

MATEMATIKAI STATISZTIKAI MÓDSZEREK ALKALMAZÁSÁNAK PROBLÉMÁI AZ EMPIRIKUS PSZICHOLÓGIAI VIZSGÁLATOKBÓL SZÁRMAZÓ ADATOK FELDOLGOZÁSÁBAN

Ma már nem kell hangsúlyoznunk a matematikai statisztikai módszerek jelentőségét a pszichológiai kutatások egy jelentős részében. A kutatási feladat megoldására sokszor egyrészt különböző statisztikai módszerek, modellek sorozatának alkalmazását, másrészt az ezeket követő, ezekre épülő fogalmi-tartalmi elemzéseket jelenti, kiindulva valamilyen adatfelvételi anyagából. Tehát egy olyan folyamatról van szó, amelyben a statisztika eszközei több ponton fontos, de nem egyedül meghatározó szerepet játszanak. Emeljük ki és nézzük meg e folyamat néhány lényeges pontját. Beszélhetnénk többről is, de a terjedelem korlátoz. Tudnunk kell azt, hogy

- milyen kérdésekre keresünk választ, és hogy
- mi az a kör amín belül a választ keressük, tehát aminek határai között megállapításaink érvényesek lesznek.

Az empirikus vizsgálatoknál az előbbi meghatározza, hogy mit kell megfigyelnünk, az utóbbi pedig azt mondja meg, hogy milyen populációra vonatkozóan végezzük a megfigyeléseket.

Folytassuk a gondolatsort a populációra vonatkozó megjegyzésekkel. Általában nem az egész populációra végezzük a megfigyeléseket, hanem annak csak egy meghatározott részén, az ún. mintán. A kutatásoknál nagyon világosnak kell lennie annak, hogy mi a populáció, hogy mi a megfigyelt minta és annak, hogy mi a kettő közötti viszony. Rengeteg hiba forrása lehet, ha ezeket nem tisztázzuk kellőképpen. A felületesség később félreértések, hibák, támadások alapja is lehet. Ha ugyanis nem a megfelelő körültekintéssel járunk el a minta és a populáció viszonyát illetően, akkor az elemzésekből és értelmezésből adódó megállapítások - bár objektíven megalapozottnak tűnnek, mégsem lesznek valójában érvényesek a populációra. Nyilvánvalóan minden kutató törekszik arra, hogy kutatásai, közleményei megalapozottak legyenek - e vonatkozásban is. Ennek ellenére, még a nagy tekintélyű, magas tudományos színvonalú folyóiratoknál sem mondhatjuk, hogy minden rendben van ezen a területen. Hivatkozhatunk itt Elisabeth R. Valentine munkájára (1), aki beszámol két amerikai és két angol szerző vizsgálatairól, amik 1964-67 közötti amerikai (USA) és 1969-70 közötti angol publikációkra vonatkoznak. A vizsgálatok tárgya egyrészt az volt, hogy a közölt cikkekben mennyire korrekt a populáció és a minta leírása ill. a minta mennyire tekinthető reprezentatívnak. Az eredmény kétségbeejtő! Az USA-ban a kutatásokhoz tartozó empirikus vizsgáldások 75-80%-át diákokon (egyetemi, főiskolai hallgatókon) hajtották végre. A

cikkek 3,6%-a egyáltalán nem közölte, hogy kik voltak a kísérleti személyek. Figyelemre méltó, hogy a kutatások 80%-át a népesség 3%-át kitevő specifikus részpopuláción belül hajtották végre. Nem szorul magyarázatra, hogy ez a 3% miért nem reprezentatív a társadalom egészére nézve. Az alapprobléma nem a "csak" 3%-ban rejlik, hanem annak számos sajátos tulajdonságában (fiatalság, olvasottság, intelligencia, társadalmi hovatartozás, nagyobb férfi arány stb.). Angliában a diákokkal végzett vizsgálatok aránya a nem klinikai közlemények kivételével majdnem azonos az USA-belivel: 76,4%. További torzítás, hogy a vizsgálatoknak több mint a fele önkéntes vállalkozókkal dolgozik, ráadásul 28%-ban az sincs föltüntetve, hogy honnan toborozták a kísérleti személyeket. A publikációknak csak kevesebb mint 15%-a számol be adekvátnak tekinthető mintavételi eljárásról. További 5% tárgyalja nyíltan a mintavételi hiányosságokat és azok lehetséges következményeit. A két angol szerző végkövetkeztetése az, hogy a vizsgált folyóirat-cikkeknek mindössze 1-2%-a alapozta eredményeit az általában vett felnőtt népesség helyesen választott mintájára. Legyünk optimisták, bízzunk abban, hogy Magyarországon ma sokkal jobb a helyzet, de talán mégsem felesleges felhívni a figyelmet arra, hogy igyekezzünk helyesen eljárni a mintavételnél és a feltételek közlésénél (annak érdekében, hogy az idézett, ideális 1-2%-ba tartozhassunk).

Cikkünk elején felvetett első kérdéshez kapcsolódik az, hogy "mit" figyelünk meg és a megfigyelések eredményeit hogyan regisztráljuk. Itt mindenekelőtt két dologra érdemes felhívni a figyelmet:

- az első a hiányzó válaszok kezelése és annak megkülönböztetése a nemleges válaszoktól,
- a második a vizsgált tulajdonság(ok) mérés- és skálaelméleti hovatartozása.

Ezek tisztázása és a tisztázás eredménye alapján a kódrendszer helyes vagy helytelen megválasztása eldönti, meghatározza, hogy később milyen elemzéseket végezhetünk az összegyűjtött anyagon.

Sokszor találkozunk olyan esetekkel, amikor az alkalmazott kódrendszer nem teszi lehetővé, de alaposan megnehezíti, vagyis csak nagy többlet munkával teszi lehetővé az adekvát statisztikai módszerek alkalmazását. Tipikus példa erre, hogy ordinális skálán mérhető jellemzőknél nem a tulajdonságnak megfelelő monotonosági kódrendszert alkalmazunk. Az már csak az értelmezést zavarja és ezért nehezíti, ha a skálák irányát nem egyeztetjük. Mindkét hibán úrrá lehetünk, de csak megfelelő többletmunka árán. Másrészt viszont a statisztikai feldolgozást végzőket is türelemre kell inteni, mert elképzelhető, hogy kérdéstechnikai okai vannak a statisztikai elemzést zavaró kódok alkalmazásának. Például a kísérleti személyeknek nem hozzuk a tudomására, hogy mi a pozitív, ill. mi a negatív válasz. Ilyenkor a számítástechnikát kell segítségül hívni, hogy a megfelelő kódrend előálljon. Lehetőleg ne kódolásra bizzuk ezt a munkát, mert az nagy hibaforrás lehet.

A fentebb említett másik szempont a mérés- és skálaelméleti hovatartozás volt. Ennek tartalmi követelményeiről is szólnunk kell. E kérdéskör részletesebb kifejtését egész cikkemben, fejezetekben, könyvekben (6) találjuk meg, itt csak egy vonatkozását érintjük. A statisztikustól sokszor csak a feldolgozást, valamilyen próba elvégzését, valamilyen eljárás alkalmazását kéri és nem ritkán csak az értelmezésnél jelentkező meglepő következtetések hívják fel a figyelmet arra, hogy amit kaptunk az a feldolgozott

szá:nok alapján csak formálisan helytálló, de tartalmat nem hordoz, mert az alkalmazott módszerek nem adekvátak. Tipikus példa e hibára a regresszió- és a faktoranalízis alkalmazása nem dichotóm nominális skála esetén.

Lehet, hogy kutatásunkban e pontig mindenben helyesen jártunk el (jó a minta, megfelelőek a skálák stb.), de további buktatók következhetnek az egy- és kétváltozós feldolgozásban és az ezzel összefüggő tesztelésben. A tesztelésnek minden esetben statisztikai feltételei is vannak, nemcsak skálaelméletiek. A már említett két angol szerző kiválasztott ötven cikket és azokat vizsgálta meg e szempontból. Azt találták, hogy a legtöbb esetben a szerzők nem közölnek elég információt ahhoz, hogy a használt próba feltételeinek teljesülését ellenőrizni lehessen és a szignifikanciapróbát közlő kutatások 1/3-ánál alapvető hibát követtek el a próba alkalmazásában. Legalább 26% súlyosan megsértette az előfeltételeket. Vigyázzunk tehát! Nemcsak a próbát kell elvégezni, hanem azt is ellenőrizni kell, hogy teljesülnek-e az alkalmazás feltételei.

Egy újabb fontos kérdés, amit legalábbis említenünk kell, a hipotézis elfogadásának és elutasításának viszonya, illetve a statisztikai és a szakmai hipotézis kapcsolata. Akár el kell utasítanom, akár nem cáfolhatom a próba eredménye alapján a nullhipotézist, a statisztikus jelleg miatt mindig követünk el hibát. Vigyázzunk kell arra, hogy ha nem tudtuk megcáfolni a nullhipotézist, az még távolról sem jelenti annak igaz, cáfolhatatlan voltát. Ha el kell utasítanom egy statisztikai nullhipotézist (ami általában valamilyen mutatók megegyezésére vonatkozik), még mindig nagy óvatosságot igényel az így kimutatott különbség értelmezése. Egy konkrét kutatásban két gyerekcsoportot vizsgálva az derült ki, hogy az egyikben a beszédhibás gyerekek száma szignifikánsan magasabb mint a másikban. Ez a statisztikai eredmény. Az már interpretáció, ha kimondom, hogy a cigány gyerekeknel gyakoribb a beszédhiba. Mielőtt azonban ebből következtetéseket vonnánk le, gondoljuk meg, hogy a magyar anyanyelvű gyerekek, akiket kevésbé beszédhibásnak találtunk, milyen "beszédhibásak" lennének angol, orosz vagy más nyelvtérületen. A cigány gyerekek jórésze ui. nem tekinthető magyar anyanyelvűnek.

Már átvezet a kétváltozós elemzés kérdéskörébe az összefüggések szorosságának vizsgálata, az, hogy mit tudunk mondani két jellemző kapcsolatáról. A korábban már idézett két angol szerző tapasztalata szerint a kutatók inkább csak az összefüggések statisztikailag szignifikáns voltát vizsgálják, az összefüggés hasznosságát azonban csak ritkán. Saját tapasztalatunk a jelenlegi magyar gyakorlatban ennek ellenkezőjét mutatja és ez sajnos már a többváltozós eredményeket is súlyosan befolyásolja. A szerzők ugyanis elsősorban csak az összefüggések szorosságával foglalkoznak. Erre a kérdésre később még visszatérünk.

Fontos tudnunk még azt is, hogy ha egyes összefüggések részmintánként fellelhetők, abból még nem következik az, hogy azokat a teljes mintában is megtaláljuk, de érdekes módon, ha a részmintákban két jellemző között nincs kapcsolat, a teljes anyagban mégis találhatunk szoros összefüggést (2). Ebből többek között az következik, hogy bizonyos tényezők egy-egy szűkebb területet (részpulációt, populációt) tekintve egymástól függetlennek tűnhetnek, de egy tágabb környezetbe helyezve kapcsolatuk már lényeges lehet.

Bár vizsgálódásaink során általában több tényezőt veszünk figyelembe, a statisztikai feldolgozáskor mégis többnyire meglegszünk az egy- és kétváltozós módszerek

használatával, magyarázatával, értelmezésével, pedig így általában nem kapunk teljes választ a feltett kérdésekre. És gyakran még ezen eljárások végigvitele sem következetes, nem mindig használjuk ki a továbblépés lehetőségeit.

Nagyon sokszor előfordul, hogy egy megfigyelt, vagy származtatott jellemző szemléletesen is jól észlelhető eltérést mutat két részmintát tekintve. Ilyenkor gyakran az eltérést kimondó megállapítást találunk csak, de hiányzik az eltérés szignifikáns voltát igazoló statisztikai próba. Ez nem engedhető el még akkor sem, ha az eltérést nyilvánvalóan nagyra gondoljuk, mert pl. valamelyik rész minta kis elemszáma miatt szemléletünk "csal"-hat és az eltérés statisztikailag nem tekinthető szignifikánsnak. Előfordulhat az is, hogy az átlagban azonos abszolútértékű különbség a szórások eltérése miatt egyik esetben szignifikánsnak minősítendő, a másikban pedig nem. (Változók szórásainak különbségéről van itt szó és nem a részmintákéről!)

Térjünk most vissza az összefüggések szorosságánál feltett kérdésre. Kisszámú tényező esetén például a korrelációs-mátrix elemeinek páronkénti összehasonlításával is feltárhatjuk a változók összefüggő blokkjait, nagy változószám esetén azonban ez már nem járható út. Ilyenkor a többváltozós módszerek az adekvát eszközök. Ezen modellek segítségével válaszolhatjuk meg többek között azt a kérdést, hogy mik azok a jellemzők, amelyek valamilyen szempont alapján tömböt, blokkot alkotnak, hogy az egyes csoportokon belül milyen mértékű a kapcsolódás, hogy az egyes csoportok között milyen mértékű a szétválás. Így e modellek lehetőséget adnak a megfigyelt jellemzők strukturális (faktoriális) elemzésére, lehetővé teszik a mögöttes, meghatározó komponensek feltárását. A strukturális elemzés két alapvető eljárása a klaszter- és a faktoranalízis (pszichológusok által hagyományosan igen gyakran használt módszerek). Ezek alkalmazására, ill. az alkalmazás során felvetődő néhány kérdésre mutatunk be két példát.

Az első példa szociális pályára felkészülő középiskolás tanulók személyiségvizsgálati anyagából és annak elemzéséből származik (3). Egy-egy populáció vagy részpopuláció személyiségstruktúrájának feltárásánál talán még izgalmasabb kérdés az, hogy a megfigyelt jellemzők kapcsolódása különbözik-e a részpopulációkban, pályakörökben, illetve kimutatható-e eltérés a struktúrában ugyanazon részpopuláció időbeni változása, fejlődése során, a pályára való felkészülés szempontjából alapvetően fontos időszakok között. A bemutatandó példában az egyik alapkérdés az volt, hogy a megfigyelt személyiségjellemzők között milyen a kapcsolat, más szavakkal az, hogy milyen struktúrában rendezhetők. Erre a kérdésre klaszteranalízissel kerestük a választ. A klaszteranalízissel kapott empirikus struktúrák egymáshoz elég közel álltak. Egyrészt megnéztük az egészségügyi tanulók önszabályozás változóinak belső szerkezetét, másrészt összehasonlítottuk a tanulmányok elején és végén nyert struktúrákat. Meglehetősen laza szerkezetet kaptunk mindkét esetben. A közöttük lévő eltérést a Boorman és Arabi által javasolt szintmértékeken alapuló távolság (4) segítségével határoztuk meg, melynek értéke 0 és 1 között változhat. Az ezzel számított struktúra-távolság 0,058-nak adódott. Vizsgáltuk az elméleti és az empirikus struktúra közti távolságot is. A tanulmányok elején ez 0,17-nak, a tanulmányok végén 0,16-nak adódott. A vendéglátóipari tanulóknál ugyancsak megnéztük a tanulmányok elején és végén nyert empirikus struktúrák közötti távolságot. Ezt 0,073-nak találtuk a tevékenység szabályozás változóira vonatkozóan. Meghatároztuk

még a tanulmányok végén az egészségügyi és vendéglátóipari tanulók empirikus struktúrái közti távolságot is, ami 0,071 volt.

A bemutatott esetekről megállapíthatjuk, hogy

- egyrészt az empirikus struktúrák közel állnak egymáshoz, de nem teljesen esnek egybe, akár a tanulmányok elején és végén kapottakat, akár a különböző pályára készülöket hasonlítjuk össze,
- másrészt ennél lényegesen nagyobb az eltérés, ha az elméleti struktúra és az empirikus struktúra távolságát tekintjük (bár ez utóbbi is kisebb a közepesnél).

Még két megjegyzést kell itt tennünk.

- Úgy fogalmaztunk, hogy "lényegesen nagyobb", de ezt a megállapítást statisztikai eszközök híján igazolni nem tudjuk ezzel a távolság-fogalommal.
- Mit jelent ez az eredmény? Jelentheti például azt, hogy a mérőeszköz korrekcióra szorul, mert nem illeszkedik az elméleti struktúra, vagy azt, hogy az elméleti struktúra szorul finomításra.

Ennek eldöntésében a statisztika és a statisztikus szakember közvetlen segítséget nem tud adni, ez a pszichológus kutató feladata.

A klaszterelemzés alkalmazásakor további szempontokra is tekintettel kell lennünk. Ezek közül még kettőt említünk meg:

- fontos az, hogy figyeljünk a korrelációs együtttható nagyságára, azaz arra, hogy a hierarchikus klaszterstruktúrából kiválasztott megoldásban csak olyan kapcsolódások szerepeljenek, amelyek mögött álló korrelációs együtttható (vagy más hasonlósági mutató) meghaladja a megkövetelt szignifikancia szintet,
- ugyancsak lényeges, hogy maga a klaszterelemzés jósági mutatója - a klaszterezés minőségét jelző mutató - elfogadható legyen.

A klaszterelemzések eredményeinek bemutatásánál általában sem a szignifikancia szintet, sem a jósági mutatót nem szokták közölni. Ha azonban ezek bármelyikét nem vesszük figyelembe, akkor következtetéseinket félrevezető eredményre építhetjük.

A második példánk a JATE BTK Oktatási Nevelési Bizottsága vizsgálatához (2) kötődik, melyben felmérte, hogy hogyan értékeli a kari képzésben alapvető tanárképző munka hatékonyságát - munkatapasztalataik tükrében - az 1979-83-ban tanárszakon, nappali tagozaton végzett egykori bölcsészek. Azaz pályakezdő fiatalok visszatekintő értékelését kértük a pályára való felkészítés komponenseire vonatkozóan (a pedagógiai, a pszichológiai, a szaktárgyi, az ideológiai stb. területek vonatkozásában).

A vizsgálat egyik részében faktoranalízissel elemeztük a véleménystruktúrát. Itt az alkalmazott módszer egy olyan bonyolult statisztikai eljárás volt, melynek lényege az, hogy a faktoranalízis eredményeként kapott faktortérben tovább vizsgáljuk a véleményeket hordozó változók halmazának belső szerkezetét. Az egymástól szignifikánsan elkülönítható változócsoportok kiválasztására egy speciális procedúrát - irányadatok elemzésére vonatkozó tesztet (5) - használtunk.

Meghatároztuk tehát (minden részletében statisztikailag alátámasztva) a véleménystruktúrát.

Tovább menve összehasonlítottuk a különböző szakaszokon végzettekre nyert eredménystruktúrákat (a korábban már említett módszerrel). Eredményül azt kaptuk, hogy jól látható különbségek vannak az egyes szakok között. Példaként említjük, hogy a

történelem és az orosz szakosok között 0,195; a történelem és az angol szakosok között 0,395; a magyar és az orosz szakosok között 0,238; az orosz és az angol szakosok között 0,295 az empirikus struktúrák távolsága. Jól látható, hogy a történelem és az angol szakosok vélemény-struktúrája tehát távol áll egymástól, az orosz és a történelem szakoské esik egymáshoz a legközelebb. A különbségek ellenére a struktúrák megegyeznek abban, hogy a fogalmilag azonos csoporthoz tartozó komponensek mindenütt összekapcsolódnak (például a pedagógiai, pszichológiai tárgyak értékelése mindenhol egy csoportba jön össze). Az viszont már változik szakonként, hogy az egyes fogalmi csoportokhoz milyen más fogalmi blokkok esnek közel. Általában elmondhatjuk, hogy az volt a fontos, hogy a kari munka melyik területéről kértük a véleményt, nem pedig az, hogy milyen szempontból kellett minősíteni az egyes oktatási, nevelési területeket.

A statisztikai vizsgálatok gyakran véget érnek a faktoranalízissel (vagy különálló faktoranalízisekkel) és az eredmények értelmezésével, magyarázatával. A második példa alapjául szolgáló kutatásban azt kaptuk, hogy a részmintákban a faktoranalízis eredményei különbözőek. Éppen ezért kellett tovább folytatnunk az elemzést, hogy kideríthessük, mi rejlik e mögött, hogyan értelmezhetők ezek a különbségek és miért, hogy a különbségeket kvantitatív eszközökkel jellemezhetjük.

A klaszteranalízisnél is gyakran megállunk ott, hogy kaptunk egy struktúrát, de innen is tovább lehet lépni, nevezetesen a bemutatott úton, például az elméleti és empirikus struktúrák viszonyának, vagy az empirikus struktúrák egymáshoz való viszonyának kvantitatív jellemzése felé.

Az előzőekben kiemeltük és példával illusztráltuk az elemzési folyamat néhány csomópontját, korántsem beszélve az összesről. Hangsúlyozni szeretnénk végül azt, hogy a bonyolultnak tűnő strukturális elemzési eljárások is továbbvihetők a pontosabb, megalapozottabb válaszok érdekében. Csak mintát mutattunk be erre a klaszter- és faktorelemzésnél egyaránt, más kutatási feladatoknál más lépések vezetnek eredményre.

IRODALOM

- 1 VALENTINE, E.R.: Fogalmak és nézőpontok a pszichológiában. Gondolat (1988) Budapest.
- 2 HUNYA Péterné: Volt hallgatóink véleménystruktúrájáról és a benne kimutatható tendenciákról. Az egyetemi oktató-nevelő munka időszerű kérdései karunkon VI., Végzett hallgatóink véleménye (1988) Szeged, pp. 62-68. Szerkesztő: Szalamin Edit.
- 3 HELEMBAI Kornélia: A középfokú szociális pályakörök identifikációs folyamatának elemzése a szakképzésben. Kandidátusi értekezés (1988) Szeged.
- 4 BOORMAN, S.A. - P. ARABIE: Structural measures and the method of sorting. Multidimensional Scaling, Vol. 1., pp. 226-249., Seminar Press (1972) New York, London.
- 5 MARDIA, K.V.: Statistics of directional data. Academic Press (1972) London.
- 6 KINDLER J. - PAPP O.: Komplex rendszerek vizsgálata. Műszaki Könyvkiadó (1977) Budapest.

PROBLEMS OF THE APPLICATION OF MATHEMATICAL STATISTICAL METHODS IN THE PROCESSING OF DATAS FROM THE EMPIRIC PSYCHOLOGICAL EXAMINATIONS

The solution of the investigational work in many cases means the application of different series of methods, followed by the conceptional-contential analysis of them, based on the data of some kind of survey. In the probe, the population, the observed model and the relation between those two should be stated very clearly. According to articles published in the U.S.A. and in England, that examine the situation, it is the least reassuring. Another important question is the object of the observation, and the means of registering the results. In this point of view the treatment of missing answers and the measuring of the examined features and their scale-theoretical belonging is fundamental. We have to deal distinctively with the hypothesis-examination and the statistical tests connected to it. Although in the course of our inquiries we deal with many factors, during the statistical processing the use of the single- or double-variabed methods is not logical, we don't always make use of the possibility of proceeding further. We show examples to present the difficulties in multi-variabed analysis, and the way of proceeding further counting on the usual structural evaluation. Besides the general conclusions models are shown to present the way of carrying further the analysis, but in different investigational tasks different techniques can be used.