

**Danka Purić¹
i Goran Opačić**

Odeljenje za
psihologiju, Filozofski
fakultet, Univerzitet u
Beogradu

PODUZORKOVANJE, SAMOUZORKOVANJE, POSTUPAK „UNIVERZALNOG NOŽA” I NJIHOVA UPOTREBA U POSTUPCIMA ZA STATISTIČKU ANALIZU MULTIVARIJACIONIH PODATAKA²

Za primenjivanje bilo kog statističkog postupka analize podataka (analize varijanse, linearne korelacije, linearne regresije itd.) potrebno je da bude ispunjen niz pretpostavki, uglavnom u vezi sa distribucijom varijabli. Ukoliko jedna ili više ovih pretpostavki nisu ispunjene, procena parametra, koju dobijamo primenom analize, nije adekvatna. Metode poduzorkovanja nude se kao rešenje za prevazilaženje ovog problema, jer je za njihovo korišćenje potrebno napraviti samo jednu pretpostavku, a to je da podaci koje imamo u razumnoj meri predstavljaju populaciju iz koje su uzeti. Metodama poduzorkovanja se iz postojećeg, dostupnog uzorka ispitanika kreira veći broj novih poduzoraka. Na ovaj način se dobijaju empirijske distribucije željenih statistika, na osnovu kojih je moguće adekvatno proceniti parametre. U tekstu će biti predstavljeni metoda samouzorkovanja (*bootstrapping*) i postupak „univerzalnog noža” (*jackknifing*), koji spadaju u širu klasu postupaka poduzorkovanja. Najčešća varijanta postupka „univerzalnog noža” ponekad se naziva i metodom „izostavi jednog” jer podrazumeva pravljenje novih poduzoraka izbacivanjem jedne po jedne jedinice iz izvornog uzorka. Na ovaj način dobija se onoliko poduzoraka koliko je bilo jedinica u izvornom uzorku. U slučaju samouzorkovanja poduzorci se kreiraju nasumičnim biranjem jedinica iz izvornog uzorka sa vraćanjem, tako da svi poduzorci budu jednake veličine kao i izvorni uzorak. Kod metode samouzorkovanja broj poduzoraka je obično veoma veliki (preko 1000). Na primeru kanoničke diskriminacione analize prikazano je kako upotreba postupaka poduzorkovanja može u značajnoj meri izmeniti dobijene ocene parametara i zaključke istraživača, a detaljno je objašnjeno i kako se obe metode izvode u statističkom paketu IBM SPSS Statistics.

¹ Adresa autora:
dpuric@f.bg.ac.rs

Primljeno: 31. 01. 2013.
Primljena korekcija:
30. 04. 2013.
Prihvaćeno za štampu:
08. 05. 2012.

Ključne reči: poduzorkovanje, samouzorkovanje, metoda „univerzalnog noža”, uzorkovanje, kanonička diskriminaciona analiza

² Rad je prikazan na naučnom skupu XIX Empirijska istraživanja u psihologiji.

Uzorkovanje i statističko zaključivanje

Jedan od osnovnih zadataka statističkog zaključivanja je da se sa što većom tačnošću oceni vrednost nekog populacijskog parametra, koji je od interesa za istraživača. Parametar predstavlja statističku meru numeričke karakteristike populacije. Do tačne vrednosti parametra moguće je doći jedino merenjem svih jedinica koje pripadaju datoj populaciji, što bi predstavljalo ne samo dugotrajan, skup i često nemoguć postupak, već i suštinski nepotreban. Naime, izvlačenjem određenih jedinica iz populacije, moguće je oceniti vrednost parametra, naravno sa određenom verovatnoćom greške. Ovaj postupak izvlačenja jedinica iz populacije naziva se uzorkovanjem i celokupna statistika zaključivanja zasniva se na podacima dobijenim na uzorcima. Statistička mera numeričke karakteristike uzorka naziva se statistik.

Svaki put kada vršimo uzorkovanje, pravimo grešku u oceni parametra na osnovu statistika. Ova greška naziva se greška uzorka i određena je razlikom između populacijskog parametra i uzoračkog statistika koji predstavlja njegovu ocenu. Naravno, s obzirom na to da nam vrednost parametra nije poznata, ne možemo utvrditi ni vrednost greške uzorka, ali na osnovu određenih postupaka možemo proceniti verovatnoću njenog javljanja i veličine. Naime, kako znamo da bi različiti uzorci izvučeni iz iste populacije (sa istom vrednošću parametra) i sačinjeni od različitih entiteta imali različite vrednosti statistika, možemo definisati raspodelu verovatnoće dobijanja određene vrednosti statistika za uzorak izvučen iz date populacije. Ova raspodela naziva se distribucijom uzorkovanja i njena standardna devijacija predstavlja standardnu grešku datog statistika. Naravno, i standardna greška je sama po sebi parametar, koji se ocenjuje na osnovu vrednosti statistika dobijenog na konkretnom uzorku. Na osnovu standardne greške moguće je odrediti intervale pouzdanosti koji određuju granice u okviru kojih, sa određenim stepenom verovatnoće, odnosno uverenosti pretpostavljamo da se nalazi populacijski parametar (Tenjović, 2002).

Nedostaci parametrijskih metoda i načini njihovog prevazilaženja: poduzorkovanje

Opšte uzev, primenjivanje metoda parametrijske statistike podrazumeva da su ispunjeni određeni uslovi, odnosno da određene pretpostavke istraživača (najčešće u vezi sa oblikom distribucije) važe za datu varijablu u populaciji. Što je analiza koju primenjujemo kompleksnija, to je i broj uslova koji treba da budu zadovoljeni, da bi njeno primenjivanje bilo adekvatno, veći.

Kod analiza nekih od najvažnijih psiholoških varijabli primena parametrijskih metoda bi bila neopravdana jer one ne zadovoljavaju osnovne pretpostavke za analizu. Npr. opšte je poznato da se samopoštovanje ne distribuira normalno već ima negativno asimetričnu raspodelu (Baumeister, Tice, & Hutton, 1989; Schmitt & Allik, 2005). Sherman i Fennema (Sherman & Fennema, 1978) su utvrdili da uslov homogenosti varijanse između polova nije zadovoljen kod vizuo-spa-

cijalne inteligencije. Kod multivarijacionih metoda, pogotovo regresionih u kojima na osnovu psiholoških varijabli pokušavamo predvideti neki spoljni kriterijum, vrlo su česte situacije u kojima prediktori nisko koreliraju sa kriterijumom, a visoko međusobno, što gotovo uvek za rezultat ima čudne supresorske efekte. Pored toga, u okolnostima kada postoje ekstremni slučajevi (*outliers*) parametrijske metode mogu dati i pogrešne rezultate (Yu, 2003).

Stoga je, u onim slučajevima kada znamo da neka od pretpostavki neophodnih za primenjivanje željenog postupka analize nije ispunjena, ili ako samo sumnjamo u njenu ispunjenost, adekvatnije upotrebiti neki od postupaka poduzorkovanja čiji je cilj, takođe, da se oceni vrednost populacijskog parametra, ali tako što će se empirijskim putem utvrditi distribucija uzorkovanja traženog statistika.

Pojam poduzorkovanja (ili reuzorkovanja, *resampling*) predstavlja najopštiji pojam cele oblasti poduzorkovanja i pod njim se podrazumeva izvlačenje N novih poduzoraka uzimanjem jedinica iz postojećeg, dostupnog uzorka. Na osnovu svih ovih N poduzoraka ocenjuje se vrednost parametra, pri čemu nije neophodno praviti bilo kakve pretpostavke o distribuciji statistika, jer se ona empirijski dobija. Jedina pretpostavka koja se pravi prilikom korišćenja metoda poduzorkovanja jeste da podaci koje imamo u razumnoj meri predstavljaju populaciju iz koje su uzeti (Rogers, 1999; Yu, 2003). Ova pretpostavka je od suštinske važnosti za sve metode reuzorkovanja, budući da se podaci dobijeni na uzorku multiplikuju, te uzorak koji se po svojim karakteristikama značajno razlikuje od populacije može dovesti do pogrešnih zaključaka o populaciji.

Metode poduzorkovanja mogu se koristiti za ocenjivanje brojnih parametara, a mogu se razlikovati i nekoliko podvrsta poduzorkovanja. Tako, postoje metode za testiranje značajnosti razlika. Ova grupa metoda uključuje i test permutacija, odnosno, egzaktni test ili test randomizacije, kako se još naziva. Takođe, postoje i metode za validiranje pretpostavljenih modela na nasumičnim poduzorcima. Ove metode se uobičajeno nazivaju metode kros-validacije (Berger, 2006; Rudner, 1992; Yu, 2003). Time se u ovom radu nećemo detaljnije baviti, već ćemo se fokusirati na dve osnovne metode poduzorkovanja koje se koriste u ocenjivanju vrednosti parametara – samouzorkovanje (*bootstrapping*) i postupak „univerzalnog noža“ (*jackknifing*).

Još jedna grupa metoda koju treba pomenuti, bliska metodama poduzorkovanja, jeste Monte-Carlo grupa metoda (neki autori je čak smatraju podvrstom metoda poduzorkovanja – Berger, 2006). U ovoj grupi metoda, neophodno je da postoji teorijska distribucija sa poznatim karakteristikama iz koje se, zatim, izvlači veliki broj uzoraka. Na osnovu ovoga, moguće je generisati uzoračku distribuciju i oceniti parametre. Isto tako, moguće je i izračunati osetljivost nekog testa na neispunjenost pretpostavki o karakteristikama populacije, tako što ćemo porediti rezultate dobijene uzorkovanjem iz teorijskih populacija sa različitim karakteristikama (Yu, 2003). Suštinska razlika između ove grupe metoda i metoda poduzorkovanja je što se u Monte-Carlo metodama koriste teorijske distribucije simuliranih podataka, dok u metodama poduzorkovanja koristimo samo podatke koji su nam već dostupni na osnovu nekog uzorka i na osnovu njih dobijenu empirijsku distribuciju statistika.

Postupak „univerzalnog noža”

Jedna od najstarijih metoda poduzorkovanja svakako je postupak „univerzalnog noža” (*jackknifing*). Njega je 1956. godine kao metodu smanjenja pristrasnosti i davanja robustnih ocena intervala pouzdanosti predložio Quenouille, ali punu razradu i današnji matematički oblik dao mu je Tukey dve godine kasnije (Tukey, 1958, prema Miller, 1974). Ime ovog postupka treba da ilustruje njegovu široku primenljivost kao zamenu za nedostupne specijalizovane alate (Mosteller & Tukey, 1977, prema Rodgers, 1999).

U postupku „univerzalnog noža” postojeći uzorak se deli na određeni broj „grupa” slučajeva. Veličina ovih grupa može varirati od polovine uzorka do veoma malih grupa od jednog ili dva slučaja. U slučaju kada se uzorak deli na dve grupe, zapravo možemo govoriti o split-half metodi, pri čemu svaka od grupa sadrži polovinu jedinica izvornog uzorka. U praksi je ipak najčešći slučaj da se uzorak deli na onoliko grupa koliko ima jedinica u uzorku, tako da je veličina grupe jednaka jedinici (Chant & Dalgleish, 1992). Novi poduzorci se prave od dostupnih podataka tako što sve jedinice osim jedne grupe ulaze u poduzorak (Rodgers, 1999; Yu, 2003). Ovo bismo jednostavnije mogli objasniti tako što bismo rekli da se poduzorci prave izbacivanjem po jedne grupe slučajeva. Tako npr. ako uzorak podelimo na tri grupe, dobijamo tri poduzorka, pri čemu svakom od njih „nedostaje” jedna trećina jedinica, i to različita za svaki od poduzoraka. Ukoliko uzorak podelimo na četiri grupe, dobijamo četiri poduzorka, pri čemu svakom od njih „nedostaje” po jedna, različita četvrtina jedinica. U najčešćem slučaju, kada je broj grupa jednak broju jedinica u uzorku, novi poduzorci se prave izbacivanjem po jedne jedinice iz uzorka. Tako dobijamo onoliko poduzoraka koliko ima članova izvornog uzorka, pri čemu je broj članova poduzoraka uvek za jedan manji od broja članova izvornog uzorka zato što je u svakom poduzorku izostavljena po jedna jedinica. Istovremeno, svaka jedinica izostavljena je samo u jednom od poduzoraka, a prisutna je u svim drugim poduzorcima. Zbog ovakvog načina formiranja poduzoraka postupak „univerzalnog noža” naziva se ponekad i metodom „izostavi jednog”.

Iako neki autori (Mooney & Duval, 1993, prema Yu, 2003) smatraju da je postupak „univerzalnog noža” danas od pretežno istorijsko-matematičkog značaja i nema širu primenu, on je veoma koristan kada treba proceniti uticaj podgrupa na konačne rezultate (Yu, 2003). Po samoj svojoj prirodi (izostavljanja određenih jedinica iz analize), ovaj postupak odgovara situacijama kada želimo da primenimo neku klasifikacionu analizu – diskriminativnu analizu ili logističku regresiju, ali se može primenjivati i generalno prilikom ocenjivanja bilo kog parametra. Ipak, smatra se da je u većini takvih slučajeva uputnije primeniti metodu samouzorkovanja.

Samouzorkovanje

Tvorac nešto mlađe metode samouzorkovanja (*bootstrapping*), čije ime je nezaobilazno kad god se govori o postupcima poduzorkovanja, jeste američki statističar Bradley Efron. Efron je još 1977. godine predložio metodu samouzorko-

vanja kao inovativnu kompjutersku statističku tehniku i dao njeno matematičko rešenje. Njegov rad se oslanja na rad Quenouilleja i Tukeyja i nudi metodu samouzorkovanja kao usavršenu, pouzdaniju i šire primenljivu varijantu postupka „univerzalnog noža“ (Efron, 1977). Naziv samouzorkovanje dat je kako bi bio u sazvučju sa imenom postupka „univerzalnog noža“ (Efron, 2000), a potiče od engleskog izraza *to pull oneself by one's bootstraps*, što, u prenesenom smislu, znači postići uspeh bez oslanjanja na pomoć spolja.³

Samouzorkovanje se može definisati kao metod kojim se na osnovu dostupnih podataka iz nekog uzorka, kreira veliki broj novih uzoraka, jednake veličine kao i izvorni uzorak, nasumičnim biranjem opserviranih podataka sa vraćanjem, odnosno sa zamenom (Berger, 2006; Rodgers, 1999; Yu, 2003). Ovo znači da svaka jedinica ima jednaku verovatnoću da uđe u uzorak, i nakon što je ušla u uzorak, ona se vraća u populaciju odakle je izvučena. To istovremeno znači i da jedna jedinica može više puta ući u uzorak, zato što se verovatnoća njenog biranja ne menja tokom procesa formiranja uzorka. Primenom metode samouzorkovanja moguće je, na osnovu dostupnih podataka, napraviti neograničen broj novih poduzoraka; najčešće je broj ovih poduzoraka veoma veliki, a smatra se da je minimalno potrebno 100 uzoraka za procenu standardne greške parametra, i 2000 do 5000 ukoliko želimo da procenimo pristrasnost i/ili intervale pouzdanosti (Lunneborg, 2000). U svakom od ovih poduzoraka, dakle, neka jedinica će se naći više puta, dok neka druga jedinica uopšte neće ući u dati poduzorak. U podvarijanti samouzorkovanja, koja se naziva balansirano samouzorkovanje, izjednačava se broj pojavljivanja svake jedinice u svim poduzorcima (Dalglish, 1994).

Osnovni cilj ove metode, kao što je već i pomenuto, jeste procena populacijskih parametara. Za svaki od poduzoraka računaju se statistici koji interesuju istraživača, a zatim se formira njihova distribucija. Ova distribucija je, dakle, empirijska i ne oslanja se ni na kakve pretpostavke istraživača. Na osnovu njenih dobijenih vrednosti računaju se statistici i intervali pouzdanosti oko njih.

Primeri primene postupka „univerzalnog noža“ i metode samouzorkovanja

U narednom odeljku kratko ćemo prikazati način primene pomenutih metoda poduzorkovanja na primeru kanoničke diskriminacione analize.

Kanonička diskriminaciona analiza se koristi kada na osnovu više kontinuiranih varijabli želimo predvideti pripadnost ispitanika nekoj grupi.⁴ To je metod izbora za utvrđivanje dijagnostičke valjanosti. Uslovi za korišćenje ove metode su sledeći (videti Green, Salkind, & Jones, 1996):

³ Bukvalno značenje izraza bilo bi „povući se za kaiševe na čizmama“ ili „izvući se povlačenjem za kaiševe na čizmama“.

⁴ Kanonička diskriminaciona analiza nije tema ovog rada pa nećemo ulaziti dublje u objašnjavanje metode.

- Multinormalna raspodela – sve kvantitativne varijable treba da budu normalno distribuirane u svakoj subpopulaciji;
- Kontinuirane varijable moraju biti linearno nezavisne, tj. nijedna varijabla ne sme da bude linearna kombinacija ostalih varijabli. Nepoželjna je i visoka kolinearnost između kontinuiranih varijabli;
- Populacione varijanse/kovarijanse kvantitativnih varijabli treba da budu jednake na svim nivoima nominalne varijable;
- Ispitanici treba da budu nezavisni, dobijeni slučajnim uzorkovanjem tako da rezultat nekog ispitanika na nekoj varijabli ni na koji način ne zavisi od rezultata drugih ispitanika;
- Kategorije moraju biti ekskluzivne i sveobuhvatne te ne smeju biti rezultat pretvaranja kontinuirane varijable u diskretnu.

Iz didaktičkih razloga prvo ćemo prikazati rezultate kanoničke diskriminacione analize izvedene na klasičan način i izvesti odgovarajuće zaključke. Zatim ćemo te rezultate i zaključke uporediti sa rezultatima i zaključcima koji se mogu izvesti nakon primene postupka „univerzalnog noža“, a potom i metode samouzorkovanja.

Metode će biti predstavljene na stvarnim rezultatima dobijenim u jednom realnom neobjavljenom istraživanju. Istraživanje je sprovedeno na uzorku od 203 ispitanika muškog pola srednjoškolskog uzrasta i to 48 iz porodica bez oca, 51 iz porodica u kojima je otac živ ali zbog razvoda adolescent živi samo s majkom i 104 iz dvoroditeljskih nuklearnih porodica. Kao uslovno nezavisne varijable poslužili su rezultati na 10 faceta self-koncepta merenih instrumentom Gensel (Opačić, 1995), verzija 3, koji meri: emocionalnost, rigidnost, mizantropiju, moralni relativizam, eksternalnost, globalnu kompetentnost, izgled, prihvaćenost od strane vršnjaka, fizičku snagu i spretnost i samoprocenu intelektualnog funkcionisanja.

Rezultati

Klasični pristup

Kanonička diskriminaciona analiza pokazala je da je od dve moguće kanoničke diskriminacione funkcije jedna značajna ($r = .39$, $\chi^2(8) = 39.26$, $p = .006$) i ona objašnjava 14.8% varijanse razlika između grupa. Ovaj nalaz dao bi nam za pravo da tumačimo koeficijente kanoničke diskriminacione funkcije (Tabela 1), koeficijente strukture (Tabela 2) i centroide kanoničke diskriminacione funkcije (Tabela 3).⁵

⁵ Tabele su namerno date u izvornom obliku. APA standard podrazumeva drugačiji način prikazivanja rezultata kanoničke diskriminacione analize.

Tabela 1

Koeficijenti kanoničke diskriminacione funkcije

Standardized Canonical Discriminant Function Coefficients^a

	Function	
	1	2
emocionalnost	-.222	-.642
rigidnost	.374	.221
mizantropija	-.332	.120
moralni relativizam	.222	.035
eksternalnost	-.086	.795
globalna kompetentnost	.209	-.082
izgled	-.001	.131
prihvaćenost od strane vršnjaka	-.743	.339
fizička snaga i spretnost	.814	-.034
samoprocena intelektualnog funkcionisanja	.342	-.146

Kao što je vidljivo iz Tabele 1, dve varijable – samoprocena fizičke snage i prihvaćenost od strane vršnjaka – „najodgovornije” su za izgradnju ove funkcije. Drugu kolonu ne bismo interpretirali jer je pripadajuća kanonička funkcija neznačajna. Na osnovu koeficijenata strukture pokušali bismo da proniknemo u „dublju psihološku prirodu” prve kanoničke diskriminacione funkcije.

Tabela 2

Koeficijenti strukture kanoničke diskriminacione funkcije

Structure Matrix^a

	Function	
	1	2
fizička snaga i spretnost	.582*	.260
globalna kompetentnost	.298*	-.184
emocionalnost	-.254*	-.179
samoprocena intelektualnog funkcionisanja	.250*	.006
mizantropija	-.232*	.183
eksternalnost	-.095	.753*
izgled	.096	.399*
rigidnost	.285	.362*
prihvaćenost od strane vršnjaka	-.164	.356*
moralni relativizam	.034	.179*

Pooled within-groups correlations between discriminating variables and standardized canonical discriminant functions

Variables ordered by absolute size of correlation within function.

**. Largest absolute correlation between each variable and any discriminant function*

Kombinujući nalaze iz Tabela 1 i 2, zaključili bismo da je psihološka priroda kanoničke diskriminacione funkcije najviše determinisana samoprocenom fizičke snage i spretnosti, a da je prihvaćenost od strane vršnjaka tipičan supresor. Uvid u Tabelu 3 daje nam položaj ispitanika na ovoj funkciji.

Tabela 3

Centroidi grupa na kanoničkoj diskriminacionoj funkciji

Functions at Group Centroids^a

Porodična struktura	Function	
	1	2
0 smrt oca/ živi s majkom	-.734	-.060
1 nuklearna intaktna porodica	.283	-.140
2 razvod/ živi s majkom	.114	.343

Unstandardized canonical discriminant functions evaluated at group means

Zaključak nakon uvida u tabelu sa centroidima bio bi da mladići iz porodica bez oca imaju lošiju sliku o sebi na tipično muškim atributima (snaga i spretnost), a u situacijama kada su izjednačeni po ostalim karakteristikama, imaju viši rezultat na tipično ženskom domenu samovrednovanja (socijalnost – prihvaćenost od strane vršnjaka). U pomoć bismo pozvali teoriju učenja po modelu ili neku sličnu i zaključili npr. da su predikcije koje slede iz psihoanalitičke teorije potpuno odbačene kao neadekvatne.

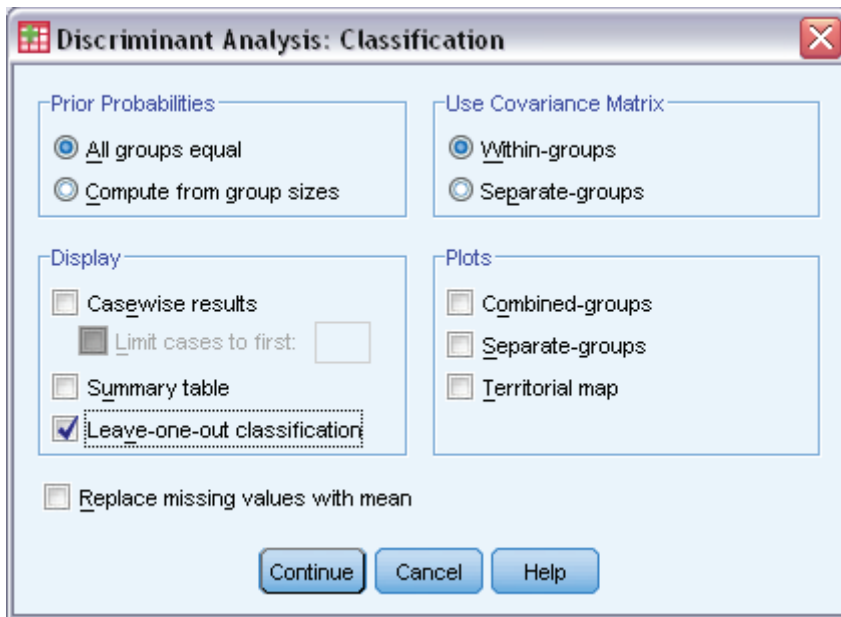
Da li je to baš tako?

Za proveru homogenosti varijansi/kovarijansi u subpopulacijama, kao jednog od uslova za kanoničku diskriminacionu analizu, korišćen je Boxov M (Box, 1949). Nulta hipoteza je da se matrice varijansi/kovarijansi u ispitivanim grupama ne razlikuju, a njegova statistička značajnost se aproksimira pomoću F -testa. Budući da je Boxov test značajan ($F(110, 57938) = 1.49, p = .001$),⁶ možemo zaključiti da se matrice varijansi/kovarijansi ispitivanih varijabli statistički značajno razlikuju. Ovaj rezultat praktično znači da je prisutna nehomogenost varijansi pojedinih varijabli i/ili da su odnosi između varijabli različiti na različitim nivoima nominalne varijable (što je u suštini definicija kvalitativnih razlika). U ovom slučaju, dakle, preporuka je da se primeni neka od metoda poduzorkovanja.

⁶ Boxov test osetljiv je kako na nehomogenost matrica kovarijansi, tako i na odstupanje multivarijantne raspodele od normalne (Tenjović, 2002).

Postupak „univerzalnog noža” u SPSS-u

U najnovijoj verziji SPSS-a (IBM SPSS Statistics 20), kao i u većini prethodnih verzija, postupak „univerzalnog noža” zastupljen je samo jednom svojom (istina najčešće korišćenom) varijantom – metodom „izostavi jednog” – i to samo u okviru diskriminacione analize. Ovo je razumljivo, imajući u vidu da je diskriminaciona analiza kao oblik klasifikacione analize najčešće polje primene metode „univerzalnog noža”. Da bi uključio ovu opciju, korisnik treba da u okviru prozora *Discriminant Analysis*, pritisne dugme *Classify*, nakon čega će se pojaviti prozor gde je moguće štiklirati opciju *Leave-one-out classification* (vidi Sliku 1). Kako bi ova opcija bila omogućena, potrebno je da se analiza vrši na unutargrupnoj matrici kovarijansi (*Within groups*). Biranjem date opcije diskriminaciona funkcija će se praviti na osnovu podataka svih ispitanika osim jednog, a zatim će taj ispitanik biti klasifikovan (na osnovu funkcije u čijem formiranju nije učestvovao).



Slika 1. Prozor za uključivanje opcije „izostavi jednog”

Primer primene postupka „univerzalnog noža”

Kao primer primene postupka „univerzalnog noža” biće korišćena kanonička diskriminaciona analiza na istim podacima koji su prikazani u okviru klasičnog pristupa – 203 adolescenta čiji su porodična struktura i skorovi na dimenzijama self-koncepta poznati.

Primenom metode „univerzalnog noža”, ono što se razlikuje u odnosu na klasičnu kanoničku diskriminacionu analizu jeste poslednja tabela SPSS-ovog ispisa (ovde označena kao Tabela 4), u kojoj su navedeni rezultati klasifikacione analize. U redu *Original*, nalaze se rezultati klasifikacije ukoliko se diskriminaciona funkcija pravi na svim ispitanicima, a zatim na istim tim ispitanicima i proverava, dok je *Cross-validated* red onaj koji se odnosi na rezultate dobijene primenom metode „izostavi jednog”.

Tabela 4

Tačnost klasifikacije upotrebom klasičnog postupka i postupka „univerzalnog noža”

Classification Results^{a,b,d}

	Porodična struktura	Predicted Group Membership			Total	
		smrt oca/ živi s majkom	nuklearna intaktna porodica	razvod / živi s majkom		
Original	Count	smrt oca/ živi s majkom	26	12	10	48
		nuklearna intaktna porodica	21	52	31	104
		razvod/ živi s majkom	14	14	23	51
	%	smrt oca/ živi s majkom	54.2	25.0	20.8	100.0
		nuklearna intaktna porodica	20.2	50.0	29.8	100.0
		razvod/ živi s majkom	27.5	27.5	45.1	100.0
Cross-validated ^c	Count	smrt oca/ živi s majkom	19	17	12	48
		nuklearna intaktna porodica	23	43	38	104
		razvod/ živi s majkom	15	17	19	51
	%	smrt oca/ živi s majkom	39.6	35.4	25.0	100.0
		nuklearna intaktna porodica	22.1	41.3	36.5	100.0
		razvod/ živi s majkom	29.4	33.3	37.3	100.0

a. pol ispitanika = muški pol

b. 49.8% of original grouped cases correctly classified.

c. Cross validation is done only for those cases in the analysis. In cross validation, each case is classified by the functions derived from all cases other than that case.

d. 39.9% of cross-validated grouped cases correctly classified.

Kao što možemo videti, procenat tačno klasifikovanih slučajeva značajno je niži kada se koristi metoda „izostavi jednog“, za čitavih 10% (39.9% tačno klasifikovanih metodom „izostavi jednog“, naspram izvornih 49.8%). Ovako velika razlika govori o tome da je stvarna mogućnost klasifikovanja adolescenata u kategorije različitih tipova porodice na osnovu njihovog self-koncepta značajno niža nego što bismo to zaključili na osnovu inicijalne analize. Ovaj podatak ima znatno veći praktični nego teorijski značaj, posebno kod upotrebe različitih dijagnostičkih postupaka gde je tačnost klasifikacije jedan od presudnih kriterijuma valjanosti.

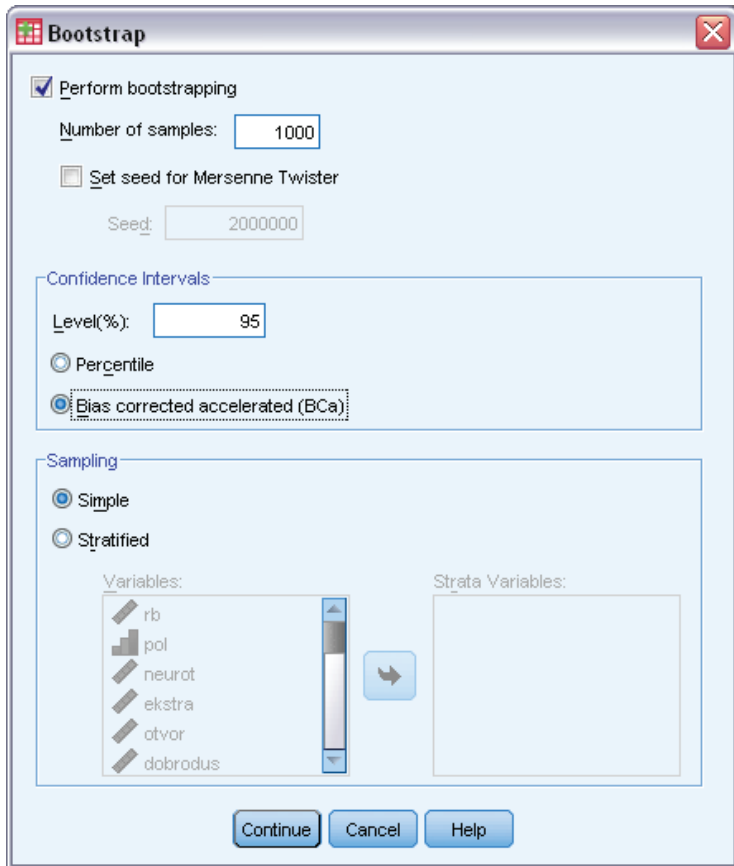
Samouzorkovanje u SPSS-u

U najnovijoj verziji SPSS-a (IBM SPSS Statistics 20) metode samouzorkovanja (*bootstrapping*) podržane su za brojne osnovne i napredne opcije koje se često koriste: *Frequencies, Descriptives, Explore, Crosstabs, Means, One-Sample T test, Independent-Samples T test, Paired-Samples T test, One-Way ANOVA, GLM Univariate, Bivariate Correlations, Partial Correlations, Linear Regression, Ordinal Regression, Discriminant Analysis, GLM Multivariate, Linear Mixed Models, Generalized Linear Models, Cox Regression, Binary Logistic Regression, Multinomial Logistic Regression*.

Bez obzira na to koju od datih analiza korisnik želi da izvrši, treba izabrati podopciju *Perform Bootstrap*, pri čemu se pojavljuje prozor za definisanje vrste bootstrapping uzorka (vidi Sliku 2).

Pre svega, treba uključiti opciju *Perform Bootstrap*, a zatim (po potrebi) načiniti odgovarajuća podešavanja. Broj uzoraka je automatski podešen na 1000, ali se ovo može izmeniti. Opcija *Set seed for Marsenne Twister* omogućava postavljanje početne tačke za generator nasumičnih brojeva i uglavnom se koristi za replikaciju već izvršenog postupka. Nivo pouzdanosti podešen je na 95% i takođe se može izmeniti. Opcija *Bias corrected accelerated (BCa)* daje intervale pouzdanosti koji su prilagođeni s obzirom na pristrasnost i ubrzanje uzoračke distribucije.⁷ Kako se teorijski i empirijski pokazalo da je ova metoda računanja intervala pouzdanosti pouzdanija od percentilnih metoda (Lunneborg, 2000), preporučujemo korisnicima da je izaberu za svoje analize. Takođe, moguće je vršiti prosto i stratifikovano uzorkovanje, pri čemu je za stratifikovano uzorkovanje neophodno definisati varijable koje određuju stratume.

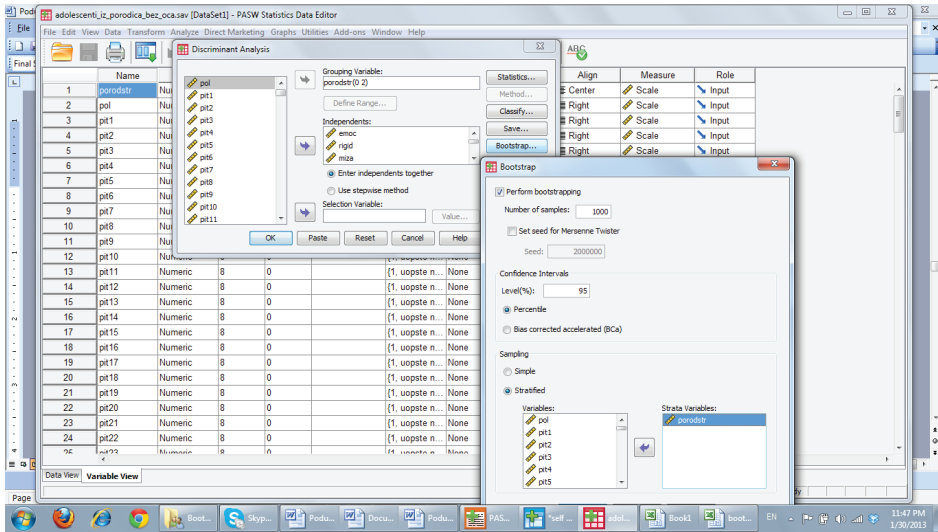
⁷ Pristrasnost se odnosi na vrednost aritmetičke sredine uzoračke distribucije, dok se ubrzanje odnosi na smanjenje standardne devijacije uzoračke distribucije sa promenom vrednosti parametra.



Slika 2. Prozor za definisanje bootstrap poduzorka

Primer primene samouzorkovanja

Kao što je vidljivo iz Slike 3, pored elemenata koji su univerzalni za sve postupke samouzorkovanja, kod metoda u kojima je bitna struktura uzorka (ANOVA, kanonička diskriminaciona analiza, logistička regresija, multinomijalna regresija itd.) potrebno je navesti i varijablu na osnovu koje se vrši stratifikacija. U našem slučaju to je varijabla porodična struktura.



Slika 3. Zadavanje samouzorkovanja unutar opcije *Discriminant Analysis*

Nakon primene metode samouzorkovanja dobijamo iste rezultate kao i korišćenjem klasičnog pristupa, ali pored navedenih parametara dobijamo procenu pristrasnosti, standardnu grešku, značajnost i intervale pouzdanosti za svaki pojedinačni standardizovani koeficijent kanoničke diskriminacione funkcije.

Pregledom vrednosti iz Tabele 5 u kojoj su prikazani standardizovani koeficijenti kanoničke diskriminacione funkcije dobijeni primenom metode samouzorkovanja, možemo videti da je pristrasnost koeficijenata mala do umereno visoka, dok je standardna greška merenja prilično visoka. Ovo na bitan način utiče i na dobijene intervale pouzdanosti za sve koeficijente, odnosno granice u kojima se najverovatnije nalazi vrednost parametara. U prikazanom primeru, intervale pouzdanosti za sve koeficijente veoma su široki, i kreću se od negativnih do pozitivnih vrednosti, obuhvatajući i vrednost nula.⁸ Ovo znači da bi ponovno uzorkovanje iz iste populacije dalo vrednosti diskriminacionih koeficijenata koje su značajno više, ali i značajno niže, pa čak i suprotnog predznaka u odnosu na dobijene. Jasno je, samim tim, da postoji velika verovatnoća da je vrednost populacijskih koeficijenata zapravo jednaka nuli. Drugim rečima, ne možemo prihvatiti rezultate dobijene kanoničkom diskriminacionom analizom, jer verovatnoća da je izdvojena diskriminaciona funkcija zapravo dobijena slučajno prelazi graničnu vrednost od .05. Naravno, posle ovog zaključka ne bismo se upuštali u tumačenje i teorijska razmatranja, a najverovatnije ni pokušali da objavimo rad.

⁸ Intervali pouzdanosti koji ne obuhvataju nulu ukazuju na realno postojeći doprinos date varijable.

Tabela 5
Standardizovani koefijenti kanoničke diskriminacione funkcije dobijeni metodom samouzorkovanja
Standardized Canonical Discriminant Function Coefficients

	Function	Coefficient	Bootstrap ^a			
			Bias	Std. Error	95% Confidence Interval	
					Lower	Upper
emocionalnost	1	-.222	.084	.314	-.703	.556
	2	-.642	.525	.502	-.830	.962
rigidnost	1	.374	-.105	.285	-.476	.691
	2	.221	-.166	.312	-.551	.642
mizantropija	1	-.332	.157	.305	-.601	.557
	2	.120	-.034	.304	-.507	.712
moralni relativizam	1	.222	-.095	.257	-.431	.581
	2	.035	-.031	.361	-.714	.695
eksternalnost	1	-.086	.101	.281	-.484	.632
	2	.795	-.439	.487	-.695	.995
globalna kompetentnost	1	.209	-.026	.214	-.232	.598
	2	-.082	.202	.292	-.423	.656
izgled	1	-.001	.006	.247	-.473	.491
	2	.131	-.073	.340	-.609	.737
prihvaćenost od strane vršnjaka	1	-.743	.288	.507	-1.033	.860
	2	.339	-.240	.401	-.712	.817
fizička snaga i spretnost	1	.814	-.297	.521	-.875	1.022
	2	-.034	.077	.309	-.531	.650
samoprocena intelektualnog funkcionisanja	1	.342	-.109	.294	-.462	.694
	2	-.146	.133	.327	-.613	.640

a. Unless otherwise noted, bootstrap results are based on 1000 bootstrap samples

Dakle, metod samouzorkovanja nas čuva od olakih zaključaka i pokušaja da teorijski osmislimo rezultate koji su u stvari proizvod slučaja.

Zaključak

Sličnosti i razlike između samouzorkovanja i postupka „univerzalnog noža“

Obe metode namenjene su istoj svrsi, a to je procena populacijskih parametara. Obe se zasnivaju na korišćenju postojećih i dostupnih podataka. To ih razlikuje od Monte-Carlo metoda, i čini ih sličnim metodama randomizacije i kros-validacije. U obe se na osnovu većeg broja poduzoraka računaju statistici, njihove distribucije i intervali pouzdanosti za date statistike. Drugim rečima, procene statistika ne zasnivaju se na pretpostavljenim, već na opserviranim, empirijskim distribucijama.

Jedna od osnovnih razlika između samouzorkovanja i postupka „univerzalnog noža“ jeste što samouzorkovanje kada se ponavlja više puta na istim podacima može dati neznatno drugačije rezultate, dok postupak „univerzalnog noža“ daje uvek identične rezultate za isti skup podataka. Ovo je posledica slučajnog biranja jedinica koje ulaze u poduzorke u slučaju metode samouzorkovanja i sistematskog izbacivanja slučajeva u okviru postupka „univerzalnog noža“. Zatim, metoda samouzorkovanja se zasniva na uzorkovanju sa vraćanjem, a postupak „univerzalnog noža“ na uzorkovanju bez vraćanja. Metoda samouzorkovanja se zato smatra preciznijom. Njom se, takođe, može produkovati teoretski neograničen broj poduzoraka na osnovu datih podataka, dok se postupkom „univerzalnog noža“ može napraviti maksimalno onoliko poduzoraka koliki je broj slučajeva u izvornom uzorku. Još jedna od prednosti metode samouzorkovanja je i ta što je veličina poduzoraka identična veličini originalnog uzorka, što može pomoći u modelovanju uticaja veličine uzorka na dobijene rezultate.

Prednosti i mane postupaka poduzorkovanja

Istraživači koji su se bavili metodama poduzorkovanja navode više razloga zbog kojih smatraju upotrebu ovih metoda opravdanim. U osnovnim crtama, njihova argumentacija zasniva se na činjenici da je za upotrebe metoda poduzorkovanja potrebno napraviti samo jednu pretpostavku – da podaci u razumnoj meri predstavljaju populaciju iz koje su izvučeni (Rogers, 1999; Yu, 2003). Nije neophodno pretpostavljati ništa o obliku distribucije varijable, ni u populaciji, ni u uzorku. Zatim, iako prvenstveno razvijene za primenu na malim uzorcima, ove metode su pogodne čak i kod velikih uzoraka, jer je verovatnoća odbacivanja tačnih, kao i prihvatanja netačnih pretpostavki manja. Na kraju, problem replikabilnosti nalaza istraživanja se umanjuje, jer metode poduzorkovanja već u sebi po definiciji sadrže niz replikacija dobijenih nalaza. Naravno, ovakva unutrašnja replikacija nikako nije zamena za spoljašnju replikaciju, dobijenu na drugim podacima, ali svakako

dosta govori o stabilnosti dobijenih rezultata (Yu, 2003). Tako čak i u slučajevima kada uzorci nisu reprezentativni, dobijamo značajne podatke o stabilnosti rezultata za slične grupe ispitanika. Drugim rečima, u ovim metodama je odnos „ulaza” i „izlaza” optimalan jer su uslovi za primenjivanje postupaka minimalni, a mogu se koristiti u svakoj situaciji, koliko god da je ona komplikovana (Efron & Gong, 1983).

Od zamerki koje se mogu uputiti ovom skupu metoda, najsnažnija je svakako ta da se sve vreme koriste isti podaci koji se samo „vrte u krug”, pa se postavlja pitanje da li zaista možemo dobiti valjanije podatke od onih koje već imamo (Rudner & Shafer, 1992). Zatim, u slučaju kada uzorak loše predstavlja populaciju iz koje je izvučen, metode poduzorkovanja će samo ponavljati i eventualno povećati ovu grešku (Rogers, 1999; Yu, 2003). Treba, međutim, primetiti da se ove zamerke odnose na situacije kada bismo i primenom drugih, klasičnih parametrijskih metoda pravili greške u zaključivanju o populacijskim parametrima. Ozbiljnija zamerka je, stoga, da u situacijama kada su istraživačke hipoteze o distribuciji tačne za populaciju, metode poduzorkovanja daju nepreciznije mere nego klasične parametrijske metode (Yu, 2003).

Za kraj, treba razmotriti i praktičnost i primenljivost ovih metoda. Činjenica da najnovija verzija SPSS-a u sebi sadrži već ugrađenu opciju samouzorkovanja za najveći broj analiza govori dosta o mogućnostima njene široke primene. Uključivanje metoda poduzorkovanja u regularne analize ne oduzima istraživaču mnogo vremena, a ovako dobijeni rezultati predstavljaju značajnu dopunu i svojevrsnu proveru rezultata dobijenih parametrijskim metodama. U našim primerima, razlike u osnovnim nalazima dobijenim primenom ovih dveju metoda bile su značajne i relevantne. Međutim, čak i u slučajevima kada metode poduzorkovanja daju slične rezultate onima koji se dobijaju bez njihove primene, istraživači mogu, osim dodatnog poverenja u svoje nalaze, dobiti informacije koje drugim metodama ne bi mogle biti dobijene. Stoga, zaključujemo da je metode samouzorkovanja najbolje koristiti u kombinaciji sa klasičnim parametrijskim metodama.

Reference

- Baumeister, R. F., Tice, D. M., & Hutton, D. G. (1989). Self-Presentational Motivations and Personality Differences in Self-Esteem. *Journal of Personality, 57*, 547–579.
- Berger, D. E. (2006). *Introduction to Resampling*. Statistics Workshop presented at the meeting of the Western Psychological Association, Palm Springs, CA.
- Box, G. E. (1949). A general distribution theory for a class of likelihood criteria. *Biometrika, 36*, 317–346.
- Chant, D., & Dalglish, L. I. (1992). A SAS Macro for Jackknifing the Results of Discriminant Analyses. *Multivariate Behavioral Research, 27*, 323–333.
- Dalglish, L. I. (1994). Discriminant Analysis: Statistical Inference Using the Jackknife and Bootstrap Procedures. *Psychological Bulletin, 116*, 498–508.

- Efron, B. (1977). *Bootstrap Methods: Another Look at the Jackknife*. California: Division of Biostatistics, Stanford University.
- Efron, B., & Gong, G. (1983). A Leisurely Look at the Bootstrap, the Jackknife and Cross-Validation. *The American Statistician*, 37, 36–48.
- Efron, B. (2000). The Bootstrap and Modern Statistics. *Journal of the American Statistical Association*, 95, 1293–1296.
- Green, S. B., Salkind, N. J., & Jones, T. M. (1996). *Using SPSS for Windows; analyzing and understanding data*. Prentice Hall PTR.
- Lunneborg, C. E. (2000). *Data Analysis by Resampling: Concepts and Applications*. Pacific Grove, CA: Brooks/Cole/Duxbury.
- Miller, R. G. (1974). The Jackknife – A Review. *Biometrika*, 61, 1–15.
- Opačić, G. Đ. (1995). *Ličnost u socijalnom ogledalu*. Beograd: Institut za pedagoška istraživanja, IPI.
- Rodgers, J. L. (1999). The Bootstrap, the Jackknife, and the Randomisation Test: A Sampling Taxonomy. *Multivariate Behavioral Research*, 34, 441–456.
- Rudner, L. M., & Shafer, M. M. (1992). *Resampling: A Marriage of Computers and Statistics*. Washington, DC: ERIC Clearinghouse on Tests Measurement and Evaluation.
- Schmitt, D. P., & Allik, J. (2005). Simultaneous administration of the Rosenberg Self-Esteem Scale in 53 nations: Exploring the universal and culture-specific features of global self-esteem. *Journal of Personality and Social Psychology*, 89, 623–642.
- Sherman, J. A., & Fennema, E. (1978). Distribution of spatial visualization and mathematical problem solving scores: A test of Stafford's X-linked hypotheses. *Psychology of Women Quarterly*, 3, 157–167.
- Tenjović, L. (2002). *Statistika u psihologiji, priručnik (Dopunjeno izdanje)*. Beograd: Centar za primenjenu psihologiju.
- Yu, C. H. (2003). Resampling methods: concepts, applications, and justification. *Practical Assessment, Research & Evaluation*, 8. Preuzeto sa <http://PAREonline.net/getvn.asp?v=8&n=19>

Danka Purić and Goran Opačić

Department of
Psychology,
Faculty of Philosophy,
University of Belgrade

RESAMPLING, BOOTSTRAPPING, JACKKNIFING AND THEIR USE IN MULTIVARIATE (STATISTICAL) DATA ANALYSES

In order to perform any data analysis procedure (analysis of variance, linear correlation, linear regression, etc.), a series of assumptions need to be true, most of which are concerned with the distribution of the variables. If one or more of these assumptions are violated, the obtained parameter estimates are inadequate. Resampling methods are offered as a means to overcome this issue, as they only require one assumption to be made - that the available data is reasonably representative of the population. Resampling methods use existing, available samples to create a large number of new subsamples. This produces an empirical distribution of the desired statistics, which forms a basis for an adequate parameter estimate. In this paper, we will discuss the methods of bootstrapping and jackknifing, which fall under the broader category of resampling.

The most common form of the jackknifing procedure is sometimes labeled "leave-one-out" procedure, because new subsamples are made by excluding one unit of the original sample at a time. Thus, the number of new subsamples is equal to the number of units in the original sample. In bootstrapping, subsamples are created by randomly picking units from the original subsample with replacement so that all subsamples are equal in size to the original sample. The number of bootstrapping subsamples is usually very large (over 1000). A canonical discriminant analysis will serve as an example to illustrate how the usage of resampling procedures can significantly alter the obtained parameter estimates and researcher's conclusions. A detailed explanation on how to perform these methods in IBM SPSS Statistics package is also given.

Keywords: resampling, bootstrapping, jackknifing, sampling, canonical discriminant analysis