

H. P. HUBER

Ludwig-Bölkzmann-Institut für Angewandte Klinische Psychologie, Graz

FEJLŐDÉSI TENDENCIÁK AZ EGYEDI ESETEK STATISZTIKAI ELEMZÉSÉBEN: HOL TARTUNK MA?

1. Bevezetés

Az 1916-tól 1933-ig Hamburgban oktató pszichológus, William STERN volt az első, aki az egyedi esetek vizsgálatát mind tartalmi, mind módszerbeli szempontból kívánatosnak tartotta. „Differentiellen Psychologie” című művének harmadik fejezete, amelyet STERN az „individualitások kutatásának” szentelt, sajátos védőbeszéd az egyedi-eset-orientált pszichológia mellett. STERN rámutatott arra, hogy az „individualizált pszichográfia, amely a nomotetikus pszichológia szükségszerű kiegészítője, mind elméleti, mind gyakorlati szükségletek miatt kívánatos...” (1921, 320. old.). Azonban azt is felismerte, hogy „a statisztikai módszer nem csak nagyszámú egyed, hanem egyetlen egyén esetében is alkalmazható” (1921, 341. old.).

STERN különbséget tett vertikális pszichogramm, amelynek feladata az egyéni fejlődés vizsgálata, és horizontális pszichogramm között, amelyben a „szinkrónia” áll az előtérben: az egyén pszichéjének nem a geneziséét, hanem a struktúráját kell feltárni. A pszichográfiai eljárásokat „módszer nélküli”, „kevert módszerű” és „elszigetelt módszerű” csoportokra osztotta. Az elszigetelt módszerű pszichogrammmok közé STERN csak azokat a vizsgálatokat sorolta, amelyekben bizonyos tulajdonságoknak vagy egy egyén magatartásmódjainak mennyiségi elemzését is elvégezték.

STERN után mindenekelőtt amerikai és angol kutatók foglalkoztak az egyedi eset elemzésének módszertani problémáival. Egy szimpóziumon, amely „Statisztikai ismeretek a klinikus számára” témában került megrendezésre, ZUBIN (1950) megfogalmazta az egyedi eset statisztikai vizsgálatának „axiómáját”. SAPHIRO angol pszichológus számos publikációban (többek között 1957, 1961) nyilvánította álláspontját a klinikai-pszichológiai egyedi esetek elméletével és gyakorlatával kapcsolatban, és meggyőző példákkal mutatott rá arra, hogy érvényes pszichológiai leírás egyedi esetben is lehetséges, ha fontos prediktív változókat sikerül megragadni. Az Egyesült Államokban CHASSAN statisztikus (lásd többek között 1960, 1967) erősen támogatta az egyénközpontú kutatást a pszichopatológiában. Mint annak idején STERN, CHASSAN is két egymást kiegészítő szempontra irányította a figyelmet. Az „extenzív modellre”, amely a csoport-statisztika vizsgálatán alapul, valamint az „intenzív modellre”, amely egy valószínűségi élmény- és viselkedés-fogalmon nyugszik és egyedi esetben is lehetővé teszi a véletlen kiküszöbölését. Az intenzív modellt GLASS, WILL-

SON és GOTTMAN (1975) időszereleméleti alapokra helyezte. Ennek a munkának az érdeme, hogy a BOX és JENKINS (1970) által kifejlesztett ARIMA-modellt (lásd később: időszerelemző modellek) a viselkedéstudományokba bevezették.

EDGAR és BILLINGSLEY már 10 évvel ezelőtt (1974) megállapították, hogy az intenzív modellel kapcsolatos, egyre növekvő számú publikációk ellenére az egyénközpontú vizsgálat a pszichológiai kutatásban még mindig a „tudománytalanság aurájával” van övezve. Ma — csaknem száz évvel EBBINGHAUS (1885) klasszikus egyedi eset vizsgálata után — fel kell tennünk a kérdést: a rendelkezésre álló statisztikai módszerek képesek-e garantálni az ilyen vizsgálatok megbízhatóságát. Arra a filozófiai kérdésre, vajon az egyén mint olyan a tudományos kutatás legitim tárgya-e, itt nem adható válasz. Egy nomotetikus NEM valószínűleg éppen olyan helytelen, mint egy idiografikus IGEN. Ezzel szemben egyedi esetek megkülönböztetése célszerűségének kérdését WESTMEYER-el (1979. 20. f. old.) egyszerűen meg tudjuk válaszolni: minden olyan hipotézis esetén (diagnosztika, terápia-ellenőrzés, vagy exporlatív kísérlet területén), amelyeknél egyénekről történnek kijelentések, az esetek elemzése a megfelelő, ahol azonban „aggregáthipotézisek” vizsgálatáról van szó, ott a választandó módszer a csoportjellemző elemzése.

2. Egyedi adatok vizuális elemzéséről

A komplex tényállások szemléltetésének egyik fontos és mind máig nélkülözhetetlen eszköze a tudományos kutatásban és érintkezésben a grafikai módszerek alkalmazása (vö. WAINER és THISSEN, 1981). Idősor adatok vizuális elemzése azonban, mint ahogy ezt nemrégén BALLARD (1983) kifejtette, csak egy rákövetkező statisztikai kiértékelés exploratív előstádiuma lehet. Sajnálatos módon a SKINNER-i iskola, amely az egyedi kísérletet kutatói stratégiája integráns részének tekinti (vö. SIDMAN, 1960), az egyedi esetelemzés statisztikai kérdéseit sohasem tekintette elsőrendű problémának. Ugyanez érvényes — néhány kivételtől eltekintve (vö. pl. YATES, 1958) — a viselkedésterápiai gyakorlatra is. A viselkedés-elmélet által inspirált értekezések és monográfiák többsége — mint például HERSEN és BALOW (1976) vagy KAZDIN (1982) könyvei — csaknem kizárólag a különböző inverz és többszörös alapgörbék bemutatásával foglalkoznak, amelyeket rendszerint leíró kritériumok alapján értékelnek. HERSEN és BALOW (1976) az induktív-statisztikai aspektusnak csak egy rövid fejezetet szentel (KAZDIN-tól); KAZDIN-nál (1982) a statisztikai módszerek egyáltalában csak a függelékben kerülnek szóba. Jellemző az értékrendre, amellyel a statisztikai módszert az egyedi-eset-orientált viselkedéskutatásban illetik, PARSONSON és BAER (1978, 162. old.) felfogása, amely szerint a rendelkezésre álló statisztikai eljárástól függetlenül, a grafikus szemléltetett adatok vizuális analízisének kell elsődlegesnek maradnia. Ez éppoly kevésbé járul hozzá ahhoz, hogy az egyedi eset vizsgálatok hitele megszilárduljon, mint BALLARD (1983, 72. old.) ajánlása, mely szerint az idősorok vizuális analízisében kell nagyobb gyakorlatot szerezni.

3. Varianciaanalízises eljárások

Az egytényezős egyedi eset kísérlet varianciaanalízissel történő feldolgozásának egyik módszerét SHINE és BOWER (1971) javasolta; ennek az eljárásnak az általánosítását többtényezős és többváltozós esetekre SHINE (1973, 1982) végezte el. Feltételezzük, hogy az egyént „response generator”-nak (válaszadó gépnek) lehet tekinteni, akinek a reakciói egy meghatározott várható érték körül, egymástól (kísérletileg) függetlenül, normális eloszlás szerint szóródnak. A SHINE-BOWER-analízis prototípusa szerint egy kísérleti személyt több kísérleti menetben meghatározott, esetleg egymást kiegyenlítő sorrendben következő feltételek között kell megfigyelni. Míg a kísérleti lépéseket, melyek egy segédtenyező különböző erősségi fokait reprezentálják, véletlen hatásokkal számolva tervezik meg, a kísérleti tényezőnek szilárd hatásúnak kell lennie. A SHINE-BOWER-analízis a szokásos varianciaanalízisektől mindenekelőtt a hibavariancia becslésében különbözik. A kísérleti menet lépései közötti konstans eltolódásokat szűri ki a menet-faktor. Ennek az eljárásnak a helyessége annak a feltételezésnek a jogosságán áll vagy bukik, hogy (a) a kísérlet lépései között fellépő ismétlési hatások enyhén emelkednek vagy csökkennek, vagy (b) stationer esetben az idősorbeli függőségeket el lehet hanyagolni. A gyakorlatban ilyen feltételeket leginkább abban az esetben lehet találni, amikor egy tanulási folyamatot leíró görbe egy egyeneshez simul hozzá és az időkülönbségek a kísérleti lépések között megfelelően nagyok. A módszer megkívánja, hogy a lépések közötti csökkenés – emelkedés visszafordítható legyen (kezelési hatás).

A SHINE-BOWER-analízis alternatíváját jelentik a COX (1951, 1952) szerinti „Szisztematikus kísérleti tervek”. A randomizált blokkok módszerével ellentétben a kezelés sorrendjét rendszeres szempontok szerint úgy választják, hogy az a várható (első vagy magasabb fokú) trendhez viszonyítva *ortogonális* legyen. A kezelés hatásnak ismét visszafordíthatónak kell lennie. Továbbá azt is feltételezik, hogy a lépések közötti enyhe emelkedő vagy csökkenő trendhatás utáni maradék csökkenés szakaszonként azonos szórású normális eloszláshoz vezet. A FISHER-féle ortogonális polinom segítségével a trend-összetevőket jól el lehet választani a kezelési hatásoktól. A szisztematikus kísérleti tervek vitathatatlan előnyét abban a körülményben látjuk, hogy a megfigyeléseknek már igen kis számával is viszonylag hatékonyan lehet dolgozni. A kezelési sorrend szükséges ortogonalitása viszont nem lényegtelen korlátozást jelent a kísérleti terv összeállításánál. Ráadásul a többfokozatú kezelési tényezővel operáló kísérleti tervek esetében hamar bekövetkezhet az az eset, hogy a tervezett mérés számhoz numerikus okokból nem képezhető ortogonális kezelési sorrend és közelítő megoldásokat kell keresni (COX, 1952).

Monte-Carlo-futtatások segítségével PHILLIPS (1983) megkísérelte a sorosan korrelált reziduumoknak az F-próbára gyakorolt hatását varianciaanalitikus egyedi eset tervek különböző típusainál tanulmányozni. Megmutatkozott, hogy a kezelési hatások megítélése nagymértékben függ a kísérlet összeállításától, de legalábbis az F-próba eredményei erősen torzítottak. A trend-összetevők szignifikanciavizsgálatánál a negatív korreláció a reziduumok között többnyire döntéseméletileg konzervatív ítéletekhez vezet, míg a sorosan pozitívan korrelált reziduumok általában a szignifikancia szintet növelik.

4. Idősorelemző módszerek

CATTELL-re (1966, 9. old.) támaszkodva célszerűnek látszik manipulatív és nem manipulatív idősortervek között különbséget tenni. Nem manipulatív egyénkísérleteket longitudinális összefüggések vizsgálatára folytatnak, ezzel szemben a manipulatív vizsgálatok különféle beavatkozások hatáselemzését szolgálják. Az idősoranalitikus esettanulmányok az ARIMA (p, d, q)-modellek általános osztályára épülnek (BOX és JENKINS, 1970). Az ARIMA az „autoregressive integrated moving average” (autoregresszív integrált mozgóátlag) rövidítése. A p, d és q strukturális paraméterek a véletlen összetevők és az idősor többi része közötti viszonyt írják le, a d paraméter a különbségképzés fokszámát, a p az autoregresszív folyamat rendjét és a q a mozgóátlag folyamatok rendjét jelöli.

A legfontosabb stratégiákat, amelyeket a nem manipulatív idősor-vizsgálatoknál használnak, CATALANO, DOOLEY és JACKSON (1983) mutatták be összefoglalóan. Voltaképpen statisztikai rutin eljárások hierarchiájáról van szó, amely többek között a következő elemeket foglalja magában: (a) „prewhitening” (az idősor véletlen összetevőinek eltávolítása), (b) előszűrés (a kritériumsor reziduumaik eltávolítása a folyamat modelljében szereplő predikátor-sor segítségével), (c) az (a) és/vagy (b) szerint előkezelt sorok keresztkorrelációs kiértékelése, valamint (d) egy transzfer-modell felállítása. A kiértékelő lépések helyes megválasztása érdekében fontos tudni, vajon vannak-e előzetes információk az időbeli előzmény struktúrájáról, vagy ezeket először meg kell szerezni.

A manipulatív idősorvizsgálat prototípusa CAMPBELL és STANLEY (1963) félbeszakított idősorozat kísérlete. Egy eleve rögzített időpont kettéosztja az idősort elő- és utókezelő fázisra. Amennyiben a függőségi paraméterek rendjéről nincs eleve értesülésünk, úgy az ARIMA-modell megbízható becsléséhez legalább 100 olyan megfigyelés szükséges, amelyek egyenlő időközönként következnek (BOX és JENKINS, 1970). Ilyen hatáselemzés csak akkor lehetséges, ha esetleges trend-összetevők, mint autoregressziós- és mozgóátlag folyamatok statisztikailag kontroll alatt tarthatóak. A kezelési hatás lappangásától és lefolyásától függően különféle modelleket lehet specifikálni. Bizonyos okokból, amelyekre ehelyütt nem térhetünk ki részletesebben, MÖBUS, GÖRICKE és KRÖH (1984); BOX és TIAO (1975) modelljét ajánlják szemben GLASS, WILLSON és GOTTMANN (1975) GLM-ével.

Az idősoros vizsgálatok hatás-elemzésébe didaktikusan kiváló bevezetést találunk McDOWALL-nál és szerzőtársainál (1980). Német nyelven REVENSTORF (1979) írt bevezetőt a témáról. Az egyének közti trend-összehasonlítás fázisspecifikus függőségi modelljeivel foglalkozott HORNE és YANG (1982).

Az ARIMA-módszer alkalmazása a mai szoftver-kínálat mellett lényegesen könnyebb (vö. a BASIC-ül írt TIMCAI interaktív programmal BROWN-tól, 1982). Az első felhasználási tapasztalatok arra utalnak, hogy élmény és viselkedés-pszichológiai idősor adatok többnyire igen egyszerű függőségi struktúrát mutatnak, legtöbbször autoregresszív és mozgóátlag folyamatokkal találkozunk (vö. DAHME, 1979). Többszörösen is indokolt a kétely, vajon a klinikai-pszichológiai egyedi eset vizsgálatokból származó adatok kielégítik-e egy ARIMA-elemzés igényes elvárásait. Kísérlet-tervezői nézőpontból emlékeztetünk arra, hogy bár az ARIMA-modellek elegendőek, az idősor félbe-

szakítása kvázi kísérleti elrendezésnek számít, CAMPBELL és STANLEY (1963). De arra nincs mód, hogy egymással vetélkedő külső tényezők hatását (pl. pszichoszociális kísérőjelenségek kezelést elősegítő hatásait) kontrolláljuk.

5. Randomizációs tesztek

EDGINGTON (1967, 1975, 1980, 1982) munkái különös figyelmet érdemelnek, a kanadai pszichológus meggyőző módon tudta kimutatni, hogy egyénvizsgálatokban a reakciók megítélésénél a randomizáló teszt lehet a célszerű módszer, ha a tervezett kezelési rendszabályok és vizsgálati időpontok véletlenszerűen rendelhetők egymáshoz. Amennyiben tartani kell kezelés-specifikus transzferhatástól, vissza nem fordítható kezelési hatásokkal kell számolni, vagy terápiás okokból a randomizált kezelési eljárás nem javasolt, úgy EDGINGTON (1975, 1980) azt ajánlja, hogy az intervenció időpontot határozzuk meg véletlenszerűen.

A randomizációs módszer FISHER-re (1936) és PITMAN-ra (1937) vezethető vissza. A randomizáló tesztek olyan nem parametrikus eljárások, amelyek lehetővé teszik, hogy a mérési értékek közötti különbségről információt nyerjünk. Amennyiben a különbség-skála minősége tekintetében kétségek merülnének fel, úgy a mérési értékeket mindenkor pótolni lehet azok rangjával. Eloszlás-elméletileg az egyedi esetben nyert intraindividuális mérési értékektor egy többdimenziós véletlen változó megvalósulásának tekinthető. Annak a null-hipotézisnek a teljesülését, hogy a mérési értékvektorok valamennyi permutációja azonos valószínűséggel fordul elő, a kísérleti koncepció randomizálásával biztosítjuk. Kezelés-specifikus mérési értékfüggőségek nem zavarják a permutációk egyenlő eloszlásának feltételezését, mindenesetre a null-hipotézis értelmében a kezelési feltételek érvényességi tartományán belül nem lehetnek különbségek az egyes időpontok korrelációi között. A randomizáló tesztek nagy előnyét abban kell látnunk, hogy az egész, ismeretlennek elfogadott, kezelés-specifikus függőségi struktúrát a null-hipotézis mellett is meg lehet tartani (vö. KRAUTH, 1981). A randomizáló tesztek egyik hátránya, hogy nem lehet táblázatba szedni. Ezt a körülményt azonban modern mikrokomputerek kapacitását figyelembe véve nem szabad túldramatizálni. EDGINGTON (1980) értékes utalásokat adott a randomizáló tesztek komputeres programjának előállításához.

6. Pszichometrikus egyedi esetelemzés

A különböző teszt-elméleti tételek abból indulnak ki, hogy egy személy vizsgálata pszichometrikus eljárással egy véletlen változót hoz létre. Bizonyos feltételek között a vizsgálati varianciákat nem kell az egyedi esetben mért teszt-értékből becsülni, hanem le lehet vezetni az alappopuláció paramétereiből. Mint ahogy HUBER (1973) rámutatott, a klasszikus teszt-elmélet alapján többféle megítélési szempontot lehet megkülönböztetni: A „megbízhatóság megítélési aspektus” esetében az a kérdés érdekel bennünket, vajon az egyedi esetben megfigyelt tesztérték-különbség az alkalmazott skálák mérési hibáira való tekintettel mint mérési vagy mint reális különbség fogható-e fel, a

„diagnosztikai érvényesség” aspektusa mellett pedig azt kérdezzük, vajon egy megvizsgált kísérleti személy még a pszichikailag egészséges populációhoz számítható-e.

A klasszikus teszt-elmélet koncepciójának átvihetőségét rangszámokra röviddel ezelőtt WOLFRUM (1984) vizsgálta. Ebben az összefüggésben többek között felmerült az ordinális skálára felvitt teszt-adatok statisztikai feldolgozásának lehetősége is.

ZIMMERMANN (1975), valamint ZIMMERMANN és WILLIAMS (1977) megfontolásaihoz csatlakozva TACK (1979) megvizsgálta az egyedi esetek analízisének tesztelméleti alapjait. Az egyedi esetek vizsgálatának jobb megalapozása érdekében azt javasolta, hogy egyes személyek helyett személyek és feltételek bizonyos kombinációit tekintsük véletlen változók által reprezentált elemzési egységeknek. Az ilyen elemzési egységek esetében a párhuzamosság, homogenitás és érvényesség teszt-elméleti koncepcióit is újra meg kell fogalmazni.

7. Az egyedi esetvizsgálatok ismételhetségiéről

A klinikai-pszichológiai gyakorlatban általában individuáldiagnosztikai okokból, vagy az egyéni terápia ellenőrzése céljából végeznek egyedi esettanulmányokat, mint ahogy WESTMEYER (1979) is mondta „kivételes-egyedi”, vagy „idiografikus” hipotéziseket vizsgálnak meg. Az egyedi esetekből leszűrt tapasztalatokkal kapcsolatban azonban gyakran merül fel a kérdés, vajon az egyik vagy másik betegen mért értékek ugyanennek a nosológiai kategóriának további páciensein is reprodukálhatók-e. Más szavakkal: arra tesznek kísérletet, hogy a tapasztalati eredmények halmozódó feldolgozásával keletkezett „kvázi-megfordíthatóság” hipotézisét a megfelelő esetre vonatkoztatott „kivételes-egyedi vagy idiografikus” hipotézisek segítségével szisztematikusan vagy véletlenszerűen kiválasztott betegek egész sorától nyert „háttér-információval” igazolják (vö. WESTMEYER, 1979, 29. old.). A PETERMANN (1982, 52. old.) által kidolgozott koncepció, az „ellenőrzött praxis” esetében a halmozódó tapasztalat felhasználása olyan munkalépést jelent, amelyben „különböző terápiákból származó tapasztalatok összevetése” zajlik. Nyilvánvaló, hogy a pszichoterápiás kiértékelő folyamatnak az az átmeneti állapota érhető tetten, amelyben az esetre orientált munka „clinical trial” formájában szisztematikus kutatási munkával találkozhat (vö. HUBER, 1983).

Az egyéni vizsgálati eredmények más egyéneknél való reprodukcióját mint az intenzív módszer következetes továbbfejlesztését SHAPIRO (1961) ajánlotta. Azt a kérdést, hogy milyen statisztikai eljárással összesítették vagy foglalták össze egyetlen aggregált kijelentésre különböző személyektől nyert adatokat, az egyedi esetek szakirodalmában — ha egyáltalában — csak periférikusan taglalták. Célszerű ehelyütt a grafikai adatelemzés módszereire visszautalni (vö. WAINER és THISEN, 1981). Azok a kiértékelési stratégiák, amelyeket LIENERT (1978) az egyes folyamatok homogenitásának, együttjárásának és hierarchikus osztályozásának megítélésére javasolt összesítési célokra, ugyancsak felhasználhatók. SHINE (1973b) interindividually kibővítette a többfaktoros SHINE—BOWER-módszert úgy, hogy az egyedi eset statisztikai kiértékelése mellett mód van a kísérleti személyek összehasonlító elemzésére is. Idevágó eljárások keresésénél gondolni kellene arra, hogy hasonló problémák a klasszikus kísérlet-tervezésben is adódnak, ha több csoportos kísérlet eredményeit

kombinálják, amelyeket különböző mintákból nyertek. Ehhez alkalmas módszerek annak idején FISHER (1932), PEARSON, K. (1933) és PEARSON, E. S. (1938) ajánlottak és újabban WINER (1971) is visszatért ezekre, előnyeit és hátrányait LANCES-TER (1949) és MAXWELL (1961) tárgyalta. Semmi nem szól az ellen, hogy a homolog esettanulmányoknál, amelyeket meghatározott populáció véletlenszerűen kiválasztott személyeinél végeznek el, a csoportvizsgálatokhoz hasonló módon járunk el és az egyedi esetek statisztikai vizsgálati eredményeit egy statisztikai összefoglalóvá foglalkoztatjuk össze (vö. HUNBER, 1978).

Amennyiben a reprodukciós eljárás egy „quasi-univerzális” hipotézist igazolt, úgy hozzá lehet kezdeni az *exploratív csoporttanulmány* előkészítéséhez. Az előzetes egyedi esetelemzéseket és reprodukciós vizsgálatokat egyfajta „szűrőnek” lehet tekinteni, „amely abban dönt, hogy egyáltalában milyen fajta kezeléseket és milyen feltételek között lehet alkalmazni a több ráfordítást igénylő és kevésbé rugalmas csoportvizsgálatokban” (WESTMEYER, 1979, 29. old.).

8. Az egyedi eset statisztikájával szembeni elvárások

Vitathatatlan, hogy az egyedi esetvizsgálatok tervezésénél és kiértékelésénél haladás mutatkozik, ez azonban nem téveszthet meg bennünket a nyitott kérdések tekintetében, amelyekből néhányról a következőkben lesz szó.

(1) Az ARIMA-modellek hozzájárultak ugyan a faktoriális szerkezetű egyedi eset kísérletek hibái sorozatos korrelációjának jobb megértéséhez, de a problémát nem oldották meg. Az a kézenfekvő gondolat, hogy az eredeti megfigyelt értékeket esetleges ARIMA-részekről szabadítsák meg, különböző okokból többnyire megvalósíthatatlan: (a) A megfigyelt sorok rendszerint túl rövidek ahhoz, hogy a függőségi paraméterek megbízható becslését el lehessen végezni; (b) az ARIMA-szűréssel együttjáró mérési értékvesztés a kezelési folyamat kiegyensúlyozottságának és ortogonalitásának korlátozását jelentené; (c) különösen rövid mérőértéksoroknál nagy a veszélye annak, hogy a vizsgálati statisztika F-eloszlását a függőségi paraméter becslése torzítja (vö. JOHNSTON, 1972).

(2) Habár egyedi esetvizsgálatoknál rendszerint többféle változót szerepeltetnek, az egyes változókat általában külön egyváltozósaként analizálják. Az egyedi eset statisztikának egy további feladata tehát a *többszörös feloldozási módszerek kifejlesztése*. A CATTELL P és O technikájának faktoranalitikus hagyományára (CATTELL, 1946) támaszkodó intraindividuális korrelációs analízisek a mért értékek függőségének problémáján buktak meg. Sokszoros idősorok analíziséhez döntő impulzusok majd csak QUENOUILLE-től (1957) származnak. Transzfermodellek és többváltozós ARIMA-modellek alkalmazására a pszichológiai egyedi eset statisztikában eddig hovatovább alig vannak tapasztalatok. Ugyanez vonatkozik a MANOVA-modellek (vö. SHINE, 1982, 1983) egyedi eset statisztikai adaptációjára is.

(3) A gyakorlatban keletkező teszt-eredmények nem jelentéktelen része pszichometrikus tesztvizsgálatokon alapszik. Az ellenőrzött esettanulmányok tervezése azonban ebben a körben mindenekelőtt azért okoz nagy nehézségeket, mert a legtöbb teszt

legjobb esetben egy párhuzamos formával rendelkezik. Elegendő számú sorozatosan sztenderdizált párhuzamos formával rendelkező teszt fejlesztése tehát egyedi eset-statisztikai szempontból a tesztkonstrukció egyik legfontosabb feladata. A sorozatos sztenderdizálás alap gondolata egyszerű: a párhuzamos formákat ugyanabban a sorrendben kell sztenderdizálni, mint ahogy később alkalmazni kell őket. Ily módon minden egyes tesztismétlés számára külön sztenderd-értékeket kaphatnánk, amelyek legalább egy meghatározott populáción belül sorozatosan egyenértékűek lennének.

(4) A kérdés, hogy a pszichológiai jellemzőket milyen finom skálán mérjük, gyakran (és eddig elhamarkodottan) az intervallum skálázhatóság javára kerül megválaszolásra. WOLFRUM (1984) ezzel szemben rámutatott, hogy akár az önmegítélésnél, akár a külső megítélésnél végbemenő ítélezési folyamatok esetében egy az ordinálisnál magasabb szintű mérhetőség feltételezése irreális. Ezen tényállásra való tekintettel kívánatos az egyedi eset kiértékelésére irányuló statisztikai erőfeszítéseket intenzívebbé tenni, megteremteni a nem parametrikus kiértékelés lehetőségét. Eloszlásmentes idősortesztek felhasználása tekintetében LIENERT (1978) és REVENSTORF és VOGEL (1979) adnak tájékoztatást.

9. Zárómegjegyzések

Az ismert biometrikus Joseph ZUBIN „A viselkedés etiológiája 2000-ben” című dolgozatában bátorkodott azt jövendőlni, hogy a véletlenül kiválasztott személyeken végzett személyiségkutatás a XXI. században elavult lesz, csoportstatisztikai vizsgálatok helyett egyedi esettanulmányokat fognak folytatni, amelyek számára a kísérleti személyek a saját magukról való tudásuk alapján Bayesianus apriori-valószínűségeket fognak tudni rendelkezésre bocsátani (ZUBIN, 1971, 20. old.). Magam nem osztom ZUBIN futurologikus optimizmusát, arról azonban meg vagyok győződve, hogy a széleskörűen elterjedt szkepszis az egyedi esetre koncentrááló módszerekkel szemben lassanként leépíthető, amennyiben (a) az adatszerzés módszerei messzemenően sztenderdizálva lesznek és a pszichometrikus kutatás színvonalával lépést tartanak, (b) a belső validitás (érvényesség) megfelelő intézkedések útján a kísérlet tervezésénél és a kísérletek végrehajtásánál biztosítva lesz, (c) a vizsgálati eredmények véletlenszerűségének biztosítása már a kísérlet tervezésénél célként szerepel, (d) ha – amennyiben exploratív kísérletekről van szó – a külső validitás követelményének az egyedi esetben kapott vizsgálati eredmények reprodukálhatóságának biztosításával eleget tesznek.

Irodalom

- BALLARD, K. D., 1983, The visual analysis of time series data: Issues affecting the assessment of behavioural interventions, *New Zealand Journal of Psychology*, 12, 69–73.
- BOX, G. E. P. and JENKINS, G. M., 1970, *Time-series analysis: Forecasting and control*, Holden Day, San Francisco.

- BOX, G. E. P. and TIAO, G. C., 1975, Intervention analysis with applications of economic and environmental problems, *Journal of the American Statistical Association*, *70*, 70–92.
- BROWN, R. L., 1983, TIMCAI: An interactive program to assist in the identification of univariate nonseasonal nonmixed ARIMA time series, *Behavior Research Methods and Instrumentation*, *15*, 101–102.
- CAMPBELL, D. T. and STANLEY, J. S., 1963, Experimental and quasi-experimental designs for research, In: GAGE, N. L. (ed.): *Handbook of research on teaching*. Rand McNally, Chicago, 171–246.
- CATALANO, R. A., DOOLEY, D. and JACKSON, R., 1983, Selecting a time-series strategy, *Psychological Bulletin*, *94*, 506–523.
- CATTELL, R. B., 1946, *The description and measurement of personality*, World Book, New York.
- CATTELL, R. B., 1966, Guest editorial: Multivariate behavioral research and the integrative challenge, *Multivariate Behavioral Research*, *1*, 4–23.
- CHASSAN, J. B., 1960, Statistical inference and the single case in clinical design, *Psychiatry*, *23*, 173–183.
- CHASSAN, J. B., 1967, *Research design in clinical psychology and psychiatry*, Appleton-Century-Crofts, New York.
- COX, D. R., 1951, Some systematic experimental designs, *Biometrika*, *38*, 312–323.
- COX, D. R., 1952, Some recent work on systematic experimental designs, *Journal of the Royal Statistical Society*, *14*, 211–219.
- DAHME, B., 1979, Statistische Analyse kurzer Zeitreihen in der klinischen Effekt-Prüfung, In: PETERMANN F. and HEHL, F. J. (eds.), *Einzelfallanalyse*, Urban and Schwarzenberg, München, 251–265.
- EBBINGHAUS, H., 1885, *Über das Gedächtnis*, Duncker and Humboldt, Leipzig
- EDGAR, E. and BILLINGSLEY, F., 1974, Believability when $N = 1$, *Psychological Record*, *24*, 147–160.
- EDGINGTON, E. S., 1967, Statistical inference $N = 1$ experiments, *Journal of Psychology*, *65*, 195–199.
- EDGINGTON, E. S., 1975, Randomization tests for one-subject operant experiments, *Journal of Psychology*, *90*, 57–68.
- EDGINGTON, E. S., 1980, *Randomization tests*, Dekker, New York.
- EDGINGTON, E. S., 1982, Nonparametric tests for single-subject multiple schedule experiments, *Behavioral Assessment*, *4*, 83–91.
- FISHER, R. A., 1932, *Statistical methods for research workers*, 4th, ed. Oliver and Boyd, London.
- FISHER, R. A., 1936, "The coefficient of racial likeness" and the future of craniometry, *Journal of the Royal Anthropological Institute of Great Britain and Ireland*, *66*, 57–63.
- GLASS, G. V., WILLSON, V. L. and GOTTMAN, J. M., 1975, *Design and analysis of time-series experiments*, Colorado Associated University Press, Boulder.
- HERSEN, M. and BARLOW, D. H., 1976, *Single-case experimental designs: Strategies for studying behavior change*, Pergamon Press, New York.

- HORNE, G. P., YANG, M. C. K. and WARE, W. B., 1982, Time-series analysis for single-subject designs, *Psychological Bulletin*, *91*, 178–189.
- HUBER, H. P., 1973, *Psychometrische Einzelfalldiagnostik*, Weinheim, Basel.
- HUBER, H. P., 1978, Kontrollierte Fallstudie, In: PONGRATZ, L. J. (ed.), *Klinische Psychologie, Band 8/2*, Hogrefe, Göttingen, 1153–1199.
- HUBER, H. P., 1983, Strategien zur Evaluation therapeutischer Prozesse, In: MINSEL, W. R. and SCHELLER, R. (eds.), *Forschungskonzepte der Klinischen Psychologie*, Kösel, München, 87–102.
- JOHNSTON, J., 1972, *Econometric methods*, 2nd ed., McGraw-Hill, New York.
- KAZDIN, A. E., 1982, *Single-case research designs*, Oxford University Press, New York.
- KRAUTH, J., 1981, Statistische Methoden zur Veränderungsmessung, In: BAUMANN, U., BERBALK, H. and SEIDENSTÜCKER G. (eds.), *Klinische Psychologie. Trends in Forschung und Praxis*, Huber Band IV., Bern, 98–131.
- LANCASTER, H. O., 1949, The combinations of probabilities arising from data in discrete distributions, *Biometrika*, *36*, 370–382.
- LIENERT, G. A., 1978, Verteilungsfreie Methoden in den Biostatistik, Ban, 2, 2. Aufl. Meisenheim: Hain.
- MAXWELL, A. E., 1961, *Analysing qualitative data*, Methuen, London.
- McDOWALL, D., McCLEARY, R., MEIDINGER, E. E. and HAY, R. A., 1980, Jr., *Interrupted time series analysis*, Sage Publications, Beverly Hills.
- MÖBUS, C., GÖRICKE, G. and KRÖH, P. A., 1983, Statistical analysis of single-case experimental designs: Conditional equivalence of the general-linear-model approach of Glass, Willson and Gottman with the intervention model of Box and Tiao, *EDV in Medizin und Biologie*, *14*, 98–108.
- PARSONSON, B. S. and BAER, D. M., 1978, The analysis and presentation of graphic data, In: KRATOCHWILL, Th. R. (ed.), *Single subject research. Strategies for evaluating change*, Academic Press, New York, 101–165.
- PEARSON, E. S., 1938, The probability integral transformation for testing goodness-of-fit and combining independent tests of significance, *Biometrika*, *30*, 134, 148.
- PEARSON, K., 1933, On a method of determining whether a sample of size n supposed to have been drawn from a parent population having a known probability integral has probably been drawn at random, *Biometrika*, *25*, 379–410.
- PETERMANN, F., 1982, *Einzelfalldiagnose und klinische Praxis*, Kohlhammer, Stuttgart.
- PHILLIPS, J. P. N., 1983, Serially correlated errors in some single-subject designs, *British Journal of Mathematical and Statistical Psychology*, *36*, 269–280.
- PITMAN, E. J. G., 1937, Significance tests which may be applied to samples from any populations, *Journal of the Royal Statistical Society*, *4*, 119–130.
- QUENOUILLE, M. H., 1957, *The analysis of multiple time series*, Hafner, New York.
- REVENSTORF, D., 1979, *Zeitreihenanalyse für klinische Daten. Methodik und Anwendungen*, Beltz, Weinheim.
- REVENSTORF, D. and VOGEL, B., 1979, Zur Analyse qualitativer Verlaufsdaten – ein Überblick, In: PETERMANN, F. and HEHL, F. J. (eds.): *Einzelfallanalyse*, Urban and Schwarzenberg, München, 229–250.

- SHAPIRO, M. B., 1957, Experimental method in the psychological description of the individual psychiatric patient, *International Journal of Social Psychiatry*, 3, 89–103.
- SHAPIRO, M. B., 1961, The single case in fundamental clinical psychological research, *British Journal of Medical Psychology*, 34, 255–262.
- SHINE, L. C., 1973a, A multi-way analysis of variance for single-subject designs, *Educational and Psychological Measurement*, 33, 633–636.
- SHINE, L. C., 1973b, A design combining the single-subject and multi-subject approaches to research, *Educational and Psychological Measurement*, 33, 763–766.
- SHINE, L. C., 1982, The multivariate case of the Shine–Bower single-subject ANOVA, *Educational and Psychological Measurement*, 42, 453–457.
- SHINE, L. C., 1983, An illustration of the multivariate case of the Shine–Bower single-subject ANOVA, *Educational and Psychological Measurement*, 43, 1033–1039.
- SHINE, L. C. and BOWER, S. M., 1971, A one-way analysis of variance for single-subject designs, *Educational and Psychological Measurement*, 31, 105–113.
- SIDMAN, M., 1960, *Tactics of scientific research*, Basic Books, New York.
- STERN, W., 1921, *Die differentielle Psychologie in ihren methodischen Grundlagen*, Unveränderter Abdruck der Ausgabe von 1911 vermehrt um ein Nachwort 1921 nebst neuer Bibliographie, Barth, Leipzig.
- TACK, W. H., 1979, Testtheoretische Grundlagen der Einzelfallanalyse, In: PETERMANN, F. and HEHL, F. J. (eds.), *Einzelfallanalyse*, Urban and Schwarzenberg, München, 49–69.
- WAINER, H. and THISSEN, D., 1981, Graphical data analysis, *Annual Review of Psychology*, 32, 191–241.
- WESTMEYER, H., 1979, Wissenschaftstheoretische Grundlagen der Einzelfallanalyse, In: PETERMANN, F. and HEHL, F. J. (eds.), *Einzelfallanalyse*, Urban and Schwarzenberg, München, 17–34.
- WINER, B. J., 1971, *Statistical principles in experimental design*, 2nd ed. McGraw–Hill, München.
- WOLFRUM, Chr., 1984, *Möglichkeiten zur Erstellung von Ordinalskalen*, Habil.-Schrift, Düsseldorf.
- YATES, A. J., 1958, The application of learning theory to the treatment of tics, *Journal of Abnormal and Social Psychology*, 56, 175–182.
- ZIMMERMANN, D. W., 1975, Probability spaces, Hilbert spaces, and the axioms of test theory, *Psychometrika*, 40, 395–412.
- ZIMMERMANN, D. W. and WILLIAMS, R. H., 1977, The theory of test validity and correlated errors of measurement, *Journal of Mathematical Psychology*, 16, 135–152.
- ZUBIN, J., 1950, Symposium on statistics for the clinician, *Journal of Clinical Psychology*, 6, 1–6.
- ZUBIN, J., 1971, The etiology of behavior in the year 2000, In: EVANS, W. O. and KLINE, N. S. (eds.), *Psychotropic drugs in the year 2000, Use by normal humans*, Thomas, Springfield, 3–24.

HELMUTH P. HUBER

**STATISTICS FOR SINGLE-SUBJECT DESIGNS:
A REVIEW OF SOME RECENT DEVELOPMENTS**

A landmark in the history of single-case analysis is EBBINGHAUS' investigation of memory. However, about 100 years after the publication of EBBINGHAUS' work an aura of "unscientific" still remains associated with single-subject research. In the view of this fact the following approaches to single-case analysis are discussed: (1) Visual analysis of time-series data, (2) ANOVA techniques, (3) time-series analysis, (4) randomization tests, (5) psychometric assessment procedures (based on classical test theory models) and (6) demonstration of generality by replication (methods of combining results from different single-subject investigations). Finally, attention is drawn to some desiderata of single-case statistics.