

AZ ADATOK JELENTÉSÉNEK TÚLBECSÜLÉSE – EGY GYAKORI TÉVEDÉS A HAZAI HUMÁN VISELKEDÉSTUDOMÁNYOKBAN

Rózsa Lajos

MTA-MTM Állatökológiai Kutatócsoport, Budapest, Ludovika t. 2., 1083

A fejlett orvosi gyakorlatban kiemelt szerepe van a megelőzésnek, melyhez feltétlen szükséges a népesség általános egészségi állapotának és életmódjának rendszeres nyomon követése; monitorozása. A klasszikus orvosi gyógyító tevékenység – mely az egyénre összpontosítja figyelmét – mellett tehát a szakirodalomban egyre növekvő szerepet kapnak a népesség (ill. az azt reprezentáló minta) egészét vizsgáló tanulmányok. Szeretjük vagy sem, a népesség valamely általános tulajdonságainak leírásában vagy e tulajdonságok népességek közti összehasonlításában legfontosabb eszközünk a statisztika. A statisztika szó kapcsán még ma is sokakban a megjegyezhetetlen, misztikus képletek kényszerű bemagolásának riasztó emléke merül fel. E rémképet bátran elfelejthetjük, hiszen ma már felhasználóbarát szoftverek gazdag kínálata segít a számítások villámgyors megvalósításában.

De éppen a gazdag kínálat jelent újabb kihívást; a lehetséges statisztikai funkciók miriádjából kell kiválasztanunk a céljainknak megfelelő egyetlen egyet. Ha tévesen választunk, akkor adatainkat félreértelmezve téves következtetésekre juthatunk. A *Magyar Epidemiológia* legújabb számában örömmel fedeztem fel, hogy a lap tág teret ad a magatartás- és pszichoszociális epidemiológiai tanulmányoknak. És mivel e tanulmányok némelyike jól példázza a hazai humán magatartás-tudományi tanulmányok egyik jellemző hibáját, megragadom itt az alkalmat, hogy felhívjam a figyelmet egy közkeletű félreértésre.

Ha adatokat elemzünk, elsőként mindig gondoljuk végig az adatsorban rejlő számok jellegét. Mit jelentenek ezek a számok, és mit nem jelentenek?

A számok végső soron egy skála mentén ábrázolt értékek. E skála három különböző típusú lehet; azaz három eltérő jelentést hordozhat.

1. Az *arányskála* jól példázható a távolság vagy tömeg jellegű adatokkal. Egy 100 kg-os ember kétszer akkora tömegű, mint az 50 kg-os ember, és 1 kg-mal nehezebb, mint a 99 kg-os embertársa. Nemcsak azt állíthatjuk tehát, hogy az egyik ember nehezebb, mint a másik, de arányokat és intervallumokat is ábrázolhatunk e skála mentén.

2. Az *intervallumskála* ennél kevesebb információt hordoz. Ha például Celsius fokban kifejezett hőmérsékleti adatokkal dolgozunk, akkor nem feltételezzük azt, hogy 2 °C kétszer olyan meleg volna, mint 1 °C. Azért hiányzik ez az információ az adatainkból, mert skálánk nullpontjával egy mesterségesen kiválasztott pontot (a víz forráspontját) választottuk. Azt azonban állíthatjuk, hogy 2 °C annyival melegebb 1 °C-nál, mint amennyivel 12 °C melegebb 11 °C-nál. Arányokat tehát nem ábrázolhatunk e skála mentén, de még mindig hiteles állításokat fogalmazhatunk meg az intervallumokról.

3. A lehetséges skálák harmadik típusa a *rangskála*. Ez csak a sorrendiségről közöl információt, amint az jól szemléltethető pl. az emberek közösségen belül elfoglalt rangjával. (Itt feltételezzük persze, hogy az emberek egy közösségen belül egymástól függetlenül szerzik rangjaikat, ami nem reális, de ettől most eltekintünk.) Tekintsük pl. az egyetemi hierarchiában

szerezhető rangokat, és rendeljük ezekhez rangszámokat. Rangban a tanársegéd (1) felett áll az adjunktus (2), felette a docens (3), aki felett rangsorrendben már csak a professzor (4) következik. E rangszámok alapján egy kívülálló is biztosan tudhatja, hogy ki a magasabb és ki az alacsonyabb rangú. Ámde ebben az esetben nem beszélhetünk arról, hogy mondjuk az adjunktus pontosan annyiival állna a tanársegéd fölött, mint amennyivel a docens áll öfölötte, és nem mondhatjuk, hogy a professzor rangja kétszerese volna az adjunktus rangjának. Ezek az információk egyszerűen nem értelmezhetők egy rangskála mentén, mert itt kizárólag a rangok egymásutánra következésének sorrendjét ábrázoltuk. E skála tehát nem hordoz információt a rangok közti intervallumokról vagy arányokról.

A fenti példában a professzor rangja (4) látszólag kétszerese az adjunktus rangjának (2). Nyilvánvaló, hogy ez csak műtermék, hiszen ha a tanársegédi rangot 0 (vagy bármi más) értékkel indítva számszerűsíténénk a rangsorrendet, akkor más látszólagos arányokat kapnánk. A rangskála mentén ábrázolt adatok tehát csak a sorrendiségről hordoznak információt, de a rangok arányairól vagy a köztük lévő intervallumokról már nem.

Mondandóm lényege, hogy ezért itt több közönséges statisztikai művelet alkalmazása is megalapozatlan, az így kapott eredmények pedig félrevezetők. Ne számoljunk pl. átlagot és szórást, mert ezek az információk csak értelmezhető intervallumok esetén relevánsak. Hiszen az 1 és a 3 átlaga csak akkor 2, ha az 1 ugyanolyan messzire van 2-től, mint a 3 a 2-től. Ezért nincs értelme azt állítani, hogy a tanszéki oktatók „átlagos rangja = adjunktus-egész-23-század”. Van értelme viszont számsorainkat medián értékekkel jellemezni, mert pl. azt elmondhatjuk, hogy a tanszéki oktatókat rang szerint sorba állítva a középső ember rangja adjunktus (tehát medián rang = 2).

Egy rosszul kiválasztott statisztikai teszt révén azonban a szerzők – akaratlanul is – olyan információt magyaráznak bele adataikba, amely azokban nem volt benne. Keresztes et al. (2005) valamint a Keresztes és

Pikó (2005) éppen ezt tették a *Magyar Epidemiológia* legutóbbi számában. Tévedésük nem kivételes, hiszen számos más példát sorolhatnánk még pl. az *Iskolapszichológia* (Módszertani füzetek, ELTE Társadalom- és Nevelépszichológiai Tanszék) füzetéből is.

Mindkét fent említett tanulmány kérdőíves adatfelvételen nyugszik, ahol az adatközlők szubjektív kategorizálás alapján közöltek „score”, vagyis rangszám-jellegű adatokat magukról. Ilyen adatok az alkoholfogyasztási kategóriák; vagyishogy az adatközlő az elmúlt 3 hónapban „nem” (1), vagy „alkalmanként” (2), vagy „rendszeresen” (3) fogyasztott alkoholt; valamint hogy soha (1), ritkán (2), időnként (3), ill. gyakran (4) tapasztal magán pszichoszomatikus tünetek; vagy pl. az étellel való elégedettség score-jai.

Az egyébként igen kiváló tanulmányok szerzői tehát félreértelmezték saját adataikat, amikor a puszta sorrendiségnél messzebb mutató, pl. az intervallumok megbízhatóságán alapuló statisztikai teszteket (ún. parametrikus teszteket) alkalmaztak elemzéseikben. Az átlag és a szórás értékek helyett medián értékekkel jellemezhetők volna mintáikat. Két minta összehasonlításában Student-féle t-teszt helyett Mann-Whitney U-tesztet alkalmazhattak volna, amely nem az átlagokat, hanem a mediánokat hasonlítja össze. Kettőnél több minta összehasonlításában varianciaanalízis (ANOVA) helyett a Kruskal-Wallis teszt kínálkozott volna megalapozott alternatívaként. Persze ezeknek az ún. nem-parametrikus teszteknek is vannak alkalmazási feltételei, melyek nem feltétlenül teljesülnek az egyes konkrét esetekben; ilyenkor azonban más, modernebb eljárások – pl. bootstrap – kínálhatnak megoldást.

Az a benyomásom, hogy az eljárások téves kiválasztásának egyik oka az lehet, hogy a szerzők nagyon drága és nagyon komplex statisztikai szoftvereket (pl. SPSS) alkalmaznak, amelyek használata szerteágazó biostatisztikai előképzettséget feltételez. A legtöbb tanulmány esetében ugyanakkor egy ingyenes, felhasználóbarát, és viszonylag szerény

kínálható programcsomagok (pl. GrahPad Instat 3, ingyenes demo verzió) segítségével sokkal körültekintőbb statisztikai döntéseket hozhatnának, mert ezek naiv felhasználót feltételezve egyszerű tanácsokat adnak arra nézve, hogy mikor melyik funkciót válasszuk.

Összegezve tehát a hazai humán viselkedéstudományokban egy gyakran felmerülő, tipikus hibának tűnik a rangskála mentén értelmezhető adatok „túlhasznosítása”. Az átlagok, szórások, valamint a parametrikus tesztek (t-teszt, ANOVA) alkalmazásával ilyenkor akaratlanul is olyan információt magyarázunk bele adatainkba, amelyet azok valójában nem hordoznak. Kutatási eredményeink, valamint az

ezebből levonható következtetések hitelességét mindez főlegesen erodálhatja. Ne hagyjuk.

Hivatkozások

1. Keresztes N, Makó M, Klembucz E, Hanusz K, Pikó B. Magatartási kockázati tényezők összehasonlító epidemiológiai vizsgálata a Dél-alföldi régió ifjúsága körében. Magyar Epidemiológia 2005; 2: 195-207.
2. Keresztes N, Pikó B. Pszichoszociális egészségfaktorok összehasonlító epidemiológiai vizsgálata a Dél-alföldi régió ifjúsága körében. Magyar Epidemiológia 2005; 2: 217-228.

PARADIGMÁKON INNEN ÉS TÚL: AZ ÉRTELMES PÁRBESZÉD LEHETŐSÉGE ÉS LEHETETLENSÉGE

Pikó Bettina

SZTE ÁOK Magatartástudományi Csoport, Szeged 6722 Szentháromság u. 5.

Érdeklődéssel olvastam Rózsa Lajos „Az adatok jelentésének túlbecslése – egy gyakori tévedés a hazai humán viselkedéstudományokban” címmel írt hozzászólását. Annál is inkább, hiszen – mint ahogy ezt leírja – a hazai tanulmányok egyik jellemző hibájáról van szó, nem kivételes „tévedésről”. A Hozzászóló Kolléga kifogásolja, hogy nem megfelelő szinten kezeltük az adatokat, s ezek alapján téves következtetéseket vontunk le. Részletesen bemutatja a skálák jellegzetességeit, és javaslatot tesz másfajta statisztikai program használatára. Köszönjük a részletes bemutatást, de remélem, a kolléga elhiszi, hogy ezek nem ismeretlenek számunkra. A továbbiakban megkísérlem röviden megválaszolni felvetéseit.

A kuhni értelemben vett paradigmaelméletből indulnék ki, amely kimondja, hogy minden tudományterület a maga szabályrendszerében mozog, e szabályok felhasználásával alkotja

elméleteit és tervezi tudományos kutatását (1). A természet- és társadalomtudományos paradigmák sokszor éles ellentétben állnak egymással, érintkezési felületük alig van. A társadalomtudomány soha nem lesz olyan mértékben „objektív”, mint a természettudományos mérés, hiszen például a magatartás „mérése” eltér egy laboratóriumi méréstől. Mégis ez szolgáltatja az alapját a társadalom és az emberi magatartás komplex megértéséhez. A társadalomtudományi mérés „puhább volta” megköveteli a szigorú elméleti megalapozást. Bármennyire is szeretnénk, és bármennyire is egyszerűbb volna az ember „szubjektív valóságát” kiiktatni, ez nem lehetséges, az ember soha nem lesz molekulák halmaza, mert biopszichoszociális lény. S ez minden körülmények között érvényes. Az orvostudományban ez különösen nagy kihívást jelent, mert tárgya az ember, a maga teljességében. Itt találkozunk, vagy

jobban mondva gyakran „ütközik” két kultúra, a természet- és társadalomtudományi, ami nem megy konfliktus nélkül (2). Aki ezen a szakterületen dolgozik, az külföldi tapasztalatok alapján is tudja ezt. Még az orvostudományon belül is vannak olyan interdiszciplináris határterületek, mint a népegészségtan (public health) és a magatartástudomány (behavioral science), ahol a társadalomtudományi paradigmák fokozottan érvényesülnek, de ide sorolhatnám a pszichiátriát is, ahol az utóbbi évtizedek biológiai pszichiátriai eredményei mellett továbbra is megmaradt a szociálpszichiátriai látásmód (3).

Itt el kell osztanom egy félreértést. A magatartástudomány nem humán viselkedéstudomány, mint ahogy az emberi magatartást sem lehet az állati viselkedéssel összehasonlítani, legalábbis bizonyos értelemben, mint amilyen a betegségmegelőzés vagy az egészségfejlesztés. Köztudott, hogy a megelőzés a miatt került előtérbe, mert a modern korban a betegségek túlnyomórészt életmódunkból, azaz komplex (biopszichoszociális) egészség-magatartásunkból keletkeznek. Mi is a magatartástudomány és ennek alkalmazott területe, a magatartásorvoslás? Az orvoslás során, annak valamennyi fázisában felvetődnek pszichológiai, szociológiai, etikai vagy kulturális antropológiai kérdések. Ezt jól tükrözi a magatartásorvoslás definíciója. Az 1977-es Yale Konferencián ugyanis a következőképpen fogalmazták meg a magatartásorvoslás lényegét: Interdiszciplináris terület, amelynek célja, hogy a magatartástudomány és a biomedicina tudáskincsét és módszertanát integrált módon alkalmazza az egészséggel és betegséggel kapcsolatos problémák során a prevenció, a diagnózis, a gyógyítás és a rehabilitáció valamennyi területén (4). A magatartásorvoslás szinte valamennyi orvostudományi területtel és specializációval kapcsolatban áll.

A magatartás-epidemiológia (illetve szociális vagy pszichoszociális epidemiológia) az egészséget befolyásoló magatartási, pszichikai, kulturális, társadalmi kockázati faktorokat vizsgálja. Ezzel segíti a megelőzés elméletét és gyakorlatát. Itt térnék vissza a paradigmához, hiszen a magatartás-epidemiológia módszertanában igen gyakori (mondhatni leggyakoribb) mérési módszer a kérdőíves adatfelvétel (ami a viselkedéstudományban, még a humán viselkedéstudományban sem bevett, hiszen itt a megfigyelés tölti be ezt a szerepet). E területnek is megvan a maga paradigmarendszere, amit téves lenne a *hazai* kutatások hibájának felróni. Annál is inkább, mert sajnos a hazai magatartástudományok csak jelentős késéssel jelenhettek meg az orvostudományokon belül az előző rendszer tudománypolitikája miatt (5). A társadalomtudományban a statisztikai módszertan alkalmazása valóban nem ideális, hiszen az emberi magatartás „mérésén” alapul, ami valójában közvetlenül nem mérhető, hanem önbevallásra épül. Mégis fontos számunkra ez az információ, hiszen e nélkül nem érthetjük meg az emberi társadalmat. A magatartást mérő skálák konstrukciók, nem ideális arány- vagy intervallumskálák, mégis a társadalom- és magatartástudományi paradigmában ezeket folytonos változóként kezeljük, éppen az interpretálhatóság kedvéért. Természetesen mindenki nagyon jól tudja, aki ezen a szakterületen dolgozik, hogy a CDI vagy SWL score 1-es és 2-es értéke nem jelenti azt, hogy valaki kétszer olyan depressziós vagy kétszer olyan elégedett az életével, mint aki az 1-es score-t kapta. Aki azonban ezen a paradigmán belül dolgozik, abban ez fel sem merül, mert ismeri a paradigma szabályait. Ez nem a hazai, talán „elmaradott” kutatóközönségre vonatkozik –mint azt a Hozzászóló feltételezi -, hanem a nemzetközire még inkább, amit a Hozzászóló Kolléga nagy

valószínűséggel nem ismer. (Csak zárójelben jegyzem meg, hogy nem is kell ismernie, hiszen szakterülete ezt nem is kívánja meg). A hazai kutatások követik a nemzetközi trendeket, amit az is bizonyít, hogy a 90-es évektől kezdve ez a terület is sikeresen bekapcsolódott a nemzetközi tudományos vérkeringésbe, művelői rendszeresen publikálnak nemzetközi, rangos folyóiratokba. Ahol mindenki folytonos változóként kezeli a skálákat, összesített score értékeket, és nemcsak t-próbát vagy ANOVA-t végeznek ezek tesztelésére, hanem korreláció- és regresszioelemzést is. Ami egy természettudományos paradigmán belül mozgó kutató számára lehet, hogy nem egészen fogadható el, mert nem ismeri e szakterület szabályrendszerét. Ezzel kapcsolatban jegyezném meg azt is, hogy a kifogásolt statisztikai programcsomag éppen a társadalomtudományi kutatások számára lett kifejlesztve (SPSS: Statistical Package for Social Science). Ha nem ezt alkalmaznánk, külföldi folyóiratok meg sem jelentenék a tanulmányainkat. Azt pedig kérem, higgye el a Tisztelt Hozzászóló, hogy alkalmazását és a hozzátartozó statisztikai elméletet több éven keresztül tanultuk egyetemi szinten. A javasolt nemparaméteres eljárások sokkal „gyengébb” tesztek, ami a statisztikai szignifikanciát jelenti, a kétmintás t-próba például sokkal „erősebb” e tekintetben, azaz ha ezzel alkalmazva is szignifikáns az eredmény, akkor az valóban megerősítő jellegű. Talán ezért is terjedtek el ezek a kifogásolt módszerek a skálák elemzésénél, főként nagy elemszámú survey típusú vizsgálatoknál, míg például el tudom képzelni, hogy a viselkedés megfigyelésének eredményeit, alacsony elemszámú mintánál érdemes ezekkel tesztelni. Ebbe a témába azonban nem szívesen megyek bele, lévén, hogy e szakterületen nem vagyok kompetens. Ráadásul a rangsorok kialakításának a skálák értelmezésénél nem is nagyon van értelme, ugyanakkor vannak olyan

elemzések, ahol igenis gyakran alkalmazott módszer valamilyen nemparaméteres eljárás.

Mindent egybevetve, köszönöm a hozzászólást, hiszen ennek kapcsán lehetőségem nyílt a magatartás-epidemiológiai módszertan felvázolására. Természetesen nem vonom kétségbe hozzászólásának jóhiszeműségét, azonban nem tartom szerencsésnek, ha nem egészen megalapozottan, csupán egyetlen tanulmányra hivatkozva messzemenő következtetéseket vonunk le, megkérdőjelezve ezzel mások munkájának eredményét.

Hivatkozások

3. Kuhn TS. The structure of scientific revolution, 2nd ed. Chicago: University Press of Chicago, 1970.
4. Snow CP. The two cultures. Cambridge: Cambridge University Press, 1999.
5. Pikó B. A társadalomtudományok illetékessége a biomedicina számára. Valóság 1999; 42(12): 22-30.
6. Pikó B. Magatartástudomány és prevenció: A preventív magatartásorvoslás jelentősége. Magyar Tudomány 2003; 48(11): 1381-1390.
7. Piko BF., Kopp MS. (Behavioral medicine in Hungary: Past, present and future. Behavioral Medicine 2002; 28(2): 72-78.