Dejan Pajić

PRIMENA TEHNIKA VIZUALIZACIJE U BAZIČNOJ STATISTICI

FILOZOFSKI FAKULTET · NOVI SAD ·

Novi Sad, 2020.

Dejan Pajić

PRIMENA TEHNIKA VIZUALIZACIJE U BAZIČNOJ STATISTICI

Novi Sad, 2020.

UNIVERZITET U NOVOM SADU
FILOZOFSKI FAKULTET
21000 Novi Sad
Dr Zorana Đinđića 2
www.ff.uns.ac.rs

Za izdavača
Prof. dr Ivana Živančević Sekeruš

Dejan Pajić
Primena tehnika vizualizacije u bazičnoj statistici

Recenzenti
dr Dragutin Ivanec
dr Oliver Tošković

Lektura i korektura
Ljiljana Ćuk

Tehnička priprema i dizajn korica
Dejan Pajić

ISBN
978-86-6065-582-2

Novi Sad, 2020.

SADRŽAJ

UVOD

Udžbenici iz oblasti statistike neretko počinju pokušajima autora da objasne otpor i animozitet studenata prema statistici, ili ubeđivanjem čitalaca da ne odustanu od čitanja nakon prvih nekoliko stranica. Škotski psiholog Gordon Rag u svojoj knjizi koju je šaljivo nazvao *Upotreba statistike: nežan uvod*, konstatuje da se pomenuti otpor javlja zato što statistiku uglavnom predstavljaju statističari čiji su mozgovi „drugačije umreženi" i koji najčešće stavljaju naglasak na matematičke osnove a ne na razumljivost, zanimljivost i interaktivnost (Rugg, 2007). Ovaj udžbenik je pokušaj da se osnovne statističke metode predstave na drugačiji, dinamičniji i čitljiviji način, prikladniji vremenu u kome živimo, odnosno komunikaciji koju karakteriše upotreba multimedijalnih (višekanalnih) sadržaja. Za početak, ponudićemo tri definicije statistike:

1. Statistika je oblast matematike koja se bavi podacima – njihovim prikupljanjem, sažimanjem, poređenjem, uopštavanjem i prikazivanjem, a zatim i zaključivanjem i predviđanjem na osnovu podataka.
2. Statistika je oblast **matematike** koja se bavi podacima – njihovim prikupljanjem, sažimanjem, poređenjem, uopštavanjem i prikazivanjem, a zatim i zaključivanjem i predviđanjem na osnovu podataka.
3. Statistika je oblast matematike koja se bavi **podacima** – njihovim prikupljanjem, sažimanjem, poređenjem, uopštavanjem i prikazivanjem, a zatim i **zaključivanjem** i predviđanjem na osnovu podataka.

Sve navedene definicije su semantički i sintaksički identične, ali vizuelno nisu. Samim tim, one čitaocu šalju drugačiju poruku. Percepciji studenta koji ima jasan pozitivan ili negativan stav prema matematici bliža je druga rečenica. Studentu koji želi da nauči i razume osnovne statističke metode i principe, trebalo bi da bude bliža treća. U ovom udžbeniku, analogno izgledu treće definicije, blago se prikriva matematika, a naglašavaju se podaci. Polazi se od stava da suština statistike, kao skupa metoda kojima se obrađuju podaci, nije njena matematička osnova, već njena praktična primena. Podaci se često

nazivaju naftom 21. veka i nalaze se svuda oko nas. Bez velike pretencioznosti, možemo reći da je samim tim i statistika svuda oko nas, jer svi ti podaci moraju da se pretvore u nešto upotrebljivo. Raspored artikala u supermarketima, grafikoni vremenske prognoze, procena bezbednosti automobila, efikasnost lekova, članci u kojima se preporučuje hrana koju treba ili ne treba da konzumiramo, izveštaji o stanju na berzama, preporuke koje nam se pojavljuju na Fejsbuk stranicama i predviđanja rezultata parlamentarnih izbora, imaju u svojoj pozadini podatke obrađene primenom različitih statističkih metoda.

Predlažemo da već na početku zaboravite česta (pogrešna) ubeđenja i aforizme vezane za statistiku. Verovatno najpoznatiji je onaj koji se pripisuje engleskom političaru iz 19. veka Bendžaminu Dizraeliju: „Postoje obične laži, odvratne laži i statistika". Jedan od ciljeva ovog udžbenika jeste da pokaže da statistika ne laže. Mogu da lažu samo ljudi. Osobe koje pogrešno primenjuju statističke metode zbog nedovoljne pismenosti, nenamerno obmanjuju javnost. To ih, naravno, ne oslobađa od odgovornosti. Sa druge strane, pristrasan izbor statističkih metoda, svesna manipulacija podacima i selektivna interpretacija rezultata obrade, predstavljaju ozbiljne etičke prekršaje u akademskoj zajednici. Ali čak ni tada nije moguće obmanuti statistički pismenu osobu koja poseduje barem osnovno znanje o tome kako se podaci obrađuju i kako treba da se interpretiraju. Ako četiri osobe imaju mesečnu platu od 25.000, a peta 250.000 dinara, onda prosečna plata tih pet osoba iznosi 70.000 dinara. To je tačan ali besmislen podatak, jer ne govori ništa o materijalnom statusu bilo koje od tih pet osoba. Ukoliko je određeni kandidat dobio 50% glasova na izbornima na kojima je izlaznost bila takođe 50%, onda je za njega glasalo (samo) 25% građana sa pravom glasa. Čak i kada bi se u nekom istraživanju utvrdilo da kod dece postoji povezanost između stepena agresivnosti i vremena provedenog u igranju „pucačkih" igara, to ne bi značilo da igre izazivaju agresivno ponašanje. I tako dalje. Statistika, dakle, uspešno odgovara samo na dobro postavljena pitanja i to samo onima koji žele i kompetentni su da je razumeju i adekvatno primene. Nažalost, to često nije slučaj. U tom kontekstu je ilustrativan aforizam škotskog pisca Endrjua Langa koji delom i objašnjava pomenuti negativan stav prema statistici: „Statistika se često koristi na način na koji pijanac koristi uličnu svetiljku: da bi se za nju pridržao, a ne da bi bolje video put ispred sebe".

Uvod ćemo završiti još jednim citatom. Karl Pirson, engleski statističar koga ćemo često pominjati u ovom udžbeniku, nazvao je statistiku *gramatikom nauke*. Statistički način razmišljanja je zaista bitan segment naučnog pristupa opisivanju i razumevanju fenomena koji nas okružuju, ali i važna komponenta

savremene akademske pismenosti. Ne u smislu poznavanja čitanja i pisanja, već u smislu posedovanja veština za 21. vek. Živimo u vremenu „seizmičkih promena" načina na koji komuniciramo i prenosimo znanje (Cope & Kalantzis, 2009). Svet koji se otkrivao kroz štampane knjige, sve više postaje svet koji spoznajemo kroz multimedijalne sadržaje prikazane na ekranima računara, televizora i mobilnih telefona. Postajemo preopterećeni podacima koji bi bez statističkih procedura zagušili naše kanale komunikacije. Rezultati pretrage koje nam nudi *Gugl*, za nas su „savladivi" upravo zato što im prethodi obrada i sažimanje milijardi podataka primenom PageRank algoritma (Brin & Page, 1998) kojim se vrši statistička procena relevantnosti sadržaja za dati upit. I ne samo to. Rezultati takve obrade podataka prikazuju se u formi teksta koji je obogaćen različitim vizuelnim karakteristikama. Kao i u definicijama iz našeg primera, različiti delovi teksta rezultata pretrage označeni su različitom bojom i debljinom, što nam govori o njihovoj važnosti, funkciji i kontekstu. Pri tome smo se veoma lako i brzo opismenili da ta vizuelna svojstva tumačimo na odgovarajući način i koristimo ih u cilju bržeg pronalaženja traženog podatka. Ovakva vizuelna komunikacija je ne samo efikasna već i univerzalna. Upravo je vizualizacija ono što omogućava da nam mozgovi budu „jednako umreženi", bez obzira na to da li smo statističari, naučnici, studenti ili umetnici. U sprezi sa statistikom, vizualizacija postaje neizostavna alatka koja unapređuje našu komunikaciju i razumevanje sveta oko nas.

Ovaj (statistički) udžbenik sastoji se iz tri poglavlja. U prvom će biti opisani osnovni statistički pojmovi i koncepti kroz prizmu vizualizacije, odnosno vizuelnog mišljenja, komunikacije i pismenosti. U drugom će naglasak biti na *deskriptivnoj statistici*, odnosno tehnikama pomoću kojih se podaci sažimaju i kojima se opisuju njihove raspodele, vrednosti oko kojih se grupišu i stepen njihove sličnosti. Treće poglavlje posvećeno je osnovnim tehnikama *induktivne statistike* ili statistike zaključivanja. Ove tehnike omogućavaju da se ode dalje od prostog opisivanja podataka i da se prave procene, donose zaključci i vrše predviđanja o fenomenima koji se analiziraju. Pošto zaključivanje obično podrazumeva uopštavanje od užeg i specifičnog (npr. stavovi grupe osoba) ka opštem i univerzalnom (npr. stavovi svih građana jedne države), ova grupa tehnika često se označava i terminom *inferencijalna statistika*. Naziv dolazi od engleskog termina *inference*, odnosno latinskog *inferre*, koji označavaju postupak izvođenja zaključaka na osnovu poznatih činjenica. U ovom udžbeniku biće opisane osnovne inferencijalne tehnike kojima se procenjuje značajnost razlika između grupa merenja i stepen povezanosti među pojavama.

Većina odeljaka sadrži interaktivni deo koji podrazumeva da čitalac učestvuje u kreiranju, menjanju i interpretaciji podataka kako bi bolje razumeo koncepte koji se objašnjavaju. U tom smislu, očekuje se da budete aktivni prilikom čitanja udžbenika i otkrivanju mogućnosti koje pruža vizualizacija podataka, ne samo u kontekstu statističke primene, već i u smislu drugačije forme komunikacije i prenošenja informacija.

POJAM, VRSTE I SVRHA VIZUALIZACIJE

Moderne tehnologije su u tolikoj meri utkane u naše svakodnevne aktivnosti da su u psihološkom smislu postale nevidljive. Potpuno je prirodno da automobilu ili kućnom bioskopu komande zadajemo glasom ili da dodirivanjem sličica na mobilnom telefonu naručujemo hranu. Mark Vajser, bivši rukovodilac istraživačkog centra PARC (engl. *Palo Alto Research Center*) kompanije Ziroks, opisao je ovaj fenomen terminom *sveprisutno računarstvo* (engl. *ubiquitous computing*) (Weiser, 1999). Upravo su istraživači ovog centra doprineli toj sveprisutnosti, jer su između ostalog zaslužni za kreiranje *WIMP* (*Windows Icons Menus Pointer*), grafičkog interfejsa bez koga se ne mogu zamisliti savremeni računarski operativni sistemi, i definisanje *WYSIWYG* (*What You See Is What You Get*), standarda koji nalaže da ono što vidimo u nekom programu za uređivanje teksta treba da izgleda isto kao i konačan proizvod koji se štampa. Očigledno je da se zbog sveprisutnog računarstva naša komunikacija preusmerila primarno na grafička okruženja u kojima vizuelni elementi kao što su slike, oblici, boje i animacije, bitno obogaćuju verbalni materijal kojim razmenjujemo informacije. Imajući u vidu činjenicu da svakodnevno primimo više podataka kroz vizuelne kanale nego kroz sva druga čula zajedno (Ware, 2004), upotreba vizuelnih elemenata često je način da se efikasnije sporazumemo i uspešnije izborimo sa velikom količinom informacija koje opterećuju naše kognitivne kapacitete i zagađuju našu komunikaciju. Iako mnogi ove promene vide kao negativne, činjenica je da upotrebom emotikona možemo da zamenimo cele rečenice i da brzo i jednostavno saopštimo drugoj osobi kako se osećamo, šta mislimo i šta radimo. Savremena komunikacija je nesporno postala multimedijalna i multimodalna, pa samim tim podrazumeva multimodalno mišljenje i multimodalnu pismenost.

1.1. Vizuelno mišljenje

Američki neuropsiholog Rodžer Speri, dobitnik Nobelove nagrade za fiziologiju ili medicinu, napisao je jednom prilikom da obrazovni sistem i nauka u savremenom društvu zapostavljaju desnu moždanu hemisferu (Sperry, 1973).

Ovaj pomalo čudan zaključak doneo je na osnovu istraživanja „podeljenog mozga" pacijenata kojima je zbog čestih epileptičnih napada hirurški prekinut *corpus callosum* – snop nervnih vlakana koji povezuje moždane hemisfere. Speri i njegovi saradnici uočili su da ovi pacijenti nisu u stanju da pročitaju reč koja im se prikazuje u levom vizuelnom polju, a da desnom rukom ne uspevaju da slože raznobojne kockice i formiraju zadati vizuelni šablon. Sa druge strane, pacijenti su uspešno rešavali verbalne zadatke kada su koristili desno oko, a neverbalne kada su koristili levu ruku. Speri je na osnovu toga zaključio da moždane hemisfere obrađuju informacije na drugačije načine, odnosno da su odgovorne za dve različite forme mišljenja – verbalno i neverbalno. Leva hemisfera je specijalizovana za jezičke funkcije i računske operacije, dok su centri desne hemisfere zaduženi za izvršenje neverbalnih zadataka kao što su crtanje, percepcija prostornih odnosa, prepoznavanje lica i formiranje vizuelnih modela objekata. Nema sumnje da je ovo otkriće bilo veoma značajno za oblast kognitivnih neuronauka, ali čini se da je ideja o postojanju dihotomije mišljenja i neurološke osnove različitih kognitivnih stilova postala jednako atraktivna, ako ne i atraktivnija, autorima naučno-popularne i kvazinaučne literature. I sam Speri je bio svestan da su rezultati njegovih istraživanja često pogrešno interpretirani (Puente, 2012). Tako se, na primer, pojavljuju tipologije prema kojima se osobe razlikuju po tome da li im je mozak pretežno „muški i racionalan" ili „ženski i intuitivan" ili, pak, po tome da li su sklonije „zapadnjačkom analitičkom" ili „istočnjačkom sintetičkom" načinu rezonovanja. Slično se desilo i sa gorepomenutom Sperijevom tvrdnjom da klasično obrazovanje, koje se u najvećoj meri bazira(lo) na čitanju, pisanju i aritmetici, potencira upotrebu leve moždane hemisfere. Kao odgovor na ovu kritiku, javlja se veliki broj pseudo-stručnih programa obrazovanja kojima se kod dece navodno razvija kreativnost i stimuliše razvoj desne hemisfere (Lindell & Kidd, 2011). Naučna zasnovanost ovakvih programa je diskutabilna, ali njihova popularnost ukazuje na potrebu na koju je Speri zapravo i želeo da skrene pažnju – klasičan model obrazovanja, koji se primarno oslanja na verbalnu komunikaciju i verbalno mišljenje, treba da se revidira i obogati. Ova potreba postaje sve izraženija u vremenu u kome se o mladim ljudima govori kao o „digitalnim urođenicima" koji se od najranijeg uzrasta susreću sa vizuelnim formama komunikacije, pa samim tim uče, razmišljaju i obrađuju informacije na potpuno drugačiji način u odnosu na starije generacije koje, u tom smislu, postaju „digitalni doseljenici". Stoga se može reći da današnji đaci i studenti više nisu osobe za koje je aktuelni model obrazovanja prvobitno osmišljen (Prensky, 2001).

Ako prihvatimo stanovište da je verbalna komunikacija dominantna forma razmene informacija i interakcije sa drugima, postavlja se pitanje u kojoj meri je to uticalo na procese mišljenja, odnosno način na koji rasuđujemo i učimo? Deo odgovora nalazi se u činjenici da je velikim brojem neuroloških i neuropsiholoških istraživanja potvrđen fenomen lokalizacije moždanih funkcija i postojanje morfološke i funkcionalne asimetrije mozga (Iaccino, 2014). Naime, asimetričnost moždanih hemisfera najizraženija je upravo u zonama koje su specijalizovane za razumevanje govora i jezičku produkciju, pa su tako kod većine osoba, posebno kod onih desnorukih, Brokaova govorna zona i *planum temporale* leve moždane hemisfere, nekoliko puta veći od analognih područja desne hemisfere. Jačanje ove asimetrije povezuje se sa porastom značaja verbalne komunikacije i sticanjem jezičkih kompetencija, kako u toku filogenetskog razvoja ljudske vrste, tako i u toku ontogenetskog razvoja čoveka kao jedinke (Toga & Thompson, 2003). U evolucionom i kulturnom smislu, dominacija govornog i pisanog jezika u odnosu na rane forme glasovne, gestikularne i slikovne komunikacije, dovela je i do promene u našem poimanju sveta. Modeli spoznaje, barem u zapadnoj civilizaciji, najvećim delom se oslanjaju na verbalne metafore i štampanu kulturu, za razliku od drevnih, urođeničkih, ali i mnogih savremenih istočnih civilizacija, čija se komunikacija dominanto bazira(la) na vizuelnim metaforama i slikovnim pismima (St. Clair, 2000). Naše razumevanje stvarnosti uglavnom se svodi na linearno dekodiranje nizova simbola pisanog jezika. Fenomene koji nas okružuju najčešće interpretiramo i opisujemo deterministički, određivanjem uzroka i posledica, slično kao što računarski programi izvršavaju određeni algoritam. Čak i kada posmatramo neko umetničko delo, najčešće ga pretačemo u tekst, odnosno priču koja ima pravolinijski tok. Stoga pojedini autori smatraju da je koncept spoznaje u zapadnim kulturama postao „agresivno lingvistički" (Stafford, 1998). U vremenu naglog razvoja računarske grafike i njenog sve većeg prisustva u oblastima telekomunikacija i obrazovanja, ova kritika postaje još aktuelnija. Ističe se važnost *vizuelne komunikacije* i *vizuelnog mišljenja* koji se često i neopravdano zanemaruju u učionicama i udžbenicima (Mathewson, 1999). Suštinu ovih zamerki sažeo je američki psiholog Stiven Rid u knjizi *Misliti vizuelno*: „Jezik je sjajna alatka za komunikaciju, ali je veoma precenjen kao alatka za razmišljanje". (Reed, 2013).

Termin *mišljenje* obično asocira na unutrašnji govor kao formu bezvučne komunikacije sa samim sobom. Taj unutrašnji govor nam olakšava planiranje aktivnosti, rešavanje problema i prisećanje događaja. Da li to znači da je

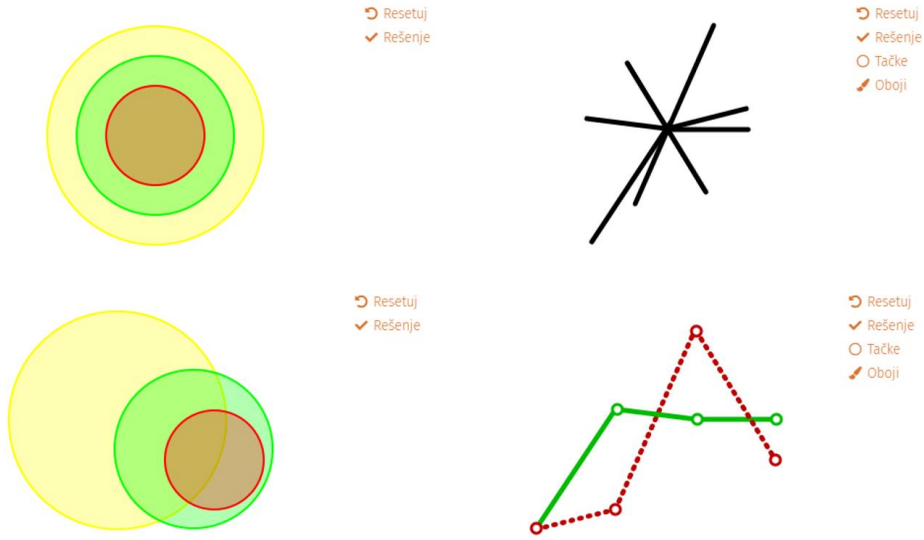
mišljenje isključivo verbalno? Novija istraživanja primenom funkcionalne magnetne rezonance (fMRI) pokazuju da je unutrašnji govor gotovo uvek praćen vizuelnim predstavama objekata i događaja o kojima razmišljamo (Amit, Hoeflin, Hamzah, & Fedorenko, 2017). Na primer, ukoliko dobijete zadatak da u sebi formulišete rečenicu koja sadrži termine *koncert* i *publika*, najverovatnije ćete zamisliti sliku nekog koncerta sa mnogo više detalja nego što nekoliko rečenica može da opiše. Pri tome će mentalni prikaz koji ste formirali najvećim delom biti zasnovan na vašem pređašnjem iskustvu. Moguće objašnjenje ovog fenomena može se pronaći u *teoriji dvojakog kodiranja* (TDK), koja ima značajne praktične implikacije upravo u oblasti obrazovanja i pedagoške psihologije (Paivio, 1986). Slično gorepomenutoj Sperijevoj dihotomiji, TDK pretpostavlja postojanje dva mentalna sistema, verbalnog i neverbalnog, koji funkcionišu kao složena asocijativna mreža jezičkih i vizuelnih predstava, odnosno mentalnih modela objekata. Ovo dvojako kodiranje ili paralelno mapiranje objekata i događaja u formi reči i slike, povećava efikasnost učenja novog materijala i prisećanje već naučenog u situacijama kada su, na primer, tekstualno gradivo ili usmena predavanja praćeni slikama, animacijama i odgovarajućim vizuelnim predstavama fenomena o kojima je reč (Clark & Paivio, 1991). Štaviše, u slučaju da tekst nije praćen slikom, ove vizuelne predstave ćemo sami formirati da bismo olakšali razumevanje i pamćenje pročitanog. Primer koji može da potkrepi ovu tvrdnju jeste eksperiment Stenfilda i Zvana (Stanfield & Zwaan, 2001) u kome su ispitanici dobijali zadatak da prepoznaju da li se određeni objekat pominje u rečenici koju su upravo pročitali. Nakon pročitane rečenice „Ona je zakucala ekser u pod“, ispitanici su brže reagovali na sliku eksera u vertikalnom položaju, za razliku od rečenice „Ona je zakucala ekser u zid“, nakon koje su brže prepoznavali sliku eksera u horizontalnom položaju. Dakle, čak i u toku verbalne komunikacije imamo pristup vizuelnim informacijama koje nisu proste predstave objekata, već mentalne simulacije njihovog položaja i odnosa sa drugim objektima i pojmovima. Ovo je u skladu sa rezultatima brojnih eksperimentalnih istraživanja koja ukazuju da proces mišljenja nije samo unutrašnja verbalizacija, već ponovno proživljavanje naših perceptivnih i motoričkih iskustava (Reed, 2013).

1.2. Vizuelna komunikacija

Najšire shvaćeno, vizuelna komunikacija predstavlja proces prenošenja informacija i ideja u formi koja se pretežno oslanja na vizuelnu percepciju i

upotrebu crteža, tipografije, animacija, grafičkog dizajna i ilustracija (Ryan, 2016). U vremenu sveprisutnog računarstva, upotreba *vizuelnog jezika* postaje gotovo neizbežna. Geografske mape, grafikoni, saobraćajna signalizacija, emotikoni i ikonice na ekranima mobilnih telefona, bitno nam olakšavaju razmenu i obradu velikih količina podataka. Grafički simboli koje viđamo na kućnim aparatima, etiketama na odeći, hemijskim preparatima, prehrambenim proizvodima, u uputstvima za upotrebu, u automobilima i restoranima, samo su dodatni korak u evoluciji potrebe prvog čoveka da bića, objekte i događaje iz svoje okoline predstavi na efikasan, jednostavan, samorazumljiv i univerzalan način. Postoje najmanje dve bitne prednosti predstavljanja objekata i fenomena odgovarajućim *vizuelnim kodovima* u odnosu na upotrebu reči. Prva je što se time omogućava paralelno ili holističko procesiranje informacija, za razliku od govora i teksta koji se obrađuju sekvencijalno, reč po reč. Uporedite, na primer, situaciju u kojoj vam neko diktira rezultate parlamentarnih izbora u procentima, nasuprot slici na kojoj visina stubića pokazuje odnos broja glasova različitih partija. Druga bitna prednost vizuelnog kodiranja je u tome što se bazira na univerzalno razumljivim simbolima koji, za razliku od verbalnih, najčešće nisu proizvoljni, jer imaju perceptivne karakteristike stvarnih objekata. Na primer, termini *pahulja*, *snowflake* i *flocon de neige* su proizvoljni nizovi slova koji označavaju isti pojam ili objekat, a nastali su konsenzusom osoba koje se sporazumevaju srpskim, engleskim, odnosno francuskim jezikom. Međutim, govornici različitih jezika će taj objekat vizuelno predstaviti na sličan, svima prepoznatljiv način. Štaviše, šestokraki zvezdasti oblik može da posluži ne samo kao vizuelni kod za pojam pahulje, već i kao simbol složenijih koncepata, npr. vremenskih prilika na grafikonima vremenske prognoze ili trenutne postavke klima uređaja. U tom slučaju govorimo o *vizuelnim metaforama* kojima se mogu predstaviti konkretni objekti, ali i apstraktni pojmovi. Vizuelne metafore mogu da budu *konkretne* i *asocijativne* (Lakoff & Núñez, 2000). Prve koriste stvarna iskustva u predstavljanju ideja, npr. upotreba novčića za objašnjavanje matematičkih operacija sabiranja i oduzimanja deci predškolskog uzrasta ili brisanje datoteke njenim prevlačenjem do ikonice u obliku kante za otpatke (engl. *recycle bin*). Sa druge strane, asocijativne metafore povezuju, predstavljaju i kodiraju koncepte iz jedne oblasti, konceptima iz druge oblasti znanja. Primer asocijativne metafore bio bi prikaz porasta profita neke kompanije iscrtavanjem rastuće linije ili opisivanje strukture atoma na osnovu analogije sa planetama Sunčevog sistema. Na taj način nam vizuelne metafore i vizuelno kodiranje bitno olakšavaju razumevanje (apstraktnih) koncepata i prenošenje informacija drugima.

Upotrebu vizuelnih metafora ilustrovaćemo **primerom geometrijskih slika** kojima ćemo predstaviti skupove objekata, njihove odnose i promene u toku vremena (Slika 1 – gore). Za početak, izmenite raspored krugova na levoj slici tako da formiraju vizuelnu metaforu grupa studenata koji su birali tri različita izborna predmeta. Najviše studenata odabralo je predmet A, a najmanje predmet C. Svi studenti koji su odabrali predmet C, odabrali su i predmet B. U grupi studenata koji su odabrali predmet B, više je onih koji su odabrali predmet A, nego onih koji nisu. Polovina studenata koji su odabrali predmet C, odabrala je predmet A, a druga polovina nije. Jedno od mogućih rešenja možete da vidite ako kliknete taster *Rešenje* (Slika 1 – dole).



Slika 1. Primer upotrebe geometrijskih slika kao vizuelnih metafora

Da li vam je lakše da na osnovu teksta formirate traženi vizuelni prikaz ili da na osnovu vizuelne metafore sastavite tekst koji opisuje grupe studenata?

Potrebno je da utvrdite da li postoje studenti koji su odabrali sva tri kursa. Da li ćete taj zaključak lakše doneti na osnovu prikazane slike ili na osnovu pročitanog teksta?

Da li vam je lakše da poredite veličine kružnica ili veličine njihovih preseka?

Na desnoj slici nalazi se šest linija raspoređenih u dvodimenzionalnom prostoru. Izmenite njihov raspored tako da prikazuje obim proizvodnje dve vrste voća u periodu od četiri godine. Na početku je proizvodnja obe vrste voća bila na istom nivou. U toku prve godine, proizvodnja prve vrste voća rasla je naglo, a proizvodnja druge vrste rasla je znatno sporije. Nakon prve godine, proizvodnja prve vrste blago opada, a proizvodnja druge veoma naglo raste. U toku treće godine, proizvodnja prve vrste voća stagnira, dok proizvodnja druge naglo opada. Kliknite taster *Rešenje* da biste prikazali traženi raspored linija (Slika 1 – dole). Linije prikazuju dva različita stepena pada, tri različita stepena porasta i jednu ravnu liniju kojom je predstavljena stagnacija. Obratite pažnju na to da četiri godine i dve vrste voća formiraju osam (zamišljenih) tačaka u dvodimenzionalnom prostoru koje su povezane pomoću šest prikazanih linija. Kliknite taster *Tačke* da biste prikazali po četiri vremenske tačke za obe vrste voća. Tačke koje prikazuju proizvodnju na početku prve godine analiziranog perioda se preklapaju, tako da je vidljiva samo jedna od njih.

> Šta predstavlja horizontalna a šta vertikalna dimenzija ove slike?
>
> Kako biste linije učinili jasnijim i informativnijim?
>
> Da li biste rekli da konačna slika prikazuje šest linija ili samo dve?
>
> Šta nedostaje slici da bi bila potpuno razumljiva i nekome ko nije pročitao ovaj tekst?

1.3. Vizuelna pismenost

U procesima razvoja čoveka kao vrste i kao jedinke, pojava vizuelne komunikacije i vizuelnog mišljenja prethodi sticanju jezičkih kompetencija. Izraženu filogenetsku osnovu neverbalne komunikacije potvrđuju rezultati istraživanja na senzorno depriviranim gluvim i slepim osobama, novorođenim bebama, blizancima i pripadnicima različitih kultura (Knapp, Hall, & Horgan, 2013). Nezavisno od društvenog konteksta, ljudi pokazuju tendenciju, ali imaju i sposobnost, da određene neverbalne signale koriste i interpretiraju na isti način. U smislu vizuelne komunikacije, ova sposobnost nije ograničena samo na gestikulaciju, facijalnu ekspresiju i prepoznavanje bazičnih emotivnih stanja, već i na opažanje i tumačenje oblika i boja. Na primer, široko je prihvaćeno

stanovište da evolucija trihromatskog sistema percepcije boja kod čoveka, odnosno postojanje receptora za crvenu, zelenu i plavu boju, može da se poveže sa vitalnom potrebom prvih ljudi da razlikuju crvenkaste plodove od okolnog zelenila ili da prepoznaju emotivna stanja i seksualne signale na osnovu suptilnih promena boje kože drugih osoba (Hurlbert & Ling, 2012). Slična filogenetska osnova uočljiva je i u načinu na koji opažamo formu. U toku filogenetskog razvoja čoveka, oblici su postali vrsta koda kojim se prenose informacije o prikladnosti objekta za određenu upotrebu, odnosno o načinu na koji treba da reagujemo kada ga uočimo (Singh & Hoffman, 2013). Čak i deca na veoma ranom uzrastu brže detektuju ovakve evolutivno bitne oblike (npr. slike zmije ili pauka) od onih nebitnih (npr. slike cveta ili žabe) (LoBue & DeLoache, 2008). Ovaj koncept poznat je u ekološkoj psihologiji kao *afordansa* (engl. *affordance*). Termin je prvi upotrebio američki psiholog Džerom Gibson (Gibson, 1977) da bi označio svojstvo okruženja koje u interakciji sa čovekom ili životinjom omogućava ili inicira izvršenje neke akcije. Iako je to svojstvo u osnovi vizuelno, afordansa prema Gibsonu ne podrazumeva nužno opažanje objekta, odnosno okruženja, jer ona postoji nezavisno od percepcije. Sa druge strane, američki psiholog *Donald Norman* stavio je naglasak upravo na vizuelnu komunikaciju funkcionalnosti objekta, upotrebivši termin ***opažena afordansa*** (Norman, 1988). Ovako shvaćena, opažena afordansa postaje ključni element savremenog dizajna jer, za razliku od objektivne afordanse prema Gibsonu, naglašava subjektivnost, tj. važnost ispravne percepcije mogućnosti koje pruža određeni objekat (Slika 2). Na primer, svaki deo internet stranice prikazane na ekranu računara može da priušti akciju kliktanja pokazivačem miša, ali samo delovi teksta drugačije boje imaju opaženu afordansu koja nam omogućava da neki objekat prepoznamo kao vezu za prelazak na drugu veb stranicu. Međutim, ispravno tumačenje ovih vizuelnih karakteristika okruženja podrazumeva određene konvencije koje bi dizajneri trebalo da poštuju, a korisnici da poznaju. Na primer, klikom na krstić u gornjem desnom uglu ekrana zatvara se aktivni prozor, širenjem prstiju na ekranu mobilnog telefona uvećava se slika, pritiskom na strelice ili oznake + i - na daljinskom upravljaču povećava se ili smanjuje jačina zvuka i tako dalje. Očigledno je da vizuelna komunikacija ima jaku filogenetsku osnovu i bazira se na našim vizuelnim *sposobnostima*, ali u digitalnom dobu ona podrazumeva i usvajanje određenih *znanja* i *veština*. Stoga se sve češće govori o *vizuelnoj pismenosti* kao važnoj kompetenciji modernog čoveka koja se uči i unapređuje.

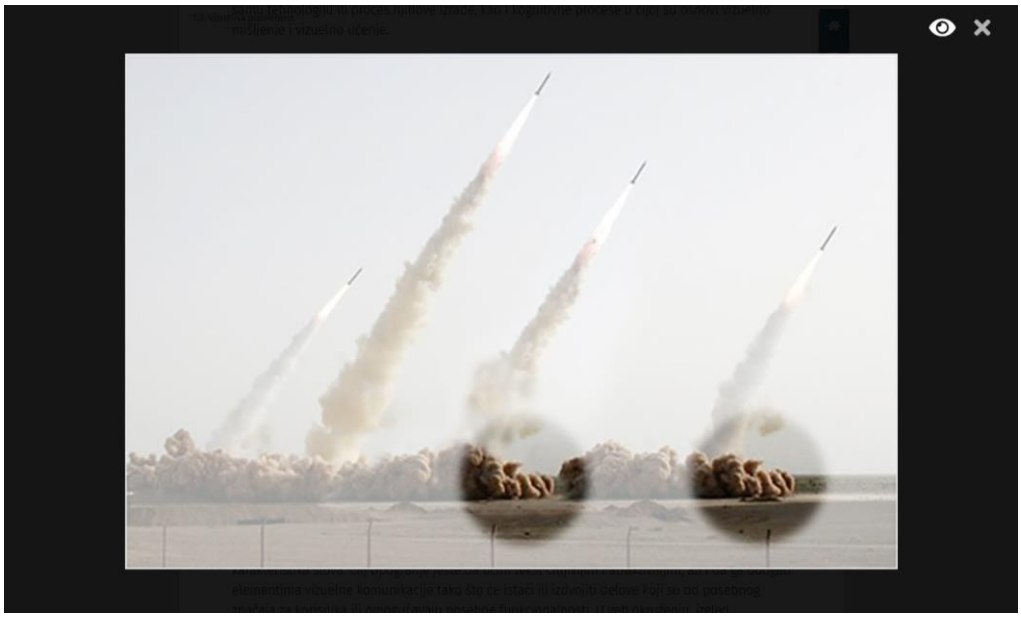
Slika 2. Primer loše dizajnirane opažene afordanse objekta (Izvor: Satu Kyröläinen)

Vizuelna pismenost je složen i još uvek nedovoljno operacionalizovan koncept koji obuhvata najmanje tri bitna elementa: *vizuelnu komunikaciju (retoriku)*, kao veštinu da se ideje prenesu u vizuelnoj formi, *vizuelno mišljenje*, kao sposobnost da se formiraju mentalni modeli i vizuelne predstave pojmova i njihovih relacija, i *vizuelno učenje*, kao kompetenciju koja omogućava uspešno interpretiranje, razumevanje i usvajanje informacija predstavljenih u vizuelnoj formi (Trumbo, 1999). Razlog zbog koga vizuelna pismenost nema solidniju teorijsku osnovu je njena multidimenzionalnost i multidisciplinarnost. Iz aspekta umetnosti, vizuelno pismena osoba sposobna je da razume i kritički analizira umetničko delo (npr. kompoziciju slike). Za dizajnera je vizuelna pismenost adekvatna upotreba osnovnih principa i elemenata dobre forme (npr. kontrasta i teksture). U učionici, nastavnik mora da poseduje određeni nivo vizuelne pismenosti da bi odabrao i kreirao adekvatnu simulaciju koncepta koji želi da opiše (npr. vizuelni prikaz magnetnog polja). Sa druge strane, učenici treba da umeju da povežu konkretnu grafičku formu i apstraktni pojam koji ona predstavlja (npr. presek i unija skupova koji su ilustrovani u prethodnom odeljku). Lekari se obučavaju za postavljanje dijagnoza na osnovu senčenja i boja na slici (npr. RTG ili PET snimci). Posao inženjera je nezamisliv bez vizuelnog mišljenja i prostornog rezonovanja (npr. povezivanje ortografskih i izometrijskih projekcija). Naučnici bi trebalo da budu kompetentni da predstave i interpretiraju podatke koje su prikupili u istraživanjima i prikazali u vizuelnoj

formi (npr. različite vrste grafikona). Verbalna pismenost nije manje važna niti prevaziđena, ali ona je u tolikoj meri protkana drugim i drugačijim modalitetima komunikacije da i sama menja prirodu i postaje deo šireg koncepta *multipismenosti* (Cope & Kalantzis, 2009). Očigledan primer je pregledanje internet stranica koje je, zbog upotrebe multimedijalnih sadržaja, različitih fontova, animacija i hiperlinkova, sličnije posmatranju slika nego čitanju teksta. Vizuelna pismenost, u tom smislu, postaje bitna komponenta pismenosti za 21. vek. Ona podrazumeva veštine razumevanja, korišćenja i kreiranja onoga što se na engleskom jeziku naziva *visuals*, a na srpskom bi obuhvatilo različite predstave koje se mogu videti u pravom smislu te reči ili u smislu zamišljanja, odnosno imaginacije: slike, fotografije, filmovi, stripovi, mape, šeme, dijagrami, grafikoni, animacije, ikonice, emotikoni, gestikulacije, mentalni modeli itd. U ovom udžbeniku ćemo kao prigodan, ali ne i potpuno prikladan prevod, koristiti termin *vizualizacija* da bismo označili sve navedene grafičke forme i sredstva vizuelne komunikacije, samu tehnologiju ili proces njihove izrade, kao i kognitivne procese u čijoj su osnovi vizuelno mišljenje i vizuelno učenje.

1.3.1. Različiti aspekti vizuelne pismenosti

Vizuelna pismenost obuhvata veoma širok spektar znanja i veština potrebnih za uspešno „čitanje" (tumačenje) i „pisanje" (kreiranje) vizuelnih sadržaja. Primer takvog kreiranja je sve učestalije manipulisanje slikama, posebno onima koje se dele na društvenim mrežama. Koristeći neologizam koji takođe upućuje na snažnu protkanost savremene komunikacije vizuelnim elementima, možemo reći da *fotošopiranje* fotografija u cilju obmanjivanja drugih postaje ozbiljan psihološki problem koji među mladim ljudima može biti povezan sa iskrivljenom slikom o sopstvenom telu i sa poremećajima u ishrani (McLean, Paxton, Wertheim, & Masters, 2015). Ipak, kao ilustraciju ovog aspekta vizuelne pismenosti ćemo upotrebiti manje dramatičan primer. U pitanju je **fotografija lansiranja iranskih krstarećih raketa** (Slika 3) na kojoj treba da uočite elemente koji ukazuju da prikazana scena ne odgovara stvarnosti. Fotografija je izmenjena kako bi se prikrila činjenica da jedna od četiri rakete nije uspešno lansirana. Ukoliko ne uspete da uočite sve izmenjene elemente, kliknite više puta na ikonicu oka u gornjem desnom uglu prozora.



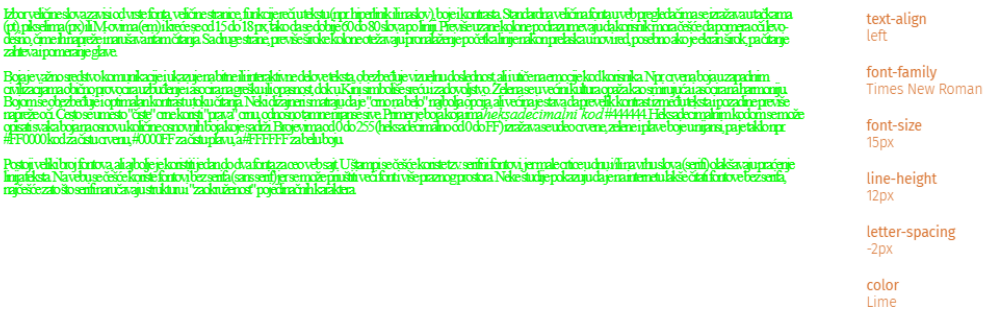
Slika 3. Izmenjena fotografija lansiranja iranskih krstarećih raketa
(Izvor: Agence France-Presse)

Na osnovu čega možete da zaključite da su neki od elemenata na slici naknadno izmenjeni?

Kako biste u obrazloženju odgovora na prethodno pitanje upotrebili termin „verovatnoća“?

Veoma bitan element vizuelnog jezika i vizuelne pismenosti iz ugla dizajna jeste *tipografija*. To je umetnost ili delatnost koja se bavi estetskim aspektima štampanog i elektronskog teksta, organizacijom i izborom fontova i uređivanjem oblika, boje, veličine, pozicije i drugih karakteristika slova. Cilj tipografije jeste da učini tekst čitljivijim i atraktivnijim, ali i da ga obogati elementima vizuelne komunikacije tako što će istaći ili izdvojiti delove koji su od posebnog značaja za korisnika ili omogućavaju posebne funkcionalnosti. U veb okruženju, izgled dokumenata i teksta definiše se upotrebom *CSS jezika* (engl. *Cascading Style Sheets*). Na taj način se sadržina teksta u potpunosti odvaja od njegovog izgleda, što olakšava dizajniranje i prilagođavanje veb stranica različitim grupama korisnika. Svojstva objekata prikazanih na veb stranici označavaju se rečima engleskog jezika, npr. *font-size* (veličina teksta), *line-height* (visina reda teksta), *color* (boja teksta), *letter-spacing* (razmak

između slova) i tako dalje. Unapredite izgled i čitljivost **teksta u kome se opisuju osnovni principi tipografije** izmenom njegovih CSS stilova, odnosno različitih tipografskih svojstava (Slika 4). Čak i ako niste pohađali kurs iz oblasti veb dizajna, vrlo je verovatno da ćete odabrati svojstva teksta koja su u skladu sa univerzalnim estetskim kriterijumima i čine test lakšim i prijatnijim za čitanje.



Slika 4. Primer uticaja css stilova na čitljivost teksta

> Na koji način tipografija utiče na čitljivost teksta?
>
> Uporedite svoje vizuelno rešenje sa stilovima koje je postavio neko drugi ili sa tekstovima koje srećete na internetu? U kojim aspektima tipografije postoji slaganje a u kojima ne?

1.3.1.1. Piktogrami i piktografici

Poseban značaj za vizuelnu komunikaciju imaju *piktogrami*, univerzalni simboli kojima se na osnovu sličnosti ili analogije sa fizičkim objektima, veoma lako prenose informacije i poruke. Zahvaljujući piktogramima, u bilo kojoj državi lako ćemo odrediti u koji toalet treba da uđemo, kada treba da vežemo pojas u avionu ili na kojim mestima ne smemo da koristimo mobilni telefon. Ovaj poslednji primer pokazuje da se piktogramima često prenose šira i složenija značenja, jer se gubi njihova sličnost sa stvarnim objektima ili se kombinuju sa apstraktnim simbolima, kao što je crveni precrtani krug koji simboliše zabranu. Tada se obično govori o *ideogramima*, simbolima koji označavaju apstraktne pojmove, koncepte ili ideje. U tom smislu, vizuelna pismenost nije samo dopuna uobičajenim veštinama pisanja i čitanja, već često i njihova efikasnija alternativa, pa čak i jedini način da se informacije prenesu specifičnim kategorijama

korisnika. Prikladnu ilustraciju predstavljaju vizuelna uputstva za korišćenje lekova pomoću kojih se nepismenim osobama može objasniti kako da koriste lek, kako da ga skladište i kakvi su efekti njegove primene (Dowse & Ehlers, 2001). Štaviše, korisnici ne moraju ni da budu upoznati sa poreklom i karakteristikama fizičkih objekata po čijoj analogiji su nastali piktogrami i ideogrami. Očigledan primer je činjenica da većina osoba koje koriste mobilne telefone i računare nikada nije imala iskustva sa starim telefonskim aparatima sa slušalicama, disketama za snimanje podataka ili slanjem pisama klasičnom poštom. Čak i pored toga, ikonice koje simbolišu započinjanje poziva (slušalica), čuvanje promena u datoteci (disketa) ili slanje imejla (pismo), razumljive su svima. Njihov dizajn se najverovatnije neće ni menjati, jer je postao deo standardnog vizuelnog jezika kojim uspešno komuniciramo u vremenu sveprisutnog računarstva. Ne samo da je ovaj jezik veoma lak za učenje i pamćenje, već je i univerzalan, jer o izgledu ikonica postoji međunarodni konsenzus. Međunarodna organizacija za standardizaciju (engl. *ISO*), na primer, propisala je izgled **preko 6.700 piktograma** koji se koriste u svim zemljama sveta kao grafički simboli različitih objekata, aktivnosti, obaveštenja, upozorenja i svojstava (Slika 5).

Slika 5. Primeri grafičkih simbola (piktograma) koje propisuje Međunarodna organizacija za standardizaciju (ISO)

U narednom primeru ćemo analizirati **efikasnost piktograma u prenošenju informacija**, ali i vašu vizuelnu pismenost u smislu tačne interpretacije njihovog značenja (Slika 6). Piktogrami su preuzeti iz besplatnog skupa ikonica popularnog **Font Awesome** fonta koji se često koristi prilikom dizajniranja veb stranica. U pitanju je kratak upitnik koji se sastoji od pet pitanja na koja treba da odgovorite tako što ćete kliknuti neku od ponuđenih opcija,

odnosno ikonica. Kada odgovorite na pitanje, vaš odgovor se beleži, brojač pored odgovora se povećava za jedan, a u redu u kome se nalazi opcija koju ste odabrali, pojavljuje se piktogram koji simboliše broj osoba koje su dale isti odgovor na pitanje. Samo vaš prvi odgovor se beleži u bazu radi kasnije uporedne analize. Zamislite da ste u ulozi drugih korisnika udžbenika koji pristupaju istoj stranici i daju svoje odgovore na dato pitanje. Nastavite da klikćete ikonice, simulirajući tačne i netačne odgovore većeg broja korisnika. Posmatrajte kako se menja broj i izgled piktograma sa desne strane svakog od odgovora. Jedan piktogram simboliše jednu osobu, ali samo dok se red ne popuni, nakon čega se značenje ikonice menja da bi se omogućilo grafičko prikazivanje brojeva većih od 19, a potom i 38, 95 i tako dalje. Sa desne strane tabele nalazi se *legenda* kojom se objašnjava značenje pojedinačne ikonice korisnika. Ukoliko želite da vidite dosadašnje odgovore čitalaca udžbenika, kliknite taster *Online podaci*. Da biste zaključili koji odgovor je tačan, kliknite taster *Tačan odgovor*. Obratite pažnju na činjenicu da time ne označavate konkretnu ikonicu, već prikazujete simulirane učestalosti različitih odgovora koje će vam ukazati na to koji je od njih tačan. Pretpostavićemo da je to onaj odgovor koji je najčešći, odnosno opcija koju je odabralo najviše (zamišljenih) ispitanika. Ilustrovani grafički prikaz numeričkih vrednosti pomoću odgovarajućeg broja piktograma, naziva se *piktografik* i omogućava lakše poređenje broja elemenata u različitim kategorijama. Da biste prešli na naredno pitanje, kliknite strelicu u gornjem desnom uglu okvira. Za svako pitanje možete da upotrebite opcije *Online odgovori* i *Tačan odgovor*.

Slika 6. Primer piktografika kojim su vizualizovani odgovori 543 ispitanika

Da li je značenje piktograma uvek nedvosmisleno?

Zbog čega u petom pitanju nije očigledno koji odgovor je tačan?

Šta označava upitnik a šta zelena štrikla u gornjem desnom uglu okvira?

Ako je na pitanja u upitniku odgovorilo 100 ispitanika koji su za svaki tačan odgovor dobijali po jedan poen, kako bi izgledao piktografik kojim se vizualizuje njihov uspeh izražen zbirom bodova?

Koje je značenje poruke ispisane na početnoj veb stranici ovog udžbenika?

1.4. Karta, mapa, dijagram, grafik, infografik

Istorija vizualizacije je delom i istorija umetnosti. Crteži pronađeni u pećini **Lasko**, koja se često naziva i praistorijskom Sikstinskom kapelom, nisu samo izuzetan primer paleolitskog slikarstva već i svojevrsno vizuelno uputstvo za izvođenje lova i magijskih rituala. Raspored crteža u pećini **Altamira** na severu Španije, za koju je Pablo Pikaso izjavio: „Posle Altamire, sve je dekadencija", ukazuje na postojanje narativnog toka i ima obeležja sintakse pisanog jezika (Wildgen, 2004). Slike stvorene urezivanjem i dubljenjem površine kamena (tzv. petroglifi) na **Pisanoj steni**, crteži na antičkoj grnčariji i ideogrami u egipatskim grobnicama, još su neki primeri beleženja i prenošenja informacija u vizuelnoj formi. Međutim, tek je pojava papirusa oko 3.000. godine pre nove ere omogućila lakše i efikasnije zapisivanje, razmenu i širenje znanja. Značaj ovog pronalaska ogleda se i u jeziku, odnosno u upotrebi termina vezanih za (vizuelnu) komunikaciju. *Papir* u srpskom jeziku, *paper* u engleskom i *papier* u francuskom jeziku, vode poreklo od grčke reči πᾶπῦρος, ali označavaju podlogu za pisanje i crtanje koja je prvi put upotrebljena tek u Kini tri hiljade godina posle pronalaska papirusa. Reči kao što su *Biblija*, *biblioteka* i *bibliografija* u svom korenu imaju grčku imenicu βῐβλῐ́ον kojom je označavana traka papirusa. Srpsko *karta* i *hartija*, englesko *chart* i italijansko *carta*, vode poreklo od grčkog χαρτί ili χάρτης što je bio naziv za list papirusa na kome je nešto zabeleženo, najčešće mapa ili vizuelni prikaz informacije. Ovi poslednji primeri govore o tesnoj vezi ranih oblika pisane komunikacije i procesa vizualizacije. U literaturi se upravo karte ili mape najčešće navode kao preteče današnje vizualizacije informacija. Gotovo tri metra duga **Torinska mapa na papirusu** iz 1160. godine p. n. e. veoma precizno i u više boja prikazuje okolinu Tebe u Egiptu sa ucrtanim podacima o nalazištima ruda i geološkim karakteristikama tog područja. Pergament iz 13. veka poznat kao **Tabula Peutingeriana** je replika starorimske mape Mediterana, Bliskog Istoka

i dela Azije iz 4. veka p. n. e. na kojoj su linijama i površinama raznih boja označeni putevi, rečni tokovi i planinski masivi, a ikonicama, tj. piktogramima, najvažnija naselja. Značajan pomak ka vizualizaciji informacija u modernom smislu te reči, predstavlja **Ptolomejeva mapa** iz 2. veka na kojoj su prvi put linijama označene geografske dužine i širine, odnosno fenomeni koji fizički ne postoje ili su nevidljivi. Na taj način karta fizičkih objekata postaje *dijagram* apstraktnih informacija.

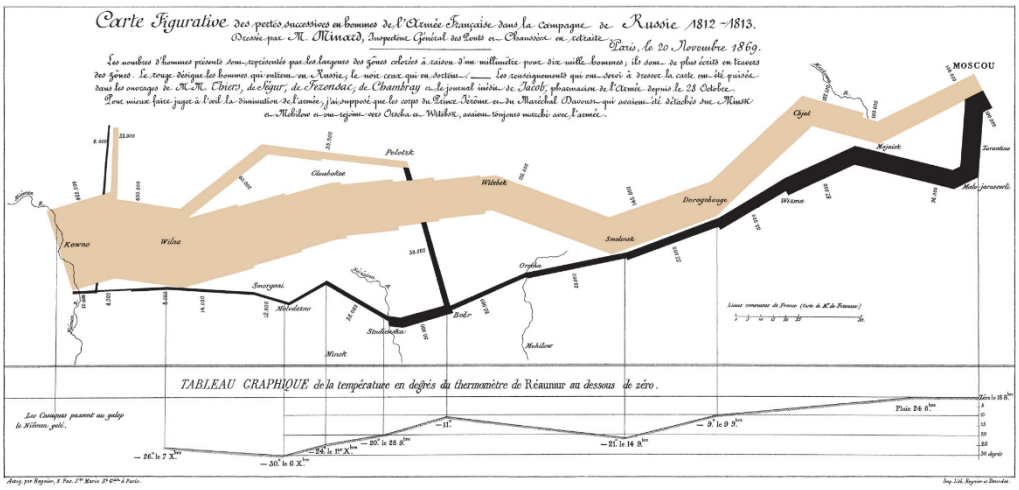
U ovom udžbeniku termin *dijagram* ćemo koristiti za označavanje bilo kog vizuelnog prikaza objekata, fenomena ili njihovih međusobnih relacija korišćenjem simbola. Mapa podzemne železnice u nekom gradu nije (samo) mapa već dijagram na kome su različitim simbolima i bojama prikazane putanje železničkih linija, njihova ukrštanja i stanice. Porodično stablo se obično prikazuje dijagramom hijerarhijski organizovanih i povezanih pravougaonika koji simbolišu osobe i njihove rodbinske veze. Primeri dijagrama su strukturne formule kojima se grafički prikazuju molekuli hemijskih jedinjenja, kao i vizuelni prikaz delova i načina rada motora. Međutim, za različite vizualizacije u oblasti statistike češće se koristi nešto uži pojam *grafici* ili *grafikoni*. Grafikoni su vrsta dijagrama kojima se upotrebom boja, oblika, linija i tačaka, na pregledan i sažet način, prikazuju skupovi podataka. Jednostavan primer grafikona predstavljaju piktografici koje smo opisali u prethodnom odeljku. O raznolikosti formi statističkih grafikona govori činjenica da se u engleskom jeziku za njihovo označavanje koriste čak tri termina: *chart*, *graph* i *plot*. Prvi ima najšire značenje i koristi se da označi grafikone sastavljene od geometrijskih oblika, npr. pravougaonika različite visine (engl. *bar chart*) ili odsečaka kružnice koji imaju različitu površinu (engl. *pie chart*). Termin *graph* se obično vezuje za iscrtavanje linija, npr. da bi se prikazao trend porasta ili pada akcija neke kompanije (engl. *line graph*), slično našem primeru iz odeljka o vizuelnim metaforama. Pojedini autori predlažu da kriterijum za razlikovanje pojmova *chart* i *graph* bude (ne)postojanje precizno definisanog pravila za određivanje rasporeda objekata na dijagramu (Börner, Maltese, Balliet, & Heimlich, 2016). Tako bi **oblak tagova** bio *chart*, jer ne postoji kriterijum na osnovu koga bi se potpuno precizno odredila veličina i položaj reči u oblaku (Slika 7). Sa druge strane, dijagrami koji se iscrtavaju u koordinatnom sistemu bili bi *graphs*, jer je pozicija elemenata grafikona tačno određena njihovim vrednostima na horizontalnoj i vertikalnoj osi. Ipak, u većini slučajeva navedeni termini upotrebljavaju se nedosledno ili kao potpuni sinonimi. Izuzetak je donekle termin *plots* koji se po pravilu koristi kao naziv za kategoriju grafikona nastalih iscrtavanjem tačaka na jednoj osi ili,

češće, u koordinatnom sistemu (npr. *box plot* ili *scatter plot*). U poslednje vreme sve više se koristi i termin *infographics,* ponekad kao sinonim za bilo koji vid vizualizacije informacija, ali češće da označi prigodne grafičke prikaze koji se ne mogu nazvati grafikonima, već pre ilustracijama čija je namena efikasno informisanje opšte populacije. Infografici se obično koriste u sredstvima javnog informisanja, kao vizualizacije koje upotpunjuju vesti, analize i novinske članke. Tipičan primer vizuelno bogatih infografikona predstavljaju ilustracije u naučno-popularnim časopisima kao što je **National geographic**.

Slika 7. Oblak tagova (najčešćih reči) iz ovog udžbenika napravljen pomoću servisa https://www.wordclouds.com

Na početku ovog odeljka pomenuli smo pećinske crteže i antičke mape kao primere rane vizualizacije. U najširem smislu, to su ujedno i prvi oblici dijagrama. Međutim, ključna razlika između ranih i savremenih formi grafičkog predstavljanja, pored tehnološkog aspekta, nalazi se u stepenu apstraktnosti, intenzitetu sažimanja i broju predstavljenih dimenzija. Kao ilustrativan istorijski primer možemo da navedemo **crtež nepoznatog astronoma iz 10. veka** koji prikazuje putanje Sunca, Meseca i nekoliko planeta. Karakteristika koja izdvaja ovaj dijagram od mapa nebeskih tela koje su pravljene još u drevnom Egiptu jeste činjenica da je na njemu prikazana dimenzija vremena kao apstraktno

svojstvo nastalo sažimanjem podataka iz različitih vremenskih perioda. Još jedan primer vizualizacije u kontekstu u kome danas koristimo taj termin, predstavlja shema **putanje Napoleonove armije** u pohodu na Rusiju 1812. godine koju je izradio francuski inženjer Šarl Minar (Slika 8). Na mapi je debljinom linije prikazan broj vojnika koji je sa oko 420.000 na početku invazije, spao na 10.000 preživelih koji su uspeli da se povuku sa bojišta. U donjem delu slike prikazane su temperature u različitim vremenskim periodima, tako da je uz pomoć dijagrama moguće dovesti u vezu porast smrtnosti sa padom temperature. Moć vizualizacije da otkrije nevidljive, ali i da jasnije prikaže vidljive relacije među fenomenima, ilustruje primer mape koju je izradio engleski lekar Džon Snou kako bi otkrio izvor epidemije kolere u Londonu 1854. godine. On je na mapi Londona iscrtao lokacije pumpi za vodu, a potom crticama označio broj smrtnih slučajeva na mestu svakog domaćinstva. Na osnovu grupisanja crtica zaključio je da se kolera širi preko zagađene vode, odnosno ručke jedne od pumpi, a ne udisanjem „lošeg vazduha" ili *mijazme*, kako je smatrala većina lekara toga doba. Sličnih primera ima još mnogo, a za detaljan pregled istorije vizuelnog mišljenja i vizualizacije informacija preporučujemo da pogledate projekat **Milestones in the history of thematic cartography, statistical graphics, and data visualization** koji su sačinili psiholozi Majkl Frendli i Denijel Dejvis (Friendly & Denis, 2001). Na ovom mestu ćemo pomenuti samo još ime Viljema Plejfera, škotskog ekonomiste iz 18. veka, začetnika vizualizacije u statistici i tvorca prvih modernih grafikona kao što su **stubičasti i pita dijagram** koji će biti opisani u narednom odeljku.



Slika 8. Mapa putanje Napoleonove armije u pohodu na Rusiju koju je izradio francuski inženjer Šarl Minar (Izvor: Wikimedia Commons)

Razvoj informacionih tehnologija omogućio je da u realnom vremenu vizualizujemo terabajte podataka. Međutim, osnovni principi vizualizacije koje su postavili ranije pomenuti pioniri u ovoj oblasti, još uvek su presudni za razumevanje svojstava uspešne vizualizacije. Suština grafičkog predstavljanja podataka nije u tome da se potpuno iskoriste raskoš i sve funkcionalnosti savremenih alatki za vizualizaciju, već da se odabere i primeni najprikladniji i najjednostavniji način pomoću koga će se nevidljivo učiniti vidljivim i lako razumljivim. Kao ilustraciju ćemo upotrebiti API (engl. *application programming interface*) servisa **OpenLayers** za kreiranje geografskih mapa na veb stranicama. Prikazaćemo postupak kojim je Džon Snou uz pomoć jednostavne vizualizacije utvrdio **izvor epidemije kolere** (Slika 9). Na mapu današnjeg Londona dodata su dva sloja. Na prvom su ucrtane lokacije tadašnjih pumpi za vodu, a na drugom lokacije smrtnih slučajeva. Prikažite pumpe a potom i lokacije obolelih.

> Na koji način je prikazano više smrtnih slučajeva na istoj adresi?
>
> Na osnovu čega je Džon Snou zaključio gde se nalazi izvor zaraze?
>
> Da li se smrtni slučajevi nalaze samo u blizini zaražene pumpe označene crvenim markerom? Kako to utiče na zaključak o zdravstvenoj ispravnosti pumpi označenih zelenim markerima?

Slika 9. Mapa epidemije kolere u Londonu 1854. godine engleskog lekara Džona Snoua izrađena uz pomoć API servisa https://openlayers.org

U ovom primeru treba obratiti pažnju na dve stvari. Prva je da crtice na originalnom crtežu Džona Snoua predstavljaju pojedinačne podatke koji sami

po sebi ne govore ništa o nekom fenomenu (npr. smrtnosti) i njegovoj vezi sa drugim pojavama (npr. izvoru zaraze). Podatke je, dakle, potrebno na neki način grupisati, sažeti i obraditi da bismo zaista razumeli neku pojavu ili uočili neku pravilnost. Drugi bitan momenat je *verovatnoća* nekog događaja. Smrtni slučajevi ne nalaze se samo u okolini zaražene pumpe već su sporadično raspoređeni i relativno daleko od nje. Međutim, najveća učestalost smrtnih slučajeva je upravo u blizini crvenog markera. Drugim rečima, veća verovatnoća smrtnog ishoda je očigledno povezana sa većom verovatnoćom da je osoba pila vodu sa zaražene pumpe. O ovim temama biće više reči u drugom poglavlju.

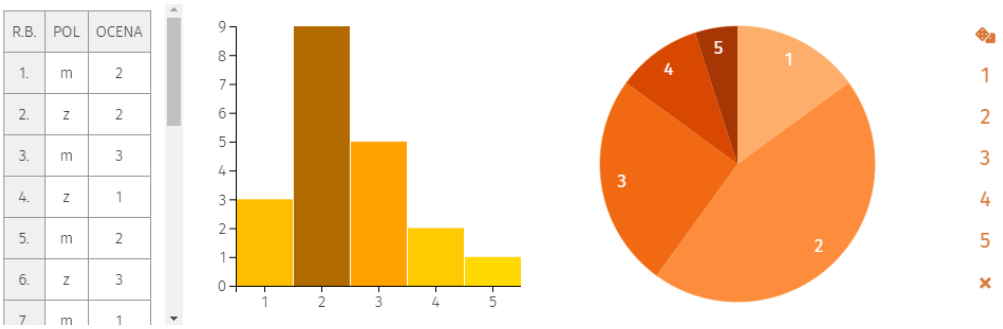
1.5. Podatak, informacija, znanje, razumevanje

U dosadašnjem tekstu smo naizmenično i nedosledno koristili termine *podatak* i *informacija*. Treba napomenuti da ovi termini nisu sinonimi. Termin *podaci* (engl. *data*) obično označava skup vrednosti, odnosno sirov materijal koji nastaje kao rezultat niza merenja. Tek nakon sažimanja, obrade i smeštanja u odgovarajući kontekst, podatak postaje *informacija* (engl. *information*). Ono što je upotrebljivo iz aspekta krajnjeg korisnika jeste informacija a ne (sirov) podatak. Na primer, mogli bismo da kažemo da Gugl obezbeđuje samo podatak o tome u kojim dokumentima se javlja termin koji tražimo, ali ne i informaciju o temi koja nas interesuje. Podatak da će sutra najverovatnije padati kiša postaje informacija tek ako dobije primenu, npr. u odluci da na posao ponesemo kišobran. Sirovi podaci mogu da imaju različite forme, ali se za potrebe statističke obrade obično skladište u obliku *tabela* ili *matrica*. U ćelije (kućice) takve tabele unose se pojedinačne vrednosti (engl. *datum*). Horizontalni nizovi ćelija (*redovi*) sadrže podatke o različitim entitetima ili objektima merenja. U psihologiji su to najčešće, mada ne i isključivo, osobe kao ispitanici. Vertikalni nizovi ćelija (*kolone* ili *vrste*) sadrže podatke o svojstvima, odnosno merljivim atributima entiteta. U statistici svojstva objekata koja mogu da imaju različite vrednosti kod različitih entiteta, nazivamo *varijablama*. Na primer, ukoliko u grupi od 100 učenika neke škole izmerite ili registrujete njihove vrednosti na tri varijable: *ocena iz matematike*, *pol* i *razred*, biće vam potrebna matrica dimenzija 100 x 3. Ovakvu tabelu ćemo nazvati *matricom sirovih podataka*. Ukoliko ne računamo red u kome se nalaze nazivi varijabli i kolonu u kojoj se nalaze npr. redni brojevi ili imena đaka, prvi red matrice će sadržati podatke o svim unetim svojstvima jednog učenika, a prva kolona će

sadržati podatke o oceni iz matematike za sve učenike. U tako organizovanoj tabeli relativno lako možemo da pronađemo bilo koji sirovi *podatak*. Međutim, korisne *informacije* dobijamo tek kada podatke iz tabele na određeni način sažmemo, obradimo, analiziramo i razumemo. Vizualizacija nam, u tom smislu, pomaže da na kontinuumu koji počinje podatkom i nastavlja se informacijom, odemo i dalje, ka znanju i uviđanju, odnosno ka primeni informacija za rešavanje praktičnih problema. U tom smislu, pojedini autori govore o *hijerarhiji znanja* (videti npr. Ackoff, 1989) ukazujući na potrebu da podaci uvek treba da dovedu do uviđanja, saznanja, razumevanja i konkretne primene. Bez toga oni ostaju samo nizovi beskorisnih brojki ili reči. Na primer, podaci o zadovoljstvu korisnika usluga nekog mobilnog operatera biće neupotrebljivi ukoliko ne dovedu do otklanjanja uočenih nedostataka.

1.5.1. Tabelarni i grafički prikaz podataka

Zamislimo da ste kao nastavnik prikupili podatke o polu i zaključenoj oceni iz matematike u grupi od **20 đaka nekog odeljenja** (Slika 10). Sa leve strane okvira prikazana je tabela dimenzija 20 x 2. Svaki red u tabeli predstavlja jednog đaka. Prva kolona sadrži redne brojeve koji su dodati samo radi preglednosti, tako da je nećemo tretirati kao varijablu, odnosno svojstvo ispitanika. Naravno, uvek se može izdvojiti jedna kolona matrice za imena osoba ili šifre na osnovu kojih bi one mogle da se identifikuju ukoliko je potrebno. Međutim, u većini istraživanja identitet ispitanika nije ni bitan jer se zaključci donose o grupi kao celini. Osim toga, lični podaci ispitanika ne smeju da se koriste i čuvaju bez njihove saglasnosti i/ili odobrenja odgovarajućih etičkih komisija. U drugu kolonu tabele već su uneti podaci o polu đaka. Treća kolona predviđena je za unos ocena iz matematike.



Slika 10. Primer matrice sirovih podataka, stubičastog i torta dijagrama

Unesite nekoliko nasumičnih vrednosti iz raspona od 1 do 5 u prazne ćelije treće kolone. Posmatrajte kako se menja grafički prikaz podataka sa desne strane. Odabrana su dva najpopularnija oblika vizualizacije podataka – *stubičasti dijagram* (engl. *bar chart*) i *kružni dijagram*, poznatiji kao *torta* ili *pita* dijagram (engl. *pie chart*). Koordinatni sistem u kome se nalazi stubičasti dijagram sadrži dve ose. Na horizontalnoj, koju nazivamo x-osa ili *apscisa*, označene su moguće vrednosti varijable koju vizualizujemo, npr. ocene od 1 do 5. Na vertikalnoj osi, koju nazivamo y-osa ili *ordinata*, označava se broj, odnosno *učestalost* ili *frekvencija* članova svake kategorije, npr. broj učenika koji su dobili ocenu 2. Odsečci na torta dijagramu predstavljaju *relativne frekvencije* ili *proporcije* određene kategorije u ukupnom broju elemenata. Pošto se računaju kao odnos učestalosti neke vrednosti i ukupnog broja merenja, proporcije mogu da se kreću u rasponu od 0 do 1. Kliknite ikonicu x da biste uklonili podatke iz tabele, a potom u nju unesite vrednosti 1, 2, 3 i 4. Obratite pažnju na to da su svi stubići iste visine i da svaka od kategorija ocena na kružnom dijagramu zauzima po četvrtinu, odnosno 0,25 delova kruga. Ako u prazne ćelije tabele unesete još jedan niz vrednosti 1, 2, 3 i 4, torta dijagram će izgledati identično kao i pre toga, jer je proporcija svake od ocena u odnosu na ukupan broj merenja (2 : 8 = 0,25) jednaka kao i u prethodnom primeru (1 : 4 = 0,25). Stubičasti dijagram se izmenio samo utoliko što maksimalna vrednost na y-osi više nije 1, već 2. Upotrebite ikonicu kockica sa desne strane da biste generisali nasumične nizove podataka i analizirajte izgled dobijenih grafikona. Primetićete da su u nekim situacijama pojedine kućice u trećoj koloni tabele prazne, čime su označeni tzv. *nedostajući podaci*.

Koja ocena je najčešća u primeru 1? Da li ćete na ovo pitanje najlakše odgovoriti pomoću tabele, stubičastog ili torta dijagrama? Zbog čega?

U čemu je prednost stubičastog u odnosu na kružni dijagram u primeru 2?

Koji zaključak o podacima u primeru 3 biste lakše doneli na osnovu kružnog nego na osnovu stubičastog dijagrama?

Da li biste rekli da je uspeh odeljenja iz primera 4, kao grupe, dobar ili loš?

Oba grafikona u primeru 5 su simetrična. Kako na procenu simetričnosti utiču upotrebljene boje? Koji kriterijum je upotrebljen za određivanje nijansi boja na stubičastom a koji na kružnom dijagramu?

> Šta je potrebno uraditi sa postojećom matricom ako želite da prikupite podatke o dodatnih 20 đaka i izmerite im još jedno svojstvo, npr. zaključnu ocenu iz hemije?

1.5.2. Deskriptivna i inferencijalna statistika

Ranije pomenuti model hijerarhije znanja poznat je i kao *DIKW piramida* (engl. *Data, Information, Knowledge, Wisdom*). Može se reći da ova piramida znanja opisuje i tok analize podataka primenom statističkih metoda. Primarni cilj svake statističke obrade je da se neke pojave opišu, tj. da se objasni šta se desilo u prirodi ili društvu. U prvoj fazi analize obično se bavimo deskripcijom *podataka* (engl. *data*), pa se statističke tehnike koje se koriste za te potrebe nazivaju *deskriptivnim*. Stubičasti dijagram je tipična tehnika deskriptivne statistike koju smo u ranijem primeru upotrebili da opišemo uspeh đaka u nekom odeljenju. Iako veoma važan deo, a može se reći i neophodan prvi korak svake statističke obrade, opisivanje fenomena samo po sebi nije dovoljno. U nauci, pa tako i u statistici, uvek se trudimo da pored odgovora na pitanje *šta* se desilo, odgovorimo i na pitanje *zbog čega* se nešto desilo. Tada govorimo o nivou *informacija* nastalih od sirovih podataka (engl. *information*). To obično podrazumeva potrebu za većom količinom podataka i većim brojem varijabli kako bismo uspešno i iscrpno analizirali neki fenomen. Na primer, da bismo objasnili loš uspeh đaka, bilo bi dobro da imamo podatke o tome kako su ocenili svog nastavnika i da li su imali neophodan edukativni materijal u toku nastave. Ako krenemo naviše u piramidi znanja, videćemo da je objašnjenje razloga za pojavu nekih fenomena (npr. zemljotresa) veoma korisno, ali da je još korisnija mogućnost da te fenomene predvidimo na osnovu postojećih informacija, odnosno da procenimo verovatnoću njihovog (ponovnog) javljanja. Ovaj nivo analize je nešto što odgovara komponenti *znanja* (engl. *knowledge*) u DIKW piramidi. U ovoj fazi se koriste naprednije statističke tehnike koje se nazivaju *induktivnim,* jer omogućavaju zaključivanje o pojavama i njihovim odnosima na osnovu početnih premisa o podacima. Često se koristi i naziv *inferencijalne tehnike*, jer se zaključci uopštavaju sa entiteta na kojima je obavljeno merenje (npr. učenici nekoliko škola), na sve slične entitete (npr. učenike svih osnovnih škola u državi). Primenom induktivnih tehnika možemo, na primer, da uporedimo uspeh dve grupe đaka koji su učili iz različitih udžbenika i da na osnovu uočene razlike zaključimo koji od ta dva udžbenika

treba preporučiti sledećim generacijama. Ukoliko smo na osnovu obrađenih podataka, izdvojenih informacija i stečenog znanja u mogućnosti da ponudimo i određene preporuke, govorimo o najvišem nivou hijerarhije – *uviđanju* ili *razumevanju* (engl. *wisdom*). Iz navedenog opisa, jasno je da statističke obrade i njihovi rezultati postaju sve vredniji što smo više pozicionirani u DIKW hijerarhiji znanja. Podaci jesu veoma dragoceni, ali postaju korisni samo ako dovedu do uviđanja i praktične primene u rešavanju aktuelnih problema.

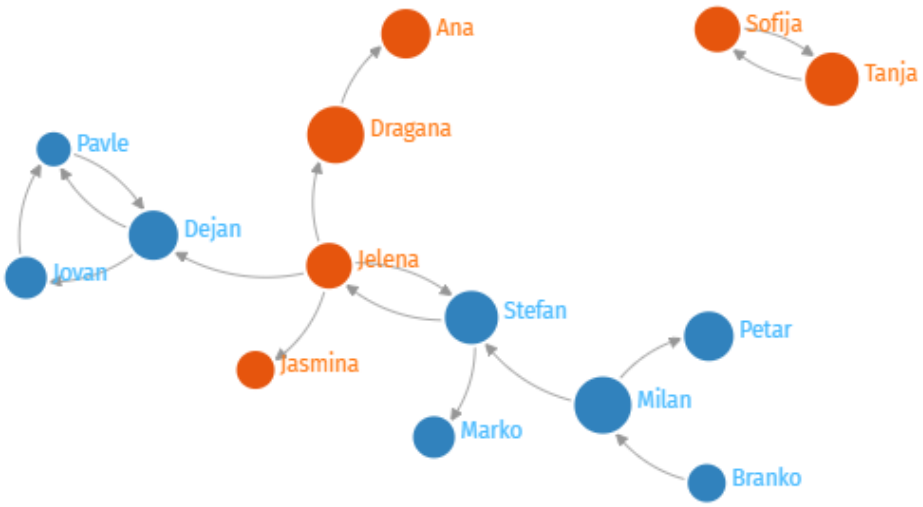
1.6. Naučna vizualizacija i vizualizacija informacija

Vizualizacija je višeznačan pojam. U oblasti psihologije obično se definiše kao proces formiranja vizuelnih predstava objekata, njihovih atributa i međusobnih relacija (Krstić, 1991). Reč je o unutrašnjem kognitivnom procesu koji se na prvom mestu može povezati sa konceptom vizuelnog mišljenja, jer podrazumeva formiranje mentalnih slika i modela kojima simuliramo ili ponovo proživljavamo događaje i pojave. Međutim, razvoj informacionih tehnologija i računarske grafike doveo je do proširenja značenja pojma vizualizacije i njegovog pomeranja sa procesa internalizacije, odnosno formiranja mentalnih slika u svesti osobe, na procese eksternalizacije. Van oblasti psihologije, vizualizacija se sve ređe posmatra kao interni kognitivni proces, a sve češće kao postupak grafičkog predstavljanja koncepata, podataka i ideja (Ware, 2004). Štaviše, u pojedinim definicijama, vizualizacija se vezuje isključivo za upotrebu interaktivnih, računarski generisanih vizuelnih reprezentacija (Card, Mackinlay, & Shneiderman, 1999). S obzirom na predmet vizualizacije, obično se pravi razlika između *naučne vizualizacije* i *vizualizacije informacija* (Card et al., 1999; Mazza, 2009; Zhang, 2010). Naučna vizualizacija odnosi se na prikazivanje fenomena i podataka koji imaju svoju analogiju u stvarnom svetu i čije relacije mogu da se opišu fizičkim i matematičkim modelima. Ova vrsta vizualizacije služi za prikazivanje onoga što realno postoji, ali najčešće nije vidljivo jer je suviše malo, udaljeno, nedostupno i nesagledivo iz prave perspektive. Primeri naučne vizualizacije su sheme hemijskih molekula, modeli Sunčevog sistema, rentgenski snimci ili prikaz strujanja vazduha u aerotunelima. S druge strane, vizualizacija informacija je primena vizuelnih predstava za prikazivanje fenomena koji su apstraktni i nemaju unapred definisani prostorni raspored niti strukturu. Ranije opisana mapa epidemije kolere je odličan primer vizualizacije informacija, jer na njoj nisu vidljivi samo konkretni objekti, kao što su pumpe i smrtni slučajevi, već i apstrahovani ili izvedeni fenomeni, kao što je učestalost

smrtnih slučajeva na određenim lokacijama ili povezanost pumpi sa većim rizikom od zaraze. Mape su veoma intuitivne grafičke predstave koje su postale široko prihvaćen model vizualizacije odnosa među objektima, jer ih korisnici veoma lako interpretiraju i razumevaju, nezavisno od toga da li se njima dočarava prostorna, tematska ili neka druga vrsta bliskosti objekata (Koshman, 2006; Chalmers, 1993; Cole, Mandelblatt, & Stevenson, 2002). Zbog svoje sličnosti sa fenomenima iz svakodnevnog života, ovakav pristup vizualizaciji je poznat kao *paradigma prirodnih situacija* (Robertson, 1991). Korišćenjem analogija sa rastojanjima, rasporedom i veličinama objekata na geografskim mapama, moguće je, na primer, dočarati socijalnu, emotivnu ili bilo koju drugu vrstu (latentne) bliskosti osoba.

Zamislite da kao školski psiholog imate potrebu da opišete odnose među učenicima nekog odeljenja. Kao pogodna alternativa ili dopuna vašem izveštaju u tekstualnoj formi, može da posluži *sociogram*. Sociogram je u osnovi *graf,* matematički objekat koji se predstavlja mrežom *čvorova* i njihovih *grana*, tj. međusobnih veza. Graf, dakle, nije isto što i grafikon, već samo jedna podvrsta dijagrama. Grafovi su intuitivna tehnika za prikazivanje strukture povezanih entiteta, kao što su npr. prijatelji na nekoj društvenoj mreži ili države koje imaju intenzivnu trgovinsku razmenu. U tehničkom smislu, iscrtavanje grafova obično zahteva primenu *algoritama ugrađenih opruga* (engl. *spring-embedded algorithms*) (Kamada & Kawai, 1989) koji matematički simuliraju fizički model u kome bi čvorovi bili kugle različite mase, povezane oprugama različite jačine. U celokupnom sistemu čvorova (kugli) i grana (opruga) deluju sile privlačenja i odbijanja, pa se ova vrsta grafičkog predstavljanja često naziva i *grafovima zasnovanim na silama* (engl. *energy-based* ili *force-directed graphs*) (Fruchterman & Reingold, 1991). Sile, odnosno napregnutost opruga, grafički se ispoljavaju tako što čvorovi (npr. učenici) koji su međusobno jače povezani (npr. sede zajedno na više časova) teže da na grafikonu budu bliži, a oni među kojima ne postoje veze ili su spojeni vezama slabijeg intenziteta, teže da se udalje. Iscrtavanje grafa najčešće započinje nasumičnim rasporedom čvorova, da bi se nakon manjeg ili većeg broja iteracija (ponavljanja) razmestili tako da „opruge" budu što manje napregnute. U svakom koraku se evaluira ukupna energija sistema ili suma sila, a postupak se završava u momentu kada nivo energije dostigne minimum. Minimum energije podrazumeva najprirodniji položaj i odnos grana i čvorova, odnosno graf na kome udaljenosti među čvorovima najvernije odražavaju jačinu veza među njima. To znači da pored osnovnog zahteva za postizanjem minimalne sume sila koje deluju na čvorove

i opruge, postoji i zahtev koji se tiče estetike ili dobre forme u kojoj će svi čvorovi i veze biti jasno uočljivi i ravnomerno raspoređeni po dostupnoj površini ekrana. Uzmimo kao primer **odgovore učenika na pitanje sa kime bi voleli da sede na času** (Slika 11). Na početnoj slici prikazane su samo pozicije, odnosno međusobna bliskost učenika, ali dijagram može da se dopuni i vrednostima drugih varijabli. Kada kliknete taster *Variraj boju*, krugovi će biti obojeni različitim bojama u zavisnosti od pola učenika. Veličinom kruga može se označiti uspeh u školi. Na kraju, uz svaki krug mogu se prikazati imena učenika. Menjate izgled grafa upotrebom opcija sa desne strane i analizirajte prednosti i nedostatke različitih prikaza. Dodajte još jednog učenika na graf klikom na taster *Dodaj čvor* i analizirajte kako je to uticalo na promenu strukture mreže i raspored čvorova.



Slika 11. Primer grafa, odnosno sociograma učenika jednog odeljenja

Kako na preglednost i intuitivnost grafa utiču variranje boje i veličine krugova?

Da li su raspored i pozicija objekata na ekranu potpuno proizvoljni ili strogo određeni?

Da li vam je lakše da uočite razlike među đacima u polu ili u uspehu?

Da li bi graf bio razumljiviji kada bi se pol učenika predstavio veličinom kruga a uspeh u školi različitim bojama? Da li je to opravdano?

> Zbog čega je, po vašem mišljenju, Milica „zvezda" odeljenja?
>
> Koje dodatne varijable bi se mogle vizualizovati na grafu i na koji način?

1.7. Vizualizacija kao eksplorativna tehnika

Pojedini autori prave razliku između vizualizacije za potrebe *prezentacije* i vizualizacije u cilju *eksploracije* (Unwin, Chen, & Härdle, 2008). Kriterijum za podelu je veoma sličan onome koji smo opisali u prethodnom odeljku, ali se u ovom slučaju tehnike vizualizacije razlikuju sa stanovišta njihove svrhe i primene, a ne s obzirom na vrstu podataka koji se vizualizuju. Čen (Chen, 2006), na primer, naglašava značaj vrste zadatka koji se postavlja pred korisnika u toku interakcije sa grafičkim prikazima. Prvi tip zadatka naziva se *vizuo-spacijalnim* i podrazumeva interpretaciju topoloških karakteristika elemenata na vizuelnoj predstavi. Drugi tip je *semantički* i odnosi se na uviđanje latentnih fenomena koji ne moraju nužno da odgovaraju fizičkim svojstvima vizuelne predstave. U oblasti statistike, vizualizacija se obično vezuje upravo za ove druge zadatke, odnosno za *eksplorativnu analizu podataka* (Tukey, 1977). Primena tehnika vizualizacije u statistici je na neki način „detektivski" posao, čiji cilj nije samo da se rezultati predstave u preglednoj formi, već i da se oni detaljnije opišu, međusobno povežu i da se u njima i na osnovu njih uoče (ne)pravilnosti. Nadmoćnost grafičkog predstavljanja je u tome što omogućava ne samo prikazivanje, već i *otkrivanje* podataka (Tufte, 1985). Dijagrami olakšavaju formiranje mentalnih predstava koje odražavaju ne samo spoljašnju stvarnost već i apstraktne relacije među podacima. Cilj svake vizualizacije je da kompleksno učini jednostavnim i da nevidljivo učini vidljivim i interpretabilnim. Vizualizacija nas podstiče da u podacima prepoznamo fenomene koje na prvi pogled nismo videli, pa čak ni očekivali. Zbog svega toga, vizualizacija kao sredstvo eksploracije unapređuje našu spoznaju, olakšavajući i ubrzavajući opažanje, učenje i zaključivanje. Prema psihologu Stjuartu Kardu, koji je takođe bio jedan od vodećih istraživača ranije pomenutog PARC centra, postoji barem pet načina na koje vizualizacija pospešuje naše razumevanje stvarnosti (Card et al., 1999):

1. *Proširenjem kognitivnih kapaciteta*. Primenom vizualizacije rasterećuju se i efikasnije koriste perceptivni i memorijski resursi korisnika. Naši kapaciteti za pamćenje vizuelnih sadržaja su praktično neograničeni i

daleko nadmašuju sposobnost prisećanja verbalnog materijala (Standing, 1973). Osim toga, grafičke metafore često imaju ulogu „spoljašnje memorije" zato što omogućavaju paralelno procesiranje većeg broja svojstava objekata za razliku od tekstualnih sadržaja koji se procesiraju serijski.

2. *Ubrzavanjem pretrage*. Velike količine podataka mogu da se prikažu grafički na malom prostoru, pri čemu različita vizuelna svojstva objekata kao što su boja, veličina ili oblik, olakšavaju njihovo pronalaženje i razumevanje u kontekstu drugih elemenata. Najjednostavniji primer je već pomenuta upotreba različitih boja i zadebljanja teksta u prikazivanju rezultata pretrage interneta ili lociranje učenika koji su najpopularniji ili najbolji u odeljenju.

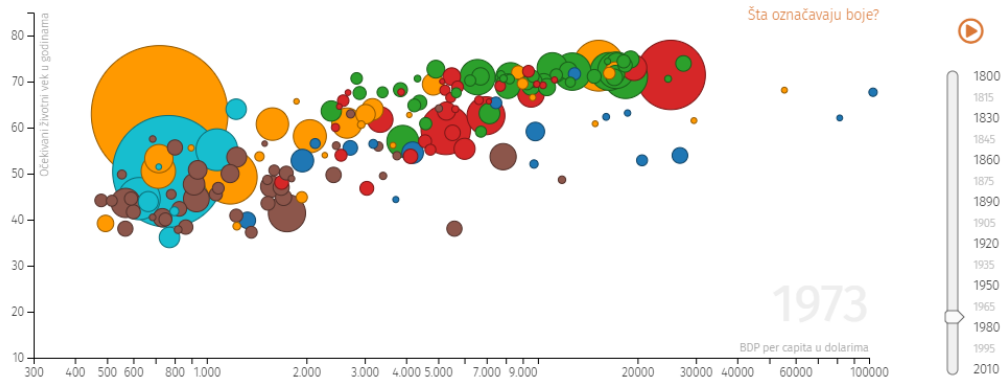
3. *Jasnijim uvidom u strukturu i promene*. Vizualizacijom postaju uočljive relacije među podacima koje su apstraktne ili nisu direktno vidljive pre nego što se velike količine podataka sažmu. Dinamičkim vizualizacijama moguće je prikazati trendove i promene u vremenu. Sjajan primer predstavlja vizuelizacija kojom se traga za mogućim **razlozima globalnog zagrevanja planete Zemlje**.

4. *Olakšavanjem zaključivanja*. Grafički prikaz podataka omogućava lakše uočavanje problema, formiranje pretpostavki i donošenje zaključaka već na nivou *ranog opažanja*, odnosno percepcije osnovnih karakteristika objekata kao što su boja, veličina, oblik i položaj. Na primeru sociograma videli smo koliko variranje svojstava grafičkih objekata pomaže njihovom uočavanju, grupisanju i povezivanju.

5. *Efikasnijim manipulisanjem podacima*. Vizualizacija omogućava interakciju sa korisnikom koji u pravom smislu te reči *istražuje* informacioni prostor upravljajući vizuelnim metaforama. Veoma lako se mogu dodavati i oduzimati elementi prikazani na slici, i posmatrati kako te promene utiču na ostale podatke.

Kao prikladnu ilustraciju navedenih prednosti vizualizacije za potrebe eksploracije podataka, upotrebićemo neznatno modifikovani grafikon koji je izradio **Majk Bostok**, autor biblioteke D3 korišćene i za izradu ovog udžbenika. Grafikon prikazuje **podatke o životnom veku i bogatstvu nacija** koje je švedski fizičar i statističar Hans Rosling, osnivač fondacije **Gepmajnder**, predstavio u svojoj **TED prezentaciji**. U pitanju je *grafikon kretanja* (engl. *motion chart*) kojim se istovremeno može prikazati veći broj varijabli i njihove

promene u toku vremena (Slika 12). Na grafikonu su prikazane vrednosti sledećih varijabli za 180 zemalja sveta: broj stanovnika, očekivani životni vek stanovnika, bruto nacionalni dohodak i region u kome se država nalazi. Vrednosti varijabli mogu da se predstave dinamički u vremenskom periodu od 1800–2010. godine pomeranjem klizača sa desne strane ili klikom na ikonicu u gornjem desnom uglu iznad klizača. Ime države prikazuje se kada se iznad kruga postavi pokazivač miša. Legenda koja objašnjava značenje boja dostupna je u gornjem desnom uglu grafikona.



Slika 12. Grafikon kretanja životnog veka i bogatstva nacija baziran na ideji i podacima Hansa Roslinga

Uporedite različite države i regione po broju stanovnika i nacionalnom dohotku.

Analizirajte promene u vremenu i povežite ih sa značajnim istorijskim događajima.

Uočite države koje se u različitim vremenskim periodima i po različitim svojstvima izdvajaju od ostalih država istog regiona ili celog sveta.

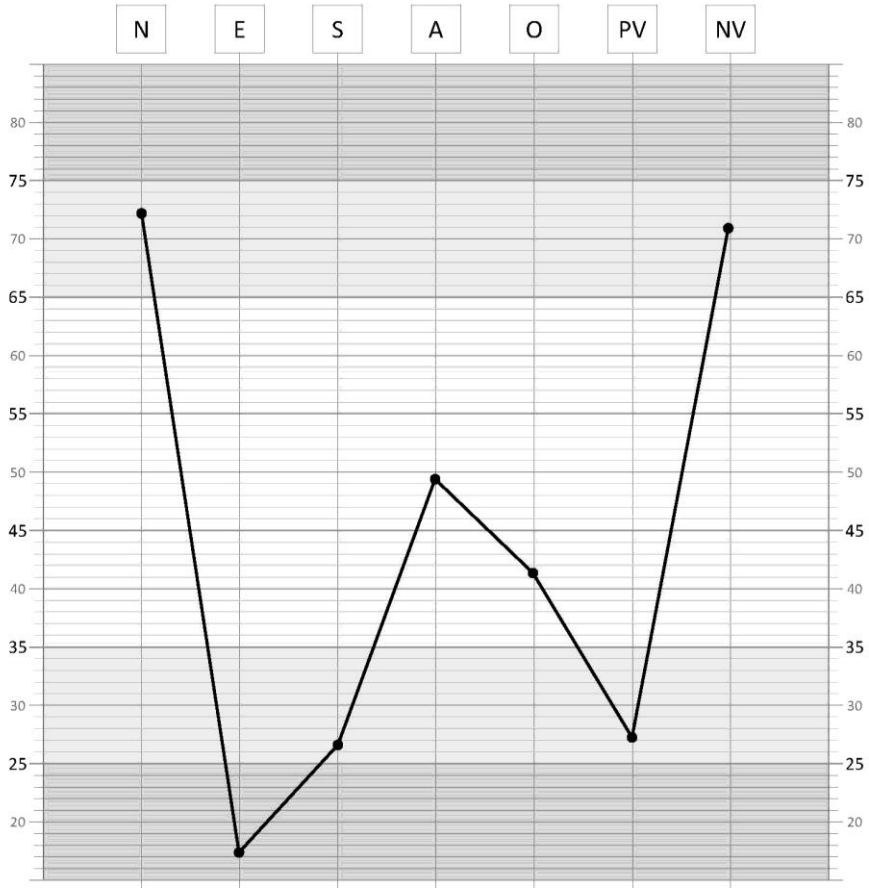
Da li su države Supsaharske Afrike u poslednjih 100 godina uspele da dostignu evropske države po kvalitetu života?

Koje države su imale najveći rast bruto nacionalnog dohotka po glavi stanovnika nakon Drugog svetskog rata?

1.8. Izbor prikladne tehnike vizualizacije

Podaci su svuda oko nas. Vizualizacija nam pomaže da ih sažmemo, strukturiramo, istražimo i pretvorimo u znanje. Termini koji se najčešće koriste za pretragu interneta mogu da se upotrebe za opisivanje interesovanja, navika i **karakteristika medijske kulture**. Emisija pozitrona iz radionuklida detektuje se uz pomoć PET skenera i pretvara u sliku na osnovu koje određeni psihološki procesi mogu da se povežu sa **aktivacijom moždanih zona**. Odgovori na stavke upitnika predstavljaju osnovu za formiranje **grafičkih profila** osobina ličnosti (Slika 13) (Kodžopeljić, Smederevac, Mitrović, Ćolović, & Pajić, 2019). **Obim trgovinske razmene** može da pokaže veze među različitim regionima i državama sveta. Učestalost zajedničkog pojavljivanja (tzv. *koincidencija*) reči u naučnim člancima omogućava mapiranje naučnih disciplina i prepoznavanje **aktuelnih istraživačkih tema**. Prevalenca bolesti u određenom vremenskom periodu ukazuje na **učinkovitost vakcina**. Na osnovu pokazatelja bliskosti pojedinaca može se zaključivati o **strukturi socijalnih mreža**. I tako dalje. Imajući u vidu toliku raznolikost formi grafičkog predstavljanja podataka, pomenute podele i tipologije deluju kao artefakt, jer gotovo da ne postoji vizualizacija koja se može nazvati „nenaučnom", kao što ni svaka naučna vizualizacija nije samo shematski prikaz stvarnosti, već omogućava istraživanje informacija pružajući dodatni kvalitet u odnosu na sirove podatke ili fizički model objekta koji se prikazuje. Stoga možemo reći da podele koje smo naveli u prethodnim odeljcima više služe kao smernice koje određuju naša očekivanja od vizualizacije, odnosno kriterijume na osnovu kojih ćemo određeni grafički prikaz proceniti kao prikladan ili ne. Prilikom odabira adekvatnog rešenja iz širokog spektra tehnika i formata vizualizacija, potrebno je uzeti u obzir više faktora koji bi grubo mogli da se podele u tri grupe. Prva grupa se tiče osobina podataka, odnosno broja varijabli koje želimo da vizualizujemo i načina na koji smo ih izmerili. U tom smislu, naučna vizualizacija najčešće ne izlazi van okvira tri osnovne dimenzije, budući da polazi od postojećih objekata i prirodno datih formi i struktura. Sa druge strane, vizualizacija informacija obično podrazumeva opisivanje višedimenzionalnog prostora, pri čemu se različitim karakteristikama objekata na grafikonu dočaravaju vrednosti različitih varijabli. Druga grupa faktora odnosi se na svrhu vizualizacije, odnosno osobine korisnika kojima je namenjena. Vizualizacija za potrebe ilustracije sadržaja teksta u dnevnim novinama ne može da ima isti oblik kao i grafikon prikazan u naučnom članku. Ovde treba uzeti u obzir ne samo kompetencije i motivaciju ciljne grupe osoba

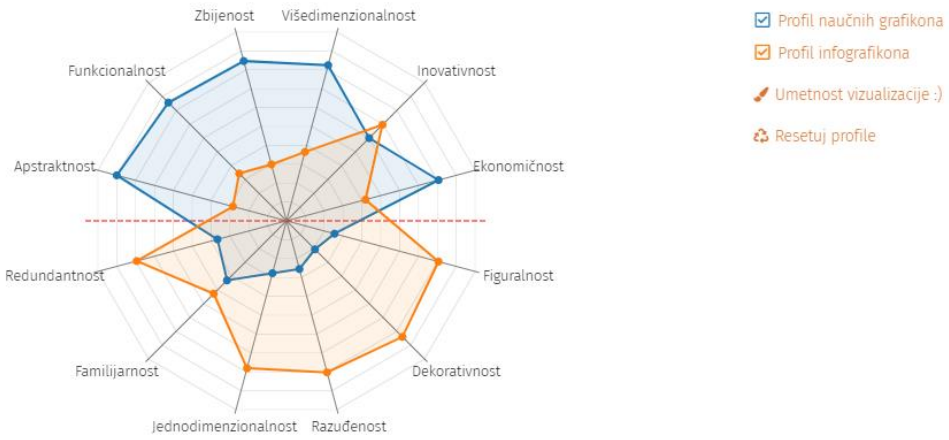
kojima su dijagrami namenjeni već i njihove osobine i potencijalne probleme, kao što su slabovidost ili nemogućnost razlikovanja boja. Na kraju, treća grupa faktora odnosi se na medijum koji se koristi za vizuelnu komunikaciju. To može da bude štampani materijal, veb stranica, aplikacija na mobilnom telefonu, ekran neke mašine, video snimak ili nešto drugo. Od konkretnog medijuma zavisi da li će vizualizacija moći da bude interaktivna, koja svojstva objekata će biti moguće prikazati, koja količina podataka će moći da se iskoristi i tako dalje.



Slika 13. Primer vizualizacije profila ličnosti formiranog pomoću upitnika Velikih pet plus dva za decu

Priprema, organizacija i prezentacija informacija postala je presudna u vremenu u kome se važne strategijske odluke donose na osnovu obrade ogromnih količina podataka. Stoga se naglašava i potreba za *dizajnerima informacija* kao novoj profesiji za 21. vek (Horn, 1999). Na ovom mestu ćemo pomenuti španskog dizajnera informacija Alberta Kaira, dugogodišnjeg

novinara i kreatora infografika za poznati list El Mundo. Kairo je u svojoj knjizi *Funkcionalna umetnost* (Cairo, 2013) predložio tzv. *točak vizualizacije* kao grafički model na osnovu koga se procenjuje prikladnost vizualizacije za određenu svrhu. Točak se sastoji od dvanaest dimenzija, odnosno šest dvopolnih skala pomoću kojih se opisuju karakteristike nekog grafičkog rešenja. Te dimenzije su: apstraktnost – figuralnost, funkcionalnost – dekorativnost, inovativnost – familijarnost, ekonomičnost – redundantnost, zbijenost – razuđenost i višedimenzionalnost – jednodimenzionalnost. Za ilustraciju primene **točka vizualizacije** upotrebićemo *pauk* ili *radar dijagram* (Slika 14). Na osnovu njegovog izgleda lako se može zaključiti kako je dobio naziv. Svaka linija koja polazi od centra grafikona predstavlja jednu dimenziju, a pozicija tačaka na tim linijama, odnosno njihova udaljenost od centra, određena je vrednošću te varijable. Spajanjem tačaka dobija se linijski profil entiteta kao kombinacija vrednosti na svim varijablama. Na taj način grafikonom je moguće prikazati vrednosti većeg broja varijabli za više entiteta istovremeno, npr. ocene đaka ili odeljenja iz različitih predmeta. U ovom primeru iskorišćeni su podaci koje je Kairo naveo kao karakteristike infografika, odnosno naučnih (statističkih) grafikona, da bi ilustrovao različite pristupe predstavljanju podataka u grafičkoj formi u zavisnosti od ciljne grupe korisnika.



Slika 14. Točak vizualizacije Alberta Kaira izrađen uz pomoć radar dijagrama

> Na kojim dimenzijama se profil infografika u najvećoj meri razlikuje od profila naučnih grafikona?

Prikažite oba profila. Da li vam je na osnovu prikazanog grafikona lakše da uočite razlike među profilima ili razlike na pojedinačnim dimenzijama?

Ako šest osa radar dijagrama posmatramo kao dvopolne (bipolarne) dimenzije na čijim su krajevima suprotne karakteristike, koje svojstvo grafikona nije intuitivno ili nije logično?

Da li bi poređenje entiteta, odnosno varijabli, bilo podjednako lako i opravdano kada bi varijable imale različite raspone vrednosti, npr. ocena iz fizičkog, visina đaka, težina đaka, vreme za koje đak pretrči 100 metara i broj zgibova koje je uspeo da uradi?

Pronađite proizvoljnu vizualizaciju na internetu i izmenite vrednosti na grafikonu tako da formiraju profil koji bi joj najviše odgovarao.

Kao što smo videli iz prethodnog primera, vizualizacije za potrebe informisanja šire javnosti obično se suštinski razlikuju od vizualizacija kojima se predstavljaju rezultati naučnih istraživanja. Prvi elementi parova pomenutih dimenzija, prikazani u gornjem delu dijagrama, predstavljaju karakteristike koje poseduju, ili bi trebalo da poseduju, vizualizacije nastale statističkim obradama, posebno primenom eksplorativnih tehnika analize podataka. Drugi elementi parova su u većoj meri karakteristični za infografike i vizualizacije namenjene opštoj populaciji. Tako će u dnevnim novinama učestalosti nekih pojava ređe biti predstavljene stubićima koji su u suštini apstraktni oblici, a češće figuralno, konkretnim slikama ili piktogramima objekata koje treba da simbolišu, npr. sličicama naslaganih novčića kojima se prikazuje prihod različitih kompanija ili brojem tenkova koji govori o količini naoružanja država koje se porede. Pored toga, infografici se često dodatno dekorišu prigodnim slikama i elementima koji nisu tipični za grafikone namenjene istraživačima i analitičarima. Na primer, uobičajena greška koja se pravi prilikom prezentovanja stubičastih i kružnih dijagrama je upotreba treće dimenzije tako da stubići postanu kvadri a krugovi valjci. Ovakva vrsta „estetski unapređenih" grafikona često se može videti u sredstvima javnog informisanja, ali ona nije prikladna za prikazivanje naučno-istraživačkih rezultata. Svaka karakteristika objekata na grafikonu trebalo bi da predstavlja samo jedno i tačno određeno svojstvo, te stoga dodavanje treće dimenzije stubiću ili kružnici ne nosi nikakvu informativnu vrednost i najčešće zbunjuje posmatrača. Korisnici svakako neće procenjivati i porediti zapreminu trodimenzionalnih figura već samo njihove osnovne dimenzije – širinu, visinu ili površinu. Takođe, infografici mogu da sadrže podatke koji su redundantni, dok

statističke grafikone treba da odlikuje funkcionalnost i ekonomičnost, odnosno racionalno trošenje površine na kojoj se podaci prikazuju. Poznati statističar Edvard Tafti naziva ovo svojstvo *odnos podaci–mastilo*, naglašavajući potrebu da prilikom izrade grafikona svako povećanje utroška mastila (u kontekstu elektronskog sadržaja to bi mogao da bude utrošak tačaka ekrana) bude praćeno povećanjem količine prikazanih podataka (Tufte, 1985). Tafti navodi devet principa koje treba poštovati prilikom izrade grafikona:

1. Pokažite podatke.
2. Podstaknite posmatrača da razmišlja o suštini informacije, a ne o dizajnu ili načinu na koji je grafikon nastao.
3. Izbegavajte bilo kakvo iskrivljenje podataka.
4. Prikažite puno podataka na malom prostoru.
5. Velike skupove podataka učinite koherentnim.
6. Podstaknite posmatrača da poredi podatke prikazane na grafikonu.
7. Prikažite podatke na nekoliko nivoa detalja, od opšteg pregleda do specifične strukture.
8. Upotrebite grafički prikaz za određenu svrhu: opisivanje, eksploracija, tabulacija ili dekoracija.
9. Pridružite grafikonu odgovarajući verbalni i statistički opis.

1.8.1. Nivoi merenja varijabli

U prethodnom odeljku naveli smo tri grupe faktora koji utiču na procenu prikladnosti neke vizualizacije za određenu svrhu. U kontekstu primene grafičkih tehnika u statistici, najznačajnija je prva grupa koja se tiče karakteristika varijabli. Pored broja varijabli koje je potrebno prikazati u grafičkoj formi, za izbor odgovarajućeg grafikona veoma bitan je i način na koji smo izmerili varijable. U prirodnim i tehničkim naukama merenje se gotovo isključivo vezuje za numeričke podatke. Međutim, u psihologiji i drugim društvenim i humanističkim disciplinama merenje može da se shvati i šire, kao *proces dodeljivanja kvantitativne vrednosti ili kvalitativne oznake svojstvu objekta u skladu sa odgovarajućim standardom i pod kontrolisanim uslovima*. Zato se u statistici često pravi razlika između dva tipa varijabli – kvantitativnih i

kvalitativnih (Slika 15). *Kvantitativne varijable* su atributi koje je moguće izmeriti u pravom smislu te reči, i to dodeljivanjem numeričke vrednosti koja pokazuje izraženost ili količinu svojstva koje objekat poseduje. Visinu osobe izražavamo numerički nakon poređenja sa usvojenim standardom (npr. metrom), kontrolišući uslove pod kojima se merenje obavlja (npr. osoba stoji uspravno i nije obuvena). Intelektualne sposobnosti kvantifikujemo pomoću broja tačno rešenih zadataka na standardizovanom testu u dobro osvetljenoj prostoriji, u vreme kada su ispitanici odmorni i niko ih ne ometa. U navedenim primerima, osobe ili ispitanici su *objekti* merenja, a metar i test su *instrumenti* kojima se merenje obavlja. Međutim, ponekad i osoba može da ima funkciju instrumenta merenja. Na primer, nastavnik ocenjuje znanje studenta na osnovu usmenog ispitivanja, a iskusan psiholog može grubo da proceni izraženost osobina ličnosti osobe već na osnovu intervjua. Jasno je da u ovakvim situacijama procena atributa nema objektivnost i preciznost merenja u užem smislu, ali neka svojstva ponekad nije ni moguće kvantifikovati. Stoga ćemo u ovom udžbeniku termin *merenje* koristi u širem značenju koje obuhvata i prosto svrstavanje objekta u određenu grupu, kategoriju ili klasu. Ukoliko određeno svojstvo ne može da se kvantifikuje, govorimo o *kvalitativnim varijablama*. Primeri kvalitativnih varijabli su pol, nacionalnost, dijagnoza bolesti, bračni status, veroispovest, mesto u kome živimo, marka telefona koji koristimo ili fakultet koji smo upisali nakon srednje škole. Ovim atributima mogu da se pripišu numeričke vrednosti (npr. 1 – žensko, 2 – muško), ali ti brojevi nemaju kvantitativno svojstvo, već služe samo kao oznaka pripadnosti klasi. O ograničenoj upotrebi brojeva kod kvalitativnih varijabli govori nam i činjenica da smo potpuno legitimno muškarce mogli da označimo brojem 123, a žene brojem 56 ili bilo kojim drugim brojem, slovom ili rečju.

Iz prethodnih primera može se uočiti da merenje ne podrazumeva uvek *direktno* poređenje svojstva objekta sa unapred datim standardom. Štaviše, merenje u psihologiji češće se obavlja *indirektno*, tj. registrovanjem manifestacija latentnih svojstava. Ne postoji „metar" kojim bismo izmerili neuroticizam osobe, već se o izraženosti ove osobine zaključuje indirektno, na osnovu odgovora na stavke upitnika ličnosti. Pri tome, osoba koja popunjava upitnik takođe obavlja indirektno merenje izražavajući svoje subjektivno slaganje sa nekom tvrdnjom, npr. na skali od 1 do 5. To znači da se varijable mogu izmeriti na različite načine i sa različitom preciznošću. Visina đaka može da se izrazi u centimetrima, ali i kao redni broj u koloni u kojoj su đaci poređani od najvišeg do najnižeg. Prvi način je precizniji od drugog zato što se *skala* centimetara suštinski razlikuje od skale kojom smo „merili" rang đaka. Prva bitna

razlika među njima je što skala mernih jedinica za visinu, odnosno dužinu, ima praktično neograničen broj podeoka. Između svaka dva podeoka koji označavaju centimetre, možete umetnuti podeok koji označava milimetre, između svaka dva milimetra mikrometar i tako dalje. Može se reći da je niz tih podeoka beskonačan i neprekidan, te se stoga ova vrsta varijabli naziva *kontinuiranim*. To naravno ne znači da visinu treba izražavati u nanometrima ili pikometrima, već samo da je to moguće kada je potrebno i kada se poseduju dovoljno precizni instrumenti. Pošto je visina kontinuirana varijabla, potpuno je opravdano osobi dodeliti vrednost 172,35 cm. Sa druge strane, nije uobičajeno da se nečiji rang u koloni označi vrednošću 12,78. Svojstvo *radna memorija* objekta *telefon* ili svojstvo *broj dece* objekta *porodica* takođe nije moguće (ili nije logično) izraziti vrednostima koje nisu celobrojne. Porodica ne može da ima 2,5 deteta, kao što se ne projektuju ni memorijski moduli kapaciteta 511,3 Mb. Takve varijable nazivamo *diskretnim* jer su podeoci na skalama kojima su merene isprekidani i jasno odvojeni „prazninama". Međutim, treba obratiti pažnju na činjenicu da se skale broja dece i memorije telefona bitno razlikuju od ranije pomenute skale rangova visine učenika. Iako su sve ove varijable diskretne, kod prve dve intervali između podeoka su jednaki u smislu veličine razlike u količini svojstva koje je mereno. Porodica sa vrednošću 4 na varijabli *broj dece* ima duplo više dece od porodice koja je dobila vrednost 2. Telefon sa vrednošću 6 Gb na varijabli RAM ima 4 Gb više memorije od telefona koji ima vrednost 2 Gb. Međutim, kod ranga visine to nije slučaj. Razlika u visini između prvog i drugog učenika u koloni jeste jedan rang ili jedna pozicija, ali ne mora da bude ista kao razlika između drugog i trećeg ili petog i šestog učenika. Učenik koji ima rang 10 nije duplo viši od učenika koji ima rang 5. Stoga ćemo smatrati da su varijable *broj dece*, *visina u cm* i *RAM* izmerene na višem *nivou merenja* nego visina izražena rangom.

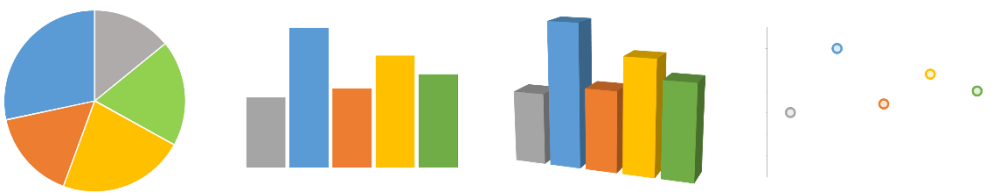
Najpoznatiju podelu varijabli na osnovu nivoa merenja predložio je američki psiholog Stenli Smit Stivens. Prema njegovoj podeli postoje četiri tipa varijabli: nominalne, ordinalne, intervalne i racio. *Nominalni* nivo merenja u suštini nije merenje u pravom smislu, pošto vrednost koja se pripisuje objektu ne govori ništa o količini svojstva, već samo imenuje (lat. *nomen*) klasu kojoj objekat pripada. To su varijable koje smo ranije nazvali kvalitativnim. Kao što smo rekli, čak i ako je vrednost varijable brojčana, to ne znači da je varijabla kvantitativna. Na primer, JMBG građana ili brojevi na dresu fudbalera jesu numeričke vrednosti, ali imaju samo funkciju označavanja, tj. klasifikovanja, i ne govore ništa o količini ili izraženosti svojstva. Varijable izmerene na *ordinalnom* ili *rang* nivou pružaju informaciju o tome da li je kod neke osobe određeno

svojstvo više ili manje izraženo, ali ne i o tome kolika je ta razlika. Na primer, rang trkača (1, 2, 3, 4...) jeste način da se numerički izrazi i uporedi njihova brzina, ali na osnovu tog broja nismo u mogućnosti da zaključimo za koliko je neki trkač brži od nekog drugog. Ukoliko nam vrednosti varijable omogućavaju da poredimo intervale, odnosno da utvrdimo ne samo da li je, već i za koliko je neko svojstvo izraženije kod jedne osobe nego kod druge, govorimo o *intervalnim* varijablama. Intervalne varijable, međutim, mere se skalom koja ne sadrži *apsolutnu nulu* kao vrednost koja označava potpuno odsustvo svojstva. IQ skala je tipičan primer varijable intervalnog nivoa jer ne sadrži nulu. Čak i ako neka osoba ne reši nijedan zadatak na testu sposobnosti, njen IQ neće dobiti nultu vrednost. Sa druge strane, vrednost 100, kao prosečan ili tipičan IQ, zapravo je određena proizvoljno, što znači da je konsenzusom čak i vrednost 0 mogla da se proglasi prosečnom, a negativne vrednosti ispodprosečnim. Slično je i sa temperaturom izraženom u stepenima Celzijusove skale koja sadrži nultu vrednost, ali ta tačka zapravo ne označava potpuno odsustvo temperature, već samo temperaturu na kojoj se zamrzava voda. Za razliku od intervalnih, *racio* ili *razmerne* skale poseduju apsolutnu nulu. To ne znači nužno da neki objekat može da dobije tu vrednost, već samo da nula postoji na skali kojom se svojstvo meri. Ne postoji osoba koja ima 0 kg, ali svojstvo *telesna težina* je varijabla razmernog nivoa merenja, jer na skali kojom se meri težina postoji nula koja označava i potpuno odsustvo tog svojstva. Razumevanje kriterijuma za podelu tipova varijabli na osnovu nivoa merenja veoma je važno zbog kasnijeg odabira prikladne tehnike vizualizacije, ali i odgovarajuće metode statističke obrade, jer se na podacima sa nižih nivoa merenja ne mogu i ne smeju obavljati određene matematičke operacije. Na primer, tek postojanje apsolutne nule, odnosno razmerni nivo merenja, omogućava da se na vrednostima neke varijable obavljaju operacije množenja i deljenja.

1.8.2. Hijerarhija vizuelnih kodova

Na osnovu nivoa merenja varijabli određuje se forma u kojoj će one biti prikazane grafički, odnosno način na koji će se *kodirati*. Francuski kartograf Žak Berten u svojoj knjizi *Semiologija grafikona* (Bertin, 1983) definiše osnovni skup atributa kojima vrednosti varijable mogu da se predstave na grafikonu: položaj, veličina, svetlina, tekstura, boja, orijentacija i oblik. Američki statističari Vilijem Klivlend i Robert Mekgil (Cleveland & McGill, 1984) dodali su ovom skupu i dužinu, ugao, površinu, zapreminu, zakrivljenje i zasićenost boje. Tako dobijena

lista nije samo popis vizuelnih karakteristika objekata, već i *hijerarhija* zadataka koje obavljamo prilikom tumačenja grafički prikazanih podataka, odnosno rang lista naše uspešnosti u poređenju objekata na osnovu navedenih atributa. Naime, Klivlend i Mekgil su nizom eksperimenata pokazali da su ispitanici uspešniji u proceni razlika u položaju ili dužini objekata, nego u proceni razlika u zapremini ili stepenu osenčenosti. Stoga preporučuju da se prilikom izrade grafikona najpre koriste svojstva koja se nalaze više u ovoj hijerarhiji. Prisetite se **primera sa stubičastim i torta dijagramom** i razmislite da li vam je bilo lakše da poredite frekvencije različitih kategorija đaka na osnovu visine stubića ili na osnovu površine, odnosno ugla odsečaka kružnice. Većina osoba lakše i tačnije procenjuje razlike u dužini objekata, nego razlike u njihovoj površini, zapremini ili uglu koji zahvataju (Slika 15). Treba, međutim, imati na umu da se zaključci Klivlenda i Mekgila odnose na kvantitativne varijable. Numeričke vrednosti ne bi mogle adekvatno da se predstave, na primer, različitom teksturom ili oblikom. Sa druge strane, kvalitativne varijable ne bi trebalo kodirati veličinom ili dužinom objekta, npr. tako što bi se studenti psihologije na grafikonu predstavili velikim, a studenti mašinstva malim krugovima. Preferirani atributi za vizuelno kodiranje nominalnih svojstava su boja i oblik. Boja je veoma korisna za kodiranje jer se uočava automatski, već na nivou *ranog opažanja* koje se odvija u deliću sekunde i ne zahteva angažovanje viših kognitivnih procesa (Treisman, 1986). Na primer, **na ranije prikazanom primeru sociograma**, veoma lako i brzo se prepoznaju i vizuelno grupišu krugovi različite boje, mnogo lakše nego krugovi različitog prečnika. Naravno, treba imati na umu ograničenje kapaciteta naše kratkotrajne memorije zbog koga bi upotreba većeg broja boja na grafikonu znatno usporila procesiranje informacija. Taj broj se obično kreće između 5 i 9, što čini poznati „magični broj" od 7±2 stimulusa koje možemo da pohranimo u radnoj memoriji (Miller, 1956). U kros-kulturnim studijama, četiri osnovne boje (crvena, zelena, žuta i plava) pokazale su se kao najprikladnije za kodiranje (Ware, 2004).



Slika 15. Grafičko kodiranje kvantitativnih vrednosti varijable: ugao, visina, zapremina i pozicija na zajedničkoj osi

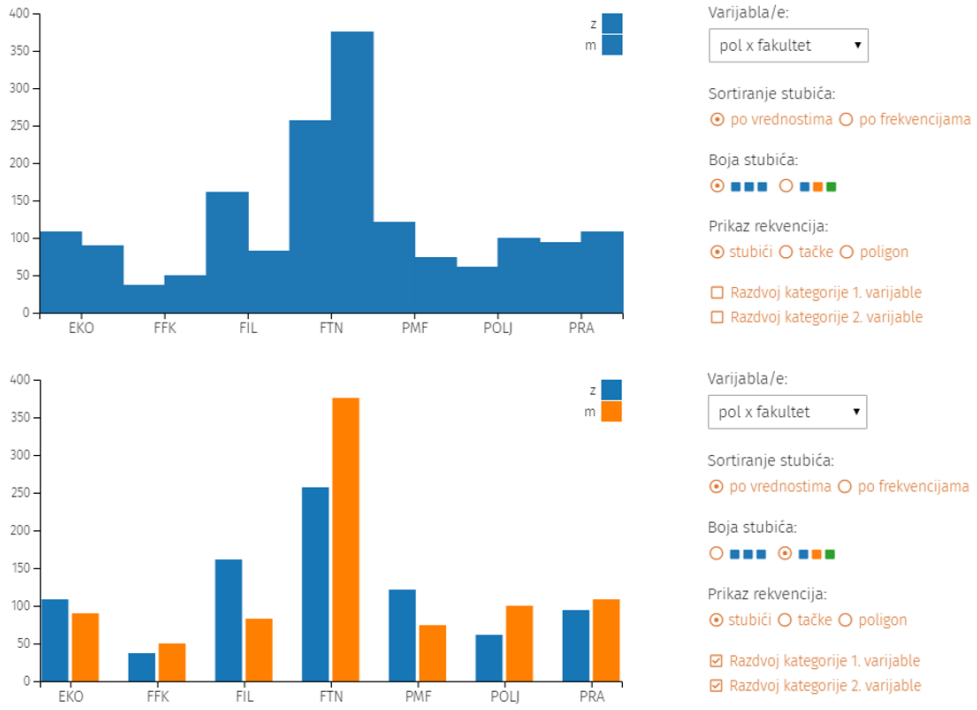
Fenomen ranog opažanja jedan je od značajnijih doprinosa psiholoških nauka oblasti vizualizacije informacija. Drugi bitan psihološki koncept vezan za temu ovog udžbenika, predstavljaju *Geštalt principi* koje su definisali nemački psiholozi Maks Verhajmer, Volfgang Keler i Kurt Kofka početkom 20. veka. U osnovi ovih principa je ideja da objekte uvek opažamo kao celinu a ne kao izolovane, pojedinačne delove (nem. *gestalt* – oblik, forma, šablon). Drugim rečima, u toku opažanja pokazujemo tendenciju da grafičke elemente organizujemo i grupišemo na osnovu njihovih vizuelnih karakteristika. Tako će, na primer, objekti koji su *bliski* biti opaženi kao grupa ili celina. Slično je i sa objektima koji su međusobno *slični* po boji, obliku, teksturi ili veličini. Prisetite se da na **primeru radar dijagrama** najverovatnije niste opažali izolovane pozicije tačaka na dimenzijama, već površine, odnosno profile kao celinu. U kontekstu primene vizuelnih kodova, to znači da različiti vizuelni atributi mogu da se upotrebe za kodiranje različitih varijabli, kako bi se korisnicima olakšalo opažanje pravilnosti i važnih karakteristika podataka. Pri tome je, naravno, potrebno poštovati ili iskoristiti razlike u hijerarhiji tih atributa. Na primer, boja ima jači efekat u kontekstu ranog opažanja i grupisanja u odnosu na oblik. Međutim, ova hijerarhija svojstava nije unapred određena i nepromenljiva, već zavisi od brojnih faktora, kao što su prethodno uputstvo koje korisnik dobije (Treisman & Gormican, 1988) ili varijacije u kontrastu i zasićenosti boja koje se koriste (Berg, Cornelissen, & Roerdink, 2008). Iako se ne pojavljuje u osnovnom setu Geštalt principa, zakon *homogene povezanosti* (Palmer & Rock, 1994) pokazao se kao snažan kriterijum grupisanja grafičkih elemenata koji, u određenim situacijama, ima čak i viši prioritet u odnosu na bliskost ili sličnost. Na primer, objekti koji su povezani linijama ili su uokvireni, biće opaženi kao grupa iako nisu bliski ili su različitog oblika i boje. Prilikom izbora linija kojima će se povezati elementi grafikona (npr. čvorovi na sociogramu) dolazi do izražaja još jedan Geštalt princip – zakon *kontinuiranosti*. Zaobljene krive linije efikasnije dočaravaju vezu između objekata od izlomljenih, jer je na grafikonu lakše pratiti njihov tok. S tim u vezi je i princip *zatvorenosti*. Ukoliko se dva kružna elementa na grafikonu preklapaju, u skladu sa principima kontinuiranosti (kružnice) i zatvorenosti (oblika), oba ćemo opaziti kao kompletne. Na kraju, veoma važan Geštalt princip koji dolazi do izražaja pri interpretaciji podataka na grafikonu je *simetričnost*. Simetrija je možda i ključni kriterijum procene estetskih aspekata vizualizacije. Uostalom, čak i lepotu osobe obično procenjujemo na osnovu simetričnosti crta njenog lica. Sposobnost da lako prepoznamo odstupanja po simetriji, u statistici nam pomaže da uočimo pravilnosti ili atipične karakteristike podataka. Na primer, stubičasti dijagram

koji prikazuje rezultate na teškom testu znanja izgledaće potpuno drugačije od onoga za test koji je bio veoma lak ili umereno težak. Prvi će imati više stubiće sa leve strane, a drugi sa desne, gde su veće vrednosti broja osvojenih poena.

1.8.3. Čitljivost grafikona

U prethodnim odeljcima doveli smo u vezu kriterijume čitljivosti grafikona sa statističkim konceptima kao što su nivoi merenja, ali i psihološkim fenomenima kao što su hijerarhija vizuelnih kodova i Geštalt principi. Opisane principe ilustrovaćemo **primerom podataka o grupi studenata** (Slika 16). O svakom studentu imamo podatak o polu, fakultetu na kome studira i broju bodova koje je osvojio na testu znanja iz statistike. Na početku je prikazana raspodela studenata po polu. Na grafikonu su prikazana dva stubića jer je varijabla *pol dihotomna*, što znači da ima dve moguće vrednosti ili dva *nivoa*. Iako su stubići potpuno spojeni, na osnovu njihove visine može se zaključiti da je prikazano nešto više od 1.700 podataka, oko 880 vrednosti m (muško) i oko 840 vrednosti z (žensko). Tačan broj studenata u ovom primeru iznosi 1.723, ali ga na osnovu stubičastog dijagrama nije moguće potpuno precizno odrediti. Za razliku od pola, varijabla *fakultet* je *politomna*, a u našem primeru ima sedam mogućih vrednosti: Ekonomski fakultet, Fakultet fizičke kulture, Filozofski fakultet, Fakultet tehničkih nauka, Prirodno-matematički fakultet, Poljoprivredni fakultet i Pravni fakultet. Prikažite varijablu *fakultet* i uočite da je najviše studenata upisano na FTN, a najmanje na FFK. Pošto su stubići spojeni, opažamo ih kao jedinstven nepravilan oblik, što otežava interpretaciju njihovih pojedinačnih visina. Čitljivost grafikona možemo da unapredimo upotrebom pomenutih vizuelnih kodova i Geštalt principa. Odaberite opciju za vizuelno razdvajanje studijskih grupa *Razdvoj kategorije 1. varijable*, a potom odaberite i drugu opciju za boju stubića. Kao što možete da primetite, vrednosti varijable na x-osi sortirane su abecednim redosledom, ali potpuno je opravdano sortirati ih po bilo kom drugom kriterijumu da bismo lakše poredili visine stubića. Odaberite opciju za sortiranje stubića prema visini, odnosno frekvenciji studenata u svakoj od kategorija. Slobodu da proizvoljno sortiramo stubiće imamo zato što je varijabla *fakultet* kvalitativna, odnosno nominalnog nivoa merenja. Međutim, ako prikažete varijablu *broj bodova*, dobićete stubičasti dijagram koji je neprihvatljiv, jer vrednosti na x-osi nisu raspoređene rastućim redosledom. U ovom primeru kriterijum za sortiranje moraju da budu vrednosti varijable, jer je u pitanju kvantitativno svojstvo razmernog nivoa merenja. Osim

toga, korišćenje različitih boja u ovom primeru nije opravdano zbog prevelikog broja kategorija. Drugim rečima, boja kao vizuelni kod postaje neinformativna i samo zbunjuje posmatrača. Ujednačite boju stubića i sortirajte ih po vrednostima, odnosno rastućem broju bodova. Sada mnogo lakše opažamo raspodelu varijable *broj bodova*. Uočavamo da se uspeh studenata kreće od 0 do 20 osvojenih bodova i da ih je najviše grupisano u rasponu od 10 do 15 bodova. Mogli bismo da zaključimo da su studenti, kao grupa, relativno dobro uradili test. Raspodele podataka na osnovu kojih se može zaključiti koje vrednosti varijable i koliko često se javljaju u nekoj grupi merenja, nazivaju se *distribucijama*. Distribucije mogu da se prikažu tabelarno ili grafički, ali najčešće je potrebno da se opišu i matematički. Analiza karakteristika distribucija podataka je prvi, a možda i najznačajniji korak u svakoj statističkoj obradi. Ovim pitanjima ćemo se baviti u poglavlju o distribucijama verovatnoća.



Slika 16. Primer teško i lako čitljivog stubičastog dijagrama

Do sada smo učestalosti različitih kategorija ispitanika, npr. onih koji su muškog pola, onih koji studiraju na Filozofskom fakultetu ili onih koji su osvojili 15 bodova na testu, označavali visinom stubića. Ponekad je, u cilju bolje

preglednosti grafikona, prikladnije vizualizovati distribucije pomoću tačaka čija pozicija na y-osi govori o učestalosti svake kategorije. Ovo je posebno korisno kada na istom grafikonu želimo da prikažemo vrednosti varijable za više grupa. Odaberite opciju *pol x broj bodova*. Ovoga puta su studenti podeljeni u dve grupe na osnovu pola, a distribucija broja bodova prikazana je posebno za svaku od grupa. Ipak, dve distribucije je teško vizuelno razdvojiti zbog toga što su stubići zbijeni i jednako obojeni. Obojite stubiće različitim bojama a potom odaberite opciju *Razdvoj kategorije 2. varijable*. Sada je jasnije da u okviru svake kategorije formirane na osnovu broja bodova, postoje dve kategorije pola. Na delu su principi sličnosti stubića (po polu) i njihove bliskosti (po vrednostima). Grafikon bismo mogli da učinimo još čitljivijim ako učestalosti po grupama prikažemo tačkama, ali tek primenom Geštalt principa zajedničke sudbine (tih tačaka), grafikon postaje potpuno razumljiv. Odaberite opciju *tačke* a potom i opciju *poligon* da biste povezali tačke linijom i kreirali dijagram koji je poznat kao *poligon frekvencija*. Poligon frekvencija koristi se isključivo za vizualizaciju kvantitativnih varijabli, a njegova početna i krajnja tačka uvek treba da dodiruju x-osu. Tek tada tačke formiraju zatvorenu izlomljenu liniju (poligon) na osnovu koje možemo da zaključimo kakve su karakteristike distribucije podataka. U našem primeru uočavamo da su studenti većinom bolje uradili test od studentkinja. Odaberite opciju *fakultet x pol* i uočite da poligon frekvencija nije prikladna tehnika vizualizacije u slučaju nominalnih varijabli. Ponovo prikažite stubiće da biste jasnije razlučili dva kriterijuma za grupisanje kategorija. Prvi je prostorna bliskost, na osnovu koje studente različitih fakulteta grupišemo po polu, a drugi je boja, na osnovu koje studente različitog pola grupišemo po tome šta studiraju. U ovom primeru bliskost je očigledno pozicionirana više u hijerarhiji kodova u odnosu na boju, jer je verovatnije da ćemo ovu sliku opisati kao dva, a ne kao sedam skupova stubića. Isključite i ponovo uključite opcije za razdvajanje kategorija da biste analizirali kako na čitljivost grafikona utiče primena Geštalt principa bliskosti. Ako isključite sve opcije razdvajanja kategorija i opciju različite obojenosti stubića, dijagram postaje nečitljiv. Iste podatke možete da vizualizujete i odabirom opcije *pol x fakultet*, ali ovoga puta grupisanje je najpre obavljeno na osnovu fakulteta na kome student studira, a tek u drugom koraku na osnovu pola studenta.

> Koji zaključak ćete lakše doneti na osnovu grafikona *pol x fakultet* a koji na osnovu grafikona *fakultet x pol*?

Da li je opravdano upotrebiti poligone frekvencija u primeru *pol x fakultet*? Sortirajte vrednosti po frekvencijama dok su poligoni vidljivi.

Menjajte opcije za generisanje grafikona i utvrdite kada je opravdano a kada ne, sortiranje po učestalostima, različita obojenost stubića i prikazivanje tačaka, odnosno iscrtavanje poligona.

VIZUALIZACIJA DISTRIBUCIJA VEROVATNOĆA

Statistički udžbenici po pravilu sadrže najmanje jedno poglavlje posvećeno osnovama teorije verovatnoće i teorije skupova. Ove dve oblasti matematike veoma su važne za razumevanje statističkih pojmova i procedura, posebno onih koje se tiču predviđanja ishoda događaja i izvođenja zaključaka. U skladu sa obećanjem datim na početku, u ovom udžbeniku ćemo prikriti matematiku (koliko je to moguće), a osnovne postavke vezane za teorije verovatnoće i skupova objasnićemo u nekoliko narednih odeljaka koji se bave tehnikama uzorkovanja i deskriptivne statistike, tj. postupcima prikupljanja, sažimanja i opisivanja podataka. Logike teorije verovatnoće i teorije skupova su sveprisutne i mnogima već poznate, barem na nekom osnovnom, intuitivnom nivou. Na primer, često koristimo opciju dovršavanja reči prilikom upotrebe internet pretraživača. Ako u okvir za unos teksta nekog pretraživača ukucate nisku *stat*, pojaviće se lista ponuđenih ključnih reči na kojoj će možda biti i reč *statistika* ali ne iznad reči *status*, jer su statusi na društvenim mrežama popularnija tema od varijabli i grafikona. Ukoliko pretraživač nema dodatne podatke o vama, vi (p)ostajete tipičan korisnik interneta, tako da vam se nudi ono što korisnici najčešće, odnosno najverovatnije traže. Dakle, pretraživač procenjuje *verovatnoću* onoga što želite da unesete na osnovu učestalosti upita velikog broja korisnika, polazeći od jednostavne pretpostavke da najverovatnije tražite ono što je tražila većina ljudi pre vas. Međutim, ukoliko ste u toku pretrage prijavljeni kao Branko Simić, građevinski inženjer, proste verovatnoće postaju manje korisne i zamenjuju ih *uslovne verovatnoće,* prilagođene činjenici da je ispunjen neki uslov, da su neki podaci unapred poznati, odnosno da su se neki događaji već desili. U ovom primeru to je podatak da ste vi građevinski a ne npr. mašinski inženjer, što povećava uslovnu verovatnoću da vas interesuje *statika materijala*, a smanjuje verovatnoću da vas interesuju *statori*. Tada se verovatnoća ne procenjuje na osnovu učestalosti pretraga svih korisnika, već na osnovu *podskupa* pretraga koje su obavili drugi građevinski inženjeri ili, u nekoj drugoj situaciji, korisnici vašeg uzrasta i/ili pola, korisnici iz mesta u kome živite, korisnici koji imaju istu marku mobilnog telefona, a zapravo najčešće vi sami u dotadašnjem korišćenju istog pretraživača. Naravno, iako je verovatnoća da ćete nakon *stat* ukucati nisku *istika* ili *us* drastično veća, može se desiti da je vaš

cilj da pronađete informacije o *statiranju u filmovima*. Ukucali ste reči, pritisnuli enter i na vrhu liste rezultata pretrage, opet na osnovu principa verovatnoće, nalaziće se sajtovi koji su vam najverovatnije korisni, tj. oni koje je posetilo najviše osoba sličnih interesovanja, oni koji su najbolje ocenjeni, oni na koje vodi najviše linkova sa drugih sajtova itd. Na kraju, dolazite do stranice na kojoj se nude stotine potencijalnih angažmana za statiste, tako da vam je potrebna pomoć teorije skupova. Označavate opcije i izjašnjavate se da želite da statirate *na otvorenom*, negde *u Vojvodini*, u filmu sa *stranim producentima* i sl. Na taj način formirate presek podskupova velikog skupa ponuda, čime vam je bitno olakšano pronalaženje potrebne informacije. Može se reći da matematički i statistički postupci omogućavaju da se velike količine podataka sažmu, istraže i pretvore u informacije na osnovu koji se, sa manjom ili većom pouzdanošću i tačnošću, mogu predvideti ishodi različitih događaja.

2.1. Pojam verovatnoće

Do sada smo više puta pominjali pojam verovatnoće oslanjajući se na njegovo kolokvijalno značenje i na intuitivnost, odnosno na opštu informisanost čitaoca. Čak je i opravdanje za ovakvo očekivanje autora povezano sa pojmom verovatnoće, jer su „velike šanse" da osoba koja čita ovaj udžbenik gotovo svakoga dana pravi ili koristi procene verovatnoće, tj. procene mogućnosti da se desi neki događaj. Te procene su nekada zasnovane na subjektivnom utisku (npr. „Imam (pred)osećaj da ćemo ih pobediti u utakmici!"), nekada na pređašnjem iskustvu u sličnim situacijama (npr. „Tamo ćeš 100% naći mesto za parkiranje."), a nekada na egzaktnim podacima o povezanosti određenih pojava (npr. „Polje niskog vazdušnog pritiska pomera se prema našoj zemlji, što će dovesti do padavina u severnim krajevima."). U statistici, naravno, brojčano izražavanje verovatnoće mora da bude objektivnije i preciznije. To možda nije bilo dovoljno očigledno iz dosadašnjih primera, ali kvantifikacija verovatnoće bazira se na učestalostima različitih događaja. Na primer, vrlo je verovatno da su slike lansiranja iranskih raketa krivotvorene, jer se identični oblici dima slučajno pojavljuju *veoma retko* ili *nikada*. Velika je verovatnoća da pumpa za vodu nije higijenski ispravna zato što su osobe koje su živele u njenoj blizini *znatno češće* oboljevale od kolere. Kada u okvir za unos teksta pretraživača unesete reč *Michael*, na vrhu liste će vam najverovatnije biti ponuđeni nastavci *Jackson*, *Douglas*, *Jordan* ili *Fassbender* zato što korisnici interneta *mnogo češće* traže podatke o tim osobama nego o Majku Bostoku.

Zamislimo da niste znali odgovor na neko od pitanja iz **testa znanja o piktogramima**. Kolika je verovatnoća da slučajno pogodite tačan odgovor? U teoriji verovatnoće, vaše davanje odgovora predstavlja jedan *eksperiment*. Eksperiment može da bude bilo koja aktivnost, kao što je bacanje kockice u igri *Jamb* ili istraživački eksperiment koji hemičar sprovodi u svojoj laboratoriji. Na svako pitanje bila su ponuđena četiri odgovora, što znači da je eksperiment mogao da ima četiri *ishoda*. Skup svih mogućih ishoda naziva se *prostor uzorka*. Nas interesuje verovatnoća jednog *događaja*, a to je izbor tačnog odgovora. Događaj je skup svih poželjnih ishoda, koji u našem primeru sadrži samo jedan element. U tom slučaju, verovatnoću slučajnog pogađanja tačnog odgovora izrazićemo kao odnos broja poželjnih ishoda i ukupnog broja ishoda:

$$\mathrm{p} = \frac{\text{broj poželjnih ishoda}}{\text{broj mogućih ishoda}} = \frac{1}{4} = 0{,}25$$

Obratite pažnju na činjenicu da smo verovatnoću označili malim latiničnim slovom p od engleskog *probability*, ali i od *proportion*, jer verovatnoća nije ništa drugo nego odnos dve učestalosti, tj. proporcija ili udeo poželjnih ishoda u skupu svih ishoda koji su mogli da se dese. Dakle, verovatnoća da ćete slučajno pogoditi tačan odgovor iznosi 0,25, što možemo da izrazimo i procentima ako dobijenu proporciju pomnožimo sa 100. Iako često umemo da kažemo da smo 200% sigurni u nešto, vrednosti p mogu da se kreću samo u intervalu od 0 do 1 ili od 0% do 100%.

> Kolika je verovatnoća slučajnog pogađanja tačnog odgovora na pitanja koja sadrže samo dve opcije – tačno i netačno?
>
> Kolika je verovatnoća pogađanja tačnog odgovora ako vam je poznato da jedan od četiri ponuđena odgovora nije tačan?
>
> Kolika je verovatnoća pogađanja tačnog odgovora ako vam je poznato da jedan od četiri ponuđena odgovora nije tačan i ako postoje dva odgovora koja se priznaju kao tačna, tj. ako događaj ima dva poželjna ishoda?

U prirodi i društvu postoje događaji koji su potpuno izvesni. Međutim, u statističkom zaključivanju nikada nećete moći da budete 100% sigurni u svoju procenu. Ako ipak jeste, onda se bavite fenomenima koji ne zavređuju pažnju istraživača i nisu naučno relevantni, jer statistika se bavi *probabilističkim* fenomenima čije ishode nije moguće predvideti sa potpunom sigurnošću.

Nasuprot tome, *deterministički* fenomeni su oni koji su potpuno predvidivi, ali takođe i izuzetno retki, naročito u društvenim naukama kao što je psihologija. Ipak, olakšavajuću okolnost za naučnike i statističare predstavlja činjenica da mogu da budu zadovoljni i nivoima sigurnosti, tj. poverenja u svoje ili tuđe zaključke koji su niži od 100%. Uostalom, to radimo i u svakodnevnom životu. Ako vremenska prognoza najavljuje 80% verovatnoće padavina, većina nas će poneti kišobran ili smisliti način da do cilja stignete sa suvom odećom na sebi. Naravno, interpretacija nivoa verovatnoće ima i određeni subjektivni momenat. Ako neki događaj nije potpuno izvestan, obično vagamo između verovatnoće jednog ishoda i verovatnoće drugih mogućih ishoda. Tipičan primer su igre na sreću koje privlače veliki broj osoba spremnih da svoja velika očekivanja baziraju na izuzetno malim verovatnoćama dobitka. Student koji nije naučio kompletno gradivo veruje u verovatnoću da neće izvući pitanje koje ne zna, mada se često desi upravo to. Slično je i u nauci. Istraživač mora da prihvati rizik da rezultat njegovog istraživanja možda nije tačan, ali, za razliku od svakodnevnih i subjektivnih situacija, taj rizik mora da svede na najmanji mogući nivo. Ovo je ključna karakteristika statističkog zaključivanja koja se često zanemaruje, jer većina ljudi ipak teži da izbegne neizvesnost i pokušava da bude ubedljivija u odbrani svojih stavova. Na primer, poruka „Pušenje ubija!" na paklici cigareta implicira uzročno-posledičnu vezu između pušenja i smrtnog ishoda. Stoga će, na primer, strastveni pušač koji poznaje drugog strastvenog pušača starog 80 godina, reći da ona nije tačna. Ali to nije korektan (statistički) način razmišljanja. Na paklicu cigareta ne možemo da smestimo rečenicu: „Pušenje značajno povećava verovatnoću oboljevanja od karcinoma pluća ili grla koji, ukoliko se ne dijagnostikuju na vreme i ne leče na adekvatan način, u većini slučajeva dovode do smrtnog ishoda". Ono što možemo da uradimo, jeste da unapredimo svoju statističku pismenost da bismo bolje razumeli i tačnije interpretirali podatke i informacije kojima smo svakodnevno izloženi.

2.2. Populacija i uzorak

U sredstvima javnog informisanja često nailazimo na izveštaje o gledanosti različitih emisija i filmova u određenoj državi, regionu ili celom svetu. Pri tome se retko zapitamo kako su ti podaci prikupljeni i obrađeni, jer sabiranje broja prodatih ulaznica ili prebrojavanje korisnika prijavljenih na neki internet servis za emitovanje sadržaja (engl. *streaming*) ne zvuči kao preterano zahtevan statistički poduhvat. Naravno, pod uslovom da su kompanije voljne i da smeju

da dele svoje podatke. Ali čak ni tada procena gledanosti ne svodi se na primenu osnovnih matematičkih operacija. Uzmimo kao primer informaciju da je gledanost jedne televizijske emisije u nekoj državi bila 25%, što znači da ju je (navodno) gledala četvrtina ukupnog broja stanovnika. Procentualni deo neke vrednosti najlakše se izračunava tako što se traženi procenat podeli sa 100 i pretvori u proporciju, a potom se proporcija pomnoži sa datom vrednošću. U našem primeru, ako država ima 6 miliona stanovnika, 25% te vrednosti je isto što i proporcija od 0,25 (25 : 100), što znači da je emisiju pratilo 6 · 0,25 ili 1,5 miliona gledalaca. Svima je jasno da anketari agencije koja je saopštila navedenu informaciju nisu išli od stanovnika do stanovnika, niti su svakome od njih postavljali pitanja telefonom ili poštom, jer bi taj postupak trajao veoma dugo i ne bi bio ekonomski isplativ. Međutim, ono što je bitnije za temu ovog poglavlja jeste da taj postupak ne bi bio ni potreban. Naime, dovoljno pouzdani zaključci o većem skupu entiteta mogu da se donesu i na osnovu posmatranja njegovog manjeg podskupa. U nauci se zaključci i pretpostavke (inferencije) o *populaciji*, najčešće donose na osnovu posmatranja karakteristika *uzorka* uzetog iz te populacije. Populacija je, dakle, skup svih entiteta o kojima želimo da donesemo neki zaključak i na koje će se odnositi naša pretpostavka ili rezultat, a uzorak je samo jedan njen podskup, odnosno deo koji nam je u istraživanju bio dostupan za merenje i analizu. Za označavanje populacije ponekad se koristi i termin *univerzum*, čime se posebno naglašava činjenica da je u pitanju teorijski neograničen skup entiteta. Neograničene ili neprebrojive populacije postoje samo u teoriji, odnosno u hipotetičkim situacijama, kao što je npr. populacija svih bacanja novčića ili kockica. Međutim, suština neograničenosti je u tome što čak i prebrojive populacije, za koje bismo mogli da odredimo ili saznamo konačan broj članova, u praksi postaju potpuno „neuhvatljive". Ako biste hteli da postavite pitanje, ili da nešto izmerite svim korisnicima neke društvene mreže ili svim ljudima koji imaju alergiju na polen, veličina i sastav ciljne populacije znatno bi se izmenili već nakon prvih par dana prikupljanja podataka. Naravno, postoje situacije u kojima je relativno lako pristupiti svim članovima populacije, npr. ako je definišete kao pacijente nekog kliničkog centra obolele od određene bolesti ili đake neke osnovne škole. Ali u tom slučaju bi vaši zaključci bili prilično ograničeni i odnosili bi se samo na članove te populacije – pacijente *tog* kliničkog centra i đake *te* škole. Međutim, suština primene statistike u naučnim istraživanjima jeste izvođenje zaključaka koji se mogu uopštiti, tj. *generalizovati* na znatno veći broj slučajeva.

2.2.1. Tehnike uzorkovanja

Pre nego što se pristupi uzorkovanju, populacija i kriterijumi pripadnosti na osnovu kojih se odlučuje da li je neki entitet deo te populacije, moraju da budu definisani iscrpno, nedvosmisleno i koncizno. U zavisnosti od ciljeva istraživanja, populaciju mogu da čine npr. svi građani Vojvodine, studenti završnih godina tehničkih fakulteta u Srbiji, klijenti svih filijala neke banke starosti između 30 i 40 godina, osobe koje su u toku prethodnih pet godina barem jednom zatražile savet psihologa, ali i svi automobili u nekom gradu, svi zasadi kukuruza u nekoj opštini ili sve bombonjere proizvedene u nekoj fabrici. U našem primeru o gledanosti emisija, populaciju bi činili svi građani neke države koji su stariji od četiri godine i imaju pristup TV prijemniku u svom domaćinstvu. Da bi nešto zaključile o ponašanju gledalaca koji čine tu populaciju, agencije prikupljaju podatke na uzorku domaćinstava u kojima su instalirani tzv. *piplmetri*, uređaji koji beleže i šalju informacije o tome na kom televizijskom programu, u kom terminu i koliko dugo su se članovi odabranih domaćinstava zadržavali. Broj tih domaćinstava u državama veličine Srbije ne prelazi 1.000, što znači da neku emisiju zapravo nije gledalo 1,5 od 6 miliona stanovnika (članova populacije), već npr. 300 od 1.200 ispitanika (članova uzorka) ili 150 od 600 ispitanika koji su u tom trenutku imali uključen televizor. Drugim rečima, obrasci uočeni na uzorku, pripisuju se celoj populaciji i uopštavaju na sve njene članove. Prilikom generalizacije zaključaka sa nekoliko stotina na nekoliko miliona ljudi, agencije se pouzdaju u *reprezentativnost* svog uzorka. Uzorak smatramo reprezentativnim ukoliko verno odražava sve bitne karakteristike populacije, odnosno sve one varijable koje na neki način mogu da budu povezane ili da utiču na ishode merenja. U našem primeru te varijable bi mogle da budu pol, uzrast, stepen obrazovanja, materijalni status i sve druge karakteristike osoba za koje se očekuje da su na neki način povezane sa time šta vole da prate na televiziji. Reprezentativan uzorak bi trebalo da bude odraz raznolikosti gledalaca u populaciji, a taj odraz se najbolje pravi ukoliko se sve prepusti zakonima verovatnoće. Stoga se najpoželjnijom tehnikom uzorkovanja smatra *jednostavno nasumično uzorkovanje* koje podrazumeva da se izbor članova uzorka iz populacije vrši potpuno nasumično, bez nekog reda i pravilnosti. To znači da ukoliko u populaciji ima više žena od muškaraca ili više osoba mlađih od 20 godina nego starijih, verovatnoća da će takve osobe biti nasumično odabirane postaje veća. Samim tim, njih će biti *proporcionalno* više i u uzorku. Postupak nasumičnog uzorkovanja podrazumeva postojanje *okvira*

za uzorkovanje, tj. iscrpnog popisa svih članova populacije, sa koga se nasumičnim izborom rednih brojeva određuju članovi koji će formirati uzorak. Alternativno, sa popisa može da se nasumično odabere samo jedan član, a nakon njega svaki n-ti, npr. svaki deseti. Tada govorimo o tehnici *sistematskog uzorkovanja*.

Upravo opisane tehnike uzorkovanja obično se nazivaju *verovatnosnim*, jer u postupku selekcije svaki član populacije ima istu verovatnoću da dospe u uzorak. U ovu grupu tehnika spadaju još i *stratifikovano* i *klasterisano* uzorkovanje. Kod prvog se populacija najpre deli na *stratume* (slojeve) iz kojih se potom nasumično izdvajaju članovi, ali tako da je proporcija svakog stratuma u uzorku približno ista kao proporcija tog stratuma u populaciji. Ukoliko, na primer, želimo da sprovedemo istraživanje koje se odnosi na studente nekog univerziteta, potrudićemo se da u uzorku budu zastupljeni studenti svih fakulteta tog univerziteta i to u proporciji koja odgovara stvarnom broju studenata na svakom fakultetu. Na taj način fakulteti sa više studenata imaće i više svojih predstavnika u uzorku, čime se obezbeđuje veća preciznost, posebno u situacijama kada je uzorak manji, pa se ne može očekivati da nakon jednostavnog nasumičnog uzorkovanja svi stratumi populacije budu pravično zastupljeni. Sa druge strane, kod klasterisanog uzorkovanja najpre se nasumično odabira klaster, tj. podgrupa neke populacije, nakon čega se u uzorak uključuju svi članovi tog klastera. Na primer, ukoliko želimo da saznamo nešto o učenicima srednjih škola u nekom gradu, možemo nasumično da odaberemo nekoliko škola, a zatim da ispitamo sve đake odabranih škola.

Iako verovatnosne tehnike obezbeđuju bolju reprezentativnost uzorka, u istraživanjima se veoma često koriste i tzv. *neverovatnosne* metode kod kojih se članovi populacije biraju na manje ili više pristrasan način, tako da neki od njih imaju veću a neki manju verovatnoću da dospeju u uzorak. Očigledno je da reprezentativnost uzorka u ovakvim situacijama postaje diskutabilna, a uopštavanje zaključaka na celu populaciju manje opravdano. Razlozi za primenu neverovatnosnih tehnika uzorkovanja mogu da budu objektivni, kao što su nedostatak preciznog okvira za uzorkovanje ili nemogućnost pristupa određenim stratumima populacije, ali su mnogo češće subjektivni, odnosno vođeni namerom da se istraživanje obavi na brži, jednostavniji i jeftiniji način. Tipičan primer su *prigodni* uzorci koji se formiraju od članova populacije koji su istraživaču najlakše dostupni ili nisu u poziciji da odbiju učešće u istraživanju. Primer prigodnog i pristrasnog uzorkovanja bi bilo prikupljanje podataka o gledanosti televizije preko digitalnog prijemnika nekog kablovskog operatera

ili anketiranje klijenata banke preko neke društvene mreže. Nekada se prigodni uzorci „poboljšavaju" primenom uzorkovanja tipa *snežne grudve* ili *lančanog* uzorkovanja, kada se od početno formirane grupe ispitanika traži da regrutuju nove ispitanike. Ovo je i dalje neverovatnosno uzorkovanje, jer veću verovatnoću ulaska u uzorak imaju osobe koje su bliskije i sličnije članovima inicijalne grupe. Treba imati na umu da pristrasnost u uzorkovanju može da ima ozbiljne posledice na validnost rezultata istraživanja. U psihologiji, na primer, već nekoliko decenija traje polemika o tome koliko je opravdano uopštavati zaključke o psihičkim fenomenima, ako se zna da su oni u ogromnom broju slučajeva donošeni na osnovu uzoraka sačinjenih prvenstveno od studenata početnih godina psihologije (Henrich, Heine, & Norenzayan, 2010; Shen et al., 2011; Sherman, Buddie, Dragan, End, & Finney, 1999; Smart, 1966). Sve češća upotreba internet platformi za prikupljanje podataka dodatno intenzivira ovu diskusiju, jer istraživač nikada ne može da bude potpuno siguran ko je, koliko puta i pod kojim uslovima popunio neki elektronski upitnik (Sharpe & Poets, 2017). Na kraju, čak i ako ove kritike ne shvatimo ozbiljno, veća dostupnost i globalizacija obrazovanja, nauke, znanja i informacija, dovela je do toga da uzorci u većini istraživanja deluju „čudno" iz aspekta država neengleskog govornog područja, posebno onih sa azijskog i afričkog kontinenta, jer su u najvećoj meri uzimani iz tzv. *WEIRD* populacija (engl. *Western, Educated, Industrialized, Rich, and Democratic*) (Henrich et al., 2010). Ovo izvrdavanje zakona verovatnoće jedan je od razloga trenutne *krize ponovljivosti* psihološke nauke, odnosno nemogućnosti da se u ponovljenim istraživanjima dođe do istog zaključka (Open Science Collaboration, 2015). Stoga je veoma važno da se u istraživanju jasno i objektivno naznači na koju populaciju se odnose zaključci doneti na osnovnu uzorka, odnosno za koju grupu entiteta taj uzorak može da se smatra reprezentativnim.

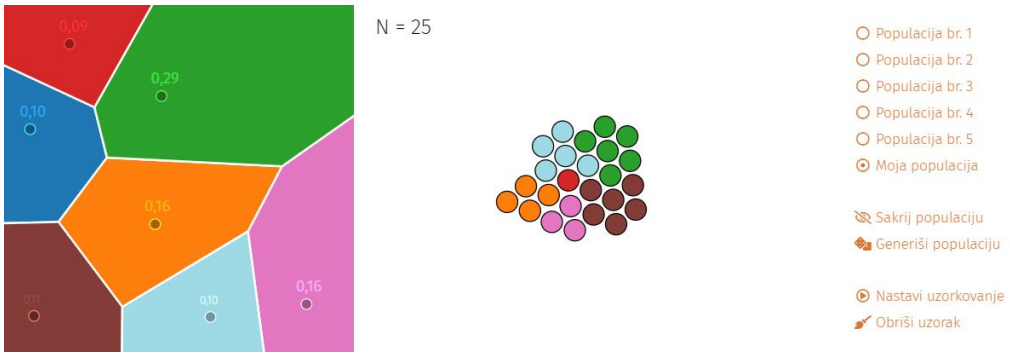
2.3. Pojam nasumičnosti ili slučajnosti

U prethodnom odeljku koristili smo termin *nasumično* za označavanje radnji i događaja koji se dešavaju bez vidljivog reda i namere, tj. za označavanje procesa čiji ishod nije moguće predvideti sa potpunom sigurnošću. Na primer, prilikom nasumičnog odabira člana neke raznovrsne populacije ljudi, nećete moći unapred da znate da li će to biti muška ili ženska, starija ili mlađa, viša ili niža, zaposlena ili nezaposlena osoba. Za označavanje nasumičnih procesa u engleskom jeziku koristi se termin *random*, koji se u domaćoj literaturi obično

prevodi kao *slučajno*. Samim tim se i teorija verovatnoće definiše kao oblast matematike koja se bavi ishodima *slučajnih* varijabli (engl. *random variables*) ili *stohastičkih* procesa kao skupa većeg broja slučajnih varijabli. Kako bismo bili dosledni sa postojećom literaturom, termine *slučajno* i *nasumično* koristićemo kao sinonime, ali predlažemo čitaocu da *nasumično* prihvati kao precizniji i prikladniji termin. U srpskom jeziku, pojam *slučajno* često ima drugačiju konotaciju i ne odražava na pravi način suštinu izvorne engleske reči. Naime, reč *random* svojim korenom i etimološkim poreklom upućuje na vezu sa trčanjem (engl. *run*, nem. *rennen*, st. fr. *randir*), tačnije sa procesima koji traju ili se ponavljaju. Upravo to je ključna karakteristika fenomena kojima se bavi i statistika. Ako neko dete *slučajno*, dakle bez namere, udari druga, sigurno neće privući pažnju školskog psihologa. Ali ako ono *nasumično* udari nekoga, to može da ukaže na postojanje namere. Štaviše, ukoliko se ti postupci ponavljaju, moguće je zaključiti da postoji „pozadinska" pravilnost u njegovom ponašanju. Psiholog i dalje neće moći da predvidi koga će dete da udari i kada, ali određeni nevidljivi obrazac postaje uočljiv nakon većeg broja ponavljanja događaja (engl. *long-run*). Uočeni obrazac ukazuje na potencijalne probleme ili karakteristike tog (agresivnog) deteta. Slično tome, nemoguće je predvideti da li će prilikom bacanja novčića pasti pismo ili glava, ali sa razlogom očekujemo da će nakon više ponovljenih bacanja, broj pisama i glava biti podjednak. U narednom primeru ilustrovaćemo ovaj fenomen, odnosno svojstvo stohastičkih procesa poznato kao *zakon velikih brojeva*.

Kao vizuelnu ilustraciju postupka uzorkovanja upotrebićemo **Voronojev dijagram**, nazvan prema ukrajinskom matematičaru Georgiju Fedosijeviču Voronoju koji ga je prvi opisao i definisao (Slika 17). Ovaj dijagram služi za podelu površine na više segmenata definisanjem težišnih tačaka u dvodimenzionalnom prostoru. Svaki segment površine, ili *Voronojeva ćelija*, predstavlja skup tačaka koje se nalaze bliže težištu te ćelije nego težištima drugih ćelija. Pomoću Voronojevog dijagrama, na veoma intuitivan način, mogu da se predstave veličine klastera i pozicije njihovih centralnih tačaka ili *centroida*. To mogu da budu zgrade koje gravitiraju ka repetitorima mobilnog operatera ili korisnici (zaražene) pumpe za vodu, kao u ranije opisanom primeru iz 19. veka. Isto tako, centroidi mogu da budu i tipični predstavnici nekih kategorija osoba, npr. studenata različitih fakulteta ili osoba koje slušaju određenu vrstu muzike. U takvim situacijama, složenijim statističkim postupcima, kao što je npr. *klaster analiza*, moguće je odrediti pozicije i karakteristike centroida čak i na osnovu vrednosti većeg broja varijabli, odnosno

svojstava grupa ispitanika. Međutim, tačna pozicija centroida u našem primeru nije bitna, jer se njome određuje samo veličina i položaj segmenata (stratuma) ukupne površine koju ćemo tretirati kao populaciju. Na početku je populacija izdeljena na sedam približno jednakih delova, od kojih svaki zauzima oko 15% ili 0,15 delova kvadrata. Zamislimo da je kriterijum za kategorizaciju bila nominalna varijabla, npr. fakultet koji osoba studira ili marka telefona koji koristi. Kliknite na taster *Počni uzorkovanje* da biste započeli proces formiranja jednostavnog nasumičnog uzorka studenata ili korisnika mobilnih telefona. Nasumičnost se ogleda u činjenici da potpuno nepristrasno birate članove populacije, odnosno da ne možete sa sigurnošću da predvidite koje boje će biti naredna izvučena loptica. Nakon 25 odabranih loptica, proces se pauzira.



Slika 17. Proporcionalni prikaz stratuma populacije uz pomoć Voronojevog dijagrama

> Kakav odnos (proporciju) broja kuglica očekujete u uzorku na osnovu izgleda populacije?
>
> Da li proporcije stratuma u uzorku veličine 25 verno odražavaju proporcije koje možete da uočite u populaciji?
>
> Kliknite taster *Obriši uzorak* i formirajte novi uzorak veličine 25. Da li su proporcije kružića različitih boja u drugom uzorku iste kao u prvom?

Kliknite taster *Nastavi uzorkovanje*. Izvlačenje kuglica nastavlja se dok njihov broj ne dostigne 100. Obratite pažnju na to da je broj entiteta u uzorku, tačnije veličina uzorka, označena velikim latiničnim slovom N. Nekada se sa N označava veličina populacije, a malim n veličina uzorka ili veličina poduzorka većeg uzorka koji čini N elemenata. U ovom udžbeniku, koristićemo oznaku N

za označavanje veličine uzorka, uz pretpostavku da tačna veličina populacije najčešće nije ni poznata.

> Da li se reprezentativnost uzorka popravila nakon što je njegova veličina povećana?
>
> Formirajte još nekoliko uzoraka veličine 100. Da li se uzorci veličine 100 međusobno više ili manje razlikuju od uzoraka veličine 25?

Kliknite taster *Sakrij populaciju* kako biste simulirali činjenicu da istraživač često ne zna kakve su prave karakteristike skupa entiteta iz koga uzima uzorak. Nakon toga odaberite opciju *Populacija br. 2* i pokrenite uzorkovanje do veličine uzorka od 25.

> Da li na osnovu uzorka veličine 25 možete da zaključite koliko stratuma postoji u populaciji?
>
> Da li na osnovu uzorka veličine 25 možete da zaključite koji od stratuma u populaciji je proporcionalno najmanji a koji najveći?

Dovršite uzorkovanje i otkrijte populaciju. Proporcije boja u skupu od 100 loptica trebalo bi verno da odražavaju odnos proporcija površina različitih boja na kvadratu. Ti odnosi verovatno nisu potpuno identični, ali nam uzorak pruža dovoljno precizne i korisne informacije o populaciji. I dalje ne možemo da pretpostavimo koje će boje biti naredna loptica uzeta iz populacije, ali sada imamo podatke na osnovu kojih možemo da procenimo verovatnoću tog ishoda, tj. da kažemo da je najverovatnije da će loptica biti zelene boje. Ta verovatnoća je 0,33 (33%) i odražava udeo zelenih loptica u ukupnom broju loptica u populaciji. Obratite pažnju na to da verovatnoća izračunata na osnovu odnosa loptica u uzorku ne mora da bude ista kao ona koju bismo izračunali na osnovu stanja u populaciji. Međutim, s obzirom na činjenicu da je primenjeno jednostavno nasumično uzorkovanje, te verovatnoće su veoma slične.

> Ponovite postupak skrivanja i uzorkovanja sa populacijama 3 i 4. Da li vam je bilo lakše i da li ste bili tačniji u proceni odnosa verovatnoća različitih ishoda na osnovu uzoraka uzetih iz Populacije 3 ili iz Populacije 4?

> U kojoj od ove dve populacije je raznolikost entiteta, tj. boja veća?

2.4. Pojam varijabilnosti

Veoma važno svojstvo svakog skupa podataka je njihova raznolikost, različitost, raspršenje ili *varijabilnost*. Varijabilnost nam govori o tome u kolikoj meri se entiteti međusobno razlikuju s obzirom na vrednost neke varijable. Na primer, visina u grupi predškolske dece nije toliko varijabilna koliko bi bila u grupi predškolske dece, dece osnovnoškolskog uzrasta, srednjoškolaca i studenata posmatranih zajedno. Nacionalnost kao svojstvo građana nekog regiona više varira u Vojvodini nego u Šumadiji. Ako se vratimo na **primere sa Voronojevim dijagramima,** poređenje varijabilnosti svojstva prikazanog na grafikonima za populacije 1 i 3, odnosno varijabilnosti uzoraka uzetih iz tih populacija, ne bi trebalo da bude težak zadatak čak ni za statističkog laika. U Populaciji 1 postoji sedam kategorija (klasa) varijable, pa su tako i loptice međusobno različitije po boji u odnosu na stanje u Populaciji 3 koja ima pet stratuma. U skladu sa tom logikom, možemo da zaključimo da bi najmanju moguću varijabilnost predstavljala situacija u kojoj su sve loptice iste boje, a površina kvadrata je jednobojna. Na primer, u populaciji studenata psihologije nema varijabilnosti svojstva *studijska grupa*. U tom slučaju, verovatnoća da ćete izvući kuglicu boje kojom su označeni budući psiholozi bila bi 1, a verovatnoća da ćete izvući kuglicu neke druge boje iznosi 0. To naravno ne znači da je boja, odnosno studijska grupa, prestala da bude varijabla i postala *konstanta*, već samo da ta varijabla uopšte ne varira u populaciji. Niska varijabilnost pojave koju smo izmerili mogla bi da ukaže na to da je odabrani uzorak pristrasan i da ne odslikava stvarno stanje u populaciji ili, pak, da je istraživač odlučio da se bavi fenomenom koji nije interesantan niti relevantan za analizu. Na primer, ukoliko svi stanovnici nekog grada voze automobile ne starije od tri godine, onda starost automobila najverovatnije neće biti relevantna za istraživanje uzroka povećanog zagađenja vazduha u tom gradu. To će pre biti učestalost vožnje, preovlađujuća vrsta goriva, prohodnost puteva ili nešto drugo.

Poređenje varijabilnosti populacija 3 i 4 (možda) predstavlja nešto teži zadatak. Obe populacije imaju isti broj stratuma, ali se proporcije tih stratuma, pa samim tim i verovatnoće odgovarajućih ishoda, bitno razlikuju. Prilikom rešavanja ovog zadatka treba krenuti od pitanja u kojoj od dve populacije je grupisanje oko iste vrednosti snažnije i očiglednije, odnosno u kojoj od ovih

populacija je verovatnoća nekog od ishoda vidljivo veća od ostalih. Boja kuglica zapravo manje varira u populaciji 4, jer ipak može da se kaže da neka kategorija dominira i da je veliki broj članova te populacije međusobno sličan, npr. studira isti fakultet. Varijabilnost bitno utiče i na pouzdanost naše procene, jer ćete sigurno mnogo lakše i tačnije rangirati boje po učestalosti, tj. verovatnoći, u slučaju uzorka uzetog iz Populacije 4 nego onog iz Populacije 3. Naravno, u statistici nije dovoljno samo vizuelno proceniti varijabilnost neke pojave, već ju je potrebno izraziti i kvantitativno. O tome će biti više reči u narednim odeljcima, ali primeri pomenutih kvalitativnih varijabli mogu da posluže kao pogodan način da se čitalac uvede u numeričko izražavanje varijabilnosti. Recimo da želimo da numerički izrazimo varijabilnost svojstva *pol*. U grupi žena, proporcija vrednosti *ž* varijable *pol* iznosi 1, a proporcija vrednosti *m* 0. Jednostavnim množenjem ovih proporcija dobijamo vrednost 0 koja govori da *pol*, kao varijabla, ne varira u grupi žena. Analogno tome, najveća varijabilnost bi postojala u situaciji da u nekoj grupi postoji jednak broj žena i muškaraca. Tada bi umnožak proporcija iznosio $0{,}5 \cdot 0{,}5 = 0{,}25$, što je i najveća moguća varijabilnost varijable *pol*. Svaki drugi odnos muških i ženskih osoba dao bi umnožak manji od ove vrednosti i ukazivao bi da su osobe međusobno više slične po polu, jer je jedna od vrednosti učestalija. Umnožak proporcija može da se iskoristi kao pokazatelj varijabilnosti i u slučajevima kada postoji više klasa, kao u našim primerima sa studijskim grupama ili mobilnim operaterima.

> Sakrijte populaciju, kliknite taster *Generiši populaciju* da biste formirali nasumičnu populaciju i potom na osnovu uzoraka različite veličine pokušajte da procenite kako ona izgleda.
>
> Pomeranjem težišnih tačaka na dijagramu promenite proporcije stratuma i kreirajte sopstvenu populaciju. Da li suma proporcija može da bude veća ili manja od 1? Da li povećavanje proporcije jednog ishoda utiče na proporcije (svih) drugih ishoda?

2.5. Osnovne tehnike sažimanja podataka

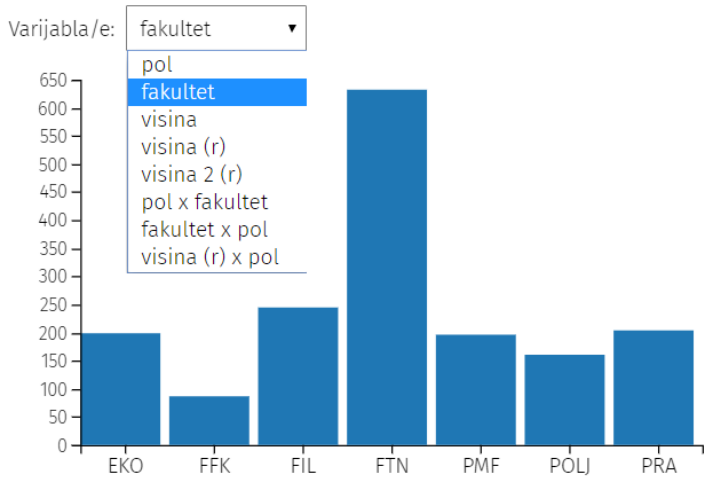
Nakon formiranja reprezentativnog uzorka i merenja odabranih svojstava članova tog uzorka, prikupljene podatke treba na neki način sažeti kako bi se bolje opisali, lakše razumeli i efikasnije saopštili drugima. Izveštaji o

gledanosti emisija ne sadrže sve sirove podatke o tome kog je pola ili uzrasta svaki gledalac i koju emisiju je kada gledao, već se ti podaci prikazuju u sažetoj formi, npr. grupisanjem po emisijama, vremenskim periodima, kategorijama stanovništva ili polu. Kao što smo videli, grafikoni su veoma pogodne alatke za sažimanje i opisivanje podataka, te bi trebalo da budu prvi korak u svakoj statističkoj obradi. Na osnovu grafikona lako se mogu uočiti pravilnosti i nepravilnosti u podacima, tačke oko kojih se oni grupišu, razlike i sličnosti među podgrupama ispitanika, varijabilnost podataka i atipični rezultati. Međutim, sve ove korisne informacije nisu potpuno precizne, jer se u suštini baziraju na procenama vizuelnih karakteristika grafikona od strane istraživača. Stoga je uobičajeno da se različita svojstva podataka, odnosno varijabli, opisuju i na numerički način, tj. kvantitativno. U nastavku odeljka ukratko ćemo opisati osnove tehnike numeričkog sažimanja i opisivanja varijabli.

2.5.1. Tabele frekvencija i tabele kontingencije

Statistička obrada podataka obično započinje njihovim razvrstavanjem u kategorije i određivanjem broja elemenata u svakoj od njih. Na taj način mogu se doneti važni zaključci o karakteristikama uzorka, npr. o tome da li je uzorak dovoljno reprezentativan za svaki od stratuma populacije ili koliko različitih kategorija entiteta je moguće formirati. Grafički rezultat ovog postupka je ranije opisani stubičasti dijagram, a njegov numerički pandan je *tabela frekvencija*. Slično grafikonu, na osnovu tabela frekvencija može da se analizira distribucija učestalosti različitih vrednosti varijable, odnosno njihova raspodela po klasama ili kategorijama entiteta. Ponovo ćemo upotrebiti raniji **primer sa studentima** i varijablama *pol* i *fakultet*, ali ćemo ovoga puta u analizu uključiti varijablu *visina* umesto varijable *broj bodova* (Slika 18). Na početku su prikazani grafikon i tabela frekvencija za varijablu pol. Na osnovu grafikona uočavamo da u našem uzorku ima više studenata (označenih slovom *m*), ali tačna razlika može da se odredi tek na osnovu informacija iz tabele. Pored klasičnih frekvencija koje se označavaju latiničnim slovom *f*, u tabelama se obično prikazuju i *relativne frekvencije* koje nisu ništa drugo do ranije pominjane proporcije, odnosno udeo svake od kategorija u ukupnom broju entiteta. U tabeli smo ovu vrstu frekvencija označili slovom *p* (engl. *proportion*) kako bismo naglasili njihovu vezu sa verovatnoćom ishoda (engl. *probability*). Na osnovu tabele može da se kaže da u uzorku ima nešto više muškaraca, ali i da je verovatnoća da će neka nasumično odabrana osoba koja studira na tom univerzitetu biti muškog pola,

veća od verovatnoće da će biti ženskog. Preciznije, te verovatnoće iznose p_m = 882 : 1723 ≈ 0,51 i p_z = 841 : 1723 ≈ 0,49. Ova dva ishoda nazivaju se *komplementarnim* jer je suma njihovih verovatnoća 1 ili 100%. U formuli smo, zbog greške nastale zaokruživanjem, upotrebili simbol ≈ (približno) a ne = (jednako). Ova razlika nije toliko bitna jer prilikom statističkog zaključivanja nije važno da li je tražena verovatnoća jednaka nekoj vrednosti, već da li je od nje veća ili manja. O tome će biti više reči u trećem poglavlju.



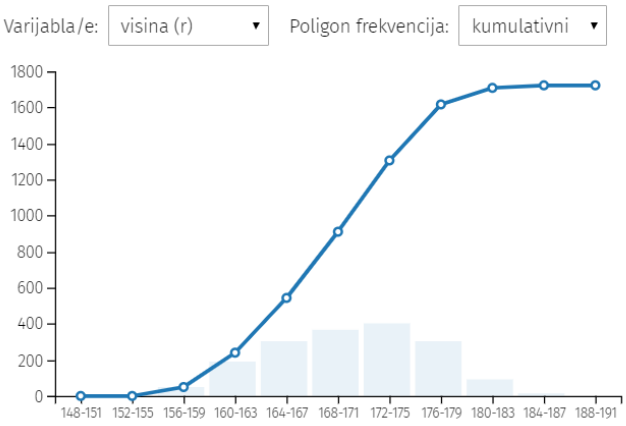
	f	p
EKO	198	0,12
FFK	87	0,05
FIL	244	0,14
FTN	633	0,37
PMF	196	0,11
POLJ	161	0,09
PRA	204	0,12
UKUPNO	1723	1,00

Slika 18. Primer stubičastog dijagrama i tabele frekvencija za varijablu fakultet

> Koristeći primer tabele frekvencija i stubičastog dijagrama za varijablu *fakultet*, odredite koje ste procene i zaključke lakše doneli na osnovu jednog a koje na osnovu drugog načina sažimanja podataka.

U slučaju kvalitativnih, odnosno nominalnih varijabli, kategorije entiteta formiraju se veoma lako. Stoga se ove varijable često nazivaju *kategorijalnim*, jer najčešće služe tome da entitete (ispitanike) razvrstamo u dve ili više grupa. Na sličan način možemo da upotrebimo i ordinalne varijable koje imaju relativno mali broj nivoa, kao što je npr. nivo stručne spreme ispitanika. Međutim, prikazivanje svih mogućih kategorija, odnosno vrednosti varijabli intervalnog ili razmernog nivoa na grafikonu, često je nepraktično. Prikažite grafikon za varijablu *visina*. Kao što vidite, tabela frekvencija je, zbog velikog broja vrednosti na x-osi, prevelika i neprikladna za sažimanje rezultata. Stoga je uobičajeno da se u slučaju kvantitativnih varijabli, čije skale imaju veliki broj

podeoka, formiraju tzv. *razredi*, odnosno intervali vrednosti. Na taj način se poboljšava preglednost grafikona i olakšava interpretacija rezultata. Primer takvog grafikona i tabele videćete kada odaberete opciju *visina (r)*. Ovoga puta (relativne) frekvencije se ne računaju za pojedinačne rezultate, već za razrede rezultata. Intervali, naravno, moraju da budu iscrpni i međusobno isključivi kako bi svaki rezultat mogao da se svrsta u samo jedan razred. Obratite pažnju na to da su u ovom primeru u tabeli prikazane dve dodatne kolone: f_c i p_c. Prva sadrži *kumulativne frekvencije* (engl. *cumulative*) koje pokazuju koliko podataka (entiteta) se akumuliralo ili nakupilo do određene tačke. Na primer, iz tabele možemo da vidimo da u našem uzorku ima 306 studenata koji su visoki između 164 i 167 cm. Zajedno sa svim prethodnim kategorijama to čini kumulativnu frekvenciju od 546 studenata i studentkinja koji su visoki između 152 i 167 cm. U poslednjem redu tabele, tačnije u poslednjoj kategoriji, kumulativne frekvencije dostižu vrednost ukupnog broja entiteta. Samim tim, računanje sume ove kolone nema smisla. U oblasti statističkog zaključivanja naročito su korisne *kumulativne relativne frekvencije* ili *kumulativne proporcije*, označene simbolom p_c. One se računaju na isti način kao i relativne frekvencije, tako da nam omogućavaju donošenje zaključaka u terminima proporcija, odnosno verovatnoća. Na primer, lako možemo da uočimo da se do srednjeg reda, odnosno intervala 168–171 nakupila približno polovina rezultata. Drugim rečima, iznad i ispod neke od vrednosti iz intervala 168–171 nalazi se oko 50% ispitanika. Ta simetričnost je lako uočljiva i na grafikonu, kako na stubičastom dijagramu, tako i na poligonu frekvencija. Ukoliko poligonom frekvencija prikažemo vrednosti f_c umesto vrednosti f, dobija se *poligon kumulativnih frekvencija* (Slika 19). Kriva koja nastaje na ovaj način naziva se *ogiva*, a visina svake tačke na njoj jednaka je sumi visine trenutnog stubića i visina svih stubića koji se nalaze ispod, tj. sa leve strane te tačke. U našem primeru ogiva je takođe simetrična, jer je njen tok od prve do šeste tačke odraz u (dvostrukom) ogledalu njenog toka od šeste do jedanaeste tačke. Njen porast je najpre blag, zato što u nižim kategorijama ima manje ispitanika, potom je mnogo oštriji, jer se oko sredine nalazi najviše rezultata, a na kraju je ponovo blag, zato što u višim kategorijama ima približno isto ispitanika koliko ih je u nižim. Odaberite opciju *visina 2 (r)* da biste videli primer grafikona koji nije simetričan, jer prikazuje uzorak u kome ima znatno više studenata koji su vrlo visoki. U ovom slučaju ogiva ima drugačiji tok rasta koji je do sedme tačke blag, a nakon nje znatno oštriji sve do kraja. Ovakav oblik krive odražavaju i vrednosti u koloni p_c. U prvih pet kategorija visine akumlirano je tek 20% rezultata, da bi u naredne četiri bilo akumulirano i preostalih 80%.



	f	p	f_c	p_c
152-155	3	0,00	3	0,00
156-159	48	0,03	51	0,03
160-163	189	0,11	240	0,14
164-167	306	0,18	546	0,32
168-171	364	0,21	910	0,53
172-175	399	0,23	1309	0,76
176-179	307	0,18	1616	0,94
180-183	94	0,05	1710	0,99
184-187	13	0,01	1723	1,00
UKUPNO	1723	1,00		

Slika 19. Kriva kumulativnih frekvencija (ogiva) varijable *visina*

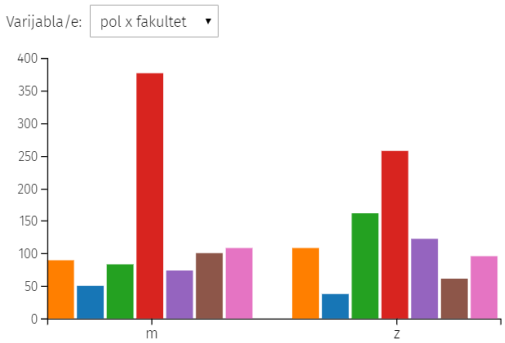
Zbog čega je broj tačaka na poligonima frekvencija za dva veći od broja redova u tabeli frekvencija?

Zbog čega ogiva ima krivolinijski oblik? U kom slučaju bi imala oblik prave?

Koju vrednost na y-osi ima najviša tačka ogive?

Do sada smo tabelama frekvencija prikazivali jednodimenzionalne ili *univarijantne* distribucije frekvencija, tj. distribucije nastale kategorizacijom entiteta na osnovu jedne dimenzije. Na sličan način mogu se prikazati i kategorije nastale kombinacijom većeg broja varijabli, odnosno *multivarijantne* distribucije. Na primer, ako odaberete opciju *pol x fakultet*, biće prikazana tabela koja ima 14 ćelija nastalih kombinovanjem dva nivoa varijable *pol* i sedam nivoa varijable *fakultet*. Pored osnovnih, tabela ima i devet *marginalnih ćelija* u kojima se nalaze sume odgovarajućih redova i kolona, odnosno *marginalne frekvencije*. U poslednjoj ćeliji tabele naveden je ukupan broj entiteta, odnosno vrednost N. Kao što se vidi iz zaglavlja poslednje kolone, u statistici sume označavamo velikim grčkim slovom *Σ (sigma)*. S obzirom na to da se formiraju ukrštanjem dve ili više varijabli, ove tabele nazivaju se i *krostabulacijama* (engl. *crosstabs*) ili, mnogo češće, *tabelama kontingencije* (lat. *contigere* – dogoditi se, pojaviti se). Tabele kontingencije nam omogućavaju da analiziramo distribuciju učestalosti u višedimenzionalnom prostoru i indirektno utvrdimo postojanje različitih veza među varijablama. U našem primeru sa studentima reč je o

dvodimenzionalnom ili *bivarijantnom* prostoru varijabli (Slika 20). Iz aspekta statističke obrade, potpuno je nebitno koja varijabla će formirati kolone a koja redove. Ako ste odabrali opciju *pol x fakultet*, prvi red tabele odgovara levom stubičastom dijagramu a drugi desnom. Iste podatke možemo da prikažemo i drugačije. Kada odaberete opciju *fakultet x pol*, broj redova u tabeli jednak je broju parova stubića na grafikonu. U odnosu na prethodni primer, tabela je samo zarotirana za 90 stepeni, što nije toliko očigledno na osnovu poređenja dva prateća stubičasta dijagrama.



	EKO	FFK	FIL	FTN	PMF	POLJ	PRA	Σ
m	90	50	83	376	74	100	109	882
z	108	37	161	257	122	61	95	841
UKUPNO	198	87	244	633	196	161	204	1723

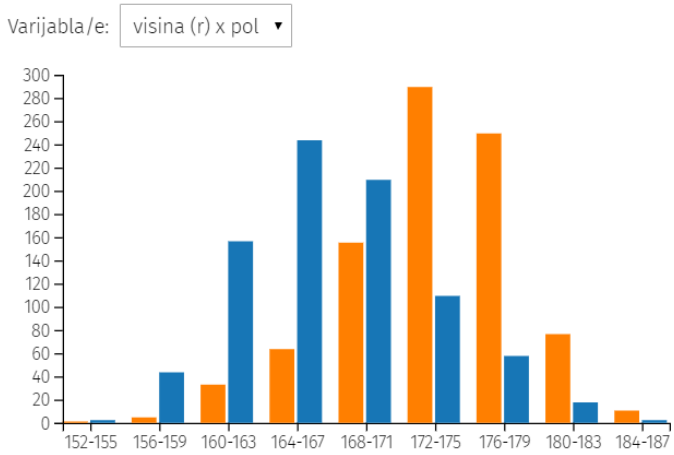
Slika 20. Stubičasti dijagram i tabela kontingencije varijabli *pol* i *fakultet*

Koja boja označava koji pol na grafikonu u primeru *fakultet x pol*?

Posmatrajući rezultate krostabulacije varijabli *fakultet* i *pol*, proverite da li raspodela studenata po polu u ukupnom uzorku odgovara raspodelama u podgrupama formiranim na osnovu varijable *fakultet*.

U jednodimenzionalnim tabelama frekvencija postoji samo jedna suma (kolone). Stoga proporcije i verovatnoće ishoda mogu da se računaju na samo jedan način, kao odnos frekvencije u odgovarajućoj ćeliji i ukupne veličine uzorka. U dvodimenzionalnim tabelama kontingencije to nije slučaj. Pođimo od primera tabele kontingencije za varijable *visina (r) x pol*, u kojoj su žutom pozadinom označene ćelije koje ćemo analizirati (Slika 21). Praktično sve kombinacije ovih ćelija mogu da nam daju informaciju o verovatnoći nekog ishoda. Kolika je, na primer, verovatnoća da je neko od članova populacije studenata muška osoba visoka između 172 i 175 cm? Odgovor je 289 : 1723 ili približno 0,17. Međutim, ukoliko broj 289 podelimo sa sumom kolone *m*, umesto sa ukupnom sumom, dobijamo potpuno drugu vrednost. Ovoga puta

odnos frekvencija govori o verovatnoći da je neki od studenata (muškog pola) visok između 172 i 175 cm. U pitanju je uslovna verovatnoća koja nije nezavisna od vrednosti u drugim ćelijama, već je određena verovatnoćom da je osoba muškog pola. Dakle, *uslovna* verovatnoća da se neko nalazi u razredu 172–175, ukoliko *znamo* da je u pitanju muška osoba, iznosi 289 : 882 ili približno 33%. Na sličan način možemo zaključiti da je verovatnoća da će osoba koja je visoka između 172 i 175 cm biti muško, mnogo veća od verovatnoće da će ona biti ženskog pola. Prva iznosi oko 0,72 (289 : 399), a druga oko 0,28 (110 : 399). Ovoga puta reč je o uslovnoj verovatnoći da je neko muško, ako znamo da je visok između 172 i 175 cm. Navedene zaključke možemo da formulišemo i tako da se odnose na tačnost predviđanja događaja. Na primer, verovatnoća da ćemo tačno pogoditi kog pola je osoba koja je visoka između 172 i 175 cm, veća je od verovatnoće da ćemo tačno pogoditi u kojoj kategoriji visine se nalazi osoba koja je muškog pola.



	m	z	Σ
152-155	1	2	3
156-159	5	43	48
160-163	33	156	189
164-167	63	243	306
168-171	155	209	364
172-175	289	110	399
176-179	249	58	307
180-183	76	18	94
184-187	11	2	13
UKUPNO	882	841	1723

Slika 21. Stubičasti dijagram i tabela kontingencije varijabli *pol* i *visina*

Odnosi frekvencija u osnovnim i marginalnim ćelijama pružaju nam informacije o verovatnoćama koje smo opazili „u praksi", tj. u istraživanju ili na osnovu ličnog iskustva. Stoga se ove verovatnoće nazivaju *empirijskim*. Sa druge strane, odnos marginalnih frekvencija i veličine uzorka pruža nam informaciju o tome šta bi trebalo ili šta bismo mogli da očekujemo „u teoriji". Vratimo se na empirijsku verovatnoću od 17% da je neko u našoj populaciji muškarac visok između 172 i 175 cm. Da li je to u skladu sa očekivanjima? U odgovoru na ovo pitanje pomoći će nam sume odgovarajućeg reda i kolone. Ako pretpostavimo da pol i visina nisu ni u kakvoj vezi, tj. da muškarci nisu nešto

viši od žena kao u našem primeru, onda sume redova i sume kolona u tabeli možemo da posmatramo kao potpuno nezavisne. Nezavisni ishodi su oni kod kojih verovatnoća jednog ishoda ne utiče na verovatnoću drugog, npr. pol ne utiče na visinu ili obratno. U tom slučaju, verovatnoća zajedničkog javljanja tih ishoda računa se kao proizvod verovatnoća svakog od njih. Na primer, verovatnoća da od 39 kuglica u igri *Loto* izvučete broj 7 (ili bilo koji drugi broj) iznosi 1 : 39 ili 0,03. Ta verovatnoća ni na koji način ne utiče na verovatnoću narednog ishoda, osim što je u bubnju ostalo manje kuglica, tako da je verovatnoća da bude izvučen broj 23 (ili bilo koji drugi broj osim 7) sada 1 : 38, što je još uvek oko 0,03. Na osnovu ovih podataka, možemo da izračunamo verovatnoću da na početku izvlačenja budu izvučeni brojevi 7 i 23. Ona iznosi 0,03 · 0,03 = 0,0009 ili 0,09%. Istu logiku možemo da primenimo i u našem primeru sa studentima. Verovatnoća da je neko u populaciji muško, bez obzira na visinu, iznosi 882 : 1723 ili oko 0,51. Verovatnoća da je neka osoba u populaciji visoka između 172 i 175 cm, bez obzira na pol, je 399 : 1723 ili oko 0,23. To znači da bismo u našoj populaciji mogli da očekujemo *teorijsku* verovatnoću od 0,23 · 0,51 ili oko 12% da je neko muškarac visine između 172 i 175 cm. S obzirom na to da smo ustanovili da je *empirijska* verovatnoća veća i iznosi 17%, zaključujemo da visina i pol nisu potpuno nezavisne varijable. Opazili smo više muškaraca a manje žena visokih između 172 i 175 cm, nego što bi se očekivalo da razlike u visini među polovima nema. Stoga imamo pravo da zaključimo da postoji razlika u visini između studenata i studentkinja. To nam, uostalom, pokazuje i stubičasti dijagram. Ovo je bila kratka ilustracija logike zaključivanja na osnovu verovatnoća različitih ishoda kojom ćemo se detaljnije baviti u trećem poglavlju.

> Koja karakteristika stubičastog dijagrama u primeru *visina (r) x pol* ukazuje na to da postoji razlika u visini među polovima?
>
> Kolika je teorijska verovatnoća da je neka osoba u populaciji iz primera, nezavisno od toga kog je pola, viša od 175 cm?
>
> Kolika je teorijska verovatnoća da je neka studentkinja visoka između 184 i 187 cm?

2.5.2. Mere grupisanja ili centralne tendencije

Skupovi podataka veoma često se sažimaju do najvišeg stepena, tako da se predstave jednim brojem koji nazivamo merom *grupisanja* ili *centralne tendencije* rezultata. Taj broj je ponekad nedovoljno informativan, npr. ako zaključke o nekom učeniku donosimo samo na osnovu prosečnog uspeha i bez uvida u njegove pojedinačne ocene. Nekada su mere centralne tendencije neprikladne ili pristrasne, npr. kada materijalni status stanovnika neke države izrazimo prosekom njihovih zarada. I pored toga, sažimanje ove vrste u nauci i statistici je neminovno, najpre zato što su sirovi podaci u tabelarnoj formi nedovoljno pregledni i teški za interpretaciju, ali i zato što većina naprednijih analiza polazi od prethodno izračunatih mera centralne tendencije koje imaju ulogu „predstavnika" skupova podataka. Tako ćemo npr. odeljenja u školi porediti po prosečnom uspehu đaka a bogatstvo država po bruto domaćem proizvodu (engl. *GDP*) po stanovniku (lat. *per capita*). Stoga je veoma važno da istraživač bude svestan prednosti i nedostataka različitih mera centralne tendencije, kao i faktora koji utiču na njihovu vrednost, a time i na njihovu prikladnost u različitim situacijama. Ti faktori se na prvom mestu odnose na nivo merenja varijable, oblik njene distribucije i njenu varijabilnost.

2.5.2.1. Aritmetička sredina, medijana i mod

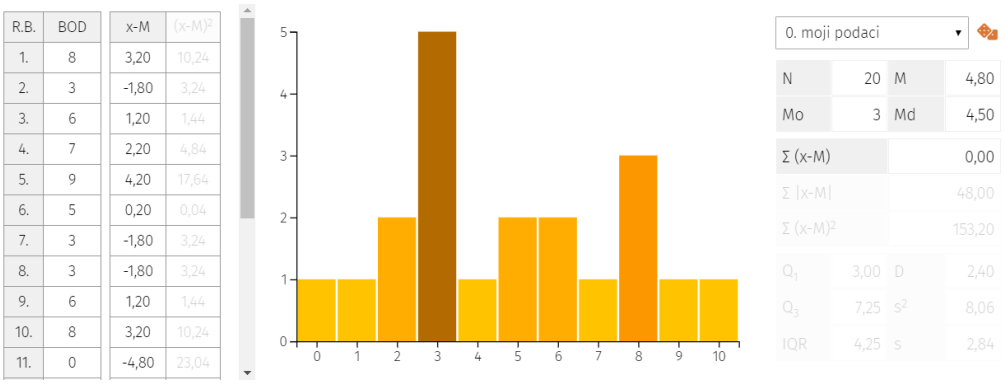
Aritmetička sredina, medijana i mod su najčešće korišćeni deskriptivni pokazatelji kojima se kvantitativno iskazuje mesto oko koga se grupišu rezultati merenja. Aritmetička sredina je *prosečna* vrednost svih rezultata. Označava se velikim slovom M (engl. *mean*) ili $\bar{X}$ (engl. *x-bar*), a računa se po formuli:

$$M = \frac{\sum x}{N}$$

gde x označava pojedinačne rezultate, odnosno izmerene vrednosti varijable čiji se prosek računa, Σ označava sumu, a N broj merenja, odnosno veličinu uzorka. U sledećem primeru **prikazani su rezultati 20 đaka na testu znanja** na kome su mogli da osvoje od 0 do 10 bodova. Stubičasti dijagram prikazuje podatke koji su uneti u tabelu sa leve strane, dok su različiti deskriptivni pokazatelji prikazani sa desne (Slika 22). Određene ćelije desne tabele su prozirne, jer će o njihovom sadržaju biti više reči tek u narednom odeljku. Obratite pažnju na činjenicu da N u formuli za računanje proseka ne označava

broj redova u matrici već broj postojećih vrednosti. Ukoliko obrišete neku od vrednosti u matrici, vrednost N će se smanjiti, ali će aritmetička sredina (M) ostati ista, jer se i suma svih vrednosti smanjila. U tom slučaju postoji jedan nedostajući podatak, što je moglo da se desi ako neki učenik nije bio prisutan na času na kome se radio test. Sada u prazno polje unesite vrednost 0 i proverite da li se promenila M. Naravno, vrednost 0 označava da jedan od učenika nije uspeo da osvoji nijedan bod na testu a ne da je bio odsutan. Stoga je N ponovo 20 a vrednost M postaje manja od 5.

> Koju vrednost treba uneti umesto neke od petica u tabeli da bi vrednost M ponovo postala 5, odnosno da bi se ponovo uspostavila simetričnost distribucije prikazane na grafikonu?



Slika 22. Tabelarni i grafički prikaz podataka uz tipične mere centralne tendencije

Izmenite nekoliko petica u tabeli u proizvoljne vrednosti iz raspona od 0 do 10 i pratite kako se menja M. Ćelije u koloni možete da birate i korišćenjem tastera - (gore) i + (dole) na numeričkom delu tastature. Primetićete da je vrednost izraza Σ(x-M), odnosno suma odstupanja pojedinačnih rezultata od aritmetičke sredine, uvek nula. Drugim rečima, aritmetička sredina je *težište* svih rezultata ili tačka na x-osi koja može da se zamisli kao oslonac klackalice u ravnotežnom položaju. Ukoliko u tabeli povećavate broj vrednosti manjih od 5, težište se pomera ulevo, a ukoliko ima više vrednosti većih od 5, težište se pomera udesno. U tabeli sa leve strane, ova pravilnost se ogleda u činjenici da će „masa" odstupanja od aritmetičke sredine biti jednaka sa obe njene strane, odnosno da će suma negativnih brojeva u koloni x-M uvek biti jednaka sumi pozitivnih, naravno ukoliko se zanemari njihov predznak.

> Formirajte više puta slučajne distribucije klikom na ikonicu kockica u gornjem desnom uglu i pokušajte na osnovu izgleda grafikona, traženjem njegovog centra ravnoteže, da procenite kolika je vrednost M dobijenih distribucija.

U prethodnom primeru računali smo aritmetičku sredinu varijable koja je očigledno diskretna. To na prvi pogled može da se učini neopravdanim, jer kao rezultat dobijamo vrednost koja se ne nalazi na skali instrumenta kojim smo izvršili merenje, npr. testa znanja. U većini slučajeva, ovo nije sporno. Na primer, često se dešava da nastavnici učenicima daju pola, trećinu ili četvrtinu boda kako bi preciznije izrazili njihovo znanje i pokušali da ga tretiraju kao kontinuiranu varijablu. Međutim, postoje i situacije u kojima takav tretman diskretnih varijabli nema opravdanja. Kada odaberete primer 2 sa liste, na grafikonu će biti prikazana distribucija varijable *broj dece u porodici*. Dve porodice imaju jedno dete, osam porodica dvoje itd. Jasno je da nije umesno reći da porodice u proseku imaju 3,11 dece, a još manje da je toliki broj dece karakteristika „prosečne" porodice. Pored toga, odmah se uočava da postoji vrednost koja je atipična, aberantna i vidljivo udaljena od preostalih stubića na grafikonu. To je porodica koja ima desetoro dece. Statistički gledano, ovaj podatak je problematičan zato što bitno menja aritmetičku sredinu distribucije i utiče na dalje zaključke o pojavi koja je izmerena. Takve vrednosti nazivamo *autlajerima* (engl. *outlier*) i trebalo bi ih na odgovarajući način tretirati pre dalje analize. Za početak, potrebno je utvrditi zbog čega su se javile, jer često mogu da budu posledica greške u merenju ili omaške prilikom unosa podataka u tabelu. Nakon toga ove vrednosti mogu da se isključe iz analize ukoliko to veličina uzorka dozvoljava. U našem primeru, nakon brisanja vrednosti 10 iz matrice, M se znatno smanjuje, ali ipak ne toliko bitno u smislu konačnog zaključka. Naime, vrednost M i dalje sugeriše da porodice u proseku imaju približno troje dece, što nije dovoljno pouzdan podatak imajući u vidu izgled distribucije podataka. Stoga se u ovakvim slučajevima, kao prikladnija mera centralne tendencije, češće koristi *mod* ili *dominantna vrednost* koja je u našem primeru označena sa Mo. Mod je najčešća vrednost u skupu podataka ili vrednost koju dobijamo kao odgovor na pitanje šta je *tipično* za grupu merenja. Prikladnije i ispravnije je reći da *tipična* porodica ima dvoje dece, nego da „prosečna" porodica ima otprilike troje. Mod je, dakle, pravičnija mera centralne tendencije u odnosu na M kada su distribucije atipične, kada postoje aberantni rezultati, a posebno onda kada su varijable merene na nižim nivoima merenja.

Štaviše, mod je jedina mera grupisanja koja može i sme da se primeni za kvalitativne (nominalne) varijable. Na primer, ukoliko zamislimo da vrednosti na našem grafikonu ne označavaju broj dece već fakultet koji neko studira, aritmetička sredina postaje potpuno besmislen pokazatelj i jedini opravdani zaključak je da dominantna vrednost varijable *fakultet* iznosi 2, odnosno da u uzorku ima najviše studenata fakulteta koji je označen tim kodom.

Pored aritmetičke sredine i moda, u statistici se, kao mera grupisanja rezultata, često koristi i *medijan, medijana* ili *srednja vrednost*. Kao što joj ime kaže, to je vrednost koja se nalazi na srednjoj poziciji u nizu svih rezultata poređanih od najmanjeg do najvećeg. Ova pozicija lako se pronalazi ako je broj rezultata neparan, a ukoliko je paran, medijana se računa kao prosek dva središnja rezultata u nizu. Medijana može da se definiše i uz pomoć ranije pomenutih kumulativnih frekvencija. Srednja vrednost je prvi razred ili podeok na x-osi kod koga kumulativna frekvencija postaje veća od N : 2 ili prosek vrednosti podeoka čija je kumulativna frekvencija jednaka N : 2 i prvog narednog podeoka čija je frekvencija veća od nule. Na prikazanom grafikonu možete da vidite kumulativne frekvencije kada pokazivačem miša prelazite preko stubića. U primeru 2 kumulativna frekvencija postaje veća od 9 (N : 2 = 18 : 2) iznad broja 2, što je ujedno srednja vrednost distribucije koja je u tabeli sa desne strane označena simbolom Md. Uklonite ponovo vrednost 10 iz matrice i uočite da se vrednosti Mo i Md ne menjaju, za razliku od vrednosti M. Obratite pažnju na to da srednja i prosečna vrednost distribucije nisu iste, te stoga ni ove termine ne treba koristiti kao sinonime. Prosečna, srednja i tipična vrednost predstavljaju suštinski različite mere centralne tendencije koje mogu, ali ne moraju da imaju istu vrednost.

> Odaberite primer 3 i analizirajte vrednosti mera centralne tendencije.
>
> Da li su M i Md u ovom primeru prikladne mere centralne tendencije?
>
> Zbog čega M i Md u ovom primeru imaju istu vrednost? Kakav oblik imaju sve distribucije kod kojih M i Md imaju istu vrednost?
>
> Zbog čega umesto broja za vrednost Mo stoji reč „više“? Zašto ovu distribuciju nazivamo bimodalnom? Kako bi mogla da izgleda neka *polimodalna* distribucija?

Distribucije 1 i 3 imaju istu M. Za koju od tih distribucija je vrednost M bolja mera centralne tendencije i zašto?

Formirajte više puta slučajne distribucije klikom na ikonicu kockica i analizirajte odnose između vrednosti različitih mera centralne tendencije.

2.5.2.2. Još neke vrste prosečnih vrednosti

U prethodnim primerima videli smo da prosek ili aritmetička sredina, iako najpopularnija, nije uvek najprikladnija mera grupisanja. Njena „pravičnost" je ujedno i njena glavna mana, jer, za razliku od medijane i moda, uzima u obzir *vrednost* svakog rezultata merenja pa tako i vrednosti eventualnih autlajera. Stoga se medijana i mod smatraju otpornijim ili *robusnijim* merama grupisanja. Čak i ako u nekom skupu podataka postoji aberantan rezultat koji je 10, 100 ili 1.000 puta veći od ostalih rezultata, on neće uticati na vrednosti medijane i moda. Međutim, robusnost u statistici obično podrazumeva i manju preciznost, te je stoga većina otpornijih tehnika ujedno i manje precizna, odnosno manje „moćna" da ukaže na postojanje određenih fenomena. Da bi se aritmetičke sredine učinile robusnijim, u statistici se koriste njene varijacije poznate kao *podrezane sredine* (engl. *trimmed means*). Prilikom računanja podrezanih sredina ne uzimaju se u obzir svi rezultati, već samo određeni procenat onih središnjih. Obično se isključuje 5% rezultata sa obe strane distribucije, uz pretpostavku da će na taj način biti isključeni i potencijalni aberantni rezultati, te će se dobiti vrednost koja tačnije procenjuje mesto oko koga se grupišu rezultati. Jedna od mera centralne tendencije ovog tipa je tzv. *triprosek* (engl. *trimean*) koju je predložio američki matematičar Džon Vajlder Tuki (Tukey, 1977). Da bi se izračunao triprosek, potrebno je pronaći vrednosti *kvartila*, odnosno tačaka koje dele rezultate u četiri grupe jednake po broju. Prvi kvartil je tačka ispod koje se nalazi približno 25% rezultata, drugi je medijana, odnosno tačka ispod i iznad koje se nalazi 50% rezultata, a treći je vrednost iznad koje se nalazi preostalih 25% rezultata. Triprosek je aritmetička sredina zbira dve vrednosti medijane, i vrednosti prvog i trećeg kvartila.

Aritmetička sredina pripada grupi tzv. *Pitagorejskih sredina* u koju još spadaju geometrijska i harmonijska sredina. U nekim oblastima nauke, kao što je npr. ekonomija, druge dve sredine koriste se češće od aritmetičke. Stoga treba biti obazriv ukoliko se neka vrednost naziva prosekom, jer taj termin može da ima drugačije značenje u različitim oblastima i kontekstima. Na ovom mestu

ćemo ukratko objasniti logiku ovih mera centralne tendencije i način njihove primene, kako bismo ukazali na činjenicu da aritmetička sredina u određenim situacijama može da pruži pogrešne informacije o podacima. *Geometrijska sredina* je prikladna mera grupisanja rezultata nastalih postupkom množenja, odnosno onih čiji se međusobni odnosi tačnije opisuju izrazom *koliko puta,* a ne *za koliko* je neka vrednost veća od neke druge. Na primer, ukoliko je proizvodnja jabuka u prvoj godini porasla 2 puta, sa 100 tona na 200 tona, a u drugoj godini čak 8 puta, sa 200 na 1.600 tona, onda prosečni godišnji porast izražen aritmetičkom sredinom nije tačan. Vrednost (2 + 8) : 2 = 5 bi sugerisala da nakon dve godine proizvodnja treba da bude 2.500 tona, jer se svake godine povećavala prosečno 5 puta. U ovakvim situacijama bi trebalo izračunati geometrijsku sredinu vrednosti kao n-ti koren njihovog proizvoda. U našem primeru to je kvadratni koren vrednosti 8 · 2 koji iznosi 4. Ovoga puta prosečni godišnji porast daje tačan krajnji rezultat: 100 · 4 · 4 = 1.600. Slično tome, zamislite da se uz brdo visoko 4 km penjete brzinom od 4 km na sat, a potom niz njega silazite brzinom od 12 km na sat. Vaša prosečna brzina nije (4 + 12) : 2 = 8 km/h, jer bi to značilo da ste celo brdo prešli za jedan sat, a zapravo ste to vreme utrošili samo za penjanje. U ovom slučaju treba upotrebiti *harmonijsku sredinu* koja je prikladnija za računanje proseka rezultata izraženih kao odnos dveju vrednosti. Harmonijska sredina računa se kao recipročna vrednost aritmetičke sredine recipročnih vrednosti niza rezultata. Recipročna vrednost nekog broja dobija se kada se 1 podeli tim brojem. U našem primeru prosečna brzina kretanja je harmonijska sredina dve vrednosti: 1 : ((1 : 4 + 1 : 12) : 2) = 6 km/h. Ovo je tačna vrednost, jer ste razdaljinu od 8 kilometara uz i niz brdo prešli za 1 sat i 20 minuta.

Na kraju ovog odeljka treba napomenuti i to da nisu retke situacije u kojima se prosek rezultata izračunava na osnovu sažetih a ne na osnovu sirovih vrednosti. Na primer, ukoliko imamo samo podatke o prosečnom uspehu učenika većeg broja škola u nekom regionu, a želimo da izračunamo prosečan uspeh svih đaka u tom regionu, nameće se vrlo jednostavno rešenje da sve proseke saberemo i podelimo brojem škola. Ovako dobijen *prosek proseka* jeste ispravno rešenje, ali može da bude veoma pristrasno ako svaka aritmetička sredina nije izračunata na istom broju rezultata. Stoga je uvek dobro znati i veličine uzoraka na kojima je prosek izračunat, kako bi se onim aritmetičkim sredinama koje su dobijene na većim uzorcima dalo i veće *opterećenje* ili *ponder* (engl. *weight*). Tako dolazimo do pokazatelja koji je poznat kao *zajednička aritmetička sredina*. Ona se računa tako što se svaki prosek najpre pomnoži

veličinom uzorka na kome je izračunat a potom se suma tako *ponderisanih* proseka podeli ukupnim brojem ispitanika. Na taj način se sirovi rezultati merenja, doduše na veoma grub način, „rekonstruišu" uz pretpostavku da se npr. prosek 3,23 izračunat u školi od 500 đaka, odnosi na sve te đake i da u obračun zajedničke aritmetičke sredine možemo da uključimo 500 vrednosti 3,23. Ipak, ovo je znatno korektnije nego da smo vrednosti 3,23 pridali isti značaj kao i npr. vrednosti 4,65 koja je dobijena u školi sa 1.500 đaka. Na sličan način aritmetička sredina može da se izračuna i na osnovu tabela frekvencija koje smo opisali u poglavlju 2.5.1., tako što se svaka vrednost u tabeli pomnoži sa njenom frekvencijom, potom se dobijeni proizvodi saberu i na kraju podele ukupnim brojem rezultata, odnosno sumom svih frekvencija.

2.5.3. Mere raspršenja ili varijabilnosti

Primeri iz prethodnog odeljka pokazuju da dve distribucije potpuno različitog oblika mogu da imaju iste vrednosti aritmetičke sredine. To znači da mere grupisanja ne pružaju dovoljno informacija potrebnih da bi se adekvatno i potpuno opisala neka pojava. Na primer, ukoliko student u proseku provodi sat vremena dnevno na društvenim mrežama, to može da znači da on svakoga dana troši jedan sat na tu aktivnost ili da radnim danima provede 10 do 15 minuta, a vikendom između 3 i 4 sata koristeći društvene mreže. U drugom slučaju, svojstvo definisano kao vreme provedeno na društvenim mrežama očigledno u većoj meri varira. Pojam varijabilnosti uveli smo u poglavlju 2.4. na primeru kvalitativne (nominalne) varijable, a u ovom odeljku ćemo objasniti logiku najčešće korišćenih mera varijabilnosti kojima se opisuju kvantitativne varijable, odnosno distribucije. Na početku treba imati na umu da mere varijabilnosti ne pokazuju samo stepen disperzije ili raspršenja individualnih podataka, već indirektno govore i o tome koliko poverenja možemo da imamo u odabranu meru grupisanja. U našem primeru veća varijabilnost vremena koje student provodi na društvenim mrežama ukazuje na to da se pojedinačne dnevne vrednosti bitno razlikuju među sobom, ali i na to da dobijeni prosek, iako tačna, nije i dovoljno pouzdana mera grupisanja. Na primer, moguće je da student zapravo nijednog dana nije proveo jedan sat koristeći društvene mreže. Dakle, ukoliko želimo kvantitativno da opišemo neku pojavu, pored odabrane mere grupisanja, biće nam potrebna i odgovarajuća mera raspršenja rezultata.

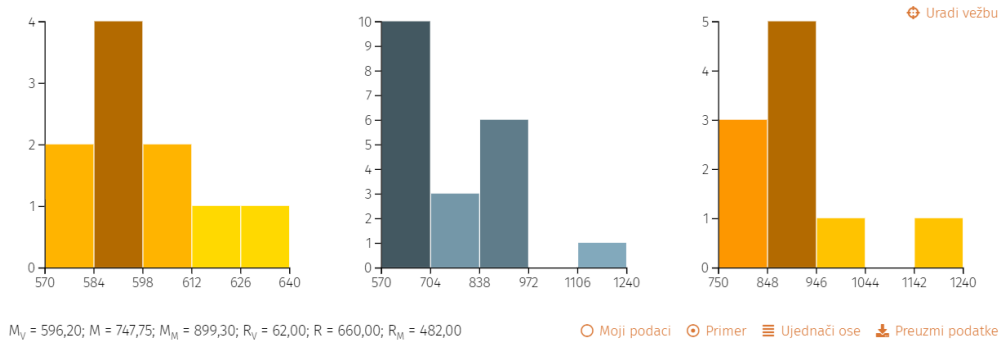
2.5.3.1. Vizuelna procena i poređenje varijabilnosti

Pre nego što pređemo na problem numeričkog iskazivanja varijabilnosti, zadržaćemo se na pitanju vizuelne procene raspršenosti rezultata na osnovu grafikona. U narednom primeru **prikupićemo podatke o brzini reakcije na vizuelni stimulus**. Potrebno je kliknuti sivi kvadrat sa leve strane, nakon čega će se na mestu mete sa desne strane pojaviti narandžasti kvadrat. Njega treba kliknuti što brže jer se vreme proteklo između klika na sivi i klika na narandžasti kvadrat beleži kao brzina reakcije u milisekundama. Nakon dva merenja koja služe za vežbu, postupak se ponavlja 20 puta. Naizmenično će biti prikazano 10 velikih i 10 malih kvadrata. Kliknite sivi kvadrat da biste uradili vežbu.

> Koliko varijabli prepoznajete u ovom eksperimentu? Koji je nivo merenja svake od njih?
>
> Koliko redova i koliko kolona treba da ima matrica sirovih podataka u koju biste zabeležili podatke prikupljene u ovom primeru?
>
> Da li se redovi matrice sirovih podataka u navedenom primeru odnose na različite osobe (ispitanike) ili na nešto drugo?
>
> Na koji način je upotrebljena varijabla *veličina kvadrata* na prikazanim grafikonima?
>
> Da li je brzina reakcije mogla da bude izražena u sekundama ili nekim drugim jedinicama umesto u milisekundama? Šta to govori o varijabli?

Rezultati svih 20 merenja prikazani su na središnjem grafikonu, a njihova aritmetička sredina označena je sa M u donjem levom uglu okvira (Slika 23). Veličina kvadrata upotrebljena je kao dihotomna grupišuća varijabla na osnovu koje su merenja podeljena u dve kategorije. Sa jedne strane nalazi se grafikon koji prikazuje 10 izmerenih brzina u situacijama kada je kvadrat bio velik, a sa druge 10 vrednosti kod kojih je kvadrat bio mali. Na osnovu proseka, ali i na osnovu raspršenosti rezultata, možemo da zaključimo koji grafikon prikazuje koju kategoriju, odnosno grupu merenja. Očekujemo da je brzina bila veća kada je narandžasti kvadrat bio veći zato što je bio bliži sivom kvadratu, a imao je i veću površinu koja je olakšavala pozicioniranje pokazivača miša. Vrednost M_V odnosi se na levi grafikon (veliki kvadrati), a vrednost M_M na desni (mali

kvadrati). Imajte na umu da manja vrednost M ukazuje na veću brzinu u milisekundama. Pored toga što je prosek desne distribucije veći, očekujemo da će biti veća i raspršenost rezultata. Raspršenost najjednostavnije možemo da izrazimo kao razliku između najvećeg i najmanjeg rezultata u nizu. Ovaj pokazatelj naziva se *raspon,* označava se slovom R i veoma lako može da se očita sa x-ose grafikona. Brzina reakcije u grupi malih kvadrata trebalo bi da ima veću vrednost raspona (R_M) ne samo zato što je brzina kliktanja više varirala zbog relativno male mete, već i zato što jedan od malih kvadrata namerno nije bio postavljen na mesto na kome je trebalo da se pojavi. Osim toga, bio je i slabije vidljiv. Na taj način je simulirana pojava autlajera, odnosno aberantno visokog rezultata merenja.



Slika 23. Primer distribucija brzine kliktanja mišem na velike (levo) i male mete (desno)

Ukoliko izračunate prosek vrednosti M_V i M_M, dobićete zajedničku aritmetičku sredinu koja je jednaka vrednosti M. Zbog čega je to tako? Kada prosek tih vrednosti ne bi dao vrednost M?

Da li je opravdano izračunavanje vrednosti R kao proseka ili zbira vrednosti R_V i R_M? Zbog čega jeste, odnosno zbog čega nije?

Kako biste izračunali raspon rezultata svih merenja na osnovu raspona grupa merenja, odnosno levog i desnog grafikona?

U nastavku teksta koristićemo grafikone simuliranih podataka koji se prikazuju kada odaberete opciju *Primer*. Analizirajte aritmetičke sredine i raspone svih merenja i uporedite ih sa merenjima po grupama. Obratite pažnju na to da bi se na osnovu izgleda grafikona, a bez uvida u vrednosti x-ose, moglo

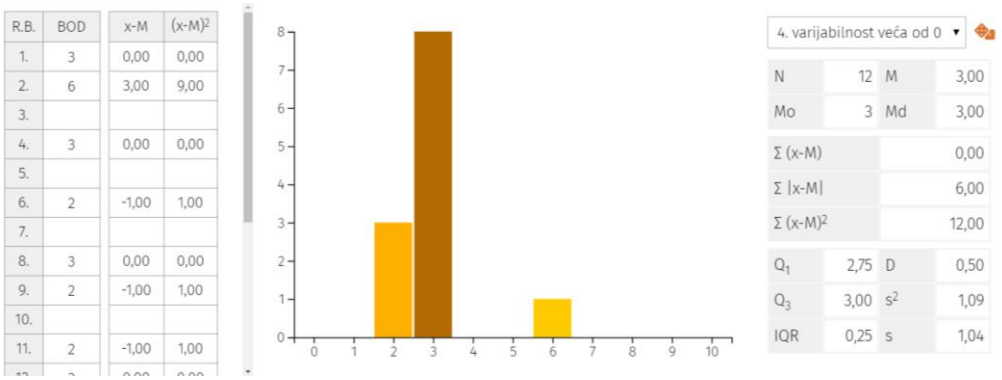
zaključiti da sva tri grafikona imaju veoma sličan oblik, a time i veoma slične mere centralne tendencije i raspršenja. To naravno nije tačno, jer su podeoci x-ose prilagođeni rasponima prikazanih rezultata. Stoga su i intervali razreda na levom grafikonu manji od onih na desnom. Kada kliknete taster *Ujednači ose*, podeoci na x-osama levog i desnog grafikona izjednačiće se sa podeocima x-ose srednjeg grafikona, uzimajući u obzir najmanju i najveću vrednost dobijenu na celokupnom skupu podataka. Sada je vizuelno poređenje varijabilnosti dve grupe merenja opravdano i mnogo lakše, a razlika u raspršenosti rezultata postaje očiglednija. Uočite da je središnji grafikon nastao preklapanjem levog i desnog grafikona, što je rezultiralo bimodalnom distribucijom brzine reakcije. Aberantni rezultat vidljiv je u grupi malih kvadrata i na srednjem grafikonu, ali ne postoji u grupi velikih kvadrata. Prikazujte naizmenično svoje podatke biranjem opcije *Moji podaci* i podatke iz primera biranjem opcije *Primer* da biste analizirali po čemu su oni slični, a po čemu se razlikuju. Ukoliko želite da ponovo uradite vežbu, kliknite taster u gornjem desnom uglu okvira.

Podatke iz primera i podatke koje ste prikupili radeći vežbu, možete da preuzmete klikom na taster *Preuzmi podatke*. Podaci su smešteni u datoteku koja ima csv format (engl. *comma-separated values*). U pitanju je tekstualna datoteka u kojoj svaki red predstavlja jednog ispitanika, jedno merenje ili jedan niz podataka. Kolone, odnosno varijable u okviru reda, razdvojene su zarezima. Ovo je veoma čest format čuvanja podataka u tabelarnom obliku. Ukoliko imate instaliran paket Microsoft Office, datoteka će verovatno biti povezana sa programom Excel. Jasno je da korišćenje csv formata za čuvanje podataka nije prikladno u situacijama kada neke od vrednosti u kolonama sadrže decimalne vrednosti odvojene zarezima. Stoga se podaci obično čuvaju u složenijim formatima kao što su Excel (xls) i Calc (ods) tabele ili matrice napravljene u statističkim paketima (npr. sta, sav, sas). Druga mogućnost je da se umesto zareza upotrebi znak koji ne može da bude deo vrednosti varijable, npr. onaj koji se dobija pritiskom na taster *Tab* na tastaturi.

2.5.3.2. Varijansa i standardna devijacija

Raspon se računa na osnovu samo dve vrednosti iz skupa rezultata, što ga čini jednostavnom, ali veoma grubom merom raspršenja. Njegova vrednost pokazuje udaljenost između najmanjeg i najvećeg rezultata, ali ne govori ništa o tome koliko rezultati variraju unutar tog intervala. Stoga se u statistici češće koriste mere varijabilnosti kojima se obuhvataju svi rezultati. Za vežbu ćemo

ponovo iskoristiti **primer iz odeljka o merama grupisanja**. Na početku je prikazano 20 rezultata koji se međusobno ne razlikuju, tako da sve mere varijabilnosti, uključujući i raspon, imaju vrednost 0. Odaberite primer 4 sa liste da biste na grafikonu prikazali rezultate koji variraju (Slika 24). U ovom skupu podataka, raspon iznosi 6 - 2 = 4 boda, a aritmetička sredina 3. Kao što smo rekli, vrednost M ne nalazi se nužno na sredini raspona, ali se uvek nalazi u težištu distribucije. U sredini raspona je vrednost 4, ali je prosek manji od toga, jer je 3 boda tačka u odnosu na koju je „masa" odstupanja ulevo (3 · -1 bod) jednaka „masi" odstupanja udesno (1 · 3 boda).



Slika 24. Tabelarni i grafički prikaz podataka uz tipične mere varijabilnosti

Pošto vrednost izraza Σ(x-M) uvek iznosi nula, tek suma apsolutnih vrednosti odstupanja, označena izrazom Σ|x-M|, daje nam informaciju o ukupnoj količini odstupanja svih rezultata od aritmetičke sredine, bez obzira na predznak. Ukoliko tu sumu podelimo veličinom uzorka (N), dobijamo pokazatelj koji se zove *prosečno apsolutno odstupanje*, a u tabeli sa desne strane označen je slovom D. Vrednost D za prikazani skup podataka iznosi 0,5, što znači da je varijabilnost rezultata pola boda. Drugim rečima, rezultati 12 đaka na testu znanja odstupaju u proseku za pola boda od 3 boda – jedan đak za tri, tri đaka za 1. Rezultati preostalih 8 đaka uopšte ne odstupaju od proseka. Unosite vrednosti 3 u prazne kućice matrice i posmatrajte kako se menja vrednost D. Dodavanjem vrednosti koje su jednake aritmetičkoj sredini, ukupna suma odstupanja ostaje ista. Iako je raspon varijable sve vreme isti, prosečno apsolutno odstupanje se smanjuje, jer se ista suma odstupanja deli većim N. Drugim rečima, M postaje sve preciznija i pouzdanija mera grupisanja, jer postaje bolji predstavnik sve većeg broja pojedinačnih rezultata. Ukoliko, pak,

vrednosti 3 zamenite nekim drugim vrednostima, na primer 5, varijabilnost će početi da se povećava, a pouzdanost aritmetičke sredine da se smanjuje.

Negativan predznak neke vrednosti moguće je ukloniti i njenim kvadriranjem. Ukoliko se na taj način izračuna količina odstupanja pojedinačnih rezultata od njihovog proseka, dobija se vrednost označena izrazom $\Sigma(x\text{-}M)^2$. Ovu vrednost takođe možemo da podelimo veličinom uzorka (brojem merenja) i da dobijemo prosek kvadriranih odstupanja rezultata od aritmetičke sredine. Tako izračunat pokazatelj varijabilnosti naziva se *varijansa* i označava se simbolom s^2. Odaberite ponovo primer 1 sa liste i zamenite bilo koje dve vrednosti 5 u levoj tabeli vrednostima 4 i 6. Uočavate da su sume apsolutnih odstupanja i kvadriranih odstupanja jednake, ali se vrednosti prosečnog apsolutnog odstupanja i varijanse razlikuju. Konkretno, D je nešto manje od s^2, a razlog je u formuli za izračunavanje varijanse:

$$s^2 = \frac{\sum(x - M)^2}{N - 1}$$

U imeniocu gornje formule nije vrednost N, već N-1, što varijansu čini nešto većom u odnosu na situaciju kada bi ona zaista bila prosek svih kvadriranih odstupanja. O razlogu ove korekcije biće više reči kasnije, a na ovom mestu je dovoljno odgovoriti na pitanje kada ta korekcija u većoj meri utiče na konačni ishod, odnosno na razliku između rezultata koji se dobija deljem sa N i onog koji se dobija deljenjem sa N - 1. Odgovor je da će razlika biti veća kada je vrednost N mala. Zato ovu korekciju možemo da shvatimo i kao neku vrstu kazne za istraživača koji želi da donese zaključak na osnovu veoma malog broja merenja. Ta kazna očigledno ima značajniji efekat kada je veličina uzorka 10, nego kada je 10.000. Ako odemo još dalje, možemo reći da kazne neće ni biti ako je varijabla izmerena u celoj populaciji, kao teorijski neograničenom skupu entiteta. Tada će formula za izračunavanje varijanse biti malo drugačija:

$$\sigma^2 = \frac{\sum(x - \mu)^2}{N}$$

Uočavamo da u gornjoj formuli nema pomenute korekcije u vidu umanjivanja veličine uzorka, jer varijablu nismo ni merili na uzorku već na celoj populaciji. Iz istog razloga upotrebljeni su i drugačiji simboli. Naime, aritmetičku sredinu varijable u populaciji ne označavamo slovom M, već malim grčkim slovom μ (mi). Na osnovu nje može da se izračuna i varijansa varijable u populaciji koju označavamo sa σ^2 (sigma na kvadrat) a ne s^2. U pitanju su dakle isti deskriptivni

pokazatelji, ali se odnose na različite skupove entiteta. Deskriptivne pokazatelje koji se odnose na celu populaciju (μ i σ^2) nazivamo *parametrima*, a pokazatelje koji se odnose na uzorak uzet iz populacije (M i s^2) nazivamo *statisticima*.

Bitan nedostatak varijanse u odnosu na prosečno apsolutno odstupanje predstavljaju jedinice u kojima se ona izražava. Na primer, ako je aritmetička sredina varijable izražena u bodovima, varijansa će biti izražena u bodovima na kvadrat. To je čini manje interpretabilnom i intuitivnom, pa se u statistici češće koristi njen kvadratni koren. Vrednost koja se dobija na taj način zove se *standardna devijacija* i predstavlja najpopularniju meru raspršenja podataka. Standardna devijacija uzorka se, dakle, računa po formuli:

$$s = \sqrt{\frac{\sum(x - M)^2}{N - 1}}$$

Na prvi pogled, može se učiniti da u gornjoj formuli korenovanje potpuno potire prethodnu operaciju kvadriranja, te da se vrednost standardne devijacije svodi na prosečno apsolutno odstupanje. To naravno nije tačno. Vrednost s izračunata na uzorku uvek će biti veća od vrednosti D za istu varijablu, ne samo zbog korekcije u imeniocu formule, već i zbog toga što postupak računanja standardne devijacije, odnosno operacija kvadriranja, dodatno naglašava velika odstupanja rezultata od proseka. Odaberite ponovo primer 4 sa liste i obratite pažnju na to da je s približno duplo veća od D. Ukoliko vrednost 6 u tabeli zamenite vrednošću 4, varijabilnost se naravno smanjuje, pa tako i s i D, ali njihova razlika sada nije toliko izražena. Ako ponovo povećamo varijabilnost rezultata tako što neku vrednost 3 u tabeli zamenimo sa 2, prosečno apsolutno odstupanje ponovo postaje 0,5, ali je sada standardna devijacija neznatno veća od njega. Dakle, za iste vrednost D, vrednosti s mogu da budu različite u zavisnosti od oblika distribucije. Razlika među njima posebno je izražena kada u skupu podataka postoje aberantni rezultati. Primere sa liste možete da poredite sa podacima koje ste sami uneli ili izmenili izborom opcije *0. moji podaci*.

2.5.3.3. Pojam matematičke funkcije

Na ovom mestu ćemo veoma kratko skrenuti pažnju čitaoca na neka od svojstava varijanse koja je čine aritmetički „poželjnijom“ merom od prosečnog apsolutnog odstupanja. Osnovni smisao ovog odeljka je podsećanje čitaoca na

logiku *matematičkih funkcija* kao veoma važnog koncepta u statistici (Slika 25). U **dvodimenzionalnom koordinatnom sistemu** prikazano je dvadesetak tačaka. Njihova pozicija u prostoru određena je uz pomoć veoma jednostavne formule:

$$f(x) = x$$

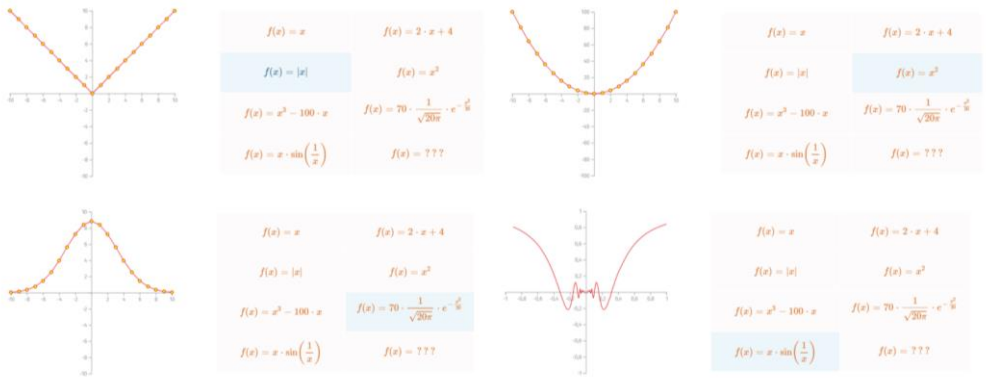
Gornji izraz predstavlja funkciju koja opisuje odnos između vrednosti dva skupa podataka. Konkretnije, za svaku vrednost koja se nalazi na x-osi, vrednost na y-osi jednaka je vrednosti x. Na primer, za vrednost 4 na x-osi, vrednost f (funkcija) od x je takođe 4. Kada kliknete bilo koju tačku na grafikonu i držite pritisnut taster miša, videćete *projekcije* te tačke na obe ose. Projekcije vam omogućavaju da lakše povežete vrednosti x i y, odnosno da lakše očitate *koordinate* svake tačke. Ako skale obe ose tretiramo kao kontinuirane, mogli bismo da iscrtamo teorijski neograničen niz tačaka koje formiraju pravu liniju prikazanu na slici. Stoga ovu vrstu funkcija nazivamo *linearnim*. Kada kliknete formulu:

$$f(x) = 2 \cdot x + 4$$

biće prikazana malo drugačija linearna funkcija koja ovoga puta ne prolazi kroz centar koordinatnog sistema, zato što za nultu vrednost x vrednost y više nije 0 već 4. Treća formula odnosi se na funkciju apsolutnih vrednosti. Kao što vidite, ona nije linearna i ima jedan oštar prelom u tački 0, jer vrednosti f(x) ne mogu da budu negativne. Za razliku od nje, funkcija kvadriranih vrednosti je takođe nelinearna, ali je glatka i postepeno menja svoj tok. Matematičkim rečnikom rečeno, funkcija apsolutnih vrednosti nema *izvod* u tački 0 i da zbog toga nije *diferencijabilna* u svakoj svojoj tački. Upravo ta razlika između treće i četvrte funkcije ilustruje prednost kvadriranih u odnosu na apsolutne vrednosti u statistici, odnosno prednost varijanse kao mere varijabilnosti u odnosu na prosečno apsolutno odstupanje.

Da bismo objasnili praktičnu prednost funkcija koje su diferencijabilne nećemo koristiti matematički jezik, već jedan mnogo banalniji primer. Zamislite da prema vama leti objekat koji treba da izbegnete. To ćete najlakše uraditi ako se on kreće pravolinijski, odnosno ako njegovu putanju možete da opišete i predvidite linearnom funkcijom. Sledeći ishod koji bi bio prihvatljiv je onaj u kome objekat menja pravac, ali to ne čini naglo i oštro, već postepeno. Upravo nam izvod funkcije u svakoj tački omogućava da predvidimo kakav će biti njen

dalji tok. Predviđanja ove vrste često pravimo potpuno intuitivno. Iako se to može shvatiti kao rizično ponašanje, sigurno vam se desilo da prelazite ulicu dok prema vama ide automobil i puštate ga da prođe na nekoliko desetina centimetara od vas. Takvu opuštenost i tačnost procene mogla je da vam pruži samo *izvesnost* pravolinijske putanje automobila. Međutim, mnogo teži zadatak je izbegavanje objekta koji se ne kreće pravolinijski, posebno ako naglo menja pravac kretanja. Slično je i u matematici, odnosno statistici. Izvod je mera osetljivosti promene funkcije u svakoj tački, a matematičke operacije lakše se obavljaju kada te promene nisu nagle. Izvod se grafički prikazuje kao tangenta funkcije u određenoj tački, odnosno prava koja dodiruje liniju funkcije ali je ne seče. U našem primeru sa funkcijama, tangente se prikazuju kao zelene duži kada pokazivačem miša prelazite preko tačaka u koordinatnom sistemu.



Slika 25. Primeri različitih matematičkih funkcija

Pređite pokazivačem miša preko svih tačaka, najpre na funkciji kvadriranih vrednosti a potom i na funkciji apsolutnih vrednosti. Uočite da ova druga nema tangentu u tački 0 jer ju je nemoguće povući. Tačnije, ima ih više, tako da je nemoguće predvideti u kom pravcu će se funkcija kretati nakon te tačke. Peta i šesta formula prikazuju još dve kontinuirane krivolinijske funkcije koje imaju izvod u svakoj tački. Od posebne važnosti je šesta, koja se često naziva *zvonastom krivom*. O njenim svojstvima biće više reči u narednom odeljku. Na kraju, sedma funkcija služi samo kao primer mogućnosti da se vizualizacija kao estetski izraz poveže sa naučnom vizualizacijom. Zbog specifičnog oblika funkcije, u ovom primeru nisu prikazane pojedinačne tačke u koordinatnom sistemu. Ako imamo na umu nameru koju smo izneli na početku udžbenika, odnosno želju da damo prioritet razumevanju statistike u

odnosu na (tehničko) usvajanje matematičkih pojmova, možemo reći da će matematičar na slici sigurno prepoznati varijantu sinusoidne funkcije, ali statističar na njoj slobodno može da uoči i slepog miša.

> Odaberite osmu funkciju i analizirajte njen oblik. Linija koju vidite sastavljena je iz dva dela, odnosno dve funkcije. Možete li da ih uočite? Zbog čega prikazana linija nije mogla da bude nacrtana uz pomoć samo jedne formule, tj. funkcije?

2.5.3.4. Interkvartilni raspon

Najmanja moguća vrednost svih mera varijabilnosti je 0. Ona se dobija kada varijabla, odnosno pojava, uopšte ne varira i kada su sva merenja jednaka, npr. kada svi ispitanici imaju istu vrednost izmerenog svojstva. Teorijska gornja granica mera raspršenja zavisi od karakteristika same varijable, najviše od merne skale i njenog teorijskog raspona. Veće vrednosti standardne devijacije mogu se očekivati na testu na kome su đaci imali mogućnost da dobiju između 0 i 100 poena, nego na testu gde je maksimalan mogući broj poena bio 10. Iako rasponi bodova koje su đaci dobili na ova dva testa mogu da budu jednaki, *teorijski* raspon prvog testa je veći, pa je samim tim i veća mogućnost da dobijemo više vrednosti odstupanja od proseka. Sada ćemo pokušati da pronađemo najveću vrednost s u našem **primeru sa testom čiji je teorijski raspon 10 bodova**. Krenite od primera 4 sa liste i menjajte vrednosti u tabeli tako da se s povećava. Ako se najmanja varijabilnost vizuelno predstavlja kao jedan stubić i potpuna koncentracija rezultata oko jedne vrednosti, onda povećavanje varijabilnosti podrazumeva potrebu da se rezultati u većoj meri raspršе po x osi, tj. da se povećava broj merenja koja (bitno) odstupaju od proseka. Mogući međukorak u ovom pokušaju je distribucija koja se dobija odabirom primera 5 sa liste. Ovu distribuciju nazivamo *uniformnom*, ali ne zbog toga što su „uniformisani" ispitanici, tj. merenja, već zato što je ujednačena verovatnoća dobijanja bilo kog rezultata iz datog (teorijskog) raspona. U našem primeru broj loših, prosečnih i odličnih rezultata na testu potpuno je isti. Varijabilnost je naravno veća nego u prethodnom slučaju i iznosi više od 3 boda, što čini dve trećine vrednosti aritmetičke sredine.

Iako uniformna distribucija vizuelno sugeriše veoma veliku varijabilnost neke pojave, vrednost s može da bude i veća. Pri njenom daljem povećavanju

polazimo od ranije pomenute logike da standardna devijacija ne opisuje samo varijabilnost pojave, već indirektno govori i pouzdanosti aritmetičke sredine. Ako vrednosti 5 u tabeli menjamo nekim drugim vrednostima, varijabilnost će nastaviti da raste. Ekstreman slučaj nepouzdanosti M je kada ona daje potpuno pogrešnu sliku tipičnog rezultata u grupi merenja. To je slika koju prikazuje primer 6. Aritmetička sredina iznosi 5, ali ne samo da niko od đaka nije osvojio toliko bodova već niko nije osvojio ni približno toliko. Podatak da je vrednost s veoma blizu vrednosti M ili čak veća od nje, govori nam da sažimanje rezultata na vrednost aritmetičke sredine nije opravdano. Pri tome čak nije ni potrebno dati odgovor na pitanje koja varijabilnost je previše velika, jer je suština u činjenici da statistički postupci jednostavno neće „dozvoliti" da se zaključci donose na pokazateljima niske pouzdanosti. Na primer, ukoliko je jedna grupa đaka na testu znanja ostvarila rezultate prikazane u primeru 7a, a druga grupa rezultate prikazane u primeru 7b, moći ćemo da kažemo da je druga grupa bolja, jer ne samo da postoji razlika između proseka grupa, nego su ti proseci i dovoljno pouzdani. Relativno mala varijabilnost unutar obe grupe pokazuje da se doneti zaključak zaista odnosi na većinu đaka. Drugim rečima, dve prikazane distribucije mogu se lako vizuelno razdvojiti jer je njihovo preklapanje relativno malo. Međutim, ako su dobijeni rezultati kao u primerima 8a i 8b, tada ne bi trebalo da tvrdimo da je druga grupa zaista bolja, iako je razlika aritmetičkih sredina ista kao u prethodnom primeru. Naime, velika varijabilnost unutar grupa ukazuje na to da bi zaključak o postojanju razlike u suštini bio pogrešan, jer postoji puno đaka iz „bolje" grupe koji su lošije uradili test od onih iz grupe sa manjom M.

Na kraju odeljka o merama varijabilnosti pomenućemo još *interkvartilni raspon* (engl. *IQR – interquartile range*). Logika ovog pokazatelja slična je ranije pomenutoj podrezanoj aritmetičkoj sredini, a sastoji se u računanju proseka, ali ne na ukupnom rasponu, već na rasponu središnjih 50% rezultata. Interkvartilni raspon na „zanemaruje" po 25% rezultata sa obe strane distribucije, a time i potencijalne autlajere. Već smo pomenuli da tačke na x-osi kojima se definišu granice tako nastalih četvrtina rezultata nazivamo *kvartilima*. Prvi kvartil (Q_1) je tačka ispod koje se nalazi četvrtina merenja, treći kvartil (Q_3) tačka iznad koje se nalazi četvrtina rezultata, a drugi kvartil (Q_2) je zapravo medijana distribucije. Vrednost interkvartilnog raspona je razlika između Q_3 i Q_1. U primeru 5 sa liste, vidimo da Q_1 iznosi 2,5, jer se približno četvrtina đaka (3 od 11) po rezultatu nalazi ispod te vrednosti. Sa druge strane, troje đaka nalazi se iznad vrednosti 7,5. Pozicije Q_1 i Q_3 su simetrične u odnosu na središte distribucije jer je i sama

distribucija simetrična, ali to ne mora uvek da bude slučaj. U primeru 2 sa liste, vrednost Q_1 je bliža levom kraju distribucije, nego Q_3 desnom. Desni kraj distribucije je razvučen, tako da je sa te strane potreban veći raspon rezultata da bi se obuhvatila četvrtina entiteta, tj. porodica, što je u ovom slučaju, njih četiri ili pet. Pošto je distribucija asimetrična, sa njene leve strane 25% entiteta nalazi se u intervalu od 0 do 2, a sa desne između 3,75 i 10. Po istom principu mogu da se izračunaju i *kvantili*, čije vrednosti dele površinu distribucije na pet jednakih delova, ili *centili* (*percentili*) koji dele distribuciju na 100 delova.

2.6. Karakteristike i važnost normalne distribucije

U dosadašnjim primerima pokazali smo da rezultati merenja mogu da se distribuiraju na različite načine. Pomenuli smo nekoliko specifičnih oblika distribucija, kao što su uniformna ili bimodalna. Takođe smo pokazali da se oblik distribucija može opisati matematičkim funkcijama. Na primer, ukoliko je distribucija ocena na nekom predmetu uniformna, to znači da je jednak broj studenata dobio svaku od ocena. Ako je u uzorku bilo 60 studenata, onda se odnos između ocena prikazanih na x-osi i njihove učestalosti prikazane na y-osi stubičastog dijagrama, može izraziti formulom:

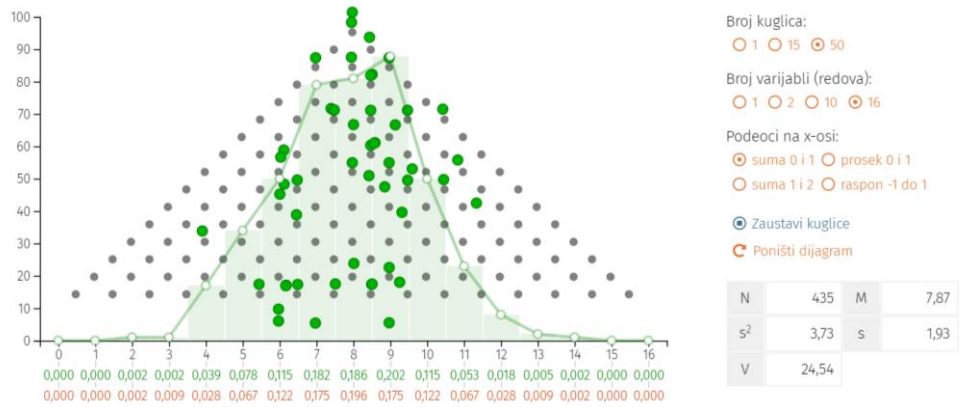
$$f(x) = 0 \cdot x + 10$$

Gornja formula označava da se za bilo koju ocenu, tj. vrednost x iz raspona od 5 do 10, očekuje da učestalost, tj. vrednost y, bude 10. Poseban značaj u statistici ima distribucija podataka koja se može opisati ranije pomenutom zvonastom krivom. U pitanju je simetrična distribucija u kojoj je najveći broj rezultata grupisan oko sredine raspona, što govori da u uzorku ili populaciji postoji najveći broj prosečnih osoba ili merenja. Izjednačavanje prosečnosti i normalnosti može se smatrati diskutabilnim, ali činjenica je da se upravo ovaj oblik distribucije u statistici naziva *normalnim*. U nastavku teksta ćemo objasniti zbog čega *normalna distribucija* ima poseban status u statistici i koje su njene karakteristike važne za postupak statističkog zaključivanja.

2.6.1. Centralna granična teorema

Svet oko nas je multivarijatan. Većina događaja i fenomena koji nas okružuju može da se posmatra kao ishod delovanja velikog broja nasumičnih

varijabli. Na prvi pogled, ta činjenica deluje obeshrabrujuće za istraživača koji želi da uoči pravilnosti u pojavama koje opisuje ili da predvidi verovatnoću ishoda nekog događaja. Srećom, pravilnost u delovanju velikog broja varijabli ipak postoji, a poznata je kao *centralna granična teorema*. Centralna granična teorema predstavlja jednu od najvažnijih postavki u oblastima statistike i teorije verovatnoće, a njenoj matematičkoj formulaciji i definiciji doprineli su brojni autori, najpre francuski matematičar Pjer-Simon Laplas početkom 19. veka, njegovi sunarodnici Abram de Moavr i Ogisten Luj Koši, a potom i ruski matematičar Aleksandar Ljapunov početkom 20. veka (Fischer, 2011).



Slika 26. Ilustracija centralne granične teoreme na primeru Goltonove table

Za potrebe ilustracije centralne granične teoreme, iskoristićemo primer jednostavnog instrumenta koji je konstruisao engleski statističar i psiholog Frensis Golton. Instrument se sastojao od drvene table u koju su ukucani ravnomerno raspoređeni klinovi. U njenom gornjem delu nalazio se levak za ubacivanje kuglica koje su se u toku pada odbijale o klinove i završavale u jednoj od pregrada u dnu table. Pojednostavljena forma ***Goltonove table*** prikazana je na početku ove vežbe i sadrži samo jedan klin, odnosno jednu nasumičnu dihotomnu varijablu koja deluje na kuglicu (Slika 26). Kada kuglica udari u klin, može da se odbije ulevo i dobije vrednost 0, ili udesno, kada dobija vrednost 1. Na osnovu teorije verovatnoće, očekujemo da će nakon većeg broja ubacivanja, polovina kuglica završiti u levoj, a druga polovina u desnoj pregradi, odnosno da će polovina „merenja" dati rezultat 0, a druga polovina rezultat 1. Očekivane ili teorijske verovatnoće prikazane su crvenim brojkama ispod svakog od mogućih ishoda na x-osi. Kada kliknete taster *Pusti kuglice*, primetićete da visine stubića u početku nisu jednake, ali da sa povećavanjem broja merenja, odnosno

veličine uzorka (kuglica), njihove visine postaju sve sličnije. To znači da na osnovu verovatnoća koje smo opazili na dovoljno velikom uzorku, možemo dosta precizno da procenimo teorijske verovatnoće ishoda nekog eksperimenta. Opažene ili empirijske verovatnoće prikazane su zelenim brojkama, a izračunate su kao proporcija, tj. odnos frekvencije svakog ishoda i ukupnog broja „merenja" (N). Sa povećavanjem veličine uzorka, empirijske verovatnoće postaju sve sličnije teorijskim. Iako su one ponekad jednake čak i na manjim uzorcima, postale bi potpuno iste, stabilne i nepromenljive, tek kada bi uzorak postao beskonačno velik, odnosno kada bismo varijablu izmerili na svim članovima populacije.

U sledećem koraku dodaćemo još jednu varijablu u sistem, biranjem opcije 2 za broj varijabli, odnosno dodavanjem još jednog reda klinova. Kliknite taster *Pusti kuglice*. Da biste brže dostigli veću vrednost N, možete da povećate broj kuglica, npr. na 15. Distribucija teorijskih verovatnoća se promenila, jer se delovanjem dve varijable povećava broj mogućih ishoda. U našem primeru ti ishodi su izraženi kao suma pojedinačnih rezultata na svakoj varijabli. Pri tome broj varijabli u sistemu nije predstavljen brojem klinova već brojem redova. Klinovi su samo vizuelni prikaz nastajanja svih mogućih permutacija koje čine određeni ishod. Ako su odabrane dve varijable, broj ishoda je 3, ali je broj permutacija 4. Ishod 0 dešava se kada se loptica oba puta odbije ulevo (0+0), 2 kada se oba puta odbije udesno (1+1), a 1 je najverovatniji ishod jer se dobija dvema permutacijama: 0+1 i 1+0. Kada broj varijabli povećamo na 10, povećava se i raspon mogućih ishoda, odnosno varijabilnost dobijenih suma. Sada postaje još očiglednije da najveći broj permutacija vrednosti pojedinačnih varijabli daje sumu koja se nalazi u centru distribucije ili sume koje su u blizini te centralne vrednosti. Upravo na ovaj fenomen odnosi se termin *centralno* u nazivu pomenute teoreme. Dobijena distribucija verovatnoća je simetrična, što znači da su ishodi koji se nalaze sa leve i desne strane medijane i proseka približno jednako verovatni, s tim da njihova verovatnoća postaje sve manja kako se sume nasumičnih varijabli približavaju krajnjim vrednostima raspona. Specifičan oblik ovakve distribucije postaje još očigledniji kada broj varijabli povećamo na 16, a postupak uzorkovanja ubrzamo povećavanjem broja kuglica na 50. Već na uzorcima većim od 100, postaje uočljivo da vrednosti 7, 8 i 9 obuhvataju više od polovine dobijenih rezultata. To možete da proverite ako saberete empirijske (zelene) verovatnoće ova tri središnja ishoda. Činjenica da se najveći broj rezultata grupiše oko proseka, vidljiva je i po tome što je linija poligona frekvencija veoma strma sa obe strane, otprilike do vrednosti 5 i 11, a

potom se znatno blaže opada do 0 i 16, kao krajnjih vrednosti teorijskog raspona distribucije. Otuda i pomenuti naziv ovakve krive koja podseća na oblik zvona. U statistici je ova zvonasta kriva poznatija kao *Gausova* ili *normalna distribucija*. Termin *granična* u nazivu teoreme odnosi se na činjenicu da sa približavanjem broja varijabli i veličine uzorka beskonačnosti, distribucija suma tih varijabli postaje sve sličnija teorijskoj normalnoj raspodeli koja se predstavlja idealno glatkom i kontinuiranom zvonastom krivom kakvu smo prikazali u odeljku o matematičkim funkcijama. Stoga se normalna distribucija smatra graničnom raspodelom ili, matematičkim rečnikom rečeno, *limesom* distribucije suma nasumičnih varijabli. Da biste jasnije videli oblik koji formira poligon frekvencija kliknite taster *Zaustavi kuglice* kada uzorak dostigne veličinu od nekoliko hiljada „merenja".

Prethodno opisani primer predstavlja samo jednu od manifestacija centralne granične teoreme. Njena važnost u statistici je mnogo veća, a oblast njene primene mnogo šira. U datom primeru sumirali smo nasumične rezultate uzete iz tzv. *Bernulijevih distribucija* kojima se opisuju dihotomne varijable, odnosno pojave koje mogu da imaju samo dva moguća ishoda (npr. tačno-netačno, uspešno-neuspešno, pismo-glava, muško-žensko). Kao rezultat, dobili smo *binomnu distribucija* koja nastaje izvođenjem većeg broja eksperimenata sa dihotomnim ishodom. Ali centralna granična teorema primenjiva je i na sve druge nasumične varijable, bez obzira na oblik njihove distribucije. Štaviše, da kada bismo umesto sume računali prosek vrednosti dobijenih na pojedinačnim varijablama, takođe bismo kao rezultat dobili približno normalnu distribuciju. Ako ponovo pustite 50 kuglica u primeru sa 16 varijabli, ali kao podeoke na x-osi odaberete proseke umesto suma nula i jedinica, primetićete da se oblik distribucije ne menja, kao ni raspodela teorijskih i empirijskih verovatnoća. Svaki rezultat na x-osi promenili (transformisali) smo na isti način, deljenjem sa brojem varijabli, koji je u ovom slučaju 16. Obratite pažnju na to da su vrednosti M i s takođe postale 16 puta manje. Podeoke na skali x-ose možete da menjate čak i u toku formiranja grafikona da biste jasnije uočili opisanu pravilnost. Uočite da varijabilnost konačnih rezultata ostaje ista čak i kada je vrednost s drugačija. Standardna devijacija menja se u apsolutnom smislu, ali to ne znači da je raspršenje rezultata drugačije u zavisnosti od toga da li računamo sumu ili prosek vrednosti 0 i 1. Drugim rečima, variranje od s = 2 u odnosu na M = 8, potpuno je isto kao variranje od s = 0,12 u odnosu na M = 0,5. To potvrđuje i pokazatelj koji se zove *koeficijent varijabilnosti*, a računa se po formuli:

$$V = \frac{s}{M} \cdot 100$$

Koeficijent varijabilnosti (V) je zapravo vrednost standardne devijacije pretvorena u procente, kao univerzalne, standardne jedinice pomoću kojih se može porediti varijabilnost različitih pojava. Ako u toku formiranja distribucije menjate način računanja konačnih rezultata između sume i proseka, primetićete da se M i s menjaju, ali da V ostaje isti, odnosno da standardna devijacija u oba slučaja iznosi približno četvrtinu ili 25% vrednosti aritmetičke sredine.

Rezultati merenja mogu da se transformišu na različite načine. Na primer, umesto vrednosti 0 i 1, ishodima nasumičnih varijabli mogli smo da dodelimo vrednosti 1 i 2. Tada bismo kao rezultat dobili sume koje se kreću u intervalu od 16 do 32, umesto od 0 do 16. Ova distribucija biće prikazana kada kao vrednosti x-ose odaberete opciju *suma 1 i 2*. Ni u ovom slučaju ne menja se oblik distribucije, već se ona samo pravolinijski „pomera" u desnu stranu koordinatnog sistema, tako da joj aritmetička sredina postaje 24, umesto prethodnih 8. Ovoga puta se standardna devijacija varijable nije promenila jer je raspon rezultata ostao potpuno isti. Međutim, koeficijent varijacije opada, jer je varijabilnost od 2, u odnosu na prosek od 24, zaista manja nego u odnosu na prosek koji iznosi 8. Opisane transformacije sirovih rezultata operacijama sabiranja, oduzimanja, množenja ili deljenja nazivamo *linearnim*, jer se relativne pozicije i relativna rastojanja među pojedinačnim rezultatima ne menjaju, tako da se ne menja ni oblik distribucije verovatnoća. Način na koji linearne transformacije utiču na vrednosti M i s možemo da prikažemo jednostavnim matematičkim formulama. Ako se vrednosti varijable X povećaju ili umanje za konstantu k, aritmetička sredina varijable povećava se ili umanjuje za istu vrednost k:

$$M_{X+k} = M_X + k$$

Ako se svaka vrednost varijable X pomnoži ili podeli konstantom k, prosek vrednosti varijable povećava se ili umanjuje k puta:

$$M_{X \cdot k} = M_X \cdot k$$

Ako se vrednosti varijable X povećaju ili umanje za konstantu k, standardna devijacija varijable ostaje ista:

$$s_{X+k} = s_X$$

Ako se vrednosti varijable X pomnože ili podele konstantom k, standardna devijacija varijable povećava se ili umanjuje k puta:

$$s_{X \cdot k} = s_X \cdot k$$

Iako se u navedenim formulama koriste operacije sabiranja i množenja, one su primenjive i na preostale dve osnovne operacije, jer je oduzimanje zapravo sabiranje sa negativnim brojem (npr. 4 - 2 = 4 + -2), a deljenje je množenje recipročnom vrednoću od 1 (npr. 4 : 2 = 4 · (1 : 2) = 4 · 0,5).

Koristeći gore navedena pravila, moguće je menjati M i s bilo koje varijable i pretvarati ih u vrednosti koje su lakše za interpretaciju, ne menjajući pri tome njihov smisao. Na primer, moguće je standardizovati sume nasumičnih varijabli iz našeg primera tako što se od dobijene sume oduzme maksimalna teorijska vrednost varijable i podeli polovinom teorijskog raspona. Distribuciju ovako dobijenih rezultata možete da vidite ako odaberete opciju *raspon -1 do 1* za vrednosti x-ose. Bez obzira na to koliko nasumičnih varijabli odaberete, konačni rezultati uvek će imati isti teorijski raspon i istu očekivanu aritmetičku sredinu. Ovako dobijene vrednosti veoma su intuitivne za interpretaciju, jer je očekivani prosek rezultata 0, a vrednosti sa leve i desne strane distribucije imaju različit predznak. Zamislite, na primer, da ste na ovaj način izrazili uspeh đaka na testu znanja. Podatak da neki đak ima vrednost 0 odmah bi vam ukazao na to da je on postigao prosečan rezultat, dok bi predznak rezultata ukazivao na to da li je neki đak bolji ili lošiji od proseka grupe kojoj pripada. Ovaj opšti princip transformacije rezultata, u cilju lakšeg poređenja ispitanika na različitim varijablama, veoma često se koristi u psihologiji i biće detaljnije objašnjen kasnije. Na ovom mestu ćemo skrenuti pažnju još samo na to da koeficijent varijabilnosti nije moguće izračunati kada je aritmetička sredina jednaka nuli. Njegove vrednosti su neprikladne i veoma nestabilne čak i kada su vrednosti M bliske nuli, što možete da zaključite na osnovu znatno većeg variranja vrednosti V kada se kao raspon podeoka na x-osi odabere raspon između -1 i 1 u odnosu na distribucije varijabli koje nastaju kao sume ili proseci vrednosti binarnih varijabli. Međutim, ako je očekivana aritmetička sredina distribucija jednaka nuli, standardizacija nije ni neophodna, jer tada vrednosti standardnih devijacija omogućavaju poređenje varijabilnosti različitih varijabli.

Na kraju ovog odeljka ukazaćemo na još jednu važnu pravilnost vezanu za sumu nezavisnih nasumičnih varijabli. Ona može da se izrazi formulom:

$$s^2_{\sum X} = \sum s^2_X$$

Naime, varijansa sume nezavisnih nasumičnih varijabli jednaka je sumi varijansi tih varijabli. Ovo pravilo možemo da demonstriramo na primeru Goltonove table. Odaberite 50 kuglica, jednu varijablu, sumu 0 i 1 kao podeoke x-ose i pustite kuglice. Nakon dovoljno velikog broja merenja, varijansa dihotomne varijable stabilizuje se oko vrednosti 0,25. Povećavanjem broja varijabli na 2, varijansa se takođe povećava dva puta. Distribucija suma 10 nasumičnih varijabli u našem primeru ima varijansu $10 \cdot 0{,}25 = 2{,}5$, dok u slučaju 16 varijabli ona iznosi $16 \cdot 0{,}25 = 4$. Pošto su varijanse svih varijabli u našem primeru iste, gornju formulu možemo da napišemo i ovako:

$$s^2_{\sum X} = n \cdot s^2_X$$

gde je n broj nasumičnih varijabli. Ukoliko varijabilnost izrazimo standardnom devijacijom, prethodna formula dobija sledeći oblik:

$$s_{\sum X} = \sqrt{n} \cdot s_X$$

Gornja formula važi za sume varijabli, ali se veoma lako može prilagoditi i primeru sa prosekom vrednosti nasumičnih varijabli:

$$s_{\frac{\sum X}{n}} = \frac{\sqrt{n} \cdot s_X}{n} = \frac{s_X}{\sqrt{n}}$$

Ovde ćemo se zaustaviti kada je u pitanju izvođenje formula, ali kasnije ćemo se vratiti na veoma važnu pravilnost opisanu poslednjom jednačinom.

2.6.2. Funkcije mase i gustine verovatnoće

U prethodnom odeljku pokazali smo da distribucija suma vrednosti nasumičnih varijabli ima oblik normalne krive kada su veličina uzorka i broj varijabli dovoljno veliki. Zbog ove pravilnosti moguće je izračunati teorijske verovatnoće svakog ishoda sumativne varijable. U slučaju Goltonove table, ne možemo tačno da znamo gde će neka od kuglica da padne, ali možemo da odredimo gde će to *najverovatnije* biti. Štaviše, očekivane verovatnoće mogu da se izračunaju veoma precizno, korišćenjem odgovarajuće formule na osnovu

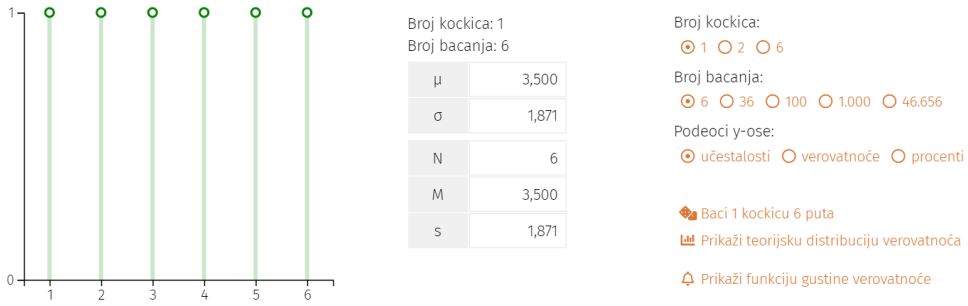
koje su određene pozicije tačaka poligona frekvencija u dvodimenzionalnom koordinatnom sistemu. U odeljku o funkcijama prikazali smo nekoliko takvih formula. Pošto pomoću njih može da se izračuna verovatnoća svakog ishoda varijable X, ove formule nazivaju se *funkcijama verovatnoće*. U slučaju binomne raspodele jedne ili više varijabli, funkcija verovatnoće ima sledeći oblik:

$$f(x) = \frac{n!}{x!\,(n-x)!} \cdot p^x \cdot (1-p)^{n-x}$$

Statistička pismenost ne podrazumeva potrebu da se ova i slične formule znaju napamet, niti da se koriste svaki put kada je potrebno izračunati verovatnoću ishoda u nekoj distribuciji. Statistički pismen istraživač treba samo da razume opštu logiku funkcija verovatnoće. Ona je, u suštini, veoma jednostavna – na osnovu poznatih parametara, kao što su npr. *x*, *n* i *p*, može se izračunati vrednost nepoznatog parametra *f(x)*. To je ujedno i osnovna logika statističkog zaključivanja i predviđanja. U gornjoj formuli *x* označava vrednost na x-osi, tj. ishod čija nas verovatnoća interesuje, *n* je broj dihotomnih varijabli, a *p* verovatnoća jednog od dva ishoda, npr. verovatnoća da kuglica odskoči ulevo ili da bacanjem novčića dobijemo pismo. Uzvičnikom se označava *faktorijel* broja, odnosno proizvod svih prirodnih brojeva koji su manji ili jednaki njemu. Izraz *f(x)* je verovatnoća određenog ishoda koja se prikazuje kao vrednost na y-osi koordinatnog sistema. Na primer, ako smo bacili 16 ispravnih novčića kod kojih verovatnoća da padne pismo iznosi 0,5, onda na osnovu formule možemo da izračunamo da verovatnoća da 5 novčića padne na pismo iznosi približno 0,067. To znači da bi oko 7% rezultata trebalo da završi u toj „pregradi", a visina stubića iznad broja 5 trebalo bi da bude 0,067. U ovom primeru vrednost *n* (broj varijabli) iznosi 16 zato što jedno merenje, odnosno jedan eksperiment, nije bio bacanje jednog već šesnaest novčića. Sa druge strane, veličina uzorka određuje koliko puta smo izveli eksperiment, odnosno bacili 16 novčića. U slučaju da je eksperiment predstavljalo bacanje jednog novčića (ponavljano više puta), distribucija verovatnoća bila bi uniformna, sa dva stubića jednake visine. Osim toga, verovatnoća da padne 5 pisama u bacanju 16 novčića ista je kao i verovatnoća da padne 11 pisama, tj. 5 glava. S obzirom na to da je varijabla koja je prikazana na x-osi diskretna, vrednosti *y* pokazuju koliko bi (proporcionalno) rezultata trebalo da se „nagomila" u svakoj tački x. Stoga ovu vrstu funkcija nazivamo *funkcijama mase verovatnoće*.

Diskusiju o teorijskim i empirijskim verovatnoćama nastavićemo na primeru varijabli sa većim brojem mogućih ishoda. Ovoga puta naglasak

stavljamo na uticaj veličine uzorka i varijabilnosti pojave na tačnost procene teorijskih verovatnoća. Kao ilustraciju ćemo upotrebiti **simulaciju bacanja kockica za igru**. Na početku je *štapićastim dijagramom* prikazana distribucija opaženih verovatnoća koja bi mogla da se dobije ako 1 kockicu bacimo 6 puta (Slika 27). Verovatnoća dobijanja bilo kog od 6 brojeva iznosi 1 : 6 ili približno 17%. U početnom primeru desilo se upravo to – svaki od brojeva dobijen je po jednom. Menjajte podeoke, odnosno jedinice u kojima su izražene vrednosti na y-osi da biste lakše uočili vezu između učestalosti ishoda i njihove verovatnoće koja može da se izrazi kao relativna frekvencija ili kao procenat. U početnom primeru distribucija empirijskih verovatnoća jednaka je distribuciji teorijskih, jer je dobijen rezultat koji u potpunosti odgovara očekivanjima.



Slika 27. Štapićasti dijagram koji prikazuje slučajnu distribuciju brojeva dobijenih bacanjem kockice šest puta

Ako ste u prvom bacanju dobili broj 3, kolika je verovatnoća da u drugom bacanju dobijete isti broj?

Ako kockicu bacite 6 puta, da li je verovatnije da ćete dobiti brojeve ovim redom: 1, 1, 1, 1, 1, 1 ili ovim redom: 2, 4, 6, 1, 3, 5?

Ako istovremeno bacite 6 kockica, da li je verovatnije da ćete dobiti brojeve 1, 1, 1, 1, 1, 1 ili 2, 4, 6, 1, 3, 5?

Da li biste rekli da je verovatnoća da dobijete distribuciju prikazanu na početku ove vežbe, nakon šest bacanja jedne kockice, mala ili velika?

Prethodna pitanja vezana su za neke od čestih grešaka u zaključivanju o verovatnoćama. Prvo se odnosi na pogrešnu pretpostavku o međusobnoj

povezanosti ishoda događaja. Ishod bacanja kockice ni na koji način ne utiče na ishode njenih narednih bacanja jer su u pitanju potpuno nezavisni događaji. Verovatnoća da dobijete broj 3 ista je u prvom bacanju, kao i u svakom narednom. Druga greška odnosi se na izjednačavanje *kombinacija* i *permutacija* elemenata iz određenog skupa ishoda. Kod ovih drugih, svaki mogući niz elemenata skupa ishoda tretira se kao drugačiji događaj, tako da 1-2-3 nije isto što i 2-3-1. Na primer, dve dihotomne varijable daju tri kombinacije ishoda (dve 0, dve 1 i jedna 0 i jedna 1), ali četiri permutacije (0-0, 1-1, 0-1 i 1-0). Ako se vratimo na primer sa kockicama, permutacije 1-1-1-1-1-1 i 2-4-6-1-3-5 imaju potpuno iste verovatnoće. Jednako je (malo) verovatno da će se u 6 bacanja kockica dobiti baš taj redosled brojeva. Međutim, ako nas interesuje samo prisustvo tih brojeva u skupu od 6 brojeva, onda je reč o kombinacijama. Tada je verovatnoća da se dobiju svi mogući brojevi zaista veća od one da se dobiju samo jedinice, jer je veći broj kombinacija koji nam daje željeni ishod. Tako dolazimo i do odgovora na poslednje pitanje. Ukoliko istovremeno bacite 6 kockica, broj permutacija može da se izračuna kao 6^6, što daje 46.656 mogućih ishoda. Sa druge strane, ako nas interesuje koliko kombinacija svih 6 brojeva postoji, rezultat ćemo dobiti izračunavanjem vrednosti 6!, što iznosi 720. Dakle, verovatnoća da nakon bacanja 6 kockica dobijemo sliku koju vidimo na grafikonu je izuzetno mala i iznosi 720 : 46.656 ≈ 1,5%. U tih 720 ishoda, na primer, spadaju permutacije 2-4-6-1-3-5, 1-2-4-5-6-3, 3-4-2-1-5-6, 5-4-6-3-1-2 i sve ostale koje sadrže svaki od 6 brojeva. Međutim, verovatnoća da dobijemo sve jedinice još je manja i iznosi 1 : 46.656, jer samo jedna permutacija daje željeni ishod: 1-1-1-1-1-1.

Kako se pravilnosti koje smo opisali odražavaju na istraživačku praksu? Zamislite da na osnovu uzorka veličine 6 želite da opišete distribuciju neke pojave u populaciji. Kada kliknete taster *Baci 1 kockicu 6 puta*, biće prikazani simulirani rezultati koje biste mogli da dobijete na odabranom uzorku veličine 6. Obratite pažnju na dve stvari. Prvo, svaki put kada kliknete taster i napravite novi uzorak, verovatno ćete dobiti drugačiju distribuciju opaženih verovatnoća. Drugo, malo je verovatno da ćete na uzorku veličine 6 dobiti distribuciju koja tačno odslikava teorijsku raspodelu verovatnoća u populaciji koja bi trebalo da je uniformna. Postavite vrednosti y-ose na verovatnoće i pređite pokazivačem miša preko tačkica da biste videli empirijske verovatnoće svakog ishoda. Kao što smo rekli, u pitanju je proporcija izražena kao odnos broja poželjnih ishoda i ukupnog broja mogućih ishoda. Empirijske verovatnoće najverovatnije se razlikuju od teorijskih koje iznose 0,167. Na osnovu prikazanog primera postaje

jasno da je svaki uzorak koji koristimo u nekom istraživanju samo jedan od velikog broja uzoraka koji mogu da se uzmu iz iste populacije. Osim toga, ukoliko je uzorak mali, vrlo je verovatno da će slika populacije, a samim tim i zaključci o njoj, biti pogrešni. Kada veličinu uzorka povećate na 36 (bacanja), primetićete da se empirijske verovatnoće približavaju teorijskim. Za početak, gotovo je nemoguće da nakon 36 bacanja zaključite da je verovatnoća nekog od ishoda 0, kao što se dešavalo u uzorcima veličine 6. Dobijena slika stanja u populaciji i dalje nije potpuno precizna, ali će postajati sve tačnija sa daljim povećavanjem veličine uzorka. U statistici i teoriji verovatnoće ovaj fenomen se naziva *zakonom velikih brojeva*. Kada 1 kockicu bacite 46.656 puta, dobićete distribuciju koja je gotovo identična teorijskoj. Tada se opažene verovatnoće razlikuju od teorijskih tek na trećoj decimali. Međutim, već na uzorku veličine 100, razlike između opaženih i očekivanih učestalosti postaju relativno male i jasno upućuju na uniformnu distribuciju verovatnoća. U praktičnom smislu, to znači da je u istraživanjima poželjno imati što veći uzorak merenja, odnosno ispitanika, ali da to povećanje ima smisla samo do određene granice. Sada odaberite opcije bacanja 2 kockice 6 puta i kliknite taster *Baci 2 kockice 6 puta*. U ovom primeru, kao rezultati varijable prikazane dijagramom, koriste se zbirovi brojeva dobijenih na dve istovremeno bačene kockice.

> Da li možete da procenite kako izgleda teorijska distribucija verovatnoća u ovom primeru?
>
> Da li je jednako verovatno da prilikom bacanja dve kockice dobijete zbir 2, 12 i 7? Koja vrednost je verovatnija i zašto?

Na osnovu uzorka veličine 6, praktično je nemoguće predvideti izgled teorijske distribucije varijable. Situacija se popravlja na uzorku veličine 36, ali tek nakon 100 bacanja kockica može se naslutiti da distribucija u populaciji nije uniformna i da su najverovatniji događaji oni koji se nalaze oko središnje vrednosti. Ovi događaji su najverovatniji jer najveći broj permutacija dobijenih brojeva daje sumu 7, a najmanji sume 2 i 12. Kada prikažete teorijsku distribuciju verovatnoća, biće vam jasnije kakav je odnos teorijskih učestalosti svake od 11 mogućih suma u rasponu od 2 (samo 1+1), preko 7 (npr. 5+2, 2+5, 4+3, 6+1), do 12 (samo 6+6). Ako nastavite dalje i odaberete bacanje 6 kockica 100 puta, primetićete da je tada teško predvideti teorijske verovatnoće čak i na uzorku koji je ranije bio dovoljno velik da pruži barem približno tačnu sliku distribucije varijable u populaciji. Ali čak ni u ovom slučaju nije neophodno praviti uzorak

veličine 46.656, jer se sasvim prihvatljiva procena dobija i na uzorku veličine 1.000. Ipak, zbog povećanja broja mogućih ishoda, odnosno povećanja raspona rezultata, potrebno je formirati veći uzorak kako bi se dobila adekvatna slika stanja u populaciji. To znači da granica na osnovu koje bi se definisao *dovoljno veliki* uzorak, prvenstveno zavisi od broja i raspona mogućih ishoda, odnosno od varijabilnosti pojave koju želimo da opišemo. Što je varijabilnost pojave veća, biće potreban veći uzorak da bi se ona pouzdano opisala. Na primer, ukoliko odaberete 6 za broj kockica i kliknete taster *Prikaži teorijsku distribuciju verovatnoća*, videćete da je aritmetička sredina te varijable u populaciji μ = 21, a njena standardna devijacija σ = 4,183 (Slika 28). To su vrednosti koje očekujemo ali koje najverovatnije nećemo dobiti na malim uzorcima. Ako odaberete 6 bacanja i nekoliko puta napravite uzorke veličine 6, primetićete da se aritmetičke sredine (M) i standardne devijacije (s) uzoraka često značajno razlikuju od populacijskih parametara. Međutim, sa povećanjem veličine uzorka, raste tačnost procene parametara populacije (μ i σ) na osnovu statistika uzorka (M i s). Malo je verovatno da će u velikom nasumičnom uzorku procene parametara biti izuzetno netačne.

Slika 28. Teorijska distribucija zbirova brojeva dobijenih bacanjem šest kockica

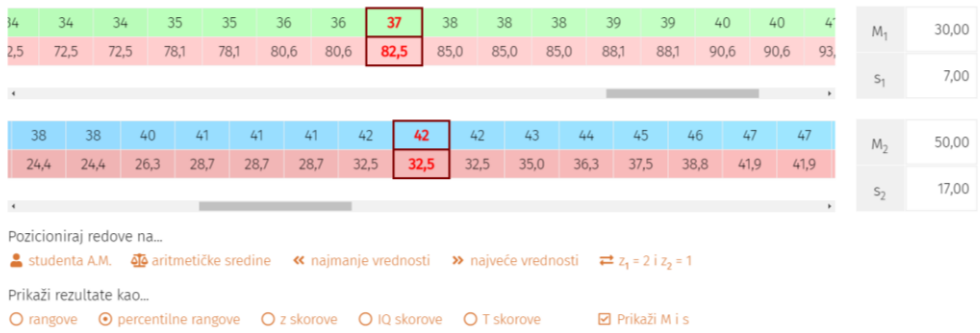
Primetili ste da se sa povećavanjem broja kockica čiji zbir računamo, povećava i raspon dobijenih rezultata, odnosno broj podeoka na x-osi. Samim tim, verovatnoća pojedinačnih ishoda postaje sve manja, jer suma verovatnoća, koja uvek iznosi 1, mora da se podeli na više delova. Na primer, verovatnoća broja 7 kao najverovatnijeg ishoda zbira brojeva dobijenih bacanjem 2 kockice, iznosi oko 0,167, dok verovatnoća vrednosti 21, kao najčešće sume brojeva dobijenih bacanjem 6 kockica, iznosi oko 0,093. To znači da bi sa daljim povećavanjem broja podeoka na x-osi, verovatnoća svake od tih vrednosti postajala sve bliža 0. Štaviše, verovatnoća pojedinačnih ishoda zaista bi postala

0 kada bi broj podeoka postao beskonačan, tj. kada bismo hteli da opišemo distribuciju varijable koja nije diskretna već kontinuirana. Na prvi pogled, ova pravilnost deluje nelogično, ali treba posmatrati njene praktične implikacije. Kada opisujemo neku varijablu, najčešće nas ne interesuje samo jedan rezultat, već *raspon* rezultata u okviru koga se nalazi najveća koncentracija verovatnoća. Stoga se matematičke funkcije kojima se opisuju raspodele verovatnoća kontinuiranih varijabli, nazivaju funkcijama *gustine* verovatnoća. Ovaj princip često se primenjuje i na diskretne varijable čije skale imaju puno podeoka. Tako ćete, na primer, češće nailaziti na tvrdnje da se najverovatnija ili prosečna inteligencija nalazi u rasponu između 90 i 110, nego da iznosi tačno 100. U tim slučajevima se zapravo pretpostavlja da je varijabla u populaciji kontinuirana. Zbog toga grafički prikaz funkcije gustine verovatnoće više ne izgleda kao niz tačaka koje formiraju poligon frekvencija, već kao neprekidna, glatka kriva koja blago i postepeno menja svoj oblik od početne do krajnje tačke raspona. U slučaju normalne raspodele, koju dobijamo kao sumu velikog broja nasumičnih varijabli, ta funkcija se predstavlja zvonastom krivom koju možete da prikažete kada kliknete taster *Prikaži funkciju gustine verovatnoće*. Na osnovu površine osenčene crvenom bojom, može se zaključiti da je većina rezultata obuhvaćena intervalom između 17 i 25, što približno odgovara rasponu koji biste dobili ako od aritmetičke sredine oduzmete vrednost jedne standardne devijacije i na nju dodate istu tu vrednost. Ako prikažete teorijsku distribuciju verovatnoća za bacanje 6 kockica i saberete pojedinačne verovatnoće vrednosti 17 do 25, dobićete sumu koja je veća od 70%. Dakle, verovatnoća da se neki ishod u normalnog distribuciji nalazi između vrednosti M - s i M + s iznosi približno 0,7. Osim toga, može se videti i da je verovatnoća dobijanja rezultata koji odstupaju od proseka za više od 2 standardne devijacije ulevo ili udesno, veoma mala ako su rezultati distribuirani normalno. Ovakav način interpretacije verovatnoća događaja čini osnovu statističkog zaključivanja, o čemu će biti više reči u trećem poglavlju.

2.6.3. Standardizacija sirovih rezultata

Standardizacija je višeznačan pojam, čak i u okviru pojedinačnih oblasti kao što su psihologija ili statistika. Načelno, standard podrazumeva postojanje opšteprihvaćenog principa, konsenzusa i uređenosti načina upotrebe nekog proizvoda ili sprovođenja nekih aktivnosti. Primeri standarda su saobraćajni znaci, internet protokoli za razmenu podataka, dozvoljena emisija ugljen-

dioksida iz motora ili sistem međunarodno priznatih mernih jedinica, kao što su metar, kilogram ili amper. Ranije pomenuti koeficijent varijacije je takođe oblik standardizacije koja omogućava poređenje varijabilnosti varijabli koje imaju različite standardne devijacije i aritmetičke sredine. Jasno je da nepostojanje standarda otežava komunikaciju i razumevanje pojava koje se opisuju. Za većinu građana kontinentalne Evrope deluje zbunjujuće informacija da je neka osoba visoka 5,5 stopa ili da se automobil kreće brzinom od 65 milja na sat. Zamislite tek kakvi bi problemi u komunikaciji nastali da se i dalje koriste sve arhaične mere za širinu i dužinu koje su različite kulture upotrebljavale kroz vekove, kao što su npr. lakat, koji uz to može da bude biblijski i grčki, hvat, furlong, liga, perč, trska ili šaka. U oblasti psihologije, pod standardizacijom se najčešće podrazumeva prilagođavanje instrumenata za korišćenje u određenoj populaciji, npr. adaptacija nekog upitnika za primenu u populaciji adolescenata ili adaptacija testa sposobnosti, nastalog u SAD za korišćenje u drugoj državi. Pitanjima standardizacije testova bavi se granična oblast statistike i psihologije koja se zove *psihometrija*. U ovom odeljku biće više reči o drugačijoj vrsti standardizacije koja podrazumeva linearno transformisanje sirovih rezultata merenja u cilju njihove lakše interpretacije i poređenja među različitim ispitanicima i/ili varijablama. U tom smislu, standardizacija podrazumeva promenu jedinica mere u kojima su izražene vrednosti varijable, npr. bodova, sekundi ili centimetara, u univerzalne i samorazumljive jedinice, a u skladu sa unapred definisanim i opšteprihvaćenim pravilima.



Slika 29. Primer sirovih rezultata i percentilnih rangova studenata na testu znanja

Zamislićemo da je **80 studenata radilo dva testa znanja**, zeleni i plavi. Bodovi svih studenata prikazani su u dva reda i sortirani su od najnižeg ka najvišem (Slika 29). Rezultati na oba testa distribuirani su normalno, što znači da je najviše vrednosti grupisano oko medijane i proseka. Student A. M.

postigao je 37 bodova na prvom testu a 42 na drugom. Kućice sa rezultatima ovog studenta možete da locirate u grupi svih rezultata pomeranjem klizača ispod oba reda rezultata ili klikom na opciju *Pozicioniraj redove na … studenta A. M.* Pokušajte da odgovorite na sledeća pitanja samo na osnovu sirovih rezultata koje je A. M. postigao na oba testa:

> Da li je A. M. dobro ili loše uradio zeleni test?
>
> Da li je A. M. bolje uradio zeleni ili plavi test?
>
> Koliko je A. M. bolji na plavom testu od nekoga ko je osvojio 25 bodova?

Očigledno je da na osnovu dva sirova rezultata nećete moći da date odgovore na postavljena pitanja, jer vam nedostaje kontekst koji bi tim podacima dao smisao, npr. podatak o tome koliki je bio teorijski raspon bodova na testovima ili koliki je prosečan učinak svih studenata. Međutim, pozicija klizača ispod svakog reda ipak nam govori nešto o učinku A. M., jer je njegov rezultat na zelenom testu pomeren više udesno, ka većim rezultatima, a na plavom je upravo suprotno. Čak i na taj način smo izvršili standardizaciju, jer smo učinak A. M. izrazili kao poziciju u nizu rezultata koji su sortirani, odnosno rangirani. Ako kliknete taster *Prikaži rezultate kao … rangove*, biće prikazane i konkretne vrednosti tih pozicija za svakog ispitanika. Kada uspeh A. M. izrazimo njegovim rangovima na listi svih studenata (15 i 55) umesto sirovim brojem bodova (37 i 42), možemo da zaključimo da je A. M. relativno dobro uradio prvi test i da ga je verovatno uradio bolje od drugog testa, naravno pod uslovom da smo rang 1 dodelili najboljem studentu u grupi. Obratite pažnju na to da neki studenti imaju vrednosti rangova koje nisu celobrojne. U pitanju je način računanja tzv. *spojenih rangova* za više istih rezultata. Ako su npr. 7, 8, 9. i 10. student ostvarili isti broj poena na testu, onda i njihov rang treba da bude jednak. Zajednički rang obično se računa kao prosek rangova koji bi inače bili dodeljeni vrednostima po redosledu na listi, u ovom slučaju 8,5.

Korišćenje rangova kao načina standardizacije vrednosti varijable ima dva nedostatka. Prvi je činjenica da rangovi nisu *ekvidistantne* jedinice, jer zavise od gustine rezultata u različitim delovima distribucije. Na primer, kada pogledate rezultat studenta A. M., primetićete da je njegov rang na plavom testu bolji za 3 pozicije od studenata koji su osvojili samo bod manje. Ako pomerite plavi red ka najmanjim vrednostima, uočićete da se poslednji i

pretposlednji student međusobno razlikuju za čak 5 bodova, iako je razlika njihovih rangova samo 1. To znači da rangovi pružaju informaciju o tome koji rezultat je bolji ili lošiji od nekog drugog, ali ne i za koliko. Drugi i bitniji nedostatak upotrebe rangova kao standardnih mera je činjenica da njihove vrednosti zavise od broja merenja. U ovom primeru, rangovi na različitim testovima uporedivi su samo zato što je oba testa radio isti broj studenata. Da je, na primer, drugi test uradilo deset studenata manje, rangovi više ne bi bili uporedivi, jer bi na prvom testu rang 40 označavao da se student nalazi otprilike na polovini rang liste, dok bi na drugom ukazivao da je on po učinku u donjoj polovini grupe. Uostalom, mi tvrdimo da je A. M. „relativno dobro" uradio test samo zato što znamo da je u grupi bilo 80 studenata. Rang 15 bi imao potpuno drugačije značenje da je ostvaren u grupi od 20 studenata. Ovaj nedostatak ispravlja se veoma lako, tako što se rang svakog rezultata podeli najvećim rangom u grupi i pomnoži sa sto. Na taj način se dobijaju *percentilni rangovi* čije vrednosti se uvek kreću u rasponu od 0 do 100 i pokazuju na kom procentualnom delu distribucije podataka se nalazi neki rezultat. Ovde treba napomenuti da smer rangiranja obično nije isti za ova dva oblika transformacije rezultata. Uobičajeno je da se rangiranje po uspehu obavlja tako što se prvi rang dodeli najboljem rezultatu. Međutim, kod percentilnih rangova logika je upravo suprotna. Najniži percentilni rang dodeljuje se najmanjem rezultatu, odnosno krajnjem levom rezultatu neke distribucije podataka. To znači da pre formiranja percentilnih rangova, bodove studenata treba sortirati obratno, tako da najbolji student dobije najniži rang. Upravo to su rezultati koje dobijate ako kliknete taster *Prikaži rezultate kao … percentilne rangove*. Najbolji student sada dobija transformisani rezultat 100, što znači da je 100% studenata u grupi po učinku jednako ili lošije od njega. Analogno tome, student koji ima percentilni rang 50, nalazi se na polovini rang liste. Student A. M. ima percentilni rang 82,5 na zelenom testu, što znači da ga je uradio bolje od približno 82% studenata. Njegov percentilni rang na plavom testu ukazuje na to da ga je uradio znatno lošije od zelenog, jer je na tom testu bolji od približno trećine studenata. Sirovi rezultati koji su povezani sa odgovarajućim percentilnim rangovima nazivaju se *percentili* ili *centili*. Kao što smo rekli u odeljku o interkvartilnom rasponu, to su podeoci koji dele distribuciju na 100 jednakih delova. Na primer, vrednost 31 na zelenom testu je 60. percentil distribucije, odnosno vrednost ispod koje se nalazi oko 60% rezultata svih ispitanika. Treba imati na umu da percentilni rangovi, kao ni rangovi, nisu ekvidistantne jedinice, tako da pomenutih 100 delova distribucije nije jednako po rasponu sirovih rezultata, već po njihovom broju. Na primer, na plavom testu, isti broj (procenat) rezultata (studenata)

nalazi se između vrednosti 40 i 44, kao i između vrednosti 23 i 34. Veći raspon koji je potreban da bi se obuhvatilo 10% rezultata u zoni nižih vrednosti ukazuje na to da je više studenata grupisano oko vrednosti 40, dok su levi i desni kraj (normalne) distribucije razvučeni, pa se u tim zonama nalazi manji broj rezultata. Da je distribucija bodova bila uniformna, rangovi bi postali ekvidistantni, jer bi isti raspon bodova, u bilo kom delu distribucije, obuhvatao isti procenat rezultata, odnosno studenata.

Položaj rezultata u grupi najbolje ćemo razumeti ako su nam poznate aritmetička sredina i standardna devijacija te grupe. Kada prikažete vrednosti M i s za oba testa, postaje jasno da je A. M. bolji od proseka grupe na zelenom, a lošiji od proseka na plavom testu. Pored toga, vidimo da je njegovo odstupanje od proseka izraženo apsolutnim brojem bodova veće na plavom testu (37 - 30 = 7), nego na zelenom (42 - 50 = -8). Međutim, kao što nam sirovi broj bodova A. M. ne govori puno o njegovom učinku van odgovarajućeg konteksta, tako nam ni njegovo odstupanje u broju bodova od proseka ne znači mnogo ako ne uzmemo u obzir varijabilnost rezultata. Kao što se vidi na osnovu raspona i standardnih devijacija, varijabilnost je veća na plavom testu. Razlika između najboljeg i najlošijeg studenta na plavom testu iznosi 83 boda, dok je raspon bodova na zelenom nekoliko puta manji. U skladu sa tim, može se reći da je odstupanje od 8 bodova na plavom testu, u relativnom smislu, manje nego odstupanje od 7 bodova na zelenom. Preciznije, prva razlika je blizu polovine vrednosti s_2, a druga je jednaka vrednosti s_1. Ovom logikom dolazimo do postupka standardizacije koji se veoma često koristi u psihologiji, a poznat je kao transformacija u *z (zet) skorove* ili *z vrednosti*. Standardizacija pretvaranjem u z vrednosti uzima u obzir kako prosek, tako i varijabilnost rezultata, a vrši se prema sledećoj formuli:

$$z = \frac{x - M}{s}$$

Svako *x* u grupi merenja pretvara se u njegovo odstupanje od proseka grupe kojoj pripada, pri čemu se to odstupanje izražava u broju standardnih devijacija. Kada kliknete opciju *Prikaži rezultate kao … z skorove*, u donjim redovima će biti prikazane standardizovani z skorovi svakog studenta. Student A. M. dobija z skor 1 na zelenom testu, jer je od proseka rezultata bolji za vrednost jedne s, dok je njegov standardizovani skor na plavom testu oko -0,5, zato što je na njemu postigao rezultat koji je za približno pola standardne devijacije lošiji od proseka svih studenata. To znači da studenti koji su ostvarili prosečan rezultat,

nakon standardizacije dobijaju z vrednost 0, pa tako i aritmetička sredina novonastale varijable postaje 0. Standardna devijacija standardizovane (z) varijable, kao i njena varijansa, uvek iznosi 1, što može da se dokaže pomoću formula za računanje z vrednosti i pravilnosti koje smo pomenuli u poglavlju o centralnoj graničnoj teoremi:

$$s_z = s_{\frac{X-M}{s}} = \frac{s_{X-M}}{s} = \frac{s_X}{s} = \frac{s}{s} = 1$$

S obzirom na to da su rezultati koje koristimo u ovom primeru distribuirani normalno, standardizovane vrednosti bodova nam omogućavaju da donesemo određene zaključke o gustini verovatnoća ispod normalne krive. Odaberite opciju *Pozicioniraj redove na ... $z_1 = 2$ i $z_2 = 1$* da biste označili kućice u kojima se nalaze rezultati studenta koji je na plavom testu bolji za 2 standardne devijacije od proseka i studenta koji je na zelenom testu bolji za jednu standardnu devijaciju od proseka. Ako standardizovane vrednosti izrazite kao percentilne rangove, možete zaključiti da je vrednosti z = 1 analogna vrednost 85. centila, a da vrednosti z = 2 odgovara vrednost 97,5. centila. Ova pravilnost je karakteristika svake (približno) normalne distribucije. Ukoliko je neka varijabla normalno distribuirana, već na osnovu z vrednosti može da se zaključi kolika je verovatnoća ishoda koji se nalaze iznad i/ili ispod te tačke. U našem primeru, student koji ima z vrednost 2 na testu čiji su rezultati normalno distribuirani, bolji je od približno 97,5%, a lošiji od približno 2,5% svojih kolega. Pošto je normalna distribucija simetrična, isti princip je primenjiv i na suprotni kraj distribucije. Ako pronađete poziciju studenta koji je osvojio 16 bodova na plavom testu, primetićete da tom broju odgovara z vrednost -2, odnosno percentilni rang 2,5. Student koji na plavom testu ima 16 bodova, bolji je od samo 2,5% svojih kolega. To znači da se približno 95% vrednosti varijable koja je normalno distribuirana nalazi između vrednosti M - 2 · s i vrednosti M + 2 · s. Ako isti princip primenimo za analizu vrednosti koje se nalaze na pozicijama z = 1 i z = -1 na oba testa, dolazimo do zaključka da interval M ± 1 · s obuhvata oko 68% rezultata. Ove pravilnosti vezane za gustinu verovatnoća, odnosno površinu ispod normalne krive, biće detaljnije opisane u narednom poglavlju.

Standardizacija sirovih rezultata merenja olakšava *intraindividualno* i *interindividualno* poređenje. Prvi termin označava mogućnost poređenja rezultata iste osobe na različitim varijablama, npr. na dva testa znanja. Drugi termin označava mogućnost poređenja različitih osoba na istim, pa čak i na različitim varijablama. Zvuči pomalo čudno, ali može se reći da je osoba koja

ima z = 1,64 na težini, a z = -0,98 na visini, više teška nego što je visoka. Pri tome uvek treba imati na umu da je u pitanju *relativna* vrednost varijable, izračunata na osnovu *proseka* grupe rezultata. U našem primeru sa testovima, student koji je osvojio 47 bodova na zelenom testu dobiće z skor 2,43, što upućuje na njegov izuzetno dobar učinak. Međutim, ukoliko pretpostavimo da je teorijski raspon na testu bio 100 bodova, njegov apsolutni učinak više nije tako dobar. Međutim, on je ipak najbolji *u svojoj grupi*, tj. u kontekstu uspeha ostalih studenata na istom testu. To je suština standardizacije u psihologiji, jer se pojam „normalnog" ili „prosečnog" najčešće definiše na osnovu rezultata referentne grupe u kojoj se merenja obavljaju. Kada bi svi stanovnici planete Zemlje naprasno postali inteligentniji i imali u proseku koeficijent inteligencije 120, onda bi se upravo ta vrednost proglasila „normalnom".

Još jedna pogodnost koju nude z skorovi jeste mogućnost da se njihov raspon prilagodi vrednostima koje su istraživaču najpogodnije za potrebe analize i interpretacije podataka. Aritmetička sredina i standardna devijacija z distribucije mogu da se izmene jednostavnim računskim operacijama, a da se pri tome očuva njihov smisao i relativna pozicija ispitanika u grupi. Na primer, odluka da se prosečna vrednost IQ skale označi sa 100, nije ništa drugo nego međunarodni konsenzus psihologa. Prosečan IQ skor mogao je isto tako da bude označen nulom ili vrednošću 50. To je jedan od razloga zbog kojih IQ skalu smatramo intervalnom a ne razmernom. Ako zamislimo da su zeleni i plavi test bili testovi sposobnosti, sve bodove možemo da transformišemo u vrednosti standardne IQ skale tako što najpre sirove rezultate pretvorimo u z skorove, a potom svaki z skor pomnožimo vrednošću željene standardne devijacije i saberemo sa vrednošću željene aritmetičke sredine. U slučaju IQ skale, te vrednosti su $\sigma = 15$ i $\mu = 100$. Ovako dobijeni IQ skorovi prikazuju se kada kliknete opciju *Prikaži rezultate kao … IQ skorove*. Biti prosečan sada ne znači imati 30 bodova, već imati IQ 100. Analogno tome, biti bolji od 97,5% studenata sada znači imati IQ skor 130. Na sličan način može se obaviti bilo koja druga vrsta linearne transformacije sirovih rezultata. Na primer, u oblasti psihometrije često se koristi skala *T skorova*, čija je aritmetička sredina 50, a standardna devijacija 10.

2.6.4. Površina ispod normalne krive

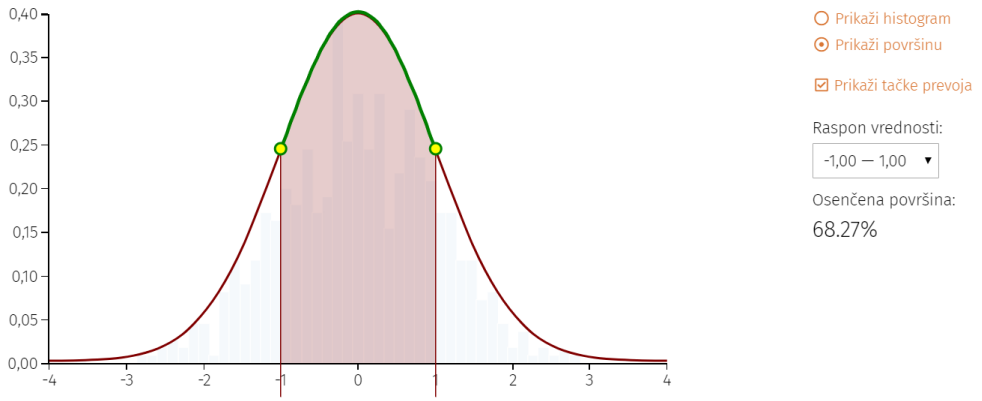
Normalna distribucija je jedna od najznačajnijih distribucija u statistici. Već smo pokazali da centralna granična teorema pruža osnov za procenu

verovatnoća ishoda koji nastaju kao suma nasumičnih varijabli. Opisane pravilnosti su posebne korisne u oblasti inferencijalne statistike. Normalna distribucija predstavlja važan fenomen i u oblasti psihologije, jer se veliki broj varijabli, npr. osobine ličnosti ili različite sposobnosti, distribuira na ovaj način. Stoga bi trebalo poznavati osnovne karakteristike normalne distribucije kao teorijskog modela i matematičke funkcije. Uzmimo kao primer **rezultate na testu sposobnosti izražene IQ skorovima** za koje se smatra da su u populaciji normalno distribuirani (Slika 30). Na početku je prikazan histogram IQ vrednosti uzorka veličine 500 uzetog iz opšte populacije. Svaki put kada kliknete opciju *Prikaži histogram*, biće prikazivani rezultati novog i drugačijeg uzorka od 500 osoba uzetih iz iste opšte populacije. Dobijene distribucije se međusobno razlikuju, ali upućuju na isti zaključak – najverovatnije je da će se iz populacije nasumično odabrati osobe koje su prosečne inteligencije. Pri tome se, kao što smo rekli, ne misli na IQ skor 100 kao očekivanu vrednost aritmetičke sredine, već na raspon IQ skorova u blizini te vrednosti koji obuhvata najveći broj članova populacije i najveći broj osoba u uzorku. Svaka od dobijenih empirijskih distribucija rezultata u uzorcima, može približno da se predstavi ili *aproksimira* teorijskom krivom koja je prikazana kao crvena linija. Ova linija predstavlja funkciju gustine verovatnoće koja se primenjuje na sve normalno distribuirane varijable. Već smo pomenuli da svaka funkcija verovatnoće može da se izrazi odgovarajućom formulom, a u slučaju normalne distribucije, ta formula izgleda ovako:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{\frac{(x-\mu)^2}{2\sigma^2}}$$

Pošto su *e* (*Ojlerov broj*) i π (*Arhimedov broj*) konstante, funkcija pokazuje da verovatnoća bilo koje vrednosti *x* normalno distribuirane varijable može da se predvidi na osnovu dva njena parametra – aritmetičke sredine μ i standardne devijacije σ. Ako u formulu uvrstimo vrednosti $x = 130$, $\mu = 100$ i $\sigma = 15$, možemo da zaključimo da je verovatnoća da neka osoba u opštoj populaciji ima IQ veći od 130 veoma mala. Štaviše, možemo da budemo još precizniji ukoliko koristimo pravilnosti vezane za površinu ispod normalne krive. Ranije smo objasnili da se, za razliku od funkcije mase verovatnoće, funkcija gustine ne koristi za opisivanje verovatnoće pojedinačnog ishoda, već za opisivanje gustine verovatnoća u određenom rasponu ishoda. Na primer, kada odaberete opciju *Prikaži površinu*, biće prikazana površina ispod normalne krive koju formira interval vrednosti između jedne standardne devijacije levo i jedne

standardne devijacije desno od aritmetičke sredine. U slučaju standardizovane varijable čija je aritmetička sredina 0, to je interval između -1 i 1, dok bi u slučaju IQ, to bio interval između 85 (100 - 15) i 115 (100 + 15). Kao što vidite, taj interval obuhvata približno 68% svih rezultata. To znači da, ukoliko je inteligencija zaista distribuirana normalno sa μ = 100 i σ = 15, približno 68% populacije ima IQ između 85 i 115. Ili, drugačije rečeno, ukoliko nasumično birate osobe iz opšte populacije, verovatnoća da odaberete nekoga ko ima IQ između 85 i 115 iznosi 68%. Ako počnemo da proširujemo osenčeni interval, njime će biti obuhvatan sve veći i veći broj ishoda, odnosno rezultata. Međutim, taj priraštaj će biti sve manji kako se približavamo krajevima distribucije. Postavite i levu i desnu granicu osenčene površine na vrednost 0, potom jednu od njih polako pomerajte ka desnom kraju distribucije. Obratite pažnju na to da do vrednosti 1, odnosno μ + 1σ, procenat raste veoma brzo, a nakon te tačke dosta sporije, da bi između 3. i 4. standardne devijacije porast postao gotovo beznačajan. Istu pravilnost uočićete dok levu granicu pomerate ka levom kraju distribucije. Ova promena iz brzog u spori priraštaj dešava se na precizno definisanim mestima koja se nazivaju *tačke prevoja* ili *infleksije*. Kada kliknete opciju *Prikaži tačke prevoja*, na crvenoj liniji biće označene tačke prevoja u kojima distribucija menja svoj tok i prelazi iz konkavnog (ispupčenog) u konveksan (udubljen) oblik. Uočite da se tačke prevoja nalaze iznad vrednosti jedne standardne devijacije levo i desno od proseka.



Slika 30. Prikaz teorijske normalne distribucije sa tačkama prevoja

Na osnovu do sada opisanih pravilnosti, možemo da sažmemo nekoliko osnovnih svojstava teorijske normalne distribucije. Najpre, ona je *simetrična*, što znači da su njena leva i desna strana potpuno jednake u odnosu na središnju osu i aritmetičku sredinu. Drugim rečima, 50% ishoda nalazi se u rasponu od 0

do levog kraja distribucije, kao i od 0 do njenog desnog kraja. U vezi sa tim, teorijska normalna distribucija ima samo jedan očigledan vrh iznad aritmetičke sredine, pa se stoga kaže da je *unimodalna*. Međutim, ukoliko pokušate da prikažete verovatnoću moda ili najčešće vrednosti postavljanjem obe granice površine na istu tačku, primetićete da je to nemoguće i da je vrednost uvek 0%. Ovo je posledica ranije pomenute *kontinuiranosti* funkcije gustina verovatnoća. Zbog specifične pozicije tačaka prevoja, obično se kaže da simetrični krajevi normalne distribucije imaju *sigmoidalan* oblik koji podseća na latinična slova S. Ovaj oblik ukazuje na to da je najveća gustina verovatnoća skoncentrisana u rasponu μ ± 1σ. Ukoliko taj raspon povećavamo, njime ćemo obuhvatati sve veći procenat mogućih vrednosti. Ako raspon povećamo za po jednu standardnu devijaciju, obuhvatićemo približno 95% rezultata. Preciznije, 95% rezultata koji potiču iz normalne distribucije naći će se u rasponu μ ± 1,96 · σ. Dodavanjem još jedne standardne devijacije, dolazimo do raspona koji obuhvata gotovo sve rezultate. Konkretno, verovatnoća da se neka vrednost uzeta iz normalne distribucije nađe u intervalu μ ± 2,58 · σ iznosi 99%. Kao što smo rekli, dalje povećavanje raspona praktično nema smisla, jer će se njime ostvarivati sve manji i manji priraštaj, a raspon vrednosti nikada neće dostići 100%. Ova karakteristika teorijske normalne distribucije zove se *asimptotičnost*. Asimptotičnost se na grafikonu ispoljava tako što crvena linija dodiruje x-osu tek u beskonačnosti. U praktičnom smislu, to znači da u statistici nikada ne možemo da budemo 100% sigurni u svoje zaključke. Na primer, ne bi trebalo da tvrdite da ste 100% sigurni da u populaciji ne postoji osoba koja bi ostvarila bolji rezultat na testu sposobnosti od najvišeg rezultata koji ste opazili u svom uzorku. Možete samo da kažete da je ta verovatnoća izuzetno mala, npr. manja od 5% ili 1%. Srećom, to je sasvim dovoljno za potrebe statističkog zaključivanja. Naravno, često postoji i ograničenje instrumenata koje koristimo, jer čak i da u populaciji postoji osoba koja ima IQ 250, to ne bismo mogli da izmerimo. U tom smislu, empirijske distribucije dobijene na uzorku samo su bolja ili lošija procena stanja u populaciji, ali na njih se mogu primeniti pravila koja se odnose na odgovarajuće teorijske distribucije. Ako ponovo prikažete histogram sirovih rezultata, primetićete da on nije ni simetričan, ni kontinuiran, ni asimptotičan, možda čak ni unimodalan, ali njegova sličnost sa normalnom distribucijom nam ipak daje za pravo da primenjujemo sva pomenuta pravila. Intervali M ± 1,96 · s i M ± 2,58 · s, obuhvataju približno iste procente ishoda koje u teorijskoj distribuciji obuhvataju intervali μ ± 1,96 · σ i μ ± 2,58 · σ.

Pomerajte granice površine ispod normalne krive i pratite promene u procentu obuhvaćenih rezultata. Interpretirajte granice u terminima z skorova pojedinačnih rezultata.

Koliki procenat rezultata se nalazi izvan raspona $\mu \pm 1{,}96 \cdot \sigma$ a koliki izvan raspona $\mu \pm 2{,}58 \cdot \sigma$?

Kolika je verovatnoća da iz populacije nasumično odaberete ispitanika koji na nekoj normalno distribuiranoj varijabli ima z skor manji od -1?

Kolika je verovatnoća da iz populacije nasumično odaberete ispitanika koji na nekoj normalno distribuiranoj varijabli ima z skor veći od 1,64?

Kolika je verovatnoća da iz populacije nasumično odaberete ispitanika koji na nekoj normalno distribuiranoj varijabli ima z skor između 1 i 2?

2.6.5. Standardna greška aritmetičke sredine

Do sada smo, u kontekstu verovatnoće ishoda, često ukazivali na razlike između onoga što se očekuje i onoga što se opazi u istraživanju. Teorijske verovatnoće vezivali smo za svojstva populacije, a empirijske za svojstva uzorka. Primeri sa novčićima i kockicama za igru bili su pogodni za ilustraciju osnovnih principa statističkog zaključivanja zato što su nam kod njih poznate teorijske verovatnoće ishoda. Međutim, u istraživanjima najčešće nisu poznati parametri distribucija varijabli u populaciji, što znači da su M i s koje se dobiju na uzorku, samo bolje ili lošije procene nama nepoznatih vrednosti μ i σ. Pokazali smo da tačnost te procene zavisi od dva faktora. Prvo, greška M kao procene μ obrnuto je proporcionalna veličini uzorka – što je uzorak manji, verovatnije je da M bitno odstupa od μ. Drugo, greška M kao procene μ direktno je proporcionalna varijabilnosti (varijansi) pojave – što je varijabilnost veća, veća je verovatnoća da se u uzorku dobije M koja značajno odstupa od prave vrednosti μ. To znači da greška M kao procene μ neke varijable X, može da se izrazi sledećim odnosom:

$$\frac{\sigma_X^2}{N}$$

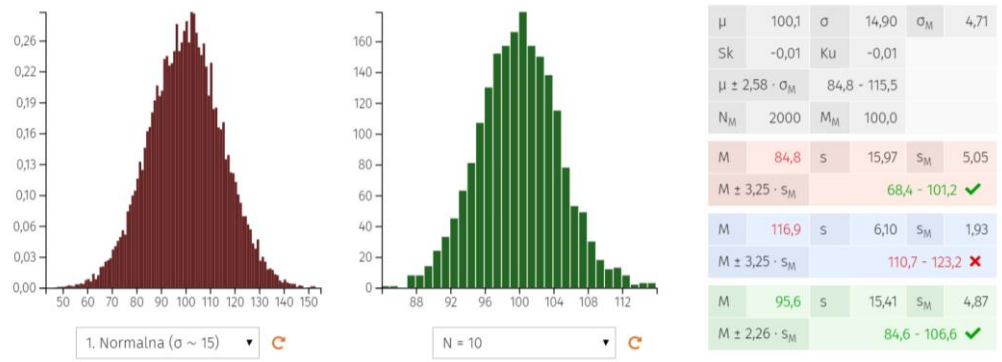
U poglavlju o centralnoj graničnoj teoremi pokazali smo da je prosek varijansi jednak varijansi proseka, te bi gornji odnos mogao da se napiše i ovako:

$$\frac{\sigma_X^2}{N} = \sigma_{\frac{X}{N}}^2 = \sigma_{\frac{\sum x}{N}}^2 = \sigma_M^2$$

To znači da je:

$$\sigma_M = \frac{\sigma_X}{\sqrt{N}}$$

Poslednja formula predstavlja izraz pomoću koga se izračunava vrednost *standardne greške aritmetičke sredine*. U nastavku teksta, detaljnije ćemo opisati logiku ovog važnog deskriptivnog pokazatelja i njegovu ulogu u donošenju zaključaka o izmerenim pojavama.



Slika 31. Distribucija aritmetičkih sredina uzoraka veličine 10 (desno) uzetih iz populacije normalno distribuiranih rezultata (levo)

U vežbi ćemo koristiti **primere simuliranih IQ skorova** (Slika 31). Za početak, odaberite prvu distribuciju sa leve padajuće liste. U pitanju su normalno distribuirani rezultati merenja, čija je aritmetička sredina oko 100, a standardna devijacija oko 15. Ove vrednosti prikazane su u kućicama sive tabele sa desne strane, a simboli μ i σ upotrebljeni su zato što prikazanu distribuciju tretiramo kao sliku ciljne populacije, npr. rezultate koje bismo dobili da smo sve studente neke države testirali testom sposobnosti. Jasno je da se iz iste, dovoljno velike populacije, može uzeti praktično neograničen broj uzoraka. Na primer, iz ciljne populacije koja ima oko 10.000 članova, može se uzeti više od 2 kvintilijarde (broj sa 33 nule) različitih kombinacija od po 10 članova. Svaki od tih uzoraka veličine 10, imao bi svoju M koja bi bila bolja ili lošija procena μ.

Istraživač, naravno, neće uzimati hiljade uzoraka iz iste populacije, ali mora da bude svestan da je uzorak koji koristi, samo jedan od velikog broja mogućih uzoraka. Samim tim, aritmetička sredina koju je dobio, samo je jedna od velikog broja (drugačijih) aritmetičkih sredina koje su mogle da budu dobijene na uzorcima iste veličine, uzetim iz iste populacije. Aritmetičke sredine izračunate na tako formiranim uzorcima možemo da prikažemo i grafički. Odaberite opciju N = 10 na desnoj padajućoj listi da biste simulirali uzimanje uzoraka veličine 10 iz prikazane populacije i iscrtali dobijene aritmetičke sredine na grafikonu. Iz populacije će biti uzeto 2.000 uzoraka. Kao što vidite, u uzorcima veličine 10 dobijaju se i vrednosti M koje se dosta razlikuju od μ. U crvenoj tabeli prikazana je aritmetička sredina nasumično formiranog uzorka koja je najudaljenija od prave aritmetičke sredine u levu stranu, a u plavoj aritmetička sredina uzorka koja je od nje najviše udaljena udesno, ka većim vrednostima. Na osnovu grafikona vidi se da je dobijanje takvih ekstremno pogrešnih vrednosti malo verovatno, ali moguće. U zelenoj tabeli prikazuju se aritmetičke sredine dobijene na svakom pedesetom nasumično formiranom uzorku. Simbol N_M u sivoj tabeli označava broj M-ova prikazanih na desnom grafikonu, odnosno broj uzoraka koji su uzeti iz populacije, a simbol M_M aritmetičku sredinu tih M-ova. Obratite pažnju na to da sa porastom broja uzoraka uzetih iz iste populacije, distribucija njihovih aritmetičkih sredina sve više podseća na normalnu krivu, a aritmetička sredina aritmetičkih sredina M_M postaje sve bliža aritmetičkoj sredini populacije μ. Ovo je još jedna od važnih i korisnih manifestacija centralne granične teoreme. Ako isti postupak ponovite na uzorcima veličine 50 (N = 50), primetićete da je distribucija velikog broja tako dobijenih M-ova takođe približno normalna, ali je njena varijabilnost manja. Ta varijabilnost nije ništa drugo do ranije pomenuta standardna greška aritmetičke sredine, koja je u sivoj tabeli prikazana simbolom σ_M. To znači da je standardna greška aritmetičke sredine zapravo standardna devijacija aritmetičkih sredina uzoraka iste veličine, uzetih iz iste populacije. U primeru sa veličinom uzoraka od 50 studenata, ona je izračunata po formuli:

$$\sigma_M = \frac{\sigma}{\sqrt{50}}$$

Sada je jasnije i zbog čega se za označavanje standardne devijacije i standardne greške aritmetičke sredine koristi isti osnovni simbol – σ. Oba pokazatelja su mere varijabilnosti, odnosno standardne devijacije, ali se u slučaju distribucija statistika, kao procena parametara populacije, obično koristi termin *standardna greška* (parametra). Vrednost σ_M uvek se izražava u istim jedinicama u kojima

se izražavaju μ i σ. U našem primeru to su IQ jedinice. Ako umesto N = 50 odaberete opciju N = 250, primetićete da se vrednost σ_M dodatno smanjuje, jer su aritmetičke sredine tako velikih uzoraka još tačnije procene aritmetičke sredine populacije. Ako bismo, na kraju, pojavu izmerili na celoj populaciji, greške aritmetičke sredine ne bi ni bilo. Kada na levoj listi odaberete drugu normalnu distribuciju, koja ima manju varijansu od prve, primetićete da na uzorcima veličine 250, σ_M postaje još manja. Vrednosti M su sada gotovo uvek odlične procene prave μ, zbog relativno male varijanse pojave u populaciji i relativno velikih uzoraka na kojima se računaju aritmetičke sredine. Ekstremno pogrešne aritmetičke sredine, koje su prikazane u crvenoj i plavoj tabeli, sada su znatno bliže vrednosti μ nego u prethodnom primeru.

Odaberite treću distribuciju sa leve liste i opciju N = 10 sa desne, kako biste simulirali uzimanje uzoraka veličine 10 iz populacije u kojoj IQ skorovi nisu distribuirani normalno, već uniformno. Varijansa distribucije u populaciji znatno je veća nego u prethodna dva primera i iznosi približno 29 jedinica. Očigledno je da ne postoji jasno grupisanje rezultata oko proseka, jer je podjednak broj osoba postigao rezultate koji su prosečni, kao i one koji su znatno iznad ili znatno ispod proseka. Međutim, bez obzira na to što varijabla nije distribuirana normalno u populaciji, aritmetičke sredine uzoraka uzetih iz te populacije ipak formiraju tipičnu zvonastu krivu prikazanu na desnom grafikonu. Što je veličina uzoraka veća, oblik distribucije njihovih aritmetičkih sredina sve više podseća na standardnu normalnu raspodelu. To znači da na distribuciju M-ova velikih uzoraka uzetih iz iste populacije, bez obzira na oblik distribucije varijable u populaciji, mogu da se primene pravilnosti vezane za površinu ispod normalne krive. Analogno rasponu $\mu \pm z \cdot \sigma$, koji obuhvata određeni procenat normalno distribuiranih podataka, zaključujemo da se isti procenat distribucije M-ova može obuhvatiti intervalom $\mu \pm z \cdot \sigma_M$. Vrednost z u ovim izrazima predstavlja bilo koju vrednost uzetu iz standardne normalne distribucije na osnovu koje se formira željeni interval. Pomenuli smo da su dve najčešće korišćene vrednosti 1,96 i 2,58. To znači, na primer, da bismo u 99% uzoraka uzetih iz populacije čija je aritmetička sredina μ, dobili M koja se nalazi u intervalu između $\mu - 2{,}58 \cdot \sigma_M$ i $\mu + 2{,}58 \cdot \sigma_M$. Ovaj interval prikazan je u trećem redu sive tabele. Odaberite veličinu uzorka N = 250 i primer uniformne distribucije. Interval $\mu \pm 2{,}58 \cdot \sigma_M$ pokazuje da se 99% aritmetičkih sredina uzoraka veličine 250 nalazi između vrednosti 95,5 i 105, iako je aritmetička sredina populacije približno 100. To znači da su u približno 1% uzoraka dobijene aritmetičke sredine koje odstupaju od μ za više od 2,58 vrednosti σ_M. Vrednosti M u crvenoj i plavoj tabeli

najverovatnije su upravo takve ekstremno pogrešne procene μ, koje se nalaze van raspona $\mu \pm 2{,}58 \cdot \sigma_M$. Na osnovu boje kojom su ispisane vrednosti M, možete da zaključite da li je zaista tako. Zelena boja označava da se M nalazi u intervalu $\mu \pm 2{,}58 \cdot \sigma_M$ a crvena da je van njega.

Sa leve padajuće liste odaberite 4. distribuciju a sa desne opciju N = 10. Distribucija M-ova uzoraka uzetih iz bimodalne distribucije takođe podseća na normalnu, a ukoliko su uzorci dovoljno veliki, na nju će moći da se primeni i pravilnost vezana za raspone $\mu \pm z \cdot \sigma_M$. Više puta smo ukazali na teorijsku važnost ovog raspona, ali kakva je njegova praktična korist u istraživanjima? Odgovor je – nikakva, jer istraživaču nisu poznate vrednosti μ i σ. Međutim, to ne znači da opisane pravilnosti ne mogu da se primene i na nivou uzorka. Naime, činjenica da svako M potiče iz teorijske distribucije M-ova, koja je normalna kada su uzorci veliki, omogućava da se uz pomoć formule:

$$s_M = \frac{s}{\sqrt{N}}$$

izračuna *standardna greška aritmetičke sredine uzorka*. U ovoj situaciji istraživač menja smer zaključivanja, pretpostavljajući da je M koje je dobio u uzorku potpuno tačno, tj. da je jednako μ. Pri tome je svestan činjenice da su u drugim uzorcima iste veličine, uzetim iz iste populacije, mogle da se dobiju drugačije vrednosti M. Ovoga puta raspon koji obuhvata određeni procenat tih vrednosti određen je intervalom $M \pm z \cdot s_M$ a ne $\mu \pm z \cdot \sigma_M$. U tako definisanom intervalu trebalo bi da se nađe i prava aritmetička sredina populacije. Verovatnoća da se to zaista desilo, određena je odabranom vrednošću z, koja se naziva *marginom greške*. Na primer, verovatnoća da se μ nalazi u intervalu $M \pm 1{,}96 \cdot s_M$ iznosi oko 95%, a da se nalazi u intervalu $M \pm 2{,}58 \cdot s_M$ oko 99%. Pošto indirektno govore o stepenu poverenja u M kao procenu μ, ovi intervali nazivaju se *intervalima poverenja* ili *intervalima pouzdanosti aritmetičke sredine*.

Većina statističara bi upravo iznetu interpretaciju intervala poverenja aritmetičke sredine smatrala netačnom. Naime, raspon vrednosti koji se dobija pomoću izraza $M \pm z \cdot s_M$ predstavlja ishod uzorkovanja, a ne nasumičnu varijablu za koju može da se veže distribucija verovatnoća. Varijablu bi činili svi mogući intervali izračunati na isti način. To znači da jedan konkretan interval može da sadrži ili da ne sadrži vrednost μ, odnosno da verovatnoća da on uključuje aritmetičku sredinu populacije iznosi 0 ili 1. Pomenutih 95% ili 99% ne odnosi se, dakle, na verovatnoću da se μ nalazi u određenom intervalu, već na verovatnoću da se u skupu svih intervala koji su formirani na isti način, dobije

onaj koji u sebi sadrži pravu aritmetičku sredinu populacije. Za potrebe ovog udžbenika ipak ćemo se složiti sa pojedinim uglednim statističarima koji smatraju da je ovakvo tradicionalističko shvatanje samo „cepanje dlake na dva dela", i da u većoj meri zbunjuje čitaoca nego što mu pomaže da shvati smisao i način primene intervala poverenja (Howell, 2012). Uz napomenu da se takav pristup može smatrati pogrešnim, intervale poverenja tretiraćemo kao raspone vrednosti u okviru kojih se, sa određenim stepenom verovatnoće, nalazi neki parametar. Pri tome treba imati na umu dve činjenice. Prvo, u statistici se uglavnom koriste dva pomenuta intervala od 95% i 99%, ali se može upotrebiti bilo koji drugi. Međutim, dodatno proširivanje intervala pouzdanosti najčešće nema smisla, jer zbog asimptotičnosti normalne distribucije, poverenje u M nikada ne može da dostigne 100%, niti verovatnoća da μ nije u definisanom intervalu može da postane 0. Tako je, na primer, verovatnoća da je μ van raspona $M \pm 1{,}96 \cdot s_M$ manja od 5%, da je van raspona $M \pm 2{,}58 \cdot s_M$ manja je od 1%, da je van raspona $M \pm 3 \cdot s_M$ manja je od 0,3%, dok je verovatnoća da μ nije u rasponu $M \pm 4 \cdot s_M$ manja od 0,007%. Vrednosti M koje se prikazuju u crvenoj, plavoj i zelenoj tabeli, ispisuju se zelenom bojom ako je μ u okvirima raspona $\mu \pm 2{,}58 \cdot \sigma_M$, a crvenom ako nije.

Druga bitna činjenica vezana za intervale poverenja aritmetičke sredine, odnosi se na vrednosti margine greške koje se koriste za njihovo izračunavanje. Ako sa desne liste odaberete manje uzorke, veličine 10 ili 50, primetićete da se za izračunavanje intervala poverenja u crvenoj, plavoj i zelenoj tabeli ne koriste vrednosti 2,58 i 1,96. Razlog je vezan za primenu zakona velikih brojeva. Naime, centralna granična teorema može da se primeni na distribucije uzoračkih aritmetičkih sredina samo ako su uzorci dovoljno veliki, nezavisno od toga kakav je oblik distribucije u populaciji. Ukoliko su uzorci mali, distribucije njihovih aritmetičkih sredina više ili manje odstupaju od normalne distribucije. Tako je, ne primer, kao margina greške za N = 10 upotrebljena vrednost 3,25 umesto 2,58, odnosno 2,26 umesto 1,96. O načinu na koji se određuju margine greške u slučaju malih uzoraka, biće više reči u narednom odeljku. Za sada je bitno da uočite da su vrednosti M i intervala poverenja M koji se prikazuju u zelenoj tabeli, najverovatnije obojeni zelenom bojom. To znači da se većina vrednosti M nalazi u rasponu $\mu \pm 2{,}58 \cdot \sigma_M$, odnosno da se na osnovu većine tih vrednosti i odgovarajuće standardne greške aritmetičke sredine, dobija raspon koji obuhvata pravu aritmetičku sredinu populacije, iako je u pitanju interval poverenja od 95%. Obratite pažnju i na to da ekstremno pogrešne vrednosti M, prikazane u crvenoj i plavoj tabeli, najčešće daju intervale poverenja koji ne

obuhvataju vrednost μ. Ipak, ove ekstremne vrednosti M dobijaju se veoma retko kada su uzorci nasumični i reprezentativni.

Za razliku od prva četiri primera distribucija u populaciji, dve poslednje raspodele sa leve padajuće liste su asimetrične, odnosno iskošene. Oblik pete distribucije nazivamo *iskošenim ulevo*, jer je njena leva strana ili *rep* (engl. *tail*) razvučenija nego desna. Takva distribucija mogla bi da se dobije ako je test sposobnosti bio relativno lak, pa je većina ispitanika ostvarila visoke rezultate. Ukoliko odaberete opciju N = 10 sa desne liste, primetićete da distribucija M-ova u ovom slučaju nema oblik normalne krive i da je takođe blago iskošena ulevo. Kao što smo rekli, razlog je u tome što se centralna granična teorema manifestuje samo u slučaju *velikog* broja aritmetičkih sredina *velikih* uzoraka iste veličine, uzetih *iz iste* populacije. Kada bismo iz populacije uzimali uzorke veličine 1, dobili bismo distribuciju koja je identična distribuciji pojave u populaciji, dok bi se sa povećavanjem veličine uzoraka njen oblik sve više približavao normalnoj. Isti slučaj je i sa šestom distribucijom koja je *iskošena udesno*. Iskoristićemo navedene primere teorijskih distribucija da bismo dodatno pojasnili logiku μ i σ kao pokazatelja kojima se opisuju raspodele rezultata merenja. Kao što vidite, prve četiri distribucije prikazane sa leve strane imaju slične aritmetičke sredine, ali različite standardne devijacije. Sa druge strane, dve poslednje (iskošene) distribucije imaju različite proseke, ali sličnu varijabilnost. Pri tome, one imaju nešto veću standardnu devijaciju od prve normalne distribucije, ali bitno manju od uniformne i bimodalne. To znači da već sama standardna devijacija pruža informacije o nekoj vrsti greške aritmetičke sredine. Dok s_M pokazuje koliko je M dobra procena μ, s pokazuje koliko je M dobra kao predstavnik svih vrednosti x. U našem primeru M ipak bolje predstavlja rezultate svih studenata kada su distribucije „samo" iskošene, nego kada bitno odstupaju od normalne, kao u primeru uniformne ili bimodalne distribucije. U prvom slučaju, očiglednije je grupisanje rezultata oko jedne vrednosti koja se nalazi relativno blizu aritmetičkoj sredini. Na kraju, uočite i to da je kod pozitivno iskošene distribucije aritmetička sredina veća od moda, dok je kod negativno iskošene upravo suprotno.

U ovom poglavlju pokušali smo da pojasnimo matematičku i logičku osnovu formule za izračunavanje s_M, ali je za čitaoca mnogo važnije da razume njenu praktičnu vrednost. Pored činjenice da nam omogućava da procenimo u kom rasponu rezultata bi mogla da se nađe μ, standardna greška aritmetičke sredine, odnosno njena formula, šalje nam nekoliko bitnih poruka. Prva je da će aritmetička sredina u uzorku biti potpuno tačna procena aritmetičke sredine u

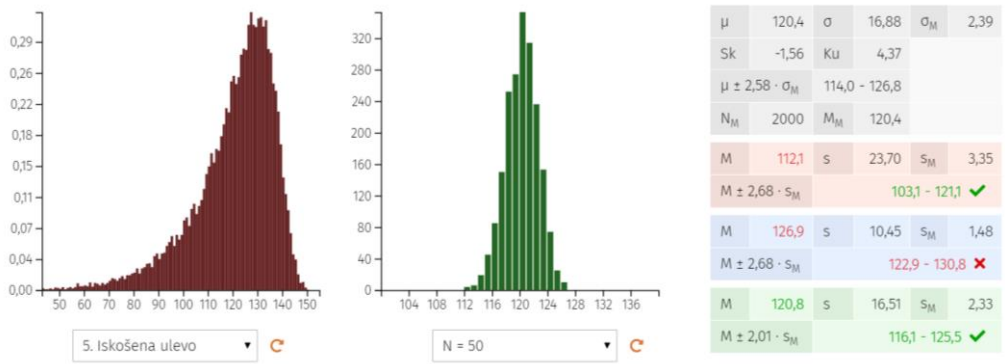
populaciji samo ako je varijabilnost merene pojave nulta ili ako je N beskonačno, tj. ako smo varijablu izmerili na celoj populaciji. Prva situacija je moguća, ali u tom slučaju reč je o varijabli koja nema nikakvu informativnu ili istraživačku vrednost. Podatak da su svi studenti na nekom testu dobili isti broj bodova potpuno je beskoristan u kontekstu statističke obrade, izuzev što nam pokazuje da test možda nije pouzdan. Druga situacija je najčešće nemoguća, obično neekonomična, a tek ponekad zaista poželjna. Dakle, istraživač mora da se pomiri sa činjenicom da M koje je izračunao na uzorku, uvek sa sobom nosi rizik greške. Srećom, centralna granična teorema omogućava nam da tu grešku kvantifikujemo.

Druga bitna poruka formule za izračunavanje s_M vezana je za koren u njenom imeniocu. Poruka glasi da nema potrebe niti smisla da se uzorak povećava u beskonačnost, jer to povećanje nije u linearnom odnosu sa smanjenjem greške. Operacije korenovanja i stepenovanja su nelinearne transformacije vrednosti, što ste mogli da vidite na primeru kvadratne jednačine u odeljku o funkcijama. U kontekstu poverenja u aritmetičku sredinu, to znači da se veći efekat postiže povećavanjem malih uzoraka, nego dodatnim povećavanjem velikih. Na primer, ako se uzorak poveća sa 5 na 105, vrednost imenioca povećaće se oko 5 puta, te će se za toliko smanjiti i greška aritmetičke sredine. Ali ako se nakon toga uzorak poveća za dodatnih 100 ispitanika, tako da njegova veličina bude 205, vrednost imenioca biće manja tek oko 1,5 puta. Drugim rečima, povećanje uzorka ima opravdanja do određenog nivoa, nakon čega vrednost s_M dostiže plato. Koji je to plato, zavisi na prvom mestu od varijabilnosti pojave koja se meri, ali o tome će biti više reči u narednom poglavlju. Ponovite prethodne vežbe na uzorcima različite veličine, za različite teorijske distribucije i posmatrajte za koliko se smanjuje standardna greška aritmetičke sredine kada se veličina uzorka poveća za 40, sa N = 10 na N = 50, a koliko kada se veličina poveća za 200 ispitanika, sa N = 50 na N = 250.

2.6.6. Skjunis i kurtozis

Varijable i njihove distribucije do sada smo opisivali koristeći dve vrste numeričkih pokazatelja. Prvu čine mere centralne tendencije, kojima se iskazuje najverovatniji ishod u nekoj distribuciji, npr. srednja, prosečna ili tipična vrednost. Ta vrednost treba što bolje da predstavlja sve podatke, odnosno da odredi tačku na x-osi oko koje se grupiše većina rezultata. Stoga se mere

grupisanja često nazivaju i *lokacijom* distribucije. Drugu grupu pokazatelja čine mere varijabilnosti, koje pokazuju koliko snažno se rezultati grupišu oko odabrane mere centralne tendencije. Međutim, pored lokacije i varijabilnosti, često je potrebno opisati i oblik distribucije. To se, naravno, najlakše postiže upotrebom grafikona, ali istraživač treba da bude upoznat i sa kvantitativnim pokazateljima koji se koriste za ove potrebe. Dva najčešća su *skjunis* i *kurtozis*. Postoji nekoliko različitih formula za njihovo izračunavanje (Panik, 2005), ali u većini statističkih programa koriste se algoritmi koji vrednosti skjunisa i kurtozisa vezuju za oblik i svojstva normalne raspodele. Naime, vrednosti ovih deskriptivnih pokazatelja uvek će iznositi 0 kada je distribucija normalna, tako da se njihova udaljenost od nulte vrednosti može upotrebiti za kvantifikaciju stepena odstupanja neke raspodele podataka od tipične zvonaste krive. Skjunis je pokazatelj iskošenosti (engl. *skewness*) ili simetričnosti distribucije. Distribucija koja ima nultu vrednost skjunisa potpuno je simetrična, što ne mora da znači i da je normalna. Na primer, **prve četiri distribucije sa leve padajuće liste** imaju skjunis blizak 0, jer su simetrične u odnosu na svoje centralne vrednosti. Sa druge strane, iskošene distribucije imaju skjunis različit od nule. Ulevo ili negativno iskošene distribucije imaju negativan skjunis, dok pozitivno iskošene distribucije imaju vrednosti skjunisa veće od 0. Skjunis distribucija prikazanih na levom grafikonu označen je simbolom *Sk* u sivoj tabeli (Slika 32).



Slika 32. Distribucija aritmetičkih sredina uzoraka veličine 50 (desno) uzetih iz populacije rezultata koji imaju negativan skjunis (levo)

Za razliku od skjunisa, logika kurtozisa je manje intuitivna, ali je njegova vrednost veoma laka za interpretaciju. Kurtozis se često definiše kao pokazatelj ispupčenosti ili spljoštenosti distribucije, ali samo prvi deo ovog objašnjenja je zapravo tačan. Ako prikažete prve dve distribucije sa leve liste, primetićete da

su njihove vrednosti kurtozisa (Ku) bliske nuli. Normalna distribucija smatra se srednje ispupčenom odnosno *mezokurtičnom*, jer njeni repovi nisu ni previše dugački, ni previše kratki u odnosu na dominantan broj središnjih vrednosti distribucije. Uniformna distribucija, međutim, ima vrednost kurtozisa manju od nule jer je spljoštena ili *platikurtična*. Tačnije, ona ni na jednom mestu nije ispupčena. Međutim, kada odaberete bimodalnu distribuciju, uočićete da vrednost kurtozisa postaje još manja. Kurtozis, dakle, treba shvatiti kao meru konkavnosti (st. gr. *κυρτός*) distribucije na mestu njene aritmetičke sredine, a u odnosu na njene krajeve, tj. repove. U tom smislu, nasuprot konkavnosti nije spljoštenost, već konveksnost, odnosno udubljenost distribucije. Iako to nije očigledno na prvi pogled, najmanju teorijsku vrednost kurtozisa (oko -2) ima Bernulijeva distribucija čija su dva ishoda jednako verovatna. U suštini, svi rezultati u ovoj distribuciji čine njene repove, jer ne postoji nijedan ishod koji je jednak proseku ili medijani. Drugim rečima, kurtozis pokazuje u kojoj meri rezultati koji su (znatno) udaljeni od aritmetičke sredine doprinose varijabilnosti pojave (Kaltenbach, 2012). Ukoliko je broj rezultata koji bitno povećavaju varijabilnost prevelik, kao kod bimodalne distribucije, kurtozis će biti negativan. Ukoliko je broj tih rezultata manji, kurtozis će biti pozitivan. To znači da kurtozis možemo da shvatimo i kao indikator postojanja aberantnih rezultata. Odaberite drugu distribuciju sa liste i obratite pažnju na to da su vrednosti 50, 100 i 150 na x-osi u stvari hiperlinkovi. Kliknite broj 100 da biste uklonili repove distribucije i zadržali samo rezultate bliske aritmetičkoj sredini. Distribucija postaje u većoj meri zaravnjena, tako da se vrednost kurtozisa smanjuje. Može se reći da sada nema dovoljno rezultata udaljenih od proseka da bi se distribucija smatrala normalnom. Ako, pak, kliknete brojeve 50 ili 150, kurtozis će se naglo povećati, jer se povećava varijabilnost zbog relativno malog broja aberantnih rezultata koji znatno odstupaju od aritmetičke sredine. Distribucija postaje ispupčena ili *leptokurtična* imajući u vidu ceo raspon x-ose. Kliknite ponovo iste brojeve da biste distribuciji vratili početni izgled.

> Zbog čega je kurtozis veći kada grupu aberantnih rezultata dodate samo sa jedne strane distribucije nego kada ih dodate sa obe?
>
> Kako na vrednost skjunisa utiče dodavanje grupe aberantnih rezultata sa leve, kako sa desne, a kako sa obe strane?

Ranije smo pokazali da značajno odstupanje distribucije od normalnosti može da obesmisli izračunate vrednosti aritmetičke sredine i standardne

devijacije kao najčešće korišćene mere grupisanja i raspršenja. Stepen i oblik odstupanja od normalnosti najlakše uočavamo na grafikonu, ali postavlja se pitanje koje vrednosti skjunisa i kurtozisa treba smatrati kritičnim u smislu bitno većih ili manjih od nule. Pošto se, kao i vrednosti drugih opisnih pokazatelja, skjunis i kurtozis računaju na uzorcima, većina statističkih paketa uz njihove vrednosti daje i vrednosti njihovih standardnih grešaka. Uz pomoć njih, a po istoj logici koju smo opisali u odeljku o standardnoj grešci aritmetičke sredine, mogu se izračunati intervali poverenja skjunisa i kurtozisa. U ovom slučaju nas takođe interesuje u kom intervalu se najverovatnije nalaze prave vrednosti pokazatelja za varijablu u populaciji, ali sada imamo i konkretnu vrednost koju tražimo, a to je nula. Ukoliko se nula nalazi unutar intervala poverenja skjunisa ili kurtozisa koji se izračunavaju kao $Sk \pm 1{,}96 \cdot s_{Sk}$, odnosno $Ku \pm 1{,}96 \cdot s_{Ku}$, zaključićemo da bi pokazatelji u populaciji vrlo verovatno mogli da budu nulti, iako smo na uzorku dobili nenulte vrednosti. Drugim rečima, iz populacije u kojoj je neka varijabla normalno distribuirana, moguće je izvući nasumični uzorak u kome ista varijabla nije normalno distribuirana. To smo, uostalom, više puta pokazali na raznim primerima uzimanja uzoraka. Naravno, sa povećanjem veličine uzorka, opada verovatnoća da će oblik distribucije znatno odstupati od slike u populaciji. Stoga pojedini autori smatraju da čak i umereno visoke apsolutne vrednosti skjunisa i kurtozisa ne utiču bitno na rezultate analize ukoliko su uzroci dovoljno veliki, npr. veći od 200 (Tabachnick & Fidell, 2014).

2.7. Još neke važne statističke distribucije

U prethodnom odeljku opisali smo karakteristike normalne distribucije i ukazali na njenu važnost u oblastima psihologije i statistike. Skrenuli smo pažnju na to da se normalna distribucija kao teorijski model može definisati matematički i tako pružiti osnov za procenu verovatnoće ishoda određivanjem pozicije nekog rezultata u odnosu na ostale vrednosti iz istog skupa podataka. Kao primer smo naveli procenu položaja prave aritmetičke sredine populacije u distribuciji velikog broja aritmetičkih sredina velikih uzoraka uzetih iz iste populacije. U statistici se koriste i druge vrste distribucija koje su podjednako važne za razumevanje pojava i donošenje zaključaka o njima. U ovom odeljku opisaćemo logiku i način nastanka tri raspodele na koje ćemo se pozivati u poglavlju o tehnikama inferencijalne statistike. Iako su u pitanju teorijske distribucije, za čitaoca nije nužno da u potpunosti razume njihovu matematičku osnovu, već prvenstveno njihovu praktičnu vrednost i način primene u oblasti statističkog zaključivanja. Ta vrednost ogleda se u činjenici da verovatnoće

ishoda koje želimo da opišemo često nisu distribuirane normalno, ali kao i u slučaju centralne granične teoreme, u tipičnim „nenormalnim" distribucijama postoje pravilnosti koje mogu da se precizno opišu i iskoriste u praksi.

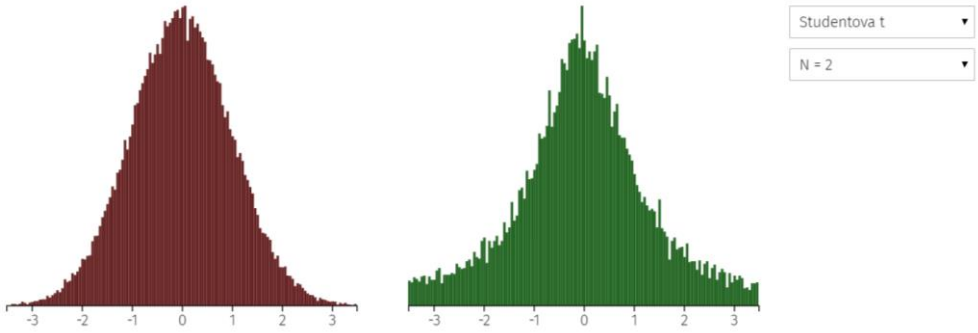
2.7.1. Studentova t distribucija

Na samom početku 20. veka engleski hemičar i matematičar Vilijam Sili Goset bio je zaposlen u pivari Ginis u Dablinu. Upoređujući količinu prinosa i kvalitet različitih sorti ječma, koristio je pravilnosti opisane u prethodnom odeljku kako bi pomoću intervala poverenja procenjivao tačnost aritmetičkih sredina izračunatih na različitim uzorcima. Nakon velikog broja eksperimenata, uočio je da te pravilnosti ne važe kada su uzorci mali. Naime, zbog činjenice da standardna devijacija koja je izračunata na malom broju merenja obično nije dobra procena prave standardne devijacije varijable u populaciji, uobičajeni intervali poverenja, bazirani na vrednostima standardne greške aritmetičke sredine uzorka, pokazali su se kao nedovoljno pouzdani. Goset je svoje otkriće predstavio u članku pod nazivom *The probable error of a mean* koji je potpisao pseudonimom *Student*, jer kompanija Ginis svojim radnicima nije dozvoljavala da objavljuju podatke koji bi mogli da ugroze njenu konkurentnost na tržištu. Goset je u navedenom članku ukazao na problem tačnosti procene vrednosti μ na osnovu M i s_M malih uzoraka, izračunao verovatnoće da se μ nalazi u određenom intervalu poverenja za uzorke različitih veličina, i matematički definisao funkciju distribucije verovatnoća odstupanja aritmetičkih sredina uzoraka od najverovatnije vrednosti aritmetičke sredine populacije (Student, 1908). Ova distribucija postala je poznata kao *Studentova*. Gosetu je u pripremi članka značajno pomogao Karl Pirson, a nakon objavljivanja vodio je intenzivnu diskusiju sa još jednim od začetnika statistike kao moderne naučne discipline, engleskim genetičarom i matematičarom Ronaldom Fišerom. Fišer je značajno doprineo promociji i unapređenju Gosetovog modela, definišući preciznije tzv. *t-odnos* (Fisher, 1925):

$$t = \frac{M - \mu}{s_M}$$

Ako se prisetite logike standardizacije sirovih podataka, uočićete da je gornji izraz veoma sličan formuli koja se koristi za izračunavanje z vrednosti. Može se reći da je t vrednost zapravo z vrednost aritmetičke sredine uzorka u distribuciji

velikog broja aritmetičkih sredina uzoraka iste veličine, uzetih iz iste populacije. Drugim rečima, t označava standardizovanu udaljenost M od μ na osnovu koje preciznije i tačnije mogu da se izračunaju intervali poverenja M kada su uzorci mali, odnosno kada vrednosti s i s_M nisu dovoljno tačne procene σ i σ_M.



Slika 33. Distribucija t-odnosa (desno) izračunatih na 50.000 uzoraka veličine 2 uzetih iz populacije normalno distribuiranih rezultata (levo)

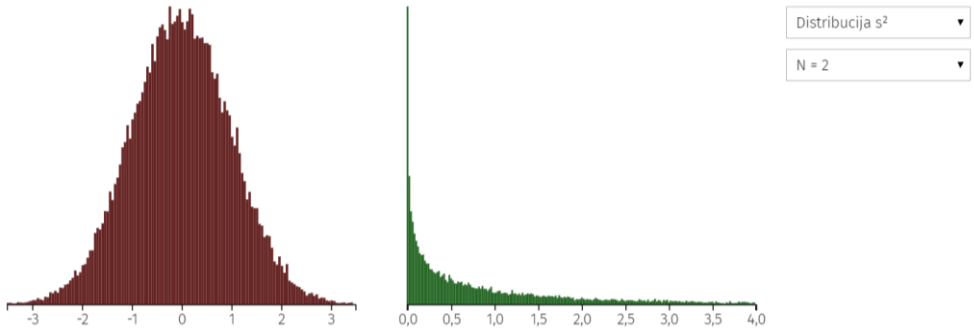
U narednom primeru poći ćemo od **normalne distribucije 50.000 merenja** prikazanih na grafikonu crvene boje. U cilju boljeg razumevanja, rezultati su standardizovani, tako da je μ = 0, a σ = 1. Zamislite da smo iz te populacije merenja uzeli 50.000 uzoraka veličine 2 i za svaki od njih izračunali M i s_M a potom vrednost t-odnosa prema ranije navedenoj formuli. Distribucija tako dobijenih vrednosti t podsećala bi na normalnu, ali bi bila nešto drugačija zbog činjenice da smo M izračunavali na osnovu samo dva merenja. Odaberite opciju *Studentova t* sa prve padajuće liste i opciju *N = 2* sa druge, da biste prikazali ovu distribuciju (Slika 33). Za razliku od normalne distribucije, t distribucija je u većoj meri platikurtična. Njeni krajevi su izduženi i „teži", što nam govori da značajan broj aritmetičkih sredina uzoraka ima t-odnos koji je po apsolutnoj vrednosti veći od 2, pa čak i od 3. To potvrđuje činjenicu da postoji velika verovatnoća da na malim uzorcima dobijemo vrednosti M koje bitno odstupaju od μ. Međutim, prikazana t distribucija govori o još jednom važnom fenomenu. Ukoliko bismo primenili ranije opisani način da izračunamo intervale poverenja aritmetičke sredine uzorka, koristeći vrednost s_M, pravilnosti vezane za površinu ispod normalne krive ne bi bile primenjive i ne bi nam dale tačnu procenu prave vrednosti μ. Na primer, interval $M \pm 1{,}96 \cdot s_M$ obuhvatio bi manji broj vrednosti M nego što bi to bio slučaj da je distribucija aritmetičkih sredina potpuno normalna. Drugim rečima, ukoliko smo M izračunali na malom uzorku, moraćemo da upotrebimo širi interval poverenja kako bismo došli do željenih

95% ili 99% sigurnosti u procenu vrednosti aritmetičke sredine u populaciji. Osnovni razlog povećanja greške procene leži u činjenici da interval poverenja M nismo računali na osnovu σ_M, već na osnovu s_M. Međutim, na većim uzorcima, ove vrednosti postaju sve sličnije, pa se tako i t distribucija menja. Ukoliko povećavate veličinu uzoraka koristeći opcije sa druge padajuće liste, primetićete da t distribucija već na uzorcima veličine 50 poprima oblik normalne krive, te da postaje skoro potpuno normalna na uzorcima od 100 merenja. Zbog toga se, između ostalog, vrednosti 50 ili 100 često pominju kao granice iznad kojih se uzorak smatra „dovoljno velikim". Međutim, čitaocu savetujemo da nikada ne određuje kriterijume „velikog" uzorka u apsolutnom smislu, već s obzirom na broj varijabli koje se koriste u istraživačkom nacrtu i stepen njihove varijabilnosti. Što je više varijabli uključeno u analizu i što je veća raspršenost rezultata na tim varijablama, to će biti potreban veći uzorak da bi se obuhvatila i opisala ukupna varijabilnost sistema.

2.7.2. Hi-kvadrat distribucija

Do sada smo se pretežno bavili opisivanjem distribucija aritmetičkih sredina uzoraka i mogućnostima procene vrednosti μ na osnovu M. U statistici je jednako važna i potreba da se opiše distribucija varijansi uzoraka u odnosu na pravu varijansu populacije. Ta distribucija ima specifičan oblik koji, kao i u slučaju t distribucije, zavisi od broja merenja na kojima se računaju vrednosti varijanse. Kada odaberete opcije ***Distribucija s^2*** **i** ***N = 1***, iz populacije sa leve strane biće uzeto 50.000 uzoraka veličine 1, a njihove varijanse biće prikazane na desnom grafikonu. Kao što vidite, histogram sadrži samo jedan stubić iznad vrednosti 0 jer je varijansa jednog rezultata merenja uvek 0. Kada uzorak povećate na 2, primetićete da varijanse veće od 0 postaju češće, ali da su i dalje vrednosti bliske 0 najverovatniji ishod koji bismo dobili na uzorku od dva merenja (Slika 34). Razlog je to što se nasumičnim izborom rezultata iz normalne distribucije najčešće dobijaju vrednosti koje su jednake ili bliske aritmetičkoj sredini. To takođe govori da će se računanjem varijanse neke pojave na malom uzorku, najverovatnije potceniti njena vrednost u populaciji. Ipak, obratite pažnju na to da je distribucija varijansi u ovom slučaju izrazito pozitivno iskošena i da smo u nekim uzorcima dobijali čak i varijanse veće od 3, iako ona u populaciji iznosi 1. Posmatrajte kako se daljim povećavanjem veličine uzorka distribucija s^2 menja iz izrazito pozitivno zakrivljene u približno normalnu. Već na uzorcima od 50 merenja aritmetička sredina distribucije

varijansi postaje bliska jedinici, odnosno vrednosti σ^2 u našem primeru. Pored toga, varijabilnost distribucije se značajno smanjuje, što pokazuje da se na velikim reprezentativnim uzorcima uglavnom dobijaju tačne procene σ^2.



Slika 34. Distribucija varijansi (desno) izračunatih na 50.000 uzoraka veličine 2 uzetih iz populacije normalno distribuiranih rezultata (levo)

Već smo naglasili da istraživač ne mora da uzima hiljade uzoraka iz iste populacije da bi ustanovio koje su vrednosti μ i σ^2. Dovoljno je da bude svestan da su M i s^2 koje je izračunao na svom uzorku, samo slučajno dobijeni elementi gotovo neograničenog skupa mogućih vrednosti koje potiču iz odgovarajućih teorijskih distribucija. U slučaju M, to je ranije opisana t distribucija koju su definisali Goset i Fišer. U slučaju s^2, to je raspodela koja veoma podseća na one koje smo videli u prethodnim primerima. Analiza funkcije gustine verovatnoće distribucije varijansi prevazilazi ambicije ovog udžbenika, ali se logika njenog nastanka može ilustrovati relativno lako. Za početak, u pitanju je očigledno raspodela kvadriranih vrednosti (s^2). Kada bismo iz standardizovane normalne raspodele prikazane na levom grafikonu, nasumično birali vrednosti, kvadrirali ih i potom sabirali, dobili bismo distribuciju koja je po obliku veoma slična distribuciji varijansi uzoraka. Nju bismo mogli da označimo sa x^2, gde je x odabrani rezultat, odnosno nasumično izvučena z vrednost. Pošto je u statistici uobičajeno da se vrednosti vezane za distribuciju merenja u uzorku označavaju latiničnim slovima (npr. M), a iste te vrednosti u teoriji, odnosno populaciji, grčkim slovima (npr. μ), u ovom slučaju se kao oznaka teorijske distribucije ne koristi latinično slovo x, već grčko slovo *χ (hi)*. Stoga se teorijska distribucija suma kvadrata z vrednosti uzetih iz normalne distribucije naziva *hi-kvadrat distribucijom*, pri čemu se vrednost χ^2 može izračunati prema formuli:

$$\chi^2(N) = \sum_{i=1}^{N} z_i^2 = \sum_{i=1}^{N} \frac{(x_i - \mu)^2}{\sigma^2}$$

gde je N broj nasumično odabranih, kadriranih i sumiranih z vrednosti iz normalne distribucije. Distribucija χ^2 vrednosti koristi se za opisivanje raspodele varijansi uzoraka, ali i drugih važnih fenomena o kojima će biti više reči u narednom poglavlju.

Kada sa liste odaberete hi-kvadrat distribuciju, primetićete da su njeni oblici za različite veličine uzoraka slični obliku distribucija s^2 u uzorcima, ali za vrednosti N – 1. Na primer, hi-kvadrat distribucija velikog broja pojedinačnih, nasumično odabranih z skorova (N = 1) ima isti oblik kao distribucija s^2 koja je dobijena na velikom broju uzoraka veličine dva. Odaberite opciju *Hi-kvadrat* χ^2 sa prve liste i *N = 2* sa druge i kliknite više puta taster *Analogna distribucija* da biste lakše uporedili analogne s^2 i χ^2 raspodele. Očigledno je da ove dve distribucije imaju identičan oblik ali drugačije raspone. Njihovi rasponi mogu čak i da se izjednače tako što se svaka s^2 pomnoži veličinom uzorka na kome je izračunata. Tada se dobija odgovarajuća χ^2 vrednost, ali za analogni uzorak z vrednosti veličine N - 1:

$$\chi^2(N-1) = N \cdot s^2$$

Treba napomenuti da gornja formula važi samo za distribuciju varijansi uzoraka uzetih iz standardizovane normalne distribucije, ali ona ipak ilustruje jednu važnu karakteristiku svih χ^2 distribucija koja doprinosi njihovoj univerzalnoj primenjivosti. Naime, kada se raspodela varijansi uzoraka opiše χ^2 distribucijom, njena aritmetička sredina biće jednaka vrednosti N, odnosno veličini uzorka na kojoj je varijansa izračunata. Pri tome treba obratiti pažnju na to da kod opcije *Distribucija* s^2 N označava broj vrednosti na kojima je računata varijansa, dok je kod opcije *Hi-kvadrat* χ^2 to broj z vrednosti koje su kvadrirane i sumirane. Na primer, hi-kvadrat distribucija za N = 50 približno je normalna i centrirana oko vrednosti 50, jer se najveći broj nasumično odabranih z vrednosti iz populacije prikazane na levom grafikonu nalazi između -1 i 1. Isti oblik imala bi distribucija varijansi za N = 51. Istraživač koji se bavi statističkom obradom podataka ne mora da poznaje matematičke osnove i način izvođenja navedenih formula, ali treba da razume njihovu važnost u opisivanju onoga što se očekuje, onoga što bi se desilo „u teoriji" i onoga što se dešava u skladu sa zakonima verovatnoće. Na primer, ako je varijansa u populaciji 1, kao u našem primeru, verovatnoća da

dobijemo $s^2 = 4$ na uzorku uzetom iz te populacije, izuzetno je mala, bez obzira na to koliko je uzorak velik. Ukoliko je uzorak dovoljno velik, ta verovatnoća je praktično nulta. Ako se u uzorku ipak dobije tolika vrednost s^2, opravdano je pretpostaviti da σ^2 zapravo nije 1. Ista logika može da se primeni i na hi-kvadrat distribuciju, ali o tome će biti više reči u narednom poglavlju.

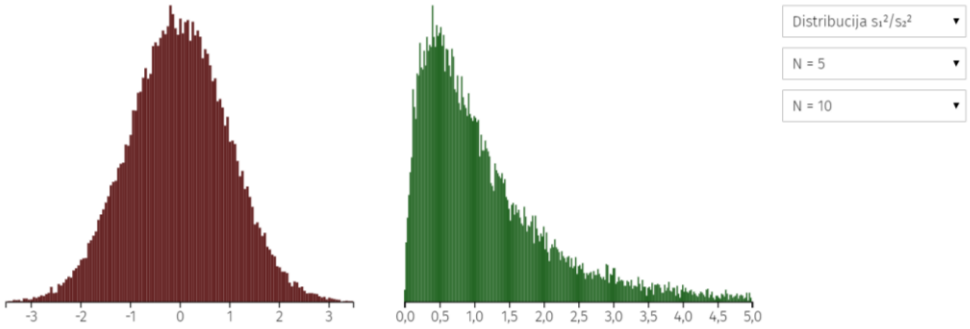
> Zbog čega je hi-kvadrat distribucija iskošena u desnu stranu?

U poslednjoj formuli za izračunavanje χ^2 vrednosti koju smo naveli, ponovo smo se susreli sa izrazom N - 1. Pre toga smo istu vrednosti upotrebili u formuli za izračunavanje standardne devijacije uzorka. Ovo je prilika da čitaocu skrenemo pažnju na razloge umanjivanja veličine uzorka u određenim formulama u statistici. Za početak, iz poslednje formule jasno je da na uzorcima veličine 1 očekujemo nultu vrednost χ^2, jer je varijabilnost jednog rezultata merenja uvek 0. Da bismo uopšte mogli da izračunamo varijabilnost neke pojave, potrebna su nam barem dva podatka. To znači da u svakom skupu od N podataka čija je varijansa s^2, samo N - 1 podataka doprinosi varijabilnosti pojave koja se meri. Ukoliko imamo uzorak veličine 2, samo jedan od njih doprinosi varijabilnosti. Za uzorak veličine 10 taj broj je 9. U uzorku veličine 100 ima 99 takvih rezultata i tako dalje. Zbog ove činjenice istraživač uvek ima manje *slobode* kada zaključke donosi na osnovu podataka u uzorku, nego što bi imao da je ispitao celu populaciju. Pokušaćemo da ilustrujemo ovu pravilnost na primeru testa znanja. Recimo da vam je cilj da proverite poznavanje glavnih gradova država u grupi učenika. Dali ste im imena pet država i tražili da na crticu pored njih napišu odgovarajuća imena glavnih gradova. Na svako od pet pitanja, odgovor može da bude bilo koja vrednost iz ogromne *populacije* naziva glavnih gradova. Ukoliko učenik odgovori tačno na svako pitanje, njegovo znanje slobodno možete da izrazite brojem 5. Sada zamislite da ste učenicima dali ponuđene odgovore i da je njihov zadatak bio da spoje naziv države sa nazivom njenog glavnog grada. U ovom slučaju formirali ste ograničeni *uzorak* naziva glavnih gradova. Upravo zbog toga znanje učenika koji je ispravno spojio sve parove, ne bi trebalo da izrazite brojem 5. Sloboda zaključivanja vam je ovoga puta ograničena, jer možete da budete sigurni samo u to da je učenik tačno znao N - 1 = 4 odgovora, a da N-ti nije morao ni da zna, jer je to poslednji par koji bi svakako bio spojen.

2.7.3. Fišer–Snedekorova F distribucija

Doprinos Ser Ronalda Fišera razvoju moderne statistike toliko je velik da bismo njegovo ime mogli da pomenemo u svakom poglavlju ovog udžbenika. U kontekstu sadržaja ovog odeljka, Fišerov stav je bio da statističke distribucije ne treba opisivati kao skupove stvarnih podataka ili rezultata merenja, već kao apstraktne fenomene definisane odgovarajućim matematičkim formulama (Salsburg, 2001). Po njegovom shvatanju, raspodele rezultata koje opažamo u istraživanjima samo su sredstvo za bolju ili lošiju procenu parametara različitih teorijskih distribucija. Jednu od njih Fišer je matematički opisao kao odnos varijansi dva uzorka uzeta iz populacije normalno distribuiranih podataka (Slika 35). Kada sa liste odaberete opciju ***Distribucija s_1^2/s_2^2***, primetićete da je potrebno da odredite dve veličine uzorka. Za vrednosti *N = 1* i *N = 1* distribuciju nije moguće prikazati jer je varijansa drugog uzorka (imenilac) nula, tako da se ne može izračunati pomenuti odnos. Ukoliko veličinu drugog uzorka povećate, svi dobijeni odnosi iznosiće 0, jer ne postoji varijabilnost u prvom uzorku. Da bi logika odnosa varijansi bila jasnija, ovoga puta nećemo postepeno povećavati veličinu uzorka, već ćemo odmah odabrati vrednost *N = 500* na obe liste. Distribucija odnosa velikog broja parova varijansi, izračunatih na velikim uzorcima, simetrična je u odnosu na vrednost 1. To znači da smo nasumičnim izborom najčešće formirali uzorke čije su varijanse identične ili veoma slične. Međutim, u određenom broju slučajeva taj odnos je manji ili veći od 1, što ukazuje na to da se iz iste populacije, potpuno slučajno, mogu uzeti uzorci drugačijih varijansi. Kao i u prethodnim primerima, prikazana distribucija pruža informacije o verovatnoćama takvih ishoda. Na primer, praktično je nemoguće da dva uzorka veličine 500, uzeta iz iste populacije, imaju varijanse koje su više od 1,5 puta veće ili manje jedna od druge. Ovo je suština *statističkog testiranja*, o čemu će biti više reči u nastavku udžbenika. Distribucija koju prikazujemo zapravo je test verovatnoće ili procena slučajnosti događaja. Na primer, ukoliko varijanse dva uzorka veličine 500 u našem istraživanju imaju odnos koji je veći od dva, zaključićemo da to verovatno nije posledica slučajnosti, već stvarne razlike među varijansama uzoraka. Drugim rečima, zaključićemo da ta dva uzorka ne potiču iz iste, već iz različitih populacija. Ukoliko su, pak, uzorci mali, npr. *N = 5* i *N = 10*, primećujemo da je distribucija odnosa varijansi izrazito pozitivno zakrivljena i da je veća verovatnoća da taj odnos bude udaljen od jedinice nego u slučaju velikih uzoraka. Tada ćemo i odnose s_1^2/s_2^2 koji su veći

od dva, smatrati mogućim ili dovoljno verovatnim, jer su mogli da se dese potpuno slučajno, čak i ako dva uzorka potiču iz iste populacije.



Slika 35. Distribucija odnosa varijansi (desno) 50.000 parova uzoraka veličine 5 i 10 uzetih iz populacije normalno distribuiranih rezultata (levo)

Većina Fišerovih radova bila je zasićena složenim formulama i opisima matematičkih osnova statističkih fenomena. To je često rezultiralo tekstovima koji nisu bili prijemčivi i dovoljno razumljivi za nematematičare. Stoga se velike zasluge za popularizaciju i ilustraciju praktičnog značaja distribucije odnosa dve varijanse, pripisuju američkom matematičaru Džordžu Snedekoru, osnivaču prve katedre za statistiku u SAD na Univerzitetu u Ajovi. Svestan da studenti sa skromnijim matematičkim znanjem neće biti oduševljeni formulama na kojima je insistirao Fišer, on je, polazeći od Fišerove poznate knjige *Statistical Methods for Research Workers*, opisao njegove ideje i otkrića na jasniji, precizniji i detaljniji način. Objavio ih je 1937. godine u jednom od najcitiranijih statističkih udžbenika pod nazivom *Statistical Methods (Applied to Experiments in Agriculture and Biology)*. U svom udžbeniku Snedekor je odnos varijansi dva uzorka nazvao *F odnosom* u čast Ronalda Fišera. Teorijska distribucija koju upravo opisujemo, od tada je poznata kao *F*, *Fišerova*, *Snedekorova* ili *Fišer–Snedekorova*. Funkcija gustine verovatnoće ove distribucije je složena i prevazilazi ambicije ovog udžbenika, ali se F vrednosti mogu jednostavno izraziti kao odnos dvaju skaliranih hi-kvadrat vrednosti kojima se, kao što smo ranije pokazali, opisuje distribucija varijansi uzoraka uzetih iz standardne normalne distribucije. Tako je:

$$F(n, m) = \frac{\chi_n^2/n}{\chi_m^2/m}$$

gde su n i m brojevi kvadriranih z skorova, nasumično uzorkovanih iz normalne distribucije, na osnovu kojih su izračunate vrednosti χ_n^2 i χ_m^2. Ukoliko sa liste distribucija odaberete *Fišer–Snedekorova F*, na desnom grafikonu biće prikazana distribucija velikog broja F odnosa, izračunatih pomoću dva skupa nasumično odabranih z vrednosti. Veličine skupova možete da postavite izborom opcija sa listi *Veličina 1. uzorka* i *Veličina 2. uzorka*. I u ovom slučaju može se uočiti da teorijska (F) distribucija ima isti oblik kao odgovarajuća empirijska distribucija (odnosa varijansi), ali za veličine uzoraka N - 1. Odaberite vrednosti *N = 2* i *N = 2* i upotrebite taster *Analogna distribucija* da biste lakše obavili poređenje. Pošto su u pitanju odnosi, raspodele će imati jednake raspone na x-osi, za razliku od primera sa hi-kvadrat distribucijom i distribucijom varijansi uzoraka.

> Zbog čega F distribucija za veličine uzoraka N = 2 i N = 100 nema isti oblik kao ona za veličine uzoraka N = 100 i N = 2?

2.8. Stepeni slobode

Stepeni slobode su veoma važan i naizgled apstraktan statistički koncept. Oni su još jedan doprinos Ronalda Fišera oblasti statističkog zaključivanja. Fišer je logiku stepeni slobode objasnio pomoću principa analitičke mehanike i geometrije višedimenzionalnog prostora (Fisher, 1922). Rečnikom fizike, broj stepeni slobode je broj nezavisnih parametara koji su potrebni da bi se opisalo stanje nekog sistema. Kao ilustraciju ove definicije upotrebićemo jednostavan primer. Zamislite situaciju u kojoj treba da definišete poziciju tri objekta koji se kreću u trodimenzionalnom prostoru: aviona, automobila i voza. Broj mogućih dimenzija na kojima ovi objekti mogu da se kreću, označićemo sa N. Pozicija aviona može da se menja levo-desno po x-osi, gore-dole po y-osi i napred-nazad po z-osi. Ovde ne uzimamo u obzir mogućnost da se avion rotira po svim osama, jer nas interesuje samo njegova pozicija u trodimenzionalnom koordinatnom sistemu. U slučaju aviona, sve navedene vrednosti mogu da variraju nezavisno, tako da nam je potrebno N podataka da bismo tačno odredili tačku u kojoj se on nalazi. Može se reći da pozicija aviona ima tri stepena slobode. Sada zamislite automobil koji vozi reli kroz pustinju. Automobil ima ograničenje u smislu kretanja po y-osi (gore-dole), tako da su nam za određivanje njegove pozicije dovoljne samo dve vrednosti, x i z, jer uvek

znamo da je njegova koordinata na y-osi vezana za najnižu tačku tla. Zato kretanje automobila ima jedan stepen slobode manje ili N - 1. Na kraju, zamislite voz koji se prugom kreće pravolinijski ka severozapadu. Koliko stepeni slobode ima kretanje voza? Tačan odgovor nije N - 1 kao u slučaju automobila, već zapravo N - 2. Naime, pored činjenice da je vrednost na y-osi uvek vezana za najnižu tačku tla, x i z koordinate nisu nezavisne ukoliko nam je poznata trasa pruge kojom se voz kreće. Za određenu tačku x, voz može da ima samo jednu poziciju na z-osi koju mu putanja pruge dozvoljava. U odnosu na početnu tačku iz koje je voz krenuo, dovoljna nam je samo informacija o njegovoj poziciji na x-osi da bismo odredili odgovarajuću vrednost na z-osi, a time i njegovu tačnu poziciju u koordinatnom sistemu. Dakle, najopštija definicija stepeni slobode nekog sistema ili neke procene, bila bi da je to broj podataka koji mogu da menjaju vrednost nezavisno jedan od drugog, a koji su dovoljni da se opiše stanje sistema ili izvrši potrebna procena. To znači da se broj stepeni slobode može izračunati tako što se od ukupnog broja elemenata koji slobodno mogu da variraju, oduzme broj ograničavajućih faktora i/ili broj elemenata čija vrednost zavisi od vrednosti nekog elementa koji je već u sistemu. Prisetite se primera sa testom znanja o glavnim gradovima država u kome je broj stepeni slobode bio jednak broju zadataka umanjenom za jedan, jer je odgovor na poslednje pitanje zavisio od odgovora na sva prethodna. Broj stepeni slobode u statistici se označava oznakom *df*, prema akronimu njihovog naziva na engleskom jeziku – *degrees of freedom*.

Geometrijska logika stepeni slobode može veoma lako da se primeni na statisticima i distribucijama koje smo opisali u dosadašnjem tekstu. Na primer, broj stepeni slobode vezan za vrednost μ je N, pošto sve vrednosti varijable u populaciji na osnovu kojih računamo aritmetičku sredinu mogu nezavisno da variraju. Isti slučaj je i sa vrednošću M kao procenom μ. Ako se nasumično odabere N - 1 vrednosti iz neke populacije, nemoguće je predvideti koja će biti vrednost N-tog rezultata. Stoga formula za računanje aritmetičke sredine glasi:

$$M = \frac{\sum x}{N}$$

Svaka od N vrednosti x u ovoj formuli može slobodno da varira i da dobije bilo koju vrednost nezavisnu od ostalih. Međutim, u slučaju računanja standardne devijacije, odnosno varijanse, ta sloboda više ne postoji. Ukoliko nam je poznata vrednost μ, varijansa populacije može da se izračuna već na osnovu jednog nasumično odabranog rezultata, jer on može potpuno slobodno da varira i da dobije vrednost koja je različita od μ. U tom slučaju, reći ćemo da postoji samo

jedan stepen slobode. Kada iz populacije uzmemo još jedan rezultat, on ni na koji način ne zavisi od prethodnog, tako da ćemo dobiti novu procenu varijanse uz koju se sada vezuju dva stepena slobode. Međutim, s obzirom na to da u istraživanjima obično ne znamo pravu vrednost μ, sve procene, uključujući i procenu varijanse populacije, vrše se uz pomoć vrednosti M. Tada je, međutim, broj stepeni slobode manji. Ukoliko u uzorku postoji samo jedan rezultat, broj stepeni slobode je nula, jer znamo da taj rezultat mora da bude jednak vrednosti M. Ukoliko je veličina uzorka dva, samo x_1 može slobodno da menja svoju vrednost, dok vrednost x_2 mora da bude takva da zajedno sa vrednošću x_1 da prosek M i sumu odstupanja Σ(x-M) nula. Stoga se za vrednost s^2 uzorka kao procenu σ^2 populacije vezuje N - 1 stepeni slobode. Ova pravilnost se ogleda u formuli:

$$s^2 = \frac{\sum(x - M)^2}{N - 1}$$

Broj stepeni slobode računa se na različite načine u slučaju različitih statističkih postupaka, ali se najčešće izražava kao veličina uzorka umanjena za broj ograničavajućih faktora, npr. za broj procena parametara populacije. U primeru sa formulom za računanje standardne devijacije, broj tih parametara je jedan (M), pa je tako i broj stepeni slobode N - 1. Ranije smo smisao stepeni slobode pokušali da ilustrujemo poređenjima teorijskih distribucija i analognih empirijskih distribucija. Broj „merenja" u slučaju teorijskih distribucija (npr. F), uvek je bio za jedan manji od odgovarajuće empirijske distribucije (npr. s_1^2/s_2^2). U primeru sa Studentovom t distribucijom, grafikon za N = 1 nije bilo moguće iscrtati, jer bi tada broj stepeni slobode bio 0, a standardnu devijaciju ne bi bilo moguće izračunati. Dakle, t distribucija za N = 2 je prva t distribucija koju je moguće formirati, jer je tada df = 1. Pojam stepeni slobode može delovati veoma komplikovano, ali za istraživača je bitno samo da razume praktične implikacije njihove primene kao oblika korekcije u većini statističkih postupaka. U odeljku o merama raspršenja pomenuli smo da korekcija u smislu primene broja stepeni slobode umesto vrednosti N, može da se shvati kao „kazna" za istraživača koji želi da donosi zaključke koristeći mali uzorak ili oslanjajući se na veliki broj procena stvarnih parametara populacije. Iz svega navedenog, mogu da se izvuku dva zaključka. Prvo, što je veći broj parametara koje procenjujemo statisticima izračunatim na uzorku, to će nam biti potreban veći uzorak da bismo dobili dovoljno pouzdane rezultate. I drugo, upotreba stepeni slobode,

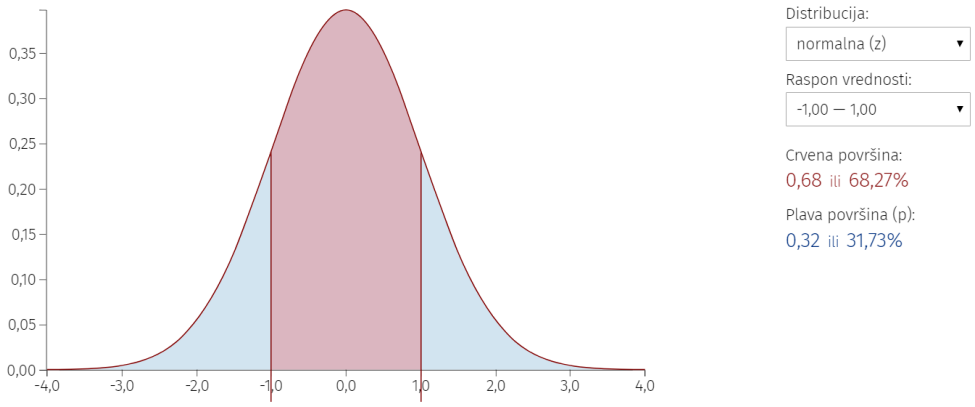
kao korekcije broja merenja u statističkim formulama, neće bitno uticati na rezultate analize ukoliko je uzorak velik.

2.9. Test-statistici, p vrednosti i nivoi značajnosti

U prethodnim odeljcima demonstrirali smo logiku nastanka t, χ^2 i F distribucija na primerima rezultata uzorkovanih iz populacije standardnih z vrednosti. Iako smo ih vizualizovali uz pomoć histograma, napomenuli smo da se ove distribucije u statistici posmatraju kao teorijske, kontinuirane raspodele koje se opisuju odgovarajućim matematičkim funkcijama gustine verovatnoće. Na osnovu oblika krivih koje te funkcije formiraju, odnosno površina koje zahvataju, moguće je procenjivati verovatnoću različitih ishoda. Na primer, moguće je izraziti verovatnoću da se na uzorcima veličine 10 i 50 slučajno dobije F odnos koji iznosi 3. Na sličan način, ali koristeći t distribuciju, moguće je utvrditi verovatnoću da uzorak veličine 10, čija je aritmetička sredina 54, a standardna devijacija 5, potiče iz populacije čija aritmetička sredina iznosi 56. Drugim rečima, opisane teorijske distribucije omogućavaju da se na osnovu podataka prikupljenih na uzorku, testira tačnost pretpostavki o pojavama u populaciji. Primer takvih pretpostavki može da bude tvrdnja da se varijanse dva uzorka ne razlikuju ili da uzorak u kome je M = 16 potiče iz populacije čija je μ = 14. Stoga se vrednosti t, χ^2 i F nazivaju *test-statisticima* (Howell, 2012), a analogne metode u kojima se koriste, poznate su kao t-test, χ^2-test i F-test. Postupak primene ovih metoda biće detaljnije opisan u narednom poglavlju, ali pre toga je potrebno objasniti dva važna statistička koncepta – *nivo verovatnoće* i *nivo značajnosti*.

Zamislite da je **visina u populaciji studenata** distribuirana normalno sa μ = 160 cm i σ = 10 cm. To znači da približno 68% studenata ima visinu u intervalu μ ± 1 · σ (crvena površina), odnosno između 150 i 170 cm. Istu pravilnost možemo da posmatramo i drugačije, ukoliko interpretiramo deo površine izvan označenog intervala μ ± 1 · σ. Zamislite da ne znamo, već samo očekujemo ili pretpostavljamo da su parametri distribucije μ = 160 cm i σ = 10 cm. Kolika je verovatnoća da iz takve populacije nasumično odaberemo osobu čija visina odstupa od μ za jednu ili više σ, tj. studenta koji je viši od 170 cm ili niži od 150 cm? Nivo verovatnoće ili *p nivo* tog ishoda prikazan je plavom površinom i iznosi oko 32% ili 0,32 (Slika 36). Naravno, oznaka *p* ne dolazi od reči *plavo*, već od latinske reči *probabilitas* i engleske *probability*. Ranije smo

pomenuli da se verovatnoće u slučaju kontinuiranih distribucija gustine verovatnoće ne odnose na jedan ishod, već na raspon ishoda. Dakle, vrednost p od 0,32 označava verovatnoću da u populaciji koja ima gore navedene očekivane parametre, postoje studenti koji od μ odstupaju za jednu ili više σ. To takođe znači da bismo, čak i da je prosek visine u populaciji 160 cm, približno jednom u tri pokušaja, nasumičnim izborom, odabirali osobe koju su visoke 150 cm i niže ili 170 cm i više.



Slika 36. Nivo verovatnoće ishoda u normalnoj distribuciji koji odstupaju za jednu ili više σ od μ (plava površina)

Kada bismo mogli da posumnjamo u tačnost početne pretpostavke o parametrima distribucije? Odaberite sa druge liste raspon -4,00 – 4,00 koji obuhvata gotovo 100% mogućih ishoda, odnosno visina studenata u populaciji. Verovatnoća da iz populacije u ovom primeru nasumično odaberemo studenta koji je visok 120 cm ($\mu - 4 \cdot \sigma$) ili manje, odnosno studenta koji je visok 200 cm ($\mu + 4 \cdot \sigma$) ili više, izuzetno je mala. Drugim rečima, ako je nasumično odabrani student visok 120 ili 200 cm, imamo razlog da posumnjamo u pretpostavljene parametre distribucije u populaciji, jer je u takvoj distribuciji ishod koji smo dobili izuzetno redak. Prilikom izvođenja ovakvog zaključka, polazimo od toga da je malo verovatno da takav rezultat dobijemo slučajno, tj. da je mnogo verovatnije da odabrani student u stvari potiče iz populacije čija μ nije 160 cm i/ili σ nije 10 cm. Tada imamo opravdanje da početnu pretpostavku smatramo netačnom, ali nikada nećemo biti potpuno sigurni da nismo pogrešili. Razlog je činjenica što, prostim slučajem, iz populacije u kojoj je varijabla distribuirana normalno, zaista možemo da dobijemo rezultat koji od μ odstupa za više od 4, 5, ili 6 σ. Mi se jednostavno vodimo logikom da su takvi slučajevi izuzetno retki

i donosimo odluku na osnovu raspoloživih empirijskih podataka. To takođe znači da kao istraživači snosimo odgovornost za moguće odbacivanje tačne pretpostavke ili nemogućnost odbacivanja netačne. Prilikom donošenja te odluke, uobičajeno ja da se koriste unapred definisane granice u odnosu na koje se procenjuje da li je neki događaj dovoljno redak da bismo posumnjali u tačnost početne pretpostavke. Te granice se nazivaju *nivoima značajnosti* i označavaju se grčkim slovom α (alfa). U primeru sa intervalom -4,00 – 4,00, nivo značajnosti smo postavili na 0,00006, jer je tolika verovatnoća, odnosno p nivo, da iz skupa normalno distribuiranih podataka izvučemo rezultat koji od proseka odstupa za 4 ili više standardnih devijacija. U statistici, naravno, ne moramo da budemo toliko strogi. Opšte prihvaćena konvencija je da se koriste dva nivoa značajnosti, blaži od 5% ili 0,05 i stroži od 1% ili 0,01. Ako promenite raspone vrednosti površine označene crvenom bojom, uočićete da su to upravo oni rasponi koji se vezuju za ranije pomenute intervale poverenja $\mu \pm 1{,}96 \cdot \sigma$ i $\mu \pm 2{,}58 \cdot \sigma$. Dakle, ishode koji se dešavaju samo pet puta u 100 merenja ili jednom u 100 merenja, u statistici smatramo dovoljno „čudnim“, „sumnjivim“, „retkim“ ili „atipičnim“, tako da imamo razlog da posumnjamo u pretpostavku od koje smo pošli. Istraživač može da se opredeli i za neki drugi nivo značajnosti, bilo da je on još blaži (npr. 10% ili 0,1) ili još stroži (npr. 0,1% ili 0,001), ali dva navedena nivoa se najčešće sreću u statističkim analizama. Poređenje p nivoa verovatnoće i α nivoa značajnosti predstavlja suštinu statističkog zaključivanja – ukoliko je p nivo dobijenog ishoda manji od postavljenog nivoa značajnosti, istraživač zaključuje da se polazna tvrdnja može smatrati netačnom. U suprotnom, smatraće je tačnom. U oba slučaja, verovatnoća greške nije nulta.

Logika statističkog zaključivanja koju smo opisali na primeru normalne distribucije može da se primeni i na ostale teorijske distribucije. Sa prve liste odaberite Studentovu t distribuciju i obratite pažnju na to da njen oblik u slučaju veoma malog broja stepeni slobode, odnosno veoma malog uzorka na kome je izračunat t-odnos, značajno odstupa od oblika Gausove krive. Interval $\mu \pm 1 \cdot \sigma$, kojim je u slučaju normalne distribucije bilo obuhvaćeno oko 68% rezultata, zahvata tek 50% površine t distribucije za df = 1. Promenite broj stepeni slobode na 3 i uočite da se sa povećanjem vrednosti df izgled krive približava obliku normalne distribucije. Sada zamislite da ste na uzorku od samo četiri osobe želeli da procenite prosečnu vrednost sistolnog ili „gornjeg“ krvnog pritiska studenata. Nakon četiri merenja ste dobili vrednosti M = 121 mmHg i s = 0,82 mmHg, pa je tako s_M = 0,41 mmHg. Vrednost df u ovom primeru iznosi 3 (N - 1), jer smo imali 4 ispitanika, odnosno merenja. Tvrdnja čiju tačnost želimo

da proverimo jeste da se prosečan krvni pritisak studenata ne razlikuje od vrednosti 120 mmHg koja se obično smatra normalnom. Odlučili smo se za α = 0,05 kao blaži nivo značajnosti. Da bismo proverili tačnost postavljenje tvrdnje, upotrebićemo formulu iz poglavlja o t distribuciji:

$$t = \frac{M - \mu}{s_M} = \frac{121 - 120}{0{,}41} \approx 2{,}44$$

Dakle, t vrednost aritmetičke sredine dobijene na uzorku od četiri studenta, uz pretpostavku da aritmetička sredina u populaciji iznosi 120, iznosi 2,44. Sa druge padajuće liste odaberite raspon -2,44–2,44, a sa treće 3 stepena slobode. Prikazani p nivo koji odgovara dobijenoj t vrednosti, odnosno dobijenoj M, iznosi 0,09. Pošto je veći od nivoa značajnosti 0,05, početnu tvrdnju možemo da smatramo tačnom. Naime, iako se vrednosti 120 i 121 razlikuju u apsolutnom smislu, one se ne razlikuju *statistički značajno*. Pri tome treba imati na umu da početna tvrdnja nije sadržala nikakvu pretpostavku o smeru razlike, tako da je na opisani način zapravo obavljeno *dvostrano* testiranje (engl. *two-tailed*). Vrednost 0,09 označava verovatnoću da se na uzorku veličine 4 slučajno dobije M koje je od μ udaljeno za 2,44 s_M u bilo koju stranu. Drugim rečima, čak i da μ iznosi 120 mmHg, postoji 9% verovatnoće da se potpuno slučajno dobije vrednost M koja je nalazi izvan raspona 120 ± 2,44 · 0,41, tj. vrednost koja je manja ili jednaka 119, odnosno veća ili jednaka 121. Upravo to se desilo u našem zamišljenom istraživanju, te zaključujemo da ishod 121 nije dovoljno redak da bismo ga smatrali statistički značajno različitim ili udaljenim od 120. Obratite pažnju na to da bi za M = 119 apsolutna vrednost t bila ista, ali bi t-odnos imao negativan predznak. Da li bi naš zaključak bio drugačiji da smo istu t vrednost dobili na nešto većem uzorku? Odaberite 9 kao broj stepeni slobode i proverite p nivo. Verovatnoća da se u uzorku veličine 10 dobije M koja odstupa od μ za 2,44 ili više s_M manja je nego u prethodnom primeru, tako da bismo vrednost 121 ovoga puta smatrali statistički značajno različitom od 120, ali samo *na nivou 0,05*. Dobijena razlika i dalje nije statistički značajna *na nivou 0,01*, jer je 0,04 veće od strožeg nivoa značajnosti koji obuhvata 1% najređih ishoda. Razlika ne bi bila značajna na tom strožem nivou čak ni na uzorku od približno 100 studenata, jer bi tada p nivo iznosio približno 0,02. Na osnovu opisanih primera možemo da zaključimo da statistički postupci bukvalno sprečavaju istraživače da na suviše malim uzorcima izvode zaključke koji bi doveli do proglašavanja netačnih tvrdnji tačnim.

Pomeranjem graničnih linija crvene površine utvrdite koja je minimalna t vrednost potrebna da bismo neku razliku smatrali značajnom na nivou 0,01 za 100 stepeni slobode.

Logika statističkog zaključivanja primenom χ^2 i F distribucija donekle se razlikuje u odnosu na prethodno opisanu primenu t distribucije. Za početak, ove distribucije ograničene su sa leve strane, jer najmanja moguća vrednost χ^2 i F iznosi nula. Osim toga, one su asimetrične i izrazito pozitivno zakrivljene kada je broj stepeni slobode mali. Specifičnost χ^2 distribucije je i u tome što, za razliku od t distribucije, sa porastom broja stepeni slobode, raste i minimalna, odnosno granična vrednost od koje izračunata χ^2 vrednost treba da bude veća da bismo fenomen koji smo uočili smatrali statistički značajnim. Odaberite hi-kvadrat distribuciju sa prve padajuće liste. Plava površina na prikazanom grafikonu prikazuje verovatnoću da vrednost χ^2 bude jednaka ili veća od 3,84 za jedan stepen slobode. Teorijski, to znači da će 95% z vrednosti koje su nasumično odabrane iz normalne distribucije, kvadriranjem dati rezultat iz intervala od 0 do 3,84. Ako ovu pravilnost povežemo sa površinom ispod normalne krive, moglo bi se reći da je χ^2 distribucija zapravo kvadrirana normalna distribucija koja je „presavijena" po vertikalnoj osi. Verovatnoća da iz normalne distribucije nasumično odaberemo z vrednost veću od 1,96 ili manju od -1,96 iznosi 0,05, što je jednako verovatnoći da dobijemo kvadrat z vrednosti koji iznosi $\pm 1{,}96^2$ ili 3,84. Sa povećanjem broja uzorkovanih z vrednosti, odnosno broja stepeni slobode, očekivano se dobijaju veće sume izražene χ^2 vrednostima. Na primer, sume devet nasumično odabiranih z vrednosti (df = 9) kretaće se u intervalu od 0 do 3,84 tek u nešto manje od 8% slučajeva. Praktična primena χ^2 distribucije je višestruka i biće objašnjena u narednom poglavlju. Na ovom mestu ćemo kratko ilustrovati samo jednu od njih. Odaberite raspon 0,00–11,34 i broj stepeni slobode 3. Prikazani grafikon omogućava testiranje pretpostavke da 3 nasumično odabrane z vrednosti potiču iz normalne populacije. Naime, ukoliko kao zbir tih vrednosti dobijemo sumu koja je veća od 11,34, možemo da, sa verovatnoćom greške manjom od 1%, odnosno sigurnošću većom od 99%, tvrdimo da one ne potiču iz normalne distribucije, jer bi se tolike sume slučajno dobijale samo u 1% uzoraka. Dakle, pomoću samo tri merenja neke pojave, može se pretpostaviti da li je ona u populaciji normalno distribuirana.

Podrazumeva se da istraživač neće na ovako banalan način testirati normalnost raspodele podataka, ali to je ipak pogodna ilustracija logike statističkog zaključivanja.

> Od koje vrednosti treba da bude veći dobijeni χ^2 za dva stepena slobode da bismo tvrdnju da odabrane z vrednosti ne potiču iz normalne distribucije smatrali tačnom, uz verovatnoću greške manju od 5%, odnosno na nivou značajnosti 0,05?

2.9.1. Jednostrano testiranje razlika

Iz gore navedenih primera može se zaključiti da prilikom interpretacije statističke značajnosti χ^2 vrednosti posmatramo samo desni kraj χ^2 distribucije za odgovarajući broj stepeni slobode. Pri tome se ipak primenjuju pomenute statističke konvencije i α nivoi od 0,01 ili 0,05. Ali u ovom slučaju, tih 1% ili 5% retkih vrednosti definiše se kao odsečak desnog repa distribucije. Stoga se ova vrsta testiranja naziva *jednostranim* (engl. *one-tailed*). Za razliku od toga, u slučaju simetrične t distribucije, koristili smo dvostrano testiranje i pomenute nivoe delili na dva dela, tako da zbir plavih površina levog i desnog odsečka daje vrednost 0,01 ili 0,05. Konkretno, proveravali smo tačnost tvrdnje da se dobijena M statistički značajno razlikuje od μ, tj. verovatnoću da je statistički značajno manja ili statistički značajno veća od nje. Međutim, istraživač ima mogućnost i slobodu da upotrebi postupak jednostranog testiranja, čak i u slučaju t distribucije. Odaberite ponovo **Studentovu t distribuciju sa prve liste**, raspon -1,76–1,76 sa druge i 100 kao broj stepeni slobode. Pretpostavka čiju tačnost želimo da proverimo je da se prosečno vreme za koje studenti Fakulteta sporta i fizičkog vaspitanja pretrče stazu od 100 m, a koje iznosi 12 sekundi, ne razlikuje statistički značajno od zadatog kriterijuma od 13 sekundi. Ukoliko, na primer, dobijena t vrednost iznosi -1,76, zaključak bi bio da su studenti brži, ali ne i statistički značajno brži od kriterijuma, jer bismo razliku čija je apsolutna t vrednost 1,76 ili veća, potpuno slučajno mogli da dobijemo u 8% uzoraka veličine od oko 100. Drugim rečima, čak i da je prosečno vreme u populaciji svih studenata fakulteta sporta 13 sekundi, u 8 od 100 uzoraka uzetih iz te populacije dobijali bismo vrednosti aritmetičke sredine koje iznose 12 sekundi ili manje, odnosno 14 sekundi ili više. Sada pomerite desnu graničnu liniju do samog desnog kraja distribucije. Preostala plava površina postaje

duplo manja, a odgovarajući p nivo pruža nam informaciju o tačnosti drugačije formulisane tvrdnje. Naime, ukoliko smo pre početka istraživanja opravdano očekivali da su studenti brži od datog kriterijuma, polazna tvrdnja nije morala da sadrži pitanje da li je M *značajno različito* od 13, već da li je *značajno manje* od 13. Tada nas ne bi ni interesovao desni kraj distribucije, tačnije onih 2,5% verovatnoće da je M - μ > 0, jer isključujemo mogućnost da su studenti sporiji od kriterijuma. Kritičnih 1% ili 5% aritmetičkih sredina tražimo samo u levom repu distribucije, tako da će naš zaključak biti da je M koje iznosi 12 sekundi, statistički značajno *manje* od 13 sekundi. Polaznu tvrdnju smo, dakle, testirali jednostrano. Naravno, istraživač ne treba da donosi odluku o tome koju vrstu testiranja će primeniti nakon što mu se rezultati koje je dobio ne dopadnu, već isključivo ako pre sprovedene analize pokaže da postoje jasne indicije da razlika može da se manifestuje samo u jednom smeru. Verovatnoća da neku razliku proglasimo statistički značajnom uvek je duplo veća ukoliko primenjujemo jednostrano testiranje umesto dvostranog.

Specifičnost F distribucije je u tome što njen oblik određuju dve vrednosti stepeni slobode, jer opisuje odnos varijansi parova uzoraka uzetih iz normalne distribucije. Testiranje tačnosti pretpostavki primenom F distribucije može da bude jednostrano i dvostrano, ali ćemo se u ovom udžbeniku kratko zadržati samo na prvom pristupu. Za početak, sa prve padajuće liste odaberite Fišer–Snedekorovu F distribuciju. Raspon prikazane distribucije pokazuje da su procene σ na malim uzorcima veoma neprecizne i da vrednosti s^2 mnogo više variraju nego u slučaju velikih uzoraka. Grafikon pokazuje da bismo u približno 50% parova nasumično formiranih uzoraka dobili veću varijansu u imeniocu, tj. vrednosti F odnosa između 0 i 1, a u preostalih 50% veću varijansu u brojiocu i F odnos veći od 1. Kriva je takođe izrazito pozitivno zakrivljena, a razlog je vezan za karakteristike raspona racionalnih brojeva, tj. onih brojeva koji se izražavaju kao razlomci. Zamislite da smo u šest merenja dobijali sledeće odnose varijansi: 1/1, 2/1, 3/1, 4/1, 5/1 i 6/1. Vrednosti F odnosa bi se tada kretale u rasponu od 1 do 6. Međutim, u slučaju da smo dobili odnose 1/1, 1/2, 1/3, 1/4, 1/5 i 1/6, taj raspon bio bi mnogo uži i kretao bi se između 0,17 i 1. Dakle, ista proporcija ishoda koncentrisana je u mnogo manjem rasponu brojeva levo od jedinice, te je stoga desni kraj F distribucije u prikazanom primeru veoma razvučen. Ukoliko povećavate broj stepeni slobode varijanse u imeniocu (df_2) na 3, 9, a potom i na 100, primetićete da se oblik krive ne menja bitno, ali da se verovatnoća dobijanja F vrednosti većih od 1 smanjuje. Ako se prisetimo opisa nastanka χ^2 distribucije, razlog ove promene možemo pronaći u činjenici da na malim

uzorcima uzetim iz standardizovane normalne distribucije postoji veća verovatnoća da se potceni vrednost σ^2. U primeru sa $df_1 = 1$ i $df_2 = 100$, veća je verovatnoća da je s^2 u brojiocu manja od s^2 u imeniocu, a ova druga je sigurno tačnija procena prave σ^2. Sa povećavanjem vrednosti df_1, verovatnoća ishoda manjih i većih od 1 polako se izjednačava, raspon F distribucije se smanjuje i ona postaje sve više simetrična, jer je mala verovatnoća da se na velikim uzorcima dobijaju s^2 koje bitno odstupaju od σ^2. Na osnovu oblika krive može da se zaključi da je skoro nemoguće da varijansa jednog relativno velikog uzorka uzetog iz normalne distribucije, bude duplo veća od varijanse drugog uzorka iste veličine uzetog iz iste populacije.

Odredite graničnu vrednost, od koje treba da bude veći F odnos, da biste varijansu za koju je df = 3 smatrali statistički značajno većom od varijanse za koju je df = 100 na nivou 0,01.

Odaberite raspon 0,00–3,92, $df_1 = 3$ i $df_2 = 9$. Kakav zaključak o varijansama biste doneli na osnovu površine koju formira taj raspon? Da li isti zaključak važi i u slučaju kada je $df_1 = 9$ a $df_2 = 3$? Zbog čega važi, odnosno ne važi?

VIZUALIZACIJA RAZLIKA I POVEZANOSTI IZMEĐU VARIJABLI

U prethodnom poglavlju bavili smo se pojmom verovatnoće i važnim teorijskim distribucijama u statistici. Opisali smo osnovne grafičke i numeričke metode koje se koriste u oblasti deskriptivne statistike i koje bi trebalo da budu osnova i priprema za svaku analizu podataka, bez obzira na njen konačni cilj ili složenost. Adekvatno opisivanje varijabli predstavlja uslov bez koga rezultati statističke obrade podataka mogu da postanu potpuno neupotrebljivi i navedu istraživača na pogrešne zaključke, npr. zbog značajne iskošenosti distribucija, postojanja aberantnih rezultata ili primene metoda koje nisu prikladne za nivo na kome su varijable izmerene. Na kraju poglavlja, dotakli smo se i osnovnih principa inferencijalne statistike, kao skupa tehnika pomoću kojih se donose zaključci, odnosno sudovi o svojstvima populacije, na osnovu karakteristika uočenih na reprezentativnom uzorku merenja. Primer statističke inferencije bila bi procena vrednosti parametra populacije, npr. aritmetičke sredine, na osnovu statistika izračunatih na uzorku uzetom iz te populacije. Tehnike inferencijalne statistike omogućavaju nam da odemo dalje od prostog opisivanja pojava i da utvrdimo u kojoj meri se pravilnosti uočene na jednom ili više uzoraka, mogu uopštiti i pripisati celoj populaciji. U ovom poglavlju detaljnije ćemo opisati nekoliko osnovnih, ali često korišćenih inferencijalnih testova pomoću kojih se određuje statistička značajnost razlika između grupa merenja, odnosno stepen povezanosti (korelacije) između varijabli.

3.1. Testiranje (ne)tačnosti nul-hipoteza

Tipičan primer statističkog zaključivanja predstavlja donošenje odluke o tačnosti pretpostavke, postavljene na početku istraživanja ili nakon preliminarne deskriptivne analize podataka. Ova pretpostavka naziva se *nulta* ili *nul-hipoteza* a označava se simbolom H_0. Njen naziv sugeriše da istraživač obično negira postojanje nekog fenomena, npr. očekuje da se studijom koju sprovodi neće utvrditi značajne razlike ili povezanosti između varijabli. U tom smislu, podaci se prikupljaju i obrađuju sa ciljem da pruže dovoljno empirijskih argumenata na osnovu kojih će se nulta hipoteza, sa određenim stepenom

sigurnosti, odbaciti. Ukoliko takvi argumenti ne postoje, istraživačka pretpostavka ne može da se smatra pogrešnom. Ovakvu logiku testiranja pretpostavki opisali smo na primerima upotrebe t, χ^2 i F vrednosti kao test-statistika. Istraživačka hipoteza mogla bi da se formuliše na sledeći način:

H_0: Prosečan krvni pritisak studenata psihologije
ne razlikuje se statistički značajno od 120 mmHg

U ovom primeru, polazna pretpostavka je da su dve vrednosti jednake, tj. da je razlika među njima nulta (engl. *null*), pa se nul-hipoteza može predstaviti i izrazom:

$$H_0: \mu = M \text{ ili } H_0: M - \mu = 0$$

U prethodnom poglavlju pokazali smo da dve vrednosti koje su različite u apsolutnom smislu, ne moraju nužno da budu drugačije i u smislu statističke značajnosti. Vrednosti 120 i 121 smatraćemo probabilistički jednakim ukoliko pokažemo da vrlo verovatno potiču iz iste populacije, tj. da p vrednost ukazuje na to da se M = 121 mogla potpuno slučajno i verovatno dobiti na uzorku uzetom iz populacije čija je μ = 120. Tada se početna nul-hipoteza ne odbacuje, a razlike se pripisuju slučajnosti, tj. posledici očekivanih varijacija uzoraka ili nesavršenosti postupka merenja.

Objašnjavajući logiku pojma „nulta hipoteza", Ronald Fišer je upotrebio analogiju sa eksperimentima u oblasti fizike, slično kao i u slučaju pojma stepeni slobode (Bennet, 1990). Poziciju iz koje polazi istraživač, opisao je kao nulto stanje galvanometra, uređaja kojim se meri jačina slabih struja. Početno stanje ne označava nužno nultu razliku, već bilo koju početnu poziciju sistema, tj. pretpostavku od koje se kreće u istraživanje. Kazaljka galvanometra, baš kao i razlika koju istraživač uoči u svom istraživanju, može da se kreće ulevo ili udesno, u zavisnosti od smera struje. Ukoliko struja nije dovoljno jaka, kazaljka se neće značajno pomeriti u odnosu na svoje početno „nulto" stanje. Analogno tome, ukoliko je nulta hipoteza tačna, vrlo je mala verovatnoća da će se u nekom eksperimentu uočiti razlika ili povezanost koja je značajna i koja bitno odstupa od nultog stanja. O toj verovatnoći govori p nivo koji je, prema Fišeru, osnovni rezultat istraživanja posle koga nije ni potrebno donositi odluku o tačnosti hipoteze. Fišer se fokusira na jednu istraživačku nul-hipotezu za koju je, po njemu, lakše dokazati da je netačna nego da je tačna (Howell, 2012). Na primer, ukoliko 1.000 nasumično odabranih muškaraca ima brkove, tvrdnja da

svi muškarci na svetu imaju brkove i dalje ne može da se smatra 100% tačnom. Sa druge strane, dovoljno je da jedan od tih 1.000 muškaraca nema brkove da bismo tvrdnju smatrali 100% netačnom.

Za razliku od Ronalda Fišera, poljski matematičar Jirži Nejman i engleski matematičar Egon Pirson, sin pomenutog Karla Pirsona, smatrali su da je za istraživača važnije da definiše α nivo značajnosti i da odluči da li hipotezu odbacuje ili ne. Pored toga, Nejman–Pirsonov pristup podrazumeva definisanje najmanje jedne *alternativne* hipoteze koja će se proglasiti tačnom ukoliko nulta hipoteza nije tačna (Neyman & Pearson, 1928). Ovaj pristup je na neki način pragmatičniji, jer stavlja naglasak na činjenicu da istraživač uvek pravi manju ili veću grešku u zaključivanju. Prema Nejmanu i Pirsonu, postoje dva tipa greške: I i II (Neyman & Pearson, 1933). Greška tipa I odnosi se na situaciju kada istraživač tačnu nultu hipotezu proglasi netačnom, tj. kada na uzorku utvrdi postojanje nekog fenomena koji u populaciji zapravo ne postoji. Stoga se ona naziva i greškom „lažnog alarma" ili „lažnog otkrića". Verovatnoća ove greške označava se simbolom α i predstavlja ranije pomenuti nivo značajnosti. Greška tipa II odnosi se na situaciju u kojoj istraživač netačnu nul-hipotezu proglašava tačnom i ne uspeva da uoči postojanje nekog fenomena. Stoga je ona poznata kao greška „propuštene detekcije" ili „propuštene šanse". Verovatnoća greške tipa II označava se simbolom β. U Tabeli 1 dat je pregled četiri moguća ishoda koji nastaju kombinacijom (stvarnog) statusa hipoteze u populaciji i odluke koju istraživač donosi na osnovu vrednosti statistika koje je izračunao na uzorku.

Tabela 1. Četiri kombinacije stvarnog statusa nulte hipoteze u populaciji i odluke donete na osnovu statistika izračunatog na uzorku

Odluka o H_0	Stvarni status H_0 u populaciji	
	Tačna	Netačna
Odbacuje se	Greška tipa I (α)	Opravdano odbacivanje (1 - β)
Ne odbacuje se	Opravdano neodbacivanje (1 - α)	Greška tipa II (β)

Iako se koncept testiranja nulte hipoteze u modernoj statistici često smatra zastarelim, prevaziđenim, pa čak i štetnim (Nickerson, 2000), veliki broj istraživača još uvek se drži postavki „nul rituala" (Gigerenzer, 2004) koje su, u najvećoj meri, definisali Nejman i Pirson. Zato je veoma važno ukazati na pogrešne interpretacije i shvatanja ovog postupka koje su česte, ne samo među studentima, već i među nastavnicima i istraživačima (Haller & Krauss, 2002). Na ovom mestu ukratko ćemo razjasniti najčešće zablude i greške vezane za tumačenje statusa nulte hipoteze na osnovu nivoa značajnosti i empirijski utvrđenih nivoa verovatnoće (Nickerson, 2000).

1. Nivo α i nivo p su dve različite vrste verovatnoće. Prva je *apriorna*, tj. prethodna ili nezavisna od iskustva. Ona se određuje kriterijumom koji istraživač postavlja pre istraživanja. Druga je *posteriorna* jer govori o verovatnoćama ishoda koje su opažene na uzorku, tj. nakon istraživanja i pod pretpostavkom da je nulta hipoteza tačna. Poređenjem ove dve verovatnoće, istraživač donosi odluku o statusu nulte hipoteze. To znači da istraživač utiče na verovatnoću pojave grešaka oba tipa i procenjuje koja od njih je manje štetna u datom kontekstu. Na primer, kada se testira efikasnost novog leka, biće poželjno da se greška tipa I smanji tako što će se odabrati nivo značajnosti od 0,01 ili čak 0,001 ukoliko primena tog leka, pored izlečenja, može da ima i negativne posledice po pacijenta. Sa druge strane, nekada ćemo biti tolerantniji u donošenju odluke, te ćemo kao značajne prihvatiti i razlike čiji je p nivo veći od 0,05, kako bismo ukazali na potrebu za dodatnim istraživanjima u toj oblasti. Međutim, tada prihvatamo rizik koji nosi veća verovatnoća greške II.
2. U vezi sa prethodnim, p nivo nije mera verovatnoće da je nulta hipoteza tačna, niti je 1 - p verovatnoća da je tačna alternativna hipoteza. Nivo p je samo verovatnoća da statistik koji smo izračunali na uzorku, slučajno dobije vrednost koja je toliko ili više udaljena od pretpostavljenog parametra u populaciji, tj. od vrednosti koja se očekuje na osnovu nulte hipoteze. Ukoliko vrednost statistika nije dovoljno retka i „čudna", tj. statistički značajno udaljena od neke referentne vrednosti, istraživač nema opravdanje da odbaci nultu hipotezu. Nivo p je samo uslovna verovatnoća pojave greške tipa α, ako je nulta hipoteza tačna. To nije isto što i verovatnoća da je hipoteza tačna ako smo dobili određenu vrednost statistika.
3. Statistička značajnost ne podrazumeva nužno teorijsku ili praktičnu značajnost. Postupak testiranja nulte hipoteze ne sme se primenjivati

mehanički, već isključivo u kontekstu logičnih objašnjenja dobijenih rezultata i njihove potencijalne primene. Ukoliko vrednost p nekog test-statistika iznosi 0,049, onda je u izveštaju o istraživanju korektnije i tačnije napisati p = 0,049, nego prosto konstatovati da je p manje od odabranog nivoa 0,05. U drugom slučaju, čitalac će biti uskraćen za informaciju da je u pitanju potencijalno značajan nivo verovatnoće. Osim toga, statistička značajnost može da bude posledica slučajnosti ili koincidencije događaja, kao što je npr. **povezanost između visine ulaganja u nauku i broja samoubistava** u SAD. Dakle, statistički značajna povezanost može da bude praktično potpuno beznačajna. O ovome će biti više reči u narednim odeljcima.

4. Određenje hipoteze kao tačne ili netačne može da se donese samo ako su poznati parametri populacije. Ukoliko se status hipoteze određuje na osnovu uzorka, mogući ishodi su njeno odbacivanje na određenom nivou značajnosti i nemogućnost njenog odbacivanja. Ovaj drugi ishod ne podrazumeva da je nulta hipoteza tačna, jer velika vrednost p ne mora da bude posledica nepostojanja nekog fenomena u populaciji. Uzroci mogu da budu nedovoljno velik uzorak, greška u merenju, postojanje aberantnog rezultata ili prosto odluka istraživača da prihvati stroži nivo značajnosti. Sama činjenica da vrednost 0,049 sugeriše jedan, a vrednost 0,051 potpuno drugi zaključak, pokazuje da se tačnost hipoteze ne sme strogo vezivati za dobijenu vrednost p.

5. U vezi sa prethodnom tačkom, još jednom ćemo se vratiti na problem ponovljivosti u naučnim istraživanjima. Stalni pritisak na istraživače da objavljuju originalne i uzbudljive rezultate doveo je do hiperprodukcije novih i „revolucionarnih" otkrića koja nisu zasnovana na raznovrsnom i bogatom pređašnjem iskustvu i koja često ne mogu da budu potvrđena u naknadnim studijama. Istraživači neretko izbegavaju da objave rezultate koji nisu u skladu sa očekivanjima ili nisu statistički značajni, fokusirajući se na efekte jakog intenziteta, npr. samo na merenja u kojima je došlo do značajnog poboljšanja zdravstvenog stanja pacijenata nakon primene nekog leka. Na taj način se prenaglašava statistički značaj određenih fenomena što dovodi do tzv. *inflacije istina* ili *greške tipa M* (engl. *magnitude*) (Reinhart, 2015). Može se reći da je insistiranje na novim, kontinuiranim, brzim i značajnim otkrićima u potpunoj suprotnosti sa ključnom idejom *opovrgljivosti* naučnih teorija poznatog austrijsko-britanskog filozofa *Karla Popera* (Popper, 2002).

Prema Poperu, najjača potvrda neke teorije u nauci je neuspešnost istraživača da je udruženim naporima i ponovljenim eksperimentima opovrgnu. Dakle, treba imati na umu da je dobijeni p nivo samo jedno od mogućih p, kao što je i aritmetička sredine uzorka samo jedna od mogućih procena aritmetičke sredine populacije. Zato p nivo, sam po sebi, ne može da bude potvrda neke hipoteze, pogotovo ako ne postoje rezultati sličnih istraživanja sa kojima bi mogao da se uporedi i sa kojima bi mogao da čini solidniji i dosledniji korpus znanja. Pored toga, mladi istraživači treba da budu svesni da velika vrednost p i nemogućnost da se hipoteza odbaci ne znače da njihovo istraživanje nije uspelo ili da je manje vredno.

6. Još jedan argument u prilog stava da p nivo nije pokazatelj tačnosti nulte hipoteze, jeste činjenica da se ista istraživačka pretpostavka može testirati različitim test-statisticima, tj. različitim statističkim postupcima. Štaviše, primena drugačijih metoda na istim podacima može da dovede i do drugačijih zaključaka. Jedan od razloga je različita *snaga statističkih testova* (Cohen, 1988) koja se izražava kao vrednost 1 - β. Snaga testa predstavlja uslovnu verovatnoću da se nul-hipoteza odbaci kada je netačna, a u Tabeli 1 označena je kao *opravdano odbacivanje*. Drugim rečima, kada primenjujemo snažnije statističke testove, verovatnije je da ćemo na uzorku utvrditi postojanje nekog fenomena koji zaista postoji u populaciji. Snaga statističkog testa zavisi od njegovih karakteristika, veličine uzorka i odabranog graničnog nivoa značajnosti. To znači da je na osnovu željenih i/ili poznatih parametara, moguće proceniti da li je uzorak koji istraživač planira da prikupi dovoljan da bi omogućio utvrđivanje statističke značajnosti neke razlike ili povezanosti. O ovom postupku biće nešto više reči kasnije.

Do sada smo se u opisivanju istraživačkih hipoteza, nivoa verovatnoće i nivoa značajnosti, u znatnoj meri oslanjali na teorijsku i matematičku osnovu ovih koncepata. To nas je donekle udaljilo od praktičnih aspekata primene statističkih metoda. Dodatno, logički puritanizam vezan za opisane zablude o testiranju hipoteza, može da zbuni čitaoca. Postavlja se pitanje zbog čega uopšte računamo vrednosti test-statistika i p nivoe, ako to ne služi proveri tačnosti nul-hipoteze. Jednostavan odgovor je da nisko p nije toliko snažan argument u prilog netačnosti nul-hipoteze koliko njegova vrednost sugeriše (Nickerson, 2000), ali ipak upućuje na veću verovatnoću da je tačna alternativna hipoteza. U tom smislu, istraživač ne treba da bude robot koji čeka da p nivo

pređe granicu od 0,01 ili 0,05, već da dobijene rezultate interpretira savesno, odgovorno i sa dovoljno pratećih argumenata u prilog iznetih tvrdnji. Jedan od tih argumenata mogu da budu ranije pomenuti intervali poverenja ili pouzdanosti izračunatih statistika koji su mnogo informativniji i korisniji od p vrednosti i mehaničke konstatacije da je neka tvrdnja tačna ili netačna. U narednih nekoliko odeljaka, postupak testiranja nul-hipoteza stavićemo u kontekst primene nekoliko bazičnih inferencijalnih statističkih metoda.

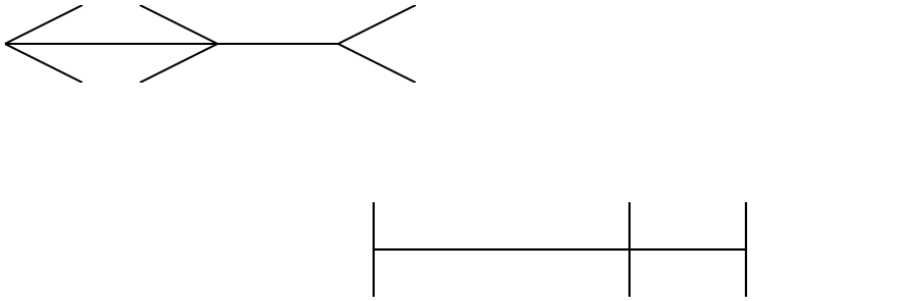
3.2. T-test za jedan uzorak

Logiku metode koja je poznata kao *t-test za jedan uzorak* opisali smo u odeljku o Studentovoj t distribuciji. Reč je o postupku kojim se uz pomoć intervala poverenja aritmetičke sredine, izračunate na jednom uzorku merenja, testira značajnost njene razlike od neke unapred date vrednosti. Za početak, pođimo od sledeće nulte hipoteze:

H_0: Efekat Miler-Lajerove iluzije nije statistički značajan

Kratkim eksperimentom proverićemo da li postoji efekat optičke iluzije koju je osmislio nemački sociolog Franc Karl Miler-Lajer. U pitanju je pojava da **horizontalnu duž opažamo kao kraću** kada je ograničena strelicama koje su usmerene ka spolja, nego kada su strelice usmerene ka unutra. U našem primeru upotrebićemo **prilagođenu verziju ove iluzije**. Kada kliknete taster *Enter*, biće prikazana jedna horizontalna duž i tri izlomljene vertikalne duži (strelice) od kojih su leva i desna usmerene ulevo, a srednja udesno (Slika 37). Vaš zadatak je da uz pomoć tastera za kontrolu kursora na tastaturi postavite središnju duž u poziciju u kojoj će seći horizontalnu duž tačno na pola. Da biste brže pomerali središnju duž, držite pritisnut taster *Ctrl*. Rastojanje od levog kraja horizontalne duži do mesta na kome se ona seče sa srednjom strelicom, treba da bude isto kao rastojanje od mesta tog preseka do tačke u kojoj desna strelica dodiruje horizontalnu duž. Nakon što ste postavili duž na željeno mesto, pritisnite taster *Enter* da biste sačuvali rezultat i prikazali novi zadatak. Pošto tačnost procene polovine duži ne zavisi samo od uticaja Miler-Lajerove iluzije, već i od drugih faktora, u vežbu ćemo uključiti skup zadataka kod kojih se umesto strelica nalaze vertikalne duži koje seku horizontalnu duž pod uglom od 90 stepeni. U grupi ovih merenja, tačnost procene bi trebalo da bude veća. Nakon što ste uradili dva zadatka za vežbu, potrebno je da obavite još 14

merenja, 7 sa strelicama, tj. izlomljenim vertikalnim dužima i 7 sa pravim. Nakon završene vežbe, podaci će biti prikazani histogramima.



Slika 37. Primer zadatka pomoću koga se testira efekat Miler-Lajerove iluzije

Koliko varijabli, odnosno svojstava koja su menjala svoju vrednost, možete da prepoznate u ovom primeru? Nabrojte ih.

Koje varijable su bile pod kontrolom eksperimentatora, tj. autora udžbenika, a koje su zavisile od ispitanika, tj. vas?

Zbog čega je eksperimentator namerno varirao dužinu horizontalne duži, poziciju duži na ekranu i početnu poziciju srednje vertikalne duži?

Zbog čega je merenje ponavljano više puta, umesto da je svaki tip vertikalnih duži bio prikazan po jednom?

Ključna varijabla u našem primeru je tačnost procene polovine duži. Na osnovu njenih vrednosti možemo da utvrdimo u kojoj meri ste kao ispitanik bili podložni efektu Miler-Lajerove iluzije. Osim toga, interesuje nas i da li tačnost zavisi od tipa vertikalnih duži, tj. da li ste pravili tačnije procene kada su se duži sekle pod pravim ili pod oštrim uglom. U istraživanju nam nije bilo bitno da li su na tačnost uticale dužina horizontalne duži, pozicija na ekranu i početni položaj srednje duži, tako da vrednosti ovih varijabli nisu ni beležene. One su namerno varirane od strane eksperimentatora kako bi se kontrolisao efekat drugih potencijalno bitnih faktora, npr. da bi se isključio efekat uvežbavanja. Stoga ćemo ove varijable nazvati *kontrolnim*. Svojstvo koje govori o ishodu merenja i reakcijama ispitanika obično se naziva *zavisnom varijablom*. U našem

primeru, to je tačnost procene. Svojstvo čiji se efekat na zavisnu varijablu proverava i čije vrednosti namerno varira osoba koja je dizajnirala istraživački nacrt, naziva se *nezavisnom varijablom*. U našem primeru, to je tip, odnosno ugao vertikalnih duži kao dihotomna varijabla. S obzirom na to da se dužina horizontalne duži nasumično menjala u 14 prikazanih zadataka, tačnost procene izračunata je kao odnos rastojanja do mesta na koje ste postavili središnju duž i ukupne dužine horizontalne duži. Klikom na ikonicu pored druge padajuće liste možete da preuzmete matricu sa sirovim podacima u kojoj je vrednosti u koloni *procena* potrebno podeliti vrednostima u koloni *duzina* da bi se dobio pokazatelj tačnosti procene. Tako dolazimo do kriterijuma od 0,5 koji ukazuje na savršeno tačnu procenu. Ukoliko uporedite vrednosti M za grupu plavih merenja (prav ugao) i M za grupu narandžastih merenja (oštar ugao), trebalo bi da uočite da je vaša procena bila tačnija, odnosno bliža vrednosti 0,5, kada su vertikalne duži bile prave. Izlomljene vertikalne duži usmerene ka unutra, odnosno strelice usmerene ka spolja, vizuelno su skraćivale levi segment horizontalne duži, tako da ste najverovatnije imali nesvesnu potrebu da ga produžite. Stoga je M za tu grupu merenja vrlo verovatno veća od 0,5. Kasnije ćemo se vratiti na podatke koje ste samostalno prikupili, a sada ćemo objasniti postupak primene t-testa za jedan uzorak na nekoliko pripremljenih skupova podataka.

Odaberite primer *M-L iluzija* sa liste ili kliknite link *Prikaži primer* ukoliko je otvoren okvir sa zadatkom. U zagradama je navedena referentna vrednost sa kojom se porede dobijene aritmetičke sredine. Radi boljeg razumevanja grafičkog prikaza, sirovi rezultati su sortirani od najmanjeg ka najvećem i navedeni na spisku koji se nalazi između grafikona i tabele sa deskriptivnim pokazateljima. Podaci su podeljeni u dve grupe od po 7 vrednosti i označeni su različitim bojama i simbolima – plavo || za grupu merenja kod kojih je ugao preseka bio prav, odnosno narandžasto >< za grupu merenja kod kojih je ugao preseka bio oštar. Aritmetička sredina plavih merenja iznosi 0,511 i razlikuje se od utvrđenog standarda. Međutim, potrebno je proveriti da li je ta razlika i statistički značajna. Ako na osnovu M i s_M izračunamo interval poverenja od 99%, dobićemo raspon vrednosti između 0,491 i 0,530. Pošto ranije utvrđeni standard pripada tom rasponu, ne smemo da tvrdimo da se 0,511 statistički značajno razlikuje od 0,5. Numeričku potvrdu ovog zaključka daje nam vrednost t. Kada uvrstimo dostupne informacije u odgovarajuću formulu:

$$t = \frac{M - \mu}{s_M} = \frac{0{,}51057 - 0{,}5}{0{,}00533} \approx 1{,}985$$

dobijamo vrednost t-testa koja iznosi približno 1,985. Obratite pažnju na to da su vrednosti deskriptivnih pokazatelja u tabeli zaokružene na tri decimale tako da nećete dobiti istu vrednost t ukoliko ih tako skraćene uvrstite u gornju formulu. Na osnovu odgovarajuće teorijske distribucije, može se utvrditi da vrednost t za 7 merenja, odnosno za 6 stepeni slobode (N - 1), treba da bude veća od 2,447 da bi se smatrala statistički značajno udaljenom od nule na nivou 0,05, a veća od 3,708 da bi se smatrala statistički značajnom na nivou 0,01. Naravno, od istraživača se ne očekuje da samostalno pronalazi ove granične vrednosti, već samo da razume logiku nivoa značajnosti i nivoa verovatnoće. U programima za statističku obradu, uz datu vrednost t-testa uvek se prikazuje i odgovarajuća p vrednost, koja u našem primeru iznosi 0,104. To znači da čak i ako prava μ svih mogućih procena polovina duži iznosi 0,5, postoji približno 10% verovatnoće da bismo u uzorku veličine 7 nasumično dobili M koje odstupa za 1,985 ili više svojih standardnih grešaka od te vrednosti. Ova verovatnoća je suviše velika da bismo 0,511 smatrali statistički značajno udaljenim, odnosno različitim od 0,5. Stoga nultu hipotezu ne treba da odbacimo. U izveštaju o istraživanju dobijeni rezultati obično se prikazuju na sledeći način:

$$t(6) = 1{,}985,\ p > 0{,}05$$

ili, češće, uz navođenje konkretne p vrednosti:

$$t(6) = 1{,}985,\ p = 0{,}104$$

Broj u zagradama označava broj stepeni slobode. Kada isti princip primenimo na uzorku od 7 merenja u kojima su horizontalne duži bile izlomljene, dobijamo interval potencijalnih vrednosti μ koji se kreće od 0,550 do 0,592 i koji ne obuhvata zadati kriterijum od 0,5. U ovom slučaju, hipotezu ćemo odbaciti kao netačnu, jer je izuzetno mala verovatnoća da slučajno dobijemo M koje toliko odstupa od 0,5. Drugim rečima, zaključujemo da je veća verovatnoća da M = 0,571 potiče iz populacije procena čija aritmetička sredina nije 0,5, odnosno da je na tačnost naše procene uticao ugao vertikalnih duži. Ovim smo statistički dokazali postojanje efekta Miler-Lajerove iluzije:

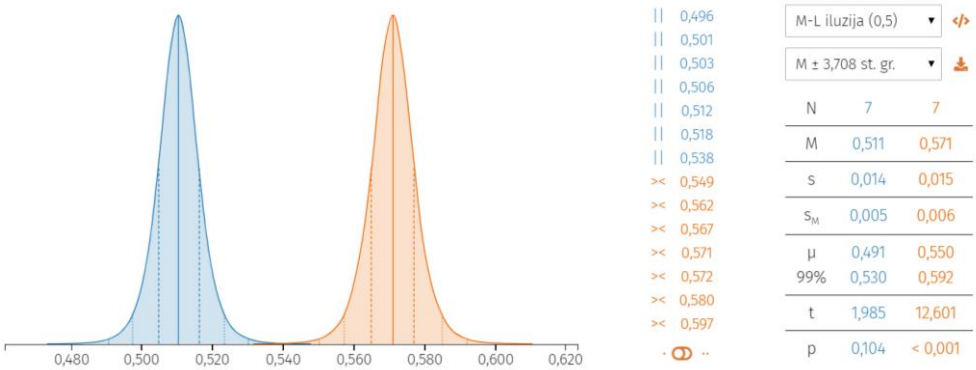
$$t(6) = 12{,}601,\ p < 0{,}001$$

Zbog čega, prilikom računanja intervala poverenja aritmetičke sredine od 99%, nije upotrebljen raspon $M \pm 2{,}58 \cdot s_M$, već raspon $M \pm 3{,}708 s_M$?

Kolika bi trebalo da bude s_M u grupi plavih merenja da bismo M = 0,511 smatrali statistički značajno različitom od vrednosti 0,5 na nivou 0,05?

Kolika bi bila vrednost t da je u grupi narandžastih merenja M iznosilo 0,429?

Kolika bi bila vrednost p da smo u grupi plavih merenja primenili postupak jednostranog testiranja razlike?



Slika 38. Primer teorijskih distribucija aritmetičkih sredina (intervala poverenja) u primeru testiranja postojanja Miler-Lajerove iluzije

Aritmetička sredina i standardna devijacija predstavljaju krajnji stepen sažimanja nekog skupa podataka. Ukoliko je varijabla približno normalno distribuirana, uz pomoć ovih vrednosti možemo da „rekonstruišemo", odnosno da aproksimiramo izgled distribucije sirovih podataka. Odaberite opciju *M ± 3 st. dev.* sa druge liste da biste videli ovu teorijsku „rekonstrukciju" sirovih podataka iz prvog primera u formi dve krive gustine verovatnoće. Krive prikazuju raspone od približno 99% sirovih rezultata dobijenih u istraživanju. One se delimično preklapaju, jer bi se uz dobijene vrednosti M i s za dve grupe merenja moglo očekivati da određeni broj procena iz narandžaste grupe bude tačniji od onih iz plave grupe. Međutim, procenat takvih rezultata je izuzetno mali, tako da već na osnovu grafikona možemo da zaključimo da dve grupe merenja, odnosno dva uzorka, ne potiču iz iste populacije. Pravu potvrdu ove tvrdnje nam, međutim, daju slike teorijskih distribucija potencijalnih vrednosti

μ, odnosno potencijalnih vrednosti M drugih uzoraka iste veličine, uzetih iz iste populacije. Odaberite opciju *M ± 3,708 st. gr.* (standardne greške M) sa druge liste da biste prikazali intervale poverenja M, odnosno raspone u kojima se, uz 99% verovatnoće, nalaze vrednosti μ (Slika 38). Vrednost 0,5 pripada rasponu plave distribucije, ali ne i rasponu narandžaste, što pokazuje da se aritmetička sredina plave distribucije (0,511) ne razlikuje statistički značajno od vrednosti 0,5, dok je aritmetička sredina narandžaste distribucije (0,571) značajno udaljena od kriterijuma. Osim toga, dve distribucije se praktično ni ne dodiruju, što znači da μ za uzorak procena u slučaju pravih duži, najverovatnije nije isto kao i μ za uzorak procena u slučaju izlomljenih. Dakle, $M \pm 3 \cdot s$ i $M \pm 3 \cdot s_M$ su dva suštinski različita intervala, od kojih nam prvi govori o tome kako izgleda očekivana distribucija sirovih podataka (x), pod pretpostavkom da su oni normalno distribuirani, a drugi kakva je očekivana distribucija velikog broja aritmetičkih sredina (M), koje se uvek distribuiraju u skladu sa normalnom raspodelom kada su uzroci veliki. Drugi interval, u zavisnosti od toga kako ga izračunamo, obuhvata određeni procenat vrednosti M koje bismo verovatno dobili na drugim uzorcima iste veličine, uzetim iz iste populacije. Istraživač jednostavno pretpostavlja da je jedno od tih M u stvari prava vrednost μ.

3.3. T-test za dva uzorka

U prethodnim primerima upotrebili smo postupak računanja t statistika da bismo proverili značajnost razlike aritmetičkih sredina od unapred definisanog kriterijuma tačnosti. Primenili smo t-test za jedan uzorak na svakom od dva skupa merenja i utvrdili da je u jednom slučaju M bila statistički značajno različita od 0,5, a u drugom nije. Na kraju nam je grafički prikaz intervala poverenja obe M sugerisao da one najverovatnije potiču iz različitih populacija. Štaviše, t-test nam omogućava da ovu pretpostavku proverimo i numerički. Početnu hipotezu koja je glasila:

H_0: Efekat Miler-Lajerove iluzije nije statistički značajan

možemo da formulišemo preciznije:

H_0: Ne postoji statistički značajna razlika u proceni polovine horizontalne duži između grupe merenja sa pravim vertikalnim dužima i grupe merenja sa izlomljenim vertikalnim dužima

ili, jednostavnije:

H_0: Ne postoji statistički značajan uticaj ugla preseka duži na tačnost procene polovine horizontalne duži

U ovom slučaju, više nam nije cilj da poredimo vrednosti M sa očekivanom vrednošću μ, već da *međusobno* uporedimo dve aritmetičke sredine i da na osnovu vrednosti t statistika utvrdimo da li se one razlikuju *jedna od druge*. Naša hipoteza se, dakle, može izraziti i na sledeći način:

$$H_0: \mu_1 = \mu_2 \text{ ili } H_0: \mu_1 - \mu_2 = 0$$

Pošto poredimo dve uzoračke aritmetičke sredine, primenićemo tehniku koja se zove *t-test za dva uzorka*. U ovom slučaju, formula za računanje t vrednosti se neznatno razlikuje od prethodne, ali je njena logika potpuno ista kao u slučaju t-testa za jedan uzorak ili u slučaju računanja z vrednosti. Suština je u tome da se neko odstupanje, npr. x - M ili M - μ, dovede u vezu sa pokazateljem intenziteta svih odstupanja, npr. s ili s_M. U slučaju t-testa za dva uzorka, reč je o odstupanju dobijene razlike M_1 - M_2 od nule koja se očekuje na osnovu nulte hipoteze. Varijabilnost tih odstupanja naziva se *standardna greška razlike* aritmetičkih sredina i označava se simbolom s_{M1-M2}. Greška razlike direktno zavisi od veličine grešaka pojedinačnih aritmetičkih sredina i računa se prema formuli:

$$s_{M_1-M_2} = \sqrt{s_{M_1}^2 + s_{M_2}^2}$$

Tako dolazimo do opšte formule za računanje t-testa za dva uzorka:

$$t = \frac{(M_1 - M_2) - 0}{s_{M_1-M_2}} = \frac{M_1 - M_2}{\sqrt{s_{M_1}^2 + s_{M_2}^2}}$$

Smisao standardne greške razlike aritmetičkih sredina mogli bismo da objasnimo na isti način kao i standardnu grešku aritmetičke sredine. U poglavlju 2.6.5. pokazali smo da se iz populacije u kojoj je aritmetička sredina μ, dobijaju uzorci čije se aritmetičke sredine distribuiraju u skladu sa t raspodelom oko vrednosti μ, sa standardnom devijacijom čija je vrednost s_M. Isti princip važi i za razlike aritmetičkih sredina. Čak i ako su aritmetičke sredine dveju populacija iste, to ne znači da će za svaki par uzoraka uzetih iz tih populacija vrednost M_1

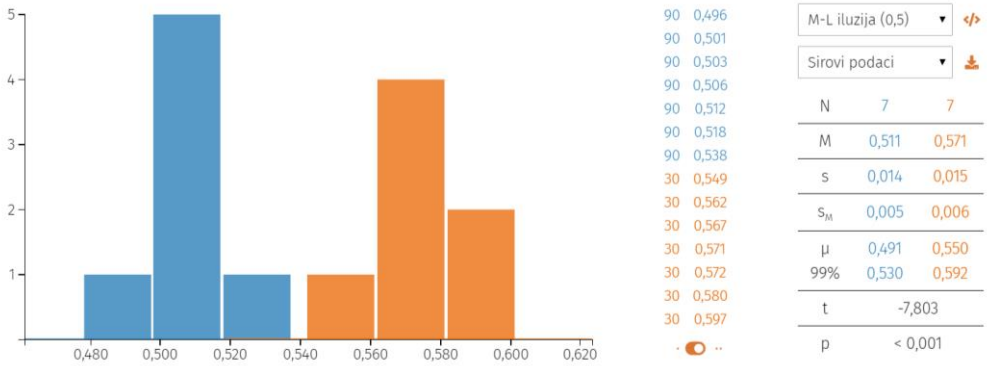
- M_2 biti 0. U stvari, dobijene razlike biće distribuirane oko vrednosti 0 (ili bilo koje prave vrednosti razlike $\mu_1 - \mu_2$) sa varijabilnošću s_{M1-M2}, takođe u skladu sa t raspodelom. Analogno vrednosti s_M, možemo da kažemo da je standardna greška razlike aritmetičkih sredina zapravo standardna devijacija velikog uzorka razlika aritmetičkih sredina koje su izračunate na parovima uzoraka uzetim iz odgovarajuće ili odgovarajućih populacija. S obzirom na to da ne znamo da li uzorci potiču iz iste ili iz različitih populacija, o poreklu uzoraka možemo da zaključujemo samo na osnovu vrednosti t-odnosa. Ukoliko je on blizak nuli, pretpostavićemo da dva uzorka, odnosno dve M, potiču iz iste populacije. Ukoliko je t vrednost dovoljno velika, zaključićemo suprotno. U slučaju velikih uzoraka, dovoljno velika apsolutna vrednost je ona iznad 1,96 ili 2,58, što znači da razlika treba da bude približno dva ili tri puta veća od svoje greške da bi se smatrala statistički značajnom. Na neki način razlika postaje naš „ispitanik", a t postaje njegova „z vrednost". Ispitanika (tj. razliku) čija je apsolutna vrednost z (tj. t) velika, posmatramo kao atipičnog i udaljenog od proseka grupe. Prosek grupe u slučaju testiranja značajnosti razlike iznosi nula jer se to pretpostavlja nultom hipotezom.

Vratimo se na **primer *M-L iluzija* iz prethodnog odeljka.** U dnu spiska sirovih podataka nalazi se prekidač pomoću koga se prikaz rezultata u tabeli menja između dva t-testa za jedan uzorak i jednog t-testa za dva uzorka (Slika 39). Kliknite prekidač da biste prikazali vrednost t-test za odabrani primer. Na osnovu vrednosti t i nivoa verovatnoće p, možemo da zaključimo da se dve aritmetičke sredine statistički značajno razlikuju ne samo na nivou 0,01, već i na nivou 0,001:

$$t(12) = -7{,}803,\ p < 0{,}001$$

Pošto smo M_1 i M_2 računali na dva skupa od po 7 merenja, broj stepeni slobode je $(N_1 - 1) + (N_2 - 1)$ ili N - 2, gde je N veličina uzorka, odnosno ukupan broj merenja koji u našem primeru iznosi 14. Obratite pažnju na to da se, osim poslednja dva reda u tabeli, delimično izmenio i spisak sirovih rezultata. Ovoga puta su u prvoj koloni navedene vrednosti, tj. kodovi grupišuće varijable – 90 za merenja u kojima su se duži sekle pod pravim uglom, odnosno 30 za merenja u kojima su se sekle pod oštrim uglom od približno 30 stepeni. To je ujedno i pojednostavljen prikaz strukture matrice sirovih podataka potrebne da bi se izračunao t-test za dva uzorka. U ćelije jedne kolone matrice potrebno je upisati vrednosti dihotomne nezavisne varijable na osnovu koje se merenja dele u dve grupe, a u ćelije druge kolone vrednosti kvantitativne zavisne varijable za koju

se računaju aritmetičke sredine po grupama. Broj redova u matrici jednak je ukupnom broju merenja, što u našem primeru iznosi 14.



Slika 39. Primer distribucija čije se aritmetičke sredine razlikuju statistički značajno na nivou 0,001 – t(12) = -7,803, p < 0,001

Kako treba da izgleda matrica sirovih podataka, na osnovu kojih biste mogli da proverite da li se visina 30 dečaka statistički značajno razlikuje od visine 25 devojčica?

Zbog čega je u slučaju t-testa za dva uzorka hipoteza postavljena u formi H_0: $\mu_1 = \mu_2$ a ne u formi H_0: $M_1 = M_2$?

Opisanu logiku t-testa za dva uzorka možemo lako preneti na bilo koji istraživački zadatak u kome treba uporediti proseke dveju grupa merenja ili ispitanika. Na drugoj listi odaberite opciju *M ± 3,708 st. gr.*, a na prvoj odaberite primer *Završni ispit 1*. Zamislimo da su u pitanju bodovi koje su učenici osmih razreda dve osnovne škole dobili na završnom ispitu. Na osnovu izračunate t vrednosti, kao i na osnovu grafički prikazanih intervala poverenja koji se potpuno preklapaju, zaključujemo da se đaci dve škole ne razlikuju značajno u uspehu. Na sličan zaključak upućuje i poređenje procenjenih distribucija sirovih rezultata koje se prikazuju izborom opcije *M ± 3 st. dev*. Međutim, sada je uočljivije da su đaci OŠ2 u proseku nešto bolji od đaka OŠ1, kao i da se varijabilnost rezultata u grupama značajno razlikuje. Ovakva *nehomogenost* varijansi može da predstavlja ozbiljan problem u interpretaciji t-testa za dva uzorka. Tek kada prikažemo grafikon sirovih rezultata, postaje jasno da je velika varijabilnost u grupi OŠ1 posledica aberantnog rezultata zbog koga su se

povećale vrednosti M i s_M a umanjila vrednost t. Ukoliko taj aberantni rezultat uklonimo iz analize izborom primera *Završni ispit 2*, ustanovićemo da se dve grupe ipak statistički značajno razlikuju:

$$t(11) = -4{,}540,\ p = 0{,}001$$

Obratite pažnju na to da se histogrami i rezultirajuće krive gustina verovatnoća za poslednja dva primera razlikuju u samo jednom rezultatu. Taj rezultat je drastično povećao standardnu devijaciju i standardnu grešku aritmetičke sredine u grupi OŠ1, a time i standardnu grešku razlike aritmetičkih sredina. Takođe, imajte na umu činjenicu da smo poredili dve grupe đaka, nezavisno od bilo kakvog kriterijuma. Obe grupe su mogle, ali nisu morale, da budu značajno bolje ili lošije od nekog kriterijuma, a da to i dalje ne utiče na njihovu međusobnu razliku. U ovom primeru, kriterijum je iznosio 50 bodova i naveden je u zagradi. Kada promenite prikaz t vrednosti uz pomoć prekidača, videćete da se učinak đaka OŠ1 ne razlikuje značajno od kriterijuma, dok su đaci OŠ2 u proseku značajno bolji od njega.

> Na osnovu kojih podataka iz prikazane tabele biste mogli da zaključite da u primeru *Završni ispit 1* možda postoji aberantan rezultat?
>
> Zbog čega vrednosti t-testova za jedan uzorak u primeru *Završni ispit 2* imaju suprotan predznak?
>
> U kom slučaju bi t-test za dva uzorka u primerima sa završnim ispitom imao istu apsolutnu vrednost ali drugačiji predznak? Kako bi to uticalo na naš zaključak o razlikama među grupama?

3.3.1. Uslovi za primenu t-testa

Prikažite histograme sirovih podatka za primer *BMI* i odaberite t-testove za jedan uzorak. Nezavisna varijabla u ovom primeru je način ishrane, na osnovu koga su ispitanici podeljeni na vegetarijance i one koji jedu meso. Zavisna varijabla je *indeks telesne mase* (engl. *BMI – body mass index*) koji se računa kao odnos mase osobe i kvadrata njene visine. Nivoi verovatnoće t-odnosa ukazuju da se proseci grupa ne razlikuju značajno od vrednosti 20 koja je odabrana kao kriterijum, tj. poželjan odnos mase i visine. To znači da ni

vegetarijanci ni mesojedi *u proseku* nisu ni gojazni ni neuhranjeni. Sa druge strane, na osnovu vrednosti t-testa za dva uzorka, zaključujemo da su osobe koje jedu meso (M) statistički značajno gojaznije od vegetarijanaca (V), ali na nivou 0,05. Drugim rečima, način ishrane vrlo verovatno utiče na povećanje telesne mase. Međutim, ovu tvrdnju možemo da smatramo validnom samo ako su ispunjeni uslovi da se promene na zavisnoj varijabli (npr. razlika u prosečnoj masi među grupama) pripišu promenama na nezavisnoj varijabli (npr. načinu ishrane), a ne drugim faktorima (npr. polu). Prvi i osnovni uslov je *slučajno i nezavisno razvrstavanje ispitanika u grupe*. Grupe moraju da budu homogene po svim relevantnim svojstvima, osim po onom koje je određeno kao nezavisna varijabla. U našem primeru, to znači da grupe ispitanika moraju da se razlikuju samo u načinu ishrane, a da istovremeno budu ujednačene po svim drugim relevantnim varijablama. Broj takvih varijabli zavisi od istraživačkog nacrta i kompleksnosti fenomena koji se opisuju. U ovom zamišljenom istraživanju, one bi trebalo da obuhvate sve faktore koji mogu da budu povezani sa telesnom masom osobe: pol, starost, bavljenje fizičkim aktivnostima, hormonski status i slično. Dakle, dve grupe moraju da se formiraju tako da sadrže podjednak broj muškaraca i žena, osoba koje se bave ili ne bave fizičkim aktivnostima, osoba različitog uzrasta itd. U suprotnom, razlika u masi mogla bi da bude posledica pristrasnosti uzorkovanja, odnosno činjenice da je, na primer, u jednoj grupi bilo više osoba koje redovno idu u teretanu, a u drugoj više osoba koje se retko bave fizičkim aktivnostima.

Povećanje broja relevantnih svojstava koje je potrebno ujednačiti po grupama i držati pod kontrolom, podrazumeva i povećanje varijabilnosti. To nas dovodi do drugog važnog uslova za primenu t-testa – *uzorci moraju da budu dovoljno veliki i približno iste veličine*. Iako se u literaturi mogu pronaći konkretni kriterijumi za razlikovanje velikih od malih uzoraka, npr. više od 30, 50 ili 100 ispitanika, potrebnu veličinu uzorka nikada ne treba definisati kao apsolutnu granicu. Ako se prisetimo formule za izračunavanje standardne greške aritmetičke sredine, primetićemo da je (razumno) povećavanje veličine uzorka način da se kompenzuje velika varijabilnost pojave koju merimo i da se smanji greška procene parametara populacije. Pri tome se varijabilnost ne odnosi samo na varijanse pojedinačnih varijabli, već i na složenost istraživačkog nacrta, odnosno ukupan broj varijabli u sistemu koji analiziramo. U našem primeru, grupe od po sedam ispitanika najverovatnije nisu dovoljne da se pokrije i opiše varijabilnost sistema koji ne čine samo nezavisna i zavisna varijabla, već i niz drugih relevantnih kontrolnih varijabli. Pitanjima procene potrebne veličine

uzorka za konkretan istraživački cilj, bave se tehnike *analize statističke snage* (engl. *power analysis*) koju smo pomenuli u odeljku o testiranju hipoteza. Većina statističkih paketa poseduje opcije pomoću kojih je moguće proceniti snagu statističkog testa na osnovu pretpostavljene veličine uzorka i željenog nivoa značajnosti ili, pak, potrebnu veličinu uzorka, na osnovu željene snage testa, tj. verovatnoće da će hipoteza biti odbačena. Pri tome, procena potrebne veličine uzorka ne odnosi se samo na određivanje minimalnog broja merenja koje je potrebno obaviti da bi se došlo do nekog zaključka. Kao što smo već nekoliko puta napomenuli, uvek je bolje da uzorak bude što veći, ali ne veći od optimalnog nivoa nakon kog se više ne postiže porast preciznosti ili snage. U tom smislu, analiza snage olakšava optimizaciju napora i materijalnih ulaganja, jer u statistici uzorci mogu da budu i preveliki, tj. nepotrebno veliki. Kao ilustraciju ćemo upotrebiti besplatnu alatku **Statulator** da bismo procenili potrebnu veličinu uzorka u primeru sa indeksom telesne mase. Ako unesemo podatak da je očekivana razlika približno 2,5 BMI jedinica, varijabilnost približno 1,5 BMI jedinica, nivo značajnosti 0,05, a željena snaga t-testa 80%, program sugeriše da je dovoljno da naš uzorak ima 12 ispitanika (Dhand & Khatkar, 2014). Međutim, kada bi očekivana varijabilnost bila veća, npr. 5 BMI jedinica, bio bi potreban 10 puta veći uzorak da bi se postigla verovatnoća odbacivanja netačne nul-hipoteze od 80%. Naravno, kada bismo želeli još veću snagu i/ili primenili stroži nivo značajnosti od 0,01, uzorak bi trebalo da bude još veći. Treba imati na umu da procenjena snaga testa govori o tome kolika je verovatnoća da dobijemo rezultat nakon koga ćemo odbaciti nultu hipotezu, ali samo pod pretpostavkom da je ona netačna u populaciji. Drugim rečima, statistički testovi ne mogu da se „nateraju" da neku razliku prikažu kao značajnu ako ona ne postoji u populaciji, čak ni ako su uzorci veoma veliki, odnosno, preciznije rečeno, *pogotovo* kada su uzorci veliki.

Dodatni uslovi za primenu t-testa proističu iz činjenice da se ova tehnika bazira na poređenju dve aritmetičke sredine. Za početak, to znači da zavisna varijabla mora da bude *kvantitativna i izmerena na intervalnom ili racio nivou merenja*. Nisu retki primeri u kojima se t-test primenjuje i na varijablama rang nivoa, npr. odgovorima ispitanika na skalama slaganja od 1 do 5, ali u takvim situacijama treba biti veoma oprezan. Računanje proseka često je *opravdano* ali *besmisleno*, jer (be)smislenost nekog zaključka nije isto što i njegova (ne)tačnost (Marcus-Roberts & Roberts, 1987). Tvrdnja da je prosek visine učenika nekog odeljenja veći od prosečne visine školskih zgrada jeste netačna ali nije besmislena, jer ispravnost te tvrdnje može da se proveri, bez obzira na to u

kojim jedinicama je izražena visina. Međutim, tvrdnja da je znanje matematike grupe učenika čija je prosečna ocena 4,86, duplo veće od znanja matematike učenika čija je prosečna ocena 2,43, iako deluje kao tačna, nema nikakvog smisla. Prvi razlog je činjenica da podaci ordinalnog nivoa govore o postojanju razlike među vrednostima, ali ne i o količini te razlike. Stoga su operacije množenja na vrednostima koje potiču sa ordinalnih skala najčešće besmislene. Drugi razlog je nedovoljna objektivnost školskih ocena, odnosno nedovoljna preciznost i nedovoljna ujednačenost načina na koji se „meri" ili izražava stepen nečijeg znanja. Dakle, istraživač prilikom odluke o primeni t-testa ne treba da se vodi samo pitanjem da li je opravdano da se test primeni, jer to zaista može da bude slučaj čak i ako podaci nisu intervalnog ili racio nivoa, već da li rezultati, odnosno zaključci do kojih će doći, imaju smisla. Stoga preporučujemo da se ne vodite toliko strogo kriterijumima primene testova vezanih za nivoe merenja, već da uvek razmislite da li to ima opravdanja. Primena t-testa može da bude besmislena čak i kada su podaci razmernog nivoa merenja.

U slučaju t-testa, aritmetičke sredine i standardne devijacije uzoraka koriste se kao procene parametara distribucije zavisne varijable u populaciji. Te procene su uvek pouzdanije ako su *distribucije varijabli približno normalne*. Osim toga, poređenje ovih procena je pouzdanije ukoliko je varijabilnost u grupama podjednaka, što znači da je *homogenost varijansi* još jedan od uslova za primenu t-testa. U primeru *Završni ispit 1* pokazali smo da odstupanje od normalnosti, a posebno postojanje aberantnih rezultata, može bitno da utiče na tačnost procene stanja u populaciji i da umanji pouzdanost dobijenih t i p vrednosti. Kao ilustraciju, prikažite sirove podatke za primer *Kurs jezika*. Možete da pređete pokazivačem miša preko stubića da biste jasnije videli oblike distribucija, s obzirom na to da se u nekim delovima histograma stubići potpuno preklapaju. Recimo da nam je cilj bio da proverimo razlike između studenata književnosti i studenata računarstva u stavu prema uvođenju drugog obaveznog kursa stranog jezika. Stav je izmeren pitanjem sa petostepenom skalom odgovora, od 1 – potpuno sam protiv, preko 3 – svejedno mi je, do 5 – potpuno se slažem. Vrednosti t-testova za jedan uzorak pokazuju da obe grupe studenata imaju neutralan stav, a vrednost t-testa za dva uzorka da se dve grupe *u proseku* ne razliku. Međutim, aritmetička sredina u ovom slučaju nije prikladna aproksimacija težišnih rezultata grupa merenja. Samim tim, ni dobijene t-vrednosti najverovatnije ne odražavaju pravo stanje u populacijama. Iako vrednost t-testa za dva uzorka nije statistički značajna, histogrami ukazuju da se raspodele, odnosno *strukture* studentskih odgovora bitno razlikuju među

grupama. Za početak, vrednost 3 nije adekvatna mera centralne tendencije ni za jednu od grupa, pa tako nije ni dobra osnova za procenu parametara u populaciji. Na primer, uočljivo je da većina studenata računarstva ima negativan stav prema uvođenju drugog jezika, za razliku od većine studenata književnosti.

Na kraju treba napomenuti da se u popularnim statističkim paketima primenjuju procedure kojima se postupak računanja t-testa koriguje kako bi se ublažio negativan uticaj nehomogenosti varijansi i bitno različitih veličina uzoraka. Pojedini autori tvrde da t-test ni izbliza nije toliko osetljiv kao što se u statističkim udžbenicima upozorava i da je dovoljno robustan, odnosno da pruža osnovu za pouzdane i validne zaključke, čak i kada su pomenuti uslovi prekršeni (Norman, 2010). Diskusija o tome da li se ordinalne skale mogu ili ne mogu tretirati kao intervalne i dalje traje, tako da je lako naći argumente u prilog oba stava (Knapp, 1990). To znači da je odgovornost na istraživačima a ne na autorima statističkih udžbenika i priručnika. Istraživači treba da poznaju potencijalne opasnosti vezane za kršenje pomenutih uslova za primenu t-testa, a posebno onih koji se tiču metodološke korektnosti. Pre nego što se podaci prikupe i obrade, potrebno je detaljno razraditi i osmisliti istraživački nacrt kako bi se uzeo u obzir efekat svih potencijalno relevantnih svojstava na vrednosti zavisne varijable. Osim toga, istraživač treba da bude upoznat i sa alternativnim tehnikama kojima može da ostvari isti cilj. U određenim situacijama, poželjno je sprovesti veći broj različitih analiza kako bi doneo ispravan i održiv zaključak.

> Prikažite ponovo podatke koje ste samostalno prikupili u vežbi i analizirajte ih u skladu sa opisanom logikom t-testa za jedan, odnosno dva uzorka.
>
> Analizirajte prikazane histograme i proverite da li je na prikupljenim podacima opravdano primeniti t-test.
>
> Ponovite vežbu tako da u nekom od zadataka namerno napravite netačan aberantan rezultat. Analizirajte distribucije i dobijene vrednosti t-testa.
>
> Na koji način može da se proveri, odnosno testira ispunjenost uslova homogenosti varijansi pre primene t-testa?

3.4. Neparametrijske alternative t-testu za dva uzorka

Statističke procedure u kojima se polazi od pretpostavke da distribucije u populaciji imaju određeni oblik i koje se baziraju na procenama parametara tih distribucija, nazivaju se *parametrijskim*. Tipičan predstavnik ove grupe je t-test. Kao što smo rekli, njegova logika zasniva se na očekivanju da je zavisna varijabla normalno distribuirana i da su aritmetičke sredine dvaju uzoraka adekvatne procene mera grupisanja u populaciji. Kada ti uslovi nisu ispunjeni, ne bi trebalo primenjivati t-test. U takvim situacijama se kao alternativa može upotrebiti neka od *neparametrijskih metoda*, odnosno statističkih tehnika kod kojih nije neophodno da distribucije budu određenog oblika, da varijable budu intervalnog ili racio nivoa, pa čak ni da varijable budu kvantitativne. Zbog manje strogih uslova primene, ove tehnike se obično smatraju robusnijim od parametrijskih. Međutim, s obzirom na činjenicu da je većina njih prilagođena ordinalnom i nominalnom nivou merenja varijabli, one su ujedno i manje precizne. Statističkim rečnikom rečeno, neparametrijske tehnike imaju *manju snagu* od parametrijskih. Prilikom opisivanja tipova grešaka u statističkom zaključivanju, rekli smo da se snaga statističkog testa odnosi na njegovu sposobnost da detektuje postojanje nekog fenomena, odnosno na verovatnoću da se odbaci pogrešna nulta hipoteza. Na neki način, parametrijski testovi su kao pinceta a neparametrijski kao klešta. Kleštima možemo da uhvatimo čipove na nekoj štampanoj ploči, ali ćemo tu operaciju uraditi grubo, a možda čak i oštetiti čip. Pincetom ćemo tu operaciju obaviti mnogo preciznije, ali zato njome ne možemo da vadimo eksere iz zida. Slično je i sa statističkim procedurama. Neparametrijske metode mogu da se primenjuju na varijablama intervalnog ili racio nivoa, ali se tako umanjuje količina iskorišćenih informacija i preciznost analize. Sa druge strane, parametrijske tehnike najčešće ne mogu i ne smeju da se primenjuju na varijablama ordinalnog i nominalnog nivoa, ali pružaju veću preciznost i povećavaju verovatnoću da se utvrdi postojanje neke pravilnosti, npr. razlike među grupama na zavisnoj varijabli.

3.4.1. Vold–Volfovicov test nizova

Vratimo se na **primer *M-L iluzija* i t-test za dva uzorka**. Razlika među merenjima je statistički značajna, što je vidljivo ne samo na osnovu udaljenosti distribucija podataka u grupama, već i na osnovu sortiranih sirovih rezultata,

koji jasno formiraju dva raznobojna kontinuirana niza od po 7 vrednosti. To pokazuje da su podaci grupisani na dve jasno razlučive „gomile" i da među njima postoji značajna razlika. Nijedan član plave grupe ne pripada rasponu rezultata narandžaste grupe niti obratno. Da bismo ilustrovali situaciju u kojoj među grupama ne postoji razlika, iskoristićemo primer *Agresivnost*. U pitanju su međupolne razlike u skorovima na skali agresivnosti nekog upitnika ličnosti. Distribucije se gotovo potpuno preklapaju a na spisku rezultata se smenjuju skorovi dečaka (1) i devojčica (2). Umesto dva niza koja smo videli u primeru sa Miler-Lajerovom iluzijom, ovde postoji čak 14 nizova od po jednog rezultata (Slika 40). Pojedinačni rezultati potpuno su izmešani, tj. nasumično raspoređeni među grupama, što znači da uzorci najverovatnije potiču iz iste populacije. Ovo je osnovna logika testiranja značajnosti razlike između dve grupe merenja primenom neparametrijske alternative t-testu koja se naziva *test homogenih nizova* i koju su osmislili mađarski statističar Abraham Vold i poljski statističar Jakob Volfovic. Po njima je tehnika nazvana *Vold–Volfovicov test nizova* (engl. *Wald–Wolfowitz runs test*). Ako je broj ovako formiranih nizova merenja manji od onoga koji bi mogao da se očekuje potpuno slučajno, može se zaključiti da rezultati ne potiču iz iste populacije. Za uzorke kod kojih je broj merenja po grupama manji od 20, očekivani broj nizova očitava se iz tablica graničnih vrednosti. Na primer, za dve grupe od po 7 merenja (ispitanika), granična vrednost iznosi 4. Ukoliko prikažete podatke iz primera *Završni ispit 1*, uočićete da se oni grupišu u tri niza – dva plava i jedan narandžasti. S obzirom na to da je broj nizova manji od date granične vrednosti, zaključićemo da razlika jeste statistički značajna. Aberantni rezultat, koji je povećao grešku aritmetičke sredine u jednoj od grupa i time umanjio vrednost t-testa u ovom primeru, nije uticao na zaključak koji smo doneli primenom testa nizova. Osim toga, zaključak bi bio isti čak i da je rezultat poslednjeg đaka na listi, umesto 75, imao vrednost 100, 150 ili 200. Razlog je to što se broj nizova ne bi promenio, jer se broj bodova u ovom slučaju tretira kao ordinalna varijabla, bez obzira na činjenicu da je izmerena na višem nivou.

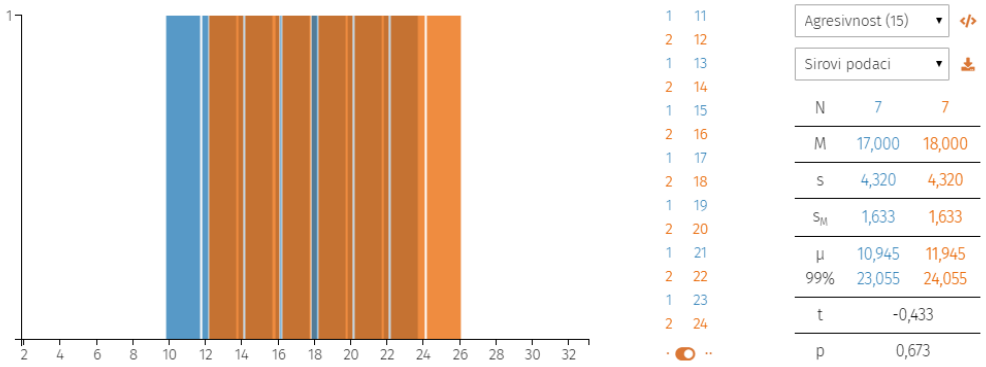
Vold–Volfovicovim testom zapravo se procenjuje da li je uočeni broj nizova mogao da se dobije potpuno slučajno, čak i da u populaciji ne postoji razlika među grupama. Ukoliko su uzorci veći od 20, distribucija očekivanog broja nizova je približno normalna, sa aritmetičkom sredinom koja se izračunava po formuli:

$$\mu = \frac{2N_1N_2}{N_1 + N_2} + 1$$

i standardnom devijacijom koja se izračunava po formuli:

$$\sigma = \sqrt{\frac{2N_1N_2(2N_1N_2 - N_1 - N_2)}{(N_1 + N_2)^2(N_1 + N_2 - 1)}}$$

gde su N_1 i N_2 veličine prvog i drugog uzorka. Potom se uz pomoć μ i σ izračuna z vrednost za broj nizova koji je dobijen u istraživanju, i interpretira se u skladu sa odabranim nivoima značajnosti i uobičajenim nivoima verovatnoće za normalnu distribuciju. Ako je, na primer, z vrednost veća od 1,96, to znači da je i broj nizova značajno drugačiji od onoga koji bi mogao da se dobije slučajno, pa možemo da kažemo da je razlika među grupama značajna na nivou 0,05. Vold–Volfovicov test nije naročito popularan, ali ga ovde navodimo kako bismo još jednom ukazali na smislenost različitih zaključaka o razlikama u zavisnosti od toga da li se zavisna varijabla tretira kao intervalna ili ordinalna.



Slika 40. Primer podataka koji formiraju 14 nizova i pokazuju da nema razlike između dve grupe od po 7 merenja

3.4.2. Kolmogorov–Smirnovljev test za dva uzorka

Odaberite ponovo primer *Kurs jezika*. Kao što smo rekli, primena t-testa u ovom slučaju nije prikladna, najpre zbog ordinalnog nivoa merenja zavisne varijable, a potom i zbog značajnog odstupanja distribucija od normalne i postojanja aberantnih rezultata. Već sama činjenica da decimalnim brojevima izražavamo mere centralne tendencije grupa ukazuje na neopravdanost donošenja zaključaka na osnovu aritmetičke sredine koja je izračunata za

ordinalnu varijablu sa samo pet mogućih vrednosti. Međutim, zaključak da se grupe studenata ne razlikuju u stavu prema uvođenju novog kursa ne bi bio drugačiji čak i da smo primenili Vold–Volfovicov test jer je najveći broj nizova koje mogu da formiraju rezultati veći od granične vrednosti 4. Ipak, grafički prikaz podataka sugeriše da se distribucije podataka po grupama bitno razlikuju i da struktura stavova studenata nije ista. U takvim situacijama poželjno je testirati istu hipotezu još nekim testom. Jedna od opcija bi mogao da bude postupak koji su osmislili ruski matematičari Andrej Kolmogrov i Nikolaj Smirnov, a kojim se testira upravo razlika *oblika* dve distribucije. Postupak primene Kolmogorov–Smirnovljevog (K–S) testa je prilično jednostavan. Najpre treba da se pronađe najveća pojedinačna razlika (D) kumulativnih empirijskih distribucija verovatnoća dve grupe. U našem primeru, najveća razlika je ona u učestalosti odgovora „nakupljenih" do vrednosti 3 i iznosi 5/7, jer u grupi studenata računarstva empirijska verovatnoća odgovora 1, 2 i 3 iznosi 6/7, a u grupi studenata književnosti 1/7. Dobijena razlika verovatnoća, u ovom slučaju 0,714, treba da bude veća od granične vrednosti za datu veličinu grupa. Granična vrednost izračunava se prema formuli:

$$D_{gr} = k(\alpha)\sqrt{\frac{N_1 + N_2}{N_1 N_2}}$$

gde su N_1 i N_2 veličine uzoraka, a k(α) konstantna koja se određuje u zavisnosti od željenog nivoa značajnosti. Nivoi se mogu odabrati i tako da se razlika testira jednostrano. Tako, na primer, za jednostrano testiranje razlike na nivou 0,05, ova vrednost iznosi 1,224, pa je vrednost D u našem primeru 0,654. Pošto je empirijska razlika veća od očekivane, zaključićemo da se distribucije dveju grupa statistički značajno razlikuju na nivou 0,05 i da studenti računarstva imaju statistički značajno *negativniji* stav prema uvođenju dodatnog obaveznog kursa stranog jezika od studenata književnosti. Time ne tvrdimo da se nužno razlikuju i mere centralne tendencije dve grupe merenja, već samo njihove distribucije, odnosno kumulativne empirijske verovatnoće različitih ishoda.

Poslednji primer ilustruje činjenicu da je u istraživanjima podjednako važno (ako ne i važnije) postaviti odgovarajuće pitanje, kao što je važno odabrati adekvatan statistički metod kojim će se dati odgovor na postavljeno pitanje. Iako testiranje razlika između dve grupe merenja obično podrazumeva poređenje njihovih mera centralne tendencije, to nije uvek i najprikladniji izbor. Štaviše, ponekad ova vrsta testiranja može da navede istraživača na pogrešan

zaključak. Kao što smo više puta pokazali u odeljku o deskriptivnim pokazateljima, dve distribucije mogu da imaju iste ili veoma bliske vrednosti aritmetičkih sredina ili medijana a da pri tome imaju potpuno drugačije oblike i različitu varijabilnost. Na primer, odeljenje u kome svi đaci imaju ocenu 3 iz matematike i odeljenje u kome polovina đaka ima 1 a druga polovina 5, neće se razlikovati u prosečnom učinku, ali će se bitno razlikovati u raspodeli ocena. Ova druga informacija može da bude važnija od prostog i površnog zaključka da je prosečna ocena u oba razreda 3, jer može da ukaže na drugačije kriterijume koje nastavnik ima prema deci ili na problem neujednačenosti đaka u odeljenju po sposobnostima i interesovanjima. Opisane alternative t-testu, kao što su Vold–Volfovicov i Kolmogorov–Smirnovljev test, prikladnije su u situacijama u kojima je, pored mera centralne tendencije, potrebno uporediti i mere varijabilnosti, odnosno oblike distribucija dve grupe merenja.

3.4.3. Men–Vitnijev test sume rangova

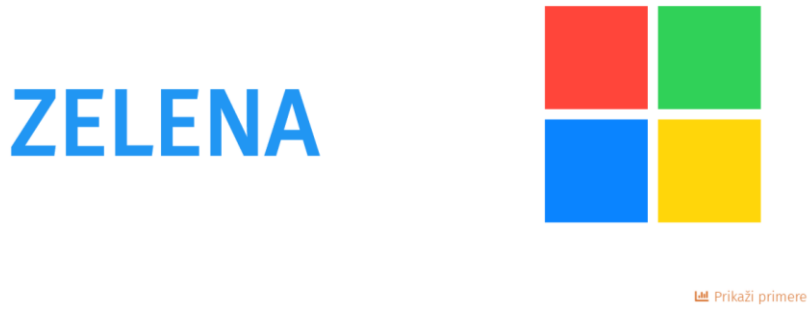
U narednom primeru testiraćemo **postojanje *Strupovog efekta***. Efekat je dobio ime po američkom psihologu Džonu Ridliju Strupu koji je nizom eksperimenata potvrdio postojanje fenomena *semantičke interferencije* (Stroop, 1935). Fenomen se ogleda u tome da (ne)slaganje između značenja reči i boje kojom su reči ispisane, utiče na brzinu njihovog prepoznavanja. Stimulusi koji su podudarni ili *kongruentni*, tj. nazivi boja koji su ispisani bojom koju označava reč, prepoznaju se brže od nepodudarnih ili *nekongruentnih*, kod kojih značenje reči i boja kojom je ona ispisana nisu usklađeni (Slika 41). Kada se reč pojavi na ekranu, vaš zadatak je da što brže kliknete kvadrat obojen bojom koju označava reč. Prvih deset stimulusa, pet podudarnih i pet nepodudarnih, služe kao vežba. Nakon toga potrebno je da uradite 15 + 15 zadataka u kojima će se vreme reakcije beležiti. Nezavisna grupišuća varijabla u ovom primeru je tip stimulusa (podudarni / nepodudarni), a zavisna varijabla je brzina reakcije na stimulus, odnosno brzina prepoznavanja značenja reči. Po završetku eksperimenta, rezultati izraženi u milisekundama biće sortirani od najmanjeg do najvećeg i prikazani u tabeli u donjem delu okvira. Boja ćelija u tabeli označava pripadnost grupama nastalim na osnovu vrednosti nezavisne varijable.

Kao grafički prikaz podataka u primeru sa Strupovim efektom, odabran je *kutijasti dijagram* (engl. *box and whiskers plot*) kojim na veoma pregledan način mogu da se prikažu mere centralne tendencije i raspršenja za jednu ili

više grupa merenja istovremeno (Slika 42). Plavi dijagram prikazuje distribuciju rezultata u grupi podudarnih stimulusa, a narandžasti u grupi nepodudarnih. Središnji kvadratić označava vrednost medijane, kutijom (engl. *box*) je predstavljen interkvartilni raspon između prvog (Q_1) i trećeg (Q_3) kvartila, a „brkovi" (engl. *whiskers*) prikazuju raspon između najmanje i najveće vrednosti, izuzimajući eventualne autlajere. Autlajeri se predstavljaju kružićima ili krstićima pozicioniranim izvan raspona „brkova". Kao što smo rekli, aberantni rezultati mogu negativno da utiču na validnost statističkih zaključaka, posebno ako su podaci prikupljeni na malim uzorcima i ako se analiza obavlja manje robusnim metodama. Ipak, oni mogu da pruže veoma dragocene informacije o postupku merenja varijabli, karakteristikama ispitanika, greškama nastalim u unosu podataka, pa čak i o metodološkim propustima napravljenim od strane istraživača. Stoga ih ne treba zanemarivati ili uklanjati iz analize bez prethodne provere. U programima za statističku obradu postoji mogućnost da se definiše kriterijum na osnovu koga će se rezultati proglasiti aberantnim. Kada je u pitanju univarijantni ili bivarijantni prostor, obično se koriste kriterijumi distance. Tuki, na primer, definiše autlajer kao rezultat koji izlazi van raspona

$$[Q_1 - k(Q_3 - Q_1), Q_3 + k(Q_3 - Q_1)]$$

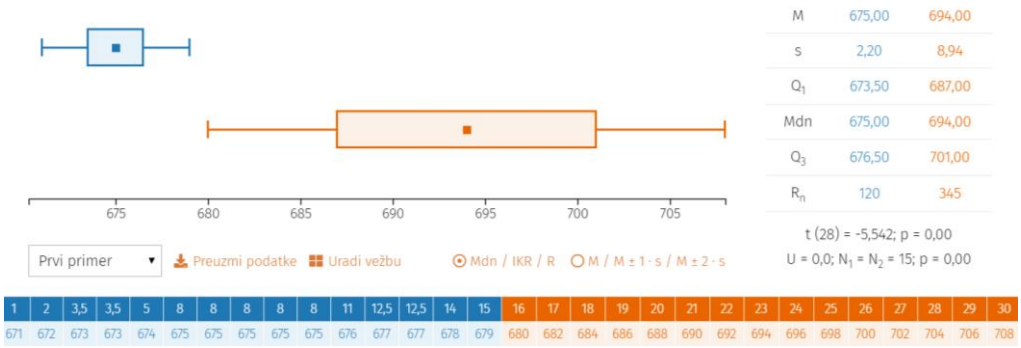
gde su Q_1 i Q_3 vrednosti prvog i trećeg kvartila, a k proizvoljan pozitivan broj. Prema Tukiju, aberantna vrednost je ona koja se nalazi van intervala dobijenog uz vrednost konstante k = 1,5, a ekstremni autlajer je ona vrednost koja izlazi van intervala dobijenog za k = 3 (Tukey, 1977).



Slika 41. Primer nekongruentnog stimulusa u testiranju postojanja Strupovog efekta

Autlajeri mogu da se uoče i na osnovu njihovih (ekstremno) visokih z vrednosti. Jedan od ovakvih postupaka je *Šoveneov kriterijum* koji je predložio francuski matematičar Vilijam Šovene. Na osnovu Šoveneovih kriterijumima, čak

i z vrednosti veće od 2 treba smatrati „čudnim“ ako su uzorci veoma mali. Na većim uzorcima te granice su liberalnije. Tako se, na primer, u skupu od 1.000 merenja, aberantnim smatraju tek oni rezultati čija je z vrednost veća od 3.5. Postupak detekcije autlajera primenom Šoveneovog kriterijuma je iterativan, jer se nakon uklanjanja jednog aberantnog rezultata, ponovo računaju z vrednosti, tako da neki drugi rezultat može da postane autlajer u izmenjenom skupu podataka. Tada kažemo da je prvi autlajer *maskirao* drugog. Moguć je i drugačiji ishod, tj. da se upravo zbog prisustva jednog rezultata, neki drugi smatra aberantnim. Ovo je tzv. *efekat preplavljivanja*. Treba imati na umu da i maskiranje i preplavljivanje mogu da dovedu do isključivanja prevelikog broja rezultata, pa se preporučuje upotreba grafičkih metoda za detekciju rezultata ili skupova rezultata koji će se smatrati aberantnim. U našem primeru, upotrebili smo Tukijev kriterijum, te smo kao aberantne označili vrednosti koje su za više od jednog interkvartilnog raspona veće od vrednosti trećeg kvartila ili manje od vrednosti prvog.



Slika 42. Kutijasti dijagrami koji ukazuju na razliku medijana i interkvartilnih raspona plave i narandžaste grupe merenja

Pomoću grafikona uporedite mere grupisanja, mere raspršenja i oblik distribucije dveju grupa merenja. Da li uočavate aberantne rezultate? Da li se oblici distribucija razlikuju s obzirom na simetričnost i/ili smer zakrivljenosti?

Da li vrednost t-testa, prikazana u tabeli sa desne strane, ukazuje na postojanje statistički značajne razlike u vremenu reakcije na podudarne i nepodudarne stimuluse?

Odaberite prvi primer sa liste. Na osnovu kutijastih dijagrama može se zaključiti da se dve grupe merenja statistički značajno razlikuju, jer se rasponi distribucija čak ni ne dodiruju. Medijana i aritmetička sredina brzine reakcije u grupi podudarnih merenja značajno su manje od onih u grupi nepodudarnih. Pored toga, vidi se da je varijabilnost u drugoj grupi znatno veća. Na postojanje značajne razlike u brzini ukazuje i vrednosti t-testa, kao i postojanje samo dva homogena niza podataka u donjoj tabeli, jer su rezultati iz plave grupe stimulusa rangirani od 1. do 15. mesta, a rezultati iz narandžaste grupe od 16. do 30. Stoga je suma rangova značajno manja u grupi podudarnih stimulusa, nego u grupi nepodudarnih. Sume rangova prikazane su u tabeli sa desne strane i označene su simbolom R_n. Poređenje vrednosti R_n između dve grupe merenja predstavlja osnovnu logiku *testova sume rangova*, a tipičan predstavnik ove grupe metoda je *Men–Vitnijev U test* koji su osmislili američki matematičari Henri Men i Donald Vitni. Test se često naziva i *Vilkokson–Men–Vitnijev test*, jer je statistički ekvivalentnu metodu analize rangiranih podataka predložio i američki hemičar *Frenk Vilkokson*. Postupak podrazumeva da se izvrši $N_1 \cdot N_2$ poređenja svakog rezultata iz jedne grupe sa svakim rezultatom iz druge, te da se prebroje slučajevi u kojima su rezultati iz jedne grupe veći od rezultata iz druge. Brojevi slučajeva računaju se prema formulama:

$$U_1 = N_1 N_2 + \frac{N_1(N_1 + 1)}{2} - R_1$$

$$U_2 = N_1 N_2 + \frac{N_2(N_2 + 1)}{2} - R_2$$

gde su N_1 i N_2 veličine uzoraka a R_1 i R_2 sume rangova rezultata po grupama. Manja od dve U vrednosti predstavlja vrednost Men–Vitnijevog testa, a njegova značajnost očitava se iz odgovarajućih tablica ukoliko su uzorci manji od 20, ili korišćenjem aproksimacije normalne distribucije na osnovu z skora dobijene U vrednosti koja se može izračunati po formuli:

$$z = \frac{U - \frac{N_1 N_2}{2}}{\sqrt{\frac{N_1 N_2 (N + 1)}{12}}}$$

Kao što vidimo u tabeli sa desne strane, u ovom primeru bi i t-test i U test pokazali da je razlika u brzini prepoznavanja značenja reči između podudarnih i nepodudarnih stimulusa statistički značajna. Vrednost U iznosi 0, što znači da

ne postoji nijedan rezultat iz plave grupe koji je veći od nekog rezultata iz narandžaste grupe.

> Na osnovu grafikona može se zaključiti da su obe distribucije simetrične. Uporedite rastojanja između najmanjeg rezultata, prvog kvartila, medijane, trećeg kvartila i najvećeg rezultata. Povežite veličinu ovih raspona sa sirovim rezultatima prikazanim u tabeli. Kako biste na osnovu ovih informacija zaključili koja od dve distribucije je uniformna a koja je približno normalna?

Odaberite drugi primer sa liste. Na osnovu grafikona je uočljivo da su u pitanju dve pozitivno zakrivljene distribucije i da ne postoji razlika u brzini reakcije s obzirom na vrstu stimulusa. Rezultati obe grupe merenja potpuno su identični, tako da su i njihove pozicije izražene *spojenim* ili *vezanim rangovima*. Prva dva rezultata dele 1. i 2. rang, te oba imaju 1,5. rang. Nakon toga slede 3. i 4. rezultat koji, s obzirom na to da su isti, dele 3,5. rang. I tako dalje. Samim tim, sume rangova su identične, kao i dobijene U vrednosti koje su jednake polovini ukupnog broja mogućih poređenja. U našem primeru, to je 15 · 15 ili 225. Obratite pažnju na to da medijana ne mora da se nalazi na polovini interkvartilnog raspona. U ovom primeru, interval kojim je obuhvaćeno prvih 25% rezultata manjih od medijane, veći je od onog u kome se nalazi 25% rezultata najbližih medijani sa desne strane. Dakle, rasponi između Q_1 i Q_2, odnosno Q_2 i Q_3, ne obuhvataju četvrtinu *raspona skale*, već četvrtinu *merenja*. Na kraju, odaberite treći primer sa liste. Ponovo se uočava da su distribucije zakrivljene, plava ulevo a narandžasta udesno. Osim toga, u grupi podudarnih stimulusa prisutan je aberantan rezultat predstavljen kružićem. Obe navedene činjenice smanjuju pouzdanost t-testa kao procene razlike među populacijama. Za razliku od t-testa, U test sugeriše da je razlika među grupama statistički značajna na nivou 0,05, jer se prilikom njegovog računanja uzimaju u obzir samo pozicije rezultata u nizu ali ne i njihove konkretne vrednosti. Možemo da zaključimo da uzorak brzine prepoznavanja podudarnih stimulusa i uzorak brzine prepoznavanja nepodudarnih stimulusa, najverovatnije ne potiču iz iste populacije.

Već smo rekli da se kutijastim dijagramom najčešće prikazuju medijana, interkvartilni raspon i raspon jedne ili više distribucija podataka. Međutim, dimenzije kutije i raspon „brkova" mogu da se odrede i drugim merama centralne tendencije, odnosno raspršenja. Na primer, centralni kvadrat može da

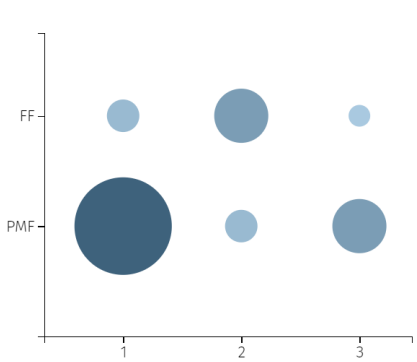
označava aritmetičku sredinu, kutija raspon obuhvaćen intervalom M ± 1 · s, a „brkovi" raspon M ± 2 · s. Rasponi kutije i „brkova" mogu da se definišu i kao intervali pouzdanosti aritmetičkih sredina ako se upotrebi vrednost standardne greške aritmetičke sredine. Naravno, upotreba vrednosti aritmetičke sredine i standardne devijacije prikladna je samo ako se sirovi rezultati mogu približno predstaviti normalnom distribucijom. Ovako dobijeni kutijasti dijagrami uvek će biti simetrični, jer ne prikazuju stvarnu distribuciju rezultata, već onu koja bi se mogla očekivati uz pretpostavku da su raspoređeni normalno. Upotrebite opcije ispod grafikona da biste menjali vrednosti na osnovu kojih se iscrtavaju kutijasti dijagrami i uporedite razlike između opcija Mdn / IKR / R i M / M ± 1 · s / M ± 2 · s za sva tri primera, kao i za podatke koje ste prikupili vežbom.

> U kojim primerima su razlike kutijastih dijagrama iscrtanih korišćenjem različitih mera grupisanja i raspršenja, veće a u kojima manje?
>
> Zbog čega se na kutijastim dijagramima iscrtanim uz opciju M / M ± 1 · s / M ± 2 · s, ne pojavljuju aberantni rezultati?
>
> Koliki procenat rezultata obuhvataju kutija i „brkovi" u slučaju M / IKR / R a koliki u slučaju M / M ± 1 · s / M ± 2 · s? Koja od opcija prikazuje procenat rezultata u različitim delovima opažene distribucije podataka?

3.5. Hi-kvadrat test

U prethodnim odeljcima opisali smo statističke metode kojima se testira značajnost razlika između dve grupe ispitanika na varijablama ordinalnog ili višeg nivoa merenja. Međutim, istraživači često imaju potrebu da porede grupe na kvalitativnim (nominalnim) varijablama ili ordinalnim varijablama sa malim brojem podeoka. U ovakvim situacijama nije prikladna primena parametrijskih testova, pa čak ni primena ranije opisanih neparametrijskih testova, npr. zbog prevelikog broja spojenih rangova. U ovom odeljku opisaćemo *Pirsonov χ^2 (hi-kvadrat) test* kao tipičnu neparametrijsku tehniku kojom se vrši poređenje grupa kada je zavisna varijabla po prirodi kategorijalna. U takvim situacijama je prikladnije, a često i jedino moguće, samo prebrojati slučajeve u okviru svake od kategorija i potom analizirati distribuciju dobijenih opaženih učestalosti. Zamislimo **istraživanje u kome je učestvovalo 60 studenata** sa Filozofskog fakulteta (FF) i isto toliko sa Prirodno-matematičkog (PMF). Cilj nam je da

uporedimo njihovo zadovoljstvo uslugama studentskog restorana izmereno trostepenom skalom: 1 – nezadovoljan/na, 2 – delimično zadovoljan/na i 3 – veoma zadovoljan/na (Slika 43). Ukrštanjem varijabli *Usluga* i *Fakultet* nastaje šest grupa studenata čije su učestalosti prikazane u tabeli kontingencije. Zadebljani brojevi označavaju *empirijske* ili *opažene frekvencije*, tj. veličinu svake od šest grupa koje smo „zatekli" u uzorku. Ove frekvencije obično se označavaju simbolom f_o. Isti podaci prikazani su i grafički uz pomoć *mehurastog dijagrama* (engl. *bubble chart*). Prečnik krugova označava relativnu veličinu grupe nastale ukrštanjem podeoka na x-osi (1, 2, 3) i y-osi (FF, PMF). Očigledno je da bi svi do sada pomenuti statistički testovi ukazali da ne postoje značajne razlike među grupama, jer su distribucije odgovora studenata FF i PMF identične. U obe grupe je broj nezadovoljnih, srednje zadovoljnih i veoma zadovoljnih isti, što rezultira i jednakim vrednostima svih mera centralne tendencije. Aritmetičke sredine i medijane bi u ovom primeru iznosile (20 · 1 + 20 · 2 + 20 · 3) : 60 = 2.



Moji podaci χ^2 (2, 105) = 21,00; p < 0,01

	1	2	3	Σ
FF	10 (16,7)	20 (10,0)	5 (8,3)	35
PMF	40 (33,3)	10 (20,0)	20 (16,7)	70
Σ	50	30	25	105

Preuzmi podatke f_o (f_t) reziduali

Slika 43. Vizualizacija tabele kontingencije primenom mehurastog dijagrama

Za razliku od ranije pomenutih metoda kojima se porede proseci ili rangovi, χ^2 test daje odgovor na pitanje da li se dobijene frekvencije razlikuju od onih koje bi mogle da se očekuju potpuno slučajno, u skladu sa (uslovnom) verovatnoćom svakog ishoda. Ovu logiku detaljnije smo opisali u odeljku 2.5.1. Na primer, ukoliko je u istraživanju učestvovalo 60 studenata FF, a ukupno je 40 studenata nezadovoljno uslugom restorana, u ćeliji FF-1 se očekuje (60 : 120) · (40 : 120) ili približno 0,17, odnosno 17% od 120 studenata, koliko ih je učestvovalo u anketi. Izraženo apsolutnim brojevima, to je upravo onoliko studenata koliko smo i opazili – (60 : 120) · (40 : 120) · 120 = 20. Drugim rečima, opažene frekvencije uopšte se ne razlikuju od *teorijskih*, odnosno *očekivanih*. Teorijske frekvencije označavaju se simbolom f_t a u ćelijama tabele prikazane

su u zagradama. Dakle, u skladu sa opisanom logikom uslovnih verovatnoća, očekivane frekvencije za svaku od ćelija tabele računaju se tako što se proizvod marginalnih frekvencija, odnosno suma odgovarajućeg reda i kolone, podeli veličinom uzorka:

$$f_t = \frac{\sum r \sum k}{N}$$

Nakon toga vrednost χ^2 testa računa se kao suma odstupanja opaženih od očekivanih frekvencija u svakoj ćeliji prema formuli:

$$\chi^2 = \sum \frac{(f_o - f_t)^2}{f_t}$$

Značajnost, tj. p nivo dobijene χ^2 vrednosti, određuje se na osnovu teorijske distribucije opisane u odeljku 2.7.2. Pošto u našem primeru χ^2 iznosi 0, zaključujemo da ne postoji statistički značajna razlika u zadovoljstvu studenata različitih fakulteta.

Sa padajuće liste odaberite primer *Usluga x fakultet 2*. Na prvi pogled, mehurasti dijagram ukazuje na to da se znatno više studenata FF izjasnilo da je zadovoljno uslugama studentskog restorana u poređenju sa studentima PMF. Međutim, uvidom u vrednosti date u tabeli kontingencije, uočavamo da su to upravo one frekvencije koje bismo mogli i da očekujemo. S obzirom na to da je u uzorku bilo više studenata FF nego PMF, ali i da se veći broj studenata oba fakulteta izjasnio pozitivno o uslugama restorana, upravo u ćeliji FF-3 očekuje se najveća frekvencija, tačnije 40 od 105 studenata. Najmanje studenata je trebalo očekivati u kategoriji PMF-1, zato što je u uzorku bilo manje studenata PMF i nezadovoljnih studenata oba fakulteta. Upravo takvo stanje smo opazili. Stoga i u ovom primeru, vrednost χ^2 testa sugeriše da ne postoje statistički značajne razlike među grupama, odnosno distribucijama odgovora studenata FF i PMF. Proporcija 10 od 70 studenata FF u koloni 1, jednaka je proporciji 5 od 35 studenata PMF u istoj koloni.

Pokušajmo sada da simuliramo raspodele koje bi ukazale da među grupama postoje statistički značajne razlike. Veličinu mehura na grafikonu možete da menjate tako što ga kliknete, držite pritisnut taster miša i pomerate pokazivač miša nagore ili nadole. Smanjite opaženu frekvenciju ćelije FF-3 na 10, a opaženu frekvenciju ćelije FF-1 povećajte na 40. Ukupan broj studenata u uzorku ostao je isti, ali smo distribuciju odgovora studenata FF „zarotirali" po

vertikali i tako promenili sume kolona koje govore o učestalosti različitih ocena usluge u ukupnom uzorku studenata. Ova promena uticala je na vrednosti f_t, a time i na sumu razlika opaženih i teorijskih frekvencija. Ćelije u kojima su vrednosti f_o veće od f_t obojene su različitim nijansama zelene boje, a ćelije u kojima je f_o manje od f_t, različitim nijansama crvene. Obratite pažnju na činjenicu da je umanjivanje vrednosti f_o u ćeliji FF-3 uticalo na očekivanu vrednost f_t u toj ćeliji, ali i na vrednost f_t u ćeliji ispod. Na osnovu ovakve distribucije odgovora, zaključujemo da se više studenata FF izjasnilo negativnije u odnosu na ono što bi se očekivalo potpuno slučajno, dok se više studenata PMF izjasnilo pozitivnije u odnosu na ono što bi sugerisale uslovne verovatnoće ćelije PMF-3. Teorijske frekvencije u srednjoj koloni tabele kontingencije ostale su potpuno iste, jer ovom intervencijom nismo izmenili marginalne sume redova, niti marginalnu sumu srednje kolone. Trenutno stanje ukazuje na to da postoji statistički značajna razlika u stavu između studenata FF i PMF, tj. da studenti PMF značajno pozitivnije ocenjuju usluge restorana. U terminima razlika f_o i f_t, ćelije FF-1 i PMF-3 sadrže znatno više studenata od onoga što bi se moglo očekivati slučajno, tj. u situaciji kada razlika u stavovima ne bi postojala. Analogno tome, u grupama FF-3 i PMF-1 opazili smo manje studenata nego što bi se to očekivalo na osnovu proste slučajnosti, tj. u situaciji da je tačna nulta hipoteza o nepostojanju razlike u stavu studenata FF i PMF.

U odeljku 2.7.2. pokazali smo da oblik χ^2 distribucije, pa tako i granične vrednosti χ^2 testa za odgovarajuće nivoe značajnosti, zavise od broja stepeni slobode. Za razliku od t-testa, broj stepeni slobode kod Pirsonovog χ^2 testa ne računa se na osnovu veličine uzorka, već na osnovu veličine kontingencijske tabele. Razlog leži u činjenici da podaci koje obrađujemo χ^2 testom nisu pojedinačna merenja, već frekvencije osnovnih i marginalnih ćelija. Rekli smo da stepeni slobode označavaju broj nezavisnih (slobodnih) rezultata, odnosno ukupan broj vrednosti na osnovu kojih se računa neki pokazatelj, umanjen za broj ograničavajućih faktora. U slučaju standardne devijacije, broj tih vrednosti je N, a broj ograničavajućih faktora je 1, jer u formuli koristimo samo M kao procenu parametra populacije μ. U slučaju χ^2 testa, distribucije suma frekvencija po redovima i kolonama su neka vrsta „standardnih devijacija“, dok je ograničavajući faktor za svaku varijablu donja desna ćelija (ukupan broj ispitanika) jer sume suma redova, odnosno sume suma kolona, moraju da budu jednake vrednosti ove ćelije. To znači da samo r - 1 suma redova i k - 1 suma kolona može potpuno nezavisno da menja svoju vrednost. Suma poslednjeg reda i suma poslednje kolone moraju da dobiju onu vrednost koja će na kraju

dati ukupnu sumu frekvencija. Pošto kontingencijske tabele nastaju ukrštanjem vrednosti dveju kategorijalnih varijabli, samo (r - 1) · (k - 1) *ćelija* može nezavisno da menja svoju vrednost. Tako dolazimo do opšte formule za računanje broja stepeni slobode kod χ^2 testa:

$$df = (r - 1)(k - 1)$$

gde je *r* broj redova, a *k* broj kolona tabele kontingencije.

Odaberite opciju *Grickanje x pol* sa liste. U ovom primeru želimo da proverimo da li je grickanje noktiju (DA / NE) učestalije kod dečaka (M) ili kod devojčica (Ž). Na uzorku veličine 36 i uz prikazanu distribuciju frekvencija po grupama, zaključak je da razlika nije statistički značajna, iako nešto veći broj devojčica gricka nokte. Rezultate analize prikazujemo u formi:

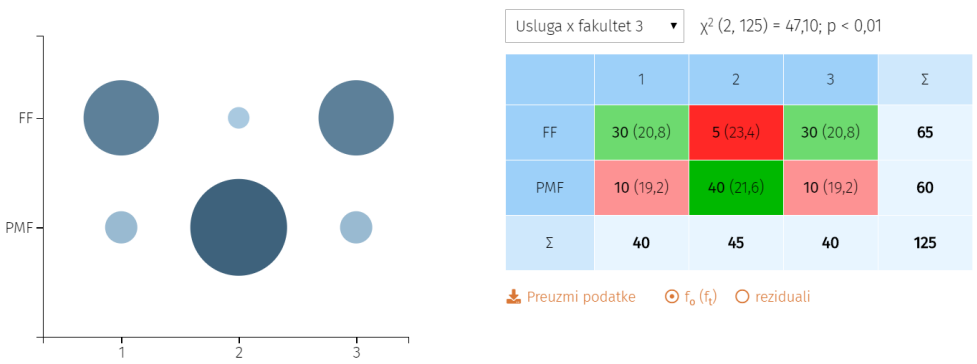
$$\chi^2 (1, 36) = 1{,}68; p = 0{,}19$$

pri čemu se u zagradama navode broj stepeni slobode i ukupna veličina uzorka. Prikazana tabela kontingencije ima samo jedan stepen slobode, jer samo jedna ćelija može da promeni svoju vrednost nezavisno od drugih, a da pri tome marginalne frekvencije i ukupna suma frekvencija ostanu iste. Drugim rečima, ako su nam poznate marginalne frekvencije, dovoljno je da znamo samo učestalost u jednoj od ćelija da bismo izračunali frekvencije svih preostalih ćelija. Ukoliko se vratite na primer *Usluga x fakultet 1*, primetićete da je broj stepeni slobode 2, što znači da možete nasumično da promenite vrednosti dveju ćelija, a sve ostale frekvencije ćete morati da „prilagodite" njima kako bi se dobile iste distribucije marginalnih frekvencija. U slučaju tabele 3 x 3, broj nezavisnih ćelija je 4, u slučaju tabele 3 x 5 je 8 i tako dalje. Prisetite se da, za razliku od t-testa, sa povećanjem broja stepeni slobode rastu i granične vrednosti χ^2 testa koje se smatraju statističkim značajnim. To je i logično, jer se u tabelama kontingencije većih dimenzija očekuju i veće vrednosti suma razlika opaženih i teorijskih frekvencija.

3.5.1. Hi-kvadrat kao test nezavisnosti

Vratimo se ponovo na **primer sa stavovima studenata prema uslugama restorana**. Odaberite opciju *Usluga x fakultet 3* sa liste. Simetrične distribucije odgovora po grupama ukazuju da studenti FF imaju polarizovane

stavove i da je jednak broj onih koji su veoma nezadovoljni i onih koji su izrazito zadovoljni. Sa druge strane, studenti PMF imaju većinom neutralan stav. To znači da su proseci grupa, ali i njihove sume rangova, jednaki. Stoga ni t-test ni Men–Vitnijev test ne bi ukazali na postojanje statistički značajne razlike. Ipak, razlika u distribucijama odgovora očigledno postoji. Upravo zato bi χ^2 test, ali i Kolmogorov–Smirnovljev test koji mu je po logici veoma sličan, ukazali na značajnu razliku *raspodela* odgovora studenata FF i PMF. Pitanje testiranja razlika među centrima i razlika među oblicima distribucija pominjali smo i ranije, a ovo je prilika da podsetimo čitaoca i na činjenicu da u postupku statističkog zaključivanja ne treba izjednačavati tačnost i smislenost. Naime, t-test i Men–Vitnijev test bi *tačno* pokazali da se proseci i medijane odgovora studenata PMF i FF ne razlikuju, jer oni u obe grupe imaju vrednost 2. Sa druge strane, χ^2 test bi ukazao da je to *besmisleno*, jer se *struktura* stavova studenata dva fakulteta značajno razlikuje. Drugim rečima, iako se *prosek* stavova studenata ne razlikuje, postoji značajna razlika između *tipičnih* stavova. Očigledno je da distribucije odgovora studenata FF i PMF ne potiču iz iste populacije odgovora. U ovom slučaju, opravdano je izračunati aritmetičku sredinu ili medijanu ordinalne varijable, ali je potpuno besmisleno koristiti je za izvođenje zaključaka i poređenje grupa (Slika 44).



Slika 44. Distribucije odgovora dve grupe studenata se značajno razlikuju iako su im medijane i aritmetičke sredine jednake

Bitna prednost χ^2 testa u odnosu na ranije pomenute testove kojima se porede grupe merenja, jeste mogućnost da se primeni i u situacijama kada postoji više od dve grupe ispitanika (merenja) i kada su varijable nominalnog nivoa. Odaberite primer *Operater x fakultet*. U analizu uključujemo i treću grupu studenata sa Fakulteta tehničkih nauka, a varijablu koja se ticala stava o

zadovoljstvu uslugama studentskog restorana, menjamo pitanjem o nazivu mobilnog operatera čije usluge student koristi – A, B ili C. Studente, dakle, poredimo na kvalitativnom svojstvu za koje nije moguće izračunati ni prosek, ni medijanu. Vrednost χ^2 testa iznosi 17,58 i ukazuje na postojanje razlika koje su značajne na nivou 0,01. Međutim, sam pojam razlike u kontekstu χ^2 testa treba posmatrati drugačije nego u slučaju t-testa. Za početak, u našem primeru ne možemo da tvrdimo da se sve tri grupe međusobno razlikuju s obzirom na odabranog operatera mobilne telefonije. Grafikon pokazuje da su studenti FF i PMF međusobno sličniji i najčešće biraju operatera A, kao i da se razlikuju od studenata FTN, kod kojih je najzastupljeniji operater C. Pored toga, ne možemo da tvrdimo ni da se preferencije studenata tri fakulteta *potpuno* razlikuju, jer je, na primer, zastupljenost operatera B približno ista u svim grupama. Zbog svega toga χ^2 test ne treba posmatrati kao tipičnu metodu za utvrđivanje statističke značajnosti razlika, već kao *test nezavisnosti* dve varijable. Ukoliko je njegova vrednost mala, distribucija opaženih frekvencija jednaka je, ili veoma slična, distribuciji koja bi se dobila da je raspored frekvencija po grupama potpuno nasumičan. Sa druge strane, značajan χ^2 test ukazuje na određenu pravilnost u promenama na obe varijable, odnosno na *međusobnu povezanost* tj. *korelaciju*. U našem primeru, ovu povezanost možemo da uočimo na osnovu grafikona ali i na osnovu tabele kontingencije. Navešćemo nekoliko primera. Veća je verovatnoća da osoba koja studira na FTN koristi usluge operatera C. Ukoliko neki student koristi usluge operatera A, najveća je verovatnoća da studira na PMF. Ukoliko student koristi usluge operatera A, relativno je mala verovatnoća da studira na FTN. Kao i u slučaju drugih statističkih testova, svaki od ovih zaključaka nosi sa sobom veću ili manju verovatnoću greške, ali ipak ukazuje na određene pravilnosti u prikupljenim podacima i potencijalne relacije među varijablama.

> Da li se broj kolona u matrici sirovih podataka za primer *Operater x fakultet* promenio u odnosu na primere *Usluga x fakultet*?

U statistici se stepen povezanosti varijabli, radi lakše interpretacije, izražava pokazateljima čije se apsolutne vrednosti kreću u intervalu od 0 do 1. Vrednost 0 označava potpunu nezavisnost varijabli, a vrednost 1 njihovu potpunu povezanost, odnosno najveću moguću korelaciju. Najmanja vrednost χ^2 testa koju je moguće dobiti iznosi 0 i ona upućuje na zaključak da nema povezanosti među varijablama. Međutim, njegova maksimalna vrednost zavisi

od veličine uzorka i veličine tabele kontingencije. U tom smislu, χ^2 nije prikladan za iskazivanje stepena povezanosti varijabli, pa se u te svrhe koriste različiti oblici njegovih standardizovanih vrednosti, transformisanih tako da se interval [0, +∞) pretvara u interval [0, 1]. Jedna od takvih standardizacija je *Pirsonov koeficijent kontingencije C* koji se izračunava prema sledećoj formuli:

$$C = \sqrt{\frac{\chi^2}{\chi^2 + N}}$$

U našem poslednjem primeru vrednost C koeficijenta iznosi 0,26, što upućuje na blagu korelaciju studijske grupe i preferencija prema operaterima mobilne telefonije. Kao i većina statističkih testova, vrednosti koeficijenata korelacije imaju svoj p nivo. U slučaju koeficijenata koji se izvode iz χ^2 testa, to je zapravo p nivo χ^2 vrednosti, tako da ćemo u našem primeru reći da je povezanost varijabli statistički značajna na nivou 0,01. Koeficijent C pogodan je za iskazivanje stepena povezanosti dve varijable kada su tabele kontingencije veće, npr. ukoliko imaju više od pet redova i kolona. Koeficijent C ne može da dostigne jediničnu vrednost, ali joj se sve više približava kako se povećava tabela kontingencije (Cohen, 1988).

Većini čitalaca verovatno je poznata osnovna logika i princip korelacije. Mnogo puta ste čuli ili pročitali informacije o povezanosti različitih pojava, najčešće u formi tvrdnji da se sa porastom vrednosti jedne varijable, povećava ili smanjuje vrednost druge. Na primer, rizik od pojave srčanih oboljenja povezan je sa povećanjem telesne mase, potražnja za proizvodom povezana je sa njegovom cenom, uspeh u školi povezan je sa nekim osobinama učenika, i tako dalje. Tabele kontingencije su dobra prilika i povod da čitaocu ukažemo na nedovoljnu preciznost ovakvog shvatanja korelacije. Odaberite sa liste primer *Uspeh x izostanci*. Prikazan je odnos između broja neopravdanih izostanaka i školskog uspeha u grupi od 154 đaka. Prva varijabla može da ima vrednosti 5 (5 ili manje izostanaka), 20 (6–20 izostanaka) i 50 (21–50 izostanaka). Uspeh je izražen kategorijalno kao dobar (3), vrlo dobar (4) i odličan (5). Koeficijent C je značajan i pokazuje da postoji veza između uspeha i učestalosti izostajanja sa časova. Prostije rečeno, đaci koji češće izostaju sa časova, imaju lošiji uspeh. To ne znači da oni imaju lošiji uspeh *zato* što izostaju sa časova, niti da više izostaju sa časova *zato* što im je uspeh loš. To samo znači da između dve varijable postoji značajna korelacija, koja u ovom primeru iznosi 0,56. Na grafikonu se ta korelacija manifestuje kao veća veličina krugova, a u tabeli

kontingencije kao veća razlika u korist opaženih frekvencija u ćelijama 50–3, 20–4 i 5–5. Navedeni parovi vrednosti dve varijable su očigledno povezani, jer sa porastom vrednosti jedne varijable, vrednosti druge opadaju.

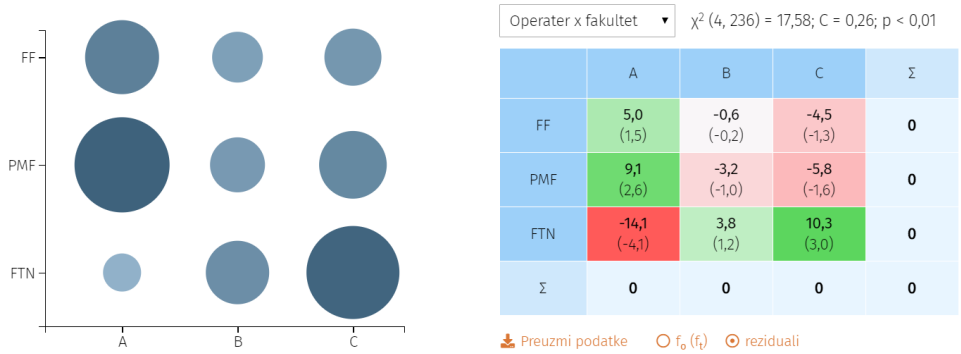
Da bismo dodatno pojasnili logiku povezanosti varijabli, iskoristićemo primer *Anksioznost x uspeh*. Ovoga puta je sa uspehom đaka ukrštena procena stepena njihove anksioznosti od strane psihologa na skali od 1 (ispod proseka), preko 2 (prosečna), do 3 (natprosečna). Korelacija je ponovo značajna, ali se ne može reći da sa porastom vrednosti jedne varijable, opada ili raste vrednost druge. Naime, visok stepen anksioznosti đaka jeste povezan sa nešto lošijim uspehom, ali đaci čiji je stepen anksioznosti ispod proseka češće postižu lošiji uspeh od onih čija je anksioznost u granicama proseka. To znači da je odnos varijabli nelinearan, za razliku od onoga u prethodnom primeru. Dakle, povezanost dve pojave ne mora da implicira njihovu linearnu vezu u smislu zajedničkog porasta ili smanjenja vrednosti. Visok koeficijent korelacije zapravo ukazuje na to da se, sa povećanjem *verovatnoće* nekog ishoda na jednoj varijabli, povećava ili smanjuje *verovatnoća* nekog ishoda na drugoj. O ovome će biti više reči u narednom odeljku.

Ukoliko se χ^2 testom utvrdi postojanje značajne povezanosti između dve varijable, trebalo bi objasniti koja su to konkretno odstupanja, odnosno koje su pojedinačne razlike opaženih i teorijskih frekvencija dovele do visoke vrednosti χ^2. To se postiže uvidom u *reziduale*, odnosno razlike između vrednosti f_o i f_t za svaku ćeliju. Odaberite primer *Operater x fakultet* i prikažite reziduale izborom opcije koja se nalazi ispod tabele kontingencije (Slika 45). Sirovi reziduali nisu naročito informativni, jer zavise od frekvencija u konkretnoj ćeliji. Na primer, razlika od 5 ispitanika između f_o i f_t nema istu težinu i važnost kada je očekivana frekvencija 5 i kada je ona 100. U prvom slučaju, ta razlika je, u relativnom smislu, veća i važnija, jer je opaženo duplo više slučajeva nego što bi se to očekivalo na osnovu proste slučajnosti. Stoga je uobičajeno da se reziduali standardizuju, na primer računanjem *Pirsonovih prilagođenih reziduala* prema formuli:

$$\frac{f_o - f_t}{\sqrt{f_t\left(1-\frac{\sum r}{N}\right)\left(1-\frac{\sum k}{N}\right)}}$$

Dobijene vrednosti reziduala distribuiraju se približno normalno i mogu se interpretirati slično kao z vrednosti. U našem primeru, standardizovani reziduali prikazani su u zagradama i ukazuju da se najveća odstupanja javljaju

u ćelijama FTN-A, FTN-C i PMF-A. Povezanost između vrste studija i izbora mobilnog operatera je statistički značajna, a sastoji se u tome što studenti FTN znatno češće biraju operatera C i znatno ređe operatera A, dok studenti PMF najčešće biraju operatera A. Obično se interpretiraju standardizovani reziduali koji su veći od 2 ili 3, ali treba imati na umu da je verovatnoća da se neki od tako velikih reziduala pojavi potpuno slučajno, veća ukoliko su tabele kontingencije veće (Agresti, 2002), odnosno ako se poredi veći broj grupa na varijabli koja ima više nivoa.



Slika 45. Standardizovani reziduali upućuje na ćelije u kojima su razlike opaženih i teorijskih frekvencija velike i značajne

Odaberite primer *Ispit x udžbenik 1* sa liste. Želimo da proverimo da li postoji veza između vrste udžbenika koji su studenti koristili (E – elektronski, K – klasičan) i prolaznosti na ispitu (1 – položio/la, 0 – nije položilo/la). Korelacija dve dihotomne varijable obično se izražava *ϕ (Fi) koeficijentom* koji se takođe može izračunati na osnovu χ^2 vrednosti:

$$\phi = \sqrt{\frac{\chi^2}{N}}$$

U našem primeru, razlike opaženih i teorijskih distribucija frekvencija ukazuju na postojanje blage korelacije koja nije statistički značajna. Iako se na osnovu gornje formule može zaključiti da vrednost ϕ koeficijenta nikada nije manja od nule, u nekim statističkim programima on može da dobije i negativnu vrednost. Razlog je primena alternativne formule, koju smo upotrebili i u našem primeru:

$$\phi = \frac{ad + bc}{\sqrt{(a + b)(c + d)(a + c)(b + d)}}$$

U gornjoj formuli vrednosti *a* i *b* su opažene frekvencije ćelija prvog reda, a *c* i *d* frekvencije u ćelijama drugog reda, posmatrano sleva na desno. U našem primeru, vrednosti a, b, c i d su 5, 8, 9 i 6. Pozitivna vrednost ϕ koeficijenta znači da su visoki pozitivni reziduali koncentrisani u ćelijama *glavne dijagonale* tabele kontingencije, tj. one koja ide od gornjeg levog ka donjem desnom uglu. Negativna korelacija ukazuje na visoke pozitivne reziduale u *sporednoj dijagonali*, tj. u ćelijama b i c. Istu vrednost koeficijenta korelacije, ali suprotnog smera, postići ćemo ako zamenimo mesta redovima ili kolonama u tabeli. Odaberite primer *Ispit x udžbenik 2* i posmatrajte da li su se i na koji način promenile vrednosti opaženih frekvencija, χ^2 testa i ϕ koeficijenta.

Izmenite opažene frekvencije u svim ćelijama tabele kontingencije za primer *Ispit x udžbenik* tako da se dobije najveća moguća ϕ koeficijenta koja u ovoj vežbi može da bude 0,78 ili -0,78.

Odaberite ponovo primer *Operater x fakultet* i izmenite vrednosti opaženih frekvencija tako da postignete najveću vrednost C koeficijenta koja u našem primeru može da bude 0,70. Koliko takvih rešenja postoji?

Kolika je vrednost koeficijenta C kada su opažene frekvencije u svim ćelijama tabele kontingencije jednake?

3.5.2. Pojam veličine efekta

Odaberite ponovo primer *Ispit x udžbenik 2*. Rezultati sugerišu da postoji blaga povezanost između vrste korišćene literature i uspešnosti na ispitu, jer među studentima koji su položili ispit, ima nešto više onih koji su ga spremali koristeći elektronski udžbenik. Ova razlika, odnosno povezanost, nije statistički značajna. Zamislimo sada da je uzorak studenata bio četiri puta veći, ali da je odnos među veličinama kategorija potpuno isti. Tabelu kontingencije i grafikon za ovaj primer možete da vidite ako odaberete opciju *Ispit x udžbenik 3*. U ovom primeru, kao i u prethodnom, oko 64% studenata koji su koristili elektronski udžbenik položilo je ispit, a približno isti procenat onih koji su koristili klasičan udžbenik nije. Međutim, ovoga puta vrednost χ^2 testa je veća i postala je

značajna na nivou 0,05. Ovaj primer ilustruje činjenicu da p nivo verovatnoće zavisi od dva faktora. Prvi je veličina uzorka a drugi je postojanje uočenog fenomena u populaciji. Dakle, ukoliko neki efekat zaista postoji u populaciji, vrlo je verovatno da ćemo uspeti da ga uočimo čak i na malim reprezentativnim uzorcima. Sa druge strane, kada prikupimo veoma veliki uzorak merenja, postoji rizik da neznatan efekat proglasimo statistički značajnim. U slučaju χ^2 testa, upravo ϕ i C koeficijenti su način da izrazimo tzv. *veličinu efekta* i ublažimo uticaj veličine uzorka na odluku o nultoj hipotezi. U prethodna dva primera, vrednost ϕ koeficijenta je identična, iako p nivoi upućuju na potpuno različite zaključke. Uzimajući u obzir veličinu efekta, čak i u primeru *Ispit x udžbenik 3,* opravdano je doneti zaključak da povezanost, iako statistički značajna, nije i *praktično značajna*, odnosno nije *statistički bitna*, jer je veličina efekta relativno niska. Najčešće korišćene smernice za tumačenje statističke bitnosti, tj. veličine efekta dao je američki psiholog i statističar Džejkob Koen (Cohen, 1988) koji je predložio da se vrednosti oko 0,10 smatraju niskim, oko 0,30 srednjim, a oko 0,50 velikim efektom. Koen je veličinu efekta izrazio kao indeks označen slovom w, a u Tabeli 2 dat je pregled ekvivalentnih vrednosti C i ϕ koeficijenta za različite veličine tabele kontingencije prema Koenu. Oznaka V u prvom redu Tabele 2 je simbol korigovanog ϕ koeficijenta koji je predložio švedski statističar *Harald Kramer* kao korekciju za tabele veće od 2 x 2. Ovaj koeficijent korelacije je poznat kao *Kramerovo V*, a izračunava se prema formuli:

$$\mathrm{V} = \sqrt{\frac{\chi^2}{\mathrm{N}(\mathrm{m}-1)}}$$

gde je *m* manja vrednost od broja redova ili broja kolona. Tako, na primer, Kramerovo V koje iznosi 0,3, a izračunato je na tabeli dimenzija 2 x 4, upućuje na jak efekat uočenog fenomena, tj. razlike ili povezanosti.

Koen je u svojoj knjizi *Statistical power analysis for the behavioral sciences* (Cohen, 1988) opisao načine računanja veličine efekta i za druge popularne statističke testove. Tako se, na primer, kao indeks veličine efekta u slučaju primene t-testa često koristi *Koenovo d*:

$$d = \frac{M_1 - M_2}{s}$$

gde je *s* zajednička standardna devijacija oba uzorka, izračunata prema formuli:

$$s = \sqrt{\frac{\sum(x_1 - M_1)^2 + \sum(x_2 - M_2)^2}{N_1 + N_2 - 2}}$$

Koen preporučuje da se kao referentni indeksi d koji upućuju na mali, srednji i veliki efekat uočene razlike koriste vrednosti 0,2, 0,5 i 0,8. Priručnik za publikovanje *Američke psihološke asocijacije* jasno sugeriše istraživačima da, kad god je to moguće, uz p nivoe navode i odgovarajuće pokazatelje veličine efekta, ali i intervale poverenja izračunatih statističkih testova (APA, 2010).

Tabela 2. Ekvivalente vrednosti C i ϕ koeficijenata za indeks efekta w (Cohen, 1988, str. 222)

w	C	ϕ (V)				
		df = 1	df = 2	df = 3	df = 4	df = 5
0,10	0,100	0,10	0,071	0,058	0,050	0,045
0,20	0,196	0,20	0,141	0,115	0,100	0,089
0,30	0,287	0,30	0,212	0,173	0,150	0,134
0,40	0,371	0,40	0,283	0,231	0,200	0,179
0,50	0,447	0,50	0,354	0,289	0,250	0,224
0,60	0,514	0,60	0,424	0,346	0,300	0,268
0,70	0,573	0,70	0,495	0,404	0,350	0,313
0,80	0,625	0,80	0,566	0,462	0,400	0,358
0,90	0,669	0,90	0,636	0,520	0,450	0,402

3.5.3. Hi-kvadrat kao test stepena poklapanja (distribucija)

U prethodnom odeljku opisali smo primere upotrebe hi-kvadrat testa za poređenje dve ili više grupa merenja i utvrđivanje stepena povezanosti dve varijable. Sličan postupak primenjuje se i kada postoji samo jedna distribucija frekvencija. Tada govorimo o χ^2 testu za jedan uzorak ili *testu stepena poklapanja distribucija* (engl. *goodness-of-fit*). Zamislimo da je cilj istraživanja da se proveri postojanje **razlike u zastupljenosti (broju) nastavnika i nastavnica** u nekoliko srednjih škola. Binomna distribucija opaženih frekvencija prikazana je sa leve strane *uporednog stubičastog dijagrama* i pokazuje da u školama radi nešto više nastavnica (45) nego nastavnika (31). Odgovor na postavljeno pitanje dobićemo ako analiziramo stepen poklapanja dobijene empirijske distribucije

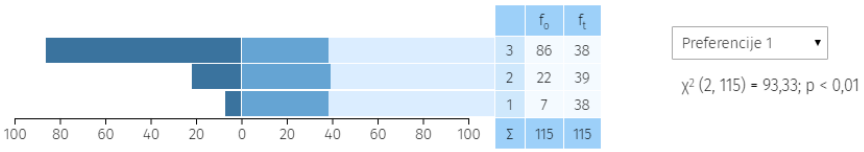
sa pretpostavljenom (teorijskom) distribucijom. Kada ne bi postojala razlika u zastupljenosti polova, očekivali bismo da je verovatnoća svakog ishoda 0,5, tj. da je jedna polovina zaposlenih muškog, a druga polovina ženskog pola. U analiziranim školama zaposleno je 76 nastavnika/ca, tako da su očekivane učestalosti polova 38 i 38. Kada vrednosti f_o i f_t uvrstimo u formulu za računanje χ^2 testa, dobijamo rezultat koji pokazuje da se distribucije u velikoj meri preklapaju, odnosno da njihovo međusobno odstupanje nije toliko da bi se mogle smatrati statistički značajno različitim:

$$\chi2\ (1, 76) = 2{,}58;\ p = 0{,}11$$

U ovom slučaju, broj stepeni slobode je broj kategorija umanjen za jedan, jer tabela ima samo jednu kolonu opaženih frekvencija. Istu polaznu hipotezu možemo da proverimo i na uzorku zaposlenih u nekoliko vrtića, izborom opcije *Vaspitači/ce*. Od 86 zaposlenih vaspitačice čine većinu od 68. Kada opažene frekvencije uporedimo sa teorijskim koje iznose 86 : 2 = 43, dolazimo do zaključka da se dve distribucije ne poklapaju, da su statistički značajno različite, odnosno da se vaspitačkim poslom u vrtićima bavi značajno veći broj žena.

Princip koji smo opisali na dva primera dihotomnih varijabli, može da se primeni na bilo koju distribuciju opaženih frekvencija. Pri tome se može koristiti bilo koja distribucija teorijskih frekvencija. Jedini uslov je da suma očekivanih frekvencija bude jednaka veličini uzorka, odnosno sumi opaženih frekvencija. Primer *Preferencije 1* prikazuje distribuciju ocena usluga studentskog restorana na trostepenoj skali (Slika 46). Dobijena vrednost χ^2 testa ukazuje na to da se distribucija opaženih frekvencija odgovora studenata značajno razlikuje od uniformne, odnosno od pretpostavljene situacije u kojoj bi podjednak broj studenata bio nezadovoljan, uglavnom zadovoljan i veoma zadovoljan uslugama. Zbog činjenice da broj 115 nije deljiv sa 3, vrednosti ćelija kolone f_t nisu iste. Međutim, naša pretpostavka ne mora da bude takva. Potpuno je legitimno da „imamo teoriju“, tj. da očekujemo da većina studenata bude veoma zadovoljna uslugama. U tom slučaju, opaženu distribuciju odgovora poredimo sa drugačijim očekivanim stanjem. Recimo da težimo tome da najmanje 3/4, odnosno oko 75% studenata bude veoma zadovoljno uslugom, a manje od 5% nezadovoljno. Stepen poklapanja dobijenih rezultata sa ovakvom očekivanom raspodelom odgovora prikazan je u primeru *Preferencije 2*. Analiza ukazuje na to da se opažena distribucija stavova uklapa u distribuciju koju želimo da postignemo, jer razlika među njima nije statistički značajna. Redosled odgovora na grafikonu i u tabeli je namerno izmenjen, kako bismo

ukazali na činjenicu da se u slučaju primene χ^2 testa varijable tretiraju kao nominalne, čak i ako su one po prirodi višeg nivoa merenja. Za razliku od Kolmogorov–Smirnovljevog testa, koji se bazira na analizi kumulativnih frekvencija i podrazumeva da rezultati mogu da se rangiraju, kod χ^2 testa je raspored kategorija u tabeli potpuno nebitan, jer se uzima u obzir svaka pojedinačna razlika između f_o i f_t.



Slika 46. Uporedni stubičasti dijagram kojim se ilustruje primena χ^2 testa za proveru stepena poklapanja distribucija

U primeru *Kockica 1* proverićemo ispravnost šestostrane kockice za igru. Leva distribucija ukazuje na potencijalno sumnjiv rezultat, jer je u čak 42 od 180 bacanja dobijena petica, duplo češće nego jedinica. Međutim, kada se opažena distribucija uporedi sa uniformnom teorijskom distribucijom koja se očekuje zato što je verovatnoća svakog od šest mogućih ishoda jednaka, zaključujemo da se opažene i očekivane verovatnoće, prikazane sa leve i desne strane uporednog stubičastog dijagrama, ne razlikuju statistički značajno jer p nivo iznosi 0,09. To znači da su uočena odstupanja mogla da se dogode potpuno slučajno, čak i kod sasvim ispravne kockice. Primer *Kockica 2*, međutim, upućuje na atipičnu distribuciju i potencijalni problem ili grešku u kockici. Broj 6 je očigledno dobijan mnogo češće nego što bi se očekivalo na osnovu zakona verovatnoće. Sa druge strane, veliko odstupanje uočljivo je i kod broja 1 koji je dobijan mnogo ređe. Ovi rezultati ukazuju na potencijalnu povezanost pojedinačnih ishoda ili ispitanika, koju bi uvek trebalo detaljnije obrazložiti prilikom interpretacije rezultata χ^2 testa. U slučaju primera sa kockicom, obrazloženje je veoma jednostavno. Naime, povećanje verovatnoće ishoda 6 nije podjednako uticalo na umanjivanje verovatnoće svih preostalih ishoda, već najviše na verovatnoću ishoda 1, jer se brojevi 6 i 1 nalaze na naspramnim stranama kockice, tako da su njihove verovatnoće obrnuto proporcionalne. Zbog ovakve zavisnosti ishoda, suma razlika f_o od f_t je dodatno povećana, jer su odstupanja na neki način računata dva puta, jednom za ishod 6 i drugi put

za ishod 1. Drugim rečima, vrednost χ^2 testa je veća nego što bi to bio slučaj da su ova dva ishoda bila potpuno nezavisna.

Primer *Visina 1* ilustruje postupak testiranja značajnosti odstupanja distribucije varijable izmerene na razmernom nivou od očekivane normalne distribucije. Vrednosti visine su najpre razvrstane u kategorije, odnosno razrede jednakih intervala. Distribucija opaženih frekvencija nije potpuno simetrična i u određenim delovima odstupa od očekivanog oblika Gausove krive. Analogne frekvencije razreda sa desne strane formirane su na osnovu aritmetičke sredine i standardne devijacije empirijskih rezultata. Za svaki interval vrednosti u centimetrima, izračunat je odgovarajući raspon standardizovanih z vrednosti, a potom i proporcija rezultata koje bi taj raspon trebalo da obuhvati. Na primer, ukoliko je M = 169,11, a s = 8,61, interval 168-171 obuhvata razliku između kumulativne frekvencije rezultata nakupljenih do 171 cm, umanjene za broj rezultata akumuliranih do vrednosti 167 cm. Izraženo u z vrednostima, to je interval od -0,24 do 0,22. Ovaj interval obuhvata približno 18,5% površine normalne krive ili, u našem primeru, $285 \cdot 0{,}185 \approx 53$ ispitanika. Poređenjem ovako dobijenih očekivanih frekvencija sa onima koje su opažene u uzorku, zaključujemo da se dve distribucije ne razlikuju značajno i da empirijsku distribuciju visine možemo da tretiramo kao normalnu. Za razliku od toga, u primeru *Visina 2* zaključujemo da se dobijena bimodalna distribucija visine statistički značajno razlikuje od pretpostavljene normalne distribucije. Vrednost standardne devijacije je relativno velika zbog bimodalnog oblika distribucije, tako da je i očekivana distribucija u većoj meri platikurtična. Zbog toga su očekivane frekvencije rezultata u najnižem i najvišem razredu više nego što bi se očekivalo kod tipične normalne distribucije. Naime, ovi razredi obuhvataju i očekivane frekvencije svih nižih, odnosno viših razreda, koji u distribuciji opaženih frekvencija nisu ni prikazani, jer ne sadrže nijedan rezultat.

3.5.4. Uslovi za primenu hi-kvadrat testa

Kao što smo rekli, χ^2 test spada u grupu neparametrijskih tehnika, tako da su uslovi za njegovu primenu blaži i liberalniji u odnosu na parametrijske metode kao što je t-test. Ipak, postoji nekoliko zahteva koji moraju da budu ispunjeni da bi rezultati χ^2 testa bili pouzdani. Već smo pomenuli da suma opaženih frekvencija uvek mora da bude jednaka sumi teorijskih. Osim toga, kategorije koje su formirane ukrštanjem varijabli moraju da budu iscrpne i

međusobno isključujuće. To znači da svako merenje, odnosno svaki ispitanik, može i mora da se nađe u samo jednoj ćeliji. Treći uslov, koji smo takođe već pomenuli, jeste međusobna nezavisnost opservacija, odnosno merenja. Nijedno merenje ili ispitanik ne sme da bude povezano sa nekim drugim i/ili da utiče na ishod tog drugog merenja. Sledeći važan uslov je da suma opaženih frekvencija svakog reda i svake kolone bude veća od nule. U suprotnom će i teorijske frekvencije nekih ćelija biti nulte, pa vrednost χ^2 neće moći da se izračuna zbog nule u imeniocu formule. Štaviše, poželjno je da očekivane frekvencije u ćelijama ne budu manje od 5. Kod tabela 2 x 2 to je apsolutni uslov, dok se kod većih tabela preporučuje da broj takvih slučajeva ne prelazi 20% ukupnog broja ćelija. Za tabele 2 x 2 takođe se preporučuje upotreba korekcije koja je poznata kao *Jejtsov χ^2 test* ili *χ^2 sa korekcijom za kontinuiranost*. Jejtsov χ^2 računa se tako što se svaka razlika f_o i f_t umanji za 0,5:

$$\chi^2 = \sum \frac{(|f_o - f_t| - 0{,}5)^2}{f_t}$$

Na ovaj način ublažava se problem upotrebe χ^2 vrednosti koje, kao što smo videli u odeljku 2.7.2., potiču iz kontinuirane teorijske distribucije, za analizu oblika diskretnih distribucija. Međutim, s obzirom na to da Jejtsova korekcija obično previše umanjuje vrednost χ^2 testa i tako povećava verovatnoću greške tipa II, pojedini autori preporučuju da se u slučaju tabela kontingencije 2 x 2, umesto Jejtsovog χ^2, koriste druge tehnike, kao što je npr. *Fišerov egzaktni test*, koji je dostupan u većini statističkih paketa (Howell, 2012).

3.6. Pirsonov produkt-moment koeficijent korelacije

U prethodnom odeljku ukratko smo objasnili pojam korelacije na primeru parova kategorijalnih varijabli. Pokazali smo da se pomoću C, ϕ i V koeficijenata može izraziti jačina zajedničkog variranja dve pojave, odnosno stepen u kome je verovatnoća ishoda na jednoj varijabli, povezana sa podjednakom verovatnoćom određenih ishoda na drugoj. Na primer, ukoliko u istraživanju utvrdimo da osobe koje se bave sportom ređe oboljevaju od srčanih oboljenja u odnosu na osobe koje nisu fizički aktivne, možemo da pretpostavimo da su fizička aktivnost i rizik od srčanih oboljenja na neki način povezani. Osim toga, primeri sa koeficijentima korelacije baziranim na χ^2 testu, najbolje ilustruju činjenicu da su pitanja postojanja razlika i postojanja

povezanosti veoma slična i da predstavljaju samo drugačiji pogled na isti istraživački problem. U oba slučaja cilj je da se utvrdi postojanje pravilnosti, odnosno fenomena koji se nisu desili potpuno slučajno i/ili nasumično. Na primer, ukoliko se dečaci i devojčice razlikuju u učestalosti grickanja noktiju, to nam govori da postoji povezanost između pola i grickanja. Međutim, s obzirom na to da se baziraju na χ^2 testu, pomenuti koeficijenti nisu najbolje rešenje ukoliko je potrebno utvrditi stepen povezanosti varijabli intervalnog ili racio nivoa. Tada se obično koristi *Pirsonov produkt–moment koeficijent korelacije*. Uzmimo kao primer **grupu studenata koji su polagali ispit iz nekog predmeta**. Za svakog studenta postoji podatak o polu, broju bodova koje je osvojio na testu i broju sati koje je, prema sopstvenoj proceni, proveo u spremanju ispita. Uz pomoć t-testa mogli bismo da proverimo da li postoji statistički značajna razlika u uspehu ili dužini učenja između studentkinja i studenata. Sa druge strane, jasno je da t-testom nije moguće proveriti da li postoji razlika između uspeha i dužine učenja, jer su varijable izražene u različitim mernim jedinicama i samim tim su njihove aritmetičke sredine neuporedive. Poređenje ne bi bilo moguće čak ni kada bismo vrednosti obe varijable transformisali u z skorove, jer bi tada njihove aritmetičke sredine bile svedene na nulu. Međutim, standardizacija varijabli omogućava donošenje jednog drugačijeg ali podjednako važnog zaključka o ovim dvema varijablama. Ukoliko nakon pretvaranja sirovih bodova u z skorove uočimo da su vrednosti na jednoj varijabli „uparene" sa sličnim vrednostima na drugoj, moći ćemo da zaključimo da su promene na tim varijablama međusobno povezane, tj. da dve varijable u značajnoj meri *kovariraju*.

Uobičajeno je da se kovariranje varijabli vizualizuje uz pomoć *dijagrama raspršenja* ili *skater-dijagrama* (engl. *scatter plot*). Dijagram se sastoji od apscise na kojoj su podeoci skale za varijablu X i ordinate sa podeocima za varijablu Y (Slika 47). U našem primeru, varijabla X je dužina učenja a varijabla Y uspeh na testu. Ispitanici se predstavljaju kružićima ili tačkama raspršenim u tako formiranom dvodimenzionalnom prostoru. Svaka tačka ima svoje koordinate ili *projekcije* na x-osu i y-osu. Klikom na prazan grafikon dodajte kružić čije su koordinate (50, 100), odnosno studenta koji se za ispit pripremao 50 sati i na testu osvojio 100 poena. Da biste lakše obavili zadatak, prikažite linije projekcija tako što ćete držati pritisnut taster *Ctrl* na tastaturi. Koordinate kursora prikazane su u gornjem desnom uglu grafikona. Kao što se vidi iz tabela sa leve i desne strane grafikona, na osnovu jednog para rezultata nije moguće izračunati standardne devijacije varijabli, pa tako ni odgovarajuće z skorove.

Sada na grafikon dodajte studenta koji je nakon 10 sati učenja postigao 20 bodova na testu, tj. kružić sa koordinatama (10, 20). Dodatni par rezultata omogućava računanje varijansi obe varijable prema poznatoj formuli:

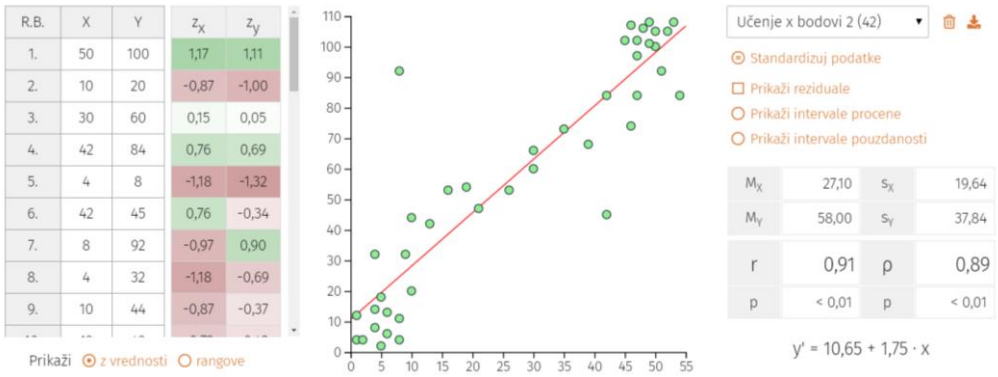
$$s_X^2 = \frac{\sum(x-M_X)^2}{N-1}, \text{ odnosno } s_Y^2 = \frac{\sum(y-M_Y)^2}{N-1}$$

što se može izraziti i na sledeći način:

$$s_X^2 = \frac{\sum(x-M_X)(x-M_X)}{N-1}, \text{ odnosno } s_Y^2 = \frac{\sum(y-M_Y)(y-M_Y)}{N-1}$$

Kombinacijom ovih formula mogli bismo da izračunamo stepen zajedničkog variranja varijabli X i Y, odnosno njihovu *kovarijansu*:

$$s_{XY}^2 = \frac{\sum(x-M_X)(y-M_Y)}{N-1}$$



Slika 47. Dijagram raspršenja koji prikazuje visoku povezanost dužine učenja u satima i broja bodova osvojenih na testu znanja

Vrednost kovarijanse govori o intenzitetu u kome su promene na jednoj varijabli praćene promenama na drugoj. U našem primeru ta vrednost bi trebalo da je visoka, jer je pozitivno odstupanje od prosečne dužine učenja (50 od 30) povezano sa pozitivnim odstupanjem u osvojenom broju bodova (100 od 60), i obratno – negativno odstupanje rezultata na jednoj varijabli (10 od 30) povezano je sa negativnim odstupanjem na drugoj (20 od 60). Prostije rečeno, student koji je duže učio, dobio je više bodova na testu. Jasno je da nam uzorak od samo dva ispitanika ne daje za pravo da donosimo zaključke o celoj populaciji, ali suština je da ova dva para rezultata sugerišu da su dužina učenja i uspeh na testu povezani, tj. da su u korelaciji.

Izražavanje povezanosti dve varijable kovarijansom je nepraktično zato što njena maksimalna vrednost, baš kao i maksimalne vrednosti varijansi, nije ograničena i zavisi od mernih jedinica skala kojima su merene varijable. Da bi se ovaj problem rešio, kovarijansa se standardizuje deljenjem sa standardnim devijacijama obe varijable. Ovaj postupak je potpuno analogan računanju z vrednosti za pojedinačne rezultate. Tako dobijena standardizovana kovarijansa predstavlja pomenuti Pirsonov koeficijent korelacije, koji se označava slovom r:

$$r = \frac{\sum(x - M_X)(y - M_Y)}{(N - 1)s_X s_Y}$$

Standardizovana vrednost kovarijanse uvek se kreće u intervalu od -1 do 1. Postoje i drugačije formule za računanje Pirsonovog r, ali se na osnovu ove jasno vidi da je u pitanju prosek proizvoda (produkata) standardizovanih odstupanja rezultata od aritmetičkih sredina (momenata) varijabli. Otuda i naziv *produkt–moment* koeficijent korelacije:

$$r = \frac{\sum z_x z_y}{N - 1}$$

Odgovarajuće z vrednosti za broj bodova i broj sati prikazane su u kolonama pored tabele sirovih rezultata sa leve strane grafikona. Obratite pažnju na to da su standardizovana odstupanja na varijablama za oba ispitanika potpuno jednaka, kao i da su proizvodi tih odstupanja pozitivni. Samim tim, suma proizvoda z vrednosti po redovima daje vrednost koja je najveća moguća za uzorak veličine 2. Stoga je i vrednost r maksimalna i iznosi +1. Ovakvu korelaciju nazivamo *potpunom*, jer svakom rezultatu na x-osi, odgovara samo jedan rezultat na y-osi. To znači da se odnos dve varijable može predstaviti funkcijom koja je na grafikonu prikazana crvenom pravom:

$$y = b \cdot x$$

gde je *b* koeficijent kojim treba pomnožiti vrednost varijable X, da bi se dobila odgovarajuća vrednost varijable Y. U našem primeru ta vrednost je 2, što znači da studenti dobijaju duplo više bodova od broja sati koje su uložili u pripremu ispita. Štaviše, ukoliko je korelacija dužine učenja i uspeha na testu potpuna, moguće je „predvideti" broj bodova koje će student osvojiti na testu, ako je poznato koliko sati je proveo u spremanju ispita. Na primer, pritisnite taster *Ctrl* i postavite kursor na pravu iznad vrednosti 30 na x-osi. Uočićete da su koordinate te tačke (30, 60), što znači da na osnovu dobijene formule

očekujemo da student koji je učio 30 sati, osvoji 60 bodova na testu. Dodajte ovu tačku na grafikon i pogledajte kako je to uticalo na z vrednosti. Dodavanje ispitanika koji je potpuno prosečan na obe varijable, nije izmenilo aritmetičke sredine varijabli ali je smanjilo standardne devijacije. U skladu sa tim, prethodna dva ispitanika sada značajnije odstupaju od vrednosti M_X i M_Y ali je prosek umnožaka svih odstupanja i dalje 1. Standardizovana odstupanja svakog ispitanika od proseka i dalje su potpuno jednaka na obe varijable, što je vizuelno predstavljeno različitim nijansama zelene i crvene boje. Dodavanje novih ispitanika na liniju možda će izmeniti aritmetičke sredine varijabli i odgovarajuće z vrednosti sirovih rezultata, ali ne i vrednost r, koja će ukazivati da između broja bodova i broja sati postoji potpuna korelacija. Dodajte, na primer, kružiće (studente) čije su koordinate (42, 84) i (4, 8). Ukoliko napravite grešku prilikom određivanja koordinata, kružić možete da uklonite ako ga kliknete dok držite pritisnut taster *Shift* na tastaturi. Svi kružići sa grafikona uklanjaju se klikom na ikonicu kante za otpatke pored padajuće liste.

3.6.1. Regresiona jednačina i regresiona prava

Potpuna povezanost varijabli sreće se veoma retko u svakodnevnom životu. U tom smislu, naš prethodni primer nije realan, jer se ne može očekivati da je vreme provedeno u učenju jedini faktor uspeha na testu. Na primer, studenti koji su redovno pohađali nastavu verovatno će uspešnije savladati ispitne zadatke, čak i ako su na učenje potrošili manje vremena od onih koji nisu bili redovni u obavljanju predispitnih aktivnosti. **Dodajte na grafikon** kružić sa koordinatama (42, 45). Ovaj student je na učenje potrošio isti broj sati kao i student predstavljen kružićem (42, 84), ali ipak nije uspeo da osvoji 84 poena već samo 45. U levoj tabeli se vidi da par z skorova poslednjeg dodatog studenta odstupa od pravilnosti na koju upućuje visoka korelacija, jer je za natprosečnu dužinu učenja osvojio ispodprosečan broj bodova. Obratite pažnju na to da su izmenjene i sve ostale z vrednosti, kao i stepen njihovog slaganja. Zbog svega toga, koeficijent korelacije više nije maksimalan, ali je i dalje veoma visok i iznosi 0,91. Dodajte još jednog studenta čiji podaci odstupaju od pomenute pravilnosti, ovoga puta sa koordinatama (8, 92). Ovaj student je, uz relativno malo utrošenog vremena, ostvario odličan učinak na testu. Posmatrajte kako se promenila vrednost r i usklađenost boja, odnosno z vrednosti u tabeli sa leve strane. Za razliku od parova z vrednosti u prvih 5 slučajeva gde su boje ćelija bile usklađene zbog odstupanja vrednosti u istom

smeru na obe varijable, kod poslednja dva dodata ispitanika boje su upravo suprotne. Pirsonov koeficijent korelacije opao je na 0,57, jer je pravilo „što više sati učenja – to više bodova" sada slabije potkrepljeno empirijskim podacima. Očigledno je da studenti za isti ili sličan broj uloženih sati učenja, mogu da dobiju bitno drugačiji broj bodova na testu, tj. da sa malo uloženog truda mogu da ostvare dobar rezultat i obratno. To znači da korelacija varijabli nije potpuna, ali i da prognoza uspeha na testu ne bi bila dovoljno pouzdana ako bi se bazirala (samo) na podatku o vremenu provedenom u učenju. Očigledno je da kružići više ne mogu da budu raspoređeni duž jedne prave linije, te je ona povučena tako da bude u najvećoj meri „fer" za sve njih. Statistički i geometrijski, „fer" znači da su položaj i pravac linije određeni tako da je udaljenost kružića od nje ista sa gornje i donje strane, odnosno da je suma tih odstupanja nula. Može se reći da je linija u stvari „dvodimenzionalna aritmetička sredina" varijabli X i Y koja aproksimira njihov međusobni odnos. Pošto ona ne mora uvek da prolazi kroz centar koordinatnog sistema, potpunija formula koja je opisuje glasi:

$$y' = a + b \cdot x$$

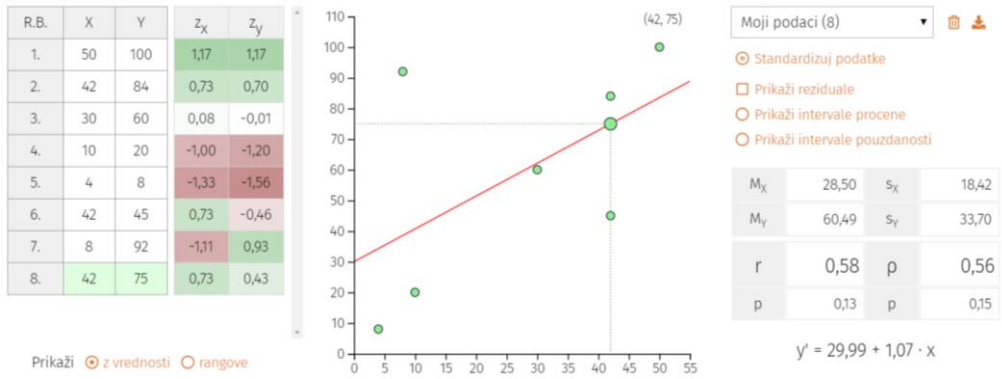
Gornji izraz je poznat kao *regresiona jednačina*, a linija koju ona definiše naziva se *regresiona prava*. Simbol ' u gornjoj formuli čita se „prim" i ne označava empirijsku vrednost y koja je dobijena u istraživanju i koja je bila uparena sa odgovarajućom vrednošću x u tabeli sirovih podataka, već onu koja mogla da se očekuje na osnovu date vrednosti x. Pomoću jednačine ili prave, algebarski ili geometrijski, može da se izračuna vrednost y' ako su poznate vrednosti x, a i b. Postupak kojim se dolazi do regresione jednačine, odnosno modela koji opisuje odnos dveju varijabli, naziva se *jednostavna linearna regresija*. Ona je uobičajena nadogradnja ili nastavak korelacione analize kojom je utvrđen stepen povezanosti dve varijable. U postupku regresione analize, koeficijent *b* se određuje kao odnos kovarijanse varijabli i varijanse varijable X, odnosno:

$$b = \frac{\sum(x - M_X)(y - M_Y)}{\sum(x - M_X)^2}$$

pošto se izraz N - 1 u brojiocu i imeniocu potire. Potom se računa konstanta *a* prema formuli:

$$a = M_Y - b \cdot M_X$$

U primeru sa satima i bodovima vrednosti *a* i *b* prikazane su u regresionoj jednačini ispod tabele sa desne strane grafikona. Ako u jednačinu uvrstimo vrednost 42 kao x, dobijamo prognoziranu vrednost y koja iznosi približno 75 bodova. Do istog zaključka došli biste ako uz pritisnut taster *Ctrl* pronađete y koordinatu tačke koja se nalazi na regresionoj pravoj, a x koordinata joj je 42 (Slika 48). Dobijena vrednost y′ manja je od stvarne vrednosti y za jednog od ispitanika koji je učio 42 sata, ali je veća od vrednosti y za drugog. Stoga se kaže da ona teži, odnosno *regresira* ka proseku varijable Y koji iznosi 58,43 bodova. Frensis Golton je prvi put upotrebio termin *regresija* u ovom kontekstu istražujući naslednost veličine semenki biljaka a potom i naslednost visine ljudi (Galton, 1886). Naime, on je utvrdio da postoji visoka korelacija visine roditelja i visine njihove dece, ali da je prosek visine dece visokih roditelja nešto niži, a prosek visine dece niskih roditelja nešto viši od prosečne visine njihovih roditelja. To znači da obe aritmetičke sredine teže ka prosečnoj visini sve dece, pa je Golton nazvao ovaj fenomen *regresija ka proseku*.



Slika 48. Grafički postupak pronalaženja prognozirane vrednosti y na osnovu date vrednosti x i regresione jednačine, tj. prave

Ekstreman primer regresije ka proseku, prilikom procene vrednosti y, ilustruje situacija u kojoj među varijablama nema nikakve povezanosti. Ovaj slučaj ćemo simulirati polazeći od prethodno formiranog uzorka od 7 studenata u kome je korelacija sati i bodova iznosila 0,57. Primer možete da prikažete ako odaberete opciju *Učenje x bodovi 1* sa padajuće liste. Za početak, dodajte 15 novih ispitanika pozicioniranih duž cele regresione linije, tako da blago odstupaju od nje, ali podjednako sa gornje i donje strane. Ovih 15 parova rezultata trebalo bi da korelaciju povećaju na 0,70 ili više, pošto njihovi odnosi govore u prilog pozitivnoj vezi dve varijable. Međutim, ispitanici koji bi još

značajnije povećali korelaciju zapravo su oni čiji su z skorovi još ekstremniji na obe varijable. Dodajte 10 kružića u krajnji donji levi ugao grafikona, ispod regresione linije, a potom 10 kružića u gornji desni ugao iznad nje. U zavisnosti od toga gde ste pozicionirali kružiće, korelacija bi mogla da se poveća čak i preko 0,90. Raspored kružića koji ste mogli da dobijete prikazan je u primeru *Učenje x bodovi 2*. U kontekstu velikog broja parova rezultata koji prate pravac regresione linije, dva studenta koji više odstupaju od nje, ne utiču bitno na visinu korelacije. Kada bi takvih studenata bilo više, korelacija bi, naravno, bila manja. Klikćite na područje donjeg desnog i gornjeg levog kvadranta grafikona kako biste smanjili korelaciju varijabli sve do vrednosti r = 0,00. Na ovaj način ste dodali parove rezultata čija odstupanja od proseka imaju suprotan predznak, tako da su proizvodi z skorova negativni, a ukupna suma proizvoda, kao i vrednost r, postaju sve manje. Odaberite opciju *Učenje x bodovi 3* da biste videli primer varijabli među kojima nema povezanosti. Pirsonov koeficijent korelacije iznosi 0, te je stoga i prognoza broja bodova na osnovu broja sati potpuno besmislena, jer će regresiona jednačina za bilo koju unetu vrednost x, dati rezultat jednak aritmetičkoj sredini varijable Y.

3.6.1.1. Smisao koeficijenta b i konstante a u regresionoj analizi

Ponovo prikažite sve **primere *Učenje x bodovi***, ali ovoga puta pokušajte da pronađete vezu između položaja regresione prave i vrednosti regresionog koeficijenta *b*. Trebalo bi da uočite da koeficijent *b* u regresionoj jednačini određuje nagib regresione prave, odnosno stepen u kome promene vrednosti y′ prate promene na varijabli X. Ukoliko je taj nagib velik, promene y′ dosledno prate promene na varijabli X. Ukoliko je nagib blag, vrednosti y′ menjaju se neznatno, čak i uz velike promene vrednosti na x-osi. Na kraju, ukoliko nagiba nema, vrednosti y′ se ne menjaju, bez obzira na intenzitet promena vrednosti x. To znači da nagib regresione prave ukazuje na visinu korelacije varijabli – što je on „strmiji", veća je i njihova korelacija. Ipak, prilikom ovakvih interpretacija treba biti veoma obazriv. Odaberite primer *Težina x visina 1* da biste prikazali korelaciju između visine ispitanika u centimetrima i njihove težine (mase) u kilogramima. Nagib regresione prave je relativno blag i naizgled nije u saglasnosti sa visokom korelacijom varijabli koja iznosi 0,80. Zabunu stvara raspon skale y-ose koji nije prilagođen stvarnom rasponu vrednosti y. Primer *Težina x visina 2* prikazuje identične podatke, ali ovoga puta uz drugačiji raspon vrednosti na y-osi. Sada je moguće napraviti grubu procenu visine

koeficijenta r, ne samo na osnovu raspršenja kružića, već i na osnovu nagiba regresione prave. Na sličan način treba posmatrati i vrednost regresionog koeficijenta *b*. Na primer, vrednost b je veća u primeru *Učenje x bodovi 1* nego *Težina x visina 2*, iako je koeficijent korelacije manji. Razlog je ponovo u vrednostima skala varijabli X i Y. U prvom slučaju, pomoću regresione jednačine broj sati se transformiše u broj bodova, a u drugom kilogrami u centimetre. Samim tim, vrednosti koeficijenta *b* nisu uporedive, jer su prilagođene različitim mernim skalama. Međutim, ukoliko se sirovi rezultati standardizuju pretvaranjem u z vrednosti, regresioni koeficijent može da pruži informaciju o tome u kolikoj meri se pomoću varijable X može opisati varijabilnost varijable Y. Odaberite primer *Težina x visina 2* i kliknite taster *Standardizuj podatke* da biste transformisali centimetre i kilograme u odgovarajuće z vrednosti. Iako su se vrednosti na x- i y-osi izmenile, raspored kružića je ostao potpuno isti, kao i koeficijent korelacije r. Promenio se *b* koeficijent koji je takođe standardizovan i sada nam pruža preciznu informaciju o stepenu povezanosti dve varijable. Standardizovano b se u statistici naziva *β (beta) koeficijentom*. U slučaju jednostavne linearne regresije, beta koeficijent jednak je koeficijentu korelacije dveju varijabli. Smisao povezanosti varijabli, odnosno β koeficijenta u regresionoj jednačini, postaće još očigledniji ako kliknete taster *Ujednači ose*, čime se izjednačavaju rasponi vrednosti x- i y-osa i olakšava njihovo poređenje. Uočite da vrednost β koeficijenta, odnosno nagib regresione prave, pokazuje za koliko će se standardnih jedinica promeniti vrednost Y varijable, ako se X varijabla promeni za jednu standardnu jedinicu. U poslednjem primeru sa učenjem i bodovima, taj odnos je 0,8. Kada bi korelacija varijabli bila potpuna, ta vrednost bi bila 1. Ako korelacija ne postoji, vrednost β koeficijenta biće 0, jer promene vrednosti varijable X ni na koji način nisu povezane sa promenama na varijabli Y.

> Da li biste na osnovu regresione jednačine $y' = 8{,}26 + 3{,}42 \cdot x$ mogli da zaključite kolika je približno korelacija varijabli X i Y?
>
> Da li biste na osnovu regresione jednačine $y' = 0 + 7{,}65 \cdot x$ mogli da zaključite da li je ona dobijena na osnovu sirovih ili standardizovanih podataka?

Iz prethodnih primera mogli smo da uočimo da linearna transformacija varijable, kao što je npr. standardizacija, ne menja njenu korelaciju sa drugim varijablama. Ilustrovaćemo tu činjenicu još jednim primerom. Odaberite opciju

Prijemni x ESPB 1. U pitanju je odnos broja bodova koje su studenti osvojili na prijemnom ispitu za upis na fakultet i uspešnosti u studiranju izraženoj sumom osvojenih ESPB kredita u toku prve dve godine studija. Na osnovu dijagrama raspršenja, koeficijenta r i p nivoa zaključujemo da je korelacija varijabli veoma niska i beznačajna što znači da studenti koji su ostvarili bolji uspeh na prijemnom ispitu nisu nužno uspešniji u studiranju u toku prve dve godine studija. Zamislimo sada situaciju u kojoj je svim studentima, zbog promene plana i programa, kurs iz stranog jezika koji nosi 3 ESPB priznat kao položen. To znači da su svi studenti jednoobrazno popravili svoj uspeh. Drugim rečima, vrednosti varijable Y su pravolinijski transformisane dodavanjem vrednosti 3. Odgovor na pitanje da li je to uticalo na koeficijent korelacije varijabli, dobićete ako sa liste odaberete primer *Prijemni x ESPB 2*. Svi kružići su pomereni nagore, zajedno sa regresionom pravom, što nije uticalo na koeficijent korelacije r. Nije se izmenio ni regresioni koeficijent *b* jer je on vezan za vrednosti varijable X. Međutim, ova linearna transformacija uticala je na vrednost aritmetičke sredine varijable Y, pa tako i na vrednost regresione konstante *a*. To znači da konstanta *a* zapravo predstavlja odsečak na y-osi, odnosno vrednost koju će varijabla Y imati kada je vrednost varijable X nula. Stoga se za nju u engleskom jeziku, a često kao tuđica i u srpskom, koristi termin *intercept*. Kao što smo videli iz ranijih primera (npr. *Učenje x bodovi 3*), ako korelacija varijabli ne postoji, *a* će biti jednako vrednosti aritmetičke sredine varijable Y. Drugim rečima, bez obzira na promene na varijabli X, očekivana vrednost varijable Y biće uvek ista. Slična stvar bi se desila i nakon transformacije vrednosti varijable X. Odaberite ponovo primer *Težina x visina 2*. Težina ispitanika izražena je u kilogramima, ali ne bi bio nikakav problem da smo je izrazili u nekim drugim jedinicama, npr. funtama. Štaviše, tu jednostavnu linearnu transformaciju možemo da obavimo i naknadno, tako što ćemo svaku vrednost x pomnožiti sa 2,2046262. Rezultati ove transformacije prikazani su u primeru *Težina x visina 3*. Vrednost konstante *a* ostala je ista jer se prosek varijable Y nije promenio. Međutim, vrednost *b* postala je manja, jer je se sada procenjuju vrednosti izražene u stotinama (centimetara) na osnovu stotina (funti) a ne na osnovu desetina (kilograma). Međutim, suština je da se koeficijent korelacije varijabli nije promenio, upravo zato što linearna transformacija varijable ne menja njen odnos i stepen povezanosti sa drugim varijablama. To važi i za standardizaciju pretvaranjem u z vrednosti. Tada nagib regresione prave postaje jednak koeficijentu korelacije a odsečak postaje nula, jer prava uvek prolazi kroz centar (0, 0) koordinatnog sistema.

Da li bi se korelacija dve varijable promenila kada bi se vrednosti svake od njih transformisale u percentilne rangove?

Da li na osnovu boja ćelija u tabeli sa leve strane možete da odredite gde se na grafikonu nalazi određeni kružić, odnosno ispitanik?

Primeri *Učenje x bodovi 3* i *Prijemni x ESPB 1* ukazuju na nisku ili nepostojeću povezanost varijabli. Šta biste mogli da zaključite o prirodi te (niske) korelacije varijabli?

3.6.2. Standardna greška procene

Odaberite **primer *Pušenje x kapacitet 1***. Prikazan je odnos „pušačkog staža" i vitalnog kapaciteta pluća petoro ispitanika. Na x-osi navedena je dužina upotrebe cigareta izražena u godinama, a na y-osi kapacitet pluća izražen u litrima. Očekivano je da ta korelacija bude negativna, što pokazuje da osobe koje duže konzumiraju cigarete, imaju manji kapacitet pluća, odnosno da sa produženjem „pušačkog staža" dolazi do smanjenja kapaciteta pluća. Iz tabele sa leve strane možete da vidite da su umnošci z vrednosti negativni, jer su pozitivna odstupanja od proseka na jednoj varijabli povezana sa negativnim odstupanjima na drugoj. U regresionoj jednačini ova pravilnost odrazila se na predznak koeficijenta *b* koji je negativan. Pošto se korelacija od -1 smatra potpunom, prognoza vrednosti varijable Y na osnovu vrednosti varijable X bila bi potpuno tačna, kao u primerima u kojima je koeficijent r bio 1. Dodajte ispitanika koji odstupa od uočene pravilnosti, npr. sa približnim koordinatama (11, 3,50). Korelacija postaje nešto niža, pa je i greška prognoze veća. Na to ukazuje veće raspršenje kružića i njihovo veće odstupanje od regresione prave. Kliknite taster *Prikaži reziduale* da biste grafički prikazali reziduale, odnosno odstupanja empirijskih vrednosti y od onih koje bi se očekivale na osnovu dobijene regresione jednačine, a koje smo ranije označili simbolom y'. Kao što smo rekli, regresiona prava mogla bi da se shvati kao dvodimenzionalna aritmetička sredina varijabli X i Y. U tom smislu, pomoću reziduala može da se izračuna „dvodimenzionalna standardna devijacija" koja govori o tome kolike su razlike opaženih i očekivanih vrednosti varijable Y. Pošto je regresiona prava povučena tako da je suma odstupanja svih vrednosti od nje uvek nula, uobičajeno je da se reziduali pre sabiranja kvadriraju. Tako se dobija vrednost

koja se u statistici obično označava sa *SS* od engleskog *sum of squares*, slično kao i u postupku računanja varijanse. Ovoga puta u pitanju je *suma kvadriranih reziduala* koja se koristi kao procena *sume kvadriranih grešaka* regresije (engl. *sum of squared errors*):

$$SS_e = \sum (y - y')^2$$

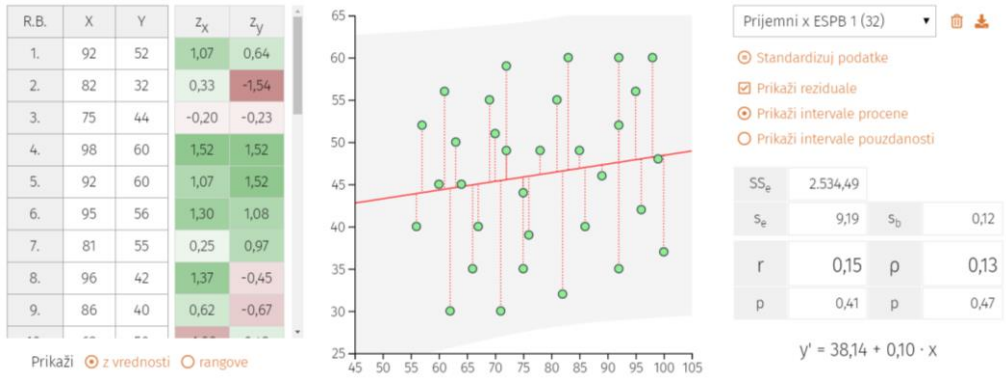
Pređite pokazivačem miša preko kružića da biste prikazali površinu kojom se jasnije vizualizuju razlike između dobijenih i prognoziranih vrednosti varijable Y. Ukoliko se kružić nalazi iznad prave, projekciju dobijene vrednosti prikazuje gornja stranica pravougaonika, a projekciju prognozirane vrednosti njegova donja stranica. Ako se kružić nalazi ispod regresione prave, projekciju vrednosti y prikazuje donja stranica, a vrednosti y′ gornja stranica pravougaonika. Kao što se vidi u tabeli sa desne strane, suma kadriranih reziduala SS_e relativno je mala. Međutim, ako dodate nekoliko ispitanika u gornji desni i donji levi kvadrant dijagrama, apsolutna vrednost koeficijenta korelacije se smanjuje, a vrednost SS_e povećava. Ukoliko nastavite da dodajete ispitanike u ove delove grafikona, smer korelacije će se promeniti, a vrednost r će postajati sve viša. Naravno, suma kvadrata odstupanja je veća u slučaju većeg broja rezultata, pa je uobičajeno da se ona deli veličinom uzorka kako bi se dobila objektivnija mera greške. Tako nastaje pokazatelj koji se naziva *prosečna kvadrirana greška*, a označava se sa MS_e od engleskog *mean squared error*:

$$MS_e = \frac{\sum (y - y')^2}{N}$$

Pošto se na ovaj način, kao i u slučaju varijanse, dobijaju kvadrirane jedinice (npr. l^2, cm^2 ili $ESPB^2$), potrebno je izračunati koren gornjeg izraza. Dodatno, s obzirom na to da se MS_e računa na uzorku a ne na celoj populaciji, u imeniocu ove formule se ne koristi vrednost N, već vrednost df. Više puta smo ponovili da se broj stepeni slobode obično računa na osnovu veličine uzorka koja je umanjena za broj procena parametara upotrebljenih u formuli. U našem slučaju, da bismo izračunali vrednost y′ bilo je potrebno proceniti vrednost parametara a i b. Stoga formula za izračunavanje pokazatelja koji se naziva *standardna greška regresije* ili *standardna greška procene*, glasi:

$$s_e = \sqrt{\frac{\sum (y - y')^2}{N - 2}}$$

Iz oznake pokazatelja (s) zaključujemo da je ponovo u pitanju standardna devijacija, ovoga puta standardna devijacija distribucije grešaka procene. Njena vrednost je približna prosečnoj dužini duži koje su na grafikonu prikazane crvenim tačkastim linijama. Ukoliko su ispunjeni uslovi, o kojima će biti više reči u nastavku teksta, greške procene distribuiraju se u skladu sa t raspodelom, tako da se vrednost s_e može upotrebiti za računanje tzv. *intervala procene*. Na primer, na osnovu dovoljno velikog uzorka merenja možemo da očekujemo da će se za neku vrednost x, odgovarajuća vrednost y u 95% slučajeva naći u intervalu $y' \pm 1{,}96 \cdot s_{ey'}$. Ipak, treba napomenuti da se za donošenje ovakvih zaključaka ne koristi uobičajena, već korigovana vrednost standardne greške procene koju nismo označili sa s_e već sa $s_{ey'}$.



Slika 49. Primer dijagrama raspršenja sa prikazanim rezidualima (isprekidane linije) i intervalima procene (siva površina)

Za čitaoca nije nužno da razume postupak kojim se standardna greška procene koriguje da bi se dobili odgovarajući intervali procene, ali je potrebno naglasiti da je korekcija neophodna zato što veličina greške nije ista u svim regionima x-ose. Naime, greška procene je manja u zoni prosečnih vrednosti, a veća ukoliko se vrednost y procenjuje na osnovu x koje znatno odstupa od aritmetičke sredine. Odaberite primer *Prijemni x ESPB 1* i prikažite intervale procene izborom opcije *Prikaži intervale procene* (Slika 49). Uočite da ivice sive površine, koja prikazuje intervale procene, nisu paralelne sa regresionom pravom, već su u njenom centralnom delu blago ugnuti ka unutra. Ovaj fenomen biće još uočljiviji kada se umesto intervala procene vizualizuju tzv. *intervali pouzdanosti* kojim se, po istom principu kao i u prethodnom primeru, procenjuje raspon proseka svih potencijalnih vrednosti y za određeno x. Ovaj interval označićemo sa $y' \pm 1{,}96 \cdot s_{eM}$, a možete da ga vidite ako odaberete

opciju *Prikaži intervale pouzdanosti*. Nešto drugačijom korekcijom standardne greške procene, za svako x moguće je izračunati interval u kome se, sa određenim stepenom sigurnosti, nalazi aritmetička sredina svih potencijalnih vrednosti y. Drugim rečima, u pitanju je uslovna verovatnoća dobijanja određenog proseka y vrednosti za dato x. Razlika između intervala procene i intervala pouzdanosti analogna je razlici između intervala M ± s i M ± s_M, koju smo pojasnili u poglavlju o standardnoj grešci aritmetičke sredine.

Na kraju treba napomenuti da je na osnovu modifikovanih standardnih grešaka procene moguće testirati statističku značajnost konstante a i koeficijenta b i izračunati njihove intervale pouzdanosti. Posebno je korisna standardna greška koeficijenta b, pa ćemo ukratko opisati logiku njene upotrebe. Ona je u tabeli sa desne strane označena simbolom s_b, a vidljiva je kada su na grafikonu prikazani reziduali i/ili intervali. Odaberite ponovo primer *Učenje x bodovi 1* i prikažite reziduale i/ili intervale. Korelacija je umereno visoka i iznosi 0,57, b koeficijent je 1,07, a njegova standardna greška iznosi 0,69. Pošto se odnosi koeficijenta b i njegove greške distribuiraju u skladu sa t raspodelom, uz pomoć formule:

$$t = \frac{b}{s_b}$$

možemo da izračunamo odgovarajuću vrednost t i interpretiramo je kao bilo koji drugi t-odnos. U našem primeru, vrednost t-testa iznosila bi 1,07 : 0,69 = 1,55 što je nedovoljno da bi se koefcijent b, a samim tim i korelacija varijabli, smatrali statistički značajnim. Po sličnom principu možemo da izračunamo i interval pouzdanosti regresionog koeficijenta. Granična vrednost t-testa za 5 stepeni slobode i nivo značajnosti 0,01 iznosi 4,03, tako da možemo da budemo 99% sigurni da bi vrednost b koeficijenta u populaciji bila negde između 1,07 - 4,03 · 0,69 i 1,07 + 4,03 · 0,69. Pošto ovaj interval obuhvata vrednost 0, ne smemo da tvrdimo da je b značajno drugačije od nulte vrednosti. Samim tim, ni koeficijent korelacije dobijen u ovom primeru ne možemo da prihvatimo kao opravdanje za prognozu broja osvojenih bodova na osnovu broja sati koje je student proveo u učenju. Obratite pažnju na to da broj stepeni slobode iznosi 5 jer smo imali uzorak veličine 7, a prilikom računanja koeficijenta b koriste se procene dva parametra – aritmetičke sredine varijable X i aritmetičke sredine varijable Y.

Zbog čega su u regresionoj analizi intervali procene uvek širi od odgovarajućih intervala pouzdanosti?

Da li je u primeru *Učenje x bodovi 2* koeficijent b značajno različit od nule?

Zbog čega intervali procene u primeru *Učenje x bodovi 2* ne obuhvataju sve prikazane rezultate već se dva kružića nalaze izvan njega?

3.6.3. Interpretacija koeficijenta korelacije

Interpretacija koeficijenta korelacije podrazumeva tumačenje njegove apsolutne vrednosti, predznaka, statističke značajnosti i praktične značajnosti. Ukratko ćemo objasniti svaki od ovih elemenata. U vezi sa interpretacijom apsolutne vrednosti, ne može se reći da u literaturi postoji konsenzus o tome koji stepen korelacije bi trebalo smatrati (dovoljno) visokim. U Tabeli 3 date su okvirne preporuke za interpretaciju visine koeficijenta korelacije koje su dali različiti autori. Delimična neslaganja, pogotovo u zoni „umerenih" vrednosti, ukazuju na to da procena intenziteta korelacije u velikoj meri zavisi od percepcije i iskustva istraživača, konteksta istraživanja, varijabli koje se analiziraju, pa čak i od oblasti u kojoj se istraživanje sprovodi. U prirodnim naukama uobičajeno je da se fenomeni opisuju zakonima koji podrazumevaju veoma jaku korelaciju varijabli. *Šarlov zakon*, na primer, definiše potpunu linearnu povezanost između temperature i zapremine gasa. Sa druge strane, zakonitosti u društvenim naukama su veoma retke, a fenomeni kao što je ljudsko ponašanje često su nepredvidivi ili ih je potrebno opisati pomoću većeg broja varijabli. Stoga su i visoke korelacije varijabli mnogo ređe nego u „tvrdim" naukama. Već smo pomenuli da se u psihologiji često koriste smernice koje je dao Džejkob Koen (Cohen, 1988), prema kojima se koeficijenti korelacije od 0,30 mogu smatrati umerenim, a već oni od 0,50 jakim. Koen polazi od logike da korelaciju treba interpretirati u skladu sa kontekstom, odnosno sa empirijskim rezultatima koji upućuju na maksimalne vrednosti koeficijenata korelacije koje je opravdano očekivati. To su, prema Koenu, rezultati istraživanja iz oblasti pedagoške psihologije u kojima su utvrđene korelacije od 0,50 između uspeha na testovima sposobnosti i školskog postignuća. Međutim, suština je da vrednost koeficijenta korelacije uvek treba interpretirati u skladu sa logičnim objašnjenjima, sličnim rezultatima prethodnih istraživanja, ali i posledicama

koje bi zaključci mogli da proizvedu u budućnosti. Korelacija od 0,30, koja se po nekima smatra niskom, nekada može da ukaže na postojanje fenomena koji bi trebalo dodatno istražiti. Isto tako, koeficijent korelacije od 0,50 između rezultata dva upitnika, kojima se navodno meri ista dimenzija ličnosti, trebalo bi smatrati niskim i proveriti da li instrumenti zaista imaju isti predmet merenja.

Tabela 3. Smernice za interpretaciju visine koeficijenta korelacije

Interpretacija korelacije	Vrednost koeficijenta korelacije		
	(Guilford, 1978)	(Evans, 1996)	(Hinkle, Wiersma, & Jurs, 2003)
veoma slaba	0,00–0,20	0,00–0,20	0,00–0,30
slaba	0,21–0,40	0,21–0,40	0,31–0,50
umerena	0,41–0,70	0,41–0,60	0,51–0,70
jaka	0,71–0,90	0,61–0,80	0,71–0,90
veoma jaka	0,91–1,00	0,81–1,00	0,91–1,00

Rasponi koeficijenata korelacije navedeni u Tabeli 3 odnose se na njihove apsolutne vrednosti. To znači da se, na primer, koeficijenti 0,83 i -0,83 mogu smatrati jednako visokim, bez obzira na predznak. Predznak se, naravno, ne sme zanemariti, jer nam govori o prirodi odnosa dve varijable. U slučaju obrnute korelacije, kao što je bio slučaj sa dužinom pušenja i kapacitetom pluća, sa porastom vrednosti na jednoj varijabli, vrednosti druge varijable se smanjuju. Sličnih primera u psihologiji i srodnim naukama ima puno. Viša inteligencija je povezana sa nižom aktivacijom kore velikog mozga prilikom rešavanja kompleksnih zadataka (Neubauer & Fink, 2009). Neke osobine ličnosti, kao npr. Savesnost i Negativna valenca, takođe su u umerenoj negativnoj korelaciji (Čolović, Smederevac, & Mitrović, 2014). Utvrđena je i obrnuta veza između agresivnosti i davanja socijalno poželjnih odgovora na upitnicima ličnosti (Banse, Messer, & Fischer, 2015). Ekstraverzija je u negativnoj korelaciji sa vremenom provedenim u korišćenju interneta (Landers & Lounsbury, 2006). Učestalost agresivnog ponašanja slabijeg intenziteta kod primata povezana je sa nižim nivoima hormona kortizola u krvi (Westergaard et al., 2003). I tako dalje. U istraživanjima u kojima se koriste upitnici, uvek treba obratiti pažnju na to da li su stavke usmerene na očekivan način, jer bi u suprotnom to moglo da

dovede do nelogičnih negativnih korelacija. Na primer, stavke „U društvu sam ćutljiv" i „Sa mnom svi vole da se druže" mere ekstravertnost, ali je pre sabiranja odgovora, prvi od njih potrebno rekodirati, tj. visoke vrednosti zameniti niskim i obratno, npr. 5 sa 1, 4 sa 2, 2 sa 4 i 1 sa 5.

Pirsonov koeficijent korelacije koji je izračunat na uzorku, predstavlja samo procenu stvarne veze među varijablama u populaciji. U tom smislu, koeficijent r može da se koristiti za testiranje nulte hipoteze da korelacija varijabli u populaciji ne postoji. Kao i kod drugih statističkih testova o kojima smo pisali, odluku o (ne)odbacivanju ove hipoteze donosimo na osnovu odgovarajućeg p nivoa. Koeficijenti korelacije velikih uzoraka uzetih iz populacije u kojoj je korelacija nulta, distribuiraju se oko nule u skladu sa t-raspodelom. Stoga se za testiranje statističke značajnosti koeficijenta korelacije r koristi t-test koji se računa prema formuli:

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

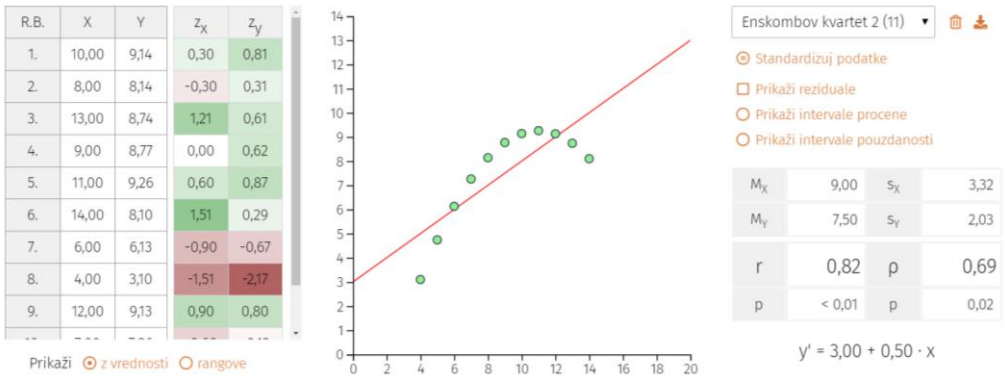
Gornja formula ukazuje na potencijalne opasnosti prilikom interpretacije koeficijenta korelacije. Jasno je da će t vrednosti biti veća ukoliko je r veće, pošto sa porastom vrednosti r, vrednost brojioca u formuli postaje sve veća, a vrednost imenioca sve manja. Međutim, u imeniocu se nalazi i veličina uzorka, što znači da čak i relativno niski koeficijenti korelacije mogu da budu postanu statistički značajni kada su izračunati na veoma velikim uzorcima. U primeru *Učenje x bodovi 1* koeficijent korelacije je relativno visok, ali nije statistički značajan, jer je izračunat na osnovu samo 7 parova rezultata. Broj stepeni slobode za izračunati t-test je N-2, jer smo prilikom računanja koeficijenta r upotrebili dve procene parametara – prosek varijable X i prosek varijable Y. Prikažite primer *Instagram x IQ* koji prikazuje povezanost broja sati koje 211 osoba dnevno provede u korišćenju Instagrama (x-osa), sa njihovim rezultatima na testu sposobnosti (y-osa). Koeficijent korelacije je statistički značajan jer je izračunat na velikom uzorku ispitanika, ali njegova vrednost, kao i raspršenje kružića na skater-dijagramu, ukazuju da bi zaključak da su ove dve varijable značajno povezane, bio besmislen. Gornja formula najbolje ilustruje ranije pomenutu pravilnost da je statistička značajnost proizvod veličine stvarnog efekta (r) i veličine uzorka (N-2), tj. da se jedan aspekt značajnosti može kompenzovati drugim. Ali to ne znači da takvu kompenzaciju treba primenjivati, jer je za istraživača mnogo važnije da pokaže stvarni efekat dobijenog rezultata. U tom smislu, u statistici se prilikom interpretacije koeficijenta korelacije koristi

i kvadrat njegove vrednosti koji se naziva *koeficijent determinacije*. Koeficijent determinacije pokazuje kolika proporcija varijanse jedne varijable može da se objasni promenama na drugoj varijabli. U našem poslednjem primeru, iako je r označen kao statistički značajan, njegova vrednost nam govori da tek nešto manje od 4% ($0{,}19^2$ = 0,036) varijabilnosti intelektualnih sposobnosti može da se objasni vremenom provedenim na Instagramu ili obratno. To je, naravno, neprihvatljivo malo i ovaj koeficijent ne možemo smatrati *praktično* značajnim. To znači da čak i u slučaju relativno visokih koeficijenata korelacije, istraživač treba da se zapita koje to druge varijable opisuju varijablu Y, ako je proporcija zajedničke varijanse relativno mala. Na primer, čak i koeficijent od 0,70 ukazuje na to da varijable imaju manje od 50% zajedničke varijanse i da bi u nacrt istraživanja trebalo uključiti još neke varijable da bi se umanjila proporcija neobjašnjene varijanse reziduala koja preostaje kada na osnovu vrednosti varijable X pokušamo da predvidimo rezultat na varijabli Y.

3.6.4. Uslovi za primenu Pirsonovog r

Činjenica da se Pirsonov koeficijent korelacije bazira na z vrednostima i poređenju odstupanja pojedinačnih rezultata od aritmetičkih sredina, svrstava ovu metodu u grupu parametrijskih testova. To znači da bi pre izračunavanja i interpretacije vrednosti r, trebalo proveriti ispunjenost uslova koji su pomenuti i u slučaju t-testa: intervalni ili racio nivo merenja varijabli, normalnost raspodele obe varijable i homogenost njihovih varijansi. Ovo su „udžbenički" uslovi koji nekada mogu da budu i prekršeni a da se to ne odrazi bitno na validnost dobijenog koeficijenta korelacije. Na primer, Pirsonov koeficijent može da bude prikladna mera povezanosti ordinalnih varijabli koje imaju veći broj nivoa, kao i varijabli koje su iskošene u istu stranu. U tom smislu, dijagram raspršenja pruža dragocene informacije o stvarnom odnosu među varijablama i trebalo bi da bude obavezni deo interpretacije koeficijenta povezanosti. Sjajnu ilustraciju važnosti grafičkog prikaza rezultata u korelacionoj analizi dao je engleski statističar Frensis Džon Enskomb u četiri primera podataka koji su poznati kao ***Enskombov kvartet*** (Anscombe, 1973). On je simulirao vrednosti četiri para varijabli koje imaju iste aritmetičke sredine, varijanse i koeficijente korelacije, ali sa potpuno drugačijim međusobnim odnosom koji se uočava samo pomoću skater-dijagrama. Prvi primer, koji je na padajućoj listi naveden kao *Enskombov kvartet 1*, prikazuje odnos dveju varijabli u visokoj korelaciji. Blago raspršenje kružića u odnosu na regresionu liniju potpuno je prihvatljivo i sugeriše da među

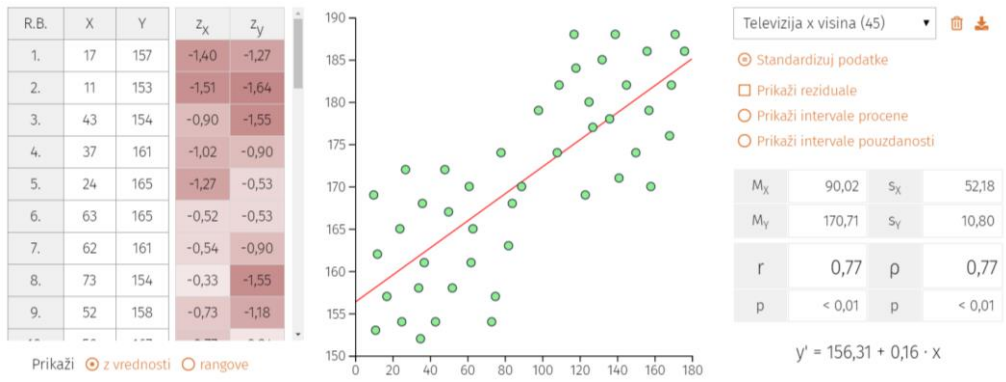
varijablama zaista postoji linearna veza pozitivnog smera jačine 0,82. *Enksombov kvartet 2* prikazuje drugi par varijabli sa istim koeficijentom korelacije, ali sa potpuno drugačijim odnosom (Slika 50). U ovom primeru odnos varijabli je problematičan jer nije linearan, pa tako ni Pirsonov koeficijent korelacije, iako veoma visok, nije prikladan za opisivanje te veze. Logika Pirsonovog r bazira se na pretpostavci pravolinijskog odnosa varijabli, ali u ovom primeru vrednosti y prate porast vrednosti x samo do određene tačke, nakon čega ta veza menja smer. Ovakav nelinearan odnos varijabli bi se, na primer, mogao očekivati u slučaju jačine osvetljenja u prostoriji i radnog učinka, ili u slučaju motivacije i postignuća. Treći primer ilustruje problem koji mogu da stvore aberantni rezultati. Samo jedan autlajer u ovom primeru znatno je povećao koeficijent korelacije, te se može reći da vrednost 0,82 ne odražava pravo stanje stvari. Štaviše, kada uklonite aberantni rezultat, primetićete da r više nije ni moguće izračunati, jer varijansa varijable X postaje nula. Na kraju, primer *Enskombov kvartet 4* ilustruje još jedan mogući uticaj autlajera na vrednost r. U ovom slučaju, aberantni rezultat je umanjio koeficijent korelacije koji bi inače bio maksimalan, što možete da vidite ako uklonite kružić koji značajno odstupa od regresione prave.



Slika 50. Primer nelinearnog odnosa varijabli (Frensis Džon Enskomb)

Pored pretpostavke o linearnom odnosu varijabli, validnost Pirsonovog koeficijenta korelacije bazira se i na očekivanju da je raspršenje rezultata podjednako duž cele regresione prave, odnosno da se verovatnoća greške procene ne razlikuje previše za različite vrednosti x-ose. Ovaj uslov je poznat kao *homoskedasticitet*. Termin je nastao od starogrčkih reči ὁμός (isto) i σκεδαστός (raspršenje). Primer *Pušenje x kapacitet 2* prikazuje situaciju u kojoj je ovaj uslov prekršen. Oblik raspršenja ukazuje na to da je koeficijent korelacije

negativan, ali intenzitet te veze nije isti za sve vrednosti X i Y varijabli. Naime, kapacitet pluća znatno više varira u grupi osoba sa kraćim „pušačkim stažom", tako da je opravdano postaviti pitanje da li jaka korelacija zaista postoji u svim potencijalnim poduzorcima. Na primer, moguće je da su osobe koje su konzumirale cigarete u kraćem periodu, u stvari mlađe osobe kod kojih se pušenje još uvek nije značajno odrazilo na umanjenje kapaciteta pluća. Drugim rečima, varijabilnost varijable Y u zoni nižih vrednosti varijable X može da predstavlja uobičajenu varijabilnost kapaciteta pluća određene uzrasne grupe, bez obzira na to da li su u pitanju pušači ili ne. Još drastičniji primer je *Televizija x visina*. Zamislimo da ste grupi ispitanika izmerili visinu i te podatke doveli u vezu sa brojem minuta koje dnevno provode u gledanju sportskih i informativnih emisija na televiziji. Dobijena korelacija je visoka i pozitivna, ali grafikon ukazuje na postojanje *heteroskedasticiteta*. Moguće objašnjenje je heterogenost uzorka, odnosno činjenica da su na istom grafikonu prikazani i muškarci i žene. Muškarci su u proseku viši a provode i više vremena gledajući sportske događaje, za razliku od žena. To znači da koeficijent korelacije za ove dve varijable najverovatnije ne bi bio značajan kada bi se izračunao u svakoj podgrupi, tj. posebno za muškarce a posebno za žene.



Slika 51. Primer heteroskedasticiteta – nejednakog raspršenja duž regresione linije

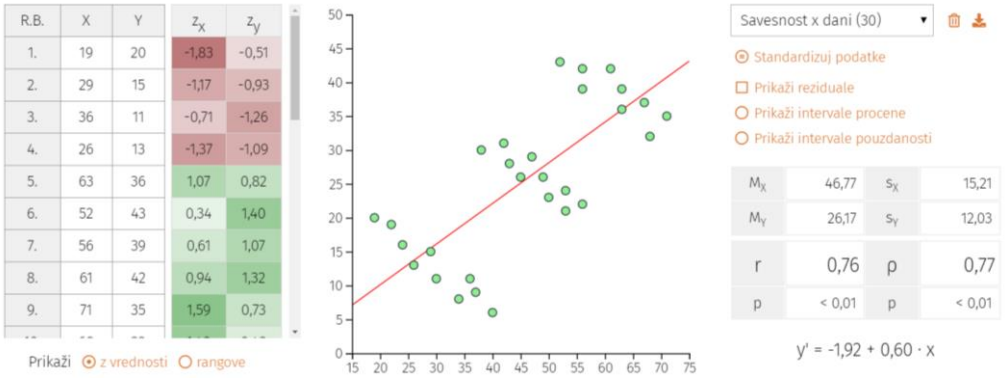
> Kako biste, na osnovu raspršenja kružića u primeru *Televizija x visina*, zaključili kakvog su oblika distribucije dveju varijabli?

Problem heteroskedasticiteta u korelacionoj analizi je dobar povod da čitaocu još jednom skrenemo pažnju na moguće logičke greške u izvođenju zaključaka na osnovu rezultata statističkih analiza. Odaberite primer *Savesnost*

x dani da biste prikazali odnos između skora na skali savesnosti nekog upitnika ličnosti i broja izostanaka sa posla za 30 zaposlenih u jednoj firmi (Slika 51). Visoka pozitivna vrednost koeficijenta r sugeriše da zaposleni koji su postigli više skorove na skali savesnosti, češće odsustvuju sa posla. Međutim, na osnovu skater-dijagrama se vidi da postoje tri stratuma ispitanika koji se očigledno razlikuju u prosečnom broju izostanaka, ali i u prosečnom skoru na skali savesnosti. Zaključak donet na ukupnom uzorku bio bi u suštini pogrešan, jer se ne odnosi ni na jedan od poduzoraka zaposlenih. Štaviše, korelacija dve varijable unutar pojedinačnih stratuma je negativna. Razlog bi mogao da bude to što su savesnije osobe zaposlene na odgovornijim funkcijama a upravo one, zbog većeg stresa koji doživljavaju na takvim pozicijama, češće odsustvuju sa posla. Ovo je, naravno, samo pretpostavka čiju opravdanost bi trebalo na neki način testirati. Pojava da su zaključci koji su doneti na podskupovima podataka drugačiji, pa čak i suprotni od onih koji su doneti na objedinjenom skupu istih podataka, poznata je kao *Simpsonov paradoks* (Simpson, 1951) (Slika 52). Ovaj tip greške u interpretaciji rezultata statističkih analiza javlja se mnogo češće nego što bi se moglo pretpostaviti na osnovu anegdotskih primera kojima se paradoks obično ilustruje (Kievit, Frankenhuis, Waldorp, & Borsboom, 2013). Pored toga, u pitanju je samo jedna od velikog broja logičkih zabluda ili grešaka (engl. *fallacy*) koje se javljaju ako su istraživači usredsređeni samo na (numeričke) rezultate statističkih analiza. Može se reći da je u osnovi ovih zabluda niska odgovornost, nedovoljno iskustvo, ali i niska metodološka i statistička obučenost istraživača. Na primer, analiza razlika među grupama pomoću t-testa može da dovede do potpuno drugačijih zaključaka u zavisnosti od toga da li se kontroliše početni nivo zavisne varijable u grupama, odnosno njena kovarijansa sa drugim bitnim atributima ispitanika. Ovaj fenomen je poznat kao *Lordov paradoks* (Lord, 1967).

Istraživači često zanemaruju činjenicu da je indukcija, tj. generalizacija zaključaka sa uzorka na populaciju, legitiman postupak inferencije, ali da sa sobom nosi verovatnoću greške koja nikada nije nulta. U tom smislu, prosta dedukcija na osnovu takvih generalizacija često je pogrešna. Ako se vratimo na poslednji primer sa izostancima, može se uočiti da druga podgrupa zaposlenih, gledano sleva nadesno, ima nešto viši prosek na skali savesnosti od prve. Međutim, to ne podrazumeva da su svi članovi druge grupe savesniji od bilo kog člana prve. Samim tim, svaki dalji zaključak baziran na prosecima grupa, mogao bi da dovede do greške kada se rezultati primenjuju i interpretiraju na individualnom nivou. Tipičan primer ove logičke greške je *Robinsonov paradoks*,

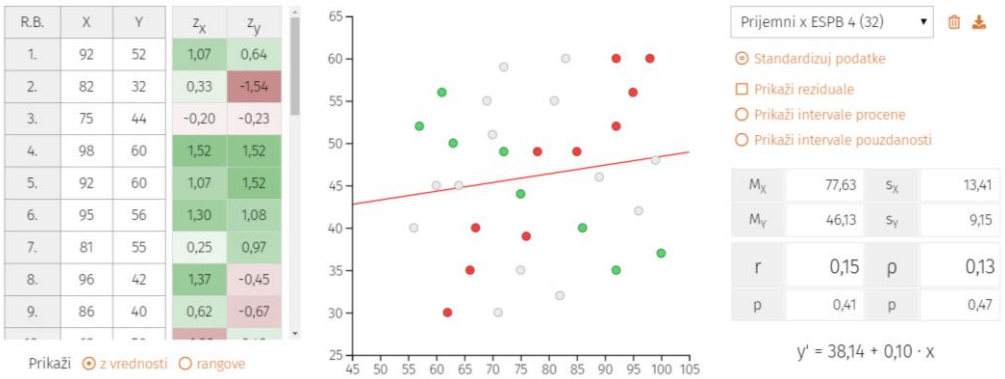
odnosno pojava da su korelacije grupa merenja drugačije od onih koje se računaju na individualnim merenjima. Na primer, u jednom istraživanju je uočeno da postoji značajna negativna korelacija procenta imigranata u državama SAD i procenta nepismenih osoba. Međutim, rezultati analize na individualnom nivou pokazali su da je udeo nepismenih osoba zapravo veći u grupi imigranata, nego u grupi starosedelaca. Objašnjenje ovog naizgled paradoksalnog rezultata leži u činjenici da su se imigranti češće doseljavali u države u kojima je udeo nepismenog stanovništva niži (Robinson, 2009). Jasno je da vizualizacija podataka ne može uvek da pomogne istraživaču da izbegne neku od logičkih grešaka u zaključivanju, ali uz dovoljnu motivaciju da se detaljnije istraže potencijalne veze među varijablama, grafički prikaz podataka je nezamenjiva eksplorativna alatka za uočavanje fenomena koji na prvi pogled nisu vidljivi.



Slika 52. Primer Simpsonovog paradoksa – korelacija na celom uzorku je pozitivna, ali je na poduzorcima negativna

Vratimo se nakratko na primer *Prijemni x ESPB 1*. Raspored kružića na dijagramu potvrđuje podatak da je korelacija između varijabli veoma niska. Međutim, ako kružiće obojimo različitim bojama na osnovu vrednosti neke kategorijalne varijable, npr. tipa srednje škole iz koje student dolazi, zaključak bi mogao da bude drugačiji. Ova opcija postoji u većini statističkih paketa, ali se često previđa jer nije lako identifikovati varijablu koja bi mogla da bude tzv. *kovarijat*. U zavisnosti od analize koja se sprovodi, kovarijat može da bude bilo koja kategorijalna ili kontinuirana, nezavisna ili kontrolna, opažena ili izmerena varijabla koja utiče na povezanost svojstava koja se analiziraju ili efekat nezavisne varijable na zavisnu. Kada odaberete opciju *Prijemni x ESPB 4*, uočićete da prvobitni zaključak o nepostojanju povezanosti važi samo za sivu

podgrupu studenata, dok je u ostalim podgrupama korelacija visoka, a uz to je i različitog smera – u crvenoj je pozitivna a u zelenoj negativna (Slika 53). U ovom primeru vrsta srednje škole je kovarijat, tačnije spoljna varijabla koja nas zbunjuje i iskrivljuje sliku o vezi broja osvojenih bodova na prijemnom ispitu i broja prikupljenih kredita. Stoga se ovakve nepoželjne spoljne varijable često nazivaju i *konfundirajućim*. Međutim, kovarijati nisu uvek neželjene ili nepredviđene varijable. Štaviše, oni se često planirano uključuju u složenije statističke postupke, kao što su npr. *višestruka regresiona analiza* ili *analiza kovarijanse*, kako bi se detaljnije i potpunije opisala priroda odnosa varijabli od interesa.



Slika 53. Efekat kovarijata na povezanost dve varijable

U kontekstu prethodnih primera, potrebno je pomenuti još jedan važan uslov koji treba da bude ispunjen da bi se koeficijent korelacije smatrao validnom merom povezanosti dve varijable. U primeru *Prijemni x ESPB 2* videli smo da ne postoji statistički značajna korelacija rezultata kandidata na prijemnom ispitu i njihovog budućeg uspeha u studiranju. Ovaj podatak deluje obeshrabrujuće, jer je osnovni smisao prijemnog ispita da se napravi izbor kandidata za koje postoji veća verovatnoća da će u toku studija ostvariti bolji uspeh i koji će, na kraju krajeva, blagovremeno završiti studije. Razlog koji stoji iza ovog nelogičnog rezultata je fenomen poznat kao *ograničenje raspona*. Naime, u navedenom primeru doneli smo zaključak na selektivnom uzorku ispitanika, tj. na osnovu ograničenog raspona vrednosti varijable X u grupi kandidata koji su uspešno položili prijemni i ostvarili više od 55 poena. Slika bi mogla da izgleda potpuno drugačije da smo svim kandidatima koji su polagali prijemni ispit, omogućili da započnu studije i tek potom izračunali korelaciju uspeha na testu i uspeha u studiranju. Tada bi, na primer, odnos varijabli mogao

da bude kao onaj prikazan u primeru *Prijemni x ESPB 3*. Sada je očigledno da među varijablama postoji visoka i statistički značajna korelacija. Verovatnoća da će kandidati koji su bolje uradili prijemni biti i bolji studenti, zaista je veća ako se posmatra celokupan uzorak kandidata. Međutim, u okviru grupe onih koji su bili bolji na prijemnom ispitu, broj osvojenih bodova prestaje da bude bitan prediktor broja osvojenih ESPB kredita.

> Kakav je oblik distribucija varijabli u primeru *Pušenje x kapacitet 2*?
>
> Da li odstupanje varijabli X i/ili Y od normalnosti podrazumeva da njihov odnos neće biti linearan?
>
> Na osnovu čega se pomoću skater-dijagrama može utvrditi da su obe varijable, barem približno, normalno distribuirane?

3.6.5. Korelacija i uzročnost

Verovatno najopasnija logička greška koja se vezuje za pojam korelacije jeste greška neopravdanog pripisivanja uzročno-posledičnih veza analiziranim pojavama. Rezultati korelacione i regresione analize često navode istraživače da donesu zaključke o kauzalnom odnosu varijabli, jer smo generalno skloni da fenomene koji nas okružuju opisujemo na takav način, npr. da se nešto desilo *zbog* određenog razloga ili da će neki postupak *izazvati* određenu reakciju. Međutim, treba biti veoma oprezan prilikom donošenja ovakvih zaključaka samo i isključivo na osnovu koeficijenta korelacije. U primeru Šarlovog zakona visoka povezanost zapremine i temperature gasa zaista potvrđuje uzročno-posledičnu vezu, jer je pretpostavka da se sve ostale varijable, kao što je npr. pritisak gasa, drže pod kontrolom ili su konstantne. Pri tome je očigledno da povišenje temperature predstavlja uzrok povećanja zapremine a ne obratno, jer promena na prvoj varijabli *vremenski prethodi* promeni na drugoj. Prilikom grafičkog prikazivanja odnosa varijabli uz pomoć skater-dijagrama, uobičajeno je da se potencijalni uzrok ili *prediktor* prikazuje na x-osi a potencijalna posledica ili *kriterijum* na y-osi. Ovaj princip je primenjen i u primerima *Prijemni x ESPB*, ali samo zato što broj osvojenih bodova na prijemnom ispitu prethodi broju ostvarenih ESPB bodova. To i dalje ne znači da je X uzrok a Y posledica. Mnogo je verovatnije da varijable koje su u korelaciji u stvari imaju zajednički

uzrok, a to bi u ovom primeru mogle da budu različite intelektualne sposobnosti studenata, njihova motivacija, pa čak i neke osobine ličnosti.

Neopravdanost zaključaka o kauzalnosti na osnovu korelacije postaje još očiglednija u primerima *Težina x visina*. Ne samo da visina osobe ne utiče na njenu težinu ili obratno, već se ne može reći čak ni da jedna od te dve varijable vremenski prethodi onoj drugoj. Greške neopravdanog pripisivanja kauzalnosti nisu retke čak ni među iskusnijim istraživačima i naučnicima. Poznat je primer istraživanja u kome je utvrđena korelacija učestalosti korišćenja noćnog svetla kod male dece i kratkovidosti u starijem dobu (Quinn, Shin, Maguire, & Stone, 1999). Međutim, u naknadnim studijama utvrđeno je da korelacija zapravo ne postoji i da istraživači u originalnoj studiji nisu kontrolisali sve relevantne varijable, kao što je npr. nasledni faktor, odnosno kratkovidost roditelja. Sličan primer predstavlja niz istraživanja u kojima je utvrđeno da hormonske terapije estrogenom kod žena smanjuju rizik od pojave koronarne bolesti srca. U ovom slučaju nisu kontrolisane bitne varijable kao što su starost ispitanica ili njihov materijalni status, tako da su u nekim naknadnim studijama dobijene čak i korelacije suprotnog smera (Sotelo & Johnson, 1997). Očigledno je da velika odgovornost za korektno sprovođenje istraživanja i tačnu interpretaciju dobijenih rezultata pripada samim istraživačima i ima veze sa njihovom etičnošću i nivoom kompetencija. Ako istraživači nisu dovoljno metodološki i statistički obučeni, a pri tome se primarno vode potrebom da proizvedu što više senzacionalističkih rezultata, postoji rizik da se naučni prostor zagadi pseudonaučnim i kvazinaučnim informacijama koje nemaju potporu u stvarnosti. Korelacione studije su, u tom smislu, posebno osetljive, jer pojave koje nemaju nikakvih dodirnih tačaka, lako mogu da se dovedu u vezu samo zbog vremenske koincidencije. U popularnoj literaturi često se navode primeri pseudokorelacija globalnog zagrevanja i učestalosti terorističkih napada ili količine prodatih sladoleda i broja ubistava u SAD.

Upozorenja koja smo upravo izneli ne znače da kauzalno zaključivanje nije dozvoljeno, a još manje da nije moguće na osnovu rezultata korelacionih analiza. Naprotiv. Osnovni uslov koji mora da bude ispunjen da bi se neke dve pojave nazvale uzrokom i posledicom, upravo je postojanje njihove značajne korelacije. Ali, kao što smo rekli, taj uslov nije dovoljan. Dodatni uslovi su da mogući uzrok vremenski prethodi posledici i da su isključene ili kontrolisane sve varijable koje bi mogle da budu pravi ili alternativni uzroci u kauzalnoj vezi. Jasno je da su zaključci o (ne)postojanju kauzalnih veza mogući samo ako su uslovi u kojima se sprovodi istraživanje strogo kontrolisani, kao u primeru u

kome smo testirali postojanje efekta Miler-Lajerove iluzije. Stoga se kauzalno zaključivanje obično, ako ne i isključivo, vezuje za eksperimentalne nacrte u kojima istraživač ima potpunu kontrolu i mogućnost da utiče na vrednosti (nezavisnih) varijabli. U korelacionim nacrtima takva vrsta kontrole ne postoji, pa tako ni zaključci korelacionih studija ne mogu da imaju istu težinu. To nikako ne znači da su oni manje vredni, već samo da imaju drugačiji smisao od onih koji se postižu klasičnim eksperimentima. U tom smislu, statističke metode ne treba posmatrati kao algoritme kojima se matrice sirovih podataka pretvaraju u testove, koeficijente i p vrednosti. Statističke metode su (završna) faza složenog procesa opisivanja prirodnih i društvenih pojava, koja zahteva razumevanje tih pojava, načina na koje su one registrovane i izmerene, kao i konteksta u kome će prikazani rezultati biti tumačeni i upotrebljeni.

3.7. Koeficijenti korelacije za rangirane podatke

Pirsonov koeficijent korelacije smatra se dovoljno robusnom metodom koja je otporna na blaža odstupanja distribucija od tipične Gausove krive. Međutim, u slučaju postojanja aberantnih rezultata ili značajnijih odstupanja podataka od normalne raspodele, prikladniju meru povezanosti predstavlja neki od neparametrijskih koeficijenata korelacije. To znači da se nivo merenja može i naknadno „spustiti“ kako bi se omogućila primena statističkih metoda manje snage. Na primer, najbolji izbor za iskazivanje stepena povezanosti visine i težine ispitanika jeste Pirsonov koeficijent korelacije, ali ukoliko nisu ispunjeni uslovi za njegovu primenu, potpuno je legitimno da se vrednosti obe varijable rasporede u kategorije (npr. viši, srednji niži, odnosno teži, srednji, lakši) i da se potom izračuna koeficijent kontingencije C. Međutim, prethodno bi trebalo proveriti da li se adekvatnija i preciznija informacija o odnosu među varijablama dobija primenom nekog od koeficijenata korelacije za varijable ordinalnog nivoa merenja. Verovatno najpopularnija metoda ovog tipa je *Spirmanov koeficijent ρ* (ro), nazvan po engleskom psihologu Čarlsu Edvardu Spirmanu, jednom od pionira u oblasti merenja i opisivanja strukture inteligencije. Spirmanov ρ može da se primeni kod bilo koje kombinacije varijabli ordinalnog ili višeg nivoa merenja (npr. ordinalni-intervalni, racio-ordinalni, intervalni-racio), ali ne i na varijablama nominalnog nivoa. Uobičajeni postupci računanja Spirmanovog ρ i Pirsonovog r potpuno su isti, izuzev što se podaci u slučaju koeficijenta ρ prethodno transformišu u rangove. Ovu transformaciju, naravno, ne obavlja istraživač već statistički program u kome se vrši obrada podataka.

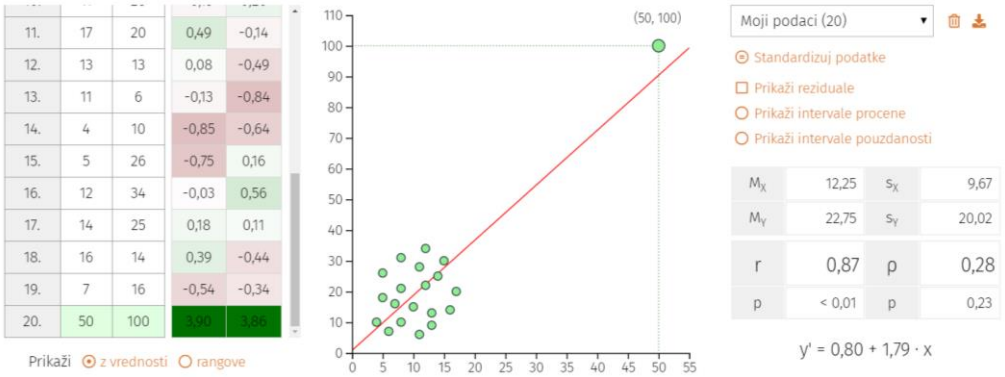
Ako ponovo prikažete grafikone za neke od **primera koje smo opisali u prethodnom odeljku**, primetićete da se vrednosti koeficijenata r i ρ uglavnom neznatno razlikuju. Najočiglednija razlika javlja se ako u grupi merenja postoje aberantni rezultati (npr. *Enskombov kvartet 3*), pa ćemo sličan primer upotrebiti za kratku ilustraciju logike Spirmanovog ρ.

Odaberite opciju *Učenje x bodovi 1* i uklonite sve kružiće sa grafikona. Sada dodajte dvadesetak ispitanika u donji levi kvadrant grafikona, tako da se njihove x vrednosti kreću u rasponu 0 do 20, y vrednosti između 0 i 40, i da je korelacija varijabli veoma niska. Spirmanov ρ bi trebalo da ima nisku i približno istu vrednost kao Pirsonov r. Sada dodajte ispitanika u gornji desni ugao grafikona, npr. sa koordinatama (50, 100). Obratite pažnju na to da se Pirsonov koeficijent korelacije značajno povećao, za razliku od Spirmanovog, koji i dalje nije statistički značajan (Slika 54). Locirajte podatke za poslednjeg dodatog ispitanika u tabeli tako što ćete postaviti pokazivač miša iznad kružića koji ga predstavlja. Već na osnovu boje ćelija u tabeli, a potom i na osnovu ekstremno visokih z vrednosti, jasno je da ispitanik na obe varijable značajno odstupa od proseka svoje grupe. Zbog toga se bitno uvećala suma proizvoda z vrednosti a time i Pirsonov koeficijent korelacije. Međutim, postojanje autlajera nije bitno uticalo na Spirmanov koeficijent, jer se prilikom njegovog računanja podaci tretiraju kao rangovi. Odaberite opciju *Prikaži rangove* ispod tabele sa leve strane da biste, umesto z vrednosti, prikazali rangove ispitanika. Sada se poslednji dodati student razlikuje od prvog narednog za samo jedan rang, a ne za tri standardne devijacije ili 30 sati, odnosno 60 bodova, kao u prethodnom slučaju. Ponovo odaberite primer *Učenje x bodovi 1* i uklonite dva ispitanika koji najviše odstupaju od regresione prave, tako što ćete kliknuti kružiće dok držite pritisnut taster *Shift*. Oba koeficijenta korelacije dobila su vrednost 1. Osim toga, rangovi studenata na obe varijable postali su isti, što znači da suma njihovih razlika sada iznosi nula. To je, ukratko, logika jedne od formula koja se koristi za računanje Spirmanovog koeficijenta korelacije:

$$\rho = 1 - \frac{6\sum \mathrm{D}^2}{\mathrm{N}(\mathrm{N}^2 - 1)}$$

Simbol D u gornjoj formuli označava razliku rangova za svaki par rezultata. Ako je suma tih razlika nula, Spirmanov ρ će imati vrednost 1. Ako je suma razlika među rangovima maksimalna za dati skup podataka, kao u primeru *Pušenje x kapacitet 1*, Spirmanov ρ će imati vrednost -1. U većini programa za statističku obradu, za računanje Spirmanovog koeficijenta korelacije koriste se formule

pomoću kojih se računa i Pirsonov r. One su prikladnije od gore navedene formule u situacijama kada postoje spojeni rangovi.



Slika 54. Uticaj abernatnog rezultata na Pirsonov i Spirmanov koeficijent korelacije

Pored Spirmanovog ρ, stepen povezanost varijabli rang nivoa često se iskazuje i *Kendalovim* τ (tau) koeficijentom koji se bazira na određivanju broja *invertovanih* ili *diskordantnih (nesaglasnih) rangova*. Vrednosti jedne varijable se najpre sortiraju a potom se obavi međusobno poređenje rangova na drugoj varijabli. Invertovanim se smatraju parovi rangova druge varijable u kojima se viši rang nalazi ispod nižeg. Odaberite primer *Enskombov kvartet 1* i opciju *Prikaži rangove*. U prvom redu postoji samo jedan invertovani rang jer se u koloni R_y vrednost 1 nalazi ispod vrednosti 2. U drugom redu broj nesaglasnih rangova je 4, jer se ispod 6 nalaze 1, 4, 5 i 3. U trećem redu ih nema jer je 1 najviši rang. I tako dalje. Ukupan broj takvih poređenja iznosi N · (N - 1) : 2, a lako se može izračunati i broj *konkordantnih (saglasnih)* rangova. Kendalov τ se nakon toga izračunava prema formuli:

$$\tau = \frac{K - D}{K + D}$$

ili prema formuli:

$$\tau = 1 - \frac{4D}{N(N-1)}$$

gde je K broj saglasnih, D broj nesaglasnih rangova, a N veličina uzorka. Kendalov τ se smatra boljim pokazateljem povezanosti od Spirmanovog ρ za podatke ordinalnog nivoa merenja, zato što daje tačniju procenu parametara,

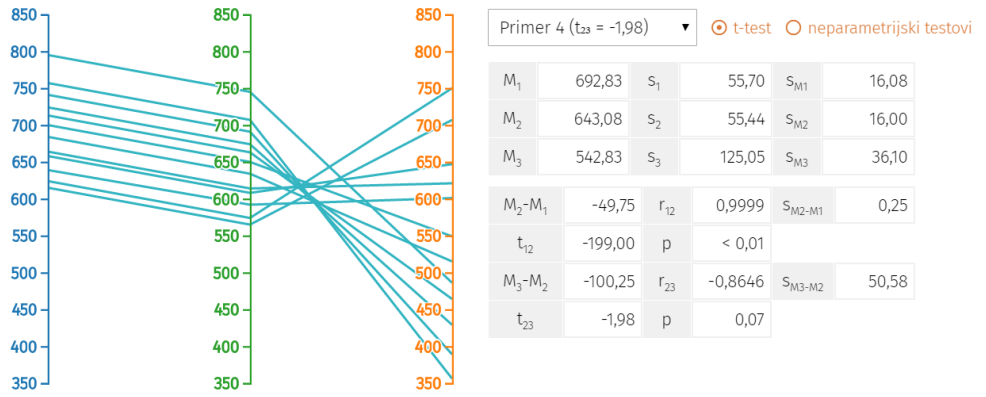
tj. stvarne korelacije u populaciji. Stoga se može izračunati i njegova standardna greška, za razliku od Spirmanovog ρ gde to nije moguće (Howell, 2012).

U literaturi se često pominje i *Gudmen–Kraskalov γ* (gama) koeficijent, ali ga ovde nećemo detaljnije opisivati jer se bazira na veoma sličnoj logici kao i τ. Različiti statistički paketi koriste različite formule za računanje Kendalovog koeficijenta, tako da se pored opisanog τ, koji se često označava sa tau-a, nude i njegove modifikacije tau-b i tau-c u kojima se vrednost koeficijenta koriguje s obzirom na broj spojenih rangova. Generalno gledano, τ i γ su bolja rešenja od Spirmanovog ρ u slučaju velikog broja spojenih rangova, što znači da se mogu primeniti čak i na tabelama kontingencije ukoliko postoji osnov da se redovi i kolone sortiraju po nekom kriterijumu. Na kraju, treba naglasiti da se opisani koeficijenti korelacije rangova ne smeju posmatrati kao podrazumevane alternative Pirsonovom r. Oni jesu bolje rešenje kada varijable nisu normalno distribuirane ili su merene na ordinalnom nivou merenja, ali su podjednako osetljivi na odstupanja koja se tiču nelinearnog odnosa među varijablama, ograničenja raspona ili heteroskedasticiteta. Na primer, ukoliko na grafikonu iscrtate kružiće u obliku slova U, videćete da su oba koeficijenta korelacije niska. To, naravno, ne znači da varijable nisu povezane, već samo da im je odnos nelinearan. O tome je već bilo reči u odeljku 3.5.1. Kada biste vrednosti varijabli izdelili u kategorije i izračunali koeficijent kontingencije C, korelacija bi najverovatnije bila visoka i značajna.

3.8. T-test za zavisne uzorke

U odeljku 3.3. opisali smo t-test za dva uzorka kao postupak kojim se procenjuje statistička značajnost razlike dve aritmetičke sredine. Istraživački nacrt u kome se ova metoda primenjuje podrazumeva postojanje barem jedne kvantitativne varijable kao zavisne i jedne dihotomne varijable kao nezavisne. Nakon što se merenja podele u dve grupe na osnovu nivoa nezavisne varijable, računa se vrednost t-testa za razliku aritmetičkih sredina zavisne varijable. Na primer, ako bi farmaceut želeo da testira efikasnost novog leka u terapiji hipertenzije, visina krvnog pritiska pacijenata bila bi zavisna varijabla a vrsta terapije nezavisna. Ukoliko je cilj da se proveri relativna efikasnost leka u odnosu na drugu vrstu terapije, jedna grupa ispitanika bi koristila novi, a druga postojeći lek. Ukoliko je, pak, potrebno proveriti apsolutnu efikasnost leka, eksperimentalna grupa bi koristila novi preparat, a kontrolna sredstvo sa

potpuno neutralnim dejstvom ili tzv. *placebo*. Ovaj drugi nacrt je problematičan sa etičkog aspekta, jer ne samo da se grupa pacijenata dovodi u zabludu, već im se uskraćuje i pravo na terapiju u određenom vremenskom periodu. Osim toga, u oba nacrta postoji problem vezan za adekvatno ujednačavanje grupa po svim potencijalno važnim svojstvima, kao što su pol, uzrast, nivo holesterola u krvi, učestalost bavljenja fizičkim aktivnostima, vrsta posla i tako dalje. Stoga se u istraživanjima često koristi nacrt u kome se obe aritmetičke sredine računaju na istoj grupi ispitanika. U pomenutom primeru, to bi značilo da se najpre izračuna prosek vrednosti krvnog pritiska u grupi ispitanika, potom se svima daje terapija u određenom vremenskom periodu, a na kraju tog perioda zavisna varijabla se ponovo izmeri i izračuna njena aritmetička sredina. Ovaj postupak poznat je kao *t-test za zavisne uzorke* ili *t-test za ponovljena merenja*. Analogno tome, potpuniji naziv metode koja je opisana u poglavlju 3.3. bio bi *t-test za nezavisne uzorke*. Pri tome prideve „zavisni" i „nezavisni" koji se odnose na uzorke, ne treba mešati sa istim terminima koji se odnose na status varijable u eksperimentu. Termin „zavisni uzorci" treba da ukaže na činjenicu da rezultati drugog merenja na neki način zavise od rezultata prvog, jer su u pitanju iste osobe, odnosno entiteti. Štaviše, u opisanom primeru ne postoje dva uzorka *ispitanika*, već dva uzorka *merenja* koja su obavljena na istoj grupi pacijenata. Samim tim, istraživač ne mora da ujednačava grupe kao u slučaju nezavisnih uzoraka, te je svaka uočena razlika, u relativnom smislu, značajnija i veća, jer se može pripisati samo promeni koja se desila između merenja, a ne drugim svojstvima ispitanika koja nisu (u dovoljnoj meri) ujednačena. Oba istraživačka nacrta imaju svoje prednosti i nedostatke vezane prvenstveno za problem ujednačavanja grupa kod nezavisnih uzoraka i problem osipanja ispitanika kod zavisnih. Stoga se kao optimalno rešenje preporučuje kombinacija ovih nacrta, odnosno ponovljena merenja na dve grupe ispitanika, eksperimentalnoj i kontrolnoj. Međutim, podatke prikupljene na ovaj način ne bi trebalo obrađivati primenom t-testa, već složenijim statističkim postupcima, kao što je npr. *analiza varijanse za ponovljena merenja*. Ovaj postupak izlazi van okvira tematike ovog udžbenika tako da ćemo se u daljem tekstu, za potrebe ilustracije postupka analize ponovljenih merenja, zadržati na t-testu za zavisne uzorke.



Slika 55. Dijagram paralelnih koordinata kojim su prikazani rezultati tri merenja na istom uzorku ispitanika (nacrt za zavisne uzorke)

Zamislićemo istraživanje u kome je **grupi od 12 pilota izmerena brzina reakcije na vizuelne stimuluse**. Nakon toga svi piloti su prošli obuku na simulatoru letenja u trajanju od mesec dana. Po isteku obuke ponovo im je izmerena brzina reakcije a rezultati su prikazani uz pomoć *dijagrama paralelnih koordinata* (Slika 55). Na plavoj osi biće zabeleženi rezultati prvog merenja a na zelenoj rezultati drugog. Podeoci na osama označavaju brzinu izraženu u hiljaditim delovima sekunde. Odaberite *Primer 1* sa liste da biste prikazali odnos rezultata prvog i drugog merenja. Svaka linija predstavlja jednog ispitanika. Aritmetička sredina prvog (plavog) merenja iznosi približno 693 ms. U dugom (zelenom) merenju, svaki pilot je reagovao za 50 ms brže nego u prvom, pa prosek brzine reakcije nakon obuke iznosi oko 643 ms. Na osnovu grafikona zaključujemo da se kod svih pilota desila identična linearna promena, tako da su standardna devijacija i standardna greška aritmetičke sredine u drugom merenju iste kao u prvom, a korelacija merenja je potpuna. Greška razlike aritmetičke sredine iznosi nula, jer je tvrdnja da se u grupi od 12 pilota brzina reakcije povećala za 50 ms potpuno tačna. Drugim rečima, ova tvrdnja se ne odnosi samo na ispitanike kao grupu već i na svakog pojedinačnog pilota u uzorku. Međutim, zbog toga nije moguće izračunati vrednost t-testa, jer se u imeniocu formule

$$t = \frac{M_1 - M_2}{s_{M_1 - M_2}}$$

nalazi vrednost 0, pa je t u stvari beskonačno. Stoga ćemo morati „veštački" da povećamo grešku, odnosno da napravimo barem malu varijaciju u promenama

koje su se desile na nivou pojedinačnih parova rezultata. Odaberite *Primer 2* i uočite da je rezultat jednog od ispitanika na drugom merenju promenjen. Ova intervencija je neznatno umanjila razliku i korelaciju, ali je povećala grešku razlike, pa je sada moguće izračunati vrednost t-testa. Nivo verovatnoće p govori nam da postoji efekat jednomesečne obuke, odnosno da je razlika između prvog i drugog merenja statistički značajna na nivou 0,01. Na ovom primeru se vidi da je standardna greška razlike aritmetičkih sredina u suštini varijabilnost individualnih promena između prvog i drugog merenja. Ona nam govori koliko poverenja možemo da imamo u apsolutnu vrednost razlike, tj. u kojoj meri je ona dobar predstavnik promena koje su se desile na nivou svakog pojedinačnog ispitanika.

Zamislimo sada da su u drugoj fazi istraživanja izmenjene određene karakteristike stimulusa, npr. njihova boja. Nakon toga je ponovo izmerena brzina reakcije, a rezultati će biti prikazani na trećoj (narandžastoj) osi kada odaberete *Primer 3*. Očigledno je da promena boje stimulusa ima mnogo jači efekat na brzinu reakcije nego obuka na simulatoru letenja, što se vidi iz razlike M_3 - M_1 koja je duplo veća od M_2 - M_1 i iznosi oko -100 ms. Standardna greška od 0,25 ms jednako je mala kao i u prethodnom primeru, tako da je vrednost t-testa duplo veća. Međutim, čak i tako velika razlika ne mora nužno da bude značajna. Na primer, piloti su mogli potpuno drugačije da reaguju na promenu boje stimulusa zbog razlike u perceptivnim sposobnostima ili zbog određenih ličnih preferencija. Kao ilustraciju ćemo upotrebiti *Primer 4*. U ovom slučaju, apsolutna razlika je potpuno ista kao i u prethodnom ali je t-test nizak i nije statistički značajan. Razlog je velika greška razlike, koja je delom posledica veće varijabilnosti podataka u trećem merenju, a delom posledica visoke negativne korelacije drugog i trećeg merenja. Kada smo u prethodnom primeru tvrdili da se desila promena od oko -100 ms, to je bilo tačno za gotovo svakog pilota. Ista tvrdnja u ovom primeru ne važi praktično ni za jednog ispitanika. Pređite pokazivačem miša preko linija da biste jasnije videli tok promene brzine reakcije pojedinaca. Uočavate da se kod nekih ispitanika desila promena od skoro 200 ms u pozitivnom smeru, kod nekih preko 300 ms u negativnom, a kod nekih gotovo da nije ni došlo do promene. Ispitanici su, dakle, na različite načine reagovali na promenu boje stimulusa. Štaviše, piloti koji su bili brži u drugom merenju, sada su sporiji, a oni koji su bili sporiji, sada su brži. Upravo ta negativna korelacija drugog i trećeg merenja bila je presudna za drastično smanjenje vrednosti t-testa. Naime, u slučaju t-testa za zavisne uzorke,

standardna greška razlike aritmetičkih sredina računa se prema nešto drugačijoj formuli:

$$s_{M_1-M_2} = \sqrt{s_{M_1}^2 + s_{M_2}^2 - 2r_{12}s_{M_1}s_{M_2}}$$

Vrednost r_{12} u formuli predstavlja korelaciju rezultata dva merenja. Obratite pažnju na to kako visina Pirsonovog koeficijenta korelacije utiče na vrednost standardne greške razlike aritmetičkih sredina. Ukoliko je $r_{12} = 0$, formula se svodi na onu koja se koristi u slučaju t-testa za nezavisne uzorke:

$$s_{M_1-M_2} = \sqrt{s_{M_1}^2 + s_{M_2}^2}$$

Drugim rečima, ako nema povezanosti između ponovljenih merenja, tretiramo ih kao da potiču od dve različite grupe ispitanika, iako su u pitanju iste osobe. Ukoliko je, pak, korelacija visoka i pozitivna, vrednost greške postaje bitno manja, a vrednost t-testa veća. Na taj način se povećava verovatnoća dobijanja statistički značajne razlike, jer njenu validnost potvrđuje činjenica da su se promene desile na nivou (gotovo) svakog pojedinca. Na kraju, ukoliko je vrednost r_{12} negativna, standardna greška postaje veća nego što bi bila da smo primenili formulu za nezavisne uzorke. Ako je nulta korelacija sugerisala da ispitanike možemo da tretiramo kao različite i nezavisne grupe, onda negativna korelacija pokazuje da oni zaista jesu bitno drugačiji, tj. da je promena koja postoji na nivou grupa, veoma loš reprezent promena koje su se desile kod svakog pojedinačnog ispitanika.

Odaberite ponovo *Primer 3* sa liste. Najsporiji pilot u grupi prikazan je najvišom linijom na grafikonu. Njegovi rezultati odstupaju od proseka grupe za približno 1,8 standardnih devijacija, a od prvog narednog rezultata za manje od jedne standardne devijacije. Iako položaj linije može da sugeriše da je u pitanju autlajer, ni Tukijev ni Šoveneov kriterijum ne ukazuju na to da ovaj podatak treba odbaciti. Locirajte najbržeg pilota u grupi i pomeranjem linije kojom su predstavljeni njegovi rezultati, smanjite mu brzinu reakcije u trećem merenju na 350 ms. Pratite kako to utiče na pokazatelje u tabeli. Ova promena je povećala varijabilnost trećeg merenja, neznatno smanjila korelaciju i drastično umanjila vrednost t-testa, odnosno značajnost razlike. Sada polako vratite liniju na prethodnu poziciju, tj. na rezultat trećeg merenja od oko 460 ms, a potom je pomerite naviše, do vrednosti 600 ms. Ovoga puta vrednost t-testa je još više opala, iako je standardna devijacija trećeg merenja postala manja. Razlog je

mnogo manja korelacija, jer za razliku od prethodne intervencije, u ovoj je promena između drugog i trećeg merenja dobila obrnuti smer. Ispitanik čije smo rezultate izmenili nije aberantan ni na jednom merenju, ali njegova *promena* postaje autlajer i može da utiče na validnost zaključka o razlici aritmetičkih sredina. Ukoliko nastavite da povećavate brzinu reakcije istog ispitanika do vrednosti 850 ms, uočićete da t-test više neće biti statistički značajan. U ovom slučaju, aberantnost rezultata trećeg merenja i aberantnost promene između merenja, doveli bi do zaključka koji se ne može smatrati validnim za većinu ispitanika. Detekcija autlajera je, naravno, najlakša ukoliko se podaci prikažu grafički, ali na prvu vrstu aberantnosti mogla bi da ukaže visoka vrednost s_{M3}, a na drugu niska vrednost r_{23}. Obe ove promene uticale su na povećanje vrednosti s_{M3-M2}.

> Prikažite *Primer 4* i pokušajte da promenama rezultata samo jednog ispitanika dobijete statistički značajnu vrednost t-testa.

Postupak koji smo upravo opisali često se naziva i *t-test za uparene rezultate* (engl. *paired samples* ili *matched-pairs*). Štaviše, ovaj naziv je prikladniji od prethodno navedenih jer direktnije upućuje na osnovnu logiku metode. Naime, t-test za zavisne uzorke bazira se na poređenju u kojem postoji mogućnost i opravdanje da se svaki rezultat iz jedne grupe, poveže (upari) sa samo jednim odgovarajućim rezultatom iz druge grupe. U slučaju ponovljenih merenja, ta veza se ostvaruje na osnovu činjenice da dva rezultata potiču od istog ispitanika. Ali to ne mora uvek da bude slučaj. Kada bismo, na primer, želeli da uporedimo grupe dečaka i devojčica u uspešnosti na nekom testu, preporučeni postupak bio bi t-test za nezavisne uzorke. Pri tome se, kao što smo rekli, mora voditi računa o ujednačavanju grupa. Najprecizniji postupak ujednačavanja bilo bi uparivanje pojedinačnih ispitanika tako da se svako dete iz jedne grupe (dečak), poveže sa jednim detetom iz druge grupe (devojčicom) sa kojim je veoma slično po svim relevantnim svojstvima, izuzev po vrednostima nezavisne i zavisne varijable, tj. polu i uspehu na testu. Ovaj postupak je dugotrajan i složen, tako da se retko koristi u istraživanjima, ali omogućava primenu nacrta za zavisne uzorke. Prednost ovakvog pristupa nije samo u povećavanju teorijske verovatnoće da se uoči neki fenomen, već i u tome što se dobijaju preciznije informacije o promenama „unutar" pojedinaca (engl. *within-subjects*) u odnosu na poređenje prosečnih rezultata ispitanika (engl. *between-subjects*). Kao što smo videli iz prethodnih primera, promena na nivou grupe ne

mora tačno niti verno da odražava karakteristike promena na nivou pojedinaca. Računanjem koeficijenta korelacije dve grupe merenja veća važnost se pridaje upravo razlikama na individualnom nivou. U tome je ključna razlika između t-testa za nezavisne i t-testa za zavisne uzorke. Druga tehnika je prikladnija za poređenje grupa međusobno povezanih ispitanika, npr. braće i sestara, muževa i žena ili roditelja i njihove dece. Naravno, samo pod uslovom da su aritmetičke sredine zavisne varijable uporedive i da potiču sa istih mernih skala, npr. IQ skor, dioptrija, skor na standardizovanom upitniku ličnosti i sl. U takvim situacijama, podatak o razlici može da bude korisna dopuna koeficijentu korelacije u opisivanju nekog fenomena. Na primer, moguće je da postoji značajna povezanost stepena ekstravertnosti bračnih partnera, ali to ne mora da znači da postoji i razlika u prosečnim skorovima. U proseku, muževi mogu da budu jednako, manje ili više ekstravertni od žena, a da korelacija dve grupe merenja bude identična.

Da bismo ilustrovali logiku poređenja uparenih rezultata, ponovo ćemo upotrebiti primer *Primer 3*. Držite pritisnut taster *Ctrl* na tastaturi i pomerajte bilo koju liniju iz druge kolone da biste linearno promenili i sve ostale rezultate. Pokušajte da izjednačite vrednosti M_2 i M_3, koliko je to moguće, tako da dobijete vrednost t-testa koja nije statistički značajna. Obratite pažnju na činjenicu da su čak i veoma male promene od 1 do 3 ms statistički značajne, jer je korelacija merenja izuzetno visoka. Tek promena koja je veoma bliska nuli prestaje da bude statistički značajna. Pomerajte sve rezultate naviše i naniže i pratite kako to utiče na vrednosti t-testa, aritmetičkih sredina i koeficijenta korelacije. Sada zamenite rezultate najbržeg i najsporijeg pilota u trećem merenju pomeranjem njihovih linija naviše, odnosno naniže. Ove dve promene drastično su smanjile koeficijent korelacije, a zbog toga i statističku značajnost razlike, mada nisu bitno uticale na varijabilnost i prosek brzine u trećem merenju. Menjajte prosek trećeg merenja tako što ćete pomerati neku od linija držeći pritisnut taster *Ctrl*. Uočite da, zbog veoma niske korelacije dva merenja, razlika ovoga puta treba bude mnogo veća da bi bila statistički značajna. Pre nego što ste zamenili rezultate najbržeg i najsporijeg pilota, čak i razlike od nekoliko milisekundi bile su značajne, ali nakon toga ni pedeset puta veće promene nisu statistički značajne. Ovi primeri ilustruju činjenicu da aritmetičke sredine dve varijable (merenja) ne govore ništa o stepenu njihove povezanosti. Isto tako, ukoliko među nekim varijablama ili merenjima postoji visoka korelacija, to nikako ne znači da su njihove aritmetičke sredine iste. Sa druge

strane, ako su razlike standardnih devijacija merenja velike, kao u četvrtom primeru sa liste, to znači da poređenje proseka verovatno nije opravdano.

U grupi ispitanika evidentiran je broj zapamćenih besmislenih slogova neposredno nakon učenja i dva sata kasnije. Šta je zavisna a šta nezavisna varijabla u ovom istraživanju?

Da li bi t-test u četvrtom primeru bio statistički značajan da je primenjeno jednostrano testiranje razlike? Koja vrsta testiranja razlike je primerenija u ovom istraživanju?

Pokušajte da povećate standardnu devijaciju trećeg merenja u trećem primeru na oko 45 ms, a da pri tome korelacija drugog i trećeg merenja ostane približno ista.

Koja vrsta greške u zaključivanju je verovatnija ako se na dve grupe ispitanika sa uparenim rezultatima primeni t-test za nezavisne uzorke?

U čemu se razlikuju matrice sirovih podataka, na osnovu kojih se računa t-test za nezavisne uzorke i t-test za zavisne uzorke?

Prisetite se primera sa Miler-Lajerovom iluzijom u kome je svih 14 merenja obavljeno na istom uzorku, tačnije na istom ispitaniku. Da li je zbog toga opravdano primeniti t-test za zavisne umesto t-testa za nezavisne uzorke?

3.9. Neparametrijske alternative t-testu za zavisne uzorke

Uslovi za primenu t-testa opisani u poglavlju 3.3.1. trebalo bi da budu ispunjeni i u slučaju nacrta za zavisne uzorke. Oni podrazumevaju intervalni ili racio nivo merenja, približno normalno distribuirane podatke, homogenost varijansi i odsustvo autlajera. Kao što smo pokazali u jednom od prethodnih primera, poslednji uslov se ne odnosi samo na značajna odstupanja pojedinih merenja od proseka, već i na moguća odstupanja smera i stepena *promene* između merenja. Kao i u slučaju t-testa za nezavisne uzorke, ukoliko je neki od pomenutih uslova ozbiljno prekršen, trebalo bi razmotriti primenu alternativnih neparametrijskih testova. Ako se rezultati merenja mogu rangirati, uobičajene alternative t-testu za uparene rezultate su *Vilkoksonov test ekvivalentnih parova* (engl. *Wilcoxon's matched-pairs test*) i *Test predznaka* (engl. *Sign test*). Odaberite

Primer 4* sa liste i opciju *neparametrijski testovi. Rezultati prvog merenja su prikazani u redu A, a rezultati drugog u redu B. Postupak računanja Testa ekvivalentnih parova počinje izračunavanjem razlike među parovima rezultata dva merenja (B-A). Nakon toga se razlike rangiraju od najmanje do najveće po svojoj apsolutnoj vrednosti (R_{B-A}) a potom se svaki rang označi predznakom razlike ($\pm R_{B-A}$). Zbog ove poslednje operacije Vilkoksonov test je poznat i kao *Test označenih rangova* (engl. *Signed-ranks test*). U narednom koraku računaju se sume pozitivnih i negativnih rangova. Manja od dve apsolutne vrednosti dobijenih suma je vrednost T, čija se značajnost određuje na osnovu z vrednosti izračunate prema formuli:

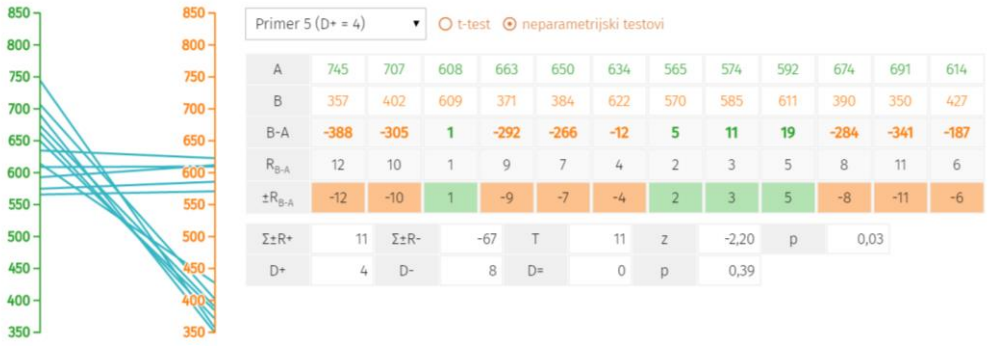
$$z = \frac{T - \frac{N(N+1)}{4}}{\sqrt{\frac{N(N+1)(2N+1)}{24}}}$$

U brojiocu formule je procena vrednosti aritmetičke sredine razlika apsolutnih vrednosti suma pozitivnih i negativnih rangova. U skladu sa nultom hipotezom, očekuje se da ta razlika ne postoji, odnosno da su sume rangiranih promena u pozitivnom i negativnom smeru jednake. U tom slučaju, vrednost brojioca u gornjoj formuli je nula. Vrednost imenioca predstavlja očekivanu varijabilnost, tj. standardnu devijaciju razlika suma rangova u uzorku, pod pretpostavkom da ta razlika u populaciji ne postoji. U našem primeru, vrednosti T i t upućuju na isti zaključak, tj. da razlika nije statistički značajna.

Odaberite *Primer 3* sa liste. Vilkoksonov test ukazuje na to da su razlike u ovom slučaju statistički značajne, jer je smer svih promena isti, a suma rangova negativnih razlika značajno je veća od sume pozitivnih, koja iznosi nula. Kao što se vidi u redu R_{B-A}, 11 od 12 razlika je potpuno isto, tako da deli prosek prvih 11 rangova – (1 + 2 + … + 11) : 11 = 6. Da podsetimo, na samom početku smo jednu promenu namerno načinili drugačijom da bi bilo moguće izračunati standardnu grešku razlike. Držeći pritisnut taster *Ctrl*, linearno menjajte rezultate svih ispitanika na trećem merenju naviše i naniže. Uočićete da se T i z vrednosti u većini pozicija uopšte ne menjaju zato što rang razlika ostaje isti. T je nula i kada se obrne smer promena jer je tada suma negativnih razlika nula. Postoje samo tri pozicije u kojima će T, odnosno z vrednosti biti nešto drugačije, ali je potrebna velika preciznost u njihovom pronalaženju. Prva je kada jednu razliku učinite nultom dok ostale razlike dele isti rang. T ostaje 0, ali z je nešto manje pošto se uzorak smanjio za jedan. Naime, uobičajeno je da se nulte

razlike prilikom računanja Vilkoksonovog testa ne uzimaju u obzir. Druga pozicija je upravo obrnuta, kada su sve promene, osim jedne, nulte. Tada je veličina uzorka 1, a razlika više nije statistički značajna. To je i logično, jer nijedan statistički metod „ne bi dozvolio" da se neki efekat proglasi značajnim na tako malom uzorku merenja. Na kraju, paralelne linije mogu da se postave i tako da postoji 1 negativna i 11 pozitivnih razlika, tako da vrednost T iznosi 1. Ona, naravno, nije dovoljno velika da bi razlika prestala da bude značajna.

Vratite se na treći primer i odaberite opciju *t-test*. Linearno pomerite sve linije tako da razlika između drugog i trećeg merenja (M_3 - M_2) bude približno -20 ms. Razlika je statistički značajna. Pomeranjem linije kojom je predstavljen najsporiji pilot, smanjite njegovu brzinu reakcije do vrednosti 850 ms na trećoj osi. Ovaj, očigledno aberantni rezultat, povećao je prosek i varijabilnost rezultata trećeg merenja a umanjio razliku između proseka drugog i trećeg merenja. Zbog toga je vrednost t-testa pala ispod granične vrednosti. Ukoliko rezultat istog ispitanika smanjujete i na kraju postavite na 350 ms, primetićete da t-test najpre raste, ali potom ponovo opada do vrednosti koja nije statistički značajna. Već smo rekli da ovoga puta nije reč samo o aberantnom rezultatu na trećem merenju već i o *aberantnoj promeni* rezultata. Sada prikažite Vilkoksonov T test i ponovite promene koje ste pravili, pomerajući rezultat istog ispitanika naviše do 850 ms, a potom naniže do 350 ms. Obratite pažnju na to da ovoga puta autlajer ne utiče na konačni zaključak da je razlika među merenjima statistički značajna. Kao što smo rekli, ova robusnost, odnosno otpornost na postojanje aberantnih rezultata, jedna je od osnovnih odlika neparametrijskih testova.



Slika 56. Postupak računanja vrednosti Testa označenih rangova i Testa predznaka

Test predznaka se takođe bazira na poređenju broja pozitivnih i negativnih promena između merenja, ali za razliku od Testa ekvivalentnih parova, ne uzima u obzir intenzitet razlika već samo njihov smer. To ga čini grubljim i manje snažnim. Odaberite *Primer 5* sa liste i opciju *t-test*. Iako vrednost t-testa ukazuje na postojanje značajne razlike između prvog i drugog merenja, na osnovu grafikona mogu da se uoče dve (pod)grupe ispitanika – jedne u kojoj je zaista došlo do povećanja brzine reakcije, i druge u kojoj je prosečna brzina ostala približno ista. U tom kontekstu, zaključak donet primenom parametrijske tehnike ne bi bio validan. Kada odaberete opciju *neparametrijski testovi*, uočićete da i Vilkoksonov test ukazuje na postojanje razlike, ali značajne samo na nivou 0,05 (Slika 56). Međutim, Test predznaka pokazuje da ova promena nije dovoljno dosledna da bi se smatrala statistički značajnom. Ovaj zaključak je donet na osnovu poređenja distribucije opaženih promena u oba smera, sa distribucijom koja bi se očekivala u skladu sa nultom hipotezom. Promene u pozitivnom smeru označene su sa D+, a promene u negativnom smeru sa D-. Kada razlike među merenjima ne bi bilo, očekivao bi se jednak broj promena u oba smera. U našem primeru, to znači da treba uporediti empirijski dobijene vrednosti 4 i 8 sa teorijskim frekvencijama 6 i 6. Poređenje može da se obavi na više načina. U većini statističkih programa, p nivo za Test predznaka određuje se kao verovatnoća ishoda u binomnoj distribuciji:

$$p = \frac{N!}{x!\,(N-x)!} \cdot 0{,}5^{x} \cdot 0{,}5^{N-x}$$

gde je N veličina uzorka, x ishod čiju verovatnoću računamo, a ! je oznaka za faktorijel broja. Vrednost 0,5 je teorijska verovatnoća pojedinačnih događaja, tj. verovatnoća promene u jednom ili u drugom smeru, pod pretpostavkom da je nulta hipoteza tačna. Kao i u slučaju Vilkoksonovog testa, parovi merenja kod kojih nije došlo do promene se ignorišu. Njihov broj je označen u tabeli simbolom D=. Dakle, Test predznaka pokazuje kolika je verovatnoća da slučajno dobijemo distribuciju promena koju smo opazili, ako pretpostavimo da razlike među merenjima u populaciji nema. U našem primeru, 0,39 označava verovatnoću da od 12 merenja, 4 ili manje bude u jednom smeru, odnosno da ih 8 ili više bude u drugom. Pošto ova vrednost nije manja od uobičajenih nivoa značajnosti, zaključujemo da bismo u približno 39 od 100 merenja mogli potpuno slučajno da dobijemo 4 i 8, čak i ako razlike nema, odnosno ako je broj promena u populaciji isti u oba smera. Treba naglasiti da je odabrani primer

poslužio samo kao ilustracija razlika u ishodima primene različitih statističkih metoda. To ne znači da je neki od tih ishoda uvek poželjniji, prihvatljiviji ili tačniji. Grafikon koji je prikazan u petom primeru ne treba da posluži istraživaču samo kao osnov za odlučivanje o tome koju statističku tehniku treba da primeni, već i kao signal da postoji atipičan obrazac u prikupljenim podacima. U ovom primeru, trebalo bi da se utvrdi razlog zbog koga su se merenja, tj. ispitanici, grupisali u dva očigledna klastera.

> Na koji način se još može proveriti značajnost odstupanja opaženih frekvencija 4 i 8 od očekivanih frekvencija 6 i 6?
>
> Pomerajte linije na grafikonu da biste utvrdili koliki broj promena bi se smatrao statistički značajnim ako je veličina uzorka 12.

U prethodnim odeljcima opisali smo neparametrijske metode kojima se obrađuju rezultati merenja ordinalnog nivoa. Napomenuli smo da se one mogu primenjivati i na podacima koji potiču sa intervalnog ili razmernog nivoa, ali da tada opada preciznost i snaga testa, jer se rezultati pre obrade automatski pretvaraju u rangove. To znači da bismo dobili iste vrednosti Spirmanovog koeficijenta korelacije ili Men–Vitnijevog U testa kada bismo ih izračunali na podacima intervalnog ili racio nivoa, kao i na istim podacima koji su prethodno rangirani. Međutim, u slučaju Vilkoksonovog testa i Testa predznaka, ova logika nije uvek opravdana, što može da zbuni istraživača a često i da dovede do neadekvatnih rezultata. Prikažite *Primer 3* i odaberite najpre prikaz t-testa, a potom i neparametrijskih testova. Sve metode upućuju na zaključak da je došlo do statistički značajne promene između drugog i trećeg merenja. Aritmetičke sredine se značajno razlikuju, kao i sume rangova pozitivnih i negativnih promena. Osim toga, sve promene su se desile u istom (negativnom) smeru, tako da je i Test predznaka statistički značajan. Sada zamislite da smo umesto brzine u milisekundama imali samo podatak o rangu ispitanika na oba merenja, odnosno informaciju o tome koji je od pilota brži od koga, ali ne i za koliko. Ovako izraženu brzinu možete da vidite ako držite pritisnut taster *R* na tastaturi dok su prikazani rezultati neparametrijskih testova. S obzirom na to da su rangovi ispitanika isti na prvom i na drugom merenju, a podatak o povećanju brzine u milisekundama više nije dostupan, oba neparametrijska testa sugerišu da razlike nema. Drugim rečima, ne može se reći da je došlo do promene, jer su nakon promene boje stimulusa brži piloti ostali brži a sporiji su ostali sporiji.

Odaberite *Primer 6* i prikažite vrednosti neparametrijskih testova. Na osnovu p nivoa možete da uočite da Test ekvivalentnih parova ima veću snagu od Testa predznaka, što je posledica primene postupka koji uzima u obzir ne samo smer promene već i njen intenzitet. Ukoliko, pak, prikažete rangirane podatke pritiskanjem tastera *R,* uočićete da nijedan od testova nije statistički značajan. Sličnu nedoslednost ćete primetiti i kada postupak ponovite na petom primeru. Uočite da u ovom slučaju p nivo za Test predznaka koji je izračunat na rangovima iznosi 1, jer je potpuno sigurno da će u uzorku od 12 parova merenja, 6 ili manje, odnosno 6 ili više promena, biti istog smera. Na prvi pogled, ovaj i prethodni primeri mogu da deluju nelogično, jer su u pitanju metode namenjene upravo obradi podataka sa ordinalnog nivoa merenja. Međutim, potrebno je razumeti da u slučaju uparenih rezultata mora da postoji način da se izrazi apsolutna promena. Ako podatke oba merenja rangiramo nezavisno, ta promena više nije uočljiva. To bi bilo analogno primeni t-testa za nezavisne uzorke na podacima koji su prethodno transformisani u z skorove za svaku grupu posebno, čime bi aritmetičke sredine obe grupe dobile vrednost 0. Da bi primena opisanih neparametrijskih postupaka bila opravdana i na rangovima, rezultati oba merenja moraju da se rangiraju objedinjeno. U našem primeru, to bi značilo da se umesto dva niza u kojima se rangovi kreću od 1 do 12, formiraju nizovi u kojima se rangovi kreću od 1 do 24. Ovaj način rangiranja biće prikazan kada držite taster *Q* na tastaturi. Prikažite rezultate za primere 3 do 6 na ovaj način i obratite pažnju na to da se rezultati Testa predznaka ne menjaju u odnosu na podatke razmernog nivoa, dok su rezultati Vilkoksonovog testa neznatno drugačiji zbog manjeg raspona rezultata. Opisani primeri treba da posluže kao opomena da u slučaju primene neparametrijskih testova za zavisne uzorke, *rangovi* nemaju uvek isti smisao i značenje kao *rangirani podaci* intervalnog ili racio nivoa. Zbog toga većina autora preporučuje da se Vilkoksonov test koristi kao alternativa t-testu kada zavisna varijabla nije normalno distribuirana, ali je ipak izmerena na najmanje intervalnom nivou.

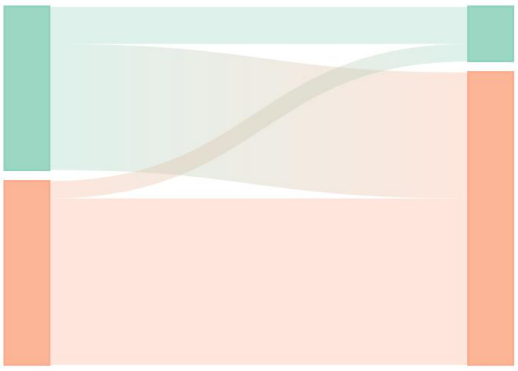
> Da li bi se rezultati t-testa bitno izmenili kada bi se on primenio na podacima koji su pretvoreni u rangove na gore opisani način, odnosno objedinjeno za obe grupe ispitanika ili oba merenja?

3.10. Značajnost razlika uparenih podataka nominalnog nivoa

Ukoliko su „merenja" varijabli obavljena na nominalnom nivou, primena prethodno opisanih neparametrijskih alternativa t-testu nije moguća. Osim toga, čak i ako su varijable ordinalnog nivoa ali imaju mali broj mogućih vrednosti, testiranje hipoteza o postojanju razlika između uparenih merenja bi trebalo da se obavlja pomoću tehnika baziranih na tabelama kontingencije, odnosno na hi-kvadrat testu. U vezi sa tim, u poglavlju 3.5.4. naglasili smo da je jedan od bitnih uslova za računanje hi-kvadrat testa nezavisnost opservacija. U slučaju zavisnih uzoraka, ovaj uslov ne može da bude ispunjen, pa je potrebno primeniti drugačije postupke za računanje χ^2 vrednosti. U ovom odeljku ćemo opisati neke od najčešće korišćenih statističkih postupaka ove vrste. Za grafičko predstavljanje podataka koristićemo ***dijagram toka***, sličan onome koji je napravio Šarl Minar da bi ilustrovao napredovanje Napoleonove vojske u pohodu na Rusiju. U primeru *Energetska pića* prikazani su efekti kampanje informisanja o zdravoj ishrani, koja je sprovedena u nekoliko srednjih škola u toku jednog polugodišta (Slika 57). Đaci su na početku i na kraju polugodišta popunjavali kratak upitnik o navikama u ishrani, pri čemu su odgovarali i na pitanje da li konzumiraju energetska pića. Na desnoj strani interaktivnog okvira prikazana je tabela kontingencije koja je formirana ukrštanjem vrednosti prvog i drugog merenja. Na levoj strani nalazi se dijagram na kome boja i visina pravougaonika označavaju učestalost odgovora DA i NE, a širina traka broj đaka u svakoj od četiri kategorije nastale ukrštanjem: DA (početak) – DA (kraj), DA – NE, NE – DA i NE – NE. Ovu vrstu grafikona popularizovao je irski inženjer *Metju Henri Senki* vizualizujući **efikasnost parne mašine**, odnosno pravce protoka energije prilikom njenog rada. Stoga su dijagrami toka poznati i kao *Senkijevi dijagrami*. U literaturi se ponekad upotrebljava i termin *aluvijalni dijagrami* zato što isprepletane trake različitih širina, kojima se prikazuje veličina promene između dve ili više tačaka, podsećaju na aluvijalne ravni, tj. rečne nanose (lat. *alluvio*) nastale izlivanjem ili promenom rečnog toka.

> Koliko kolona i koliko redova ima matrica sirovih podataka iz koje je nastala prikazana tabela kontingencije? Kako biste nazvali kolone u matrici?

Zbog čega tabela kontingencije u ovom primeru ne bi mogla i ne bi trebalo da se formira ukrštanjem vremena merenja (PRE / POSLE) i odgovora na pitanje (DA / NE)?



Energetska pića ▾ T = 94; z_T = 8,44; p < 0,01

		KRAJ		κ = 0,14
		DA	NE	Σ
POČETAK	DA	32	109	141
	NE	15	144	159
	Σ	47	253	300

Slika 57. Primena dijagrama toka za vizuelizaciju povezanosti dva merenja nominalnog nivoa (dihotomna varijabla)

3.10.1. Maknimarov test

Na osnovu visine pravougaonika na levoj strani Senkijevog dijagrama za primer *Energetska pića*, može se zaključiti da je broj đaka koji su konzumirali energetska pića pre kampanje, bio gotovo jednak broju onih koji nisu. Taj broj je vidljivo opao u drugom merenju, tj. na kraju polugodišta. Najširа traka na grafikonu prikazuje kategoriju NE – NE, odnosno 144 đaka koji ni pre ni posle kampanje nisu pili energetska pića. Sledeća kategorija po veličini je grupa od 109 đaka koji su u toku polugodišta prestali da konzumiraju energetska pića (DA – NE). Širina ove trake sugeriše da je kampanja verovatno uticala na promenu ponašanja, ali da bismo utvrdili da li je ova promena i statistički značajna, potrebno je izračunati χ^2 test. Pri tome nas ne interesuje da li je svaki pojedinačni ispitanik promenio svoj stav, već da li su se promenile proporcije odgovora DA i NE u drugom merenju. To znači da nam nisu bitne ćelije tabele u kojima se nalazi broj đaka koji nisu promenili stav (DA – DA i NE – NE), nego one koje govore o promeni. Ako ćelije sleva nadesno i od gore ka dole označimo slovima a, b, c i d, to su frekvencije u ćelijama b (promena iz DA u NE) i c (promena iz NE u DA). Kada ne bi bilo promene u proporciji odgovora između prvog i drugog merenja, pravougaonici sa leve i desne strane bili bi iste visine, a vrednosti u ćelijama b i c bile bi jednake. Pošto je u našem primeru 109 + 15

= 124 đaka promenilo navike, u skladu sa nultom hipotezom bismo očekivali da ih je 124 : 2 = 62 prešlo iz grupe DA u grupu NE, a 62 iz grupe NE u grupu DA. Značajnost odstupanja opaženih vrednosti 109 i 15 od teorijskih 62 i 62 možemo da testiramo u odnosu na verovatnoće ishoda u binomnoj distribuciji, što je analogno ranije opisanom Testu predznaka. Alternativno, može se primeniti i opšta formula za χ^2 test:

$$\chi^2 = \sum \frac{(f_o - f_t)^2}{f_t}$$

U slučaju zavisnih uzoraka, vrednosti f_o su frekvencije u ćelijama b i c, a f_t očekivane frekvencije koje se izračunavaju kao (b + c) : 2. Kada ove podatke uvrstimo u gornji izraz i svedemo ga na jednostavniju formu, dobijamo formulu za izračunavanje hi-kvadrat testa za zavisne uzorke. Nju je predložio američki psiholog Kvin Maknimar (McNemar, 1947), te se stoga ovaj postupak često naziva i *Maknimarov χ^2 test*:

$$\chi^2 = \frac{(b - c)^2}{b + c}$$

Slova b i c u gornjoj formuli označavaju frekvencije ćelija nastalih ukrštanjem diskordantnih vrednosti varijable. To ne moraju da budu druga i treća ćelija tabele. Da smo tabelu napravili tako da je prva kolona bila NE, a druga DA, tada bi o promeni govorile ćelije a i d. Ako uz to primenimo i Jejtsovu korekciju, koja se preporučuje u slučaju tabela dimenzija 2 x 2, dolazimo do najčešće korišćene formule za Maknimarov χ^2 test sa korekcijom za kontinuitet:

$$\chi^2 = \frac{(|a - d| - 1)^2}{a + d}$$

Širinu trake na prikazanom Senkijevom grafikonu možete da menjate tako što je kliknete, zadržite pritisnut taster miša i pomerate pokazivač nagore ili nadole. Na opisani način postavite frekvenciju ćelije b (DA - NE) u primeru *Energetska pića* na 15. Sada su veličine promena u oba pravca iste, što znači da se one međusobno potiru i da ne postoji značajna promena na nivou grupe. Prikazana χ^2 vrednost ipak nije nulta zato što je za njeno izračunavanje primenjena formula sa Jejtsovom korekcijom. Sume redova i kolona, odnosno marginalne frekvencije, jednake su po kategorijama odgovora – 159 đaka je odgovorilo NE a 47 đaka DA, kako na početku, tako i na kraju polugodišta. Koristeći statističku terminologiju, može se reći da su distribucije marginalnih

frekvencija homogene. Otuda naziv *testovi marginalne homogenosti* za grupu metoda kojima se testira značajnost promena uparenih rezultata merenja nominalnog nivoa kojima pripada i Maknimarov test. Smanjite širinu najšire trake na Senkijevom dijagramu, odnosno frekvenciju ćelije NE – NE, na 5 i obratite pažnju na dve stvari. Prvo, ova promena uopšte nije uticala na vrednost χ^2 testa jer je marginalna homogenost ostala očuvana. Drugo, sve ostale trake na dijagramu postale su šire, ali ne zato što se povećala frekvencija odgovarajućih ćelija, već zato što se povećala njihova *relativna proporcija* u odnosu na ukupnu veličinu uzorka. U tom smislu, treba obratiti pažnju na to da Senkijev dijagram pruža informacije o relativnom odnosu među kategorijama, ali ne o broju merenja ili veličini uzorka, slično kao kružni (torta) dijagram. Sa druge strane, ako povećavate učestalost u ćeliji c (NE - DA), primetićete da sa sve većim narušavanjem marginalne homogenosti, vrednost χ^2 testa postaje sve veća a razlika sve značajnija. Prostije rečeno, proporcija odgovora više nije ista ako se uporede prvo i drugo merenje.

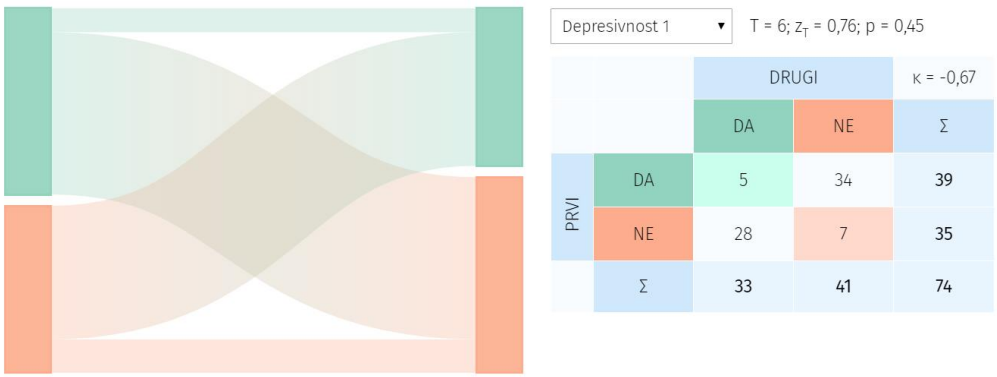
> Kako bi trebalo izmeriti konzumaciju energetskih pića da bi se omogućila primena t-testa za zavisne uzorke? Objasnite zbog čega bi u toj situaciji Maknimarov test imao manju snagu od t-testa za zavisne uzorke?
>
> Pomeranjem traka napravite primer u kome je 200 od 210 ispitanika promenilo svoje ponašanje ali razlika između prvog i drugog merenja nije statistički značajna. Da li je ovaj rezultat, po vašem mišljenju, logičan i opravdan?
>
> Da li bi neki od postupaka pomenutih u ranijim odeljcima ipak mogao da ukaže da u primeru koji ste formirali postoji statistički značajna pravilnost?

3.10.2. Koenova kapa

Odaberite primer *Depresivnost 1*. U grupi ispitanika primenjena su dva upitnika kojima se dijagnostikuje depresivni poremećaj. U pitanju su upareni rezultati, tako da je opravdano primeniti nacrt za zavisne uzorke, odnosno ponovljena merenja. Kao što smo više puta naglasili, neparametrijski testovi mogu da se računaju i na varijablama višeg nivoa, ako se njihove vrednosti pre toga transformišu u rangove ili kategorije. Na primer, skorovi na upitniku mogu da se *veštački dihotomizuju*, podelom ispitanika u grupe onih koji imaju kliničke

simptome depresije (DA) i onih kod kojih ti simptomi nisu izraženi u značajnoj meri (NE). Ovoga puta nas interesuje da li postoji razlika u proceni simptoma depresije na osnovu dva upitnika. Kada bismo za proveravanje ove hipoteze upotrebili Maknimarov test, zaključak bi bio da nema razlike, zato što se u diskordantnim kategorijama (DA – NE i NE – DA) nalazi podjednak broj ispitanika, te je vrednost χ^2 niska. Ovaj zaključak je, naravno, neopravdan, jer je očigledno da su dijagnoze donete uz pomoć dva upitnika zapravo potpuno suprotne, ako se posmatraju ishodi na nivou pojedinaca (Slika 58). Ova nelogičnost proizilazi iz činjenice da Maknimarov test govori o proporciji događaja na nivou grupe, a ovoga puta je potreban podatak o *saglasnosti* dva upitnika na nivou individualnih promena. Jedan od najčešće korišćenih indikatora saglasnosti je *Koenova kapa* (Cohen, 1960). Kapa koeficijent označava se grčkim slovom κ i pokazuje stepen slaganja dva procenjivača ili dva instrumenta za procenu, na varijablama nominalnog nivoa. Zamislimo da su procenu postojanja simptoma depresije obavila dva klinička psihologa na osnovu intervjua sa klijentom. Kapa koeficijent računa se tako što se porede opažene i očekivane frekvencije u jednoj od dijagonala tabele kontingencije, obično glavnoj, tj. onoj koja se proteže od gornje leve do donje desne ćelije:

$$\kappa = \frac{\sum f_o - \sum f_t}{N - \sum f_t}$$



Slika 58. Primer u kome Maknimarov test sugeriše da nema razlike između dva merenja dok Koenova kapa pokazuje da ne postoji saglasnost

U našem primeru, vrednosti f_o su 5 i 7, a vrednosti f_t se dobijaju na način koji smo opisali u odeljku o hi-kvadrat testu. Rezultat je broj između -1 i 1 koji se interpretira na sličan način kao i drugi koeficijenti povezanosti. Smernice za interpretaciju Koenove kape date su u Tabeli 4. U primeru sa procenom

depresivnosti, vrednost κ je manja od nule, što znači da saglasnost ne postoji. Štaviše, njena apsolutna vrednost toliko je visoka, da bismo mogli da kažemo da postoji visok stepen *neslaganja* među procenama, iako je Maknimarov test pokazao da razlika među merenjima ne postoji. Slična prividna nelogičnost postoji i u primeru *Depresivnost 2*. Maknimarov test upućuje na postojanje značajnih razlika među procenama, ali uvidom u tabelu kontingencije i vrednost kapa koeficijenta, ipak možemo da zaključimo da je slaganje relativno visoko, uzimajući u obzir ukupnu veličinu uzorka. Ako dodatno povećavate frekvencije u ćelijama a i d, proširivanjem najširih traka na Senkijevom dijagramu, primetićete da vrednost κ raste.

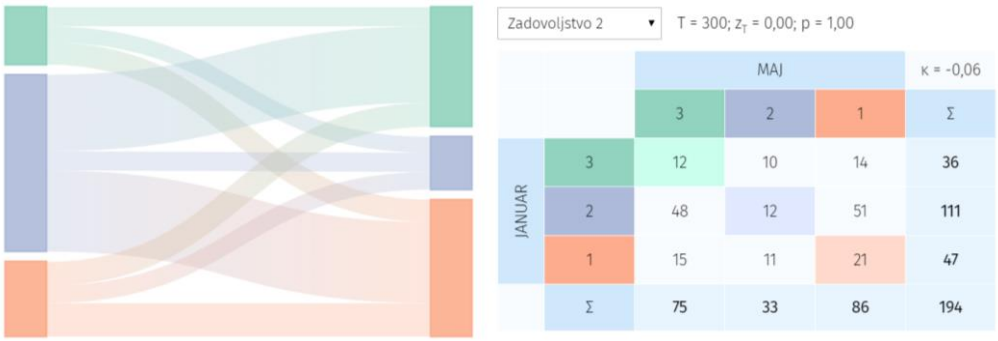
Tabela 4. Smernice za interpretaciju visine kapa koeficijenta saglasnosti

Vrednost κ	Interpretacija κ	Koeficijent determinacije
< 0,00	saglasnost manja od slučajne	
0,01–0,20	beznačajna saglasnost	0–4%
0,21–0,39	minimalna saglasnost	4–15%
0,40–0,59	slaba saglasnost	16–35%
0,60–0,79	umerena saglasnost	36–63%
0,80–0,90	visoka saglasnost	64–81%
0,91–1,00	(gotovo) potpuna saglasnost	82–100%

3.10.3. Testovi marginalne homogenosti za politomne varijable

Maknimarov test namenjen je poređenju dveju dihotomnih varijabli. Ukoliko varijable imaju više od dva nivoa, potrebno je primeniti neki od testova marginalne homogenosti za tabele kontingencije većih dimenzija. U ovom odeljku ćemo opisati dva postupka koji predstavljaju prikladne predstavnike različitih pristupa analizi homogenosti marginalnih distribucija u tabelama kontingencije većim od 2 x 2. Postupke njihovog izračunavanja nećemo prikazivati, jer zahtevaju primenu nešto složenijih procedura matrične algebre. Odaberite primer *Zadovoljstvo 1*. U pitanju su rezultati ankete o zadovoljstvu kupaca rasporedom proizvoda u lancu supermarketa. Distribucija marginalnih

frekvencija odgovora prikupljenih u januaru (poslednja kolona) govori da je najveći broj potrošača bio neopredeljen (53), a da je podjednak broj njih ocenio raspored artikala pozitivno (21), kao i negativno (23). Ova distribucija se očigledno razlikuje od marginalne distribucije odgovora u maju (poslednji red) kada je najviše kupaca dalo negativnu ocenu. Dakle, značajan broj anketiranih osoba promenio je mišljenje i „prešao" iz kategorije neodlučnih u kategoriju nezadovoljnih. Ova promena vidi se kao najširа traka na dijagramu. Vrednost χ^2 testa iznosi 20,24, tako da je razlika statistički značajna.



Slika 59. U prikazanom primeru može se doći do potpuno suprotnih zaključaka u zavisnosti od toga da li se varijable tretiraju kao kvalitativne ili kvantitativne

Za računanje χ^2 vrednosti u prethodnom primeru, primenjen je *Stjuart–Maksvelov test* (Maxwell, 1970; Stuart, 1955) koji predstavlja generalizaciju Maknimarovog testa na tabele kontingencije dimenzija većih od 2 x 2. To znači da Maknimarov i Stjuart–Maksvelov test imaju iste vrednosti kada se izračunaju na paru dihotomnih varijabli, odnosno merenja. Međutim, treba obratiti pažnju da se za obradu uparenih rezultata merenja na kategorijalnim varijablama upotrebljavaju i algoritmi koji tretiraju odgovore kao rangove (Agresti, 1983), što može bitno da utiče na ishod zaključivanja o postojanju razlike. Ilustrovaćemo ovaj problem primerom *Zadovoljstvo 2*. Ponovo su u pitanju odgovori anketiranih potrošača, ali smo ih ovoga puta označili brojevima: 3 – zadovoljan, 2 – neodlučan i 1 – nezadovoljan (Slika 59). Distribucije marginalnih frekvencija očigledno su potpuno drugačije, tako da Stjuart–Maksvelov test pokazuje da je razlika između odgovora iz januara i onih koji su prikupljeni u maju, statistički značajna. Međutim, ako se primeni algoritam koji se primenjuje u nekim od popularnih statističkih paketa (IBM, 2016), zaključak može da bude potpuno suprotan. Dok držite pritisnut taster *Ctrl*, umesto Stjuart–Maksvelovog χ^2 testa biće vidljiva T vrednost ili tzv. *MH*

statistik. MH statistik predstavlja neku vrstu aritmetičke sredine marginalnih frekvencija i ukazuje na to da razlika nije statistički značajna. Naime, ukoliko izračunamo aritmetičke sredina marginalnih frekvencija odgovora, uočićemo da su one iste – $(36 \cdot 3 + 111 \cdot 2 + 47 \cdot 1) : 194 = (75 \cdot 3 + 33 \cdot 2 + 86 \cdot 1) : 194 \approx 1{,}94$. Iako se distribucije odgovora potpuno razlikuju, njihove aritmetičke sredine su iste. Dakle, oba zaključka su opravdana i imaju potporu u rezultatima statističke obrade. Ali odluku o tome koji od njih je smisleniji, mora da donese istraživač u kontekstu daljih implikacija tog zaključka. U ovom primeru, verovatno je opravdaniji zaključak da se struktura odgovora ispitanika značajno izmenila u periodu od januara do maja.

> Menjajte frekvenciju kategorija 1 – 1 a potom i 1 – 2 u primeru *Zadovoljstvo 2*. Posmatrajte kako promene utiču na vrednost χ^2, odnosno T testa. Zbog čega promene frekvencija u glavnoj dijagonali tabele kontingencije ne utiču na vrednosti testova homogenosti marginalnih frekvencija?

U vezi sa prethodno pomenutim situacijama u kojima dva testa mogu da dovedu do potpuno suprotnih zaključaka, trebalo bi naglasiti da ranije opisani kapa koeficijent nije koeficijent korelacije poput C ili ϕ koeficijenata, iako na to upućuje raspon njegovih vrednosti. Odaberite primer *Agresivnost* sa liste. U pitanju je tabela kontingencije koja prikazuje saglasnost roditelja u proceni agresivnosti njihove dece na skali od 1 do 3. Činjenica da roditelji ocenjuju isto dete, omogućava uparivanje njihovih odgovora i primenu postupaka namenjenih poređenju zavisnih uzoraka. Kapa koeficijent od 0,16 ukazuje na zanemarljivo slaganje roditelja u proceni agresivnosti deteta, jer se veliki broj parova rezultata nalazi van glavne dijagonale. Sa druge strane, vrednost χ^2 testa je statistički značajna, što na prvi pogled nije u suprotnosti sa prethodnim zaključkom, pošto je opravdano reći da postoji razlika u proceni agresivnosti deteta od strane oba roditelja. Međutim, to istovremeno pokazuje da u tabeli kontingencije postoji određena pravilnost, tj. veza između odgovora očeva i odgovora majki. Intenzitet te veze može da se iskaže C koeficijentom kontingencije, koji bi u ovom slučaju iznosio oko 0,42 i bio bi statistički značajan. Dakle, iako ne postoji saglasnost između ocena roditelja, može se reći da je njihova korelacija umerena. Ukoliko podrobnije analiziramo frekvencije u tabeli kontingencije, uočićemo da ćelije 1 – 2, 2 – 3 i 3 – 3 imaju najveće učestalosti. Povezanost odgovora roditelja zapravo se ne ogleda u slaganju, odnosno

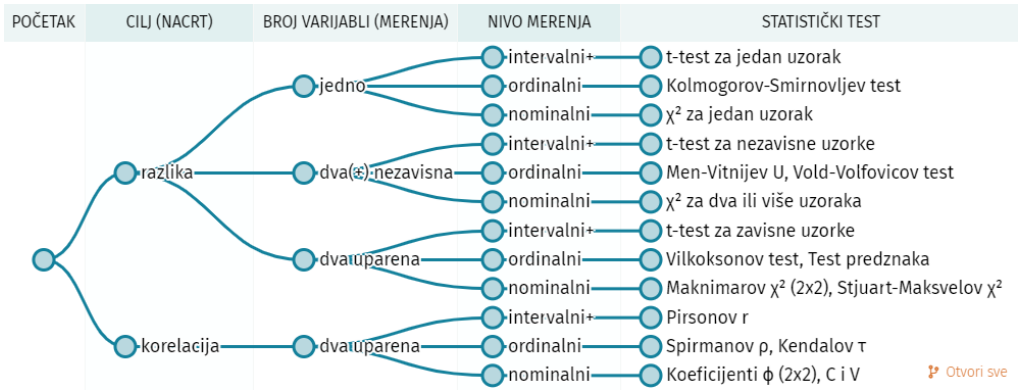
sličnosti njihovih procena, već u pojavi da majke sistematski daju nešto više ocene od očeva. Moguće objašnjenje ovakvog rezultata je to što su očevi možda u većoj meri tolerantni na agresivno ponašanje svoje dece ili su majke osetljivije po tom pitanju. Ovo je još jedna ilustracija koja govori u prilog upozorenju da podatke treba analizirati iz više uglova i na različite načine, kako bi se doneo potpun i validan zaključak o fenomenima koji se istražuju.

ZAVRŠNE NAPOMENE

Čitaocima je sigurno poznata izreka da jedna slika vredi više od 1.000 reči. Nadamo se da smo ovim udžbenikom pokazali da grafikoni i dijagrami vrede više od 1.000 p vrednosti, kao što je to rekao američki psiholog Džefri Loftus (Loftus, 1993). Sa jedne strane, cilj nam je bio da buduće istraživače upoznamo sa prednostima različitih tehnika vizualizacije podataka, a sa druge, da ih motivišemo i obučimo da pojavama koje žele da razumeju, pristupaju odgovorno i kompetentno sa statističkog i metodološkog aspekta. U smislu obuhvatnosti udžbenika, oblast inferencijalne statistike samo je načeta, ali su obrađeni pojmovi i koncepti koji su presudni za razumevanje i primenu većine složenijih statističkih metoda. Završna poruka čitaocu mogla bi da glasi da je uz pomoć statistike moguće uspešno dati odgovor na (istraživačko) pitanje, samo ako je to **pitanje ispravno, mudro i precizno postavljeno**. Da bi se pitanje precizno postavilo, potrebno je na adekvatan način prikupiti, sažeti i opisati podatke. Nakon toga izbor odgovarajuće statističke metode može da se shvati kao kretanje kroz **hijerarhijsko stablo odlučivanja** u kome se daju odgovori i prave kompromisi između pitanja o tome šta želimo da postignemo i šta nam podaci omogućavaju, odnosno „dozvoljavaju" (Slika 60). Kliknite kružić koji se nalazi u koloni POČETAK. Ako se zadržimo na skupu tehnika koje su obrađene u ovom udžbeniku, prva odluka koju treba doneti jeste da li želimo da utvrdimo postojanje razlika ili postojanje povezanosti među merenjima. Recimo da nam je cilj ovo prvo. Dalji tok zavisi od odgovora na pitanje da li smo neku varijablu izmerili jednom, dvaput u različitim grupama ili dvaput u istoj grupi, odnosno tako da rezultate merenja možemo da uparimo. Odaberite poslednju opciju. U zavisnosti od toga na kom nivou je obavljeno merenje, odlučujemo koja statistička metoda je najprikladnija za obradu podataka. Na primer, za nominalne varijable, to bi bio neki od testova marginalne homogenosti. Ako krenemo od drugačijeg cilja, npr. da odredimo povezanost među merenjima, jasno je da nemamo opciju da dovodimo u vezu dva potpuno nezavisna merenja, tako da će izbor testa biti određen isključivo nivoom merenja varijabli.

Prolazeći kroz prikazani dijagram i stablo odlučivanja, dolazi se do nekog od petnaestak bazičnih testova obrađenih u ovom udžbeniku. Skup statističkih metoda koje će istraživaču biti korisne i potrebne je, naravno, mnogo veći, ali apsolutni uslov za njihovo savladavanje jeste poznavanje i

razumevanje pojmova i koncepata koje smo opisali u prethodnim poglavljima. Na primer, ukoliko u istraživanju treba uporediti više od dve grupe ispitanika, prikladan izbor ne bi bila kombinacija t-testova već tehnika koja se zove *analiza varijanse (ANOVA)*. U tom slučaju, statistička značajnost razlika interpretira se na osnovu poređenja varijansi grupa, tj. odgovarajućih vrednosti F testa koji smo opisali u odeljku o statističkim distribucijama. Uz pomoć tehnika kao što su *kanonička korelaciona analiza* ili *višestruka regresiona analiza*, analizira se povezanost više varijabli istovremeno, ali i tada je bitno da se razume smisao standardizovanih (β) regresionih koeficijenta i logika linearnih jednačina. Logika regresione jednačine nalazi se i u osnovi *faktorske analize* i *analize glavnih komponenata*, tehnika bez kojih teško mogu da se zamisle istraživanja strukture ličnosti i (intelektualnih) sposobnosti. Uz pomoć χ^2 testa može da se proceni adekvatnost rezultata dobijenih nekom od tehnika *strukturalnog modeliranja*, kojima se opisuje latentni prostor skupa manifestnih varijabli. I tako dalje. U tom smislu, *obrazovni* cilj ovog udžbenika nije bio da čitalac savlada što veći broj tehnika i metoda kako bi ih (mehanički) primenjivao, već da nauči da te tehnike primenjuje sa razumevanjem. Štaviše, mnogo je važnije da udžbenik ostvari svoj *vaspitni* cilj, tako što će pažnju (budućih) istraživača usmeriti na podatke i njihov smisao, a ne na metode, rezultate i pokazatelje koji treba da impresioniraju ili zadovolje nastavnike, recenzente, kolege, političare ili neku drugu populaciju potencijalnih korisnika. Upravo zbog toga, smatramo da je vizualizacija pravi način da istraživač pokaže svoje statističko obrazovanje i vaspitanje. Brojevi i reči su (još uvek) neizbežna alatka za *opisivanje* pojava, ali vizualizacija postaje nezamenjiva u procesu njihovog razumevanja i *otkrivanja*.



Slika 60. Stablo odlučivanja o primeni inferencijalnih statističkih metoda opisanih u ovom udžbeniku

PRILOZI

Prilog 1: Odgovori na pitanja za vežbu

1.2. Vizuelna komunikacija

> Da li vam je lakše da na osnovu teksta formirate traženi vizuelni prikaz ili da na osnovu vizuelne metafore sastavite tekst koji opisuje grupe studenata?

Očekuje se da ćete lakše i efikasnije obaviti drugi zadatak. Informacije u tekstu navedene su sekvencijalno, pa se tako i slika formira u nekoliko koraka, sve dok se ne dobije prikaz koji zadovoljava sve zadate kriterijume. Sa druge strane, vizuelni prikaz nudi celovit (holistički) pristup informacijama, tako da treba samo izdvojiti one koje su bitne. Verovatno vam je teže da naučite navedeni tekst, nego da upamtite sliku i da je koristite kao osnovu za generisanje „priče", tj. opisa svih relacija vizualizovanih objekata. Slika olakšava izdvajanje informacija koje su (trenutno) bitne, ignorišući one koje su irelevantne.

> Potrebno je da utvrdite da li postoje studenti koji su odabrali sva tri kursa. Da li ćete taj zaključak lakše doneti na osnovu prikazane slike ili na osnovu pročitanog teksta?

I ovaj zadatak je lakše obaviti uz pomoć slike. Dovoljno je da se proveri da li postoji deo površine koju prekrivaju sva tri kruga. Zaključak da taj presek postoji mnogo je teže doneti na osnovu teksta. U pokušaju da odgovorimo na postavljeno pitanje, formiramo više različitih modela koji su kompatibilni sa ponuđenim premisama. Taj proces opterećuje radnu memoriju, što negativno utiče na njegovu brzinu i efikasnost, jer generišemo i modele koji su netačni. U slučaju vizuelnog modela, manipulišemo vizuelnim metaforama i na taj način rasterećujemo memoriju. To nam omogućava da lakše testiramo alternativne

mogućnosti i odbacimo one koji su netačni. Ukratko, slika nam pomaže u procesu rasuđivanja (Bauer & Johnson-Laird, 1993).

> Da li vam je lakše da poredite veličine kružnica ili veličine njihovih preseka?

Sigurno ste bili u prilici da prilikom kupovine procenjujete zapreminu različitih pakovanja prehrambenih ili hemijskih proizvoda. Svesni ste da izdužene flaše sadrže istu, a često i manju zapreminu tečnosti, ali one jednostavno deluju kao veće. Istraživanja pokazuju da prilikom procene razlika u površini i zapremini objekata, često imamo tendenciju da pojednostavljujemo zadatak tako što se fokusiramo na jednu, dominantnu dimenziju objekta (Krider, Raghubir, & Krishna, 2001). U našem primeru, lakše je proceniti veličinu kružnice na osnovu njenog prečnika, nego presek koga određuju dve dimenzije. Ovakvi zadaci se obavljaju sporije kada je potrebno kombinovati više elementarnih procena.

> Šta predstavlja horizontalna a šta vertikalna dimenzija ove slike?

Horizontalna dimenzija bi predstavljala vreme, odnosno vremenske tačke u kojima je poređena proizvodnja dve vrste voća. Vertikalna dimenzija je obim proizvodnje.

> Kako biste linije učinili jasnijim i informativnijim?

Linije se preklapaju, tako da ih je potrebno na neki način vizuelno odvojiti da bi se naglasio kontinuitet svake od njih. To se može uraditi promenom debljine, boje i/ili teksture linije. Kliknite taster *Oboji* da biste videli jedno od rešenja.

> Da li biste rekli da konačna slika prikazuje šest linija ili samo dve?

Šest linija sa početka vežbe sada su spojene „zajedničkom sudbinom“ tako da je ispravnije reći da su u pitanju dve izlomljene linije.

> Šta nedostaje slici da bi bila potpuno razumljiva i nekome ko nije pročitao ovaj tekst?

Slici nedostaje *legenda*, odnosno tumačenje upotrebljenih simbola i vizuelnih karakteristika, npr. boje linija. Poželjno je i da se na odgovarajući način označe i imenuju horizontalna i vertikalna dimenzija slike.

1.3.1. Različiti aspekti vizuelne pismenosti

> Na osnovu čega možete da zaključite da su neki od elemenata na slici naknadno izmenjeni?

Očigledno je da se neki elementi na slici ponavljaju u istoj formi.

> Kako biste u obrazloženju odgovora na prethodno pitanje upotrebili termin „verovatnoća"?

Verovatnoća da se različiti elementi na slici pojave u identičnom obliku, izuzetno je mala. Praktično je nemoguće da izduvni gasovi raketa formiraju iste oblike dima i prašine koji se uočavaju na slici.

> Na koji način tipografija utiče na čitljivost teksta?

Važnost čitljivosti teksta prepoznata je davno pre pojave digitalnih medija (Dale & Chall, 1949). Ključna dimenzija čitljivosti je brzina, tj. efikasnost kojom se tekst procesira. Jasno je da brzina zavisi od opšte „prijatnosti" koju čitalac oseća u toku čitanja. Tekst koji je ispisan malim fontom ili jarkim bojama negativno utiče na čitljivost, kao i previše zbijena ili previše razmaknuta slova.

> Uporedite svoje vizuelno rešenje sa stilovima koje je postavio neko drugi ili sa tekstovima koje srećete na internetu? U kojim aspektima tipografije postoji slaganje a u kojima ne?

Vrlo je verovatno da će slaganje postojati u veličini fonta, visini linije i boji, jer su u pitanju bazične karakteristike stimulusa. Sa druge strane, vrsta fonta ili postojanje serifa, najverovatnije se opažaju drugačije u zavisnosti od iskustva i ličnih preferencija čitalaca (McCarthy & Mothersbaugh, 2002).

> Da li je značenje piktograma uvek nedvosmisleno?

Ne. Razumljivost piktograma u velikoj meri se zasniva na njihovoj intuitivnosti, ali u velikoj meri zavisi od konteksta i prethodnih iskustava korisnika. Prva ikonica u četvrtom pitanju asocirala bi korisnika na šrafove kada bi se nalazila pored šrafcigera. Međutim, u nekom medicinskom priručniku ili uz ikonicu koja prikazuje špric, verovatnije je da bi se interpretirala kao prikaz tableta.

> Zbog čega u petom pitanju nije očigledno koji odgovor je tačan?

Tačan odgovor u ovom pitanju ne postoji. Reč je o apstraktnom pojmu koji za svaku osobu verovatno ima drugačije značenje. Ovim pitanjem zapravo nismo proveravali poznavanje značenja piktograma, već percepciju značenja pojma „prestiž" kod ispitanika.

> Šta označava upitnik a šta zelena štrikla u gornjem desnom uglu okvira?

Upitnik označava da niste odgovorili na pitanje. Zelena štrikla prikazuje se kada date odgovor na pitanje.

> Ako je na pitanja u upitniku odgovorilo 100 ispitanika koji su za svaki tačan odgovor dobijali po jedan poen, kako bi izgledao piktografik kojim se vizualizuje njihov uspeh izražen zbirom bodova?

S obzirom na to da se test sastojao od 5 pitanja, minimalan broj bodova koji je neko mogao da osvoji je 0 a maksimalan 5. Piktografik bi, dakle, imao 6 redova označenih ciframa 0 do 5. U svakom redu, piktogramima bi bio prikazan broj studenata koji su osvojili određeni broj bodova. Suma brojeva po kategorijama odgovora, odnosno redovima, treba da bude 100.

> Koje je značenje poruke ispisane na početnoj veb stranici ovog udžbenika?

„Dobrodošli u otvoreni univerzitetski udžbenik iz oblasti statistike, koji će vam pomoći da vidite i lakše razumete kako od podataka nastaju ideje". Poruku će

teže protumačiti osobe koje nisu upoznate sa serijalom „Zvezdane staze“. Sa druge strane, većini čitalaca verovatno je jasna asocijacija na univerzitet, iako u našoj sredini nije uobičajena tradicija uniformi i bacanja studentskih kapa.

1.4. Karta, mapa, dijagram, grafik, infografik

> Na koji način je prikazano više smrtnih slučajeva na istoj adresi?

Zumirajte mapu da biste videli više detalja. Svaka osoba predstavljena je crticom, a „nagomilane“ crtice formiraju stubić koji govori o učestalosti smrtnih slučajeva na različitim lokacijama.

> Na osnovu čega je Džon Snou zaključio gde se nalazi izvor zaraze?

Oko jedne od pumpi očigledno ima mnogo više crtica, odnosno smrtnih slučajeva, nego oko ostalih.

> Da li se smrtni slučajevi nalaze samo u blizini zaražene pumpe označene crvenim markerom? Kako to utiče na zaključak o zdravstvenoj ispravnosti pumpi označenih zelenim markerima?

Činjenica da se smrtni slučajevi pojavljuju i relativno daleko od pumpe koja je očigledno zaražena, ne bi trebalo da utiče na zaključke o ispravnosti ostalih pumpi. Prostije rečeno, verovatnije je da su sporadični smrtni slučajevi oko drugih pumpi posledica korišćenja vode sa zaraženog izvora, nego da su sve pumpe zaražene. Da su i druge pumpe bile zaražene, učestalost smrtnih slučajeva u njihovoj blizini bila bi mnogo veća.

1.5.1. Tabelarni i grafički prikaz podataka

> Koja ocena je najčešća u primeru 1? Da li ćete na ovo pitanje najlakše odgovoriti pomoću tabele, stubičastog ili torta dijagrama? Zbog čega?

Najviše đaka dobilo je ocenu 5. Ovo je teško zaključiti iz tabele, jer prethodno treba obaviti dodatne operacije prebrojavanja. Procena se najlakše obavlja na osnovu stubičastog dijagrama, jer je lakše preceniti visinu, kao osnovnu karakteristiku objekata, nego njihovu površinu ili ugao koji zahvataju.

> U čemu je prednost stubičastog u odnosu na kružni dijagram u primeru 2?

Pored toga što je na osnovu torta dijagrama teže uporediti učestalosti različitih ocena i proceniti razliku u njihovom broju, njime nije moguće prikazati prazne kategorije vrednosti, odnosno ocene koje se nisu pojavile u grupi đaka.

> Koji zaključak o podacima u primeru 3 biste lakše doneli na osnovu kružnog nego na osnovu stubičastog dijagrama?

Iako teže procenjujemo razlike u površinama nego u visinama objekata, kružni dijagram olakšava (vizuelno) sabiranje tih površina. Na osnovu njega ćemo lakše utvrditi da je polovina đaka dobila ocenu četiri, dok bi kod stubičastog dijagrama to podrazumevalo teži zadatak sabiranja visina stubića.

> Da li biste rekli da je uspeh odeljenja iz primera 4, kao grupe, dobar ili loš?

Iako u odeljenju postoji troje đaka koji su dobili ocene 4 ili 5, većina ih je dobila ocenu 2, pa bi se moglo reći da kao grupa imaju loš učinak.

> Oba grafikona u primeru 5 su simetrična. Kako na procenu simetričnosti utiču upotrebljene boje? Koji kriterijum je upotrebljen za određivanje nijansi boja na stubičastom a koji na kružnom dijagramu?

Pored oblika, stubičasti dijagram je simetričan i u odnosu na upotrebljene boje. Kod njega su nijanse određene visinom stubića, tj. brojem đaka u svakoj kategoriji. Nijanse odsečaka kružnog dijagrama označavaju vrednosti varijable.

> Šta je potrebno uraditi sa postojećom matricom ako želite da prikupite podatke o dodatnih 20 đaka i izmerite im još jedno svojstvo, npr. zaključnu ocenu iz hemije?

Pošto se redovi u matrici sirovih podataka vezuju za ispitanike, tj. entitete, a kolone za varijable, tj. svojstva koja su im izmerena, postojećoj matrici treba dodati dvadeset novih redova i jednu kolonu.

1.6. Naučna vizualizacija i vizualizacija informacija

> Kako na preglednost i intuitivnost grafa utiču variranje boje i veličine krugova?

Odgovor zavisi od cilja vizualizacije. Svako povećanje broja varijabli koje se predstavljaju grafom, utiče na njegovu preglednost, ali ne nužno negativno. Boja omogućava da se lakše grupišu elementi koji nisu prostorno bliski ali su slični po nekom svojstvu, npr. polu. Veličina kruga omogućava lako poređenje objekata i lociranje onih elemenata koji odudaraju od većine drugih. Iako je variranjem boje i veličine na grafu prikazana veća količina informacija, to ne narušava njegovu preglednost i čini ga intuitivnijim i lakšim za interpretaciju.

> Da li su raspored i pozicija objekata na ekranu potpuno proizvoljni ili strogo određeni?

Ni jedno ni drugo. Položaj kružića nema veze sa pozicijama na horizontalnoj i vertikalnoj dimenziji slike, ali njihov odnos nije potpuno proizvoljan. Primenjeni algoritam ne dozvoljava da se bliski entiteti nađu na različitim stranama grafa, jer moraju da budu očuvani njihovi međusobni odnosi. Ručnim pomeranjem kružića po ekranu možete da im promenite apsolutni položaj, ali time se neće bitno promeniti njihov relativni položaj u odnosu na ostale elemente. U tom smislu se donose i zaključci. Na primer, Tanja i/ili Sofija nisu udaljene od Ane, Petra, Marka ili Jovana. One su kao dijada udaljene od grupe ostalih đaka koji mogu da se posmatraju kao koherentna celina.

> Da li vam je lakše da uočite razlike među đacima u polu ili u uspehu?

Boja je svojstvo koje se razlikuje već na nivou ranog opažanja. Boje se lakše i brže upoređuju nego oblici ili njihove veličine. Naravno, ova svojstva uvek mogu da se upotrebe tako da oblik (npr. trougao i krug) bude dominantan u odnosu na boju (npr. svetlo- i tamnoplava nijansa).

> Da li bi graf bio razumljiviji kada bi se pol učenika predstavio veličinom kruga a uspeh u školi različitim bojama? Da li je to opravdano?

Ovo nije opravdano zato što su u pitanju varijable različitog tipa. Ocene su numeričke vrednosti među kojima može da se vrši poređenje po veličini. Pol je varijabla koja (uglavnom) ima dve jasno odvojene kategorije. U tom smislu, upotreba veličine za oznaku pola a boje za ocenu, ne bi bilo intuitivno.

> Zbog čega je, po vašem mišljenju, Milica „zvezda" odeljenja?

Milica je „zvezda" odeljenja jer većina veza drugih đaka vodi ka njoj. Razlog nije moguće videti na osnovu grafikona ali se može pretpostaviti da ima veze sa njenim dobrim uspehom. Druga deca možda žele da se druže sa njom zato što je cene i poštuju njene sposobnosti, a možda jednostavno žele da sede pored nje da bi mogli da prepisuju ili traže pomoć.

> Koje dodatne varijable bi se mogle vizualizovati na grafu i na koji način?

Đacima bi moglo da se postavi pitanje o tome koliko često se druže van škole, tako da bi o njihovoj bliskosti mogla da govori debljina linije. Oblikom bi mogao da se prikaže podatak da li se đak bavi nekim sportom ili ne. Treba imati na umu da se povećanjem broja vizuelnih svojstava narušava razumljivost i intuitivnost grafa, jer postaje sve teže porediti i grupisati elemente po raznim atributima ili njihovim kombinacijama. Na slici je mnogo lakše utvrditi odnos broja crvenih i plavih objekata, nego odnos crvenih trouglova i plavih kvadrata.

1.7. Vizualizacija kao eksplorativna tehnika

> Uporedite različite države i regione po broju stanovnika i nacionalnom dohotku.

Prilikom prvog prikazivanja grafikona najverovatnije ste prvo uočili razlike u veličini krugova, odnosno broju stanovnika, zbog dve najmnogoljudnije države na slici (Kina i Indija). Poređenje po regionima zahteva „prebacivanje" u drugi režim filtriranja podataka, ignorisanje veličine krugova kao primarnog svojstva i grupisanje entiteta po boji.

> Analizirajte promene u vremenu i povežite ih sa značajnim istorijskim događajima.

Najočiglednije negativne promene kod većeg broja država su one u periodu Prvog i Drugog svetskog rata. Na grafikonu mogu da se vide i posledice velike gladi u Irskoj u drugoj polovini 19. veka, Ukrajini 30-tih godina i Kini 60-tih godina 20. veka. Jedna od prekretnica je poslednja dekada 19. veka, kada SAD preuzimaju poziciju najjače ekonomije sveta od Ujedinjenog Kraljevstva kao najveće imperijalne sile do tada.

> Uočite države koje se u različitim vremenskim periodima i po različitim svojstvima izdvajaju od ostalih država istog regiona ili celog sveta.

Obratite pažnju na to da je moguće grupisati države po različitim atributima. Na kraju analiziranog perioda, u gornjem desnom uglu nalazi se veliki broj evropskih država, SAD, nekoliko azijskih i nekoliko bliskoistočnih država koje međusobno nisu slične po veličini, odnosno broju stanovnika.

> Da li su države Supsaharske Afrike u poslednjih 100 godina uspele da dostignu evropske države po kvalitetu života?

Ne. To je jedan od osnovnih problema na koje je Hans Rosling želeo da skrene pažnju koristeći ovu zanimljivu vizualizaciju.

> Koje države su imale najveći rast bruto nacionalnog dohotka po glavi stanovnika nakon Drugog svetskog rata?

Pored SAD i nekolicine evropskih zemalja koje su i pre toga beležile stalan rast ekonomije, najjači posleratni razvoj uočljiv je kod država bogatih naftom: Irak, Kuvajt, Libija, Saudijska Arabija, Katar i Ujedinjeni Arapski Emirati.

1.8. Izbor prikladne tehnike vizualizacije

> Na kojim dimenzijama se profil infografika u najvećoj meri razlikuje od profila naučnih grafikona?

Naučni grafikoni treba da budu funkcionalni, svedeni i ekonomični. Infografici su namenjeni širem auditorijumu i drugačijim medijumima za komunikaciju, tako da mogu da sadrže određene dekorativne elemente, slike konkretnih objekata koji se predstavljaju ili da zauzimaju više prostora na ekranu ili u novinskim člancima.

> Prikažite oba profila. Da li vam je na osnovu prikazanog grafikona lakše da uočite razlike među profilima ili razlike na pojedinačnim dimenzijama?

Verovatnije je da kombinaciju vrednosti na dimenzijama opažate kao celinu zbog fizičke povezanosti tačaka i osenčenosti površine. Stoga vam je lakše da uočite razlike među profilima, nego da izolujete duži koje govore o razlikama na pojedinačnim atributima (info)grafika.

> Ako šest osa radar dijagrama posmatramo kao dvopolne (bipolarne) dimenzije na čijim su krajevima suprotne karakteristike, koje svojstvo grafikona nije intuitivno ili nije logično?

Ako su dimenzije bipolarne, povećanje vrednosti na jednoj od njih trebalo bi da bude povezano sa smanjenjem vrednosti na onoj koja joj je naspramna. Kod radar dijagrama to ne mora da bude slučaj, tako da naučni grafikoni

imaju (navodno) veće vrednosti od naučnih grafikona na inovativnosti, ali i na naspramnoj familijarnosti.

> Da li bi poređenje entiteta, odnosno varijabli, bilo podjednako lako i opravdano kada bi varijable imale različite raspone vrednosti, npr. ocena iz fizičkog, visina đaka, težina đaka, vreme za koje đak pretrči 100 metara i broj zgibova koje je uspeo da uradi?

Ako bi svaki profil predstavljao jednog đaka, njihovo međusobno poređenje bilo bi opravdano i lako, jer bi se dovodile u vezu vrednosti na istim varijablama. Međutim, poređenje različitih dimenzija nije opravdano zato što su njihove vrednosti izražene u različitim jedinicama. Ipak, varijable mogu da se transformišu tako da je moguće njihovo poređenje i donošenje zaključka da, na primer, neki đak ima veću vrednost na visini, nego neki drugi đak na težini. O tome će biti više reči u narednim odeljcima.

> Pronađite proizvoljnu vizualizaciju na internetu i izmenite vrednosti na grafikonu tako da formiraju profil koji bi joj najviše odgovarao.

Za pretragu interneta upotrebite ključne reči „data visualisation examples".

1.8.3. Čitljivost grafikona

> Koji zaključak ćete lakše doneti na osnovu grafikona *pol x fakultet* a koji na osnovu grafikona *fakultet x pol*?

Na osnovu grafikona *pol x fakultet* lakše ćete utvrditi razlike u ukupnom broju studenata po fakultetima, kao i razlike u broju studenata različitog pola na svakom fakultetu. Sa druge strane, grafikon *fakultet x pol* olakšava zaključivanje o obrazovnom profilu, odnosno strukturi odabranih usmerenja po polu.

> Da li je opravdano upotrebiti poligone frekvencija u primeru *pol x fakultet*? Sortirajte vrednosti po frekvencijama dok su poligoni vidljivi.

Nije. Poligon frekvencija koristi se za prikazivanje kvantitativnih varijabli. Ako menjate kriterijum sortiranja sa vrednosti na frekvencije i obratno, to ne utiče na razumljivost i smisao stubičastog dijagrama. Međutim, poligon frekvencija se bitno menja, što može (neopravdano) da sugeriše da se raspodele vrednosti varijable razlikuju.

> Menjajte opcije za generisanje grafikona i utvrdite kada je opravdano a kada ne, sortiranje po učestalostima, različita obojenost stubića i prikazivanje tačaka, odnosno iscrtavanje poligona.

Osnovni kriterijum prilikom procene opravdanosti trebalo bi da bude razlika između kvalitativnih i kvantitativnih (kategorijalnih) varijabli. Ove prve ne treba predstavljati poligonom frekvencija. Sa druge strane, kod kvantitativnih varijabli, ali i kod kvalitativnih koje imaju previše kategorija, upotreba različitih boja postaje besmislena.

2.1. Pojam verovatnoće

> Kolika je verovatnoća slučajnog pogađanja tačnog odgovora na pitanja koja sadrže samo dve opcije – tačno i netačno?

Verovatnoća slučajnog pogađanja je 50% ili 0,5.

> Kolika je verovatnoća pogađanja tačnog odgovora ako vam je poznato da jedan od četiri ponuđena odgovora nije tačan?

Verovatnoća slučajnog pogađanja je 1 : 3 ili približno 33%. Eliminisanje nekih od mogućih ishoda povećava verovatnoću dobijanja onih preostalih.

> Kolika je verovatnoća pogađanja tačnog odgovora ako vam je poznato da jedan od četiri ponuđena odgovora nije tačan i ako postoje dva odgovora koja se priznaju kao tačna, tj. ako događaj ima dva poželjna ishoda?

Verovatnoća je 2 · 33% ili 2 : 3, odnosno oko 67%.

2.3. Pojam nasumičnosti ili slučajnosti

> Kakav odnos (proporciju) broja kuglica očekujete u uzorku na osnovu izgleda populacije?

Očekuje se da odnos broja kuglica različitih boja u uzorku bude približno isti kao i u populaciji. Verovatnoća izvlačenja kuglice bilo koje boje je podjednaka.

> Da li proporcije stratuma u uzorku veličine 25 verno odražavaju proporcije koje možete da uočite u populaciji?

Najverovatnije ne. Da bi se zakon velikih brojeva manifestovao, uzorci moraju da budu dovoljno veliki.

> Kliknite taster *Obriši uzorak* i formirajte novi uzorak veličine 25. Da li su proporcije kružića različitih boja u drugom uzorku iste kao u prvom?

Najverovatnije ne. Iz iste populacije moguće je uzeti praktično neograničen broj uzoraka koji međusobno ne moraju da budu isti, pa čak ni slični, posebno ako se radi o uzorcima male veličine.

> Da li se reprezentativnost uzorka popravila nakon što je njegova veličina povećana?

Najverovatnije da. Povećavanjem veličine uzorka, povećava se i verovatnoća da će on biti reprezentativan u odnosu na populaciju iz koje je uzet.

> Formirajte još nekoliko uzoraka veličine 100. Da li se uzorci veličine 100 međusobno više ili manje razlikuju od uzoraka veličine 25?

Veći uzorci trebalo bi da budu međusobno sličniji jer daju tačniju i precizniju sliku populacije. Prostije rečeno, kada u nečemu grešimo, to može da bude na veliki broj različitih načina. Skup tačnih odgovora obično je mnogo manji.

> Da li na osnovu uzorka veličine 25 možete da zaključite koliko stratuma postoji u populaciji?

Moguće je pretpostaviti, ali bi ta pretpostavka verovatno bila pogrešna u većini slučajeva, tj. u većini zaključaka donetih na osnovu malog uzorka.

> Da li na osnovu uzorka veličine 25 možete da zaključite koji od stratuma u populaciji je proporcionalno najmanji a koji najveći?

Ukoliko su razlike u veličinama stratuma u populaciji veće, veća je verovatnoća da će se one manifestovati i na malim uzorcima. Ipak, ove procene će biti mnogo tačnije ako su uzorci veći.

> Ponovite postupak skrivanja i uzorkovanja sa populacijama 3 i 4. Da li vam je bilo lakše i da li ste bili tačniji u proceni odnosa verovatnoća različitih ishoda na osnovu uzoraka uzetih iz Populacije 3 ili iz Populacije 4?

Verovatno ste tačno predvideli da je u Populaciji 4 najveći narandžasti stratum, ali vam je bilo teško da odredite kakav je odnos preostalih boja zbog veoma male verovatnoće tih ishoda.

> U kojoj od ove dve populacije je raznolikost entiteta, tj. boja veća?

Iako je broj stratuma u obe populacije isti, raznolikost je veća u Populaciji 3. Naime, u Populaciji 4 očigledno dominira jedna boja, te možemo da kažemo da se u njoj nalazi veći broj međusobno sličnih entiteta.

2.4. Pojam varijabilnosti

> Sakrijte populaciju, kliknite taster *Generiši populaciju* da biste formirali nasumičnu populaciju i potom na osnovu uzoraka različite veličine pokušajte da procenite kako ona izgleda.

Obratite pažnju na to da prilikom analize uzoraka i uopštavanja zaključaka na celu populaciju, najčešće nećete biti podjednako sigurni u tvrdnje koje iznosite. Verovatno je lakše proceniti da li je neka boja dominantna, nego napraviti rang listu boja po učestalosti.

> Pomeranjem težišnih tačaka na dijagramu promenite proporcije stratuma i kreirajte sopstvenu populaciju. Da li suma proporcija može da bude veća ili manja od 1? Da li povećavanje proporcije jednog ishoda utiče na proporcije (svih) drugih ishoda?

Suma verovatnoća, odnosno proporcija međusobno isključujućih ishoda jednog događaja, uvek mora da bude 100, tj. 1. U našem primeru, to znači da je potpuno izvesno da će svaka izvučena kuglica imati neku boju. Kada se povećava verovatnoća jednog ishoda, verovatnoća nekog drugog ili svih ostalih se smanjuje i obratno.

2.5.1. Tabele frekvencija i tabele kontingencije

> Koristeći primer tabele frekvencija i stubičastog dijagrama za varijablu *fakultet*, odredite koje ste procene i zaključke lakše doneli na osnovu jednog a koje na osnovu drugog načina sažimanja podataka.

Grafikon je nezamenjiv kada treba napraviti grube procene odnosa među različitim vrednosti. Na primer, na osnovu stubičastog dijagrama ćete mnogo lakše i brže utvrditi da je najviše studenata na FTN ili da ih je više na FIL nego na PRA. Međutim, ako treba da se prikažu precizne proporcije ili fine razlike među vrednostima, tabela frekvencija je preglednija i informativnija. Raspored elemenata na grafikonu može da utiče na njegovu čitljivost, tako da ćete razliku između EKO i PRA teško uočiti ako ne konsultujete podatke iz tabele.

> Zbog čega je broj tačaka na poligonima frekvencija za dva veći od broja redova u tabeli frekvencija?

Uobičajeno je da se poligon frekvencija iscrtava tako da linija počinje i završava se na x-osi. Štaviše, to je karakteristika koja ovu vrstu grafikona čini

poligonom a ne prostom izlomljenom linijom. Pored razreda navedenih u tabeli, na grafikonu su prikazani najniži i najviši razred u kojima nema rezultata. Ovo je još jedan argument koji pokazuje da poligoni frekvencija nisu prikladni za vizualizaciju kvalitativnih varijabli kod koji je redosled kategorija na x-osi, u suštini, nebitan.

> Zbog čega ogiva ima krivolinijski oblik? U kom slučaju bi imala oblik prave?

Krivolinijski oblik je posledica različitih učestalosti po razredima. Kumulativni priraštaj je manji u nižim i višim razredima, a veći u onima koji se nalaze oko srednje vrednosti. Kada bi frekvencije svih razreda bile podjednake, ogiva bi imala oblik prave linije jer bi priraštaj u svim tačkama bio isti.

> Koju vrednost na y-osi ima najviša tačka ogive?

Vrednost na y-osi u najvišoj tački ogive jednaka je veličini uzorka.

> Koja boja označava koji pol na grafikonu u primeru *fakultet x pol*?

Na osnovu poređenja visine stubića na grafikonu i vrednosti datih u tabeli, zaključujemo da plava boja predstavlja studentkinje a narandžasta studente.

> Posmatrajući rezultate krostabulacije varijabli *fakultet* i *pol*, proverite da li raspodela studenata po polu u ukupnom uzorku odgovara raspodelama u podgrupama formiranim na osnovu varijable *fakultet*.

Na osnovu raspodele u primeru *pol*, zaključujemo da u ukupnom uzorku ima više muškaraca. Međutim, ovaj odnos ne važi za pojedinačne fakultete, kako u smislu razlike u korist muškaraca, tako i u smislu veličine te razlike.

> Koja karakteristika stubičastog dijagrama u primeru *visina (r) x pol* ukazuje na to da postoji razlika u visini među polovima?

Narandžasti stubići pomereni su više udesno, ka većim vrednostima x-ose.

> Kolika je teorijska verovatnoća da je neka osoba u populaciji iz primera, nezavisno od toga kog je pola, viša od 175 cm?

Da bi se dao odgovor na ovo pitanje, treba sabrati učestalosti razreda iznad vrednosti 175 cm i podeliti ih veličinom uzorka. Dakle, verovatnoća iznosi (307 + 94 + 13) : 1723 ≈ 0,24.

> Kolika je teorijska verovatnoća da je neka studentkinja visoka između 184 i 187 cm?

Ako znamo da je u pitanju studentkinja, verovatnoća da je ona visoka između 184 i 187 cm iznosi 2 : 841 ≈ 0,002. Ako nam to nije poznato, verovatnoća da iz populacije nasumično odaberemo studentkinju te visine duplo je manja – 2 : 1723 ≈ 0,001.

2.5.2.1. Aritmetička sredina, medijana i mod

> Koju vrednost treba uneti umesto neke od petica u tabeli da bi vrednost M ponovo postala 5, odnosno da bi se ponovo uspostavila simetričnost distribucije prikazane na grafikonu?

Potrebno je uneti vrednost koja jednako odstupa od 5 kao i 0, ali u suprotnom smeru. Kada unesete vrednost 10, aritmetička sredina ponovo postaje 5 jer je (10 + 0) : 2 = 5.

> Formirajte više puta slučajne distribucije klikom na ikonicu kockica u gornjem desnom uglu i pokušajte na osnovu izgleda grafikona, traženjem njegovog centra ravnoteže, da procenite kolika je vrednost M dobijenih distribucija.

Vrednost aritmetičke sredine uvek se nalazi u težištu distribucije, bez obzira na njen oblik. Sa druge strane, samo kod simetričnih distribucija težište se nalazi na polovini raspona raspona, odnosno na poziciji srednje vrednosti distribucije.

> Odaberite primer 3 i analizirajte vrednosti mera centralne tendencije.

Ukoliko je distribucija simetrična, vrednosti aritmetičke sredine i medijane biće približno iste. Velika razlika između moda, sa jedne strane, i medijane, odnosno aritmetičke sredine, sa druge, sugeriše da je verovatno u pitanju izrazito asimetrična raspodela.

> Da li su M i Md u ovom primeru prikladne mere centralne tendencije?

Ne. I jedna i druga su podjednako loše, tj. pogrešne.

> Zbog čega M i Md u ovom primeru imaju istu vrednost? Kakav oblik imaju sve distribucije kod kojih M i Md imaju istu vrednost?

Vrednost aritmetičke sredine nalazi se u sredini niza rezultata, jer je odstupanje rezultata sa leve i desne strane medijane jednako. Sve simetrične distribucije imaju jednake vrednosti medijane i aritmetičke sredine.

> Zbog čega umesto broja za vrednost Mo stoji reč „više“? Zašto ovu distribuciju nazivamo bimodalnom? Kako bi mogla da izgleda neka *polimodalna* distribucija?

Očigledno je da postoji više najčešćih vrednosti, odnosno više vrednosti čija je učestalost jednaka. U ovom primeru postoje dva moda pa se i distribucija naziva bimodalnom. Polimodalne distribucije imaju više modova ali oni ne moraju da imaju istu učestalost.

> Distribucije 1 i 3 imaju istu M. Za koju od tih distribucija je vrednost M bolja mera centralne tendencije i zašto?

Aritmetička sredina je bolja mera centralne tendencije u prvom primeru jer tačno predstavlja sve rezultate. U drugom primeru ona ne govori ništa o stvarnom učinku đaka na testu znanja.

> Formirajte više puta slučajne distribucije klikom na ikonicu kockica i analizirajte odnose između vrednosti različitih mera centralne tendencije.

Obratite posebnu pažnju na primere u kojima je medijana veća od proseka i obratno. Imajte na umu da isti rasponi rezultata na osi, ne moraju da obuhvate i isti broj rezultata. Na primer, u rasponu od 0 do 5 bodova može da bude duplo manje đaka nego u rasponu od 5 do 10. Tada će aritmetička sredina verovatno da bude manja od moda i medijane.

2.5.3.1. Vizuelna procena i poređenje varijabilnosti

> Koliko varijabli prepoznajete u ovom eksperimentu? Koji je nivo merenja svake od njih?

Dve ključne varijable u ovom eksperimentu su veličina kvadratića koji su služili kao meta i vaša brzina reakcije, odnosno brzina pronalaženja mete. Prva varijabla je nominalnog nivoa a druga razmernog.

> Koliko redova i koliko kolona treba da ima matrica sirovih podataka u koju biste zabeležili podatke prikupljene u ovom primeru?

Pošto je obavljeno 20 merenja, tabela treba da ima 20 redova. U svakom merenju zabeležene su vrednosti dveju varijabli – veličine kvadrata i vaše brzine. Matrica sirovih podataka, dakle, treba da ima dve kolone.

> Da li se redovi matrice sirovih podataka u navedenom primeru odnose na različite osobe (ispitanike) ili na nešto drugo?

Redovi se odnose na 20 merenja. U eksperimentu je učestvovao samo jedan ispitanik, tako da se analizira brzina na nivou pojedinca a ne grupe. Merenje je ponavljano kako bi se smanjila mogućnost greške i dobila stabilnija i tačnija procena brzine. Očekuje se da će zbog toga pojedinačne greške merenja

vezane za uslove, npr. pomeranje miša ili odvlačenje pažnje, manje uticati na konačne rezultate.

> Na koji način je upotrebljena varijabla *veličina kvadrata* na prikazanim grafikonima?

Veličina kvadrata upotrebljena je kao grupišuća varijabla na osnovu koje je 20 merenja podeljeno u dve grupe od po 10 vrednosti.

> Da li je brzina reakcije mogla da bude izražena u sekundama ili nekim drugim jedinicama umesto u milisekundama? Šta to govori o varijabli?

Da, i to ni na koji način ne bi uticalo na naše zaključke. Međutim, promena iz jednih jedinica u druge mora da se obavi po jasnim pravilima, jer je u pitanju kvantitativna varijabla razmernog nivoa. Sa druge strane, veličinu kvadrata mogli smo da izrazimo njihovom površinom, veličinom stranice, ciframa 0 i 1 ili slovima A i B. Ovakvu slobodu imamo zato što varijablu tretiramo kao nominalnu, tj. kao kvalitativni kriterijum na osnovu koga se merenja razvrstavaju u dve grupe.

> Ukoliko izračunate prosek vrednosti M_V i M_M, dobićete zajedničku aritmetičku sredinu koja je jednaka vrednosti M. Zbog čega je to tako? Kada prosek tih vrednosti ne bi dao vrednost M?

Ovakav rezultat se dobija jer su veličine grupa na kojima su izračunati proseci iste. Kada to ne bi bio slučaj, prosek aritmetičkih sredina po grupama i prosek objedinjenih rezultata ne bi bili isti. Tada bi se vrednosti M tretirale kao ravnopravne, ali bi zapravo jedna od njih trebalo da ima veću težinu.

> Da li je opravdano izračunavanje vrednosti R kao proseka ili zbira vrednosti R_V i R_M? Zbog čega jeste, odnosno zbog čega nije?

To nije moguće. Za izračunavanje R potreban je podatak o najmanjem i najvećem rezultatu u grupi merenja, što nije moguće zaključiti na osnovu vrednosti raspona u pojedinačnim grupama merenja.

> Kako biste izračunali raspon rezultata svih merenja na osnovu raspona grupa merenja, odnosno levog i desnog grafikona?

Najmanji rezultat prikazan na levom ili desnom grafikonu predstavlja donju granicu raspona srednjeg grafikona. Najveći rezultat sa levog ili desnog grafikona je njegova gornja granica.

2.5.3.3. Pojam matematičke funkcije

> Odaberite osmu funkciju i analizirajte njen oblik. Linija koju vidite sastavljena je iz dva dela, odnosno dve funkcije. Možete li da ih uočite? Zbog čega prikazana linija nije mogla da bude nacrtana uz pomoć samo jedne formule, tj. funkcije?

Prvom funkcijom opisan je oblik gornjeg dela srca, tj. dva polukruga, a drugom njegov donji deo. Prikazani oblik ne bi mogao da se definiše jednom formulom, jer za istu vrednost na x-osi mogu da se dobiju različite vrednosti na y-osi. U prvoj funkciji za x = 0, y je 1, a u drugoj je približno -2.

2.6.2. Funkcije mase i gustine verovatnoće

> Ako ste u prvom bacanju dobili broj 3, kolika je verovatnoća da u drugom bacanju dobijete isti broj?

Ista kao i u prvom – oko 17%.

> Ako kockicu bacite 6 puta, da li je verovatnije da ćete dobiti brojeve ovim redom: 1, 1, 1, 1, 1, 1 ili ovim redom: 2, 4, 6, 1, 3, 5?

Verovatnoća oba ishoda je jednaka. Raznovrsnost brojeva u drugom ishodu sugeriše da bi on mogao češće da se dobije slučajno, ali verovatnoća da

brojevi padnu tim redom potpuno je ista kao i da padne šest jedinica zaredom.

> Ako istovremeno bacite 6 kockica, da li je verovatnije da ćete dobiti brojeve 1, 1, 1, 1, 1, 1 ili 2, 4, 6, 1, 3, 5?

Ovoga puta je verovatniji drugi ishod, jer nas ne interesuje tačan redosled, već samo postojanje brojeva u skupu od šest kockica. Drugi ishod može da se dobije ako redom padnu brojevi 1, 2, 3, 4, 5, 6 ili 6, 5, 4, 3, 2, 1 ili 1, 3, 2, 4, 6, 5 i tako dalje. Prvi se dobija samo ako se na svim kockicama dobije broj 1.

> Da li biste rekli da je verovatnoća da dobijete distribuciju prikazanu na početku ove vežbe, nakon šest bacanja jedne kockice, mala ili velika?

Verovatnoća je veoma mala. U prvom bacanju je nebitno koji broj ćemo dobiti. Dakle, verovatnoća povoljnog ishoda je 6 : 6. U drugom treba da dobijemo neki od preostalih pet brojeva, pa je verovatnoća 5 : 6. I tako dalje. To znači da je verovatnoća da dobijemo prikazanu sliku $(6 : 6) \cdot (5 : 6) \cdot (4 : 6) \cdot (3 : 6) \cdot (2 : 6) \cdot (1 : 6)$ ili približno 1,5%.

> Da li možete da procenite kako izgleda teorijska distribucija verovatnoća u ovom primeru?

Imajte na umu da različiti parovi brojeva mogu da daju iste sume. Samim tim, neke sume su verovatnije od drugih.

> Da li je jednako verovatno da prilikom bacanja dve kockice dobijete zbir 2, 12 i 7? Koja vrednost je verovatnija i zašto?

Ne, jer različite sume mogu da se dobiju različitim brojem permutacija brojeva. Najverovatniji ishod je 7 jer se dobija najvećim brojem permutacija. Najmanje verovatni ishodi (sume) su 2 i 12.

2.6.3. Standardizacija sirovih rezultata

> Da li je A. M. dobro ili loše uradio zeleni test?

Na osnovu pozicije u nizu, može se reći da je A. M. dobro uradio test. Međutim, to i dalje ne znači da ga je uradio dovoljno dobro jer nemamo podatke o karakteristikama testa.

> Da li je A. M. bolje uradio zeleni ili plavi test?

Na ovo pitanje nije moguće dati odgovor jer nemamo dovoljno podataka o testovima. Veći broj bodova u apsolutnom smislu ne znači da je student bolje uradio neki test. Moguće je da se razlikuju teorijski rasponi rezultata testova.

> Koliko je A. M. bolji na plavom testu od nekoga ko je osvojio 25 bodova?

Odgovor može da se da samo u apsolutnom smislu ali ne i u relativnom. Razlika od 17 bodova može da bude velika ili mala, u zavisnosti od ukupnog raspona rezultata i karakteristika distribucije.

2.6.4. Površina ispod normalne krive

> Pomerajte granice površine ispod normalne krive i pratite promene u procentu obuhvaćenih rezultata. Interpretirajte granice u terminima z skorova pojedinačnih rezultata.

Obratite pažnju na to da isti rasponi z skorova ne obuhvataju isti procenat rezultata, tj. istu proporciju površine ispod normalne krive. Na primer, rasponom rezultata između proseka i prve standardne devijacije, obuhvaćeno je duplo više rezultata nego rasponom od z vrednosti 1 do desnog kraja distribucije.

> Koliki procenat rezultata se nalazi izvan raspona $\mu \pm 1{,}96 \cdot \sigma$ a koliki izvan raspona $\mu \pm 2{,}58 \cdot \sigma$?

Izvan raspona $\mu \pm 1{,}96 \cdot \sigma$ nalazi se 100 - 95 ili oko 5% rezultata. Izvan raspona $\mu \pm 2{,}58 \cdot \sigma$ nalazi se približno 1% svih rezultata.

> Kolika je verovatnoća da iz populacije nasumično odaberete ispitanika koji na nekoj normalno distribuiranoj varijabli ima z skor manji od -1?

Verovatnoća je približno 16%.

> Kolika je verovatnoća da iz populacije nasumično odaberete ispitanika koji na nekoj normalno distribuiranoj varijabli ima z skor veći od 1,64?

Verovatnoća je približno 5%.

> Kolika je verovatnoća da iz populacije nasumično odaberete ispitanika koji na nekoj normalno distribuiranoj varijabli ima z skor između 1 i 2?

Verovatnoća je približno 14%.

2.6.6. Skjunis i kurtozis

> Zbog čega je kurtozis veći kada grupu aberantnih rezultata dodate samo sa jedne strane distribucije nego kada ih dodate sa obe?

Središnji rezultati više „štrče" ako se aberantni rezultati dodaju samo sa jedne strane. Kada se dodaju i autlajeri sa desne strane, aberantnost više nije toliko izražena, jer veliki broj atipičnih rezultata prestaje da bude atipično. Štaviše, ako biste nastavili da dodajete rezultate samo sa leve ili desne strane, kurtozis bi se najpre približavao nuli, a nakon toga bi postao negativan.

> Kako na vrednost skjunisa utiče dodavanje grupe aberantnih rezultata sa leve, kako sa desne, a kako sa obe strane?

Dodavanje rezultata samo sa jedne strane, zakrivljuje distribuciju u tu stranu, tako da se vrednost skjunisa menja u zavisnosti od smera iskošenosti. Ako se

aberantni rezultati dodaju sa obe strane, distribucija je simetrična, te je i njen skjunis blizak nuli.

2.7.2. Hi-kvadrat distribucija

> Zbog čega je hi-kvadrat distribucija iskošena u desnu stranu?

Hi-kvadrat distribucija je distribucija kvadriranih z vrednosti, što znači da ona nema negativnu stranu. Osim toga, operacija kvadriranja predstavlja nelinearnu transformaciju kojom se veće vrednosti više povećavaju od manjih. Stoga je desni kraj distribucije „razvučen" u odnosu na raspon kvadriranih središnjih vrednosti normalne distribucije.

> Zbog čega F distribucija za veličine uzoraka N = 2 i N = 100 nema isti oblik kao ona za veličine uzoraka N = 100 i N = 2?

U prvom primeru veća je verovatnoća da varijansa u brojiocu bude bliska nuli, jer je uzorak manji. U drugom je situacija obrnuta, tako da je prosek distribucije pomeren udesno, ka većim vrednostima.

2.9. Test-statistici, p vrednosti i nivoi značajnosti

> Pomeranjem graničnih linija crvene površine utvrdite koja je minimalna t vrednost potrebna da bismo neku razliku smatrali značajnom na nivou 0,01 za 100 stepeni slobode.

U odgovoru na pitanje upotrebite podatak o procentu koji obuhvata površina, pošto se vrednost p, odnosno proporcija površine izvan graničnih linija, zaokružuje na dve decimale, te nije dovoljno precizna na krajevima distribucije. Obratite pažnju na to da tražimo simetričnu površinu koja obuhvata obe strane distribucije. Da biste istovremeno pomerali obe granične linije, držite pritisnut taster *Ctrl*. Tražena vrednost iznosi približno 2,63, što je

neznatno veće od odgovarajuće vrednosti koja za normalnu distribuciju iznosi 2,58.

> Od koje vrednosti treba da bude veći dobijeni χ^2 za dva stepena slobode da bismo tvrdnju da odabrane z vrednosti ne potiču iz normalne distribucije smatrali tačnom, uz verovatnoću greške manju od 5%, odnosno na nivou značajnosti 0,05?

Tražena vrednost iznosi približno 6.

2.9.1. Jednostrano testiranje razlika

> Odredite graničnu vrednost, od koje treba da bude veći F odnos, da biste varijansu za koju je df = 3 smatrali statistički značajno većom od varijanse za koju je df = 100 na nivou 0,01.

Tražena vrednost iznosi približno 4.

> Odaberite raspon 0,00–3,92, $df_1 = 3$ i $df_2 = 9$. Kakav zaključak o varijansama biste doneli na osnovu površine koju formira taj raspon? Da li isti zaključak važi i u slučaju kada je $df_1 = 9$ a $df_2 = 3$? Zbog čega važi, odnosno ne važi?

U prvom slučaju F odnos veći od 3,92 bio bi značajan na nivou 0,05. Isti zaključak ne važi u drugom slučaju, jer je tada veća verovatnoća dobijanja veće vrednosti F zbog očekivane veće varijabilnosti u brojiocu.

3.2. T-test za jedan uzorak

> Koliko varijabli, odnosno svojstava koja su menjala svoju vrednost, možete da prepoznate u ovom primeru? Nabrojte ih.

Dužina horizontalne duži, položaj duži na ekranu, inicijalni položaj srednje duži, vreme koje vam je bilo potrebno da rešite zadatak i vaš odgovor, tj. krajnji položaj srednje duži.

> Koje varijable su bile pod kontrolom eksperimentatora, tj. autora udžbenika, a koje su zavisile od ispitanika, tj. vas?

Ispitanik je mogao da utiče na krajnji odgovor, tj. poziciju srednje duži, kao i na vreme potrebno za rešavanje zadatka. Eksperimentator je dizajnirao aplikaciju, tako da se namerno i nasumično variraju vrednosti svih ostalih varijabli.

> Zbog čega je eksperimentator namerno varirao dužinu horizontalne duži, poziciju duži na ekranu i početnu poziciju srednje vertikalne duži?

Ova svojstva stimulusa namerno su varirana da ne bi došlo do uvežbavanja ispitanika, rutinskog davanja odgovora ili korišćenja markera na ekranu ili monitoru (npr. logotipa) za pronalaženje srednje tačke duži. Svrha je bila da se kontrolišu sve varijable koje bi mogle da utiču na tačnost procene.

> Zbog čega je merenje ponavljano više puta, umesto da je svaki tip vertikalnih duži bio prikazan po jednom?

U eksperimentu je učestvovao samo jedan ispitanik, ali je broj merenja veći da bi se umanjio efekat potencijalnih grešaka. Ispitanik je slučajno mogao da pritisne taster ili napravi previd u jednom ili dva pokušaja, što ne bi trebalo da ima velik uticaj na konačni rezultat ako je broj merenja dovoljno velik. Uzet je veći uzorak procena da bi se došlo do pouzdanije procene tačnosti u populaciji svih odgovora.

> Zbog čega, prilikom računanja intervala poverenja aritmetičke sredine od 99%, nije upotrebljen raspon $M \pm 2{,}58 \cdot s_M$, već raspon $M \pm 3{,}708 s_M$?

Zbog veličine uzorka. Upotrebili smo graničnu vrednost t za 6 stepeni slobode.

> Kolika bi trebalo da bude s_M u grupi plavih merenja da bismo M = 0,511 smatrali statistički značajno različitom od vrednosti 0,5 na nivou 0,05?

Pošto granična vrednost t-testa za 6 stepeni slobode iznosi približno 2,447, vrednost s_M treba da bude manja od 0,0045 da bi razlika bila značajna.

> Kolika bi bila vrednost t da je u grupi narandžastih merenja M iznosilo 0,429?

Apsolutna vrednost t bila bi ista ali bi njen predznak bio negativan. Vrednost 0,429 jednako je udaljena od 0,5 kao i 0,571, ali se nalazi na suprotnoj strani očekivane distribucije. Obe vrednosti su značajno različite od 0,5, ali je prva značajno manja a druga značajno veća.

> Kolika bi bila vrednost p da smo u grupi plavih merenja primenili postupak jednostranog testiranja razlike?

Vrednost p bi bila duplo manja – 0,052.

3.3. T-test za dva uzorka

> Kako treba da izgleda matrica sirovih podataka, na osnovu kojih biste mogli da proverite da li se visina 30 dečaka statistički značajno razlikuje od visine 25 devojčica?

Matrica treba da ima 55 redova, za 30 dečaka i 25 devojčica, i 2 kolone, jednu za varijablu *pol* i drugu za varijablu *visina*.

> Zbog čega je u slučaju t-testa za dva uzorka hipoteza postavljena u formi H_0: $\mu_1 = \mu_2$ a ne u formi H_0: $M_1 = M_2$?

Statistički testovi koriste se da bi se na osnovu statistika koji su izračunati na uzorku, procenile vrednosti parametara populacije. Isto tako, istraživačke

hipoteze su pretpostavke o fenomenima u populaciji a ne u uzorku ili uzorcima. Da bi se proverila tačnost izraza $M_1 = M_2$, ne mora ni da se primenjuje t-test, već je dovoljno uporediti dve dobijene vrednosti. Međutim, takav zaključak nema nikakvu praktičnu vrednost jer se odnosi samo na entitete koji su činili uzorak. Smisao t-testa nije da se utvrdi postojanje razlike među aritmetičkim sredinama, već da se proceni statistička značajnost te razlike. Ukoliko je razlika statistički značajna, najverovatnije postoji i razlika između aritmetičkih sredina populacija.

> Na osnovu kojih podataka iz prikazane tabele biste mogli da zaključite da u primeru *Završni ispit 1* možda postoji aberantan rezultat?

Standardna devijacija ima visoku vrednost u poređenju sa prosekom.

> Zbog čega vrednosti t-testova za jedan uzorak u primeru *Završni ispit 2* imaju suprotan predznak?

Zato što je jedna grupa u proseku lošija a druga bolja od datog kriterijuma. Predznak, međutim, ne utiče na zaključak o tome da li je neka razlika statistički značajna ili nije.

> U kom slučaju bi t-test za dva uzorka u primerima sa završnim ispitom imao istu apsolutnu vrednost ali drugačiji predznak? Kako bi to uticalo na naš zaključak o razlikama među grupama?

Vrednost t imala bi drugačiji predznak kada bismo u formuli za njeno računanje zamenili mesta vrednostima M_1 i M_2. Ova promena ne bi uticala na konačni zaključak da li se grupe značajno razlikuju ili ne. Naravno, uvek je potrebno obratiti pažnju na predznak t-testa kako bi se doneo ispravan zaključak o tome koja grupa je bolja ili lošija.

> Prikažite ponovo podatke koje ste samostalno prikupili u vežbi i analizirajte ih u skladu sa opisanom logikom t-testa za jedan, odnosno dva uzorka.

Pokušajte da zamislite sve moguće ishode, odnosno kombinacije zaključaka u ovoj vežbi. Na primer, moguće je da ste u obe grupe merenja davali pogrešnu procenu sredine duži, ali da razlika među merenjima nije značajna jer ste grešili u istu stranu. Da ste u jednoj grupi potcenjivali a u drugoj precenjivali polovinu duži, razlika među merenjima bi verovatno bila značajna. U ovom poslednjem slučaju, može se desiti čak i to da razlika među merenjima bude značajna, iako se proseci merenja po grupama ne razlikuju značajno od datog kriterijuma.

> Analizirajte prikazane histograme i proverite da li je na prikupljenim podacima opravdano primeniti t-test.

Prvenstveno obratite pažnju na to da li postoje aberantni rezultati i značajna iskošenost distribucija.

> Ponovite vežbu tako da u nekom od zadataka namerno napravite netačan aberantan rezultat. Analizirajte distribucije i dobijene vrednosti t-testa.

Vrlo je verovatno da razlike među merenjima u ovoj situaciji neće biti statistički značajne zbog veće varijabilnosti rezultata i veće standardne greške razlike.

> Na koji način može da se proveri, odnosno testira ispunjenost uslova homogenosti varijansi pre primene t-testa?

U odeljku o važnim statističkim distribucijama objasnili smo da se značajnost razlika među varijansama može testirati pomoću F testa.

3.4.3. Men–Vitnijev test sume rangova

> Pomoću grafikona uporedite mere grupisanja, mere raspršenja i oblik distribucije dveju grupa merenja. Da li uočavate aberantne rezultate? Da li se oblici distribucija razlikuju s obzirom na simetričnost i/ili smer zakrivljenosti?

Kutijasti dijagram je verovatno najpregledniji način poređenja mera grupisanja i raspršenja većeg broja grupa istovremeno. Veličina kutije i raspon „brkova" su veoma lake za poređenje, jer se posmatra samo jedna dimenzija grafikona. Isto tako, različita dužina „brkova" sa obe strane ili izmeštenost kvadratića u odnosu na centar kutije, na veoma intuitivan način govori o asimetričnosti distribucije.

> Da li vrednost t-testa, prikazana u tabeli sa desne strane, ukazuje na postojanje statistički značajne razlike u vremenu reakcije na podudarne i nepodudarne stimuluse?

Očekuje se da će razlika biti značajna, ali u zavisnosti od toga koju strategiju primeni ispitanik, moguće je da Strupov efekat ne bude izražen.

> Na osnovu grafikona može se zaključiti da su obe distribucije simetrične. Uporedite rastojanja između najmanjeg rezultata, prvog kvartila, medijane, trećeg kvartila i najvećeg rezultata. Povežite veličinu ovih raspona sa sirovim rezultatima prikazanim u tabeli. Kako biste na osnovu ovih informacija zaključili koja od dve distribucije je uniformna a koja je približno normalna?

U obe grupe raspon između Q_1 i Mdn isti je kao raspon između Mdn i Q_3. To ukazuje na to da je distribucija najverovatnije simetrična. Međutim, IKR plave distribucije očigledno je proporcionalno manji u odnosu na ukupan raspon. Rastojanje od kvadratića do ivice kutije, manje je nego rastojanje od ivice kutije do kraja raspona. Kod narandžaste distribucije ovo rastojanje je približno isto. To pokazuje da je 50% rezultata u plavoj distribuciji obuhvaćeno proporcionalno manjim rasponom vrednosti nego u narandžastoj. Stoga je verovatnije da je plava distribucija normalna a da je narandžasta uniformna.

> U kojim primerima su razlike kutijastih dijagrama iscrtanih korišćenjem različitih mera grupisanja i raspršenja, veće a u kojima manje?

Ove razlike su manje izražene kod simetričnih distribucija. Ako su distribucije iskošene, nije prikladna upotreba aritmetičke sredine i standardne devijacije za određivanje dimenzija kutije i „brkova".

> Zbog čega se na kutijastim dijagramima iscrtanim uz opciju M / M ± 1 · s / M ± 2 · s, ne pojavljuju aberantni rezultati?

Na ovaj način se prikazuje teorijska distribucija koja ima unapred zadate parametre. U teorijskim distribucijama ne postoje aberantni rezultati.

> Koliki procenat rezultata obuhvataju kutija i „brkovi" u slučaju M / IKR / R a koliki u slučaju M / M ± 1 · s / M ± 2 · s? Koja od opcija prikazuje procenat rezultata u različitim delovima opažene distribucije podataka?

U prvom slučaju kutija obuhvata 50% a „brkovi" 100% *empirijski dobijenih*, tj. opaženih rezultata, izuzimajući eventualne autlajere. U drugom slučaju, kutijom je obuhvaćeno oko 68% a „brkovima" oko 95% *teorijskih* rezultata, tj. onih koji bi mogli da se očekuju uz pretpostavku da je distribucija normalna.

3.5.1. Hi-kvadrat kao test nezavisnosti

> Da li se broj kolona u matrici sirovih podataka za primer *Operater x fakultet* promenio u odnosu na primere *Usluga x fakultet*?

Nije. Matrica se izmenila utoliko što ima više redova, jer je uzorak veći. Broj kolona ostao je isti, jer je i broj varijabli isti. Opet se ukrštaju dve varijable, ali sada obe imaju po tri nivoa, odnosno tri moguće vrednosti. Preuzmite podatke za različite primere da biste videli u čemu se ove matrice razlikuju.

> Izmenite opažene frekvencije u svim ćelijama tabele kontingencije za primer *Ispit x udžbenik* tako da se dobije najveća moguća ϕ koeficijenta koja u ovoj vežbi može da bude 0,78 ili -0,78.

Potrebno je smanjiti frekvencije ćelija jedne od dijagonala i povećati frekvencije u ćelijama one druge.

> Odaberite ponovo primer *Operater x fakultet* i izmenite vrednosti opaženih frekvencija tako da postignete najveću vrednost C koeficijenta koja u našem primeru može da bude 0,70. Koliko takvih rešenja postoji?

Dva moguća rešenja nameću se odmah, a to je da se vrednosti ćelija jedne od dijagonala postave na 40 a vrednosti svih ostalih ćelija na 5. Međutim, tačno je svako rešenje kod koga se ćelije sa visokim učestalostima ne nalaze u istom redu ili u istoj koloni, npr. FF-A, FTN-B i PMF-C. Ukupan broj rešenja je 6. Povezanost među varijablama je veća ako su vrednosti na jednoj od njih povezane samo sa određenim vrednostima na drugoj.

> Kolika je vrednost koeficijenta C kada su opažene frekvencije u svim ćelijama tabele kontingencije jednake?

Vrednost koeficijenta C biće nula jer je vrednost χ^2 testa nula.

3.6.1.1. Smisao koeficijenta b i konstante a u regresionoj analizi

> Da li biste na osnovu regresione jednačine $y' = 8,26 + 3,42 \cdot x$ mogli da zaključite kolika je približno korelacija varijabli X i Y?

Ne. Regresiona jednačina predstavlja samo „uputstvo" na koji način treba da se transformiše vrednost x da bi se dobila vrednost y, ali uz pretpostavku da korelacija varijabli postoji i da je njihov odnos linearan.

> Da li biste na osnovu regresione jednačine $y' = 0 + 7,65 \cdot x$ mogli da zaključite da li je ona dobijena na osnovu sirovih ili standardizovanih podataka?

Ako je regresiona jednačina formirana na osnovu standardizovanih varijabli, vrednost koeficijenta β biće jednaka koeficijentu korelacije. S obzirom na to da vrednosti koeficijenta korelacije mogu da se kreću samo u rasponu od -1 do 1, gornja jednačina je očigledno dobijena na osnovu sirovih podataka.

> Da li bi se korelacija dve varijable promenila kada bi se vrednosti svake od njih transformisale u percentilne rangove?

Da. Za razliku od standardizacije pretvaranjem u z vrednosti, transformacijom podataka u percentilne rangove, menja se oblik distribucije. Bez obzira na njen prvobitni oblik, distribucija transformisanih vrednosti biće slična uniformnoj. Distance među vrednostima više ne odražavaju stvarnu razliku u izraženosti svojstva već samo razliku u rangu.

> Da li na osnovu boja ćelija u tabeli sa leve strane možete da odredite gde se na grafikonu nalazi određeni kružić, odnosno ispitanik?

Postoje četiri kombinacije crvene i zelene boje koje govore u kom delu grafikona, tačnije u kom kvadrantu se nalazi ispitanik. Parovi vrednosti čija je kombinacija zelena – zelena, nalaze se u gornjem desnom kvadrantu. Parovi crvenih i zelenih kućica odnose se na rezultate koji se nalaze u gornjem levom kvadrantu. Intenzitet boje govori o udaljenosti rezultata od proseka.

> Primeri *Učenje x bodovi 3* i *Prijemni x ESPB 1* ukazuju na nisku ili nepostojeću povezanost varijabli. Šta biste mogli da zaključite o prirodi te (niske) korelacije varijabli?

Iako je korelacija u oba primera jednako niska, u prvom ipak mogu da se uoče određene pravilnosti, odnosno klasteri ispitanika. Stoga bi trebalo detaljnije istražiti potencijalne razloge ovakvog raspršenja. U drugom primeru zaista ne postoji nikakva povezanost između varijabli.

3.6.2. Standardna greška procene

> Zbog čega su u regresionoj analizi intervali procene uvek širi od odgovarajućih intervala pouzdanosti?

Zato što se na osnovu prvih procenjuje varijabilnost prognoziranih vrednosti y varijable, a na osnovu drugih varijabilnost njihovog proseka. Razlika je analogna onoj koja postoji između intervala $M \pm s$ i $M \pm s_M$.

> Da li je u primeru *Učenje x bodovi 2* koeficijent b značajno različit od nule?

Da. Nula se gotovo sigurno nalazi izvan raspona koji se može grubo izračunati kao $b \pm 3 \cdot s_b$.

> Zbog čega intervali procene u primeru *Učenje x bodovi 2* ne obuhvataju sve prikazane rezultate već se dva kružića nalaze izvan njega?

Svi intervali procene ili pouzdanosti koje smo do sada pominjali baziraju se na očekivanju da će se njima obuhvatiti određeni procenat rezultata. Taj procenat nikada ne može da dostigne vrednost 100%. Dakle, verovatnoća da postoji (aberantna) vrednost koja se nalazi izvan definisanih intervala, nikada nije nulta.

3.6.4. Uslovi za primenu Pirsonovog r

> Kako biste, na osnovu raspršenja kružića u primeru *Televizija x visina*, zaključili kakvog su oblika distribucije dveju varijabli?

Raspršenje i broj rezultata u zoni središnjih vrednosti obe varijable veoma su mali u odnosu na raspone niskih i visokih vrednosti. To bi moglo da znači da su u pitanju dve bimodalne distribucije podataka.

> Kakav je oblik distribucija varijabli u primeru *Pušenje x kapacitet 2*?

S obzirom na to da raspršenje i gustina rezultata nisu iste u zoni visokih i niskih vrednosti, možemo da zaključimo da je varijabla X iskošena u desnu a varijabla Y u levu stranu.

> Da li odstupanje varijabli X i/ili Y od normalnosti podrazumeva da njihov odnos neće biti linearan?

Ne. Linearnost je karakteristika odnosa dve varijable a ne svojstvo njihovih distribucija. Varijable mogu da imaju linearan ili nelinearan odnos, bez obzira na to kako su distribuirane.

> Na osnovu čega se pomoću skater-dijagrama može utvrditi da su obe varijable, barem približno, normalno distribuirane?

Ukoliko su obe varijable normalno distribuirane, gustina kružića bi trebalo da bude najveća u zoni središnjih vrednosti.

3.8. T-test za zavisne uzorke

> Prikažite *Primer 4* i pokušajte da promenama rezultata samo jednog ispitanika dobijete statistički značajnu vrednost t-testa.

Ovo možete da postignete tako što ćete, na primer, brzinu najsporijeg pilota u trećem merenju, koja iznosi oko 750 ms, promeniti na 350 ms. Prosek trećeg merenja svakako je manji od proseka drugog, a ovom promenom ta razlika se dodatno povećava.

> U grupi ispitanika evidentiran je broj zapamćenih besmislenih slogova neposredno nakon učenja i dva sata kasnije. Šta je zavisna a šta nezavisna varijabla u ovom istraživanju?

Zavisna varijabla je broj tačno reprodukovanih slogova koji je meren u dva navrata. Nezavisna varijabla je vreme, odnosno protok vremena.

> Da li bi t-test u četvrtom primeru bio statistički značajan da je primenjeno jednostrano testiranje razlike? Koja vrsta testiranja razlike je primerenija u ovom istraživanju?

Vrednost t-testa bila bi značajna na nivou 0,05, jer bi p nivo, za istu vrednost t-testa, bio duplo manji. U ovom slučaju jednostrano testiranje ima više opravdanja zato što se uvežbavanje, odnosno promena karakteristika stimulusa, vrši upravo da bi se postiglo povećanje brzine reakcije. Negativna razlika između drugog i trećeg merenja je očekivan rezultat.

> Pokušajte da povećate standardnu devijaciju trećeg merenja u trećem primeru na oko 45 ms, a da pri tome korelacija drugog i trećeg merenja ostane približno ista.

Rešenje podrazumeva da se linije razvuku u oblik lepeze, tako da je razmak između prve i poslednje dovoljno velik da se postigne tražena varijabilnost, a da odnos razlika među njima ostane očuvan u odnosu na početne pozicije.

> Koja vrsta greške u zaključivanju je verovatnija ako se na dve grupe ispitanika sa uparenim rezultatima primeni t-test za nezavisne uzorke?

Ako među merenjima postoji pozitivna korelacija, veća je verovatnoća da nećemo utvrditi postojanje razlike, tj. da ćemo napraviti grešku tipa β. Sa druge strane, ako je korelacija negativna, raste verovatnoća greške tipa α. U prvom slučaju, isključivanje koeficijenta korelacije povećava standardnu grešku razlike a u drugom je umanjuje.

> U čemu se razlikuju matrice sirovih podataka, na osnovu kojih se računa t-test za nezavisne uzorke i t-test za zavisne uzorke?

U slučaju t-testa za nezavisne varijable potrebna je jedna kategorijalna i jedna kvantitativna varijabla intervalnog ili racio nivoa. Na primer, u prvu kolonu bi se unosio podatak o polu ispitanika a u drugu podatak o njihovoj visini. U slučaju t-testa za zavisne, obe kolone sadrže kvantitativne vrednosti zavisne varijable. U prvoj su merenja obavljena pre uvođenja tretmana a u drugoj posle njega. Kod t-testa za uparene rezultate, ne postoji grupišuća varijabla, već su merenja razvrstana po kolonama.

> Prisetite se primera sa Miler-Lajerovom iluzijom u kome je svih 14 merenja obavljeno na istom uzorku, tačnije na istom ispitaniku. Da li je zbog toga opravdano primeniti t-test za zavisne umesto t-testa za nezavisne uzorke?

Iako činjenica da su sva merenja obavljena na istom ispitaniku implicira nacrt za zavisne uzorke, to nije u potpunosti opravdano. Naime, primena ove metode podrazumeva da se svako merenje iz jedne grupe upari sa tačno određenim merenjem iz druge grupe. Kao kriterijum povezivanja mogao bi da se upotrebi redosled izlaganja stimulusa, npr. prvo sa drugim, treće sa četvrtim i tako dalje, ali ovakvo uparivanje je jednako (ne)legitimno kao i uparivanje po bilo kom drugom kriterijumu. U tom smislu, primena t-testa za zavisne uzorke mogla bi da dovede do pogrešnog zaključka zbog uključivanja koeficijenta korelacije koji potpuno slučajno može da bude visok ili nizak. Najverovatnije je, međutim, da primena različitih modela t-testa ne bi dovela do bitno drugačijih rezultata.

3.9. Neparametrijske alternative t-testu za zavisne uzorke

> Na koji način se još može proveriti značajnost odstupanja opaženih frekvencija 4 i 8 od očekivanih frekvencija 6 i 6?

Značajnost razlike između opaženih i očekivanih frekvencija mogla bi da se testira primenom χ^2 testa za jedan uzorak.

> Pomerajte linije na grafikonu da biste utvrdili koliki broj promena bi se smatrao statistički značajnim ako je veličina uzorka 12.

Da bi razlika bila značajna na nivou 0,05, broj promena u jednom smeru mora da bude barem 10.

> Da li bi se rezultati t-testa bitno izmenili kada bi se on primenio na podacima koji su pretvoreni u rangove na gore opisani način, odnosno objedinjeno za obe grupe ispitanika ili oba merenja?

Ne bi. Razlike između aritmetičkih sredina rangova bile bi takođe značajne.

3.10. Značajnost razlika uparenih podataka nominalnog nivoa

> Koliko kolona i koliko redova ima matrica sirovih podataka iz koje je nastala prikazana tabela kontingencije? Kako biste nazvali kolone u matrici?

Matrica ima 300 redova i dve kolone. U svakoj koloni pojavljuju se vrednosti DA i NE, a varijable bi mogle da se nazovu *pre* i *posle* ili *početak* i *kraj* (polugodišta).

> Zbog čega tabela kontingencije u ovom primeru ne bi mogla i ne bi trebalo da se formira ukrštanjem vremena merenja (PRE / POSLE) i odgovora na pitanje (DA / NE)?

U tabelama kontingencije svaki ispitanik, odnosno svako merenje, može i mora da se nađe u samo jednoj od ćelija. U slučaju zavisnih uzoraka, to znači da pozicija u ćeliji mora da bude određena vrednostima na prvom i drugom merenju, npr. DA - DA ili NE - DA. Kada bi tabela bila formirana na način koji je naveden u pitanju, isti ispitanik bi se našao u dve ćelije, što nije dozvoljeno.

3.10.1. Maknimarov test

> Kako bi trebalo izmeriti konzumaciju energetskih pića da bi se omogućila primena t-testa za zavisne uzorke? Objasnite zbog čega bi u toj situaciji Maknimarov test imao manju snagu od t-testa za zavisne uzorke?

Ako bismo, na primer, upotrebu energetskih pića izrazili kao količinu ispijenu u toku nedelju dana, mogli bismo da primenimo t-test za zavisne uzorke. Ako se konzumacija izrazi dihotomno, sve „fine“ razlike ostaju skrivene, pa đaci koji popiju jedno piće nedeljno i dva pića dnevno spadaju u istu kategoriju. Merenje varijable na razmernom nivou povećava verovatnoću utvrđivanja kvantitativne razlike među grupama.

> Pomeranjem traka napravite primer u kome je 200 od 210 ispitanika promenilo svoje ponašanje ali razlika između prvog i drugog merenja nije statistički značajna. Da li je ovaj rezultat, po vašem mišljenju, logičan i opravdan?

Frekvencije ćelija a, b, c i d treba da se postave na vrednosti 5, 100, 100, 5. Razlika nije značajna na nivou cele grupe pošto je odnos broja đaka koji piju i onih koji ne piju energetska pića posle kampanje isti kao i pre nje. Međutim, očigledno je da se na nivou pojedinaca desila drastična promena, jer su svi đaci listom promenili svoje ponašanje. Ovaj fenomen bi sigurno trebalo dodatno istražiti.

> Da li bi neki od postupaka pomenutih u ranijim odeljcima ipak mogao da ukaže da u primeru koji ste formirali postoji statistički značajna pravilnost?

Koeficijent korelacije ϕ ukazao bi na postojanje visoke povezanosti, odnosno pravilnosti u odnosu promena na prvom i drugom merenju.

3.10.2. Testovi marginalne homogenosti za politomne varijable

> Menjajte frekvenciju kategorija 1 – 1 a potom i 1 – 2 u primeru *Zadovoljstvo 2*. Posmatrajte kako promene utiču na vrednost χ^2, odnosno T testa. Zbog čega promene frekvencija u glavnoj dijagonali tabele kontingencije ne utiču na vrednosti testova homogenosti marginalnih frekvencija?

Promena vrednosti u ćeliji 1 - 1 ne utiče na vrednosti testova, jer se na taj način menjaju iste marginalne frekvencije vrednosti prvog i drugog merenja. Logika testova marginalne homogenosti bazira se na analizi ćelija van glavne dijagonale, na osnovu kojih se može zaključiti da je velika verovatnoća ishoda na prvom merenju, povezana sa velikom verovatnoćom drugačijeg ishoda na drugom. Stoga, promena u ćeliji 1 - 2 utiče na vrednosti testova i povećanje njihove značajnosti.

LITERATURA

Ackoff, R. L. (1989). From data to wisdom. *Journal of Applied Systems Analysis*, *16*, 3–9.

Agresti, A. (1983). Testing marginal homogeneity for ordinal categorical variables. *Biometrics*, *39*(2), 505–510.

Agresti, A. (2002). *Categorical Data Analysis*. https://doi.org/10.1007/978-3-642-04898-2_161

Amit, E., Hoeflin, C., Hamzah, N., & Fedorenko, E. (2017). An asymmetrical relationship between verbal and visual thinking: Converging evidence from behavior and fMRI. *NeuroImage*, *152*, 619–627. https://doi.org/10.1016/j.neuroimage.2017.03.029

Anscombe, F. J. (1973). Graphs in Statistical Analysis. *The American Statistician*, *27*(1), 17–21. https://doi.org/10.1080/00031305.1973.10478966

APA. (2010). *Publication Manual of the American Psychological Association, 6th Edition*. Washington, DC: APA.

Banse, R., Messer, M., & Fischer, I. (2015). Predicting aggressive behavior with the aggressiveness-IAT. *Aggressive Behavior*, *41*(1), 65–83. https://doi.org/10.1002/ab.21574

Bauer, M. I., & Johnson-Laird, P. N. (1993). How Diagrams Can Improve Reasoning. *Psychological Science*, *4*(6), 372–378. https://doi.org/10.1111/j.1467-9280.1993.tb00584.x

Bennet, J. H. (Ed.). (1990). *Statistical inference and analysis: Selected correspondence of RA Fisher*. Oxford: Clarendon Press.

Berg, R. V., Cornelissen, F. W., & Roerdink, J. B. T. M. (2008). Perceptual dependencies in information visualization assessed by complex visual search. *ACM Transactions on Applied Perception*, *4*(4), Article 22.

Bertin, J. (1983). *Semiology of Graphics: Diagrams, Networks, Maps*. University of Wisconsin Press.

Börner, K., Maltese, A., Balliet, R. N., & Heimlich, J. (2016). Investigating aspects of data visualization literacy using 20 information visualizations and 273 science museum visitors. *Information Visualization*, *15*(3), 198–213. https://doi.org/10.1177/1473871615594652

Brin, S., & Page, L. (1998). The anatomy of a large-scale hypertextual Web search engine. *Computer Networks and ISDN Systems*, *30*(1), 107–117. https://doi.org/10.1016/S0169-7552(98)00110-X

Cairo, A. (2013). *The Functional Art: An introduction to information graphics and visualization*. Berkeley: New Riders.

Card, S. K., Mackinlay, J. D., & Shneiderman, B. (1999). *Readings in information visualization: Using vision to think*. San Francisco: Morgan Kaufmann.

Chalmers, M. (1993). Using a landscape metaphor to represent a corpus of documents. *Proceedings of the European Conference on Spatial Information Theory, Elba, September 1993*, 377–390.

Chen, C. (2006). *Information visualization: Beyond the horizon*. London: Springer.

Clark, J. M., & Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review*, *3*(3), 149–210. https://doi.org/10.1007/BF01320076

Cleveland, W. S., & McGill, R. (1984). Graphical perception: Theory, experimentation, and application to the development of graphical methods. *Journal of the American Statistical Association*, *79*(387), 531–554.

Cohen, J. (1960). A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, *20*(1), 37–46. https://doi.org/10.1177/001316446002000104

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. New York: Lawrence Erlbaum Associates.

Cole, C., Mandelblatt, B., & Stevenson, J. (2002). Visualizing a high recall search strategy output for undergraduates in an exploration stage of researching a term paper. *Information Processing & Management*, *38*(1), 37–54.

Čolović, P., Smederevac, S., & Mitrović, D. (2014). Velikih pet plus dva: Validacija skraćene verzije. *Primenjena psihologija*, *7*(3–1), 227–254. https://doi.org/10.19090/pp.2014.3-1.227-254

Cope, B., & Kalantzis, M. (2009). "Multiliteracies": New Literacies, New Learning. *Pedagogies: An International Journal*, *4*(3), 164–195. https://doi.org/10.1080/15544800903076044

Dale, E., & Chall, J. S. (1949). The Concept of Readability. *Elementary English*, *26*(1), 19–26. Retrieved from JSTOR.

Dhand, N. K., & Khatkar, M. S. (2014). Statulator: An online statistical calculator. Sample Size Calculator for Comparing Two Independent Means. Retrieved July 29, 2019, from http://statulator.com/SampleSize/ss2M.html

Dowse, R., & Ehlers, M. S. (2001). The evaluation of pharmaceutical pictograms in a low-literate South African population. *Patient Education and Counseling*, *45*(2), 87–99. https://doi.org/10.1016/S0738-3991(00)00197-X

Evans, J. D. (1996). *Straightforward statistics for the behavioral sciences*. Pacific Grove, CA: Brooks/Cole Publishing.

Fischer, H. (2011). *A history of the central limit theorem: From classical to modern probability theory* (H. Fischer, Ed.). https://doi.org/10.1007/978-0-387-87857-7_2

Fisher, R. A. (1922). On the Interpretation of χ2 from Contingency Tables, and the Calculation of P. *Journal of the Royal Statistical Society*, *85*(1), 87–94. https://doi.org/10.2307/2340521

Fisher, R. A. (1925). Applications of "Student's" distribution. *Metron*, *5*(3), 90–104.

Friendly, M., & Denis. (2001). Milestones in the History of Thematic Cartography, Statistical Graphics, and Data Visualization. Retrieved February 27, 2018, from http://datavis.ca/milestones/

Fruchterman, T. M. J., & Reingold, E. M. (1991). Graph drawing by force-directed placement. *Software: Practice and Experience*, *21*(11), 1129–1164.

Galton, F. (1886). Regression Towards Mediocrity in Hereditary Stature. *The Journal of the Anthropological Institute of Great Britain and Ireland*, *15*, 246–263. https://doi.org/10.2307/2841583

Gibson, J. J. (1977). The theory of affordances. In R. E. Shaw & J. Bransford (Eds.), *Perceiving, acting, and knowing* (pp. 127–143). Hillsdale, NJ: Lawrence Erlbaum and Associates.

Gigerenzer, G. (2004). Mindless statistics. *The Journal of Socio-Economics*, *33*(5), 587–606. https://doi.org/10.1016/j.socec.2004.09.033

Guilford, J. P. (1978). *Fundamental statistics in psychology and education*. New York: McGraw-Hill.

Haller, H., & Krauss, S. (2002). Misinterpretations of significance: A problem students share with their teachers. *Methods of Psychological Research*, *7*(1), 1–20.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, *33*(2–3), 61–83. https://doi.org/10.1017/S0140525X0999152X

Hinkle, D. E., Wiersma, W., & Jurs, S. G. (2003). *Applied Statistics for the Behavioral Sciences*. Boston: Houghton Mifflin.

Horn, R. E. (1999). Information design: Emergence of a new profession. In R. E. Jacobson, *Information design* (pp. 15–33). Cambridge, MA: MIT Press.

Howell, D. C. (2012). *Statistical Methods for Psychology*. Belmont: Wadsworth, Cengage Learning.

Hurlbert, A., & Ling, Y. (2012). Understanding colour perception and preference. In J. Best (Ed.), *Colour Design* (pp. 129–157). https://doi.org/10.1533/9780857095534.1.129

Iaccino, J. F. (2014). *Left brain – right brain differences: Inquiries, evidence, and new approaches*. New York: Psychology Press.

IBM. (2016). *IBM SPSS Statistics 24 Algorithms*. Retrieved from ftp://public.dhe.ibm.com/software/analytics/spss/documentation/statistics/24.0/en/client/Manuals/IBM_SPSS_Statistics_Algorithms.pdf

Kaltenbach, H.-M. (2012). *A Concise Guide to Statistics*. Retrieved from https://www.springer.com/la/book/9783642235016

Kamada, T., & Kawai, S. (1989). An algorithm for drawing general undirected graphs. *Information Processing Letters*, *31*(1), 7–15.

Kievit, R., Frankenhuis, W. E., Waldorp, L., & Borsboom, D. (2013). Simpson's paradox in psychological science: A practical guide. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00513

Knapp, M. L., Hall, J. A., & Horgan, T. G. (2013). *Nonverbal Communication in Human Interaction*. Boston: Wadsworth Cengage Learning.

Knapp, T. R. (1990). Treating ordinal scales as interval scales: An attempt to resolve the controversy. *Nursing Research*, *39*(2), 121–123.

Kodžopeljić, J., Smederevac, S., Mitrović, D., Ćolović, P., & Pajić, D. (2019). *Velikih pet plus dva – Verzija za decu: Primena i interpretacija*. Beograd: Centar za primenjenu psihologiju.

Koshman, S. (2006). Visualization-based information retrieval on the web. *Library & Information Science Research*, *28*(2), 192–207.

Krider, R. E., Raghubir, P., & Krishna, A. (2001). Pizzas: π or Square? Psychophysical Biases in Area Comparisons. *Marketing Science*, *20*(4), 405–425. https://doi.org/10.1287/mksc.20.4.405.9756

Krstić, D. (1991). *Psihološki rečnik*. Beograd: Savremena administracija.

Lakoff, G., & Núñez, R. E. (2000). *Where Mathematics Comes From: How the Embodied Mind Brings Mathematics Into Being*. Retrieved from https://www.maa.org/press/maa-reviews/where-mathematics-comes-from-how-the-embodied-mind-brings-mathematics-into-being

Landers, R. N., & Lounsbury, J. W. (2006). An investigation of Big Five and narrow personality traits in relation to Internet usage. *Computers in Human Behavior*, *22*(2), 283–293. https://doi.org/10.1016/j.chb.2004.06.001

Lindell, A. K., & Kidd, E. (2011). Why right-brain teaching is half-witted: A critique of the misapplication of neuroscience to education. *Mind, Brain, and Education*, *5*(3), 121–127. https://doi.org/10.1111/j.1751-228X.2011.01120.x

LoBue, V., & DeLoache, J. S. (2008). Detecting the Snake in the Grass: Attention to Fear-Relevant Stimuli by Adults and Young Children. *Psychological Science*, *19*(3), 284–289. https://doi.org/10.1111/j.1467-9280.2008.02081.x

Loftus, G. R. (1993). A picture is worth a thousand p values: On the irrelevance of hypothesis testing in the microcomputer age. *Behavior Research Methods, Instruments, & Computers*, *25*(2), 250–256. https://doi.org/10.3758/BF03204506

Lord, F. M. (1967). A paradox in the interpretation of group comparisons. *Psychological Bulletin*, *68*(5), 304.

Marcus-Roberts, H. M., & Roberts, F. S. (1987). Meaningless Statistics. *Journal of Educational Statistics*, *12*(4), 383–394. https://doi.org/10.3102/10769986012004383

Mathewson, J. H. (1999). Visual-spatial thinking: An aspect of science overlooked by educators. *Science Education*, *83*(1), 33–54. https://doi.org/10.1002/(SICI)1098-237X(199901)83:1<33::AID-SCE2>3.0.CO;2-Z

Maxwell, A. E. (1970). Comparing the Classification of Subjects by Two Independent Judges. *The British Journal of Psychiatry*, *116*(535), 651–655. https://doi.org/10.1192/bjp.116.535.651

Mazza, R. (2009). *Introduction to information visualization*. New York: Springer.

McCarthy, M. S., & Mothersbaugh, D. L. (2002). Effects of typographic factors in advertising-based persuasion: A general model and initial empirical tests. *Psychology & Marketing*, *19*(7–8), 663–691. https://doi.org/10.1002/mar.10030

McLean, S. A., Paxton, S. J., Wertheim, E. H., & Masters, J. (2015). Photoshopping the selfie: Self photo editing and photo investment are associated with body dissatisfaction in adolescent girls. *International Journal of Eating Disorders*, *48*(8), 1132–1140. https://doi.org/10.1002/eat.22449

McNemar, Q. (1947). Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika*, *12*(2), 153–157. https://doi.org/10.1007/BF02295996

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*(2), 81–97.

Neubauer, A. C., & Fink, A. (2009). Intelligence and neural efficiency. *Neuroscience & Biobehavioral Reviews*, *33*(7), 1004–1023. https://doi.org/10.1016/j.neubiorev.2009.04.001

Neyman, J., & Pearson, E. S. (1928). On the Use and Interpretation of Certain Test Criteria for Purposes of Statistical Inference: Part I. *Biometrika*, *20A*(1/2), 175–240. https://doi.org/10.2307/2331945

Neyman, Jerzy, & Pearson, E. S. (1933). On the problem of the most efficient tests of statistical hypotheses. *Philosophical Transactions of the Royal Society of London. Series A*, *231*(694–706), 289–337. https://doi.org/10.1098/rsta.1933.0009

Nickerson, R. S. (2000). Null hypothesis significance testing: A review of an old and continuing controversy. *Psychological Methods*, *5*(2), 241–301.

Norman, D. A. (1988). *The Psychology of Everyday Things*. New York: Basic Books.

Norman, G. (2010). Likert scales, levels of measurement and the "laws" of statistics. *Advances in Health Sciences Education*, *15*(5), 625–632. https://doi.org/10.1007/s10459-010-9222-y

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251), aac4716. https://doi.org/10.1126/science.aac4716

Paivio, A. (1986). *Mental Representations: A Dual Coding Approach*. New York: Oxford University Press.

Palmer, S., & Rock, I. (1994). Rethinking perceptual organization: The role of uniform connectedness. *Psychonomic Bulletin & Review*, *1*(1), 29–55. https://doi.org/10.3758/BF03200760

Panik, M. (2005). *Advanced Statistics from an Elementary Point of View*. Amsterdam: Elsevier Academic Press.

Popper, K. (2002). *The Logic of Scientific Discovery*. London and New York: Routledge.

Prensky, M. (2001). Digital Natives, Digital Immigrants Part 1. *On the Horizon*, *9*(5), 1–6. https://doi.org/10.1108/10748120110424816

Puente, A. E. (2012). Roger W. Sperry: From Neuro-Science to Neuro-Philosophy. In A. Y. Stringer, E. L. Cooley, & A.-L. Christensen (Eds.), *Pathways to Prominence in Neuropsychology: Reflections of Twentieth-Century Pioneers* (pp. 63–76). New York: Psychology Press.

Quinn, G. E., Shin, C. H., Maguire, M. G., & Stone, R. A. (1999). Myopia and ambient lighting at night. *Nature*, *399*(6732), 113. https://doi.org/10.1038/20094

Reed, S. K. (2013). *Thinking Visually*. New York: Psychology Press.

Reinhart, A. (2015). *Statistics Done Wrong: The Woefully Complete Guide*. San Francisco: No Starch Press.

Robertson, P. K. (1991). A methodology for choosing data representations. *Computer Graphics and Applications, IEEE*, *11*(3), 56–67.

Robinson, W. S. (2009). Ecological Correlations and the Behavior of Individuals. *International Journal of Epidemiology*, *38*(2), 337–341. https://doi.org/10.1093/ije/dyn357

Rugg, G. (2007). *Using Statistics: A Gentle Introduction*. McGraw-Hill Education.

Ryan, L. (2016). Visual communication and literacy. In *The visual imperative* (pp. 109–130). https://doi.org/10.1016/B978-0-12-803844-4.00006-6

Salsburg, D. (2001). *The lady tasting tea: How statistics revolutionized science in the twentieth century*. New York: W. H. Freeman and Company.

Sharpe, D., & Poets, S. (2017). Canadian psychology department participant pools: Closing for the season? *Canadian Psychology/Psychologie Canadienne*, *58*(2), 168–177. https://doi.org/10.1037/cap0000090

Shen, W., Kiger, T. B., Davies, S. E., Rasch, R. L., Simon, K. M., & Ones, D. S. (2011). Samples in applied psychology: Over a decade of research in review. *The Journal of Applied Psychology*, *96*(5), 1055–1064. https://doi.org/10.1037/a0023322

Sherman, R. C., Buddie, A. M., Dragan, K. L., End, C. M., & Finney, L. J. (1999). Twenty Years of PSPB: Trends in Content, Design, and Analysis. *Personality and Social Psychology Bulletin*, *25*(2), 177–187. https://doi.org/10.1177/0146167299025002004

Simpson, E. H. (1951). The Interpretation of Interaction in Contingency Tables. *Journal of the Royal Statistical Society: Series B (Methodological)*, *13*(2), 238–241. https://doi.org/10.1111/j.2517-6161.1951.tb00088.x

Singh, M., & Hoffman, D. D. (2013). Natural Selection and Shape Perception. In *Advances in Computer Vision and Pattern Recognition*. *Shape Perception in Human and Computer Vision* (pp. 171–185). https://doi.org/10.1007/978-1-4471-5195-1_12

Smart, R. G. (1966). Subject selection bias in psychological research. *Canadian Psychologist/Psychologie Canadienne*, *7a*(2), 115–121. https://doi.org/10.1037/h0083096

Sotelo, M. M., & Johnson, S. R. (1997). The effects of hormone replacement therapy on coronary heart disease. *Endocrinology and Metabolism Clinics of North America*, *26*(2), 313–328. https://doi.org/10.1016/S0889-8529(05)70249-8

Sperry, R. W. (1973). Lateral specialization of cerebral function in the surgically separated hemispheres. In F. J. McGuigan & R. A. Schoonover, *The psychophisiology of thinking: Studies of covert processes*. New York: Academic Press.

St. Clair, S. (2000). Visual Metaphor, Cultural Knowledge, and the New Rhetoric. In J. Reyhner, J. Martin, L. Lockard, & W. Sakiestewa Gilbert

(Eds.), *Learn In Beauty: Indigenous Education for a New Century*. Retrieved from https://eric.ed.gov/?id=ED445871

Stafford, B. M. (1998). *Good Looking: Essays on the Virtue of Images*. Cambridge, MA: MIT Press.

Standing, L. (1973). Learning 10000 pictures. *Quarterly Journal of Experimental Psychology*, *25*(2), 207–222. https://doi.org/10.1080/14640747308400340

Stanfield, R. A., & Zwaan, R. A. (2001). The Effect of Implied Orientation Derived from Verbal Context on Picture Recognition. *Psychological Science*, *12*(2), 153–156. https://doi.org/10.1111/1467-9280.00326

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*(6), 643–662. https://doi.org/10.1037/h0054651

Stuart, A. (1955). A Test for Homogeneity of the Marginal Distributions in a Two-Way Classification. *Biometrika*, *42*(3/4), 412–416. https://doi.org/10.2307/2333387

Student. (1908). The Probable Error of a Mean. *Biometrika*, *6*(1), 1–25. https://doi.org/10.2307/2331554

Tabachnick, B. G., & Fidell, L. S. (2014). *Using Multivariate Statistics*. Retrieved from https://scholar.google.com/scholar?hl=sr&as_sdt=0%2C5&authuser=1&q=Using+Multivariate+Statistics&btnG=

Toga, A. W., & Thompson, P. M. (2003). Mapping brain asymmetry. *Nature Reviews Neuroscience*, *4*(1), 37–48. https://doi.org/10.1038/nrn1009

Treisman, A. (1986). Preattentive Processing in Vision. In A. Rosenfeld (Ed.), *Human and Machine Vision II* (pp. 313–334). https://doi.org/10.1016/B978-0-12-597345-8.50017-0

Treisman, A., & Gormican, S. (1988). Feature analysis in early vision: Evidence from search asymmetries. *Psychological Review*, *95*(1), 15.

Trumbo, J. (1999). Visual literacy and science communication. *Science Communication*, *20*(4), 409–425. https://doi.org/10.1177/1075547099020004004

Tufte, E. R. (1985). *The visual display of quantitative information*. Cheshire: Graphics Press.

Tukey, J. W. (1977). *Exploratory data analysis*. Reading, MA: Addison-Wesley.

Unwin, A., Chen, C., & Härdle, W. K. (2008). Introduction. In Chun-houh Chen, W. K. Härdle, & A. Unwin (Eds.), *Handbook of data visualization* (pp. 3–12). Berlin: Springer.

Ware, C. (2004). *Information Visualization: Perception for design*. San Francisco: Morgan Kaufmann.

Weiser, M. (1999). The computer for the 21st century. *Mobile Computing and Communications Review*, *3*(3), 3–11.

Westergaard, G. C., Suomi, S. J., Chavanne, T. J., Houser, L., Hurley, A., Cleveland, A., … Higley, J. D. (2003). Physiological Correlates of Aggression and Impulsivity in Free-Ranging Female Primates. *Neuropsychopharmacology*, *28*(6), 1045. https://doi.org/10.1038/sj.npp.1300171

Wildgen, W. (2004). The Paleolithic Origins of Art, its Dynamic and Topological Aspects, and the Transition to Writing. In M. Bax, B. van Heusden, & W. Wildgen (Eds.), *Semiotic Evolution and the Dynamics of Culture* (pp. 117–153). Bern: Peter Lang.

Zhang, J. (2010). *Visualization for information retrieval*. Berlin: Springer.

UNIVERZITET U NOVOM SADU
FILOZOFSKI FAKULTET
21000 Novi Sad
Dr Zorana Đinđića 2
www.ff.uns.ac.rs

Interaktivna verzija
http://psihologija.ff.uns.ac.rs/viz/

CIP - Каталогизација у публикацији
Библиотеке Матице српске, Нови Сад

159.9:311(075.8)

ПАЈИЋ, Дејан, 1973-
Primena tehnika vizualizacije u bazičnoj statistici / Dejan Pajić. - Novi Sad : Filozofski fakultet, 2020 ([s. l. : s. n.]). - 283 str. : ilustr. ; 26 cm

Bibliografija.

ISBN 978-86-6065-582-2

а) Психологија -- Технике визуализације -- Статистика

COBISS.SR-ID 332998663