
Hunyadi László

professor emeritus

E-mail: hunyadi44@gmail.com

Vita László

professor emeritus

E-mail: laszlo.vita@uni-corvinus.hu

**Száműzött
szignifikanciatesztek**

DOI: 10.20311/stat2016.04.hu435

Érdekesítő cikk jelent meg a *Statisztikai Szemle* 2016. januári számában. *Bárdits Anna, Németh Renáta és Terplán Győző* statisztika területén dolgozó szociológusok „Egy régi probléma újra előtérben: a nullhipotézis szignifikanciateszt téves gyakorlata” című műhelytanulmányukban¹ ismertetik és kommentálják egy vezető nemzetközi szociálpszichológiai folyóirat szerkesztőségének állásfoglalását, miszerint az, elsődlegesen a sok módszertani visszaélés elkerülése érdekében, a továbbiakban nem kíván olyan írásokat közölni, amelyek a szerzők állításait vagy azok egy részét a szokásos szignifikanciatesztekkel igazolják.

A hír valóban meglepő, sőt sokkoló hatású lehet a statisztikusok és a statisztikát alkalmazó, elsősorban társadalomtudományi kutatók számára; azonban nagyon körültekintően kell bánnunk a reflexiókkal. Egyrészt kerülni kell még a látszatát is, hogy valaki kívülről be akarna avatkozni egy szakma vagy egy folyóirat belügyeibe. Másrészt a cikk, amely egyebek között ezt a hozzászólást is indukálta, nagyrészt leíró, ismertető jellegű: a szerzők bemutatják a folyóirat szerkesztőségének állásfoglalását, idéznek ezzel kapcsolatos álláspontokat, de persze valamilyen szinten leírják véleményüket is.

Hozzászólásunkkal kapcsolatban mindenekelőtt hangsúlyozni szeretnénk, hogy a kérdést elsősorban módszertani oldalról közelítjük meg, s mivel nem ismerjük a szociálpszichológia statisztikai alkalmazásának nemzetközi gyakorlatát, véleményünket csak a saját (elsősorban magyar társadalomtudományi és közgazdaságtudo-

¹ BÁRDITS A. – NÉMETH R. – TERPLÁN GY. [2016]: Egy régi probléma újra előtérben: a nullhipotézis szignifikanciateszt téves gyakorlata. *Statisztikai Szemle*. 94. évf. 1. sz. 52–75. old. <http://dx.doi.org/10.20311/stat2016.01.hu0052>

mányi-gazdálkodási statisztikai alkalmazásokra épülő) tapasztalataink alapján fogalmazzuk meg. Ezáltal természetesen azzal a feltevéssel élünk, hogy a két terület problémái, legalábbis jórészt, hasonlóak, és a kibontakozási lehetőségekben is bőven van közös elem. Végül, utalva a vitaindító cikk leíró jellegére, a továbbiakban alapvetően nem a cikk szerzőivel, hanem az általuk közvetített vélemények némelyikével kívánunk vitatkozni, s mindez a szerzőket csak annyiban érinti, amennyire azonosulni tudnak az általuk tolmácsoltakkal.

Ebben a vitacikkben nagyjából az eredeti írás szerkezetét követjük: a bevezetés után először azzal foglalkozunk, hogy milyen okok válhatták ki a szokatlan döntést, és ezek fennállnak-e az általunk ismert területeken is. Majd megpróbáljuk összegyűjteni a lehetséges megoldás elemeit. A rövid hozzászólás végén szükségszerűen újból felvetődik a kérdés, vajon a mi szemszögünkből helyeselhető-e ez a drasztikus elhatározás.

1. Szűkebb szakmai bajok a szignifikanciatesztekkel

A cikk részletesen sorra veszi az említett döntés lehetséges indokait, amelyekkel nagyrészt egyet lehet érteni. Így itt nem térünk ki egyebek között a szignifikanciaszint értelmezésére vagy annak fontosságára, hogy az egyes tesztek használatának feltételeit szem előtt kell tartani az alkalmazás során, de arra a közismert tényre sem, hogy nagyon nagy minták esetén ezek a tesztek szisztematikusan húznak a nullhipotézis elutasítása felé, azaz a mintában meglevő egészen kis különbségeknek indokolatlanul nagy jelentőséget tulajdonítanak. Megjegyezzük, hogy amennyiben egy teljes sokaság megfigyelésére kerül sor, a tesztek használata per definitionem értelmetlen, ám bizonyos esetekben a teljes sokaság egy nagyobb ún. szuperpopuláció részeként is felfogható, s akár mintaként is kezelhető. A nagyon nagy mintákkal kapcsolatos megjegyzés természetesen ez esetben is érvényes marad.

Nem térünk ki továbbá a cikk által felsorolt többi jellegzetes hibára, mivel ezekkel komolyan és nagyobb számban nem találkoztunk.

Két elemre azonban egy kicsit részletesebben is reagálunk. Ezek egyike a szakmai és a statisztikai szignifikancia szembeállítás. Véleményünk szerint bár e kétféle szignifikancia eltérhet egymástól, a szembeállítással óvatosan kell bánni. A „szignifikáns” elnevezés maga is okozhat zavart, de nem hisszük, hogy ez lényeges lenne. Persze máshogy merül fel ez, ha valaki anyanyelvén nem tudja egymástól megkülönböztetni a *szignifikáns* és a *lényeges, jelentős* kifejezéseket, illetve máshogy akkor, ha arra gondolunk, hogy a jelenlegi magyar nyelvben mennyire divatosá váltak olyan szavak, mint az *installáció*, az *applikáció* vagy a *desztináció*. Ilyen erővel a szignifikáns kifejezés is elterjed(het) akár köznapi, akár tudományos környezetben

(és így ugyan okozhat félreértést, de ezzel nem lehet mit tenni), de nem statisztikai értelemben.

A probléma inkább az, hogy a statisztikai modell szignifikanciája nem szükségszerűen jelenti azt, hogy az eredmény tárgyszerűen, szakmailag is jelentős. Ez persze könnyen igaz lehet, hiszen előfordulhat, hogy egy statisztikai próba *szignifikáns* a maga értelmében, közben valami olyan részletkérdésben mond ki állítást, ami az egész problémakört illetően nem elsőrendű fontosságú, azaz *nem szignifikáns*. Hasonló eset adódik például akkor, amikor nagy mintaelemszámok mellett az eredmény erősen szignifikáns, de e mögött csak a H_0 -beli helyzettől való jelentéktelen, szakmailag teljesen érdektelen eltérés húzódik meg. Annak érdekében, hogy elkerüljük a statisztikai és a szakmai szignifikancia téves azonosítását, (hacsak lehet) érdemes ún. hatásvizsgálatokat is végezni. Ez azt jelenti, hogy a „Lehet-e ekkora?” kérdés mellett mindig érdemes a „Mekkora lehet?” kérdésnek is kellő figyelmet szentelni. A hatásvizsgálatokra még akkor is szükség van, ha valamely hipotézisvizsgálattal kapott két vagy több eredmény mindegyike erősen szignifikáns. Ezekben és a hasonló esetekben semmiféle lényegi ellentmondást nem érzünk a kétféle szignifikanciafogalom között. Értelmes kutató-alkalmazó ezt természetesnek tekinti, legfeljebb alkalmas szóhasználattal és/vagy az eredmények gondos megfogalmazásával segít eloszlatni a félreértéseket.

Roszbab a helyzet akkor, ha a kutató-alkalmazó felül akarja bírálni a statisztikai eredményeket, mondván, „Ez elképzelhetetlen, a modell nem tudja figyelembe venni azt, hogy...”, „Szakmai tapasztalatom azt diktálja, hogy...” Szerintünk ez a kutatói magatartás megengedhetetlen: ha valaki hisz a modellben, higgyen az eredményekben is; ha pedig nem hisz benne, ne csinálja, vagy ne azt csinálja. Bár azt, hogy ez talán nem teljen nyilvánvaló, igazolni látszik *Henri Theil* ([1971]² vi. old.) egy megjegyzése is: „Bizonyos érettséget követel meg az, hogy valaki megértse: a modellek arra valók, hogy használjuk őket, nem pedig arra, hogy higgyünk bennük.”

Nyilván vannak bonyolultabb modellek, amelyek megengedik, hogy a kutató bizonyos módon (korlátozásokkal, feltételezésekkel stb.) előre, az eredmények ismerete nélkül közölje saját (külső) tapasztalatait. Az így készült modellekre alkalmazott szignifikanciatesztek már nehezen bírálhatók felül szubjektív „fontosságérzet” alapján. Ebben az értelemben tehát a statisztikai és a szakmai szignifikancia között nem látunk érdemi különbséget.

A másik elem, amire reagálni szeretnénk, a próbák ereje. Az erő valójában egy kicsit nehezen megfogható, bár kétségkívül fontos tulajdonsága egy tesztnek. (Megjegyezzük, hogy egy másik hasonlóan fontos tulajdonság, a torzítatlanság szinte teljesen ismeretlen az alkalmazók körében.) Egy próba ereje egy több tényezőtől függő függvény. A részletek elhagyásával most csak annyit említünk meg, hogy amennyiben a próba ellenhipotézise összetett hipotézis, az erő egyebek között az

² THEIL, H. [1971]: *Principles of Econometrics*. Wiley. New York.

ellenhipotézis függvénye. Ha arra gondolunk, hogy egy egyszerű regressziós modell valamely együtthatóját teszteljük a szokásos

$$H_0 : \beta_j = 0 \text{ és } H_1 : \beta_j \neq 0$$

hipotézispárral, az ismert t -próba ereje attól (is) függ, hogy az ellenhipotézisek halmazából melyik egyszerű hipotézist választjuk. Ha az összes ellenhipotézishez tartozó erőt meg akarjuk adni, akkor kapjuk az erőfüggvényt. Így van ez nemcsak ebben a példában, hanem a gyakorlatban alkalmazott próbák jó része esetében is. Ebből pedig az következik, hogy nem olyan egyszerű sem meghatározni, sem megadni egy próba erejét, egy átlagos, a statisztikában nem járatos felhasználótól pedig nem igazán lehet azt követelni, hogy minden teszthez erőfüggvényt és erőfüggvény-értékelést mellékeljen.

Ezen a statisztikai próbák előre megtervezett végrehajtásával lehet segíteni, ami azt jelenti, hogy mindkét fajta hiba elkövetési valószínűségét valamilyen elfogadható szinten próbáljuk tartani. Az elsőfajú hibáét például 5, a másodfajúét pedig mondjuk 10-20 százalékos szinten. Ehhez előre meg kell adni egy szakmailag már lényegesenek tekinthető minimális eltérést a H_0 -beli helyzettől, és meg kell határozni azt a legkisebb mintanagyságot, amelyiken már teljesül a másodfajú hiba elkövetési valószínűségére vonatkozó elvárásunk. Ennek elvégzéséhez ma már többféle számítógépes program is rendelkezésre áll. (Erről lásd *Vargha* ([2007]³ 163. old.).)

A cikkben felsorolt problémák és hiányosságok között nem szerepel ugyanakkor a következtetések alapjául szolgáló minta vizsgálata. Az általunk valamelyest ismert társadalomtudományi alkalmazások többnyire kimondják, hogy mintájuk reprezentatív (ami azonban némiképp obskurus, nem tisztán definiált fogalom), de gyakran még csak ezt sem feltételezik, hanem bevezetik a „kényelmi” vagy „hozzáférhetőségi” minták fogalmát, ami egyfelől szakmai okokból érthető, másfelől azonban ellene van minden statisztikai megfontolásnak. Az már valóban csak „hab a tortán”, ha valaki nem veszi figyelembe azt, hogy véges (kis) sokaságok esetén valós, visszatevés nélküli kiválasztások mellett a klasszikus statisztika eredményei korrekcióra szorulnak. Ismét csak nem elmélyedve a részletekben, meg kell említeni, hogy az egész klasszikus következtető statisztika módszertana arra a feltételezésre épül, hogy aktuális mintánk egy a sok lehetséges minta közül, és kísérletünk hasonló körülmények között tetszés szerinti számban megismételhető. Ebből adódik a klasszikus statisztika egyik nagy ellentmondása, nevezetesen az, hogy jóllehet csak egyetlen mintánk van, amikor következtetünk, feltételezzük a megismételhetőséget, és eredményeink (például a becslés torzítatlansága, a konfidencia-intervallumok, a tesztek p -értékei stb.) csak ezen ismétlések kereteiben értelmezhetők. Bár ezeket az alapokat

³ VARGHA A. [2007]: *Matematikai statisztika*. Pólya Kiadó. Budapest.

már a legelső induktív statisztikai kurzuson tanítjuk, azt tapasztaltuk, hogy a hallgatóság érdemben nem sajátítja el a következtető statisztikának ezt az alapgondolatát, és még inkább féltő, hogy más területeken, ahol a matematikai és általában a módszertani képzésre érhető okokból kevesebb figyelmet, erőt, időt fordítanak, mint a közgazdászoknál, a hallgatók és a későbbi kutatók még kevésbé értik és érzik a következtetési statisztika lényegét.

2. Egyéb okok

Miután a szűkebb szakmai hiányosságokat áttekintettük, azonnal felmerül a kérdés, hogy ezek miből adódhatnak, azaz, ha egy kicsit mélyebbre ásunk, találunk-e még olyan tényezőket, amelyeket érdemes megemlíteni. Úgy véljük, hogy a kérdésre igen a válasz, amit azonnal meg is indokolunk. Az első tény, ami felbukkan az, hogy a rendelkezésre álló vagy inkább kényelmesen elérhető statisztikai eszközök tárháza elég szűkös, nem mindig igazodik a társadalomtudomány problémáihoz, rendelkezésre álló mintáihoz. A dolog talán ott kezdődik, hogy a klasszikus statisztikai eszköztár kialakulása a múlt század első harmadára-felére tehető, amikor *Ronald Fisher*, *Jerzy Neyman*, *Karl Pearson* és munkatársaik elsősorban természettudományos és mérési, méréseméleti problémák kapcsán kidolgozták a klasszikus statisztika alapelveit és módszereit. Ezek az ismételt mintavételen, elsősorban folytonos változókon, jellemző módon nagy mintákon, a normális eloszlás központi szerepén, a likelihood függvényen és annak domináns alkalmazásán alapulnak a becslések és tesztek körében. Ez a tetszetős és konzisztens elmélet hamar igazolást és alkalmazást nyert egy sor természettudományos területen, és döntő hatással volt az 1930-as években fellendülő közgazdasági és ökonometriai kutatásokra is. Bár a lényegi hasonlóságok mellett viszonylag hamar kiütöztek a természettudományos gondolkodás és a társadalomtudományok problémái, lehetőségei közötti eltérések és ellentmondások, ez a klasszikus statisztika vált uralkodóvá a közgazdasági kutatásokban és a valamivel később induló társadalomtudományi, pszichológiai, orvosi stb. alkalmazásokban. A módszertan persze finomodott, sokban igazodott a speciális igényekhez, de alapvető változás, valamiféle paradigmaváltás nem következett be: ma is a közel százéves alapokra épülő statisztikát használjuk, olykor megújuló köntösben. Ez azonban nagyon sok esetben nem megfelelő a felmerülő problémák helyes kezelésére. Új, adekvát eszközök vagy nincsenek, vagy vannak ugyan, de a kutatók nem ismerik őket, vagy vannak, és ismerik is őket, de nincs meg hozzájuk a megfelelő szoftver. Mindez alapot adhat és gyakran ad is arra, hogy a kutatók a jól bevált, könnyen elérhető módszereket alkalmazzák, akár helytelenül is.

Az előbbieken kifejtettük, hogy a helytelen alkalmazások egyik lehetséges oka az, hogy nem áll minden esetben rendelkezésre a problémának megfelelő módszer. Emellett fontos oknak tartjuk azt is, hogy szűklátókörű tudománypolitikai indíttatásból egy sor olyan kutatásra kényszerítenek szakembereket, amelyek meghaladják képzettségüket és képesítésüket. Sok beosztásban – elsősorban egyetemeken – a szakmai előrelépés, de gyakran a tartós alkalmazás feltétele a tudományos fokozat megléte, aminek megszerzéséhez egyre több területen statisztikailag megalapozott és kiértékelt kutatások szükségeltetnek. Nem nehéz belátni, hogy az egyetemeken, főiskolákon egy sor olyan terület van, ahol az elmélet alapjait ismerő, kiváló pedagógiai és gyakorlati adottságokkal, felkészültséggel rendelkező szakember magas szintű oktatást tud ellátni még akkor is, ha nem jeleskedik a kutatásokban. Vagy gondoljunk arra, hogy a szinte mindenütt tömegesedő oktatás olyan oktatókat igényel, akik ezeket a tömeges igényeket megfelelően magas színvonalon tudják kielégíteni. Ehhez ismét nem sok köze van a kutatói kompetenciáknak és a kutatási eredményeknek. Mindezek alapján azt bátorkodjuk állítani, hogy a Magyarországon jelenleg érvényes szabályozás arra kényszerít sok, a maga területén értékes szakembert, hogy ott is próbáljon „lépni”, ahol nincs meg a kellő felkészültsége, ezért nagy az esélye annak, hogy munkáikban a módszertani alkalmazások, gyakran éppen a statisztika, azon belül pedig a tesztek kifogásolhatók, megalapozatlanok, hibásak lesznek. Alighanem más országokban is ehhez hasonló a helyzet.

Végül még egy problémát, a publikációs lehetőségek egészségtelen növekedését szeretnénk kiemelni. Az igény, mint láttuk, megvan, hiszen a felsőoktatás tömegesen nő, a felsőoktatók száma majdnem hasonlóan, a formális követelmények (elvárt publikációk száma) nem különben. Ez a nyomás oda vezetett, hogy kialakult egy üzlet, nevezzük publikációs üzletnek, amelyben a költségek és a nyereség megfizetése mellett bárki megkaphatja a lehetőséget a publikálásra.⁴ Hangzatos elnevezésű folyóiratok, sőt nem kevésbé jól csengő nevű könyvkiadók jöttek létre azzal a nem is titkolt céllal, hogy üzleti alapon publikációs lehetőséget nyújtsanak mindazoknak, akik arra áldozni tudnak. Ehhez nem kell más, mint kifizetni a benyújtási díjat, vagy vállalni azt, hogy a megjelenő könyvből a szerző garantáltan elad (megvesz) 100-150 példányt, és máris látványos helye van a kiadványnak. Na persze, az itt leírtak egy kicsit (de valóban csak egy kicsit) eltúlzottak, ugyanis vannak színvonalas, komolyan lektorált és referált folyóiratok, a kiemelkedő szakmai eredmények pedig nyilván ezekben jelennek meg; ám kétségtelen a publikációs üzlet létezése és az is, hogy annak hatása igen negatív a (statisztikai) módszerek korrekt alkalmazása szempontjából.

⁴ *Umberto Eco* híres regényében a *Foucault-ingájában* ironikusan már felvázolja ennek az üzletnek az alapelveit, de ezt ő még úgy képzelte el, hogy párhuzamosan létezik egy komoly tudományos, valamint egy üzleti alapon álló kiadó, mely utóbbi a publikálni kívánó szerzők pénzén tartja el az igazi tudományt képviselő másikat. Úgy tűnik, e téren a valóság még a dús fantáziájú szerző elgondolásait is túlszárnyalta.

3. Mi lehet a megoldás?

Az előzőkben felvázoltuk azt, hogy milyen okok vezethettek ahhoz a súlyos helyzethez, hogy egy vezető folyóirat a sok visszaélés miatt teljesen elfordul a következtetési statisztika alapvető eszközeitől. A következőkben azokat az elemeket próbáljuk összeszedni, amelyek, ha ezt a döntést nem is, de a statisztika alkalmazásaiban terjedő hibás gyakorlatot képesek lehetnek visszafordítani. Természetesen ennek a vitairatnak a szerzői nincsenek semmiféle döntési pozícióban, ezért legfeljebb ajánlásokat, kéréseket fogalmazhatnak meg, illetőleg tehetnek hozzá mindahhoz, amit a bevezetésben szereplő cikk szerzői már megtettek.

Az említett írás sok lehetséges megoldást vet fel, ezek nagy részével ismételtlen csak egyet lehet érteni; ezért közülük csak azokra fogunk kitérni, amelyeknek különös fontosságot tulajdonítunk. Ezekben túl azonban van néhány olyan probléma is, amelyekre a szerzők nem tértek ki cikkükben.

Ha az elején kezdjük, azt kell mondani, hogy a tudománypolitika sokat tudna javítani a helyzeten azzal, amennyiben nem követelné meg válogatás nélkül az akadémiai szféra szinte minden szereplőjétől a tudományos fokozatot, és ezzel együtt a nagy mennyiségű, sokszor vitatható értékű publikációt. Ez együtt járhatna a gombamód szaporodó doktori iskolák tevékenységének ésszerű korlátok között tartásával, a tudósjelöltek számának drasztikus csökkentésével, a formális követelmények racionális korlátozásával, ami jócskán segítene egy sor tudományág (köztük a statisztika) módszertani megtisztulásában. Az ilyen és hasonló kérdésekben alighanem több jogosítványt lehetne és talán kellene adni az akadémiai tudományos bizottságoknak.

Áttérve a folyóiratok és a könyvkiadók tevékenységére, úgy gondoljuk, bőven van mit tenni a kiadók szakmai színvonalának emelése terén. Ezen most konkrétan és elsősorban a szakmai bírálatokat értjük. A nevezett és sok vitát kiváltó döntést alighanem az előzte meg, hogy a folyóirat lektorai nem tudták jól szétválasztani a közölhető cikkeket a közölhető, de alaposan átdolgozandóktól, valamint a közlésre alkalmatlanoktól. Ez valóban igen nehéz, gyakran kellemetlen, és sok hibalehetőséget magában rejtő munka, de ha színvonalas publikációkat szeretnénk, nem lehet eltekinteni a lektorok fontos szerepétől sem, és fel kell vetni a tudományos lektori tevékenység elismerésének kérdését. A lektorálást a mai magyar (de nem csupán magyar) gyakorlat szerint az adott szakterület felkért képviselője kevés kivétellel minden erkölcsi és anyagi elismerés nélkül vagy megalázó anyagi feltételek mellett végzi. Mindez igaz a tudományos élet más – nem nyilvános publikációra kerülő – anyagainak (például a doktori értekezéseknek a) bírálóira és bírálataira is.⁵ Egyes folyóiratok azzal próbálják értékelni lek-

⁵ Egy nemzetközileg is elismert tudós például volt mikor mindössze annyi honoráriumot kapott egy nehéz, közel 200 oldalas akadémiai doktori értekezés bírálatáért, hogy abból éppen csak ki tudta fizetni az utcai parkolási díjat az értekezés vitájának idejére.

toraik tevékenységét, hogy évente egyszer nyilvánosan közzéteszik azok nevét, megköszönve tevékenységüket, de ez csak a legrangosabb folyóiratok esetén lehet mozgósító erejű. Véleményünk szerint egy-egy cikk vagy szakkönyv felelősségteljes lektorálása igen komoly szakmai teljesítmény, ezért azt a jelenleginél sokkal jobban kellene honorálni. Már több helyen felvetettük, hogy módot kellene találni a színvonalas lektorálási tevékenység önálló tudományos eredményként való elismerésére a különböző szintű tudományos minősítő eljárásokban. Ha emellett esetenként anyagi díjazásban is tudnák részesíteni a lektorokat, elképzelhető, hogy lényegesen emelkedne a lektorálás színvonala, ami a javuló szelekción keresztül segíthetne elkerülni a most tárgyalthoz hasonló kellemetlen tiltásokat.

Visszatérve a szűkebb szakmai kérdésekre, mindenekelőtt hangsúlyoznunk kell, hogy az egyik legfontosabb dolognak azt tartjuk, hogy mindenki, aki következtető statisztikai módszereket használ, legyen tisztában az elméleti alapokkal, vagyis azzal, amit korábban kifejtettünk. Ha az alkalmazók tudják, mi a valószínűségi minta, mi a mintavétel lényege, hogyan kell értékelni egy becslés vagy egy teszt eredményeit, nyilvánvalóan sokkal ritkábban követik el a korábban részletezett hibákat. Természetesen nem gondoljuk, hogy tisztában kell lenniük a becslések és a tesztek matematikai-valószínűség-számítási háttérével; de azoknak, akik tesztek terveznek, végeznek vagy értékelnek, nem szabad nekikezdeniük a munkának anélkül, hogy tudnák, milyen logika, milyen feltételezések állnak mögötte, és pontosan mit jelentenek az eredmények. Ha mégis valaki ilyen feladatra vállalkozik, mindenképpen keresse az együttműködést hozzáértő statisztikussal!⁶ Mindebben az oktatás fontossága megkérdőjelezhetetlen. Sajnos az a helyzet, hogy még a közgazdászképzésben is, ahol az egzakt tárgyak (matematika, statisztika, informatika) oktatásának nemcsak nagy hagyománya van, de az jelenleg is színvonalas és hangsúlyos, távol állunk attól, hogy minden végzett hallgató biztos tudással rendelkezzen az induktív statisztika terén. Így elképzelhető, hogy olyan szakterületeken, ahol mindezek a tárgyak meg se jelennek, vagy egyesek (például a statisztika) megjelennek ugyan, de a hallgatói értékítélet szerint marginálisak, nagy a hallgatók tájékozatlansága ezekről a kérdésekről. A megoldás konkrét lépéseit pillanatnyilag nem tudnánk felvázolni, de nekünk, akik közel állunk a statisztika oktatásához, mindent meg kell tennünk azért, hogy ezek a statisztikai ismeretek minél szélesebb körben elterjedjenek. Ehhez persze egyebek között arra is szükség lenne, hogy más szaktárgyak (és itt nem csupán a közgazdasági felsőoktatás tárgyaira gondolunk) érdemben és minél jobban építsenek arra az ismeretanyagra, amit a hallgatók a módszertani tárgyak megismerése során elsajátítanak.

A tesztek helyes alkalmazását nagyban elősegítené, ha a felhasználók nagyobb kínálatból választhatnának. Amennyiben a leggyakrabban használt tesztek alkalmazási feltételei nem állnak fenn, a felhasználók kereshetnének és találhatnának egy, a

⁶ Olyan ez, mint amikor egy gyógyszer szedését szigorú orvosi felügyelethez kötik.

saját problémájuk megoldásához jobban illeszkedőt, mert azért ilyenek nagy számban vannak. Ezzel kapcsolatban természetesen első feladat a megismerés és a megismertetés: már az egyetemi oktatásban törekedni kell arra, hogy a hallgatók minél több tesztnek legalább a gondolatát, alkalmazási feltételeit megismerjék. Be kell vallani, hogy ezen a téren a (magyar) oktatásnak bőven van teendője, hiszen – amennyire tudjuk – a lehetséges és alkalmas tesztek közül csak a leggyakrabban alkalmazott néhány próbát ismertetik a társadalomtudomány statisztikáját tárgyaló tankönyvek. Kiváltképp nagy az adósságunk a robusztus, sok helyzetben használható nemparaméteres próbákkal. Persze a megfelelő próbák sem alkalmasak minden feladat megoldására, nyilván vannak, lehetnek olyan esetek, amelyeket a mai eszközökkel nem lehet korrekt módon megoldani. Az ideális eset az lenne, ha folynának olyan módszertani jellegű kutatások, amelyek éppen arra irányulnának, hogy a társadalomtudományi alkalmazásokban felmerülő, speciális statisztikai problémákra megoldást adjanak. Ez persze nem egyszerű és gyorsan teljesíthető feladat, és főleg – úgy tűnik – nincs meg még hozzá se az szándék, se az intézményi keret.⁷

Itt kell említést tenni a szoftverekről is, ahogy ezt a kérdést a vita alapját képező tanulmány is tárgyalja. Úgy véljük, el kell fogadni azt a többnyire ki nem mondott, de általánosan vallott nézetet, hogy az alkalmazók számára csak az a statisztikai módszer létezik, amely valamely közismert statisztikai szoftverben (az általunk vizsgált területeken leginkább az SPSS-ben) elérhető. Persze az SPSS-be is alighanem jóval több olyan eljárás van beépítve, mint amit az alkalmazók döntő többsége valóban használ, de nyilván az is igaz, hogy messze nem minden szükséges eljárás, próba érhető el így. Ezért tartjuk nagyon fontosnak azt, hogy készüljenek (lehetőleg kényelmesen kezelhető formában) olyan szoftverek, amelyek fejlesztésekor a tervezők már messzemenően figyelembe veszik az alkalmazási terület speciális vonásait. Mindenképpen fel kell hívni a figyelmet például a magyar fejlesztésű RopStat-csomagra, amelynek ötletadója és kidolgozója éppen a pszichológia és egyéb, hasonló eszközöket igénylő tudományágak statisztikájának kiváló ismerője. A szoftverekkel kapcsolatban meg kell még azt is említeni, hogy a szabad hozzáférésű és igen rugalmas R-nyelv kiválóan alkalmas ezeken a területeken alkalmazható próbák végrehajtására, sőt azok tulajdonságainak ellenőrzésére is. Természetesen a felhasználóktól nem várható el, hogy R-kódokat fejlesszenek, de az egyáltalán nem elképzelhetetlen, hogy a módszereket jól ismerő és az R-t valamilyen szinten alkalmazni tudó statisztikusok együttműködjenek bizonyos speciális helyzetekre a tesztek kidolgozásában és értékelésében. Amennyiben az ezek során szerzett tapasztalatokat a szakmában még ki is cserélnék, sokat léphetnének előre a korrekt társadalomtudományi alkalmazások terén.

⁷ Már több fórumon szóvá tettük, hogy a statisztikának nincs saját kutatószervezete. Vannak közgazdaságtudományi matematikai és pszichológiai kutatóintézetek, és van még sok más tudományágnak is kutatóbázisa, de tudásunk szerint az még szóba se került, hogy statisztikai kutatóhely is létezhessen.

Mégis, mi van akkor, ha ezeken a tudományterületeken olyan probléma adódik, amelyet nem lehet korrekten megoldani a rendelkezésre álló eszközökkel például és jellemző módon azért, mert a használt minta (nagysága, nem véletlen jellege stb.) nem alkalmas arra, hogy belőle általánosítható következtetéseket lehessen levonni? Szerintünk ebben az esetben nem lehet, nem szabad erőltetni a következtetést, hiszen az csak formailag adhat esetleg értékelhető eredményt. Ilyen esetekben talán meg kell elégedni a leíró statisztikákkal, illetve azok korrekt használatával és szakmai ismeretekkel való kiegészítésével kell törekedni minél általánosabb és szélesebb körben használható eredményekre.

Persze azért van még lehetőség keresni a továbblépés útját, és erre a kiinduló műhelytanulmány is helyesen rámutat. Ennek egy módja, amelyet az alapcikk is említ, és amellyel tudomásunk szerint eddig gyakorlatilag csak az orvosi alkalmazásokban lehetett találkozni, a metaanalízis. A módszer alkalmazási lehetősége persze még nagyon korlátozott, de sok területen utat nyitnak felé a nemzetközileg egyre inkább standardizálódó indikátorok és indikátorrendszerek, az egymással egyre inkább harmonizáló európai statisztikai felvételek.

4. Helyes lépés volt-e a tiltás?

A tiltás utóéletét az alapcikk is megpróbálta elemezni. A kérdésre a végleges választ nyilván az idő adja meg. Addig csak jósolgatni lehet. Úgy gondoljuk azonban, hogy egy ilyen drasztikus lépés nemcsak szokatlan, de nem is célravezető, amitől – ha tanácsot kellene adni – óva intenénk mindenkit. Szakmai oldalról persze érthető, hogy mindent – szó szerint mindent – megtesznek a hibás és félrevezető alkalmazások elterjedése ellen. Ugyanakkor látni kell, hogy ezzel az esetleges és bizonyára létező helyes alkalmazásokat is lehetetlenné teszik. Ez pedig mindenképpen a haladás ellen dolgozik. Úgy véljük, hogy a kérdést inkább széleskörű felvilágosító munkával, az alternatív eszközök tájának fejlesztésével és hozzáférhetővé tételével, valamint a jelenleginél odafigyelőbb szelektív lektorálással lehetne megoldani, mert ilyen és hasonló tiltásokkal aligha oldódik meg a probléma.