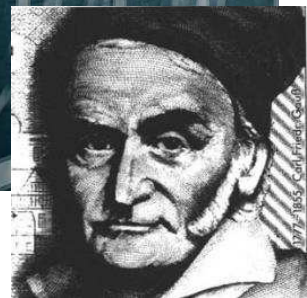


SELMAN  
REPIŠTI

# ZBIRKA ZADATAKA IZ STATISTIKE U DRUŠTVENIM NAUKAMA



Elektronska verzija (v1.0) | 2017

Selman Repišti

**ZBIRKA ZADATAKA IZ  
STATISTIKE U DRUŠTVENIM NAUKAMA**

Elektronska verzija / 2017

### ***Važna napomena***

Zbirku mogu koristiti nastavnici i saradnici u nastavi statistike tokom predavanja i vježbi, kao i studenti tokom spremanja ispita. Zapravo, ova zbirka je i napisana kao pomoć studentima da još bolje nauče gradivo iz statistike, te provježbaju karakteristične tipove zadataka i odgovore na s njima povezana pitanja. Stoga je nekomercijalna upotreba sadržaja dozvoljena.

Strogo je zabranjeno korištenje dijelova teksta, grafikona, tabela ili cijele knjige u svrhu lične zarade. Zabranjeno je i prikazivanje dijelova ili čitavog teksta kao autorskog vlasništva drugih lica (plagijarizam).

©2017 *Selman Repišti*

## SADRŽAJ

|   |    |
|---|----|
| Predgovor .....   | 1  |
| O notaciji, formatu pitanja/zadataka i jeziku .....                 | 3  |
| <i>Teorijska racionala, pitanja i zadaci:</i>                       |    |
| Tabelarno i grafičko organizovanje i prikazivanje podataka .....    | 5  |
| Mjere centralne tendencije i varijabilnosti .....                   | 11 |
| Z-vrijednosti i normalna kriva .....                                | 19 |
| Standardna greška i intervali pouzdanosti aritmetičke sredine ..... | 23 |
| T-testovi .....   | 26 |
| Analiza varijanse (ANOVA) .....                                     | 34 |
| Korelacija .....  | 42 |
| Regresijska analiza .....   | 48 |
| Neki neparametrijski statistički postupci .....                     | 55 |
| Osnove teorije vjerovatnoće i kombinatorike .....                   | 61 |
| <i>Odgovori na pitanja i rješenja zadataka:</i>                     |    |
| Tabelarno i grafičko organizovanje i prikazivanje podataka .....    | 67 |
| Mjere centralne tendencije i varijabilnosti .....                   | 68 |
| Z-vrijednosti i normalna kriva .....                                | 70 |
| Standardna greška i intervali pouzdanosti aritmetičke sredine ..... | 71 |
| T-testovi .....   | 71 |
| Analiza varijanse (ANOVA) .....                                     | 73 |
| Korelacija .....  | 74 |
| Regresijska analiza .....   | 75 |
| Neki neparametrijski statistički postupci .....                     | 77 |
| Osnove teorije vjerovatnoće i kombinatorike .....                   | 78 |
| Literatura .....  | 80 |
| (Za)bilješka o autoru .....   | 81 |

## PREDGOVOR

Znam da niste upisali fakultet na kojem se izučava neka od društvenih nauka da biste ponovo kuburili sa numeričkim i geometrijskim konceptima koji su vas pratili tokom osnovnoškolskog i srednjoškolskog obrazovanja. Zapravo, željeli ste izbjeći matematiku, upisujući fakultet za koji ste pretpostavljali da je sterilno-siguran kada je riječ o ovoj oblasti.

Ipak, *dočekala vas je statistika!* No, nemojte očajavati! Statistika je jedna od grana primijenjene matematike, tj. sastoji se u aplikaciji znanja iz ove discipline na probleme i fenomene iz oblasti koja vas zanima (pedagogija, psihologija, sociologija, menadžment, kriminalistika/kriminologija, politologija...). Uostalom, statistika (uz metodologiju koja je provjereno suhoparnija oblast) doprinosi naučnom kredibilitetu društvenih nauka. Puko-raspravljачku, filozofsku tradiciju (koja oduvijek prati društvene nauke) u dobrom dijelu je moguće suzbiti logičkim, na rezultatima empirijskih studija zasnovanim interpretacijama i zaključcima.

Statistika *nalazi primjenu i u svakodnevnom životu*. Kritičko mišljenje zasnovano na poznavanju ovog područja može vam pomoći da procijenite koliko su pouzdani, objektivni i validni rezultati popisa, izbora i anketa koji su plasirani putem pisanih i usmenih medija. Dobro poznavanje statistike olakšava vam lično finansijsko planiranje (npr. prihoda i rashoda za tekući i svaki sljedeći mjesec). Ako se bavite poljoprivredom, uz pomoć statistike možete kvalitetnije i uspješnije planirati sadnju i ubirati plodove svoga rada. Statistika vam olakšava i planiranje nešto dužeg putovanja (procjena prosječnog zadržavanja na nekom graničnom prelazu, gužve na putevima, uzimanje u obzir udaljenosti različitih mjesta i najbržih načina stizanja do njih). Uz objektivno i dosljedno bilježenje podataka o vlastitoj težini (tjelesnoj masi) i dnevnoj fizičkoj aktivnosti (lagani, umjereni i teški rad), možete raditi na gubljenju kilograma. Prateći svoj napredak u sportu, izražen bročano ili grafički, imaćete bolji pregled svojih postignuća, te uspješnije planirati dalje fizičke aktivnosti. *Statistička pismenost* je, ustvari, oblik *sekundarne matematičke pismenosti*. Statističke kompetencije postaju korisne u procesu preživljavanja, snalaženja u okolini, svakodnevnog funkcionisanja i unapređenja vlastitog zdravstvenog i socioekonomskog statusa.

Možete se radovati i činjenici da se statistika može učiti s razumijevanjem, za razliku od mnogih oblasti društvenih nauka, gdje se podosta materijala treba jednostavno memorisati (tj. naučiti napamet ili "nabubati"). *Pozitivan stav* prema nekoj oblasti (ovdje statistici)

olakšava savladavanje gradiva i sticanje novih vještina u toj i sličnim oblastima. Nakon poželjnih postignuća, ovakav stav se učvršćuje, te dodatno potkrepljuje vaš napredak.

Nije mi namjera natjerati vas da volite statistiku onoliko koliko je ja volim, već da ne krenete s negativnim predrasudama u savladavanje njenih sadržaja. Statistika je vaš saveznik u učenju onoga što vas interesuje i samo vam može pomoći da posložite određena saznanja u smislenu mentalnu strukturu koja će vam zatrebati kada budete rješavali teorijske i praktične probleme u vlastitoj struci. Dakle, zadržite pozitivan mentalni stav, jer se bolje, brže i kvalitetnije uči kada smo u dobrom raspoloženju.

Srdačno,

*Selman Repišti*

*P. S.* Ako se pitate ko su lica koja vas gledaju sa tri crno-bijele slike naslovnice (korice)... U gornjem desnom uglu je Sir Ronald Aylmer Fisher (1890 – 1962), zaslužan za iscrpljujući postupak ANOVA-e koji vam zagorčava početak studentskog života. U donjem lijevom uglu je Karl Pearson (1857 – 1936), nešto milostiviji statističar koji je, između ostalog, obogatio vaše dosadašnje znanje korelacijom i hi-kvadrat testom. U donjem desnom uglu je Johann Carl Friedrich Gauss (1777 – 1855), mojoj i starijim generacijama poznat sa novačnice od 10 "dojč" maraka. Imamo mu zahvaliti na normalnoj krivoj, bez koje bi vam mnogi statistički koncepti prikazani u ovoj zbirci bili nejasni i nekako "strani".

## O NOTACIJI, FORMATU PITANJA/ZADATAKA I JEZIKU

Korištena notacija uključuje standardne skraćenice u statistici (npr. aritmetička sredina dobijena na uzorku će se označavati sa  $M$ , premda se može koristiti i simbol  $X$  sa crtom iznad, te oznaka  $AS$ ; za uzoračku standardnu devijaciju najčešće je korištena oznaka  $SD$ , mada se ovaj statistik može označiti i slovom  $s$  ...). Kao decimalni separator, korištena je tačka, što je i preporuka *Američke psihološke asocijacije* (APA, 2010). Dakle,  $t = 2.42$  isto je što i  $t = 2,42$ . Za koeficijente čija vrijednost ne može biti veća od jedan, niti manja od minus jedan, nula je izostavljena, takođe slijedeći preporuke APA-e. Stoga je  $r = .3$  isto što i  $r = 0.3$ , odnosno  $r = 0,3$ .

Najzastupljeniji format pitanja i zadataka je *MCQ* (eng. *multiple-choice questions*). Riječ je o pet ponuđenih odgovora u okviru svakog ovakvog pitanja/zadatka. Jedan od njih je tačan, a do njega se dolazi na neki od sljedeća tri načina: poznavanjem teorijske osnove odgovarajućeg statističkog koncepta, jednostavnim/kraćim izračunom (često bez potrebe za papirom i olovkom) ili složenijim/dužim izračunom (koji zahtijeva više vremena i mentalnog napora). Druge vrste pitanja/zadataka zahtijevaju da student dopuni ono što nedostaje u nekoj tablici ili u tekstu. Rješenja (sa više-manje dovoljnim objašnjenjima) nalaze se u drugom dijelu priručnika. Smatram da nije fer prema korisnicima neke zbirke da njen autor navodi rješenja za samo dio postavljenih problema (zadataka). Zato sam ponudio rješenja svih zadataka, uz već spomenuta objašnjenja. Uostalom, poznato je da *učimo i tokom rješavanja zadataka* (ispita, drugih provjera znanja i slično).

Statističku značajnost dobijenih koeficijenata, omjera i ostalih vrijednosti možete provjeravati uz pomoć tablica koje se nalaze na kraju većine statističkih udžbenika/priručnika. Popularni su i online (Internet) kalkulatori, gdje možete unijeti zadate (dobijene) vrijednosti, a kalkulator će vam izbaciti tražene vrijednosti (npr. vrijednost  $t$ -omjera, stepene slobode, značajnost, intervale pouzdanosti...). Svakako, mogu se koristiti i statistički softveri poput *SPSS*-a, *R*-a, *STATA*-e i *SAS*-a, te mogućnosti (funkcije) koje se nude unutar *MS Excela*.

Napominjem da za neke statističke procedure (naročito one parametrijske prirode) postoje veoma strogi preduslovi njihove primjene. O njima se možete najbolje informisati u kvalitetnim, iscrpnim udžbenicima iz statistike.

Ovo nije udžbenik statistike, pa su na početku svakog poglavlja prikazane najnužnije formule (obrasci). Dakle, statistiku ćete učiti iz dostupnih udžbenika na našem i stranim

jezicima, a svoje znanje možete provjeravati uz pomoć sadržaja ove zbirke. Za razliku od većine udžbenika, zbirke zadataka imaju svojstvo *interaktivnosti*, tj. njihovi korisnici (učenici i studenti) se potiču da razmisle o nekom problemu, izvedu računski postupak, donesu pravi zaključak, ispune prazan prostor ili zaokruže tačan odgovor na neko pitanje... U modernopremljenim udžbenicima statistike (zasnovanim na savremenom pedagoškom pristupu) objavljenim u svijetu takođe prisutan kvalitet interaktivnosti.

Zbirka je pisana jezikom koji je spoj bosanskog, crnogorskog, hrvatskog i srpskog (*BCHS*). Ovaj jezik je nekada (na vrijeme!) mogao biti nazvan svojim pravim imenom – jugoslavenskim jezikom.

Ako primijetite neku grešku koja se potkrala u rješenjima ili objašnjenjima, slobodno mi se javite putem e-maila [selman9r@yahoo.com](mailto:selman9r@yahoo.com).

Hvala i sretno,

*S. R.*



## TABELARNO I GRAFIČKO ORGANIZOVANJE I PRIKAZIVANJE PODATAKA

### Teorijska racionala

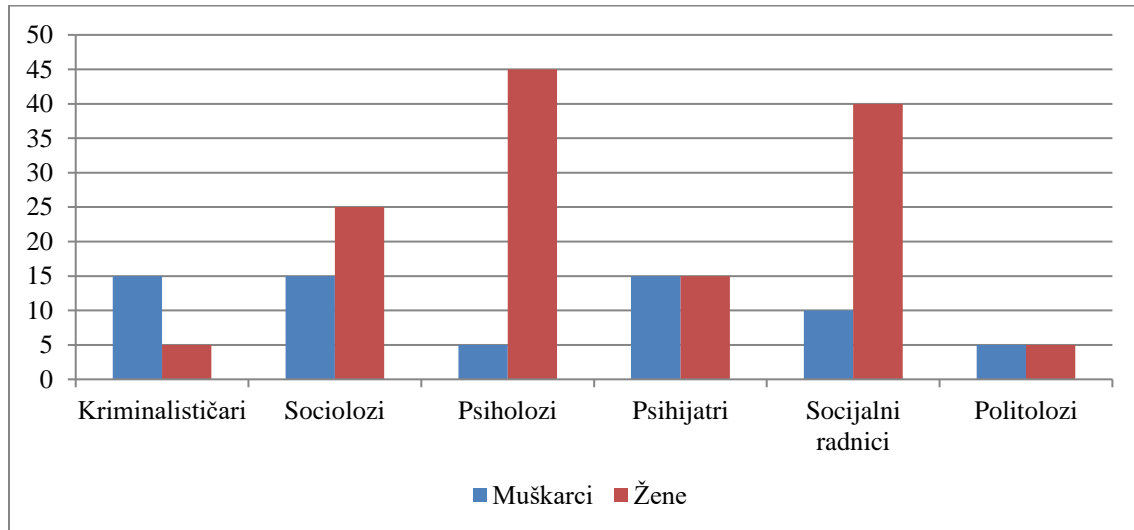
- ✓ Grafikoni (grafici, dijagrami) i tabele trebaju jasno i jednostavno ("čitljivo") prikazati podatke do kojih smo došli u nekom istraživanju.
- ✓ Svaka tabela treba imati svoje zaglavlje, a svaki grafikon legendu ili neki vid pojašnjenja korištenih simbola, linija različite boje i slično.
- ✓ *Kružnim* (pita) dijagramima prikazuje se broj ispitanika po kategorijama neke varijable (broj ispitanika u nekoj grupi, razredu ili kategoriji naziva se i frekvencija –  $f$ ) i/ili procenti (%).
- ✓ *Histogrami* su pogodni za prikazivanje kontinuiranih varijabli, npr. podijeljenih u jednake intervale (u svakom se nalazi određeni broj rezultata).
- ✓ *Stupčasti* dijagrami (eng. *bar charts*) se koriste za prikazivanje kategorijalnih varijabli (npr. broja muškaraca i žena u uzorku).
- ✓ Nekada u tabeli (na kraju reda i/ili kolone) imamo total (sumu), što je ukupan broj ispitanika ( $N$  je oznaka za veličinu uzorka, tj. ukupan broj ispitanika) ili ukupna procentualna vrijednost (u većini slučajeva je to 100%).
- ✓ Relativna frekvencija ( $rf$ ) je broj ispitanika u nekom poduzorku (grupi ili tzv. "razredu") podijeljen ukupnim brojem ispitanika. Zbir relativnih frekvencija je jedan.
- ✓ Kumulativna frekvencija ( $cf$ ) je ukupan broj ispitanika u datom razredu (grupi) i svim prethodnim razredima (grupama).
- ✓ Relativna kumulativna frekvencija ( $rcf$ ) jednaka je kumulativnoj frekvenciji podijeljenoj ukupnim brojem ispitanika.
- ✓ Sve vrste frekvencija se, naravno, mogu izraziti u procentima.

### Pitanja i zadaci

1. U tabelu koja slijedi upišite podatke koji nedostaju:

| Stručna sprema               | $f$  | %  |
|------------------------------|------|----|
| Osnovna škola                | 150  | ?  |
| Srednja stručna sprema (SSS) | 500  | 50 |
| Viša škola (VŠS)             | ?    | ?  |
| Visoka stručna sprema (VSS)  | 250  | ?  |
| Magisterij ili doktorat      | 30   | ?  |
| Total                        | 1000 | ?  |

2. Provedeno je istraživanje o polnoj strukturi pripadnika šest profesija. Na osnovu priloženog grafikona, popunite dio teksta koji slijedi (napomena: u svakoj kategoriji je cijeli broj ispitanika; ako procenti nisu cijeli brojevi, zaokružite vrijednosti na dvije decimale):



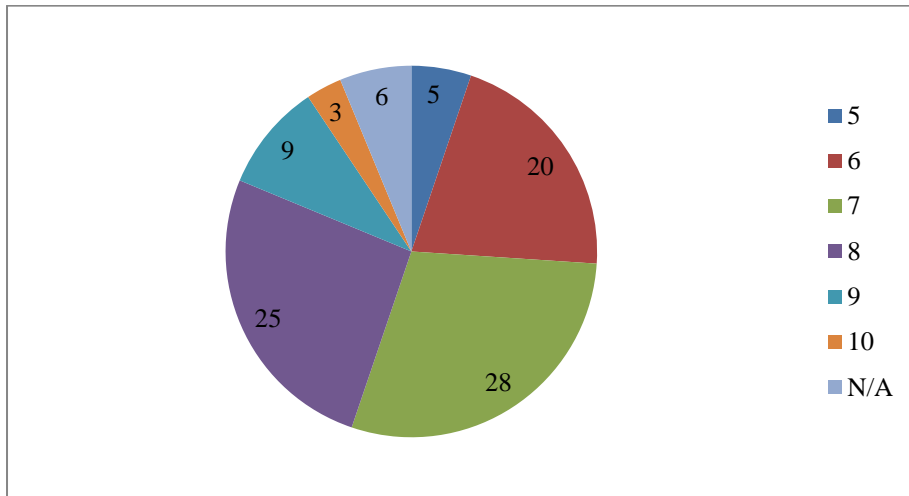
Uzorak je činilo  $N =$  \_\_\_\_\_ ispitanika, od toga \_\_\_\_\_ muškaraca (što je \_\_\_\_\_ % uzorka) i \_\_\_\_\_ žena (tj. \_\_\_\_\_ % uzorka). U uzorku je bilo najmanje politologa, i to njih \_\_\_\_\_ (odnosno, \_\_\_\_\_ % ukupnog uzorka). Pored toga, imali smo jednak broj psihologa i socijalnih radnika (po \_\_\_\_\_, odnosno po \_\_\_\_\_%). Među kriminalističarima, bilo je \_\_\_\_\_ muškaraca. Isti broj muškaraca bio je među \_\_\_\_\_ i \_\_\_\_\_. Podjednak broj muškaraca i žena bio je među \_\_\_\_\_. Najviše žena bilo je među psiholozima, i to njih \_\_\_\_\_, što je \_\_\_\_\_ % ukupnog broja žena u našem uzorku.

3. *Pretpostavimo da je u nekom uzorku od 50 ljudi: 12 neoženjenih/neudatih osoba, duplo više oženjenih/udatih osoba, petina uzorka su razvedeni i ostatak su udovci/udovice. Dopunite sljedeću rečenicu:*

Dakle, u uzorku je \_\_\_\_\_ oženjenih/udatih osoba (što je \_\_\_\_\_ % ukupnog broja ispitanika), \_\_\_\_\_ razvedenih osoba (tj. \_\_\_\_\_ % uzorka) i \_\_\_\_\_ udovaca/udovica (\_\_\_\_\_ % uzorka).

4. *Prikazan je pita-dijagram sa ocjenama studenata prve godine studija ekonomije jednog državnog fakulteta (ocjene su od 5 do 10) na jednom od uvodnih kolegija. Broj studenata koji su dobili neku od ocjena je unutar kruga. Svi studenti su bili na predavanju na kojem ih je profesor pitao da navedu svoje ocjene iz datog kolegija.*

*N/A znači da ne postoji podatak za određeni broj studenata. Dopuniti tekst koji je interpretacija ovog dijagrama (odgovore koji nisu cijeli brojevi zaokružiti na jednu decimalu):*



Na prvoj godini ekonomije na datom fakultetu ima ukupno \_\_\_\_\_ studenata. Najveći broj njih dobio je ocjenu \_\_\_\_\_. Procenat studenata sa ovom ocjenom u ukupnom uzorku bio je \_\_\_\_\_ %. Najmanje studenata (tj. njih \_\_\_\_\_) dobilo je ocjenu \_\_\_\_\_. Devetku je dobilo \_\_\_\_\_ studenata, što je \_\_\_\_\_ % ukupnog uzorka. Devetku i desetku je dobilo ukupno \_\_\_\_\_ studenata, tj. \_\_\_\_\_ % ukupnog uzorka. Procenat studenata koji su dali podatak o svojoj ocjeni je \_\_\_\_\_ %.

5. *Direktor je popisao broj učenika koji pohađaju osnovnu školu. Podatke je organizovao po razredima. Dopuniti tabelu (brojeve koji nisu cijeli zaokružiti na dvije decimale):*

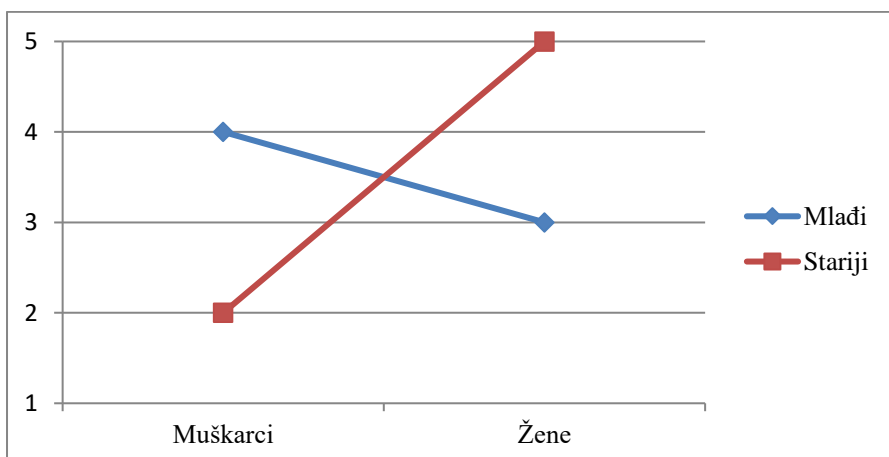
| Razred  | <i>f</i> | <i>f</i> (%) | <i>rf</i> | <i>cf</i> | <i>cf</i> (%) | <i>rcf</i> |
|---------|----------|--------------|-----------|-----------|---------------|------------|
| 1. – 3. | 85       | ?            | ?         | 85        | ?             | ?          |
| 4. – 6. | 60       | ?            | ?         | ?         | ?             | ?          |
| 7. – 9. | ?        | ?            | ?         | ?         | ?             | ?          |
| Total   | 180      | ≈100         | ≈1        |           |               |            |

6. *Iz popunjene tabele u zadatku br. 5 može se zaključiti sljedeće (dopuniti): neki od prvih šest razreda pohađa \_\_\_\_\_ djece, što je \_\_\_\_\_% ukupnog broja osnovnoškolaca; neki od prva tri razreda pohađa \_\_\_\_\_% osnovnoškolaca; a relativna frekvencija broja učenika u sedmom do devetog razreda je \_\_\_\_\_.*

7. *Kružni dijagram pogodan je za prikazivanje:*

a) frekvencija, tj. broja ispitanika po određenim grupama (kategorijama) i/ili pripadajućih procenata

- b) prosječnih vrijednosti varijabli
  - c) frekvencija i prosječnih vrijednosti
  - d) procenata i prosječnih vrijednosti
  - e) razlika u prosječnim vrijednostima dvije ili više analiziranih grupa
8. *Relativna frekvencija je omjer:*
- a) broja ispitanika u nekom razredu (kategoriji) i broja ispitanika u sljedećem razredu (kategoriji)
  - b) procenta ispitanika u jednom i procenta ispitanika u sljedećem razredu (ili kategoriji)
  - c) broja ispitanika u nekom razredu (kategoriji) i ukupnog broja ispitanika
  - d) procenta ispitanika u nekom razredu (kategoriji) i ukupnog broja ispitanika
  - e) broja ispitanika u nekoj kategoriji (razredu) i broja 100
9. *Zbir svih relativnih frekvencija nekog skupa podataka (zanemarujući nezatno nepodudaranje usljed zaokruživanja) je:*
- a)  $N$  (tj. jednaka ukupnom broju ispitanika)
  - b) -1
  - c) 0
  - d) 1
  - e) 100
10. *Ispitivano je da li na zadovoljstvo roditeljstvom utiče starost i pol ispitanika. Ispitanici su po starosti podijeljeni na mlađe i starije. Zadovoljstvo roditeljstvom je procjenjivano na skali od 1 do 5 (gdje je 1 značilo "jako nezadovoljan", 2 – "nezadovoljan", 3 – "niti zadovoljan niti nezadovoljan", 4 – "zadovoljan" i 5 – "izuzetno zadovoljan"). Rezultati su prikazani na grafikonu ispod (prosječne vrijednosti su, za potrebe lakše čitljivosti grafikona, zaokružene na cijele brojeve):*



*Dopuniti interpretaciju grafikona:*

Zadovoljstvo roditeljstvom u poduzorku mlađih muškaraca iznosilo je \_\_\_\_\_, dok je zadovoljstvo roditeljstvom u poduzorku starijih žena bilo \_\_\_\_\_. Procjena ispitanika u grupi \_\_\_\_\_ iznosila je dva, a u grupi \_\_\_\_\_ tri.

*11. Dostupni su vam podaci o broju kriminalnih djela različite vrste u četiri vremenska perioda na nekom području (vrijednosti koje nisu cijeli brojevi zaokružiti na dvije decimale):*

| Kriminalna djela | 2013. | 2014. | 2015. | 2016. |
|------------------|-------|-------|-------|-------|
| Provalništvo     | 164   | 153   | 128   | 141   |
| Podmićivanje     | 23    | 30    | 38    | 42    |
| Prostitucija     | 12    | 10    | 14    | 15    |
| Silovanje        | 6     | 8     | 7     | 5     |

*Dopuniti dio kriminalističkog izvještaja koji se odnosi na navedenu tabelu:*

Najviše slučajeva podmićivanja zabilježeno je \_\_\_\_\_ godine. Najmanji broj silovanja bio je tokom \_\_\_\_\_ godine i iznosio je \_\_\_\_\_. Broj provalništava u 2014. godini iznosi \_\_\_\_\_ % od ukupnog broja provalništava u sve četiri godine. U odnosu na 2014. godinu, broj djela iz kategorije "provalništvo" 2015. godine bio je manji za \_\_\_\_\_ %. Ukupan broj silovanja u prikazanom četvorogodišnjem periodu iznosio je \_\_\_\_\_. Najveći broj prikazanih kriminalnih djela (bez obzira na njihovu vrstu) desio se u \_\_\_\_\_ godini.

*12. U tabeli koja slijedi dati su neki demografski podaci iz 2016. godine koji se odnose na stanovništvo jednog grada:*

| <i>P</i> | <i>N</i> | <i>M</i> | <i>I</i> | <i>E</i> | <i>n</i> | <i>m</i> | <i>VI</i> | <i>S</i> | <i>SPS (1000)</i> |
|----------|----------|----------|----------|----------|----------|----------|-----------|----------|-------------------|
| 80026    | 241      | 163      | 140      | 1200     | ?        | ?        | ?         | ?        | ?                 |

*Dopuniti tabelu, rukovodeći se sljedećim informacijama: 1) P je populacija u 2015. godini, 2) n je opšta stopa nataliteta (broj živorođene djece na 1000 stanovnika), a N je ukupan broj živorođene djece 2016. godine; 3) m je opšta stopa mortaliteta (broj umrlih na 1000 stanovnika), a M je ukupan broj umrlih u 2016. godini; 4) I je broj imigranata u datoj godini; 5) E je broj osoba koje su emigrirale u datoj godini; 6) VI je oznaka koju sam koristio da označim tzv. "vitalni indeks" (broj živorođenih se dijeli brojem umrlih u datoj godini, nakon čega se dobijeni rezultat množi sa 100), 7) S je oznaka za migracioni saldo ( $S = \text{broj imigranata} - \text{broj emigranata}$ ) i 8) SPS (1000) je oznaka koju sam koristio da označim stopu prirasta stanovništva na 1000*

stanovnika ( $SPS(1000) = 1000 \cdot (N - M + I - E) / P$ ). Vrijednosti koje nisu cijeli brojevi zaokružite na dvije decimale.

13. Na osnovu uvrštenih vrijednosti iz zadatka 12 i dopunjene tabele, kakav je zaključak?

U odnosu na 2015. godinu, populacija pomenutog mjesta:

- a) se povećala za otprilike 12 stanovnika.
- b) se povećala za otprilike 982 stanovnika.
- c) ostala je ista.
- d) se smanjila za otprilike 12 stanovnika.
- e) se smanjila za otprilike 982 stanovnika.

## MJERE CENTRALNE TENDENCIJE I VARIJABILNOSTI

### Teorijska racionala

- ✓ Mjere centralne tendencije su: mod ili dominantna vrijednost ( $D$ ), medijana (centralna vrijednost,  $Mdn$  ili  $C$ ) i aritmetička sredina ( $M$ ).
- ✓ Mod je najčešća vrijednost u nekom nizu podataka.
- ✓ Medijana je vrijednost koja pada u sredinu, kada se podaci (vrijednosti, brojevi) poredaju po veličini.
- ✓ Aritmetička sredina je prosječan rezultat, odnosno zbir svih rezultata podijeljen brojem ( $N$ ) tih rezultata.
- ✓ Najčešće mjere varijabilnosti su: broj kategorija (grupa, vrsta...), "obični" raspon (minimalni rezultat oduzet od maksimalnog rezultata nekog niza podataka) interkvartilni raspon ( $IQR$ ) i standardna devijacija ( $SD$ ).
- ✓ Broj kategorija (možemo ga označiti sa  $k$ ) se koristi uz dominantnu vrijednost i ova dva statistička indikatora vezana su za *nominalne* (*kategorijalne*) varijable (pol, vrste nekog proizvoda, tipovi kazne, odsjeci na nekom fakultetu...)
- ✓ Za *ordinalne* skale mjerenja i pripadajuće varijable koriste se medijana i interkvartilni raspon, premda se mogu koristiti i nešto "primitivniji" pokazatelji, poput moda i broja kategorija.
- ✓ Interkvartilni raspon je razlika između rezultata koji pada u 75. i onog koji pada u 25. centil. Kada ga podijelimo sa dva, dobijamo *poluinterkvartilni* (*semiinterkvartilni*) raspon.
- ✓ Za *intervalne* skale mjerenja i varijable koje im odgovaraju računaju se aritmetička sredina i standardna devijacija (mada se mogu koristiti i ostale mjere koje smo već spomenuli).
- ✓ Standardna devijacija populacije ( $\sigma$ ) je korijen prosjeka kvadriranih odstupanja rezultata od aritmetičke sredine:

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

- ✓ Kada nazivniku (imeniocu) razlomka ispod korijena oduzmemo jedinicu (tj. umjesto  $N$  napišemo  $N - 1$ ), dobijamo uzoračku standardnu devijaciju ( $SD$  ili  $s$ ).
- ✓ Inače, *parametri* je naziv za statističke vrijednosti *populacije*. Populacijska aritmetička sredina se obilježava  $\mu$ , a ona izračunata na uzorku  $M$ . Populacijska standardna

devijacija je  $\sigma$ , a ako je izračunata na uzorku, onda je  $SD$  ili  $s$ . Statistički pokazatelji (vrijednosti) izračunate na *uzorku* nazivaju se *statistici*.

- ✓ "Postrizižena" aritmetička sredina (eng. *trimmed mean*) je  $M$  izračunata tako da npr. izbacimo 20% najnižih i 20% najviših vrijednosti. Ostaje 60% centralnih vrijednosti (rezultata) čiju aritmetičku sredinu računamo. U našem primjeru, riječ je o 20%-no "postriziženoj" aritmetičkoj sredini.
- ✓ Još jedna od mjera varijabiliteta jeste *koeficijent varijacije* ( $CV$ ). Ovaj koeficijent omogućava da uporedimo varijabilitet varijabli izraženih u različitim mjernim jedinicama (Boslaugh i Watters, 2008). Obrazac za njegovo izračunavanje je:

$$CV = \frac{s}{M} \times 100$$

### ***Pitanja i zadaci***

1. *Pisac je utvrdio da u svojim kratkim pričama najmanje koristi imenice, potom glagole, pa priloge i zamjenice. Isto tako, utvrdio je da najviše koristi pridjeve. Aki imenice označimo brojem 1, glagole brojem 2, priloge 3, zamjenice 4 i pridjeve 5, koja je dominantna vrijednost (mod) ovih podataka?*
  - a)  $D = 1$
  - b)  $D = 2$
  - c)  $D = 3$
  - d)  $D = 4$
  - e)  $D = 5$
2. *Jedan glumac volio se hvaliti svojim prethodnim vezama. Zapravo, o njima je vodio detaljnu evidenciju. Zabilježio je da je bio sa deset crnki, jedanaest plavuša, osam brineta, dvije ridokose djevojke i tri djevojke koje su bile ofarbane zelenom, ljubičastom i rozom bojom. Mod boje kose njegovih bivših djevojaka je:*
  - a)  $D = \text{"crna"}$
  - b)  $D = \text{"plava"}$
  - c)  $D = \text{"braon"}$
  - d)  $D = \text{"narandžasta/crvena"}$
  - e)  $D = \text{"zeleno" ili "roza" ili "ljubičasto"}$
3. *Psijatar je danas imao 20 pacijenata. Od toga, četvero ih je bilo s dijagnozom opsesivno-kompulzivnog poremećaja, po petoro s dijagnozom sumanutog poremećaja i poremećaja ličnosti, troje s dijagnozom poremećaja vezanih za ovisnost o*



*psihoaktivnim supstancama, dvoje sa poremećajem spavanja i jedno sa poremećajem ishrane. U ovim podacima, imamo:*

- a) jednu dominantnu vrijednost, i to je dijagnoza opsesivno-kompluzivnog poremećaja
  - b) jednu dominantnu vrijednost, i to je poremećaj hranjenja
  - c) dvije dominantne vrijednosti – sumanuti poremećaj i poremećaj ličnosti
  - d) dvije dominantne vrijednosti – dvije osobe sa poremećajem spavanja
  - e) tri dominantne vrijednosti – tri osobe koje su ovisne o psihoaktivnim supstancama
4. *Sociolog je podijelio naseljena mjesta na: metropole, velike gradove, gradove srednje veličine, manje gradove, prigradska mjesta i sela. Metropolama je pripisan rang 1, velikim gradovima 2, gradovima srednje veličine 3... do sela, kojima je pripisan rang 6. Koja je medijana (centralna vrijednost) ovih podataka?*
- a) 2.5
  - b) 3
  - c) 3.5
  - d) 4
  - e) 4.5
5. *Kada raspodjela skupa podataka izraženih na intervalnoj skali mjerenja statistički značajno odstupa od normalne distribucije, ispravno je računati:*
- a) samo mod
  - b) samo medijanu
  - c) samo aritmetičku sredinu
  - d) mod i medijanu, ali ne i aritmetičku sredinu
  - e) medijanu i aritmetičku sredinu, ali ne i mod
6. *Šestoro dobrovoljaca učestvovalo je u jednom testu zapamćivanja riječi sa pet slogova. Njihovi rezultati bili su: 2, 3, 14, 5, 7 i 5. Medijana i aritmetička sredina ovog niza podataka su:*
- a)  $C = 5, M = 6$
  - b)  $C = 5, M = 8$
  - c)  $C = 6, M = 5$
  - d)  $C = 6, M = 5$
  - e)  $C = 8, M = 6$
7. *Osam popisivača danas je popisalo sljedeći broj osoba: 4, 16, 16, 15, 17, 29, 17 i 15. Vidimo da je prvi popisivač značajno podbacio, dok je šesti pokazao zavidnu*

učinkovitost. Za potrebe ovog zadatka, želimo eliminisati rezultate ova dva popisivača, tako što ćemo izračunati "postrizenu" aritmetičku sredinu. Koliko iznosi ova aritmetička sredina i za koliko % je smanjen broj podataka?

- a)  $M = 12$ , za 12.5%
- b)  $M = 16.0$ , za 25%
- c)  $M = 16.125$ , za 12.5%
- d)  $M = 16.5$ , za 25%
- e)  $M = 21.5$ , za 12.5%

8. Mod, medijana i aritmetička sredina niza podataka u zagradi (0, 50, 80, 100, 100, 120, 150 i 200) su:

- a)  $D = C = M = 100$
- b)  $D = M = 100, C = 90$
- c)  $D = M = 100, C = 110$
- d)  $D = C = 90, M = 100$
- e)  $D = C = 110, M = 100$

9. Po materijalnom statusu, stanovništvo jedne regije smo podijelili u sljedećih pet kategorija: A. loše materijalne prilike (mjesečni prihodi manji od 200 eura), B. ispodprosječne materijalne prilike (mjesečni prihodi od 200 do 350 eura), C. prosječne materijalne prilike (mjesečni prihodi veći od 350, a manji od 650 eura), D. iznadprosječne materijalne prilike (mjesečni prihodi u iznosu od 650 do 1300 eura) i E. izvrsne materijalne prilike (mjesečni prihodi iznad 1300 eura). U prvoj grupi bilo je 2000, u drugoj 8000, u trećoj 10000, u četvrtoj 750, a u petoj 80 porodica. Dominantna vrijednost i varijabilnost varijable "kategorija" su:

- a)  $D = "A"$  i varijabilnost je 5
- b)  $D = "B"$  i varijabilnost je 1100
- c)  $D = "C"$  i varijabilnost je 5
- d)  $D = "D"$  i varijabilnost je 1100
- e)  $D = "E"$  i varijabilnost je 9920

10. Raspon niza podataka u zagradi (1, 1, 7, 7, 3, 5, 6, 6, 4) je:

- a) 3
- b) 4.44
- c) 5
- d) 6
- e) 7

11. *Aritmetička sredina rezultata na testu inteligencije petnaest studenata filozofije je 125, a trideset studenata psihologije 115. Zajednička aritmetička sredina je:*
- a) 118.33
  - b) 120
  - c) 121.67
  - d) 142.5
  - e) 240
12. *Na jednom uzorku dobijeni su sljedeći rezultati: 2, 4, 4, 3, 6 i 5. Standardna devijacija ovih podataka iznosi:*
- a)  $SD = 1.29$
  - b)  $SD = 1.41$
  - c)  $SD = 1.67$
  - d)  $SD = 2.00$
  - e)  $SD = 4.00$
13. *Ako su podaci iz zadatka br. 12 dobijeni na nekoj populaciji, njihova standardna devijacija iznosi:*
- a)  $SD = 1.29$
  - b)  $SD = 1.41$
  - c)  $SD = 1.67$
  - d)  $SD = 2.00$
  - e)  $SD = 4.00$
14. *Svi studenti su na testu znanja iz antropologije postigli maksimalan broj bodova (=100 poena). Aritmetička sredina i standardna devijacija njihovih rezultata su:*
- a)  $M = SD = 100$
  - b)  $M = 0, SD = 100$
  - c)  $M = 100, SD = 0$
  - d)  $M = SD = 0$
  - e)  $M = 100, SD = 10$
15. *Na trećoj godini Odsjeka za sociologiju je samo jedanaest studenata. Iz kolegija "Savremene sociološke teorije" postigli su sljedeći broj poena: 30, 45, 62, 100, 70, 40, 75, 65, 93, 97 i 82. Rezultati koji padaju u 25, 50 i 75. centil su (navedenim redoslijedom):*
- a) 62, 40 i 93
  - b) 45, 70 i 93

- c) 30, 70 i 100
- d) 40, 62 i 100
- e) 45, 75 i 97

16. Medijana i interkvartilni raspon podataka iz zadatka br. 15 su:

- a)  $C = 70, IQR = 48$
- b)  $C = 72.5, IQR = 48$
- c)  $C = 70, IQR = 25$
- d)  $C = 72.5, IQR = 25$
- e)  $C = 70, IQR = 23$

17. Između 25. i 75. centila nalazi se:

- a) 25% rezultata
- b) 33.3% rezultata
- c) 50% rezultata
- d) 66.7% rezultata
- e) 75% rezultata

18. Ako je medijana niza podataka  $Mdn = 138$ , aritmetička sredina  $M = 145$ , a standardna devijacija  $SD = 28$ , koji rezultat pada u 50. centil?

- a)  $X = 28$
- b)  $X = 100$
- c)  $X = 141$
- d)  $X = 138$
- e)  $X = 145$

19. Za pozitivno asimetričnu distribuciju vrijedi:

- a)  $D = C = M$
- b)  $D > C, C = M$
- c)  $M > C, C = D$
- d)  $D > C > M$
- e)  $D < C < M$

20. Za negativno asimetričnu distribuciju karakteristično je sljedeće:

- a) otklon rezultata prema višim vrijednostima
- b) otklon rezultata prema nižim vrijednostima
- c)  $D = C = M$
- d) otklon rezultata prema višim vrijednostima i  $D = C = M$
- e) otklon rezultata prema nižim vrijednostima i  $D = C = M$

21. Ako je standardna devijacija nekog niza podataka  $SD = 25$ , njegova varijansa je:
- a) 5
  - b) 25
  - c) 50
  - d) 250
  - e) 625
22. S povećanjem broja ispitanika, standardna devijacija:
- a) opada
  - b) ostaje ista
  - c) raste
  - d) postaje bliža aritmetičkoj sredini
  - e) postaje bliža varijansi
23. Jedan istraživač dobio je sljedeće podatke: 12, 16, 20, 22, 24, 23, 21, 22, 26, 30, 123. Šta je najkorektnije da on uradi kako bi mjera centralne tendencije bila validan pokazatelj grupisanja podataka?
- a) Izračunati aritmetičku sredinu na svim podacima, jer je ona najpreciznija mjera centralne tendencije.
  - b) Izbaciti posljednji podatak, pa izračunati aritmetičku sredinu ostalih vrijednosti ili izračunati medijanu svih podataka.
  - c) Izračunati dominantnu vrijednost podataka.
  - d) Izračunati "postrizenu" aritmetičku sredinu, izbacujući rezultate  $X = 12$  i  $X = 123$ .
  - e) Izbaciti posljednji podatak, potom izračunati aritmetičku sredinu preostalih vrijednosti na način da se njihov zbir podijeli sa 11.
24. Aritmetička sredina jedne skale procjene kojom se mjeri savjesnost kao crta ličnosti je  $M = 32$ . Ovu skalu procjene ispunilo je deset studenata pedagogije. Istraživač je izgubio rezultat petog studenta, pa sada raspolaže sljedećim podacima: 25, 27, 28, 29,  $X_5$ , 34, 35, 37, 38 i 40. Koji je skor na pomenutoj skali postigao peti student?
- a)  $X_5 = 27$
  - b)  $X_5 = 29$
  - c)  $X_5 = 30$
  - d)  $X_5 = 31$
  - e)  $X_5 = 32$
25. Koji je raspon podataka u zagradi (6, -3, 8, 14, -10, 2.5, 3.756, -4.21)?
- a) 0

- b) 4
- c) 8
- d) 11.5
- e) 24

26. *Koji je raspon vrijednosti neke konstante?*

- a)  $-\infty$
- b) -1
- c) 0
- d) 1
- e)  $+\infty$

27. *Koliki je koeficijent varijacije ako je poznato da aritmetička sredina jedne varijable iznosi  $M = 23$ , a standardna devijacija  $SD = 6.5$ ?*

- a) 3.54%
- b) 8.25%
- c) 16.50%
- d) 28.26%
- e) 56.52%

## Z-VRIJEDNOSTI I NORMALNA KRIVA

### Teorijska racionala

- ✓ Normalna (zvonolika, Gaussova) kriva je unimodalna (ima jednu dominantnu vrijednost) i simetrična kriva (u odnosu na vertikalu koja prolazi njenom sredinom).
- ✓ Standardna (standardizovana, normirana) normalna kriva ima aritmetičku sredinu 0, a standardnu devijaciju 1.
- ✓ Z-vrijednosti nastaju standardizovanjem "običnih" rezultata ispitanika (z nekog rezultata koji smo obilježili sa  $i$  jednak je njegovom odstupanju od aritmetičke sredine podataka, podijeljenom standardnom devijacijom tih istih podataka kojima pripada):

$$z_i = \frac{X_i - M}{SD}$$

- ✓ Za idealnu normalnu krivu (bez odstupanja) vrijedi:

| <i>U rasponu:</i> | <i>Nalazi se:</i> |
|-------------------|-------------------|
| $M \pm 1 SD$      | 68.26% rezultata  |
| $M \pm 2 SD$      | 95.44% rezultata  |
| $M \pm 3 SD$      | 99.73% rezultata  |

### Pitanja i zadaci

1. Na osnovu podataka dobijenih na reprezentativnom uzorku, obavljena je standardizacija jednog testa intelektualnih sposobnosti u našoj državi. Utvrđeno je da aritmetička sredina iznosi  $M = 52$ , a standardna devijacija  $SD = 12$ . Ako rezultate izrazimo u z-vrijednostima, prosječan skor na pomenutom testu iznosiće:
  - a) 0
  - b) 1
  - c) 12
  - d) 40
  - e) 52
2. Standardna devijacija standardizovane normalne krive (standardne raspodjele) je:
  - a) uvijek nula
  - b) uvijek jedan

- c) nekada nula, a nekada jedan
  - d) uvijek jednaka njenoj aritmetičkoj sredini
  - e) nekada jednaka njenoj aritmetičkoj sredini, a nekada maksimalnoj vrijednosti
3. *Aritmetička sredina rezultata na nekoj skali procjene je  $M = 22$ , a standardna devijacija  $SD = 5$ . Utvrđeno je da se ovi rezultati distribuiraju po normalnoj krivoj. Koji je rezultat postigao ispitanik M. B., ako je njegov standardizovani skor  $z = -2$ ?*
- a) 11
  - b) 12
  - c) 17
  - d) 22
  - e) 24
4. *Ispitanik I. N. ostvario je skor  $X = 43$  na testu za koji je utvrđeno da je  $M = 50$ , a  $SD = 7$ . Koliko iznosi njegov rezultat, izražen kao  $z$ -vrijednost?*
- a) -2
  - b) -1
  - c) 0
  - d) 1
  - e) 2
5. *U rasponu od  $z_1 = -3$  do  $z_2 = 3$ , nalazi se:*
- a) manje od 50% rezultata
  - b) između 50 i 65% rezultata
  - c) između 65 i 80% rezultata
  - d) između 80 i 95% rezultata
  - e) više od 95% rezultata
6. *Aritmetička sredina nekog niza rezultata koji se distribuiraju u skladu sa normalnom krivom je  $M = 77$ . Koliko % rezultata je nižih od aritmetičke sredine?*
- a) 25%
  - b) 33.3%
  - c) 50%
  - d) 66.7%
  - e) 75%
7. *Raspoložete sljedećim podacima:  $M = 10$ ,  $SD = 2$ ,  $N = 200$ . Distribucija rezultata na varijabli za koju su dobijeni ovi podaci ne razlikuje se od zvonolike krive. Koje su odgovarajuće  $z$ -vrijednosti rezultata ispitanika A ( $X_A = 4$ ) i ispitanika B ( $X_B = 16$ )?*



- a) oba su jednaka nuli
  - b)  $z_A = -1, z_B = 1$
  - c)  $z_A = -2, z_B = 2$
  - d)  $z_A = -3, z_B = 3$
  - e) oba su jednaka jedinici
8. *Uzimajući u obzir rezultate distribuirane po Gaussovoj (zvonolikoj) krivoj, unutar raspona  $M \pm 1SD$  nalazi se:*
- a) manje od 10% rezultata
  - b) između 37.5 i 50% rezultata
  - c) više od 55, a manje od 65% rezultata
  - d) više od 68, a manje od 70% rezultata
  - e) više od 90% rezultata
9. *Rezultati dva ispitanika na istoj varijabli, izraženi u z-vrijednostima ( $z_1 = 1.26$  i  $z_2 = -1.26$ ), su:*
- a) ispodprosječni rezultati
  - b) jednako udaljeni od aritmetičke sredine
  - c) izuzetno niski rezultati
  - d) izuzetno visoki rezultati
  - e) jednaki
10. *Za rezultat izražen u obliku z-vrijednosti ( $z = 4$ ), može se reći da je:*
- a) znatno ispod prosjeka
  - b) prosječan
  - c) znatno iznad prosjeka
  - d) teorijski nemoguć
  - e) jednak aritmetičkoj sredini standardizovanih rezultata na ovoj varijabli
11. *Između rezultata kojem odgovara z-vrijednost od -1 i rezultata čija je z-vrijednost +2 nalazi se:*
- a) 13.59% rezultata
  - b) 20.54% rezultata
  - c) 52.28% rezultata
  - d) 68.26% rezultata
  - e) 81.85% rezultata

12. Aritmetička sredina rezultata distribuiranih u skladu sa normalnom raspodjelom je  $M = 15$ , a standardna devijacija  $SD = 4.5$ . Ispitanik je postigao rezultat  $X = 24$ . Koliko % ispitanika je postiglo rezultat veći od ovog ispitanika?

- a) 0.14%
- b) 2.28%
- c) 4.56%
- d) 7.94%
- e) 15.87%

## STANDARDNA GREŠKA I INTERVALI POUZDANOSTI ARITMETIČKE SREDINE

### Teorijska racionala

- ✓ Standardna greška *aritmetičke sredine* računa se po obrascu:

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

- ✓ Zapravo, kako nam nije poznata standardna devijacija populacije, koja bi trebala stajati u ovom obrascu (dakle,  $\sigma$  umjesto  $s$ ), koristimo standardnu devijaciju uzorka (Petz, 2007).
- ✓ Intervali pouzdanosti aritmetičke sredine u principu se poklapaju sa intervalima koje smo naveli za procenat rezultata pod normalnom krivom. Međutim, najčešće nam treba 95%-ni i 99%-ni interval pouzdanosti:

| <i>Raspon</i>   | <i>Interval</i> |
|-----------------|-----------------|
| $M \pm 1.96s_M$ | 95%-ni          |
| $M \pm 2.58s_M$ | 99 %-ni         |

### Pitanja i zadaci

1. Na uzorku od 115 učenika ispitali smo njihovu motivaciju za savladavanje školskog gradiva. Aritmetička sredina njihovih procjena bila je  $M = 3.05$  (teorijski prosjek skale procjene koju smo koristili bio je 3). Standardna devijacija iznosila je  $SD = 0.6$ . Koliko iznosi standardna greška aritmetičke sredine?
  - a) 0.005
  - b) 0.007
  - c) 0.056
  - d) 0.072
  - e) 0.164
2. Između kojih rezultata se sa vjerovatnoćom od 68.26% može tvrditi da se nalazi prava aritmetička sredina iz zadatka 1? Drugim riječima, koje su granice ovog intervala?
  - a) 3.045 – 3.055
  - b) 3.043 – 3.057
  - c) 2.994 – 3.106
  - d) 2.978 – 3.122

- e) 2.886 – 3.214
3. U kojem intervalu se sa 95.44%-nom vjerovatnoćom može tvrditi da se nalazi prava aritmetička sredina iz zadatka 1?
- a) 3.04 – 3.06  
b) 3.036 – 3.064  
c) 2.938 – 3.162  
d) 2.906 – 3.194  
e) 2.722 – 3.378
4. Ako uzmemo duplo veći uzorak ispitanika i ako standardna devijacija ostane ista, kako se i da li se uopšte promijeni standardna greška aritmetičke sredine? "Nova" standardna greška je:
- a) duplo manja od "stare"  
b)  $\sqrt{2}$  puta manja od "stare"  
c) duplo veća od "stare"  
d)  $\sqrt{2}$  puta veća od "stare"  
e) jednaka "staroj"
5. Ako veličina uzorka ostaje ista, a "nova" standardna devijacija je tri puta veća od postojeće, kako se promijeni standardna greška aritmetičke sredine?
- a) smanji se tri puta  
b) smanji se devet puta  
c) poveća se tri puta  
d) poveća se devet puta  
e) ostaje ista
6. Standardna greška aritmetičke sredine je .25, a dobijena je na uzorku  $N = 64$ . Kolika je standardna devijacija rezultata ispitanika?
- a) 2  
b) 4  
c) 8  
d) 12  
e) 16
7. Koji je 95%-ni interval pouzdanosti aritmetičke sredine koja iznosi  $M = 9$ , a za koju su vezani podaci iz zadatka br. 6?
- a) 5 – 13  
b) 7 – 11

- c)  $8 - 10$
  - d)  $8.5 - 9.5$
  - e)  $8.51 - 9.49$
8. *Koji je 99%-ni interval pouzdanosti aritmetičke sredine koja iznosi  $M = 9$ , a za koju su vezani podaci iz zadatka br. 6?*
- a)  $3 - 15$
  - b)  $6 - 12$
  - c)  $7.5 - 10.5$
  - d)  $8.25 - 9.75$
  - e)  $8.355 - 9.645$

## T-TESTOVI

### Teorijska racionala

- ✓ T-test za jedan uzorak (eng. *one-sample t-test*) koristi se kada nas zanima da li neki uzorak dolazi iz populacije sa određenom (ili pretpostavljenom, očekivanom) aritmetičkom sredinom:

$$t = \frac{M - \mu}{\sqrt{\frac{s^2}{N}}}$$

- ✓ T-test za nezavisne uzorke (eng. *independent-samples t-test*) koristi se kada želimo uporediti dvije grupe po određenoj varijabli. Obrazac za velike uzorke (tj. one čija je veličina  $N > 30$ , Mordkoff, 2016) je:

$$t = \frac{M_1 - M_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

- ✓ Slijedi obrazac za male uzorke, s tim što se prije njegove primjene treba provjeriti jednakost varijansi (kvadriranih standardnih devijacija) rezultata dvije komparirane grupe. Dakle, za male nezavisne uzorke  $t$ -statistik se računa ovako:

$$t = \frac{M_1 - M_2}{\sqrt{\frac{s_1^2(n_1-1) + s_2^2(n_2-1)}{n_1 + n_2 - 2} \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

- ✓ T-test za uparene (zavisne) uzorke (eng. *paired-samples t-test*) koristi se kada namjeravamo uporediti jednu grupu ispitanika u dvije vremenske tačke (npr. prije i poslije uvođenja neke eksperimentalne intervencije). Takođe se može računati za velike i male uzorke. Za velike uzorke, obrazac je:

$$t = \frac{M_1 - M_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2} - 2r_{12} \frac{s_1 s_2}{\sqrt{n_1 n_2}}}}$$

- ✓ Za male, uparene uzorke, obrazac se razlikuje, budući da se temelji na tzv. *metodi diferencije*:

$$t = \frac{M_D}{\frac{s_D}{\sqrt{N}}}$$

- ✓  $M_D$  je aritmetička sredina razlika između rezultata istih ispitanika u dvije tačke (situacije, uslova) mjerenja.  $S_D$  je standardna devijacija ovih razlika, a  $N$  je broj ispitanika (dakle, ne ukupan broj rezultata, kojih je duplo više od broja ispitanika, jer svaki ispitanik ima rezultat u prvom i u drugom mjerenju).
- ✓ Standardna (po)greška razlike aritmetičkih sredina ( $S_{M_1-M_2}$ ) u okviru računanja  $t$ -omjera je, ustvari, navedena u nazivniku (imeniocu)  $t$ -omjera.

### ***Pitanja i zadaci***

1. *Ako je standardna greška razlike aritmetičkih sredina jedan,  $t$ -omjer za nezavisne uzorke je:*
  - a) uvijek nula
  - b) uvijek  $M_1 - M_2$
  - c) uvijek  $M_1 + M_2$
  - d) uvijek jednak jednoj od aritmetičkih sredina
  - e) nemoguće izračunati
2. *Prikupili smo uzorak od 230 osoba. Želimo uporediti mlade ( $N = 100$ ) i starije ( $N = 130$ ) osobe iz ovog uzorka po stepenu njihovog zadovoljstva životom. Koji ćemo statistički postupak koristiti za ovu svrhu?*
  - a)  $t$ -test za male zavisne uzorke
  - b)  $t$ -test za male nezavisne uzorke
  - c)  $t$ -test za velike zavisne uzorke
  - d)  $t$ -test za velike nezavisne uzorke
  - e)  $t$ -test za jedan uzorak
3. *Koliko iznose stepeni slobode koji se koriste u statističkom postupku iz zadatka br. 2?*
  - a)  $df = 115$
  - b)  $df = 130$
  - c)  $df = 228$
  - d)  $df = 229$
  - e)  $df = 230$
4. *Muškarci ( $N = 120$ ) i žene ( $N = 150$ ) iz jednog uzorka upoređeni su po nivou opšteg samopoštovanja. Aritmetička sredina rezultata muškaraca na ovoj varijabli bila je  $M = 40$  ( $SD = 6$ ), a žena  $M = 32$  ( $SD = 7$ ). Da li je razlika u njihovim prosječnim rezultatima statistički značajna? Odaberi opciju koja odgovara dobijenom rezultatu:*
  - a)  $t(270) = 10.11, p < .01$

- b)  $t(270) = 8, p < .01$
  - c)  $t(268) = 8, p > .05$
  - d)  $t(270) = 8, p > .05$
  - e)  $t(268) = 10.11, p < .01$
5. *Zaključak u zadatku broj 4 je:*
- a) Nema statistički značajnih polnih (rodnih) razlika u opštem samopoštovanju.
  - b) Dobijena je razlika na pomenutoj varijabli u korist muškaraca, ali ona nije dovoljno velika da bi bila statistički značajna.
  - c) Dobijena je razlika u korist žena, ali nedovoljno velika da bi bila proglašena statistički značajnom.
  - d) Dobijena je statistički značajna razlika u opštem samopoštovanju, i to u korist muškaraca.
  - e) Dobijena je statistički značajna razlika u opštem samopoštovanju, i to u korist žena.
6. *Grupa od 108 osoba prisustvovala je treningu koji je imao za cilj unapređenje njihove emocionalne inteligencije. Mjerenje ove varijable izvršeno je prije i po okončanju treninga. Koji statistički postupak bi istraživač koristio da ispita ima li statistički značajnih pomaka u njihovoj emocionalnoj inteligenciji?*
- a)  $t$ -test za male zavisne uzorke
  - b)  $t$ -test za male nezavisne uzorke
  - c)  $t$ -test za velike zavisne uzorke
  - d)  $t$ -test za velike nezavisne uzorke
  - e)  $t$ -test za jedan uzorak
7. *Koliko iznose stepeni slobode u zadatku br. 6?*
- a)  $df = 54$
  - b)  $df = 106$
  - c)  $df = 107$
  - d)  $df = 108$
  - e)  $df = 216$
8. *Od grupe studenata žurnalistike ( $N = 82$ ) je na početku prve godine zatraženo da na petostepenoj skali procijene koliko vjeruju elektronskim medijima (1 – "uopšte ne vjerujem", 2 – "ne vjerujem", 3 – "niti vjerujem niti ne vjerujem", 4 – "vjerujem", 5 – "u potpunosti vjerujem"). Prosjek njihovih odgovora bio je  $M = 3.78$  (uz  $SD = 1.01$ ). Nakon završetka treće godine, ponovo su procijenili povjerenje u sadržaje plasirane*



putem elektronskih medija. Sada su rezultati bili sljedeći:  $M = 2.31$  i  $SD = 0.78$ . Povezanost (korelacija) između rezultata dobijenih u ove dvije procjene bila je  $r = .52$ . Da li je došlo do statistički značajne promjene u procjenama vjerovanja elektronskim medijima? Odaberite neki od ponuđenih statističkih zapisa, koji bi bio u skladu sa vašim zaključkom:

- a)  $t(81) = 14.80, p < .01$
- b)  $t(80) = 1.47, p < .01$
- c)  $t(82) = 1.47, p > .05$
- d)  $t(80) = 1.47, p > .05$
- e)  $t(81) = 14.80, p > .05$

9. Rezultate iz zadatka br. 8 interpretiramo na sljedeći način:

- a) Studenti žurnalistike statistički značajno manje vjeruju elektronskim medijima u drugoj tački mjerenja.
- b) Studenti žurnalistike podjednako vjeruju elektronskim medijima na početku studiranja i na kraju treće godine.
- c) Studenti žurnalistike u prvoj tački mjerenja statistički značajno manje vjeruju medijima nego u drugoj tački mjerenja.
- d) Studenti žurnalistike na kraju treće godine nešto manje vjeruju medijima, ali nedovoljno, da bi razlika bila statistički značajna.
- e) Studenti žurnalistike na početku prve godine vjeruju nešto više medijima, međutim, razlika između prve i druge tačke mjerenja nije statistički značajna.

10. S povećanjem standardne greške razlike aritmetičkih sredina:

- a)  $t$ -omjer raste
- b)  $t$ -omjer ostaje isti
- c)  $t$ -omjer opada
- d) povećava se vjerovatnoća da razlike budu statistički značajne
- e) mjerenje obje aritmetičke sredine je preciznije

11. Ako želimo ispitati da li kvantni fizičari imaju statistički značajno veći koeficijent inteligencije od populacijskog prosjeka, korist ćemo:

- a)  $t$ -test za male zavisne uzorke
- b)  $t$ -test za male nezavisne uzorke
- c)  $t$ -test za velike zavisne uzorke
- d)  $t$ -test za velike nezavisne uzorke
- e)  $t$ -test za jedan uzorak

12. Nakon što ste odabrali statistički postupak u pitanju br. 11, dobijate sljedeće informacije: populacijski prosjek inteligencije je 100, uzorak fizičara je  $N = 49$ , aritmetička sredina njihovog nivoa (koeficijenta) inteligencije je 119, a standardna devijacija 10. Da li se njihov prosječni IQ zaista statistički značajno razlikuje od populacijskog prosjeka?

- a) Da ( $t(48) = 13.3, p < .01$ )
- b) Ne ( $t(48) = 13.3, p < .01$ )
- c) Da ( $t(48) = 9.31, p < .01$ )
- d) Ne ( $t(48) = 0.08, p > .05$ )
- e) Da ( $t(48) = 0.08, p < .05$ )

13. Devet ispitanika sa mjesečnim primanjima iznad 700 eura i 11 ispitanika sa primanjima ispod 700 eura procijenili su kvalitet života na skali od 1 do 10 (od 1 – "jako je teško i neprijatno živjeti i sadržaji koji me zanimaju nisu mi lako dostupni", do 10 – "veoma je lako i ugodno živjeti, a gotovo svi sadržaji koji me interesuju su mi dostupni"). Njihovi rezultati prikazani su u tabeli:

| Prva grupa<br>(primanja > 700 eura) | Druga grupa<br>(primanja < 700 eura) |
|-------------------------------------|--------------------------------------|
| 6                                   | 9                                    |
| 7                                   | 6                                    |
| 5                                   | 5                                    |
| 8                                   | 4                                    |
| 7                                   | 5                                    |
| 9                                   | 6                                    |
| 10                                  | 7                                    |
| 7                                   | 8                                    |
| 8                                   | 5                                    |
|                                     | 3                                    |
|                                     | 4                                    |

Koji postupak primjenjujemo da ispitamo ima li statistički značajnih razlika u kvalitetu života ove dvije grupe?

- a)  $t$ -test za jedan uzorak
- b)  $t$ -test za male nezavisne uzorke
- c)  $t$ -test za velike nezavisne uzorke
- d)  $t$ -test za male uparene (zavisne) uzorke
- e)  $t$ -test za velike uparene (zavisne) uzorke

14. Koliko iznose stepeni slobode u zadatku br. 13?

- a) 9

- b) 10
- c) 11
- d) 18
- e) 19

15. Koliko iznosi razlika između aritmetičkih sredina dvije grupe iz zadatka br. 13?

- a) 1.80
- b) 3.60
- c) 5.64
- d) 6.54
- e) 7.44

16. Kakav je zaključak, na osnovu rezultata provedenog t-testa za podatke iz zadatka br. 13?

- a) Ispitanici koji zarađuju više od 700 eura su statistički značajno zadovoljniji kvalitetom svoga života.
- b) Ispitanici koji zarađuju manje od 700 eura su statistički značajno zadovoljniji kvalitetom svoga života.
- c) Nema statistički značajnih razlika u zadovoljstvu kvalitetom života između ove dvije grupe.
- d) Zadovoljstvo kvalitetom života ispitanika koji zarađuju manje od 700 eura ne razlikuje se statistički značajno od populacijskog prosjeka.
- e) Zadovoljstvo kvalitetom života ispitanika koji zarađuju više od 700 eura je statistički značajno veće nego kada su isti ispitanici zarađivali manje od 700 eura.

17. Pretpostavite da smo ispitali stavove 12 mladih iz naše države prema jednoj od susjednih zemalja (na skali od 1 do 10, gdje 1 znači "izrazito negativan stav", a 10 "veoma pozitivan stav"). Nakon toga, ispitanici su učestvovali u dvosatnoj radionici koja je imala za cilj informisati mlade o toj susjednoj državi i omogućiti im direktan kontakt sa njenim državljanima. Poslije radionice, ponovo smo ispitali stavove ovog uzorka naših državljana. Rezultati su bili sljedeći:

| Ispitanik | Prije radionice | Nakon radionice |
|-----------|-----------------|-----------------|
| G. A.     | 2               | 4               |
| S. Z.     | 5               | 6               |
| N. B.     | 3               | 4               |
| V. I.     | 5               | 5               |
| A. V.     | 6               | 7               |
| Ž. S.     | 7               | 9               |
| M. M.     | 4               | 6               |

|       |   |    |
|-------|---|----|
| P. A. | 8 | 8  |
| Z. B. | 6 | 7  |
| T. R. | 3 | 5  |
| S. P. | 5 | 7  |
| Š. S. | 9 | 10 |

---

*Koju statističku proceduru koristite kako biste provjerili je li došlo do statistički značajne promjene u stavovima prema narodu susjedne države?*

- a) *t*-test za jedan uzorak
- b) *t*-test za male nezavisne uzorke
- c) *t*-test za velike nezavisne uzorke
- d) *t*-test za male zavisne uzorke
- e) *t*-test za velike zavisne uzorke

*18. Koliko je stepeni slobode za t-omjer iz zadatka br. 17?*

- a) 11
- b) 12
- c) 22
- d) 23
- e) 24

*19. Koliki je t-omjer izračunat na osnovu podataka iz zadatka br. 17?*

- a) Između 1 i 2
- b) Između 2 i 3
- c) Između 3 i 4
- d) Između 5 i 6
- e) Između 6 i 7

*20. Kakav je zaključak, nakon što se provjeri značajnost t-testa dobijenog iz podataka zadatka br. 17?*

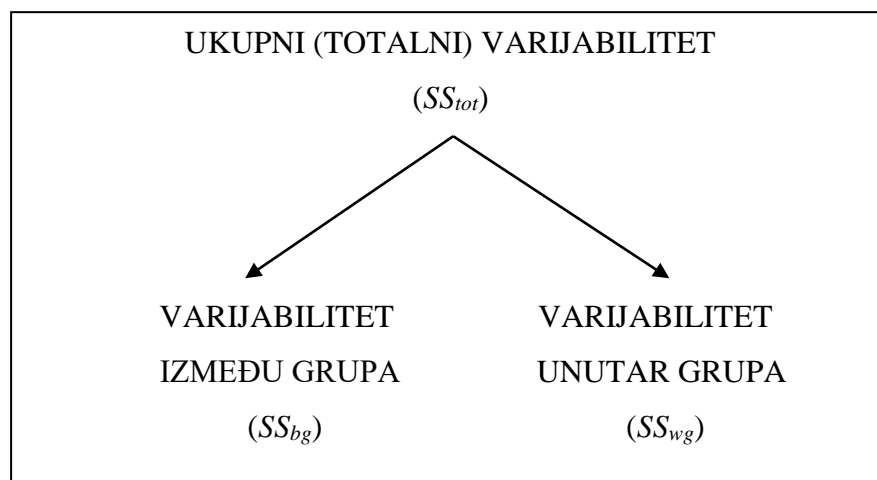
- a) Dvije grupe ispitanika se ne razlikuju statistički značajno po stavovima prema stanovništvu susjedne države.
- b) Dvije grupe ispitanika se razlikuju statistički značajno po stavovima prema stanovništvu susjedne države.
- c) Ispitanici u drugoj tački mjerenja imaju statistički značajno pozitivnije stavove prema stanovništvu susjedne države.
- d) Stavovi ispitanika prema stanovništvu susjedne države u dvije tačke mjerenja se međusobno ne razlikuju statistički značajno.

- e) Prosječna vrijednost (pozitivnosti ili negativnosti) stavova ispitivane grupe prema stanovništvu susjedne države se ne razlikuje statistički značajno od populacijske aritmetičke sredine ove varijable.

## ANALIZA VARIJANSE (ANOVA)

### Teorijska racionala

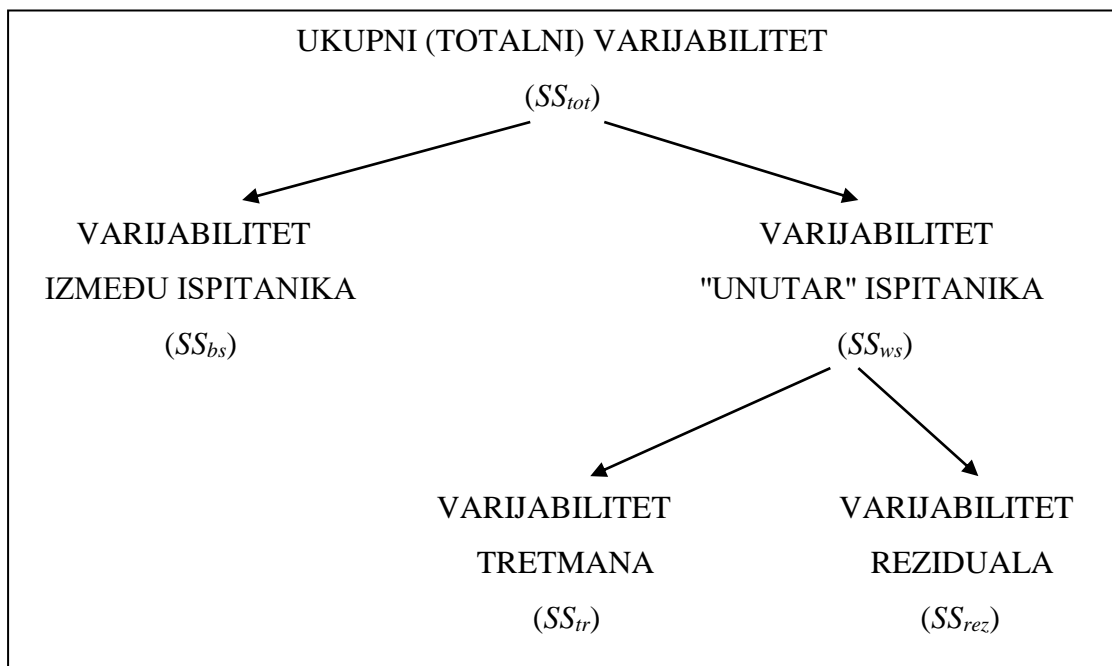
- ✓ Analiza varijanse (ANOVA) temelji se, kako joj i sam naziv kaže, na *varijansi* i povezanim konceptima.
- ✓ Analiza varijanse koristi se kada imamo dvije ili više grupa koje upoređujemo, dva ili više eksperimentalnih uslova (i tačaka mjerenja), te kada proučavamo djelovanje jedne ili više nezavisnih varijabli na jednu ili više zavisnih varijabli.
- ✓ Ako imamo jednu nezavisnu i jednu zavisnu varijablu, govorimo o *jednosmjernoj univarijatnoj analizi varijanse*.
- ✓ Ako imamo dvije ili više nezavisnih varijabli, riječ je o *faktorijalnoj (složenoj) ANOVA-i*. U njoj se nezavisne varijable zovu *faktori*, a njihove vrijednosti (kategorije) *nivoi*.
- ✓ Ako, npr. imamo dvije nezavisne i jednu zavisnu varijablu, to je *dvosmjerna, odnosno dvofaktorijalna univarijatna ANOVA*.
- ✓ Ako imamo tri nezavisne i dvije zavisne varijable, govorimo o *trofaktorijalnoj bivarijatnoj analizi varijanse*.
- ✓ ANOVA sa više zavisnih varijabli je *multivarijatna analiza varijanse (MANOVA)*.
- ✓ Varijansa se može rastaviti na nekoliko načina i u nekoliko kategorija, u zavisnosti od vrste analize varijanse.
- ✓ *Dekompozicija* (rastavljanje, dijeljenje) varijabiliteta u jednosmjernoj univarijatnoj ANOVA-i:



- ✓ Dekompozicija varijabiliteta u dvosmjernoj ANOVA-i je ista kao u prethodnom slučaju, s tim što se varijabilitet između grupa dijeli na varijabilitet svakog od faktora i

njihove interakcije. Naprimjer, ako imamo dva faktora ( $A$  i  $B$ ),  $SS_{bg} = SS_A + SS_B + SS_{AxB}$ . Za tri faktora, situacija se dramatično usložnjava:  $SS_{bg} = SS_A + SS_B + SS_C + SS_{AxB} + SS_{AxC} + SS_{BxC} + SS_{AxBxC}$  (Dancey i Reidy, 2011).

- ✓ U dvosmjernoj ANOVA-i imamo osam mogućih scenarija: 1) samo je glavni efekat prvog faktora statistički značajan, 2) samo je glavni efekat drugog faktora statistički značajan, 3) samo je interakcija statistički značajna, 4) oba glavna efekta su značajna (interakcija nije), 5) interakcija je značajna (glavni efekti nisu), 6) značajni su jedino glavni efekat prvog faktora i interakcija, 7) značajni su jedino glavni efekat drugog faktora i interakcija i 8) ni glavni efekti ni efekat interakcije nisu statistički značajni (Goodwin, 2010).
- ✓ Dekompozicija varijabiliteta u ANOVA-i za ponovljena mjerenja (pandanu  $t$ -testa za zavisne uzorke, s tim što ANOVA za ponovljena mjerenja nije ograničena samo na dva, već na i više vremenskih tačaka u kojima se vrši mjerenje, procjena i slično):



- ✓ Obrazac za računanje prosječnih kvadrata (primjer za varijansu između grupa -  $bg$ ):

$$MS_{bg} = \frac{SS_{bg}}{df_{bg}}$$

- ✓ Obrazac za računanje  $F$ -omjera (primjer: za faktor  $B$  u dvosmjernoj ANOVA-i):

$$F_A = \frac{MS_A}{MS_{wg}}$$

✓ Obrazac za računanje  $F$ -omjera u ANOVA-i za ponovljena mjerenja:

$$F = \frac{MS_{tr}}{MS_{rez}}$$

### **Pitanja i zadaci**

1. Želimo ispitati u kojoj mjeri radni status (zaposlen, nezaposlen i penzionisan), pol i mjesto stanovanja (grad, prigradsko naselje i selo) utiču na nečije zadovoljstvo podrškom koju primaju od svoje okoline. Riječ je sljedećem nacrtu (dizajnu) ANOVA-e:
  - a) 1 x 1 x 1
  - b) 3 x 1 x 3
  - c) 3 x 3 x 3
  - d) 3 x 2 x 3
  - e) 1 x 2 x 1
2. Ako ispitujemo uticaj nivoa obrazovanja, dobne grupe kojoj ispitanici pripadaju i materijalnog statusa na mentalno zdravlje, korist ćemo:
  - a) jednosmjernu univarijatnu ANOVA-u
  - b) dvosmjernu univarijatnu ANOVA-u
  - c) trosmjernu univarijatnu ANOVA-u
  - d) jednofaktorijalnu multivarijatnu ANOVA-u
  - e) dvofaktorijalnu bivarijatnu ANOVA-u
3. Ako je suma kvadrata za faktor A  $SS_A = 25$ , suma kvadrata unutar grupa 15, stepeni slobode za faktor A  $df_A = 1$  i  $df$  unutar grupa 67, koliko iznosi prosječni kvadrat faktora A?
  - a) 0.22
  - b) 2.68
  - c) 4.47
  - d) 25
  - e) 67
4. Koliko nivoa ima faktor A iz zadatka br. 3?
  - a) 1
  - b) 2
  - c) 3



- d) 4
- e) 5

5. Koliki je  $F$ -omjer za faktor  $A$  iz zadatka br. 3?

- a) 5.6
- b) 15
- c) 25
- d) 67
- e) 111.67

6. Statistički zapis  $F_{P \times Q}(2, 207) = 2.70, p > .05$  znači:

- a) glavni efekat faktora  $P$  je statistički značajan
- b) glavni efekat faktora  $Q$  je statistički značajan
- c) efekat interakcije faktora  $P$  i  $Q$  je statistički značajan
- d) i glavni efekti i efekat interakcije faktora  $P$  i  $Q$  su statistički značajni
- e) efekat interakcije faktora  $P$  i  $Q$  nije statistički značajan

7. Ovo je statistički zapis dobijen na osnovu neke jednosmjerne univarijatne ANOVA-e:

$F(4, 138) = 3.81, p < .05$ . On se interpretira na sljedeći način:

- a) Između bar dvije od četiri grupe ispitanika postoji statistički značajna razlika u prosječnim rezultatima na zavisnoj varijabli.
- b) Između bar dvije od pet grupa ispitanika postoji statistički značajna razlika u prosječnim rezultatima na zavisnoj varijabli.
- c) Nema statistički značajne razlike između ove četiri grupe ispitanika u prosječnim rezultatima na zavisnoj varijabli.
- d) Nema statistički značajne razlike između ovih pet grupa ispitanika u prosječnim rezultatima na zavisnoj varijabli.
- e) Razlike u aritmetičkim sredinama rezultata četiri ili pet grupa ispitanika nije moguće utvrditi jednosmjernom ANOVA-om.

8. Suma kvadrata između grupa u nekoj dvofaktorijalnoj analizi varijanse je 165. Suma kvadrata prvog faktora je 80, a drugog 42. Kolika je suma kvadrata njihove interakcije?

- a) 38
- b) 43
- c) 122
- d) 203
- e) 287

9. Ako su prosječni kvadrat tretmana i prosječni kvadrat reziduala jednaki, vrijedi sljedeće:

- a) efekat tretmana je statistički neznačajan
- b) efekat tretmana je statistički značajan
- c)  $F$ -omjer je statistički značajan
- d)  $F$ -omjer je nula
- e)  $F$ -omjer je nemoguće izračunati

10. Ako je u jednosmjernoj ANOVA-i omjer prosječnog kvadrata između grupa i prosječnog kvadrata unutar grupa manji od jedan, onda:

- a) nema potrebe da provjeravamo značajnost  $F$ -omjera, jer je statistički neznačajan
- b) nema potrebe da provjeravamo značajnost  $F$ -omjera, jer je statistički značajan
- c) trebamo provjeriti značajnost  $F$ -omjera, jer iz postavke zadatka ne možemo izvesti nikakav zaključak
- d) ne smijemo računati vrijednost  $F$ -omjera
- e) vrijednost  $F$ -omjera postaje nepouzdan statistički pokazatelj

11. Data je tablica u kojoj je prikazan nivo stresa kod socijalnih radnika, vaspitačica i policajaca. Viši rezultat znači veću percepciju stresa na radnom mjestu:

| Socijalni radnici | Vaspitačice | Policajci |
|-------------------|-------------|-----------|
| 25                | 20          | 28        |
| 35                | 28          | 32        |
| 30                | 26          | 31        |
| 18                | 23          | 25        |
| 42                | 21          | 30        |
| 32                | 19          | 36        |
| 31                | 16          | 34        |
| 23                | 12          | 26        |
| 22                | 13          | 25        |
| 27                | 15          | 29        |

Koju vrstu ANOVA-e treba primijeniti, kako bi se ispitalo ima li statistički značajnih razlika između ovih grupa u percipiranom nivou stresa na poslu?

- a) jednosmjernu univarijatnu ANOVA-u
- b) dvosmjernu univarijatnu ANOVA-u
- c) trosmjernu univarijatnu ANOVA-u
- d) složenu univarijatnu ANOVA-u
- e) ANOVA-u za ponovljena mjerenja

12. Referirajući se na podatke iz zadatka br. 11, koliko iznose stepeni slobode između grupa?

- a) 0
- b) 1
- c) 2
- d) 3
- e) 4

13. *Koliko iznose stepeni slobode totala za podatke iz zadatka br. 11?*

- a) 27
- b) 28
- c) 29
- d) 30
- e) 31

14. *Koliki je F-omjer, izračunat za podatke iz zadatka br. 11?*

- a) 2.6
- b) 5.2
- c) 7.8
- d) 10.4
- e) 13

15. *Kakav je rezultat ANOVA-e koju ste upravo proveli na osnovu podataka iz zadatka br. 11?*

- a) Sve grupe se statistički značajno razlikuju po nivou percipiranog stresa na poslu.
- b) Socijalni radnici i policajci percipiraju statistički značajno viši nivo stresa na poslu nego vaspitačice.
- c) Socijalni radnici i vaspitačice percipiraju statistički značajno viši nivo stresa na poslu od policajaca.
- d) Policajci i vaspitačice percipiraju statistički značajno viši nivo stresa na poslu od socijalnih radnika.
- e) Nema statistički značajnih razlika u percipiranom nivou stresa na poslu između ove tri grupe profesionalaca.

16. *U jednom eksperimentalnom istraživanju, rezultati 37 ispitanika na zavisnoj varijabli mjereni su četiri puta (prije prve intervencije, između prve i druge intervencije, između druge i treće intervencije i nakon treće intervencije). Dopuniti tabelu ANOVA-e koja je rezultat ovog istraživanja (u posljednjoj koloni treba navesti  $p > .05$  ili  $p < .05$ ):*

| Izvor variranja   | <i>SS</i> | <i>df</i> | <i>MS</i> | <i>F</i> | <i>p</i> |
|-------------------|-----------|-----------|-----------|----------|----------|
| Između ispitanika | 63        | ?         |           |          |          |
| Tretman           | 89        | ?         | ?         | ?        | ?        |
| Rezidual          | 309       | ?         | ?         |          |          |
| Unutar ispitanika | ?         | ?         |           |          |          |
| Total             | ?         | ?         |           |          |          |

17. Kakav je zaključak koji se temelji na rezultatima dobijenim u zadatku br. 16?

- Nema statistički značajne razlike u rezultatima dobijenim u četiri tačke mjerenja.
- Dobijena je bar jedna statistički značajna razlika između rezultata u četiri tačke mjerenja.
- Dobijene su bar dvije statistički značajne razlike između rezultata u četiri tačke mjerenja.
- Dobijene su bar tri statistički značajne razlike između rezultata u četiri tačke mjerenja.
- Sve razlike u rezultatima između četiri tačke mjerenja su statistički značajne.

18. Eksperimentator je ispitivao djelovanje dvije varijable (sa po tri nivoa) na jednu zavisnu varijablu. O kojem nacrtu ANOVA-e, odnosno eksperimentalnom dizajnu je riječ?

- 2 x 1
- 2 x 3
- 3 x 3
- 2 x 3 x 1
- 3 x 3 x 1

19. Na osnovu eksperimenta iz zadatka br. 18, dobijena je sljedeća tabela ANOVA-e.

Dopunite je (u kolonu "p" upisujte  $p > .05$  ili  $p < .05$ ):

| Izvor variranja | <i>SS</i> | <i>df</i> | <i>MS</i> | <i>F</i> | <i>p</i> |
|-----------------|-----------|-----------|-----------|----------|----------|
| Unutar grupa    | 185       | 54        | ?         |          |          |
| NZV1            | 16        | ?         | ?         | ?        | ?        |
| NZV2            | 18        | ?         | ?         | ?        | ?        |
| NZV1 x NZV2     | ?         | ?         | ?         | ?        | ?        |
| Između grupa    | 42        | ?         | ?         |          |          |
| Total           | ?         | ?         |           |          |          |

20. Kakav je zaključak nakon što ste ispunili tabelu iz zadatka br. 19?

- Glavni efekti su statistički značajni, ali interakcija nije.
- Interakcija je statistički značajna, ali glavni efekti nisu.
- Ni glavni efekti ni interakcija nisu statistički značajni.

- d) I glavni efekti i interakcija su statistički značni.
- e) Glavni efekat varijable NZV2 je statistički značajan, ali ostalo nije.

21. *Koliki je uzorak ispitanika koji su učestvovali u eksperimentu iz zadataka 18 i 19?*

- a) 62
- b) 63
- c) 67
- d) 70
- e) 71

## KORELACIJA

### *Teorijska racionala*

- ✓ Međusobna povezanost dvije ili više pojava (karakteristika, varijabli, fenomena) naziva se korelacijom ili njihovim sukladnim variranjem.
- ✓ *Pozitivna* korelacija: s porastom vrijednosti jedne varijable rastu i vrijednosti druge; s opadanjem vrijednosti jedne varijable opadaju i vrijednosti druge.
- ✓ *Negativna* korelacija: sa smanjenjem rezultata na jednoj varijabli povećavaju se vrijednosti u drugoj; sa povećanjem rezultata na prvoj opadaju rezultati na drugoj varijabli.
- ✓ Povezanost varijable sa samom sobom je maksimalna jer je riječ o istoj varijabli.
- ✓ Povezanost između varijable i konstante je nula, jer konstanta ne varira.
- ✓ *Pearsonov* (čita se: "Pirsonov") koeficijent označava se  $r$  i po svojoj prirodi i namjeni je koeficijent *linearne* korelacije:

$$r_{xy} = \frac{\sum[(x - M_x)(y - M_y)]}{\sqrt{\sum(x - M_x)^2 \sum(y - M_y)^2}}$$

- ✓ Korelacija je standardizovana kovarijansa, tj. jednaka je omjeru kovarijanse dvije varijable ( $x$  i  $y$ ) i umnoška njihovih standardnih devijacija:

$$r_{xy} = \frac{cov_{xy}}{s_x s_y}$$

- ✓ Korelacija upućuje samo na povezanost između nekih pojava, svojstava i slično. Ona *nije isto što i uzročno-posljedična veza*, ali je jedan od (pred)uslova za ovakav vid veze između varijabli.
- ✓ Korelacija se po svojoj snazi (tj. veličini koeficijenta korelacije) dijeli na nisku (slabu), osrednju (umjerenu) i visoku (jaku, čvrstu povezanost). Autori nisu jedinstveni oko ove podjele, te oko raspona koeficijenata korelacije u svakoj od kategorija formiranih po veličini koeficijenta korelacije.
- ✓ Koeficijenti korelacije  $r = -.5$  i  $r = .5$  su iste visine, samo drugačijeg predznaka (u prvom slučaju, riječ je o negativnoj, a u drugom o pozitivnoj povezanosti).
- ✓ Raspon koeficijenta korelacije:

$$-1 \leq r \leq 1$$

- ✓ Za potrebe ovog priručnika (rješavanje postavljenih zadataka, odnosno odgovaranje na pitanja), prikazujem vam podjelu visine povezanosti (u apsolutnim vrijednostima, dakle, nije bitno je li predznak + ili -), koju navodi Evans (1996):

| <i>Stepen povezanosti</i> | <i>Raspon koeficijenta korelacije (r)</i> |
|---------------------------|---|
| Veoma slaba               | 0 - .19                                   |
| Niska (slaba)             | .20 - .39                                 |
| Umjerena                  | .40 - .59                                 |
| Jaka (visoka)             | .60 - .79                                 |
| Veoma jaka (veoma visoka) | .80 - 1                                   |

### ***Pitanja i zadaci***

- Ako je korelacija između depresivnosti i anksioznosti (tjeskobe, strepnje, zabrinutosti)  $r = .61$ , povezanost između anksioznosti i depresivnosti je:*
  - 0.61
  - 0.39
  - 0.39
  - 0.61
  - 1
- Povezanost varijable sa samom sobom je:*
  - 1
  - 0
  - 1
  - jednaka aritmetičkoj sredini te varijable
  - jednaka standardnoj devijaciji te varijable
- Povezanost neke varijable i neke konstante je:*
  - 1
  - 0.5
  - 0
  - 0.5
  - 1
- Grafički prikaz korelacije između jedne varijable i jedne konstante je:*
  - prava linija
  - kriva linija
  - trougao

- d) kvadrat
  - e) pravougaonik
5. Što je veća raspršenost tačaka na skater-dijagramu, to je korelacija:
- a) niža
  - b) viša
  - c) negativnija
  - d) pozitivnija
  - e) bliža jedinici
6. Ako je povezanost između varijabli  $D$  i  $W$ , izražena Pearsonovim koeficijentom korelacije,  $r = .79$ , to bi značilo:
- a) s povećanjem vrijednosti varijable  $D$  ne mijenjaju se vrijednosti varijable  $W$
  - b) sa povećanjem rezultata na varijabli  $W$  ne mijenjaju se rezultati varijable  $D$
  - c) sa smanjenjem vrijednosti varijable  $W$  smanjuju se vrijednosti varijable  $D$
  - d) varijabla  $D$  je uzrok varijable  $W$
  - e) varijabla  $D$  je posljedica varijable  $W$
7. Ako smo dobili da je korelacija između dvije varijable  $r = -.45$ , ova povezanost je:
- a) negativna i niska
  - b) negativna i umjerena
  - c) negativna i visoka
  - d) minimalna moguća
  - e) maksimalna moguća
8. Korelacija između dvije varijable je veoma visoka i negativna. Od ponuđenih koeficijenata korelacije, na nju bi se odnosio sljedeći od njih:
- a)  $r = -.52$
  - b)  $r = -.33$
  - c)  $r = -.6$
  - d)  $r = -.82$
  - e)  $r = -.007$
9. Na istim podacima smo dobili sljedeća dva koeficijenta korelacije:  $r = -.53$  i  $r = .53$ . Za njih vrijedi sljedeće:
- a) drugi je viši od prvog
  - b) prvi je viši od drugog
  - c) jednaki su po svojoj visini
  - d) oba su veoma niska



- e) oba su veoma visoka
10. *Od punuđenih tvrdnji jedna je tačna. To je:*
- a) korelacija je standardizovana kovarijansa
  - b) korelacija je standardizovana varijansa
  - c) kovarijansa je standardizovana korelacija
  - d) varijansa je standardizovana korelacija
  - e) korelacija je drugi naziv za kovarijansu
11. *Korelacija između dvije varijable je  $r = .4$ . Koliki je koeficijent determinacije ovog para varijabli?*
- a) kvadrat koeficijenta korelacije, dakle, .16
  - b) duplo niži, dakle, .2
  - c) takođe .4
  - d) koeficijent korelacije oduzet od jedinice, dakle, .6
  - e) duplo viši, dakle, .8
12. *Ako je izračunati koeficijent korelacije iznosio  $r = .7$ , zajednička varijansa dvije varijable na koje se odnosi ovaj koeficijent je:*
- a) 7%
  - b) 14%
  - c) 30%
  - d) 49%
  - e) 70%
13. *Procenat neobjašnjene varijanse varijable B, na osnovu poznatog koeficijenta korelacije između varijabli A i B ( $r_{AB} = .3$ ), je:*
- a) .09%
  - b) 3%
  - c) 30%
  - d) 70%
  - e) 91%
14. *Ako je koeficijent korelacije  $r = .66$ , a uzorak na kojem je izračunat  $N = 98$ , stepeni slobode su:*
- a) 32
  - b) 64
  - c) 65
  - d) 96

e) 97

15. Statistički zapis:  $r(116) = .41, p < .01$  se interpretira na sljedeći način:

- a) Dobijena je umjerena pozitivna i statistički značajna korelacija (uz nivo značajnosti .01 i  $df = 116$ ).
- b) Dobijena je niska pozitivna i statistički značajna korelacija (uz nivo značajnosti .01 i  $df = 116$ ).
- c) Dobijena je umjerena pozitivna i statistički značajna korelacija (uz nivo značajnosti .01 i  $N = 116$ ).
- d) Dobijena je niska pozitivna korelacija koja nije bila statistički značajna.
- e) Dobijena je umjerena pozitivna i statistički neznačajna korelacija.

16. Statistički zapis:  $r(35) = .18, p > .05$  se interpretira ovako:

- a) Dobili smo nisku pozitivnu i statistički značajnu korelaciju na nivou .05, uz  $df = 35$ .
- b) Dobili smo nisku pozitivnu i statistički neznačajnu korelaciju na nivou .05, uz  $df = 35$ .
- c) Dobili smo umjerenu pozitivnu i statistički značajnu korelaciju na nivou .05, uz  $df = 35$ .
- d) Dobili smo umjerenu pozitivnu i statističku neznačajnu korelaciju na nivou .05, uz  $N = 37$ .
- e) Dobili smo nisku pozitivnu i statističku neznačajnu korelaciju na nivou .05, koja bi bila statistički značajna na nivou .01 (uz  $N = 37$ ).

17. U tabeli koja slijedi dati su rezultati ispitanika na varijablama "optimizam" i "generalno samopoštovanje":

| Ispitanik | Optimizam | Generalno samopoštovanje |
|-----------|-----------|--------------------------|
| B. P.     | 12        | 8                        |
| K. G.     | 10        | 10                       |
| A. X.     | 9         | 8                        |
| O. L.     | 5         | 6                        |
| L. W.     | 10        | 7                        |
| Q. Y.     | 6         | 8                        |
| H. J.     | 8         | 5                        |
| L. L.     | 14        | 13                       |
| Z. R.     | 11        | 9                        |

Koliko iznosi Pearsonov koeficijent korelacije?

- a) .432
- b) .514
- c) .649

d) .721

e) .802

18. Šta možete reći za statističku (ne)značajnost ovog koeficijenta?

a) Značajan je na nivou .05, ali nije na nivou .01.

b) Značajan je na nivou .01, ali nije na .05.

c) Nije značajan niti na nivou .05 niti na .01.

d) Značajan je na nivou .01.

e) Nije značajan na nivou .05.

## REGRESIJSKA ANALIZA

### Teorijska racionala

- ✓ Jednačina linearne regresije sa dva prediktora ( $x_1$  i  $x_2$ ) i jednim kriterijumom ( $y$ ) je ( $b_0$  je odsječak na  $y$ -osi /intercept/, a  $b_1$  i  $b_2$  su nestandardizovani regresijski koeficijenti):

$$y' = b_0 + b_1x_1 + b_2x_2$$

- ✓ Jednačina za predviđanje rezultata kriterijuma na osnovu jednog prediktora može se prikazati na sljedeći način:  $y' = b_0 + bx$ . Povezani obrasci su:

$$b = \frac{n \sum(xy) - \sum x \sum y}{n \sum x^2 - (\sum x)^2}$$

$$b_0 = M_y - bM_x$$

- ✓ Ako raspoložemo aritmetičkim sredinama, standardnim devijacijama i korelacijom između prediktora i kriterijuma, koristimo sljedeći obrazac:

$$y' = r_{xy} \frac{s_y}{s_x} (X - M_x) + M_y$$

- ✓ Koeficijent *multiple determinacije* za dva prediktora (1 i 2) i kriterijum ( $y$ ):

$$R^2 = \frac{r_{y1}^2 + r_{y2}^2 - 2r_{y1}r_{y2}r_{12}}{1 - r_{12}^2}$$

- ✓ Koeficijent *alijenacije* (proporcija neobjašnjene varijanse):

$$1 - R^2$$

- ✓ Koeficijent *multiple korelacije* jednak je korijenu koeficijenta multiple determinacije:

$$R = \sqrt{R^2}$$

- ✓ Koeficijent multiple determinacije na osnovu beta-pondera (za dva prediktora):

$$R^2 = \beta_1 r_{y1} + \beta_2 r_{y2}$$

- ✓ Pretvaranje beta-pondera prediktora  $i$  u nestandardizovani regresijski koeficijent:

$$b = \beta_i \frac{s_y}{s_i}$$

- ✓ Procenat objašnjene varijanse kriterijuma dobija se množenjem koeficijenta multiple determinacije sa 100.
- ✓ Procenat neobjašnjene varijanse tj. procenat ukupne varijanse kriterijuma koji otpada na rezidual dobija se množenjem koeficijenta alijenacije sa 100.
- ✓ Standardna greška prognoze rezultata varijable  $y$  na osnovu rezultata varijable  $x$  je (kako uzorak postaje veći, vrijednost razlomka pod korijenom teži jedinici):

$$s_{y(x)} = s_y \sqrt{1 - r_{xy}^2} \sqrt{\frac{N-1}{N-2}}$$

### **Pitanja i zadaci**

1. *U regresijskoj analizi, varijabla koja nam služi za predviđanje rezultata neke druge varijable naziva se:*
  - a) prediktorska varijabla
  - b) kriterijumska varijabla
  - c) nezavisna varijabla
  - d) zavisna varijabla
  - e) konfundirajuća varijabla
2. *Ako je korelacija između dva prediktora nula, njihovi beta-ponderi u kontekstu predviđanja treće varijable su uvijek:*
  - a) jednaki nuli
  - b) međusobno jednaki
  - c) takvi da su suprotnih predznaka, a jednake veličine
  - d) jednaki koeficijentima korelacije ovih prediktora i kriterijske varijable
  - e) statistički neznačajni
3. *Stav prema fizici je u korelaciji sa ocjenom iz ovog predmeta, a ova povezanost iznosi  $r = .4$ . Generalna inteligencija je sa ocjenom iz fizike u nešto većoj korelaciji ( $r = .5$ ). Pod pretpostavkom da stavovi prema fizici i generalna inteligencija nisu u međusobnoj vezi, koliko iznosi njihov koeficijent multiple determinacije prilikom predviđanja ocjena iz fizike?*
  - a)  $R^2 = .10$

- b)  $R^2 = .20$
  - c)  $R^2 = .41$
  - d)  $R^2 = .59$
  - e)  $R^2 = .90$
4. *Koeficijent multiple korelacije izračunat na osnovu podataka iz zadatka br. 3 iznosi:*
- a)  $R = .02$
  - b)  $R = .10$
  - c)  $R = .36$
  - d)  $R = .64$
  - e)  $R = .90$
5. *Koeficijent alijenacije, određen na osnovu podataka iz zadatka 3, jeste:*
- a)  $R = .10$
  - b)  $R = .36$
  - c)  $R = .41$
  - d)  $R = .59$
  - e)  $R = .64$
6. *Prediktori su radno iskustvo i motivacija za posao. Kriterijum je zadovoljstvo poslom. Beta-ponder radnog iskustva je  $\beta_1 = .252$ , a motivacije za posao  $\beta_2 = .465$ . Z-vrijednost ispitanikovog radnog iskustva je  $-0.5$ , a njegove motivacije  $1.75$ . Koji je njegov najvjerojatniji rezultat vezan za zadovoljstvo poslom, izražen u obliku z-vrijednosti?*
- a)  $-0.10$
  - b)  $0.69$
  - c)  $1.25$
  - d)  $1.97$
  - e)  $2.25$
7. *U principu, beta-ponder može imati sljedeće vrijednosti:*
- a) od  $-\infty$  do  $+\infty$
  - b) od  $-100$  do  $100$
  - c) od  $-10$  do  $10$
  - d) od  $-.5$  do  $.5$
  - e) od  $-1$  do  $1$

8. U nastavku je matrica interkorelacija ekstraverzije, saradljivosti i neuroticizma (emocionalne nestabilnosti) kao crta ličnosti. Na osnovu ekstraverzije i neuroticizma pokušavamo predvidjeti stepen u kome je neko saradljiv.

|               | ekstraverzija | saradljivost | neuroticizam |
|---------------|---------------|--------------|--------------|
| ekstraverzija | 1             | .350         | -.390        |
| saradljivost  |               | 1            | -.410        |
| neuroticizam  |               |              | 1            |

Beta-ponderi za ekstraverziju i neuroticizam su, redom:

- a)  $\beta_1 = .224$  i  $\beta_2 = -.323$   
 b)  $\beta_1 = -.224$  i  $\beta_2 = .323$   
 c)  $\beta_1 = -.224$  i  $\beta_2 = -.323$   
 d)  $\beta_1 = .224$  i  $\beta_2 = .323$   
 e) jednaki odgovarajućim koeficijentima korelacije ( $\beta_1 = r_1$  i  $\beta_2 = r_2$ )
9. Koeficijent multiple determinacije izračunat na osnovu podataka iz zadatka br. 8, je:  
 a)  $R^2 = -.450$   
 b)  $R^2 = .211$   
 c)  $R^2 = .383$   
 d)  $R^2 = .459$   
 e)  $R^2 = 1.150$
10. Koeficijent multiple korelacije, dobijen za podatke iz zadatka br. 8, iznosi:  
 a)  $R = -.450$   
 b)  $R = .211$   
 c)  $R = .383$   
 d)  $R = .459$   
 e)  $R = 1.150$
11. Standardna devijacija prediktora je 5, a kriterijuma 8. Beta ponder ovog prediktora je .504. Nestandardizovani regresijski koeficijent ovog prediktora iznosi:  
 a)  $b = -0.806$   
 b)  $b = -0.315$   
 c)  $b = 0.315$   
 d)  $b = 0.504$   
 e)  $b = 0.806$
12. Data je regresijska jednačina:  $y = -0.4 - 3.1x_1 + 2.9x_2$ . Ako je ispitanik postigao rezultat 12 na prvom, a 22 na drugom prediktoru, koliki je njegov prognozirani rezultat na kriterijumu?

- a) -101.4
- b) 26.2
- c) 27
- d) 34
- e) 100.6

13. Dva ispitanika su na prediktoru postigli sljedeće rezultate:  $X_1 = 18$  i  $X_2 = 23$ . Koliki su njihovi prognozirani rezultati na kriterijumu? Poznato je sljedeće:  $M_{\text{prediktora}} = 20$ ,  $M_{\text{kriterijuma}} = 28$ ,  $S_{\text{prediktora}} = 5$ ,  $S_{\text{kriterijuma}} = 6$ , a korelacija između kriterijuma i prediktora je .65.

- a)  $Y_1 = 26.44$ ,  $Y_2 = 30.34$
- b)  $Y_1 = 29.56$ ,  $Y_2 = 30.34$
- c)  $Y_1 = 30.34$ ,  $Y_2 = 26.44$
- d)  $Y_1 = 30.34$ ,  $Y_2 = 29.56$
- e)  $Y_1 = 26.44$ ,  $Y_2 = 29.56$

14. Ako podacima iz prethodnog zadatka dodamo onaj o veličini uzorka ( $N = 42$ ), koliko iznosi standardna greška prognoze rezultata kriterijuma, na osnovu rezultata na prediktoru?

- a) 0.59
- b) 0.77
- c) 3.55
- d) 3.59
- e) 4.62

15. Ako sada pokušamo predvidjeti rezultate prediktora iz zadatka br. 13, na osnovu varijable koja nam je bila kriterijum, kolika je "nova" standardna greška prognoze?

- a) 0.59
- b) 0.77
- c) 3.55
- d) 3.85
- e) 4.62

16. Ispitivana je povezanost između socijalne inteligencije i radne učinkovitosti. Slijede podaci za jedan uzorak zaposlenika:

| R. br. ispitanika | Socijalna inteligencija | Radna učinkovitost |
|-------------------|-------------------------|--------------------|
| 1.                | 14                      | 17                 |
| 2.                | 12                      | 30                 |
| 3.                | 24                      | 29                 |



|     |    |    |
|-----|----|----|
| 4.  | 27 | 40 |
| 5.  | 17 | 30 |
| 6.  | 24 | 25 |
| 7.  | 28 | 27 |
| 8.  | 21 | 30 |
| 9.  | 10 | 15 |
| 10. | 12 | 26 |
| 11. | 8  | 14 |
| 12. | 14 | 26 |
| 13. | 25 | 30 |

---

*Koliko iznosi koeficijent korelacije između ove dvije varijable i da li je on statistički značajan? Odaberite odgovarajući statistički zapis:*

- a)  $r(12) = .407, p > .05$
- b)  $r(12) = .456, p > .05$
- c)  $r(12) = .552, p > .05$
- d)  $r(12) = .675, p < .05$
- e)  $r(12) = .697, p < .05$

17. *Na osnovu podataka iz zadatka br. 16, regresijska jednačina izgleda ovako:*

- a)  $y' = 13.43 + 0.697x$
- b)  $y' = 13.43 + 0.675x$
- c)  $y' = 0.675 + 13.43x$
- d)  $y' = 0.697 + 13.43x$
- e)  $y' = 13.43 + 0.697x$

18. *Koristeći se rješenjima zadataka br. 16 i 17., jasno je da, ako se nečiji rezultat na mjeri socijalne inteligencije poveća za 10, njegova radna učinkovitost:*

- a) smanji se za 6.97
- b) smanji se za 6.75
- c) ostaje ista
- d) poveća se za 6.75
- e) poveća se za 6.97

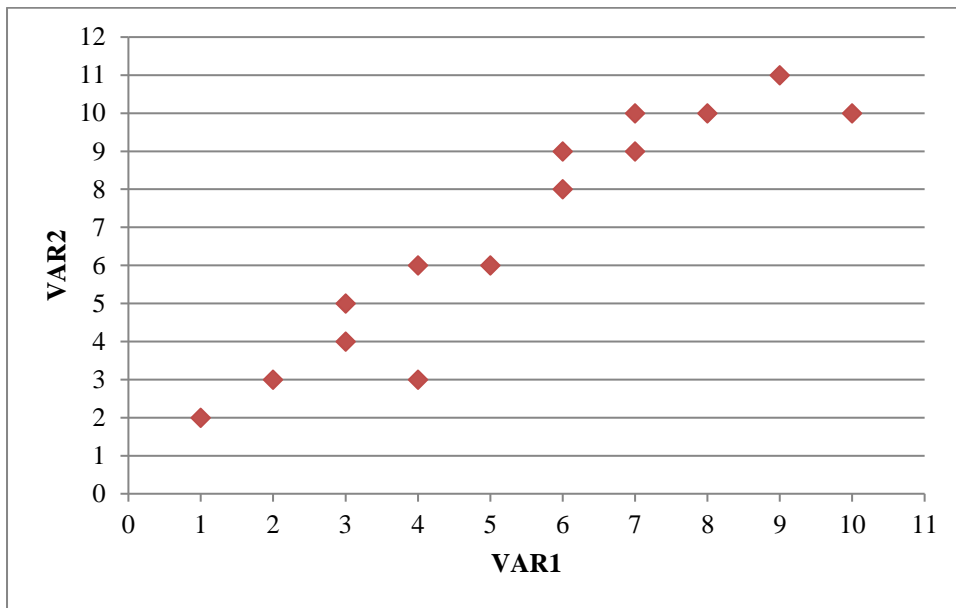
19. *Kolika je standardna greška prognoze radne učinkovitosti na osnovu socijalne inteligencije (potrebni podaci su u zadatku br. 16)?*

- a) 5.21
- b) 5.36
- c) 5.37
- d) 5.53
- e) 5.66

20. Ako je zaposlenik na korištenoj mjeri socijalne inteligencije u zadatku br. 16 postigao rezultat  $X = 20$ , koji je njegov prognozirani rezultat na mjeri radne učinkovitosti?

- a)  $Y' = 18.15$
- b)  $Y' = 20$
- c)  $Y' = 26.08$
- d)  $Y' = 26.93$
- e)  $Y' = 27.37$

21. Na raspolaganju vam je skater-dijagram:



Kako se može primijetiti, korelacija između VAR1 i VAR2 je:

- a) veoma niska i negativna
- b) veoma niska i pozitivna
- c) osrednje visine
- d) veoma visoka i negativna
- e) veoma visoka i pozitivna

22. Ako su sve vrijednosti na grafikonu iz zadatka br. 21 cijeli brojevi, koliko iznosi Pearsonov koeficijent korelacije između pomenute dvije varijable?

- a) -.94
- b) -.74
- c) .50
- d) .74
- e) .94

***Teorijska racionala***

- ✓ Od "mora" neparametrijskih statističkih procedura, odabrao sam tri koje se, po svim prilikama, najčešće koriste: hi-kvadrat test, Mann-Whitney *U* test i Spearmanov koeficijent korelacije.
- ✓ *Hi-kvadrat test* za jedan uzorak koristi se u kontekstu nominalnih varijabli, tačnije da se ispita ima li statistički značajne razlike u broju ispitanika između određenih kategorija date varijable. Hi-kvadrat test za *dva ili više nezavisnih uzoraka* izračunava se za podatke prikazane u *kontingencijskim* tabelama. U obrascu koji slijedi,  $f_o$  su opažene (od eng. *observed*) frekvencije, a  $f_e$  (od eng. *expected*) očekivane frekvencije:

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

- ✓ *Mann-Whitney U test* koristi se umjesto t-testa za nezavisne uzorke kada su podaci s ordinalne skale mjerenja ili sa intervalne skale, a da pritom njihova distribucija statistički značajno odstupa od normalne. *U*-statistik se za svaku kompariranu grupu računa na sljedeći način:

$$U = \sum R - \frac{n(n+1)}{2}$$

Za dovoljno velike uzorke ( $n_1 + n_2 > 30$ , Freund i Wilson, 2003), u okviru Mann-Whitney *U* testa se može računati *z*-vrijednost i to uz pomoć ovog obrasca ( $n_1$  i  $n_2$  je broj ispitanika u prvoj i drugoj grupi koje se porede, a u obrazac se uvrštava manja *U*-vrijednost dobijena pomoću prethodne formule):

$$z = \frac{U - \frac{n_1 n_2}{2}}{\sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}}$$

- ✓ *Spearmanov* (čita se: "Spirmanov") *koeficijent (rang) korelacije* može se koristiti umjesto Pearsonovog koeficijenta, kada su podaci na ordinalnoj skali ili kada njihova distribucija statistički značajno odstupa od normalne krive. Obično se označava  $\rho$  (čita se: "ro") ili  $r_s$ . Obrazac izgleda ovako (*D* je razlika /diferencija/ između rangova dva skupa podataka čija se međusobna korelacija provjerava na ovaj način):

$$r_s = \frac{1 - 6 \sum D^2}{n(n^2 - 1)}$$

### Pitanja i zadaci

1. *Proizvođač igrački želi plasirati na tržište jednu od pet igrački koje je osmislio. Djeca od tri do šest godina birali su koja im se igračka najviše sviđa. Rezultati su sljedeći:*

|   | Igračka<br>A | Igračka<br>B | Igračka<br>C | Igračka<br>D | Igračka<br>E |
|---|--------------|--------------|--------------|--------------|--------------|
| Broj djece kojima se najviše sviđa određena igračka | 12           | 18           | 35           | 24           | 21           |

*Koliko iznosi očekivana frekvencija za svaku vrstu igračke, ako su sve očekivane frekvencije jednake?*

- 12
  - 16
  - 18
  - 22
  - 26
2. *Koji je broj stepeni slobode u zadatku br. 1?*
- $df = 1$
  - $df = 2$
  - $df = 3$
  - $df = 4$
  - $df = 5$
3. *Na osnovu podataka iz zadatka br. 1, koliko iznosi vrijednost hi-kvadrat testa i da li je statistički značajna na nivou .05?*
- 1.38, nije
  - 1.83, nije
  - 3.81, jeste
  - 8.31, jeste
  - 13.18, jeste
4. *Zanima nas da li rezultati izbora (glasanja) u jednom malom mjestu odstupaju od očekivanja jedne organizacije koja je provela predizbornu anketu. U "igri" su tri politička rivala, što je uz ostale podatke prikazano u sljedećoj tabeli:*

|                                    | Isaac<br>Maslow | Abraham<br>Fromm | Erich<br>Newton |
|------------------------------------|-----------------|------------------|-----------------|
| Stvarni rezultati izbora*          | 308             | 154              | 162             |
| Očekivanja anketarske organizacije | 312             | 137              | 175             |

\* na osnovu slučajnog uzorka glasačkih listića

*Za početak, koliko iznose stepeni slobode?*

- a)  $df = 0$
- b)  $df = 1$
- c)  $df = 2$
- d)  $df = 3$
- e)  $df = 4$

5. *Ako očekivanja anketarske organizacije iz zadatka br. 4 shvatimo kao očekivane frekvencije, da li ima statistički značajne razlike između očekivanih i stvarnih rezultata?*

- a) Nema, jer je dobijen  $\chi^2 = 1.33$  koji je statistički neznačajan.
- b) Nema, jer je dobijen  $\chi^2 = 3.13$  koji je statistički neznačajan.
- c) Ima, jer je dobijen  $\chi^2 = 3.13$  koji je statistički značajan.
- d) Nema, jer je dobijen  $\chi^2 = 13.31$  koji je statistički neznačajan.
- e) Ima, jer je dobijen  $\chi^2 = 13.31$ , koji je statistički značajan.

6. *Kakav je zaključak baziran na statističkom rezultatu iz pitanja br. 5?*

- a) Anketarska organizacija je u principu obavila dobar posao.
- b) Anketarska organizacija je statistički značajno precijenila broj glasova za kandidata Isaaca Maslowa.
- c) Anketarska organizacija je statistički značajno precijenila broj glasova za Ericha Newtona.
- d) Anketarska organizacija je statistički značajno potcijenila broj glasova za kandidata Abrahama Fromma.
- e) Anketarska organizacija je obavila loš posao, budući da je "promašila" broj glasova za sva tri kandidata.

7. *Istraživač je ispitao određeni broj muškaraca i žena o njihovim stavovima prema odlasku iz vlastite države u prosperitetnije zemlje. Dobijeni rezultati prikazani su u tabeli kontingencije:*

|                | Muškarci | Žene |
|----------------|----------|------|
| Za odlazak     | 78       | 52   |
| Protiv odlaska | 49       | 80   |

*Koja je očekivana (teorijska) frekvencija za polje tabele u kojem je naveden broj muškaraca koji su protiv odlaska u prosperitetnije zemlje?*

- a) 63.25
- b) 63.75
- c) 64.75
- d) 65.75
- e) 66.25

8. *Koliko iznose stepeni slobode za podatke iz zadatka br. 7?*

- a) 1
- b) 2
- c) 4
- d) 6
- e) 8

9. *Koji statistički zapis je rezultat provjere značajnosti hi-kvadrat statistika iz zadatka br. 7?*

- a)  $\chi^2(1, N = 259) = 12.56, p > .05$
- b)  $\chi^2(1, N = 259) = 12.56, p < .01$
- c)  $\chi^2(1, N = 259) = 2.56, p > .05$
- d)  $\chi^2(259, N = 1) = 2.56, p < .05$
- e)  $\chi^2(259, N = 1) = 12.56, p < .01$

10. *Koji je najvjerojatniji zaključak na osnovu pregleda tabele iz zadatka br. 7 i rezultata iz zadatka br. 9?*

- a) Muškarci i žene se ne razlikuju statistički značajno po stavovima o odlasku iz svoje zemlje u države većeg prosperiteta.
- b) Muškarci imaju pozitivnije stavove prema odlasku, ali nedovoljno da bi razlika bila proglašena statistički značajnom.
- c) Žene imaju negativnije stavove prema odlasku, ali nedovoljno da bi razlika bila smatrana statistički značajnom.
- d) Muškarci i žene se statistički značajno razlikuju po svojim stavovima o pomenutoj temi, pri čemu žene imaju pozitivnije stavove.
- e) Muškarci i žene se statistički značajno razlikuju po svojim stavovima o pomenutoj temi, pri čemu muškarci imaju pozitivnije stavove.

11. *Dvije grupe učenika (u svakoj je bilo po osmoro djece) su procjenjivane po učešću i zalaganju na nastavi. Jedna grupa bila je izložena tradicionalnom načinu*

podučavanja, a druga nekom od savremenih oblika podučavanja (npr. onom koje je omogućavalo veću interakciju učenika i učiteljice). Rezultati su prikazani u tabeli ispod:

| Prva grupa<br>(tradicionalno podučavanje) | Druga grupa<br>(savremeno podučavanje) |
|---|--|
| 7   | 10                                     |
| 9   | 8                                      |
| 14  | 16                                     |
| 3   | 5                                      |
| 11  | 17                                     |
| 5   | 9                                      |
| 4   | 8                                      |
| 6   | 12                                     |

Koliko iznosi Mann-Whitney U statistik za prvu, a koliko za drugu grupu učenika?

- a)  $U_1 = 53, U_2 = 83$
- b)  $U_1 = 17, U_2 = 47$
- c)  $U_1 = 17, U_2 = 53$
- d)  $U_1 = 47, U_2 = 83$
- e)  $U_1 = 47, U_2 = 53$

12. Da smo u zadatku br. 11 imali veći uzorak, mogli bismo izračunati i z-vrijednost U-statistika. Ako je u prvoj grupi bilo 145, a u drugoj 152 učenika (pod uslovom da U-statistik ostaje isti kao u zadatku br. 11), kolika je z-vrijednost pomenutog statistika?

- a)  $z = 0.005$
- b)  $z = 5.98$
- c)  $z = -4.29$
- d)  $z = -14.87$
- e)  $z = -25.71$

13. Ocjene deset studenata na kolegijima Psihologija ličnosti i Socijalna psihologija prikazane su u tabeli koja slijedi.

| Student | Psihologija ličnosti | Socijalna psihologija |
|---------|----------------------|-----------------------|
| A. N.   | 8                    | 9                     |
| K. S.   | 7                    | 8                     |
| S. R.   | 6                    | 7                     |
| I. I.   | 7                    | 7                     |
| B. Z.   | 9                    | 8                     |
| M. O.   | 10                   | 10                    |
| C. D.   | 5                    | 5                     |
| E. H.   | 6                    | 7                     |
| G. T.   | 6                    | 6                     |
| T. V.   | 8                    | 6                     |

*Kako su ocjene, u biti, ordinalna varijabla, zanima nas koliki je Spearmanov koeficijent korelacije između ocjena dobijenih na ova dva kolegija. Njegova vrijednost je:*

a)  $r_s = .736, p < .05$

b)  $r_s = .746, p < .05$

c)  $r_s = .756, p > .05$

d)  $r_s = .766, p > .05$

e)  $r_s = .776, p < .05$

14. *Ako rangiramo studente po ocjenama (rang jedan pridružimo najvišoj ocjeni, itd.) iz zadatka br. 13, Spearmanov koeficijent dobijen na ovakvim rangovima iznosiće:*

a)  $r_s = .736, p < .05$

b)  $r_s = .746, p < .05$

c)  $r_s = .756, p > .05$

d)  $r_s = .766, p > .05$

e)  $r_s = .776, p < .05$



*Teorijska racionala*

- ✓ Vjerovatnoća da se desi događaj  $A$  je  $P(A)$ , pri čemu vrijedi:  $0 \leq P(A) \leq 1$ .
- ✓ *Matematički eksperiment* je npr. bacanje kocke ili novčića, te izvlačenje kuglice iz kutije.
- ✓  $S$  je skup (prostor) svih mogućih elementarnih događaja (ishoda) ovakvog eksperimenta (npr. kada bacimo standardnu, pravilnu šestostranu kocku, možemo dobiti neki od brojeva od jedan do šest).
- ✓ *Elementarni događaj* (ishod) je npr. dobijanje broja četiri prilikom bacanja kocke ili "pisma" bacanjem novčića.
- ✓ Vjerovatnoća da će se skupa dogoditi dva ili više *nezavisnih događaja* je (obrazac za slučaj tri nezavisna događaja):

$$P(A \cap B \cap C) = P(A) \times P(B) \times P(C)$$

- ✓ Vjerovatnoća da se desi bar neki od *međusobno isključivih događaja* (ishoda):

$$P(A \cup B) = P(A) + P(B)$$

- ✓ Vjerovatnoća događaja koji se međusobno *ne isključuju*:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

- ✓ Uslovna vjerovatnoća odnosi se na odvijanje izvjesnog događaja (nastupanje nekog ishoda) pod uslovom da se već desio neki događaj (ishod). Za događaje  $A$  i  $B$  (nastupanje događaja  $B$  pod uslovom da se desio događaj  $A$ ) vrijedi:

$$P(B | A) = \frac{P(A \cap B)}{P(A)}$$

- ✓ Uzorkovanje sa vraćanjem/zamjenom (eng. *sampling with replacement*) je vrsta "eksperimenta vjerovatnoće" u kojem prije svakog sljedećeg izvlačenja (npr. kuglica iz kutije, karti iz špila...) vratimo/zamijenimo ono što smo izvukli (Howell, 2010). Naravno da je vjerovatnoća određenog događaja u ovakvom "eksperimentu" drugačija sa i bez vraćanja/zamjene.
- ✓ Iz *kombinatorike*, kao još jedne zanimljive discipline primijenjene matematike, upoznaću vas sa osnovnim obrascima za računanje permutacija, varijacija i kombinacija. Prije obrazaca, prisjetite se da je npr.  $5!$  ("pet faktorijel") =  $5 \cdot 4 \cdot 3 \cdot 2 \cdot 1$ .

- ✓ Ako neki skup podataka ima  $n$  elemenata, broj *permutacija bez ponavljanja* ovih elemenata je:

$$P(n) = n!$$

- ✓ Ako se  $i$ -ti element niza pojavljuje  $r_i$  puta, a dužina niza je  $n$ , riječ je o *permutacijama sa ponavljanjem*:

$$P_{r_1, \dots, r_n}(n) = \frac{n!}{r_1! r_2! \dots r_n!}$$

- ✓ *Varijacije  $k$ -te klase bez ponavljanja* su skupovi od nekog broja elemenata koji je manji ili jednak ukupnom broju elemenata u skupu ( $n$ ):

$$V_n^k = \frac{n!}{(n-k)!}$$

- ✓ *Varijacijama  $k$ -te klase sa ponavljanjem*, iz skupa  $n$  elemenata, odgovara sljedeći obrazac:

$$\bar{V}_n^k = n^k$$

- ✓ *Kombinacije  $k$ -te klase bez ponavljanja* (tj. njihov broj) matematički su definisane na sljedeći način (*napomena*: kod kombinacija nije bitan poredak elemenata prilikom njihovog grupisanja ili izvlačenja):

$$C_n^k = \frac{V_n^k}{k!} = \frac{n!}{k!(n-k)!}$$

- ✓ Za *kombinacije  $k$ -te klase sa ponavljanjem* (o ovdje je riječ o slučaju kada nije bitan redoslijed biranja nekog elementa ili više njih, već broj biranja) vrijedi:

$$\bar{C}_n^k = \frac{(n+k-1)!}{k!(n-1)!}$$

**Pitanja i zadaci**

1. *Kolika je vjerovatnoća da u jednom bacanju kocke dobijete broj šest?*
  - a) 1/6
  - b) 2/6
  - c) 3/6
  - d) 4/6
  - e) 5/6
2. *Kocku bacate dva puta. Kolika je vjerovatnoća da oba puta dobijete broj jedan?*

- a)  $1/6$
  - b)  $1/12$
  - c)  $1/18$
  - d)  $1/24$
  - e)  $1/36$
3. *Kolika je vjerovatnoća da u oba bacanja iz zadatka br. 2 dobijete peticu?*
- a)  $1/6$
  - b)  $1/12$
  - c)  $1/18$
  - d)  $1/24$
  - e)  $1/36$
4. *Ako je u jednoj zemlji nepismenost 8.5%, koja je vjerovatnoća da slučajnim odabirom nekog člana populacije on bude pismen?*
- a) -8.5%
  - b) 8.5%
  - c) 85%
  - d) 9.15%
  - e) 91.5%
5. *Ako jednom bacite kocku, pa dva puta bacite novčić, koja je vjerovatnoća da redom dobijete trojku, "pismo" i "glavu"?*
- a)  $1/2$
  - b)  $1/6$
  - c)  $1/12$
  - d)  $1/18$
  - e)  $1/24$
6. *Ako dvaput bacite novčić, pa onda jednom kocku, koja je vjerovatnoća da redom dobijete "pismo", "glavu" i trojku?*
- a)  $1/24$
  - b)  $1/18$
  - c)  $1/12$
  - d)  $1/6$
  - e)  $1/2$

7. *Koja je vjerovatnoća da psiholog, bacajući kocku jednom, dobije paran broj? Poslije njega, kocku jednom baca sociolog. Koja je vjerovatnoća da sociolog dobije paran broj?*
- Vjerovatnoća je u oba slučaja ista i iznosi  $1/6$ .
  - Vjerovatnoća je u oba slučaja ista i iznosi  $1/2$ .
  - Vjerovatnoća se razlikuje. Pošto psiholog baca prvi, vjerovatnoća pomenutog ishoda je  $1/2$ , dok je u slučaju sociologa  $1$ .
  - Vjerovatnoća se razlikuje. Pošto psiholog baca prvi, vjerovatnoća pomenutog ishoda je  $1$ , dok je u slučaju sociologa  $1/2$ .
  - Vjerovatnoća se razlikuje. Pošto psiholog baca prvi, vjerovatnoća pomenutog ishoda je  $1/6$ , dok je u slučaju sociologa  $1/2$ .
8. *Igrač iz špila od 52 karte izvlači jednu kartu (koju ne vraća u špil prije drugog izvlačenja), potom drugu (koju takođe ne vraća u špil prije trećeg izvlačenja), pa treću. Koja je vjerovatnoća da je, redom, izvukao keca tref, kralja herc i desetku karo?*
- $1/52 * 1/52 * 1/52$
  - $1/50 * 1/50 * 1/50$
  - $1/52 * 1/51 * 1/50$
  - $1/52 * 1/51 * 1/51$
  - $1/51 * 1/50 * 1/50$
9. *Antropolog namjerava ugostiti u svojoj kući dva Italijana, tri Francuza i četiri Engleza. Svi se nalaze ispred ulaznih vrata njegove kuće. Koja je vjerovatnoća da će drugi ući jedan Englez, pod uslovom da je prvi ušao jedan Francuz?*
- $1/9$
  - $1/8$
  - $1/4$
  - $1/3$
  - $1/2$
10. *Ako su događaji D i F međusobno nezavisni i ako je vjerovatnoća prvog događaja  $1/7$  i drugog  $1/5$ , koja je vjerovatnoća nastupanja događaja D pod uslovom da se desio događaj F?*
- $1/5$
  - $1/7$
  - $1/12$
  - $12/35$

e)  $1/35$

11. *Profesor geografije napravio je test od 20 pitanja sa po četiri ponuđena odgovora (samo jedan od njih je tačan). Ako učenik nasumice zaokruži neki odgovor unutar svakog pitanja, koji je njegov najvjerojatniji (očekivani) rezultat na testu (napomena: svaki tačan odgovor nosi po jedan bod)?*

a) 0

b) 2

c) 4

d) 5

e) 10

12. *U ruletu je 18 crnih i isto toliko crvenih polja, a postoji i polje "nula", te u nekim ruletima uz polje "nula" i polje "dvije nule" (00). Koja je vjerovatnoća da kuglica stane na polju crvene boje?*

a) između 0 i .45

b) veća od .45, a manja od .50

c) tačno .50

d) veća od .50, a manja od .55

e) veća od .55

13. *Koliko se riječi (bez obzira jesu li smislene ili besmislene) od pet slova možete napisati ako raspolazete samo ovim slovima: A, D, K, E i S? Napomena: slova se ne smiju ponavljati.*

a) 20

b) 25

c) 60

d) 120

e) 240

14. *Koliko šestocifrenih brojeva možete sastaviti od cifri iz skupa u zagradi {1, 1, 3, 3, 3, 5}?*

a) 30

b) 60

c) 180

d) 360

e) 720

15. U kutiji su tri iste knjige iz pedagogije, dvije iste knjige iz psihologije i četiri iste knjige iz ekonomije. Na koliko načina možete uzeti (izvući) dvije knjige (poredak je bitan)?

- a) 9
- b) 12
- c) 15
- d) 18
- e) 24

16. Ako u zadatku br. 15 redosljed izvučenih elemenata nije bitan, koji ćete obrazac primijeniti? Obrazac za računanje broja:

- a) permutacija s ponavljanjem
- b) varijacija bez ponavljanja
- c) varijacija sa ponavljanjem
- d) kombijacija bez ponavljanja
- e) kombinacija sa ponavljanjem

17. Koliko je mogućih ishoda izvlačenja primjenom obrasca iz zadatka br. 16?

- a) 2
- b) 3
- c) 4
- d) 5
- e) 6

## ODGOVORI NA PITANJA I RJEŠENJA ZADATAKA

### TABELARNO I GRAFIČKO ORGANIZOVANJE I PRIKAZIVANJE PODATAKA

1.

| Stručna sprema               | <i>f</i>  | %          |
|------------------------------|-----------|------------|
| Osnovna škola                | 150       | <b>15</b>  |
| Srednja stručna sprema (SSS) | 500       | 50         |
| Viša škola (VŠS)             | <b>70</b> | <b>7</b>   |
| Visoka stručna sprema (VSS)  | 250       | <b>25</b>  |
| Magisterij ili doktorat      | 30        | <b>3</b>   |
| Total                        | 1000      | <b>100</b> |

2. Uzorak je činilo  $N = 200$  ispitanika, od toga **65** muškaraca (što je **32.5 %** uzorka) i **135** žena (tj. **67.5 %** uzorka). U uzorku je bilo najmanje politologa, i to njih **10** (odnosno, **5 %** ukupnog uzorka). Pored toga, imali smo jednak broj psihologa i socijalnih radnika (po **50**, odnosno po **25 %** uzorka). Među kriminalističarima, bilo je **15** muškaraca. Isti broj muškaraca bio je među **sociolozima** i **psihijatrima**. Podjednak broj muškaraca i žena bio je među **psihijatrima**. Najviše žena bilo je među psiholozima, i to njih **45**, što je **33.3 %** ukupnog broja žena u našem uzorku.

3. Dakle, u uzorku je **24** oženjenih/udatih osoba (što je **48 %** ukupnog broja ispitanika), **10** razvedenih osoba (tj. **20 %** uzorka) i **4** udovca/udovice (**8 %** uzorka).

4. Na prvoj godini ekonomije na datom fakultetu ima ukupno **96** studenata. Najveći broj njih dobio je ocjenu **7**. Procenat studenata sa ovom ocjenom u ukupnom uzorku bio je **29.2 %**. Najmanje studenata (tj. njih **troje**) dobilo je ocjenu **10**. Devetku je dobilo **9** studenata, što je **9.4 %** ukupnog uzorka. Devetku i desetku je dobilo ukupno **12** studenata, tj. **12.5 %** ukupnog uzorka. Procenat studenata koji su dali podatak o svojoj ocjeni je **93.8 %**.

5.

| Razred  | <i>f</i>  | <i>f</i> (%) | <i>rf</i>  | <i>cf</i>  | <i>cf</i> (%) | <i>rcf</i> |
|---------|-----------|--------------|------------|------------|---------------|------------|
| 1. – 3. | 85        | <b>47.22</b> | <b>.47</b> | 85         | <b>47.22</b>  | <b>.47</b> |
| 4. – 6. | 60        | <b>33.33</b> | <b>.33</b> | <b>145</b> | <b>80.56</b>  | <b>.81</b> |
| 7. – 9. | <b>35</b> | <b>19.44</b> | <b>.19</b> | <b>180</b> | <b>100</b>    | <b>1</b>   |
| Total   | 180       | ≈100         | ≈1         |            |               |            |

6. Neki od prvih šest razreda pohađa **145** djece, što je **80.56** % ukupnog broja osnovnoškolaca; neki od prva tri razreda pohađa **47.22** % osnovnoškolaca; a relativna frekvencija broja učenika u sedmom do devetog razreda je **.19**.

7. a) Kružnim dijagramima se prikazuju frekvencije (broj ispitanika) i/ili procenti (%) po određenim kategorijama (npr. broj/procentat muškaraca, odnosno žena u nekom uzorku).

8. c) Relativna frekvencija (*rf*) je omjer broja ispitanika u nekom razredu (kategoriji) i ukupnog broja ispitanika (tj. cijelog uzorka).

9. d) Zbir svih relativnih frekvencija (zanemarujući sitne razlike oko zaokruživanja) nekog niza podataka je 1.

10. Zadovoljstvo roditeljstvom u poduzorku mlađih muškaraca iznosilo je **četiri**, dok je zadovoljstvo roditeljstvom u poduzorku starijih žena bilo **pet**. Procjena ispitanika u grupi **starijih muškaraca** iznosila je dva, a u grupi **mlađih žena** tri.

11. Najviše slučajeva podmićivanja zabilježeno je **2016.** godine. Najmanji broj silovanja bio je tokom **2016.** godine i iznosio je **pet**. Broj provalništava u 2014. godini iznosi **26.11** % od ukupnog broja provalništava u sve četiri godine. U odnosu na 2014. godinu, broj djela iz kategorije "provalništvo" 2015. godine bio je manji za **16.34** %. Ukupan broj silovanja u prikazanom četvorogodišnjem periodu iznosio je **26**. Najveći broj prikazanih kriminalnih djela (bez obzira na njihovu vrstu) desio se u **2013.** godini.

12.

| <i>P</i> | <i>N</i> | <i>M</i> | <i>I</i> | <i>E</i> | <i>n</i>    | <i>m</i>    | <i>VI</i>     | <i>S</i>     | <i>SPS (1000)</i> |
|----------|----------|----------|----------|----------|-------------|-------------|---------------|--------------|-------------------|
| 80026    | 241      | 163      | 140      | 1200     | <b>3.01</b> | <b>2.04</b> | <b>147.85</b> | <b>-1060</b> | <b>-12.27</b>     |

(Napomena:  $n = 1000 * N/P$ ,  $m = 1000 * M/P$ )

13. e) Na svakih 1000 stanovnika, došlo je do smanjenja stanovništva za 12.27 (približno 12 stanovnika). Kako je populacija sa zaključenjem 2015. godine iznosila 80026, do kraja 2016. godine smanjila se za ukupno 981.92 (približno 982) stanovnika.

### ***MJERE CENTRALNE TENDENCIJE I VARIJABILNOSTI***

| <i>Red. br.</i> | <i>Tačno rješenje</i> | <i>Objašnjenje</i>  |
|-----------------|-----------------------|---|
| 1.              | e)                    | Pisac koristi najviše pridjeva, a ove vrste riječi smo označili brojem 5. |
| 2.              | b)                    | Plavuše su bile najbrojnije među bivšim djevojkama                        |



|     |    |  |
|-----|----|--|
|     |    | pomenutog glumca. Stoga je plava boja kose dominantna vrijednost.  |
| 3.  | c) | Među pacijentima je najviše bilo onih sa dijagnozom sumanutog poremećaja i poremećaja ličnosti. Kako ih je bilo po pet, imamo dvije dominantne vrijednosti (moda).   |
| 4.  | c) | Vrstama naseljenih mjesta pripisani su rangovi od 1 do 6. Medijana pada između ranga 3 i 4. Prosjek ovih rangova je 3.5, što je, stoga, ujedno medijana podataka.  |
| 5.  | d) | Mod i medijana računaju se za rezultate s intervalne skale čija distribucija statistički značajno odstupa od normalne. Pored toga, ove vrijednosti se mogu računati za rezultate sa ordinalne skale mjerenja. Mod je karakterističan i za nominalne skale.   |
| 6.  | a) | Primijenite obrasce za računanje medijane i aritmetičke sredine!   |
| 7.  | b) | Izbacite rezultate 4 i 29, te izračunajte aritmetičku sredinu preostalih podataka. Ova dva podatka su 25% od osam navedenih podataka.  |
| 8.  | a) | 100 je najčešća vrijednost, pa je to mod datog niza podataka. Preostale dvije mjere centralne tendencije računaju se po njihovim obrascima.  |
| 9.  | c) | U trećoj kategoriji ("C") je najviše stanovnika, pa je to dominantna vrijednost pet navedenih kategorija materijalnog statusa. Varijabilnost kategorijalne/nominalne varijable jednaka je broju njenih kategorija. Dakle, varijabilnost je 5.  |
| 10. | d) | Maksimalni rezultat ( $X_{\max} = 7$ ) minus minimalni rezultat ( $X_{\min} = 1$ ) daje raspon od 6.   |
| 11. | a) | Zajednička aritmetička sredina je: $(15 \cdot 125 + 30 \cdot 115) / (15 + 30) = 118.33$ .  |
| 12. | b) | Prilikom računanja standardne devijacije na nekom uzorku, broj ispitanika se umanjuje za 1 ( $N - 1$ ).  |
| 13. | a) | Za razliku od prethodnog zadatka, prilikom računanja standardne devijacije na cijeloj populaciji, broj ispitanika u pripadajućoj formuli se <i>ne</i> umanjuje za 1.   |
| 14. | c) | Kako svi studenti postižu isti rezultat ( $X = 100$ ), nema varijabilnosti u njihovim skorovima. Dakle, standardna devijacija je nula. Aritmetička sredina je, razumljivo, $M = 100$ .   |
| 15. | b) | Rezultati se poredaju po veličini (od najnižeg do najvišeg). Rezultat koji je u sredini (šesti rezultat) pada u 50. centil. S lijeve strane ostaje pet brojeva, a isto toliko ostaje i s desne strane medijane. Rezultat u sredini skupa rezultata s lijeve strane (treći po redu) pada u 25. centil, a rezultat u sredini skupa s desne strane pada u 75. centil. |
| 16. | a) | Medijana se poklapa sa rezultatom koji pada u 50. centil, a interkvartilni raspon je, jednostavno, razlika između rezultata koji pada u 75. i onog koji pada u 25. centil.   |
| 17. | c) | Između 75. centila i 25. centila je $75\% - 25\% = 50\%$ centralnih rezultata.   |
| 18. | d) | Rezultat koji pada u 50. centil jednak je medijani ( $C$ ).  |
| 19. | e) | Pozitivno asimetrična distribucija ima "rep" s desne strane. Njen "brijeg" je slijeva, a njegov vrh je dominantna vrijednost.  |

|     |    |  |
|-----|----|--|
|     |    | Nakon nje (slijeva-nadesno) slijedi medijana, te aritmetička sredina. Zato vrijedi $D > C > M$ .   |
| 20. | a) | Kod negativne asimetrije, "rep" distribucije je lijevo, a "brijeg" desno. "Pod brijegom" je više rezultata nego "ispod repa".  |
| 21. | e) | Varijansa je po definiciji kvadrirana standardna devijacija.   |
| 22. | a) | S porastom uzorka, povećava se imenilac (nazivnik) razlomka u okviru obrasca za računanje standardne devijacije. Kada neki broj ili izraz dijelimo sve većim brojem, dobijeni rezultat se smanjuje. Zato ovdje standardna devijacija opada.  |
| 23. | b) | Posljednji rezultat je ekstremna vrijednost (eng. <i>outlier</i> ), koja se izbacuje, pa se aritmetička sredina računa na osnovu ostalih podataka. Alternativa je računanje medijane na osnovu svih podataka, jer ova mjera centralne tendencije nije osjetljiva na ekstremne vrijednosti. |
| 24. | a) | Polazimo od obrasca $M = \sum X_i / N$ . $M$ je 32, a $N$ je 10. Potom, u brojiocu (brojniku) razlomka saberemo sve poznate rezultate, dodajući $X_5$ . Ovdje su nam poznati $M$ , $N$ i svi drugi rezultati, pa se jednačina svodi na jednu nepoznatu ( $X_5$ ).                          |
| 25. | e) | Raspon je $R = X_{max} - X_{min} = 14 - (-10) = 24$  |
| 26. | c) | Konstanta ne varira, dakle, raspon njenih vrijednosti (koje su sve iste) je nula. I njena standardna devijacija je, jasno, nula.   |
| 27. | d) | Podaci se, jednostavno, uvrste u formulu za izračunavanje koeficijenta varijacije.   |

### Z-VRIJEDNOSTI I NORMALNA KRIVA

| Red. br. | Tačno rješenje | Objašnjenje   |
|----------|----------------|---|
| 1.       | a)             | Z-vrijednost aritmetičke sredine je uvijek 0.   |
| 2.       | b)             | Standardna devijacija standardizovane normalne krive je uvijek 1.   |
| 3.       | b)             | Od $M = 22$ oduzmete dvije standardne devijacije (jer je $z = -2$ ). Standardna devijacija je $SD = 5$ , dakle, $22 - 2 * 5 = 12$ .                                       |
| 4.       | b)             | Rezultat $X = 43$ ima z-vrijednost od -1 jer je za jednu standardnu devijaciju ( $SD = 7$ ) niži od aritmetičke sredine ( $M = 50$ ).                                     |
| 5.       | e)             | U ovom rasponu je 99.73% rezultata.   |
| 6.       | c)             | 50% je rezultata nižih od aritmetičke sredine, a 50% viših od nje, budući da je normalna kriva simetrična.  |
| 7.       | d)             | Prvi rezultat je tri standardne devijacije ispod, a drugi isto toliko standardnih devijacija iznad aritmetičke sredine.   |
| 8.       | d)             | U navedenom rasponu je 68.26% rezultata.  |
| 9.       | b)             | Kako se z-vrijednosti ova dva rezultata razlikuju samo po predznaku, oni su podjednako udaljeni od aritmetičke sredine (jedan je ispodprosječan, a drugi iznadprosječan). |
| 10.      | c)             | Z-vrijednost od 4 znači da je pripadajući rezultat za čak četiri standardne devijacije viši od prosječnog.  |
| 11.      | e)             | Od $z = -1$ do $z = 0$ nalazi se pola od 68.26% rezultata. To je 34.13%. Od $z = 0$ do $z = 2$ nalazi se pola od 95.44% rezultata.  |

|     |    |  |
|-----|----|--|
|     |    | To je 47.72%. Dakle, između $z = -1$ i $z = 2$ nalazi se ukupno 81.85% svih rezultata.   |
| 12. | b) | Riječ je o rezultatu čija je $z$ -vrijednost $= (24 - 15)/4.5 = 2$ . Unutar raspona od $z = -2$ do $z = 2$ poznato nam je da se nalazi 95.44% rezultata. Ispod i iznad ovih $z$ -vrijednosti je, jasno, $100\% - 95.44\% = 4.56\%$ rezultata. Kako je normalna kriva simetrična, jednak je broj rezultata ispod $z = -2$ i iznad $z = 2$ . Zato je tačan odgovor $4.56\%/2 = 2.28\%$ . |

### STANDARDNA GREŠKA I INTERVALI POUZDANOSTI ARITMETIČKE SREDINE

| Red. br. | Tačno rješenje | Objašnjenje  |
|----------|----------------|--|
| 1.       | c)             | Jednostavno uvrstite vrijednosti u obrazac za računanje standardne greške aritmetičke sredine.   |
| 2.       | c)             | Od aritmetičke sredine oduzmite njenu standardnu grešku (i dobijate donju granicu intervala pouzdanosti). Slično, dodajte joj standardnu grešku, kako biste dobili gornju granicu naznačenog intervala u kome se nalazi prava aritmetička sredina. |
| 3.       | c)             | Ponovite postupak rješavanja zadatka 2, s tim što aritmetičkoj sredini oduzimate i dodajete dvije vrijednosti standardne greške (dakle, $2*s_M$ ).   |
| 4.       | b)             | Iz obrasca za njeno računanje, primjećuje se da standardna greška aritmetičke sredine s povećanjem uzorka $X$ puta opada korijen od $X$ puta. U postavci zadatka imamo dupli uzorak $\Rightarrow$ ova pogreška se smanji korijen iz dva puta.      |
| 5.       | c)             | Iz obrasca za računanje $s_M$ jasno je da ako standardna devijacija poraste $Y$ puta, toliko puta poraste i $s_M$ . Dakle, tri puta veća standardna devijacija $\Rightarrow$ tri puta veća $s_M$ .   |
| 6.       | a)             | Iz obrasca za računanje $s_M$ slijedi $s = s_M * \sqrt{N}$ .   |
| 7.       | e)             | 95% interval pouzdanosti je $M \pm 1.96*s_M$ , u našem slučaju, $9 \pm 1.96*0.25$ .  |
| 8.       | e)             | Od aritmetičke sredine oduzmite $2.58*s_M$ za dobijanje donje granice 99%-nog intervala, a dodajte joj $2.58*s_M$ kako biste dobili gornju granicu naznačenog intervala pouzdanosti.   |

### T-TESTOVI

| Red. br. | Tačno rješenje | Objašnjenje   |
|----------|----------------|---|
| 1.       | b)             | Kada se jedinica uvrsti u nazivnik (imenilac) razlomka u $t$ -test obrascu, ostaje samo brojnik (brojilac) koji je jednak $M_1 - M_2$ .   |
| 2.       | d)             | Kako su u pitanju dvije grupe ispitanika (mlađi i stariji), znamo da se koristi $t$ -test za nezavisne uzorke. Uzorak je veći od 30, dakle, riječ je o $t$ -testu za velike nezavisne uzorke. |
| 3.       | c)             | Stepeni slobode u $t$ -testu za velike nezavisne uzorke su $df = N - 2 \Rightarrow df = 230 - 2 = 228$ .  |
| 4.       | e)             | Uvrstite podatke u obrazac za $t$ -test za velike nezavisne uzorke i  |

|     |    |   |
|-----|----|---|
|     |    | odrediti je li $t$ -omjer statistički značajan uz pomoć dostupnih/odgovarajućih tabela.   |
| 5.  | d) | Kako smo dobili statistički značajnu razliku u zadatku br. 4, vidimo i da je aritmetička sredina rezultata muškaraca veća od aritmetičke sredine žena. Otuda ovaj odgovor.  |
| 6.  | c) | Ista grupa je bila izložena treningu, dakle, riječ je o zavisnim uzorcima. U uzorku je 108 osoba, pa ovakav uzorak spada u velike. Odgovor je, jasno, $t$ -test za velike zavisne uzorke.   |
| 7.  | c) | Kod $t$ -testa za zavisne uzorke, stepeni slobode su: $df = N - 1$ , dakle, $108 - 1 = 107$ .   |
| 8.  | a) | Podatke uvrstite u obrazac za računanje $t$ -omjera za velike zavisne uzorke. Na kraju se dobija statistički značajan rezultat (uz $df = 81$ ).   |
| 9.  | a) | Kako je aritmetička sredina procjena studenata niža u drugom mjerenju (nakon treće godine) u odnosu na prvo mjerenje (na početku prve godine fakulteta), te kako je ova razlika statistički značajna, odgovor je razumljiv sam po sebi.   |
| 10. | c) | Povećanjem standardne greške razlike aritmetičkih sredina raste i nazivnik (imenilac) $t$ -omjera. Posljedično, vrijednost razlomka opada.  |
| 11. | e) | Kako neku aritmetičku sredinu izračunatu na uzorku upoređujemo s populacijskom, koristimo $t$ -test za jedan uzorak.  |
| 12. | a) | Uvrštavanjem datih vrijednosti u obrazac računanja $t$ -testa za jedan uzorak dolazimo do vrijednosti $t$ -omjera za koju se ispostavlja da je statistički značajna bar na nivou $p < .01$ .  |
| 13. | b) | Imamo dvije grupe ispitanika, u prvoj je devet, a u drugoj 11 ispitanika. Dakle, riječ je o $t$ -testu za male nezavisne uzorke.  |
| 14. | d) | $df = N - 2 = 20 - 2 = 18$  |
| 15. | a) | $M_1 - M_2 = 7.44 - 5.64 = 1.80$  |
| 16. | a) | Nakon računanja aritmetičkih sredina i standardnih devijacija, podaci se uvrste u obrazac za $t$ -test za male nezavisne uzorke, te provjeri statistička značajnost $t$ -omjera. Ispostavlja se da je $t$ -omjer statistički značajan, a pošto je aritmetička sredina grupe koja zarađuje više od 700 eura mjesečno veća, onda ova razlika ide u prilog pomenute grupe. |
| 17. | d) | Ista grupa ispitana je prije i poslije radionice. Grupa je mala, pa je riječ o $t$ -testu za male zavisne uzorke.   |
| 18. | a) | $df = N - 1 = 12 - 1 = 11$  |
| 19. | d) | $T$ -test, izračunat metodom diferencije, iznosi 5.77 (ako se rezultati iz prve tačke mjerenja oduzimaju od odgovarajućih rezultata iz druge tačke mjerenja). U suprotnom (ako se rezultati iz druge tačke mjerenja oduzimaju od rezultata iz prve tačke mjerenja), $t$ -test je -5.77.   |
| 20. | c) | Izračunati $t$ -omjer je statistički značajan. Možete primijetiti i da je aritmetička sredina rezultata ispitanika nakon radionice veća od one izračunate prije radionice. Dakle, stavovi su, nakon radionice, statistički značajno pozitivniji nego prije.   |

## ANALIZA VARIJANSE (ANOVA)

| Red. br. | Tačno rješenje | Objašnjenje   |
|----------|----------------|---|
| 1.       | d)             | Radni status ima tri nivoa, pol dva (muški i ženski) i mjesto stanovanja tri. Dakle, nacrt (dizajn) je $3 \times 2 \times 3$ .  |
| 2.       | c)             | Imamo tri nezavisne varijable (nivo obrazovanja, dobna grupa i materijalni status), a jednu zavisnu varijablu (mentalno zdravlje). Stoga je odgovor trosmjerna (trofaktorijalna) univarijatna ANOVA.  |
| 3.       | d)             | $MS_A = SS_A / df_A = 25 / 1 = 25$  |
| 4.       | b)             | Stepeni slobode za faktor A su 1. Iz formule $df_A = a - 1$ (gdje je $a$ broj nivoa datog faktora) slijedi $a = df_A + 1 = 1 + 1$ . Dakle, faktor A ima dva nivoa.  |
| 5.       | e)             | Poznate vrijednosti uvrstite u sljedeći obrazac: $F_A = MS_A / MS_{wg}$ .   |
| 6.       | e)             | U indeksu simbola za $F$ -omjer je naznačeno da se ovaj omjer odnosi na interakciju faktora $P \times Q$ . Kako je $p > .05$ , efekat interakcije pomenutih faktora nije statistički značajan.  |
| 7.       | b)             | Prvi broj u zagradi označava stepene slobode između grupa. Oni su jednaki broju grupa ( $k$ ) minus 1 ( $4 = k - 1$ ). Zato u našem zadatku imamo $k = 5$ grupa. Kako je $p < .05$ , rezultat je statistički značajan. Naravno, ne mora značiti da se svaka grupa statistički značajno razlikuje od svake, već da postoji ovakva razlika između bar dvije grupe. Za utvrđivanje koje se sve grupe međusobno razlikuju (statistički značajno), koriste se <i>post-hoc</i> testovi.                     |
| 8.       | b)             | $SS_{A \times B} = SS_{bg} - SS_A - SS_B$   |
| 9.       | a)             | $F = MS_{tr} / MS_{rez} = 1$ , što znači da su varijanse tretmana i reziduala jednake. Dakle, efekat tretmana nije statistički značajan.  |
| 10.      | a)             | Ako je $F$ -omjer manji od jedan, onda je $MS_{wg}$ veće od $MS_{bg}$ . Kako bi rezultat bio statistički značajan, cilj je dobiti što veće $MS_{bg}$ u odnosu na $MS_{wg}$ . Stoga $F < 1$ upućuje na statistički neznačajan rezultat.  |
| 11.      | a)             | Kako je nezavisna varijabla (faktor) zanimanje, a zavisna percipirani stres na radnom mjestu (poslu), riječ je o jednosmjernoj univarijatnoj ANOVA-i.   |
| 12.      | c)             | $df_{bg} = k - 1$ , gdje je $k = 3$ , jer imamo tri grupe zanimanja   |
| 13.      | c)             | $df_{tot} = N - 1$ ( $N$ – ukupan broj ispitanika koji je u našem slučaju 30)   |
| 14.      | d)             | Provedete proceduru računanja za jednosmjernu univarijatnu ANOVA-u i dođete do $F$ -omjera.   |
| 15.      | b)             | Nekim od <i>post hoc</i> testova (npr. <i>Scheffeovim</i> testom) dolazite do statistički značajnih razlika između socijalnih radnika i vaspitačica, te između policajaca i vaspitačica. Suprotno, socijalni radnici i policajci se međusobno ne razlikuju statistički značajno po percipiranom stresu na poslu. Iz podataka (ili računanjem aritmetičke sredine rezultata za svaku grupu) vidi se da socijalni radnici i policajci procjenjuju svoj stres na poslu kao viši u odnosu na vaspitačice. |

16.

| Izvor variranja   | SS         | df         | MS           | F            | p               |
|-------------------|------------|------------|--------------|--------------|-----------------|
| Između ispitanika | 63         | <b>36</b>  |              |              |                 |
| Tretman           | 89         | <b>3</b>   | <b>29.67</b> | <b>10.37</b> | <b>&lt; .05</b> |
| Rezidual          | 309        | <b>108</b> | <b>2.86</b>  |              |                 |
| Unutar ispitanika | <b>398</b> | <b>111</b> |              |              |                 |
| Total             | <b>461</b> | <b>147</b> |              |              |                 |

*Napomena:*  $df_{tr} = t - 1$  ( $t$  – broj tretmana, u našem slučaju mjerenja);  $df_{rez} = (N-1)(t-1)$ ;  $df_{bs} = N - 1$ . Ostali dio tabele rješava se uz pomoć dekompozicije varijabiliteta u ANOVA-i za ponovljena mjerenja, te obrazaca za računanje prosječnih kvadrata (*MS*) i *F*-omjera. Na kraju se uz pomoć odgovarajuće tablice provjeri značajnost *F*-omjera.

17. *b)* Dobili smo statistički značajan *F*-omjer, ali ne znamo između kojih tačno i koliko vremenskih tačaka mjerenja ima statistički značajne razlike.

18. *c)* Dvije su nezavisne varijable (faktora) sa po tri nivoa, dakle, nacrt je 3 x 3.

19.

| Izvor variranja | SS         | df        | MS          | F           | p              |
|-----------------|------------|-----------|-------------|-------------|----------------|
| Unutar grupa    | 185        | 54        | <b>3.43</b> |             |                |
| NZV1            | 16         | <b>2</b>  | <b>8</b>    | <b>2.33</b> | <b>&gt;.05</b> |
| NZV2            | 18         | <b>2</b>  | <b>9</b>    | <b>2.62</b> | <b>&gt;.05</b> |
| NZV1 x NZV2     | <b>8</b>   | <b>4</b>  | <b>2</b>    | <b>0.58</b> | <b>&gt;.05</b> |
| Između grupa    | 42         | <b>8</b>  | <b>5.25</b> |             |                |
| Total           | <b>227</b> | <b>62</b> |             |             |                |

20. *c)* Iz posljednje kolone tabele koju je trebalo popuniti u zadatku br. 19 jasno je da ni glavni efekti ni interakcija nisu statistički značajni (u svim slučajevima je  $p > .05$ ).

21. *b)*  $df_{tot} = N - 1 = 62$ ,  $N = 63$ .

## KORELACIJA

| Red. br. | Tačno rješenje | Objašnjenje  |
|----------|----------------|--|
| 1.       | d)             | Vrijedi reciprociitet: korelacija varijable <i>A</i> sa varijablom <i>B</i> je jednaka korelaciji varijable <i>B</i> sa varijablom <i>A</i> .  |
| 2.       | c)             | Korelacija varijable sa samom sobom je uvijek maksimalna i pozitivna, dakle, $r = 1$ .   |
| 3.       | c)             | Korelacija varijable sa konstantom je uvijek 0, jer konstanta <i>ne</i> varira.  |
| 4.       | a)             | Rezultati konstante su isti, tj. svi ispitanici postižu isti rezultat. S druge strane, rezultati varijable, logično, variraju. Grafički prikaz je prava linija (npr. vertikalna ako jedini rezultat konstante prikažemo na <i>x</i> -osi, a rezultate varijable na <i>y</i> -osi). |
| 5.       | a)             | Veća raspršenost tačaka na skater-dijagramu podrazumijeva  |

|     |    |  |
|-----|----|--|
|     |    | slabije sukladno variranje, tj. nižu korelaciju.   |
| 6.  | c) | Korelacija nije isto što i uzročno-posljedični odnos između varijabli. Pozitivna korelacija je rast vrijednosti jedne varijable praćen rastom u vrijednostima druge varijable. Takođe, smanjenje vrijednosti jedne varijable praćeno je smanjenjem vrijednosti druge u slučaju pozitivne korelacije. |
| 7.  | b) | Prikazani koeficijent korelacije je blizu vrijednosti $r = -.50$ . Dakle, po predznaku je negativan, a po veličini osrednji (umjerena korelacija).   |
| 8.  | d) | Maksimalna negativna korelacija je $r = -1$ . Ovoj vrijednosti najbliži je koeficijent $r = -.82$ .  |
| 9.  | c) | Koeficijenti su jednaki po svojoj apsolutnoj vrijednosti, odnosno, po veličini. Razlikuju se samo po predznaku.  |
| 10. | a) | Po jednom od obrazaca za računanje koeficijenta korelacije, korelacija je standardizovana kovarijansa.   |
| 11. | a) | Koeficijent determinacije jednak je kvadratu koeficijenta korelacije. Dakle, $r^2 = .4^2 = .16$ .  |
| 12. | d) | Zajednička varijansa, izražena u procentima, računa se tako što se koeficijent determinacije pomnoži sa 100 i doda znak za procenat (%): $.7^2 * 100 \Rightarrow 49\%$   |
| 13. | e) | Procenat neobjašnjene varijanse dobija se oduzimanjem procenta objašnjene varijanse od 100%: $100\% - 9\% = 91\%$ .  |
| 14. | d) | U slučaju korelacije, stepeni slobode se računaju po sljedećem obrascu: $df = N - 2$ .   |
| 15. | a) | Korelacija je statistički značajna jer je $p < .01$ , stepeni slobode su u zagradi (odmah iza oznake za koeficijent korelacije) i iznose 116, a riječ je o umjerenoj pozitivnoj korelaciji (jer je njena vrijednost $r = .41$ ).   |
| 16. | b) | Korelacija je niska i pozitivna, a iz podatka $p > .05$ vidimo da je i statistički neznačajna. Ako korelacija nije statistički značajna na nivou .05, nije značajna ni na nivou .01 (zato odgovor <i>e</i> ) nije tačan).  |
| 17. | b) | Uvrstite podatke u obrazac za računanje Pearsonovog koeficijenta korelacije.   |
| 18. | a) | Značajnost ovog koeficijenta je $p = .029$ , dakle, značajan je na nivou .05, ali nije na nivou .01.   |

## REGRESIJSKA ANALIZA

| Red. br. | Tačno rješenje | Objašnjenje  |
|----------|----------------|--|
| 1.       | a)             | Prediktorska varijabla ili prediktor, koja je pandan nezavisnoj varijabli u eksperimentalnim istraživanjima.   |
| 2.       | d)             | Jednaki su pripadajućim koeficijentima korelacije jer to slijedi iz obrasca za računanje beta-pondera (kada uvrsti njihova međusobna korelacija za koju je u postavci zadatka rečeno da je nula). Uvrstite, pa se uvjerite i sami. |
| 3.       | c)             | Kvadriranjem koeficijenata korelacije i sabiranjem ovako dobijenih rezultata dolazite do koeficijenta multiple   |

|     |    |   |
|-----|----|---|
|     |    | determinacije: $R^2 = .4^2 + .5^2 = .41$  |
| 4.  | d) | Koeficijent multiple korelacije je korijen koeficijenta multiple determinacije, dakle $R = \sqrt{.41} = .64$  |
| 5.  | d) | Koeficijent alijenacije je, ustvari, neobjašnjena varijansa. Od jedinice (ukupne varijanse) oduzme se koeficijent multiple determinacije: $1 - .41 = .59$ .   |
| 6.  | b) | Najvjerovatniji rezultat, izražen u obliku z-vrijednosti, je: $z_y' = \beta_1 * z_1 + \beta_2 * z_2 = -0.5 * .252 + 1.75 * .465 = 0.69$   |
| 7.  | e) | Vrijednosti beta-pondera, kao i koeficijenata korelacije, kreću se u rasponu od -1 do 1.  |
| 8.  | a) | Uvrstite koeficijente korelacije u obrazac za računanje beta-pondera, posebno za ekstraverziju i posebno za neuroticizam.   |
| 9.  | b) | Primjenom obrasca za računanje koeficijenta multiple determinacije dobija se tačno rješenje.  |
| 10. | d) | Korijen koeficijenta multiple determinacije jednak je koeficijentu multiple korelacije.   |
| 11. | e) | Po obrascu $b = \beta * s_y/s_x$ vrši se pretvaranje standardizovanih ( $\beta$ ) u nestandardizovane regresijske vrijednosti ( $b$ ).  |
| 12. | b) | Samo uvrstite rezultate u regresijsku jednačinu za dva prediktora.  |
| 13. | a) | Uvrstite svaki rezultat posebno u regresijsku jednačinu za jedan prediktor (jednačina u kojoj su aritmetičke sredine, standardne devijacije i korelacija između prediktora i kriterijuma).  |
| 14. | e) | Uvrstite poznate vrijednosti u formulu za računanje standardne greške prognoze.   |
| 15. | d) | U formuli za standardnu grešku prognoze, umjesto standardne devijacije kriterija iz zadataka br. 13. i 14., uvrstite standardnu devijaciju prediktora.  |
| 16. | d) | Izračunajte odgovarajuće vrijednosti na osnovu podataka u tabeli i uvrstite ih u obrazac za Pearsonov koeficijent korelacije.   |
| 17. | a) | Odredite odsječak na y-osi i nagib regresijske crte, te napišite regresijsku jednačinu.   |
| 18. | e) | $13.43 + 0.697(x + 10) = 13.43 + 0.697x + 0.697 * 10 = y' + 6.97$ . Dakle, prognozirani rezultat se povećava za 6.97.   |
| 19. | d) | Uvrstite uzorak i standardnu devijaciju radne učinkovitosti u obrazac za računanje standardne greške prognoze.  |
| 20. | e) | Uvrstite rezultat $X = 20$ u jednačinu regresije dobijenu u zadatku br. 17.   |
| 21. | e) | Tačke na skater-dijagramu su usko grupisane oko prave linije čiji je pravac od donjeg lijevog do gornjeg desnog ugla. Dakle, korelacija je veoma visoka i pozitivna.  |
| 22. | e) | Svaka tačka na grafikonu predstavlja uređen par rezultata svakog od ispitanika. Jedan rezultat je na prvoj, a drugi na drugoj varijabli (očitaljavu se sa x i y-ose grafikona). Izračunajte vrijednosti potrebne za uvrštavanje u obrazac za računanje Pearsonovog koeficijenta korelacije. |



## NEKI NEPARAMETRIJSKI STATISTIČKI POSTUPCI

| Red. br. | Tačno rješenje | Objašnjenje   |
|----------|----------------|---|
| 1.       | d)             | Očekivane frekvencije su po postavci zadatka jednake. Dobijaju se kada se ukupni broj ispitanika (u zadatku: djece) podijeli brojem vrsta igrački. Dakle, $N/k = 110/5 = 22$  |
| 2.       | d)             | Stepeni slobode se dobijaju oduzimanjem jedinice od ukupnog broja kategorija varijable. Dakle, $df = k - 1 = 5 - 1 = 4$   |
| 3.       | e)             | Opažene i očekivane frekvencije se uvrste u obrazac za računanje hi-kvadrat testa i dobije se vrijednost za koju se ispostavi (npr. korištenjem odgovarajuće tablice za značajnost hi-kvadrat testa) da je statistički značajna.  |
| 4.       | c)             | $df = k - 1 = 3 - 1 = 2$  |
| 5.       | b)             | Očekivanja anketarske organizacije uvrste se kao očekivane frekvencije, a stvarni rezultati izbora (glasanja) kao opažene (empirijske) frekvencije. Kasnije se utvrdi da vrijednost hi-kvadrat testa nije statistički značajna.   |
| 6.       | a)             | Kako nije bilo statistički značajne razlike između očekivanih i stvarnih rezultata, možemo reći da je anketarska organizacija dobro procijenila ("predvidjela") rezultate. Dakle, uradila je dobar posao.   |
| 7.       | a)             | Očekivana frekvencija za dato polje tabele se dobija kada se pomnože sume reda i kolone u kojoj se to polje nalazi i podijele ukupnim brojem ispitanika. Dakle, $127 \cdot 129 / 259 = 63.25$   |
| 8.       | a)             | Stepeni slobode se u slučaju tabele kontingencije dobijaju po obrascu: $df = (m - 1)(n - 1)$ , gdje je $m$ broj redova, a $n$ broj kolona ove tabele. Zapravo, to je broj nivoa svake od kategorijalnih varijabli. U našem slučaju je: $df = (2 - 1) \cdot (2 - 1) = 1$ .   |
| 9.       | b)             | Standardni statistički zapis za rezultat hi-kvadrat testa obuhvata oznaku za hi-kvadrat statistik, broj stepeni slobode i uzorak u zagradi, te statističku značajnost ( $p$ ).  |
| 10.      | e)             | Među muškarcima, najviše ih je bilo za odlazak iz zemlje, a najviše žena je bilo protiv odlaska iz svoje države. Rezultat je i statistički značajan, otuda je odgovor na ovo pitanje samorazumljiv.   |
| 11.      | b)             | Podaci se organizuju (transformišu) tako da se najvišoj vrijednosti pripiše rang 16 (jer je ukupan broj učenika 16), te se smanjuje rang kako se smanjuju rezultati ispitanika. Ako su dva rezultata jednaka i trebala bi biti npr. 10. i 9. rang, svakom od njih se pripisuje rang 9.5. Nakon toga, rangovi se sumiraju, posebno za prvu i posebno za drugu grupu učenika. Ovi rezultati se, uz broj ispitanika po grupama, unesu u formulu za računanje $U$ -statistika. Kako imamo dvije grupe, računaju se dva $U$ -statistika. |
| 12.      | d)             | Iz zadatka br. 11 u formulu za računanje $z$ -vrijednosti se uvrsti manja dobijena vrijednost $U$ -statistika. U istu formulu uvrštavamo i "novi" broj ispitanika koji smo dobili u ovom zadatku.   |

|     |    |   |
|-----|----|---|
| 13. | a) | Izračunaju se razlike u ocjenama između ova dva predmeta, pa se uvrste u obrazac za računanje Spearmanovog koeficijenta korelacije. Iz odgovarajuće tabele ili na neki drugi način (u okviru nekog statističkog programa/softvera) utvrdi se da je ovaj koeficijent statistički značajan. |
| 14. | a) | Kada se podaci pretvore u rangove (npr. rang 1 se pripíše najboljem rezultatu), dobijamo istu vrijednost Spearmanovog koeficijenta koju smo dobili u zadatku br. 13.  |

### OSNOVE TEORIJE VJEROVATNOĆE I KOMBINATORIKE

| Red. br. | Tačno rješenje | Objašnjenje  |
|----------|----------------|--|
| 1.       | a)             | Bacanjem kocke, vjerovatnoća dobijanja jednog broja iz skupa $\{1, \dots, 6\}$ je $1/6$ , dakle, jednaka je za svaki broj iz navedenog skupa.  |
| 2.       | e)             | Kako su ishodi dva bacanja kocke nezavisni događaji, njihova vjerovatnoća je $1/6 * 1/6 = 1/36$  |
| 3.       | e)             | Vrijedi isto kao za podatke u zadatku br. 2.   |
| 4.       | e)             | Ukupna populacija je 100%. Broj nepismenih je 8.5%, dakle, broj pismenih je $100\% - 8.5\% = 91.5\%$   |
| 5.       | e)             | Vjerovatnoća dobijanja jednog od brojeva na kocki u jednom bacanju je $1/6$ , a vjerovatnoća dobijanja "pisma" bacanjem novčića $1/2$ . Ista je vjerovatnoća dobijanja "glave". Kako je riječ o nezavisnim događajima, izračun je sljedeći: $1/6 * 1/2 * 1/2 = 1/24$ .   |
| 6.       | a)             | Rješenje je isto kao u zadatku br. 5, jer ne zavisi od redoslijeda matematičkih eksperimenata (drugim riječima, opet je posrijedi nezavisnost događaja).   |
| 7.       | b)             | Parni brojevi na kocki su 2, 4 i 6 i oni čine 50% mogućih brojeva koji se mogu dobiti bacanjem kocke. 50% je $1/2$ , što je vjerovatnoća nezavisna od toga ko i kojim redoslijedom baca kocku.   |
| 8.       | c)             | U punom špil, vjerovatnoća izvlačenja određene karte je $1/52$ . Kada se izvuče neka karta, u špil ih ostane 51. Dakle, vjerovatnoća izvlačenja neke karte je u ovom slučaju $1/51$ . Kada se izvlači treći put (a izvučene karte u prethodnim izvlačenjima nisu vraćene u špil), vjerovatnoća je $1/50$ . Dakle, vjerovatnoća koja se traži u zadatku je $1/52 * 1/51 * 1/50$ . |
| 9.       | e)             | Ako je već ušao jedan Francuz, ispred vrata ostaje osmoro ljudi: dva Italijana, dva Francuza i četiri Engleza. Englezi su 50% ljudi koji su ostali ispred vrata. Dakle, vjerovatnoća da sljedeći uđe Englez je $1/2$ .   |
| 10.      | b)             | Kako su događaji $D$ i $F$ nezavisni, nastupanje jednog nije zavisno (uslovljeno) nastupanjem drugog. Dakle, svaki ima svoju vjerovatnoću pojavljivanja. Tako je vjerovatnoća događaja $D$ : $P(D) = 1/7$ .  |
| 11.      | d)             | Ovdje je riječ o matematičkom očekivanju (tj. o očekivanoj vrijednosti). Vjerovatnoća da je pogođen tačan odgovor za   |

|     |    |  |
|-----|----|--|
|     |    | jedno pitanje je $1/4$ jer imamo četiri ponuđena odgovora (a učenik bira jedan od njih). Kako je 20 pitanja, očekivani broj bodova za učenika koji samo pogađa odgovore je $20 * 1/4 = 5$ .  |
| 12. | b) | Rezultat svake "runde" ruleta je nezavisan događaj, dakle, šanse da se kuglica zaustavi na crvenoj, odnosno crnoj boji su iste. Kako još imamo polje "0" ili dato polje i polje "00", vjerovatnoća da kuglica ostane na crvenom polju nije 50% (ili $1/2$ ), već nešto manja od toga (tačnije, $18/37$ ili $18/38$ , što je .49, odnosno .47). |
| 13. | d) | Kako je na raspolaganju pet slova koja se ne ponavljaju i kako sastavljate riječi od po pet slova, riječ je o permutacijama bez ponavljanja: $5! = 5*4*3*2*1 = 120$  |
| 14. | b) | Na raspolaganju je šest elemenata skupa, od kojih se dva elementa ponavljaju (dvije jedinice i tri trojke). Dakle, $6!/(2!3!1!) = 60$  |
| 15. | a) | Imate po nekoliko primjeraka knjiga iz tri oblasti ( $n$ je, stoga, tri). Izvlačite po dvije knjige ( $k = 2$ ). Elementi se ponavljaju, a skup koji se izvlači je manji od ukupnog skupa elemenata. Dakle, riječ je o varijacijama s ponavljanjem. Zato nakon uvrštavanja u odgovarajuću formulu dobijamo $3^2 = 9$ .                         |
| 16. | e) | Kombinacije s ponavljanjem, jer imamo ponovljene elemente (npr. četiri knjige iz ekonomije), izvučeni broj elemenata manji je od njihovog ukupnog broja, a poredak nije bitan.   |
| 17. | e) | Uvrštavanjem u obrazac za kombijacije s ponavljanjem dobijamo: $4!/(2!2!) = 6$ .   |

## LITERATURA

- American Psychological Association (2010). *Publication manual of the American Psychological Association* (6. izd.). Washington, DC: APA.
- Boslaugh, S. i Watters, P. A. (2008). *Statistics in a nutshell: A desktop quick reference*. Cambridge: O'Reilly.
- Dancey, C. i Reidy, J. (2011). *Statistics without maths for psychology* (5. izd.). New York: Pearson.
- Evans, J. (1996). *Straightforward statistics for the behavioral sciences*. Pacific Grove, CA: Brooks/Cole Publishing.
- Freund, R. i Wilson, W. (2003). *Statistical methods* (2. izd.). New York: Academic Press.
- Goodwin, J. (2010). *Research in psychology: Methods and design* (6. izd.). Hoboken, NJ: Wiley.
- Howell, D. (2010). *Statistical methods for psychology* (7. izd.). Belmont, CA: Cengage Wadsworth.
- Mordkoff, T. (2016). *The assumption(s) of normality* (<http://www2.psychology.uiowa.edu/faculty/mordkoff/GradStats/part%201/I.07%20normal.pdf>).
- Petz, B. (2007). *Osnovne statističke metode za nematematičare* (6. izd.). Jastrebarsko: Naklada Slap.

## (Za)bilješka o autoru



Rođen sam 26. maja 1987. godine u Cetinju (Crna Gora). Osnovnu i srednju školu završio sam u Kotoru, a magistrirao psihologiju na Filozofskom fakultetu u Sarajevu 2011. godine. Usljed nezadovoljavajuće finansijske situacije, nisam uspio nastaviti svoje formalno obrazovanje i usavršavanje.

Autor sam pet knjiga i preko 20 naučnih radova. Pored toga, pišem poeziju i kratke priče. Pišem i tekstove popularnog karaktera, koji imaju za cilj pomoći u nadvladavanju stresa, anksioznosti, depresije, osjećaja niskog samopoštovanja i ostalih psihosocijalnih poteškoća.

Uspješno sam okončao preko 70 kurseva putem različitih Internet-platforni koje omogućavaju učenje na daljinu. Kursevi su pokrivali oblasti kao što su: matematika, fizika, medicina, psihologija, antropologija, sociologija, ekonomija...

Glavne oblasti mojeg interesovanja su: psihologija ličnosti, statistika u društvenim naukama, psihopatologija, forenzička psihologija i matematičko obrazovanje.

Nezaposlen sam i nisam član nijedne političke stranke. Nadam se da će se prvo promijeniti, a drugo ostati kakvo i jeste.