FEJLESZTŐPROGRAMOK HATÁSVIZSGÁLATÁNAK MATEMATIKAI STATISZTIKAI ALAPFOGALMAI

Szerzők:

Máth János
Debreceni Egyetem

Mező Ferenc
Debreceni Egyetem

Abari Kálmán
Debreceni Egyetem

Mező Katalin
Debreceni Egyetem

Első szerző e-mail címe:
math.janos@arts.unideb.hu

Lektorok:

Demetrovics János
Eötvös Lóránd Tudományegyetem

Nagy Dénes
Nemzetközi Szimmetria Társaság

Varga Imre
Szegedi Tudományegyetem

Koncz István
Professzorok az Európai Magyarországért

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

Absztrakt

Pedagógusok, óvodapedagógusok, gyógypedagógusok, fejlesztőpedagógusok, pszichológusok – egy csokorra való azon szakemberek köréből, akik képesség- és/vagy személyiségfejlesztő programokat hozhatnak létre, valósíthatnak meg. E programokkal szemben elvárható azonban, hogy az általuk ígért fejlesztőhatást akár matematikai statisztikai elemzéssel is alá lehessen támasztani. Ez azonban nem mindig történik meg – ennek hátterében pedig részben a matematikai statisztikai fogalmi és módszertani hiányosságok, valamint a drága statisztikai szoftverek állnak. E tanulmány a különleges bánásmódot igénylők számára készült fejlesztőprogramok hatásvizsgálatának matematikai statisztikai elemzéséhez nyújt elméleti összefoglalót.

Kulcsszavak: fejlesztőprogram, hatásvizsgálat, statisztika

Diszciplinák: matematika, pszichológia, gyógypedagógia, pedagógia

Abstract

BASIC TERMS IN MATHEMATICAL STATISTICS OF IMPACT STUDIES FOR DEVELOPMENTAL PROGRAMS

Pedagogues, kindergarten pedagogues, special needs teacher (also known as remedial teachers) and psychologists – just a few of those people who can create and use development *programmes. These programmes are expected to be proven in their effects by using mathematical analysis. However, it doesn't happen in every case – partly due to shortage of definitions of mathematical statistics, methodology problems and expensive softwares. This paper is a theoretical summary about mathematical statistical analysis of effectiveness studies of remedial courses designed for those who live with special needs.*

Key words: development programmes, effectiveness studies, statitics

Disciplines: mathematics, psychology, special education, pedagogy

A sajátos nevelési igényű és/vagy beilleszkedési, tanulási és magatartászavaros és/vagy tehetséges tanulók felé irányuló különleges bánásmód többek között egyéni vagy csoportos fejlesztő-programokban nyilvánulhat meg. E fejlesztések központi kérdése, hogy vajon hatékonyak-e, elérik-e a fejlesztési tervben kitűzött céljukat; illetve, hogy valóban bizonyítható-e, hogy fejlesztés történt (vö.: Berényi és Katona, 2013, de lényegében ugyanezt a kérdést járja körbe medicinálisabb aspektusból: Guyatt, Cook és Haynes, 2004).

Egyéni fejlesztés hatékonyságának vizsgálata ≠ fejlesztő program hatásvizsgálata

Ha egy adott gyermek fejlesztésének hatékonyságáról kívánunk képet kapni, akkor a fejlesztő folyamat adott pontján tapasztalható eredményeit vethetjük egybe a saját korábbi (pl. fejlesztés előtti) eredményeivel és/vagy a fejlesztési tervben kitűzött célértékkel és/vagy a kortársak eredményeivel. Ez a helyzet nem igényel különösebb statisztikai felkészültséget. Ha egyáltalán számszerűsítenünk kell az eredményeket, akkor a vizsgáló eljárás leírásában szereplő útmutató révén összesítenünk kell a vizsgálati eredményeket (szükséges matematikai műveletek általában: összeadás, kivonás, szorzás, osztás, átlag- és százalékszámítás), majd össze kell hasonlítanunk a kapott végeredményeket a fejlesztési tervben meghatározott célértékkel, korábbi eredménnyel, mások eredményével (elvégzendő művelet: két érték között a kisebb, nagyobb, egyenlő reláció megállapítása). De: néhány egyéni, eseti jó/rossz eredmény alapján még nem biztos, hogy valóban általánosítható, ténylegesen a programról szóló információkat kapunk.

Az előzőnél bonyolultabb matematikai statisztikai feladat előtt állunk akkor, ha nem egy adott gyermek fejlesztésének a hatékonyságáról, hanem magának a fejlesztő programnak a hatásáról kívánunk képet alkotni. Egyrészt ez olyan esetben történhet meg, amikor ugyan-olyan fejlesztési terv alapján megvalósult sok egyéni/csoportos fejlesztés eredményének összesítése után szeretnénk a program eredményességére következtetni. Másrészt, ha egy adott tulajdonságokkal rendelkező csoport (nevezzük ezt a csoportot: vizsgálati csoportnak vagy mintának!) vizsgálati eredményei alapján szeretnénk következtetni arra, hogy a hasonló adottságokkal rendelkező többi személyre (egy szóval: a „populáció"-ra) is általánosíthatók lehetnek-e a fejlesztő program hatásáról szóló megállapításaink. Ilyen esetekben tehát statisztikai következtetéseket kell tennünk (Bolla és Krámli, 2012; Obádovics, 2003).

A fejlesztőprogramok hatásvizsgálata alkalmával arra vagyunk kíváncsiak, hogy a program következtében *jelentős különbség* adódik-e két vizsgálat eredményei között - pl. az elő- és utóvizsgálat(ok) és/vagy a fejlesztésbe különböző módon bevont csoportok eredményei, vagy ezen eredmények változásai között -, illetve egy vizsgálat és egy általunk meghatározott kritériumszint (pl. tehetségküszöb vagy a fejlesztési célban meghatározott érték) között. Az ilyen összehasonlításokat összefoglaló néven különbségvizsgálatoknak nevezzük.

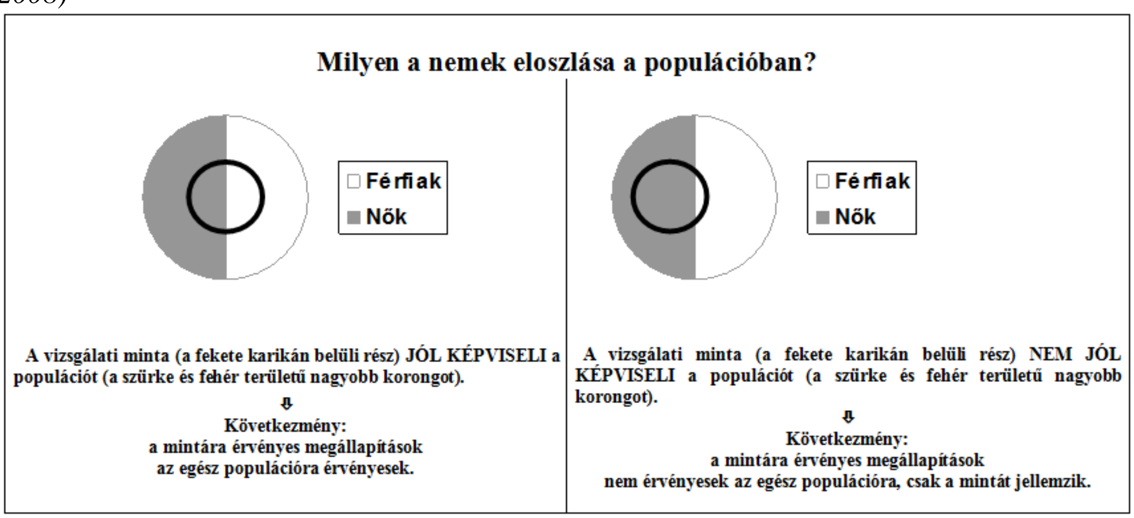
Ha a kapott különbség, eltérés jelentős (tudományosan fogalmazva szignifikáns), akkor nagy valószínűséggel valós hatást jeleznek. Ellenkező esetben a kapott eltérések véletlenszerű, eseti jelenségnek tekinthetők, melyek semmit sem bizonyítanak.

A továbbiakban vegyük szemügyre közelebbről a minta, a populáció, a változó, a skála, a statisztikai próba, a leíró statisztika és a matematikai statisztika, a hipotézisvizsgálat és a szignifikancia, a különbségvizsgálat és a kapcsolatvizsgálat fogalmát!

Minta és populáció

Egy konkrét fejlesztőprogram hatásvizsgálata kapcsán elsősorban az érdekli a szervezőket és a résztvevőket, hogy a programban résztvevő személyek milyen eredményt értek el. Más esetben – például a fejlesztőprogram tágabb értelemben vett tudományos jellegű hatásvizsgálata során – arra a kérdésre is szeretnénk választ kapni, hogy amennyiben a programban résztvevőket egy nagyobb közösség (populáció) tipikus képviselőinek tekintjük, akkor általánosítható-e a populációra (vagyis a nagy többségre) a résztvevők eredményei alapján kirajzolódó kép. A kérdésfeltevéseink utóbbi típusa tehát az emberek tágabb körére vonatkozik, mint azokra, akikről adatokat tudunk gyűjteni. Ha egy fejlesztőmódszer hatékonyságát vizsgáljuk egy csoportban, valójában azt kérdezzük, a csoporttagokra általában, vagy bizonyos csoporttagokra általában miként hat ez a módszer? Ebben az esetben tehát nem a konkrét csoport érdekel minket, hanem a módszernek egy bizonyos népességre gyakorolt hatása. Ezt a „szélesebb” népességet szokás *populációnak* nevezni, a konkrétan megvizsgált személyeket pedig *mintának*. Fontos megérteni, hogy a statisztikai eljárások mindig a mintát vizsgálják, de a populációról kívánnak állításokat megfogalmazni. Ez az általánosítás statisztikailag akkor megalapozott, ha a minta a releváns változók tekintetében „hasonló” a populációhoz (vö.: 1. ábra).

1. ábra: a populáció és a minta két lehetséges viszonya (Forrás: Mező, Máth és Abari, 2008)



Változók, skálák és statisztikai próbák

Ha emberek tulajdonságaival kapcsolatban – mint pl. nem, életkor, kreativitás, IQ, agresszió, feladatmegoldás pontszáma, irodalom jegy – méréseket végzünk, akkor a mért tulajdonságokat *változóknak* szokás nevezni. E változók megítélhetők abból a szempontból is, hogy a lehetséges értékeik mennyire rendelkeznek a számok tulajdonságaival. Ez alapján mondhatjuk egy változóról, hogy nominális, ordinális vagy kvantitatív skálát képvisel:

- *Nominális skála* esetén a lehetséges értékek között nem értelmezhető a kisebb/nagyobb fogalma sem, így sorrend sem kínálkozik közöttük. Példa lehet nominális változóra diákok esetén a kedvenc tantárgyuk.
- *Ordinális skála* esetén a lehetséges értékek között van sorrend, de ezen értékek különbsége nem értelmezhető. Példaként tekintsük az iskolai végzettséget a következő 4 értékkel: *legfeljebb 8 általános, szakmunkásképző, érettségi, felsőfokú végzettség*. Bár a végzettség fokozatai egyértelműen sorba rakhatók, az nem jelenthető ki, hogy a *főiskola* éppen annyival jelent magasabb végzettséget az *érettséginél*, mint a *szakmunkásképző* a *8 általánosnál*. Valójában a kérdés ennyiből nem is értelmezhető. További példák ordinális változóra: a diákoknak rangsorolni kell, hogy adott dolog hol helyezkedik el számukra egy fontossági sorban, vagy hogy mekkora erőfeszítést tesznek egy ügy érdekében és a lehetséges választások pl. a következők: *semmilyet, épphogy csak, amennyit szükséges, mindent beleadok*. Egy másik példa, hogy a diákok órai teljesítményét a következő fokozatokkal ítélik meg: *gyenge, megfelelő, kiváló*.
- *Kvantitatív (vagy: Intervallum) skáláról* akkor beszélünk**,** ha a változó értékei között nem csak a sorrend értelmes, de ezen értékek különbsége is értelmezhető. Az ilyen változókat *kvantitatívnak* is szokás nevezni. Ilyennek tekinthető pl. a legtöbb teljesítményteszt, képességteszt (pl. IQ teszt), ami tehát azzal a plusz tulajdonsággal jár, hogy kijelenthetjük: a 120 pont éppen annyival jobb eredmény a 100-nál, mint a 140 a 120-nál.

A változók egy másik szempont szerint lehetnek *diszkrétek* és *folytonosak*. A diszkrét változó lehetséges értékei jól elkülönülnek egymástól, mint például a kedvenc tantárgy vagy az iskolai végzettség. Az olyan változót, amely eredendően folytonos jelenséget mér és a lehetséges értékei között kellően kicsi különbség van (vagyis a mérés kellő pontossággal rendelkezik), folytonosnak nevezzük. Például a legtöbb IQ teszt folytonosnak tekinthető, mert az intelligencia (vélhetően) folytonos és a mérés kellően árnyalt. Ugyanakkor nem folytonos az intelligenciának az a mérése, ahol csak átlag alatti, átlagos és átlag feletti kategóriákat használjuk. A folytonosság kérdése azért fontos, mert ha egy változó legalább ordinális, általában egyszerűbb és hatékonyabb az elemzés, ha a változót folytonosnak tekinthetjük. Minél finomabb a skála, annál inkább megtehetjük ezt. A statisztikai gyakorlatban az ötfokú skálákra már szokták alkalmazni a folytonos változókra vonatkozó eljárásokat.

Leíró és matematikai statisztikák

A leíró statisztikák a minták számszerű jellemzésére valók. A minta „közepét” (kvantitatív változó esetén) általában az *átlaggal* vagy (legalább ordinális változó esetén) a *mediánnal* szokás megragadni. A minta átlaga az, amit mindenki ért ezen: a minta értékeinek átlagát. A medián talán rövid magyarázatra szorul. Némi leegyszerűsítéssel azt az értéket jelenti, amely megfelezi a mintát abban az értelemben, hogy annyi érték van felette, mint alatta. Ha pl. 5 diák IQ értékei: 80, 86, 101, 140, 90, akkor ennek az 5 elemű mintának a mediánja 90 (két érték – a 80 és a 86 – kisebb, mint 90; két érték – a 101 és a 140 – pedig nagyobb, mint 90). Ha a mintát 6 eleműre bővítjük: 80, 85, 101, 140, 90, 70, akkor a medián (a növekvő sorrendbe állított 70, 80, 85, 86, 90, 101, 140 számsor) a két középső értékének, vagyis a 86-nak és a 90-nek az átlaga, tehát: 88. Az átlag intervallum skálák esetén használatos, míg a medián csupán ordinális skálát „igényel” (Falus és Ollé, 2000; Szűcs, 2002).

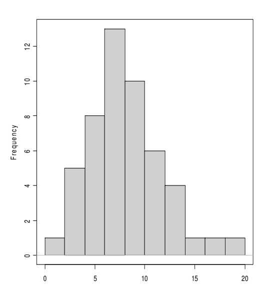
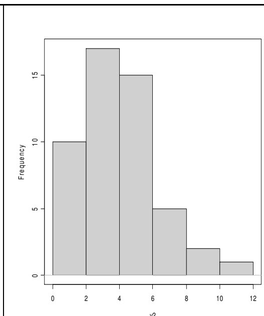
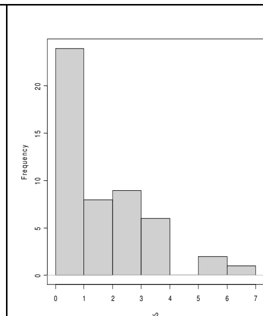
Gyakran felmerül a közép körüli ingadozás nagysága. Ennek leggyakoribb mérőszáma a *szórás*. Némi leegyszerűsítéssel azt mondhatjuk, a szórás az átlagtól való eltérésnek egy olyan középértéke, ami közel áll az átlaghoz: az átlagtól való átlagos eltérés.

A matematikai statisztikai eljárások a leíró statisztikák adatait felhasználó, illetve azok alapján megválasztható számítások (Vargha, 2000). A vizsgálatok egy része olyan eszközökkel (tesztek, feladatsorok, osztályzatok) történik, melyek lehetséges értékei *kvantitatív-skálát* képviselnek és lényegében folytonosak is (vagyis elég „finom” beosztásúak). Itt a további elemzés azon múlik, hogy a minta eloszlása normálisnak tekinthető-e? Ha igen, akkor az átlagok eltérését vizsgáló eljárások (t-próbák, varianciaanalízis) alkalmazhatók, melyek a legerősebb eszközök a különböző (csoportok közötti, időpontok közötti) eltérések kimutatására. A „normalitás” vizsgálata (többek között) a Shapiro-Wilk-próba segítségével történhet. Ha a Shapiro-Wilk-próba eredményeként kapott szignifikancia érték (*p-value*) 0,05 alatt van, akkor a normalitást el szokás vetni. A 2. ábrán néhány példát láthatunk, ahol az első minta még normálisnak tekinthető, a második kettő már nem. (Megjegyezzük, hogy a varianciaanalízis nem túlságosan érzékeny a normalitás hiányára, ezért ha az említett *p-value* néhány századdal 0,05 alatt marad, a varianciaanalízis által adott eredmények még jól használhatók)

Ha azt kaptuk, hogy a *normalitás nem áll fenn*, vagy a skála *ordinális*, akkor az ún. rangokon alapuló eljárásokat alkalmazhatjuk pl. (Wilcoxon-próba), melyek az adatok tényleges értéke helyett a nagysági sorrendben elfoglalt helyükkel számol. Némi leegyszerűsítéssel azt mondhatjuk, itt az átlagok helyett a mediánok eltérését vizsgálják.

Ha a minta (csoportonkénti) eloszlása a 2. ábra nem normális eloszlást mutató eseteiben látható eloszlásnál is „csúnyább” képet mutat (pl. több „púpja” van), érdemes statisztikussal konzultálni (kapcsolat: www.kockakor.hu), mert sok minden lehet a dolog hátterében.

2. ábra: példák normális és nem normális eloszlásra (Forrás: Mező, Máth és Abari, 2008)

		
Shapiro-Wilk normality test data: y1 W = 0.9615 *p-value* = 0.103	Shapiro-Wilk normality test data: y2 W = 0.9409 *p-value* = 0.015	Shapiro-Wilk normality test data: y3 W = 0.8556, *p-value* = 0.000
Eloszlás: normális	Eloszlás: nem normális	Eloszlás: nem normális

Az is lehetséges, hogy a mérés *nominális skálán* történik. Hatásvizsgálatokban talán a *kétértékű nominális változó* fordul elő leggyakrabban, melynek tipikus példája az igen/nem, a siker/kudarc, fiú/lány. Kétértékű változók esetén több eljárás kínálkozik a fejlődés detektálására, de többértékű esetben is vizsgálható a változás. Ha több csoportot egyszerre vizsgálunk, a változások értelmezése bonyolulttá válik, melynek elméleti hátterét a loglineáris elemzés adja. Ennek ismertetése túlmutat e mű keretein (Máth, 2004).

A matematikai statisztikai vizsgálatok két átfogóbb csoportját képezik a különbség-vizsgálatok, illetve a kapcsolatvizsgálatok. Amíg az előzőek összehasonlítanak egy vagy több adatsort, addig az utóbbiak az adatsorok közötti bejósolhatóságot, együtt járást vizsgálják.

Hipotézisvizsgálat

A hipotézisvizsgálat lényege, hogy legalább két „versengő” hipotézis között (*H_0* és *H_1* hipotézis) kell döntenünk (Borovkov, 2012). *H_0* (más szóval: *nullhipotézis*) jelentése: az eltérés (a csoportok, időpontok között) nem jelentős, véletlenszerű, csak a vizsgált mintában van jelen, az egész populáció esetében már nem áll fenn. *H_1* jelentése: az eltérés a véletlent meghaladó, lényeges különbség, az egész populációra vonatkozó, azaz *„szignifikáns”*. A szignifikancia annak a valószínűségét jelenti, hogy a kapott (vagy annál is nagyobb) eltérés H_0 esetén is bekövetkezhet. Minél kisebb a szignifikancia, annál nagyobb meglepetést jelent az adott eltérés *H_0*-t feltételezve. A pedagógiai kutatásokban a 0,05-nél kisebb szignifikanciaérték (a továbbiakban: *p-érték* vagy *p-value*) esetén szokás úgy dönteni, hogy *H_1*-et választjuk, vagyis hogy a kapott eltérés lényeges, nem a véletlen műve, azaz szignifikáns (2. ábra).

2. ábra: a szignifikancia értékének (a p-értéknek az) értelmezési lehetőségei

HA a szignifikanciát számszerűsítő p-érték (p-value) =	**0 0,05**	**… 1**
AKKOR:		
Mekkora a valószínűsége, hogy véletlen folytán legalább ekkora különbség keletkezzen két vizsgálati eredmény között (feltéve, hogy a módszer teljesen hatástalan)?	**0% - 5%**	**100%**
Mekkora a valószínűsége (=1-p), hogy véletlen esetén legfeljebb ekkora különbség mutatkozik (feltéve, hogy a módszer teljesen hatástalan)?	**100% 95%**	**0%**
Az eredmény szignifikáns?	**IGEN**	**NEM**
H_0 (=a különbség nem jelentős, véletlenszerű, csak a vizsgált mintában van jelen, a populációra nem érvényes) igaz?	**NEM**	**IGEN**
H_1 (=a különbség jelentős, a populáció egészére is érvényes, nemcsak a mintára) igaz?	**IGEN**	**NEM**

Példaként tegyük fel, arra vagyunk kíváncsiak, hogy egy csoport év végi átlagteljesítménye (55,6) fejlődést mutat-e az év elejivel (52,7) összevetve egy képzeletbeli tesztben? A szignifikanciavizsgálat eredménye lehet:

- a szignifikancia értéke (a p-érték) *1*: ez arra utal, hogy a két átlag „hajszálra” megegyezik és semmi okunk bármilyen fejlődésre gyanakodni.
- a kapott szignifikancia érték (p-érték) nagyobb, mint 0,05. Ha például: *p=0,16*, ez azt jelenti, hogy ekkora (vagyis 55,6-52,7=2,9 pontos) vagy ennél is nagyobb eltérés 0,16 (azaz 16%-os) valószínűséggel akkor is előfordulhat, ha valójában semmilyen érdemi változás nem történt a csoportban, ezért ebben az esetben nem tudtuk igazolni, hogy fejlődés történt volna: *tehát az eltérés nem szignifikáns*.
- a kapott szignifikancia érték (p-érték) kisebb 0,05 vagy pontosan 0,05. Például: ha *p=0,02*, akkor kijelenthetjük, érdemi fejlődés áll fenn, az év eleji és az év végi vizsgálat eredményei között *az eltérés szignifikáns*.

Meg kell jegyezni, hogy a hipotézisvizsgálat logikája sem zárja ki a tévedés lehetőségét. Példánknál maradva előfordulhat, hogy a csoport – noha semmit sem fejlődött – mégis szignifikánsan jobban teljesít, mert az első mérés eredménye volt túl rossz (pl. a vizsgálati személyek mással voltak elfoglalva a méréskor) vagy a második túl jó (pl. könnyűek voltak a feladatok, jó napot fogtak ki, stb…). Ilyenkor tévesen H_1 – et választjuk. E hiba elkövetésének valószínűségét maximalizálja a 0,05-ös szint. Ennek az ellenkezője is megeshet, amikor tényleges fejlődés van az osztályban, de az első mérés sikerült „túl jól" vagy a második „túl rosszul" és az átlagok nem térnek el eléggé. Ebben az esetben tévesen azt mondjuk, nem tudtunk érdemi, azaz szignifikáns fejlődést kimutatni. Különösen nagy e hiba esélye akkor, ha túl kicsi a vizsgált minta és/vagy túl kicsi a kapott eltérés és/vagy túl nagy a mérés hibája.

Különbségvizsgálatok

A különbségvizsgálatokon alapuló hatásvizsgálatok kilenc jellegzetes esetét különböztethetjük meg aszerint, hogy hány csoportról hány alkalommal nyert eredményeket vetjük össze egymással (1. táblázat).

1. táblázat: A legalapvetőbb különbségvizsgálatokra vonatkozó példák

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?	Pl.: a fejlesztőprogramba bevont csoport program előtti (pl. 2005-ös) és utáni (pl. 2015-ös) teljesítménye különbözik-e egymástól?	Pl.: a fejlesztett csoport 2005-ben, 2010-ben, 2015-ben stb. nyújtott teljesítménye különbözik-e egymástól?
2	Pl.: a fejlesztett (vizsgálati) csoport 2015-ös teljesítménye különbözik egy fejlesztésbe be nem vont (kontroll-)csoport teljesítményétől?	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 a az egyes csoportokban?

A különbségvizsgálatokban résztvevő minták jellegzetességei alapján egymintás, összetartozó mintás, illetve független mintás eseteket különböztethetünk meg. *Egymintás* különbségvizsgálatról van szó abban az esetben, ha egy csoport eredményét egy (pl. fejlesztési tervben) kitűzött célértékhez (kritériumszinthez) hasonlítjuk. *Összetartozó mintás* vizsgálatoknak tekinthetjük azokat a helyzeteket, amelyek alkalmával ugyanannak a mintának két (e speciális esetben: páros vagy paired vizsgálat megnevezéssel is találkozhatunk) vagy több adatgyűjtés során nyert eredményeit vetjük össze egymással. A *független minták* összehasonlításakor két vagy több csoport teljesítményét hasonlítjuk össze azonos változó(k) mentén – ilyen változók lehetnek például: IQ, szófluencia stb..

Kapcsolatvizsgálatok

Fejlesztőprogramok hatásvizsgálata során felmerülhet annak igénye is, hogy a vizsgálatok során gyűjtött adatok közötti összefüggéseket (például két változó közötti egyenes/fordított arányosságot, együtt járást stb.) is feltárjuk. Ez történik olyan esetekben, amikor arra keressük a választ, hogy az egyik változó (például: általános intelligencia, motoros fejlettség, kreativitás stb.) értékének növekedése esetén a másik változó (például önértékelés önismeret, tanulmányi átlag stb.) értéke vajon szisztematikusan nő, csökken, vagy teljesen bejósolhatatlanul változik-e. A kapcsolatvizsgálatok alkalmazásakor tehát arra vagyunk kíváncsiak, hogy egy(néhány) változó aktuális értékének ismeretében érvényes és megbízható következtetéseket tehetünk-e más változó(k) értékére.

Ha csak két változó kapcsolatát vizsgáljuk, akkor a leggyakrabban használt mutatók folytonos változók esetén (tehát ha a teljesítményt egy árnyalt pontszámmal mérjük) a Pearson-féle és a Spearman-féle korreláció, illetve az egyváltozós regressziószámítás.

Ha a mérésünk csak néhány fokozatot jelent, esetleg csak nominális skálán mérünk (tehát a lehetséges értékek még sorban sem rendezhetők), akkor használhatjuk a Kendall tau-b mutatót, vagy a Khi-négyzet próbával együtt a lambda együtthatót.

Ha egyszerre több változónk is van, melyekből következtetni kívánunk a megfelelő, teljesítmény változóra (például önértékelés önismeret, tanulmányi átlag), akkor a következő többváltozós eljárásokat alkalmazhatjuk (vö.: Vargha, 2000):

- Parciális korrelációszámítás;
- Többváltozós lineáris regressziószámítás;
- Diszkriminancia analízis;
- Logisztikus regressziószámítás.

A kapcsolatvizsgálatok sajátos esete a faktoranalízis, amikor nagy számú vizsgálati változónk mögött meghúzódó (rejtett, latens módon megnyilvánuló) háttérváltozókat keresünk.

Zárógondolatok

A fentiekben áttekintettük néhány jellegzetes esetét és alapfogalmát a bizonyítékokon alapuló hatásvizsgálatok matematikai statisztikai világának.

A matematikai statisztikai számítások végrehajtására napjainkban már szoftverek is rendelkezésre állnak - ezek többsége azonban olyan összegbe kerül, hogy megvásárlásuk sokszor intézményi szinten is kihívást jelenthet. Utolérhetők azonban az internetről ingyenesen letölthető, matematikai statisztikai számítások végzésére alkalmas programcsomagok is. Az egyik ilyen ingyenesen elérhető, sokoldalú és magyar nyelven is utolérhető szakirodalmi segédletekkel (Abari, 2008; Mező, Máth és Abari, 2008; Solymosi, 2005) rendelkező szoftver az úgynevezett R-nyelv. E lap következő számaiban igyekszünk közzé tenni a jövőben az R-nyelvvel és a hatásvizsgálatokkal kapcsolatos további módszertani tanulmányokat, amelyeket ezúton is ajánlunk az érdeklődők figyelmébe!

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

Berényi Marianne - Katona Ferenc (2013): Fejlesztések és terápiák. Fogalomzavar, vagy vetélkedés a mindennapokért? *Gyógypedagógiai Szemle*, 2013, XLI. évf. 3., 174-185.

Bolla M. és Krámli A. (2012): *Statisztikai következtetések elmélete.* Typotex Kiadó, Budapest.

Borovkov, A. A. (2012): *Matematikai statisztika. Paraméterek becslése, hipotézisvizsgálat*. Typotex, Budapest.

Falus I. és Ollé J. (2000): *Statisztikai módszerek pedagógusok számára*. Okker Kiadói Kft., Budapest.

Guyatt, G., Cook, D. és Haynes, B. (2004): Evidence based medicine has come a long way. *BMJ*, 2004. Oct. 30. 329(7473), 990-991.

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.

Obádovics J. Gy. (2003): *Valószínűségszámítás és matematikai statisztika*. Scolar Kiadó, Budapest.

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

Szűcs István (2002): *Alkalmazott statisztika*. Agroinform Kiadó, Budapest.

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