

FEJLESZTŐPROGRAMOK EGYMINTÁS LONGITUDINÁLIS VIZSGÁLATÁNAK MATEMATIKAI STATISZTIKAI HÁTTERE

Szerzők:

Mező Ferenc
Debreceni Egyetem

Máth János
Debreceni Egyetem

Abari Kálmán
Debreceni Egyetem

Mező Katalin
Debreceni Egyetem

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

Lektorok:

Demetrovics János
Eötvös Lóránd Tudományegyetem

Koncz István
Professzorok az Európai Magyarorszáért

Szilágyi Barnabás
Debreceni Egyetem

Nemes Magdolna
Debreceni Egyetem

Mező F., Máth J., Abari K. és Mező K. (2016): Fejlesztőprogramok egymintás longitudinális vizsgálatának matematikai statisztikai háttere. *Különleges Bánásmód*, II. évf., 2016/1. szám, 63-72. DOI 10.18458/KB.2016.1.63

Absztrakt

E tanulmány a különleges bánásmódot igénylő tanulók számára készült fejlesztőprogramok egymintás, longitudinális (például ugyanazt a tanulócsoportot évente felmérő, nyomon követéses jellegű) hatásvizsgálatának matematikai statisztikai elemzéséhez kínál módszertani útmutatót. Praktikus segítséget nyújtunk a különbségvizsgálatok matematikai statisztikai számításainak kiválasztásához, a számítások elvégzésére alkalmas R statisztikai szoftverbe írható parancssorokhoz, az R által végrehajtott számítások eredményeinek értelmezéséhez, szövegbe foglalásához.

Kulcsszavak: fejlesztőprogram, hatásvizsgálat, statisztika

Diszciplínák: matematika, pszichológia, gyógypedagógia, pedagógia

Abstract

MATHEMATICAL STATISTICAL BACKGROUND OF
ONE SAMPLED LONGITUDINAL STUDY OF DEVELOPMENT PROGRAMMES

This paper offers a methodological guidance to mathematical statistical analysis of one sampled longitudinal studies of development programmes. After summarising the basic terms we: a) show a simple algorithm to choice adequate statistic tests, b) present R commands of these statistic tests (note: R is a downloadable free software), and c) give suggestions about how we can interpret the results of statistical calculations done by this software.

Keywords: development programmes, effectiveness study, statistics, R language

Disciplines: mathematics, psychology, special education, pedagogy

A különleges bánásmódot igénylő (sajátos nevelési igényű és/vagy beilleszkedési, tanulási, magatartási zavaros és/vagy tehetséges) tanulókra fókuszáló fejlesztőprogramok hatásvizsgálata (kontrollcsoport alkalmazása nélkül is) megvalósulhat:

- a fejlesztési tervben foglalt adott kritériumszinthez történő viszonyítással (Mező és tsai, 2015a);
- elővizsgálat-fejlesztés-utóvizsgálat elrendezésben (Mező és tsai, 2015b);
- nyomon követéses módon, longitudinális vizsgálat keretében (1. ábra) – jelen tanulmányban erre az esetre fókuszálunk a továbbiakban.

1. ábra: különbségvizsgálaton alapuló egymintás (vagy: összetartozó mintás) nyomon követéses hatásvizsgálat sémája: ugyanazzal a csoporttal ugyanazokat a vizsgálatokat vesszük fel különböző alkalmakkor (forrás: Mező F.)



Nyomon követéses hatásvizsgálattal találkozunk például olyankor, amikor az alábbiakhoz hasonló kérdésekre kell matematikai statisztikai eljárásokkal bizonyított választ adnunk:

- Vizsgálatunkban arra voltunk kíváncsiak, hogy az „IPOO – minimum program” (Mező, 2011, Dávid, 2006) tanulás módszertani tréning előtt és után, illetve egy évvel a tréning után mért tanulás módszertani teljesítményben (a Jupiterbolha-próba összpontszámában) jelentős különbség tapasztalható-e? Különösen: kedvező (itt: teljesítményt növelő) hatása volt-e a tréningnek, vagy sem?
- Vizsgálatunkban arra voltunk kíváncsiak, hogy tanulóink az „IPOO-minimum program” előtti osztályzataikhoz képest lényegesen jobb osztályzatokat kapnak-e a tréning után közvetlenül, illetve egy év múlva, olyan tananyagokból adott feleleteikre, amelyekkel kapcsolatban vélhetően minimális előzetes tudásuk volt, s saját maguknak kellett a tananyagot feldolgozni?
- Egy fejlesztőprogramban három éven keresztül kísérjük figyelemmel a ló-asszisztált foglalkozásokra (Schéder, 2015) koncentráltó táborba jelentkezők számát. A diákok például 2014-ben kerültek be a programba, s ekkor regisztráltuk először, hogy jelentkeztek-e a táborba vagy sem. 2015-ben és 2016-ban is regisztráltuk a jelentkezőket. Azt vizsgáltuk, a jelentkezések gyakoriságai mutatnak-e változást a három év folyamán?

E kérdések közös vonása: egy csoport kettőnél több vizsgálati alkalommal nyert eredményét hasonlítjuk össze (nyomon követéses vizsgálatot végzünk). Az alábbiakban egyrészt a különböző alkalmakkal végzett mérések eredményeinek összehasonlításához szükséges matematikai statisztikai próbák kiválasztásának algoritmusára fókuszálunk, másrészt e próbák elvégzését segítő, ingyenesen letölthető szoftver (az R statisztikai programcsomag) idevágó parancsaira és a szoftver outputjaként megjelenő statisztikai eredmények interpretációjára nyújtunk példákat. A tanulmány olvasásához szükséges minimális matematikai statisztikai

ismeretek (pl. minta, hipotézis, szignifikancia, eloszlás, nominális, ordinális, kvantitatív változók) összefoglalója megtalálható Máth és tsai. (2015) művében. A tudományos és statisztikai következtetés alapjaival kapcsolatban Dienes (2013) ad átfogó elemzést. Az R-nyelvvel kapcsolatos összefoglalót pedig Abari és tsai. (2015) és Solymosi (2005) közölnek.

E módszertani útmutatóban az R parancsokkal kapcsolatban a következő jelöléseket vezetjük be (vö.: Abari, 2008) az adatbázisok (lásd: 2. ábra) kapcsán:

2. ábra: széles és hosszú szerkezetű adattáblák (forrás: a Szerzők)

<i>d</i> adattábla (széles szerkezet):	<i>d.h</i> adattábla (hosszú szerkezet):
<pre> SZEMELY X1 X2 X3 1 sz.1 1 11 21 2 sz.2 2 12 22 3 sz.3 3 13 23 stb... 10 sz.10 10 20 30 </pre>	<pre> SZEMELY ALKALOM X 1 sz.1 X1 1 2 sz.2 X1 2 3 sz.3 X1 3 stb... 10 sz.10 X1 10 11 sz.1 X2 11 12 sz.2 X2 12 13 sz.3 X2 13 stb... 20 sz.10 X2 20 21 sz.1 X3 21 22 sz.2 X3 22 23 sz.3 X3 23 stb... 30 sz.10 X3 30 </pre>
<p>Az adatokat általában a <i>d</i> adattábla formában szokás rögzíteni, ám olykor szükséges lehet a <i>d.h</i> adattábla formátum alkalmazása.</p> <p>A <i>d</i> adattáblából a <i>d.h</i> adattáblába történő átalakítást az R-ben a <code>reshape2</code> csomag <code>melt()</code> függvényével lehet megoldani:</p>	
<pre> library(reshape2) d.h <- melt(d, variable.name = "ALKALOM", value.name = "X") </pre>	

- *d*: egy *d* nevű 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” stb. változók) adatait tartalmazzák. Az ilyen szerkezetű adattáblákat szélesnek nevezzük. Példa: 2. ábra *d* adatbázisa, melynek oszlopai a vizsgált változók (*X1*, *X2*, *X3* stb. mint mondjuk az 1., 2., 3. stb. osztályos korban mért eredmények) adatait tartalmazzák, a *SZEMELY* oszlopban pedig a vizsgált személyek azonosítói jelennek meg. Az azonos személyek különböző mérési időpontok közötti összetartozását *X1*, *X2*, *X3* stb. nevekkel jelöljük, így ezek a változók az R parancsokban *d\$X1*, *d\$X2*, illetve *d\$X3* módon jelennek meg (pl. *X1* = első osztályos korban mért teljesítmény, *X2* = második osztályos korban mért teljesítmény, *X3* = harmadik osztályos korban mért teljesítmény; a *d\$X1* jelentése pedig: *d* adatbázis *X1* változójára hivatkozunk az R parancsban, stb.).

- *d.h.*: néhány statisztikai próba (például a Bartlett-próba vagy az összetartozó mintás egyszempontos varianciaanalízis) esetében a *d* adatbázist át kell alakítanunk jelen tanulmányban *d.h.*-nak nevezett ún. hosszú formátumba, annak érdekében, hogy az egy személyhez tartozó *n* számú mérés *n* számú külön sorban szerepeljen (lásd: 2. ábra *d.h* adatbázisát). A *d.h* adatbázis formátumában az az elv érvényesül, hogy „az azonos változóhoz tartozó adatértékek kerüljenek azonos oszlopba”. Az R-ben a `reshape2` csomag `melt()` függvényével van lehetőség arra, hogy a *d* adattábla szerkezetét átranzszformáljuk át a *d.h* adattábla hosszú struktúrájába.

Statisztikaválasztás és végrehajtás az R-nyelvből

Az egymintás nyomon követéses jellegű, s a különböző időpontokban gyűjtött adatok különbségeire fókuszáló hatásvizsgálatokhoz szükséges matematikai statisztikai próba kiválasztásának menete röviden a következő (vö.: 3. ábra):

1. Döntenünk kell arról, hogy nominális, ordinális, vagy kvantitatív változókkal kell-e számolnunk (vö.: Máth és tsai, 2015).
2. Kvantitatív változók esetében az R-nyelv segítségével döntjük el, hogy: a) normális eloszlásúnak tekinthető (harang alakú görbét mutató) változókról van-e szó; b) a változók szórása közel egyformának tekinthető-e.
3. Kizárólag e két lépés végrehajtását követően kerülhet megválasztásra, végrehajtásra a megfelelő statisztikai próba.
4. Időnként azonban további elemzésekre is szükség lehet.

A továbbiakban a változók kvantitatív, ordinális, nominális jellege alapján tekintjük át a szükséges R parancsokat, s szignifikancia szinttől függő interpretációjukat.

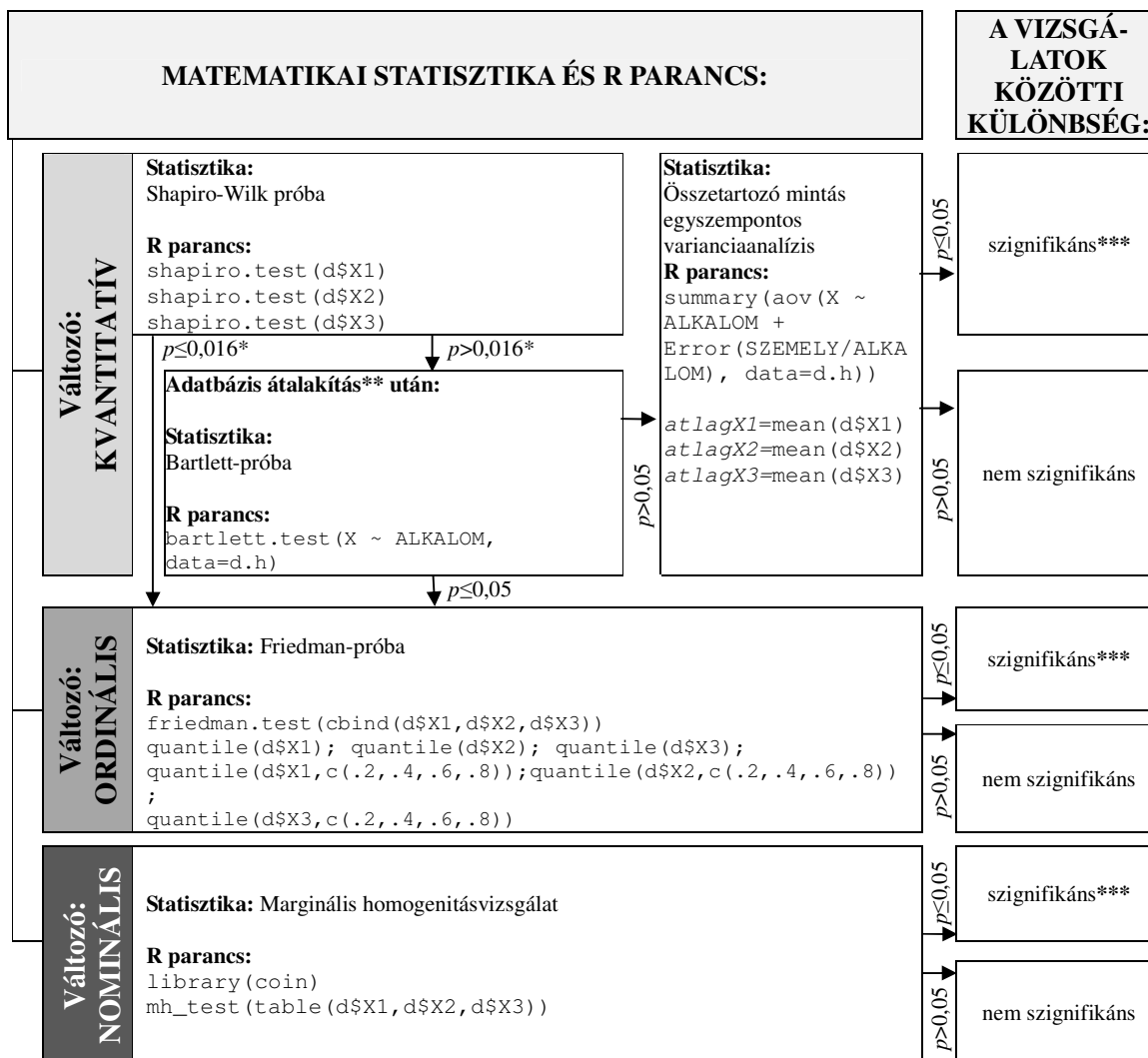
Kvantitatív változók a nyomon követéses vizsgálatokban

A nyomon követéses vizsgálatokban szereplő kvantitatív változók esetében ellenőriznünk kell, hogy esetükben teljesül-e a normális eloszlás, illetve a szóráshomogenitás feltétele. Amennyiben e feltételek teljesülnek, akkor valóban építhetünk e változók kvantitatív jellegére, s (összetartozó mintás egyszempontos) varianciaanalízissel elemezhetjük adatainkat. Ha e két feltételnek csak egyike teljesül vagy egyike sem, akkor változóinkkal legfeljebb az ordinális változókkal végezhető matematikai statisztikai próba (esetünkben például a Friedman-próba) hajtható végre (lásd: jelen tanulmány következő alfejezetét). E két feltétel ellenőrzésének, illetve a varianciaanalízis alkalmazásának módja dióhéjban a következő:

Az *eloszlásvizsgálat* során arra a kérdésre kell választ kapnunk hogy „Az ugyanahhoz a dologhoz/személyhez tartozó, több (legalább 3) különböző időpontban/helyzetben mért *X1* (például a fejlesztés elején mért teljesítmény), *X2* (mondjuk: a fejlesztés végén mért teljesítmény), *X3* (például fejlesztés után egy évvel mért teljesítmény) *stb.* kvantitatív (függő) változók normális eloszlásúnak tekinthetők-e?”. E kérdés megválaszolására alkalmas egyik statisztikai próba a Shapiro-Wilk-próba, melynek R-parancsa a következő:

```
shapiro.test(d$X1)
shapiro.test(d$X2)
shapiro.test(d$X3)
```

3. ábra: Statisztikaválasztás egymintás, nyomon követéses hatásvizsgálat esetében ($d\$X1$ = első vizsgálat, $d\$X2$ = második vizsgálat, $d\$X3$ = harmadik vizsgálat adatait tartalmazó változók a „d” adatbázisban). Forrás: Mező, Máth és Abari (2008) alapján Mező F.



* Három változó esetén a 0,05-öt hárommal osztjuk, és ehhez az értékhez viszonyítjuk a p -értéket: $0,05/3=0,016$.

Értelemszerűen: n darab változó (itt: vizsgálati alkalom) esetén a p -értéket a $0,05/n$ értékhez viszonyítjuk.

**Adatbázis átalakítás R-parancsai:

```
library(reshape2)
d.h <- melt(d, variable.name = "ALKALOM", value.name = "X")
```

***Utóvizsgálatokat, páronkénti összehasonlítással célszerű folytatni az adatelemzést (lásd: Mező és tsai, 2015b)

Mindez például így foglalható össze egy tanulmányban: „Vizsgálatunkban arra voltunk kíváncsiak, hogy az „IPOO – minimum program” tanulás módszertani tréning előtt és után, illetve egy évvel a tréning után mért tanulás módszertani teljesítményben (a Jupiterbolha-próba összpontszámában) jelentős különbség tapasztalható-e? Különösen: kedvező (itt: teljesítményt növelő) hatása volt-e a tréningnek, vagy sem? Az adekvát statisztikai próba kiválasztásához mindenekelőtt ellenőriznünk kellett azt, hogy a három mérési alkalom adatait tartalmazó változóink normális eloszlásúak-e. Ezt a Shapiro-Wilk próbával ellenőriztük, amelynek eredménye szerint...”

A Shapiro-Wilk-próba eredményei és az R által megadott szignifikancia szint (p -érték) alapján döntünk a további teendőinkről (fontos: a statisztikaválasztó döntésünket befolyásoló kritériumértéknek $0,05/n$ értéket válasszunk, ahol n = a vizsgált változók száma. Példánkban 3 mérési alkalom adatait tartalmazó 3 változó eloszlását vizsgáljuk, így kritériumszintünk $0,05/3=0,016$ lesz):

- a) Ha a Shapiro-Wilk próba legalább az egyik változó esetében $p\text{-value} \leq 0,016$ értéket ad, akkor a változók között van nem normális eloszlású változó, és a fent megkezdett mondat pedig így folytatódhat: „...nem teljesül, hogy mindegyik változó normális eloszlású (az első vizsgálat esetén $p = p\text{-value}$; a második vizsgálat esetén $p = p\text{-value}$; a harmadik vizsgálat esetén $p = p\text{-value}$).” A „ $p\text{-value}$ ” helyére az R által megadott p -értéket kell beírni. Mindez azzal a következménnyel is jár, hogy adataink, változóink nem alkalmasak varianciaanalízis végrehajtására, s az ordinális változók esetében alkalmazható statisztikai próbákhoz kell folyamodnunk (lásd e tanulmány következő alfejezetét).
- b) Ha a Shapiro-Wilk próba mindhárom változó esetében $p > 0,016$, akkor a változók eloszlása normálisnak tekinthető, s imént félbe hagyott mondatunk így fejezhető be: „...($p=p\text{-value}$) mindhárom változót normális eloszlásúnak tekinthetjük (az első vizsgálat esetén $p=p\text{-value}$; a második vizsgálat esetén $p=p\text{-value}$; a harmadik vizsgálat esetén $p=p\text{-value}$). A továbbiakban meggyőződünk arról is, hogy a különböző mérések alkalmával a szórások közel egyformák-e.” A „ $p\text{-value}$ ” helyére az R által megadott p -értéket kell beírni. Ebben az esetben az adatelemző munka következő lépéseként a szóráshomogenitás vizsgálata következhet.

A szóráshomogenitás azt jelenti, hogy az $X1$, $X2$ és $X3$ változók szórása közel egyforma. A szóráshomogenitás vizsgálatára alkalmas egyik statisztikai próba a Bartlett-próba, amely esetében az R az úgynevezett hosszú szerkezetű (a további példáinkban *d.h* néven hivatkozott) adattáblát kívánja meg (ennek az adattáblának a szerkezetét és előállításának R-parancsát a 2. ábra foglalja össze). A Bartlett-próba példánkhoz tartozó R parancsa a következő:

```
bartlett.test(X ~ ALKALOM, data=d.h)
```

A Bartlett-próba esetében a 0,05-ös szignifikancia szinthez viszonyítjuk a p -értékeket, melyek az alábbiak lehetnek:

- a) Ha a Bartlett-próba eredménye szerint $p\text{-value} \leq 0,05$, akkor változóink szórása nem tekinthető közel sem egyformának. Ezt így foglalhatjuk szövegbe például: „A szóráshomogenitás vizsgálatára Bartlett-próbát alkalmaztunk – ennek eredménye szerint ($p=p\text{-value}$), a szóráshomogenitás feltétele nem teljesül.” Ez azzal jár, hogy adataink nem alkalmasak a varianciaanalízissel történő további elemzésre, s ehelyett inkább az ordinális változók esetében használható statisztikai próbák jöhetnek szóba (lásd: az ordinális változókkal kapcsolatos következő alfejezetet!).
- b) Ha a Bartlett-próba eredménye szerint $p\text{-value} > 0,05$, akkor a szórások közel egyformának tekinthetők, s ezt a következőképpen interpretálhatjuk: „A szóráshomogenitás vizsgálatára Bartlett-próbát alkalmaztunk – ennek eredménye szerint ($p=p\text{-value}$), a szóráshomogenitás feltétele teljesül.” Tekintettel arra, hogy ezt a próbát csak akkor végeztük el, ha a normalitásra vonatkozó feltétel is teljesült, ez az eredmény azzal jár, hogy adataink alkalmasak a varianciaanalízissel történő további adatelemzésre.

Összetartozó mintás egyszempontos varianciaanalízis alkalmával példánkban most lényegében azt ellenőrizzük, hogy az $X1$, $X2$, $X3$ változók populáció átlagai egyenlők-e. A szükséges R-parancsok:

```
summary(aov(X ~ ALKALOM + Error(SZEMELY/ALKALOM), data=d.h))
atlagX1=mean(d$X1); atlagX2=mean(d$X2); atlagX3=mean(d$X3)
```

Lehetséges eredmények:

- a) Ha az összetartozó mintás egyszempontos varianciaanalízis eredménye szerint $p\text{-value} \leq 0,05$, akkor az $X1$, $X2$, $X3$ változók populáció átlagai eltérők (különbség van a mérések között). Szövegbe foglalva: „Ennek ismeretében összetartozó mintás egyszempontos varianciaanalízist végeztünk, amelynek eredménye szerint: különbség van a különböző időpontokban mért teljesítmények között ($p=p\text{-value}$). A szignifikánsan eltérő átlagok a következők: *atlagX1*, *atlagX2*, *atlagX3*.” A változás növekvő vagy csökkenő jellegét a változók átlagai mutatják meg. Statisztikailag egzaktabb vizsgálat elvégzéséhez lásd (Vargha, 2000; Falus és Ollé, 2000; Reiczigel, Harnos és Solymosi, 2007).
- b) Ha az összetartozó mintás egyszempontos varianciaanalízis eredménye $p\text{-value} > 0,05$ értéket ad, akkor az $X1$, $X2$, $X3$ változó populáció átlagai egyenlők (másképp: nem különböznek egymástól lényegesen). Lehetséges interpretáció: „Ennek ismeretében összetartozó mintás egyszempontos varianciaanalízist végeztünk, amelynek eredménye szerint: nincs különbség a különböző időpontokban mért teljesítmények között ($p=p\text{-value}$)”

Ordinális változók a nyomon követéses vizsgálatokban

Az ordinális változókkal folytatható vizsgálatokra akkor lehet szükségünk, ha vizsgált változóink: a) eleve ordinális skálájúak (például ötfokú skálán adott osztályzatok); b) kvantitatív skálájúak ugyan, ám nem felelnek meg a normalitás és a szóráshomogenitás fentebb közölt feltételeinek. Általános értelemben véve a következő kérdésre keressük a választ a nyomon követéses vizsgálat ilyen eseteiben: „Az ugyanahhoz a dologhoz/személyhez tartozó X legalább ordinális (függő) változó értéke változik-e különböző időpontokban/helyzetekben (alkalom = független változó) történő mérésekkor?”. Ide vonatkozó példánk legyen a következő: „Vizsgálatunkban arra voltunk kíváncsiak, hogy tanulóink az „IPOO-minimum program” előtti osztályzataikhoz képest lényegesen jobb osztályzatokat kapnak-e a tanulás módszertani tréning után közvetlenül, illetve egy év múlva, olyan tananyagokból adott feleleteikre, amelyekkel kapcsolatban vélhetően minimális előzetes tudásuk volt, s saját maguknak kellett a tananyagot feldolgozni?”

Az adatelemzéshez használható itt tárgyalt statisztikai próba a Friedman-próba, mely arra a kérdésre ad választ, hogy az $X1$, $X2$, $X3$ változók populáció mediánjai egyenlők-e. A Friedman-próba R-parancsa esetünkben:

```
friedman.test(cbind(d$X1, d$X2, d$X3))
```

A Friedman-próba lehetséges eredményei az alábbiak lehetnek:

- a) Ha a Friedman-próba eredménye szerint $p\text{-value} > 0,05$, akkor az $X1$, $X2$, $X3$ változók populáció mediánjai egyenlők. Vagyis nyomon követéses vizsgálatunkról szóló prezentációnkban megfogalmazhatjuk, hogy: „A Friedman-próba eredménye szerint nincs szignifikáns különbség ($p=p\text{-value}$) a különböző alkalmakkal felvett tesztek eredménye között.”
- b) Ha a Friedman-próba eredményeként $p\text{-value} \leq 0,05$ értéket kapunk, akkor az arra utal, hogy az $X1$, $X2$, $X3$ változók populáció mediánjai nem egyenlők. Ezt az eredményt például ekképpen foglalhatjuk szavakba: „A Friedman-próba eredménye szerint

szignifikáns különbség ($p=p\text{-value}$) van a különböző alkalmakkal felvett tesztek eredménye között.” Ismertessük és értelmezzük a különbséget! Az értelmezésben segíthet, ha leíró statisztikákat kérünk az $X1$, $X2$, $X3$ változókról. Kérhetünk kvartiliseket (melyek a mintát negyedelik), de kérhetünk részletesebb ábrázolást is. A kvartiliseket vagy a mintát ötödölő értékeket (kvartiliseket) a következő R parancsokkal kaphatjuk meg:

```
quantile(d$X1)
quantile(d$X2)
quantile(d$X3);
quantile(d$X1,c(.2,.4,.6,.8))
quantile(d$X2,c(.2,.4,.6,.8))
quantile(d$X3,c(.2,.4,.6,.8))
```

Nominális változók a nyomon követéses vizsgálatokban

A nyomon követéses vizsgálatokban a nominális (tehát az egymástól megkülönböztethető, de nem sorbarendezhető értékekkel rendelkező) változók esetében feltehető általános jellegű kérdés lehet például: „Ugyanahhoz a személyekhez/dologhoz tartozó $X1$, $X2$, $X3$ nominális változó megfigyelt gyakoriságai eltérnek-e egymástól, azaz van-e változás?”.

Tegyük fel például, hogy jelen tanulmány elején felvetett ló-asszisztált foglalkozásokat (Schéder, 2015) tartalmazó nyári táborba jelentkezők számának alakulását kívánjuk nyomon követni 2014-2016. között, három éven keresztül. Jelentse $X1$ = a nyári táborba történő jelentkezések 2014. évi számát (1=jelentkezett; 2=nem jelentkezett), $X2$ = a nyári táborba történő jelentkezések 2015. évi számát (1=jelentkezett; 2=nem jelentkezett), $X3$ = a nyári táborba történő jelentkezések 2016. évi számát (1=jelentkezett; 2=nem jelentkezett). Ezt megfogalmazhatjuk így is: „Egy ló-asszisztált foglalkozásokra építő nyári tábor jellegű fejlesztő programban három éven keresztül kísértük figyelemmel a táborba jelentkezők számát. A diákok 2014-ben kerültek be a programba, s ekkor regisztráltuk először, hogy jelentkeztek-e a táborba vagy sem. 2015-ben és 2016-ban is regisztráltuk a jelentkezőket. Azt vizsgáltuk, hogy a jelentkezések gyakoriságai mutatnak-e változást a három év folyamán?”.

Az $X1$, $X2$ és $X3$ nominális változók gyakoriságának hasonlóságára vonatkozó kérdés megválaszolására alkalmas statisztikai próba lehet a marginális homogenitásvizsgálat, melynek R parancsa (mivel e próba végrehajtásához szükséges függvény az R programon belül a `coin` csomagban található, az `mh_test()` függvény hívása előtt ki kell adnunk a `library(coin)` parancsot):

```
library(coin)
mh_test(table(d$X1, d$X2, d$X3))
```

Ha a marginális homogenitásvizsgálat eredménye $p\text{-value} > 0,05$, az arra utal, hogy az $X1$, $X2$, $X3$ változók gyakoriságai hasonlóak. Ennek egy lehetséges megfogalmazása az adott példa kontextusában: „Marginális homogenitás teszttel vetettük össze a három időpontban megállapított jelentkezési gyakoriságokat és lényeges eltérést nem találtunk ($p=p\text{-value}$).”

Amennyiben a marginális homogenitásvizsgálat $p\text{-value} \leq 0,05$ szinten szignifikáns különbséget jelez $X1$, $X2$, $X3$ változók gyakoriságai között, akkor egyrészt szövegbe foglalva ilyesmit állapíthatunk meg: „Marginális homogenitás teszttel vetettük össze a három időpontban megállapított jelentkezési gyakoriságokat és lényeges eltérést találtunk ($p=p\text{-value}$).”

value); máriszt közöljük mindhárom mérés esetén a változók gyakoriságait a `table(d$X1)`, `table(d$X2)` és `table(d$X3)` parancsokkal hogy a változás lényegét demonstrálhassuk és értelmezhesük!

Kategorikus változók elemzésével kapcsolatban lásd még: Máth (2004).

Zárógondolatok

Jelen tanulmányban a fejlesztőprogramok hatásvizsgálatának arra a speciális esetére fókuszáltunk, amelyben egy csoport három vagy több alkalommal azonos módon/eszközzel gyűjtött adatait vetjük össze egymással. Ezzel kapcsolatban öt matematikai statisztikai próba és R-nyelvbeli parancsaikat tekintettük át:

- Shapiro-Wilk-próbát alkalmaztunk a három (vagy több) vizsgálat során gyűjtött adatok normális eloszlásának vizsgálata céljából.
- Bartlett-próbával teszteltük a három (vagy több) vizsgálati alkalom során gyűjtött adatok szóráshomogenitását.
- Összetartozó mintás egyszempontos varianciaanalízist alkalmaztunk a normális eloszlás és a szóráshomogenitás feltételeinek megfelelő kvantitatív változók elemzése érdekében.
- Friedman-próbával teszteltük a három (vagy több) alkalommal gyűjtött adatok esetében a mediánok hasonlóságát.
- Marginális homogenitásvizsgálatot használtunk a nominális változók gyakoriságainak hasonlóságának elemzése céljából.

Fontos tudni, hogy más statisztikai próbák, módszerek (lásd: Varga, 2000) és vizsgálati elrendezések (Eccles és tsai, 2003) is léteznek hasonló kérdések eldöntésére.

A nyomon követéses, longitudinális vizsgálatok eredményeinek értelmezésekor (s a hatásvizsgálat korlátai között) célszerű figyelembe venni olyan lehetséges torzító tényezőket, mint a sorozathatások és az idői hatások (Szokolszky, 2004). A sorozathatások eredményei lehetnek például a gyakorlatlanságból eredő kezdeti relatív gyenge kiinduló eredmények, vagy az első vizsgálat tapasztalatai alapján kialakult válaszadói beállítódások. Az idői hatások megnyilvánulhatnak többek között az adatgyűjtések közötti időszak ellenőrizetlen történéseiben (például a vizsgált fejlesztőprogram mellett öt másik hasonló célú fejlesztőprogramban történő részvétel esetében), a spontán érés hatásában (nem a fejlesztőprogram eredménye a különböző alkalmakkal mért teljesítménybeli változás, hanem a belső érési folyamatoké), vagy a vizsgálati személyek időközben történő (kilépés, utolérhetetlenség, esetleg elhalálozás miatt történő) lemorzsolódásában.

A fejlesztőprogram hatását a fentiekén túlmenően jellemezheti az is, hogy a fejlesztendő terület különböző kezdeti jellemzőivel bíró személyek (például az eleve gyengébb vagy jobb képességűek – például a vizsgálati csoporto alsó és felső 25%-ába sorolható személyek) eltérő módon reagálnak-e a programra. Ebből a szempontból a csoportok közötti különbségvizsgálatok (Mező, Máth és Abari, 2008) mellett kapcsolatvizsgálatokat (például korrelációs számítást) is célszerű végezni a kezdeti értékek és a fejlődés mértéke (= a fejlesztés végén tapasztalt értékekből kivont kezdeti értékek) között. Ilyen esetben például az erős pozitív korreláció arra utalhat, hogy a kezdetben eleve magasabb pontszámmal jelezhető személyek esélyesebbek a (pontszámokban kimutatható) nagyobb fejlődésre. Egy erős negatív korreláció pedig azt jelezheti, hogy a kezdeti mérés során gyenge eredményt elérők esetében valószínűbb a nagyobb értékű pontszámnövekedés.

A fejlesztőprogramok hatásvizsgálatába mindössze egyetlen csoportot bevonó esetek – lásd: Mező és tsai (2015a,b), illetve jelen tanulmányt – mellett léteznek két vagy több csoportot egy vagy több alkalommal vizsgáló esetek is, melyeket a Különleges Bánásmód folyóirat további számaiban foglalkozunk össze.

Irodalom

- Abari K. (2008). A tehetségdiagnosztika adatkezelésbeli alapjai R környezetben. In Mező F. (szerk.). *Tehetségdiagnosztika*. Kocka Kör & Faculty of Central European Studies, Constantine the Philosopher University in Nitra, Debrecen. pp 105-130.
- Abari K., Mező F., Mező K. és Máth J. (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
- Dávid M. (2006): A tanulási kompetencia fejlesztése - elméleti háttér. *Alkalmazott pszichológia*, 2006. VIII. évfolyam, 1. szám, 51-64.p
- Dienes Z. (2013). *Mitől tudomány a pszichológia? A tudományos és statisztikai következtetés alapjai*. Akadémiai Kiadó, Budapest.
- Eccles, M., Grimshaw, J., Cambell, M. és Ramsay, C. (2003). Research designs for studies evaluating the effectiveness of change and improvement strategies. *Quality and Safety in Health Care*, 2003, 12, 47-52 DOI: 10.1136/qhc.12.1.47
- Falus I. és Ollé J. (2000). *Statisztikai módszerek pedagógusok számára*. Okker Kiadói Kft., Budapest.
- Máth J. (2004). Kategórikus változók elemzése (loglineáris modell látens változókkal). *Alkalmazott Pszichológia*, VI/1, pp 57-81. Letöltés: 2015.09.14. Web: http://psycho.unideb.hu/archivum/kategorikus_valtozok.pdf
- Máth J., Mező F., Abari K. és Mező K. (2015). Fejlesztőprogramok hatásvizsgálatának matematikai statisztikai alapfogalmai. *Különleges Bánásmód*, I. évf. 2015/1. szám, 69-77. DOI 10.18458/KB.2015.1.69
- Mező F. (2011). *Tanulás: diagnosztika és fejlesztés az IPOO-modell alapján*. K+F Stúdió Kft., Debrecen.
- 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.
- Mező F., Máth J., Abari K. és Mező K. (2015a). Fejlesztőprogramok egymintás, kritériumorientált hatásvizsgálatának matematikai statisztikai háttere. *Különleges Bánásmód*, I. évf. 2015/3. szám, 69-78. DOI 10.18458/KB.2015.3.69
- Mező F., Máth J., Abari K. és Mező K. (2015b): Fejlesztőprogramok önkontrollos hatásvizsgálatának matematikai statisztikai háttere. *Különleges Bánásmód*, I. évf. 2015/4. szám, 65-75. DOI 10.18458/KB.2015.4.65
- Reiczigel J., Hamos A. és Solymosi N. (2007). *Biostatisztika nem statisztikusoknak*. Pars Kft., Nagykovácsi.
- Schéder V. (2015). Különleges bánásmód ló-asszisztált foglalkozásokon. *Különleges Bánásmód*, I. évf. 2015/1. szám, 79-89. DOI 10.18458/KB.2015.1.79
- Solymosi N. (2005). *R<-...erre, erre...! Internetes R-jegyzet*. Letöltés: 2015.09.14. Web: <http://cran.r-project.org/doc/contrib/Solymosi-Rjegyzet.pdf>
- Szokolszky Á. (2004). *Kutatómunka a pszichológiában*. Osiris Kiadó, Budapest.
- Vargha A. (2000). *Matematikai statisztika pszichológiai, nyelvészeti és biológiai alkalmazásokkal*. Pólya Kiadó, Budapest.