

Közzététel: 2020. május 12.

A tanulmány címe:

Vargha András: Többváltozós statisztika dióhéjban: Változó-orientált módszerek

Szerző:

VARGA BALÁZS ANDRÁS, az Eötvös Loránd Tudományegyetem PhD-hallgatója

E-mail: varga.balazs@ppk.elte.hu

DOI: <https://doi.org/10.20311/stat2020.5.hu0467>

Az alábbi feltételek érvényesek minden, a Központi Statisztikai Hivatal (a továbbiakban: KSH) *Statisztikai Szemle* c. folyóiratában (a továbbiakban: Folyóirat) megjelenő tanulmányra. Felhasználó a tanulmány vagy annak részei felhasználásával egyidejűleg tudomásul veszi a jelen dokumentumban foglalt felhasználási feltételeket, és azokat magára nézve kötelezőnek fogadja el. Tudomásul veszi, hogy a jelen feltételek megszegéséből eredő valamennyi kárért felelősséggel tartozik.

1. A jogszabályi tartalom kivételével a tanulmányok a szerzői jogról szóló 1999. évi LXXVI. törvény (Sztj.) szerint szerzői műnek minősülnek. A szerzői jog jogosultja a KSH.
2. A KSH földrajzi és időbeli korlátozás nélküli, nem kizárólagos, nem átadható, térítésmentes felhasználási jogot biztosít a Felhasználó részére a tanulmány vonatkozásában.
3. A felhasználási jog keretében a Felhasználó jogosult a tanulmány:
 - a) oktatási és kutatási célú felhasználására (nyilvánosságra hozatalára és továbbítására a 4. pontban foglalt kivétellel) a Folyóirat és a szerző(k) feltüntetésével;
 - b) tartalmáról összefoglaló készítésére az írott és az elektronikus médiában a Folyóirat és a szerző(k) feltüntetésével;
 - c) részletének idézésére – az átvevő mű jellege és célja által indokolt terjedelemben és az eredetihez híven – a forrás, valamint az ott megjelölt szerző(k) megnevezésével.
4. A Felhasználó nem jogosult a tanulmány továbbértékesítésére, haszonszerzési célú felhasználására. Ez a korlátozás nem érinti a tanulmány felhasználásával előállított, de az Sztj. szerint önálló szerzői műnek minősülő mű ilyen célú felhasználását.
5. A tanulmány átdolgozása, újra publikálása tilos.
6. A 3. a)–c.) pontban foglaltak alapján a Folyóiratot és a szerző(ke)t az alábbiak szerint kell feltüntetni:

„*Forrás: Statisztikai Szemle* c. folyóirat 98. évfolyam 5. számában megjelent, **Varga Balázs András** által írt, **'Vargha András: Többváltozós statisztika dióhéjban: Változó-orientált módszerek'** című tanulmány (link csatolása)”

7. A Folyóiratban megjelenő tanulmányok kutatói véleményeket tükröznek, amelyek nem esnek szükségképpen egybe a KSH vagy a szerzők által képviselt intézmények hivatalos álláspontjával.

Varga Balázs András

Vargha András: **Többváltozós statisztika dióhéjban: Változó-orientált módszerek***

VARGA BALÁZS ANDRÁS, az Eötvös Loránd Tudományegyetem PhD-hallgatója
E-mail: varga.balazs@ppk.elte.hu

Ez a hiánypótló pszichometriai kötet gyakorlati szempontból közelíti meg azokat a napjainkban nélkülözhetetlen változó-orientált statisztikai eljárásokat, melyek elengedhetetlenek a XXI. századi pszichológiai – és egyéb módszertanilag rokonítható társtudományok körében alkalmazott – kutatások minőségi kivitelezéséhez. A könyv egyszerre célozza meg az elmélethez tartozó fogalmak gyakorlati szinten szükséges tisztázását, illetve az eljárások lépéseinek könnyen követhetővé tételét. A lépések kivitelezéséhez mindig a legjobban illeszkedő programo(ka)t kínálja a szerző, többször akár párhuzamosan bemutatva a megközelítéseket az egyes felületeken (Excel, SPSS, RopStat, Mplus, R). A könyv azok számára nyújt könnyen használható információkat, akik a társadalomtudományi módszertanban használt statisztikai alapfogalmakkal készség szinten tisztában vannak, és tudásukat bővíteni szeretnék a többváltozós eljárások irányába, de nem törekcszenek e statisztikai módszerek szükségesnél mélyebb matematikai/statisztikai megértésére. A könyv tömör rövideége is jelzi, hogy a szerző célja a gyakorlati szempontok megválasztásához nélkülözhetetlen információk átadása. A kötet *Vargha András* „Matematikai statisztika” című könyve (2000) méltó folytatásának, kiterjesztésének tekinthető. Az első rész a kvantitatív változók közötti kapcsolatokat tárgyalja és mutatja be, kiemelve a többszörös lineáris és nemlineáris regresszió-, valamint a varianciaanalízis azonosságait, illetve különbségeit. Míg a második rész a kvantitatív változók hierarchikus struktúrájával foglalkozik, azon belül is a dimenziócsökkentő eljárásokra helyezve a hangsúlyt.

Az első fejezetben a szerző kifejti a regresszióanalízis egyváltozós modelljét, majd erre fűzi fel a többváltozós gyakorlati alkalmazás során szükséges alapfogal-

* Pólya Kiadó. Budapest. 2019. 160 old.

makat és alkalmazási feltételeket. A feltételek teljesülésének lényegi megértése – tapasztalataim szerint – a felhasználók körében a regresszióelemzés vakfoltja, inkább szükséges rutinként tekintenek rá a módszertani elemzések során, ezért ezek tisztázása minden kutató számára nélkülözhetetlen lenne; ezt tökéletesen segíti ez a fejezet.

A második fejezet a nemlineáris kapcsolatok ábrázolási és magyarázó modelljeinek a bemutatásával foglalkozik. Az ábrázolás módját három program (Excel, RopStat, SPSS) lehetőségein és lépésein keresztül demonstrálja. Ezután az olvasó a polinomiális regresszió egyenletével és elméleti háttérével ismerkedhet meg. A lineáris regressziónál szemléltetett lépésenkénti modellépítéshez hasonlóan itt is információt kaphatunk a különböző módszerekről és mutatókról. Sajnos – kivételesen – egy konkrét modellépítési folyamat bemutatására nem került sor ebben a fejezetben, pedig tovább növelte volna a megközelítés biztos használatát azok számára, akik innen tájékozódnak majd az elemzést illetően.

A harmadik fejezet a varianciaanalízis és a kovarianciaanalízis használatát és szempontrendszerét tárja az olvasó elé, aki számára nagy segítség, hogy a varianciaanalízis egyenletalapú megközelítését is részletesen ismerteti a szerző. Így könnyebben érthetővé válik, hogy ez valójában lineáris modellnek tekinthető-e, a varianciaanalízis és más lineáris modellek matematikai háttére valójában nem tér-e el nagymértékben egymástól. Ehhez hasonlóan nagy segítség a hatásmagnaságok szempontjainak a megfeleltetése is a regressziós modelleknél bemutatottakkal. Az effajta analógiák jelentősen támogatják a megértést, és elősegítik a statisztika értő használatát. Különös figyelmet érdemel az összetartozó mérések többtényezős vizsgálatának lépésenkénti bemutatása, az ilyen jellegű tesztelés részletes ismertetését ritkán érinti efféle megközelítésben a magyar nyelvű szakirodalom.

A második rész a gyakorlati alkalmazás során is jellemzően összekapcsolódó ívet ölel fel: a tételek dimenziószámának csökkentése, az így létrehozott redukált modell illeszkedésének vizsgálata, illetve annak opcionális további csoportosítása klaszteranalízissel.

A főkomponens-elemzés és a faktorelemzés fogalmi és gyakorlati szétválasztása nem csak az egyetemi hallgatók számára okoz problémát a tapasztalataim szerint. A legtöbbször kihangsúlyozott, modellekben használt oksági irányok különbségén túl hiánypótló módon ismerteti a szerző a két megközelítés egyéb fogalmi és gyakorlati, matematikai különbségeit. A fejezet célja többek között a mutatók és a számértékek összefüggésbeli tisztázása, ami az alkalmazott eljárás során hüvelykujj-szabályként szokott felmerülni. Hogyan függnek össze a sajátérték, a megmagyarázott variancia, a kommunalitás és a faktorsúlyok fogalmai? A szerző a legadekvátabb módszer kiválasztását sokszor a kutatóra bízta, de ezt minden alkalommal jól elkülönülő szempontok bemutatása után teszi; ebben nagy segítséget nyújt a feltáró elemzések típusainak részletes ismertetése is.

Napjainkban a korábbi gyakorlaton túl szükséges lépéssé vált a feltáró faktoranalízis után a pszichometriai tesztek konfirmatív elemzése is. Ennek lépéseire és a modell struktúrájának kialakítási szempontjaira eddig kevés magyar nyelvű szakirodalom fordított nagy figyelmet. A szerző Mplus és R szoftverek segítségével is ismerteti az elemzés folyamatát, külön kitérve a parancssor megírására és az eredménytábla értelmezésére. Részletekbe menően tekinti át az eredménytáblán megjelenő adekvációs mutatók jelentését, illetve a jellemzően SEM- (structural equation modeling – strukturális egyenletek modellezése) felületeket használó, látens és manifeszt változókat tartalmazó ábra értelmezését.

A lehetséges megfigyelésalapú, illetve változóalapú klaszterezési eljárások közül – a könyv címének is megfelelően – a változók csoportosítását célzó eljárást mutatja be a szerző, kitérve a különböző távolságszámításokra és klaszterképzési eljárásokra is. Mérsékli a fejezet teljességét, hogy csak a többváltozós megközelítést írja le, míg a személyközpontú klaszterezés ismertetése a kötetben sajnos elmarad. Ahogy Vargha András a bevezető fejezetben említi, ennek az eljárásnak a részletekbe menő kifejtését egy másik, remélhetőleg a közeljövőben megjelenő kötetbe tartogatja, amelyben emellett más minőségi függőváltozós eljárásokkal is foglalkozik majd.

A mű rövidege ellenére a szerző úgy mutatja be az egyes fogalmakat és megközelítéseket, hogy azok beilleszthetővé válnak egy holisztikus képbe, így nemcsak az adott elemzési módszerekkel ismerkedhet meg az olvasó, hanem a könyv elolvasásával érthetővé válhatnak számára a módszerek különbségei, strukturális és fogalmi azonosságai, illetve összefüggései is. Saját tapasztalatom, hogy a statisztikával foglalkozó szakemberek egyre tágítják tudásukat, szívesen használnak párhuzamosan több különféle szoftvert is, miután megismerték azok előnyeit és hátrányait. Vargha András könyve hiánypótló mű, ajánlom azoknak, akik nemcsak elemzéseket akarnak végezni, de könnyebb tájékozódásra is vágnak a társadalomtudományi statisztikai eljárások hatalmas „rengetegében”.