



NAJNOVIJA 2020/2021 god **skripta iz statistike za pismeni**

autor: Časlav Pejdić

(teorija potrebna za zadatke i primeri ispitnih rokova)



SA NAMA MOŽETE SPREMITI ISPIT IZ:

- MATEMATIKE
- OSNOVA EKONOMIJE
- STATISTIKE
- FINANSIJSKOG RAČUNOVODSTVA
- TEORIJE CENA
- UPRAVLJAČKOG RAČUNOVODSTVA
- MODELA
- MATEMATIKE 2
- TEORIJSKE STATISTIKE
- LSE (MATEMATIKE I STATISTIKE)
- RAČUNOVODSTVA TROŠKOVA
- FINANSIJSKE MATEMATIKE
- ENGLESKOG

22 godine sa vama, 50 m od fakulteta,
iskusni predavači, kvalitetne skripte,
klimatizovane učionice, **male grupe (do 8 polaznika)**
BESPLATNO SAVETOVALIŠTE

www.smartbasic.edu.rs, smartbasic@gmail.com

Lomina 5, Beograd, **064/123-09-10**

grupa na fejsu: www.facebook.com/groups/statistikaekof/
online prijava na: smartbasic.edu.rs/online-prijava/

KOPIRNICA " MINA" Gavрила Principa 29

0. Uvod

Da bi uspešno razumeli tekstove zadataka moramo se u svakoj lekciji upoznati sa osnovnim pojmovima vezanim za tu lekciju. Što se tiče uvoda bitno je shvatiti kako u zadacima razlikovati da li je u pitanju skup ili uzorak. Zbog toga evo definicije skupa i uzorka:

Osnovni skup se sastoji od svih elemenata ili jedinica posmatranja čije karakteristike ispitujuemo – pojedina, stvari ili predmeta.

Deo osnovnog skupa koji je izabran u svrhe statističke analize naziva se uzorkom.

Dakle, ako u zadatku kažu da su „slučajnim putem izabrali ili anketirali 100 studenata Ekonomskog fakulteta“, sigurno da je u pitanju uzorak, dok ako se kaže da su „ispitivanjem svih 3000 studenata Ekonomskog fakulteta dobijeni sledeći rezultati ...“, onda je u pitanju skup. Uglavnom fraza svi, sve, ... ukazuje da je to skup, a ako je samo deo te celine onda je to uzorak.

Svaki skup ili uzorak sastoji se iz jedinica posmatranja ili elemenata. U gornjem primeru to je student.

A da bi jedinice skupa posmatrali biramo jednu njihovu osobinu i to se zove obeležje ili promenljiva. Na primer ako posmatramo skup studenata Ekonomskog fakulteta jedno njihovo obeležje je broj položenih ispita. To je promenljiva jer se menja od jedne do druge jedinice posmatranja. Vrednost promenljive koja se odnosi na jednu jedinicu naziva se opservacijom ili podatkom.

Promenljive se inače dele na:

1. Kvantitativne ili numeričke
 - prekidne (diskretne)
 - neprekidne

2. Kvalitativne ili kategorijske

Kvantitativna promenljiva je promenljiva koja se može brojčano izraziti. Prekidna je ako može da uzme samo cele vrednosti (npr. Broj golova na utakmici, broj dece, broj osnovnih škola u nekom mestu, ...), a neprekidna ako može da uzme bilo koju vrednost iz nekog intervala (npr. Cena, vreme za rešavanje nekog zadatka, visina, težina, ...).

Kvalitativna promenljiva ne može numerički da se iskaže ali može da se klasifikuje u dve ili više kategorija (npr. pol, marka automobila, boja očiju, ...).

I na kraju ove uvodne priče definisaćemo još pojam serije podataka.

Seriya podataka je skup podataka koji se odnosi na jednu ili više promenljivih.

Serije se dele na:

1. Strukturne serije
2. Vremenske serije

Vremenske serije se dele na momentne i intervalne. Kod momentnih (zalihe, broj zaposlenih, štednja, ...) godišnji podatak se poklapa sa poslednjim kvartalom ili mesecom, a kod intervalnih godišnji podatak se dobija sabiranjem kvartalnih ili mesečnih podataka.

strukturne serije $\rightarrow$

x	f

 x – vrednosti obeležja (opservacije)
 f – frekvencije (učestanosti) javljanja

vremenske serije $\rightarrow$

<i>God.</i>	y

 y – vrednosti obeležja

U knjizi nema priče o mernim skalama, a pominju se na predavanjima i na vežbama, pa evo ukratko priče o tome.

Postoje 4 merne skale: nominalna, ordinalna, intervalna i skala odnosa.

Nominalna se koristi za kvalitativna obeležja koja ne mogu da se rangiraju, kao na primer pol(muško, žensko), lokacija za odmor (more, planina, banja), vrsta automobila (audi, bmw, opel, ...).

Ordinalna se koristi za kvalitativna obeležja koja mogu da se rangiraju, kao na primer stepen stručne spreme (osnovna, srednja, viša, visoka), uspeh u srednjoj školi (odličan, vrlo dobar, dobar, dovoljan).

Intervalna se koristi za kvantitativna obeležja za koja ne može da se utvrdi odnos i kod njih pojava postoji i pri vrednosti nula (temperatura, godina rođenja, godina venčanja, geografska širina,...).

Skala odnosa se koristi za kvantitativna obeležja za koja može da se utvrdi odnos i kod njih nula znači da pojave nema (težina, visina, plata, ...).

NAPOMENA: Ako su samo dva stanja (muško/žensko, zdrav/bolestan, udoban/neudoban, ...) onda je uvek nominalna skala.

1. Sređivanje i grafičko prikazivanje podataka

Vrste raspodela:

1. Raspodela frekvencija
2. Raspodela relativnih frekvencija i procentualna raspodela
3. Raspodela kumulativnih frekvencija
4. Raspodela kumulativnih relativnih frekvencija i kumulativnog učešća

Grafički prikazi:

1. Štapićasti dijagram
2. Strukturni krug
3. Histogram
4. Poligon
5. Kriva frekvencija
6. Prikaz u obliku stabljike i lista
7. Tačkasti dijagrami
8. Aritmetički dijagram
9. Polulogaritamski dijagram

2. Numeričke deskriptivne mere

1. Mere centralne tendencije

- a) Aritmetička sredina
- b) Medijana
- c) Modus

Podrazumevana mera centralne tendencije je aritmetička sredina. Izuzetak je ako je raspored znatno asimetričan jer se onda koristi medijana (jer je neosetljiva na ekstremne vrednosti). Ako je obeležje na nominalnoj skali od mera centralne tendencije može da se izračuna samo modus, a ako je obeležje na ordinalnoj skali mogu da se izračunaju samo modus i medijana. Pozicione mere su modus i medijana, a izračunata je aritmetička sredina.

Aritmetička sredina

$$\mu = \frac{\sum x_i}{N} \quad \text{- skup, negrupisani podaci}$$

$$\mu = \frac{\sum f_i x_i}{N} \quad \text{- skup, grupisani podaci}$$

$$\bar{x} = \frac{\sum x_i}{n} \quad \text{- uzorak, negrupisani podaci}$$

$$\bar{x} = \frac{\sum f_i x_i}{n} \quad \text{- uzorak, grupisani podaci}$$

μ - aritmetička sredina skupa

—

$\bar{x}$ - aritmetička sredina uzorka

N - veličina skupa

n - veličina uzorka

Kod grupisanih podataka važi: $N = \sum f_i$,

a ako je u pitanju uzorak $n = \sum f_i$.

Medijana

Medijana je jednaka vrednosti središnjeg člana serije podataka rangirane po rastućem redosledu.

Na primer:

$$5, 9, 11, 14, 3 \rightarrow 3, 5, 9, 11, 14 \rightarrow M_e = 9.$$

Neparan broj podataka

$$3, 5, 9, 11, 14, 15 \rightarrow M_e = \frac{9+11}{2} = 10.$$

Paran broj podataka

Kod grupisanih podataka medijana je onaj podatak čija kumulativna frekvencija je prva veća od $N/2$. Izuzetak je ako postoji podatak čija kumulativna frekvencija je tačno $N/2$. U tom slučaju medijana je aritmetička sredina tog i sledećeg podatka.

Modus

Modus (M_O) je vrednost koja se javlja sa najvećom frekvencijom u seriji podataka..

$$2 \ 3 \ 3 \ 4 \ 5 \ 9 \rightarrow M_O = 3$$

$$2 \ 3 \ 3 \ 4 \ 4 \ 5 \ 9 \rightarrow M_{O1} = 3, M_{O2} = 4.$$

$$2 \ 2 \ 3 \ 3 \ 4 \ 4 \ 5 \ 5 \rightarrow \text{nema } M_O.$$

2. Mere disperzije

- a) Interval varijacije
- b) Varijansa
- c) Standardna devijacija
- d) Koeficijent varijacije
- e) Interkvartilna razlika
- f) Standardizovano odstupanje

Podrazumevana mera disperzije je standardna devijacija. Izuzeci su:

1. Ako poredimo dve serije sa različitim aritmetičkim sredinama onda je podrazumevana mera koeficijent varijacije,
2. Ako je raspored znatno asimetričan računa se interkvartilna razlika.

Kod kvalitativnih obeležja ne mogu da se računaju mere disperzije.

Kad u zadatku kažu apsolutna disperzija misli se na standardnu devijaciju, relativna znači da je u pitanju koeficijent varijacije, a pozicione mere su interval varijacije i interkvartilna razlika. Inače, apsolutne mere disperzije su: interval varijacije, varijansa, standardna devijacija, interkvartilna razlika, a relativne su: koeficijent varijacije i standardizovano odstupanje.

Interval varijacije

$$i = x_{\max} - x_{\min}$$

Varijansa

$$\sigma^2 = \frac{\sum (x - \mu)^2}{N} = \frac{\sum x^2 - \frac{(\sum x)^2}{N}}{N} \quad \text{- skup, negrupisani}$$

$$\sigma^2 = \frac{\sum f(x - \mu)^2}{N} = \frac{\sum fx^2 - \frac{(\sum fx)^2}{N}}{N} \quad \text{- skup, grupisani}$$

$$s^2 = \frac{\sum (x - \bar{x})^2}{n-1} = \frac{\sum x^2 - \frac{(\sum x)^2}{n}}{n-1} \quad \text{- uzorak, negrupisani}$$

$$s^2 = \frac{\sum f(x - \bar{x})^2}{n-1} = \frac{\sum fx^2 - \frac{(\sum fx)^2}{n}}{n-1} \quad \text{- uzorak, grupisani}$$

Varijansa pokazuje prosek kvadrata odstupanja podataka od aritmetičke sredine. Najveći nedostatak varijanse je što je iskazana u jedinici mere na kvadrat.

Standardna devijacija

$$\sigma = \sqrt{\sigma^2} \text{ - skup} \quad s = \sqrt{s^2} \text{ - uzorak}$$

Standardna devijacija pokazuje prosek odstupanja podataka od aritmetičke sredine.

Koeficijent varijacije

$$CV = \frac{\sigma}{\mu} \cdot 100 \text{ - skup} \quad CV = \frac{s}{\bar{x}} \cdot 100 \text{ - uzorak}$$

Koeficijent varijacije pokazuje koliki je prosek odstupanja podataka od aritmetičke sredine izraženo u procentima aritmetičke sredine.

Kvartili

Kvartili su vrednosti obeležja koji seriju uređenu po rastućem redosledu dele na 4 jednaka dela.

Na primer:

$$2,4,6,7,8,9,12,15 \Rightarrow Q_1 = \frac{4+6}{2} = 5, Q_2 = M_e = \frac{7+8}{2} = 7,5, Q_3 = \frac{9+12}{2} = 10,5.$$

$$2,4,6,7,8,9,12 \Rightarrow Q_1 = 4, Q_2 = M_e = 7, Q_3 = 9$$

$$2,4,6,7,8,9,12,15,16 \Rightarrow Q_1 = \frac{4+6}{2} = 5, Q_2 = M_e = 8, Q_3 = \frac{12+15}{2} = 13,5$$

$$2,4,6,7,8,9,12,15,16,20 \Rightarrow Q_1 = 6, Q_2 = M_e = \frac{8+9}{2} = 8,5, Q_3 = 15$$

Kod grupisanih podataka Q_1 je onaj podatak čija kumulativna frekvencija je prva veća od $N/4$.

Kod grupisanih podataka Q_3 je onaj podatak čija kumulativna frekvencija je prva veća od $3N/4$.

Interkvartilna razlika

$$IQR = Q_3 - Q_1$$

Standardizovano odstupanje

$$Z = \frac{X - \mu}{\sigma} \text{ - skup}$$

$$Z = \frac{X - \bar{x}}{s} \text{ - uzorak}$$

Z je jedina mera disperzije koja se ne odnosi na celu seriju, već samo na 1 njen podatak. Z pokazuje koliko taj podatak odstupa od aritmetičke sredine, izraženo u standardnim devijacijama.

3. Mere oblika rasporeda

a) Mere asimetrije

Postoje tri načina za utvrđivanje simetričnosti rasporeda:

1) Pomoću srednjih vrednosti

Izračunamo aritmetičku sredinu, modus i medijanu i postoje 3 slučaja:

- 1) $\bar{x} = M_e = M_o \Rightarrow$ raspored je simetričan
- 2) $\bar{x} < M_e < M_o \Rightarrow$ raspored je asimetričan ulevo (negativno asimetričan)
- 3) $M_o < M_e < \bar{x} \Rightarrow$ raspored je asimetričan udesno (pozitivno asimetričan)

2) Pomoću koeficijenta asimetrije

α_3 - koeficijent asimetrije

$$\alpha_3 = \frac{M_3}{s^3}, \text{ gde je } M_3 = \frac{\sum f(x - \bar{x})^3}{n} - \text{treći centralni moment.}$$

Postoji 5 slučajeva:

- 1) $\alpha_3 = 0 \Rightarrow$ raspored je simetričan
- 2) $0 < \alpha_3 < 0,5 \Rightarrow$ umerena pozitivna asimetrija
- 3) $\alpha_3 \geq 0,5 \Rightarrow$ znatna pozitivna asimetrija
- 4) $-0,5 < \alpha_3 < 0 \Rightarrow$ umerena negativna asimetrija
- 5) $\alpha_3 \leq -0,5 \Rightarrow$ znatna negativna asimetrija

3) Pomoću boks-plot dijagrama

Videćemo na kursu kako se crta boks-plot dijagram i na osnovu njega tumači asimetrija.

b) Mere spljoštenosti

Pomoću koeficijenta spljoštenosti

α_4 - koeficijent spljoštenosti

$$\alpha_4 = \frac{M_4}{s^4}, \text{ gde je } M_4 = \frac{\sum f(x - \bar{x})^4}{n} - \text{četvrti centralni moment.}$$

Postoje 3 slučaja:

- 1) $\alpha_4 = 3 \Rightarrow$ normalna spljoštenost
- 2) $\alpha_4 > 3 \Rightarrow$ spljoštenost manja od normalne
- 3) $\alpha_4 < 3 \Rightarrow$ spljoštenost veća od normalne

3. Verovatnoća

Ne dolazi na pismenom i najvećim delom izbačeno i sa usmenog.
Na kolokvijumu često dolazi, pa stoga evo par najvažnijih teorijskih postavki iz ove lekcije.

Eksperiment je proces, čiji rezultat izvođenja je jedna i samo jedna od mnogih opservacija. Te opservacije se nazivaju ishodi eksperimenta. Skup svih ishoda jednog eksperimenta se naziva prostor uzorka.

Tko npr. Bacanje kocke jednom je eksperiment. Ishodi su 1,2,3,4,5,6, a prostor uzorka je $S = \{1,2,3,4,5,6\}$.

Događaj je skup jednog ili više ishoda eksperimenta. Ako sadrži samo jedan ishod onda je u pitanju elementarni događaj i on se najčešće obeležava sa E_i . Ako sadrži više od jednog ishoda onda je u pitanju složen događaj.

Tako npr. ako je događaj A-pad parnog broja, onda je to složen događaj, jer mu odgovaraju 3 ishoda: $A = \{2,4,6\}$, a događaj B – pad broja manjeg od 2 je prost događaj, jer mu odgovara samo 1 ishod: $B = \{1\}$.

Verovatnoća je numerička mera mogućnosti realizacije određenog događaja.

Za verovatnoću važi:

$$P(E_i) = \frac{1}{\text{Ukupan broj ishoda eksperimenta}}$$

$$P(A) = \frac{\text{Broj povoljnih ishoda za A}}{\text{Ukupan broj ishoda eksperimenta}}$$

$$0 \leq P \leq 1$$

Događaji koji ne mogu da se realizuju istovremeno zovu se međusobno isključivi događaji.

Ako realizacija jednog događaja ne utiče na verovatnoću realizacije drugog onda su događaji nezavisni. Kod njih važi:

$$P(A|B) = P(A) \quad \text{ i } \quad P(B|A) = P(B).$$

Događaj koji sadrži sve ishode eksperimenta koje ne sadrži A se naziva A komplement.

Presek događaja A i B sadrži sve ishode koji su zajednički za A i B. Obeležava se sa AB.

Ako su događaji nezavisni važi: $P(AB) = P(A) \cdot P(B)$.

Ako su događaji isključivi važi: $P(AB) = 0$.

Unija događaja A i B je skup svih ishoda koji se nalazi ili u A ili u B. Obeležava se sa $A \cup B$.

$$P(A \cup B) = P(A) + P(B) - P(AB).$$

Ako su događaji isključivi važi: $P(A \cup B) = P(A) + P(B)$.

4. Diskretne slučajne promenljive i njihove raspodele

Slučajna promenljiva je promenljiva čija je vrednost određena ishodom slučajnog eksperimenta (npr: X – broj grbova pri bacanju 3 novčića; Y – broj poena na ispitu; Z – zbir na dve kocke; W – vreme potrebno za rešavanje nekog problema). Slučajne promenljive se dele na diskretne i neprekidne. Diskretne mogu da uzmu samo izolovane cele vrednosti (X, Y, Z), a neprekidne bilo koju vrednost iz nekog intervala (W).

A. Raspored verovatnoće diskretne slučajne promenljive

X – broj grbova pri bacanju 3 novčića

x	$P(x)$	x – vrednosti prekidne slučajne promenljive
0	1/8	$P(x)$ – verovatnoće javljanja tih vrednosti
1	3/8	$\sum P(x) = 1$
2	3/8	
3	1/8	
Σ	1	

$$P = \frac{\text{broj povoljnih ishoda}}{\text{ukupan broj ishoda}}$$

P P P $\rightarrow$ 0 grbova

P P G $\rightarrow$ 1 grb

P G P $\rightarrow$ 1 grb

P G G $\rightarrow$ 2 grba

G P P $\rightarrow$ 1 grb

G P G $\rightarrow$ 2 grba

G G P $\rightarrow$ 2 grba

G G G $\rightarrow$ 3 grba

$$P(X=2) = 3/8$$

$$P(X>2) = P(X=3) = 1/8$$

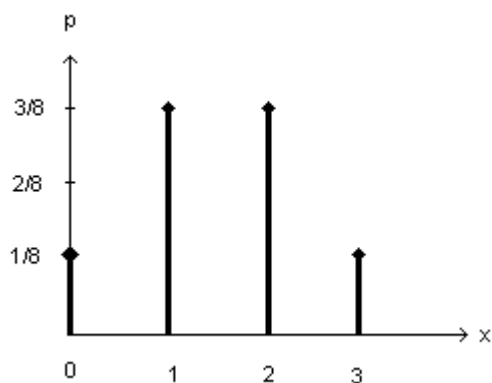
$$P(X<2) = P(X=0) + P(X=1) = 4/8$$

$$P(X<7) = P(X=0) + P(X=1) + P(X=2) + P(X=3) = 1$$

$$P(1<X<2) = 0$$

$$P(1<X\leq 2) = P(X=2) = 3/8$$

Grafičko predstavljanje je uvek pomoću štapićastog dijagrama.



B. Očekivana vrednost diskretne slučajne promenljive

$$E(X) = \mu = \sum xP(x)$$

x	$P(x)$	$xP(x)$
0	1/8	0
1	3/8	3/8
2	3/8	6/8
3	1/8	3/8
Σ	1	1,5

$$E(X) = 1,5$$

Očekivana vrednost nije ono što mi očekujemo u statističkom eksperimentu, već predstavlja prosečnu vrednost neke slučajne promenljive ili populacije.

C. Standardna devijacija diskretne slučajne promenljive

$$\sigma = \sqrt{\sum x^2 P(x) - \mu^2}$$

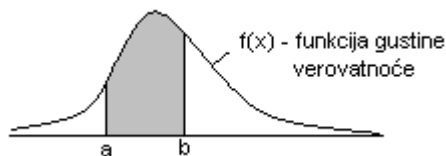
x	$P(x)$	$xP(x)$	$x^2 P(x)$
0	1/8	0	0
1	3/8	3/8	3/8
2	3/8	6/8	12/8
3	1/8	3/8	9/8
Σ	1	1,5	3

$$\sum x^2 P(x) = 3$$

$$\sigma = 0,866$$

5. Nепrekidne slučajne promenljive i normalna raspodela

A. Raspored verovatnoće nепrekidne sluč. promenljive



Površina koju funkcija gustine verovatnoće zaklapa sa x-osom je uvek jednaka 1.

$$P(X = a) = 0$$

$$P(X \leq a) = \int_{-\infty}^a f(x)dx = F(a)$$

$$P(X \geq a) = \int_a^{\infty} f(x)dx = 1 - F(a)$$

$$P(a \leq X \leq b) = \int_a^b f(x)dx = F(b) - F(a)$$

B. Normalna raspodela

Normalan raspored je raspored čija funkcija verovatnoće glasi:

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

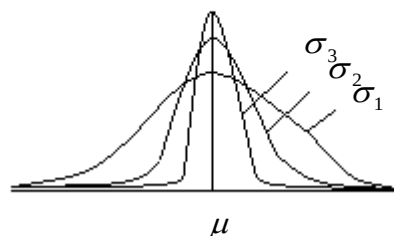
Poslednja relacija se skraćeno zapisuje: $X : N(\mu; \sigma^2)$

Grafički prikaz normalnog rasporeda:



$$\begin{aligned} \alpha_3 &= 0 \\ \alpha_4 &= 3 \end{aligned}$$

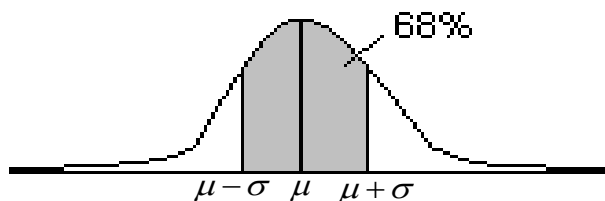
Parametar μ pokazuje samo gde je x-koordinata vrha rasporeda, a parametar σ^2 određuje oblik rasporeda. Što je σ veće to je raspored više spljošten. Primer tri normalna rasporeda sa istom aritmetičkom sredinom, ali različitim varijansama:



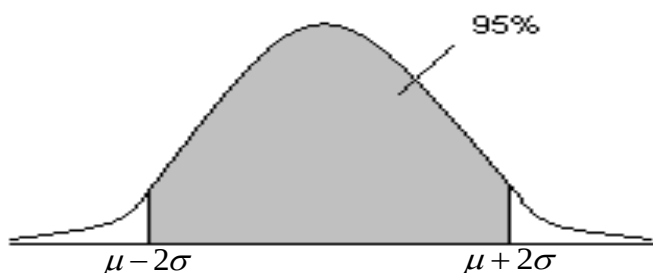
$$\sigma_1 > \sigma_2 > \sigma_3$$

Vezano za normalan raspored bitno je zapamtiti sledeće 3 osobine:

1. 68% normalnog rasporeda se nalazi na rastojanju do jedne standardne devijacije od aritmetičke sredine:

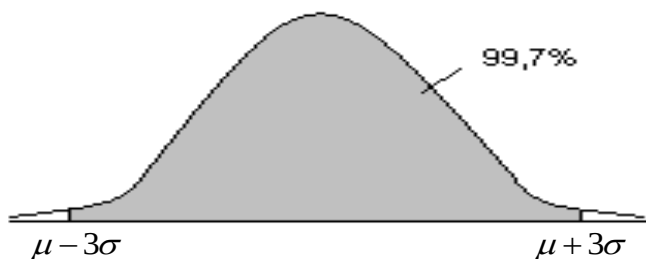


2. 95% normalnog rasporeda se nalazi na rastojanju do dve standardne devijacije od aritmetičke sredine:



Vrednosti na rastojanju manjem od dve devijacije od proseka se zovu veoma verovatne vrednosti, a vrednosti na većem odstojanju malo verovatne vrednosti.

3. 99,7 % normalnog rasporeda se nalazi na rastojanju do tri standardne devijacije od aritmetičke sredine:



Kod normalnog rasporeda važno je upamtiti još sledeće osobine:

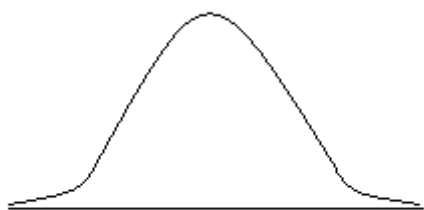
Ako su $X_1 : N(\mu_1; \sigma_1^2)$ i $X_2 : N(\mu_2; \sigma_2^2)$ tada važi:

1. $X_1 + X_2 : N(\mu_1 + \mu_2; \sigma_1^2 + \sigma_2^2)$

2. $X_1 - X_2 : N(\mu_1 - \mu_2; \sigma_1^2 + \sigma_2^2)$

3. $n \cdot X_1 : N(n\mu_1; n^2\sigma_1^2)$

C. Standardizovan normalan raspored ($Z:N(0,1)$)



1. $P(Z \leq a) = F(a)$
2. $P(Z \geq a) = 1 - F(a)$
3. $P(a \leq Z \leq b) = F(b) - F(a)$
4. $P(-a \leq Z \leq a) = 2F(a) - 1$

Tablica 4

$$Z = \frac{X - \mu}{\sigma} \text{ - konverzija normalnog u standardizovani normalan raspored.}$$

D. Primena normalnog rasporeda na računanje proporcije

$$p = \frac{N_1}{N}$$

$$\hat{p} = \frac{X}{n}$$

p – proporcija u skupu
 N_1 – broj „uspeha“ u skupu
 N – veličina skupa
 $\hat{p}$ – proporcija u uzorku
 X – broj „uspeha“ u uzorku
 n – veličina uzorka

Proporcija predstavlja verovatnoću da je slučajno izabrana jedinica „uspeh“.

6. Uzoračka raspodela

Zadatke iz ove oblasti prepoznavamo po tome što se pitanje uvek odnosi na uzorak, a parametri skupa su uglavnom poznati.

A. Raspored aritmetičkih sredina uzoraka ($\bar{X}$)

$$P(\bar{X} \geq a), P(\bar{X} \leq a), P(a \leq \bar{X} \leq b)$$

$\bar{X}$ ima normalan raspored ako je ispunjen jedan od sledeća dva uslova:

1. Osnovni skup je normalno raspoređen

2. $n \geq 30$ (centralna granična teorema)

$$\bar{X} : N(\mu_{\bar{X}}; \sigma_{\bar{X}}^2)$$

$$\mu_{\bar{X}} = \mu$$

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N-1}}$$

Izraz $\sqrt{\frac{N-n}{N-1}}$ se zove popravni faktor za konačne skupove

(pfks) i koristi se u formuli samo ako su ispunjena sledeća 3 uslova:

1. uzorak je bez ponavljanja
2. osnovni skup je konačan (zna se N)
3. stopa izbora je veća od 5% ($\frac{n}{N} > 0,05$)

NAPOMENA: Uslov broj 1 se podrazumeva da je uvek ispunjen osim ako nije drugačije naglašeno.

$\sigma_{\bar{X}}^2$ – varijansa rasporeda aritmetičkih sredina i pokazuje prosek kvadrata odstupanja aritmetičkih sredina uzoraka od aritmetičke sredine skupa.

$\sigma_{\bar{X}}$ – standardna greška rasporeda aritmetičkih sredina i pokazuje prosek odstupanja aritmetičkih sredina uzoraka od aritmetičke sredine skupa.

Ponekad u zadacima $\sigma_{\bar{X}}$ se računa po sledećoj formuli: $\sigma_{\bar{X}} = \sqrt{\sum \bar{x}^2 P(\bar{x}) - (\mu_{\bar{x}})^2}$, pogledati u knjizi na strani 327.

B. Raspored proporcija uzoraka ($\hat{P}$)

$$P(\hat{P} \geq a), P(\hat{P} \leq a), P(a \leq \hat{P} \leq b)$$

Da bi $\hat{P}$ imao normalan raspored moraju biti ispunjeni sledeći uslovi:

1. $n \cdot p > 5$

2. $n \cdot q > 5$

$$\hat{P} : N(\mu_{\hat{P}}; \sigma_{\hat{P}}^2)$$

$$\mu_{\hat{P}} = p$$

$$\sigma_{\hat{P}} = \sqrt{\frac{p \cdot q}{n}} \cdot \sqrt{\frac{N-n}{N-1}}$$

$\sigma_{\hat{P}}$ – standardna greška ocene proporcija i pokazuje prosek odstupanja proporcija uzoraka od proporcije skupa.

Šabloni

1. Da se $\bar{X}$ razlikuje od μ za manje od 2

$$P\left(\left|\bar{X} - \mu\right| < 2\right)$$

2. Da se $\bar{X}$ razlikuje od μ za više od 2

$$P\left(\left|\bar{X} - \mu\right| > 2\right)$$

3. Da je $\bar{X}$ veći od μ za manje od 2

$$P(\bar{X} - \mu < 2)$$

4. Da je $\bar{X}$ veći od μ za više od 2

$$P(\bar{X} - \mu > +2)$$

5. Da je $\bar{X}$ manji od μ za manje od 2

$$P(\bar{X} - \mu > -2)$$

6. Da je $\bar{X}$ manji od μ za više od 2

$$P(\bar{X} - \mu < -2)$$

7. Ocenjivanje aritmetičke sredine i proporcije

Statističko zaključivanje je postupak donošenja zaključaka o parametrima osnovnog skupa na osnovu informacija dobijenih iz uzorka. Postoje dva vida statističkog zaključivanja:

1. Statističko ocenjivanje
2. Testiranje statističkih hipoteza

A. Ocenjivanje aritmetičke sredine osnovnog skupa

1. σ je poznata

$$\bar{x} - z_{\alpha/2} \cdot \sigma_{\bar{x}} \leq \mu \leq \bar{x} + z_{\alpha/2} \cdot \sigma_{\bar{x}}$$

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N-1}}$$

$\bar{x}$ – aritmetička sredina uzorka
 $\sigma_{\bar{x}}$ – standardna greška ocene aritmetičke sredine

α – rizik greške (nivo značajnosti)

$1 - \alpha$ – nivo pouzdanosti (koeficijent poverenja)

$\alpha = 5\%$ – podrazumevana vrednost

Interval poverenja možemo smanjiti na dva načina:

1. Povećanjem rizika greške α ,
2. Povećanjem veličine uzorka n .

Uži interval znači preciznija ocena.

$$F(z_{\alpha/2}) = 1 - \frac{\alpha}{2}$$

Na primer:

$$\alpha = 5\% \Rightarrow F(z_{\alpha/2}) = 1 - \frac{0,05}{2} = 0,9750 \Rightarrow z_{\alpha/2} = 1,96.$$

Uslovi:

1. σ je poznata
2. a) osnovni skup je normalno raspoređen
 b) $n \geq 30$

2. σ nije poznata

$$\bar{x} - t_{n-1; \alpha/2} \cdot s_{\bar{x}} \leq \mu \leq \bar{x} + t_{n-1; \alpha/2} \cdot s_{\bar{x}}$$

$$s_{\bar{x}} = \frac{s}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N-1}}$$

$s_{\bar{x}}$ – ocena za $\sigma_{\bar{x}}$

t – studentov t test $\rightarrow$ tablica 5.

primer:

$$n = 16, \alpha = 0,05$$

$$t_{n-1; \alpha/2} = t_{15; 0,025} = 2,131$$

Uslovi:

1. σ nije poznata
2. a) Osnovni skup je normalno raspoređen
b) $n \geq 30$

B. Ocenjivanje proporcije osnovnog skupa

$$\hat{p} - z_{\alpha/2} s_{\hat{p}} \leq p \leq \hat{p} + z_{\alpha/2} s_{\hat{p}}$$

$$\hat{p} = \frac{X}{n} \quad \hat{p} - \text{proporcija u uzorku}$$

$$s_{\hat{p}} = \sqrt{\frac{\hat{p} \cdot \hat{q}}{n}} \cdot \sqrt{\frac{N-n}{N-1}} \quad s_{\hat{p}} - \text{ocena } \sigma_{\hat{p}}$$

Uslovi:

1. $n \cdot \hat{p} > 5$
2. $n \cdot \hat{q} > 5$

8. Testiranje hipoteza o aritmetičkoj sredini i proporciji

A. Osnovni pojmovi

H_0 : nulta hipoteza (nevin)

H_1 : alternativna hipoteza (kriv)

Ono što ispitujemo stavljamo u H_1 , a negaciju toga u H_0 . Pokušavamo da odbacimo H_0 i ako uspemo prihvatamo H_1 kao tačno, dok u suprotnom ne možemo prihvatiti H_1 .

Pri testiranju hipoteza postoje dva tipa grešaka:

1. Greška I vrste (α) – da odbacimo tačnu H_0 (osudimo nevinog),
2. Greška II vrste (β) – da ne odbacimo netačnu H_0 (pustimo krivca).

Moć (jačina) testa ($1 - \beta$) je verovatnoća da odbacimo netačnu H_0 (osudimo krivca).

Postoje 3 tipa hipoteza:

$$A) \quad H_0 : \theta = \theta_0 \quad H_1 : \theta \neq \theta_0$$

$$B) \quad H_0 : \theta \leq \theta_0 \quad H_1 : \theta > \theta_0$$

$$C) \quad H_0 : \theta \geq \theta_0 \quad H_1 : \theta < \theta_0$$

$$\theta \in \{\mu, p\}$$

Postupak testiranja hipoteza je sledeći:

I korak: postavljanje nulte i alternativne hipoteze

II korak: izbor odgovarajuće statistike testa

III korak: slika

IV korak: izračunavanje statistike testa

V korak: zaključak

B. Testiranje hipoteze o aritmetičkoj sredini osnovnog skupa

$$A) \quad H_0 : \mu = \mu_0 \quad H_1 : \mu \neq \mu_0$$

$$B) \quad H_0 : \mu \leq \mu_0 \quad H_1 : \mu > \mu_0$$

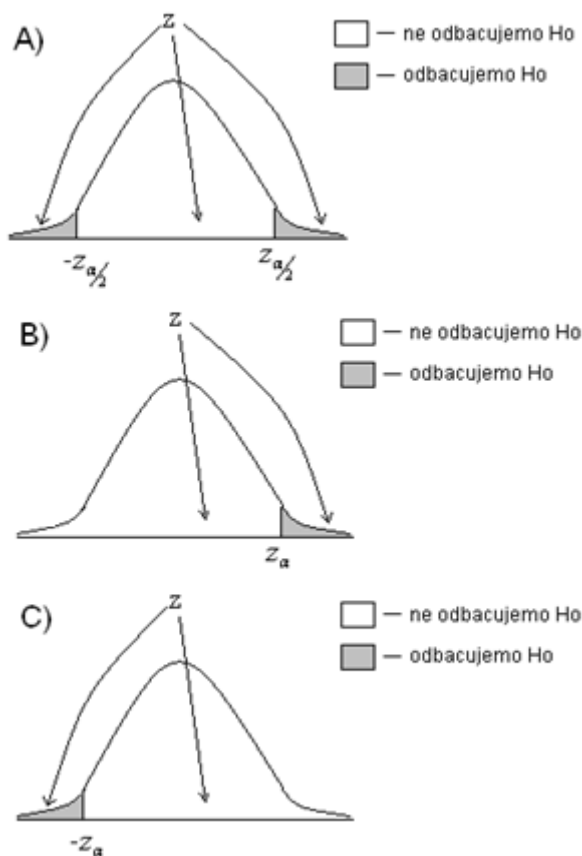
$$C) \quad H_0 : \mu \geq \mu_0 \quad H_1 : \mu < \mu_0$$

1. σ je poznata

$$z = \frac{\bar{x} - \mu_0}{\sigma_x} \quad \sigma_x = \frac{\sigma}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N-1}}$$

Uslovi:

1. σ je poznata
2. a) osnovni skup je normalno raspoređen
- b) $n \geq 30$



NAPOMENA:

Ako se testiranje radi pomoću p-vrednosti postupak je sledeći:

$$p = 2P(Z > |z|)$$

$$p = P(Z > z)$$

$$p = P(Z < z)$$

Ako je $p \leq \alpha$ odbacujemo H_0 , a ako je $p > \alpha$ ne odbacujemo H_0 .

2. σ nije poznata

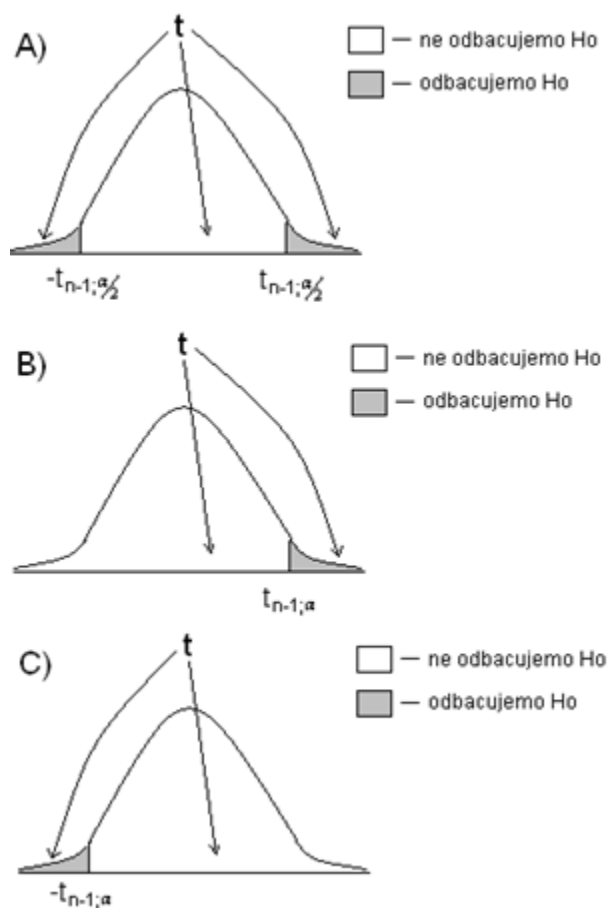
$$t = \frac{\bar{x} - \mu_0}{s_x} \quad s_x = \frac{s}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N-1}}$$

Uslovi:

1. σ nije poznata
2. a) Osnovni skup je normalno raspoređen

$$b) n \geq 30$$

$$df = n - 1 \quad df - \text{broj stepeni slobode}$$



NAPOMENA: Ako se testiranje radi pomoću p-vrednosti postupak je sledeći:

$$p = 2P(T > |t|)$$

$$p = P(T > t)$$

$$p = P(T < t)$$

Ako je $p \leq \alpha$ odbacujemo H_0 , a ako je $p > \alpha$ ne odbacujemo H_0 .

C. Testiranje proporcije skupa

$$A) \quad H_0 : p = p_0 \quad H_1 : p \neq p_0$$

$$B) \quad H_0 : p \leq p_0 \quad H_1 : p > p_0$$

$$C) \quad H_0 : p \geq p_0 \quad H_1 : p < p_0$$

$$z = \frac{\hat{p} - p_0}{\sigma_{\hat{p}}} \quad \sigma_{\hat{p}} = \sqrt{\frac{p_0 q_0}{n}} \cdot \sqrt{\frac{N-n}{N-1}}$$

Uslovi:

$$1. \quad np_0 > 5$$

$$2. \quad nq_0 > 5$$

D. Zaključivanje o razlici aritmetičkih sredina dva osnovna skupa na osnovu nezavisnih uzoraka kada su σ_1 i σ_2 poznate

TESTIRANJE

A) $H_0 : \mu_1 = \mu_2$ $H_1 : \mu_1 \neq \mu_2$

B) $H_0 : \mu_1 \leq \mu_2$ $H_1 : \mu_1 > \mu_2$

C) $H_0 : \mu_1 \geq \mu_2$ $H_1 : \mu_1 < \mu_2$

$$z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

Uslovi:

1. Uzorci su nezavisni
2. Standardne devijacije skupova su poznate
3. Oba uzorka su velika ($n_1 \geq 30, n_2 \geq 30$) ili oba skupa imaju normalne raspodele

E. Zaključivanje o razlici aritmetičkih sredina dva osnovna skupa na osnovu nezavisnih uzoraka kada su σ_1 i σ_2 nepoznate ali jednake

TESTIRANJE

A) $H_0 : \mu_1 = \mu_2$ $H_1 : \mu_1 \neq \mu_2$

B) $H_0 : \mu_1 \leq \mu_2$ $H_1 : \mu_1 > \mu_2$

C) $H_0 : \mu_1 \geq \mu_2$ $H_1 : \mu_1 < \mu_2$

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s_{\bar{x}_1 - \bar{x}_2}}$$

gde je $s_{\bar{x}_1 - \bar{x}_2} = s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$

$$s_p = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

$$df = n_1 + n_2 - 2$$

Uslovi:

1. Uzorci su nezavisni
2. Standardne devijacije skupova su nepoznate, ali jednake
3. Oba uzorka su velika ($n_1 \geq 30, n_2 \geq 30$) ili oba skupa imaju normalne raspodele

F. Zaključivanje o razlici proporcija dva osnovna skupa na osnovu velikih i nezavisnih uzoraka

TESTIRANJE

A) $H_0 : p_1 = p_2$ $H_1 : p_1 \neq p_2$

B) $H_0 : p_1 \leq p_2$ $H_1 : p_1 > p_2$

C) $H_0 : p_1 \geq p_2$ $H_1 : p_1 < p_2$

$$z = \frac{\hat{p}_1 - \hat{p}_2}{s_{\hat{p}_1 - \hat{p}_2}} \quad \text{gde je} \quad s_{\hat{p}_1 - \hat{p}_2} = \sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}}$$

Uslovi:

1. Uzorci su nezavisni
2. $n_1 \hat{p}_1 > 5$, $n_1 \hat{q}_1 > 5$, $n_2 \hat{p}_2 > 5$, $n_2 \hat{q}_2 > 5$

9. χ^2 test

A. Test prilagođenosti

H_0 : empirijski raspored odgovara teorijskom

H_1 : empirijski raspored ne odgovara teorijskom

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

O - ostvarena frekvencija (ono što smo dobili u uzorku)
E - očekivana frekvencija (šta je trebalo da dobijemo) $E=np$

Uslovi:

1. $E \geq 5$

$df = k - 1 \rightarrow$ tablica 6

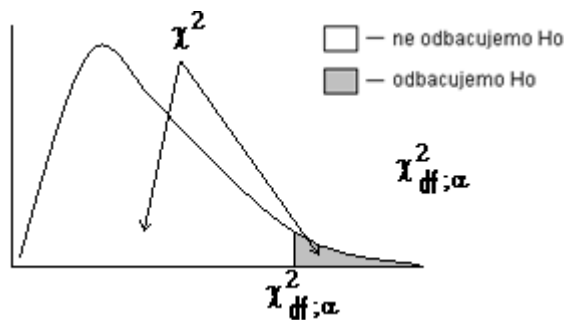
k - broj mogućih ishoda

B. Test nezavisnosti

Služi za testiranje nezavisnosti dva obeležja.

H_0 : obeležja su nezavisna

H_1 : obeležja su zavisna



Testiranje se vrši pomoću tabele kontingencije koja predstavlja kombinovanu tabelu u čijim poljima se uvek nalaze frekvencije. Ako su u postavci zadatka dati procenti treba ih pretvoriti u frekvencije.

$$E = \frac{(\text{Suma reda})(\text{Suma kolone})}{\text{vel. uzorka}}$$

$df = (R - 1)(K - 1)$

R - broj redova
K - broj kolona

10. Analiza varijanse

A. Analiza varijanse sa jednim faktorom

Služi za analizu jednakosti aritmetičkih sredina više od dva skupa.

$$H_0 : \mu_1 = \mu_2 = \mu_3 \dots \quad H_1 : \text{bar dva se razlikuju}$$

$$S_T = S_A + S_R$$

S_T – ukupan varijabilitet ili ukupna suma kvadrata ili ukupno odstupanje svih opservacija

S_A – faktorski varijabilitet ili faktorska suma kvadrata ili ukupno odstupanje između uzoraka

S_R – rezidualni varijabilitet ili rezidualna suma kvadrata ili ukupno odstupanje unutar uzoraka

$$S_A = \left(\frac{T_1^2}{n_1} + \frac{T_2^2}{n_2} + \frac{T_3^2}{n_3} + \dots \right) - \frac{(\sum x)^2}{n}$$

$$S_R = \sum x^2 - \left(\frac{T_1^2}{n_1} + \frac{T_2^2}{n_2} + \frac{T_3^2}{n_3} + \dots \right)$$

$$V_A = \frac{S_A}{k-1}$$

V_A – faktorska varijansa (prosek odstupanja između uzoraka)
k - broj uzoraka

$$V_R = \frac{S_R}{n-k}$$

V_R – rezidualna varijansa (prosek odstupanja unutar uzoraka)
n – ukupan broj podataka u svim uzorcima

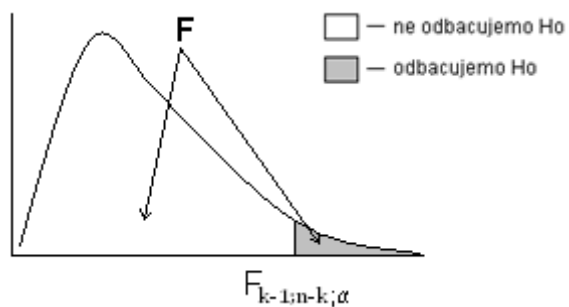
$$F = \frac{V_A}{V_R}$$

F – test (tablica 7)
 $df_1 = k-1, \quad df_2 = n-k$

IZVOR	STEPENI SLOBODE	SUMA KVADRATA	SREDNJI KVADRAT	F	p-vrednost
FAKTOR	$k-1$	S_A	V_A	V_A/V_R	
GREŠKA	$n-k$	S_R	V_R		
TOTAL	$n-1$	S_T			

Primer za traženje p-vrednosti:

$$\left. \begin{array}{l} k = 3, \\ n = 15, \\ F = 6,28 \end{array} \right\} \Rightarrow \begin{array}{l} df_1 = k - 1 = 2 \\ df_2 = n - k = 12, \\ 0,01 < p < 0,05 \end{array}$$



Uslovi:

1. Uzorci su slučajni i međusobno nezavisni
2. Osnovni skupovi su normalno raspoređeni, a njihove nepoznate varijanse su jednake

11. Prosta linearna regresija

1. Osnovni pojmovi

Cilj regresije je da se utvrdi priroda veze, odnosno oblik zavisnosti između posmatranih pojava. To se postiže pomoću odgovarajućeg regresionog modela koji pokazuje prosečno slaganje varijacija posmatranih pojava.

X – nezavisna (objašnjavajuća) slučajna promenljiva

Y – zavisna (objašnjena) promenljiva

$Y = f(X)$ $Y = \beta_0 + \beta_1 \cdot X$ - regresiona linija u skupu

β_0 – odsečak na Y osi (pokazuje kolika je prosečna vrednost Y -a kada je $X=0$)

β_1 – koeficijent nagiba (pokazuje prosečnu promenu Y -a kada se X poveća za 1)

Npr: $Y = 20 - 3X$ X – cena u 00 din, Y – tražnja u 000 t.

Pri ceni od 0 dinara očekuje se prosečna tražnja od 20000 tona, a sa svakim povećanjem cene za 100 dinara, tražnja u proseku opadne za 3000 tona.

$\hat{y} = b_0 + b_1 x$ - regresiona linija u uzorku

b_0 – ocena za β_0 , b_1 – ocena za β_1 .

Koeficijente b_0 i b_1 nalazimo po metodu najmanjih kvadrata. Metod se zasniva na minimiziranju kvadrata vertikalnih odstupanja podataka od regresione linije. Ta odstupanja se zovu reziduali (e_i).

$$e_i = y - \hat{y}$$

$$b_1 = \frac{SP_{xy}}{SK_{xx}}, \quad SP_{xy} = \sum xy - \frac{\sum x \sum y}{n}$$

$$b_0 = \bar{y} - b_1 \bar{x}, \quad SK_{xx} = \sum x^2 - \frac{(\sum x)^2}{n}$$

$$SK_{yy} = \sum y^2 - \frac{(\sum y)^2}{n}$$

U zadacima se koriste sledeći nazivi:

SK_{yy} – varijabilitet zavisne (objašnjene) promenljive ili suma kvadrata odstupanja objašnjene promenljive od aritm. sredine

SK_{xx} – varijabilitet nezavisne (objašnjavajuće) promenljive ili suma kvadrata odstupanja objašnjavajuće promenljive od aritm. sredine

s_x^2 - varijansa nezavisne (objašnjavajuće) promenljive ili kažu prosek kvadrata odstupanja objašnjavajuće promenljive od aritmetičke sredine

s_y^2 - varijansa zavisne (objašnjene) promenljive ili kažu prosek kvadrata odstupanja objašnjene promenljive od aritmetičke sredine

s_x - devijacija nezavisne (objašnjavajuće) promenljive ili kažu prosek odstupanja objašnjavajuće promenljive od aritmetičke sredine

s_y - devijacija zavisne (objašnjene) promenljive ili kažu prosek odstupanja objašnjene promenljive od aritmetičke sredine

s - standardna greška regresije ili apsolutna mera reprezentativnosti ili devijacija slučajne greške ili prosečno odstupanje objašnjene promenljive (y) od ocenjene vrednosti (regresione linije).

SKN – Suma kvadrata odstupanja objašnjene promenljive (y) od ocenjene vrednosti (regresione linije).

s^2 - varijansa regresije prosek kvadrata odstupanja objašnjene promenljive od ocenjene vrednosti (regresione linije).

Između njih postoje sledeće veze: $s_x^2 = \frac{SK_{xx}}{n-1}$; $s_y^2 = \frac{SK_{yy}}{n-1}$ $s^2 = \frac{SKN}{n-2}$.

2. Mere reprezentativnosti regresione linije

$$SKU = SKO + SKN$$

SKU – ukupna suma kvadrata (ukupan varijabilitet)

SKO – objašnjena suma kvadrata (objašnjen varijabilitet)

SKN – neobjašnjena suma kvadrata (neobjašnjen varijabilitet)

a) Apsolutna mera reprezentativnosti (s)

S – standardna greška regresije (pokazuje prosečno odstupanje podataka od regresione linije iskazano u jedinicama mere Y -a).

$$0 \leq s < \infty$$

$s = 0$ – funkcionalna veza (sve tačkice su na liniji)

Sa porastom S opada reprezentativnost.

$$s = \sqrt{\frac{SKN}{n-2}} = \sqrt{\frac{SK_{yy} - b_1 SP_{xy}}{n-2}}$$

Nedostatak je što je iskazana u jedinici mere.

b) Relativna mera reprezentativnosti (r^2)

r^2 – koeficijent determinacije (pokazuje udeo objašnjenog varijabiliteta u ukupnom)

$$r^2 = \frac{SKO}{SKU} = \frac{b_1 SP_{xy}}{SK_{yy}}$$

$$0 \leq r^2 \leq 1$$

Uslov reprezentativnosti je $r^2 > 0,5$.

$r^2 = 0 \rightarrow$ odsustvo veze

$r^2 = 1 \rightarrow$ funkcionalna veza

r^2 se izražava u procentima. Sa porastom r^2 raste reprezentativnost.

$$1 - r^2 = \frac{SKN}{SKU} - \text{udeo neobjašnjenog varijabiliteta u ukupnom.}$$

3. Testiranje značajnosti regresione linije

(Da li X utiče na Y?)

$$H_0 : \beta_1 = 0 \quad H_1 : \beta_1 \neq 0$$

$$H_0 : \beta_1 = 0 \quad H_1 : \beta_1 > 0$$

$$H_0 : \beta_1 = 0 \quad H_1 : \beta_1 < 0$$

$$t = \frac{b_1}{s_{b1}}$$

s_{b1} - standardna greška ocene koeficijenta nagiba (pokazuje za koliko u proseku b_1 odstupa od β_1)

$$s_{b1} = \frac{s}{\sqrt{SK_{xx}}}, \quad df = n - 2$$

Ukoliko se testiranje vrši pomoću p-vrednosti, onda važi: ako je $p \leq \alpha \Rightarrow X$ utiče na Y, a ako je $p > \alpha \Rightarrow X$ ne utiče na Y.

$$b_1 - t_{n-2; \alpha/2} s_{b1} \leq \beta_1 \leq b_1 + t_{n-2; \alpha/2} s_{b1}$$

Ako 0 pripada gornjem intervalu to znači da X ne utiče na Y.

Npr:

$$-3 \leq \beta_1 \leq 2, \quad x - \text{cena u 00 din}, \quad y - \text{tražnja u 000 tona}$$

Komentar: cena ne utiče na tražnju.

$$-3 \leq \beta_1 \leq -2, \quad x - \text{cena u 00 din}, \quad y - \text{tražnja u 000 tona}$$

Komentar: Pri povećanju cene za 100 dinara, tražnja u proseku opadne između 2000 tona i 3000 tona.

Kad god u zadatku pitaju za koliko se promeni Y kada se X poveća za n jedinica, treba formirati interval za β_1 pa ga pomnožiti sa n.

4. Korišćenje regresione funkcije za ocenjivanje i predviđanje Y-a

Uslovi:

- 1) Reprezentativnost ($r^2 > 0,5$)
- 2) Značajnost ($\beta_1 \neq 0$)
- 3) Da nije ekstrapolacija (ekstrapolacija je predviđanje Y-a za vrednosti X-a van opsega uzorka; dozvoljena je samo bliska ekstrapolacija)

Najuži interval kod predviđanja se dobija ako je $x_p = \bar{x}$, a najširi za vrednost x-a iz uzorka koja se najviše razlikuje od $\bar{x}$ (najveća ili najmanja vrednost uzorka).

Pri predviđanju individualne vrednosti dobijaju se širi intervali nego pri predviđanju prosečne vrednosti y-a.

U zadacima ispitujemo gornje uslove tek na kraju zadatka. Dakle i kad nisu ispunjeni radimo zadatak.

a) Ocenjivanje prosečne vrednosti Y-a

$$\hat{y}_p - t_{n-2;\alpha/2} s_{\hat{y}_p} \leq \mu_{Y/X} \leq \hat{y}_p + t_{n-2;\alpha/2} s_{\hat{y}_p}$$

$$\hat{y}_p = b_0 + b_1 x_p$$

x_p - vrednost X-a data u zadatku za koju vršimo predviđanje

$$s_{\hat{y}_p} = s \cdot \sqrt{\frac{1}{n} + \frac{(x_p - \bar{x})^2}{SK_{xx}}}$$

$s_{\hat{y}_p}$ - standardna greška ocene prosečne vrednosti Y-a.

b) Predviđanje individualne vrednosti Y-a

$$\hat{y}_p - t_{n-2;\alpha/2} s_{y_p} \leq Y_p \leq \hat{y}_p + t_{n-2;\alpha/2} s_{y_p}$$

$$s_{y_p} = s \cdot \sqrt{1 + \frac{1}{n} + \frac{(x_p - \bar{x})^2}{\sum x^2 - n\bar{x}^2}}$$

s_{y_p} - standardna greška predviđanja individualne vrednosti Y-a.

12. Prosta linearna korelacija

1. Osnovni pojmovi

Cilj korelacione analize je da se utvrdi da li između varijacija posmatranih pojava postoji kvantitativno slaganje (korelaciona veza) i ako postoji u kom stepenu.

ρ – koeficijent korelacije u skupu

r – Pirsonov koeficijent korelacije u uzorku

$r = \pm\sqrt{r^2}$ (+ se koristi kada je direktna veza tj. $b_1 > 0$, a – kada je inverzna veza tj. $b_1 < 0$).

$$r = \frac{SP_{xy}}{\sqrt{SK_{xx}SK_{yy}}}$$

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

$r = 0 \Rightarrow$ između pojava ne postoji linearna veza.

2. Testiranje korelacije

$$H_0 : \rho = 0 \quad H_1 : \rho \neq 0$$

$$H_0 : \rho = 0 \quad H_1 : \rho > 0$$

$$H_0 : \rho = 0 \quad H_1 : \rho < 0$$

$$t = r\sqrt{\frac{n-2}{1-r^2}} \quad df = n-2$$

Ako se testiranje radi pomoći p-vrednosti, onda važi: ako je $p \leq \alpha \Rightarrow$ u skupu postoji linearna veza između X i Y, a ako je $p > \alpha \Rightarrow$ u skupu ne postoji linearna veza između X i Y.

13. Višestruka linearna regresija i korelacija

A. Višestruka linearna regresija

1. Osnovni pojmovi

$Y = \beta_0 + \beta_1 \cdot X_1 + \beta_2 \cdot X_2$ - regresiona ravan u skupu

β_0 – pokazuje kolika je prosečna vrednost Y -a kada su $X_1 = 0$ i $X_2 = 0$.

β_1 – pokazuje prosečnu promenu Y -a kada se pri nepromenjenom X_2 , X_1 poveća za 1.

β_2 – pokazuje prosečnu promenu Y -a kada se pri nepromenjenom X_1 , X_2 poveća za 1.

$\hat{y}_i = b_0 + b_1 x_{1i} + b_2 x_{2i}$ - regresiona ravan u uzorku

b_0 – ocena za β_0 , b_1 – ocena za β_1 , b_2 – ocena za β_2 .

2. Mere reprezentativnosti regresione funkcije

a) Apsolutna mera reprezentativnosti (s)

S – standardna greška regresije (pokazuje prosečno odstupanje podataka od regresione ravni iskazano u jedinicama mere Y -a).

$$0 \leq s < \infty$$

$S = 0$ – funkcionalna veza

Sa porastom S opada reprezentativnost.

$$s = \sqrt{\frac{SKN}{n-3}}$$

b) Relativna mera reprezentativnosti (R^2)

R^2 – koeficijent višestruke determinacije (pokazuje udeo objašnjenog varijabiliteta u ukupnom)

$$0 \leq R^2 \leq 1$$

Uslov reprezentativnosti je $R^2 > 0,5$.

$R^2 = 0 \rightarrow$ odsustvo veze

$R^2 = 1 \rightarrow$ funkcionalna veza

R^2 se izražava u procentima. Sa porastom R^2 raste reprezentativnost.

$1 - R^2$ – udeo neobjašnjenog varijabiliteta u ukupnom.

Kod malih uzoraka koristimo korigovani koeficijent višestruke determinacije $\bar{R}^2 = 1 - \frac{n-1}{n-3}(1 - R^2)$

3. Testiranje značajnosti regresione linije

a) Da li X_1 utiče na Y

$$H_0 : \beta_1 = 0 \quad H_1 : \beta_1 \neq 0$$

$$H_0 : \beta_1 = 0 \quad H_1 : \beta_1 > 0$$

$$H_0 : \beta_1 = 0 \quad H_1 : \beta_1 < 0$$

$$t = \frac{b_1}{s_{b1}}, df = n - 3, \quad p = 2P(T > |t|)$$

$$p \leq \alpha \Rightarrow X_1 \text{ utiče na } Y$$

$$b_1 - t_{n-3;\alpha/2} \cdot s_{b1} \leq \beta_1 \leq b_1 + t_{n-3;\alpha/2} \cdot s_{b1}$$

b) Da li X_2 utiče na Y

$$H_0 : \beta_2 = 0 \quad H_1 : \beta_2 \neq 0$$

$$H_0 : \beta_2 = 0 \quad H_1 : \beta_2 > 0$$

$$H_0 : \beta_2 = 0 \quad H_1 : \beta_2 < 0$$

$$t = \frac{b_2}{s_{b2}}, df = n - 3, \quad p = 2P(T > |t|)$$

$$p \leq \alpha \Rightarrow X_2 \text{ utiče na } Y$$

$$b_2 - t_{n-3;\alpha/2} \cdot s_{b2} \leq \beta_2 \leq b_2 + t_{n-3;\alpha/2} \cdot s_{b2}$$

U zadacima često pitaju da li datim podacima više odgovara prost ili višestruki model. U tom slučaju treba izvršiti oba prethodna testa, pa ako oba x -a utiču na y to znači da podacima više odgovara višestruki model, a ako samo jedan od x -ova utiče na y , to znači da podacima više odgovara prost regresioni model.

4. Korišćenje regresione funkcije za ocenjivanje i predviđanje Y -a

Uslovi:

- 1) Reprezentativnost ($R^2 > 0,5$)
- 2) Značajnost ($\beta_1 \neq 0, \beta_2 \neq 0$)
- 3) Da nije ekstrapolacija.

a) Ocenjivanje prosečne vrednosti Y -a

$$\hat{y}_p - t_{n-3;\alpha/2} s_{\hat{y}_p} \leq \mu_{Y/X} \leq \hat{y}_p + t_{n-3;\alpha/2} s_{\hat{y}_p}$$

$$\hat{y}_p = b_0 + b_1 x_{1p} + b_2 x_{2p}$$

x_{1p}, x_{2p} - vrednosti za X_1 i X_2 date u zadatku, za koje vršimo predviđanje.

$s_{\hat{y}_p}$ - standardna greška ocene prosečne vrednosti Y -a.

b) Predviđanje individualne vrednosti Y-a

$$\hat{y}_p - t_{n-3;\alpha/2} s_{y_p} \leq Y_p \leq \hat{y}_p + t_{n-3;\alpha/2} s_{y_p}$$

s_{y_p} - standardna greška predviđanja individualne vrednosti Y-a.

B. Višestruka linearna korelacija

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

$$0 \leq R \leq 1$$

R - koeficijent višestruke linearne korelacije (pokazuje stepen kvantitativnog slaganja – korelacione veze između varijacija Y-a sa jedne strane i X_1 i X_2 sa druge).

PRIMER:

Na osnovu podataka iz zadatka br. **29.3.** sa strane 46. zbirke, dobijeni su sledeći rezultati, korišćenjem statističkog programa **MINITAB**:

Regression Analysis

The regression equation is

Promet = 98,3 - 1,59 Zaposleni + 9,92 Prostor

	Koef	St. Greška	T	P
Konstanta	98,281	6,801	14,45	0,001
Zaposleni	-1,592	1,381	-1,15	0,332
Prostor	9,924	2,267	4,38	0,022

S = 4,299 $R^2 = 97,6\%$ R^2 (korigovani) = 96,0%

a) Objasnite ekonomski smisao ocenjenih regresionih koeficijenata.

b) Kako tumačite dobijeni koeficijent višestruke determinacije? Objasnite zašto uvodimo korigovani koeficijent višestruke determinacije i interpretirajte njegovu vrednost u ovom primeru.

c) Testirajte da li broj zaposlenih utiče na obim prometa. Da li veličina poslovnog prostora utiče na obim prometa? Na osnovu dobijenih rezultata zaključite da li datim podacima više odgovara prost ili višestruki linearni regresioni model.

d) Predvidite prosečan nivo prometa u prodavnicama sa deset zaposlenih i poslovnim prostorom od 2000m², ako je poznato da standardna greška ocene prosečne vrednosti obima prometa iznosi 26,52. Da li je dobijena vrednost validna? Objasnite!

e) Izračunajte i objasnite koeficijent višestruke linearne korelacije.

f) Ocenite za koliko se promeni promet ako se pri nepromenjenom broju zaposlenih veličina poslovnog prostor poveća za 200 m².

14. Spirmanova korelacija ranga (r_s)

Računa se samo ako je bar jedan od podataka na ordinalnoj skali ili ako je bar jedan od podataka rangiran, ili ako skup nije normalan.

$$r_s = 1 - \frac{6 \sum d_i^2}{n \cdot (n^2 - 1)} \quad d_i = R_X - R_Y$$

$$-1 \leq r_s \leq 1$$

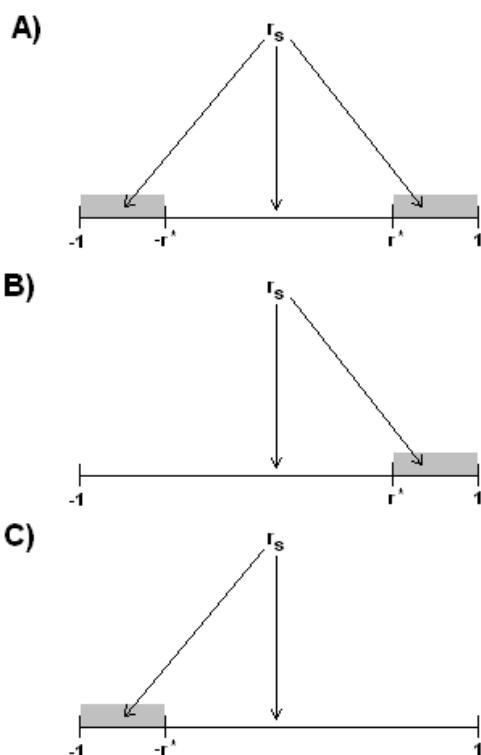
Testiranje Spirmanove korelacije:

A) $H_0 : \rho_s = 0 \quad H_1 : \rho_s \neq 0$

B) $H_0 : \rho_s = 0 \quad H_1 : \rho_s > 0$

C) $H_0 : \rho_s = 0 \quad H_1 : \rho_s < 0$

A) je dvosmerna, a B) i C) su jednosmerne hipoteze.



Tablica 11.

Ako je $n > 30$ testiranje se vrši pomoću Z - testa, gde je $Z = r_s \cdot \sqrt{n-1}$.

15. Indeksni brojevi

A. Individualni indeksi

1. Lančani indeksi

$$L_t = \frac{y_t}{y_{t-1}} \cdot 100$$

y_t - tekući podatak
 y_{t-1} - prethodni podatak

L_t se tumači u odnosu na 100, pa tako na primer $L_t = 101,1$ znači da je podatak za 1,1% veći od prethodnog, a $L_t = 98,3$ da je podatak za 1,7% manji od prethodnog.

2. Bazni indeksi

$$I_t = \frac{y_t}{y_0} \cdot 100$$

y_t - tekući podatak
 y_0 - podatak iz baznog perioda

Najpoznatiji bazni indeksi:

$$I_p = \frac{p_t}{p_0} \cdot 100 \rightarrow \text{indeks cena}$$

$$I_q = \frac{q_t}{q_0} \cdot 100 \rightarrow \text{indeks količina}$$

$$I_{pq} = \frac{p_t q_t}{p_0 q_0} \cdot 100 \rightarrow \text{indeks vrednosti}$$

p_0 - cena iz baznog perioda
 p_t - cena iz tekućeg perioda
 q_0 - količina iz baznog perioda
 q_t - količina iz tekućeg perioda
 $p_0 q_0$ - vrednost iz baznog perioda
 $p_t q_t$ - vrednost iz tekućeg perioda

3. Preračunavanje individualnih indeksa

a) Preračunavanje baznih indeksa sa bazom u jednoj godini na bazne indekse sa bazom u drugoj godini i lančane indekse

Princip: zamislimo da su bazni indeksi podaci, pa na osnovu njih izračunavamo druge indekse.

Primer:

Godina	I _t 1983=100	I _t 1985=100	L _t
83	100		
84	101,1		
85	101,3		
86	105,0		
87	103,2		
88	101,9		

b) Preračunavanje lančanih indeksa u bazne indekse

$$I_t = \frac{L_t \cdot I_{t-1}}{100} \downarrow \quad I_{t-1} = \frac{I_t}{L_t} \cdot 100 \uparrow$$

Primer:

Godina	L _t	I _t 1985=100
83	/	
84	101,1	
85	100,1	
86	103,7	
87	98,3	
88	98,7	

4. Geometrijska stopa rasta

Geometrijska stopa rasta (r_g) se računa samo za vremenske serije i pokazuje prosečan rast pojave u posmatranom periodu iskazan u procentima. Računa se na 2 načina:

1. Pomoću originalnih podataka

$$r_g = \left(\sqrt[T]{\frac{y_T}{y_1}} - 1 \right) \cdot 100$$

y_1 - prvi podatak, y_T - poslednji podatak, T – broj podataka

2. Pomoću lančanih indeksa

$$r_g = G - 100, \quad G = \sqrt[T]{L_2 L_3 \cdots L_T}$$

r_g se tumači u odnosu na nulu. Tako na primer $r_g = 57,7\%$ znači da je pojava u proseku godišnje rasla za 57,7%, a $r_g = -32,2\%$ znači da je pojava u proseku godišnje opadala za 32,2%.

Za razliku od njega lančani indeksi se tumače u odnosu na 100. Tako npr. $L_t = 123$ znači da je pojava za 23% veća u odnosu na prethodni period, a $L_t = 92$ znači da je pojava za 8% manja u odnosu na prethodni period.

B. Grupni indeksi

1. Metod agregata

$$I_p = \frac{\sum p_t}{\sum p_0} \cdot 100 \quad I_q = \frac{\sum q_t}{\sum q_0} \cdot 100 \quad I_{pq} = \frac{\sum p_t q_t}{\sum p_0 q_0} \cdot 100$$

Metod prosečnih odnosa - Aritmetički indeksi

$$\bar{I}_p = \frac{\sum I_p}{n} \quad \bar{I}_q = \frac{\sum I_q}{n} \quad \bar{I}_{pq} = \frac{\sum I_{pq}}{n}$$

C. Ponderisani grupni indeksi

1. Metod agregata

$$\left. \begin{aligned} I_p &= \frac{\sum p_t q_0}{\sum p_0 q_0} \cdot 100 \\ I_q &= \frac{\sum q_t p_0}{\sum q_0 p_0} \cdot 100 \\ I_{pq} &= \frac{\sum p_t q_t}{\sum p_0 q_0} \cdot 100 \end{aligned} \right\} \text{Ponder iz baznog perioda (Laspejresov metod)}$$

$$\left. \begin{aligned} I_p &= \frac{\sum p_t q_t}{\sum p_0 q_t} \cdot 100 \\ I_q &= \frac{\sum q_t p_t}{\sum q_0 p_t} \cdot 100 \\ I_{pq} &= \frac{\sum p_t q_t}{\sum p_0 q_0} \cdot 100 \end{aligned} \right\} \text{Ponder iz tekućeg perioda (Paševov metod)}$$

$$\left. \begin{aligned} I_p &= \sqrt{\frac{\sum p_t q_0}{\sum p_0 q_0} \cdot \frac{\sum p_t q_t}{\sum p_0 q_t}} \cdot 100 \\ I_q &= \sqrt{\frac{\sum q_t p_0}{\sum q_0 p_0} \cdot \frac{\sum q_t p_t}{\sum q_0 p_t}} \cdot 100 \\ I_{pq} &= \frac{\sum p_t q_t}{\sum p_0 q_0} \cdot 100 \end{aligned} \right\} \text{Fišerovi "idealni" indeksi}$$

2. Metod prosečnih odnosa (aritmetički indeksi)

$$\left. \begin{aligned} \bar{I}_p &= \frac{\sum I_p p_0 q_0}{\sum p_0 q_0} \\ \bar{I}_q &= \frac{\sum I_q p_0 q_0}{\sum p_0 q_0} \\ \bar{I}_{pq} &= \frac{\sum I_{pq} p_0 q_0}{\sum p_0 q_0} \end{aligned} \right\} \text{Ponder bazni}$$

16. Analiza vremenskih serija

Vremenske serije su nizovi statističkih podataka koji pokazuju varijacije pojava tokom vremena. Cilj ovog poglavlja je da izvršimo analizu vremenskih serija i utvrdimo zašto postoji varijacija podataka tokom vremena.

Uzmimo na primer broj turista na Crnogorskom primorju. Taj broj se tokom godina menja. 1998. broj turista je npr. 185670 a 2000. broj turista je 190248. Zašto postoje ove razlike?

Na varijacije podataka u vremenskim serijama utiču 4 faktora:

T - uticaj trenda

S - uticaj sezone

C - ciklična komponenta

R - rezidualni uticaj.

Uticaj trenda predstavlja dugoročni uticaj na neku pojavu, tako da pojava neprestano raste ili opada po nekom zakonu. Tako npr. broj turista na Crnogorskom primorju iz godine u godinu je sve veći i veći, dakle postoji dugoročni rastući trend ove pojave.

Uticaj sezone je pojava da u određenim delovima godine neka pojava je izrazito veća od proseka za tu godinu, a u drugim delovima godine izrazito manja od proseka. Tako npr. broj turista na Crnogorskom primorju je u julu i avgustu mnogo veći od proseka za tu godinu, a u januaru i februaru mnogo manji od proseka.

Ciklična komponenta predstavlja pojavu da je tokom niza godina neka pojava naglo veća od očekivanja, a onda niz godina naglo manja. Tako na primer u periodu rata i sankcija broj stranaca na Crnogorskom primorju je značajno opao, a u periodu posle će verovatno naglo da poraste zbog znatizelje starih posetilaca.

Rezidualna komponenta predstavlja slučajni uticaj na varijaciju pojave.

Postoji više teorija kako ove komponente utiču na varijaciju pojave.

Ako podatke serije označimo sa Y po multiplikativnoj teoriji važiće $Y=T \cdot C \cdot S \cdot R$, po aditivnoj $Y=T+C+S+R$, po kombinovanim $Y=T+C \cdot S \cdot R$ ili $Y=T \cdot C+S \cdot R$ itd. Ako su podaci godišnji multiplikativni model je $Y=T \cdot C$. U ovom poglavlju izučava se samo multiplikativni model.

Kod multiplikativnog modela T je dato u jedinici mere Y -a, a S, C i R su relativni brojevi. Kod aditivnog modela i T i S i C i R su dati u jedinici mere Y -a.

U svakom zadatku pre nego što se krene sa bilo kakvom analizom vrši se prenumeracija podataka. Prenumeracija podataka znači da godine i kvartali nisu t kolona već se t kolona formira prostom prenumeracijom od 1 do n .

A. *Uticaj trenda*

Postoje tri funkcije trenda:

- 1) Linearni trend
- 2) Parabolični trend
- 3) Eksponencijalni trend

Prvi problem pri određivanju funkcije trenda je odrediti koja funkcija trenda je najbolja za konkretni slučaj.

Postoje četiri načina za izbor funkcije trenda:

1. Metod diferencija (razlika) - Najpre se formira nova kolona u tabeli koja se označava sa Δ_1 i u koju se upisuju apsolutne razlike susednih podataka iz kolone y . Ukoliko su te razlike približno jednake, onda je najbolji linearni trend, a ako nisu jednake formira se nova kolona u tabeli Δ_2 u koju se upisuju apsolutne razlike susednih podataka iz kolone Δ_1 . Ako su te razlike približno jednake onda je u pitanju parabolični trend, a ako nisu formira se nova kolona $\Delta_{\log y}$ u koju se upisuju apsolutne razlike svih susednih podataka iz kolone $\log y$, pa ako su te razlike približno jednake onda je to eksponencijalni trend. Najbolji primer kako se ovo radi dat je u rešenju zadatka 6.2. Obratiti pažnju da su razlike uvek pozitivne. Znači nije bitno koji je podatak veći.

2. Metod pokretnih sredina (proseka) - Najpre se formira kolona u koju se upisuju pokretni proseci. Pokretni proseci mogu da budu dvogodišnji, trogodišnji, četvorogodišnji, itd. Ukoliko se pokretni proseci prave sa neparnim brojem podataka, npr. trogodišnji pokretni proseci, vrši se sabiranje svake tri susedne godine i deljenje sa tri. Rezultat koji se dobije se piše naspram srednje od te tri godine, znači naspram druge. Primer pravljenja trogodišnjih i petogodišnjih pokretnih proseka dat je na strani 403. Ukoliko se pokretni proseci prave sa parnim brojem podataka, npr. četvorogodišnji pokretni proseci, vrši se sabiranje svake četiri susedne godine i deljenje sa 4. Rezultat koji se dobije se piše na mestu između drugog i trećeg podatka. Međutim, kako sad pokretni proseci nisu centrirani, vrši se njihovo

centriranje, tako što se svaka dva susedna podatka sabere, rezultat podeli sa dva i napiše između ta dva podatka. Primer pravljenja pokretnih proseka od 4 podatka dat je na str 428. Kada su formirani pokretni proseki, na te pokretne proseke se primenjuje grafički metod.

3. Srednja kvadratna greška - Najpre se pronađu sve tri funkcije trenda, pa se za svaku od tih funkcija izračuna

srednja kvadratna greška po sledećoj formuli: $SKG = \frac{\sum (y - \hat{y}_t)^2}{T}$.

Bira se ona funkcija trenda koja ima najmanju srednju kvadratnu grešku.

4. Grafički metod – Serija se prikaže na aritmetičkom dijagramu. Ako liči na liniju onda je to linearni trend, a ako liči na eksponencijalnu funkciju onda je to eksponencijalni trend.

1. Linearni trend

Funkcija linearnog trenda je $\hat{y}_t = b_0 + b_1 t$, gde je $b_0 = \bar{y} - b_1 \bar{t}$, $b_1 = \frac{\sum t y - \frac{\sum t \sum y}{T}}{\sum t^2 - \frac{(\sum t)^2}{T}}$.

b_1 – pokazuje srednji apsolutni porast za posmatrani period.

2. Parabolični trend (izbačeno)

Funkcija paraboličnog trenda je: $\hat{y}_t = b_0 + b_1 t + b_2 t^2$, gde se koeficijenti b_0, b_1 i b_2 dobijaju rešavanjem sledećeg sistema linearnih jednačina:

$$\begin{aligned}\sum y &= T \cdot b_0 + b_1 \sum t + b_2 \sum t^2 \\ \sum t y &= b_0 \sum t + b_1 \sum t^2 + b_2 \sum t^3 \\ \sum t^2 y &= b_0 \sum t^2 + b_1 \sum t^3 + b_2 \sum t^4\end{aligned}$$

3. Eksponencijalni trend

Funkcija eksponencijalnog trenda je: $\hat{y}_t = b_0 \cdot b_1^t$, gde je

$$\log b_0 = \frac{\sum \log y}{T} - \log b_1 \cdot \frac{\sum t}{T}, \quad \log b_1 = \frac{\sum t \log y - \frac{\sum t \sum \log y}{T}}{\sum t^2 - \frac{(\sum t)^2}{T}},$$

$$b_0 = 10^{\log b_0}, b_1 = 10^{\log b_1}.$$

Vezano za eksponencijalni trend definiše se i eksponencijalna stopa rasta: $r_e = (b_1 - 1) \cdot 100$.

b_1 – srednji relativni porast

$b_1 \cdot 100$ – srednji tempo rasta

$\log \hat{y}_t = \log b_0 + t \cdot \log b_1$ – linearni logaritamski trend.

1. Tip zadatka: Daju nam funkciju trenda, ili je mi sami nađemo, pa treba da je iskoristimo za predviđanje podataka u budućnosti, a ne spominje se nigde u zadatku sezone.

2. Tip zadatka: Isključiti trend (naći procentualno odstupanje podataka od linije trenda).

$$\frac{y}{\hat{y}_t} \cdot 100 \text{ - formula za isključenje trenda}$$

B. Sezonska komponenta

Uticaj sezone je pojava da podaci variraju u okviru godine. U svim zadacima vezano za uticaj sezone podaci će uvek biti dati po kvartalima.

Uticaj sezone se radi na sledeći način: (strana 428)

Najpre se formiraju pokretni proseci i centrirani pokretni proseci od po četiri podatka na ranije objašnjeni način.

Zatim se formira kolona $\frac{y}{\bar{y}_c}$ koja se zove specifični sezonski indeksi. Naziv specifični sezonski indeksi potiče od toga

da je u koloni dat količnik podataka iz nekog kvartala y sa prosekom za celu godinu $\bar{y}_c$. Dakle, dobija se relativni uticaj kvartala na podatke. Ako je indeks veći od 1 znači da su u tom kvartalu podaci veći od godišnjeg proseka i obrnuto. Međutim, problem je što npr. za prvi kvartal mi imamo u tabeli dato 6 različitih vrednosti. To se rešava pomoću pomoćne tabele na strani 429. U njoj se za svaki kvartal upišu svi specifični sezonski indeksi i nađe njihov prosek. Tako se dobija kolona koja se zove kvartalne sredine (a ne kao što piše u knjizi I_S). Množenjem kvartalnih sredina sa 100 dobijamo tipične sezone indekse I_S koji predstavljaju uticaje pojedinih kvartala na podatke. Tako npr. za prvi kvartal I_S je 47 što znači da su podaci u prvom kvartalu za 53% manji od proseka, a u trećem kvartalu I_S je 131 što znači da su podaci u trećem kvartalu za 31% veći od proseka.

Kvartal sa najvećim I_S se zove živa sezona, a kvartal sa najmanjim I_S mrtva sezona.

Kvartal čiji I_S se najviše razlikuje od 100 je kvartal u kojem je sezona najviše izražena.

Kod kvartalnih podataka važi: $\sum I_S = 400$.

Kod mesečnih podataka važi: $\sum I_S = 1200$.

Vezano za sezone indekse postoje dva tipa zadataka:

3. Tip zadatka: Predvideti nivo pojave u budućnosti uzimajući u obzir uticaj sezone.

Formula za predviđanje nivoa pojave koristeći uticaj sezone je: $\hat{y}_t^* = \hat{y}_t \cdot \frac{I_S}{100}$.

4. Tip zadatka: Isključiti uticaj sezone iz nekih podataka. (Izvršiti desezoniranje)

Formula za isključenje uticaja sezone je $\frac{y}{I_S} \cdot 100$. Primer:

1992	Y	I _S	y/I _S ·100
I	129	47	274,47
II	239	88	271,59
III	361	131	275,57
IV	367	134	273,88

C. *Ciklična komponenta*

Ciklična komponenta je u knjizi totalno zanemarena, tako da dalji tekst možete slobodno preskočiti. Ukoliko ipak želite da znate i ponešto o tome pročitajte dalji tekst.

Ovde postoje dva tipa zadatka:

5. Tip zadatka: Uticaj nepostojanih faktora

Izolacija nepostojanih faktora vrši se po obrascu: $CR = \frac{y_t}{\hat{y}_t^*} \cdot 100$, ako su u pitanju kvartali ili meseci, a za godišnje podatke formula je $CR = \frac{y_t}{\hat{y}_t} \cdot 100$.

rokovi

- Na osnovu raspoloživih podataka o ceni (u 000 din) i ponudi (u 000 kom) u uzorku od 7 prodavnica dobijeni su sledeći podaci: 96,55% varijacija u ponudi je objašnjeno promenama cene, suma kvadrata reziduala je 28950 (000 kom)². Cene u proseku odstupaju od prosečne cene za 6.8 hilj. din. Najmanja cena u uzorku je 3000 dinara, a najveća 22000 din. Prosečna cena u uzorku je 10,29 hiljada dinara, a prosečna ponuda je 700 hiljada dinara.
 - Oceniti prost regresioni model i interpretiraj vrednost koeficijenta nagiba.
 - Sa pouzdanošću od 95% oceni prosečnu ponudu, ako je cena 20000 dinara.
- Iz skupa od 800 ispitanika odabran je uzorak od 60 i zabeleženo je:

Mesečne zarade (u 000 din)	Radni staž u godinama	
	Do 15	Preko 15
20-40	8	9
40-60	9	13
60 i više	15	6

- Formirati 95% interval poverenja za učešće zaposlenih sa zaradama do 60000 dinara
- Na nivou značajnosti od 5% ispitati da li je učešće zaposlenih sa stažom do 15 godina veće od 60%

- Vrednosti prodaje proizvoda A, B i C u 2014. su 15, 18 i 20 mil. din. Cene ova tri proizvoda u 2016. su 8, 9 i 12. Vrednost prodaje proizvoda A se u 2016. povećala za 15%, proizvoda C za 22%, a proizvoda B se smanjila za 8%. Cene sva tri proizvoda su porasle za 13%, 11% i 16%. Metodom agregata izračunati grupne indekse fizičkog obima prodaje i cene koristeći pondere iz tekućeg perioda.
- Cilj jednog istraživanja je da se ispita razlika u prosečnom uspehu studenata koji pohađaju 5 različitih metoda nastave jednog predmeta. Od svakog od 5 metoda odabrani su uzorci i zabeležene su ocene koje su studenti dobili na ispitu:

I	II	III	IV	V
6	8	8	10	9
6	8	9	6	7
7	9	7	10	9
9		8	9	
10			8	
			6	

Na nivou značajnosti od 10% ispitati da li metodi nastave izazivaju sistematske varijacije u uspehu studenata.

- Slučajna promenljiva X ima normalnu raspodelu, pri čemu važi da je $P(X > 6) = 0,4013$. Slučajna promenljiva Y takođe ima normalnu raspodelu kod koje je $P(Y < 4) = 0,1469$. Poznato je da je varijansa slučajne promenljive Y 4 puta manja od varijanse slučajne promenljive X. Relativna disperzija promenljive X je 2,76 puta veća od relativne disperzije promenljive Y. Odrediti varijanse slučajnih promenljivih X i Y.

1. U prostom slučajnom uzorku od 7 stanovnika jedne opštine dobijeni su podaci o mesečnoj plati (u hiljadama dinara) i mesečnim izdacima za telefon (u stotinama dinara):

Mesečna plata	17	28	31	39	45	52	66
Mesečni izdaci	6	9	18	15	20	29	23

- a) Izračunati i proumačiti meru linearnog slaganja između promenljivih u uzorku **(5)**
 b) Izračunati i protumačiti apsolutnu meru reprezentativnosti linije regresije, a zatim i učešće neobjašnjenog varijabiliteta u ukupnom. **(10)**
 c) Sa pouzdanošću od 95% oceniti za koliko se u proseku promene mesečni izdaci za telefon, ako se mesečna plata poveća za 6 hiljada dinara. **(10)**
2. U sledećoj tabeli su dati podaci o izvozu tri vrste prehrambenih proizvoda u 2013. i 2016. godini:

	Individualni indeksi u 2016. (2013=100)			Vrednost izvoza u 2013. (u dinarima)
	Cene	Količine	Vrednosti	
A	116,67	124	144,67	22.500
B	115,38	92,86	107,14	36.400
C	120,69	131,82	159,09	31.900

Metodom agregata izračunati grupne indekse izvoznih cena i količina izvoza sve tri vrste prehrambenih proizvoda u 2016. godini (2013=100), koristeći pondere iz baznog perioda. **(20)**

3. Na jednom fakultetu studira 500 studenata. Od tog broja 45% studenata ima prosečnu ocenu ispod 7, 25% studenata ima prosečnu ocenu između 7 i 8, njih 20% ima prosečnu ocenu između 8 i 9, dok ostali studenti imaju prosečnu ocenu iznad 9. U prostom slučajnom uzorku od 40 studenata posmatranog fakulteta struktura je sledeća:

Prosečna ocena	Učešće studenata
Ispod 7	30%
7-8	40%
8-9	...
preko 9	20%

Na nivou značajnosti od 5% ispitati da li se raspodela studenata prema visini prosečne ocene značajno razlikuje od očekivane raspodele. **(15)**

4. Standardna devijacija normalno raspoređenog osnovnog skupa iznosi 3, a relativna disperzija je 12,5% sredine. Iz osnovnog skupa od 500 elemenata izabran je prost slučajni uzorak i zabeležene su sledeće vrednosti: 13, 17, 19, 22, 18, 20, 14, 17, 19, 12, 15, 17, 21, 23, 14, 20.
 a) Izračunati tačno učešće i ukupan broj elemenata u skupu sa vrednošću većom od 17. **(5)**
 b) Uporediti interval varijacije osnovnog skupa i uzorka. **(5)**
 c) Sa pouzdanošću od 95% utvrditi da li je uzorak reprezentativan pri oceni sredine skupa. **(10)**
5. Jedno istraživanje je pokazalo da je 56% ispitanika trgovalo akcijama na berzi. Neka se ovaj rezultat odnosi na populaciju. Pretpostavimo da su 4 osobe izabrane slučajnim putem. Neka je X broj osoba koje trguju na berzi.
 a) Odrediti raspodelu verovatnoća za X i nacrtati stablo ishoda. **(15)**
 b) Izračunati verovatnoću da od 4 izabrane osobe tačno njih dve trguju akcijama na berzi **(5)**

1. U sledećoj tabeli je prikazana raspodela 60 slučajno izabranih radnika jednog preduzeća prema ostvarenoj proizvodnji i smeni:

Smena	Broj proizvedenih proizvoda		
	Ispod 10	10-20	Preko 20
I	6	9	5
II	10	6	4
III	3	11	6

Na nivou značajnosti od 5%, ispitati da li postoji razlika u prosečnoj proizvodnji između radnika druge i treće smene (sve potrebne pretpostavke su ispunjene). **(15)**

2. U prostom slučajnom uzorku od 7 stanovnika jedne opštine dobijeni su podaci o mesečnoj plati (u hiljadama dinara) i mesečnim izdacima za posete sportskim dešavanjima (u stotinama dinara):

Mesečna plata	22	27	36	41	49	60	74
Mesečni izdaci	8	11	17	20	28	19	30

a) Za koliko mesečni izdaci pojedinih ispitanika u proseku odstupaju od ocenjenih mesečnih izdataka, a za koliko od prosečnih mesečnih izdataka? **(10)**

b) Sa pouzdanošću od 95%, oceniti prosečan nivo mesečnih izdataka za posete sportskim dešavanjima ispitanika sa mesečnom zaradom od 76 hiljada dinara. Da li je ocenjena vrednost validna? **(15)**

3. Iz osnovnog skupa od 600 ispitanika koji redovno idu na rekreaciju izabran je prost slučajni uzorak od njih 80. Raspodela ispitanika po starosti data je u sledećoj tabeli:

Starost (u godinama)	Broj ispitanika
20-30	19
30-40	28
40-50	23
Preko 50	10

a) Sa pouzdanošću od 95% oceniti učešće ispitanika starijih od 40 godina koji redovno idu na rekreaciju. **(10)**

b) Na nivou značajnosti od 5% ispitati da li je učešće ispitanika mlađih od 30 godina koji svakodnevno treniraju veće od 30% **(10)**

4. U sledećoj tabeli je data serija baznih indeksa prodaje jedne vrste laptop računara u jednoj prodavnici u period od 2011-2016. godine:

Godina	2011.	2012.	2013.	2014.	2015.	2016.
Bazni indeksi (2010=100)	125,60	155,20	172	184,80	231,20	241,60

a) Polazeći od lančanih indeksa, izračunati prosečnu procentualnu promenu prodaje laptop računara u posmatranom periodu. **(5)**

b) Na osnovu srednje kvadratne greške, izabrati i oceniti odgovarajuću funkciju trenda u period 2011. do 2016. godine (SKG lin. trenda iznosi 128,52 komada²), ako je prodaja u 2010. iznosila 125 komada. **(15)**

5. Vreme koje svi studenti provedu gledajući filmove u toku jedne nedelje sledi normalnu raspodelu sa sredinom 8 časova. Učešće studenata koji gledaju filmove između 6 i 10 časova iznosi 89,04%. Verovatnoća da proporcija studenata koji gledaju filmove manje od 11 časova u prostom slučajnom uzorku veličine n bude veća od proporcije u osnovnom skupu za više od 0,05 iznosi 0,0150. Odrediti veličinu prostog slučajnog uzorka (n).

1. Dati su indeksi fizičkog obima prodaje jedne vrste voća, po kvartalima 2014, 2015 i 2016. godine (prosek 2015=100): 42,55; 102,13; 164,54; 59,57; 48,23; 99,29; 178,72; 73,76; 36,88; 110,64; 190,07; 51,06. Prodaja pomenutog voća u II kvartalu 2015. godine iznosila je 35 hiljada tona. Sezonski indeksi prodaje po kvartalima u posmatrane tri godine iznosili su redom: 43,5; 106; 179 i 69,5. Predvideti prosečan obim prodaje pomenutog voća u kvartalu sa „živom“ sezonom u 2017. godini, uzimajući u obzir uticaj sezonskih faktora. **(25)**
2. Od 250 studenata jednog fakulteta, 40 studenata nije položilo nijedan ispit u toku školske godine. Verovatnoća da učešće studenata koji nisu položili nijedan ispit u prostom slučajnom uzorku veličine n bude manje od učešća svih studenata koji nisu položili nijedan ispit za više od 0,06 iznosi 0,1635. Odrediti veličinu prostog slučajnog uzorka. **(15)**
3. Iz tri košarkaška kluba u Srbiji slučajnim putem je izabrano po 5 igrača i zabeleženo vreme koje nedeljno provedu trenirajući (u časovima):

I klub: 5 8 7 10 9

II klub: 9 7 11 15 12

III klub: 10 6 13 16 9

Na nivou značajnosti od 5% ispitati da li se prosečno vreme treniranja igrača razlikuje između bar dva kluba (sve potrebne pretpostavke su ispunjene). **(15)**

4. U jednom istraživanju je učestvovalo 100 kupaca, koji su dali odgovore na pitanja o starosti i učestalosti kupovine u prodavnicama zdrave hrane u toku jedne nedelje. Rezultati su prikazani u sledećoj tabeli:

		Učestalost kupovine u toku nedelje		
		Nikad	Jednom nedeljno	Više od jednom nedeljno
Godine starosti kupca	do 20	4	5	15
	20-40	5	9	10
	40-60	9	11	13
	preko 60	14	3	2

- a) Sa pouzdanošću od 95% formirati interval poverenja za učešće kupaca koji su stariji od 40 godina. **(10)**
 - b) Na nivou značajnosti od 5% ispitati da li postoji veza između starosti kupaca i učestalosti kupovine **(15)**
5. U prostom slučajnom uzorku od 7 prodavnica na teritoriji Beograda zabeleženi su podaci o izdacima za reklamiranje (u 000 din) i broju prodatih proizvoda (u 00 kom) i dobijeni su sledeći rezultati: broj prodatih proizvoda u proseku odstupa od prosečnog broja prodatih proizvoda za 4,57 stotina komada; ocenjena standardna greška ocene regresionog parametra uz objašnjavajuću promenljivu je 0,166 hiljada dinara; prosečan broj prodatih proizvoda iznosi 8,29 stotina komada; izdaci za reklamiranje iznose: 2, 4, 4, 5, 7, 8, 10.
 - a) Oceniti prost linearni regresioni model i objasniti dobijeni regresioni koeficijent nagiba. **(10)**
 - b) Sa pouzdanošću od 95% predvideti broj prodatih proizvoda, ako izdaci za reklamiranje iznose 12 hiljada dinara. Da li je dobijena predviđena vrednost validna? **(10)**

1. Struktura 150 anketiranih osoba po starosti i polu data je sledećom tabelom:

Starost	Pol	
	Muški	Ženski
do 20	9	17
20-30	14	21
30-40	12	33
Preko 40	20	24

a) Odrediti relativnu meru disperzije starosti osoba muškog pola. **(10)**

b) Sa pouzdanošću od 95% formirati interval poverenja za prosečnu starost osoba ženskog pola. **(10)**

2. Dati su lančani indeksi proizvodnje jednog proizvoda u svim mesecima 2015. godine: 103,70; 104,46; 90,60; 92,45; 96,94; 95,79; 108,79; 106,06; 104,76; 103,64; 107,89; 104,88. Prosečna vrednost proizvodnje (u hiljadama tona) tog proizvoda u decembru 2014. godine iznosila je 108 hiljada tona. Ocenjena funkcija eksponencijalnog trenda je $\hat{y}_t = 20,65 \cdot 1,16^t$, a srednja kvadratna greška eksponencijalnog trenda je 115,12 hiljada tona².

a) Odabrati funkciju trenda koja se najbolje prilagođava podacima u posmatranom period (kao kriterijum koristiti srednju kvadratnu grešku). **(15)**

b) Isključiti trend komponentu u junu 2015. Objasniti rezultat. **(5)**

3. U jednom preduzeću radi 1200 radnika. Od tog broja 300 radnika radi u preduzeću manje od 10 godina, 600 radnika radi od 10-20 godina, 180 radnika radi od 20-30 godina, a ostali rade preko 30 godina. U prostom slučajnom uzorku od 150 radnika struktura je sledeća:

Radno iskustvo (u godinama)	Učešće radnika (u %)
0-10	30
10-20	...
20-30	20
preko 30	10

Na nivou značajnosti od 5% ispitati da li se raspodela radnika prema radnom iskustvu razlikuje od očekivane raspodele. **(15)**

4. Slučajna promenljiva X ima normalnu raspodelu sa parametrima μ i σ^2 . Poznato je da se 75% površine ispod krive nalazi ulevo od broja 6,4 i da je $P(X > \mu - 1,2) = 0,9772$. Verovatnoća da prosek slučajne promenljive X u prostom slučajnom uzorku veličine n bude veći od μ za više od 0,1 iznosi 0,1587. Izračunati verovatnoću da proporcija vrednosti koje su manje od 5,7 u prostom slučajnom uzorku veličine n bude veća od proporcije tih vrednosti u osnovnom skupu za najviše 20%. **(20)**

5. Sprovedeno je istraživanje sa ciljem da se utvrdi da li starost gledalaca (u godinama) i slobodno vreme (u časovima) utiču na gledanost televizijskog programa (u časovima). Na osnovu prostog slučajnog uzorka od 8 gledalaca dobijeni su sledeći rezultati:

	Koeficijent	Standardna greška
Konstanta	14,2	2,8
Starost gledalaca	2,56	2,12
Slobodno vreme	3,84	1,16

94,5% varijacija gledanosti televizijskog programa objašnjeno je zajedničkim uticajem starosti gledalaca i slobodnog vremena.

a) Na nivou značajnosti od 5%, ispitati da li starost gledalaca i slobodno vreme utiču na gledanost televizijskog programa **(10)**

b) Koliko iznosi procenat varijacija gledanosti televizijskog programa koji nije objašnjen zajedničkim uticajem starosti gledalaca i slobodnog vremena? **(5)**

c) Sa pouzdanošću od 95%, oceniti za koliko će se u proseku promeniti gledanost televizijskog programa, ako se starost poveća za 3 meseca, a slobodno vreme ostane nepromenjeno? **(10)**

- Fizički obim izvoza proizvoda (u hiljadama tona) sledi normalnu raspodelu sa sredinom 15 i varijansom 4. U prostom slučajnom uzorku od 25 proizvoda zabeležen je sledeći fizički obim izvoza: 11, 10, 19, 35, 29, 25, 19, 12, 16, 29, 35, 9, 10, 11, 25, 14, 15, 34, 30, 14, 12, 9, 12, 14, 10. Stopa izbora iznosi 10%.
 - Izračunati za koliko u proseku fizički obim izvoza pojedinih proizvoda odstupa od prosečnog fizičkog obima izvoza u uzorku. Interpretirati rezultat i navesti jedinice mere. **(10)**
 - Izračunati učešće i ukupan broj proizvoda u skupu čiji je fizički obim izvoza veći od 18 hiljada tona. **(5)**
 - Sa pouzdanošću od 95% formirati interval poverenja za učešće i ukupan broj svih proizvoda čiji je fizički obim izvoza veći od 18 hiljada tona. Da li je izabrani uzorak reprezentativan? **(15)**
- Cena proizvoda (u hiljadama dinara) sledi normalnu raspodelu. 62,93% svih proizvoda je jeftinije od 2,7 hiljada dinara, dok je 79,67% svih proizvoda skuplje od 2 hiljade dinara. Verovatnoća da se proporcija proizvoda koji su skuplji od 3 hiljade dinara u prostom slučajnom uzorku veličine n razlikuje od proporcije u osnovnom skupu za više od 0,02 iznosi 0,8026. Odrediti veličinu prostog slučajnog uzorka (n). **(20)**
- U jednom istraživanju je učestvovalo 200 ispitanika, koji su dali odgovore na pitanja o zaradi (u hiljadama dinara) i učestalosti polaganja depozita na deviznu štednju. Rezultati su prikazani u sledećoj tabeli:

		Učestalost polaganja depozita na deviznu štednju		
		Nikad	Jednom mesečno	Jednom godišnje
Zarada ispitanika	Ispod 40	16%	10%	6%
	40-80	8%	12%	16%
	Preko 80	6%	14%	12%

Na nivou značajnosti od 10% ispitati da li postoji povezanost između zarade ispitanika i učestalosti polaganja depozita na deviznu štednju. **(12)**

- Na osnovu raspoloživih podataka o uvozu sirovina (u hiljadama dinara) i ostvarenoj proizvodnji (u hiljadama tona) jednog proizvoda u prostom slučajnom uzorku od devet preduzeća dobijeni su sledeći rezultati: 1,15% varijacija proizvodnje nije objašnjeno regresionim modelom, prosečno kvadratno odstupanje uvoza sirovina pojedinih proizvoda od prosečnog uvoza sirovina iznosi 680,50 hiljada dinara², proizvodnja pojedinih proizvoda u proseku odstupa od ocenjene proizvodnje za 3.100 tona, prosečna proizvodnja iznosi 94,33 hiljade tona, prosečan uvoz sirovina je 64 hiljade dinara.
 - Izračunati sumu kvadrata odstupanja ocenjene proizvodnje od prosečne proizvodnje **(5)**
 - Oceniti parametre prostog linearnog regresionog modela. **(5)**
 - Sa pouzdanošću od 95% oceniti prosečnu proizvodnju ako uvoz sirovina iznosi 70 hiljada dinara. **(10)**
- U sledećoj tabeli su dati bazni indeksi prihoda od prodaje jednog proizvoda:

Godina	2012.	2013.	2014.	2015.	2016.	2017.	2018.
$I_g(2015=100)$	71,91	79,78	85,39	100,00	121,35	108,99	137,08

- U kojoj godini je ostvarena najveća procentualna promena prihoda u odnosu na 2017. godinu? Interpretirati dobijeni rezultat. **(8)**
- Koristeći geometrijsku sredinu lančanih indeksa, izračunati prosečnu stopu rasta prihoda u period 2012-2018. Interpretirati dobijeni rezultat. **(10)**

1. Raspored klijenata prema visini depozita na deviznu štednju (u €) u dve banke dat je u sledećoj tabeli:

Visina depozita	do 500	500-1000	1000-1500	preko 1500
Banka 1	10	12	9	5
Banka 2	5	14	11	6

- a) Uporediti varijabilitet visine depozita u obe banke **(10)**
- b) Ako visina depozita u prvoj sledi normalnu raspodelu sa sredinom 600€, sa pouzdanošću 95% formirati interval poverenja za prosečnu visinu depozita i ispitati da li je izabrani uzorak reprezentativan. **(12)**
- c) Izračunati koliko deposit u iznosu od 700€ u drugoj banci odstupa od prosečnog depozita u uzorku izraženo u standardnim devijacijama. **(4)**
2. a) Težina osoba koje praktikuju poseban režim ishrane sledi normalnu raspodelu. Najveći broj osoba ima 110 kg. Verovatnoća da težina bude veća od 120 kg iznosi 0,0228. Verovatnoća da učešće osoba čija je težina manja od 108 kg u prostom slučajnom uzorku veličine n bude veća od proporcije skupa za više od 0,2 iznosi 0,0174. Izračunaj n. **(14)**
- b) Analizira se potrošnja voća domaćinstva na mesečnom nivou. Prost slučajan uzorak od 80 domaćinstava, dao je sledeći rezultat:

Potrošnja (u kg)	Ispod 10	10-20	20-30	Preko 30
Broj domaćinstava	17	31	20	12

- Na nivou značajnosti od 5% ispitati da li je učešće svih domaćinstava čija je potrošnja veća od 20 kg., manje od 45%. **(10)**
3. Poznato je da od 1200 osoba koje se rekreativno bave sportom u sportskom centru, 36 osoba posećuje časove joge, 300 osoba ide na plivanje, 360 vežba u teretani, dok se preostale osobe bave aerobikom. Slučajan uzorak od 150 osoba pokazao je da 10% osoba posećuje jogu, 20% se bavi plivanjem, 26% vežba u teretani, a ostali idu na aerobik. Na nivou značajnosti od 1% ispitati da li se raspodela osoba prema sportu kojim se bave razlikuje od očekivane raspodele. **(12)**
4. Na osnovu raspoloživih podataka o ceni proizvoda (u 000 dinara) i tražnji za tim proizvodom (u 000 t) u prostom slučajnom uzorku od 8 proizvoda dobijeni su sledeći rezultati: tražnja za pojedinim proizvodima u proseku odstupa od prosečne tražnje za 25,47 hiljada tona, ocenjena standardna devijacija ocene regresionog parametra uz objašnjavajuću promenljivu iznosi 0,124 hiljada dinara, Pirsonov koeficijent korelacije između posmatranih promenljivih iznosi -0,964, ocenjena tražnja pri ceni od 45 hiljada dinara iznosi 67,42 hiljada tona, prosečna tražnja iznosi 72,63 hiljada tona.
- a) Izračunati procenat varijacije zavisne promenljive koji je pod uticajem rezidualnih faktora **(4)**
- b) Ocenite parametre prostog linearnog regresionog modela **(10)**
- c) Sa pouzdanošću od 95% predvideti tražnju za proizvodom, ako cena iznosi 55 hiljada dinara. **(6)**
5. Dati su lančani indeksi prodaje jednog proizvoda u period januar-oktobar 2018: 127,45; 113,85; 109,46; 109,88; 107,87; 114,58; 105,45; 107,76; 107,20; 111,19. Prodaja tog proizvoda (u hiljadama kg) u maju 2018. iznosila je 96 hiljada kg. Ocenjena funkcija linearnog trenda je $\hat{y}_t = 54,18 + 9,04t$, a srednja kvadratna greška linearnog trenda je 4,83 hiljada kg².
- a) Odabrati funkciju trenda koja se najbolje prilagođava podacima u posmatranom period (kao kriterijum koristiti srednju kvadratnu grešku). **(14)**
- b) Isključiti trend komponentu u mesecu u kom je ostvarena najveća stopa rasta. Objasniti rezultat. **(4)**

1. U jednom istraživanju je učestvovalo 200 ispitanika koji su dali odgovore na pitanja o zaradi (evri) i učestalosti polaganja depozita na deviznu štednju. Rezultati su dati u sledećoj tabeli:

		Učestalost polaganja depozita na deviznu štednju		
		Nikad	Jednom mesečno	Jednom godišnje
Zarada ispitanika (u €)	Ispod 1000	16%	10%	6%
	1000-2000	8%	12%	16%
	Preko 2000	6%	14%	12%

- a) Izračunati za koliko evra zarada ispitanika koji jednom godišnje polažu deposit na deviznu štednju odstupa od njihove prosečne zarade **(6)**
- b) Sa pouzdanošću od 95% oceniti učešće ispitanika sa zaradom ispod 1000€ koji ne polažu deposit na deviznu štednju u ukupnom broju ispitanika koji ne polažu deposit. **(10)**
- c) Na nivou značajnosti od 5% ispitati da li učešće ispitanika sa zaradom preko 2000€ koji više puta godišnje polažu deposit na deviznu štednju u ukupnom broju ispitanika sa zaradom preko 2000€ manje od 30% **(10)**
2. a) Težina osoba koje praktikuju poseban režim ishrane sledi normalnu raspodelu. Verovatnoća da težina bude manja od 115 kg iznosi 0,2033. Verovatnoća da težina bude veća od 123 kg iznosi 0,3085. Verovatnoća da prosečna težina osoba u prostom slučajnom uzorku veličine n bude veća od prosečne težine svih osoba za manje od 2 kg iznosi 0,9082. Izračunati veličinu prostog slučajnog uzorka (n). **(14)**
- b) Prost slučajan uzorak od 100 osoba dao je sledeće rezultate o njihovoj težini:

Težina (kg)	Broj osoba
Ispod 50	10
50-70	45
70-90	35
preko 90	10

Sa pouzdanošću od 95% formirati interval poverenja za prosečnu težinu svih osoba. **(10)**

3. Ispituje se razlika u prosečnom vremenu trajanja usluga u 4 osiguravajuće kuće. Iz svake kuće je izabran slučajan uzorak od 4 klijenta i zabeleženo je vreme koje je potrebno da svaki klijent bude uslužen:

I	II	III	IV
10	30	14	28
15	22	7	10
19	19	20	15
8	13	16	19

Na nivou značajnosti od 10% ispitati da li postoji razlika u prosečnom vremenu trajanja usluga između bar 2 kuće (sve potrebne pretpostavke su ispunjene). **(12)**

4. Na osnovu raspoloživih podataka o ceni (u 000 din) i obimu ponude (u 000 komada) jednog proizvoda u prostom slučajnom uzorku od 8 proizvoda dobijeni su sledeći rezultati: ocenjena standardna devijacija ocene regresionog parametra uz objašnjavajuću promenljivu je 164 dinara; prosek kvadrata odstupanja cene pojedinih proizvoda od prosečne cene iznosi 785,14 hiljada din^2 ; 16,68% varijacija zavisne promenljive je pod uticajem rezidualnih faktora; prosečna cena iznosi 55 hiljada dinara; prosečan obim ponude je 89,63 hiljada komada.
- a) Izračunati sumu kvadrata odstupanja ponude pojedinih proizvoda od ocenjene ponude. **(4)**
- b) Oceniti parametre prostog regresionog modela. **(10)**
- c) Sa pouzdanošću od 95% oceniti prosečnu ponudu proizvoda ako cena iznosi 50000 dinara. **(6)**
5. Dati su indeksi fizičkog obima proizvodnje jednog proizvoda po kvartalima 2016., 2017. i 2018. godine (prosek 2017=100): 89,69; 111,21; 159,64; 39,46; 86,10; 116,59; 150,67; 46,64; 98,65; 121,97; 168,61; 52,02. Proizvodnja pomenutog proizvoda u II kvartalu 2018. godine iznosila je 68000 tona. Sezonski indeksi proizvodnje po kvartalima u posmatrane 3 godine, iznosile su redom: 153,5; 89,5; 114,5; 42,5.
- a) Predvideti prosečan obim proizvodnje pomenutog proizvoda u kvartalu sa "mrtvom" sezonom u 2019. godini, uzimajući u obzir uticaj sezonskih faktora **(14)**
- b) Odrediti uticaj nepostojanih faktora u kvartalu sa "živom" sezonom. **(4)**

1. Iz jedne korporacije u kojoj radi 1500 zaposlenih, izabran je prost slučajni uzorak od 150 zaposlenih i dobijeni su sledeći rezultati:

zarada (u 000 din)	učestće zaposlenih
do 40	22%
40-80	...
80-120	34%
preko 120	18%

- a) Izračunati za koliko u proseku zarada pojedinih zaposlenih odstupa od njihove prosečne zarade u relativnom smislu.
 b) Na nivou značajnosti od 10% ispitati da li je učestće zaposlenih čija je zarada manja od 80 hiljada veće od 50%.
 c) sa poudanošću od 95% oceniti ukupnu i prosečnu zaradu svih zaposlenih.

2. Vreme koje je potrebno da bi klijent u banci A bio uslužen (izraženo u min) je normalno raspoređena slučajna promenljiva, a učestće klijenata kod kojih vreme čekanja traje manje od 7 minuta iznosi 84,13%. Vreme koje je potrebno da bi klijent u banci B bio uslužen (takođe izraženo u minutima) je normalno raspoređena slučajna promenljiva, a učestće klijenata kod kojih vreme čekanja traje više od 5 minuta iznosi 77,34%. Varijansa vremena čekanja u banci A je 4 puta manja od varijanse vremena čekanja u banci B. Relativna disperzija vremena čekanja u banci A je 1,25 puta manja od relativne disperzije vremena čekanja u banci B. Izračunati interkvartilnu razliku vremena čekanja u banci A i banci B.
3. Iz tri rukometna kluba u Srbiji slučajnim putem je izabrano po 6 igrača i zabeleženo vreme koje nedeljno provedu trenirajući (u časovima):

I klub: 14 12 16 18 20 16

II klub: 8 10 12 8 16 10

III klub: 6 8 6 10 8 7

Na nivou značajnosti od 5% ispitati da li se prosečno vreme treniranja igrača razlikuje između bar dva rukometna kluba (sve potrebne pretpostavke su ispunjene).

4. Ispituje se uticaj kretanja cene jednog proizvoda na ponudu tog proizvoda u prostom slučajnom uzorku od šest regiona. Raspolaze se podacima o ceni (u hiljadama dinara): 5, 9, 13, 18, 26, 38 i ponudi (u komadima): 12, 22, 30, 45, 61, 84.
- a) Izračunati sumu kvadrata odstupanja ocenjene ponude proizvoda u pojedinim regionima od prosečne ponude. Izračunati procenat varijacije ponude koji je pod uticajem rezidualnih faktora.
 b) Sa pouzdanošću od 95% oceniti prosečnu ponudu, ako cena iznosi 25 hiljada dinara. Da li je dobijena vrednost validna?
5. Dati su indeksi fizičkog obima proizvodnje jedne vrste voća, po kvartalima 2016., 2017. i 2018. godine (prosek 2017=100): 78,74; 119,69; 144,88; 56,69; 72,44; 125,98; 138,58; 62,99; 85,04; 135,43; 163,78; 50,39. Fizički obim proizvodnje voća u III kvartalu 2017. godine iznosio je 44 hiljade tona. Sezonski indeksi fizičkog obima proizvodnje, po kvartalima u posmatrane 3 godine, iznosili su redom: 75,5; 125; 141; 58,5.
- a) Predvideti prosečan fizički obim proizvodnje pomenutog voća u kvartalu sa "živom" sezonom u 2019. uzimajući u obzir uticaj sezonskih faktora.
 b) Odrediti uticaj nepostojanih faktora u kvartalu sa "mrtvom" sezonom.

1. Zarada zaposlenih u jednom preduzeće sledi normalnu raspodelu sa sredinom 55 hiljada dinara i standardnom devijacijom 2 hiljade dinara. Slučajnim putem je izabrano 15 zaposlenih i zabeležena je visina njihove zarade u prošlom mesecu (u hiljadama dinara): 35, 31, 96, 40, 54, 37, 48, 28, 42, 57, 77, 64, 51, 84, 39.

a) Izračunati mere centralne tendencije u izabranom uzorku.

(5)

b) Odrediti verovatnoću da se proporcija zaposlenih čija je plata veća od 60 hiljada dinara u prostom slučajnom uzorku od 15 zaposlenih razlikuje od proporcije osnovnog skupa za više od 0,04.

(10)

c) Ako je zarada jednog zaposlenog u uzorku 40 hiljada dinara, koliko iznosi odstupanje te zarade od prosečne zarade u uzorku izraženo u standardnim devijacijama?

(5)

2. U jednom istraživanju je učestvovalo 100 kupaca, koji su dali odgovore na pitanja o starosti i potrošnji (u hiljadama dinara) u toku jedne nedelje. Rezultati su prikazani u sledećoj tabeli:

		Potrošnja (u 000 din)		
		Ispod 15	15-30	Preko 30
Godine starosti kupca	10-30	20	8	4
	30-50	15	19	5
	preko 50	9	12	8

a) Izračunati za koliko u proseku starost pojedinih kupaca koji su potrošili više od 30 hiljada dinara odstupa od prosečne starosti te grupe kupaca, u apsolutnom i relativnom iznosu.

(10)

b) Sa pouzadnošću od 95% oceniti učešće kupaca, u ukupnom broju od 100 ispitanika, starijih od 30 godina koji su u toku jedne nedelje potrošili više od 15 hiljada dinara.

(10)

c) Na nivou značajnosti od 5% ispitati da li je prosečna potrošnja kupaca starosti od 10-30 godina manja od 10 hiljada dinara.

(10)

3. U jednom istraživanju je učestvovalo 100 kupaca, koji su dali odgovore na pisanja o starosti i potrošnji (u hiljadama dinara) u toku jedne nedelje. Rezultati su prikazani u sledećoj tabeli:

		Vrednost kupovine (u 000 din)		
		Ispod 15	15-30	Preko 30
Godine starosti kupca	10-30	20	8	4
	30-50	15	19	5
	preko 50	9	12	8

Na nivou značajnosti od 10% Ispitati da li postoji povezanost između starosti ispitanika i potrošnje.

(15)

4. U prostom slučajnom uzorku od 12 studenata dobijeni su sledeći rezultati o visini mesečne stipendije (u hiljadama dinara) i mesečnim troškovima (u hiljadama dinara): prosečni mesečni troškovi su 14,33 hiljade dinara; prosečne mesečne stipendije iznose 24,58 hiljada dinara, suma kvadrata odstupanja ocenjenih mesečnih troškova od prosečnih mesečnih troškova iznosi 689,07 hiljada dinara²; mesečna stipendija pojedinih studenata u proseku odstupa od prosečne mesečne stipendije za 10.440 dinara; suma kvadrata odstupanja mesečnih troškova pojedinih studenata od ocenjenih mesečnih troškova iznosi 187,6 hiljada dinara².

a) Izračunati procenat varijacija mesečnih troškova koji je objašnjen mesečnom stipendijom

(4)

b) Oceniti parametre prostog linearnog regresionog modela i objasniti dobijeni regresioni koeficijent nagiba.

(8)

c) Sa pouzadnošću od 95% predvideti mesečne troškove ako mesečna stipendija iznosi 20 hiljada dinara.

(8)

5. Sledeća tabela sadrži kvartalne podatke o stanju na zalihama televizora u magacinu jedne prodavnice u periodu od 2014-2017. godine:

Godina	2014.		2015.				2016.				2017.	
Kvartal	III	IV	I	II	III	IV	I	II	III	IV	I	II
Stanje na zalihama	25	20	16	21	24	15	13	19	21	18	14	22

Ispitati da li posmatranu seriju karakterišu izražene sezonske varijacije i isključiti uticaj sezone u kvartalu sa "mrtvom" sezonom.

(15)

1. Dati su podaci o broju pročitanih knjiga u slučajnom uzorku putnika na dve destinacije:

Broj pročitanih knjiga	1	2	3	4	5
I destinacija	13	10	4	6	3
II destinacija	15	8	6	5	2

a) Odrediti pozicione mere centralne tendencije za obe destinacije i uporediti relativne mere disperzije broja pročitanih knjiga ove dve destinacije **(13)**

b) Sa pouzdanosti od 95% oceniti prosečan broj knjiga koje su pročitali ispitanici na prvoj destinaciji. **(10)**

2. Težina jednog proizvoda sledi normalnu raspodelu sa sredinom 36 kg i standardnom devijacijom 4 kg.

a) Od 500 proizvoda slučajnim putem je izabran uzorak od 36 proizvoda čija raspodela je prikazana u sledećoj tabeli:

Težina proizvoda (u kg)	Relativna frekvencija
Ispod 15	0,25
15-30	0,30
30-45	0,15
preko 45	0,30

Na nivou značajnosti od 5% ispitati da li je učešće proizvoda čija je težina veća od 30 kg manje od 40%. **(10)**

b) Naći verovatnoću da je učešće proizvoda sa težinom manjom od 40 kg u uzorku od 36 proizvoda veće od 85%. **(5)**

c) Naći veličinu uzorka (n), ako je verovatnoća da se prosečna težina u uzorku razlikuje od prosečne težine svih za više od 1 kg jednaka 0,2112. **(12)**

3. Iz tri prodavnice slučajnim putem je izabrano po pet proizvoda i zabeleženo vreme potrebno za njihovu izradu (u časovima)

I prodavnica	II prodavnica	III prodavnica
10	13	11
12	12	13
8	14	9
8	15	8
10	11	12

Na nivou značajnosti od 1% ispitati da li se prosečno vreme potrebno za izradu proizvoda razlikuje između bar dve prodavnice (sve potrebne pretpostavke su ispunjene). **(15)**

4. Na osnovu raspoloživih podataka o ceni (u hiljadama dinara) i tražnji za tim proizvodom (u hiljadama komada) u prostom slučajnom uzorku od osam proizvoda dobijeni su sledeći rezultati: prosečna tražnja iznosi 43,63 hiljada komada, a prosečna cena 26,88 hiljada dinara; ocenjena standardna greška ocene koeficijenta nagiba je 164 dinara, a 12,61% varijacija tražnje je pod uticajem rezidualnih faktora, dok cena pojedinih proizvoda u proseku odstupa od prosečne cene za 12,87 hiljada dinara.

a) Izračunati sumu kvadrata odstupanja tražnje za pojedinim proizvodima od linije regresije u uzorku. **(4)**

b) Oceniti parametre prostog linearnog regresionog modela i objasniti dobijeni regresioni koeficijent nagiba. **(8)**

c) Sa pouzdanosti od 95% oceniti prosečnu tražnju ako cena iznosi 15 hiljada dinara. **(8)**

5. Dati su lančani indeksi proizvodnje jedne marke računara u period 2011.-2017. godine: 110; 127,27; 121,43; 117,65; 115; 119,57; 127,27.

a) Oceniti eksponencijalnu funkciju trenda u period 2011.-2017. godine, ako je proizvodnja u 2015. Godini iznosila 4,6 miliona komada. **(10)**

b) Predvideti proizvodnju računara u 2019. godini. Da li je dobijena vrednost validna? **(5)**

1.

- a) Troškovi proizvodnje proizvoda slede normalnu raspodelu sa sredinom 35 hiljada dinara. Učešće proizvoda čiji su troškovi proizvodnje između 32 hiljade i 38 hiljada dinara iznosi 68,26%.
- 1) Za koliko dinara troškovi pojedinih proizvoda u proseku odstupaju od prosečnih troškova proizvodnje? **(8)**
 - 2) Odrediti oblik uzoračke raspodele statistike $\hat{P}$, gde $\hat{P}$ predstavlja proporciju onih proizvoda u uzorku veličine 25 čiji su troškovi proizvodnje veći od 36 hiljada dinara. **(4)**
 - 3) Izračunati verovatnoću da proporcija onih proizvoda u uzorku veličine 25 čiji su troškovi proizvodnje veći od 38 hiljada dinara bude manja od 30%? **(4)**
- b) U prostom slučajnom uzorku od 25 proizvoda zabeležene su sledeće vrednosti troškova proizvodnje (u hiljadama dinara):

Troškovi proizvodnje (u hiljadama dinara)	Broj proizvoda
Ispod 20	7
20-40	10
40-60	8

Izračunati apsolutnu meru disperzije raspodele broja proizvoda prema troškovima proizvodnje. Interpretirajte dobijeni rezultat.

2. Površina novih stanova (u m²) sledi normalnu raspodelu sa sredinom 65 m² i standardnom devijacijom 3 m². U prostom slučajnom uzorku od 16 stanova zabeležena je njihova površina (u m²): 90, 35, 65, 42, 78, 55, 58, 105, 38, 74, 50, 80, 86, 30, 62, 120.
- a) Na nivou značajnosti od 5% ispitati da li se učešće stanova koji su manji od 60 m² razlikuje od 45%. **(10)**
 - b) Koliko iznosi realizovana vrednost statistike $\bar{X}$, gde $\bar{X}$ predstavlja prosečnu površinu stanova u uzorku veličine 16? **(4)**
 - c) Sa pouzdanosti od 95% formirati interval poverenja za prosečnu površinu stanova u skupu svih stanova. **(5)**
3. Iz četiri tržna centra slučajnim putem je izabano po pet ispitanika i zabeležen je iznos novca koji su potrošili u tom tržnom centru u toku jedne sedmice (u hiljadama dinara). Na osnovu dobijenih podataka izračunata je faktorska suma kvadrata koja iznosi 600 hiljada dinara² i rezidualna varijansa koja iznosi 40 hiljada dinara². Popunite datu tabelu i na nivou značajnosti od 5% ispitajte da li se prosečna potrošnja ispitanika razlikuje između bar dva tržna centra (sve potrebne pretpostavke su ispunjene). **(14)**

H_0 : _____
 H_1 : _____

Izvor varijacija	Broj stepeni slobode	Suma kvadrata	Varijansa	Realizovana vrednost statistike testa
Između uzoraka				
Unutar uzoraka				
Ukupno				

Odgovor:

4. Raspoložemo podacima o godinama radnog iskustva i mesečnim zaradama (u hiljadama dinara) u prostom slučajnom uzorku od deset ispitanika (ispitanik koji radi godinu dana ima najmanje radnog iskustva, dok ispitanik koji radi 25 godina ima najviše radnog iskustva; najniža i najviša mesečna zarada iznose 50.000 i 120.000 dinara). Na osnovu raspoloživih podataka dobijeni su sledeći međurezultati:
 $\sum x = 102, \sum y = 750, \sum x^2 = 1604, \sum y^2 = 60600, \sum xy = 9200$.
- a) Na nivou značajnosti od 5% ispitati da li godine radnog iskustva pozitivno utiču na mesečnu zaradu u skupu. **(13)**
 - b) Sa pouzdanosti od 95% predvideti mesečnu zaradu slučajno odabranog ispitanika sa radnim iskustvom od 20 godina. Da li je dobijena vrednost validna? **(5)**
5. Poznati su lančani indeksi fizičkog obima prodaje jednog proizvoda u periodu od marta do decembra 2019. godine: 112,50; 106,67; 108,33; 113,46; 103,39; 109,84; 110,45; 110,81; 109,76; 106,67.
- a) Na osnovu serije lančanih indeksa, izračunati prosečnu mesečnu procentualnu promenu fizičkog obima prodaje u periodu od marta do decembra 2019. godine. **(5)**
 - b) U kom mesecu je ostvarena najveća, a u kom najmanja procentualna promena fizičkog obima prodaje u odnosu na maj 2019. godine? **(12)**

1. Poznata je vrednost prodaje jednog proizvoda (u hiljadama dinara) u svim mesecima 2019. godine: 21, 27, 35, 41, 55, 86, 93, 109, 115, 120, 135, 139. Ocenjena funkcija eksponencijalnog trenda u periodu od januara do decembra 2019. godine je $\bar{y}_i = 4,66 \cdot 1,15^i$, a srednja kvadratna greška 53,62 hiljada dinara².

a) Odabrati funkciju trenda koja se najbolje prilagođava podacima o posmatranom periodu (kao kriterijum koristiti srednju kvadratnu grešku)

b) Isključiti trend komponentu iz date vremenske serije za svaki mesec drugog kvartala 2019. godine i prokomentarisati rezultat.

c) Predvideti vrednost prodaje posmatranog proizvoda u decembru 2020. godine. Da li je dobijena prognoza validna?

2. a) U sledećoj tabeli prikazana je raspodela slučajno izabranih porodica prema broju automobila:

Broj automobila	0	1	2	3
Broj porodica	2	10	7	6

75% porodica ima _____ ili manje automobila (prikazati postupak izračunavanja).

b) Vreme koje je potrebno ispitanicima da popune upitnik je normalno raspodeljeno sa sredinom 100 minuta i standardnom devijacijom 5 min.

(1) Neka se u prostom slučaju uzorku od n ispitanika analiziraju oni koji su popunjavali upitnik duže od 97 minuta. Verovatnoća da procenat takvih bude manji od 55% iznosi 0,0244. Odrediti veličinu prostog slučajnog uzorka (n).

(2) Odrediti interval varijacije promenljive vreme koje je potrebno ispitanicima da popune upitnik.

3. U tri grupe studenata kod kojih su nastavnici koristili različite metode pripreme ispita uspeh je meren ocenom na ispitu. Iz prve grupe na slučajan način je izabrano 6 studenata, iz druge grupe 4 studenata, dok je iz treće grupe izabrano 5 studenata. Na osnovu zabeleženih ocena izračunata je rezidualna suma kvadrata koja iznosi 400 i faktorska varijansa koja iznosi 80. Popunite datu tabelu i na nivou značajnosti od 5% ispitajte da li između bar dve grupe studenata postoji razlika u prosečnom uspehu (sve potrebne pretpostavke su ispunjene).

H_0 : _____

H_1 : _____

Izvor varijacija	Broj stepeni slobode	Suma kvadrata	Varijansa	Realizovana vrednost statistike testa
Između uzoraka				
Unutar uzoraka				
Ukupno				

4. a)

1) Mesečni izdaci za hranu (u hiljadama dinara) četvoročlanih domaćinstava slede normalnu raspodelu sa nepoznatom sredinom i varijansom. Konstruisan je 95% interval poverenja za prosečne mesečne izdatke svih porodica i njegovih širina iznosi 3.248 dinara. Odrediti standardnu devijaciju prostog slučajnog uzorka.

2) Na osnovu informacija iz dela zadatka 4. a) 1) izračunati marginalnu grešku za nivo pouzdanosti od 90%.

b) Neto gotovinski priliv (u milionima dolara) sledi normalnu raspodelu sa sredinom 2,5 miliona dolara i varijansom 0,4 miliona dolara². U prostom slučajnom uzorku od 16 investicionih fondova zabeleženi su sledeći neto gotovinski prilivi: 1,5; 1,6; 1,9; 2,0; 2,0; 2,2; 2,4; 2,5; 2,6; 2,6; 2,8; 3,1; 3,1; 3,3; 3,5; 3,7. Na nivou značajnosti od 5%, ispitati da li je učešće investicionih fondova čiju su neto gotovinski prilivi manji od 2,6 miliona dolara veće od 45%.

I etapa

II etapa

III etapa

IV etapa

V etapa

5. Raspoložemo podacima o ceni (u hiljadama dinara) i tražnji za proizvodom u prostom slučajnom uzorku od deset regiona. Na osnovu raspoloživih podataka dobijeni su sledeći međurezultati:

$$\sum x = 385, \sum y = 433, \sum x^2 = 18279, \sum y^2 = 24425, \sum xy = 12434.$$

a) Na nivou značajnosti od 5% ispitati da li u skupu postoji linearno kvantitativno slaganje između cene i tražnje za proizvodom.

b) Sa pouzdanošću od 95% predvideti tražnju za proizvodom u regionu u kojem je cena posmatranog proizvoda 22.000 dinara.

1.

- a) Mesečni troškovi za hranu tročlanih porodica (u hiljadama dinara) slede normalnu raspodelu sa sredinom $\mu=35$ hiljada dinara. Verovatnoća da mesečni troškovi za hranu budu veći od μ za više od 2 hiljade dinara iznosi 0,1587
- 1) Za koliko dinara mesečni troškovi za hranu pojedinih porodica u proseku odstupaju od prosečnih mesečnih troškova za hranu.
 - 2) Odrediti oblik uzoračke raspodele statistike $\bar{X}$, gde $\bar{X}$ predstavlja prosečne mesečne troškove za hranu u uzorku veličine 25. Obrazložiti odgovor.
 - 3) Izračunati verovatnoću da se prosečni mesečni troškovi za hranu u uzorku veličine 25 razlikuju od prosečnih mesečnih troškova za hranu u osnovnom skupu za manje od 5 hiljada dinara.
- b) U prostom slučajnom uzorku od 36 porodica zabeleženi su sledeći mesečni troškovi za hranu (u hiljadama dinara):

Mesečni troškovi za hranu (u hiljadama dinara)	Broj porodica
Ispod 20	9
20-40	12
40-60	15

Izračunati relativnu meru disperzije raspodele broja porodica prema mesečnim troškovima za hranu interpretirati dobijeni rezultat.

2. Iznos koji vozači automobila potroše na kupovinu goriva (u hiljadama dinara) sledi normalnu raspodelu prostom slučajnom uzorku od 16 vozača zabeleženi su sledeći iznosi: 8, 20, 15, 13, 16, 5, 3, 25, 16, 10, 12, 14, 6, 8, 7, 11.

Poznato je da je $\sum x^2=2739$

- a) Na nivou značajnosti od 3% ispitati da li je prosečan iznos koji vozači automobila potroše na kupovinu goriva veći od 10 hiljada dinara

I etapa

II etapa

III etapa

IV etapa

V etapa

- b) Koliko iznosi realizovana vrednost statistike $\hat{P}$ gde $\hat{P}$ predstavlja proporciju onih vozača u uzorku veličine 16 koji na kupovinu goriva potroše manje od 10 hiljada dinara?
- c) Koliko iznosi ocenjena vrednost standardne greške statistike $\hat{P}$ gde $\hat{P}$ predstavlja proporciju vozača u uzorku veličine 16 koji na kupovinu goriva potroše manje od 10 hiljada dinara?
- d) Sa pouzdanošću od 95% formirati interval poverenja za učešće svih vozača koji na kupovinu goriva troše manje od 10 hiljada dinara.

3. Istraživanje sprovedeno u skupu od 500 studenata jednog smera na jednom fakultetu u 2017. godini pokazalo je da je 50 studenata ocenilo kvalitet tog smera ocenom dva, 150 studenata je dalo ocenu tri, 250 studenata je dalo ocenu četiri, dok su preostali studenti dali ocenu pet. Istraživanje sprovedeno u 2020. godini na osnovu slučajnog uzorka od 50 studenata istog smera pokazalo je da je 10 studenata ocenilo kvalitet smera ocenom dva, 12 studenata je dalo ocenu tri, 20 studenata je dalo ocenu četiri dok su preostali studenti dali ocenu pet. Na nivou značajnosti od 5% ispitati da li raspodela studenata u 2020. prema oceni kojom vrednuju smer odstupa od raspodele studenata u 2017. godini.

I etapa

II etapa

III etapa

IV etapa

V etapa

4. Raspoložimo podacima o površini stana (u m²) i mesečnom zakupu (u evrima) u prostom slučajnom uzorku od deset stanova. Na osnovu raspoloživih podataka dobijeni su sledeći međurezultati:
 $\sum x = 559$, $\sum y = 2010$, $\sum x^2 = 35759$, $\sum y^2 = 482500$, $\sum xy = 130680$
- Na nivou značajnosti od 5% ispitati da li u skupu postoji pozitivno linearno kvantitativno slaganje između površine stana i mesečnog zakupa.
 - Sa pouzdanošću od 95% oceniti prosečni mesečni zakup svih stanova čija je površina 60m².

5. Poznati su bazni indeksi fizičkog obima prodaje jednog proizvoda u periodu od marta do decembra 2019. godine (jul 2019=100): 46,51; 55,81; 72,09; 81,40; 100,00; 116,28; 137,21; 158,14; 172,09; 204,65.
- Koristeći seriju baznih indeksa izračunati prosečnu mesečnu procentualnu promenu fizičkog obima prodaje u periodu od marta do decembra 2019. godine.
 - Koliko iznosi procentualnu promenu fizičkog obima prodaje u svakom mesecu u periodu od marta do decembra 2019. godine u odnosu na prethodi mesec?

ПРИМЕРИ ИСПИТНИХ ЗАДАТАКА

(Дескриптивне мере, расподела вероватноћа, узорачка расподела,
интервално оцењивање, тестирање хипотеза на једном и на два узорка)

1. (Jun 2015) Stipendije studenata jednog fakulteta (u hiljadama dinara) slede normalnu raspodelu sa sredinom 12 i varijansom 2. U prostom slučajnom uzorku od 25 studenata zabeleženi su iznosi stipendija koje primaju: 7, 5, 15, 30, 25, 20, 15, 7, 7, 25, 30, 5, 5, 7, 20, 10, 10, 30, 25, 10, 7, 5, 7, 10, 5. Stopa izbora iznosi 6,25%.

a) Izračunati učešće i ukupan broj studenata u skupu čija je stipendija veća od 15 hiljada. (5)

b) Sa pouzdanošću od 95% formirati interval poverenja za učešće svih studenata čija je stipendija veća od 15 hiljada. (10)

2. (Jun 2015) Slučajnim putem je izabrano 70 ispitanika i zabeležen je broj članova u njihovim porodicama:

Broj članova u porodici	1	2	3	4	5
Broj ispitanika	7	16	25	19	3

a) Izračunati i interpretirati apsolutnu i relativnu meru disperzije gornje raspodele. (10)

b) Sa pouzdanošću od 95% oceniti prosečan broj članova u porodicama svih ispitanika. (10)

3. (Jun 2015) Raspodela 200 kupaca u jednoj prodavnici zdrave hrane tokom jedne nedelje, data je sledećom tabelom:

Starost kupca	Pol	
	Muški	Ženski
do 20	15	23
20-30	20	28
30-40	18	39
preko 40	27	30

a) Uporediti relativnu disperziju starosti kupaca muškog i ženskog pola. (10)

b) Na nivou značajnosti od 5%, ispitati da li je prosečna starost kupaca muškog pola veća od prosečne starosti kupaca ženskog pola. (10)

4. (Jul 2017) У једном истраживању је учествовало 80 испитаника, који су дали одговоре на питања о месечним примањима (у 000 динара) и о учесталости посећивања биоскопа у току једног месеца:

		Месечна примања			
		до 20	20-40	40-60	преко 60
Број посета биоскопу у току месеца	Ниједном	12	10	7	4
	1-3 пута	4	5	9	7
	Више од 3 пута	2	3	6	11

a) Са поузданошћу од 90% оценити учешће испитаника чија су примања од 20-60 хиљада динара и који посећују биоскоп у току месеца. (10)

b) На нивоу значајности од 5%, испитати да ли се учешће испитаника чија су месечна примања до 60 хиљада динара, а који ниједном нису посетили биоскоп у току месеца разликује од 40%. (10)

5. (Jun 2015) U jednom istraživanju je učestvovalo 200 kupaca, koji su dali odgovore na pitanja o starosti i učestalosti kupovine u prodavnicama zdrave hrane u toku jedne nedelje. Rezulati su prikazani u sledećoj tabeli:

		Učestalost kupovine u toku nedelje		
		Nikad	Jednom nedeljno	Više od tri puta nedeljno
Godine starosti kupca	do 20	8	10	31
	20-30	11	18	24
	30-40	17	25	16
	preko 40	28	7	5

a) Sa pouzdanošću od 95% oceniti prosečnu starost kupaca koji u prodavnicama zdrave hrane kupuju više od tri puta nedeljno. (10)

b) Na nivou značajnosti od 5%, ispitati da li postoji povezanost između starosti kupaca i učestalosti kupovine. (15)

6. (Септембар 2017) Зараде запослених у једном предузећу следе нормалну расподелу са средином 38 хиљада динара. Од 120 запослених изабран је прост случајан узорак од 15 запослених чија је зарада у претходном месецу (у хиљадама динара) износила: 28, 14, 62, 37, 41, 19, 30, 36, 65, 70, 24, 44, 57, 48, 23.

a) Са поузданошћу од 95% утврдити да ли је изабрани узорак репрезентативан при оцени просечне зараде свих запослених. (10)

b) Израчунати за колико у просеку зараде појединих запослених посматраног узорка одступају од њихове просечне месечне зараде у апсолутном и релативном износу? (10)

7. (Jun 2015) Jedno istraživanje je pokazalo da bi 46% ispitanika češće posećivalo pozorište da su ulaznice jeftinije. Pretpostavimo da ovaj rezultat važi za populaciju odraslih osoba, iz koje su slučajnim putem izabrane tri odrasle osobe. Neka je X broj odraslih osoba koji imaju gore navedeno mišljenje.

- a) Nacrtati stablo ishoda i odrediti raspodelu verovatnoća za X . (10)
b) Koliko iznosi verovatnoća da će od tri izabrane osobe tačno jedna imati gore navedeno mišljenje? (10)

8. (Jul 2017) Анализирају се месечна примања домаћинстава у једној општини. Из основног скупа од 1000 домаћинстава изабран је прост случајан узорак од 80 домаћинстава и добијени су следећи резултати:

Примања (у 000 динара)	испод 30	30-60	60-90	преко 90
Број домаћинстава	17	29	21	13

- a) Израчунати и интерпретирати апсолутну и релативну меру дисперзије горње расподеле. (10)
b) Са поузданошћу од 95% формирати интервал поверења за просечна и укупна примања свих домаћинстава. (10)

9. (Jul 2015) Прост случајан узорак од 150 особа дао је следећи резултат:

Visina (u cm)	Broj osoba
ispod 150	15
150-170	65
170-190	60
preko 190	10

- a) На нивоу значајности од 5%, испитати да ли је уčešће свих особа висине до 190 cm веће од 85%. (10)
b) Са поузданошћу од 95% формирати интервал поверења за просечну висину свих особа. (10)

10. (Jul 2016) 1. U sledećoj tabeli je prikazana raspodela slučajno izabranih studenata jednog fakulteta prema oceni na ispitu iz Ekonometrije i Teorijske statistike:

Ocena	5	6	7	8	9	10
Ekonometrija	11	5	9	7	4	2
Teorijska statistika	7	8	8	4	3	1

Na nivou značajnosti od 5%, ispitati da li je prosečna ocena studenata koji su položili ispit iz Ekonometrije manja od prosečne ocene studenata koji su položili ispit iz Teorijske statistike (sve potrebne pretpostavke su ispunjene). (15)

11. (Jul 2015) U sledećoj tabeli je prikazana raspodela slučajno izabranih studenata 1 godine jednog Ekonomskog fakulteta prema oceni na ispitu iz Matematike i Osnova statističke analize:

Ocena	5	6	7	8	9	10
Matematika	9	8	14	10	6	3
Osnovi statističke analize	15	9	13	6	5	2

- a) На нивоу значајности од 5%, испитати да ли је уčešће студента који су положили испит из Matematike веће од 75%. (10)
b) Uporediti interkvartilnu razliku ocena iz Matematike i ocena iz Osnova statističke analize. (10)

12. (Jul 2015) U sledećoj tabeli je prikazana raspodela 150 slučajno izabranih radnika jednog preduzeća prema dužini radnog staža (u godinama) i broju ostvarenih bonusa:

Dužina radnog staža	Broj ostvarenih bonusa		
	ispod 5	od 5 do 15	preko 15
ispod 10	8	12	30
od 10 do 20	15	7	19
preko 20	22	26	11

- a) Са поузданошћу од 95%, формирати интервал поверења за уčešће радника који у поменутом предузећу раде више од 10 година. (10)
b) На нивоу значајности од 5%, испитати да ли постоји веза између дужине радног стажа и броја остvарених бонуса. (15)

13. (Jul 2015) Plate zaposlenih (u 000 dinara) u jednom preduzeću slede normalnu raspodelu sa sredinom 40 hiljada dinara. Verovatnoća da plata radnika bude manja od prosečne plate za više od 4 hiljade dinara iznosi 0,1587. Verovatnoća da se u prostom slučajnom uzorku veličine n prosečna plata razlikuje od prosečne plate u osnovnom skupu za više od hiljadu dinara iznosi 0,3174. Izračunati verovatnoću da u prostom slučajnom uzorku veličine n proporcija zaposlenih čija je plata manja od 38 hiljada dinara bude veća od 80%. (20)

14. (Jun 2017) Време које је потребно за израду производа А (изражено у часовима) је нормално расподељено случајна променљива, а учешће производа А за чију је израду потребно више од 6 часова износи 40,13%. Време које је потребно за израду производа Б (у часовима) је нормално расподељено, а учешће производа Б за чију је израду потребно мање од 4 часа је 14,69%. Варијанса времена израде производа Б је четири пута мања од варијансе времена израде производа А, док је релативна дисперзија времена израде тог производа 2,76 пута већа него за производ Б. Одредити просечна времена потребна за израду производа А и Б. (20)

15. (Januar 2017) Slučajnim putem izabrano je 50 zaposlenih osoba u jednoj banci i u sledećoj tabeli prikazana je raspodela njihovih zarada:

Zarada (u hiljadama dinara)	Učešće zaposlenih
do 20	24%
20-40	40%
40-60	...
preko 60	16%

- a) Na nivou značajnosti od 5% ispitati da li je učešće zaposlenih sa zaradom većom od 40000 dinara manje od 40%. (10)
- b) Sa pouzdanošću od 95% formirati interval poverenja za prosečnu zaradu svih zaposlenih. (10)

16. (Septembar 2015) Ispituje se sklonost kupaca prema kupovini jedne vrste bezalkoholnog pića. Na osnovu prostog slučajnog uzorka od 50 ispitanika dobijeni su sledeći rezultati:

Mesečni prihod (u hiljadama dinara)	Sklonost kupovini		
	svaki dan	veoma retko	nikad
do 20	4%	12%	18%
20-40	10%	12%	14%
40-60	16%	8%	6%

Na nivou značajnosti od 5%, ispitati da li postoji razlika u prosečnom mesečnom prihodu između kupaca koji kupuju svaki dan i onih koji veoma retko kupuju pomenutu vrstu bezalkoholnog pića (pretpostavlja se da su ispunjene sve neophodne pretpostavke). (15)

17. (Октобар 2017) Zarada zaposlenih (u 000 dinara) u jednom preduzeću sledi normalnu raspodelu sa sredinom 40 hiljada dinara. Verovatnoća da zarada zaposlenih bude veća od prosečne zarade za više od 6 hiljada dinara iznosi 0,0228. Verovatnoća da se u prostom slučajnom uzorku veličine n prosečna zarada razlikuje od prosečne zarade u osnovnom skupu za više od 1,5 hiljada dinara iznosi 0,0124. Izračunati verovatnoću da u prostom slučajnom uzorku veličine n proporcija zaposlenih čija je zarada manja od 43 hiljade dinara bude veća od 70%. (20)

18. (Октобар 2015) Дневна вредност појединачне куповине у једној продавници је нормално расподељена случајна променљива са аритметичком средином 980 динара и стандардном девијацијом 20 динара. Изабран је прост случајан узорак од 12 купаца и забележене су вредности њихових куповина (у стотинама динара): 16, 9, 13, 18, 6, 4, 8, 12, 17, 9, 25, 11.

- a) Ако у основном скупу има 300 купаца, израчунати тачно учешће и укупан број купаца чија је вредност куповине већа од 1000 динара. (10)
- b) Са поузданошћу од 95% утврдити дали је узорак репрезентативан при оцени средине скупа. (10)
- в) За колико куповина у износу од 1800 динара одступа од просечне вредности куповине у узорку изражено у стандардним девијацијама? (5)

19. (Јун 2016) Slučajna promenljiva X ima normalnu raspodelu sa parametrima μ i σ . Poznato je da se 75% površine ispod krive nalazi ulevo od broja 6,4 i da je $P(X > \mu - 1,2) = 0,9772$.

Verovatnoća da prosek slučajne promenljive X u prostom slučajnom uzorku veličine n bude veći od μ za više od 0,1 iznosi 0,1587. Izračunati verovatnoću da proporcija vrednosti koje su manje od 5,7 u prostom slučajnom uzorku veličine n bude veća od proporcije tih vrednosti u osnovnom skupu za najviše 20%. (20)

20. (Октобар 2015) Prečnik kuglica za novogodišnju jelku ima normalnu raspodelu. U prostom slučajnom uzorku od 36 kuglica dobijen je sledeći rezultat:

Prečnik kuglica (cm)	Broj proizvoda
2-4	10
4-6	12
6-8	8
8-10	6

- a) Na nivou značajnosti od 5%, ispitati da li je učešće kuglica u osnovnom skupu sa prečnikom 4-8 cm manje od 60%. (10)
- b) Učešće kuglica u osnovnom skupu sa prečnikom većim od 8cm iznosi 10,56%. Verovatnoća da u prostom slučajnom uzorku od 36 kuglica prosečan prečnik bude manji od 7 cm iznosi 0,9938. Odrediti prosečan prečnik svih kuglica. (10)

21. (Јул 2017) Једно истраживање је показало да 28% испитаника послује на берзи. Нека се овај резултат односи на тренутну популацију. Претпоставимо да су четири особе изабране случајним путем. Нека је X број особа које послују на берзи.

- а) Одредити расподелу вероватноћа за X и нацртати стабло исхода. (15)
 б) Израчунати вероватноћу да ће од четири изабране особе тачно две пословати на берзи. (5)

22. (Септембар 2017) Од 150 студената једног факултета изабран је прост случајан узорак од 15 студената и забележено је време које су посветили учењу страног језика у току једне недеље (у часовима): 8, 10, 13, 4, 12, 7, 16, 5, 2, 9, 7, 3, 11, 6, 5.

- а) Са поузданошћу од 95% оценили учење и укупан број студената који су у току једне недеље провели мање од 7 часова у учењу страног језика. (10)
 б) На нивоу значајности од 5% испитати да ли је учење студената који су у току једне недеље провели више од 8 сати у учењу страног језика веће од 35%. (10)

23. (Септембар 2017) Плата менаџера у једном предузећу је нормално расподељена случајна променљива, а учење менаџера чије су плате веће од 1140 евра износи 97,72%. Плата радника у истом предузећу је такође нормално расподељена, а учење радника чије су плате мање од 430 евра износи 93,32%. Варијанса плата менаџера је 2,25 пута већа од варијансе плата радника, док је релативна дисперзија плата менаџера два пута мања него у случају плата радника.

- а) Израчунати средине и варијансе плата менаџера и радника. (15)
 б) Одредити облик расподеле разлике плата менаџера и радника. (5)

24. (Јул 2016) Тежина једног производа следи нормалну расподелу. Половина производа је тежа од 15kg. Вероватноћа да тежина производа буде између 12kg и 18kg износи 0,8664. Вероватноћа да се просечна тежина производа у простом случајном узorku величине n разликује од просечне тежине свих производа за више од 1kg износи 0,0124. Одредити величину простог случајног узorka (n). (20)

25. (Јул 2016) У следећој табели је приказана расподела случајно изабраних студената једног факултета према оцени на испиту из Економетрије и Теоријске статистике:

Оцена	5	6	7	8	9	10
Економетрија	11	9	12	9	5	4
Теоријска статистика	7	12	13	7	6	5

- а) На нивоу значајности од 5%, испитати да ли је просечна оцена студената који су положили испит из Економетрије већа од 8,2. (10)
 б) Са поузданошћу од 95% формирати интервал поверења за учење студената који су положили испит из Теоријске статистике. (10)
 в) Одредити медијане оцена из Економетрије и Теоријске статистике. (5)

26. (Јул 2016) Расподела 100 испитаника у простом случајном узorku према заради и радном искуству приказана је у следећој табели:

Зарада (у 000 динара)	Радно искуство	
	До 10 година	преко 10 година
до 30	17	11
30-60	22	25
преко 60	10	15

- а) Са поузданошћу од 95% формирати интервал поверења за просечну зараду свих испитаника. (10)
 б) На нивоу значајности од 5%, испитати да ли се учење испитаника чије је радно искуство преко 10 година разликује од 60%. (10)

27. (Јул 2016) Тежина једног производа следи нормалну расподелу. Вероватноћа да тежина производа буде мања од 17kg износи 0,1587, а вероватноћа да тежина производа буде већа од 26kg износи 0,0228. Вероватноћа да просечна тежина производа у простом случајном узorku величине n буде мања од просечне тежине свих производа за мање од 1,2 износи 0,9918. Одредити величину простог случајног узorka (n). (20)

28. (Септембар 2013) Расподела плата запослених у једном предузећу је нормална. Просечно одступање плата од просека износи 200 динара. Плата случајно одабраног радника од 36000 динара већа је од просека за 5 стандардних девијација. Колико износи вероватноћа да у простом случајном узorku од 36 радника просечна плата буде већа од просечне плате у основном скупу за најмање 100 динара? (20)

29. (Септембар 2016) Плате радника у сектору А (у hiljadama динара) су нормално расподељене, а учење радника чија је плата мања од 53 hiljade динара износи 97,72%. Плате радника у сектору В (у hiljadama динара) такође су нормално расподељене, а учење радника чија је плата већа од 64,5 hiljada динара је 0,62%. Варијанса плата у сектору В је 1,44 пута већа од варијансе плата у сектору А. Релативне мере дисперзије плата у секторима А и В су једнаке. Упоредити треће квилте плате у секторима А и В. (20)

ИСПИТНИ ЗАДАЦИ – РЕШЕЊА ЗАДАТАКА

(дескриптивне мере, расподела вероватноћа, узорачка расподела,
интервално оцењивање, тестирање хипотеза на једном и на два узорка)

1. a) Na osnovu stope izbora $N=400$

$$P(X > 15) = 1 - P(X < 15) = 1 - P\left(\frac{X - \mu}{\sigma} < \frac{15 - 12}{\sqrt{2}}\right) = 1 - P(Z < 2,13) = 0,0166 \quad \text{Ukupan broj studenata: } 0,0166 \cdot 400 \approx 7$$

b) Uslovi su ispunjeni:

$$\begin{aligned} n\hat{p} &= 8 & \hat{p} - z_{\alpha/2} \hat{p} \leq p \leq \hat{p} + z_{\alpha/2} \hat{p} & \quad 0,32 - 1,96 \cdot 0,09 \leq p \leq 0,32 + 1,96 \cdot 0,09 \\ n\hat{q} &= 17 & 0,144 \leq p \leq 0,496 & \end{aligned} \quad s_{\hat{p}} = \sqrt{\frac{\hat{p}\hat{q}}{n} \cdot \frac{N-n}{N-1}} = 0,09$$

Sa pouzdanošću od 95% tvrdimo da se učešće svih studenata čija je stipendija veća od 15 hiljada dinara nalazi u intervalu od 14,4% do 49,6%.

2. a)

$$s = \sqrt{\frac{\sum x^2 f - \frac{(\sum x f)^2}{n}}{n-1}} = 1,04 \quad CV = \frac{s}{\bar{x}} \cdot 100 = \frac{1,04}{2,93} \cdot 100 = 35,49\%$$

b) uslovi:

standardna devijacija skupa je nepoznata; uzorak je veći od 30 – koristi se t raspodela

$$\begin{aligned} \bar{x} - t_{s\bar{x}} \leq \mu \leq \bar{x} + t_{s\bar{x}} & \quad 2,93 - 1,995 \cdot 0,124 \leq \mu \leq 2,93 + 1,995 \cdot 0,124 \\ & \quad 2,68 \leq \mu \leq 3,18 \end{aligned}$$

$$3. a) \bar{x}_1 = \frac{\sum x f_1}{n_1} = 32,13 \quad \bar{x}_2 = \frac{\sum x f_2}{n_2} = 31,33 \quad s_1 = 11,27 \quad s_2 = 10,61 \quad CV_1 = \frac{s_1}{\bar{x}_1} \cdot 100 = 35,08\% \quad CV_2 = \frac{s_2}{\bar{x}_2} \cdot 100 = 33,87\%$$

Veća je relativna disperzija starosti kupaca muškog pola.

b) I Formulisanje nulte i alternativne hipoteze

$$H_0: \mu_1 \leq \mu_2 \quad H_1: \mu_1 > \mu_2$$

II Izbor raspodele koja se koristi

uslovi: uzorci su slučajni i nezavisni, osnovni skupovi imaju normalnu raspodelu, standardne devijacije skupova su nepoznate ali jednake, uzorci su veći od 30 – koristi se t raspodela

III Određivanje oblasti odbacivanja i neodbacivanja

$$t_{n_1+n_2-2; \alpha} = t_{198; 0,05} = 1,645$$

IV Izračunavanje realizovane vrednosti statistike testa

$$\begin{aligned} s_p &= \sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}} = 10,88 & s_{\bar{x}_1 - \bar{x}_2} &= s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} = 1,57 & t &= \frac{\bar{x}_1 - \bar{x}_2}{s_{\bar{x}_1 - \bar{x}_2}} = 0,51 \end{aligned}$$

V Donošenje odluke Na nivou značajnosti od 5%, nulta hipoteza se ne odbacuje, pa se ne može zaključiti da je prosečna starost kupaca muškog pola veća od prosečne starosti kupaca ženskog pola

4. a) USLOVI su ispunjeni:

$$\begin{aligned} n\hat{p} &= 23,2 & \hat{p} - z_{\alpha/2} \hat{p} \leq p \leq \hat{p} + z_{\alpha/2} \hat{p} & \quad 0,29 - 1,64 \cdot 0,051 \leq p \leq 0,29 + 1,64 \cdot 0,051 \\ n\hat{q} &= 56,8 & 0,21 \leq p \leq 0,37 & \end{aligned}$$

Sa pouzdanošću od 90% tvrdimo da se učešće ispitanika čija su primanja od 20-60 hiljada dinara, a koji posećuju bioskop u toku meseca nalazi u intervalu od 21% do 37%.

b)

I Formulisanje nulte i alternativne hipoteze

$$H_0: p = 0,40 \quad H_1: p \neq 0,40$$

II Izbor raspodele koja se koristi

USLOVI su ispunjeni:

$$np = 80 \cdot 0,40 = 32$$

$$nq = 80 \cdot 0,60 = 48$$

III Određivanje oblasti odbacivanja i neodbacivanja

$$z = 1,96$$

IV Izračunavanje realizovane vrednosti statistike testa

$$z = \frac{\hat{p} - p}{\sigma \hat{p}} = \frac{0,36 - 0,4}{0,055} = -0,73$$

V Donošenje odluke Na nivou značajnosti od 5%, nulta hipoteza se ne odbacuje, pa se ne može zaključiti da se učešće ispitanika čija su mesečna primanja do 60 hiljada dinara, a koji nijednom nisu posetili bioskop razlikuje od 40%.

5.a)

$$\bar{x} = 24,34 \quad s = 9,43$$

Uslovi:

standardna devijacija skupa nepoznata, uzorak veći od 30 – koristi se t raspodela

$$\bar{x} - t_{s\bar{x}} \leq \mu \leq \bar{x} + t_{s\bar{x}} \quad 24,34 - 1,992 \cdot 1,08 \leq \mu \leq 24,34 + 1,992 \cdot 1,08$$

$$22,19 \leq \mu \leq 26,49$$

Sa pouzdanošću od 95% tvrdimo da se prosečna starost kupaca koji u prodavnicama zdrave hrane kupuju više od tri puta nedeljno nalazi u intervalu od 22,19 do 26,49 godina.

b)

I Formulisanje nulte i alternativne hipoteze

H_0 : između starosti kupaca i učestalosti kupovine ne postoji povezanost

H_1 : između starosti kupaca i učestalosti kupovine postoji povezanost

II Izbor raspodele koja se koristi

Pošto se ispituje nezavisnost obeležja koristi se hi-kvadrat test nezavisnosti

III Određivanje oblasti odbacivanja i neodbacivanja

$$\chi^2_{(R-1)(K-1); \alpha} = \chi^2_{6; 0,05} = 12,592$$

IV Izračunavanje realizovane vrednosti statistike testa

O	E	O-E	(O-E) ²	(O-E) ² /E
8	15.68	-7.68	58.98	3.76
11	16.96	-5.96	35.52	2.09
17	18.56	-1.56	2.43	0.13
28	12.8	15.20	231.04	18.05
10	14.7	-4.70	22.09	1.50
18	15.9	2.10	4.41	0.28
25	17.4	7.60	57.76	3.32
7	12	-5.00	25.00	2.08
31	18.62	12.38	153.26	8.23
24	20.14	3.86	14.90	0.74
16	22.04	-6.04	36.48	1.66
5	15.2	-10.20	104.04	6.84
200	200	0.00	745.92	48.69

$$\text{Npr. } E_1 = \frac{\text{sumareda} \cdot \text{sumakolone}}{n} = \frac{49 \cdot 64}{200} = 15,68 \quad \text{Uslov: } E \geq 5 \text{ je ispunjen} \quad \chi^2 = 48,69$$

V Donošenje odluke

Na nivou značajnosti od 5%, nulta hipoteza se odbacuje i zaključuje se da između starosti kupaca i učestalosti kupovine postoji povezanost.

6. a) USLOVI:

standardna devijacija skupa nepoznata, uzorak manji od 30, skup ima N raspodelu – koristi se t raspodela
koristi se pfks

$$\bar{x} - t_{s\bar{x}} \leq \mu \leq \bar{x} + t_{s\bar{x}} \quad 39,87 - 2,145 \cdot 4,25 \leq \mu \leq 39,87 + 2,145 \cdot 4,25$$

$$30,75 \leq \mu \leq 48,99$$

Uzorak je reprezentativan.

$$\text{b) } s = 17,55 \quad CV = \frac{s}{\bar{x}} \cdot 100 = 44,02\%$$

8.

a)

$$\bar{x} = 56,25 \quad s = 29,95 \quad CV = \frac{s}{\bar{x}} \cdot 100 = 53,24\%$$

b)

standardna devijacija skupa nepoznata, uzorak veći od 30 – koristi se t raspodela

koristi se pfs: $\frac{n}{N} = 0,08$

$$\begin{aligned} 56,25 - 1,96 \cdot 3,21 &\leq \mu \leq 56,25 + 1,96 \cdot 3,21 \\ \bar{x} - t_{s\bar{x}} &\leq \mu \leq \bar{x} + t_{s\bar{x}} \quad 49,96 \leq \mu \leq 62,54 \\ 49960 &\leq UP \leq 62540 \end{aligned}$$

9.

I Formulisanje nulte i alternativne hipoteze

$$H_0: p \leq 0,85 \quad H_1: p > 0,85$$

II Izbor raspodele koja se koristi

$$np = 150 \cdot 0,85 = 127,5$$

USLOVI su ispunjeni: $nq = 150 \cdot 0,15 = 22,5$

III Određivanje oblasti odbacivanja i neodbacivanja

$$z = 1,64$$

IV Izračunavanje realizovane vrednosti statistike testa

$$z = \frac{\hat{p} - p}{\sigma \hat{p}} = \frac{0,93 - 0,85}{0,029} = 2,76$$

V Donošenje odluke Na nivou značajnosti od 5%, nulta hipoteza se odbacuje i zaključuje se da je učešće svih osoba nižih od 190cm veće od 0,85.

b) USLOVI:

standardna devijacija skupa nepoznata, $n > 30$ – koristi se t raspodela

$$\begin{aligned} \bar{x} - t_{s\bar{x}} &\leq \mu \leq \bar{x} + t_{s\bar{x}} \quad 168,67 - 1,96 \cdot 1,25 \leq \mu \leq 168,67 + 1,96 \cdot 1,25 \\ 166,22 &\leq \mu \leq 171,12 \end{aligned}$$

Sa pouzdanošću od 95% tvrdimo da se prosečna visina svih osoba nalazi u intervalu od 166,22 do 171,12 cm.

10. I Formulisanje nulte i alternativne hipoteze

$$H_0: \mu_1 \geq \mu_2 \quad H_1: \mu_1 < \mu_2$$

II Izbor raspodele koja se koristi

uslovi:

uzorci su slučajni i nezavisni, osnovni skupovi imaju normalnu raspodelu, standardne devijacije skupova su nepoznate ali jednake, uzorci su manji od 30 – koristi se t raspodela

III Određivanje oblasti odbacivanja i neodbacivanja $t_{n_1+n_2-2; \alpha} = t_{49; 0,05} = 1,677$

IV Izračunavanje realizovane vrednosti statistike testa

$$\bar{x}_1 = 7,59; \quad \bar{x}_2 = 7,21 \quad s_1^2 = 1,4 \quad s_2^2 = 1,39 \quad s_p = 1,18 \quad s_{\bar{x}_1 - \bar{x}_2} = 0,33 \quad t = \frac{\bar{x}_1 - \bar{x}_2}{s_{\bar{x}_1 - \bar{x}_2}} = 1,15$$

V Donošenje odluke

Na nivou značajnosti od 5%, nulta hipoteza se ne odbacuje i zaključuje se da prosečna ocena studenata koji su položili ispit iz Ekonometrije nije manja od prosečne ocene studenata koji su položili ispit iz Teorijske statistike.

11.

a)

I

$$H_0: p \leq 0,75 \quad H_1: p > 0,75$$

II USLOVI su ispunjeni:

$$np = 50 \cdot 0,75 = 37,5$$

$$nq = 50 \cdot 0,25 = 12,5$$

III $z = 1,64$

$$IV \quad z = \frac{\hat{p} - p}{\sigma_{\hat{p}}} = \frac{0,82 - 0,75}{0,061} = 1,148$$

V **Donošenje odluke** Na nivou značajnosti od 5%, nulta hipoteza se ne odbacuje, pa se ne može zaključiti da je učešće studenata koji su položili ispit iz Matematike veće od 75%.

b)

$$\frac{N+1}{4} = 12,75$$

$$\frac{3(N+1)}{4} = 38,25$$

Matematika

$$Q_1 = 6$$

$$Q_3 = 8$$

$$IQR = 2$$

Osnovi statističke analize

$$Q_1 = 5$$

$$Q_3 = 8$$

$$IQR = 3$$

12.

a) USLOVI su ispunjeni:

$$n\hat{p} = 100$$

$$n\hat{q} = 50$$

$$\hat{p} - z_{\alpha} \hat{p} \leq p \leq \hat{p} + z_{\alpha} \hat{p}$$

$$0,67 - 1,96 \cdot 0,038 \leq p \leq 0,67 + 1,96 \cdot 0,038$$

$$0,596 \leq p \leq 0,744$$

Sa pouzdanosću od 95% tvrdimo da se učešće radnika koji u pomenutom preduzeću rade više od 10 godina nalazi u intervalu od 59,6% do 74,4%.

b)

I Formulisanje nulte i alternativne hipoteze

H_0 : između dužine radnog staža i broja ostvarenih bonusa ne postoji povezanost

H_1 : između dužine radnog staža i broja ostvarenih bonusa postoji povezanost

II Izbor raspodele koja se koristi

Pošto se ispituje nezavisnost obeležja koristi se hi-kvadrat test nezavisnosti

III Određivanje oblasti odbacivanja i neodbacivanja

$$\chi^2_{(R-1)(K-1); \alpha} = \chi^2_{4; 0,05} = 9,488$$

IV Izračunavanje realizovane vrednosti statistike testa

O	E	O-E	(O-E) ²	(O-E) ² /E
8.00	15.00	-7.00	49.00	3.27
12.00	15.00	-3.00	9.00	0.60
30.00	20.00	10.00	100.00	5.00
15.00	12.30	2.70	7.29	0.59
7.00	12.30	-5.30	28.09	2.28
19.00	16.40	2.60	6.76	0.41
22.00	17.70	4.30	18.49	1.04
26.00	17.70	8.30	68.89	3.89
11.00	23.60	-12.60	158.76	6.73
150.00	150.00	0.00	446.28	23.82

$$\text{Npr. } E_1 = \frac{\text{sumareda} \cdot \text{sumakolone}}{n} = \frac{50 \cdot 45}{150} = 15$$

USLOV:

$$E \geq 5 \text{ je ispunjen}$$

$$\text{Realizovana vrednost: } \chi^2 = 23,82$$

V Donošenje odluke Na nivou značajnosti od 5%, nulta hipoteza se odbacuje i zaključuje se da između dužine radnog staža i broja ostvarenih bonusa postoji povezanost.

13.

$$P(\mu - X > 4) = P(X - \mu < -4) = P\left(\frac{X - \mu}{\sigma} < \frac{-4}{\sigma}\right) = 0,1587$$

$$\frac{-4}{\sigma} = -1$$

$$\sigma = 4$$

$$P(|\bar{X} - \mu| > 1) = 1 - P(|\bar{X} - \mu| < 1) = 0,3174$$

$$P(|\bar{X} - \mu| < 1) = P(-1 < \bar{X} - \mu < 1) = P\left(\frac{-1}{\frac{\sigma}{\sqrt{n}}} < \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} < \frac{1}{\frac{\sigma}{\sqrt{n}}}\right) = 0,6826$$

$$2F\left(\frac{1}{\frac{\sigma}{\sqrt{n}}}\right) - 1 = 0,6826$$

$$2F\left(\frac{\sqrt{n}}{4}\right) - 1 = 0,6826$$

$$\frac{\sqrt{n}}{4} = 1$$

$$n = 16$$

$$p = P(X < 38) = P\left(\frac{X - \mu}{\sigma} < \frac{38 - 40}{4}\right) = P(Z < -0,5) = 0,3085$$

Uslovi nisu ispunjeni za aproksimaciju proporcije uzorka normalnom raspodelom:

$$np = 4,936$$

$$nq = 11,064$$

14.

$$P(X > 6) = 1 - P(X < 6) = 0,4013$$

$$P\left(\frac{X - \mu}{\sigma} < \frac{6 - \mu_X}{\sigma_X}\right) = 0,5987$$

$$\sigma_X^2 = 4\sigma_Y^2$$

$$\sigma_X = 2\sigma_Y$$

$$CV_X = 2,76CV_Y$$

$$\frac{\sigma_X}{\mu_X} = \frac{2,76\sigma_Y}{\mu_Y}$$

$$\frac{2\sigma_Y}{\mu_X} = \frac{2,76\sigma_Y}{\mu_Y}$$

$$\mu_Y = 1,38\mu_X$$

$$\frac{6 - \mu_X}{\sigma_X} = 0,25$$

$$6 - \mu_X = 0,5\sigma_Y$$

$$\mu_X = 6 - 0,5\sigma_Y$$

$$P(Y < 4) = 0,1469$$

$$\frac{4 - \mu_Y}{\sigma_Y} = -1,05$$

$$4 - \mu_Y = -1,05\sigma_Y$$

$$4 - 1,38(6 - 0,5\sigma_Y) = -1,05\sigma_Y$$

$$\sigma_X = 4,92$$

$$\mu_X = 4,77$$

$$\sigma_Y = 2,46$$

$$\mu_Y = 6,58$$

15. a)

I Formulisanje nulte i alternativne hipoteze

$$H_0: p \geq 0,4 \quad H_1: p < 0,4$$

II Izbor raspodele koja se koristi

USLOVI su ispunjeni: $np = 50 \cdot 0,4 = 20$
 $nq = 50 \cdot 0,6 = 30$

III Određivanje oblasti odbacivanja i neodbacivanja

$$z = -1,64$$

IV Izračunavanje realizovane vrednosti statistike testa

$$z = \frac{\hat{p} - p}{\sigma_{\hat{p}}} = \frac{0,36 - 0,4}{0,069} = -0,58$$

V **Donošenje odluke** Na nivou značajnosti od 5%, nulta hipoteza se ne odbacuje.

b) USLOVI:

standardna devijacija skupa nepoznata, $n > 30$ – koristi se t raspodela

$$\bar{x} - t_{S\bar{x}} \leq \mu \leq \bar{x} + t_{S\bar{x}} \quad 35,6 - 2,01 \cdot 2,86 \leq \mu \leq 35,6 + 2,01 \cdot 2,86$$

$$29,85 \leq \mu \leq 41,35$$

16.I Formulisanje nulte i alternativne hipoteze

$$H_0: \mu_1 = \mu_2 \quad H_1: \mu_1 \neq \mu_2$$

II Izbor raspodele koja se koristi: uzorci su slučajni i nezavisni, osnovni skupovi imaju normalnu raspodelu, standardne devijacije skupova su nepoznate ali jednake, uzorci su manji od 30 – koristi se t raspodela

III Određivanje oblasti odbacivanja i neodbacivanja $t_{n_1+n_2-2; \frac{\alpha}{2}} = t_{29; 0,025} = 2,045$

IV Izračunavanje realizovane vrednosti statistike testa

$$\bar{x}_1 = \frac{\sum x' f_1}{n_1} = \frac{570}{15} = 38; \quad \bar{x}_2 = \frac{\sum x' f_2}{n_2} = \frac{440}{16} = 27,5 \quad s_1^2 = 217,14 \quad s_2^2 = 260$$

$$s_p = \sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}} = 15,47 \quad s_{\bar{x}_1 - \bar{x}_2} = s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} = 5,56 \quad t = \frac{\bar{x}_1 - \bar{x}_2}{s_{\bar{x}_1 - \bar{x}_2}} = 1,89$$

V **Donošenje odluke:** Na nivou značajnosti od 5%, nulta hipoteza se ne odbacuje.

17.

$$P(X - \mu > 6) = 1 - P(X - \mu < 6) = 0,0228$$

$$P(X - \mu < 6) = P\left(\frac{X - \mu}{\sigma} < \frac{6}{\sigma}\right) = 0,9772$$

$$\frac{6}{\sigma} = 2$$

$$\sigma = 3$$

$$P(|\bar{X} - \mu| > 1,5) = 1 - P(|\bar{X} - \mu| < 1,5) = 0,0124$$

$$P(|\bar{X} - \mu| < 1,5) = P(-1,5 < \bar{X} - \mu < 1,5) = P\left(\frac{-1,5}{\frac{\sigma}{\sqrt{n}}} < \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} < \frac{1,5}{\frac{\sigma}{\sqrt{n}}}\right) = 0,9876$$

$$2F\left(\frac{1,5}{\frac{\sigma}{\sqrt{n}}}\right) - 1 = 0,9876$$

$$F\left(\frac{1,5}{\frac{\sigma}{\sqrt{n}}}\right) = 0,9938$$

$$\frac{1,5}{\frac{\sigma}{\sqrt{n}}} = 2,5$$

$$n = 25$$

$$p = P(X < 43) = P\left(\frac{X - \mu}{\sigma} < \frac{43 - 40}{3}\right) = P(Z < 1) = 0,8413$$

Nisu ispunjeni uslovi za aproksimaciju raspodele proporcije uzorka normalnom raspodelom:

$$np = 21,03$$

$$nq = 3,97$$

18. a)

$$P(X > 1000) = 1 - P(X < 1000) = 1 - P(Z < 1) = 0,1587$$

$$0,1587 \cdot 300 \approx 48$$

b) Uslovi:

standardna devijacija skupa poznata, N raspodela skupa – koristi se z raspodela

$$\bar{x} - z\sigma_{\bar{x}} \leq \mu \leq \bar{x} + z\sigma_{\bar{x}} \quad 12,33 - 1,96 \cdot 5,77 \leq \mu \leq 12,33 + 1,96 \cdot 5,77$$

$$1,02 \leq \mu \leq 23,64$$

$$12,22 \leq \mu \leq 12,44$$

Uzorak je reprezentativan.

$$z = \frac{x - \mu}{s} = \frac{18 - 12,33}{5,88} = 0,96$$

c)

19.

$$P(X < 6,4) = 0,75$$

$$P\left(\frac{X - \mu}{\sigma} < \frac{6,4 - \mu}{\sigma}\right) = 0,75$$

$$\frac{6,4 - \mu}{\sigma} = 0,67$$

$$P(\mu - X < 1,2) = 0,9772$$

$$P(X - \mu < -1,2) = 0,0228$$

$$\frac{-1,2}{\sigma} = -2$$

$$\sigma = 0,6$$

$$\mu = 6$$

$$P(\bar{X} - \mu > 0,1) = 0,1587$$

$$P(\bar{X} - \mu < 0,1) = 0,8413$$

$$\frac{0,1}{\frac{0,6}{\sqrt{n}}} = 1$$

$$\sqrt{n} = 0,6$$

Uslov: $\bar{X}$ ima normalnu raspodelu

$$p = P(X < 5,7) = 0,3085$$

$$q = 0,6915$$

$$P(\hat{p} - p < 0,2) = P(Z < 2,6) = 0,9953$$

Uslovi su ispunjeni za aproksimaciju proporcije uzorka normalnom raspodelom:

$$np = 11,11$$

$$nq = 24,89$$

20. a)

I Formulisanje nulte i alternativne hipoteze

$$H_0: p \geq 0,6 \quad H_1: p < 0,6$$

II Izbor raspodele koja se koristi

$$\text{Ispunjeni uslovi:} \quad np = 36 \cdot 0,6 = 21,6$$

$$nq = 36 \cdot 0,4 = 14,4$$

III Određivanje oblasti odbacivanja i neodbacivanja

$$F(z) = \alpha = 0,05$$

$$z = -1,64$$

IV Izračunavanje realizovane vrednosti statistike testa

$$z = \frac{\hat{p} - p}{\sigma_{\hat{p}}} = \frac{0,56 - 0,6}{0,082} = -0,49$$

V Donošenje odluke:

Na nivou značajnosti od 5%, nulta hipoteza se ne odbacuje.

b)

$$P(X > 8) = 1 - P(X < 8) = 0,1056 \quad P(\bar{X} < 7) = 0,9938$$

$$P(X < 8) = 0,8944$$

$$\frac{7 - \mu}{\frac{\sigma}{6}} = 2,5$$

$$\frac{8 - \mu}{\sigma} = 1,25$$

$$\sigma = 1,2$$

$$\mu = 8 - 1,25\sigma$$

$$\mu = 6,5$$

22.

a)

USLOVI su ispunjeni:

$$n\hat{p} = 6$$

$$n\hat{q} = 9$$

Koristi se pfs

$$\hat{p} - z_{\alpha} \hat{p} \leq p \leq \hat{p} + z_{\alpha} \hat{p}$$

$$0,4 - 1,96 \cdot 0,12 \leq p \leq 0,4 + 1,96 \cdot 0,12$$

$$0,16 \leq p \leq 0,64$$

$$24 \leq UB \leq 96$$

b)

I

$$H_0: p \leq 0,35 \quad H_1: p > 0,35$$

II USLOVI su ispunjeni:

$$np = 5,25$$

$$nq = 9,75$$

III

$$z = 1,64$$

IV

Koristi se pfs

$$z = \frac{\hat{p} - p}{\sigma_{\hat{p}}} = \frac{0,4 - 0,35}{0,117} = 0,43$$

V Donošenje odluke

Na nivou značajnosti od 5%, nulta hipoteza se ne odbacuje, pa se ne može zaključiti da je učešće studenata koji su u toku nedelje proveli više od 8 sati u učenju stranog jezika veće od 35%.

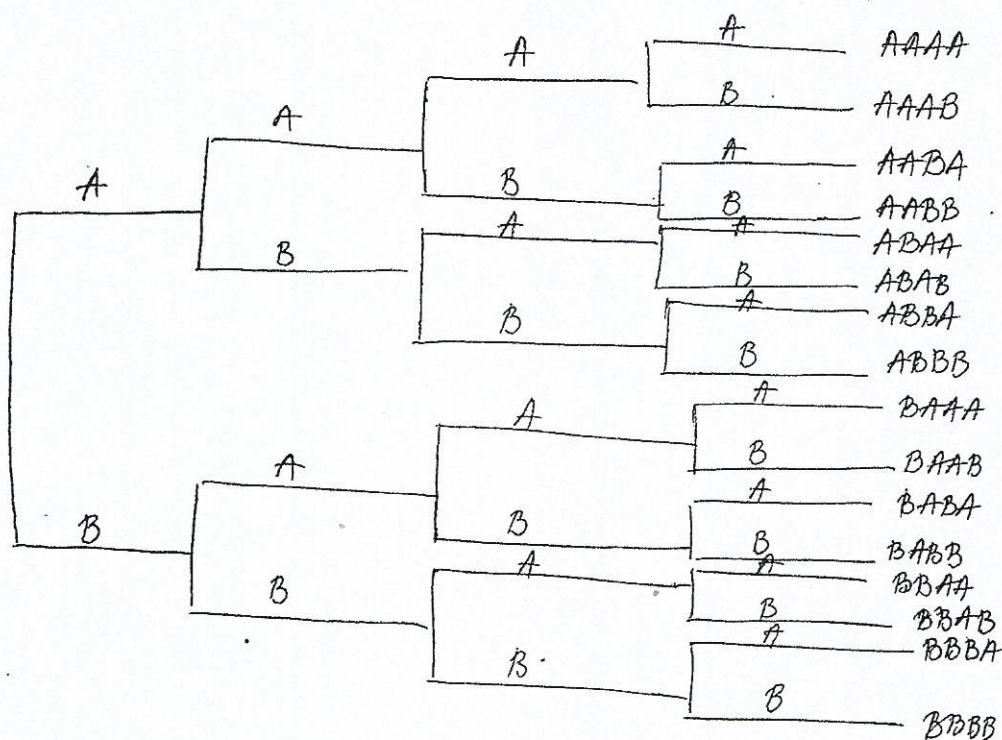
21.

$$p=0,28$$

$$q=0,72$$

A - IZABRANI ISPITANIK POSLUJE NA BERZI

B - IZABRANI ISPITANIK NE POSLUJE NA BERZI



x_i	0	1	2	3	4
p_i	0,2687	0,4180	0,2439	0,0632	0,0061

d) $P(X=2) = 0,2439$

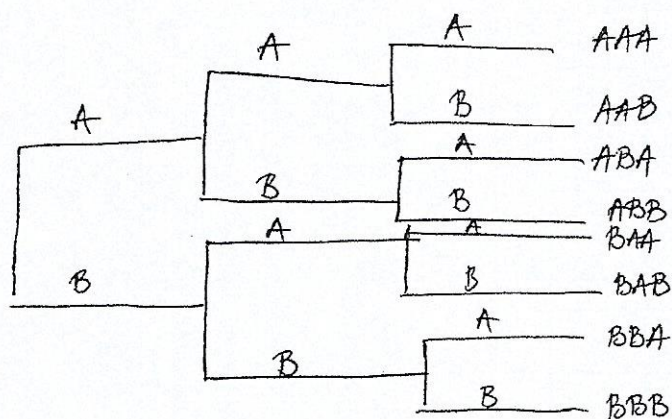
7. a)

$$p = 0,46$$

$$q = 0,54$$

A - IZABRANI ISPITANIK BI ČEJDE POSEĐIVAO POŽORIŠTE DA SU
ULAŽNICE JEFTINIJE

B - IZABRANI ISPITANIK NE BI ČEJDE POSEĐIVAO POŽORIŠTE DA SU
ULAŽNICE JEFTINIJE



x_i	0	1	2	3
p_i	0,1575	0,4024	0,3428	0,0973

$$P(X=0) = P(BBB) = 0,54^3 = 0,1575$$

$$P(X=1) = P(ABB) + P(BAB) + P(BBA) = 0,46 \cdot 0,54^2 + 0,54 \cdot 0,46 \cdot 0,54 + 0,54^2 \cdot 0,46 = 0,4024$$

$$P(X=2) = P(AAB) + P(ABA) + P(BAA) = 0,3428$$

$$P(X=3) = P(AAA) = 0,0973$$

δ) $P(X=1) = 0,4024$

KOLOKVIJUM IZ PREDMETA: OSNOVI STATISTIČKE ANALIZE

GRUPA I

Ime i prezime _____ Broj indeksa _____

1. Pored sledećih iskaza napisati da li su tačni ili netačni:
 - a. Uzoračka raspodela statistike predstavlja skup parova vrednosti koje statistika može uzeti i odgovarajućih frekvencija.
 - b. Greška I vrste se javlja kada se istinita nulta hipoteza ne odbaci.
 - c. Kocka za igru se baca dva puta. Ako je A događaj da je zbir palih brojeva 4, 5 ili 6, a B događaj da je zbir palih brojeva neparan broj, onda su događaji A i B međusobno isključivi.
 - d. Raspodela kumulativnih frekvencija grafički se može predstaviti štapčastim dijagramom.
 - e. Modus se može izračunati i za kvalitativne i za kvantitativne podatke.

(10)

2. Iz osnovnog skupa od 360 kuglica izabran je prost slučajni uzorak od 36 kuglica i zabeležen je njihov prečnik (u cm):

Prečnik	1-3	3-5	5-7
Broj kuglica	10	20	6

- a) Izračunati i komentarisati apsolutnu i relativnu meru disperzije navedene raspodele.
- b) Na nivou značajnosti od 10% ispitati da li je prosečan prečnik svih kuglica manji od 4 cm.

(10)

(10)

3. a) Raspodela verovatnoća slučajne promenljive X prikazana je tabelom:

X	a	0	1	2
$P(X=x_i)$	0.4	0.1	0.2	0.3

Ako je $E(X)=0$, naći a .

(10)

- b) U sledećoj tabeli je prikazana raspodela slučajno izabranih studenata I godine jednog Ekonomskog fakulteta prema oceni koju su dobili u junskom roku na ispitu iz Osnova statističke analize:

Ocena	5	6	7	8	9	10
Osnovi statističke analize	9	7	8	6	4	2

Sa pouzdanosti od 95% oceniti učešće svih studenata koji su u junskom roku položili ispit iz Osnova statističke analize.

(10)

1. N, N, N, N, T

2. a) $N=360$
 $n=36$

$$s = \sqrt{\frac{\sum x^2 f - \frac{(\sum x f)^2}{n}}{n-1}} = \sqrt{\frac{576 - \frac{136^2}{36}}{35}} = 1,33$$

$$\bar{x} = \frac{\sum x f}{n} = \frac{136}{36} = 3,78$$

$$CV = \frac{s}{\bar{x}} \cdot 100 = 35,19\%$$

x	f	x'	x' ²	x' ² f
1-3	10	2	20	40
3-5	20	4	80	320
5-7	6	6	36	216
Σ	36		136	576

Δ - ^{pojedinačnih} Prečnik kuglica u proseku odstupa od prosečnog prečnika za 1,33 cm.
CV - Prosečno odstupanje iznosi 35,19% aritmetičke sredine

b) I Formulisanje H_0 i H_1

$$H_0: \mu \geq 4$$

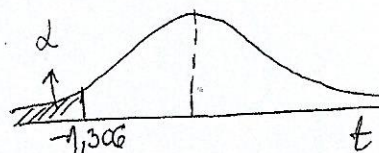
$$H_1: \mu < 4$$

II Izbor raspodele

o nepoznatu σ raspodela
 $n \geq 30$

III Određivanje oblasti odbacivanja i neodbacivanja

$$t = t_{n-1; \alpha} = t_{35; 0,1} = -1,306$$



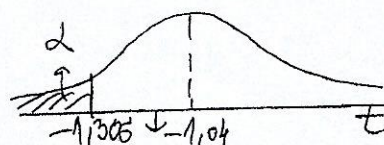
IV Izračunavanje vrednosti statistike testa

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{3,78 - 4}{0,211} = -1,09$$

$$\frac{n}{N} = 0,1 > 0,05 \quad s/\sqrt{n} = \frac{s}{\sqrt{N}} \sqrt{\frac{N-n}{N-1}} = \frac{1,33}{\sqrt{36}} \sqrt{\frac{360-36}{360-1}} = 0,211$$

V Donošenje odluke

Na nivou značajnosti od 10% H_0 se ne odbacuje i ne može se zaključiti da je prosečan prečnik svih kuglica manji od 4 cm.



$$3. a) E(X) = \sum x_i p_i = a \cdot 0,4 + 0 \cdot 0,1 + 1 \cdot 0,2 + 2 \cdot 0,3 = 0$$

$$a = -2$$

$$b) n = 36$$

$$f = 27$$

$$\hat{p} = \frac{f}{n} = 0,75$$

uslovi:

$$np = 27 > 5$$

$$n\hat{q} = 9 > 5$$

$$\hat{p} - z\Delta p \leq p \leq \hat{p} + z\Delta p$$

$$F(z) = 1 - \frac{\alpha}{2} = 0,975 \quad z = 1,96$$

$$\Delta p = \sqrt{\frac{\hat{p}\hat{q}}{n}} = 0,072$$

$$0,75 - 1,96 \cdot 0,072 \leq p \leq 0,75 + 1,96 \cdot 0,072$$

$$0,609 \leq p \leq 0,891$$

Sa pouzdanoscu od 95% tvrdimo da se mace svih studenata koji su položili ispit iz OSA nalazi u intervalu od 60,9% do 89,1%.

KOLOKVIJUM IZ PREDMETA: OSNOVI STATISTIČKE ANALIZE

GRUPA II

Ime i prezime _____ Broj indeksa _____

1. Koji od sledećih iskaza su tačni (T), a koji netačni (N)? (Zaokružiti odgovarajuće slovo posle svakog iskaza.)
- a. Nulta hipoteza je tvđenje o nekom parametru osnovnog skupa koje se smatra istinitim sve dok se ne dokaže suprotno. T N
 - b. t raspodela je simetrična raspodela koja je više izdužena i više raspršena od standardizovane normalne raspodele. T N
 - c. Kocka za igru se baca dva puta. Ako je A događaj da je zbir palih brojeva 3, 6 ili 8, a B događaj da je zbir paran broj, onda su događaji A i B međusobno isključivi. T N
 - d. Modus je vrednost koja se javlja sa najvećom frekvencijom u seriji podataka. T N
 - e. Grupisani kvantitativni podaci (prikazani celobrojnim vrednostima) mogu se grafički prikazati pomoću poligona frekvencija. T N

(10)

2. Raspodela 100 ispitanika u prostom slučajnom uzorku prema zaradi i radnom iskustvu prikazana je u sledećoj tabeli:

Zarada (u 000 dinara)	Radno iskustvo	
	Do 10 godina	preko 10 godina
do 40	17	8
40-60	20	21
60-80	10	12
preko 80	5	7

- a) Izračunati i protumačiti relativnu meru disperzije ispitanika čije je radno iskustvo do 10 godina. (10)
- b) Sa pouzdanošću od 95% oceniti prosečnu zaradu svih ispitanika čije je radno iskustvo preko 10 godina. (10)

3. Težina proizvoda sledi normalnu raspodelu sa sredinom 8kg i varijansom 4kg^2 . U prostom slučajnom uzorku od 15 proizvoda zabeležene su sledeće težine proizvoda: 5, 13, 8, 14, 5, 7, 7, 10, 12, 8, 9, 14, 3, 15, 7.
- a) Izračunati verovatnoću da proporcija proizvoda u uzorku čija je težina veća od 5 kg bude manja od 80%. (10)
 - b) Na nivou značajnosti od 10% ispitati da li je učešće proizvoda čija je težina manja od 9kg veće od 60%. (10)

ПРИМЕРИ ИСПИТНИХ ЗАДАТАКА

(хи-квадрат тест, анализа варијансе, корелација, регресија, индексни бројеви, анализа временских серија)

1. (Jun 2015) U jednom istraživanju je učestvovalo 50 ispitanika, koji su dali odgovore na pitanja o starosti i učestalosti plivanja. Rezultati su prikazani u sledećoj tabeli:

		Učestalost plivanja u toku nedelje		
		Nikad	Jednom nedeljno	Više od tri puta nedeljno
Godine starosti ispitanika	Ispod 20	8%	12%	14%
	20-40	10%	10%	14%
	Preko 40	14%	16%	2%

Na nivou značajnosti od 5%, ispitati da li postoji povezanost između starosti ispitanika i učestalosti plivanja. (20)

2. (Jun 2015) U tabeli su dati podaci o uvozu sirovina (u milionima dinara) i ostvarenoj proizvodnji (u hiljadama tona) u prostom slučajnom uzorku od devet preduzeća:

Uvoz	25	37	44	52	68	74	82	93	101
Proizvodnja	54	63	78	84	97	102	114	122	135

- a) Izračunati sumu kvadrata odstupanja proizvodnje pojedinih preduzeća od ocenjene vrednosti proizvodnje. (10)
 b) Na nivou značajnosti od 5%, ispitati da li uvoz sirovina značajno utiče na proizvodnju. (5)
 c) Sa pouzdanošću od 95% oceniti za koliko će se u proseku promeniti proizvodnja, ako se uvoz poveća za 3 miliona dinara. (10)

3. (Jul 2015) Vreme (u časovima) koje 15 slučajno odabranih ispitanika provede nedeljno u teretani s obzirom na godine starosti dato je u sledećoj tabeli:

Godine starosti		
do 20	20-40	preko 40
14	12	7
8	9	8
12	8	4
6	2	11
10	5	6

Na nivou značajnosti od 5%, ispitati da li godine starosti ispitanika sistematski utiču na varijabilitet vremena koje ispitanici provedu u teretani. (15)

4. (Jul 2015) Ispituje se uticaj kretanja plate na kretanje izdataka za garderobu. Raspolaze se podacima o plati (u 000 dinara): 26, 34, 23, 38, 45 i 52 i izdacima za garderobu (u 000 dinara): 2, 6, 3, 8, 10 i 14 u prostom slučajnom uzorku od šest ispitanika.

- a) Izračunati i interpretirati relativnu meru reprezentativnosti linije regresije (sve pretpostavke prostog linearnog regresionog modela su ispunjene). (10)
 b) Sa pouzdanošću od 95% oceniti za koliko će se u proseku promeniti izdaci za garderobu, ako se plata poveća za 10 hiljada dinara. (10)

5. (Jul 2015) Vrednost prodaje proizvoda A, B, C i D u jednoj grani prehrambene industrije u 2010. godini bila je sledeća: 2800; 2520; 2688 i 2480. Indeksi vrednosti prodaje ovih proizvoda u 2014. godini (2010=100) su iznosili: 146; 178,1; 170,57 i 203,23. Relativan rast fizičkog obima prodaje ovih proizvoda u 2014. godini u odnosu na 2010. godinu iznosio je redom: 30,36%; 48,41%; 55,95% i 93,55%. Cena proizvoda A je u 2014. godini povećana za 12% u odnosu na 2010. godinu, proizvoda B je povećana za 20%, proizvoda C je povećana za 9,38% i proizvoda D je povećana za 5%. Metodom agregata odrediti grupni indeks fizičkog obima prodaje i cene u 2014. godini (2010=100), koristeći pondere iz tekućeg perioda. (20)

6. (Jul 2015) 500 studenata jednog fakulteta prima stipendiju. 250 studenata prima stipendiju do 5000 dinara, 125 studenata prima stipendiju od 5000 do 10000 dinara, 75 studenata prima stipendiju od 10000 do 15000 dinara i 50 studenata prima stipendiju preko 15000 dinara. U slučajnom uzorku od 50 studenata dobijeni su sledeći rezultati:

Visina stipendije	Broj studenata
do 5000	20
5000 - 10000	15
10000 - 15000	10
preko 15000	5

Na nivou značajnosti od 5%, ispitati da li se raspodela studenata prema visini stipendije značajno razlikuje od očekivane raspodele. (15)

7. (Jul 2015) U tabeli su dati podaci o ceni proizvoda (u stotinama dinara) i tražnji za tim proizvodom (u kg) u prostom slučajnom uzorku od šest proizvoda:

Cena	4	7	13	19	25	32
Tražnja	90	75	51	38	22	13

- a) Izračunati sumu kvadrata odstupanja tražnje za pojedinim proizvodima od prosečne tražnje i sumu kvadrata odstupanja tražnje za pojedinim proizvodima od ocenjene tražnje. Koliko je učešće neobjašnjenog varijabiliteta? (10)
b) Sa pouzdanošću od 95% oceniti za koliko će se u proseku promeniti tražnja, ako se cena poveća za 2 stotine dinara. (10)

8. (Jul 2015) Na osnovu originalne serije obima prodaje jednog proizvoda u periodu od 2007. do 2014. godine, utvrđeno je da je prosečna prodaja iznosila 54,13 kg, a da se prodaja u proseku godišnje povećavala po stopi od 11,71%. Poznato je da je u 2007. godini prodaja za 4,76% bila veća od ocenjene funkcije linearnog trenda, a da je u 2014. godini prodaja iznosila 76 kg.

- a) Izračunati apsolutnu i relativnu promenu prodaje u 2014. u odnosu na 2007. godinu. (15)
b) Predvideti vrednost prodaje u 2016. godini. Da li je dobijena vrednost validna? (5)

9. (Septembar 2015) Dati su indeksi fizičkog obima prodaje jedne vrste voća, po kvartalima 2012., 2013. i 2014. godine (prosek 2013=100): 42,55; 102,13; 164,54; 59,57; 48,23; 99,29; 178,72; 73,76; 36,88; 110,64; 190,07; 51,06. Prodaja pomenutog voća u II kvartalu 2013. godine iznosila je 35 hiljada tona. Sezonski indeksi prodaje, po kvartalima u posmatrane tri godine, iznosili su redom: 43,5; 106; 179 i 69,5. Predvideti prosečan obim prodaje pomenutog voća u kvartalu sa "živom" sezonom u 2015. godini, uzimajući u obzir uticaj sezonskih faktora. (25)

10. (Septembar 2015) Iz tri košarkaška kluba u Srbiji slučajnim putem je izabrano po pet igrača i zabeleženo vreme koje nedeljno provedu trenirajući (u časovima):

Prvi klub	Drugi klub	Treći klub
5	9	10
8	7	6
7	11	13
10	15	16
9	12	9

Na nivou značajnosti od 5%, ispitati da li se prosečno vreme treniranja igrača razlikuje između bar dva kluba (sve potrebne pretpostavke su ispunjene). (15)

11. (Septembar 2015) Menadžere Telenora interesuje da li se prosečne vrednosti ostvarenog telefonskog saobraćaja njihovih korisnika sa Voždovca, Banovog brda, Zemuna i Čukarice međusobno razlikuju. Na slučajan način odabrano je po 4 korisnika iz svake opštine i zabeležena vrednost telefonskog saobraćaja (u stotinama dinara) svakog od njih:

Vrednost saobraćaja			
Voždovac	Banovo brdo	Zemun	Čukarica
7	15	19	4
18	23	26	22
14	31	29	17
9	12	13	8

Na nivou značajnosti od 5%, ispitati da li se prosečna vrednost telefonskog saobraćaja u Telenor mreži značajno razlikuje između bar dve opštine. (15)

12. (Septembar 2015) U prostom slučajnom uzorku od 7 prodavnica na teritoriji Beograda zabeleženi su podaci o izdacima za reklamiranje (u 000 dinara) i broju prodatih proizvoda (u 00 komada) i dobijeni su sledeći rezultati: broj prodatih proizvoda u proseku odstupa od prosečnog broja prodatih proizvoda za 4,57 stotina komada; ocenjena standardna greška ocene regresionog parametra uz objašnjavajuću promenljivu je 0,166 hiljada dinara; prosečan broj prodatih proizvoda iznosi 8,29 stotina komada; izdaci za reklamiranje iznose: 2, 4, 4, 5, 7, 8, 10.

- a) Oceniti prost linearni regresioni model i objasniti dobijeni regresioni koeficijent nagiba. (10)
b) Sa pouzdanošću od 95% predvideti broj prodatih proizvoda, ako izdaci za reklamiranje iznose 12 hiljada dinara. Da li je dobijena predviđena vrednost validna? (15)

13. (Jun 2016) Sprovedeno je istraživanje sa ciljem da se utvrdi da li starost gledalaca (u godinama) i slobodno vreme (u časovima) utiču na gledanost televizijskog programa (u časovima). Na osnovu prostog slučajnog uzorka od 8 gledalaca dobijeni su sledeći rezultati:

	Koeficijent	Standardna greška
Konstanta	14,2	2,8
Starost gledalaca	2,56	2,12
Slobodno vreme	3,84	1,16

94,5% varijacija gledanosti televizijskog programa objašnjeno je zajedničkim uticajem starosti gledalaca i slobodnog vremena.

- a) Na nivou značajnosti od 5% ispitati da li starost gledalaca i slobodno vreme utiču na gledanost televizijskog programa. (10)
- b) Koliko iznosi procenat varijacija gledanosti televizijskog programa koji nije objašnjen zajedničkim uticajem starosti gledalaca i slobodnog vremena? (5)
- c) Sa pouzdanošću od 95% oceniti za koliko će se u proseku promeniti gledanost televizijskog programa, ako se starost gledalaca poveća za 3 meseca, a slobodno vreme ostane nepromenjeno. (10)

14. (Jun 2016) Dati su lančani indeksi proizvodnje jednog proizvoda u svim mesecima 2015. godine: 103,70; 104,46; 90,60; 92,45; 96,94; 95,79; 108,79; 106,06; 104,76; 103,64; 107,89; 104,88. Prosečna vrednost proizvodnje (u hiljadama tona) tog proizvoda u decembru 2014. godine iznosila je 108 hiljada tona. Ocenjena funkcija eksponencijalnog trenda je $\hat{y}_t = 20,35 \cdot 1,16^t$, a srednja kvadratna greška eksponencijalnog trenda 115,12 hiljada tona.

- a) Odabрати funkciju trenda koja se najbolje prilagođava podacima (kao kriterijum koristiti srednju kvadratnu grešku). (15)
- b) Isključiti trend komponentu iz date vremenske serije u junu 2015. godine. Objasniti rezultat. (5)

15. (Jun 2016) Poznato je da od 1500 radnika jednog preduzeća 600 radnika ima platu do 20.000 dinara, 300 radnika ima platu od 20.000 do 40.000 dinara, 150 radnika ima platu od 40.000 do 60.000 dinara, dok preostali radnici imaju platu preko 60.000 dinara. Slučajan uzorak od 200 radnika pokazao je da 45% radnika ima platu do 20.000 dinara, 15% radnika ima platu od 20.000 do 40.000 dinara, 10% radnika ima platu od 40.000 do 60.000 dinara, dok preostali radnici imaju platu preko 60.000 dinara. Na nivou značajnosti od 10% ispitati da li se raspodela radnika prema plati razlikuje od očekivane raspodele. (15)

16. (Jun 2016) U sledećoj tabeli su dati bazni indeksi izvoza jednog proizvoda:

Godina	2009.	2010.	2011.	2012.	2013.	2014.
$I_{(2011=100)}$	72,13	92,46	100	109,18	118,36	134,43

- a) U kojoj godini je ostvarena najmanja relativna promena izvoza proizvoda u odnosu na 2009. godinu? (10)
- b) Ako je izvoz proizvoda u 2012. iznosio 333 hiljade dolara, kolika je apsolutna a kolika relativna promena izvoza proizvoda u 2013. godini u odnosu na 2012. godinu? (5)
- c) Po kojoj stopi se izvoz proizvoda u periodu od 2010-2013. godine u proseku godišnje menjao? (5)

17. (Jul 2016) Vrednost prodaje četiri proizvoda u jednoj grani industrije u 2013. godini bila je sledeća: 2700; 2929; 2320 i 4480 dinara. Relativan rast vrednosti prodaje ovih proizvoda u 2015. godini u odnosu na 2013. godinu iznosio je redom: 41,78%; 33,36%; 107,67% i 54,53%. Fizički obim prodaje prvog proizvoda je u 2015. godini povećan za 22,22% u odnosu na 2013. godinu, drugog proizvoda je povećan za 24,75%, trećeg proizvoda je povećan za 25,86% i četvrtog proizvoda je povećan za 15%. Cena prvog proizvoda je u 2015. godini povećana za 16% u odnosu na 2013. godinu, drugog proizvoda je povećana za 6,90%, trećeg proizvoda je povećana za 65% i četvrtog proizvoda je povećana za 34,38%. Metodom agregata odrediti grupni indeks fizičkog obima prodaje i cene u 2015. godini (2013=100), koristeći pondere iz tekućeg perioda. (20)

18. (Jul 2016) Na osnovu raspoloživih podataka o ceni (u hiljadama dinarima) i obimu ponude (u hiljadama komada) jednog proizvoda u prostom slučajnom uzorku od 6 preduzeća prehrambene industrije dobijeni su sledeći rezultati: suma kvadrata odstupanja cene pojedinih proizvoda od prosečne cene iznosi 538,83 hiljada dinara², ocenjena standardna devijacija ocene regresionog parametra uz objašnjavajuću promenljivu je 230 dinara, učešće neobjašnjene varijabilnosti u ukupnom varijabilitetu iznosi 3,54%, prosečan obim ponude proizvoda iznosi 84,17 hiljada komada, a prosečna cena 40,17 hiljada dinara.

- a) Oceniti prost linearni regresioni model i objasniti dobijeni regresioni koeficijent nagiba. (10)
- b) Sa pouzdanošću od 95% predvideti obim ponude proizvoda, ako cena iznosi 46 hiljada dinara. (10)

19. (Jul 2016) Posmatra se prodaja jednog proizvoda (u hiljadama komada) i prihodi od prodaje (u hiljadama evra). U prostom slučajnom uzorku od 6 tržišnih segmenata zabeleženi su sledeći rezultati: prosečan prihod je 8.330 evra; prosečna prodaja iznosi 35,5 hiljada komada; prihodi po pojedinim tržišnim segmentima u proseku odstupaju od ocenjenog prihoda za 2.440 evra; prosečno kvadratno odstupanje prodaje proizvoda po pojedinim tržišnim segmentima od prosečne prodaje iznosi 193,5 hiljada komada²; prihodi od prodaje po pojedinim tržišnim segmentima u proseku odstupaju od prosečnog prihoda za 4.590 evra.

- a) Oceniti prost linearni regresioni model i objasniti dobijeni regresioni koeficijent nagiba. (10)
- b) Sa pouzdanošću od 95% oceniti prosečan prihod od prodaje u osnovnom skupu svih tržišnih segmenata, ako prodaja iznosi 40 hiljada komada. (10)

20. (Septembar 2016) U sledećoj tabeli dati su bazni indeksi uvoza jednog prehrambenog proizvoda (prosek 2013=100) i sezonskoj i rezidualnoj komponenti u periodu 2012-2015. godine, po kvartalima:

Godina	2012.		2013.				2014.				2015.	
Kvartal	III	IV	I	II	III	IV	I	II	III	IV	I	II
I_t (prosek 2013=100)	46,60	69,90	174,76	89,32	54,37	81,55	182,52	108,74	66,02	97,09	205,83	124,27
Sezonska i rezidualna komponenta	-	-	1,82	0,91	0,54	0,78	1,69	0,97	0,57	0,80	-	-

a) Uvoz posmatranog proizvoda je u I kvartalu 2014. godine iznosio 47 hiljada tona. Predvideti uvoz posmatranog proizvoda u I kvartalu 2016. godine uzimajući u obzir uticaj sezonskih faktora. (20)

b) Izračunati prosečnu stopu rasta uvoza posmatranog proizvoda od IV kvartala 2012. do IV kvartala 2014. godine. (5)

21. (Septembar 2016) Dati su podaci o ulaganju u istraživanje i razvoj (u hiljadama dinara) i prihodima od prodaje jednog proizvoda (u hiljadama dinara) u prostom slučajnom uzorku od šest kompanija.

Ulaganje	60	78	42	84	101	125
Prihod	81	93	55	106	118	95

a) Izračunati sumu kvadrata objašnjenog varijabiliteta. Koliko iznosi prosečno odstupanje prihoda od prodaje proizvoda od linije regresije u uzorku? (10)

b) Ako zajednička raspodela posmatranih promenljivih odstupa značajno od normalne, na nivou značajnosti od 5% ispitati da li između ulaganja u istraživanje i razvoj i prihoda od prodaje datog proizvoda postoji kvantitativno slaganje. (10)

22. (Jun 2015) Raspoložemo podacima o uvozu sirovina (u hiljadama dinara): 20, 33, 39, 48, 56, 65 i vrednosti proizvodnje (u milionima dinara): 97, 105, 112, 124, 135, 142 u prostom slučajnom uzorku od šest preduzeća.

a) Sa pouzdanosću od 95% oceniti za koliko će se u proseku promeniti vrednost proizvodnje, ako se uvoz poveća za 5 hiljada dinara. (10)

b) Pod pretpostavkom da zajednička raspodela uvoza sirovina i vrednosti proizvodnje ne odstupa značajno od normalne, na nivou značajnosti od 5%, ispitati da li između posmatranih promenljivih postoji kvantitativno slaganje. (10)

ИСПИТНИ ЗАДАЦИ – РЕШЕЊА ЗАДАТАКА

(хи-квадрат тест, анализа варијансе, корелација, регресија, индексни бројеви, анализа временских серија)

1.

I Formulisanje nulte i alternativne hipoteze

H_0 : između starosti ispitanika i učestalosti plivanja ne postoji povezanost

H_1 : između starosti ispitanika i učestalosti plivanja postoji povezanost

II Izbor raspodele koja se koristi: Pošto se ispituje nezavisnost obeležja koristi se hi-квадрат test nezavisnosti

III Određivanje oblasti odbacivanja i neodbacivanja $\chi^2_{(R-1)(K-1); \alpha} = \chi^2_{4; 0,05} = 9,488$

IV Izračunavanje realizovane vrednosti statistike testa

%	O	E	O-E	(O-E) ²	(O-E) ² /E
8	4	5.44	-1.44	2.07	0.38
10	5	5.44	-0.44	0.19	0.04
14	7	5.12	1.88	3.53	0.69
12	6	6.46	-0.46	0.21	0.03
10	5	6.46	-1.46	2.13	0.33
16	8	6.08	1.92	3.69	0.61
14	7	5.10	1.90	3.61	0.71
14	7	5.10	1.90	3.61	0.71
2	1	4.80	-3.80	14.44	3.01
100	50	50	0	33.4912	6.500137

Npr. $E_1 = \frac{\text{sumareda} \cdot \text{sumakolone}}{n} = \frac{17 \cdot 16}{50} = 5,44$

Uslov: $E \geq 5$ nije ispunjen za $O=1$.

1) Redovi se spajaju

Kritična vrednost: $\chi^2_{(R-1)(K-1); \alpha} = \chi^2_{2; 0,05} = 5,991$ Realizovana vrednost: $\chi^2 = 1,70$

2) Kolone sa spajaju

Kritična vrednost: $\chi^2_{(R-1)(K-1); \alpha} = \chi^2_{2; 0,05} = 5,991$ Realizovana vrednost: $\chi^2 = 1,63$

V Donošenje odluke: Na nivou značajnosti od 5%, nulta hipoteza se odbacuje i zaključuje se da između starosti ispitanika i učestalosti plivanja postoji povezanost.

2. a)

$b_1 = 1,03$

$SKN = SK_{yy} - b_1 \cdot SP_{xy} = 67,46$

$b_0 = 28,41$

b) $H_0: \beta_1 = 0$ $H_1: \beta_1 \neq 0$

kritična vrednost: $t_{n-2; \frac{\alpha}{2}} = t_{7; 0,025} = 2,365$ realizovana vrednost: $t = \frac{b_1}{s_{b_1}} = \frac{1,03}{0,042} = 24,52$

Na nivou značajnosti od 5% zaključuje se da uvoz sirovina značajno utiče na proizvodnju.

c)

$b_1 - t \cdot s_{b_1} \leq \beta_1 \leq b_1 + t \cdot s_{b_1}$

$1,03 - 2,365 \cdot 0,042 \leq \beta_1 \leq 1,03 + 2,365 \cdot 0,042$

$0,931 \leq \beta_1 \leq 1,129$

$2,793 \leq 3\beta_1 \leq 3,387$

Ako se uvoz poveća za 3 miliona dinara, ostvarena proizvodnja će se u proseku promeniti u intervalu od 2,793 do 3,387 hiljada tona.

3.

$$H_0: \mu_1 = \mu_2 = \mu_3; H_1: \mu_i \neq \mu_j$$

I Formulisanje nulte i alternativne hipoteze

II Izbor raspodele koja se koristi: osnovni skupovi imaju normalnu raspodelu; varijanse skupova su jednake; uzorci su slučajni i nezavisni – koristi se F raspodela

III Određivanje oblasti odbacivanja i neodbacivanja $F_{k-1, n-k, \alpha} = F_{2; 12; 0,05} = 3,89$

IV Izračunavanje realizovane vrednosti statistike testa

$$V_A = \frac{S_A}{k-1} = 13,07$$

$$V_R = \frac{S_R}{n-k} = 10,47$$

$$F = \frac{V_A}{V_R} = 1,25$$

V Donošenje odluke

Na nivou značajnosti od 5%, nulta hipoteza se ne odbacuje i zaključuje se da godine starosti ispitanika ne utiču sistematski na varijabilitet vremena koje ispitanici provedu u teretani.

4. a)

$$b_1 = 0,4$$

$$b_0 = 7,36$$

$$r^2 = \frac{b_1 \cdot SP_{xy}}{SK_{xx}} = 97,06\%$$

97,06% varijacija izdataka za garderobu je objašnjeno varijacijama plate.

b)

$$b_1 - t \cdot s_{b_1} \leq \beta_1 \leq b_1 + t \cdot s_{b_1}$$

$$0,4 - 2,776 \cdot 0,035 \leq \beta_1 \leq 0,4 + 2,776 \cdot 0,035$$

$$0,303 \leq \beta_1 \leq 0,497$$

$$3,03 \leq 10\beta_1 \leq 4,97$$

Ako se plata poveća za 10 hiljada dinara, izdaci za garderobu će se u proseku povećati u intervalu od 3,03 do 4,97 hiljada dinara.

5.

poqo	Ipq	Ip	Iq
2800	146.00	112.00	130.36
2520	178.10	120.00	148.41
2688	170.57	109.38	155.95
2480	203.23	105.00	193.55

$$p_t q_t = \frac{I_{pq} \cdot P_0 q_0}{100}$$

$$q_0 p_t = \frac{p_t q_t \cdot 100}{I_q}$$

$$p_0 q_t = \frac{p_t q_t \cdot 100}{I_p}$$

$$I_q = \frac{18201}{11704} \cdot 100 = 155,51$$

$$I_p = \frac{18201}{16382} \cdot 100 = 111,1$$

6.

I Formulisanje nulte i alternativne hipoteze

H_0 : raspodela studenata prema visini stipendije se statistički značajno ne razlikuje od očekivane raspodele

H_1 : raspodela studenata prema visini stipendije se statistički značajno razlikuje od očekivane raspodele

II Izbor raspodele koja se koristi: $k > 2$ – koristi se hi-kvadrat test prilagođenosti

III Određivanje oblasti odbacivanja i neodbacivanja $\chi^2_{k-1;\alpha} = \chi^2_{3;0,05} = 7,815$

IV Izračunavanje realizovane vrednosti statistike testa

Proporcije u osnovnom skupu su redom: 0,5; 0,25; 0,15 i 0,1.

Npr. $E_1 = n p_1 = 50 \cdot 0,5 = 25$

USLOV: $E \geq 5$ je ispunjen Realizovana vrednost: $\chi^2 = 2,33$

V Donošenje odluke

Na nivou značajnosti od 5%, nulta hipoteza se ne odbacuje i zaključuje se da se raspodela studenata prema visini stipendije statistički značajno ne razlikuje od očekivane raspodele.

7. a)

SKxx	577.33
SKyy	4502.83
SPxy	-1580.67

$$b_1 = -2,74$$

$$b_0 = 93,85$$

$$SKN = SK_{yy} - b_1 \cdot SP_{xy} = 171,79$$

$$\frac{SKN}{SKU} \cdot 100 = \frac{171,79}{4502,83} \cdot 100 = 3,82\%$$

b)

$$b_1 - t \cdot s_{b_1} \leq \beta_1 \leq b_1 + t \cdot s_{b_1}$$

$$-2,74 - 2,776 \cdot 0,27 \leq \beta_1 \leq -2,74 + 2,776 \cdot 0,27$$

$$-3,49 \leq \beta_1 \leq -1,99$$

$$-6,98 \leq 2\beta_1 \leq -3,98$$

Ako se cena poveća za 2 stotine dinara, tražnja će se u proseku smanjiti u intervalu od 3,98 do 6,98 kg.

8. a)

$$r_g = \left(8 - \sqrt{\frac{76}{y_{2007}}} - 1\right) \cdot 100 = 11,71\%$$

$$y_{2007} = 35$$

$$\frac{y_{2007}}{\hat{y}_{2007}} \cdot 100 = \frac{35}{\hat{y}_{2007}} \cdot 100 = 104,76$$

$$\hat{y}_{2007} = 33,41$$

$$\hat{y}_{2007} = b_0 + b_1 \cdot t \quad b_0 = \bar{y} - b_1 \cdot \bar{t}$$

$$33,41 = b_0 + b_1 \cdot 1 \quad b_0 = 54,13 - b_1 \cdot 4,5$$

$$b_0 = 27,49$$

$$b_1 = 5,92$$

$$I_t = \frac{y_{2014}}{y_{2007}} \cdot 100 = \frac{76}{35} \cdot 100 = 217,14$$

$$y_{2014} - y_{2007} = 41$$

Relativna promena iznosi 117,14%, a apsolutna 41 kg.

b)

$$\hat{y}_{2016} = 27,49 + 5,92 \cdot 10 = 86,69$$

Predviđena vrednost je validna.

9.

$$a) I_t = \frac{y_t}{y_0} \cdot 100$$

$$y_0 = \frac{y_{II,2013} \cdot 100}{I_{II,2013}} = 35,25$$

$$y_t = \frac{I_t \cdot y_0}{100} \quad y_{st} = \frac{y_t}{I_s} \cdot 100$$

	It	Yt	Yst	t	t ²	t*Yst
	42.55	15	34.48276	1	1	34.48276
	102.13	36	33.96226	2	4	67.92453
	164.54	58	32.40223	3	9	97.2067
	59.57	21	30.21583	4	16	120.8633
	48.23	17	39.08046	5	25	195.4023
	99.29	35	33.01887	6	36	198.1132
	178.72	63	35.19553	7	49	246.3687
	73.76	26	37.41007	8	64	299.2806
	36.88	13	29.88506	9	81	268.9655
	110.64	39	36.79245	10	100	367.9245
	190.07	67	37.43017	11	121	411.7318
	51.06	18	25.89928	12	144	310.7914
SUMA			405.775	78	650	2619.055

Ocenjena linija trenda na desezoniranim podacima: $\hat{y}_{st} = 34,66 - 0,129 \cdot t$

$$\hat{y}_{III,2015}^* = \frac{\hat{y}_{III,2015} \cdot I_s}{100} = \frac{32,72 \cdot 179}{100} = 58,57$$

10.

$$H_0: \mu_1 = \mu_2 = \mu_3 \quad H_1: \mu_i \neq \mu_j$$

I Formulisanje nulte i alternativne hipoteze

II Izbor raspodele koja se koristi: osnovni skupovi imaju normalnu raspodelu; varijanse skupova su jednake; uzorci su slučajni i nezavisni – koristi se F raspodela

$$F_{k-1, n-k, \alpha} = F_{2; 12; 0,05} = 3,89$$

III Određivanje oblasti odbacivanja i neodbacivanja

IV Izračunavanje realizovane vrednosti statistike testa

$$V_A = \frac{S_A}{k-1} = 15$$

$$V_R = \frac{S_R}{n-k} = 9,2$$

$$F = \frac{V_A}{V_R} = 1,63$$

V Donošenje odluke: Na nivou značajnosti od 5%, nulta hipoteza se ne odbacuje.

11.

$$H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4 \quad H_1: \mu_i \neq \mu_j$$

I Formulisanje nulte i alternativne hipoteze

II Izbor raspodele koja se koristi: osnovni skupovi imaju normalnu raspodelu; varijanse skupova su jednake; uzorci su slučajni i nezavisni – koristi se F raspodela

$$F_{k-1, n-k, \alpha} = F_{3; 12; 0,05} = 3,49$$

III Određivanje oblasti odbacivanja i neodbacivanja

IV Izračunavanje realizovane vrednosti statistike testa

$$V_A = \frac{S_A}{k-1} = 101,06$$

$$V_R = \frac{S_R}{n-k} = 54,19$$

$$F = \frac{V_A}{V_R} = 1,86$$

V Donošenje odluke: Na nivou značajnosti od 5%, nulta hipoteza se ne odbacuje.

$$12. \quad a) \quad s_y^2 = \frac{SK_{yy}}{n-1} \quad SK_{yy} = 125,31 \quad s_{b1} = \frac{s}{\sqrt{SK_{xx}}} \quad s = 1,12 \quad s = \sqrt{\frac{SK_{yy} - b_1 \cdot SP_{xy}}{n-2}} = \sqrt{\frac{SK_{yy} - \frac{SP_{xy}^2}{SK_{xx}}}{n-2}}$$

$$SP_{xy} = 73,54 \quad SK_{xx} = 45,43 \quad \bar{x} = 5,71 \quad b_1 = 1,62 \quad b_0 = -0,96$$

$$b) \quad \hat{y}_p - t_{s_{y_p}} \leq Y_p \leq \hat{y}_p + t_{s_{y_p}}$$

$$18,48 - 2,571 \cdot 1,59 \leq Y_p \leq 18,48 + 2,571 \cdot 1,59$$

$$14,39 \leq Y_p \leq 22,57$$

Uslovi su ispunjeni:

$$1) \quad r^2 = 95,07\%$$

2) ekstrapolacija je u neposrednoj blizini podataka iz uzorka

3) $\beta_1 \neq 0$ (kritična vrednost je $t=2,571$, realizovana vrednost je $t=9,76$)

13. a)

I Formulisanje nulte i alternativne hipoteze

$$H_0: \beta_1 = 0$$

$$H_1: \beta_1 \neq 0$$

$$H_0: \beta_2 = 0$$

$$H_1: \beta_2 \neq 0$$

II Izbor raspodele koja se koristi: Pošto je σ nepoznato – koristi se t raspodela

III Određivanje oblasti odbacivanja i neodbacivanja $t_{n-k-1, \frac{\alpha}{2}} = t_{5, 0,025} = 2,571$

IV Izračunavanje realizovane vrednosti statistike testa

$$t = \frac{b_1}{s_{b_1}} = \frac{2,56}{2,12} = 1,21$$

$$t = \frac{b_2}{s_{b_2}} = \frac{3,84}{1,16} = 3,31$$

$$1 - \frac{2}{5} (1 - 0,945) = 0,23$$

V Donošenje odluke: Na nivou značajnosti od 5%, zaključuje se da starost gledalaca ne utiče, a da slobodno vreme utiče na gledanost televizijskog programa.

b)

$$\bar{R}^2 = 1 - (1 - R^2) \left(\frac{n-1}{n-k-1} \right) = 0,923$$

$$1 - \bar{R}^2 = 1 - 0,923 = 0,077$$

Procenat varijacija gledanosti televizijskog programa koji nije objašnjen zajedničkim uticajem starosti gledalaca i slobodnog vremena iznosi 7,7%.

c)

$$b_1 - t \cdot s_{b_1} \leq \beta_1 \leq b_1 + t \cdot s_{b_1}$$

$$-2,89 \leq \beta_1 \leq 8,01$$

$$-0,72 \leq 0,25\beta_1 \leq 2$$

Ako se starost gledalaca poveća za 3 meseca, a slobodno vreme ostane nepromenjeno, gledanost televizijskog programa se u proseku neće promeniti, jer starost ne utiče na gledanost televizijskog programa.

14. a) Linearni trend: $\hat{y}_t = 99,28 + 1,38 \cdot t$ $SKG = 98,46$ SKG linearnog trenda je manja, bira se linearni trend.

$$b) \quad \frac{y_t}{\hat{y}_t} \cdot 100 = \frac{91}{107,56} \cdot 100 = 84,6$$

Fizički obim proizvodnje u junu 2015. odstupa od ocenjene funkcije trenda za -15,4%.

15.

I Formulisanje nulte i alternativne hipoteze

H_0 : raspodela radnika prema plati se ne razlikuje od očekivane raspodele

H_1 : raspodela radnika prema plati se razlikuje od očekivane raspodele

II Izbor raspodele koja se koristi: $k > 2$ – koristi se hi-kvadrat test prilagođenosti

III Određivanje oblasti odbacivanja i neodbacivanja

$$\chi^2_{k-1;\alpha} = \chi^2_{3;0,1} = 6,251$$

IV Izračunavanje realizovane vrednosti statistike testa

Proporcije u osnovnom skupu su redom: 0,4; 0,2, 0,1 i 0,3.

$$\text{Npr. } E_1 = np_1 = 200 \cdot 0,4 = 80$$

Uslov: $E \geq 5$ je ispunjen

$$\text{Realizovana vrednost: } \chi^2 = 3,75$$

V Donošenje odluke: Na nivou značajnosti od 10%, nulta hipoteza se ne odbacuje i zaključuje se da se raspodela radnika prema plati ne razlikuje od očekivane.

16. a)

It 2011=100	It 2009=100
72.13	100
92.46	128.19
100	138.64
109.18	151.37
118.36	164.09
134.43	186.37

U 2010. godini u odnosu na 2009. je ostvarena najmanja relativna promena (28,19%).

b)

$$y_{2012} = 333$$

$$\frac{333}{y_0} \cdot 100 = 109,18$$

$$y_0 = 305$$

$$\frac{y_{2013}}{305} \cdot 100 = 118,36$$

$$y_{2013} = 361$$

Apsolutna promena 28 hiljada dolara.

Relativna promena 8,41%.

c)

$$r_g = \left(\sqrt[3]{\frac{164,09}{128,19}} - 1 \right) \cdot 100 = 8,58\%$$

17.

poqo	Ipq	Ip	Iq	ptqt	poqt	ptqo
2700.00	141.78	116	122.22	3828.06	3300.052	3132.106
2929.00	133.36	106.9	124.75	3906.114	3653.989	3131.154
2320.00	207.67	165	125.86	4817.944	2919.966	3828.018
4480.00	154.53	134.38	115	6922.944	5151.767	6019.951
				19475.06	15025.77	16111.23

$$p_t q_t = \frac{I_{pq} \cdot p_0 q_0}{100}$$

$$q_0 p_t = \frac{p_t q_t \cdot 100}{I_q}$$

$$p_0 q_t = \frac{p_t q_t \cdot 100}{I_p}$$

$$I_q = \frac{19475,05}{16111,23} \cdot 100 = 120,88$$

$$I_p = \frac{19475,05}{15025,76} \cdot 100 = 129,61$$

18. a)

$$s_{b1} = \frac{s}{\sqrt{SK_{xx}}} \rightarrow s = 5,34 \rightarrow SKN = 114,06$$

$$\frac{SKN}{SKU} = 3,54\% \rightarrow SKU = 3222,03$$

$$r^2 = 96,46\% \rightarrow r = 0,98 \quad SP_{xy} = 1291,27$$

$$b_1 = 2,4$$

$$b_0 = -12,24$$

b)

$$\hat{y}_p - t_{s_{y_p}} \leq Y_p \leq \hat{y}_p + t_{s_{y_p}}$$

$$98,16 - 2,776 \cdot 5,92 \leq Y_p \leq 98,16 + 2,776 \cdot 5,92$$

$$81,73 \leq Y_p \leq 114,59$$

19. a)

$$s = \sqrt{\frac{SKN}{n-2}} = 2,44 \rightarrow SKN = 23,81$$

$$s_X^2 = \frac{SK_{xx}}{n-1} \rightarrow SK_{xx} = 967,5$$

$$SKO = 81,53$$

$$s_Y^2 = \frac{SK_{yy}}{n-1} \rightarrow SK_{yy} = 105,34$$

$$r^2 = 77,4\% \rightarrow r = 0,88$$

$$SP_{xy} = 280,93$$

$$b_1 = 0,29$$

$$b_0 = -1,97$$

b)

$$\hat{y}_p - t_{s_{\hat{y}_p}} \leq \mu_{Y/X} \leq \hat{y}_p + t_{s_{\hat{y}_p}}$$

$$9,63 - 2,776 \cdot 1,057 \leq \mu_{Y/X} \leq 9,63 + 2,776 \cdot 1,057$$

$$6,7 \leq \mu_{Y/X} \leq 12,56$$

20. a)

$$I_t = \frac{y_t}{y_0} \cdot 100 \quad y_0 = \frac{y_{I,2014} \cdot 100}{I_{I,2014}} = 25,75$$

$$y_t = \frac{I_t \cdot y_0}{100}$$

It	Yt	SR
46.60	12	-
69.90	18	-
174.76	45	1.82
89.32	23	0.91
54.37	14	0.54
81.55	21	0.78
182.52	47	1.69
108.74	28	0.97
66.02	17	0.57
97.09	25	0.8
205.83	53	-
124.27	32	-

Is
175.5
94
55.5
79

$$\hat{y}_{st} = 21,8 + 2,53 \cdot t$$

$$\hat{y}_{I,2016} = 21,8 + 2,53 \cdot 15 = 59,75$$

$$\hat{y}_{I,2016}^* = \frac{\hat{y}_{I,2016} \cdot I_s}{100} = \frac{59,75 \cdot 175,5}{100} = 104,86$$

$$b) r_g = \left(\sqrt[8]{\frac{25}{18}} - 1 \right) \cdot 100 = 4,19\%$$

21. a)

$$b_1 = 0,55$$

$$SKO = 1302,22$$

$$s = 16,33$$

b)

$$I \quad H_0: \rho_s = 0 \quad H_1: \rho_s \neq 0$$

II Uslov: pošto zajednička raspodela promenljivih odstupa od normalne koristi se postupak testiranja Spirmanovog koeficijenta korelacije ranga

$$III \quad r_{s,6;0,05} = \pm 0,886$$

IV

$$r_s = 1 - \frac{6 \cdot 6}{n(n^2 - 1)} = 0,83$$

V Na nivou značajnosti od 5% nulta hipoteza se ne odbacuje i zaključuje se da između promenljivih ne postoji kvantitativno slaganje.

22.a)

$$b_1 - t \cdot s_{b_1} \leq \beta_1 \leq b_1 + t \cdot s_{b_1}$$

$$b_1 = 1,07$$

$$1,07 - 2,776 \cdot 0,072 \leq \beta_1 \leq 1,07 + 2,776 \cdot 0,072$$

$$b_0 = 72,63$$

$$0,87 \leq \beta_1 \leq 1,27$$

$$4,35 \leq \beta_1 \leq 6,35$$

Ako se uvoz poveća za 5 hiljada dinara, vrednost proizvodnje će se u proseku promeniti u intervalu od 4,35 do 6,35 hiljada dinara.

b)

I Formulisanje nulte i alternativne hipoteze

H_0 : između posmatranih promenljivih ne postoji linearno kvantitativno slaganje

H_1 : između posmatranih promenljivih postoji linearno kvantitativno slaganje

II Izbor raspodele koja se koristi: Pod pretpostavkom da zajednička raspodela uvoza sirovina i vrednosti proizvodnje ne odstupa značajno od normalne – koristi se t raspodela

III Određivanje oblasti odbacivanja i neodbacivanja $t_{n-2;0,025} = 2,776$

IV Izračunavanje realizovane vrednosti statistike testa

$$r = \frac{SP_{xy}}{\sqrt{SK_{xx}SK_{yy}}} = 0,99 \quad t = r \sqrt{\frac{n-2}{1-r^2}} = 14$$

V Donošenje odluke

Na nivou značajnosti od 5%, nulta hipoteza se odbacuje i zaključuje se da između posmatranih promenljivih postoji linearno kvantitativno slaganje.