Osnove psihološke statistike

Vladimir Hedrih, Filozofski fakultet, Univerzitet u Nišu, Srbija

Andjelka Hedrih, Matematički institut SANU, Beograd, Srbija

PREPRINT

Ovo je preliminarna verzija srpske verzije knjige, koja još nije zvanično objavljena niti katalogizovana, te nema još ni pune bibliografske podatke. Ukoliko se u nekom svom tekstu pozivate na neki od sadržaja ove knjige, molimo citirajte inicijalnu, zvanično objavljenu englesku verziju knjige:

Hedrih, V., & Hedrih, A. (2022). *Interpreting statistics for beginners : a guide for behavioural and social scientists*. Routledge, Taylor&Francis Group. https://doi.org/10.4324/9781003107712

PREPRINT

Dragi čitaoci,

živimo u vremenu nauke i to je razlog zašto većina ljudi koji radi na poslovima koji zahtevaju fakultetsko obrazovanje ima, makar s vremena na vreme, potrebu da čita i interpretira statističke podatke i rezultate statističkih proračuna. Sa jednakom učestalošću pojavljuje se i potreba da se proveri validnost zaključaka i različitih tvrdnji izvedenih iz statističkih podataka. Veština izvođenja ispravnih zaključaka u ovakvim situacijama može značiti razliku između uspeha i neuspeha u poslu koji osoba radi. Većina ljudi, većinu vremena, su korisnici statističkih rezultata koje su drugi uradili, dok je broj situacija gde ljudi sami rade statičke analize mnogo manji. Međutim, većina osnovne statističke literature je napisana sa ciljem da nauči čitaoca kako da sam radi statističke analize uz srazmerno mnogo manje posvećene pažnje interpretaciji rezultata. Pored toga, statistika je nauka koja se neprekidno razvija, a isti je slučaj i sa onim njenim delovima koji su postali centralni deo metodologije naučnih istraživanja i obrade podataka. Dok su u decenijama pre nas, shvatanja naučnika iz različitih oblasti nauke, a pre svega iz oblasti društvenih nauka bila takva da su favorizovala upotrebu određenih statističkih postupaka na određene načine, danas su ta shvatanja u velikoj meri drugačija, a može se očekivati da će trend promena i razvoja dominantne statističke metodologije koja se primenjuje u istraživanjima nastaviti i u budućnosti.

Ova knjiga predstavlja pregled osnovnih statističkih postupaka, ali datih na način koji je pre svega fokusiran na interpretaciju rezultata. Na taj način, ona predstavlja novi pristup predstavljaju statističkih tehnika koji je fokusiran na način na koji se one koriste u praksi istraživačkog rada i na to kako se rezultati dobijeni njima valjano interpretiraju kada na njih naiđemo u naučnoj i stručnoj literaturi. S jedne strane, ovo je knjiga koja predstavlja uvod u oblast statistike, ali s druge strane koja se ne koncentriše na navođenje formula, već na to da čitaocima predstavi osnovne koncepte i ideje na kojima su zasnovane statističke analize, načine na koji se one koriste u savremenim istraživanjima i valjano interpretiraju, kao i opšti naučni i epistemološki okvir u kom se statističke analize rade. Pisana je kako za ljude koji tek počinju da uče statistiku ili ljude koji se ne razumeju u statistiku uopšte, ali koji imaju potrebu da mogu čitaju i razumeju statsitičke podatke u svom radu, ali i za ljude koji imaju iskustva u primeni statistike, a potreban im je pregled osnovnih metoda u upotrebi u nauci i to kako ih valjano interpretirati. Iz ovog razloga, ova knjiga sadrži puni primera statističkih podataka i njihovih intepretacija preuzetih iz različitih naučnih oblasti u

PREPRINT

okviru polja društvenih nauka i nauka o ponašanju. Ovo je takođe i knjiga koja pokriva sadržaje uvodnog kursa iz statistike koji prvi autor predaje na svom univerziteti, tako da je na neki način i više od 20 godina iskustva u razvoju ovog predmeta kondenzovano u stranice ove knjige.

Prvu verziju ove knjige izdali smo 2022. godine za međunarodnu publiku kod britanskog/američkog izdavača Routledge-a pod imenom Interpreting Statistics for Beginners: A Guide for Behavioral and Social Sciences, a sada se pred vama nalazi izmenjeno i dopunjeno izdanje namenjeno srpskoj publici. U odnosu na englesku verziju, ova verzija knjige sadrži niz dopuna i izmena kojima smo hteli da bolje prilagodimo sadržaje korisnicima, ali i unapredimo kvalitet teksta.

Uživajte!

Autori.

PREPRINT

Sadržaj

Sadržaj 5
Pre nego što krenemo dalje: Opšta uputstva za vežbe 7
Poglavlje 1. Naučne paradigme i naučna objašnjenja, pseudonauka 9
1.1. Nauka, naučne paradigme i statistika 10
1.2. Priroda uzročno posledične mreže univerzuma – stohasticizam naspram determinizma. 23
1.3. Naučna objašnjenja 27
1.4. Samoispunjujuća proročanstva i kvarenje statističkih pokazatelja 32
1.5. Nauka i pseudonauka 37
Poglavlje 2. Osnovni koncepti statistike 45
2.1. Slučajni događaj 45
2.2. Verovatnoća 47
2.3. Entitet 49
2.4. Varijable i konstante 50
2.5. Organizacija podataka, matrica, vektor 51
2.6. Populacija i uzorak, parametri i statistici 56
2.7. Uzorkovanje 61
2.8. Vrste uzoraka 65
2.9. Čega treba da budemo posebno svesni u vezi uzorkovanja? 74
2.10. Nivoi merenja 76
2.11. Kontinualne i diskretne varijable 82
2.12. Hajde da primenimo to što smo naučili do sada! 83
Poglavlje 3. Deskriptivna statistika 88
3.1. Distribucija 88
3.2. Percentili i ostali kvantili 90
3.3. Mere centralne tendencije 92
3.4. Mere varijabilnosti 96
3.5. Kako se distribucija može predstaviti? 100
3.6. Hajde da primenimo šta smo do sada naučili! 108
Poglavlje 4. Distribucije 115
4.1. Teorijske i empirijske distribucije 115
4.2. Normalna distribucija 116

PREPRINT

4.3. Puasonova distribucija 120
4.4. Binomna distribucija 122
4.5. Uniformna distribucija 124
4.6. Odstupanja od normalne distribucije 125
4.7. Standardni skorovi i standardizacija 143
4.8. Ipsatizacija. 147
Poglavlje 5. Statistika zaključivanja, osnovni pojmovi 157
5.1. Centralna granična teorema 158
5.2. Pristup zaključivanju o vrednostima parametara preko butstrepinga 163
5.3. Procena parametara populacije, tačkasta i intervalna procena. 165
5.4. Nulta hipoteza 170
5.5. Bajesov faktor i testiranje statističkih hipoteza 191
5.6. Parametrijski i neparametrijski statistici 194
Poglavlje 6. Korelacije 203
6.1. Povezanost između varijabli 203
6.2. Vrste povezanosti između varijabli 206
6.3. Sketergram 213
6.4. Koeficijent korelacije 217
6.5. Vrste koeficijenata korelacije 222
6.6. Hajde da primenimo šta smo do sada naučili! 234
Poglavlje 7. Statistički testovi za poređenje dva uzorka 241
7.1. Zavisni i nezavisni uzorci 242
7.2. Poređenje dve aritmetičke sredine – t test 246
7.3. Poređenje centralnih tendencija ordinalnih podataka na nezavisnim uzorcima – Men-Vitnijev U test i Vilkoksonov test sume rangova 250
7.4. Poređenje centralnih tendencija zavisnih uzoraka – test znaka i Vilkoksonov test rangova sa predznakom 252
7.5. Poređenje standardnih devijacija/varijansi – Levenov test. 257
7.6. Poređenje dve distribucije – Kolmogorov-Smirnov test, Hi kvadrat, Vald-Volfovic test 258
7.7. Hajde da primenimo šta smo do sada naučili! 267
Poglavlje 8. Vežbe – hajde da primenimo šta smo naučili u ovoj knjizi! 275
Reference 285

PREPRINT

Pre nego što krenemo dalje: Opšta uputstva za vežbe

Cilj ove knjige nije samo da upozna čitaoca sa osnovnim aspektima savremene statistike već i da pomogne čitaocu da nauči da valjano čita i interpretira rezultate statističkih proračuna koji se sreću u naučnoj literaturi. Iz ovog razloga, na kraju svakog poglavlja ove knjige (osim prvog), stavili smo vežbe koje pružaju čitaocu priliku da primeni statistička znanja koja su pokrivena u knjizi do tog mesta. Tu je i poslednje poglavlje koje se sastoji isključivo od takvih vežbi. Ove vežbe se sastoje od isečaka iz pravih naučnih publikacija, kao i iz priča u kojima se opisuju istraživački planovi. O ovima ćemo zajedno govoriti kao o isečcima. Svaki isečak praćen je određenim brojem tvrdnji koje se odnose na njegov sadržaj, a koje čitalac treba da kategoriše u različite kategorije na osnovu njihovih svojstava i toga u kakvom su odnosu sa sadržajima isečka.

Molimo vas da pažljivo proučite svaki isečak, a da potom pražljivo pročitate svaku tvrdnju i da je svrstate u jednu od četiri kategorije:

1 – Tačno – ako se može videti ili zaključiti iz isečka da je tvrdnja tačna.

2 – Netačno – ako je tvrdnja smislena (vidite opis besmislenih tvrdnji), ali se može zaključiti iz isečka da tvrdnja nije tačna.

3 – Besmisleno – ako je tvrdnja besmislena tj. ako:

- savremeni statistički postupci i naučni metodu ne mogu, čak ni teorijski, proizvesti dati rezultate koji bi mogli da je potvrde i/ili
- ako je tvrdnja logički nemoguća ili gramatički besmislena i/ili
- ako se u tvrdnji pominje neki nepostojeće ili besmisleni statistički pojam ili koncept
- ako tvrdnja tvrdi da neki postojeći/smisleni statistik ima neku nemoguću vrednost (vrednost koju ne može da ima)

4 – Nepoznato / ne može se zaključiti iz dostupnih podataka – ako je tvrdnja smislena tj. takva da postojeći statistički i naučni postupci mogu (makar i samo teorijski) dati rezultate koji bi je potvrdili, ali u isečku nema o tih podataka ili podataka o tome i stoga se ne može na osnovu isečka (ili samo na osnovu isečka) zaključiti da li je tvrdnja tačna ili ne.

Sledeća pravila će takođe važiti za sve vežbe:

PREPRINT

- Tvrdnje neće sadržati smislene-postojeće statističke pojmove koji nisu objašnjeni u ovoj knjizi. Eventualni izuzeci od ovog pravila će biti objašnjeni u opisu isečka.
- Osim ako suprotno nije očigledno iz podataka datih u isečku, treba pretpostaviti da su svi uslovi za korišćenje postupka koji je predstavljen u isečku ispunjeni tj. da su predstavljeni statistički postupci upotrebljeni ispravno.
- Statistički značajno, ako nije eksplicitno navedeno drugačije, znači da treba koristiti 0,05 kao kritični prag.
- Osim ako suprotno nije eksplicitno navedeno u određenim tvrdnjama, kada se govori o varijablama u isečku (na tako uopšten način), varijabla koja deli uzorak u grupe neće biti uključena u te varijable (ako takva varijabla postoji tj. ako je uzorak podeljen u grupe), čak n i ako je njeno ime dato u isečku.
- Ako suprotno nije eksplicitno navedeno, treba pretpostaviti da su vrednosti varijable pozitivno skorovane (tako da viši skorovi pokazuju više nivoe svojstva ili konstrukta na koji se varijabla odnosi).

Ovo je opšte uputstvo koje važi za sve vežbe ovog tipa u svim delovima knjige. Tačni odgovori i kratka objašnjenja svakog odgovora su dati nakon svake vežbe. Možete ih pročitati da bi videli koliko ste uspešno prošli vežbe i razjasnili eventualne greške.

PREPRINT

Poglavlje 1. Naučne paradigme i naučna objašnjenja, pseudonauka

Apstrakt. Ovo uvodno poglavlje predstavlja delove konceptualnog okvira na kome se zasniva upotreba statistike kao alata nauke. Poglavlje počinje razmatranjima toga šta je nauka, predstavljanjem koncepta naučne paradigme – skupa neproverenih i često neproverljivih pretpostavki koje čine konceptualni okvir naučnih istraživanja. Nakon toga sledi kratko razmatranje najznačajnijih pretpostavki aktuelne naučne paradigme i diskusija o tome kako one oblikuju funkcionisanje nauke. Pretpostavke o kojima se govori su pretpostavka o homogenosti prostorvremena, koncept objektivnosti, oslanjanja na čula, pitanja kompleksnosti našeg univerzuma i druge, a razmatrana je i uloga statistike u nauci koja je zasnovana na ovim pretpostavkama. Posebno podpoglavlje je posvećeno predstavljanju i diskusiji koncepata determinističkog i stohastičkog/indeterminističkog univerzuma jer su ovi koncepti direktno povezani sa upotrebom statistike u naučnim istraživanjima. Sledi deo koji se bavi vrstama naučnih objašnjenja. Predstavljene su različite vrste naučnih objašnjenja, uključujući statističko objašnjenje i zahtevi i karakteristike svakog od tih objašnjenja su razmatrani. Naredni deo je posvećen fenomenu samoispunjujućih proročanstava i pojavi kvarenja statističkih indikatora. Ova dva fenomena su razmatrana jer je važno uzeti ih u obzir kad god se oslanjamo na statističke generalizacije za donošenje odluka, tj. kada god odluke donosimo na osnovu rezultata statističkih analiza. Poslednji deo poglavlja upoznaje čitaoca sa fenomenom pseudonauke i pitanjem razlikovanja pseudonauke od validne nauke.

Ključne reči: naučna paradigma, determinizam, stohastički procesi, kvarenje statističkih indikatora, pseudonauka.

PREPRINT

U savremenom svetu, statistiku koristimo na puno različitih načina. Svaki put kada treba da opišemo neki veliki skup podataka ili kada treba da izvedemo zaključke o temama koje uključuju veliki broj slučajeva, oslanjamo se na statistiku. Sama **statistika je definisana kao nauka, oblast matematike, koja se bavi skupljanjem i analizom velikih skupova podataka.** Ona nam omogućava da opišemo takve skupove i da izvodimo zaključke o njima i na osnovu njih. Kao deo nauke, dobra praksa statistike se oslanja na adekvatnu primenu naučnih metoda, kao i na

razumevanju uslova za primenu statističkih metoda i njihovih ograničenja. Ovo je razlog zašto je prvo poglavlje ove knjige, poglavlje čiji je cilj da upozna čitaoca sa upotrebom statistike, posvećeno razmatranju načina na koji funkcioniše nauka i razmatranju mesta statistike u njoj. U ovom poglavlju ćemo povući i granicu između dobre naučne prakse i zloupotreba i loših postupaka koji pokušavaju da se okoriste o dobru reputaciju nauke, ali uopšte nisu ni nauka niti su naučni.

1.1. Nauka, naučne paradigme i statistika

Nauka se može definisati kao **sistematski poduhvat koji gradi i organizuje znanja u obliku proverljivih objašnjenja i predviđanja o univerzumu** (*Science (80-.)*., 2020). Nauka, kao svoj glavni rezultat, stvara **naučna uopštavanja ili generalizacije, naučne zakone ili teorije koji se svi mogu proveriti**. U idealnom slučaju, ovi proizvodi nauke su zasnovani na tumačenju odnosa između naučnih opažanja tj. empirijski ustanovljenih činjenica prikupljenih u postupku naučnog istraživanja (empirijski znači da su podaci dobijeni opažanjem i da se na takav način mogu proveriti, nasuprot toga da su nastali zaključivanjem, izvedeni iz neke teorije ili prosto izmišljeni). Međutim, to šta tačno predstavlja naučno opažanje, a šta ne tj. šta se može smatrati empirijski ustanovljenom činjenicom, kao i to gde će i kako naučnici tražiti naučna opažanja i koje će tačno vrste opažanja tražiti je nešto što definiše, i to često implicitno, naučna paradigma na koju se naučnici oslanjaju. **Naučna paradigma je filozofski i teorijski okvir određene škole naučnog mišljenja ili naučne discipline koja služi kao osnova svih teorija formulisanih u okviru te discipline, svih naučnih zakona i generalizacija, ali je i okvir za organizaciju naučnih opažanja.** Ovaj pojam je u filozofiju nauke ušao tokom 20. veka (npr. Walker, 2010; Wray, 2011), a njegovoj popularizaciji je veoma doprineo Tomas Kun (Thomas Kuhn) koji je naučnu paradigmu definisao kao "univerzalno prihvaćeno naučno dostignuće koje, tokom određenog vremena, obezbeđuje modele problema u rešenja zajednici praktičara (naučnika koji se praktično bave naukom, prim. prev)" (Kuhn, 1970, p. viii). Paradigma je obično skup verovanja koja su zajednička za naučnike u određenoj oblasti, a koja se obično ne mogu uopšte proveriti, iako ta verovanja definišu i vrste naučnih opažanja koje će naučnici tražiti, ali i to šta će iz tih opažanja zaključiti i kako. Ova verovanja su obično implicitna i naučnici ih vrlo često uzimaju zdravo za gotovo, smatrajući ih za jedini validni pogled na svet. Zbog toga što su same paradigme najčešće

PREPRINT

neproverene i neproverljive, a uprkos tome su, ipak, osnova svih naučnih znanja, veoma je važno da naučnici i ljudi koji rade sa naučnim podacima budu svesni pretpostavki paradigme ili paradigmi na koje se oslanjaju, kako bi mogli da prepoznaju karakteristike i organičenja naučnih znanja s kojima rade. Kun je verovao, a mnogi savremeni filozofi nauke će se sa tim složiti (Kuhn, 1970; Walker, 2010; Wray, 2011) da paradigme imaju tendenciju da se menjaju vremenom i ovo se tipično dešava onda kada više nije moguće ignorisati naučna zapažanja koja ne mogu da budu adekvatno objašnjena u okviru postojeće paradigme, a kada se, istovremeno, pojavi druga paradigma koja može dati bolja objašnjenja ovih zapažanja. Vremena u kojima nauka radi u okviru paradigme, Kun je nazvao **periodima "normalne nauke**", dok je kratke periode tokom kojih dolazi do promene paradigme nazvao **"naučnim revolucijama"**. Kun je verovao da se istorija nauke može opisati kao smena dugih perioda "normalne nauke" i kratkih i dinamičnih perioda "naučnih revolucija". Nakon što se naučna revolucija desi, ponovo nastupa period normalne nauke u kom nauka nastavlja da radi u okviru te nove paradigme.

A koje su pretpostavke sadašnje naučne paradigme? Većina naučnih polja ima svoje sopstvene skupove pretpostavki koje čine specifične paradigme u datim poljima. Ovo je posebno istinito ako uzmemo u obzir Kunovu definiciju paradigme kao skupa prototipnih problema i rešenja, jer su ovi zaista specifični za pojedinačne naučne oblasti. Ova specifičnost paradigme u pojedinačnim naučnim oblastima je zapravo ono što je dovelo do Kunovog otkrića naučnih paradigmi. Kun navodi da je shvatio da su paradigme važne onda kada je uočio da postoje velike razlike između naučnika u oblasti društvenih nauka u odnosu na pretpostavke o prirodi naučnih problema kojima su se bavili (Kuhn, 1970). Međutim, ako situaciju pogledamo šire, možemo primetiti da **pored elemenata paradigme koji su specifični za svako naučno polje, postoje i (neproverene i neproverljive) pretpostavke na kojima je nauka u celini zasnovana i koje su zajedničke za različite naučne oblasti**. Ovde ćemo probati da damo kratki pregled nekih od tih pretpostavki, a posebno onih koje su najvažnije za razumevanje praktične upotrebe statistike koje će biti predstavljeno u ovoj knjizi. S druge strane, čitaoci treba da budu svesni da, kako su elementi paradigme tipično implicitni, ovaj spisak ne treba posmatrati kao iscpnu listu svih zajedničkih karakteristika savremene naučne paradigme. Evo liste:

- **Postoji objektivna realnost van našeg uma i mi tu realnost možemo adekvatno opaziti koristeći naša priznata čula**;

PREPRINT

- **Određeni fenomen** je deo objektivne realnosti tj. **stvarno postoji, samo ako ga može opaziti više osoba**. Ako fenomen može da opazi samo jedna osoba i niko više, taj fenomen nije stvaran;
- Naš **um je odvojen od objektivne realnosti i ne utiče na nju, osim kroz aktivnosti naših naučno priznatih organa**;
- **Prostorvreme je homogeno** – **ni mesto u prostoru ni vreme nisu faktori ničega sami za sebe**. **Svi naučni zakoni podjednako rade u svim tačkama vremena i prostora**. **Zakoni prirode se ne menjaju sa prostorom i vremenom, nego rade svuda i sve vreme podjednako**;
- **Funkcionisanje univerzima je moguće u potpunosti opisati skupom pravila koji se dovoljno jednostavan da može stati u naš univerzum**;
- **Vreme je posebna dimenzija – kroz vreme je moguće kretanje samo unapred**, od prošlosti ka budućnosti, za razliku od prostornih dimenzija po kojima je moguće kretanje u oba smera.
- **Uzročno-posledični odnosi funkcionišu samo u jednom vremenskom smeru**, ne funkcionišu u suprotnom – **prošli događaji mogu uticati na buduće događaje, ali sadašnji ili budući događaji ne mogu uticati na prošle događaje**. Sadašnjost i budućnost ne mogu promeniti prošlost;
- **Fenomeni koje posmatramo imaju uzroke** (koji dovode do njihovog nastanka) **i u principu je moguće ustanoviti šta su ti uzroci.**
- Itd.

Objektivna realnost van našeg uma. U svojoj osnovi, naučna istraživanja su zasnovana na opažanjima tj. opservacijama, a opažanja su smislena samo ako postoji nešto što bi bilo opaženo i ako postoji način da se to opazi. Zbog toga je **neophodna pretpostavka takvog poduhvata to da postoji nešto van našeg uma što se može opaziti tj. objektivna realnost**, kao i to da su naša čula odgovarajući alati za opažanje i otkrivanje karakteristika te spoljašnje realnosti. Međutim, kao što pokazuje veliki broj radova u oblasti filozofije nauke, ali i interesovanje za ovu temu u oblasti umetnosti, mi ne raspolažemo zaista neospornim dokazima da je ono što opažamo stvarno objektivna realnost van našeg uma opažena našim čulima, a ne rezultat nečeg drugog što bi takođe

PREPRINT

moglo da stvori tok svesnih iskustava u našem umu, poput simulacije, proizvoda našeg uma ili nečeg drugog čega trenutno nismo svesni. Ovu situaciju verovatno najbolje opisuju filozofski stav solipsizma i intenzivne rasprave koje je solipsistički pogled na svet izazivao tokom istorije (e.g. Baldwin & Bell, 1988; Fodor, 1991). Da stvari budu još komplikovanije, **mi ni ne smatramo sva opažanja koja doživimo opažanjima spoljašnje realnosti**. Svi mi imamo opažaje koje doživljavamo, ali koje ne smatramo opažanjima spoljašnje realnosti. Takvi opažaji uključuju snove, halucinacije, "vizije", proizvode mašte, ali takođe i netačna opažanja. Kada se osoba bavi naučnim istraživanjima, važno je da može da razlikuje opažaje koji su posledica opažanja spoljašnje realnosti i one koji to nisu, a važno je i da se naučnici međusobno slažu oko kriterijuma za razlikovanje ove dve vrste opažaja. Iz tog razloga, trenutna "konvencija" u nauci je da se samo opažaji dobijeti kroz upotrebu čulnih organa koji su deo ljudskog tela, organa koje je nauka proučila, priznaje ih za čulne organe i za koje razume kako rade, mogu smatrati opažanjima. Ovi organi se mogu oslanjati na upotrebu različitih instrumenata, ali da bi određeni fenomen bio predmet naučnog opažanja, njegovo opažanje na kraju mora biti izvršeno priznatim senzornim organom. Na primer, naučnici mogu koristiti radar ili mikroskop ili spektrometar ili neki drugi naučni instrument kao pomoć, ali njihove rezultate na kraju moraju očitati koristeći oči, sluh, taktilni osećaj ili neki od ostalih senzornih organa ljudskog tela. Oni opažaji koji nisu dobijeni korišćenjem priznatih senzornih organa ne smatraju se opažanjima i ne mogu biti osnova naučnog rada. Pri ovom naglašavamo reči "priznata" čula tj. senzorni organi, jer kroz istoriju, postojala su različita viđenja toga koji metodi dolaženja do opažanja se mogu smatrati validnim. Tokom istorije je bilo mesta i vremena kada su, na primer, halucinatorna iskustva dobijena pod uticajem određenih droga smatrana za izvore valjanih opservacija (npr. Wallace, 1959). Takođe, kako je nauka napredovala i kako su se znanja o funkcionisanju našeg senzornog aparata uvećavala, tako se povećavao i broj naših priznatih čula – ljudi su verovatno oduvek bili svesni toga čemu služe oči i uši, ali su funkcije nekih drugih organa, poput sistema za propriocepciju (opažanje kretanja i pozicije tela i njegovih delova) ili poput funkcije unutrašnjeg uveta, otkriveni relativno skoro. Zbog ovoga, broj senzornih organa koje nauka priznaje povećavao se sa napretkom nauke. **Implicitna pretpostavka nauke je da ako se opažanje nije desilo putem priznatog čula, ono nije validno, nije realno**. Samo opažanja koja su dobijena putem (priznatnih) čula mogu ući u razmatranje da se smatraju validnim.

PREPRINT

Princip intersubjektivne saglasnosti. Posebna tema je pitanje toga šta se dešava kada se načini na koji dve različite osobe opažaju isti fenomen razlikuju. Ili šta se dešava kada fenomen uspeva da opazi jedna osoba, ali ne i druga? **Najprostije pravilo za razlikovanje valjanih opažanja od netačnih opažanja, halucinacija i sličnih fenomena je ono koje kaže da će validna opažanja do kojih dođu različite osobe biti slična**. Fenomeni koje jedna osoba tvrdi da je opazila, ali koje drugi ne uspevaju da opaze, iako su u poziciji da to mogu da urade, nisu stvarni. Pretpostavka koja stoji u osnovi nauke je da ako je nešto „realno", „stvarno" tj. ako postoji u fizičkom svetu, onda će različite osobe moći da to nešto opaze. Ako to nešto ne može opaziti niko osim jedne pojedinačne osobe ili određene specifične grupe ljudi, tada ne možemo da budemo sigurni da li se radi o tome da ti ljudi prosto izmišljaju, da li su nešto opazili pogrešno, da li imaju halucinacije ili neki sličan fenomen ili šta je već u pitanju, ali naučna pretpostavka o tome šta se tu dešava je da nešto nije u redu sa njihovim opservacijama. Naučno validnim se mogu smatrati samo fenomeni koje može opaziti više različitih osoba, tj., u idelnom slučaju, svi oni koji su zainteresovani da ih opaze (naravno, pretpostavka je da oni urade šta je potrebno da bi to opažanje bilo moguće – koriste potrebnu opremu ili instrumente ako su potrebni, dođu na potrebno mesto i vreme sa kog se to može opaziti isl.). Ovaj stav naziva se **principom intersubjektivne saglanosti** i on predstavlja suštinski deo naučne objektivnosti i replikabilnosti – ponovljivosti naučnih nalaza, a obe ove stvari su ključne komponente dobro sprovedenog naučnog rada.

Međutim, koliko god da je princip intersubjektivne saglasnosti praktično koristan, pretpostavka da stvarno postoje samo fenomeni koje može da opazi više osoba stvara svoj posebni skup problema za nauku. Prema ovom principu, **fenomeni koji su se desili samo jednom, tj. jedinstveni – jednokratni fenomeni nisu naučno opazivi.** Drugim rečima, takvi fenomeni, fenomeni koji su se desili samo jednom i nikad više ne bi bili smatrani realnim. Ista je situacija sa **fenomenima koje može da opazi samo jedna osoba, jer ni jedna druga osoba nije u poziciji da ih opazi.** Ali da li ovakvi (jednokratni, jedinstveni) fenomeni zaista postoje? Iako se itekako može raspravljati o tome da li su različiti jedinstveni – jednokratni događaji o kojima različiti ljudi govore realni ili izmišljeni, u istoriji nauke ima puno primera fenomena koje je opazila i o kojima je izvestila samo jedna osoba i koji su u početku smatrani izmišljotinama ili prevarama, da bi se tek kasnije pokazalo da su ti fenomeni realni, nakon što su se pojavili načini da ih drugi članovi naučne zajednice opaze. Relativno svež primer ovakve situacije je fenomen kuglastih munja, prirodni fenomen koji su mnogi veoma dugo smatrali za izmišljotinu ili čak prevaru uprkos postojanju više izveštaja o

PREPRINT

javljanju ovog fenomena kroz istoriju. Kuglaste munje smatrane su manje-više za izmišljotinu sve do trenutka kada je primerak kuglaste munje kamerom i drugim naučnim instrumentima snimila grupa naučnika (Cen et al., 2014). Međutim, fenomen kuglaste munje je zapravo fenomen koji je redak, ali nije jedinstven niti jednokratan. Kuglaste munje se po svoji prilici javljaju retko, ali se javljaju. Da je kuglasta munja zaista bila jedinstveni fenomen, nešto što se desilo jednom i nikad više ponovo, vrlo je verovatno da njeno postojanje nikada ne bi postalo deo naučnoh znanja.

Što se tiče **fenomena koji su opazivi samo za jednu osobu i ni za kog više**, jednu ogromnu i veoma važnu klasu tih fenomena **čine svesna iskustva ili kvalije**, kako ih neki istraživači nazivaju (e.g. Jackson, 1982). Iako smo svi verovatno svesni svoje sopstvene svesti tj. svojih sopstvenih ličnih iskustava onako kako ih sve vreme doživljavamo, ta iskustva predstavljaju kategoriju fenomena koje, u ovom trenutku razvoja nauke, ne može da opazi direktno niko osim nas samih. Naravno, mi pretpostavljamo da svi imaju svesna iskustva onako kako ih mi imamo, ali niko ne može direktno da opazi svesna iskustva bilo kog drugog osim sebe. Drugim rečima, **iako ja znam da ja imam svesna iskustva jer ih opažam sve vreme, kako da znam da ih i drugi ljudi imaju**? Ja ne mogu da direktno opazim svesna iskustva bilo kog drugog osim sebe. A kad smo već kod toga, ko i šta tačno sve ima svesna iskustva? Samo ljudi? Da li ih imaju i životinje? Biljke? Neživi objekti? Ovaj problem poznat je u naučnoj literaturi kao **„problem drugih umova“,** a veoma lepa diskusija aktuelnih znanja i shvatanja o ovoj temi može se naći u knjizi Marka Solmsa „Skriveni izvor“ (Solms, 2021). Svesna iskustva kao kategorija fenomena dolaze u konflikt sa pretpostavkom nauke koju smo ovde opisali i to je verovatno glavni razlog zašto je naučni status svesnih iskustava jako dugo bio upitan i tema brojnih oštrih diskusija, a takva situacija je i danas. Šta više, **iako su svesna iskustva nešto što verovatno svako od nas opaža sve vreme, za postojanje ove klase fenomena sadašanja nauka nema ama baš nikakvo objašnjenje**. Zapravo, **prema prihvaćenim biološkim i fizičkim modelima , svesna iskustva ne bi trebalo da postoje uopšte, jer ne postoji ni jedna naučna teorija ili objašnjenje koje bi uopšte mogli da objasne kako nastaje i funkcioniše tačno fizički proces kojim nastaju i prenose se svesna iskustva**. Kako je to lepo formulisao David Čalmers u svojoj knjizi iz 1996. koja se bavila ovom temom: „Ne samo da nam nedostaje precizna teorija; mi ne znamo ni kako bi uopšte teorija svesti trebala da izgleda (We do not just lack a detailed theory; we are in the dark about what a theory of cosciousness would even look like)“ (Chalmers, 1996, p. ix).

PREPRINT

Problem svesti je, s druge strane, važna tema filozofije, još od nastanka filozofije. Filozofi su razmatrali i razmatraju i dalje ovu temu pod različitim imenima poput – odnos tela i uma, odnos tela i duše, teški i laki problemi svesti i drugim. Stanovišta koja su zauzimali istraživači po ovom pitanju veoma su se razlikovala kroz istoriju. I među današnjim teoretičarima možemo videti one koji problem svesnih iskustava smatraju za važnu naučnu temu, ali i one koji negiraju da svesna isustva uopšte postoje. A postoje i oni među naučnicima koje se prave da nemamo „slona na sred sobe“ i da ovaj problem uopšte ne postoji. Ovo se takođe održava i na status svesnih iskustava u različitim društvenim naukama. S jedne strane, svesna iskustva se uzimaju praktično zdravo za gotovo u oblasti prava i pravnih nauka. U ovoj oblasti ne samo da se postojanje svesnih iskustava priznaje, već se ta iskustva, bez ikakve zadrške, tretiraju kao uzroci ljudskog ponašanja i koriste se razne praktične metode za zaključivanje o svesnim iskustvima koje osoba ima. S druge strane, istorija nauka poput psihologije se u velikoj meri sastoji od ozbiljnih debata o statusu i samom postojanju svesnih fenomena (e.g. Chomsky, 1959; Watson, 1913). Ovaj konflikt između prirode svesnih iskustava i naučnih zahteva za intersubjektivnom saglasnošću je glavni razlog zašto su različiti autori u različitim trenucima kroz istoriju osporavali to da su društvene nauke, a posebno psihologija, uopšte nauke. Međutim, tu treba primetiti da iako naučnici ni danas nemaju iole validnu teoriju koja bi objasnila postojanje svesnih doživljaja, to ne znači da u budućnosti neće biti moguće stvoriti takvu teoriju, kao ni to da će zauvek ostati nemoguće da druge osobe opaze lična svesna iskustva jedne osobe. U trenutku kada je ova knjiga pisana, takve mogućnosti su tema dela naučne fantastike, iako brojni istraživači iz različitih naučnih oblasti prepoznaju postojanje ovog problema i rade na tome da ponude moguća naučna objašnjenja ovog fenomena (e.g. Atmanspacher, 2004; Gao, 2008; Hunt & Schooler, 2019). Vrlo je čvrsto uverenje autora ove knjige da će otkriće fizičkih procesa koji rezultiraju nastankom svesnih događaja, kada se desi, kao i valjano rasvetljavanje ostalih problema vezanih za svesna iskustva, počev od problema drugih umova, biti pokretači velike naučne revolucije, te da će buduća istorija nauke pamtiti upravo tu revoluciju kao najvažniju naučnu revoluciju u dotadašnjoj istoriji nauke. Međutim, kada će se to desiti u ovom trenutku ne možemo da predvidimo.

Naš um utiče na spoljašnji svet samo kroz naučno priznate delove našeg tela. Ovaj princip kaže da na objektivnu realnost možemo uticati jedino kroz interakciju delova našeg tela sa tom spoljašnjom realnošću. Ako hoćemo da razgovaramo sa nekim, moramo da za to iskoristimo odgovarajući deo tela kako bi napravili poruku koju ta druga osoba može da opazi i razume. Na

PREPRINT

primer, moramo da koristimo naše glasne žice da bi proizveli zvuke koji čine govor ili moramo da koristimo prste da otkucamo poruku na kompjuteru. Ako hoćemo da pomerimo neki predmet, moramo za to da iskoristimo ruke ili neki drugi deo tela da taj predmet podignemo ili pomerimo ili da upravljamo opremom i instrumentima koji će predmet pomeriti. Ovo takođe znači i da su sve navodne interakcije sa spoljašnjim svetom koje se ne odvijaju preko naših naučno priznatih delova tela naučno nemoguće. Na primer, telepatija i telekineza su, prema ovom shvatanju, nemoguće zato što mi nemamo ni jedan organ niti deo tela koji može da proizvede efekte od kojih se ovi navodni fenomeni sastoje. Ovaj princip takođe znači i da naše misli same za sebe nemaju nikakav efekat na spoljašnji svet, ako na osnovu tih misli ne uradimo nešto koristeći svoje telo. To što ćemo nekome samo poželeti dobro ili loše neće proizvesti nikakav efekat, ako ne uradimo ništa dodatno da te svoje želje sprovedemo u delo. Zadatak o kome stalno mislimo se neće uraditi ako neko ne ode i uradi ta. Verovanje da naše misli same za sebe mogu uticati na spoljašnji svet – objektivnu realnost je vrsta magijskog mišljenja i ovaj princip je direktno suprotan takvim verovanjima.

Homogenost prostorvremena. Jedna od ključnih postavki naučnog metoda je zahtev za replikabilnošću tj. ponovljivošću naučnih opažanja. Nalazi nauke moraju biti ponovljivi tj. ako jedna osoba izvesti da određene radnje pod određenim uslovima dovode do određenih rezultata, drugi ljudi koji urade iste radnje pod istim uslovima moraju dobiti iste rezultate. Udžbenici iz naučne metodologije redovno navode da **vreme i mesto, sami za sebe, nisu faktori koji mogu uticati na događaje i da se ne mogu smatrati uzrocima bilo čega**. Uzročni faktori mogu biti različiti procesi koji se dešavaju kroz vreme, ali ne vreme samo za sebe. Isti je i slučaj sa mestom. Sva ova očekivanja zasnovana su na verovanju da je **prostorvreme homogeno tj. da su sve tačke u prostoru i vremenu jednake i da univerzum funkcioniše jednako i po istim pravilima svuda i sve vreme**. Iako većina naučnika veoma čvrsto veruje u ovo, valja primetiti da ova pretpostavka (do sada) nije ozbiljnije ili obimnije proveravana, niti za sada može biti ozbiljnije proveravana. Svi naučni podaci kojima trenutno raspolažemo prikupljeni su unutar vrlo kratkog vremenskog okiva (u poredjenju sa ukupnom starošću univerzuma i vremena koje će univerzum verovatno još postojati) i unutar veoma veoma malog dela prostora tj. na Zemlji i njenoj neposrednoj okolini, kao i u prostoru koji je tokom ovog kratkog vremenskog perioda zauzimala krećući se kroz svemir. A čak je i ovo rađeno samo na određenim tačkama Zemlje na kojima je sprovođeno prikupljanje specifičnih podataka (na primer, na mestima gde se vrše naučna opažanja i merenja). Uprkos ovome, naučnici obično veoma čvrsto veruju da je prostorvreme homogeno i da su svi naučni

PREPRINT

zakoni izvedeni iz naučnih opservacija jednako validni u svim tačkama prostora, da su bili i da će biti jednako validni tokom čitave večnosti, od početka do kraja svemira i ranije i kasnije. Ako pogledamo bilo koju naučnu teoriju ili generalizaciju, nećemo naći da i jedna od njih specifikuje tačne vremenske ili prostorne koordinate u kojima je validna. I zaista, bez pretpostavke o homogenosti prostorvremena, nauka kakvu poznajemo ne bi verovatno ni postojala i verovatno bi morala da bude veoma različita. Bez homogenosti prostorvremena, bilo kakva pravilnost koju bi naučnici uočili morala bi da bude ispitana u različitim tačkama prostorvremena. Ne bismo znali da li su pravilnosti koje danas opažamo postojale i ranije, kao ni da li će važiti i na drugim mestima. Ako bi iste radnje dovele do različitih rezultata u dva različita ispitivanja, mi ne bismo mogli da znamo da li je to zato što su zakoni prirode drugačiji u tačkama prostorvremena u kojima smo radili ispitivanja ili zato što smo napravili negde neku grešku ili zato što postoje neki drugi faktori koje nismo uzeli u obzir. Izvođenje bilo kakvih zaključaka o prošlim događajima bi bilo mnogo teže, a zaključci o dalekoj prošlosti ili dalekoj budućnosti ili o dalekim mestima bi bili potpuno nemogući ako prethodno ne bi stekli dovoljno detaljna znanja o tome kako se različite tačke prostorvremena razlikuju i i zašto. Bilo koja naučna teorija ili generalizacija o zakonima prirode bi u ovakvoj situaciji morala da uključuje i ograničenja o prostorvremenskim koordinatama u kojima se može smatrati validnom. Predviđanje budućnosti bi takođe bilo mnogo teže ili nemoguće, jer ne bi mogli da znamo da li će sadašnji zakoni prirode ostati isti i u budućnosti i koliko tačno dugo. Srećom, pretpostavka o homogenosti prostorvremena štiti nauku od svega ovoga i dozvoljava da se bilo koje pravilnosti koje opazimo uopšte (generalizuju) na sva vremena i sva mesta. Međutim, svaki naučnik bi trebao da stalno ima na umu da pretpostavka o homogenosti prostorvremena nije nešto što je nauka stvarno mogla u iole većem obimu da proveri, niti je stvarno može ozbiljno proveriti, bar koristeći trenutne mogućnosti civilizacije na planeti Zemlji.

Dešavanja u univerzumu određuje skup jednostavnih pravila. Jedan od načina na koji možemo opisati ciljeve nauke je da kažemo da je nauka sistematska aktivnost usmerena na to da se razumeju pravila po kojima priroda i univerzum funkcionišu. I zaista, ono što skoro svaki istraživač koji se bavi osnovnim naučnim istraživanjima radi, ma u kojoj oblasti nauke to bilo, može se dobro opisati kao serija pokušaja da se otkriju pravila po kojima funkcionišu fenomeni koje dati istraživač proučava. Ali koliko su zapravo kompleksna ta pravila koja tražimo? Da li su veoma, veoma prosta ili veoma komplikovana tj. kompleksna? Dosta poznat primer zagonetke ovog tipa imamo u slušaju fizike gde su Njutnove, relativno jednostavne, zakone kretanja zamenila dosta

PREPRINT

kompleksnija Ajnštajnova pravila relativnosti, jer ta pravila bolje opisuju i predviđaju opservirano funkcionisanje univerzuma i objekata u njemu. U oblasti društvenih nauka, možemo primetiti kako se kompleksnost objašnjenja koja naučnici daju menja. Na primer, psihologija je manje-više počela sa skupom modela koji su poznati kao mentalna hemija, a kojim su ti rani istraživači pokušali da u ovoj oblasti primene modele objašnjenja koji su slični modelima koji su bili popularni u hemiji tog vremena (e.g. Gentner & Grudin, 1985). Nakon njih su se pojavili psihodinamski modeli koji su psihičko funkcionisanje čoveka predstavljali na način koji je sličan funkcionisanju parnih kotlova i motora sa unutrašnjim sagorevanjem koji su bili popularni početkom 20. veka, dok u psihološkoj nauci u trenutku pisanja ove knjige, dominiraju modeli koji opisuju psihološko funkcionisanje preko sistema linearnih strukturalnih jednačina. Međutim, odabrani tipovi objašnjenja, bar u oblasti društvenih nauka, su do sada vrlo retko bili posledica suštinskog razumevanja prirode fenomena koji se opisuju i stoga retko kad izabrani zato što su istraživači smatrali da je baš ta vrsta modela najadekvatnija za objašnjavanje posmatranih fenomena. Mnogo češće, zapravo tipično, radilo se o tome da su društveni naučnici koristili one vrste modela koje su im bile lako dostupne – u 19. veku došlo je do naglog razvoja oblasti hemije, pa su društveni naučnici pokušali to da iskopiraju. Može se osnovano tvrditi da je sadašnja popularnost modela zasnovanih na linearnim strukturnim jednačinama (primenjenih kroz statističke postupke poput regresije, faktorske analize, strukturnog modelovanja itd.) zapravo motivisana samo time što je pravljenje ovakvih modela relativno lako korišćenjem dostupnog komercijalnog statističkog softvera u kojima su široko raspoložive različite statističke procedure zasnovane na linearnim strukturalnim jednačinama. Kao argument u prilog ovoj tezi ide i to da se pregledom postojeće naučne literature u oblasti društvenih nauka jedva uopšte mogu naći tekstovi u kojima se diskutuje ili objašnjava zašto su baš linearne strukturne jednačine odabrane kao osnova za modele koji se koriste za objašnjavanje društvenih fenomena ili zašto istraživači koji ih primenjuju baš takve modele smatraju najpogodnijim.

Razmatrajući ovaj problem nameće se i jedno očigledno pitanje – ako su ovi eksplanatorni modeli (modeli objašnjenja) u društvenim naukama izabrani samo zato što ih je lako napraviti, a ne zato što su najadekvatniji, kako bi izgledali najbolji mogući modeli za objašnjenje društvenih fenomena? Ovo pitanje je jednako validno i u svim drugim oblastima nauke i zahvaljujući njemu je nastala oblast nauke poznata kao **teorija kompleksnosti** (Byrne, 1998; Manson, 2001). Važna tema kod razmatranja kompleksnosti je pitanje toga šta je tačno kompleksnost i kako razlikovati

PREPRINT

različite nivoe kompleksnosti. Kako razlikovati koji modeli su više, a koji manje kompleksni? Veoma koristan doprinos ovoj temi je pojam algoritamske kompleksnosti, koja je takođe poznata i kao Kolmogorovljeva kompleksnost ili Kolmogorov-Chaitinova kompleksnost (Kolmogorov-Čejtinova kompleksnost) (Chaitin, 2005, 1974; Delahaye & Zenil, 2008). **Algoritamska kompleknost definiše se kao „najprostiji računski algoritam koji može proizvesti ponašanje sistema“** (Manson, 2001, p. 404). Rečeno jednostavnijim rečnikom, sistem čije se ponašanje može proizvesti ili simulirati kratkim algoritmom se smatra jednostavnim tj. prostim. S druge strane, sistem čije se ponašanje ne može reprodukovati niti simulirati ni na jedan način koji je jednostavniji od toga da prosto popišemo i precizno opišemo sve specifične događaje koji čine ponašanje tog sistema, smatra se kompleksnim. Na primer, ako bi naš sistem bio niz znakova poput:

ABABABABABABABABABABABABABABABABABABAB…. i tako dalje do beskonačnosti, mi bi mogli da ovaj sistem opišemo navodeći da je to niz znakova AB ponovljen beskonačan broj puta. Iako je niz koji opisujemo beskonačan, mi ga možemo proizvesti koristeći vrlo prosto pravilo. Ovakav niz je stoga prost tj. jednostavan. S druge strane, možemo da zamislimo i niz znakova koji je takav da ne prati nikakvo pravilo i da zbog toga ne postoji drugi način da se rekreira ovaj niz osim da se zapiše niz tačno takav kakav je. Ovakav niz znakova bi smatrali kompleksnim. Ako bi uporedili ova dva slučaja, možemo primetiti da su instrukcije koje stvaraju prvi niz mnogo kraće od samog niza. S druge strane, instrukcije za pisanje drugog niza su bar podjednako dugačke kao i sam niz, što taj niz čini kompleksnim, zapravo onoliko kompleksnim koliko sistem uopšte može da bude kompleksan.

Ako bi isti način rezonovanja pokušali da primenimo na funkcionisanje univerzuma, mogli bi da primetimo da je **krajnji cilj nauke da otkrije skup instrukcija koje opisuju funkcionisanje celog univerzuma i to u situaciji kada su i sama nauka i naučnici deo tog univerzuma**. Pitanje koje se logično postavlja je da li je to moguće. Ovo je zapravo pitanje **toga da li su pravila koja upravljaju našim univerzumom dovoljno jednostavna da instrukcije potrebne za njihovu realizaciju mogu da stanu u univerzum ili su ta pravila toliko kompleksna da su instrukcije potrebne za realizaciju funkcionisanja univerzuma previše velika da bi stala u univerzum.** U prvom slučaju, cilj nauke da se stvori savršeni model univerzuma je sprovodiv, dok u drugom slučaju nije. Ali kako da znamo koji je od ova dva slučaja u pitanju kada su u pitanju odnos nauke

PREPRINT

i funkcionisanja univerzuma? Drugim rečima, kako da znamo da li je naš univerzum kompleksan ili jednostavan? S jedne strane, možemo da primetimo da je nauka u mnogim oblastima do sada uspevala da prilično dobro predviti ponašanje velikih prirodnih sistema i to korišćenjem relativno prostih jednačina i matematičkih modela. Ovo može da znači da prava kompleksnost univerzuma nije blizu maksimalne moguće i da je verovatno da je prava kompleksnost univerzuma negde u nivou nižih nivoa kompleknosti. Ali koliko je zapravo naš univerzum jednostavan? Neki autori, kao na primer Volfram, izneli su ideju da bi prava kompleksnost univerzuma mogla biti veoma niska tj. da je naš univerzum zapravo veoma jednostavan i da se verovatno može rekonstruisati korišćenjem veoma prostih pravila (Wolfram, 1984). Međutim, ovaj pristup proceni kompleksnosti našeg univerzuma se nije do sada bitnije razvio dalje od ove opšte ideje i zbog toga je bio meta jake kritike (e.g. Chaitin, 2005; Manson, 2001). Iako su razmatranja pitanja kompleksnosti relativno nova u nauci uopšte, a i naučnici iz oblasti društvenih nauka vrlo retko diskutuju i razmatraju pravu kompleksnost fenomena koje proučavaju, možemo zaključiti da **naučnici implicitno veruju da su fenomeni koje proučavaju dovoljno prosti da postojanje nauke ima smisla**. Drugim rečima, veruju da je kompleksnost fenomena koje proučavaju dovoljno niska da se oni mogu rekreirati kroz model (ili skup instrukcija) koji je dovoljno mali da može stati u naš univerzum. Kada to ne bi bilo tako, nauka ovakva kakva je danas ne bi bila moguća. Imajući u vidu ovu situaciju, bilo bi dobro da naučnici, pogotovo u oblasti društvenih nauka počnu da redovno razmatraju verovatni nivo kompleksnosti pravila koja upravljaju fenomenima koje proučavaju, umesto da slepo grade nauku samo na osnovu modela koji su im dostupni u statističkom softveru koji koriste. Takođe treba pomenuti da koncept kompleksnosti već ima svoje prve primene u oblasti društvenih nauka (e.g. Gauvrit et al., 2017).

Vreme je posebna dimenzija – fenomeni se kroz vreme kreću samo unapred. Iako prihvatamo da je prostorvreme homogeno tj. da su sve tačke prostora i vremena iste i da u njima važe isti zakoni prirode, vreme je posebno u smislu da se fenomeni kroz njega mogu kretati samo unapred. Za razliku od dimenzija prostora kroz koje se fenomeni mogu kretati u svim smerovima, **u vremenu se objekti iz prirode mogu kretati samo unapred** (bar što se tiče naučnih istraživanja izvan oblasti fizike). Iz ovog razloga, **nauka priznaje postojanje samo onih uzročno-posledičnih odnosa u kojima uzrok vremenski prethodi posledici ili eventualno u kojima se uzrok i posledica dešavaju istovremeno**. **Uzroci ne mogu proizvoditi efekte u prošlosti, samo u budućnosti**. Većina epistemoloških rasprava o uzročno-posledičnim odnosima eksplicitno

PREPRINT

navodi zahtev da uzrok mora vremenski prethoditi posledici ili biti istovremen s posledicom kao jedan od ključnih uslova za utvrđivanje da li je određeni odnos između dve pojave zaista odnos uzroka i posledice. Drugim rečima, mi pretpostavljamo da prošlost može uticati na sadašnjost i budućnost, ali da budućnost ne može uticati na prošlost. Na primer, jedan od najistaknutijih filozofa modernog doba – Dejvid Hjum u svom čuvenom delu „Traktat o ljudskoj prirodi [A treatise of human nature]“ navodi da da bi neka pojava X mogla da bude uzrok pojave Y, X mora prethoditi pojavi Y u vremenu. Ovakvo shvatanje deli još jedan drugi čuveni filozof nauke – Džon Stjuart Mil [John Stuart Mill], kao i većina kasnijih naučnika koji su proučavali uzročnost. Ovaj princip je takođe i jedan od ključnih principa za utvrđivanje uzročnosti koji se može naći u savremenim udžbenicima metodologije naučnih istraživanja i literaturi koja opisuje eksperimentalne procedure (e.g. Milas, 2009). To je i glavni razlog zbog čega naučnici smatraju uspešnost u predviđanju budućih događaja mnogo boljim pokazateljom kvaliteta naučnih znanja nego od uspešnosti u objašnjavanju prošlih događaja.

Uzroci dostupni spoznaji, koji se mogu saznati. Pojave koje opažamo imaju uzroke koji dovode do njihovog javljanja i u principu je moguće saznati/otkriti šta su tačno ti uzroci. Ne samo da su pravila koja upravljaju našim univerzumom dovoljno jednostavna da je moguće dovoljno precizan model našeg univerzuma smestiti u naš univerzum, još jedna dodatna pretpostavka nauke je to da **smo mi u stanju da saznamo šta su uzroci pojava u prirodi i da smo sposobni da razumemo pravila koja upravljaju univerzumom**. Ovo je još jedno osnovno verovanje koje daje smisao naučnim istraživanjima. Kada ovo ne bi bio slučaj, ne bi bilo moguće da ljudi ikada otkriju i razumeju mreže uzroka i posledica koje čine univerzum, a naučna istraživanja usmerena ka ovom cilju bi bila uzaludna. Takođe, ovo znači i da **ako u ovom trenutku nismo u stanju da identifikujemo uzroke određene pojave, ovo verovanje zahteva da zaključimo da je to zbog toga što datu pojavu nismo dovoljno ili nismo adekvatno istražili, a ne zato što su uzroci ove pojave u principu nesaznatljivi**. To znači da samo treba da nastavimo da istražujemo i dalje tražimo uzroke i kada jednom otkrijemo pravi pristup istraživanju, pravu metodologiju istraživanja ili prave instrumente za merenje, otkrićemo i uzroke proučavane pojave, zato što su svi uzroci koji postoje u prirodi takvi da se mogu saznati. **Nema u prirodi pojava čiji se uzroci ne mogu otkriti.** U filozofiji, varijanta ovog principa je poznata kao **princip dovoljnog razloga** (eng. principle of sufficient reason) i definisana je kao verovanje da „sve ima razlog, uzrok ili osnovu“ ili drugim

PREPRINT

rečima „za svaku činjenicu F, mora postojati dovoljan razlog zašto je baš F činjenica [For every fact F, there must be a sufficient reason why F is the case.]" (e.g. Melamed & Lin, 2021).

1.2. Priroda uzročno posledične mreže univerzuma – stohasticizam naspram determinizma.

Jedan važan aspekt načina na koji gradimo naučno razumevanje sveta oko nas je pogled na prirodu uzročno-posledične mreže univerzuma tj. pogled na prirodu odnosa između uzroka i posledica. Ovaj pogled na prirodu odnosa između uzroka i posledica je nešto po čemu se različite naučne oblasti u praksi razlikuju i gde takođe možemo često videti da se dva ne sasvim kompatibilna pogleda na prirodu univerzuma naizmenično koriste u naučnom radu. Jedan od ovih pogleda na svet naziva se determinizam (e.g. Hoefer, 2016; James, 1884), a drugi, suprotan njemu, nazvaćemo stohasticizam.

Osnovna ideja determinizma, uzročnog determinizma ili „tvrdog determinizma", kako ga neki zovu, je da univerzum zapravo predstavlja veoma složenu mrežu uzroka i posledica, takvu da je svaki događaj u njemu u potpunosti određen skupom uzroka. Svaki od tih uzroka je pak potpuno određen nekim drugim skupom uzroka i ta mreža uzroka i posledica proteže se u večnost, kako ka budućnosti, tako i u prošlost prema samom početku univerzuma, ako je takva tačka u vremenu ikada postojala. Ovo znači da ako bi znali sve početne uslove koji su vladali u univerzumu i sve zakone prirode, bili bismo u stanju da potpuno precizno, bez ijedne greške predvidimo sve događaje koji će se desiti u univerzumu tokom celog njegovog postojanja. Pretpostavka da je priroda našeg univerzuma deterministička tj. deterministički pogled na svet ima neke vrlo važne implikacije:

- Kako je univerzum po prirodi deterministički mehanizam uzroka i posledica, koji je takav da uzroci potpuno određuju posledice, **naša nemogućnost da sa potpunom preciznošću predividimo bilo koji događaj je posledica našeg nedovoljnog znanja o zakonima prirode i uzrocima događaja**, a ne toga što je dati događaj u principu nepredvidiv.
- Iz istog razloga, **sve što će se ikada desiti u univerzumu određeno je uslovima koji su postojali na početku** (postojanja univerzuma) i **nema načina da se na taj tok događaja**

bilo kako utiče. Sudbina postoji, realna je i ne možemo je promeniti. Desiće se to što će se desiti i to se ne može promeniti.

- Ovaj pogled na svet takođe implicira da **nema slobodne volje** tj. da je slobodna volja samo iluzija jer je sve unapred određeno.
- Kada se uzme zajedno sa ostalim pretpostavkama nauke, pre svega sa pretpostavkom da je skup pravila koji upravlja našim univerzumom dovoljno jednostavan tj. da savršeni model našeg univerzuma može stati u univerzum, deterministički pogled na svet takođe implicira i da je **u principu moguće napraviti teoriju svega**, savršeni model univerzuma koji će potpuno precizno da opiše i objasni funkcionisanje univerzuma.

Pitanje toga da li je priroda univerzuma deterministička je vekovima bilo predmet filozofskih razmatranja. Filozofi su razmatrali različite varijante determinističkog pogleda na svet koji je uključivao i stvari poput „mekog determinizma“ koji bi dozvolio postojanje slobodne volje kao kontrast determinizmu koji je ovde predstavljen, a koji bi bio nazvan „tvrdi determinizam“ (e.g. James, 1884). Naučnici van oblasti filozofije veoma retko otvoreno pričaju ili se uopšte izjašnjavaju o tome kakva veruju da je priroda univerzuma kada formulišu svoje naučne teorije i objašnjenja. Međutim, iz toga kako su teorije tipično formulisane, posebno u oblasti prirodnih i društvenih nauka, može se zaključiti da većina naučnika u ovim oblastima smatra da deterministički pogled na svet valjano opisuje prirodu našeg univerzuma, bez obzira da li su dati naučnici svesni postojanja i detalja ove filozofske rasprave ili ne. Valja takođe naglasiti i da **u determinističkom univerzumu nema mesta za slučajne događaje**, pa samim tim **nema mesta ni za pretpostavke na kojima se zasniva statistika kao nauka**, kao što ćemo kasnije videti.

Osnovna ideja stohasticizma je da je priroda našeg univerzuma stohastička što znači da stohastički procesi mogu da postoje i realno postoje. Stohastički procesi su procesi gde isti skup uzorka pod istim skupom uslova može dovesti do više različitih ishoda od kojih svaki ima određenu verovatnoću javljanja. **Stohastički univerzum je univerzum u kom postoje pojave koje se ne mogu precizno predvideti čak ni kada znamo sve uzroke i sve zakone prirode tj. pojave koje su u principu nepredvidive**. To je univerzum u kom događaji nisu unapred određeni i u kome poznavanje početnih uslova univerzuma ne omogućava precizno određivanje toga kakvi će biti uslovi na kraju. Ideja da je priroda našeg univerzuma stohastička ima takođe neke važne implikacije:

PREPRINT

- **Nemogućnost dovoljno preciznog predviđanja neke pojave ne znači nužno da mi tu pojavu nismo dovoljno upoznali**. Sasvim je moguće da je priroda pojave koju ne uspevamo da predvidimo takva da se ona, u principu, ne može predvideti i to zbog toga što, u stohastičkom univerzumu, isti procesi mogu dovesti do različitih ishoda, prosto jer je priroda univerzuma takva.
- Iz istog razlog, **način na koji se događaji u univerzumu odvijaju se ne može predvideti sa potpunom preciznošću ili se ne može predvideti uopšte**. Kako isti skup uzroka može voditi različitim posledicama, nećemo moći sa potpunom sigurnošću znati koji ishod će se desiti u nekom pojedinačnom slučaju (možemo samo da govorimo o verovatnim ishodima). Poznavanje početnih uslova univerzuma i svih zakona prirode nije dovoljno da omogući predviđanje konačnih uslova na kraju univerzuma niti šta će se sve tačno desiti u međuvremenu između početka i kraja svemira. Univerzum je u osnovi nepredvidiv ili bar nepotpuno predvidiv. Ovo takođe znači i da **ne postoji sudbina i da ishodi događaja nisu predodređeni.**
- U stohastičkom univerzumu je **moguće da postoji slobodna volja u nekom obliku**, zavisno već od tačne prirode tih stohastičkih procesa. Međutim, može se zamisliti i stohastički univerzum u kome postoji nepredvidivost bez slobodne volje! Ovo u velikoj meri zavisi od toga kakva bi priroda slobodne volje mogle da bude, jer je slobodna volja fenomen za čije postojanje ili objašnjenje nauka u trenutku pisanja ove knjige nema dovoljno dobar model.
- Kako je univerzum nepredvidiv, bez obzira koliko naraste naše znanje o njemu, **nikada nećemo napraviti model univerzuma koji dozvoljava potpuno precizno predviđanje svih događaja**. Najbolje što možemo uraditi je procena verovatnoća određenih ishoda, a čak i to može biti jalov napor ako se ove verovatnoće mogu menjati, zavisno od toga kako tačno funkcioniše stohastički univerzum.

Iako nema mnogo naučnih teorija niti teoretičara koji će eksplicitno izjaviti da veruju da je priroda našeg univerzuma stohastička u smislu koji je gore naveden, **ima mnogo pristupa u različitim oblastima nauke koji se u praksi oslanjaju na metodologiju i naučne alate stvorene za jedan stohastički univerzum**. **Najšire korišćen takav naučni alat je** onaj koji je i tema ove knjige – **statistika**. Možemo lako uočiti da je statistika jedan od osnovnih alata naučnih istraživanja u svim oblastima nauke počev od fizike, pogotovo kvantne fizike, pa preko biologije, posebno oblasti kao

PREPRINT

što su genetika, medicine (studije populacije, epidemiologija...), hemije i drugih, pa sve do društvenih nauka. Ako pogledamo naučne tekstove koji predstavljaju rezultate empirijskih istraživanja u savremenoj ekonomiji, psihologiji, sociologiji, pedagogiji, zapravo u svim društvenim naukama, videćemo da je tu statistike na pretek i da se rad u tim oblastima u ogromnoj meri oslanja na stohastičke modele. Ovo je glavni razlog zašto je povlačenje razlike između determinističkog i stohastičkog pogleda na svet i prirodu univefzuma važno za jednu uvodnu knjigu o statistici. To **predstavlja veoma veliku i važnu kontradikciju u modernoj nauci koje svi koji rade u nauci treba da budu veoma svesni**. U većini slučajeva, videćemo naučnike **kako pokušavaju da formulišu determinističke teorije koje objašnjavaju pojave koje su proučavali, pri tom sve vreme za to koristeći statistiku tj. naučne postupke koji su zasnovani na verovanju da je priroda našeg svemira stohastička**, da bi bi obradili podatke o pojavama koje su proučavali. Dodatno, **mnogi će probati da negiraju postojanje ove kontradikcije** tvrdeći da oni koriste statističke procedure jer su korisne, što je nesumnjivo, a da se pojave u prirodi ponašaju na način sličan stohastičkom zato što nemamo dovoljno znanja o njima. Oni će nekada verovati da je tretiranje pojava kao da im je priroda stohastička prvi korak u njihovom proučavanju i da će, kada budu saznali dovoljno, biti u mogućnosti da naprave precizne determinističke modele, nakon kojih im statistika više neće trebati. Na ovaj način, **statistika i postupci zasnovani na verovanju da je priroda univerzuma stohastička, smatraju se neizbežnim zlom, neophodnim prvim korakom u proučavanju sveta**, nečim što će biti zamenjeno determinističkim modelima onda kada steknemo dovoljno znanja. Ovaj stav se možda može najbolje ilustrovati rečima velikog fizičara Alberta Ajnštajna koji je, razmatrajući tada novu kvantnu teoriju u fizici koja je usvojila staohastički pristup, rekao „Kvantna teorija daje mnogo, ali teško da nas dovodi blizu tajnama Starog (misli na Boga, prim. prev.). Ja sam, u svakom slučaju, ubeđen da se On ne kocka sa univerzumom. [Quantum theory yields much, but it hardly brings us close to the Old One's secrets. I, in any case, am convinced He does not play dice with the universe.]" (*Physics and Beyond: "God Does Not Play Dice", What Did Einstein Mean?*, 2021). Ove reči dosta jasno izražavaju verovanje ovog naučnika, **verovanje koje implicitno većina naučnika deli, da stohastičke metode, iako nesumnjivo korisne, ne odražavaju pravu prirodu univerzuma koja je, u stvari, deterministička**. Bilo kako bilo, ostaje činjenica da je **prava priroda našeg univerzuma**, u trenutku pisanja ove knjige, **nepoznata** i da su sve što imamo samo različita verovanja o njegovoj prirodi.

PREPRINT

Činjenica da možemo da predvidimo mnoge prirodne pojave sa ogromnom, a često i potpunom, preciznošću znači samo da su te pojave bile predvidljive do sada, ali ne znači da su sve pojave svuda predvidljive niti da nema pojava koje su u principu nepredvidljive. Drugim rečima, to nije dokaz da je priroda univerzuma deterministička, jer i u stohastičkom univerzumu mogu postojati skupovi uzroka koji uvek dovode do iste posledice tj. koji su predvidljivi. Činjenica da imamo mnogo veoma dobrih determinističkih modela ne znači da ćemo moći da napravimo takve modele i za pojave za koje ih trenutno nemamo. I što je najvažnije, ovo ne umanjuje niti eliminiše činjenicu **da korišćenje procedura zasnovanih na pretpostavci da je priroda univerzuma stohastička za pravljenje i proveru determinističkih teorija, teorija koje ne ostavljaju prostora za stohastičke procese, predstavlja važnu kontradikciju savremene nauke**. Ova kontradikcija je nešto čega bi trebalo da budu veoma svesni svi koji primenjuju statističke procedure u naučnim istraživanjima.

1.3. Naučna objašnjenja

Kada govorimo o sastavnim delovima nauke koji su bitni za jedan uvodni tekst o statistici, neizbežno dolazimo do teme naučnih objašnjenja. Jedan od ključnih ciljeva nauke je to da objasni pojave sa kojima se susrećemo tj. da da odgovore na pitanja koja počinju sa „zašto“. Naučna objašnjenja i njihova priroda su vekovima bili tema intenzivnih diskusija filozofa nauke i iz tih rasprava je proizašlo više filozofskih modela naučnih objašnjenja (e.g. Woodward & Ross, 2021). Objašnjenje se obično definiše kao skup tvrdnji koje opisuju skup činjenica i objašnjavaju njihove uzroke i posledice. Hempel & Oppenheim (1948) navode da se naučno objašnjenje sastoji od dva dela – eksplananduma i eksplanansa. Eksplanandum je iskaz kojim se opisuje pojava koju treba objasniti, dok eksplanans čine iskazi koji su dodati da bi objasnili tu pojavu (Hempel & Oppenheim, 1948, p. 137). Da bi objašnjenje bilo adekvatno, eksplanandum mora biti logički izvodiv iz informacija koje su sadržane u eksplanansu, eksplanans mora sadržati opšte naučne zakone i sadržaj eksplanansa mora biti empirijski tj. mora biti takav da se može proveriti empirijskim istraživanjem. Iako će se većina naučnika složiti da su najbolja objašnjenja u nauci ona koja pojave objašnjavaju preko uzroka i posledica i da za krajnji cilj nauke možemo smatrati stvaranje uzročno-posledičnog modela celog univerzuma, na taj način objasnivši univerzum i sve u njemu na ovakav način, činjenica je da mnoga široko prihvaćena naučna objašnjenja ne uspevaju

PREPRINT

da dostignu ovaj način objašnjenja zato što mnoge pojave naučnici ne umeju, za sada, da objasne preko uzroka i posledica. Neke od tih nemogućnosti objašnjenja su verovatno zbog nedovoljnog znanja, ali da li su neke od neobjašnjenih pojava suštinski neobjašnjive ili nepredvidive, u ovom trenutku ne možemo da znamo (pogledajte diskusiju o determinističkoj i stohastičkoj prirodi univerzuma u prethodnom poglavlju ili filozofske rasprave o determinizmu i indeterminizmu).

Ovo znači da nauka radi sa različitim tipovima objašnjenja, a za valjano razumevanje teme ove knjige – statistike, veoma je važno da razumemo kada u igru stvaranja naučnih objašnjenja ulaze zaključci izvedeni iz statističkih procedura, kao i to kojim vrstama naučnih objašnjenja one mogu da doprinesu i kako. Jedan način da se naučna objašnjenja podele prema vrstama je da ih klasifikujemo u sledeće tipove:

- **Uzročna ili kauzalna objašnjenja** – su vrsta objašnjena u kojima se **pojava objašnjava tako što se navode uzroci koji dovode do nje i/ili posledice koja ta pojava proizvodi**. Objašnjenja ovog tipa se oslanjaju na prihvaćene naučne teorije, prirodne zakone i opšta pravila da bi objasnila interakciju između različitih varijabli i elemenata koji su uključeni u objašnjenje. U idealnoj situaciji, pojava se objašnjava podvođenjem pod zakone prirode ili pod određenu teoriju (e.g. Hempel & Oppenheim, 1948; C. R. Hitchcock, 1975). Može se reći da je krajnji cilj nauke to da objasni sve što postoji u univerzumu kroz mrežu uzročnih objašnjenja. Ova vrsta objašnjenja važi za **najbolju vrstu naučnih objašnjenja**. Međutim, davanje dobrog kauzalnog objašnjenja zahteva duboko, detaljno i organizovano poznavanje materije, znanje kome će samo objašnjenje doprineti i čiji će postati intergalni deo, a to nije nešto što je dostupno za mnoge pojave i klase pojava. Uzročna objašnjanja, na primer, uključuju objašnjenja zašto planete našeg Sunčevog sistema kruže oko Sunca i ona to objašnjavaju oslanjajući se na zakone fizike. U ovu kategoriju spadaju i objašnjenja toga zašto pada kiša koja se oslanjaju na interakciju atmosferskih uslova koja dovodi do pretvaranja vodene pare u kapljice vode i dovode do toga da te kapljice padaju na zemlju. U biološkim naukama, na primer, objašnjenje toga kako izlaganje očiju svetslosti izaziva impulse u optičkom nervu, a koje uključuje fizička, hemijska i biološka svojstva oka koja omogućavaju da svetlo izazove takve efekte spadaju takođe u ovu vrstu objašnjenja.

PREPRINT

- **Statistička ili probabilistička objašnjenja** – su objašnjenja u kojima se **nešto objašnjava** ili, još češće, **predviđa sa određenom verovatnoćom na osnovu toga što se konstatuje da je određena pojava u prošlosti težila da se javlja zajedno sa nekom drugom pojavom**. U prošlosti smo opazili da dve vrste događaja pokazuju tendenciju da se javljaju zajedno i iz toga zaključujemo da bi ovo moglo da se nastavi tako i u budućnosti. Moguće je da uopšte ne znamo zašto se javljaju zajedno, takođe možemo da ne znamo ni da li je jedan od tih događaja uzrok drugog ili obrnuto; moguće je i da znamo da se oni nisu uvek javljaji zajedno, ali smo primetili da kada se jedan od tih događaja desi, verovatnije je da će se desiti i drugi, nego da se neće desiti. Iz toga onda zaključujemo, ili bolje reći – nadamo se, da će se tako nastaviti i u budućnosti. Veoma važna ideja za koncept statističkih objašnjenja je **Rajhenbahov princip zajedničkog uzroka** (C. Hitchcock & Rédei, 2020; Reichenbach, 1956) koji kaže da **kada dva događaja pokazuju tendenciju da se stalno javljaju zajedno (kada su probabilistički povezani), to je ili zbog toga što je jedan od njih uzrok onog drugog ili zato što su oba uzrokovana nekim trećim faktorom (ili grupom trećih faktora) koji vremenski prethode oba razmatrana događaja**. To znači da kada pravimo generalizaciju ili objašnjenje samo na osnovu opažanja da su dva događaja težila da se dešavaju zajedno u prošlosti, mi ne znamo da li je to možda zbog toga što su oba uzrokovana nekim trećim faktorom koji nismo uopšte opazili. Ovo takođe znači i da ako se taj treći faktor u budućnosti promeni bez našeg znanja, naša dva razmatrana događaja mogu da prestanu da se dešavaju zajedno. Ovo je nešto što je veoma poznato, na primer, ljudima koji se bave predviđanjima kretanja tržišta. Zbog ovih svojstava, **statistička objašnjenja se obično smatraju najgorom vrstom naučnih objašnjenja**. Međutim, ona su takođe **vrsta naučnih objašnjenja koja zahteva najmanje teorijskih znanja i razumevanja problema**. Statistički zaključci o verovatnoćama da se dva događaja dese zajedno mogu da se donesu bez razumevanja dubljih zakona prirode koji upravljaju ovim događajima, kao i bez raspolaganja dobrom ili bilo kakvom teorijom koja bi objasnila pojave o kojima zaključujemo. Ovo čini **statistička objašnjenja prvom vrstom objašnjenja i načina izvođenja zaključaka o ranije nepoznatim pojavama**, s obzirom da se **statistička objašnjenja mogu formulisati samo na osnovu empirijskih opservacija**. Može se reći da **razvoj naučnog razumevanja počinje statističkim objašnjenjima** i da onda napreduje ka kauzalnim objašnjenjima kako istraživači sve više

PREPRINT

upoznaju proučavane pojave i razvijaju sistematska znanja o njima. Ova knjiga predstavlja statističke koncepte i procedure koje su potrebne za dolaženje do statističkih zaključaka koji dozvoljavaju stvaranje ove vrste objašnjenja. To je razlog zašto je važno da čitalac razume poziciju statističkih objašnjenja u tipologiji naučnih objašnjenja, kao i činjenicu da ona predstavljaju najgoru, ali takođe i prvu vrstu naučnog objašnjenja koja se može dati, a i **da stvaranje bilo kakvog objašnjenja koje je bolje od čisto statističkog zahteva stvari koje nisu statistika** tj. da **za stvaranje drugih vrsta objašnjenja statistički podaci i statistički postupci nisu dovoljni**.

Pored ove dve vrste objašnjenja koje možemo uslovno smatrati najboljom i najgorom vrstom naučnih objašnjenja, gde bi uzročno objašnjenje bilo najbolje, a statističko najgore, u nauci se primenjuju i druge vrste objašnjenja, objašnjenja čija je epistemološka vrednost negde između ova dva, objašnjenja koja se s jedne strane oslanjaju na više znanja o proučavanoj pojavi od čistih statističkih/probabilističkih, ali koja s druge strane ne dosežu do niova kauzalnih objašnjenja i nisu u mogućnosti da objasne uzročne veze između pojava koje treba objasniti. Ove vrste objašnjenja uključuju:

- **Razvojna objašnjenja** (e.g. Kitchener, 1983; Woodward, 1980) – takođe poznata i kao **genetička objašnjenja** (Nagel, 1961) su vrsta objašnjenja koja **objašnjavaju zašto je određeni sistem u određenoj fazi razvoja tako što se oslanjaju na zakon koji opisuje niz stadijuma kroz koje sistemi ovog tipa prolaze** (Kitchener, 1983). Woodward (1980) govori o tri vrste razvojnih objašnjenja – a) kada se ponašanje određenog organizma objašnjava pozivanjem na razvojni stadijum u kome se organizam nalazi (na primer, objašnjavanje činjenice da beba plače navođenjem činjenice da se radi o bebi tj. o ljudskoj individui u ranom stadijumu razvoja; b) kada objašnjavamo da je organizam ili objekat na određenom stadijumu ili da će ući u određeni stadijum pozivajući se na naučni zakon o sledu razvojnih stadijuma u razvoju date vrste objekata ili organizama i c) kada objašnjavamo zakon o sledu razvojnih stadijuma pozivajući se na neke šire razvojne principe ili mehanizme koji objašnjavaju napredovanje kroz razvojne faze. Razvojna objašnjenja se mogu naći u svim oblasti nauke, ali su najverovatnije najpoznatija ona koja nalazimo u biološkim naukama, a koja objašnjavaju faze razvoja organizama ili ona u

PREPRINT

psihologiji koja objašnjavaju psihološki razvoj pojedinaca kroz detinjstvo i tokom celog životnog veka.

- **Funkcionalna objašnjenja** – su objašnjenja u kojima se **određena jedinica nekog sistema objašnjava navođenjem funkcije koju ta jedinica obavlja u većem sistemu i to obično tako što se navodi svojstvo ili osobina sistema čije održavanje data jedinica omogućava**. Primer ove vrste objašnjenja bi bila situacija kada objašnjavamo ulogu srca u ljudskom organizmu tako što navodimo da je funkcija srca da pumpa krv kroz telo. Primer iz društvenih nauka bi bilo objašnjenje postojanja škola u društvu navođenjem da škole postoje da bi obrazovale decu i tako doprinele opštem nivou obrazovanja u društvu.
- **Teleološka objašnjenja ponašanja živih organizama** – su posebna vrsta objašnjenja gde se **ponašanje osobe, grupe osoba ili određenog živog organizma objašnjava navođenjem ciljeva tog ponašanja**. Na primer, tako možemo objasiti činjenicu da student uči navođenjem da je njegov/njen cilj da položi ispit (ispit za koji uči). Ovim putem smo ponašanje objasnili navodeći ciljeve ili svrhu koju osoba ima i zbog koje se ta osoba ponaša onako kako se ponaša. Iako neki autori smatraju da funkcionalna i teleološka objašnjenja spadaju u istu kategoriju naučnih objašnjenja (e.g. Nagel, 1961), naše je mišljenje da postoji jasna razlika između objašnjenja gde istraživač objašnjava funkciju koju određeni deo ima u radu nekog šireg sistema (kao što je slučaj sa srcem koje pumpa krv kroz telo ili sa atmosferom koja omogućava disanje živim organizmima i tako omogućava postojanje ekosistema Zemlje) i situacije gde objašnjavamo namerno ponašanje ljudi i drugih živih organizama kroz navođenje namera, onako kako je to, na primer, slučaj u psihologiji i u pravnim naukama kada je namera važan faktor objašnjavanja i procene ponašanja određene osobe.

Dok su uzročna objašnjenja vrsta objašnjenja koja se već vekovima skoro pa podrazumeva u naukama poput fizike i hemije, u drugim naukama, poput biologije, nauka o životnim procesima, a pogotovo društvenim naukama, objašnjenja ove vrste su retka. Međutim, sa razvojem nauke, naučne metodologije, ali i tehnologije naučnih merenja, te razvojem računarskih sistema koji omogućavaju istraživačima da složene proračune rade daleko jednostavnije i brže nego da to rade „peške“, može se primetiti sve više pomaka ka tome da se kauzalna objašnjenja uvedu u nauke u kojima ih je ranije bilo srazmerno malo. Dok su se u mnogim oblastima biologije i nauka o životu nekada naučna objašnjenja svodila na funkcionalna, razvojna ili čak statistička, zadnjih decenija

PREPRINT

primetan je nagli razvoj naučnih modela koji pretenduju da daju kauzalna objašnjenja važnih bioloških procesa. Na primer, jedan od autora ovog teksta veliki deo svoje karijere posvetio je razvoju i predlaganju matematičkih modela koji na kauzalan način objašnjavaju različite fiziološke procese poput procesa oplodnje (e.g. Andjelka Hedrih, 2014; Andjelka Hedrih et al., 2015; Andjelka Hedrih & Banić, 2016; Andjelka Hedrih & Ugrčić, 2012) ili mehanike ponašanja lanaca dezoksiribonukleinske kiseline (DNK) tokom procesa kopiranja i prepisivanja (e.g. Andjelka Hedrih, 2011; Andjelka Hedrih & Hedrih, 2014; K. Hedrih & Hedrih, 2010) ili detalja ponašanja stabala drveća pod dejstvom vetra (Andjelka Hedrih, 2021b) i uloge mehaničkih oscilacija u biološkim procesima generalno (Andjelka Hedrih, 2021a; Andjelka Hedrih & Hedrih, 2016), što su sve procesi čija su se objašnjenja u ranijim decenijama uglavnom svodila na razvojna ili funkcionalna. Na sličan način, u oblasti psihologije možemo videti porast broja istraživanja u kojima se eksperimentalne metode (koje su praktično najsigurniji način dolaženja do saznanja koja omogućava kauzalna objašnjenja) primenjuju na složenim društvenim pojavama (e.g. Djoric, 2021) ili se testiraju odnosi koji uključuju kauzaciju poput prenosa dejstva jednog faktora na drugi, preko nekog trećeg, tzv. medijacija ili situacija gde jedan faktor menja odnos neka druga dva, tzv. moderacija (e.g. Lee & Ok, 2014; Liao et al., 2017; Okech et al., 2018)

1.4. Samoispunjujuća proročanstva i kvarenje statističkih pokazatelja

Kada razmatramo naučne zakone, generalizacije i predviđanja u kontekstu ponašanja ljudi, valja primetiti da **primena naučnih znanja u svrhu donošenja odluka koje će se primeniti kroz intervencije u praksi može dovesti do promena ponašanja kod ljudi na koje se te odluke odnose i do dodatnih intervencija koje mogu biti ili u skladu sa očekivanjima onih koji su pomenute odluke doneli ili usmerene ka tome da se tim odlukama suprotstave ili da promene njihove efekte**. Ovo se dešava zato što kada su ljudi svesni koncepata koje oni koji imaju moć nad njima koriste da donose odluke i te odluke su ljudima važne, ljudi će težiti da stvore aktivan odnos prema tim odluka i intervencija i mogu odlučiti da shodno tome promene i svoje ponašanje. Ovako može doći do takozvanih samoispunjavajućih proročanstava. **Do samoispunjavajućih proročanstava dolazi onda kada ljudi, koji bi se inače ponašali drugačije, počnu da se ponašaju da način koji se od njih očekuje nakon što su saznali da se takvo ponašanje od njih**

PREPRINT

očekuje i počeli da veruju u takvo očekivanje. Na primer, student koji nije imao dobar uspeh na studijama, ali se trudio da uspeh popravi može odustati od daljih pokušaja da uspeh popravi i „prihvatiti svoju sudbinu“ nakon što mu nakon, na primer, kognitivnog testiranja kažu da su njegove kognitivne sposobnosti slabe. On će tako ispuniti očekivanja onih koji su mu to saopštili - da studenti slabih kognitivnih sposobnosti imaju slabe rezultate na studijama. Osoba koja razmišlja o tome da započne novi posao, može da odustane od te ideje nakon što mu/joj kažu da će ekonomija na području u kome je on/ona razmišljao/la da pokrene posao stagnirati. Na ovaj način, budući doprinos njegovog/njenog posla ekonomiji se neće realizovati, tako itekako pomažući da se očekivanje da će ekonomija stagnirati ostvari.

Samoispunjujuća proročanstva su takođe veoma važna tema kada treba predvideti kretanja organizovanih finansijkih tržišta – izjave o očekivanjima različitih analitičara i važnih ličnosti o tome kako će se cene deonica i/ili drugih finansijskih instrumenata kretati, mogu podstaći ljude da kupuju ili prodaju deonice (i/ili druge finansijske instrumente) i tako dovesti do toga da se kretanje cene u predviđenom smeru zaista i desi. Poznata je strategija „šortera“ (ljudi i institucija koje šortuju tj. prodaju pozajmljene deonice sa očekivanjem da će ih kupiti i tako vratiti pozajmicu onda kada njihova cena bude pala) da pokušavaju da formiraju očekivanje javnosti da će se cena deonica koje su šortovali sniziti, tako što stvaraju negativan publicitet kompaniji čije deonice su šortovali uvek uključujući u negativnu priču koju promovišu i očekivanje da će se cene sniziti. Ko je na primer pratio dešavanja na finansijskim tržištima u SAD u zimu 2020-2021 mogao je da primeti kako su cene akcija jedne kompanije, čije je poslovanje već duže vreme bilo na jasnoj i produženoj silaznoj putanji, narasle desetostruko za svega nekoliko meseci, bez bilo kakvih razloga koji bi imali veze sa promenama u poslovanju te kompanije, samo zbog očekivanja ljudi koja su formirana kroz društvene mreže i reklamu, a koje su podsticale ljude da kupuju akcije te kompanije, najavljujući eksplozivan rast njihove cene. Ovaj konkretni događaj u trenutku pisanja ove knjige postao je poznat kao „Gamestop ludilo [Gamestop frenzy]“.

Još jedan primer samoispunjujućih proročanstava su različite situacije u kojima se beleže **placebo efekti**, tj. **situacije kada postupci koji, sami za sebe, nemaju nikakve efekte ipak dovode do određenih rezultata zato što je kod ljudi na kojima su ovi postupci primenjeni stvoreno očekivanje da će efekata biti**. Jedan klasični primer ovakve pojave je takozvani **Hautorn efekat** (Howthorne effect), koji je ime dobio po seriji istraživanja produktivnosti radnika pod različitim

PREPRINT

fizičkim uslovima radne sredine, a koja su sprovedena početkom 20. veka. U ovoj seriji istraživanja, istraživači su ispitivali kako različiti uslovi rada, npr. razlike u osvetljenju prostorije, pauzama u toku rada, broju radnih sati, utiču na produktivnost radnika. Rezultati su pokazali da je produktivnost grupa radnika koji su bili uključeni u ova istraživanja generalno bila bolja u odnosu na uobičajenu produktivnost radnika koji nisu bili uključeni u studiju manje-više nezavisno od toga kako su istraživači menjali uslove. Ovo je objašnjeno time što su radnici znali da istraživači istražuju produktivnost. Npr. kada su ispitivali efekte osvetljenja, produktivnost radnika je rasla čak i kada su istraživači počeli da smanjuju stepen osvetljenosti prostorije sve više i više. Radnici su počeli da se žale tek kad je nivo osvetljenja sveden na onaj koji odgovara svetlosti koju daje mesečina tokom noći. Slična stvar se dešavala i kada su ostale varijable bile u pitanju (e.g. Wickstrom & Bendix, 2000)

Još jedan tema koju treba razmotriti kada se radi o zaključivanju o pravilnostima sa ciljem da se na osnovu njih donesu odluke i preduzmu intervencije je **fenomen korupcije ili kvarenja statističkih indikatora**. U oblasti javnog upravljanja, upravljanja preduzećima i u različitim organizacionim situacijama, ljudi koji su zaduženi za upravljanje ovim organizacijama se često oslanjaju na različite načine merenja događaja i pojava koji su za njih važni i onda na tim merenjima zasnivaju svoje odluke. Na primer, za procenu kognitivnih sposobnosti, ljudima se zadaju određeni zadaci u okviru psiholoških testova; za merenje inflacije, uzimaju se u obzir cene određenih proizvoda i onda se na osnovu njih računa indeks promene cena; da bi se procenila pismenost ili kvalitet obrazovanja određene zemlje prati se uspeh učenika iz određenih škola ili se testiraju ti učenici; da bi se pratile promene cena deonica na berzi, prate se cene deonica određenih preduzeća. Na sličan način nastavnici ispituju stepen u kome su njihovi učenici savladali gradivo tako što im daju zadatke ili postavljaju pitanja koja ispituju usvojenost određenih gradiva, ali ne celokupnog gradiva jer bi takav postupak bio preobiman, već praveći izbor pitanja koja će zadati u okviru testova ili procesa ispitivanja. Ovo su sve primeri situacija gde se odluke donose na osnovu vrednosti određenih (statističkih) pokazatelja, odluke koje utiču na različite načine na ljude i za koje različiti ljudi mogu biti zainteresovani. **Kvarenje statističkih indikatora (pokazatelja) se dešava u situacijama kada ljudi na koje utiču vrednosti određenih statističkih indikatora i/ili koji su zainteresovani za ishode odluka koje su zasnovane na tim pokazateljima, preduzimaju radnje da promene vrednosti ovih pokazatelja na način koji im ide u prilog, ali samo da promene vrednost pokazatelja, bez promene svojstva na koje ti pokazatelji treba da**

PREPRINT

ukazuju. Ovakvo njihovo ponašanje zasnovano je na tome što **oni razumeju kako ti pokazatelji rade i stoga razumeju i šta je potrebno uraditi da bi se promenili**. Kada dođe do ovakvog ponašanja, **statistički pokazatelji postaju nevalidni**, a onda će **i odluke koje budu donesene na osnovu njih biti nevalidne.** Biće nevalidne jer su zasnovane na vrednostima pokazatelja koji su sada postali nevalidni. Na primer, ako studenti znaju koja će tačno pitanja nastavnik da im postavi na ispitu mogu odlučiti da nauče samo odgovore na ta konkretna pitanja, a da zanemare svo ostalo gradivo. Takvi studenti će onda dobiti odličnu ocenu i položiti ispit iako je njihovo poznavanje gradiva zapravo vrlo malo tj. iako ne znaju ništa osim tih konkretnih pitanja čije su odgovore naučili. A ispitni test će postati nevalidan, iako bi, da studenti nisu unapred saznali njegov sadržaj, bio sasvim validan pokazatelj stepena u kom su savldali gradivo. Na sličan način, vlada neke zemlje koja je zainteresovana za to da „pokaže“ kako je inflacija u njihovoj zemlji niska može preduzeti mere da snizi zvanične cene odabranih proizvoda na osnovu kojih se računa indeks potrošačkih cena (čak i po cenu da takvim mešanjem u cene dovedu do nestašice!!), a da istovremeno dozvoli da cene ostalih proizvoda drastično porastu. Tako bi dobili situaciju da je statistički indeks koji ukazuje na inflaciju nizak, iako je realna inflacija možda ogromna. Na ovaj način, statistički indeks koji se koristio kao pokazatelj inflacije postao je nevalidan, iako je bio sasvim dobar pokazatelj inflacije pre ove vladine intervencije. Isto tako, preduzeće čijim se deonicama trguje na berzi može odlučiti (a takve primere uopšte nije teško naći) da naručuje sama svoje sopstvene proizvode i da onda te narudžbine sama sebi plaća, kako bi naduvala podatke o prometu preduzeća. Investitori koji prate promenu prometa preduzeća kao jedan od pokazatelja rasta bi onda (pogrešno) zaključili da promet preduzeća raste i bili bi spremniji da skuplje plate deonice kompanije, što bi dovelo do rasta cena deonica tog preduzeća na berzi. Još jedan veoma poznat istorijski primer fenomena kvarenja statističkih pokazatelja je bila situacija sa proizvodnjom čelika u komunističkoj Kini tokom Velikog Skoka Napred (https://en.wikipedia.org/wiki/Great_Leap_Forward) sredinom 20. veka. Tada je stanovnicima sela širom Kine naređeno da treba da proizvode čelik, iako nisu imali ni resurse, ni industrijsku infrastrukturu niti opremu potrebnu za industrijsku proizvodnju čelika. Ovo je dovelo do toga da su sela proizvodila beskorisne grumenove čelika lošeg kvaliteta (“pig steel” na engleskom) pretapajući razne predmete koji su im bili dostupni, a koji su sadržali čelik. Jedini cilj ovakve aktivnosti je bio da se zabeleži da u Kini raste proizvodnja čelika, jer je u to vreme količina proizvedenog čelika u zemlji tokom određenog perioda korišćena kao pokazatelj stepena razvoja

PREPRINT

određene zemlje. Ovo je bilo moguće jer su u to vreme industrijalizovane, dakle razvijene zemlje, proizvodile mnogo više čelika nego nerazvijene i to zato što je čelik bio neophodan za razvoj industrijske infrastrukture, kao i za proizvodnju različitih industrijskih proizvoda. Međutim, ništa od toga nije bio slučaj sa ovim pomenutim grumenovima čelika koji su bili potpuno beskorisni, učinivši tako količinu proizvedenog čelika potpuno nevalidnim pokazateljem razvoja u slučaju Kine u datom trenutku. Naravno, veoma brzo je postao jasno da veštačko "naduvavanje" statističkog pokazatelja razvoja nije isto što i pravi razvoj i posledice toga su u ovom slučaju bile vrlo teške.

Verovatno najobuhvatiniji i najdugotrajniji primer ovog efekta je pojava takozvanog **Flinovog efekta u oblasti testiranja inteligencije i kognitivnih sposobnosti** uopšte. Naime, primećeno ja da su tokom 20. veka i prvih godina 21. veka postignuća ljudi širom razvijenog sveta na testovima inteligencije rasla (e.g. V. Hedrih, 2020). Različiti psiholozi iznosili su različite pretpostavke o poreklu ovog fenomena (e.g. Sundet et al., 2004; Teasdale & Owen, 2005), ali se na kraju pokazalo, kako je to objasnio sam Flin [Flynn], novozelandski psiholog po kom je ovaj efekat i dobio ime (Flynn, 2007), da **kako su testovi inteligencije i drugi testovi slični testovima inteligencije sve šire i šire korišćeni za donošenje različitih odluka od značaja za živote ljudi, tako su i ljudi postajali sve upoznatiji i upoznatiji sa zadacima kakvi se sreću u testovima inteligencije i postajali sve veštiji u njihovom rešavanju (i tako dobijali više skorove), iako se inteligencija, psihološka crta koju ti zadaci treba da mere, nije promenila već je ostala manje-više na istom nivou**. Na ovaj način, sticanje nove, relativno beskorisne, veštine (veštine rešavanja zadataka poput onih koji se koriste u testovima inteligencije), pokvarila je pokazatelje važnih kognitivnih sposobnosti i tako naterala psihologe da inoviraju i da menjaju pravila za interpretaciju testova kognitivnih sposobnosti da bi očuvali njihovu validnost.

Svi ovi primeri pokazuju zašto su i samoispunjujuća proročanstva i kvarenje statističkih pokazatelja fenomeni koji moraju da se uzmu u obzir prilikom korišćenja naučnih objašnjenja, teorija i sličnih instrumenata nauke u oblasti društvenih nauka za donošenje odluka koje utiču na ljude, kao i zašto je neophodno stalno biti svestan da primena određenih pravila na ljude može dovesti do toga da ti ljudi promene svoje ponašanje kao odgovor na ta nova pravila i odluke, a zavisno od toga kako procene ta pravila i odluke. Takođe treba da budemo svesni da **mnoge pravilnosti koje su nađene u oblasti nauka o ponašanju nastavljaju da budu pravilnosti samo**

PREPRINT

zato što ljudi i dalje opažaju situacije koje se tiču tih pravilnosti na isti određeni način i aktivno donose odluke da se ponašaju na isti način. Međutim, to je **nešto što se može lako promeniti kao rezultat bilo kakve intervencije zasnovane na tim pravilnostima**. Prema tome, za razliku od situacije koja postoji u prirodnim naukama, gde neživi objekti ne mogu da promene svoje ponašanje na osnovu toga kako procenjuju neku odluku koja ih se tiče, **primena nalaza i uočenih pravilnosti u društvenim naukama na ljude može lako dovesti do toga da ti isti nalazi i pravilnosti prestanu da budu validni tj. prestanu da važe**.

1.5. Nauka i pseudonauka

Savremeni svet sagrađen je na tekovinama nauke. Svuda oko nas, većina predmeta koje koristimo, predmeta od kojih zavisi održanje našeg načina života, pa i održanje samog života, proizvodi su nauke. Najveći broj njih proizvod je vekova naučnih istraživanja i razvoja koji su doveli do veština i znanja koji su omogućili da znamo i umemo kako da te predmete napravimo. Zbog ovoga, nauka u društvu ima veoma veliki ugled i ljudi joj veruju toliko da na dnevnoj bazi oslanjaju svoje živote na to poverenje. To poverenje je s druge strane, veoma zasluženo, jer je nauka rezultat združenih napora naučnika iz celog sveta koji veoma pažljivo koriste naučni metod i testiraju, proveravaju i ponovo proveravaju sve svoje rezultate i nalaze. Međutim, kao i kod svih vrednih stvari, postoje i ljudi koji bi želeli da se okoriste o to poverenje da bi ostvarili svoje lične ciljeve, često na štetu drugih ljudi. I to je put kojim nastaje pseudonauka. **Pseudonauku** možemo definisati kao **bilo koji skup verovanja, tvrdnji ili postupaka koji se predstavlja** (ili ga predstavljaju) **kao nauku, a da nije nastao kao posledica upotrebe naučnog metoda, kome nedostaju dokazi ili logička konzistencija i koji ne može biti potvrđen ili opovrgnut (Popper, 1963) ili koji se iz drugih razloga ne može smatrati naučno validnim**. Iako je ova definicija psuedonauke dosta jasna na apstraktnom nivou, postavlja se pitanje praktičnog razlikovanja nauke i pseudonauke i to je veoma važno pitanje. Ako se uradi pogrešno, takva pogrešna demarkacija može dovesti ili do toga da se prave, validne naučne aktivnosti proglase za pseudonaučne ili da se pseudonaučne prakse i sistemi uključe u nauku tako je obezvređujući i šteteći naporima naučnika da dovedu do spoznaje sveta oko nas (e.g. Lakatos, 1978).

Kako onda da razlikujemo nauku od pseudonauke? Dejvidu Hjumu, poznatom filozofu iz 18. veka pripisuju se sledeće reči: „Ako u ruku uzmemo bilo koje izdanje; o božanstvenom ili o školi

PREPRINT

metafizike na primer; zapitajmo se, da li ono sadrži apstraktna razmišljanja koja se tiču količine ili brojeva? Ne. Da li sadrže bilo kakvo eksperimentalno razmišljanje koje se tiče činjenica i postojanja? Ne. Bacimo ih onda u vatru. Jer ne sadrže ništa osim sofizma (reč „sofizam“ se obično odnosi na nelogičan ili nevalidan argument napravljen sa ciljem da zavara ili prevari onog kome se iznosi, prim. prev.) i iluzija“ (Lakatos, 1978, p. 1). Iako je ovaj stav u osnovi tačan, njegova praktična primena je nešto sasvim drugo. Naime, s jedne strane je sasvim jasno da postoje pseudonaučni sistemi koji uključuju i brojeve i apstraktno rezonovanje (npr. numerologija, astrologija), istorija nauke beleži i situacije kada su legitimne naučne aktivnosti i istraživanja smatrani pseudonaučnim od strane delova društva, kao što je to na primer bio slučaj sa psihologijom (e.g. Ferguson, 2015; Skinner, 1990) ili sa naučnicima koji su prihvatali Mendeljejevsku genetiku, a koji su tokom jednog perioda bili tretirani kao pseudonaučnici u nekadašnjem Sovjetskom Savezu (Lakatos, 1978). S druge strane, postojale su takođe situacije kada su sistemi koji su jasno pseudonaučni bili prihvatani kao nauka. Primer za to je slučaj iz 17. veka kada je veštičarenje tretirano kao uzorni naučni pristup od strane zvaničnog filozofa britanskog Kraljevskog društva (Lakatos, 1978). Dobar primer je i savremeno prihvatanje „medicinskih postupaka“ i „tehnika lečenja“ koje su zasnovane na „kosmičkoj energiji“ od strane vlada različitih zemalja, uključujući i Srbiju. Tako se na primer u arhivi web stranica Ministarstva zdravlja Republike Srbije (http://arhiva.zdravlje.gov.rs/showpage.php?id=216) od 16.4.2009. godine može naći i tekst prikazan na slici 1.

Slika 1.1. Snimak ekrana arhivirane stranice na sajtu Ministarstva zdravlja Republike Srbije, na kojoj se predstavlja jedna od „tehnika lečenja“ zasnovana na „isceljujućoj kosmičkoj energiji“

PREPRINT

Република Србија
МИНИСТАРСТВО ЗДРАВЉА

Датум:
08.12.2021.

Језик:
Српски English

Напредна претрага

Навигација
Министарство
Документа
Актуелности
Међународни пројекти
Препоручујемо
Линкови
Архива
Контакт подаци
Јавне набавке
Услуге
Најчешћа питања

Анкета
Да ли бисте потписали донаторску картицу ?
Да
Не
Гласај

Почетна страна : Препоручујемо : Традиционална медицина : Методе у Србији : Реики

Са подацима од: 16.04.2009.

ТРАДИЦИОНАЛНА (КОМПЛЕМЕНТАРНА И АЛТЕРНАТИВНА) МЕДИЦИНА

Почетна | У развијеним земљама | Развојни пут | Систематизација метода | Преглед метода у Србији

РЕИКИ

Реики је исцељујућа космичка енергија и систем холистичког лечења који се на тој енергији заснива. Током третмана терапеут/исцељитељ полаже руке на пацијента и тиме преноси ову благотворну енергију на његово биоелектромагнетно поље. Према класификацији Националног центра за комплементарну и алтернативну медицину при америчком Националном институту за здравље, Реики спада у медицину биопоља. Савремена примена Реикија почиње 20-тих година прошлог века у Јапану када Микао Усуи, оснивач Реикија, формулише и први успешно примењује и подучава овај систем природног лечења. Наследник Усуија, Цујиро Хауаши отвара прву Реики клинику у Токију и успоставља сарадњу са здравственим институцијама. После другог светског рата Реики се проширио по целом свету и данас је, осим у Јапану и другим земљама Далеког Истока, најраспрострањенији у Сједињеним Државама, Канади, скоро свим европским земљама, Јужној Америци и Аустралији. Широм света постоје многе Реики организације и удружења Реики терапеута и инструктора. Постоји више научних хипотеза о механизму дејства Реикија, мада јединствена теорија још увек није усаглашена. Познато је да поље Реики енергије индукује биоелектромагнетне токове унутар ћелија и ткива и тиме појачава способности организма за самолечење и отпорност према стресу. Од многобројних позитивних ефеката, најпознатији су: смањење напетости, стреса, опште психофизичко опуштање, повећање

Imajući sve ovo u vidu i svesni da je pitanje razlikovanja nauke od pseudonauke bilo i i danas je tema velika debate među filozofima nauke i naučnicima uopšte, mogu se primetiti neke **tendencije i pravilnosti koje se mogu koristiti u praksi da bi prepoznali da li imamo posla sa pravom naukom ili pseudonaukom koja se predstavlja kao nauka**. Ove razlike se mogu videti u:

- **Objavljivanju rezultata** – naučnici po pravilu svoje rezultate prvo objavljuju u naučnim časopisima i sličnim naučnim publikacijama, publikacijama koje recenziraju drugi naučnici i eksperti u oblasti. Te publikacije održavaju striktne standardne naučnog poštenja, tačnosti i preciznosti u predstavljanju rezultata istraživanja. Rezultati se prikazuju opštoj, nenaučnoj javnosti tek ako (i nakon što) uspešno prođu ocenjivanje i preispitivanje od strane drugih eksperata, a prvenstveno onih koji su nezavisni od autora istraživanja. Tek onda se opštoj javnosti ovi rezultati prikazuju u vidu takozvanih sekundarnih publikacija (sekundardnih u smislu da to nisu publikacije u kojima se prvi put prikazuju dati naučni

rezultati) u kojima se rezultati prikazuju na način koji je prilagođen opštoj javnosti. Takve publikacije uključuju knjige, monografije, udžbenike, popularne naučne članke, video klipove i druge vrste medijskih sadržaja. Za razliku od ovog pristupa, pseudonaučnici **svojim publikacijama prvenstveno ciljaju na opštu javnosti, a idealno na ljude koji nisu dovoljno stručni i nemaju znanja koja su potrebna da bi valjano procenili kvalitet i sadržaj materijala koji im pseudonaučnici predstavljaju**. Dok na ovakve ljude ciljaju, pseudonaučnici po pravilu **izbegavaju prave stručnjake za oblast** kako god mogu, a takvo svoje ponašanje pokušavaju da objasne tvrdnjama da oni samo „izbegavaju zavere“ koje „protiv njih organizuju“ eksperti ili tvrdeći da su njihove publikacije zabranjene (iako nisu) i na druge slične načine. **Pseudonaučne publikacije** se ne podnose na recenziranje ekspertima za datu naučnu oblast, **ne prolaze kroz bilo kakve provere** i njihovi **autori ne stavljaju pred sebe bilo kakve zahteve koji se tiču tačnost ili istinitosti tvrdnji** koje iznose u takvim publikacijama. Autori i distributeri takvih publikacija će, na primer, probati da svoje publikacije distribuiraju na mestima poput lokalnog frizera, po kafeima ili će pokušavati da ih prodaju naivnim ljudima, ali će aktivno izbegavati bilo kakvo proveravanje ili evaluaciju od strane ljudi koji se zapravo razumeju u materiju koju obrađuje pseudonaučna publikacija.

- **Ponovljivi rezultati** – nauka zahteva rezultate koji se mogu ponoviti i tako proveriti. Iz ovog razloga, u naučnim publikacijama i izveštajima o istraživanjima svi sprovedeni postupci se predstavljaju i opisuju sa dovoljno detalja da drugi naučnici mogu da ih ponove tačno onako kako su urađeni, tj. da ponove tačno taj isti postupak koji su uradili autori publikacije. Drugi naučnici tako mogu da sami provere da li će i oni dobiti iste rezultate. S druge strane, **pseudonaučna „istraživanja“ se ne mogu ni ponoviti ni proveriti**. Ako i ima bilo kakvog opisa „istraživačkog“ postupka, taj postupak je samo vrlo maglovito opisan, tako da je nemoguće zasigurno utvrditi šta je tačno rađeno i kako tačno. Na primer, jedan veoma raširen pseudonaučni sistem opisuje poreklo svojih svojih „znanja“ preko priče o tome kako je taj sistem osnovalo nekoliko ljudi koji su otkrili sva znanja i elemente koji čine sistem zato što su ti ljudi „bili vrlo talentovani sa veoma posebnim uvidima“ i eksplicitno negiraju da bi bilo ko mogao da ponovi njihove istraživačke postupke, zato što niko drugi nema „talente“ niti „specijalne sposobnosti“ koje su imali njihovi osnivači. Šta god bilo ko uradio da ponovi pseudonaučni „istraživački postupak“, ako dobije negativne

PREPRINT

rezultate, pseudonaučnik će mu prosto reći da je on nešto pogrešio i da nije kako treba ponovio proceduru! Naravno, ovakav pristup nema ništa zajedničko sa time kako nauka treba da radi.

- **Ponašanje prema greškama** – **nauka aktivno traži greške**. Kada jedna grupa naučnika sprovede istraživanje i objavi rezultate, drugi naučnici će pregledati njihove nalaze i, u tom postupku, aktivno tražiti greške koje su napravili, alternativna objašnjenja koja su propuštena, kao i druge neregularnosti. Naučnici znaju da, kao što i pokvaren sat 2x dnevno pokazuje tačno vreme, tako i pogrešne teorije nekada mogu (slučajno) dati tačna predviđanja. Iz ovog razloga, naučnici će testirati pretpostavke naučnih teorija u novim situacijama i u novim uslovima, znajući da ako je teorija tačna ona neće davati pogrešna predviđanja. S druge strane, **pseudonaučnici ignorišu greške, opravdavaju ih, lažu da ne postoje, odbijaju da priznaju njihovo postojanje, pokušavaju da ih objasne neadekvatnim ad-hoc objašnjenjima, trude se da zaborave na njih, da ih sakriju ili da urade bilo šta drugo kako bi minimizovali svest drugih o njima, te kako bi preusmerili pažnju javnosti na nešto drugo umesto na greške u njihovim predviđanjima**. Ovo rade zato što je cilj pseduonaučnika to da ubede druge da su njihove tvrdnje tačne, a ne da provere da li su im tvrdnje zaista tačne.
- **Napredak znanja** – **nauka napreduje**. Kako se sprovodi sve više i više naučnih istraživanja, naučnici saznaju sve više i više o fenomenima koje istražuju, a to znači i da fizički procesi koji stoje u osnovi fenomena koje naučnici proučavaju postaju sve bolje i bolje istraživani i sve je više znanja o njima. Naučna istraživanja nekog fenomena mogu početi od nule, ali svako novo istraživanje povećava količinu znanja koja postoji o tom fenomenu. Doprinos naučnom znanju je osnovni zahtev koji se postavlja prilikom evaluacije bilo kog naučnog rada. Naučni rad koji ne doprinosi postojećim znanjima, ne obezbeđuje nikakva nova znanja, bilo kroz proveru postojećih generalizacija i teorija ili kroz nova otkrića se u principu smatra bezvrednim. Nasuprot ovome, **pseudonauka ne napreduje i vremenom se ništa novo ne saznaje**. Fizički proces koji je u osnovi fenomena koji je tema pseudonaučnog učenja se ne istražuje. Šta god da je bilo poznato onda kada je sistem napraviljen, po pravilu ostaje isto tokom celokupnog vremena postojanja pseudonaučnog sistema.

PREPRINT

- **Oslanjanje na dokaze i podatke** – nauka ubeđuje ljude u valjanost tvrdnji koje iznosi iznošenjem dokaza, korišćenjem logičkih argumenata ili matematičkih odnosa, modela i opisa. Nauka pokušava da od postojećih podataka izvuče najbolje što može. Kada se pojave novi podaci i dokazi, takvi da pokazuju da prethodne interpretacije podataka nisu bile ispravne, to se prihvata, te ranije interpretacije se napuštaju i traže se nove koje su bolje i koje se uklapaju u postojeće podatke. S druge strane, **pseudonauka pokušava da navede ljude da prihvate njene tvrdnje bez postavljanja bilo kakvih pitanja**. Pseudonaučnici pokušavaju da preobrate ljude dajući pseudoreligiozni karakter svojim verovanjima i trude se da ljudi prihvate njihove tvrdnje uprkos činjenicama koje pokazuju da su tvrdnje pseudonaučnika neistinite. **Pseudonaučnici najčešće nikada ne napuštaju svoju izvornu ideju bez obzira šta podaci pokazuju**.
- **Način finansiranja** – **nauka na tržište izbacuje samo proizvode i postupke koji su uspešno prošli rigorozno provere**. Svaki proizvod nauke prolazi kroz proces rigoroznog i temeljnog testiranja pre nego što bude ponuđen javnosti. U ovom postupku, ispituju se i bezbednost i efikasnost datog proizvoda. Ako proizvod ne prođe ove provere i pokaže se da ne može ni da se unapredi dovoljno da bi ih prošao, takav proizvod se napušta i nikad ne bude izbačen na tržište. S druge strane, **glavni izvor prihoda pseudonauke i pseudonaučnika je prodaja raznoraznih problematičnih proizvoda**, proizvoda koji tipično dolaze sa velikim obećanjima, ali, u boljim slučajevima, isporučuju slabe ili nikakve rezultate, a u gorim slučajevima čak i nanose štetu kupcu. Takvi proizvodi uključuju stvari poput „lekova“ koji ne mogu da izleče ništa, knjiga i „obrazovnih“ kurseva koji obećavaju fantastične rezultate ili pseudonaučne usluge poput proricanja sudbine, komunikacije sa duhovima, mrtvima ili natprirodnim silama, izrada horoskopa isl. Ovakvi proizvodi u najboljem slučaju imaju efekat jednak placebu, a nekada mogu biti i štetni za osobu koja ih koristi. Placebo efekat je pojava da primena proizvoda ili tretmana za koje se, na osnovu njihovog sadržaja, ne bi moglo očekivati da imaju ikakav efekat na osobu, pokažu određeni efekat kao posledicu verovanja osobe da ti proizvodi/usluge „rade“ i da će proizvesti efekat kakav se dobije.

PREPRINT

Tabela 1.2. Kako razlikovati nauku od pseudonauke, pregled heurističkih smernica

Kriterijum razlikovanja	Nauka	Pseudonauka
Objavljivanje rezultata	Prvo se objavljuju u publikacijama za eksperte i stručnu javnost, prolaze recenzije i intenzivna ispitivanja i provere od strane drugih eksperata. Za opštu publiku se objavljuju tek nakon što uspešno prođu evaluaciju od strane stručnjaka.	Usmereni na opštu javnost, a posebno na ljude bez stručnih znanja koja su potrebna da valjano procene tvrdnje koje iznose pseudonaučnici. Recenzija ili bilo kakva evaluacija od strane stručnjaka se aktivno izbegava.
Replikabilnost / ponovljivost rezultata	Nauka zahteva ponovljive rezultate i zato se postupci opisuju u dovoljno detalja kako bi drugi naučnici mogli da ponove istraživački postupak i tako sami provere rezultate.	Rezultati ne mogu da se provere. Postupci se ne opisuju ili se opisuju samo maglovito tako da ne mogu da se ponove, jer se iz opisa ne može videti šta je tačno rađeno.
Ponašanje prema greškama	Greške se aktivno traže da bi se ispravile.	Greške se ignorišu, kriju, prikrivaju lažnim objašnjenjima, o njima se laže… Početni stavovi se nikada ne napuštaju, niti se koriguju.
Napredak znanja	Nauka napreduje. Sa novim istraživanjima količina znanja o pojavi koja je predmet istraživanja i fizičkim procesima koji stoje u njenoj osnovi raste. Sa svakim istraživanjem saznajemo nešto novo.	Napretka znanja nema. Pseudonaučni sistem zauvek ostaje manje-više isti kakav je bio kad je stvoren. Nova "istraživanja" ne stvaraju nova znanja.
Upotreba dokaza i empirijskih nalaza	Nauka ubeđuje predstavljanjem dokaza, empirijskih nalaza, logičkom diskusijom, matematičkim razmatranjima, funkcijama i opisima.	Pseudonauka pokušava da ljude ubedi da veruju u nju, oslanjajući se na zahtev za slepim verovanjem. Pokušava da preobrati, a ne da ubedi.

PREPRINT

		Zahteva da ljudi veruju u nju uprkos postojećim dokazima i argumentima, a ne zbog njih.
Kako se izdržava?	Ne prodaje neispitane ili prevarantske proizvode. Proizvodi nauke prolaze rigorozna ispitivanja i testiranja pre nego što se plasiraju na tržište.	Prodaja problematičnih proizvoda i usluga je glavni izvor prihoda. Reklame za te proizvode i usluge su pune velikih obećanja, ali proizvodi i usluge ne ispunjavaju ta obećanja, imaju vrlo slabe ili nikakve efekte, a nekada mogu biti i štetni za korisnika.

PREPRINT

Poglavlje 2. Osnovni koncepti statistike

Apstrakt. Ovo poglavlje upoznaje čitaoca sa osnovnim konceptima statistike. Počinje sa definicijom i diskusijom pojmova slučajnog događaja i verovatnoće. Sledi upoznavanje čitaoca sa time šta su entiteti i svojstva entiteta predstavljena preko konstanti i varijabli. Iza ovog se obrađuje tema organizacija podataka za potrebe statističkih računanja sa predstavljanjem koncepta matrice podataka, posebnih vrsta matrica kao što su dijagonalna matrica, matrica identiteta i vektor. Deo o populacijama i uzorcima predstavlja ove pojmove i njihove odnose uvodeći i razliku između parametara i statistika. Slede delovi u kojima se govori o osnovama uzorkovanja, uzorkovanju sa i bez vraćanja, probabilističkom i neprobabilističkom uzorkovanju, reuzorkovanju, a od ovih postupaka se predstavalja butstreping, džeknajfing i deljenje uzorka za potrebe krosvalidacije. Postupci uzorkovanja koji se najčešće sreću u litaraturi su prikazani nakon toga i ovi uključuju prosti slučajni uzorak, prigodni uzorak, stratifikovani i kvotni uzorak, uzorak snežne grudve – tzv. snoubol uzorak, namerni uzorak, klaster uzorak i druge. Nakon toga sledi predstavljanje četiri nivoa merenja – nominalnog, ordinalnog, intervalnog i racio nivoa, a potom sledi diskusija razlika između diskretnih i kontinualnih varijabli i diskretnih i kontinualnih mera.

Ključne reči: varijabla, matrica podataka, uzorak, populacija, uzorkovanje, nivoi merenja, kontinualna i diskretna merenja.

PREPRINT

Počećemo ovu uvodnu priču predstavljajući osnovne pojmove i koncepcije na kojima je statistika zasnovana. Dok smo u prethodnom poglavlju predstavili filozofske osnove nauke uopšte, a posebno statistike, da bi pomogli čitaocu da razume mesto koje statistika zauzima u svetu nauke, ovo poglavlje počinje upoznavanjem čitaoca sa samom statistikom preko predstavljanja i opisivanja ključnih ideja i koncepcija na kojima je statistika zasnovana.

2.1. Slučajni događaj

U statističkoj teoriji, **slučajni događaj je događaj čiji su ishodi nepredvidivi u principu**, događaj **koji pod istim skupom uslova može da dovede do različitih ishoda**. Slučajni događaji su

gradivni elementi stohastičkog sveta (vidi poglavlje 1.2.) jer su oni ona komponenta koja stohastički svet čini mogućim tako što stvara situacije u kojima isti skup uzroka dovodi do različitih posledica. Iako se slučajni događaj definiše kao događaj koji je u osnovi nepredvidiv, u načelu se smatra da verovatnoće tj. učestalosti sa kojima se različiti ishodi javljaju može da se opazi, a da ove učestalosti mogu da se iskoriste da se predvide učestalosti sa kojima će se ovi ishodi javljati u budućnosti. Imajući ovo u vidu, kao i diskusiju iz poglavlja 1.2. o tome kako u ovom trenutku ne znamo sa sigurnošću da li je priroda našeg univerzuma stohastička ili deterministička, treba reći da su slučajni događaji u ovom trenutku prvenstveno teorijski koncept, instrument statističke teorije koji je potreban da bi se stvorili i primenili statistički postupci. U ovom trenutku, nema nijednog poznatog načina koijim bi se proizveo događaj za koji bismo bili potpuno sigurni da je slučajan, a takođe ne postoji ni bilo koja kategorija događaja za koje možemo biti potpuno sigurni da su slučajni. Da, postoji puno klasa fenomena i klasa događaja koje tretiramo kao da su slučajni zato što ne znamo kako da ih precizno predividimo, ali nema događaja za koje možemo sa sigurnošću da znamo da neće biti moguće predvideti ih u budućnosti kada razvoj nauke dovede do novih i boljih saznanja o njihovoj prirodi.

Pa onda, ako je statistika zasnovana na slučajnim događajima, ali mi slučajne događaje ne možemo ni da proizvedemo niti umemo da ih prepoznamo, kako statistika uopšte radi? Radi tako što primenjujemo najbolju dostupnu aproksimaciju – takozvane pseudoslučajne događaje. **Pseudoslučajni događaji su događaji koji nisu stvarno slučajni ili koji verovatno nisu stvarno slučajni, ali su takvi da ih ne možemo predvideti ili se pak namerno suzdržavamo od toga da ih predvidimo** (jer bi to bilo „varanje“ i poništilo samu svrhu za koju ih koristimo) **i koji nisu uzročno povezani sa onim za šta su nam potrebni (barem kada su u pitanju pseudoslučajni brojevi koji se koriste u statističkim simulacijama)**. Najčešće korišćena vrsta pseudoslučajnih događaja su **pseudoslučajni brojevi** koji se dobijaju korišćenjem različitih uređaja koji se zajedno zovu **generatori pseudoslučajnih brojeva**. Ovi uređaji su međusobno vrlo različiti – od veoma prostih, poput npr. kockica za igre na sreću (kocke sa različitim brojem ispisanim na svakoj strani koje se koriste u igrama na sreću) ili izvlačenja loptica sa brojevima iz posude, pa sve do veoma složenih uređaja poput onih zasnovanih na radioaktivnom raspadu atomskih jezgara, fluktuacijama prirodnih emisija radio talasa ili drugih fizičkih fenomena, pa sve do složenih matematičkih postupaka (e.g. Akhshani et al., 2014; Desai et al., 2011; Liu et al., 2021). **Generatori pseudoslučajnih brojeva se u principu sastoje od dve komponente – semena** (eng. seed**) tj.**

PREPRINT

početnih vrednosti koje se unose u generator i algoritma ili matematičke funkcije koja transformiše te početne vrednosti u neki konačni broj koji je u skladu sa zahtevima generatora (npr. da broj bude u određenom opsegu). Ova činjenica da generator pseudoslučajnih brojeva koristi algoritam ili skup matematičkih operacija da bi proizveo konačnu vrednost, ponovo nas podseća da generatori pseudoslučajnih brojeva nisu pravi generatori slučajnih brojeva. Moguće je naravno, da seme bude slučajni događaj, ali do sada nije poznat ni jedan način kako bi se mogle napraviti te početne vrednosti na način koji je stvarno i potpuno sigurno slučajan. Korišćenje prirodnih događaja koje trenutno ne umemo da predvidimo ili onih koji su teško predvidivi može te početne vrednosti da učini nepredvidivim za nas, ali to nije dokaz da su te vrednosti u principu nepredvidive. Imajuću ovo u vidu, **najvažnije svojstvo generatora pseudoslučajnih brojeva su statistička svojstva brojeva koje daju i stepen do kog se svojstva tih skupova brojava slažu sa očekivanjima zasnovanim na statističkoj teoriji o tome kako bi slučajni događaji i kako bi skupovi pravih slučajnih brojeva trebalo da izgledaju**. Međutim, ma koliko skupovi pseudoslučajnih brojeva bili slični teorijskim očekivanjima o tome kako bi trebalo da izgledaju slučajni događaji ili skupovi slučajnih događaja, ne bismo smeli zaboraviti da se ovde radi samo o pseudoslučajnim brojevima niti ovakve brojeve izjednačiti sa pravim slučajnim događajima ili stvarno slučajnim skupovima brojeva koji bi bili mogući samo ako bi mogli da proizvodimo stvarno slučajne događaje. Kako je poznati matematičar i fizičar Džon Fon Nojman (John Von Neumann) svojevremeno rekao „Svako ko razmatra aritmetičke metode stvaranja slučajnih brojeva je, naravno, u stanju greha“.

2.2. Verovatnoća

Još jedan ključni koncept koji je zasnovan na pretpostavci o postojanju slučajnih događaja je verovatnoća. **Verovatnoća** se uobičajeno definiše kao stepen u kom je izražena mogućnost da će određeni slučajni događaj imati određeni ishod. Obično **se izražava kao broj od 0 do 1**, tako da **označava proporciju slučajeva kada je dati slučajni događaj imao ishod na koji se odnosi verovatnoća**. Verovatnoća od 0 znači da se dati ishod ne javlja nikad kao ishod tog slučajnog događaja, dok verovatnoća od 1 znači da dati ishod uvek nastupa kao posledica datog slučajnog događaja. Brojke između te dve krajnosti označavaju učestalost javljanja koja je između. Tako npr. verovatnoća od 0,2 znači da će se, prilikom velikog broja ponavljanja datog slučajnog događaja, u 20% slučajeva javiti dati konkretni ishod.

PREPRINT

Formulisana na ovaj način, **verovatnoća predstavlja zapravo očekivanje o tome kako će se budući događaji odvijati.** Međutim, **verovatnoća se računa na osnovu prošlih događaja i na osnovu očekivanja da će budućnost biti ista kakva je bila prošlost**. Verovatnoća se računa tako što se napravi veliki broj opservacija određenog događaja, događaja koji imamo osnova da smatramo (u dovoljnoj meri) slučajnim, i onda se broji koliko puta je dati događaj imao koji od ishoda. Na kraju se broj puta kada se javio određeni ishod podeli sa ukupnim brojem opaženih ishoda (svih vrsta), tj. sa ukupnim brojem javljanja datog slučajnog događaja i onda proglasimo da proporcija koju smo dobili na taj način predstavlja verovatnoću tog konkretnog ishoda. Kada se matematički predstavi to izgleda ovako:

Verovatnoća ishoda A = broj puta koliko smo opazili da je slučajni događaj imao ishod A / ukupan broj opaženih ishoda datog slučajnog događaja

Glavna stvar koju treba ovde uočiti je to da **ukupan broj opaženih ishoda treba da bude veliki**, zapravo veoma veliki, **da bi ovako mogli da procenjujemo verovatnoću**. Ovaj zahtev je povezan sa matematičkom teoremom poznatom kao **Zakon velikih brojeva**, a koja kaže da će **odnosi ishoda posmatranog slučajnog događaja biti utoliko bliži pravim verovatnoćama ukoliko je broj posmatranih događaja** (koji se u ovoj teoriji zovu „pokušaji“ (eng. trials)) **veći**. Kada je broj posmatranih ishoda mali, mnogo je verovatnije da će odnosi broja ishoda odstupati (bitnije) od pravih verovatnoća. Na primer, kada bacamo kockice (za igre na sreću) mnogo je verovatnije da ćemo dobiti broj 6 u 2 uzastopna bacanja, nego u 10 uzastopnih bacanja. Aki bi kockicu bacili milion puta, bilo bi praktično nemoguće da svih milion puta dobijemo broj 6. Isto tako bi bilo skoro nemoguće da 6 ne dobijemo ni jednom. A najverovatnije je da bi broj padanja kockice na svaku od njenih stranica bio približno podjednak. Ovo, naravno, pod pretpostavkom da kockica ima jednaku verovatnoću da padne na svaku od stranica (a svaka stranica je, podsetimo se, označena različitim brojem od 1 do 6).

Treba ovde reći i da, kao što je diskutovano u prethodnim poglavljima, procenjivanje verovatnoće na ovaj način **zahteva da se oslonimo na dva verovanja**, odnosno na dve pretpostavke **za koje, po pravilu, nemamo dokaze da su tačne**:

- **da je događaj koji posmatramo zaista slučajan, iako znamo da on to verovatno nije ili da nije sigurno da je slučajan.** Međutim, ovde se ipak možemo osloniti na prošle opservacije istog događaja iz kojih možemo videti da li se on u prošlosti ponašao onako

PREPRINT

kako bi očekivali da se ponaša slučajni događaj. Ako je ovo slučaj, onda možemo objaviti da je ponašanje datog događaja dovoljno slično slučajnom događaju da se događaj može tretirati kao da je slučajan.

- **da će se posmatrani događaj ili klasa događaja i u budućnosti ponašati na isti način na koji se ponašao u rošlosti**, što znači da očekujemo da će **učestalosti javljanja određenih ishoda tj. njihove verovatnoće ostati iste**. U vezi s ovim, ljudi koji primenjuju statistiku u praksi imaju običaj da kažu da su prošli trendovi često loš pokazatelj budućih trendova i uspeha budućih predviđanja. Verovanje da će se pojave ponašati u budućnosti na isti način na koji su se ponašale u prošlosti, bez dovoljnog poznavanja njihove prirode koja bi omogućila procenu valjanosti ove pretpostavke je u najmanju ruku rizično. Međutim, kada ništa bolje nemamo na raspolaganju, ovaj pristup ostaje najbolja raspoloživa opcija (pogledajte tekst o statističkim objašnjenjima u prethodnom poglavlju).

Proučavanje verovatnoća je tema kojom se bavi **grana matematike** koja se zove **teorija verovatnoće**.

Treba napomenuti da u statistici postoji i alternativni koncept verovatnoće koji je predložio Tomas Bajes [Thomas Bayes], engleski statističar iz 18. veka. Prema ovoj koncepciji, koja je poznata kao Bajesova interpretacija verovatnoće, verovatnoća predstavlja stepen uverenosti (istraživača, osobe koja verovatnoću procenjuje) u određeni događaj odnosno ishod. Ovako definisana verovatnoća centralni je koncept pristupa statistici koji se zove Bajesijanska statistika. O ovome će biti više reči u kasnijem delu ove knjige.

2.3. Entitet

Statističke opservacije i statistički proračuni se obično rade na određenim svojstvima određenih objekata. Prirode ovih objekata mogu da budu različite i one se i veoma i razlikuju u različitim oblastima nauke u kojima se koristi statistika. U društvenim naukama, ovi objekti mogu biti osobe tj. ljudi, grupe ljudi ili organizacije. U biologiji to mogu biti organizmi, biljke, životinje. A mogu biti npr. i umetničke slike, komadi opreme, subatomske čestice ili delovi objekata itd. Ovakvi objekti se obično proučavaju u skupovima tj. proučavaju se skupovi objekata i to je osnovni način kako statistika radi.

PREPRINT

Kad god hoćemo da kažemo nešto o pojedinačnom članu nekog statističkog skupa, a nećemo da određujemo prirodu tog člana statističkog skupa, nazvaćemo člana tog statističkog skupa entitetom. Drugim rečima**, entitet je naziv za pojedinačnog člana statističkog skupa bez obzira koja je priroda tog člana.** Entitet može biti osoba, životinja, objekat, deo nekog predmeta ili bilo šta. Dok god je dati objekat uključen u određeni statistički skup, možemo ga nazivati entitetom.

2.4. Varijable i konstante

Kao što je ranije rečeno, statističke opservacije i statistički proračuni rade se na određenim svojstvima određenih objekata. Ova svojstva mogu biti takva da imaju različite vrednost za različite entitete ili mogu imat istu vrednost za sve entitete. Na primer, ako je svojstvo koje posmatramo boja kose koju osoba ima, njene vrednosti mogu biti različite postojeće boje kose poput crne, plave, crvene, narandžaste, zelene itd. Ako je svojstvo koje posmatramo visina osobe, onda vrednosti tog svojstva mogu biti različite visine koje osoba može da ima, npr. 180 cm, 165cm, 182 cm, 190 cm itd. Ova svojstva mogu biti takva da svi posmatrani entiteti imaju istu vrednost na njima ili takva da različiti entiteti imaju različite vrednosti tog svojstva. **Ako svi posmatrani entiteti imaju istu vrednost nekog svojstva, takvo svojstvo nazivamo konstantom za datu grupu entiteta**. Primer konstante bi bio ako bi brojali koliko glava svaka od posmatranih osoba ima. Ako imamo u vidu to da svaka osoba ima tačno jednu glavu, broj glava po osobi bi bio primer konstante[1]. **Ako je svojstvo takvo da različiti entiteti imaju različite vrednosti tog svojstva, takvo svojstvo se zove varijabla**. Malopre pomenuta visina i boja kose su primeri varijabli. Određene dimenzije ličnosti kao, na primer, ekstraverzija takođe mogu biti primeri varijabli jer ljudi mogu biti ekstravertni u različitom stepenu i taj stepen se može proceniti psihološkim testiranjem. Cena određene robe može takođe biti varijabla ako, na primer, različite dane posmatramo kao entitete, pa onda za svaki dan beležimo cenu koju je data roba imala tog dana. I tako dalje. Ono što je važno primetiti je da se **statistički postupci prvenstveno rade na varijablama**, dok je **vrlo malo toga što se statističkim postupcima može uraditi sa konstantama**. U principu, **što se vrednosti varijabli više razlikuju, što više variraju, više je i korisnih stvari koje možemo uraditi primenom statističkih postuapaka**. **Skoro svi postojeći**

[1] Treba primetiti da iako postoje retki slučajevi u svetu gde dve glave žive povezane na isto telo, takvi slučajevi se zvanično tretiraju kao dve različite osobe, kao sijamski blizanci koji dele isto telo, a ne kao jedna osoba sa dve glave. Prema tome, broj glava po osobi je konstanta čak i ako bi takve slučajeve uzeli u obzir.

PREPRINT

statistički postupci, a posebno svi postupci koji se razmatraju u ovoj knjizi, **namenjeni su radu sa varijablama, a ne sa konstantama**. Oni daju smislene i korisne rezultate samo kada se koriste na varijablama! Kada se koriste na konstantama, obično ništa više ne možemo saznati njihovim korišćenjem od toga da imamo posla sa konstantom, a to smo znali i ranije. Prema tome, **kada govorimo o svojstvima entiteta koje posmatramo, jedna od glavnih stvari koje treba uočiti je to da li imamo posla sa konstantom ili sa varijablom.**

Još jedna stvar koju treba naglasiti je to da je važno naglasiti razliku između toga šta je varijabla, a šta je svojstvo određene varijable. Ovo zato što ova dva pojma studenti koji tek ulaze u oblast statistike često i lako pomešaju. **Varijabla je,** dakle**, svojstvo entiteta koje može da ima različite vrednosti za različite entitete ili koje može da se menja,** dok su **vrednosti varijable vrednosti koje varijabla može da ima**. Na primer, jedna varijabla može da bude masa, a njene različite vrednosti su različite moguće mase u kilogramima (ili bilo kojoj drugoj jedinici) koje bi osoba mogla da ima. Slično tome, ako je grupa studenata radila neki test koji se ocenjuje na skali od 5-10, varijabla bi mogla da bude ocena, a moguće vrednosti te varijable bi bile 5, 6, 7, 8, 9 i 10 (što su moguće ocene koje student može da dobije).

2.5. Organizacija podataka, matrica, vektor

Sada kada znamo da se statistički postupci sprovode na vrednostima varijabli čije vrednosti procenjujemo ili merimo na grupama entiteta, postavlja se pitanje kako ćemo organizovati te podatke da bude moguće raditi računanje na njima. Tipični statistički postupci koji se koriste u savremenoj nauci uključuju korišćenje prilično velikih količina podataka i upotrebu softverskih alata za izvođenje tih proračuna. U stvari, situacije u kojima se statistički proračuni koji su i praktično korisni mogu u današnje vreme razumno uraditi bez upotrebe kompjutera i statističkog softvera su veoma, veoma retke. Zbog ovoga, veoma važno pitanje je toga kako organizovati podatke da se na njima mogu raditi proračuni korišćenjem najčešćih softverskih alata.

Najtipičniji **način na koji se organizuju statistički podaci** tj. podaci o vrednostima grupe entiteta na nizu varijabli **za potrebe sprovođenja statističkih računanja** je tako što ih **predstavimo u obliku matrice podataka**. **Matrica podataka je obična tabela koja ima određenim broj redova i određeni broj kolona.** Imenuje se prema broju redova i broju kolona koje ima. Na

PREPRINT

primer, matrica sa 100 redova i 200 kolona zvaće se 100x200 matrica (ili 200x100 matrica, zavisno od pravila imenovanja koje se koristi u grupi ljudi, naučnoj oblasti ili organizaciji koja radi imenovanje). Najučestalija praksa u trenutku pisanja ove knjige je da se **entiteti u matrici predstavljaju kao redovi, a varijable kao kolone**. Na taj način, ako pratimo određeni red matrice, možemo da pročitamo vrednosti entiteta koji je predstavljen tim redom na svim varijablama. Na isti način, jedna kolona pokazuje vrednosti svih entiteta na varijabli koja je predstavljena tom kolonom. Ima, naravno, i slučajeva kada se podaci organizuju drugačije, ali najčešći način predstavljanja, onaj koji se trenutno sreće u praktično svom komercijalnom statističkom softveru je upravo ovaj – entiteti po redovima, varijable po kolonama.

Tabela 2.1. Primer matrice podataka. Entiteti su ljudi koji su radili jedan test ličnosti, varijable su njihova imena (varijabla ime), kao i njihovi skorovi na merama različitih crta ličnosti (izmišljeni podaci).

Ime	Neuroticizan N	Ekstraverzija E	Saradljivost A	Otvorenost za iskustvo O	Savesnost C
Leposava	45	55	55	65	70
Anita	22	67	50	72	62
Vladislava	37	25	45	65	45
Maida	55	32	65	42	35
Marko	25	25	32	28	60
Radoslav	52	12	42	35	28
Filip	50	70	48	47	32
Vladimir	38	65	51	59	65
Jovan	40	25	65	61	65
Goran	25	30	22	45	50
Ilona	32	45	45	68	65
Petar	35	55	65	42	58

Tabela 2.2. Još jedan primer matrice podataka. Entiteti su deonice različitih kompanija, a varijable su berzanski symbol njihovih akcija (Kompanija/simbol) i različiti pokazatelji finansijskog stanja kompanije i cene njihovih deonica, prema podacima koji su bili javno dostupni u trenutku pisanja knjige (podaci su izneti samo kao ilustracija matričnog prikaza podataka i više nisu tačni).

Kompanija/ simbol	Cena u trenutku poslednjeg	EPS	P/E odnos	P/S odnos	P/B odnos	Beta

PREPRINT

	zatvaranja berze					
AMD	107.56	2.8	38.4	9.78	18.5	2.01
BABA	173.73	8.3	20.9	3.97	3.1	0.81
MSFT	292.52	8.1	36.3	13.08	15.5	0.78
ARCB	68.37	5.1	13.3	.52	1.9	1.78
ENPH	163.48	1.27	128.9	19.87	36.1	1.17
HLX	3.63	.06	60.3	.81	.3	3.39
PLAN	59	-1.18	-50.3	18.03	32.5	
QCOM	144.41	8.01	18	5	19.9	1.32
CLVT	22.61	-0.6	-37.65	8.99	1.4	
FSLY	40.96	-1.58	-26	14.78	4.6	
EXEL	19	.29	66	5.18	2.9	1.05
W	298.38	3.06	97.7	2.09	-20	3.09
IQ	8.75	-1.09	-8.05	1.47	5.5	.8
JD	64.26	5.04	12.7	.82	3.4	.76
Skraćenice: EPS – profit po deonici; P/E – odnos cene i profita kompanije; P/S – odnos cene i ukupnih prihoda kompanije; P/B – odnos cene kompanije i knjigovodstvene vrednosti njene imovine; Beta – Beta koeficijent promenljivosti cene deonica kompanije						

Takođe, pored predstavljanja empirijskih podataka, **u obliku matrice se mogu predstaviti i rezultati statističkih proračuna**, pri čemu konkretan oblik kako će izgledati predstavljanje statističkih rezultata u obliku matrice može varirati i zavisi od sadržaja koje treba predstaviti u matrici i ličnih preferencija osobe koja te podatke predstavlja. Na primer, u tabeli 2.3. predstavljene su mere povezanosti između varijabli koje su korišćene u istraživanju i to predstavljanje je urađeno u obliku matrice. U ovoj matrici, i redovi i kolone predstavljaju varijable, a brojevi su koeficijenti koji ukazuju na stepen zajedničkog variranja tj. korelacije između varijabli. Tabela je preuzeta iz Tošić Radev & Hedrih (2017).

Tabela 2.3. Primer matrice podataka iskorišćene za predstavljanje rezultata statističkih proračuna (Tošić Radev & Hedrih, 2017). I redovi i kolone predstavlaju varijable, a brojevi u preseku redova i kolona su korelacije (mere povezanosti / statistički pokazatelji stepena zajedničkih promena vrednosti varijabli) varijabli predstavljenih u redovima i onih predstavljenih u kolonama. Broj u preseku određenog reda i određene kolone je korelacija varijable koja je predstavljena datim redom, sa varijablom koja je predstavljena datom kolonom. Podaci u matrici su prevedeni na srpski, originalna publikacija uključujući i ovu matricu su na engleskom.

PREPRINT

Koeficijenti korelacije između Multidimenzionalne skale ljubomore i procenjivanih eksternih varijabli

	Skala	Kognitivna ljubomora	Emocionalna ljubomora	Ponašajna ljubomora	Ukupna ljubomora
BFI	Globalno samopoštovanje	-0,32*	-0,11*	-0,13*	-0,25*
	Neuroticizam	0,32*	0,27*	0,27*	0,36*
	Ekstraverzija	-0,12*	0,01	0,03	-0,05
	Otvorenost za iskustvo	-0,12*	-0,14*	-0,14*	-0,16*
LAS	Saradljivost	-0,10*	-0,07	-0,13*	-0,12*
	Savesnost	-0,20*	-0,04	-0,09*	-0,15*
	Eros	-0,17*	0,05	0,05	-0,05
	Ludus	0,16*	-0,07	-0,03	0,04
	Storge	-0,02	-0,05	0,01	-0,02
	Pragma	0,07	0,06	0,13*	0,10
	Mania	0,39*	0,37*	0,44*	0,50*
	Agape	-0,09*	0,03	0,11*	0,01

Svi koeficijenti korelacije veći od 0,09 su statistički značajni bar na nivou 0,05.

Matrica koja ima isti broj redova i kolona naziva se kvadratna matrica. Kvadratne matricu se često koriste kao metod predstavljanja odnosa između varijabli, pri čemu su iste varijable predstavljene i redovima i kolonama, a u ćelije matrice su upisane mere odnosa između varijabli koje autori žele da predstave. **Matrica u kojoj broj redova i kolona nije jednak, naziva se pravougaonom matricom.**

Postoje i određeni posebni tipovi matrica koji su važni delovi različitih statističkih postupaka, gde se obično koriste za poređenje sa dobijenim strukturom podataka (dobijenih bilo računanjem, bilo empirijski) i koje zato imaju svoja posebna imena. Takve vrste matrica uključuju:

- **dijagonalne matrice su kvadratne matrice u kojima ćelije duž glavne dijagonale sadrže različite brojeve, dok je u svim ostalim ćelijama matrice broj 0.** Dijagonalna matrica se obično koristi kao **referentna matrica za opisivanje odnosa između varijabli gde ona predstavlja situaciju u kojoj su mere odnosa različitih varijabli 0, dok svaka varijabla može imati odnos sa samom sobom ili nekom drugom varijablom sa kojom je sparena koji nije nula** (sparene varijable su onda predstavljene korespondencijom reda i kolone). Uobičajeno je da u ovakvim tipovima matrica iste varijable budu predstavljene istim redom i kolonom ili da postoje dva seta sparenih varijabli, gde iz svakog para jednu varijablu predstavlja red, a drugu odgovarajuća kolona (obično kolona istog rednog broja

PREPRINT

kao i red koji predstavlja njen par). Dijagonalna matrica je takođe međuproizvod u različitim složenijim statističkim postupcima.

- **Matrica identiteta** – je **varijanta dijagonalne matrice**. **To je kvadratna matrica gde je u svim ćelijama duž glavne dijagonale broj 1, dok je u svim ostalim ćelijama matrice 0.** Slično dijagonalnoj matrici i matrica identiteta se koristi za poređenje kod opisivanja odnosa između varijabli. U takvim primerima, matrica identiteta opisuje stanje gde je mera odnosa između različitih varijabli 0, dok je mera odnosa varijable prema sebi 1.

Tabela 2.4. Dijagonalna matrica, primer. Svi elementi matrice su nule, osim elemenata na glavnoj dijagonali koji mogu biti bilo koji broj (izmišljeni podaci).

.22	0	0	0	0	0	0	0
0	45	0	0	0	0	0	0
0	0	11	0	0	0	0	0
0	0	0	22	0	0	0	0
0	0	0	0	-25	0	0	0
0	0	0	0	0	35	0	0
0	0	0	0	0	0	-2	0
0	0	0	0	0	0	0	55

Tabela 2.5. Matrica identiteta, primer. Svi elementi su nule, osim elemenata na glavnoj dijagonali koji su 1. Slično dijagonalnoj matrici, matrica identiteta se koristi kao referentno stanje za opisivanje odnosa između varijabli kako bi označila situaciju gde su mere odnosa između različitih varijabli 0 (otud nule na dijagonali), dok je pokazatelj odnosa varijable sa samom sobom 1.

1	0	0	0	0	0	0	0
0	1	0	0	0	0	0	0
0	0	1	0	0	0	0	0
0	0	0	1	0	0	0	0
0	0	0	0	1	0	0	0
0	0	0	0	0	1	0	0
0	0	0	0	0	0	1	0
0	0	0	0	0	0	0	1

PREPRINT

Matrica koja se sastoji samo od jedne kolone (bez obzira na broj redova) ili od jednog reda (bez obzira na broj kolona) zove se vektor. Vektor obično predstavlja vrednosti jedne varijable na svim entitetima ili vrednost jednog entiteta na svim varijablama, ali može predstavljati i odnose jedne varijable sa svim ostalim varijablama u posmatranom skupu, ali i različite druge stvari. Formalno gledano, to je matrica veličine 1 x n ili n x 1, gde je n bilo koji prirodni broj. U takozvanoj informacionoj geometriji, što je posebna vrsta primene geometrije u statistici, vektor se može koristiti kao način predstavljanja koordinata jednog entiteta u posmatranom statističkom prostoru i to tako što vektor sadrži vrednosti entiteta na svim varijablama, koje su zapravo koordinate u statističkom prostoru.

Tabela 2.6. Dva primera vektora. Vektori su matrice koje se sastoje od jednog reda ili jedne kolone. Prvi primer vektora sadrži Vladislavine podatke na svim varijablama (osenčeni red). Drugi primer vektora sadrži vrednosti svih osoba čiji su podaci u matrici na osobini ličnosti Otvorenost za iskustvo (osenčena kolona).

Ime	Neuroticizan N	Ekstraverzija E	Saradljivost A	Otvorenost za iskustvo O	Savesnost C
Leposava	45	55	55	65	70
Anita	22	67	50	72	62
Vladislava	37	25	45	65	45
Maida	55	32	65	42	35
Marko	25	25	32	28	60
Radoslav	52	12	42	35	28
Filip	50	70	48	47	32
Vladimir	38	65	51	59	65
Jovan	40	25	65	61	65
Goran	25	30	22	45	50
Ilona	32	45	45	68	65
Petar	35	55	65	42	58

2.6. Populacija i uzorak, parametri i statistici

Statistički postupci se generalno sprovode sa ciljem da se izvedu zaključci ili da se nešto otkrije u vezi velike grupe ili klase entiteta. **Velika grupa entiteta koju želimo da proučimo naziva se**

PREPRINT

populacija. Na primer, populacija mogu biti svi ljudi koji žive u određenoj zemlji ili u određenoj oblasti, svi ljudi svuda, sve životinje određene vrste u određenoj oblasti ili na celoj planeti. Ali populacije mogu biti i neživi objekti, kao na primer, sva voda u reci ili sva voda u određenom jezeru ili moru, sve slike naslikane određenim stilom slikarstva, svi organi određene vrste kod organizama određene vrste, sve cene određene robe na različite dane ili u različitim vremenskim trenucima, sve cene različitih vrsta robe, sve mere određenog svojstva objekata itd. Populacije mogu da budu veoma različite po prirodi, ali **da bi se određena grupa smatrala populacijom u statističkom smislu, potrebno je da ona bude precizno definisana**, odnosno da kada posmatramo bilo koji pojedinačni entitet, **bude moguće potpuno jasno odrediti da li dati entitet pripada datoj populaciji** ili ne**. Populacije po svojoj prirodi mogu biti ograničene ili mogu biti (praktično) neograničene.** Na primer, ako defiinišemo našu populaciju kao sve učenike koji idu u određeno odeljenje određene škole, to je jasno određena populacija jer je jasno da tu možemo napraviti listu svih članova populacije i da se ta lista najverovatnije neće menjati dok ta populacija postoji tj. dok se školski program ne završi, a odeljenje bude raspušteno. Moguće je, naravno, da novi učenici uđu u odeljenje ili da neki promene odeljenje ili školu, ali u bilo kojoj konkretnoj vremenskoj tački, tačno se zna koji učenici pripadaju tom odeljenju i ovo se po pravilu ne menja ili se menja vrlo malo. Nasuprot ovome, možemo takođe da odredimo ciljnu populaciju kao „svi ljudi“ ili „svi ljudi koji žive na određenoj teritoriji“. Ako uzmemo „sve ljude“ za populaciju koja nas interesuje, to onda uključuje sve ljude koji su ikad živeli, kao i one koji tek treba da se rode. To je populacija za koju je sasvim jasno da ne možemo napraviti spisak svih članova, jer i kada bi imali spisak svih ljudi koji su ikada živeli, nemamo nikakvog načina da znamo ko će se tačno sve roditi u budućnosti. Međutim, kako ćemo videti kasnije, populacija koja bi bila tako definisana bi bila previše široka da bi je zaista bilo moguće proučavati statističkim metodama, te bi je za istraživačke svrhe svakako bilo potrebno suziti i specifikovati. Mogli bismo npr. da definišemo populaciju koja nas interesuje kao „svi ljudi koji žive na određenoj teritoriji (državi, regionu..) u trenutku sprovođenja istraživanja“. Iako je ovo dosta uže određenje, pravljenje spiska članova ovakve populacije bi takođe bilo praktično nemoguće (osim ako teritorija nije vrlo mala), zato što pravljenje spiska zahteva vreme, a tokom vremena koje je potrebno da se napravi spisak, sadržaj tog spiska bi se promenio, jer bi neki ljudi umrli, neki se rodili, neki se odselili, neki doselili itd. I dok bi mi uspeli da sastavimo spisak, taj spisak već više ne bi bio sasvim tačan. Takođe je sasvim verovatno da bi, bez nekog fantastičnog uređaja koji trenutno ne postoji, bilo nemoguće ili bar

PREPRINT

veoma, veoma teško da uspešno detektujemo sve ljude, bez izuzetka, koji žive na nekoj većoj teritoriji, zato što je vrlo verovatno da nisu svi ti ljudi registrovani i bilo bi veoma teško i nemoguće da nađemo baš sve ljude. Stvari postaju još komplikovanije ako je naša populacija „voda u određenoj reci“ ili „insekti određene vrste koji žive u određenoj velikoj oblasti“ itd. Međutim, dobra vest je da **istraživanje možemo da sprovedemo čak i ako ne raspolažemo spiskom svih članova populacije dokle god je populacija definisana dovoljno precizno da možemo jasno da razlikujemo entitete koji pripadaju datoj populaciji od onih koji ne pripadaju**. Razlika između ograničenih i neograničenih populacija ulazi u igru prilikom izvođenja statističkih zaključaka o njima, a koji su zasnovani na proučavanju dela populacije.

U idealnoj situaciji, kada želimo da proučimo neko svojstvo populacije tj. da proučimo neke varijable u populaciji, mi bi izmerili vrednosti tih varijabli na svim članovima populacije, a onda zaključke zasnovali na tim rezultatima. Ovo je praktično moguće izvesti u situacijama kada su populacije koje proučavamo ograničene i dovoljno male. Ako populacija nije mala, uzimanje mera (odnosno procena) vrednosti varijabli od svih članova populacije vrlo brzo postaje nepraktičan i skup poduhvat, a ako populacija nije ograničena, postaje i praktično nemoguć.

Međutim, uzimanje mera/procena svih članova populacije najčešće i nije neophodno. **Ako je naš istraživački cilj da dođemo do saznanja o svojstvima populacije kao celine i nismo zainteresovani za individualne mere** tj. vrednosti svakog pojedinačnog člana populacije, **proučavanje svih članova populacije postaje nepotrebno**. Na primer, ako želimo da saznamo koja je tipična cena paradajza na određenoj pijaci u praksi nije neophodno da pitamo svakog prodavca paradajza za cenu paradajza koji prodaje. Može da bude sasvim dovoljno da zabeležimo cene paradajza kod određenog broja nasumice izabranih prodavaca i da iz toga dobijemo prilično dobru procenu toga u kom se rasponu kreće cena paradajza na toj pijaci. Ako želimo da saznamo koji je udeo, na primer, nekog hemijskog jedinjenja u određenom jezeru, ne moramo da prebrojimo svaki molekul tog jedinjenja u tom jezeru. Obično je dovoljno da uzmemo po flašu vode iz različitih tačaka na jezeru i iz toga možemo dobiti prilično dobru procenu toga koliko supstance koju ispitujemo ima u jezeru. Ako nam je cilj da procenimo neki opšti nivo znanja određenog stranog jezika u populaciji, to možemo da uradimo tako što ćemo proceniti znanje stranog jezika grupe ljudi koje smo slučajnim putem izabrali iz populacije. Statističkim rečnikom rečeno – proučavamo uzorak da bi na osnovu njega izveli zaključke o populaciji.

PREPRINT

Deo populacije koji smo odabrali za sprovođenje istraživanja odnosno ispitivanja naziva se uzorak. Opšta ideja u osnovu proučavanja uzorka je ta da ćemo proučavajući uzorak naučiti ono što nas interesuje o populaciji zato što je uzorak dovoljno sličan populaciji da na osnovu toga što smo saznali proučavanjem uzorka možemo da izvedemo zaključke o stanju stvari u populaciji. Naravno, svojstva uzorka mogu ponekad da ne budu skroz ista kao svojstva populacije, ali možemo očekivati da u većini slučajeva neće biti previše različita. Ako uzmemo uzorak koji je dovoljno veliki i ako uradimo šta možemo da izbegnemo da namerno odaberemo uzorak koji je različit od populacije, verovatno je da će uzorak biti dovoljno sličan populaciji da na osnovu njega možemo valjano izvesti puno korisnih zaključaka o populaciji.

Kakav bi trebao da bude dobar uzorak? **U idealnom slučaju, svojstva uzorka bi trebalo da budu ista kao svojstva populacije posmatrane kao celina**, **s tim da jedina razlika bude to što je uzorak dovoljno mali da ga je moguće proučiti** (za razliku od populacije, u većini slučajeva). **Uzorak čija su sva svojstva ista kao svojstva populacije** (osim veličine tj. broja entiteta koji ga čine) naziva se **reprezentativni uzorak**. **Kada radimo istraživanje** sa ciljem da otkrijemo nešto o svojstvima populacije, **u idealnom slučaju bi želeli da to uradimo tako što bi ispitali uzorak koji je što je moguće više reprezentativan za tu populaciju**. Ali da li možemo zaista da znamo da li je uzorak koji smo odabrali reprezentativan za populaciju i u kojoj meri? Ovo je donekle kao ono pitanje o tome da li je starije kokoška ili jaje – da bi sa sigurnošću znali da li je naš uzorak reprezentativan, mi bi morali da ga uporedimo sa populacijom. Međutim, da bi ga uporedili sa populacijom, morali bi da posedujemo podatke o populaciji na onim istim varijablama koje hoćemo da ispitamo na uzorku. Ali kada bi imali podatke na populaciji, nama uopšte ne bi bio potreban uzorak, pa bi ovakvo pitanje bilo sasvim suvišno. To u praktičnim situacijama sprovođenja istraživanja znači da **nikad ne možemo zasigurno da znamo da li je naš uzorak reprezentativan za populaciju koju želimo da reprezentuje.** Naravno, u većini slučajeva biće **na raspolaganju određeni heuristici[2] koji mogu da nam pomognu da primetimo situacije kada podaci sa uzorka baš, baš mnogo „promašuju“ stanje u populaciji**. Na primer, ako bi sastavili uzorak ljudi i videli da su svi u tom uzorku mladi i ako, pri tom, znamo da populacija koju želimo da proučavamo ima mnogo veću varijabilnost u pogledu starosti, bilo bi očigledno da naš uzorak nije reprezentativan za populaciju. Ista bi situacija bila i ako bi sastavili uzorak u kom su

[2] Heuristici su elementi, podaci ili postupci koji omogućavaju da se brzo i bez dovoljno podataka izvede zaključak u nečemu, ali ne sa sigurnošću. Zaključci izvedeni korišćenjem heuristika mogu da ne budu tačni.

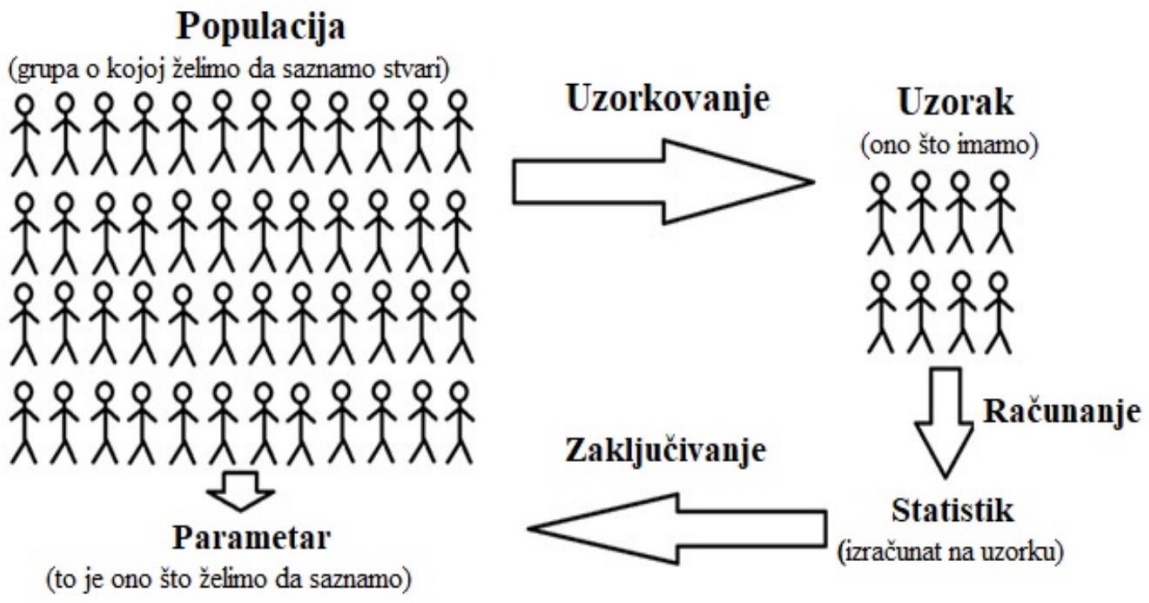
PREPRINT

samo muškarci, ali kroz neformalno posmatranje populacije uočimo da naša populacija sadrži ljude oba pola. Međutim, u slučajevima kada razlike između populacije i uzorka nisu tako ekstremne, očigledne ili izražene, ne bi bilo načina da ustanovimo da li je naš uzorak reprezentativan. Ovo znači da, **iako postoje slučajevi kada možemo uočiti da uzorak nije reprezentativan, ne postoji način kojim bi nedvosmisleno potvrdili da uzorak jeste reprezentativan**. Postoje različiti načini kako istraživači mogu probati da maksimizuju verovatnoću da se dobije reprezentativan uzorak, kao što je npr. upotreba određenih tehnika prikupljanja uzorka koje su u prošlosti često rezultirali uzorcima na osnovu kojih je bilo moguće uspešno predviđanje ponašanja populacije ili kao što je merenje različitih svojstava uzorka i njihovo poređenje sa tim istim svojstvima ranije skupljanih uzoraka, **ali ni jedan od ovih postupaka ne garantuje da će uzorak koji je skupljen na određeni način ili koji je prošao određenu statističku procenu biti reprezentativan**. Ništa osim direktnog poređenja uzorka sa populacijom u pogledu vrednosti varijabli koje nas interesuju nije dovoljno da zasigurno potvrdi da je uzorak reprezentativan, a, kao što je napred već rečeno, u praktičnim situacijama istraživanja izvođenje ovakvog poređenja nije raspoloživo.

To znači da **kako god napravili uzorak, uvek postoji šansa da će se taj uzorak manje ili više razlikovati od populacije koju treba da reprezentuje**. Pa ako prihvatimo da se mogućnost javljanja ovakve razlike ne može izbeći, to znači da ta mogućnost mora na ovaj ili onaj način da bude uključena u način na koji procenjujemo svojstva populacije na osnovu uzorka. Iz ovog razloga, statističari koriste pojmove **statistika i parametara. Kada procenjujemo određena svojstva uzorka**, na primer, kada računamo prosečnu vrednost određene varijable na uzorku ili kada brojimo koliko entiteta u uzorku ima određenu vrednost na nekoj varijabli, rezultati ovakvih postupaka zovu se **„statistici“** (jednina je „statistik“). **Kada to isto radimo na populaciji, rezultat se zove „parametar“.** Ono kako se tipično radi je da se **statistici računaju direktno iz uzorka, pa se onda zaključuje o vrednosti parametara na osnovu vrednosti tih statistika**. Obično **parametre nije moguće direktno izračunati iz podataka**, jer bi njihovo računanje zahtevalo da imamo podatke o svim članovima populacije. Metodologija zaključivanja o vrednostima parametara na osnovu statistika je tema oblasti statistike koja se zove **statistika zaključivanja**. Osnove statistike zaključivanja će biti predstavljene u drugoj delu ove knjige.

PREPRINT

Slika 2.7. Populacija i uzorak, kako stvari rade u statistici.



2.7. Uzorkovanje

Postupak izbora uzorka entiteta koji će biti korišćen u proučavanju populacije zove se **uzorkovanje**. Kako je ranije pomenuto, da bi istraživanje na uzorku bilo uspešno i da bi podaci koji se na uzorku dobiju bili korisni za upoznavanje svojstava populacije, potrebno je da uzorak ima određena svojstva od kojih je obično najvažnije to da bude koliko god je to moguće reprezentativan za populaciju o kojoj želimo da izvodimo zaključke[3]. Da bi postigli ovaj cilj postoje različite tehnike za odabir uzorka – tehnike uzorkovanja koje se razlikuju po načinima na koji se entiteti biraju za uzorak, po tome koliko je lako/teško formirati uzorak na taj način, a najčešće i po tome koliko je izgledno da će uzorak dobijen na taj način biti reprezentativan. Kada su u pitanju tehnike uzorkovanja, možemo povući razliku između opštih pristupa uzorkovanju i posebnih tehnika uzorkovanja.

[3] Treba primetiti da postoje situacije kada je cilj studije da se detaljnije prouče određeni delovi populacije i u tim situacijama idealni uzorak može da bude i uzorak koji se po svojim opštim karakteristikama razlikuje od populacije u celini i to obično u pogledu udela entiteta koji predstavljaju određene delove populacije i to tako što će neke kategorije entiteta biti više zastupljene nego što je to slučaj kod opšte populacije, ali i ovo je slučaj gde se od uzorka očekuje da bude reprezentativan za neki veći skup o kom želimo da izvodimo zaključke na osnovu uzorka. U takvim situacijama samo treba ostati svestan toga za koju smo populaciju tačno želeli da uzorak koji pravimo bude reprezentativan.

Kada su u pitanju opšti pristupi uzorkovanju, prvo treba podvući razliku između uzorkovanja bez vraćanja i uzorkovanja sa vraćanjem.

Uzorkovanje bez vraćanja je pristup uzorkovanju u kome se entiteti izabrani za uzorak prebacuju iz populacije u uzorak i onda ne mogu biti ponovo izabrani u uzorak u narednom koraku uzorkovanja, jer nakon uključivanja u uzorak više nisu u populaciji iz koje se uzorak bira. **To znači da svaki entitet iz populacije može da bude uključen u jedan isti uzorak samo jednom**. Glavni efekat ovakvog pristupa je **pojava** takozvanih **„rastućih verovatnoća“.** Ako entitete biramo u uzorak slučajnim izborom, takvim da svaki entitet ima jednaku verovatnoću da bude izabran u uzorak, kako uzorkovanje napreduje verovatnoća svakog pojedinačnog entiteta iz populacije da bude izabran u uzorak raste kako se ukupan broj entiteta od kojih se bira smanjuje. Do ovog smanjivanja dolazi jer su u koracima pre datog koraka neki entiteti već uključeni u uzorak i nisu više dostupni za biranje. Na primer, ako imamo populaciju od 1000 entiteta i slučajnim putem biramo one koji će ući u uzorak tako da svaki entitet iz populacije ima jednaku šansu da bude odabran, u trenutku kada biramo prvi entitet za uzorak, svaki entitet u uzorku ima verovatnoću 1/1000 da bude odabran (zato što biramo 1 od 1000, a svaki entitet ima istu verovatnoću da bude odabran, ta verovatnoća za svaki entitet je 1/1000). Međutim, već u sledećem koraku biramo 1 entitet od 999, što verovatnoću svakog entiteta da bude odabran povećava na 1/999. U trećem koraku već biramo 1 od 998, pa 1 od 997 i tako dalje. Ovo je posebno važno u situacijama kada je veličina populacije ograničena, a veličina uzorka je supstantivan deo veličine populacije. S druge strane, **ovaj efekat je praktično zanemarljiv u situacijama kada je populacija neograničena** ili **kada je veličina uzorka zanemarljivo mala u odnosu na populaciju**. Još jedna važna posledica primene ove metode je to da **ona sprečava da se napravi uzorak koji je veći od populacije** (tj. koji ima više entiteta nego što ih ima u populaciji).

Uzorkovanje sa vraćanjem je pristup uzorkovanju u kome, **nakon što entitet bude izabran za uključivanje u uzorak** (tj. nakon što bude uzorkovan), **taj isti entitet „vraćamo“ u populaciju tako da ima priliku da bude opet izabran**. Na ovaj način **isti entitet može da bude odabran više puta u isti uzorak**. Rezultat ovog je to da isti entitet iz populacije, može biti predstavljen kao više različitih elemenata uzorka (kao više različitih entiteta u uzorku). Iako ovo na prvi pogled izgleda kao veoma čudna ideja, ovaj pristup uzorkovanju ima različita korisna svojstva. Prvo i verovatno najočiglednije je to da, za razliku od uzorkovanja vez vraćanja, kada radimo slučajno

PREPRINT

uzorkovanje sa jednakim verovatnoćama izbora za sve entitete, **ne dolazi do pojave rastućih verovatnoća**, odnosno verovatnoća svakog entiteta da bude izabran u uzorak ostaje ista u svim koracima. Kada radimo slučajno uzorkovanje bez vraćanja iz, na primer, populacije od 1000 ljudi i sa tim da svi entiteti imaju jednake verovatnoće da budu izabrani u uzorak, svaki entitet će imati 1/1000 verovatnoću da bude izabran i u prvom koraku uzorkovanja i u svim kasnijim koracima. Do ovoga dolazi zbog činjenice da **entiteti koji su izabrani u uzorak u jednom koraku, učestvuju u izboru i u narednim koracima, tako da se u svakom koraku bira iz punog sastava populacije**. Još jedno svojstvo ovog pristupa je to da nam **on omogućava da iz bilo koje veličine populacije dobijemo uzorak bilo koje veličine**. To znači da **možemo da pravimo i uzorke koji su veći od populacije**, kao i da **možemo da imamo veliki, neograničeni broj uzoraka iz iste ograničene populacije**, a da se pri tom sastav tih uzoraka međusobno razlikuje više nego što bi to mogao da bude slučaj kod uzorkovanja bez vraćanja. Na primer, ako bi napravili 100 uzoraka koji su iste veličine kao i populacije uzorkovanjem bez vraćanja, svi ti uzorci bi nužno bili identični jer bi se sastojali od potpuno istih entiteta. Međutim, ako bi ovo uradili uzorkovanjem sa vraćanjem, ovi uzorci bi se mogli razlikovati jer bi verovatno u svakom od tih uzoraka bilo višestrukih kopija istih entiteta, ali bi bilo različito to kojih tačno entiteta i u kom broju, dok bi takođe bilo moguće da neki entiteti uopšte ne budu uključeni u neke uzorke, što sve stvara prilike da se uzorci razlikuju međusobno.

Još jedno bitno opšte svojstvo postupaka uzorkovanja je to da li je postupak izbora entiteta probabilistički ili neprobabilistički. Kod **probabilističkih** postupaka uzorkovanja, izbor entiteta je zasnovan na nekom stohastičkom procesu (ili, u praksi, dostupnoj aproskimaciiji takvog procesa, vidite poglavlje 1.), gde svaki entitet u populaciji ima određenu verovatnoću da bude izabran u uzorak. Ove verovatnoće mogu da se razlikuju za različite entitete ili mogu da budu jednake, ali da bi se postupak uzorkovanja smatrao probabilističkim, uzorkovanje mora da bude zasnovano na verovatnoćama. Nasuprot tome, **neprobabilistički** postupci uzorkovanja koriste različite procedure uzorkovanja u kojima izbor entiteta nije zasnovan na verovatnoći.

U ovoj diskusiji valja pomenuti i koncept **reuzorkovanja tj. ponovnog uzorkovanja.** Reuzorkovanje tj. ponovno uzorkovanje odnosi se na **skup postupaka gde se novi uzorak ili novi uzorci prave iz postojećeg uzorka,** a to se **tipično radi da bi se simuliralo to šta bi se desilo ako bi dodatni uzorci bili uzorkovani iz iste populacije**. Kako znamo da će se, bez obzira na to

PREPRINT

kako pravimo uzorak, njegova svojstva manje ili više razlikovati od svojstava populacije, reuzorkovanje može biti zgodan način da se procene verovatne razmere tih razlika, te da se tako dobiju precizniji zaključci o tome šta možemo očekivati da se dobije u budućim istraživanjima na istu temu na novim uzorcima. Iako metoda reuzorkovanja ima puno, u naučnim istraživanjima i statističkom softveru se najčešće sreću butstreping (eng. bootstrapping), džeknajfing (eng. jackknifing) i krosvalidacija.

Butstreping je naziv za skup postupaka gde **se jedan (obično veoma veliki) broj novih uzoraka uzorkuje sa vraćanjem iz uzorka (jednog) koji nam je na raspolaganju** (e.g. Good, 2006). U osnovi**, postojeći uzorak se tretira kao da je populacija, a onda se potrebni broj novih uzoraka uzorkuje iz njega**. Ovo se obično radi postupcima slučajnog uzorkovanja (videti kasnije), a zbog toga što se radi uzorkovanjem sa vraćanjem, nema ograničenja u broju novih uzoraka koji se mogu dobiti, kao ni u pogledu toga koliko ti novi uzorci mogu da budu veliki. Butstreping se najtipičnije koristi u postupcima zaključivanja o parametru (tj. o vrednostima statističkih mera u populaciji) na osnovu statistika dobijenih na uzorku. Na ovaj način, butstreping se posmatra kao simulacija onoga što bi se desilo ako bi veliki broj uzoraka bio uzet iz iste populacije, s tom razlikom da se tu uzorkovanje ne radi iz ciljne populacije, nego samo iz uzorka koji je uzet iz te ciljne populacije. O ovome će biti više reči u delu knjige posvećenom statistici zaključivanja. Postupak bustrepinga je prvi predložio Efron (1979), a onda ga je razvio niz istraživača u različitim oblicima. U trenutku pisanja ove knjige, butstreping polako postaje jedan od uobičajenih postupaka u statistici zaključivanja koji se sve više i više sreće u statističkom softveru koji je u širokoj upotrebi.

Džeknajfing je postupak u kom se veći broj uzoraka formira tako što se iz uzorka koji imamo isključuje određeni deo ispitanika. Ovo je u suštini isti postupak kao da uzorkovanjem bez vraćanja formiramo više uzoraka iz originalnog uzorka koji smo prikupili. Ovaj postupak je 1949 predložio Ouenouille, a dalje ga je razvio Tukey 50ih godina 20. veka (Miller, 1974). Ideja u osnovi ovog postupka je da se proceni koliko se statistici uzorka menjaju kada se isključi deo uzorka i, na taj način, da se napravi procena šta bi se dobilo kada bi se neki drugi uzorak uzeo iz iste populacije i statistici izračunali na njemu.

Krosvalidacija je postupak u kome se uzorak deli na dva dela. Statistički postupci i zaključci na njima zasnovani se onda izvode na jednom delu uzorka, a na drugom delu uzorka se ispituje da li

PREPRINT

ti zaključci koji su izvedeni iz prvog dela važe i tu. To je još jedan način da se proceni koliko su validne generalizacije sa uzorka na populaciju ili, praktično, koliko se može očekivati da bi zaključci izvedeni iz ispitivanog uzorka bili validni i na drugim uzorcima dobijenim iz iste populacije.

2.8. Vrste uzoraka

Sada ćemo proći kroz posebne vrste uzoraka tj. kroz posebne postupke uzorkovanja i ukratko prodiskutovati svaki od njih. Vrste uzoraka se razlikuju prema načinima na koji su entiteti selektovani u uzorak i, kroz to, oni se razlikuju i prema tome koliko je teško napraviti određenu vrstu uzorka, kao i tome koliko je verovatno da se desi da se svojstva tako odabranog uzorka bitnije razlikuju od populacije. Ovo nije iscrpna lista svih mogućih vrsta uzoraka, zato što bi takva lista bila neograničana, jer se svaki pojedinačni način na koji neko odluči da sprovede uzorkovanje zapravo može tretirati kao posebna tehnika uzorkovanja. Ovo što predstavljamo na ovom mestu su samo neki od najčešće sretanih postupaka uzorkovanja:

Prosti slučajni uzorak se pravi tako što **se krene od spiska svih entiteta u populaciji koja se izučava, a onda se generatorom slučajnih brojeva biraju oni koji će ući u uzorak.** Ovaj postupak generisanja slučajnih brojeva se organizuje tako da svi entiteti u populaciji imaju jednaku verovatnoću da budu izabrani u uzorak. Prosti slučajni uzorak je ono na šta ljudi obično misle kada govore o slučajnim uzorcima. Na neki način, ova vrsta uzorka je i "uzor" uzorka u statistici, jer je osnova većine statističkih teorema i modela zaključivanja, pri čemu se sve ostale vrste uzorkovanja posmatraju kao bolje ili gore aproksimacije ove vrste uzorka. Međutim, u praktičnim situacijama, **ovu vrstu uzorka je uglavnom nemoguće napraviti**. Prvi razlog za ovo je to da prosti slučajni uzorak zahteva da možemo da generišemo slučajne brojeve. Ovo je, naravno, nemoguće, te zato koristimo pseudoslučajne brojeve umesto slučajnih. Takođe, ova vrsta uzorka zahteva da populacija koju proučavamo bude ograničena i da imamo na raspolaganju kompletnu listu svih entiteta u populaciji. To praktično znači da prosti slučajni uzorak nije metoda koju možemo da primenimo u situacijama kada imamo neograničenu populaciju ili kada nemamo spisak svih članova populacije. Konačno, ova vrsta uzorka zahteva da imamo moć da obezbedimo da svaki entitet koji smo odabrali procesom generisanja slučajnih brojeva zaista i postane deo uzorka. Kada su ovi entiteti ljudi, kao što je na primer slučaj u društvenim naukama, to je, naravno, nemoguće,

PREPRINT

jer ljudi i mogu i znaju da odbiju da učestvuju u istraživanju. Ako entiteti nisu ljudi, već neki drugi objekti iz prirode, kao što su npr. molekuli vode, populacija može da ne bude ograničena, ali i ako je ograničena, kao npr. ako je naša populacija drveće u određenoj šumi, obično nije previše praktično da se pravi spisak svih entiteta (npr. popis svog pojedinačnog drveća u nekoj velikoj šumi) samo za potrebe uzorkovanja. Imajući ovo sve u vidu sledi da, iako je prosti slučajni uzorak, vrsta "uzornog uzorka" u statistici, takav uzorak se u praksi retko kada može zaista napraviti. Međutim, ovo i nije preterano veliki problem jer postoje druge tehnike uzorkovanja koje je mnogo jednostavnije realizovati u praksi, a koje tipično daju uzorke koji nisu mnogo gori u pogledu reprezentativnosti u odnosu na slučajne uzorke.

Prigodni uzorak se pravi tako što se **prosto u uzorak uključe oni entiteti koje je najzgodnije uzeti u uzorak, koji su istraživačima najlakše dostupni**. Nasuprot prostom slučajnom uzorku, **prigodno uzorkovanje se smatra za najgoru tehniku uzorkovanja**. Glavni problem sa prigodnim uzorkovanjem je to što **kvalitet uzorka** koji se dobije tj. to koliko će dobijeni uzorak biti reprezentativan za populaciju, **može jako puno da varira**. Nekada prigodno uzorkovanje može proizvesti uzorke koji su vrlo slični populaciji, ali u drugim situacijama može proizvesti uzorke koji veoma mnogo "promašuju" karakteristike populacije. A takođe je često teško ili nemoguće utvrditi koja od ove dve situacije je u pitanju samo na osnovu posmatranja pojedinačnog uzorka. S druge strane, u različitim društvenim naukama, prigodno uzorkovanje zna da bude potpomognuto ličnim iskustvom istraživača na koje se istraživač oslanja da bi poboljšao sličnost između uzorka i populacije. Na primer, iskusni istraživači će često imati neki opšti uvid u to kakva su svojstva populacije koju proučavaju i onda na osnovu toga moći da znaju i u kom lako dostupnom delu populacije se mogu naći tipični predstavnici populacije u celini, a takođe mogu imati i uvide o tome odakle se tačno može lako uzeti uzorak koji neće biti previše različit od populacije. Postoji, naravno, rizik i od toga da takva znanja budu iskorišćena sa suprotnim ciljem – istraživač čije postupke vodi cilj koji nije naučni (nego je politički, ideološki, pseudonaučni, obmanjivački....), "istraživač" koji ne želi stvarno da istraži stvari već je samo zainteresovan da dobije rezultate koji mu/joj odgovaraju (sa ciljem da obmane naučnu javnost) može odabrati prigodni uzorak za koji zna da se razlikuje od populacije, ali da ga ipak izabere zato što očekuje da će na njemu dobiti rezultate kakve priželjkuje. Na ovaj način, u najvažnijim aspektima, upotreba i kvalitet prigodnog uzorka zavise od iskustva istraživača, ali i njegovog/njenog naučnog poštenja i integriteta. Ipak, i pored svega ovoga, **prigodno uzorkovanje je daleko najjeftinija,**

PREPRINT

najjednostavnija i najmanje zahtevna tehnika uzorkovanja, tehnika koja je metoda izbora kad god su resursi kojima raspolažu istraživači skromni. To je i glavni razlog zašto je prigodno uzorkovanje **metoda koja se sreće u daleko najvećem broju istraživanja u društvenim naukama**. Međutim, to je takođe i razlog zašto se ona ređe sreće u studijama sa najvećim naučnim uticajem u odnosu na to koliko je zastupljena u tipičnim naučnim istraživanjima.

Stratifikovani uzorak se pravi tako što se **populacija podeli u stratume tj. u subpopulacije koje su zasnovane na određenoj karakteristici, a onda se entiteti koji će biti uključeni u uzorak biraju iz svakog stratuma nekom drugom tehnikom uzorkovanja** (slučajno uzorkovanje, prigodno itd.). Ideja u osnovi ovog postupka je da, ako se populacija sastoji od jasno različitih podgrupa koje su važne za temu istraživanja, stratifikovano uzorkovanje osigurava da sve ove podgrupe budu adekvatno zastupljene u uzorku. U idealnom slučaju, podela na stratume će pratiti neku očiglednu podelu u podgrupe, tako da bude lako sprovesti uzorkovanje unutar svakog stratuma (primer mogu biti različiti razredi unutar škole ili različite opštine ili regioni unutar države). Stratifikovani uzorci mogu biti **proporcionalni**, kada je udeo eniteta iz svake populacije u uzorku proporcionalan udelu tog stratuma u populaciji ili **disproporcionalni**, kada udeo entiteta iz pojedinačnih stratuma u uzorku nije proporcionalan njihovom udelu u populaciji. Disproporcionalno uzorkovanje može biti korisno u situacijama kada je potrebno da imamo više entiteta iz određenog stratuma, kao na primer onda kada su neki stratumi suviše mali, pa bi, ako bi uzorkovanje bilo proporcionalno, u konačnom uzorku bilo premalo entiteta iz datog stratuma za valjano sprovođenje istraživanja. Treba takođe napomenuti da, iako je disproporcionalno uzorkovanje sasvim legitimna tehnika uzorkovanja, važno je da istraživači koji realizuju studiju, kako oni koji rade na prikupljanju podataka, tako i oni koji ih interpretiraju, sve vreme tokom svog rada imaju na umu da rade sa disproporcionalnim uzorkom i da se u skladu s tim uzdrže od toga da na osnovu takvog uzorka izvode zaključke o populaciji u celini onako kako bi to radili da je uzorak proporcionalan. Kada se radi sa disproporcionalnim uzorkom, činjenica da je uzorak disproporcionalan mora da se uzme u obzir i da se ta disproporcionalnost kompenzuje prilikom izvođenja zaključaka o populaciji u celini na osnovu takvog uzorka. Ovo se može učiniti npr. tako što će se različitim stratumima pripisati različiti "ponderi" ili "težine" prilikom računanja ukupnih rezultata (praktično – da se uticaj prezastupljenih stratuma na ukupan rezultat smanji na nivo koji odgovara njihovom udelu u populaciji, a onih manje zastupljenih u uzorku da se poveća na nivo koji odgovara njihovom udelu u populaciji).

PREPRINT

Kvotni uzorak nastaje **korišćenjem sistema kvota pri odabiru uzorka**. Kvote su brojevi ispitanika sa određenim vrednostima na varijablama koje su važne istraživačima koje treba uključiti u uzorak. **Kvote se određuju prema tome koliki je udeo u populaciji entiteta sa određenim vrednostima na posmatranim varijablama tako da se obezbedi da i u uzorku proporcije entiteta sa tim vrednostima budu jednake kao u populaciji**. Ideja u osnovi stvaranja ovakvog uzorka je to da ako se uzorak napravi tako da liči na populaciju u pogledu raspodele vrednosti određenih ključnih varijabli, onda je verovatnije da će ličiti na populaciju i u pogledu svih ostalih varijabli, uključujući i one koje treba da budu proučavane na tom uzorku. Naravno, da bi kvotni uzorak bilo moguće napraviti, distribucije vrednosti varijabli koje će biti korišćene za stvaranja kvota u populaciji moraju biti poznate. Kada se kvote definišu. svaki dostupni entitet koji ispunjava zahteve neke od kvota biva izabran u uzorak (dok se kvote ne popune). Na primer, ako znamo da se populacija koju želimo da proučimo sastoji od 15% studenata univerziteta i 85% ljudi koji nisu studenti, možemo sa tim podatkom napraviti kvotni uzorak od 100 ljudi tako što ćemo definisati da on treba da se sastoji od 15 studenata i od 85 ljudi koji nisu studenti. Ovi brojevi studenata i ljudi koji nisu studenti koje treba uključiti u uzorak predstavljaju kvote za dati uzorak. Ovo je najprostiji primer kvotnog uzorka, primer u kom se za formiranje kvota koriste samo dve kategorije i jedna varijabla. Nasuprot tome, tipični kvotni uzorci uključuju kvote zasnovane na više varijabli i obično imaju veći broj kategorija. U takvim situacijama, kada se kvote formiraju na osnovu više varijabli, one mogu biti vezane ili nevezane. **Nevezane kvote** su **kvote koje se određuju za svaku varijablu na osnovu koje se prave kvote posebno, bez ukrštanja**. Na primer, ako bi pravili kvote zasnovane na tome da li je osoba student univerziteta ili nije (jedna varijabla) i na tome da li je osoba iznad ili ispod 40 godina starosti (druga varijabla), nevezane kvote bi određivale koliko studenata i koliko ljudi koji nisu studenti treba da imamo u uzorku, a onda i koliko ljudi ispod 40 godina i koliko onih iznad 40 godina treba da imamo, ali ne bi određivale koliko studenata ispod 40, a koliko iznad 40 godina treba da imamo, niti bi to određivale za ljude koji nisu studenti. Dakle, sa nevezanim kvotama, u ovom primeru, to da li je osoba student ili nije i to da li ima više ili manje od 40 godina posmatraju se posebno. S druge strane, **vezane kvote su kvote gde se varijable kombinuju prilikom formiranja kvota**. Ako bi hteli da napravimo vezane kvote u prethodnom primeru, bilo bi potrebno da definišemo 4 kominovane kvote – koliko ćemo imati studenata mlađih od 40 godina, koliko studenata starijih od 40 godina,

PREPRINT

koliko ljudi koji nisu studenti mlađih od 40 godina i koliko ljudi koji nisu studenti starijih od 40 godina.

Tabela 2.8. Primer vezanih i nevezanih kvota – vrednosti varijable u populaciji iz primera (izmišljeni podaci)

Vrednosti varijable u populaciji (% ukupne populacije)				
		Da li je student?		Ukupno
		Students	Nije student	
Starost	Do 40 godina	14%	36%	50%
	Preko 40 godina	1%	49%	50%
Ukupno		15%	85%	100%

Tabela 2.9. Primer vezanih i nevezanih kvota – primer planova kvotnog uzorkovanja (izmišljeni podaci).

Kvote Veličina uzorka = 100 entiteta (učesnika u istraživanju)	
Primer nevezanih kvota	Primer vezanih kvota
Pravimo uzorak od 100 učesnika (entiteta) koji se sastoji od: - 15 studenata, 85 nestudenata - 50 osoba do 40 godina starosti, 50 osoba iznad 40 godina starosti. (Obratite pažnju kako se vrednosti dve varijable posmatraju nezavisno jedne od drugih, tj. nisu vezane)	Pravimo uzorak od 100 učesnika (entiteta) koji se sastoji od: - 14 studenata do 40 godina starosti - 1 studenta preko 40 godina starosti - 36 nestudenata do 40 godina starosti - 49 nestudenata preko 40 godina starosti (Obratite pažnju kako su vrednosti dve varijable kombinovane tj. vezane)

PREPRINT

Ako uporedimo vezane i nevezane kvote, možemo primetiti da je jedna od prednosti vezanih kvota to što pravi uzorke koji su sličniji svojstvima populacije na varijablama koje su korišćene za izradu kvota, te tako verovatno popravlja i šanse da svojstva uzorka nalikuju svojstvima populacije i u pogledu svih drugih varijabli. S druge strane, vezane kvote nameću dosta veće zahteve ljudima koji rade na prikupljanju podataka. Kada uzorkovanje počne, svaki entitet koji je dostupan (odnosno svaki potencijalni učesnik u istraživanju u slučaju društvenih nauka) spada u neku od kvota. Ali kako uzorkovanje napreduje, neke kvote će se popuniti brzo, dok će neke ostati nepopunjene. Tako kada dođemo pred kraj formiranja uzorka tj. postupka uzorkovanja može lako nastati situacija gde su nam potrebni još samo ispitanici sa kombinacijama karakteristika koje je teško naći. Ova situacija postaje utoliko verovatnija i utoliko teža što je broj varijabli koje se kombinuju za stvaranje kvota veći i što je veći ukupan broj kvota na koje se oslanja naše uzorkovanje. To u praksi, u ekstremnim slučajevima, može stvoriti motivaciju da se ili falsifikuju podaci iz nedostajućih kategorija ili da se uzorkovanje završi bez njih, pri čemu je ova prva situacija opasnija jer predstavlja naučnu prevaru, a takođe se i često dešava bez znanja istraživača koji rukovode istraživanjem, koji je takođe često ne mogu uočiti kada se dešava, niti rekonstruisati iz podataka da se to desilo (već mogu samo sumnjati bez čvrstih dokaza). Iz ovog razloga, **kada se planira kvotno uzorkovanje, treba voditi računa o tome da se napravi balans između želje da se obezbedi što veća reprezentativnost uzorka i potrebe da procedura uzorkovanja bude realno izvodiva**.

Stratifikovano naspram kvotnog uzorkovanja. Valja primetiti da **neki autori smatraju stratifikovano i kvotno uzorkovanje varijantama iste metode uzorkovanja**, pri čemu kao jedinu razliku navode ono što se radi nakon što se uzorak podeli u stratume odnosno nakon što se definišu kvote (ovde se kvote tretiraju kao analogne stratumima). Ako se nakon ove podele izbor entiteta (za popunjavanje kvota odnosno unutar strauma) radi na probabilistički način, npr. kroz slučajno uzorkovanje, onda je u pitanju stratifikovano uzorkovanje (neki bi tu dodali – stratifikovano slučajno uzorkovanje), dok ako se uzorkovanje entiteta radi na neprobabilistički način, najčešće kroz prigodno uzorkovanje, onda je u pitanju kvotno uzorkovanje. Sreću se takođe i mišljenja da je stratifikovano uzorkovanje postupak koji se primenjuje kada imamo podgrupe koje su administrativno, fizički ili na neki drugi način jasno razdvojene u posebne grupe, tako da

PREPRINT

entiteti iz različitih stratuma nisu pomešani tokom uzorkovanja, dok kvotno uzorkovanje uključuje izbor entiteta na osnovu njihovih svojstava, ali koji su u populaciji pomešani. Na primer, mogli bi da biramo učenike različitih razreda (stratuma) koje bi ispitivali tokom trajnja nastave u školi, a onda bi uzorkovanje iz stratuma mogli da radimo prosto tako što bi odabrali koju učionicu ćemo da posetimo, jer su svi učenici koje zateknemo u istoj učionici učenici istog razreda (stratuma). Nasuprot tome, kod kvotnog uzorka, potencijalni učesnici u istraživanju bi bili prvo ispitani da ustanovimo u koju kvotu spadaju, jer su ljudi koji spadaju u različite kvote na terenu pomešani. Može se primetiti da oba ova načina razlikovanja stratifikovanih i kvotnih uzoraka impliciraju da populacija za stratifikovano uzorkovanje mora da bude ograničena i organizovana na neki način, jer bez toga ne bi bilo spiska članova populacije na osnovu kog bi radili slučajno uzorkovanje, a ne bi bilo moguće ni striktno odeljivanje članova različitih stratuma. Ali bilo kako bilo, za potrebe ove knjige, konstatovaćemo da postoji konceptualno preklapanje između postupaka stratifikovanog i kvotnog uzorkovanja i predstaviti oba, istovremeno priznajući da bi moglo biti podjednako validno da se ova dva metoda predstave i kao različite varijante istog postupka uzorkovanja.

Namerno uzorkovanje se radi u situacijama kada se **tačno određeni entiteti namerno biraju za uključivanje u uzorak**. Sastav uzorka i način na koji se formira se utvrđuju unapred na način koji ne mora nužno da prati distribuciju relevantnih kategorija entiteta u populaciji. **Ideja u osnovi ovakvog uzorkovanja je to da se obezbedi da se istraživanje sprovede na tačno određenim entitetima koji su iz nekog razloga interesantni za proučavanje**. Takvi entiteti mogu biti, na primer, slučajevi koji su posebno informativni (na primer, u studiji slučaja), ljudi koji imaju neka posebna svojstva koja su bitna za temu studije, pripadnici određene grupe za koju se u ranijim studijama pokazalo da se na osnovu njih mogu dobro predvideti parametri populacije isl. Zgodan primer ovakvog uzorkovanja je studija iz 2016. godine u kojoj je grupa istraživača namerno odabrala za proučavanje dvoje ljudi – oca i ćerku koji su imali sposobnost da brzo i lako pričaju unatrag. Ovi istraživači su proučavali različite biološke i psihološke faktore kod ovo dvoje ljudi kako bi probali da ustanove šta im to daje ovu jedinstvenu sposobnost (Prekovic et al., 2016). Ove dve osobe su izabrane za ovo istraživanje upravo zato što je bilo poznato da poseduju tu jedinstvenu sposobnost koju su istraživači želeli da proučavaju. Cilj istraživanja ne bi zadovoljilo biranje nekih drugih ljudi koji nemaju ovu sposobnost. Ipak, iako su istraživači bili zainteresovani da identifikuju faktore koji su zajednički za sve ljude koji imaju ovu sposobnost (populacija), oni

PREPRINT

su namerno izabrali ovo dvoje ljudi, jer su to jedini ljudi sa takvim sposobnostima za koje su znali. Druge ne bi znali gde da nađu.

Klastersko uzorkovanje se sprovodi tako što se populacije podeli u grupe koje se zovu klasteri i onda se biraju klasteri (obično slučajno) koji će biti uključeni u uzorak. Ovo može da se uradi u jednom stadijumu – da se cela populacija podeli na klastere, a da se onda neki klasteri, sa svim svojim članovima uključe u uzorak ili u više stadijuma tako što se populacija podeli u klastere, onda se svaki od tih klastera podeli u podklastere, koji se onda takođe podele u podklastere, sve dok se ne dođe do nivoa podklastera koji su dovoljno mali da mogu da se ispitaju celi za potrebe sprovođenja studije. U klasterskom uzorkovanju koje uključuje više stadijuma, nakon što se populacija podeli na dovoljno male klastere, pravi se izbor klastera koji će biti uključeni u istraživanje (obično slučajnim putem), a onda se svi članovi odabranih klastera (ili u nekim slučajevima deo članova odabran na određeni način) uključuje u uzorak. Jedan mogući primer ovoga bi bila situacija kada bi želeli da dobijemo klasterski uzorak populacije određene države, pa bi onda tu državu podelili na opštine, a onda svaku opštinu na neke manje oblasti, nakon čega bi odrabrali oblasti čije bi stanovnike uključili u uzorak. Ovakav uzorak bi predstavljao geografski klaster uzorak. Jedan **važan zahtev kod klasterskog uzorkovanja je to da klasteri budu međusobno homogeni, a da članovi klastera budu heterogeni**. Drugim rečima, **ne bi trebalo da bude klastera čija su svojstva suviše različita od svojstava populacije celini tj. svaki klaster bi za sebe morao da bude populacija u malom.** Ovo je suštinski različito od stratuma i kvota (kod stratifikovanog i kvotnog uzorkovanja) gde se grupe koje predstavljaju stratume ili kvote biraju zbog zajedničkih svojstava njihovih članova i koji su po tim svojstvima međusobno različiti. Klastersko uzorkovanje je obično jeftinije i lakše za sprovođenje u odnosu na druge tehnike uzorkovanja, ali i pored toga ono uključuje postupke koji služe da obezbede da se uzorak uzima iz različitih delova populacije smanjujući tako šanse da se dobije uzorak koji je previše različit od populacije. Ovaj postupak uzorkovanja se može koristiti na populacijama koje nisu ograničene ili za koje nema spiska članova populacije, dokle god je moguće takvu populaciju podeliti u klastere koji pokrivaju sve članove populacije.

Snoubol uzorkovanje ili uzorak "snežne grudve" se pravi **tako što se prvo u uzorak uključe učesnici (ljudi) koristeći prigodno ili namerno uzorkovanje ili već neki drugi metod kojim se može doći do prvih učesnika u uzorku, a onda istraživači te prve učesnike pitaju da im**

PREPRINT

preporuče naredne učesnike u istraživanju. Nakon što ispitaju te druge učesnike u istraživanju, onda pitaju njih da ih upute na nove učesnike i tako redom dok se ne nakupi potrebna veličina uzorka. Kao što je i očigledno, **snoubol uzorkovanje je primenljivo samo u situacijama kada su entiteti ljudi**, jer je neophodno da istraživači mogu da razgovaraju sa njima i da ih ovi upute na druge moguće učesnike u uzorku. Snoubol uzorkovanje je **metoda izbora onda kada je cilj istraživanja određena specifična populacija čiji su članovi retki u opštoj populaciji ljudi, ali kada se može razumeno očekivati da će članovi te populacije biti u kontaktu jedni s drugima**. Na primer, ako bi hteli da proučavamo populaciju bajkera (ljudi koji su članovi grupa ili klubova okupljenih oko aktivnosti sa vožnjom motora), ne bi imalo puno smisla ići od kuće do kuće pitajući ljude da li su možda slučajno bajkeri, jer bi, u većini slučajeva, bilo potrebno da obiđemo baš puno domaćinstava dok ne naletimo na nekog bajkera. S druge strane, kako su bajkeri međusobno u kontaktu, kroz okupljanja na kojima učestvuju ili preko klubova ili društvenih mreža, svaki bajker će verovatno poznavati još nekoliko drugih bajkera, koji će pak poznavati bar još nekoliko i tako redom. Na sličan način, snoubol uzorkovanje se može koristiti da se sakupi uzorak ronilaca, peintbolera, ljudi koji se bave određenim sportovima koji nisu popularni, ali takođe i ljudi koji pate od određenih retkih bolesti, a koje su takve da zahtevaju specifične vrste podrške koje dovode obolele u međusobni kontakt, ovako je moguće naći i ljude koji su preživeli određene katastrofalne događaje itd.

Sistematsko uzorkovanje se radi tako što se **napravi spisak članova populacije, a onda se u uzorak uzme svaki n-ti entitet na spisku**. Na primer, zavisno od veličine uzorka koja nam treba, možemo da u uzorak uzmemo svaki 100-i ili svaki 50-i ili svaki 20i entietet i slično. Ovaj **broj kojim se određuje koji entitet na spisku će biti odabran za uzorak** se zove **"korak" sistematskog uzorka**. Na primer, pravljenje sistematskog uzorka sa korakom 20 znači da ćemo u uzorak odabrati svaki 20-i entitet sa spiska članova populacije. Kada se radi sistematsko uzorkovanje, generalno je dobra praksa da se ne krene od prvog entiteta na spisku, nego da se prvo generiše jedan slučajni (odnosno pseudoslučajni) broj koji je manji od veličine koraka, pa da se uzorkovanje onda počne od njega. Na primer, ako bi sistematski uzorak pravili sa korakom 30, prvo bi bilo potrebno da generišemo slučajni broj između 1 i 30 (dakle manji od veličine koraka – širina ovog raspona je 29, jer je 30-1=29). Ako bi tako, na primer, dobili broj 12, to znači da bi se naš uzorak sastojao od 12-og, 42-og, 72-og, 102-og i tako dalje entiteta sa spiska članova populacije. **Opšte pravilo je da biramo entitete čiji je redni broj na spisku populacije jednak**

PREPRINT

slučajnom broju koji smo izvukli na početku kome onda dodajemo sve moguće proizvode koraka i prirodnih brojeva, sve dok ne iscrpimo spisak populacije. U principu, ako redosled entiteta na spisku članova populacije nije povezan sa varijablama koje proučavamo, sistematsko uzorkovanje može proizvesti rezultate koji su prilično slični slučajnom uzorkovanju, a pri tom je lakše za sprovođenje. S druge strane, imajući u vidu da sistematski uzorak zahteva da imamo i spisak članova populacije i način da generišemo slučajne ili pseudoslučajne brojeve, svi slučajevi u kojima možemo da radimo sistematsko uzorkovanje su istovremeno i slučajevi u kojima možemo da radimo i prosto slučajno uzorkovanje.

2.9. Čega treba da budemo posebno svesni u vezi uzorkovanja?

U većini slučajeva, optimalni uzorak je uzorak koji je najviše koliko je moguće reprezentativan za populaciju. Međutim, **postoje i situacije u kojima reprezentativni uzorak nije optimalan.** Nekada je cilj istraživača da detaljno prouče različite delove populacije, a neki od tih delova mogu da budu previše mali da bi bili predstavljeni u uzorku sa dovoljnim brojem entiteta za proučavanje ako se uzorak u celini napravi da bude reprezentativan za populaciju. U takvim situacijama, istraživači mogu odlučiti da naprave uzorak u kom će ti delovi populacije biti **prezastupljeni tj. više zastupljeni** nego što bi to bilo opravdano veličinom udela tog dela populacije u populaciji u celini. Iako je takvo pravljenje uzoraka sasvim validan istraživački postupak, **veoma je važno da istraživači koji rade sa takvim uzorcima ni u jednom trenutku ne zaborave da rade sa uzorkom u kome delovi populacije nisu zastupljeni proporcionalno i da su određene grupe prezastupljene**. **Ovo je posebno važno imati na umu prilikom izvođenja zaključaka o populaciji u celini,** jer tretiranje uzorka u kome su određene kategorije prezastupljene kao da je proporcionalan može dovesti do veoma ozbiljnih grešaka u zaključcima.

Kada pričamo o tome **kako izveštavati o vrsti uzorka koji je korišćen u istraživanju**, **autori uvek treba da svoj uzorak nazovu prema tehnici uzorkovanja koja najviše odgovara postupku koji su primenili za sastavljanje uzorka**. Iako procedure uzorkovanja u praksi mogu biti veoma različite i iako je tipično da se ti postupci, onako kako se primenjuju u praksi, više ili manje razlikuju u detaljima od postupaka uzorkovanja koji su opisani u literaturi, uključujući i ovu

PREPRINT

knjigu, **kada se o postupku izveštava, autori treba da ili postupak opišu detaljno ili da ga imenuju prema tehnici uzorkovanja na koju najviše nalikuje.** Uprkos tome, neretko se mogu sresti situacije u kojima istraživači nazivaju (ili pokušavaju da nazovu pa ih recenzenti spreče pre objavljivanja takvog izveštaja!) svoj uzorak „reprezentativnim“. Kao što je opisano u prethodnim poglavljima, **ne postoji tehnika uzorkovanja koja se zove „reprezentativno uzorkovanje“**, pa prema tome, ako isključimo neke sasvim trivijalne slučajeve (vrlo mala populacija, uzorak jednak populaciji), **ne postoji način kako bi istraživač mogao da zna da li je tehnika uzorkovanja koju je primenio/primenila proizvela reprezentativni uzorak ili ne i koliko reprezentativan.** To znači da je praktično u svim situacijama kada autori svoj uzorak opisuju kao „reprezentativan“, takva tvrdnja suštinski obmanjujuća, jer autori iznose tvrdnje o nečemu što ne mogu da znaju. Iz ovog razloga, naša je **preporuka čitaocima da budu veoma oprezni kada nalete na izveštaj o istraživanju u kome se uzorak opisuje kao „reprezentativan**“, a posebno ako ovakvo imenovanje ne prati detaljno objašnjenje tehnike uzorkovanja koja je primenjena. Dešava se da autori, pogotovo oni koji rade komercijalna istraživanja ili imaju komercijalne interese za ishod istraživanja, prosto proglase da je njihov uzorak reprezentativan u nadi da će tako povećati šanse da neko kupi njihovo istraživanje ili da ga finansira, a taj kupac ili finansijer obično bude strana sa samo najosnovnijim znanjima iz statistike ili čak bez ikakvih znanja iz statistike. Ovakvi autori se onda nadaju da će kupac njihovo istraživanje videti kao pouzdanije ili vrednije ako napišu ili kažu da im je uzorak reprezentativan. Međutim, koliko god procedura uzorkovanja koja je primenjena bila složena ili kvalitetna, imenovanje uzorka „reprezentativnim“, a bez mogućnosti da se uzorak zaista uporedi sa populacijom i tako ta tvrdnja proveri, predstavlja uvek obmanjujuću praksu.

Još jedna greška koja se često uočava kod istraživača sa ograničenim znanjem iz statistike, ali i kod onih koji pokušavaju da svoje istraživanje prikažu kao vrednije nego što jeste je nazivanje uzorka slučajnim u situacijama kada on nije takav. Ono što se najčešće sreće je to da **ljudi koji su zapravo koristili prigodno uzorkovanje svoje postupak uzorkovanja nazivaju slučajnim uzorkovanjem**. Zbog toga što ne znaju na šta se slučajno uzorkovanje zapravo odnosi, ovakvi ljudi će za svoje uzorke reći da su slučajni zato što su ih napravili tako što su „ispitivali slučajne prolaznike na ulici“, „ispitivali studente koje su slučajno našli u lokalnom kafeu“, „ispitivali pacijente koje su slučajno zatekli na klinici na kojoj rade“ isl. Kada čitamo takve primere, moramo da budemo svesni da postoji razlika između toga kako se reč „slučajno“ koristi u statistici i toga

PREPRINT

kako se obično koristi u svakodnevnom govoru. Uzorak koji se sastoji od ljudi koje smo sreli na ulici nije slučajni nego prigodni uzorak. Isti je slučaj i sa ljudima iz naše lokalne kafeterije ili sa pacijentima sa klinike na kojoj istraživač radi. Zato treba uvek imati na umu razliku između slučajnog i prigodnog uzorka, kako bi se, kada naletimo na studiju u kojoj autori pišu kako su istraživanje radili na slučajnom uzorku, setili da pogledamo opis postupka uzorkovanja koji je primenjen i da onda proverimo da li je zaista u pitanju slučajno uzorkovanje ili se ipak radi o situaciji da su autori prigodni uzorak pogrešno nazvali slučajnim

2.10. Nivoi merenja

Iz matematike za osnovnu i srednju školu smo naučili različite matematičke operacije koje se mogu raditi sa brojevima. Međutim, svako od nas je verovatno imao prilike da naleti na situaciju u kojoj se koriste brojevi, ali na takav način da izvođenje nekih matematičkih operacija sa njima nema smisla. Na primer, brojevi se uobičajeno koriste da označe različite stanove u stambenoj zgradi ili različite sobe u hotelu. Kada se brojevi upotrebe na takav način, možemo, na primer, reći da broj 1 i broj 2 označavaju različite sobe u hotelu, što znači da ta dva broja ne označavaju istu sobu. Međutim, ne bi imalo mnogo smisla sabrati brojeve dve sobe, zato što soba broj 1 + soba broj 2 nisu jednaki sobi broj 3. Takvo sabiranje je besmisleno, kao i sama postavka. S druge strane, kada su brojevi s kojima radimo, na primer, stepeni Celzijusove temperaturne skale, složićemo se da je 1 ^{o}C hladnije od 2 ^{o}C, ali ćemo takođe primetiti i da 1 ^{o}C nije duplo hladnije od 2 ^{o}C. Međutim, tačno je reći da je 2 ^{o}C za 1 ^{o}C toplije od 1 ^{o}C. Drugim rečima, ima smisla oduzeti jednu temperaturu izraženu u ^{o}C od druge i tako dobiti njihovu razliku u ^{o}C, iako nema smisla ove vrednosti deliti jedne sa drugima, niti ih množiti. S druge strane, ako su brojevi koje koristimo metri ili grami, te mere možemo lako i deliti i množiti – masa od 2g je duplo veća od mase od 1g. Trupac od 2 metra dužine je duplo duži od trupca od 1 metra dužine. **Ova pravila koja određuju koje matematičke operacije možemo, a koje ne možemo smisleno primeniti na brojevima** zavise od onoga što se u statistici naziva **nivoima merenja.** Verovatno najpoznatija i najšire primenjivana klasifimkacija nivoa merenja je ona koju je predložio Stevens (1946). Stevens je predložio 4 nivoa merenja – nominalni, ordinalni, intervalni i racio nivo merenja. U osnovi, ovi nivoi merenja se razlikuju po tome koji matematički odnosi se mogu smisleno uspostaviti između brojeva na svakom nivou merenja. Oni grade hijerarhiju takvu da svaki naredni nivo merenja dozvoljava sve matematičke odnose koji su bili primenljivi na prethodnom nivou plus još neke nove.

PREPRINT

Na **nominalnom** nivou merenja **jedini odnos koji se može smisleno uspostaviti između brojeva je odnos jednakosti/nijednakosti**. Brojevi su na ovom nivou prosto oznake za pojedinačne objekte ili za kategorije objekata i zato, za svaka dva broja, možemo samo ustanoviti to da li označavaju isti objekat tj. istu klasu objekata ili različit/različitu. Primeri nominalnog nivoa merenja uključuju situacije kada koristimo brojeve da bi označili sobe u hotelu ili različite stanove u stambenoj zgradi, kada koristimo brojeve da označimo igrače u nekom timskom sportu koji igraju na različitim pozicijama, ali i situacije kada koristimo brojeve u statističkoj matrici podataka da označimo vrednosti varijabli poput etničke pripadnosti (gde svaka etnička grupa ima sopstveni broj različit od drugih), pol (svaki pol dobija različit broj), različite profesije u kojima ljudi rade (svaka profesija označena različitim brojem) ili drugih sličnih kategorijalnih varijabli koje imaju kategorije između kojih se ne može ustanoviti neki drugi kvantitativni odnos. Takođe, podatke možemo tretirati kao da su nominalni čak i onda kada objekti ili kategorije objekata nisu označeni brojevima nego skupovima slova ili znakova (na primer, pogledajte varijablu ime u tabeli 2.6.), zato što se odnos jednakosti/nejednakosti može jednako uspostaviti i sa takvim oznakama jer nominalni nivo merenja ne zahteva zaista bilo koje svojstvo koje je specifično za brojeve. Varijable na nominalnom nivou merenja nazivaju se **nominalnim varijablama.**

Na **ordinalnom** nivou merenja **možemo ustanoviti i da li je određena vrednost veća, manja ili jednaka bilo kojoj drugoj vrednosti**, pored toga što što možemo i dalje ustanoviti da li broj označava isti objekat/vrstu objekata ili ne. Na ovom nivou merenja se **znakovi <, > i = mogu smisleno koristiti**. Sve vrste rangova i rezultata rangiranja podataka za proizvod daju podatke na ordinalnom nivou merenja – rangovi/redni brojevi u trci (tu iz rangova možemo ustanoviti ko je za kraće vreme prešao stazu od koga, ali ne i za koliko kraće odnosno duže) ili nekom drugom takmičenju, činovi u vojsci (tu možemo ustanoviti na primer da je kapetan rangiran iznad poručnika, a možemo konstatovati i da dve osobe u činu kapetana imaju isti čin, ali ne možemo npr. valjano ustanoviti da li je razlika između poručnika i kapetana iste veličine kao razlika između majora i potpukovnika) itd. Ordinalni nivo merenja se takođe koristi u nekim od široko primenjivanih skala poput Merkalijeve skale intenziteta seizmičkih aktivnosti (Wood & Neumann, 1931), u različitim skalama za procenu ukusa, a treba pomenuti i tekuću debatu o tome da li su skale samoprocene koje se obilato koriste u društvenim naukama na ordinalnom nivou merenja ili na onom narednom. Varijable čije su vrednosti na ordinalnom nivou merenja nazivamo **ordinalnim varijablama.**

PREPRINT

Na **intervalnom** nivou merenja, možemo i **da poredimo veličine intervala između dve vrednosti,** što znači da možemo da koristimo i sabiranje i oduzimanje (+ i -) na određene načine. **Mere na intervalnom nivou merenja imaju fiksnu jedinicu mere** i, zbog ovog, ne samo da možemo da utvrdimo koja mera je od koje veća, već možemo da ustanovimo i veličinu razlike između te dve vrednosti. Međutim, **kod intervalnih mera je nula skale ili početna vrednost arbitrarno određena, što znači da vrednost 0 ne ukazuje na potpuno odsustvo merene osobine**, već je to jednostavno početni nivo koji je proizvoljno određen od strane tvorca skale. Na primer, moguće je smisleno oduzeti jednu vrednost od druge da bi dobili njihovu razliku, a moguće je i dodati ili oduzeti određeni broj mernih jedinica nekoj meri da bi je povećali ili umanjili. Međutim, na ovom nivou merenja **ne možemo smisleno sabrati vrednosti dva entiteta da bi dobili vrednost entiteta čija je izraženost merene osobine jednaka zbiru stepena izraženosti te osobine na ta dva objekta.** Primer intervalnog nivoa merenja mogu biti Celzijusova ili Farenhajtova temperaturna skala. Obe ove skale imaju fiksne jedinice mere koje možemo iskoristiti da bi smisleno izračunali koliko je, u stepenima date skale, jedan objekat topliji ili hladniji od drugog. S druge strane, njihove nule su proizvoljne – za Celzijusovu skalu je odabrano da 0°C predstavlja temperaturu na kojoj se voda pretvara u led pod uobičajenim atmosferskim pritiskom, dok je 0°F temperatura na kojoj se koncentrovani rastvor napravljen od 50% soli i 50% leda topi. Da su tvorci ovih skala odabrali neke druge temperature umesto ovih da predstavljaju 0 stepeni skale, ove skale bi jednako dobro funkcionisale kao što funkcionišu sada. Takođe, ako je temperatura u nekoj prostoriji, na primer, 20°C, tu temperaturu je moguće povećati ili smanjiti za npr. 2°C i tako dobiti novu temperaturu prostorije od 22°C ili 18 ^{o}C i to povećanje/smanjenje je ispravno tako izraziti. S druge strane, ne bi bilo ispravno reći da je objekat čija je temparature npr. 4°C duplo topliji od objekta čija je temperatura 2°C (jer to nije tačno!), a isto pravilo važi i za Farenhajtovu skalu**. Intervalni nivo merenja dozvoljava da se odredi jednakost intervala ili razlika, ali ne i odnosa**. Na ovom nivou merenja nije moguće smisleno množiti li deliti mere (ali jeste razlike između mera, koje su na narednom nivou merenja gde ove operacije funkcionišu!). U psihologiji i drugim društvenim naukama, **uobičajena je praksa da se rezultati psiholoških testova** (posebno kada su izraženi u vidu standardnih skorova), **kao i rezultati sa skala procene i samoprocene smatraju za mere na intervalnom nivou merenja**, iako u naučnoj zajednici postoji debata o tome da li ovakve mere u potpunosti ispunjavaju zahteve intervalnog nivoa merenja ili ih je ipak pravilnije tretirati kao ordinalne. Glavna tačka ove diskusije je to da li ovakve

PREPRINT

mere ispunjavaju zahtev da merne jedinice[4] budu fiksne veličine ili ne. Varijable koje su na intervalnom nivou merenja nazivaju se **intervalne varijable.**

Na **racio** nivou merenja, **možemo smisleno porediti odnose između mera, što znači da je moguće smisleno množiti ili deliti mere** (koristiti znakove množenja i deljenja u radnjama sa ovim merama) da bi ustanovili jednakost odnosa, a i sabiranje i oduzimanje vrednosti je moguće bez ograničenja koja postoje kod intervalnih mera. Pored svojstava intervalnih mera, **mere na racio nivou merenja imaju i tzv. realnu nulu tj. realne početne vrednosti, što znači da vrednost 0 ili početna vrednost racio skale označava potpuno odsustvo merene osobine**. Zbog ovog svojstva racio skala, kod njih možemo smisleno računati i to koliko je puta jedna vrednost veća ili manja od druge. Na primer, sasvim je ispravno konstatovati da je objekat koji je dugačak 2 metra duplo duži od objekta koji je dugačak 1 metar. Objekat čija je masa 4 kilograma je duplo teži od objekta čija je masa 2 kilograma. Takođe, ako u kesu u kojoj se nalazi 3 kg krušaka, izručimo sadržaj kese u kojoj je bilo 2 kg krušaka, imaćemo u kesi ukupno 5kg krušaka. Ove mere su na racio nivou merenja zato što imaju realne nule – 0 grama označava potpuno odsustvo mase, kao što i 0 metara označava da nema udaljenosti od između 2 tačke na koje se ta dužina odnosi. Ako se nalazimo 0 metara od neke tačke, to znači da se nalazimo baš na toj tački i da joj nikako ne možemo biti bliži. Primeri racio nivoa merenja uključuju izražavanje mase u gramima, mere dužine u metrima (ili sličnim ekvivalentnim merama), izražavanje temperature u stepenima Kelvina (zato što 0°K, temperatura koja je poznata i kao "apsolutna nula" zaista predstavlja potpuno odsustvo toplote!), a na racio nivou merenja su i rezultati dobijeni prebrojavanjem objekata – na primer broja vrata u nečijoj kući, broja grla stoke koje neko poseduje, broja ljudi u autobusu isl.

Treba napomenuti i to da **je razlika između dve intervalne mere takođe na racio nivou merenja**. Na primer, ako oduzmemo dve vrednosti u ^{o}C jednu od druge, njihova razlika biće na racio nivou merenja, iako su same mere intervalne. To možda najlakše možemo videti po tome što kada oduzmemo dve identične intervalne mere jednu od druge, npr. kada oduzmemo 16 ^{o}C od 16

[4] Veličina merne jedinice odnosi se na veličinu jednog podeoka na skali, na primer, veličinu razlike u merenom svojstvu između objekata kojima odgovaraju susedni brojevi na skali (npr. 2 i 3, 45 i 46 ili bilo koji par susednih brojeva). Da bi mera bila intervalna, razlike u izraženosti merenog svojstva moraju biti jednake između svih objekata koji se razlikuju za po jednu jedinicu, odnosno za po 1 na skali. To znači, na primer, da razlika između vrednosti 2 i 3, mora biti jednaka kao razlika između 45 i 46, a onda obe ove moraju biti jednake veličini razlike između bilo koja dva druga susedna broja tj. broja koji se razlikuju za 1.

PREPRINT

^{o}C, kao rezultat ćemo dobiti razliku od 0 ^{o}C, za koju ćemo se lako složiti da označava to da se ove dve intervalne mere ne razlikuju uopšte, odnosno da predstavlja odsustvo bilo kakve razlike između mera. A potpuno odsustvo merenog svojstva je upravo ono što nazivamo realnom nulom, što je svojstvo koje izdvaja racio skale od intervalnih (racio skale, kao što je gore navedeno imaju realnu nulu, dok je nula kod intervalnih arbitrarna). Takođe ćemo se lako složiti da je razlika između dve intervalne mere, npr. opet u ^{o}C, od 1 ^{o}C, duplo manja od razlike od 2 ^{o}C, što je opet svojstvo racio skala. Na primer, složićemo se lako da je razlika između 18 ^{o}C i 20 ^{o}C duplo veća od razlike između 18 ^{o}C i 19 ^{o}C. Dakle, razlike između intervalnih mera su na racio nivou merenja!

Postoje autori koji predlažu i druge dodatne nivoe merenja, pored ova četiri predstavljena ovde, kao što je npr. **apsolutni** nivo merenja (na kom bi bili rezultati prebrojavanja, to je kao racio nivo, samo što vrednosti mogu biti samo prirodni brojevi) ili loglinearni nivo merenja (kao racio nivo, samo što uključuje mere koje su u jedinicama koje predstavljaju odnos/količnik dve druge jedinice, kao što su npr. m/s ili N/m^2). Međutim, kako na ovakvim merama nisu stvarno moguće neke dodatne matematičke operacije (u odnosu na racio nivo), u ovoj knjizi ćemo se zadržati na Stivensovoj sistematizaciji nivoa merenja i razlikovati samo navedena četiri nivoa. Varijable čije su vrednosti na racio nivou merenja zovu se racio varijable.

Jedan važan izuzetak od pravila koja određuju nivoe merenja su **binarne ili dihotomne varijable**. **Binarne tj. dihotomne varijable su varijable koje imaju samo dve moguće vrednosti**. Zato što imaju samo dve vrednosti **one se mogu posmatrati kao da su na bilo kom nivou merenja**. Naime, očigledno je da te dve moguće vrednosti nisu jedna drugoj jednake i da tako ispunjavaju uslove za nominalni nivo merenja. Ako na primer te dve vrednosti označimo sa A i B (iako binarne varijable mogu imati bilo kakve dve vrednosti – 0 i 1, tačno i netačno, C i D itd.) i složimo se da su to dve različite vrednosti, postoje samo dva načina na koje ih možemo rangirati – AB i BA. Ova dva načina rangiranja predstavljaju zapravo isti redosled vrednosti samo u suprotnim smerovima, te tako binarne varijable ispunjavaju i zahtev ordinalnog nivoa da se vrednosti mogu poređati od najmanje do najveće ili obrnuto. Intervalni nivo merenja zahteva fiksnu jedinicu mere, što znači da na ovom nivou mora da bude moguće da se raspon varijable podeli na intervale jednakih veličina. Kod binarnih varijabli, postoji samo jedan takav interval i on je uvek jednak samom sebi, te tako binarne varijable mogu da se tretiraju i kao da ispunjavaju zahteve intervalnog nivoa merenja. Konačno, racio nivo zahteva da postoji realna ili apsolutna nula tj. da nema vrednosti

PREPRINT

manjih od nule. Imajući u vidu da binarna skala ima samo dve vrednosti, možemo arbitrarno odlučiti da je jedna od njih niža (osim ako jedna nije i prirodno niža, kao što je često slučaj kada je jedna od vrednosti nula, a druga je neki pozitivni broj) od druge i onda će ta vrednost istovremeno biti i najniža moguća vrednost (jer nema nižih od nje), te tako binarne varijable ispunjavaju uslove i racio nivoa merenja. Iako je, naravno, jasno da binarne varijable nisu zaista prave ordinalne, intervalne ili racio varijable, **navedena svojstva binarnih varijabli zgodno omogućavaju da u određenim, odgovarajućim situacijama koristimo statističke i matematičke postupke namenjene podacima na ordinalnom, intervalnom ili racio nivou i na binarnim varijablama.** Naravno, **rezultate tih postupaka na binarnim varijablama treba uvek interpretirati uzimajući u obzir da su računanja rađena na binarnim varijablama, a ne na pravim intervalnim ili racio varijablama**. Neki statistički postupci, uključujući i neke predstavljene u ovoj knjizi, čak imaju i drugačija imena kada se primenjuju na binarnim varijablama kako bi odrazila ove potrebne razlike u interpretacijama rezultata.

****

Prepoznavanje nivoa merenja na kom su podaci sa kojima radimo je od kitične važnosti za statističke postupke, zato što različiti nivoi merenja dozvoljavaju (ili ne dozvoljavaju) korišćenje različitih statističkih postupaka. U principu, što je viši nivo merenja, veći je broj statističkih postupaka koji se mogu primeniti na tim podacima. Postoje statistički postupci koji se mogu primeniti na svim nivoima merenja, ali postoje i statistički postupci koji su ograničeni samo na podatke koji su na određenom nivou merenja ili na nekom nivou merenja koji je viši od tog traženog. **Za svaki statistički postupak postoji minimalni nivo merenja na kom podaci moraju da budu da bi se taj postupak na njima mogao smisleno primeniti**. Opšte pravilo je da je **minimalni nivo merenja koji je potreban za neki statistički postupak onaj nivo merenja koji dozvoljava sve matematičke operacije koje su potrebne za izračunavanje statističke mere koja je rezultat tog postupka.** Zbog ovoga, određivanje nivoa merenja podataka je jedna od prvih aktivnosti koje treba uraditi kada se radi sa skupom podataka i za svaki statistički postupak koji će biti predstavljen u ovoj knjizi biće pomenut i minimalni nivo merenja na kom podaci moraju da budu da bi se taj statistički postupak mogao smisleno[5] primeniti.

[5] Kada govorimo o uslovima za sprovođenje statističkih postupaka često koristimo termin „smisleno“ zbog toga što su, nezavisno od nivoa merenja, podaci uvek izraženi kao brojevi, a „brojevi trpe sve“, odnosno mi statističku

PREPRINT

2.11. Kontinualne i diskretne varijable

Kada imamo u vidu to da varijable mogu imati različite vrednosti i nakon što smo prodiskutovali pitanje nivoa merenja, pitanje se može postaviti o tome **koliko različitih vrednosti varijabla može da ima**. S tim je povezano i pitanje toga koliko različitih vrednosti varijabla može da ima u odnosu na broj entiteta u uzorku koji opisujemo na toj varijabli. Ako je broj različitih vrednosti varijable mali u odnosu na veličinu uzorka, onda će više entiteta imati istu vrednost varijable, te onda može biti smisleno da se uzorak opisuje brojanjem koliko entiteta ima svaki od mogućih vrednosti. Ako je pak broj različitih vrednosti veliki u odnosu na veličinu uzorka, onda će biti manje entiteta koji imaju istu vrednost na varijabli ili se čak može desiti da svaki entitet ima različitu vrednost i da, u tom slučaju, brojanje entiteta koji imaju svaku od mogućih vrednosti ne bi bio naročito efikasan način opisivanje uzorka, jer bi za većinu ili čak za sve moguće vrednosti prosto ustanovili da datu vrednost ima samo jedan entitet. Drugo pitanje koje je sa ovim povezano je to **da li postoji iscrpna i konačna lista svih mogućih vrednosti varijable ili je ta lista beskonačna**, kao i pitanje toga koje se vrste brojeva mogu koristiti za izražavanje ovih vrednosti. Ovo pitanje se sreće najčešće kao pitanje toga **šta su moguće vrednosti varijable**. Kada su podaci na **nominalnom nivou merenja**, odgovor na ovo pitanje je jasan – s obzirom da su brojeve samo oznake za objekte ili kategorije objekata, različitih vrednosti – **različitih brojeva koji izražavaju vrednosti varijable ima tačno onoliko koliko je onaj ko je operacionalizovao varijablu odredio**. Nasuprot tome, kada imamo posla sa **intervalnim ili racio varijablama** ovo pitanje postaje komplikovanije. **Teorijski, između svake dve tačke intervalne ili racion skale moguće je zamisliti beskonačan broj različitih vrednosti** (koje se izražavaju u vidu decimala ili razlomaka) i ovaj broj je ograničen samo preciznošću mernog instrument koji koristimo. U praksi smo, međutim, **često slobodni da smanjimo broj mogućih različitih vrednosti varijabli da bi uprostili merenje**, kao i obradu podataka i to tako što **vrednosti zaokružujemo** ili tako što **istu**

---

proceduru možemo matematički sprovesti i kada podaci nisu na potrebnom nivou merenja, samo što rezultati koje dobijemo neće biti smisleni, iako će ih npr. statistički softver svejedno izračunati ako mu tako naložimo. To istovremeno znači i **da rezultati statističkih proračuna mogu da ne budu smisleni ako nivo merenja podataka nije uzet u obzir, iako će korisnik koji takve proračune zahteva od statističkog softvera najčešće dobiti nekakve brojeve kao rezultate**.

PREPRINT

vrednost dodeljujemo određenim intervalima vrednosti varijable umesto pojedinačnim vrednostima.

Da bi se odredili prema ovim pitanjima, potrebno je uvesti koncepte diskretnih i kontinualnih varijabli i njima odgovarajućih diskretnih i kontinualnih merenja. **Diskretne varijable** su **varijable koje mogu imati samo određene, obično unapred propisane, vrednosti**. Primer takvih varijabli su one koje mogu imati samo određene vrednosti koje je osoba koja vrši merenje unapred odredila ili čije vrednosti mogu biti samo prirodni brojevi ili samo celi brojevi ili samo desetine itd. Nasuprot tome, vrednosti **kontinualnih varijabli** mogu biti **svi realni brojevi**. **Diskretne varijable mogu biti na bilo kom nivou merenja, dok kontinualne varijable svoj pun smisao dobijaju tek na intervalnom ili racio nivou merenja.**

Imajući ovo u vidu, postavlja se pitanje toga da li je u praksi zaista moguće imati varijablu čija vrednost može da bude bilo koji prirodni broj tj. da li u praksi zaista možemo imati kontinualna merenja koja daju prave kontinualne varijable. Dosta je očigledno da je odgovor na to pitanje – ne. **Kontinualne varijable predstavljaju teorijsku koncepciju, ali smo u praksi uvek ograničeni preciznošću instrumenta koji koristimo da bismo dobili mere**. Kako ne postoje instrumenti čija je preciznost apsolutna – neograničena, tako nije ni moguće da svi realni brojevi budu moguće vrednosti bilo koje varijable, već smo ograničeni na minimalnu razliku u vrednostima varijable koju instrument koji koristimo može da detektuje. Iz ovog razloga, **u praksi ne postoje kontinualne varijable, već samo diskretne varijable sa širim ili užim kategorijam**a tj. sa manjim ili većim brojem kategorija tj. mogućih diskretnih vrednosti. Međutim, u situacijama **kada broj mogućih kategorija postane dovoljno veliki** i kada pri tom imamo posla sa **intervalnom ili racio varijablom, možemo sasvim valjano tretirati takve mere kao dovoljno dobre aproksimacije kontinualnih merenjenja**, te onda takva merenja tretirati kao kontinualna.

2.12. Hajde da primenimo to što smo naučili do sada!

Hajde da probamo sada da primenimo stvari koje smo predstavili u ovom poglavlju kroz nekoliko vežbi. Molimo vas da pogledate opšte uputstvo za ovakve vežbe koje možete naći na početku knjige. Naša preporuka je da prvo pročitate svaki isečak i tvrdnje date u njemu i da onda date svoj

PREPRINT

odgovor. Odgovor možete upisati u kolonu za odgovore, a posle toga pročitajte odgovore i uporedite svoje odgovore sa njima.

Vežba A. Varijable, nivoi merenja, uzorkovanje.

Jovan radi istraživanje u kom planira da istraži da li se osobine ličnosti ljudi koji rade kao fudbalske sudije razlikuju od osobina ličnosti opšte populacije. To planira da uradi tako što će porediti rezultate grupe fudbalskih sudija na HEXACO PI-R testu ličnosti sa normama ovog testa za opštu populaciju. Njegov plan je da skupi uzorak fudbalskih sudija tako što će uzeti spisak svih fudbalskih sudija u zemlji i onda iskoristiti generator slučajnih brojeva da odabere one koje će uključiti u uzorak. Njegov plan je da sastavi uzorak od 600 fudbalskih sudija na ovaj način. U postupku uzorkovanja, sudija koji jednom bude uključen u uzorak, neće više biti uključen u izvlačenja daljih učesnika (tj. jedan sudija može da bude u uzorku samo jednom).

Samo testiranje ovih sudija će se raditi korišćenjem HEXACO PI-R inventara ličnosti koji daje rezultate na 6 različitih dimenzija ličnosti. Za ove testne skorove se može smatrati da su na skali koja ima fiksnu jedinicu mere, ali arbitrarnu 0 tj. ispitanici ne mogu da imaju vrednost 0, a najmanja moguća vrednost je jednaka broju stavki koje ima određena skala. Jovan će takođe zabeležiti podatak o zanimanju sudije van sporta (većina sudija ima i neku drugo zanimanje kojim se bavi, pored sudijskog), njihov pol (sa opcijama muško, žensko i drugo), kao i broj fudbalskih utakmica u kojima je sudija sudio/la tokom svoje karijere.

Jelena radi istražianje u isto vreme kada i Jovan. Jelena planira da svoje istraživanje sprovede sama i da ispita samo fudbalske studije koji žive u njenoj zemlji, a koji su sudili u barem jednoj međunarodnoj utakmici u zadnjih godinu dana. Ona planira da uzorak skupi tako što će naći nekoliko sudija u fudbalskom savezu i onda ih zamoliti da joj preporuče svoje kolege – druge sudije koji ispunjavaju Jelenin uslov (da su sudili u međunarodnoj utakmici u zadnjih godinu dana). Ona će onda testirati ove preporučene sudije, a onda zamoliti i njih da preporuče svoje kolege koji ispunjavaju uslove za učešće u istraživanju i tako će nastaviti sve dok ne sakupi potrebnu veličinu uzorka. Sve sudije koje je Jelena zamolila da učestvuju u njenom istraživanju su pristale.

Molimo pretpostavite da su u svim opisanim situacijama distribucije jednake teorijskim distribucijama koje se očekuju u situacijama datog tipa.

A	**Statement:**	**Answer**
A1.	Ako sve sudije koje je Jovan pozvao da učestvuju u istraživanju pristanu da učestvuju, Jovanov uzorak će biti prigodni uzorak.	
A2.	Jelena planira da svoje istraživanje sprovede na uzorku „snežne grudve“.	
A3.	Svih 600 fudbalskih sudija koje je Jovan odabrao je pristalo da da sve podatke koje je Jovan tražio.	
A4.	Jovanov uzorak će biti reprezentativan za populaciju.	
A5.	HEXACO PI-R skorovi su na racio nivou merenja.	

PREPRINT

A6.	Ako se ispostavi da su sve sudije u Jeleninom uzorku muškarci, varijabla pol će u njenom uzorku biti konstanta.	
A7.	Jovan radi uzorkovanje bez vraćanja.	
A8.	Broj fudbalskih utakmica u kojima je sudija sudio tokom svoje karijere je diskretna varijabla.	
A9.	Jelenin postupak uzorkovanja garantovano daje reprezentativan uzorak.	
A10.	U Jovanovom istraživanju, pol je binarna varijabla.	

Vežba B. Varijable, nivoi merenja, uzorkovanje.

Katarina sprovodi istraživanje čiji je cilj da otkrije koliko su stanovnici njene opštine zadovoljni novom aplikacijom za izdavanje zvaničnih dokumenata koju su opštinske vlasti počele da koriste. Ova aplikacija treba da služi svim stanovnicima opštine i, prema tome, su svi stanovnici opštine ciljna grupa njenog istraživanja. U pitanju je gradska opština sa otprilike 20 000 stanovnika. Katarina je istraživanje sprovela tako što je otišla u park koji je blizu njene kuće i tamo zaustavljala prolaznike sa molbom da popune njen upitnik. Ovaj postupak je nastavila sve dok nije skupila odgovore 200 ljudi, što je broj ljudi koji je planirala da ima u uzorku.

Njena koleginica Isidora je sprovela isto istraživanje, ali je uzorak napravila tako što je uzela zvaničnu bazu sa spiskom svih stanovnika opštine, a onda zamolila svakog 100og stanovnika da učestvuje u istraživanju.

I Katarina i Isidora su tražile od učesnika u istraživanju da rangiraju različite usluge koje opština pruža i pri tom beležile posebno rang koji je dobila aplikacija (koja ih zapravo interesuje) u odnosu na ostale usluge koje pruža opština. One su beležile i starost učesnika u istraživanju (ispitanika), u godinama. Na kraju su podelile opštinu u nekoliko oblasti, svaku od tih oblasti obeležile različitim brojem i onda beležile oblast opštine u kojoj ispitanik živi.

B	**Statement:**	**Answer**
B1.	Katarinin uzorak je reprezentativan za populaciju.	
B2.	Katarina je napravila slučajni uzorak.	
B3.	Isidora je svoj uzorak napravila postupkom sistematskog uzorkovanja.	
B4.	Izabelin uzorak je reprezentativniji od Katarininog.	
B5.	Starost ispitanika je varijabla na racio nivou merenja.	
B6.	Varijabla preko koje se procena kvaliteta aplikacije koja je predmet istraživanja poredi sa procenom kvaliteta ostalih usluga koje opština pruža je na intervalnom nivou merenja.	

PREPRINT

B7.	Varijabla koja pokazuje oblast unutar opštine u kojoj živi učesnik u istraživanju je na nominalnom nivou merenja.	
B8.	Isidora će imati otprilike 200 učesnika u istraživanju, ako svi koje odabere odluče da učestvuju.	
B9.	U Isidorinom uzorku ima više muškaraca nego žena.	
B10.	Starost ispitanika je primordijalna varijabla.	

Hajde sada da vidimo odgovore:

A1-netačno. Priča kaže da Jovan planira da skupi uzorak korišćenjem generatora slučajnih brojeva radi biranja članova uzorka sa spiska članova populacije. To je opis prostog slučajnog uzorka, a ne prigodnog.

A2 – tačno. Jelena će pitati svoje inicijalne učesnike u istraživanju da preporuče dalje učesnike i to je način kako se pravi uzorak „snežne grudve“, tj. snoubol uzorak.

A3 – nepoznato. Tvrdnja se odnosi na dešavanja tokom prikupljanja podataka, a ovo nije objašnjeno u priči.

A4 – nepoznato. Iako je postupak prostog slučajnog uzorkovanja koju Jovan sprovodi dobra procedura uzorkovanja, nikada ne možemo da znamo da li je naš uzorak zaista ispao reprezentativan ako nismo u mogućnosti da ga uporedimo sa populacijom. Međutim, kada bi imali podatke o populaciji, uzorak nam ne bi ni trebao.

A5 – netačno. U priči se jasno kaže da ove skale nemaju realnu 0, pa stoga ne mogu biti na racio nivou merenja.

A6 – tačno. Da, ako svi entiteti u uzorku imaju istu vrednost određenog svojstva, to svojstvo je konstanta, a ne varijabla. Da bi svojstvo bilo varijabla u uzorku, entiteti treba da imaju različite vrednosti na njemu.

A7 – tačno. Priča kaže da kada je sudija jednom uključen u uzorak, onda se on ne razmatra ponovo za uključivanje u uzorak. To je definicija uzorkovanja bez vraćanja.

A8 – tačno. Iako će različite sudije imati različit broj utakmica u kojima su sudili, ove vrednosti mogu da budu samo prirodni brojevi, što broj utakmica čini diskretnom varijablom.

PREPRINT

A9 – netačno. Ne postoji procedura uzorkovanja koja sigurno daje reprezentativni uzorak.

A10 – netačno. U priči se kaže da je pol operacionalizovan kao varijabla sa 3 kategorije. Binarna varijabla je varijabla sa 2 kategorije.

B1 – nepoznato. Ništa u priči ne kaže da li je reprezentativan ili ne, a nijedna tehnika uzorkovanja ne garantuje da će dobijeni uzorak biti reprezentativan.

B2 – netačno. Katarinin uzorak je prigodan. Iako bi verovatno, u svakodnevnom govoru, rekli da je ona ispitivala „slučajne prolaznike“, takav pristup uzorkovanju se zove prigodno uzorkovanje i nema nikakve veze sa slučajnim uzorkovanjem.

B3 – tačno. Primenila je sistematsko uzorkovanje sa spiska stanovnika, sa korakom 100.

B4 – nepoznato. U priči nema nikakvih elemenata na osnovu kojih bi mogli da izvedemo zaključak o tome koji uzorak je reprezentativniji. Opet, primena bolje tehnike uzorkovanja nije garancija da će dobijeni uzorak biti reprezentativniji od nekog drugog uzorka.

B5 – tačno. Da, starost je racio varijabla. Osoba koja ima 20 godina je duplo starija od osobe koja ima 10 godina. To je racio nivo.

B6 – netačno. Ta varijabla se dobija rangiranjem. Rangovi su na ordinalnom nivou merenja.

B7 – tačno. Oblasti su označene brojevima i jedina stvar koju možemo smisleno da kažemo na osnovu ove varijable je da li osobe koje poredimo žive u istoj oblasti opštine ili u različitim oblastima. Brojevi označavaju kategorije. To je nominalni nivo merenja.

B8 – tačno. Ako opština ima 20 000 stanovnika, a Isidora koristi sistematsko uzorkovanje se korakom 100, to je 20 000 / 100 = 200. Tu možemo još oduzmemo ili dodamo nekog stanovnika za slučaj da nije počela od stanovnika broj 1, ili da ima zapravo nešto manje ili više od 20 000 stanovnika opštine, s obzirom da priča kaže da ih je otprilike 20 000 i otud ono otprilike.

B9 – nepoznato. Nema nikakvih podataka o tome kakav je odnos muškaraca i žena u uzorcima u tesktu.

B10 – besmisleno. Ne postoje nikakve primordijalne varijable.

PREPRINT

Poglavlje 3. Deskriptivna statistika

Apstrakt. Ovo poglavlje je posvećeno predstavljaju statističkih postupaka koji se koriste za opisivanje uzorka. U prvom delu je predstavljen pojam distribucije, kao i osnovne statističke mere za opisivanje distribucije – frekvencija, proporcija i procenat. Nakon toga sledi deo u kome se predstavljaju percentili, percentilni rangovi, kao i kvantili. Podpoglavlje o merama centralne tendencije je sledeće i u njemu su predstavljeni aritmetička sredina, medijana i mod. Pored njih se ukratko govori i o harmonijskoj i geometrijskoj sredini. Sledi deo o merama varijabilnosti – standardnoj devijaciji i varijansi, medijani, apsolutnoj medijani odstupanja, kvartilnoj devijaciji i rasponu. Poslednji deo ovog poglavlja sadrži kratak uvod u različite načine grafičkog predstavljanja distribucije.

Ključne reči: distribucija, mere centralne tendencije, mere varijabilnosti, grafičko predstavljanje distribucije

Statistička obrada podataka prikupljenih na uzorku tipično počinje opisivanjem uzorka, a nakon toga, izvođenjem zaključaka o populaciji na osnovu ovih opisa. **Statistički postupci i pokazatelji koji se koriste za opisivanje uzorka čine deo statistike koji se zove deskriptivna statistika**. Najosnovniji postupci i statističke mere deskriptivne statistike će biti predstavljeni u ovom poglavlju.

3.1. Distribucija

Spisak ili opis vrednosti entiteta iz uzorka na određenoj varijabli naziva se distribucijom. Distribucija varijable sastoji se od spiska svih vrednosti svih entiteta u uzorku, a može se sastojati i od slikovnog prikaza svih mogućih vrednosti sa informacijama o tome koliko se često entiteti koji imaju svaku pojedinačnu vrednost susreću u uzorku. Kada su varijable diskretne ili, u praksi, kada je broj mogućih različitih vrednosti varijable mali, distribucija može da se opiše navođenjem koliko entiteta ima svaku od vrednosti. Takve distribucije zovu se **diskretne distribucije.** Ako je

PREPRINT

broj mogućih različitih vrednosti veoma veliki, preveliki da bi se predstavio na napred pomenuti način, onda se vrednosti mogu grupisati u šire kategorije, a distribucija se onda može predstaviti navođenjem koliko entiteta spada u svaku od kategorija. Ako želimo da opišemo distribuciju vrednosti neke kontinualne varijable na uzorku, onda imamo posla sa **kontinualnom distribucijom.** Kontinualna distribucija može da bude predstavljena preko **funkcije gustine verovatnoće,** tj. funkcije čija vrednost u bilo kojoj tački može da se interpretira kao **relativna verovatnoća da slučajno izabrani entitet iz uzorka ima vrednost blizu vrednosti kojoj odgovara ta tačka**. Opis kontinualne distribucije preko funkcije gustine verovatnoće se obično pravi tako što se raspon mogućih vrednosti kontinualne varijable podeli u određeni broj intervala, a onda se izračuna koliko entiteta ima u svakom od intervala. Relativni odnosi ovih brojeva (proporcije entiteta u svakoj kategoriji u odnosu na ukupan broj entiteta u uzorku) se onda tretiraju kao verovatnoće da se slučajno odabrani entitet nađe unutar određenog intervala i na osnovu tih verovatnoća se onda konstruiše funkcija gustine verovatnoće.

Za **osnovno opisivanje distribucije**, potrebno je da uvedemo sledeće pojmove:

- **Frekvencija** – je **broj entiteta koji imaju određenu vrednost varijable**. Frekvenciju određene vrednosti varijable tj. određene kategorije varijable dobijamo tako što izbrojimo ukupan broj entiteta u uzorku koji imaju datu vrednost tj. tako što izbrojimo koliko entiteta pripada kategoriji čiju frekvenciju računamo.
- **Proporcija** – predstavlja **udeo ispitanika koji imaju određenu vrednost** tj. **koji spadaju u određenu kategoriju u ukupnom broju entiteta u uzorku.** Proporcija se dobija tako što se **frekvencija podeli sa ukupnim brojem entiteta u uzorku**: proporcija = frekvencija/ukupan broj entiteta u uzorku). Proporcije se izražavaju na skali od 0 do 1. Ako je proporcija neke vrednosti 0, to znači da u uzorku nema entiteta sa tom vrednošću. Ako je proporcija 1, to znači da svi entiteti u uzorku imaju datu vrednost tj. da je ispitivana varijabla konstanta, a ne varijabla, jer da bi neko svojstvo bilo varijabla potrebno je da entiteti imaju različite vrednosti tog svojstva. Ako svi entiteti imaju istu vrednost, u pitanju je konstanta.
- **Procenat (%)** - je **proporcija pomnožena sa 100**. Procenat je isti kao proporcija, samo što je prebačen na raspon od 0 do 100. Procenat se označava znakom %. Podjednako kao i proporcija, i procenat ukazuje na relativni udeo entiteta koji spadaju u određenu kategoriju

PREPRINT

ili koji imaju određenu vrednost u ukupnom uzorku, samo što procenti to rade koristeći raspon brojeva koji je zgodniji za praktičnu komunikaciju i upotrebu (mali, često celi brojevi, naspram decimalnih brojeva manjih od 1).

3.2. Percentili i ostali kvantili

Nekada postoji potreba da se označe određene pozicije na distribuciji varijable koja nije nominalna. Ovo se najčešće radi korišćenjem percentila i percentilnih rangova.

- **Percentili** predstavljaju **tačku na distribuciji ispod koje ili u nivou s kojom se nalaze vrednosti određenog procenta entiteta iz uzorka**. Drugim rečima, to je vrednost koja je veća ili jednaka od tačno određenog procenta vrednosti entiteta iz uzorka. Percentil se imenuje na osnovu procenta entiteta koji imaju vrednosti niže od datog percentila ili jednake njemu. Na primer, 50i percentil je vrednost koja je veća ili jednaka vrednosti tačno 50% entiteta u uzorku na posmatranoj varijabli. 100i percentil je najveća vrednost u uzorku. 0 (nulti) percentil je najniža vrednost u uzorku. 20i percentil je vrednost ispod koje ili u nivou s kojom se nalazi tačno 20% entiteta u uzorku (dok 80% entiteta ima vrednosti veće od tog percentila). **Percetili se obično računaju tako što se entiteti iz uzorka poređaju u rastući ili opadajući niz na osnovu njihovih vrednosti na posmatranoj varijabli i onda se krene od najmanje vrednosti i ide na gore dok se ne stigne do vrednosti od koje ne manje od potrebnog procenta entiteta ima niže vrednosti i u odnosu na koju bar traženi procenat entiteta ima manje ili jednake vrednosti**. Ova metoda računanja percentila se zove **metoda najbližeg ranga.**
- **Percentilni rang** je **procenat entiteta u uzorku koji imaju niže ili jednake vrednosti od posmatranog entiteta**. Prema tome, percentilni rang je svojstvo određenog entiteta (svojstvo koje je relativno u odnosu na uzorak kom entitet pripada, ali nešto što se pripisuje entitetu). Percentilni rang je **blisko povezan sa percentilima, ali percentil je vrednost varijable, tačka na distribuciji, dok je percentilni rang svojstvo koje se pripisuje pojedinačnom entitetu**. U praksi, rekli bismo, na primer, da je 30i percentil određena vrednost (određeni broj), ali bi za entitet čija je vrednost na toj varijabli jednaka 30om percentilu rekli da je njegov percentilni rang 30. Na primer: “30i percentil našeg uzorka na

PREPRINT

varijabli X je 42 (42 je ovde vrednost na varijabli)", ali "Milicin skor na testu je 42, što znači da je njen percentilni rang 30".

Percentili i percentilni rangovi su najčešće korišćeni markeri pozicija na distribuciji (tj. toga koliko je visoka neka vrednost u odnosu na entitete u uzorku) i oni su zasnovani na procentima tj. na podeli uzorka na stote delove (1/100). To znači da bi i bilo koja druga podela na delove tj. deljenje distribucije na neki drugi broj jednakih delova, bila jednako validna. Opšti naziv za takve delove dobijene podelom distribucije na određeni broj jednakih delova je **kvantili** (Harding et al., 2014)**,** a u literaturu se sreće i naziv **n-tili**, pri čemu "n" označava taj broj delova na koje se cela distribucija deli. Percentili su tako podela distribucije na 100 jednakih delova, ali podjednako validno to može da bude i bilo koji drugi broj jednakih delova. Drugi često **korišćeni kvantili**, pored percentila su:

- **Kvartili** – dele **distribuciju na 4 jednaka dela**. Zbog toga imamo 4 kvartila. Prvi kvartil odgovara 25om percentilu, drugi kvartil odgovara 50om percentilu, treći kvartil odgovara 75om percentilu i četvrti kvartil odgovara 100om percentilu. Treba primetiti i da **neki autori koriste kvartile ne kao pokazatelje pozicije na distribuciji, već da označe intervale na distribuciji**. U takvom sistemu, prvi kvartil se odnosi na četvrtinu uzorka sa najnižim vrednostima, drugi kvartil se odnosi na entitete čije su vrednosti između 25og i 50og percentila, treći kvartil se odnosi na vrednosti između 50og i 75og percentila i četvrti kvartil čini četvrtina ispitanika sa najvišim vrednostima na posmatranoj varijabli.
- **Kvintili** – dele **distribuciju na 5 jednakih delova.** Svi detalji koji su opisani za kvartile važe jednako i za kvintile s tim da je jedina razlika to što je uzorak sada podeljen na 5 jednakih grupa umesto na 4.
- **Decili** – dele **distribuciju na 10 jednakih grupa**. Prvi decil odgovara 10om precentilu, drugi odgovara 20om i tako dalje. I decili se nekada koriste da označe intervale umesto tačaka na distribuciji i u tom slučaju se prvi decil odnosi na 10% entiteta sa najnižim vrednostima u uzorku, drugi decil se odnosi na one sa vrednostima između 10og i 20og percentila i tako dalje.

Percentili, ali i drugi kvantili se koriste kada god postoji potreba da se označi pozicija pojedinačnog entiteta na distribuciji određene varijable koja je bar na ordinalnom nivou merenja. Na primer, percentili se koriste u psihologiji kao primarni statistici za interpretaciju nivoa izraženosti različitih

PREPRINT

psiholoških osobina merenih psihološkim testovima u okviru takozvanog normativnog pristupa interpretaciji rezultata testiranja (za više detalja videti npr. Hedrih (2020)). U obrazovanju, kvantili se mogu koristiti da se grupa učenika podeli na nekoliko grupa prema postignuću na testiranju. U kliničkoj proceni, kvantili se koriste za razlikovanje klinički relevantnih rezultata od rezultata koji se smatraju "normalnim" ili tipičnim. U ekonomiji, kvantili se često koriste da podele kompanije koje rade u određenoj oblasti u kategorije prema poslovnom uspehu. U mnogim drugim oblastima često postoji potreba za podelom uzorka na više grupa na osnovu vrednosti određenih važnih varijabli ili za označavanjem određenih pozicija i to je potreba koju kvantili zadovoljavaju.

Upotreba kvantila zahteva da podaci budu **barem na ordinalnom nivou merenja**.

****

Dok je računanje frekvencija i procenata i mera izvedenih iz njih sasvim dovoljno za opisivanje distribucija nominalnih varijabli ili varijabli sa malim brojem kategorija, kod viših nivoa merenja uobičajenije je to da se uzorak opisuje računanjem statistika koji označava najtipičniju ili neku centralnu vrednost (ili vrednosti) varijable kao i statistika koji opisuje veličinu razlika između vrednosti entiteta u uzorku. Ovi prvi statistici nazivaju se merama centralne tendecnije, a ovi drugi merama varijabilnosti ili merama disperzije.

3.3. Mere centralne tendencije

Mere centralne tendencije su kategorija statistika čija je svrha da označe centralnu tendenciju entiteta u uzorku u odnosu na vrednosti posmatrane varijable. U idealnom slučaju, oni predstavljaju tačku na distribuciji tj. vrednost varijable za koju je najverovatnije da je entiteti u uzorku imaju ili oko koje teže da se grupišu vrednosti (na višim nivoima merenja). Najpoznatije mere centralne tendencije su aritmetička sredina, poznata još i kao "prosek", medijana i mod.

Aritmetička sredina, **prosečna vrednost** ili samo **prosek** se računa tako što se **saberu vrednosti svih entiteta u uzorku i taj rezultat se onda podeli brojem entiteta u uzorku**:

$$Mean = \frac{\sum_{i=1}^{n} X_i}{N}$$

U naučnoj literaturi, aritmetička sredina se **tipično obeležava sa slovom M ili sa grškim slovom μ**. U litaraturi **na srpskom jeziku** i drugim srodnim jezicima, često se sreće i oznaka **AS**. Aritmetička sredina se **može smisleno računati za podatke koji su na intervalnom ili racio nivou merenja.** Međutim, **formula za računanja aritmetičke sredine** se **može nekada koristiti i na ordinalnim podacima, ali statistik dobijen na ovakav način se zove prosek rangova ili prosečan rang.** Prosek rangova može biti koristan statistik u situacijama kada imamo više grupa koje su rangirane zajedno na zajedničkoj rang listi i u tom slučaju prosečan rang označava da li članovi određene grupe teže da se nalaze u višem ili nižem delu rang liste i tako pokazuje da li vrednosti članova te grupe teže da budu među višim ili među nižim vrednostima u odnosu na ostatak uzorka. U najvećem broju slučajeva, ima malo smisla računati prosek rangova za ceo uzorak, zato što ako su naši ordinalni podaci zapravo rang lista, prosek rangova će prosto biti jednak polovini broja entiteta u uzorku (ako su entiteti rangirani od prvog do poslednjeg) ili broju koji ne znači ništa posebno ako sistem ordinalnih vrednosti radi na neki drugačiji način, a u odsustvu fiksne merne jedinice.

Jedno veoma **važno svojstvo aritmetičke sredine** kao mere centralne tendencije je to da **na nju utiču vrednosti svih entiteta u uzorku**, te su, na taj način, ovom merom svi entiteti uzeti u obzir. Međutim, ovo isto svojstvo ima važnu negativno stranu koja se sastoji u tome da **na vrednost aritmetičke sredine mogu veoma mnogo uticati pojedinačne ekstremne vrednosti** ili **vrlo male grupe entiteta sa neobičnim vrednostima**. U praksi, to znači da, na primer, ako bismo hteli da izračunamo prosečne prihode jedne grupe ljudi i ta grupa se sastoji od 100 ljudi koji zarađuju 100 dolara godišnje i jedne osobe koja zarađuje milion dolara godišnje, prosečni prihodi grupe bi bili 10 000 dolara godišnje. I tu je očigledno da je prosečna vrednost od 10 000 dolara, kao opis centralne tendencije ove grupe, prilično beskorisna i obmanjujuća jer omašuje primarnu svrhu mere centralne tendencije, a to je da ukaže da tipične vrednosti ili na centralnu tendenciju grupe. Poznata je i stara šala popularna među ljudima koji uče statistiku, a koja kaže da ako imamo dve grupe ljudi od kojih jedna za jelo ima samo kupus, a druga za jelo ima šnicle, da te dve grupe u proseku jedu kupus s mesom. Takvo zaključivanje je očigledno pogrešno. Postoje takođe i poznati primeri različitih vlada širom sveta koje (zlo)upotrebljavaju aritmetičku sredinu da bi prikazali plate u zemlji većim nego što stvarno jesu (za većinu ljudi) tako što prikazuju prosečnu platu kao zvanični statistik, pri tom zanemarujući ili umanjujući činjenicu da je prosek plata pomeren dosta ka višim vrednostima na distribuciji plata zbog malog broja veoma visokih plata

PREPRINT

(slično kao u malopređašnjem primeru). U takvim situacijama, prosečna plata je broj koji je veći (nekada i dosta veći) od onog što zapravo većina ljudi zarađuje. To je razlog zbog kog je **dodatni zahtev za računanje aritmetičke sredine kao mere centralne tendencije to da vrednosti entiteta budu raspodeljene tako da je većina grupisana oko jedne centralne tačke i da ima sve manje i manje entiteta kako se vrednosti udaljavaju od te centralne tačke**. U kasnijem delu ove knjige diskutovaćemo o teorijskim distribucijama i videćemo da mnogi autori smatraju takozvanu **„normalnu distribuciju" podataka za neophodni preduslov da bi aritmetička sredina imala smisla** kao mera centralne tendencije.

Mere centralne tendencije slične aritmetičkoj sredini su i harmonijska sredina i geometrijska sredina, ali ćemo ove mere centralne tendencije retko sresti u društvenim naukama i naukama o ponašanju. **Harmonijska sredina se računa tako što se broj entiteta u uzorku podeli sa sumom recipročnih vrednosti entiteta u uzorku:**

$$Harmonijska\ AS = \frac{N}{\sum_{i=1}^{n} \frac{1}{X_i}}$$

gde je N broj entiteta u uzorku, i je redni broj entiteta, a X_i je vrednosti i-tog ispitnika.

Geometrijska sredina je Nti koren proizvoda vrednosti svih entiteta u uzorku, pri čemu N predstavlja broj entiteta u uzorku. Dobro je biti svestan postojanja ovih mera centralne tendencije, ali je takođe činjenica da ćemo ih veoma retko, ako ikad, sresti primenjene u praksi društvenih nauka i nauka o ponašanju. Ove mere, ipak, imaju široku primenu u drugim naukama.

Medijana predstavlja centralnu tačku distribucije, tačku ispod i iznad koje se nalaze vrednosti po 50% entiteta iz uzorka (50% ispod medijane, 50% iznad medijane). Drugim rečima, **medijana je 50i percentil** ili 2. kvartil. Da bi računanje medijane bilo smisleno, potrebno je da **podaci budu barem na intervalnom nivou merenja**, iako postoje situacije kada je medijanu moguće smisleno izračunati i na ordinalnom nivou merenja (npr. neke situacije računanja medijane na podgrupama radi poređenja).

U nekim aspektima, svojstva medijane su suprotna od svojstava aritmetičke sredine – i **na medijanu utiču svi entiteti u uzorku, ali ne svojim vrednostima, već samo svojim pozicijama u odnosu na centralni deo distribucije.** Na ovaj način, **na medijanu ne utiču ekstremne**

PREPRINT

vrednosti pojedinačnih entiteta, bez obzira koliko da su ekstremne. To je, s jedne strane, dobra stvar jer će vrednost medijane uvek biti vrednost nekog entiteta u distribuciji. U prethodnom primeru, gde 100 ljudi zarađuje po 100 dolara godišnje i jedna osoba zarađuje 1 000 000 dolara godišnje, medijana bi bila 100 dolara zato što je to centar distribucije tj. 50i percentil. U takvoj situaciji, 100 dolara je svakako bolja procena centralne tendencije zarada ljudi u tom uzorku, zato što ova vrednost jasno ukazuje na nivo prihoda koju većina ljudi u uzorku ima. Međutim, ako hoćemo da realistično procenimo grupu poput ove, grupu u kojoj, pored većine koja je praktično bez prihoda, ima i jedan koji zarađuje milion, moramo da zaključimo da zanemarivanje postojanja takve osobe takođe predstavlja pogrešnu procenu ove grupe – grupa od 101 osobe u kojoj svako zarađuje po 100 dolara godišnje je verovatno dosta različita od grupe u kojoj je ta osoba sa milion dolara godišnje i to upravo zbog postojanja te jedne osobe. Takođe, kada bi se naš uzorak sastojao od 51 osobe koja zarađuje 100 dolara godišnje i još 49 ljudi koji zarađuju milion dolara godišnje, medijana bi i dalje bila 100 dolara i tu bi se već vrlo lako složili da je 100 dolara još gora procena centralne tendencije prihoda ove grupe od one iz prethodnog primera. Zbog svega ovog, najsmislenija upotreba medijane je takođe u situacijama kada su vrednosti većine entiteta iz uzorka koncentrisane oko jedne centralne vrednosti, pri čemu broj entiteta postaje sve manji i manji kako se vrednosti udaljavaju od te grupe centralnih vrednosti. Zvanično, zahtevi za računanje medijane su manje strogi nego oni za računanje aritmetičke sredine (kao što ćemo videti u kasnijim poglavljima), međutim, svrha i korisnost medijane kao mere centralne tendencije se brzo gubi sa udaljavanjem oblika distribucije od ovog malopre opisanog.

Mod je **najčeća vrednost u uzorku** tj. vrednost sa najvećom frekvencijom. Uzorak može da ima više od jednog moda. Takva situacija, da uzorak ima više od jednog moda, se dešava onda kada dve ili nekoliko najčešćih vrednosti imaju jednake frekvencije. Takođe, **kada se izveštava o modu, a pri tom postoji dve ili više vrednosti koje su upadljivo česte u odnosu na frekvencije ostalih vrednosti, mnogi autori će izvestiti o svim tim vrednostima, čak iako njihove frekvencije nisu skroz identične tj. čak i kada je, tehnički gledano, samo jedna od njih mod**. Razlog za ovo je to što će autori u takvim situacijama smatrati da izveštavanje o svojstvima uzorka ne bi bilo valjano ako bi propustili da navedu i ostale najfrekventnije vrednosti pored one koja matematički predstavlja mod. Mod **se može računati na podacima koji su na nominalnom nivou** merenja i on je zapravo **jedina široko korišćena mera centralne tendencije koja se može valjano računati na nominalnim podacima.** Mod **se može računati i na višim nivoima**

PREPRINT

merenja, ali takvi postupci su **smisleni samo ako radimo sa diskretnim merama sa malim brojem različitih vrednosti u odnosu na veličinu uzorka**, odnosno ako su mere spojene u relativno mali broj grupa na osnovu intervala vrednosti (navođenje raspona vrednosti koje će predstavljati istu kategoriju, a pri tom obezbediti da ukupan broj takvih kategorija bude mali). **Ako je broj različitih vrednosti varijable koje entiteti imaju veliki u odnosu na veličinu uzorka**, onda možemo imati situaciju u kojoj većina vrednosti ima frekvenciju 1 ili će ta frekvencija biti neki nizak jednocifreni broj i **to je onda situacija kada će to koja kategorija ima najveću frekvenciju biti stvar slučaja, a ne pokazatelj bilo kakve centralne tendencije uzorka**. Primer ovakve situacije bi bilo kada bi imali, na primer, grupu ljudi čiju bi visinu merili veoma precizno i beležili je u mikronima (1 mikron je 1/1000 milimetra, hiljaditi deo milimetra). Sa takvom preciznošću i veoma malim uzorkom, vrlo je verovatno da ne bismo našli čak ni dve osobe koje imaju identičnu visinu do u mikron, a i da nađemo dve takve osobe i da to proglasimo za mod, takav mod teško da bi ukazivao na bilo kakvu opštu centralnu tendenciju uzorka. Međutim, ako bi zaokružili vrednosti visina na decimetre – na primer tako da oni ispod 1,5 metara budu prva kategorija, oni od 1,5 do 1,6 metara druga kategorija, oni između 1,7 i 1,8 metara treća kategorija i tako dalje, mod bi bio mnogo smisleniji pokazatelj centralne tendencije u oblasti visine.

Mod se nekad naziva i **modalna vrednost**. To možemo da sretnom u literaturi u formulacijama poput “Mod uzorka je 23”, ali i “modalna vrednost je 23”.

3.4. Mere varijabilnosti

Mere varijabilnosti pokazuju nivo varijabilnosti entiteta u uzorku tj. veličinu razlika između entiteta u uzorku. Najčešće korišćene mere varijabilnosti su standardna devijacija, varijansa, kvartilna, devijacija, raspon i koeficijent varijacije.

Standardna devijacija se računa tako što se aritmetička sredina uzorka oduzme od vrednosti svakog entiteta, tako dobijena razlika se digne na kvadrat, zatim se izračuna prosek tako dobijenih kvadrata razlika i iz tog proseka izvadi kvadratni koren. U naučnoj literaturi, **standardna devijacija se tipično označava slovima SD ili grčkim slovom σ.** Standardnu devijaciju je **moguće smisleno izračunati samo ako su podaci na intervalnom ili racio nivou merenja**. Za opisivanje uzorka, ona se **tipično uparuje sa aritmetičkom sredinom** – u situacijama kada

PREPRINT

koristimo aritmetičku sredinu da opišemo uzorak, standardnu devijaciju koristimo da opišemo njegovu varijabilnost.

Formula za računanje standardne devijacije je:

$$SD = \sqrt{\frac{\sum_{i=1}^{n}(X_i - M)^2}{N}}$$

ili

$$SD = \sqrt{\frac{\sum_{i=1}^{n}(X_i - M)^2}{N - 1}}$$

gde je M aritmetička sredina uzorka, sa X je vrednost pojedinačnog entiteta, *i* se odnosi na redni broj entiteta, ∑ označava sumu, što znači da nakon što smo oduzeli vrednosti svakog pojedinačnog entiteta od aritmetičke sredine uzorka i kvadrirali rezultat, treba da izračunamo sumu svih tih kvadriranih razlika. N i *n* se odnose na broj entiteta u uzorku. Često se N u ovoj formuli menja za N-1 što predstavlja vrednost koja se zove **broj stepeni slobode.** U statistici, **stepeni slobode predstavljaju broj vrednosti koje su slobodne da variraju, a da zadati uslovi ne budu narušeni**. U ovom slučaju se računa tako što se broj odnosa oduzima od broja opservacija (tj. broja entiteta u uzorku). Za standardnu devijaciju, broj stepeni slobode je uvek N-1. Iako formula sa brojem stepeni slobode važi za matematički ispravniju, sa praktične strane društevnih, kao i većine drugih nauka, a imajući u vidu tipične veličine uzorka, **obično nema nikakve praktične razlike između toga da li ćemo za računanje standardne devijacije koristiti u formuli broj entiteta ili broj stepeni slobode** tj. da li ćemo koristiti prvu ili drugu formulu. U najvećem broju realnih slučajeva, deljenje nečega sa 200 ne dovodi do bitno različitog rezultata u odnosu na to kad delimo sa 199 (za uzorak od 200 entiteta) ili, još bolje, vrlo je mala razlika u rezultatu ako podelimo sa 1300 u odnosu na ono što bi dobili kada delimo sa 1299 (za uzorak od 1300 entiteta). Ova razlika može biti vidnija kada su uzorci vrlo mali, ali takvi uzorci se u naučnim istraživanjima sreću retko ili nikad, a i kada se sretnu standardna devijacija obično nije najbolji način da se opiše njihova varijabilnost. Na primer, ne bi imalo puno smisla opisivati uzorak od 5 ili 10 entiteta standardnom devijacijom. Sa tako malim uzorkom bilo bi mnogo smislenije pobrojati njihove frekvencije njihovih vrednosti.

PREPRINT

Kao i kod aritmetičke sredine, **važno svojstvo standardne devijacije je to da na njenu veličinu utiče veličina odstupanja svih entiteta u uzorku od aritmetičke sredine**, kao i to da **odstupanja povećavaju veličinu standardne devijacije bez obzira na koju stranu je to odstupanje** (bez obzira da li je vrednost entiteta veća ili manja od aritmetičke sredine, zbog toga što se predznak razlike gubi kada se rezultat kvadrira, te su zato sva kvadrirana odstupanja pozitivna bez obzira na to u kom smeru je bila sama razlika. Kao i na aritmetičku sredinu, i **na standardnu devijaciju veoma mnogo utiču entiteti sa ekstremnim vrednostima** – pojedinačne ekstremne vrednosti mogu vidljivo povećati veličinu standardne devijacije, tako da sve prednosti i mane aritmetičke sredine, koje su već pomenute, važe i za standardnu devijaciju. **Zahtev o obliku distribucije** koji je potreban da bi računanje aritmetičke sredine imalo smisla**, važi podjednako i za računanje standardne devijacije**. Treba reći i da, **kada je svojstvo koje ispitujemo konstanta, onda će standardna devijacija biti 0.**

Varijansa – računanje standardne devijacije uključuje vađenje kvadratnog korena iz proseka kvadriranih odstupanja vrednosti entiteta od aritmetičke sredine. Ako preskočimo taj poslednji korak i ne izvadimo kvadratni koren iz tog rezultata, već ostanemo samo na računanju proseka kvadriranih odstupanja vrednosti entiteta od aritmetičke sredine, dobijamo meru varijabilnosti koja se zove varijansa. Drugim rečima, **varijansa je kvadrat standardne devijacije** i obrnuto – standardna devijacija je kvadratni koren varijanse. To je razlog zašto se **varijansa često i označava kao kvadrat standardne devijacije - σ^2, ali se varijansa često označava i slovom V ili slovima VAR**. Kao mera varijabilnosti, varijansa funkciniše jednako kao i standardna devijacija. Imajući u vidu da je u pitanju samo kvadrat standardne devijacije, standardna devijacija i varijansa su u potpunom skladu kada su u pitanju zaključci o varijabilnosti koji se mogu izvesti iz njih. Međutim, **varijansa je statistik koji se veoma često koristi u različitim multivarijatnim statističkim tehnikama** (nisu predstavljeni u ovoj knjizi), kao i u različitim složenim statističkim postupcima.

Medijana apsolutnih odsupanja je **medijana apsolutnih vrednosti odstupanja entiteta u uzorku od medijane uzorka**. Drugim rečima, to je medijana razlika između vrednosti entiteta i medijane uzorka, pri računanju koje smo zanemarili smer tih razlika (tj. to da li je vrednost određenog entiteta viša ili niža od medijane). Računa se tako što se **vrednost medijane uzorka oduzme od vrednosti svakog pojedinačnog entiteta, ukloni se predznak između tako dobijene**

razlike (+ ili -, drugim rečima pretvori se razlika u apsolutnu vrednost), a onda se izračuna **medijana tih razlika** i to se onda zove medijana apsolutnih odstupanja. Odnos između medijane i medijane apsolutnih odstupanja je sličan odnosu između aritmetiče sredine i standardne devijacije. Međutim, u trenutku pisanja ove knjige, **medijana apsolutnih odstupanja se prilično retko sreće** u naučnim publikacijama, posebno u oblasti društvenih nauka i nauka o ponašanju. Medijana apsolutnih odstupanja se može smisleno računati ako su podaci bar na intervalnom nivou merenja.

Kvartilna devijacija je **polovina razlike između 75og i 25og percentila** ili, drugim rečima, to je polovina razlike između 3eg i 1og kvartila (otud i ime – kvartilna devijacija). **Razlika između trećeg i prvog kvartila** zove se **interkvartilni opseg**, te se onda kvartilna devijacija može definisati i kao **polovina interkvartilnog opsega**. Kvartilna devijacija u grubim crtama ukazuje na veličinu raspršenja centralnog dela distribucije. Tipično se koristi zajedno sa medijanom. Upotreba kvartilne devijacije je u punoj meri **smislena ako su podaci bar na intervalnom nivou merenja**, jer je fiksna jedinica mere potrebna da bi podatak o veličini raspršenja bio smislen. Međutim, postoje i situacije kada se ova mera smisleno može koristiti i na ordinalnom nivou merenja.

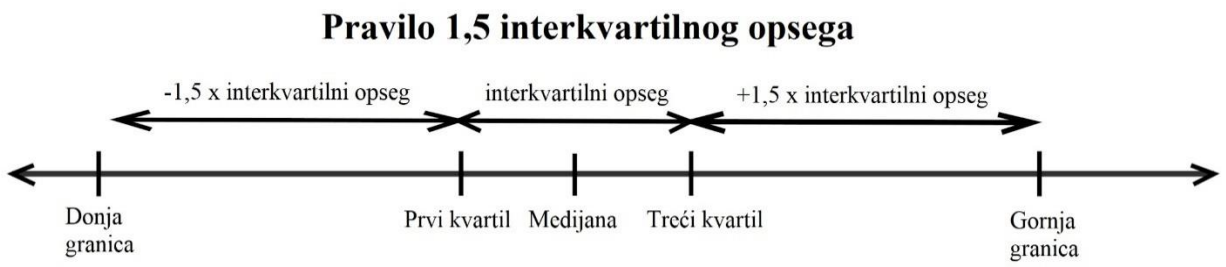
Kao i na medijanu, na kvartilnu devijaciju utiče odnos vrednosti entiteta i vrednosti drugih entiteta u uzorku, ali ne vrednost entiteta sama za sebe. Zbog toga, za razliku od standardne devijacije, **na kvartilnu devijaciju ne utiču ekstremne vrednosti** entieta na posmatranoj varijabli.

Raspon je razlika između najveće i najmanje vrednosti u uzorku. Raspon pokazuje veličinu intervala unutar kog se nalaze vrednosti svih entiteta u uzorku. Iako je poznavanje veličine ovog intervala svakako korisno, **problem sa rasponom je to što raspon isključivo određuju dva najekstremnija slučaja u uzorku** – najniža vrednost i najviša vrednost. Problem sa ovim merama u istraživačkoj praksi je to što su **one često posledice loših ili nevalidnih merenja**, a nekada, kada istraživači nisu dovoljno pažljivi, i **greške pri unosu podataka**. Kada je takva situacija u pitanju, raspon kao mera postaje besmislen. Zbog toga je nekada korisno računati takozvani **skraćeni raspon** tj. **raspon izračunat na uzorku iz kog je uklonjeno po nekoliko entiteta sa najvišim i isto toliko sa najnižim vrednostima. Nekada se skraćeni raspon računa tako što se isključe vrednosti entiteta koje se upadljivo razlikuju od vrednosti ostalih entiteta u uzorku** tj. čije se vrednosti za više od neke unapred definisane vrednosti razlikuju od centralnog "tela distribucije"

PREPRINT

ili od određene tačke na distribuciji. Takvi entiteti nazivaju se u statistici **autlajerima** (od engleskog – outlier). Na primer, **jedan način da se odluči o tome koji entiteti su autlajeri je takozvano pravilo 1,5 interkvartilnog opsega (pravilo jednog ipo interkvartilnog opsega)** ili skraćeno 1,5 IQR pravilo (engleski 1.5 IQR rule). Ovo pravilo kaže da **treba da interkvartilni opseg pomnožimo sa 1,5 i da onda tu vrednost dodamo na 3. kvartil i da je oduzmemo od 1. kvartila. Svi entiteti čije se vrednosti nalaze van intervala koji definišu ove dve vrednosti predstavljaju autlajere prema ovom pravilu.**

Slika 3.1. Ilustracija pravila 1,5 interkvartilnog opsega na brojčanoj pravoj. Jedan ipo (1,5) interkvartilni opseg se dodaje na 75i percentil (tj. na treći kvartil) i ista vrednost se oduzima od 25og percentila (tj. prvog kvartila). Entiteti van opsega definisanog ovim dvema vrednostima smatraju se za autlajere.



3.5. Kako se distribucija može predstaviti?

Distribucija se može predstaviti tako što se prosto napravi spisak svih entiteta sa njihovim vrednostima tj. u vidu matrice, distribucija se tipično u literaturi predstavlja i u vidu tabela u kojima su predstavljene sve moguće vrednosti varijable sa njihovim frekvencijama, proporcijama ili procentima ili preko različitih grafičkih prikaza.

Diskretne distribucije se mogu predstaviti kroz pita-dijagrame, linijske dijagrame i slične grafičke prikaze, dok se kontinualne distribucije tipično predstavljaju korišćenjem histograma, boksplotova i grafičkih predstava funkcije gustine verovatnoće

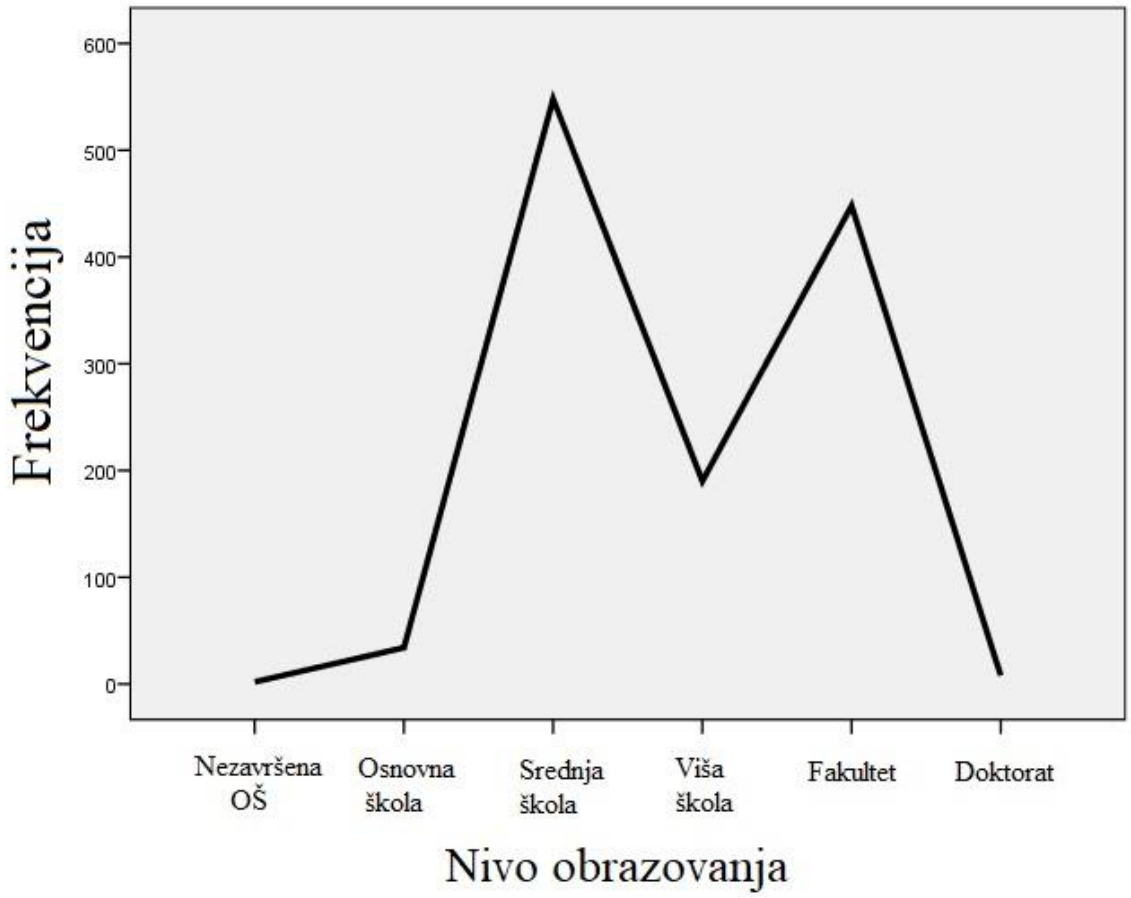
Tabela 3.2. i slike 3.3-3.9. (Anđelka Hedrih & Hedrih, 2012; V. Hedrih, 2017b; V. Hedrih et al., 2018; Kwon & Lee, 2020; Sobotka et al., 2020)

Tabela 3.2. Predstavljanje distribucije preko tabele sa podacima o procentima entiteta u svakoj kategoriji. Podaci prikazani u tabeli su odgovori učesnika u istraživanju Hedrih & Hedrih (2012) o potencijalnim donatorima sperme (donatorima za potrebe veštačke oplodnje) koji su pitani o svojstvima ljudi kojima bi bili spremni da doniraju spermu. Prva kolona su varijable, tj. pitanja koja su postavljena, a kasnije kolone u tom redu vrednosti ovih varijabli – mogući ponuđeni odgovori i procenti učesnika u istraživanju koji su dali svaki od odgovora. Na primer, možemo videti da je 47,8% odgovorilo da poznanstvo (sa primaocem donacije) nije važno za njihovu odluku od davanju donacije sperme. Takođe, iz trežeg reda, možemo videti da je 69,9% učesnika u istraživanju odgovorilo da bi dali saglasnost da njihova donacije sperme bude iskorišćena od strane (nevenčanog) heteroseksualnog para, dok je 90,3% odgovorilo da bi dali saglasnost da venčani par bude primalac njihove donacije sperme. Izvorna tabela je na engleskom, sadržaj je ovde preveden na srpski od strane autora knjige.

Pitanje	%					
Kome biste dali donaciju?	Osobe koje poznajem	Osobe koje ne poznaje m	I jedni i drugi	Poznanst vo nije od značaja		
	5,5	26,5	20,2	47,8		
Kojim kategorijama biste dali donaciju? (% štiklirao)	Bračni par	Heterose ksualni par	Lezbej ski par	udovica	Ženska osoba bez partnera	Razvede na žena
	72,3	32,0	12,9	27,7	40,3	29,0
Da li biste dali saglasnost da spermu koju ste donirali upotrebi....(% da)	Bračni par	Heterose ksualni par	Lezbej ski par	Ženska osoba bez partnera	Razvede na žena	
	90,3	69,9	22,2	61,2	54,1	

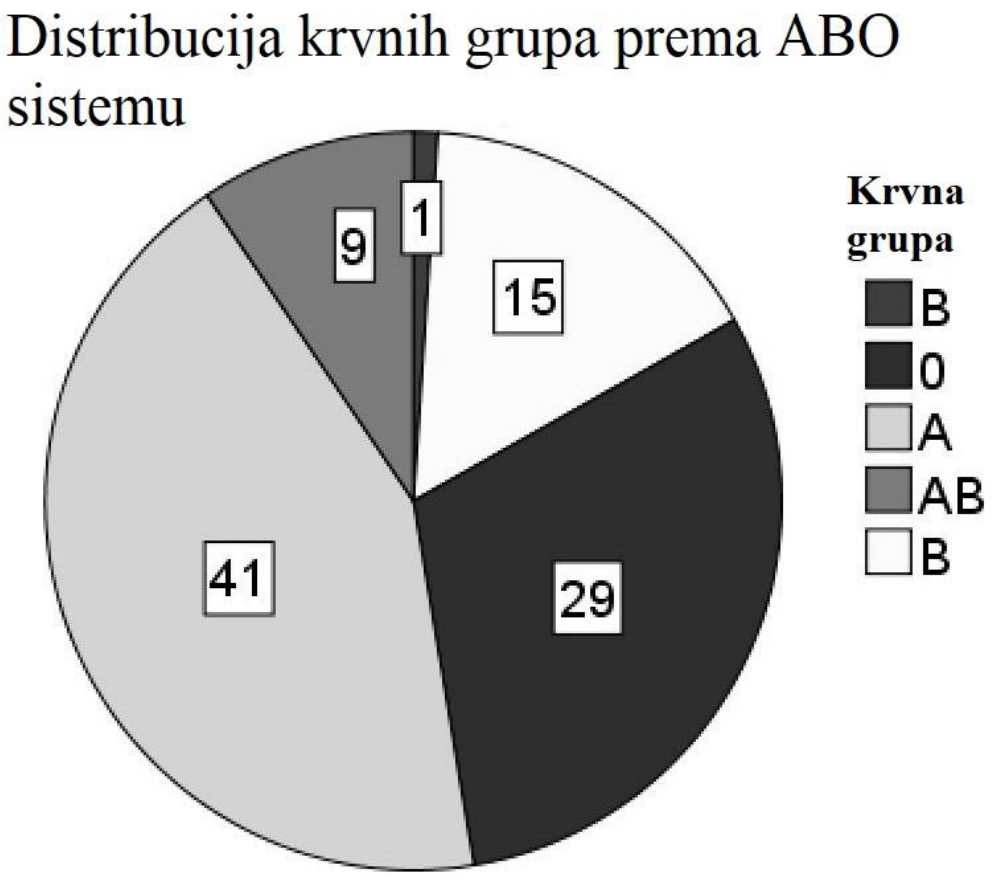
PREPRINT

Slika 3.3. Predstavljanje distribucije preko linijskog grafikona – distribucija nivoa obrazovanja na uzorku iz studije odnosa na poslu i u porodici u Srbiji. Vertikalna osoba predstavlja frekvencije tj. broj učesnika u istraživanju u određenoj kategoriji, dok su različiti nivoi obrazovanja predstavljeni na horizontalnoj osi (V. Hedrih, 2017a). Sa slike možemo da vidimo da je najčešći nivo obrazovanja u uzorku srednja škola, sa više od 500 učesnika koji su rekli da imaju ovaj nivo obrazovanja, što srednju školu čini modom varijable Nivo obrazovanja.

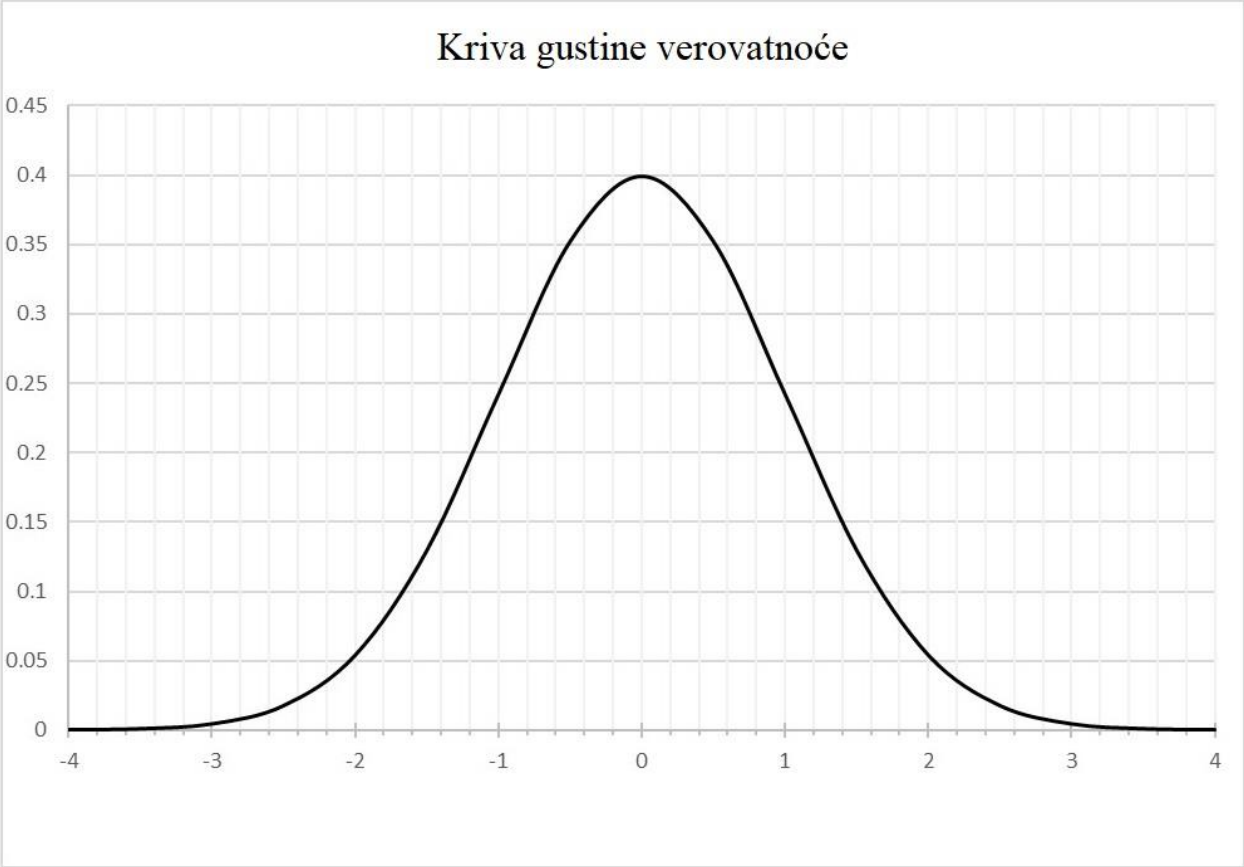


Slika 3.4. Predstavljanje distribucije preko pitastog dijagrama – distribucija krvnih grupa prema ABO sistemu, nominalna varijabla. Parčići pite predstavljaju udeo učesnika u istraživanju sa određenom krvnom grupom u uzorku. Brojevi na parčićima pite predstavljaju frekvencije tj. broj učesnika u istraživanju u datoj kategoriji. Ovaj pitasti dijagram je napravljen na osnovu podataka iz istraživanja Hedrihove i sar. (Andjelka Hedrih et al., 2018) o uticaju mikrotraume šake na patogenezu i razvoj osteoartroze šake i vrata. Sa ove konkretne slike možemo videti da najveći

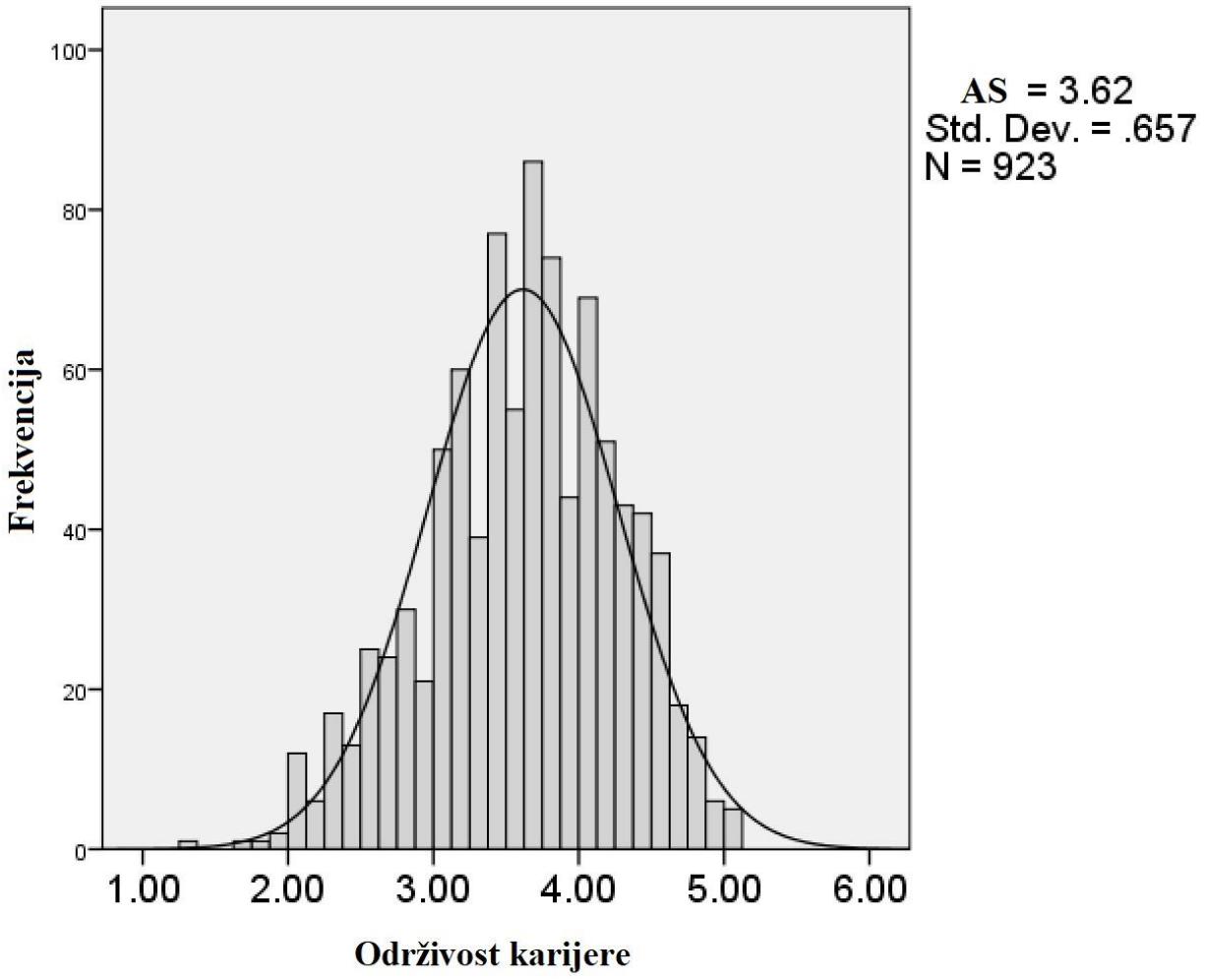
broj učesnika – 41, pripada krvnoj grupi A, što čini A krvnu grupu modom ovog uzorka. Najređa krvna grupa je B, sa samo jednim učesnikom u ovom istraživanju čija krv pripada ovoj grupi.



Slika 3.5. Prikaz distribucije varijable na intervalnom nivou merenja preko funkcije gustine verovatnoće – horizontalna osa predstavlja vrednosti varijable, dok vertikalna osa predstavlja verovatnoće (date kao proporcije) da slučajno odabrani entitet iz populacije ima određenu vrednost.

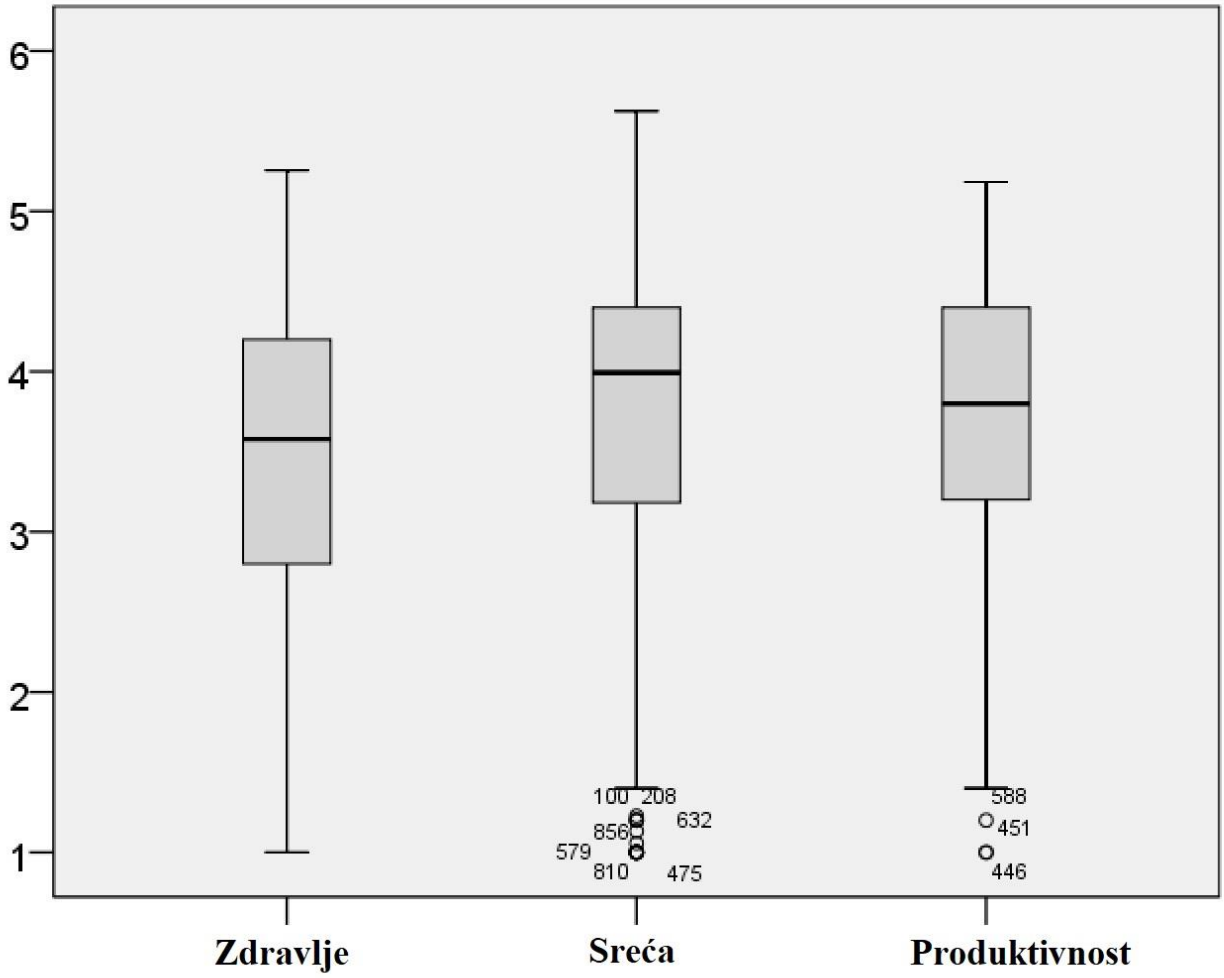


Slika 3.6. Prikaz distribucije varijable na intervalnom nivou merenja preko histograma – slika predstavlja distribuciju uzorka iz jednog neobjavljenog rada prvog autora (Hedrih&Milić) na percepciji održivosti karijere, što je jedan od psiholoških konstrukata koji meri Skala percepcije održivosti karijere, koju su razvili autori ovog istraživanja. Ovakav grafički prikaz se dobija pretvaranjem kontinualne varijable u skup intervala tj. skup diskretnih vrednosti, tako da je moguće izbrojati entitete unutar svakog intervala. Vertikalna osa predstavlja broj entiteta tj. učesnika u istraživanju u ovom slučaju, a visina svakog stuba odgovara broju entiteta u intervalu vrednosti varijable koji je predstavljen datim stubom. Preko stubova je nacrtana funkcija gustine verovatnoće zasnovana na obliku distribucije na koji ukazuju relativni odnosi veličina stubova histograma. Iz ove konkretne slike možemo da vidimo da učesnici u istraživanju teže da imaju najčešće vrednosti u centralnom delu, a da broj učesnika u istraživanju postaje sve manji i manji kako se vrednosti udaljavaju od ove centralne zone distribucije.



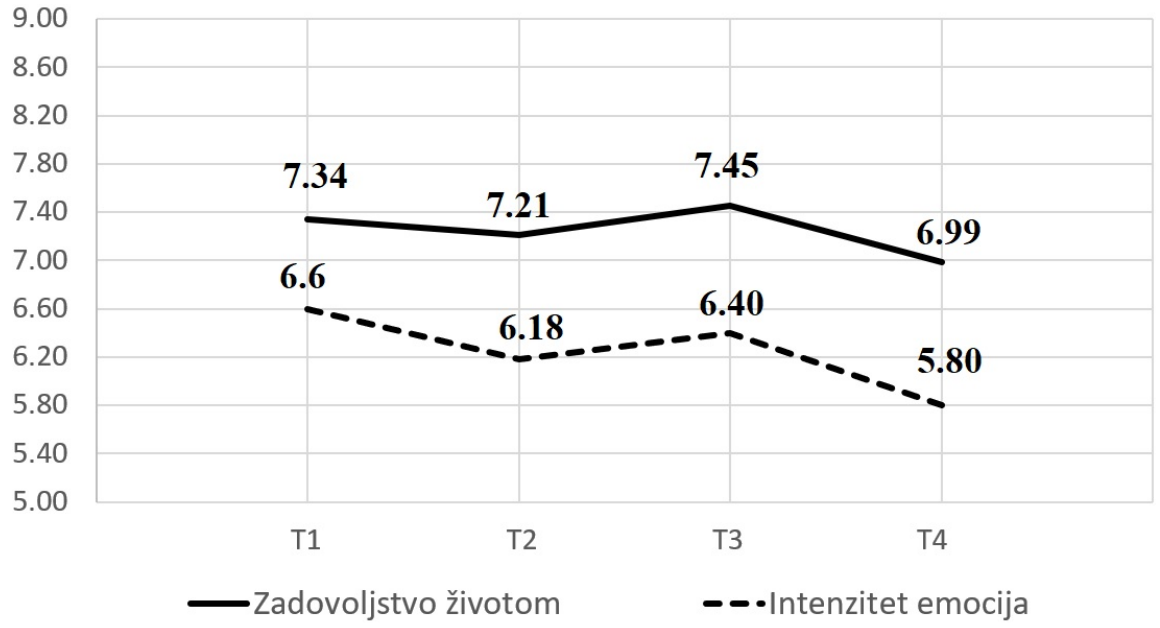
Slika 3.7. Prikaz tri distribucije varijable na intervalnom nivou merenja preko boksplotova – slika predstavlja distribucije skorova entiteta iz uzorka na tri skale percepcije održivosti karijere iz neobjavljenog istraživanja Hedriha i Milićeve. Vertikalna osba predstavlja vrednosti varijable (skale ovih varijabli su uporedive), na horizontalnoj osi su tri varijable i njihova imena su navedena. Po jedan boksplot je za svaku varijablu. Kutijasti deo boksplota predstavlja raspon vrednosti varijable unutar kog je grupisan najveći deo uzorka. Debela linija koja ide preko sredine svake kutije je aritmetička sredina uzorka u ovom slučaju, iako se boksplotovi mogu praviti i tako da debela linija pokazuje vrednost neke druge mere centralne tendencije. „Kanapi“ tj. tanke linije

koje idu na gore i na dole od kutije predstavljaju raspon vrednosti u kojima ima entiteta, ali su ređi. Konačno, brojevi van raspona vrednosti koji su pokriveni tankim linijama (i kutijama) predstavljaju autlajere. Na ovim boksplotovima možemo videti da ima više autlajera sa neobično niskim vrednostima na Sreći, kao i da postoje 3 autlajera sa veoma niskim vrednostima na Produktivnosti (ima samo 2 kruga zato što 2 od 3 autlajera imaju istu vrednosti, pa im se krugovim poklapaju). Dužina boksplota, a posebno dužina kutija pokazuje varijabilnost grupe koja je predstavljena boksplotom na posmatranoj varijabli – što je boksplot duži, a pogotovo što je kutija duža, veća je varijabilnost.



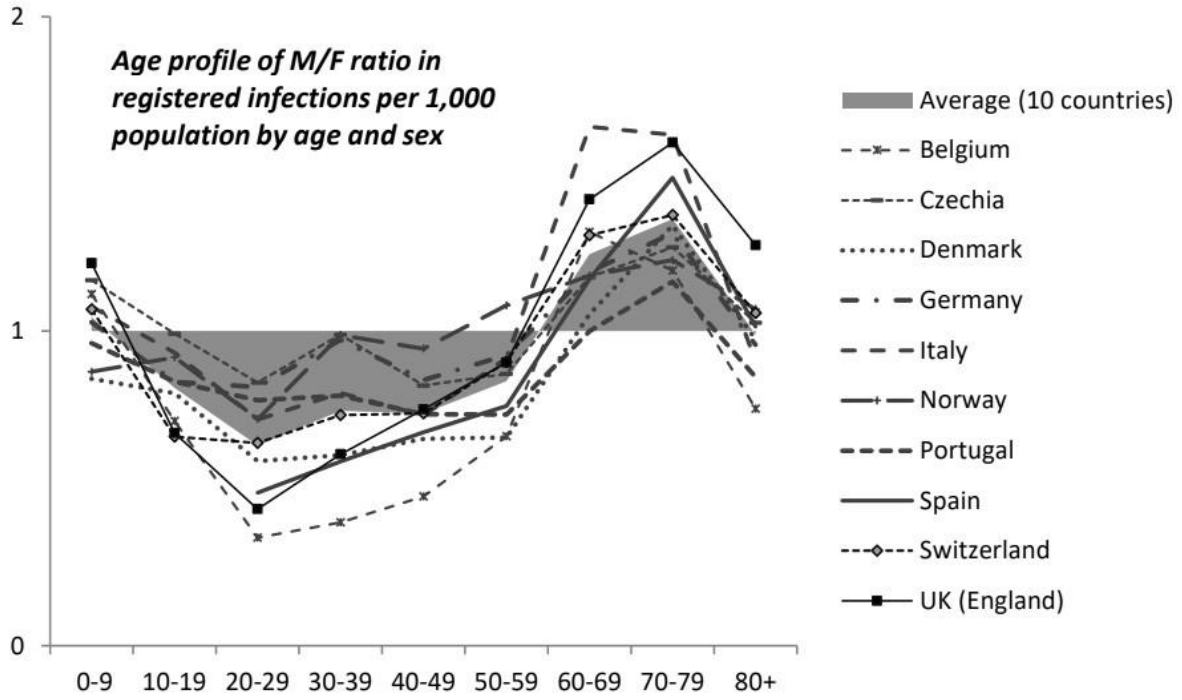
Slika 3.8. Linijski dijagram iskorišćen da prikaže centralne tendencije više uzoraka na dve različite varijable. Tačke na horizontalnoj osi su 4 različite vremenske tačke kada su mere uzimane, dok

vertikalna osa sadrži vrednosti varijable. Svaka linija predstavlja jednu varijablu. Tačke na liniji iznad T1, T2, T3 i T4, kao i brojevi napisani iznad linija predstavljaju aritmetičke sredine uzorka na dve varijable u tim vremenskim tačkama. Na primer, možemo videti da u vremenskoj tački T3, aritmetička sredina uzorka na varijabli Intenzitet emocija je 6,40, dok je 7,45 na Zadovoljstvu životom. Slika je zasnovana na izmišljenim podacima koji su inspirisani realnim istraživanjima.



Slika 3.9. Složeniji grafički prikaz poređenja više distribucija. Grafik predstavlja odnos broja inficiranih muškaraca i žena u različitim zemljama zavisno od pola. Svaka linija predstavlja različitu zemlju. Vertikalna osa predstavlja odnos broja muškaraca i žena inficiranih COVID-19 virusom, koji je dobijen tako što je broj registrovanih COVID-19 infekcija muškaraca podeljen brojem COVID-19 infencija žena iz iste starosne grupe, dok horizontalna osa predstavlja različite starosne grupe u godinama. Na slici možemo videti da u starosnoj grupi 20-29 godina, ima više žena nego muškaraca kod kojih su registrovane COVID-19 infekcije (zato što je odnos manji od 1, što znači da je delilac veći od deljenika u razlomku), dok u starosnim grupama 60-69 i 70-79 godina ima mnogo više zaraženih muškaraca nego žena. Slika je preuzeta iz preprinta Sobotke i sar (Sobotka et al., 2020). Preštampano na osnovu dozvole autora.

Figure : Relative ratio of male to female COVID-19 infection rates (M/F ratio) per 1,000 population by age; ten European countries



3.6. Hajde da primenimo šta smo do sada naučili!

Hajde da probamo sada da primenimo stvari koje smo predstavili u ovom poglavlju kroz nekoliko vežbi. Molimo vas da pogledate opšte uputstvo za ovakve vežbe koje možete naći na početku knjige. Naša preporuka je da prvo pročitate svaki isečak i tvrdnje date u njemu i da onda date svoj odgovor. Odgovor možete upisati u kolonu za odgovore, a posle toga pročitajte odgovore i uporedite svoje odgovore sa njima.

Vežba C. Distribucija, mere centralne tendencije, mere varijabilnosti.

Deskriptivna statistika			
		N=261	%
Pol	Ženski	89	34,1
	Muški	172	65,9
Starost u godinama	20-25	53	20,3
	26-30	109	41,8
	31-35	73	28,0
	36-40	26	10,0
Nivo obrazovanja	Osnovna škola	15	5,7
	Srednja škola	200	76,6
	Fakultet	46	17,6
Bračno stanje	U braku	60	23,0
	Nije u braku	201	77,0
Mesečni prihodi	do 40 000 RSD	25	9,6
	40 001 – 60 000 RSD	48	18,4
	60 001 – 90 000 RSD	44	16,9
	90 001 – 120 000 RSD	49	18,8
	Više od 120 000 RSD	107	36,4
Radni status	Zaposlen	184	70,5
	Student	40	15,3
	Ostalo	37	14,2

Podaci su izmišljeni.

C	Tvrdnja:	Odgovor
C1.	Brojevi u krajnje desnoj koloni su procenti.	
C2.	Pol je u ovoj tabeli operacionalizovan kao binarna varijabla.	
C3.	Većina ljudi u uzorku ima fakultetsko obrazovanje.	
C4.	U uzorku ima više od 200 žena.	
C5.	Većina muškaraca u uzorku ima manje od 30 godina.	
C6.	U uzorku ima više ljudi koji zarađuju više od 120 000 RSD, nego ljudi koji zarađuju manje od 60 000 RSD.	
C7.	Više od 2/3 (dve trećine) učesnika u istraživanju je nezaposleno.	
C8.	Bračno stanje je ovde operacionalizovano kao nominalna varijabla sa 4 kategorije.	
C9.	Medijana Prihoda ispitanika iz ovog uzorka je između 40 001 – 60 000 RSD.	
C10.	Apsolutna aritmetička sredina uzorka je iznad praga medijane.	

Vežba D. Distribucija, mere centralne tendencije, mere varijabilnosti.

PREPRINT

			Stavka broj 20					Ukupno
			1	2	3	4	5	
Pol	Muško	Frekvencija	50	20	22	12	9	113
		% unutar pola	44,2%	17,7%	19,5%	10,6%	8,0%	100%
		% unutar odgovora	41,0%	44,4%	53,7%	63,2%	69,2%	47.1%
		% od ukupnog br.	20,8%	8,3%	9,2%	5,0%	3,8%	47.1%
	Žensko	Frekvencija	72	25	19	7	4	127
		% unutar pola	56,7%	19,7%	15,0%	5,5%	3,1%	100%
		% unutar odgovora	59,0%	55,6%	46,3%	36,8%	30,8%	52.9%
		% od ukupnog br.	30,0%	10,4%	7,9%	2,9%	1,7%	52.9%
Ukupni uzorak		Frekvencija	122	45	41	19	13	240
		% unutar pola	50,8%	18,8%	17,1%	7,9%	5,4%	100%
		% unutar odgovora	100.0%	100,0%	100,0%	100%	100%	100%
		% od ukupnog br.	50.8%	18,8%	17,1%	7,9%	5,4%	100%

Tabela prikazuje distribuciju odgovora učesnika u istraživanju (ljudi) na stavku broj 20 upitnika. Na stavku se odgovara preko skale procene gde su mogući odgovori brojevi između 1 i 5, pri čemu 1 znači potpuno neslaganje sa tvrdnjom stavke, a 5 znači potpuno slaganje. Tabela predstavlja distribuciju celog uzorka i posebno na poduzorcima žena i muškaraca, kao i procentualnu distribuciju po polu unutar svakog odgovora. Podaci potiču iz istraživanja koje su sproveli autori.

D	**Tvrdnja:**	**Odgovor**
D1.	Ima više žena koje su odgovorile 1 na stavku broj 20 nego muškaraca koji su dali isti odgovor.	
D2.	Od svih žena iz uzorka, 44,2% je odgovorilo 1 na stavku broj 20.	
D3.	Od svih muškaraca iz uzorka, 17,7% je odgovorilo 2 na stavku broj 20.	
D4.	Od svih ljudi iz uzorka koji su odgovorili 1 na stavku broj 20, 59% čine žene.	
D5.	Od svih žena, manje od 5% je odgovorilo 5 na stavku broj 20.	
D6.	U uzorku ima više žena nego muškaraca.	
D7.	Mod stavke broj 20 na celom uzorku je odgovor broj 3.	
D8.	U celom uzorku ima manje od 200 ljudi.	
D9.	Broj žena koje su odgovorile 4 na stavku broj 20 je veći od broja muškaraca koji su dali isti odgovor na tu stavku.	
D10.	Mod (modalna vrednost) varijable pol na ovom uzorku je žensko.	

Vežba E. Distribucija, mere centralne tendencije, mere varijabilnosti.

Prosečna frekvencija mikronukleusa i indeksi deobe ćelijskog jedra u ispitivanim starosnim grupama. Ukupna veličina uzorka = 133.

	Starosne grupe u godinama	**MN (AS±SD)**	**NDI (AS±SD)**
1	**0 (n=29)**	0,56 ± 0,71	1,52 ± 0,30
2	**21-40 (n=30)**	0,82 ± 0,78	1,72 ± 0,26
3	**41-60 (n=29)**	1,26 ± 1,48	1,68 ± 0,35
4	**61-80 (n=26)**	5,48 ± 3,65	1,55 ± 0,24
5	**81-92 (n=19)**	4,48 ± 2,69	1,62 ± 0,34

Legenda: n – broj ispitanika u grupi. MN – frekvencija mikronukleusa na 1000 dvojedarnih ćelija. NDI – indeks deobe jedra. Predstavljene su aritmetičke sredine i standardne devijacije za svaku od pet starosnih grupa.

Tabela je napravljena na osnovu podataka iz istraživanja autora koje je objavljeno u Hedrih et al. (2018)

PREPRINT

E	Tvrdnja:	Odgovor
E1.	Modalni prosek je veći od 31 u svim grupama.	
E2.	Frekvencija mikronukleusa (varijabla MN) je najveća u starosnoj grupi 61-80 godina.	
E3.	Najveća varijabilnost frekvencije mikronukleusa (varijabla MN) je u grupi novorođenčadi (starost 0).	
E4.	Starosna grupa sa najvećom aritmetičkom sredinom na varijabli MN ima u isto vreme i najveću aritmetičku sredinu na varijabli NDI.	
E5.	Frekvencija mikronukleusa (varijabla MN) je na nominalnom nivou merenja.	
E6.	Varijansa varijable MN za grupu novorođenčadi (starost 0) je veća od 1.	
E7.	Nema ni jedne osobe u starosnoj grupi 61-80 godina u uzorku sa vrednošću MN ispod 1.	
E8.	Sva novorođenčad (starost 0) u uzorku imaju vrednosti NDI-a ispod 1,5.	
E9.	Na varijabli MN, dve starosne grupe koje su ispod 40 godina imaju manju varijabilnost nego grupe od 41 godina i više.	
E10.	U uzorku ima više od 200 učesnika/ispitanika.	

Pogledajmo sada odgovore:

C1 - tačno. Ovi brojevi su zaista procenti, kao što se može videti iz znaka % u imenu kolone.

C2 – tačno. Možemo da vidimo da su prikazane samo dve kategorije pola, što znači da je pol operacionalizovan kao binarna varijabla.

C3 – netačno. Možemo videti da 46 osoba ima fakultetsko obrazovanje, dok 200 ljudi ili 76,6% ima srednju školu.

C4 – netačno. Možemo videti da u uzorku ima 89 žena, što je manje od 200.

C5 – nepoznato. Iako imamo podatke i o broju muškaraca i o broju osoba ispod 30 godina, ne možemo iz tih podataka zaključiti koliko tačno ima muškaraca sa manje od 30 godina.

C6 – tačno. Možemo videti da je 36,4% ljudi iz uzorka izjavilo da ima prihode veće od 120 000 RSD, dok dve kategorije prihoda ispod 60 000 RSD sadrže 18,4% i 9,6% uzorka. Ove dve kategorije zajedno čine 28,0% uzorka, što je manje od 36,4%.

C7 – netačno. Možemo videti da je broj učesnika u istraživanju koji su naveli kao zanimanje „zaposlen“ 70,5%, što je više od dve trećine uzorka. Ako smatramo da su ljudi koji su kao zanimanje naveli da su zaposleni, zaista zaposleni, onda čak i ako bi svi ostali učesnici u istraživanju bili nezaposleni, to bi i dalje značilo da je manje od jedne trećine uzorka nezaposleno.

PREPRINT

C8 – netačno. Iz tabele možemo videti da postoje samo dve kategorije bračnog stanja – u braku i nije u braku.

C9 – netačno. Da bi odgovorili na ovo pitanje posmatramo medijanu tj. 50i percentil. Ako bi sabrali procente po kategorijama na varijabli Prihod, videli bi da je suma procenata ispitanika u kategorijama do 60 000 RSD manja od 50%. Čak i da dodamo kategoriju 60 001- 90 000 RSD, zbir procenata i dalje ne bi bio 50%. Da prebacimo 50%, potrebno je da dodamo i kategoriju 90 001 – 120 000 RSD, što znači da se medijana nalazi u toj kategoriji.

C10 – besmisleno. Ne postoji standardni statistički pojam koji se zove „apsolutna aritmetička sredina“. Isto važi i za „medijanu praga“. Ako bi autor rada hteo da uvede ove pojmove, morao bi prvo da ih definiše.

D1 – tačno. Možemo videti iz tabele da su 72 žene odgovorile sa 1, naspram 50 muškaraca.

D2 – netačno. Možemo videti da je % žena (% unutar pola) koje su odgovorile 1 jednak 56,7%. Podatak 44,2% odnosi se na procenat muškaraca.

D3 – tačno. Možemo videti da je % unutar pola za muškarce na odgovoru 2 zaista 17,7%.

D4 – tačno. Možemo videti da je % unutar odgovora za žene 59%. To znači da od svih ljudi koji su odgovorili sa 1, 59% čine žene (dok su 41% muškarci).

D5 – tačno. Možemo videti da su samo 4 žene odgovorile brojem 5 i da je to 3,1% svih žena u uzorku.

D6 – tačno. Možemo videti iz tabele da su ovi brojevi 127 žena i 113 muškaraca.

D7- netačno. Možemo videti da je mod tj. najčešći odgovor na stavku 20 u ukupnom uzorku odgovor 1, a ne 3. 122 osobe ili 50,8% ukupnog uzorka je odgovorilo sa 1.

D8 – netačno. Ne, ukupna veličina uzorka je 240, kao što možemo videti iz polja gde je navedena ukupna veličina uzorka.

D9 – netačno. Možemo videti da je 7 žena dalo odgovor 4, a 12 muškaraca. Prema tome, broj muškaraca koji je dao ovaj odgovor je veći, što čini ovu tvrdnju netačnom.

PREPRINT

D10 – tačno. Možemo videti da pol u ovoj tabeli ima dve vrednosti i da je frekvencija žena veća, što ih čini modom – žene čine 52,9% uzorka, dok muškarci čine 47,1% uzorka, što znači da je žensko modalna vrednost ovog uzorka.

E1 – besmisleno. Ne postoji nikakav „modalni prosek“.

E2 – tačno. Da, možemo videti da je prosečna frekvencija mikronukleusa u ovoj grupi 5,48, što je više od bilo koje druge prosečne vrednosti u ovoj koloni (uporedite ovo sa svim ostalim prvim brojevima u MN koloni).

E3 – netačno. Možemo videti da je mera varijabilnosti – standardna devijacija u ovom slučaju, za grupu novorođenčadi (starost 0) jednaka 0,71 i ovo nije najveća vrednost u MN koloni (tj. za MN varijablu). U stvari, to je najniža varijabilnost od svih grupa.

E4 – netačno. Možemo videti da je najveća varijablnost na varijabli MN u grupi 4 tj. u grupi ljudi između 61 i 80 godina starosti, ali da je najveća standardna devijacija na varijabli NDI u grupi između 41 i 60 godina starosti.

E5 – netačno. Vidimo da su autori računali aritmetičku sredinu i standardnu devijaciju za tu varijablu, što znači da ona mora da je bar na intervalnom nivou merenja, jer je to minimalni nivo merenja za ove statistike. Takođe, varijabla se zove frekvencija mikronukleusa, što znači da se dobija brojanjem mikronukleusa, a mere dobijene prebrojavanjem su na racio nivou merenja.

E6 – netačno. Varijansa je standardna devijacija na kvadrat (standardna devijacija pomnožena samom sobom). Kako možemo videti da je standardna devijacija ove grupe 0,71, tj. manja je od 1, varijansa ove grupe će takođe biti manja od 1, zato što su kvadrati brojeva manjih od 1 manji od broja čiji su to kvadrati, što takođe znači i da su manji od 1.

E7 – nepoznato. Iako vidimo da je aritmetička sredina ove grupe 5,48, što je mera centralne tendencije, nemamo prikaz cele distribucije uzorka. To znači da je moguće da ima nekog u grupi sa vrednošću manjom od 1 (tj. sa 0 mikronukleusa), ali to ne znamo. Takođe, imajući u vidu da je standardna devijacija ove grupe 3,65, vrednost 1 je svakako moguća, a i 0.

E8 – netačno. Možemo videti da je aritmetička sredina grupe novorođenčadi na varijabli NDI 1,52. Prema tome, da bi ova tvrdnja bila tačna, bilo bi potrebno da svi učesnici u istraživanju iz

PREPRINT

ove grupe imaju vrednosti ispod aritmetičke sredine grupe, a to, znajući kako se računa aritmetička sredina, možemo zaključiti da nije moguće.

E9 – tačno. Možemo videti da su standardne devijacije prve dve grupe obe ispod 1, dok su standardne devijacije ostale tri grupe veće. Kako je standardna devijacija mera varijabilnosti, sledi da je tvrdnja o varijabilnostima grupa tačna.

E10 – netačno. Možemo pročitati iz gornjeg dela tabele da je veličina uzorka 133, što je manje od 200.

PREPRINT

Poglavlje 4. Distribucije

Apstrakt. Ovo poglavlje počinje određivanjem razlike između teorijskih i empirijskih distribucija, da bi potom bila predstavljena normalna, Puasonova, binomna i uniformna distribucija. Diskutuje se o tome kada se može očekivati da se koja od ovih distribucija dobije i za koje procese se smatra da stoje u osnovu distribucija podataka koje imaju pomenute oblike. Slede delovi koji predstavljaju pitanja horizontalnih i vertikalnih odstupanja od normalne distribucije, te predstavljaju i objašnjavaju pojmove skjunesa i kurtozisa, kao i neke tipične situacije kada se pozitivno i negativno asimetrične, kao i one koje vertikalno odstupaju od oblika normalne distribucije – „šiljate“ i „spljoštene“ distribucije mogu očekivati. Ovo poglavlje posebno ističe različita specifična svojstva kurtozisa i podvlači razliku između „šiljatih“ i „spljoštenih“ distribucija tj. distribucija koje po vertikali odstupaju od normalne s jedne strane i pojmova leptokurtične, mezokurtične i platikurtične distribucije s druge strane. Pojmovi bimodalne i polimodalne distribucije su takođe predstavljeni. Poslednji deo ovog poglavlja posvećen je postupcima standardizacije i ipsatizacije, predstavljaju z skorova i z skala, objašnjavanju prevođenja na i sa z skale i svojstava z skale.

Ključne reči: teorijska distribucija, skjunes, kurtozis, standardizacija, ipsatizacija

4.1. Teorijske i empirijske distribucije

Kako je navedeno u prethodnom poglavlju, distribucija je spisak ili funkcija koja opisuje sve postojeće (ili čak sve moguće) vrednosti varijable i to koliko često se javljaju. Obično se pravi tako što se navedu sve vrednosti varijable čiju distribuciju predstavljamo i frekvencije s kojima se javljaju. Alternativno, distribucija se može predstaviti i predstavljanjem intervala vrednosti za kontinualne varijable ili širih kategorija za varijable sa puno različitih mogućih vrednosti (kao što je slučaj sa kontinualnim varijablama), a potom navođenjem frekvencija entiteta u svakom od ovih intervala ili širih kategorija. Distribucija se može predstaviti i funkcijom gustine verovatnoće.

Pored davanja prostog činjeničnog opisa vrednosti entiteta iz uzorka na posmatranoj varijabli, **distribucije se u statistici koriste i za izvođenje različitih zaključaka o uslovima vezanim za svojstva varijable, uzorka i merenja, a na osnovu oblika distribucije**. Ovo se radi tako što se

PREPRINT

distribucija koja je dobijena iz empirijski prikupljenih podataka (dakle – na uzorku) **poredi sa određenih poznatim oblicima distribucija,** distribucija za koje je poznato da nastaju kao rezultat nekih poznatih tipova procesa. **Ako oblik distribucije uzorka liči na oblik sa kojim se poredi, onda se može zaključiti,** pod uslovom da svojstva uzorka i merenja to dozvoljavaju, **da su i na uzorku koji posmatramo dejstvovali procesi za koje se zna da dovode do takvih oblika distribucije.** A možemo, u suprotnom, zaključiti i da naša distribucija verovatno nije nastala kao rezultat procesa koji stvaraju distribucije datog oblika. Iz ovog razloga, važno je uvesti razliku između empirijskih i teorijskih distribucija.

- **Empirijska distribucija je realna distribucija dobijena iz realnih podataka** tj. realno načinjenih opservacija ili merenja. Može da ima bilo koji oblik, u zavisnosti od toga kakvi su podaci i kakva su svojstva uzorka na kom je dobijena.
- **Teorijske distribucije imaju određeni oblik i ime i zasnovane su na teoriji ili drugom sličnom naučnom instrumentu koji opisuje situaciju u kojoj je možemo očekivati**. Ove teorije pružaju uvid u to koji tip distribucije očekujemo sa kakvim tipom podataka i one su onda osnova za predviđanje verovatnoća budućih događaja.

4.2. Normalna distribucija

Najšire korišćena i najpoznatija teorijska distribucija je normalna distribucija koja se takođe zove i **Gausova distribucija ili Gaus-Laplasova distribucija**. To je **kontinualna distribucija** koja se **obično predstavlja preko funkcije gustine verovatnoće koja se zove normalna kriva, Gausova kriva ili zvonasta kriva**. Normalna distribucija je **simetrična,** pri čemu su **entiteti najgušće skoncentrisani oko centra distribucije, a frekvencije opadaju sa udaljavanjem od tog centra u oba smera (i na gore i na dole) podjednako.** Kriva normalne distribucije se **asimptotski približava frekvenciji/verovatnoći od nula,** što znači da sa udaljenošću vrednosti varijable od centra distribucije, verovatnoća da se entitet nađe na datoj poziciji na distribuciji se sve više i više smanjuje, ali nikada ne dostiže nulu, samo postaje sve manja i manja u beskonačnost.

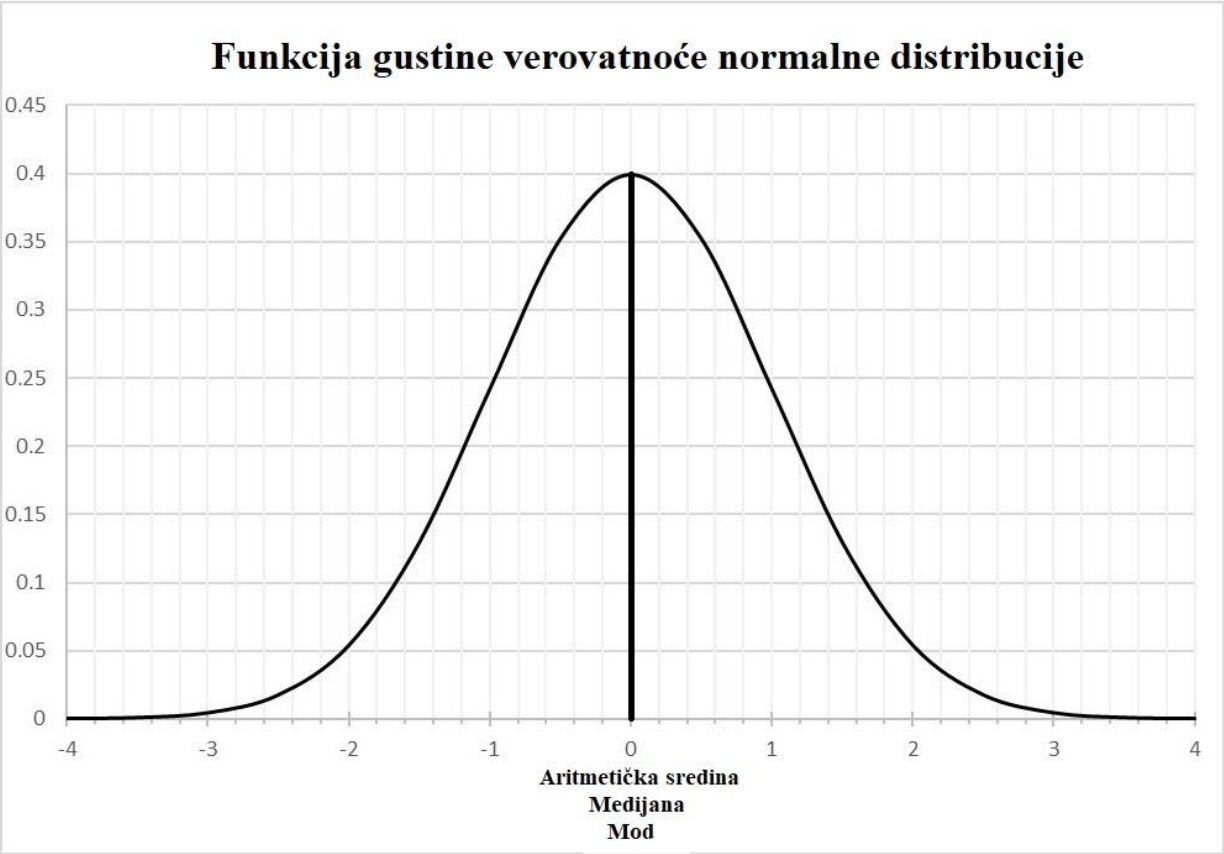
Normalna distribucija je očekivana distribucija podataka u velikom rasponu situacija koje se mogu opisati kao **situacije u kojima na populacione vrednosti utiče jedan veoma uticajni faktor koji je konstanta za sve entitete i veći broj drugih faktora koji su pojedinačno mnogo manje**

PREPRINT

uticajni i međusobno su nezavisni, te se mogu smatrati slučajnim faktorima koji su takvi da im se vrednosti razlikuju od entiteta do entiteta i koji su onda uzroci razlika među entitetima. Situacije poput ove opisane uključuju pre svega **individualne razlike između ljudi i drugih organizama na različitim varijablama**, postignuće na testovima obrazovnog postignuća (u situacijama kada su ljudi sličnih sposobnosti prošli kroz isti obrazovni program), postignuće na psihološkim testovima grupa sličnih osoba, rezultati ponovljenih merenja istog fenomena instrumentima čija preciznost je manja od savršene tj. koji proizvode nenulte greške merenja, različite pokazatelje promena na tržištu isl. Pojam normalne distribucije je poznat i u upotrebi vekovima i nije poznato ko je prvi osmislio ovaj pojam, iako se prvi doprinos razvoju koncepta normalne distribucije često pripisuje francuskom matematičaru iz 18. veka Abrahamu de Moavru i njegovoj knjizi o verovatnoći koja je na engleskom objavljena po naslovom "The Doctrine of Chances: Or, A Method of Calculating the Probabilities of Events and Play" (de Moivre, 1756)

Slika 4.1. Prikaz normalne distribucije preko funkcije gustine verovatnoće. Tačka označena sa 0 je aritmetička sredina, medijana i mod, jer sve ove tri mere centralne tendencije imaju istu vrednost na idealnoj normalnoj distribuciji. Vertikalna osa predstavlja verovatnoće, dok horizontalna osa predstavlja vrednosti varijable.

PREPRINT



Kako se pojam normalne distribucije upotrebljava u praksi? Kao što je napred navedeno, normalna distribucija je očekivana distribucija u situacijama kada su vrednosti varijable rezultat jednog veoma uticajnog faktora koji je konstanta i velikog broja nezavisnih slučajnih faktora koji su varijable. Najčešće, to je **očekivana situacija onda kada imamo posla sa individualnim razlikama između entiteta u populaciji u odnosu na određenu varijablu.** Pretpostavku o tome kako normalna distribucija nastaje možemo iskoristiti da testiramo li ona važi za određenu populaciju koju ispitujemo na određenoj varijabli. Na primer, možemo ispitivati da li distribucija individualnih razlika u visini stanovnika određene oblasti nalikuje obliku normalne distribucije tj. da li su njihove individualne razlike samo posledica različitih slučajnih faktora malog pojedinačnog uticaja ili su u igri i neki uticajniji faktori. Mogli bismo tako da uzmemo uzorak ljudi iz te populacije, da im izmerimo visine i da onda tako dobijenu empirijsku distribuciju uporedimo sa teorijskom normalnom distribucijom. Ako bi se pokazalo da se te dve distribucije poklapaju, to bi značilo da posmatramo jednu homogenu populaciju kada je u pitanju visina. Ako bi se pokazalo

da se ove dve distribucije (empirijska distribucija uzorka i teorijska normalna distribucija) ne poklapaju, mogli bismo da ispitamo te razlike kako bi saznali šta se desilo. Ono što bi, u slučaju poput ovog, verovatno otkrili je to da je biološki pol važan faktor koji određuje visinu i da bi, ako bi hteli da dobijemo normalnu distribuciju, morali da istražujemo ljude istog pola posebno. Razlog za ovo je to što je telesna visina polno dimorfna[6] osobina kod ljudi. Međutim, u zavisnosti od toga koju smo populaciju proučavali, ako unutar populacije postoje grupe koje su na primer hronično gladovale tokom detinjstva ili na koje je sistematski uticao neki drugi faktor od značaja za visinu, to bi verovatno bilo takođe vidljivo kroz oblik distribucije, pod pretpostavkom, naravno, da je grupa na koju je dati faktor uticao dovoljno velika. Na sličan način, ako test koji ispituje usvojenost gradiva iz određenog programa učenja na grupi učenika, možemo očekivati da rezultati budu normalno distribuirani, ako su svi učenici prošli taj isti program (i otprilike se jednako spremali za test, bili jednako motivisani i svi drugi kritični faktori bili slični kod svih). Ako bi se pokazalo da rezultati nisu normalni to bi mogli da ukazuje na postojanje nekih sistematskih faktora odgovornih za odstupanje od normalne distribucije. Na primer, mogli bi da otkrijemo da neki učenici nisu uopšte pratili taj program, nego su samo došli da polažu test, a nekada bi rezultati koji odstupaju od normalne distribucije mogli da ukazuju i na probleme sa valjanošću samog testa. Na sličan način, ako bi gledali nekog sportistu koji trenira za trku i merili vreme koje mu/joj je potrebno da pretrči stazu, takođe bi mogli da očekujemo da distribucija vremena koje mu je bilo potrebno prati oblik normalne distribucije. Ako bi otkrili da ne prati, ovo bi moglo da ukazuje na neke sistematske faktore – da nije postajao vremenom sve umorniji kako je trening odmicao, pa zato bivao sve sporiji? Da mu/joj se nije promenila motivacija za trku? Da li je možda bilo nekih prepreka na stazi tokom određenih prelazaka? Možda je tokom nekih prelazaka staze bio/bila povređen/a ili je imao bolove itd.

*****

Ako uzmemo **dve međusobno nezavisne, normalno distribuirane varijable i izračunamo njihov količnik tj. podelimo vrednosti dve varijable na svakom entitetu jednu sa drugom**, **ovako dobijena varijabla imaće distribuciju koja se zove Košijeva distribucija (eng. Couchy)**, koja je ime dobila prema francuskom matematičaru iz 18.-19. veka Augustinu-Luisu Košiju

[6] Reč dimorfizam označava to da se određeno svojstvo javlja u dva oblika kod iste vrste. Polni dimorfizam se odnosi na to da se jedinke različitog pola iste vrste razlikuju na određeni način u osobini o kojoj je reč, odnosno da se osobina javlja u jednom obliku kod jednog pola, a u drugom kod drugog pola.

PREPRINT

(Augustin-Louis Couchy). Košijeva distribucija se nekada takođe naziva i **Lorencovom distribucijom** (eng. Lorenz distribution), **Koši-Lorencovom distribucijom** (eng. Cauchy-Lorenz distribution) ili **Breit-Vignerovom distribucijom** (eng. Breit-Wigner distribution).

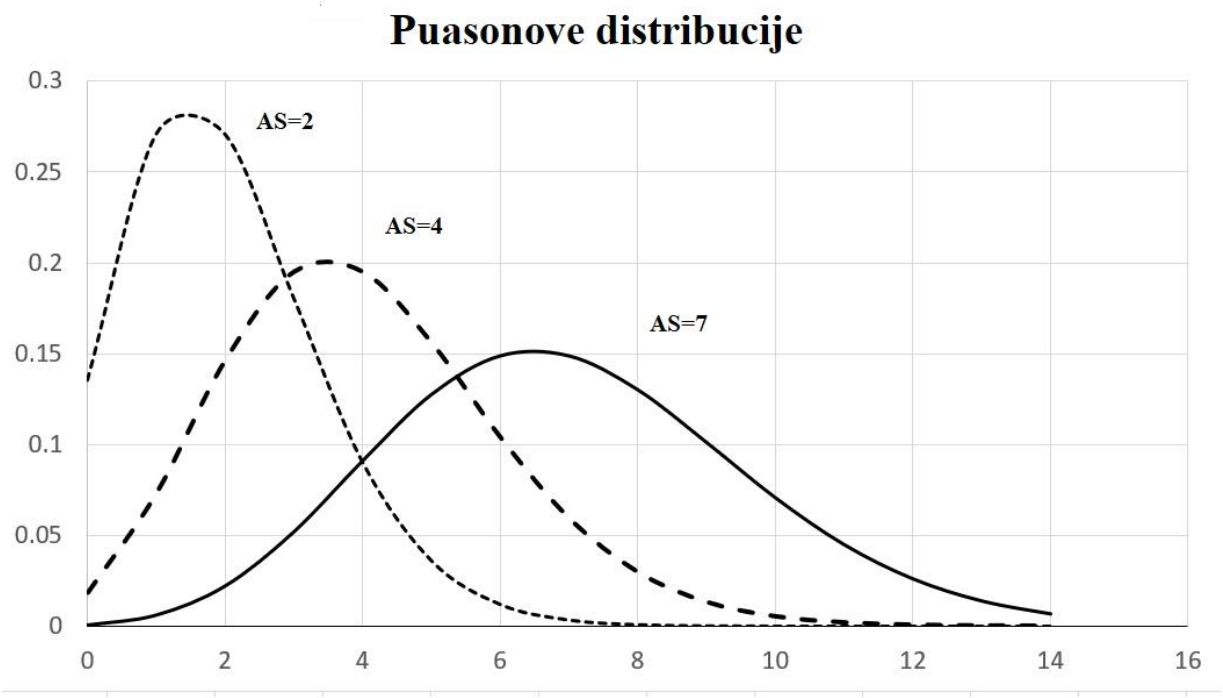
4.3. Puasonova distribucija

Puasonova distribucija (eng. Poisson) opisuje **koliko se puta neki slučajni događaj desio unutar jedinice merenja – perioda vremena ili jedinice prostora.** Ovaj slučajni događaj je takav da ima **fiksnu prosečnu stopu javljanja u svakoj jedinici merenja**. Prema ovoj koncepciji, **verovatnoća javljanja takvog događaja je stalna u svakoj jedinici merenja, ne menja se, a javljanja događaja su nezavisna jedna od drugih**. Jedinica merenja može biti određeni **vremenski period, jedinica prostora, određena udaljenost ili površina, ali može biti i pojedinačna osoba**. U društvenim naukama, Puasonova distribucija se često koristi za opisivanje toga koliko se puta svakom članu populacije desio određeni događaj tokom određenog vremenskog perioda. Pretpostavka je da je verovatnoća javljanja ovog događaja fiksna unutar jedinice vremena za sve članove populacije.

Puasonova distribucija je **diskretna distribucija** tj. njene vrednosti mogu biti samo prirodni brojevi, jer predstavlja broj puta koliko se neki događaj desio, a to može biti samo prirodan broj, dakle broj koji je ceo i nije manji od 0. Ova distribucija je ime dobila po francuskom matematičaru Simonu Denisu Puasonu (Simon Denis Poisson).

Slika 4.2. Oblici Puasonove distribucije zavisno od aritmetičke sredine distribucije. Kada je aritmetička sredina dovoljno visoka tj. kada je centar distribucije dovoljno daleko od 0, oblik Puasonove distribucije liči na oblik normalne distribucije. Puasonova distribucija je diskretna distribucija i njene vrednosti su brojevi puta koliko se događaj desio po entitetu, dakle na racio nivou merenja i mogu biti samo prirodni brojevi. Na vertikalnoj osi su verovatnoće, dok su na horizontalnoj osi brojevi puta koliko se događaj desio (po entitetu). Kriva distribucije predstavlja verovatnoće da će slučajno odabrani entitet iz populacije u kojoj važe parametri distribucije imati određenu vrednost na horizontalnoj osi.

PREPRINT



Kako se Puasonova distribucija koristi u praksi? **Puasonova distribucija je korisna u različitim situacijama gde je cilj da se odredi da li su određeni događaji samo posledica slučaja** (tj. da li su u pitanju slučajni događaji) **ili možda do njih dovode neki sistematski faktori**. Na primer, ako želimo da ispitamo da li se greške u nekom industrijskom proizvodnom procesu dešavaju slučajno ili zbog toga što, na primer, neki radnici ne poštuju proceduru, ovo ispitivanje možemo početi sa očekivanjem da do grešaka dolazi prosto zbog toga što je proizvodni proces takav da neizbežno dovodi do određenog procenta slučajnih grešaka. U tom slučaju, prosečan broj grešaka tokom fiksnog intervala vremena će biti sličan za sve radnike koji rade istu vrstu aktivnosti. Ovo, međutim, ne znači da će nakon određenog intervala vremena svi radnici imati isti broj grešaka, jer neće. To znači da će distribucija grešaka po radniku imati oblik Puasonove distribucije, pri čemu će neki radnici imati više grešaka, neki manje, ali će prosek grešaka biti jednak osnovnoj očekivanoj stopi grešaka za taj postupak. Kad izračunamo taj broj grešaka po radniku, onda možemo pravu, empirijski dobijenu distribuciju broja grešaka po radniku da uporedimo sa Puasonovom distribucijom sa istom prosečnom vrednošću javljanja grešaka i videti da li se te dve distribucije poklapaju. Ako su te dve distribucije slične, mogli bi da zaključimo da su greške zaista posledica slučaja. Tada, ako bi hteli da smanjimo stopu javljanja grešaka bilo bi potrebno da redizajniramo proizvodni proces koji smo proučavali tako da rezultira manjim brojem grešaka.

Međutim, ako bi se ispostavilo da se ove dve distribucije ne poklapaju, već da postoje upadljive razlike između njih, bilo bi potrebno da proučimo prirodu tih razlika, tj. kako one izgledaju. Onda bismo, na primer, mogli da otkrijemo da postoji grupa radnika koja pravi mnogo manje grešaka nego što bi se očekivalo na osnovu Puasonove distribucije. Ovo bi onda moglo da znači da ti radnici u praksi rade po nekom drugom postupku koji je različit od standardnog. Onda bismo mogli da proučimo šta je to što oni rade i da to onda primenimo svuda. Alternativno, bilo bi moguće da otkrijemo da postoje radnici koji prave mnogo više grešaka nego što bi se očekivalo na osnovu Puasonove distribucije. To bi onda značilo da verovatno postoji neki neželjeni sistematski faktor koji dejstvuje u toj grupi, koji bi onda morali da proučimo. Tu bi onda očekivali da otkrijemo stvari poput nedovoljne ili neadekvatne veštine radnika ili nedovoljne ili neadekvatne obuke, nedostatak radne discipline, nedostatak sposobnosti ili motivacije da se prate propisane procedure ili neki drugi neželjeni faktor koji je specifičan za datu grupu. Na ovaj način, **Puasonova distribucija nam omogućava da odredimo koje bi razlike u učestalosti javljanja događaja koji ispitujemo** (taj događaj su, u malopređašnjem primeru, greške u proizvodnom procesu) **mogle da budu samo posledica slučaja i stoga očekivane, a gde bi bilo smisleno sumnjati da su u igri neki sistematski faktori.**

Na sličan način, poređenje empirijske distibucije sa Puasonovom distribucijom nam može pomoći da odredimo stvari poput, na primer, toga da li postoje sati u danu u radu kol centra kada je učestalost poziva viša ili niža od uobičajene, da procenimo da li do većeg ili manjeg broja saobraćajnih nezgoda na određenim rutama ili lokacijama dolazi čisto slučajno ili postoje neke rute koje su bezbednije ili opasnije ili da su neki vozači veštiji ili manje vešti vozači, da procenimo kada je broj pacijenata koji stiže na neku kliniku unutar normalnih slučajnih varijacija u učestalosti, a kada bi promena učestalosti mogla da bude posledica nekih posebnih faktora itd.

4.4. Binomna distribucija

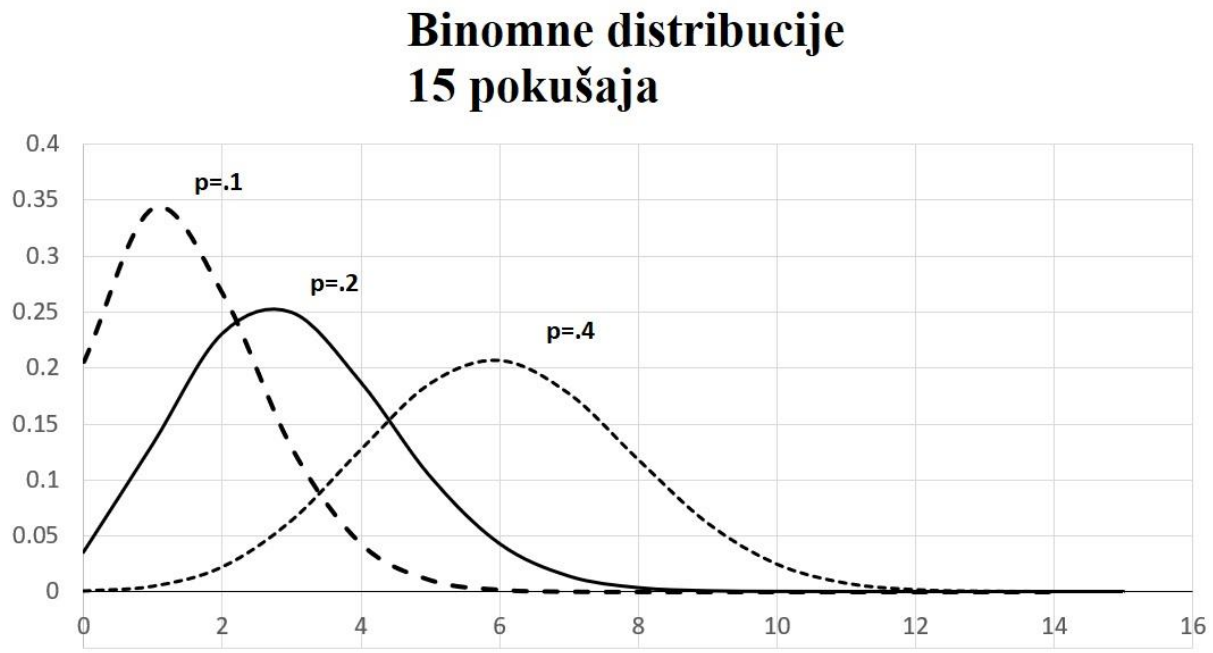
Binomna distribucija opisuje koliko se puta slučajni događaj desio po entitetu u populaciji u okviru nekog unapred određenog broja pokušaja kroz koje je prošao svaki od posmatranih entiteta. Drugim rečima, to je diskretna distribucija verovatnoća broja događaja koji su se desili po entitetu u sklopu nekog unapred određenog broja pokušaja. **Događaj** koji je u pitanju je takav da **ima određenu fiksnu stopu/verovatnoću javljanja u svakom pokušaju**, a za svaki pokušaj

beležimo da li se u njemu događaj desio ili nije. **Pokušaj o kom se ovde radi je određena aktivnost koja može da za rezultat ima javljanje pomenutog slučajnog događaja ili da nema to za rezultat**. **Ishod ovih pokušaja je uvek binaran** (događaj se desio ili se nije desio), a **broj pokušaja je isti za svaki entitet u uzorku.** Ovi pokušaji nazivaju se **Bernulijevim pokušajima**. **Poseban slučaj binomne distribucije je slučaj kada je broj pokušaja (po entitetu) jednak 1 i takva distribucija se zove Bernulijeva distribucija.**

Može se primetiti da su uslovi pod kojima se očekuje binomna distribucija slični onima pod kojima se očekuje Puasonova distribucija pri čemu je glavna razlika u tome da je Puasonova distribucija zasnovana na jedinici mere – vremenskom intervalu, prostornoj jedinici itd. i tome da se događaj može desiti više puta unutar iste jedinice mere, dok je binomna distribucija zasnovana na Bernulijevim pokušajima, a sam kritični događaj se može desiti jednom u jednom pokušaju. Bernulijevi pokušaji, kao i Bernulijeva distribucija, koja je poseban slučaj binomne distribucije ime su dobili prema švajcarskom matematičaru iz 17. veka Jakobu Bernuliju (Jacom Bernoulli).

Slika 4.3. Primeri oblika binomne distribucije sa 15 Bernulijevih pokušaja i sa različitim verovatnoćama javljanja događanja tokom pokušaja. Kada je centar binomne distribucije dovoljno udaljen od 0, oblik binomne distribucije lični na oblik normalne distribucije. Brojevi iznad krive pokazuju verovatnoću da se događaj za distribuciju koja je opisana datom krivom. Binomna distribucija je diskretna distribucije i njene vrednosti su prirodni brojevi koji ne mogu biti veći od ukupnog broja pokušaja tj. maksimum binomne distribucije je situacija kada se događaj desio u svakom pokušaju. Na vertikalnoj osi su verovatnoće, dok horizontalna osa predstavlja broj puta koliko se događaj desio (po entitetu). Kriva distribucije predstavlja verovatnoće da slučajno odabrani entitet iz populacije u kojoj važe parametri date distribucije ima određenu vrednost na horizontalnoj osi.

PREPRINT



Praktična upotreba binomne distribucije je **slična načinu na koji se može koristiti Puasonova distrubicija s tom razlikom da situacije moraju da uključuju procese koji se mogu adekvatno predstaviti preko Bernulijevih pokušaja** tj. procesa sa binarnim ishodima, a ti binarni ishodi treba da budu ono što interesuje istraživača. Na primer, železnička kompanija može da odluči da poredi distribuciju broja zakasnelih dolazaka u stanicu po mašinovođi nakon određenog broja vožnji duž standardne rute sa oblikom binomne distribucije sa ciljem da zaključe da li su kašnjenja voza posledica slučaja i toga što prosto postoje slučajni faktori koji jednako utiču na sve mašinovođe i dovode do toga da zakasne ili možda postoje sistematski faktori koji deluju na određene vozače (npr. nedovoljna veština upravljanja vozom, loše organizovanje vremena) ili na vozove (npr. tehnički nedostaci). Mogli bi takođe, na primer, da koristimo binomnu distribuciju kao referentni okvir da bismo saznali da li je trenutno nivo uspeha (da li je uspeo da preskoči cilj ili ne) sportiste/sportistkinje koji/a se bavi skokom s motkom u nivou njegovog/njenog redovnog uspeha ili mu/joj se uspeh popravlja ili opada kao posledica nekih sistematskih faktora itd.

4.5. Uniformna distribucija

Uniformna distribucija opisuje situacije u kojima sve vrednosti varijable imaju istu frekvenciju tj. situaciju kada su verovatnoće javljanja svih kategorija/vrednosti varijable jednake. Ona se najčešće

koristi onda kada su podaci na nominalnom nivou merenja, ali postoje i posebne situacije kada se smisleno može koristiti na podacima na višim nivoima merenja. Uniformna distribucija zahteva da podaci budu diskretni, a najbolje je ako su grupisani u manji broj širih kategorija.

Uniformna distribucija se najčešće koristi u situacijama kada je potrebno ustanoviti da li uzorak potiče iz populacije u kojoj sve vrednosti razmatrane varijable imaju jednake verovatnoće javljanja tj. u kojoj su im frekvencije jednake. Takve situacije uključuju testiranje različitih generatora pseudoslučajnih brojeva, poput onih koji se koriste u igrama na sreću, ali i različite istraživačke situacije u kojima je potrebno ustanoviti da li su svi posmatrani ishodi jednako verovatni. Na primer, možemo testirati da li kocka za igru (jedna od onih sa stranicama koje su označene različitim brojevima, mogu se posmatrati kao jednostavan generator pseudoslučajnih brojeva) radi kako treba tako što ćemo je baciti nekoliko stotina puta i onda uporediti dobijenu distribuciju ishoda (to koji broj je „pao“) sa uniformnom distribucijom. Ako su te dve distribucije dovoljno slične, to znači da kocka za igru funkcioniše kako treba. Ako te dve distribucije nisu dovoljno slične, možemo zaključiti da je kockica neispravna, tj. da češće pada na neke strane nego na neke druge. Kocka za igru koja radi kako treba, trebalo bi da pada na svaku stranu podjednako često. Na sličan način, možemo da se pitamo da li kupci dva ili više proizvoda kupuju podjednako često ili pak kupci više vole/više kupuju neki od tih proizvoda nego neki drugi. Ovo možemo da proverimo tako što ćemo uporediti empirijsku distribuciju broja kupljenih proizvoda tokom određenog vremena, sa uniformnom distribucijom. Ako su te dve distribucije dovoljno slične, možemo zaključiti da kupci ne pokazuju sklonosti da kupuju više neke proizvode od drugih (iz grupe posmatranih proizvoda), odnosno da posmatrane proizvode kupuju jednako. Ako ove dve distribucije nisu dovoljno slične, onda možemo zaključiti da kupci imaju veću sklonost ka kupovini jednih proizvoda nego drugih. Onda bi mogli da pregledamo empirijsku distribuciju i iz nje vidimo koji su to proizvodi koje kupci kupuju češće, a koje kupuju ređe.

4.6. Odstupanja od normalne distribucije

Normalna distribucija je teorijska distribucija koja se najčešće koristi u istraživanjima. To je podrazumevana očekivana distribucija individualnih razlika na čitavom nizu varijabli i, zbog toga,

PREPRINT

precizno poznavanje svojstava odstupanja empirijski dobijenih podataka od ove distribucije nam može pomoći u donošenju različitih zaključaka o podacima. Ovo je razlog zašto istraživači ne samo da procenjuju da li distribucija koju su dobili liči na normalnu ili ne, već i opisuju kakva su tačno dobijena odstupanja. Ranije smo rekli da je teorijska normalna distribucija simetrična i da njena kriva ima određeni oblik. Krajevi normalne distribucije se asimptotski približavaju nuli i, zbog toga što je distribucija simetrična, sa najvećom koncentracijom entiteta na sredini distribucije, njena aritmetička sredina, medijama i mod su jednaki (imaju jednake vrednosti) i nalaze se tačno na sredini distribucije. Distribucije empirijskih podataka mogu od ovog stanja odstupati po vertikali ili po horizontali, ali i na različite kompleksnije načine.

Horizontalna odstupanja od normalne distribucije imamo onda kada je **empirijska distribucija asimetrična** tj. kada, posmatrajući distribuciju, možemo uočiti da **entiteti teže da budu gušće koncentrisani na jednoj strani distribucije, a razređeniji na drugoj strani distribucije**. Kada se to grafički predstavi, dobijamo sliku distribucije na kojoj je **jedna strana kraća nego što bi bila na normalnoj distribuciji, dok je druga strana izdužena**. Ovi asimetrični oblici distribucije nazivaju se **pozitivnim ili negativnim zavisno od toga da li je izduženi kraj distribucije iznad centra ili ispod centra** tj. da li su razlike između vrednosti entiteta na izduženom kraju distribucije i aritmetičke sredine te distribucije pozitivne ili negativne. **Normalna distribucija koja nije asimetrična naziva se simetričnom distribucijom.** Prema tome, zavisno od toga koja vrsta asimetrije je u pitanju, distribucija može biti:

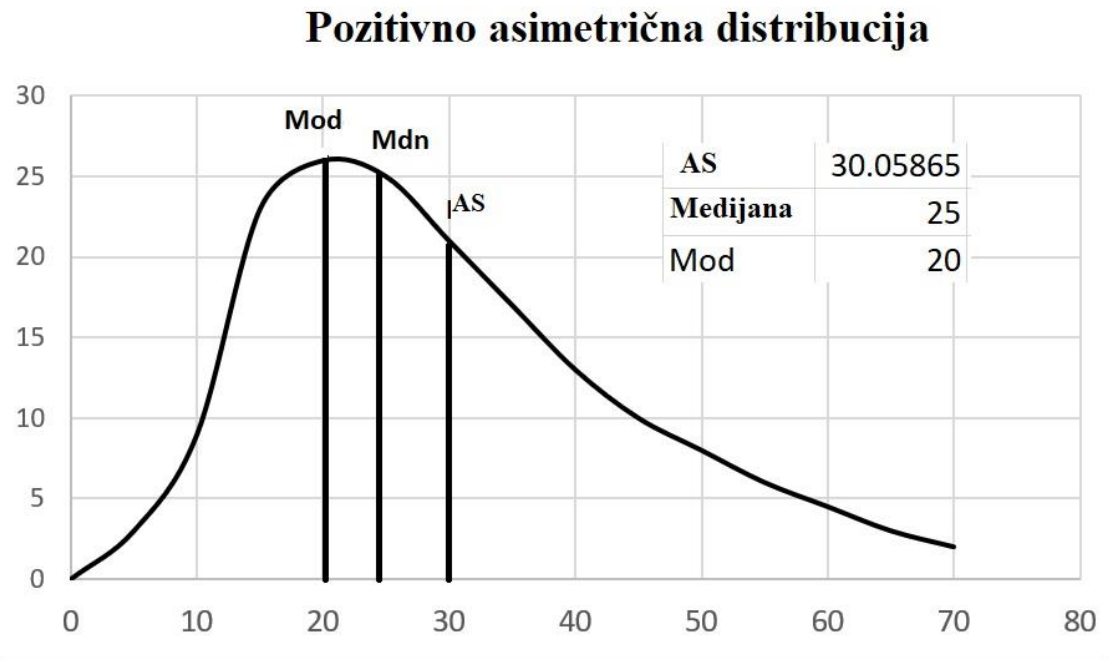
- **Pozitivno asimetrična distribucija je oblik distribucije na kojoj je izduženi kraj distribucije onaj gornji**, tj kraj na kome se nalaze više ili (na skali na kojoj je AS=0) pozitivne vrednosti varijable. **Entiteti su koncentrisani na donjem delu distribucije**, dok je **gornji deo distribucije**, deo čije su vrednosti više od vrednosti centra distribucije, **izdužen**. Na pozitivno asimetričnoj distribuciji, **mod je niži od medijane**, a **medijana je niža od aritmetičke sredine**. Pozitivno asimetrična distribucija obično ukazuje na prisustvo faktora koji sprečavaju entitete da dobiju više vrednosti. Koji tačno faktori su u pitanju, naravno, zavisi od oblasti nauke o kojoj se radi. Na primer, u psihološkom ili pedagoškom testiranju, **pozitivno asimetrična distribucija tipično ukazuje na to je da test težak** tj. da se test sastoji od stavki čija težina je iznad nivoa osobine koja se procenjuje kod osoba iz uzorka. Ako se ovakva distribucija dobije na testu znanja, to tipično ukazuje

PREPRINT

na to da većina ispitanika (iz uzorka) nema dovoljno znanja o temi koju test ispituje odnosno da je test pretežak za tipičan nivo znanja ispitanika. Ako je u pitanju kognitivni test ili test fizički sposobnosti ili set fizičkih vežbi, ovakva distribucija tipično ukazuje da su zadaci iz testa previše teški i iznad nivoa sposobnosti ili veština većine ispitanika. Ako je u pitanju test stavova ili neki druga vrsta konativnog testa, ovakva distribucija ukazuje da su stavovi ili preferencije ispitanika veoma različite od onih na koje ukazuju stavke testa. **Ako cilj testiranja nije da se fino razlikuju ispitanici u gornjem delu distribucije, oni sa najvišim vrednostima, pozitivno asimetrična distribucija po pravilu nije dobra stvar**, jer ona znači da će razlike u vrednostima na varijabli (npr. u skorovima na testu) između entiteta u centru distribucije biti smanjene. Kako se većina entiteta nalazi na centru distribucije, to znači da će **varijabilnost individualnih vrednosti biti smanjena** (u odnosu na normalnu distribuciju) **za većinu entiteta**. Imajući u vidu to da se stvari mere ili procenjuju zato da bi se saznale razlike u vrednostima između njih, ovakva situacija je direktno suprotna ovakvom cilju i zato nije dobra.

Slika 4.4. Grafički prikaz pozitivno asimetrične distribucije. Možemo videti na slici da je aritmetička sredina viša od medijane, a da je medijana viša od moda. Izduženi, debeli rep distribucije pruža se ka pozitivnim vrednostima, dok se na negativnom kraju distribucija naglo završava. Brojevi na vertikalnoj osi predstavljaju frekvencije, dok brojevi na horizontalnoj osi predstavljaju vrednosti varijable.

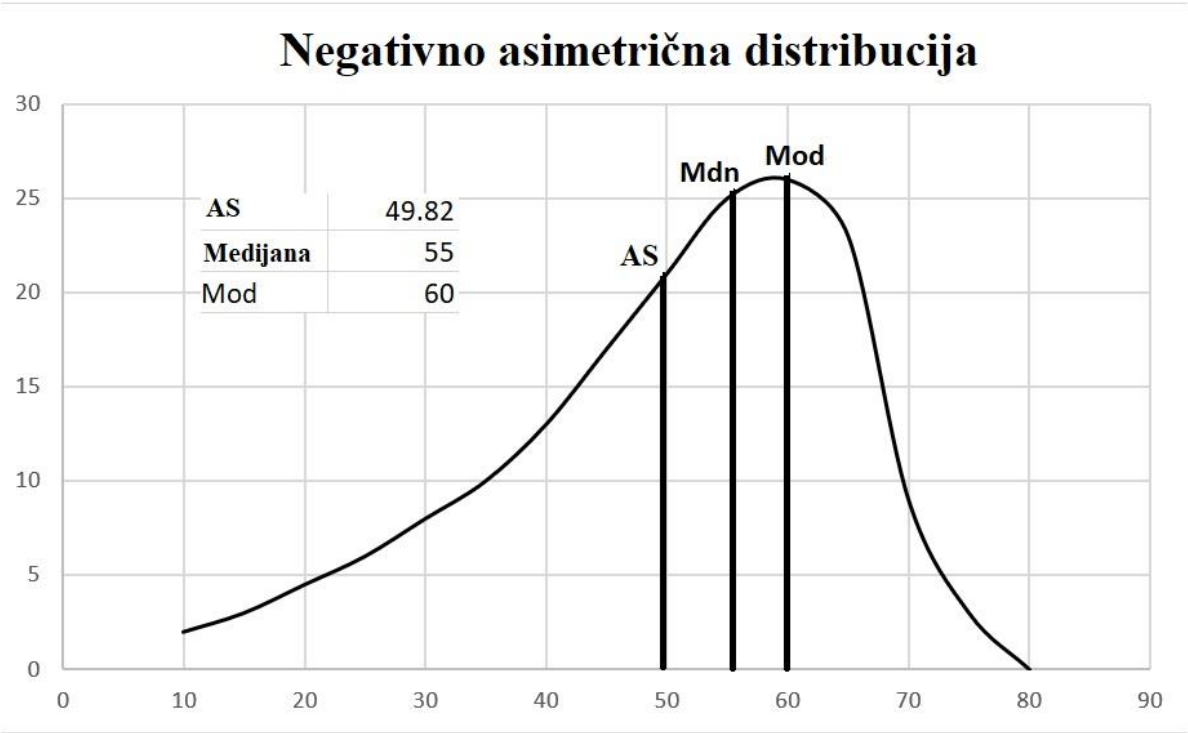
PREPRINT



- **Negativno asimetrična distribucija** je oblik distribucije u kom **je izdužena strana ona na donjem kraju distribucije**. Kod ovog oblika distribucije, **entiteti su koncentrisaniji na gornjem delu distribucije, a međusobno udaljeniji i razređeniji na donjem delu distribucije**. Kod ovog oblika distribucije, **mod je viši od medijane, a medijana je viša od aritmetičke sredine**. Kod psihološkog ili pedagoškog testiranja, negativno asimetrična distribucija obično ukazuje da je **test previše lak**. Opšta ideja testiranja je da se testovi prave tako da neki ispitanici uspeju da uspešno reše tek nekoliko stavki, neki da uspeju da reše većinu stavki, a da neki reše uspešno određeni srednji broj stavki. Na ovaj način se postiže to da ispitanici sa različitim nivoima izraženosti ispitivane osobine imaju različite skorove na testu. Nasuprot tome, negativno asimetrična distribucija ukazuje na to da većina ispitanika ima visoke skorove na testu i , prema tome, umesto da im skorovi budu raspršeni po celom rasponu mogućih vrednosti testovnih skorova, veliki broj njih je koncentrisan u relativno malom intervalu (visokih) testnih skorova otežavajući time to da ih se razlikuje u pogledu stepena izraženosti merene osobine. Na primer, ako imamo grupu učenika u kojoj skoro svi imaju najvišu moguću ocenu, unutar te grupe nećemo moći da ustanovimo razlike u stepenu savladanosti gradiva, osim u odnosu na onu nekolicinu učenika koji nemaju maksimalnu ocenu (to zbog toga što koristimo ocene za ovo razlikovanje, a svi imaju istu

ocenu – najvišu), a to je, ako nam je cilj da razlikujemo učenike po postignuću, podjednako loše kao i da su svi dobili loše ocene. Sa stanovišta razlikovanja postignuća ispitanika, negativno i pozitivno asimetrične distribucije su podjednako loše, s tim da je jedina razlika između njih u tome da jedna otežava razlikovanje postignuća najslabijih učenika, dok druga otežava razlikovanje postignuća najboljih učenika.

Slika 4.5. Grafički prikaz negativno asimetrične distribucije. Sa slike možemo videti da je aritmetička sredina niža od medijane, a da je medijana niža od moda. Izduženi, debeli rep distribucije pruža se prema pozitivnim vrednostima, dok se na pozitivnom kraju distribucija naglo završava. Brojevi na vertikalnoj osi su frekvencije, dok su brojevi na horizontalnoj osi vrednosti varijable.



Statistik koji pokazuje nivo horizontalnog odstupanja od normalne distribucije zove se skjunes. Različiti autori su ponudili različite postupke za računanje skjunesa, ali je, u većini formula**, računanje skjunesa zasnovano na odnosu između aritmetičke sredine i neke druge**

mere centralne tendencije, na broju entiteta ispod i iznad aritmetičke sredine ili na veličini pozitivnih i negativnih razlika između vrednosti entiteta i aritmetičke sredine kod dva vrste asimetričnih distribucija, a takođe i na činjenici da **što je distribucija asimetričnija, to će ove razlike biti veće**. Na primer, jedna relativno jednostavna formula za računanje skjunesa zasnovana na poređenju aritmetičke sredine i medijane je:

Skjunes=(3*(Aritmetička sredina – Medijana)/Standardna devijacija

Prema tome, **kada je aritmetička sredina veća od medijane, ovaj statistik skjunesa će biti pozitivan i tako pokazivati da je distribucija pozitivno asimetrična**. Kada je **aritmetička sredina manja od medijane, statistik skjunesa će biti negativan i tako pokazati da je distribucija negativno asimetrična.** Naravno, **kada je skjunes nula (ili dovoljno blizu nule) distribucija je simetrična**. Koeficijent izračunat po ovoj formuli zove se **drugi Pirsonov koeficijent skjunesa ili Pirsonov medijanski koeficijent skjunesa**. Još jedan široko korišćen koeficijent skjunesa zove se Pirsonov moment koeficijent skjunesa i zasnovan je na odstupanju pojedinačnih vrednosti od aritmetičke sredine (podignutih na treći stepen). Postoji i koeficijent skjunesa zasnovana na razlici između aritmetičke sredine i moda, tzv. Pirsonov prvi koeficijent skjunesa, a postoje i različiti drugi predloženi od strane različitih autora. Međutim, u osnovi je interpretacija svih ovih koeficijenata ista – **pozitivan skjunes znači da je distribucija pozitivno asimetrična, negativan skjunes znači da je distribucija negativno asimetrična**.

Sledeće pitanje koje se postavlja je **kako interpretirati veličinu skjunesa**? Koliko asimetrična treba distribucija da bude da bismo zaključili da se ne može smatrati normalnom? U trenutku pisanja ove knjige, nema nekog opšte prihvaćenog odgovora na ovo pitanje i različiti autori koriste i predlažu različite smernice za interpretaciju veličine skjunesa. Na primer, pojedini autori (e.g. Field, 2009; Tošković, 2020) predlažu da se **skjunes interpretira deljenjem koeficijenta skjunesa sa statistikom koji se zove standardna greška skjunesa** (ovaj statistik će biti objašnjem u poglavlju o statistici zaključivanja), a da se onda, ako je vrednost koja se dobije **iznad 1,96 ili ispod -1,96** (odnosno, u slobodnijoj varijanti, 2,56 i -2,56) **smatra da distribucija nije normalna**, dok ako je ovako dobijeni statistik unutar raspona navedenih vrednosti da se smatra da distribucija nije previše asimetrična. Ovu proceduru možemo **donekle uprostiti** tako što ćemo **1,96 da zaokružimo na 2** i da onda kažemo da koeficijenti skjunesa koji su preko dva puta veći

PREPRINT

od svoje standardne greške (dakle veći od standardne greške skjunesa pomnožene sa 2) opisuju distribucije koje nisu normalne. Ovaj postupak u neku ruku procenjuje verovatnoću da se distribucija sa skjunesom kakav je taj koji hoćemo da interpretiramo dobije na uzorku uzetom iz populacije čija je distribucija normalna. Međutim, **standardna greška je funkcija broja entiteta u uzorku i kada je broj entiteta veoma veliki, ona postaje veoma mala**. Tako onda deljenje skjunesa brojem koji je veoma malo daje rezultat koji je veliki i zbog toga, kako veličina uzorka raste, ova procedura postaje sve manje i manje korisna, jer veličine skjunesa koje prelaze malopre pomenuti kritični prag postaju sve manje i manje. To na kraju dovodi do situacije gde kod veoma velikih uzoraka, čak i vrlo mala odstupanja od simetričnog oblika prevazilaze ove granice. Neki drugi autori su stoga predložili različite i često vrlo proizvoljne kritične nivoe različitog nivoa strogosti. Na primer, **Hahs-Vaughn & Lomax (2020) navode da je raspon skjunesa od +3 do -3 vrlo liberalan raspon za to da se distribucija ne smatra previše asimetrično, da bi umereno veliki raspon skjunesa bio od +2 do -2, a da bi konzervativni raspon bio od +1 do -1.**

Vertikalna odstupanja od normalne distribucije se dešavaju onda **kada je distribucija „previše šiljata“ sa previše entiteta skoncentrisanih u centru i sa vrednostima nedovoljno razuđenim** ili **kada su entiteti vrednosti entiteta previše razuđene, više raštrkane oko aritmetičke sredine nego što je to slučaj kod normalne distribucije**. Ove ove situacije se tipično dešavaju kada su uslovi na kojima je zasnovano očekivanje da distribucija bude normalna narušeni/nisu ispunjeni. Na primer, ako glavni faktor koji bi trebalo da bude konstanta i koji određuje centralnu vrednost distribucije nije konstantna, već ima vrednosti koje su dovoljno bliske da distribucija i dalje izgleda kao jedinstveni kontinuum (umesto dve diskretne grupe vrednosti), to se može manifestovati kao distribucija koja je razuđenija od normalne distribucije. S druge strane, kada ima premalo faktora pored tog glavnog ili su oni preslabi da dovedu do potrebnog nivoa individualnih razlika ili kada postoje faktori koji sprečavaju javljanje razlika u vrednostima entiteta, rezultat može biti distribucija koja je „šiljatija“ od normalne distribucije, sa vrednostima entiteta koje su više zgusnute oko sredine nego što je to slučaj kod teorijske normalne distribucije. Na primer, u kontekstu psihološkog ili pedagoškog testiranja, kada je test previše lak ili previše težak, pa tako dovodi do toga da svi ispitanici budu grupisani na maksimumu ili na minimumu, takva distribucija, pored toga što će biti asimetrična, biće vrlo verovatno i „šiljata“ tj. vrednosti ispitanika biće više koncentrisane nego što bi se očekivalo kod normalne distribucije. Slična situacija se može desiti i onda kada učenici neki test rade zajedno, tako što sarađuju ili prepisuju jedni od drugih, ali podnose

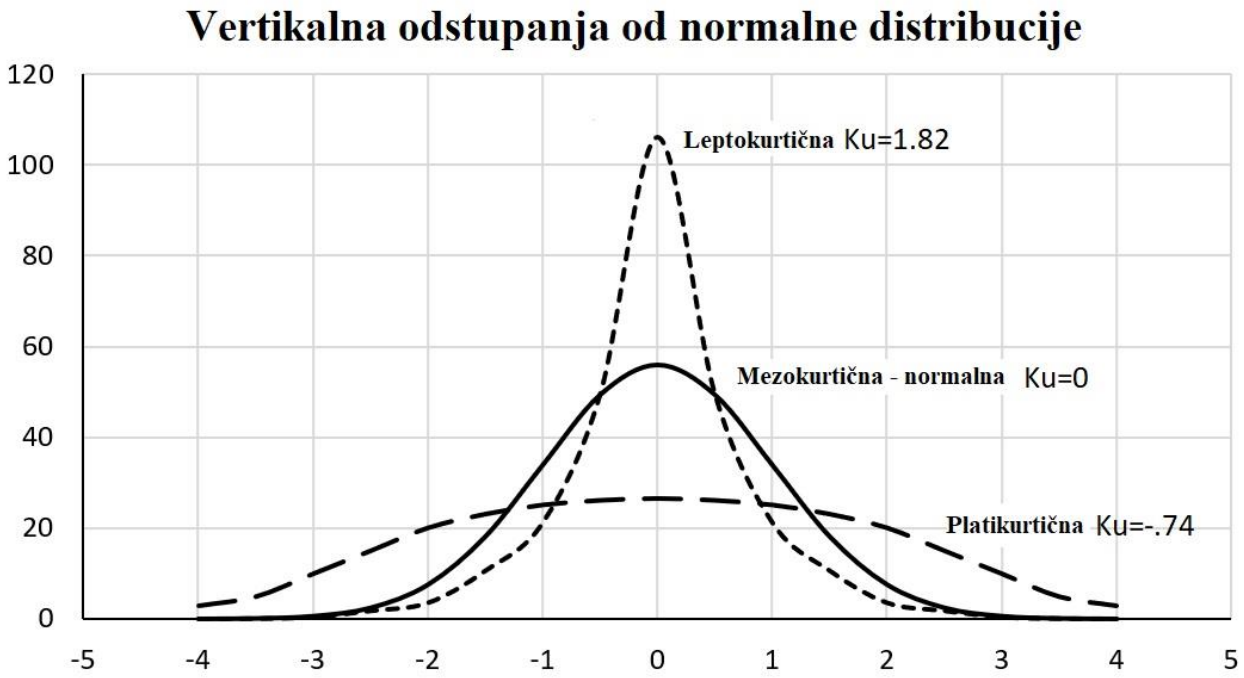
PREPRINT

testove individualno, kao da su ih sami radili. Kada se to desi, može se očekivati da varijabilnost bude manja od normalne, čak i u slučajevima kada učenici na neka pitanja odgovore pogrešno, i to će onda dovesti do „šiljate“ distribucije. S druge strane, kada na primer, neko društvo počne da se polarizuje oko neke političke teme i to tako da populacija počne da se deli na dve veoma različite grupe kada je u pitanju stav oko date teme, populacija će vrlo verovatno u jednom trenutku tog procesa proći kroz fazu gde će distribucija stavova o toje temi biti „spljoštena“, imati razuđen oblik, pre nego što se, kako se polarizacija nastavi, podeli u dve posebne grupe u pogledu ovih stavova.

U statističkoj literaturi ćemo često sresti termine **platikurtična i leptokurtična distribucija** u kontekstu opisivanja vertikalnih odstupanja od normalne distribucije. U ovoj terminologiji, **normalna distribucija je mezokurtična distribucija**. Najčešće, **termin platikurtična distribucija se povezuje sa „spljoštenim distribucijama“, gde su vrednosti entiteta međusobno razuđenije, dok se termin leptokurtična distribucija tipično povezuje sa „šiljatim“ distribucijama, kod kojih su entiteti više koncentrisani nego što je to slučaj sa normalnom distribucijom.**

Slika 4.6. Grafički prikaz vertikalnih odstupanja od normalne distribucije. Na slici možemo videti da „šiljata“, leptokurtična distribucija ima pozitivnu vrednost viška kurtozisa, da normalna distribucija ima nultu vrednost i da je vrednost viška kurtozisa „spljoštene“, platikurtične distribucije negativna. Treba međutim da imamo u vidu da dodavanje malog broja autlajera, entiteta na krajeve normalne distribucije (na rep distribucije) povećava kurtozis, a da će dodavanje entiteta na mesta na distribuciji koja su umereno udaljena od aritmetičke sredine (na ramena distribucije) težiti da smanji vrednost kurtozisa.

PREPRINT



Međutim, **takva povezivanja nisu uvek tačna jer su koncepti platikurtične i leptokurtične distribucije dosta složeniji od toga**. Naime, **statistik koji se koristi da bi se ustanovilo da li je distribucija platikurtična, leptokurtična ili mezokurtična zove se kurtozis**. **Formula za računanje kurtozisa zasnovana je na proseku razlika između vrednosti pojedinačnih entiteta i aritmeiičke sredine podignute na četvrti stepen i podeljene sa standardnom devijacijom podignutom na 4 stepen.** Ta formula izgleda ovako:

$$Ku = \frac{\frac{\sum_{i=1}^{n}(X_i - \mathrm{AS})^4}{N}}{SD^4} - 3$$

U ovoj formuli X se odnosi na vrednosti pojedinačnih entiteta iz uzorka, AS je aritmetička sredina uzorka, a SD je standardna devijacija. N je broj entiteta u uzorku. **Deo formule bez -3 na kraju predstavlja tzv. četvrti standardizovani momenat, a kada se toj formuli doda -3, statistik dobijen na taj način zove se višak kurtozisa**. **Ako je vrednost kurtozisa izračunata na ovaj način veća od nula tj. pozitivna, distribucija je leptokurtična. Ako je vrednost kurtozisa negativna, distribucija je platikurtična. Vrednost kurtozisa mezokurtične distribucije je 0,** ako koristimo formulu koja je prikazana – formulu za višak kurtozisa. Kada se koristi normalna

formula za četvrti standardizovani momenat, dakle formula bez -3 na kraju, vrednost mezokurtične distribucije je 3. Formula za višak kurtozisa koriguje ovu vrednost da bude 0, tako što oduzima broj 3 od formule za četvrti standardizovani moment, čineći tako rezultat zgodnijim za interpretaciju tj. omogućavajući da znak statistika predstavlja smer odstupanja od normalne distribucije.

Ovde valja primetiti da **podizanje razike od aritmetičke sredine na četvrti stepen za rezultat ima brisanje predznaka date razlike, tako da i pozitivna i negativna odstupanja od aritmetičke sredine povećavaju vrednost kurtozisa.** Iz istog razloga **velika odstupanja od aritmetičke sredine tj. ekstremne vrednosti utiču na ukupnu vrednost kurtozisa mnogo više nego mala odstupanja**. Važno je primetiti da **se zbog ovih svojstava kurtozis ne može stvarno tretirati kao mera „šiljatosti" centralnog dela distribucija jer na njegovu vrednost u velikoj meri utiče i veličina krajeva tj. „repova" distribucije** (e.g. Darlington, 1970; Westfall, 2014).

Ako pogledamo formulu za kurtozis i proučimo to kako ona radi, možemo zaključiti sledeće stvari:

- **Dodavanje u uzorak entiteta čije su vrednosti oko centra distribucije će povećati vrednost kurtozisa i to tako što će smanjiti standardnu devijaciju uzorka** (koja je pak delilac u formuli za kurtozis). Ovo se dešava zato što ovi dodatni entiteti povećavaju veličinu uzorka, koja je delilac u formuli za standardnu devijaciju, a u isto vreme u dosta manjoj meri povećavaju sumu odstupanja od aritmetičke sredine jer su blizu aritmetičke sredine, pa su vrednosti odstupanja male.
- **Dodavanje malog broja entiteta na same krajeve distribucije, a pogotovo dodavanje autlajera takođe povećava kurtozis**. Ovo se dešava zbog toga što ovi entiteti povećavaju prosek odstupanja entiteta od aritmetičke sredine podignut na četvrti stepen mnogo više nego što povećavaju standardnu devijaciju uzorka (čak i kad se ova podigne na četvrti stepen). **Ovo dodavanje autlajera povećava vrednost kurtozisa samo dok je broj entiteta blizu aritmetičke sredine mnogo puta veći od broja dodatih autlajera** tj. dokle god je broj entiteta u uzorku dovoljan da uticaj ovih dodatnih autlajera na vrednost standardne devijacije bude mali. **Ako se doda previše autlajera, dalje dodavanje autlajera će smanjivati vrednost kurtozisa umesto da je povećava.**
- **Ako je naša varijabla konstanta** tj. ako svi entiteti imaju istu vrednost, **standardna devijacija će biti nula**, što će onda značiti da je **vrednost kurtozisa beskonačna ili**

PREPRINT

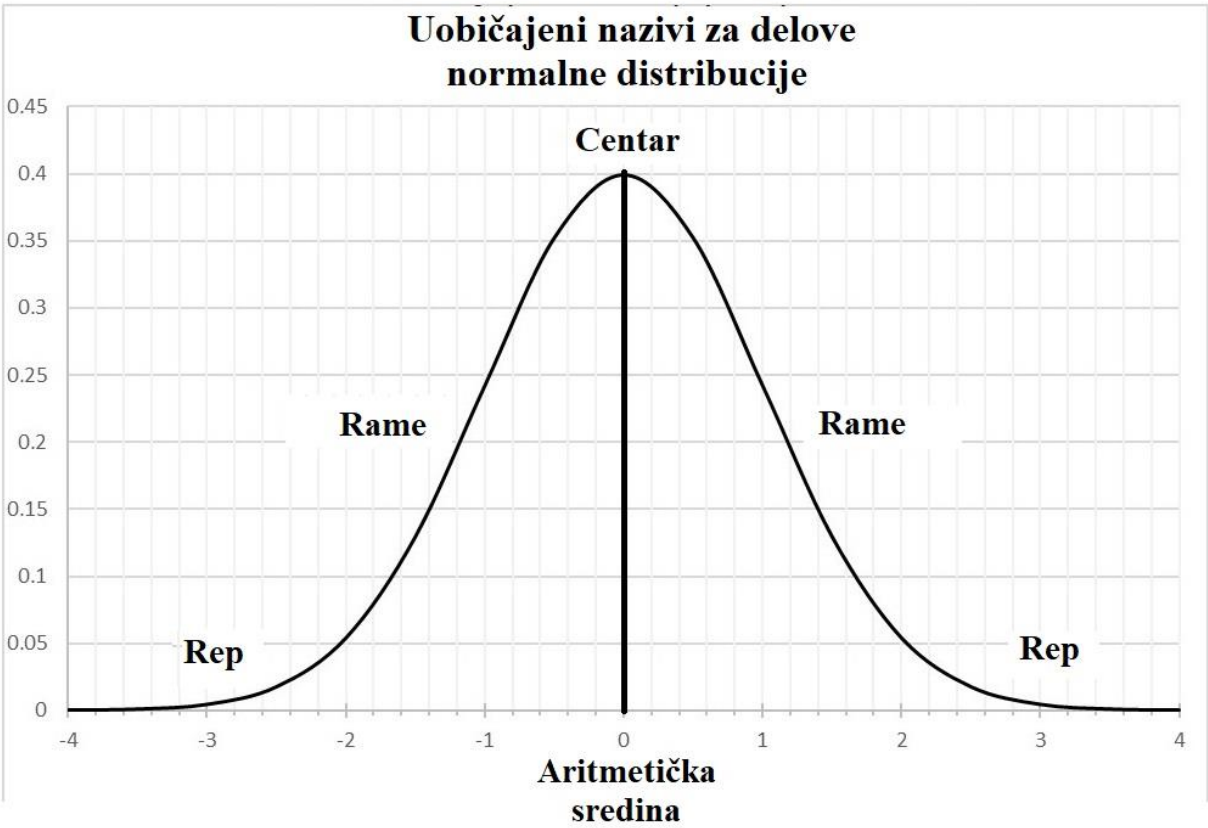
neizračunljiva zato što obuhvata deljenje sa nulom (0/0). **Ako bismo imali samo jedan entitet ili par njih koji imaju vrednost različitu od vrednosti svih ostalih, vrednost kurtozisa bi bila veoma visoka umesto beskonačna/neizračunljiva**. Međutim, **najmanja moguća vrednost viška kurtozisa, a to je -2, može se dobiti u situacijama kada varijabla ima samo 2 različite vrednosti** i polovina entiteta ima jednu, a druga polovina drugu vrednost.

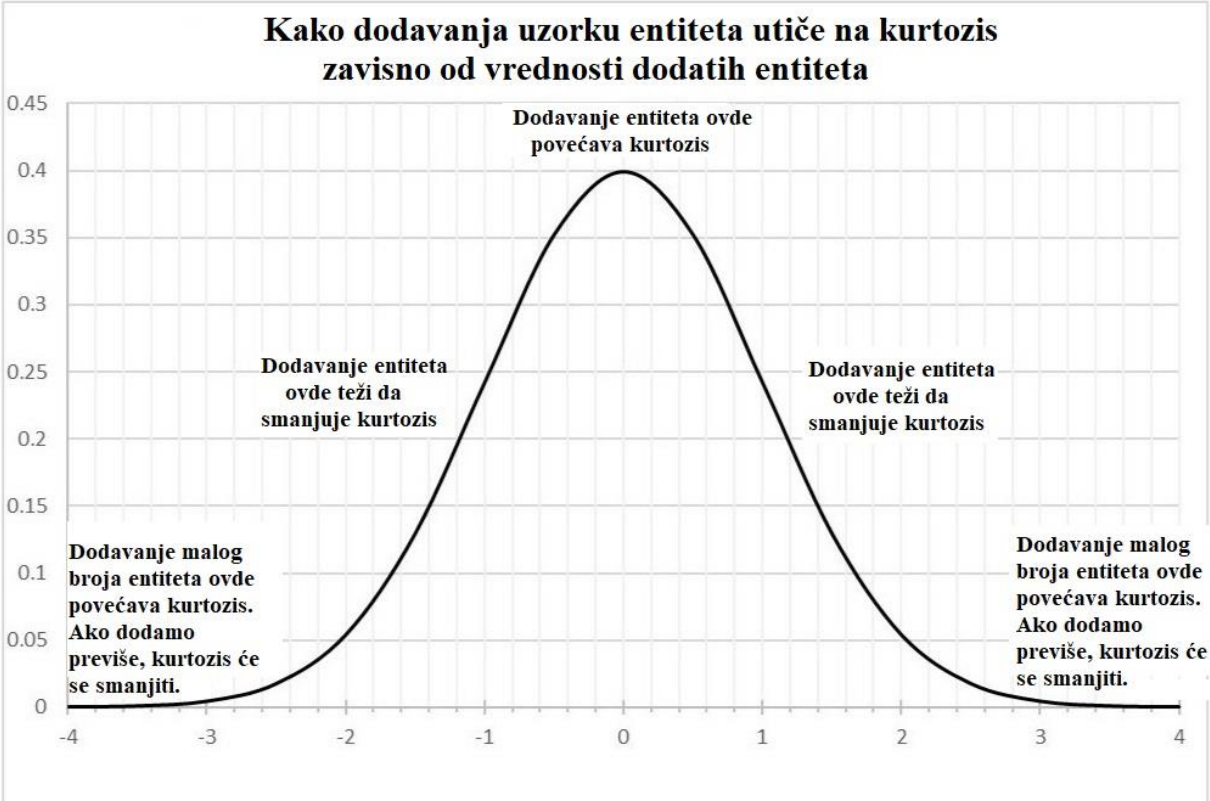
U naučnoj i statističkoj literaturi, **nazivi koji se tipično koriste za ove različite delove normalne distribucije** (ili distribucije koja je slična normalnoj) su:

- **Centar** – to je oblast oko **centralnog dela/sredine distribucije.**
- **Repovi** – to su **oblasti na samim krajevima distribucije**, kako na levoj, tako i na desnoj strani tj. i na donjoj i na gornjoj strani u odnosu na vrednosti varijable.
- **Ramena** – to su **delovi distribucije između repova i centra distribucije**.

Slika 4.7. Delovi normalne distribucije, onako kako se tipično nazivaju u literaturi (gornja slika). **Centralni deo distribucije naziva se centrom**, **krajevi distribucije**, sa obe strane, **zovu se repovi**, a **oblast između centra i repova se naziva ramenima distribucije**. Donja slika ilustruje orijentaciono kako se kurtozis menja kada se entiteti dodaju na različitim delovima distribucije (donja slika). Dodavanje entiteta čije su vrednosti blizu aritmetičke sredine uzorka će povećati kurtozis. Dodavanje entiteta sa vrednostima koje odgovaraju samim krajevima distribucije će do jednog momenta (dok je broj dodatih entiteta srazmerno mali) povećavati kurtozis, a dodavanje dodatnih entiteta preko toga će početi da smanjuje kurtozis. Dodavanje entiteta čije vrednosti odgovaraju ramenima distribucije težiće da smanji kurtozis. Ovi efekti se odnose na uzorke koji su normalno distribuirani. Ako je distribucija previše različita od normalne i ovi efekti se mogu razlikovati. Podaci na vertikalnoj osi su frekvencije, dok su oni na horizontalnoj osi vrednosti varijable.

PREPRINT

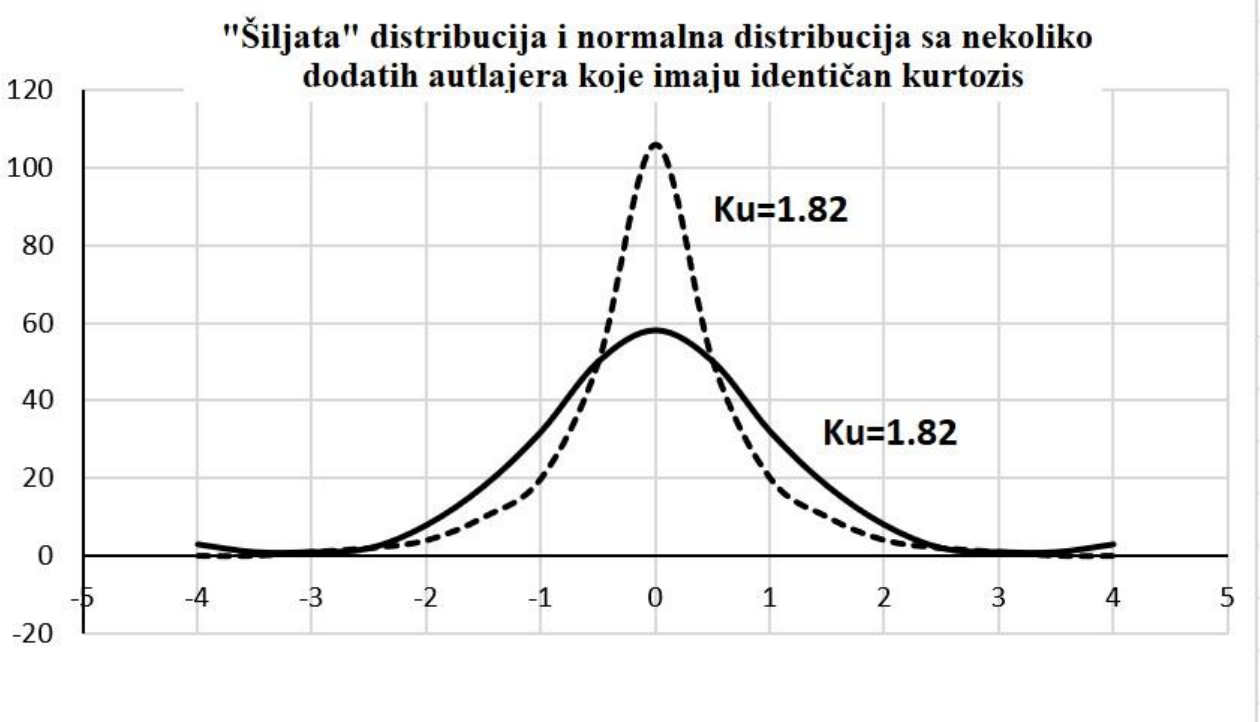

Uobičajeni nazivi za delove normalne distribucije
Centar
Rame
Rame
Rep
Rep
Aritmetička sredina
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
-4
-3
-2
-1
0
1
2
3
4


Kako dodavanja uzorku entiteta utiče na kurtozis zavisno od vrednosti dodatih entiteta
Dodavanje entiteta ovde povećava kurtozis
Dodavanje entiteta ovde teži da smanjuje kurtozis
Dodavanje entiteta ovde teži da smanjuje kurtozis
Dodavanje malog broja entiteta ovde povećava kurtozis. Ako dodamo previše, kurtozis će se smanjiti.
Dodavanje malog broja entiteta ovde povećava kurtozis. Ako dodamo previše, kurtozis će se smanjiti.
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
-4
-3
-2
-1
0
1
2
3
4

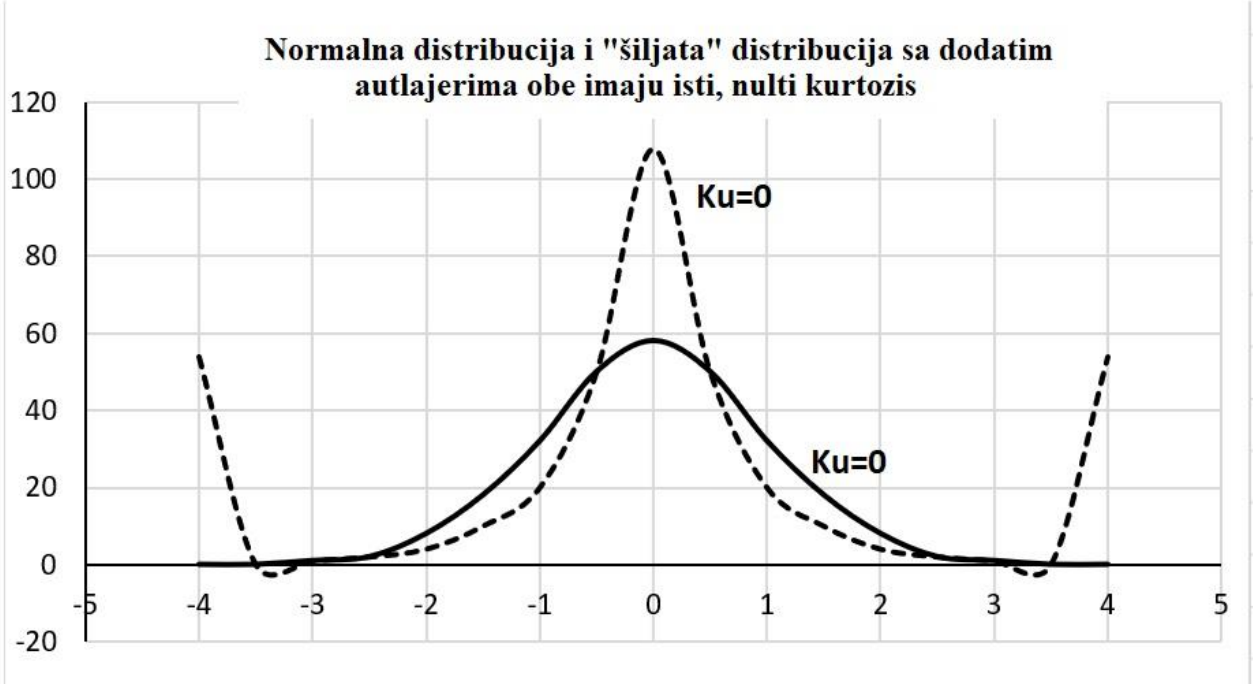
Ova svojstva kurtozisa dovode do toga da **distribucije sa vrlo različitim oblicima mogu da imaju iste vrednosti kurtozisa, što može da dovede istraživače do pogrešnih zaključaka**. Na primer, mezokurtična distribucija ili čak i blago platikurtična distribucija sa relativno malim brojem autlajera može imati višu vrednost kurtozisa od autentično „šiljate“ distribucije (distribucije čiji su entiteti visoko koncentrisani oko sredine). Jasno „šiljata distribucija“ koja ima i značajan broj autlajera može imati kurtozis nula i tako navesti istraživače da poveruju da je pred njima normalna, mezokurtična distribucija. Postoje i razne druge kombinacije, a svima im je zajedničko to da mogu dovesti istraživače koji ne razumeju dovoljno svojstva kurtozisa do pogrešnih zaključaka o strukturi podataka, kako je pokazano u primerima, ako istraživači ne bi izvršili inspekciju oblika distribucije na drugi način (npr. vizuelno).

Slika 4.8. Grafičko poređenje između „šiljate distribucije“ i normalne distribucije kojoj je dodato nekoliko entiteta na oba kraja distribucije i koje sada imaju isti kurtozis. Obe distribucije su napravljene na uzorcima od 280 entiteta i pre ovog dodavanja, ona sa nižim centrom je imala vrednost viška kurtozisa od 0, što znači da je bila mezokurtična tj. normalna, dok je „šiljata“ distribucija imala vrednost viška kurtozisa od 1,82. Dodavanje svega 8 entiteta na repove ove normalne distribucije dovelo je do toga da vrednosti viška kurtozisa ove dve distribucije postanu jednake tj. da obe distribucije budu leptokurtične. Ako bi dodali još entiteta na njene repove, kurtozis ove distribucije sa nižim centrom (koja je u startu bila normalna) bi postao veći od kurtozisa prikazane „šiljate“ distribucije. Međutim, dodavanje mnogo većeg broja entiteta na repove distribucije bi brzo ponovo smanjilo vrednost kurtozisa (vidi sliku 4.7.). Brojevi na vertikalnoj osi su frekvencije, dok brojevi na horizontalnoj osi predstavljaju vrednosti varijable.

PREPRINT



Slika 4.9. Grafičko poređenje normalne distribucije i „šiljate“ distribucije sa visokim kurtozisom kojoj je dodat značajan broj entiteta na repove čime je njen (višak) kurtozisa smanjen na 0. Obe disitribucije su u startu napravljene na uzorcima od 280 entiteta, ali je dodatnih 110 entiteta dodato na repove „šiljate“ distribucije da bi joj smanjili kurtozis na 0, čime je ova distribucija suštinski pretvorena u distribuciju sa tri tačke koncentracije entiteta. Ovo takođe pokazuje kako oblici distribucije koji su veoma različiti od normalne distribucije i dalje mogu imati vrednost viška kurtozisa od 0 tj. vrednost identičnu onoj koju ima normalna distribucija. Brojevi na vertikalnoj osi su frekvencije, dok su na horizontalnoj osi vrednosti varijable.



Imajući sve ovo na umu, treba istaći da je **obično glavni i ,najčešće, jedini razlog za računanje kurtozisa želja da se utvrdi da li je uzorak normalno distribuiran ili nije**. A **u slučaju da nije, ono što obično želimo da saznamo korišćenjem kurtozisa je to da li je distribucija uzorak „šiljata“ ili „spljoštena“**. Oblici distribucije koje zapravo hoćemo da prepoznamo su onaj u kom su entiteti gusto raspoređeni, visoko koncentrisani, kao i onaj gde su međusobno raspršeniji nego što je to slučaj na normalnoj distribuciji. U takvoj situaciji, **dobijanje visokog, pozitivnog kurtozisa samo zbog nekoliko outlajera na distribuciji koja je inače po svom obliku normalna ili dobijanje nultog kurtozisa u situaciji kada je distribucija „šiljata“, ali ima puno autlajera je veoma, veoma nepoželjan ishod**. Takav ishod postaje još više nepoželjan ako imamo ju vidu da su autlajeri često posledica lošeg ili nevalidnog merenja (na primer, ljudi koji nisu ozbiljno pristupili testiranju ili koji su znali odgovore unapred) ili čak grešaka prilikom unosa podataka (na primer, kada se unese 55, umesto 5 ili kada se unese 7 umesto 4, zato što je 7 na numeričkoj tastaturi odmah iznad četvorke!). Zbog svega ovog, **pre nego što pristupimo interpretaciji vrednosti kurtozisa, važno je da vizuelno pregledamo oblik distribucije i ustanovimo da li su i u kojoj meri na vrednost kurtozisa koju smo dobili uticali autlajeri**. Takođe **može biti korisno i da izračunamo vrednost kurtozisa nakon što mali broj entiteta sklonimo sa svakog kraja distribucije i da onda uporedimo tako dobijeni rezultat sa kurtozisom na celom uzorku.** Međutim, ovo uklanjanje entiteta sa krajeva nikako ne bi smelo da

PREPRINT

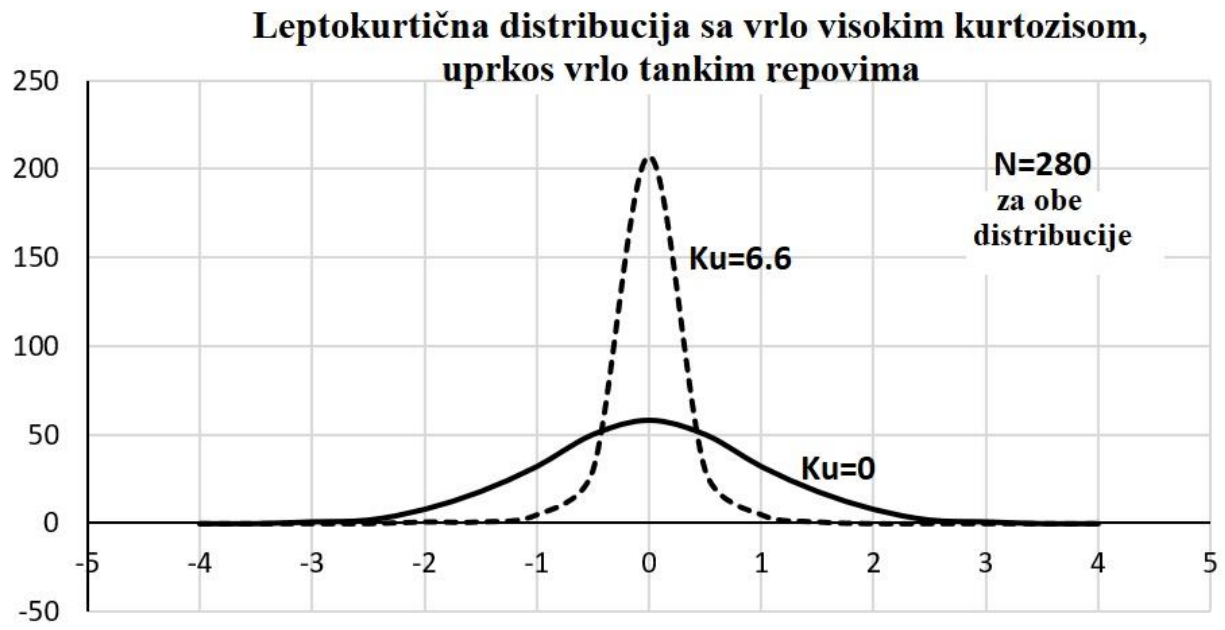
zameni vizuelni pregled distribucije, jer i mnogi mnogo neobičniji oblici distribucije od onih koji su ovde razmatrani mogu da daju normalan kurtozis (a i skjunes). Na primer, distribucija sa tri jasna vrha može veoma lako da ima skjunes od 0 (u suštini, to je u izvesnoj meri ono što prikazuje slika 4.9). Imajući sve ovo u vidu, možemo zaključiti da uprkos popularnosti i širokoj upotrebi, **kurtozis nije baš naročito dobar indikator stepena vertikalnog odstupanja distribucije od oblika normalne distribucije.**

Još jedna **tema koju valja prodiskutovati je tvrdnja koja se često sreće među istraživačima**, pa i u literaturi, a koja kaže da je kurtozis „mera debljine repove distribucije". Ova tvrdnja se može naći u brojnim naučnim i obrazovnim tekstovima, uključujući i udžbenike statistike. Međutim, **kurtozis nije mera debljine repova distribucije!!** Da je kurtozis mera debljine repova distribucije, sve distribucije sa visokim kurtozisom imale bi debele repove. Ali to nije slučaj i to se može jasno videti i u mnogim tekstovima koji iznose ovakve tvrdnje, a potom leptokurtične distribucije predstavljaju slikama distribucija sa veoma tankim repovima ili skoro bez repova. Iako, kako je ranije pomenuto, dodavanje entiteta repovima distribucije tj. zadebljavanje repova, može povećati kurtozis, to tako funkcioniše samo u jednoj vrlo ograničenoj meri. Dodavanje više od toga broja entiteta (malog u odnosu na veličinu uzorka!) repovima distribucije dovešće do toga da kurtozis počne da se smanjuje. To znači da postoji obrnuti U odnos između broja entiteta dodatih repovima distribucije i promene kurtozisa. Pokazatelj čiji se smer promene obrne u određenoj tački sa daljim rastom vrednosti osobine čiji pokazatelj treba da bude, a koji takođe može imati i veoma visoke vrednosti kada je osobina čiji je to pokazatelj vrlo niska ili je 0 nije zapravo nikakav pokazatelj te osobine. Takođe, treba imati na umu i to da je činjenica da na kurtozis utiču autlajeri i entiteti na samim krajevima distribucije u većini slučajeva vrlo nepoželjno svojstvo formule za kurtozis, a ne nešto što kurtozis zapravo želimo da meri. To su sve razlozi zašto su **tvrdnje da je kurtozis pokazatelj debljine repova distribucije jednostavno pogrešne**.

Slika 4.10. Kurtozis nije pokazatelj debljine repova distribucije! Ovaj primer pokazuje distribuciju koja skoro da i nema repove uopšte, a koja, uprkos tome, ima veoma visok kurtozis. Poređenja radi je prikazana i normalna distribucija. Ovaj primer govori protiv tvrdnji koje se često mogu čuti od istraživača i nekada naći u literaturi o tome kako je kurtozis pokazatelj debline repova distribucije. On to nije, iako debljina repova distribucije tj. broj entiteta na krajevima distribucije

PREPRINT

utiče na kurtozis na jedan složen način koji je objašnjen u ovom poglavlju. Brojevi na vertikalnoj osi su frekvencije, dok horizontalna osa predstavlja vrednosti varijable.



Nakon što smo pregledali oblik distribucije i zaključili da vrednost kurtozisa nije posledica nekog neželjenog efekta tj. da se može smatrati za validan pokazatelj vertikalnog odstupanja distribucije od normalne, **kako ćemo interpretirati veličinu kurtozisa**? Kao i kod skjunesa, ne postoji univerzalno prihvaćeni metod interpretacije veličine kurtozisa. Kao i kod skjunesa, jedan način da se interpretira vrednost kurtozisa je da se **vrednost kurtozisa podeli sa standardnom greškom kurtozisa** (ovaj statistik je objašnjem u poglavlju o statistici zaključivanja**), a da se onda proveri da li je dobijena vrednost između -1,96 i +1,96** (odnosno između -2,56 i +2,56) (e.g. Field, 2009; Tošković, 2020). Ovo se takođe može **uprostiti** tako što ćemo **1,96 zaokružiti na dva i reći da distribucija nije normalna ako je vrednost kurtozisa više nego 2 puta veća od standardne greške kurtozisa**. Ponovo, ovaj metod procene **radi dobro za male uzorke, ali ne i za velike**, jer je standardna greška funkcija veličine uzorka (veći uzorak – manja standardna greška). Zbog ovoga, neki autori predlažu da se **vrednosti (viška) kurtozisa između +1 i -1 uzmu kao granice**, pa da se **distribucije čiji je kurtozis unutar ovog raspona tretiraju kao dovoljno bliske normalnoj**, a da **se za one čiji je kurtozis van tog raspona smatra da ne odgovaraju obliku normalne distribucije** (e.g. Hair et al., 2016). Valja se podsetiti da, kao što je napred rečeno, **vrednost (viška) kurtozisa ne može biti manja od -2.**

****

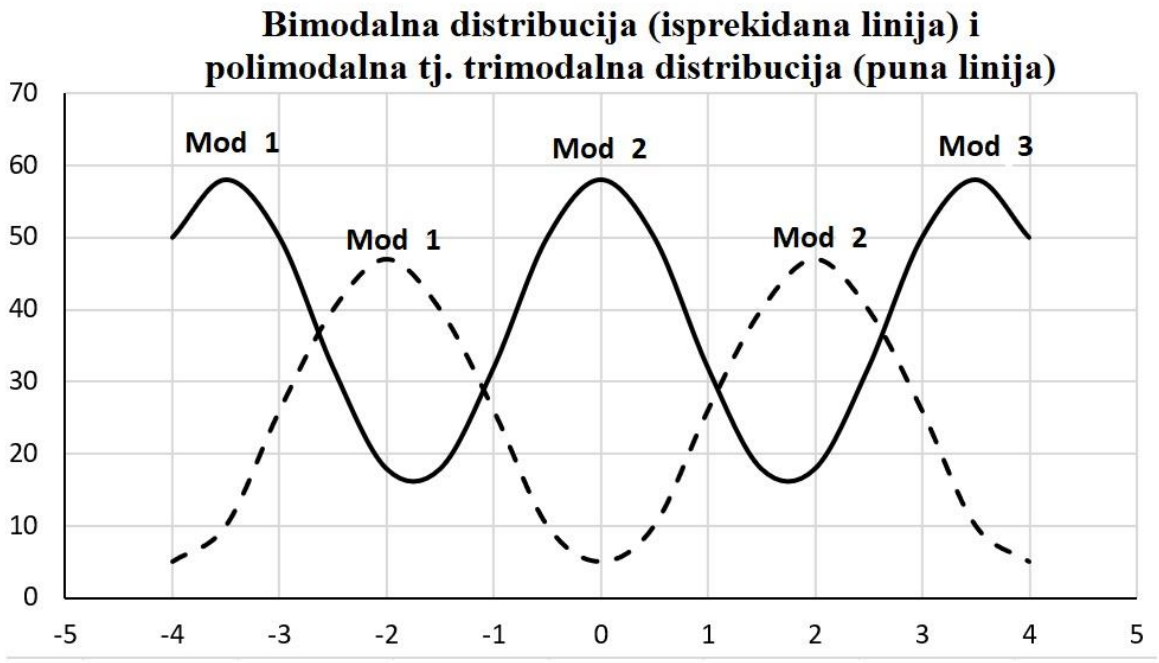
Pored opisanih horizontalnih i vertikalnih odstupanja od normalne distribucije, **empirijske distribucije mogu da odstupaju od oblika normalne distribucije i na različite kompleksnije načine**. Jedna takva situacija je ona kada empirijska distribucija ima više od jednog moda. Takve distribucije nazivaju se **bimodalne, kada imaju dva moda i polimodalne, kada imaju više od dva moda.** Ovakve distribucije se **nekada imenuju i prema tačnom broju modova koje imaju**, pa tako imamo **trimodalne distribucije, kvartimodalne distribucije i druge**. **Bimodalne i polimodalne distribucije se tipično dobijaju onda kada je pretpostavka o jednom glavnom faktoru koji je konstanta za celu populaciju narušena.** Ovo takođe može da se odnosi i na situacije kada je uzorak uzet iz 2 ili više različitih populacija. Na primer, jedna distribucija za koju se često očekuje da bude bimodalna u ljudskoj populaciji je distribucija visina, pri čemu imamo mod visina žena i poseban mod visina muškaraca. Ako pogledamo zarade zaposlenih u određenoj grupi široko rasprostranjenih profesija (na primer, konobara, prodavaca u maloprodaji), ali napravimo uzorak tako da se sastoji zajedno od ljudi iz neke visoko razvijene oblasti sa snažnom ekonomijom i iz druge slabo razvijene oblasti čija ekonomija jedva preživljava, distribucija zarada tih ljudi će vrlo verovatno imati dva moda – jedan za ljude iz siromašne oblasti i drugi za ljude iz bogate oblasti. U situacijama jakih političkih sukoba ili građanskog rata, ako bi ispitivali stavove ljudi iz suprotstavljenih grupa prema glavnoj temi političkog sukoba, distribucija njihovih stavova bi takođe verovatno imala dva ili više moda (svaka grupa bi bila poseban mod). Zapravo, u situacijama poput ove, primećivanje da se distribucija stavova transformisala u bimodalnu ili polimodalnu može biti rani znak da je politički sukob „odmah iza ugla“ i može ukazivati na postojanje pretnje da dođe do političkog nasilja ili građanskog rata u društvu ako je tema oko koje je došlo do polarizacije dovoljno važna, a njeno potencijalno nepovoljno razrešenje opaženo kao dovoljno preteće za blagostanje ili identitet neke od grupa (Hamburger et al., 2021).

Polimodalnost distribucije se najbolje može otkriti pregledom grafičkog ili tabelarnog prikaza distribucije i prosto prebrojavanjem broja vrhova distribucije i generalno njenog oblika. **Veoma niska vrednost kurtozisa može da bude i pokazatelj polimodalne distribucije**. Međutim, treba napomenjuti da **polimodalna distribucija nekada može imati i skroz neupadljive vrednosti skjunesa i kurtozisa,** pa se istraživačima preporučuje da uvek obave

PREPRINT

vizuelni pregled grafičkih prikaza distribucije podataka i da se ne oslanjaju samo na vrednosti skjunesa i kurtozisa.

Slika 4.11. Primer bimodalne distribucije (isprekidana linija) i polimodalne distribucije sa tri moda (puna linija).



4.7. Standardni skorovi i standardizacija

U istraživačkoj praksi često postoji potreba da se uporede pozicije na distribuciji varijable. Nekada postoji potreba da se uporede pozicije dva entiteta na distribuciji iste varijable, a nekada istog entiteta na distribucijama dve različite varijable. Dok postoji niz načina da se uporede vrednosti dva entiteta na istoj varijabli, samo poređenje sirovih vrednosti (vrednosti varijable onakve kakve su u startu izmerene) najčešće ne govori samo od sebe o poziciji na distribuciji. Ovo pitanje postaje još teže kada treba uporediti pozicije na distribucijama dve različite varijable a pri tom su, na primer, te varijable merene na različitim skalama i koriste različite jedinice mere. Jedno takvo pitanje, koje zahteva poređenje različitih skala, moglo bi da bude, na primer, - da li je u odnosu na druge ljude određena osoba više visoka ili više teška? Naravno, ne bi imalo nikakvog smisla

porediti kilograme sa metrima. Međutim, sasvim bi imalo smisla uporediti relativne pozicije date osobe na distribucijama težine (odnosno telesne mase) i visine i ustanovili da li je pozicija date osobe u odnosu na druge ljude iz uzorka viša na visini ili na težini. To zapravo i jeste način kako naučnici, a i ljudi generalno, dolaze do velikog broja praktičnih zaključaka – na primer, za osobu čija je pozicija na distribuciji visine niska, ali na distribuciji težine visoka, mogli bi da zaključimo da je gojazna, dok bi za osobu koja je oko sredine na distribuciji težine, a visoko pozicionirana na distribuciji visine, mogli da zaključimo da je mršava i tako dalje. Da bi izveli zaključke koristeći statistiku, mogli bi da koristimo percentile ili percentilne rangove i ovi statistici se zaista i koriste za donošenje ovakvih i sličnih zaključaka. Međutim, problem sa percentilima i percentilnim rangovima je u tome što su oni na ordinalnom nivou merenja, pa bi brojne dalje računice s njima, poput, na primer, onih koje zahtevaju sabiranje i oduzimanje bile uglavnom nemoguće (ili ne bi bile smislene – brojevi tolerišu sve, ali rezultati gube smisao ako se stvari ne urade kako treba). Određeni stepen poboljšanja u odnosu na precentile za ove svrhe obezbeđuje upotreba standardnih skala i standardnih skorova. **Standardne skale su skale sa unapred definisanim svojstvima**, što uglavnom znači da **imaju unapred definisanu aritmetičku sredinu i standardnu devijaciju**, a često i **unapred definisan oblik distribucije**. Sirove vrednosti varijable (vrednosti dobijene postupkom merenja ili procene) se onda transformišu u skorove na standardnoj skali. **Proces transformacije sirovih vrednosti varijabli (odnosno sirovih skorova) u skorove standardne skale zove se standardizacija.** Zato što su svojstva standardne skale unapred određena**, poznato je uvek koji skor na standardnoj skali (ili samo – standardni skor) odgovara kojoj poziciji na distribuciji tj. kom percentilu**. Međutim, za razliku od percentila, **standardni skorovi su na intervalnom nivou merenja**, što dozvoljava izvođenje različitih računanja koja ne bi bilo moguće smisleno izvesti na percentilima i percentilnim rangovima (zato što su percentili i percentilni rangovi ordinalne mere!) Standardni skorovi se široko koriste u stručnoj praksi psihologa. Na primer, široko poznati IQ skorovi (koji se koriste u proceni inteligencije i drugih kognitivnih sposobnosti) su vrsta skorova na standardnoj skali, čija je unapred definisana aritmetička sredina 100, a standardna devijacija 15 ili 16 (V. Hedrih, 2020), a ima i različitih drugih primera.

U statistici, termin „standardna skala“ se obično odnosi na z skalu, a **standardizacija se tipično odnosi na transformaciju sirovih vrednosti u z skorove** (tj. na z skalu). **Z skala je standardna skala čija je aritmetička sredina 0, a standardna devijacija 1.** Za z skorove je takođe vezano i teorijsko očekivanje da su **normalno distribuirani** tj. računanje z skorova obično

implicira da oni grade normalnu distribuciju (iako ćemo u praksi sretati z skorove dobijene na empirijskim podacima koji grade distribuciju koja nije oblika normalne!). **Podaci se prevode na z skalu oduzimanje aritmetičke sredine uzorka od vrednosti pojedinačnog entiteta, a onda deljenjem dobijenog rezultata standardnom devijacijom**. Dobijeni rezultat je **z skor:**

Z skor=(Sirova vrednost varijable – aritmetička sredina uzorka) / standardna devijacija

A ako hoćemo da **konvertujemo z skorove natrag u sirove vrednosti varijable**, sve što treba da **uradimo je da pomnožimo z skor standardnom devijacijom, a da onda dodamo aritmetičku sredinu uzorka**:

Sirova vrednost varijable = z skor * standardna devijacija + aritmetička sredina

U poređenju sa sirovim vrednostima varijable i sa percentilima, kao pokazateljima pozicije na distribuciji, z skorovi imaju veći broj prednosti:

- **Z skorovi su na intervalnom nivou merenja** (za razliku od percentila);
- **Z skorovi nemaju jedinicu mere**, tj. jedinica mere se gubi prilikom transformacije na z skalu u trenutku kada se deli sa standardnom devijacijom (koja je u istim mernim jedinicama u kojima su i sirovi skorovi), tako da se jedinice mere u vrednosti varijable i u standardnoj devijaciji međusobno potiru. To što z skorovi nemaju jedinicu mere omogućava nam da **smisleno poredimo z skorove različitih varijabli, varijabli čiji su sirovi skorovi u različitim mernim jedinicama** (npr. jedna je u kilogramima, a druga u metrima).
- **Z skor pokazuje veličinu razlike od aritmetičke sredine uzorka, izražene u standardnim devijacijama**. **Jedan z skor je jedna standardna devijacija**. Z skor **pokazuje koliko standardnih devijacija se vrednost entiteta nalazi ispod ili iznad aritmetičke sredine uzorka**. Na primer, z skor od +2 znači da se vrednost entiteta nalazi 2 standardne devijacije iznad aritmetičke sredine uzorka.
- Ako je distribucija normalna (ili uopšte ako je njen oblik poznat), **možemo precizno odrediti procenat entiteta u uzorku čije su vrednosti manje od određenog z skora tj. možemo pretvoriti bilo koji z skor u percentile na normalnoj** (ili bilo kojoj teorijskoj) **distribuciji.** Prema tome, **z skorovi su kao percentili, samo na intervalnom nivou merenja**.

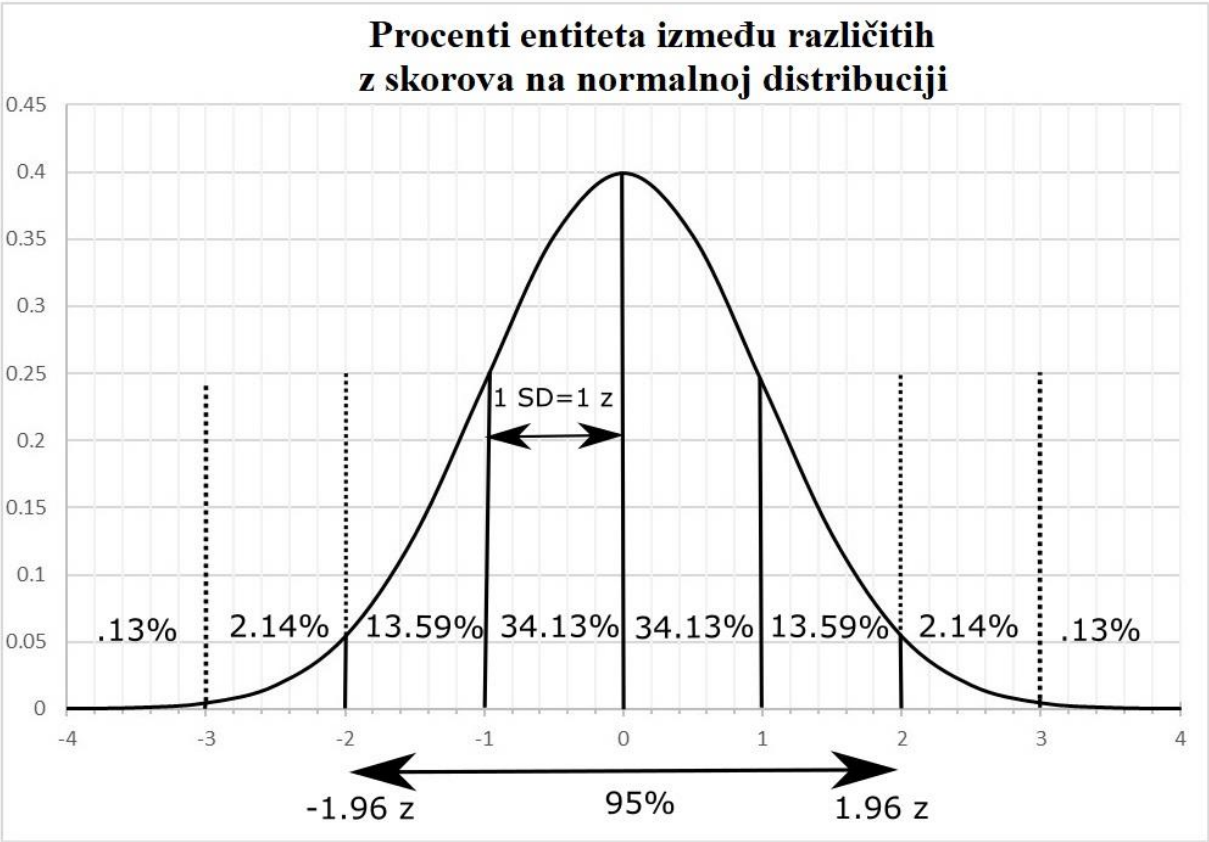
PREPRINT

- Z skorovi se lako čitaju – **pozitivne vrednosti znače da je vrednost entiteta iznad proseka tj. aritmetičke sredine, negativne vrednosti znače da je ispod proseka. Prosek tj. aritmetička sredina je 0.** Za razliku od sirovih vrednosti varijable, nije neophodno da tražimo koliko je aritmetička sredina uzorka da bi mogli da ustanovimo da li je određena vrednost ispod ili iznad te aritmetičke sredine i koliko se razlikuje od nje.

Z skorovi su takođe **važne komponente različitih složenijih statistika** (neki od najosnovnijih ovakvih statistika su predstavljeni u kasnijem delu ove knjige). Oni su zapravo jedna od ključnih komponenata mnogih složenih statističkih postupaka. Valja takođe primetiti da **interval između z skorova -1,96 i + 1,96 sadrži 95% entiteta iz normalno distribuiranog uzorka, a da interval između -2,56 i + 2,56 sadrži 99% entiteta iz takvog uzorka**.

Slika 4.12. Procenti entiteta na normalnoj distribuciji između različitih z skorova. Jedinice na horizontalnoj osi su z skorovi tj. standardne devijacije. Brojevi na vertikalnoj osi su verovatnoće. Sa slike možemo videti da, na normalnoj distribuciji, 34,13% entiteta leži između aritmetičke sredine (z skora 0) i -1 z (tačke koja je 1 standardnu devijaciju ispod aritmetičke sredine). Takođe, 13,59% entiteta leži između z skorova 1 i 2, a isti procenat entiteta je između z skorova -1 i -2. Važno je primetiti da se 95% entiteta nalazi između z skorova -1,96 i 1,96. Brojevi za interval od 99% entiteta (nije prikazan) su -2,56 i 2,56 z. Možemo dobiti i druge procente dodajući i oduzimajući procente između ovih intervala. Na primer, između z skorova -1 i 1, nalazi se 34.13%+34.13=68.26% svih entiteta tj. nešto iznad dve trećine celog uzorka (ili populacije, u zavisnosti od toga šta razmatramo).

PREPRINT



4.8. Ipsatizacija.

Postupak standardizacije koji je gore opisan odnosi se na situacije u kojima se vrednosti svih entiteta na određenoj varijabli prevode u z skorove. Kada se takav postupak sprovede, kažemo da je data varijabla prevedena na z skalu ili da je standardizovana. Međutim, isto tako je moguće sprovesti postupak standardizacije vrednosti jednog entiteta na svim merenim varijablama ili na odabranoj grupi varijabli. Naravno, da bi takav postupak bio smislen, vrednosti ovih različitih varijabli moraju da budu uporedive. Ovaj postupak se zove ipsatizacija. Drugim rečima, **ipsatizacija je postupak u kom se vrednosti pojedinačnog entiteta na nizu (uporedivih varijabli) standardizuju tj. prevode na z skalu**. Nakon što se uradi ipsatizacija**, taj pojedinačni entitet, čije vrednosti su ipsatizovane, će imate prosečnu vrednost 0 na grupi varijabli koje su bile uključene u postupak ipsatizacije, a standardna devijacija njegovih vrednosti na tim varijablama će biti 1.** Postupak ipsatizacije je identičan postupku standardizacije, samo je razlika u tome što se standardizacija radi na vrednostima svih entiteta na jednoj varijabli, dok se

ipsatizacija radi na vrednostima svih varijabli (ili grupe odabranih varijabli) na jednom entitetu.

Kada se ipsatizacija može upotrebiti u praksi? **Nekada može da se desi da nam nije važno kakvu tačno vrednost entitet ima na nekoj određenoj varijabli, već da nam je važno kakva je ta vrednost u odnosu na vrednosti koje taj entitet ima na drugim uporedivim varijablama**. Na primer, kada se proučavaju stavovi prema određenoj temi, ipsatizacija **može nekada pomoći da se ponište efekti stilova odgovaranja ispitanika.** Zamislimo na primer da smo tražili od grupe učesnika u istraživanju da ocene grupu avio kompanija na skali od 1 do 5 prema tome koliko su zadovoljni njima. Zamislimo sada da su neki od tih učesnika u istraživanju zaista uživali u iskustvima leta sa tim kompanijama. Oni će onda svim tim avio kompanijama dati četvorke i petice tj. oceniće ih visoko. Zamislimo sada da u uzorku ima i ljudi koji u principu mrze letenje, koji su imali baš nezadovoljavajuća iskustva sa ovim kompanijama ili koji su prosto veoma zahtevni ljudi. Oni će verovatno avio kompanijama davati jedinice i dvojke. Sada možemo da se zapitamo – da li ocena 4 dobijena od nekog ko je svim ostalim avio kompanijama dao ocenu 5 zaista pokazuje da on/ona ima pozitivniji stav prema toj konkretnoj avio kompaniji od nekog ko je toj kompaniji dao ocenu 2, ali je svim ostalim kompanijama dao ocenu 1? A to pitanje možemo da postavimo i na još praktičniji način – kada te osobe budu morale da lete avionom, ko će radije izabrati datu aviokompaniju – onaj koji joj je dao dvojku, ali svima ostalima jedinice ili onaj koji joj je dao 4, a svima ostalima petice? Odgovor je verovatno da će tu avio kompaniju izabrati ovaj prvi, uprkos činjenici da joj je dao nižu ocenu (dao joj je 2) nego ovaj drugi učesnik u istraživanju (koji joj je dao 4).

Tabela 4.13. Primer 12 učesnika u istraživanju koji su ocenjivali kvalitet 5 različitih avio kompanija (izmišljeni podaci). Možemo primetiti da, iako su izvorni odgovori učesnika u ovom istraživanju različiti, nakon ipsatizacije, njih pet (boldirani) imaju iste odgovore. Ovo je zbog toga što postupak ipsatizacije pretvara sirove skorove u z skorove koji ukazuju na relativnu poziciju odgovora ispitanika u odnosu na sve ostale odgovore tog ispitanika. Kako su svi ovi ispitanici dali jednake ocene za 4 avio kompanije, a petoj dali nižu ocenu nego ostalima, ovo je rezultiralo jednakim pozicijama na distribucijama odgovora bez obzira na konkretne ocene koje su odabrali. Takođe treba primentiti i da onaj jedan ispitanik koji je dao petice svim avio kompanijama, sada ima nule na svima odgovorima, zbog toga što su njegovi odgovori konstanta, dakle jednaki na

PREPRINT

svim pitanjima. Valja primetiti i da, iako su prosečne ocene avio kompanija koje su dali ispitanici različite, nakon ipsatizacije, sve prosečne ocene pojedinačnog ispitanika su izjednačene tako da su 0. To je zbog toga što ipsatizacija standardizuje vrednosti ispitanika na različitim varijablama, dovodeći do toga da je aritmetička sredina svakog ispitanika na različitim varijablama uključenim u ipsatizaciju 0 (a standardna devijacija je 1).

Kako učesnici u istraživanju ocenjuju avio kompanije, sirovi podaci Ocene sun a skali od 1 do 5, pri čemu je 1 najgora ocean, a 5 najbolja						
Ime učesnika u istraživanju	Avio kompanija A	Avio kompanija B	Avio kompanija C	Avio kompanija D	Avio kompanija E	Aritmetička sredina
Leposava	3	5	4	3	2	3.4
Anita	5	5	4	5	5	4.8
Vladislava	4	2	5	5	2	3.6
Maida	2	2	1	2	2	1.8
Marko	5	5	5	5	5	5
Radoslav	1	5	2	5	4	3.4
Filip	5	4	4	5	4	4.4
Vladimir	3	1	3	5	4	3.2
Jovan	1	5	1	5	1	2.6
Goran	5	5	1	5	5	4.2
Ilona	5	5	3	5	5	4.6
Petar	4	4	2	4	4	3.6
Kako učesnici u istraživanju ocenjuju avio kompanije, ipsatizovani podaci						
Ime učesnika u istraživanju	Avio kompanija A	Avio kompanija B	Avio kompanija C	Avio kompanija D	Avio kompanija E	Aritmetička sredina
Leposava	-.35	1.40	.53	-.35	-1.23	0.0
Anita	**.45**	**.45**	**-1.79**	**.45**	**.45**	**0.0**
Vladislava	.26	-1.06	.92	.92	-1.06	0.0
Maida	**.45**	**.45**	**-1.79**	**.45**	**.45**	**0.0**
Marko	*.00*	*.00*	*.00*	*.00*	*.00*	*0.0*
Radoslav	-1.32	.88	-.77	.88	.33	0.0
Filip	1.10	-.73	-.73	1.10	-.73	0.0
Vladimir	-.13	-1.48	-.13	1.21	.54	0.0
Jovan	-.73	1.10	-.73	1.10	-.73	0.0
Goran	**.45**	**.45**	**-1.79**	**.45**	**.45**	**0.0**
Ilona	**.45**	**.45**	**-1.79**	**.45**	**.45**	**0.0**
Petar	**.45**	**.45**	**-1.79**	**.45**	**.45**	**0.0**

PREPRINT

Pored ovoga, **postoje i različiti psihološki testovi kod kojih je princip ipsatizacije ugrađen u postupke za računanje ukupnih skorova**. Kod takvih testova **suma skorova na svim varijablama koje test meri je fiksna, a vrednosti pojedinačnih varijabli mogu da variraju unutar tog opšteg ograničenja** (e.g. Plutchik, 1989; Plutchik & Kellerman, 1974). Ovakvi testovi se zovu **ipsativni testovi.** Postoje i drugi primeri, ali bez obzira, valja primetiti da, dok je standardizacija rutinska procedura, koja se praktično koristi svuda, **primena ipsatizacije se mnogo ređe sreće u praksi.**

4.9. Hajde da primenimo šta smo do sada naučili!

Hajde da probamo sada da primenimo stvari koje smo predstavili u ovom poglavlju kroz nekoliko vežbi. Molimo vas da pogledate opšte uputstvo za ovakve vežbe koje možete naći na početku knjige. Naša preporuka je da prvo pročitate svaki isečak i tvrdnje date u njemu i da onda date svoj odgovor. Odgovor možete upisati u kolonu za odgovore, a posle toga pročitajte odgovore i uporedite svoje odgovore sa njima.

Vežba F. Teorijske i empirijske distribucije, odstupanja od normalne distribucije, standardizacija i ipsatizacija

PREPRINT

Jelena radi kao konsultant u fabrici za preradu voća. Fabrika trenutno kupuje i oprema novu hladnjaču u kojoj će se raditi sortiranje, pakovanje i skladištenje malina. Jelena treba da odluči koju od dve linije za sortiranje malina koje su u ponudi kompanija treba da kupi.

Prva linija svakom radniku koji radi na sortiranju izlaže jednu po jednu mailnu, a radnik onda treba brzo da je pregleda i skloni sa linije ako nije dobra. Radnik prosečne pažljivosti koji radi na ovoj liniji ima 70% šanse da uoči lošu malinu onda kada mu/joj bude izložena (tj. kada loša malina bude izložena radniku, radnik ima 70% šanse da uoči da je ta malina loša). Tokom jednog sata, ova linija izloži 5000 malina svakom radniku od kojih su tipično 1000 loše.

Druga linija koristi široku pokretnu traku koja se polako pomera i na kojoj su rasprostrte maline. Radnici stoje pored pokretne trake i prebiraju po malinama koje prolaze pokušavajući da prepoznaju one koje su loše. Radnik tipične pažljivosti će, u proseku, uočiti 15 loših malina po minutu. U jednom satu, 5000 malina prođe pored svakog radnika, od kojih su 1000 loše.

Predstavljeni podaci su validni za radnike prosečne pažljivosti. Međutim, Jelena je utvrdila da postoje izražene individualne razlike između radnika u pogledu njihove pažljivosti.

Prilikom odgovaranja na tvrdnje treba smatrati da sve varijable koje se u tekstu pominju ili koje se mogu izvesti iz podataka imaju distribucije koje su jednake idealnim teorijskim distribucijama koje se očekuju za takve varijable u situacijama u kojima su dobijene. U jednom satu ima 60 intervala od jednog minuta. Na svakoj liniji svaku malinu pregleda samo jedan radnik. Nema situacija u kojima ista malina biva izložena ili prolazi pored više od jednog radnika. Takođe, pretpostaviti da nema situacija kada radnici dobre maline pogrešno identifikuju kao loše, tu opciju ne treba razmatrati prilikom odgovaranja.

F	**Tvrdnja:**	**Odgovor**
F1.	Pažljivost radnika ženskog pola ima oblik Hongove distribucije.	
F2.	Ako bi svi radnici bili prosečno pažljivi, distribucija broja uočenih loših malina po radniku na prvoj liniji bi imala oblik Puasonove distribucije.	
F3.	Prva linija je efikasnija u eliminisanju loših malina od druge linije (eliminiše veći procenat loših malina).	
F4.	Druga linija obrađuje više malina na sat nego prva linija.	
F5.	Pažljivost radnika je uniformno distribuirana (ima oblik uniformne distribucije).	
F6.	Ako bi svi radnici imali prosečnu pažljivost, prosečan broj uočenih loših malina na sat po radniku na prvoj liniji bi bio veći od 650.	
F7.	Ako bi svi radnici imali prosečnu pažljivost, prosečan broj uočenih loših malina po radniku na drugoj liniji bi bio veći od 650.	
F8.	Ako bi pratili rad grupe radnika prosečne pažljivosti tokom dva sata rada na drugoj liniji, distribucija broja uočenih loših malina po radniku bi imala oblik uniformne distribucije.	
F9.	Distribucija broja uočenih loših malina po radniku na prvoj liniji, nakon što je samo jedna malina izložena svakom od njih, bi imala oblik Bernulijeve distribucije (treba pretpostaviti da svi radnici imaju prosečnu pažljivost).	
F10.	Mlađi radnici su u proseku pažljiviji nego stariji radnici.	

PREPRINT

Vežba G. Teorijske i empirijske distribucije, odstupanja od normalne distribucije, standardizacija i ipsatizacija (Popov et al., 2021)

Table 1
Descriptive statistics for variables in the study

	Theoretical range	Achieved range	*M*	*SD*	*Skewness*	*Kurtosis*
Prior physical activity (GLTEQ)	0–119	0–119	38.83	28.88	.97	.48
Avoidant coping (Brief COPE)	12–48	12–43	21.40	4.90	.74	1.03
Problem focused coping (Brief COPE)	12–48	12–48	33.17	7.45	-.36	-.18
Emotion focused coping (Brief COPE)	6–24	6–24	15.48	4.24	-.22	-.02
Current physical exercise	1–4	1–4	2.57	.97	-.41	-.89
Depression (DASS–21)	0–21	0–21	4.07	4.87	1.50	1.75
Anxiety (DASS–21)	0–21	0–21	3.11	4.49	1.81	2.81
Stress (DASS–21)	0–21	0–21	7.38	5.64	.53	-.59

Značenje termina: Depression – Depresivnost, Anxiety – Anksioznost, Prior physical activity – Prethodna fizička aktivnost, Current physical exercise – Trenutno fizičko vežbanje, Avoidant coping – Izbegavajuće suočavanje sa streson, Stress – Stres, Skewness – skjunes, Kurtosis – kurtozis, Achieved range – dobijeni raspon skorova na ovom uzorku (najmanja i najveća dobijena vrednost), Theoretical range – teorijski raspon vrednosti varijable (najviša i najniža moguća vrednost na datoj varijabli bez obzira da li u uzorku ima entiteta sa tim vrednostima ili ne).

Tabela preštampana iz: Popov, S., Sokić, J., & Stupar, D. (2021). Activity Matters: Physical Exercise and Stress Coping during COVID-19 State of Emergency. Psihologija, 54(3), 307–322. https://doi.org/10.2298/psi200804002p . Preštampano na osnovu dozvole autora.

G	Tvrdnja:	Odgovor
G1.	Distribuciija Depresivnosti je pozitivno asimetrična	
G2.	Distribucija Anksioznosti je negativno asimetrična.	
G3.	Distribucija Anksioznosti je leptotkurtična.	
G4.	Aritmetička sredina Prethodne fizičke aktivnosti je niža od medijane ove varijable.	
G5.	Aritmetička sredina Trenutnog fizičkog vežbanja je niža od medijane ove varijable.	
G6.	Distribucija Anksioznosti je šiljata, tj. učesnici u istraživanju su koncentrisaniji (oko sredine) nego što je to slučaj kod normalne distribucije.	
G7.	Distribucija Stresa je platikurtična.	
G8.	Učesnici u istraživanju su međusobno razređeniji na Trenutnom fizičkom vežbanju nego što bi to bio slučaj na normalnoj distribuciji.	

PREPRINT

G9.	50i percentil varijable Izbegavajuće suočavanje sa stresom je veći od 21,40.	
G10.	80i percentil varijable Stres je veći od 22.	

Vežba H. Teorijske i empirijske distribucije, odstupanja od normalne distribucije, standardizacija i ipsatizacija (V. Hedrih, 2011; Holland, 1959)

N=360	R	I	A	S	E	C
Skjunes	1.38	-.44	.225	-.01	.30	1.51
Kurtozis	1.88	-.29	-.91	-.80	-.77	3.01
25i percentil	.18	.61	.69	.96	.69	.36
50i percentil	.44	1.05	1.21	1.45	1.19	.58
75i percentil	.80	1.50	1.82	1.99	1.76	.87

Tabela je napravljena na osnovu podataka korišćenih u Hedrih (2011).
R, I, A, S, E i C su mere tipova profesionalnih interesovanja iz Holandove teorije profesionalnih interesovanja (Holland, 1959)

H	**Tvrdnja:**	**Odgovor**
H1.	Ako osoba iz ovog uzorka ima skor 1,5 na varijabli I, njegov/njen z skor na toj varijabli bi bio pozitivan.	
H2.	U uzorku nema ljudi sa skorom većim od 2 na varijabli E.	
H3.	U uzorku ima više od 400 ispitanika.	
H4.	Distribucija varijable S je simetrična i platikurtična.	
H5.	Distribucija varijable C je pozitivno asimetrična i platikurtična.	
H6.	Distribucija varijable R ima oblik Hansenove distribucije.	
H7.	Medijanski skorovi kurtozisa su viši na 25om nego na 75om percentilu.	
H8.	Aritmetička sredina varijable S je manja od 1,6.	
H9.	Ako osoba iz ovog uzorka ima skor 1 na varijabli S, njegov/njen z skor na ovoj varijabli bi bio pozitivan.	
H10.	Najniži skor na varijabli A je 1.	

Pogledajmo sada odgovore:

F1 – besmisleno. Hongova distribucija ne postoji.

F2 – netačno. Ne, funkcionisanje prve linije opisano je u terminima verovatnoće javljanja događaja (događaj je uočavanje loše maline) po pokušaju (izlaganje loše maline) i ovo je postavka za binomnu distribuciju. Dakle, binomina distribucija bi bila očekivana distribucija u ovom slučaju.

PREPRINT

F3 – netačno. Obe linije obrađuju 5000 malina na sat po radniku i 1000 malina od tih 5000 je loše. Sa prvom linijom, radnik ima 70% šanse da uoči lošu malinu kada mu bude izložena. To znači da će od 1000 loših marlina koje budu izložene radnik u proseku uočiti 70%, što je 700. S druge strane, sa drugom linijom radnik će u proseku uočavati 15 loših malina u minutu. U jednom satu ima 60 minuta. To znači da će radnik u proseku uočiti 15x60 loših malina na sat, što je 900. 900 je više od 700, što znači da je druga linija efikasnija jer eliminiše 90% loših malina u odnosu na prvu liniju koja eliminiše svega 70%.

F4 – netačno. Kao što je već objašnjeno u F3, obe linije obrađuju jednaku količinu malina na sat.

F5 – netačno. Pažljivost (radnika) je crta individualnih razlika i prema tome tu očekujemo da distribucija bude normalna, a ne uniformna.

F6 – tačno. U F5 smo već objasnili računicu da je to 700 malina u proseku. 700 je više od 650, što znači da je tvrdnja tačna.

F7 – tačno. Rad druge linije, kako je objašnjeno u F3, rezultira u 900 uočenih loših malina na sat po radniku. 900 je više od 650, što ovu tvrdnju čini tačnom.

F8 – netačno. Rad druge linije je objašnjen u terminima prosečnog broja loših malina po minutu rada i to je postavka za Puasonovu distribuciju, a ne za uniformnu.

F9 – tačno. Rad prve linije je postavka za binomnu distribuciju. Bernulijeva distribucija je poseban slučaj binomne distribucije kada je broj pokušaja 1. Prema tome, tvrdnja je tačna.

F10 – nepoznato. Iako je to svakako moguće, nema podataka u tekstu na osnovu kojih bi mogli da zaključimo da li je tačno ili nije. Prema tome, ne znamo.

G1 – tačno. Skjunes Depresivnosti je 1,5, dakle pozitivan. Pozitivan skjunes pokazuje da je distribucija pozitivno asimetrična.

G2 – netačno. Skjunes Anksioznosti je 1,81, dakle pozitivan. Pozitivan skjunes pokazuje da je distribucija pozitivno asimetrična, a ne negativno.

G3 – tačno. Leptokurtične distribucije imaju pozitivan kurtozis. Za Anksioznost, to je 2,81, dakle pozitivan, što znači da je distribucija zaista leptokurtična.

PREPRINT

G4 – netačno. Možemo videti da Prethodna fizička aktivnost ima pozitivan skjunes, što znači da je aritmetička sredina viša od medijane.

G5 – tračno. Možemo videti da Trenutno fizičko vežbanje ima negativan skjunes, što znači da je na toj varijabli aritmetička sredina niža od medijane.

G6 – verovatno tačno. Ova tvrdnja je donekle problematična. S jedne strane, možemo videti da Anksioznost ima dosta visoku vrednost kurtozisa, što znači da je leptokurtična. S druge strane, znamo da visok kurtozis može biti i posledica autlajera, a ne samo ekstremne koncentracije entiteta na sredini. Međutim, ako pogledamo dobijeni raspon skorova (raspon skorova u uzorku, predstavljen kao raspon od najveće do najmanje vrednosti u tabeli – Achieved range) možemo da vidimo da je aritmetička sredina manje od 1 standardne devijacije iznad donje granice ovog raspona, što znači da je veliki deo uzorka koncentrisan u relativno malom intervalu između donje granice ovog raspona i aritmetičke sredine. Možemo zaključiti iz ovog da zaista postoji ekstremna koncentracija ispitanika u oblasti niskih skorova, iako verovatno postoje i debelo rame i rep distribucije usmereni ka pozitivnoj strani. Tvrdnja je, prema tome, verovatno tačna.

G7 – tačno. Kurtozis Stresa je -0,59 tj. negativan, što znači da je distribucija platikurtična.

G8 – tačno. „Međusobno razređeniji“ implicira platikurtičnu distribuciju, a ovo znači da je kurtozis negativan. Kurtozis je -0,89, dakle negativan, što znači da je tvrdnja tačna.

G9 – netačno. 50i percentil je medijana, skjunes ove varijable je pozitivan, što znači da je medijana manja od aritmetičke sredine. Aritmetička sredina je 21,40 a medijana mora onda biti manja od toga, iz čega sledi da je tvrdnja netačna.

G10 – netačno. Iako nemamo podatke za baš 80i percentil, možemo videti da su gornje granice i teorijskog i dobijenog raspona Stresa niže od 21. To znači da nema vrednosti većih od 21, što znači da 80i percentil ne može biti veći od 22.

H1 – tačno. Možemo videti da je skjunes varijable i pozitivan, što znači da je aritmetička sredina ispod medijane. Medijana je 50i percentil, što je 1,05 u ovom slučaju, a aritmetička sredina onda mora da bude manja od toga. Svi z skorovi iznad aritmetičke sredine su pozitivni, što znači i da skor 1,5 mora da odgovara pozitivnoj z vrednosti.

PREPRINT

H2 – nepoznato. Možemo videti da je 75i percentil na ovoj varijabli 1,76, ali ne znamo koliko visoko iznad toga ima skorova. Moguće je da ima entiteta sa skorom 2, a moguće i da nema, to se iz tabele ne vidi.

H3 – netačno. Piše N=360. Sa N se obično označava broj entiteta u uzorku i to je tako označeno i ovde. Dakle, u uzorku ima 360 entiteta, što je manje od 400.

H4 – tačno. Skjunes varijable S je praktično 0, što znači da je distribucija simetrična. Negativan kurtozis pokazuje da je distribucija platikurtična.

H5 – netačno. Pozitivan skjunes pokazuje da je distribucija zaista pozitivno asimetrična, međutim pozitivan kurtozis pokazuje da nije platikurtična, već leptokurtična.

H6 – besmisleno. Ne postoji nikakva „Hansenova distribucija".

H7 – besmisleno. Ne postoje nikakvi „medijanski skorovi kurtozisa", a nije jasno ni šta bi moglo da znači to da su viši na 25om, nego na 75om percentilu. Cela tvrdnja je besmislena.

H8 – tačno. Varijabla S ima simetričnu distribuciju, što znači da aritmetička sredina i medijana tj. 50i percentil imaju istu vrednost. Vrednost medijane je 1,45, što je niže od 1,6.

H9 – netačno. Možemo videti da je distribucija varijable S simetrična i da ima skjunes 0, što znači da njena aritmetička sredina ima istu vrednost kao 50i percentil tj. medijana. Kako je medijana 1,45, to znači da je vrednost 1 ispod vrednosti aritmetičke sredine, te tako odgovara negativnom z skoru.

H10 – netačno. Možemo videti da je 25i percentil 0,69, što je niže od 1. To znači onda da 1 ne može biti najniži skor na ovoj varijabli.

PREPRINT

Poglavlje 5. Statistika zaključivanja, osnovni pojmovi

Apstrakt. Poglavlje o osnovnim pojmovima statistike zaključivanja počinje predstavljanjem opšte ideje statistike zaključivanja koju prati predstavljanje ključne teoreme u ovoj oblasti – centralne granične teoreme. Prvi deo poglavlja predstavlja centralnu graničnu teoremu i to kako se ova koristi za procenu parametara populacije. Predstavljen je postupak procene parametara preko butstrepinga. Dve metode procene parametara – tačkasta procena i intervalna procena su predstavljenje u okviru oba pomenuta postupka procene parametara populacije, a opisani su i pojmovi distribucije uzorkovanja i standardne greške. U sledećem delu se diskutuje pojam nulte hipoteze i statističke značajnosti i ovaj deo teksta upoznaje čitaoca sa testiranjem nulte hipoteze preko postupaka zasnovanih na centralnoj graničnoj teoremi, butstrepingu i korišćenjem Bajesovog faktora i Bajesijanske interpretacije verovatnoće. Poslednji deo ovog poglavja posvećen je predstavljanju pojmova parametrijskog i neparametrijskog statističkog postupka.

Ključne reči: procena parametara, centralna granična teorema, butstreping, Bajesov faktor, parametrijski i neparametrijski postupci

Statistika zaključivanja je naziv za skup metoda koje se koriste da bi se izveli zaključci o svojstvima populacije na osnovu statistika izračunatih na uzorku tj. da se izvedu zaključci o vrednostima parametara na osnovu vrednosti statistika. U prethodnom delu ove knjige predstavili smo različite postupke za opisivanje uzorka, kako za opisivanje celog uzorka, tako i za opisivanje vrednosti pojedinačnih entiteta u poređenju sa ostatkom uzorka. Predstavili smo i postupke sa prikupljanje uzorka. Ali, većinu vremena, **smisao korišćenja statističkih postupaka nije da se otkriju stvari o nekom konkretnom uzorku, nego da se taj uzorak iskoristi da bi se izveli zaključci o populaciji.** Važna stvar u vezi ovoga je to što znamo da nikad nije potpuno sigurno da će uzorak uzet iz populacije, čak i ako primenimo najbolji mogući postupak uzorkovanja, biti skroz reprezentativan tj. ne možemo garantovati da će sva svojstva uzorka biti potpuno ista kao svojstva populacije. Upravo suprotno, **vrlo je verovatno da će se svojstva uzorka bar donekle razlikovati od svojstava populacije**. To je razlog zašto, **kada izvodimo zaključke o populaciji, ova mogućnost da svojstva uzorka budu manje ili više različita od svojstava populacije mora**

PREPRINT

da se uzme u obzir. To je jedan od razloga zašto je **važno povući jasnu razliku između statističkih pokazatelja izračunatih na uzorku i tih istih pokazatelja u populaciji. Ovi prvi, računati na uzorku, se zovu statistici, a ovi drugi, koji se odnose na populaciju, se zovu parametri**.

U trenutku pisanja ove knjige, dva su pristupa proceni parametara na osnovu vrednosti statistika u najčešćoj upotrebi: pristup zasnovan na centralnoj graničnoj teoremi i pristup zasnovan na upotrebi reuzorkovanja, pre svega na upotrebi butstrepinga.

5.1. Centralna granična teorema

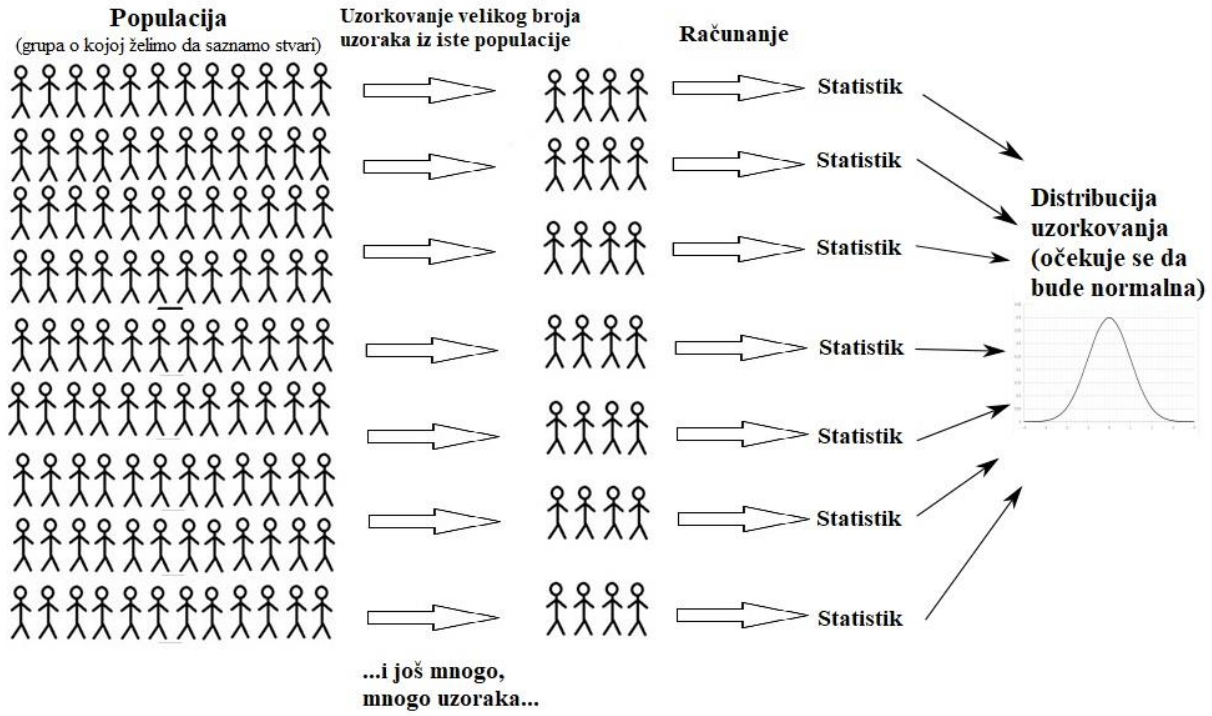
Centralna granična teorema govori o tome šta se dešava kada uzmemo više uzoraka iz iste populacije. To je zapravo tipična situacija u kojoj se proveravaju istraživački nalazi – u jednom istraživanju istraživači uzmu uzorak iz populacije koju žele da prouče i objave svoje nalaze, a onda, kasnije, u drugom istraživanju, istraživači uzmu uzorak iz iste populacije i objave nalaze o tome da li se nalazi koje su oni dobili slažu sa nalazima ovog prvog istraživanja.

Centralna granična teorema kaže da ako uzmemo veliki broj slučajnih uzoraka iz populacije i onda izračunamo isti statistik iz svakog od tih uzoraka, distribucija vrednosti tog statistika na svim tim uzorcima će imati oblik normalne distribucije. Uz ovo se takođe pretpostavlja i da je **populacija iz koje se radi uzorkovanje neograničena ili veoma velika u odnosu na uzorak** (tako da je praktično situacija ista kao da je neograničena) ili da se uzorkovanje radi sa vraćanjem (pa je tako verovatnoća svakog pojedinačnog entiteta da bude uzet u uzorak ostaje ista tokom celog procesa uzorkovanja). Na primer, ovo znači da bi mogli da uzmemo, na primer, 2000 slučajnih uzoraka iz iste populacije, izmerimo onda određenu varijablu na svakom uzorku i da onda izračunano odgovarajući statistik iz svakog od tih uzoraka. Taj statistik bi, na primer, mogla da bude aritmetička sredina varijable koja se meri. Mi bi onda izračunali aritmetičku sredinu date varijable na svakom od ovih 2000 slučajnih uzoraka. Potom bi pogledali distribuciju koju pravi ovih 2000 aritmetičkih sredina (1 aritmetička sredina iz svakog od 2000 uzoraka znači da imamo ukupno 2000 aritmetičkih sredina!) i otkrili bi, prema centralnoj graničnoj teoremi, da ta distribucija ima oblik normalne distribucije.

PREPRINT

Distribucija vrednosti (istih) statistika izračunatih iz ovakvog velikog broja slučajnih uzoraka (uzetih iz iste populacije) naziva se distribucija uzorkovanja. Prema centralnoj graničnoj teoremi, **aritmetička sredina distribucije uzorkovanja je jednaka parametru.**

Slika 5.1. Grafički prikaz ključnih postavki centralne granične teoreme. Veliki broj uzoraka se uzima iz ista populacije i isti statistik se računa iz svakog od njih. Distribucijastatistika dobijenih na ovaj način se zove distribucija uzorkovanja i teorija kaže da će ta distribucija imati oblik normalne distribucije i da će joj aritmetička sredina biti jednaka parametru. Standardna devijacija distribucije uzorkovanja dobijene na ovaj način se zove standardna greška.



Iz pretpostavke ove teorije da je aritmetička sredina distribucije uzorkovanja jednaka parametru sledi da **iako statistici pojedinačnih uzoraka mogu više ili manje da se razlikuju od parametra, oni će i dalje težiti da se grupišu oko njega.** A **parametar je ujedno i centar distribucije uzorkovanja** imajući u vidu da je parametar aritmetička sredina, a da je distribucija uzorkovanja normalna i, prema tome, simetrična. **Što se statistik nekog uzorka više razlikuje od parametra, to je manje verovatno da se takav uzorak dobije slučajnim uzorkovanjem**. Takođe, **parametar je na distribuciji uzorkovanja lociran u tački sa najvećom verovatnoćom.**

Drugim rečima, **ako bismo morali da pogađamo gde se na distribuciji uzorkovanja nalaze statistici različitih slučajnih uzoraka koje smo slučajnim uzorkovanjem uzeli iz populacije, ukupno bi nam greške bile najmanje ako bismo pretpostavili da se nalaze u centru, tj. da je statistik čiju lokaciju pogađamo jednak parametru**. Naravno, **pogađajući na takav način, i dalje bismo češće bili u krivu nego u pravu, ali bi sa ovakvim pretpostavkama naše greške bile manje, nego što bi bile da smo pretpostavili da se statistik koji pogađamo nalazi na bilo kojoj drugoj lokaciji na distribuciji**. Ovo zato što je centar distribucije tačka sa najvećom verovatnoćom tj. tačka na distribuciji oko koje su entiteti najgušće skoncentrisani. Ta tačka je takođe i tačka sa najmanjom prosečnom udaljenošću od drugih tačaka na distribuciji (zato što je u centru), tako da bi **prosečna veličina naše greške predviđanja bila najmanja ako bi pretpostavili da se statistik uzorka, čija pozicija na distribuciji uzorkovanja nam nije poznata, nalazi u centru distribucije i da je jednak parametru**.

Pored aritmetičke sredine tj. mere centralne tendencije, opis distribucije zahteva navođenje i neke mere varijabilnosti. Ta mera varijabilnosti je u ovom slučaju standardna devijacija. **Standardna devijacija distribucije uzorkovanja zove se standardna greška**. U osnovi, **standardna greška nam pokazuje koliko možemo da očekujemo da se statistici pojedinačnih uzoraka razlikuju od parametra** (vrednosti statističkog pokazatelja koji posmatramo na populaciji). Takođe, ako se vratimo na poglavlje o standardizaciji, videćemo da postoji jasan odnosi između odstupanja od aritmetičke sredine izraženog u standardnim devijacijama i pozicije na normalnoj distribuciji. Ovo znači da nam **veličina standardne greške govori i to koliko su verovatne različite veličine razlika (u vrednosti ispitivanog statističkog pokazatelja) između uzorka i populacije**. Međutim, da bi mogli da standardnu grešku koristimo na ovaj način, moramo prvo da je izračunamo. Već znamo da nam za računanje standardne devijacije trebaju pojedinačne vrednosti entiteta iz uzorka. U ovom slučaju, te pojedinačne vrednosti bi bile vrednosti sa velikog broja uzoraka uzetih iz iste populacije na način kako to postavlja centralna granična teorema, a to je nešto što nam zapravo nije dostupno u praksi – ono što u praktičnim situacijama tipično imamo je jedan uzorak koji smo skupili za potrebe istraživanja koje radimo i to je sve. Zbog ovog, **u okviru pristupa zasnovanog na centralnoj graničnoj teoremi, standardna greška se ne može stvarno izračunati iz podataka** (zato što imamo samo jedan uzorak skupljen za datu određenu studiju i sa kojim radimo naša računanja, a nemamo veliki broj uzoraka!), **već se procenjuje na osnovu formula**. Postoje posebne formule za računanje standardne greške za svaki statistički pokazatelj,

PREPRINT

a svima njima je zajedničko to **da veličina standardne greške zavisi od veličine uzorka**. **Što je veći uzorak, to je manja standardna greška**. Takođe, **na standardne greške jednog broja statistika utiče i standardna devijacija i to tako da što je veća standardna devijacija, to je veća i standardna greška.** Što je veća varijabilnost uzorka (a kada je veća varijabilnost uzorka, verovatno je da je veća i varijabilnost populacije) veće su i prilike da se pojedinačni uzorak razlikuje od parametra. Da navedemo očigledan primer, ako bi u populaciji svi bili visoki tačno po 2 metra, ne bi bilo nikakve šanse da se iz takve populacije uzme uzorak koji bi bio od nje različit. Kakav god uzorak uzeli iz takve populacije, u njemu bi uvek svi bili tačno 2 metra visoki! S druge strane, ako visine članova populacije variraju, itekako je moguće izabrati uzorak koji ima više niskih ljudi nego visokih i obrnuto. Na sličan način, ako slučajno biramo ljude iz populacije, ali, na primer, biramo samo jednu osobu, postoji određena verovatnoća da, čisto slučajno, izaberemo osobu koja je veoma visoka i koja, zbog toga, nije uopšte reprezentativna za tipične visine ljudi u populaciji. Međutim, ako bi nasumično uzimali iz populacije veći broj ljudi, verovatnoća da svi izabrani budu veoma, veoma visoki, a tako i veoma, veoma nereprezentativni za tipičnu visinu populacije je mnogo, mnogo manja (treba reći da nikada nije 0 ako su populacije neograničene i/ili se radi uzorkovanje sa vraćanjem). To je slična situacija kao ona sa osvajanjem premije na lutriji ili na Lotou. Iako uglavnom nije nemoguće naći osobu koja je osvojila premiju na Lotou (to je veoma redak događaj, ali s vremena na vreme, neko osvoji tu premiju), vrlo je verovatno da ne postoji osoba koja je osvojila 100 premija u nizu na istoj lutriji (što je događaj koji je toliko malo verovatan da se verovatno nikada neće desiti). Ovo je razlog zašto povećavanje veličine uzorka smanjuje standardnu grešku.

Formule za računanje standardnih grešaka različitih statistika se u principu izvode na osnovu postulata centralne granične teoreme i razlikuju se od statistika do statistika. Na primer, formula za **standardnu grešku aritmetičke sredine podrazumeva deljenje standardne devijacije kvadratnim korenom broja entiteta u uzorku** (e.g. Harding et al., 2014):

$$SG_{AS} = \frac{SD}{\sqrt{N}}$$

Standardna greška se tipično označava sa SG u srpskoj literaturi (odnosno sa SE u engleskoj), a uz nju se obično upisuje i ime statistika na koji se data standardna greška odnosi u vidu subskripta.

PREPRINT

Tako u ovoj formuli SG_{AS} označava standardnu grešku aritmetičke sredine. SD je standardna devijacija uzorka, a N je broj entiteta u uzorku.

Formula za računanje **standardne devijacije medijane** (Harding et al., 2014) je:

$$SG_{SD} = \frac{SD}{\sqrt{2(N-1)}}$$

Iz ove formule možemo videti da, **sa istom aritmetičkom sredinom i standardnom devijacijom, standardna greška standardne devijacije je manja od standardne greške aritmetičke sredine**. Imajući u vidu da je $1/\sqrt{2}$ otprilike 0,71, a da je ostatak formule praktično isti kao formula za standardnu grešku aritmetičke sredine, možemo zaključiti da je veličina standardne greške standardne devijacije za nijansu više od 70% standardne greške aritmetičke sredine.

Formule za **standardnu grešku skjunesa i kurtozisa** su donekle različite od formula koje su do sada predstavljene i ta razlika se sastoji u tome da se **ove standardne greške procenjuju samo na osnovu veličine uzorka**. **U njihovo računanje ne ulazi standardna devijacija uzorka**. To znači da će **vrednosti ovih statistika biti jednake za sve uzorke koji su jednake veličine** tj. sastoje se od jednakog broja entiteta. Standardna greška skjunesa se može proceniti preko formule (Harding et al., 2014):

$$SG_{skjunes} = \sqrt{\frac{6N(N-1)}{(N-2)(N+1)(N+3)}}$$

U ovoj formuli je N broj entiteta u uzorku tj. veličina uzorka. Standardna greška kurtozisa može se proceniti prema formuli (Harding et al., 2014):

$$SG_{Kurtozis} = 2 * SG_{Skjunes} * \sqrt{\frac{N^2-1}{(N-3)(N+5)}}$$

Još jedan važan statistik koji treba ovde pomenuti je **koeficijent** korelacije (Pirsonov koeficijent korelacije, predstavljen u kasnijem delu ove knjige). **Formula za procenu standardne greške**

PREPRINT

Pirsonovog koeficijenta korelacije (koji se tipično označava sa r) **takođe ne sadrži standardnu devijaciju, ali uključuje sam koeficijent korelacije** i izgleda ovako:

$$SG_r = \frac{(\sqrt{(1-r^2}}{(\sqrt{(N-2)}}$$

U ovoj formuli, r je veličina koeficijenta korelacije. Treba primetiti da je **u najčešćoj varijanti praktične primene ove formule,** o kojoj će biti reči i u kasnijem delu ove knjige**, koeficijent r jednak 0, tako da je onda standardna greška koeficijenta korelacije (ako zanemarimo ono -2 u formuili) jednaka recipročnoj vrednosti kvadratnog korena broja ispitanika u uzorku.**

Tokom celog prethodnog veka, ali i u vreme pisanja ove knjige, **pristup preko centralne granične teoreme je najšire korišćeni pristup u statistici zaključivanja**. On se primenjuje u skoro svim testovima i statističkim postupcima za izvođenje zaključaka o vrednostima populacije, za izvođenje zaključaka o razlikama između populacija, a pretpostavka o normalnoj distribuciji uzorkovanja se koristi čak i u situacijama kada se ocenjuju rezultati većeg broja istraživanja sprovedenih na istu temu da bi se izvukli zaključci o tome da li postoje osnovi za sumnju u neregularnosti u objavljivanju naučnih rezultata. Pretpostavka je da će, kada je sve u redu, statistici dobijeni u velikom broju istraživanja istog fenomena dati normalnu distribuciju vrednosti (istog) statistika računatog iz uzoraka različitih istraživanja. Neregularnost za koju se najčešće smatra da je odgovorna za situacije kada se ne dobije normalna distribucija je takozvani „efekat fioke“. **„Efekat fioke“ je pojava da se rezultati istraživanja koji se smatraju nepoželjnim ne objavljuju** (bilo zato što istraživači sami reše da ih ne objave, bilo zato što izdavači odbijaju da ih objave), **dok se poželjni rezultati objavljuju**. I onda se zbog toga primenjuju statistički postupci koji analiziraju oblik distribucije koju daju rezultati većeg broja istraživanja da bi ustanovili da li neki delovi distribucije uzorkovanja nedostaju. Odstupanje oblika distribucije statistika iz većeg broja istraživanja od oblika normalne distribucije se u ovakvim situacijama tretira kao pokazatelj mogućih takvih neregularnosti. Takođe, **praktično svi najpopularniji statistički softverski paketi u vreme pisanja ove knjige prvenstveno koriste postupke statistike zaključivanja koji se oslanjaju na pretpostavke centralne granične teoreme.**

5.2. Pristup zaključivanju o vrednostima parametara preko butstrepinga

PREPRINT

Uprkos širokoj upotrebi postupaka zasnovanih na **centralnoj graničnoj teoremi, jedan važan problem s njima je to što oni polaze od pretpostavke da je distribucija uzorkovanja oblika normalne distribucije, bez da to zaista i provere u svakom pojedinačnom slučaju primene ovih pretpostavki.** Drugim rečima, problem sa centralnom graničnom teoremom je to što se ona relativno retko empirijski proverava, a ipak se široko koristi. Zbog ovog nedostatka, različiti autori predlažu alternativne metode računanja standardne greške i izvođenja zaključaka o populaciji. Pregled nekih od ovih metoda mogu se naći u tekstu Efrona (Efron, 1981).

U trenutku pisanja ove knjige, najširu upotrebu u naučnoj zajednici i kod autora statističkog softvera (od postupaka koji nisu zasnovani na centralnoj graničnoj teoremi) ima postupak izvođenja zaključaka o populaciji koji je zasnovan na butstrepingu. **Prednost postupaka statistike zaključivanja koji su zasnovani na butstrepingu u odnosu na one koji se oslanjaju na centralnu graničnu teoremu je u tome što se ne oslanjaju ni na kakve pretpostavke o obliku distribucije uzorkovanja.** Umesto toga, procedura butstrepinga, onako kako se tipično koristi za ove svrhe, problem nepostojanja distribucije uzorkovanja rešava tako što uzima veliki broj uzoraka iz postojećeg uzorka (onog čiji se podaci obrađuju i iz kog želimo da izvedemo zaključke) umesto iz populacije. Dakle, **u postupku butstrepinga, veliki broj uzoraka** (npr. 10 000 uzoraka, ovaj broj je ograničen samo brzinom računara koji se koristi za ove postupke i spremnošću istraživača da čeka na rezultate) **se uzima sa vraćanjem iz uzorka na kom radimo analize.** Na ovaj način, zato što se uzorkovanje radi sa vraćanjem, moguće je uzeti neograničen broj uzoraka iz istog, ograničenog uzorka istraživanja koji analiziramo. **Uzorak koji obrađujemo u ovom postupku služi kao zamena za populaciju, dok se veliki broj uzoraka uzetih postupkom butstrepinga iz tog uzorka tretira kao zamena za veliki broj uzoraka uzetih iz populacije. Isti statistik se onda računa iz svakog od tih uzoraka stvorenih postupkom butstrepinga, a distribucija vrednosti tog statistika u uzorcima dobijenim butstrepingom se smatra za distribuciju uzorkovanja. Standardna devijacija ovakve distribucije uzorkovanja predstavlja standardnu grešku.** Obično se očekuje da statistik uzorka (uzorka koji analiziramo – onog iz kog smo stvorili ovaj veliki broj uzoraka) bude negde oko sredine distribucije uzorkovanja dobijene butstrepingom, ali ovo nekad ne bude slučaj. Kada sprovodimo postupke statistike zaključivanja na osnovu butstrepinga, razlika između aritmetičke sredine distribucije uzorkovanja koju smo napravili postupkom butstrepinga i aritmetičke sredine uzorka koji obrađujemo se takođe računa i navodi u opisu sprovedenog postupka. **Ova razlika (između AS**

PREPRINT

uzorka i AS distribucije uzorkovanja koja je dobijena butstrepingom) se obično zove pristrasnost ili bias (u literaturi na engleskom jeziku – bias), ali se **u literaturi sreću i drugi nazivi poput devijacija ili samo – razlika** (eng. deviation, difference).

Sve u svemu, za razliku od situacije primene pristupa zasnovanog na centralnoj graničnoj teoremi, gde samo procenjujemo svojstva distribucije uzorkovanja (pre svega standardnu grešku) korišćenjem formula, **u pristupu zasnovanom na butstrepingu, distribucija uzorkovanja se pravi na način koji je gore opisan i standardna greška se direktno računa iz nje**.

5.3. Procena parametara populacije, tačkasta i intervalna procena.

Postoje dva glavna pristupa proceni parametara – tačkasta procena i intervalna procena. Oba ova pristupa proceni vrednosti parametara se mogu primeniti i u okviru primene centralne granične teoreme i kroz postupak zasnovan na bustrepingu. **U okviru tačkaste procene, računa se i beleži procena vrednosti parametra i standardna greška. U okviru intervalnog pristupa proceni parametara, parametar se procenjuje definisanjem intervala poverenja i navođenjem verovatnoće da se unutar tog intervala nalazi parametar.**

Prilikom tačkaste procene vrednosti parametara na osnovu centralne granične teoreme, parametar se smatra jednakim statistiku, a računa se i prikazuje i standardna greška statistika. Naravno, postavke centralne granične teoreme nam kažu da statistik koji smo izračunali iz našeg uzorka može biti bilo gde na distribuciji uzorkovanja. Međutim, nisu sve pozicije na distribuciji jednako verovatne, a najverovatnija je ona koja se nalazi tačno na centru distribucije tj. na mestu gde se, po ovoj teoriji, nalazi parametar. To je razlog zašto se prilikom tačkaste procene vrednosti parametara u okviru pristupa zasnovanog na centralnoj graničnoj teoremi pretpostavlja da je statistik našeg uzorka jednak parametru. Kako valjan opis uzorka, po pravilu, treba da uključi meru centralne tendencije i meru varijabilnosti, i ovde je potrebno da se obezbedi mera varijabilnosti koja bi dala procenu verovatnu veličinu greške procene (veličinu greške koju pravimo kada pretpostavimo da je naš statistik jednak parametru). Ovaj zahtev se ispunjava procenom i navođenjem standardne greške. Prema tome, valjana tačkasta procena parametra u okviru pristupa zasnovanog na centralnoj graničnoj teoremi radi se tako što pretpostavimo da je statistik koji smo izračunali iz našeg uzorka jednak parametru i tako što, pored toga, izvestimo i o

PREPRINT

standardnoj greški tog statistika. Kao što je pomenuto u prethodnom poglavlju, **u okviru pristupa zasnovanog na centralnoj graničnoj teoremi, standardna greška se procenjuje korišćenjem formula.**

Tačkasta procena zasnovana na butstrepingu sprovodi se tako što se navedu vrednosti:

- razlike između statistika uzorka i aritmetičke sredine distribucije uzorkovanja koja je dobijena postupkom butstrapinga (ova razlka se zove pristrasnost ili bias) i

- standardna devijacija distribucije uzorkovanja tj. standardna greška.

Jednostavnije rečeno, **u okviru pristupa proceni parametra preko butstrepinga, tačkasta procena parametra se radi tako što se navedu statistik uzorka, aritmetička sredina distribucije uzorkovanja koja je dobijena butstrepingom, pristrasnost i standardna greška.**

Intervalna procena parametra se sprovodi tako što se **definiše interval poverenja za koji postoji određena, definisana, verovatnoća da se unutar intervala nalazi parametar**. Ovo se radi tako što se **definiše i proceni interval koji uključuje određenu, unapred definisanu, proporciju slučajeva iz distribucije uzorkovanja**. U okviru pristupa preko **centralne granične teoreme, procena intervala poverenja je zasnovana na očekivanju da je distribucija uzorkovanja normalna** (iako u praktičnim situacijama mi zapravo ne raspolažemo distribucijom uzorkovanja!), a **intervali poverenja se definišu kao intervali čija su gornja i donja granica jednako udaljene od aritmetičke sredine i unutar kojih se nalazi željeni procenat entiteta na normalnoj distribuciji** čija je aritmetička sredina zapravo aritmetička sredina uzorka, a čija je standardna devijacija standardna greška. Kako je objašnjeno u prethodnom poglavlju**, odstupanja od aritmetičke sredine izražena u standardnim devijacijama su z skorovi, a z skorovi se uvek mogu pretvoriti u percentile**. To znači da m**ožemo definisati interval na normalnoj distribuciji, koji sadrži željenu proporciju entiteta i čije su granice jednako udaljene od aritmetičke sredine, jednostavno tako što ćemo odrediti z skorove** (tj. broj standardnih devijacija ispod i iznad aritmetičke sredine) **između kojih se nalazi određena proporcija entiteta kada je distribucija normalna**.

U okviru pristupa zasnovanog na **butstrepingu**, **procedura bustrepinga nam dozvoljava da napravimo distribuciju uzorkovanja** (tj. efektivno – njenu simulaciju), a onda ostaje samo da se **definiše interval na toj distribuciji uzorkovanja za koji želimo da bude interval poverenja**.

PREPRINT

Najčešće korišćene propocije entiteta sa distribucije uzorkovanja koje se koriste za pravljenje intervala poverenja su **95% i 99%.** Iako je sasvim jednako moguće koristiti i bilo koji drugi procenat, postoji vrlo ukorenjen običaj među istraživačima da se koriste baš ova dva procenta.

U okviru centralne granične teoreme, da bi napravili 95% interval poverenja, potrebno je da standardnu grešku pomnožimo sa 1,96, **dok za 99% interval standardnu grešku množimo sa 2,56**. Ovo je zbog toga što se na normalnoj distribuciji 95% entiteta nalazi između z skorova -1,96 i + 1,96, a 99% entiteta je između z skorova -2,56 i +2,56. Interval od 1 z skora jednak je jednoj standardnoj devijaciji, ili jednoj standardnoj grešci distribucije uzorkovanja, te onda možemo reći da se 95% entiteta na normalnoj distribuciji uzorkovanja nalazi u intervalu koji počinje od 1,96 standardnih grešaka ispod aritmetičke sredine i završava se u tački koja je 1,96 standardnih grešaka iznad aritmetičke sredine. Prema tome, da bi odredili 95% interval poverenja oko statistika, pod pretpostavkama centralne granične teoreme, koristimo sledeće formule:

$$CI_{gornja\ granica} = Statistik + 1.96\ SG_{statistik}$$

$$CI_{donja\ granica} = Statistik - 1.96\ SG_{statistik}$$

U ovoj formuli CI (od englekog Confidence Interval) se odnosi na gornju odnosno donju granicu intervala poverenja i ovo je generalno skraćenica koja se veoma često koristi u naučnoj literaturi za označavanje intervala poverenja. Statistik u ovoj formuli može da bude bilo koji statistik koji koristimo da na osnovu njega izvedemo zaključke o parametru populacije. Taj statistik tako može da bude aritmetička sredina, medijana, standardna devijacija, varijansa, skjunes, kurtozis ili bilo koji drugi statistik računat na uzorku. SG je standardna greška tog konkretnog statistika.

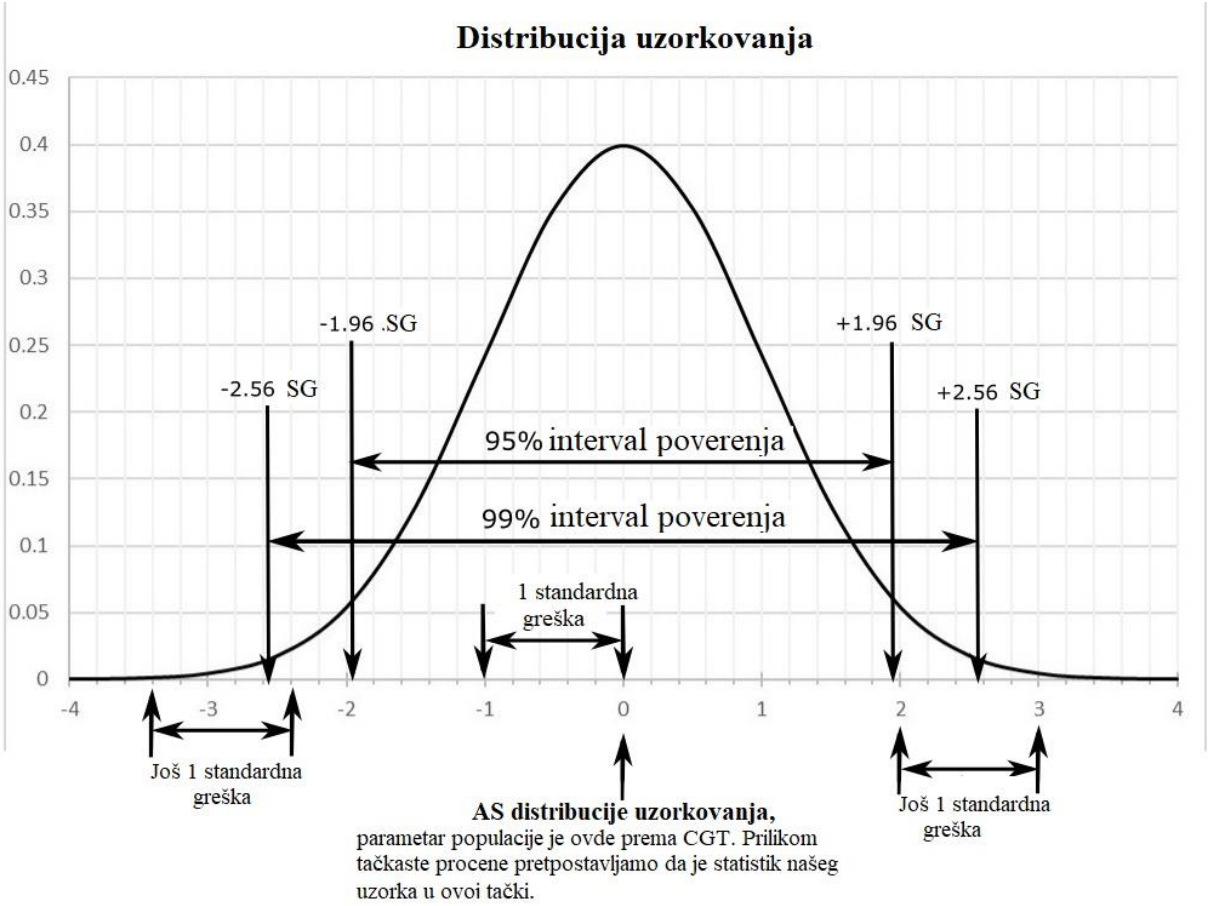
Da bi izračunali 99% interval poverenja u okviru pristupa zasnovanog na centralnoj graničnoj teoremi, koristimo iste formule, samo što umesto 1,96 stoji 2,56. Isto tako je moguće izračunati i bilo koji drugi interval (u smislu da zahvata bilo koji drugi procenat entiteta), samo što bi tada ovaj koeficijent 1,96, koji koristimo da dobijemo 95% interval, zamenili za koeficijent koji je odgovarajući za interval koji želimo. Ovaj koeficijent se može izračunati u statističkom softveru koristeći funkcije koje pretvaraju željene pozicije na normalnoj distribuciji u z skorove koji im odgovaraju.

PREPRINT

U okviru pristupa proceni parametara koji je zasnovan na bustrepingu, distribucija uzorkovanja nam je na raspolaganju (u stvari, zamena za distribuciju uzorkovanja – distribucija statistika iz velikog broja uzoraka dobijenih iz istraživačkog uzorka na kom radimo analizu, kako je prethodno opisano), tako da se interval poverenja dobija prosto tako što se identifikuje interval koji sadrži željeni procenat entiteta tj. statistika iz distribucije uzorkovanja. Ovaj interval se pravi tako da su njegova gornja i donja granica podjednako udaljene od centra distribucije uzorkovanja koja je dobijena butstrepingom.

Slika 5.2. Grafički prikaz distribucije uzorkovanja i povezanih pojmova. Prikazane su pozicije 95% i 99% intervala poverenja na distribuciji uzorkovanja, kao i veličina jedne standardne greške u odnosu na distribuciju i intervale poverenja. Brojevi na vertikalnoj osi su verovatnoće, dok su brojevi na horizontalnoj osi vrednosti varijable izražene u z skorovima. Veličina z skora jednaka je jednoj standardnoj devijaciji i, s obziom na to da je ovo distribucija uzoprkovanja, standardna devijacija se zove standardna greška, tako da možemo reći i da su jedinice na horizontalnoj osi zapravo standardne greške.

PREPRINT



A **šta se dešava ako ispadne da smo puno pogrešili** kada smo prihvatili pretpostavku da je statistik našeg uzorka jednak parametru? Ovo je posebno važno imajući u vidu da, kada pravimo procenu vrednosti parametra, mi nećemo znati koliko smo pogrešili kada smo prihvatili tu pretpostavku. Šta ako vrednost našeg statistika nije blizu vrednosti parametra? Da, itekako je moguće i često se i dešava da vrednosti razmatranog statistik istraživačkog uzorka koji analiziramo i vrednosti parametra nisu baš međusobno blizu. Centralna granična teorema opisuje upravo takve mogućnosti. Nekad će se takođe i desiti da je statistik našeg uzorka zapravo negde na samim krajevima distribucije uzorkovanja, a kako ne znamo vrednost parametra, samo na osnovu tog jednog istraživanja/uzorka koji imamo, nećemo ni moći da znamo da je to tako ispalo. Međutim, to je ona situacija u kojoj standardna greška i intervali poverenja ulaze u igru. Tu samo treba da se podsetimo onoga što je očigledno – udaljenost statistika od parametra je jednaka udaljenost parametra od statistika. Tako da, **ako postavimo interval poverenja oko parametra** (zamislimo za trenutak da nekako znamo vrednost parametra; postavljanje intervala poverenja oko parametra

znači onda da je parametar u centru takvog intervala) **i taj interval poverenja je dovoljno širok da obuhvati i statistik iz našeg uzorka, onda je i interval poverenja iste širine koji bi postavili oko statistika** (tako da statistik bude u centru intervala) **dovoljno širok da obuhvati parametar**! Prema tome, dok god je naš interval poverenja dovoljno širok, sve bi trebalo da bude u redu u većini slučajeva, a ako radimo tačkastu procenu, veličina standardne greške nam daje informaciju o veličini verovatne greške u procenu parametra, te onda na dalje radimo sa razumevanjem o tome koliki je nivo preciznosti sa kojim procenjujemo vrednost parametra. Imajući ovo u vidu, **treba biti svestan i toga da će u retkim sluačejvima naša procena parametra biti veoma pogrešna i ta mogućnost je već uključena u metodologiju koju opisujemo – 95% interval poverenja koji funkcioniše kako treba i dalje znači da će u 5% slučajeva parametar biti van našeg intervala poverenja (ili u 1% slučajeva, ako je reč o 99% intervalu poverenja).**

5.4. Nulta hipoteza

Kada istražujemo određeni fenomen, pored toga što ga opisujemo, ili u slučaju statistike, pored toga što izvodimo zaključke o svojstvima populacije na osnovu uzorka ili uzoraka koji/koje smo proučili, želećemo obično i da izvedemo određene zaključke o vezama između ispitivanog fenomena i nekih drugih poznatih ili sličnih fenomena. Na primer, kada proučavamo određena svojstva određene populacije, nećemo naše istraživanje ograničiti samo na procenu svojstava te populacije na osnovu uzorka, već se takođe možemo pitati i da li su svojstva populacije koju proučavamo različita od svojstava neke druge poznate populacije. Da li su ljudi iz zemlje A viši od ljudi iz zemlje B? Da li su stavovi ljudi u ovom gradu različiti od stavova ljudi u nekom drugom gradu? Da li deca različitog uzrasta imaju različite nivoe izraženosti određene psihološke crte/osobine? Možemo biti i zainteresovani za odnose između različitih fenomena. Na primer, može nas interesovati da li određena crta ličnosti čini ljude sklonijim da donose određene vrste poslovnih odluka ili da li ih čini sklonijim da kupuju i čuvaju određene predmete. Može nas interesovati da li određena radnja proizvodi određene posledice. Na primer, možemo se pitati da li određeni psihoterapijski postupak koji primenjujemo dovodi do promena kod ljudi koji su mu bili podvrgnuti. Ili da li je znanje o određenoj temi poraslo u grupi studenata nakon što su pohađali kurs na tu temu? Ili da li određeni lek radi tj. smanjuje simptome određene bolesti ili je leči u potpunosti.

PREPRINT

Na sva ova pitanja se tipično odgovara tako što se prvo postavi hipoteza, koja je, u suštini, specifikacija naših očekivanja o tome kakvi su odnosi između fenomena koje proučavamo (varijabli, grupa/populacija), a onda se ta hipoteza testira i ustanovi da li je u skladu sa podacima koje smo dobili. Međutim, da bi testirali određenu hipotezu, potrebno je da ta hipoteza bude veoma precizna. Nasuprot tome, naša očekivanja o tome kakvi su odnosi između varijabli ili grupa entiteta po pravilu uopšte nisu precizna. Na primer, naše očekivanje da će ljudi imati više znanja o određenoj temi nakon što su pohađali kurs koji obrađuje tu temu obično ne uključuje informaciju o tome koliko će tačno nivo znanja grupe koja je pohađala kurs da poraste (u odnosu na nivo koji su imali pre kursa ili neku drugu grupu koja nije pohađala kurs). Naše očekivanje da će neki postupak lečenja da radi ne znači nužno da očekujemo da će 100% ljudi sa određenom bolešću koji budu podvrgnuti tom postupku da bude izlečeno. Postupak lečenja koji uspešno leči 80% bolesnika se u mnogim slučajevima može smatrati sasvim efikasnim. Na sličan način možemo zaključiti da se stavovi dve grupe razlikuju i u situaciji kada su stavovi poređenih grupa potpuno različiti, kao i u situaciji kada se tek donekle razlikuju. Iz ovih razloga često nije moguće tačno specifikovati kolika očekujemo da bude razlika između grupa koje poredimo ili koliko tačno očekujemo da bude jaka veza između varijabli čiju povezanost proučavamo. Na primer, ako naša hipoteza kaže da primena određenog postupka lečenja dovodi do izlečenja kod 100% osoba koje boluju od određene bolesti i ako, nakon testiranja te hipoteze, otkrijemo da 100% osoba koje boluju od te bolesti i koje su podvrgnute tretmanu nije izlečeno, da li to znači da ovaj postupak ne radi ili pak da radi, samo ne u 100% slučajeva? Ili ako otkrijemo da pohađanje određenog kursa nije dovelo do prosečnog povećanja skora na testu znanja od tačno 55 poena, da li to znači da ljudi nisu naučili ništa na tom kursu ili da njihova promena nivoa znanja odgovara nekom drugom nivou promene skora na testu znanja? Oba ova ishoda su moguća, naravno. Ovo su vrste situacija u kojima se nulta hipoteza pokazuje kao veoma koristan alat.

U društvenim naukama, naukama o ponašanju i medicinskim naukama, najčešće se sreće određenje nulte hipoteze po kome je **nulta hipoteza bilo koja hipoteza koja kaže da je vrednost proučavanog parametra 0 (ili razlika između parametara, koja je sama za sebe takođe parametar).** Nulta hipoteza se **tipično označava sa H_0.** Nulta hipoteza može da kaže da dva uzorka potiču iz populacija čije se aritmetičke sredine na ispitivanoj varijabli razlikuju za 0 (tj. koje se ne razlikuju). Može da kaže i da je vrednost određenog statistika u populaciji 0. Ili da veći broj uzoraka potiče iz populacija čiji se ispitivani statistici razlikuju za 0 (tj. da ispitivani statistik

PREPRINT

ima istu vrednost u svim posmatranim populacijama). Vrednost 0 je važna zato što su dve vrednosti iste ako im je razlika nula, ali ako razlika između njih nije 0, to znači da se te vrednosti razlikuju. Takođe, ako je veličina određenog efekta 0, to znači da nema efekta. Ako je veličina/intenzitet nekog efekta bilo koja vrednost koja nije 0, to znači da taj efekat postoji. **Opovrgavajući nultu hipotezu pokazujemo da veličina o kojoj nulta hipoteza govori nije 0.** Na primer, da postoji razlika između posmatranih populacija, da ispitivani efekat postoji (da nije 0), da se parametri koje posmatramo razlikuju itd. **Najčešća upotreba nulte hipoteze je za ispitivanje da li dva ili više uzoraka potiče iz iste populacije** (ili iz populacija između kojih je razlika 0 tj. koje su sve iste u pogledu ispitivanog svojstva) **ili da li je vrednost nekog statističkog parametra 0.**

Međutim, onako kako je izvorno zamišljena, **H_0 ili nulta hipoteza uopšte ne mora da tvrdi da je neka vrednost nula, već to može da bude bilo koja precizna statistička hipoteza. Drugim rečima nulta hipoteza može da bude hipoteza koja specifikuje bilo koju konkretnu vrednost ili bilo koji konkretan odnos, dokle god je potpuno precizna. Nulta hipoteza se zove nulta hipoteza zato što je zamišljena kao hipoteza koja treba da bude opovrgavana tj. nulifikovana (eng. nullified) empirijskim podacima** (e.g. Gigerenzer, 2004), a ne zbog toga što kaže da je nešto nula. Zbog ovoga, neki autori naglašavaju razliku između nulte hipoteze, koju definišu kao bilo koju preciznu statističku hipotezu koja je formulisana zato da bi statističkim testiranjem pokušali da je opovrgnemo od nil hipoteze (eng. nil hypothesis) koju definišu kao podskup nultih hipoteza koji čine nulte hipoteze koje kažu da je posmatrani parametar 0 ili da je odnos između parametara jednak nuli. U suštini, **ono što tipični istraživači obično nazivaju nultom hipotezom, ovi autori zovu nil hipotezom**.

Ali šta će nam uopšte nulta hipoteza? Kao što je opisano u prethodnim primerima, glavna prednost nulte hipoteze je u tome što je precizna – kaže da je nešto jednako nuli. S druge strane, **takozvane alternativne hipoteze, hipoteze koje opisuju naša realna očekivanja, po pravilu nisu precizne.** Alternativna hipoteza može biti očekivanje da se statistici razlikuju, da posmatrani statistik ima neku vrednost koja nije nula, da proučavani efekat postoji ili nešto slično, ali kada govorimo o razlici, razlika može značiti puno različitih stepena razlike, vrednost koja nije nula može biti puno različitih vrednosti, a i efekat koji postoji može imati mnogo različitih intenziteta, od veoma niskih do veoma visokih. **Nulta hipoteza i alternativna hipoteza su suprotne jedna drugoj, tako da ako je jedna tačna, druga nije.** To znači da **testiranjem nulte hipoteze,** koja je precizna i zbog

PREPRINT

toga i testabilna, **mi testiramo i alternativnu hipotezu**. Ako testiramo nultu hipotezu koja kaže da je efekat određenog postupka lečenja koji nas interesuje jednak nuli, tj. da on nema efekta, pa rezultati pokažu da je to istina, to istovremeno znači i da naša alterantivna hipoteza, tj. očekivanje da dati tretman ima efekta nije tačna. S druge strane, ako zaključimo da naša nulta hipoteza nije tačna, to znači da je naša alternativna hipoteza, naše pravo očekivanje opravdano. U prethodnom primeru sa testiranjem efekata postupka lečenja, zaključak da nulta hipoteza nije tačna tj. odbacivanje nulte hipoteze istovremeno imlicira i da efekat tretmana nije nula.

Kako ovo primenjujemo u praksi? **Opšta procedura za testiranje nulte hipoteze polazi od postavljanja pretpostavke da je nulta hipoteza tačna u populaciji**. Sećamo se iz prethodnog poglavlja da statistici uzorka mogu više ili manje da odstupaju od parametara populacije. Stoga ova procedura počinje pretpostavljanjem da je nulta hipoteza tačna u populaciji bez obzira kakve su vrednosti statistika koje smo dobili na našem uzorku. Nakon što smo to pretpostavili **pristupamo računanju verovatnoće da dobijemo na uzorku rezultate kakve smo dobili ili rezultate koji još više odstupaju od stanja definisanog nultom hipotezom u situaciji kada je nulta hipoteza tačna u populaciji**. Ovu računicu radimo **ili primenjujući centralnu graničnu hipotezu ili direktno kroz neki od metoda reuzorkovanja**, kao što je opisano u prethodnom poglavlju. Konačno, **na osnovu te verovatnoće tj. na osnovu toga koliko je velika ta verovatnoća, donosimo odluku o tome da li smatramo verovatnim da je nulta hipoteza tačna u populaciji ili ne.**

Na primer, ako bi želeli da testiramo da li dva uzorka potiču iz iste populacije, ili još preciznije, iz populacija koje imaju iste vrednosti aritmetičke sredine na varijabli koja nas interesuje, počeli bi od pretpostavke da je razlika između parametara dve populacije tj. njihovih aritmetičkih sredina na datoj varijabli 0. Onda izračunamo verovatnoću, bilo na osnovu centralne granične teoreme, bilo kroz reuzorkovanje, da razlika između aritmetičkih sredina dva uzorka koja je dobija na našem uzorku ili veća bude dobijena na dva uzorka u situaciji kada je nulta hipoteza tačna u populaciji. Na kraju, donosimo odluku od tome da li je verovatnoća koju smo izračunali dovoljno mala da bi zaključili da nije verovatno da je nulta hipoteza tačna u populaciji, nakon čega nultu hipotezu odbacujemo i prihvatamo alternativnu, ili je dovoljno velika da proglasimo da je izgledno da je nulta hipoteza tačna u populaciji te da na osnovu toga prihvatimo nultu hipotezu, a odbacimo alternativnu.

PREPRINT

Ključni statistik u ovim razmatranjima je verovatnoća da se dobiju rezultati kakvi su dobijeni na uzorku ili još ekstremniji (ekstremniji u smislu da još više odstupaju od stanja definisanog nultom hipotezom od stanja dobijenog na uzorku) **u situaciji kada je nulta hipoteza tačna u populaciji.** Ovaj statistik **zove se statistička značajnost. Statistička značajnost se tipično predstavlja kao proporcija, sa vrednostima između 0 i 1**. Obično se **označava sa p. ili sig. ili "znač" u srpskoj literaturi i često se o njoj govori kao o "p vrednosti" ili samo o "značajnosti"**.

Na primer, ako bismo hteli da izvedemo zaključak o tome da li dva uzorka potiču iz populacija sa identičnim aritmetičkim sredinama na posmatranoj varijabli, mogli bi da izračunamo verovatnoću da se razlika između aritmetičkih sredina dva uzorka veličine kao što je ona koju smo dobili na našim uzorcima ili veća dobije u situaciji kada je razlika između aritmetičkih sredina populacija nula. Ta verovatnoća bi bila statistička značajnost. Isto bi bilo i ako bi računali verovatnoću da određeni statistik uzorka (na primer skjunes ili kurtozis) ima vrednost koju ima na našem uzorku ili veću (u stvari – više različitu od nule, bilo u pozitivnom ili u negativnom smeru) u situaciji kada je vrednost parametra populacije nula (u ovom primeru – skjunesa odnosno kurtozisa).

Statistička značajnost se koristi u odlučivanju da li da odbacimo ili da prihvatimo nultu hipotezu. To se tipično radi tako što **poredimo vrednost statističke značajnosti sa kritičnim nivoom (statističke značajnosti) za odlučivanje koji smo prihvatili**. **Kritični nivoi statističke značajnosti koji se najčešće sreću u naučnoj i stručnoj literaturi su 5% i 1% odnosno 0,05 i 0,01.** Primena 5% ili 0,05 kao kritičnog nivoa statističke značajnosti znači da **ako je verovatnoća 5% ili ispod 5% (p<0,05) da se rezultati kakvi su dobijeni na našem uzorku ili ekstremniji dobiju u situaciji kada je nulta hipoteza tačna u populaciji, onda nultu hipotezu odbacujemo**. S druge **strane, ako je verovatnoća da dobijemo rezultate kakve smo dobili na uzorku ili ekstremnije u situaciji kada je nulta hipoteza tačna u populaciji veća od 5% (p>0,05), onda prihvatamo nultu hipotezu**. **"Ekstremniji rezultati" ovde znači da se rezultati više razlikuju od situacije koja je pretpostavljena nultom hipotezom nego rezultati dobijeni na uzorku**. Postupak za odlučivanje o tome da li prihvatiti ili odbaciti nultu hipotezu koristeći nivo statističke značajnosti od 0,01 ili 1% kao prag za odluku je isti, samo što je u tom slučaju kritični nivo statističke značajnosti 0,01 umesto 0,05.

PREPRINT

Statistička značajnost se u principu računa slično kao što se računaju intervali poverenja kod procene parametara (molim pogledajte prethodno poglavlje) – **napravimo distribuciju uzorkovanja kakva bi bila dobijena ako bi nulta hipoteza bila tačna ili izračunamo njene parametre, a onda označimo oblast distribucije koja sadrži vrednosti statistika koji su dobijeni na uzorku i one koji se više od toga razlikuju od vrednosti određenih nultom hipotezom (tj. od 0). Proporcija slučajeva na distribuciji uzorkovanja koja spada u tu oblast jednaka je statističkoj značajnosti**. Na primer, ako bismo hteli da uporedimo aritmetičke sredine dve populacije i naša nulta hipoteza je da je razlika između aritmetičkih sredina populacija nula, mi bi onda napravili ili izračunali svojstva distribucije uzorkovanja razlika između aritmetičkih sredina uzoraka uzetih iz ove dve populacije. Ako je razlika između aritmetičkih sredina uzoraka koje poredimo, recimo, 2, mi bi onda izračunali proporciju slučajeva na distribuciji uzorkovanja razlika između aritmetičkih sredina koji su jednaki 2 ili više (a u nekim slučajevima i -2 ili manje) i ova proporcija bi predstavljala nivo statističke značajnosti razlike između uzoraka.

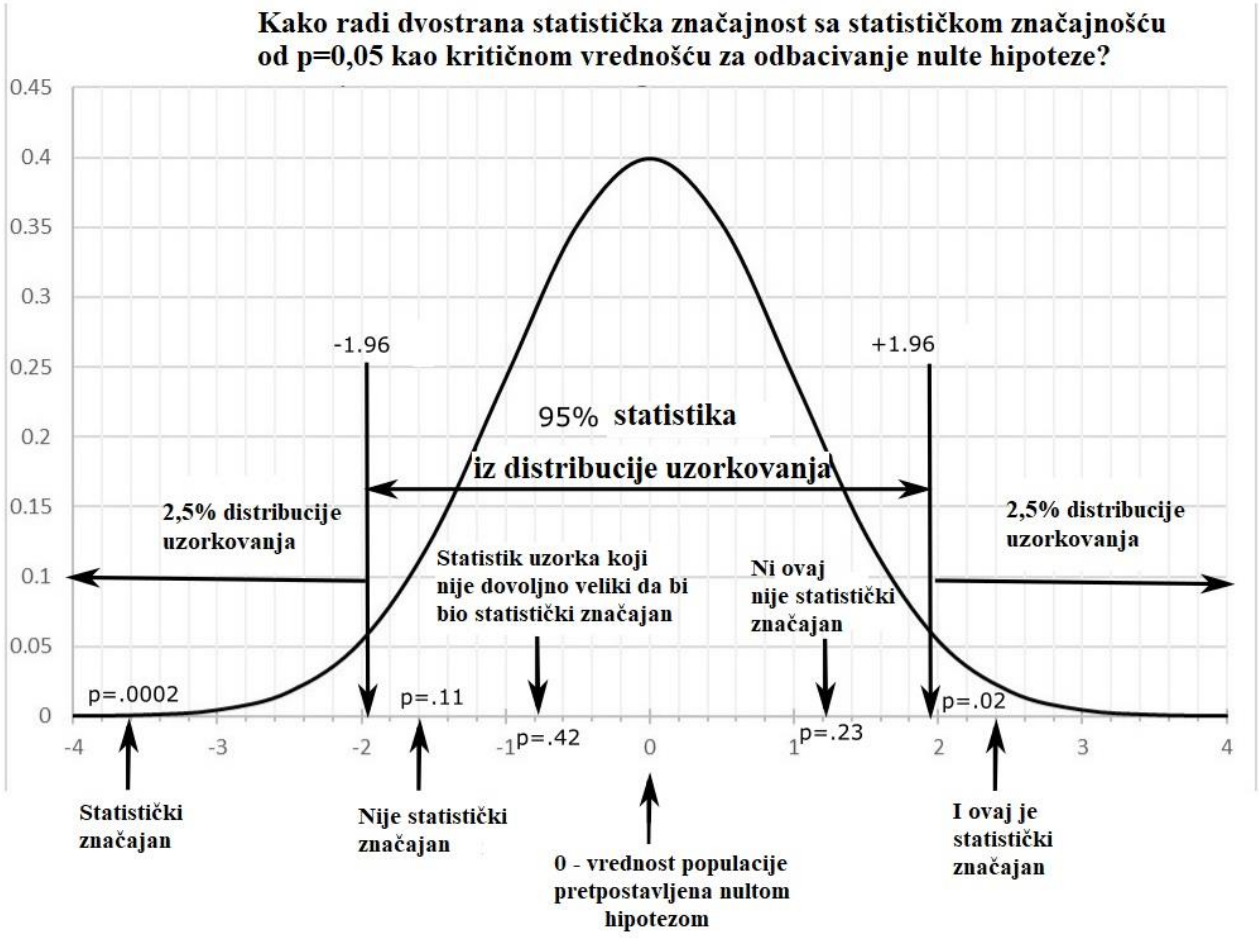
U vezi sa ovim, postoje u principu **dva načina da se označi ova oblast prilikom računanja statističke značajnosti – možemo računati jednostranu statističku značajnost ili dvostranu statističku značajnost** (eng. **one-tailed, two-tailed**). **Jednostrana statistička značajnost se računa tako što se u oblast na distribuciji uzorkovanja koja se koristi za procenu statističke značajnosti uključuju samo entiteti koji su na istoj strani distribucije kao i rezultat dobijen na uzorku, dok se dvostrana statistička značajnost računa tako što se uključuju eniteti sa istim ili višim nivoom razlike od vrednosti određene nultom hipotezom sa obe strane distribucije**. U prethodnom primeru, onom gde nulta hipoteza kaže da je razlika između aritmetičkih sredina dve populacije 0, a gde je dobijena razlika između aritmetičkih sredina dva uzorka dobijena iz te dve populacije bila 2, jednostranu statističku značajnost bi dobili tako što bi uzeli u obzir samo oblast distribucije uzorkovanja koja sadrži vrednost 2 i vrednosti veće od 2. Dvostranu statističku značajnost bi dobili tako što bi uzeli u obzir vrednost 2 i vrednosti veće od dva (jednu stranu distribucije – onu iznad aritmetičke sredine), ali takođe i vrednost od -2 i vrednosti manje od -2 (drugu stranu distribucije – onu ispod aritmetičke sredine). Ovo je zbog toga što vrednosti 2 i -2 predstavljaju jednak nivo odstupanja od situacije specifikovane nultom hipotezom. Jedina razlika je u tome što su to razlike različitog smera – jedna je pozitivna, a druga negativna. To znači da, imajući u vidu simetričnu distribuciju uzorkovanja, nivo dvostrane

PREPRINT

statističke značajnosti će uvek biti dva puta veći od nivoa jednostrane statističke značajnosti – ako je jednostrana statistička značajnost 0,1, onda će dvostrana statistička značajnost biti 0,2.

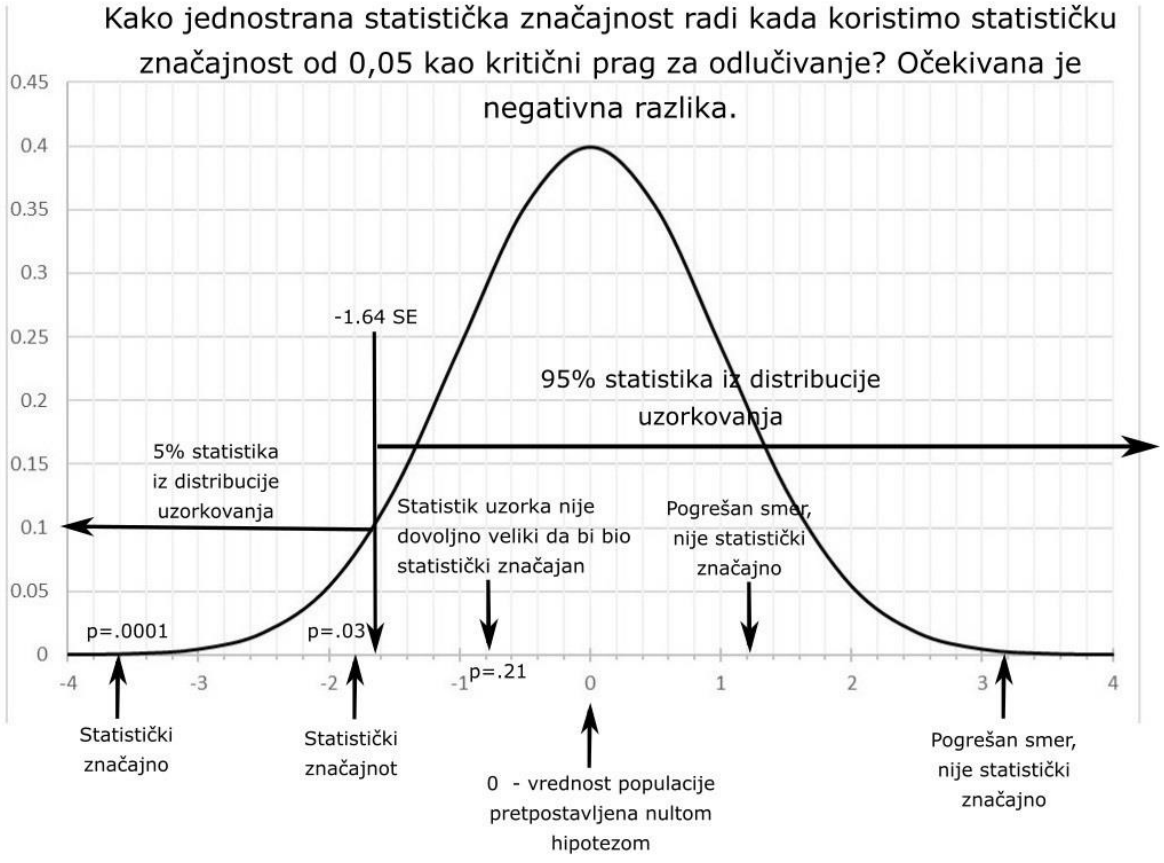
Slika 5.3. Grafički prikaz načina na koji se statistička značajnost koristi da bi se donela odluka o tome da li prihvatiti ili odbaciti nultu hipotezu koristeći dvostranu statističku značajnost. Ovaj primer koristi 0,05 kao kritični prag statističke značajnosti za donošenje odluke. Možemo videti da je 95% interval poverenja napravljen oko populacijske vrednosti koja je pretpostavljena nultom hipotezom, što je nula u ovom slučaju, i prikazan je raspon vrednosti koje nisu statistički značajne, između -1,96 i +1,96 standardnih grešaka oko nule tj. centra distribucije. Kako ovde koristimo dvostranu statističku značajnost, uzimamo u obzir odstupanja od nulte hipoteze u oba smera i to znači da ovaj prag od 0,05, ostavlja van intervala statistički neznačajnih vrednosti po 2,5% statistika sa distribucije uzorkovanja sa svake strane distribucije (i gornje i donje). Na slici možemo videti i da izabrani statistici uzorka (označeni kratkim strelicama i njihovim p vrednostima tj. nivoima statističke značajnosti) koji su unutar intervala poverenja koji je predstavljen nisu statistički značajni, što znači da bi, kada bi dobili takve statistike na uzorku, nultu hipotezu prihvatili (za sve čije su p vrednosti iznad 0,05), a vidimo i statistike van intervala poverenja (one čije su p vrednosti manje od 0,05) koje su statistički značajne i za koje bi odbacili nultu hipotezu. Treba ovde primetiti kako se p vrednosti smanjuju u oba smera (i u pozitivnom i u negativnom) sa udaljavanjem od centra distribucije uzorkovanja postavljene oko vrednosti koju pretpostavlja nulta hipoteza.

PREPRINT



Slika 5.4. Grafički prikaz načina na koji se statistička značajnost koristi za odlučivanje o tome da li prihvatiti ili odbaciti nultu hipotezu kada se koristi jednostrana statistička značajnost. Ovaj primer takođe koristi 0,05 kao kritični prag statističke značajnosti. Kako je ovde u pitanju pristup preko jednostrane statističke značajnosti, u postavci je navedeno i očekivanje da će razlike od stanja definisanog nultom hipotezom, ako ih bude, biti u negativnom smeru. To znači da se sve pozitivne razlike od stanja definisanog nultom hipotezom (koje su, dakle, suprotne očekivanom smeru), tretiraju kao statistički neznačajne, odnosno nisu osnov za odbacivanje nulte hipoteze. Ovakav pristup bi bilo ispravno primeniti u praksi samo ako bi znali da su razlike u pozitivnom smeru nemoguće i da su, ako se i jave, sigurno greška (ako su odstupanja u oba smera moguća, tada treba da koristimo dvostranu statističku značajnost, a ne ovo!!!). Ovde govorimo o tome da

su pozitivne razlike nemoguće zato što smo naveli primer gde je očekivana razlika negativna. Da je očekivana razlika pozitivna, onda bi bilo neophodno da negativne razlike budu nemoguće da bi bilo ispravno koristiti jednostranu statističku značajnost. Na slici možemo videti da razlike koje nisu dovoljno velike da njihove p vrednosti budu niže od prihvaćenog praga statističke značajnosti nisu statistički značajne. Međutim, kako je ovde u pitanju jednosmerna statistička značajnost, van oblasti statističke neznačajnosti je samo 5% statistika na jednom kraju distribucije, dok svi slučajevi na drugom kraju distribucije spadaju u oblast statističke neznačajnosti. Ovo je razlog zašto je granica ove oblasti označena samo sa jedne strane, kao i zašto je bliža aritmetičkoj sredini tj. centru distribucije (na 1,64 standardne greške od AS) nego što je to bio slučaj kod dvostrane statističke značajnosti. Možemo primetiti i kako ovde statistici uzorka koji su jednako udaljeni od aritmetičke sredine kao oni na slici 5.3. ovde imaju duplo manje p vrednosti, nego na slici 5.3.



Kada koristimo jednostranu, a kada dvostranu distribuciju? **U praksi, istraživači će računati dvostranu statističku značajnost uvek, osim u slučajevima kada su odstupanja od stanja pretpostavljenog nultom hipotezom moguća samo u jednom smeru, odnosno kada su odstupanja od vrednosti pretpostavljenih nultom hipotezom u drugom smeru nemoguća.** Ovo može biti slučaj samo onda kada postoje jasni teorijski razlozi zbog kojih razlike u drugom smeru nisu moguće ili u situacijama kada nije tehnički moguće da vrednosti odstupaju od stanja pretpostavljenog nultom hipotezom u drugom smeru. Na primer, ako imamo kurs/obuku na određenu temu u nekom programu studija, i to na temu o kojoj studenti ništa nisu znali pre kursa, tada bi mogli smisleno da računamo jednostrani nivo statističke značajnosti prilikom poređenja skorova na testu znanja o toj temi pre i posle kursa, zato što znamo da je nivo znanja mogao da se promeni samo u jednom smeru – naviše. Međutim, ako bismo hteli da poredimo dve populacije studenata, jednu koja je pohađala jednu verziju nekog kursa i drugu koja je pohađala neku drugu verziju tog kursa, morali bismo da računamo dvostranu statističku značajnost, jer ne bi mogli da budemo sigurni da je razlika između grupa moguća samo u jednom smeru, čak i da smo krenuli od ideje da je jedna od te dve verzije tog kursa bolja. Imajući sve to u vidu, ovde se treba i podestiti na to da bi u primeru koji je opisan, statističku značajnost mogli smisleno da računamo (a i generalno da koristimo statistiku zaključivanja) samo ako bi hteli da saznamo kakvi bi bili efekti kursa na ljude uopšte koji su slični ovim studentima ako bi ti ljudi pohađali kurs. Ako smo testirali sve studente koji su pohađali kurs i samo nas interesuje uspeh tih konkretnih studenata, mogli bi da uporedimo njihove rezultate samo posmatrajući deskriptivne statističke mere i direktno ih poredeći bez računanja statističke značajnosti ili bilo kakve statistike zaključivanja. Kao što je već rečeno, statistika zaključivanja se koristi isključivo u situacijama kada hoćemo da izvedemo zaključke o populaciji na osnovu podataka o uzorku. Da bi izvukli samo zaključke o uzorku, dovoljan nam je samo taj uzorak!

Kako je računanje statističke značajnosti zasnovano na distribuciji uzorkovanja, **na vrednost statističke značajnosti utiču isti oni faktori koji određuju širinu intervala poverenja – varijabilnost podataka i veličina uzorka tj. broj entiteta u uzorku**. **Što su vrednosti entiteta iz uzorka uniformnije (tj. što se manje međusobno razlikuju) i što je uzorak veći, niže će biti vrednosti statističke značajnosti tj. p vrednosti**. Ovo znači da će **istom nivou odstupanja od vrednosti pretpostavljenih nultom hipotezom (tj. od 0) odgovarati manje p vrednosti ako se veličina uzorka poveća, a isto će se desiti ako se varijabilnost varijable (na kojoj su podaci)**

PREPRINT

smanji. Ako se ova dva faktora održavaju fiksnim, vrednost statističke značajnosti će zavisiti od veličine razlike između dobijenog statistika na uzorku i vrednosti koje bi očekivali na osnovu nulte hipoteze (tj. u većini praktičnih slučajeva upotrebe nulte hipoteze - koliko se vrednosti dobijene na uzorku razlikuju od nule). **Što je ova razlika veća, to će vrednosti statističke značajnosti biti manje.**

Kada pričamo o statističkoj značajnosti, valja biti svestan i **jedne kontradikcije** koja postoji **između načina na koji se o statističkoj značajnosti govori u svakodnevnom naučnom statisitčkom žargonu i samih brojčanih vrednosti statističke značajnosti**. Naime, **u naučnom/stručnom žargonu, „viši nivo statističke značajnosti“ znači da su vrednosti statističke značajnosti niže!** Na primer, naučnici će tipično reći da je vrednost statističke značajnosti od 0,01 statistički značajnija tj. da ukazuje na viši nivo statističke značajnosti od 0,02, koja opet predstavlja viši nivo statističke značajnosti od, na primer, 0,2. To znači da, **u naučnom/stručnom žargonu, veći ili viši nivo statističke značajnosti tipično znači manju brojčanu vrednost statističke značajnosti.** Veoma je važno biti svestan ovog kurioztiteta naučnog statističkog žargona, pogotovo za početnike u oblasti istraživačke statistike, jer njegovo nerazumevanje može dovesti do velikih zabuna prilikom čitanja naučnih publikacija i praćenja naučnih diskusija.

Još jedna važna stvar koje treba biti svestan je da kada se oslanjamo na nivo statističke značajnosti prilikom odlučivanja o tome da li prihvatiti ili odbaciti nultu hipotezu **i kada biramo određeni kritični nivo statističke značajnosti kao prag za odbacivanje/prihvatanje nulte hipoteze, mi takođe prihvatamo i to da će naši zaključci o tome da li se nulta hipoteza može prihvatiti ili ne biti pogrešni u određenom broju slučajeva.** Ovo će se desiti ili zato što smo prihvatili pogrešnu nultu hipotezu ili zato što smo odbacili nultu hipotezu koja je zapravo bila tačna. Ovakve greške se zovu statističke greške:

- **Greška koju pravimo onda kada odlučimo da odbacimo nultu hipotezu koja je tačna zove se statistička greška tipa 1**. Greška tipa 1 se takođe zove i **lažno pozitivni nalaz** – proglašavanje da rezultat (koji je pretpostavljen alternativnom hipotezom) postoji, u situaciji kada zapravo ne postoji u populaciji.

PREPRINT

- **Grešku koju pravimo onda kada prihvatimo nultu hipotezu koja nije tačna zove se statistička greška tipa 2.** Greška tipa 2 se takođe naziva i **lažno negativni nalaz** – proglašavanje da očekivani efekat ne postoji u situaciji kada dati efekat postoji u populaciji.

Statistička greška tipa 1 se dešava onda **kada odlučimo, na osnovu statističke značajnosti, da nije verovatno da naš uzorak dolazi iz populacije u kojoj je nulta hipoteza tačna**, zato što su vrednosti statistika koje posmatramo na uzorku previše različite od vrednosti koje pretpostavlja nulta hipoteza. Međutim, čak i malo verovatni događaji se nekada dese, tako da **ako uzmemo kao kritični nivo statističke značajnosti 5% ili 0,05, to znači da ćemo za 5% uzoraka, uzetih iz populacije** (ili populacija) u kojima je nulta hipoteza tačna, **da zaključimo (pogrešno) da ne potiču iz takve populacije**. Ako snizimo kritični nivo statističke značajnosti (tj. u statističkom žargonu – ako zahtevamo veću statističku značajnost za odbacivanje nulte hipoteze) na 1% odnosno 0,01, onda možemo očekivati da imamo 1% situacija kada odbacujemo nultu hipotezu koja je tačna u populaciji. Dakle, **povećanje statističke značajnosti potrebne za odbacivanje nulte hipoteze** (što znači da smanjujemo numeričku vrednost kritičnog nivoa statističke značajnosti, smanjujemo kritičnu p vrednost) **će smanjiti i verovatnoću pravljenja statističke greške tipa 1.** Međutim, **u isto vreme, dok smanjujemo verovatnoću statističke greške tipa 1, raste verovatnoća činjenja statističke greške tipa 2**. Do ovog dolazi zbog toga što, kako je već pomenuto, **zahtevanje niže p vrednosti** (odnosno veće statističke značajnosti) **za odbacivanje nulte hipoteze povećava raspon vrednosti uzorka za koje ćemo prihvatiti nultu hipotezu**. To onda **znači da će biti više mogućih vrednosti populacije koje su različite od onih koje predviđa nulta hipoteza, ali koje nisu dovoljno različite od vrednosti predviđenih nultom hipotezom da statistici dobijeni na uzorcima uzetim iz njih dovoljno često pređu prag statističke značajnosti** (dovoljno često, zato što i kada nema razlike između populacija, razlike između uzoraka uzetih iz njih nekada pređu graničnu vrednost, to treba da zapamtimo!). Drugim rečima, **da bi istraživači mogli tačno i pouzdano da otkriju razlike između pravih vrednosti populacije i vrednosti populacije koje su očekivane na osnovu nulte hipoteze, razika između tih vrednosti treba da bude dovoljno velika**. Mala odstupanja od nulte hipoteze su utoliko teža da se otkriju preko računanja statističke značajnosti što su manja. Prema tome, ustrožnjavanje zahtevanog praga statističke značajnosti za odbacivanje nulte hipoteze (tako što bi tražili numerički manje p vrednosti za odbacivanje nulte hipoteze) smanjuje verovatnoću činjenja statističke greške tipa 1, ali povećava verovatnoću činjenja statističke greške tipa 2 i obrnuto.

PREPRINT

I šta onda možemo da uradimo oko ovih grešaka? Prvo jednu stvar treba da razjasnimo – **šta god radili, nikada ne možemo u potpunosti eliminisati mogućnost greške, uvek će postojati nenulta verovatnoća da se napravi greška** (tj. da se donese pogrešan zaključak). To je razlog zašto **generalno nije mnogo pametno da se odluke sa važnim praktičnim posledicama donose na osnovu rezultata koji su dobijeni na samo jednom uzorku**. Ako je odluka koju donosimo na osnovu statističkih zaključaka važna ili može imati ozbiljne posledice, takva odluka bi trebala, kad god je moguće, da bude zasnovana na seriji nezavisnih studija koje su sve sprovedene na međusobno nezavisnim uzorcima.

Sledeća stvar koju treba razmotriti je to **koliko nas košta greška tipa 1, a koliko greška tipa 2**. Da li je veća šteta da se ne uoči efekat koji stvarno postoji ili da se objavi da efekat postoji kada ga zapravo nema? Cena odnosno šteta od pravljenja jedne ili druge greške zavisi od prirode fenomena koji proučavamo. Na primer, ako testiramo nove načine da se izleče neke uobičajene bolesti, bolesti za koje već postoje efikasni lekovi i kada mi poredimo efikasnost novih postupaka lečenja sa efikasnošću onih koji već postoje, verovatno je da bismo hteli da verovatnoća greške tipa 1 bude minimalna. Ovo je tbog toga što nas greška tipa 1 može dovesti do zaključka da je neefikasni postupak lečenja efikasniji od nekog već postojećeg. Mi bi onda investirali novac, i to značajne količine novca, u razvoj ovog postupka i njegovu primenu u kliničkoj praksi. Međutim, uskoro bi postalo poznato da nova procedura u stvari nije efikasnija od postojeće (ali je, recimo, skuplja). Kada ovo postane poznato javnosti, izgubili bismo klijente, naš postupak bi bio napušten i investicija bi nam propala. Prema tome, u takvim situacijama, bilo bi manje opasno da nam promakne novi postupak lečenja koji realno jeste efikasan, zbog greške tipa 2, nego da investiramo u neefikasan postupak zbog greške tipa 1.

S druge strane, da smo kojim slučajem operater radara protivvazdušne odbrane tokom rata, i da je naš posao da uzbunimo protivvazdušnu odbranu kada otkrijemo neprijateljske avione, dronove ili rakete, bilo bi mnogo opasnije da napravimo grešku tipa 2, tj da ne primetimo nailazeće neprijateljske avione, nego da napravimo grešku tipa 1 tj. da podignemo lažnu uzbunu, objavljujući da smo otkrili neprijateljske avione kada aviona realno nema.

Činjenje greške tipa 2 bi u ovoj situaciji verovatno rezultiralo materijalnom štetom i smrću naših kolega, prijatelja ili članova porodice, pa i našom sopstvenom, dok bi činjenje greške tipa 1 rezultiralo samo time da nekoliko kolega bude bespotrebno uznemireno i zato ne baš zadovoljno.

PREPRINT

Prema tome, **kada odlučujemo o pragu statističke značajnosti, veoma je važno proceniti koju od ovih vrsta grešaka je važnije da izbegnemo**. **Ako odlučimo da je važnije da izbegnemo greške tipa 1, onda kritičnu vrednost statističke značajnosti treba da pomerimo ka nižim p vrednostima. Ako je važnije da izbegnemo grešku tipa 2, onda bi trebalo da pomerimo prag statističke značajnosti prema višim p vrednostima.**

Tabela 5.5. Shematski prikaz dve vrste statističkih grešaka i dve vrste ispravnih zaključaka

Statističke greške		Mislimo li da postoji efekat?	
		Da	Ne
Postoji li efekat stvarno?	Da	Odlično! **Naš zaključak je tačan!!** Odbacili smo pogrešnu nultu hipotezu!	**Statistička greška tipa 2.** Efekat stvarno postoji, ali mi mislimo da ga nema. **Prihvatili smo nultu hipotezu koja nije tačna!**
	Ne	**Statistička greška tipa 1.** Efekta nema, ali mi mislimo da ga ima. **Odbacili smo nultu hipotezu koja je tačna!!**	Odlično! **Naš zaključak je tačan!!** Prihvatili smo tačnu nultu hipotezu!

Konačno, **ono što možemo da uradimo da smanjimo verovatnoću obe vrste grešaka je da suzimo distribuciju uzorkovanja u celini, da snizimo standardnu grešku**. Kao što je ranije rečeno, relativna širina distribucije uzorkovanja zavisi od varijabilnosti varijable koju razmatramo i od veličine uzorka. Kako je varijabilnost tolika kolika je, tj. kako ne možemo da smanjimo varijabilnost bez da narušimo reprezentativnost uzorka, **ono što možemo da uradimo je da povećamo veličinu uzorka**. **Povećavanje veličine uzorka će smanjiti standardnu grešku**, a time će i suziti distribuciju uzorkovanja i **tako učiniti da manja odstupanja od stanja definisanog nultom hipotezom pređu prag statističke značajnosti** (kritičnu vrednost statističke značajnosti), **smanjujući javljanje i greške tipa 1 i greške tipa 2** (ali ne menjajući relativne odnose njihovih učestalosti – ne menjajući to koliko se puta jedna javlja češće nego druga). Sve u

PREPRINT

svemu, **ako hoćemo da smanjimo učestalost javljanja statističkih grešaka, treba da povečamo veličinu uzorka.** S druge strane, povećavanje veličine uzorka takođe povećava i cenu realizacije istraživanja. Ako povećanje veličine uzorka nije zapravo bilo potrebno, samo smo bespotrebno protraćili novac na skupljanje podataka od dodatnih entiteta koji nam nisu stvarno trebali. Međutim, ako se pokaže da je naš uzorak bio previše mali da bi se efekat koji želimo da proučimo mogao pouzdano detektovati (zato što, setimo se, razlika između vrednosti populacije i vrednosti pretpostavljenih nultom hipotezom treba da bude određene veličine da bi bilo dovoljno verovatno da statistici uzorka pređu usvojeni prag statističke značajnosti) i onda zbog toga počinimo statističku grešku tipa 2 (bez da to znamo, naravno, nevolja sa ovim greškama što kada ih činimo, ne znamo da to radimo!), onda su novac i trud koji smo uložili u celu studiju propali. To je razlog zašto istraživači **koriste seriju postupaka koje zovu analizom snage (eng. power analysis)** sa **ciljem utvrđivanja minimalne veličine uzorka koja je potrebna da bismo imali određenu verovatnoću (obično 80% ili 90%) da detektujemo efekat** (tj. razliku od stanja definisanog nultom hipotezom) **određene očekivane veličine koristeći određeni prag statističke značajnosti** (određenu kritičnu vrednost za odbacivanje nulte hipoteze). **Ova veličina se naziva snagom studije i tipično se izražava kao proporcija testova/istraživanja koji bi dali statistički značajan rezultat** (tj. koji bi prešli usvojeni nivo statističke značajnosti, opravdavajući tako odbacivanje nulte hipoteze) **ako očekivatni efekat postoji u populaciji**. Na ovaj način, snaga od 0,92 znači da će 92 od 100 testova biti statistički značajno ako očekivani efekat (efekat očekivane veličine) postoji u populaciji. **Postupci analize snage tipično kombinuju informacije o veličini očekivanog efekta, željenom pragu statističke značajnosti** (za odlučivanje o prihvatanju/odbacivanju nulte hipoteze) **i željenoj verovatnoći da uzorak iz populacije u kojoj postoji efekat očekivane veličine pređe prag statističke značajnosti da bi izračunali minimalnu potrebnu veličinu uzorka.** Tipična pretpostavka u postupcima analize snage je da je uzorak izvučen iz populacije postupkom slučajnog uzorkovanja. Treba reći i da su veličine uzoraka koje se tipično viđaju u objavljenim istraživanjima u vodećim naučnim publikacijama u oblasti društenih nauka i nauka o ponašanju uglavnom dosta, dosta veće od minimalnih veličina na koje bi ukazivali rezultati analize snage. Zbog toga su postupci analize snage najkorisniji onda kada nije moguće skupiti veliki uzorak, bilo zbog toga što su entiteti (ili učesnici u istraživanju, onda kada su ti entiteti ljudi) koji nam trebaju retki ili zato što su postupci merenja koje koristimo komplikovani, skupi ili vrlo ekstenzivni. Drugim rečima, **analiza snage je najkorisnija onda**

PREPRINT

kada smo ograničeni na istraživanja na veoma malim uzorcima. **Kada je uzorak veliki, čak i veoma mali efekti će lako preći prag statističke značajnosti, a onda se javlja problem odlučivanja o tome koji od detektovanih efekata** (tj. detektovanih odstupanja od stanja pretpostavljenog nultom hipotezom) **zaista vrede daljeg razmatranja**, a koji se mogu zanemariti jer su previše mali da bi imali ikakav praktičan značaj.

****

Iako je statistička značajnost bila centralna tačka statistike zaključivanja i testiranja hipoteza u društvenim naukama i naukama o ponašanju tokom većeg dela 20. veka, a tu poziciju je zadržala i u prvim decenijama 21. veka**, statistička značajnost je bila i meta mnogih kritika** koje su dovele do toga da istraživači počnu da istražuju i primenjuju alternative isključivom oslanjanju na statističku značajnost kod izvođenja zaključaka o rezultatima istraživanja.

Jedan od **glavnih nedostataka** koncepta statističke značajnosti je činjenica da **njena vrednost zavisi od veličine uzorka**. Zbog toga, kada je uzorak mali, čak i velika odstupanja od nulte hipoteze (tj. veliki efekti) mogu da ostanu statistički neznačajni (odnosno da ne pređu kritičnu vrednost statističke značajnosti za odbacivanje nulte hipoteze), a tako i neotkriveni. S druge strane, kada je uzorak veliki, čak i zanemarljive razlike, razlike bez ikakovog praktičnog značaja, lako postaju statistički značajne. Ovaj nedostatak je toliko važan da su neki istraživači uvredljivo počekli da računanje statističke značajnosti nazivaju „testiranje veličine uzorka“ ili „nulti ritual“ (e.g. Gigerenzer, 2004). Ovakva situacija stvara i savrešenu priliku za nepoštene pristupe istraživanju – ako bi istraživač varalica hteo da pokaže da određeni efekat ne postoji, on/ona bi mogao/mogla da uradi istraživanje na uzorku koji je previše mali i da onda izvesti kako efekat nije detektovan (što bi onda bilo interpretirano kao da „ne postoji“). S druge strane, ako bi istraživač hteo da istakne efekat zanemarljive veličine, iz razloga koji nemaju veze sa stvaranjem naučnih znanja, on/ona bi mogao/mogla da uradi studiju na veoma, veoma velikom uzorku, a na takvom uzorku bi, dodajući tome i mogućnost da se radi veliki broj poređenja, bilo sasvim verovatno da bi nešto, neko poređenje prešlo prihvaćeni prag statističke značajnosti. Takav istraživač bi onda mogao da zgodno „zaboravi“ da prokomentariše veličinu odstupanja od nulte hipoteze i da se fokusira samo na činjenicu da je dobijen statistički značajan efekat.

PREPRINT

Ovaj poslednji primer dovodi nas do **drugog važnog nedostatka isključivog oslanjanja na statističku značajnost prilikom izvođenja zaključaka o rezultatima istraživanja – statistička značajnost, kada se koristi onako kako je napred opisano, za rezultat daje isključivo binarne odluke – da se nulta hipoteza odbaci ili prihvati.** Međutim, odbacivanje nulte hipoteze nam ne kaže ništa o verovatnoj veličini razlike između proučavane populacije i vrednosti pretpostavljenih nultom hipotezom, već nam samo omogućava da zaključimo da razlike ima (ili da nema, ako nultu hipotezu odlučimo da prihvatimo). Ovo nekada nije dovoljno zato što, kako je već pomenuto, istraživanja na velikim uzorcima mogu da dovedu do toga da usvojeni prag statističke značajnosti pređu razlike koje su previše male da bi bile od ikakvog praktičnog značaja.

Zbog toga je, u današnjoj nauci, **veoma široko prihvaćena preporuka da se uz statističku značajnost prikaže i neka mera veličine efekta** (tj. mera veličine razlike između dobijenih vrednosti i onih koje su pretpostavljene nultom hipotezom). U trenutku pisanja ove knjige, 2022. godine, korišćenje mera veličine efekta je široko prihvaćeno, a pri tom su u naučnoj literaturi prisutni i pozivi da se u potpunosti napusti korišćenje statističke značajnosti za potrebe izvođenja zaključaka o rezultatima istraživanja. Ipak, uprkos ovim pozivima, korišćenje statističke značajnosti je i dalje glavni postupak izvođenja zaključaka o rezultatima istraživanja. Naravno, kako su istraživači sve svesniji ovih kritika, **zaključci se sve ređe zasnivaju samo na statističkoj značajnosti, već se sve češće kombinuju sa merama veličine efekta.**

Postoji i jedan pravac kritike koji gađa konceptualnu osnovu testiranja nulte hipoteze preko statističke značajnosti, na način koji je upravo opisan i koji testiranje statističke značajnosti pogrdno naziva „testiranje značajnosti nulte hipoteze“ ili na engleskom NHST, od „null hypothesis significance testing“ (e.g. Perezgonzalez, 2015). Ovi autori valjano primećuju da je postupak testiranja nulte hipoteze postupak koji se često koristi u društvenim naukama, biologiji, psihologiji i drugim oblastima tj. postupak koji je opisan u ovoj knjizi, zapravo kombinacija dva pristupa testiranju podataka – jednog koji je predložio Fišer (Fischer) i jednog koji su predložili Nejman i Pirson (Neyman, Pearson) u prvom polovini 20. veka (Gigerenzer, 2004; Haig, 2017; Lakić, 2019; Perezgonzalez, 2015). Oni primećuju da su, u okviru ovog pristupa, neka svojstva ova dva teorijska pristupa kombinovana na način za koji oni smatraju da nije u skladu sa karakteristikama ova dva pristupa, a da su neke druge stvari veoma uprošćene, počevši od same nulte hipoteze, koja je, manje-više, svedena na hipotezu koja kaže da je parametar nula, da alternativna hipoteza obično

PREPRINT

nije ni formulisana, a da istraživači ovaj pristup često posmatraju kao da je u pitanju neki ritual i primenjuju ga bez da stvarno razumeju koncepte na kojima su ove procedure zasnovane. Kažu da ovi istraživači kritične nivoe statističke značajnosti (to je najčešće nivo od 0,05) tretiraju, u praksi i u onom kako uče studente i istraživače početnike, potpuno nefleksibilno, kao krute pragove, umesto kao manje ili više arbitrarne prelomne tačke, kao i da su mnoge stvari koje istraživači rade prosto pogrešne (kao, na primer, interpretiranje statistički neznačajnih rezultata kao „tendencije" prema stanju stvari koje bi dalo osnova za odbacivanje nulte hipoteze). Vrlo detaljan pregled sličnosti i razlika originalnog Fišerovog i Nejman-Pirsonovg pristupa s jedne strane i „testiranja značajnosti nulte hipoteze", onako kako se ovaj pristup primenjuje u istraživačkoj praksi dao je Perezgonzalez (2015). U datom radu autor daje detaljni pregled sva tri pristupa i tabelu u kojoj detaljno navodi sličnosti i razlike između njih po ključnim tačkama.

Iako su mnoge tačke ove kritike sasvim validne na konceptualnom nivou, ono što se možemo zapitati je da li one zaista diskvalifikuju „testiranje značajnosti nulte hipoteze" kao pristup koji je „dovoljno dobar" za praktičnu upotrebu u istraživanjima. Po mišljenju autora ove knjige, odgovor na to pitanje je veoma jasno ne. Iako, zaista, opisani postupci imaju puno nedostataka, od kojih su najvažniji već pomenuti u ovoj knjizi, oni su takođe zaslužni za uvođenje i širenje upotrebe statističkih i matematičkih postupaka u oblastima nauke u kojima su pre toga bile mnogo, mnogo manje pristune, čime su u potpunosti promenile metodološki pristup istraživanjima u ovim oblastima. Njihova srazmerna jednostavnost je omogućila da integraciju ovih postupaka u studijske programe naučnih oblasti koje ranije nisu praktično koristile nikakvu matematiku i dozvolile naučnicima iz oblasti koje tradicionalno nisu mnogo bliske matematici da usvoje i koriste relativno složene statističke postupke u svojim istraživanjima. Uopšte nije preterivanje ako kažemo da je uvođenje statističkih postupaka, počevši od onih koje su opisane u ovoj knjizi, napravilo revoluciju u načinu na koji se sprovode istraživanja u društvenim naukama, u psihologiji, ali takođe i u biologiji i mnogim drugim oblastima nauke. Ovo se najbolje može videti kada uporedimo naučni kvalitet i naučnu vrednost istraživanja koja rade istraživači koji dolaze iz sredina gde se ova integracija statistike u istraživačku metodologiju desila pre više decenija ili čak pre više od veka sa istraživanjima koja rade istrživači koji dolaze iz sredina gde se ova integracija nije desila uopšte ili u kojima je početak te integracije relativno novijeg datuma. Iako ima glasova koji iznose mišljenje da su događaji poput krize replikabilnosti nalaza u psihologiji (ovu krizu je pokrenula serija skorašnjih istraživanja koja su imala za cilj da ponove klasične/čuvene

PREPRINT

istraživačke postupke iz oblasti psihologije i koja nisu uspela da dobiju rezultate kakve su prijavili njihovi istorijski uzori, a i oni koji jesu, prijavili su uglavnom da su dobijeni efekti slabiji nego u istraživanjima koja su im bili uzori) (e.g. Świątkowski & Dompnier, 2017) delom i posledica manjkavosti u kvalitetu primene statistike zaključivanja u psihološkim istraživanjima i uopšte manjkavosti u razumevanju statističkih koncepata od strane psihologa (Lakić, 2019), mi dosta čvrsto verujemo da bez uprošćavanja koja su uvedena u praktičnu primenu statistika od strane istraživača, ne bi bilo ni široke primene statistike u istraživanjima u ovim oblastima, pa onda ne bi bilo ni replikabilnosti. Valja primetiti da da bi imali krizu replikabilnosti, prvo moramo da imamo rezultate istraživanja koji su replikabilni, a ovakvih rezultata ne bi bilo da istraživači nisu usvojili statistiku. Iz ovog sledi da je, uprkos postojećim i potencijalnim nedostacima statističkih postupaka koji su trenutno u širokoj upotrebi, jednostavno neosporna činjenica to da je uvođenje statistike napravilo revoluciju i u velikoj meri unapredilo kvalitet istraživanja u ovim oblastima nauke. I pored ovoga, statistički postupci su stvari koje se razvijaju i mesta za unapređenje definitivno ima. Neka od mogućih unapređenja su predstavljena i ovde. Takođe, treba reći i da su neke od kritika koje su upućene ovim statističkim postupcima prosto neopravdane. Na primer, kritika o rigidnosti postupaka koji se primenjuju u naučnim istraživanjima, a koja kaže da se od svih naučnika traži da primenjuju iste postupke i ista pravila za donošenje odluka, uključujući i kritični nivo statističke značajnosti, a koji neki kritičari pogrdno nazivaju „nulti ritual“ ili bezumnom upotrebom statistike (Gigerenzer, 2004), zapravo predstavljaju poželjnu karakteristiku standardizacije naučnih postupaka i podsticanje objektivnosti. Standardizovane naučne procedure i objektivnost su svojstva koja su od izuzetne važnosti u nauci uopšte, ali njihova važnost još više dolazi do izražaja u oblasti društvenih nauka, oblasti koja je kroz svoju istoriju bila često pod velikim uticajem, a nekada i sasvim „oteta“ od strane različitih političkih ideologija i planova, neobjavljenih ličnih verovanja i opšteg subjektivizma do te mere da je to zaustavljalo svaki naučni napredak i dovodilo do potpuno bezvrednih istraživanja (e.g. V. Hedrih, 2020; Reyna, 2017). U takvom okruženju, standardizacija istraživačkih postupaka na način koji obezbeđuje da svi istraživači izvedu iste zaključke iz istog skupa podataka i istog skupa statističkih postupaka je važna prednost, a ne nedostatak, čak i kada je cena koju treba platiti za to nešto povećana teorijska neusklađenost u odnosu na nivo pre toga. Ovo, naravno, ne treba nikako shvatiti kao stav da nema prostora za unapređivanje postupaka testiranja nulte hipoteze koji su trenutno u širokoj upotrebi. Upravo

PREPRINT

suprotno – mnoga takva unapređenja su veća postala ili uveliko postaju deo dominantne naučne prakse.

Testiranje nulte hipoteze reuzorkovanjem. Postupci za računanje statističke značajnosti su se tradicionalno oslanjali na centralnu graničnu teoremu, što znači da su svojstva distribucije uzorkovanja procenjivana formulama, kao što je opisano u prethodnom poglavlju. Jedna alternativa ovom pristupu, koja u trenutku pisanja ove knjige polako ulazi u široku upotrebu preko različitih statističkih softverskih paketa je testiranje nulte hipoteze preko postupka butstrepinga. **Testiranje nulte hipoteze pomoću butstrepinga** radi na način koji je sličan načinu na koji koriste intervali poverenja da bi se procenila vrednost parametra:

- U prvom koraku se **uzorkovanjem sa vraćanjem uzima željeni broj uzoraka iz izvornog uzorka istraživanja,** a onda se, **na svakom od tih uzoraka, izračunaju statistici na kojima treba da se testira nulta hipoteza.**
- **Od ovako izračunatih statistika formira se distribucija uzorkovanja.**
- **Granične tačke intervala poverenja distribucije uzorkovanja definišu se na osnovu željenog praga tj. na osnovu toga koliki procenat distribucije uzorkovanja želimo da obuhvatimo intervalom poverenja.** U naučnoj i stručnoj literaturi se **najčešće koriste 95% i 99% interval**. Za ove intervale se može smatrati da **odgovaraju nivoima statističke značajnosti od 5% i 1% tj. od 0,05 i 0,01**, zato što kada je 95% distribucije uzorkovanja uključeno u interval poverenja, 5% distribucije nije uključeno u njega, a taj procenat/proporcija entiteta koja je van intervala poverenja je u suštini ono što statistička značajnost predstavlja. Treba primetiti i da se ovaj interval poverenja pravi tako da su gornja i donja granica jednako udaljene od aritmetičke sredine distribucije uzorkovanja. Ovaj interval poverenja se obično naziva **butstrep interval poverenja uz navođenje procenta distribucije uzorkovanja koji taj interval uključuje**. Na primer, ako napravimo butstrep interval poverenja tako da uključuje 95% entiteta iz distribucije uzorkovanja, takav interval se zove 95% butstrep interval poverenja, a ako se napravi tako da uključuje 99% entiteta, onda se zove 99% butstrep interval poverenja i tako dalje. U ovom postupku se **obično računa i razlika između aritmetičke sredine distribucije uzorkovanja i aritmetičke sredine izvornog istraživakog uzorka** (dakle uzorka koji smo

PREPRINT

empirijski prikupili i iz kog smo radili uzorkovanje u ovkiru postupka butstrepinga) i ova razlika se zove **bias ili pristrasnost.**

- **Nulta hipoteza** se onda procenjuje **uočavanjem da li se vrednost koju pretpostavlja nulta hipoteza (a to je obično 0) nalazi unutar intervala poverenja koji je napravljen u prethodnom koraku ili se nalazi van njega**. **Ako se vrednost određena nultom hipotezom nalazi unutar intervala poverenja, onda se nulta hipoteza prihvata. Ako se vrednost specifikovana nultom hipotezom nalazi van intervala poverenja, onda se nulta hipoteza odbacuje.**

Na primer, ako bi želeli da testiramo nultu hipotezu da dva uzorka potiču iz populacija sa istom vrednošću aritmetičke sredine (na nekoj varijabli koju smo proučavali), prvo bi postavili nultu hipotezu da je razlika između aritmetičkih sredina dve populacije iz kojih smo izvadili naše uzorke jednaka 0. Onda bi mogli da sprovedemo postupak butstrepinga i napravimo, recimo, 10 000 parova uzoraka iz ova izvorna dva uzorka, a onda da izračunamo razliku između aritmetičkih sredina za svaki od ovih 10 000 parova uzoraka. Tako bismo dobili 10 000 razlika između aritmetičkih sredina i one bi činile našu distribuciju uzorkovanja razlika između aritmetičkih sredina. Onda bismo definisali, recimo, 95% butstrep interval poverenja, takav da su njegova gornja i donja granica podjednako udaljene od centra distribucije uzorkovanja. Pretpostavimo sada da smo napravili ovaj interval i da on obuhvata vrednosti između, recimo, -1 i 3, pri čemu je aritmetička sredina distribucije uzorkovanja 1. Kako naša nulta hipoteza kaže da je razlika između aritmetičkih sredina populacija 0, mi bi onda pregledali interval poverenja koji smo napravili i uočili da se 0 nalazi unutar našeg intervala poverenja. Na osnovu ovoga, mi bismo zaključili da nulta hipoteza nije opovrgnuta, što znači da bismo je prihvatili. Ako bi, međutim, ispalo da naš interval poverenja obuhvata vrednosti između, recimo -3 i -1 ili na primer između 2 i 5, onda bi zaključili da 0 (vrednost pretpostavljena nultom hipotezom) nije unutar butstrep intervala poverenja i, na osnovu toga, odbacili nultu hipotezu. Ovde je još važno ponovo istaći da, iako su nulte hipoteze koje kažu da je neka vrednost nula (tj. „nil“ hipoteze) praktično jedina vrsta nultih hipoteza koja se sreće u stručnoj i naučnoj literaturi i statističkom sofveru, bilo koja precizna statistička hipoteza (tj. bilo koja hipoteza koja precizno pretpostavlja vrednosti parametara) može podjednako dobro da odigra ulogu nulte hipoteze.

PREPRINT

Kao i kod korišćenja butstrep procedure za procenu vrednosti parametara, **prednost butstrepinga u poređenju sa računanjem statističke značajnosti na osnovu centralne granične teoreme sastoji se u tome da se ovaj postupak ne oslanja na pretpostavke o svojstvima distribucije uzorkovanja, već direktno računa svojstva iz simulirane distribucije uzorkovanja napravljene izvlačenjem uzoraka iz izvornog uzorka datog istraživanja**.

5.5. Bajesov faktor i testiranje statističkih hipoteza

Jedno važno svojstvo testiranja nulte hipoteze preko računanja statističke značajnosti, a takođe i jedna od glavnih tačaka kritike ovog postupka je činjenica da se, onako kako se tipično koristi, **postupak testiranja statističke značajnosti ni na koji način ne oslanja na postojeća teorijska znanja o temi koja se istražuje** (e.g. Dienes, 2014; Dienes & Mclatchie, 2018; Gigerenzer, 2004; Lakić, 2019). Istraživač postavlja nultu hipotezu, hipotezu za koju očekuje da je pogrešna, ali, u okviru statističkog postupka, nema nikakve specifikacije onoga što istraživač zapravo očekuje. Iako je ovaj postupak svakako ispravan kada nemamo nikakvih prethodnih znanja o fenomenu koji se proučava, pa tako ni osnova za neka posebna očekivanja, kada su u pitanju pojave koje su dobro proučene, istraživači često itekako imaju osnova da formulišu i testiraju specifične hipoteze, samo što se tako nešto radi vrlo retko. Takođe, **kada nivo statističke značajnnosti ukazuje da treba da prihvatimo nultu hipotezu, to ne znači da je verovatno da je ta hipoteza tačna, nego samo to da nije bilo moguće odbaciti je na osnovu rezultat primenjenog statističkog postupka**. Da bi prevazišli ovaj problem, neki istraživači predlažu da se, **umesto računanja statističke značajnosti nulte hipoteze, porede konkurentske hipoteze,** a jedan način da se to uradi je **računanje Bajesovog faktora.** Postupak za računanje Bajesovog faktora je deo pristupa statistici koji se zove **Bajesijanska statistika.** Bajesijanska statistika je zasnovana na **Bajesovoj interpretaciji verovatnoće, gde se verovatnoća posmatra kao izraz stepena uverenosti u određeni događaj**. Bajesijanska statistika je dobila ima po engleskom statističaru iz 18. veka Tomasu Bajesu (Thomas Bayes), koji je autor koncepta koji danas nazivamo Bajesovom teoremom. Ovo polje je kasnije dodatno razvijeno radovima francuskog matematičara Pjer-Simona Laplasa (Pierre-Simon Laplace).

Ovaj konkretni pristup testiranju statističkih hipoteza **počinje formulisanjem konkurentskih hipoteza, hipoteza od kojih se jedna može smatrati za nultu hipotezu, a druga za alternativnu**

PREPRINT

hipotezu, ali te dve konkurentske hipoteze mogu prosto i biti dve konkurentske hipoteze proistekle iz nesaglasnih rezultata ranijih istraživanja (gde su neka istraživanja pokazivala jedno, a druga nešto drugo). **Ove hipoteze pretpostavljaju da je parametar jednak određenoj vrednosti ili da može imati određeni raspon vrednosti**. Na osnovu toga, **za svaku hipotezu se pravi distribucija verovatnoća dobijanja različitih vrednosti statistika na uzorku ako je data hipoteza tačna.** Ove distribucije verovatnoća mogu napraviti istraživači, a mogu se praviti i određenim statističkim postupcima i **one se prave pre nego što se počne sa prikupljanjem podataka** tj. ispitivanjem uzorka našeg istraživanja. **Zbog toga što se prave pre prikupljanja podataka, ove verovatnoće se zovu apriorne verovatnoće**. Ovo potiče od latinskog izraza „a priori", koji znači „od prethodnog", a koristi se da označi situacije kada se zaključci izvode bez prethodno pribavljenih empirijskih podataka, kada su zaključci unapred izvedeni rezonovanjem naspram toga da budu izvedeni iz prikupljenih podataka.

Nakon prikupljanja podataka, istraživač računa tačnu verovatnoću dobijanja rezultata koji je dobijen na uzorku u situaciji kada je prva hipoteza tačna (u populaciji), **a onda verovatnoću dobijanja takvih rezultata kada je druga hipoteza tačna** (u populaciji). **Deljenje prve verovatnoće** (verovatnoće dobijanja rezultata koji je dobijen na uzorku u situaciji kada je prva hipoteza tačna) **drugom verovatnoćom** (verovatnoća dobijanja na uzorku rezultata koji je dobijen ako je druga hipoteza tačna) **rezultira veličinom koja se zove Bajesov faktor**, a tipično se označava sa **BG**. **Vrednost Bajesovog faktora od 1 označava da su obe hipoteze podjednako verovatne imajući u vidu raspoložive podatke** (podatke iz istraživačkog uzorka koji smo ispitali). **Vrednosti Bajesovog faktora koje su veće od jedan ukazuju da je prva hipoteza** (ona koja ima ulogu imenioca u brojioca u deljenju za računanje BF) **verovatnija, a vrednosti Bajesovog faktora manje od 1 ukazuju da je verovatnije da je druga hipoteza tačna** (ona hipoteza koja ima ulogu imeniocu u deljenju za računanje BF). Kada je u pitanju **prag za interpretaciju rezultata**, jedna često sretana preporuka je da se **vrednosti BF iznad 1, pa do 3 tretiraju kao neubedljive**, kao vrednosti iz kojih se ne mogu izvući zaključci o tome koja od hipoteza je verovatnija, v**rednosti između 3 i 10 se smatraju za umereno snažan rezultat u prilog prve hipoteze**, rezultat kakav bi se očekivao u ranim fazama istraživanja. Vrednosti **10-30 se smatraju za snažan rezultat u prilog prve hipoteze**, vrednosti **od 30-100 su veoma snažan rezultat u korist prve hipoteze**, dok se vrednosti BF **preko 100 smatraju za ekstremno jake**, **krucijalne nalaze u prilog prve hipoteze**. Na sličan način, **vrednosti BF ispod 1, pa sve do 1/3**

(ili 0,33) smatraju se za neubedljive, nalaze iz kojih se ne mogu izvući zaključci o preimućstvu neke od hipoteza, **vrednosti između 0,33 i 0,1 se smatraju umerenim nalazima u korist hipoteze 2** (hipoteze koja predstavlja imenilac prilikom računanja Bajesovog faktora), **vrednosti između 0,1 i 0,033 predstavljaju snažen nalaze, između 0,033 i 0,01 veoma snažne nalaze i vrednosti ispod 0,01 predstavljaju ekstremno snažne, krucijalne nalaze u korist druge hipoteze** (Lakić, 2019; Nicenboim et al., 2021).

Ključno svojstvo ovog pristupa je formulacija hipoteza, tj. određivanje apriornih verovatnoća. Ovo se može uraditi na različite načine, od toga da proglasimo sve moguće vrednosti jednako verovatnim (što znači da napravimo uniformnu distribuciju apriornih verovatnoća), pa do veoma preciznih distribucija verovatnoća zasnovanih na prethodnom znanju. Ovo, po samoj svojoj prirodi, **zahteva oslanjanje na sud istraživača zarad formulisanja hipoteza i samim tim možemo napraviti situacije gde različiti istraživači ove verovatnoće formulišu na različite načine, sve imajući u vidu iste podatke dostupne pre početka istraživanja, što ovaj postupak čini subjektivnim.** Kako Lakić primećuje u svom radu o Bajesovom faktoru (Lakić, 2019), računanje Bajesovog faktora se može uraditi i na objektivan način, recimo dodeljivanjem jednakih verovatnoća svim mogućim vrednostima razmatranog statistika. Međutim, može se reći da ovakav postupak poništava osnovnu ideju ovog pristupa, a to je uključivanje prethodnog znanja o temi koja se proučava u proračune. Ova neophodna subjektivnost ovog pristupa može takođe, potencijalno, biti zloupotrebljena od strane istraživača previše motivisanih da pokažu da je njihova hipoteza tačna, a koji bi svoje hipoteze formulisali tek nakon što prikupe podatke (i dobro ih prouče), a ne pre toga kako ovaj pristup zahteva (a pravili se da su ih formulisali pre!!). Međutim, **pristup zasnovan na Bajesovom faktoru vidi istraživanje kao dinamički proces, a ne kao jednokratni događaj**. Ovo znači da čak iako bi neki istraživač uradio ovako nešto, budući istraživači i buduća istraživanja koja budu proučavala isti skup prethodnih podataka bi lako dovela u pitanje validnost ovakvih nalaza. Još jedna važna stvar koju treba istaći je to da **nam Bajesov faktor sam za sebe ne može reći koja je hipoteza najverovatnija, već samo, imajući u vidu nalaze našeg istraživanja i podatke iz prethodnih istraživanja, koliko bi i kako trebalo da ažuriramo stepen svoje relativne uverenosti u razmatrane konkurentske hipoteze** (Nicenboim et al., 2021).

PREPRINT

Imajući sve ovo u vidu, **glavno svojstvo Bajesijanskog pristupa – oslanjanje na sud istraživača prilikom procene apriornih verovatnoća izgleda u isto vreme i kao njegov glavni nedostatak, jer ono čini ovaj pristup kompleksnijim i težim za istraživače koji nisu previše vični statistici**. Upravo to je verovatno i jedan od razloga za relativno skroman nivo upotrebe ovog pristupa u trenutno objavljenim istraživanjima u oblasti društvenih nauka i nauka o ponašanju.

5.6. Parametrijski i neparametrijski statistici

U ovoj tački, bilo bi korisno da povučemo razliku između dve vrste statističkih postupaka koji se razlikuju po tome da li polaze od pretpostavke o tome kako su podaci distribuirani ili ne polaze. **Statistički posutpci koji polaze od pretpostavke o tome kakva je distribucija podataka, a moguće i od drugih pretpostavki nazivaju se parametrijskim postupcima. Oni koji ne polaze od pretpostavki o obliku distribucije podataka nazivaju se neparametrijskim postupcima.**

Parametrijski postupci koji se najčešće sreću u naučnoj literaturi u oblasti društvenih nauka i nauka o ponašanju **polaze od pretpostavke da je nešto normalno distribuirano**. To može biti **distribucija uzorkovanja**, ali, u složenijim statističkim postupcima, ova pretpostavka se može odnositi i na druge stvari. Pored ovog, ovi postupci obično takođe polaze i od sledećih pretpostavki:

- **Da su podaci bar na intervalnom nivou merenja** – ovo očekivanje ide zajedno sa očekivanjem da distribucija bude normalna. Normalna distribucija je distribucija podataka koji su bar na intervalnom nivou merenja, jer se oslanja na postojanje fiksne jedinice mere. Bez fiksne jedinice mere, ne bi bilo moguće ustanoviti da li je oblik distribucije normalan ili nije tj. da li su vrednosti entiteta grupisane onako kako predviđa normalna distribucija. Isto tako, ako podaci ne bi bili bar intervalni, ne bi bilo moguće izračunati deskriptivne statističke pokazatelje za normalno distribuirane podatke – i aritmetička sredina i standardna devijacija zahtevaju da podaci budu bar na intervalnom nivou merenja. Iz tog razloga, ako su podaci na ordinalnom ili nominalnom nivou merenja, ne možemo govoriti o normalnoj distribuciji.
- **Da podaci iz uzorka potiču iz nezavisnih posmatranja/merenja**. Ovo u osnovi znači da vrednosti jednog entiteta ne utiču na vrednosti bilo kog drugog entiteta u uzorku i da

PREPRINT

različiti podaci potiču iz različitih merenja, te da nema višestrukih kopija istog merenja predstavljenih kao da su različita merenja. U slučaju društvenim nauka i nauka o ponašanju, onda kada su ljudi entiteti od kojih se skupljaju podaci, a pri čemu su podaci odgovori na različita pitanje ili testovne stimuluse, ovo najčešće treba interpretirati kao zahtev da su učesnici u istrživanju odgovarali samostalno i nezavisno jedan od drugog bez dogovaranja o tome šta da odgovore, ali i bez da su imali instrukcije šta da odgovore.

- **Da su varijanse homogene u posmatranim grupama** – ova pretpostavka u osnovi znači da se **razlike između grupa, ako postoje, sastoje u razlikama u prosečnim nivoima izraženosti merenih varijabli, a ne u razlikama u varijabilnosti ovih varijabli**. Ovo takođe imlicira i da, kada testiramo efekte različitih faktora, pretpostavljamo da ovi faktori utiču na sve entitete na otprilike isti način, stvarajući tako razlike između aritmetičkih sredina grupa koje su bile i onih koje nisu bile izložene tim faktorima, a ne između njihovih varijansi. Ako posmatramo više grupa, ova pretpostavka znači da sve grupe imaju istu varijansu. Ako posmatramo odnose između varijabli, ova pretpostavka znači da je varijansa jedne varijable ista na svim nivoima druge varijable.

Treba primetiti i da neki parametrijski statistički testovi imaju i varijante koje ne polaze od pretpostavki o jednakim varijansama u svim grupama, te kada primenjujemo ovakve varijante tih statističkih testova, običaj je da to naglasimo prilikom predstavljanja dobijenih rezultata.

Kada diskutujemo statističke postupke u kasnijem delu ove knjige, **uvek ćemo naglasiti da li je dati statistički postupak parametrijski ili neparametrijski**. Za većinu slučajeva obrade podataka, postoje i parametrijski i neparametrijski postupci. Na primer, postupak za procenu parametara populacije na osnovu podataka sa uzorka koji je zasnovan na centralnoj graničnoj teoremi se oslanja na formule za procenu standardne greške, koje se pak oslanjaju na pretpostavku da je distribucija uzorkovanja normalna, što ih čini parametrijskim postupcima. S druge strane, butstrep postupak za procenu parametara se ne oslanja na takve pretpostavke, pa je prema tome neparametrijski.

5.7. Hajde da primenimo šta smo do sada naučili!

PREPRINT

Hajde da probamo sada da primenimo stvari koje smo predstavili u ovom poglavlju kroz nekoliko vežbi. Molimo vas da pogledate opšte uputstvo za ovakve vežbe koje možete naći na početku knjige. Naša preporuka je da prvo pročitate svaki isečak i tvrdnje date u njemu i da onda date svoj odgovor. Odgovor možete upisati u kolonu za odgovore, a posle toga pročitajte odgovore i uporedite svoje odgovore sa njima.

Vežba I. Procena parametara, testiranje nulte hipoteze.

Tip profesionalnih interesovanja		AS	SD	t statistik	Statistička značajnost
Realni tip (R)	**Sportisti**	2,83	1,19	1,41	0,16
	Nesportisti	2,71	1,26		
Istraživački tip (I)	**Sportisti**	3,49	1,17	-0,68	0,50
	Nesportisti	3,55	1,23		
Umetnički tip (A)	**Sportisti**	3,35	1,52	-2,71	0,01
	Nesportisti	3,63	1,57		
Društveni tip (S)	**Sportisti**	4,04	1,12	-3,44	0,00
	Nesportisti	4,29	1,08		
Preduzetnički tip (E)	**Sportisti**	3,85	1,08	-0,15	0,88
	Nesportisti	3,86	1,03		
Konvencionalni tip (C)	**Sportisti**	3,20	1,10	1,58	0,11
	Nesportisti	3,08	1,14		

Nulta hipoteza testirana u ovim postupcima je da dva uzorka potiču iz populacija sa jednakim aritmetičkim sredinama! Postupak je parametrijski.

Molimo koristite 0,05 kao prag statističke značajnosti prilikom odgovaranja na tvrdnje!

U tabeli su predstavljene aritmetičke sredine i standardne devijacije dve grupe ljudi (sportista i nesportista) na seriji varijabli iz oblasti profesionalnih interesovanja. Statistički test koji je ovde sproveden testira nultu hipotezu da su aritmetičke sredine populacija iz kojih dve poređene grupe potiču na ispitivanoj varijabli jednake, a statistička značajnost je navedena u poslednjoj koloni (krajnja desna).

Tabela je zasnovana na podacima objavljenim u Hedrih et al. (2017)

I	Tvrdnja:	Odgovor
I1.	U ovom postupku bi nultu hipotezu trebalo odbaciti za varijablu R.	
I2.	U ovom postupku bi nultu hipotezu trebalo odbaciti za varijablu I.	
I3.	U ovom postupku bi nultu hipotezu trebalo odbaciti za varijablu A.	
I4.	U ovom postupku bi nultu hipotezu trebalo odbaciti za varijablu S.	

PREPRINT

I5.	U ovom postupku bi nultu hipotezu trebalo odbaciti za varijablu E.	
I6.	U ovom postupku bi nultu hipotezu trebalo odbaciti za varijablu C.	
I7.	Gornja granica 95% intervala poverenja aritmetičke sredine varijable R na grupi sportista bi bila viša od 6.	
I8.	Donja granica 95% intervala poverenja aritmetičke sredine varijable I na grupi sportista bi bila niža od 8.	
I9.	Skjunes distribucija i jedne i druge grupe na varijabli R je niži od 0,5.	
I10.	Bavljenje sportom dovodi do smanjenja izraženosti i umetničkog i društvenog tipa profesionalnih interesovanja i to je razlog zašto postoje statistički značajne razlike između aritmetičkih sredina dve grupe na ovim varijablama.	

Vežba J. Procena parametara, testiranje nulte hipoteze.

Descriptives

			Statistic	Std. Error
p39	Mean		3.93	.084
	95% Confidence Interval for Mean	Lower Bound	3.76	
		Upper Bound	4.09	
	5% Trimmed Mean		4.03	
	Median		4.00	
	Variance		1.691	
	Std. Deviation		1.300	
	Minimum		1	
	Maximum		5	
	Range		4	
	Interquartile Range		2	
	Skewness		-1.198	.157
	Kurtosis		.336	.312

Mean – aritmetička sredina. 95% Confidence Interval for Mean – 95% interval poverenja aritmetičke sredine. Lower Bound – donja granica, Upper Bound – gornja granica, Median – medijana. Variance – varijansa, Std. Error – standardna greška, Std. Deviation – standardna devijacija, Skewness – skjunes, Kurtosis – kurtozis, Statistic – vrednost statistika
U tabeli su predstavljeni deskriptivni statistički podaci i standardne greške odgovora na jednu stavku upitnika iz istraživanja koje su sproveli autori.

J	Tvrdnja:	Odgovor
J1.	Postoji 95% šanse da se vrednost aritmetičke sredine populacije na ovom uzorku na prikazanoj varijabli nalazi između 3,75 i 4,09.	
J2.	Standardna greška aritmetičke sredine varijable p39 je niža od 0,1.	

J3.	Standardna greška standardne devijacije varijable p39 je niža od 1.	
J4.	Varijansa varijable p39 je viša od 2.	
J5.	Aritmetička sredina žena na varijabli p39 je viša od aritmetičke sredine muškaraca.	
J6.	Percentilna devijacija varijable p39 je 4.	
J7.	Kvartilni srednji raspon varijable p39 je 2.	
J8.	Gornja granica 99% intervala poverenja aritmetičke sredine varijable p39 je niža od 4.	
J9.	P39 ima negativno asimetričnu distribuciju.	
J10.	50i percentil varijable p39 je 4.	

Vežba K. Procena parametara, testiranje nulte hipoteze.

			Rezultati butstrepinga, 1000 uzoraka			
	Statistik	Vrednost statistika uzorka	Bias	Standardna greška	95% confidence interval lower boundary	95% confidence interval upper boundary
WFC	Aritmetička sredina	2,62	0,00	0,04	2,54	2,69
	Medijana	2,4	0,06	0,09	2,4	2,6
	Standardna devijacija	1,19	0,00	0,02	1,15	1,23
	Skjunes	0,30	0,00	0,05	0,21	0,40
	Kurtozis	-0,97	0,00	0,06	-1,08	-0,85
FWC	Aritmetička sredina	1,72	0,00	0,03	1,67	1,77
	Medijana	1,4	0,02	0,06	1,4	1,6
	Standardna devijacija	0,83	0,00	0,02	0,78	0,87
	Skjunes	1,27	0,00	0,08	1,12	1,41
	Kurtozis	1,16	0,00	0,03	0,59	1,77

U tabeli su predstavljeni deskriptivni statistički podaci i intervali poverenja konflikta posao-porodica (WFC, od work-family conflict) i konflikta porodica-posao (FWC, od family work conflict) zasnovani na delu podataka iz Hedrih (2017a)

K	**Tvrdnja:**	**Odgovor**
K1.	Standardna devijacija distribucije uzorkovanja medijane varijable WFC je 0,06.	
K2.	Nema razlike između aritmetičke sredine uzorka na varijabli WFC i aritmetičke sredine distribucije uzorkovanja aritmetičke sredine ove varijable koja je dobijena postupkom butstrepinga.	
K3.	Vrlo je moguće da je vrednost kurtozisa populacije iz koje je uzet uzorak na WFC jednaka 0.	

PREPRINT

K4.	Glavni sredinski kvartil je veći za WFC nego za FWC.	
K5.	U uzorku ima više od 500 entiteta.	
K6.	Vrlo je verovatno da je distribucija FWC u populaciji iz koje je ovaj uzorak pozitivno asimetrična.	
K7.	Možemo očekivati da će aritmetička sredina varijable WFC biti između 2,54 i 2,69 u 95% ponovljenih studija na uzorcima iz populacije iz koje je i ovaj uzorak.	
K8.	Nema nikakve šanse da u bilo kom budućem istraživanju na uzorcima iz populacije iz koje je ovaj uzorak aritmetička sredina FWC bude niža od 1,7.	
K9.	Distribucija populacije iz koje je ovaj uzorak je verovatno platikurtična na WFC.	
K10.	Postupak koji je ovde predstavljen za zaključivanje o vrednostima parametra je zasnovan na centralnoj graničnoj teoremi.	

Pogledajmo sada i odgovore:

I1 – netačno. Možemo videte da je nivo statističke značajnosti 0,16. Ovo je više od našeg praga statističke značajnosti od 0,05 što ukazuje da ne bi trebalo da odbacimo nultu hipotezu.

I2 – netačno. Isto rezonovanje kao u I1. Nivo statističke značajnosti ovde je 0,5, što je više od 0,05.

I3 – tačno. Nivo statističke značajnosti od 0,01 je niže od 0,05, što znači da bi trebalo da odbacimo nultu hipotezu. U tipičnom naučnom žargonu, rekli bi da je rezultat "statistički značajniji" od prihvaćenog praga što opravdava odbacivanje nulte hipoteze.

I4 – tačno. Isto rezonovanje kao za I3, a rezultat postupka je "još statistički značajniji" nego što je to slučaj bio kod I3.

I5 – netačno. Nije statistički značajno. Statistička značajnost od 0,88 je daleko iznad praga od 0,05.

I6 – netačno. Statistička značajnost od 0,11 nije dovoljno niska za odbacivanje nulte hipoteze, imajući u vidu da nam je prag 0,05.

I7 – netačno. Iako nemamo intervale poverenja aritmetičke sredine koja je ovde prikazana, znamo da se gornja granica intervala poverenja određuje tako što se otprilike 2 standardne greške (1,96 ako hoćemo da budemo skroz precizni!!) dodaju na vrednost aritmetičke sredine. Takođe znamo i da se standardna greška aritmetičke sredine dobija tako što se standardna devijacija podeli kvadratnim korenom broja entiteta u uzorku. To znači da standardna greška aritmetičke sredine mora da bude manja od standardne devijacije. Takođe možemo da vidimo da je aritmetička sredina grupe sportista na ovoj varijabli 2,83, dok je standardna devijacija otprilike 1,2 (tj. 1,19). Dve standardne devijacije su otprilike 2,4. Ako dodamo 2,4 na 2,83 rezultat će zasigurno biti manji od 6, što znači da je nemoguće da gornja granica 95% intervala poverenja aritmetičke sredine bude veća od 6.

PREPRINT

I8 – tačno. Možemo videte da je aritmetička sredina grupe sportista na varijabli I 3,49. Donja granica intervala poverenja aritmetičke sredine mora da bude niža od aritmetičke sredine uzorka, pa kako je aritmetička sredina uzorka niža od 8, tako i donja granica intervala aritmetičke sredine mora takođe da bude niža od 8.

I9 – nepoznato. Nema podataka u tabeli koji bi nam omogućili da izvedemo zaključke o skjunesu. Da, postupak koji je korišćen (biće obrađen u kasnijem delu knjige) je parametrijski, što znači da zahteva normalnu distribuciju (doduše distribuciju uzorkovanja, ali to se tipično interpretira kao zahtev da i podaci budu normalno distribuirani!), ali je veličina skjunesa o kom je reč i dalje u okviru veličine odstupanja u kom se upotreba parametrijskih postupaka smatra prihvatljivom (ne ukazuje na odstupanje od normalne distribucije koje je dovoljno veliko da zbog toga ne mogu da se koriste parametrijski postupci).

I10 – nepoznato. Podaci u tabeli samo pokazuju da postoji razlika između dve grupe na ove dve varijable i da bi bilo opravdano zaključiti da razlike između aritmetičkih sredina postoje i u populacijama iz kojih su ove dve grupe uzorkovane. Međutim, nema podataka o uzrocima ovih razlika. Možemo zamisliti da bude moguće da bavljenje sportom menja profesionalna interesovanja, ali može biti u pitanju i to da se ljudi sa različitim profesionalnim interesovanjima razlikuju u pogledu toga koliko su zainteresovani za bavljenje sportom. Takođe je moguće i da neki drugi dovode do razlika između ljudi kako u bavljenju sportom, tako i u pogledu profesionalnih interesovanja. Podaci koji su predstavljeni u tabeli ne omogućavaju da se ustanovi koja od ovih mogućnosti je tačna, ako ijedna. Prema tome, ne možemo znati da li je ova tvrdnja tačna i ne možemo doneti sud o njenoj tačnosti na osnovu podataka koji su predstavljeni u tabeli.

J1 – tačno. To što tvrdnja opisuje je 95% interval poverenja aritmetičke sredine. Ovaj interval je predstavljen u tabeli i njegov raspon je tačan, što se može videte iz tabele.

J2 – tačno. Možemo videte da je standardna greška aritmetičke sredine 0,084, što je niže od 0,01.

J3 – tačno. Ovde nemamo prikazanu standardnu grešku standardne devijacije, ali imamo standardnu grešku aritmetičke sredine. Kako smo videlu u ovom poglavlju, standardna greška standardne devijacije je uvek niža od standardne greške aritmetičke sredine. Imajući to u vidu, ako je standardna greška aritmetičke sredine 0,084, što je mnogo niže od 1, isto mora biti slučaj i sa standardnom greškom standardne devijacije.

J4 – netačno. Iz tabele možemo da pročitamo da je varijansa tačno 1,691, što nije više od 2.

J5 – nepoznato. U tabeli nemamo podatke na osnovu kojih bi mogli da izvedemo bilo kakve zaključke o statisticima ili parametrima muškaraca i žena.

J6 – besmisleno. Šta bi tačno bila "percentilna devijacija"? Koliko nam je poznato, takav statistik ne postoji.

J7 – besmisleno. Šta bi tačno bio "kvartilni srednji raspon"? Koliko nam je poznato, takav statistik ne postoji.

PREPRINT

J8 – netačno. Možemo videte da je gornja granica 95% intervala poverenja viša od 4. Gornja granica 99% intervala poverenja bi morala da bude još viša (pogledajte formulu za računanje ovih intervala poverenja), što znači da ne može biti niža od 4.

J9 – tačno. Možemo videte da je vrednost skjunesa negativna i prilično velika, što znači da je distribucija negativno asimetrična.

J10 – tačno. 50i percentil je medijana, a iz tabele možemo pročitati da je medijana zaista 4.

K1 – netačno. Standardna devijacija distribucije uzorkovanja zove se standardna greška, a iz tabele možemo da pročitamo da je ona 0,09 za medijanu WFC. Vrednost 0,06 je bias, a to je nešto drugo.

K2 – tačno. Razlika o kojoj se govori u ovoj tvrdnji zove se bias i možemo videte da je ona ovde zaista 0.

K3 – netačno. Možemo videti da je vrednost kurtozisa na uzorku -0,97, a da interval poverenja kurtozisa koji je predstavljen u tabeli takođe ne uključuje vrednost 0, što znači da nije verovatno da je to vrednost u populaciji. U stvari, možemo videte da je vrednost 0 čak 15 standardnih grešaka udaljena od vrednosti uzorka. To znači da je vrednost 0 u populaciji veoma, veoma, veoma, veoma malo verovatna. Ovo je, naravno, tako ako pretpostavimo da radimo sa slučajnim uzorkom iz populacije o kojoj izvodimo zaključke ili uzorku koji je dovoljno dobra aproksimacija slučajnog uzorka, što je opšta premisa svih postupaka statistike zaključivanja koji su predstavljeni u ovoj knjizi.

K4 – besmisleno. Ne postoji statistik koji se zove “glavni sredinski kvartil”.

K5 – nepoznato. Nema podataka o veličini uzorka u ovoj tabeli. O tome bi eventualno mogli da probamo da izvedemo neke zaključke na osnovu odnosa veličina standardne greške i standardne devijacije, ako bi pretpostavili da bi vrednost standardne greške dobijena butstrepingom bila slična vrednosti standardne greške dobijene na osnovu centralne granične teoreme, ali ovo nikako nije sigurno i takva računica bi bila previse komplikovana da bi je očekivali od tipičnog čitaoca naučnog teksta. Mnogo je lakše pitati autore o tome ili pogledati neki drugi deo rada (u situacijama kada čitamo pun tekst naučnog rada naravno, ne samo pojedinačnu tabelu u vežbanju iz statistike!!).

K6 – tačno. Možemo videte da je skjunes pozitivan, kao i da je ceo raspon intervala poverenja u oblasti pozitivnog skjunesa, a ovo je interval za koji postoji 95% šanse da obuhvata parametar. Takođe, možemo videte i da je donja granica ovog intervala vrlo, vrlo daleko od 0. Pozitivan skjunes znači da je distribucija pozitivno asimetrična.

K7 – tačno. Da, tvrdnja se odnosi na drugi način da se interpretira interval poverenja aritmetičke sredine i vrednosti su tačne.

K8 – netačno. Interval poverenja je 1,72 i možemo videte da je vrednost od 1,7 unutar tog intervala. To znači da možemo u punoj meri očekivati da određeni broj budućih istraživanja ima aritmetičke sredine od 1,7 ili niže. Ali i da ovo nije slučaj, tvrdnja bi i dalje bila pogrešna, jer bilo koji rezultat, šta god on bio, može da učini samo malo verovatnim da se dobije aritmetička sredina određene vrednosti, ali nikako nemogućim, imajući u vidu da normalna distribucija asimptotski prilazi 0,

PREPRINT

nikad je ne dostižući (a da se vrednost o kojoj je reč nalazi u domenu vrednosti koje se mogu dobiti na skali koja je u pitanju).

K9 – tačno. Možemo videte da je vrednost kurtozisa visoko negativna, što znači da je distribucija platikurtična.

K10 – netačno. Ne, ovo je postupak za izvođenje zaključaka o vrednostima populacije zasnovan na butstrepingu, a ne na centralnoj graničnoj teoremi.

PREPRINT

Poglavlje 6. Korelacije

Apstrakt. Ovo poglavlje predstavlja koncepte povezanosti između varijabli i statističke postupke za opisivanje nivoa i vrste povezanosti između varijabli. Prvi deo poglavlja pokriva koncepte povezanosti i vrste povezanosti između varijabli, opisujući razlike između monotonih i nemonotonih veza, te povlačeći, unutar monotonih veza, razliku između linearnih i nelinearnih povezanosti. Nakon toga sledi kratak deo u kome se čitalac upoznaje sa sketergramima kao zgodnim grafičkim predstavljanjima veza između varijabli koje omogućavaju vizuelnu procenu vrste veze između varijabli. U kasnijem delu poglavlja, upoznajemo se sa konceptom koeficijenta korelacije, opisuje se kako se koeficijenti korelacije koriste i interpretiraju, a diskutuje se o rasponu vrednosti ovog koeficijenta i kako interpretirati različite vrednosti koeficijenta korelacije. Testiranje nulte hipoteze o veličini koeficijenta korelacije u populaciji tj. statističke značajnosti koeficijenta korelacije je predstavljeno u narednom delu. Poslednji deo ovog poglavlja predstavlja najpopularnije koeficijente korelacije kao što su Pirsonov produkt-moment koeficijent korelacije, Spirmanov koeficijent korelacije rangova, point-biserijski koeficijent, biserijski koeficijent korelacije, eta, fi, koeficijent kontingencije i Kramerov V.

Ključne reči: povezanost između varijabli, korelacije monotona/nemonotona, linearna/nelinearna.

6.1. Povezanost između varijabli

Jedna od prvih stvari koju treba uraditi onda kada se počinje istraživanje nove oblasti ili neke nove teme je identifikovanje ključnih varijabli, a potom sprovođenje posmatranja da bi se postavile hipoteze o njihovim međusobnim odnosima. Važan deo ovih aktivnosti je utvrđivanje koje varijable teže da se menjaju zajedno, a koje ne tj. koje varijable su povezane jedne s drugima.

U statistici, **statistik koji izražava snagu povezanosti između varijabli zove se korelacija**. **Da bi mogli da kažemo da su dve varijable povezane, neophodno je da možemo da uočimo da promene vrednosti jedne varijable prate, manje ili više precizne, promene vrednosti druge varijable. Ovo zajedničko menjanje vrednosti posmatranih varijabli je stvar stepena** – u nekim slučajevima će promene jedne varijable potpuno precizno pratiti promene vrednosti druge varijable, ali će u većini slučajeva ove zajedničke promene biti manje precizne, sve do tačke kada postoji samo vrlo slaba veza između dve varijable.

PREPRINT

Treba da imamo u vidu i da **korelacija tj. povezanost dve varijable ne znači da su one u uzročno-posledičnom odnosu!** Ako su dve varijable u korelaciji to ne znači da je jedna uzrok drugoj (tj. da promene jedne izazivaju promene ove druge). Ako jedna varijabla uzrokuje drugu tj. ako postoji uzročno-posledična veza između dve varijable, onda je verovatno da će one biti u korelaciji odnosno povezane, ali **varijable mogu biti korelirane/povezane iz čitavog niza različitih razloga od kojih je uzročno-posledična veza samo jedan od mogućih**. Takođe je moguće i da promene u obe varijable izaziva neka treća varijabla (koju ne posmatramo u datoj situaciji) i da to onda bude razlog za dobijenu korelaciju/povezanost.

Priroda veze između dve varijable može takođe da bude i mnogo kompleksnija, na primer, da obe budu deo istog uzročno-posledičnog lanca ili mreže, pa da onda veći broj drugih varijabli utiče i na jednu i na drugu i da ti raznorodni uticaji dovode do korelacije. U situacijama poput ove, **gde su varijable povezane, ali njihova veza nije uzročno-posledična** (u smislu da nije jedna uzrok drugoj), **namerne promene vrednosti jedne varijable neće dovesti do premene vrednosti druge varijable**, **a ovako napravljene promene će zapravo smanjiti nivo povezanosti između te dve varijable koje posmatramo ili će čak potpuno poništiti njihovu povezanost**. Namerno menjanje vrednosti jedne varijable neće dovesti do promena kod druge varijable ni u situaciji kada su dve varijable uzročno-posledično povezane, ali je ona koju menjamo posledica one druge, a ne njen uzrok. **Jedina situacija u kojoj će namerno menjanje vrednosti jedne varijable dovesti do promene vrednosti druge je situacija gde je varijabla koju menjamo uzrok, a ona druga je posledica** (ili je negde dalje u uzročnom lancu posledica ove prve varijable), a takva situacija je samo jedna od onih koje srećemo u naučnim istraživanjima. Zato je **veoma važno da čitaoci uvek imaju u vidu da korelacija nije kauzacija, tj. da povezanost dve varijable ne znači uzročno-posledičnu vezu**, iako **neke korelacije mogu biti posledica uzročno-posledičnih veza između varijabli.** Ovo je važno naglasiti zbog toga što je istorija nauke, a pogotovo istorija primene naučnih znanja puna situacija kada su obični odnosi povezanosti tj. korelacije pogrešno interpretirani kao uzročno-posledični i kada su, zbog takvih pogrešnih interpretacija, ljudi pogešno verovali da će promena jedne varijable iz poznatog korelacionog odnosa dovesti do promena druge. Neki od istorijskih primera su već pomenuti u poglavlju o samoispunjujućim proročanstvima i kvarenju statističkih indikatora, a još jedna veoma česta greška koju valja pomenuti je kroz istoriju veoma učestala tendencija da se razni prehrambeni artikli ili navike u ponašanju koje su u datom trenutku popularne među bogatijim slojevima društva opisuju kao

PREPRINT

zdrave ili korisne za unapređenje zdravlja, samo na osnovu opaženih korelacija između konzumiranja tih prehrambenih artikala ili upražnjavanja takvih navika i različitih pokazatelja zdravstvenog statusa. Iako je vrlo izgledno da postoje navike u ponašanju i vrste hrane koje pomažu unapređenju zdravlja, vrlo često su ove povezanosti prosto posledica činjenice da bogatiji delovi stanovništva imaju istovremeno i bolje životne uslove i pristup boljoj zdravstvenoj zaštiti (što sve vodi boljem zdravlju!), ali i pomenutoj egzotičnoj hrani/navikama u ponašanju zbog svog boljeg finansijskog stanja, tako stvarajući ovakvu korelaciju. Međutim, ljudi koji nemaju pristup takvim finansijskim sredstvima, a onda zbog toga ni takvom kvalitetu zdravstvene zaštite i života uopšte, a koji samo usvoje navike bogatih, vrlo brzo otkriju da očekivani zdravstveni rezultati izostaju. Tu je takođe i veoma jednostavan primer koji možemo naći u mnogim uvodnim udžbenicima iz statistike – ako bi računali korelaciju između broja učitelja i broja lopova u nizu gradova, videli bi da između broja učitelja i broja lopova u gradu postoji vrlo jasna povezanost. Na osnovu ovoga, osoba koja ne razlikuje korelaciju od uzročno-posledične veze bi mogla da zaključi da učitelji nekako privlače ili proizvode lopove (ili obrnuto?). Međutim, razlog za ovu korelaciju je prosto to što postoji treća varijabla koja je odgovorna za tu vezu, a to je veličina grada. Veći gradovi će imati više ljudi, a to znači i više učitelja i više lopova, te zapravo i više ljudi bilo koje drugo profesije.

Još jedno svojstvo korelacionog odnosa je to da, **bez dodatnog teorijskog znanja, postojanje korelacije između varijabli ne znači da će takva povezanost nastaviti da postoji i u budućnosti**. Kao što smo već pomenuli u prvom poglavlju, u delu o statističkim objašnjenjima i kvarenju statističkih indikatora, **da bi mogli da pretpostavimo da li će određena povezanost nastaviti da postoji i u budućnosti, potrebno je da znamo zašto data povezanost uopšte postoji sada** tj. moramo imati validno i detaljno naučno objašnjenje povezanosti između varijabli. Idealno bi bilo da to bude kauzalno objašnjenje, ali često i druge vrste objašnjenja mogu da budu dovoljno dobre. Ako nemamo valjano objašnjenje veze između varijabli, korelacija koja je nekada postojala može prosto da prestane da postoji u bilo kom trenutku u budućnosti, a isto tako se korelacija može pojaviti u budućnost između varijabli koje u prošlosti nisu bile u korelaciji. A onda ta novonastala korelacija može ponovo nestati kasnije ili promeniti svoj intenzitet. Ovaj fenomen je posebno dobro poznat stručnjacima koji rade u oblasti predviđanja cena različite robe, a pogotovo onima koji se bave predviđanjem cena finansijskih instrumenata na organizovanim finansijskim tržištima kao što su na primer berze. Kako berzanski trgovci znaju da kažu – „prošli

PREPRINT

dobici su slabi pokazatelji budućih dobitaka“, a isto važi i za korealcije. Ovo je posebno slučaj kada su varijable u pitanju zapravo elementi ponašanja ljudi.

6.2. Vrste povezanosti između varijabli

Kako sve mogu da izgledaju veze između varijabli? Iako često mislimo da su povezanosti između varijabli prosta stvar – npr. da kad jedna raste, raste i druga, u praksi **povezanosti mogu da imaju puno različitih oblika, od vrlo prostih odnosa do veoma složenih**. Na primer, ako posmatramo dva plesača koji zajedno izvode kompleksnu koreografiju, možemo zaključiti da su njihovi pokreti (tokom izvođenja plesa) povezani iako su sasvim različiti. Ako znamo kako koreografija izgleda i znamo koje pokrete jedan od plesača izvodi u datom trenutku, mogli bismo vrlo precizno da predvidimo pokrete koje drugi plesač izvodi u datom trenutku, čak iako ne bi mogli da vidimo šta taj drugi plesač radi. Na drugoj krajnosti kompleksnosti veza između varijabli imamo vrlo proste odnose poput, recimo, nivoa pritiska na pedalu gasa u automobilu i brzine automobila (na ravnoj podlozi, u istoj brzini...) ili između različitih sličnih osobina ličnosti, kao na primer u slučaju srednje povezanosti između otvorenosti za iskustvo i umetničkih profesionalnih interesovanja (e.g. V. Hedrih, 2009). Između ovih krajnosti su umereno kompleksni odnosi poput, na primer, odnosa između toga kako osoba procenjuje svoju stručnost za neku oblast i toga kakva je objektivno njena stručnost za datu oblast. Ovaj odnos opisuje poznati Daning-Krugerov efekat, imenovan po autorima istraživačke studije čiji su rezultati pokazali da su ljudi u najnižem kvartilu po skorovima na testu koji su istraživači zadavali precenjivali svoje prave skorove mnogo češće nego ljudi u višim kvartilima, pri čemu su učesnici u istraživanju sa najvišim skorovima potcenjivali svoj uspeh (Kruger & Dunning, 1999). Isto tako zgodan primer je i logaritamski odnos između fizičkog intenziteta vizuelnih ili zvučnih stimulusa i njihovog opaženog intenziteta, poznat i kao Veber-Fehnerov zakon (e.g. Portugal & Svaiter, 2011).

Kada razmatramo povezanost između varijabli još jedan važan pojam koji treba razumeti je pojam smera povezanosti između varijabli. **Smer povezanosti između varijabli odnosi se na smer u kom se vrednosti jedne varijable menjaju kada se menjaju vrednosti druge varijable. Kada porast vrednosti jedne varijable prati porast vrednosti druge varijable, govorimo o pozitivnoj povezanosti**. Povezanost je takođe pozitivna i ako smanjenje vrednosti jedne varijable prati smanjenje vrednosti druge varijable. Drugim rečima, kad god se dve varijable menjaju

PREPRINT

zajedno u istom smeru, u pitanju je pozitivna povezanost. **Negativna povezanost postoji onda kada porast vrednosti jedne varijable prati smanjenje vrednosti druge varijable** tj. kada su odgovarajuće promene vrednosti dve varijable u suprotnim smerovima. Valja naglasiti i to da **da bi povezanost između varijabli uopšte imala smer, te varijable moraju da budu bar na ordinalnom nivou merenja tj. na nivou merenja koji ima smerove** (tj. niže i više vrednosti). **Varijable na nominalnom nivou merenja nemaju smer povezanosti** zato što, na ovom nivou merenja, **nema manjih i većih vrednosti** te stoga vrednosti ne mogu da rastu niti da opadaju.

Prema tome, povezanost između varijabli može da ima puno različitih oblika, a za potrebe računanja statističkih pokazatelja njihovih intenziteta, **možemo ih podeliti prema tome da li se smer povezanosti menja ili je stalan**. S obzirom na to svojstvo, povezanosti između varijabli mogu biti:

- **Monotone povezanosti** – kada je **promena vrednosti jedne varijable povezana sa promenama druge varijable, ali uvek u istom smeru**. Kada vrednost jedne varijable raste, vrednost druge varijable ili raste ili opada, ali taj smer ostaje isti za sve vrednosti prve varijable. Ako vrednosti druge varijable rastu sa porastom vrednosti prve varijable, to ostaje tako za sve vrednosti prve varijable. Ako se vrednosti druge varijable smanjuju sa porastom vrednosti prve varijable, to ostaje tako za sve vrednosti prve varijable. Na grafiku, najbolja predstava povezanosti ovog tipa je linija koja nikada ne menja smer (ni u horizontalnom ni u vertikalnom meru).
- **Nemonotone povezanosti** – kada su **promene vrednosti jedne varijable praćene promenama vrednosti druge varijable, ali ove promene imaju različite smerove za različite vrednosti prve varijable govorimo o nemonotonoj povezanosti.** Ovo znači da kada vrednosti prve varijable rastu, taj porast prati porast vrednosti druge varijable na jednom delu raspona prve varijable, a pad vrednosti druge varijable na drugom delu ili delovima raspona vrednosti prve varijable. Drugim rečima - **smer povezanosti se menja**. Na grafiku, najbolja predstava nemonotonog odnosa je linija koja menja smer bilo u horizontalnom, bilo u vertikalnom pravcu ili u oba pravca.

Još jedna stvar koju treba razmotriti u priči o vrstama povezanosti, posebno monotonih povezanosti je **brzina promene prve varijable kada se druga varijabla menja**. Ova brzina može biti konstantna ili se može menjati. **Konstantna brzina promene znači da kad god se vrednosti**

PREPRINT

jedne varijable promene za određenu vrednosti, dođe do promene vrednosti druge varijable odgovarajuće veličine i taj partitet promena između toga koliko se jedna varijabla promeni naspram toga koliko se druga varijabla promeni je konstantan u celom rasponu varijable. To **ne znači da su brzine promene dve varijable jednake**, već znači da se **brzina promene jedne varijable može predstaviti kao brzina promene druge pomnožena nekim fiksnim koeficijentom** (npr. jedna se menja 2 puta ili 3 puta ili 0,5 puta brže/sporije od druge). Ako bi ovakav odnos između dve varijable predstavili na grafiku, **takav odnos bi najbolje opisala prava linija.** S druge strane, kada brzina **promene nije konstantna to znači da je veličina promene jedne varijable koja odgovara fiksnom iznosu promene druge varijable različita za različite pozicije na rasponu vrednosti druge varijable**. U odnosu na tempo promene, monotone povezanosti između varijabli mogu biti:

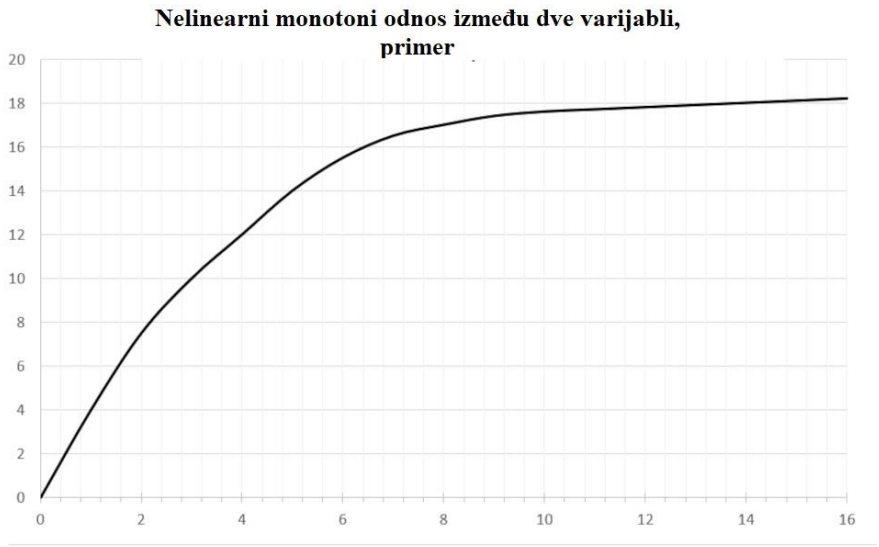
- **Linearne – kada je brzina promene jedne varijable u odnosu na promenu druge varijable konstantna**. Na primer, ako promena vrednosti jedne varijable za 2, dovodi do promene druge varijable za 6, ovaj paritet važi za ceo raspon vrednosti dve varijable. To znači da od koje god vrednosti da krenemo, povećavanje prve varijable za 2 će pratiti povećanje druge varijable za 6 (odnosno za oko 6, ako intenzitet povezanosti nije savršen). Na grafiku, najbolja reprezentacija linearnog odnosa između dve varijable je **prava linija.**
- **Nelinearne** – kada je **brzina promene jedne varijable u odnosu na promene druge varijable promenljiva.** Na primer, sa porastom vrednosti jedne varijable, druga varijable raste prvo vrlo sporo, a onda sve brže i brže kako vrednosti prve varijable rastu. Ili, recimo, kada sa konstantom stopom promene prve varijable, druga se menja prvo vrlo brzo, a onda posle sve sporije i sporije.

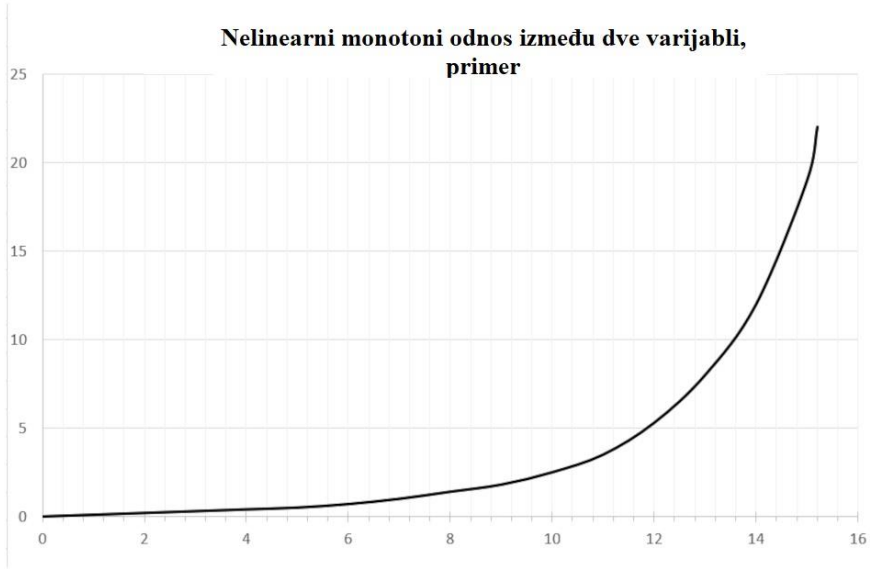
Da bi monotona povezanost između dve varijable mogla da bude linearna ili nelinearna, varijable o čijoj povezanosti je reč moraju da budu bar na intervalnom nivou merenja. Ovo je neophodno zato što da bi bilo moguće odrediti da li je veličina promene varijabli konstantna ili se menja, mora da bude moguće proceniti/izračunati veličinu promene, a da bi to bilo moguće potrebna je jedinica mere fiksne veličine, a to je nešto što postoji tek na intervalnom nivou merenja. **Na ordinalnom nivou merenja se ne može smisleno razmatrati linearnost/nelinearnost odnosa između dve varijable**, jer na ovom nivou merenja nije moguće upoređivati veličine promene vrednosti zato što nema fiksne jedinice mere.

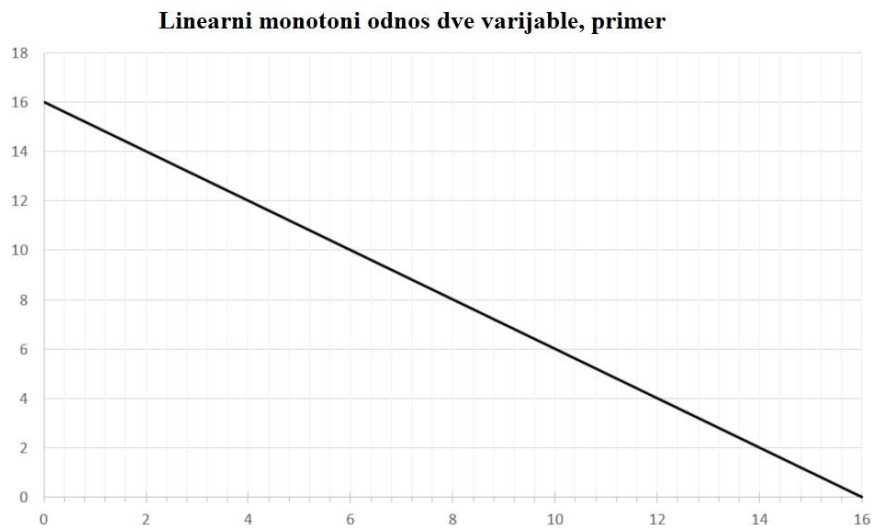
PREPRINT

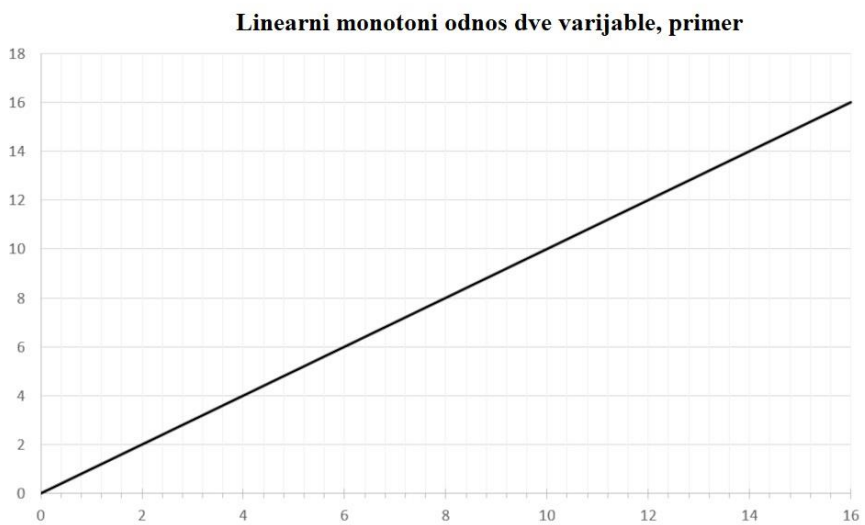
Slika 6.1. Grafički prikaz nekih od mogućih oblika moonotonih odnosa. Horizontalna osa predstavlja vrednosti jedne varijable, dok vertikalna osa predstavlja vrednosti druge varijable. Linija predstavlja vrednosti dve varijable koje jedna drugoj odgovaraju. Na primer, na gornjoj levoj slici možemo videti da vrednost 6 na varijabli koja je predstavljena horizontalnom osom odgovara vrednosti koja je nešto manja od 16 na drugoj varijabli. Dva gornja grafika predstavljaju nelinearne odnose. Možemo videti da je linija kriva što ukazuje da se brzina promene jedne varijable sa promenama druge menja. Dva donja grafika predstavljaju linearne odnose. Donji levi grafik predstavlja negativnu povezanost između varijabli, gde povećanje vrednosti jedne varijable prati smanjenje vrednosti druge varijable. S druge strane, donji desni grafik predstavlja pozitivnu povezanost između dve varijable, jer povećanje vrednosti jedne varijable prati povećanje vrednosti druge varijable. Dve nelinearne povezanosti koje su predstavljene na gornja dva grafika takođe predstavljaju pozitivne povezanosti između varijabli. Možemo videti da povećanje vrednosti jedne varijable takođe prati povećanje vrednosti druge, samo što se brzina promene menja, jer je odnos nelinearan.

PREPRINT


Nelinearni monotoni odnos između dve varijabli, primer
20
18
16
14
12
10
8
6
4
2
0
0 2 4 6 8 10 12 14 16

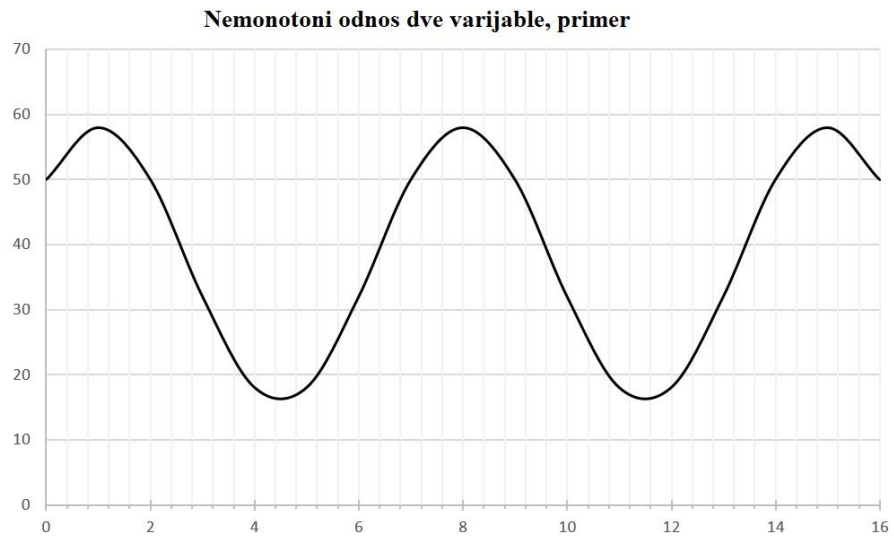

Nelinearni monotoni odnos između dve varijabli, primer
25
20
15
10
5
0
0 2 4 6 8 10 12 14 16

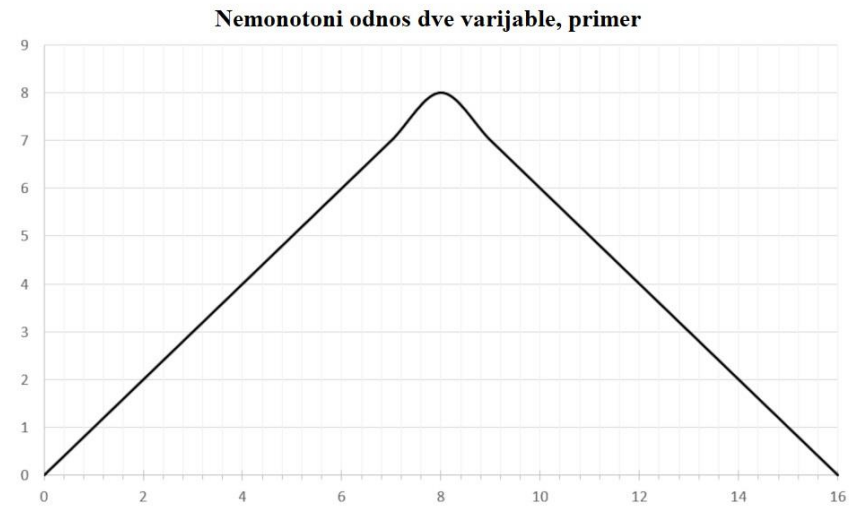

Linearni monotoni odnos dve varijable, primer
18
16
14
12
10
8
6
4
2
0
0 2 4 6 8 10 12 14 16

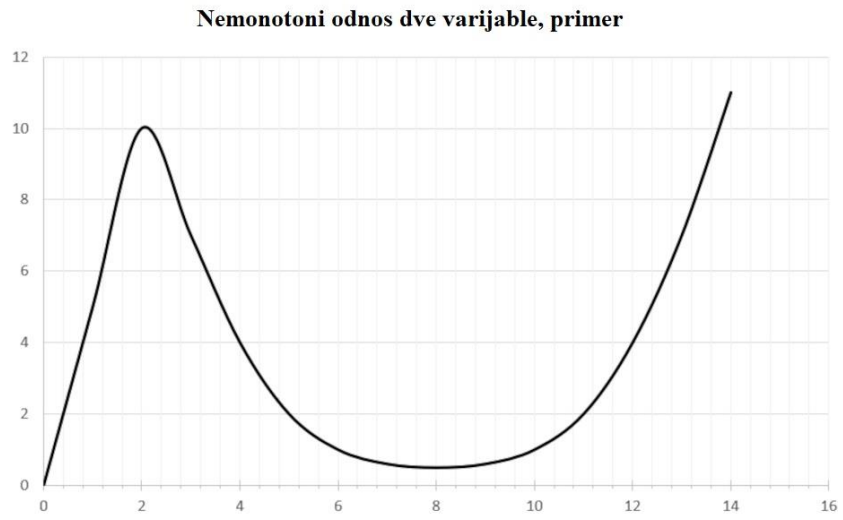

Linearni monotoni odnos dve varijable, primer
18
16
14
12
10
8
6
4
2
0
0 2 4 6 8 10 12 14 16

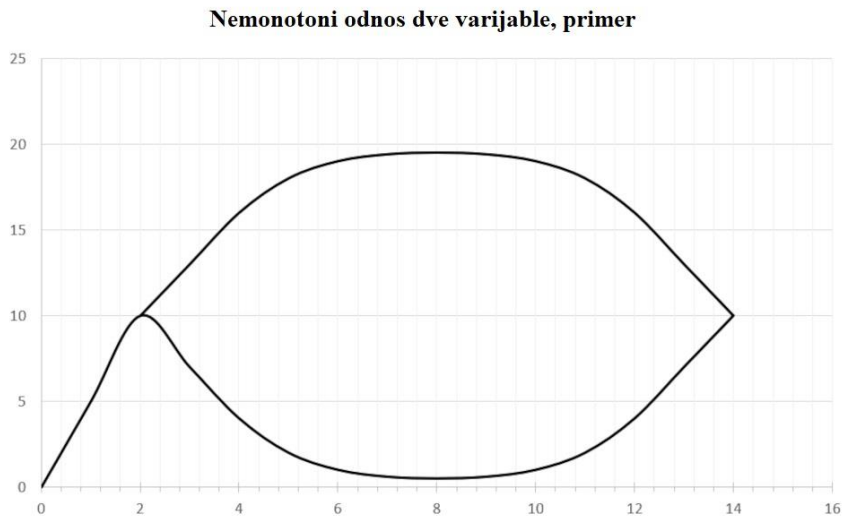
Slika 6.2. Grafički prikaz nekih mogućih oblika nemonotonih odnosa. Iz ovih tipova odnosa možemo videti da bar jedna varijabla ima dve ili više vrednosti druge varijable koje joj odgovaraju. Zbog ovoga, preciznost sa kojom se vrednosti jedne varijable mogu predvideti na osnovu druge može da bude različita zavisno od toga koja varijabla se koristi za predviđanje koje. Na primer, ako pogledamo gornji levi grafik, možemo videti da ako probamo da predvidimo vrednosti varijable predstavljene vertikalnom osom, na osnovu vrednosti varijable koja je predstavljena horizontalnom osom, to možemo da uradimo savršeno jer za svaku vrednost na horizontalnoj osi postoji tačno i samo jedna odgovarajuća vrednost na vertikalnoj osi. S druge strane, većina vrednosti varijable na vertikalnoj osi ima više vrednosti na horizontalnoj osi koje joj odgovaraju. Na primer, ako pogledamo vrednost 40 na vertikalnoj osi, možemo videti da su vrednosti na horizontalnoj osi koje joj odgovaraju 2,3, 6,2 i 12,7 (ovo su samo gruba očitavanja vrednosti na osnovu vizuelnog pregleda grafika, tako da možda nisu sasvim precizna). Situacija je slična na gornjem desnom grafiku (onom koji ima linearne delove), kao i na donjem levom. S druge trane, kod odnosa koji je prikazan na donjem desnom grafiku obe varijable imaju više od jedne odgovarajuće vrednosti druge varijable, ali ne duž celog raspona. Možemo videti na grafiku da vrednosti ispod 2 na varijabli koja je predstavljena horizontalnom osom imaju tačno jednu odgovarajuću vrednost na vertikalnoj osi.

PREPRINT


Nemonotoni odnos dve varijable, primer
70
60
50
40
30
20
10
0
0 2 4 6 8 10 12 14 16


Nemonotoni odnos dve varijable, primer
9
8
7
6
5
4
3
2
1
0
0 2 4 6 8 10 12 14 16


Nemonotoni odnos dve varijable, primer
12
10
8
6
4
2
0
0 2 4 6 8 10 12 14 16


Nemonotoni odnos dve varijable, primer
25
20
15
10
5
0
0 2 4 6 8 10 12 14 16

Još jedna stvar koju treba razmotriti kada je u pitanju povezanost varijabli je to da **intenzitet povezanosti između varijabli ne mora da bude stalan u celom rasponu vrednosti posmatranih varijabli**. Iako itekako može biti slučajeva kada je intenzitet veze između varijabli (to koliko precizno promene jedne prate promene druge varijable) stalan u celom rasponu vrednosti varijabli, postoje i slučajevi gde se intenzitet veze između varijabli razlikuje u različitih delovima raspona vrednosti dve varijable, a mogu postojati i situacije gde postoje intervali vrednosti varijabli u kojima su varijable povezane i intervali vrednosti u kojima nisu povezane.

Dok se različite vrste povezanosti između varijabli uzimaju u obzir tako što se računaju pokazatelji povezanosti koji su odgovarajući za vrstu povezanosti koja postoji između konkretnih varijabli čiju povezanost računamo, **promenljivost intenziteta veze između varijabli se uzima u obzir prvenstveno tako što se računaju pokazatelji povezanosti za različite delove raspona vrednosti varijabli posebno** (za svaki deo gde je vidno različit intenzitet povezanosti računa se poseban pokazatelj povezanosti).

6.3. Sketergram

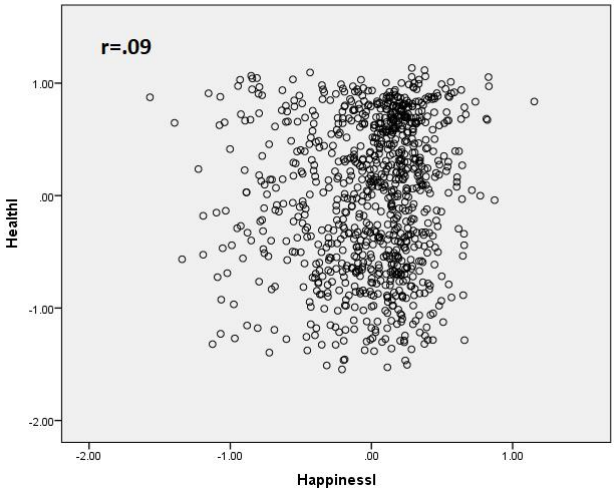
Jedan veoma koristan alat za ispitivanje povezanosti dve varijable je sketergram. **Sketergram je grafički prikaz odnosa dve varijable**. To je, najčešće, dvodimenzionalni grafik na kom je jedna varijabla predstavljena kao jedna dimenzija – koordinatna osa na grafiku, a druga kao druga dimenzija – koordinatna osa. Entiteti su predstavljeni kao tačke čije su koordinate na grafiku njihove vrednosti na te dve varijable. Kada se stvari tako predstave, dobija se skup tačaka na grafiku, a iz načina na koji su tačke grupisane moguće je izvesti zaključke o verovatnoj vrsti odnosa između dve varijable, smeru njihove povezanosti, a i otprilike proceniti intenzitet njihove povezanosti.

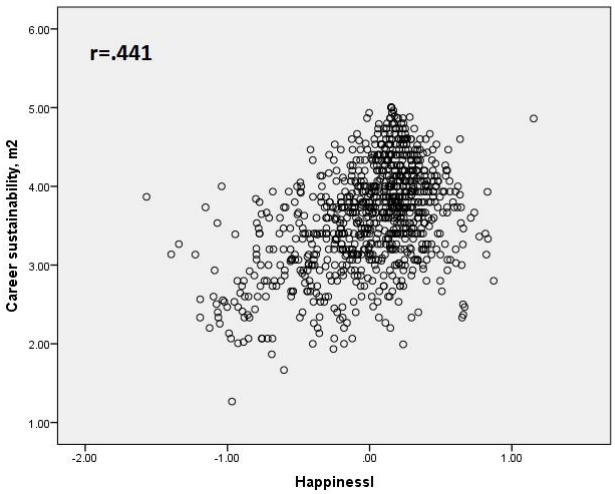
Slika 6.3. Primeri sketergrama koji prikazuju korelacije između dve varijable različitog intenziteta i smera. Možemo videti da kako korelacija između varijabli postaje jača, tako i opšti oblik koji grade tačke na sketergramu postaje tanji i, u suprotnom smeru, kako korelacije postaju slabije, tako i oblik koji gradi raspored tačaka postaje okruglastiji. Gornja 4 sketergrama prikazuju pozitivne korelacije, dok donja dva prikazuju negativne. Možemo videti da, kada su korelacije negativne,

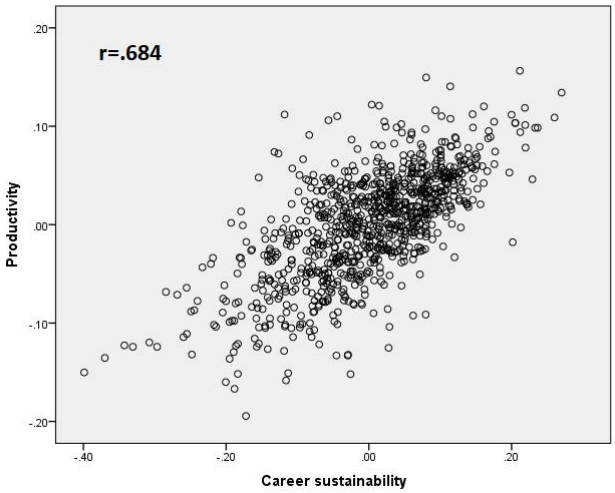
PREPRINT

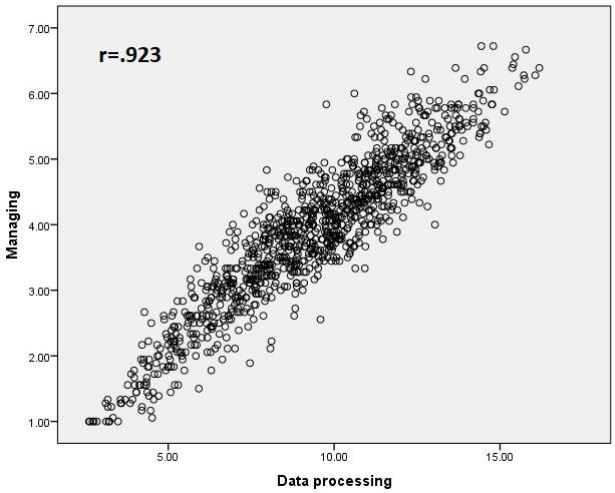
visoke vrednosti jedne varijable odgovaraju niskim vrednostima druge varijable. Podaci potiču iz različitih istraživanja koja su sproveli autori.

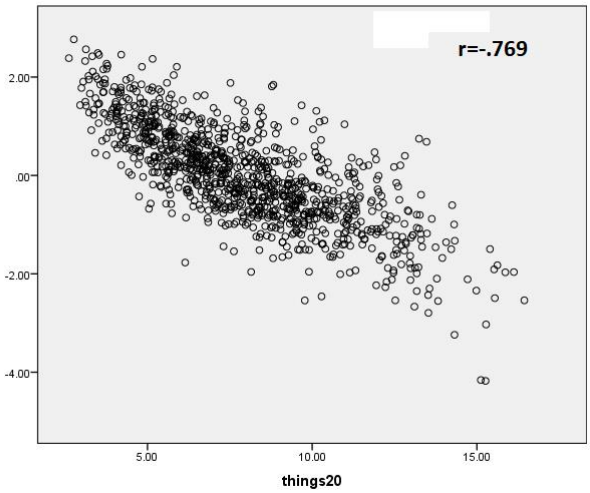
PREPRINT

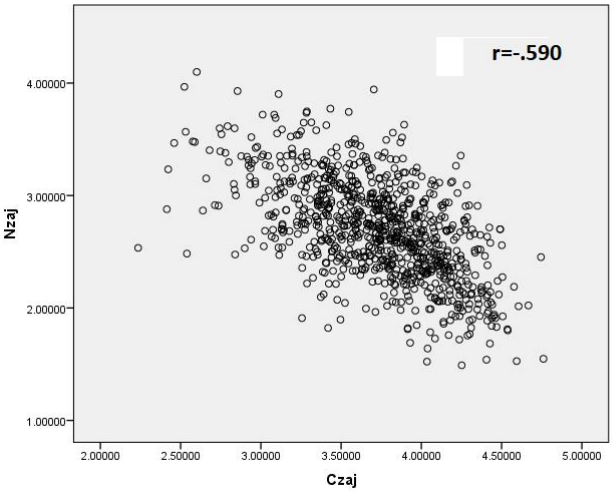

r=.09
Healthl
Happinessl


r=.441
Career sustainability, m2
Happinessl

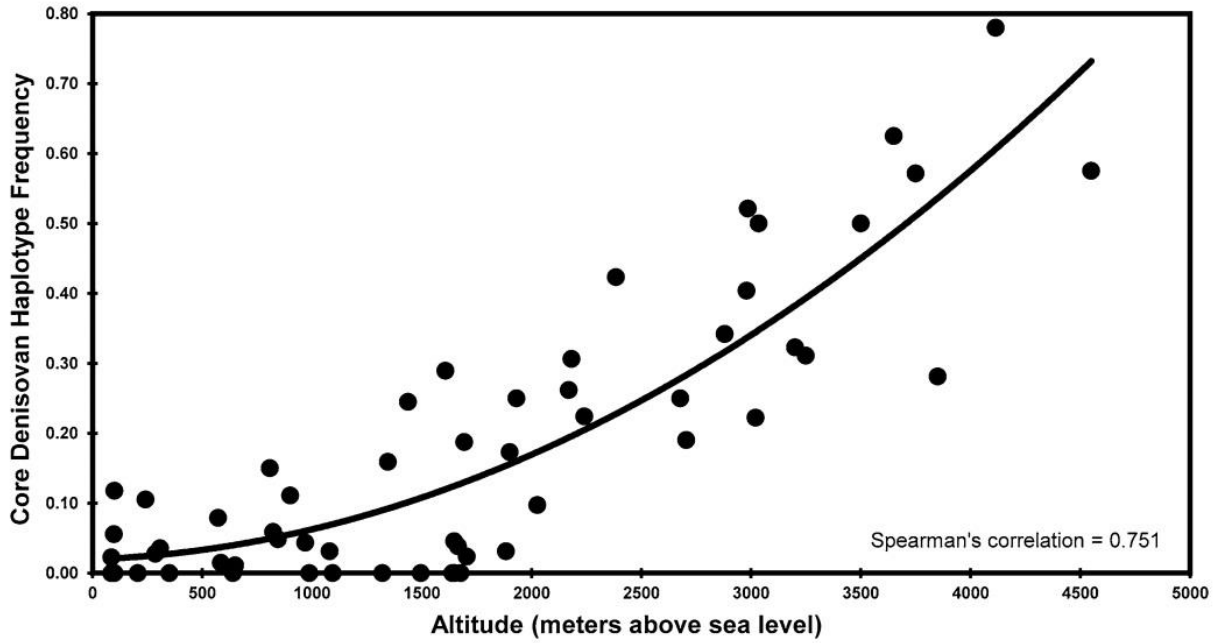

r=.684
Productivity
Career sustainability


r=.923
Managing
Data processing


r=-.769
things20


r=-.590
Nzaj
Czaj

Slika 6.4. Primer sketergrama koji predstavlja blago nelinearnu, monotonu, pozitivnu korelaciju. Sliku prenosimo na iz Hackinger et al. (2016) na osnovu dozvole autora i objavljuje se u skladu sa Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/). Pored uklanjanja izvornog broja slike, koji je uklonjen da bi sprečili zabunu oko označavanja slika u ovoj knjizi i grafičkog formatiranja za potrebe uključivanja u ovu knjigu, nikakve druge promene nisu unošene u sadržaj ove slike.



U principu, pravila su sledeća – **ako je korelacija nula ili blizu nule, tj. kada nema povezanost ili je ona vrlo, vrlo slaba, distribucija entiteta na sketergramu će težiti da ima otprilike kružan oblik. Što se više razlikuje od kružnog oblika, što postaje uža, više nalik liniji, to je povezanost jača. Ako grupisanje podseća na pravu liniju, po svoj prilici je u pitanju linearni odnos. Ako izgleda kao kriva linija, ali ona koja ne menja smer, onda je u pitanju nelinearna monotona povezanost. Konačno, ako distribucija tačaka podseća na krivu liniju, ali liniju koja menja smer (ide na gore pa na dole ili na levo, pa na desno) radi se o nemonotonoj povezanosti između varijabli.**

U delu srpske naučne/stručne literature ova vrsta grafičkog prikaza se naziva i **dijagram raspršenja.**

6.4. Koeficijent korelacije

Statistici koji se koriste kao pokazatelji snage (ili intenziteta) povezanosti između dve varijable zovu se koeficijenti korelacije. Koeficijent korelacije **tipično ima vrednosti između 0 i 1** i **može da bude pozitivan ili negativan tj. + ili -. Vrednost koeficijenta korelacije i njegov predznak se tumače nezavisno jedan od drugog**. Koeficijent korelacije **tipično pokazuje dve stvari:**

- **Intenzitet korelacije tj. koliko je jaka povezanost između dve varijable i to pokazuiije (apsolutna) vrednost koeficijenta**. Ova vrednost **ide od 0 do 1. Korelacija koja je 0 pokazuje da nema povezanosti između varijabli, dok korelacija 1 pokazuje da postoji potpuna ili maksimalna povezanost između varijabli.** Drugim rečima, kada je korelacija 0, to znači da predviđanje vrednosti jedne varijable na osnovu druge nije ništa uspešnije od slučajnog pogađanja, tj. ništa uspešnije nego kada bismo varijabli koju predviđamo prosto dodelili vrednosti slučajnim putem. S druge strane, kada je korelacija 1 to znači da vrednosti jedne varijable možemo predvideti na osnovu vrednosti druge varijable savršeno, bez ikakvih grešaka. **Korelacija između 0 i 1 znači da ćemo napraviti određene greške prilikom predviđanja vrednosti jedne varijable na osnovu vrednosti druge, ali će ove greške biti sve manje i manje što je koeficijent korelacije veći.**
- **Smer korelacije pokazuje predznak koeficijenta korelacije** tj. to da li je koeficijent korelacije pozitivan ili negativan:
    - **Pozitivni koeficijent korelacije** pokazuje da je **korelacija između dve varijable pozitivna.** Pozitivna korelacija znači da se **vrednosti varijabli menjaju zajedno u istom smeru.** Porast vrednosti jedne varijable prati porast vrednosti druge varijable i obrnuto – smanjenje vrednosti jedne varijable prati smanjenje vrednosti druge varijable. Ili drugim rečima – entiteti koji imaju više vrednosti na jednoj varijabli, težiće da imaju više vrednost i na drugoj varijabli, a entiteti koji imaju niže vrednosti na jednoj varijabli, težiće da imaju niže vrednosti i na drugoj varijabli.
    - **Negativni koeficijent korelacije** pokazuje da je **korelacija između varijabli negativna. Negativna korelacija znači da se varijable zajedno menjaju u suprotnim smerovima.** Povećanje vrednosti jedne varijable prati smanjenje

PREPRINT

vrednosti druge varijable. Ili drugim rečima – entiteti koji imaju više vrednosti na jednoj varijabli teže da imaju niže na drugoj i obrnuto.

Konačno**, nulta (0) korelacija pokazuje da nema korelacije između varijabli**. Promene vrednosti jedne varijable ne prate bilo kakve određene ili predvidljive promene vrednosti druge varijable tj. **varijable se menjaju nezavisno jedna od druge**. Isto tako, poznavanje vrednosti entiteta na jednoj varijabli, ne omogućava nam da predvidimo kakvu će vrednost dati entitet imati na drugoj varijabli. Međutim, valja primetiti da **postoje različiti koeficijenti korelacije** (predstavljeni u narednom delu ove knjige) koji su **namenjeni opisivanju intenziteta različitih vrsta povezanosti između varijabli**. Iz ovog razloga, **nulta korelacija može takođe da znači i to da vrsta povezanosti između varijabli nije ona one vrste koju može da detektuje/valjano opiše koeficijent korelacije koji je izračunat**. Veoma česta situacija ovog tipa je situacija gde je veza između varijabli nemonotona, a mi računamo koeficijent korelacije koji je namenjen opisivanju linearne povezanost između varijabli (linearne korelacije). U takvom sluačju koeficijent linearne korelacije će vrlo verovatno pokazivati nultu korelaciju, kada u realnosti postoji nemonotona korelacija određenog intenziteta.

Takođe treba imati u vidu i to da **nemaju svi koreficijent korelacije smer, pa prema tome ni predznak**:

- **Koeficijenti korelacije koji se koriste za izražavanje intenziteta nemonotone povezanosti između varijabli po pravilu nemaju predznak** tj. smer, zato što se, u nemonotonom odnosu, smer povezanosti menja, nije stalan. To znači da je na jednom delu raspona vrednosti varijable pozitivan, na drugom delu je negativan, a moguće je da ima i delova gde je nula, te zbog toga jedan jedinstveni pokazatelj smera povezanosti ne bi bio adekvatan.
- **Koeficijent korelacije koji se koristi da izrazi intenzitet povezanosti nominalnih varijabli takođe nema smer odnosno predznak.** Kao što smo pomenuli ranije, smer povezanosti zahteva podatke za koje se može ustanoviti koje vrednosti su više, a koje niže (te da tako ima smisla govoriti o povećanju ili smanjenju vrednosti). Ovo nije slučaj sa nominalnim varijablama kod kojih za bilo koje dve vrednosti možemo samo da ustanovimo da li su jednake ili ne, ali ne možemo da smisleno poredimo njihove veličine. Zbog toga, koeficijenti korelacije ne mogu da budu pozitivni ili negativni za ove varijable.

PREPRINT

Jedna stvar koju valja primetiti kada su u pitanju pokazatelji smera korelacije je i to da se, prema postojećim pravilima za pisanje brojeva, pozitivan predznak ne piše ispred broja, te da kada, recimo hoćemo da napišemo +3, mi napišemo samo 3. Predznak korelacije se piše samo za negativne brojeve. Iz ovog razloga, **kada tumačimo rezultate statističkih postupaka koji su predstavljeni u literaturi, često je potrebno da obratimo dodatnu pažnju na detalje predstavljenih rezultata kako bi razlikovali pozitivne koeficijente korelacije od koeficijenata korelacije koji nemaju predznak** tj. koji nemaju smer. S druge strane, kada vidimo negativan koeficijent korelacije, znamo da se defitivno radi o koeficijentu korelacije koji ima smer.

Za **interpretaciju intenziteta koeficijenta korelacije** nema nekih fiksnih ili prirodnih pravila koja bi odredila šta tačno predstavlja malu korelaciju, a koliko tačno su visoki visoki koeficijenti korelacije. Međutim, postoje različite preporuke o tome kako ih interpretirati. Jedan od verovatno najšire upotrebljavanih skupova preporuka u naučnoj i stručnoj literaturi su preporuke koje je predložio Koen (Cohen, 1988). **Koenova, široko prihvaćena, preporuka je da se koeficijenti korelacije od oko 0,1 smatraju za slabe/niske, korelacije oko 0,3 za umerene, a korelacije od oko 0,5 za visoke. Ako konvertujemo ove preporuke u invervale, mogli bi da kažemo da su korelacije do 0,2 niske/slabe, one između 0,2 i 0,4 su umerene/srednje, a one preko 0,4 su visoke.** Međutim, to šta je veliko, a šta se može smatrati malim zavisi u velikoj meri i od pozicije iz koje posmatramo stvari. U svom pregledu iz 1990. godine, **Tejlor** (Taylor, 1990) pominje istraživačke studije u oblasti medicine u kojima su dobijeni koeficijenti korelacije iznad 0,9 i navodi preporuke za interpretaciju veličine koeficijenta korelacije prema kojima se **korelacije ispod 0,35 smatraju slabim, one između 0,36 i 0,67 se smatraju umerenim, visokim se smatraju korelacije iznad 0,68, dok su korelacije preko 0,9 veoma visoke**. S druge strane, **pregled nivoa korelacija koje su dobijene u psihologiji** objavio je Hemfil u svom tekstu iz 2003. godine (Hemphill, 2003) i ono što sledi iz njih je donekle različito od obe ove prepreke. Hemfil je razmatrao veliki broj metaanalitičkih studija sprovedenih **o istraživanjima u psihologiji** koje je podelio u tri kategorije jednake veličine prema veličinama uzoraka i našao je da je **otprilike jedna trećina istraživanja dobila korelacije oko 0,2, još jedna trećina je dobila korelacije između 0,2 i 0,3, a u trećini studija u kojima su dobijene najveće korelacije te korelacije su išle od otprilike 0,3 – 0,35, pa do 0,6 ili 0,78** zavisno već od grupe studija koja je posmatrana. Prema tome, **praktične smernice za interpretaciju veličine koeficijenta korelacije se mogu razlikovati zavisno od oblasti u kojoj radimo**. Imajući ovo u vidu, u praksi bi bilo **najmudrije**

PREPRINT

koristiti pravila za interpretaciju veličine koeficijenta korelacije koja je prihvaćena u oblasti u kojoj osoba radi.

Kada hoćemo da **procenimo verovatnu veličinu koeficijenta korelacije u populaciji, to se najčešće radi tako što testiramo statističku značajnost nulte hipoteze da je korelacija u populaciji jednaka nuli (0)** tj. da je uzorak uzet iz populacije u kojoj je korelacija nula. U okviru pristupa zasnovanog na centralnoj graničnoj teoremi, ovo testiranje se radi tako što se **prvo izračuna standardna greška koeficijenta korelacije**, a onda se **napravi interval poverenja koji obuhvata željeni deo teorijske distribucije uzorkovanja prema ovoj teoremi**. Najčešće korišćena formula za računanje **standardne greške koeficijenta korelacije** (Pirsonovog produkt-moment koeficijenta korelacije, opisan u narednom poglavlju) je:

$$SG_r = \sqrt{\frac{1 - r^2}{N - 2}}$$

U ovoj formuli, SG_r je standardna greška koeficijenta korelacije, r je veličina koeficijenta korelacije, a N je veličina uzorka tj. broj entiteta u uzorku. Ovo je opšta formula za računanje standardne greške koeficijenta korelacije, ali **kada standardnu grešku koeficijenta korelacije računamo za potrebe testiranja nulte hipoteze koja kaže da je koeficijent korelacije u populaciji jednak nuli, vrednost r u jednačini je 0** (zato što interval poverenja pravimo oko nule). Takođe, iz praktičnih razloga i kada imamo posla sa velikim uzorcima kakvi se tipično sreću u društvenim naukama, **zarad jednostavnosti, možemo skloniti i ono -2 iz jednačine**, zato što deljenje neke vrednosti sa 202 i sa 200 (ili čak sa 1002 naspram deljenja sa 1000) ne pravi nikakvu vidljivu praktičnu razliku i tada smo ovu jednačinu uprostili na:

$$SG_r = \frac{1}{\sqrt{N}}$$

Potom ovu **standardnu grešku pomnožimo odgovarajućim koeficijentom zavisno od kritičnog nivoa statističke značajnosti koji želimo da iskoristimo** (1,96 za nivo 0,05 odnosno

PREPRINT

2,56 za 0,01) i tako **dobijemo minimalnu veličinu koju koeficijent korelacije treba da ima da bi prešao prag statističke značajnosti.**

Korišćenje ove jednačine je **vrlo korisno kada treba da napravimo brzu procenu** toga koliko veliki treba da bude koeficijent korelacije da bi prošao kritični prag statističke značajnosti na uzorku određene veličine, međutim **naučni tekstovi u kojima se predstavljaju rezultati istraživanja obično ili jasno navode minimalnu veličinu koeficijenta korelacije koja prelazi prag statističke značajnosti** (minimalna veličina korelacije koja je potrebna da bi se odbacila nulta hipoteza) tj. **koji su statistički značajni, kako se to obično zove u naučnom žargonu ili navode precizno nivo statističke značajnosti za svaki koeficijent korelacije koji prikazuju**. U ovom drugom slučaju, **nultu hipotezu** (koja obično kaže da je korelacije u populaciji 0) **prihvatamo ili odbacujemo na osnovu toga da li je nivo statističke značajnosti** (kako je objašnjeno u prethodnom delu ove knjige) **iznad ili ispod našeg praga statističke značajnosti za ovu odluku.** Na primer, ako je statistička značajnost koeficijenta korelacije 0,023 a naš prihvaćeni prag statističke značajnost je 0,05, nultu hipotezu ćemo odbaciti (jer je 0,023 manje od 0,05). Ako je, pak, statistička značajnost našeg koeficijenta korelacije 0,22, nultu hipotezu ćemo prihvatiti. **Ako odlučimo da nultu hipotezu prihvatimo, onda smo obavezni da korelaciju o kojoj se radi tretiramo kao da je nulta, odnosno da je nema**, ili preciznije, treba da prihvatimo da nam podaci koje imamo ne omogućavaju da zaključimo da u populaciji postoji nenulta korelacija. **Ako odlučimo da nultu hipotezu odbacimo, onda tretiramo nivo korelacije u populaciji kao da je otprilike jednak vrednosti koeficijenta korelacije koji smo dobili na našem uzorku**. Takođe, za dodatnu preciznost, **može se napraviti i interval poverenja oko koeficijenta korelacije koji smo dobili na uzorku**. To se radi na način koji je objašnjen u delu knjige o intervalima poverenja.

Još jedan način da **testiramo nultu hipotezu o veličini koeficijenta korelacije u populaciji** je **kroz postupak butstrepinga**, kako je to opisano u delu knjige o statističkoj značajnosti. U tom slučaju se **pravi butstrep interval poverenja željene veličine** (to je najčešće 95% ili 99%) tako što se izvlači određeni veliki broj uzoraka sa vraćanjem iz uzorka istraživanja i onda na osnovu toga **pravi distribucija uzorkovanja koeficijenta korelacije**. **Ako se vrednost koja je pretpostavljena nultom hipotezom (tj. nula) nalazi unutar intervala poverenja napravljenog na ovaj način, onda prihvatamo nultu hipotezu**. **Ako se nalazi van tog intervala onda nultu hipotezu odbacujemo.** Na primer, ako bi ovim postupkom dobili interval poverenja od -0,15 do

PREPRINT

0,20, nultu hipotezu bi prihvatili jer se vrednost 0 nalazi unutar intervala čije su granice -0,15 i 0,20. Onda bi koeficijent korelacije tretirali kao da je 0 tj. da korelacije nema. Ako bi pak interval poverenja išao od npr. -0,35 do -0,20, nultu hipozu bi odbacili jer se vrednost 0 ne nalazi unutar tog intervala, te bi tretirali koeficijent korelacije koji smo dobili na uzorku kao valjanu procenu vrednosti parametra populacije.

6.5. Vrste koeficijenata korelacije

Ranije je rečeno da koeficijent korelacije nije jedan jedinstveni tip statistika, već da postoji puno različitih vrsta koeficijenata korelacije. Ovakva situacija je kako posledica toga da različiti autori predlažu različite načine računanja korelacije između varijabli, tako i posledica toga da postoje različite vrste povezanosti između varijabli (kako smo već pomenuli ranije), a ove različite vrste povezanosti se ne mogu valjano obuhvatiti jednom jedinom formulom, ili je to bar slučaj u ovom trenutku razvoja statistike kao nauke. U ovoj knjizi ćemo predstaviti neke od najčešće korišćenih koeficijenata korelacije, ali čitaoci treba da budu svesni da postoji mnogo, mnogo različitih vrsta koeficijenata korelacije koje su predložili i i danas predlažu različiti autori, tako da ovo što predstavljamo u ovoj knjizi nikako ne treba smatrati za iscrpnu listu mera povezanosti između varijabli.

Pirsonov produkt-moment koeficijent korelacije ili samo **Pirsonov koeficijent korelacije**, je verovatno najpoznatiji i najšire primenjivani koeficijent korelacije. Njegova upotreba je zasnovana na pretpostavci da je odnos između dve varijable, čija se povezanost izražava ovim koeficijentom, linearan, a **ovaj koeficijent onda izražava snagu linearne povezanosti između te dve varijable**. U literaturi ovaj koeficijent nalazimo i kao samostalni pokazatelj linearne povezanosti između varijabli, ali i kao deo mnogih kompleksnijih statističkih postupaka. Tipično se obeležava sa **„r“.** To je **parametrijska statistička mera** koja zahteva da varijable budu bar na intervalnom nivou merenja, a da podaci budu normalno distribuirani. U osnovi, on se računa kao **suma proizvoda z skorova entiteta na varijablama između kojih se računa korelacija koja je podeljena brojem stepeni slobode** (ili, ako hoćemo da uprostimo stvari – brojem entiteta u uzorku, razlika između broja entiteta i broja stepeni slobode je samo 1 i stoga je praktično potpuno zanemarljiva kod velikih uzoraka):

PREPRINT

$$r = \frac{\sum_{i=1}^{N} z1_i * z2_i}{N-1}$$

U ovoj jednačini, r je Pirsonov produkt-moment koreficijent korelacije, z1 i z2 su vrednosti varijabli čije korelacije računamo izraženi kao z skorovi, a N je broj entiteta u uzorku tj. broj parova podataka u z1 i z2. Treba naglasiti da ovde predstavljamo ovu verziju formule Pirsonovog koeficijenta zarad jednostavnosti, ali da **postoje verzije formule za računanje ovog koeficijenta** (i mogu se izvesti iz ove formule), **kod kojih se podaci unose u jednačinu u svoj sirovom obliku**, bez prethodne konverzije u z skorove.

Kao što je i inače slučaj sa koeficijentima korelacije, **raspon mogućih vrednosti Pirsonovog koeficijenta korelacije je od 0 do ±1** (+1 ili -1). Ako pogledamo formulu za računanje ovog koeficijenta korelacije, možemo da primetimo nekoliko stvari:

- Kako je ovaj koeficijent zasnovan na z skorovima, to znači da **jedinice merenja** dve varijable čiju povezanost računamo, kao i skale na kojima su **nisu ni od značaja za veličinu koeficijenta**. U postupku računanja koeficijenta korelacije (ili pre početka računanja, ako koristimo formulu koja je gore predstavljena) vrednosti varijabli se konvertuju u z skorove, a to znači da one u postupak računanja koeficijenta korelacije ulaze sa identičnim standardnim devijacijama (tj. varijansama) i identičnim aritmetičkim sredinama. To znači da se Pirsonov koeficijent može računati za bilo koje dve varijable dokle god su njihove distribucije normalne (ili dovoljno blizu normalnim).
- Kako se radi o proizvodu z skorova entiteta na dve varijable između kojih računamo korelaciju, kada su oba z skora pozitivna i kada su oba z skora negativna, njihov proizvod će imati pozitivnu vrednosti. Međutim, kada je jedan od z skorova pozitivan, a drugi je negativan, njihov proizvod će imati negativnu vrednost. Iz toga sledi da ako je suma proizvoda z skorova entiteta koji imaju vrednosti na istoj polovini dsitribucije (gornjoj ili donjoj polovini) veća od sume proizvoda z vrednosti entiteta koji imaju vrednosti na suprotnim polovinama distribucije dve varijable, koeficijent korelacije će biti pozitivan. Ako ova druga suma nadmašuje ovu prvu, koreficijent korelacije će biti negativan. Međutim, ako su ove dve sume jednake, koeficijent korelacije će biti 0. To znači da **čak i u situacijama kada je korelacija između dve varijable pozitivna, mogu postojati (neobični) slučajevi koji imaju visoke vrednosti na jednoj varijabli, a niske na drugoj**. Isto tako, čak i kada je korelacija negativna, mogu postojati slučajevi kod koji imaju

PREPRINT

vrednosti sa sličnim pozicijama na distribucijama obe varijable (tj. pozitivne vrednosti na obe ili negativne vrednosti na obe). Međutim, **takvi slučajevi će smanjiti koeficijent korelacije** i što je niži koeficijent korelacije, to će biti više prilike za postojanje takvih slučajeva.

- Napred izneseno takođe znači i da ako naš uzorak dolazi iz dve različite populacije, jedne u kojoj postoji pozitivna korelacija i druge sa negativnom korelacijom između varijabli, na uzorku sastavljenom od entiteta iz ove dve populacije zajedno, imaćemo nultu korelaciju (ili mnogo nižu korelaciju u jednom ili drugom smeru, zavisno od toga koliko su visoke korelacije unutar svake grupe i koliki je odnos veličina grupa tj. broja entiteta u jednoj i drugoj grupi).
- Proizvodi brojeva većih od 1 su veći od brojeva koji se množe, dok su proizvodi brojeva manjih od 1 manji od brojeva koji se množe. Kako je brojilac u formuli za Pirsonov koeficijent korelacije suma ovih proizvoda, to znači da će proizvod većih vrednosti doprinositi više ukupnoj sumi nego proizvod malih vrednosti. To je razlog zašto **entiteti sa ekstremnim vrednostima na obe varijable mogu imati vrlo snažan i disproporcionalno jak uticaj na vrednost koeficijenta korelacije**. Jedan ekstremni primer sa odnosom vrednosti u jednom smeru može anulirati efekte većeg broja slučajeva čiji odnos vrednosti ukazuje na drugi smer povezanosti. Na primer, ako bi imali jedan jedini entitet koji ima vrednosti z skorova od +7 na obe varijable, proizvod ta dva skora bi bio 49. Da dovedemo ukupnu sumu na nulu tj. da poništimo efekat tog jednog entiteta, trebalo bi nam 196 entitea čiji z skorovi na jednoj varijabli iznose 0,5, a na drugoj -0,5. Naravno, ako bi tako izgledao ukupni uzorak, mi verovatno ne bi ni računali Pirsonov koeficijent korelacije, zato što tada uslov da distribucija bude normalna ne bi bio ispunjen, čak i da su ovakvi entiteti samo deo ukupnog uzorka (naravno, supstantivni – veliki deo. Ako bi to bio neki vrlo ogroman uzorak gde su ovih skoro 200 entiteta zanemarljivi deo, to bi bila druga priča). Međutim, ovaj primer koristimo samo da bi ilustrovali kako visoke vrednosti z skorova mogu da disproporcionalno svom značaju promene vrednost Pirsonovog koeficijenta. Ovo je **posebno važno imati u vidu zato što su ekstremne vrednosti** u istraživanjima dosta **često posledica grešaka u unosu podataka ili nevalidnog merenja**, te bi ovakve pojave trebalo rešiti pre početka računanja korelacije, čak i da ne budu primećene prilikom pregleda oblika distribucije.

PREPRINT

Kvadrat Pirsonovog koeficijenta korelacije zove se koeficijent determinacije i interpretira se kao proporcija varijanse koju dve varijable dele (pod pretpostavkom da je njihova povezanost potpuno linearna).Treba da imamo u vidu da korelacija znači da se dve varijable menjaju zajedno tj. da promene vrednosti jedne varijable prate promene vrednosti druge varijable, a da je ovo stvar stepena. Kada se ovaj stepen predstavi kao proporcija ukupne varijabilnosti koja je izražena kao varijansa, onaj deo te ukupne varijabilnosti koji se dešava zajedno u dve varijable izražen je koeficijentom determinacije. Kako su varijanse dve varijable za koje se računa Pirsonova korelacije izjednačene pre ili u procesu računanja korelacije kroz konverziju u z skorove, **proporcija vartijanse koju jedna varijabla deli sa onom drugo je jednaka proporciji varijanse koju druga deli sa prvom**. Drugim rečima, ova proporcija je jednaka za obe varijable. **Koeficijent determinacije se često koristi kao pokazatelj intenziteta veze između dve varijable u sklopu različitih multivarijatnih statističkih postupaka** (koji nisu predstavljeni u ovoj knjizi, ali će neke od njih biti predstavljene u našoj knjizi posvećenoj multivarijatnim statističkim tehnikama, koja je u ovom trenutku već ugovorena sa izdavačem!), ali se nekada koristi i kao samostalni pokazatelj intenziteta povezanosti između varijabli, jer ima dosta istraživača koji vole ovu meru baš zato što je u pitanju proporcija, te je **njena veličina nekad zahvalnija za interpretaciju nego što je to slučaj sa koeficijentom korelacije**.

Pirsonov koeficijent korelacije se takođe **računa i u sklopu različitih multivarijatnih statističkih postupak**a, gde se koristi kao pokazatelj korelacije između varijabli koje su izvedene iz varijabli dobijenih na uzorku. U takvim situacijama **ovaj koeficijent ima različita drugačija imena, u zavisnosti od prirode varijabli za koje se računa korelacija i u zavisnosti od toga kako je izveden.** Ova imena uključuju „kanoničku korelaciju“, „koeficijent multiple korelacije“, „faktorsko zasićenje“, „koeficijent strukture“, „koeficijent krosstrukture“, „parcijalna korelacija“, „semiparcijalna korelacija“ itd.

Spirmanov koeficijent korelacije rangova je koeficijent korelacije koji je namenjen varijablama koje su bar **na ordinalnom nivou merenja**, a **zasnovan je na pretpostavci da je povezanost između dve varijable monotona.** To je **neparametrijski statistik**, što znači da se njegovo korišćenje ne oslanja na bilo kakvu pretpostavku o obliku distribucije varijabli između kojih se korelacija računa. Postoje različite formule za računanje ovog koeficijenta, međutim verovatno je najjednostavniji način to da se rangiraju podaci tj. da se **podaci konvertuju u rangove na svakoj**

PREPRINT

varijabli posebno i da se onda izračuna Pirsonov koeficijent korelacije između ova dva niza rangova. Rangovi se kreiraju tako što se prosto dodeli rang 1 entitetu koji ima najnižu vrednosti, rang 2 entitetu sa drugom najnižom vrednošću i tako redom, do najviše vrednosti. Ovo se radi za svaku od dve varijable posebno**. Spirmanov koeficijent korelacije rangova se tipično označava sa grčkim slovom ro – ρ,** mada bi tu čitaoci trebali da budu pažljivi, jer se **isto slovo koristi i da označi vrednost Pirsonovog koeficijenta korelacije u populaciji** (kao kontrast vrednosti ovog koeficijenta na uzorku koja se označava obično sa r), a moguće je da se ovim znakom označe i drugi statistici.

Point-biserijski koeficijent korelacije je Pirsonov produkt-moment koeficijent korelacije izračunat između jedne binarne i jedne intervalne varijable. Podsetimo se – binarne varijable su varijable koje imaju samo dve moguće vrednosti. Zovu se i dihotomne varijable. **Point-biserijski koeficijent korelacije se tipično koristi da prikaže stepen preklapanja između distribucija dve grupe i, u isto vreme, veličinu razlike između aritmetičkih sredina tih grupa.** To je razlog zašto se on često u literaturi pominje samo pod imenom **"veličina efekta"** (kada je razlog za njegovo računanje procena veličine razlike između aritmetičkih sredina dve grupe), mada treba imati u vidu da se i različite druge statističke mere u literaturi označavaju kao veličine efekta te onda čitalac treba da obrati pažnju na to o kom statistiku se zapravo radi. U ovoj situaciji, kada se point-biserijski koeficijent korelacije koristi kao pokazatelj veličine razlike između aritmetičkih sredina dve grupe, binarna varijabla sadrži informacije o tome kojoj grupi (od dve poređene) pripada entitet tj. jedna grupa se označava jednom vrednošću binarne varijable, a druga drugom vrednošću binarne varijable. Računanje point-biserijskog koeficijenta korelacija onda implicira da će jedna jedna grupa biti obeležena sa npr. 0, a druga sa 1 (ili sa bilo koja dva različita broja), a kako se binarne varijable mogu tretirati kao da su na bilo kom nivou merenja, i intervalna varijabla (ona na kojoj su entiteti mereni ili procenjivani) i binarna varijabla mogu biti uneti u formulu za računanje Pirsonove korelacije, samo što se statistik koji se tako dobije neće zvati Pirsonov koeficijent korelacije, nego point-biserijski koeficijent korelacije.

Kao što je to slučaj sa Pirsonovim koeficijentom korelacije, pointbiserijska korelacija takođe ima raspon vrednosti od -1 do +1, samo što se interpretacija donekle razlikuje. **Pojmovi pozitivne i negativne korelacije nemaju mnogo smisla u situaciji kada je jedna od varijabli binarna te to da li će korelacije biti pozitivna ili negativna zavisi isključivo od toga kojoj ćemo od dve**

PREPRINT

grupe dodeliti veći broj. Ako ispadne da smo grupu sa većom aritmetičkom sredinom označili i većim brojem, onda će point-biserijski koeficijent korelacije biti pozitivan. Ako grupa označena manjim brojem ima višu aritmetičku sredinu, point-biserijski koeficijent korelacije će biti negativan. **Ako preokrenemo to kojoj smo grupi dodelili koji broj**, pa grupu koja je pre bila označena većim brojem sada označimo manjim i obrnuto, **predznak koeficijenta point-biserijske korelacije će se takođe promeniti**. Na primer, ako imamo grupu ljudi iz Leskovca i grupu ljudi iz Surdulice i želimo da im poredimo prosečnu telesnu masu mogli bi, recimo, da Surdulicu obeležimo sa 0, a Leskovac sa 1, pa ako bi se desilo da Leskovčani imaju više prosečne telesne mase (podsetnik: prosek je isto što i aritmetička sredina!) od Surduličana, koeficijent point-biserijske korelacije između grada u kome osoba živi (varijabla 1) i telesne mase (varijabla 2) bi bio pozitivan. Ako bi pak ispalo da veću prosečnu telesnu masu imaju Surduličani, koeficijent pointbiserijske korelacije između grada u kome osoba živi i telesne mase bi bio negativan. Ovo je zbog toga što su u ovom primeru stanovnici Surdulice, koji (u ovom primeru, izmišljenom) imaju višu aritmetičku sredinu na merenoj varijabli, označeni manjim od dva broja kojima je označavano kojoj od dve grupe (Leskovac je jedna, Surdilica druga) osoba pripada. Na sličan način, ako bi dobili point-biserijski koeficijent korelacije od, recimo, -0,75 u istraživanju koje poredi dve grupe u pogledu prosečnih vrednosti na nekoj intervalnoj varijabli, to bi značilo da grupa koja je označena manjim brojem ima veću aritmetičku sredinu od te dve poređene grupe.

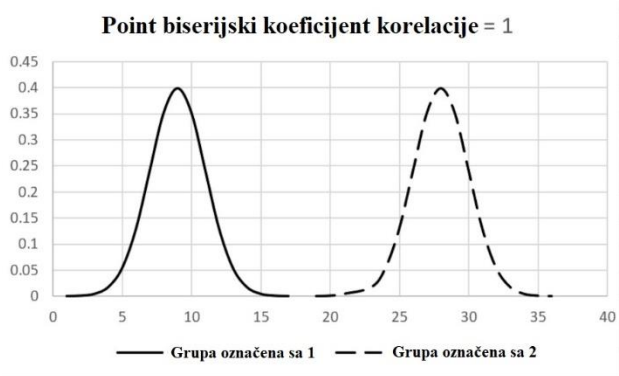
Kada je u pitanji intenzitet koeficijenta point-biserijske korelacije, ranije smo pomen uli da on predstavlja stepen preklapanja distribucija dve poređene grupe na intervalnoj varijabli (ovo, treba primetiti, podrazumeva da su obe grupe normalno distribuirane, jer ovo je parametrijski postupak!). **Kada je koeficijent point-biserijske korelacije 0, to znači da da se distribucije dve grupe u potpunosti preklapaju tj. da nema razlike ni između aritmetičkih sredina dve grupe ni između njihovih distribucija.** Potpuno se preklapaju. **Kada je point-biserijski koeficijent korelacije 1 (ili -1), to znači da nema uopšte preklapanja između distribucija dve grupe.** Drugim rečima, **ako znamo da vrednosti bilo kog entiteta na intervalnoj varijabli, kada je korelacija 1 ili -1, možemo bez ikakvih grešaka, da predvidimo kojoj grupi na posmatranoj binarnoj varijabli pripada dati entitet tj. kojoj od dve grupe pripada. Obrnuto ne važi, jer entiteti iz iste grupe mogu imati raspon različitih vrednosti na intervalnoj varijabli.** Kada je point-biserijska korelacija 1 (ili -1) možemo samo predvideti raspon vrednosti unutar kog se nalazi dati entitet ako znamo to kojoj grupi pripada, odnosno ako znamo njegovu vrednost na binarnoj

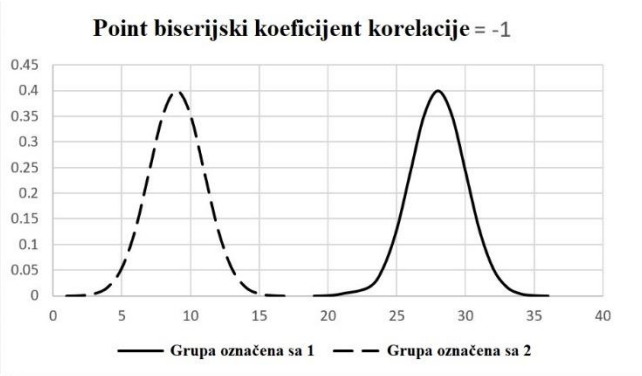
PREPRINT

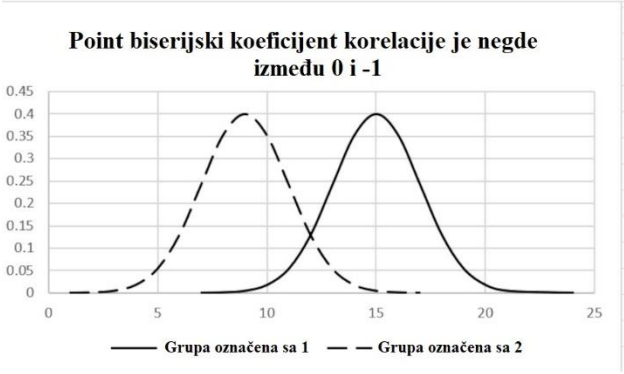
varijabli. Naravno, kada je point-biserijska korelacija između 0 i 1 (ili između 0 i -1, što je isto!) moći ćemo da predvidimo kojoj grupi pripada entitet na osnovu vrednosti intervalne varijable bolje nego što bi to uspeli slučajnim pogađanjem, ali što je korelacija niža (bliža nuli), više ćemo grešaka napraviti i one će biti veće. Ovo je zbog toga što će preklapanje distribucija dve posmatrane grupe na intervalnoj varijabli biti sve veće što je point-biserijska korelacija niža.

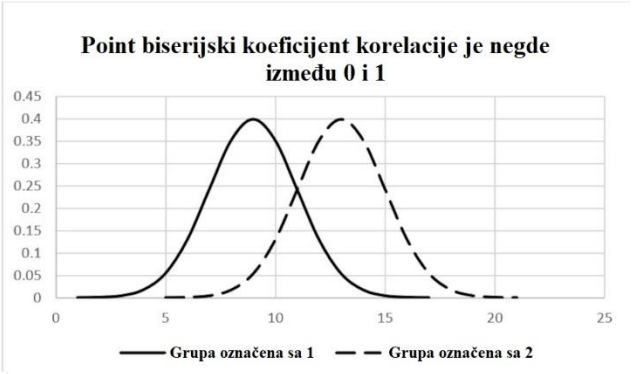
Slika 6.5. Grafički prikaz različitih stepena preklapanja distribucija dve grupe i njima odgovarajući point biserijski koeficijenti korelacije. Na slici možemo videti da je, kada se distribucije ne preklapaju, korelacija 1 ili -1. Korelacija je 1 onda kada grupa koja je označena većim brojem ima veću aritmetičku sredinu. Korelacija je -1 onda kada grupa označena većim brojem ima manju aritmetičku sredinu. Kada se dve distribucije preklapaju, ali ne potpuno, koeficijent korelacije ima vrednosti između 0 i 1 (odnosno 0 i -1, zavisno od toga koja grupa je označena većim brojem prilikom računanja). Konačno, kada se distribucije dve grupe poklapaju, koeficijent je 0.

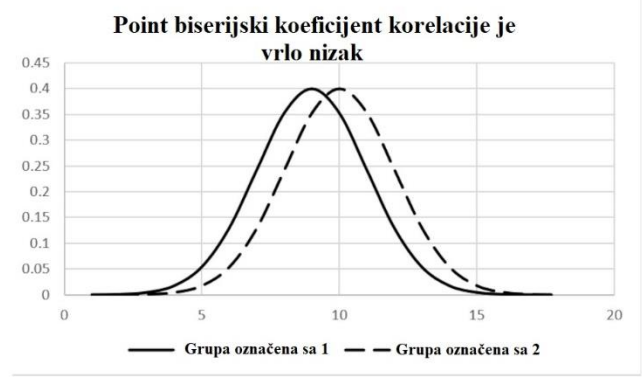
PREPRINT

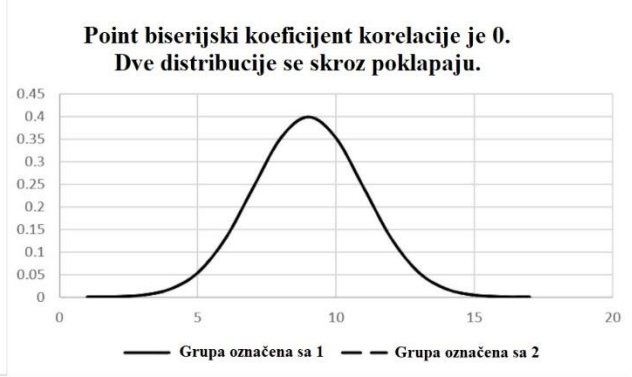












Jedan važan zahtev kada se računanje korelacije radi na ovaj način je da **binarna varijabla mora da bude prirodno binarna** tj. da dve vrednosti binarne varijable nisu samo široke kategorije u koje je podeljen neki kontinuum različitih vrednosti. Na primer, ako bismo hteli da poredimo Leskovčane (grupa 1) i Surduličane (grupa 2) u pogledu njihove telesne mase, mogli bismo da sakupimo uzorak Leskovčana i uzorak Surduličana, da im izmerimo telesnu masu i da onda računamo point-biserijski koeficijent korelacije između mesta življenja tj. binarne varijable koja sadrži podatak o tome da li osoba živi u Leskovcu ili u Surdulici, i telesne mase, koja je varijabla na racio nivou merenja. Ovakva računica bi bila adekvatna zato što je mesto življenja, sa mogućim vrednostima Surdulica i Leskovac prava binarna varijabla. S druge strane, ako bi, recimo, hteli da ispitamo da li studenti koji su pali pismeni ispit iz nekog fakultetskog predmeta i oni koji su ga

položili imaju različite stavove prema tom predmetu (zamislimo da se stav meri korišćenjem nekog validnog testa za merenje stavova i da je to normalno distribuirana mera na intervalnom nivou merenja) i onda da bi to uradili napravili binarnu varijablu koja sadrži informaciju o tome da li je osoba pala ili položila ispit izvedenu iz njegovog/njenog skora na pismenom testu, to ne bi bila adekvatna postavka za point-biserijski koeficijent korelacije. Ovo je zbog toga što binarna varijabla koja sadrži informacije o tome da li je osoba položila ili pala test nije prava binarna varijabla tj. nije prava dihotomija, jer je izvedena iz kontinuuma koji sadrži različite testovne skorove. Studenti koji su pali test su imali čitav niz različitih testovnih skorova, kojima je samo zajedničko to da su svi ispod praga potrebnog da bi se položio test. A slična je situacija i sa onima koji su test položili – oni isto tako imaju čitav niz različitih skorova (a to znači i različitih nivoa znanja o temi kojom se predmet bavi), samo što su svi ti skorovi iznad nivoa potrebnog da se položi ispit. U takvim sitaucijama koeficijent korelacije koji treba računati **je biserijski koeficijent korelacije**. Iako se zove slično, **formula za računanje biserijskog koeficijenta korelacije je dosta različita od one za računanje point-biserijskog koeficijenta korelacije**. A sada, kada to znamo, treba reći i da je u situacijama kao što je ova iz prethodnog primera, mnogo bolji izbor prosto izračunati Pirsonov ili Spirmanov koeficijent korelacije između dve intervalne ili ordinalne varijable i uopšte ne komplikovati stvari pravljenjem veštačke dihotomije.

Eta koeficijent korelacije (η) je koeficijent koji **služi kao mera povezanosti jedne nominalne i jedne intervalne varijable, ali i kao mera monotone nelinearne ili nemonotone povezanosti dve intervalne varijable**. Nominalna varijabla koja ulazi u računanje eta koeficijenta može da bude prava nominalna, **ali može da bude i diskretna varijabla koja nema previše kategorija** (na primer, intervalna koja je svedena na nekoliko grupa). **Kada je odnos između dve varijable linearan, eta i Pirsonov koeficijent korelacije su jednaki.** Međutim, **kada odnos nije linearan, eta će biti viša od Pirsonovog koeficijenta korelacije.** Takođe, **kada je odnos između dve varijable nemonoton, računaju se dve različite vrednosti eta koeficijenta – jedna za predviđanje jedne varijable na osnovu druge i jedna za predviđanje druge varijable na osnovu prve**. Ovo je zbog toga što kod nemonotonih odnosa, zbog promena smera povezanosti između varijabli, preciznost sa kojim se jedna varijabla može predvideti na osnovu druge može da bude različita od preciznosti sa kojom se druga može predvideti na osnovu prve. Na primer, u ranije već pominjanom primeru sa odnosom poznatim kao Daning-Kruger efekat (Kruger & Dunning, 1999) između samoprocene kompetencija i pravog nivoa kompetencija, bili bismo

PREPRINT

verovatno uspešniji u predviđanju kako će osoba samoproceniti svoje kompetencije ako znamo pravu kompetenciju osobe nego što bi bili ako bi trebali na osnovu samoprocene kompetencija da procenimo prave kompetencije osobe. Ovo je zato što kako prava kompetentnost raste, samoprocena kompetencija prvo počinje da raste donekle, onda posle određenog nivoa prave kompetencije samoprocena kompetencije počela da se smanjuje, a onda, sa daljim rastom prave kompetencije samoprocena ponovo počinje da raste (ovo se odnosi na promene u opaženom testnom skoru kao funkciji pravog testnog skora, kako je to predstavljeno u istraživanjima 2 i 3 rada Krugera i Daninga koji je predstavio prvi ovaj efekat). U takvoj situaciji, visoka samoprocenjena kompetencija odgovara i visokim i niskim nivoima prave kompetencije, ali takva dupla korespondencija ne postoji kada predviđamo samoprocenu kompetencije na osnovu prave kompetencije. **Kada se računa eta, varijabla koja se koristi za predviđanje druge varijable zove se nezavisna varijabla, dok se varijabla koju predviđamo zove zavisna varijabla**. U dva računanja eta koeficijenta, prema tome, varijable menjaju mesta tako da je svaka jednom zavisna i jednom nezavisna varijabla. Međutim, **kada se eta koristi za računanje povezanosti jedne nominalne i jedne intervalne varijable, tipično se eta koeficijent prikazuje samo za nominalnu varijablu kao nezavisnu, a za intervalnu varijablu kao zavisnu**. Ovo **svojstvo, da statistik ima različite vrednosti zavisno od toga koja se varijabla tretira kao zavisna, a koja kao nezavisna zove se asimetrija** (u tom pogledu eta predstavlja suprotnost npr. Pirsonovom koeficijentu korelacije, koji je simetrični statistik). Međutim, iako su imena nezavisna i zavisna varijabla ista ona imena za varijable koja se sreću u eksperimentalnim istraživanjima, treba još jednom da ponovimo da korelacije same za sebe ne mogu da ustanove postojanje uzročno-posledičnih odnosa između varijabli i to važi i za eta koeficijent.

Kada se koristi da izrazi stepen povezanosti između nominalne i intervalne varijable, eta se može interpretirati kao pokazatelj stepena preklapanja između distribucija grupa entiteta sa različitim vrednostim na nominalnoj varijabli na intervalnoj varijabli. Dakle, entiteti koji imaju istu vrednost na nominalnoj varijabli čine istu grupu, i onda za svaku vrednost imamo grupu entiteta koji imaju datu vrednost. Svaka od tih grupa ima svoju distribuciju vrednosti na intervalnoj varijabli (to je ona druga varijabla koja ulazi u računanje korelacije), za koju očekujemo da ima oblik normalne distribucije. U toj situaciji eta između date nominalne varijable i date intervalne varijable se može tretirati kao mera stepena preklapanja distribucija ovih grupa varijabli na intervalnoj varijabli. U istoj situaciji se eta koeficijent može interpretirati i kao pokazatelj veličine

PREPRINT

razlike između aritmetičkih sredina ovih grupa. Pošto se svaka grupa sastoji od entiteta koji na nominalnoj varijabli imaju istu vrednost, grupa će biti onoliko koliko nominalna varijabla ima vrednosti (ponovo – jedna vrednost na nominalnoj varijabli određuje jednu grupu). **Kada je eta 0, aritmetičke sredine svih grupa su jednake aritmetičkoj sredini celog uzorka (celi uzorak čine sve grupe zajedno), a njihove distribucije se poklapaju. Kako eta postaje veće, tako se i nivo preklapanja između distribucija grupa (na intervalnoj varijabli) smanjuje.**

Eta je parametrijski statistik što znači da očekujemo da je intervalna varijabla normalno distribuirana. Kada **se eta koristi za računanje povezanosti između nominalne i intervalne varijable, ne bi trebalo da bude vrednosti nominalne varijable sa previše niskim frekvencijama**. Šta tačno znači previše niska, u principu zavisi od konteksta određene oblasti istraživanja, ali u principu **ne bi trebalo da bude grupa u kojima je broj entiteta jednocifren**. Zato što je eta koeficijent nemonotone povezanosti (a i zato što je jedna od varijabli nominalna), **eta koeficijent nema predznak** tj. ne može biti pozitivan ili negativan.

Za izražavanje veličine razlike između aritmetičkih sredina više grupa, eta kvadrat se često prikazuje pored eta koeficijenta ili umesto eta koeficijenta. **Eta kvadrat se dobija tako što se eta koeficijent podigne na kvadrat tj. na drugi stepen** (eta koeficijent se pomnoži sa samim sobom – eta x eta). Eta kvadrat se interpretira na sličan način kao kvadrat Pirsonovog koeficijenta korelacije – koeficijent determinacije tj. kao proporcija varijanse koja je zajedničkaa za dve varijable između kojih je računata povezanost ovim koeficijentom.

Fi koeficijent korelacije je Pirsonov koeficijent korelacije koji se **računa između dve binarne varijable**. Kod ovog koeficijenta korelacije, **predznak označava koja kategorija jedne varijable teži da bude povezana sa kojom kategorijom druge varijable** tj. da se javlja zajedno sa njom. S obzirom da su obe varijable binarne i da su njihove kategorije označene brojevima, kod svake varijable će jedna kategorija biti označena većim brojem, a druga manjim brojem. Ako entiteti koji pripadaju kategoriji koja je označena većim brojem na jednoj varijabli, teže da pripadaju kategoriji koja je takođe označena većim brojem na drugoj varijabli, predznak fi koeficijenta će biti pozitivan. U obrnutoj situaciji – kada entiteti koji pripadaju kategoriji koja je označena većim brojem na jednoj varijabli teže da pripadaju kategoriji koja je označena manjim brojem na drugoj varijabli, fi koeficijent će biti negativan. Slično kao kod interpretacije point-biserijske korelacije, **predznak ovog koeficijenta ne bi trebalo tumačiti kao pozitivnu ili**

PREPRINT

negativnu korelaciju, već samo kao pokazatelj toga koje kategorije dve posmatrane binarne varijable međusobno odgovaraju. Način na koji se brojevi dodeljuju dvema kategorijama binarne varijable tj. to koja kategorija će biti označena većim, a koja manjim brojem je najčešće potpuno proizvoljno što znači da je kategorija označena manjim brojem podjednako validno mogla da bude označena i većim brojem i obrnuto. Zbog toga, nema puno smisla da se predznak fi koeficijenta korelacije tumači kao smer korelacije jer se to može lako promeniti samo ako označimo kategorije brojevima na drugačiji način. Na primer, ako bi imali dve binarne varijable – jednu koja sadrži informaciju o tome da li je dete učenik osnovne ili srednje škole i još jednu varijabli koja označava da li dete više voli da nosi plave ili bele majice, recimo da smo tu varijablu nazvali „majica“, ne bi imalo nikakvog smisla reći da je korelacije između varijabli „škola“ i „majica“ pozitivna ili negativna. Bilo bi, s druge strane, itekako smisleno reći da, recimo, učenici osnovne škole teže da više vole bele majice, a da učenici srednje škole teže da više vole plave majice, naravno ako bismo zaista dobili koeficijent korelacije koji tako nešto pokazuje. Na sličan način, ako bi naše dve varijable bili to da li osoba pripada grupi A ili grupi B u nekoj školi (varijabla „grupa“) i to da li je osoba pala ili položila neki test (varijabla „test“), ne bi bilo mnogo informativno da kažemo da postoji pozitivna ili negativna korelacija između „grupe“ i „testa“. Međutim, bilo bi sasvim smisleno da kažemo da je više ljudi iz jedne grupe položilo test nego iz druge grupe (ako je, naravno, to ono što se zaista desilo i što korelacija pokazuje).

Koeficijent kontingencije je **koeficijent korelacije namenjen opisivanju nivoa povezanosti dve nominalne varijable**. On se, u neku ruku, može smatrati prilagođenom verzijom fi koeficijenta za (nominalne) varijable sa više od dve kategorije. Međutim, za razliku od fi koeficijenta, **koeficijent kontingencije nema predznak** tj. vrednost mu je uvek pozitivna. U principu, koeficijent kontingencije od 0 ukazuje da dve varijable za koje je računat nisu povezane tj. da pripadanje određenoj kategorije jedne varijable ne ukazuje na višu ili nižu verovatnoću pripadanja nekoj određenoj kategoriji druge varijable. S druge strane, što je veći koeficijent korelacije, to je jača povezanost između dve varijable što znači da će entiteti koji pripadaju određenoj kategoriji jedne varijable imati veću verovatnoću da pripadaju nekoj određenoj kategoriji (ili grupi kategorija) druge varijable. Za razliku od situacije sa fi koeficijentom, kada računamo koeficijent kontingencije, **informacija o tome koja kategorija jedne varijable je povezana sa kojom kategorijom druge varijable se ne može izvesti samo iz koeficijenta, već tipično zahteva pregled tabele sa ukršetnenim frekvencijama kategorija dve varijable.**

PREPRINT

Takođe je važno napomenuti i da **koeficijent kontingencije nikada neće imati vrednost 1**, čak i kada postoji potpuna povezanost dve varijable, već će samo imati vrednosti koje su blizu 1. Takođe, **veličina koeficijenta kontingencije zavisi od broja kategorija dve varijable** – koeficijent kontingencije će **težiti da ima veće vrednosti kada je veći broj kategorija dve varijable**, ali će težiti i da **ima manje vrednosti za nominalne varijable sa manjim brojem kategorija**. Često se može čuti da istraživači **preporučuju da se koeficijent kontingencije računa tek za nominalne varijable od kojih svaka ima bar po 5 kategorija.**

Kramerov V je još jedan popularni koeficijent korelacije koji opisuje **nivo povezanosti dve nominalne varijable**. Kramerov V se **računa po formuli koja dozvoljava da gornja granica mogućih vrednosti ovog koeficijenta bude 1.** To znači da će u slučaju savršene povezanosti između varijabli, vrednost Kramerovog V koeficijenta biti 1 (za razliku od koeficijenta kontingencije koji ne može da dostigne 1) i to je razlog zašto mnogi istraživači više vole da koriste ovaj koeficijent kao meru povezasnosti dve nominalne varijable nego koeficijent kontingencije.

6.6. Hajde da primenimo šta smo do sada naučili!

Hajde da probamo sada da primenimo stvari koje smo predstavili u ovom poglavlju kroz nekoliko vežbi. Molimo vas da pogledate opšte uputstvo za ovakve vežbe koje možete naći na početku knjige. Naša preporuka je da prvo pročitate svaki isečak i tvrdnje date u njemu i da onda date svoj odgovor. Odgovor možete upisati u kolonu za odgovore, a posle toga pročitajte odgovore i uporedite svoje odgovore sa njima.

Vežba L. Korelacije.

Varijable		Poverenje	Bezbednost	Nezavisnost/ autonomija	Drugarstvo
Kvalitet života	Društveni odnosi	.209*	.225*	.126*	.205*
	Fizičko zdravlje	.063*	-.086*	-.052*	-.052*
	Uslovi života	.034	.056*	.096*	.124*
	Mentalno zdravlje	.062*	.073*	.077*	.099*

* - statistički značajno bar na nivou 0,05.
Tabela je napravljena na osnovu podataka iz Andelković et al. (2012).

PREPRINT

L	Tvrdnja:	Odgovor
L1.	Kako raste Poverenje, tako i Uslovi života teže da rastu (da se popravljaju).	
L2.	Kako Nezavisnost/autonomija raste, Fizičko zdravlje teži da se smanjuje.	
L3.	Kada se Fizičko zdravlje smanjuje, Drugarstvo takođe teži da se smanjuje.	
L4.	Niže vrednosti na Mentalnom zdravlju teže da budu povezane (da korespondiraju) sa nižim vrednostima na Poverenju.	
L5.	Bezbednost je na nominalnom nivou merenja.	
L6.	Sve prikazane korelacije su statistički značajne.	
L7.	Intenzitet svih prikazanih statistički značajnih korelacija je niska, ako se vodimo Koenovim preporukama.	
L8.	Poverenje nije povezano sa Uslovima života.	
L9.	Korelacije varijabli predstavljenih u kolonama teže da budu više sa Uslovima života nego sa bilo kojom drugom varijablom od onih predstavljenih u redovima.	
L10.	Niži nivoi Fizičkog zdravlja teže da budu povezani (da korespondiraju) sa višim nivoima Bezbednosti.	

Vežba M. Korelacije

	R	I	A	S	E	C
R	-					
I	.45	-				
A	.26	.60	-			
S	.14	.40	.48	-		
E	.37	.25	.26	.62	-	
C	.64	.30	.14	.20	.54	-

Sve korelacije više od 0,07 su statistički značajne bar na nivou 0,05. Prikazane korelacije su Pirsonovi produkt-moment koeficijenti korelacije. Tabela je zasnovana na podacima iz istraživanja V. Hedrih (2011)

M	Tvrdnja:	Odgovor
M1.	R i I su u korelaciji umerenog intenziteta, ako sudimo po Tejlorovim (Taylor) preporukama.	
M2.	Ako bi na glavnoj dijagonali bili prikazani koeficijenti korelacije umesto crtica, sve te korelacije bi bile 0.	
M3.	Svi predstavljeni koeficijenti korelacije su statistički značajni.	
M4.	Korelacija između A i C je niska, sudeći prema Koenovim preporukama.	
M5.	A i S nisu povezani.	
M6.	E je uzrok C.	
M7.	Kada R raste, raste i C.	
M8.	Kada se C smanjuje, smanjuje se i I.	
M9.	Kada bi korelacije bile prikazane u gornjem trouglu tabele, koeficijenti bi bili isti kao oni koji su prikazani u donjem trouglu.	

PREPRINT

M10.	Prikazani koeficijenti korelacije su Spirmanovi koeficijenti korelacije rangova.	

Vežba N. Korelacije

Descriptive statistics and intercorrelations of variables in the study

	M	*SD*	1	2	3	4	5	6	7	8	9	10	11
Children													
1. CovFear_C	2.83	0.83											
2. Age_Children	12.78	3.57	-.15**										
Parents													
3. Distress NSE	2.76	1.00	.03	.04									
4. Distress CF	3.02	0.67	-.03	.05	.06								
5. Distress PD	2.68	0.92	.37**	.02	.21**	.07							
6. SelfEf	5.62	0.99	.10*	-.14**	-.18**	.11*	-.15**						
7. ER_CR	5.19	1.25	.10	-.07	-.19**	-.02	-.06	.32**					
8. ER_ES	3.36	1.40	.10*	-.04	-.18**	.03	.09	.00	.30**				
9. SomAnx	1.39	0.57	.19**	.08	.17**	.04	.40**	-.21**	-.09	.10			
10. CogAnx	1.70	0.63	.17**	.06	.21**	.07	.49**	-.22**	-.13*	.07	.70**		
11. CovFear_P	2.73	0.74	.49**	.03	.03	.08	.66**	-.07	-.02	.07	.37**	.44**	
12. PanParent	3.98	0.63	.24**	-.27**	-.04	-.02	.08	.24**	.34**	.03	.03	.04	.12*

*$p < .05$; **$p < .01$.

Note. CovFear_C = Children's Fear of COVID-19; Age_Children = Children's age; Distress NSE = Parental distress due to the national state of emergency; Distress CF = Parental distress due to the curfew; Distress PD = Parental distress due to the pandemic; ER_CR = Emotional regulation (cognitive reappraisal); ER_ES = Emotional regulation (expressive suppression); SomAnx = Somatic anxiety; CogAnx = Cognitive anxiety; CovFear_P = Parents' Fear of COVID-19; SelfEf = Self efficacy; PanParent = Quality of parental pandemic practices.

Prevodi imena varijabli, odnosno šta koja skraćenica znači na srpskom su dati u tekstu tvrdnji (varijabla je imenovana na srpskom, a u zagradi je data skraćenica kojom je ta varijabla označena u tabeli).

Tabela preštampana iz: Radanović, A., Micić, I., Pavlović, S., & Krstić, K. (2021). Pandemic Parenting: Predictors of Quality of Parental Pandemic Practices during COVID-19 Lockdown in Serbia. Psihologija, 54(3), 323–345. https://doi.org/10.2298/psi200731040r . Preštampano na osnovu dozvole autora

N	Tvrdnja:	Odgovor
N1.	Roditeljski distress zbog pandemije (Distress PD) izaziva Strah dece od COVIDa (CovFear_C).	
N2.	Samoefikasnost (varijabla SelfEF) izaziva Roditeljski distress zbog pandemije (Distress PD)	

N3.	Roditeljski distress zbog vanrednog stanja (Distress NSE) izaziva Strah dece od COVIDa (CovFear_C).	
N4.	Distress PD i Distress NSE imaju više od 20% zajedničke varijanse.	
N5.	Koeficijenti korelacije koji su ovde prikazani su koeficijenti kontingencije.	
N6.	Ne možemo zaključiti iz podataka da su Distress CF i CovFear_C povezani u populaciji.	
N7.	Prema Tejlorovim (Taylor) preporukama, korelacija između Distress PD i CovFear_C se može smatrati umerenom.	
N8.	U populaciji nema korelacije između Distress PD i Distress CF.	
N9.	U populaciji nema korelacije između SelfEf i Distress CF	
N10.	Kvalitet roditeljskog pandemijskog postupanja (PanParent) raste sa uzrastom dece (Age_Children).	

Pogledajmo sada odgovore:

L1- netačno. Možemo videti da je korelacija između ove dve varijable 0,034 i da nije statistički značajna (korelacije koje su statistički značajne su označene u tabeli sa *, kao što i piše u samoj tabeli), što znači da prihvatamo nultu hipotezu koja kaže da je korelacija u populaciji nula i ovu korelaciju tako i tretiramo.

L2 – tačno. Možemo videti da iako je korelacija od -0,052 veoma slaba, ona je statistički značajna (uzorak istraživanja je veoma veliki!), tako da odbacujemo nultu hipotezu i zaključujemo da u populaciji postoji nenulta korelacija, te koeficijent korelacije tretiramo kao da je toliko koliko je. Tvrdnja govori o suprotnim smerovima promena dve varijable (jedna raste, druga se smanjuje), što je negativna korelacija, a u tabeli vidimo da je i koeficijent korelacije negativan.

L3 – netačno. Možemo videti da ova korelacija ima istu vrednost kao ona iz L2, ali tvrdnja kaže da je smer promene dve varijable isti (jedna se smanjuje, druga se smanjuje). Tvrdnja kaže da je korelacija pozitivna, a mi možemo iz vrednosti koeficijenta koji stoji u tabeli (-0,052) da vidimo da je korelacija između ove dve varijable negativna, što znači da je tvrdnja netačna.

L4 – tačno. Da, korelacija između ove dve varijable je pozitivna i statistički značajna (označena sa *). Ona je 0,062. Tvrdnja opisuje promene dve varijable u istom smeru, dakle pozitivnu korelaciju. Znači da je tvrdnja tačna.

L5 – netačno. Nema u tabeli nikakvih podataka o tome koji koeficijent korelacije je računat, ali možemo da vidimo da je Bezbednost u negativnoj korelaciji sa Fizičkim zdravljem, a to bi bilo nemoguće da je Bezbednost nominalna varijabla. Koeficijenti korelacije nominalnih varijabli nemaju smer (na način na koji korelacije varijabli na višim nivoima merenja imaju). Još jedna napomena – kada čitamo statističke tekstove i vidimo da su autori računali korelacije, ali da nisu napisali koje tačno koeficijente, u većini slučajeva ćemo biti u pravu ako pretpostavimo da su u pitanju Pirsonovi produkt-moment koeficijenti korelacije, ako nema elemenata u tekstu koji bi

PREPRINT

ukazivali da su ovi neodgovarajući za date podatke. Ova pretpostavka, međutim, neće biti tačna u 100% slučajeva, te treba biti oprezan u zaključivanju.

L6 – netačno. Ima jedna korelacija koja nije statistički značajna – ona između Poverenja i Uslova života.

L7 – netačno. Koen preporučuje da se korelacije oko 0,1 tretiraju kao niske. Ako koristimo interpretaciju njegove preporuke iz ove knjige koja kaže da se korelacije ispod 0,2 smatraju za niske, možemo videti da je ovo pogrešno, jer postoje tri koeficijenta korelacije veća od 0,2.

L8 – tačno. Korelacija između Poverenja i Uslova života nije statistički značajna. Prema tome, prihvatamo nultu hipotezu i smatramo da ove dve varijable nisu povezane.

L9 – netačno. Možemo da vidimo da su sve prikazane korelacije sa Društvenim odnosima više od onih sa Uslovima života.

L10 – tačno. Tvrdnja implicira da je korelacija negativna (niže vrednosti na jednoj varijabli odgovaraju višim vrednostima na drugoj varijabli). Možemo videti u tabeli da je koeficijent korelacije -0,086, statistički značajan i negativan. Tvrdnja je, prema tome, tačna.

M1 – tačno. Tejlorova preporuka je da se korelacije između 0,36 i 0,67 smatraju umerenim, a korelacija između R i I je 0,45, dakle unutar opsega umerenih korelacija.

M2 – netačno. Duž dijagonale bi bile korelacije varijable sa samom sobom. Kako je svaka varijabla identična samoj sebi, to bi bila korelacija između dva identična skupa vrednosti, što bi bilo 1, a ne 0. Ona, međutim, nije prikazana u tabeli zato što računanje korelacije između skupa podataka i njene identične kopije je naučno bezvredno.

M3 – tačno. Tekst ispod tabele kaže da su sve korelacije više od 0,07 statistički značajne i možemo videti da nema korelacija u tabeli koje su niže od toga. Treba da zapamtimo da je moguće prosto odrediti prag veličine korelacije na ovaj način dokle god je veličina uzorka ista za sve posmatrane koeficijente korelacije, kao što je slučaj na primer onda kada su svi računati na istom uzorku. Ovo je moguće zbog toga što je standardna greška koeficijenta korelacije, onda kada testiramo nultu hipotezu da je korelacija u populaciji 0, samo funkcija veličine uzorka (1 podeljeno sa kvadratnim korenom iz broja ispitanika u uzorku). Kako je ovo upravo to što je urađeno u ovoj tabeli, određivanje za koje koeficijente možemo da odbacimo nultu hipotezu, minimalna veličina koeficijenta korelacije za ovo je ista za sve koeficijente u tabeli.

M4 – tačno. Koen preporučuje da koeficijente oko 0,1 smatramo niskim. Mi smo ovo interpretirali kao koeficijente ispod 0,2. Ovaj konkretna koeficijent je 0,14, dakle nizak, što znači da je tvrdnja tačna.

M5 – netačno. Korelacija između A i S, kao što možemo videti u tabeli, je 0,48 i to je statistički značajna korelacija, ali i prilično supstantivna.

M6 – nepoznato. U tabeli su prikazane korelacije. Korelacija nije kauzacija i iz korelacija ne možemo da izvedemo nikakve zaključke o tome šta je uzrok čemu. E može da bude uzrok C, ali i ne mora. Samo na osnovu podataka koji su dati u tabeli to ne možemo da odredimo.

PREPRINT

M7 – tačno. To da varijable zajedno rastu znači da je korelacija pozitivna i iz tabele možemo videti da je korelacija 0,64 i da je zaista pozitivna.

M8 – tačno. To da se obe varijable smanjuju ukazuje na isti smer zajedničke promene, dakle na pozitivnu korelaciju. Iz tabele možemo videti da je korelacija 0,30, što je pozitivna korelacija.

M9 – tačno. Da, kolone i redovi ove tabele sadrže iste varijable. I kako su ovo Pirsonovi koeficijenti korelacije, koeficijenti koji su simetrični, korelacija varijable A sa varijablom B je ista kao korelacija varijable B sa varijablom A. U ovom slučaju, na primer, korelacija R i I je ista kao korelacija I i R. Treba primetiti da da su ovo, na primer, eta koeficijenti ili neka druga asimetrična mera povezanosti, ovo ne bi morao da bude slučaj.

M10 – netačno. Ne, to su Pirsonovi produkt-moment koeficijenti korelacije. Iako je Spirmanov koeficijent korelacije rangova zapravo varijanta Pirsonovog koeficijenta korelacije, u literaturi je uobičajeno da se, ako su računati Spirmanovi koeficijenti, to tako i napiše, a to ovde nije slučaj.

Tabela N – opšti komentari. Ovo je malo složeniji prikaz statističkih podataka od onih koje smo do sada koristili, tako da ćemo dati kratko objašnjenje pre nego što pređemo na odgovore na tvrdnje. Možemo videti da ova tabela integriše podatke o aritmetičkim sredinama i standardnim devijacijama varijabli sa korelacijama između varijabli. Prve dve kolone sa brojevima su podaci o aritmetičkim sredinama i standardnim devijacijama varijabli čija imena su data sa leve strane. Svaki red je jedna varijabla, a pri tom se neke varijable odnose na neka svojstva dece (prve dve varijable), dok se ostale odnose na svojstva njihovih roditelja. Možemo da primetimo da je svaka od ovih varijabli označena brojem. Taj broj je ponovo iskorišćen kod imenovanja kolona u gornjem delu tabele, da označi te iste varijable za svrhe prikazivanja korelacija između varijabli. Na primer, možemo da vidimo da je korelacija između varijable SelfEf i varijable Age_Children -0,14. Ovo možemo da vidimo zato što je kolona u kojoj se nalazi ova vrednost označena brojem 2, što se odnosi na varijablu koja je takođe označena brojem 2 u redovima gde se nalaze imena varijabli. Na sličan način, kolona imenovana brojem 4 se odnosi na varijablu Distress CF, kolona 5 na varijablu Distress PD i tako dalje. Konačno, zvezdica (*) stavljena pored koeficijenta korelacije pokazuje da je vrednost statističke značajnost ovog koeficijenta manja od 0,05 tj. da je on statistički značajniji od 0,05. Dve zvezdice (**) označavaju da je vrednost statističke značajnosti koeficijente manja od 0,01 (tj. da je statistički značajniji od 0,01). Ovo sve možemo pročitati iz napomene ispod tabele. Konačno, čitaoci mogu da primete, da uz tabelu postoji vrlo detaljna napomena (na engleskom) u kojoj su navedena puna imena varijabli, dok su u tabeli samo skrećenice.

N1 – nepoznato. Možemo da vidimo da je korelacija između ovih varijabli pozitivna i statistički značajna (0,37), tako da je kauzacija svakako moguća. Međutim, korelacija nije kauzacija i samo na osnovu korelacije ne možemo da znamo da li je jedna varijabla uzrok druge ili ne. N2 – nepoznato. I ovde možemo da vidimo da postoji negativna korelacija između ovih varijabli (-0,15), korelacija koja se statistički značajna, što znači da je kauzacija moguća. Međutim, korelacija nije kauzacija i to da li je jedna varijabla uzrok druge nije poznato (samo na osnovu podataka iz tabele).

N3 – netačno. Možemo da vidimo da ove dve varijable nisu u korelaciji, tj. da je njihova korelacija skoro 0 i da nije statistički značajna. To isključuje mogućnost da jedna izaziva drugu, jer bi

PREPRINT

uzročno-posledični odnos implicirao postojanje nekog efekta jedne varijable na drugu (ako jedna varijabla nema nikakav efekat na drugu, ne može ni da joj bude uzrok), a to znači korelaciju. Naravno, i dalje je moguće da su obe ove varijable deo nekog kompleksnijeg kauzalnog sistema, ili da bi jedna mogla da bude uzrok druge ako bi supresivni efekti neke treće varijable bili uklonjeni, ali u situaciji na koje se ovi podaci odnose, to nije slučaj.

N4 – netačno. Setimo se da se proporcija zajedničke varijanse dobija tako što se računa koeficijent determinacije tj. kvadrat koeficijenta korelacije. Iz tabele možemo videti da je korelacija između ove dve varijable 0,21, a kvadrat broja 0,21 (tj. 0,21 X 0,21) je oko 0,04, što je 4%, a to je daleko ispod 20% koji se pominju u tvrdnji.

N5 – netačno. Koeficijent kontingencije ne može da ima negativne vrednosti, a ovi koeficijenti imaju negativne vrednosti. Pored toga, aritmetičke sredine i standardne devijacije su računate, što znači da ovo nisu nominalne varijable.

N6 – tačno. Korelacija između ove dve varijable, koja je -0,03) nije statistički značajna (skoro je nula i nema *).

N7 – tačno. Prema Tejloru, korelacije između 0,36 i 0,67 su umerene, a ovo je korelacija od 0,37, dakle umerena.

N8 – tačno. Možemo videti da njihova korelacija od 0,07 nije statistički značajna. Dakle, prihvatamo nultu hipotezu koja kaže da nema korelacije u populaciji.

N9 – netačno. Iz tabele možemo pročitati da je korelacija 0,11. Ona je svakako niska, ali je označena sa *, što znači da je statistički značajna. Prema tome, odbacujemo nultu hipotezu i zaključujemo da postoji nenulta korelacija u populaciji između ove dve varijable.

N10 – netačno. Možemo pročitati iz tabele da je korelacija između ove dve varijable negativna (-0,27). Dakle, viši nivo roditeljskih panedmijskih postupanja (PanParent) odgovara manjoj starosti dece (Age_Children), a ne višoj.

PREPRINT

Poglavlje 7. Statistički testovi za poređenje dva uzorka

Apstrakt. Ovo poglavlje predstavlja neke od najpopularnijih statističkih testova za poređenje dve grupe. Počinje tako što objašnjava razliku između zavisnih i nezavisnih uzoraka i značaj te razlike za razumevanje načina na koji statistički testovi funkcionišu. U narednom delu poglavlja predstavljen je i objašenjen t test, zajedno sa neparametrijskim alternativama za t test, kako za nezavisne uzorke - Men-Vitnijev U test i Vilkoksonov test sume rangova, tako i za zavisne uzorke – test znaka i Vilkoksonov test rangova sa predznakom. Iako nije u pitanju test za poređenje striktno dva uzorka, Levenov test je predstavljen sledeći jer je u pitanju test čiji se rezultati u literaturi često predstavljaju zajedno sa drugim testovima iz ovog poglavlja kada postoji potreba da se ispita homogenost varijansi poređenih grupa, što je za neke od testova iz ovog poglavlja ili preduslov ili faktor koji menja način na koji se računaju. Poslednji deo poglavlja posvećen je testovima za poređenje distribucija – Kolmogorov-Smirnov testu, Pirsonovom Hi kvadrat testu i Vald-Volfovic testu. Fišerov egzaktni test i binomni test su takođe ukratko predstavljeni. Pored testova, prikazane su i mere veličine efekta koje se obično prikazuju uz date testove. Pored koeficijenata korelacije koji su već razmatrani u prethodnim poglavljima, ovde su opisani Koenovo d, Somersovo D, odnos varijansi i odnos varovatnoća (odds ratio).

Ključne reči: statistički testovi, zavisni i nezavisni uzorci, mere veličine efekta, t test, K-S test.

U poglavlju o statistici zaključivanja razmatrali smo postupke i pojmove koji su važni za izvođenje zaključaka o vrednostima populacije (parametrima) na osnovu vrednosti dobijenih na uzorku (statistika). Ti opšti postupci predstavljaju glavni okvir za izvođenje takvih zaključaka. Međutim, kada je naš cilj da uporedimo specifična svojstva dve različite populacije, da izvedemo zaključak o tome da li dva uzorka potiču iz populacija koje su jednake u pogledu nekog tačno određenog statističkog svojstva ili da li su određeni statistički parametri dve populacije u tačno određenoj vrsti odnosa kakvu pretpostavljamo, potrebno je da primenimo odgovarajuće statističke testove, napravljene upravo za sprovođenje takvih vrsta poređenja. Iako se koncepti statističke značajnosti i/ili intervala poverenja zasnovanih na reuzorkovanju i drugi odgovarajući postupci koriste u svim ovim testovima, sami testovi uključuju proračune koji su potrebni da bi se ovi opšti statistici izračunali.

PREPRINT

Do dana današnjeg, različiti statističari su predlagali i i dalje predlažu različite statističke testove za testiranje različitih statističkih hipoteza. Postojeći statistički testovi mogu da se koriste za poređenje dva uzorka, za poređenje više uzoraka od jednom, ali i za poređenje jednog jedinog empirijski dobijenog uzorka sa određenim teorijskim očekivanjem izraženim bilo kao određeni oblik distribucije ili određena vrednost statistika koja se unosi u postupak testiranja. Broj predloženih statističkih testova koji se mogu naći u literaturi je u stotinama, a verovatno i u hiljadama. Međutim, kako je cilj ove knjige da čitaoce upozna sa osnovama istraživačke statistike, u ovom poglavlju ćemo se fokusirati na najčešće upotrebljavane i najpoznatnije testove za poređenje dva uzorka ili za poređenje jednog empirijski dobijenog uzorka (tj. uzorka koji se sastoji od pravih entiteta na kojima su prava merenja izvođenja, čiji su rezultati onda uneseni u matricu podataka koju analiziramo kao podaci o merenim varijablama) sa različitim teorijskim očekivanjima, bilo da su u pitanju teorijske distribucije ili određene očekivane vrednosti parametra.

7.1. Zavisni i nezavisni uzorci

Jedna važna tema koju treba prodiskutovati na početku poglavlja o statističkim testovima je razlika između zavisnih i nezavisnih uzoraka. Ovo je važno zbog toga što je većina testova namenjena ili primeni na zavisnim uzorcima ili na nezavisnim uzorcima uz svega nekoliko testovima koji imaju varijante i za jednu i za drugu vrstu uzoraka. A čak i kada testovi imaju varijante i za zavisne i za nezavisne uzorke, te varijante uključuju donekle različite proračune, koji se samo nazivaju istim testom. Postoje statistički postupci koji se međusobno razlikuju manje nego varijante istog testa za zavisne i nezavisne uzorke, a da se smatraju različitim testovima! To da li su uzorci zavisni ili nezavisni određuje i kako će matrica sa podacima biti pripremljena tj. kako će podaci biti organizovani u matrici.

Zavisni uzorci su uzorci kod kojih su entiteti iz različitih uzoraka upareni tako da za svaki entitet iz jednog uzorka postoji njemu odgovarajući entitet u drugom uzorku (ili drugim uzorcima, za slučaj da ima više zavisnih uzoraka!). To znači da za svaki pojedinačni entitet u jednom uzorku znamo koji mu tačno entitet iz drugog uzorka odgovara. **Ova korespondencija između entiteta se uspostavlja bilo tako što se mere uzimaju na entitetima koji već odgovaraju jedan drugom po nekom svojstvu i onda su po tom svojstvu upareni ili preko**

višestrukih merenja vrednosti varijable istih entiteta (i u tom slučaju entitet odgovara samom sebi). Kada se testiraju hipoteze o odnosima između dva uzorka, **računanja se tipično rade na način koji pravi poređenja između vrednosti uparenih/korespondirajućih entiteta**. Situacije u istraživanja kada se koriste zavisni uzorci uključuju na primer:

- **Istraživanja koja zahtevaju dva ili više merenja istih entiteta**, kao što su istraživanja gde se **mere rade pre i posle određenog postupka,** kao i **longitudinalne studije** gde se mere istih entiteta uzimaju u više navrata tokom trajanja studije. U ovim slučajevima, vrednosti entiteta iz prvog merenja odgovara vrednost istog entiteta iz drugog merenja (i svih drugi merenja, ako ima više od dva). Ako, na primer, testiramo efekte određenog postupka na grupu ljudi, vrednost svake osobe iz prvog merenja odgovara vrednosti iste osobe iz drugog merenja (i svih narednih merenja).
- **Istraživanja u kojima se uporedive varijable mere na istim entitetima**. U ovom slučaju **vrednost entiteta na jednoj varijabli odgovara vrednosti istog entiteta na drugoj varijabli**. Primer ovoga su situacije gde se mere stavovi prema određenim objektima i onda uparujemo mere stavova iste osobe prema različitim objektima prema kojima su ispitivani stavovi. Ili situacije u kojima poredimo postignuća osobe na testovima iz različitih predmeta (ovo pod pretpostavkom da su ti testovi pravljeni tako da budu smisleno uporedivi na neki način, recimo nekom metodom poređenja pozicija na distribuciji!).
- **Istraživanja u kojima se ispituju različiti, ali na neki način upareni entiteti.** Ovo uključuje situacije kada istu varijablu merimo na ljubavnim parovima ili na članovima porodice ili na roditeljima i njihovoj deci, nastavnicima i njihovim učenicima isl. U takvim situacijama vrednost varijable kod jednog partnera ljubavnog para se uparuje sa vrednostima varijable kod njegovog/njenog partnera. Ili se vrednosti roditelja uparuju sa vrednostima njihove dece na istoj varijabli ili se uparuju vrednosti članova iste porodice koji učestvuju u israživanju ili nastavnika i njihovih učenika.

U matrici podataka, zavisne/uparene vrednosti se tipično stavljaju u isti red, jednako kao da su dobijene na istom entitetu, čak i u situacijama kada se radi o vrednostima koje su dobijene na različitim uparenim entitetima.

PREPRINT

Tabela 7.1. Primer zavisnih uzoraka. Grupa učesnika u istraživanju (imena su data u prvoj koloni s leve strane) je ocenjivala kvalitet usluga 5 avio kompanija (ovde označene kao avio kompanije A, B, C, D i E). Kako su ovo sve ocene na skali od 1-5 koje su pravljene sa namerom da se porede, možemo tretirati ocene različitih avio kompanija kao zavisne uzorke i zaista i tako napraviti poređenja. Na primer, možemo upariti ocene koje su učesnici dali avio kompaniji A sa ocenama koje su ovi isti ljudi dali avio kompaniji B. U ovom poređenju, Leposavina ocena avio kompanije A bi bila uparena sa njenom ocenom avio kompanije B. Anitina ocena avio kompanije A bi bila uparena sa Anitinom ocenom aviokompanije B i tako bi to bilo urađeno za sve učesnike u istraživanju. Na isti način bi uparili ocene svih pet aviokompanija u ovom istraživanju i uporedili bi ih. Dakle, ovde bi imali 5 zavisnih uzoraka, prvi bi činile ocene koje je ova grupa ljudi dala avio kompaniji A, druge ocene koje su dali avio kompaniji B i tako sve do avio kompanije E. Prikazani podaci su izmišljeni.

Kako učesnici u istraživanju ocenjuju avio kompanije, sirovi podaci Ocene sun a skali od 1 do 5, pri čemu je 1 najgora ocena, a 5 najbolja.					
Ime učesnika u istraživanju	**Avio kompanija A**	*Avio kompanija B*	Avio kompanija C	Avio kompanija D	Avio kompanija E
Leposava	**3**	*5*	4	3	2
Anita	**5**	*5*	4	5	5
Vladislava	**4**	*2*	5	5	2
Maida	**2**	*2*	1	2	2
Marko	**5**	*5*	5	5	5
Radoslav	**1**	*5*	2	5	4
Filip	**5**	*4*	4	5	4
Vladimir	**3**	*1*	3	5	4
Jovan	**1**	*5*	1	5	1
Goran	**5**	*5*	1	5	5
Ilona	**5**	*5*	3	5	5
Petar	**4**	*4*	2	4	4

Nezavisni uzorci su uzorci kod kojih entiteti iz jednog uzorka nemaju odgovarajuće entitete iz drugog uzorka tj. entiteti iz različitih uzoraka su nezavisni. Iako ovakvi uzorci mogu biti, i obično jesu, uporedivi kao celina, **ne postoji nikakva utvrđena korespondencija između pojedinačnih entiteta iz jednog uzorka i pojedinačnih entiteta iz drugog uzorka.**

PREPRINT

U matrici podataka, dva (ili više) nezavisna uzorka se tipično predstavljaju kao posebni redovi podataka (posebni unosi), a u matricu se uključuje i varijabla koje sadrži podatke o tome koji entitet (tj. koji red podataka) pripada kom uzorku.

Tabela 7.2. Primer nezavisnih uzoraka. Učesnici u istraživanju su podeljeni u dva uzorak po polu. Ovde treba primetiti varijablu pol, na kojoj su učesnici u istraživanju koji su rekli za sebe da su ženskog pola označeni brojem 1, a oni koji su saopštili da su muškog pola označeni brojem 2. Boldirali smo jednu grupu, dok su za drugu korišćena normalna slova. Možemo videti da ovde imamo varijablu (ta varijabla je Pol u ovom slučaju) koja sadrži informaciju o tome kojoj grupi pripada dati učesnik u istraživanju, ali nema nikakve korespondencije između entiteta, nema načina da se ustanovi koji entitet/učesnik u istraživanju iz jedne grupe odgovara kom entitetu/učesniku u istraživanju iz druge grupe, jer takve uparenosti nema. Nezavisni uzorci m ogu biti i različitih veličina (za razliku od zavisnih) – možemo videti da u ovom primeru ima 5 osoba ženskog pola naspram 7 osoba ženskog pola. Prikazani podaci su izmišljeni.

Ime učesnika u istraživanju	Pol	Neuroticizam N	Ekstraverzija E	Saradljivost A	Otvorenost za iskustvo O	Savesnost C
Leposava	**1**	**45**	**55**	**55**	**65**	**70**
Anita	**1**	**22**	**67**	**50**	**72**	**62**
Vladislava	**1**	**37**	**25**	**45**	**65**	**45**
Maida	**1**	**55**	**32**	**65**	**42**	**35**
Marko	2	25	25	32	28	60
Radoslav	2	52	12	42	35	28
Filip	2	50	70	48	47	32
Vladimir	2	38	65	51	59	65
Jovan	2	40	25	65	61	65
Goran	2	25	30	22	45	50
Ilona	**1**	**32**	**45**	**45**	**68**	**65**
Petar	2	35	55	65	42	58

Iako od ovoga mogu postojati odstupanja jer se polje statistike stalno razvija, **najčešće statistički testovi za zavisne uzorke zahtevaju manje razlike između uzoraka tj. manje veličine efekta za odbacivanje nulte hipoteze** ili su potrebni manji uzorci da bi ista veličina efekta dostigla kritični prag statističke znajačnosti nego što je to slučaj kod nezavisnih uzoraka. Drugim rečima, **testovi za zavisne uzorke obično imaju veću snagu** nego testovi za nezavisne uzorke.

PREPRINT

7.2. Poređenje dve aritmetičke sredine – t test

T test je parametrijski test namenjen testiranju hipoteze o tome da li dva uzorka potiču iz populacija sa istim vrednostima aritmetičke sredine. Ovo se radi tako što se testira **nulta hipoteza da je razlika između aritmetičkih sredina dve populacije nula**. Testiranje ove hipoteze se zasniva na računanju standardne greške razlike između aritmetičkih sredina i onda deljenju dobijene razlike između aritmetičkih sredina uzoraka tom standardnom greškom razlike između aritmetičkih sredina. **Veličina standardne greške razlike između aritmetičkih sredina zavisi od** ona ista dva faktora koje smo već pominjali kao faktore koji određuju veličine standardnih grešaka – **varijabilnosti varijabli na kojoj su aritmetičke sredine na uzorku na kom su računate** (tj. od njihovih standardnih devijacija/varijansi) i od **veličine uzorka**. **Što je uzorak veći i što je manja varijabilnost vrednosti entiteta u oba uzorka, biće manja i standardna greška razlike između aritmetičkih sredina**. Nakon što se izračuna ova standardna greška, **računa se razlika između aritmetičkih sredina i ta razlika se podeli standardnom greškom razlike između aritmetičkih sredina.** Statistik koji se dobije ovim deljenjem zove se **t statistik. T statistik** (t statistik se inače obeležava malim „t“, ovde je veliko t jer je početak rečenice) **se onda koristi za procenu verovatnoće da se dobije razlika između aritmetičkih sredina dva uzorak kolika je dobijena ili veća kada je razlika između aritmetičkih sredina populacija 0 i ta verovatnoća predstavlja statističku značajnost t testa**. **Ako je rezultat t testa statistički značajan** tj. ako je p vrednost manja od prihvaćenog praga statističke značajnosti (a to je, podsetimo se, najčešće 0,05**), nultu hipotezu možemo da odbacimo i da zaključimo da uzorci ne dolaze iz populacija sa istim vrednostima aritmetičke sredine na ispitivanoj varijabli** (tj. da zaključimo da aritmetičke sredine populacija iz kojih su uzorci nisu jednake). S druge strane, ako je vrednost statističke značajnosti iznad prihvaćenog praga za odbacivanje nulte hipoteze, nultu hipotezu prihavatamo i zaključujemo da nema dovoljno dokaza da aritmetičke sredine populacija iz kojih su uzorci nisu jednake (tj. tretiraćemo ih na dalje kao da jesu jednake).

T test je parametrijski test što znači, u ovom slučaju, da se zasniva na pretpostavci da je **distribucija uzorkovanja razlika između aritmetičkih sredina normalna**. Ako računamo statističku značajnost t testa oslanjajući se na centralnu graničnu teoremu, nemamo načina da testiramo ovu pretpostavku. Međutim, formule za ove proračune se već oslanjaju na očekivanje da

PREPRINT

je distribucija uzorkovanja normalna (normalna=da ima oblik normalne distribucije). **Kako se normalnost distribucija uzrokovanja razlika između aritmetičkih sredina ne može testirati**, jer imamo samo jedan uzorak, a ne celu distribuciju uzorkovanja**, ono što istraživači obično rade je da testiraju normalnost distribucije podataka na dva uzorka** (tj. provere da li su distribucije dva uzorka na datoj varijabli dovoljno slične obliku normalne distribucije) **i onda pretpostave da ako su distribucije uzoraka normalne, da će i distribucija uzorkovanja njihovih razlika biti takođe normalna** (podsetimo se ovde da je distribucija uzorkovanja distribucija koji bi činili statistici izračunati iz velikog broja uzoraka uzetih iz iste populacije ili iz iste dve populacije u slučaju distribucije uzorkovanja razlike između aritmetičkih sredina o kojoj govorimo ovde). Međutim, Field (2009), na primer, ispravno konstatuje da je **moguće da imamo dva uzorka na kojima podaci nisu normalno distribuirani, ali da distribucija njihovih razlika bude normalna**. U svakom slučaju, uobičajena praksa istraživača je da t test koriste pod uslovom da oblik distribucije podataka sa dva uzorka ne odstupa previše od oblika toerijske normalne distribucije.

Još jedna važna pretpostavka t testa je **pretpostavka da su varijanse/standardne devijacije dva uzorka jednake i da svi podaci potiču od posebnih/nezavisnih merenja** (to ne znači da uzorci treba da budu nezavisni – na zavisnim uzorcima se takođe mogu raditi nezavisna merenja!). Kada je u pitanju prva pretpsotavka, ova o jednakosti varijansi grupa, treba primetiti da to jeste osnovno očekivanje kada je u pitanju t test, ali i **da postoje korekcije za formulu za računanje rezultata t testa za situacije kada varijanse dva uzorka nisu jednake.** Tu je samo važno da se varijanse uporede pre sprovođenja t testa i da se onda primeni odgovarajuća korekcija za računanje rezultata u situaciji kada varijanse nisu jednake. Što se tiče drugog uslova, to da mere u uzorku treba da budu **nezavisne znači da unutar svakog uzorka, svaki podatak treba da potiče od posebnog entiteta tj. ne treba da bude entiteta čije su vrednosti unesene više puta u uzorak**. Ako su naši entiteti ljudi to znači da svaka osoba treba da bude predstavljena u uzorku samo jednom. **Ovo ne treba mešati sa situacijama kada imamo zavisne uzorke koji se sastoje od različitih skupova mera napravljenih na istoj grupi entiteta**. Ovaj uslov samo znači da unutar jednog uzorka tj. jedne grupe, na jednoj varijabli, ne treba da bude višestrukih merenja istog entiteta niti višestrukih kopija istog merenja.

PREPRINT

Drugi način da **se testira nulta hipoteza poput one koja se testira t testom** je kroz **reuzorkovanje,** što tipično znači **butstreping.** U ovom postupku, pravi se distribucija uzorkovanja tako što se uzme unapred određeni, veliki broj uzoraka kroz uzorkovanje sa vraćanjem iz izvornog uzorka, a onda se definiše interval poverenja koji uključuje željeni procenat distribucije uzorkovanja i centriran je oko aritmetičke sredine distribucije uzorkovanja (što znači uglavnom da su granice intervala poverenja jednako udaljene od aritmetičke sredine distribucije uzorkovanja). Nakon toga, **istraživač proverava da li je 0** (ili koja je već vrednost određena nultom hipotezom) **unutar intervala poverenja ili nije. Ako jeste unutar intervala poverenja, tada prihvatamo nultu hipotezu, a ako nije unutar tog intervala, onda nultu hipotezu odbacujemo**. Treba primetiti da se ovakav pristup testiranju razlikuje u pogledu zahteva od t testa u onom delu koji kaže da merenja treba da budu nezavisna, jer se u okviru postupka uzorkovanja sa vraćanjem nužno prave višestruke kopije istog entiteta tokom postupka reuzorkovanja. Uprkos tome, ovaj postupak je, u trenutku pisanja ove knjige, sve češće u upotrebi.

T test ima varijante i za zavisne i za nezavisne uzorke, što znači da se može koristiti na obe vrste uzoraka. Međutim, formule za računanje statistika ovog testa se razlikuju u ova dva slučaja. **T test se može računati i za samo jedan uzorak, onda kada hoćemo da testiramo nultu hipotezu da uzorak potiče iz populacije sa tačno određenom vrednošću aritmetičke sredine.**

Kod t testa, varijabla koja deli uzorak u dve grupe, koja sadrži podatke o tome koji entitet pripada kojoj od dve grupe naziva se nezavisna varijabla, dok je intervalna ili racio varijabla na kojoj su izračunate aritmetičke sredine koje se porede zavisna varijabla. Isti plan podele varijabli na nezavisne i zavisne važi i za sve ostale testove koji će biti predstavljeni u ovom poglavlju u kojima se vrši poređenje dve grupe.

Rezultati t testa su osnova za odlučivanje da li nultu hipotezu da prihvatimo ili da odbacimo, međutim **odbacivanje nulte hipoteze i usvajanje zaključka da aritmetičke sredine dve populacije iz kojih su uzorci nisu jednake, ne znači nužno i da su te razlike dovoljno velike da se praktično vredi njima uopšte baviti**. Iz ovog razloga**, uobičajeno je da se izračuna i prikaže i neka mera veličine efekta zajedno sa rezultatima t testa. Mere veličine efekta koje se tipično koriste zajedno sa t testom su point-biserijski koeficijent korelacije ($r_{p.bis}$) i Koenovo d.**

Pointbiserijski koeficijent korelacije ($r_{p.bis}$) je već razmatran u poglavlju o korelacijama. On se računa kao **Pirsonov koeficijent korelacije između jedne (prave) binarne varijable i jedne intervalne varijable**. U slučaju t testa tj. poređenja dva uzorka, **binarna varijabla sadrži informaciju o tome koji entitet pripada kom uzorku, a intervalna varijabla je varijabla na kojoj su računate aritmetičke sredine koje se porede**. Point-biserijski koeficijent korelacije se **može izračunati i direktno iz t statistika preko formule**. Kao što je pomenuto već u prethodnom poglavlju, jedna vrlo korisna preporuka za interpretaciju veličine korelacija je ona koju je predložio Koen (Cohen, 1988), a koja kaže da koeficijente korelacije od 0,1 treba smatrati za slabe/male, koeficijente korelacije od 0,3 treba smatrati za umerene, a korelacije od 0,5 treba smatrati za visoke.

Koenovo d je standardizovana razlika između aritmetičkih sredina. Računa se tako što se **razlika između aritmetičkih sredina dva uzorka podeli njihovom zajedničkom/združenom standardnom devijacijom**. Združena standardna devijacija data dva uzorka je prosto aritmetička sredina njihove dve standardne devijacije (od svakog uzorka po jedna) ponderisana veličinama dve grupe. Ako se dva uzorka za koja ovo računamo sastoje od podjednakog broja entiteta, onda je njihova združena standardna devijacija obična aritmetička sredina njihovih standardnih devijacija tj. (SD1 + SD2)/2. Ako su standardne devijacije dva uzorka jednake, onda je zajednička standardna devijacija prosto standardna devijacija uzoraka (npr. ako je standardna devijacija i jednog i drugog uzorka 30, onda je i združena standardna devijacija 30). Ako standardne devijacije dva uzorka nisu jednake, a ni uzorci nisu jednake veličine, onda se svaka standardna devijacija deli brojem entiteta u uzorku iz kog je izračunata, tako dobijeni proizvodi (standardna devijacija x broj entiteta u uzorku iz kog je izračunata) se saberu, a suma se podeli sa brojem entiteta u oba uzorka zajedno. Na ovaj način, **Koenovo D nam pokazuje koliko je velika razlika između aritmetičkih sredina u jedinicama standardnih devijacija,** što je vrlo slično z skorovima. **Koen predlaže da se d vrednosti od 0,2 smatraju niskim, d vrednosti od 0,5 da se smatraju srednjim, a d vrednosti od 0,8 i preko da se smatraju velikim. Vrednosti Koenovog d koje su manje od 0,2 u principu predstavljaju vrlo male razlike koje su u većini slučajeva mogu smatrati zanemarljivim** i tako ih i treba tretirati osim ako nema nekog posebno dobrog teorijskog opravdanja da se uradi suprotno, čak i ako su rezultati t testa u takvim slučajevima statistički značajni.

PREPRINT

7.3. Poređenje centralnih tendencija ordinalnih podataka na nezavisnim uzorcima – Men-Vitnijev U test i Vilkoksonov test sume rangova

Jedna popularna neparametrijska alternativa t testu za nezavisne uzorke je Men-Vitnijev U test (eng. Mann Whitney's U test). **U test poredi centralne tendencije dva nezavisna uzorka**. Ono što **ovaj test efektivno poredi su prosečni rangovi dva uzorka na zajedničkoj rang listi**. **Nulta hipoteza je da dva uzorka potiču iz iste populacije** (ili iz populacija sa istim svojstvima) i da, zbog toga, kada se podaci iz oba uzorka zajedno rangiraju, **proseci rangova jedne i druge grupe će biti jednaki.** Neki istraživači ovo interpretiraju kao da U test poredi medijane dve grupe. Međutim, činjenica je da je ono što se poredi proseci rangova, što ne mora uvek da bude jednako medijanama. Može se smatrati da ovaj test poredi medijane eventualno pod uslovom da su distribucije dve grupe identične, što ne mora da bude slučaj. U literaturi se takođe može naći i to da on **testira hipotezu da postoji 50% šanse da je nasumice izvučena vrednost iz jednog uzorka veća od nasumice izvučene vrednosti iz drugog uzorka.**

U test se sprovodi tako što se prvo napravi zajednička rang lista za oba uzorka. To se radi tako što se dva uzorka spoje u jedan, a onda se rang 1 dodeljuje entitetu sa najmanjom vrednošću na posmatranoj varijabli, rang 2 onoj sa drugom najmanjom itd. Kada dva entiteta imaju jednake vrednosti, onda se i jednom i drugom takvom entitetu dodeljuje aritmetička sredina dva susedna ranga, ona koja bi zauzeli po redu. Na primer, ako dva entiteta imaju vrednost 5 na ispitivanoj varijabli i to je 10. (deseta) najniža vrednost u uzorku, što znači da bi ta dva entiteta zauzela 10. i 11. mesto na rang listi, tada se i jednom i drugom entitetu dodeljuje rang 10,5, zato što je aritmetička sredina brojeva 10 i 11 jednaka 10,5. Takve situacije**, kada dva ili više entiteta imaju istu vrednost prilikom rangiranja, pa zato dobiju isti rang na način koji je upravo opisan nazivaju se situacijama vezanih rangova**. Nakon što se napred opisano rangiranje uradi, primenjuje se formula za računanje U koeficijenta koja je zasnovana na broju entiteta iz dva uzorka i sumi rangova u jednom od uzoraka. U se računa za oba uzorka i manji od tako dobijena dva rezultata smatra se za U statistik. Taj statistik se onda koristi za računanje statističke značajnosti na osnovu koje odlučujemo da li da nultu hipotezu prihvatimo ili da je odbacimo.

Vilkoksonov test sume rangova (eng. Wilcoxon's sum-rank test) je **sličan U testu i ova dva testa zapravo daju isti krajnji rezultat**. Testni statistik ovog testa označava se **sa W** i on je **suma rangova uzorka čija je suma rangova manja**. Prosek sume rangova dve grupe se onda oduzima od ovog statistika i rezultat se deli standardnom greškom W statistika. Obe ove kasnije vrednosti

PREPRINT

se računaju na osnovu broja entiteta u dve grupe. **Rezultat koji se dobije je po svojoj prirodi z skor** (sirova vrednost (W) minus aritmetička sredina (aritmetička sredina sume rangova) i to podeljeno standardnom devijacijom (standardna greška kojom se deli je standardna devijacija distribucije uzorkovanja!)) i **pokazuje poziciju našeg skora na distribuciji uzorkovanja koju bi dobili ako bi populacije iz kojih su naše dve grupe bile jednake**. Ovo onda možemo da **uporedimo sa z vrednostima za koje znamo da predstavljaju prihvaćene kritične vrednosti za odlučivanja o prihvatanju/odbacivanju nulte hipoteze da bi ustanovili da li je W statistika statistički značajan** (podsetimo se – ako usvojimo 0,05 kao kritičnu vrednost za odbacivanje/prihvatanje nulte hipoteze, z skor koji se nalazi na toj kritičnoj granici je 1,96).

Iako ovi proračuni mogu izgledati dosta složeno, ideja u osnovi oba testa je to da ako dve grupe dolaze iz jednakih populacija, kada napravimo jedinstvenu rang listu koja sadrži entitete iz obe grupe, biće u svakoj od dve poređene grupe otprilike podjednak broj entiteta sa visokim i sa niskim rangovima (na toj zajedničkoj rang listi). Ako visoki skorovi/visoke vrednosti na varijabli teže da budu koncentrisane u jednu grupu, a niske u drugu, tada dve grupe nisu jednake. Na primer, zamislimo dva tima koji trče trku i nas kako beležimo redosled kojim je svaki trkač prošao kroz cilj tj. njihove rangove u trci. Ako su dva tima otprilike jednake brzine, onda će biti ljudi koji su završili među prvima i ljudi koji su završili među poslednjima u obe grupe (a i onih koji su završili među srednjima!). Ako primetimo da trkači iz jednog tima generalno pokazuju tendenciju da stignu na cilj pre trkača iz druge grupe tj. da najviše zauzimaju prvi deo rang liste, dok drugi tim uglavnom zauzima zadnji deo rang liste, iz toga možemo zaključiti da je jedan tim brži od drugog.

Kada interpretiramo rezultate Men-Vitnijevog U testa i Vilkoksonovog testa sume rangova, **najvažniji statistik za svrhe interpretacije je, kao i kod ostalih testova, statistička značajnost**. Opet, **ako je vrednost statističke značajnosti veća od prihvaćenog praga (što je obično 0,05), nultu hipotezu prihvatamo, zaključujući da nema dokaza da dva uzorka dolaze iz različitih populacija, te ih onda tretiramo kao da dolaze iz iste populacije. Ako je vrednost statističke značajnosti manja od prihvaćenog praga (dakle obično ispod 0,05), nultu hipotezu odbacujemo i zaključujemo da dva poređena uzorka ne potiču iz iste populacije odnosno iz populacija sa jednakim svojstvima.**

Iako može biti relativno nejasno koje tačno statistike ova dva testa porede, njihovi rezultati nedvosmisleno govore o centralnim tendencijama populacija iz kojih su dva uzorka koja se porede

PREPRINT

uzeti. Ono što možemo naći u literaturi je to da istraživači često teže da tumače rezultate U testa kao da govore o odnosu između aritmetičkih sredina (mada će obično izbegavati da to tako direktno napišu).

I Men-Vitni U test i Vilkoksonov test sume rangova zahtevaju da podaci budu bar na ordinalnom nivou merenja, ali, kako su u pitanju neparametrijski testovi, ne postavljaju nikakve zahteve oko oblika distribucije podataka.

Veličine efekta koje se mogu koristiti sa ovim podacima su različite. U literaturi često možemo naići na **rang-biserijski koeficijent korelacije korišćen kao mera veličine efekta uz Men-Vitnijev U test**. Sreće se i **Spirmanov koeficijent korelacije rangova** (koji je dosta sličan rang-biserijskom koeficijentu) **izračunat između binarne varijable koja pokazuje koji entitet je iz kog uzorka i varijable na kojoj se uzorci porede u ulozi mere veličine efekta**. Takođe, Fild (Field, 2009) se u svojoj knjizi poziva na Rozentala i predlaže da se **z skor dobijen u okviru Vilkoksonovog testa pretvori u koeficijent korelacije tako što će se podeliti sa kvadratnim korenom iz ukupne veličine uzorka** (oba uzorka zajedno). Kako su sve ovo koeficijenti korelacije, pravila za njihovu interpretaciju su ista kao i kod svih drugih koeficijenata korelacije.

7.4. Poređenje centralnih tendencija zavisnih uzoraka – test znaka i Vilkoksonov test rangova sa predznakom

Najpoznatije alternative t testu za zavisne uzorke su test znaka i Vilkoksonov test rangova sa predznakom.

Test znaka testira nultu hipotezu da dva zavisna uzorka dolaze iz populacija sa jednakim medijanama. Sprovodi se tako što se **podaci iz dva uzorka rasporede u dve kolone, tako da parovi podataka koji odgovaraju jedan drugom iz dva uzorka budu u istom redu** (dakle podaci sa entiteta koji su jedan drugom par). Onda **za svaki entitet iz jednog uzorka zabeležimo da li od njemu odgovarajućeg entiteta iz drugog uzorka ima manju, veću ili jednaku vrednost na varijabli na kojoj se vrši poređenje**. Ako ima veću vrednost, to se obično obeležava sa +, ako je manja, to se obeležava sa -, a ako su vrednosti jednake, to se označava znakom =. Očekivanje je da će, **ako je nulta hipoteza tačna, verovatnoće javljanja + i – znakova biti jednake** tj. da će ukupan broj + i ukupan broj – znakova biti otprilike podjednak. Što je veća razlika između

PREPRINT

ukupnog broja + i – znakova, to će rezultat testa znaka biti statistički značajniji (tj. p vrednost će biti manja). Na primer, zamislimo da imamo parove studenata koji su napravljeni prema njihovim akademskim sposobnosti, poznavanju statistike i motivaciji da se pohađa kurs iz statistike (tako da par čine studenti koji imaju jednake vrednosti na ovim varijablama). Njih onda možemo da podelimo u dve grupe i to tako što ćemo staviti jednog studenta iz jednog para studenata u jednu grupu, a drugog u drugu. Onda organizujemo da jedna od te dve grupe studenata pohađa jedan kurs iz statistike, a druga grupa da pohađa drugi kurs iz statistike. Sačuvamo podatke o tome koji student iz jedne grupe odgovara kom studentu iz druge grupe, tako da ova dva uzorka možemo da organizujemo kao zavisne uzorke. Kako su studenti upareni, „kvalitet“ obe grupe studenata je jednak u pogledu njihove sposobnosti da iskoriste priliku da pohađaju kurs iz statistike. Imajući ovo u vidu, ako su oba kursa jednakog kvaliteta, očekivali bi da posle pohađanja kurseva znanje iz statistike i jedne i druge grupe bude jednako unapređeno. To znači da ako bi poredili vrednosti studenata na nekom testu iz statistike koji bi studenti radili nakon završenog kursa, očekivali bi da broj situacija u kojima je student iz prve grupe postigao bolji rezultat od svog para iz druge grupe bude otprilike jednak broju situacija gde je student iz druge grupe postigao bolji rezultat od svog para iz prve grupe. Ako ovi brojevi nisu otprilike jednaki, tj. ako je jasno da ima više slučajeva gde su studenti iz jedne grupe postigli bolje rezultate na testu nego studenti iz druge grupe, nego što je obrnutih slučajeva, to bi moglo da znači da kvalitet data dva kursa nije jednak, pod pretpostavkom da možemo razumno da isključimo bilo koji drugi faktor koji bi mogao da izazove razliku između grupa. Ovo je takozvani eksperimentalni istraživački nacrt i tu bi mogli da iskoristimo test znaka da izvedemo zaključke o njegovim rezultatima ako bi se pokazalo da iz bilo kog razloga ne možemo da koristimo t test za zavisne uzorke.

U stručnoj literaturi postoji određena **nejasnoća oko toga koji je nivo merenja potreban za korišćenje testa znaka. Minimalni nivo merenja koji je potreban za korišćenje testa znaka je ordinalni nivo.** Međutim, činjenica je i to da **ako podaci predstavljaju dve nezavisne rang liste**, tako da je svaki od dva para uzoraka rangiran rangovima od 1 do broja entiteta u tom uzorku, **test znaka neće biti upotrebljiv**, jer će broj znakova + i – biti nužno jednak. Međutim, **ako su vrednosti dobijene tako što su oba uzorka združeno rangirana ili ako je korišćen bilo koji drugi sistem merenja ili procene koji stavlja oba uzorka u isti referentni okvir**, to je **dovoljno za korišćenje testa** znaka. **Test znaka ne postavlja nikakve zahteve oko oblika distribucije podataka.**

PREPRINT

Vilkoksonov test rangova sa predznakom zasnovan je na istom konceptu kao i test znaka, sa tom razlikom da **on uzima u obzir i veličinu razlike između parova vrednosti koje se porede, a ne samo koja je veća/manja.** To je takođe neparametrijska alternativa t testu, ali, za razliku od testa znaka, **njegovi rezultati pun smiso dobijaju tek ako su podaci bar na intervalnom nivou merenja**. Ovo je z**bog toga što sprovođenje ovog testa zahteva računanje i poređenje razlika između vrednosti, a da bi takvo poređenje bilo smisleno neophodna je fiksna jedinica mere**, a najniži nivo merenja na kom to postoji je intervalni.

Postavka za sprovođenje Vilkoksonovog testa rangova sa predznakom je ista kao za test znaka, samo što je računica malo složenija:

- **Podaci se organizuju u dve paralelne kolone, tako da parovi vrednosti koje odgovaraju jedna drugoj budu u istom redu**, isto kao kod testa znaka.
- Potom se **računaju razlike između parova odgovarajućih vrednosti** (parova vrednosti koje smo stavili u isti red). Ove razlike mogu biti pozitivne, negativne ili nulte, zavisno od toga koja vrednost unutar para je veća ili da li su možda jednake.
- U sledećem koraku, **pozitivni i negativni predznaci se uklanjaju i sve razlike se potom rangiraju po veličini** (dakle – bez obzira na to da li je razlika izvorno bila pozitivna ili negativna) – najmanja razlika dobija rang 1, sledeća najmanja rang 2 i tako dalje. Ako je razlika između dva ili više parove vrednosti jednaka, sve takve vrednosti dobijaju rang koji je aritmetička sredina mesta na rang listi koja bi te razlike zauzimale. Na primer, ako postoje tri identične vrednosti razlika koje bi imale rangove 4, 5 i 6, onda sva tri para (odnosno sve tri razlike u vrednostima ovih parova) dobijaju rang 5, zato što je (4+5+6)/3=5.
- Nakon ovoga se **predznaci iz drugog koraka** (oni koje smo obrisali, da bi radili rangiranje) **pripisuju ovim rangovima**. Na ovaj način dobijamo znakove sa predznakom. Na primer, ako je razlika između dva para vrednosti dobila rang šest, a druga vrednost u paru je bila veća od prve, onda ovaj rang postaje -6. Ako je prva vrednost bila veća, onda rang postaje 6 (u stvari +6, ali obično ne pišemo predznak ispred pozitivnih vrednosti).
- **Rangovi sa predznakom se onda sabiraju i na osnovu ove sume se računa statistička značajnost ovog testa.**

PREPRINT

Nulta hipoteza Vilkoksonovog testa rangova sa predznakom je da dva zavisna uzorka dolaze iz iste populacije odnosno iz populacija sa istom centralnom tendencijom. **Kada je nulta hipoteza tačna, očekujemo da suma rangova sa predznakom bude oko 0. Što se više razlikuje od 0, to je rezultat Vilkoksonovog testa rangova sa predznakom statistički značajniji** (tj. vrednost statističke značajnosti postaje manja).

Kao i kod svih ostalih testova, tako i kod ovih testova, **kada je vrednost statističke značajnosti niža od prihvaćenog praga (a to je, uglavnom, 0,05), tj. kad je rezultat statistički značajan, nultu hipotezu odbacujemo i zaključujemo da centralne tendencije populacija iz kojih su dva uzorka nisu jednake. Ako je vrednost statističke značajnosti iznad prihvaćenog praga tj. ako rezultat nije statistički značajan, nultu hipotezu prihvatamo i zaključujemo da nema dokaza da se centralne tendencije populacija iz kojih su dva poređena uzorka različite.**

Tabela 7.3. Primer računanja testa znaka i Vilkoksonovog testa rangova sa predznakom. Prve tri kolone predstavljaju podatke, četvrta sa leva predstavlja deo računanja testa znaka. Četiri kolone sa desne strane predstavljaju postavku za računanje Vilkoksonovog testa rangova sa predznakom. Ako je nulta hipoteza tačna, broj + i - znakova kod testa znaka će biti otprilike podjednak. Ako je nulta hipoteza Vilkoksonovog testa rangova sa predznakom tačna, ukupna suma rangova će biti nula. Podaci su izmišljeni.

Podaci – vrednosti parova u dva zavisna uzorka			Test znaka	Vilkoksonov test rangova sa predznako,			
Parovi učesnika u istraživanju (osoba iz grupe 1 – osoba iz grupe 2)	Testni skorovi grupe 1	Testni skorovi grupe 2	Predznak razlike između grupa 1 i 2 (1-2)	Razlika između vrednosti 1 i 2	Razlika bez predznaka /apsolutna vrednost razlike	Rang apsolutne vrednosti razlike	Rang sa predznakom
Leposava – Ana	45	55	-	-10	10	2	-2
Anita – Snežana	22	67	-	-45	45	11	-11
Vladislava – Julija	37	25	+	12	12	3	3
Katarina – Vanja	55	32	+	23	23	8	8

PREPRINT

Isailo – Janko	25	25	=	0	0		
Jovan – Milan	52	12	+	40	40	10	10
Stojan – Damjan	50	70	-	-20	20	6.5	-6.5
Vladimir – Dušan	38	65	-	-27	27	9	-9
Gorčilo – Stevan	40	25	+	15	15	5	5
Marko – Viktor	25	30	-	-5	5	1	-1
Jelena – Mihaela	32	45	-	-13	13	4	-4
Petar - Filip	35	55	-	-20	20	6.5	-6.5
Rezultati testa znaka: Ukupan broj - :7 , ukupan broj + :4 Rezultati Vilkoksonovog test rangova sa predznakom: Suma negativni rangova: -40; Suma pozitivnih rangova: 26. Ukupna suma rangova: -14							

Mere veličine efekta koje mogu da se primene zajedno sa testom znaka i Vilkoksonovim testom uključuju **rang-biserijski koeficijent korelacije**, ali se takođe sreće i upotreba **Spirmanovog koeficijenta korelacije**. Neki autori radije računaju **Somersovo D, koje je mera veličine efekta zasnovana na poređenju broja parova entiteta kod kojih je vrednost prvog entiteta iz para veća od vrednosti drugog entiteta sa brojem parova kod kojih je vrednost drugog veća od vrednosti prvog** tj. **na poređenju broja + sa brojem – parova** (tipično se kod opisa postupka računanja ovog statistika ovi parovi zovu **konkordantnim i diskordantnim parovima**). **Somersovo D može da ima vrednosti između -1 i 1 i to zavisi od toga da li je više parova kod kojih je razlika između vrednosti pozitivna** (+ parovi, konkordantni parovi) **ili negativna** (- parovi, diskordantni parovi), tj. da li je poređenje vrednosti entiteta unutar parova rezultiralo sa više + ili sa više – znakova. Na primer, **ako svi parovi imaju - znakove** tj. kada je vrednost drugog entiteta u paru uvek veća od vrednosti prvog**, tada će D biti -1**. Međutim, slično načinu kako interpretiramo point-biserijski koeficijent korelacije, **predznak Somersovog D ne treba interpretirati kao da ima ikakvo substantivnije ili stalnije značenje od toga da nam pokazuje da li prva ili druga grupa teže da imaju više rangove**. Treba da imamo stalno u vidu da je to koji je uzorak prvi, a koji drugi, odnosno koji je entitet u paru prvi, a koji drugi, potpuno proizvoljno i da bi drugačiji način obeležavanja uzoraka doveo do drugačijeg rezultata kada je upitanju predznak D statistika.

PREPRINT

7.5. Poređenje standardnih devijacija/varijansi – Levenov test.

Test koji se u naučnoj i stručnoj literaturi često sreće kada postoji potreba za poređenjem varijansi ili standardnih devijacija je Levenov test. **Nulta hipoteza Levenovog testa je da sve poređene grupe dolaze iz populacija sa jednakim varijansama tj. da su razlike između varijansi populacija koje poredimo jednake 0.** Da, za razliku od ostalih testova predstavljenih u ovom poglavlju i suprotno od naslova ovog poglavlja**, Levenov test može da uporedi više od dve grupe istovremeno.** Kao i kod svih testova, **kada je rezultat Levenovog testa statistički značajan** (tj. kada je vrednost statističke značajnosti niža od prihvaćenog praga, koji je obično 0,05), **nultu hipotezu odbacujemo i zaključujemo da poređene grupe ne dolaze iz populacija sa jednakim varijansama**. Kada se ovaj test koristi za poređenje dva uzorka, to znači da populacije iz kojih su data dva uzorka imaju različite varijanse. Međutim, **kada poredimo više od dva uzorka, statistički značajan rezultat Levenovog testa znači da varijanse populacija o kojima je reč nisu iste, ali kako ih ima više od 2, ne znamo koja populacija verovatno ima različitu varijansu od koje tačno populacije od poređenih populacija**. **Statistički značajan rezultat Levenovog testa koji je sproveden da bi uporedili više grupa istovremeno ne znači automatski da se varijansa svake od poređenih populacija razlikuje od varijanse svake druge populacije**. On samo pokazuje da nisu varijanse svih grupa/uzoraka jednake, a što je više uzoraka koji se porede, to je više stvari na koje ovaj nalaz može konkretno da ukazuje. On, naravno, može zaista da znači da se varijansa svake grupe razlikuje od varijanse svake druge grupe, ali može da znači i da je verovatno da samo dve od poređenih populacija imaju jasno različite varijanse ili da se varijansa jedne populacije razlikuje od varijanse svih ostalih populacija, koje pak imaju jednake varijanse. Međutim, samo na osnovu jednog računanja Levenovog test na većem broju grupa, ne možemo da znamo koja od ovih mogućnosti se ostvarila. **Da bi ustanovili precizno varijanse kojih populacija se verovatno razliku od varijansa kojih populacija**, morali bi da ponovo sprovedemo veći broj Levenovih testova gde bi poredili svaki uzorak sa svakim uzorkom, tj. da **uzorke poredimo dva po dva ovim testom**. S druge strane, **kada rezultat Levenovog testa nije statistički značajan to znači da možemo da prihvatimo nultu hipotezu i onda zaključimo da nema dokaza da populacije iz kojih su uzeti poređeni uzorci imaju različite varijanse.**

PREPRINT

Levenov test je **parametrijski test,** što znači da pretpostavlja normalnu distribuciju uzorkovanja i da podaci budu bar na intervalnom nivou merenja. Takođe, imajući u vidu da je standardna devijacija zapravo koren iz varijanse, **rezultati Levenovog testa se jednako odnose i na standardne devijacije populacija iz kojih su poređeni uzorci**.

Levenov test se može koristiti da se testira pretpostavka o razlikama u varijabilnosti populacija iz kojih su uzorci, ali se v**erovatno najčešće sreće upotrebljen pre sprovođenja drugih statističkih testova** koji se oslanjaju na pretpostavku da su varijanse populacija iz kojih su poređeni uzorci jednake tj. **koji se oslanjaju na pretpostavku o homogenosti varijansi**. Takva je, na primer, situacija sa t testom koji je opisan u ovoj knjizi. **Levenov test se često računa pre računanja t testa da bi se proverilo da li su varijanse populacija iz kojih su dva uzorka koje želimo da poredimo jednake**. Ako Levenov test bude statistički značajan, tada zaključujemo da varijanse populacija iz kojih su poređeni uzorci nisu jednake i onda se t test radi bez pretpostavke o homogenosti varijansi, što znači da se primenjuje korekcija načina na koji se računa statistička značajnost t testa.

Jedan problem koji se javlja kod primene Levenovog testa za procenu homogenosti varijansi pre upotrebe nekog drugog testa nastaje **kada je veličina uzorka veoma velika**. Kao i kod svih računanja statističke značajnosti, kako uzorak postaje veći, tako standardna greška postaje manja. Zbog ovoga, **kada uzorak postane veoma veliki, veoma male razlike između poređenih vrednosti počinju da bivaju statistički značajne.** U tom slučaju, to znači da ćemo **kada radimo sa velikim uzorcima često sretati situacije kada je Levenov test statistički značajan, ali kada su razlike između varijansi grupa praktično zanemarljive**. U takvim situacijama, možemo prosto konstatovati da se ta situacija desila, pa zaključiti da su razlike između varijansi premale da bismo na osnovu njih odbacili pretpostavku o homogenosti varijansi. Međutim, bolji i objektivniji pristup je da izračunamo statistik poznat kao **odnos varijansi (eng. variance ratio)**, koji zapravo predstavlja **količnik najveće i najmanje od poređenih varijansi**. Kritične vrednosti odnosa varijansi zavise od broja entiteta u uzorku i broja uzoraka koji se porede i mogu se naći npr. kod Filda (Field, 2009).

7.6. Poređenje dve distribucije – Kolmogorov-Smirnov test, Hi kvadrat, Vald-Volfovic test

PREPRINT

Verovatno najpopularniji test za poređenje dve distribucije je Kolmogorov-Smirnov test, poznatiji i po svom skraćenom nazivu, K-S test. **Kolmogorov-Smirnov test testira nultu hipotezu da dva uzorka potiču iz populacija sa jednakim distribucijama**. Ovaj test se može koristiti za poređenje dve empirijske distribucije, ali u naučnoj literaturi, mnogo češće se sreće njegova upotreba za **poređenje jedne empirijske distribucije sa određenom teorijskom distribucijom** (koja ima istu aritmetičku sredinu i standardnu devijaciju kao empirijska distribucija s kojom se poređenje radi) tj. da se uporedi oblik distribucije koja je dobijena na uzorku datog istraživanja sa oblikom određene teorijske distribucije. Poređenje ovog tipa koje se najčešće sreće u literaturi je poređenje sa teorijskom normalnom distribucijom. Drugim rečima, **K-S test se najčešće sreće u situacijama gde se koristi kao test normalnosti distribucije**, tj. kada se koristi **za testiranje nulte hipoteze da uzorak istraživanja potiče iz populacije koja ima normalnu distribuciju na ispitivanoj varijabli.** To se **najčešće radi tokom preliminarnog ispitivanja empirijski dobijenh podataka, kada istraživači žele da procene da li mogu da koriste parametrijske testove** tj. testove koji se oslanjaju na pretpostavku o normalnoj distribuciji podataka ili na pretpostavku da je distribucija uzorkovanja normalna (a onda se testira normalnost distribucije podataka, kao zamena za ovo, pošto distribucija uzorkovanja nije dostupna, pa ne može onda ni da se testira hipoteza o njenom obliku).

Kolmogorov-Smirnov test **zahteva da podaci budu bar na intervalnom nivou merenja**. On takođe zahteva da **distribucija podataka bude kontinualna** (zbog čega na primer, Puasonova i binomna distribucija nisu adekvatne za ovaj test) i **potpuno određena.** Test se radi tako što se **računa serija razlika između dve distribucije, a onda je statistik K-S testa, koji se zove Kolmogorov-Smirnovljevo D, zasnovan na razlici između dve distribucije u tački gde je ta razlika najveća** (tj. na najvećoj razlici između dve distribucije). To znači da će rezultat K-S testa za dve distribucije koje se razlikuju u svim delovima i za dve distribucije koje su jednake svuda osim u jednom delu biti jednak, dokle god je najveća razlika između poređenih distribucija jednake veličine.

Kada je K-S test statistički značajan tj. kada je vrednost statističke značajnosti niža od prihvaćenog praga (što je uglavnom 0,05), nultu hipotezu odbacujemo i zaključujemo da poređeni uzorci ne potiču iz populacija sa jednakim distribucijama (to je zaključak kada

PREPRINT

poredimo dva uzorak) ili **da uzorak ne potiče iz populacije čija distribucija po obliku odgovara teorijskoj distribuciji sa kojom je poređena**. U najčešćoj varijanti upotrebe K-S testa, onda **kada se ovaj test koristi za poređenje distribucije uzorka sa normalnom distribucijom tj. kao test normalnosti, statistički značajan K-S test pokazuje da uzorak ne potiče iz populacije koja je normalno distribuirana na ispitivanoj varijabli**. S druge strane, **kada rezultati K-S testa normalnosti distribucije nisu statistički značajni tj. kada je vrednost statističke značajnosti iznad prihvaćenog praga (što je obično 0,05), možemo zaključiti da nema dokaza da distribucija populacije nije normalna, te se može tretirati kao da je u pitanju normalna distribucija**. Treba da budemo svesni i toga da i ovde, kao i kod svih drugih računanja statističke značajnosti, **kako raste veličina uzorka, tako se smanjuju veličine razlika koje prelaze prag statističke značajnosti**. To znači da će **na veoma velikim uzorcima čak i veoma mala odstupanja od normalne distribucije rezultirati statistički značajnim K-S testom**. Iz ovog razloga, **kada je uzorak veliki**, a cilj nam je da odlučimo da li da koristimo test koji se oslanja na pretpostavku o normalnoj distribuciji podataka (ili pretpostavku o normalnoj distribuciji podataka koristimo kao zamenu za pretpostavku o normalnoj distribuciji uzorkovanja!), **mnogo je razumnije oslanjati se na statistike koji pokazuju veličinu odstupanja od normalne distribucije – skjunes i kurtozis, umesto na K-S test**, zato što će, na dovoljno velikim uzorcima, čak i potpuno zanemarljiva odstupanja od normalne distribucije rezultirati statistički značajnim K-S testom.

Pirsonov hi kvadrat test (eng. chi square test), koji se u literaturi često pominje i samo kao hi **kvadrat ili hi kvadrat test** je **postupak za poređenje empirijske distribucije sa teorijskom distribucijom, pri čemu te distribucije moraju biti takve da se mogu predstaviti u obliku tabele frekvencija tj. kao skup diskretnih kategorija.** Ovo se radi sa ciljem da se **testira hipoteza da uzorak potiče iz populacije čija je distribucija ista kao teorijska distribucija s kojom vršimo poređenje.** Test se onda sprovodi tako što se **računaju frekvencije svake kategorije empirijski procenjene/merene varijable tj. broji se koliko entiteta spada u svaku od kategorija**. Frekvencije koje se dobiju na ovaj način zovu se **opažene frekvencije.** Na osnovu ukupnog broja entiteta u uzorku, **potom se pravi distribucija frekvencija po kategorijama koja je u skladu sa teorijskom distribucijom sa kojom hoćemo da poredimo našu empirijsku distribuciju**. Drugim rečimo, u ovom drugom koraku procenjujemo kolika bi bila frekvencija svake od kategorija (tj. koliko bi entiteta bilo u svakoj od kategorija) ako bi varijabla koju

PREPRINT

razmatramo imala onu teorijsku distribuciju sa kojom želimo da izvršimo pređenje. **Frekvencije koje napravimo na ovaj način zovu se očekivane frekvencije, a cela distribucija koju čine zove se očekivana distribucija**. Statistik ovog testa, koji se **zove hi kvadrat statistik se onda računa na osnovu sume kvadrata razlika između ova dva skupa frekvencija** (opaženih i očekivanih, oduzimaju se opažene i očekivane frekvencije za svaku kategoriju date varijable). **Što su veće razike između opažene i očekivane distribucije i što je više kategorija sa razlikama između opažene i očekivane frekvencije, to će hi kvadrat statistik biti veći i statistički značajniji** (naravno, za istu veličinu uzorka!). **Nulta hipoteza hi kvadrat testa je da uzorak potiče iz populacije čija je distribucija identična očekivanoj distribuciji.** Kada **je hi kvadrat statistički značajan, nultu hipotezu odbacujemo i zaključujemo da distribucija populacije iz koje je uzorak nije jednaka očekivanoj distribuciji**. Ako **hi kvadrat nije statistički značajan, zaključujemo da nema dovoljno dokaza da se distribucija populacije iz koje je uzorak razlikuje od teorijske distribucije na kojoj smo zasnovali očekivane frekvencije**.

Jedno važno svojstvo hi kvadrat testa je to da on **može da se koristi na nominalnim podacima** (radi sa kategorijama!), ali **hi kvadrat test može da se računa i za izvorno kontinualne podatke koji su pretvoreni u određeni broj fiksnih kategorija.**

Najčešće korišćena teorijska distribucija s kojom se porede empirijske distribucije računanjem hi kvadrat testa je **uniformna distribucija**, ali se **hi kvadrat test može koristiti za poređenje empirijskih podataka sa bilo kojom teorijskom distribucijom dokle god se na osnovu te distribucije mogu napraviti očekivane frekvencije**. Može se koristiti i kao test normalnosti distribucije i, u toj ulozu, se hi kvadrat razlikuje od K-S testa u činjenici da na njegovu vrednost utiču sve razlike između poređenih rezultata, a ne samo najveća razlika, kao što je to slučaj kod K-S testa.

Još **jedan čest način upotrebe hi kvadrat testa je da se testira nulta hipoteza o da dve nominalne varijable** (ili bar kategorijalne ili svedene na fiksni broj kategorija) **potiču iz populacija u kojima nisu povezane**. Ovo se proverava tako što se prvo napravi krostabulacija dve varijable tj. naprave se kategorije koje su definisane vrednostima na obe varijable. Na primer, ako su naše dve varijable pol i boja kosa i pol ima vrednosti muško i žensko, a boja kose ima moguće vrednosti crna, braon, plava i druga, mi bi napravili po jednu kategoriju za svaku moguću kombinaciju vrednosti dve varijable. Tako bi imali jednu kategoriju za muškarce sa crnom kosom,

PREPRINT

za žene sa crnom kosom, muškarce sa braon kosom, žene sa braon kosom i tako redom dok ne iscrpimo sve kombinacije kategorija. Potom izračunamo tj. prebrojimo frekvencije entiteta u svakoj od ovih (ovako dobijenih, kombinovanih) kategorija u uzorku i to su onda opažene frekvenceije. Na kraju, napravimo procenu toga kolike bi bile frekvencije u uzorku naše veličine za svaku kategoriju, ako ove dve varijable ne bi bile povezane i te procene nam služe kao očekivane frekvencije za računanje hi kvadrat statistika. **U slučaju korišćenja hi kvadrata za testiranje povezanosti između varijabli, odbacivanje nulte hipoteze** (tj. hi kvadrat koji je statistički značajan) **pokazuje da su te dve posmatrane varijable verovatno povezane u populaciji**. **Prihvatanje nulte hipoteze pokazuje da nema dovoljno dokaza da su date dve varijable povezane u populaciji, što znači da treba da ih tretiramo kao da nisu povezane u populaciji.**

Jedan **zahtev Pirsonovog hi kvadrat testa je to da opservacije tj. merenja budu nezavisna**. To znači da hi kvadrat test nije odgovarajući statistik u situacijama kada u uzorku imamo višestruke mere dobijene na istom entitetu. Još jedan važan zahtev ovog testa je to da **nema očekivanih frekvencija koje su manje od 5**. Kada se tako nešto desi, da očekivana frekvencija neke kategorije bude manja od 5), istraživači treba da razmotre da, recimo, spoje neke kategorije, ako se to može smisleno uraditi da bi sve očekivane frekvencije bile iznad 5 ili da iskoristi neki alternativni test koji je odgovarajući za situacije kada je uzorak mali. Na primer, jedna **često pominjana alternativa Pirsonovom hi kvadrat testu** namenjena situacijama kada treba **testirani povezanost dve nominalne varijable, a uzorak je mali je Fišerov egzaktni test.** Takođe, **kada se Pirsonov hi kvadrat koristi na tabelama 2x2** (tj. za ispitivanje povezanosti dve binarne varijable)**,** sugestija je (e.g. Field, 2009) da hi kvadrat u takvim situacijama teži da daje previše male vrednosti statističke značajnosti (tj. lako postaje statistički značajan) i **da bi tada trebalo primeniti Jejtsovu korekciju za kontinutitet.** Ova korekcija efektivno **smanjuje veličinu razlike između svakog para očekivanih i opaženih frekvencija pre nego što se ta razlika digne na kvadrat prilikom računanja hi kvadrat statistika i tako smanjuje veličinu hi kvadrat statistika**.

Različiti statistici se mogu koristiti kao **mere veličine efekta** zajedno sa Pirsonovim hi kvadrat testom. Kada se hi kvadrat test koristi za testiranje povezanosti dve nominalne varijable, **koeficijenti korelacije namenjeni nominalnim varijablama mogu da budu adekvatne mere veličine efekta**. Ovi koeficijenti uključuju koeficijent kontingencije, fi koeficijent i Kramerov V. Ovi koeficijenti se sami mogu izvesti iz vrednosti hi kvadrat statistika ili se njihovo računanje

PREPRINT

zasniva na sličnim podacima. Još jedna moguća mera veličine efekta koja se može koristiti zajedno sa hi kvadrat testom je **količnik verovatnoće ili ods racio**. **Količnik verovatnoće (ods racio) je prosto količnik dve verovatnoće i pokazuje koliko puta je verovatnoća jednog ishoda veća od verovatnoće drugog ishoda.** Dobija se tako što se podele dve verovatnoće (jedna se podeli drugom). Količnik verovatnoće je najlakše intepretirati onda kada imamo posla sa tabelom frekvencija 2x2. Na primer, ako jedna varijabla sadrži podatke o tome da li je učesnik u istraživanju učenik osnovne ili srednje škole, a druga da li je učenik pao ili položio određeni test, mogli bi da izraćunamo verovatnoću da učenik srednje škole položi test i verovatnoću da učenik osnovne škole položi taj test. To bi uradili tako što bi na uzorku izračunali proporciju srednjoškolaca koji su položili test u odnosu na ukupan broj srednjoškolaca u uzorku i proporciju učenika osnovne škole iz uzorka u odnosu na ukupan broj učenika osnovne škole u uzorku i te proporcije proglasili za verovatnoće (da osnovnoškolac odnosno srednjoškolac položi test). Onda bi te dve, tako dobijene, verovatnoće podelili jednu s drugom i tako dobili na primer, koliko je puta verovatnije da srednjoškolac položi test nego osnovnoškolac.

U situaciji **kada imamo samo jednu binarnu varijablu i želimo da testiramo da li uzorak potiče iz populacije sa određenom proporcijom entiteta u svakoj od kategorija** (na primer, populacije u kojoj su obe kategorije jednako verovatno ili gde postoji određena, poznata verovatnoća svake od dve kategorije), **test koji se može upotrebiti za to je binomni test**. On **testira nultu hipotezu da uzorak potiče iz populacije u kojoj su verovatnoće javljanja svake od dve kategorije onakve kakve smo definisali**. Ovaj test je isključivo namenjen izvođenju zaključaka o proprocijama u populaciji dve kategorije jedne jedine binarne varijable.

Vald-Volfovic test (eng. Wald-Wolfovitz test) **je test za poređenje distribucija dva nezavisne uzorka koji se može primeniti onda kada su podaci bar na ordinalnom nivou merenja.** On **testira nultu hipotezu da dva nezavisna uzorka potiču iz populacija sa identičnim distribucijama**. Pretpostavka je da su distribucije identične i u pogledu oblika i u pogledu centralne tendencije tj. **da se one potpuno preklapaju**. Premisa na kojoj je test zasnovan je ta da **ako imamo zajedničku rang listu zasnovanu na vrednostima entiteta iz oba uzorka, ako su dve distribucije identične, entiteti će na rang listi biti izmešani** tj. neće biti mnogo neprekinutih serija/nizova entiteta iz istog uzorka kako idemo niz rang listu. Drugim rečima, ako krenemo od početka liste, naići ćemo na jedan ili par entiteta iz jednog uzorka, pa na jedan ili par entiteta iz

PREPRINT

drugog, pa na jedan ili par iz prvog, pa na jedan ili par iz drugog i tako do kraja. S druge strane, ako, recimo, svi entiteti iz jednog uzorka imaju veće vrednosti od svih entiteta iz drugog uzorka, što je najekstremnija razlika koju može da detektuje ovaj test, ako gledamo od početka rang liste, prvo ćemo videti niz koji čine svi entiteti iz jednog uzorka, a za njim niz koji čine svi entiteti iz drugog uzorka. **Vald-Volfovic test se radi tako što se krene od početka rang liste i broje se neprekinute serije/nizovi entiteta iz istog uzorka**. **Broj takvih nizova se onda poredi sa maksimalnim mogućim brojem takvih nizova, a na osnovu toga se onda računa statistička značajnost ovog testa.** Kada su **dva uzorka jednake veličine, maksimalani broj nizova jednak je broj entiteta u oba uzorka zajedno**. Međutim, **kada dva uzorka koja poredimo nisu jednake veličine, maksimalni broj nizova je = 2x broj entiteta u manjem uzorku +1**. Ovo je tako zato što kada se ispred i iza pozicije svakog entiteta iz manjeg uzorka nalazi po jedan entitet iz većeg uzorka, i dalje ćemo imati entitete iz većeg uzorka koji su jedan pored drugog na rang listi prosto zato što nema dovoljno entiteta iz manje grupe da prekinu sve moguće nizove entiteta iz veće grupe, prosto zato što u manjoj grupi ima manje entiteta. Ovo možemo da zamislimo kao zadatak deljenja jedne jedinstvene linije određenim brojem tačaka. Na primer, ako imamo 4 tačke koje možemo da rasporedimo duž linije, na koliko delova možemo liniju da podelimo koristeći ove 4 tačke? Odgovore je, naravno, 5 delova. Ako zamislimo da su te tačke u stvari pojedinačni entiteti iz manje grupe, dolazimo da formule za maksimalan broj nizova u situaciji kada grupe nisu jednake – 2x4+1=9 (5 delova linije i 4 tačke).

Sve u svemu, što je veća razlika između maksimalnog mogućeg broja neprekinutih nizova i opaženog broja nizova, to će Vald-Volfovic test biti statistički značajniji. **Kada je vrednost statističke značajnosti Vald-Volfovic testa niža od prihvaćenog praga statističke značajnosti (što je obično 0,05) tj. kada je rezultat statistički značajan, nultu hipotezu odbacujemo i zaključujemo da dva poređena uzorka ne potiču iz populacija sa jednakim distribucijama**. Međutim, **kada je vrednost statističke značajnosti iznad prihvaćenog praga tj. kada test nije statistički značajan, nultu hipotezu prihvatamo i zaključujemo da nema dovoljno dokaza da uzorci potiču iz dve različite populacije ili iz populacija koje se razlikuju i onda distribucije tretiramo kao da su jednake.**

Jedna važna stvar koju treba primetiti je to da **Vald-Volfovic test radi najbolje onda kada nema situacija da entiteti iz različitih uzoraka imaju istu vrednost** tj. **kada nema vezanih rangova**

PREPRINT

entiteta iz različitih uzoraka. Kada ima puno vezanih rangova entiteta iz različitih uzoraka, test postaje neupotrebljiv. Ovo je zbog toga što u situacijama u kojima više entiteta iz različitih uzoraka ima istu vrednost moramo da odlučimo koliko će neprekinutih nizova biti na takvim mestima. Na primer, ako imamo 10 entiteta koji imaju istu vrednost na ispitivanoj varijabli, to znači da će svi deliti isti rang. Zamislimo sada da je, od tih 10 entiteta, 5 iz jednog uzorka, a 5 iz drugog uzorka. Mi onda možemo da smatramo da tih 10 entiteta čini svega 2 niza - prvo 5 entiteta iz jednog uzorka, pa onda 5 entiteta iz drugog uzorka tj. 2 niza. A možemo, podjednako valjano i da smatramo da ovi entiteti čine 10 nizova – jedan iz prvog uzorka, pa jedan iz drugog, pa jedan iz prvog, pa jedan iz drugog i tako sve do 10. Imajući u vidu da svi entiteti imaju istu vrednost, obe ove interpretacije su jednako valjane. Međutim, u prvom slučaju ćemo tu imati 2 niza od 10 entiteta, što dosta doprinosi tome da test postane statistički značajan, a u drugom slučaju ćemo imati 10 nizova od 10 slučajeva, što onda test u celini udaljava od statistički značajnog rezultata. Dok su takvi vezani rangovi samo mali deo uzorka, razlika između rezultata testa ako usvojimo jedno od ova dva tumačenja i ako usvojimo drugo neće biti mnogo različita. Međutim, kada je broj vezanih rangova veći tj. kada oni uključuju mnogo više entiteta, toliko mnogo da je veliki deo uzorka uključen u vezane rangove na ovaj ili onaj način, test postaje neupotrebljiv zato što su zaključci koji slede iz jedne i druge interpretacije broja nizova kod vezanih rangova potpuno suprotni.

Tabela 7.4. Primeri različitih brojeva nizova izbrojanih u okviru Vald-Volfovic testa i predstavljanje kako se broje nizova računa. Primer 1 je najmanji mogući broj nizova, situacija kada rezultati test dostižu najveću moguću statističku značajnost (tj. najmanju moguću p vrednost), to je najveća moguća razlika od stanja pretpostavljenog nultom hipotezom. Primeri 2 i 3 predstavljaju situacije pretpostavljene nultom hipotezom – najveći mogući broj nizova što ukazuje na potpuno poklapanje distribucija, s tom razlikom sto primer 2 prikazuje situaciju kada su veličine grupa jednake, a primer 3 situaciju kada veličine grupa nisu jednake. Primer 4 predstavlja situaciju kada postoji određeno preklašanje između distribucija, ali ono nije potpuno tj. broj nizova je negde između minimumalnog i maksimalnog mogućeg broja. Nijedan od ovih primera ne predstavlja situaciju sa vezanim rangovima u kojoj bi test kao rezultat dao dva broja nizova, jedan ako

PREPRINT

računamo da su entiteti na vezanim rangovima izmešani, a drugi ako bi entitete sa vezanim rangovima računali kao da nisu izmešani.

Rangovi sa zajedničke rang liste dve grupe koje treba porediti (postavka za Vald-Volfovic test)	Ovde imamo dve grupe – 1 i 2 i brojevi u ovim kolonama predstavljaju grupu iz koje je entitet koji je postigao rang koji je označen datirm redom (leva kolona u svakom primeru) i prebrojavanje nizova entiteta iz iste grupe (desna kolona u svakom od primera).							
	Primer 1 – nema mešanja između grupa, jedna grupa ima jasno više vrednosti, distribucije su potpuno različite, bez preklapanja		Primer 2 – savršena pomešanost između grupa, distribucije se potpuno poklapaju		Primer 3 – savršena pomešanost između grupa, ali grupa 1 je 3x manja od grupe 2, distribucije se potpuno poklapaju		Primer 4 – grupe su pomešane, ali samo delimično, nejednake grupe, distribucije se delimično preklapaju	
	Grupa kojoj entitet pripada	Niz po redu	Grupa kojoj entitet pripada	Niz po redu	Grupa kojoj entitet pripada	Niz po redu	Grupa kojoj entitet pripada	Niz po redu
1	1	1	1	1	2	1	2	1
2	1	1	2	2	2	1	2	1
3	1	1	1	3	2	1	2	1
4	1	1	2	4	1	2	2	1
5	1	1	1	5	2	3	1	2
6	1	1	2	6	2	3	1	2
7	1	1	1	7	1	4	2	3
8	1	1	2	8	2	5	1	4
9	1	1	1	9	2	5	1	4
10	1	1	2	10	1	6	1	4
11	2	2	1	11	2	7	2	5
12	2	2	2	12	2	7	2	5
13	2	2	1	13	1	8	1	6
14	2	2	2	14	2	9	1	6
15	2	2	1	15	2	9	2	7
16	2	2	2	16	2	9	1	8
17	2	2	1	17	1	10	2	9
18	2	2	2	18	2	11	2	9
19	2	2	1	19	2	11	2	9
20	2	2	2	20	2	11	2	9
Ukupan broj nizova	**2**		**20**		**11**		**9**	
Maksimalni mogući broj nizova imajući u vidu veličine grupa	**20**		**20**		**11**		**17**	

PREPRINT

7.7. Hajde da primenimo šta smo do sada naučili!

Hajde da probamo sada da primenimo stvari koje smo predstavili u ovom poglavlju kroz nekoliko vežbi. Molimo vas da pogledate opšte uputstvo za ovakve vežbe koje možete naći na početku knjige. Naša preporuka je da prvo pročitate svaki isečak i tvrdnje date u njemu i da onda date svoj odgovor. Odgovor možete upisati u kolonu za odgovore, a posle toga pročitajte odgovore i uporedite svoje odgovore sa njima.

Vežba O. Statistički testovi.

Pol	AS	SD	t	p	Veličina efekta (r_{pbis})
Muški	40.03	16.03	2.196	.03	.171
Ženski	35.08	12.27			
Bračno stanje	AS	SD	t	p	Veličina efekta (r_{pbis})
U braku	33.14	13.01	-2.193	.03	.159
Nije u braku	38.15	13.99			
Zavisna varijabla – Prekomerna upotreba interneta					

Tabela je zasnovana na podacima prikazanim u istraživanju Rakić-Bajić & Hedrih (2012)

O	Tvrdnja:	Odgovor
O1.	Postoji statistički značajna razlika između aritmetičkih sredina muškaraca i žena na Prekomernoj upotrebi interneta.	
O2.	Postoji statistički značajna razlika između aritmetičkih sredina ljudi koji su u braku i onih koji nisu na Prekomernoj upotrebi interneta.	
O3.	Veličina razlike između muškaraca i žena na Prekomernoj upotrebi interneta je visoka.	
O4.	Veličina razlike između prosečnih vrednosti osoba koje su u braku i onih koje nisu na Prekomernoj upotrebi interneta je srednja.	
O5.	Test koji je ovde primenjen je parametrijski.	
O6.	Prekomerna upotreba interneta je na nominalnom nivou merenja.	
O7.	U populaciji postoji vrlo supstantivna razlika između standardnih devijacija osoba koje su u braku i onih koje nisu u braku na Prekomernoj upotrebi interneta.	
O8.	Razlika između standardnih devijacija osoba koje su u braku i onih koje nisu u braku na Prekomernoj upotrebi interneta nije statistički značajna.	
O9.	Ovde je mogao da bude primenjen i U test za sličnu svrhu umesto t testa.	
O10.	Prekomerna upotreba interneta je izraženija kod muškaraca nego kod žena.	

PREZENT

Vežba P. Statistički testovi.

PREPRINT

Differences scores between groups (passed the driving test vs. did not pass the driving test or passed it with adjustments) on different tests and demographic variables

		N	*U*	*df*	*p*	*r*
Demographic variables	Age	63	461	61	.780	0.04
	Education (in years)	63	587	61	.119	0.20
	Amnesia (in weeks)	33	74	31	**.027**	0.39
	Time from injury (in months)	63	285	61	**.006**	0.35
	Coma duration (in weeks)	40	82	38	**.001**	0.50
	GCS	40	317	38	**.001**	0.51
Mediatester variables	RTAV (*t*-value)	63	729	61	**< .001**	0.44
	ART Total RT (in seconds)	62	386	60	**.272**	0.14
	CRT – correct reactions – (M in seconds)	63	361	61	.092	0.21
	18 LRT (in seconds)	63	315	61	**.020**	0.29
	VRT Total RT (in seconds)	62	404	60	.357	0.12
Neuropsychological assessment variables	TOL – Total Move Score	63	385	61	.180	0.17
	CVLT (sum of all recalled words in all the trials)	63	629	61	**.039**	0.26
	HOVT (sum)	63	599	61	.100	0.43
	D2	63	228	61	**< .001**	0.46
	CTMT (composite index – T)	63	743	61	**.000**	0.45
	COG (*t*-value)	63	637	61	**.029**	0.27
	LVT (*t*-value)	63	665	61	**.009**	0.33
Driving-related variables	Average of all the ratings from the instructor	47	458	45	**< .001**	0.58

Note. The results where $p < .05$ are shown in bold. The variables included in the final regression model are shown in italic (GCS, 18 LRT, CTMT).
N = number of participants; *U* = Wilcoxon-Mann-Whitney U statistic; *df* = degrees of freedom; *r* = size of effect measure.

U tabeli su prikazane razlike na nizu varijabli između grupe osoba koje su položile test vožnje nakon traumatske povrede mozga i grupe osoba koje ga nisu položile ili su ga položile sa prilagođavanjima. N se odnosi na broj osoba u uzorku, U je U statistik Men-Vitnijevog U testa, r je mera veličine efekta. U tvrdnjama će imena varijabli biti ili u originalu na engleskom ili će njihovo ime u originalu na engleskom biti dato u zagradi pored imena na srpskom.
Tabela preštampana iz: Čižman Štaba, U., Klun, T. R., & Robida, R. (2021). Predicting Factors of Driving Abilities after Acquired Brain Injury through Combined Neuropsychological and Mediatester Driving Assessment. Psihologija, 54(2), 137–154. https://doi.org/10.2298/PSI200408024C . Preštampano na osnovu dozvole autora.

P	Tvrdnja:	Odgovor
P1.	Osobe koje uspeju da polože test vožnje nakon traumatske povrede mozga teže da budu starije od onih koje ne uspeju.	

P2.	Razlika između ove dve grupe na varijabli RTAV je statistički značajna.	
P3.	Razlika između poređenih grupa je veća na 18 LRT nego na RTAV.	
P4.	Mera veličine efekta koja je prikazana je rang-biserijski koeficijent korelacije.	
P5.	Veštine vožnje direktno uzrokuju razlike između grupa na varijabli D2.	
P6.	Skor medijanske značajnosti je visok samo za varijablu D2.	
P7.	Statistički test koji je ovde primenjen za poređenje dve grupe je parametrijski.	
P8.	Ove dve grupe potiču iz populacija sa otprilike jednakom starošću (varijabla Age).	
P9.	Nultu hipotezu ovde primenjenog testa bi trebalo odbaciti za varijablu Trajanje kome (varijabla Coma Duration (in weeks)).	
P10.	Nultu hipotezu ovde primenjenog testa treba prihvatiti za varijablu GCS.	

Vežba Q. Statistički testovi

			Bračno stanje		Ukupno
			U braku	Nije u braku	
Nivo obrazovanja	Univerzitet	% unutar obrazovanja	42.7%	57.3%	100%
		% unutar bračnog stanja	83.7%	39.6%	51.1%
		% od ukupnog broja	21.8%	29.3%	51.1%
	Srednja škola	% unutar obrazovanja	8.2%	91.8%	100%
		% unutar bračnog stanja	14.3%	56.1%	45.2%
		% od ukupnog broja	3.7%	41.5%	45.2%
	Drugo	% unutar obrazovanja	14.3%	85.7%	100%
		% unutar bračnog stanja	2.0%	4.3%	3.7%
		% od ukupnog broja	.5%	3.2%	3.7%
	Ukupno	% unutar obrazovanja	26.1%	73.9%	100%
		% unutar bračnog stanja	100%	100%	100%
		% od ukupnog broja	26.1%	73.9%	100

Naziv statistika	Vrednost statistika	Statistička značajnost
Pirsonov Hi kvadrat	28.325	p<.001
Kramerov V koeficijent	.388	p<.001
N	188	

Podaci su iz sopstvenog istraživanja autora.

Q	Tvrdnja:	Odgovor

PREPRINT

Q1.	Povezanost između bračnog stanja i nivoa obrazovanja nije statistički značajna.	
Q2.	Intenzitet povezanosti između bračnog stanja i nivoa obrazovanja je umeren, ako se vodimo Koenovim preporukama za interpretaciju.	
Q3.	Varijabla Nivo obrazovanja je na nominalnom nivou merenja.	
Q4.	U ovom uzorku, srednjoškolsko obrazovanje je učestalije kod učesnika u istraživanju koji su u braku nego kod onih koji nisu u braku.	
Q5.	U uzorku ima više od 200 učesnika.	
Q6.	Koeficijent korelacije koji je ovde računat može da ima negativne vrednosti.	
Q7.	Mod varijable Nivo obrazovanja na grupi ljudi koji su u braku je Univerzitet.	
Q8.	Varijabla Nivo obrazovanja je binarna (dihotomna).	
Q9.	Varijabla Bračno stanje je binarna (dihotomna).	
Q10.	U uzorku ima više muškaraca nego žena.	

Pogledajmo sada odgovore:

O1 – tačno. Možemo videti da je t statistik za poređenje muškaraca i žena statistički značajan, što ukazuje da je razlika između ove dve populacije statistički značajna. Možemo videti da je vrednost statističke značajnosti – p vrednosti 0,03. Takođe možemo da pročitamo da je Prekomerna upotreba interneta zavisna varijabla, što znači da je to varijabla na kojoj su izračunate aritmetičke sredine grupa i u odnosu na koju je rađeno poređenje.

O2 – tačno. Slično kao kod O1, samo što treba da pogledamo p vrednost za bračno stanje.

O3 - netačno. Veličina koeficijenta korelacije koji se koristi kao mera veličine efekta je 0,171, a to se smatra niskom korelacijom prema svim preporukama za interpretaciju veličine koeficijenta korelacije koje su date u ovoj knjizi.

O4 – netačno. Veličina razlike između aritmetičkih sredina dve grupe na bračnom stanju je 0,159, kao što možemo da pročitamo iz tabele, a to je niska korelacija.

O5 – tačno. Test koji je ovde korišćen je t test, koji je zaista parametrijski test.

O6 – netačno. S obzirom na to da možemo da vidimo da je sproveden t test, a da su za Prekomernu upotrebu interneta računate aritmetičke sredine i standardne devijacije, to ne može da bude nominalna varijabla, već mora da bude bar na intervalnom nivou merenja.

O7 – nepoznato. Iako se homogenost varijansi često navodi kao uslov koji treba da bude ispunjen da bi računali t test, postoji varijanta za sprovođenje ovog postupka kada varijanse nisu jednake. Ono što važi za varijanse, važi i za standardne devijacije. U tabeli nemamo informacije o populacijskim vrednostima standardnih devijacija dve grupe, niti je sproveden bilo koji test koji bi testirao neku hipotezu oko toga.

O8 – nepoznato. Ova tvrdnja je formulisana donekle različito, na način koji nalikuje onome što će čitaoci verovatno sresti u naučnim diskusijama ili u literaturi, ali ona kaže isto što i O7, samo za

PREPRINT

drugu varijabli. Odgovor je isti. Nema podataka o odnosima standardnih devijacija populacija iz kojih su ove dve grupe.

O9 – tačno. Ovo su nezavisni uzorci, što znači da bi bilo ispravno koristiti Men-Vitnijev U test. Kad god se koristi t test za nezavisne uzorke, može da se koristi i U test.

O10 – tačno. Iz tabele možemo videti da muškarci imaju veću aritmetičku sredinu i da je razlika između aritmetičkih sredina statistički značajna. Prema tome, možemo zaključiti da verovatno postoji razlika između aritmetičkih sredina populacija iz kojih su ove dve grupe i da ona ima onaj smer koji je naveden u tvrdnji. Ta razlika je veoma mala.

P – opšti komentari. Gledajući tabelu, možemo videti da ona sadrži rezultate Men-Vitnijevog U testa koji je sproveden da uporedi grupe ljudi koji su položili test vožnje i onih koji nisu uspeli da polože test nakon traumatske povrede mozga. To je razlog zašto su varijable poput Amnezija (Amnesia), Vreme od povrede (Time from Injury), Trajanje kome (Coma duration) i slične u tabeli. Možemo videti U statistik u koloni koja se zove U, p vrednost tj. statističku značajnost u posebnoj koloni i nešto što je verovatno koreficijent korelacije označen sa r u krajnje desnoj koloni koja je upotrebljena kao mera veličina efekta. U tabeli se ne pominje koji tačno koeficijenti korelacije su upotrebljeni, ali r je standardna oznaka za koeficijente korelacije u tabeli poput ove, tako da ćemo ga tretirati tako.

P1 – netačno. Starost je varijabla u prvom redu (prvi red posle imena statistika) i možemo videti da razlika izmđu dve grupe nije statistički značajna. P vrednost je 0,78, što je daleko od tipično prihvaćenih pragova statističke značajnosti. Prema tome, nultu hipotezu prihvatamo i zaključujemo da nema razlike u starosti između dve grupe. Da, ovo su rezultati U testa, tako da nije baš da se odnose na aritmetičke sredine, ali se može protumačiti kao da se odnosi na razlike između centralnih tendencija dve grupe. Takođe, postoji i korelacija između grupe i starosti koja takođe nije statistički značajna.

P2 – tačno. Možemo da vidimo da je p vrednost za razliku između dve grupe na RTAV veoma niska (što ukazuje na visoko statistički značajan rezultat!) i autori su takođe označili razliku boldovanim slovima tako naglašavajući da je statistički značajna.

P3 – netačno. Trebalo bi da uporedimo mere veličine efekta za dve greupe (kolonu r) i tu možemo videti, suprotno od onog što kaže tvrdnja, da je veličina efekta veća na RTAV (0,44 naspram 0,29).

P4 – netačno. Iako možemo iz oznake – r da zaključimo da je u pitanju neka vrsta koeficijenta korelacije, nema podataka o tome koji tačno. Mogao bi da bude rang-biserijski, mogao bi da bude Spirmanov ili neki drugi – ne znamo zasigurno iz podataka koji su prikazani.

P5 – nepoznato. Iako možemo videti da se ove dve grupe razlikuju na datoj varijabli i da su grupe formirane na osnovu toga da li je osoba položila test vožnje ili nije, ne znamo da li je razlika na ovoj određenoj varijabli posledica veština vožnje ili nečeg drugog.

P6 – besmisleno. „Skor medijanske značajnosti“ ne postoji, što znači da ne može ni da bude visok za varijablu.

P7 – netačno. Korišćen je Men-Vitnijev U test i to je neparametrijski test.

PREPRINT

P8 – tačno. Možemo videti da razlika između dve grupe na varijabli Age nije statistički značajna.

P9 – tačno. Možemo videti da su rezultati U testa statitički značajni za ovu varijabli i da je p vrednost veoma, veoma mala, kao i da je veličina efekta veoma supstantivna. To znači da bi trebalo da odbacimo nultu hipotezu. Osobe koje su uspele da polože test vožnje nakon traumatske povrede mozga i one koje nisu, teže da se razlikuju u priličnoj meri u pogledu prosečnog vremena koje su proveli u komi, kao što prikazani rezultati i pokazuju.

P10 – netačno. Možemo videti da je U statistik statistički značajan za ovu varijablu, što ukazuje da bi trebalo da odbacimo nultu hipotezu, nasuprot onome što tvrdnja kaže.

Q – opšti komentar. Tabela prikazuje rezultate hi kvadrat testa kojim je testirana nulta hipoteza da nema povezanosti između varijabli Bračno stanje i Nivo obrazovanja. Možemo vdieti i da je ovde bračno stanje binarna varijabla, dok je Nivo obrazovanja nominalna varijabla sa 3 kategorije. Pored hi kvadrat statistika, izračunat je i prikazan i Kramerov V koeficijent korelacije.

Q1 – netačno. Možemo videti da su i Kramerov V koeficijent i hi kvadrat statitički značajni, što ukazuje da su dve varijable za koje su ove mere računate povezane tj. da je njihova povezanost statistički značajna.

Q2 – tačno. Prema Koenovim preporukama korelacija od 0,388 je umerena.

Q3 – tačno. Možemo videti da su njene tri kategorije Univerzitet, Srednja škola i Drugo. Iako se može opravdano smatrati da je Univerzitet viši nivo obrazovanja od Srednje škole, takav odnose ne stoji za kategoriju Drugo, pa je zbog toga ovo nominalna varijabla (a ne ordinalna!). Kao nominalna varijabla je i korišćena u prikazanim proračunima.

Q4 – netačno. Možemo videti da od svih učesnika u istraživanju koji su u braku, samo 14,3% ima Srednju školu, ali da je ovaj procenat 56,1% među učesnicima u istraživanju koji nisu u braku.

Q5 – netačno. Možemo pročitati iz poslednjeg reda dole da ima 188 entiteta (učesnika u istraživanju u uzorku).

Q6 – netačno. Kramerov V je računat, a to je koeficijent korelacije za nominalne varijable, pa prema tome nema smislene negativne vrednosti.

Q7 – tačno. Zaista, Univerzitet je najčešća kategorija Nivoa obrazovanja u grupi učesnika u istraživanju koji su u braku. Možemo videti da je 83,7% učesnika u istraživanju koji su u braku navelo da ima Univerzitet kao obrazovanje.

Q8 – netačno. Možemo videti Nivo obrazovanja u ovoj tabeli ima 3 kategorije. Prema tome, to nije binarna varijabla, jer binarne varijable imaju po 2 kategorije.

Q9 – tačno. Možemo videti da u tabeli postoje samo dve kategorije Bračnog stanja, što znači da je to ovde binarna varijabla. Treba naglasiti da ova tvrdnja važi isključivo za ovaj konkretni slučaj koji je prikazan u ovoj konkretnoj tabeli i da se bračni status može osmisliti i tako da ima drugačiji broj kategorija od broja koji je predstavljen ovde.

Q10 – nepoznato. U tabeli nema podataka o polu učesnika u istraživanju.

PREPRINT

PREPRINT

Poglavlje 8. Vežbe – hajde da primenimo šta smo naučili u ovoj knjizi!

Apstrakt. Ovo poslednje poglavlje sadrži grupu od 5 vežbi, koje su isečci iz naučnih radova i koje pružaju čitaocu priliku da testira znanje koje je stekao u ovoj knjizi. Četiri vežbe obuhvataju statističke postupke koji su predstavljeni u ovoj knjizi, a psolednja, peta vežba, zahteva od čitaoca da primeni opšte statističke principe koji su predstavljeni u ovoj knjizi da interpretira statistike koji nisu obrađeni u knjizi.

Ključne reči: tumačenje statistike, vežbe.

Nakon što smo prošli sve sadržaje ove knjige, hajde da probamo da primenimo ono što smo u okviru nje prošli kroz seriju vežbi od kojih se sastoji ovo poglavlje. Kao i sa svim prethodnim vežbama, molimo da pogledate početak knjige za opšte uputstvo za rešavanje ovih vežbi. Naša sugestija je da prvo pročitate isečak, a onda tvrdnje i da onda date odgovore. Odgovore možete zapisati u kolonu Odgovor, a onda posle toga možete pogledati tačna rešenja i uporediti svoje odgovore s njima.

Vežba R. Opšta vežba

PREPRINT

Table 2

Correlation between compromise and aggressive conflict style of parents and adolescent

		Compromise				Aggression			
		father	a – f	mother	a – m	father	a – f	mother	a – m
Compromise	father	-	.663**	.389**	.425**	-.527**	-.446**	-.318**	-.296**
	a-f		-	.486**	.662**	-.352**	-.371**	-.303**	-.313**
	mother			-	.701**	-.222**	-.271**	-.561**	-.475**
	a-m				-	-.157**	-.215**	-.333**	-.307**
Aggression	father					-	.719**	.457**	.498**
	a-f						-	.560**	.689**
	mother							-	.788**
	a-m								-

*Note. ** p <0.01; a-f – adolescent – father conflict; a-m – adolescent – mother conflict;*

Podaci su prikupljeni od 514 adolescenata koji su odgovarali na upitnik o stilovima konflikta sa majkom i ocem. Kompromis (Compromise) i Agresija (Agression) su dva stila konflikta, a varijable otac (father) i majka (mother) koje su im dodate odnose se na to koliko adolescenti opažaju da njihov otac odnosno njihova majka primenjuju dati stil konflikta u konfliktima sa njima, dok se varijable a-f i a-m odnose na to koliko adolescent opaža da primenjuje dati stil konflikta u konfliktu sa majkom (a-m) odnosno ocem (a-f). Na primer, Compromise a-f se odnosi na to koliko adolescent opaža da koristi kompromisni stil konflikta u konfliktu sa svojim ocem.

Tabela je preštampana iz Konrad, S. Č. (2016). Parent-adolescent conflict style and conflict outcome: Age and gender differences 2. Psihologija, 49(3), 245–262. https://doi.org/10.2298/PSI1603245C na osnovu dozvole autora.

R	Tvrdnja:	Odgovor
R1.	Varijabla a-f je na nominalnom nivou merenja (za oba stila).	
R2.	Adolescenti koji koriste kompromisni stil konflikta sa svojim očevima teže da više koriste agresivni stil konflikta u konfliktima sa svojim majkama.	
R3.	Korelacija između Aggression father i Aggression mother je niža nego korelacija između Compromise mother i Aggression mother.	
R4.	Adolescenti teže da saopštavaju da je njihov stil konflikta sa roditeljem sličan stilu konflikta koji koristi taj roditelj.	
R5.	U istraživanju je učestvovalo više žena nego muškaraca.	
R6.	Adolescenti iz uzorka teže da češće koriste kompromisni stil konflikta (Compromise) nego agresivni stil konflikta (Aggression).	
R7.	Sve korelacije koje su ovde predstavljene su statistički značajne.	
R8.	Varijable koje se odnose na kompromisni stil konflikta (Compromise) su sve pozitivno korelirane sa varijablama koje se odnose na agresivni stil konflikta (Aggression).	
R9.	Koeficijenti korelacije ispod 0,3 nisu statistički značajni na ovom uzorku.	
R10.	Očevi (fathers) koriste agresivni stil konflikta (Aggression) češće nego majke (mothers).	

Vežba T. Opšta vežba.

PREPRINT

Table **Pearson correlations of the creativity tasks, age, and intelligence.**

p-value	Flexibility	Originality	RAT	CAQ	Age	Intelligence
Fluency	0.63**	0.90**	0.02	0.23**	−0.14*	0.18**
Flexibility		0.47**	0.02	0.14*	−0.16**	0.11
Originality			−0.01	0.20**	−0.15*	0.18**
RAT				−0.01	−0.19*	0.19**
CAQ					−0.11	−0.01
Age						−0.07

Notes:
* $p < 0.05$.
** $p < 0.01$.

Tabela preštampana iz: Han, W., Zhang, M., Feng, X., Gong, G., Peng, K., & Zhang, D. (2018). Genetic influences on creativity: An exploration of convergent and divergent thinking. PeerJ, 2018(7). https://doi.org/10.7717/PEERJ.5403, na osnovu dozvole autora.

T	Tvrdnja:	Odgovor
T1.	Barlovljev koeficijent je visoko značajan i za CAQ i za Age.	
T2.	Svih 5 varijabli koje su pobrojane se smanjuje sa starošću (Age).	
T3.	Originalnost (Originality) i Fluentnost (Fluency) su u visokoj korelaciji.	
T4.	Inteligencija (Intelligence) je u visokoj korelaciji sa RAT.	
T5.	Povezanost između Fluentnosti (Fluency) i CAQ je nemonotona.	
T6.	RAT nema statistički značajne korelacije sa drugim varijablama iz tabele.	
T7.	Osobe sa višim vrednostima na Inteligenciji (Intelligence) teže da imaju i više vrednosti na Fluentnosti (Fluency).	
T8.	Podaci koji su ovde prikazani su na ordinalnom nivou merenja.	
T9.	Nivo Fluentnosti (Fluency) u ovom uzorku je viši od nivoa Originalnosti (Originality).	
T10.	Nulta hipoteza koja je ovde testirana kaže da je koeficijent korelacije u populaciji jednak 0.	

Vežba U. Opšta vežba.

PREPRINT

Table 1

Means, standard deviations and metric properties of scales

	i	*M*	*SD*	*Md*	*Min*	*Max*	*K-S (p)*	*Sk*	*Ku*	α	*KMO*	*H1*	*H5*
HTQ Part I	64	23.44	8.32	24	1	53	.06 (.39)	.00	.90	.89	.93	.11	.39
HTQ Part I short	48	21.40	7.60	23	0	42	.09 (.06)	-.37	.45	.87	.92	.13	.47
HTQ Part III	6	1.38	1.33	1	0	6	.24 (.00)	1.09	1.08	.56	.71	.21	.70
HTQ Part IV Total	40	93.80	22.27	92	42	149	.07 (.28)	.03	-.69	.92	.96	.22	.55
HTQ Part IV PTSD	16	40.24	9.55	40	16	61	.06 (.36)	-.20	-.57	.82	.91	.22	.57
HTQ Part IV Functioning	24	53.56	14.31	53	24	93	.08 (.16)	.21	-.63	.89	.95	.25	.61
HSCL-25 Total	25	57.67	17.06	58.8	25	96	.06 (.44)	.03	-.86	.93	.98	.34	.68
HSCL-25 Anxiety	10	22.01	7.77	22	10	40	.08 (.14)	.14	-.75	.87	.96	.41	.79
HSCL-25 Depression	15	35.76	10.36	37	15	58	.08 (.15)	-.10	-.86	.88	.96	.38	1
BDI-II	21	23.04	12.58	22.5	0	54	.06 (.39)	.18	-.78	.91	.97	.32	.66

Note. Md – Median, *K-S* – Kolmogorov-Smirnov *D* statistic, *df* was 226 for all variables, *α* – Cronbach's alpha internal consistency coefficient, *KMO* – Kaiser-Mayer-Olkin measure of sampling adequacy, *H1* – average item intercorrelation, *H5* – Knežević-Momirović homogeneity measure (Knezević & Momirović, 1996)

M – aritmetička sredina. Md -medijana, K-S – Kolmogorov Smirnov D statistik, Sk – skjunes, Ku – kurtozis.

Imena varijable se u tvrdnjama koriste onako kako su napisana u tabeli.

Table reprinted from: Vukčević, M., Momirović, J., & Purić, D. (2016). Adaptation of Harvard Trauma Questionnaire for working with refugees and asylum seekers in Serbia. Psihologija, 49(3), 277–299. https://doi.org/10.2298/PSI1603277V, with the permission from authors.

U	Tvrdnja:	Odgovor
U1.	HTQ part III ima platikurtičnu distribuciju.	
U2.	Distribucije svih varijabli koje su ovde predstavljene su normalne osim HTQ Part III	
U3.	Hi kvadrat test je ovde korišćen za testiranje normalnosti distribucije.	
U4.	Gornja granica 95% intervala poverenja standardne devijacije varijable HTQ Part I bi bila veća od 30.	
U5.	Depression BDI-II i HTQ Part I su u visokoj korelaciji.	
U6.	Aritmetička sredina HTQ PART III je viša od medijane ove varijable.	
U7.	Distribucija Anxiety HSCL-25 je normalna u populaciji.	
U8.	HTQ Part III ima pozitivno asimetričnu distribuciju.	
U9.	50i percentil varijable HTQ Part I je manji od 30.	
U10.	Varijabla HTQ Part 1 short ima viši umereni raspon nego što je njen medijanski harmonik.	

A kao konačnu vežbu, hajde da probamo da rastumačimo tabelu koja prikazuje statistike koji nisu obrađeni u ovoj knjizi! Statistički postupci stalno evoluiraju i autori naučnih tekstova znaju često da uvedu u upotrebu nove statističke postupke i statističke mere koje nisu uobičajene ili uopšte ne postoje mimo primarnih publikacija iz statistike. Ali isti principi statistike rade i na njima.

Statistik primenjen u ovoj vežbi je količnik rizika (racio rizika, eng. hazard ratio), koji je veoma sličan količniku verovatnoća, tj. ods raciu, a koji je predstavljen u ovoj knjizi. Količnik rizika je količnik dve stope rizika. Pokazuje koliko puta je verovatnije da se događaj desi jednoj grupi nego drugoj unutar određenog intervala (vremena). Način na koji je ovaj statistik upotrebljen u ovom primeri je za poređenje verovatnoće javljanja dijabetesa u određenim godinama starosti u grupama sa različitim navedenim svojstvima. Jedna grupa ima osnovnu vrednost od 1, a količnici ostalih grupa pokazuju koliko se češće ili ređe ovaj događaj dešava datoj grupi. Nulta hipoteza na koje se p vrednosti odnose je da sve poređene grupe imaju jednaku verovatnoću ovog događaja, a to da li se ova hipoteza može prihvatiti možemo videti i posmatranjem intervala poverenja količnika rizika jedne grupe i videti da li se vrednosti drugih grupa (s kojima se poredi) nalaze unutar tog intervala. Viši količnici rizika u ovom slučaju pokazuju da postoji tendencija da dijabetes bude prvi put dijagnostifikovan na ranijem uzrastu u datoj grupi.

Grupa V. Opšta vežba.

PREPRINT

Table 2. Adjusted hazard ratios for the association between family history and age of onset for type 2 diabetes mellitus

		HR	CI
Family history of T2D	No	1.00	
	Yes	1.37	(1.20–1.56)***
Sex	Women	1.00	
	Men	1.11	(0.93–1.33)
BMI	Underweight	1.00	
	Normal	0.98	(0.54–1.79)
	Overweight	1.03	(0.56–1.89)
	Obese	1.20	(0.65–2.20)
Exercise	No	1.00	
	Yes	1.23	(1.08–1.40)**
Smoking status	Never	1.00	
	Past	0.87	(0.72–1.05)
	Current	1.62	(1.32–1.99)***
Drinking	No	1.00	
	Yes	1.32	(1.13–1.54)**

HR: Hazard ratio; CI: Confidence interval; T2D: Type 2 diabetes mellitus.
***p<0.001; **p<0.01; *p<0.05; ns: not significant.

Učesnici u istraživanju su pacijenti koji primaju tretmane za regulaciju nivoa šećera u krvi. Zavisna varijabla su uzrast tj. godine starosti koje je osoba imala kada joj je prvi put dijagnostifikovan dijabetes tipa 2 (age of onset for type 2 diabetes mellitus).
Značenje oznaka: Family history of T2D – porodična istorija dijabetesa tipa 2, Sex – pol, BMI – indeks telesne mase, Exercise – da li pacijent radi fizičke vežbe, da li vežba, Smoking status – da li puši duvan, Drinking – da li pije alkohol, No – ne, Yes – da, Underweight – telesna masa niža od normalne (BMI niži od normalnog), Normal – Normalan indeks telesne mase, Overweight – prekomerna telesna masa (BMI viši od normalnog), Obese – ekstremno visoka telesna masa (BMI mnogo viši od normalnog). Women – žene, Men – muškarci, Never – Nikad, Past – U prošlosti, Current – trenutno, HR – količnik rizik, hazard racio, CI – interval poverenja.
Tabela preštampana iz: Noh, J.-W., Hee Jung, J., Eun Park, J., Hwa Lee, J., Hee Sim, K., Park, J., Hee Kim, M., & Yoo, K.-B. (2018). The relationship between age of onset and risk factors including family history and life style in Korean population with type 2 diabetes mellitus. The Journal of Physical Therapy Science, 30, 201–206, na osnovu dozvole dobijene od the Society of Physical Therapy Science.

V	Tvrdnja:	Odgovor

V1.	Pacijentima sa porodičnom istorijom dijabetesa je u proseku dijagnostifikovan dijabetes na istom uzrtastu kao i onima koji nemaju porodičnu istoriju dijabetesa.	
V2.	U populaciji, muškarcima, u proseku, dijabetes bude prvi put dijagnostifikovan sa istim godinama starosti kao i ženama.	
V3.	Postoji tendencija da kod pacijenata koji imaju istoriju pijenja alkohola dijabetes bude dijagnostifikovan na kasnijem uzrastu nego kod osoba koje nemaju istoriju pijenja alkohola.	
V4.	Postoji tendencija da kod pacijenata koji su nekada pušili duvan dijabetes bude dijagnostifikovan na ranijem uzrastu nego kod ljudi koji nikada nisu pušili.	
V5.	Postoji tendencija da kod pacijenata koji trenutno puše duvan dijabetes bude dijagnostifikovan na ranijem uzrastu nego kod ljudi koji nisu nikada pušili duvan.	
V6.	Postoji tendencija da kod pacijenata koji kažu da vežbaju (fizički) dijabetes bude dijagnostifikovan na ranijem uzrastu nego kod ljudi koji kažu da ne vežbaju (da ne rade fizičke vežbe).	
V7.	Postoji tendencija da kod pacijenata sa normalnim indeksom telesne mase dijabetes bude dijagnostifikovan na ranijem uzrastu nego kod pacijenata koji imaju BMI niži od normalnog tj. telesnu masu nižu od normalne (Underweight).	
V8.	Prekomerna telesna masa uzrokuje dijabetes.	
V9.	Pijenje alkohola uzrokuje dijabetes.	
V10.	Pušenje duvana uzrokuje dijabetes.	

Pogledajmo sada rešenja:

R1 – netačno. Ima negativnih korelacija, prema tome ove varijable ne mogu da budu nominalne.

R2 – netačno. Korelacija između Compromise a-f i Aggression a-m je negativna, a ne pozitivna (-0,313). Ovo pokazuje da ovi adolescenti koristi agresivni konfliktni stil manje u konfliktu sa svojim majkama.

R3 – netačno. Prva korelacija je 0,457, a druga je -0,561. Prva je veći broj. Međutim, kada poredimo veličine koeficijenata korelacije ne uzimamo predznak u obzir, jer on samo pokazuje smer povezanosti, a ne njen intenzitet. Kad to tako postavimo 0,457 je manja korelacije od 0,561.

R4 – tačno. Možemo videti da su sve korelacije između stila konflikta koji koristi adolescent u konfliktu sa roditeljem i stila konflika koji primenjuje taj roditelj pozitivne. To su, u stvari, najviše korelacije u tabeli. Na primer, Compromise father i Compromise a-f su u korelaciji od 0,663, Compromise mother i Compromise a-m su u korelaciji od 0,701 itd.

PREPRINT

R5 – nepoznato. Nema podataka o distribuciji polova u tabeli.

R6 – tabela sadrži korelacije, nema podataka o nivoima izraženosti stilova konflikata.

R7 – tačno. Sve su označene sa ** i najveći broj njih je prilično visok.

R8 – netačno. Iz tabele možemo videti da su ove korelacije negativne.

R9 – netačno. Možemo videti da je onih nekoliko koeficijenata korelacije koji su niži od 0,3 takođe statistički značajno.

R10 – nepoznato. Podaci sadrže korelacije, mere koje pokazuju koje varijable teže da se menjaju zajedno. Ne znamo njihove intenzitete. Ako jedna varijabla ima vrlo nisku izraženost, a druga vrlo visoku, one i dalje mogu biti u visokoj korelaciji. Na korelaciju stepen izraženosti varijable ne utiče.

T1 – besmisleno. Ne postoji nikakav Barlovljev koeficijent koji bi mogao da bude statistički značajan.

T2 – netačno. Poslednja korelacija, sa CAQ, iako je takođe negativna, nije statistički značajna. To znači da prihvatamo nultu hipotezu i tretiramo je kao da je 0 u populaciji.

T3 – netačno. Da, možemo pročitati da je njihova korelacija 0,9.

T4 – netačno. Iako je njihova korelacija 0,19 i statistički je značajna, ona nije visoka.

T5 – netačno. Pirsonov koeficijent korelacije koji je ovde primenjen je koeficijent linearne korelacije, tako da, osim ako ga autori nisu primenili na način koji nije odgovarajući, odnos mora biti linearan (uputstvo za ove vežbe je da pretpostavimo da su predstavljeni statistički postupci valjano upotrebljeni).

T6 – netačno. Ima sa Age i Intelligence. Iako niske, statistički su značajne. Iako niske, statistički su značajne.

T7 – tačno. Korelacija je niska, ali statistički značajna.

T8 – netačno. Pirsonov koeficijent korelacije koji je ovde primenjen zahteva bar intervalni nivo merenja.

PREPRINT

T9 – nepoznato. Podaci u tabeli su korelacije, nema prikazanih podataka o nivou izraženosti (moguće da ih ima negde drugde u celom članku, ali ih nema u ovom isečku).

T10 – tačno. To je zaista nulta hipoteza kod računanja statističke značajnosti koeficijenta korelacije.

U1 – netačno. Kurtozis je 1,08, što ukazuje da je distribucija leptokurtična.

U2 – tačno. Rezultati KS testa su statistički značajni za HTQ Part III, a nisu ni za jednu drugu varijablu. Mada su za HTQ Part I short blizu kritičnog nivoa statističke značajnosti, p vrednost ove varijable od 0,06 je i dalje veća od 0,05. Kako u tabeli nije navedeno koja teorijska distribucija je korišćena za poređenje kod računanja K-S testa, pretpostavljamo da je u pitanju normalna distribucija, jer je to najčešće ktestirana distribucija u literaturi i očekivana distribucija individualnih razlika poput onih koje predstavljaju ove varijable.

U3 – netačno. Ne, upotrebljen je K-S test.

U4 – netačno. Standardna devijacija je 8,32, tako da standardna greška standardne devijacija mora da bude mnogo, mnogo manja, a 30 je više od 2 standardne devijacije iznad vrednosti standardne devijacije uzorka (gornja granica 95% intervala poverenja standardne devijacije je otprilike 2 standardne greške iznad standardne devijacije uzorka).

U5 – nepoznato. Nema u tabeli podataka o korelacijama između navedenih varijabli.

U6 – tačno. Da, aritmetička sredina je 1,38, a medijana je 1. A čak i da ovi nisu navedeni, mogli bi da zaključimo to iz činjenice da je skjunes pozitivan (što ukazuje da je aritmetička sredina viša od medijane).

U7 – tačno. Možemo videti da K-S test za ovu varijablu nije statistički značajan, p vrednost testa je 0,15. Dakle, prihvatamo nultu hipotezu da je distribucija normalna u populaciji.

U8 – tačno. Skjunes je pozitivan što ukazuje da je distribucija pozitivno asimetrična.

U9 – tačno. 50i percentil je medijana i to je 24, što je niže od 30.

U10 – besmisleno. Ne postoji nikakav „viši umereni raspon“, niti „medijanski harmonik“.

PREPRINT

V1 – netačno. Možemo videti da pacijenti sa porodičnom istorijom dijabetesa tipa 2 (T2D) imaju viši količnik rizika (HR) i da je razlika statistički značajna. Vrednost Ne grupe, koja je 1, je takođe van intervala poverenja Da grupe koji je između 1,20 i 1,56.

V2 – netačno. Možemo videti da razlika između količnika rizika (HR) nije statistički značajna, prema tome prihvatamo nultu hipotezu.

V3 – netačno. Količnik rizika za grupu koja pije alkohol je viši i ta razlika je statistički značajna. Međutim, količnik rizika je viši za grupu koja pije, što znači da je njima dijagnostifikovan na ranijem uzrastu.

V4 – netačno. Možemo videti da je njihov količnik rizika (HR) niži nego onih koji nikad nisu pušili. Međutim, razlika nije statistički značajna.

V5 – tačno. Možemo videti da je razlika statistički značajna i da je HR onih koji nikad nisu pušili van intervala poverenja količnika rizika trenutnih pušača. Razlika je takođe u smeru koji je naveden.

V6 – tačno. Možemo videti da je razlika statistički značajna i u navedenom smeru.

V7 – netačno. Razika nije statistički značajna i njihove HR vrednosti na uzorku su skoro iste.

V8 – nepoznato. Podaci pokazuju razlike na uzorku pri čemu grupa sa ekstremno visokim BMI (Obese) ima viši HR od ostalih kategorija, ali razlika nije statistički značajna u ovom istraživanju, a možemo videti da su statistici ostalih grupa unutar intervala poverenja grupe sa ekstremno visokim BMI (Obese), koji je takođe i veoma širok.

V9 – nepoznato. Moguće, ali iz ovih podataka ne možemo izvesti kauzalne zaključke.

V10 – nepoznato. Takođe moguće, ali situacija je ista kao kod V9.

PREPRINT

Reference

Akhshani, A., Akhavan, A., Mobaraki, A., Lim, S. C., & Hassan, Z. (2014). Pseudo random number generator based on quantum chaotic map. *Communications in Nonlinear Science and Numerical Simulation*, *19*(1), 101–111. https://doi.org/10.1016/j.cnsns.2013.06.017

Andelković, V., Vidanović, S., & Hedrih, V. (2012). Relationship between perceptions of children's needs importance, quality of life and family and work roles. *Ljetopis Socijalnog Rada*, *19*(2).

Atmanspacher, H. (2004). Quantum theory and consciousness: An overview with selected examples. *Discrete Dynamics in Nature and Society*, *1*, 51–73. https://doi.org/10.1155/S102602260440106X

Baldwin, T., & Bell, D. (1988). Phenomenology, Solipsism and Egocentric Thought. *Proceedings of the Aristotelian Society, Supplementary Volumes*, *62*, 27–43.

Byrne, D. (1998). *Complexity Theory and the Social Sciences: An Introduction*. Routledge.

Cen, J., Xue, S., & Groningen, H. (2014). Observation of the Optical and Spectral Characteristics of Ball Lightning. *Physical Review Letters*, *112*(035001), 035001-1-035001–035005. https://doi.org/10.1103/PhysRevLett.112.035001

Chaitin, G. (2005). *Epistemology as Information Theory: From Leibniz to* Ω *.

Chaitin, G. (1974). Information-Theoretic Computational Complexity. *IEEE Transactions on Information Theory*, 10–15. https://pdfs.semanticscholar.org/9a4f/05562a59842ce2107e2e858c8c8a7302329c.pdf

Chalmers, D. J. (1996). *The Conscious Mind: In Search of a Theory of Conscious Experience*. Oxford University Press.

Chomsky, N. (1959). A Review of B.F. Skinner's Verbal Behavior. *Language*, *35*(1), 26–58. http://cogprints.org/1148/1/chomsky.htm

PREPRINT

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences, 2nd Edition*.

Darlington, R. B. (1970). Is kurtosis really "peakedness?" *American Statistician*, *24*(2), 19–22. https://doi.org/10.1080/00031305.1970.10478885

de Moivre, A. (1756). *The Doctrine of Chances: Or, A Method of Calculating the Probabilities of Events and Play*. Chelsea Publishing Company.

Delahaye, J.-P., & Zenil, H. (2008). Towards a stable definition of Kolmogorov-Chaitin complexity. *Fundamenta Informaticae*, *XXI*, 1–15.

Desai, V. V, Deshmukh, V. B., & Rao, D. H. (2011). Pseudo random number generator using Elman neural network. *2011 IEEE Recent Advances in Intelligent Computational Systems*, 251–254. https://doi.org/10.1109/RAICS.2011.6069312

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in Psychology*, *0*, 781. https://doi.org/10.3389/FPSYG.2014.00781

Dienes, Z., & Mclatchie, N. (2018). Four reasons to prefer Bayesian analyses over significance testing. *Psychonomic Bulletin & Review*, *25*(1), 207–218. https://doi.org/10.3758/S13423-017-1266-Z

Djoric, S. (2021). *Socijalno odbacivanje i saradnja u situaciji socijalne dileme*. Filozofski fakultet.

Efron, B. (1979). Bootstrap Methods: Another Look at the Jackknife. *The Annals of Statistics*, *7*(1), 1–26. https://doi.org/10.1214/aos/1176344552

Efron, B. (1981). Nonparametric estimates of standard error: The jackknife, the bootstrap and other methods. *Biometrika*, *68*(3), 589–599.

Ferguson, C. J. (2015). "Everybody knows psychology is not a real science": Public perceptions of psychology and how we can improve our relationship with policymakers, the scientific community, and the general public. *American Psychologist*, *70*(6), 527–542. https://doi.org/10.1037/a0039405

Field, A. (2009). *Discovering Statistics Using SPSS*. SAGE Publications Ltd.

Flynn, J. (2007). *What is intelligence? Beyond the Flynn effect*. Cambridge University Press.

PREPRINT

Fodor, J. (1991). Methodological Solipsism Considered as a Research Strategy in Cognitive Psychology. In R. Boyd, P. Gasper, & J. D. Trout (Eds.), *The Philosophy of Science* (pp. 651–669). The MIT press. http://www.cog.brown.edu/courses/cg2000/Papers/Fodor1991PhilSci.pdf

Gao, S. (2008). A Quantum Theory of Consciousness. *Mind & Machines*, *18*, 39–52. https://doi.org/10.1007/s11023-007-9084-0

Gauvrit, N., Zenil, H., Soler-Toscano, F., Delahaye, J.-P., & Brugger, P. (2017). Human behavioral complexity peaks at age 25. *PLOS Computational Biology*, *13*(4), 1–14. https://doi.org/10.1371/journal.pcbi.1005408

Gentner, D., & Grudin, J. (1985). The Evolution of Mental Metaphors in Psychology : A 90-Year Retrospective. *American Psychologist*, *40*(2), 181–192.

Gigerenzer, G. (2004). Mindless statistics. *The Journal of Socio-Economics*, *33*, 587–606. https://doi.org/10.1016/j.socec.2004.09.033

Good, P. I. (2006). *Resampling Methods A Practical Guide to Data Analysis Third Edition*. Birkhauser. www.birkhauser.com

Hackinger, S., Kraaijenbrink, T., Xue, Y., Mezzavilla, M., Asan, Van Driem, G., Jobling, M. A., De Knijff, P., Tyler-Smith, C., & Ayub, Q. (2016). Wide distribution and altitude correlation of an archaic high-altitude-adaptive EPAS1 haplotype in the Himalayas. *Human Genetics*, *135*, 393–402. https://doi.org/10.1007/s00439-016-1641-2

Hahs-Vaughn, D. L., & Lomax, R. G. (2020). *An Introduction to Statistical Concepts - 4th Edition*. Routledge.

Haig, B. D. (2017). Tests of Statistical Significance Made Sound. *Educational and Psychological Measurement*, *77*(3), 489–506. https://doi.org/10.1177/0013164416667981

Hair, J. J., Hutt, T. G., Ringle, C., & Sarstedt, M. (2016). *A Primer on Partial Least Squares Structural Equation Modeling PLS-SEM*. SAGE Publications, Inc. https://www.amazon.de/gp/product/B01K0RVE0C/ref=as_li_tl?ie=UTF8&camp=1638&creative=6742&creativeASIN=B01K0RVE0C&linkCode=as2&tag=httpwwwsma079-21

PREPRINT

Hamburger, A., Hancheva, C., & Volkan, V. (Eds.). (2021). *Social Trauma - An Interdisciplinary Textbook*. Springer Nature Switzerland AG. https://doi.org/10.1007/978-3-030-47817-9

Harding, B., Tremblay, C., & Cousineau, D. (2014). Standard errors: A review and evaluation of standard error estimators using Monte Carlo simulations. *The Quantitative Methods for Psychology*, *10*(2), 107–123. https://doi.org/10.20982/tqmp.10.2.p107

Hedrih, Anđelka, & Hedrih, V. (2012). Attitudes and motives of potential sperm donors in Serbia. *Vojnosanitetski Pregled*, *69*(1). https://doi.org/10.2298/VSP1201049H

Hedrih, Andjelka. (2011). Mechanical models of the double DNA. *International Journal of Medical Engineering and Informatics*, *3*(4), 394–410. https://doi.org/10.1504/IJMEI.2011.044753

Hedrih, Andjelka. (2021a). Biological Oscillators. In K. Hedrih (Ed.), *Dynamics of hybrid systems of comples structures* (Non-period, pp. 425–466).

Hedrih, Andjelka. (2021b). Oscillations and stability of dynamics of hybrid biostructure. *European Physical Journal Special Topics*, *230*, 3573–3580. https://doi.org/https://doi.org/10.1140/EPJS/S11734-021-00240-8

Hedrih, Andjelka. (2014). Transition in oscillatory behavior in mouse oocyte and mouse embryo through oscillatory spherical net model of mouse zona pellucida. In J. an Awrejcewicz (Ed.), *Applied non-linear dynamical systems* (pp. 295–303). SPRINGER-VERLAG BERLIN. https://doi.org/10.1007/978-3-319-08266-0_21

Hedrih, Andjelka, & Banić, M. (2016). The effect of friction and impact angle on the spermatozoa - oocyte local contact dynamics. *Journal of Theoretical Biology*, *393*, 32–42.

Hedrih, Andjelka, & Hedrih, K. (2014). Modeling Double DNA Helix Main Chains of the Free and Forced Fractional Order Vibrations. In M. Lazarević & N. Mastorakis (Eds.), *Advanced Topics on Applications of Fractional Calculus on Control Problems, System Stability and Modeling* (pp. 145–183). WSEAS Press.

Hedrih, Andjelka, & Hedrih, K. (2016). Phenomenological mapping and dynamical absorptions in chain systems with multiple degrees of freedom. *Journal of Vibration and Control*, *1*(18–

PREPRINT

36). https://doi.org/10.1177/1077546314525984

Hedrih, Andjelka, Lazarević, M., & Mitrović-Jovanović, A. (2015). Influence of Sperm Impact Angle on Successful Fertilization Through mZP Oscillatory Spherical Net Model. *Computers in Biology and Medicine*, *59*, 19–29. https://doi.org/10.1016/j.compbiomed.2015.01.009

Hedrih, Andjelka, Najman, S., Hedrih, V., & Milošević-Đorđević, O. (2018). Structure of relations between the frequency of micronuclei in peripheral blood lymphocytes and age, gender, smoking habints and socio-demographic factors in south-east region of Serbia. *Facta Universitatis. Series Medicine and Biology.*, *20*(2), 47–54. https://doi.org/https://doi.org/10.22190/FUMB180102008H

Hedrih, Andjelka, & Ugrčić, M. (2012). Vibrational properties characterization of mouse embryo during microinjection. *Theoretical and Applied Mechanics*, *40*, 189–202.

Hedrih, K., & Hedrih, A. (2010). Eigen modes of the double DNA chain helix vibrations. *Journal of Theoretical and Applied Mechanics*, *48*(1), 219–231.

Hedrih, V. (2009). Profesionalna interesovanja i osobine ličnosti [Vocational interests and personality traits]. *Godišnjak Za Psihologiju*, *6*(8), 155–172. http://www.psihologijanis.rs/clanci/67.pdf

Hedrih, V. (2011). Provera konvergentne i diskriminativne validnosti analizom multiosobinske-multimetodske matrice na primeru PGI testa profesionalnih interesovanja zadatom uzorku iz Republike Makedonije [Assessment of convergent and discriminant validity through MTMM analy. *Primenjena Psihologija*, *4*, 393–408. http://primenjena.psihologija.ff.uns.ac.rs/index.php/pp/article/view/1138/1152

Hedrih, V. (Ed.). (2017a). *Work and Family Relations at the Beginning of the 21st Century*. Filozofski fakultet, Niš.

Hedrih, V. (2020). *Adapting Psychological Tests and Measurement Instruments for Cross-Cultural Research: An Introduction (1st Edition)*. Routledge, Taylor&Francis Group.

Hedrih, V. (2017b). Family and work relations at the beginning of the 21st century. In V. Hedrih (Ed.), *Work and Family relations at the beginning of the 21st century*. Filozfski fakultet,

PREPRINT

Niš.

Hedrih, V., Ristic, M., & Randjelovic, K. (2017). Vocational Interests of Recreational Athletes. *Facta Universitatis: Series Physical Education and Sports*, *15*(1), 37–48. https://doi.org/10.22190/FUPES1701037H

Hedrih, V., Šverko, I., & Pedović, I. (2018). Structure of vocational interests in Macedonia and Croatia - Evaluation of the spherical model. *Facta Universitatis, Series: Philosophy, Sociology, Psychology and History*, *17*(1), 19–36. https://doi.org/10.22190/FUPSPH1801019H

Hempel, C. G., & Oppenheim, P. (1948). Studies in the Logic of Explanation. *Philosophy of Science*, *15*(2), 135–175.

Hemphill, J. F. (2003). Interpreting the Magnitudes of Correlation Coefficients. *American Psychologist*, *58*(1), 78–80.

Hitchcock, C. R. (1975). Causal Explanation and Scientific Realism. *Erkenntnis*, *37*(2), 151–178.

Hitchcock, C., & Rédei, M. (2020). Reichenbach's Common Cause Principle. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy (Spring 2020 Edition)*. https://plato.stanford.edu/archives/spr2020/entries/physics-Rpcc/

Hoefer, C. (2016). Causal Determinism. In *The Stanford Encyclopedia of Philosophy* ((Spring 20). https://plato.stanford.edu/archives/spr2016/entries/determinism-causal

Holland, J. L. (1959). A Theory of Vocational Choice. *Journal of Counseling Psychology*, *6*(1).

Hunt, T., & Schooler, J. W. (2019). The Easy Part of the Hard Problem: A Resonance Theory of Consciousness. *Frontiers in Human Neuroscience*, *13*, 378. https://doi.org/10.3389/fnhum.2019.00378

Jackson, F. (1982). Epiphenomenal Qualities. *The Philosophical Quarterly*, *32*(127), 127–136.

James, W. (1884). *The Dilemma of Determinism*. Kessinger Publishing.

Kitchener, R. F. (1983). Developmental Explanations. *The Review of Methaphysics*, *36*(4), 791–817.

PREPRINT

Kruger, J., & Dunning, D. (1999). Unskilled and Unaware of It: How Difficulties in Recognizing One's Own Incompetence Lead to Inflated Self-Assessments. *Journal of Personality and Social Psychology*, *77*(6), 121–1134.

Kuhn, T. S. (1970). The structure of Scientific Revolutions. In *International Encyclopedia of Unified Science* (Issue 2). The University of Chicago Press.

Kwon, J., & Lee, H. (2020). Why travel prolongs happiness: Longitudinal analysis using a latent growth model. *Tourism Management*, *76*, 103944. https://doi.org/10.1016/j.tourman.2019.06.019

Lakatos, I. (1978). Science and Pseudoscience. *Philosophical Papers*, *1*, 1–7.

Lakić, S. (2019). Bayesov faktor: Opis i razlozi za upotrebu u psihološkim istraživanjima[Bayes Factor: What it is and Why it Should it be Used in Psychological Research]. *Godišnjak Za Psihologiju*, *16*, 39–58.

Lee, J. H. (Jay), & Ok, C. M. (2014). Understanding hotel employees' service sabotage: Emotional labor perspective based on conservation of resources theory. *International Journal of Hospitality Management*, *36*, 176–187. https://doi.org/10.1016/j.ijhm.2013.08.014

Liao, C., Wayne, S. J., Liden, R. C., & Meuser, J. D. (2017). Idiosyncratic deals and individual effectiveness: The moderating role of leader-member exchange differentiation. *Leadership Quarterly*, *28*(3), 438–450. https://doi.org/10.1016/j.leaqua.2016.10.014

Liu, J., Liang, Z., Luo, Y., Cao, L., Zhang, S., Wang, Y., & Yang, S. (2021). A hardware pseudo-random number generator using stochastic computing and logistic map. *Micromachines*, *12*(1), 1–12. https://doi.org/10.3390/mi12010031

Manson, S. M. (2001). Simplifying complexity: a review of complexity theory. *Geoforum*, 405–414.

Melamed, Y. Y., & Lin, M. (2021). Principle of Sufficient Reason. In *The Stanford Encyclopedia of Philosophy* (Summer 202). https://plato.stanford.edu/archives/sum2021/entries/sufficient-reason/

PREPRINT

Milas, G. (2009). *Istraživačke metode u psihologiji i drugim društvenim znanostima [Research methods in psychology and other social sciences]*. Naklada Slap.

Miller, R. (1974). The Jackknife - A Review. *Biometrika*, *61*(1), 1–15.

Nagel, E. (1961). *The Structure of Science: Problems in the Logic of Scientific Explanation*. Harcourt, Brace & World Inc.

Nicenboim, B., Schad, D., & Vasishth, S. (2021). *An Introduction to Bayesian Data Analysis for Cognitive Science*. Bookdown. https://vasishth.github.io/bayescogsci/book/

Okech, D., Howard, W., & Anarfi, J. K. (2018). Social Support, Dysfunctional Coping, and Community Reintegration as Predictors of PTSD Among Human Trafficking Survivors. *Behavioral Medicine*, *44*(3), 209–218. https://doi.org/10.1080/08964289.2018.1432553

Perezgonzalez, J. D. (2015). *Fisher, Neyman-Pearson or NHST? A tutorial for teaching data testing*. https://doi.org/10.3389/fpsyg.2015.00223

Physics and Beyond: "God does not play dice", What did Einstein mean? (2021). https://www.stmarys.ac.uk/news/2014/09/physics-beyond-god-play-dice-einstein-mean/

Plutchik, R. (1989). Measuring Emotions and their Derivatives. In *The Measurement of Emotions* (pp. 1–35). Elsevier. https://doi.org/10.1016/B978-0-12-558704-4.50007-4

Plutchik, R., & Kellerman, H. (1974). *Emotion Profile Index*. Western Psychological Services.

Popov, S., Sokić, J., & Stupar, D. (2021). Activity Matters: Physical Exercise and Stress Coping during COVID-19 State of Emergency. *Psihologija*, *54*(3), 307–322. https://doi.org/10.2298/psi200804002p

Popper, K. (1963). Science as Falsification. *Conjectures and Refutations*, *1*, 33–39.

Portugal, R. D., & Svaiter, B. F. (2011). Weber-Fechner Law and the Optimality of the Logarithmic Scale. *Mind&Machines*, *21*(1), 73–81. https://doi.org/10.1007/s11023-010-9221-z

Prekovic, S., Filipović Đurđević, D., Csifcsák, G., Šveljo, O., Stojković, O., Janković, M., Koprivšek, K., Covill, L. E., Lučić, M., Van Den Broeck, T., Helsen, C., Ceroni, F., Claessens, F., & Newbury, D. F. (2016). Multidisciplinary investigation links backward-

PREPRINT

speech trait and working memory through genetic mutation. *Scientific Reports*, *6*, 1–15. https://doi.org/10.1038/srep20369

Rakić-Bajić, G., & Hedrih, V. (2012). Prekomjerna upotreba interneta, zadovoljstvo životom i osobine ličnosti [Excessive use of the internet, life satisfaction and personality factors]. *Suvremena Psihologija*, *15*(1), 119–131.

Reichenbach, H. (1956). *The Direction of Time*. University of California Press.

Reyna, C. (2017). Scale Creation, Use, and Misuse : How Politics Undermines Measurement. In *The Politics of Social Psychology* (pp. 79–98). Psychology Press. https://doi.org/10.4324/9781315112619-6

Science. (2020). Wikipedia. https://en.wikipedia.org/wiki/Science

Skinner, B. F. (1990). Can Psychology Be a Science of Mind? *American Psychologist*, *45*(11), 1206–1210. https://doi.org/10.1037/0003-066X.45.11.1206

Sobotka, T., Brzozowska, Z., Muttarak, R., Zeman, K., & Lego, V. di. (2020). Age, gender and COVID-19 infections. *MedRxiv*, 2020.05.24.20111765. https://doi.org/10.1101/2020.05.24.20111765

Solms, M. (2021). *The Hidden Spring: A Journey to the Source of Consciousness* . Profile Books.

Stevens, S. S. (1946). On the Theory of Scales of Measurement. *Science*, *103*(2684), 677–680.

Sundet, J. M., Barlaug, D. G., & Torjussen, T. M. (2004). The end of the Flynn effect? A study of secular trends in mean intelligence test scores of Norwegian conscripts during half a century. *Intelligence*, *32*, 349–362. https://doi.org/10.1016/j.intell.2004.06.004

Świątkowski, W., & Dompnier, B. (2017). Replicability Crisis in Social Psychology: Looking at the Past to Find New Pathways for the Future. *International Review of Social Psychology*, *30*(1), 111–124. https://doi.org/10.5334/irsp.66

Taylor, R. (1990). Interpretation of the Correlation Coefficient: A Basic Review: *Journal of Diagnostic Medical Sonography*, *6*(1), 35–39. https://doi.org/10.1177/875647939000600106

PREPRINT

Teasdale, T. W., & Owen, D. R. (2005). A long-term rise and recent decline in intelligence test performance: The Flynn Effect in reverse. *Personality and Individual Differences*, *39*, 837–843. https://doi.org/10.1016/j.paid.2005.01.029

Tošić Radev, M., & Hedrih, V. (2017). Psychometric properties of the Multidimensional Jealousy Scale (MJS) on a Serbian sample. *Psihologija*, *50*(4), 521–534. https://doi.org/10.2298/PSI170121012T

Tošković, O. (2020). *Autostoperski vodič kroz statistiku: Uvod u primenjenu psihologiju [Hitchhikers guide through statistics: Introduction to Applied Statistics]*. Centar za primenjenu psihologiju.

Walker, T. C. (2010). The Perils of Paradigm Mentalities: Revisiting Kuhn, Lakatos, and Popper. *Perspectives on Politics*, *8*(2), 433–451. https://doi.org/10.1017/S1537592710001180

Wallace, A. F. C. (1959). Cultural determinants of response to hallucinatory experience. *A.M.A. Archives of General Psychiatry*, *1*(1), 58–69. https://doi.org/10.1001/archpsyc.1959.03590010074009

Watson, J. (1913). Psychology as the Behaviorist Views it. *Psychological Review*, *20*, 158–177. http://psychclassics.yorku.ca/Watson/views.htm

Westfall, P. H. (2014). Kurtosis as Peakedness, 1905–2014. R.I.P. *American Statistician*, *68*(3), 191–195. https://doi.org/10.1080/00031305.2014.917055

Wickstrom, G., & Bendix, T. (2000). The "Hawthorne effect"-what did the original Hawthorne studies actually show? In *Stand J Work Environ Health* (Vol. 26, Issue 4).

Wolfram, S. (1984). Universaility and Complexity in Cellular Automata. *Physica*, *10D*, 1–35.

Wood, H., & Neumann, F. (1931). Modified Mercalli Intensity Scale of 1931. *Bulletin of the Seismological Society of America*, *27*(4), 277–283. https://scits.stanford.edu/sites/g/files/sbiybj13751/f/277.full_.pdf

Woodward, J. (1980). Developmental explanation. *Synthese*, *44*(3), 443–466. https://doi.org/10.1007/BF00413471

Woodward, J., & Ross, L. (2021). Scientific explanation. In E. N. Zalta (Ed.), *The Stanford*

PREPRINT

Encyclopedia of Philosophy (Summer 2021 Edition).

Wray, K. B. (2011). Kuhn and the Discovery of Paradigms. *Philosophy of the Social Sciences*, *41*(3), 380–397. https://doi.org/10.1177/0048393109359778

PREPRINT