánsan változnak a különböző növekedési stádiumok között.

13.5. táblázat. Összevont varianciaanalízis az időbeni mérésekre (RCB elrendezés), a 13.3. táblázat adataira

Variáció forrása	Szabadság fok	Eltérés négyzet összeg	Közepes négyzetes eltérés	Számított F-érték	Táblázati F-érték	
					5%	1%
Ismétlés	3	0.845742	0.281914			
Kezelés (T)	7	1.265833	0.180833	4.94**	2.49	3.65
Hiba (a)	21	0.769492	0.036642			
Növekedési stádium (P)	2	52.042858	26.021429	715.88**	3.19	5.08
T × P	14	3.566442	0.254746	7.01**	1.90	2.48
Hiba (b)	48	1.744767	0.036349			
Összes	95	60.235134				

**Szignifikáns 1%-os szinten.

4. lépés. A pár összehasonlításhoz kövessük a split-plot standard eljárást az összevont varianciaanalízis eredménye alapján (3. lépés).

Példánkban, mivel az interakció a kezelés és a növekedési stádium között erősen szignifikáns, a pár összehasonlításokat a kezelések között csak ugyanabban a növekedési stádiumban végezhetjük. Két kezelés közötti különbség standard hibáját (szórását) ugyanabban a növekedési stádiumban a következőképpen számítjuk ki:

$$s_d = \sqrt{\frac{2[(p-1)E_b + E_a]}{ra}} = \sqrt{\frac{2[(2)0.036349 + 0.03642]}{(4)(3)}} = 0.13$$

ahol E_a a hiba (a) MS és E_b a hiba (b) MS az összevont varianciaanalízisben. Az $SzD_{5\%} = 0.2711$ (d.f. = 48).

A Duncan többszörös terjeselem teszt eredményét a kezeléscsoportok összehasonlítására mindegyik növekedési stádiumban a 13.6. táblázat tartalmazza.

5. lépés. Ha az interakció a kezelés és a megfigyelés ideje között szignifikáns, alkalmazzuk a megfelelő átlag-összehasonlítási módszert az interakció vizsgálatára. A módszer kiválasztása függ attól, hogy a megfigyelés ideje kvantitatív faktor-e.

Ha a megfigyelés ideje kvantitatív faktor (mint a növény kronológiai kora, kifejezve a vetés utáni napok számában, zárt térben nevelt rovarok kora stb.) és legalább három stádiumot tartalmaz, a következő eljárások közül bármelyik vagy mindkettő alkalmazható:

13.6. táblázat. Duncan többszörös terjedelem teszt (DMRT) a nyolc műtrágya kezeléscsoport összehasonlítására mindegyik növekedési stádiumra

Kezelések száma	Átlagos nitrogén tartalom, %		
	P ₁	P ₂	P ₃
1	2.95de	1.85de	1.33b
2	3.51a	1.96cde	1.30b
3	3.35ab	2.36b	1.44ab
4	3.25abcd	2.20bc	1.37ab
5	3.30abc	1.80e	1.24b
6	3.17bcd	2.14bcd	1.27b
7	2.80e	2.76a	1.64a
8	3.04cde	2.40b	1.39ab

Négy ismétlés átlaga. Egy oszlopon belül, az azonos betűvel jelzett átlagok 5%-os szinten nem különböznek szignifikánsan egymástól.

A) Az interakció SS felosztása, az időbeni megfigyelések trend összehasonlításának megfelelő módszere alapján.

B) Megfelelő regresszióanalízis alkalmazása, hogy becsüljük a függvénykapcsolatot a reakció (Y) és a megfigyelés ideje (X) között, külön-külön mindegyik kezelésre és összehasonlítsuk a regressziókat a kezeléseken keresztül.

Ha a megfigyelés ideje nem kvantitatív, vagy ha háromnál kevesebb megfigyelési stádiumot tartalmaz, az interakció SS felbontása a kezelések közötti összehasonlítások megfelelő sorozatán alapul, akár a kezelések, akár a növekedési stádiumok vagy mindkettő alapján.

Példánkban, mivel az interakció a kezelés és a megfigyelés ideje között szignifikáns, az interakció SS megfelelően felbontható. Minthogy a megfigyelés idejét a növény növekedési stádiumainak diszkrét sorozata reprezentálja (azaz, egyik megfigyelési stádium reprezentálja a kronológiai időt és egy másik reprezentálja a növény fiziológiai korát), a csoportok közötti összehasonlításokat kell elvégezni a kezelésekre vagy a növekedési stádiumokra.

13.7. táblázat. A kezelés \times növekedési stádium interakció SS felbontása az ANOVA-ban

Variáció forrása	Szabadságfok	Eltérés négyzet összeg	Közepes négyzetes eltérés	Számított F-érték	Táblázati F-érték	
					5%	1%
T \times P	14	3.566442	0.254746	7.01**	1.90	2.48
(T ₇ vs T ₁ -T ₆ és T ₈) \times P	(2)	(2.136675)	(1.068338)	29.39**	3.19	5.08
(T ₈ vs T ₁ -T ₆) \times P	(2)	(0.536011)	(0.268006)	7.37**	3.19	5.08
(T ₂ , T ₅ vs T ₁ , T ₃ , T ₄ , T ₆) \times P	(2)	(0.642872)	(0.321436)	8.84**	3.19	5.08
(T ₂ vs T ₅) \times P	(2)	(0.021700)	(0.010850)	< 1	–	–
(T ₁ , T ₃ , T ₄ , T ₆) \times P	(6)	(0.229184)	(0.038197)	1.05 ^{ns}	2.30	3.20
Hiba (b)	48	1.744767	0.036349			

Az interakció SS felbontására a csoportok közötti (between group) összehasonlítást választottuk a kezelésekre. Az eredményeket a 13.7. táblázat tartalmazza. Az eredmények mutatják, hogy a nyolc műtrágyakezelés a nitrogén tartalom időbeni változásának hasonlósága alapján négy csoportba sorolható: az első csoport tartalmazza a T₇-et, a második csoport a T₈-at, a harmadik tartalmazza a T₂-t és a T₅-öt és a negyedik tartalmazza a T₁, T₃, T₄ és T₆ kezeléseket.

13.3. Varianciaanalízis GenStat programmal

A számítógépes adatfájlban az Ismétlés, Kezelés, Növekedési stádium képezik a faktorokat, Nitrogén tartalom a változó. Az adatok analízise a következő lépésekben történt: (i) Növekedési stádiumonkénti (P) egytényezős RCB varianciaanalízis, (ii) Kombinált varianciaanalízis: kéttényezős osztott parcellás elrendezésnek megfelelően, kezelés a főparcella, növekedési stádium az alparcella, (iii) Kezelés \times növekedési stádium interakció SS felbontása a kontrasztok módszerével. Négy kontraszt számítását mutatjuk be, az ötödik kontraszt logikusan következik az 1-4. kontrasztból a 13.7. táblázatnak megfelelően.

Növekedési stádiumonkénti varianciaanalízis

1. Növekedési stádium (P_1)**Analysis of variance**

Variate: Nitrogen talaj

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Ismetles stratum	3	0.90724	0.30241	7.59	
Ismetles.*Units* stratum					
Kezeles	7	1.48144	0.21163	5.31	0.001
Residual	21	0.83671	0.03984		
Total	31	3.22539			

Duncan's multiple range test**Kezeles**

	Mean	
7	2.803	a
1	2.948	ab
8	3.040	abc
6	3.168	bc
4	3.253	bcd
5	3.303	cd
3	3.345	cd
2	3.508	d

2. Növekedési stádium (P_2)**Analysis of variance**

Variate: Nitrogentalaj

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Ismetles stratum	3	0.05896	0.01965	0.50	
Ismetles.*Units* stratum					
Kezeles	7	2.90504	0.41501	10.54	<.001
Residual	21	0.82649	0.03936		
Total	31	3.79049			

Duncan's multiple range test

Kezeles

	Mean	
5	1.800	a
1	1.845	ab
2	1.955	abc
6	2.140	bcd
4	2.200	cd
3	2.358	d
8	2.395	d
7	2.763	e

3. Növekedési stádium (P₃)**Analysis of variance**

Variate: Nitrogentalaj

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Ismetles stratum	3	0.18438	0.06146	2.36	
Ismetles.*Units* stratum					
Kezeles	7	0.44580	0.06369	2.45	0.053
Residual	21	0.54622	0.02601		
Total	31	1.17640			

Duncan's multiple range test

Kezeles

	Mean	
5	1.238	a
6	1.265	a
2	1.298	a
1	1.333	a
4	1.365	a
8	1.390	ab
3	1.435	ab
7	1.638	b

Kombinált varianciaanalízis:

Analysis of variance

Variate: Nitrogentalaj

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Ismetles stratum	3	0.84574	0.28191	7.69	
Ismetles.Kezeles stratum					
Kezeles	7	1.26583	0.18083	4.94	0.002
Residual	21	0.76949	0.03664	1.01	
Ismetles.Kezeles.Novekstadium stratum					
Novekstadium	2	52.04286	26.02143	715.87	<.001
Kezeles.Novekstadium	14	3.56644	0.25475	7.01	<.001
Residual	48	1.74477	0.03635		
Total	95	60.23513			

Kezelés × növekedési stádium interakció SS felbontása:

Contrast 1	1	1	1	1	1	1	-7	1
Contrast 2	1	1	1	1	1	1	0	-6
Contrast 3	1	-2	1	1	-2	1	0	0
Contrast 4	0	1	0	0	-1	0	0	0

Analysis of variance

Variate: Nitrogentalaj

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Ismetles stratum	3	0.84574	0.28191	7.69	

Ismetles.Kezeles stratum					
Kezeles	7	1.26583	0.18083	4.94	0.002
Contrast 1	1	0.35109	0.35109	9.58	0.005
Contrast 2	1	0.04552	0.04552	1.24	0.278
Contrast 3	1	0.02275	0.02275	0.62	0.440
Contrast 4	1	0.11760	0.11760	3.21	0.088
Residual	21	0.76949	0.03664	1.01	
Ismetles.Kezeles.Novekstadium stratum					
Novekstadium	2	52.04286	26.02143	715.87	<.001
Kezeles.Novekstadium	14	3.56644	0.25475	7.01	<.001
Contrast 1.Novekstadium	2	2.13667	1.06834	29.39	<.001
Contrast 2.Novekstadium	2	0.53601	0.26801	7.37	0.002
Contrast 3.Novekstadium	2	0.64287	0.32144	8.84	<.001
Contrast 4.Novekstadium	2	0.02170	0.01085	0.30	0.743
Residual	48	1.74477	0.03635		
Total	95	60.23513			



13.1. kép. A N-műtrágyázás hatása a kukorica növekedésére a martonvásári tartamkísérletben (N-kezelések balról jobbra: 0, 80, 160 és 240 kg ha⁻¹)

13.4. Kísérlet leírása

A N-műtrágyázás hatását tanulmányoztuk a kukorica hibridek növekedésére tartamkísérletben (13.1. kép). A N-műtrágya kezeléseket a következők voltak: 0, 80, 160 és 240 kg ha⁻¹. A négy ismétlésben, osztott parcellás (split-plot) elrendezésben beállított kísérlet főparcellája az N-kezelés, alparcellája a kukorica hibrid volt. A vizsgálatokat 2001 és 2002 években, három eltérő tenyészidejű hibriddel végeztük: *Mv TC 272* (FAO 280), *Mv 355 SC* (FAO 390) és *Maraton SC* (FAO 450). A növekedésanalízishez a növényminták vételét a vetéstől számított 28–35. napon (a kukorica 4-leveles fejlettségénél) kezdtük meg és a fiziológiai érésig folytattuk. Mindkét évben 8 mintavétel történt, 14 napos intervallumokban. Az alábbiakban a növénymagasság időbeni mérési adatait értékeljük (Berzsenyi, 2008).

13.5. Az ANOVA és az REML eredményének összehasonlítása

Az évenkénti adatokra elvégeztük a háromtényezős, split-split-plot elrendezésnek megfelelő varianciaanalízist, melyben az idő tényező volt az al-parcella. Az REML variancia komponens analízisben az autoregresszív és ante-dependence korrelációs modelleket választottuk, illetve a fix hatásokat Wald-teszttel értékeltük. A növényenkénti növénymagasság időbeni dinamikáját kísérleti kezelésenként regresszianalízissel is vizsgáltuk. A növénymagasság időbeni dinamikáját logisztikus függvényvel jellemeztük.

13.8. táblázat. A növénymagasság időbeni mérésének összehasonlító vizsgálata varianciaanalízissel (ANOVA) és vegyes modellel (REML)

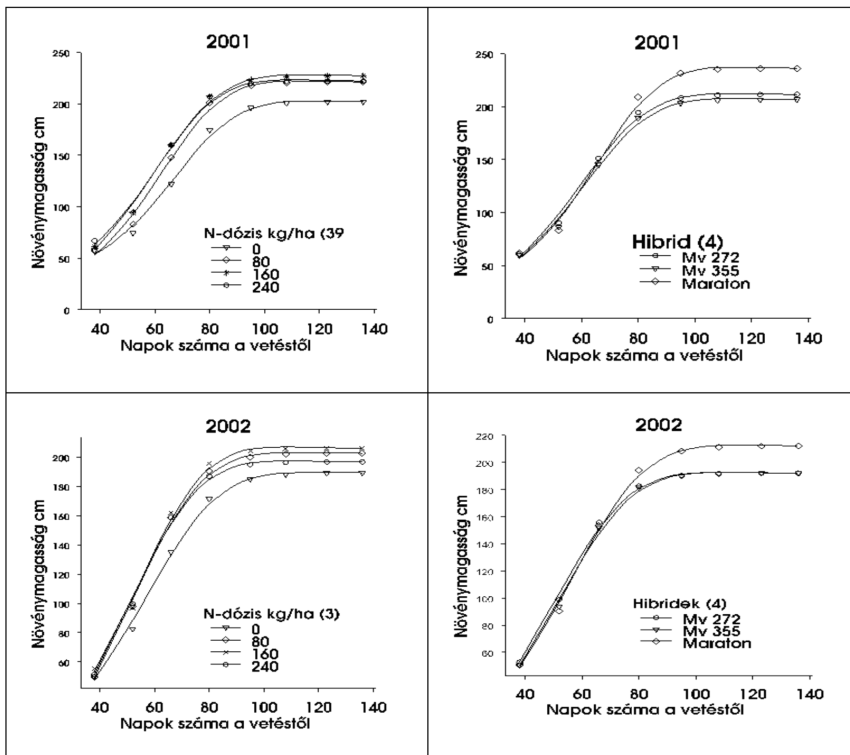
Variancia komponensek (fix hatások)	d.f.	2001		2002	
		F-teszt	Wald-teszt	F-teszt	Wald-teszt
————— Növénymagasság —————					
Idő (I)	7	421.3***	421.3***	329.1***	329.1***
Nitrogén (N)	3	67.0***	81.5***	21.0***	33.6***
Hibrid (H)	2	71.7***	65.8***	35.2***	27.3***
I x N	21	2.24***	2.72***	<1	1.33***
I x H	14	8.38***	7.70***	5.74***	4.45***
N x H	6	3.77***	3.46***	<1	<1
I x N x H	42	<1	<1	<1	<1

*** P = 0.1%-os szinten szignifikáns, ^{NS} = nem szignifikáns

A növénymagasságra a háromtényezős varianciaanalízis alapján a fő hatások mindkét évben szignifikánsak voltak (13.8. táblázat). Az MS értékek alapján legjelentősebb volt az idő tényező hatása, sorrendben a hibrid és a N-műtrágyázás hatása következett. Ugyancsak szignifikáns volt a N-műtrágyázás × idő és a hibrid × idő kölcsönhatás, vagyis az N-műtrágyázás és a hibrid hatása a növénymagasságra mintavételi időpontonként változott. 2001-ben szignifikáns volt a N-műtrágyázás × hibrid kölcsönhatás. A 13.8. táblázatban az ANOVA F-értéke mellett feltüntettük a vegyes modell

(REML) alapú *Wald-teszt* eredményét is. Megállapítható, hogy a fix hatások és kölcsönhatások jelentőségét, illetve szignifikanciáját tekintve az *F-teszt* és a *Wald-teszt* azonos eredményt adott.

A növénymagasság időbeni dinamikája S alakú, szigmoid típusú görbével jellemezhető (13.2. ábra). A mérési adatokhoz illesztett logisztikus függvény paraméterei pontosan leírták a növénymagasság időbeni mintázatát. Az értelmezett variancia (korigált R^2) értéke 97 és 99.5% között változott és a regresszió *F*-értéke $P = 0.1\%$ -os szinten szignifikáns volt. Az inflexiós pont és a növekedési ráta értékei jól mutatták, hogy kedvezőtlen N-ellátottságnál kisebb a növény relatív növekedési sebessége és az inflexiós pont 3–5 nappal később következik be. Ez utóbbi megmutatkozott a virágzási időpontokban is (Berzsényi, 2008).



13.2. ábra. A kukorica növénymagasságának időbeni dinamikája a N-műtrágyázástól és a hibridtől függően, a logisztikus függvény illesztése alapján (Berzsényi, 2008)

14. Kísérletsorozatok kombinált varianciaanalízise

A mezőgazdasági kutatási programoknak lényegében két fő célja van: (1) információt kapni, amely segít megmagyarázni annak a módját, ahogy a növények és állatok választ adnak a változó körülményekre (*alapkutatás*) és (2) információt kapni, amellyel megoldunk problémákat vagy javaslatokat adunk a termelőknek (*alkalmazott kutatás*).

Többnyire az alapkutatás magában foglalja egy-egy kísérlet sorozatát, amelyeket egymás után állítunk be. A kutató javasol egy hipotézist arra, hogy várhatóan hogyan válaszol egy biológiai rendszer a változásokra. Elvégez egy kísérletet, hogy adatokat kapjon a hipotézis teszteléséhez. Ennek a kísérletnek az eredménye alapján a kutató módosíthatja a hipotézist és lefuttat egy másik kísérletet, hogy tesztelje a módosított hipotézist. Ez a folyamat folytatódik egymás után, váltakozva a hipotézis és a kísérlet között, egészen egy megfelelő végpont eléréséig. A kísérleteket a befejezés után analizálják, és a folyamat folytonos természete miatt az adatok kombinált analízise két vagy több kísérletből ritkán szükséges.

A javaslatok (szaktanácsok) a termelőknek, másrésztől, ritkán lehetnek hatékonyak, ha egyetlen kísérlet adatainak analízisén alapulnak. A kísérletet bármennyire jól tervezték és hajtották végre, az csak egyetlen termőhelyről ad információt egyetlen évre vonatkozóan. Ez rendszerint nem elég információ a termesztési javaslatok megfogalmazásához. Például, egy növénynemesítő gyakran érdekelt új fajták kiválasztásában, amelyek jó termést adnak eltérő körülmények között a termőhelyek bizonyos körében. Az agronómus rendszerint érdekelt az új eljárásra adott reakcióban eltérő talajtípusokon. Ahhoz, hogy megfeleljünk ezeknek a céloknak, szükséges, hogy a kísérleteket több éven (legalább három éven) keresztül, illetve több termőhelyen végezzük, amelyre a javaslatok irányulnak.

Ahhoz, hogy egy alkalmazott kutatási program maximális értékét megkapjuk, számos kísérlet eredményének kombinált analízise válik fontossá. Egy kombinált analízis információt adhat a kezelés-hatásokról és a kezeléshatásokban levő különbségekről, a reakció változásáról a talajban és az időjárásban levő különbségekre az egész termőtájon.

Általánosságban, a kombinált becslések csak akkor lesznek érvényesek, ha a termőhelyek, amelyeken a kísérleteket végezték, az összes lehetséges hely random mintái az adott termőtájon. Ha minden lehetséges helynek egyenlő az esélye, hogy a kísérleti helyek közé kerüljön, akkor a kombinált becslések torzítatlanok lesznek arra a területre, amelyre javasolják. A kutatási programban gyakran nem praktikus ragaszkodni a szigorú random kiválasztási követelményhez. Olyan termőhelyeket kell használni költség és más szempontok miatt, amelyek rendelkezésre állnak, könnyen elérhetők stb. Ebben az esetben az itt bemutatott módszerek még használhatók, azonban az eredményeket a korlátozott hely kiválasztás figyelembe vételével kell interpretálni.

A kísérletsorozat értékelésének alapjai

Kísérletsorozatról akkor beszélünk, ha ugyanazt a kísérletet különböző helyeken, illetve időpontokban többször megismételjük. Általában két kérdésre várunk választ: (i) mekkora kezeléskülönbségek vannak az eltérő kísérleti körülmények átlagában, vagyis mekkora a kezelések főhatása; (ii) a kezeléskülönbségek változnak-e az eltérő kísérleti körülmények hatására, vagyis van-e kezelés \times kísérleti körülmény kölcsönhatás. A kísérletsorozat értékelésekor általános irányelvnek tekinthető, hogy a kísérletenkénti kezelésátlagokból végezzük a számítást. Formailag minden egyes kísérlet megannyi ismétlést, blokkot jelent (Sváb, 1981). A kísérletsorozat varianciaanalízisének szerkezetét egytényezős véletlen blokkelrendezésű kísérletekben az 14.1. táblázat mutatja.

14.1. táblázat. Kísérletsorozat varianciaanalízisének szerkezete, ha a helyek, illetve az időpontok képezik az ismétléseket

Tényező	df	Tényező	df
Összes	hv-1	Összes	iv-1
Hely	h-1	Idő	i-1
Kezelés	v-1	Kezelés	v-1
Kezelés \times Hely	(v-1)(h-1)	Kezelés \times Idő	(v-1)(i-1)

A véletlen blokkelrendezés és az ilyen jellegű kísérletsorozat közötti különbség az, hogy a kísérleti hibátényezőt a kezelés \times hely, illetve a kezelés \times idő kölcsönhatás helyettesíti. A kölcsönhatás számszerű értéke magában foglalja ugyan a kísérleti hibát, egy része azonban valódi, értelmezhető kezelés \times kísérleti körülmény kölcsönhatás lehet.

Ha ismerjük a kísérletenkénti hiba szórásnégyzetet vagy valamilyen módon becsülni tudjuk, akkor a kísérletsorozat varianciaanalízisét az összevont hiba (pooled error) tényezővel kiegészítjük. Az összevont hibával ellenőrizhetjük, hogy a kölcsönhatás számszerű értékét csak kísérleti hibának tulajdonítsuk vagy valódi kölcsönhatás is van. A teljes varianciaanalízis ekkor a 14.2. táblázat szerint alakul.

14.2. táblázat. Kísérletsorozat varianciaanalízisének szerkezete, ha az egyes kísérletekben a kísérleti hiba meghatározható

Tényező	df		Tényező	df
Összes	hv-1		Összes	iv-1
Hely	h-1		Idő	i-1
Kezelés	v-1		Kezelés	v-1
Kezelés \times Hely	(v-1)(h-1)		Kezelés \times Idő	(v-1)(i-1)
Összevont hiba	Σ (hiba df)		Összevont hiba	Σ (hiba df)

Az összevont hiba MS-t a kísérletenkénti hiba MS összevonásából kapjuk. Az összevont hiba MS szabadságfokainak száma mindig a kísérletenkénti hiba MS szabadságfokainak összege.

Ha a kísérletek elrendezése miatt több hibátényező van, minden hibátényezőt külön vonunk össze. Így pl. osztott parcellás elrendezésű kísérletek összevonásakor a hiba (a) MS és hiba (b) MS értékeket külön-külön vonjuk össze (14.3. táblázat).

14.3. táblázat. Kísérletsorozat varianciaanalízisének szerkezete, ha több hibátényező van (osztott parcellás elrendezésű kísérletsorozat összevonása)

Tényező	df
Összes	hab-1
Hely	h-1
A	a-1
A × Hely	(a-1)(h-1)
Összevont hiba (a)	$\sum(\text{Hiba}(a)\text{df})$
B	b-1
B × Hely	(b-1)(h-1)
A × B	(a-1)(b-1)
A × B × Hely	(a-1)(b-1)(h-1)
Összevont hiba (b)	$\sum(\text{Hiba}(b)\text{df})$

Kísérletsorozat értékelésekor az összevont hiba MS értékét az egyes kísérletek hiba MS értékéből határozzuk meg. Nem egyértelmű azonban ennek kiszámítása. Az összevont hiba szórásnégyzet (s^2) alapképletében kísérletenként r -rel, azaz az ismétlések számával elosztott hiba szórásnégyzeteket összegezzük, és az összeget osztjuk a kísérletek számával, k -val:

$$s^2 = \frac{1}{k} \left[\frac{s_1^2}{r_1} + \frac{s_2^2}{r_2} + \dots + \frac{s_k^2}{r_k} \right]$$

Ha az ismétlések száma minden kísérletben azonos, akkor az összevont hiba szórásnégyzet képlete:

$$s^2 = \frac{s_1^2 + s_2^2 + \dots + s_k^2}{kr}$$

Ha az s^2 értékek kísérletenként nagyon eltérőek, vagyis a kísérletsorozat nagyon inhomogén hibaszórású, akkor elég problematikus az értékelés. Gyakorlatilag legcélszerűbb eljárás az, hogy a kiugróan nagy hibaszórású kísérletet, főleg, ha a kezelések eredményének a sorrendje a többi kísérletnek ellentmond, az értékelésből kizárjuk (Sváb, 1981).

14.1. Kísérletsorozat: évek analízise

A növény teljesítménye függ a genotípustól, a környezettől, amelyben a növényt termesztjük, valamint a genotípus és a környezet közötti kölcsönhatástól (interakciótól). A genotípust és a környezet néhány faktorát, mint a műtrágya dózist, a növény-számot és a károsítók elleni védekezést szabályozni tudja a kutató. A környezet más faktorai azonban, mint a napsugárzás, csapadék és bizonyos talajtulajdonságok, általában fixek és módosításuk bonyolult egy adott helyen és tenyészidőszakban. Tehát egy kutató egy kísérlettel egyetlen helyen csak a szabályozható faktorokat változtathatja és értékelheti, azonban nem a környezeti faktorokat, amelyek szabályozásán kívül esnek.

A nem-szabályozható környezeti faktorok hatása a növény teljesítményére legalább olyan fontos, ha nem fontosabb, mint a szabályozható faktorok hatása; és hatásaik értékelése és kvantifikálása éppen olyan lényeges. Mivel a nem-szabályozható faktorokról feltételezzük, hogy változnak szezonálisan és a termőhellyel, és mivel ezek a változások mérhetőek, hatásuk a kezelések teljesítményére értékelhető. A növénytermesztési kutatásban a legáltalánosabb módja annak, hogy értékeljük a nem-szabályozható környezeti faktorok hatásait a növény reakcióra az, hogy megismételjük a kísérletet több helyen, illetve több éven keresztül.

A több helyen és több éven keresztül végzett kísérleteket négy fő csoportba sorolhatjuk céljaik szerint. Ezek a következők:

- *Előkísérletek* (preliminary evaluation experiments), melyeket azért tervezünk, hogy – nagy számú új technológiából – azonosítsunk néhány technológiát, melyeknek következetesen kiváló a teljesítményük abban a térségben (helyen), ahol kifejlesztették.
- *Technológiai adaptációs kísérletek*, melyeket azért terveztek, hogy meghatározzák – korábban az előkísérletekben azonosított – néhány kiváló technológia geográfiai alkalmazkodóképességének tartományát.
- *Tartamkísérletek*, melyeknek célja, hogy jellemezzenek egy új technológiát, figyelemmel a produktivitásra gyakorolt hosszú távú hatásukra.
- *Reakciót előrejelző kísérletek*, melyeknek célja, hogy meghatározzunk egy függvénykapcsolatot a növényreakció és néhány környezeti faktor között és előrejelezzük a produktivitást a környezet tág tartományára.

Az elsődleges célja ennek a fejezetnek, hogy leírjuk a kísérleteknek ezt a négy típusát és a megfelelő adatelemzést, amely kritikus az előírt célok megválaszolásában. Feltételezzük, hogy mindegyik kísérlet megfelelő analízise elkészült és csupán több kísérlet kombinált analízisét mutatjuk be. Hangsúlyozzuk a specifikus statisztikai eljárások használatát, melyek megfelelnek a kísérlet specifikus céljainak.

Előkísérletek

A növénytermesztési technológia fejlesztése rendszerint magában foglalja az eliminációs (kizárási) eljárások sorozatát, kezdve az új technológiák nagy számával és befejezve kevés kiváló technológiával, melyeket javasolnak gyakorlati felhasználásra. A kizárási folyamatok többségét egy helyen végzik, ahol az új technológiákat kifejlesztették és összegyűjtötték.

Az előkísérlet része a kizárási eljárásnak egyetlen helyen, ahol a kísérleteket ismétlésekben állították be és megismételték több tenyészidőben. Az elsődleges cél az, hogy azonosítsunk egy technológiát (vagy technológiákat), melyek következetesen kiválóbbak legalább egy helyen, ahol a következetes kiválóságot úgy határozzák meg, mint élvonal teljesítmény legalább két tenyészidőszakban (évben).

Évek analízise

Az előrejelezhető mintázat hiánya miatt az éveket általában random változónak tekintjük: a kiváló technológiának következetes kiválóságot kell mutatni több éven át. Tehát az évenkénti adatok kombinált analízisének elsődleges célja olyan technológiák azonosítása, melyek átlagos hatása az évek folyamán stabil és magas. A kezelés és az év közötti interakciónak itt kisebb az agronómiai jelentősége.

14.1.1. Kísérlet leírása

Az évenkénti adatok kombinált analízisét egy olyan kísérlettel mutatjuk be, melyben hét rizsfajtát vizsgáltak két éven keresztül RCB elrendezésben és három ismétlésben (Gomez és Gomez, 1984). A szemtermés adatokat a 14.4. táblázat és az évenkénti varianciaanalízist a 14.5. táblázat mutatja. A fajtaátlagok mindegyik évben a 14.6. táblázatban találhatók.

14.4. táblázat. Hét rizsfajta szemtermése RCB elrendezésű kísérletben, három ismétlésben, két egymást követő évben

Fajta sorszám	1. év				2. év			
	ism. 1	ism. 2	ism. 3	Összes	ism. 1	ism. 2	ism. 3	Összes
1	3.036	4.177	3.884	11.097	1.981	3.198	3.726	8.905
2	1.369	1.554	1.899	4.822	3.751	2.391	3.714	9.856
3	5.311	5.091	4.839	15.241	3.868	3.134	3.487	10.489
4	2.559	3.98	3.853	10.392	2.729	2.786	2.598	8.113
5	1.291	1.705	2.13	5.126	3.222	3.554	2.452	9.228
6	3.452	3.548	4.64	11.640	4.25	4.134	3.339	11.723
7	1.812	2.914	0.958	5.684	3.336	4.073	2.885	10.294
Összes				64.002				68.608

14.5. táblázat. A hét fajtát és három ismétlést tartalmazó RCB kísérlet évenkénti varianciaanalízise az 14.4. táblázat adatai alapján

Variáció forrása	Szabadság fok	Eltérés négyzet összeg	Közepes négyzetes eltérés	Számított F-érték	Táblázati F-érték	
					5%	1%
1. év						
Ismétlés	2	1.3855	0.6927			
Fajta	6	31.8569	5.3095	16.11**	3.00	4.82
Hiba	12	3.9559	0.3297			
2. év						
Ismétlés	2	0.0970	0.0485			
Fajta	6	2.7980	0.4663	1.16 ^{ns}	3.00	4.82
Hiba	12	4.8435	0.4036			

**szignifikáns 1%-os szinten, ^{ns}nem szignifikáns.

14.6. táblázat. Hét rizsfajta átlagtermése két vizsgált évben, az 14.4. táblázat adataiból számítva

Fajta sorszám	Átlagtermés t ha ⁻¹			
	1. év	2. év	Átlag	Különbség
1	3.699	2.968	3.334	-0.731
2	1.607	3.285	2.446	1.678
3	5.080	3.496	4.288	-1.584
4	3.464	2.704	3.084	-0.760
5	1.709	3.076	2.392	1.367
6	3.880	3.908	3.894	0.028
7	1.895	3.431	2.663	1.536
Átlag	3.048	3.267	3.158	0.219

14.1.2. Varianciaanalízis lépései

Az évenkénti adatok kombinált varianciaanalízisének lépésenkénti eljárásai a következők:

1. lépés. Szerkesszük meg a kombinált varianciaanalízis táblázatát. Példánkban, az RCB elrendezésre, az évek kombinált varianciaanalízisét a 14.7. táblázat mutatja. A kezelés-hatás szignifikancia vizsgálatának F-tesztjében a hiba tag az év × kezelés interakció MS.

2. lépés. Számítsuk ki a különböző SS és MS értékeket, melyek szükségesek a kombinált varianciaanalízishez. Az adatösszegek (total) alapján a különböző SS értékek kiszámítása követi a standard eljárást:

$$C.F. = \frac{\left(\sum_{i=1}^s G_i \right)^2}{srt}$$

$$S\ SS = \sum_{i=1}^s \frac{G_i^2}{rt} - C.F.$$

$$T\ SS = \sum_{j=1}^t \frac{T_j^2}{sr} - C.F.$$

$$S \times T\ SS = \sum_{i=1}^s \sum_{j=1}^t \frac{(ST)_{ij}^2}{r} - C.F. - S\ SS - T\ SS$$

ahol G_i az i -edik év főösszege, $(ST)_{ij}$ a j -edik kezelés összege az i -edik évben, és T_j a j -edik kezelés összege az összes s évben.

14.7. táblázat. A kombinált varianciaanalízis szerkezet s éven át, t kezelést és r ismétlést tartalmazó RCB elrendezésű kísérlet alapján

Variáció forrása	Szabadságfok	Közepes négyzetes eltérés	Számított F-érték
Év (Y)	$s-1$	Y MS	Y MS/R MS
Ismétlés éven belül	$s(r-1)$	R MS	
Kezelés (T)	$t-1$	T MS	T MS/ Y \times T MS
Y \times T	$(s-1)(t-1)$	Y \times T MS	Y \times T MS/E MS
Összevont hiba	$s(r-1)(t-1)$	E MS	
Összes	$srt-1$		

Példánkra a számítások a következők:

Ismétlés éven belül Y SS = 1.38549 + 0.09698 = 1.48247

Összevont hiba SS = 3.95588 + 4.84346 = 8.79934

$$C.F. = \frac{(64.002 + 68.608)^2}{42} = 418.70029$$

$$Y\ SS = \frac{(64.002)^2 + (68.608)^2}{(3)(7)} - 418.70029 = 0.50512$$

$$T\ SS = \frac{\left[(20.002)^2 + (14.678)^2 + \dots + (15.978)^2 \right]}{(3)(2)} - 418.70029 = 19.15891$$

$$Y \times T\ SS = \frac{\left[(11.097)^2 + (4.822)^2 + \dots + (10.294)^2 \right]}{3} - 418.70029 - 0.50512 - 19.15891 = 15.496$$

Az MS értékeket mindegyik SS-re a szokásos módon számítjuk ki. Az eredményeket az 14.8. táblázat tartalmazza.

3. Lépés. Teszteljük az s hiba MS homogenitását az évenkénti varianciaanalízisből az F-teszttel vagy a khi-négyzet teszttel. Példánkban az F-tesztet használjuk:

$$F = \frac{\text{nagyobbik hiba MS}}{\text{kisebbik hiba MS}} = \frac{0.40362}{0.32966} = 1.22$$

A számított F-érték kisebb, mint a megfelelő táblázati F-érték (2.69) $f_1=f_2=12$ szabadságfoknál, 5%-os szignifikancia szintnél. Tehát a homogén variancia hipotézisét nem utasítjuk vissza.

14.8. táblázat. Két éven keresztül folytatott RCB kísérlet kombinált varianciaanalízise, a 14.4. és 14.5. táblázat adataiból számítva

Variáció forrása	Szabadság fok	Eltérés négyzet	Közepes négyzetes eltérés	Számított F-érték	Táblázati F-érték	
					5%	1%
Év	1	0.5051	0.5051			
Ismétlés éven belül	4	1.4825	0.3706			
Fajta (V)	6	19.1589	3.1932	1.24 ^{ns}	4.28	8.47
Y × V	6	15.4960	2.5827	7.04 ^{**}	2.51	3.67
Összevont hiba	24	8.7993	0.3666			
Összes	41	45.4418				

**szignifikáns 1%-os szinten, ^{ns}nem szignifikáns.

4. Lépés. Az F-értékek kiszámítását a 14.7. táblázat formulái szerint végezzük. Mivel az ismétlések éven belül SS szabadságfoka csupán 4, nem számítunk ki F-értéket az év fő hatásának tesztelésére. A másik két F-értéket a következőképpen számítjuk ki:

$$F(V) = \frac{V \text{ MS}}{Y \times V \text{ MS}} = \frac{3.19315}{2.58267} = 1.24$$

$$F(Y \times V) = \frac{Y \times V \text{ MS}}{\text{Összevont hiba MS}} = \frac{2.58267}{0.36664} = 7.04$$

A számított F-értékeket összehasonlítva a megfelelő táblázati F-értékekkel a szokásos módon, mutatja, hogy F(V) nem szignifikáns és F(Y × V) szignifikáns 1%-os szinten. Tehát az interakció hatás a fajta és az év között erősen szignifikáns, azonban az átlagos fajta hatás nem szignifikáns.

5. Lépés. Ha a variancia homogenitásának tesztje a 4. lépésben szignifikáns, akkor az összevont hiba SS felbontása válik szükségessé. Példánkban, mivel a hiba variancia homogenitását bizonyítottuk, az összevont hiba SS felbontása nem szükségszerű. Azaz, az Y × V hatás, valamint bármely komponense tesztelhető, felhasználva az összevont hiba MS-t, mint hiba tagot.

6. Lépés. Ha a kezelés × év interakció szignifikáns, vizsgáljuk meg nagyságát a kezelések átlagos hatásához viszonyítva. Ha az interakció hatás kicsi az átlagos hatáshoz

viszonyítva, a kezelések rangsora a különböző években várhatóan stabil (pl. A kezelés jobban teljesít mint B kezelés minden évben, jóllehet bizonyos években a különbség kissé nagyobb, mint más években) és az interakció figyelmen kívül hagyható.

Másrésztől, ha az interakció hatás viszonylag nagy és a kezelések sorrendje változik az évek folyamán (pl. amikor A kezelés jobban teljesít mint B kezelés bizonyos években és nem jobban, vagy akár rosszabbul más években), ekkor az interakció természetének vizsgálata hasznos lehet.

Példánkban, az interakció oly nagy, hogy a fajta átlagos hatása nem szignifikáns. Az évek közötti átlagos különbség vizsgálatok a hét fajta mindegyikére (14.6. táblázat, utolsó oszlop) a fajták három csoportja azonosítható. Az első csoportba tartozik a 2, 5 és 7 fajta, melyeknek nagyobb a termése a 2. évben, mint az 1. évben; a második csoport az 1, 3 és 4 fajtából áll, melyek jobban teljesítettek az 1. évben, mint a 2. évben; és a harmadik csoport a 6. fajtát foglalja magában, amely hasonló termést adott mindkét évben.

Jóllehet világos, hogy egy következetesen kiváló fajta nem azonosítható ebben a kísérletben, a jövőbeni kísérletekben a hangsúlyt valószínűleg a 6. fajtára kellene helyezni, amely magas fokú következetességet mutatott; és a 3. fajtára, amelyik a legnagyobb átlagtermést (4.288 t ha^{-1}) adta a vizsgált években és rendkívül jó volt a teljesítménye az első évben (5.08 t ha^{-1}).

14.1.3. Varianciaanalízis GenStat programmal

Split-plot design

Treatment Structure: Year * Variety

Block Structure: Year/Replication

Whole plot: Year

Sub-plot: Variety

Analysis of variance

Variate: RiceYield

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Year.Replication stratum					
Year	1	0.5051	0.5051	1.36	0.308
Residual	4	1.4825	0.3706	1.01	
Year.Replication.Variety stratum					
Variety	6	19.1589	3.1932	8.71	<.001
Year.Variety	6	15.4960	2.5827	7.04	<.001
Residual	24	8.7993	0.3666		
Total	41	45.4418			

Tables of means

Variate: RiceYield

Grand mean 3.157

Year	1	2
	3.048	3.267

Variety	1	2	3	4	5	6	7
	3.334	2.446	4.288	3.084	2.392	3.894	2.663
Year	Variety	1	2	3	4	5	6
1		3.699	1.607	5.080	3.464	1.709	3.880
2		2.968	3.285	3.496	2.704	3.076	3.908
Year	Variety	7					
1		1.895					
2		3.431					

Standard errors of differences of means

Table	Year	Variety	Year Variety
rep.	21	6	3
s.e.d.	0.1879	0.3496	0.4948
d.f.	4	24	28
Except when comparing means with the same level(s) of			
Year			0.4944
d.f.			24

Least significant differences of means (5% level)

Table	Year	Variety	Year Variety
rep.	21	6	3
l.s.d.	0.5216	0.7215	1.0135
d.f.	4	24	28
Except when comparing means with the same level(s) of			
Year			1.0204
d.f.			24

Year 1

RCB Design

Analysis of variance

Variate: RiceYield

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Replication stratum	2	1.3855	0.6927	2.10	
Replication.*Units* stratum					
Variety	6	31.8569	5.3095	16.11	<.001
Residual	12	3.9559	0.3297		
Total	20	37.1983			

Tables of means

Variate: RiceYield

Grand mean 3.05

Variety	1	2	3	4	5	6	7
	3.70	1.61	5.08	3.46	1.71	3.88	1.89

Standard errors of differences of means

Table	Variety
rep.	3
d.f.	12
s.e.d.	0.469

Year 2

RCB Design

Analysis of variance

Variate: RiceYield

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Replication stratum	2	0.0970	0.0485	0.12	
Replication.*Units* stratum					
Variety	6	2.7980	0.4663	1.16	0.390
Residual	12	4.8435	0.4036		
Total	20	7.7384			

Tables of means

Variate: RiceYield

Grand mean 3.27

Variety	1	2	3	4	5	6	7
	2.97	3.29	3.50	2.70	3.08	3.91	3.43

Standard errors of differences of means

TableVariety

rep.	3
d.f.	12
s.e.d.	0.519

Least significant differences of means (5% level)

TableVariety

rep.	3
d.f.	12
l.s.d.	1.130

Megjegyezzük, hogy a két év kombinált varianciaanalízisének a GenStat program futtatásakor a fajta (Variety) F-értéke szignifikáns, mivel kiszámításakor a program automatikusan a fajta MS értékét a maradék MS értékhez viszonyította. A korrekt értékelést a 14.8. táblázat mutatja, melyben a fajta MS-t az $Y \times V$ interakció MS-sel osztottuk, és ezáltal a fajta F-értéke nem szignifikáns.

14.2. Technológiai adaptációs kísérletsorozat: különböző helyek analízise

A technológiai adaptációs kísérleteket azért tervezik, hogy becsüljék az új természeti technológiák adaptálhatóságának tartományát. A technológia adaptálhatóságát egy adott helyen más – ugyanazon a helyen vizsgált – technológiák feletti kiválóságában határozzák meg. Az ilyen kísérlet elsődleges célja, hogy javaslatot tegyünk egy vagy több új eljárásra, amelyek a farmerek által jelenleg használt eljárások fejlesztései vagy azok helyettesítői. A technológia adaptációs kísérletnek három elsődleges sajátossága van:

1. Az elsődleges cél az, hogy meghatározzuk egy technológia adaptálhatóságának tartományát. Egy sajátos technológiáról akkor mondjuk, hogy adaptálódott egy sajátos helyhez, ha a csúcsteljesítményt nyújtók között van azon a helyen. Továbbá, alkalmazkodóképességének tartománya magában foglalja a teszt helyek által reprezentált területeket, amelyekben a technológia kiváló teljesítményt mutatott.

2. A teszthelyek kiválasztásának elsődleges alapja egy geográfiai terület reprezentálása. A specifikus helyeket a technológiai adaptáció kísérletekre célszerűen választjuk ki, hogy reprezentálják a geográfiai területet vagy a környezetnek egy tartományát, amelyben a technológia adaptálhatóságának tartományát azonosították. Az ilyen területeket nem random választjuk ki. A legtöbb esetben, ezek a teszthelyek kutató állomások különböző geográfiai helyeken. Amikor azonban ilyen kutató állomások nem állnak rendelkezésre néha a farmerek tábláit használják teszt helyeknek.

3. A kezelések főként ígéretes technológiákból állnak. Csak azokat a technológiákat tesztelik, amelyek kiválóan mutatkoztak legalább egy környezetben (pl. kiválasztás egy előkísérletből). Ezen kívül, a vizsgált kezelések közül legalább egy rendszerint *kontroll*, amely vagy egy nem-technológia kezelést (mint amilyen a trágyázás nélküli vagy védekezés nélküli kontroll) vagy egy jelenleg használt technológiát (mint amilyen a helyi fajta) reprezentál.

A technológiai adaptációs kísérletekre két gyakori példa a fajtakísérletek vagy trágyázási kísérletsorozatok különböző kutató állomásokon egy régióban vagy országban. A fajtakísérletekben egy növényfaj kevés, ígéretes új fajtáját teszteljük számos teszthelyen és több tenyészidőben, együtt a legnagyobb területen termesztett fajtával egy adott területen. Az ilyen kísérletek eredményeit használjuk fel elsődlegesen a legjobb fajták azonosítására, valamint e fajták alkalmazkodóképességi tartományának meghatározására. A trágyázási kísérletekben másrésről több műtrágyadózis tesztelhető különböző vizsgálati helyeken és több éven keresztül abból a célból, hogy azonosítsuk a helyek csoportjait, amelyek hasonló műtrágya reakcióval rendelkeznek.

Mivel a technológia adaptációs kísérleteket általában több helyen állítjuk be, mind egyik kísérlet rendszerint kevés kezelést tartalmaz és elrendezése egyszerű. Ha faktoriális kísérleteket használnak, a faktorok száma általában nem haladja meg a kettőt. A két leggyakrabban használt elrendezés a randomizált teljes blokk és a split-plot elrendezés.

A különböző helyeken beállított technológia adaptációs kísérletek általában ugyanazokat a kezeléseket tartalmazzák és azonos elrendezéssel, amely nagyban megkönnyíti az analízist. Az ilyen kísérletsorozatokat általában együtt analizáljuk a tenyészidőszak végén, hogy megvizsgáljuk a kezelés \times hely interakció hatást és a

kezelések átlagos hatását a homogén kísérleti helyeken. Ezek a hatások képezik az elsődleges alapját annak, hogy identifikáljuk a legjobb teljesítményűeket, alkalmazkodóképességük tartományát a vizsgált különböző technológiák között.

14.2.1. Kísérlet leírása

A fajta adaptációs kísérlet kombinált adatelemzését m helyen mutatjuk be egy olyan kísérletsorozat adataival, amely hét rizsfajtát foglal magában RCB elrendezésben, kilenc helyen (Gomez és Gomez, 1984).

Minden kísérleti hely varianciaanalízisének eredményét a szemtermés adatok alapján az 14.9. táblázat mutatja. A fajta átlagterméseket mindegyik helyen a 14.10. táblázat mutatja.

14.9. táblázat. A kilenc kísérleti hely (L_1 – L_9) egyenkénti varianciaanalízise (RCB), az adatok hét rizsfajtát magában foglaló fajtakísérletből származnak

Variáció forrás	Szabadságfok	Eltérésnégyzet összegek (SS)								
		L_1	L_2	L_3	L_4	L_5	L_6	L_7	L_8	L_9
Ismétlés	2	0.0158	1.3602	0.2625	0.1399	0.0330	0.0969	4.3398	0.7083	0.0832
Fajta	6	6.7577*	5.9147	7.4418	8.6001**	8.3456**	2.7980	7.7674	3.0798	3.4924**
Hiba	12	3.2507	4.2536	20.582	1.3392	0.5177	4.8435	5.3392	4.1597	0.5272
CV (%)		10.9	16.2	17.0	6.5	14.7	19.4	9.0	14.5	4.4

**szignifikáns 1%-os szinten, *szignifikáns 5%-os szinten

14.2.2. Varianciaanalízis lépései

1. Lépés. Alkalmazzuk a khi-négyzet tesztet az m hiba variancia homogenitásának vizsgálatára az egyedi varianciaanalízisekből. Ha a teszt szignifikáns (amely általában várható), minden hely, amelynek variációs koefficiense rendkívül nagy (pl. CV értéke > 20%) kizárható a kombinált analízisből.

14.10. táblázat. A hét rizsfajta (V_1 – V_7) átlagtermése a kilenc kísérleti helyen (L_1 – L_9)

Fajta	Átlagtermés t ha ⁻¹									
	L_1	L_2	L_3	L_4	L_5	L_6	L_7	L_8	L_9	Átlag
V_1	4.835	4.288	7.882	5.219	2.052	2.968	8.066	4.568	5.122	5.000
V_2	4.412	3.694	8.110	5.036	1.042	3.285	6.584	3.889	4.445	4.500
V_3	4.888	3.963	7.495	5.054	1.342	3.496	8.387	4.136	4.945	4.856
V_4	3.717	2.675	6.568	3.725	2.643	2.704	7.182	4.136	5.200	4.283
V_5	5.635	3.351	7.475	5.858	0.772	3.076	7.679	3.765	4.159	4.641
V_6	4.808	4.288	7.672	5.403	1.042	3.908	6.796	4.506	5.012	4.826
V_7	5.271	3.495	8.652	5.645	1.000	3.431	7.261	3.395	4.224	4.708
Átlag	4.795	3.679	7.693	5.134	1.413	3.267	7.422	4.056	4.730	4.688

Példánkra, a khi-négyzet teszt a kilenc hiba MS-re (14.9. táblázat) $\chi^2 = 59.65$ értéket eredményez, amely 1%-os szinten szignifikáns. Mivel azonban egyik helyen sem volt nagyobb a CV értéke 20%-nál, mind a kilenc helyet bevontuk a kombinált analízisbe.

2. Lépés. Szerkesszük meg a kombinált variancia táblázatot az m hely figyelembevételével és számítsuk ki a különböző SS és MS értékeket.

A számítások az előző példában az összegértékeken alapultak. Ebben a példában bemutatott számítások a 14.10. táblázat átlagos értékein és az 14.9. táblázat SS értékein alapulnak.

Ismétlések helyen belül SS = 0.01581 + 1.36015 + ... + 0.08317 = 7.03960

Összevont hiba SS = 3.25070 + 4.25357 + ... + 0.52724 = 44.81283

C.F. = (3)(7)(9)(4.688) = 4 153.71802

L SS = (3)(7)[(4.795)² + ... + (4.729)²] - 4 153.71802 = 647.87421

V SS = (3)(9)[(5.000)² + ... + (4.708)²] - 4 153.71802 = 8.85162

V × L SS = (3)[(4.835)² + (4.288)² + ... + (4.224)²] - 4 153.71802 - 647.87421 - 8.85162 = 45.63398

Az MS értékeket a szokásos módon számítjuk ki. Az eredményeket a 14.11. táblázat mutatja.

3. Lépés. Azonosítsuk a kezelések alkalmazkodóképességének tartományát a kezelés × hely interakció SS megfelelő felbontásán keresztül. Ez az alábbi megközelítések egyikével vagy kombinációjával végezhető el.

- *Homogén hely megközelítés.* Csoportosítsuk a helyeket homogén csoportokba úgy, hogy az interakció a kezelés és a hely között ne legyen szignifikáns. A fajták elkülönített csoportja azonosítható, mint javasolható adaptálódott fajták a homogén helyek mindegyik sorozatára.
- *Homogén kezelés megközelítés.* Azonosítsunk egy vagy több kezelést, melyeknek nagy a hozzájárulása a kezelés és hely közötti interakcióhoz. Azok a kezelések, amelyeknek teljesítménye jelentősen ingadozik kísérleti helyenként, legjobban hozzájárulnak a kezelés × hely interakcióhoz és várhatóan szűk az alkalmazkodási (adaptációs) tartományuk.

14.11. táblázat. A kísérleti helyek előzetes kombinált varianciaanalízise, a 14.9. és 14.10. táblázat adataiból számolva

Variáció forrása	Szabadságfok	Eltérésnégyzet összeg	Közepes négyzetes eltérés
Hely (L)	8	647.87	80.98
Ismétlés helyen belül	18	7.0396	0.391
Fajta (V)	6	8.85	1.475
V × L	48	45.63	0.951
Összevont hiba	108	44.81	0.415
Összes	188	754.21	

Annak eldöntéséhez, hogy melyik megközelítést válasszuk, vizsgáljuk meg az adatokat: (1) minden fajta átlagos értékét ábrázoljuk az Y-tengelyen és a hely kódokat (pl. L_1, L_2, L_3, \dots) az X-tengelyen; (2) ábrázoljuk mindegyik hely átlagos értékét az Y-tengelyen és a fajta kódokat az X-tengelyen; (3) mindkettő ábrázolása. Ha az (1) ábra feltár egy vagy több fajtát, amelyek egyértelműen különböznek a többitől, a homogén kezelés-megközelítés javasolt. Ha azonban a (2) ábra vizuális megtekintése a helyek egyértelmű csoportosítását jelzi hasonló fajta reakcióval, akkor a homogén hely-megközelítés javasolt.

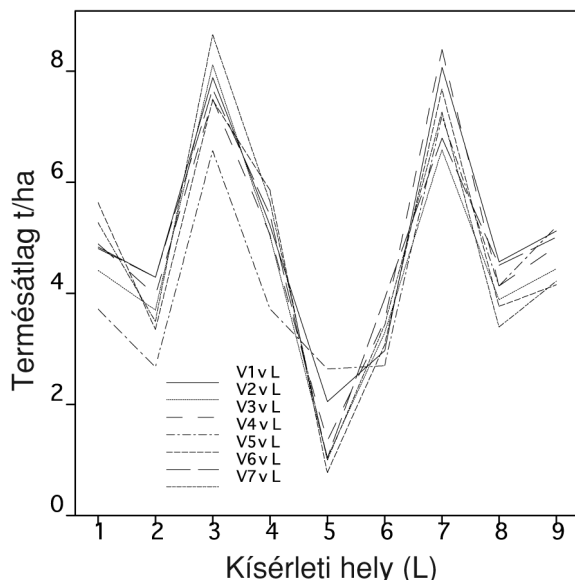
Példánkra, az (1) típusú megközelítést az 14.1. ábra szemlélteti. Jól látható, hogy a V_4 fajta a többi fajtától elkülönülő trendet mutat. Tehát a homogén kezelés-megközelítés javasolt, a V SS és $V \times L$ SS felbontása javasolt a következő komponensekre:

E_1 : V_4 vs. többi fajta

E_2 : többi fajta között

A felbontás eredményét a 14.12. táblázat mutatja.

4. Lépés. A specifikus eljárás az F-érték kiszámítására, V SS, $V \times L$ SS és komponenseik szignifikanciájának tesztelésére függ a kísérleti helyek szerinti hiba variancia homogenitás vizsgálatának szignifikanciájától, melyet az 1. lépésben számítottunk ki. Példánkban, a hiba varianciák homogenitásának vizsgálata az 1. lépésben szignifikáns volt. Ezáltal az összevont hiba SS felbontása megfelel a 3. lépésben bevezetett fajta kontrasztnak. Az eredményt a 14.12. táblázat mutatja.



14.1. ábra. A fajta \times hely interakció vizsgálata, mind a hét fajta (V_1 – V_7) átlagtermését ábrázolva a kísérleti helyek (L) (1–9) szerint

14.12. táblázat. Az SS felbontása az előzetes kombinált varianciaanalízisben (14.11. táblázat), a helyek közötti hiba variancia heterogenitás vizsgálatára

Variancia forrása	Szabadságfok	Eltérés Négyzet-összeg	Közepes négyzetes eltérés	Számított F-érték	Táblázati F-érték	
					5%	1%
Hely (L)	8	647.87	80.98	207.01**	2.51	3.71
Ismétlés helyen belül	18	7.0396	0.391			
Fajta (V)	6	8.852	1.475	3.56**	2.18	2.96
E1: V4 vs. többi	(1)	5.159	5.159	12.43**	3.93	6.88
E2: a többi között	(5)	3.692	0.738	1.78 ^{ns}	2.30	3.19
V × L	48	45.634	0.951	2.29**	1.48	1.76
E ₁ × L	(8)	21.22	2.653	6.39**	2.02	2.68
E ₂ × L	(40)	24.41	0.610	1.47 ^{ns}	1.50	1.78
Összevont hiba	108	44.81	0.415			
Ismétlés helyen belül × E ₁	(18)	5.055	0.281			
Ismétlés helyen belül × E ₂	(90)	39.758	0.442			

Az összevont hiba SS két komponense homogenitásának F-tesztje adja az $F = 1.57$ számított értéket, amely nem szignifikáns. Tehát az összevont hiba MS az F-érték nevezőjeként használható a V hatás, a $V \times L$ interakció hatás, vagy komponenseik szignifikanciájának tesztelésére. Például,

$$F(V) = \frac{V \text{ MS}}{\text{Összevont hiba MS}} = \frac{1.475}{0.415} = 3.56$$

amely szignifikáns 1%-os szinten. Minden teszt eredménye (14.12. táblázat) mutatja, hogy a V_4 fajta kivételével, nem volt szignifikáns különbség a fajták között és ez következetes volt minden helyen. Ez azt jelenti, hogy a V_4 kivételével minden fajta egyformán adaptálódott a kilenc kísérleti helyhez. A V_4 teljesítménye (14.10. táblázat) azonban jelentősen ingadozott a kísérleti helyeken, jelezve alkalmazkodóképességének szűk tartományát. Továbbá, mivel átlagtermése is szignifikánsan alacsonyabb a többinél, ajánlása kétséges még olyan kiválasztott helyekre is, ahol viszonylag nagy termést ért el.

14.3. Technológiai adaptációs kísérletsorozat: Műtrágyázási kísérlet

14.3.1. Kísérlet leírása

A kombinált varianciaanalízis lépésenkénti eljárásait az m helyen beállított osztott parcellás kísérlet adataival mutatjuk be. A faktoriális kísérlet magában foglal két rizsfajtát és hat nitrogén dózist, melyet három helyen teszteltek. Mindegyik helyen három ismétlés volt. Nitrogén dózis képezte a főparcellát és a fajta az alparcella faktort. A szemtermés adatokat a 14.13. táblázat, a kísérleti helyenkénti varianciaanalízist a 14.14. táblázat tartalmazza.

14.13. táblázat. Két rizsfajta (V_1 és V_2) szemtermése, hat nitrogén dózissal (N_1 -től N_6 -ig) és három helyen (L_1 -től L_3 -ig) vizsgálva, osztott parcellás elrendezésben és három ismétlésben

Hely	Nitrogén	Szemtermés kg ha ⁻¹							
		I. ismétlés		II. ismétlés		III. ismétlés		Összes	
		V_1	V_2	V_1	V_2	V_1	V_2	V_1	V_2
L_1	N_1	1979	5301	1511	1883	3664	3571	7154	10755
	N_2	4572	5655	4340	5100	4132	5385	13044	16140
	N_3	5630	6339	6780	6622	4933	6332	17343	19293
	N_4	7153	8108	6504	8583	6326	7637	19983	24328
	N_5	7223	7530	7107	7097	6051	6667	20381	21294
	N_6	7239	7853	6829	7105	5874	7443	19942	22401
	Összes							97847	114211
L_2	N_1	3617	3447	3580	3560	3939	3516	11136	10523
	N_2	6065	5905	5463	5969	5435	6026	16963	17900
	N_3	6092	5322	6571	5883	6084	6489	18747	17694
	N_4	5916	6513	6982	6556	7145	7853	20043	20922
	N_5	7191	8153	6109	7208	7967	6685	21267	22046
	N_6	5805	7290	6890	6564	7113	7401	19808	21255
	Összes							107964	110340
L_3	N_1	4320	4891	4068	2577	3856	4541	12244	12009
	N_2	5862	6009	4626	6625	4913	5672	15401	18306
	N_3	5136	6712	5836	6693	4898	6799	15870	20204
	N_4	6336	6458	5456	6675	5663	6636	17455	19769
	N_5	5571	5683	5854	6868	5533	5692	16958	18243
	N_6	6765	6335	5263	6064	3910	5949	15938	18348
	Összes							93866	106879

14.14. táblázat. Termőhelyenkénti varianciaanalízis az 1. táblázat adatai alapján

Variáció forrása	d.f.	L ₁			L ₂			L ₃		
		SS	MS	F-érték	SS	MS	F-érték	SS	MS	F-érték
Ismétlés	2	1984474	992237		1040418	520209		1520023	760012	
Nitrogén(N)	5	85511456	17102291	25.21**	51671993	10334399	37.35**	18257312	3651462	8.50**
Hiba (a)	10	6782715	678271		2766814	276681		4294863	429486	
Fajta (V)	1	7438347	7438347	17.93**	156816	156816	< 1	4703838	4703838	11.79**
N × V	5	1247352	249470		815824	163165	< 1	1978116	395623	< 1
Hiba (b)	12	4979194	414933	< 1	3667249	305604		4788886	399074	
CV (a) %			14.0			8.7			11.8	
CV (b) %			10.9			9.1			11.3	

14.3.2. Varianciaanalízis lépései

1. lépés. Vizsgáljuk meg a hiba (a) MS és hiba (b) MS homogenitását. Példánkban a három hiba (a) MS homogenitásának tesztelésére kiszámított χ^2 értéke 1.90; a hiba (b) MS χ^2 értéke 0.31; egyik sem szignifikáns.

2. lépés. Állítsuk össze a kombinált varianciaanalízist m hely figyelembevételével, az osztott parcellás elrendezés alapján az a főparcella kezelésekkel, b alparcella kezelésekkel és r ismétléssel, mint a 14.15. táblázat mutatja.

14.15. táblázat. A kombinált varianciaanalízis táblázata helyek figyelembevételével, osztott parcellás elrendezés alapján

Variáció forrása	Szabadságfok ^a	Közepes négyzetes eltérés	Számított F-érték
Hely (L)	$m-1$	L MS	L MS/R MS
Ismétlés helyen belül	$m(r-1)$	R MS	
Főparcella faktor (A)	$a-1$	A MS	A MS/ E _a MS
L × A	$(m-1)(a-1)$	L × A MS	L × A MS/ E _a MS
Összevont hiba (a)	$m(r-1)(a-1)$	E _a MS	
Alparcella faktor (B)	$b-1$	B MS	B MS/ E _b MS
L × B	$(m-1)(b-1)$	L × B MS	L × B MS/ E _b MS
A × B	$(a-1)(b-1)$	A × B MS	A × B MS/ E _b MS
L × A × B	$(m-1)(a-1)(b-1)$	L × A × B MS	L × A × B MS/ E _b MS
Összevont hiba (b)	$ma(r-1)(b-1)$	E _b MS	
Összes	$mra-1$		

^aa = a főparcella kezelések száma, b = alparcella kezelések száma, r = ismétlések száma és m = kísérleti helyek száma

3. lépés. Számítsuk ki az ismétlések kísérleti helyen belül SS-t, az összevont hiba (a) SS-t és az összevont hiba (b) SS-t a következőképpen:

$$\text{Ismétlések kísérleti helyen belül SS} = \sum_{i=1}^m (\text{Ismétlés SS})_i$$

$$\text{Összevont hiba (a) SS} = \sum_{i=1}^m [\text{Hiba (a) SS}]_i$$

$$\text{Összevont hiba (b) SS} = \sum_{i=1}^m [\text{Hiba (b) SS}]_i$$

ahol i alsó index az i -edik helyre vonatkozik.

Példánkra, kiszámítjuk a három SS értéket, felhasználva a 14.14. táblázat adatait:

Ismétlések kísérleti helyen belül SS = 4544916

Összevont hiba (a) SS = 13844392

Összevont hiba (b) SS = 13435330

4. lépés. Számítsuk ki a fennmaradó SS értékeket, amelyek a 14.15. táblázatba szükségesek az összeg értékek alapján. Az átlagértékek szerinti számítás formuláit Gomez és Gomez (1984) ismerteti. A különböző SS értékek összes (total) adatok szerinti kiszámítása a standard eljárást követi:

$$C.F. = \frac{\left(\sum_{i=1}^m G_i \right)^2}{mabr}$$

$$L \text{ SS} = \frac{\sum_{i=1}^m G_i^2}{abr} - C.F.$$

$$A \text{ SS} = \frac{\sum_{j=1}^a A_j^2}{mbr} - C.F.$$

$$L \times A \text{ SS} = \frac{\sum_{i=1}^m \sum_{j=1}^a (LA)_{ij}^2}{br} - C.F. - L \text{ SS} - A \text{ SS}$$

$$B \text{ SS} = \frac{\sum_{k=1}^b B_k^2}{mar} - C.F.$$

$$L \times B \text{ SS} = \frac{\sum_{i=1}^m \sum_{k=1}^b (LB)_{ik}^2}{ar} - C.F. - L \text{ SS} - B \text{ SS}$$

$$A \times B \text{ SS} = \frac{\sum_{j=1}^a \sum_{k=1}^b (AB)_{jk}^2}{mr} - \text{C.F.} - A \text{ SS} - B \text{ SS}$$

$$L \times A \times B \text{ SS} = \frac{\sum_{i=1}^m \sum_{j=1}^a \sum_{k=1}^b (LAB)_{ijk}^2}{r} - \text{C.F.} - L \text{ SS} - A \text{ SS} - B \text{ SS} - L \times A \text{ SS} - L \times B \text{ SS} - A \times B \text{ SS}$$

ahol G_i az i -edik hely főösszege, A_j az A faktor j -edik szintjének összege, $(LA)_{ij}$ az A faktor j -edik szintjének összege az i -edik helyen, B_k a B faktor k -edik szintjének összege, $(LB)_{ik}$ a B faktor k -edik szintjének az összege az i -edik helyen, $(AB)_{jk}$ kezelés-összeg, amely magában foglalja az A faktor j -edik szintjét és a B faktor k -edik szintjét, és $(LAB)_{ijk}$ kezelés-összeg, amely magában foglalja az A faktor j -edik szintjét és a B faktor k -edik szintjét az i -edik helyen.

14.16. táblázat. A hely \times nitrogén ($L \times N$) összegek, az 14.13. táblázat adataiból számítva

Nitrogén	Termés-összeg			Nitrogén-összeg
	L_1	L_2	L_3	
N_1	17909	21659	24253	63821
N_2	29184	34863	33707	97754
N_3	36636	36441	36074	109151
N_4	44311	40965	37224	122500
N_5	41675	43313	35201	120189
N_6	42343	41063	34286	117692
Összesen	212058	218304	200745	631107

14.17. táblázat. A nitrogén \times fajta ($N \times V$) összegek, az 14.13. táblázat adataiból számolva

Nitrogén	Termés-összeg	
	V_1	V_2
N_1	30534	32287
N_2	45408	52346
N_3	51960	57191
N_4	57481	65019
N_5	58606	61583
N_6	55688	62004
Fajta-összeg	299677	331430

Példánkban a számítás a 14.13. táblázat összeg-értékei alapján történik. Először az $L \times N$ és $N \times V$ táblázat-összegeket számítjuk ki, amelyeket a 14.16. és 14.17. táblázat mutat. Ezután a különböző SS-ek kiszámítása történik standard formában:

$$C.F. = \frac{(212058 + 218304 + 200745)^2}{(3)(6)(2)(3)} = 3678926346$$

$$L\ SS = \frac{(212058)^2 + (218304)^2 + (200745)^2}{(6)(2)(3)} - 3678926346 = 4401065$$

$$N\ SS = \frac{(63821)^2 + (97754)^2 + \dots + (117692)^2}{(3)(2)(3)} - 3678926346 = 136849095$$

$$L \times N\ SS = \frac{(17909)^2 + (21659)^2 + \dots + (117692)^2}{(2)(3)} - 3678926346 - 4401065 - 136849095 = 18591664$$

$$V\ SS = \frac{(299677)^2 + (331430)^2}{(3)(6)(3)} - 3678926346 = 9335676$$

$$L \times V\ SS = \frac{(97847)^2 + (114211)^2 + \dots + (106879)^2}{(6)(3)} - 3678926346 - 4401065 - 9335676 = 2963325$$

$$N \times V\ SS = \frac{(30534)^2 + (33287)^2 + \dots + (62004)^2}{(3)(3)} - 3678926346 - 136849095 - 9335676 = 1145104$$

$$L \times N \times V\ SS = \frac{(7154)^2 + (10755)^2 + \dots + (18348)^2}{3} - 3678926346 - 4401065 - 136849095 - 9335676 - 18591664 - 2963325 - 1145104 = 2896188$$

5. lépés. A variáció minden forrására számítsuk ki a közepes négyzetes eltérés értékeit, az SS adatokat elosztva a megfelelő szabadságfokokkal. Példánkban, az eredményeket a 14.18. táblázat mutatja.

6. lépés. Számítsuk ki az F-értéket, hogy teszteljük a hely hatás szignifikanciáját:

$$F(L) = \frac{L\ MS}{\text{Ismétlések helyen belül MS}} = \frac{2200532}{757486} = 2.91$$

Mivel a számított F-érték kisebb, mint a megfelelő táblázati F-érték (5.14) $f_1=2$ és $f_2=6$ szabadságfoknál 5%-os szignifikancia szinten, a helyhatás nem szignifikáns.

7. lépés. Ha megállapítottuk a hiba (a) MS homogenitását a különböző helyeken (1. lépés), akkor számítsuk ki az F-értéket, hogy teszteljük az A hatás és az $L \times A$ hatás szignifikanciáját:

$$F(A) = \frac{A\ MS}{\text{Összevont hiba (a) MS}}$$

$$F(L \times A) = \frac{L \times A\ MS}{\text{Összevont hiba (a) MS}}$$

Példánkban, mivel megállapítottuk a három hiba (a) MS homogenitását, kiszámítjuk a két F-értéket a formulák alapján:

$$F(N) = \frac{27369819}{461480} = 59.31$$

$$F(L \times N) = \frac{1859166}{461480} = 4.03$$

A megfelelő táblázati F-értékek az F(N)-re 2.53 az 5%-os és 3.70 az 1%-os szignifikancia szinten; az F(L × N)-re 2.16 és 2.98. Ezáltal mind a nitrogén hatás és mind interakciója a hellyel erősen szignifikáns. A kísérleti bizonyíték világosan jelzi, hogy jöllehet az átlagos nitrogén reakció szignifikáns, a reakció változott a kísérleti helyek között.

14.18. táblázat. Kombinált varianciaanalízis kísérleti helyek figyelembe vételével (osztott parcellás elrendezés) az 14.13. és 14.14. táblázatok adataira

Variáció forrása	d.f.	SS	MS	Számított F-érték	Táblázati F-érték	
					5%	1%
Hely (L)	2	4401065	2200532	2.91 ^{ns}	5.14	10.92
Ismétlés helyen belül	6	4544916	757486			
Nitrogén dózis (N)	5	136849095	27369819	59.31**	2.53	3.70
L × N	10	18591664	1859166	4.03**	2.16	2.98
Összevont hiba (a)	30	13844392	461480			
Fajta (V)	1	9335676	9335676	25.01**	4.11	7.39
L × V	2	2963325	1481662	3.97*		
N × V	5	1145104	229021	< 1		
L × N × V	10	2896188	289619	< 1		
Összevont hiba (b)	36	13435330	373204			
Összes	107					

**szignifikáns 1%-os szinten, *szignifikáns 5%-os szinten, ^{ns} nem szignifikáns.

8. lépés. Ha megállapítottuk a hiba (b) MS homogenitását a kísérleti helyek szerint (1. lépés), számítsuk ki az F-értékeket minden hatásra, amely magában foglalja a B alparcella faktort:

$$F(B) = \frac{B \text{ MS}}{\text{Összevont hiba (b) MS}}$$

$$F(L \times B) = \frac{L \times B \text{ MS}}{\text{Összevont hiba (b) MS}}$$

$$F(A \times B) = \frac{A \times B \text{ MS}}{\text{Összevont hiba (b) MS}}$$

$$F(L \times A \times B) = \frac{L \times A \times B \text{ MS}}{\text{Összevont hiba (b) MS}}$$

Példánkban, mivel megállapítottuk a hiba (b) homogenitását a három helyre, az F-értékeket kiszámítjuk a formulák alapján:

$$F(V) = \frac{9335676}{373204} = 25.01$$

$$F(L \times V) = \frac{1481662}{373204} = 3.97$$

$$F(N \times V) = \frac{229021}{373204} \langle 1$$

$$F(L \times N \times V) = \frac{289619}{373204} \langle 1$$

9. lépés. Írjuk be minden számítás eredményét a 3.-tól a 8.-ik lépésig a kombinált varianciaanalízis táblázatba. Példánkban az eredményeket a 14.18. táblázat tartalmazza. Az eredmények a következőképpen összegezhetők:

- Nem volt szignifikáns interakció a nitrogén és a fajta között, és ez az eredmény következetes volt minden vizsgált helyen (azaz, $L \times N \times V$ nem szignifikáns).
- Mind a nitrogén és mind a fajta fő hatása erősen szignifikáns, azonban a nitrogén reakció és a fajta különbség egyaránt változott a vizsgált helyek között (azaz, mind az $L \times N$ és mind az $L \times V$ szignifikáns).
- Nem volt szignifikáns különbség a hely átlagok között.

10. lépés. Vizsgáljunk meg minden olyan interakció hatást, amely magában foglalja a helyet és szignifikáns. Példánkban két ilyen interakció van, nevezetesen az $L \times N$ és az $L \times V$ interakció.

Az $L \times N$ interakció esetében bontsuk fel az $L \times N$ SS-t és illesszünk regressziós egyenletet, amely leírja a termésreakciót a nitrogénre mindegyik helyen. Az $L \times N$ SS felbontásának eredményét a 14.19. táblázat mutatja, és jelzi, hogy a nitrogén reakció görbének csupán a lineáris szakasza változott a helyek között. A becült másodfokú regressziókat, egyet mindegyik helyre, grafikusan a 14.2. ábra mutatja.

Az $L \times V$ interakcióra, mivel a fajta egy diszkrét faktor, és mivel csak két fajta van, a helyek szerinti fajtaátlag különbségek táblázatát a 14.19. táblázat mutatja. Az eredmények jelzik, hogy V_2 nagyobb termést adott, mint V_1 minden helyen, azonban csak az L_1 és L_2 helyen volt szignifikáns.

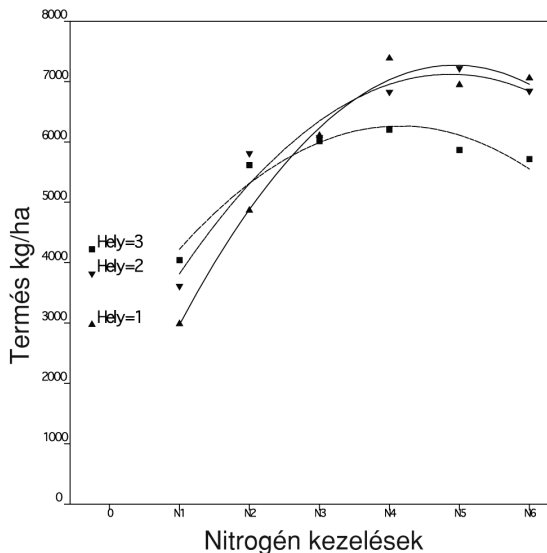
14. 19. táblázat. Fajta különbség mindegyik kísérleti helyen, az 1. táblázat adataiból számolva

Kísérleti hely	Átlagtermés kg ha ⁻¹		
	V ₁	V ₂	Különbség
L ₁	5436	6345	909**
L ₂	5998	6130	132 ^{ns}
L ₃	5215	5938	723**
Átlag	5550	6138	

**szignifikáns 1%-os szinten, *szignifikáns 5%-os szinten, ^{ns} nem szignifikáns.

11. lépés. Értékeljük az eddigi eredményeket, hogy összegezhessük a megállapításokat és levonhassunk következtetéseket, továbbá javaslatokat tevhessünk, ha lehetséges. Példáinkra az eredmények a következőképpen összegezhetők:

- A fajta vonatkozásában, mivel V₂ jobb vagy egyenlő volt V₁ fajtával mindegyik vizsgált helyen, V₂ a potenciális fajta a szaktanácsadásban olyan területeken, mint amelyeket a három vizsgálati hely reprezentál.
- Nitrogén reakciót tekintve, a reakciógörbe lineáris szakasza különbözött a kísérleti helyek között. A becsült másodfokú reakció egyenleteket a három helyre a 14.2. ábra mutatja. A nitrogén dózisok, amelyek maximalizálják a termést, 118 kg N ha⁻¹ az L₁ helyen, 116 kg N ha⁻¹ az L₂ helyen és 96 kg N ha⁻¹ az L₃ helyen.



14.2. ábra. Termésreakció a nitrogén műtrágyázásra három helyen

14.3.3. Kísérletsorozat számítógépes értékelése GenStat programmal

1. Termőhely:**Analysis of variance**

Variate: Yield

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Replication stratum	2	1984474.	992237.	1.46	
Replication.Nitrogen stratum					
Nitrogen	5	85511456.	17102291.	25.21	<.001
Residual	10	6782715.	678271.	1.63	
Replication.Nitrogen.Variety stratum					
Variety	1	7438347.	7438347.	17.93	0.001
Nitrogen.Variety	5	1247352.	249470.	0.60	0.700
Residual	12	4979194.	414933.		
Total	35	107943537			

Tables of means

Variate: Yield

Grand mean 5890.

	1	2	3	4	5	6
Nitrogen	2985.	4864.	6106.	7385.	6946.	7057.
Variety	5436.	6345.				
Nitrogen Variety		1	2			
1		2385.	3585.			
2		4348.	5380.			
3		5781.	6431.			
4		6661.	8109.			
5		6794.	7098.			
6		6647.	7467.			

Standard errors of differences of means

Table	Nitroge	Variety	Nitrogen Variety
rep.	6	18	3
s.e.d.	475.5	214.7	603.7
d.f.	10	12	19.80
Except when comparing means with the same level(s) of			
Nitrogen			525.9
d.f.			12

Least significant differences of means (5% level)

Table	Nitrogen	Variety	Nitrogen Variety
rep.	6	18	3
l.s.d.	1059.5	467.8	1260.0
d.f.	10	12	19.80
Except when comparing means with the same level(s) of			
Nitrogen			1145.9
d.f.			12

**2. termőhely:
Analysis of variance**

Variate: Yield

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Replication stratum	2	1040418.	520209.	1.88	
Replication.Nitrogen stratum					
Nitrogen	5	51671993.	10334399.	37.35	<.001
Residual	10	2766814.	276681.	0.91	
Replication.Nitrogen.Variety stratum					
Variety	1	156816.	156816.	0.51	0.487
Nitrogen.Variety	5	815824.	163165.	0.53	0.747
Residual	12	3667249.	305604.		
Total	35	60119114.			

Tables of means

Variate: Yield

Grand mean 6064.

Nitrogen	1	2	3	4	5	6
	3610.	5810.	6074.	6828.	7219.	6844.
Variety	1	2				
	5998.	6130.				
Nitrogen	Variety	1	2			
1		3712.	3508.			
2		5654.	5967.			
3		6249.	5898.			
4		6681.	6974.			
5		7089.	7349.			
6		6603.	7085.			

Standard errors of differences of means

Table	Nitrogen	Variety	Nitrogen
		Variety	
rep.	6	18	3
s.e.d.	303.7	184.3	440.6
d.f.	10	12	21.96
Except when comparing means with the same level(s) of			
Nitrogen			451.4
d.f.			12

Least significant differences of means (5% level)

Table	Nitrogen	Variety	Nitrogen
		Variety	
rep.	6	18	3
l.s.d.	676.7	401.5	913.8
d.f.	10	12	21.96
Except when comparing means with the same level(s) of			
Nitrogen			983.5
d.f.			12

3. termőhely:**Analysis of variance**

Variate: Yield

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Replication stratum	2	1520023.	760012.	1.77	
Replication.Nitrogen stratum					
Nitrogen	5	18257312.	3651462.	8.50	0.002
Residual	10	4294863.	429486.	1.08	
Replication.Nitrogen.Variety stratum					
Variety	1	4703838.	4703838.	11.79	0.005
Nitrogen.Variety	5	1978116.	395623.	0.99	0.463
Residual	12	4788886.	399074.		
Total	35	35543039.			

Tables of means

Variate: Yield

Grand mean 5576.

Nitrogen	1	2	3	4	5	6
	4042.	5618.	6012.	6204.	5867.	5714.
Variety	1	2				

	5215.	5938.
Nitrogen		
Variety	1	2
1	4081.	4003.
2	5134.	6102.
3	5290.	6735.
4	5818.	6590.
5	5653.	6081.
6	5313.	6116.

Standard errors of differences of means

Table	Nitrogen	Variety	Nitrogen Variety
rep.	6	18	3
s.e.d.	378.4	210.6	525.5
d.f.	10	12	21.64
Except when comparing means with the same level(s) of			
Nitrogen			515.8
d.f.			12

Least significant differences of means (5% level)

Table	Nitrogen	Variety	Nitrogen Variety
rep.	6	18	3
l.s.d.	843.1	458.8	1090.9
d.f.	10	12	21.64
Except when comparing means with the same level(s) of			
Nitrogen			1123.8
d.f.			12

1–3 termőhely kombinált varianciaanalízise: Analysis of variance

Variate: Yield

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Location.Replication stratum					
Location	2	4401064.	2200532.	2.91	0.131
Residual	6	4544915.	757486.	1.64	
Location.Replication.Nitrogen stratum					
Nitrogen	5	136849095.	27369819.	59.31	<.001
Location.Nitrogen	10	18591666.	1859167.	4.03	0.001
Residual	30	13844392.	461480.	1.24	

Location. Replication. Nitrogen. Variety stratum

Variety	1	9335676.	9335676.	25.01	<.001
Location.Variety	2	2963325.	1481663.	3.97	0.028
Nitrogen.Variety	5	1145104.	229021.	0.61	0.690
Location.Nitrogen.Variety	10	2896188.	289619.	0.78	0.651
Residual	36	13435330.	373204.		

Total 107 208006754.

Tables of means

Variate: Yield

Grand mean 5844.

Location	1	2	3
	5890.	6064.	5576.

Nitrogen	1	2	3	4	5	6
	3546.	5431.	6064.	6806.	6677.	6538.

Variety	1	2
	5550.	6138.

Location Nitrogen	1	2	3	4	5	6
1	2985.	4864.	6106.	7385.	6946.	7057.
2	3610.	5810.	6074.	6828.	7219.	6844.
3	4042.	5618.	6012.	6204.	5867.	5714.

Location Variety	1	2
1	5436.	6345.
2	5998.	6130.
3	5215.	5938.

Nitrogen Variety	1	2
1	3393.	3699.
2	5045.	5816.
3	5773.	6355.
4	6387.	7224.
5	6512.	6843.
6	6188.	6889.

Location Nitrogen Variety			1	2
	1	1	2385.	3585.
		2	4348.	5380.
		3	5781.	6431.
		4	6661.	8109.
		5	6794.	7098.

		6	6647.	7467.
2		1	3712.	3508.
		2	5654.	5967.
		3	6249.	5898.
		4	6681.	6974.
		5	7089.	7349.
		6	6603.	7085.
3		1	4081.	4003.
		2	5134.	6102.
		3	5290.	6735.
		4	5818.	6590.
		5	5653.	6081.
		6	5313.	6116.

Standard errors of differences of means

Table	Location	Nitrogen	Variety	Location Nitrogen
rep.	36	18	54	6
s.e.d.	205.1	226.4	117.6	412.6
d.f.	6	30	36	34.40
Except when comparing means with the same level(s) of				
Location				392.2
d.f.				30

Table	Location Variety	Nitrogen Variety	Location Nitrogen Variety
rep.	18	9	3
s.e.d.	250.6	304.5	542.8
d.f.	12.85	63.52	68.22
Except when comparing means with the same level(s) of			
Location	203.6		527.5
d.f.	36		63.52
Nitrogen		288.0	
d.f.		36	
Location.Nitrogen			498.8
d.f.			36
Location.Variety			527.5
d.f.			63.52

Least significant differences of means (5% level)

Table	Location	Nitrogen	Variety	Location Nitrogen
rep.	36	18	54	6
l.s.d.	502.0	462.5	238.4	838.2
d.f.	6	30	36	34.40
Except when comparing means with the same level(s) of				
Location				801.0
d.f.				30
Table	Location	Nitrogen	Location	

	Variety	Variety	Nitrogen Variety
rep.	18	9	3
l.s.d.	542.1	608.5	1083.2
d.f.	12.85	63.52	68.22
Except when comparing means with the same level(s) of			
Location	413.0		1053.9
d.f.	36		63.52
Nitrogen		584.1	
d.f.		36	
Location.Nitrogen			1011.6
d.f.			36
Location.Variety			1053.9
d.f.			63.52

Kombinált varianciaanalízis

Termőhelyek közötti kontrasztok:

(1): -1 +1 0

(2): -1 -1 +2

Analysis of variance

Variate: Yield

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Location.Replication stratum					
Location	2	4401064.	2200532.	2.91	0.131
Contrast 1	1	541841.	541841.	0.72	0.430
Contrast 2	1	3859224.	3859224.	5.09	0.065
Residual	6	4544915.	757486.	1.64	
Location.Replication.Nitrogen stratum					
Nitrogen	5	136849095.	27369819.	59.31	<.001
Location.Nitrogen	10	18591666.	1859167.	4.03	0.001
Contrast 1.Nitrogen	5	4613887.	922777.	2.00	0.107
Contrast 2.Nitrogen	5	13977779.	2795556.	6.06	<.001
Residual	30	13844392.	461480.	1.24	
Location.Replication.Nitrogen.Variety stratum					
Variety	1	9335676.	9335676.	25.01	<.001
Location.Variety	2	2963325.	1481663.	3.97	0.028
Contrast 1.Variety	1	2717558.	2717558.	7.28	0.011
Contrast 2.Variety	1	245768.	245768.	0.66	0.422
Nitrogen.Variety	5	1145104.	229021.	0.61	0.690
Location.Nitrogen.Variety	10	2896188.	289619.	0.78	0.651

Contrast 1.Nitrogen.Variety	5	990138.	198028.	0.53	0.752
Contrast 2.Nitrogen.Variety	5	1906050.	381210.	1.02	0.420
Residual	36	13435330.	373204.		
Total	107	208006754.			

Regression analysis

Response variate: L1
 Fitted terms: Constant + Nitrog_n
 Submodels: POL(Nitrog_n; 2)

Summary of analysis

Source	d.f.	s.s.	m.s.	v.r.	F pr.
Regression	2	13991976.	6995988.	81.21	0.002
Residual	3	258430.	86143.		
Total	5	14250406.	2850081.		

Percentage variance accounted for 97.0

Standard error of observations is estimated to be 294.

Estimates of parameters

Parameter	estimate	s.e.	t(3)	t pr.
Constant	2972.	266.	11.17	0.002
Nitrog_n Lin	72.88	8.34	8.74	0.003
Nitrog_n Quad	-0.3088	0.0534	-5.79	0.010

Fitted values and residuals

Unit	Response	Standardized		Leverage
		Fitted value	residual	
1	2985.	2972.	0.10	0.82
2	4864.	4881.	-0.07	0.31
3	6106.	6233.	-0.55	0.37
4	7385.	7030.	1.53	0.37
5	6946.	7271.	-1.33	0.31
6	7057.	6956.	0.82	0.82
Mean	5890.	5891.	0.08	0.50

Regression analysis

Response variate: L2
 Fitted terms: Constant + Nitrog_n
 Submodels: POL(Nitrog_n; 2)

Summary of analysis

Source	d.f.	s.s.	m.s.	v.r.	F pr.
Regression	2	8210154.	4105077.	30.62	0.010
Residual	3	402192.	134064.		
Total	5	8612345.	1722469.		

Percentage variance accounted for 92.2
 Standard error of observations is estimated to be 366.

Estimates of parameters

Parameter	estimate	s.e.	t(3)	t pr.
Constant	3818.	332.	11.50	0.001
Nitrog_n Lin	57.0	10.4	5.47	0.012
Nitrog_n Quad	-0.2454	0.0666	-3.69	0.035

Fitted values and residuals

Unit	Response	Standardized		Leverage
		Fitted value	residual	
1	3610.	3818.	-1.34	0.82
2	5811.	5305.	1.66	0.31
3	6074.	6351.	-0.95	0.37
4	6828.	6955.	-0.44	0.37
5	7219.	7118.	0.33	0.31
6	6844.	6839.	0.03	0.82
Mean	6064.	6064.	-0.12	0.50

Regression analysis

Response variate: L3
 Fitted terms: Constant + Nitrog_n
 Submodels: POL(Nitrog_n; 2)

Summary of analysis

Source	d.f.	s.s.	m.s.	v.r.	F pr.
Regression	2	2828660.	1414330.	19.78	0.019
Residual	3	214465.	71488.		
Total	5	3043125.	608625.		

Percentage variance accounted for 88.3

Standard error of observations is estimated to be 267.

Estimates of parameters

Parameter	estimate	s.e.	t(3)	t pr.
Constant	4223.	242.	17.43	<.001
Nitrog_n Lin	43.29	7.60	5.70	0.011
Nitrog_n Quad	-0.2295	0.0486	-4.72	0.018

Fitted values and residuals

Unit	Response	Fitted value	Standardized residual	Leverage
1	4042.	4223.	-1.60	0.82
2	5618.	5315.	1.36	0.31
3	6012.	5994.	0.08	0.37
4	6204.	6260.	-0.26	0.37
5	5867.	6112.	-1.10	0.31
6	5714.	5552.	1.44	0.82
Mean	5576.	5576.	-0.02	0.50

14.4. 2ⁿ Faktoriális kísérletsorozat kombinált varianciaanalízise

A kombinált analízis megkezdése előtt szükséges néhány előzetes analízis.

(1) A varianciaanalízist el kell végezni minden kísérletre.

(2) Számos kísérletnek a kísérleti hibáját meg kell vizsgálni heterogenitásra.

(3) Egy előzetes becslés szükséges a kezelés interakciójáról a környezettel vagy az évvel vagy mindkettővel.

Ezeknek az előzetes lépéseknek az eredményétől és a szóban forgó kísérlet típusától függően, a kombinált analízis elvégezhető. Az előzetes analízis célja annak meghatározása, hogy az összes kísérlet bevonható-e ugyanabba a kombinált analízisbe vagy a kísérleteket először szétválasszuk-e többé-kevésbé homogén csoportokba. Továbbá, az előzetes analízis bizonyos jelzést adhat arról, hogy vajon általánosan alkalmazható javaslatok lehetségesek-e vagy különböző javaslatok szükségesek, hogy megfeleljenek az eltérő körülményeknek.

Amikor számba vesszük az adatok kombinált analízisét számos termőhelyről, az első követelmény az, hogy becsüljük a hiba variancia homogenitását a különböző termőhelyeken. Ha a hibák homogének, akkor az analízis folytatható. Ha nem homogének a hibák, akkor az adatokat transzformálni kell, hogy homogén varianciát állítsunk elő vagy a termőhelyeket csoportokba kell elkülöníteni, amelyeken belül a variancia homogén. A kombinált varianciaanalízis módszerét *Petersen* (1994) alapján ismertetjük.

14.4.1. Kísérlet leírása

Több helyen beállított kísérletben vizsgálták a lencse termesztésére ható faktorokat 1985–1986-ban Szíriában (14.20. táblázat) (*Petersen*, 1994). A kísérleteket az ICARDA (Nemzetközi Mezőgazdasági Kutatási Központ Száraz Területekre) koordinálta. A kísérlet 2³ faktoriális kísérlet nyolc termőhelyen, randomizált blokk elrendezésben, két blokkban minden termőhelyen. A kezelés faktorok a következők:

1. Sitona ormányosbogár elleni védekezés (S)

S_0 = nincs védekezés S_1 = 1 kg/ha carbofuran

2. Műtrágya (F)

F_0 = 0 műtrágya F_1 = 50 kg/ha P_2O_5

3. Gyomirtás (W)

W_0 = nincs gyomirtás W_1 = herbicid kijuttatása

A szemtermés varianciaanalízisét mindegyik termőhelyen a 14.21. táblázat tartalmazza. A hiba MS (közepes négyzetes eltérés) értékek a 2. táblázatban e kísérletekre vonatkozó hiba varianciák (s^2). Ezek a becslések nyújtanak adatokat a variancia homogenitásának vizsgálatához.

14.20. táblázat. A korai vetésű lencse átlagos szemtermése (kg ha⁻¹) nyolc termőhelyen Szíriában, 1985–1986. években

Keze- lés*	Termőhely							
	1	2	3	4	5	6	7	8
(1)	976.5	1193.5	549.0	513.5	148.0	770.5	562.5	843.0
S	1166.5	1111.0	751.0	751.0	233.0	809.5	710.0	1133.5
F	1074.5	1021.5	666.0	667.0	305.0	920.0	579.5	1074.0
SF	1260.0	1058.0	843.5	1239.5	656.5	878.0	649.0	1153.5
W	1091.0	1138.5	626.0	860.0	410.5	1005.5	744.5	963.0
SW	1193.0	1176.5	1156.5	1445.0	888.0	1120.0	814.5	1323.5
FW	1107.0	1132.5	593.5	841.0	791.5	880.0	929.0	935.5
SFW	1310.0	1214.5	1077.0	1334.0	1024.5	1093.5	919.5	1069.5

*(1): Kontroll; S: 1 kg ha⁻¹ karbofuran; F: 50 kg ha⁻¹ P₂O₅; SF: karbofuran + P₂O₅; W: herbicid; SW: karbofuran + herbicid; FW: P₂O₅ + herbicid; SFW: karbofuran + P₂O₅ + herbicid.

14.21. táblázat. A korai vetésű lencse szemtermésének varianciaanalízise nyolc termőhelyen Szíriában, 1985–1986. években

Forrás	d.f.	MS termőhelyenként			
		1	2	3	4
Blokk	1	105	4970	2522	62750
Kezelés	7	23113*	8779	102750	221985**
Hiba	7	5260	3494	34406	31031
Forrás	d.f.	MS termőhelyenként			
		5	6	7	8
Blokk	1	100172	176400	3220	155433
Kezelés	7	214352**	32464	39984	44523
Hiba	7	27704	20190	21276	32759

*szignifikáns 5%-os; **szignifikáns 1%-os szinten.

14.4.2. A variancia homogenitásának vizsgálata

A variancia homogenitásának gyors tesztjét adja a kísérletsorozatban a legnagyobb és a legkisebb s^2 aránya. A kritikus F-értékek táblázatával (lásd melléklet) összehasonlítva gyakran lehetséges levonni egy következtetést a variancia homogenitásáról, további vizsgálatok nélkül. A korai vetésű lencse adatokra

$$\frac{s^2(\max)}{s^2(\min)} = \frac{34,406}{3494} = 9.85$$

A táblázati érték 5%-ra, $a = k = 8$, $\gamma = 7$ értékekre 12.7. Mivel $9.85 < 12.7$, levonjuk azt a következtetést, hogy a varianciák nem heterogének.

Egy alternatív eljárás, amely érzékenyebb mint az arány teszt, a **Bartlett próba** a variancia homogenitására (Bartlett, 1937). Ezt a tesztet, amely a minta varianciák természetes logaritmusán alapul, *Snedecor* és *Cochran* (1989) írta le. Ennek a tesztnek a végzéséhez, legyen $\gamma =$ hiba szabadságfok az egyes kísérletekre, = hiba MS az i termőhelyen, a = termőhelyek száma, akkor

$$M = \gamma \left[a \left(\ln \bar{s}^2 \right) - \sum_i \ln s_i^2 \right] \quad \bar{s}^2 = \frac{\sum_i s_i^2}{a} \quad \text{és}$$

$$C = 1 + \frac{a+1}{3a\gamma}$$

Az M/C arány egy teszt mutató a null hipotézisre, hogy minden s^2 egy becslése a σ^2 -nek. A M/C arány χ^2 eloszlást mutat $(a-1)$ szabadságfokkal.

Számításokat a Bartlett teszthez a 14.22. táblázat tartalmazza. A 14.22. táblázatból kiszámítjuk

$$M = \gamma \left[a \left(\ln \bar{s}^2 \right) - \sum_i \ln s_i^2 \right]$$

$$= (7) \left[(8)(9.994) - 77.90 \right] = 13.734$$

$$\text{és } C = 1 + \frac{a+1}{3a\gamma} = 1 + \frac{8+1}{(3)(8)(7)} = 1.0536$$

$$\chi^2 = \frac{M}{C} = \frac{13.734}{1.0536} = 13.064$$

Mivel $\chi^2_{0.05(7)} = 14.067 > 13.064$, levonjuk azt a következtetést, hogy a varianciák nem heterogének.

14.22. táblázat. A variancia homogenitásának Bartlett próbája a korai lencse szemtermésére, 1985–1986. években

Termőhely	s^2	$\ln(s^2)$
1	5260	8.568
2	3494	8.159
3	34406	10.446
4	30131	10.313
5	27704	10.299
6	20190	9.913
7	21270	9.965
8	32759	1.397
Összes	175214	77.990

$\gamma = 7$ df a hibára és $a = k = 8$ termőhely, $s^2 = 175\,214/8 = 21\,902$; $\ln = 9.994$

14.4.3. Előzetes analízis

Mivel következtetésünk mindkét tesztből az, hogy a varianciák homogének, folytathatjuk a kombinált analízissel, felhasználva az adatokat mind a nyolc termőhelyről. Először elvégzünk egy előzetes kombinált varianciaanalízist. Ezzel az analízissel áttekintjük a termőhelyek közötti variáció nagyságát, a variációt a kezelések között, és különösen a termőhely \times kezelés interakciót. A korai vetésű lencse adatokra a varianciaanalízist a 14.23. táblázat mutatja. (Jegyezzük meg, hogy az összevont hiba MS erre az analízisre a termőhelyenkénti analízisek hiba MS értékének átlaga.)

14.23. táblázat. A korai vetésű lencse szemtermésének előzetes varianciaanalízise, kombinálva a nyolc termőhelyen, 1985–1986. években

Forrás	d.f.	MS	F-érték
Termőhely	7	701 579	32.03**
Blokk (termőhelyen belül)	8	63 197	
Kezelés	7	386 166	17.63**
Termőhely × Kezelés	49	43 112	1.97 ^{NS}
Összevont hiba	56	21 902	

**Szignyifikáns 1%-os szinten; NS nem szignifikáns. Az összevont hiba MS-t használtuk minden F teszt nevezőjében.

A termőhely × kezelés interakció tesztje egy jelzést ad arról, hogy a kezelések ugyanúgy viselkednek-e egyik termőhelytől a másikig. Egy szignifikáns interakció azt jelenti, hogy a kezelés-hatások változnak egyik termőhelyről a másik termőhelyig. Ebben az esetben az adatok kombinált analízisének mindegyik termőhelyről kevés értelme van. Megtálalhatjuk azonban a termőhelyek alcsoportjait, amelyekre az interakció nem probléma és amelyekre ajánlások adhatók.

Egy nem szignifikáns termőhely × kezelés interakció, másrésztől, nem szükségszerűen jelenti azt, hogy minden értelmes összehasonlítás a kezeléseik között független a termőhelytől. Ennek vizsgálatához a termőhely × kezelés interakciót fel kellene osztani komponensekre, amelyek megfelelnek olyan interakcióknak, amelyek értelmes kezelés összehasonlítások a termőhellyel. Ezeknek az interakció komponenseknek a szignifikanciáját meg kell vizsgálni.

14.4.4. Részletes analízis

A korai vetésű lencse kísérletben a nyolc kezelés $2 \times 2 \times 2$ faktoriális sorozatát alkotja három faktornak: Sitona elleni védekezés, S; műtrágya, F; és gyomok elleni védekezés, W. Használhatjuk a szokásos faktoriális analízist, hogy megvizsgáljuk a főhatásokat és az interakciókat a kezeléseik között és a faktoriális hatások interakcióját a termőhellyel. A varianciaanalízist erre a megosztásra a 14.24. táblázat tartalmazza. Jegyezzük meg, hogy a főhatások és a kezeléseik közötti interakciók mindegyike 1 szabadságfokot foglal le. E faktoriális hatások interakciói a termőhellyel mind 7 szabadságfokot tartalmaznak.

Az interakció komponenseket az összevont hiba MS ellenében teszteljük, hogy meghatározzuk, melyik (ha van ilyen) kezelés összehasonlítás különbözik az egyik termőhelytől a másikig. A korai vetésű lencse adatokra (14.24. táblázat) vonatkozóan látjuk, hogy a termőhely × Sitona és a termőhely × gyom interakciók egyaránt szignifikánsak. Ez valószínűleg az eltérő fertőzések eredménye a Sitona és a gyomok nyolc termőhelyén.

Amikor egy interakció komponens a termőhely × kezelés variancia forrásban szignifikáns, akkor ez arra utal, hogy a kezelés komponens nemcsak a kísérleti hiba variációnak van kitéve, hanem a termőhelynek tulajdonítható variációnak is. Ebben az esetben indokolt használni az interakció komponens MS értékét az F arány nevezőjében a

kezelés összehasonlítás tesztelésére. Ez két dolgot csinál: közel korrekt hiba varianciát használ a tesztben, és csökkenti a szabadságfokot a számlálóban. A hatás az, hogy a teszt konzervatívabbá válik.

14.24. táblázat. A korai vetésű lencse szemtermésének (kg ha^{-1}) részletes varianciaanalízise, kombinálva a nyolc termőhelyen, Szíriában, 1985–1986. években

Forrás	d.f.	MS	F-érték
Blokk (termőhelyen belül)	8	63 197	
Termőhely (L)	7	701 579	
Kezelés	7		
Sitona (S)	1	1 545 280	11.93* F1
Műtrágya (F)	1	94 395	4.31*
SF	1	8 001	0.36
Gyomirtás (W)	1	965 703	13.34** F2
SW	1	44 179	2.02
FW	1	40 970	1.87
SFW	1	4 632	0.21
Termőhely \times Kezelés	49		
LS	7	129 563	5.92*
LF	7	37 948	1.73
LSF	7	8 762	0.40
LW	7	72 395	3.30**
LSW	7	17 715	0.81
LFW	7	24 882	1.14
LSFW	7	10 519	0.48
Összevont hiba	56	21 902	

*Szignyifikáns 5%-os szinten, **Szignyifikáns 1%-os szinten. MS (LS) volt a nevező az F1-nél, MS (LW) volt a nevező az F2-nél; összevont hiba MS volt a nevezőben az összes többi F próbánál.

A korai vetésű lencse adatoknál a nevező MS a Sitona elleni védekezés hatásának tesztelésére $L \times S$ MS, amely ad egy MS értéket 1 és 7 df szabadságfokkal. Hasonlóan az $L \times W$ MS-t használjuk a gyomirtás hatásának tesztelésére. Mivel nincs több szignifikáns interakció a termőhellyel, az összevont hiba MS-t használhatjuk a többi faktoriális hatás tesztelésére. Ezekhez az F értékekhez 1 és 56 df szabadságfokok tartoznak.

A varianciaanalízis eredményét használjuk annak meghatározására, hogy mely átlagokat kell használni a kombinált adatok összegezéséhez. A kombinált analízis interpretációja ugyanazt követi, mint a faktoriális analízis. Ha kezelés \times termőhely interakció áll fenn, akkor kétirányú táblázatokat kell használni. Ha az interakciók a termőhellyel nem szignifikánsak, akkor az átlagok egyirányú táblázatát kell használni a szignifikáns főhatásokra.

Több termőhelyről származó adatoknál felmerül az a probléma, hogyan a legjobb kombinálni az adatokat, hogy becsüljük az átlagokat. Cochran és Cox (1957) arra

a következtetésre jutott, hogy kivéve azokat az eseteket, amelyekben kiugró variáció van a kísérletek közötti pontosságban, a számtani átlagot termőhelyek szerint célszerű előnyben részesíteni bármely súlyozott átlaggal szemben. A kombinált analízis eredményeinek értelmezésénél ezt használjuk.

14.4.5. A statisztikai analízis eredményének értelmezése

Termesztési kísérleteket végeztek nyolc termőhelyen, Szíriában 1985–1986-ban. A kísérletek célja volt, hogy értékeljék három faktor – Sitona elleni védekezés, foszfor trágyázás és gyomirtás – hatását a korai vetésű lencse szemtermésére. A fontosabb eredményeket a 6–8. táblázatban foglaltuk össze. Az $50 \text{ kg ha}^{-1} \text{ P}_2\text{O}_5$ kijuttatása átlagosan 54 kg ha^{-1} terménynövekedést eredményezett. Ez a kismértékű növekedés függetlennek tűnik a termőhelytől (14.25. táblázat).

14.25. táblázat. A korai vetésű lencse átlagos szemtermése (kg ha^{-1}) foszfor műtrágyázással és foszfor műtrágya nélkül, Szíriában, 1985–1986. években

Foszfor	0	50 kg ha^{-1}	Standard hiba
Termés (kg ha^{-1})	892.5	946.8	18.50

14.26. táblázat. A korai vetésű lencse átlagos szemtermése (kg ha^{-1}) Sitona elleni védekezés esetén és védekezés nélkül nyolc termőhelyen, Szíriában, 1985–1986. években

Sitona elleni védekezés	Termőhely							
	1	2	3	4	5	6	7	8
Kontroll	1062.2	1121.2	608.6	720.4	413.8	894.0	703.9	953.9
Karbofuran	1232.4	1140.0	957.0	1287.9	700.5	975.2	773.2	1170.0

Standard hiba = 127.26

A Sitona elleni védekezés hatása termőhelyenként változott (14.26. táblázat). A legnagyobb különbség, 567.5 kg ha^{-1} a kontroll és a nem kontroll között a 4. termőhelyen következett be. A legkisebb különbség, 18.8 kg/ha a 2. termőhelyen volt. A Sitona elleni védekezéshez hasonlóan, a gyomirtás hatása is termőhelyenként eltérő volt (14.27. táblázat). A legnagyobb különbség a herbicid kezelés és a kontroll között 443.0 kg ha^{-1} volt az 5. termőhelyen; a legkisebb különbség 21.9 kg ha^{-1} volt a 8. termőhelyen. Nincs jelzés arra, hogy a három faktor bármelyike a másik két faktor szintjétől függ.

14.27. táblázat. A korai vetésű lencse átlagos szemtermése (kg ha^{-1}) gyomirtás nélkül és herbicid alkalmazásakor, nyolc termőhelyen, Szíriában, 1985–1986. években

Gyomok elleni védekezés	Termőhely							
	1	2	3	4	5	6	7	8
Kontroll	1119.4	1096.0	702.4	888.2	335.6	844.5	625.2	1051.0
Herbicid	1175.2	1165.5	863.2	1120.0	778.6	1024.8	851.9	1072.9

Standard hiba = 95.13

14.5. Ültetvény kísérletsorozatok varianciaanalízise

A kezelések összehasonlításának egy nagyobb programja több kísérletet tartalmaz, melyek mindegyikből kezelésátlagokat kapunk. Ezek az átlagok összegezhetők egy kétirányú táblázatban, amely tartalmazza kísérleti helyeket és kezeléseket további analízisre. További adatok a kísérletenkénti analízisből magukban foglalják a maradék és kezelés MS-eket.

A következő lépés a kétirányú táblázatok analízise abból a célból, hogy genotípus átlagok több sorozatát reprezentálhatja-e az átlagok egyetlen sorozata. Más szavakkal, szükségünk van egy kétfaktoros (hely és faj) varianciaanalízisre, hogy megvizsgáljuk az interakció szignifikanciáját. Ha a hely×faj interakció nem szignifikáns, a kétirányú táblázatot reprezentálhatják sor és oszlop összegeik (faktor főhatások) és egy nagyon egyszerű összefoglalásunk lehet a kísérletsorozatról, a fajok relatív teljesítményében kifejezve. A hely × faj interakció azonban statisztikailag majdnem mindig szignifikáns, magában foglalva, hogy a fajok relatív teljesítménye változik kísérleti helyenként.

A regresszió analízis egyike a nagyszámú módszernek, amely értelmezni tudja a struktúrát a kétirányú táblázatban, amikor az interakció szignifikáns. Ha a kétirányú táblázat komplett, a regresszió analízis módszerét általánosabban *Yates* és *Cochran* (1938) vagy *Finlay-Wilkinson* (1963) módszerként ismerik. Gyakran a genotípusokat nem állítják be minden helyen és a táblázat inkomplett lesz. Ekkor *Yates* és *Cochran* módszerét általánosítani kell, mint erre *Digby* (1979) rámutatott.



14.1. kép. *Acacia auriculiformis* szaporítóanyag-kísérlet Pápua Új-Guineában. A fotó az állomány 3.5 éves korában készült (*Williams et al.*, 2002)

14.5.1. Komplet kétirányú táblázatok

Lehetnek kísérleteink l helyen (termőhelyen), melyeknek mindegyikén v fajtát hasonlítunk össze, és egyesíteni akarjuk az információt minden helyről egyetlen varianciaanalízisben. Ha minden kísérlet RCB elrendezésű r ismétléssel, a variancia táblázat minden helyen a következő formájú:

Variáció forrása	d.f.	MS
Ismétlés	$r-1$	
Fafaj	$v-1$	
Maradék	$(r-1)(v-1)$	s_j^2

Feltételezzük, hogy az analízis parcella átlagadatakon alapul minden kísérletben. A becsült faját átlagokat megkapjuk minden egyes analízisből. Ezeket ezután összegyűjtjük az l kísérleti helyen, amely a becsült faját átlagok $v \times l$ táblázatát adja, amely ezután analízisre bontható kétirányú varianciaanalízisként a következő formában:

Variáció forrása	d.f.	MS
Hely	$l-1$	
Fafaj	$v-1$	
Hely \times faj	$(l-1)(v-1)$	

Szignifikáns-e a Hely \times faj interakció? A hiba tag (s^2) az interakció tesztelésére az s_j^2 -k átlaga az egyedi analízisekből, osztva az ismétlések számával minden kísérleti helyen. Az átlag kiszámítása előtt ellenőrizzük, hogy az s_j^2 -k homogének-e és kombinálhatók-e egymással. *Patterson és Silvey* (1980) javasolta, hogy a közepes négyzetes eltérések 10-szeres nagyságrendű eltérése tolerálható, amikor az összevont maradék MS-t képezzük a kísérleti helyeken át. Ha az arány ennél nagyobb, akkor a kétirányú táblázat súlyozott analízisét kell megfontolni, ahol a súly a becsült faját átlagra a j -edik helyen $1/s_j$. Rendszeresen nem foglalkozunk ilyen bonyolult eljárással (*Williams et al.*, 2002).

Rendszerint a több helyen beállított kísérletekben az ismétlések száma ugyanaz mindegyik kísérleti helyen. Ha nem így van, a kétirányú táblázatok súlyozott analízisét kell elvégezni, ahol a súlyok az ismétlések száma az adott kísérleti helyen. A variancia táblázat formája a megfelelő F-próbához a hely \times faj interakció tesztelésére a következő:

Variáció forrása	d.f.	MS
Hely	$l-1$	
Fafaj	$v-1$	
Hely \times faj	$(l-1)(v-1)$	
Maradék	$l(r-1)(v-1)$	$s^2 = \sum s_j^2 / (lr)$

14.5.2. Kísérlet leírása

1985-ben fafaj kísérleteket állítottak be hat helyen Thaiföldön, az ACIAR projekt részeként több éven keresztül, hogy teszteljék az ausztráliai többféle hasznosítású fafajokat (14.1–14.2. kép). A kísérleti elrendezés minden helyen RCB volt három ismétléssel és a fajok száma 30 és 42 között változott. A parcellák 5×5 fát tartalmaztak, $2 \text{ m} \times 2 \text{ m}$ térállásban. A fagasság méréseket a 24. hónapban végezték. Az analízis a parcella átlagadatokon alapult minden helyen (Williams *et al.*, 2002).

Az 14.28. táblázat tartalmazza a fajok magasságának átlagos értékeit. Mint látható, a táblázatban vannak hiányzó adatok, nem minden fajt természetek minden helyen. A hely \times faj táblázatokban gyakori, hogy hiányzó adatokat tartalmaznak, azonban az első analízisnél a komplett táblázat varianciaanalízisét kívánjuk bemutatni, úgyhogy kihagyjuk a két utolsó kísérleti helyet és azokat a fajokot, amelyek nincsenek jelen mind a négy helyen. Az adatokat a 14.29. táblázat tartalmazza. A kétváltozós varianciaanalízis eredményét a 14.30. táblázat tartalmazza. Figyeljük meg, hogy a táblázatban nincs sor a maradék közepes négyzetes eltérésre (MS). Mint már szó volt róla, ezt a mutatót (s^2) minden egyes analízis összevont maradék MS-ből kapjuk és osztjuk 3-mal, mivel az 14.28. táblázat adatai három ismétlés átlagai.

14.28. táblázat. A fagasság (cm) 24. hónapon végzett mérésének átlagadatai hat helyen Thaiföldön (Williams *et al.*, 2002)

Fajok száma	Hely						
	Fajnév	Ratchaburi	Sai Thong	Si Sa Ket	Sakaerat	Chantaburi	Huai Bong
13877	ACAAUL	334	503	318	304	163	124
13866	ACAAUL	348	413	399	251	205	196
13689	ACAAUL	424	742	606	463	288	254
13688	ACAAUL	439	802	570	375	363	200
13861	ACAAUR	465	858	661	523	379	345
13854	ACAAUR	428	899	658	527	391	316
13686	ACAAUR	567	884	–	–	–	–
13684	ACAAUR	520	941	641	538	363	293
13864	ACACIN	326	571	353	440	166	–
13863	ACACRA	416	799	677	453	–	–
13683	ACACRA	600	1083	738	658	285	–
13681	ACACRA	571	920	679	610	–	272
13680	ACACRA	584	1073	674	–	248	–
14623	ACADIF	608	695	–	–	–	–
14175	ACAFLA	436	624	421	271	137	–
14660	ACAHOL	552	685	415	433	336	251
13691	ACALEP	553	853	735	–	248	167
13653	ACALEP	520	796	658	392	306	189
13846	ACAMAN	473	483	521	283	143	–
13621	ACAMAN	419	497	453	166	187	102
14176	ACAMEL	–	259	158	161	174	134
13871	ACAPOL	256	392	261	177	141	96
14622	ACASHI	467	–	392	205	–	–

14.28. táblázat folytatása

Fajok száma	Hely						
	Fajnév	Ratchaburi	Sai Thong	Si Sa Ket	Sakaerat	Chantaburi	Huai Bong
13876	ALLIT	–	444	311	226	84	136
13519	CASCUN	500	434	420	334	186	176
13514	CASCUN	523	349	330	273	143	123
13148	CASCUN	383	267	220	223	132	158
13990	CASEQU	406	263	252	206	171	–
14537	EUCAM	764	884	839	632	569	337
14106	EUCAM	764	939	809	575	555	397
12013	EUCPEL	673	894	566	332	382	327
14130	EUCTOR	490	475	414	310	415	237
14485	MELBRA	216	116	97	63	80	75
14166	MELDEA	209	–	338	174	–	119
11935	MELDEA	218	319	363	191	76	100
14170	MELSYM	248	–	353	235	–	142
14152	MELVIR	–	–	288	90	111	113

A 2. példában részletesen megvizsgáljuk az s^2 kiszámítását, most azonban az $s^2 = 1040$ számított értéket használjuk 208 d.f.-nél. Így a varianciaanalízis a 14.30. táblázatban manuálisan módosítható, amely a 14.31. táblázatot adja, amelyben a variancia táblázat komplett.

Az F-érték a 14.31. táblázatban erősen szignifikáns, amikor a hely×faj interakció MS-t viszonyítjuk az s^2 -hez (maradék MS). Az interakciót tovább fogjuk vizsgálni, hogy megértsük, mi okozza az interakciót és interpretáljuk az eredményt a *Finlay-Wilkinson (1963)* stabilitásanalízissel. Először azonban bemutatjuk az inkomplett kétirányú táblázat analízisét.

14.29. táblázat. A fajok átlagos magasságának értékei négy helyre és 27 fajra a 14.28. táblázat adatai alapján

Hely No.	Faj No.	Mag. cm	Hely No.	Faj No.	Mag. cm	Hely No.	Faj No.	Mag. cm	Hely No.	Faj No.	Mag. cm
1	1	334	2	1	503	3	1	318	4	1	304
1	2	348	2	2	413	3	2	399	4	2	251
1	3	424	2	3	742	3	3	606	4	3	463
1	4	439	2	4	802	3	4	570	4	4	375
1	5	465	2	5	858	3	5	661	4	5	523
1	6	428	2	6	899	3	6	658	4	6	527
1	7	520	2	7	941	3	7	641	4	7	538
1	8	326	2	8	571	3	8	353	4	8	440
1	9	416	2	9	799	3	9	677	4	9	453
1	10	600	2	10	1083	3	10	738	4	10	658
1	11	571	2	11	920	3	11	679	4	11	610
1	12	436	2	12	624	3	12	421	4	12	271

14.29. táblázat folytatása

Hely No.	Faj No.	Mag. cm	Hely No.	Faj No.	Mag. cm	Hely No.	Faj No.	Mag. cm	Hely No.	Faj No.	Mag. cm
1	13	552	2	13	685	3	13	415	4	13	433
1	14	520	2	14	796	3	14	658	4	14	392
1	15	473	2	15	483	3	15	521	4	15	283
1	16	419	2	16	497	3	16	453	4	16	166
1	17	256	2	17	392	3	17	261	4	17	177
1	18	500	2	18	434	3	18	420	4	18	334
1	19	523	2	19	349	3	19	330	4	19	273
1	20	383	2	20	267	3	20	220	4	20	223
1	21	406	2	21	263	3	21	252	4	21	206
1	22	764	2	22	884	3	22	839	4	22	632
1	23	764	2	23	939	3	23	809	4	23	575
1	24	673	2	24	894	3	24	566	4	24	332
1	25	490	2	25	475	3	25	414	4	25	310
1	26	216	2	26	116	3	26	97	4	26	63
1	27	218	2	27	319	3	27	363	4	27	191

14.5.3. Inkomplett kétirányú táblázatok

Egy kétirányú táblázatot adathiányokkal inkomplett vagy nem-ortogonális kétirányú táblázatnak hívjuk. Nehézségek adódnak az inkomplett kétirányú táblázatok analízisének. A becslt átlagokat nem lehet képezni egyszerű sor vagy oszlop átlagként. Ennek oka egészen nyilvánvaló. Tételezzük fel, hogy egy speciális faj magját minden helyen elvetették, kivéve a legjobb helyet. Nem volna korrekt becsülni ennek a fajnak az átlagát oly módon, hogy átlagoljuk az összes helynek az eredményét, ahol jelen volt. Azok a fajok, amelyeknek magjaik jelen voltak a legjobb helyen, előnyt élveznének ily módon. Jobb felállítani egy statisztikai modellt a kétirányú táblázatra, azután becsülni a főhatásokat, felhasználva a legkisebb négyzetek módszerét. A lineáris modell feltételezi, hogy egy adat a kétirányú táblázatban három komponensre bontható fel: első a főátlag; a második az átlagnak egy kiigazítása, figyelembe véve a sajátos fajajt; a harmadik egy másik kiigazítás a sajátos helyre. Ezt így írhatjuk fel:

(megfigyelés) = (főátlag) + (hely hatás) + (fafaj hatás) + (hely × faj hatás)
vagy szimbolikusan:

$$Y_{ij} = \mu + \theta_i + \tau_j + \theta\tau_{ij}$$

ahol Y_{ij} ($i = 1, 2, \dots, v$) az adatok a hely × faj táblázatban; μ egy paraméter a főátlagra; θ_i és τ_j paraméterek a helyre és a fajra; $\theta\tau_{ij}$ interakció hatások a helyek és a fajok között.

Az ANOVA a GenStatban használható az inkomplett táblázatokra, helyettesítve az adathiányokat a táblázatban becsült értékekkel. A varianciaanalízis eredménye ily módon csupán megközelítés, és különösen azzá válik, ha az adathiányok száma nő. Nem volna bölcs dolog használni az ANOVA-t, ha kettőnél vagy háromnál több hiányzó értéket kell becsülni.



14.2. kép. *Casuarina junghuniana* hét éves állománya Thaiföldön. A mély vízvezető árkokat figyelembe kell venni a kísérlet tervezésnél (Williams et al., 2002)

14.5.4. Kísérlet leírása

A 14.28. táblázat kéttényezős analízisének a GenStat program feltételezi, hogy az adatok a javasolt formában vannak, mint ahogy a 14.29. táblázat mutatja. A különbség itt az, hogy jöllehet az adatfájl $37 \times 6 = 222$ méretű, 30 átlagos magasság adat hiányzik, melyet a * szimbólum jelöl a GenStatban. Az eredményként kapott varianciaanalízis táblázatot az 14.32. táblázat tartalmazza. Hasonlóan a 14.30. táblázathoz, az összevont maradék közepes négyzetes eltérést az egyes kísérletek analíziséből kézzel (manuálisan) írjuk be a 14.32. táblázatba. A maradék MS részletes adatok a magasság adatok analíziséből a hat hely mindegyikén a 14.33. táblázatban találhatók. Ez a táblázat egy általános összefoglalását adja az eredményeknek az egyedi varianciaanalízisből. A 14.33. táblázat segít megmagyarázni, hogy a variancia homogenitásának feltétele teljesül-e. A maradék MS (s_f^2) tartománya 891 cm^2 -től 5852 cm^2 -ig változik. Mint korábban említettük, a maradék MS tízszeres tartománya tolerálható, mielőtt egy súlyozott analízist szükségesnek tartanánk. Tehát a 14.32. táblázat nem súlyozott analízisét támogatja a 14.33. táblázat eredménye:

14.30. táblázat. A varianciaanalízis eredménye az átlagos fmagasság értékekkel, a 14.29. táblázat adatai alapján

Analysis of variance							
Variate: Height							
Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.		
Site	3	919585.	306528.				
Seedlot	26	3176289.	122165.				
Site. Seedlot	78	707957.	9076.				
Total	107	4803831.					
Tables of means							
Variate: Height							
Grand mean	488.46						
Site	1	2	3	4			
	461.63	627.70	494.04	370.48			
Seedlot	1	2	3	4	5	6	7
	364.75	352.75	558.75	546.50	626.75	628.00	660.00
Seedlot	8	9	10	11	12	13	14
	422.50	586.25	769.75	695.00	438.00	521.25	591.50
Seedlot	15	16	17	18	19	20	21
	440.00	383.75	271.50	422.00	368.75	273.25	281.75
Seedlot	22	23	24	25	26	27	
	779.75	771.75	616.25	422.25	123.00	272.75	

14.31. táblázat. Variancia táblázat a 14.29. táblázat adatai alapján, manuálisan kiegészítve a maradék közepes négyzetes eltérés értékével

Analysis of variance					
Variate: Height					
Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Site	3	919585.	306528.		
Seedlot	26	3176289.	122165.		
Site. Seedlot	78	707957.	9076.	8.73	<.001
Residual	208		1040.		

A következőkben ki kell számítani az összevont maradék MS-t (s^2), mely hozzáilleszhető a 14.32. táblázathoz. A legtöbb növényfaj kísérleti programban, ahol fajokat hasonlítunk össze különböző helyeken, a maradék MS mindegyike ugyanazzal a d.f.-fel rendelkezik, tehát könnyebb egyszerűen a maradék MS értékek átlagát venni. Kicsit óvatosabbnak kell lennünk, ha a szabadságfokok száma a maradék közepes négyze-

tes eltérésekre nagyon kicsi, pl. kevesebb, mint 20. Példánkban elegendő a maradék MS szabadságfoka minden helyen, tehát elfogadható, hogy vegyük az átlagát a 14.33. táblázatban található maradék közepes négyzetes eltéréseknek. Ez az érték 2891 cm^2 , melyet ezután elosztunk az ismétlések számával minden helyen (azaz hárommal), amely $s^2 = 964 \text{ cm}^2$ értéket ad. A szabadságfokok száma (384) az s^2 -re a maradék szabadságfokok számának összege minden helyre. A végső varianciaanalízist a 14.34. táblázat mutatja, amely tartalmazza az összevont maradék MS-t a 14.33. táblázat alapján. A hely \times faj faj interakció F-értéke 150 és 384 d.f.-re 9.35, amely erősen szignifikáns. Ezt az interakciót tovább vizsgáljuk a későbbiekben.

14.32. táblázat. A famagasság adatok varianciaanalízise az 14.29. táblázat adatai alapján

Analysis of an unbalanced design using GenStat regression

Variate: Heightcm

Accumulated analysis of variance

Change	d.f.	s.s.	m.s.
+ SiteNo	5	4157543.	831509.
+ SeedlotNo	36	4425296.	122925.
+ SiteNo.SeedlotNo	150	1351054.	9007.
Total	191	9933893.	52010.

14.33. táblázat. Az egyes kísérleti helyek varianciaanalízisének összefoglaló táblázata az átlagos famagasságra, a közepes négyzetes eltérés adatokkal

Hely	Közepes négyzetes eltérés (MS)			
	Ismétlés	Fafaj	Maradék	Parcellán belül
Ratchaburi	34376	63700	3068	364
Shai Thong	85252	216149	5852	846
Si Sa Ket	13521	112267	1866	430
Sakerat	61971	82155	1804	381
Chanthaburi	128691	58430	3865	526
Huaia Bong		23893	891	274

14.34. táblázat. A famagasság adatok varianciaanalízise az 14.29. táblázat adatai alapján, hozzáírva manuálisan a maradék közepes négyzetes eltérést

Analysis of an unbalanced design using GenStat regression

Variate: Heightcm

Accumulated analysis of variance

Change	d.f.	s.s.	m.s.	v.r.	F pr.
+ SiteNo	5	4157543.	831509.		
+ SeedlotNo	36	4425296.	122925.		
+ SiteNo.SeedlotNo	150	1351054.	9007.	9.35	<.001
Residual	384		964.		

14.5.5. Regresszió analízis a komplett kétirányú táblázatokkal

Az 1. példa varianciaanalízisének eredménye (4.31. táblázat) tipikus, mivel a faj \times genotípus interakció rendszerint erősen szignifikáns.

Számos módszer áll rendelkezésre, hogy értelmezzük a kétirányú táblázatok struktúráját. Az általános terminológia az ilyen interakcióra a genotípus \times környezet interakció. Az 1. példa esetében a genotípusok és környezetek megfelelnek a faj és hely faktoroknak. A következőkben a regresszió analízis módszerével vizsgáljuk a hely \times faj interakció okát. Ez az egyik legegyszerűbb módszer és az interakciót minden fajra kiszámított regressziós koefficienssel jellemzi. A modell a következő:

(megfigyelés) = (főátlag) + (genotípus regressziós koefficiens) \times (hely hatás) + (genotípus hatás) + (maradék)

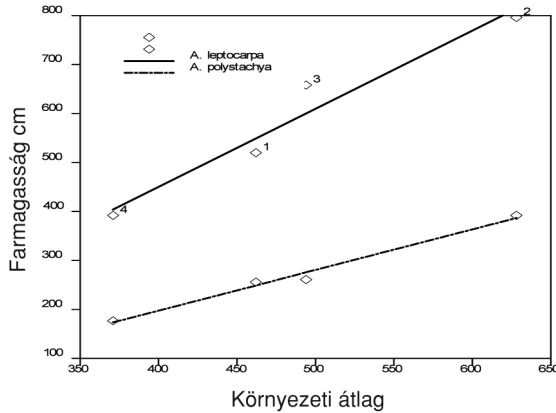
vagy szimbolikusan:

$$Y_{ij} = \mu + \gamma_j \theta_i + \tau_j + \varepsilon_{ij}$$

ahol: μ : főátlag; θ_i : hely hatás; τ_j : genotípus hatás; γ_j : genotípus regressziós koefficiense, a hely hatás szorzója.

Ha nincs interakció, akkor az összes genotípus regressziós koefficiens egyenlő eggyel és a modell leegyszerűsödik a korábbi verzióra. Rendszerint van valamekkora interakció és γ_j egy körül ingadozik, egyes genotípusoké nagyobb, másoké kisebb egynél.

A paraméterek becslése nem bonyolult a komplett kétirányú táblázatra és ezt mutatjuk be a következőkben. Eltekintve a főátlagtól, a paraméterek három sorozatát kell becsülni; nevezetesen a hely hatásokat, a genotípus hatásokat és a genotípus regressziós koefficienseket. Egy komplett kétirányú táblázatra, a hely és genotípus hatásokat megkapjuk a táblázatok oszlop és sor összegeiből. Minden helyen a genotípusok és a környezeti átlag közötti regressziót számítjuk ki és minden genotípusra kapunk egy regressziós koefficienset. A módszert az 1. példa adataival mutatjuk be. A becsült hely és genotípus átlagok a 14.29. táblázat adatai alapján a 14.30. táblázatban található, a varianciaanalízis részeként. A regressziós koefficiensek úgy határozhatók meg, hogy kiszámítjuk a regressziót az egyes genotípusok kísérleti helyenkénti értékei és a kísérleti helyenként a genotípusok átlagértékei között. A 14.30. táblázatból a kísérleti hely átlagok 462, 628, 494 és 371 cm. Most a 14.28. táblázatban válasszuk ki a 13653 számú genotípust (*Acacia leptocarpa*), melynek kísérleti helyenkénti értéke 520, 796, 658 és 392 cm az első négy hely alapján. Az *Acacia leptocarpa* értékek és a kísérleti hely átlagértékek közötti lineáris regresszió analízis alapján a regressziós koefficiens 1.59. Másrésről, a 13871 számú *Acacia polystachya* kísérleti helyenkénti értékei 256, 392, 261 és 177, a regressziós koefficiens értéke 0.83. E két genotípus adatait az 14.3. ábra mutatja, ahol a hely szerinti átlagok vannak a horizontális tengelyen. A regressziós egyeneseket és a mérési adatokat mutatja az ábra. Az *Acacia leptocarpa* esetében a regressziós egyenes meredeksége nagyobb, mint egy; ez azt jelenti, hogy a genotípus nagy különbségeket mutat teljesítményben a helyek között, gyengén nő a gyenge termőhelyen és erőteljesen nő a jó termőhelyen. Az ilyen genotípust „nem stabilnak” tekintjük. Ezzel szemben, az *Acacia polystachya* regressziós koefficiense egyenél kisebb, amely azt jelenti, hogy a genotípus következetesebben teljesít az átlagosnál gyengébb és jó termőhelyeken egyaránt. Az ilyen genotípust „stabilnak” nevezzük.



14.3. ábra. Két faj átlagos magasságának ábrázolása versus környezeti átlag (minden faj átlagos magassága mindegyik kísérleti helyen)

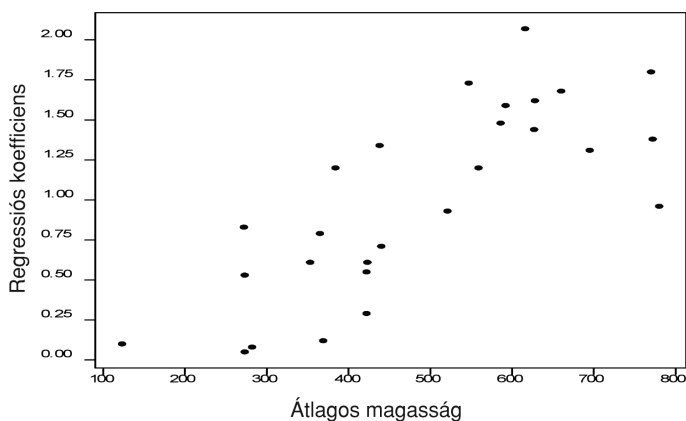
A szántóföldi kísérletsorozat eredménye alapján a megfelelő genotípus kiválasztásához a genotípus stabilitásának ismerete különösen fontos. Hasonló növekedési rátát feltételezve, egy stabil genotípust előnyben részesíthetünk eltérő minőségű helyeken, míg egy nem-stabil genotípust jobb termőhelyeken indokolt előnyben részesíteni. A 14.35. táblázat tartalmazza a számított regressziós koefficienseket mind a 27 fajra, továbbá feltüntettük a becsült átlagos magasságukat is a 14.30. táblázatból.

14.35. táblázat. Átlagos famagasságok és regressziós koefficiensek a 14.29. táblázat adataira

Faj száma	Átlagos magasság, cm	Regressziós koefficiens
13877	365	0.79
13866	353	0.61
13688	559	1.20
13689	547	1.73
13861	627	1.44
13854	628	1.62
13684	660	1.68
13864	423	0.61
13863	586	1.48
13683	770	1.80
13681	695	1.31
14175	438	1.34
14660	521	0.93
13653	592	1.59

Faj száma	Átlagos magasság, cm	Regressziós koefficiens
13846	440	0.71
13621	384	1.20
13871	272	0.83
13519	422	0.29
13514	369	0.12
13148	273	0.05
13990	282	0.08
14537	780	0.96
14106	772	1.38
12013	616	2.07
14130	422	0.55
14485	123	0.10
11935	273	0.53

A regresszió analízis modell szükségessé teszi genotípusok regressziós koefficiensein kívül a hely \times faj interakció okának feltárását is. A regressziós koefficienseknek kevés az értéke, ha nem segítik elő az interakció értelmezését. Az interakció eltérésnégyzet összegét a 14.31. táblázatban felbonthatjuk két komponensre: (i) regresszió négyzetösszeg és (ii) eltérés a regressziótól. Ha a regresszió közepes négyzetes eltérés szignifikáns és a közepes négyzetes eltérés a regressziótól nem szignifikáns, akkor a regressziós modell sikeres volt az interakció értelmezésében. A regresszió variancia táblázatát példánkra a 14.36. táblázat mutatja. A közepes négyzetes eltérés a regresszióra 11 866, összehasonlítva az eltérés 7682 értékével. Nyilvánvalóan az eltérés még erősen szignifikáns, ami azt jelenti, hogy a regresszió analízis modell nem volt teljesen sikeres az interakció értelmezésében. Mindamellet a fajok regressziós koefficiensei a 14.35. táblázatban egy magyarázatot adnak bizonyos interpretációkra. Nagyon hasznos ábrázolni a fajok regressziós koefficienseit a fajok átlagos értékei szerint, ahogy a 14.4. ábra mutatja. A nem-stabil fajok magas átlagértékekkel az ábra felső jobboldali részén vannak, a stabil fajok magas átlagértékekkel (az ábra alsó jobboldala) ugyancsak fontosak lehetnek.



14.4. ábra. A regressziós koefficiensek ábrázolása növénymagasságtól függően az egyes fajokra

14.36. táblázat. A variancia regresszió analízise a 14.29. táblázat adatai alapján, manuálisan kiegészítve a maradék közepes négyzetes eltérés értékével

Analysis of variance

Variate: Height

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Site	3	919585.	306528.		
Seedlot	26	3176289.	122165.		
Site.Seedlot	78	707957.	9076.		
Regressions	26	308503	11866	11.41	<.001
Deviations	52	399454	7682	7.39	<.001
Residual	208		1040		

14.5.6. Regresszió analízis inkomplett kétirányú táblázatokkal

Gyakran előfordul a kísérletsorozatoknál, hogy az átlagokból képzett hely×faj táblázat hiányos. Ennek több oka lehet: (i) A fajt nem minden kísérleti helyen állították be, mivel korábbról ismert volt, hogy a fajnak rossz lesz a teljesítménye, pl. fagyérzékeny nagy magasságú helyen. (ii) A faj be volt állítva kísérletbe minden helyen, azonban nem minden helyen maradt életben. (iii) Néhány fajnak nem volt elegendő szaporítóanyaga, hogy beállítsák minden helyen.

Bonyolultabb alkalmazni a regressziós modellt az inkomplett kétirányú táblázatok esetében. A modell ugyanaz marad, mint az előzőekben, a problémát a három paraméter becslése okozza. Másrészről, egy faj, amelyik csak egy helyen szerepel a kétirányú táblázatban, valószínűleg gyenge teljesítményű, és nem maradt életben a többi helyen. Előnyösebb lehet kizárni ezt a fajt a regresszió analízisből, mivel számítási nehézségeket okozhat és kevés információval járul hozzá az eredményekhez. Egy elfogadott vélemény az, hogy azokat a fajokat tartsuk meg az analízisben, amelyek legalább három helyen előfordulnak. Ha ennél kevesebb helyen előforduló fajokat hagyunk meg az analízisben, akkor a fajok átlagértékei és regressziós koefficiensei érzékenyek lesznek.

14.37. táblázat. Átlagos famagasságok és regressziós koefficiensek a 14.28. táblázat adataira

No	Fafaj száma	Kód	Species	Átlag	Regr. koeff.
1	13877	a	<i>Acacia aulacocarpa</i>	291	0.86
2	13866	a	<i>Acacia aulacocarpa</i>	302	0.60
3	13689	a	<i>Acacia aulacocarpa</i>	463	1.18
4	13688	a	<i>Acacia aulacocarpa</i>	458	1.30
5	13861	b	<i>Acacia auriculiformis</i>	539	1.20
6	13854	b	<i>Acacia auriculiformis</i>	537	1.29
7	13686	b	<i>Acacia auriculiformis</i>	510	1.63
8	13684	b	<i>Acacia auriculiformis</i>	549	1.46
9	13864	c	<i>Acacia cincinnata</i>	336	0.92
10	13863	d	<i>Acacia crassicsrpa</i>	465	1.49
11	13683	d	<i>Acacia crassicsrpa</i>	594	2.04
12	13681	d	<i>Acacia crassicsrpa</i>	572	1.46
13	13680	d	<i>Acacia crassicsrpa</i>	519	2.25
14	14623	e	<i>Acacia difficilis</i>	592	0.45
15	14175	f	<i>Acacia flavescens</i>	327	1.33
16	14660	g	<i>Acacia holosericea</i>	445	0.92
17	13691	h	<i>Acacia leptocarpa</i>	500	1.73
18	13653	h	<i>Acacia leptocarpa</i>	477	1.46
19	13846	i	<i>Acacia mangium</i>	342	1.01
20	13621	i	<i>Acacia mangium</i>	304	1.02
21	14176	j	<i>Acacia melanoxylon</i>	179	0.23

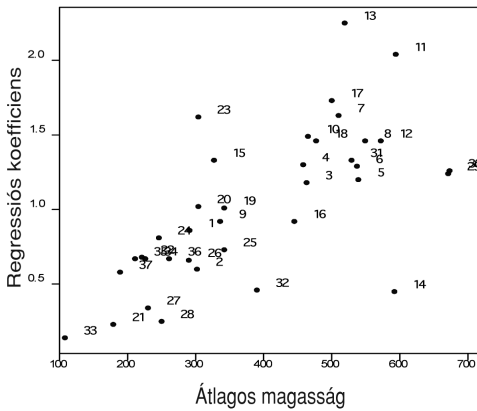
14.37. táblázat folytatása

No	Fafaj száma	Kód	Species	Átlag	Regr. koef.
22	13871	k	<i>Acacia polystachya</i>	221	0.68
23	14622	l	<i>Acacia shirleyi</i>	304	1.62
24	13876	m	<i>Allocasuarina littoralis</i>	246	0.81
25	13519	n	<i>Casuarina cunningghaiana</i>	342	0.73
26	13514	n	<i>Casuarina cunningghaiana</i>	290	0.66
27	13148	n	<i>Casuarina cunningghaiana</i>	230	0.34
28	13990	o	<i>Casuarina equisetifolia</i>	250	0.25
29	14537	p	<i>Eucalyptus camaldulensis</i>	671	1.24
30	14106	p	<i>Eucalyptus camaldulensis</i>	673	1.26
31	12013	q	<i>Eucalyptus pellita</i>	529	1.33
32	14130	r	<i>Eucalyptus torellina</i>	390	0.46
33	14485	s	<i>Melaleuca bracteata</i>	108	0.14
34	14166	t	<i>Melaleuca dealbata</i>	226	0.67
35	11935	t	<i>Melaleuca dealbata</i>	211	0.67
36	14170	u	<i>Melaleuca symphyocarpa</i>	261	0.67
37	14152	v	<i>Melaleuca viridiflora</i>	189	0.58

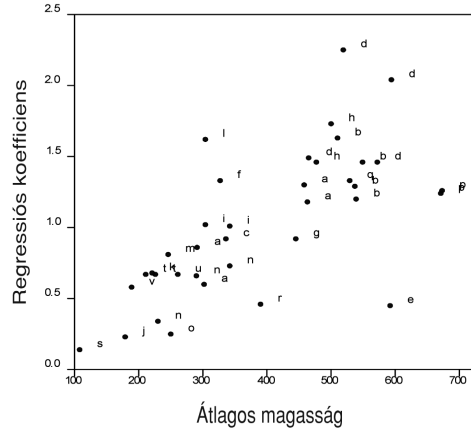
Kísérleti hely	Átlagmagas.	Kísérleti hely	Átlagmagas.	Kísérleti hely	Átlagmagas.
<i>Ratchaburi</i>	425	<i>Si Sa Ket</i>	481	<i>Chanthaburi</i>	258
<i>SaiThong</i>	620	<i>Sakaerat</i>	359	<i>Huai Bong</i>	198

Végső soron, szükségünk van egy módszerre, hogy becsüljük a regresszió analízis paramétereit az inkomplett kétirányú táblázatok esetében. A GenStat program eredménye az 14.37. táblázatban található. Figyeljük meg, hogy bevontuk az analízisbe azokat a fajokat, melyek csupán két helyen voltak a 14.28. táblázatban, nevezetesen a 13686 és 14623 jelűeket. Mivel a módszernek nehézségei lehetnek a fajátlagok és a regressziós koeficiensek elkülönítésében, a hiányosan reprezentált fajoknál szükséges ellenőrizni, hogy ne kapjunk helytelen eredményt. Valóban, a regressziós koeficiensek ábrája a fajátlagadatok szerint (14.5. ábra) mutatja, hogy a 14623 faj (14-es számú az ábrán) valamelyest elkülönül és valószínűleg megfelelő lenne újból elvégezni az analízist e faj kizárásával. A számok a 14.5. ábrán megfelelnek a 14.37. táblázat első oszlopában levő számoknak, hogy lehetővé tegyék a fajok könnyű azonosítását. Az ábra rávilágít arra, hogy a 11 és 13 számú fajok (mindkettő *Acacia crassicarpa*) nem stabilak. A legmagasabb fajok a 29 és 30 számúak (mindkettő *Eucalyptus camaldulensis*). Nyilvánvaló a 14.5. ábrából, hogy ugyanabból a fajtól származó utódok együtt csoportosulnak. Rávilágítunk erre az ábra újbóli elkészítésével, felhasználva faj kódokat (a-v) a 14.37. táblázatban (lásd 14.6. ábrát). Az ábra nagyon világosan mutatja, hogyan összegezhetők a kísérletek fajonkénti eredményei a kísérleti helyeken keresztül. Az *Acacia crassicarpa* (d kód) négy származéka csoportosul egy kívánatos pozícióban az ábrán (magas faj átlag és regressziós koeficiens).

A 14.5. és 14.6. ábrák a teljes adatsor megfelelő összegezését adják. Ellenőrizni kell azonban, hogy a regresszió analízis felelős-e elegendő interakcióért. A varianciaanalízis táblázatot (nem-lineáris analízis módszer alkalmazása Ng és Williams, 2001 alapján) a 14.38. táblázat mutatja.



14.5. ábra. A regressziós koefficiensek ábrázolása az egyes fajok átlagos magasságától függően



14.6. ábra. A regressziós koefficiensek ábrázolása az egyes fajok átlagos magasságától függően, minden faj kódolásával

14.38. táblázat. A variancia regresszió analízise az 14.28. táblázat adatai alapján, az összevont maradék közepes négyzetes eltérés manuális beírásával

Analysis of an unbalanced design using GenStat regression

Variate: Heightcm

Accumulated analysis of variance

Variáció forrása	d.f.	s.s.	m.s.	.v.r.	F pr.
+ Site	5	4157543.	831509.		
+ Seedlot	36	4425296.	122925.		
+ SiteNo.Seedlot	150	1351054.	9007.		
Regression	36	764676	21241	4.13	<.001
Deviations	114	586416	5144	5.34	<.001
Residual	384		964		

Összehasonlíthatjuk a regressziós modellt egy olyan modellel, amelyik csak a főhatásokat illeszti, megvizsgálva a regresszió és eltérés MS arányát. Az F-értéke 4.13 lesz, 36 és 113 d.f.-nél, amely erősen szignifikáns ($P < 0.001$), vagyis jelentős hatást értelmez a regressziós modellel. Az eltérés és az összevont maradék MS aránya alapján az F-próba azonban még erősen szignifikáns. Jóllehet a 14.5. és 14.6. ábrák az interakció nagyon értékes interpretációját adják, tudatában kell lennünk annak, hogy megvan a lehetősége, hogy további információt kapjunk az interakcióból.

15. Üzemi kísérletek

A mezőgazdasági kutatás hagyományosan a kutató intézetekben és egyetemeken történik, ahol a kísérleti eszköz-ellátottság jó és a kutatók is rendelkezésre állnak. Gyakran feltételezik, hogy a legjobb technológia a kutató intézetekben egyúttal a legjobb a gazdaság tábláin is. A technológiai teljesítmény következetességének vagy megismételhetőségének feltétele azonban a kutatóintézetek és a gazdaságok táblája között nem tekinthető általánosnak.

Ha nincs következetesség a technológiai teljesítményben a kutatóintézetek és a gazdaságok táblái között, a legjobb technológia kiválasztása a termelőknek nem alaphoz járhat kizárólagosan kutató intézeti kísérleteken. Ilyen szelekciós eljárásnak valójában termelői kísérleteken kellene alapulni, amelyekben az új technológiát összehasonlítjuk a termelői gyakorlattal a gazdaság termelési feltételei között.

A kísérleteket a gazdaság tábláin osztályozhatjuk, mint technológia-fejlesztési és technológia-összehasonlítási kísérletek. A technológia-fejlesztési kísérletek célja új termesztéstechnológiák fejlesztése, amelyek növelik a termést vagy csökkentik a termelési költségeket. A technológia-összehasonlítási kísérletek célja, hogy összehasonlítsuk az új technológiák fölényét – melyek ígéretesnek bizonyultak a technológia-fejlesztési kísérletekben – a termelő jelenlegi gyakorlatával.

A standard kísérleti és statisztikai eljárások, melyeket a korábbi fejezetekben áttekintettünk, elsősorban a kutatóintézetekben érvényesek a technológia-fejlesztési kísérletekre. Minthogy határozott különbség van a kutatóintézetek és a gazdaság táblája, mint tesztelő hely között, nem minden eljárás megfelelő a termelő tábláján végzett kísérletekre. Mivel azonban a technológia-fejlesztési kísérlet elsődleges célja ugyanaz a kutatóintézetekben és a gazdaság tábláján, az eddig tárgyalt statisztikai eljárások kismértékű módosítása szükséges a technológia-fejlesztési kísérletekben történő alkalmazáshoz. Másoldalról, a technológia-összehasonlítási folyamat jelentősen eltér a technológia-fejlesztéstől és ezáltal teljesen eltérő biometriaival megközelítést igényel.

A gazdaság alapvetően termelési egység, melynek elsődleges célja a termeléyenység, a jövedelem és a család jólétének fokozása. Következésképpen, a kutatónak, aki kísérletet végez a termelő tábláján, meg kell ismerni a kísérleti terület sajátosságait, ahol a termelés, és nem a kutatás a fő prioritás. Néhány fontos megkülönböztető sajátossága a termelő táblájának, mint tesztelő helynek, a kutatóintézethez viszonyítva, a következő:

- Kísérleti lehetőségek (vízszabályozás, a növényvédelem), kísérleti eszközök (parcella vetőgép, kombájn, stb.) hiánya vagy korlátozott megléte.
- Nagy variáció a gazdaságok (farmok) között és a táblák között egy gazdaságban.
- Nehéz hozzáférhetőség, amely problémát okoz a kísérlet ellenőrzésében.
- A kísérleti tábla talaj és klimatikus adottságaira vonatkozó adatok hiánya.
- A termelő és eszközeinek rendelkezésre állása a kísérletekhez.

15.1. Technológia-fejlesztési kísérletek

15.1.1. A tesztelő hely kiválasztása

A technológia-fejlesztési kísérlet tesztelő helyét úgy választjuk ki, hogy a fizikai és biológiai feltételek megfelelő sorozatát adja, amelyekre a technológiát kifejlesztették. Tehát a hely kiválasztásának módszere szándékos, nem véletlen. A kiválasztási eljárások a következők:

1. Lépés. Világosan specifikáljuk a kívánatos tesztkörnyezetet specifikus fizikai és biológiai jellemzőkben kifejezve, mint a talaj, a klíma, a topográfia, táj, vízellátottság stb.

2. Lépés. A specifikált környezeti jellemzők mindegyikét csoportosítsuk az alábbiak szerint:

- Összefüggő terület relatív nagysága, amelyben egy adott jellemző homogenitása várható.
- A szükséges információ rendelkezésre állása, vagy beszerzésének relatív könnyősége, a kívánt jellemzőre vonatkozóan. Például, klimatikus adatok rendszerint könnyebben hozzáférhetőek, mint az információk a vízellátottságra és vetésforgóra.

3. Lépés. Válasszunk ki egy nagy összefüggő területet, amely megfelel azoknak a környezeti sajátosságoknak, amelyek rendszerint homogének egy nagy területen; és ezen a területen belül, azonosítsunk al-területeket (gazdaságokat), amelyek megfelelnek azoknak a környezeti feltételeknek, amelyek variábilisabbak. Például, egy nagy összefüggő terület először kiválasztható, hogy megfeleljen a szükséges klímának és talajnak. Ezek alapulhatnak meteorológiai állomások adatain és talajtérképeken. A kiválasztott területen belül azonosítsuk azokat a gazdaságokat, amelyeknek topográfiája, tájtérképe, vízellátottsága és vetésforgója legjobban közelít a szükséges teszt környezethez.

4. Lépés. Mindegyik kiválasztott gazdaságban válasszunk ki egy területet vagy táblát, amely elég nagy a kísérlet beállításához és legkisebb a talaj heterogenitása.

15.1.2. A kísérletek elrendezése

A kísérletek tervezésének célja a termelő tábláján, hogy tartsuk korlátok között a kísérlet méretét. Ezt elérhetjük a kezelések számának és az ismétlések számának minimumon tartásával. A kezelések számának csökkentése elérhető frakcionális faktoriális kísérlet tervezéssel (*Gomez és Gomez, 1984*). Az ismétlések számának meghatározásakor két ellentétes szempontot kell megvizsgálni:

- A hibák rendszerint fokozódnak az üzemi kísérletekben, mivel a terület kevésbé hozzáférhető és nehéz a kísérlet ellenőrzése. Ezen kívül, a különböző kártételek gyakrabban előfordulnak gazdaságokban, amelyek növelik a kísérleti hibát.
- Másrészt a gyakran kisebb rovar és betegség kártétel és az egységesebben művelt táblák kisebb kísérleti hibát eredményezhetnek. Tehát az ismétlések számának meghatározása függ e két ellentétes hatás relatív fontosságától.

A parcellák elhelyezésénél a termelő tábláján általában ugyanazt az eljárást alkalmazzuk, mint a kutatóintézetben. Az alábbi szempontok figyelembevétele javasolható:

- A parcella alakja igazodjék a tábla alakjához.
- Ha egy tábla nem elegendő az egész kísérlet beállításához, akkor a tábla legalább egy teljes ismétlést foglaljon magában.

15.1.3. Adatgyűjtés

A kutatóintézetben szokásos mérések és felvételezések az üzemi kísérletekben is elvégezhetők. Mivel a technológia-fejlesztési kísérletek elsődleges célja a termelő tábláin a környezet tág körének megmintázása, általános a technológia-fejlesztési kísérleteknél, hogy több gazdaságban állítjuk be. Ahhoz, hogy hatékonyan becsüljük az interakció hatást a kezelés és a gazdaság környezete között, azonos adatsorozatot kell begyűjteni mindegyik gazdaságban.

15.1.4. Adatelemzés

A technológia-fejlesztési kísérletek célja ugyanaz marad, akár kutatóintézetben, akár a gazdaság tábláján állítjuk be. Következésképpen, a kísérletek adatelemzése azonos. Amikor egy kísérlet egynél több gazdaságban van beállítva, az eljárás követi a több helyen beállított kísérletsorozat analízisét. Mivel a gazdaságok különböző környezetet reprezentálnak, a hangsúlyt a kezelés és gazdaság közötti interakció értelmezésére kell helyezni, hogy megértsük a környezeti faktorok befolyását a vizsgált kezelésekre hatékonyására.

Kísérlet leírása

Az eljárás illusztrálására egy gyomirtási kísérlet adatait használjuk, amelyet négy gazdaságban állítottak be, RCB elrendezésben, öt kezeléssel és három ismétlésben gazdaságonként (15.1. táblázat).

15.1. táblázat. Rizs szemtermése egy RCB elrendezésű, három ismétlést és öt kezelést tartalmazó gyomirtási kísérletben, melyet négy gazdaságban állítottak be (Gomez és Gomez, 1984)

Kezelések	Szemtermés t ha ⁻¹			
	I. ismétlés	II. ismétlés	III. ismétlés	Átlag
1. Gazdaság				
T1	5.5	4.8	5.3	5.2
T2	0.6	2.5	1.2	1.4
T3	0.2	0.1	0.4	0.2
T4	5.7	5.2	4.9	5.3
T5	0.3	0.2	0.3	0.3

15.1. táblázat folytatása

Kezelések	Szemtermés t ha ⁻¹			
	I. ismétlés	II. ismétlés	III. ismétlés	Átlag
2. Gazdaság				
T1	5.6	5.8	6.0	5.8
T2	6.8	3.1	4.4	4.8
T3	5.7	5.7	5.1	5.5
T4	5.4	4.8	4.0	4.7
T5	0.4	0.6	0.6	0.5
3. Gazdaság				
T1	1.8	2.4	3.1	2.4
T2	3.3	3.7	2.9	3.3
T3	2.1	3.0	2.4	2.5
T4	2.9	1.1	4.0	2.7
T5	1.4	0.7	1.6	1.2
4. Gazdaság				
T1	5.1	5.4	5.6	5.4
T2	3.7	4.6	5.0	4.4
T3	3.6	4.2	4.5	4.1
T4	5.1	4.7	5.2	5.0
T5	2.8	4.2	5.8	4.3

T₁ = Butaklór + proponil, T₂ = Butralin + proponil, T₃ = Proponil, T₄ = kézi gyomirtás és T₅ = kezeletlen (gyomos) kontroll

A lépésenkénti eljárások a kezelés és a gazdaság közötti interakció vizsgálatára következők:

1. Lépés. Végezzük el a kombinált varianciaanalízist minden gazdaság adataiból, követve a kombinált adatelemzés eljárásait. Példánkra, az eredményt a 15.2. táblázat mutatja. A gazdaság és a kezelés közötti interakció hatás erősen szignifikáns, jelezve, hogy a gyomirtási kezelések relatív teljesítménye változott a gazdaságok között.

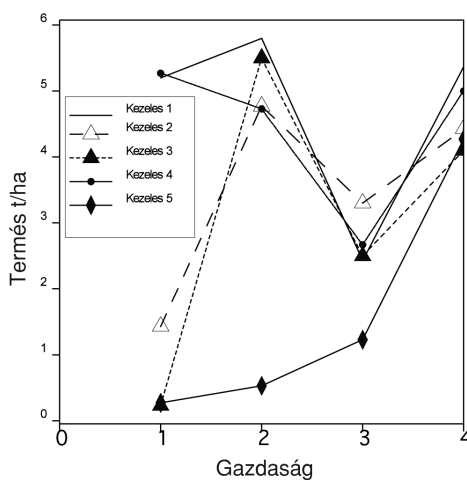
2. Lépés. Ha a gazdaság × kezelés interakció az 1. lépésben nem szignifikáns, jelezve, hogy a kezelés teljesítménye következetes volt minden vizsgált gazdaságban, végezzünk el egy kezelésközvetlen (minden gazdaság átlaga) összehasonlítást, követve a standard eljárást. Máskülönbén vizsgáljuk meg az interakció természetét az interakció SS megfelelő felbontásával.

15.2. táblázat. Kombinált varianciaanalízis a gazdaságok figyelembevételével, adatok a 15.1. táblázatból

Variáció forrása	Szabadságfok	Eltérésnégyzet összeg (SS)	Közepes négyzetes eltérés (MS)	Számított F-érték
Gazdaság (F)	3	60.8298	20.2766	25.05**
Ismétlés gazdaságon belül	8	6.476	0.8095	
Kezelés (T)	4	73.7773	18.4443	35.07**
F × T	12	69.8161	5.818	11.06**
Összevont hiba	32	16.8306	0.526	
Összes	59	227.7298		

CV % = 21.0, ** = szignifikáns 1%-os szinten

A) Ábrázoljuk minden kezelés átlagos értékét (Y tengely) a gazdaságok (X tengely) szerint, mint az 15.1. ábra mutatja. Példáinkra az eredmény mutatja:



15.1. ábra. A gazdaság × kezelés interakció vizsgálata, minden kezelés termését gazdaságok szerint ábrázolva

(i) A kezelések három eltérő csoportját tudjuk elkülöníteni relatív teljesítményük alapján, a vizsgált gazdaságokban. Az egyik csoport tartalmazza a T₁ és T₄, egy másik csoport a T₂ és T₃, és az utolsó csoport a T₅ (kontroll) kezelést. A fő különbség az első két csoport között az 1-es gazdaság teljesítménye: T₁ és T₄ nagy termést, míg T₂ és T₃ kis termést adott. Vagyis, T₁ és T₄ nagyon hatékony volt a gyomok szabályozásában az 1-es gazdaságban, míg T₂ és T₃ nem volt hatékony.

(ii) A 4-es gazdaságban, T₅ (kezeletlen kontroll) termése nem különbözött a többi kezelés termésétől, jelezve, hogy a gyomok nem jelentettek súlyos gondot ebben a gazdaságban.

(iii) Mindegyik gyomirtási eljárás (T₁, T₂, T₃ és T₄) hatékony volt a 2-es és 3-as gazdaságban, azonban a reakció sokkal nagyobb volt a 2-es és 3-as gazdaságban, mint a 3-asban.

15.3. táblázat. Öt gyomirtási kezelés termésének összehasonlítása négy gazdaságban

Kezelések	Átlagtermés t ha ⁻¹				
	1. Gazdaság	2. Gazdaság	3. Gazdaság	4. Gazdaság	Átlag
T1	5.20 a	5.80 a	2.43 ab	5.37 a	4.70
T2	1.43 b	4.77 a	3.30 a	4.43 a	3.48
T3	0.23 b	5.50 a	2.50 ab	4.10 a	3.08
T4	5.27 a	4.73 a	2.67 a	5.00 a	4.42
T5 (kontroll)	0.27 b	0.53 b	1.23 b	4.27 a	1.58
Átlag	2.48	4.27	2.43	4.63	3.45

Gazdaságonként három ismétlés átlaga. Mindegyik gazdaságban, azok az átlagok, melyeket azonos betű követ, szignifikánsan nem különböznek 5%-os szinten.

B) Bizonyítsuk az A-ban tett vizuális megfigyeléseket egy megfelelő átlag összehasonlítással vagy a kezelés \times gazdaság interakció SS felbontásán keresztül. Példánkban, (ii) megfigyelést igazolja a pár összehasonlítás nem szignifikáns eredménye az öt kezelésátlag között a 4-es gazdaságban (15.3. táblázat). Az (i) és (iii) megfigyeléseket bizonyítja az interakció SS felbontásának eredménye (15.4. táblázat).

15.4. táblázat. A kezelés \times gazdaság interakció SS (15.1. táblázat) felbontása az 15.1. ábra alapján tett következtetések igazolására

Variáció forrása	Szabadságfok	SS	MS	Számított F-érték
(T ₅ vs. többi) \times Gazdaság	3	23.6535	7.8845	14.99**
(T ₅ vs. többi) \times (2.Gazd. \times 3. Gazd.)	(1)	12.0968	12.0968	23.00**
(T ₅ vs. többi) \times többi Gazd	(2)	11.5567	5.7784	10.99**
(T ₁ , T ₄ vs. T ₂ , T ₃) \times Gazd.	3	41.5142	13.8381	26.31**
(T ₁ vs. T ₄) \times Gazd.	3	1.5150	0.5050	< 1
(T ₂ vs. T ₃) \times Gazd.	3	3.1333	1.0444	1.99 ^{ns}

** = szignifikáns 1%-os szinten, ^{ns} = nem szignifikáns.

3. Lépés. Azonosítsuk a környezeti faktorokat, amelyek nagy valószínűséggel a kezelés és gazdaság közötti interakciót okozzák, melyet a 2. lépésben kaptunk. Példánkban, feltételezhetjük, hogy a gyomok a legfőbb okozói. Tehát, először megvizsgáltuk az összes gyomtömeget a gyomirtási kezelések teljesítményével összefüggésben gazdaságonként (15.5. táblázat).

A gyomtömeget a kelés után a 62. napon mérték és átlagolták a három ismétlést gazdaságonként; a termés adatok az 15.1. táblázatban vannak.

Az adatok mutatják:

- A 4-es gazdaságban a gyomok tömege nagyon kevés volt, mint a kezeletlen parcellák adatai mutatják. Ez valószínűleg az oka, miért volt a kezeletlen parcellák termése ebben a gazdaságban nagy és miért nem volt szignifikáns reakció egyetlen gyomirtási eljárásra sem (15.3. táblázat).
- Az 1. gazdaságban a gyomtömeg a kezeletlen parcellában a legnagyobb volt (majdnem kétszer annyit volt, mint a 2. és 3. gazdaságban). Ez valószínűleg az

oka a nagy variációnak a kezelt parcellák között, azaz T₂ és T₃ termése sokkal kisebb volt, mint a T₁ és T₄ kezeléseké (15.3. táblázat). Nyilvánvalóan, a fűfélék és a sásfélék voltak a domináns gyomtípusok ebben a gazdaságban; és míg T₁ és T₄ hatékony volt e gyomok szabályozásában, T₂ és T₃ nem volt hatékony (15.6. táblázat).

15.5. táblázat. Átlagos összes gyomtömeg a gyomirtási kísérletben négy gazdaságban

Kezelések	Átlagos összes gyomtömeg, g m ²			
	1. Gazdaság	2. Gazdaság	3. Gazdaság	4. Gazdaság
T1	28.0	44.4	176.8	7.2
T2	290.6	93.8	197.4	12.2
T3	528.6	19.2	351.6	45.2
T4	16.4	39.2	32.4	5.8
T5 (kontroll)	828.2	421.8	430.4	49.6
Átlag	338.4	123.7	237.7	24.0

15.6. táblázat. A gyomok átlagos tömege, típusok (kétszikűek, fűfélék és sásfélék) szerint osztályozva egy gyomirtási kísérletben, négy gazdaságban

Kezelések	Gyomok tömege g m ²			
	1. Gazdaság	2. Gazdaság	3. Gazdaság	4. Gazdaság
Kétszikűek				
T1	7.8	16.8	0.2	0.2
T2	9.6	26.8	4.2	1.2
T3	4.0	5.8	9.4	0.0
T4	0.4	0.2	3.4	0.4
T5	0.0	13.8	32.2	2.0
Fűfélék				
T1	20.2	20.8	173.0	6.4
T2	224.0	27.2	164.0	8.0
T3	461.2	10.8	328.0	28.4
T4	6.2	30.0	25.6	4.8
T5	696.0	390.6	387.0	40.8
Sásfélék				
T1	0.0	6.8	3.6	0.6
T2	57.0	39.8	29.2	0.6
T3	63.4	2.6	14.2	8.0
T4	9.8	8.8	3.4	0.4
T5	132.2	17.6	11.2	6.6

- A viszonylag kisebb termés és a kisebb reakció a gyomirtási kezelésekre a 3. gazdaságban, viszonyítva a 2. gazdasághoz, nem kielégítően magyarázható meg az 15.5. és 15.6. táblázat adataival. Először, az összes gyomtömeg a kezeletlen parcellákban hasonló volt ebben a két gazdaságban (15.5. táblázat) és jöllehet a 3. gazdaságban több kétszikű gyom volt a kezeletlen parcellákban, mint a 2. gaz-

daságban, e gyomtípus irtása a 3. gazdaságban még jobb volt, mint a 2. gazdaságban (15.6. táblázat). Másrésztől, jóllehet a fűfélék összes tömege ugyanaz volt a 2. és 3. gazdaságban, a három vegyszeres kezelés (T_1 , T_2 és T_3) hatékonyabb volt a gyomok irtásában a 2. gazdaságban, mint a 3. gazdaságban. Azonban, a hatékonyság ellenére a fűfélék mennyiségének csökkentésében a kézi gyomirtással (T_4) a 3. gazdaságban, a termés nem volt szignifikánsan különböző a vegyszeres kezelésektől. Tehát úgy tűnik, más termés-csökkentő faktorok voltak jelen a 3. gazdaságban a gyomokon túlmenően. Ezeket a faktorokat tovább kell vizsgálni.

15.1.5. Varianciaanalízis a GenStat programmal

Az adatok bevitelekor a Gazdaság, Ismétlés és Kezelés képezi a faktorokat, Termés a változó. A gazdaságnak 4, az ismétlésnek 3, a kezelésnek 5 szintje van. Az adattáblázat szerkezete megegyezik a korábban leírtakéval.

A varianciaanalízishez a kiválasztott menük sorrendje: Stats – Analysis of Variance – General Analysis of Variance. Y-variate: Termés, Treatment Structure: Kezelés × Gazdaság, Block Structure: Gazdaság/Ismétlés.

A varianciaanalízis outputja, amely megegyezik a 15.2. táblázat eredményével, a következő:

Analysis of variance					
Variate: Termes					
Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Gazdasag.Ismetles stratum					
Gazdasag	3	60.8298	20.2766	25.05	<.001
Residual	8	6.4760	0.8095	1.54	
Gazdasag.Ismetles.*Units* stratum					
Kezeles	4	73.7773	18.4443	35.07	<.001
Gazdasag.Kezeles	12	69.8160	5.8180	11.06	<.001
Residual	32	16.8307	0.5260		
Total	59	227.7298			
Tables of means					
Variate: Termes					
Grand mean 3.452					
Gazdasag	1	2	3	4	
	2.480	4.267	2.427	4.633	
Kezeles	1	2	3	4	5
	4.700	3.483	3.083	4.417	1.575

Gazdasag	Kezeles	1	2	3	4	5
	1	5.200	1.433	0.233	5.267	0.267
	2	5.800	4.767	5.500	4.733	0.533
	3	2.433	3.300	2.500	2.667	1.233
	4	5.367	4.433	4.100	5.000	4.267

Standard errors of differences of means

Table	Gazdasag	Kezeles	Gazdasag Kezeles
rep.	15	12	3
s.e.d.	0.3285	0.2961	0.6233
d.f.	8	32	38.54
Except when comparing means with the same level(s) of Gazdasag			0.5921
d.f.			32

Least significant differences of means (5% level)

Table	Gazdasag	Kezeles	Gazdasag Kezeles
rep.	15	12	3
l.s.d.	0.7576	0.6031	1.2611
d.f.	8	32	38.54
Except when comparing means with the same level(s) of Gazdasag			1.2062
d.f.			32

A varianciaanalízis 2. szakaszában a gazdaság × kezelés interakció SS-t felbontottuk a 15.4. táblázatnak megfelelően, a kontrasztok (először a kezelés, ezt követően a gazdaság kontrasztok) segítségével. Az analízis outputja az alábbi:

Kezelések szerinti kontraszt:

Contrast 1	-1	-1	-1	-1	4
Contrast 2	1	-1	-1	1	0
Contrast 3	1	0	0	-1	0
Contrast 4	0	1	-1	0	0

Analysis of variance

Variate: Termes

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Gazdasag.lsmetles stratum					
Gazdasag	3	60.8298	20.2766	25.05	<.001
Residual	8	6.4760	0.8095	1.54	

Gazdasag.Ismetles.*Units* stratum					
Kezeles	4	73.7773	18.4443	35.07	<.001
Contrast 1	1	52.8282	52.8282	100.44	<.001
Contrast 2	1	19.5075	19.5075	37.09	<.001
Contrast 3	1	0.4817	0.4817	0.92	0.346
Contrast 4	1	0.9600	0.9600	1.83	0.186
Gazdasag.Kezeles	12	69.8160	5.8180	11.06	<.001
Gazdasag.Contrast 1	3	23.6535	7.8845	14.99	<.001
Gazdasag.Contrast 2	3	41.5142	13.8381	26.31	<.001
Gazdasag.Contrast 3	3	1.5150	0.5050	0.96	0.423
Gazdasag.Contrast 4	3	3.1333	1.0444	1.99	0.136
Residual	32	16.8307	0.5260		
Total	59	227.7298			

Tables of means

Variate: Termes

Grand mean 3.452

Gazdasag		1	2	3	4		
		2.480	4.267	2.427	4.633		
Kezeles		1	2	3	4	5	
		4.700	3.483	3.083	4.417	1.575	
Gazdasag	Kezeles	1	2	3	4	5	
1		5.200	1.433	0.233	5.267	0.267	
2		5.800	4.767	5.500	4.733	0.533	
3		2.433	3.300	2.500	2.667	1.233	
4		5.367	4.433	4.100	5.000	4.267	

Standard errors of differences of means

Table	Gazdasag	Kezeles	Gazdasag Kezeles
rep.	15	12	3
s.e.d.	0.3285	0.2961	0.6233
d.f.	8	32	38.54

Except when comparing means with the same level(s) of

Gazdasag	0.5921
d.f.	32

Least significant differences of means (5% level)

Table	Gazdasag	Kezeles	Gazdasag Kezeles
rep.	15	12	3
l.s.d.	0.7576	0.6031	1.2611
d.f.	8	32	38.54

Except when comparing means with the same level(s) of

Gazdasag	1.2062
d.f.	32

Gazdaságok szerinti kontraszt:

Contrast	1:	0	1	-1	0
----------	----	---	---	----	---

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Contrast 1.Contrast 1	1	12.0967	12.0967	23.00	<.001

15.2. Technológia összehasonlítási kísérletek rendszere

A technológia összehasonlítási kísérletek célja összehasonlítani a termelő technológiájának és az új technológiának a teljesítményét a gazdaság tábláján. A teljesítmény elsődleges alapja a biológiai hozam (termés) és a jövedelmezőség. Tehát, szemben a technológia-fejlesztési kísérlettel, a technológia-összehasonlítási kísérlet megkülönböztető sajátosságai a következők:

- (i). A gazdaság táblája a vizsgálat helye.
- (ii). A termelő technológiája az összehasonlítás alapja.
- (iii). Minden agrotechnikai eljárást, kivéve a teszt faktorokat (az új technológia komponenseit, melyek különböznek a termelő gyakorlatától) a termelő szintjén tartjuk fenn.

A teszt gazdaságoknak a technológia összehasonlítására reprezentálni kell a gazdaságokat a megcélzott területen (ökológiai régióban), ahol az új technológia adaptációja várhatóan javasolható. A megcélzott területet azonosítja egy vagy több specifikus környezeti komponens (fizikai, biológiai, társadalmi és környezeti), amelyek kritikusnak tekinthetők az új technológia teljesítményéhez és végleges adaptációjához. Annak biztosítására, hogy a gazdaságok reprezentálják a vizsgált területet, megfelelő számú teszt gazdaságot kell kiválasztani. A mintaméret növekedhet a gazdaságok és a mintaterület közötti variabilitás növekedésével. Diverzitás a növényi sorrendben jó indikátora a gazdaságok közötti variabilitásnak. Általános szabály, hogy a teszt gazdaságok számának kb. tízszeresének kellene lenni a célterületen domináns növényi sorrend számának.

A leggyakrabban használt mintavételi terv a teszt gazdaságok kiválasztására a technológia összehasonlítási kísérletben a random mintavétel. A gazdaságok az elsődleges mintavételi egységek és a táblák a gazdaságon belül a másodlagos mintavételi egységek. A növényi sorrend és a gazdaság mérete a leggyakrabban használt osztályozási kritérium.

15.2.1. Kísérleti elrendezés

A technológia összehasonlítási kísérlet k teszt faktorialis úgy tekinthető, mint egy 2^k faktorialis kísérlet, amelyben minden teszt faktor két szintje reprezentálja az előírt szintet az új technológiában és a termelői gyakorlat (technológia) szintjét. Vegyünk egy példát, ahol a meglévő termelői gyakorlat a rizs termesztése a nedves évszakban, melyet kukorica követ a száraz évszakban, és az új technológia ugyanazt a növényi sorrendet követi, azonban különböző fajtákkal, műtrágya dózissal és károsítók elleni védekezéssel rizsben, és különböző károsítók elleni védekezéssel és talajműveléssel kukoricában. Az új technológia különbözik a meglévő termelői gyakorlatától hat technológiai komponensben – négy a rizsben és kettő a kukoricában, ezáltal hat teszt faktor van (azaz $k = 6$). Teszt faktoronként két szinttel a technológia összehasonlítási kísérlet 2^6 faktorialis kísérlet.

Amikor k értéke nagy, mint általában a technológia-összehasonlítási kísérletekben, 2^k faktorialis kezelések teljes sorozata túlzottan nagy lesz. A kísérlet méretének csökkentésére általában a kezelések három különböző sorozatát tesztelik. Ezek a következők:

1. **X sorozat**, két kezelést tartalmaz: az új technológia (a teszt faktorok mindegyike az új technológia szintjén) és a termelői gyakorlat (a teszt faktorok mindegyike a termelő szintjén).

2. **Y sorozat**, $(k+2)$ kezelést tartalmaz: az X sorozat két faktora, plusz k közbenső kezelés, melynek mindegyike képvisel egy kezeléskombinációt, melyben egy kivételével mindegyik teszt faktor az új technológia szintjén van. [Ez megfelel Győrffy martonvásári Komplex I. kísérletének (10.2. fejezet).]

3. **Z sorozat**, tartalmazza vagy a 2^k komplett (teljes) kezeléskombinációt vagy egy megfelelő frakcionális sorozatát. Példaként vegyünk egy búza technológia-összehasonlítási kísérletet, melyben az új technológia egy fejlesztést reprezentál a termelői gyakorlathoz képest négy komponensben: műtrágya (F), fajta (V), gyomirtás (W) és kártevők elleni védekezés (I). Tehát négy teszt faktorunk van és a technológia-összehasonlítási kísérlet egy 2^4 faktoriális kísérlet. [Ugyanilyen típusú Győrffy martonvásári Komplex II. kísérlete, melyet a 12.3. fejezetben analizáltunk].

A három kezeléssorozat (X, Y és Z) a technológia-összehasonlítási kísérlettel összefüggésben az 15.7. táblázat mutatja. Mivel a kezelések száma legkisebb az X sorozatban és legnagyobb a Z sorozatban, a gazdaságok száma a kezelések X sorozatának tesztelésére (n_x) rendszerint a legnagyobb, melyet n_y követ, azaz a gazdaságok száma a kezelések Y sorozatának tesztelésére, és végül n_z , a gazdaságok száma a kezelések Z sorozatának tesztelésére. Az n mérete függ a teszt faktorok között várható interakció fontosságától. Ha az interakció várhatóan nagy és nagyszámú teszt faktort foglal magában, akkor n_z is nagy lesz. Az n_x : n_y : n_z aránya a technológia-összehasonlítási kísérletekben általában 3:1:1.

15.7. táblázat. A három kezeléssorozat (X, Y és Z), összefüggésben a technológia-összehasonlítási kísérlettel rizsben, mely magában foglal négy faktort: műtrágya (F), fajta (V), gyomirtás (W) és kártevők elleni védekezés (I) (Gomez és Gomez, 1984)

Kezelés szám	Faktor szint ^a				Kezelés ^b			
	F	V	W	I	X sorozat	Y sorozat	Z sorozat	
							Teljes faktoriális	Frakcionális faktoriális ^c
1	n	n	n	n	*	*	*	*
2	f	n	n	n	–	*	*	–
3	n	f	n	n	–	*	*	–
4	f	f	n	n	–	–	*	*
5	n	n	f	n	–	*	*	–
6	f	n	f	n	–	–	*	*
7	n	f	f	n	–	–	*	*
8	f	f	f	n	–	–	*	–
9	n	n	n	f	–	*	*	–
10	f	n	n	f	–	–	*	*
11	n	f	n	f	–	–	*	*
12	f	f	n	f	–	–	*	–
13	n	n	f	f	–	–	*	*
14	f	n	f	f	–	–	*	–
15	n	f	f	f	–	–	*	–
16	f	f	f	f	*	*	*	*

^an = az új technológia szintje és f = termelői szint; ^b– = nem vizsgálták és * vizsgálták;

^c2⁴ frakcionális faktoriális elrendezés 1/2-e.

Kísérleti technika

A kísérleti technikának a technológia-összehasonlítási kísérletben le kell győzni számos potenciális problémát:

- Viszonylag nagy variáció a gazdaságok között és gazdaságon belül.
- Diverzitás a termelők gyakorlatában.
- A termelői gyakorlat megvalósítása kezelésenként.
- A termelői gyakorlat, mint alapul szolgáló teszt környezet.

15.2.2. Adatgyűjtés

Az adatgyűjtést a technológia-összehasonlítási kísérletekhez elsődlegesen úgy tervezzük, hogy összehasonlítsuk az új technológiát és a termelő gyakorlatát (technológiáját) és meghatározzuk az új technológia befogadásának lehetőségét. Néhány fontos adat, amit be kell gyűjteni a technológia összehasonlítási kísérletekben:

- Táblánkénti termés.
- A gazdaság fizikai és biológiai környezete, mint amilyenek a talajtulajdonságok, a termesztett növények története, az alkalmazott agrotechnika és a feljegyzett klimatikus és időjárási adatok.
- Társadalmi és ökonómiai adatok, amelyek lényegesek az új technológia adaptálhatóságának értékelésében, mint amilyen a gazdaság mérete, bevételek és kiadások, munkaerő foglalkoztatottság, a mezőgazdasági inputok költsége, a gazdaság termékeinek ára.
- Jelenlegi termelői gyakorlat és termelékenysége.
- Adatok, amelyek segítenek értelmezni mindegyik teszt faktor teljesítményét. Például, ha a gyomirtás a teszt faktor, akkor adatokat kell gyűjteni a gyomosodásról.

Az adatelemzés két típusát a technológia-összehasonlítási kísérletben a terméshiány elemzés és a költségelemzés képezi. A következőkben a terméshiány elemzést mutatjuk be.

15.2.3. Terméshiány elemzés

A terméshiány elemzés méri a különbséget a termésben (vagy a termelékenység más mutatójában) az új technológia és a termelő gyakorlata (technológiája) között, és felosztja ezt a különbséget komponensekre, amelyek reprezentálják az egyes faktorok hozzájárulását. Mivel legtöbb esetben csak a termés hasonlítható össze azonos növénynél vagy növényeknél, a terméshiány elemzés általában kultúrnövényenkénti alapon végezhető. Amikor az új technológia magában foglal egynél több kultúrnövényt, a terméshiány analízis csak arra a növényre végezhető el, amely mind az új technológiában és mind a termelői gyakorlatban közös.

Bemutatjuk a technológia-összehasonlítási kísérlet számítási eljárását egy rizs kísérlet terméshianyával (Gomez és Gomez, 1984), amely magában foglal három teszt-faktort: műtrágya (F), kártevők elleni védekezés (I) és gyomirtás (W). A kísérlet felölel 19 teszt gazdaságot, az X sorozat kezeléseit 12 gazdaságban tesztelték (15.8. táblázat), az Y kezeléssorozatot 3 gazdaságban (15.9. táblázat) és a Z sorozat kezeléseit 4 gazda-

ságban (azaz 2^3 komplett faktoriális kezelések) (15.10. táblázat). Mindegyik gazdaságban a kezeléseket két ismétlésben vizsgálták.

15.8. táblázat. Az új technológia és a termelői gyakorlat szemtermése, melyet 12 gazdaságban teszteltek, két ismétlésben gazdaságonként a technológia-összehasonlítási kísérletben rizsben

Gazdaságok száma	Szemtermés t ha ⁻¹					
	Új technológia			Termelői gyakorlat		
	I. ismétlés	II. ismétlés	Átlag	I. ismétlés	II. ismétlés	Átlag
1	6.60	6.64	6.62	3.65	4.94	4.25
2	6.50	5.96	6.23	3.88	4.62	4.25
3	7.61	6.78	7.20	6.02	5.30	5.66
4	7.20	7.02	7.11	5.26	5.15	5.20
5	3.68	3.62	3.65	3.13	2.73	2.93
6	4.18	4.47	4.32	3.00	3.05	3.02
7	6.54	5.98	6.26	5.36	5.10	5.23
8	7.26	6.20	6.73	5.76	4.69	5.22
9	7.30	6.28	6.79	5.35	5.17	5.26
10	7.60	7.60	7.60	5.96	5.74	5.85
11	6.71	6.29	6.50	5.65	5.89	5.77
12	7.14	6.16	6.65	5.02	4.56	4.79

15.9. táblázat. A technológia-összehasonlítási kísérlet Y kezeléssorozatának rizstermése, három faktor (F, I, W) bevonásával, melyet három gazdaságban vizsgáltak, gazdaságonként két ismétlésben (I. és II.)

Kezelés				Szemtermés t ha ⁻¹								
No.	Szintje ^a			13. gazdaság			14. gazdaság			15. gazdaság		
	F	I	W	I.	II.	Átl.	I.	II.	Átl.	I.	II.	Átl.
1	n	n	n	6.83	7.51	7.17	6.76	7.02	6.89	6.93	6.91	6.92
2	f	n	n	6.26	5.92	6.09	6.74	6.23	6.48	6.06	6.06	6.06
3	n	f	n	6.16	6.29	6.22	5.92	6.14	6.03	6.21	6.72	6.46
4	n	n	f	6.97	7.15	7.06	7.57	6.73	7.15	6.50	7.27	6.88
5	f	f	f	5.36	4.84	5.10	5.05	4.37	4.71	4.53	4.99	4.76

^aMűtrágya (F), kártevők elleni védekezés (I), gyomirtás (W); n = új technológia és f = termelői technológia.

15.10. táblázat. A technológia-összehasonlítási kísérlet Z kezelés sorozatának szemtermése rizsben, három tesztfaktor (F, I és W) bevonásával, melyet négy gazdaságban vizsgáltak, gazdaságonként két ismétlésben

Kezelés				Szemtermés t ha ⁻¹											
No.	Szint ^a			16. Gazdaság			17. Gazdaság			18. Gazdaság			19. Gazdaság		
	F	I	W	I.	II.	Átl.	I.	II.	Átl.	I.	II.	Átl.	I.	II.	Átl.
1	n	n	n	8.76	6.90	7.83	7.21	7.05	7.13	6.98	6.36	6.67	6.28	5.98	6.13
2	n	n	f	4.98	6.60	5.79	6.20	6.62	6.41	6.43	6.23	6.33	6.14	6.00	6.07
3	n	f	n	6.50	6.50	6.50	5.74	5.76	5.75	6.00	6.16	6.08	4.96	4.44	4.70
4	n	f	f	7.29	6.31	6.80	4.74	5.54	5.14	5.87	6.03	5.95	4.96	4.21	4.58
5	f	n	n	7.63	6.73	7.18	5.14	5.52	5.33	5.80	5.81	5.80	4.86	4.39	4.62
6	f	n	f	7.02	6.27	6.64	4.93	5.12	5.02	5.47	5.57	5.52	4.88	4.30	4.59
7	f	f	n	5.58	5.12	5.35	4.40	5.06	4.73	4.69	4.67	4.68	3.80	3.16	3.48
8	f	f	f	5.44	5.01	5.22	3.49	3.88	3.68	4.64	4.70	4.67	3.49	3.04	3.26

^aMűtrágya (F), kártevők elleni védekezés (I), gyomirtás (W); n = új technológia és f = termelői szint.

A terméshiány elemzés lépésenkénti eljárásai a következők:

1. Lépés. Mindegyik gazdaságra számítsuk ki a terméshiányt, mint az új technológia és a termelői gyakorlat termése közötti különbséget. Ezután számítsuk ki ezeknek a terméshiányoknak az átlagát az n számú gazdaságra:

$$G = \frac{\sum_{i=1}^n (\bar{P}_i - \bar{Q}_i)}{n}$$

ahol \bar{P}_i az új technológia átlagtermése és \bar{Q}_i a termelői gyakorlat (technológia) átlagtermése az i-edik gazdaságban.

Példánkra, az eredményeket a 19 teszt gazdaságra a 15.11. táblázat mutatja. A terméshiány értékek tartománya 0.72 t ha⁻¹-től (az 5. gazdaságban) 3.45 t ha⁻¹-ig (a 17. gazdaságban) változik és a terméshiány az összes teszt gazdaság átlagában 1.87 t ha⁻¹.

2. Lépés. Értékeljük az interakciót a teszt faktorok között úgy, hogy kiszámítjuk az adatok kombinált varianciaanalízisét az összes n_z gazdaság bevonásával (ahol a Z sorozat kezeléseit tesztelték), követve az RCB kísérlet kombinált adataira kidolgozott eljárást az évek figyelembevételével. Azaz, a gazdaságokat a technológia-összehasonlítási kísérletben random változónak tekintjük.

Ha nem áll rendelkezésre adat az n_z gazdaságból, mivel egyetlen gazdaság sem tesztelte sikeresen a Z sorozat kezeléseit vagy nincs n_z gazdaság, tekintsünk el ettől a lépéstől és folytassuk az I. analízis 3. lépésénél.

Példánkra, négy n_z gazdaság van (15.10. táblázat). A négy gazdaság kombinált varianciaanalízisének eredményét, az RCB elrendezés alapján a 15.12. táblázat mutatja. A kezelés SS megfelelő faktoriális felbontása fő hatásokra és a köztük levő interakciókra nem mutat szignifikáns interakció hatásokat a három teszt faktor között, ugyanakkor mindhárom faktor főhatása szignifikáns.

15.11. táblázat. A terméshiány kiszámítása, mint az új technológia és a termelői gyakorlat közötti különbség, a 15.8.–15.10. táblázatok adatai alapján

Gazdaság száma	Átlagtermés t ha ⁻¹		
	Új technológia	Termelői technológia	Terméshiány
1	6.62	4.25	2.37
2	6.23	4.25	1.98
3	7.20	5.66	1.54
4	7.11	5.20	1.91
5	3.65	2.93	0.72
6	4.32	3.02	1.30
7	6.26	5.23	1.03
8	6.73	5.22	1.51
9	6.79	5.26	1.53
10	7.60	5.85	1.75
11	6.50	5.77	0.73
12	6.65	4.79	1.86
13	7.17	5.10	2.07
14	6.89	4.71	2.18
15	6.92	4.76	2.16
16	7.83	5.22	2.61
17	7.13	3.68	3.45
18	6.67	4.67	2.00
19	6.13	3.26	2.87
Átlag	6.55	4.68	1.87

3. Lépés. Számítsuk ki mindenegyres teszt faktor hozzájárulását a terméshiányhoz a kétféle analízis egyikével: I. Analízis, amikor az interakció a teszt faktorok között nem szignifikáns, és II. analízis amikor szignifikáns.

Példánkra, az interakció a három teszt faktor között nem szignifikáns (2. lépés), és ezáltal az I. analízis megfelelő. Bemutatjuk azonban mindkét analízis számítási lépéseit, felhasználva példánk adatait.

15.12. táblázat. A 15.10. táblázat adatainak kombinált varianciaanalízise négy gazdaságra; mindegyik kísérlet magában foglalt 2^3 teljes faktoriális kezeléskombinációt

Variáció forrása	Szabadságfok	Eltérés négyzetösszeg	Közepes négyzetes eltérés	Számított F-érték
Gazdaság	$n_z - 1 = 3$	24.85006	8.28335	
Ism. gazdaságon belül	$n_z(r-1) = 4$	2.27329	0.56832	
Kezelés	$2^k - 1 = 7$	41.25957	5.89422	15.35**
F	(1)	20.38522	20.38522	53.10**
I	(1)	16.99501	16.99501	44.27**
W	(1)	2.45706	2.45706	6.40*
F × I	(1)	0.48302	0.48302	1.26 ^{ns}
F × W	(1)	0.08410	0.08410	< 1
I × W	(1)	0.35106	0.35106	< 1
F × I × W	(1)	0.50410	0.50410	1.31 ^{ns}
Kezelés × Gazdaság	$(2^k-1)(n_z-1)=21$	8.06227	0.38392	
Összevont hiba	$n_z(2^k-1)(r-1)=28$	4.23951	0.15141	
Összes	$(n_z)(r)(2^k)-1=63$	80.68470		

** = szignifikáns 1%-os szinten, * = szignifikáns 5%-os szinten, ^{ns} = nem szignifikáns.

15.2.3.1. Átlagos hozzájárulás

(i) Mindegyik n_y gazdaságra (mindegyik teszteli az Y sorozat kezeléseket) számítsuk ki mindegyik teszt faktor hozzájárulását, pl. A faktorét, mint:

$$C_A = \bar{P} - \bar{T}_A$$

ahol \bar{P} az új technológia átlagtermése és \bar{T}_A annak a kezelésnek az átlagtermése, amelyben csak az A faktor szintje van a termelői gyakorlat szintjén (azaz, minden más teszt faktor az új technológia szintjén).

Példánkra, felhasználva a 15.9. táblázat adatait, például az F faktor hozzájárulását a 13. gazdaságban így számítjuk ki:

$$\begin{aligned} C_F &= \bar{P} - \bar{T}_A \\ &= 7.17 - 6.09 = 1.08 \text{ t ha}^{-1} \end{aligned}$$

és, a W faktor hozzájárulását a 14. gazdaságban így számítjuk ki:

$$\begin{aligned} C_W &= \bar{P} - \bar{T}_W \\ &= 6.89 - 7.15 = -0.26 \text{ t ha}^{-1} \end{aligned}$$

Az eredményt mindhárom teszt faktorra és mindhárom teszt gazdaságra (13, 14 és 15) a 15.13. táblázat első három sora mutatja.

(ii) Mindegyik n_z gazdaságra (mindegyik teszteli a Z sorozat kezeléseket) számítsuk ki mindenegyestesztfaktor hozzájárulását, mondjuk az A faktorét, mint a különbség minden kezelés átlagtermése az A faktorial az új technológia szintjén \bar{A}_n és minden kezelés átlagtermése az A faktorial a termelői gyakorlat szintjén \bar{A}_f .

Példánkra, felhasználva a 15.10. táblázat adatait, például az F faktor hozzájárulását a 16. gazdaságban így számítjuk ki:

$$C_F = \bar{F}_n - \bar{F}_f$$

$$= \frac{7.83 + 5.79 + 6.50 + 6.80}{4} - \frac{7.18 + 6.64 + 5.35 + 5.22}{4} = 6.73 - 6.10 = 0.63 \text{ t ha}^{-1}$$

és a W faktor hozzájárulása a 17. gazdaságban a következő:

$$C_W = \bar{W}_n - \bar{W}_f$$

$$= \frac{7.13 + 5.75 + 5.33 + 4.73}{4} - \frac{6.41 + 5.14 + 5.02 + 3.68}{4} = 5.74 - 5.06 = 0.68 \text{ t ha}^{-1}$$

Az eredményeket mindhárom tesztfaktorra és mind a négy tesztgazdaságra (16, 17, 18 és 19 gazdaság) a 15.13. táblázat 4–7. sora mutatja.

15.13. táblázat. Mindhárom faktor egyenkénti átlagos hozzájárulása, melyet a technológia-összehasonlítási kísérletben teszteltek, a 15.9. és 15.10. táblázat adataiból számolva

Gazdaság száma	Átlagos hozzájárulás t ha ⁻¹		
	F	I	W
13	1.08	0.95	0.11
14	0.41	0.86	-0.26
15	0.86	0.46	0.04
16	0.63	0.89	0.60
17	1.42	1.15	0.68
18	1.09	0.74	0.19
19	1.38	1.35	0.11
Átlag	0.98	0.91	0.21

Műtrágya (F), Kártevők elleni védekezés (I) és Gyomirtás (W)

(iii) Mindegyik teszt faktorra számítsuk ki az átlagot az összes gazdaság (n_y és n_z) alapján, minden hozzájárulásra, melyet az (i) és (ii) pontban kiszámoltunk. Példánkra az átlagos hozzájárulást mindegyik tesztfaktorra kiszámítjuk az $n_y + n_z = 3 + 4 = 7$ gazdaságra. Az eredményeket a 15.13. táblázat utolsó sora mutatja. Az átlagos hozzájárulás műtrágyára 0.98 t ha^{-1} , a kártevők elleni védekezésre 0.91 t ha^{-1} és a gyomirtásra 0.21 t ha^{-1} .

(iv) Számítsuk ki mindenegyestesztfaktorra a kiigazított (korrigált) átlagos hozzájárulást, mint a kiigazítási faktor (g) és a kiszámított átlagos hozzájárulás szorzata.

Kiigazított átlagos hozzájárulás = (g) (átlagos hozzájárulás)

$$g = 1 - \frac{(S - G)}{S}$$

ahol S az átlagos hozzájárulás összege mindegyik tesztfaktorra és G az átlagos termés-hiány, melyet az 1. lépésben számoltunk ki.

Jegyezzük meg, hogy ha $S=G$, akkor kiigazítás nem szükséges. A különbség S és G között kicsi, amikor nincs interakció a teszt faktorok között, és nagy máskülönben. Abban az esetben, amikor nincs adat az n_z gazdaságból és ezáltal nem becsülhető az interakció (2. lépés), a különbség, $D = S - G$, amely nagyobb, mint az átlagos terméskülönbség 20%-a, tekinthető az interakció megléte indikátorának. Tehát, ha D nagyobb, mint $(0.2)(G)$, az I. analízis elhagyható és a II. analízist alkalmazzuk helyette.

Példánkra, a D különbséget kiszámítjuk, mint:

$$\begin{aligned} D &= S - G \\ &= (0.98 + 0.91 + 0.21) - 1.87 \\ &= 0.23 \text{ t ha}^{-1} \\ g &= 1 - \frac{0.23}{2.1} = 0.89048 \end{aligned}$$

D értéke kisebb, mint az átlagos termés hiány 20%-a. Ez egybevág azzal 2. lépésben kapott eredménnyel, hogy nincs szignifikáns interakció a tesztfaktorok között. Ezáltal, mindegyik tesztfaktor kiigazított átlagos hozzájárulását a következőképpen számítjuk ki:

A műtrágyázás kiigazított átlagos hozzájárulása	= $(0.89048)(0.98) = 0.87$
A kártevők elleni védekezés kiigazított átlagos hozzájárulása	= $(0.89048)(0.91) = 0.81$
A gyomirtás kiigazított átlagos hozzájárulása	= $(0.89048)(0.21) = 0.19$

Figyeljük meg, hogy a kiigazított átlagos hozzájárulások összege egyenlő az átlagos termés hiánnyal.

15.2.3.2. Egyedi és együttes hozzájárulások

Amikor a tesztfaktorok között egy vagy több interakció szignifikáns (2. lépés), számítsuk ki minden egyes tesztfaktor *egyedi hozzájárulását* az n_z gazdaság adatai alapján, amely tesztelte a komplett faktoriális kezeléseket és két vagy több tesztfaktor minden egyes kombinációjának *együttes hozzájárulását* az n_y gazdaságok és az n_z gazdaságok adatai alapján.

(i) Egy tesztfaktor, mondjuk az A faktor egyedi hozzájárulása, méri a hozzáadott termést a termelői gyakorlaton felül, melyet az A faktor új technológiai szinten történő

alkalmazásából kaptunk, miközben az összes többi teszt faktort a termelői szinten tartunk. Tehát ezt úgy számítjuk ki, mint a termelői gyakorlat termése és annak a kezelésnek a termése közötti különbség, amelyben csak az A faktor van az új technológia szinten (azaz minden más teszt faktor a termelői gyakorlat szintjén). Például, vegyük a Z sorozatot a komplett faktoriális kezelésekkel 15.7. táblázatban. Itt az F faktor egyedi hozzájárulását úgy számítjuk ki, mint:

$$H_F = \bar{Y}_{15} - \bar{Y}_{16}$$

\bar{Y}_{15} és \bar{Y}_{16} a 15. és 16. kezelés átlagtermése.

Példánkra, felhasználva a 15.10. táblázat adatait, az F faktor egyedi hozzájárulása a 16. gazdaságban:

$$H_F = 6.80 - 5.22 = 1.58 \text{ t ha}^{-1}$$

(ii) Az együttes hozzájárulás méri a hozzáadott termést, amely felülmúlja a termelői gyakorlatot, és megkapjuk két vagy több teszt faktor új technológiai szinten történő kombinációjának alkalmazásából. Például, két teszt faktor, mondjuk F és W faktor sorozatának együttes hozzájárulását a 15.7. táblázat kezelés listája alapján így számítjuk ki:

$$J_{FW} = \bar{Y}_{11} - \bar{Y}_{16}$$

ahol \bar{Y}_{11} és \bar{Y}_{16} a 11. és 16. kezelés átlagtermései. Az F, W és I három teszt faktor sorozatának együttes hozzájárulását a 15.7. táblázat alapján így számítjuk ki:

$$J_{FWI} = \bar{Y}_3 - \bar{Y}_{16}$$

ahol \bar{Y}_3 és \bar{Y}_{16} a 3. és 16. kezelés átlagtermései.

Példánkra, felhasználva a 15.10. táblázat adatait, az F és W teszt faktorok együttes hozzájárulását a 16. gazdaságban a következőképpen számítjuk ki:

$$J_{FW} = 6.50 - 5.22 = 1.28 \text{ t ha}^{-1}$$

15.2.4. Varianciaanalízis GenStat programmal

A Z sorozat adatainak analízise a következő:

Az adatok bevitelekor a gazdaságok (farm), az ismétlés (replication), a trágyázás (F), a rovarok elleni védekezés (I) és a gyomirtás (W), illetve a kezelés képezik a faktorokat, a termés (Yield) a változó. A gazdaságoknak négy, a kezeléseknek nyolc szintje van, míg az ismétlésnek, az F, I és W faktoroknak két szintje van. A számítógépes adatfájl szerkezetét a következő táblázat mutatja:

Farm	lsm.	F	I	W	Kezel.	Termés	Farm	lsm.	F	I	W	Kezel.	Termés
1	1	1	1	1	1	8.76	3	1	1	1	1	1	6.98
1	1	1	1	2	2	4.98	3	1	1	1	2	2	6.43
1	1	1	2	1	3	6.5	3	1	1	2	1	3	6
1	1	1	2	2	4	7.29	3	1	1	2	2	4	5.87
1	1	2	1	1	5	7.63	3	1	2	1	1	5	5.8
1	1	2	1	2	6	7.02	3	1	2	1	2	6	5.47
1	1	2	2	1	7	5.58	3	1	2	2	1	7	4.69
1	1	2	2	2	8	5.44	3	1	2	2	2	8	4.64
1	2	1	1	1	1	6.9	3	2	1	1	1	1	6.36
1	2	1	1	2	2	6.6	3	2	1	1	2	2	6.23
1	2	1	2	1	3	6.5	3	2	1	2	1	3	6.16
1	2	1	2	2	4	6.31	3	2	1	2	2	4	6.03
1	2	2	1	1	5	6.73	3	2	2	1	1	5	5.81
1	2	2	1	2	6	6.27	3	2	2	1	2	6	5.57
1	2	2	2	1	7	5.12	3	2	2	2	1	7	4.67
1	2	2	2	2	8	5.01	3	2	2	2	2	8	4.7
2	1	1	1	1	1	7.21	4	1	1	1	1	1	6.28
2	1	1	1	2	2	6.2	4	1	1	1	2	2	6.14
2	1	1	2	1	3	5.74	4	1	1	2	1	3	4.96
2	1	1	2	2	4	4.74	4	1	1	2	2	4	4.96
2	1	2	1	1	5	5.14	4	1	2	1	1	5	4.86
2	1	2	1	2	6	4.93	4	1	2	1	2	6	4.88
2	1	2	2	1	7	4.4	4	1	2	2	1	7	3.8
2	1	2	2	2	8	3.49	4	1	2	2	2	8	3.49
2	2	1	1	1	1	7.05	4	2	1	1	1	1	5.98
2	2	1	1	2	2	6.62	4	2	1	1	2	2	6
2	2	1	2	1	3	5.76	4	2	1	2	1	3	4.44
2	2	1	2	2	4	5.54	4	2	1	2	2	4	4.21
2	2	2	1	1	5	5.52	4	2	2	1	1	5	4.39
2	2	2	1	2	6	5.12	4	2	2	1	2	6	4.3
2	2	2	2	1	7	5.06	4	2	2	2	1	7	3.16
2	2	2	2	2	8	3.88	4	2	2	2	2	8	3.04

A Z sorozat adatainak varianciaanalízise két lépésben történik. Először a nyolc kezelés együttes hatását analizáljuk. A menük sorrendje: Stats – Analysis of Variance – General Analysis of Variance. Y-variate: Termés (Yield), Treatment Structure: Treatment (Kezelés)× Farm, Block Structure: Replication

Az analízis outputja az alábbi:

Analysis of variance

Variate: Yield

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Farm.Replication stratum					
Farm	3	24.8501	8.2834	14.58	0.013
Residual	4	2.2733	0.5683	3.75	
Farm.Replication.Treatment stratum					
Treatment	7	41.2596	5.8942	38.93	<.001
Farm.Treatment	21	8.0623	0.3839	2.54	0.011
Residual	28	4.2395	0.1514		
Total	63	80.6847			

Tables of means

Variate: Yield

Grand mean 5.552

Farm	1	2	3	4			
	6.415	5.400	5.713	4.681			
Treatment	1	2	3	4	5	6	7
	6.940	6.150	5.757	5.619	5.735	5.445	4.560
Treatment	8						
	4.211						
Farm Treatment	1	2	3	4	5	6	
1	7.830	5.790	6.500	6.800	7.180	6.645	
2	7.130	6.410	5.750	5.140	5.330	5.025	
3	6.670	6.330	6.080	5.950	5.805	5.520	
4	6.130	6.070	4.700	4.585	4.625	4.590	
Farm Treatment	7	8					
1	5.350	5.225					
2	4.730	3.685					
3	4.680	4.670					
4	3.480	3.265					

Standard errors of differences of means

Table	Farm	Treatment	Farm Treatment
rep.	16	8	2
s.e.d.	0.2665	0.1946	0.4511
d.f.	4	28	21.93
Except when comparing means with the same level(s) of			
Farm			0.3891
d.f.			28
Least significant differences of means (5% level)			

Table	Farm	Treatment	Farm Treatment
rep.	16	8	2
l.s.d.	0.7400	0.3985	0.9358
d.f.	4	28	21.93
Except when comparing means with the same level(s) of			
Farm			0.7971
d.f.			28

Az analízis második részében az F, I és W faktorok egyenkénti hatását és kölcsönhatását vizsgáljuk. Ehhez a General Analysis of Variance menüben a varianciaanalízis szerkezete a következő: Y-variate: Yield, Treatment Structure: $F \times I \times W$, Block Structure: Farm/Replication.

A varianciatáblázatot a 15.12. táblázatnak megfelelően állítjuk össze. Megjegyezzük, hogy a számítógépes futtatás outputjában a számított F-értékek eltérnek a 15.12. táblázatban számított F-értékektől. A számítógépes futtatásnál ugyanis az F-érték nevezőjében az összevont hiba MS értéke, míg a 15.12. táblázatban a Kezelés \times Gazdaság MS értéke szerepel.

Az analízis 2. lépésének outputja a következő:

Analysis of variance

Variate: Yield

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Farm stratum	3	24.8501	8.2834	14.58	
Farm.Replication stratum	4	2.2733	0.5683	2.26	

Farm.Replication.*Units* stratum					
F	1	20.3852	20.3852	81.20	<.001
I	1	16.9950	16.9950	67.69	<.001
W	1	2.4571	2.4571	9.79	0.003
F.I	1	0.4830	0.4830	1.92	0.172
F.W	1	0.0841	0.0841	0.33	0.565
I.W	1	0.3511	0.3511	1.40	0.243
F.I.W	1	0.5041	0.5041	2.01	0.163
Residual	49	12.3018	0.2511		
Total	63	80.6847			

Tables of means

Variate: Yield

Grand mean 5.552

F	1	2				
	6.117	4.988				
I	1	2				
	6.067	5.037				
W	1	2				
	5.748	5.356				
F	I	1	2			
1		6.545	5.688			
2		5.590	4.386			
F	W	1	2			
1		6.349	5.884			
2		5.147	4.828			
I	W	1	2			
1		6.338	5.797			
2		5.159	4.915			
F	I	1	2			
	W	1	2	1	2	
1		6.940	6.150	5.757	5.619	
2		5.735	5.445	4.560	4.211	

Standard errors of differences of means

Table	F	I	W	F
rep.	32	32	32	16
d.f.	49	49	49	49
s.e.d.	0.1253	0.1253	0.1253	0.1771

Table	F	I	F
	W	W	I
			W
rep.	16	16	8
d.f.	49	49	49
s.e.d.	0.1771	0.1771	0.2505

Least significant differences of means (5% level)

Table	F	I	W	F
				I
rep.	32	32	32	16
d.f.	49	49	49	49
l.s.d.	0.2517	0.2517	0.2517	0.3560

Table	F	I	F
	W	W	I
			W
rep.	16	16	8
d.f.	49	49	49
l.s.d.	0.3560	0.3560	0.5035

15.3. Termesztési rendszerek vizsgálata (FSR)

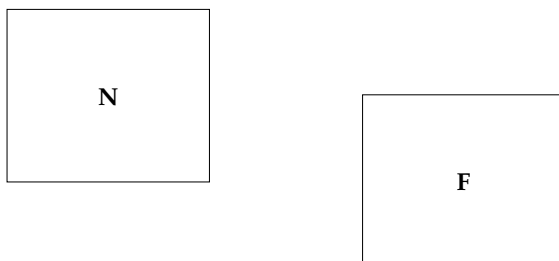
Az 1970-es évek óta az üzemi termesztési rendszerek vizsgálata (FSR, Farming Systems Research) egyre nagyobb figyelmet kapott világszerte, különösen a fejlődő országokban. Alapvető célja, hogy összehasonlítsuk a javított technológia eredményét a termelő által használt termesztési rendszer eredményével. Az FSR kísérletek termelői orientáltak, interdiszciplinárisak és miközben kiegészítik a mezőgazdasági kutatás fő áramlatát, nem helyettesítik azt. Az FSR magában foglalja a termelő (farmer) részvételét az új technológia adaptálásában a gazdaság körülményei között. Az üzemi FSR kísérlet lényegében csak két kezelést tartalmaz: a farmer technológiája (F) és új javított technológia (N).

A javított technológiának tulajdonítható termésnövekedés szignifikanciájának tesztje a kísérleti adatok varianciaanalízisén alapul. Az analízis felhasználható a minimális növekedés (LSI) meghatározásához, amely szükséges a szignifikancia eléréséhez. Következésképpen, az LSI számításához az egyoldalú t-próba értékeit használjuk fel. Az üzemi FSR kísérleteket több ismétlésben állítjuk be, mint a kutató intézetek kísérleteit. Mindegyik gazdaságban mindegyik kezelés egyetlen ismétlését állítjuk be. A kísérleti adatokat randomizált blokkelrendezésként analizáljuk, és a gazdaságok képezik a blokkokat. A kísérlet tervezése függ attól, hogy az új technológia egy vagy több komponensű. Az elemi üzemi FSR kísérletek csupán két kezelést tartalmaznak: termelői technológia (F) és javított technológia (N). Az FSR kísérletek analízisének módszerét *Petersen* (1994) alapján ismertetjük.

15.3.1. Elemi üzemi FSR kísérletek

A kísérlet elhelyezése

A kísérleti terv tartalmaz két nagy parcellát mindegyik gazdaságban, melyeket kiválasztottunk a célterületen. Ezeket a parcellákat reprezentatív táblákon kell elhelyezni a gazdaságban. A két parcellának ugyanolyan nagyságúnak és alakúnak kell lenni, azonban nem szükséges egymás mellett lenniük. Elegendő, ha mindkettő reprezentálja ugyanazokat a körülményeket talajtípusban, topográfiában, vízgazdálkodásban és így tovább. A parcelláknak elég nagyoknak kell lenni ahhoz, hogy lehetővé tegyék a normális termelési műveleteket. A parcellák egyikét kijelöljük véletlenül a javított technológiára, N-re, míg a másikat kijelöljük a farmer által használt technológiára, F-re. A növénykísérlet szántóföldi tervét az 15.2. *ábra* szemlélteti. A kísérletet beállítjuk és a termést és más adatokat feljegyezzük mindegyik parcelláról mindegyik kísérleti helyen.



15.2. *ábra.* Az elemi két kezeléssel üzemi FSR kísérlet szántóföldi elhelyezése

Adatelemzés

Tételezzük fel, hogy egy két kezeléssel kísérletet állítottunk be mindegyik kísérleti helyen. Reprézentalja y_{ij} a termést (vagy más jellemzőt, mint amilyen a nettó megtérülés) egységnyi területen a j -edik kezelésre és az i -edik helyre vonatkozóan. Az eredményeket táblázatba foglaljuk, mint az 15.14. táblázat mutatja.

15.14. táblázat. Az elemi két kezeléssel üzemű kísérlet (FSR) adattáblázata^a

Hely	Kezelés		Különbség
	F	N	
1	y_{11}	y_{12}	d_1
2	y_{21}	y_{22}	d_2
.	.	.	.
.	.	.	.
.	.	.	.
n	y_{1n}	y_{2n}	d_n
Összesen			$\frac{G}{d}$
Átlag			$\frac{G}{d}$

$$^a d_i = y_{i2} - y_{i1}$$

= a javított gyakorlat és a termelői gyakorlat közötti különbség

$$G = \sum_i d_i = \text{a különbségek főösszege}$$

$$\bar{d} = G/n = \text{átlagos különbség}$$

Az adatelemzésnek viszonylag egyszerű formája van:

1. Számítsuk ki MSE-t = $\left[\frac{1}{(n-1)} \left(\sum_i d_i^2 - \frac{G^2}{n} \right) \right]$

2. A legkisebb szignifikáns növekedést (LSI) a következőképpen számítjuk ki:

$$LSI_{0.05} = t_{0.05} \sqrt{\frac{MSE}{n}} \text{ az 5\%-os szinten}$$

$$LSI_{0.01} = t_{0.01} \sqrt{\frac{MSE}{n}} \text{ az 1\%-os szinten}$$

Jegyezzük meg, hogy $t_{0.05}$ és $t_{0.01}$ az 5%-os és 1%-os egyoldalú táblázati t-érték $n-1$ d.f. értékkel.

Ha $\bar{d} > LSI_{0.05}$, akkor a növekedés a farmer gyakorlathoz viszonyítva szignifikáns 5%-os szinten.

Ha $\bar{d} > LSI_{0,01}$, akkor a növekedés a farmer gyakorlathoz viszonyítva szignifikáns 1%-os szinten.

Ha $\bar{d} < LSI_{0,05}$, akkor a javított gyakorlat szignifikánsan nem jobb, mint a farmer által alkalmazott gyakorlat.

Kísérlet leírása

Beállítottak egy kísérletet, hogy összehasonlítsák az új búzafajtát, N-t, az általánosan termesztett fajtával, F-fel. A kísérletet nyolc helyen állították be. Két parcellát helyeztek el mindegyik táblán minden kísérleti helyen. A régi fajtát random helyezték el az egyik parcellára és az új fajtát jelölték ki a másik parcellára. A kapott termést (kg parcella⁻¹) a 15.15. táblázat mutatja.

15.15. táblázat. Az új (N) és régi (F) búzafajták termése az üzemi kísérletben

Hely	Fajta		d = N – F
	F	N	
1	30.2	32.2	2.0
2	31.5	35.6	4.1
3	37.1	41.9	4.8
4	30.8	35.4	4.6
5	23.2	25.4	2.2
6	24.8	27.1	2.3
7	26.7	33.8	7.1
8	26.7	31.1	4.4
Összesen			31.5
Átlag			3.94

A számítás menete a következő:

$$1. \text{MSE} = \left[\frac{1}{(n-1)} \right] \left[\sum_i d_i^2 - G^2/n \right]$$

$$= \left[\frac{1}{(8-1)} \right] \left[144.9100 - 31.5^2/8 \right]$$

$$= \left(\frac{1}{7} \right) (144.9100 - 124.0312)$$

$$= 20.8788/7 = 2.9827$$

$$2. LSI_{0,05} = 1.90 \sqrt{2.9827/8} = (1.90)(0.6106) = 1.16$$

$$LSI_{0,01} = 3.00 \sqrt{2.9827/8} = (3.00)(0.6106) = 1.83$$

Mivel $3.94 > 1.83$, az új fajta termése szignifikánsan nagyobb az általánosan termesztett fajtánál. Az új fajta közel $4 \text{ kg parcella}^{-1}$ -val nagyobb termést ad, mint a régi fajta.

15.3.2. Üzemi FSR kísérlet kiegészítő kezelésekkel

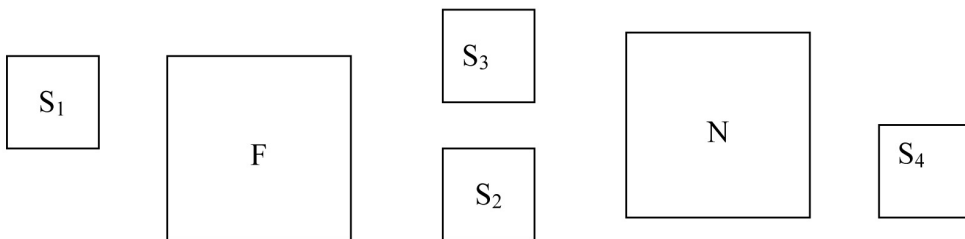
Néha indokolt bevonni más kezeléseket is a termelői gyakorlaton (technológia) és a javított gyakorlaton (technológia) kívül az üzemi FSR kísérletbe. Például, érdekes lehet megvizsgálni néhány új fajtát vagy néhány eltérő herbicidet, amelyek nincsenek beépítve a javított technológiába. Továbbá, érdekes lehet látni, hogyan teljesítenek ezek az új kezelések üzemi feltételek mellett.

A kísérlet elhelyezése

Az alapterv két kezelésre, F-re és N-re, módosítható, hogy beépítsünk akár további hat kiegészítő kezelést, S_1, S_2, \dots, S_6 . Úgy, mint az elemi növénykísérleti tervnél, számos gazdaságot (farm) választottunk ki a célterületen. A gazdaságok minimális száma, melyet használni kívánunk, függ a kísérletbe bevont kiegészítő kezelések számától. Ha csak egy vagy két extra kezelést használunk, a gazdaságok minimális száma hét vagy nyolc. Ha öt vagy hat kiegészítő kezelést vonunk be, a gazdaságok minimális száma négy vagy öt lehet.

Mindegyik gazdaságban a kísérleti terv tartalmaz két nagy és több kis parcellát. A termelői gyakorlatot, F-et jelöljük ki random a nagy parcellák egyikére, míg a javított gyakorlat N, kerül a másik parcellára. A kiegészítő kezeléseket, S_j -t, random jelöljük ki a kis parcellákra, egy kezelést mindegyik parcellára.

Egy kísérlet szántóföldi tervét négy kiegészítő kezeléssel a 15.3. ábra mutatja. A parcellák elhelyezhetők ugyanabban a táblában, de nem szabad egymással határosnak lenniük. Tartsuk emlékeztünkben azonban, hogy egy üzemi FSR kísérlet célja, hogy demonstrálja a javított gyakorlatot, összehasonlítva a termelő gyakorlatával. Ezen okból a nagy parcelláknak viszonylag egymáshoz közel kell elhelyezkedni. Minden parcellát úgy kell kiválasztani, hogy amennyire lehet, hasonlóak legyenek a kezeléseik beállítása előtt.



15.3. ábra. Az üzemi FSR kísérlet szántóföldi elhelyezése négy kiegészítő kezeléssel

Adatelemzés

Tételezzük fel, hogy a kísérletet n helyen állították be és a kezelések a következők: F , N és p ($0 < p < 6$) kiegészítő kezelés. Reprézentalja y_{ij} a termést (vagy más mért tulajdonságot) egységnyi területen a j -edik kezelésre és az i -edik helyre vonatkozóan. (Jegyezzük meg: mivel különböző parcellaméretekkel használunk, a termést egységnyi területre vetítve kell kifejezni, hogy összehasonlíthatók legyenek). A kísérlet eredményeit táblázatba foglaljuk, mint a 15.16. táblázat mutatja. A varianciaanalízis (ANOVA) szerkezetét az eredmények analizésére a 15.17. táblázat mutatja. A 15.17. táblázat adatait a következőképpen számítjuk ki:

$$\text{Összes SS} = \sum_i \sum_j y_{ij}^2 - \frac{G^2}{n(p+2)}$$

$$\text{SSL} = \frac{1}{p+2} \sum_i L_i^2 - \frac{G^2}{n(p+2)}$$

$$\text{SST} = \frac{1}{n} \sum_j T_j^2 - \frac{G^2}{n(p+2)}$$

$$\text{SSE} = \text{SSÖsszes} - \text{SSL} - \text{SST}$$

$$\text{MSE} = \frac{\text{SSE}}{(n-1)(p+1)}$$

15.16. táblázat. Az adatok összegező táblázata üzemi kísérletre p számú kiegészítő kezeléssel^a

Hely	F	N	S ₁	S ₂	...S _p	Összesen
1	y ₁₁	y ₁₂	y ₁₃	y ₁₄	... y _{1, p+2}	L ₁
2	y ₂₁	y ₂₂	y ₂₃	y ₂₄	... y _{2, p+2}	L ₂
.
.
.
n	y _{n1}	y _{n2}	y _{n3}	y _{n4}	... y _{n, p+2}	L _n
Összesen	T ₁	T ₂	T ₃	T ₄	... T _{p+2}	G
Átlag	\bar{y}_1	\bar{y}_2	\bar{y}_3	$\bar{y}_4 \dots$	\bar{y}_{p+2}	

^aL_i = $\sum_j y_{ij}$ = összes termés i helyen

T_j = $\sum_i y_{ij}$ = a j kezelés összes termése

G = $\sum_i L_i = \sum_j T_j$ = főösszeg termés

y_j = T_j/n

15.17. táblázat. Az adatok varianciaanalízisének formája egy üzemi FSR kísérletre, kiegészítő kezelésekkel

Forrás	d.f.	SS	MS
Összes	$n(p+2)+1$	SSÖsszes	
Hely	$n-1$	SSL	
Kezelés	$p+1$	SST	
Hiba	$(n-1)(p+1)$	SSE	MSE

Egy kezelésátlag szórása s_y a következő: $s_y = \sqrt{\frac{MSE}{n}}$

Két kezelésátlag közötti különbség szórása, $s_d = \sqrt{\frac{2MSE}{n}}$

A többi kezelés átlagos növekedését \bar{d}_j összehasonlítva a termelői gyakorlattal úgy számítjuk ki, hogy kivonjuk a termelői gyakorlat átlagát, \bar{y}_1 -t a többi kezelés \bar{y}_j átlagából.:

$$\bar{d}_j = \bar{y}_j - \bar{y}_1$$

A legkisebb szignifikáns növekedést (LSI) úgy kapjuk meg:

$$LSI_{0.05} = t_{0.05} \sqrt{\frac{2MSE}{n}} \text{ az } 5\% \text{-os szinten}$$

$$LSI_{0.01} = t_{0.01} \sqrt{\frac{2MSE}{n}} \text{ az } 1\% \text{-os szinten}$$

ahol $t_{0.05}$ és $t_{0.01}$ a 0.05-ös és 0.01-es egyoldalú táblázati t-érték $(n-1)(p+1)$ szabadságfoknál.

Ha $\bar{d}_j > LSI_{0.05}$, akkor a növekedés szignifikáns 5%-os szinten

Ha $\bar{d}_j > LSI_{0.01}$, akkor a növekedés erősen szignifikáns 1%-os szinten

Ha $\bar{d}_j < LSI_{0.05}$, akkor a növekedés nem szignifikáns

Kísérlet leírása

Egy üzemi FSR kísérletet állítottak be, hogy összehasonlítsák a javított gyakorlati csomagot, N-t a szokásos termelői gyakorlattal, F-fel egy búzatermesztési régióban. Egyidejűleg négy új búzafajtát, S_1, S_2, S_3 és S_4 , állítottak be a kísérletbe, hogy lássák teljesítményüket üzemi körülmények között. A kísérletet öt helyen állították be. A búza-termést a 15.18. táblázat tartalmazza és a termés varianciaanalízisét a 15.19. táblázat mutatja.

15.18. táblázat. A búza termése (t/ha) az üzemi kísérletben négy kiegészítő kezeléssel

Termőhely	F	N	S ₁	S ₂	S ₃	S ₄	Összesen
1	3.21	4.45	4.50	4.46	3.68	4.20	24.50
2	3.31	4.89	5.10	4.79	4.21	4.31	26.61
3	3.86	5.68	6.10	5.67	4.72	4.89	30.92
4	3.74	5.47	6.40	5.21	4.80	4.77	30.39
5	3.99	6.01	7.18	5.79	5.28	5.39	33.64
Összesen	18.11	26.50	29.28	25.92	22.69	23.56	146.06
Átlag	3.62	5.30	5.86	5.18	4.54	4.71	

A kezelésátlag szórása, s_y :

$$s_y = \sqrt{\frac{MSE}{n}} = \sqrt{\frac{0.0722}{5}} = 0.1201$$

15.19. táblázat. A búzatermés (t ha⁻¹) varianciaanalízise.
Az adatok egy üzemi kísérletből négy kiegészítő kezeléssel

Forrás	df	SS	MS
Összes	29	24.9997	
Hely	4	8.8142	
Kezelés	5	14.7419	
Hiba	20	1.4436	0.0722

és két kezelésátlag közötti különbség szórása:

$$s_{\bar{d}} = \sqrt{\frac{2MSE}{n}} = \sqrt{\frac{(2)(0.0722)}{5}} = 0.1699$$

A többi kezelés termésnövekedését a termelői gyakorlathoz viszonyítva a 15.20. táblázat tartalmazza.

A búza termésadatakra a legkisebb szignifikáns növekedés (LSI) a következő:

$$LSI_{0.05} = 1.72 \sqrt{\frac{(2)(0.0722)}{5}} = 0.29 \quad 5\text{-os szinten}$$

$$LSI_{0.01} = 2.53 \sqrt{\frac{(2)(0.0722)}{5}} = 0.43 \quad 1\text{-os szinten}$$

15.20. táblázat. A többi kezelés átlagos terméshozvétele (t ha⁻¹), összehasonlítva a termelői gyakorlattal^a

Kezelés	N	S ₁	S ₂	S ₃	S ₄
Növekedés	1.68	2.24	1.56	0.92	1.09

^a szórás = 0.17

Összehasonlítva a termelői gyakorlat átlagtermésével, mindegyik kezelés terméshozvétele nagyobb, mint 0.43. Ezáltal a javított gyakorlati csomag és az új fajták szignifikánsan nagyobb termést eredményeztek, mint a termelői gyakorlat. Az átlagok áttekintése a 15.18. táblázatban és a terméshozvétele a 15.20. táblázatban ahhoz a következtetéshez vezet, hogy az S₁ fajta látszik legígéretesebbnek a jövőbeni vizsgálatokra.

15.3.3. Üzemi FSR kísérlet két komponensű javított gyakorlattal

A javított technológia csomag gyakran tartalmaz két típusú ráfordítást (input). Például, a ráfordítások tartalmazhatnak egy új növényfajtát és egy herbicid használatát a gyomok irtására. Ebben az esetben érdekes lehet meghatározni, hogy az új gyakorlat terméshozvétele az egyik vagy a másik komponens eredménye-e, vagy vajon mindkét komponens hatással van-e egymásra. Kaphatunk bizonyos információt erre a kérdésre az alapterv módosításával. Ezért két további kezelést adunk az alaptervhez, és hatásában a kezeléseknél egy 2×2-es faktoriális sorát alakítjuk ki. Ezután megnézzük a komponensek fő hatásait és az interakciót (közös hatást), amikor analizáljuk az adatokat.

A kísérlet elrendezése

Tételezzük fel, hogy a javított gyakorlat két komponenset tartalmaz: *A* és *B*. Az *A* komponens például lehet egy új fajta és a *B* komponens lehet egy herbicid. A kísérletet négy kezeléssel állítjuk be:

F = termelői gyakorlat

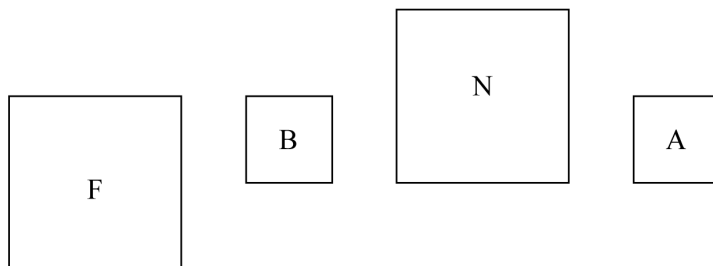
N = javított gyakorlat = új fajta + herbicid

A = komponens *A* egyedül = új fajta herbicid nélkül

B = komponens *B* egyedül = herbicid + termelői fajta

Több farmot választunk ki a célterületen. Ennél a tervnél a helyek minimális száma ötre csökkenthető. Minden helyen a kísérleti terv tartalmaz két nagy parcellát, amelyekre az *F* és *N* kezeléseket alkalmazzuk, és két kis parcellát, melyeken az *A* és *B* kezelést állítjuk be. A két komponensű elrendezés szántóföldi tervét a 15.4. ábra mutatja.

Mindegyik parcellának ugyanabban a táblában kell elhelyezkedni, de nem szükséges egymással határosnak lenniük. A parcellák elhelyezésének célja, hogy úgy válasszuk ki a parcellákat, hogy amennyire lehetséges, hasonlóak legyenek a kezeléseknél beállítását megelőzően. (Jegyezzük meg: Mivel különböző parcella méretek vannak a kísérletben, a termést egységnyi területre kell kifejezni, hogy összehasonlíthatók legyenek.)



15.4. ábra. Egy üzemi FSR kísérlet szántóföldi elhelyezése, a javított gyakorlat két komponensű

Adatelemzés

Tételezzük fel, hogy két-komponensű kísérletet állítunk be n termőhelyen (üzemben). Reprezentálja y_{ij} a termést (vagy más tulajdonságot) egységnyi területen a j -edik kezelésre és az i -edik helyre vonatkozóan. Az adatokat táblázatba foglaljuk (15.21. táblázat). Az adatokat varianciaanalízisnek vetjük alá, melynek formáját a 15.22. táblázat mutatja.

15.21. táblázat. Az üzemi FSR kísérlet adattáblázata, két komponensű javított gyakorlat esetén^a

Hely	Kezelések				Összesen
	F	N	A	B	
1	y_{11}	y_{12}	y_{13}	y_{14}	L_1
2	y_{21}	y_{22}	y_{23}	y_{24}	L_2
.
.
.
n	y_{n1}	y_{n2}	y_{n3}	y_{n4}	L_n
Összesen	T_1	T_2	T_3	T_4	G
Átlag	\bar{y}_1	\bar{y}_2	\bar{y}_3	\bar{y}_4	

$${}^a L_i = \sum_j y_{ij} = \text{összes termés } i \text{ helyen}$$

$$T_j = \sum_i y_{ij} = \text{a } j\text{-edik kezelés összes termése}$$

$$G = \sum_j L_i = \sum_j T_j = \text{a termés } f \text{ összeg}$$

$$\bar{y}_j = T_j / n = \text{a } j\text{-edik kezelés átlagtermése}$$

15.22. táblázat. Az adatok varianciaanalízisének formája egy üzemi FSR kísérlet két komponensű javított gyakorlata esetén

Forrás	d.f.	SS	MS
Összes	4 n-1	SSTOT	
Hely	n-1	SSL	
Kezelés	3	SST	
Hiba	3(n-1)	SSE	MSE

A táblázat tényezőit a következőképpen számítjuk ki:

$$\text{Összes SS} = \sum_i \sum_j y_{ij}^2 - \frac{G^2}{4n}$$

$$\text{SSL} = \frac{1}{4} \sum_i L_i^2 - \frac{G^2}{4n}$$

$$\text{SST} = \frac{1}{n} \sum_j T_j^2 - \frac{G^2}{4n}$$

$$\text{SSE} = \text{Összes SS} - \text{SSL} - \text{SST}$$

$$\text{MSE} = \frac{\text{SSE}}{3(n-1)}$$

Jegyezzük meg, hogy az ANOVA kiszámításának fő oka az, hogy megkapjuk az MSE-t, a kísérleti hiba MS-t. Ennél az analízisnél nincs szükség a többi MS kiszámítására.

A kezelésátlag szórását, s_y -t a következőképpen számítjuk ki:

$$s_y = \sqrt{\frac{\text{MSE}}{n}}$$

Hasonlóan, kiszámítjuk a két kezelés közötti különbség szórását, -t:

$$s_{\bar{d}} = \sqrt{\frac{2\text{MSE}}{n}}$$

A többi kezelés átlagos növekedését \bar{d}_j -t, összehasonlítva a termelői gyakorlattal, úgy számítjuk ki, hogy a termelői gyakorlat átlagtermését \bar{y}_1 kivonjuk a többi kezelés átlagából, \bar{y}_j :

$$\bar{d}_j = \bar{y}_j - \bar{y}_1$$

A legkisebb szignifikáns növekedést, LSI-t a következőképpen számítjuk ki:

$$LSI_{0,05} = t_{0,05} \sqrt{\frac{2MSE}{n}} \quad 5\text{-os szinten}$$

$$LSI_{0,01} = t_{0,01} \sqrt{\frac{2MSE}{n}} \quad 1\text{-os szinten}$$

ahol $t_{0,05}$ és $t_{0,01}$ a 0.05, 0.01 táblázati egyoldalú t-értékek 3 (n-1) d.f. értéknél.

Ha $\bar{d}_j > LSI_{0,05}$, a növekedés szignifikáns 5%-os szinten.

Ha $\bar{d}_j > LSI_{0,01}$, a növekedés erősen szignifikáns 1%-os szinten.

Ha $\bar{d}_j < LSI_{0,05}$, a növekedés nem szignifikáns.

A javított gyakorlat komponensei

A javított gyakorlat két komponense hathat egymástól függetlenül vagy kölcsönhatásban (interakció) lehetnek egymással és így fejtik ki hatásukat. Annak meghatározásához, hogy melyik esetről van szó, a variáció kezelés forrását feloszthatjuk három részre: mindegyik komponens fő hatása, és kölcsönhatásuk. Ebből a célból a kontrasztok fogalmát használjuk. A négy kezelés egy 2×2 -es faktoriális sort képez és a k_{ij} koeficienseket használjuk a 15.23. táblázatban, a felosztás elvégzéséhez.

15.23. táblázat. Kontrasztok, amelyek definiálják a kezelés eltérés négyzet összegek megoszlását fő hatásokra és interakcióra, egy üzemi FSR kísérlet két komponensű javított gyakorlata esetén^a

Hatás \ Kezelés	Kezelés					
	F	N	A	B		
Összes	T_1	T_2	T_3	T_4	E_i	S_i^2
A	-1	+1	+1	-1	E_A	S_A^2
B	-1	+1	-1	+1	E_B	S_B^2
AB	+1	+1	-1	-1	E_{AB}	S_{AB}^2

^a T_j = összes termés j kezelésre

E_i = i-edik faktoriális hatás = $\sum_j k_{ij} T_j$

S_i^2 = négyzetösszegek (1-d.f.), összefüggésben i hatással = $E_i^2 / 4n$

Ezeket a koeficienseket használva, a számítás a következő:

$$E_A = -T_1 + T_2 + T_3 - T_4 \quad S_A^2 = \frac{E_A^2}{4n}$$

$$E_B = -T_1 + T_2 - T_3 + T_4 \quad S_B^2 = \frac{E_B^2}{4n}$$

$$E_{AB} = +T_1 + T_2 - T_3 - T_4 \quad S_{AB}^2 = \frac{E_{AB}^2}{4n}$$

Most már tesztelni tudjuk a főhatások és az interakciók szignifikanciáját. Erre a célra használjuk az F -vizsgálatot.

$F_A = s_A^2 / MSE$, 1 és $3(n-1)$ d.f.-nél teszteli az önmagában ható A komponens szignifikanciáját.

$F_B = s_B^2 / MSE$, 1 és $3(n-1)$ d.f.-nél teszteli az önmagában ható B komponens szignifikanciáját.

$F_{AB} = s_{AB}^2 / MSE$, 1 és $3(n-1)$ d.f.-nél teszteli az A és B komponens interakciójának szignifikanciáját.

Ha F_{AB} szignifikáns, akkor a két komponens hatása nem független egymástól. Ebben az esetben a javított gyakorlatból származó növekedés függ a csomagban levő mindkét komponenstől.

Ha F_{AB} nem szignifikáns, akkor megvizsgáljuk F_A -t és F_B -t. Ha ezek közül bármelyik szignifikáns, megnézzük a szignifikáns komponensből származó növekedést és nem szabad foglalkoznunk a másik komponens meglétével vagy hiányával.

Kísérlet leírása

Egy üzemi FSR kísérletet állítottak be Szíriában, hogy értékeljenek egy két komponensű javított technológiai csomagot a búzatermesztésben. A két komponens tartalmazott (1) egy új javított búzafajtát, és (2) herbicidet a gyomszabályozásra. A kísérletet hat gazdaságban állították be, melyeket véletlenül választottak ki a célterületen. Mindegyik gazdaságban volt két nagyparcella, melyeket az F és N kezeléseknek jelölték ki, és két kis parcella, melyeken az A és B kezeléseket állították be. A kezelések a következők voltak:

F = termelő fajtája, herbicid nélkül,

N = új fajta + herbicid,

A = új fajta, herbicid nélkül,

B = termelő fajtája + herbicid.

A kísérletben kapott búzaterméseket a 15.24. táblázat tartalmazza. A termésadatok varianciaanalízise a 15.25. táblázatban található. Az új gyakorlat (technológia) terméshozását és komponenseit, összehasonlítva a termelői gyakorlattal, a 15.26. táblázat mutatja.

15.24. táblázat. Egy üzemi FSR kísérlet búzatermése (kg ha^{-1}), amely magában foglal két komponensű javított technológiát

Termőhely	Kezelés				Összesen
	F	N	A	B	
1	4100	4030	4365	4030	16525
2	3760	4165	4030	4030	15985
3	3895	3895	4100	3965	15855
4	3030	4570	4100	3360	15060
5	3695	4760	4030	3630	16115
6	2890	4500	3360	3025	14755
Összesen	21370	25920	23985	22040	93315
Átlag	3562	4320	3998	3673	

15.25. táblázat. Az üzemi FSR kísérlet (két javított komponenssel) búzatermésének (kg ha⁻¹) varianciaanalízise

Forrás	d.f.	SS	MS
Összes	23	5300691	
Hely	5	1235697	
Kezelés	3	2107136	
Hiba	15	1957858	130524

15.26. táblázat. A javított gyakorlatból és komponenseiből származó termésvnövekedés, összehasonlítva a termelői gyakorlattal

Kezelés	N	A	B
Növekedés	758	436	111

A kísérletben a legkisebb szignifikáns növekedés (LSI) a következő:

$$LSI_{0,05} = t_{0,05} \sqrt{\frac{2MSE}{6}} = 1.75 \sqrt{\frac{(2)(130524)}{6}} = 365$$

$$LSI_{0,01} = t_{0,01} \sqrt{\frac{2MSE}{6}} = 2.60 \sqrt{\frac{(2)(130524)}{6}} = 542$$

Mivel $758 > 542$, a termésvnövekedés az új gyakorlatnak (technológiának) tulajdoníthatóan erősen szignifikáns, és mivel $542 > 436 > 365$, az új fajtának tulajdonítható termésvnövekedés szignifikáns. Továbbá, mivel $111 < 365$, a herbicid hatására bekövetkező termésvnövekedés nem volt szignifikáns.

Ezután vizsgáljuk meg a javított gyakorlat komponenseinek fő hatását és az interakciót a kontrasztok módszerével, melyet a 15.27. táblázat mutat.

15.27. táblázat. Kontrasztok, amelyek meghatározzák az új fajta és a herbicid fő hatását és az interakciót a két komponensű üzemi FSR kísérletben

Összes	Kezelés				E	s ²
	F	N	A	B		
Hatás	21370	25920	23985	22040		
A	-1	+1	+1	-1	6495	1757709
B	-1	+1	-1	+1	2605	282751
AB	+1	+1	-1	-1	1265	66676

A 15.27. táblázat adatainak kiszámítása példaként a következő:

$$E_A = (-21370) + 25920 + 23985 + (-22040) = 6495$$

$$s_A^2 = \frac{6495^2}{(6)(4)} = \frac{42185025}{24} = 1757709$$

Jegyezzük meg: $1757709 + 282751 + 66676 = 2107136 = \text{SST}$. E hatások szignifikanciájának tesztelésére használjuk az F-próbát.

$$F_A = \frac{1757709}{130524} = 13.47$$

$$F_B = \frac{282751}{130524} = 2.17$$

$$F_{AB} = \frac{66676}{130524} = 0.51$$

$$F_{0.01(1,15)} = 8.68 \quad F_{0.05(1,15)} = 4.54$$

Mivel $0.51 < 4.54$, az interakció a fajta és a herbicid között nem szignifikáns, és mivel $2.17 < 4.54$, a herbicid fő hatása nem szignifikáns. Mivel azonban $13.47 > 8.68$, a fajta fő hatása erősen szignifikáns.

15.28. táblázat. A búza átlagtermése (kg ha⁻¹) a termelő fajtájával és a javított fajtával

Fajta	Termelő fajtája	Javított fajta	Standard hiba
Termésátlag	3618	4159	104.3

Standard hiba = $\text{MSE}/2n = 130542/12 = 104.3$

A két komponensű üzemi FSR kísérlet eredményét a 15.28. táblázatban összegeztük. A táblázatban látható, hogy átlagosan a javított fajta a termelő fajtájának termését mintegy 540 kg ha⁻¹-ral felülmúlta. Másrészt a herbicidnek nem volt kimutatható hatása a termésre, hasonlóan nem volt hatása az új fajta termésnövekedéséből, attól függően, hogy herbicidet használtunk vagy nem.

15.3.4. Számítógépes értékelés GenStat programmal

Az Analysis of Variance menüből kiválasztjuk: One-Way ANOVA (in randomized blocks), Y-variate: Yield (termés), Treatments: Treatment (Kezelés), Blocks: Location, Contrasts: Contrast Factor: Treatment, Number of Contrasts: 3, Contrast Type: Comparison. Ezt követően kitöltjük a kontraszt táblázatot a 15.27. táblázat szerint. Options: bejelöljük az LSDs-t. Esetünkben a t-érték egyoldalú, ezért a valószínűség kétszeresét kell venni (5%-nál 10%-ot, 1%-nál 2%-ot). LSD Significance Level (%): 10, illetve 2. A Run-ra kattintással a program elvégzi a varianciaanalízist. Az output a következő:

Analysis of variance

Variate: Yield

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Location stratum	5	1235697.	247139.	1.89	
Location.*Units* stratum					
Treatment	3	2107136.	702379.	5.38	0.010
Contrast 1	1	1757709.	1757709.	13.47	0.002
Contrast 2	1	282751.	282751.	2.17	0.162
Contrast 3	1	66676.	66676.	0.51	0.486
Residual	15	1957857.	130524.		
Total	23	5300691.			

Tables of means

Variate: Yield

Grand mean 3888.

Treatment	1	2	3	4
	3562.	4320.	3998.	3673.

Standard errors of differences of means

Table	Treatment
rep.	6
d.f.	15
s.e.d.	208.6

Least significant differences of means (2% level)

Table	Treatment
rep.	6
d.f.	15
l.s.d.	542.8

Least significant differences of means (10% level)

Table	Treatment
rep.	6
d.f.	15
l.s.d.	365.7

16. Tartamkísérletek

16.1. A tartamkísérletek jelentősége és sajátosságai

A tartamkísérletek (újabban agroökoszisztéma tartamkísérletek) úgy definiálhatók, mint nagyméretű, több, mint 20 éves szántóföldi kísérletek, amelyekben tanulmányozzuk a növénytermesztési tényezőket, a tápelem ciklusokat és a mezőgazdaság környezeti hatásait. Forrásként szolgálnak a mezőgazdasági fenntarthatóság biológiai, biogeokémiai és környezeti dimenzióinak értékelésére; a jövőbeni globális változások előrejelzésére; és a modell alkalmasság és teljesítmény ellenőrzésére (*Rasmussen et al.*, 1998). Tartamkísérletek a világ számos országában vannak és a jelenleg rendelkezésre álló legnagyobb időbeni és térbeni adatbázist képezik az ökoszisztéma változás meghatározására. Az 50 évnél hosszabb ideje fenntartott tartamkísérleteket klaszter tartamkísérleteknek hívjuk.

A tartamkísérletek közös jellemzője, hogy ugyanazokat a kezeléseket ugyanazon parcellákon állítjuk be minden évben, a kezelések időbeni, ún. tartamhatásának tanulmányozása céljából. A tartamkísérletek időtartama függ a kutatás céljától és változik néhány évtizedtől egészen akár több, mint 150 évig. A legrégebbi hazai tartamkísérletek (pl. Martonvásáron, Keszthelyen stb.) napjainkban voltak 50 évesek, az újabbak (Debrecen, Gödöllő) több, mint 20 évesek (*Berzsenyi*, 2009b; *Kismányoky és Jolánkai*, 2009; *Nagy*, 2009; *Pepó*, 2009). A tartamkísérletek nélkülözhetetlenek fontos agronómiai és ökológiai, valamint ökonómiai kérdések megválaszolásában. Összehasonlításképpen, a rövid időtartamú (2–3 éves) kísérletek prediktív ereje nagyon gyenge. E tartamkísérletek alapján lehet megalapozottan értékelni a földművelési és növénytermesztési rendszerek hatékonyságát. A tartamkísérletek élő szabadföldi kísérleti laboratóriumok és a nemzeti vagyoni részének tekintendők. A világ különböző országaiban beállított legrégebbi tartamkísérleteket az 16.1.–16.6. kép mutatja.

Kizárólag tartamkísérletek adnak megfelelő indikátorokat a természet fenntarthatóságáról. A növénytermesztés kérdéseinek eldöntésére leginkább az egzakt tartamkísérletek alkalmasak. A talaj termékenységének változását, a befolyásoló tényezők szerepét csak több évtizeden át végzett tartamkísérletekben lehet tanulmányozni, ezért nélkülözhetetlenek a természetési eljárások és technológiák tartamhatásának tanulmányozására (*Árendás és Csathó*, 2002).

A tartamkísérletekben tanulmányozzuk a kezelések tartamhatásait a lassan változó folyamatokra (talaj szervesanyag tartalma, pH, talaj eredetű betegségek stb.). Betelekinthetünk a változás mechanizmusába. A tartamkísérletek ideálisak a tápelem ciklusok tanulmányozására is. Polifaktoriális tartamkísérletekben folynak az adott ökológiai körzetben maximálisan fenntartható termés meghatározására irányuló kutatások. Ezek célja a különböző tápanyagok, illetve tápanyagok és az agrotechnikai tényezők közötti interakciók identifikálása, azok okainak feltárása és megfelelő szabályozása. A tartamkísérletek adatokat szolgáltatnak a termés és az időjárás közötti összefügg-



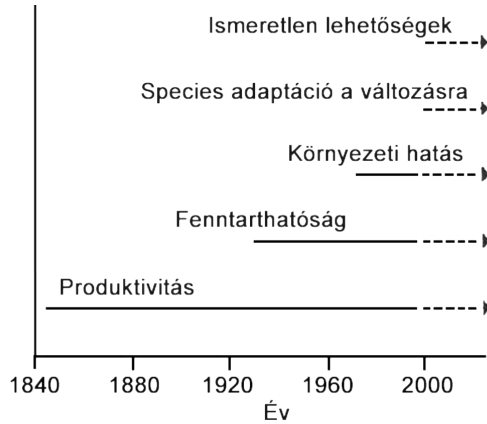
16.1. kép. Ökológiai tartamkísérlet (LTER) parcellái. Helye: Hickory Corners, Michigan, USA. Beállítás éve: 1989.

gés hosszú távú tanulmányozásához, tesztelő helyet képeznek a modern kísérleti módszerek alkalmazásához. A tartamkísérletekben végzett részletes vizsgálatok (termés és terméskomponensek, növény- és talajvizsgálatok, növekedésanalízis) alapul szolgálhatnak a predikcióhoz, másrészt ahhoz, hogy az eredményeket interpretáljuk a talaj- és klimatikus viszonyok szélesebb körére (pl. szimulációs termésmodellek). A tartamkísérletek értékesek a lassú ökológiai folyamatok vizsgálatában, a ritka vagy epizódikus jelenségek tanulmányozásában, a nagy variabilitású folyamatok vizsgálatában, a komplex jelenségek értelmezésében (*Debreczeni és Németh, 2009*).

Az MTA Mezőgazdasági Kutatóintézetében, Martonvásáron, az 1950-es évek végén és az 1960-as évek elején Győrfly Béla állította be azokat a tartamkísérleteket, amelyek az országban a legrégebbiek közé tartoznak, és teljes mértékben megfelelnek a módszertani követelményeknek. A mai napig folyamatosan fenntartott és nemzetközileg is nyilvántartott martonvásári tartamkísérletek parcellaszáma több mint 700, területük mintegy 15 ha. A legfontosabb tartamkísérletek: vetésforgó vs. monokultúra, trágyázási rendszerek összehasonlító vizsgálata, szerves- és műtrágyák kölcsönhatásának és utóhatásának vizsgálata, műtrágyázási dóziskísérletek, valamint polifaktoriális kísérletek.

A növénytermesztés színvonalával (fajta, trágyaadagok, növényvédelem, agro-technika) lépést tartó, de alapelveiben (kezelések, növény- és talajvizsgálatok) maradandó, sok évtizedes hazai tartamkísérletek továbbvitelének a kutatási feladatok között prioritást kell kapniuk. A tartamkísérletek költségesek. A martonvásári tartamkísérletek fenntartását az tette lehetővé, hogy a kukorica- és búzakutatások integrált részét képezték és tesztelő helyet biztosítottak a martonvásári kukorica hibridek és búzafajták agronómiai reakcióinak vizsgálatához.

Az 16.1. ábra mutatja az előrehaladó (progresszív) változást a tartamkísérletek funkciójában és céljában az elmúlt 150 évben és a jövőbeni ismeretlen lehetőségeket. A tág környezeti tartományt képviselő tartamkísérletek döntőek a regionális előre-vetésekben, a mezőgazdasági fenntarthatóságra és a globális klímaváltozás hatásaira vonatkozóan.



16.1. ábra. A tartamkísérletek szerepének időbeni változása (Rasmussen et al., 1998)

16.2. A tartamkísérletek értékelésének módszerei

16.2.1. A tartamkísérletek standard és javított tervezése

Minden tartamkísérlet közös meghatározó jellemvonása, hogy a kezeléseket minden parcellán beállítjuk a kísérlet kezdésekor és ezt követően a kezelt parcellákon ismétlődő méréseket végzünk sok éven keresztül. Ez olyan elrendezést eredményez, mint amely az 16.2. ábrán látható. A kezeléseket randomizálva vannak a parcellákra minden blokkban az 1. évben és ennek az évnek a végén a méréseket az első időre (1-es idő) elvégezzük minden parcellán. A méréseket a 2-es időben elvégezzük a 2-es év végén, a 3-as időbeni méréseket a 3. év végén, és így tovább a vizsgálat egész időtartamában.

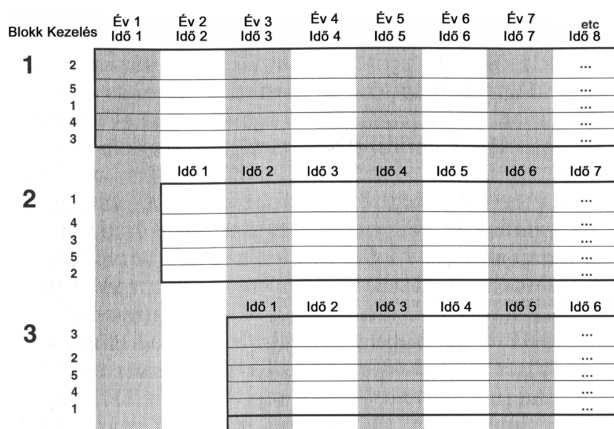
Blokk	Kezelés	Év 1	Év 2	Év 3	Év 4	Év 5	etc.
		Idő 1	Idő 2	Idő 3	Idő 4	Idő 5	etc.
1	2	■					...
	5		■				...
	1			■			...
	4				■		...
	3					■	...
2	1	■					...
	4		■				...
	3			■			...
	5				■		...
	2					■	...
3	3	■					...
	2		■				...
	5			■			...
	4				■		...
	1					■	...

16.2. ábra. Egy tartamkísérlet standard tervezésének vázlatja (Loughin, 2006)

Általánosan elfogadott, hogy a szabadföldi kísérletekben végzett méréseket nem szabályozható környezeti faktorok befolyásolhatják. Ráadásul, ezek a faktorok néhány vizsgált kezelésre eltérően hatnak, mint a többire. Ezáltal egészen valószínű, hogy a mért reakcióban jelen levő bármely fix, ismétlődő idő és idő \times kezelés hatáson kívül, vannak nem szabályozható, random évenkénti fluktuációk és év \times kezelés random hatások (Loughin, 2006). (Az időt használjuk, hogy jelentse azt az időszakot, amely eltelt a kezelés kezdetétől egy kísérleti parcellában, és az évet használjuk, hogy jelentse a naptári évet, amelyben a méréseket végezzük.) Amikor elkezdünk analizálni egy hosszú adatsorozatot, fel kell ismernünk, hogy az idő tényező benne van az adatokban, éspedig random és fix hatásként egyaránt.

A tartamkísérletek tervezésénél és analizésénél általánosan használt módszerek nem veszik megfelelően figyelembe a random hatásokat, amelyek kapcsolatban vannak az évekkel vagy az év \times kezelés interakcióval. Valóban, a tartamkísérletek tervezésének standard módszere – a kezelések kezdete minden parcellán azonos időben – nem különíti el a fix hatásokat, amelyek a kísérleti célokat reprezentálják, a random hatásoktól, amelyek zavaró hatások.

Elérni valódi időbeni ismétlést a szokásos módon – megismételni az egész kísérletet – nem megvalósítható tartamkísérletekkel, melyeknek időtartam sok éve. Egy jó kompromisszum, amely elér ismétlést kevés idő ráfordítással, a kezelések lépcsőzetes kezdése a különböző blokkokban az évek során. Azaz, a blokkokat a kísérleten belül úgy állítjuk be, hogy kezdjük az 1-es blokkal az első évben, a 2-es blokkal a második évben és így tovább, ahány ismétlés van tervezve a kísérletben (16.3. ábra). Ily módon mindegyik ismétlés különböző beállítás-évi, különböző második éves és így tovább környezetnek van kitéve. A különböző blokkokban végzett mérések ugyanabban az évben ki vannak téve az évhez kapcsolódó random környezeti hatásoknak. Fontos különbség a lépcsőzetes kezdetű és a standard kezelés között az, hogy ezek a nem szabályozott ingadozások hatással vannak a különböző időben végzett mérésekre és ezáltal a fix idő hatásoktól elkülönítve mérhető.



16.3. ábra. Egy tartamkísérlet lépcsőzetes kezdetű tervezésének vázlata (Loughin, 2006)

Ezt lépcsőzetes kezdetű elrendezést már többször javasolták a növénytermesztésben és más tudományágakban, ahol tartamkísérleteket végeznek. A növénytermesztés irodalmának áttekintése után úgy tűnik, nem sokan vagy egyáltalán nem alkalmazták.

Az adatok (pl. termés) egy tartamkísérletből minden kezelt parcellán végzett ismételt mérések formájában vannak. Mivel a talajtermékenység és a parcellára jellemző más faktorok a különböző években korrelálnak, ugyanazon a parcellán mért termésekről logikusan nem feltételezhető, hogy független hibákkal rendelkeznek. A reziduumok közötti korrelációkat egyik évről a másikra, *sorozat korrelációnak* (serial correlation) hívjuk. A sorozat korreláció fontosságának becslése a tartamkísérletek analízisének elválaszthatatlan része. Még a tartamkísérletekkel összefüggő egyedülálló probléma (a fix és random hatások keveredése) nélkül is, az időben ismételt mérési adatok egy speciális kihívást jelentenek a statisztikai analízisben (Józsa, 1988).



16.2. kép. „Sanborn Field” tartamkísérlet. Helye: Columbia, Missouri, USA.
Beállítás éve: 1888.

Ugyanazon a parcellán végzett ismételt mérések analízisére számos lehetséges megközelítés történt. *Webster és Payne* (2002) áttekintést ad az ismételt mérési adatok érvényes statisztikai analízisééről. *Loughin* (2006) és *Loughin et al.* (2007) bemutatja a tartamkísérletek javított statisztikai analízisét, felhasználva a standard elrendezést, továbbá a tartamkísérletek javított tervezését és analízisét. A legáltalánosabb módszerek közé tartoznak: (i) összegező statisztikai mutatók, (ii) többváltozós ANOVA és (iii) a korrelációs struktúra modellezése.

16.2.2. Összegező statisztikai mutatók

Az ismételt mérésekből származó adatokat analizálhatjuk először úgy, hogy minden egyes parcellára valamilyen összegező statisztikai mutatót számítunk ki. Például, tételizzük fel, hogy a reakció változó a növény termése. Ekkor az átlagos (vagy az

összes) termést a vizsgált időszakban használjuk fel arra, hogy összehasonlítsuk a különböző kezelésben részesült parcellák általános produktivitását. A termés időbeni változása felhasználható annak meghatározására, hogy a kezelések hasonló trendeket követnek-e. Az ilyen mutatók úgy kezelhetők, mint reakció változók és analizálhatók a hagyományos módszerekkel, mint amilyen a varianciaanalízis, regresszióanalízis, vagy az általánosított lineáris modellek (generalized linear models, GLM). Az összefoglaló mutatók hibájának becslése csupán a kísérletben alkalmazott randomizáláson alapszik; nem szükségesek feltételek vagy ismeretek az ismételt mérések kovariancia struktúrájáról. Jegyezzük meg, hogy ha több összegező mutatót választunk, valószínűtlen, hogy egymástól függetlenek legyenek. Tehát óvatosság indokolt az eredmények interpretálásakor vagy vegyük számba e mutatók többváltozós varianciaanalízisét (MANOVA).

A *többváltozós varianciaanalízis (MANOVA)* egy másik tradicionális módszer, amelyben az évenkénti adatok a többszörös reakcióváltozók. Szemben az ANOVA-val, nem támaszt követelményeket a kovariancia struktúrára, és ezáltal nagyon általános. Ez azt is jelenti, hogy a szignifikáns különbségek feltárásában kevésbé hatékony, mint a többi specifikus módszer.



16.3. kép. „Morrow Plots” tartamkísérlet parcellái (baloldal) és a tartamkísérlet alapítási adatai bronztáblán (jobboldal). Helye: Urbana-Champaign, Illinois, USA. Kísérlet beállítása: 1876.

16.2.3. Évenkénti és kombinált varianciaanalízis

Gyakran az ismételt méréseket úgy analizálják, hogy elkülönítetten analizálják az adatokat minden évben. A varianciaanalízisben először kísérleti kezelésekre gyakorolt hatását a kísérlet típusának megfelelő varianciaanalízissel, évenként vizsgáljuk. Ez érvényes, azonban az elkülönített tesztek a kezeléskülönbségekről csak a megfigyelt időre adnak információt. Nem használhatók arra, hogy becsljük az időben bekövetkezett változásokat. Lényegében az analízis korrekt, de nem különösen informatív. A második lépésben az évek figyelembevételén alapuló, ún. kombinált kéttényezős varianciaanalízissel értékeljük a kísérleti kezelésekre fő hatásait és a köl-

csönhatásokat. Ennek módszerét részletesen tárgyalja *Gomez és Gomez (1984)*, továbbá *Petersen (1994)*.

Már *Cochran (1939)* felismerte, hogy a tartamkísérlet adatok olyanok, mint az osztott parcellás elrendezésből származó adatok. Ha egy tulajdonság ismételt méréseit tekintjük különböző időben, az hasonló a főparcellán végzett ismételt mérésekkel az alparcella faktor különböző szintjein. Feltárta továbbá, hogy a sorozat korrelációk a mérések között, amelyeket viszonylag közeli időpontokban végeztünk, valószínűleg nagyobbak, mint a távolabbi időpontokban végzett mérések között.

Ezáltal ténylegesen nem javasolta az ismételt mérések analízisét úgy, mintha osztott parcellás elrendezésből származnának, mivel az utóbbi értelemszerűen feltételezi, hogy a korreláció minden idő-pár között egyenlő. Az időben megismételt kísérletek értékelési módszerének formai átvételét *Sváb (1981)* is hibásnak tartotta. Ennek ellenére, az ismételt mérési adatok osztott parcellás analízise, amelyben az idő (év) az alparcella faktor (és természetesen nem randomizált), virágkorát éli (*Steel és Torrie, 1980; Petersen, 1994*). Ez a megközelítés jogosan kritizálható, főleg azon az alapon, hogy a randomizációs teóriával nem igazolható (pl. *Yates, 1954*). Népszerűsége abból ered, hogy az analízis viszonylag egyszerű. A split-plot modell nagyon egyszerű korrelációs struktúrát tartalmaz az ismételt mérésekre, ezért használata a tartamkísérletek adatsorozatainak analízisére kevésbé javasolható.



16.4. kép. Szántás nélküli (no-till) szója-búza vetésforgó tartamkísérlet. Helye: INTA Experiment Station, Córdoba, Argentina. Beállítási éve: 1975.

16.2.4. Kumulatív terméselemzés

Hazai viszonylatban nagyon népszerű és általánosan elfogadott *Sváb (1981)* által a tartamkísérletekre kidolgozott kumulatív terméselemzési módszer. A kumulatív módszer lényege az, hogy (a) évenként képezzük a kezelések és báziskezelések közötti terméskülönbséget, (b) majd ezeket a különbségeket ugyancsak évenként kumulál-

jük. Értékelési eredményünk helyes értelmezhetősége döntő mértékben attól függ, hogy melyik kezelést választottuk báziskezelésnek (standardnak). Az a jó báziskezelés, amelyik körülbelül úgy reagál az évenkénti környezethatásokra, mint a kezelések zöme. Ne válasszunk labilis, érzékeny báziskezelést, mert azáltal, hogy a báziskezeléshez viszonyítunk, a báziskezelés labilitása a kezeléskülönbségekben, tehát a többi kezelés hatását jellemző különbségekben fog mutatkozni. A nem kumulált különbségek a kezelések egyenkénti és évenkénti hatását fejezik ki a báziskezeléshez képest. A kumulált terméskülönbségek azt mutatják meg, hogy a t -edik évben mennyi valamely kezelés összes terméskülönbsége a báziskezeléshez viszonyítva.

Nagyon szemléletes a tartamkísérlet kumulált terméskülönbségeinek grafikus szemléltetése. Az ábra vízszintes tengelye az időt jelzi, beosztását a tartamkísérlet szakaszai adják, általában naptári időbeosztásban. A függőleges tengelyen a báziskezeléstől való eltéréseket fejezzük ki. Ennek megfelelően pozitív és negatív is lehet az előjelük. Az ábrán a báziskezelés vonala a vízszintes tengellyel azonos. A kezelések kumulált hatását a vízszintes tengelytől való pozitív és negatív távolságok fejezik ki.



16.5. kép. Fejtrágyázási legelő tartamkísérlet. Helye: Rutherglen Centre, Victoria, Australia. Beállítás éve: 1912.

A kumulált terméskülönbségek ábráján, ha valamely kezelés vonala (1) párhuzamos a bázist jelentő főteneggellyel, akkor a kezelés nem különbözik a báziskezeléstől az adott évben, (2) ha egyenes vonalúan növekszik vagy csökken, akkor a kezeléshatás többlete vagy hiánya évről évre konstans, (3) ha a vonala hajlik, akkor a hatáskülönbség a bázishoz viszonyítva évről évre változik. A báziskezeléshez való viszonyításra azért van szükség, hogy a termés mennyiségét befolyásoló évhatást kiküszöböljük. A kezelés \times év kölcsönhatás azonban nincs kiküszöbölve.

Rowell és Walters (1976) javasolta a polinomiális egyenletek illesztését minden egyes parcella időbeni mérési adataihoz, és ezt követően az egyenletek becsült koefficienszeinek reakcióváltozóként történő felhasználását. Valószínűtlen ennek megfelelése a

koefficiensek közötti korreláció miatt. Mindazonáltal összefoglaló statisztikai mutatók használata elkerüli a kísérleti egységen (parcellán) belüli korrelációk modellezésének problémáját. Valóban, a polinomokra egy alternatíva a *random koefficiens modell illesztése*, amely becsüli mind a koefficienseket, mind pedig a köztük levő korrelációt. Mindkét megközelítés azonban feltételezi, hogy az időbeni reakció megfelelően modellezhető polinomokkal.

16.2.5. Stabilitásanalízis

A termesztés fenntarthatósága csak meghatározott időperiódus relációjában mérhető, és a rendszer fenntarthatóságára ható fontos trendek nyilvánvalóvá válnak az első 20–40 évben (Barnett *et al.* 1995). A termesztés fenntarthatóságának fontos mutatója a stabilitás. A termésstabilitás időbeni mérése magában foglal legalább három komponenst: (1) átlagos termésszint, (2) a termés variabilitása és (3) a termés összefüggése a helyi környezettel (Mead *et al.* 1986). Az átlag (fix, azaz szisztematikus hatás) és a variancia (random elem) a két fő komponens, amely leírja egy termesztési rendszer reakció mintázatát. Berzsényi (2009b) a kísérleti kezelések stabilitását egyváltozós (variancia és regresszió mutatók) és többváltozós (AMMI modell) módszerével vizsgálta.

Piepho (1998) a stabilitás mutatókat a fix és random komponenseket egyaránt tartalmazó modell keretében tárgyalja. Az alapmodell a következő formában írható fel:

$$y_{ij} = \mu_i + e_{ij} \quad (1)$$

ahol y_{ij} az i -edik rendszer teljesítménye a j -edik környezetben, μ_i az átlag (a várható érték), az i -edik rendszer szisztematikus hatása és e_{ij} a random eltérés az i -edik rendszer átlagától a j -edik környezetben.

A stabilitásanalízis különböző megközelítései különböznek abban, hogy milyen struktúrában írják le a fix és a random komponenseket. Mindegyik rendszernek megvan a saját varianciája. Az i -edik rendszer varianciája kifejezhető:

$$\text{Var}(y_{ij}) = \text{Var}(e_{ij}) = \sigma_{ii} \quad (2)$$

A σ_{ii} varianciát környezeti varianciának hívjuk. Minél nagyobb egy rendszer varianciája, annál kisebb a stabilitása. A variancia gyakran kapcsolatban van az átlaggal, ezért előnyös a variancia négyzetgyökét standardizálni az átlaggal, amely a variációs koefficiens (CV_{*i*}) eredményezi. Feltételezhető, hogy van bizonyos kovariancia (rövidítése: Cov) a különböző rendszerek között ugyanabban a környezetben. Az i és i' rendszer között a kovariancia következőképpen írható fel:

$$\text{Cov}(y_{ij}, y_{i'j}) = \sigma_{ii'} \quad (3)$$

A σ_{ii} és $\sigma_{ii'}$ (2 és 3. egyenlet) teljesen meghatározza az adatok variancia-kovariancia struktúráját. A stabilitás mutatók különböznek abban, hogyan parametrizálják σ_{ii} és $\sigma_{ii'}$ -t.



16.6. kép. „Blevins” talajművelési tartamkísérlet. Helye: Lexington, Kentucky, USA.
Beállítás éve: 1970.

A varianciaanalízisben (ANOVA) az e_{ij} random komponenst felosztjuk környezeti fő hatásra és interakcióra:

$$y_{ij} = \mu + u_j + e_{ij} \quad (4)$$

ahol u_j : random környezeti fő hatás és e_{ij} : random reziduális hatás, amely magában foglalja a rendszer \times környezet interakciót és a hibát egyaránt. A varianciaanalízis feltételezése az, hogy az u_j és e_{ij} random hatások független eloszlásúak és $\text{Var}(u_j) = \sigma_u^2$ és $\text{Var}(e_{ij}) = \sigma^2$.

Egy sokkal általánosabb feltételezés az, hogy minden rendszernek megvan a saját varianciája az e_{ij} komponensre, azaz $\text{Var}(e_{ij}) = \sigma_i^2$. Ezzel a feltétellel, a kovariancia a megfigyelések között egy környezetben ugyanaz minden megfigyelés-párra (Piepho 1998):

$$\text{Cov}(y_{ij}, y_{i'j}) = \sigma_{ii'} = \sigma_u^2 \quad (5)$$

azonban egy megfigyelés varianciája változik a rendszerek között a következőképpen:

$$\text{Var}(y_{ij}) = \sigma_{ii} = \sigma_u^2 + \sigma_i^2 \quad (6)$$

A σ_i^2 variancia megfelel Shukla (1972) stabilitás varianciájának és nagyon szoros kapcsolatban van Wricke (1962) ökovalencia indexével (Lin és Binns, 1988). Egy alacsony stabilitás varianciával rendelkező rendszer stabilnak tekinthető.

A stabilitásanalízis gyakori megközelítése az, hogy kiszámítjuk a regressziót az y_{ij} rendszer teljesítménye és egy környezeti index között, mely utóbbit egy környezetben minden megfigyelés (kezelés) átlagaként számítunk ki. Az index a környezet produktivitás mutatójának tekinthető. A viszonylag nagy regressziós meredekséggel rendelkező rendszerek átlagon felüli reakciót mutatnak a környezeti feltételek javulására, amelyet a környezeti index jelez. A regressziós módszert először Yates és Cochran

(1938) javasolta, a módszer további kidolgozása *Finlay és Wilkinson (1963)*, *Eberhart és Russell (1966)* nevéhez fűződik.

A regressziós megközelítés a stabilitásanalízis egyik legszélesebb körben használt módszere. A regressziós modell a következőképpen írható fel:

$$y_{ij} = \mu_i + \beta_i u_j + d_{ij} \quad (7)$$

ahol β_i : az i -edik rendszer regressziós koefficiense, u_j : a j -edik környezet hatása és d_{ij} : a random eltérés a regressziós egyenestől. Hangsúlyozni szükséges, hogy már nem a környezet fő hatása, hanem egy látens változó, amely nem mérhető közvetlenül, de amelyre a rendszerek feltételezhetően lineáris reakciót adnak, különbségekkel a reakció erősségében, amelyet a β_i érzékenységi paraméterek jeleznek. A modell feltételezése szerint u_j és d_{ij} random hatások függetlenek, $\text{Var}(u_j) = \sigma_u^2$ és $\text{Var}(d_{ij}) = \sigma_{di}^2$ varianciákkal (*Piepho 1998*).

A stabilitásanalízisben a termesztési rendszereket gyakran stabilnak tekintik, amikor bizonyos variancia komponens (σ_{iv} , σ_i^2 , σ_{di}^2) kicsi. Ez magában foglalja, hogy a variabilitás nem kívánatos, függetlenül a termés szintjétől. Gyakran azonban nem a variabilitás önmagában a fő gond, hanem az alacsony termés bekövetkezésének kockázata (*Mead et al. 1986*). A kockázat egyaránt függ az átlagtól és a varianciától. Számos megközelítés van, amely számba veszi egy rendszernek az átlagát és varianciáját (*Kang és Magari 1995*). *Kang (1993)* kifejlesztett egy ún. termésstabilitás (YS_i) mutatót, amely a termést és σ_i^2 -et egyetlen szelekciós kritériumba egyesítette és felhasználható a termesztési rendszerek termésszintjének és stabilitásának egyidejű értékelésére.

A stabilitásanalízist korábban ritkán alkalmazták a növénynevelés területén kívül. *Hildebrand (1984)* és *Raun et al. (1993)* felhasználták a stabilitásanalízist tartamkísérletben a trágyázási kezelések értékelésére. *Berzsenyi és Gyórfy (1995, 1996, 1997)* a stabilitásanalízis variancia- és regressziós módszerével tanulmányozta a különböző termesztési tényezők, valamint a vetésforgó és a trágyázás hatását a kukorica, illetve a búza termésstabilitására tartamkísérletekben. *Nagy et al. (2003)* a stabilitásanalízis regressziós módszerével vizsgálta a műtrágyázás és a talajművelés hatását a kukorica termésstabilitására.

Az AMMI (additív fő hatás és multiplikatív interakció) modell integrálja a varianciaanalízist és a főkomponensanalízist (*Crossa, 1990*). Az AMMI analízis első részében a varianciaanalízis az összes variációt három ortogonális forrásra bontja fel: genotípus (G), környezet (E) és genotípus \times környezet interakció (G \times E). Az AMMI analízis második részében a főkomponensanalízis (PCA) a G \times E interakciót több ortogonális főkomponens-változóra (PCA tengelyre) bontja fel.

16.2.6. A tartamkísérletek elemzése vegyes (mixed) modellekkel

Az ismételt mérések analízisének figyelembe kell venni a korrelációt a mérések között ugyanazon a kísérleti parcellán. Több különböző típusú modell van, amelyek a korrelációs struktúrát leírják (pl. az autoregresszív, compound symmetry, antedependence stb.). Ezek a korrelációs struktúrák specifikálhatók a *mixed modellben*, amely a fix hatásokat és a random hatásokat egyaránt modellezi (*Payne, 2000*). Az utóbbi időben megnőtt a népszerűsége az ismételt mérésekkel kapcsolatos korrelációs struktúra modellezésének a mixed modellek keretében. Három fő előnye: (i) a korrelációs modellek jelenléte, (ii) számos hiba tényező beépítésének lehetősége és (iii) az inkomplett

blokk elrendezések adatainak analízise. Számos módszer van a variancia komponensek becslésére. A Residual vagy Restricted Maximum Likelihood (REML) a standard módszer a legtöbb mixed modell programban. Az REML alapú mixed modell eljárások rendszerint nagyszámú statisztikai mutatót (pl. Wald teszt) tartalmaznak, amelyek kevésbé ismertek azoknak, akik a klasszikus ANOVA módszert használják, jóllehet több a hasonlóság, mint az eltérés (Piepho et al., 2003). Napjainkban modern statisztikai software-k állnak rendelkezésre, így az alapelvek megértését követően várhatóan széleskörű lesz a mixed modellek használata a hazai növénytermesztési kutatásban, különösen a tartamkísérletek adatainak analízisében.

A mixed modellek kiterjeszthetők, hogy figyelembe vegyék a nem normális eloszlású adatokat. A kiterjesztés általánosított lineáris mixed modellként (generalized linear mixed model) ismert és rendkívül fontos problémás adatoknál, amilyenek a megszámlolt adatok (pl. gyomok és rovarok) és a százalékban kifejezett méréseknél (pl. betegség előfordulása, gyomborítottság, kelési arány stb.) (Piepho et al., 2003). A statisztikusok a tartamkísérletek többéves adatsorozatainak analízisére elsősorban a mixed modellt javasolják és tartják elfogadhatónak.

16.3. A martonvásári tartamkísérletek értékelésének eredményei

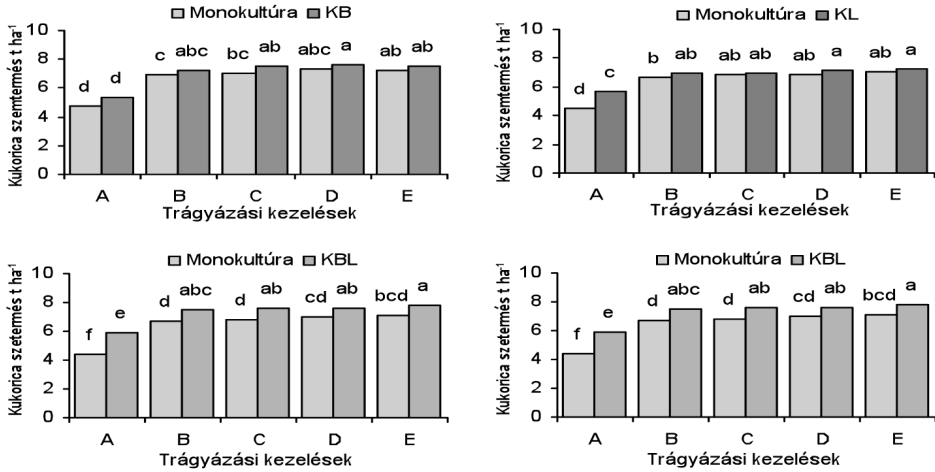
16.3.1. Kukorica és búza vetésforgó kísérlet eredményei

A vetésforgó kísérletet Győrffy Béla és munkatársai 1961-ben állították be az intézet kísérleti területén kéttényezős, osztott parcellás elrendezésben, négy ismétlésben. A főparcellákat a növényi sorrendek, az alparcellákat a trágyakezelések képezik. A főparcella 7 növényi sorrendet foglal magában, így kukorica és búza monokultúrát, dikultúrát, trikultúrát és Norfolkli típusú forgót. A kísérlet 7 kezelése az alábbi: 1. Kukorica monokultúra, 2. Búza monokultúra, 3. 3 év lucerna – 5 év kukorica (KL), 4. 3 év lucerna – 5 év búza (BL), 5. 2 év búza – 2 év kukorica (KB), 6. 3 év lucerna – 3 év kukorica – 2 év búza (KBL), 7. Kukorica – tavaszi árpa – borsó – búza (NF). A kukorica, illetve a búza részaránya a vetésforgótól függően 25, 37.5, 50, 62.5 és 100%.

A kísérlet alparcellái 5 eltérő trágyázási rendszert képviselnek: A: Kontroll, trágyázás nélkül, B: 60 t ha⁻¹ istállótrágya 4-évenként + NPK kiegészítés, C: 5 t ha⁻¹ szalma, illetve 7 t ha⁻¹ kukoricaszár évente + NPK kiegészítés, D: a növény által felvett NPK műtrágya, E: Felvett NPK, 15 t ha⁻¹ kukorica- és 10.5 t ha⁻¹ búzaterméshez. Az összes parcellaszám 140.

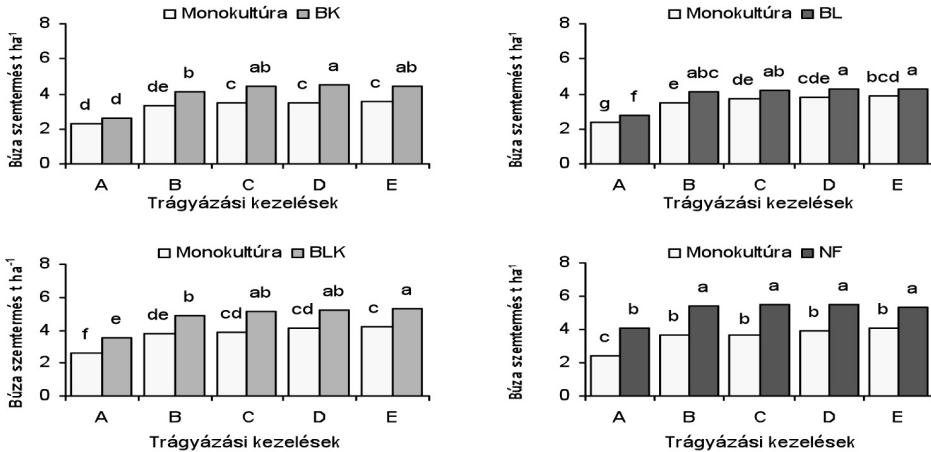
A kísérletben összehasonlítjuk a különböző rotációk hatásait a monokultúrával és egymással. A különböző növényfajok képezik a kezeléseket és a kísérlet szerkezetéből adódóan az összehasonlítandó parcellák nem mindig ugyanazokat a növényeket tartalmazták ugyanabban az évben. Következésképpen a kísérlet több évtizedes fenntartása után van mód arra, hogy a kezeléseket megfelelő számú évben összehasonlíthassuk.

A kukorica és búza termése monokultúrában minden esetben alacsonyabb volt, mint vetésforgóban, másrészt a terméskülönbség monokultúrában nagyobb volt búzában, mint kukoricában (16.4.–16.5. ábra). A vetésforgó termésmenvelő hatása for-



16.4. ábra. A vetésforgó és a trágyázás hatása a kukorica szemtermésére, összehasonlítva a monokultúrával (1961–2000)

KB: kukorica–búza; KL: kukorica–lucerna; KBL: kukorica–búza–lucerna; NF: norfolki típusú; A: kontroll, trágyázás nélkül; B: 60 t/ha istállótrágya + NPK; C: 5 t/ha szalma, illetve 7 t/ha kukoricaszár + NPK; D: növény által felvett NPK; E: NPK műtrágya szükséglet 15 t/ha kukorica- és 10.5 t/ha búzaterméshez. A vetésforgón és trágyázáson belül az azonos betűk nem szignifikáns hatásokat jelölnek P<0.05 szinten a Duncan teszt alapján.



16.5. ábra. A vetésforgó és a trágyázás hatása a búza szemtermésére, összehasonlítva a monokultúrával (1961–2000)

BK: búza-kukorica; BL: búza-lucerna; BLK: búza-lucerna-kukorica; NF: norfolki vetésforgó. Az ábra további jelölései megegyeznek az 1. ábra jelöléseivel.

dított arányban volt a kukorica, illetve a búza részarányával a vetésforgóban. A trágyakezelések (A–E) átlagában legnagyobb volt a termésmenvelő hatás a norfolki típusú forgóban (0.904 t ha⁻¹ kukoricában és 1.664 t ha⁻¹ búzában). Sorrendben következett a lucerna – kukorica – búza trikulturúra (0.853 t ha⁻¹ kukoricában és 1.223 t ha⁻¹ búzában), a búza kukorica dikulturúra (0.490 t ha⁻¹ kukoricában és 0.732 t ha⁻¹ búzában), lucerna – kukorica (0.376 t ha⁻¹) és a lucerna – búza (0.471 t ha⁻¹) forgó. Műtrágyázás nélkül a vetésforgó termés-növelő hatása (t ha⁻¹) szignifikánsan nagyobb volt (kukoricában: KB: 0.715, KL: 1.254, KBL: 1.401, NF: 1.357; búzában: BK: 0.375, BL: 0.446, KBL: 0.923, NF: 1.666). Az istállótrágyázás és a növényi maradványok (kukoricaszár, búzaszalma) visszajuttatása, NPK kiegészítéssel egyaránt hatékony módszerei a kukorica és a búza trágyázásának. Az istállótrágyázás egyúttal fokozza a termés-stabilitást is. Szignifikánsan nagyobb termést kaptunk magas NPK műtrágya dózissnál, különösen, ha a búza, illetve a kukorica részaránya a vetésforgóban $\geq 50\%$. Megállapítottuk, hogy a búza vetésforgók termésmenvelő hatását nem módosította a trágyázás. Ezzel szemben, kukorica vetésforgókban a trágyázás közel 50%-kal csökkentette a rotációs hatást (Berzsenyi et al., 2000; Berzsenyi és Dang, 2002a).

16.1. táblázat. Kukorica és búza monokultúra vs. Norfolki vetésforgó termésmenvelési paraméterei a különböző trágyázási kezelésekben (1961-2000)

Trágyázási kezelések	Termés t ha ⁻¹	CV%	σ^2	YS	IPCA	Termés t ha ⁻¹	CV%	σ^2	YS	IPCA
	Kukorica monokultúra					Búza monokultúra				
A	4.715	15.5	5.48 ^{***}		-1.59	2.441	22.6	1.20 ^{***}		1.00
B	7.092	5.0	0.20 ^{NS}	+	0.27	3.702	11.8	0.19 ^{NS}	+	-0.18
C	7.145	9.2	1.48 ^{***}		0.62	3.728	14.8	0.45 ^{***}		0.14
D	7.220	6.1	0.14 ^{NS}	+	0.29	3.934	7.9	0.21 ^{NS}	+	-0.36
E	7.200	8.2	1.30 ^{***}		0.41	4.075	6.7	0.45 ^{***}	+	-0.61
SzD _{5%}	0.237					0.141				
	Norfolki vetésforgó					Norfolki vetésforgó				
A	6.072	8.8	2.50 ^{***}		-1.29	4.055	13.4	0.79 ^{***}		-0.96
B	8.199	5.4	0.33 ^{NS}	+	0.18	5.360	4.5	0.13 ^{NS}	+	0.46
C	8.208	5.8	0.40 ^{NS}	+	0.28	5.470	6.6	0.13 ^{NS}	+	0.40
D	7.821	3.7	0.08 ^{NS}	+	0.21	5.487	8.2	0.19 ^{NS}	+	0.17
E	7.576	6.0	0.72 ^{NS}		0.63	5.338	7.8	0.32 ^{NS}	+	-0.08
SzD _{5%}	0.260					0.159				

CV%: variációs koefficiens, σ^2 : stabilitás variancia, YS: termésmenvelés, +: kiválasztott kezelések, IPCA: interakció főkomponens értékek. ^{***} Szignifikáns $P \leq 0.01$ szinten, ^{NS} Nem szignifikáns $P > 0.05$ szinten.

A stabilitásanalízis különböző módszerei kimutatták, hogy a vetésforgók stabilitása szignifikánsan különbözik a monokultúrákétól. A variancia paraméterek (CV%, σ^2 , YS) és az interakció főkomponens értékek (IPCA) tendenciájukban magasabbak voltak kukorica és búza monokultúrákban, mint vetésforgókban (16.1. táblázat). A stabilitásanalízis megfelelő módszernek bizonyult a kezelés \times környezet interakció értel-

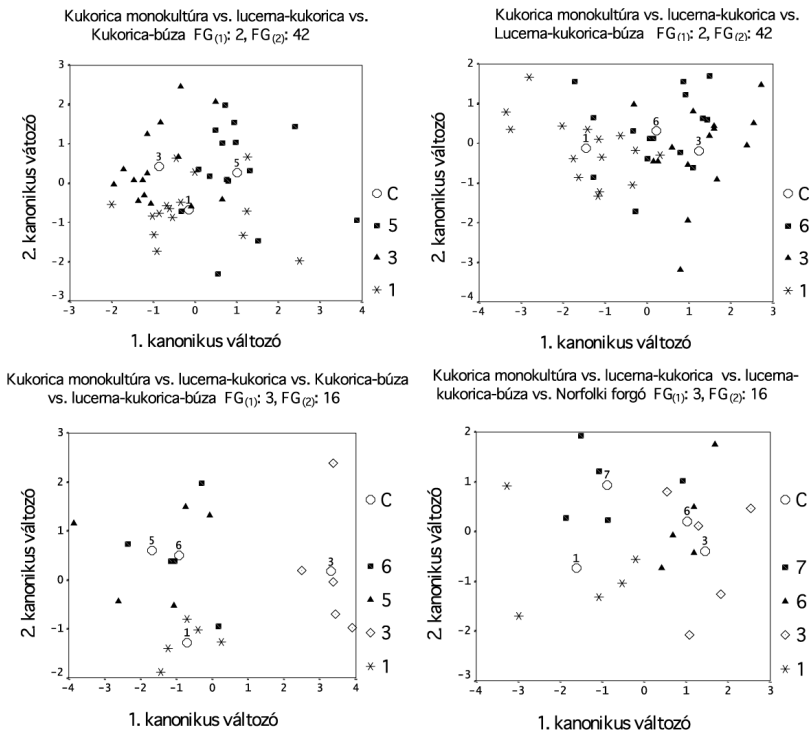
mezéséhez és a kezelések termésstabilitásának becsléséhez vetésforgó tartamkísérletekben (Berzsenyi *et al.*, 2000).

Búza és kukorica vetésforgók elkülönítése diszkriminanciaanalízissel

A kutatás célja volt a búza és kukorica vetésforgók vs. monokultúra többváltozós elkülönítése diszkriminanciaanalízissel. Minden vetésforgót és monokultúrát az öt trágyázási kezelésre (A–E) adott termésreakcióval, mint kvantitatív változóval jellemeztük (Berzsenyi és Dang, 2002a).

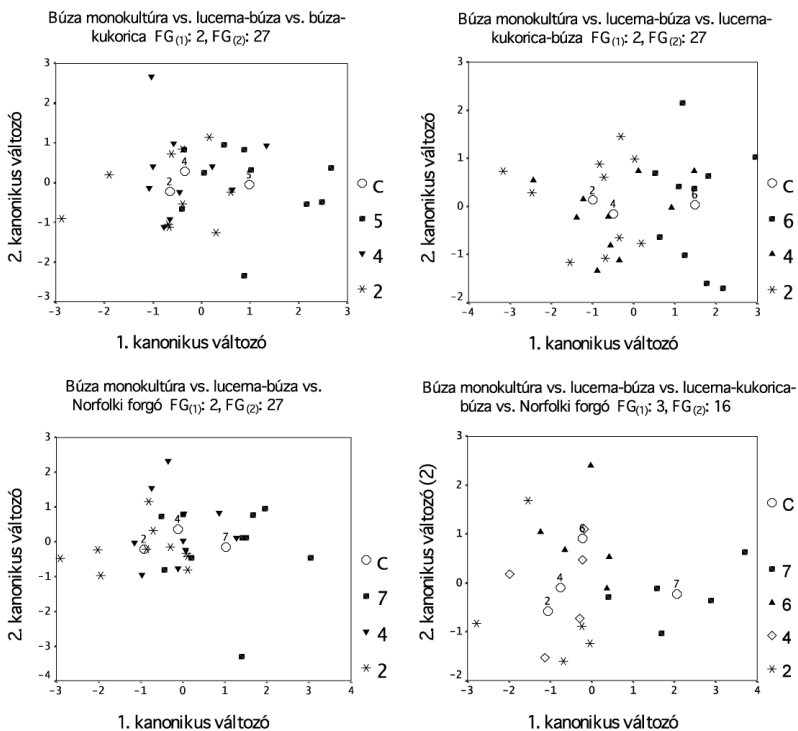
A kétszempontos és többszempontos összehasonlításban a kukorica vetésforgó és monokultúra, illetve a búza vetésforgó és monokultúra csoportcentroidok szignifikánsan különböztek, az elkülönítésért felelős változók azonban eltérőek voltak. A csoportcentroidok közötti távolság arányos volt a vetésforgók rotációs hatásával, legnagyobb volt a monokultúra és a trikkultúra, illetve a norfolki forgó között.

Megállapítottuk, hogy a kukorica vetésforgókban az eltérés a pillangósvirágú előveteményt (lucerna, borsó) magában foglaló vetésforgók és a monokultúra között elsősorban a trágyázás hatásának tulajdonítható. Feltételezhető, hogy a kukorica-búza dikultúra és a kukorica monokultúra csoport-centroidok szignifikáns elkülönítésében a trágyázási változókon kívül más agronómiai és ökológiai tényezők voltak jelentős hatással (16.6. ábra).



16.6. ábra. Kukorica vetésforgók vs. monokultúra többváltozós elkülönítése diszkriminanciaanalízissel. Jelölések: C: csoportcentroidok, 1: kukorica monokultúra, 3: lucerna-kukorica, 5: kukorica-búza, 6: lucerna-kukorica-búza, 7: norfolki forgó

Búza vetésforgókban (16.7. ábra) az ún. rotációs hatás nem elsősorban a trágyázás hatásával magyarázható, azért más faktorok, folyamatok és mechanizmusok is felelősek. A diszkriminanciaanalízis ún. stepwise módszere különösen hasznosnak bizonyult a vetésforgó- és monokultúra-csoportok elkülönítésében legfontosabb prediktív változók kiválasztásában és a páronkénti csoportcentroidok közötti távolság szignifikanciájának meghatározásában.



16.7. ábra. Búza vetésforgók vs. monokultúra többváltozós elkülönítése diszkriminanciaanalízissel. Jelölések: C: csoportcentroidok, 2: búza monokultúra, 4: lucerna-búza, 5: búza-kukorica, 6: lucerna-kukorica-búza, 7: norfolki forgó.

16.3.2. Az istállótrágya és műtrágya hatása a kukorica termésére tartamkísérletekben

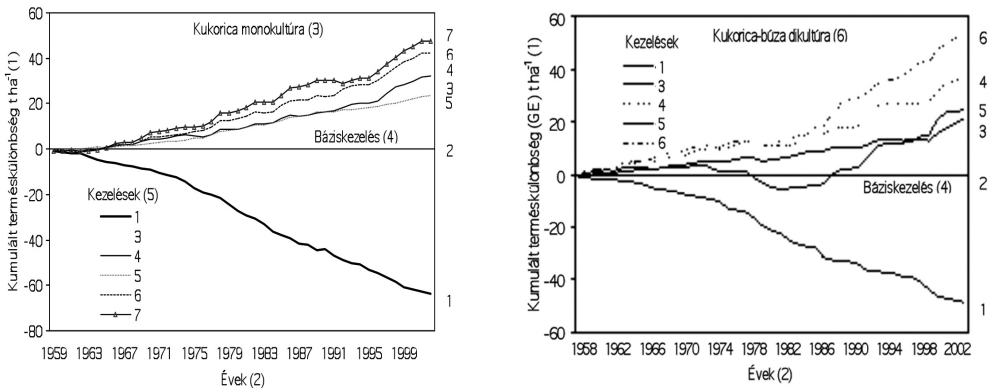
A trágyázási kísérleteket kukorica-búza dikultúrában 1958-ban, kukorica monokultúrában 1959-ben, egymás mellett állította be Györfly Béla, az MTA Mezőgazdasági Kutatóintézetének kísérleti területén, Martonvásáron. A kukorica-búza dikultúra kísérletben a növényi sorrend 2 év kukorica, 2 év búza. Mindkét kísérletet latin négyzet elrendezésben állították be, a monokultúra kísérletet 7 kezeléssel, 80 m²-es parcellákon, a dikultúra kísérletet 6 kezeléssel, 56 m²-es parcellákon.

A kukorica monokultúra kísérlet kezeléseit az alábbiak (2–7. kezelés 4 évenként): 1. Kontroll, trágyázás nélkül, 2. 35 t ha⁻¹ istállótrágya, 3. 17.5 t ha⁻¹ istállótrágya + NPK

műtrágya kiegészítés ($N_{1/2}P_{1/2}K_{1/2}$), 4. 35 t ha⁻¹ istállótrágya hatóanyagának megfelelő mennyiségben NPK műtrágya ($N_1P_1K_1$), 5. 70 t ha⁻¹ istállótrágya, 6. 35 t ha⁻¹ istállótrágya + NPK műtrágya kiegészítés ($N_1P_1K_1$), 7. 70 t ha⁻¹ istállótrágya hatóanyagának megfelelő mennyiségben NPK műtrágya ($N_2P_2K_2$).

A kukorica-búza *dikultúra kísérlet kezelése*i az alábbiak (2-5. kezelés 4 évenként): 1. Kontroll, trágyázás nélkül, 2. 35 t ha⁻¹ istállótrágya, 3. 17.5 t ha⁻¹ istállótrágya + NPK műtrágya kiegészítés ($N_{1/2}P_{1/2}K_{1/2}$), 4. 35 t ha⁻¹ istállótrágya hatóanyagának megfelelő mennyiségben NPK műtrágya ($N_1P_1K_1$), 5. 35 t ha⁻¹ istállótrágya N-tartalmának megfelelő N-műtrágya (N_1) 1981-ig, 1982-től $N_2P_2K_2$, 6. 35 t ha⁻¹ istállótrágyában levő N-nek kétszeres mennyisége, a P_2O_5 és K_2O tartalmának fele ($N_2P_{1/2}K_{1/2}$) 1981-ig, 1982-től $N_{160}P_{320}K_{320}$ évente.

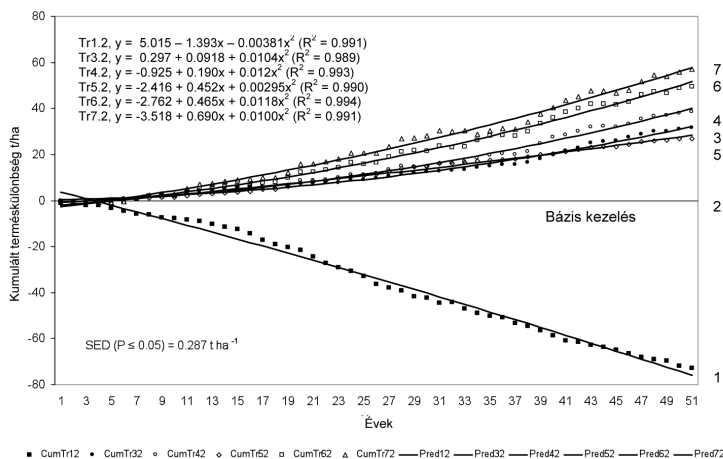
Az istállótrágya és a műtrágya tartamhatását kukorica monokultúrában és kukorica-búza dikultúrában a báziskezeléshez (35 t ha⁻¹ istállótrágya négyévenként) viszonyított kumulált terméskülönbségek alapján a 16.8. ábra szemlélteti. A kukorica termése kukorica-búza dikultúrában, mindegyik összehasonlítható kezelésben és évben, szignifikánsan nagyobb volt, mint monokultúrában (6.221 vs. 5.014 t ha⁻¹). Szignifikánsan a legnagyobb termést azokban a kezelésekben kaptuk, amelyekben az istállótrágya hatóanyagtartalmát fele arányban, illetve teljes egészében NPK műtrágya formájában juttattuk ki (Berzsényi és Gyórfly, 1997; Berzsényi és Dang, 2004).



16.8. ábra. A trágyázási kezelések kumulatív hatása a kukorica termésére monokultúrában és dikultúrában

Kezelések:: 1: kontroll, 2: 35 t ha⁻¹# (báziskezelés), 3: 17.5 t ha⁻¹# + $N_{1/2}P_{1/2}K_{1/2}$, 4: $N_1P_1K_1$, 5: 70 t ha⁻¹#, 6: 35 t ha⁻¹# + $N_1P_1K_1$, 7: $N_2P_2K_2$. (2.–7. kezelés 4 évenként). Kukorica-búza dikultúrában: 1: kontroll, 2: 35 t ha⁻¹# (báziskezelés), 3: 17.5 t ha⁻¹# + $N_{1/2}P_{1/2}K_{1/2}$, 4: $N_1P_1K_1$, 5: N_1 , 1982-től $N_2P_2K_2$, (2.–5. kezelés 4 évenként), 6: $N_2P_{1/2}K_{1/2}$, 1982-től $N_{160}P_{320}K_{320}$ évente. (7)

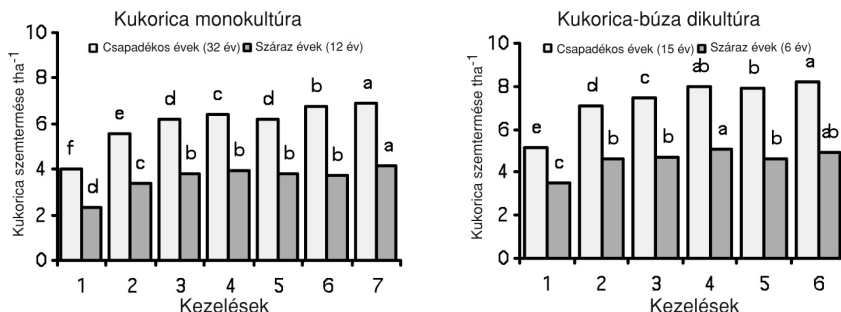
Az elemzés következő szakaszában a kumulált terméskülönbségek statisztikai analíziséhez az REML (linear mixed model) random coefficient regression módszert választottuk (Payne, 2000). A módszer egyidejűleg modellezi az egyes kezelésekre adott reakciók időbeni változását lineáris vagy másodfokú, illetve magasabb fokú polinom függvény illesztésével. Az analízis megadja az SED és a becsült LSD értéket a vizsgált időszakokra, illetve minden előre jelzett évre (16.9. ábra).



16.9. ábra. Az istállótrágya és a műtrágya termésmenvelő hatásának összehasonlító értékelése kukorica monokultúrában a kumulált terméselemzés módszerével (1959–2009)

Kukorica monokultúrában a kumulatív terméskülönbségek statisztikai analízisekor az REML analízis erősen szignifikáns ($P < 0.001$) kezelés, idő (év), illetve kezelés \times idő kölcsönhatásokat mutatott ki. A kezelések időbeni változását másodfokú modell írta le. Az LSD ($P \leq 0.05$) értéke a kumulált terméskülönbségekre és a vizsgált időszakra $0.564 t ha^{-1}$ volt.

A trágyázási kezelések hatását a kukorica termésére száraz és csapadékos években, kukorica monokultúrában és kukorica-búza dikultúrában a 16.10. ábra szemlélteti. Monokultúrában, a vizsgált kezelések átlagában a kukorica szemtermése száraz években $3.586 t ha^{-1}$, csapadékos években $6.010 t ha^{-1}$ volt, vagyis kedvező évjáratban a termésmenvekedés $2.424 t ha^{-1}$ volt. Dikultúrában, a kezelések átlagában a kukorica szemtermése száraz években $4.561 t ha^{-1}$, csapadékos években $7.306 t ha^{-1}$ volt, vagyis kedvező évjáratban a termés $2.745 t ha^{-1}$ -al volt nagyobb. Az adatokból megállapítható a dikultúra magasabb termésszintje, az évjárat-hatás azonban monokultúrában és dikultúrában hasonló nagyságrendű volt (16.10. ábra).



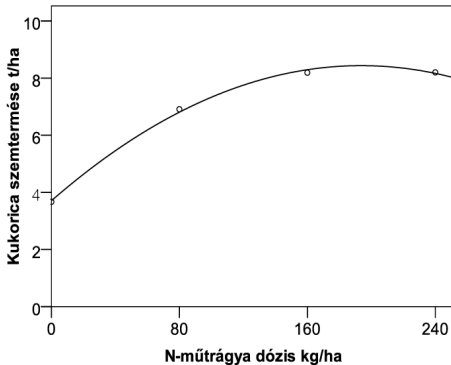
16.10. ábra. Az istállótrágya és a műtrágya hatása a kukorica termésére száraz és csapadékos években, kukorica monokultúrában (44 év) és kukorica-búza dikultúrában (21 év) Egy növényi sorrenden és évjáraton belül az azonos betűjelölést tartalmazó kezelések szignifikánsan nem különböznek egymástól a Duncan-teszt alapján.

16.3.3. Kukorica hibridek N-műtrágya reakciójának vizsgálata tartamkísérletekben

A martonvásári kukorica hibridek N-műtrágya reakcióját két eltérő környezetben, 50 éves kukorica monokultúra tartamkísérletben (stressz-környezet) és norfolki típusú vetésforgó kísérletben (optimális környezet) vizsgáltuk. Mindkét kísérletben különböző N-dózisokat állítottunk be, azonos P- és K-ellátottságnál. A kukorica monokultúra kísérletben a N-műtrágya dózisa a következő volt (kg ha^{-1}): 0, 80, 160 és 240 (jelölésük a továbbiakban: N_0 , N_{80} , N_{160} és N_{240}). A P- és K-műtrágya dózisa azonos volt (160 kg ha^{-1}). Vetésforgóban (kukorica, tavaszi árpa, borsó, őszi búza) a kukorica hibridek N-műtrágya reakcióját 0 és 280 kg ha^{-1} N-dózis tartományban, 40 kg/ha kezelésenkénti különbséggel vizsgáltuk. A P- és K-műtrágya mennyisége azonos (120 kg ha^{-1}) volt. 1995-2007 között 13 év adatai alapján mindkét kísérletbe beállított 5 kukorica hibrid N-műtrágya reakcióját 4 azonos N-műtrágyaszinten hasonlítottuk össze.

Optimális N-dózis meghatározása tartamkísérletekben

A kukorica hibridek N-műtrágya reakciójának jellemzésére először a 40 év termésadatainak átlagos értékeihez másodfokú függvényt illesztettünk. A függvény jól illeszkedett az adatokhoz. A probléma az, hogy úgy gondolhatjuk, a kukorica N-műtrágya reakciója olyan, mint amit a 16.11. ábra szemléltet a 40 év átlagadatai alapján.

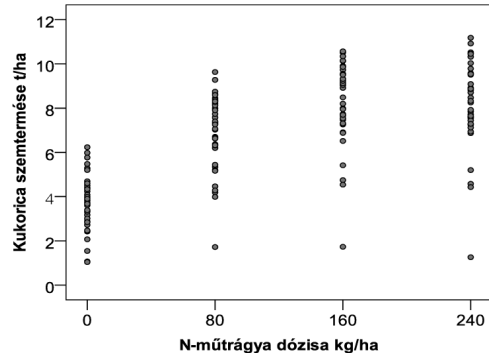


16.11. ábra. Kukorica N-műtrágya reakciója

(1970–2009):

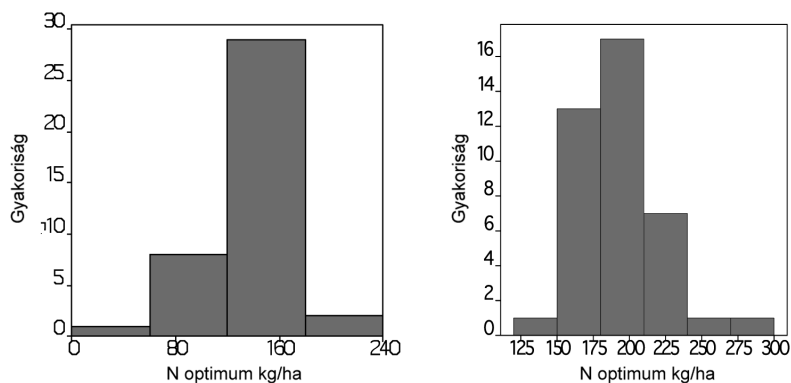
$$Y = 3.707 + 0.0489X - 0.000126X^2$$

$$r^2 = 99.5, n = 4$$



16.12. ábra. Kukorica N-műtrágya reakciójának évenkénti szórásdiagramja (1970–2009)

A 16.12. ábra az évenkénti N-műtrágya reakció adatok feltüntetésével mutatja a reaktívást, a 40 éves adatsorozat szórásdiagramját. A másodfokú függvény ($Y = a + bx - cx^2$) a vizsgált 40 év mindegyikében jól illeszkedett a N-műtrágyázási kísérlet kukorica termésadataihoz.



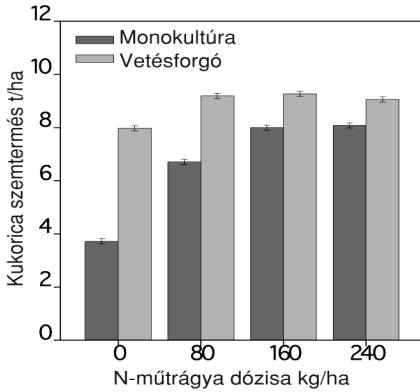
16.13. ábra. A kísérletben mért N-optimumok (bal oldali ábra) és a függvényillesztés alapján becsült optimális N-dózisok gyakorisági eloszlása (1970–2009)

A függvények paramétereit alapján kiszámítottuk az optimális N-dózist és összehasonlítottuk a kísérleti adatok évenkénti varianciaanalízise alapján szignifikáns terménynövekedést eredményező legnagyobb N-dózissal. A kísérletben mért N-optimumok és a függvényillesztés alapján becsült optimális N-dózisok gyakorisági eloszlását a 16.13. ábra mutatja. A kísérleti adatok alapján legnagyobb gyakorisággal az optimális N-dózis 160 kg ha^{-1} volt. A másodfokú függvény túlbecsülte az N-optimumot, amely legnagyobb gyakorisággal 196 kg ha^{-1} volt.

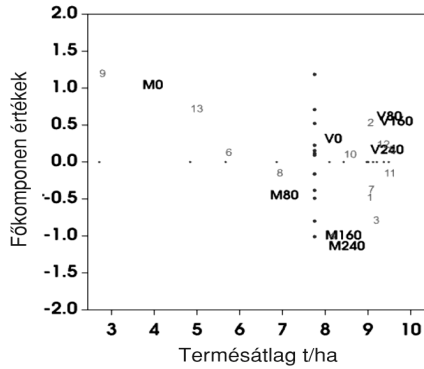
N-műtrágya reakció és termésstabilitás vetésforgóban és monokultúrában

A 16.14. ábra oszlopdiagramjai jól szemléltetik, hogy a kukorica hibridek N-műtrágya reakciója eltérő vetésforgóban és monokultúrában. Vetésforgóban 13 év átlagában 2.246 t/ha -ral nagyobb volt a kukorica termése és a maximális termést alacsonyabb, évjáratától függően, $80\text{--}120 \text{ kg/ha}$ N-dózissal érték el (Berzsenyi, 2009).

A kísérletekben jelentős volt az évjárat hatás, ezért az $\text{év} \times \text{kezelés}$ kölcsönhatást a stabilitásanalízis többváltozós módszerei közül az AMMI analízissel vizsgáltuk. A 16.15. ábra szemlélteti a termésátlagot az X tengelyen és a termésstabilitást kifejező főkomponens interakció értékeket az Y tengelyen. Az M_0 , M_{80} , M_{160} és M_{240} jelöli a monokultúrában, a V_0 , V_{80} , V_{160} és V_{240} a vetésforgóban beállított N-műtrágya kezeléseket. Minél nagyobb egy-egy kezelés főkomponens értéke, annál nagyobb a kezelés hozzájárulása az interakcióhoz, vagyis annál kisebb a termésstabilitása. Megállapítható, hogy vetésforgóban a termés stabilitása minden műtrágyaszinten nagyobb volt, mint monokultúrában. Monokultúrában legkisebb volt a termésstabilitás N-műtrágyázás nélkül és magas N-dózisok esetén, míg legnagyobb stabilitása az N_{80} kezelésnek volt.



16.14. ábra. A kukorica hibridek N-műtrágya reakciója monokultúrában és vetésforgóban (1995–2007)



16.15. ábra. A kukorica termésstabilitás eltérő N-szinteken monokultúrában (M) és vetésforgóban (V)

16.3.4. Növénytermesztési tényezők hatása a kukorica termésére és termésstabilitására

1960-ban közép-kötött vályogtalajon, 4 ismétlésben beállított tartamkísérletben 5 növénytermesztési faktor hatását vizsgáltuk 7 kezeléskombinációban (16.2. táblázat). A vizsgált faktorok: talajművelés, trágyázás, növényesség, fajta, ápolás. Mind-egyik tényezőnek két fokozata van, egy minimális és egy optimális (16.3. táblázat). Az 1. kezelésben minden tényező minimális, a 2. kezelésben minden tényező optimális fokozatban van. A 3–7. kezelésben úgy alakítottuk ki a kezeléskombinációkat, hogy egy növénytermesztési tényezőt mindig minimális fokozatban hagytunk, a többi pedig optimális fokozatban. Így a 3. kezelésben a talajművelés, a 4. kezelésben a trágyázás, az 5. kezelésben a növényesség, 6. kezelésben a fajta és 7. kezelésben az ápolás van minimumban.

16.2. táblázat. A tartamkísérlet kezelései

Kezelések	Talajművelés (T)	Trágyázás (F)	Növényesség (D)	Fajta (V)	Ápolás (W)
1	0	0	0	0	0
2	1	1	1	1	1
3	0	1	1	1	1
4	1	0	1	1	1
5	1	1	0	1	1
6	1	1	1	0	1
7	1	1	1	1	0

0: minimális szint; 1: optimális szint

16.3. táblázat. A tartamkísérletben vizsgált tényezők és fokozataik

Tényező	Minimális szint	Optimális szint
Talajművelés	sekélyszántás 12–14 cm-re	mélyszántás 24–28 cm-re
Trágyázás [†]	trágyázás nélkül	60 t ha ⁻¹ istállótrágya 4 évenként, N ₁₄₀ , P ₆₀ , K ₇₀ évente
Növényszám	35 000 tó ha ⁻¹	70 000 tó ha ⁻¹
Fajta	szabadelvirágzású: Aranyözön és Mindszentpusztai sárga	adott évben a két legjobb hibrid
Ápolás	gyenge, késői egyelés és 2 megkésített kapálás	jó, teljes, gyommentesség, időben egyelve

[†] 1980-tól a műtrágya dózísát duplájára növeltük.

[‡] 1980-ig a minimális növényszám 20 ezer, az optimális növényszám 40 ezer tó ha⁻¹ volt.

Az évenkénti varianciaanalízis kimutatta, hogy a kísérleti kezelések hatása minden évben $P = 0.1\%$ -os szinten szignifikáns volt. A 42 év kombinált varianciaanalízise alapján megállapítottuk, hogy az év hatás és a kezelés hatás egyaránt szignifikáns volt, az MS értékek azonban kimutatták, hogy a kezelés hatás több, mint 16-szor erősebb volt, mint az év hatás (16.4. táblázat).

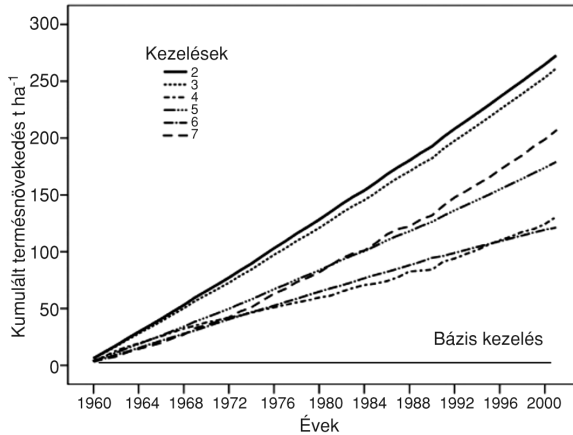
16.4. táblázat. A tartamkísérlet kombinált varianciaanalízise (1960–2001)

Forrás	df	SS	MS	F-érték
Összes	1175	8540		
Környezet (évek)	41	2091	51.0	92.5***
Ismétlés	3	16.6	5.54	10.1
Év (ismétlés)	123	67.8	0.55	
Kezelés	6	5054	842.3	643.9***
Maradék	1002	1311	1.31	

***szignifikáns $P=0.01\%$ -os szinten

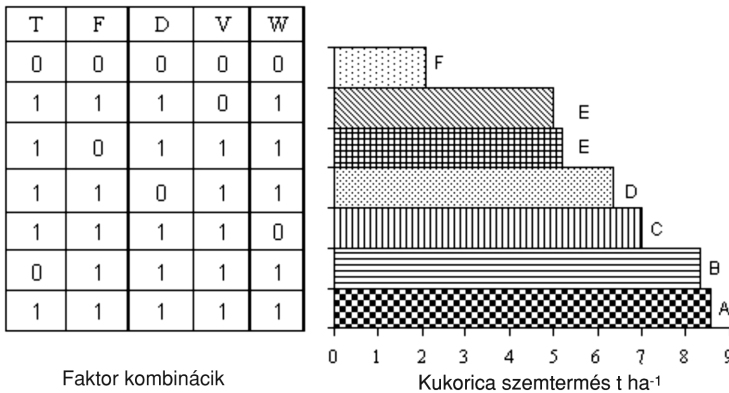
A termesztési tényezők kumulatív termésmenvelő hatását az 1960-2001. évekre, összehasonlítva a kedvezőtlen (minimális) kombinációval, a 16.16. ábra illusztrálja. A tartamkísérlet 42 éves termés adatsorozatát elemezve megállapítottuk, hogy kísérleti körülményeink között a talaj mélyművelésének hatása kismértékű, 2,4 és 3,0% között változott. A tartamkísérletben határozott trágyahatás-növekedést állapítottunk meg, a százalékos hatás a kezdeti 18,6%-ról 30,7%-ra nőtt. A növényszám-növekedés hatása a kukorica termésmenvekedésében kevésbé változott a 42 év során, 19,9 és 23,3% között alakult.

A trágyázás mellett a kukorica termésmenvekedésében legnagyobb hatása a hibrid vetőmagnak (genotípus) van, a hatás 26,6%-ról 32,6%-ra nőtt. A termésmenvelés egyik alapfeltétele a kukorica gondos ápolása, a gyomnövények kompetíciójának megelőzése, mert e nélkül sem a nagyobb növényszám, sem a nagyobb trágyaadag hatása nem tud kellően érvényesülni. A növényápolás, a gyomirtás hatása a kukorica termés-növekedésében kezdetben 23–29% között, az utóbbi 10 évben 14 és 18% között változott (Berzsenyi és Dang, 2008a).



16.16. ábra. Kumulatív termésnövekedés a különböző kezelésekben, összehasonlítva a faktorok kedvezőtlen kombinációjával (1960–2001)

A növénytermesztési tényezők hatását a kukorica termésnövekedésében, 42 év átlagában a 16.17. ábra szemlélteti. Minden tényező minimális, illetve optimális szintjének összehasonlításakor legnagyobb a különbség az átlagos termésreakcióban (2.09 és 8.59 t ha⁻¹). Egy-egy termesztési tényező minimális szintjének hatását vizsgálva (a többi tényező optimuma esetén), megállapítható, hogy a talajművelés mélységének hatása kismértékű (átlagos reakció: 8.32 t ha⁻¹). Megállapítható továbbá, hogy amikor a trágyázás, illetve a genotípus a minimumban levő tényező, a termés csökkenés – az optimumhoz viszonyítva – jelentős, > 3.0 t ha⁻¹ (átlagos termésreakció 5.21, illetve 4.98 t ha⁻¹). A növényszám, illetve a gyomirtás minimális szintjénél a termés csökkenés 2.2, illetve 1.6 t ha⁻¹ (átlagos reakció 6.36 és 7.01 t ha⁻¹). A tartamkísérlet 42 éves adata alapján a vizsgált növénytermesztési tényezők a kukorica termésnövekedéséhez a következő arányokban járultak hozzá: fajta 32.6%, trágyázás 30.6%, növényszám 20.2%, gyomirtás 14.2 és talajművelés 2.4% (Berzsenyi és Dang, 2008a).



16.17. ábra. A növénytermesztési tényezők hatása a kukorica termésnövekedésére tartamkísérletben (1960–2001)

A kezeléseket az 1. táblázat tartalmazza. Azok a kezelések (fektetett oszlopok), melyeket azonos betű követ, szignifikánsan nem különböznek a Duncan MRT alapján.

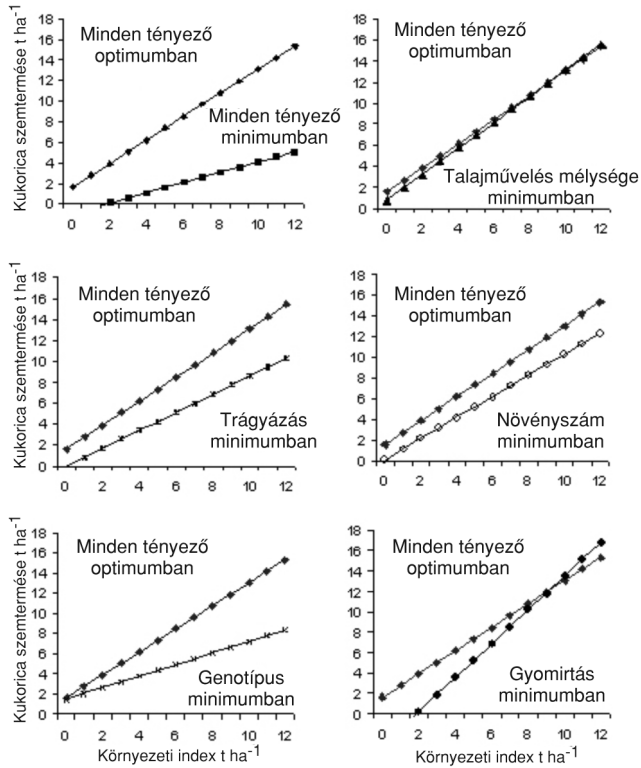
16.5. táblázat. A kísérleti kezelések stabilitását jellemző értékek az 1960-2001. évek adatai alapján

Stabilitás mutató	Minden tényező minimumban	Minden tényező optimumban	Minimumban levő tényező				
			Talajművelés	Trágyázás	Növény-szám	Fajta	Ápolás
CV%	45.7	19.5	21.9	34.8	24.5	25.1	36.6
W ²	117.7	51.0	53.5	219.2	57.1	163.2	325.9
σ ²	3.22	0.94	1.02	6.68	1.15	4.77	10.32
YS		+	+		+		+
r	0.796	0.947	0.962	0.713	0.919	0.690	0.899
b	0.500	1.145	1.236	0.860	1.020	0.580	1.659
Termés t ha ⁻¹	2.094 F	8.585A	8.315B	5.205E	6.355D	4.979E	7.014C

CV%: variációs koefficiens; W²: ökovalencia; σ²: stabilitás variancia; YS: termésstabilitás; r: korrelációs koefficiens; b: regressziós koefficiens; +: Kang (1995) STABLE programja által kiválasztott kezelések.

A különböző kísérleti kezelések stabilitásanalízisének variancia és regresszió paramétereit a 16.5. táblázatban foglaltuk össze. A termésátlagok évenkénti szórását kifejező CV% értéke legnagyobb minden tényező minimális szintjénél (45.5%) és legkisebb a növénytermesztési tényezők optimális kombinációjánál (19.5%). Egy-egy tényező minimuma esetén (a többi tényező optimumban) a növényápolás, illetve a trágyázás minimuma eredményezte a legnagyobb CV%-ot (36.6%, illetve 34.8%). Az ökovalencia (W²) és a stabilitás variancia (σ²) mutatók nagysága kifejezi a kezelések hozzájárulásának mértékét a szignifikáns év × kezelés kölcsönhatáshoz. Az analízis eredménye szerint az interakcióért legnagyobb mértékben azok a kezelések felelősek, amelyekben a növényápolás, trágyázás, genotípus, illetve minden tényező minimumban volt. Az átlagos termésreakció és a stabilitás alapján, Kang (1993) termésstabilitás (YS) paramétere kimutatta, hogy a 2, 3, 5 és 7 jelű kezelés volt a legkedvezőbb (16.5. táblázat).

Amikor a kezelések hatását eltérő környezetben vizsgáltuk a stabilitásanalízis regresszió módszerével (16.18. ábra), legnagyobb különbséget az átlagos termésreakcióban (2.094 és 8.585 t ha⁻¹), valamint a két lineáris függvény regressziós koefficiensének (b) értékében (0.500 és 1.145) akkor kaptunk, amikor minden tényező minimális szintjének kombinációját minden tényező optimális szintjének kombinációjával hasonlítottuk össze. Eltérő környezetben vizsgálva a két kezelést, látható, hogy minden tényező minimuma esetén 2 t ha⁻¹ környezeti átlagnál a termesztés teljesen meghiúsul. Másrészt, magas környezeti átlag felé haladva, nő a két kezelés közötti különbség (melyet az egyenesek távolsága fejez ki). Látható, hogy a talajművelés mélységének kismértékű hatása csak alacsony környezeti átlagnál mutatható ki. A trágyázás minimális szintjénél a termés-csökkenés nagysága jelentős és magas környezeti átlag felé haladva fokozódik. A termés-csökkenés nagysága és trendje hasonló a genotípus minimális szintjénél is. Ugyancsak hasonló a tendencia a növény-szám minimális szintjénél, azonban a termés-csökkenés mértéke, különösen magas környezeti átlagnál, kisebb (melyet a regressziós koefficiens is kifejeznek: b_i: 0.860, b_n: 1.020, b_g: 0.580). A többi tényezőhöz viszonyítva, eltérő a gyomirtás minimumának környezeti függése, ugyanis a gyomosodás termés-csökkentő hatása alacsony környezeti átlagnál a legnagyobb és 10 t ha⁻¹ környezeti indexnél minimális (Berzsenyi és Győrffy, 1995).

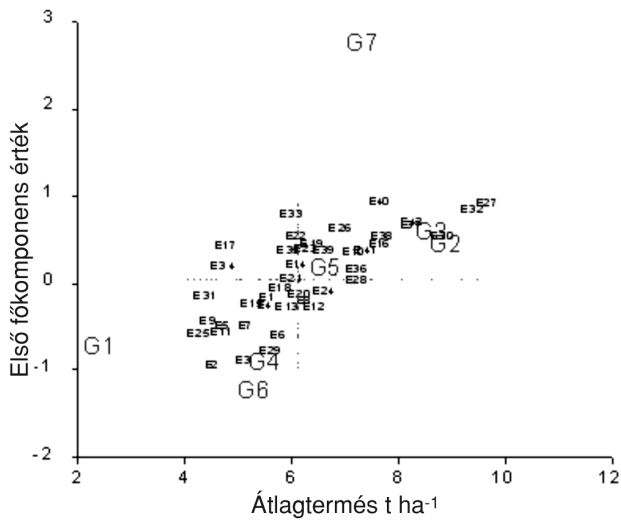


16.18. ábra. A kukorica termésreakciója egy-egy termesztési tényező kedvezőtlen (minimális) szintjére, összehasonlítva minden faktor kedvező (optimális) szintjével, eltérő környezetben (1960–2001)

16.16. táblázat. A tartamkísérlet adatainak (1960-2001) AMMI analízise

Forrás	df	SS	MS	F-érték
Összes	1175	8540		
Kezelés	6	5054	842.3	1969.7***
Környezet	41	2091	51.0	76.1***
Blokk	126	84	0.67	
Interakció	246	988	4.0	9.39***
IPCA1	46	467	10.2	23.8***
IPCA2	44	238	5.4	12.7***
Maradék	156	282	1.8	4.23***
Hiba	756	323	0.43	

***szignifikáns P=0.01%-os szinten



16.19. ábra. A 7 növénytermesztési kezeléskombináció (G1-G7) és a 42 környezet (E1-E42) átlagtermésének és I. főkomponens értékének AMMI diagramja

Az évek és a kezelések fő hatását, valamint kölcsönhatását az AMMI módszerrel analizáltuk, amely magában foglalja a varianciaanalízist és a főkomponensanalízist (16.6. táblázat). Az AMMI analízis eredményét a 16.19. ábra illusztrálja. Az X tengelyen a termésátlag, az Y tengelyen a főkomponens értékek láthatók a 7 kezelésre a 42 környezetben (évben). Minél nagyobb a főkomponens értéke, annál nagyobb a kezelés hozzájárulása az interakcióhoz, azaz annál kisebb a termésstabilitás. Látható, hogy a 7., 6., 4. és 1. kezelés (gyomirtás, fajta, trágyázás, illetve minden tényező minimumban) járult hozzá legnagyobb mértékben az interakcióhoz, míg a 2., 3. és 5. kezeléseknek (minden tényező optimumban, illetve a talajművelés, a növényszám minimumban) legnagyobb a termésstabilitása. Együtt értékelve az átlagos teljesítményt és a termésstabilitást, a legjobb volt a 2. kezelés, amely minden termesztési tényező optimális kombinációját tartalmazta. Sorrendben a 3. kezelés következett, ahol a talajművelés volt minimumban és a többi faktor optimumban (Berzsenyi és Dang, 2008a).

17. Nem-paraméteres próbák

17.1. Bevezetés

Az eddig tárgyalt minta összehasonlítási szignifikancia vizsgálatok paraméteres tesztek (próbák) voltak, amennyiben kísérleteket analizáltunk, ahol a minta átlagokat és varianciákat használtuk a populáció átlagok és varianciák becslésére. Az ilyen tesztek matematikai eljárása azon a feltételezésen alapul, hogy a minták normális eloszlású populációkból származnak, melyeknek egyenlő a varianciája. Olyan esetekben, ahol ezek a feltételek nem érvényesülnek, például, amikor az adatok ferde eloszlásúak, amikor az adatok számolások vagy százalékok formájában vannak, amikor az összehasonlítandó minták varianciái statisztikailag nem homogének, egy alternatív megközelítés szükséges. Egy olyan tesztre van szükség, amely a populáció adat eloszlásától független statisztikai mutatót eredményez. Az ilyen tesztet, mivel nem teszik szükségessé a populáció paraméterek becslését, *nem-paraméteres teszteknek* nevezzük.

Az általános módszer, melyet arra használunk, hogy nem az adat eloszlásra építve, statisztikailag leírjuk az adatokat, az adatok rangsor értékét használja. A rangsorra, amely az adatok sajátos sorrendjére vonatkozik, nincs hatással az adat értékek nagysága vagy az értékek tartománya és eloszlása. Például, egy 20 értéket tartalmazó mintában, a legnagyobb értéknek a rangja 1 lesz és a legkisebb kapja a 20-as rangsort, tekintet nélkül az értékek közötti tényleges számszerű különbségek nagyságrendjétől. A nem-paraméteres tesztek, amelyek az adatok rangsorán alapulnak, és nem a tényleges adatokon, ezáltal nem tesznek szükségessé semmilyen feltételt az adatok típusára vagy az adatok eloszlására. A valóságban, az ilyen tesztek megvizsgálják a *mintámediánok közötti különbséget*, sokkal inkább, mint a populáció átlagokat, és a null hipotézist a mediánokra fogalmazzuk meg.

Az adatok rangsorba állításával az adatokra vonatkozó bizonyos mennyiségű információ elveszik, a rangsorba állítás ugyanis nem veszi figyelembe az adatok nagyságrendjét vagy gyakorisági eloszlását. Következésképpen, a nem-paraméteres tesztek sokkal kevésbé érzékenyek az adatsorozatok közötti kis különbségekre, vagyis *a nem-paraméteres teszteknek kisebb a statisztikai ereje*, mint a paraméteres teszteknek. Ezáltal, ahol az adatokról kimutatható, hogy megfelelnek a paraméteres teszt feltételeinek, akkor a paraméteres tesztet kell előnyben részesíteni, mivel mindig nagyobb a valószínűsége, hogy feltárunk kisebb különbségeket szignifikánsnak, szemben a nem-paraméteres megfelelőjével.

17.2. Két független minta közötti különbség analízise

A leggyakrabban használt nem-paraméteres teszt két független minta közötti különbség vizsgálatára a **Mann-Whitney U teszt**. Ennek a tesztnek háttere az, hogy ha két minta homogén, akkor a két minta egyesítésekor és rangsorba állításakor, a

rangsorok jelentős átfedése következik be, mindegyik minta adatértékeinek tulajdoníthatóan. Minél nagyobb azonban a különbség a minták között, annál kisebb lesz az átfedés a két minta értékeinek rangsora között. Ezáltal szükség van egy statisztikai mutatóra, amely méri a rangsor értékek átfedésének mértékét a két minta minden értékéből, és amelynek van egy ismert valószínűségi eloszlása.

A Mann-Whitney statisztikai mutatót U szimbólummal jelöljük. Az U érték kiszámításakor először, a két minta adatait egy adatsorba egyesítjük és rangsorba állítjuk, figyelve arra, hogy bármely azonos (kapcsolt) értéknek ugyanazt a megosztott rangsor értéket adjuk. Mindegyik adat rangsor értékét mindegyik mintán belül összegezzük, amely a rangértékek két összegét, $\sum R_A$ és $\sum R_B$ adja.

A két alternatív statisztikai mutatót, U_A -t és U_B -t ezután kiszámítjuk az alábbi formulákkal:

$$U_A = \sum R_A - \frac{n_A(n_A + 1)}{2} \text{ és } U_B = \sum R_B - \frac{n_B(n_B + 1)}{2}$$

ahol n_A és n_B az értékek száma az A és B mintában.

A számítás korrektségének matematikai ellenőrzése elvégezhető ebben a stádiumban, mivel

$$\sum R_A + \sum R_B = N(N + 1)/2, \text{ ahol } N = n_A + n_B, \text{ és } U_A = (n_A \times n_B) - U_B.$$

Mivel két lehetséges statisztikai mutató van, U_A és U_B , szükséges tudni, melyiket használjuk a Mann-Whitney tesztnél. A kétoldali tesztnél, ahol a különbség aránya nem specifikált, a két érték közül a nagyobbat használjuk statisztikai mutatóként és $U_{\text{megfigyelt}}$ -nek hívjuk. Az egyoldali tesztnél az U_A vagy az U_B kiválasztása függ a különbség irányától, melynek bekövetkezését előrejeleztük. Ha tesztelni szükséges, hogy minta A nagyobb-e mint minta B , akkor az U_A -t választjuk statisztikai mutatóként, ha azonban a hipotézis az, hogy a B minta nagyobb a kettő közül, akkor U_B válik statisztikai mutatóvá.

A rangszám meghatározása a következőképpen történik: Egyesítjük a két mintát, majd az így kapott $n_A + n_B$ elemű minta elemeit rangsorba állítjuk. Ezután minden mintaelemhez hozzárendeljük annak rangsorbeli sorszámát, az ún. rangszámot. Ha esetleg a rangsorban egymással egyenlő mintaelemek is előfordulnak, akkor ezek mindegyike egységesen az azokhoz folyamatosan hozzárendelhető sorszámok számtani átlagát kapja rangszámként. Ez az ún. kapcsolt rangok esete. Például, ha két azonos érték van, amelyek az ötödik legnagyobb értéket reprezentálják egy adatsorban, akkor mindkettő az 5.5 rangszámot kapja. A következő legnagyobb érték a 7-es rangszámot kapja.

Azokban az esetekben, amikor két vagy több adatérték azonos, nagyon fontos, hogy korrekt eljárást használjunk a rangsor értékekre. Minden azonos értékhez ugyanazt a rangsor értéket rendeljük hozzá. Ahol több, mint két érték kapcsolt, akkor ezeknek hasonlóan ugyanabban a rangsorban kell osztozniuk, amely megegyezik az elfoglalt értékek legalacsonyabb és legmagasabb rangsora közötti számtani átlagnak. Így, ha az ötödik legnagyobb érték négy adat között oszlik meg, akkor felhasználják az 5., 6., 7. és 8. sorrendet és ezáltal mindegyik 6.5-ös rangsor értéket kap, míg a következő legnagyobb adat rangsor értéke 9 lesz.

Példa. Egy levélmoly lárva által okozott levélfertőzés (százalékban) szignifikáns különbségének meghatározása két eltérő egres fajtán (17.1. táblázat) (Ireland, 2010)

17.1. táblázat. Növényenkénti %-os levélfertőzés két eltérő egresfajtán

% -os levélfertőzés növényenként			
A fajta		B fajta	
Érték (%)	Rangsor	Érték (%)	Rangsor
25	14	15	7
38	21	32	19
31	18	20	11
29	17	8	2
40	22	12	6
26	15	24	12.5 ^a
24	12.5 ^a	18	10
16	8	11	5
10	4	5	1
35	20	28	16
		9	3
		17	9
$n_A = 10$	$\sum R_A = 151.5$	$n_B = 12$	$\sum R_B = 101.5$ (N = 22)

^a Figyeljük meg, hogy két 24-es érték van. Mivel a 12. és 13. rangsort foglalják el, mindkettő 12.5 rangszámot kap. A következő érték rangsora ezáltal 14 lesz.

(Matematikai ellenőrzés: $\sum R_A + \sum R_B = 253 = N(N + 1)/2$)

A kétoldali kritikus értékek a Mann-Whitney U eloszlásra az eltérő minta méretek alapján a *Melléklet 6. táblázatban* található; jegyezzük meg, hogy mindkét minta méretét, n_A és n_B figyelembe kell venni. Abból a célból, hogy egyoldali Mann-Whitney U tesztet végezzünk, az U megfigyelt statisztikai mutatót össze kell hasonlítani az egyoldali kritikus értékkel; ezt megkaphatjuk a kétoldali U táblázatból, a szükséges valószínűség kétszeresét véve. Ebből következik, hogyha az U megfigyelt értéke egyenlő vagy nagyobb, mint az U kritikus értéke, a valószínűség kiválasztott kritikus szintjén, akkor a H_0 visszautasítható és a minták közötti különbség szignifikánsnak tekinthető.

Null hipotézis: Nincs különbség a két fajta fertőzött levél százalékának medián értéke között.

$$U_A = \sum R_A - \frac{n_A(n_A + 1)}{2} = 151.5 - 55 = 96.5$$

$$U_B = \sum R_B - \frac{n_B(n_B + 1)}{2} = 101.5 - 78 = 23.5$$

Matematikai ellenőrzés: $U_A = 96.5 = (n_A \times n_B) - U_B$

Mivel $U_A > U_B$, ezért U_A -t használjuk statisztikai mutatóként a Mann-Whitney kétoldali szignifikancia vizsgálatban.

A kritikus kétoldali U érték $\alpha = 0.05$ -nél és $n_A = 10$ és $n_B = 12$ mintaméretnél 91. Mivel U_A nagyobb, mint 91.0, a null hipotézis visszautasítható.

Következtetés: Szignifikáns különbség van a levélmoly lárva által okozott levélfertőzés mértékében az egres két fajtája között ($P < 0.05$).

Számítógépes értékelés GenStat programmal

A megfigyelési adatokat külön oszlopokba visszük be (példánkban: Egres1 és Egres 2). A menük kiválasztásának sorrendje: Stats – Statistical Tests – Two-sample Nonparametric Test. A legördülő menükből kiválasztjuk: Mann-Whitney U Test. A menü dobozban: Data Arrangement: Two sets. Data set 1: Egres1, Data set2: Egres2. Options-re kattintva, Display: Ranks bejelölésével a rangsort is megkapjuk. A Run-ra kattintva az output a következő:

Mann-Whitney U (Wilcoxon rank-sum) test

Variates: Egres1, Egres2.

Value of U: 23.5 (first sample has higher rank sum).

Exact probability (adjusted for ties): 0.014
(under null hypothesis that Egres1 is equal to Egres2).

Sample sizes: 10, 12.

Ranks for individual variables

Egres1	14.0	21.0	18.0	17.0	22.0	15.0	12.5	8.0	4.0	20.0		
Egres2	7.0	19.0	11.0	2.0	6.0	12.5	10.0	5.0	1.0	16.0	3.0	9.0

Mann-Whitney U teszt kiszámítása nagy mintaméretre ($n > 20$)

A Mann-Whitney kritikus U értékek közötti statisztikai táblázatainak többsége nem ad értékeket 20-nál nagyobb mintaméretre. Ha azonban a mintaméret nagyok ($n > 20$), akkor az U eloszlás nagyon közelít a normális eloszláshoz és a normális eloszlás használható az U értékek szignifikanciájának tesztelésére (Ireland, 2010).

Egy átlagot az U eloszlásra először meghatározza:

$$\mu_U = \frac{n_A n_B}{2}$$

A normális eloszlású statisztikai mutató értékét a standard módszerrel határozzuk meg, vagyis a megfigyelt érték és az elméleti átlag különbségét osztjuk a standard hibával. Ebben az esetben a következőképpen kapjuk meg:

$$Z_{\text{megfigyelt}} = \frac{U_{\text{megfigyelt}} - \mu_U}{\sqrt{\frac{n_1 n_2 (N + 1)}{12}}}$$

ahol: $N = n_1 + n_2$
 $U_{\text{megfigy}} = \text{a nagyobb } U_A \text{ vagy } U_B$

Sajnos, ha kapcsolt rangok vannak az adatokban, bonyolultabb korrekció szükséges, hogy a normális eloszlás jobb megközelítését kapjuk. Ebben az esetben, a $\sum t$ meghatározása szükséges először:

$$\sum t = \sum (t_i^3 - t_1)$$

ahol: $t_i = \text{a kapcsolt rangok száma a kapcsolt értékek egyik csoportjában}$

A $z_{\text{megfigyelt}}$ teszt mutató meghatározása ekkor:

$$z_{\text{megfigyelt}} = \frac{U - \mu_U}{\sqrt{\frac{n_1 n_2}{N^2 - N} \times \frac{N^3 - N - \sum t}{12}}}$$

A $z_{\text{megfigyelt}}$ értékét ezután összehasonlítjuk a z kritikus értékével (Melléklet 15. táblázat) az indokolt valószínűségnél, és ettől függően a null hipotézist vagy visszautasítjuk vagy elfogadjuk.

17.3. Páros, nem-független minták közötti különbségek analízise

Számos nem-paraméteres teszt áll rendelkezésre két nem-normális eloszlású páros minta közötti különbség analízisére. A két leggyakrabban használt az **előjel teszt** és a **Wilcoxon előjel és rangteszt**.

17.3.1. Az előjel teszt

Az előjel teszt egy rendkívül egyszerű teszt annak ellenőrzésére, hogy van-e szignifikáns különbség a páros minták között. Csupán az adat-párok közötti különbség irányán alapul, azaz vajon a különbségek pozitívak vagy negatívak, és nem szükséges semmilyen kvantitatív érték input. Ezáltal különösen hasznos, amikor a vizsgálni kívánt változót bonyolult kvantifikálni. Ha azonban az adatok numerikusak, akkor egyszerűen az adatpárok közötti különbség előjele az, amelyet felhasználunk a tesztben.

A null hipotézis az, hogy nincs különbség a két populáció között, amelyekből a mintákat vették. Ha a null hipotézis igaz, akkor feltételezhető, hogy a pozitív különbségek száma és a negatív különbségek száma a minta-párok között nagyjából ugyanaz. Ha azonban szignifikáns különbség volt jelen, akkor eltérés várható a pozitív és negatív különbségek között. A statisztikai mutatót az előjel tesztre S -sel vagy $S_{\text{megfigyelt}}$ -tel jelöljük, és egyszerűen a pozitív különbségek száma vagy a negatív különbségek száma, amelyik a nagyobb. Mivel minden adatpár összehasonlításának kimenete csak „-” vagy „+” lehet, és a null hipotézis szerint egyenlő az esélye „-” vagy „+” érték megkapásának, akkor lényegében az adatok binárisak és várhatóan binomiális eloszlást követnek $p = 0.5$ -tel. Ezáltal a statisztikai mutató összehasonlítható az elméleti

értékeknek egy sorozatával, amelyek binomiális eloszláson alapulnak, abból a célból, hogy meghatározzuk a szignifikancia valószínűségét (Melléklet 8. táblázat).

Hipotézisek

H_0 : Valószínűség ($X > Y$) = Valószínűség ($X < Y$) = 0.5 Más jelöléssel: Prob ($X > Y$) = Prob ($X < Y$) = 0.5

A null hipotézis kimondja, hogy X populáció ugyanaz, mint Y populáció. Vagyis kvantitatív adatokra a medián különbség X és Y között zéró.

Három alternatív hipotézis van:

H_1 : Prob ($X > Y$) \neq 0.5

Az X medián értéke nem egyenlő Y medián értékével kvantitatív adatokra

H_1 : Prob ($X > Y$) $>$ 0.5

Az X medián értéke nagyobb, mint Y medián értéke kvantitatív adatokra vagy X populáció „kiválóbb” az Y populációhoz viszonyítva, kvantitatív adatokra.

H_1 : Prob ($X > Y$) $<$ 0.5

Az X medián értéke kisebb, mint Y medián értéke kvantitatív adatokra vagy X populáció „rosszabb”, mint Y populáció kvalitatív adatokra.

Következtetések

Ha az eltérés nem volt szignifikáns, akkor a null hipotézist kell elfogadni. A következtetés az, hogy nincs bizonyíték arra, hogy a két populáció nem azonos. Ha az eltérés szignifikáns volt, akkor a null hipotézist visszautasítjuk. Egy kétoldali tesztben a következtetés az, hogy a két populáció nem egyforma. Egy egyoldali tesztben a konklúzió az, hogy a populációk különböznek a vizsgált irányban.

Példa. Egy vizsgálatban pontszámmal értékelték a tehenek tőgyén az elülső és hátsó tőgybimbók állapotát és így a minták párosak, mivel minden állatot felhasználtak, hogy megadják az egymásnak megfelelő saját mintájukat. A cél az volt, hogy megtalálják, hogy bármilyen különbség a tőgybimbók kondíciójában összefüggésbe hozható-e olyan körülményekkel, mint az tőgygyulladás. Az elülső tőgybimbó kondícióját pontozták, úgy mint akár rosszabb, akár jobb vagy egyenlő a hátsó tőgybimbókkal. Ez a becslési módszer lehetővé tette, hogy több tehenet minősítsenek minimális idő alatt. Az adatok a 17.2. táblázatban találhatók.

17.2. táblázat. 20 tehen elülső tőgybimbója kondíciójának jelölése, annak megfelelően, hogy rosszabb volt (-), jobb volt (+), mint a hátsó, vagy az első tőgybimbók kondíciója egyenlő volt (0) a hátsó tőgybimbók kondíciójával

Tehén	1	2	3	4	5	6	7	8	9	10
Pontkülönbség	+	0	0	+	-	+	-	0	+	+
Tehén	11	12	13	14	15	16	17	18	19	20
Pontkülönbség	+	0	-	+	0	+	0	+	-	+

H_0 : Nincs különbség elülső és hátsó tőgybimbók kondíciójában.

H_1 : Az elülső és hátsó tőgybimbók kondíciója különböző.

Szignifikancia szint, $\alpha = 0.05$

Párok száma, amely „+”-t vagy „-”-t eredményezett = 14

Hányszor fordult elő a leggyakoribb jel ($S_{\text{megfigyelt}} = 10$)

Adatpárok száma, $n = 20$

S kritikus értéke ($\alpha = 0.05$): 15 (táblázati érték)

Mivel $S_{\text{megfigyelt}} < S_{\text{kritikus}}$, a nullhipotézist elfogadjuk.

Következtetés

A következtetés az, hogy nincs bizonyíték a null hipotézis elutasítására, vagyis nincs szignifikáns különbség az elülső és a hátsó tőgybimbók állapotában.

17.3.2. Wilcoxon előjelteszt

Ha kvantitatív értékek állnak rendelkezésre a páros mintáknál, akkor az előjel teszt használata helyett előnyben részesíthető egy olyan teszt használata, amely a kvantitatív információt is felhasználja. A Wilcoxon előjel teszt eléri ezt azért, hogy átalakítja a minta párok közötti különbség nagyságrendjét rangsor értékre.

A Wilcoxon előjel teszt statisztikai mutatóját a T szimbólummal jelöljük. Ahhoz, hogy meghatározzuk a T értéket, először minden egyes minta-pár közötti numerikus különbség irányát kell feljegyezni (azaz, az érték nőtt vagy csökkent) és ezután a páros értékek közötti abszolút különbséget határozzuk meg. Ezeket a különbségeket ezután rangsoroljuk növekvő sorrendben, a legkisebb különbségnek adjuk az 1-es rangsort, a következő legnagyobb különbségnek 2-es rangsor értéket adunk, stb. Ha nincs különbség egy adat-pár között értékben, akkor a különbség zéró és nem adunk rangsor értéket; az ilyen adat-párok ezáltal nem járulnak hozzá az analízishez. Ha két vagy több különbség érték azonos, akkor ugyanazt a középérték rangsort kapják, felhasználva a Mann-Whitney próbánál leírt eljárást. Azokat a rangsor értékeket, amelyek pozitív különbségekkel kapcsolódnak, és azokat a rangsor értékeket, amelyek negatív különbségekkel kapcsolódnak, ezután külön-külön összegezzük, amely két rangsor érték összeget eredményez. Ha nem volt szignifikáns különbség a két minta között, akkor várható, hogy ez a két érték megközelítőleg azonos lesz, míg minél nagyobb az eltérés köztük, annál nagyobb annak valószínűsége, hogy szignifikáns különbség van jelen. A teszt végzésekor a két rangsor érték összeg közül a nagyobbikat használjuk $T_{\text{megfigyelt}}$ statisztikai mutatóként és ezt az értéket ezután összehasonlítjuk a T kritikus értékével a választott valószínűségi szinten. Ha $T_{\text{megfigyelt}}$ nagyobb, mint T_{kritikus} , akkor a null hipotézist elutasítjuk és szignifikáns különbséget igazolunk a két minta között. A kritikus T értékeket különböző mintaméretekre a T valószínűségi táblázat tartalmazza (Melléklet 8. táblázat).

Hipotézisek

H_0 : A különbségek mediánja (d_i) = 0

Három alternatív hipotézis van:

H_1 : medián $d_i \neq 0$

Az X mediánja nem egyenlő Y mediánjával, így a két populáció különböző.

H_1 : medián $d_i > 0$

Az X mediánja nagyobb mint Y mediánja.

H_1 : medián $d_i < 0$

Az X mediánja kisebb mint Y mediánja.

Következtetések

Ha az eltérés nem volt szignifikáns, akkor a null hipotézist fogadjuk el. A következtetés az, hogy nincs bizonyíték azt mondani, hogy a két populáció nem azonos. Ha az eltérés szignifikáns, akkor a null hipotézist visszautasítjuk. Egy kétoldali tesztben a konklúzió az, hogy a két populáció nem volt egyenlő. Egy egyoldali tesztben a következtetés az, hogy a populációk különböztek a vizsgált irányban.

Példa: Egy vizsgálatot végeztek a tehenek új istállórendszerében annak megállapítására, hogy az új rendszer növelte-e a tehenek lábának károsodását. Egy pontozási rendszert használtak, amely pontozta a tehenek lábán található sérüléseket, amelyből minden állatra egy összesített pontszámot kaptak. Minden állatot pontoztak, mielőtt átvitték azokat az új istállórendszerbe és ezt követően, az új istállórendszerben eltöltött bizonyos időszak után. A kapott pontokat a 17.3. táblázat tartalmazza.

17.3. táblázat. A tehenek lábkárosodásának pontszámai az új istállórendszerbe vitelt megelőzően és az új istállórendszerben eltartott bizonyos időszak után

Tehenek száma	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Előtt	27	20	6	14	12	14	9	22	36	6	0	35	40	18	20
Után	61	35	61	36	69	22	24	58	36	16	33	10	21	22	16
d_i	-34	-15	-55	-22	-57	-8	-15	-36	0	-10	-33	25	19	-4	4
Rangsor	11	5.5	13	8	14	3	5.5	12	-	4	10	9	7	1.5	1.5

H_0 : medián $d_i = 0$

Nincs különbség a sérüléseket jellemző pontszámokban a vizsgálat elején és végén.

H_1 : median $d_i < 0$

A sérülés pontszámok növekedtek a vizsgálat idején.

Szignifikancia szint, $\alpha=0.05$

Rangsorok száma = 14

Pozitív rangsorok összege, $T=17.5$

Negatív rangsorok összege, $T = 87.5$

Kritikus T érték ($\alpha=0.05$): 84 (táblázati érték)

Következtetés

A próbastatisztika $T_{\text{megfigyelt}} > T_{\text{kritikus}}$. A következtetés az, hogy az új istállórendszer növelte a tehenek lábának károsodását.

Számítógépes értékelés GenStat programmal

Wilcoxon Matched-Pairs Test

Variate: Labkarosodas2-Labkarosodas1

Test Statistic: 17.50 (rank sum positive)

Sample size: 14 (zero values have been excluded)

Probability: 0.025 (two-sided test)

Signed Ranks of Differences

11.0	5.5	13.0	8.0	14.0	3.0	5.5	12.0	0.0	4.0	10.0
-9.0	-7.0	1.5	-1.5							

17.4. Nem-paraméteres többmintás analízis

Amikor egy kísérletben három vagy több mintát képezünk, amelyek összehasonlítást igényelnek, de amelyekről nem feltételezhető, hogy normális eloszlásúak, beleértve az adatok ferde eloszlását vagy amikor az adatok számolások vagy százalékok formájában vannak, nem érvényes a paraméteres ANOVA teszt használata. Ilyen körülmények között, többmintás, eloszlás-szabad analízisre van szükség, közülük leggyakrabban a **Kruskal-Wallis tesztet** használják.

17.4.1. Kruskal-Wallis teszt

A teszt lényegében többmintás analízis, amely a rangsorolt adatokon alapszik és ezáltal vizsgálja, hogy a minták azonos mediánú populációkból származnak-e. Hatásában, kiterjesztése a nem-paraméteres Mann-Whitney tesztnek, melyet két minta mediánjai közötti különbség vizsgálatára használunk. A valóságban, két mintára a Kruskal-Wallis és a Mann-Whitney tesztek azonos eredményt adnak.

A kezdő lépés a Kruskal-Wallis tesztben, hogy egyesítünk (kombinálunk) és azután numerikusan rangsorolunk minden adatot a kísérletben. (Két vagy több kapcsolt rangok esetében, mindegyiknek ugyanazt a rangsor értéket adjuk, amely számtani átlaga az általuk elfoglalt rangsor pozícióknak, pl. ha két érték kapcsolt a tízes rangsorhoz, akkor mindegyiknek a 10.5 rangsort adjuk és a következő érték kapja a 12 rangsor értéket.) A rangsorok összegét ($\sum R$) ezután kiszámítjuk mindegyik mintára. A teszt alapelve, hogy ha a minták mediánjai szignifikánsan nem különböznek (azaz H_0 korrekt), akkor a rangsor értékek egyenletesen oszlanak el a minták között és a rangsorok összege ($\sum R$) hasonló lesz mindegyik mintára. Ha a minta mediánok szignifikánsan különböznek (azaz H_0 nem korrekt), akkor egy vagy több minta tartalmazni fogja az alacsony rangsorok többségét és egy vagy több minta tartalmazni fogja a magas rangsorok többségét. A rangsorok összege nem lesz ekkor hasonló mindegyik mintában. Szükségünk van egy statisztikai mutatóra, amely teszteli a variációt $\sum R$ -ben és értéke alacsony lesz, amikor a null hipotézis korrekt és magas lesz az értéke amikor a null hipotézis téves.

A Kruskal-Wallis teszt H mutatóját az alábbi formulával határozzuk meg:

$$H = \left[\frac{12}{N(N+1)} \times \sum \frac{(\sum R)^2}{n} \right] - 3(N+1)$$

ahol: N = az adatok összes száma minden mintában;

$\sum R$ = a rangsorok összege a mintán belül;

n = adatértékek száma a mintában.

A gyakorlatban először a $(\sum R)^2/n$ kifejezést határozzuk meg minden mintára; minden $(\sum R)^2/n$ értéket ezután összegezzük, hogy megadjuk a $\sum(\sum R)^2/n$ -t és a számítás ezután elvégezhető a fenti formulának megfelelően.

Miután meghatároztuk a H értéket az adatsorra, ezt össze kell hasonlítani a kritikus értékkel annak eldöntésére, hogy elfogadjuk vagy visszautasítjuk a null hipotézist. A kritikus H értékek a *Melléklet, 12. táblázatban* találhatóak. A táblázat megkérdezi a kritikus valószínűséget (α), a minták számát a kísérletben (k) és a minták méretét (m). Ha a H mutató megfigyelt értéke egyenlő vagy nagyobb, mint a kritikus érték, akkor a null hipotézist visszautasítjuk a valószínűség adott szintjén.

A H valószínűségi eloszlása nagyon hasonló egy másik statisztikai mutató, a chi-négyszet (χ^2) valószínűségi eloszlásához. Valójában, mihelyt az összes minta mérete eléri az ötöt és az összehasonlítandó minták (k) összes száma eléri az ötöt, a H eloszlása annyira közel van a chi-négyszetéhez (χ^2), hogy χ^2 -et megbízhatóan használhatjuk a H kritikus értékei helyett. A megfelelő χ^2 értékek a *Melléklet 5. táblázatban* találhatóak.

Példa. Abból a célból, hogy megvizsgáljuk egy károsító lepke eloszlását egy gyümölcsös állományában, megfelelő ismétlésben csapdákat helyeztek el a lombzat alsó, középső és felső zónájában és feljegyezték egy 24 órás időszak alatt begyűjtött lepkék számát (17.4. táblázat). A Kruskal-Wallis analízist alkalmazták, hogy megvizsgálják a szignifikáns különbséget a begyűjtött lepkék számában, a három állomány zónában.

17.4. táblázat. A gyümölcsös lombzatának különböző zónáiban begyűjtött lepkék száma

Alsó lombzat		Középső lombzat		Felső lombzat	
Lepkék száma	(Rangsor érték)	Lepkék száma	(Rangsor érték)	Lepkék száma	(Rangsor érték)
14	(15)	11	(13)	7	(8.5)
12	(14)	5	(4.5)	7	(8.5)
9	(11)	6	(6.5)	2	(1)
8	(10)	6	(6.5)	5	(4.5)
10	(12)	3	(2)	4	(3)

$\sum R$	=	62		32.5		25.5
n	=	5		5		5
\bar{R}	=	12.4		6.5		5.1
$(\sum R)^2/n$	=	768.8		211.25		130.05
$\sum(\sum R)$	=	120	Ellenőrzés: N(N+1)/2 = 120			
N	=	15				
$\sum((\sum R)^2/n)$	=	1110.1				

Null hipotézis: nincs különbség a lombozat különböző zónáiból begyűjtött lepkék számának medián értékei között.

$$H = \left(\frac{12}{N(N+1)} \right) \times \left(\sum \frac{(\sum R)^2}{n} \right) - 3(N+1) = \left(\frac{12}{15(15+1)} \times 1110.1 \right) - 3(15+1) = 7.505$$

H_{krit} a táblázatból ($\alpha = 0.05$, $n = 5$, $k = 3$) = 5.78

Mivel $H_{megfigyelt} > H_{krit}$ a null hipotézist visszautasítjuk.

Következtetés

Szignifikáns különbség van a fák lombzatának különböző zónáiból begyűjtött lepkék száma között ($P < 0.05$).

Korrekció alkalmazása a kötődött rangsorokra

Mivel két pár kötődött értékünk van, korrekciót alkalmazhatunk H-ra.

$$\sum t = \sum (t_i^3 - t_i) = (2^3 - 2) + (2^3 - 2) = 12$$

$$H_{korrigált} = \frac{H}{1 - \left[\frac{\sum t}{N^3 - N} \right]} = \frac{5.78}{1 - \left[\frac{12}{15^3 - 15} \right]} = 5.80$$

Következtetés

Szignifikáns különbség van a fák lombzatának különböző zónáiból begyűjtött lepkék száma között ($P < 0.05$).

Számítógépes értékelés GenStat programmal

Kruskal-Wallis one-way analysis of variance

Variate: Lepkeszam

Group factor: Lombkorona

Value of H = 7.505

Adjusted for ties = 7.545

Sample	Size	Mean rank
Group 1	5	12.40
Group 2	5	6.50
Group 3	5	5.10

Degrees of freedom = 2

Chi-square probability = 0.023

Ranks for individual variables

Group 1					
	15.0	14.0	11.0	10.0	12.0
Mean rank:	12.40				
Group 2					
	13.0	4.5	6.5	6.5	2.0
Mean rank:	6.500				
Group 3					
	8.5	8.5	1.0	4.5	3.0
Mean rank:	5.100				

17.4.2. Nem-paraméteres többszörös összehasonlítási teszt

Ha a Kruskal-Wallis analízis arra a következtetésre vezet, hogy szignifikáns különbségek vannak jelen, a különbségek pontos helyét nem-paraméteres többszörös összehasonlítási teszttel állapítjuk meg. A legegyszerűbb eljárás egy nem-paraméteres Tukey teszt használata, felhasználva a rangsor összegeket, hogy megállapítsuk a kritikus minimális különbséget, amelynek meg kell lennie két minta medián között ahhoz, hogy a szignifikáns különbséget igazoljuk. Az átlagos rangsorok közötti különbség standard hibája a következő:

$$SED = \sqrt{\frac{N(N+1)}{12} \left(\frac{1}{n_a} + \frac{1}{n_b} \right)}$$

A kritikus különbséget a következőképpen kapjuk meg: $Q_{krit} \times SED$, ahol Q_{krit} értéket a *Melléklet 7. táblázat* kritikus Q értékéből olvassuk ki, α választott valószínűségnél és k számú mintánál a kísérletben. Bármely minta-párra, ha a különbség az átlagos rangsoraik között nagyobb, mint a számított kritikus minimális különbség, akkor az a két minta szignifikánsan különbözik, a valószínűség megállapított szintjén. A példa illusztrálja a számítást és az eredmények megfelelő prezentációját, felhasználva a Kruskal-Wallis teszt adatait a példából.

Példa. Egy nem-paraméteres Tukey típusú többszörös összehasonlító teszt

Egy teszt annak meghatározására, hogy van-e szignifikáns különbség egy fa lomboszatának különböző zónáiból gyűjtött károsító lepkék számában. Az adatokat először Kruskal-Wallis teszttel analizáltuk:

$$SED = \sqrt{\frac{N(N+1)}{12} \left(\frac{1}{n_a} + \frac{1}{n_b} \right)} = \sqrt{\frac{15(15+1)}{12} \left(\frac{1}{5} + \frac{1}{5} \right)} = 2.82$$

Q_{krit} (táblázatból, $\alpha = 0.05$, $k = 3$) = 2.394

Kritikus minimális különbség az átlagos rangsorok között = $2.82 \times 2.394 = 6.75$

Következtetés

bármely minta-pár, melynek számszerű különbsége az átlagos rangsorok között ≥ 6.75 , szignifikánsan különbözik ($P < 0.05$).

A minták összehasonlítása:

Minták	$\bar{R}_a - \bar{R}_b$	Szignifikancia
Alsó lombzat – Középső lombzat	$12.4 - 6.5 = 5.9$	n.s.
Alsó lombzat – Felső lombzat	$12.4 - 5.1 = 7.3$	$P < 0.05$
Középső lombzat – Felső lombzat	$6.5 - 5.1 = 1.4$	n.s.

17.5. Randomizált blokk elrendezésű kísérlet nem-paraméteres analízise

Az esetek többségében feltételezzük a randomizált blokk elrendezésű kísérletekből származó adatokról, hogy normális eloszlásúak, homogén minta varianciával és analízisük a paraméteres ANOVA-val történik. Ahol ezek a feltételek nem teljesülnek, a leggyakoribb eljárás, hogy matematikailag transzformáljuk az adatokat úgy, hogy közelítsenek a normál eloszláshoz és/vagy a minta varianciák elfogadhatóan homogénné váljanak, és ezután ANOVA-val analizálhatók a transzformált adatok. Az egyszerű egytényezős randomizált blokk elrendezésű kísérlet esetében azonban rendelkezésre áll egy nem-paraméteres teszt, a **Friedman teszt**, melynek nincsenek feltételei az adatok eloszlására vonatkozóan.

Mint a legtöbb más nem-paraméteres analízisnél, a Friedman teszt is az adatok rangsorán alapul és ezáltal lényegében a minta mediánok közötti teszt. Feltételezzük, hogy minden kezelésnek egyenlő az ismétlése (n) mindegyik blokkban, és az adatok rangsorolva vannak a blokkon belül. A rangsorolást követően, a rangsorok összegét ($\sum R$) meghatározzuk minden kezelésre a blokkokon át és ezután kiszámítjuk a χ^2 statisztikai mutatót a következő formulával:

$$\chi_r^2 = \frac{12}{abn^2(na+1)} \sum (\sum R)^2 - 3b(na+1)$$

ahol: a = a kezelések száma;
 b = a randomizált blokkok száma;
 n = minden kezelés ismétléseinek száma blokkonként;
 $\sum(\sum R)^2$ = a rangsor összegek négyzet összege.

A χ_r^2 megfigyelt értékét összehasonlítjuk a táblázati χ_r^2 értékkel (Melléklet, 13. táblázat) az indokolt valószínűségnél, a kezelések (a) számára és a blokkok (b) számára. Ahol $\chi_{\text{megfigyelt}}^2 \geq \chi_{\text{kritikus}}^2$ a null hipotézist, vagyis hogy nincs különbség a kezelésekek között, elutasítjuk. Ahol az a és b értéke nagyobb, mint 5, illetve 10, a χ_r^2 eloszlása nagyon hasonlóvá válik a chi-négyzet (χ^2) teszt statisztikához, és a chi-négyzet valószínűségi eloszlás (Melléklet, 5. táblázat), $a-1$ szabadságfokkal használható, hogy megkapjuk a kritikus χ_r^2 értéket. A számításokat az alábbi példa mutatja.

Példa. Egy kísérletben vizsgálták négy takarmány kiegészítő hatását a sertések jó közérzeti viselkedésére. Mindegyik takarmány-kiegészítőt a négy sertés egyikének adták, és a sertéseket hat ismétlésben karámokban (blokkok) tartották. Feljegyezték annak gyakoriságát, amikor egy-egy sertés elkülöníthető, 1 percnél hosszabb ideig tartó túrásba kezdett, 12 órás időszak alatt (17.5. táblázat).

17.5. táblázat. Takarmány kiegészítők hatása a sertések túrási viselkedésének gyakoriságára

A túrási viselkedés gyakorisága 12 óra alatt					Rangsor értékek karámonként			
Kiegészítő takarmányok (a = 4)					Kiegészítő takarmányok (a = 4)			
Karámok (b = 6)	A	B	C	D	A	B	C	D
1	7	11	12	6	2	3	4	1
2	8	10	13	9	1	3	4	2
3	4	6	18	13	1	2	4	3
4	8	13	10	6	2	4	3	1
5	6	7	4	5	3	4	1	2
6	2	8	14	7	1	3	4	2
$\Sigma R =$					10	19	20	11
$(\Sigma R)^2 =$					100	361	400	121
$\Sigma(\Sigma R^2) =$					982			

Null hipotézis: nincs különbség a túrási viselkedés gyakoriságában a sertések között, melyeket négy különböző takarmány-kiegészítővel etettek.

Következtetés

$$\chi_r^2 = \frac{12}{abn^2(na+1)} \sum (\sum R)^2 - 3b(na+1) = \left(\frac{12}{6 \times 4 \times 5} \right) \times 982 - (3 \times 6 \times 5) = 8.2$$

$$\chi_{rkritikus}^2(\alpha=0.05) \text{ (táblázatból, } a=4, b=6) = 7.6$$

$\chi_r^2 > \chi_{rkritikus}^2$ ezáltal a null hipotézist elutasítjuk; a különböző takarmány-kiegészítővel etetett sertések szignifikáns különbséget mutattak a túrási viselkedés gyakoriságában ($P < 0.05$).

Számítógépes értékelés GenStat programmal

A menük kiválasztásának sorrendje: Stats – Statistical Tests – Friedman’s Nonparametric ANOVA.

Friedman’s test

Data variate: Turasszam
 Blocks: Ismetles
 Treatments: Takarmany

Based on 6 blocks of 4 treatments

Friedman’s statistic = 8.20

Adjusted for ties = 8.20

P-value using chi-square approximation (3 d.f.) = 0.042

Based on 3 degrees of freedom

Warning: P-value is approximate-check with values below if borderline.

5% point = 7.60

1% point = 10.00

Ismetles	Takarmany	Turasszam	Ranks
1	1	7.000	2.0
2	1	8.000	1.0
3	1	4.000	1.0
4	1	8.000	2.0
5	1	6.000	3.0
6	1	2.000	1.0
1	2	11.000	3.0
2	2	10.000	3.0
3	2	6.000	2.0
4	2	13.000	4.0
5	2	7.000	4.0
6	2	8.000	3.0
1	3	12.000	4.0
2	3	13.000	4.0
3	3	18.000	4.0
4	3	10.000	3.0
5	3	4.000	1.0
6	3	14.000	4.0
1	4	6.000	1.0
2	4	9.000	2.0
3	4	13.000	3.0
4	4	6.000	1.0
5	4	5.000	2.0
6	4	7.000	2.0

17.6. Paraméteres és nem-paraméteres módszerek összehasonlítása

A paraméteres eljárások számos feltételt állítanak a populációval szemben, amelyből az adatok származnak. Ezeket arra használjuk, hogy becsljük és összehasonlítsuk a populáció paramétereit és a populáció eloszlását (annak típusát) ismerni kell. Azon statisztikai eljárásokat, amelyekben a teszt alkalmazása nem függ a változók eloszlásától, nem-paraméteres próbáknak nevezik. A nem-paraméteres próbákat két csoportba sorolják. A valódi nem-paraméteres próbákban a próbastatisztika számításához nem szükséges a populáció egyetlen paraméterének (pl. átlagnak, varianciának) ismerete sem, mivel a próbastatisztika számítása egy eloszlás alapján történik (pl. illesztés-vizsgálat). A nem-paraméteres próbák másik csoportját az eloszlás-független eljárások alkotják. Az utóbbi csoportba tartozó próbák nem követelik meg a változók bizonyos típusú eloszlását, ilyenek pl. a rangok sorrendjén alapuló próbák (Précényi *et al.*, 1995).

A nem-paraméteres eljárások robusztusabbak, mint a paraméteres eljárások, mivel kevesebb feltételt tesznek az adatok természetére. Kevésbé erőteljesek, mint a paraméteres eljárások, ha az adatok alapjául szolgáló eloszlás megfelelően közel van egy standard eloszláshoz, mivel kevesebb információt vesznek az adatokból és így kevésbé valószínű, hogy feltárnak egy különbséget, ha az létezik. Ha az adatok transzformálhatók egy olyan eloszlássá, amely lehetővé teszi a paraméteres eljárások használatát, több információt nyerünk, mint a nem-paraméteres eljárás használatakor a nem transzformált adatokkal.

Ha a paraméteres módszer feltételei teljesülnek, akkor a nem-paraméteres tesztnek több megfigyelésre van szüksége, hogy ugyanolyan hatékony legyen. Ha a paraméteres módszerek valamennyi feltétele teljesül, akkor ezek a módszerek mindig erőteljesebbek, mint a nem-paraméteres megfelelőjük, jöllehet a nem-paraméteres módszerek közel megegyeznek a paraméteres módszerekkel kis minta méreteknél.

A nem-paraméteres tesztek eredményeit nem mindig olyan könnyű interpretálni, mint a paraméteres teszt eredményeit. Kevésbé megfelelőek komplex modellek tesztelésére, különösen, ha azok tartalmaznak hierarchikus adatstruktúrát vagy interakciókat. Ezek a módszerek azonban általában robusztusabbak és az eloszlások tág körét lefedik, míg számos paraméteres módszer igényli, hogy az adatok normális eloszlásúak legyenek. Következésképpen, nem-paraméteres módszerek hasznosak olyan adatok analízisére, amelyek nem normális eloszlásúak, vagy az eloszlások nem könnyen specifikálhatók vagy ismeretlenek.

A nem-paraméteres statisztikák használhatók számos adattípuson. Használhatók nominális adatokkal, mivel az ilyen adatok, természetüknél fogva, rendszerint megsértik a paraméteres módszerek feltételeit. Használhatók számolt adatokon és rangsorolt vagy rendezett adatokon, mint amilyenek a fertőzési pontszámok. Ha csak kategorizált adatokat lehet gyűjteni (pl. a földfelszín takarásának típusát lehet úgy osztályozni, mint gyeperdő), akkor a nem-paraméteres módszereket kell használni, mivel a paraméteres módszerek nem megfelelőek. Nem-paraméteres eljárásokat szükséges használni, ha a változót intervallumskálánál kisebb pontosságú skálán mértük vagy a változónk eloszlása nem elégíti ki a paraméteres próbák követelményeit.

Ha megfelelő kategorizált adatok gyűjthetők, akkor a nem-paraméteres teszt ereje jól megfelelő lehet, a gyenge mérési skála ellenére. Ha a folytonos adatok nem

megfelelőek a paraméteres eljárással történő analízisre, akkor a nem-paraméteres módszerek használhatók. Ez bekövetkezhet, amikor a mintaméret kicsi, vagy amikor az adatok ferde eloszlásúak. Az alábbi táblázat több jellemző alapján összehasonlítja a paraméteres és nem-paraméteres próbákat (*Collins és Seeney, 1999* alapján).

Jellemzők	Paraméteres	Nem-paraméteres
Feltételek	Sok	Kevés
Erő	Erős	Gyenge
Hipotézis	Paraméter alapú	Eloszlás alapú vagy eloszlás-független
Adattípusok	Kvantitatív	Minden típusú
A teszt eredmények interpretálása	Viszonylag könnyen értelmezhető	Bonyolult lehet
Adatstruktúra	Bármilyen	Korlátolt
Populáció	Ismerni szükséges	Lehet ismeretlen
Minta méret	Korlátok a minimális méretre	Lehet kicsi

18. Korrelációs számítás és regresszió-analízis

A növénytermesztési kísérletekben a változóknak három csoportját különböztetjük meg:

(1) **Kezelések**, mint amilyenek a műtrágya dózisok, fajták, és a gyomirtási módszerek, melyeket egy vagy több agrotechnikai gyakorlatból fejlesztünk ki, és a kísérlet elsődleges célja;

(2) **Környezeti faktorok**, mint amilyen a csapadék és a napsugárzás, amelyek a környezet részét képezik és nem állnak a kutatás szabályozása alatt.

(3) **Reakciók**, amelyek a kísérleti egységek biológiai és fizikai sajátosságait képezik, amelyekre a vizsgált kezelések várhatóan hatással vannak.

A *kezelésekre adott reakciók* kimutathatók vagy a növényen, olyan biológiai tulajdonságok változásában kifejezve, mint a szemtermés és a növénymagasság (*növény reakcióknak*, crop response hívjuk), vagy a közvetlen környezetben olyan változásokban, mint a rovarok jelenléte egy entomológiai kísérletben és talaj tápelem tartalma egy trágyázási kísérletben (*nem növényi reakciók*, noncrop response).

Mivel a mezőgazdasági kutatás elsődlegesen a biológiai szervezetek viselkedésére fókuszál egy specifikus környezetben, az *összefüggések a kezelések, a környezeti faktorok és a reakció között*, amelyeket rendszerint értékelünk a növénytermesztési kutatásban, a következők:

Összefüggés a reakció változók között. A növény teljesítménye (pl. termése) számos növényi és nem növényi tulajdonság terméke. Sorjában, mindegyikre hatással vannak a kezelések. Mindezeket a tulajdonságokat rendszerint egyidejűleg mérjük, és egymás közötti összefüggésük hasznos információt adhat arról, hogyan befolyásolják a kezelések a növény reakcióját. Például, egy kísérletben, ahol a növényesség hatását kívánjuk meghatározni a búza termésére, az összefüggés a termés és a terméskomponensek között, mint amilyen a produktív hajtások száma vagy a kalász tömege, egy jó indikátora a kezelések indirekt hatásának; a szemtermés nő a nagyobb hajtásszám, vagy a nagyobb kalázméret, vagy a kettő kombinációja eredményeként.

Egy másik példa, hogy egy nemesítési programban nagy terméssel és magas fehérjetartalommal rendelkező búzafajtákat kívánnak előállítani. Egy pozitív összefüggés a két jellemző között azt jelezné, hogy nagy terméssel és magas fehérjetartalommal rendelkező fajtákat könnyű találni, míg egy negatív összefüggés jelezné a kívánt fajták alacsony gyakoriságát.

Összefüggés a reakció és a kezelés között. Ha a kezelések kvantitatívak, mint a ha-onként kijuttatott kg műtrágya és a m^2 -enkénti növényesség, akkor lehetséges leírni az összefüggést a kezelés és a reakció között. Egy ilyen összefüggés meghatározásával, a kezelés és a reakció közötti összefüggést specifikáljuk nemcsak a ténylegesen vizsgált kezelésszintekre, hanem minden más közbenső pontra a vizsgált kezelések tartományán belül. Például, egy műtrágyázási kísérletben, a hol a növény termését

értékeljük 0, 30, 60 és 90 kg ha⁻¹ N dózissal, a termés és a nitrogén dózis közötti összefüggés specifikálja a termést nemcsak a négy ténylegesen vizsgált nitrogén dózusra, hanem minden más dózusra 0 és 90 kg ha⁻¹ N között.

Összefüggés a reakció és a környezet között. Ahhoz, hogy egy új növénytermesztési technológiát elfogadhassunk, előnyét fenn kell tartani eltérő környezetben. Tehát a mezőgazdasági kísérleteket rendszerint megismételjük eltérő környezetben vagy különböző években. Az ilyen kísérletekben, az összefüggés fontos a környezeti faktorok (napsugárzás, csapadék, hőmérséklet, talaj tápelem tartalma) és a növény reakciók között.

Ahhoz, hogy a tulajdonságok közötti összefüggést jellemezzük, olyan statisztikai eljárásokra van szükség, amelyek egyidejűleg több változót kezelnek. Ha két növényi tulajdonságot mérünk a növényi reakció jellemzésére, a varianciaanalízis és az átlag összehasonlítás (t próba) eljárásai csak egy tulajdonságot tudnak értékelni egy időpontban, még akkor is, ha a reakció az egyik tulajdonságban hatással lehet a másikra vagy a kezeléshatások egyidejűleg befolyásolják mindkét tulajdonságot. A *regresszió-analízis* és a *korrelációs számítás* lehetővé teszi a kutatóknak, hogy a fentiekben leírt három típusú összefüggésnek bármelyikét vagy kombinációját megvizsgálja, feltéve, hogy a változók kvantitatív formában vannak kifejezve.

A *regresszió-analízis* leírja egy vagy több változó (független változó) hatását egyetlen változóra (függő változó), kifejezve az utóbbit az előző függvényeként. Ehhez az analízishez fontos egyértelműen különbséget tenni a függő és független változó között, jóllehet a megkülönböztetés nem mindig nyilvánvaló. Például, amikor a kísérletben a termés reakcióját vizsgáljuk nitrogénre, a termés nyilvánvalóan a függő változó, és a nitrogén dózis a független változó. Másrészről, a szemtermés és a fehérjetartalom példában, a változók indentifikálása nem egyértelmű. Általában azonban a fontosabb tulajdonság, mondjuk a szemtermés lesz a függő változó és azok a faktorok vagy tulajdonságok, amelyek befolyásolják a szemtermést, válnak független változókká.

A *korrelációs számítás*, másrészről, az összefüggés mértékének mérőszámát adja két változó között, vagy egy összefüggés illesztésének jóságát.

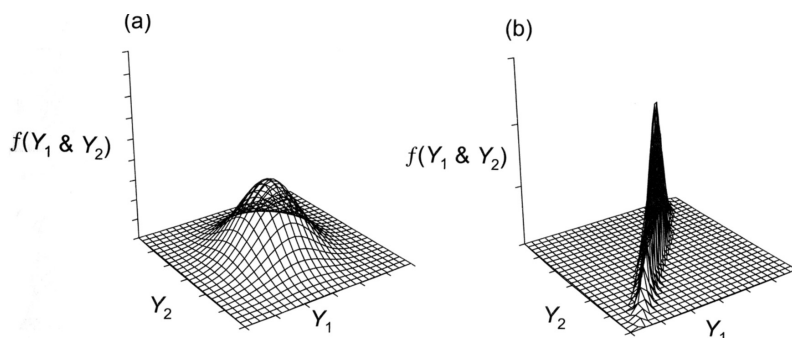
A regresszió- és korrelációanalízis osztályozható a változók száma és a függő és független változók közötti függvénykapcsolat formája alapján. Az analízist **egytényezősnek** mondjuk, ha csak két változó (egy függő és egy független) van bevonva és **többszörösnek** egyébként. Az analízist **lineárisnak** hívjuk, ha az összefüggés egyenes vonalú, és **nem-lineárisnak**, ha az összefüggés görbe vonalú. Tehát a regresszió- és korrelációanalízis négy típusba sorolható: egyszerű (egytényezős) lineáris regresszió és korreláció; egyszerű nem-lineáris regresszió és korreláció; többszörös lineáris regresszió és korreláció; többszörös nem-lineáris regresszió és korreláció. Közülük az első hárommal foglalkozunk röviden.

18.1. Korrelációs számítás

18.1.1. Bevezetés a korreláció különböző típusaihoz

A korrelációanalízis során azt vizsgáljuk, hogy két változó függ-e egymástól. Eltérően a regresszió analízistől, nem tételezünk fel köztük ok-okozati viszonyt és függvénykapcsolatot. Egyszerűen azt vizsgáljuk, hogy van-e a két változó között függő viszony. Ha a két változó nem független, akkor az egyik ismeretében jóslásokat tehetünk a másikra. A változók közötti viszony egyenrangú, nem különböztetünk meg függő és független változót.

A korrelációanalízis feltétele, hogy a két változó együttes eloszlása kétdimenziós normális eloszlás legyen. Ha a két változó között laza korreláció van, akkor a kétdimenziós normál eloszlás képe közel harang alakú (18.1.a. ábra). Amennyiben szoros kapcsolat van köztük, akkor az eloszlás képe torzul, elnyúlt taraj alakúvá válik (18.1.b. ábra). A taraj annál élesebb, minél erősebb a függő viszony a két változó között.

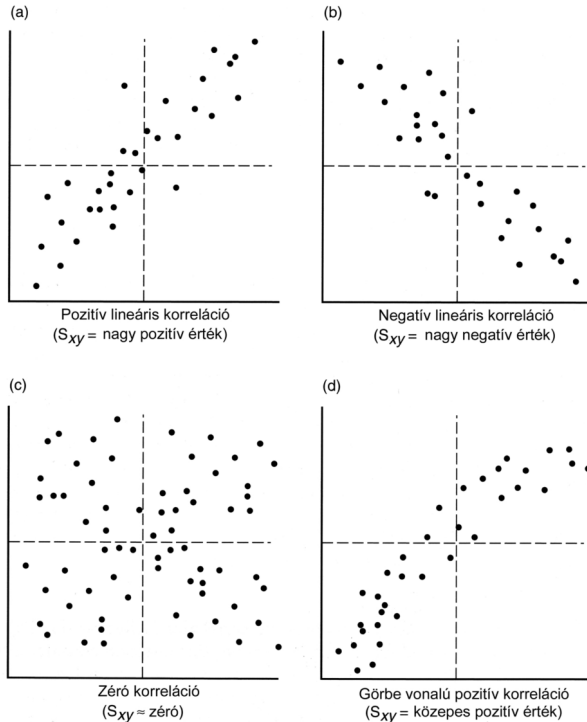


18.1. ábra. Bivariáns normális eloszlás (a) két változóra laza korreláció esetén és (b) két változó szoros pozitív korrelációjakor (Quinn és Keough, 2010)

A biometriában a korrelációt specifikusan alkalmazzuk olyan helyzetekben, ahol az egyik változó kvantitatív értéke összefügg egy másik változó kvantitatív értékével. Például, egy növénytermesztési kísérletben szükséges lehet megállapítani, hogy van-e korreláció a búza szemtermése és a kalászolás utáni levélterület tartósság (LAD) között. Egy állattenyésztési kísérletben szükséges lehet meghatározni, hogy egy állat tömeggyarapodása milyen mértékben korrelál az elfogyasztott takarmány fehérje tartalmával.

Két kvantitatív változó többféle módon korrelálhat. Leggyakrabban, két változó pozitív korrelációt mutathat, ahol egyik változó növekedése tendenciában arányos a másik változó növekedésével. Gyakran azonban, egyik változó növekedése egy másik változó csökkenését eredményezi, és az ilyen változók ezáltal negatív korrelációt mutatnak. Például, az összefüggés az állati takarmány lignintartalma és az emészthetőség között feltételezhetően egy negatív korreláció. Ha az ilyen változókat egy grafikonon ábrázoljuk, egy egyenes vonalat kapunk és azt mondjuk, hogy lineárisan korrelálnak. Gyakran azonban, két szorosan korreláló változó, amikor grafikonon ábrázoljuk, görbe-vonalú összefüggést mutat az egyenes helyett. A korreláció más típusai magukba foglalják a biológiában a log-lineáris összefüggést, továbbá az

exponenciális összefüggést. Amikor két változót grafikusán ábrázolunk egymás ellenében, az ábrát szórásvázlatnak hívjuk és az ábrázolt adatértékeket pedig bivariánssoknak. Különböző szórásvázlatokat mutat a 18.2. ábra. Míg gyakran nyilvánvalóan látszik egy szórásvázlaton, amikor két változó korrelál, számos esetben azonban nem biztos, hogy valódi korreláció áll fenn két változó között, továbbá szükséges lehet becsülni a korreláció szorosságát.



18.2. ábra. A szórásvázlat mutatja két változó közötti korreláció különböző típusait. Minden ábrázolt adatpont két változó méréséből származik és bivariáns értéknek hívjuk

18.1.2. Pearson-féle korrelációs koefficiens (r)

A kapcsolat erősségének a mérésére szolgál az ún. *Pearson-féle* korrelációs együttható (r). Értéke -1 és $+1$ között változhat. Nulla érték a két változó közötti kapcsolat hiányát, míg -1 ill. $+1$ a teljes meghatározottságot jelzi. Az előjel a *kapcsolat irányát* mutatja: pozitív együttható esetén az egyik változó növekedésével nő a másik változó, míg negatív előjelnél az egyik változó növekedésével a másik csökken.

A Pearson korrelációs koefficient r szimbólummal jelöljük és a következő formulával számítjuk ki:

$$r = \frac{\sum (x - \bar{x})(y - \bar{y})}{\sqrt{\sum (x - \bar{x})^2 \times \sum (y - \bar{y})^2}}$$

Kissé egyszerűbb jelölés használható, ha a számlálóban lévő eltérésszorzat összeget SP-vel (S_{xy} -nal, 18.2. ábra) jelöljük, továbbá a $\sum (x - \bar{x})^2$ és $\sum (y - \bar{y})^2$ kifejezéseket a nevezőben, amelyek x és y eltérésnégyzet összegei, SS_x és SS_y jelöli. Az r korrelációs koefficiens ekkor így fejezhető ki:

$$r = \frac{SP}{\sqrt{SS_x \times SS_y}}$$

Minél nagyobb az r abszolút értéke, azaz minél közelebb van $+1$ -hez vagy -1 -hez, annál erősebb a korreláció a két változó között. Pontosan $+1$ érték tökéletes pozitív korrelációt jelez, míg a pontosan -1 érték tökéletes negatív korrelációt jelez.

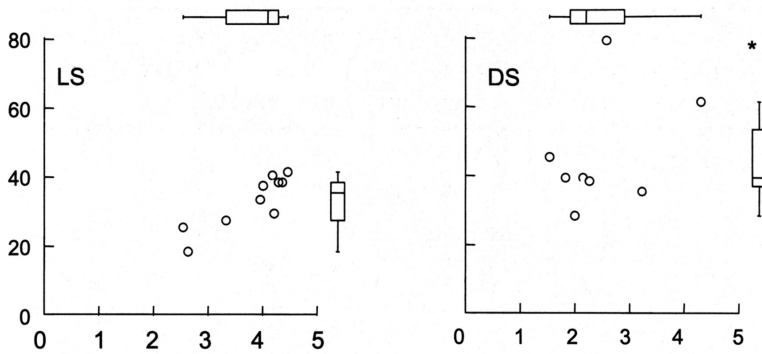
Példa. Vörös szárazföldi rákok a Karácsony Szigeten

A Karácsony Sziget az északkeleti Indiai Óceánnál híres az őshonos (endemikus) vörös szántóföldi rákjairól (*Gecarcoidea natalis*), amelyek egy sajátos tömeges vándorlással visszatérnek a tengerbe minden évben, hogy ott elhelyezzék tojásaikat. A rákok a sziget őserdőiben élnek, ahol fák magvaival táplálkoznak. A rákok ökológiájával foglalkozó egyik tanulmányban, Green (1997) azt vizsgálta, hogy van-e összefüggés a vörös szántóföldi rákok összes biomaszja tömege és a földbe ásott lyukak sűrűsége között 25 m²-es kvadráton belül (mintavételi egység) öt őserdő helyen a szigeten. Két ilyen helyet vizsgálunk meg: tíz kvadrát volt a Lower Site (LS) és nyolc kvadrát a Drumsite (DS) nevű helyeken. A szórásábrákat és a dobozábrákat a 18.3. ábra szemlélteti. Enyhe negatív ferdeség volt biomaszára és lyuk sűrűségére az LS helyen és kiugró adat volt a lyuk sűrűségére a DS helyen, de nem volt bizonyíték a linearitás hiányára. A Pearson-féle korreláció megfelelő volt ezekre az adatokra, jóllehet más korrelációkat is kiszámítottak összehasonlításképpen (Quinn és Keough, 2002).

Hely	Korreláció típusa	Mutató értéke	P (valószínűség) értéke
DS (n=8)	Pearson	0.392	0.337
	Spearman	0.168	0.691
	Kendall	0.036	0.901
LS (n=10)	Pearson	0.882	0.001
	Spearman	0.851	0.002
	Kendall	0.719	0.004

A H_0 hipotézist, hogy nincs lineáris összefüggés a rákok összes biomaszája és a lyukak száma között DS jelű helyen, nem lehet visszautasítani. Ugyanez a következtetés érvényes a Spearman- és Kendall-féle koefficiensekre. (A Spearman korrelációval részletesen foglalkozunk a 18.1.3. fejezetben.) Tehát nem volt bizonyíték a lineáris vagy még általánosabb monoton összefüggésre a lyuk sűrűség és a rákok összes biomaszája között a DS jelű helyen.

Visszautasítjuk a H_0 nullhipotézist, vagyis lineáris összefüggés hiányát a rákok összes biomaszája és a lyukak száma között, az LS jelű helyen. Ugyanez a következtetés vonatkozik a Spearman- és Kendall-féle koefficiensekre. Erős bizonyíték van a lineáris és még általánosabb monoton összefüggésre a lyuksűrűség és a rákok összes biomaszája között az LS helyen.



18.3. ábra. A szórásdiagram mutatja az összefüggést a vörös rákok földbe ásott lyukainak száma és az összes rák biomassza között 25 m² kvadráton két helyen (LS, DS) a Karácsony-szigeten (Green, 1997). Mindegyik ábra peremén a doboz ábrák vannak mindegyik változóra külön-külön.

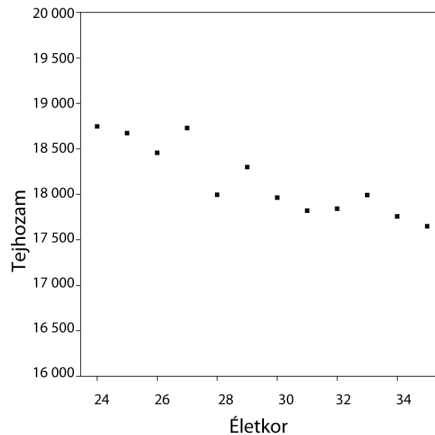
A korrelációs koefficiens kiszámításának lépéseit az alábbi példán keresztül mutatjuk be.

Példa. Pearson korrelációs koefficiens (r) meghatározása (18.1. táblázat) (Ireland, 2010 alapján)

18.1. táblázat Összefüggés a tejelő tehének élettartam alatti tejhozama és a tehének első elléskori életkora között

Életkor az első borjú születésekor (hónap)	Élettartam alatti tejhozam (kg)
24	18747
25	18673
26	18456
27	18730
28	17995
29	18300
30	17964
31	17820
32	17842
33	17991
34	17758
35	17650

Az adatok ábrázolása szórásdiagramon egyértelműen negatív korrelációt mutat (18.4. ábra):



18.4. ábra. A tehenek élettartam alatti tejhozama és a tehenek első elléskori életkora közötti összefüggés szórásdiagramja

18.1.2.1. Az r korrelációs koefficiens meghatározása

$$\begin{aligned} \sum X &= 354 & \sum Y &= 217926 & n &= 12 \\ \sum X^2 &= 10586 & \sum Y^2 &= 3959417604 & \sum XY &= 6414388 \\ SP &= \sum XY - \frac{(\sum X)(\sum Y)}{n} = -14429 \\ SS_x &= \sum X^2 - \frac{(\sum X)^2}{n} = 143; & SS_y &= \sum Y^2 - \frac{(\sum Y)^2}{n} = 1772481 \\ r &= \frac{SP}{\sqrt{SS_x \times SS_y}} = -0.906 \text{ (d.f. = 10)} \end{aligned}$$

Következtetés: a magas negatív érték egy viszonylag erős negatív korrelációt jelez az első elléskori élettartam és az élettartam alatti tejhozam között.

18.1.2.2. A korrelációs koefficiens (r) megbízhatóságának meghatározása

Mivel az r korrelációs koefficiens a valódi populáció korrelációs koefficiensének (ρ) egy becslése, jogosan feltehetjük a kérdést, mennyire megbízható r, mint egy becslés. Ennek megválaszolására lehetséges meghatározni a minta korrelációs koefficiensének a standard hibáját és konfidencia határait. A következőkben tárgyalt r tesztek ugyanazon megoldás különböző oldali megfogalmazását jelentik.

Az r standard hibája

A minta korrelációs koefficiensének standard hibáját (SE_r) a következőképpen számítjuk ki:

$$SE_r = \sqrt{\frac{1 - r^2}{n - 2}}$$

ahol: r = a minta korrelációs koefficiens; n = az x, y adat-párok száma.

Az r konfidencia határai

Számos alternatív módszer van a konfidencia határok meghatározására, a legáltalánosabb az F-mutató kétoldali kritikus értékét használja fel.

Tételezzük fel, hogy 95%-os konfidencia határt akarjuk megkapni, akkor az alsó konfidencia határt ($CL_{\text{alsó}}$) megkapjuk:

$$CL_{\text{alsó}} = \frac{(1 + F_{0.05})r + (1 - F_{0.05})}{(1 + F_{0.05}) + (1 - F_{0.05})}r$$

és a felső konfidencia határt ($CL_{\text{felső}}$) megadja:

$$CL_{\text{felső}} = \frac{(1 + F_{0.05})r - (1 - F_{0.05})}{(1 + F_{0.05}) - (1 - F_{0.05})}r$$

ahol: r = a minta korrelációs koefficiense;

$F_{0.05}$ = a kétoldali F értéke $P = 0.05$ -nél és $V_1 = V_2 = n - 2$ szabadságfoknál.

A standard hibának, a konfidencia határoknak, a lineáris korreláció szignifikanciájának és a determinációs koefficiensnek a meghatározása

A korrelációs koefficiens standard hibája (SE), valamint alsó és felső 95%-os konfidencia határai (95% CL):

$$SE_r = \pm \sqrt{\frac{1 - r^2}{n - 2}} = \pm \sqrt{\frac{1 - 0.906^2}{12 - 2}} = \pm 0.13$$

$$95\% CL_{\text{alsó}} = \frac{(1 + 3.72)0.906 + (1 - 3.72)}{(1 + 3.72) + (1 - 3.72)}0.906 = 0.690$$

$$95\% CL_{\text{felső}} = \frac{(1 + 3.72)0.906 - (1 - 3.72)}{(1 + 3.72) - (1 - 3.72)}0.906 = 0.974$$

18.1.2.3. A korreláció szignifikanciájának tesztelése

Korrelációs koefficiens (r) tesztje

Az r kritikus értékét táblázatból (Melléklet 4. táblázat) kapjuk meg, a valószínűség adott szintjén és a megfelelő szabadságfoknál. A szabadságfokok számát az adatpárok száma (n), mínusz a korreláló változók száma, azaz $n - 2$ adja meg. A korreláció null hipotézisét visszautasítjuk, ahol $r > r_{\text{kritikus}}$ és levonható az a következtetés, hogy szignifikáns lineáris korreláció van a két változó között, α valószínűségi szinten. Jegyezzük meg, hogy a tesztet rendszerint kétoldali tesztként végezzük, és ha a korreláció negatív, az előjeltől eltekintünk a szignifikancia vizsgálatnál.

Összefüggés a tejelő tehenek élettartam alatti tejhozama és a tehenek első elléskori életkora között (1. példából)

$$r = 0.906 \quad (n = 12) \quad r_{\text{kritikus}}(0.05, DF = 10) = 0.576 \quad (\text{táblázati érték})$$

Következtetés: az r abszolút értéke nagyobb, mint a kritikus érték. A null hipotézist visszautasítjuk és levonjuk azt a következtetést, hogy szignifikáns negatív korreláció van az első ellés életkora és az élettartam alatti tejtermelés között ($P < 0.05$).

Student t-teszt

A Student t-tesztet megkapjuk, ha a korrelációs koefficiens elosztjuk a standard hibával, SE_r -rel:

$$t_{\text{megfigyelt}} = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}}$$

A $t_{\text{megfigyelt}}$ értéket a táblázati Student t-értékhez viszonyítjuk (Melléklet 1. táblázat), a választott kritikus valószínűségeknél és $n - 2$ szabadságfoknál. Ha $t_{\text{megfigyelt}} > t_{\text{kritikus}}$, a null hipotézis visszautasítható és levonható az a következtetés, hogy elegendő bizonyíték támogatja a szignifikáns lineáris korreláció érvényességét.

A korreláció Student t-tesztje:

$$t_{\text{megfigyelt}} = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}} = \frac{-0.906}{\sqrt{\frac{1-0.906^2}{12-2}}} = -6.769$$

$$t_{\text{krit}(0.05, DF=10)} = 2.228 \text{ (táblázati érték)}$$

Következtetés: a $t_{\text{megfigyelt}}$ abszolút értéke nagyobb, mint a kritikus érték. A null hipotézist visszautasítjuk és levonjuk azt a következtetést, hogy szignifikáns negatív korreláció van az első ellés életkora és az élettartam alatti tejhozam között ($P < 0.05$).

Determinációs koefficiens, r^2

Miután megállapítottuk, hogy szignifikáns lineáris korreláció áll fenn két változó között, gyakran hasznos meghatározni továbbá a változásnak az arányát az egyik változóban, amely közvetlenül felelős a másik változóval való korrelációért. Nagyon ritkán fordul elő, hogy az egyik változóban bekövetkezett változás teljes mértékben a másik változóban bekövetkezett változásnak tulajdonítható és más faktor nem befolyásolja az összefüggést. A példában feltételezhető, hogy a tehének élettartam alatti tejhozamát számos más faktor befolyásolja az első ellés életkorán kívül, pl. a takarmányozás, a tartási körülmények és random változók. Abból a célból, hogy becsüljük annak mértékét, hogy egyik változó változása mennyiben tulajdonítható a másik változóban bekövetkezett változásnak, meghatározzuk a determinációs koefficiens (r^2). Értéke 0 és +1 közé esik, kifejezhető százalékban is.

Az első ellés életkora és az élettartam alatti tejhozam közötti összefüggés determinációs koefficiense:

$$\text{Determinációs koefficiens} = r^2 = (-0.906)^2 = 0.821$$

Következtetés: az élettartam alatti tejhozam variációjának 82.1%-a megmagyarázható az első ellés életkorával való korrelációval; a maradék más, nem azonosított faktoroknak tulajdonítható.

18.1.2.4. Nem-lineáris korrelációk analízisének problémája

A korrelációs számítás csak azokat az összefüggéseket tárja fel megbízhatóan, amelyek lineárisak. Amikor a két változót egymástól függően ábrázoljuk szórásdiagramon, gyakran szoros összefüggés nyilvánvaló, azonban nem lineáris, hanem görbe-vonalú. Mint már említettük, log-lineáris és exponenciális összefüggések különösen általánosak a biológiai tudományokban és görbe vonalú reakciót mutatnak egy szórásdiagramon. Más összefüggések lineáris reakciót mutathatnak az adatoknak egy részére, azonban a lineáris megtörhet, ahogy a változók értéke nő. Például az összefüggés a növény levélterülete és termése között rendszerint lineáris egy bizonyos pontig és ezután fokozatosan görbe vonalú lesz, ahogy a levélterület eléri egy optimumot és ténylegesen csökkenhet a levélterület további növelésekor. Az ilyen összefüggéseket *görbe vonalúnak* hívjuk. Az ilyen adatok korrelációanalízise csak gyenge vagy nem szignifikáns korrelációt tárhat fel, amikor nyilvánvaló a változók tényleges szoros összefüggése.

Ilyen esetekben megfelelő analízálni csupán az adatoknak egy részét, amelyek egyes vonalra esnek egy szórásdiagramon. Alternatív módon, a görbe vonalú összefüggések matematikailag átalakíthatók lineárisra *adat-transzformációval*, melynek során az egyik vagy mindkét változót átalakítjuk. Például, görbe vonalú adatok könnyen lineárisra alakíthatók \log_{10} transzformációval. Ha adat-transzformációval nem oldható meg a probléma, egy másik lehetséges megközelítés az adatanalízis nem-paraméteres módszerének alkalmazása, melyet a következőkben ismertetünk.

18.1.2.5. Korrelációs számítás GenStat programmal

A korrelációs számításhoz a kiválasztás sorrendje a következő: Stats – Summary Statistics – Correlations. A változók az Available Data ablakban jelennek meg, ahonnan a változóra kattintással a Data ablakba kerülnek. Választhatunk Correlation és Partial Correlation között. A korreláció tesztelésére a One-sided (egyoldali próba) és Two-sided (kétoldali próba) lehetőségek között választhatunk. A parciális korrelációt kettőnél több változó esetén két változó közötti r jellemzi, amennyiben ezen két változó összefüggéséből a többi változó hatását kiiktatja.

Correlations

Eletkor	1	–	
Tejhozam	2	–0.9063	–
		1	2

Number of observations: 12

Two-sided test of correlations different from zero

Eletkor	1	–	
Tejhozam	2	<0.001	–
		1	2

18.1.3. Nem-paraméteres korrelációs számítás: Spearman rangkorreláció

Spearman rangkorrelációja egy alternatív nem-paraméteres módszert képvisel a korrelációs számításban. A legtöbb nem-paraméteres statisztikai analízishez hasonlóan, az adatok rangsorán alapul. Következésképpen, nem alapul olyan feltételeken, amelyek az adatértékek eloszlására vonatkoznak, és az sem feltétel, hogy a korreláció szükségszerűen lineáris legyen. Ezáltal, különböző helyzetek tág tartományában alkalmazható, beleértve mind a számolt és mért adatok analízisét, valamint a rangsorolt (azaz ordinális) adatokat. Jóllehet biztonságosan használható normális eloszlású adatoknál, mégis ha a változók normális eloszlásúak és lineáris összefüggés várható, akkor a Spearman rangkorrelációnak kisebb a statisztikai ereje, mint a Pearson korrelációs számításnak, mivel kisebb a valószínűsége szignifikáns korreláció feltárásának, amikor az valós.

Az analízis célja egy statisztikai vizsgálat elvégzése az adatok rangsora alapján, amely felhasználható arra, hogy leírjuk és meghatározzuk két változó közötti korreláció szignifikanciáját. Először, mindegyik változó adatait külön-külön rangsoroljuk és meghatározzuk a különbséget (d) a rangsorban mindegyik x és y adat-pár között. A teszt feltételezése viszonylag egyszerű: ha két változó szorosan korrelál, mindegyik adat-pár x értékének és y értékének hasonló rangsora van; minél gyengébb a korreláció, annál nagyobb lesz a különbség mindegyik x , y adat-pár rangsor értékében. Ha a különbségeket (d) minden x , y adat-pár rangértéke között négyzetre emeljük, hogy eltávolítsuk a negatív értékeket és ezután összegezzük, ($\sum d^2$) értéket kapjuk meg, amely összefügg a korreláció fokával. Egy tökéletes korrelációnál, mindegyik x -nek és a megfelelő y értéknek pontosan ugyanaz lesz a rangsora, és $\sum d^2$ ezáltal zéró lesz. A $\sum d^2$ legnagyobb értékét egy tökéletes negatív korrelációnál kapjuk. Ezt illusztrálja a 11.5. ábra.

A $\sum d^2$ megfelelő standardizálása után megkapjuk a Spearman rangkorrelációt (r_s), melynek értéke -1 (tökéletes negatív korreláció) és $+1$ (tökéletes pozitív korreláció) fordul elő. Képlete:

$$r_s = 1 - \frac{2 \sum d^2}{(n^3 - n)} = 1 - \frac{6 \sum d^2}{n^3 - n}$$

ahol: d = különbség a rangsorban minden x , y adat-pár között; n = x , y adatpárok száma.

Az r_s kritikus értékei a *Melléklet 11. táblázatban* találhatóak. A táblázat a kritikus értékeket eltérő valószínűségi szinteken és mintaméretnél adja meg. Ha abszolút értékben $r_s \geq$ kritikus érték, akkor a null hipotézist (nincs korreláció) visszautasítjuk. A Spearman rangkorreláció kiszámítását az alábbi példa mutatja.

Példa. Spearman rangkorreláció (r_s) meghatározása

Korreláció a vizes oldat elektromos vezetőképessége (konduktivitása) és a babmagok csírázási rátája között (18.2. táblázat) (Ireland, 2010 alapján)

18.2. táblázat. A vizes oldat elektromos vezetőképessége és a bab csírázása közötti összefüggés

Magtétel	Konduktivitás (μs)	Rangsor	Csírázás (%)	Rangsor	d	d ²
1	2100	19.5	6	5	14.5	210.25
2	900	11	38	12	-1	1
3	550	3	60	21.5	-18.5	342.25
4	910	12	40	13	-1	1
5	225	1	58	19.5	-18.5	342.25
6	1480	15	18	8	7	49
7	750	8	49	14	-6	36
8	885	10	56	17	-7	49
9	2800	23	0	1	22	484
10	565	5	64	23	-18	324
11	1625	17	21	9	8	64
12	1510	16	24	10	6	36
13	670	7	52	16	-9	81
14	640	6	57	18	-12	144
15	2950	24	2	2.5	21.5	462.25
16	500	2	60	21.5	-19.5	380.25
17	2100	19.5	9	6	13.5	182.25
18	1750	18	12	7	11	121
19	930	14	32	11	3	9
20	825	9	58	19.5	-10.5	110.25
21	2350	21	5	4	17	289
22	2395	22	2	2.5	19.5	380.25
23	555	4	71	24	-20	400
24	915	13	50	15	-2	4
n = 24	$\sum d^2 = 4502$					

A mag öregedésével, a magrészek és a belső membránok degradációja a magból ionok kibocsátásához vezet, amely növeli a folyadék elektromos vezetőképességét. Ez egy lehetséges laboratóriumi tesztet ad a mag életképességének meghatározására. Abból a célból, hogy megállapítsuk a módszer érvényességét (validitását), számos különböző korú magot merítettek el vízben és mérték az oldatok elektromos konduktivitását. Ezt követően meghatározták ezeknek a magtételnek a csírázási rátáját.

Null hipotézis: nincs korreláció az oldat konduktivitása (vezetőképesség) és a bab magok csírázási rátája között

Spearman rangkorreláció (r_s) kiszámítása:

$$r_s = 1 - \frac{6 \sum d^2}{n^3 - n} = 1 - \frac{6 \times 4502}{24^3 - 24} = -0.96$$

$r_{s \text{ kritikus}}$ a táblázatból ($\alpha = 0.05, n = 24$) = 0.409

Következtetés

Szignifikáns negatív korreláció van a csirázási ráta és az oldat konduktivitása között ($P \leq 0.05$)

18.1.3.1. Spearman rangkorreláció számítás GenStat programmal

Az analízis sorrendje a következő: Stats – Statistical Tests – Nonparametric Correlations. A legördülő dobozból kiválasztjuk a Spearman's Rank Correlation-t. A változók az Available Data ablakban jelennek meg, ahonnan rákattintással vizsgáljuk a változókat a List of data ablakba. A Display által felkínált lehetőségek közül választhatunk: Test, Ranks és Correlations.

Spearman's rank correlation coefficient

Correlation -0.957 , adjusted for ties -0.959

Sample size 24

Exact probability < 0.001 (approximate as ties ignored in calculation)

t approximation -15.89 on 22 d.f., probability < 0.001

Ranks for individual variables

Konduktivitás	Csirázás
19.5	5.0
11.0	12.0
3.0	21.5
12.0	13.0
1.0	19.5
15.0	8.0
8.0	14.0
10.0	17.0
23.0	1.0
5.0	23.0
17.0	9.0
16.0	10.0
7.0	16.0
6.0	18.0
24.0	2.5
2.0	21.5
19.5	6.0
18.0	7.0
14.0	11.0
9.0	19.5
21.0	4.0
22.0	2.5
4.0	24.0
13.0	15.0

18.2. Lineáris regresszió-analízis

Számos kísérletben egy változó szintjeit szabályozzuk és mérjük a reakció változót abból a célból, hogy megállapítsuk az összefüggést a két változó között és ezáltal lehetőség legyen a jövőbeni reakció előrejelzésére. Például, mérhetjük egy feldarabolt vessző gyökeresedését egy alkalmazott növekedési hormon eltérő koncentrációjára adott reakcióban vagy egy állati takarmány emészthetőségét mérhetjük a lignin tartalom változtatására adott reakcióban. Egy független változó és egy reakció változó közötti összefüggés származhat megfigyelési tanulmányokból, mint amilyen a tehén kora az első borjú születésekor és az élettartam alatti tejhozam közötti összefüggés.

Egy ilyen vizsgálatot követve, rendszerint szükség lesz ábrázolni a reakciót a szabályozott vagy független változóval szemben egy szórásdiagramon, hogy megállapítsuk, létezik-e összefüggés. Ezután ábrázoljuk a legjobban illeszkedő trend egyenest az adat-párokon keresztül, hogy bemutassuk az összefüggés formáját.

Ha az összefüggés a két változó között lineáris, akkor az egyszerű lineáris regresszió-analízisnek hívott eljárás használható az összefüggés modellezésére és a legjobban illeszkedő egyenes egyenletének meghatározására, melyet ábrázolhatunk a kétváltozós adatokon keresztül egy szórásdiagramban. A legjobban illeszkedő egyenest a lineárisan összefüggő pontok között, melyet ezzel az eljárással kapunk meg, regressziós egyenesnek hívunk.

A bonyolultabb esetekben, egy mért változót, mint amilyen a növény termése, meghatározhatja számos különböző egyidejűleg ható változóra adott reakció, mint pl. a talajban levő különböző tápelemek szintjei. Feltételezve, hogy az összefüggés lineáris marad, a többszörös regresszió-analízist alkalmazhatjuk, hogy kapjunk egy olyan modellt, amely leírja az összefüggést a függő változó és több független változó kombinációja között.

18.2.1. A kétváltozós lineáris modell

A statisztikai analízisben a lineáris kifejezés az analízis alapjául szolgáló matematikai egyenlet formájára vonatkozik, sokkal inkább, mint a nyilvánvaló összefüggésre a mért változók között. A biológusok általában úgy tekintik a lineáris összefüggést, hogy bizonyít egy helyzetet, ahol a kvantitatív reakció egy faktorra lineáris és egy egyenes vonalat eredményez a grafikonon. Minden egyenes vonal grafikonon ábrázolva, leírható egyszerű lineáris függvényvel:

$$y = \alpha + \beta x$$

ahol: α az egyenes metszéspontja az Y-tengellyel és β az egyenes meredeksége.

Ahol a két változó között függvényszerű lineáris összefüggés van, mindegyik x , y adat az egyenes vonalra esik; azonban a valóságban a random variáció általában azt okozza, hogy az értékek szóródnak az adatokhoz legjobban illeszkedő egyenes bármelyik oldalán, mint azt az 18.5. ábra mutatja. Ezáltal, ahol lineáris összefüggés áll fenn két változó között, az összefüggés a mért Y változó és a független X változó között kifejezhető úgy, mint:

$$y_i = \alpha + \beta x_i + \varepsilon_i$$

ahol: y_i = az Y változó értéke, amely bekövetkezik az x változó x_i értékére adott reakcióként;

α = az egyenes metszéspontja az Y-tengellyel (azaz, y előrejelzett értéke, amikor $x = \text{zéró}$);

β = az egyenes meredeksége;

ε_i = random hiba az y_i mérésekor.

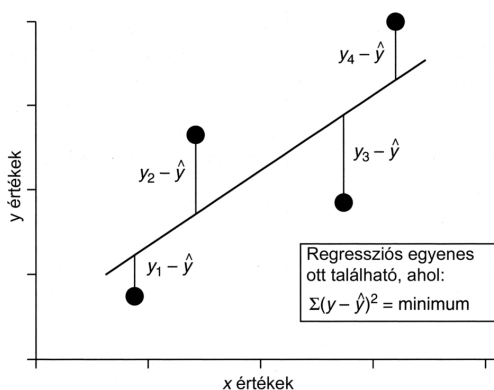
A minta regressziós egyenlete, amely lehetővé teszi, hogy előrejelezzük y értékét az x adott értékére:

$$\hat{y} = a + bx$$

ahol: \hat{y} = a függő Y változó előrejelzett értéke az X változó adott x értékére;

a = az egyenes becsült metszéspontja az Y-tengellyel;

b = az egyenes becsült meredeksége.



18.5. ábra. A regressziós egyenes meghatározása a legkisebb négyzetek módszerével

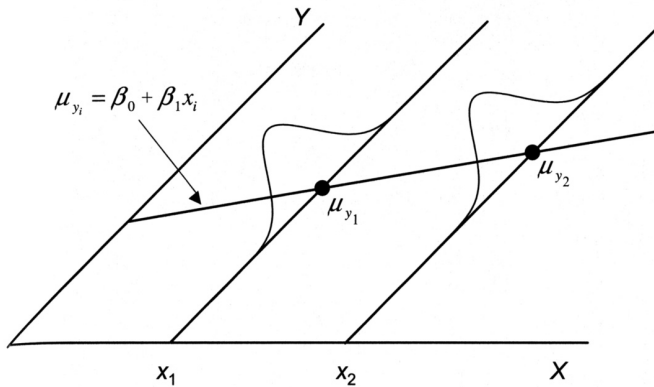
Az egyszerű lineáris regresszió-analízis célja, hogy meghatározzuk a a és b értékeit a minta adatokból, amelyek megadják az α és β populáció paraméterek legjobb becslését és ezáltal meghatározzuk a két változó között legjobban illeszkedő egyenes egyenletét.

Kétváltozós lineáris regresszió a legkisebb négyzetek módszerével

A leggyakoribb módszer, melyet a regressziós egyenes meghatározására használunk, a legkisebb négyzetek módszere (18.5. ábra). A módszer számos feltételen alapul és csak akkor használható, ha bizonyos körülmények fennállnak. Ezek a következők:

1. Az egyik változót függetlennek tekintjük és a horizontális X-tengelyen ábrázoljuk. A másik változó, amely az előző változására adott reakciót méri, a függő változó és a függőleges Y-tengelyen ábrázoljuk.

2. Egy határozott összefüggésnek kell fennállnia az Y változó és az X változó között (jóllehet ez nem szükségszerűen foglalja magában, hogy az összefüggés okozati). Ezen okból a Pearson-féle korrelációs számítást rendszerint elvégezzük a regressziós egyenlet meghatározását megelőzően.



18.6. ábra. A lineáris regressziós modell mutatja a populáció y_i értékét az x_i két értékénél.

A 18.6. ábra a populáció lineáris regressziós modell néhány fontos tulajdonságát mutatja:

1. Az Y -értékek mindenegyes x_i -nél normális eloszlásúak.
2. Az Y értékekről mindenegyes x_i -nél feltételezzük, hogy ugyanaz a varianciájuk (σ^2), melyet a variancia homogenitásának hívunk.
3. A populáció regressziós egyenese egyesíti az y -értékek átlagát

A legkisebb négyzetek módszere egy parametrikus eljárás, melyben a populáció regressziós egyenesének meredekségét a minta adatokból becsüljük, hasonlóan a parametrikus statisztikai eljárásokhoz. Ha a fenti feltételek bármelyike hiányzik, a legkisebb négyzetek módszere érvénytelenné válik.

18.2.2. A regressziós egyenlet meghatározása: a legkisebb négyzetek módszere

A legkisebb négyzetek módszere minden ábrázolt adatnak a legjobban illeszkedő egyenestől való vertikális eltérésén alapul. A legjobban illeszkedő egyenest ott határozza meg, ahol ezek az eltérések a lehető legkisebbek. Definíció szerint azonban a legjobban illeszkedő egyenestől a vertikális eltérések fele pozitív és fele negatív lesz, úgy hogy numerikus összegük zéró. Ezért az eltérések négyzetét alkalmazzuk, hogy eltávolítsuk a negatív előjeleket az összegezés előtt. Az X -változó minden értékére, az Y -változó megfelelő értékét, melyről előrejelezzük, hogy a regressziós egyenesen helyezkedik el, \hat{y} -vel jelöljük és a $\sum(y-\hat{y})^2$ kifejezést reziduális (maradék) eltérésnégyzet-összegnek hívjuk. A regressziós egyenest tehát ott kapjuk meg, ahol a reziduális eltérésnégyzet-összeg, $\sum(y-\hat{y})^2$ minimális. Ezt illusztrálja a 18.5. ábra.

A legjobban illeszkedő regressziós egyenes b meredekségét *regressziós koefficiensnek* hívjuk. A b regressziós koefficiens a minta x , y értékeire, amely teljesíti azt a követelményt, hogy a maradékok összege minimális legyen, megkapjuk, ha x , y eltérésszorzat-összegét (SP) osztjuk az x eltérésnégyzet-összegével (SS_x):

$$b = \frac{\sum (x - \bar{x})(y - \bar{y})}{\sum (x - \bar{x})^2} = \frac{\sum XY - \frac{(\sum X)(\sum Y)}{n}}{\sum X^2 - \frac{(\sum X)^2}{n}} = \frac{SP}{SS_x}$$

ahol: SP: x és y változók eltérésszorzat-összege;

SS_x: az x változó eltérésnégyzet-összege.

A regressziós egyenes *a* tengelymetszetének értékét megkaphatjuk a regressziós egyenes egyenletének átrendezésével és a *b*, \bar{x} és \bar{y} ismert értékeinek behelyettesítésével, tehát:

$$a = \bar{y} - (b \times \bar{x})$$

Mihelyt *a* és *b* ismertek, az *x* bármely értéke behelyettesíthető a regressziós egyenletbe és meghatározható az előrejelzett érték, \hat{y} , amelyik a regressziós egyenesen helyezkedik el. Ha ezt az eljárást elvégezzük két *x*, \hat{y} párra, a megfigyelt adattartomány ellenkező oldalán, akkor ezt a két pontot összekötő egyenes vonal reprezentálni fogja az adatokhoz legjobban illeszkedő regressziós egyenest.

A lineáris regresszió teljes számítását a legkisebb négyzetek módszerével az alábbi példa mutatja. Ábrázolást követően, a regressziós egyenes felhasználható arra, hogy előrejelezzük a függő változó értékeit a független változó megadott értékeire. Például, a regressziós egyenes a példában alapját képezheti a különböző ellési időponttal jellemezhető tehének tejhozamának előrejelzésére. Érvénytelen azonban előrejelzést végezni a megfigyelt értékek tartományán kívüli extrapoláció alapján, mivel nincs információ arra, hogy a kiszámított regressziós egyenes érvényes e tartományon kívül.

Példa. A kétváltozós lineáris regresszió meghatározása a legkisebb négyzetek módszerével

Összefüggés az első ellés kora és a tehének élettartam alatti tejhozama között (adatok azonosak a 18.1. táblázat adataival). Adattáblázat:

Életkor az első borjú születésekor (hónap)	Élettartam alatti tejhozam (kg)
24	18747
25	18673
26	18456
27	18730
28	17995
29	18300
30	17964
31	17820
32	17842
33	17991
34	17758
35	17650

Az adatpontok ábrázolása egy szórásdiagramon negatív lineáris összefüggést jelez; azonban szükséges tudni, hogy a korreláció szignifikáns-e és a legjobban illeszkedő egyenes hol helyezkedik el.

Regressziós mutatók számítása: az r regressziós koefficiens meghatározható úgy, mint azt korábban leírtuk.

$$\begin{array}{lll} \sum x = 354 & \sum y = 217\,926 & n = 12 \\ \sum x^2 = 10\,586 & \sum y^2 = 3\,959\,417\,604 & \sum xy = 6\,414\,388 \\ SS_x = 143 & SS_y = 1\,772\,481 & SP = -14\,429 \end{array}$$

$$r = \frac{SP}{\sqrt{SS_x \times SS_y}} = -0.906$$

Következtetés: negatív korreláció van az első ellés kori életkor és az élettartam alatti tejhozam között. A legjobban illeszkedő egyenes meredekségét, b -t következőképpen számítjuk ki:

$$b = \frac{SP}{SS_x} = \frac{-14429}{143} = -100.90$$

Az Y -tengely metszet, a , kiszámítása: $a = \bar{y} - (b \times \bar{x}) = 21137.11$

A legjobban illeszkedő egyenes egyenlete tehát: $y = 21140 - 101x$

A regressziós egyenes most ábrázolható az adatpontokkal egy *szórásdiagramon*.

18.2.3. A regressziós egyenes szignifikanciájának tesztelése

Érthetően, lehetséges kiszámítani egy regressziós egyenest bármely kétváltozós adatsorozatra, melyet egy szórásdiagramon ábrázolunk, akár szoros lineáris összefüggést mutatnak az adatok, akár nem. Számos esetben, amikor az adatok nagymértékben randomok és szóródnak, a regressziós egyenes illesztés értelmetlen gyakorlat; lesznek azonban olyan esetek, amikor a regressziós egyenes illesztésének validitása nem egyértelmű. Ezért hasznos tesztelni a regressziós egyenes statisztikai szignifikanciáját, más szavakkal, megvizsgálni, hogy az Y -változó mutat-e szignifikáns lineáris függést az X -változótól. Ez elvégezhető a Student t teszttel és a varianciaanalízissel (ANOVA, VA), amely ugyanannak az eljárásnak két reprezentációja. A varianciaanalízis tartalmazza a t próbát is ($F = t^2$). A variancia táblázat egyszerre teszteli a b együtthatót, az r -t és a determinációt.

A regresszió szignifikanciájának t -próbája

A t teszt használható annak meghatározására, hogy a megfigyelt regressziós koefficiens, b , szignifikánsan eltér-e zérótól. A t mutatót általában meghatározza a mutató becslt értéke és a hipotetikus érték közötti különbség, osztva a standard hibával.

A regressziós koefficiens standard hibája $\sqrt{RMS / SS_x}$, a t érték tehát:

$$t = \frac{b - 0}{\sqrt{\frac{RMS}{SS_x}}}$$

ahol: RMS = maradék közepes négyzetes eltérés;

SS_x = x eltérésnégyzet-összege.

A null hipotézist, mely szerint nincs különbség a megfigyelt regressziós koefficiens, b , és 0 érték között, úgy tesztljük, hogy összehasonlítjuk a megfigyelt t értéket a táblázati kritikus t értékkel, a szükséges valószínűségi szinten és $n - 2$ szabadságfoknál. Abban az esetben, amikor $t \geq t_{kritikus}$, levonjuk azt a következtetést, hogy az Y -változó szignifikáns lineáris függést mutat az X -változótól. (Jegyezzük meg, hogy a t próba felhasználható továbbá annak meghatározására, hogy a megfigyelt regressziós koefficiens, b , szignifikánsan függ-e β bármely más ismert vagy teoretikus értékétől.)

A regresszió varianciaanalízise

A regressziós egyenes ANOVA tesztjének null hipotézise az, hogy Y -változó nem függ lineárisan az X -változótól. (Alternatív módon, ez úgy tekinthető, mint az az állítás, hogy β regressziós koefficiens egyenlő 0-val.) Az ANOVA teszt követi ugyanazt az alapeljárást, mint az egytényezős ANOVA, amelyben az összes variációt felosztjuk az Y -változónak az X -változótól való lineáris függőségének tulajdonítható variáció (amely megfelel a kezeléshatásnak az egytényezős ANOVA-ban) és a maradék (reziduális) hatásnak tulajdonítható variáció között. A kettőnek az aránya egy F -értéket ad, amely felhasználható a lineáris regresszió szignifikanciájának tesztelésére.

Az adatok összes variációját megadja a függő Y -változó eltérésnégyzet-összege:

$$SS_{\text{összes}} = SS_y = \sum (y - \bar{y})^2 = \sum Y^2 - \frac{(\sum Y)^2}{n}$$

Az y értékekben levő variabilitás nagyságát, melyet közvetlenül az X -változótól való lineáris függés eredményez, a regresszió eltérésnégyzet-összege ($SS_{\text{regresszió}}$) méri. A regresszió SS az eltérés négyzetösszege minden \hat{y} pontnak a regressziós egyenesen az \bar{y} -tól, azaz:

$$SS_{\text{regresszió}} = \sum (\hat{y} - \bar{y})^2 = \frac{SP^2}{SS_x}$$

A fenti formulával számolás nagyon hosszantartó lenne; szerencsére kimutatható, hogy $SS_{\text{regresszió}} = (SP)^2 / SS_x$, ahol SP = eltérésszorzat-összeg. Mivel a regressziós meredekség, $b = S_{xy} / SS_x$, behelyettesítve a formula a következő:

$$SS_{\text{regresszió}} = b \cdot SP$$

A maradék SS -t ($SS_{\text{maradék}}$) megadja minden y pontnak a regressziós egyenesen előrejelzett \hat{y} értékétől való eltérésnégyzet-összege, azaz $SS_{\text{maradék}} = \sum (y - \hat{y})^2$. $SS_{\text{maradék}}$ megkapható sokkal könnyebben a regresszió SS -nek az összes SS -ből való kivonásával:

$$SS_{\text{maradék}} = SS_{\text{összes}} - SS_{\text{regresszió}}$$

Jegyezzük meg, amikor mindenegybes adatpont a regressziós egyenesre esik, az összes SS egyenlő lesz a regresszió SS-sel és a maradék SS zéró lesz.

A szabadságfok (d.f.) az összes variabilitással összefüggésben $n - 1$. A d.f. a regresszió SS-sel összefüggésben $p - 1$, ahol p a paraméterek száma a regressziós modellben, amelyet becsültünk. Két változó közötti egyszerű lineáris regressziónál két modell paraméter van, α és β , ezáltal a regresszió d.f. mindig 1. A maradék d.f.-et $n - p$ adja meg, és ezáltal $n - 2$ az egyszerű lineáris regressziónál.

Ezután kiszámítjuk a regresszió és a maradék közepes négyzetes eltéréseket úgy, hogy a megfelelő eltérésnégyzet-összegeket osztjuk a kapcsolódó d.f. értékekkel (azaz $MS = SS/d.f.$). Az F-arányt (értéket) megadja:

$$F\text{-arány} = \frac{\text{Regresszió MS}}{\text{Maradék MS}}$$

A lineáris regresszió ANOVA táblázatát az 18.3. táblázat mutatja.

Null hipotézis fennállásakor, amikor nincs lineáris összefüggés, a populáció regressziós koefficiense egyenlő zéróval, és az F-arány értéke közel van 1-hez, míg minél nagyobb az F-értéke 1 fölött, annál kevésbé valószínűvé válik, hogy a null hipotézis korrekt. A regresszió szignifikanciáját úgy teszteljük, hogy összehasonlítjuk a megfigyelt F-értéket a táblázati kritikus F-értékkel, az indokolt valószínűségnél, valamint a regresszió és a maradék szabadságfoknál. Ha a megfigyelt F-érték egyenlő vagy meghaladja a kritikus F-értéket, akkor a null hipotézist visszautasítjuk.

18.3. táblázat. Az ANOVA táblázat formája az egyszerű lineáris regresszió szignifikanciájának vizsgálatára

Variáció forrása	Eltérésnégyzet-összeg (SS)	Szabadság fokok (d.f.)	Közepes négyzetes eltérés (MS)	F-arány
Regresszió	$SS_{\text{regresszió}}$	1	$\frac{SS_{\text{regresszió}}}{d.f._{\text{regresszió}}}$	$\frac{MS_{\text{regresszió}}}{MS_{\text{maradék}}}$
Maradék	$SS_{\text{maradék}}$	$n - 2$	$\frac{SS_{\text{maradék}}}{d.f._{\text{maradék}}}$	
Összes	SS_y	$n - 1$		

A regresszió szignifikanciájának számítása

Összefüggés az első elléskori életkor és a tehének élettartam alatti tejhozama között (adatok a az 1. példából)

$$SS_x = 143; SS_y = 1772481; S_{xy} = -14429; b = -100.902; n = 12$$

Student t teszt: H_0 : a regressziós egyenes egyenes $\beta = 0$

$$t = \frac{b - 0}{\frac{RMS}{\sqrt{SS_x}}} = \frac{100.902}{\sqrt{\frac{31656.5}{143}}} = 6.78$$

$$t_{\text{kritikus}} (P=0.01; DF=10) = 3.17$$

Következtetés: szignifikáns különbség van a megfigyelt regressziós egyenes b és zéró érték között ($P < 0.01$), amely arra utal, hogy az y (tejhozam) szignifikáns lineáris összefüggésben van x -szel (az első borjú születésének életkorával).

ANOVA teszt: H_0 : az Y -változó (élettartam alatti tejhozam) nem függ lineárisan az X -változótól (első borjú születésekor a tehén életkora).

Eltérésnégyzet-összegek:

Összes $SS = SS_y = 1772481.00$

Regresszió $SS = b \times SP = 1455914.96$

Maradék $SS = SS_{\text{összes}} - SS_{\text{regresszió}} = 316566.04$

ANOVA táblázat

Variáció forrása	SS	<i>d.f.</i>	<i>MS</i>	<i>F-arány</i>	Kritikus <i>F</i> -érték ($P=0.01$)
Regresszió	1455916.37	1	1455916.37	45.99	10.04
Maradék	316564.63	10	31656.46		
Összes	1772481.00	11			

Következtetés: szignifikáns lineáris összefüggés van az élettartam alatti tejhozam és a tehének első borjújásának életkora között ($P < 0.01$).

18.2.4. A regressziós egyenes illeszkedésének tesztelése: determinációs egyenes (r^2)

Miután meghatároztuk az x, y adatsorozathoz a legjobban illeszkedő regressziós egyenest és kimutattuk, hogy a regressziós modell szignifikáns, a következő nyilvánvaló kérdés, ténylegesen mennyire közel illeszkednek az adatpontok a regressziós egyeneshez. Az adatpontok illeszkedésének jósága a regressziós egyeneshez mérhető az $SS_{\text{regresszió}}/SS_y$ arányával. Matematikailag kimutatható, hogy $SS_{\text{regresszió}}/SS_y$ aránya egyenlő r^2 -tel, azaz a Pearson korrelációs egyenes négyzetével, melyet *determinációs egyenesnek* hívunk. Az r^2 értékének 0 és 1 között kell lennie (gyakran százalékban fejezik ki), és megadja a variancia arányát a függő Y -változó értékeiben, melyet megmagyaráz a regresszió. A példában az r^2 értéke 0.82. Ez úgy értelmezhető, hogy a variabilitás 82%-a a megfigyelt tehének élettartam alatti tejhozamának variációjában értelmezhető a tejhozamnak a tehének első borjújásakor életkorától való lineáris függőségével, míg a megmaradó variancia más, nem meghatározott faktoroknak tulajdonítható.

18.2.5. A regressziós egyenes megbízhatóságának tesztelése

Ha kiszámítottuk a regressziós modellt a minta adatokból, és különösen, ha a regressziós egyenletet predikcióra kívánjuk használni, fontos hogy képesek legyünk becsülni, mennyire közel helyezkedik el a minta regressziós koefficiense (b) és regressziós állandója (a) a valódi populáció regressziós koefficienséhez (β) és regressziós állandójához (α). Ezt a mutatók standard hibájának (szórásának) és konfidencia határainak kiszámításával érjük el.

A b regressziós koefficiens standard hibája (hibaszórása) és konfidencia határai

A gyakorlatban a populáció varianciáját (σ^2) a minta statisztikai mutatóból kell becsülnünk. A random variációt az y értékek között, mihelyt a regresszió hatását eltávolítjuk, $\sum(y-\hat{y})^2 / d.f.$ adja meg, melyet maradék (reziduális) közepes négyzetes eltérésnek (RMS) hívunk, és ezt az értéket használjuk fel a σ^2 becslésére. (Az RMS értéket megkapjuk a regresszió varianciaanalíziséből.) A b regressziós koefficiens standard hibáját (SE_b) vagy hibaszórása:

$$SE_b = \sqrt{\frac{RMS}{SS_x}}$$

A minta regressziós koefficiens (b) körüli konfidencia határok megadják azt a tartományt, amelyen belül a populáció β koefficiensének valódi, azonban ismeretlen értéke várhatóan előfordul, a valószínűség adott szintjén. A b körüli konfidencia határokat megkapjuk a szokásos módon, vagyis az SE-t megszorozzuk a Student t -értékkel, a szükséges valószínűségnél és $n - 2$ szabadságfoknál. Tehát a regressziós koefficiens 95%-os konfidenciahatárai:

$$95\% \text{ CL} = b + SE_b \times t_{(P=0.05, DF=n-2)}$$

A regressziós állandó (a) standard hibája és konfidencia határai

A regressziós állandó (a) standard hibáját (hibaszórását) megadja:

$$SE_a = \pm \sqrt{RMS \times \left(\frac{1}{n} + \frac{\bar{X}^2}{SS_x} \right)}$$

és a 95%-os konfidencia határokat megadja:

$$95\% \text{ CL} = a + SE_a \times t_{(P=0.05, d.f. = n-2)}$$

18.2.6. Az előrejelzések (predikciók) megbízhatóságának tesztelése

Az előzőeknél talán még fontosabbak az y értékére vonatkozó becslések egy adott x értéknél, a regressziós mutatók alapján. Három különböző helyzet van, amikor ilyen becsléseket végzünk:

1. *Először*, az X -változó adott értékére szükséges lehet előrejelezni a populáció átlagát az Y -változó minden lehetséges, előforduló értékére. Kérdés, hogy Y -változó előrejelzett átlagos értékére vonatkozó becslésünk mekkora hibával terhelt. A példában, mennyi az átlagos élettartam alatti tejhozama minden olyan tehénnek a populációban, amelyik az első borjút 30 hónapos korban szülte meg. Az Y -változó előrejelzett átlagos értékének ($\hat{y}_{(\mu)}$) standard hibáját (szórását) az X -változó adott értékére (X_i) a következő egyenlet adja meg:

$$SE \hat{y}_{(\mu)} = \sqrt{RMS \times \left(\frac{1}{n} + \frac{(X_i - \bar{X})^2}{SS_X} \right)}$$

ahol: $SE \hat{y}_{(\mu)}$ = az y előrejelzett átlagos értékének standard hibája minden adatra, melynek x_i az X -változó értéke;
 RMS = maradék közepes négyzetes eltérés = $\sum(y - \hat{y})^2 / d.f.$;
 SS_X = az x eltérésnégyzet-összege.

2. *Másodszor*, szükséges lehet előrejelezni y átlagos értékét egy adott nagyságú mintára és az x egy adott értékére, a populációból. A példában, mennyi az átlagos élettartam alatti tejhozama a populációból kiválasztott tíz tehénből álló mintának, melyek 30 hónapos korban születték első borjójukat. Ha a kiválasztott minta nagyságát m -mel jelöljük, akkor az Y -változó előrejelzett átlagos értékének ($\hat{y}_{(m)}$) standard hibáját (szórását) az X -változó adott x_i mintájára a következő egyenlet adja meg:

$$SE \hat{y}_{(m)} = \sqrt{RMS \times \left(\frac{1}{m} + \frac{1}{n} + \frac{(x_i - \bar{x})^2}{SS_X} \right)}$$

ahol: $SE \hat{y}_{(m)}$ = az y előrejelzett átlagos értékének standard hibája m tagból álló mintára, melyek az X -változó x_i értékével rendelkeznek.

3. *Harmadszor*, szükséges lehet előrejelezni az y értékét egyetlen egyedre, amely a populáció adott x értékével rendelkezik. Kérdés, hogy az egyedre vonatkozó becslésünk mekkora hibával terhelt. A példában, mennyi az előrejelzett átlagos élettartam alatti tejhozama egyetlen tehénnek a populációban, amely 30 hónapos korban szülte első borjóját. Ez egyenlő a 2. helyzettel, azonban ahol a minta mérete, m egyenlő 1-gyel. A standard hiba ebben az esetben:

$$SE \hat{y} = \sqrt{RMS \times \left(1 + \frac{1}{n} + \frac{(x_i - \bar{x})^2}{SS_X} \right)}$$

ahol: $SE \hat{y}$ = az y előrejelzett értékének standard hibája bármely adatra, melynek X -változó értéke x_i .

Mindegyik esetben a konfidencia határokat az előrejelzett \hat{y} átlagos értékre megkapjuk a standard hiba és a Student t -érték szorzatából a szükséges valószínűségnél és $d.f.$ = $n - 2$ szabadságfoknál. Számos számítógépes statisztikai program megad egy ábrát,

amelyben a felső és az alsó konfidencia határok az y átlagos értékének előrejelzésén alapulnak (1. helyzet) és a regressziós egyenes felett és alatt vannak ábrázolva a szórásvázdiagramon. Ez hasznos, mivel grafikusán megmutatja az intervallumot mindegyik előrejelzett érték körül, amelyen belül 95%-os a valószínűsége, hogy a valódi y érték elhelyezkedik és ezáltal jelzi az előrejelzés megbízhatóságát. Nem meglepő, hogy a konfidencia intervallumok minimálissá válnak ott, ahol x egyenlő, és egyre nagyobbakká válnak, ahogyan az előrejelzett y értékek az x átlagtól távolodó x értékekből származnak. Ez látható a bemutatott szórásvázdiagramon. Megjegyezzük, hogy a konfidencia intervallumok az y átlagra egy adott x -nél (2. helyzet) szűkebbek lesznek, mint az előrejelzett intervallum egyetlen y -ra egy adott x -nél (3. helyzet) (18.11. ábra).

A lineáris regresszió konfidencia határainak meghatározása

Összefüggés az első ellés életkora és a tehének élettartam alatti tejhozama között
 $SS_x = 143$; $SS_y = 1772481$; $S_{xy} = -14\,429$; $b = -100.902$;
 RMS (az ANOVA-ból) = 31656.60

A regressziós koefficiens standard hibája = $\pm \sqrt{\frac{RMS}{SS_x}} = \pm \sqrt{\frac{31656.60}{143}} = \pm 14.88$

A regressziós koefficiens konfidencia határa =

$$95\% \text{ CL} = SE \times t_{(P=0.05; DF=n-2)} = 14.88 \times 2.23 = \pm 33.15$$

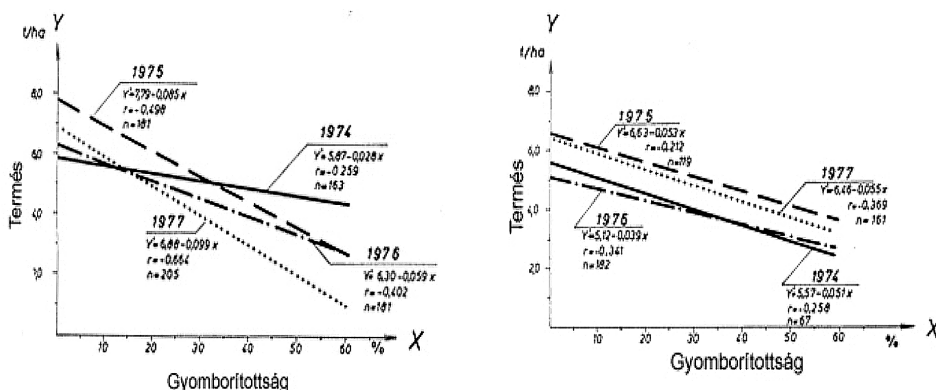
Az y előrejelzett értékének konfidencia határai:

Életkor első elléskor (hónap)	Előrejelzett élettartam alatti tejhozam (regressziós modell alapján)	$\pm SE^\dagger$	95% CL = SE \times t	95%-os felső CL határ	95%-os alsó CL határ
24	18715.46	96.62	451.09	18930.72	18500.20
25	18614.56	84.39	438.74	18802.57	18426.55
26	18513.66	73.14	428.60	18676.62	18350.70
27	18412.76	63.42	420.84	18554.05	18271.46
28	18311.85	56.00	415.58	18436.62	18187.08
29	18210.95	51.90	412.93	18326.58	18095.32
30	18110.05	51.90	412.93	18225.68	17994.42
31	18009.15	56.00	415.58	18133.92	17884.38
32	17908.24	63.42	420.84	18049.54	17766.95
33	17807.34	73.14	428.60	17970.30	17644.38
34	17706.44	84.39	438.74	17894.45	17518.43
35	17605.54	96.62	451.09	17820.80	17390.28

$$^\dagger SE \hat{y}_{(\mu)} = \sqrt{RMS \times \left(\frac{1}{n} + \frac{(X_i - \bar{X})^2}{SS_x} \right)}$$

18.2.7. A gyomborítottság és a kukorica szemtermése közötti lineáris összefüggés

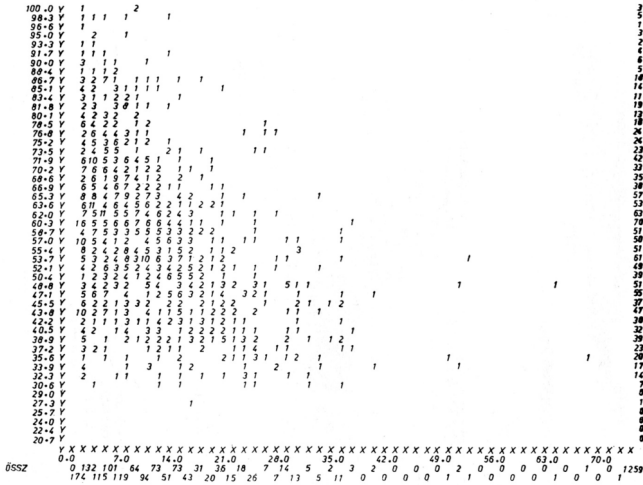
Az üzemi kukoricatermesztés táblasoros adatait biometriai módszerekkel elemeztük (Bertzsenyi, 1979a, 1980). A vizsgálat 51 mezőgazdasági nagyüzem adatain alapult. A gazdaságokat a termesztési rendszerekhez való tartozásuk alapján két csoportba soroltuk: IKR és Nem-IKR (KSZE, KITE, BKR). A vizsgálat megfigyelési egységét egy-egy kukoricatábla képezte. A gyomborítottság és a kukoricatermés összefüggését elemezve, hétféle függvényvel történt közelítés és szignifikancia vizsgálat alapján kimutattuk, hogy az összefüggés lineáris függvényvel jellemezhető (18.7. ábra).



18.7. ábra. A kukorica termése és a gyomborítottság közötti összefüggés az „IKR” (balra) és „NEM-IKR” (jobbra) csoportokban az 1974–1977. években (n = táblák száma) (Bertzsenyi, 1979b)

A gyomnövények nagyon káros hatására hívta fel figyelmünket a teljes adathalmaz összevont elemzése és számítógépes ábrázolása (Sváb és Bertzsenyi, 1979). A 18.8. ábra derékszögű koordináta-rendszerben mutatja be az 1259 kukoricatábla gyomborítottság %-a (X -tengely) és a szemtermés (Y -tengely) közötti összefüggést. A számok azt fejezik ki, hogy a gyomborítottság és a termésmennyiség adott kombinációjához hány tábla tartozik. A 18.8. ábrából kitűnik, hogy a gyomosodás a nagyüzemi kukoricatermesztés egyik legfontosabb gátló tényezője. Megfigyelhető, hogy 5% és 65% gyomborítottság között 10 t ha⁻¹-től 3 t ha⁻¹ termésig gondolatban egyenes húzható, amely felett mintegy korlátként a gyom megakadályozhatja a termést. E korlát alatt a kukorica termését az egyéb tényezők alakítják ki. Az összefüggésnek ezt a jellegét semmilyen függvényvel nem tudjuk kifejezni.

Az összefüggést az összes kukoricatábla ($n = 1259$) viszonylatában az $Y = 6.47 - 0.065X$ lineáris függvény írta le (Y = termésátlag, t ha⁻¹, X = gyomborítottság %, $r = -0.404^{***}$). A regressziós koefficiens $b_{yx} = -0.068$ értéke azt jelenti, hogy a gyomborítottság 1%-os növekedése a kukorica termését 0.068 t-val csökkentette ha-onként. Az átlagos 10.41%-os gyomborítottságot figyelembe véve, négy év átlagában a gyomborítottság a kukorica termését 0.707 t-val csökkentette ha-onként. A determinációs koefficiens ($r^2 = 16.39\%$) kimutatta, hogy a termésátlag varianciáját a vizsgálat átlagában 16.39%-ban a gyomborítottsággal lehetett megmagyarázni és átlagosan 83.61% volt az egyéb tényezők hatása.

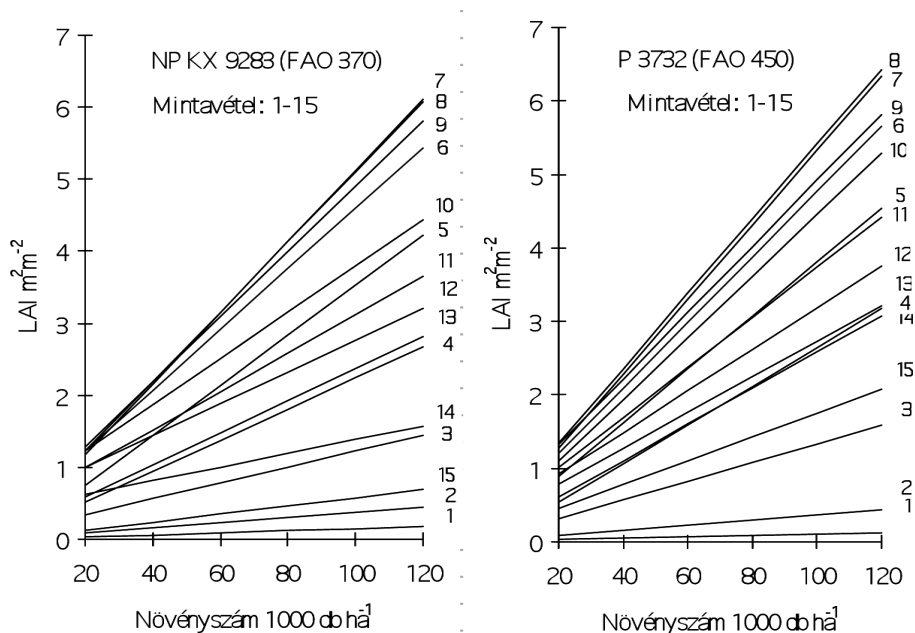


18.8. ábra. A gyomborítotttság (%) és kukorica szemtermése (q ha⁻¹) közötti összefüggés számítógépes ábrája (Sváb és Berzsényi, 1979)

18.2.8. A növényszám és a levélterület index (LAI) közötti lineáris összefüggés

Kísérlet leírása. A növényszám hatását a kukorica levélterületére ún. folytonos sűrítési kísérletben, 20 –120 · 10³ tő ha⁻¹ egyedszám tartományban, hat különböző növény számnál (20, 40, 60, 80, 100 és 120 · 10³) vizsgáltuk, az MTA Mezőgazdasági Kutatóintézetének kísérleti területén, 1987. és 1988. években. A kísérletet két eltérő tenyészidejű hibriddel állítottuk be: NK PX 9283 (370) SC és Pioneer 3732 (450) SC. Az ún. destruktív növekedésanalízishez a növény minták vételét a kukorica 4-leveles fejlettségénél kezdtük meg és a fiziológiai érésgig folytattuk, 10–14 napos intervallumokban. 1987-ben 13, 1988-ban 15 alkalommal vettünk növény mintákat. Az egyes levelek területét Delta-T típusú levélterület-mérő műszerrel határoztuk meg és ezután növényenként összegeztük. A növényenkénti levélterület alapján növény szám kezelé senként kiszámítottuk a levélterület index (LAI) nagyságát, amely kifejezi a növény állomány levélterületét egységnyi területre vetítve (m m⁻², dimenzió nélküli mutató).

Regresszió-analízissel a növény szám és a LAI nagyon fontos kapcsolatát sikerült feltárni. Mintavételként, vagyis a kukoricánövény fejlődésének különböző időpontjaiban vizsgálva a növény szám és a LAI összefüggését, lineáris kapcsolatot állapítottunk meg, vagyis a LAI értéke lineárisan emelkedett a növény szám növekedésével (18.9. ábra). A LAI lineáris függvény szerinti változása a kukoricánövény egyedfejlődésének különböző időpontjaiban jól jellemezhető az egyenes meredekségével, azaz a függvény „b” regressziós koeficiensével. A „b” értéke a virágzás időpontjáig (LAI_{max} eléréséig, 7–8. mintavétel, a vetés utáni 71–78. nap) fokozatosan növekedett, ezt követően pedig csökkent. Az összefüggés minden esetben szignifikáns volt. Biológiai megalapozottságát bizonyítja a determinációs koeficiens (r²) következetesen magas értéke. Az r² nagyságrendje a levelek fokozott szenescenciájának kezdetéig 0.96–0.99 között alakult, ezt követően kis mértékben csökkent (0.83–0.96) (Berzsényi, 1989).

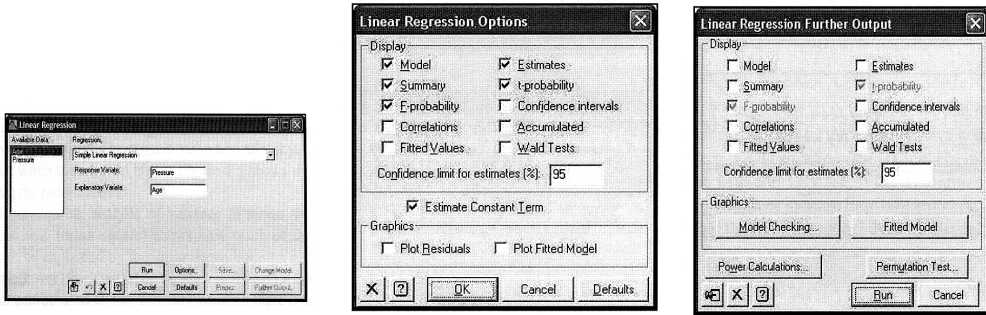


18.9. ábra. A kukoricánövény levélfelület indexének (LAI) lineáris függvény szerinti változása a növénytétel és a mintavétel időpontjának függvényében 1988. évben (Berzsenyi, 1989)

18.2.9. Lineáris regresszióanalízis GenStat programmal

Az egyszerű lineáris regresszióanalízishez két változó között, az adatokat beviszük a GenStat számológéplébe, két adatoszlopba. Ezt megtehetjük egy Excel számológéplébe fájl importálásával vagy az adatok beírásával közvetlenül a GenStat számológéplébe. A *Stat* lehetőséget választjuk ki a GenStat eszköztárban és a *Regression Analysis*-t választjuk az ezt követően legördülő menüből. A regressziós eljárások széles körét tartalmazó menü jelenik meg; az egyszerű lineáris regresszió eléréséhez kattintsunk a *Linear Models*-re, hogy megkapjuk a lineáris regresszió adatbeviteli dobozt. A lineáris regresszió dobozban válasszuk ki a *Simple Linear Regression*-t (egyszerű lineáris regressziót) és vigyük be a függő változó (ábrázolva az y-tengelyen) oszlop nevét a *Response Variate* dobozba és a független változó (ábrázolva az x-tengelyen) oszlop nevét az *Explanatory Variable* dobozba. Kattintsunk rá az *Option*-re, hogy kiválasszunk számos további statisztikát, ha szükséges, beleértve az adatok szórásdiagramját, és kattintsunk a *Run*-ra, hogy megtörténjen az analízis.

Ha kértük az adatok ábráját a *Further Output* menüben, ez kezdetben a GenStat grafika ablakban jelenik meg. Az illesztett regressziós egyenes ábrázolásán kívül, a program ábrázolja alapértelmezés szerint a felső és alsó 95%-os regresszió konfidencia intervallumokat is. A grafika szerkeszthető a GenStat-on belül és ezt követően elmenthetően importálható a Word dokumentumokba. A regresszióanalízis számszerű eredménye a GenStat kimeneti ablakban jelenik meg rendesen. A GenStat menü dobozokat a lineáris regresszióanalízishez a 18.10. ábra mutatja.



18.10. ábra. Genstat menü dobozok a lineáris regresszióanalízishez

A tehén életkora és a tejhozam közötti összefüggés vizsgálata GenStat programmal

Regression analysis

Response variate: Tejhozam
 Fitted terms: Constant, Eletkor

Summary of analysis

Source	d.f.	s.s.	m.s.	v.r.	F pr.
Regression	1	1455916.	1455916.	45.99	<.001
Residual	10	316565.	31656.		
Total	11	1772481.	161135.		

Percentage variance accounted for 80.4
 Standard error of observations is estimated to be 178.

Estimates of parameters

Parameter	estimate	s.e.	t(10)	t pr.	lower 95%	upper 95%
Constant	21137.	442.	47.83	<.001	20152.	22122.
Eletkor	-100.9	14.9	-6.78	<.001	-134.1	-67.75

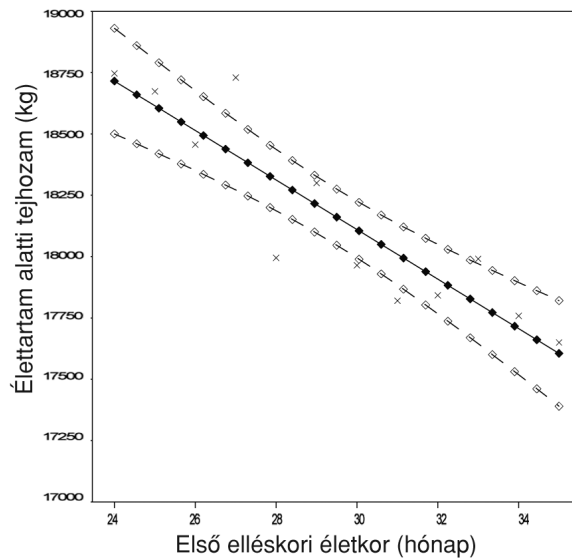
Correlations between parameter estimates

Parameter	ref	correlations	
Constant	1	1.000	
Eletkor	2	-0.993	
		1	
			2

Fitted values and residuals

Unit	Response	Standardized		
		Fitted value	residual	Leverage
1	18747.	18715.	0.21	0.29
2	18673.	18615.	0.37	0.22
3	18456.	18514.	-0.36	0.17
4	18730.	18413.	1.91	0.13
5	17995.	18312.	-1.88	0.10
6	18300.	18211.	0.52	0.09
7	17964.	18110.	-0.86	0.09
8	17820.	18009.	-1.12	0.10
9	17842.	17908.	-0.40	0.13
10	17991.	17807.	1.13	0.17
11	17758.	17706.	0.33	0.22
12	17650.	17606.	0.30	0.29
Mean	18160.	18160.	0.01	0.17

Az y előrejelzett értékeinek alsó és felső 95%-os konfidencia határait a regressziós egyenessel a 18.11. ábra mutatja:



18.11. ábra. A tehén életkora az első borjú születésekor és az élettartam alatti tejhozam lineáris összefüggésének ábrázolása (konfidencia sávokkal)

18.3. Nem-lineáris modellek

18.3.1. Lineáris és nem lineáris modellek

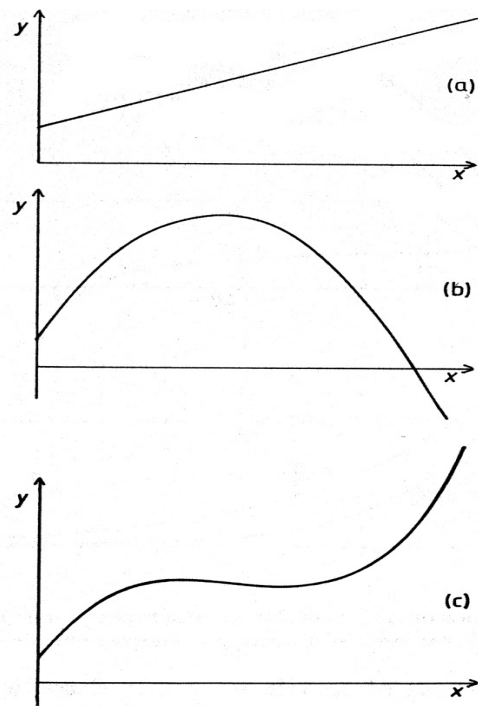
A lineáris modellek a domináns statisztikai módszerek az összefüggés vizsgálatokban, nem azért, mert mindig a legmegfelelőbbek, hanem azért, mert az adatokhoz ilyen modellek illesztésének elmélete nagyon egyszerű. A paraméterek kiszámítása a lineáris modellekben viszonylag egyszerű, elvégezhető számítógép nélkül, jóllehet számítógéppel gyorsan kiszámítható. Ezzel szemben más, lehetséges módon reálisabb modell formák, amelyek magukba foglalnak paramétereket nem lineáris módon, nem illeszthetők olyan egyszerűen, számítógép használata nélkül.

Ahhoz, hogy lássuk, miért hasznosak a nem lineáris modellek, szükséges ismerni, miért nem alkalmasak a lineáris modellek a biológiai helyzetek modellezésére. Ha figyelembe vesszünk egy összefüggést egy inger (stimulus) változó, x , és az eredményváltozó, y , között, akkor a paramétereiben lineáris modell három legegyszerűbb formája az egyenes vonal, $y = a + bx$; a másodfokú, $y = a + bx + cx^2$; és a harmadfokú, $y = a + bx + cx^2 + dx^3$. A három modellt a 18.12. ábra mutatja. Az egyenes vonal nyilvánvalóan nagyon korlátozott összefüggés. Nagyon kevés az olyan biológiai összefüggés, amely akár megközelítőleg is, lineáris az x értékek indokolt tartományában. Az egyenes vonalú összefüggés leggyakoribb formája talán az allometrikus összefüggés egy növényi vagy állati rész (szerv) tömegének logaritmusa és az egész növény vagy állat tömegének logaritmusa között.

A másodfokú (kvadratikus) modell lehetővé teszi a görbületet, azonban korlátozott két kritikus módon. Először, szimmetrikus y emelkedése az x növekedésével a maximumig pontosan ugyanolyan formájú, mint az y azt követő csökkenése az x további növekedésével. A második hátrány, amikor a másodfokú modell c paramétere negatív, hogy az y értékének negatívnak kell lennie, amikor x akár nagy, akár kicsi és ez biológiailag rendszerint nem megfelelő.

A harmadfokú és még magasabb fokú polinomok legyőzik a szimmetria hátrányát, de nem lesznek úrrá azon a hátrányon, hogy nem reális negatív vagy nagyon nagy y értéket eredményeznek az x nagy vagy kis értékénél. Ezen kívül, a maximum és minimum értékek nagyon merev struktúrájával rendelkeznek.

A polinom modell család egyik görbéje sem tesz lehetővé olyan összefüggést, melynél y aszimptotikussá válik, amikor x



18.12. ábra. Polinomiális összefüggések:
(a) $y = a + bx$, (b) $y = a + bx + cx^2$
és (c) $y = a + bx + cx^2 + dx^3$

értéke nagy, vagy olyan összefüggést, ahol y szükségszerűen pozitív. Ezen túlmenően, több általánosan használt nem lineáris modell levezethető egyszerű biológiai fogalmakból, amelyek addig a mértékig, amíg megfelelőek, igazolják a nem lineáris modell használatát.

A nem-lineáris regressziós modellek fontosak, mivel számos növény és talaj folyamatot jobban reprezentálnak a nem-lineáris modellek, mint a lineáris modellek. A mezőgazdaságban használt nem-lineáris modelleknek hat csoportját különítette el *Archontoulis* és *Miguez* (2015): exponenciális, szigmoid, fotoszintézis, hőmérséklet-függés, harang alakú görbék, egyéb nem-lineáris függvények (erő, hiperbola stb.).

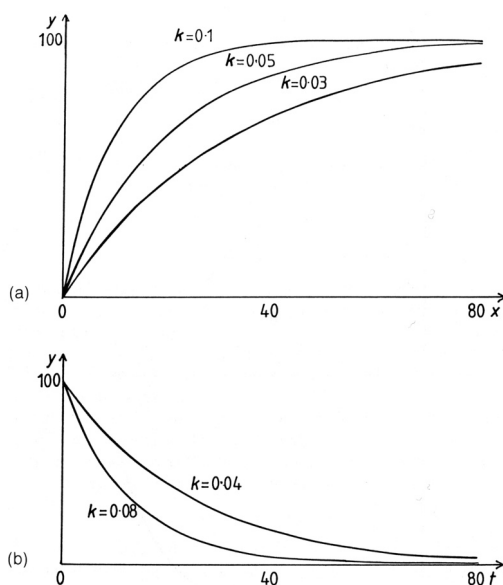
A nem lineáris modellek három típusát mutatjuk be az alábbiakban (*Mead et al.*, 1993). Az első exponenciális kifejezéseket használ, hogy reprezentáljon egy aszimptóta tendenciát. A modell formája:

$$y = A(1 - e^{-kx}),$$

amely reprezentálja pl. a termés fokozatos közeledését egy felső limithez, A -hoz, melyet a környezet határoz meg. Ezt a modellt illusztrálja a 18.13. (a) ábra a k különböző értékeire. Az x stimulus (inger) tipikusan lehet az alkalmazott kemikália koncentrációja egy élő anyagra egy laboratóriumi kísérletben, vagy az alkalmazott műtrágya dózisa a szántóföldi parcellákra egy növénytermesztési kísérletben. Egy ilyen formájú másik modell:

$$y = y_0 e^{-kt},$$

amely egy időbeni (t) csökkenést reprezentál egy kezdeti y_0 értéktől. Ezt a modellt illusztrálja a 18.13. (b) ábra.



18.13. ábra. Egyszerű exponenciális modellek: (a) aszimptotikus reakció, változó rátájú közelítéssel egy felső limithez és (b) csökkenési (hanyatlási) görbék

A nem lineáris modellek második típusa, melyet általánosan használnak és a reciprok összefüggésen alapul, tipikusan a derékszögű hiperbola. Egyszerű enzim és kémiai kinetikus (mozgás) összefüggések gyakran kifejezhetők a következő formában:

$$y = Ax/(k+x),$$

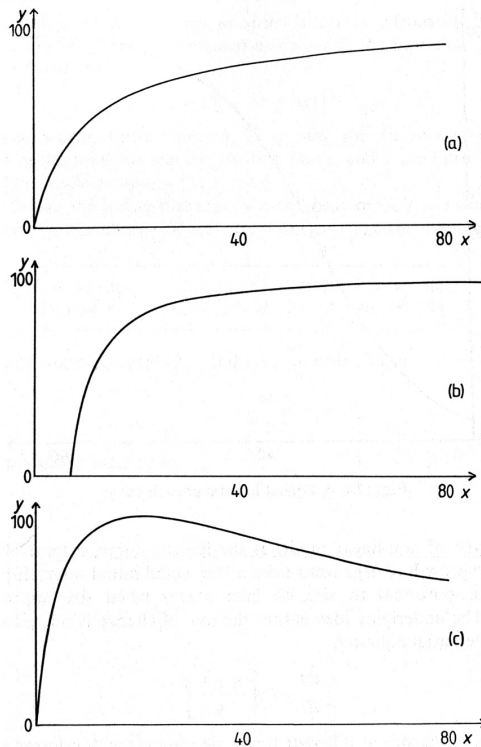
ahol k a Michaelis-Menten függvény:

$$\frac{1}{y} = \frac{1}{A} \left[1 + \frac{1}{k(x - x_0)} \right]$$

melyet felhasználnak a biokémiai reakciók sebességének leírására. Ezeket a modelleket illusztrálja a 18.14. (a) és 18.14. (b) ábra. Ennek a reciprok összefüggésnek a kiterjesztése vezet a reakciógörbék egy csoportjához, melyet agronómiai reakciók kifejezésére használunk a növényszámra vagy a műtrágya vagy kemikália inputra (x):

$$x/y = a + bx + cx^2$$

Amikor $c = 0$, ez a modell egy aszimptotikus görbére redukálódik (18.14. (a) ábra), amikor azonban c nem zéró, a modell nem-szimmetrikus reakciógörbét eredményez (18.14. (c) ábra), amely, ha a , b és c pozitív, csak pozitív y értéket enged meg.



18.14. ábra. Reciprok modellek: (a) és (b) aszimptotikus és (c) egy optimális stimulus (inger) szinttel

A levél fotoszintetikus ráta (p) reakcióját a fény intenzitására (I) a következő derékszögű hiperbola függvény írja le:

$$p = \frac{1}{a + b/I}$$

Az a és b konstansoknak specifikus értékei vannak, a növényfajtól és más külső faktorok szintjétől függően. A görbének két aszimptótája van, a horizontális $p = 1/a$ -nál, és a vertikális aszimptóta $I = -b/a$ -nál. Az utóbbinak nincs biológiai jelentősége, mivel negatív megvilágításnál fordul elő, azonban a $p = 1/a$ aszimptóta reprezentálja a maximálisan elérhető fotoszintetikus rátát (fotoszintetikus ráta fénytelítődésnél). Ha a fény nem limitáló, akkor a fotoszintetikus rátát $1/a$ határozza meg. Nyilvánvalóan, minél kisebb az a értéke, annál nagyobb lesz a fotoszintézis maximális rátája és vice versa. Az a azonosítható a széndioxid molekula külső légtérből a karboxiláció helyére történő mozgásával szembeni rezisztenciával.

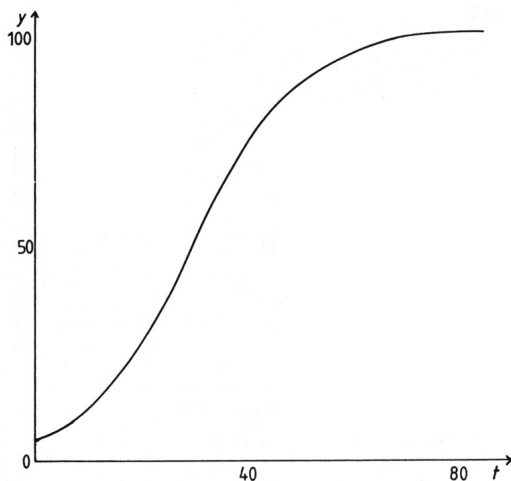
A nem lineáris modellek harmadik típusa a logisztikus görbe, melyet gyakran használunk egy organizmus növekedésének reprezentálására egy nagyon kicsi kezdeti stádiumtól (melynek során a növekedés arányos a mérettel) későbbi stádiumokig, amikor a méret (tömeg) egy aszimptóthoz közelít. Az alapul szolgáló gondolat az, hogy a változás sebessége a jelenlegi mérethez az alábbi differenciál egyenlet szerint viszonyul:

$$dy/dt = ky[(A-y)/A]$$

és méretben kifejezve különböző időben, egy nem lineáris modellt kapunk:

$$y = \frac{A}{1 + e^{-(a-kt)}}$$

A modellt a 18.15. ábra illusztrálja.



18.15. ábra. Tipikus logisztikus növekedési görbe

18.3.2. Kukorica hibridek műtrágya-reakciójának vizsgálata különböző függvényekkel

Kísérlet leírása

1967-ben beállított kéttényezős, osztott parcellás elrendezésű tartamkísérletben, 26 év adatai alapján vizsgáltuk különböző műtrágya kezelések hatását a kukorica hibridek termésére az MTA Mezőgazdasági Kutatóintézet kísérleti területén (*Berzsényi és Dang*, 2002b; 2006). A kísérlet hét kezelése (NPK aránya 2:1:1) a következő (dózis kg ha⁻¹-ban): 1. kontroll (műtrágyázás nélkül), 2. 100 kg NPK, 3. 200 kg NPK, 4. 300 kg NPK, 5. 400 kg NPK, 6. 600 kg NPK, 7. 800 kg NPK.

A kukorica hibridek műtrágya-reakcióját a kísérletben mért évenkénti termés-adatokhoz illesztett négy függvénnyel (másodfokú, négyzetgyök, exponenciális, lineáris-plató) jellemeztük. A kukorica hibridek műtrágya-reakciójának jellemzésére általánosan használt másodfokú és négyzetgyök függvényt az alábbi egyenletek írják le:

$$Y = a + bx + cx^2$$

$$Y = a + bx + cx^{1/2}$$

ahol: Y: szemtermés (t ha⁻¹), x: a műtrágya dózisa (kg ha⁻¹), a: regressziós állandó, b és c: regressziós koeficiens.

Overman et al. (1994) inverz exponenciális függvénye, mely módosított Mitscherlich függvényként is ismert, a következőképpen írható fel:

$$Y = m(1 + e^{-(cx-b)})^{-1}$$

ahol: Y: szemtermés (t ha⁻¹), m: a maximális hozam, X: a műtrágya dózisa (kg ha⁻¹), c: műtrágya-reakció koeficiens.

Anderson és Nelson (1975) által javasolt lineáris-plató modellt az alábbi egyenletek írják le:

$$Y = a + bX, \text{ ha } X < C$$

$$Y = P, \text{ ha } X \geq C$$

ahol: Y: szemtermés (t ha⁻¹), X: a műtrágya dózisa (kg ha⁻¹), a: regressziós állandó, b: regressziós koeficiens, C: a műtrágya kritikus dózisa, amely a lineáris regresszió és a plató egyenes metszéspontjában van, P: plató termés, C és P: konstansok.

Mindegyik függvénnyel 26 adatillesztést végeztünk, melynek során megkaptuk a regresszió varianciaanalízisét, a függvényparaméterek becsült értékeit, hibáit, konfidencia-intervallumát, illetve szignifikanciáját és a determinációs koeficiens (R²). Páros t-próbával vizsgáltuk a regressziótól való eltérések szignifikanciáját a függvények között. A regressziótól való eltéréseket kvartilis ábra formájában szemléltetjük.

A másodfokú és négyzetgyök függvényvel előrejelzett maximális termést

$$Y_{\max} = a - b^2/4c, \text{ illetve}$$

$$Y_{\max} = a - c^2/4b$$

formulával számítottuk ki. A lineáris-plató és az inverz exponenciális függvényenél a plató termést (P), illetve a maximális értéket (m) tekintettük maximális termésnek. Az inverz exponenciális és a lineáris-plató függvényre az optimális műtrágyadózist az egyenlet megoldásával határoztuk meg.

18.4. táblázat. A műtrágyadózis és a kukoricatermés közötti összefüggést leíró függvények determinációs koefficiensei (R^2) a vizsgált években[†]

Év	R^2 értékek				Év	R^2 értékek			
	Másodfokú	Négyzetgyök	Inverz exponenciális	Lineáris-plató		Másodfokú	Négyzetgyök	Inverz exponenciális	Lineáris-plató
1967	0.665 ^{NS}	0.358 ^{***}	0.866 ^{***}	0.958 ^{***}	1987	0.851 [†]	0.962 ^{***}	0.971 ^{***}	0.975 ^{***}
1968	0.649 ^{NS}	0.919 ^{***}	0.976 ^{***}	0.977 ^{***}	1988	0.634 ^{NS}	0.880 ^{***}	0.718 ^{***}	0.719 ^{**}
1969	0.913 ^{**}	0.966 ^{***}	0.991 ^{***}	0.993 ^{***}	1989	0.819 [†]	0.917 ^{***}	0.846 ^{***}	0.864 ^{***}
1970	0.965 ^{**}	0.976 ^{***}	0.995 ^{***}	0.967 ^{***}	1991	0.705 ^{NS}	0.971 ^{***}	0.848 ^{***}	0.849 ^{**}
1971	0.820 [†]	0.978 ^{***}	0.978 ^{***}	0.976 ^{***}	1992	0.831 [†]	0.808 ^{***}	0.606 ^{***}	0.747 ^{**}
1972	0.664 ^{NS}	0.962 ^{***}	0.958 ^{***}	0.959 ^{***}	1993	0.585 ^{NS}	0.498 ^{***}	0.413 ^{***}	0.683 ^{***}
1973	0.875 [†]	0.997 ^{***}	0.994 ^{***}	0.981 ^{***}	1994	0.256 ^{NS}	0.313 ^{***}	0.157 ^{***}	0.665 ^{***}
1974	0.759 ^{NS}	0.986 ^{***}	0.745 ^{***}	0.745 ^{***}	1995	0.500 ^{NS}	0.797 ^{***}	0.695 ^{***}	0.930 ^{***}
1975	0.872 [†]	0.970 ^{***}	0.992 ^{***}	0.993 ^{***}	1996	0.869 [†]	0.926 ^{***}	0.930 ^{***}	0.936 ^{***}
1983	0.697 ^{NS}	0.958 ^{***}	0.333 ^{**}	0.573 ^{**}	1997	0.916 ^{**}	0.876 ^{***}	0.928 ^{***}	0.962 ^{**}
1984	0.580 ^{NS}	0.900 ^{***}	0.756 ^{***}	0.900 ^{***}	1998	0.811 [†]	0.936 ^{***}	0.937 ^{***}	0.970 ^{***}
1985	0.845 [†]	0.946 ^{***}	0.947 ^{***}	0.973 ^{***}	1999	0.975 ^{**}	0.921 ^{***}	0.984 ^{***}	0.984 ^{***}
1986	0.739 ^{NS}	0.874 ^{***}	0.820 ^{***}	0.869 ^{***}	2000	0.874 [†]	0.851 ^{***}	0.888 ^{***}	0.889 ^{**}

[†]A szignifikancia szinteket F-próba alapján határoztuk meg. Szignifikancia szintek: ^{NS} = nem szignifikáns
^{***}P = 0.1%-os, ^{**}P = 1%-os, ^{*}P = 5%-os szinten szignifikáns

A 18.4. táblázat a műtrágya kezelések és a kukoricatermés közötti összefüggést leíró függvények determinációs koefficienseit (R^2) tartalmazza a vizsgált években. Az R^2 értékeknél feltüntettük a szignifikancia szinteket az F-próba alapján. Látható, hogy a leggyengébb volt a másodfokú függvény illeszkedése. A vizsgált 26 évből 12 évben nem volt szignifikáns, 10 évben 5%-os, 4 évben 1%-os szinten volt szignifikáns az illeszkedés. A négyzetgyök függvény illeszkedése az R^2 alapján minden évben 0.1%-os szinten szignifikáns volt. Az inverz exponenciális és lineáris-plató függvények R^2 értékei hasonló nagyságrendűek voltak és egy, illetve három év kivételével 0.1%-os szinten szignifikáns illeszkedést mutattak.

A 18.5. táblázat adatai mutatják, hogy a függvények hasonló maximális termést jeleztek előre, jóllehet a másodfokú függvény alapján minden évben a legnagyobb maximális termést (26 év átlagában 7.02 t ha⁻¹) kaptuk, amely kismértékben megha-

ladta a kísérletben 26 év átlagában mért legnagyobb termést (6.76 t ha⁻¹). A négyzetgyök függvény illesztése alapján meghatározott maximális termés a kísérletben mért adatoktól kevésbé tért el, 26 év átlagában 6.50 t ha⁻¹ volt. Az inverz exponenciális és lineáris-plató függvények egymástól kevésbé eltérő termésmaximumot (26 év átlagában 6.65 és 6.56 t ha⁻¹) mutattak.

18.5. táblázat. A különböző függvények által előrejelzett maximális termés a vizsgált években

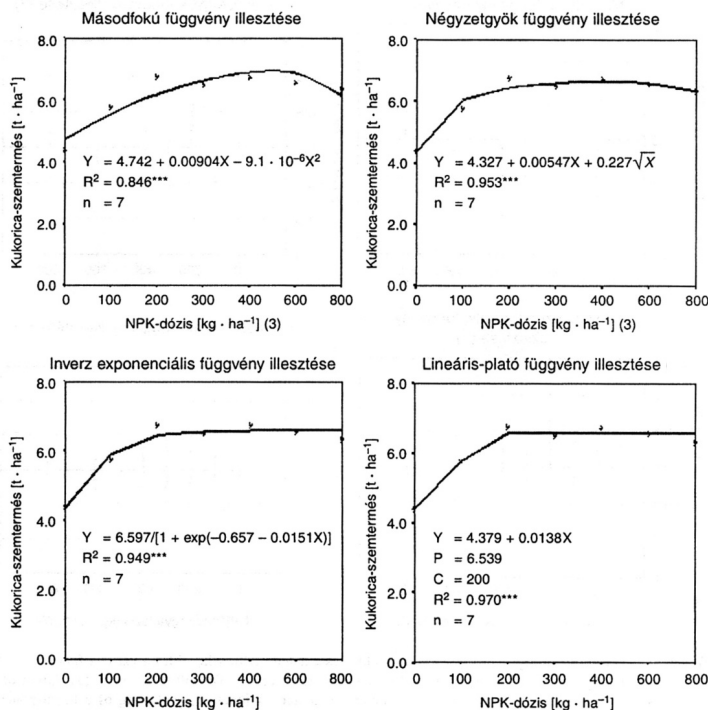
Év	Előrejelzett maximális termés (t ha ⁻¹)				Év	Előrejelzett maximális termés (t ha ⁻¹)			
	Másodfokú	Négyzetgyök	Inverz exponenciális	Lineáris-plató		Másodfokú	Négyzetgyök	Inverz exponenciális	Lineáris-plató
1967	6.18	6.02	6.02	5.97	1987	7.92	7.50	7.41	7.40
1968	5.50	5.39	5.30	5.28	1988	2.83	2.78	2.62	2.63
1969	8.96	8.55	8.52	8.43	1989	6.46	6.10	5.94	5.93
1970	6.25	6.18	6.00	5.94	1991	9.22	8.99	8.61	8.62
1971	6.22	6.02	5.94	5.93	1992	5.04	4.81	4.61	4.52
1972	7.06	6.95	6.83	6.81	1993	4.60	4.41	4.33	4.24
1973	6.41	6.25	6.21	6.21	1994	3.14	3.13	3.00	2.85
1974	5.99	5.92	5.73	5.73	1995	7.65	7.57	7.36	7.20
1975	9.40	8.95	8.87	8.80	1996	7.11	6.82	6.69	6.57
1983	4.00	4.09	3.59	3.32	1997	7.98	7.61	7.65	7.52
1984	7.25	7.13	6.78	6.57	1998	8.33	7.92	7.81	7.72
1985	7.66	7.23	7.12	7.06	1999	12.31	12.17	11.97	11.71
1986	13.28	12.71	12.36	12.17	2000	5.73	5.63	5.59	5.45

18.6. táblázat. A különböző függvények alapján előrejelzett optimális NPK műtrágyadózis a vizsgált években

Év	Előrejelzett optimális NPK dózis (kg ha ⁻¹)				Év	Előrejelzett optimális NPK dózis (kg ha ⁻¹)			
	Másodfokú	Négyzetgyök	Inverz exponenciális	Lineáris-plató		Másodfokú	Négyzetgyök	Inverz exponenciális	Lineáris-plató
1967	510	590	230	200	1987	510	480	310	200
1968	530	450	190	200	1988	440	320	170	170
1969	550	640	350	300	1989	460	350	260	180
1970	610	–	430	350	1991	440	300	150	150
1971	540	530	250	200	1992	430	290	230	250
1972	480	360	200	200	1993	430	300	250	250
1973	520	500	210	200	1994	380	220	110	250
1974	420	270	100	150	1995	440	290	200	250
1975	540	590	300	250	1996	570	760	390	250
1983	350	200	100	280	1997	530	670	430	250
1984	430	280	170	250	1998	500	440	280	250
1985	490	420	250	200	1999	590	–	620	350
1986	470	360	230	250	2000	570	–	430	250

A 18.6. táblázat a függvények alapján előrejelzett optimális NPK műtrágyadózisokat mutatja évenként. 26 év átlagában az optimum 490 kg NPK ha⁻¹ a másodfokú függvénynél, amely több mint kétszerese a lineáris-plató modell által előrejelzett optimumnak (234 kg ha⁻¹ NPK) és 66%-kal nagyobb az inverz exponenciális függvény által előrejelzett optimumnál (295 kg ha⁻¹). A négyzetgyök függvény által előrejelzett optimális NPK dózis 23 év átlagában 418 kg ha⁻¹ volt. Három évben (1970, 1999, 2000) a mért termésadatokban nem volt csökkenés, ezáltal a négyzetgyök függvényből számított optimum távoli extrapolált érték (1000 és 1500 kg ha⁻¹) lett volna.

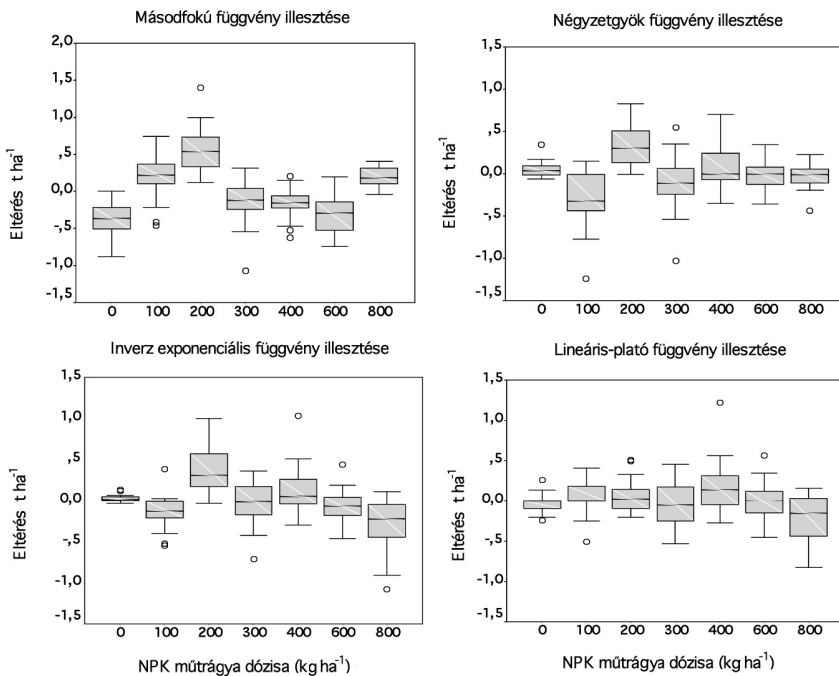
A 18.16. ábrán jól megfigyelhető a függvények eltérő illeszkedése a 26 év átlagadataihoz. Az inverz exponenciális és a lineáris-plató függvény egyaránt jól illeszkedik, műtrágyázás nélkül kisebb termést jeleznek előre, mint a másodfokú függvény és nagyobb, valós terméshövedekedést mutatnak 200–250 kg NPK dóziséig. Ugyanakkor a másodfokú függvény irreális terméshövedekedést mutat egészen 400–600 kg ha⁻¹ NPK dóziséig. Látható, hogy a másodfokú függvény jelentősen felülbecsüli a műtrágyázás optimális dózisát. Nagyon jól illeszkedik az adatokhoz a négyzetgyök függvény.



18.16. ábra. A kukoricatermés NPK műtrágyadózistól függő változásának előrejelzése négy különböző függvény illesztésével (26 év átlaga). A csillaggal a mért termést jelöltük.

A 18.16. ábra illusztrálja annak fontosságát, hogy a műtrágyázási kísérletben legalább 5–8 kezelés (dózis) legyen több éven keresztül, illetve több kísérleti helyen a műtrágya-reakció analíziséhez és a műtrágyázás optimális dózisának meghatározásához.

Az eltérést a regressziótól (mért termés – számított termés) a különböző függvények illesztésekor kvartilis grafikon formájában a 18.17. ábra szemlélteti. Megállapítható, hogy egy adott függvényillesztésnél műtrágya dózisonként is szignifikánsan különböztek a regressziótól való eltérések (reziduumok). A *másodfokú függvény*nél a regressziós eltérés átlagos értéke legnagyobb volt a 200 kg ha⁻¹ dózisinál (0.567), ezt követte a kontroll kezelés (-0.365) és a 600 kg ha⁻¹ NPK dózis (-0.316); míg legkisebb volt a 300 kg ha⁻¹ dózisinál (-0.126). A *négyzetgyök* függvénynél legnagyobb volt a regressziós eltérés a 200 kg ha⁻¹ (0.315) és a 100 kg ha⁻¹ dózisinál (-0.292), míg legkisebb a 600 kg ha⁻¹ (-0.0191) és 800 kg ha⁻¹ dózisinál (-0.252) volt. Az *inverz exponenciális függvény*nél legnagyobb volt a regressziós eltérés a 200 kg ha⁻¹ dózisinál (0.383), ezt követte a 800 kg ha⁻¹ dózis (-0.294). Legkisebb volt az eltérés átlagos értéke 300 kg ha⁻¹ dózisinál (-0.0399). A *lineáris-plató* függvény illesztésénél legnagyobb regressziós eltérést 800 kg ha⁻¹ (-0.224) és 400 kg ha⁻¹ dózisinál (0.181) kaptuk. Legkisebb volt az eltérés átlagos értéke 300 kg ha⁻¹ dózisinál (-0.0282). Legtöbb esetben a regressziós függvények műtrágyadózis szerinti eltérés- (maradék) értékei random eloszlást mutattak. A másodfokú modell eltérés értékei műtrágya dózisonként azonban szinuszcörbe típusú mintázatot mutattak.



18.17. ábra. Eltérés a regressziótól (mért termés – számított termés) a különböző függvények illesztésekor, kvartilis ábra formájában (n = 26)

Vizsgálatainkban a lineáris-plató függvény adta a legjobb eredményt. Kutatásaink rámutattak arra, hogy a függvények kiválasztásánál a determinációs koefficiens (R^2) mellett alapvető kritériumnak kell tekinteni a reziduális eltérések (mért termés – számított termés) elemzését.

18.3.3. Aszimptotikus növekedési függvények

A biológiai növekedés bármely mintázatát tekintve, egy szigmoid típusú matematikai függvény (azaz egy függvény, amelyet két horizontális aszimptota határol és mindenütt növekszik) empirikusan képes leírni a növekedést, mivel még a nem determinált növekedés is bizonyos stádiumban leáll. Ha azonban kevésbé empirikusan használunk ilyen függvényt (amely önmagában egy egyszerű, de értelmezhető biológiai modellen alapszik), akkor a módszernek szigorúbbnak kell lenni és különösen a szigmoid növekedési függvényt csak olyan növényfajokhoz vagy azok részeihez lehetne használni, amelyek növekedése determinált mintázatot mutat (Causton és Venus 1981). [A determinált növekedésű növényekre (pl. gabonafélék) jellemző, hogy a reprodukzív szervek (virágzat) megjelenésekor a vegetatív szervek (levél, szár) növekedése befejeződik.]

A szigmoid függvények lehetnek *szimmetrikusak* és *aszimmetrikusak*, annak megfelelően, hogy az inflexió pont a két aszimptótától egyenlő távolságban van-e vagy sem. A növekedésanalízisben használt aszimptotikus függvények közül bemutatjuk a *monomolekuláris*, a *logisztikus*, a *Gompertz* és a *Richards* függvényt (Hunt, 1982).

Monomolekuláris függvény

A monomolekuláris függvény a fizikai kémiából került át a biológiába, eredetileg az irreverzibilis kémiai reakció leírására használták:

$$W = a(1 - be^{-cT})$$

ahol a az aszimptóta, b méri a kezdeti tömeget és c a ráta konstans.

A monomolekuláris az egyik legegyszerűbb aszimptotikus függvény. Nincs inflexió pontja és meredeksége, amely konvex az időtengelyre, arányos a még hátralevő növekedés mértékével, $(a - W)$ -vel. A növekedési ráta folyamatosan csökken a W növekedésével és a csökkenés mértékét a c koefficiens határozza meg, amely a növény örökletes *tulajdonsága* (18.18.a. ábra).

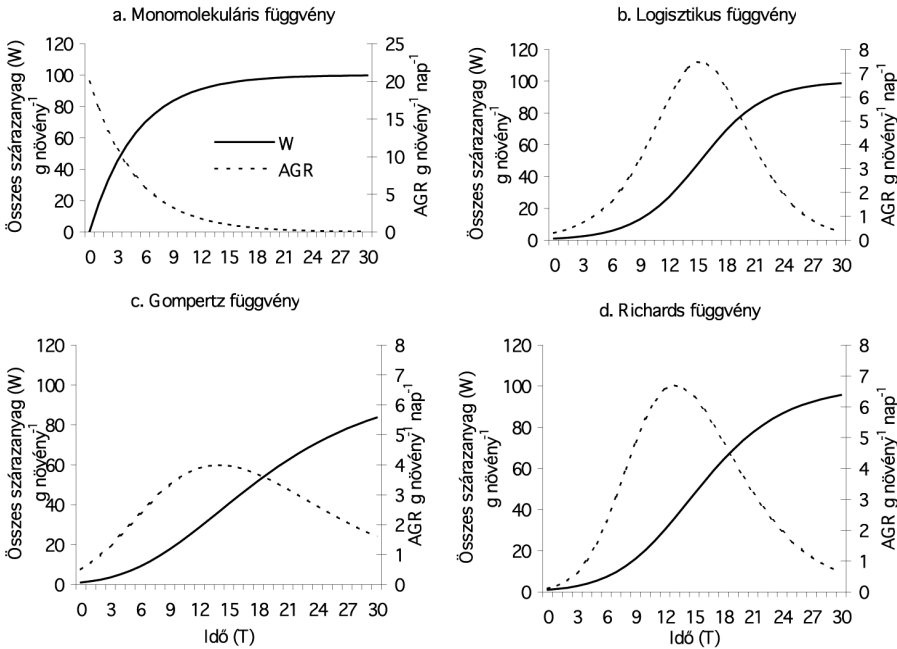
Az inflexió hiánya minden bizonnyal korlátozza a monomolekuláris függvény használhatóságát, még a növényi szervek determinált növekedésének modellezésekor is. Azokban az esetekben, amikor a növekedés a növény életében későn indul meg és ezáltal nincs inflexió, a monomolekuláris függvény megfelelő lehet.

Logisztikus függvény

A logisztikus a legismertebb szigmoid vagy S alakú függvény, amely autokatalitikus függvényként is ismert. Három paramétert tartalmaz és a következő formában írható fel:

$$W = a(1 + be^{-cT})^{-1} = a(1 + e^{b-cT})^{-1}$$

ahol „ a ” a szárazanyag-produkció maximális értéke. A másik két függvény paraméter közül „ c ” a ráta paraméter – egy magas érték jelzi a függvény gyors emelkedését a két aszimptóta között, és vice versa –, „ b ” a kiindulási tömeg és b/c meghatározza T értékét az inflexió pontnál. Az alsó aszimptóta $W(T) = 0$ és a felső aszimptóta $W(T) = a$.



18.18. ábra. Az összes szárazanyag (W) és az abszolút növekedési sebesség (AGR) időbeni dinamikája a monomolekuláris, a logisztikus, a Gompertz és Richard függvény illesztésekor (Berzsenyi, 2002)

Az S-alakú görbét a növény életciklusa során a növekedés eltérő sebességei (rátái) eredményezik. A korai, exponenciális növekedés, kiindulva egy nagyon kicsi, kezdeti méretből, relatíve nagyon gyors; a növekedés rátája arányos az organizmus méretével. Időtartama szántóföldi növényeknél mindössze néhány nap. Mielőtt a növényállomány záródik, a lineáris fázis következik viszonylag hosszú időszakra, melynek során a szárazanyag konstans rátával növekszik. A lineáris fázis időtartama és meredeksége írja le legjobban a termést. A harmadik szakaszban a növekedés tovább csökken, az organizmus mérete aszimptotikus, egy felső méretbeli limithez közelít. A növekedés negyedik és végső fázisa a növekedés antitézise, a szenescencián keresztül (Berzsenyi, 2000).

A logisztikus függvény görbéjét a 18.18.b. ábra mutatja és látható, hogy a gradiens (abszolút növekedési sebesség) mindenütt pozitív. A gradiens értéke kezdetben nő és a növekedési periódus vége felé csökken, tehát van egy inflexió pont $W = a/2$ -nél, vagyis ha az összes növekedés fele végbement. Az abszolút növekedési rátát harang alakú görbe írja le, jól definiálható maximum ponttal. A növekedés mértéke arányos a W száraz tömeggel; a növekedés a szubsztrátum mennyiségével arányos rátával megy végbe; a növekedés irreverzibilis (France és Thornley 1984).

A logisztikus függvényt gyakran alkalmazták, kezdetben mint az autokatalitikus kémiai reakció függvényét, később az állatökológia területén, a populáció egyedszám változásának modellezésére. Biológiai jelentőséget tulajdonítanak a c paraméternek, a populáció növekedés valódi rátájának és az a aszimptótának, a környezet eltartó képességének.

Újabbban, e két paraméter további alkalmazását jelenti, hogy az r -, illetve K -stratégia alapja (az r a c paraméter, és a K az a paraméter szokásos jelölése az ökológia területén). Az r stratégia esetében a populáció nagysága (N) exponenciálisan nő a $dN/dT = rN$ egyenlet szerint. A görbe meredekségét az r együttható vezérli. A növényvilágban ilyenek az egyéves gyomok, és ez a stratégia a hirtelen felszabaduló források gyors kihasználására alkalmas. A populációk másik csoportjának szaporodását a környezet eltartó képessége, a K -tényező szabályozza, mely a logisztikus görbével írható le. A K -stratégisták a fák, a cserjék és az évelő lágyszárúak, melyek az erőforrásokkal való tartós gazdálkodásra képesek (Borhidi 2002).

Gompertz-függvény

Ez egy nem-szimmetrikus szigmoid függvény, melyet Gompertz (1825) német matematikus fejlesztett ki. Három paramétert tartalmaz, kettős kitevő elrendezésben:

$$W = ae^{-be^{-CT}}$$

ahol a paramétereknek ugyanaz az általános értelmezésük, mint a logisztikus függvénynél. A W értéke $T=0$ -nál ae^{-b} . A felső aszimptóta $W(T) = 0$ és az alsó aszimptóta $W(T) = a$. A logisztikushoz hasonlóan, a Gompertz-függvény inflexiós pontja $T = b/c$ -nél van, azonban a $W(T)$ értéke itt a/e , az $a/2$ helyett.

A Gompertz-függvényt és az abszolút növekedési sebesség dinamikáját a 18.18.c. ábra szemlélteti. A logisztikus és a Gompertz-függvényt összehasonlítva látható, hogy azonos kezdeti és végső tömegnél, és hasonló inflexiós időpontnál, a Gompertz-függvény gyorsabb kezdeti növekedést, azonban lassúbb közeledést mutat az aszimptótához, hosszabb lineáris periódussal az inflexiós pont környezetében.

Többen úgy találták, hogy a Gompertz-függvény megfelelőbb a biológiai vizsgálatokban, mint bármely más szigmoid függvény. A Gompertz-függvény leggyakoribb alkalmazása a növekedésanalízisben a növényi szervek (különösen a levelek) növekedésének modellezéséhez kapcsolódik. Ugyanis a Gompertz-függvény asszimetriája sokkal jobban illeszkedik a levélnövekedési adatokhoz, mint a logisztikus függvény szimmetriája.

Richards függvény

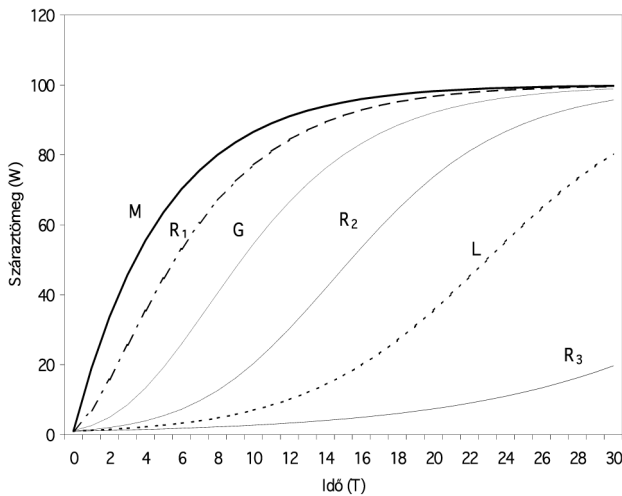
A növénytan tudományterületén Richards (1959) függvényét von Bertalanffy (1957) által az állatok növekedésének leírására kifejlesztett növekedési függvényből vezette le. A növekedésanalízisbe Causton és Venus (1981) vezette be a Richards függvényt (18.18.d. ábra). A függvény négy paramétert tartalmaz:

$$W = a(1 \pm e^{b-cT})^{-1/v}$$

a második kitevő, $-1/v$ magában foglalja a „ v ” kiegészítő paramétert. A negatív előjelet akkor használjuk, ha „ v ” negatív és fordítva. Az „ a ” paraméter adja az aszimptotikus maximumot és „ v ” írja le a görbe alakját. A „ b ” konstansnak nincs biológiai jelentősége, a görbe helyzetére vonatkozik az idő viszonylatában. A „ c ” ráta konstans, azonban értelmezése függ a „ v ” értékétől.

Richards (1959) alapján a paraméterek három különböző kombinációjának tulajdoníthatunk biológiai jelentőséget: $c/(v+1)$, a súlyozott átlagos relatív növekedési ráta a növekedés teljes időszakában; $ac/[2(v+2)]$, a súlyozott átlagos abszolút növekedési ráta; és $[2(v+2)]/c$, a növekedés fő időszaka. A negyedik paraméter, v , szabályozza, hogy a függvénynek van-e inflexiós pontja, és ha igen, hol van.

Míg a Richards függvény egy empirikus konstrukciónak (modellnek) tekinthető, bizonyos általánosíthatósággal rendelkezik, amely esetenként előnyös lehet: a kiegészítő v paraméter bizonyos eseteire magában foglalja a korábban tárgyalt három növekedési függvényt (18.19. ábra). A $v = -1$ és $v = 1$ esetében a Richards-függvény két speciális esetre korlátozódik – a monomolekuláris és a logisztikus függvényre. A $v = 0$ érték a Gompertz-függvényt képviseli (France és Thornley, 1984).



18.19. ábra. A Richards növekedési függvény a „ v ” paraméter hat különböző értékére. M jelöli a monomolekuláris ($v = -1$), G a Gompertz ($v = 0$), L a logisztikus ($v = 1$), R a Richards (R_1 : $v = -0.5$, R_2 : $v = 0.5$, R_3 : $v = 2$) függvényt (France és Thornley (1984) alapján Berzsényi, 2002)

Causton és Venus (1981) rámutatott arra, hogy a Richards-függvény egy egyszerű biológiai modellen alapszik: a vizsgált növekedési jellemző relatív növekedési rátája matematikailag egyszerű formában csökken a jellemző növekvő méretével, azonban megfelelő flexibilitás van a Richards-függvényben, hogy figyelembe vegye a kezdeti, közel konstans relatív növekedési ráták változó időtartamait (azaz közelítés az exponenciális növekedéshez). Causton és Venus (1981) könyve csaknem teljesen a Richards függvény alkalmazásán alapuló hosszú monográfia és e függvény potenciális használoinak alapvető, modern információ forrása.

Causton és Venus (1981) összehasonlította a Richards függvényt az exponenciális polinomokkal. Megállapították, hogy a Richards függvény biológiai szempontból érzékenyebb trendeket adott, ugyanakkor ritkán volt statisztikailag szignifikáns különbség a két módszer között. Causton és Venus (1981) javasolta a Richards függvény használatát a polinomokkal szemben, ahol lehetséges.

A Richards függvényt egyre nagyobb számban, egyenlő arányban alkalmazzák az egész növény és a növényi részek növekedésanalízisében. Számos növényfaj fejlődési

morfológiáját, az egyes levelek növekedését jellemezték a Richards függvénnyel. Először *Causton* (1969) írt számítógépes programot a Richards-függvény illesztésére. *Nath* és *Moore* (1992) programot dolgozott ki a Richards-függvény alkalmazására a növekedésanalízisben.

18.3.4. A vetésidő hatása a kukorica hibridek növekedésére

Kísérlet leírása

A vetésidő hatását a kukorica hibridek növekedésének és növekedési jellemzőinek dinamikájára 1995-ben és 1996-ban kisparcellás kísérletben vizsgáltuk az MTA Mezőgazdasági Kutatóintézetének (Martonvásár) kísérleti területén. A kísérlet típusa kéttényezős volt, osztott parcellás elrendezésben, négy ismétlésben. A főparcellát a vetésidő, az alparcellát a kukorica hibrid képezte. Öt vetésidő hatását vizsgáltuk (a vetésidőket a koraitól kiindulva A, B, C, D és E betűvel jelöljük: 1995-ben: A: április 13, B: április 20, C: április 29, D: május 6, E: május 16; 1996-ban: A: április 13, B: április 23, C: május 3, D: május 16, E: május 24. A kísérletben öt eltérő tenyészidejű kukorica hibrid vetésidő reakcióját tanulmányoztuk: Mara (FAO 290), Norma (FAO 370), Maya (FAO 430), DK 524 (FAO 460) és Florencia (FAO 530). A kísérletben a szemtermés-termelés növekedését hat alkalommal mértük minden évben. A teljes kukoricánövény növekedésanalízisének részeként a mintavételt a szemtermés lemorzsolhatóságakor kezdtük meg és a fiziológiai érésig folytattuk (*Berzsenyi et al.* 1999).

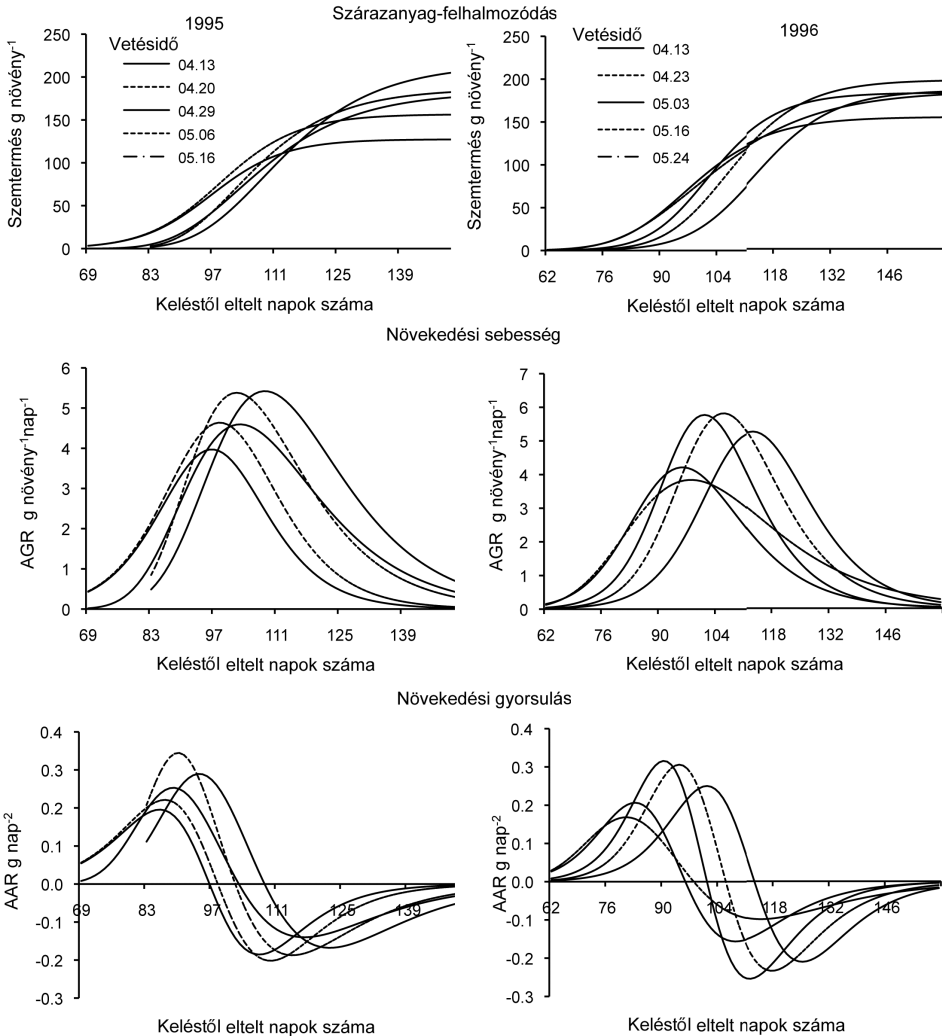
A növekedésanalízis ún. funkcionális módszerében a Richards függvényt illesztettük a mintavételi adatokhoz. Kutatásainkban a *Nath* és *Moore* (1992) által kifejlesztett növekedésanalízis programot használtuk. A program az ismétléses kísérletből származó mintavételi adatokhoz illeszti a Richards függvényt. Meghatározza a Richards függvény paramétereit, elvégzi az illesztés statisztikai próbáit. Kiszámítja az Y becslést értékeit a konfidencia intervallumokkal, továbbá a következő növekedési mutatók naponkénti értékeit: abszolút növekedési sebesség (AGR), relatív növekedési sebesség (RGR), abszolút gyorsulási sebesség (AAR), relatív gyorsulási sebesség (RAR) (*Berzsenyi*, 2000; *Berzsenyi*, 2002).

A kukoricánövény szemtermésének időbeni növekedéséhez (a szemtelítődéshez) illesztett Richards függvénnyel meghatározott szárazanyag-felhalmozódás, növekedési sebesség és növekedési gyorsulás dinamikáját a különböző vetésidőkben és években a Florencia (FAO 530) kukorica hibrid adatai alapján a 18.20. ábrán szemléltetjük. A felső ábramezőben látható, hogy a vetésidőnek jelentős hatása van a szárazanyag szemtermésbe történő beépülésének dinamikájára, másrészt a vetésidő hatása eltérő volt a vizsgált két évben. Az is jól látható, hogy a szárazanyag-beépülés a növekedési időszak során is változott a különböző vetésidőkben. Kezdetben nagyobb volt a növekedés a késői vetésidőben, azonban a tenyészidőszak második felében 1995-ben a korai, 1996-ban az átlagos vetésidő került fölénybe.

A 18.20. ábra középső ábramezője az abszolút növekedési sebesség (AGR) időbeni dinamikáját mutatja a különböző vetésidőkben és években. Az AGR jellegzetes szezonális dinamikáját harang alakú görbe jellemzi, vagyis az AGR fokozatosan nő egy maximumig, és ezt követően fokozatosan csökken. Megállapítható, hogy az AGR a maximális értéket mindkét évben korábban érte el a késői vetésidőben, majd az átlagos

gos és a korai vetésidő következett. A legfontosabb vetésidő szerinti eltérések az AGR maximális értékében figyelhetők meg.

A növekedés fő időszakát az abszolút gyorsulási sebesség (AAR) görbe írja le, amelyben jól elkülönül a gyorsulás és lassulás időszaka (18.20. ábra alsó részébe). A gyorsulási/lassulási görbe mutatja, hogy ténylegesen mennyire dinamikus a szemtermés növekedése.



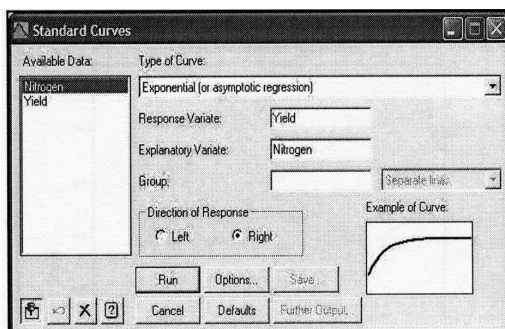
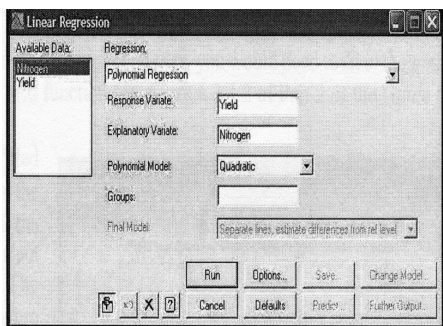
18.20. ábra. A Florencia kukorica hibrid (FAO 530) szemtermésének időbeni növekedéséhez illesztett Richards függvény alapján meghatározott szárazanyag-felhalmozódás, növekedési sebesség és gyorsulás különböző vetésidőkben 1995-ben és 1996-ban (Berzsenyi et al., 1999)

A Richards függvény paramétereivel, első, második és harmadik differenciálhányadosa alapján meghatározott növekedési mutatókkal, növekedési időpontokkal és időtartamokkal megbízhatóan tudtuk jellemezni a vetésidő, a genotípus és az évszám hatását a kukorica szemtermésének időbeni dinamikájára (Berzsenyi et al., 1999).

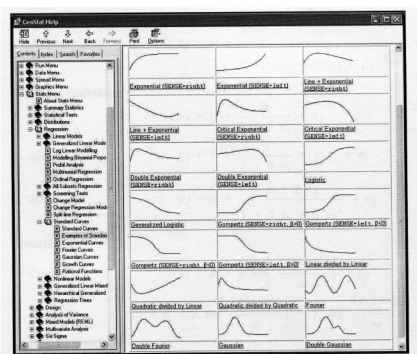
18.3.5. Nem-lineáris regresszióanalízis GenStat programmal

Polinomok illesztésekor a Linear Regression menüben kiválasztjuk a Polynomial Regression-t a legördülő listából és ezt követően a Polynomial Model-ből kiválasztjuk az illeszteni kívánt modellt (pl. Quadratic – másodfokú). Ezután specifikáljuk a reakció változót (Response Variate) és a magyarázó változót (Explanatory Variate). Végül rákattintunk a Run-ra, hogy lefusson a modell. Az illesztett modell ábráját is megkapjuk, ha a Regression Further Output menüben rákattintunk a Fitted Model gombra.

A GenStat megad továbbá több standard nem-lineáris görbét, melyeket automatikusan illeszthetünk a Standard Curves menü felhasználásával. Ehhez követjük a következő sorrendet: Stats – Regression – Standard Curves. A görbe típusát a legördülő lista dobozból (Type of Curve) választjuk ki. A kiválasztott görbe megjelenik a menü doboz jobb alsó sarkában. (18.21. ábra). A standard görbék magukba foglalják az exponenciális, logisztikus, racionális, Fourier és Gauss függvényeket (18.22. ábra).



18.21. ábra. Polinom modellek és standard görbék kiválasztása a nem-lineáris regresszióanalízisben



18.22. ábra. A standard görbék típusai a nem-lineáris regresszióanalízisben

18. 4. Többszörös regresszió-analízis

Az előző fejezetben áttekintettük két változó közötti összefüggés leírását a lineáris regresszió módszerével. A valóságban csak kontrollált (szabályozott) kísérletben áll fenn, hogy a mért változó csupán egyetlen faktorra ad reakciót; legtöbb esetben a mért változó számos, egyidejűleg ható faktorra ad reakciót. Például, állítsunk be egy szántóföldi kísérletet, hogy azonosítsuk a burgonya gumótermésének fő meghatározóit. Miközben számba vehető, hogy a növény számnak fontos szerepe van a gumótermés meghatározásában, szántóföldön várható, hogy a termést további faktorok befolyásolják, így a nitrogén- és vízellátottság, az átlagos napi hőmérséklet és így tovább, és e faktorok szintjeinek kombinációja határozza meg a végső reakciót. Ha az ilyen faktorokat kísérleti kezelésként fix szinteken állítjuk be, akkor egyedi hatásuk a reakció változóra analizálható többtenyezős ANOVA-val. Ha azonban a faktorok folytonos prediktor változók formájában vannak és megkíséreljük leírni az összefüggést a mért reakció változó és a prediktor (előrejelző) változók kombinációja között, akkor a *többszörös regresszió-analízis* módszerére van szükség.

Jóllehet a többszörös regresszió-analízis felhasználható az összefüggések elég komplex sorozatának leírására, mégis feltételezzük, hogy az összefüggés a független X-változók mindegyike és az Y reakció változó között lineáris. Gyakran az összefüggés a független X-változók és az Y reakció változó között inkább görbevonalú mint lineáris, amikor szórásdiagramon ábrázoljuk. Ekkor a *nem-lineáris regresszió* módszert alkalmazzuk.

Amikor megkíséreljük feltárni az összefüggést egy reakció változó és több lehetséges prediktor változó között a többszörös regresszió-analízissel, általában két kérdés merül fel: először, van-e szignifikáns és prediktív összefüggés a mért változó (Y) és a több prediktív változó ($X_A, X_B, X_C, \dots, X_K$) kombinációja között; és másodsor, ha van összefüggés, mely prediktor változóknak van számottevő befolyása az Y-változó megfigyelt értékének meghatározásában. Amennyiben az adatok megfelelnek bizonyos feltételeknek, akkor a többszörös regresszió-analízissel választ kaphatunk ezekre a kérdésekre.

18.4.1. Többszörös lineáris regressziós modell

Ha azok a feltételek, amelyek ahhoz szükségesek, hogy lehetséges legyen egy kétváltozós adatsorozat analízise egyszerű lineáris regresszióval, alkalmazhatók a többváltozós adatsorozathoz is, akkor a többszörös lineáris regresszió az egyszerű lineáris regressziós modell logikus kiterjesztésének tekinthető. A fő feltételek, hogy az X prediktor változók értékeinek bármely adott kombinációjára a függő Y változó elméleti populációja normális eloszlású legyen és az Y változó különböző populációinak varianciái statisztikai értelemben homogének legyenek. A lineáris regressziós modell a kétváltozós populációra a következő:

$$y = \alpha + \beta x + \varepsilon$$

Az egyenlet kiterjeszthető, hogy leírja további prediktor változók additív hatását. Tehát, ahol K prediktor változó van, az egyenlet:

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_K X_K + \varepsilon$$

ahol: Y = a függő Y -változó; X_1, X_2, \dots, X_K a prediktor (vagy magyarázó) változók; α = metszéspont az Y -tengelyen; β_i = regressziós koefficiens; ε = random hibatag.

A regressziós koefficienseket ($\beta_1, \beta_2, \dots, \beta_k$) *parciális regressziós koefficienseknek* hívjuk és együttesen meghatározzák a regressziós egyenes meredekségét. Kifejezik az X_i prediktor változó egységnyi változásának hatását az Y eredményváltozóra, a többi prediktor változó változatlansága mellett. Mivel azonban nem csupán két változó, hanem három vagy több változó közötti összefüggést veszünk figyelembe, három vagy több dimenzióban kell gondolkodnunk, és nem regressziós egyenest, hanem legjobban illeszkedő regressziós síkot határozunk meg. A háromváltozós regresszió függvény tehát egy síkot határoz meg a háromdimenziós térben. α = az Y -változó átlagos értéke, amikor minden X prediktor változó értéke zéró. A random hiba ε komponens variációját megadja az Y változó populációjának varianciája minden X változó adott értékére. Ha a minta varianciájáról feltételezhető, hogy homogén, akkor a minta *többszörös regressziós egyenlete* a következő:

$$\hat{y} = a + b_1 X_1 + b_2 X_2 + \dots + b_K X_K$$

ahol: \hat{y} = az y -változó előrejelzett értéke az X_1, X_2, \dots, X_K prediktor változók x_i értékére; a = tengelymetszet az Y tengellyel; b_i = parciális regressziós koefficiens, összefüggésben mindegyik megfelelő X prediktor változóval.

Meg kell jegyezni, hogy a modell a prediktor változókat (x_A, x_B, x_C stb.) független faktorokként kezeli, melyeknek additív hatásuk van az Y reakció változóra. Ha azonban ezek a faktorok korrelálnak egymással és/vagy egymással kölcsönhatásban vannak hatással a reakció változóra, akkor a többszörös regressziós lineáris modell értelmezése nagyon bonyolulttá válik.

Az egyenlet paramétereinek értékeit komplikált számítással lehet meghatározni, amely a legkisebb négyzetek módszerén alapul és magában foglalja a mátrix algebrát. A legtöbb modern statisztikai software csomag megadja a parciális regressziós koefficienseket (b_1, b_2, \dots, b_k) és az a konstans a standard hibáikkal együtt (b és a becült értékei a β és α populáció paramétereknek). Megadnak rendszerint egy mátrixot, amely mutatja a korrelációs koefficienseket (r) és/vagy az r^2 -t a független változók között, melyeket bevontunk a többszörös regresszió-analízisbe. Hasonlóan az egyváltozós lineáris regresszió-analízishez, mihelyt b_1, b_2, \dots, b_k és a értékei ismertek, az Y -változónak egy új értéke előrejelezhető az X prediktor változók értékeinek bármely adott kombinációjára, behelyettesítve az új X értékeket az egyenletbe. Ugyanúgy, mint amikor az egyváltozós lineáris regressziós modellt predikcióra használjuk, az exploratív predikciókat, amelyek olyan X értékeket foglalnak magukban, melyek a megfigyelt adatok tartományán kívül esnek, határozottan el kell kerülni, mivel nincs bizonyíték arra, hogy a regressziós modell a megfigyelt adatok tartományán kívül is érvényes.

A többszörös lineáris regresszió-analízisre Ireland (2010) alapján egy viszonylag egyszerű példát mutatunk be, amelyben a burgonya termése összefügg három változóval.

zóval, nevezetesen a talaj nitrogén (N), foszfor (P) és kálium (K) ellátottsági szintjeivel. Az output magában foglalja a varianciaanalízist, a regressziós modell szignifikanciájának meghatározásához.

Példa. A burgonya gumótermés reakciója a talaj nitrogén, foszfor és kálium tartalmára (18.7. táblázat)

A burgonyatermés és a talaj nitrogén-, foszfor- és kálium- (N:P:K) tartalma közötti összefüggés megállapításához a gumótermést meghatározták eltérő relatív N:P:K tartalmú tenyészedényben nevelt növényeknél (Ireland, 2010).

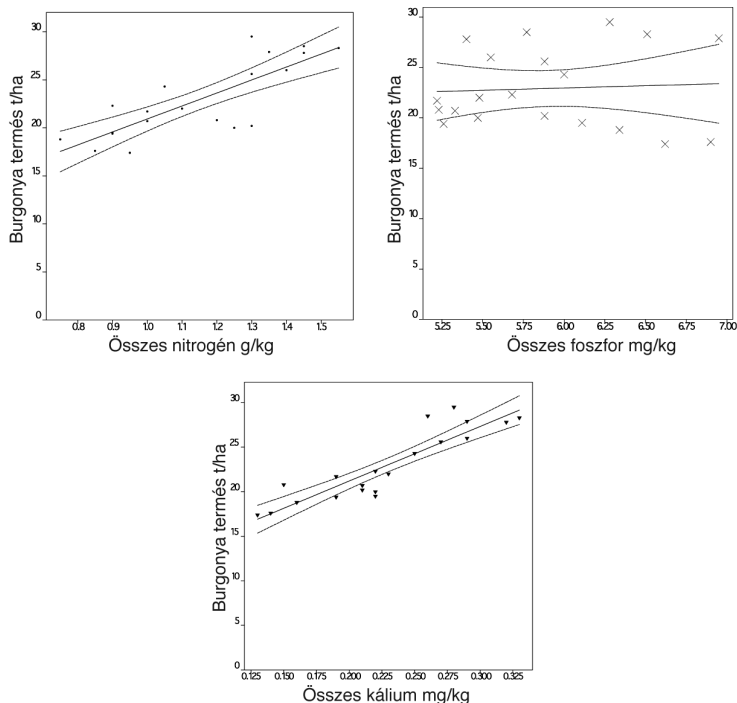
18.7. táblázat. A burgonyatermés és a talaj nitrogén-, foszfor- és káliumtartalma közötti összefüggés

Parcella	Összes talaj nitrogén (g/kg)	Összes talaj foszfor (mg/kg)	Összes talaj kálium (mg/kg)	Gumó termés (t/ha)
1	0.95	6.62	0.13	17.4
2	1.55	6.51	0.33	28.3
3	1.20	5.23	0.15	20.8
4	0.85	6.90	0.14	17.6
5	1.30	5.88	0.27	25.6
6	1.30	6.28	0.28	29.5
7	1.35	6.95	0.29	27.9
8	1.45	5.77	0.26	28.5
9	0.90	5.68	0.22	22.3
10	1.00	5.22	0.19	21.7
11	0.75	6.34	0.16	18.8
12	1.25	5.47	0.22	20.0
13	1.30	5.88	0.21	20.2
14	0.90	5.26	0.19	19.4
15	1.45	5.40	0.32	27.8
16	1.05	6.00	0.25	24.3
17	1.10	5.48	0.23	22.0
18	1.00	5.33	0.21	20.7
19	0.90	6.11	0.22	19.5
20	1.40	5.55	0.29	26.0

Többszörös lineáris regresszió (számítógépes analízisből, felhasználva a GenStat programot)

A prediktor változók közötti páronkénti korrelációs koefficiensek (r) táblázata

	Nitrogén	Foszfor	Kálium
Nitrogén	1.000		
Foszfor	0.078	1.000	
Kálium	-0.789	-0.065	1.000



18.23. ábra. Szórásdiagramok és egyváltozós lineáris regressziók a konfidencia sávokkal a burgonya gumótermése és a talaj összes N-, P- és K-tartalma közötti összefüggésre

Becsült konstans (a) = 4.141 (SE = ± 4.771)

A parciális regressziós koefficiensek (b) és statisztikai mutatóik táblázata

	b	± SE	Student t-érték	Szignifikancia	Tolerancia ^a
Nitrogén	4.44	2.754	1.613	0.126	0.377
Foszfor	0.50	0.723	0.694	0.498	0.994
Kálium	47.02	11.158	4.214	0.001	0.378

^aEzt a statisztikai mutatót a kollinearitás tesztelésére használjuk.

A többszörös lineáris modell:

$$\hat{y} = 4.14 + 4.44 x_{\text{nitrogén}} + 0.50 x_{\text{foszfor}} + 47.02 x_{\text{kálium}}$$

Null hipotézis: a burgonya termése nincs lineáris összefüggésben a három prediktor változóval ($\beta_1 = \beta_2 = \beta_3 = 0$).

ANOVA táblázat:

Variáció forrása	SS	DF	MS	F-arány	Szignifikancia
Regresszió	250.033	3	83.344	27.533	<0.01
Maradék	48.433	16	3.027		
Összes	298.465	19			

Korrigált $r^2 = 0.807$

Következtetés: a burgonya gumótermése szignifikáns lineáris összefüggésben van ($P < 0.01$) a három prediktív változóval (talaj nitrogén, foszfor és kálium ellátottsági szintje), melyet jól leírt a többszörös regressziós modell.

18.4.2. A többszörös lineáris regressziós modell szignifikanciájának tesztelése

Mielőtt felhasználnánk a kiszámított többszörös regressziós modellt a függő változó értékeinek előrejelzésére, tudni szeretnénk, mennyire jól illeszkedik a modell az adatokhoz. Ha a modell illeszkedése a megfigyelt adatokhoz nem megfelelő, akkor a modell alkalmazása a változók közötti összefüggés leírására érvénytelen. Az egyszerű lineáris regresszió tesztelésére használt Student t-teszt és az ANOVA ugyanúgy alkalmazható a többszörös lineáris regressziós modellre. Ha nincs összefüggés a függő változó és egy adott prediktor változó között, akkor a parciális regressziós koefficiens erre a prediktor változóra zéró lesz. Ezáltal a Student t-teszt felhasználható annak vizsgálatára, hogy melyik parciális regressziós koefficiens különbözik szignifikánsan zérótól. A Student t-értéket a következőképpen számítjuk ki:

$$t_{\text{megfigyelt}} = \frac{b_i}{SE_{b_i}}$$

ahol: b_i = az i-edik prediktor változó parciális regressziós koefficiense;
 SE_{b_i} = a b_i parciális regressziós koefficiens standard hibája

A szabadságfokokat $n - (p + 1)$ adja meg, ahol p a független változók összes száma. Első lépésként az ANOVA lehetővé teszi, hogy minden regressziós koefficiensre teszteljünk. Ha az F-érték szignifikáns, szükség van a t-próbákra (vagy a konfidencia intervallumokra). A null hipotézis az, hogy a becsült meredekségek, $\beta_1, \beta_2, \beta_3 \dots \beta_k$ mindegyike egyenlő zéróval. Az ANOVA ugyanazokon az elveken alapul, mint az egyváltozós lineáris regressziónál, azaz az összes variációt felosztjuk a lineáris regresszió által megmagyarázott és a random maradék hatásoknak tulajdonítható variancia között. Ha a null hipotézis igaz, akkor a regresszió variancia ($MS_{\text{regresszió}}$) és a maradék variancia ($MS_{\text{maradék}}$) értéke hasonló lesz és az F-arány ($MS_{\text{regresszió}} / MS_{\text{maradék}}$) megközelíti az 1-et. Ha a null hipotézis téves, akkor a parciális regressziós meredekségek közül legalább egy nem lesz egyenlő zéróval, úgy hogy a regresszió variancia nagyobbá válik, mint a maradék variancia és az F-érték > 1 lesz. A regresszió MS és a maradék MS meghatározásának formuláját és az ANOVA táblázat szerkezetét az 18.8. táblázat mutatja. Azt, hogy a megfigyelt F-arány elég nagy-e a null hipotézis visszautasítására adott valószínűségi szinten, megkapjuk ha összehasonlítjuk a megfelelő táblázati kritikus F-értékkel.

18.8. táblázat. Az ANOVA táblázat formája a többszörös lineáris regressziós modell szignifikanciájának tesztelésére

Variáció forrása	Eltérésnégyzet- összeg (SS)	Szabadságfokok száma (DF)	Közepes négyzetes eltérés (MS)	F-arány
Regresszió	$SS_{\text{regresszió}} = \sum (\hat{y} - \bar{y})^2$	p	$\frac{SS_{\text{regresszió}}}{DF_{\text{regresszió}}}$	$\frac{MS_{\text{regresszió}}}{MS_{\text{maradék}}}$
Maradék (reziduális)	$SS_{\text{maradék}} = \sum (y - \hat{y})^2$	n - p - 1	$\frac{SS_{\text{maradék}}}{DF_{\text{maradék}}}$	
Összes	$SS_{\text{összes}} = SS_y$	n - 1		

A fenti példában a Student t-tesztet és az ANOVA-t használtuk fel arra, hogy becsljük a többszörös regressziós modell szignifikanciáját, amely magában foglalta az összefüggést a burgonya gumótermése és a talaj N-, P- és K-tartalma között. Megjegyezzük, hogy csupán a kálium parciális regressziós koefficiense szignifikáns ($P < 0.01$), amely arra utal, hogy a kálium a fő prediktor változó a burgonyatermés szabályozásában; azonban a többszörös regressziós modell erősen szignifikáns ($P < 0.01$), utalva arra, hogy ez egy megbízható modell a termés előrejelzésére az N, P és K adott kombinált szintjeinél.

18.4.3. A többszörös regressziós modell illeszkedésének tesztelése

A kétváltozós adatsorozat megfelelő illeszkedését a regressziós modellhez becslhetjük a determinációs koefficienssel (r^2), amely úgy interpretálható, mint az Y-változó variancia-arányának mérőszáma, melyet a regresszió értelmez. Az r^2 értéket megadja a $SS_{\text{regresszió}}/SS_{\text{összes}}$ hányados. Az r^2 azonban általában nem változik, amikor több prediktor változót vonunk be a modellbe, jóllehet az extra prediktor változók bevonása javíthatja vagy csökkentheti az illeszkedés jóságát. Ezen okból nem használható arra, hogy összehasonlítsunk többszörös regressziós modelleket különböző számú prediktor változóval és kevésbé használjuk, amikor azt kívánjuk megállapítani, hogy számos lehetséges modell közül melyik adja a legjobb illeszkedést. Az illeszkedés jóságának megfelelő mutatóját a többszörös regresszió-analízisre megkaphatjuk azonban, ha a maradék MS/összes MS hányadost kivonjuk 1-ből, és ezt a mutatót korrigált (adjusted) r^2 -nek hívjuk. A korrigált r^2 -t megkapjuk:

$$r^2_{\text{korrigált}} = 1 - \frac{MS_{\text{maradék}}}{MS_{\text{összes}}} = 1 - \frac{SS_{\text{maradék}} / n - p - 1}{SS_{\text{összes}} / n - 1}$$

A korrigált r^2 értéke mindig zéró és egy között helyezkedik el és százalékos formában is kifejezhetjük. A fenti példában a korrigált $r^2 = 0.807$, amely úgy értelmezhető, hogy a burgonyatermés variációjának 80.7%-a a nitrogén, foszfor és kálium szintjeitől való többszörös lineáris függésnek tulajdonítható.

Az r^2 -től eltérően, a korrigált r^2 nő, ha a további prediktor változó modellbe történő bevonása javítja az illeszkedést és csökken, ha a további prediktor változó bevonása csökkenti az illeszkedést. Ezáltal, ha adatokat gyűjtöttünk be számos lehetsé-

ges prediktor változóra, amelyek hatással lehetnek a reakció változóra, akkor a korrigált r^2 felhasználható annak megállapítására, hogy a prediktor változók mely kombinációja adja a legjobb illeszkedést az adatokhoz a többszörös regressziós modellben. A prediktor változók hozzáadhatók vagy eltávolíthatók a modelltől, egészen addig, amíg a korrigált r^2 maximális értékét megkapjuk. Ezt az eljárást lépcsőzetes (stepwise) többszörös regresszióknak hívjuk.

18.4.4. A többszörös lineáris regressziós modell feltételeinek tesztelése

A többszörös regresszió-analízis feltételeinek érvényességét különböző diagnosztikai módszerekkel ellenőrizhetjük, hasonlóan az egyszerű lineáris regresszió-analízishez. Az a követelmény, hogy az Y -változó értékei mindegyik X -változó minden értékére normális eloszlású populációkból származzanak, homogén varianciával, becsülhető a maradék (reziduális) értékek mintázatán keresztül, amely minden Y érték eltérése a modell egyenlet által előrejelzett (becsült) értéktől ($y - \hat{y}$). Amikor a maradék értékeknek a hisztogramját ábrázoljuk, az adatok eltérése a normális elosztástól felismerhető a nem-szimmetrikus, ferde eloszlású görbe alapján. Továbbá, a kiugró (extrém) értékek meglete túlságosan nagy befolyással van a modellre, amely azonosítható a *leverage* mutatóval, minden megfigyeléssel összefüggésben. Általában, ha a leverage értéke $> 2(p/n)$, akkor a prediktor változónak azt az értékét, úgy azonosíthatjuk, hogy túlságosan nagy befolyása van a többszörös regressziós modellre.

A prediktor változók közötti kollinearitás problémája a többszörös regresszió analízisben

A többszörös regresszió-analízisnél kívánatos lenne, hogy a prediktor változók függetlenek legyenek egymástól (de ez sohasem teljesül, ezért bonyolódik az eredmények értelmezése, ezért kell folytatni az analízist. További kíváncságot, hogy additív hatásuk legyen a reakció változóra. A növény- és állattudományban ez gyakran nem áll fenn. Vegyük az előző példát, amelyben a burgonya termése összefüggött a három prediktív változóval, vagyis a talaj nitrogén, foszfor és kálium tartalmával. Miközben nagyon pontosan szabályozott tenyészedény vagy hidropon kísérletben lehetséges szabályozni mindegyik tápelemet egymástól függetlenül, szántóföldön, ahol a talaj többéves vetésforgó és teljes hatóanyag-tartalmú trágyázás hatásának van kitéve, nagyon valószínűtlen, hogy az N, P és K szintjei teljesen függetlenek legyenek egymástól. Más szavakkal, ha a talaj-ellátottság hiányos nitrogénben, akkor nagy valószínűséggel hiányos foszforban és káliumban is. A többszörös regresszió-analízisben, két vagy több prediktor változó közötti kölcsönös korreláció meglétét *kollinearitásnak* hívjuk. A kollinearitás meglete pontatlanságot okoz a többszörös regressziós modell meghatározásában és különösen azt okozza, hogy a parciális regressziós meredekségek becslése megbízhatatlanná válik. Egyik legnyilvánvalóbb hatása ennek az, hogy az adatok kismértékű változása vagy egy prediktor változó elhagyása/hozzáadása jelentősen megváltoztathatja a regressziós koefficienseket és még az előjelük is megváltozhat.

Két módja van a prediktor változók közötti kollinearitás feltárásának. Először, megvizsgáljuk a prediktor változók közötti korrelációs koefficiensek mátrixát (ezeket a sta-

tisztikai software csomagok megadják). A kollinearitást jelzi a prediktor változó-párok közötti szoros korreláció. Egy másik módszer a *tolerancia* érték meghatározása mindegyik prediktor változóval összefüggésben. A toleranciát $1 - r^2$ határozza meg, ahol r a korrelációs koeficiens, melyet az egyszerű lineáris regresszióanalízissel kapunk meg a minket érdeklő prediktor változó és a modellben lévő más prediktor változók között. A modern statisztikai software csomagok standard outputként megadják. A 0.1-nél kisebb tolerancia értékek (megfelel $r > 0.95$ -nek) általában úgy tekinthetők, hogy jelzik a szignifikáns kollinearitást. Az előző példában, míg a megfigyelt korrelációs koeficiens a nitrogén és kálium között elég magas, 0,789, mindegyik tolerancia érték jóval 0.1 felett van és ezáltal nem támogatja a kollinearitás meglétét.

Amikor kollinearitás van jelen, ennek kezelése nagyon problematikus. A legegyszerűbb megoldás eltávolítani az analízisből bármelyik prediktor változót, amely szoros korrelációt mutat bármely másik prediktor változóval.

18.4.5. Interakció a prediktor változók között a többszörös regresszió-analízisben

Annak a lehetőségét, hogy két kezelés faktor közötti interakció meghatározhatja a reakció változó nagyságrendjét, először a többtenyezős varianciaanalízis tárgyalásakor állapítottuk meg. Rámutattunk arra, hogy két kezelés faktor, amely kiválthat vagy nem válthat ki reakciót önmagában, ugyanakkor kombinációban képes extra interaktív hatást kiváltani. Hasonlóan, a többszörös regresszió-analízisben, két prediktor változó, amelynek mindegyike független hatással van a reakció változóra, kölcsönhatásban lehet egymással, hogy egy extra szintű reakciót váltsanak ki. A legegyszerűbb megközelítést véve, az interakció hatás kezelhető úgy, mint egy további független prediktor változó a többszörös regressziós modellben. Tehát, ha van két interakcióban levő prediktor változónk, X_A és X_B , interakciójuk jelölhető úgy, mint $X_A \times X_B$ és beépíthető a többszörös regresszió modellbe, mint további additív komponens:

$$y_i = \alpha + \beta_1 x_{A,i} + \beta_2 x_{B,i} + \beta_3 (x_{A,i} \times x_{B,i}) + e_i$$

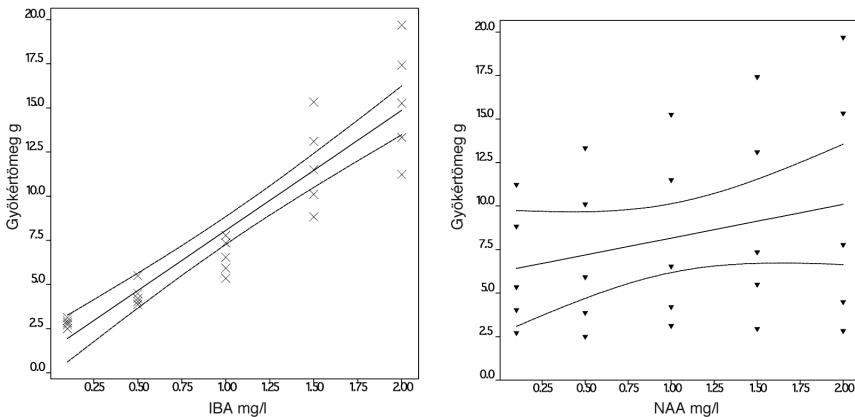
Az interakció tag parciális regressziós koeficiensének (azaz β_3 -nak) statisztikai szignifikanciája meghatározható Student t-tesztel, ahol a null hipotézis az, hogy β_3 egyenlő zéróval. A null hipotézis visszautasítása jelzi, hogy az X_A és X_B változók közötti interakció szignifikáns. Ez magában foglalja azt, hogy az Y reakció változó parciális regressziós meredeksége az első prediktor változóra (X_A) függ a második prediktor változó (X_B) értékétől. Hasonlóan, az Y reakció változó parciális regressziós meredeksége a második prediktor változóra (X_B) függ az első prediktor változó (X_A) értékétől. A varianciaanalízis bizonyíthatja a modell egyenlet általános illeszkedésének szignifikanciáját a megfigyelt adatokhoz.

Példa. Egy *Euonymus* (kecskerágó) szaporítás gyökeresedési kísérletben meghatározták a gyökerek tömegét 2 hónappal a gyökeresedést elősegítő indol butrinsav (IBA) és naftalin-ecetsav (NAA) kombinációk alkalmazását követően és a következő eredményt kapták (18.9. táblázat) (Ireland, 2010).

18.9. táblázat. A kecskerágó gyökértömege és az IBA, NAA kombinációk közötti összefüggés

IBA (mg/l)	NAA (mg/l)	Gyökértömeg(g)	IBA (mg/l)	NAA (mg/l)	Gyökértömeg(g)
0.1	0.1	2.71	1.5	0.1	8.83
0.1	0.5	2.50	1.5	0.5	10.11
0.1	1.0	3.12	1.5	1.0	11.50
0.1	1.5	2.95	1.5	1.5	13.10
0.1	2.0	2.83	1.5	2.0	15.33
0.5	0.1	4.03	2.0	0.1	11.23
0.5	0.5	3.87	2.0	0.5	13.33
0.5	1.0	4.21	2.0	1.0	15.26
0.5	1.5	5.50	2.0	1.5	17.42
0.5	2.0	4.49	2.0	2.0	19.69
1.0	0.1	5.35			
1.0	0.5	5.92			
1.0	1.0	6.54			
1.0	1.5	7.35			
1.0	2.0	7.78			

Szórásdiagramok és egyszerű lineáris regresszió (18.24. ábra):



18.24. ábra. Szórásdiagramok és lineáris regresszió konfidencia sávokkal a kecskerágó gyökértömege, valamint a gyökeresedést elősegítő IBA, illetve NAA kombinációk között

18.4.6. Többszörös lineáris regresszió a GenStat programmal

Becsült konstans (a) = 1.74 (SE = 0.631)

A parciális regressziós koefficiensek (b) és statisztikai mutatóik

	b	± SE	Student t-érték	Szignifikancia
IBA	4.4	0.515	8.53	< 0.01
NAA	-0.47	0.515	0.91	n.s.
IBA × NAA (interakció)	2.36	0.420	5.62	< 0.01

Többszörös lineáris regressziós modell:

$$\hat{y} = 1.74 + 4.4x_{\text{IBA}} - 0.47x_{\text{NAA}} + 2.36(x_{\text{IBA}} \times x_{\text{NAA}})$$

A többszörös regresszió varianciaanalízise:

Null hipotézis: nincs lineáris összefüggés a gyökértömeg és a prediktor változók között ($\beta_1 = \beta_2 = \beta_3 = 0$).

ANOVA táblázat

Variáció forrása	SS	DF	MS	F	Szignifikancia
Regresszió	607.3	3	202.44	215.06	< 0.001
Maradék	19.8	21	0.94		
Összes	627.1	24			

Következtetés: a gyökértömeg szignifikáns lineáris összefüggést mutat a prediktor változókkal (IBA, NAA, IBA \times NAA), melyet jól leír a többszörös regressziós modell.

18.4.7. Az eredmények értelmezése

A fenti példa bemutatta az *Euonymus europaeus* (kecskerágó) szaporítás gyökere-
sedési reakcióját két gyökere-
sedést elősegítő hormon, az indol butrinsav (IBA) és a naftalin-ecetsav (NAA) alkalmazására. A számítógépes statisztikai programot felhasználva, az általános lineáris modell: $\hat{y} = 1.74 + 4.4x_{\text{IBA}} - 0.47x_{\text{NAA}} + 2.36(x_{\text{IBA}} \times x_{\text{NAA}})$, ahol x_{IBA} és x_{NAA} képviseli a prediktor változókat (IBA, NAA), míg $x_{\text{IBA}} \times x_{\text{NAA}}$ képviseli az interakciójukat. Az output mutatja, hogy csak az IBA-nak van független szignifikáns hatása; azonban erősen szignifikáns interakció van; azaz a gyökere-
sedési reakció az első gyökere-
sedést elősegítő hormonra (IBA) függ a második gyökere-
sedést elősegítő hormon (NAA) szintjétől. Az output megadja a többszörös regresz-
ziós modell paramétereit és az ANOVA táblázat mutatja, hogy a többszörös regresz-
zió szignifikáns.

Regression analysis

Response variate: RootWeight

Fitted terms: Constant + IBA + NAA + IBA.NAA

Summary of analysis

Source	d.f.	s.s.	m.s.	v.r.	F pr.
Regression	3	607.31	202.4358	215.06	<.001
Residual	21	19.77	0.9413		
Total	24	627.08	26.1281		

Percentage variance accounted for 96.4

Standard error of observations is estimated to be 0.970.

Message: the following units have large standardized residuals

Unit	Response	Residual
15	7.780	-2.35

Message: the following units have high leverage

Unit	Response	Leverage
1	2.710	0.32
5	2.830	0.35
21	11.230	0.35
25	19.690	0.38

Estimates of parameters

Parameter	estimate	s.e.	t(21)	t pr.
Constant	1.736	0.631	2.75	0.012
IBA	4.396	0.515	8.53	<.001
NAA	-0.468	0.515	-0.91	0.374
IBA.NAA	2.360	0.420	5.62	<.001

18.4.8. Változók kiválasztása a regresszió-analízisben

A változó-kiválasztási (szelektációs) módszerek hasznosak olyan esetekben, amikor számos független változó van és nem rendelkezünk standard modellel. Bármely változó-kiválasztási eljárás kritériumokat tesz szükségessé annak eldöntésére, hogy mennyi változót és mely változókat válasszuk ki.

Ha a többszörös korrelációs koefficiens (R) kritériumot használnánk, az elemző mindig kiválasztaná az összes független változót, mivel az R értéke sohasem csökken további változók bevonásával az egyenletbe. Az elemzőt azonban félrevezetheti a túl sok változó bevonása. A *korrigált (adjusted) többszörös korrelációs koefficiens* (\bar{R}) kritérium választása csökkentheti a torzítást. Az elemző most kiválasztja azokat a független változókat, amelyek maximalizálják az $(\bar{R})^2$ -et. Egy másik javasolt módszer a *reziduális közepes négyzetes eltérés (RMS) minimalizálása*. Az RMS minimalizálása egyenlő az $(\bar{R})^2$ maximalizálásával. Egy másik mutató, melyet használnak a változók kiválasztásánál és a számítógépes programok tartalmazzák, az ún. *Cp kritérium*. Javasolják kiválasztani azokat a független változókat, amelyek minimalizálják a Cp értékét. További javasolt kritérium az *Akaike információs kritérium (AIC)*. Az AIC-nek egy minimális értéke kívánatos. A változók kiválasztásának módszereivel részletesen foglalkozik pl. *Afifi et al.* (2004).

A változó-kiválasztás problémája leírható úgy, mint számba venni a független változók bizonyos alsorozatát és kiválasztani azt az alsorozatot, amely vagy maximalizál vagy minimalizál egy megfelelő kritériumot. A két nyilvánvaló alsorozat a legjobb egyedüli változó és a független változók teljes sorozata. A probléma a közbelső alsorozat kiválasztásában van, amely jobb lehet mindkét extrémnél. A független változó

zóknek egy sorozatát kívánjuk kiválasztani, amelyek jó predikciót adnak és egyúttal a lehető legkevesebb számúak.

A számítógépes programokban a szignifikáns paraméterekkel rendelkező ún. optimális reakciófüggvény kiválasztására a „Forward” vagy „Enter”, a „Backward” és a „Stepwise” eliminációs módszereket alkalmazhatjuk. A *forward szelekciós módszer* egyetlen változó kiválasztásával kezdődik, kiválasztja azt a változót, amelynek korrelációja az Y-nal a legnagyobb abszolút értékben. A *forward szelekciós módszer* folytatódik oly módon, hogy hozzáad egy változót a már kiválasztott változókhoz, egészen addig, amíg egy specifikált megállási szabálynak (stopping rule) megfelelnek. A legáltalánosabban használt megállási szabály a minimális F-érték a belépéshez. A változó-kiválasztás alternatív stratégiája a *backward eliminációs módszer*. Ez a módszer azzal kezdődik, hogy minden változót figyelembe vesz a függvényben és folytatódik a legkevesbé hasznos változók egyenkénti kiszűrésével. Az F mutató megnevezése itt a számított F-érték az eltávolításhoz. A felhasználónak specifikálni kell ennek az F-értéknek a maximumát. Egy nagyon általánosan használt eljárás, amely kombinálja mindkét fenti módszert, a *stepwise eljárás*.

A számítógépek lehetővé teszik, hogy a többszörös R^2 -et (vagy más kritériumokat) és a regressziós egyenletet kiszámítsuk a változók minden lehetséges alcsoportjára (*regresszió a változók alcsoportjával*) (*subset regression*). Egy specifikált változó-méretre a változók legjobb alsorozata az, amelyiknél a legnagyobb a többszörös R^2 (vagy más választási kritérium). Az elemző így összehasonlíthatja a különböző változók legfőbb kombinációit és kiválaszhatja az előnyben részesítettet.

18.4.9. Kukorica hibridek növekedésének vizsgálata többszörös regresszióanalízissel

Kísérlet leírása

Szabadföldi kisparcellás kísérletben 1990-ben, 1991-ben, 1993-ban és 1994-ben a növekedésanalízis módszerével vizsgáltuk évente 10–12, különböző tenyészedő csoportokba tartozó kukoricahibrid növekedésének dinamikáját (*Berzsenyi és Dang, 2000b*). A kukoricahibridek jellemzésére az alábbi növekedési mutatókat számítottuk ki: abszolút növekedési sebesség (AGR), levélterület abszolút növekedési sebessége (ALGR), relatív növekedési sebesség (RGR), levélterület relatív növekedési sebessége (RLGR), levélterület index (LAI), nettó asszimilációs ráta (NAR), levélterület arány (LAR), levélterület tartósság (LAD) és biomassa tartósság (BMD). A növekedési mutatók FAO csoportok (tenyészedő) szerinti átlagos értékei eltérő irányú (csökkenő vagy emelkedő tendenciájú) és erősségű változásokat fejeztek ki.

A többszörös regresszióanalízisben a szemtermés volt a függő változó és a 9 növekedési mutató a független változó. A *regresszió-függvény* kiválasztására az „Enter”, a „Backward” és a „Stepwise” eliminációs módszereket alkalmaztuk. A „Stepwise” eljárás különösen előnyös akkor, ha a független változók egymással korrelálnak. A *varianciát infláló faktor* (VIF) mutató segítségével határoztuk meg, hogy a multikollinearitás (vagyis két vagy több független változó lineáris korrelációs kapcsolata) jelenléte milyen mértékben növeli a becsült paraméterek varianciáját. Ahogy a VIF mutató nő, úgy fokozódik a regressziós koefficiens varianciája.

18.10. táblázat. A kukorica szemtermése (függő változó) és a növekedési mutatók (független változók) közötti többszörös regresszióanalízis különböző módszereinek eredménytáblázata (n = 46)

Növekedési mutatók	Enter [†]			Backward [†]			Stepwise [†]		
	p_i^{\ddagger}	t-érték	VIF [§]	p_i^{\ddagger}	t-érték	VIF [§]	p_i^{\ddagger}	t-érték	VIF [§]
Biomassza max.	0.71	2.9 ^{xx}	11.0	0.52	3.3 ^{xx}	4.5	0.93	10.8 ^{xxx}	1.2
\overline{RGR}	-0.55	-2.2 ^x	12.0	-0.56	-2.5 ^x	9.4			
LA_{\max}	1.17	2.9 ^{xx}	30.2	0.83	3.1 ^{xx}	13.8			
\overline{ALGR}	-0.88	-3.2 ^{xx}	13.8	-0.71	3.7 ^{xxx}	7.1	-0.23	-2.7 ^{xx}	1.1
\overline{RLGR}	0.65	3.6 ^{xxx}	5.9	0.63	3.6 ^{xxx}	5.8	0.33	4.0 ^{xxx}	1.1
\overline{NAR}	0.29	1.6 ^{NS}	5.7	0.36	2.4 ^x	4.4			
LAD	-0.43	-1.2 ^{NS}	23.2						
BMD	0.05	0.2 ^{NS}	12.3						
AGR_{\max}	-0.05	-0.4 ^{NS}	2.6						
	R ² = 0.803 F-érték = 16.3 ^{xxx}			R ² = 0.794 F-érték = 25.0 ^{xxx}			R ² = 0.741 F-érték = 40.0 ^{xxx}		

[†] Eliminációs eljárások, [‡] Standardizált parciális regressziós koeficiens, [§] Varianciát infláló faktor, Szignifikancia szintek: ^{NS} = nem szignifikáns, ^xP = 5%-os, ^{xx}P = 1%-os, ^{xxx}P = 0.1%-os szinten szignifikáns.

A standardizált parciális regressziós koeficienseket a biometriában path-koeficienseknek (p_i) hívjuk. Egymáshoz viszonyított arányuk a változók jelentőségét is kifejezi. Értéke elvben -1 és +1 között változhat, előfordulnak azonban ettől eltérő értékek is (Sváb, 1979). Kiszámítását a 18.4.11. fejezetben ismertetjük.

A többszörös regresszióanalízisben az R² többszörös determinációs koeficiens kimutatta, hogy az eliminációs modellekben figyelembe vett növekedési mutatók a szemtermés varianciájának 74.1–80.3%-át magyarázzák meg (18.10. táblázat). A varianciaanalízis F-próbája mindhárom esetben P = 0.1%-os szinten szignifikáns volt. Mint-hogy a path-koeficiensek (p_i) standardizáltak, az egymáshoz viszonyított arányuk a változók jelentőségét is kifejezte. Az „Enter” eliminációs eljárásnál szignifikáns volt a hatása a maximális biomassza, az \overline{RGR} , a LA_{\max} , az \overline{ALGR} és az \overline{RLGR} mutatóknak. A LA_{\max} >1.0 értékű path-koeficiense a független változók közötti szignifikáns korrelációra utal, melyet a multikollinearitás vizsgálatában a magas VIF érték (30.2) is megerősített. A „Backward” módszernél az előző eljárásban kiválasztott szignifikáns path-koeficienseken kívül a \overline{NAR} mutató is szignifikáns volt. A path-koeficiensek számszerű értéke alapján legnagyobb volt a jelentősége a LA_{\max} , az \overline{RLGR} és a maximális biomassza mutatóknak, és ezekkel ellentétes volt a hatása az \overline{ALGR} és \overline{RGR} mutatóknak. A VIF mutató alapján a LA_{\max} path-koeficiensének varianciája volt a legnagyobb. A független változók közötti korrelációk kiszűrése után a „Stepwise” módszer mindössze három szignifikáns path-koeficienset tartalmaz, a biomassza maximumot, az \overline{ALGR} -t és az \overline{ALGR} -t. Közülük a biomassza produkció hatása a legjelentősebb a szemtermésre (0.93-as path-koeficiens).

18.4.10. A búzatermés és a növekedési mutatók regresszióanalízise

Kísérlet leírása

Kéttényezős, osztott parcellás elrendezésű tartamkísérletben vizsgáltuk a N-műtrágyázás hatását az őszi búza termésére és terméskomponenseire az MTA ATK Mezőgazdasági Intézetében, Martonvásáron (Sugár, 2014). E kutatási téma részeként többszörös regresszióanalízissel vizsgáltuk a termés, terméskomponensek, a növekedési mutatók és az egyes morfológiai paraméterek közötti összefüggéseket. A független változóknak a termésre kifejtett önálló, illetve együttes hatását a változók kiválasztásának stepwise módszerével határoztuk meg. A változók kiválasztásának kritériumaként az R^2 , a korrigált R^2 (\bar{R}^2), a C_p , valamint az AIC (Akaike információs kritérium) mutatókat használtuk. A program a R^2 és korrigált R^2 mutatók maximális, illetve a C_p , valamint AIC mutatók minimális értéke alapján választja ki a változókat. A módszer célja azoknak a változóknak az identifikálása és kiválasztása, amelyek legjobban megmagyarázzák a függő változó varianciáját (Afifi et al., 2004).

A növényenkénti termés és a növényegyed növekedési mutatói, valamint a terméskomponensek közötti „All possible subset selection” többszörös regresszióanalízis eredménye a 18.11. táblázatban látható.

A regresszió-analízis során 9 független változót vizsgáltunk, melyek közül a növényenkénti termést külön-külön meghatározó független változók (csökkenő R^2 szerint) a kalászonkénti szemszám, az $RGR_{\text{átl}}$, az $AGR_{\text{átl}}$, a zászlóslevél LAD értéke, a $LAR_{\text{átl}}$, az ezerszemtömeg és a $NAR_{\text{átl}}$. Megállapítható, hogy az R^2 alapján a termést egy-egy független változó bevonásakor legjobban meghatározza a kalászonkénti szemszám ($R^2 = 0,936$), az $RGR_{\text{átl}}$ ($R^2 = 0,793$) és az $AGR_{\text{átl}}$ ($R^2 = 0,583$). A termést együttesen legjobban meghatározó két független változó a kalászonkénti szemszám és az ezerszemtömeg.

A regresszió egyenlete: $Y = -1,136 + 0,0434_{\text{Szemszám}} + 0,02492_{\text{Ezerszemtömeg}}$.

Ezt követi (csökkenő R^2 szerint) a kalászonkénti szemszám és az $RGR_{\text{átl}}$, a kalászonkénti szemszám és az $AGR_{\text{átl}}$, valamint a kalászonkénti szemszám és a harvest index. A termést együttesen legjobban meghatározó három független változó a kalászonkénti szemszám, az ezerszemtömeg és a $RGR_{\text{átl}}$. Az összefüggés regressziós egyenlete:

$$Y = -1,486 + 0,03841_{\text{Szemszám}} + 0,02280_{\text{Ezerszemtömeg}} + 0,1907_{\text{RGRátl}}$$

Ezt követi (csökkenő R^2 szerint) a kalászonkénti szemszám, az $AGR_{\text{átl}}$ és a harvest index. A termést együttesen legjobban meghatározó négy független változó a kalászonkénti szemszám, az ezerszemtömeg, a $NAR_{\text{átl}}$ és a $LAR_{\text{átl}}$. A regresszió egyenlete:

$$Y = -1,856 + 0,04079_{\text{Szemszám}} + 0,02584_{\text{Ezerszemtömeg}} + 0,1433_{\text{NARátl}} + 0,00334_{\text{LARátl}}$$

Ezt követi (csökkenő R^2 szerint) a kalászonkénti szemszám, az ezerszemtömeg, az $AGR_{\text{átl}}$ és a BMD.

18.11. táblázat. A növényenkénti termés, a terméskomponensek és a növekedési mutatók közötti többszörös regresszióanalízis eredménye a változók kiválasztásának „stepwise” módszerével

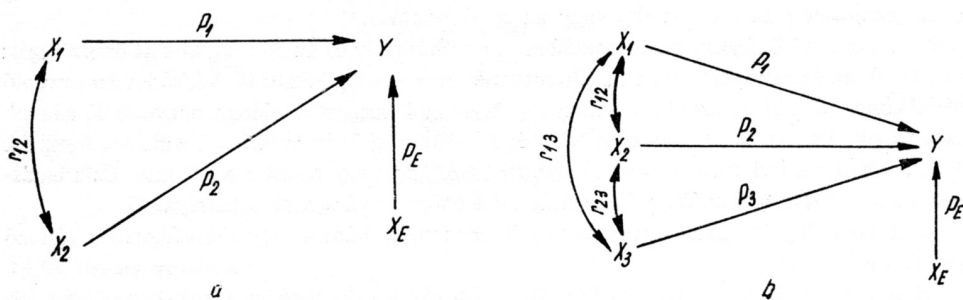
Változók száma	Változó neve	R ²	\bar{R}^2	C _p	AIC
1	Szemszám kalász ⁻¹	0,936***	0,934***	139,88	175,9
1	RGR _{átl}	0,793***	0,787***	521,78	557,8
1	AGR _{átl}	0,583***	0,571***	1082,32	1118,3
1	LAD _{zászlóslevél}	0,309***	0,288***	1815,88	1851,9
1	LAR _{átl}	0,222**	0,199**	2047,62	2083,6
1	Ezerszemtömeg	0,160*	0,135*	2213,09	2249,1
1	NAR _{átl}	0,116*	0,091*	2329,88	2365,9
1	HI	0,089 ^{NS}	0,063 ^{NS}	2402,2	2438,2
2	Szemszám kalász ⁻¹ , Ezerszemtömeg	0,982***	0,981***	18,10	54,1
2	Szemszám kalász ⁻¹ , RGR _{átl}	0,951***	0,948***	100,74	136,7
2	Szemszám kalász ⁻¹ , AGR _{átl}	0,950***	0,947***	103,89	139,9
2	Szemszám kalász ⁻¹ , HI	0,944***	0,941***	118,92	154,9
3	Szemszám kalász ⁻¹ , Ezerszemtömeg, RGR _{átl}	0,988***	0,987***	4,86	40,9
3	Szemszám kalász ⁻¹ , AGR _{átl} , HI	0,961***	0,958***	75,15	111,1
4	Szemszám kalász ⁻¹ , Ezerszemtömeg, NAR _{átl}				
	,LAR _{átl}	0,986***	0,985***	10,60	46,6
4	Szemszám kalász ⁻¹ , Ezerszemtömeg, AGR _{átl} , BMD	0,986***	0,984***	12,55	48,5

R²: többszörös determinációs együttható, \bar{R}^2 : korrigált R² érték, C_p: C_p kritérium, AIC érték: Akaike információs kritérium. Szignifikancia szintek: *P=5%, **P=1%, ***P=0,1%, ^{NS}=nem szignifikáns

18.4.11. Path-analízis

A két vagy több független változós modellben a független változók nemcsak a függő változóval, hanem egymással is kapcsolatban vannak. Ezért egy-egy független (magyarázó) változó hatása két részből, a direkt, közvetlen hatásból és az indirekt, más változókön keresztül begyűrűző hatásból tevődik össze. A regressziós koefficiensek összefüggéseiből következtetések vonhatók le a független változók közötti kapcsolatok jelentőségére. Minél erősebb a változók közötti kapcsolat, annál nagyobb a változó közvetett hatásának aránya. Feladatunk a függő és független változók közötti összefüggés felbontása, és pedig a független változó direkt hatására, valamint a többi változó indirekt hatására. A következőkben a kétváltozós korrelációs koefficiens felbontását mutatjuk be Sváb (1981) alapján. A felbontás vonatkozhat továbbá a determinációs koefficiensre és a többszörös korrelációs koefficiensre is. Megjegyezzük, hogy ez a path-analízisnek csak a bejárata, az ennél sokkal bonyolultabb és színesebb, több függő változót is megenged, stb.

A felbontás szerkezetét, két és három független változós path-diagram (út-diagram) formájában a 18.25. ábra szemlélteti. A diagramok akárhány független változóra is kiterjeszthetők. A változókat nyilak kötik össze egymással. A függő változóhoz minden független változótól nyíl vezet. A nyíl számszerű értéke a path-koefficiens, a független változó közvetlen hatását méri a függő változóra. Jele p_i. A független változókat egymással kétirányú nyíl köti össze. A kétirányú nyilak értékei a független változók közötti kétváltozós r_{ij} korrelációs koefficiensek. A path-diagram végül kiegészül az X_E eltérés (hiba) komponenssel.



18.25. ábra. Path-diagram (út-diagram) két független változóval (a) és három független változóval (b)

Kísérlet leírása

Aspergillus niger gomba micéliumtömegének alakulását vizsgálták a táptalaj N-, P- és K-tartalmától függően (modell példa). Kérdés, hogy melyik tápanyag növelése mekkora hatással van a micélium tömegének alakulására. Az alapadatokat a 18.12. táblázat tartalmazza.

18.12. táblázat. N, P és K különböző kombinációinak hatása a az *Aspergillus niger* micéliumának tömegére (Sváb, 1981)

N mg X_1	P mg X_2	K mg X_3	Micélium tömege g Y
12	72	4	1.23
12	48	4	1.17
6	72	32	1.12
6	24	16	1.11
24	8	20	1.08
30	32	12	1.02
30	32	20	1.01
18	16	12	0.99
24	16	24	0.98
18	40	28	0.93
36	24	28	0.84
42	8	8	0.78
48	56	36	0.74
48	40	32	0.74
54	56	36	0.69
Összesen 408 Átlag 27.2	544 36.3	312 20.8	14.43 0.962

A GenStat programmal végzett értékelés outputja a következő:

Regression analysis

Response variate: Micelium
Fitted terms: Constant, N, P, K

Summary of analysis

Source	d.f.	s.s.	m.s.	v.r.	F pr.
Regression	3	0.36175	0.120584	31.22	<.001
Residual	11	0.04249	0.003862		
Total	14	0.40424	0.028874		

Percentage variance accounted for 86.6

Standard error of observations is estimated to be 0.0621.

Estimates of parameters

Parameter	estimate	s.e.	t(11)	t pr.	lower 95%	upper 95%
Constant	1.2381	0.0466	26.58	<.001	1.136	1.341
N	-0.00825	0.00125	-6.60	<.001	-0.01100	-0.005496
P	0.001093	0.000835	1.31	0.217	-0.0007451	0.002930
K	-0.00439	0.00180	-2.43	0.033	-0.008366	-0.0004212

Correlations between parameter estimates

Parameter	ref	correlations			
Constant	1	1.000			
N	2	-0.468	1.000		
P	3	-0.569	0.248	1.000	
K	4	-0.211	-0.525	-0.326	1.000
		1	2	3	4

Fitted values and residuals

Unit	Response	Fitted value	Standardized residual	Leverage
1	1.2300	1.2002	0.72	0.55
2	1.1700	1.1740	-0.08	0.29
3	1.1200	1.1267	-0.16	0.53
4	1.1100	1.1445	-0.64	0.25
5	1.0800	0.9610	2.17	0.22
6	1.0200	0.9729	0.82	0.14
7	1.0100	0.9378	1.21	0.07
8	0.9900	1.0544	-1.14	0.17
9	0.9800	0.9522	0.50	0.19
10	0.9300	1.0103	-1.42	0.18
11	0.8400	0.8444	-0.08	0.14
12	0.7800	0.8653	-1.78	0.41
13	0.7400	0.7453	-0.10	0.29
14	0.7400	0.7453	-0.10	0.21
15	0.6900	0.6958	-0.12	0.37
Mean	0.9620	0.9620	-0.01	0.27

További alapszámítások:

Kiszámítjuk minden változó eltérésnégyzet-összegét (SS) és minden változó-pár eltérésszorzat-összegét (SP). A számításokat nem részletezzük.

$$\begin{array}{lll} SS_1 = 3446 & SP_{12} = -445 & SS_Y = 0.4042 \\ SS_2 = 6255 & SP_{13} = 1186 & \\ SS_3 = 1734 & SP_{23} = 781 & \end{array}$$

Mint már említettük, az Y függő változó és az X független változó közötti r_{yx} kétváltozós korrelációs koefficiens két részre bontható: az X független változó direkt hatása + indirekt hatások. Ezt a felbontást elvégezhetjük Y és bármely X független változó közötti páros korrelációs koefficiensre.

Az X változók direkt hatását a p_i path-koefficiensek fejezik ki, amelyek standardizált regressziós koefficiensek. A standardizálás képlete:

$$p_i = b_i \sqrt{\frac{SS_i}{SS_Y}}$$

Példánkban a parciális regressziós koefficienseket a számítógépes output megadta:

$$\begin{array}{l} b_1 = -0.00825 \\ b_2 = 0.00109 \\ b_3 = -0.00439 \end{array}$$

és ezekből a path-koefficiensek:

$$\begin{array}{l} p_1 = -0.00825 \sqrt{\frac{3446}{0.4042}} = -0.7618 \\ p_2 = 0.00109 \sqrt{\frac{6255}{0.4042}} = 0.1356 \\ p_3 = -0.00439 \sqrt{\frac{1734}{0.4042}} = -0.2875 \end{array}$$

Az X_E hibakomponens p_E path-értékének kiszámításához a többszörös regresszió-analízis varianciatáblázatából az eltérés SS tényezőt használjuk:

$$p_E = \sqrt{\frac{SS_E}{SS_Y}} = \sqrt{\frac{0.0424}{0.4042}} = 0.323$$

A független változók közötti kétváltozós korrelációs koefficiensek kiszámítása következik:

$$\begin{array}{l} r_{12} = \frac{SP_{12}}{\sqrt{SS_1 SS_2}} = \frac{-445}{\sqrt{3446 \cdot 6255}} = -0.0959 \\ r_{13} = \frac{SP_{13}}{\sqrt{SS_1 SS_3}} = \frac{1186}{\sqrt{3446 \cdot 1734}} = 0.4851 \\ r_{23} = \frac{SP_{23}}{\sqrt{SS_2 SS_3}} = \frac{781}{\sqrt{6255 \cdot 1734}} = 0.2371 \end{array}$$

A független változók közötti korrelációs koeficienset közvetlenül kiszámíthatjuk a GenStat programmal, melynek outputja a következő:

Correlations

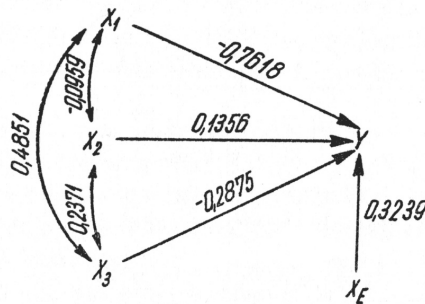
K	1	–		
N	2	0.4849	–	
P	3	0.2371	-0.0958	–
		1	2	3

Number of observations: 15

Two-sided test of correlations different from zero (P-értékek)

K	1	–		
N	2	0.0669	–	
P	3	0.3949	0.7341	–
		1	2	3

A négy path-értéket és az r_{ij} korrelációs koeficienset a 18.25.b. ábrának megfelelően feltüntetjük a path-diagramon (18.26. ábra).



18.26. ábra. Három független változós regresszióanalízis path-diagramja

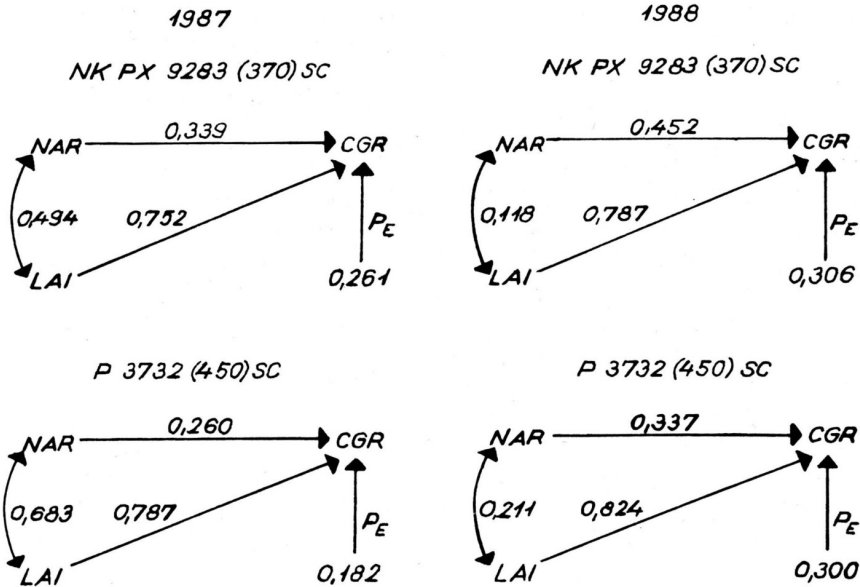
Az eredmények értékelése

Az N szoros negatív korrelációban van a micéliumtömeggel. Ez az összefüggés túlnyomóan az N direkt hatásának tulajdonítható. A P-nak elhanyagolható, a K-nak azonban figyelemre méltó az indirekt hatása. A P hatása kicsi. A K hatása negatív, és főként az N indirekt hatásának tulajdonítható. A számítások további részleteit, valamint a kétváltozós és a többszörös determinációs koeficiens felbontását Sváb (1981) ismerteti.

18.4.12. A növekedési mutatók közötti összefüggések path-diagramja

Kísérlet leírása. Az MTA Mezőgazdasági Kutatóintézetben, Martonvásáron, tartamkísérletben vizsgáltuk az N-műtrágyázás hatását a kukorica hibridek növekedési jellemzőire (Berzsenyi, 1993). A kísérleteket 1987–1988. években két egysze-

res keresztezésű, eltérő genotípusú hibriddel [NK PX 9283 (370) SC és Pioneer 3732 (450) SC] állítottuk be. A tenyésztidőszakban 1987-ben 331 mm, 1988-ban 253 mm csapadék hullott le. A vízhiány-stressz 1988-ban jelentős terméslimitáló tényező volt. A növekedésanalízishez a növényminták vételét a kukorica 4-leveles fejlettségénél kezdtük meg és a fiziológiai érésig folytattuk. A vizsgált növekedési mutatók közül, a termésnövekedési sebesség (CGR), a levélterület index (LAI) és a nettó asszimilációs ráta (NAR) között a következő összefüggés áll fenn: $CGR = NAR \times LAI$ (Hunt, 1982).

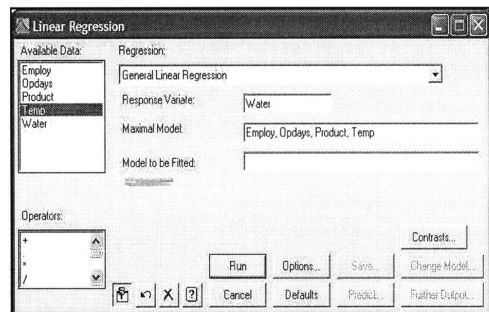
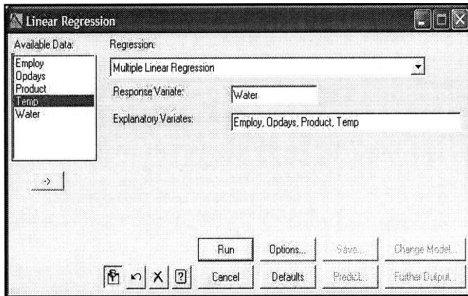


18.27. ábra. A termésnövekedési sebesség (CGR), a nettó asszimilációs ráta (NAR) és a levélterület index (LAI) közötti összefüggés path-diagramja

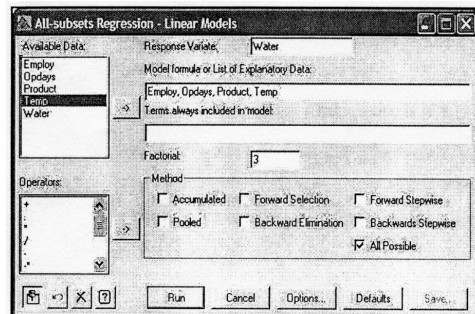
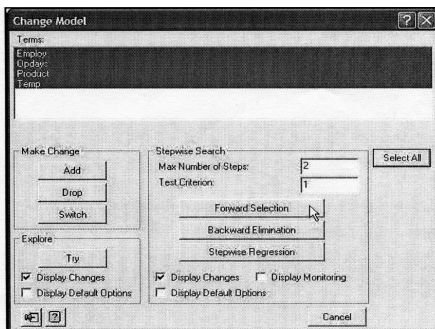
Az összefüggés-vizsgálatokkal először arra kívántunk választ kapni, hogy milyen a NAR és a LAI szerepe a CGR variabilitásában. A korrelációs számítás minden esetben szignifikáns, pozitív összefüggést mutatott a CGR és NAR, illetve a CGR és a LAI között. Különösen jelentős volt a CGR és a LAI közötti szoros, pozitív összefüggés (az r korrelációs koefficiens értéke 0.840 és 0.965 között változott). A NAR és a LAI közötti pozitív összefüggés csak kedvező évjáratban volt szignifikáns. A többszörös regresszió-analízis és a path-analízis fontos eredménye, hogy a NAR és a LAI mutatók hatásaival a CGR változását több mint 90%-ban sikerült értelmezni (az R^2 többszörös determinációs koefficiens 91–97%-ban alakult). A %-os hatások a LAI nagyobb direkt hatását igazolták (56–68% között). Csapadékos évjáratban jelentős volt a $NAR \times LAI$ kölcsönhatás is (25–28% vs. 8–12% a száraz években), ugyanakkor száraz évjáratban a NAR direkt hatása volt nagyobb (11–20% vs. 7–12% a csapadékos években). Az összefüggések felbontásának path-diagramját a 18.27. ábra szemlélteti (Berzsenyi, 1993).

18.4.13. Többszörös regresszió analízis GenStat programmal

Ha kiválasztjuk a „Multiple Linear Regression”-t a legördülő listából a „Linear Regression” menüből, specifikálhatunk számos magyarázó változót (Explanatory Variates), valamint egyetlen reakcióváltozót (Response Variate), mint a 18.28. ábra mutatja. Ahelyett azonban, hogy egyetlen lépésben illesztenénk az egész modellt, a regressziós modellek egész sorozatát illeszthetjük. Ehhez felhasználjuk a „General Linear Regression” választási lehetőséget. Ez lehetővé teszi, hogy módosítsuk a modellt, használva a „Change Model” gombot a Linear Regression menüben.



18.28. ábra. A többszörös regresszióanalízis választási lehetőségei a GenStat programban



18.29. ábra. Stepwise és All Subsets Regression választási lehetőségek a GenStat-ban

A szekvenciális illesztési módszer nagyon munkagényes lehet, ha sok változónk van. A Change Mode menü felajánlja a Stepwise regression választást, amely lehetővé teszi a modell automatikus felépítését. Minden részhalmaz regresszió elvégezhető az All Subsets Regression menü használatával. Ezt megkapjuk, ha kiválasztjuk a Stats menüből a Regresszion Analysis-t, rákattintunk az All Subsets Regression-re és ezután a Linear Models-re (18.29. ábra). Az output megadja a legjobb részhalmazt 1 változóval, 2 változóval, 3 változóval stb.

Felhasznált irodalom

- Afifi, A., Clark, V.A., May, S.: 2004. Computer-aided multivariate analysis. Fourth edition. CRC Press, Boca Raton, Florida.
- Anderson, R.L., Nelson, L.A.: 1975. A family of models involving intersecting straight lines and concomitant experimental design useful in evaluating response to fertilizer nutrients. *Biometrics* 31: 303–318.
- Archontoulis, S.V., Miguez, F.E.: 2015. Nonlinear regression models and applications in agricultural research. *Agron. J.* 107: 786–798.
- Árendás, T., Csathó, P.: 2002. Comparison of the effect of equivalent nutrients given in the form of farmyard manure or fertilizers in Hungarian long-term field trials. *Commun. Soil Sci. Plant Anal.* 30: 2861–2878.
- Baráth Csné – Itzés A. – Ugrósdý Gy.: 1996. *Biometria*. Mezőgazda Kiadó, Budapest
- Barnett, V., Payne, R., Steiner, R. (Eds.): 1995. *Agricultural sustainability. Economic, environmental and statistical considerations*. Wiley, New York.
- Bartlett, M.S.: 1937. Some examples of statistical methods of research in agriculture and biology. *J. Roy. Stat. Soc. Suppl.* 4: 137–183.
- Bertalanffy, L. von: 1957. Quantitative laws for metabolism and growth. *Q. Rev. Of Biol.* 32: 217–231.
- Berzsenyi, Z.: 1979a. A kukoricavetések gyomboritottsága és termésmennyisége közötti összefüggés. *Növénytermelés* 28: 417–426.
- Berzsenyi, Z.: 1979b. A kukoricatermelés biometriai elemzése 1974–1977. Kandidátusi értekezés, Budapest.
- Berzsenyi, Z.: 1980. A kukoricatermesztés tényezőinek összefüggésvizsgálata főkomponens-analízissel. *Növénytermelés* 29: 325–334.
- Berzsenyi, Z.: 1989. A növényszám hatása a kukorica (*Zea mays* L.) növekedésének és növekedési jellemzőire dinamikájára I. *Növénytermelés* 38: 395–407.
- Berzsenyi, Z.: 1992. A kukoricavetések egy- és kétszikű gyomnövényei elleni posztemergens védekezés újabb lehetőségeinek vizsgálata ICIA-0051 herbiciddel az 1986–1991. években. *Növénytermelés* 41: 223–236.
- Berzsenyi, Z.: 1993. A N-műtrágyázás és az évjárat hatása a kukoricahibridek (*Zea mays* L.) szemtermésére és N-műtrágyareakciójára tartamkísérletben az 1970–0991. években. *Növénytermelés* 42: 49–62.
- Berzsenyi, Z.: 2000. Növekedésanalízis a növénytermesztésben. Egyetemi jegyzet PhD hallgatóknak. Veszprémi Egyetem Georgikon Mezőgazdaságtudományi Kar, Keszthely
- Berzsenyi, Z.: 2002. A növekedésanalízis funkcionális módszere. *Növénytermelés* 51: 449–467.
- Berzsenyi, Z.: 2008. N-műtrágyázás hatása a kukorica (*Zea mays* L.) hibridek levélterületének és növénymagasságának növekedési dinamikájára tartamkísérletben. *Növénytermelés* 57: 195–210.
- Berzsenyi, Z.: 2009a. Új kihívások és módszerek a növénytermesztési kutatásban. *Növénytermelés* 58: 77–91.
- Berzsenyi, Z.: 2009b. Az ötven éves martonvásári tartamkísérletek jelentősége a növénytermesztés fejlesztésében. In: Berzsenyi, Z., Árendás, T. (szerk.): 2009. Tartamkísérletek jelentősége a növénytermesztés fejlesztésében. Jubileumi Tudományos Konferencia, Martonvásár, 2009. október 15. pp. 37–49.
- Berzsenyi, Z.: 2010. Növénytermesztési kísérletek tervezése és értékelése varianciaanalízissel. In: Radics, L. (szerk.): Fenntartható szemléletű szántóföldi növénytermesztéstan. *Agroinform*, Budapest, 419–445.

- Berzsenyi, Z., Dang, Q., L.: 2000. Különböző tenyészedejű kukorica (*Zea mays* L.) hibridek növekedésének jellemzése Richards függvényel eltérő évjáratokban. *Növénytermelés* 49: 95–116.
- Berzsenyi, Z., Dang, Q., L.: 2000b. Különböző tenyészedejű kukorica (*Zea mays* L.) hibridek növekedésanalízise Hunt–Parsons modellel és többváltozós módszerekkel. *Növénytermelés* 49: 623–640.
- Berzsenyi, Z., Dang, Q., L.: 2002a. Búza és kukorica vetésforgók elkülönítése diszkriminancia-analízissel tartamkísérletben. *Növénytermelés* 51: 21–37.
- Berzsenyi, Z., Dang, Q., L.: 2002b. Kukorica- (*Zea mays* L.) hibridek műtrágya-reakciójának vizsgálata különböző függvényekkel tartamkísérletben. *Növénytermelés* 51: 171–187.
- Berzsenyi, Z., Dang, Q., L.: 2003. A N-műtrágyázás hatása a kukorica- (*Zea mays* L.) hibridek szemtermésére és N-műtrágyareakciójára tartamkísérletben. *Növénytermelés* 52: 389–408.
- Berzsenyi, Z., Dang, Q. L.: 2004. Az istállótrágya és a műtrágya hatása a kukorica (*Zea mays* L.) termésére és termésstabilitására monokultúrás és dikultúrás tartamkísérletekben. *Növénytermelés* 53: 119–139.
- Berzsenyi, Z., Dang, Q. L.: 2006. Use of various functions to analyse the fertiliser responses of maize (*Zea mays* L.) hybrids in long-term experiments. *Acta Agronomica Hungarica* 54: 1–14.
- Berzsenyi, Z., Dang, Q.L.: 2008a. Effect of various crop production factors on the yield and stability of maize in a long-term experiment. *Cereal Research Communications* 36: 167–76.
- Berzsenyi, Z., Dang, Q.L.: 2008b. Effect of sowing date and fertilisation on the yield and yield stability of maize (*Zea mays* L.) hybrids in a long-term experiment. *Acta Agronomica Hungarica* 56: 247–267.
- Berzsenyi, Z., Györfly, B.: 1989. Comparative study of the phytotoxicity of acetanilide herbicides on maize (*Zea mays* L.) as affected by temperatures and antidotes. *Acta Agronomica Hungarica* 38: 371–384.
- Berzsenyi, Z., Györfly, B.: 1995. Különböző növénytermesztési tényezők hatása a kukorica termésére és termésstabilitására. *Növénytermelés* 44: 507–517.
- Berzsenyi, Z., Györfly, B.: 1996. A vetésforgó és a trágyázás hatása a kukorica termésére és termésstabilitására tartamkísérletben. *Növénytermelés* 45: 281–296.
- Berzsenyi, Z., Györfly, B.: 1997. Az istállótrágya és a műtrágya hatása a kukorica (*Zea mays* L.) termésére és termésstabilitására monokultúra tartamkísérletben. *Növénytermelés* 46: 509–527.
- Berzsenyi, Z., Györfly, B., Dang, Q.L.: 2000. Effect of crop rotation and fertilisation on maize and wheat yields and yield stability in a long-term experiment. *European Journal of Agronomy* 13: 225–244.
- Berzsenyi, Z., Ragab, A.J., Dang, L.Q.: 1999. A vetésidő hatásának vizsgálata a kukorica (*Zea mays* L.) szemtermésének növekedési dinamikájára Richards-függvényel. *Növénytermelés* 48: 167–187.
- Berzsenyi, Z., Varga, K., Berényi, Gy.: 1994. A növényszám és az évjárat hatása a kukorica (*Zea mays* L.) szemtermésének és terméskomponenseinek alakulására az 1981–1992. években. *Növénytermelés* 43: 61–75.
- Berzsenyi, Z., Bónis, P., Árendás, T., Berényi, Gy.: 1994. Comparative investigations on the efficacy and selectivity of different herbicides in maize. *Z. PflKrankh. PflSchutz, Sonderh. XIV*: 457–466.
- Berzsenyi, Z., Árendás, T., Bónis, P., Micskei, Gy., Sugár, E.: 2011. Long-term effect of farmyard manure and mineral fertiliser on the yield and yield stability of maize (*Zea mays* L.) in dry and wet years. *Acta Agronomica Hungarica* 59: 303–315.
- Berzsenyi, Z., Györfly, B., Árendás, T., Bónis, T., Dang, Q.L.: 1997. Studies on the phytotoxicity of herbicides in maize (*Zea mays* L.) as affected by temperature and antidotes. *Acta Agronomica Hungarica* 45: 443–448.

- Borhidi, A.: 2002. Gaia zöld ruhája. Magyarország az ezredfordulón. Stratégiai kutatások a Magyar Tudományos Akadémián. MTA, Budapest.
- Cady, F.B.: 1991. Experimental Design and Data Management of Rotation Experiments. *Agron. J.* 83: 50–56.
- Cady, F.B., Mason, D.D.: 1964. Comparison of fertility treatments in a crop rotation experiment. *Agron. J.* 56: 476–479.
- Carroll, L.: 1871. *The Adventures of Alice through the looking glass*. Macmilan, U.K.
- Casler, M.D.: 2015. Fundamentals of experimental design: Guidelines for designing successful experiments. *Agron. J.* 107: 692–705.
- Causton, D.R.: 1969. A computer program for fitting the Richards function. *Biometrics* 25: 401–9.
- Causton, D.R., Venus, J.C.: 1981. *The biometry of plant growth*. Edward Arnold, London.
- Cochran, W.G.: 1939. Long-term agricultural experiments. *J. R. Stat. Soc. [Ser A]* 6 (Suppl.): 104–148.
- Cochran, W.G., Cox, G.M.: 1957. *Experimental Design*, 2nd edn. Wiley, New York.
- Collins, C., Seeney, F.: 1999. *Statistical experiment design and interpretation*. John Wiley & Sons, Chichester, UK.
- Crossa, J.: 1990. Statistical Analysis of Multilocation Trials. *Advances in Agronomy* 44: 55–85.
- Cserhádi, S.: 2005. *Általános és különleges növénytermesztés. Második bővített kiadás*. Nitsmann József Könyvnyomdája, Győr.
- Debreczeni, B.-né, Németh, T. (szerk.): 2009. *Az Országos Műtrágyázási Tartamkísérletek (OMTK) kutatási eredményei (1967–2001)*. Akadémiai Kiadó, Budapest.
- Digby, P.G.N.: 1979. Modified joint regression analysis for incomplete variety \times environment data. *J. Agric. Sci.* 93: 81–6.
- Eberhart, S.A., Russell, W.A.: 1966. Stability parameters for comparing varieties. *Crop Sci.* 6: 36–40.
- Eisenhart, C.: 1947. The assumptions underlying the analysis of variance. *Biometrics* 3: 1–21.
- Finlay, K.W., Wilkinson, G.N.: 1963. The analysis of adaption in a plant-breeding programme. *Austral. J. Agric. Res.* 14: 742–54.
- Fisher, R.A.: 1925. *Statistical methods for research workers*. Oliver and Boyd, London.
- Fisher, R.A.: 1926. Arrangement of field experiments. *J. Ministr. Agric. G. B.* 33: 503–513.
- France, J., Thorney, J.H.M.: 1984. *Mathematical models in agriculture*. Butterworths, London.
- Gbur, E.E., Stroup, W.W., McCarter, K.s., Durham, S., Young, L.J., Christman, M., West, M., Kramer, M.: 2012. *Analysis of generalized linear mixed models in the agricultural and natural resources sciences*. American Society of Agronomy, Soil Science Society of America, Crop Science Society of America, Madison, USA.
- Gomez, K.A., Gomez, A.A.: 1984. *Statistical Procedures for Agricultural Research*. John Wiley & Sons, New York.
- Gompertz, B.: 1825. On the nature of the function expressive of the law of human mortality. *Philosophical Transactions of the Royal Society, London* 36: 513–585.
- Green, P.T.: 1997. Red crabs in rain forest on Christmas Island, Indian Ocean: activity patterns, density and biomass. *Journal of Tropical Ecology* 13: 17–38.
- Győrffy, B.: 1969. Különböző növénytermesztési tényezők hatása a kukorica termésére, komplex I. pp. 54–60. In: Ísó I. (szerk.), *Kukoricatermesztési kísérletek 1965–1968*. Akadémiai Kiadó, Budapest.
- Győrffy, B.: 1979a. Fajta-, növényszám és műtrágyahatás a kukoricatermesztésben. *Agrártudományi Közlemények* 38: 309–331.
- Győrffy, B.: 1979b. Istálló- és műtrágya hatásának értékelése kukorica-monokultúrában a hatóanyag-azonosság elve alapján, Martonvásár 1959–1974. pp. 279–289. In: Bajai, J. (szerk.), *Kukoricatermesztési kísérletek 1968–1974*. Akadémiai Kiadó, Budapest.
- Győrffy, B., Szabó, J. L., O'sváth, J.: 1969. A kukorica termésére ható növénytermesztési tényezők interakcióinak vizsgálata polyfaktorális kísérletekben, komplex II. pp. 61–77. In: Ísó I. (szerk.), *Kukoricatermesztési kísérletek 1965–1968*. Akadémiai Kiadó, Budapest.

- Hajtman, B.: 2012. Bevezetés a biostatistikába – nem csak orvosoknak. Edge 2000 Kiadó, Budapest
- Harnos Zs. – Ladányi M.: 2005. Biometria agrártudományi alkalmazásokkal. Aula Kiadó, Budapest
- Hunt, R.: 1982. Plant growth curves. The functional approach to plant growth analysis. Edward Arnold, London, UK.
- Hildebrand, P.E.: 1984. Modified stability analysis of farmer managed on-farm trials. *Agron. J.* 76: 271–274.
- Hunyadi, L., Mundruczó, Gy., Vita, L.: 2000. Statisztika. III. jav. Kiadás. Aula Kiadó, Budapest.
- Hurlbert, S.H.: 2013. Affirmation of the classical terminology for experimental design via a critique of Casella's statistical design. *Agron. J.* 105: 412–418.
- Ireland, C.: 2010. Experimental statistics for agriculture & horticulture. CABI, Wallingford, UK.
- Izsák, J., Juhász-Nagy, P., Varga, Z.: 1981. Bevezetés a biomatematikába. Tankönyvkiadó, Budapest
- Jánossy, A., Muraközi, T., Aradszky, G-né: 1966. Biometriai értelmező szótár. Mezőgazdasági Kiadó, Budapest.
- Jones, H.G.: 1992. Plants and microclimate. Second edition. Cambridge University Press, Cambridge.
- Józsa, S.: 1988. Értelmezési bizonytalanságok háttere a többváltozós elemző módszerek gyakorlati alkalmazásánál. Kandidátusi érekezés. Agrártudományi Egyetem, Keszthely
- Kang, M.S.: 1993. Simultaneous selection for yield and stability in crop performance trials: consequences for growers. *Agron. J.* 85: 754–757.
- Kang, M.S., Magari, R.: 1995. STABLE: a basic program for calculating stability and yield-stability statistics. *Agron. J.* 87: 276–277.
- Kismányoky, T., Jolánkai, M.: 2009. A magyarországi tartamkísérletek. In: Debreczeni B.-né, Németh, T. (szerk.): Az Országos Műtrágyázási Tartamkísérletek (OMTK) kutatási eredményei (1967–2001). Akadémiai Kiadó, Budapest, 25–34.
- Lin, C.S., Binns, M.R.: 1988. A method of analyzing cultivar \times location \times year experiments: a new stability parameter. *Theor. Appl. Genet.* 76: 425–430.
- Loughin, T. M.: 2006. Improved Experimental Design and Analysis for Long-Term Experiments. *Crop Sci.* 46: 2492–2502.
- Loughin, T. M., Roediger, M.P., Milliken, G.A., Schmidt, J.P.: 2007. On the analysis of long-term experiments. *J. R. Statist. Soc. A.* 170, Part 1: 29–42.
- McIntosh, M.S.: 1982. Analysis of combined experiments. *Agron. J.* 75: 153–155.
- McIntosh, M.S.: 2015. Can analysis of variance be more significant? *Agron. J.* 107: 706–717.
- McRae, K.B., Ryan, D.A.J.: 1996. Design and planning of long-term experiments. *Can. J. Plant Sci.* 76: 595–601.
- Mead, R., Curnow, R.N., Hasted, A.M.: 1993. Statistical methods in agriculture and experimental biology. Second Ed. Chapman & Hall, London, UK
- Mead, R., Curnow, R.N., Hasted, A.M.: 2003. Statistical methods in agriculture and experimental biology. Third Ed. Chapman & Hall, London, UK
- Mead, R., Riley, J., Dear, K., Sing, S.P.: 1986. Stability comparison of intercropping and monocropping systems. *Biometrics* 42: 253–266.
- MSTAT-C: 1991. User's guide to MSTAT-C. Michigan State University.
- Nagy, J.: 2009. A tápanyagellátás és a vízgazdálkodás kölcsönhatásának értékelése debreceni tartamkísérlet alapján. In: Berzsenyi, Z., Árendás, T. (szerk.): 2009. Tartamkísérletek jelentősége a növénytermesztés fejlesztésében. Jubileumi Tudományos Konferencia, Martonvásár, 2 009. október 15. pp. 21–30.
- Nagy, J., Pakurár, M., Farkas, I., Lakatos, L.: 2003. A műtrágyázás hatása a kukorica (*Zea mays* L.) termésére eltérő talajművelési változatokban. *Növénytermelés* 52: 139–146.
- Nath, S.R., Moore II.F.D.: 1992. Growth analysis by the first, second, and the third derivatives of the Richards function. *Growth, Development and Aging* 56: 237–247.

- Neyman, J., Pearson, E.S.: 1933. On the problem of the most efficient tests of statistical hypotheses. *Philos. Trans. R. Soc. London, Ser. A* 231: 289–337.
- Nielsen, D.R., Wendroth, O.: 2003. *Spatial and temporal statistics. Geoecology textbook.* Catena Verlag, Reiskirchen, Germany.
- Overman, A.R., Wilkinson, S.R., Wilson, D.M.: 1994. An extended model of forage grass response to applied nitrogen. *Agron. J.* 86: 617–620.
- Patterson, H.D.: 1964. Theory of cyclic rotation experiments. *J. R. Stat. Soc. B* 26: 1–45.
- Patterson, H.D., Silvey, V.: 1980. Statutory and recommended list trials of crop varieties in the United Kingdom (with discussion). *J. Roy. Statist. Soc. A* 143: 219–252.
- Payne, R.W. (ed.): 2000. *The Guide to GenStat: Part 2, Statistics.* VSN International, Oxford.
- Payne, R.W., Murray, D., Harding, S., Baird, D., Soutar, D., Lane, P.: 2003. *GenStat for Windows (7th Edition)* Introduction. VSN International, Oxford.
- Pepó, P.: 2009. Debreceni növénytermesztési tartamkísérletek. In: Berzsényi, Z., Árendás, T. (szerk.): 2009. *Tartamkísérletek jelentősége a növénytermesztés fejlesztésében. Jubileumi Tudományos Konferencia, Martonvásár, 2009. október 15.* pp. 133–138.
- Petersen, R.G.: 1994. *Agricultural Field Experiments. Design and Analysis.* Marcel Dekker, New York.
- Piepho, H.P.: 1998. Methods for comparing the yield stability of cropping systems – review. *J. Agron. Crop Sci.* 180: 193–213.
- Piepho, H.P., Büchse, A., Emrich, K.: 2003. A Hitchhiker's Guide to Mixed Models for Randomized Experiments. *J. Agronomy & Crop Science* 189, 310–322.
- Piepho, H.P., Büchse, A., Richter, C.: 2004. A mixed modelling approach for randomized experiments with repeated measures. *J. Agronomy & Crop Science* 190: 230–247.
- Podani, J.: 1997. *Bevezetés a többváltozós biológiai adatfeldtárás rejtelmeibe.* Scientia Kiadó, Budapest
- Précsényi, I., Barta, Z., Karsai, I., Székely, T.: 1995. Alapvető kutatásszervezési, statisztikai és projektértékelési módszerek a szupraindividuális biológiában. *Viselkedéskökológiai Kutatócsoport KLTE Evolúciós Állattani és Humánbiológiai Tanszék, Debrecen.* p. 163.
- Prékopa, I.: 2011. 25 évvel ezelőtt halt meg Sváb János, a biometria hírneves tudósa. *Magyar Tudomány:* 1140–1141.
- Quinn, G.P., Keough, M.J.: 2002. *Experimental design and data analysis for biologists.* Cambridge University Press, Cambridge, UK.
- Rasmussen, P.E., Goulding, K.W.T., Brown, J.R., Grace, P.R., Janzen, H.H., Körschens, M.: 1998. Long-term agroecosystem experiments: Assessing agricultural sustainability and global change. *Science* 282: 893–896.
- Raun, W.R., Barreto, H.J., Westerman, R.L.: 1993. Use of stability analysis for long-term soil fertility experiments. *Agron. J.* 85: 159–167.
- Richards, F.J.: 1959. A flexible growth function for empirical use. *Journal of Experimental Botany* 10: 290–300.
- Rowell, J.G., Walters, D.E.: 1976. Analysing data with repeated observations on each experimental unit. *J. Agric. Sci., Camb.* 87: 423–432.
- Saville, D.J., Rowarth, J.S.: 2008. Statistical measures, hypotheses and test sin applied research. *J. Nat. Resour. Life Sci. Educ.* 37: 74–82.
- Shukla, G.K.: 1972. Some statistical aspects of partitioning genotype-environmental components of variability. *Heredity* 29: 237–245.
- Snedecor, G.W., Cochran, W.G.: 1989. *Statistical Methods (Eighth Ed.)* Iowa State Univ. Press, Ames.
- Soós, P., E. Maajzik, P.: 1980. *Cukorrépa-hozamkísérletek értékelése kovariancia-analízissel.* Akadémiai Kiadó, Budapest.
- Steel, R.G.D., Torrie, J.H.: 1980. *Principles and Procedures of Statistics: A Biometrical Approach.* McGraw Hill, New York.

- Stevens, S.S.: 1951. Mathematics, measurement and psychophysics. In: Handbook of Experimental Psychology (Ed. S. Stevens), Wiley, New York, 1–49.
- Sugár, E.: 2014. Martonvásári búza genotípusok növekedésdinamikájának és termésprodukciónak vizsgálata különböző N-tápelem szinteken. PhD értekezés. Szent István Egyetem, Gödöllő.
- Sváb, J.: 1971. A populációgenetika alapjai. Mezőgazdasági Kiadó, Budapest.
- Sváb, J.: 1979. Többváltozós módszerek a biometriában. Mezőgazdasági Kiadó, Budapest.
- Sváb, J.: 1981. Biometriai módszerek a kutatásban. Mezőgazdasági Kiadó, Budapest.
- Sváb, J., Berzsényi, Z.: 1979. Kukoricatermés és gyomborítottság összefüggése. Esettanulmány összefüggés számítógépes ábrázolására. Növénytermelés 28: 413–416.
- Tomcsányi P.: 2000. Általános kutatómódszertan. SZIE, Gödöllő és MMI, Budapest kiadványa.
- Várallyay, Gy.: 1952. Kísérleti eredmények átvitele a gyakorlatba. Növénytermelés 1: 124–126.
- Webster, R., Payne, R.W.: 2002. Analysing repeated measurements in soil monitoring and experimentation. European Journal of Soil Science 53, 1–13.
- Williams, E.R., Matheson, A.C., Harwood, C.E.: 2002. Experimental design and analysis for tree improvement. Second Ed. CSIRO Publishing, Collingwood, Australia
- Wricke, G.: 1962. Über eine Methode zur Erfassung der ökologischen Streubreite in Feldversuchen. Z. Pflanzenzücht 47: 92–96.
- Yates, F.: 1937. The design and Analysis of Factorial Experiments. Technical Communication No 35, Imperial Bureau of Soil Science, Harpenden.
- Yates, F.: 1954. The analysis of experiments containing different crop rotations. Biometrics 10: 324–346.
- Yates, F., Cochran, W.G.: 1938. The analysis of groups of experiments. J. Agric. Sci. 28: 556–80.

A tartamkísérleti képek (16.1–16.6) forrása: Long-term experiments around the world. Special supplement to celebrate 100 years of Agronomy Journal. 2008 calander.

Melléklet

1. táblázat. Kritikus kétoldalú Student t-értékek
2. táblázat. Az F-eloszlás 5%-os (felső sor) és 1%-os (alsó sor) egyoldali kritikus értékei
3. táblázat. Kétoldalú kritikus F-értékek ($\alpha=0.05$) a variancia arány F-próbájához, két minta variancia homogenitásának teszteléséhez
4. táblázat. A Pearson korrelációs koefficiens (r) kritikus értékei
5. táblázat. A Khi-négyzet (χ^2) kritikus egyoldali értékei
6. táblázat. Kritikus kétoldalú U-értékek a Mann-Whitney próbához
7. táblázat. Kritikus Q-értékek a Kruskal-Wallis próbához (k a minták összes száma)
8. táblázat. Kritikus kétoldalú S és T értékek a páros előjelteszthez (S értékek), illetve a Wilcoxon előjelteszthez (T értékek)
9. táblázat. Kritikus q értékek ($\alpha = 0.05$) a Tukey HSD és Student-Newman-Keuls (SNK) többszörös összehasonlítási próbához
10. táblázat. Kritikus kétoldalú q' értékek ($\alpha = 0.05$) a Dunnett kétoldalú többszörös összehasonlítási próbához
11. táblázat. Spearman rangkorrelációs koefficiens (r_s) kritikus értékei
12. táblázat. Kritikus H értékek a Kruskal-Wallis többszörös összehasonlítási próbához
13. táblázat. Kritikus Friedman χ_r^2 értékek a Friedman nem-paraméteres többszörös összehasonlítási próbához
14. táblázat. Szignifikáns studentizált terjedelmek 5%-os és 1%-os szintjei a Duncan terjedelem próbához
15. táblázat. Kumulatív normális gyakorisági eloszlás (a standard normális görbe alatti terület 0-tól z-ig)

Táblázatok forrásai:

- 1–2, 4–5, 15. táblázat: Snedecor, G.W., Cochran, W.G.: 1989. Statistical Methods (Eighth Ed.) Iowa State Univ. Press, Ames.
- 3, 6–13. táblázat: Ireland, C.: 2010. Experimental statistics for agriculture & horticulture. CABI, Wallingford, UK.
14. táblázat: Gomez, K.A., Gomez, A.A.: 1984. Statistical Procedures for Agricultural Research. John Wiley & Sons, New York.

Statisztikai táblázatok

1. táblázat

Kritikus kétoldali Student t értékek

Használjuk ezeket az értékeket a konfidencia határok számítására egymintás ($n \leq 30$), kétmintás és páros t próbákhoz. (Egyoldali t próbánál a kritikus valószínűség kétszeresével számoljunk.)

Szabadság fok	A nagyobb érték valószínűsége								
	0.500	0.400	0.200	0.100	0.050	0.025	0.010	0.005	0.001
1	1.000	1.376	3.078	6.314	12.706	25.452	63.657		
2	0.816	1.061	1.886	2.920	4.303	6.205	9.925	14.089	31.598
3	0.765	0.978	1.638	2.353	3.182	4.176	5.841	7.453	12.941
4	0.741	0.941	1.533	2.132	2.776	3.495	4.604	5.598	8.610
5	0.727	0.920	1.476	2.015	2.571	3.163	4.032	4.773	6.859
6	0.718	0.906	1.440	1.943	2.447	2.969	3.707	4.317	5.959
7	0.711	0.896	1.415	1.895	2.365	2.841	3.499	4.029	5.405
8	0.706	0.889	1.397	1.860	2.306	2.752	3.355	3.832	5.041
9	0.703	0.883	1.383	1.833	2.262	2.685	3.250	3.690	4.781
10	0.700	0.879	1.372	1.812	2.228	2.634	3.169	3.581	4.587
11	0.697	0.876	1.363	1.796	2.201	2.593	3.106	3.497	4.437
12	0.695	0.873	1.356	1.782	2.179	2.560	3.055	3.428	4.318
13	0.694	0.870	1.350	1.771	2.160	2.533	3.012	3.372	4.221
14	0.692	0.868	1.345	1.761	2.145	2.510	2.977	3.326	4.140
15	0.691	0.866	1.341	1.753	2.131	2.490	2.947	3.286	4.073
16	0.690	0.865	1.337	1.746	2.120	2.473	2.921	3.252	4.015
17	0.689	0.863	1.333	1.740	2.110	2.458	2.898	3.222	3.965
18	0.688	0.862	1.330	1.734	2.101	2.445	2.878	3.197	3.922
19	0.688	0.861	1.328	1.729	2.093	2.433	2.861	3.174	3.883
20	0.687	0.860	1.325	1.725	2.086	2.423	2.845	3.153	3.850
21	0.686	0.859	1.323	1.721	2.080	2.414	2.813	3.135	3.819
22	0.686	0.858	1.321	1.717	2.074	2.406	2.819	3.119	3.792
23	0.685	0.858	1.319	1.714	2.069	2.398	2.807	3.104	3.767
24	0.685	0.857	1.318	1.711	2.064	2.391	2.797	3.090	3.745
25	0.684	0.856	1.316	1.708	2.060	2.385	2.787	3.078	3.725
26	0.684	0.856	1.315	1.706	2.056	2.379	2.779	3.067	3.707
27	0.684	0.855	1.314	1.703	2.052	2.373	2.771	3.056	3.690
28	0.683	0.855	1.313	1.701	2.048	2.368	2.763	3.047	3.674
29	0.683	0.854	1.311	1.699	2.045	2.364	2.756	3.038	3.659
30	0.683	0.854	1.310	1.697	2.042	2.360	2.750	3.030	3.646
35	0.682	0.852	1.306	1.690	2.030	2.342	2.724	2.996	3.591
40	0.681	0.851	1.303	1.684	2.021	2.329	2.704	2.971	3.551
45	0.680	0.850	1.301	1.680	2.014	2.319	2.690	2.952	3.520
50	0.680	0.849	1.299	1.676	2.008	2.310	2.678	2.937	3.496
55	0.679	0.849	1.297	1.673	2.004	2.304	2.669	2.925	3.476
60	0.679	0.848	1.296	1.671	2.000	2.299	2.660	2.915	3.460
70	0.678	0.847	1.294	1.667	2.994	2.290	2.648	2.899	3.435
80	0.678	0.847	1.293	1.665	2.989	2.284	2.638	2.887	3.416
90	0.678	0.846	1.291	1.332	2.986	2.279	2.631	2.878	3.402
100	0.677	0.846	1.290	1.661	2.982	2.276	2.625	2.871	3.390
120	0.677	0.845	1.289	1.658	2.980	2.270	2.617	2.860	3.373
∞	0.6745	0.8416	1.2816	1.6448	2.9600	2.2414	2.5758	2.8070	3.2905

2. táblázat

Az F-eloszlás 5%-os (felső sor) és 1%-os (alsó sor) egyoldaltú kritikus értékei
(Használjuk ezeket az értékeket a varianciaanalízis teszteléséhez)

V ₂ *	V ₁ df a számlálóban																				∞		
	1	2	3	4	5	6	7	8	9	10	11	12	14	16	20	24	30	40	50	75		100	200
1	161	200	216	225	230	234	237	239	241	242	243	244	245	246	248	249	250	251	252	253	253	254	254
	4.052	4.999	5.403	5.625	5.764	5.859	5.928	5.981	6.022	6.056	6.082	6.106	6.142	6.169	6.208	6.234	6.261	6.286	6.302	6.334	6.334	6.352	6.366
2	18.51	19.00	19.16	19.25	19.30	19.33	19.36	19.37	19.39	19.39	19.40	19.41	19.43	19.44	19.44	19.45	19.46	19.47	19.47	19.48	19.49	19.49	19.50
	98.49	99.00	99.17	99.25	99.30	99.33	99.36	99.37	99.39	99.40	99.41	99.42	99.43	99.44	99.45	99.46	99.47	99.48	99.48	99.49	99.49	99.50	99.50
3	10.13	9.55	9.28	9.12	9.01	8.94	8.88	8.84	8.81	8.78	8.76	8.74	8.71	8.69	8.66	8.64	8.62	8.60	8.58	8.57	8.56	8.54	8.53
	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.34	27.23	27.13	27.05	26.92	26.83	26.69	26.60	26.50	26.41	26.35	26.27	26.23	26.18	26.12
4	7.71	6.94	6.69	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.93	5.91	5.87	5.84	5.80	5.77	5.74	5.71	5.70	5.68	5.66	5.65	5.63
	21.20	18.00	16.89	15.98	15.52	15.21	14.98	14.80	14.66	14.54	14.45	14.37	14.24	14.15	14.02	13.93	13.83	13.74	13.69	13.61	13.57	13.52	13.48
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.78	4.74	4.70	4.68	4.64	4.60	4.56	4.53	4.50	4.46	4.44	4.42	4.40	4.38	4.37
	16.26	13.27	12.06	11.39	10.97	10.67	10.45	10.29	10.15	10.05	9.96	9.89	9.77	9.68	9.55	9.47	9.38	9.29	9.24	9.17	9.13	9.07	9.04
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.03	4.00	3.96	3.92	3.87	3.84	3.81	3.77	3.75	3.72	3.71	3.69	3.68
	13.74	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87	7.79	7.72	7.60	7.52	7.39	7.31	7.23	7.14	7.09	7.02	6.99	6.94	6.90
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.63	3.60	3.57	3.52	3.49	3.44	3.41	3.38	3.34	3.32	3.29	3.28	3.25	3.24
	12.25	9.55	8.45	7.85	7.46	7.19	7.00	6.84	6.71	6.62	6.54	6.47	6.35	6.27	6.15	6.07	5.98	5.90	5.85	5.78	5.75	5.70	5.67
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.34	3.31	3.28	3.23	3.20	3.15	3.12	3.08	3.05	3.03	3.00	2.98	2.96	2.94
	11.26	8.65	7.59	7.01	6.63	6.37	6.19	6.03	5.91	5.82	5.74	5.67	5.56	5.48	5.36	5.28	5.20	5.11	5.06	5.00	4.96	4.91	4.88
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.13	3.10	3.07	3.02	2.98	2.93	2.90	2.86	2.82	2.80	2.77	2.76	2.73	2.72
	10.56	8.02	6.99	6.42	6.06	5.80	5.62	5.47	5.35	5.26	5.18	5.11	5.00	4.92	4.80	4.73	4.64	4.56	4.51	4.45	4.41	4.36	4.33
10	4.93	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.97	2.94	2.91	2.86	2.82	2.77	2.74	2.70	2.67	2.64	2.61	2.59	2.56	2.55
	10.04	7.56	6.55	5.99	5.64	5.39	5.21	5.06	4.95	4.85	4.78	4.71	4.60	4.52	4.41	4.33	4.25	4.17	4.12	4.05	4.01	3.96	3.93
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.86	2.82	2.79	2.74	2.70	2.65	2.61	2.57	2.53	2.50	2.47	2.45	2.42	2.41
	9.65	7.20	6.22	5.67	5.32	5.07	4.88	4.74	4.63	4.54	4.46	4.40	4.29	4.21	4.10	4.02	3.94	3.86	3.80	3.74	3.70	3.66	3.62
12	4.75	3.88	3.49	3.26	3.11	3.00	2.92	2.85	2.80	2.76	2.72	2.69	2.64	2.60	2.54	2.50	2.46	2.42	2.40	2.36	2.35	2.32	2.31
	9.33	6.93	5.95	5.41	5.06	4.82	4.65	4.50	4.39	4.30	4.22	4.16	4.05	3.98	3.86	3.78	3.70	3.61	3.56	3.49	3.46	3.41	3.38
13	4.67	3.80	3.41	3.18	3.02	2.92	2.84	2.77	2.72	2.67	2.67	2.60	2.55	2.51	2.46	2.42	2.38	2.34	2.32	2.28	2.26	2.24	2.22
	9.07	6.70	5.74	5.20	4.86	4.62	4.44	4.30	4.19	4.10	4.02	3.96	3.85	3.78	3.67	3.59	3.51	3.42	3.37	3.30	3.27	3.21	3.18
14	4.60	3.74	3.34	3.11	2.96	2.85	2.77	2.70	2.65	2.60	2.56	2.53	2.48	2.44	2.39	2.35	2.31	2.27	2.24	2.21	2.19	2.16	2.14
	8.86	6.51	5.55	5.03	4.69	4.46	4.28	4.14	4.03	3.94	3.86	3.80	3.70	3.62	3.51	3.43	3.34	3.26	3.21	3.14	3.11	3.06	3.02
15	4.54	3.68	3.29	3.06	2.90	2.79	2.70	2.64	2.59	2.55	2.51	2.48	2.43	2.39	2.33	2.29	2.25	2.21	2.18	2.15	2.12	2.10	2.08
	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80	3.73	3.67	3.56	3.48	3.36	3.29	3.20	3.12	3.07	3.00	2.97	2.92	2.89
16	4.49	3.62	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.45	2.42	2.37	2.33	2.28	2.24	2.20	2.16	2.13	2.09	2.07	2.04	2.02
	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78	3.69	3.61	3.55	3.45	3.37	3.25	3.18	3.10	3.01	2.96	2.88	2.86	2.80	2.77
17	4.45	3.59	3.20	2.96	2.84	2.70	2.62	2.55	2.50	2.45	2.41	2.38	2.33	2.29	2.23	2.19	2.15	2.11	2.08	2.04	2.02	1.99	1.97
	8.40	6.11	5.18	4.67	4.34	4.10	3.93	3.79	3.68	3.59	3.52	3.45	3.35	3.27	3.16	3.08	3.00	2.92	2.86	2.79	2.76	2.70	2.67

*df a nevezőben Anova-nál V₁ a kezelés és V₂ a maradék szabadságfoka.

A 2. táblázat folytatása

Vz*	vi df a számlálóban																			Vz*						
	1	2	3	4	5	6	7	8	9	10	11	12	14	16	20	24	30	40	50		75	100	200	500	∞	
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.37	2.34	2.29	2.25	2.19	2.15	2.11	2.07	2.04	2.00	1.98	1.95	1.93	1.92	1.81	
	8.28	6.01	5.09	4.58	4.25	4.01	3.85	3.71	3.60	3.51	3.44	3.37	3.27	3.19	3.07	3.00	2.91	2.83	2.78	2.71	2.68	2.62	2.59	2.57	1.88	
19	4.38	3.52	3.13	2.90	2.74	2.63	2.55	2.48	2.43	2.38	2.34	2.31	2.26	2.21	2.15	2.11	2.07	2.02	2.00	1.96	1.94	1.91	1.90	1.88	1.78	
	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52	3.43	3.36	3.30	3.19	3.12	3.00	2.92	2.84	2.76	2.70	2.63	2.60	2.54	2.51	2.49	1.91	
20	4.35	3.49	3.10	2.87	2.71	2.60	2.52	2.45	2.40	2.35	2.31	2.28	2.23	2.18	2.12	2.08	2.04	1.99	1.96	1.92	1.90	1.87	1.85	1.84	1.84	
	8.10	5.58	4.94	4.43	4.10	3.87	3.71	3.56	3.45	3.37	3.30	3.23	3.13	3.05	2.94	2.86	2.77	2.69	2.63	2.56	2.53	2.47	2.44	2.42	2.0	
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.28	2.25	2.20	2.15	2.09	2.05	2.00	1.96	1.93	1.89	1.87	1.84	1.82	1.81	1.81	
	8.02	5.78	4.87	4.37	4.04	3.81	3.65	3.51	3.40	3.31	3.24	3.17	3.07	2.99	2.88	2.80	2.72	2.63	2.58	2.51	2.47	2.42	2.38	2.36	2.1	
22	4.30	3.44	3.05	2.82	2.66	2.55	2.47	2.40	2.35	2.30	2.26	2.23	2.18	2.13	2.07	2.03	1.98	1.93	1.91	1.87	1.84	1.81	1.80	1.78	1.78	
	7.94	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35	3.26	3.18	3.12	3.02	2.94	2.83	2.75	2.67	2.58	2.53	2.46	2.42	2.37	2.33	2.31	2.2	
23	4.28	3.42	3.03	2.80	2.64	2.53	2.45	2.38	2.32	2.28	2.24	2.20	2.14	2.10	2.04	2.00	1.96	1.91	1.88	1.84	1.82	1.79	1.77	1.76	1.76	
	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30	3.21	3.14	3.07	2.97	2.89	2.78	2.70	2.62	2.53	2.48	2.41	2.37	2.32	2.28	2.26	2.3	
24	4.26	3.40	3.01	2.78	2.62	2.51	2.43	2.36	2.30	2.26	2.22	2.18	2.13	2.09	2.02	1.98	1.94	1.89	1.86	1.82	1.80	1.76	1.74	1.73	1.73	
	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.25	3.17	3.09	3.03	2.93	2.85	2.74	2.66	2.58	2.49	2.44	2.36	2.33	2.27	2.23	2.21	2.4	
25	4.24	3.38	2.99	2.76	2.60	2.40	2.41	2.34	2.28	2.24	2.20	2.16	2.11	2.06	2.00	1.96	1.92	1.87	1.84	1.80	1.77	1.74	1.71	1.71	1.71	
	7.77	5.57	4.68	4.18	3.86	3.63	3.46	3.32	3.21	3.13	3.05	2.99	2.89	2.81	2.70	2.62	2.54	2.45	2.40	2.32	2.29	2.23	2.19	2.17	2.5	
26	4.22	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.18	2.15	2.10	2.05	1.99	1.95	1.90	1.85	1.82	1.78	1.76	1.72	1.70	1.69	1.69	
	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.17	3.09	3.02	2.96	2.86	2.77	2.66	2.58	2.50	2.41	2.36	2.28	2.25	2.19	2.15	2.13	2.6	
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.30	2.25	2.20	2.16	2.13	2.08	2.03	1.97	1.93	1.88	1.84	1.80	1.76	1.74	1.71	1.68	1.67	1.67	
	7.68	5.49	4.60	4.11	3.79	3.56	3.39	3.26	3.14	3.06	2.98	2.93	2.83	2.74	2.63	2.55	2.47	2.38	2.33	2.25	2.21	2.16	2.12	2.10	2.7	
28	4.20	3.34	2.95	2.71	2.56	2.44	2.36	2.29	2.24	2.19	2.15	2.12	2.06	2.02	1.96	1.91	1.87	1.81	1.78	1.75	1.72	1.69	1.67	1.65	1.65	
	7.64	5.45	4.57	4.07	3.76	3.53	3.36	3.23	3.11	3.03	2.95	2.90	2.80	2.71	2.60	2.52	2.44	2.35	2.30	2.22	2.18	2.13	2.09	2.06	2.8	
29	4.18	3.33	2.93	2.70	2.54	2.43	2.35	2.28	2.22	2.18	2.14	2.10	2.05	2.00	1.94	1.90	1.85	1.80	1.77	1.73	1.71	1.68	1.65	1.64	1.64	
	7.60	5.42	4.54	4.04	3.73	3.50	3.33	3.20	3.08	3.00	2.92	2.87	2.77	2.68	2.57	2.49	2.41	2.32	2.27	2.19	2.15	2.10	2.06	2.03	2.9	
30	4.17	3.32	2.92	2.69	2.53	2.42	2.34	2.27	2.21	2.16	2.12	2.09	2.04	1.99	1.93	1.89	1.84	1.79	1.76	1.72	1.69	1.66	1.64	1.62	1.62	
	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.06	2.98	2.90	2.84	2.74	2.66	2.55	2.47	2.38	2.29	2.24	2.16	2.13	2.07	2.03	2.01	3.0	
31	4.15	3.30	2.90	2.67	2.51	2.40	2.32	2.25	2.19	2.14	2.10	2.07	2.02	1.97	1.91	1.86	1.82	1.76	1.74	1.69	1.67	1.64	1.61	1.59	1.59	
	7.50	5.34	4.46	3.97	3.66	3.42	3.25	3.12	3.01	2.94	2.86	2.80	2.70	2.62	2.51	2.42	2.34	2.25	2.20	2.12	2.08	2.02	1.98	1.96	3.1	
34	4.13	3.28	2.88	2.65	2.49	2.38	2.30	2.23	2.17	2.12	2.08	2.05	2.00	1.95	1.89	1.84	1.80	1.74	1.71	1.67	1.64	1.61	1.59	1.57	1.57	
	7.44	5.29	4.42	3.93	3.61	3.38	3.21	3.08	2.97	2.89	2.82	2.76	2.66	2.58	2.47	2.38	2.30	2.21	2.15	2.08	2.04	1.98	1.94	1.91	3.2	
36	4.11	3.26	2.86	2.63	2.48	2.36	2.28	2.21	2.15	2.10	2.06	2.02	1.98	1.93	1.87	1.82	1.78	1.72	1.69	1.65	1.62	1.59	1.56	1.55	1.55	
	7.39	5.25	4.38	3.89	3.58	3.35	3.18	3.04	2.94	2.86	2.78	2.72	2.62	2.54	2.43	2.35	2.26	2.17	2.12	2.04	2.00	1.94	1.90	1.87	3.3	
38	4.10	3.25	2.85	2.62	2.46	2.35	2.26	2.19	2.14	2.09	2.05	2.02	1.96	1.92	1.85	1.80	1.76	1.71	1.67	1.63	1.60	1.57	1.54	1.53	1.53	
	7.35	5.21	4.34	3.86	3.54	3.32	3.15	3.02	2.91	2.82	2.75	2.69	2.59	2.51	2.40	2.32	2.22	2.14	2.08	2.00	1.97	1.90	1.86	1.84	1.84	
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.07	2.04	2.00	1.95	1.90	1.84	1.793	1.74	1.69	1.66	1.61	1.59	1.55	1.53	1.51	1.51	
	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.88	2.80	2.73	2.66	2.56	2.49	2.37	2.29	2.20	2.11	2.05	1.97	1.94	1.88	1.84	1.81	1.81	4.0

*df a nevezőben

A.2. táblázat folytatása

v ₂ *	v ₁ df a számlálóban																			v ₂ *					
	1	2	3	4	5	6	7	8	9	10	11	12	14	16	20	24	30	40	50		75	100	200	500	∞
42	4.07	3.22	2.83	2.59	2.44	2.32	2.24	2.17	2.11	2.06	2.02	1.99	1.94	1.89	1.82	1.78	1.73	1.68	1.64	1.60	1.57	1.54	1.51	1.49	1.48
	7.27	5.15	4.29	3.80	3.49	3.26	3.10	2.96	2.86	2.77	2.70	2.64	2.54	2.46	2.35	2.26	2.17	2.08	2.02	1.94	1.91	1.85	1.80	1.78	1.78
44	4.06	3.21	2.82	2.58	2.43	2.31	2.23	2.16	2.10	2.05	2.01	1.98	1.92	1.88	1.81	1.76	1.72	1.66	1.63	1.58	1.56	1.52	1.50	1.48	1.48
	7.24	5.12	4.26	3.78	3.46	3.24	3.07	2.94	2.84	2.75	2.68	2.62	2.52	2.44	2.32	2.24	2.15	2.06	2.00	1.92	1.88	1.82	1.78	1.75	1.75
46	4.05	3.20	2.81	2.57	2.42	2.30	2.22	2.14	2.09	2.04	2.00	1.97	1.91	1.87	1.80	1.75	1.71	1.65	1.62	1.57	1.54	1.51	1.48	1.46	1.46
	7.21	5.10	4.24	3.76	3.44	3.22	3.05	2.92	2.82	2.73	2.66	2.60	2.50	2.42	2.30	2.22	2.13	2.04	1.98	1.90	1.86	1.80	1.76	1.72	1.72
48	4.04	3.19	2.80	2.56	2.41	2.30	2.21	2.14	2.08	2.03	1.99	1.96	1.90	1.86	1.79	1.74	1.70	1.64	1.61	1.56	1.53	1.50	1.47	1.45	1.45
	7.19	5.08	4.22	3.74	3.42	3.20	3.04	2.90	2.80	2.71	2.64	2.58	2.48	2.40	2.28	2.20	2.11	2.02	1.96	1.88	1.84	1.78	1.73	1.70	1.70
50	4.03	3.18	2.79	2.55	2.40	2.29	2.20	2.13	2.07	2.02	1.98	1.95	1.90	1.85	1.78	1.74	1.69	1.63	1.60	1.55	1.52	1.48	1.46	1.44	1.44
	7.17	5.06	4.20	3.72	3.41	3.18	3.02	2.88	2.78	2.70	2.62	2.56	2.46	2.39	2.26	2.18	2.10	2.00	1.94	1.86	1.82	1.76	1.71	1.68	1.68
55	4.02	3.17	2.78	2.54	2.38	2.27	2.18	2.11	2.05	2.00	1.97	1.93	1.88	1.83	1.76	1.72	1.67	1.61	1.58	1.52	1.50	1.46	1.43	1.41	1.41
	7.12	5.01	4.16	3.68	3.37	3.15	2.98	2.85	2.75	2.66	2.59	2.53	2.43	2.35	2.23	2.15	2.06	1.96	1.90	1.82	1.78	1.71	1.66	1.64	1.64
60	4.00	3.15	2.76	2.52	2.37	2.25	2.17	2.10	2.04	1.99	1.95	1.92	1.86	1.81	1.75	1.70	1.65	1.59	1.56	1.50	1.48	1.44	1.41	1.39	1.39
	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63	2.56	2.50	2.40	2.32	2.20	2.12	2.05	1.93	1.87	1.79	1.74	1.68	1.63	1.60	1.60
65	3.99	3.14	2.75	2.51	2.36	2.24	2.15	2.08	2.02	1.98	1.94	1.90	1.85	1.80	1.73	1.68	1.63	1.57	1.54	1.49	1.46	1.42	1.39	1.37	1.37
	7.4	4.95	4.08	3.62	3.31	3.09	2.93	2.79	2.70	2.61	2.54	2.47	2.37	2.30	2.18	2.09	2.00	1.90	1.84	1.76	1.71	1.64	1.60	1.56	1.56
70	3.98	3.13	2.74	2.50	2.35	2.23	2.14	2.07	2.01	1.97	1.93	1.89	1.84	1.79	1.72	1.67	1.62	1.56	1.53	1.47	1.45	1.40	1.37	1.35	1.35
	7.01	4.92	4.08	3.60	3.29	3.07	2.91	2.77	2.67	2.59	2.51	2.45	2.35	2.28	2.15	2.07	1.98	1.88	1.82	1.74	1.69	1.62	1.56	1.53	1.53
80	3.96	3.11	2.72	2.48	2.33	2.21	2.12	2.05	1.99	1.95	1.91	1.88	1.82	1.77	1.70	1.65	1.60	1.54	1.51	1.45	1.42	1.38	1.35	1.32	1.32
	6.96	4.88	4.04	3.56	3.25	3.04	2.87	2.74	2.64	2.55	2.48	2.41	2.32	2.24	2.11	2.03	1.94	1.84	1.78	1.70	1.65	1.57	1.52	1.49	1.49
100	3.94	3.09	2.70	2.46	2.30	2.19	2.10	2.03	1.97	1.92	1.88	1.85	1.79	1.75	1.68	1.63	1.57	1.51	1.48	1.42	1.39	1.34	1.30	1.28	1.28
	6.90	4.82	3.98	3.51	3.20	2.99	2.82	2.69	2.59	2.51	2.43	2.36	2.26	2.19	2.06	1.98	1.89	1.79	1.73	1.64	1.59	1.51	1.46	1.43	1.43
125	3.92	3.07	2.68	2.44	2.29	2.17	2.08	2.01	1.95	1.90	1.86	1.83	1.77	1.72	1.65	1.60	1.55	1.49	1.45	1.39	1.36	1.31	1.27	1.25	1.25
	6.84	4.78	3.94	3.47	3.17	2.95	2.79	2.65	2.56	2.47	2.40	2.33	2.23	2.15	2.03	1.94	1.85	1.75	1.68	1.59	1.54	1.46	1.40	1.37	1.37
150	3.91	3.06	2.67	2.43	2.27	2.16	2.07	2.00	1.94	1.89	1.85	1.82	1.76	1.71	1.64	1.59	1.54	1.47	1.44	1.37	1.34	1.29	1.25	1.22	1.22
	6.81	4.75	3.91	3.44	3.14	2.92	2.76	2.62	2.53	2.44	2.37	2.30	2.20	2.12	2.00	1.91	1.83	1.72	1.66	1.56	1.51	1.43	1.37	1.33	1.33
200	3.89	3.04	2.65	2.41	2.26	2.14	2.05	1.98	1.92	1.87	1.83	1.80	1.74	1.69	1.62	1.57	1.52	1.45	1.42	1.35	1.32	1.26	1.22	1.19	1.19
	6.76	4.71	3.88	3.41	3.11	2.90	2.73	2.60	2.50	2.41	2.34	2.28	2.17	2.09	1.97	1.88	1.79	1.69	1.62	1.53	1.48	1.39	1.33	1.28	1.28
400	3.86	3.02	2.62	2.39	2.23	2.12	2.03	1.96	1.90	1.85	1.81	1.78	1.72	1.67	1.60	1.54	1.49	1.42	1.38	1.32	1.28	1.22	1.16	1.13	1.13
	6.70	4.66	3.83	3.36	3.06	2.85	2.69	2.55	2.46	2.37	2.29	2.23	2.12	2.04	1.92	1.84	1.74	1.64	1.57	1.47	1.42	1.32	1.24	1.19	1.19
1000	3.85	3.00	2.61	2.38	2.22	2.10	2.02	1.95	1.89	1.84	1.80	1.76	1.70	1.65	1.58	1.53	1.47	1.41	1.36	1.30	1.26	1.19	1.13	1.08	1.08
	6.66	4.62	3.80	3.34	3.04	2.82	2.66	2.52	2.43	2.34	2.26	2.20	2.09	2.01	1.89	1.81	1.71	1.61	1.54	1.44	1.38	1.28	1.19	1.11	1.11
∞	3.84	2.99	2.60	2.37	2.21	2.09	2.01	1.94	1.88	1.83	1.79	1.75	1.69	1.64	1.57	1.52	1.46	1.40	1.35	1.28	1.24	1.17	1.11	1.00	1.00
	6.63	4.60	3.78	3.32	3.02	2.80	2.64	2.51	2.41	2.32	2.24	2.18	2.07	1.99	1.87	1.79	1.69	1.59	1.52	1.41	1.36	1.25	1.15	1.00	1.00

*df a nevezőben

Kétoldali kritikus F-értékek ($\alpha=0.05$) a variancia arány F-próbához két minta variancia homogenitásának teszteléséhez

V_2	V_1																		∞							
	1	2	3	4	5	6	7	8	9	10	12	15	20	25	30	35	40	45		50	60	70	80	90	100	120
1	647.8	799.5	864.2	899.6	921.8	937.1	948.2	956.6	963.3	968.6	976.7	984.9	993.1	998.1	1001	1004	1006	1007	1008	1010	1011	1012	1013	1014	1014	1018
2	38.51	39.00	39.17	39.25	39.30	39.33	39.36	39.37	39.39	39.40	39.41	39.43	39.45	39.46	39.46	39.47	39.47	39.48	39.48	39.48	39.48	39.48	39.49	39.49	39.49	39.50
3	17.44	16.04	15.44	15.10	14.88	14.73	14.62	14.54	14.47	14.42	14.34	14.25	14.17	14.12	14.08	14.06	14.04	14.02	14.01	13.99	13.98	13.97	13.96	13.96	13.96	13.90
4	12.22	10.65	9.98	9.60	9.36	9.20	9.07	8.98	8.93	8.88	8.84	8.75	8.66	8.58	8.46	8.33	8.30	8.28	8.26	8.25	8.24	8.23	8.23	8.23	8.23	8.16
5	10.01	8.43	7.76	7.39	7.15	6.98	6.85	6.76	6.68	6.62	6.52	6.43	6.27	6.23	6.20	6.18	6.16	6.14	6.12	6.11	6.10	6.09	6.08	6.07	6.07	6.02
6	8.81	7.26	6.60	6.23	5.99	5.82	5.70	5.60	5.52	5.46	5.37	5.27	5.17	5.11	5.07	5.04	5.01	4.99	4.98	4.96	4.94	4.93	4.92	4.92	4.90	4.85
7	8.07	6.56	5.89	5.52	5.29	5.12	4.99	4.90	4.82	4.78	4.67	4.57	4.40	4.36	4.33	4.31	4.29	4.28	4.25	4.24	4.23	4.22	4.22	4.22	4.22	4.14
8	7.57	6.06	5.42	5.05	4.82	4.65	4.53	4.43	4.36	4.30	4.20	4.10	4.00	3.94	3.89	3.86	3.84	3.82	3.81	3.78	3.77	3.76	3.75	3.74	3.73	3.67
9	7.21	5.71	5.08	4.72	4.48	4.32	4.20	4.10	4.03	3.96	3.87	3.77	3.67	3.60	3.56	3.53	3.51	3.49	3.47	3.45	3.43	3.42	3.41	3.40	3.39	3.33
10	6.94	5.46	4.83	4.47	4.24	4.08	3.95	3.85	3.78	3.72	3.62	3.52	3.42	3.35	3.31	3.28	3.26	3.24	3.22	3.20	3.18	3.17	3.16	3.15	3.14	3.08
11	6.72	5.25	4.63	4.28	4.04	3.88	3.76	3.66	3.59	3.53	3.43	3.33	3.23	3.16	3.12	3.09	3.06	3.04	3.02	3.00	2.99	2.97	2.96	2.95	2.94	2.88
12	6.55	5.10	4.47	4.12	3.88	3.73	3.61	3.51	3.44	3.37	3.28	3.18	3.07	3.01	2.96	2.93	2.91	2.89	2.87	2.85	2.83	2.82	2.81	2.80	2.79	2.72
13	6.41	4.97	4.35	4.00	3.77	3.60	3.48	3.39	3.31	3.25	3.15	3.05	2.95	2.88	2.84	2.80	2.78	2.76	2.74	2.72	2.70	2.69	2.68	2.67	2.65	2.60
14	6.30	4.86	4.24	3.89	3.66	3.50	3.38	3.29	3.21	3.15	3.05	2.95	2.84	2.78	2.73	2.70	2.67	2.65	2.64	2.61	2.60	2.58	2.57	2.56	2.55	2.49
15	6.20	4.77	4.15	3.80	3.58	3.41	3.29	3.20	3.12	3.06	2.96	2.86	2.76	2.69	2.64	2.61	2.59	2.56	2.55	2.52	2.51	2.49	2.48	2.47	2.46	2.40
16	6.12	4.69	4.08	3.73	3.50	3.34	3.22	3.12	3.05	2.99	2.89	2.79	2.68	2.61	2.57	2.54	2.51	2.49	2.47	2.45	2.43	2.42	2.40	2.40	2.38	2.32
17	6.04	4.62	4.01	3.66	3.44	3.28	3.16	3.06	2.98	2.92	2.82	2.72	2.62	2.55	2.50	2.47	2.44	2.42	2.41	2.38	2.36	2.35	2.34	2.33	2.32	2.25
18	5.98	4.56	3.95	3.61	3.38	3.22	3.10	3.01	2.93	2.87	2.77	2.67	2.56	2.49	2.44	2.41	2.38	2.36	2.35	2.32	2.30	2.29	2.28	2.27	2.26	2.19
19	5.92	4.51	3.90	3.56	3.33	3.17	3.05	2.96	2.88	2.82	2.72	2.62	2.51	2.44	2.39	2.36	2.33	2.31	2.30	2.27	2.25	2.24	2.23	2.22	2.20	2.13
20	5.87	4.46	3.86	3.51	3.28	3.13	3.01	2.91	2.84	2.77	2.68	2.57	2.46	2.40	2.35	2.31	2.29	2.27	2.25	2.22	2.20	2.19	2.18	2.17	2.16	2.09
21	5.83	4.42	3.82	3.48	3.25	3.09	2.97	2.87	2.80	2.73	2.64	2.53	2.42	2.36	2.31	2.27	2.25	2.23	2.21	2.18	2.15	2.14	2.13	2.11	2.10	2.04
22	5.79	4.38	3.78	3.44	3.22	3.05	2.93	2.84	2.76	2.70	2.60	2.50	2.39	2.32	2.27	2.24	2.21	2.19	2.17	2.14	2.11	2.11	2.10	2.09	2.08	2.00
23	5.75	4.35	3.75	3.41	3.18	3.02	2.90	2.81	2.73	2.67	2.57	2.47	2.36	2.29	2.24	2.20	2.18	2.15	2.14	2.11	2.09	2.08	2.07	2.06	2.04	1.97
24	5.72	4.32	3.72	3.38	3.15	2.99	2.87	2.78	2.70	2.64	2.54	2.44	2.33	2.26	2.21	2.17	2.15	2.12	2.11	2.08	2.06	2.05	2.03	2.02	2.01	1.94
25	5.69	4.29	3.69	3.35	3.13	2.97	2.85	2.75	2.68	2.61	2.51	2.41	2.30	2.23	2.18	2.15	2.12	2.10	2.08	2.05	2.03	2.02	2.01	2.00	1.98	1.91
26	5.66	4.27	3.67	3.33	3.10	2.94	2.82	2.73	2.65	2.59	2.49	2.39	2.28	2.21	2.16	2.12	2.09	2.07	2.05	2.03	2.01	1.99	1.98	1.97	1.95	1.88
27	5.63	4.24	3.65	3.31	3.08	2.92	2.80	2.71	2.63	2.57	2.47	2.36	2.25	2.18	2.13	2.10	2.07	2.05	2.03	2.01	1.98	1.96	1.94	1.93	1.92	1.85
28	5.61	4.22	3.63	3.29	3.06	2.90	2.78	2.69	2.61	2.55	2.45	2.34	2.23	2.16	2.11	2.08	2.05	2.03	2.01	1.98	1.96	1.94	1.92	1.91	1.89	1.81
29	5.59	4.20	3.61	3.27	3.04	2.88	2.76	2.67	2.59	2.53	2.43	2.32	2.21	2.14	2.09	2.06	2.03	2.01	1.99	1.96	1.94	1.92	1.91	1.90	1.89	1.81
30	5.57	4.18	3.59	3.25	3.03	2.87	2.75	2.65	2.57	2.51	2.41	2.31	2.20	2.12	2.07	2.04	2.01	1.99	1.97	1.94	1.92	1.90	1.89	1.88	1.87	1.79
40	5.42	4.05	3.46	3.13	2.90	2.74	2.62	2.53	2.45	2.39	2.29	2.18	2.07	1.99	1.94	1.90	1.88	1.85	1.83	1.80	1.78	1.76	1.75	1.74	1.72	1.64
50	5.34	3.97	3.39	3.05	2.83	2.67	2.55	2.46	2.38	2.32	2.22	2.11	1.99	1.92	1.87	1.88	1.80	1.77	1.75	1.72	1.70	1.68	1.67	1.66	1.64	1.55
60	5.29	3.93	3.34	3.01	2.79	2.63	2.51	2.41	2.33	2.27	2.17	2.06	1.94	1.87	1.82	1.78	1.74	1.72	1.70	1.67	1.64	1.63	1.61	1.60	1.58	1.48
80	5.22	3.86	3.28	2.95	2.73	2.57	2.45	2.35	2.28	2.21	2.11	2.00	1.88	1.81	1.75	1.71	1.68	1.65	1.63	1.60	1.57	1.55	1.54	1.53	1.51	1.40
100	5.18	3.83	3.25	2.92	2.70	2.54	2.42	2.32	2.24	2.18	2.08	1.97	1.85	1.77	1.71	1.67	1.64	1.61	1.59	1.56	1.53	1.51	1.50	1.48	1.46	1.35
120	5.15	3.80	3.23	2.89	2.67	2.52	2.39	2.30	2.22	2.16	2.05	1.94	1.82	1.75	1.69	1.65	1.61	1.59	1.56	1.53	1.50	1.48	1.47	1.45	1.43	1.31
∞	5.02	3.69	3.12	2.79	2.57	2.41	2.29	2.19	2.11	2.05	1.94	1.83	1.71	1.63	1.57	1.52	1.48	1.45	1.43	1.39	1.36	1.33	1.31	1.30	1.27	1.00

Két minta varianciájának összehasonlításakor V_1 a nagyobb varianciájú, V_2 a kisebb varianciájú minta df-je.

A Pearson korrelációs koefficiens (r) kritikus értékei

df*	10%	5%	2%	1%
3	0.805	0.878	0.934	0.959
4	0.729	0.811	0.882	0.917
5	0.669	0.754	0.833	0.874
6	0.622	0.707	0.789	0.834
7	0.582	0.666	0.750	0.798
8	0.549	0.632	0.716	0.765
9	0.521	0.602	0.685	0.735
10	0.497	0.576	0.658	0.708
11	0.476	0.553	0.634	0.684
12	0.458	0.532	0.612	0.661
13	0.441	0.514	0.592	0.641
14	0.426	0.482	0.574	0.623
15	0.412	0.412	0.558	0.606
16	0.400	0.468	0.542	0.590
17	0.389	0.456	0.528	0.575
18	0.378	0.444	0.516	0.549
19	0.369	0.433	0.503	0.503
20	0.360	0.423	0.492	0.537
25	0.323	0.381	0.445	0.487
30	0.295	0.349	0.409	0.449
35	0.275	0.325	0.381	0.418
40	0.257	0.304	0.358	0.393
45	0.243	0.288	0.338	0.372
50	0.231	0.273	0.322	0.354
60	0.211	0.250	0.295	0.325
70	0.195	0.232	0.274	0.302
80	0.183	0.217	0.256	0.283
90	0.173	0.205	0.242	0.267
100	0.164	0.195	0.230	0.254
150	0.134	0.160	0.189	0.208
200	0.116	0.138	0.164	0.181
300	0.095	0.113	0.134	0.148
400	0.082	0.098	0.116	0.128
500	0.073	0.088	0.104	0.115

*n mintaméretre df= n-2

A Khi-négyzet (χ^2) kritikus egyoldalú értékei
 (Használjuk ezeket az értékeket az illeszkedés Khi-négyzet tesztjéhez és a kontingenciatablázatok tesztjéhez)

Szabadság fok	Egy nagyobb érték valószínűsége														
	0.995	0.990	0.975	0.950	0.900	0.750	0.500	0.250	0.100	0.050	0.025	0.010	0.005		
1	0.02	0.10	0.45	1.32	2.71	3.84	5.02	6.63	7.88		
2	0.01	0.02	0.05	0.10	0.21	0.58	1.39	2.77	4.61	5.99	7.38	9.21	10.60		
3	0.07	0.11	0.22	0.35	0.58	1.21	2.37	4.11	6.25	7.81	9.35	11.34	12.84		
4	0.21	0.30	0.48	0.71	1.06	1.92	3.36	5.39	7.78	9.49	11.14	13.28	14.86		
5	0.41	0.55	0.83	1.15	1.61	2.67	4.35	6.63	9.24	11.07	12.83	15.09	16.75		
6	0.68	0.87	1.24	1.64	2.20	3.45	5.35	7.84	10.64	12.59	14.45	16.81	18.55		
7	0.99	1.24	1.69	2.17	2.83	4.25	6.35	9.04	12.02	14.07	16.01	18.48	20.28		
8	1.34	1.65	2.18	2.73	3.49	5.07	7.34	10.22	13.36	15.51	17.53	20.09	21.96		
9	1.73	2.09	2.70	3.33	4.17	5.90	8.34	11.39	14.68	16.92	19.02	21.67	23.59		
10	2.16	2.56	3.25	3.94	4.87	6.74	9.34	12.55	15.99	18.31	20.48	23.21	25.19		
11	2.60	3.05	3.82	4.57	5.58	7.58	10.34	13.70	17.28	19.68	21.92	24.72	26.76		
12	3.07	3.57	4.40	5.23	6.30	8.44	11.34	14.85	18.55	21.03	23.34	26.22	28.30		
13	3.57	4.11	5.01	5.89	7.04	9.30	12.34	15.98	19.81	22.36	24.74	27.69	29.82		
14	4.07	4.66	5.63	6.57	7.79	10.17	13.34	17.12	21.06	23.68	26.12	29.14	31.32		
15	4.60	5.23	6.27	7.26	8.55	11.04	14.34	18.25	22.31	25.00	27.4*9	30.58	32.80		
16	5.14	5.81	6.91	7.96	9.31	11.91	15.34	19.37	23.54	26.30	28.85	32.00	34.27		
17	5.70	6.41	7.56	8.67	10.09	12.79	16.34	20.49	24.77	27.59	30.19	33.41	35.72		
18	6.26	7.01	8.23	9.39	10.86	13.68	17.34	21.60	25.99	28.87	31.53	34.81	37.16		
19	6.84	7.63	8.91	10.12	11.65	14.56	18.34	22.72	27.20	30.14	32.85	36.19	38.58		
20	7.43	8.26	9.59	10.85	12.44	15.45	19.34	23.83	28.41	31.41	34.17	37.57	40.00		
21	8.03	8.90	10.28	11.59	13.24	16.34	20.34	24.93	29.62	32.67	35.48	38.93	41.40		
22	8.64	9.54	10.98	12.34	14.04	17.24	21.34	26.04	30.81	33.92	36.78	40.29	42.80		
23	9.26	10.20	11.69	13.09	14.85	18.14	22.34	27.14	32.01	35.17	38.08	41.64	44.18		
24	9.89	10.86	12.40	13.85	15.66	19.04	23.34	28.24	33.20	36.42	39.36	42.98	45.56		
25	10.52	11.52	13.12	14.61	16.47	19.94	24.34	29.34	34.38	37.65	40.65	44.31	46.93		
26	11.16	12.20	13.84	15.38	17.29	20.84	25.34	30.43	35.56	38.89	41.92	45.64	48.29		
27	11.81	12.88	14.57	16.15	18.11	21.75	26.34	31.53	36.74	40.11	43.19	46.96	49.64		
28	12.46	13.56	15.31	16.93	18.94	22.66	27.34	32.62	37.92	41.34	44.46	48.28	50.99		
29	13.12	14.26	16.05	17.71	19.77	23.57	28.34	33.71	39.09	42.56	45.72	49.59	52.34		
30	13.79	14.95	16.79	18.49	20.60	24.48	29.34	34.80	40.26	43.77	46.98	50.89	53.67		
40	20.71	22.16	24.43	26.51	29.05	33.66	39.34	45.62	51.80	55.76	59.34	63.69	66.77		
50	27.99	29.71	32.36	34.76	37.69	42.94	49.33	56.33	63.17	67.50	71.42	76.15	79.49		
60	35.53	37.48	40.48	43.19	46.46	52.29	59.33	66.98	74.40	79.08	83.30	88.38	91.95		
70	43.28	45.44	48.76	51.74	55.33	61.70	69.33	77.58	85.53	90.53	95.02	100.42	104.22		
80	51.17	53.54	57.15	60.39	64.28	71.14	79.33	88.13	96.58	101.88	106.63	112.33	116.32		
90	59.20	61.75	65.65	69.13	73.29	80.62	89.33	98.64	107.56	113.14	118.14	124.12	128.30		
100	67.33	70.06	74.22	77.93	82.36	90.13	99.33	109.14	118.50	124.34	129.56	135.81	140.17		

Kritikus kétoldali U-értékek a Mann-Whitney próbához

n _b	n _a = 1				n _a = 2				n _a = 3									
	α =	0.1	0.05	0.02	0.01	n _b	α =	0.1	0.05	0.02	0.01	n _b	α =	0.1	0.05	0.02	0.01	
1																		
2						2												
3						3						3		9				
4						4						4		12				
5						5						5		14				
6						6		10				6		16				
7						7		12				7		17				
8						8		14				8		19		21	22	24
9						9		15	16			9		21	22	24		
10						10		17	18			10		23	25	26	26	27
11						11		19	20			11		26	27	29	30	30
12						12		21	22			12		28	30	32	32	33
13						13		22	23			13		31	32	34	34	35
14						14		24	25	26		14		33	35	37	37	38
15						15		25	27	28	26	15		35	37	40	41	41
16						16		27	29	30	30	16		38	40	42	42	43
17						17		29	31	32	32	17		40	42	45	45	46
18						18		31	32	34	34	18		42	45	47	47	49
19						19		32	34	36	36	19		45	50	50	50	52
20						20		34	36	37	38	20		47	50	53	53	54
21						21		36	38	39	39	21		49	52	55	55	57
22						22		37	39	41	41	22		52	55	58	58	60
23						23		39	41	43	43	23		54	57	60	60	62
24						24		41	43	45	45	24		56	60	63	63	65
25						25		42	45	47	48	25		59	62	66	66	68
26						26		44	47	49	50	26		61	65	68	68	70
27						27		46	48	51	52	27		63	67	71	71	73
28						28		47	50	52	53	28		66	70	74	74	76
29						29		49	52	54	55	29		68	72	76	76	79
30						30		51	54	56	57	30		71	74	79	79	81
31						31		53	55	58	59	31		73	77	81	81	84
32						32		54	57	60	61	32		75	79	84	84	87
33						33		56	59	62	63	33		77	82	87	87	89
34						34		58	61	64	65	34		82	87	92	92	95
35						35		59	63	65	67	35		84	89	94	94	97
36						36		61	64	67	69	36		87	92	97	97	100
37						37		63	66	69	71	37		89	94	100	100	103
38						38		66	70	73	75	38		91	97	102	102	105
39						39		68	71	75	76	39		94	99	105	105	108
40						40		69	73	77	78	40		96	102	107	107	111

A 6. táblázat folytatása

n _b	n _a = 7					n _a = 8					n _a = 9						
	α=	0.1	0.05	0.02	0.01	n _b	α=	0.1	0.05	0.02	0.01	n _b	α=	0.1	0.05	0.02	0.01
7		38	41	43	45												
8		43	46	49	50	8		49	51	55	57						
9		48	51	54	56	9		54	57	61	63	9		60	64	67	70
10		53	56	59	61	10		60	63	67	69	10		66	70	74	77
11		58	61	65	67	11		65	69	73	75	11		72	76	81	83
12		63	66	70	72	12		70	74	79	81	12		78	82	87	90
13		67	71	75	78	13		76	80	84	87	13		84	89	94	97
14		72	76	81	83	14		81	86	90	94	14		90	95	100	104
15		77	81	86	89	15		87	91	96	100	15		96	101	107	111
16		82	86	91	94	16		92	97	102	106	16		102	107	113	117
17		86	91	96	100	17		97	102	108	112	17		108	114	120	124
18		91	96	102	105	18		103	108	114	118	18		114	120	126	131
19		96	101	107	111	19		108	114	120	124	19		120	126	133	138
20		101	106	112	116	20		113	119	126	130	20		126	132	140	144
21		106	111	117	122	21		119	125	132	136	21		132	139	146	151
22		110	116	123	127	22		124	131	138	142	22		138	145	153	158
23		115	121	128	132	23		130	136	144	149	23		144	151	159	164
24		120	126	133	138	24		135	142	150	155	24		150	157	166	171
25		125	131	139	143	25		140	147	155	161	25		156	163	172	178
26		129	136	144	149	26		146	153	161	167	26		162	170	179	181
27		134	141	149	154	27		151	159	167	173	27		168	176	185	191
28		139	146	154	160	28		156	164	173	179	28		174	182	192	198
29		144	151	160	165	29		162	170	179	185	29		179	188	198	205
30		149	156	165	170	30		167	175	185	191	30		185	194	205	212
31		153	161	170	175	31		172	181	191	197	31		191	201	211	218
32		158	166	175	181	32		178	187	197	203	32		197	207	218	225
33		163	171	181	187	33		183	192	203	209	33		203	213	224	232
34		168	176	186	192	34		188	198	208	215	34		209	219	231	238
35		172	181	191	198	35		194	203	214	221	35		215	226	237	245
36		177	186	196	203	36		199	209	220	228	36		221	232	243	252
37		182	191	202	208	37		205	215	226	234	37		227	238	250	258
38		187	196	207	214	38		210	220	232	240	38		233	244	257	265
39		191	201	212	219	39		215	226	238	246	39		239	250	263	272
40		196	206	217	225	40		221	231	244	252	40		245	257	270	279

A 6. táblázat folytatása

n_b	$n_a = 13$					$n_a = 14$					$n_a = 15$						
	$\alpha =$	0.1	0.05	0.02	0.01	n_b	$\alpha =$	0.1	0.05	0.02	0.01	n_b	$\alpha =$	0.1	0.05	0.02	0.01
13		118	124	130	135												
14		126	132	139	144	14		135	141	149	154						
15		134	141	148	153	15		144	151	159	164	15		153	161	169	174
16		143	149	157	163	16		153	160	168	174	16		163	170	179	185
17		151	158	166	172	17		161	169	178	184	17		172	180	189	195
18		159	167	175	181	18		170	178	187	194	18		182	190	200	206
19		167	175	184	190	19		179	188	197	203	19		191	200	210	216
20		176	184	193	200	20		188	197	207	213	20		200	210	220	227
21		184	193	202	209	21		197	206	216	223	21		210	219	230	237
22		192	201	211	218	22		206	215	226	233	22		219	229	240	248
23		201	210	220	227	23		215	224	235	243	23		229	239	251	258
24		209	218	229	237	24		223	233	245	253	24		238	249	261	269
25		217	227	238	246	25		232	243	255	263	25		247	258	271	279
26		225	236	247	255	26		241	252	264	272	26		257	268	281	290
27		234	244	256	264	27		250	261	274	282	27		266	278	291	300
28		242	253	265	273	28		259	270	283	292	28		276	288	301	311
29		250	261	274	283	29		268	279	293	302	29		285	297	312	321
30		258	270	283	292	30		276	289	302	312	30		294	307	322	331
31		267	278	292	301	31		285	298	312	321	31		304	317	332	342
32		275	287	301	310	32		294	307	321	331	32		313	327	342	352
33		283	296	310	319	33		303	316	331	341	33		323	336	352	363
34		291	304	319	328	34		312	325	341	351	34		332	346	362	373
35		299	313	328	338	35		320	334	350	361	35		341	356	372	383
36		308	321	337	347	36		329	343	360	370	36		351	366	382	394
37		316	330	346	356	37		338	353	369	380	37		360	375	393	404
38		324	338	355	365	38		347	362	379	390	38		369	385	403	415
39		332	347	363	374	39		356	371	388	400	39		379	395	413	425
40		341	355	372	384	40		364	380	398	410	40		388	404	423	435

A 6. táblázat folytatása

n_b	$n_a = 16$					$n_a = 17$					$n_a = 18$								
	$\alpha=$	0.1	0.05	0.02	0.01	n_b	$\alpha=$	0.1	0.05	0.02	0.01	n_b	$\alpha=$	0.1	0.05	0.02	0.01		
	16		173	181	190	196													
17		183	191	201	207	17		193	202	212	219								
18		193	202	212	218	18		204	213	224	231	18		215	225	236	243		
19		203	212	222	230	19		214	224	235	242	19		226	236	248	255		
20		213	222	233	241	20		225	235	247	254	20		237	248	260	268		
21		223	233	244	252	21		236	246	258	266	21		248	259	272	280		
22		233	243	255	263	22		246	257	269	278	22		260	271	284	292		
23		243	253	266	274	23		257	268	281	289	23		271	282	296	305		
24		253	264	276	285	24		267	279	292	301	24		282	294	308	317		
25		263	274	287	296	25		278	290	303	313	25		293	305	320	329		
26		273	284	298	307	26		288	301	315	324	26		304	317	332	341		
27		283	295	309	318	27		299	312	326	336	27		315	328	344	354		
28		292	305	319	329	28		309	323	337	348	28		326	340	355	366		
29		302	315	333	340	29		320	333	349	359	29		337	351	367	378		
30		312	326	341	351	30		330	344	360	371	30		348	363	379	390		
31		322	336	352	362	31		341	355	371	382	31		359	374	391	403		
32		332	346	362	373	32		351	366	383	394	32		370	386	403	415		
33		342	357	373	384	33		362	377	394	406	33		382	397	415	427		
34		352	367	384	395	34		372	388	405	417	34		393	409	427	439		
35		362	377	395	406	35		383	399	417	429	35		404	420	439	451		
36		372	388	405	417	36		393	410	428	440	36		415	432	451	464		
37		382	398	416	428	37		404	420	439	452	37		426	443	463	476		
38		392	408	427	439	38		414	431	451	464	38		437	456	475	488		
39		402	418	437	450	39		425	442	462	475	39		448	466	486	500		
40		412	429	448	461	40		435	453	473	487	40		459	477	498	512		

A 6. táblázat folytatása

$n_a = 19$						$n_a = 20$					
n_b	$\alpha =$	0.1	0.05	0.02	0.01	n_b	$\alpha =$	0.1	0.05	0.02	0.01
19		238	248	260	268						
20		250	261	273	281	20		262	273	286	295
21		261	273	286	294	21		274	286	299	308
22		273	285	298	307	22		286	299	313	322
23		285	297	311	320	23		299	311	326	335
24		296	309	323	333	24		311	324	339	349
25		308	321	336	346	25		323	337	352	362
26		320	333	348	359	26		335	349	365	376
27		331	345	361	371	27		348	362	378	389
28		343	357	373	384	28		360	374	391	403
29		355	369	386	397	29		372	387	404	416
30		366	381	398	410	30		384	400	418	430
31		378	393	411	423	31		396	413	431	443
32		390	405	423	436	32		409	425	444	456
33		401	417	436	448	33		421	438	457	470
34		413	429	448	461	34		433	450	470	483
35		425	441	461	474	35		445	463	483	497
36		436	453	473	487	36		457	475	496	510
37		448	465	486	500	37		469	488	509	523
38		459	477	498	512	38		482	501	522	537
39		471	489	511	525	39			513	535	550
40		482	502	523	538	40			526	548	563

7. táblázat

Kritikus Q-értékek a Kruskal-Wallis próbához (k a minták összes száma)

k	Probability (α)					
	0.1	0.05	0.02	0.01	0.002	0.001
2	1.645	1.960	2.327	2.576	3.091	3.291
3	2.128	2.394	2.713	2.936	3.403	3.588
4	2.394	2.639	2.936	3.144	3.588	3.765
5	2.576	2.807	3.091	3.291	3.719	3.891
6	2.713	2.936	3.209	3.403	3.820	3.988
7	2.823	3.038	3.304	3.494	3.902	4.067
8	2.914	3.124	3.384	3.870	3.972	4.134
9	2.992	3.197	3.453	3.635	4.031	4.191
10	3.059	3.261	3.512	3.692	4.083	4.241
11	3.119	3.317	3.565	3.743	4.129	4.286
12	3.172	3.368	3.613	3.789	4.171	4.326
13	3.220	3.414	3.656	3.830	4.209	4.363
14	3.264	3.456	3.695	3.868	4.244	4.397
15	3.304	3.494	3.731	3.902	4.276	4.428
16	3.342	3.529	3.765	3.935	4.305	4.456
17	3.376	3.562	3.796	3.965	4.333	4.483
18	3.409	3.593	3.825	3.993	4.359	4.508
19	3.439	3.622	3.852	4.019	4.383	4.532
20	3.467	3.649	3.878	4.044	4.406	4.554
21	3.494	3.675	3.092	4.067	4.428	4.575
22	3.519	3.699	3.925	4.089	4.48	4.595
23	3.543	3.722	3.947	4.110	4.468	4.614
24	3.566	3.744	3.968	4.130	4.486	4.632
25	3.588	3.765	3.988	4.149	4.504	4.645

8. táblázat

**Kritikus kétoldalú S és T-értékek a páros előjelteszthez (S-értékek),
illetve a Wilcoxon előjelteszthez (T-érték)**

n	Valószínűség (α)							
	Kritikus S-értékek				Kritikus T-értékek			
	0.1	0.05	0.02	0.01	0.1	0.05	0.02	0.01
5	5				15			
6	6	6			19	21		
7	7	7	7		25	26	28	
8	7	8	8	8	31	33	35	36
9	8	9	9	9	37	40	42	44
10	9	9	10	10	45	47	50	52
11	9	10	10	11	53	56	59	61
12	10	10	11	11	61	65	69	71
13	10	11	12	12	70	74	79	82
14	11	12	12	13	80	84	90	93
15	12	12	13	13	90	95	101	105
16	12	13	14	14	101	107	113	117
17	13	13	14	15	112	119	126	130
18	13	14	15	15	124	131	139	144
19	14	15	15	16	137	144	153	158
20	15	15	16	17	150	158	167	173
21	15	16	17	17	164	173	182	189
22	16	17	17	18	178	188	198	205
23	16	17	18	19	193	203	214	222
24	17	18	19	19	209	219	231	239
25	18	18	19	20	225	236	249	257
26	18	119	20	20	241	253	267	276
27	19	20	20	21	259	271	286	295
28	19	20	21	22	276	290	305	315
29	20	21	22	22	295	309	325	335
30	20	21	22	23	314	328	345	356
31	21	22	23	24	333	349	366	378
32	22	23	24	24	353	369	388	400
33	22	23	24	25	374	391	410	423
34	23	24	25	25	395	413	433	447
35	23	24	25	26	417	435	457	471
36	24	25	26	27	439	458	481	495
37	24	25	27	27	462	482	505	521
38	25	26	27	28	485	506	530	547
39	26	27	28	28	509	531	556	573
40	26	27	28	29	534	556	582	600
45	29	30	31	32	664	692	723	744
50	32	33	34	35	8009	841	878	902

A nullhipotézist visszautasítjuk, ha az S vagy T megfigyelt értéke egyenlő vagy nagyobb, mint a kritikus érték.

Kritikus q-értékek ($\alpha=0.05$) a Tukey HSD és Student-Newman-Keuls (SNK) többszörös összehasonlítási próbához
(k az összehasonlítandó minták összes száma. v_{res} a maradék szabadságfok)

v_{res}	k													
	2	3	4	5	6	7	8	9	10	11	12	13	14	
1	17.97	26.98	32.82	37.08	40.41	43.12	45.40	47.36	49.07	50.59	51.96	53.20	54.33	
2	6.085	8.331	9.798	10.880	11.740	12.440	13.030	13.540	13.990	14.390	14.750	15.080	15.380	
3	4.501	5.910	6.825	7.502	8.037	8.478	8.853	9.177	9.462	9.717	9.946	10.150	10.350	
4	3.927	5.040	5.757	6.287	6.707	7.053	7.347	7.602	7.862	8.027	8.208	8.373	8.525	
5	3.635	4.602	5.218	5.673	6.033	6.330	6.582	6.802	6.995	7.168	7.327	7.466	7.596	
6	3.461	4.339	4.896	5.305	5.628	5.895	6.122	6.319	6.493	6.649	6.789	6.917	7.034	
7	3.344	4.165	4.681	5.060	5.359	5.606	5.815	5.998	6.158	6.303	6.431	6.550	6.658	
8	3.261	4.041	4.529	4.886	5.167	5.399	5.597	5.767	5.918	6.054	6.175	6.287	6.389	
9	3.199	3.949	4.415	4.756	5.024	5.244	5.432	5.595	5.739	5.867	5.983	6.089	6.186	
10	3.151	3.877	4.327	4.654	4.912	5.124	5.305	5.461	5.599	5.722	5.833	5.935	6.028	
11	3.113	3.820	4.256	4.574	4.823	5.028	5.202	5.353	5.487	5.605	5.713	5.811	5.901	
12	3.082	3.773	4.199	4.508	4.751	4.950	5.119	5.265	5.395	5.511	5.615	5.710	5.798	
13	3.055	3.735	4.151	4.453	4.690	4.885	5.049	5.192	5.318	5.431	5.533	5.625	5.711	
14	3.033	3.702	4.111	4.407	4.639	4.829	4.990	5.131	5.254	5.364	5.463	5.54	5.637	
15	3.014	3.674	4.076	4.367	4.595	4.782	4.940	5.077	5.198	5.306	5.404	5.493	5.574	
16	2.998	3.649	4.046	4.333	4.557	4.741	4.897	5.031	5.150	5.256	5.352	5.439	5.520	
17	2.984	3.628	4.200	4.303	4.524	4.705	4.858	4.991	5.108	5.212	5.307	5.392	5.471	
18	2.971	3.609	3.997	4.277	4.495	4.673	4.824	4.956	5.071	5.174	5.267	5.352	5.429	
19	2.960	3.593	3.770	4.253	4.469	4.645	4.794	4.924	5.038	5.140	5.231	5.315	5.391	
20	2.9650	3.576	3.958	4.232	4.445	4.620	4.768	4.896	5.008	5.108	5.199	5.282	5.357	
24	2.919	3.532	3.901	4.166	4.373	4.541	4.684	4.807	4.915	5.012	5.099	5.179	5.251	
30	2.888	3.486	3.845	4.302	4.464	4.646	4.602	4.720	4.824	4.917	5.001	5.077	5.147	
40	2.858	3.442	3.791	4.039	4.232	4.389	4.521	4.635	4.735	4.824	4.904	5.077	5.044	
60	2.829	3.390	3.737	3.977	4.163	4.314	4.441	4.550	4.646	4.732	4.808	5.878	4.942	
120	2.800	3.365	3.685	3.917	4.096	4.241	4.363	4.468	4.560	4.641	4.714	5.781	4.842	
∞	2.772	3.314	3.633	3.858	4.030	4.170	4.286	4.387	4.474	4.552	5.622	5.685	4.743	

A 9. táblázat folytatása

V _{res}	k													
	15	16	17	18	19	20	22	24	26	28	30	34	40	50
1	55.360	55.320	57.220	58.070	58.830	59.560	60.910	62.120	63.220	64.230	65.150	66.810	68.920	71.73
2	15.620	15.910	16.140	16.370	16.570	16.770	17.130	17.450	17.750	18.020	18.270	18.720	19.28	20.05
3	15.530	10.690	10.840	10.980	11.110	11.240	11.470	11.680	11.870	12.050	12.210	12.500	12.87	13.36
4	8.664	8.794	8.914	9.028	9.134	9.233	9.418	9.584	9.736	9.875	10.000	10.230	10.53	10.93
5	7.717	7.828	7.932	8.030	8.122	8.208	8.368	8.512	8.643	8.764	8.875	9.075	9.330	9.674
6	7.143	7.244	7.338	7.426	7.508	7.587	7.730	7.861	7.979	8.088	8.189	8.370	8.601	8.913
7	6.759	6.852	6.939	7.020	7.097	7.170	7.303	7.423	7.533	7.634	7.728	7.895	8.110	8.400
8	6.483	6.571	6.653	6.729	6.802	6.870	6.995	7.109	7.212	7.307	7.395	7.554	7.756	8.029
9	6.276	6.359	6.437	6.510	6.579	6.644	6.673	6.871	6.970	7.061	7.145	7.295	7.488	7.749
10	6.114	6.194	6.269	6.339	6.405	6.467	6.582	6.686	6.781	6.868	6.948	7.093	7.279	7.529
11	5.984	6.062	6.134	6.202	6.265	6.326	6.436	6.536	6.628	6.712	6.790	6.930	7.110	7.352
12	5.878	5.953	6.023	6.089	6.151	6.209	6.317	6.414	6.503	6.585	6.660	6.796	6.970	7.205
13	5.789	5.862	5.931	5.995	6.055	6.112	6.217	6.312	6.398	6.478	6.551	6.684	6.854	7.083
14	5.714	5.786	5.852	5.915	5.974	6.029	6.132	6.224	6.309	6.387	6.459	6.588	6.754	6.979
15	5.649	5.720	5.785	5.846	5.904	5.958	6.059	6.149	6.233	6.309	6.379	6.506	6.669	6.888
16	5.593	5.662	5.727	5.786	5.843	5.897	5.995	6.084	6.166	6.241	6.310	6.434	6.594	6.810
17	5.544	5.612	5.675	5.734	5.790	5.842	5.940	6.027	6.107	6.181	6.249	6.372	6.529	6.741
18	5.501	5.568	5.630	5.688	5.743	5.794	5.890	5.977	6.055	6.128	6.195	6.316	6.471	6.680
19	5.462	5.528	5.589	5.647	5.701	5.752	5.846	5.932	6.009	6.081	6.147	6.267	6.419	6.626
20	5.427	5.493	5.553	5.610	5.663	5.714	5.807	5.891	5.968	6.039	6.104	6.222	6.373	6.576
24	5.319	5.381	5.439	5.494	5.545	5.594	5.683	5.764	5.838	5.906	5.968	6.081	6.226	6.421
30	5.211	5.271	5.327	5.379	5.429	5.475	5.561	5.638	5.709	5.774	5.833	5.941	6.080	6.267
40	5.106	5.163	5.216	5.266	5.313	5.358	5.439	5.513	5.581	5.642	5.700	5.803	5.934	6.112
60	5.001	5.056	5.107	5.154	5.199	5.241	5.319	5.389	5.453	5.512	5.566	5.664	5.789	5.958
120	5.898	4.950	4.998	5.044	5.086	5.126	5.200	5.266	5.327	5.382	5.434	5.526	5.644	5.802
∞	5.796	4.845	4.891	4.934	4.974	5.012	5.081	5.144	5.201	5.253	5.301	5.388	5.498	5.646

Kritikus kétoldali q' értékek ($\alpha=0.05$) a Dunnett kétoldali többszörös összehasonlítási próbához
 (k az összehasonlítandó minták összes száma, V_{res} a maradék szabadságfok)

V_{res}	k																			
	2	3	4	5	6	7	8	9	10	11	12	13	16	21						
5	2.57	3.03	3.29	3.48	3.62	3.73	3.82	3.90	3.97	4.03	4.09	4.14	4.25	4.42						
6	2.45	2.86	3.10	3.25	3.39	3.49	3.57	3.64	3.71	3.76	3.81	3.86	3.97	4.11						
7	2.36	2.75	2.97	3.12	3.24	3.33	3.41	3.47	3.53	3.58	3.63	3.67	3.78	3.91						
8	2.31	2.67	2.88	3.02	3.13	3.22	3.20	3.25	3.41	3.46	3.50	3.54	3.64	3.76						
9	2.26	2.61	2.81	2.95	3.05	3.14	3.29	3.26	3.32	3.36	3.40	3.44	3.53	3.65						
10	2.23	2.57	2.76	2.89	2.99	3.07	3.14	3.19	3.24	3.29	3.33	3.36	3.45	3.57						
11	2.20	2.53	2.72	2.84	2.94	3.02	3.08	3.14	3.19	3.23	3.27	3.30	3.39	3.50						
12	2.18	2.50	2.68	2.81	2.90	2.98	3.04	3.09	3.14	3.18	3.22	3.25	3.34	3.45						
13	2.16	2.48	2.65	2.78	2.87	2.94	3.00	3.06	3.10	3.14	3.18	3.21	3.29	3.40						
14	2.14	2.46	2.63	2.75	2.84	2.91	2.97	3.02	3.07	3.11	3.14	3.18	3.26	3.36						
15	2.13	2.44	2.61	2.73	2.82	2.89	2.95	3.00	3.04	3.08	3.12	3.15	3.23	3.33						
16	2.12	2.42	2.59	2.71	2.80	2.87	2.92	2.97	3.02	3.06	3.09	3.12	3.20	3.30						
17	2.11	2.41	2.58	2.69	2.78	2.85	2.90	2.95	3.00	3.03	3.07	3.10	3.18	3.27						
18	2.10	2.40	2.56	2.68	2.76	2.83	2.89	2.94	2.98	3.01	3.05	3.08	3.16	3.25						
19	2.09	2.39	2.55	2.66	2.75	2.81	2.87	2.92	2.96	3.00	3.03	3.06	3.14	3.23						
20	2.09	2.38	2.54	2.65	2.73	2.80	2.86	2.90	2.95	2.98	3.02	3.05	3.12	3.22						
24	2.06	2.35	2.51	2.61	2.70	2.76	2.81	2.86	2.90	2.94	2.97	3.00	3.07	3.16						
30	2.04	2.32	2.47	2.58	2.66	2.72	2.77	2.82	2.86	2.89	2.92	2.95	3.02	3.11						
40	2.02	2.29	2.44	2.54	2.62	2.68	2.73	2.77	2.81	2.85	2.87	2.90	2.97	3.06						
60	2.00	2.27	2.41	2.51	2.58	2.64	2.69	2.73	2.77	2.80	2.83	2.86	2.92	3.00						
120	1.98	2.24	2.38	2.47	2.55	2.60	2.65	2.69	2.73	2.76	2.79	2.81	2.87	2.95						
∞	1.96	2.21	2.35	2.44	2.51	2.57	2.61	2.65	2.69	2.72	2.74	2.77	2.83	2.91						

Spearman rangkorrelációs koefficiens (r_s) kritikus értékei

<i>n</i>	Valószínűség (α)			
	0.10	0.05	0.02	0.01
5	0.900			
6	0.829		0.943	
7	0.714	0.886	0.843	0.929
8	0.643	0.738	0.833	0.881
9	0.600	0.683	0.783	0.833
10	0.564	0.648	0.745	0.794
11	0.523	0.623	0.736	0.818
12	0.497	0.591	0.703	0.780
13	0.475	0.586	0.673	0.745
14	0.457	0.545	0.648	0.716
15	0.441	0.525	0.623	0.689
16	0.425	0.507	0.601	0.666
17	0.412	0.490	0.582	0.645
18	0.399	0.476	0.564	0.625
19	0.388	0.462	0.549	0.608
20	0.377	0.450	0.534	0.591
21	0.368	0.438	0.521	0.576
22	0.359	0.428	0.508	0.562
23	0.351	0.418	0.496	0.549
24	0.343	0.409	0.485	0.537
25	0.336	0.400	0.475	0.525
26	0.329	0.392	0.465	0.515
27	0.323	0.384	0.456	0.505
28	0.317	0.377	0.448	0.496
29	0.311	0.370	0.440	0.487
30	0.305	0.364	0.432	0.478
31	0.300	0.358	0.424	0.470
32	0.295	0.352	0.418	0.462
33	0.291	0.346	0.409	0.454
34	0.287	0.341	0.402	0.446
35	0.283	0.336	0.396	0.439
36	0.279	0.331	0.390	0.432
37	0.275	0.326	0.384	0.425
38	0.271	0.322	0.378	0.418
39	0.267	0.318	0.373	0.411
40	0.264	0.314	0.368	0.405
45	0.248	0.294	0.347	0.382
50	0.235	0.279	0.329	0.363

12. táblázat

Kritikus H-értékek a Kruskal-Wallis többszörös összehasonlítási próbához

Minta mérete					α		Minta mérete					α	
n1	n2	n3	n4	n5	0.05	0.01	n1	n2	n3	n4	n5	0.05	0.01
2	2	2					2	2	1	1			
3	2	1					2	2	2	1		5.679	
3	2	2			4.714		2	2	2	2		6.167	6.667
3	3	1			5.143		3	1	1	1			
3	3	2			5.361		3	2	1	1			
3	3	3			5.600	7.200	3	2	2	1		5.833	
4	2	1					3	2	2	2		6.333	7.133
4	2	2			5.333		3	3	1	1		6.333	
4	3	1			5.208		3	3	2	1		6.244	7.200
4	3	2			5.444	6.444	3	3	2	2		6.527	7.635
4	3	3			5.791	6.745	3	3	3	1		6.600	7.400
4	4	1			4.967	6.667	3	3	3	2		6.726	8.105
4	4	2			5.455	7.036	3	3	3	3		7.000	8.538
4	4	3			5.598	7.144	4	1	1	1			
4	4	4			5.692	7.654	4	2	1	1		5.833	
5	2	1			5.000		4	2	2	1		6.133	7.000
5	2	2			5.160	6.533	4	2	2	2		6.545	7.391
5	3	1			4.960		4	3	1	1		6.178	7.067
5	3	2			5.251	6.909	4	3	2	1		6.309	7.455
5	3	3			5.648	7.079	4	3	2	2		6.621	7.871
5	4	1			4.985	6.955	4	3	3	1		6.545	7.758
5	4	2			5.273	7.205	4	3	3	2		6.795	8.333
5	4	3			5.656	7.445	4	3	3	3		6.964	8.659
5	4	4			5.657	7.760	4	4	1	1		5.945	7.909
5	5	1			5.127	7.309	4	4	2	1		6.386	7.866
5	5	2			5.338	7.338	4	4	2	2		6.731	8.346

A 12. táblázat folytatása

Minta mérete					α		Minta mérete					α	
n1	n2	n3	n4	n5	0.05	0.01	n1	n2	n3	n4	n5	0.05	0.01
5	5	5			5.705	7.578	4	4	3	1		6.635	8.231
5	5	4			5.666	7.823	4	4	3	2		6.874	8.621
5	5	5			5.780	8.000	4	4	3	3		7.038	8.876
6	2	1			4.822		4	4	4	1		6.725	8.588
6	2	2			5.345	6.982	4	4	4	2		6.957	8.871
6	3	1			4.855		4	4	4	3		7.142	9.075
6	3	2			5.348	6.970	4	4	4	4		7.235	9.287
6	3	3			5.615	7.410	2	1	1	1	1		
6	4	1			4.947	7.106	2	2	1	1	1		
6	4	2			5.340	7.340	2	2	2	1	1	6.750	
6	4	3			5.610	7.500	2	2	2	2	1	7.133	7.533
6	4	4			5.681	7.795	2	2	2	2	2	7.418	8.291
6	5	1			4.990	7.182	3	1	1	1	1		
6	5	2			5.338	7.376	3	2	1	1	1	6.583	
6	5	3			5.602	7.590	3	2	2	1	1	6.800	7.600
6	5	4			5.661	7.936	3	2	2	2	1	7.309	8.127
6	5	5			5.729	8.028	3	2	2	2	2	7.682	8.682
6	6	1			4.945	7.121	3	3	1	1	1	7.111	
6	6	2			5.410	7.457	3	3	2	1	1	7.200	8.073
6	6	3			5.625	7.725	3	3	2	2	1	7.591	8.576
6	6	4			5.724	8.000	3	3	2	2	2	7.910	9.115
6	6	5			5.765	6.124	3	3	3	1	1	7.576	8.424
6	6	6			5.801	8.222	3	3	3	2	1	7.769	9.051
7	7	7			5.819	8.378	3	3	3	2	2	8.044	9.505
8	8	8			5.805	8.465	3	3	3	3	1	8.000	9.451
							3	3	3	3	2	8.200	9.876
							3	3	3	3	3	8.333	10.200

13. táblázat

Kritikus Friedman χ^2 értékek a Friedman nem-paraméteres többszörös összehasonlítási próbához
(„a” a kezelések száma, „b” a blokkok száma)

a	b	0.05	0.01
3	3	6.000	
3	4	6.500	8.000
3	5	6.400	8.400
3	6	7.000	9.000
3	7	7.143	8.857
3	8	6.250	9.000
3	9	6.222	9.556
3	10	6.200	9.600
3	11	6.545	9.455
3	12	6.167	9.500
3	13	6.000	9.385
3	14	6.143	9.000
3	15	6.400	8.933
4	2	6.000	
4	3	7.400	9.000
4	4	7.800	9.600
4	5	7.800	9.960
4	6	7.800	10.200
4	7	7.800	10.371
4	8	7.650	10.350
4	9	7.800	10.867
4	10	7.800	10.800
4	11	7.909	11.073
4	12	7.900	11.100
4	13	7.985	11.123
4	14	7.886	11.143
4	15	8.040	11.240
5	2	7.600	8.000
5	3	8.533	10.133
5	4	8.800	11.200
5	5	8.960	11.680
5	6	9.067	11.867
5	7	9.413	12.114
5	8	9.300	12.300
5	9	9.244	12.444
5	10	9.280	12.480
6	2	9.143	9.714
6	3	9.857	11.762
6	4	10.266	12.714
6	5	10.486	13.229
6	6	10.571	13.619
6	7	10.674	13.857
6	8	10.714	14.000
6	9	10.778	14.143
6	10	10.800	14.299

Szignifikáns Studentizált terjedelmek 5%-os és 1%-os szintjei a Duncan többszörös terjedelem próbához

Hiba d.f.	Szign. szint	p = átlagok száma a vizsgált tartományokra																	
		2	3	4	5	6	7	8	9	10	12	14	16	18	20				
1	0.05	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0				
	0.01	90.0	90.0	90.0	90.0	90.0	90.0	90.0	90.0	90.0	90.0	90.0	90.0	90.0	90.0				
2	0.05	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09				
	0.01	14.0	14.0	14.0	14.0	14.0	14.0	14.0	14.0	14.0	14.0	14.0	14.0	14.0	14.0				
3	0.05	4.50	4.50	4.50	4.50	4.50	4.50	4.50	4.50	4.50	4.50	4.50	4.50	4.50	4.50				
	0.01	8.26	8.50	8.60	8.70	8.80	8.90	8.90	9.00	9.00	9.00	9.10	9.20	9.30	9.30				
4	0.05	3.93	4.01	4.02	4.02	4.02	4.02	4.02	4.02	4.02	4.02	4.02	4.02	4.02	4.02				
	0.01	6.51	6.80	6.90	7.00	7.10	7.10	7.20	7.20	7.30	7.30	7.40	7.40	7.50	7.50				
5	0.05	3.64	3.74	3.79	3.83	3.83	3.83	3.83	3.83	3.83	3.83	3.83	3.83	3.83	3.83				
	0.01	5.70	5.96	6.11	6.18	6.26	6.33	6.40	6.44	6.50	6.60	6.60	6.70	6.70	6.80				
6	0.05	3.46	3.58	3.64	3.68	3.68	3.68	3.68	3.68	3.68	3.68	3.68	3.68	3.68	3.68				
	0.01	5.24	5.51	5.65	5.73	5.81	5.88	5.95	6.00	6.00	6.10	6.20	6.20	6.30	6.30				
7	0.05	3.35	3.47	3.54	3.58	3.60	3.61	3.61	3.61	3.61	3.61	3.61	3.61	3.61	3.61				
	0.01	4.95	5.22	5.37	5.45	5.53	5.61	5.69	5.73	5.80	5.80	5.90	5.90	6.00	6.00				
8	0.05	3.26	3.39	3.47	3.52	3.55	3.56	3.56	3.56	3.56	3.56	3.56	3.56	3.56	3.56				
	0.01	4.74	5.00	5.14	5.23	5.32	5.40	6.47	5.51	5.50	5.60	5.70	5.70	5.80	5.80				
9	0.05	3.20	3.34	3.41	3.47	3.50	3.52	3.52	3.52	3.52	3.52	3.52	3.52	3.52	3.52				
	0.01	4.60	4.86	4.99	5.08	5.17	5.25	5.32	5.36	5.40	5.50	5.60	5.60	5.70	5.70				
10	0.05	3.15	3.30	3.37	3.43	3.46	3.47	3.47	3.47	3.47	3.47	3.47	3.47	3.47	3.47				
	0.01	4.48	4.73	4.88	4.99	5.06	5.13	5.20	5.24	5.28	5.36	5.42	5.48	5.54	5.55				
11	0.05	3.11	3.27	3.35	3.39	3.43	3.44	3.45	3.46	3.46	3.46	3.46	3.46	3.47	3.48				
	0.01	4.39	4.63	4.77	4.86	4.94	5.01	5.06	5.12	5.15	5.24	5.28	5.34	5.38	5.39				
12	0.05	3.08	3.23	3.33	3.36	3.40	3.42	3.44	3.44	3.45	3.46	3.46	3.46	3.47	3.48				
	0.01	4.32	4.55	4.68	4.76	4.81	4.92	4.96	5.02	5.07	5.13	5.17	5.22	5.24	5.26				
13	0.05	3.06	3.21	3.30	3.35	3.38	3.41	3.42	3.44	3.45	3.45	3.46	3.46	3.47	3.47				
	0.01	4.26	4.48	4.62	4.69	4.74	4.84	4.88	4.94	4.98	5.04	5.08	5.13	5.14	5.15				
14	0.05	3.03	3.18	3.27	3.33	3.37	3.39	3.41	3.42	3.44	3.44	3.45	3.46	3.47	3.47				
	0.01	4.21	4.42	4.55	4.63	4.70	4.78	4.83	3.87	3.87	3.44	3.45	3.46	3.47	3.47				
15	0.05	3.01	3.16	3.25	3.31	3.36	3.38	3.40	3.42	3.43	3.44	3.44	3.45	3.47	3.47				
	0.01	4.17	4.37	4.50	4.58	4.64	4.72	4.77	4.81	4.84	4.90	4.94	4.97	4.99	5.00				

A 14. táblázat folytatása

Hiba d.f.	Szign. szint	p = átlagok száma a vizsgált tartományokra																	
		2	3	4	5	6	7	8	9	10	12	14	16	18	20				
16	0.05 0.01	3.00 4.13	3.15 4.32	3.23 4.45	3.30 4.54	3.34 4.60	3.37 4.67	3.39 4.72	3.41 4.76	3.43 4.79	3.44 4.84	3.45 4.88	3.46 4.91	3.47 4.93	3.47 4.94				
17	0.05 0.01	2.98 4.10	3.13 4.30	3.22 4.41	3.28 4.50	3.33 4.56	3.36 4.63	3.38 4.68	3.40 4.72	3.42 4.75	3.44 4.80	3.45 4.83	3.46 4.86	3.47 4.88	3.47 4.89				
18	0.05 0.01	2.97 4.07	3.12 4.27	3.21 4.38	3.27 4.46	3.32 4.53	3.35 4.59	3.37 4.64	3.39 4.68	3.41 4.71	3.43 4.76	3.45 4.79	3.48 4.82	3.47 4.84	3.47 4.85				
19	0.05 0.01	2.96 4.05	3.11 4.24	3.19 4.35	3.26 4.43	3.31 4.50	3.35 4.56	3.37 4.61	3.39 4.64	3.41 4.67	3.44 4.72	3.44 4.76	3.46 4.79	3.47 4.81	3.47 4.82				
20	0.05 0.01	2.95 4.02	3.10 4.22	3.18 4.33	3.25 4.40	3.30 4.47	3.34 4.53	3.36 4.58	3.38 4.61	3.40 4.65	3.43 4.69	3.44 4.73	3.46 4.76	3.46 4.78	3.47 4.79				
22	0.05 0.01	2.93 3.99	3.08 4.17	3.17 4.28	3.24 4.36	3.29 4.42	3.32 4.48	3.35 4.53	3.37 4.57	3.39 4.60	3.42 4.65	3.44 4.68	3.45 4.71	3.46 4.74	3.47 4.75				
24	0.05 0.01	2.92 3.96	3.07 4.14	3.15 4.24	3.22 4.33	3.28 4.39	3.31 4.44	3.34 4.49	3.37 4.53	3.38 4.57	3.41 4.62	3.44 4.64	3.45 4.67	3.46 4.70	3.47 4.72				
26	0.05 0.01	2.91 3.93	3.06 4.11	3.14 4.21	3.21 4.30	3.27 4.36	3.30 4.41	3.34 4.46	3.36 4.50	3.38 4.53	3.41 4.58	3.43 4.62	3.45 4.65	3.46 4.67	3.47 4.69				
28	0.05 0.01	2.90 3.91	3.04 4.08	3.13 4.18	3.26 4.28	3.26 4.34	3.30 4.39	3.33 4.43	3.35 4.47	3.37 4.51	3.40 4.56	3.43 4.60	3.45 4.62	3.46 4.65	3.47 4.67				
30	0.05 0.01	2.89 3.89	3.04 4.06	3.12 4.16	3.20 4.22	3.25 4.32	3.29 4.36	3.32 4.41	3.35 4.45	3.37 4.48	3.40 4.54	3.43 4.58	3.44 4.61	3.46 4.63	3.47 4.65				
40	0.05 0.01	2.86 3.82	3.01 3.99	3.10 4.10	3.17 4.17	3.22 4.21	3.27 4.30	3.30 4.34	3.33 4.37	3.35 4.41	3.39 4.46	3.42 4.51	3.44 4.54	3.46 4.57	3.47 4.59				
60	0.05 0.01	2.83 3.76	2.98 3.92	3.08 4.03	3.14 4.12	3.20 4.17	3.24 4.23	3.28 4.27	3.31 4.31	3.33 4.34	3.37 4.39	3.40 4.44	3.43 4.47	3.45 4.50	3.47 4.53				
100	0.05 0.01	2.80 3.71	2.95 3.86	3.05 3.98	3.12 4.06	3.18 4.11	3.22 4.17	3.26 4.21	3.29 4.25	3.32 4.29	3.36 4.35	3.40 4.38	3.42 4.42	3.45 4.45	3.47 4.48				
∞	0.05 0.01	2.77 3.64	2.92 3.80	3.02 3.90	3.09 3.98	3.15 4.04	3.19 4.09	3.23 4.14	3.26 4.17	3.29 4.20	3.34 4.26	3.38 4.31	3.41 4.34	3.44 4.38	3.47 4.41				

Tárgymutató

- adatok 76
 - statisztikai analízis célja 14-15
- adattípusok 75-76
- adatelemzés 79
- alparcella 59, 266
- alternatív hipotézis 34
- AMMI modell 450-452, 461, 467
- ANOVA (varianciaanalízis) feltételei 126-134
- aszimptotikus függvények 523-528
 - monomolekuláris 523
 - logisztikus 523-525
 - Gompertz 525
 - Richards 525-528

- Bartlett próba 338-339, 382-383
- binomiális eloszlás 27
- biometria 3
- blokk-képzés 17-18, 41-43

- χ^2 (χ^2)-eloszlás 28
- Cserháti S. 5
- csoportosított blokk-elrendezés 53-54, 190-204
- csúcosság 100

- determinációs koefficiens 493, 505
- diszkrét adatok 76
- diszkriminancia analízis 456-457
- Duncan DMR próba 145-151
- Dunnett próba 141-142

- egymintás t-próba 112-114
- egytenyezős kísérletek 46-57, 152-230
 - teljesen véletlen elrendezés (CRD) 47-48, 152-171
 - egyenlő ismétlésszám 157-165
 - nem egyenlő ismétlésszám 165-171
 - véletlen blokkelrendezés (RCB) 48-49, 172-179
 - latinnégyzet elrendezés 49-53, 180-190
 - csoportosított blokk-elrendezés 53-54, 190-204
 - latintégla elrendezés 56-57, 205-221
 - kiegyensúlyozott 206-217
 - részben kiegyensúlyozott 217-221
- elrendezés, kísérleti 46-74
- eltérésnégyzet összeg (SS) 23, 130
- exponenciális függvény 515, 518-522

- faktoriális kísérlet 24, 57, 326, 381
- F-eloszlás 28
- Fisher, R.A. 126
- fő hatások és interakciók 246-248
- főkomponensanalízis (PCA) 452, 455, 461, 467
- főparcella 59, 262-267
- F-próba 23, 121, 127, 130
- függvény, aszimptotikus 523-529
 - exponenciális 515, 518-522
 - hiperbola 516
 - lineáris plató 518-522
 - másodfokú 514, 518-522
 - négyzetgyök 518-522
 - polinomiális 514

- GenStat
 - ANOVA 163, 170, 177, 186, 199, 216, 218, 227, 243, 259, 276, 290, 304, 322, 330, 341, 355, 371, 393, 408, 421, 440
 - kísérletek tervezése 82-86
 - leíró statisztika 102
 - lineáris regresszió 511-513
 - nem-lineáris modellek 529
 - nem-paraméteres próbák 471, 476, 478, 482

- Student t-próba 123-125
többszörös regresszió analízis
538-541
grafikus ábrázolás 102-108
- Harnos Zs. 7**
hatás
additív 23
kezelés 21
kölcson- 21, 246
háromtényezős kísérletek 301-334
teljesen véletlen elrendezés 303-306
kétszeresen osztott parcellás 307-325
2ⁿ faktoriális 326-334
helyzet-mutatók 87-90
számtani átlag 87
medián 88
módusz 88
kvartilisek 89
hiba, kísérleti 16
elsőfajú hiba 35
hipotézis 33
null- 34-35
alternatív 34
hipotézisvizsgálat 34-35
hisztogram 105
- időbeni mérések analízise 335-346
inkomplett blokk elrendezés 55-57
interakció 10, 21, 231, 537
ismétlés 17, 38-41
- kezelés 21
- hatás 21
kiválasztása 37
kontroll 21
standard 21
struktúra 24, 222-230
kétmintás t-próba 114-118
kéttényezős kísérletek 231-300
teljesen véletlen (CRD) 238-245
véletlen blokk (RCB) 252-261
osztott parcellás (split-plot) 262-279
sávós (strip-plot) 280-292
hierarchikus (nested) 293-297
ismétlés nélküli 298-301
- kísérlet 16
kísérleti egység 18
kísérleti elrendezés 46-74
égytényezős 24, 46-57
faktoriális, 24, 57-69
kéttényezős 57-61
háromtényezős 61-69
kísérletsorozat 69-71
tartamkísérlet 72-73
üzemi (demonstrációs) 73-74
kísérleti hiba 16
kísérletek tervezése 29-45
számítógépes 79-86
kísérletsorozat 25, 69-71, 347-400
évek analízise 350-357
helyek analízise 358-362
technológiai kísérlet 363-380
2ⁿ faktoriális 381-386
ültetvények 387-400
koeficiens
determinációs 493, 505
korrelációs 488-493
regressziós 500-506
path- 541
többszörös determinációs 540, 542
variációs 94
kollinearitás 536-537
konfidencia határok 98, 506
korrelációs számítás 487-497
Pearson-féle 488-494
Spearman rangkorreláció 495-497
kovariánsok 44-45
kölcsonhatás 21, 246-248
közepes négyzetes eltérés
(MS) 23, 130
középtérték hibája 94-97
kumulatív elemzés 448-450, 458-459,
463-464
különbség standard hibája (SED) 96
kvartilisek 89
- latin négyzet elrendezés 49-53, 180-190
többszörös 51-53
latintégla elrendezés 56-57, 205-221
legkisebb négyzetek módszere
499-502

- legkisebb szignifikáns különbség (SzD) 137
- legkisebb szignifikáns növekedés (LSI) 427
- leíró statisztika 87-110
GenStat programmal 100-103
- lineáris regresszió 498-513
- logisztikus függvény 517, 523-525
- másodfajú hiba** 35
- medián 88
- minta 31-33
- módusz 88
- nem-lineáris modellek** 514-529
- nem-paraméteres próbák** 468-484
előjel teszt 472-474
Friedman teszt 480-482
Kruskal-Wallis teszt 476-479
Mann-Whitney U teszt 468-472
Spearman rangkorreláció 495-497
többszörös összehasonlítás 479-480
Wilcoxon 474-476
- normális eloszlás 26
- növekedési függvények 523-529
sebesség 523-528
gyorsulás 527-528
- nullhipotézis 34-35
- osztott parcellás** 59-60, 262-279, 335-336
- összefüggés-vizsgálat** 485-486
- összehasonlítás kezeléscsoportok között** 135-151
legkisebb szignifikáns különbség 137-138
legkisebb szignifikáns növekedés 427
Duncan próba 145-151
Dunnett próba 141-142
Student-Newman-Keuls próba 139-141
Tukey próba 138-139
számítógépes értékelés 143-145, 151
- paraméteres próbák** 111-125
egymintás t-próba 112-114
kétmintás t-próba 114-118
- páros t-próba 118-120
F-próba 120-122
számítógépes értékelés 123-125
parcella 19
mérete és alakja 19
bruttó és nettó 20
szegélyhatás 20
parciális regressziós koeficiens 531
pathanalízis 544-549
pathdiagram 544-545
pathkoeficiens 547
polinomiális függvény 514, 518-522
populáció 31-33
próba, ereje 24
F-próba 23
t-próba 111-125
 χ^2 -próba 480-482
- randomizáció 17, 43-44
rangkorreláció 495-497
reakció struktúra 22
regresszió analízis 485-550
lineáris 498-513
nem-lineáris 514-529
többszörös 530-550
- sávok elrendezése 60-61, 280-292
stabilitásanalízis 450-452
standard kezelés 21
statisztikai analízis célja 14-15
statisztikai jellemző eloszlása 27-28
F-eloszlás 28
Student t- eloszlás 27
z vagy normális eloszlás 27
 χ^2 (khi-négyzet)-eloszlás 28
Student-Newman-Keuls próba 139-141
Sváb J. 5-6
- szabadságfok 22, 130
szigmoid görbe 523
szignifikanciaszint 24
szignifikáns differencia 24
kiszámítása 161-163, 168-170, 176, 186, 198, 215, 242, 257-258, 272-274, 288-289, 319-321

- szóródás mutatók 90-98
eltérésnégyzet-összeg 92
interkvartilis terjedelelem 91
terjedelelem 91
szórás 22, 93
szórásnégyzet (variancia) 22, 93
variációs koefficiens 94
- tartamkísérlet 25, 72, 442-467
értékelési módszerek 444-453
martonvásári kísérletek 453-467
- táblázatok 109-110
statisztikai (F, H, q, Q, r, r_s , t, T, S, U, χ^2) 558-583
- teljesen véletlen elrendezés (CRD) 47-48, 152-171
- termesztési rendszerek (FSR) 427-444
- többszörös korrelációs koefficiens 542
- többszörös regresszió analízis 530-547
REML módszer 345, 453, 458
ANOVA 534-535
GenStat programmal 539-541
interakció 537-539
kollinearitás 536-537
lineáris modell 530-531
pathanalízis 544-549
Student t-próba 534-535
- változók kiválasztása 540-544
- többszörös terjedelelem vizsgálatok 135-151
- többszörös kísérlet 61-69
- t-próba 111-125
egymintás 112-114
kétmintás 114-118
páros 118-120
GenStat programmal 123-125
- ültetvény kísérletsorozat 387-400
- üzemi (demonstrációs) kísérletek 401-441
technológia fejlesztési 402-411
technológia összehasonlítás 412-426
termesztési rendszerek (FSR) 427-441
- variancia 22
- varianciaanalízis (ANOVA) 22-24
modelljei 23, 235-237
alapelvei 126-151
feltételei 131-134
varianciát infláló faktor (VIF) 541
- változók 36-37
eloszlása 26-28
típusai 76-78
- varianciaanalízis GenStat programmal I.
GenStat ANOVA

