10.12663/PSYHUNG.7.2019.2.7.

# HIBABECSLÉSI ELJÁRÁSOK VÉLETLEN JELENSÉGEK PARAMÉTEREINEK BECSLÉSÉRE

T. Kárász Judit[1]

Levelező szerző: T. Kárász Judit (takacsne.karasz.judit@oh.gov.hu )

## Kivonat

Statisztikai számításokban a kiszámított paraméterek standard hibájának becslésére számos eljárás létezik. Cikkünkben ismertetünk a független, azonos eloszlású mintavételekben használható, zárt formulás eljárások közül néhányat. Az általánosabb szimulációs eljárások közül a bootstrap és jackknife eljárásokat, a BRR-Jackknife eljárást, valamint a cikk végén egy, az Országos kompetenciamérésben használt, bootstrap és jackknife eljárások ötvözéseként kifejlesztett eljárásunkat.

**Kulcsszavak**: paraméterbecslés ▪ hibabecslés ▪ bootstrap ▪ jackknife ▪ Országos kompetenciamérés

## Abstract

In a statistical process it is important to make estimation on the standard error of our calculated parameters. There are many different algorithms to do it. In our article, we discuss closed formulas on some parameters. We describe the bootstrap and the jackknife general simulation algorithms. We describe the Fay's BRR jackknife method too. Finally we describe a mixed bootstrap-jackknife algorithm which applied in the National Assessment of Basic Competencies (NABC).

**Keywords**: point estimation ▪ standard error ▪ bootstrap ▪ jackknife ▪ National Assessment of Basic Competencies (NABC)

## BEVEZETŐ

Különböző jelenségek megfigyelésekor, akár természeti, akár társadalmi vagy éppen ipari jellegűekről legyen szó, mérésekkel gyűjtünk adatokat. Ezekből számításokkal, statisztikai módszerekkel igyekszünk a mintabeli jellemzőket ki-

[1] Oktatási Hivatal, Budapest, 1055, Szalay utca 10-14.

számítani – becsülve ezzel a populációbeli változataikat. Ez lehet akár az évi középhőmérséklet, villanykörték várható élettartama vagy a félévi matematika osztályzat. Ezek a becslések leggyakrabban számértékek (pontbecslések), azonban nem mindegy, hogy milyen pontosak ezek a becslések, ezért elvárható, hogy a becslés hibáját is kiszámítsuk. Ez speciális feltételek mellett képlet, általánosan alkalmazható függvények alapján számítható, gyakrabban azonban egyéb statisztikai eljárásokat kell alkalmazni.

Az alábbiakban bemutatunk néhány hibabecslési eljárást. Az első fejezetben képlettel számított, független mintavételen alkalmazható módszereket mutatunk be (Lehmann & Casella, 1998). A második fejezettől különböző szimulációs módszereket tekintünk át, melyek közül az első egy determinisztikus változat lesz (Miller, 1974). Második szimulációs módszerünk (Efron & Gong, 1983) szintén egy általánosan használt és ismert eljárás – mely véletlen mintavételezésen alapul. Utolsó fejezetünkben az Országos kompetenciamérés (OKM) példáján keresztül ismertetünk egy determinisztikus és véletlen elemeket ötvöző eljárást (Balázsi és mtsai., 2017) annak felhasználási módjaival, mely e méréshez kifejlesztett, a második és a harmadik módszer ötvözésével keletkezett.

## STANDARD HIBA ÉS KONFIDENCIA-INTERVALLUM, ZÁRT KÉPLETEKKEL

Tegyük fel, hogy egy normális eloszlású valószínűségi változó várható értékét szeretnénk megbecsülni, s ehhez n elemű független mintát venni (független, azonos eloszlású mintavétel történik). Hangsúlyozzuk, hogy a várható érték (populáció béli átlag) becslésére a számtani átlag akkor megfelelő mutató, ha egymástól független, azonos eloszlású mintával rendelkezünk (tehát ehhez nem szükséges a normalitás, mint feltétel, lásd például (Lehmann & Casella, 1998)).

Ekkor tehát a várható értéket becsülhetjük a minta átlagával, azaz a mintaelemek értékét összeadjuk, és elosztjuk az elemszámmal.

(1) $$\hat{\mu} = \overline{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

A mintaelemek eltérésének átlagos mértékét a várható értéktől a szórás jellemzi. Általában nem ismerjük a valódi szórást, ilyenkor erre is becslést kell alkalmaznunk. Mivel a szórás négyzete (más néven variancia) a mintaelemek várható értéktől való átlagos négyzetes eltérése, ezért a becslése a mintaelemek átlagtól való átlagos négyzetes eltérése lesz.

(2) $$\hat{\sigma}^2 = \frac{\sum_{i=1}^{n} (X_i - \bar{X})^2}{n-1}$$

Tegyük fel, hogy megismételjük a kísérletet, ismét veszünk egy n elemű mintát, jelölje ezeket $X_1^{(2)}$, ..., $X_n^{(2)}$, és kiszámítjuk az átlagot.

(3)
$$\hat{\mu}_{(2)} = \overline{X}_{(2)} = \frac{\sum_{i=1}^{n} X_i^{(2)}}{n}$$

Azt fogjuk tapasztalni, hogy a két átlag (igen nagy valószínűséggel) nem egyezik meg pontosan, azaz az általunk vizsgált változó várható értékét nem tudjuk pontosan meghatározni: ahány minta, annyi átlag. Azonban, ha elég sok mintát veszünk, akkor a mintaátlagok értéke a várható érték körül fog csoportosulni. Ebben az értelemben láthatjuk, hogy a standard hiba négyzete nem más, mint a mintaátlagok varianciája a várható érték körül, ami éppen $\sigma^2/n$.

A becslésnek ezt a bizonytalanságát – vagy éppen pontosságát/pontatlanságát – fejezi ki a standard hiba értéke. Gyakran nincs lehetőség elég sok mintát venni ahhoz, hogy a mintaátlagok átlagával elég pontosan meghatározzuk a várható értéket (vagy populációs átlagot), de normális eloszlás esetében, ismert szórás mellett a hiba zárt képlettel is megadható: a szórás és az elemszám négyzetgyökének hányadosa.

(4)
$$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}}$$

Ha a szórás sem ismert, akkor a becslését alkalmazzuk.

(5)
$$\hat{\sigma}_{\overline{X}} = \frac{\hat{\sigma}}{\sqrt{n}}$$

A (4) képlet alapján látható, hogy a becslés hibája két módon "csökkenthető". Egyrészt a szórás csökkentésével, ami felett általában kevéssé van hatalmunk, hiszen az az eloszlás saját jellemzője. A másik lehetőség, hogy a minta elemszámát növeljük. Ez alapvetően egyfajta intuícióval is belátható: minél több mintaelemet veszünk, annál pontosabban becsülhetjük meg a változó kívánt paraméterét. (Megjegyezzük, hogy ehhez felhasználjuk a számtani átlagnak, mint a várható értéknek a konzisztens voltát – és úgy általában is, a konzisztens (aszimptotikusan konzisztens) becslések sajátossága az, hogy a hiba tetszőlegesen csökkenthető a mintaelemszám növelésével.)

Itt kell szót ejtenünk a konfidenciaintervallum fogalmáról is, lásd például (Neyman, 1937; Lehmann & Casella, 1998). A normális valószínűségi változók esetében az értékkészlet végtelen, azonban a mintaelemek 95%-ban a ($\mu$-1.96*$\sigma$; $\mu$+1.96*$\sigma$) tartományból származnak. Mivel az átlag maga is egy véletlen változó, így lényegében normálisok konvex kombinációjaként kezelhető. Ezért az adott valószínűségi változóhoz tartozó minták átlagai szintén normális eloszlásból származó mintaelemek, ahol a várható érték az eredeti változó várható értéke (lásd korábban a konzisztencia fogalmát). Az egymás utáni mintavételekből számított átlagok szórása pedig maga a standard hiba. Így az előbbihez hasonló gondolatmenettel, a különböző lehetséges mintákhoz számított átlagokat 95%-ban ennek a $(\mu\text{-}1.96^*\hat{\sigma}_{\overline{X}};\ \mu+1.96^*\hat{\sigma}_{\overline{X}})$ tartománynak kell tartalmaznia. Szimmetrikusan, a számtani átlaghoz rendelt 95%-os konfidencia-intervallum 95%-os valószínűséggel tartalmazza a vizsgált változó várható értékét.

Természetesen más paraméterekhez, véletlen jellemzőköz is számítható konfidenciaintervallum (Joanes & Gill, 1998; Bruce Turner, 1986), mely tehát definíció szerint adott valószínűséggel tartalmazza az elméleti paramétert – de azt nem állíthatjuk, hogy bármely paraméter esetében a fentiekhez hasonló módon, zárt alakban megadható lenne a konfidencia-intervallum sugara, szélessége – azaz a becslés standard hibája (Chan, Golub, & Leveque, 1983).

## DETERMINISZTIKUS HIBABECSLÉSI, SZIMULÁCIÓS ELJÁRÁSOK

Az előző részben egy speciális eloszlás, a normális eloszlás paramétereinek becsléséről volt szó. Más eloszlások esetén más paramétereket egyéb módon becslünk, azok hibájának kiszámítására más módszerre van szükség.

A jackknife eljárás alapötlete az, hogy az eredeti független, n elemű mintából n db (n-1) elemű mintát képezünk, és ezek mindegyikére kiszámítjuk a kérdéses paraméter becslését (Efron & Gong, 1983). Az i. minta legyen az, amikor az eredeti mintaelemek közül éppen az i. elemet hagyjuk ki.

(6) $$X_{(i)} = (X_1, X_2, \dots X_{i-1}, X_{i+1}, \dots X_{n-1}, X_n)$$

Ekkor a paraméter i. jackknife részbecslése vagy replikánsa a paraméter becslése a részmintán,

(7) $$\hat{\theta}_{(i)} = \hat{\theta}(X_{(i)})$$

a paraméter jackknife becslése pedig e replikánsok átlaga.

(8) $$\hat{\theta}_{(\cdot)} = \frac{\sum_{i=1}^{n} \hat{\theta}_{(i)}}{n}$$

Ebben az esetben a jackknife becslés hibája

(9) $$\hat{\sigma}_{(\cdot)} = \sqrt{\frac{n-1}{n} \sum_{i=1}^{n} (\hat{\theta}_{(i)} - \hat{\theta}_{(\cdot)})^2}$$

Látható, hogy az eljárás adott minta esetén mindig ugyanúgy történik, ez a determinisztikus tulajdonság, azaz a számítás reprodukálható. A módszer erőssége, hogy nem szükséges a vizsgált eloszlás normalitása, bármilyen paraméterbecslés esetén, kis elemszám esetén is alkalmazható, és csökkenti a paraméterbecslés torzítását az eredeti becsléshez képest. Hátránya, hogy nem minden eloszlás esetén alkalmas hipotézisvizsgálatra. Egy egyszerűbb matematikai trükkel azonban ezt a hátrányát is meg lehet oldani: pszeudoparaméterek bevezetésével, melynek segítségével úgy módosítható a jackknife eljárás mintavétele, hogy a keletkezett replikánsokból számított eredmények már egymástól függetlennek legyenek tekinthetők. Miller, illetve a módszer kidolgozója, Tukey (Tukey, 1958) szerint igaz, hogy ezek a pszeudoértékek közelítőleg függetlenek, akárcsak a normális eloszlás független mintaelemei, és a fenti (4), illetve (5) hibaszámításhoz hasonlóan a becslés hibája ugyanúgy, lényegében tehát determinisztikus módon megadható.

## EGY VÉLETLENEN ALAPULÓ HIBABECSLÉSI ELJÁRÁS A BOOTSTRAP ALGORITMUS

A részminták használatának ötletét használja a bootstrap algoritmus is, de ezeket a részmintákat – pl. a jackknife eljárással szemben – nem előre rögzített módon választjuk (Efron & Gong, 1983). A módszer lényege, hogy az eredeti n elemű mintára, mint teljes populációra tekint, és ebből a populációból visszatevéses módszerrel választ n elemű (vagy akár ettől eltérő számú) véletlen mintákat.

*Itt tehát arra kell gondolni, ami az egyszerű, visszatevéses, véletlen mintavétel alapkoncepciója: adott egy „doboz”, benne a különböző vizsgálati esetekkel. Véletlenszerűen választunk belőlük, feljegyezzük őket, majd visszatesszük. Újra és újra ezt a kísérletet elvégezve a doboz „tartalmát”, jellemzőit szeretnénk megismerni. Ezen alapul a bootstrap mintavételezés technikája.*

Ennek következménye, hogy egy ilyen részmintában az eredeti minta bármely eleme akár többször is szerepelhet, mások kimaradnak belőle, és a tapasztalati eloszlás tulajdonságai megőrződnek az újabb és újabb mintavételek során. Miután kellő számú részminta keletkezett, mindegyikre számítható az eredetileg vizsgált paraméter egy becslése, amik az eredeti paraméter eloszlásának tapasztalati eloszlásfüggvényét alkotják.

Ekkor, az első részben megfogalmazottak szerint, a hiba a bootstrap becslések szórása lesz (tehát a számított paraméterekre, mint mintára fogunk tekinteni), de bármilyen konfidencia-intervallum is meghatározható: egyszerűen a tapasztalati percentilisek által.

Pl a 95%-os konfidencia-intervallum határai a 2,5%-os és 97,5%-os tapasztalati percentilisek. Tehát 95%-os szinten, ha például 200 véletlen mintát választunk, akkor a keletkezett, számított paramétereket nagyság szerinti sorba rendezve az 5. legkisebb és az 5. legnagyobb becslés értéke adja a 95%-os konfidencia-intervallum széleit.

A szakirodalom szerint a részminták számát 200 és 500 között szokás meghatározni. 200 szükséges a pontossághoz, 500 felett már nem érünk el számottevően javulást a becslésben. Ezzel együtt fontos kiemelni például azt, hogy az IBM SPSS automatikus beállítása a részminták (replikánsok) számát alap beállításként 10.000-ben határozza meg (Efron & Gong, 1983).

## EGY ÚJ, KOMBINÁLT MEGVALÓSÍTÁS AZ ORSZÁGOS KOMPETENCIAMÉRÉS (OKM) HIBABECSLÉSI ELJÁRÁSA

Az Országos kompetenciamérés kialakításakor szellemiségében és módszertanában is a PISA mérések metodikájához (Fay, 1984) kívánt igazodni, ezért egyértelmű volt, hogy valamifajta hibabecslésre a Kompetenciamérés esetében is szükség van. A teljes, pontos leírást az Oktatási Hivatal honlapján el lehet érni (Auxné, Bánfi és mtsai., 2015).

A PISA mérés rétegeket (strátumokat) határoz meg, és azon belül iskolákat tekint ekvivalensnek (egymással felcserélhetőnek), vagyis csak ezekből választ ki személyeket a 15 évesek vizsgálatára (Rutkowski, 2015). A rétegzett mintavételhez egy BRR jackknife eljárást használ (lásd például Takács, 2010), ami figyelembe veszi a rétegzett mintavételből és az egy osztályban tanulók eredményeinek összefüggéséből eredő statisztikai nehézségeket.

A Kompetenciamérés 2008 évtől teljes körű, ezért a mintavétel okozta hiba elhanyagolható (lényegében nem mintát veszünk, hanem a teljes populációt vizsgáljuk adott évfolyamokon). Az OKM-ben használt súlyozás nem a mintavétel és a teljes populáció közötti eltérést korrigálja, csupán a mérésből kimaradt, aktuálisan hiányzó tanulók hiányát pótolja. A mérés hibájának becslésére bootstrap algoritmust választottunk, azonban figyelembe kellett vennünk a populáció rétegzettségét is: a tanulók osztályokba, azok feladatellátási helyekbe, azok pedig iskolákba tartoznak.

A bootstrap mintavételek alapjául választhattuk volna a teljes populációt, de mivel feladatellátási helyek szintjén is visszajelzés készül a FIT-jelentések formájában (Balázsi és mtsai., 2017), és egy teljes populációs bootstrap minta esetén előfordulhat, hogy egy adott feladatellátási hely egyetlen tanulója sem kerül bele egy részmintába, így a feladatellátási hely eredményének hibabecslése sem lehet sikeres. Ebből következik, hogy a mintavétel alapja nem lehet nagyobb, mint a feladatellátási hely szintje, azaz minden feladatellátási helyhez saját mintavételi „doboz” tartozik, melyből az adott bootstrap mintába tartozó tanulókat választjuk ki.

A fenti módszerből következik, hogy a hibabecslés alapját jelentő bootstrap mintákban egy tanuló egy (vagy több) másik tanulót (vagy tanulókat) helyettesít. Tehát akkor járunk el legjobban, ha egy tanuló helyettesítésére a hozzá – környezeti tényezőkben – leghasonlóbb tanulók közül választunk. Ez az oka annak, hogy az újramintavételezés alapja az osztály, mert így a képességpontokból számított eredmények hibabecsléséből a lehető legjobban kizártuk az iskolától függő, mindenkire egyformán ható, a képességpontok összefüggését leginkább okozó hatást.

*A fenti eljárásokat súlyozott átlag számítási eljárásoknak is fel lehet fogni: az eredeti átlagszámítás esetében minden elem súlya 1. Ezzel szemben a jackknife eljárásban egy olyan súlyozott átlagot tekintünk, ahol egyetlen mintaelem súlya 0, minden más mintaelemé 1. De megtehetjük, hogy csak annyit kötünk ki: minden mintaelemnek lehet tetszőleges súlya úgy, hogy a súlyok összege n legyen. A bootstrap esetében olyan súlyozást használunk, ahol a súlyozás alapja az osztály, és a súlyok egész számok: ha egy osztály létszáma M, akkor az adott bootstrap mintában az adott osztályba járó gyerekek súlyainak összege M lesz (tehát „egymást helyettesítik", „egymás értékeivel keverednek össze" egy-egy átlag számításakor).*

Mivel a Kompetenciamérés adatait, eredményeit kutathatóvá szerettük volna tenni, ezért nyilvánvaló volt, hogy a számítások ismételhetővé kellett, hogy váljanak. A bootstrap eljárás lényege éppen a véletlen, azonban ha minden hibabecslés esetén új bootstrap mintákat választottunk volna, akkor az eredmények sem lettek volna ismételhetők, és a számításigény is túl nagy lett volna.

Ezért azt a megoldást választottuk, hogy egy alkalommal kiválasztunk 101 bootstrap mintát, majd a tanulói adatbázissal együtt elmentjük a mintavétel eredményét, a tanulói súlyokkal módosítva. Az előző részben jeleztük, hogy 200 minta már elegendő. A kompetenciamérés esetében a minták alacsony számát kapacitási korlátok indokolták. Ettől a pillanattól fogva bármely statisztika és a hozzá tartozó hibabecslés ugyanúgy, az alábbi sematikus számítás szerint történik:

- kiszámítjuk a paraméter becslését az eredeti súlyozással
  - Tehát a diákok eredeti súlya szerint (egy-egy osztályban/telephelyen/intézményben a hiányzók számával kell korrigálni a diákok értékeit)
- kiszámítjuk ugyanezt a becslést minden bootstrap minta esetén
  - Ezt foghatjuk fel úgy, hogy a különböző osztályokat „megkeverjük", belőlük újabb és újabb replikáns paraméterket hozva létre
- a részminták paraméterbecsléseiből kiszámítjuk a standard hibát
  - Vagyis kiszámítjuk a becslések szórását
- a mintákhoz tartozó 101 becslést nagyság szerint sorba rendezzük
  - Miként azt korábban ismertettük a bootstrap eljárás leírásakor
- a 2. és 3. becslés átlaga képezi a 2.5%-os percentilist, azaz a 95%-os konfidencia intervallum alsó határát
  - Hiszen ez felelne meg például egy 200-as minta esetében az 5. értéknek
- a 99. és 100. becslés átlaga képezi a 97.5%-os percentilist, azaz a 95%-os konfidencia intervallum felső határát.
  - Hasonlóan, mint az alsó konfidencia-intervallum határnál.

Látható, hogy a paraméterek meghatározásában, a hibabecslések megvalósításban hozott döntések folyamatosan praktikus, alapvetően takarékossági okokra vezethetők vissza, mellette azonban végig szem előtt tartottuk, hogy a lehető legjobban illeszkedjen a feladathoz és a mérés céljához, ugyanakkor fejleszthető, szükség szerint módosítható is maradjon az eljárás.

Az alábbi táblázatban néhány számítási eredményt láthatunk arra vonatkozóan, hogy 6. évfolyam esetében matematika és szövegértés területen, az átlagos teljesítmény becslésekor milyen eltéréseket tapasztalhatunk a különböző módszerek között. Az átlagos teljesítmény mindhárom módszer esetében azonos – azonban a hibák között már jelentős eltérések lehetnek. Minden esetben az IBM SPSS® program eljárásait alkalmaztuk. Első esetben a már ismertetett zárt formulát. Második esetben az IBM SPSS® programcsomag beépített bootstrap becslését. Harmadik esetben a saját eljárásunkat – mely az osztályok szintjére súlyozott és iskolai sajátosságokat figyelembe vevő bootstrap eljárás becslése.

1. táblázat. Teljesítmény változók standard hibái, átlagra: 6. évfolyamon

| Teljesítmény | Zárt formula | Bootstrap | OKM eljárás |
|---|---|---|---|
| Matematika, országos | 0,637 | 0,629 | 0,429 |
| Szövegértés, országos | 0,682 | 0,683 | 0,545 |
| Matematika, lányok | 0,865 | 0,941 | 0,683 |
| Szövegértés, lányok | 0,952 | 1,021 | 0,732 |
| Matematika, fiúk | 0,934 | 0,966 | 0,720 |
| Szövegértés, fiúk | 0,960 | 0,949 | 0,823 |

Megfigyelhető, hogy nagyságrendileg a beépített bootstrap eljárás és a zárt formula hasonló értékeket ad (bár általában a bootstrap nagyobb hibát számít), azonban ez sem feltétlenül teljesül (lásd: fiúk szövegértés teljesítménye). Azonban az osztályonkénti súlyozás melletti bootstrap eljárás az, amit elsősorban most mint hibát elfogadhatunk – hiszen itt a strátumok azok, melyek a diákok közötti összefüggőséget leginkább át tudják ültetni a számítások szintjére.

A fenti eljárás már csak azért is indokolt, mert az IBM SPSS® program beépített bootstrap eljárása nem működőképes akkor, ha a diákokat a már említett súlyokkal látjuk el – azaz, az eredeti program esetleg alkalmazható variánsa jelen helyzetben technikailag nem megoldható.

## KITEKINTÉS, ALKALMAZÁS

A jelen tematikus számban az alábbi alkalmazásokat láthatjuk a fenti módszertanra. Nyitrai és munkatársainak két cikkében a szülői bevonódás vizsgálatában (Nyitrai és mtsai., 2019a), illetve (Nyitrai és mtsai., 2019b) alkalmazásra kerültek a teljesítményekre, valamint az esetszámokra vonatkozó konfidencia intervallumok meghatározásai. A különböző módozatú bevonódások esetében más és más teljesítmény mutatók, illetve diákokra vonatkozó, becsült esetszámok, arányok adódtak. Szintén a szülői bevonódás – de az iskolai életbe, nem a közvetlen tanulási metódusokba – kérdését taglalták Koltói és munkatársai (Koltói és mtsai., 2019a), illetve (Koltói és mtsai., 2019b). Ebben az esetben szintén az esetszámokra és a két tudásterületi kompetenciában nyújtott teljesítményekre vonatkozó megbízhatósági tartományokat számítottuk ki.

A szülői bevonódás mellett vizsgáltuk az iskolai élet különböző aspektusait is – a családi háttér egészére koncentrálva. Harsányi és munkatársainak tanulmányai (Harsányi és mtsai., 2019a), illetve (Harsányi és mtsai., 2019b) a családi háttér hatásának aspektusainál a különböző családi körülmények között élő diákok arányainak meghatározása mellett e csoportok teljesítményének becslését is a fenti módszerekkel vizsgálták.

További két anyag különböző diákcsoportok teljesítményét helyezte fókuszba úgy, hogy ezen specifikus csoportok arányait, esetszámait is becsülték. Egyik csoport a rendszeresen sportolók csoportja, akiket Smohai és munkatársai vizsgáltak (Smohai és mtsai., 2019a), valamint (Smohai és mtsai., 2019b) tanulmányukban. Emellett az SNI és BTM diákok arányával, valamint teljesítményével foglalkoztak Kövesdi és munkatársai (Kövesdi és mtsai., 2019), illetve Kovács és munkatársai (Kovács és mtsai., 2019) tanulmányaikban.

## BIBLIOGRÁFIA

Auxné, Bánfi, I., Balázsi, I., Balkányi, P., Gyapay, J., Lak, Á. R., Ostorics, L. I., … Vadász, C. (2015). *Országos kompetenciamérés – Technikai leírás*. Oktatási Hivatal.

Balázsi I., T. Kárász J., Lak Ágnes R., Ostorics L., Szabó Lívia D., & Vadász C. (2017). *Országos kompetenciamérés – 2016* (o. 106).

Bruce Turner, D. (1986). Comparison of Three Methods for Calculating the Standard Deviation of the Wind Direction. *Journal of Climate and Applied Meteorology*, *25*(5), 703–707. https://doi.org/10.1175/1520-0450(1986)025<0703:COTMFC>2.0.CO;2

Chan, T. F., Golub, G. H., & Leveque, R. J. (1983). Algorithms for Computing the Sample Variance: Analysis and Recommendations. *The American Statistician*, *37*(3), 242–247. https://doi.org/10.1080/00031305.1983.10483115

Efron, B., & Gong, G. (1983). A Leisurely Look at the Bootstrap, the Jackknife, and Cross-Validation. *The American Statistician*, *37*(1), 36–48. https://doi.org/10.2307/2685844

Fay, R. E. (1984). Some Properties of Estimates of Variance Based on Replication Methods. *American Statistical Association*, 495–500.

Joanes, D. N., & Gill, C. A. (1998). Comparing Measures of Sample Skewness and Kurtosis. *Journal of the Royal Statistical Society. Series D (The Statistician)*, *47*(1), 183–189. Elérés forrás JSTOR.

Lehmann, E. L., & Casella, G. (1998). *Theory of point estimation* (2nd ed). New York: Springer.

Miller, R. G. (1974). The Jackknife--A Review. *Biometrika*, *61*(1), 1–15. https://doi.org/10.2307/2334280

Neyman, J. (1937). Outline of a Theory of Statistical Estimation Based on the Classical Theory of Probability. *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences*, *236*(767), 333–380.

Rutkowski, D. (2015). The OECD and the local: PISA-based Test for Schools in the USA. *Discourse: Studies in the Cultural Politics of Education*, *36*(5), 683–699. https://doi.org/10.1080/01596306.2014.943157

Takács S. (2010). Egy nem hagyományos statisztikai eljárás bemutatása az OECD Pisa adatbázison – Esettanulmány. *Alkalmazott Matematikai Lapok*, *27*(1), 157–174.

Tukey, J., W. (1958). Bias and Confidence in Not-Quite Large Sample. *Annals of Mathematical Statistics*, *29*, 614.

E tematikus szám hivatkozott cikkei

Harsányi Szabolcs Gergő, Koltói Lilla, Kovács Dóra, Kövesdi Andrea, Nagybányai-Nagy Olivér, Nyitrai Erika, Simon Gabriella, Smohai Máté, Takács Nándor, Takács Szabolcs: Születni tudni kell? – Az Országos kompetenciamérés eredményeinek vizsgálata a szülők munkájának rendszeressége, észlelt társadalmi helyzet és a lakókörnyezet vonatkozásában

Harsányi Szabolcs Gergő, Koltói Lilla, Kovács Dóra, Kövesdi Andrea, Nagybányai-Nagy Olivér, Nyitrai Erika, Smohai Máté, Simon Gabriella, Takács Nándor, Takács Szabolcs: The relationship of school achievement with parents' employment status, perceived social status, and living environment as reflected in findings of the 2017 National Assessment of Basic Competencies (NABC)

Koltói Lilla, Harsányi Szabolcs Gergő, Kovács Dóra, Kövesdi Andrea, Nagybányai-Nagy Olivér, Nyitrai Erika, Simon Gabriella, Smohai Máté, Takács Nándor, Takács Szabolcs: A szülők tanulmányokba való bevonódásának összefüggése az iskolai teljesítménnyel

Koltói Lilla, Harsányi Szabolcs Gergő, Kovács Dóra, Kövesdi Andrea, Nagybányai-Nagy Olivér, Nyitrai Erika, Simon Gabriella, Smohai Máté, Takács Nándor, Takács Szabolcs: The relationship between school achievement and paternal involvement in children's school activities as judged by headmasters in the 2017 National Assessment of Basic Competencies (NABC)

Kovács Dóra, Kövesdi Andrea, Harsányi Szabolcs Gergő, Koltói Lilla, Nagybányai-Nagy Olivér, Nyitrai Erika, Simon Gabriella, Smohai Máté, Takács Nándor, Takács Szabolcs: Results of the 2017 National Assessment of Basic Competencies (NABC) –amongst 6th 8th and 10th grade pupils diagnosed with SEN and BTM disorder

Kövesdi Andrea, Kovács Dóra, Harsányi Szabolcs Gergő, Koltói Lilla, Nagybányai-Nagy Olivér, Nyitrai Erika, Simon Gabriella, Smohai Máté, Takács Nándor, Takács Szabolcs: A 2017. évi Országos kompetenciamérés eredményei Magyarországon- Az SNI és BTM-mel diagnosztizált 6, 8, 10-ik évfolyamos gyermekek körében

Nyitrai Erika, Harsányi Szabolcs Gergő, Koltói Lilla, Kovács Dóra, Kövesdi Andrea, Nagybányai-Nagy Olivér, Simon Gabriella, Smohai Máté, Takács Nándor, Takács Szabolcs: Iskolai teljesítmény és szülői bevonódottság

Nyitrai Erika, Harsányi Szabolcs Gergő, Koltói Lilla, Kovács Dóra, Kövesdi Andrea, Nagybányai-Nagy Olivér, Simon Gabriella, Smohai Máté, Takács Nándor, Takács Szabolcs: Relations between Parental

Involvement and School Performance in the Light of Data from National Assessment of Basic Competencies (NABC) 2017

Smohai Máté, Harsányi Szabolcs Gergő, Koltói Lilla, Kovács Dóra, Kövesdi Andrea, Nagybányai-Nagy Olivér, Nyitrai Erika, Simon Gabriella, Takács Nándor, Takács Szabolcs: Az Országos kompetenciamérés során felvett szabadidős sporttevékenységre irányuló adatok elemzése és standardjainak közlése

Smohai Máté, Harsányi Szabolcs Gergő, Koltói Lilla, Kovács Dóra, Kövesdi Andrea, Nagybányai-Nagy Olivér, Nyitrai Erika, Pulai-Kottlár Gabriella , Simon Gabriella, Takács Nándor, Takács Szabolcs: Relationship between leisure & sports activities and school performance in the light of data from the National Assessment of Basic Competencies (NABC)2017 survey