FEJLESZTŐPROGRAMOK EGYMINTÁS, KRITÉRIUMORIENTÁLT HATÁSVIZSGÁLATÁNAK MATEMATIKAI STATISZTIKAI HÁTTERE

Szerzők:

Mező Ferenc
Debreceni Egyetem

Máth János
Debreceni Egyetem

Abari Kálmán
Debreceni Egyetem

Mező Katalin
Debreceni Egyetem

Lektorok:

Demetrovics János
Eötvös Lóránd Tudományegyetem

Nagy Dénes
Nemzetközi Szimmetria Társaság

Varga Imre
Szegedi Tudományegyetem

Koncz István
Professzorok az Európai Magyarországért

Első szerző e-mail címe:
ferenc.mezo1@gmail.com

Mező Ferenc, Máth János, Abari Kálmán és Mező Katalin (2015): Fejlesztőprogramok egymintás, kritériumorientált hatásvizsgálatának matematikai statisztikai háttere. *Különleges Bánásmód*, I. évf. 2015/3. szám, 69-78. DOI 10.18458/KB.2015.3.69

Absztrakt

E tanulmány a különleges bánásmódot igénylő tanulók számára készült fejlesztőprogramok egymintás, kritérium szinthez viszonyító (a fejlesztési tervben meghatározott célérték elérését tesztelő) hatásvizsgálatának matematikai statisztikai elemzéséhez kínál módszertani útmutatót. Praktikus segítséget nyújtunk az egymintás különbségvizsgálatok matematikai statisztikai számításainak kiválasztásához (vö.: Abari és tsai, 2015), a számítások elvégzésére alkalmas R statisztikai szoftverbe írható parancssorokhoz, az R által végrehajtott számítások eredményeinek értelmezéséhez, szövegbe foglalásához.

Kulcsszavak: fejlesztőprogram, hatásvizsgálat, statisztika, R nyelv

Diszciplina: matematika, pszichológia, gyógypedagógia, pedagógia

Abstract

MATHEMATICAL STATISTICAL BACKGROUND OF ONE SAMPLE, CRITERIUM ORIENTED EFFECTIVENESS STUDY OF DEVELOPMENT PROGRAMMES

This paper offers a methodological guidance to mathematical statistical analysis of criterium oriented effectiveness studies (in case of examining one sample, when we compare the observed values to a criterium value of a development plan) of development programmes for children who need special treatment. We provide a practical help to choose the adequate mathematical statistical test for examiing differences in case of one sample (see: Abari et all, 2015), to create commands for R statistical software, as well as to interpretation of R results.

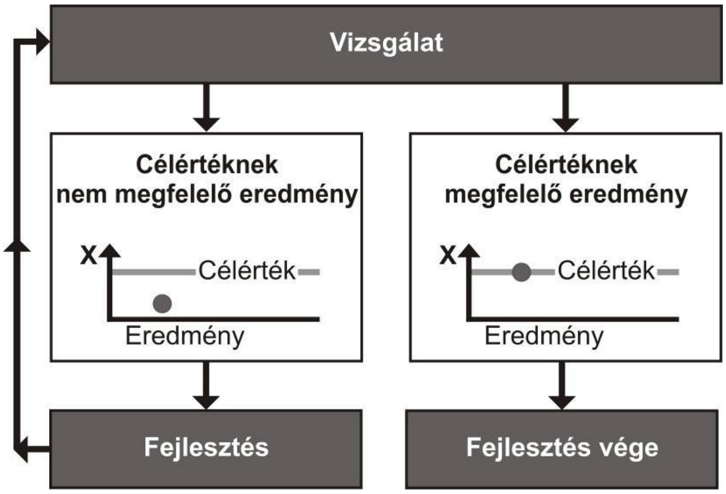
Key words: development programme, effectiveness study, statistic, R language

Disciplines: mathematics, psychology, special education, pedagogy

A különleges bánásmódot igénylők tervszerű fejlesztése feltételezi, hogy létezik egy előre meghatározott cél, aminek elérésére törekszünk a fejlesztés során. Egy fejlesztő program hatékonyságát pedig az is jelezheti, hogy e célértéket (kritériumot) megközelítették, elérték vagy túlhaladták-e a programba bevont személyek.

Egymintás, kritérium szinthez viszonyító hatásvizsgálatnak nevezzük azokat az eseteket, amikor (például különleges bánásmódot igénylő tanulók számára alkotott) fejlesztőprogramok hatékonyságát aszerint ítéljük meg, hogy a résztvevők elértek-e a program hatására egy előre kitűzött, fejlesztési tervben rögzített célértéket. A fejlesztőprogramba ágyazott ilyen hatásvizsgálatok sémáját az 1. ábra foglalja össze.

1. ábra: egymintás, kritériumszinthez (pl. fejlesztési tervben kitűzött célértékhez) viszonyított hatásvizsgálattal minőségbiztosított fejlesztőprogram sémája (forrás: a Szerző)



E vizsgálatokban ehhez hasonló kérdések fogalmazhatók meg (Mező, Máth és Abari, 2008 alapján):

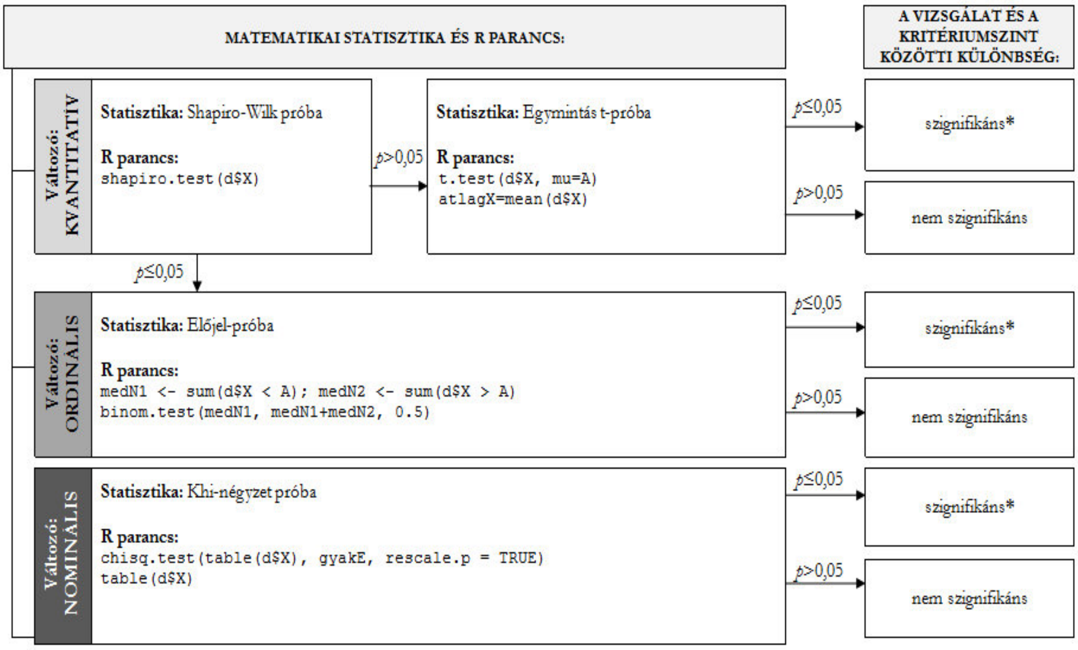
a) A fejlesztőprogramba bevont tanulók IQ-ja lényegesen eltér az átlagtól (a 100-as IQ-tól)?

b) Egy fejlesztőprogramban résztvevő diákok szövegértést felmérő dolgozatra kapott érdemjegyei összességében közepesnél (3-asnál) kisebbek vagy nagyobbak?

c) Egy művészetterápiás módszereket használó fejlesztőprogramban résztvevő tanulóközösség minden tagja az alábbi négy pályaorientációs lehetőség közül egyet választhatott ki saját életpályaként: a) tudományos pálya, b) művészeti pálya, c) sport pályafutás, d) nem tudom. Kérdés, hogy a négy lehetséges választ egyenlő gyakorisággal (az esetek 25-25%-ban) választották-e a vizsgálati személyek? Vagy: kérdés, hogy a művészeti pályát nagyobb eséllyel választották-e a diákok, mint bármelyik másikat?

E kérdések közös vonása, hogy egy csoport egyetlen vizsgálati alkalommal nyert eredményét viszonyítjuk egy adott értékhez (100-as IQ-hoz, 3-as osztályzathoz, vagy éppen a 25%-os előfordulási gyakorisághoz). Amennyiben a kérdésre adható válaszokat nemcsak az adott tanulókra, hanem általában a hasonló (pl. azonos korosztályba tartozó) tanulókra is általánosítani szeretnénk, illetve, ha nincs meg ez az általánosítási törekvés, de a csoport eredményét valószínűségi alapon is elemezni szeretnénk, akkor matematikai statisztikai módszereket kell felhasználnunk. A 2. ábra mutatja be, hogy milyen döntések meghozatala után választhatjuk ki az adekvát matematikai statisztikai eljárást, amennyiben egy vizsgálati csoport egyetlen alkalommal mért vizsgálati eredményét hasonlítjuk egy adott kritérium-szinthez.

2. ábra: Statisztikaválasztás kritériumorientált hatásvizsgálat esetében (d$X = vizsgálati adatokat tartalmazó változó a „d" adatbázisban; A = kritériumszint; gyakE = várt gyakoriságok; R parancs = a statisztikai számítások elvégzéséhez javasolt ingyenes szoftvernek, az R-nek adható parancs). Forrás: Mező, Máth és Abari (2008) alapján Mező F.



*** Ha X változó magasabb értéke jelenti a kedvezőbb eredményt (pl. jobb képességet), akkor a fejlesztési cél(érték) eléréséről vagy túlteljesítéséről d$X ≥ A esetben beszélhetünk; d$X < A viszony esetében a fejlesztési cél(érték) elérése még nem történt meg. Ha X változó alacsonyabb értéke jelenti a kedvezőbb eredményt (pl. kisebb mértékű szorongást), akkor d$X ≤ A eredmény jelenti a fejlesztési cél(érték) elérését vagy meghaladását; míg d$X > A eredmény jelzi, hogy a kitűzött fejlesztési cél(érték) elérése még nem történt meg.**

A 2. ábrán látható, hogy a statisztikaválasztás első szempontja, hogy kvantitatív (értékei között azonos intervallumokat alkalmazó, átlagolható), ordinális (rangsorolható értékekkel rendelkező) vagy csak nominális (csak megkülönböztethető, de nem rangsorolható értékekkel rendelkező) skálájú változókkal dolgozunk-e. A fenti kérdések példájánál maradva: az IQ például kvantitatív, az osztályzat ordinális a pályaorientációs választás ebben az esetben pedig nominális változónak tekinthető. A továbbiakban e három skálatípus szerint alfejezetekbe rendezve nyújtunk módszertani útmutatót az egymintás különbségvizsgálatok matematikai statisztikai számításainak kiválasztásához, a számítások elvégzésére alkalmas R statisztikai szoftverbe írható parancssorokhoz (vö.: Abari és tsai, 2015; Reiczigel és Solymosi, 2007; Solymosi, 2005), az R által végrehajtott számítások eredményeinek értelmezéséhez, szövegbe foglalásához. Jelen módszertani útmutatóban az R parancsokkal kapcsolatban a következő jelöléseket vezetjük be (vö.: Abari, 2008):

- *d*: egy adattábla neve, melyben valamennyi személy adata egyetlen sorban szerepel, még akkor is, ha több különböző mérés tartozik hozzá. Az adattábla oszlopai a vizsgált változók (pl. *X1* vagy „IQ" változók) adatait tartalmazzák.
- *d$X*: a *d* adatbázis *X* változójára történő hivatkozás. Ha a parancsban szerepel data=d argumentum, akkor a '*d$*' minősítés elmaradhat, elegendő az *X*-t megadnunk.

Kvantitatív változók a kritériumorientált hatásvizsgálatban

Az olyan kérdések, mint „Egy IQ teszt középértéke = 100?" vagy „A csoport által egy vizsgálatban elért teljesítményszázalék lényegesen eltér 50%-tól?" vagy „A vizsgált gyermekek centiméterben kifejezett testmagassága megfelel az életkori átlagnak (120 cm-nek)?" vagy „A divergens gondolkodást fejlesztő programban résztvevők fluencia pontszámbeli teljesítménye a program végére elérte-e a fejlesztési tervben kitűzött min. 20 pont célértéket?" általánosabban fogalmazva az alábbi kérdésre keresik a választ:

„X kvantitatív változó középértéke egyenlő-e egy bizonyos A számmal?"

Az IQ-ra vonatkozó fenti kérdés esetében tehát X jelöli az intelligenciahányadost; az A-érték pedig a 100-as IQ-nak feleltethető meg. A másik kérdésben X = teljesítmény százalék és A = 50%. A testmagasság (= X változó) vizsgálata során az A-érték az életkori átlagnak tekintett 120 cm. Végül: a fluencia (= X változó) A-értéke pedig a 20 pontban megállapított célérték.

Amikor e példákban közöltekhez hasonló helyzetben egy (például: fejlesztési tervben kitűzött) célérték elérését vizsgáljuk matematikai statisztikai módszerekkel, akkor a 2. ábra kvantitatív változókra utaló sorában látható módon járhatunk el. Vagyis: kvantitatív változók használatakor először (pl. Shapiro-Wilk próbával) meg kell vizsgálnunk, hogy a változó normális eloszlást követ-e. Normális eloszlás esetén egymintás t-próbát használhatunk a továbbiakban, ellenkező esetben az ordinális változóknál bemutatásra kerülő előjel-próbát alkalmazhatunk. Részletesebben:

A Shapiro-Wilk-próbát a fenti esetben az „X kvantitatív változó normális eloszlású?" kérdés eldöntésére használjuk (a próba nullhipotézise az, hogy „X változó normális eloszlású"), s az alábbi R parancsot kell használnunk:

```
shapiro.test(d$X)
```

Publikációnkban mindez az alábbihoz hasonlóan is megfogalmazható: *„Vizsgálatunkban arra voltunk kíváncsiak, hogy a csoport általános intellektuális teljesítménye (IQ-ban kifejezve) lényegesen eltér-e a 100-as IQ-értéktől. Az adekvát matematikai statisztikai próba megválasztásához először az IQ normális eloszlását Shapiro-Wilk próbával ellenőriztük, amelynek eredménye szerint…"* A mondat folytatása pedig a normalitásvizsgálat szignifikancia szintjének (p-értékének) nagyságától függ. Két lehetőség adódik:

1) Ha a Shapiro-Wilk próba eredménye szerint $p > 0{,}05$, akkor az X változót normális eloszlásúnak tekinthetjük, aminek gyakorlati következménye az, hogy lehetőség adódik egymintás t-próbát alkalmaznunk, s az X változó átlagát vethetjük össze az A-értékkel. Az imént megkezdett mondat pedig ehhez hasonlóan folytatódhat: *„…az IQ-t normális*

eloszlásúnak tekinthetjük (p=p-value). Ennek fényében lehetőségünk nyílt egymintás t-próba alkalmazására." Az egymintás t-próba (melynek nullhipotézise: „az X változó átlaga = A") a következő R paranccsal végezhető el:

```
t.test(d$X, mu=A)
atlagX=mean(d$X)
```

Amennyiben az egymintás t-próba eredménye szerint $p > 0{,}05$, akkor az „*X* változó populáció átlaga = A-érték" nullhipotézist megtarthatjuk (másképp: nincs szignifikáns eltérés az X változó átlaga és az A-érték között). Szövegbe foglalva pedig ezt akár így is megfogalmazhatjuk: *„Az egymintás t-próba eredménye szerint az általános intellektuális teljesítmény (IQ) nem tér el szignifikánsan a 100-as értéktől (átlag=atlagX; p=p-value)."*
Ellenben, amikor az egymintás t-próba eredménye az, hogy $p \leq 0{,}05$ (vagyis az X változó átlaga szignifikánsan eltér az A-értéktől), akkor írásunkban ezt így foglalhatjuk össze: *„Az egymintás t-próba eredménye szerint az általános intellektuális teljesítmény (IQ) jelentősen eltér a 100-as értéktől (átlag=atlagX; p=p-value)."* A mondat (elemzés) folytatásaként pedig fejtsük ki, hogy a vizsgált személyek átlaga nagyobb vagy kisebb-e az adott A-értéknél: ha az *átlagX* > 100, akkor a diákok IQ-ja átlag feletti; ha az *átlagX* < 100, akkor a diákok IQ-ja átlag alatti.

2) Ha a Shapiro-Wilk próba eredménye: $p \leq 0{,}05$, akkor azt állapíthatjuk meg, hogy *„…(p=p-value), e változót nem tekinthetjük normális eloszlásúnak."*. Ebben az esetben nem lehetséges a vizsgált változó átlagát hasonlítani egy adott célértékhez; de lehetőség van az ordinális változóknál alkalmazható előjel-próba alkalmazására (lásd: következő fejezet).

A normális eloszlást követő kvantitatív változók értékeit általában táblázattal, hisztogrammal, grafikonnal, oszlopdiagrammal szemléltethetjük.

Ordinális változók a kritériumorientált hatásvizsgálatban

A rangsorolható értékekkel rendelkező ordinális változók (illetve nem normális eloszlású kvantitatív változók) esetében előjel-próbával vizsgálhatjuk az ehhez hasonló jellegű kérdéseket (vö.: 2. táblázat):

„X (legalább ordinális) változó mediánja egyenlő egy bizonyos A értékkel?"

Ez az általános jellegű kérdésfelvetés olyan konkrét helyzetekkel hozható párhuzamba, mint: „Egy fejlesztő programban résztvevő diákok szövegértésre kapott érdemjegyei összességében közepesnél (3-asnál) kisebbek vagy nagyobbak?", vagy: „Az intelligenciatesztben a 100 IQ-pontnál többet elérők vannak-e többen, vagy a kevesebb IQ-pontot elérők?" Az előjel-próba (melynek nullhipotézise szerint „X változó mediánja = A") R parancsa:

```
medN1 <- sum(d$X < A)
medN2 <- sum(d$X > A)
binom.test(medN1, medN1+medN2, 0.5)
```

A fenti R utasításokban *medN1* változó fogja tartalmazni a feltételezett mediánnál, vagyis *A*-nál kisebb értékek számát; *medN2* változó pedig a feltételezett mediánnál, vagyis *A*-nál nagyobb értékek számát fogja jelölni.
Az előjel-próba $p > 0{,}05$ szignifikancia szintje arra utal, hogy az *A* értéktől nagyobb értékek előfordulásának valószínűsége ugyanakkora, mint a kisebb értékek előfordulásának

valószínűsége – vagyis: a nullhipotézist megtartjuk. Eredményünkről ehhez hasonló módon számolhatunk be a szövegértésre vonatkozó fenti példa esetében: *„Előjel-próbával ellenőriztük a kapott eredményeket. Az előjel próba szerint az érdemjegyek nem különböznek jelentősen a 3-as értéktől (p=p-value).”* Az IQ-ra vonatkozó példa esetében pedig így fogalmazhatunk: *„Előjel-próbát alkalmaztunk annak eldöntésére, hogy a 100 IQ-pontnál nagyobb vagy kisebb pontszámot elérők vannak-e többségben. Az előjel próba szerint az IQ mediánja nem különbözik jelentősen a 100-as értéktől (p=p-value).”*

Ezzel szemben az előjel-próba $p \leq 0,05$ szignifikancia szintje azt jelzi, hogy az *A* értéktől nagyobb értékek előfordulásának valószínűsége nem ugyanakkora, mint a kisebb értékek előfordulásának valószínűsége (tehát: az előjel-próba nullhipotézisét el kell vetnünk). A szövegértésre kapott érdemjegyek esetében publikációnkban ilyen jellegű közlést tehetünk: *„Előjel-próbával ellenőriztük a kapott eredményeket. Az előjel próba szerint az érdemjegyek jelentősen különböznek a 3-as értéktől (p=p-value): 3 alatt medN1 érték, 3 felett medN2 érték található.”* Az IQ-ra vonatkozó esetben is hasonló sablon-mondatot alkalmazhatunk: *„Előjel-próbával ellenőriztük, hogy a 100 IQ-pontnál nagyobb vagy kisebb pontszámot elérők vannak-e többségben. Az előjel-próba szerint az IQ mediánja jelentősen különbözik a 100-as értéktől (p=p-value). 100-as IQ alatt medN1 fő teljesített, 100-as IQ feletti eredményt pedig medN2 tanuló ért el.”*

Az ordinális szintű változók érétkeinek alakulását dobozdiagrammal (boxplot-ábrával) szemléltethetjük, mely ábrázolja a mediánt, a középmezőny alsó és felső 25%-át, a legjobb/leggyengébb 25%-ba tartozó értékek alakulását, valamint az extrém magas/alacsony eredménynek tekinthető értékeket (ha vannak ilyenek egyáltalán).

Nominális változók a kritériumorientált hatásvizsgálatban

Az egymástól megkülönböztethető, de nem rangsorolható értékeket tartalmazó (röviden: nominális) változók esetében a lehetséges értékek előfordulási gyakoriságait vethetjük össze egy adott értékkel. Általános formában tehát a következő kérdésre kereshetünk választ ilyenkor:

„X nominális változó értékeinek megfigyelt (x1, x2…xn) gyakoriságai egyenlők-e adott A=(n1, n2…n3) gyakoriságokkal?”

Erre az absztrakt kérdésre fókuszál a korábban felvetett pályaorientációs témájú példa: vajon a tudományos, a művészeti, a sport pálya, illetve a „nem tudom” választ egyenlő eséllyel választották? Itt X = a válasz (lehetséges értékei: tudományos, művészeti,sport, nem tudom), s A=(25, 25, 25, 25) – tekintve, hogy négy lehetséges válasz esetén 25-25% az esélye egy-egy válasz véletlenszerű (azaz egyenlő esélyű) előfordulásának. Értelemszerűen ez az esély: két válaszlehetőség esetén 50-50%, vagyis A = (50, 50); öt lehetséges válasz esetén 20-20% (A= (20, 20, 20, 20, 20) stb. E kérdés eldöntésére szolgáló matematikai statisztikai eljárás a khi-négyzet-próbával (χ^2-próbával) végzett eloszlás vizsgálat lehet, melynek nullhipotézise: az *X* változó populációra vonatkozó gyakoriságai közel egyenlők a várt gyakoriságokkal. Az R parancs - megjegyzés: *gyakE*=(n1, n2, n3, n4) a várt gyakoriságokat tartalmazza – most a következő lesz:

```
chisq.test(table(d$X), p = gyakE, rescale.p = TRUE)
```

A khi-négyzet-próba szignifikancia szintjének függvényében dönthetünk a nullhipotézis elfogadásáról, vagy elutasításáról. Ha a khi-négyzet-próba esetén a $p > 0,05$ eredmény arra utal, hogy nincs szignifikáns különbség a megfigyelt és a várható gyakoriság között (a nullhipotézist elfogadhatjuk). Szövegbe ágyazottan közreadhatjuk, hogy: *„A művészetterápiás módszereket használó fejlesztőprogram végén a tanulók négyféle válasz (»tudományos pálya«, »művészeti pálya«, »sport pályafutás«, »nem tudom«) közül egyet jelölhettek meg, amikor pályairányultságukról kérdeztük őket. Vizsgálatunkban arra voltunk kíváncsiak, hogy a négy választ egyforma (25%-os) gyakorisággal választották-e a diákok. A khi-négyzet próba szerint az egyes válaszok gyakoriságai között nincs lényeges különbség (p=p-value). Így például a művészeti pályát sem választották lényegesen gyakrabban, mint a többit."*

Ezzel szemben a khi-négyzet-próba eredményeként tapasztalt $p \leq 0,05$ érték jelzi, hogy szignifikáns különbség van a megfigyelt és a várható gyakoriságok között, s a nullhipotézist elvethetjük. Szövegszerűen pedig így foglalhatjuk össze tapasztalatunkat: *„A művészetterápiás módszereket használó fejlesztőprogram végén a tanulók négyféle válasz (»tudományos pálya«, »művészeti pálya«, »sport pályafutás«, »nem tudom«) közül egyet jelölhettek meg, amikor pályairányultságukról kérdeztük őket. Vizsgálatunkban arra voltunk kíváncsiak, hogy a négy választ egyforma (25%-os) gyakorisággal választották-e a diákok. A Khi-négyzet próba szerint az egyes válaszok gyakoriságai között lényeges különbség van (p=p-value). A megfigyelt gyakoriságok a következők: m1, m2, m3, m4".* Az *m1*, *m2*, *m3*, *m4* értékeit a következő R-parancs révén határozhatjuk meg:

```
table(d$X)
```

Ha a művészeti pályát választók gyakorisága lényegesen nagyobb mint a többi, akkor a szöveg így folytatódhat: *„A művészeti pályát lényegesen gyakrabban választották, mintha csak véletlenszerűen tippeltek volna."* A gyakoriságokat táblázattal és/vagy oszlopdiagrammal, kördiagrammal vagy 100%-ig halmozott oszlopdiagrammal szemléltethetjük.

Zárógondolatok

E tanulmányban a fejlesztő programokban számszerűen is megfogalmazott célértékhez (például egy adott tesztpontszámhoz) viszonyított hatásvizsgálatok matematikai statisztikai alapjait foglaltuk össze. Négy matematikai statisztikai próbát használtunk mindehhez:

- Shapiro-Wilk próbát alkalmaztunk a kvantitatív változók normális eloszlásának vizsgálatára;
- Egymintás t-próbát javasoltunk (a normális eloszlású) kvantitatív változók átlagának egy adott értékkel történő összehasonlítására;
- Előjel-próba használható az ordinális változók (illetve a nem normális eloszlású kvantitatív változók) mediánjának egy adott értékkel történő összevetésére;
- Khi-négyzet-próbával (χ^2-próbával) végzett eloszlásvizsgálattal teszteltük a nominális változók lehetséges értékei gyakoriságának egyenlőségét.

Lényeges, hogy az itt tárgyaltakon túl további statisztikai próbák, módszerek is léteznek hasonló kérdések eldöntésére (lásd: Varga, 2000).

Noha a fenti példákban nem emeltük ki külön, lényeges azonban tudatosítani, hogy a matematikai statisztikai próbák eredményei további feltáró jellegű elemzéseket igényelnek. Az alkalmazott statisztikai próbától és annak eredményétől függetlenül minden esetben célszerű olyan kérdésekre is választ keresni, mint:

- Vizsgálati eredményeink megfelelnek vagy ellentmondanak a hétköznapi tapasztalatainknak, illetve a szakirodalomban olvasható egyéb kutatási eredményeknek? Sablon mondatok lehetnek például: *„Eredményeink alátámasztják XY1 (2010), XY2 (2012), valamint XY3 és XY4 (2015) kutatási tapasztalatait. Ugyanakkor ellentmondanak XY5 (1998) eredményyeinek."*

- Vajon mi állhat eredményeink hátterében (mi lehet az oka, hogy éppen ilyen, s nem eltérő eredményeket kaptunk), s ezzel összefüggésben mi lehet az oka, hogy vizsgálati tapasztalataink megfelelnek, vagy éppen ellentmondanak mindennapi tapasztalatainknak, illetve a szakirodalomban közölteknek? Szövegbe fogalmazva például: *„Eredményeink hátterében általában véve, illetve mások eredményeivel való kapcsolatuk hátterében a következő okok feltételezhetők: a)…; b) …; c) …”*

- Kutatási eredményeinknek milyen gyakorlati következményei vannak? Eredményeink gyakori következményei lehetnek például: a) további kutatások alapjai lehetnek; b) diagnosztikai jelentőséggel bírhatnak; c) fejlesztő módszerek/programok minőségbiztosításához vagy éppen létrehozásához vezethetnek (esetleg: hatástalanságuk vagy káros hatásuk miatt e programok elvetésére indíthatnak); d) felhasználhatók a graduális, posztgraduális, tanfolyami jellegű képzések tervezésében, létrehozásában. Ennek felvezető mondata így kezdődhet: *„Vizsgálati eredményeink gyakorlati következménye lehet például: …”*

- Milyen korlátai lehetnek vizsgálatunknak, s e korlátok miként orvosolhatók a jövőbeli hasonló témájú kutatások alkalmával? Felvezető mondat lehet például: *„Vizsgálatunk korlátjának tekinthető, hogy…”* Vizsgálatok gyakori korlátjai lehetnek például:
  a) a vizsgálatban használt fogalmak pontatlansága (megoldás: hatékonyabb szakirodalmi feldolgozás, pontosabb fogalom meghatározások alkalmazása);
  b) a kérdésfeltevés pontatlansága (megoldás: vizsgálati kérdések pontosítása – általában konkretizálása);
  c) a kérdésekkel kapcsolatban megfogalmazott hipotézisek pontatlansága (megoldás: hipotézisek körülírtabb megfogalmazása. Célszerű a hipotézisek alapját – pl. saját tapasztalatot, szakirodalmi forrásokat – is megjelölni, s a megfogalmazott hipotézist indokolni);
  d) a vizsgálati minta megválasztásának problémája (pl. miért pont az adott korosztályt vizsgáltuk), kiválasztásának megválasztása (pl. hibásan gondoltuk azt, hogy véletlenszerűen összeállított mintáról volt szó a vizsgálatban); reprezentativitásának problémája (pl. a vizsgálati minta nem reprezentálja jól a populációbeli – mondjuk: nemi, életkori stb. – viszonyokat); méretének megválasztása (pl. túl kevés vizsgálati személyt vontunk be a vizsgálatba). Megoldás a jövőbeli vizsgálatok során a mintaválasztás e kérdéseinek gondos mérlegelése, figyelembe vétele.
  e) a vizsgálatban mérni kívánt fogalmak operacionalizálása (mérhetővé tétele) nem megfelelő. Megoldás: az a) pontban megfogalmazottakkal összhangban szakirodalmi kutatás e fogalmak operacionalizálásával kapcsolatos próbálkozásokkal kapcsolatban és/vagy az operacionalizálásra tett korábbi kísérletek hibáit kiküszöbölő újszerű megközelítés kidolgozása.
  f) a vizsgálati módszerek és eszközök objektivitásának (tárgyilagosságának), validitásának (érvényességének), reliabilitásának (megbízhatóságának) problémái. Megoldás objektívebb, érvényesebb, megbízhatóbb eljárások választása, létrehozása.
  g) az adatfelvétel időpontjának (nem megfelelő napszakban, hónapban, évben történő adatfelvétel), időtartamának (túl sok vagy túl kevés időtartam), környezeti feltételeinek (pl. zavaró hang és fényeffektusok közötti adatfelvétel), pszichés feltételeinek (hiba lehet például, ha extrém módon szorongó vagy álmos vizsgálati személlyel végzett képességvizsgálat eredményeit nem szorongó és nem álmos személyek teljesítményére általánosítjuk), tartalmának (túl kevés/sok adat felvétele

történt meg), személyi feltételeinek (pl. nem megfelelően képzett adatfelvevő alkalmazásának) problémája. Megoldás: az adatfelvétel körülményeinek precízebb meghatározása, kontrollálása a jövőbeli vizsgálatokban.

h) az adatrögzítés problémás (sokáig tart, pontatlan stb.). Megoldás: adatrögzítés időtartamának, személyi feltételeinek, informatikai hátterének átgondolása, praktikusabb megtervezése a jövőben.

i) a statisztikai adatelemzés során inadekvát (az adott vizsgálati kérdésnek, hipotézis-tesztelésnek, változó típusnak és eloszlásnak, mintaszámnak nem megfelelő) matematikai statisztikai próba alkalmazása. Megoldás: szakirodalom tanulmányozása – pl. Falus és Ollé (s.a.), Mező, Máth és Abari (2008), Varga (2000); matematikai statisztikához szakember bevonása (pl. Különleges Bánásmód szerkesztőségén keresztül – e-mail: kb@ped.unideb.hu).

j) adatok értelmezésével, okok és következmények bizonyításának hiányával kapcsolatos korlátok. Megoldás: alternatív értelmezések felvetése, illetve a bizonyítás jövőbeli lehetőségének felvázolása

k) végül, de nem utolsó sorban etikai korlátok és megoldási lehetőségeik is feltüntethetők.

- Milyen erősségei lehetnek a vizsgálatnak? Aktualitása, újszerűsége, eredményeinek praktikus alkalmazási lehetősége említhető itt meg például. Szövegbeli felvezető mondat lehet: *„A vizsgálatunk erősségének tekinthetők a következők: …”*

- Milyen további kutatási lehetőségek vetődtek fel a vizsgálat kapcsán? Ezt egy ehhez hasonló mondattal vezethetjük fel írásunkban: *„E vizsgálat tapasztalatai alapján a következő kutatási témák, lehetőségek fogalmazódtak meg: …”*

Lényeges tehát megérteni azt, hogy a matematikai statisztikai elemzés nem (vég)célja a fejlesztő programok hatására fókuszáló különbségvizsgálatoknak, hanem pusztán eszköz e vizsgálatok adatelemzésének szolgálatában.

Végül hangsúlyoznunk kell, hogy a fejlesztőprogramok hatásvizsgálatának egyéb vizsgálati elrendezései is ismertek (ezekkel kapcsolatos R parancsokat lásd: Mező, Máth és Abari, 2008) – ilyenek például: az önkontrollos hatásvizsgálat, a nyomon követéses hatásvizsgálat, a kontrollcsoportos hatásvizsgálat. Ezekről, valamint a hozzájuk köthető R parancsokról jelen cikk Szerzői a Különleges Bánásmód című szakmai lap 2015-ben, illetve 2016-ban megjelenő számaiban közölnek majd további útmutatókat (Máth és tsai, 2015; Abari és tsai, 2015).

Irodalom

Abari Kálmán, Mező Ferenc, Mező Katalin és Máth János (2015): Fejlesztőprogramok hatásvizsgálatát szolgáló adatbázisok szerkezete egy ingyenes statisztikai szoftverben: az R-ben. *Különleges Bánásmód*, I. évf. 2015/2. szám, 37-47. DOI 10.18458/KB.2015.2.37

Falus I. & Ollé J. (s.a.): *Statisztikai módszerek pedagógusok számára.* Okker Kiadói Kft., Budapest.

Máth János, Mező Ferenc, Abari Kálmán és Mező Katalin (2015): Fejlesztőprogramok hatásvizsgálatának matematikai statisztikai alapfogalmai. *Különleges Bánásmód*, I. évf. 2015/1. szám, 69-77. DOI 10.18458/KB.2015.1.69

Mező F., Máth J. és Abari K. (2008): A különbségvizsgálatokon alapuló tehetségdiagnosztika matematikai statisztikai alapjai (adatelemzési útmutató). In Mező F. (Szerk.): *Tehetségdiagnosztika.* Kocka Kör & Faculty of Central European Studies, Constantine the Philosopher University in Nitra, Debrecen. pp 131-207.

Reiczigel J., Harnos A. és Solymosi N. (2007): *Biostatisztika nem statisztikusoknak.* Pars Kft., Nagykovácsi.

Solymosi N. (2005): *R<-…erre, erre…! Internetes R-jegyzet.* Letöltés: 2015.09.14. Web: http://cran.r-project.org/doc/contrib/Solymosi-Rjegyzet.pdf

Vargha A. (*2000*): *Matematikai statisztika pszichológiai, nyelvészeti és biológiai alkalmazásokkal.* Pólya Kiadó, Budapest.