VIZSGÁLATI ÉS KONTROLLCSOPORTOT EGY ADATFELVÉTELI ALKALOM ALAPJÁN ÖSSZEHASONLÍTÓ HATÁSVIZSGÁLATOK 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

Koncz István
Professzorok az Európai Magyarországért

Szilágyi Barnabás
Debreceni Egyetem

Nemes Magdolna
Debreceni Egyetem

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

Mező F., Máth J., Abari K., Mező K. (2016): Vizsgálati és kontrollcsoportot egy adatfelvételi alkalom alapján összehasonlító hatásvizsgálatok matematikai statisztikai háttere. *Különleges Bánásmód*, II. évf., 2016/2. szám, 73-82. DOI 10.18458/KB.2016.2.73

Absztrakt

A fejlesztő programok hatásvizsgálatának egyik fajtája egy vizsgálati csoport és egy kontrollcsoport eredményeit hasonlítja össze (amikor az adatgyűjtés csak egyetlen alkalommal történt meg). E tanulmány módszertani útmutatót kínál nyújtani az ilyen vizsgálatok adekvát matematikai statisztikai elemzésének kiválasztásához, alkalmazásához és értelmezéséhez. Az R statisztikai programcsomag (egy ingyenesen letölthető szoftver) ajánlott a statisztikai eredmények kiszámolásához.

Kulcsszavak: fejlesztőprogram, hatásvizsgálat, statisztika, R
Diszciplinák: matematika, pszichológia, gyógypedagógia, pedagógia

Abstract

MATHEMATICAL STATISTIC BACKGROUND OF EFFECTIVENES STUDIES WHICH COLLECT DATA FROM AN EXPERIMENTAL AND A CONTROLL GROUP ONLY ONCE

One kind of effectiveness studies of develpmental programmes compares results of an experimental group and a control group (when the data collecting happened only once). This paper offers a methodological guidance to choose, to apply and to interpret of adequate mathematical statistic analysis of these studies. Using R environment (this is a downable free software) is recommended to compute statistical results.

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 hatásáról különbségvizsgálatokon keresztül tájékozódhatunk. Az ilyenkor szükséges adatelemzéshez megfelelő matematikai statisztikai próba kiválasztása többek között attól függ, hogy hány csoportot (egyet, kettőt vagy többet) vonunk be a hatásvizsgálatba, s hány (egy, két vagy több) alkalommal felvett adatokat hasonlítunk össze (1. táblázat).

1. táblázat: A legalapvetőbb különbségvizsgálatokra vonatkozó példák. Forrás: a Szerzők.

Vizsgált csoportok száma	**Egy vizsgálat:**	**Két vizsgálat: elő- és utóvizsgálat**	**Több vizsgálat: nyomon követéses vizsgálat**
1	Pl.: egy fejlesztőprogramba bevont csoport 2015-ben nyújtott teljesítménye különbözik-e egy meghatározott (pl. fejlesztési tervben kitűzött) célértéktől? Lásd: Mező és tsai (2015a)	Pl.: a fejlesztőprogramba bevont csoport program előtti (pl. 2005-ös) és utáni (pl. 2015-ös) teljesítménye különbözik-e egymástól? Lásd: Mező és tsai (2015b)	Pl.: a fejlesztett csoport 2005-ben, 2010-ben, 2015-ben stb. nyújtott teljesítménye különbözik-e egymástól? Lásd: Mező és tsai (2016)
2	Pl.: a fejlesztett (vizsgálati) csoport 2015-ös teljesítménye különbözik egy fejlesztésbe be nem vont (kontroll-)csoport teljesítményétől? Lásd: jelen tanulmány	Pl.: a fejlesztőprogramba bevont csoport és a kontrollcsoport teljesítménye különbözik-e önmagától, illetve egymástól program előtt (pl. 2005-ben) és után (pl. 2015-ben) ill. az esetleges változás mértéke eltér-e a két csoportban?	Pl.: a fejlesztőprogramba bevont csoport és a kontrollcsoport teljesítménye különbözik-e önmagától, illetve egymástól a program előtt (pl. 2005-ben) és után (pl. 2010-ben), s a további években (2015-ben stb.) ill. az esetleges változás görbéje eltér-e a két csoportban?
3 vagy több	Pl.: a különböző fejlesztésben részesülő csoportok teljesítménye különböző 2015-ben?	Pl.: a csoportok teljesítménye különbözik-e a saját, illetve a többi csoport teljesítményétől a program előtt (pl. 2005-ben) és után (pl. 2015-ben) ill. az esetleges változás mértéke eltér-e az egyes csoportokban?	Pl.: a csoportok teljesítménye különbözik-e a saját, illetve a többi csoport teljesítményétől a program előtt (pl. 2005-ben) és után (pl. 2010-ben), s a további években (2015-ben stb.) ill. az esetleges változás görbéje eltér-e az egyes csoportokban?

Az 1. táblázatban is látható, egyetlen vizsgálati csoporttal kapcsolatos eseteket a *Különleges Bánásmód folyóirat* korábbi számaiban már megtárgyaltuk (Mező és tsai, 2015a, b; 2016), miként a jelen tanulmány értő olvasásához is szükséges matematikai statisztikai alapfogalmakat is (Máth és tsai, 2015 – lásd még: Mező, Máth és Abari, 2008; Falus és Ollé, 2000; Vargha (2000), s az R-nyelv (ingyenesen letölthető statisztikai szoftver) alapjait is (Abari és tsai, 2015). Az R-nyelvvel kapcsolatban lásd még: Solymosi (2005), Abari (2008). A tudományos interpretációval, módszertannal kapcsolatban lásd: Dienes (2013), Szokolszky (2004), Eccles és tsai (2003).

Jelen tanulmányban arra a vizsgálati elrendezésre fókuszálunk, amikor a hatásvizsgálat a vizsgálati csoport mellett kontrollcsoport bevonásával valósul meg. A fejlesztőprogramok kontrollcsoportos hatásvizsgálatának legalapvetőbb esete pedig az, amikor egyetlen adatfelvételi alkalommal gyűjtünk információkat a vizsgálati csoportba és a kontrollcsoportba sorolt személyektől, s e csoportok teljesítményeit hasonlítjuk össze (lásd: 1. táblázatban a két vizsgált csoport és egy vizsgálat metszetében lévő cella tartalmát, illetve az 1. ábrát!).

1. ábra: vizsgálati és kontrollcsoport egyszeri adatfelvétel alapján történő összehasonlításának sémája. Forrás: a Szerzők.

Az 1. ábrán látható vizsgálati elrendezés esetében az alábbiakhoz hasonló kérdésekre kell matematikai statisztikai eljárásokkal bizonyított választ adnunk:

a) Vajon az IPOO-minimum (kreatív) tanulásfejlesztő programba (Mező, 2011) bevont és a kontrollcsoportba sorolt tanulóknak a Jupiterbolha-próba tanulás módszertani tesztben elért teljesítménye különbözik?
b) A tehetséggondozó programba bevont vizsgálati csoport és a programba be nem vont kontrollcsoport tanulói témazáró dolgozatot írtak. A tehetséggondozó program hatékonyságát jelezné, ha a programban résztvevők lényegesen jobb érdemjegyeket kapnának, mint a kontrollcsoportba tartozó társaik.
c) A magyarországi magyar és a határon túli magyar, tehetséggondozó programba 2009-ben bevont diákokat vizsgáltunk. Arra voltunk kíváncsiak, hogy a magyarországi magyarok között hasonló-e a táborba jelentkezők aránya, mint a határon túli magyarok között?

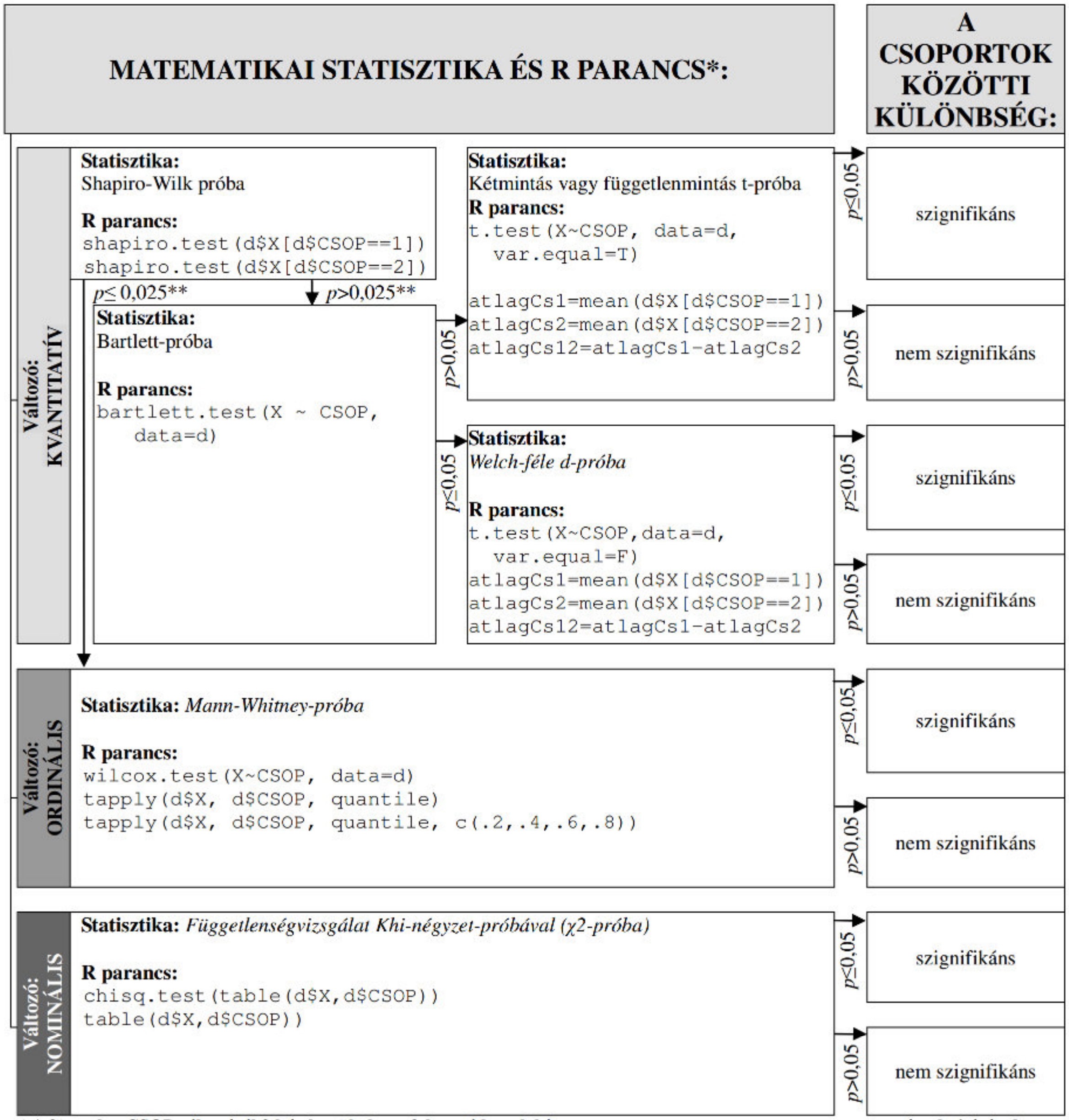
E kérdések közös vonása: két csoport egyetlen vizsgálati alkalommal nyert eredményét hasonlítjuk össze (kontrollcsoportos hatásvizsgálatot végzünk).

Statisztikaválasztás és végrehajtás az R-ben

Az egy adatfelvételi alkalommal járó kontrollcsoportos hatásvizsgálatokhoz szükséges matematikai statisztikai próba kiválasztásának menete röviden a következő:

1. Döntenünk kell arról, hogy nominális, ordinális, vagy kvantitatív változókkal kell-e számolnunk (vö.: Máth és tsai, 2015).
2. Kvantitatív változók esetében az R (Abari és tsai, 2015) segítségével döntjük el, hogy: a) normális eloszlásúnak tekinthető 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 (vö.: 2. ábra).
4. Időnként azonban további elemzésekre is szükség lehet.

2. ábra: Statisztikaválasztás egy adatfelvételi alkalommal járó kontrollcsoportos hatásvizsgálat esetében. Forrás: Mező, Máth és Abari (2008) alapján Mező F.



* A független CSOP változóról feltételezzük, hogy faktorrá lett alakítva a `d$CSOP <- factor(d$CSOP)` parancs végrehajtásával

** Két változó esetén a 0,05-öt kettővel osztjuk, és ehhez az értékhez viszonyítjuk a *p*-értéket: 0,05/2=0,025.

A továbbiakban a változók kvantitatív, ordinális vagy nominális jellege alapján tekintjük át a szükséges R parancsokat, s szignifikancia szinttől függő interpretációjukat. A példákban szereplő R-parancsok esetén a független CSOP változóról feltételezzük, hogy faktorrá lett alakítva a d$CSOP <- factor(d$CSOP) parancs végrehajtásával.

Kvantitatív változók esete

A konkrét példa kedvéért képzeljük el, hogy az alábbi kérdésre keressük a választ: „Vajon a tehetséggondozó programba bevont és a kontrollcsoportba sorolt tanulóknak a Jupiterbolha-próba tanulás módszertani tesztben elért összpontszámbeli teljesítménye különbözik?” A kérdés megválaszolásához szükséges adatokat tartalmazza egy *d* névre keresztelt adatbázis, ami legalább két változót tartalmaz. Az egyik változó legyen a *CSOP* névre keresztelt csoportosító változó (értékei: 1 = tehetséggondozó programba bevont diákok, 2 = kontrollcsoport), a másik változó legyen *X (ez tartalmazza* a Jupiterbolha-próba összpontszámra vonatkozó adatokat százalékban; értékei: 0-100). Az adatbázis tehát így néz ki:

```
CSOP    X
1       35
2       32
1       87
1       95
2       25
```

A két csoport kvantitaív változó mentén történő összehasonlításakor általánosságban véve tehát ilyen kérdés fogalmazható meg: „Az ugyanahhoz a dologhoz/személyhez tartozó *X* kvantitatív (függő) változó értéke különbözik-e (a *CSOP* független változó értékeiben kódolt) különböző csoportokban?”. A kérdés megválaszolásához szükséges statisztikai próba kiválasztásához a kvantitatív változók esetében ellenőriznünk kell, hogy normális eloszlásúak-e, s teljesül-e a szóráshomogenitás feltétele (itt: a két csoport szórása hasonló-e). E feltételek teljesülése setén *kétmintás vagy függetlenmintás t-próba* segítségével történhet a további adatelemzés. Ha a két feltételközül csak a normális eloszlásra vonatkozó teljesül, akkor *Welch-féle d-próba* alkalmazására kerülhet sor. Amennyiben a két feltétel közül egyik sem teljesül, akkor az ordinális változókkal végezhető statisztikai tesztekkel operálhatunk a továbbiakban (lásd: alább).

Az eloszlásvizsgálat alkalmával általános értelemben véve arra a kérdésre kell választ kapnunk, hogy: „Az *X*-változó mindkét csoportban normális eloszlású-e?”. A normális eloszlás vizsgálata megoldható például az almintákra alkalmazott Shapiro-Wilk próbával – ennek R-parancsa:

```
shapiro.test(d$X[d$CSOP==1])
shapiro.test(d$X[d$CSOP==2])
```

Publikációnkban mindezt így fogalmazhatjuk meg: „Az adekvát statisztikai próba kiválasztásához mindenekelőtt ellenőriznünk kellett azt, hogy a Jupiterbolha-próba összpontszáma mindkét csoportban normális eloszlásúnak tekinthető-e. Ezt a Shapiro-Wilk próbával ellenőriztük, amelynek eredménye szerint:…”

A Shapiro-Wilk próba eredménye a szignifikancia (*p*-érték) szempontjából kétféle lehet (megjegyzés: két csoport esetén a pedagógiai-pszichológiai vizsgálatokban általános p=0,05 értéket kettővel osztjuk, és ehhez az értékhez viszonyítjuk a vizsgálat eredményeként kapott *p*-értéket: 0,05/2=0,025):

a) Ha *p-value* ≤ 0,025, akkor ez arra utal, hogy *X* eloszlása legalább az egyik csoportban nem normális, s a továbbiakban az ordinális skálájú változóknál tárgyalt Mann-Whitney-féle U-próbát kell használnunk. A Shapiro-Wilk-próbával kapcsolatban fent megkezdett mondatot így folytathatjuk: „…nem teljesül, hogy e változó mindkét csoportban normális eloszlást követ (az első csoport esetén *p=p-value*; a második csoport esetén *p=p-value*). A továbbiakban ezért Mann-Whitney-próbát alkalmaztunk a két csoport teljesítményének összehasonlítására.”

b) Ha *p-value* > 0,025, akkor az azt jelenti, hogy *X* eloszlása mindkét csoportban normális, s az imént félbehagyott mondatot így folytathatjuk: „…e változót mindkét csoportban normális eloszlásúnak tekinthetjük (az első csoport esetén *p=p-value*; a második csoport esetén *p=p-value*). A továbbiakban meggyőződtünk arról is, hogy e változó szórása a két független mintában (a két csoportban) közel egyforma-e." A „*p-value*" helyére az R által megadott p-értéket kell beírni. Ebben az esetben az adatelemző munka következő lépéseként a szóráshomogenítás vizsgálata következhet.

A szóráshomogentiásról esetünkben akkor beszélhetünk, ha az almintákban (tehát a vizsgálati és a kontrollcsoportban) a vizsgált függőváltozó (itt: *X*) szórása közel egyforma. A szóráshomogenitás vizsgálatára alkalmazott matematikai statisztikai próba az almintákra alkalmazott Bartlett-próba, melynek R-parancsa:

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

A Bartlett-próba lehetséges eredményei a szignifikancia (*p*-érték) szempontjából:

a) Ha a Bartlett-próba eredménye szerint $p\text{-}value \leq 0{,}05$, akkor *X* szórása az almintákban különbözőnek tekinthető, s a két csoportot az *X* változó átlaga szempontjából a Welch-féle d-próvával hasonlíthatjuk össze. Publikációban ezt ilyesmi módon foglalhatjuk össze: „A szóráshomogenitás vizsgálatára Bartlett-próbát alkalmaztunk – ennek eredménye szerint (*p*=*p-value*) a szóráshomogenitás feltétele nem teljesül. Ezért az almintákat a továbbiakban Welch-féle d-próbával hasonlítjuk össze.". A Welch-féle d-próba R-parancsa a következő:

```
t.test(X~CSOP, data=d, var.equal=F)
atlagCs1=mean(d$X[d$CSOP==1])
atlagCs2=mean(d$X[d$CSOP==2])
atlagCs12=atlagCs1-atlagCs2
```

Az *X* változó átlagát a két független mintában hasonlónak tekinthetjük, ha a Welch-féle d-próba eredménye nem jelez szignifikáns különbséget (*p-value* > 0,05). Szövegbe ágyazva: „A Welch-féle d-próba eredménye szerint nincs jelentős különbség a két csoport között (*p*=*p-value*). A vizsgálati csoport átlaga: *atlagCs1*; a kontrollcsoport átlaga: *atlagCs2*. Az átlagok különbsége: *atlagCs12*."

A *p-value* ≤ 0,05 eredmény ellenben azt tükrözi, hogy **a**z *X* változó átlaga a két független mintában lényegesen különbözik: „A Welch-féle d-próba eredménye szerint szignifikáns (p=*p-value*) különbség van a két csoport között. A vizsgálati csoport átlaga: *atlagCs1*; a kontrollcsoport átlaga: *atlagCs2*. Az átlagok különbsége: *atlagCs12*."

b) Ha a Bartlett-próba eredménye szerint *p-value* > 0,05, akkor *X* szórása az almintákban hasonló, s a vizsgálati csoportot és a kontrollcsoportot *X* szempontjából kétmintás t-próbával (szokás függetlenmintás t-próbának is nevezni) vethetjük össze. Mindez mondatba foglalva: „A szóráshomogenitás vizsgálatára Bartlett-próbát alkalmaztunk – ennek eredménye szerint (*p*=*p-value*) a szóráshomogenitás feltétele teljesül, s lehetőség van a csoportok összehasonlításakor kétmintás t-próba alkalmazására.". A kétmintás t-próba arra a kérdésre ad választ, hogy az *X* változó átlaga a két független mintában egyenlő-e. A kétmintás t-próba R-parancsa ebben az esetben:

```
t.test(X~CSOP, data=d, var.equal=T)
atlagCs1=mean(d$X[d$CSOP==1])
atlagCs2=mean(d$X[d$CSOP==2])
atlagCs12=atlagCs1-atlagCs2
```

Most is a szignifikanciát jelző *p*-értéket vesszük figyelembe. Ha a kétmintás t-próba eredménye szignifikáns (*p-value* ≤ 0,05), az azt jelzi, hogy az *X* változó átlaga a két független mintában lényegesen különböző. Ezt eképpen fogalmazhatjuk meg: „A két mintás független t-próba eredménye szerint különbség van a két csoport között (*p=p-value*). A vizsgálati csoport átlaga: *atlagCs1*; a kontrollcsoport átlaga: *atlagCs2*. Az átlagok különbsége: *atlagCs12*."

Amennyiben a kétmintás t-próba eredménye nem szignifikáns (*p-value* > 0,05), akkor az *X* változó átlaga a két független mintában hasonló. Szövegesen: „A két mintás független t-próba eredménye szerint nincs jelentős különbség a két csoport között (*p=p-value*). A vizsgálati csoport átlaga: *atlagCs1*; a kontrollcsoport átlaga: *atlagCs2*. Az átlagok különbsége: *atlagCs12*."

Ordinális változók esete

Amennyiben a vizsgált változóink eleve ordinális skálájúak (például ötfokú skálán adott osztályzatok) vagy kvantitatív skálájúak ugyan, ám nem felelnek meg a normalitás és a szóráshomogenitás fentebb közölt feltételeinek, akkor nem paraméteres (nem az átlag és a szórás alapján „működő") statisztikai próbákkal dolgozhatunk. Általános értelemben véve a következő kérdésre keressük a választ a két csoport (legtöbbször, de nem kizárólag: a vizsgálati és a kontrollcsoport) ordinális skálát feltételező összehasonlítása során: a csoportosító változóban (például 1 és 2 számokkal kódolt) két minta a (legalább ordinális skálájú) *X* változó szempontjából sztochasztikusan egyenlő? Tekintsük át e kérdés vizsgálatát egy konkrét példán keresztül! Az alaphelyzet legyen a következő: „A tehetséggondozó programba bevont vizsgálati csoport és a programba be nem vont kontrollcsoport tanulói témazáró dolgozatot írtak. A tehetséggondozó program hatékonyságát jelezné, ha a programban résztvevők lényegesen jobb érdemjegyeket kapnának, mint a kontrollcsoportba tartozó társaik." tegyük fel, hogy az összegyűjtött adatokat egy *d* nevű adatbázisba szerkesztjük, mely legalább a következő három változót tartalmazza: DIAK (a tanuló neve vagy kódja), CSOP (csoportosító változó, értékei: 1 = vizsgálati csoport, 2 = kontrollcsoport), X (érdemjegy, értékei: 1-5), s így néz ki:

```
DIAK    CSOP   X
DIAK1   1      2
DIAK2   1      1
DIAK3   2      1
…       …      …
DIAKn   1      2
```

A javasolt statisztikai próba a Mann-Whitney-próba, amely azt teszteli, hogy az X változó szempontjából a két csoport sztochasztikusan egyenlőnek tekinthető-e. A Mann-Whitney-próba nullhipotézise: $X_{CSOP=1}$ éppolyan gyakran nagyobb $X_{CSOP=2}$-nél, mint fordítva; vagyis: *X* változó tekintetében a két minta sztochasztikusan ugyanakkora. R-parancs:

```
wilcox.test(X~CSOP, data=d)
```

A szignifikancia szempontjából megkülönböztethető két jellegzetes eredmény:

a) A nullhipotézist megtarthatjuk, ha a Mann-Whitney-próba eredménye szerint $p > 0{,}05$. Szövegbe ágyazva: „A Mann-Whitney-próba eredménye szerint a két csoport a vizsgált változónk szempontjából nem különbözik (sztochasztikusan egyenlők – *p=p-value*)."

b) A nulhipotézist elvetjük, ha a Mann-Whitney-próba eredménye: *p-value* ≤ 0,05. Ezt az eredményt például így fogalmazhatjuk meg: „A Mann-Whitney-próba eredménye szerint a két csoport a vizsgált változónk szempontjából különbözik (*p=p-value*)."

Az eredmények értelmezésében segíthet, ha leíró statisztikákat kérünk a vizsgált változókról minden csoportban. Kérhetünk kvartiliseket (melyek a mintát negyedelik), de kérhetünk részletesebb ábrázolást is. A kvartiliseket vagy a mintát ötödölő értékeket (kvantiliseket) a következő R parancsokkal kaphatjuk meg:

```
tapply(d$X, d$CSOP, quantile)
tapply(d$X, d$CSOP, quantile, c(.2,.4,.6,.8)
```

Nominális változók esete

Nominális (tehát csak megkülönböztethető, de nagyság szerinti sorrendbe nem állítható értékű) változók esetében általánosságban fogalmazva azt vizsgálhatjuk meg, hogy „*X* nominális változó megfigyelt gyakoriságai egyenlők-e a különböző csoportokban?”.

Lássunk egy konkrét példát! Tegyük fel például, hogy: „A magyarországi magyar és a határon túli magyar tehetséggondozó programba 2016-ban bevont diákokat vizsgáltunk. Arra voltunk kíváncsiak, hogy a magyarországi magyarok között hasonló-e a táborba jelentkezők aránya, mint a határon túli magyarok között?” Adatbázisunk (nevezzük röviden d-nek) ebben az esetben a következő három változót tartalmazza: DIAK (a tanuló neve vagy kódja), CSOP (csoportosító változó, értékei: 1 = magyarországi magyar, 2 = határon túli magyar), X (a jelentkezésről szóló információt tartalmazó változó, értékei: 1 = jelentkezett; 2 = nem jelentkezett), s így néz ki:

```
DIAK    CSOP   X
DIAK1   1      2
DIAK2   1      1
DIAK3   2      1
…       …      …
DIAKn   1      2
```

A javasolt statisztikai próba ilyen esetben a *függetlenségvizsgálat Khi-négyzet-próbával (χ2-próba)*, melynek nullhipotézise szerint az X változó gyakoriságai *CSOP* = 1 és *CSOP* = 2 esetén azonosak. A khi-négyzet-próba és az értelmezést segítő gyakorisági táblázat R-parancsa a következő lesz:

```
chisq.test(table(d$X,d$CSOP))
table(d$X,d$CSOP))
```

A khi-négyzet-próba eredménye a szignifikanciaszint (*p*-érték) szempontjából most is kétféle lehet. Ha $p > 0{,}05$, akkor a két csoportban hasonló a vizsgált (X) változó gyakorisága – a példánk szövegkörnyezetébe foglalva: „A khi-négyzet-próba eredménye szerint a két csoport között nincs szignifikáns különbség a jelentkezések arányát illetően (*p* = *p-value*). A Magyarországon élő és a határon túl élő magyar diákok hasonló arányban jelentkeztek/nem jelentkeztek a tehetséggondozó táborba. Az egyes csoportokban a jelentkezők és a nem jelentkezők száma a gyakorisági táblázatban látható.”

Ellenben, ha a $p \leq 0{,}05$, akkor a vizsgált *X* változó gyakoriságai a két csoportban lényegesen eltérnek. Szövegbe foglaltan például: „A khi-négyzet-próba eredménye szerint a két csoport között szignifikáns különbség volt a jelentkezések arányát illetően (*p=p-value*). A Magyarországon élő és a határon túl élő magyar diákok eltérő arányban jelentkeztek/nem jelentkeztek a tehetséggondozó táborba. Az egyes csoportokban a jelentkezők és a nem jelentkezők száma a gyakorisági táblázatban látható.”

Zárógondolatok

E tanulmány egy vizsgálati csoporttal és egy kontrollcsoporttal kapcsolatban egyetlen mintavételi alkalommal gyűjtött adatok összehasonlításának matematikai statisztikai hátterét foglalja össze – a matematikai statisztikai próbák kiválasztásától azok R-nyelvben történő végrehajtásán keresztül a kapott eredmények interpretációjáig terjedően.

A következő matematikai statisztikai próbákról esett szó:

- Shapiro-Wilk-próbát alkalmaztunk a két csoporttól gyűjtött adatok normális eloszlásának vizsgálata céljából.
- Bartlett-próbával teszteltük a két csoporttól gyűjtött adatok szóráshomogenitását.
- Kétmintás (másképp: független mintás) t-próbával hasonlítottuk össze a két csoport átlagát a normális eloszlás és a szóráshomogenitás feltételeinek megfelelő kvantitatív változók különbségvizsgálata céljából.
- Welch-féle d-próbát használtunk a két csoport átlagának összevetésére az almintánkénti normális eloszlás teljesülése, ám a szóráshomogenitás hiánya esetében.
- Mann-Whitney-féle U-próbával teszteltük az ordinális skálájú (vagy a normalitás és szóráshomogenitás kritériumának nem megfelelő esetekben a kvantitatív skálájú) változók sztochasztikus egyenlőségét.
- Khi-négyzet-próbával (χ2-próbával) végzett függetlenségvizsgálatot alkalmaztunk, ha nominális skálájú változó alapján történik a vizsgálati csoport és a kontrollcsoport összehasonlítása.

Korábbi tanulmányokban már összefoglaltuk azokat az eseteket, amelyekben mindössze egyetlen csoporttól (1, 2 vagy több alkalommal) gyűjtött adatok révén tehetünk kísérletet a fejlesztő program hatásvizsgálatának matematikai statisztikai elemzésére (Mező és tsai, 2015a,b, 2016). Jelen tanulmány 2 csoporttól egyetlen alkalommal gyűjtött adatok elemzésére fókuszált. A Különleges Bánásmód folyóirat további számaiban foglaljuk majd össze azokat az eseteket, amikor 2 csoportot 2 vagy több alkalommal gyűjtött adat alapján hasonlíthatunk össze. Hasonló a helyzet a 3 vagy több csoport bevonásával végzett hatásvizsgálatok közlésével is.

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 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

Dienes Z. (2013): *Mitől tudomány a pszichológia? A tudományos és statisztikai következtetés alapjai.* Akadémiai Kiadó, Budapest.

Eccles, M., Grimshaw, J., Cambell, M. & 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 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. (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 Ferenc, Máth János, Abari Kálmán és Mező Katalin (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

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

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.