10.12663/PsyHung.2.2014.1.4

EGY EGYSZERŰ, ÚJFAJTA KERESZT-VALIDÁLÁSI ELJÁRÁS BEMUTATÁSA

Takács Szabolcs[1], Kárász Judit[2]

[1] Károli Gáspár Református Egyetem, Pszichológiai Intézet, Általános Lélektani és Módszertani Tanszék, egyetemi tanársegéd; Budapest Főváros Kormányhivatala Munkaügyi Központ, Szervezési Osztály, statisztikai főreferens, email: tretarkhon@gmail.com

[2] Oktatási Hivatal Mérésértékelési Központ, statisztikai elemző

Absztrakt

Sűrűn előfordul kutatásokban, vizsgálatokban, hogy olyan kérdőívekkel szeretnénk dolgozni, amelyeket még nem tudtunk korábban bevizsgálni, tesztelni. Arra a célra, hogy kapott válaszaink megbízhatóságát a lehetőségekhez képest mégis a legjobb módon – ugyanakkor egyszerűen – tudjuk biztosítani, nincs túl sok módszerünk.
A most bemutatásra kerülő eljárásban olyan lehetőséget vázolunk, melyben apróbb, egyenként egyszerű ellenőrző rutinok felhasználásával számos ponton biztosíthatjuk az általunk mért adatok biztonságát, stabilitását – növelve ezzel elemzéseink, eredményeink hitelességét.
Lényegét tekintve a *split-half* eljárást és a *bootstrap* algoritmust kombináltuk egyetlen eljárásban, s így olyan vizsgálati eszközhöz jutottunk, mely két egyszerű részeljárás elsajátítása után bárki által könnyen alkalmazható, akár az általánosításokat is beleértve.
A dolgozatban annak megállapítására kívánunk egyszerű megoldást adni, hogy egy kérdőíves felmérést kitöltők közül az elsők válaszai (vagy akár nyugodtan általánosíthatjuk oly módon is az eljárást, hogy a kitöltők tetszőleges hányadának válaszai) és a további kitöltők válaszai között detektálhatók-e eltérések, vagy a kérdőívet kitöltők e tekintetben homogénnek tekinthetők?

Kulcsszavak: Kereszt-validitás ▪ split-half ▪ bootstrap

Abstract

The problem of working with untested questionnaires for establishing research is well known: the possibilities for verifying the validity of the collected data is limited.
In this paper we propose an algorithm constructed from multiple simple steps that helps to establish the validity and stability of the results.
We combine the bootstrap and the split-half methods in a simple way allowing the algorithm to be applied in many situations.
The primary concern of the paper is to check whether the characteristics of the first responses are the same as those of later responses, or some differences can be detected among them.

Keywords: Cross-validity ▪ Split-half ▪ bootstrap

BEVEZETŐ

A kérdőívek minőségbiztosítási folyamata olykor hosszadalmas, fáradságos munka, mely nem is mindig valósítható meg egyetlen felmérés során (kérdések újrafogalmazása, hibás, rosszul mérő itemek elhagyása, jól mérő tesztekkel való összehasonlítások stb.).

Cikkünkben egy olyan szituációt szeretnénk bemutatni, mely hasonló helyzetekre tetszés szerint általánosítható:

- Szeretnénk egy kutatáshoz gyors vizsgálatot végezni (például internetes felmérés);
- Saját magunk megfogalmazunk kérdéseket, melyek szakmai hozzáértésünk szerint egy (vagy akár több) jelenséget vizsgálnak;
- A felmérést idő vagy egyéb források hiányában egyetlen alkalommal tudjuk lebonyolítani;
- A kérdések átfogalmazására vagy egyéb, strukturális változatásra nincsen módunk – abból a kérdéssorból kell gazdálkodnunk, amit az adott idő alatt sikerült lekérdezni.

Ilyen helyzetek gyakran előfordulhatnak – nem csak egy kisebb kutatás során, hanem akár nagyobb volumenű kutatásokban is, amikor bármely váratlan, előre ki nem számítható helyzet miatt kénytelenek vagyunk gyors kérdőíves felméréssel információhoz jutni.

Ilyen esetekben, hacsak nem áll rendelkezésünkre éppen a problémára adaptált már bemért kérdőív, kénytelenek vagyunk magunk összeállítani a kérdéseket. Az esetek döntő többségében nem várható el, hogy validált, többszörösen bevizsgált és éppen a számunkra fontos specifikus populációra is alkalmazható kérdéssorral rendelkezzünk.

A következő fejezetekben bemutatunk egy könnyen, gyorsan adaptálható eljárást. A folyamat során egy általános validálási eljárás lépéseit fogjuk alkalmazni – nem szokványos módon. Egy validálási eljárás lépéseit több forrásból is áttekinthetjük: egy új hangulati kérdőív validálásának mozzanatait ismerhetjük meg például Halmai és munkatársai dolgozatában (Halmai, 2008), melynek eljárásait, illetve mutatóit adaptálni tudjuk a fent vázolt helyzetekre.

AZ ELJÁRÁSBAN HASZNÁLT MUTATÓK

A most bemutatásra kerülő eljárásban több statisztikai mutatót is használhatunk, melyeknek mind specifikus funkciója lesz. Bizonyos szabadságot tudunk magunknak biztosítani továbbá annak érdekében, hogy maga a módszer adott keretek között, de tetszőlegesen legyen módosítható. Erre azért van szükség, mert szeretnénk olyan algoritmust vázolni, amely tág körben használható és alkalmazható – így biztosítva azt, hogy változatos körülmények között tudjunk magunk számára könnyen adaptálható és formálható, mégis biztonságot garantáló eszköztárt kialakítani.

STABILITÁSI MUTATÓK

A kérdőíves felmérések során gyakori elvárás, hogy bizonyos indexeket, skálákat alakítsunk ki az általunk feltett kérdésekből. Ennek egyik leggyakrabban használt, általánosan elfogadott mutatója a Cronbach által megalkotott alfamutató (lásd például (Cronbach, 1951).

Ennek a mutatónak számtalan értelmezési lehetősége van, melyekre nem is szándékoznánk kitérni. Amit fontos az eljárás szempontjából figyelembe vennünk: elfogadott, hogy azon indexek vagy skálák, amelyek legalább 0,7-es alfa mutatóval bírnak, megbízhatónak, stabilnak tekinthetők[1].

Ebben a mutatóban sok mozgásterünk nincsen: széles körben elfogadott, sőt elvárt mutató, mely nélkül tesztünk megbízhatóságát nehézkes lehet elfogadtatni.

A soron következő fejezetben bemutatunk néhány olyan mutatót, eljárást, számítást, melyek szintén fontosak lehetnek egy-egy kérdőív stabilitását illetően – és amelyek stabilitási vizsgálata fontos szempont lehet.

Középértékek és szóródási mutatók

A különböző módokon[2] kialakított skálák (vagy akár az itemek) középértékeit és szóródási mutatóit nem csak jellemzési szempontból lehet vizsgálni. Ezen azt értjük, hogy e mutatóknak alapvetően kétfajta jelentése lesz e folyamatban – és érdemes is lehet kihasználni ezt a kétfajta szerepet.

Egyik oldalról a skáláinkat, indexeinket jellemzik a kialakított középértékek (pl. átlag, medián, módusz), illetve szóródási mutatók (pl. szórás, átlagos abszolút eltérés, terjedelem). Ezek a mutatók mind hasznos információval szolgálnak a skáláink vizsgálatakor.

[1] Megjegyeznénk, hogy a 0,7-es érték amolyan ökölszabály: elfogadható lehet bizonyos helyzetekben ennél kisebb érték is (meglátásunk szerint 0,5-nél kisebb értékek már valóban megbízhatatlanok); illetve a nagyon magas alfa értékek (0,9 felett) egyfajta redundanciát is jelezhetnek (a skálában szereplő egyes itemek többször vannak figyelembe véve, többször kérdeztünk rá azonos jelenségekre).

[2] Különböző módnak gondolhatjuk az átlagolást, összegzést, vagy akár súlyozott átlagolást vagy összegzést. A bemutatásra kerülő módszer szempontjából lényegtelen kérdés, hogy milyen módszerrel alakítjuk ki a skáláinkat.

Másik oldalról azonban a skáláink stabilitását jellemezheti, hogy a vizsgált, számunkra fontosnak ítélt középértékek és szóródási mutatók (hiszen nem minden indexhez vagy skálához használunk minden mutatót) milyen módon ingadozhatnak egy – vagy akár több, további – újonnan elkészített felmérésben, vagy akár a felmérésünk időbeni lefolyása során (komoly gondot okozhat, ha például a kérdőívet először kitöltők lényegesen eltérő mutatókkal, paraméterekkel jellemezhetők, mint a későbbi válaszadók).

Kapcsolat a középértékek és a Cronbach-alfa között

A Cronbach által kifejlesztett (Cronbach, 1951) alfa-mutató az összes tesztfelezéssel kapott korrelációs együttható számtani átlaga – így lényegében egyfajta belsőkonzisztencia-mutatóként fogható fel.

A középértékektől és szóródási mutatóktól ezt a fajta stabilitást szintén elvárhatjuk: ha újra és újra felmérnénk a vizsgált populációt, akkor az általunk fontosnak ítélt paraméterek nem ingadozhatnak tetszőleges mértékben – egy bizonyos, általunk is elfogadható tartományon belül kellene maradniuk.

ELOSZLÁSVIZSGÁLATOK

Miután a középértékekkel és a szóródási mutatókkal az egyes vizsgálati populációk vagy részpopulációk centruma és variabilitása górcső alá került, vizsgálhatjuk az egész populációnkban vagy részpopulációnkban a változó egészének viselkedését is.

A vizsgált itemek valamilyen előre megadott módon való összegzéséből skálákat, indexeket alkotunk. Matematikailag igazolható, hogy jól definiált és jól specifikált adatok összegzéséből, átlagolásából a normális eloszlást jól közelítő véletlen változók alkothatók.

Ennek technikai megvalósítását olvashatjuk például Rowe munkájában (Rowe, 2002), mely dolgozatban többek között normális és egyenletes eloszlás generálására alkalmas technikai lépéseket, algoritmusokat ismertet a szerző, annak matematikai hátterével és viselkedésével egyetemben.

Érthető tehát, hogy az eloszlástól elvárhatunk egyfajta stabilitást – ugyanolyan módon, mint az első két mutató esetén. Elfogadható azonban az is, hogy ez a vizsgálat nem feltétlenül *szükséges* a számunkra, ugyanis az eloszlást viszonylag ritkán teszteljük utólagosan (azaz a kérdőív használatakor).

Egy kérdőív használatánál a belső stabilitás (például minden megkérdezett személy várhatóan ugyanazt értse egy-egy kérdésen) és utána a középértékek, szóródási mutatók stabilitása elengedhetetlen (például a vizsgált populáció tagjainak átlagos és átlagostól eltérő viselkedésének jellemzése). Az eloszlás pontos ismerete nélkül is[3] számos hasznos információra tehetünk szert, így ezt a bonyolultabb vizsgálatot jelen cikkünkből kihagyjuk[4].

BELSŐ ÖSSZEFÜGGÉSEK

A középértékek, szóródási mutatók és az egész eloszlás vizsgálata után a kérdőív különböző itemeinek, vagy még inkább skáláinak egymáshoz való viszonya is vizsgálatunk tárgyát képezheti – annak stabilitása is elengedhetetlen lehet egy-egy kérdőív állandóságának, megbízhatóságának vizsgálatakor.

A kérdőívben feltett kérdések egymáshoz való viszonya, viselkedése így szintén fontos lehet. A Cronbach-alfa vizsgálat egyfajta szempontból teszteli ezt a jelenséget, tehát e mostani lépést szintén ki is hagyhatjuk akár az eljárásból – azonban egyszerűsége okán mégis bemutatjuk annak érdekében, hogy az eljárás általánosításának módszertanát jobban láthatóvá tegyük.

A feltett kérdések közötti korrelációs együtthatók (Pearson-féle, vagy akár egyéb monotonitási együtthatók, mint a Spearman-féle, vagy a Kendall-féle mutatók) szintén számíthatók – és e mutatóktól is elvárható az a fajta stabilitás, mint a már említett mutatók, paraméterek, jellemzők esetén.

A Pearson-féle korreláció elmozdulásáról, viselkedéséről, illetve jelentéstartalmáról és használhatóságáról folyamatosan változott a felfogás és hozzáállás (bár természetesen bizonyos stabil, általánosan elfogadott álláspontok megmaradtak). Erről a folyamatról ír például Aldrich, dolgozatában arról értekezve, hogy a Karl Pearson eredeti munkájában megfogalmazott jelenségtől Yule munkásságáig milyen átalakulásokon, változásokon és fejlődésen ment keresztül a korrelációs együtthatók alkalmazása (Aldrich, 1995).

Azért fontos, hogy tisztában legyünk a különböző kapcsolati mutatókkal, mert ahogy a középértékeknél és a szóródási mutatóknál, úgy az összefüggések vizsgálatánál is elengedhetetlen, hogy a változók viselkedéséhez leginkább illeszkedő kapcsolati mutatókat válasszuk, és teszteljük azok stabilitását.

Mint az majd e dolgozatban felhozott példában is látható lesz, eljárásunk alkalmas arra, hogy bármely, szabadon (és megfelelően, a változókhoz jól illeszkedő módon) kiválasztott kapcsolati mutató tesztelhető legyen, így annak stabilitását szintén vizsgálni tudjuk.

---

[3] A következő megfontolás áll a mögött, hogy magát az eloszlást nem feltétlenül kell vizsgálnunk. A kérdőív bevizsgálása után egy újabb személy által kitöltött kérdőív esetén a belső stabilitást és az átlagos vagy átlagostól eltérő viselkedést jogosan értelmezzük és vesszük figyelembe – de az eloszlást ilyen helyzetben nem tudjuk, és nem is nagyon szándékozunk vizsgálni. Ettől függetlenül – mint majd látható lesz – az eljárásban akár ez a szempont is vizsgálhatóvá válik, de ezt nem érezzük szükségszerűnek, létfontosságúnak – ahogy például ezt a vizsgálatot Halmai és munkatársai is mellőzik már említett validálási tanulmányukban (Halmai és munkatársai, 2002).

[4] Az eljárás általános voltának köszönhetően az eloszlások vizsgálatára is adaptálható.

AZ ELJÁRÁS LÉPÉSEI

A fenti szempontok alapján – mely tehát egy általános eljárást vázol – egy konkrét vizsgálat megtervezésének általános lépéseit, algoritmusát szeretnénk bemutatni.

Tételezzük fel, hogy kérdőíves felmérést helyezünk el az interneten, melynek kitöltésére egy hét áll a megkérdezettek rendelkezésére. Feltehető, hogy a kérdőív nem túl hosszú (bár az eljárást semmilyen módon nem befolyásolja a kérdőív hossza), gyorsan kitölthető – a megkérdezettek köre széles, és biztosítani tudjuk a folyamatos felmérést.

1. Készítsük elő a kérdőívet úgy, hogy a **kitöltés idejét**, illetve a kitöltés sorrendjét menteni tudjuk – a feldolgozás során ez kulcsfontosságú.
2. Az első 1-2 napon[5] kitöltött kérdőíveket mentsük el.
3. A **számunkra fontos paramétereket (mutatókat)** számítsuk ki a meglévő kérdőívekből (például: Cronbach-alfa, átlag, szórás).
4. *Folytatódik a kitöltés.*
5. A kitöltés második felében beérkező kérdőívekből a lezárást követően **véletlenszerűen**[6] válasszunk ki tetszőleges számú (legalább 200), **nagyobb elemszámú (50-100 fő itt is javasolt)** részmintát.[7]
6. A részminták **mindegyikén** számítsuk ki a korábbi mutatókat.
7. Így a számított mutatók olyan mintáját nyerjük, melyen vizsgálható, hogy az eredetileg kiszámított értékektől eltérnek-e vagy sem a randomizálásból származó minták mutatói.

Ezen a ponton döntést lehet hozni:

– amennyiben létrejött szignifikáns eltérés, úgy az a helyzet állt elő, hogy a kérdőívet kitöltők első hányada (vagy tetszőleges, véletlen része), illetve a válaszadók második (másik) hányada nem rendelkezik ugyanazokkal a jellemzőkkel. *Ez azt jelenti, hogy a kérdőívünk vélelmezhetően nem megfelelő, a különböző paraméterek eltérnek, változnak – instabil kérdőívvel dolgozunk, abból messzemenő következtetések levonása veszélyes;*

– amennyiben a paraméterek nem változnak szignifikánsan, úgy elmondható, hogy az első és második szakaszban kitöltők összességében homo-

---

[5] Az időbeli korlátozás helyett dönthetünk úgy is, hogy egy adott darabszámú kitöltő válaszait várjuk meg. A különböző mutatók számítási stabilitása miatt érdemes legalább 50-100 kitöltő adatait megvárni, ha erre módunk van. Amennyiben ennél kevesebb kitöltővel tudunk csak dolgozni, úgy a számított paraméterek, mutatók becslésének hibája drasztikusan befolyásolhatja módszerünk működését, tehát a módszer alkalmazásához szükséges, hogy kellően nagy mintával rendelkezzünk (100 fő felett a módszer már alkalmazható).

[6] A véletlenszerű kiválasztás módszeréről és egy tetszőleges paraméter számításáról az 1. függelékben található SPSS-ben megírt SYNTAX leírásában ejtünk szót, illetve megadunk erre vonatkozóan egy lehetséges megvalósítási metódust.

[7] 30-50 részminta akár elég is lehet, de Efron és Gong szerint 200-nál nagyobb részminta már bizonyosan elegendőnek tekinthető (Efron és Gong, 1983).

gén módon töltötték ki a kérdőívet, *a számunkra fontos és későbbiekben használt mutatók stabilnak mondhatók.*

A különböző tesztek gyorsak és egyszerűek, így igen kényelmes eljárást tudunk készíteni annak érdekében, hogy szinte bármely mutatót meg tudjunk vizsgálni, ami számunkra fontos lehet a kérdőív alkalmazása során.

Így még különböző statisztikai teszteket, próbákat sem kell alkalmaznunk, hiszen az úgynevezett **bootstrap eljárással**[8] bármelyik paramétert tesztelni tudjuk anélkül, hogy statisztikai próbáink alkalmazási feltételeit biztosítanunk kéne.

Tegyük fel, hogy van időnk, illetve lehetőségünk a kérdőív második feléből egy 200-as részmintát készíteni.

– **Rögzítsük az eredetileg** kiszámított, vizsgálandó **mutatót** (T);
– **Minden részmintából** számított mutatót rendezzünk nagyság szerinti sorrendbe;
– A 200 számított mutató nagyság szerinti sorrendjének 5. és 195. tagja (legyenek ezek *P(5)* és *P(195))* határozzon meg egy intervallumot.

Ekkor a következő döntést kell meghozni: amennyiben az eredetileg vizsgált paraméter (T) ebben az intervallumban van, ha tehát

$$P(5) < T < P(195),$$

úgy elmondható, hogy a paraméter 95%-os valószínűséggel nem mozdult el szignifikánsan a kérdőív két kitöltési szakaszában.

Ellenkező esetben viszont, ha tehát $T < P(5)$ vagy $T > P(195)$, akkor szignifikáns eltérés tapasztalható az eredetileg számított paraméterhez képest a randomizált mintákban, azaz instabil a T paraméter.

EGY PÉLDA AZ ELJÁRÁS ALKALMAZÁSÁRA

Hogy könnyebben átlátható legyen a fenti algoritmus, egy példával illusztráljuk eljárásunkat[9]. A soron következőkben egy olyan példát vizsgálunk, ahol matematikai területeken mértük fel diákok teljesítményét, majd azokból egy „összesített" matematikateljesítményt számítottunk.

A kérdésünk az, hogy az „összevonás", az összesített matematikai jártasság elkészítése jogos-e a különböző területeken mért részeredményekből, vagy ezek a vizsgált területek olyannyira különbözőek, hogy őket egyetlen mutatóban összevonni vétség lenne?

A 4 vizsgált terület: *tér és alakzat; változás és reláció; bizonytalanság;*

[8] Az algoritmus 5-7. lépéseinek együttese tekinthető a bootstrap eljárásnak.

[9] Az eredeti adatok az OECD PISA 2003 (OECD, 2003) felmérésből származnak, mely adatokat nem az eredeti kiértékelési módszer szerint dolgozzuk fel.

mennyiség. Hogy jobban lássuk, milyen területeket lehetne ide sorolni, vegyünk 1-1 példát az adott területhez:

- *(1): Tér és alakzat*: e feladatok esetén gondolhatunk olyan példára, hogy egy közterület díszburkolattal való fedéséhez mennyi burkolóanyagra van szükség, ha megadjuk a tér szélességét és hosszúságát, illetve kikötjük azt is, hogy a tér téglalap alakú.
- *(2): Változás és reláció*: olyan példára érdemes itt gondolni, melyben egy gyermek 3 éves és 6 éves kori testmagasságát kell összehasonlítanunk, és kíváncsiak vagyunk rá, hogy mennyivel nagyobb ruhákat kell vásárolni a gyermeknek, vagy akár (ha a ruhák méretétől függ azok ára) mennyivel drágább az öltöztetése.
- *(3): Bizonytalanság*: klasszikus statisztikai feladatokat lehet ide sorolni – egy diagramról le kell olvasni adatokat, vagy éppen fordítva: adott adatokhoz kell diagramot szerkeszteni.
- *(4): Mennyiség*: tegyük fel, hogy egy tanuló osztálykirándulásra külföldre utazik és a szülőknek valutát kell váltaniuk – de a diákoknak csak a helyi árviszonyokat tudták megmondani (a szállás ára euróban). Meg kell mondani, hogy az milyen mértékben terheli a családi költségvetést, ha az eurót adott árfolyamon tudják átváltani.

A változókon elért teljesítményekből szeretnénk egy együttes teljesítményértéket készíteni – amennyiben a vizsgálataink azt mutatják, hogy ezek a területek összevonhatók egy matematikateljesítményt mérő skálaváltozóban.

Vizsgálandó értékek:

– A 4 változót összefogó *Cronbach-alfa* mutató;
– Az összesített változóra jellemző átlag és *szórás*;
– A 4 változó közötti *korrelációs együtthatók*;

Erre a 4 értékre a későbbiekben szükségünk lesz, ha teljesítményt szeretnénk mérni: a Cronbach-alfa mutató melletti korrelációs vizsgálatot kizárólag azért tartottuk indokoltnak, mert a Cronbach-alfa mutató értéke függ a skála tételeinek számától: a tételek számának növelésével növekszik a Cronbach-alfa értéke is (Cronbach, 1951).

Másik oldalról a belső összefüggéseket árnyaltabban tudjuk vizsgálni és tesztelni a korrelációs együtthatók segítségével.

Minden vizsgálatot 5%-os szinten végzünk (ez azt jelenti, hogy az 5. és a 195. értéket nézzük a nagyság szerinti sorba rendezésnél). Azaz az első 300 eset értékeit alapul véve az érdekel minket, hogy az utánuk kitöltők közül 100 fős részmintákat vizsgálva találunk-e szignifikáns elmozdulást bármelyik vizsgált paraméterben (200 darab véletlen mintát választottunk ki[10]).

[10] Azaz 200 darab 100 fős részminta eredményeit jegyeztük fel, és az összehasonlításhoz a bázisértékeket az első 300 kitöltő adatai szolgáltatták.

AZ EREDETI MUTATÓK ÉS PARAMÉTEREK BEMUTATÁSA

Az első 300 kitöltő eredményeit az alábbi táblázatokban foglaljuk össze.

MUTATÓ	MUTATÓ ÉRTÉKE (ELSŐ 300 KITÖLTŐ)
Cronbach-alfa	0,974
Képzett skála átlaga	483,99
Képzett skála szórása	89,39
Korreláció (1,2)	0,924
Korreláció (1,3)	0,916
Korreláció (1,4)	0,908
Korreláció (2,3)	0,925
Korreláció (2,4)	0,902
Korreláció (3,4)	0,902

1. táblázat Az első 300 kitöltő adataiból számított paraméterek

- A Cronbach-alfa értéket tehát a korábbiakban felsorolt 4 matematikai alterületre számított indexekre vizsgáltuk.
- Az átlagot és a szórást a 4 index átlagaként meghatározott matematika-teljesítmény-változóra határoztuk meg.
- A korrelációs együtthatók mögötti számok jelentése: a fenti felsorolásban a 4 matematikai terület kódját jelzi a tömörebb táblázat-kitöltés érdekében. A korrelációk minden esetben szignifikánsak lettek (minden esetben $p < 0{,}001$ szinten).

A RANDOMIZÁLT MINTA PARAMÉTEREINEK KONFIDENCIA-INTERVALLUMA

A randomizált értékek esetén tehát intervallumot határoztunk meg, így tesztelhetővé vált, hogy az első 300 kitöltő által számított paraméterek (mutatók) milyen mértékben mozdultak el – mennyire tekinthető stabilnak a fenti skálarendszer.

MUTATÓ	INTERVALLUM ALSÓ HATÁRA	INTERVALLUM FELSŐ HATÁRA
Cronbach-alfa	0,967	0,979
Képzett skála átlaga	478,32	504,71
Képzett skála szórása	80,27	97,59
Korreláció (1,2)	0,896	0,941
Korreláció (1,3)	0,886	0,94
Korreláció (1,4)	0,869	0,921
Korreláció (2,3)	0,891	0,94
Korreláció (2,4)	0,882	0,943
Korreláció (3,4)	0,889	0,941

2. táblázat A 200 darab 100 fős részmintából számított 95%-os konfidencia-intervallumok

A mutatókat azonos körülmények között, azonos beállításokkal számítottuk[11] az első 200 kitöltővel megegyező metodika szerint.

VÉGSŐ ADATOK, A TELJES KITÖLTÉS UTÁNI MUTATÓK ÉRTÉKE

A végső adatok során vizsgáltuk azt, hogy az előzetesen számított adatokhoz képest a későbbiekben randomizált minták értékei mutatnak-e valamifajta elmozdulást – illetve a végső lépésben az összes kitöltő által figyelembe vett értékeket is meghatároztuk.

Ezen utolsó lépésben ellenőrizzük, hogy kérdőívünk stabilitása megfelelő-e, illetve meghatározzuk azokat a mutatókat, paramétereket, amelyeket a teljes minta alapján végső paramétereknek tekinthetünk a kérdőív esetén:

MUTATÓ	INTERVALLUM ALSÓ HATÁRA	INTERVALLUM FELSŐ HATÁRA	VÉGSŐ PARAMÉTER	STABILITÁS
Cronbach-alfa	0,967	0,979	0,975	0
Képzett skála átlaga	478,32	504,71	488,6	0
Képzett skála szórása	80,27	97,59	89,66	0
Korreláció (1,2)	0,896	0,941	0,925	0
Korreláció (1,3)	0,886	0,94	0,916	0
Korreláció (1,4)	0,869	0,921	0,899	0
Korreláció (2,3)	0,891	0,94	0,925	0
Korreláció (2,4)	0,882	0,943	0,907	0
Korreláció (3,4)	0,889	0,941	0,907	0

3. táblázat Az első 300 kitöltő paramétereivel való összehasonlítás

A táblázat számadatai tehát a végső paraméter-értékeket tartalmazzák.

- A stabilitásnál lévő érték mutatja, hogy az adott paraméter benne volt a randomizálással meghatározott intervallumban (0), vagy elmozdult onnan.
- Amennyiben elmozdulás lenne, úgy a „+1”-es érték mutathatná, hogy az intervallum felső határa felett volt az első 200 kitöltő értéke (pozitív irányban tér el az első 200 kitöltő értéke a randomizált mintákhoz képest).
- A „-1”-es érték azt jelezné, hogy az intervallum alsó határa alatt található az első 200 kitöltő értéke – negatív irányban tér el az első 200 kitöltő a randomizált mintákhoz képest.

Esetünkben megfigyelhető, hogy minden mutató a randomizált minták intervallumában található (mind a teljes mintán számítva, mind az első 200 kitöltő által számított értékek esetén).

[11] A számításokhoz SPSS programcsomagot használtunk, de ahogy látszik a számításokból, lényegében bármely más programmal is megoldható lenne – tetszőleges táblázatkezelővel akár.

ÖSSZEFOGLALÁS

Az esettanulmány során olyan eljárást mutattunk be, mely alapvetően teljesen egyszerű részfeladatokból építkezik, mégis: ezeknek az egyszerű rutinoknak a kombinációja, egyszerre történő végrehajtása, tesztelése biztosíthatja számunkra azt, hogy egy kérdőív belső konzisztenciájának, kereszt-validálásának vizsgálati lehetőségét egyetlen mérésből megteremtsük.

Az alkalmazott rutinok végrehajtásához nincs szükség nagyobb számítógépes felszereltségre, hiszen mindezek a számítások egy táblázatkezelő segítségével is végrehajthatók – így egy internetes kérdőív felvétel során akár egyes ingyenesen elérhető, bármely operációs rendszerrel kompatibilis táblázatkezelőt felhasználva is biztosíthatjuk kérdőívünk stabilitását[12].

[12] Fontos kiemelni, hogy ez az eljárás nem helyettesíti (nem is helyettesítheti) az egyes validálási eljárások azon lépését, amikor más, már bemért kérdőívekkel való pszichometriai összevetéseknek kell megfelelni.

FÜGGELÉK

Az alábbiakban egy SPSS programban megírt SYNTAX állományt mutatunk be. A SYNTAX állományt elláttuk megjegyzésekkel annak érdekében, hogy részint érthető legyen annak minden fontosabb blokkja, sora; részint pedig adott esetben átformázható legyen, saját igényeink szerint legyen alakítható. Az alábbi eljárást egymás után mondjuk 200-szor futtatva, az eredményeket például EXCEL programcsomagba feljegyezve tudjuk a dolgozatban közölt eljárást megvalósítani.

Amennyiben menüvezérelt módon szeretnénk megoldani, az eljárást a SELECT CASES menüpontban találhatjuk meg.

```
USE ALL.     ***** Ez azt jelenti, hogy a teljes mintából válogatunk ****

do if $casenum = 1.
  compute #s_$_1=200.
  compute #s_$_2=4600.
  end if.
  do if #s_$_2 > 0.
  compute filter_$ = uniform(1)* #s_$_2 < #s_$_1.
  compute #s_$_1 = #s_$_1 - filter_$.
  compute #s_$_2 = #s_$_2 - 1.
  else.
  compute filter_$ = 0.
  end if.

*** E fenti szakasz egy makró, melyben egy új változót illeszt a program a meglévő
*** adatállomány VÉGÉRE filter_$ néven, mely filterről azt tudjuk, hogy az első 4600 főt veszi
*** alapul, és innen választ ki 200 főt
*** Vegyük észre: ha az s_$_1=200 sorban a 200-as számot 300-ra írjuk át, akkor 300 főt fog
*** választani – az s_$_2=4600 sorban lévő 4600-as érték átírásakor tudjuk megadni, hogy az
*** első hány darab esetet használja a program.

VARIABLE LABEL filter_$ '200 from the first 4600 cases (SAMPLE)'.
FORMAT filter_$ (f1.0).
FILTER BY filter_$.
EXECUTE .

*** E fenti részben megadjuk, hogy mi legyen az új változó neve, formátuma – illetve, hogy ez
*** a változó filterként működjön, tehát amely esetben ez az érték 1-es, úgy figyelembe
*** veszi a program az elemzéskor – míg ha ez az érték 0, úgy kimarad az adott egyén az
*** elemzésből.

RELIABILITY
  /VARIABLES=mean_m_1 mean_m_2 mean_m_3 mean_m_4
  /SCALE('ALL VARIABLES')  ALL/MODEL=ALPHA.

*** E fenti blokkban a kiválasztott egyénekre számítja ki a Cronbach-alfa mutató értékét.

DESCRIPTIVES
  VARIABLES=mean_m
  /STATISTICS=MEAN STDDEV MIN MAX SEMEAN KURTOSIS SKEWNESS .

*** Kiszámítjuk a mean_m változó átlagát, szórását, minimumát, maximumát, az átlag
*** standard hibáját, valamit a változó ferdeségét és csúcsosságát.

FILTER OFF.
USE ALL.
EXECUTE .
*** Kikapcsoljuk a filtert, újra az egész adatállományon dolgozhatunk – és futtatjuk az
*** eljárást. Az egész blokkot többször egymás után futtatva elérhetjük a dolgozatban közölt
*** eredményeket. ***
```

IRODALOMJEGYZÉK

Aldrich, J. (1995). Correlations genuine and spurious in Pearson and Yule. *Statistical Science*, *10*, 364-376.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*, 297-334.

Efron, B., & Gong, G. (1983). A leisurely look at the bootstrap, the jackknife, and cross-validation. *The American Statistician*, *37*, 36-48.

Halmai Zs., Dömötör E., Balogh G., Sárosi A., Faludi G., & Székely A. (2008). Egy új hangulati kérdőív validálása egészséges mintán. *Neuropsyhopharmacologia Hungarica*, *10*, 151-157.

Rowe, P. (2002). *Contributions to metric number theory: Technical report*. Letöltve 2013. 09. 08-án: http://digirep.rhul.ac.uk/items/369924aa-e6aa-8ca8-691b-c74c8375fa7a/1/

OECD PISA 2003 felmérés. Letöltve 2013. 11. 06-án: www.oecd.org/pisa/pisaproducts/pisa2003/