

Slobodan Cvejić  
Filozofski fakultet  
Beograd

UDK: 303. 723. 039. 3  
*Pregledan naučni rad*  
Primljen: 21. 11. 1997.

## UZROČNA ANALIZA KATEGORIJSKIH PODATAKA - PREDISTORIJA I POČETAK MULTIVARIJACIONE ANALIZE U SOCIOLOGIJI

*Durkheim je u statistici video jedno od sredstava za izdvajanje društvenih pojava iz pojedninačnih oblika njihovog ispoljavanja. On je ovladao konceptom multidimenzionalnosti u sociološkom smislu, ali ne i u statističkom smislu. Šta je nedostajalo Durkheimovom pristupu uzročnoj analizi neeksperimentalnih podataka da bi bio savršeniji? Ne samo višestruka regresiona ili korelaciona analiza, nego i niz postupaka za analizu tabela kontingencije koje su tih godina razvijali Pearson i Yule. Dalje, Durkheim nije uočio da prividnost veza može imati i obrnuti princip. Mnoge veze koje se učine beznačajnim mogu biti uzročnost u dubljim determinističkim slojevima. Ipak, pomak kojeg je načinio Durkheim je bio ogroman, tako da se više od pola veka u sociologiji nije pojavio bolji postupak za ispitivanje uzročnosti na kategorijalnim podacima. Lazarsfeld je u sociologiju uveo pojam multivarijacione analize. Osim toga, on je bio jedan od kreatora i nosilaca talasa empirizma. Za Lazarsfeld-ovog radnog veka se desila smena nekoliko generacija računara, usavršen je niz metoda na kojima je radio veliki broj naučnika, sociološka istraživanja su postala još brojnija i još rasprostranjenija. Iza njega je ostao ogroman doprinos kojeg je dao razvoju analize latentnih struktura i istraživačkoj metodologiji uopšte. Nakon što su Durkheim i Lazarsfeld 'pripremili' sociologiju epistemološki i metodološki za ulaz multivarijacione analize stvorili su se uslovi za interaktivni razvoj metoda za analizu višesmernih tabela kontingencije.*

*Ključne reči: uzročna analiza, empirijska istraživanja, varijable, determinizam.*

### *Samoubistva u uzročnom spletu - E. Durkheim*

Paul Lazarsfeld je uneo koncept multivarijacione analize u sociološku metodologiju. Ipak, niz bitnih problema je bio razrešen još u Durkheim-ovo vreme. S obzirom na to da su i Quetelet i Durkheim radili analizu na popisnim podacima, očigledno je da je do Durkheim-ovog doba sazrela svest o društvenom totalitetu. Radi se, naime, o tome da Durkheim-a nije mučilo pitanje može li proučavati samoubistva kao celovitu društvenu pojavu, bez obzira na to što ona iskazuje varijacije po različitim karakteristikama kao što su starost, pol, religija i sl. ('Keverbergova dilema'<sup>1</sup>). Bilo je jasno da su podaci na kojima je radeno istraživa-

---

<sup>1</sup> Ovaj problem homogenosti se naziva 'Keverbergova dilema' po izvesnom baronu Keverberg-u koji je komentarisao Quetelet-ovu analizu popisnih podataka i konstatovao da se društvene pojave različito raspoređuju po geografskim područjima, starosnim kategorijama i sl. On

nje prikupljeni za celu populaciju, a budući da Durkheim ovom pitanju nije prilazio sa stanovišta teorije verovatnoće, nije ni morao lično da prevazilazi 'Keverbergovu dilemu' da bi došao do adekvatnog rešenja. U Stigler-ovoj 'Istoriji statistike' možemo naći svedočanstvo da je baš negde u vreme kada su u Francuskoj prikupljeni podaci na osnovu kojih je Durkheim sačinio svoju studiju o samoubistvu ova dilema u SAD bila razrešena i na opštem saznavnom planu<sup>2</sup>, kao i da je važnu ulogu u tome odigrao tehnološki napredak u obradi popisnih podataka: '... Kada je postojalo mnoštvo podataka (na primer popis), bilo je, u principu, moguće posmatrati sve vrste kategorija; ali u praksi to nije bilo moguće. U vezi popisa u SAD Herman Hollerith je napisao: 'Do popisa 1890., mi nikada nismo znali čak ni proporcije samaca, onih koji su u braku ili onih koji su udovci u našem stanovništvu... Sa metodima iz 1880. bilo je praktično nemoguće podeliti one koji su rođeni u zemlji na one koji imaju domaće poreklo i one koji imaju strano poreklo. O klasifikovanju stanovništva prema starosti, polu i rodnom mestu majke se nije ni razmišljalo' (Hollerith, 1894, p.678). 'Ovaj direktan napad na de Keverbergovu dilemu je zahtevao modernu opremu za tabeliranje, čak i kada su podaci bili obimni i pri ruci.' (Stigler, 1986: 360-361). A Hollerith-ov tekst koji citira Stigler je objavljen u *Journal of the Royal Statistical Society* i ima naslov 'Elektronska mašina za tabeliranje'!

Durkheim je u statistici video jedno od sredstava za izdvajanje društvenih pojava iz pojedninačnih oblika njihovog ispoljavanja (Milić, 1978: 111). Dakle, bez obzira na to što je svako samoubistvo individualan čin, stabilna stopa samoubistava tokom različitih godina govori o tome da se kroz nju ispoljavaju neka kolektivna stanja. Statistika je pribavljala onakve podatke kakve je Durkheim najviše cenio u prikupljanju iskustvene građe, jer je pojave definisala na osnovu opažajnih osobina, obuhvatala ih u celosti i iskazivala preko zbirnih pokazatelja.

Nas više zanima Durkheim-ova analiza popisnih podataka bazirana na njegovom shvatanju da je otkrivanje uzročnih odnosa osnova objašnjenja društvenih pojava. Durkheim je verovao da konstantni odnosi stabilnosti i varijacije između dve pojave ukazuju na postojanje uzročne veze. Zbog toga stabilne korelacije predstavljaju polaznu opisnu osnovu u analizi. Ali, Durkheim je bio svestan da se uzročni odnosi ne mogu konstatovati samo na osnovu stabilnih korelacija, jer postoji mogućnost javljanja prividnih korelacija, tako da je moguće da su dve pojave koje se čine povezanim u stvari proizvod nekog zajedničkog uzroka. Uzmimo primer utvrđene veze između stepena obrazovanja i stope samoubistva. Podaci su pokazali da je viši stepen obrazovanja praćen povećanom stopom samoubistva. No, Durkheim-ovim teorijskim shvatanjima nije odgovaralo da u visokom

---

nije mogao da shvati kako se pouzdana informacija o proučavanom problemu može dobiti na osnovu tako 'rasutih' podataka. Ova dilema danas zvuči trivijalno, ali ona odlično opisuje predstavu o totalitetu proučavanih pojava koja je postojala u to vreme. Iz istog razloga 'Keverbergova dilema' može da doprinese objašnjenju tempa razvoja naučne misli o društvu.

<sup>2</sup> Nema razloga za sumnju da je Francuska u tom pogledu znatno zaostajala iza SAD.

obrazovanju prepozna uzrok sklonosti ka samoubistvu, jer obrazovanje zahvata pretežno periferne delove ličnosti i zbog toga ne može da utiče na tako odsudnu životnu odluku kakvu podrazumeva samoubistvo. Stoga je on uzrok obema pojavama prepoznao u opadanju tradicionalnih shvatanja. Gubitak vere u tradicionalna shvatanja podstiče čoveka da kroz veće obrazovanje potraži odgovore na pitanja koja mu nameće život. Ali, istovremeno, '... gubitak vere u tradicionalna shvatanja u društvu, koje nije na novi način uspostavilo svoju moralnu integraciju, izaziva labavljenje unutardruštvenih odnosa koji su, po Durkemu, temelj na kojem pojedinac gradi svoju društvenu i psihičku ravnotežu...' (Milić, 1978: 128). Ovde nije bitno da li je Durkheim-ovo teorijsko obrazloženje tačno. Bitno je da se on ne povodi za opisom pojava, nego da traži širi smisao uočenih veza i njihovo teorijsko obrazloženje. Razvrstavanje podataka po različitim tabelama je samo instrument koji se aktivira teorijskom instrukcijom i čiji sadržaj dobija smisao kroz teorijsko objašnjenje.

U primeni ovakvog metodološkog postupka pri izučavanju ovog problema (samoubistva) se može uočiti osobenost sociološkog pristupa i naslutiti sav Durkheim-ov sociološki i istraživački talenat. Samo na taj način je Durkheim u jednom duboko individualnom činu mogao da prepozna elemente društvenog, i ne samo da ih prepozna, nego i da odredi u kojoj meri određeni društveni činioци utiču na sklonost ka samoubistvu. Durkheim kaže: '...Nijedan opis pojedinačnih slučajeva, ma koliko dobar bio, nam nikad neće reći koja su (samoubistva - prim. S. C.) imala sociološki karakter. Ako neko želi da zna različite efekte samoubistva kao kolektivne pojave, mora ga obuhvatiti od početka u njegovoj kolektivnoj formi, tj. kroz statističke podatke. Društvena stopa mora biti uzeta od početka kao predmet analize; napredak mora biti od celine ka delovima. Jasno, ono može biti analizirano jedino u odnosu na svoje različite uzroke, jer u njima su jedinice koje ga sačinjavaju **homogene** (moj bold), bez kvalitativne razlike. Stoga moramo odmah otkriti njegove uzroke a kasnije razmatrati njihove posledice na pojedince.' (Durkheim, 1951: 148)

Pogledajmo sada konkretan primer analize 'egoističnih samoubistava', još preciznije, onaj deo gde Durkheim proučava uticaj bračnog stanja i starosti na stopu samoubistava, a potom uvodi u analizu i pol. On ukazuje da apsolutni brojevi ostavljaju utisak da osobe koje nisu u braku čine manje samoubistava nego one koje su u braku. Ova činjenica je bila uzrokom brojnih pogrešnih interpretacija, a već površna analiza je pokazivala da je obeležje koje se krilo iza ove konstatacije starost. Jednostavno, među osobama koje nisu u braku se nalazi najveći broj onih koji su mlađi od 16 godina, a deca su daleko manje sklona samoubistvima nego sredovečni ili stariji. Situacija je, zapravo, obrnuta. Stopa samoubistva je manja među osobama koje su u braku (Durkheim, 1951: 171-172). Uvlačeći starosne kategorije u analizu, Durkheim je već pokazao i mogućnost simulacije eksperimenta u neeksperimentalnoj situaciji. On ukazuje da će poređenje oženjenih i neoženjenih u okviru svake starosne kategorije (25-30, 30-35, itd.) omogućiti izolaciju

faktora bračnog stanja od drugih obeležja i iskazivanje pune varijacije ovog faktora.

Ovakav postupak je ispravljao greške koje je donosilo zaključivanje preko proseka. Naime, modus osoba koje nisu u braku se nalazio u starosnoj kategoriji između 15 i 20 godina (oko 55%), dok je najveći broj osoba koje su u braku skoncentrisan u starosnoj kategoriji '40-45 godina'. S druge strane, podaci o samoubistvima su pokazivali da je u starosnoj kategoriji '21-30 godina' stopa samoubistava 97.9 na milion stanovnika, dok je u kategoriji '41-50' ta stopa 164.4 na milion stanovnika. Kombinujući ovakve informacije mogao bi se dobiti utisak da oni koji nisu u braku ne mogu imati stopu samoubistava veću od 97.9/mil., a da se ona za osobe u braku kreće između 114.5/mil. (vrednost za starosnu kategoriju 31-40) i 164.4/mil.

Zaključivanje na osnovu proseka je više ličilo na Quetelet-ovo konstruisanje 'prosečnog čoveka' kao pseudorekonstrukciju homogenosti podataka, dok je Durkheim svojim pristupom tragao za strukturom pojave. U tabeli 2. se može videti kako je Durkheim izložio podatke u svojoj knjizi, ali bez koeficijenata očuvanosti (preservation) koje je on koristio kao indeksni pokazatelj odnosa između varijabli.

*Tabela 2. Francuska (1889-1891): Samoubistva počinjena na 1,000,000 stanovnika za svako doba i grupu bračnog statusa, prosečna godina*

Starost	Muškarci			Žene		
	Neoženjeni	Oženjeni	Udovci	Neudate	Udate	Udovice
15-20	113	500	*	79.4	33	333
20-25	237	97	142	106	53	66
25-30	394	122	412	151	68	178
30-40	627	226	560	126	82	205
40-50	975	340	721	171	106	168
50-60	1434	520	979	204	151	199
60-70	1768	635	1166	189	158	257
70-80	1983	704	1288	206	209	248
80 <sup>+</sup>	1571	770	1154	176	110	240

Izvor: Durkheim (1951): 178.

Iz table se može zaključiti da je učešće samoubistava osoba u braku manje nego isto učešće osoba koje nisu u braku i to skoro u svakoj starosnoj kategoriji. Na osnovu ovog podatka Durkheim je zaključio da se za '...brak može reći da smanjuje opasnost od samoubistva...' (Durkheim, 1951: 173). Durkheim je, naravno, doneo niz zaključaka o starosti, polu i bračnom stanju kao uzrocima samoubistava, a oni su iskazani u nekoliko zakona (v. Durkheim, 1951: 178-180).

Kod Durkheim-a imamo veoma razrađen koncept analize tabela kontingencije sa više 'ulaza'. To je elementarni postupak uzročne analize neeksperimentalnih podataka i začetak multivarijacione analize. Ovakvim postupkom istraživači '...prevazilaze multivarijacioni problem 'u glavi!...' (Fajgelj, 1995: 113). Rekli smo da je u Durkheim-ovo vreme, za razliku od Quetelet-ovog, formirana svest o celini društva. Durkheim je iskazao jasnu potrebu za istraživanjem uzajamne povezanosti različitih osobina koje taj totalitet nosi u sebi. Takode, on je bio svestan da među brojnim vezama koje je moguće uočiti mnoge nisu od značaja, dok je određen broj onih koje se učine značajnim - samo prividno uzročne. Postupak uvođenja dodatnih varijabli u analizu je imao za cilj da razreši ovaj problem. Dalje, uzročnost kojom se Durkheim bavio nije bila determinističkog tipa - pre bi se moglo reći da je Durkheim o uzrocima razmišljao kroz kategoriju verovatnih dovoljnih uslova. On, međutim, nije tragao za većom preciznošću koju su uskoro doneli probablistički pristup i savršenije metode multivarijacione analize. Durkheim je, dakle, ovladao konceptom multidimenzionalnosti u sociološkom smislu, ali ne i u statističkom smislu. Ipak, i ovakva preciznost je bila dovoljna za ono vreme; uostalom, decenijama istraživači nisu uspevali da prevaziđu Durkheim-ovu iscrpnost u ovom pogledu.

Šta je nedostajalo Durkheim-ovom pristupu uzročnoj analizi neeksperimentalnih podataka da bi bio savršeniji? Ne samo višestruka regresiona ili korelaciona analiza, nego i niz postupaka za analizu tabela kontingencije koje su tih godina razvijali Pearson i Yule. Dalje, Durkheim nije uočio da prividnost veza može imati i obrnuti princip. Mnoge veze koje se učine beznačajnim mogu kriti uzročnost u dubljim determinističkim slojevima. Selvin navodi primer u kojem Durkheim kreće od uočene visoke stope samoubistava u Nemačkoj i testira tezu o uticaju karakteristika nemačke civilizacije na ovu pojavu. Da bi proverio ovu tvrdnju, on je proučavao kretanje stope samoubistava kod Nemaca koji žive izvan Nemačke, u Austro-Ugarskoj, u čijim provincijama oni čine između 1.9% i 100 % stanovništva. Durkheim je otkrio da nema ni najmanjeg traga uticaja nemačke kulture na stopu samoubistava, jer se u provincijama sa apsolutnom nemačkom većinom javlja niža stopa samoubistava nego u provincijama gde Nemci čine relativnu manjinu (Durkheim, 1951: 86-87). Nakon takvog zaključka Durkheim je napustio dalju analizu.

Selvin, međutim, pokazuje upotrebom Spearman-ovog koeficijenta korelacije ranga da se može uočiti čvrsta statistička veza između stope samoubistava i proporcije nemačkog stanovništva ( $\rho=0.57$ ) (Selvin, 1976: 41), pogotovu ako se iz analize isključe zapadne austrijske provincije u kojima ima mnogo Nemaca i neproporcionalno malo samoubistava ( $\rho=0.97$ ) (Selvin, 1965: 119). S obzirom na istorijske aspekte naše teme nije bez značaja ni Selvin-ov komentar da je pišući ovaj tekst o Durkheim-u otkrio da je problem sličan ovome Yule prepoznao i razrešio u Booth-ovim podacima o siromaštvu (Selvin, 1965: 119-120). Još jedan nedostatak je i to što se Durkheim-ova analiza morala kretati kroz onaj set varijabli koje su dostupne kroz podatke popisa, dok je kasnije sa primenom uzoraka u pri-

kupljanju podataka povećavana mogućnost operacionalne raznovrsnosti. Durkheim je i sam naglašavao ovaj problem, žaleći se da se u mnogim zemljama ne beleži dovoljno podataka o samoubicama.

Sve u svemu, pomak kojeg je načinio Durkheim je bio ogroman, tako da se više od pola veka u sociologiji nije pojavio bolji posutpak za ispitivanje uzročnosti na kategorijalnim podacima. Postojala su, međutim, dalja usavršavanja tehnike i logike uzročne analize neeksperimentalnih podataka u sociološkim istraživanjima koja su postupno pripremala teren za ulazak metoda multivarijacione analize u sociologiju i za njihovu transformaciju.

*Standardizacija uzročne analize u anketnim istraživanjima, početak multivarijacione analize u sociologiji - P.F. Lazarsfeld*

O P. Lazarsfeld-u i njegovim doprinosima razvoju sociološke metodologije je dosta pisano i u stranim i u domaćim sociološkim publikacijama. Tema ovog teksta ne zahteva obuhvat celokupnog Lazarsfeldovog radnog opusa. Za nas su bitne njegove aktivnosti kojima su obeležena dva bitna elementa u razvoju upotrebe multivarijacione analize u sociološkim istraživanjima, jedan element koji predstavlja opšti okvir naše teme i drugi koji odslikava njegov konkretan doprinos metodološkom napretku ovog pristupa.

Što se tiče prvog elementa, početak najživljih Lazarsfeld-ovih istraživačkih aktivnosti pada u vreme kada se na planu praktičnih rezultata konvergencije statističke i sociološke metodologije odigrava verovatno najbitinija sinteza. Radi se o tome da je 30-ih godina u Sjedinjenim Američkim Državama došlo do koncentracije niza bitnih činilaca koji su u velikoj meri oblikovali onaj način iskustvenih socioloških istraživanja koji je poznat pod imenom anketna istraživanja. Još 1839. godine je osnovano Američko statističko društvo i krenulo se sa redovnim statističkim popisima koji se, uz one koji su sprovedeni u Švedskoj, smatraju prvim modernim popisima (Stigler, 1986; Duncan, 1984). Ova vrsta podataka je predstavljala izvor informacija za brojna sociografska istraživanja, dok se u drugoj fazi društvena istraživanja okreću podacima prikupljenim na terenu za specifične potrebe same nauke (Bogdanović, 1981: 38). Do tridesetih godina društvena istraživanja su stekla već stabilnu poziciju i dugogodišnju praksu. Njihova egzistencija je izvirala iz realnih društvenih potreba, a počivala je na reformatorskom društvenom pokretu. Štaviše, do tog trenutka ovakva slika se pretočila u čvrstu institucionalizovanu formu<sup>3</sup>.

Univerzitetska promocija sociologije je tih godina već bila okončana, a kvantitativna orijentacija je u tome imala posebnu ulogu. Uvođenje intezivnih kurseva statistike na sociološka odeljenja američkih univerziteta bilo je pitanje pres-

---

<sup>3</sup>Proces institucionalizacije američke sociologije je podrobno analiziran u (Oberschall, ed., 1972).

tiža. Statistika je u takvom trenutku bila pogodna za jačanje pozicije sociologije u sistemu nauka na dva načina. S jedne strane, ona joj je davala dovoljno formalne 'čvrstine', zadovoljavala je prohteve za egzaktnošću koji su stajali pred svim naukama. S druge strane, ona je svojom neutralnošću bila podobna za prilagodavanje osobenostima predmeta istraživanja u sociologiji<sup>4</sup>. Jasno je da takvo prilagodavanje ne može biti definitivno i da je ono proces kojem smo i mi svedoci. Takođe je jasno da se često dešavalo da se predmet istraživanja 'prilagodavao' statističkim normama, a ne obrnuto. Ali, kroz nabrojane naučne i institucionalne okolnosti statistika se čvrsto usadila u sam koncept empirijskih istraživanja u američkoj sociologiji. U vreme kada Lazarsfeld razvija svoje naučne aktivnosti britanska statistička dostignuća su se rasprostrla po različitim naučnim poljima (podsetimo se da je to vreme kada i nova statistička nauka mukotrno probija svoj put u sistem nauka). Jedno od njih koje je posebno bitno za situaciju o kojoj govorimo je teorija uzoraka koju je njen začetnik J. Neyman iz Britanije doneo u povelju, a potpuno je razvio i iskustveno afirmisao u SAD. Ovaj momenat je bitan zato što su uzorci kroz koncept reprezentativnosti omogućavali brža i produbljenija sociološka istraživanja širih populacija, na većem broju simultano merenih osobina. Sve navedene okolnosti su bile na okupu u vreme kada je Lazarsfeld zajedno sa Merton-om dosegao najviše domete u razvoju empirizma (Bogdanović, 1981: 42).

Ovaj opšti momenat je bitan za temu ovog rada utoliko što od tog trenutka analiza višestruke povezanosti društvenih obeležja biva usmerena ka drugačijem referentnom okviru i na drugačiji način. Iako je i ranije bilo istraživanja društvenih problema koja su počivala na parcijalnim popisima, u trenutku o kojem govorimo anketna istraživanja sprovedena na uzorcima postaju najrasprostranjeniji oblik empirijskih istraživanja u sociologiji. To je omogućilo fokusiranje istraživanja na one segmente nosilaca istraživanih pojava kod kojih su te pojave imale najbolju ispoljenost. Na taj način je planom uzorka bivalo omogućeno očuvanje homogenosti podataka uz jasnije ispoljavanje istraživanih osobina. Ovakav način smanjenja elementa greške u stohastičkom modelu ne utiče na pouzdanost zaključaka o ustanovljenim vezama između pojava (može, eventualno, da utiče na intezitet njihovog ispoljavanja), a doprinosi lakšem otkrivanju uzročnih odnosa. S druge strane, kroz anketu istraživač u analizu uvlači one varijable koje on smatra bitnim, osim ako se radi o sekundarnoj analizi podataka prikupljenih u drugu svrhu. Na taj način je omogućena kako problemska raznovrsnost, tako i dubina uzročnih odnosa kakve podaci prikupljeni opštim popisom ne mogu da obezbede. Shodno navedenim činjenicama, multivarijaciona analiza je mogla da načini dublje i samim tim, značajnije prodore u determinističku strukturu istraživanih pojava, a da značajnost njenih zaključaka ostane u prihvatljivim granicama. Već u Lazarsfeld-ovo vreme ovim okolnostima u istraživanju su se pridružili i savršeniji metodi multivarija-

---

<sup>4</sup>Ovu dvostruku funkciju statistike, funkciju konformizma i funkciju diferencijacije, u ustanovljenju novih društvenih nauka: sociologije, antropologije i ekonomije, Camic i Xie nazivaju 'dilema novopridošlih' (Camic, and Xie, 1994: 776).

cione analize, tako da je nakon pedesetih godina celokupna istraživačka situacija imala znatno drugačiji izgled od one u Durkheim-ovo vreme. Osnovna logika multivarijacije kod analize tabela višestrukih ukrštanja jeste ostala ista, ali su implikacije analize bile i šire i produbljenije. No, to ne znači da su one nužno bile i dalekosežnije, jer je taj momenat zavisio od teorijskog pristupa i kvaliteta operacionalizacije. Brzina koju je omogućavao novi koncept istraživanja je često dovela do površnosti i nekoherentnosti zaključaka izvedenih na osnovu savršenijih metoda multivarijacione analize. Na taj način su mnogi empiristi koji su se koristili razrađenim Durkheim-ovim modelom uzročne analize neeksperimentalnih podataka, krijući se ispod plašta velikog sociologa u sociološku nauku uneli pored transformisanog koncepta multivarijacione analize i breme jednostranosti, površnosti i formalizma.

Drugi element razvoja upotrebe multivarijacione analize u sociološkim istraživanjima koji nosi Lazarsfeld-ov beleg jeste njegov lični doprinos na ovom polju. Ovaj doprinos nije bio samo plod navedenih okolnosti. U dugom periodu Lazarsfeld-ovih aktivnosti u sociološkoj metodologiji i istraživanjima (preko četrdeset godina) došlo je do niza promena na polju multivarijacione analize kojima je i on doprineo. Već smo pominjali analizu latentnih struktura i za nju vezano usavršavanje faktorske analize nominalnih varijabli. Takođe, u Lazarsfeld-ovo vreme su u empirijska sociološka istraživanja najpre ušli 'veliki' (mainframe) računari nakon Drugog svetskog rata, da bi se pred kraj njegove karijere pojavila i prva generacija personalnih računara sa kojima je upotreba metoda multivarijacione analize postala ekstenzivna. Da ovo nije bez značaja potvrđuje i činjenica da je Lazarsfeld za operacionalni jezik (jezik varijabli) koristio i naziv IBM jezik (po IBM računaru), a da je računarsku bušenu karticu koristio kao predstavu međupovezanosti (vid. Lazarsfeld and Rozenberg, eds., 1957: 111). Ali, da bi se održao kontinuitet sa dosadašnjim tokom rada, pažnja će biti poklonjena Lazarsfeld-ovoj primeni, razradi i sistematizaciji onog načina uzročne analize neeksperimentalnih podataka kakva je predstavljena u poglavlju koje se bavi Durkheim-ovim istraživanjem samoubistava.

Lazarsfeld je svoj koncept pregledno izložio zajedno sa Patriciom Kendall u tekstu koji govori o metodološkim aspektima obimnog proučavanja organizacije, grupne strukture i morala američke vojske u Drugom svetskom ratu (Kendall and Lazarsfeld, 1950). Ovaj njihov tekst je bio osnova za niz priručnika o anketnim i drugim preglednim istraživanjima koji su se pojavili pedesetih godina.

Kendall i Lazarsfeld odmah iznose konstataciju da su podaci na kojima je obavljena ova naknadna analiza prikupljeni u druge svrhe, tako da prilozi u studiji čiji je on suurednik (uz Merton--a) predstavljaju sekundarnu analizu. To znači da varijable sa kojima se operiše u ovakvoj analizi (pitanja iz upitnika) nisu smišljene tako da bi proveravale unapred postavljene hipoteze, pa se iz njih uglavnom mogu crpsti samo smernice za konstrukcije određenih teorijskih objašnjenja. No, treba napomenuti da zbog složene prirode društvenog determinizma čak i čvršće usme-



ren istraživački pristup podrazumeva traganja za multivarijacionim relacijama koje nisu pretpostavljene, podrazumeva stalnu eksploraciju i povremena *ad hoc* rešenja.

Ovakve okolnosti bitno razlikuju anketno istraživanje od kontrolisanog eksperimenta, ali je ono, ipak, najbolja osnova za simulaciju eksperimentalnih istraživanja u sociologiji. To je upravo omogućeno multivarijacionim pristupom u analizi, pri situaciji gde se relacije između varijabli sa pretpostavljenim uzročno-posledičnim dejstvom posmatraju pri konstantnim vrednostima (kategorijama) ostalih varijabli u izdvojenom determinističkom spletu. Verovatno je to razlog zbog kojeg su se u metodološkim sistematizacijama ove vrste javljali komentari na temu sličnosti i odstupanja ovakve analize od one koja se obavlja u uslovima kontrolisanog eksperimenta (Kendall and Lazarsfeld, 1950: 136-147; Hyman, 1957: 243-247).

Kendall i Lazarsfeld naglašavaju da je razjašnjenje veza između dve ili više varijabli svakako važniji zadatak nego opis tih varijabli. Ovaj proces se odvija u dve etape: u prvoj nastojimo da odredimo koliko je osnovano zaključivati o uzročno-posledičnim odnosima između varijabli, a u drugoj je cilj da se prouči sam uzročno-posledični proces. Naglasimo još jednom da se kao osnovni instrument za analizu koriste tabele kontingencije.

U tekstu se kreće od primera u kojem je na opisnom nivou konstatovana veza između lokacije službovanja vojnika i njihove prilagodivosti vojničkom životu. Podaci su pokazivali da vojnici koji su stacionirani peko okeana pokazuju slabije raspoloženje od onih koji su stacionirani u SAD. No, upravo zbog toga što se ovdje ne radi o kontrolisanom eksperimentu, po čijoj logici bi slučajno odabrane grupe vojnika sa odgovarajućim karakteristikama bile poslone na različite lokacije da bi se proverio uticaj tog faktora na stanje njihovog duha, ne sme se doneti zaključak o uzročnoj povezanosti ovih varijabli dok se ne proverí uticaj drugih relevantnih obeležja na njihov međusobni odnos. Testiranje prividnih veza se obavlja tako što se tabela kontingencije usložnjava uvođenjem novih varijabli. U ovom primeru se moglo desiti da dužina službovanja utiče na prividnost veze između lokacije i zadovoljstva, jer vojnici stacionirani izvan SAD službuju duže. Stoga je uzorak ispitanih vojnika podeljen na podgrupe na osnovu različitih dužina vremena službovanja u vojsci da bi se u okviru tih kategorija analizirala veza između lokacije i morala. Ako je osnovna veza bila ispravna, ona je morala da se ispolji. Prividnost se, međutim, ne mora otkriti već na prvom koraku, tako da je nekad potrebno osloniti se na veći broj kontrolnih varijabli. Tu se javlja i problem veličine uzorka. Naime, svaka nova varijabla u tabeli višestrukih ukrštanja značajno uvećava broj polja u tabeli i verovatnoću da će određeni broj tih polja ostati prazan, što slabi efekte analize. Već sa šest dihotomnih varijabli (npr. 1 - služi izvan SAD, 0 - služi u SAD) bismo dobili tabelu od  $(2 \times 2 \times 2)^2 = 64$  polja. Uvećanje broja varijabli bi se desilo već sa povećanjem broja kategorija u postojećim varijablama. Analitičari podataka o američkim vojnicima su načinili 138 malih homogenih podgrupa na osnovu nekoliko kontrolnih varijabli i ustanovili da u 113 od njih vojnici izvan SAD pokazuju slabiji moral nego oni u SAD, čime je oborena pret-

postavka o prividnosti osnovne veze, barem u odabranom segmentu determinističkog spleta.

Drugi problem koji se može pojaviti pri ovakvoj analizi jeste određivanje vremenskog sleda varijabli. U kontrolisanom eksperimentu se testira unapred definisan uzročni odnos, ali u anketnim istraživanjima to često ne može biti slučaj. Istraživač u određenom broju situacija može znati koja varijabla prethodi kojoj, ali se može desiti da za ustanovljenu vezu treba odrediti smer (da li vrednosne orijentacije kod maloletnika utiču na izbor vršnjačke grupe, ili pripadnost grupi utiče na vrednosne orijentacije?). Kendall i Lazarsfeld daju primer odnosa između bračnog stanja i čina kod kojeg su ustanovili da je verovatnije da oženjeni imaju viši čin, čak i kada su faktori starosti i staža u službi kontrolisani. Ali, da li bračno stanje utiče na unapredenje ili unapredenje podstiče na odluku o ženidbi? Odgovor su ponudili podaci o tome da li je ispitanik bio u braku pre stupanja u službu ili se to desilo nakon pristupanja vojsci. Kod onih koji su prvo ušli u vojsku, a potom u brak je primećena jača veza između bračnog stanja i ranga, tako da se pre može suditi da je brak rezultat unapredenja, nego njegova predispozicija. (Moja pretpostavka u vezi ovog primera je da uočeni redosled nije dovoljno čvrsto potvrđen na ovaj način i da on jednostavno zavisi od trećeg faktora koji je zajednički uzrok obe pojave, a to je posvećenost pozivu). Kao moguća rešenja za problem utvrđivanja vremenskog sleda među pojavama Kendall i Lazarsfeld navode panel istraživanja i retrospektivna pitanja.

U knjizi V. Milića su navedene etape u anlizi neeksperimentalnih podataka pedantno izložene (Milić, 1978: 717-723). On razlikuje: 1) *opis* u kojem se dobija polazna predstava o mogućim vezama između varijabli koja, opet, daje odgovarajući značaj tim varijablama u daljoj analizi i 2) *specifikaciju* u kojoj se početni opis rasčlanjava da bi se utvrdila uzročna struktura veze između obeležja. U okviru specifikacije Milić razlikuje dve vrste problema i izlaže ih propraćene formalnim grafičkim prikazima (Milić, 1978: 723-741) koji su ubrzo nakon Merton-ove i Lazarsfeld-ove studije postale standard u modeliranju uzročne analize<sup>5</sup>. Navešćemo samo nekoliko dijagrama da bi smo ilustrovali način razmišljanja koji će se kasnije ponoviti u odeljku o loglinearnim modelima.

#### 2a) Prividne korelacije i neutvrđeni vremenski redosled

- podudaranje nezavisnih naporednih nizova

$$\begin{array}{ccc}
 t_1 & y & z \\
 & \downarrow & \downarrow \\
 t_2 & x \text{-----} y & 
 \end{array} \quad (1)$$

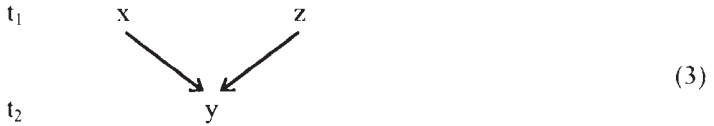
<sup>5</sup> Mora se naglasiti da su 'modu' grafičkog predstavljanja doneli oni metodolozi koji su problem uzročne analize rešavali koeficijentima višestruke korelacije i višestruke regresije. Primer su H.Blalock i H.Simon. Blalock se u tom smislu i poziva na Simon-a (vidi Blalock, 1964; Simon, 1957).

- zajednički uzrok



2b. Višestruki deterministički spletovi

- više nezavisnih uzroka



- uključenost uzročne veze u duži uzročni niz



- povratna veza



U nastavku Kendall i Lazarsfeld daju shemu elaboracije ustanovljenih veza između tri varijable. Oni vezu između dve osnovne varijable definišu kao uslovni

zbir njihove direktne veze na različitim nivoima kontrolne varijable i direktne veze svake od njih sa kontrolnom varijablom. Za dve od njih su dali i shematski prikaz u stilu gorenavedenih. Detaljno objašnjenje ove elaboracije se može naći i kod Supeka (Supek, 1981: 406-422) i kod Đurića (Đurić, 1962: 233-236). U ovom radu će biti prikazani samo formalni iskazi elaboracije (Kendall and Lazarsfeld, 1950: 153-155) i to iz tri razloga koji će biti izloženi odmah nakon prikaza.

- osnovna jednačina

$$[xy] = [xy:t^+] \oplus [xy:t] \oplus [xt] [ty]$$

- tumačenje simbola

[xy]	direktna veza između dve osnovne varijable
[xy:t <sup>+</sup> ] i [xy:t]	dve parcijalne veze pri različitim kategorijama kontrolne varijable
⊕	znak za uslovni zbir
[xt]	veza između petpostavljenog uzorka i kontrolne varijable
[ty]	veza između pretpostavljene posledice i kontrolne varijable

- izvedene jednačine

$$[xy] = [xy:t^+] \oplus [xy:t] + (0) [ty] \quad \text{predstavljena dijagramom (3)}$$

$$[xy] = [xy:t^+] \oplus [xy:t] + [xt] (0) \quad \text{predstavljena dijagramom (4)}$$

$$[xy] = (0) + (0) + [xt] [ty] \quad \text{predstavljena dijagramom (2)}$$

Prvi razlog zbog kojeg su izložene ove jednačine je da bi se pokazalo da su one kao formalni pokazatelj povezanosti varijabli lošije rešenje nego ranije prikazani dijagrami. Druga dva razloga zbog kojih sam smatrao bitnim njihovo navođenje su obrazložena u fusnotama teksta Kendall-ove i Lazarsfeld-a. U prvoj oni kažu da: '...Puna jednačina za slučaj tri atributa se može naći u Yule i Kendall, *Uvod u teoriju statistike...*' (Kendall and Lazarsfeld, 1950: 153), a u drugoj da je '...Sličnost (osnovne - prim. S.C.) jednačine (1) standardnoj formuli parcijalne korelacije (je) očigledna...' (Kendall and Lazarsfeld, 1950: 154). Pozivanje na Yule-a pokazuje da aktivnosti statističara koji su najaktivnije radili na razvoju početaka multivarijacione analize dobijaju odjeka u sociološkoj metodologiji. Nije, dakle, samo Neyman-ova teorija uzoraka ušla u splet okolnosti koje su obeležile početak interaktivne konvergencije sociološke i statističke metodologije, nego se to desilo i sa multivarijacionom analizom. Drugi citat jasno pokazuje da se u formalizaciji postupka uzročne analize neeksperimentalnih podataka usvaja logika jednog metoda multivarijacione analize, višestruke korelacije, koji na ovakve

podatke nije mogao biti primenjen zbog njihove kategorijalne prirode. Konačno su i kontinuitet razvojne linije statistike i kontinuitet razvojne linije sociološke istraživačke metodologije našli ishodište u radovima istih autora, u pristupu istog istraživačkog pravca, baveći se pojavama koje su tipične za sociološki naučni vidokrug. Ako smo za Quetelet-a ili Yule-a mogli reći da su statističari koji su se bavili i društvenim istraživanjima, od Lazarsfeld-a na dalje možemo sve češće da uočimo naučnike koji rade na oba polja, gradeći novi trend u pristupu sociološkim istraživanjima (vid. Clogg 1992). Lazarsfeld-ov statistički rad na tabelama kontigencije je ostao zabeležen i kao etapa u razvoju loglinearne analize.

Sve okolnosti navedene u uvodu ovog odeljka su se skoncentrisale u empiristički talas i svakako je 'kvantitativni ekstenzionalizam' (Wiley) sedamdesetih posebno pogodovao nekontrolisanom prodoru metoda multivarijacione analize u sociologiju, ali celokupna logika analize koja je odražena kroz standarde koje je začeo Durkheim, a koje su metodološki uobličili Kendall i Lazarsfeld, pokazuje da postoji 'zdrava' osnova za razvoj metoda multivarijacione analize koji mogu na adekvatan način da odgovore prohtevima socioloških istraživanja. Ne radi se samo o tome da je nakon njihovog doprinosa statističko ispitivanje višestruke međusobne povezanosti istraživanih pojava i traganje za uzročnošću postalo rasprostranjeno i podrazumevajuće u anketnim istraživanjima. Jednako je bitno i to da je pred istraživače i metodologe postavljen imperativ daljeg unapređenja ove procedure u cilju postizanja preciznijih 'specifikacija' i pouzdanijih 'elaboracija'. To je bio krupan metodološki pomak nakon kojeg su sledila postupnija methodska usavršavanja.

Na sledećih nekoliko strana će biti napravljeno malo odstupanje u odnosu na primere uzročne analize, ali će se ostati na području multivarijacionih metoda, ili, bolje reći, multivarijacione logike. Ovo odstupanje se pravi sa namerom da se pokaže da je u sociološkim istraživanjima postojao prostor za razvoj i onih metoda multivarijacione analize čiji je cilj razvrstavanje i klasifikacija objekata merenja. U tu svrhu biće poklonjena pažnja prvom koeditoru ove studije Robert-u Merton-u, koji je zajedno sa Alice Kitt priložio tekst o teoriji referentnih grupa (Merton and Kitt, 1950). Pored prethodne, ovaj rad Merton-a i Kitt-ove ima još jednu važnu crtu. Svojim temeljnim teorijsko-metodološkim pristupom Merton i Kitt oživljavaju epistemološki referentni okvir koji sam razvio na prethodnim stranicama. Njihov tekst, za razliku od gorenavedenih primera iz ovog zbornika, pokazuje kontakt između teorije i istraživanja 'na delu'. U takvom svetlu i njihov analitički postupak, u kojem sam ja prepoznao multivarijacione odlike, dobija veći značaj i smisao.

Merton i Kitt počinju konstatacijom da postoji dvosmerna komunikacija između društvene teorije i empirijskog istraživanja, jer sistematski prikupljeni podaci unapređuju teoriju postavljajući pred nju zadatak i omogućujući joj zaključivanje na neplanirane načine, dok teorija usmerava istraživanje i uvećava predviđajuću moć podataka (Merton and Kitt, 1950: 40). Autori smatraju da činjenica da se radi o sekundarnoj analizi ne predstavlja problem, jer se induktivnim tretmanom podataka uz stalan pogled na apstraktnija uopštavanja prvih analitičara

ovog empirijskog materijala mogu '... konsolidovati trenutno rasuti fragmenti teorije' (Merton and Kitt, 1950: 41).

Na posrednom nivou između teorije referentnih grupa i podataka autori tragaju za pojmom 'relativne uskraćenosti'<sup>6</sup>, jer su kreatori prvih izveštaja iz ovog istraživanja najčešće koristili taj pojam u svojim generalizacijama. Pošto je ovaj pojam bio noseći u objašnjenjima, trebalo ga je 'rekonstruisati' na empirijskom nivou, čime bi se dobila iskustvena potpora za 'konsolidaciju' teorije. Budući da operacionalizacija ovog pojma nije unapred ugrađena u merni instrument, Merton i Kitt obavljaju preliminarnu pretragu materijala da bi pronašli varijable koje bi najbolje odgovarale toj svrsi. Nakon ove pretrage oni najpre konstatuju da se u odnosu između odabranih zavisnih i nezavisnih varijabli (pretpostavljeni uzročni odnos) javljaju izvesne neočekivane relacije kao što je primer da oženjeni muškarci iskazuju veću ozlojedenost zbog toga što su mobilisani. Tu se u tumačenje uključuje pojam relativne uskraćenosti, pa je model analize otprilike sledeći: '...oženjeni muškarci (nezavisna varijabla) češće dovode u pitanje opravdanost sopstvene mobilizacije (zavisna varijabla), jer oni ocenjuju situaciju u referentnom okviru (interpretativna varijabla) stvorenom poređenjem sebe sa drugim oženjenim muškarcima koji su još uvek civili, koji su potpuno izbegli služenje u trupi, ili sa neoženjenim ljudima u vojsci kojima mobilizacija donosi neuporedivo manje žrtvovanja...'<sup>7</sup> (Merton and Kitt, 1950: 46). Na ovaj način je na teorijskom planu obavljena procedura koju su sistematizovali Kendall i Lazarsfeld. Ali, trebalo je pronaći empirijske varijable kojima će biti predstavljen pojam relativne uskraćenosti. Upravo to je drugi rezultat preliminarne pretrage podataka. Merton i Kitt zaključuju da se ovaj pojam može iskazati, s jedne strane, kroz to da li vojnici za referentni okvir biraju neku grupu kojoj trenutno pripadaju ili pojedinca koji je u njihovoj grupi, ili grupu/pojedinca sa kojim trenutno ne dele pripadnost, a, s druge strane, da li su njihova poređenja situacije u kojoj se nalaze usmerena ka grupama ili pojedincima koji imaju različit društveni položaj od njihovog (veći ili manji), ili ka onima koji su istog društvenog položaja (Merton and Kitt, 1950: 48). Simultano razvrstavanje ispitanika na osnovu ove dve varijable je na korak od konstruisanja treće varijable čije bi kategorije u prvom mahu bile formirane kombinovanjem kategorija prethodne dve (1. sopstvena grupa + isti društveni položaj; 2. sopstvena grupa + drugačiji društveni položaj; 3. drugačija grupa + isti društveni položaj; 4. drugačija grupa + drugačiji društveni položaj). Autori ne čine ovaj formalni korak koji je danas, primenom jednostavne računarske manipulacije matricama podataka, još jedan oblik 'prevazilaženja multivarijacionog problema u glavi', ali im to ne

---

<sup>6</sup> Autori konstatuju odsustvo eksplicitne definicije 'relativne uskraćenosti' kod pisaca izveštaja o ovom istraživanju, ali smatraju da to nije veliki problem, jer se pojam može rekonstruisati preko nekoliko stavova i osećanja američkih vojnika koji iskazuju stanje osujećenosti i nezadovoljstva u odnosu na referentnu grupu (Merton, and Kitt, 1950: 42-45).

<sup>7</sup> Po svoj prilici su autori očekivali da oženjeni muškarci pokažu manje nezadovoljstva sopstvenom mobilisanošću, možda zato što se od njih očekuje veća odgovornost i požrtvovanje.

uskraćuje mogućnost da konstatuju da '...Proučavanje ove matrice varijabli koja podrazumeva pojam relativne uskraćenosti odjednom usmerava pažnju na nekoliko empirijskih i teorijskih problema...' (Merton and Kitt, 1950: 49). Nećemo se upuštati u empirijske i teorijske implikacije ovakvog analitičkog postupka. Konstatujemo dve stvari:

1. Jedan, u suštini, multivarijacioni postupak je omogućio da se iz empirijskih podataka simultanim tretmanom dve varijable (što je bitno drugačiji postupak od konstrukcije kompozitnih indeksa) rekonstruiše apstraktan pojam kojeg nije moguće direktno meriti.

2. Određenje ovog pojma 'stabilizuje' 'srednji princip' i vodi dalje ka odgovorima na suštinska pitanja teorije referentnih grupa.

Savršeniji metodi multivarijacione analize za razvrstavanje objekata ispitivanja omogućuju još preciznije razvrstavanje na osnovu više varijabli. Na taj način kriterijum razvrstavanja postaje potpuniji i precizniji, a objašnjenja koja su zasnovana na ishodima razvrstavanja imaju veću pouzdanost.

\* \* \*

Za Lazarsfeld-ovog radnog veka se desila smena nekoliko generacija računara, usavršen je niz metoda na kojima je radio veliki broj naučnika, sociološka istraživanja su postala još brojnija i još rasprostranjenija. Pregled istraživačkih radova objavljenih od sedamdesetih na ovamo u dva vodeća američka sociološka časopisa *AJS* i *ASR* pokazuje da se u velikom broju radova pored rezultata istraživanja izlaže i 'modeliranje' multivarijacione metodologije.

#### LITERATURA:

- Agresti, A. (1990): *Categorical Data Analysis*. New York: Wiley.
- Barton, A. H. (1957): The Concept of Property-Space in Social Research, pp. 40-53 in P. F. Lazarsfeld and M. Rozenberg (eds.). *The Language of Social Research*. Glencoe: Free Press.
- Blalock, H. M. Jr. (1964): *Causal Inferences in Nonexperimental Research*. Chapel Hill: The University of North Carolina Press
- (1968a): The Measurement Problem: A Gap Between the Languages of Theory and Research, pp. 5-27 in H.M. Blalock Jr. (ed.). *Methodology in Social Research*. New York: McGraw-Hill.
- (1968b): Theory Building and Causal Inferences, pp. 155-198 in H.M. Blalock Jr. (ed.). *Methodology in Social Research*. New York: McGraw-Hill.
- (1988): The Real and Unrealized Contributions of Quantitative Sociology. *American Sociological Review*, 54/3.
- Bogdanović, M. (1981): *Kvantitativni pristup u sociologiji*. Beograd: Službeni list.
- (1994): *Metodološke studije*. Beograd: Institut za političke studije.
- Bottomore, T. and R. Nisbet [eds.] (1978): *A History of Sociological Analysis*. London: Heinemann.

- Camic, C. and Xie, Y. (1994): "The Statistical Turn in American Social Science: Columbia University, 1890 to 1915", *American Sociological Review*, 59.
- Clogg, C. C. (1992): The Impact of Sociological Methodology on Statistical Methodology. *Statistical Science*, 7/2.
- Durić, M. (1962): *Problemi sociološkog metoda*. Beograd: Savremena škola.
- Duncan, O. D. (1984): *Notes on Social Measurement*. New York: Russel Sage Foundation.
- Durkheim, E. (1951): *Suicide*. New York: The Free Press.  
(1963): *Pravila sociološkog metoda*. Beograd: Savremena škola.
- Esthopy, G. (1974): *History of Social Research Methods*. London: Longman.
- Fajgelj, S. (1995): Evaluacija statističkih obrada u ispitivanjima javnog mnjenja. *Sociološki pregled*, XXIX/1.
- Festinger, L., S. Schachter and K. Back (1957): Matrix Analysis of Group Structures. pp. 358-367 in P. F. Lazarsfeld and M. Rozenberg (eds.), *The Language of Social Research*, Glencoe: Free Press Heinemann.
- Horwitz, H. and E. Smith (1957): The Interchangeability of Socio-Economic Indices. pp. 73-77 in P. F. Lazarsfeld and M. Rozenberg (eds.), *The Language of Social Research*, Glencoe: Free Press.
- Hyman, H. (1957): *Survey Design and Analysis*. Glencoe: Free Press.  
*International Encyclopedia of the Social Sciences* (1968), New York: Macmillan and Free Press.
- Jobson, J.D. (1992): *Applied Multivariate Data Analysis: Volume II*. New York: Springer-Verlag.
- Kaplan, A. (1957): "Definition and Specification of Meaning", pp. 527-532 in P. F. Lazarsfeld and M. Rozenberg, eds., *The Language of Social Research*, Glencoe: Free Press.
- Kendall, M.G. and W.R. Buckland (1965): *A Dictionary of Statistical Terms*, Edinburgh: Oliver & Boyd.
- Kendall, P.L. and P.F. Lazarsfeld (1950): "Problems of Survey Analysis". pp. 133-197 in Merton, R. K. and P. F. Lazarsfeld (eds.), *Continuities in Social Research*, Glencoe, Ill: The Free Press.
- Kendall, P. L. (1957): "A Review of Indicators Used in 'The American Soldier'", pp. 27-39 in P. F. Lazarsfeld and M. Rozenberg (eds.), *The Language of Social Research*, Glencoe: Free Press.
- Kovačić, Z. J. (1994): *Multivarijaciona analiza*. Beograd: Ekonomski fakultet.
- Lazarsfeld, P. F. (1957): Interpretation of Statistical Relations as a Research Operation. pp. 115-125 in P. F. Lazarsfeld and M. Rozenberg (eds.), *The Language of Social Research*, Glencoe: Free Press.  
(1961): Notes on the History of Quantification in Sociology - Trends, Sources and Problems. pp. 147-203 in H. Wolf, ed., *Quantification: A History of the Meaning of Measurement in the Natural and Social Sciences*, Indianapolis and New York: Bobbs-Merrill.  
(1970): *Main Trends in Sociology*. London: George Allen & Unwin Ltd.
- Lazarsfeld, P.F. and M. Rosenberg [ed.] (1957): *The Language of Social Research*, Glencoe: The Free Press.
- Lazarsfeld, P.F. and N. W. Henry [eds.] (1966): *Readings in Mathematical Social Science*, Chicago: SRA.



- Merton, R.K. (1957): A Paradigm for the Study of the Sociology of Knowledge. pp. 498-510, u P. F. Lazarsfeld and M. Rozenberg (eds.), *The Language of Social Research*. Glencoe: Free Press.  
(1979): *O teorijskoj sociologiji*. Zagreb: CDD.
- Merton, R.K. and P. F. Lazarsfeld [eds.] (1950): *Continuities in Social Research: Studies in the Scope and Method of 'The American Soldier'*. Glencoe, IL: The Free Press.
- Merton, R.K. and A.S. Kitt (1950): Contributions to the Theory of Reference Group Behavior. pp. 40-106 in Merton, R.K. and P.F. Lazarsfeld, eds., *Continuities in Social Research: Studies in the Scope and Method of 'The American Soldier'*. Glencoe, IL: The Free Press.
- Milić, V. (1978): *Sociološki metod*, Beograd: Nolit.
- Mrkšić, D. (1974): *Axiomata media socioloških teorija*. Magistarski rad. Beograd: Filozofski fakultet, biblioteka Odeljenja za sociologiju.
- Oberschall, A. [ed.] (1972): *The Establishment of Empirical Sociology: Studies in Continuity, Discontinuity and Institutionalization*. New York: Harper and Row.
- Selvin, H.C. (1965): Durkheim's *Suicide*: Further Thoughts on a Methodological Classic. pp. 113-137 in R.A. Nisbet: *Emile Durkheim*. Englewood Cliffs, NJ: Prentice-Hall.  
(1976): "Durkheim, Booth and Yule: the non-diffusion of an intellectual innovation". *Archives europeennes de sociologie*. XVII/1.
- Simon, H.A. (1956): *Models of Man*, New York: John Wiley & Sons.  
(1965): Mathematical Constructions in Social Science. pp. 83-98 in D. Bray brooke (ed.), *The Philosophical Problems of the Social Sciences*. Toronto: Collier-Macmillan Canada Ltd.
- Stigler, S.M. (1986): *The History of Statistics*. Cambridge, MA: Harvard University Press.
- Suchman, E.A. and H. Menzel (1957): "The Interplay of Demographic and Psychological Variables In the Analysis of Voting Surveys". pp. 148-155 in P. F. Lazarsfeld and M. Rozenberg, eds., *The Language of Social Research*. Glencoe: Free Press.
- Supek, R. (1981): *Ispitivanje javnog mnjenja*. Zagreb: SNL.
- Wiley, N. (1985): The Current Interregnum in American Sociology. *Social Research* 1/52.
- Zetterberg, H. (1957): On Axiomatic Theories in Sociology. pp. 533-539 in P. F. Lazarsfeld and M. Rozenberg (eds.), *The Language of Social Research*. Glencoe: Free Press.

*Summary*

Slobodan Cvejčić

### **CAUSAL ANALYSIS OF CATEGORICAL DATA - PREHISTORY AND BEGGING OF MULTIVARIATE ANALYSIS IN SOCIOLOGY**

Durkheim saw statistics as a mean for extraction of social phenomena from single forms of their manifestation. He took over the concept of multidimensionality sociologically, but not statistically. What was that Durkheim missed to make his causal inference on non-experimental data better? Not just multiple regression or correlation, but a set of procedures for contingency table analysis, developed these days by Pearson and Yule.

Moreover, Durkheim couldn't spot that illusive relations could be turned another way. Many of the relations that seemed irrelevant could hide causality in deeper deterministic levels. Yet, Durkheim made enormous improvement and it took more than half of century to construct better method for causal analysis of categorical data.

Lazarsfeld brought notion 'multivariate analysis' into sociology. Besides, he was one of those who created wave of empirism. During Lazarsfeld's age few generations of computers appeared, few methods were developed by great number of scientists, sociological investigations became more numerous and spread. He left big contribution in latent structure analysis and research methodology in general.

*Key words:* causal analysis, empirical investigations, variables, determinism.