

# Sadržaj

<b>1</b>	<b>Prikazivanje statističkih podataka</b>	<b>5</b>
1.1	Kvantitativni podaci . . . . .	5
1.1.1	Diskretnog tipa . . . . .	5
1.1.2	Neprekidnog tipa . . . . .	7
1.2	Kategorijalni podaci . . . . .	8
1.2.1	Nominalni . . . . .	8
1.2.2	Ordinalni . . . . .	9
	Vežbe za samostalni rad . . . . .	10
<b>2</b>	<b>Deskriptivna statistika</b>	<b>11</b>
2.1	Mere centralne tendencije . . . . .	11
2.2	Mere odstupanja . . . . .	16
	Vežbe za samostalni rad . . . . .	20
<b>3</b>	<b>Intervali poverenja za sredinu populacije</b>	<b>23</b>
	Vežbe za samostalni rad . . . . .	31
<b>4</b>	<b>t-test za sredinu populacije</b>	<b>35</b>
	Vežbe za samostalni rad . . . . .	42
<b>5</b>	<b>Analiza varijanse (ANOVA)</b>	<b>45</b>
	Vežbe za samostalni rad . . . . .	50
<b>6</b>	<b>Linearna korelacija</b>	<b>53</b>
	Vežbe za samostalni rad . . . . .	59
<b>7</b>	<b>Linearna regresija</b>	<b>61</b>
	Vežbe za samostalni rad . . . . .	69

<b>8 Primeri kolokvijuma</b>	<b>73</b>
<b>A <math>z</math>-tablica</b>	<b>79</b>
<b>B <math>t</math>-tablica</b>	<b>83</b>
<b>C <math>F</math>-tablica, <math>\alpha = 0.05</math></b>	<b>87</b>
<b>D <math>F</math>-tablica, <math>\alpha = 0.01</math></b>	<b>89</b>

# Predgovor

Ova zbirka zadataka po strukturi i po sadržaju prati udžbenik *Ljiljana Cvetković: Poslovna statistika, Futura, Novi Sad, 2006.*

Nakon rešenih zadataka, u svakom poglavlju navedeni su i zadaci za samostalni rad.

Ukoliko ova zbirka doprinese boljem razumevanju osnovnih načela poslovne statistike i osposobi čitaoca za statistički način razmišljanja, sa ciljem donošenja dobrih poslovnih odluka, biće zadovoljni...

*Autori*



# Prikazivanje statističkih podataka

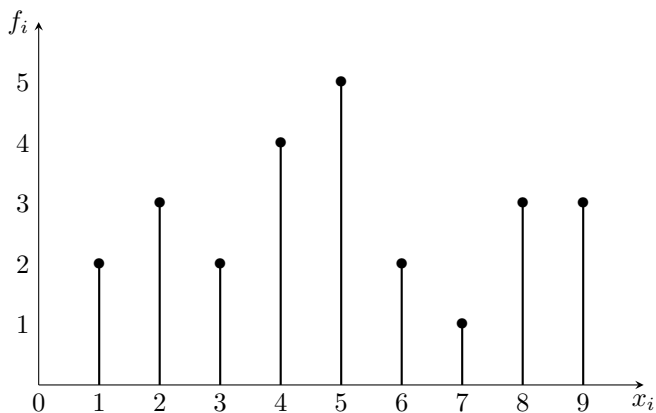
## 1.1 Kvantitativni podaci

### 1.1.1 Diskretnog tipa

**Vežba 1.1** U 25 slučajno odabranih ulica u gradu izbrojano je redom 1, 2, 3, 5, 4, 9, 4, 9, 4, 9, 5, 4, 3, 1, 2, 5, 6, 2, 7, 8, 8, 5, 6, 5, 8 kafića. Ove podatke prikazati tabelarno i grafički.

*Rešenje:* Obeležje koje posmatramo je broj kafića u ulici u gradu. U tabeli navodimo, u prvoj koloni, njegove varijetete, poredane u rastućem redosledu, a u drugoj koloni broj pojavljivanja (frekvencije) svakog od varijeteta u datom uzorku.

$x_i$	$f_i$
1	2
2	3
3	2
4	4
5	5
6	2
7	1
8	3
9	3
$\Sigma$	25

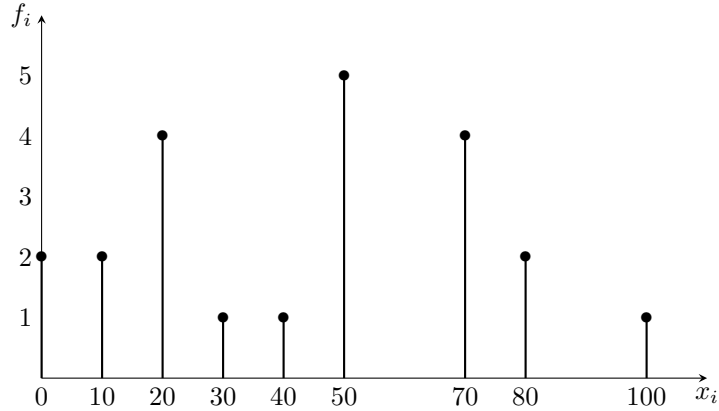


S obzirom da je posmatrano obeležje diskretnog tipa, grafička reprezentacija podataka je data skupom tačaka i vertikalnih linija do svake od njih.

**Vežba 1.2** *Studenti koji su polagali ispit iz statistike, u septembarskom ispitnom roku, osvojili su sledeće poene: 70, 80, 70, 50, 40, 20, 50, 100, 10, 20, 70, 0, 0, 70, 10, 30, 20, 20, 50, 50, 50, 80. Ove podatke prikazati tabelarno i grafički.*

Rešenje:

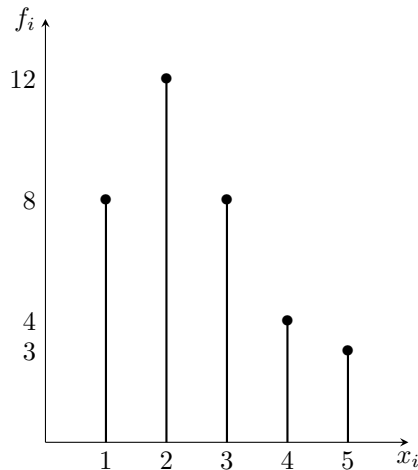
$x_i$	$f_i$
0	2
10	2
20	4
30	1
40	1
50	5
70	4
80	2
100	1
$\Sigma$	22



**Vežba 1.3** *U okviru istraživanja životnog ciklusa jedne vrste sisara ispitivan je broj mladunaca iz jednog legla. Na uzorku od 35 legala zabeležen je sledeći broj mladunaca po leglu: 1, 1, 2, 4, 3, 2, 2, 2, 2, 2, 3, 4, 5, 5, 1, 2, 3, 2, 3, 2, 3, 2, 2, 1, 1, 1, 4, 3, 1, 5, 4, 3, 3, 2, 1. Podatke prikazati grafički i tabelom sa relativnim frekvencijama.*

Rešenje:

$x_i$	$f_i$	$p_i$
1	8	23%
2	12	34%
3	8	23%
4	4	11%
5	3	9%
$\Sigma$	35	100%



Relativne frekvencije predstavljaju odnos svake od frekvencija spram veličine uzorka (procentualni udeo svakog varijeteta). Računaju se po sledećoj formuli i iskazuju u

procentima

$$p_i = \frac{f_i}{n}, \quad i = 1, 2, \dots, k,$$

pri čemu je  $k$  broj varijeteta.

### 1.1.2 Nепrekidnog tipa

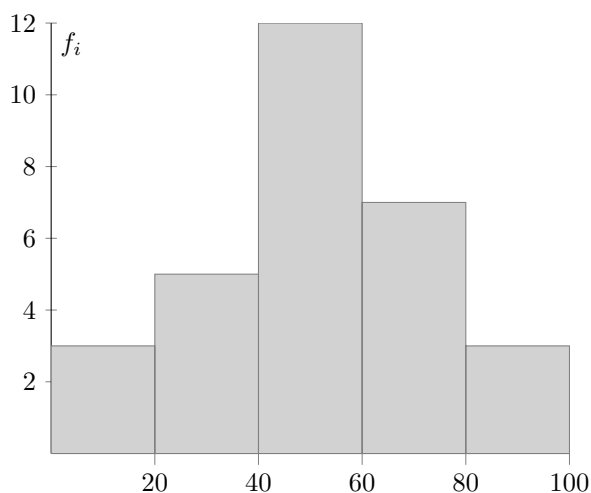
**Vežba 1.4** U okviru istraživanja životnog standarda stanovnika jednog grada, registrovan je prihod u 30 slučajno izabranih domaćinstava. Dobijeni su sledeći podaci (u dinarima):

12111, 13211, 15008, 23330, 54002, 81900, 76901, 51222, 45332, 78652, 91999, 98233, 53988, 65444, 45873, 42390, 76449, 76122, 34233, 76338, 34765, 43573, 37854, 74282, 29112, 47876, 54343, 59867, 54334, 45432.

Podatke najpre grupisati u 5 jednakih intervala u rasponu od 0 do 100, izraženo u hiljadama dinara, a zatim ih prikazati tabelarno i grafički.

Rešenje:

Interval	$f_i$
[ 0 , 20 )	3
[ 20 , 40 )	5
[ 40 , 60 )	12
[ 60 , 80 )	7
[ 80 , 100 ]	3
$\Sigma$	30



Budući da su sada podaci neprekidnog tipa, grafik je predstavljen unijom pravougaonika, pri čemu je svaki od njih širine kao i pripadajući interval, dok je visina određena veličinom frekvencije. Ovaj tip grafika poznat je pod nazivom *histogram*.

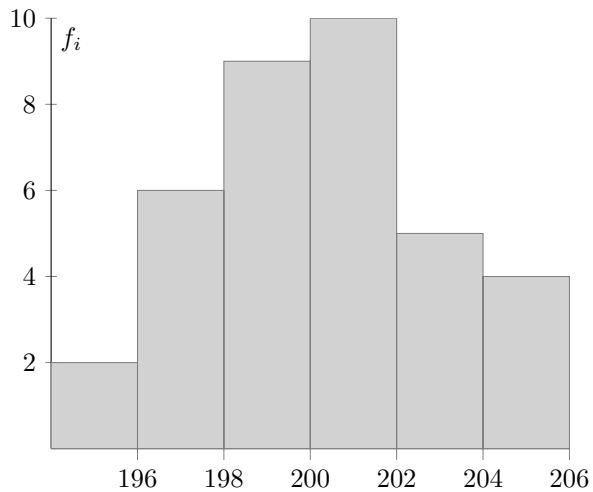
**Vežba 1.5** Radi kontrole kvaliteta vitaminskog napitka pakovanog u flašice od 100 ml, merena je količina  $C$  vitamina u 36 slučajno izabranih napitaka. Dobijeni su sledeći rezultati izraženi u mg:

197, 195, 194, 196, 196, 198, 197, 199, 201, 200, 201, 204, 206, 205, 203, 202, 202, 201, 201, 203, 196, 199, 197, 198, 200, 198, 199, 199, 198, 200, 202, 205, 200, 199, 200, 201.

Podatke najpre grupisati u 6 jednakih intervala, a zatim ih prikazati tabelarno i grafički.

Rešenje:

Interval	$f_i$
[ 194 , 196 )	2
[ 196 , 198 )	6
[ 198 , 200 )	9
[ 200 , 202 )	10
[ 202 , 204 )	5
[ 204 , 206 ]	4
$\Sigma$	36



## 1.2 Kategorijalni podaci

### 1.2.1 Nominalni

**Vežba 1.6** U jednom rasadniku botaničar je odlučio da izvrši analizu popularnog cveta Božićna zvezda (*Poinsettia*). Na raspolaganju ima 17 rasada u tri različite boje: crvena ( $C$ ), roza ( $R$ ) i bela ( $B$ ). Nakon što je obišao rasadnik, botaničar je redom zabeležio svako pojavljivanje na sledeći način:

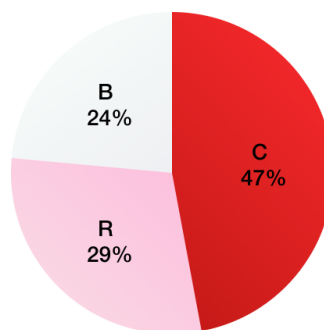
$R, C, C, B, R, R, B, C, C, R, C, C, C, B, B, R, C.$

Ove podatke prikazati najpre tabelarno sa relativnim frekvencijama, a potom i grafički.



Rešenje:

$x_i$	$f_i$	$p_i$
C	8	47%
R	5	29%
B	4	24%
$\Sigma$	17	100%



U ovom slučaju, podaci su dati u vidu neuređenih kategorija, pa smo ih prikazali *pitičastim* grafikonom.

### 1.2.2 Ordinalni

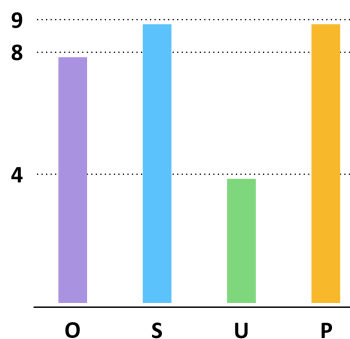
**Vežba 1.7** Poznato je da je *Trichoptera caddisfly* jedan od bioindikatora kvaliteta voda. Povodom svetskog dana voda, grupa ekologa je odlučila da ispita njen kvalitet na dotičnom podneblju, mereći prisustvo bioindikatora *Trichoptera caddisfly* na 30 kontrolnih tačaka na sledeći način: P (prisustvo), U (umereno prisustvo), S (slabo prisustvo), O (odsustvo). Rezultati su sledeći:

P, P, U, S, U, P, O, O, S, U, O, P, S, P, S, U, O, O, S, S, P, O, S, O, S, P, P, O, S, P.

Dobijene podatke prikazati najpre tabelarno, a zatim i grafički.


Rešenje:

$x_i$	$f_i$
O	8
S	9
U	4
P	9
$\Sigma$	30




S obzirom da je u pitanju ordinalno obeležje, moguće je uspostaviti poredak između varijeteta, te smo podatke prikazali *stubičastim* grafikonom.

## Vežbe za samostalni rad

 **Vežba 1.8** Nakon što je na terenskom radu prikupio 17 uzoraka cveta crvene deteline (*Trifolium pratense*), mladi biolog odlučio je da za potrebe dalje analize izurši prebrojavanje cvetova u okviru cvasti i, pri tome, registrovao sledeće brojeve:


55, 63, 65, 52, 68, 80, 50, 53, 57, 72, 70, 59, 62, 64, 67, 74, 62.

Podatke grupisati u 6 jednakih intervala, a potom ih prikazati tabelarno i grafički.

 **Vežba 1.9** Zoolog je, izučavajući jedinke žirafa na podneblju afričkog kontinenta, dokumentovao podatke o visini (u cm) novorođenih jedinki pomenute vrste na jednom lokalitetu u proteklih 5 godina:


178, 182, 159, 151, 168, 178, 187, 185, 169, 180, 181, 175, 184, 190, 158, 182, 150, 188, 183, 170, 165, 166, 177, 184, 162, 183, 167, 158, 162, 160, 157, 163, 170, 184, 159, 162.

Date podatke najpre grupisati u 4 jednaka intervala, a zatim ih prikazati tabelarno i grafički.

 **Vežba 1.10** Na redovnom sistematskom pregledu radnika jednog preduzeća, laborant je prikupio serum 36 osoba muškog pola i za potrebe analize posmatrao nivo kreatin-fosfokinaze (CK) u krvi. Tim putem, dobijene su sledeće vrednosti:


121, 82, 100, 151, 68, 58, 95, 145, 64, 201, 101, 163, 84, 57, 139, 60, 78, 94, 119, 104, 110, 113, 118, 203, 62, 83, 67, 93, 92, 110, 25, 123, 70, 48, 95, 42.

Podatke najpre grupisati u 5 jednakih intervala u rasponu od 20 do 220, a potom ih prikazati tabelarno i grafički.

 **Vežba 1.11** Registrujući boju automobila tokom 24 časa prilikom prolaska naplatne rampe na autoputu, pri čemu su boje sledeće: crna (C), bela (B), crvena (CR), srebrna (S), plava (P), registrovan je sledeći slučajni uzorak:

C, B, B, S, P, C, C, S, CR, S, CR, B, C, S, P, B, S, S, C, C, CR, C, CR, P.

Prikupljene podatke prikazati tabelarno i grafički.

 **Vežba 1.12** Uspeh koji učenik može ostvariti u toku školske godine može primiti sledeće vrednosti: nedovoljan (N), dovoljan (D), dobar (DO), vrlo dobar (VD) i odličan (O). Na kraju polugodišta, razredni starešina je, za potrebe analize uspešnosti svojih učenika, prikupio sledeće podatke:

O, N, D, DO, VD, O, O, D, DO, DO, VD, O, VD, O, N, D, VD, VD, D, VD, D, DO, DO, O, O, O, VD, D, O, D.

Date podatke prikazati tabelom sa relativnim frekvencijama, a potom i grafički.

# Deskriptivna statistika

## 2.1 Mere centralne tendencije


**Sredina** (srednja vrednost, aritmetička sredina) *uzorka*

$$\bar{x} = \frac{\sum_{i=1}^k x_i f_i}{n} = \frac{x_1 f_1 + x_2 f_2 + \dots + x_k f_k}{n},$$

- varijeteti :  $x_i$ ,  $i = 1, 2, \dots, k$
- njihove frekvencije :  $f_i$ ,  $i = 1, 2, \dots, k$
- veličina (obim) uzorka :  $n = f_1 + f_2 + \dots + f_k$

**Medijana** je vrednost koja se nalazi tačno na sredini, kada se varijeteti poredaju u rastućem redosledu i svaki navede onoliko puta kolika mu je frekvencija.

**Mod** je varijetet koji ima najveću frekvenciju.

 **Vežba 2.1** U okviru istraživanja životnog ciklusa jedne vrste sisara ispitivao se broj mladunaca iz jednog legla. Prikupljeni podaci o 800 slučajno izabranih legala svrstani su u tabelu:

broj mladunaca	1	2	3	4	5
broj legala	145	240	194	126	95

Odrediti mod, medijanu i aritmetičku sredinu broja mladunaca po leglu.


*Rešenje:* Najpre ćemo definisati šta je obeležje i šta su njegovi varijeteti. Obeležje je broj mladunaca iz jednog legla, a njegovi varijeteti su 1,2,3,4 i 5. Njih smeštamo u prvu

kolonu pomoćne tabele (levo), u kojoj vršimo sva neophodna izračunavanja, pre upotrebe formula. Poslednju vrstu tretiramo kao zbirnu, i u njoj računamo sumu svake kolone, osim poslednje. Nakon kompletiranja pomoćne tabele, prelazimo na izračunavanja pomoću formula.

$x_i$	$f_i$	$x_i f_i$	$F_i$		
1	145	145	145	$n$	800
2	240	480	385	mod	2
3	194	582	579	med	3
4	126	504	705	$\bar{x}$	2.7325
5	95	475	800		
$\Sigma$	800	2186			

Obim uzorka dobijamo sabiranjem frekvencija i on iznosi 800. Najveća frekvencija iznosi 240 i ona odgovara varijetetu 2, stoga je mod posmatranog obeležja 2. Postupak za određivanje medijane je sledeći. Najpre ćemo, u pomoćnoj tabeli, u poslednjoj koloni sračunati *kumulativne* frekvencije ( $F_i$ ) i konstatovati da je obim uzorka paran broj. To znači da su centralne pozicije u rastućem varijacionom nizu pozicije 400. i 401. Preostaje još da očitamo vrednosti varijeteta na ovim pozicijama koristeći podatke o kumulativnim frekvencijama. Tako nalazimo da se na 400. poziciji nalazi 3, a isti je slučaj i sa 401. pozicijom. Medijana je njihova aritmetička sredina, što iznosi 3. Za kraj, upotrebom formule za sredinu, nalazimo da ona iznosi

$$\bar{x} = \frac{2186}{800} = 2.7325.$$

 **Vežba 2.2** Prema podacima Zavoda za statistiku, godišnji prosek mesečnih zarada za 2019. godinu, po oblastima delatnosti, iznosio je:

	br. zaposl. u hiljadama	zarade u RSD
industrija i rudarstvo	643	68319
poljoprivreda, šumarstvo i vodoprivreda	67	42639
ribarstvo	6	42660
prerađivačka industrija	14	42972
građevinarstvo	88	39763
saobraćaj i veze	109	55055
trgovina	112	36542
ugostiteljstvo i turizam	36	27149
proizvodnja el. energije, gasa i vode	26	59849
stambeno komunalna delatnost	47	45083
finansijsko posredovanje	61	99771
obrazovanje i kultura	153	50634
zdravstvo i socijalna zaštita	161	50447
državna uprava	79	59336

*Odrediti prosečan lični dohodak zaposlenih u 2019. godini, kao i mod i medijanu.*

*Rešenje:* Kao i obično, najpre konstruišemo pomoćnu tabelu, u kojoj vršimo sva neophodna preliminarna izračunavanja.

$x_i$	$f_i$	$x_i f_i$	$F_i$	
27149	36	977364	36	
36542	112	4092704	148	
39763	88	3499144	236	
42639	67	2856813	303	
42660	6	255960	309	
42972	14	601608	323	
45083	47	2118901	370	
50447	161	8121967	531	
50634	153	7747002	684	
55055	109	6000995	793	
59336	79	4687544	872	
59849	26	1556074	898	
68319	643	43929117	1541	
99771	61	6086031	1602	
$\Sigma$	1602	92531224		

$n$	1602
mod	68319
med	59336
$\bar{x}$	57759.815

Dakle, mod godišnjeg proseka dohotka građana za 2019. godinu iznosio je 68319 RSD, medijana je bila 59336 RSD, dok je prosečan lični dohodak zaposlenog lica bio

$$\bar{x} = \frac{92531224}{1602} = 57759.815 \text{ RSD.}$$

**Vežba 2.3** Radi ispitivanja broja izostanaka sa nastave u toku jednog dana u jednoj školi sa 700 učenika, izabran je uzorak od 100 učenika i dobijeni su podaci dati u vidu sledeće tabele:

broj izostanaka	0	1	2	3	4	5	6
broj učenika	30	20	7	15	18	6	4

*Odrediti mod, medijanu i prosečan broj izostanaka.*

*Rešenje:* Prateći postupak prezentovan u prethodnom primeru, postepeno nalazimo mere centralne tendencije. Pre svega, konstruisaćemo pomoćnu tabelu za sva neophodna izračunavanja.

$x_i$	$f_i$	$x_i f_i$	$F_i$
0	30	0	30
1	20	20	50
2	7	14	57
3	15	45	72
4	18	72	90
5	6	30	96
6	4	24	100
$\Sigma$	100	205	

$n$	100
mod	0
med	1.5
$\bar{x}$	2.05

Prateći postupak prezentovan u prethodnom primeru, postepeno nalazimo mere centralne tendencije. Dakle, mod iznosi 0, medijana je 1.5, a prosečna vrednost izostanaka je

$$\bar{x} = \frac{205}{100} = 2.05.$$

**Vežba 2.4** Fabrika šećera pakuje šećer u kockama u kutije težine pola kilograma. Da bi se proverila težina kutija, uzima se uzorak od 50 kutija. Pošto je izmerena težina svake kutije, podaci su grupisani u vidu sledeće tabele:

težina u gramima	[496, 498)	[498, 500)	[500, 502)	[502, 504]
broj kutija	8	25	12	5

Odrediti mere centralne tendencije za posmatrani uzorak.


*Rešenje:* Primetimo da su u ovom slučaju vrednosti obeležja predstavljene intervalno, pa ih tako unosimo u pomoćnu tabelu, s tim da kolonu  $x_i$  računamo kao sredine tih intervala. Frekvencije prepisemo. Ostatak procedure ostaje nepromenjen.

Interval	$x_i$	$f_i$	$x_i f_i$	$F_i$
[ 496 , 498 )	497	8	3976	8
[ 498 , 500 )	499	25	12475	33
[ 500 , 502 )	501	12	6012	45
[ 502 , 504 ]	503	5	2515	50
$\Sigma$		50	24978	

$n$	50
mod	499
med	499
$\bar{x}$	499.56

U posmatranom uzorku najviše je kutija sa težinom od 499 g (to je mod ovog obeležja), 50% ukupnog broja kutija je teže a 50% lakše od 499 g (to je medijana ovog obeležja), dok je prosečna težina kutije šećera koju proizvodi posmatrana fabrika

$$\bar{x} = \frac{24978}{50} = 499.56 \text{ g.}$$

 **Vežba 2.5** U tabeli su dati podaci o broju stanara po stanu u jednom bloku zgrada:

broj stanova	5	1	3	6	4	2	7
broj stanara	1	7	5	4	3	6	2

Odrediti mere centralne tendencije za broj stanara po stanu u tom bloku zgrada.


*Rešenje:* Pre svega, primetimo da mogući brojevi stanara u stanu zapravo predstavljaju varijetete, dok su frekvencije date brojem stanova u kojima toliki broj stanara živi.

$x_i$	$f_i$	$x_i f_i$	$F_i$		
1	5	5	5		
2	7	14	12		
3	4	12	16		
4	6	24	22		
5	3	15	25		
6	2	12	27		
7	1	7	28		
$\Sigma$	28	89			

$n$	28
mod	2
med	3
$\bar{x}$	3.1786

Zaključujemo da je, u posmatranom bloku zgrada, najviše stanova u kojima žive dve osobe, dok se u oko 50% ukupnog broja stanova nalazi najviše 3 stanara. Prosečan broj stanara u stanu iznosi 3 (zaokružena vrednost od 3.1786).

 **Vežba 2.6** Nakon registrovanja broja auto-kuća u 60 zemalja, dobijeni podaci su sređeni u vidu tabele:

broj auto kuća	[0, 20)	[20, 40)	[40, 60)	[60, 80]
broj zemalja	18	12	19	11

Izračunati prosečan broj auto-kuća u jednoj zemlji, kao i mod i medijanu tog obeležja.

*Rešenje:* Budući da su intervali dati u pogodnom obliku, prepisujemo ih u pomoćnu tabelu, dok vrednosti broja zemalja predstavljaju frekvencije.

Interval	$x_i$	$f_i$	$x_i f_i$	$F_i$		
[ 0 , 20 )	10	18	180	18		$n$ 60
[ 20 , 40 )	30	12	360	30		mod 50
[ 40 , 60 )	50	19	950	49		med 40
[ 60 , 80 ]	70	11	770	60		$\bar{x}$ 37.6667
$\Sigma$		60	2260			

Dakle, najveći broj zemalja na svojoj teritoriji ima registrovanih 50 auto-kuća, dok 50% ukupnog broja zemalja ima najviše 40 auto-kuća koje posluju u njihovim okvirima. Prosečan broj auto kuća u jednoj zemlji iz posmatranog uzorka je 38.

## 2.2 Mere odstupanja

$$\text{Varijansa (disperzija): } s^2 = \frac{\sum_{i=1}^k (x_i - \bar{x})^2 f_i}{n} = \frac{\sum_{i=1}^k x_i^2 f_i}{n} - \bar{x}^2$$

$$\text{Standardna devijacija } s: \quad s = \sqrt{s^2}$$

$$\text{Koeficijent varijacije: } V = \frac{s}{\bar{x}} \cdot 100\%.$$

**Vežba 2.7** U okviru istraživanja životnog ciklusa jedne vrste sisara ispitivao se broj mladunaca iz jednog legla. Prikupljeni podaci o 800 slučajno izabranih legala svrstani su u tabelu:

broj mladunaca	1	2	3	4	5
broj legala	145	240	194	126	95

Odrediti aritmetičku sredinu posmatranog obeležja i procentualno odstupanje od nje.

*Rešenje:* Konstrukcija pomoćne tabele slična je ranijoj proceduri, uz dodatak nove kolone  $x_i^2 f_i$ , koju koristimo da odredimo mere odstupanja obeležja.

$x_i$	$f_i$	$x_i f_i$	$x_i^2 f_i$		
1	145	145	145	$n$	800
2	240	480	960	$\bar{x}$	2.7325
3	194	582	1746	$s^2$	1.5859
4	126	504	2016	$s$	1.2593
5	95	475	2375	$V$	46.09%
$\Sigma$	800	2186	7242		

U ovom primeru navešćemo detaljno postupak izračunavanja:

$$\begin{aligned} \bar{x} &= \frac{2186}{800} = 2.7325 \\ s^2 &= \frac{7242}{800} - 2.7325^2 = 1.5859 \\ s &= \sqrt{1.5859} = 1.2593 \\ V &= \frac{1.2593}{2.7325} \cdot 100\% = 46.09\% \end{aligned}$$

Dakle, prosečan broj mladunaca u jednom leglu iznosi 3, uz procentualno odstupanje od ove vrednosti u iznosu od 46.09%.



**Vežba 2.8** Prema podacima Zavoda za statistiku, godišnji prosek mesečnih zarada za 2019. godinu, po oblastima delatnosti, iznosio je:

	br. zaposl. u hiljadama	zarade u RSD
industrija i rudarstvo	643	68319
poljoprivreda, šumarstvo i vodoprivreda	67	42639
ribarstvo	6	42660
prerađivačka industrija	14	42972
građevinarstvo	88	39763
saobraćaj i veze	109	55055
trgovina	112	36542
ugostiteljstvo i turizam	36	27149
proizvodnja el. energije, gasa i vode	26	59849
stambeno komunalna delatnost	47	45083
finansijsko posredovanje	61	99771
obrazovanje i kultura	153	50634
zdravstvo i socijalna zaštita	161	50447
državna uprava	79	59336

Odrediti prosečan lični dohodak zaposlenih u 2019. godini, kao i procentualno odstupanje od te vrednosti.

*Rešenje:* Zbog jednostavnosti, podatke o registrovanoj prosečnoj zaradi za svaku od ponuđenih grana privrede iskazaćemo u 10000 RSD (primera radi, prosečnu zaradu od 27149 prikazaćemo kao 2.7149).

$x_i$	$f_i$	$x_i f_i$	$x_i^2 f_i$	
2.7149	36	97.7364	265.3446	
3.6542	112	409.2704	1495.5559	
3.9763	88	349.9144	1391.3646	
4.2639	67	285.6813	1218.1165	
4.266	6	25.596	109.1925	
4.2972	14	60.1608	258.5230	
4.5083	47	211.8901	955.2641	
5.0447	161	812.1967	4097.2887	
5.0634	153	774.7002	3922.6170	
5.5055	109	600.0995	3303.8478	
5.9336	79	468.7544	2781.4011	
5.9849	26	155.6074	931.2947	
6.8319	643	4392.912	30011.9334	
9.9771	61	608.6031	6072.0940	
$\Sigma$	1602	9253.122	56813.8380	
				$n$ 1602
				$\bar{x}$ 5.7760
				$s^2$ 2.1024
				$s$ 1.45
				$V$ 25.10%

Prosečan lični dohodak zaposlenih u 2019. godini iznosio je 57760 RSD, dok je procentualno odstupanje od ove vrednosti iznosilo 25.10%.

**Vežba 2.9** Radi ispitivanja broja izostanaka sa nastave u toku jednog dana u jednoj školi sa 700 učenika, izabran je uzorak od 100 učenika i dobijeni su podaci dati u vidu sledeće tabele:

broj izostanaka	0	1	2	3	4	5	6
broj učenika	30	20	7	15	18	6	4

Odrediti sredinu uzorka, standardnu devijaciju i koeficijent varijacije.

Rešenje:

$x_i$	$f_i$	$x_i f_i$	$x_i^2 f_i$		
0	30	0	0		
1	20	20	20		$n$ 100
2	7	14	28		$\bar{x}$ 2.05
3	15	45	135		$s^2$ 3.4475
4	18	72	288		$s$ 1.8567
5	6	30	150		$V$ 90.57%
6	4	24	144		
$\Sigma$	100	205	765		

Prosečan broj izostanaka u toku posmatranog dana po učeniku iznosi 2, dok je standardna devijacija 1.8567. Koeficijent varijacije je 90.57%.

**Vežba 2.10** Fabrika šećera pakuje šećer u kockama u kutije težine pola kilograma. Da bi se proverilo da li je težina kutija ista, uzima se uzorak od 50 kutija. Pošto je izmerena težina svake kutije, podaci su grupisani u vidu sledeće tabele:


težina u gramima	[496, 498)	[498, 500)	[500, 502)	[502, 504]
broj kutija	8	25	12	5

Izračunati sredinu uzorka, standardnu devijaciju i koeficijent varijacije. Da li fabrika šećera poštuje standard dozvoljenog odstupanja od prosečne težine do 1% ?

Rešenje:

Interval	$x_i$	$f_i$	$x_i f_i$	$x_i^2 f_i$		
[ 496 , 498 )	497	8	3976	1976072		$n$ 50
[ 498 , 500 )	499	25	12475	6225025		$\bar{x}$ 499.56
[ 500 , 502 )	501	12	6012	3012012		$s^2$ 2.8864
[ 502 , 504 ]	503	5	2515	1265045		$s$ 1.6989
$\Sigma$		50	24978	12478154		$V$ 0.34%

Prosečna težina kutija šećera iznosi 499.56 g. Standardna devijacija je 1.6989, dok je prosečno odstupanje od srednje vrednosti 0.34%. Na osnovu toga zaključujemo da fabrika poštuje standard dozvoljenog odstupanja od prosečne težine do 1%.

 **Vežba 2.11** U tabeli su dati podaci o broju stanara po stanu u jednom bloku zgrada:


broj stanova	5	1	3	6	4	2	7
broj stanara	1	7	5	4	3	6	2

Odrediti mere odstupanja za broj stanara po stanu u tom bloku zgrada.

Rešenje:

$x_i$	$f_i$	$x_i f_i$	$x_i^2 f_i$		
1	5	5	5		
2	7	14	28	$n$	28
3	4	12	36	$\bar{x}$	3.1786
4	6	24	96	$s^2$	2.7895
5	3	15	75	$s$	1.6702
6	2	12	72	$V$	52.55%
7	1	7	49		
$\Sigma$	28	89	361		

Varijansa, odnosno srednje kvadratno odstupanje od sredine uzorka iznosi 2.7895, dok je standardno odstupanje 1.6702. Procentualno odstupanje od srednje vrednosti je 52.55%.

 **Vežba 2.12** Nakon registrovanja broja auto-kuća u 60 zemalja, dobijeni podaci su sređeni u vidu tabele:

broj auto-kuća	[0, 20)	[20, 40)	[40, 60)	[60, 80]
broj zemalja	18	12	19	11


Izračunati prosečan broj auto-kuća u jednoj zemlji, kao i procentualno odstupanje od te vrednosti.

Rešenje:

Interval	$x_i$	$f_i$	$x_i f_i$	$x_i^2 f_i$		
[ 0 , 20 )	10	18	180	1800	$n$	60
[ 20 , 40 )	30	12	360	10800	$\bar{x}$	37.6667
[ 40 , 60 )	50	19	950	47500	$s^2$	481.2222
[ 60 , 80 ]	70	11	770	53900	$s$	21.9368
$\Sigma$		60	2260	114000	$V$	58.24%


Na osnovu datog uzorka zaključujemo da je prosečan broj auto-kuća 38, dok je procentualno odstupanje od ove vrednosti 58.24%.

## Vežbe za samostalni rad

 **Vežba 2.13** *Studenti biologije su posejali zrna pasulja u zasebne plastične posude i postavili ih pod izvor svetlosti. Nakon 14 dana, izvršili su merenje visine (u mm) izraslih biljaka, i pritom su registrovali sledeće vrednosti:*


*75, 101, 83, 98, 57, 103, 92, 87, 50, 66, 67, 110, 105, 86, 88, 59, 62, 81, 94, 54, 102, 89, 57, 61, 73, 72, 79, 88, 89, 67, 76, 55, 59, 65, 68, 71.*

*Podatke grupisati u 6 jednakih intervala, a zatim odrediti mere centralne tendencije i mere odstupanja za posmatrani uzorak.*

 **Vežba 2.14** *Radi kontrole kvaliteta vitaminskog napitka pakovanog u bočice od 100 ml izmerena je količina C vitamina u 35 slučajno izabranih bočica. Dobijeni su sledeći rezultati izraženi u mg:*


<i>kol. C</i>	[194, 196)	[196, 198)	[198, 200)	[200, 202)	[202, 204)	[204, 206]
<i>br. boč.</i>	2	6	9	10	4	4

*Za ovako prikupljenje podatke, odrediti mere centralne tendencije, kao i mere odstupanja.*

 **Vežba 2.15** *Zoolog je, izučavajući jedinke žirafa na podneblju afričkog kontinenta, dokumentovao podatke o visini (u cm) novorođenih jedinki pomenute vrste na jednom lokalitetu u proteklih 5 godina:*


<i>visina žirafe</i>	[150, 160)	[160, 170)	[170, 180)	[180, 190]
<i>broj jedinki</i>	7	10	7	14

*Za podatke koje je prikupio zoolog, odrediti mere centralne tendencije i mere odstupanja.*

 **Vežba 2.16** *Istraživač koji se bavi proučavanjem ponašanja miševa sproveo je eksperiment, u kojem je merio vreme (u sekundama) potrebno da bi miš pronašao hranu u lavirintu. Vreme je merio na uzorku od 20 miševa i tim povodom prikupio sledeće podatke:*

*31, 33, 46, 23, 20, 28, 29, 33, 27, 34, 35, 28, 32, 50, 41, 31, 48, 22, 30, 39.*

*Dobijene rezultate najpre grupisati u 5 jednakih intervala, a potom odrediti mere centralne tendencije i mere odstupanja.*

 **Vežba 2.17** Za potrebe analize potrošnje goriva na 100km jedne kategorije putničkih automobila u uslovima vožnje na otvorenom putu, redakcija auto časopisa izvršila je merenja (u litrima), i zabeležila sledeće vrednosti:

6.2, 4, 7.8, 9, 8.9, 5.4, 5.2, 7.1, 8.7, 6.5, 5.9, 4.8, 4.2, 8.3, 8.2, 7.5, 4.4, 5.6, 8.8.  
Podatke grupisati u 5 jednakih intervala, a potom odrediti mere centralne tendencije, kao i mere odstupanja.



## Intervali poverenja za sredinu populacije

### Veličina populacije nije poznata

#### 1. Dati ili prethodno sračunati podaci:

$n$  - veličina uzorka     $\bar{x}$  - sredina uzorka     $1 - \alpha$  - nivo poverenja

#### 2. Računanje korekcije $\varepsilon$ :

<i>slučaj</i>	$\varepsilon$
$\sigma$ poznato	$z_{\frac{1-\alpha}{2}} \frac{\sigma}{\sqrt{n}}$
$\sigma$ nepoznato i $n > 30$	$z_{\frac{1-\alpha}{2}} \frac{s}{\sqrt{n-1}}$
$\sigma$ nepoznato i $n \leq 30$	$t_{n-1, \alpha} \frac{s}{\sqrt{n-1}}$

- $\sigma$  - standardna devijacija cele populacije
- $s$  - standardna devijacija uzorka
- $z_{\frac{1-\alpha}{2}}$  - broj koji se određuje iz  $z$ -tablice
- $t_{n-1, \alpha}$  - broj koji se određuje iz  $t$ -tablice

#### 3. Interval poverenja za sredinu je: $(\bar{x} - \varepsilon, \bar{x} + \varepsilon)$

## Veličina populacije je poznata

- ako je  $\frac{n}{N} < 0.04 = 4\%$  koristi se gore navedeni postupak
- ako je  $\frac{n}{N} \geq 0.04 = 4\%$  dodatno se računa *korektivni faktor*

$$\kappa = \sqrt{\frac{N-n}{N-1}}$$

**Interval poverenja za sredinu:**  $(\bar{x} - \kappa\varepsilon, \bar{x} + \kappa\varepsilon)$

**Vežba 3.1** U okviru istraživanja životnog ciklusa jedne vrste sisara ispitivao se broj mladunaca iz jednog legla. Prikupljeni podaci o 800 slučajno izabranih legala svrstani su u tabelu:

broj mladunaca	1	2	3	4	5
broj legala	145	240	194	126	95

- (a) Odrediti interval poverenja za srednju vrednost na nivou poverenja 95%, ako se ne zna veličina populacije.
- (b) Sa sigurnošću od 95% proceniti ukupan broj mladunaca, ako je posmatrani uzorak uzet na teritoriji regiona na kojoj je procenjeno da postoji dve hiljade legala ove vrste sisara. Da li ste primenjivali korektivni faktor i zašto?

*Rešenje:* Najpre ćemo pristupiti računanju neophodnih statistika na osnovu raspoloživih podataka.

$x_i$	$f_i$	$x_i f_i$	$x_i^2 f_i$
1	145	145	145
2	240	480	960
3	194	582	1746
4	126	504	2016
5	95	475	2375
$\Sigma$	800	2186	7242

Uzorak		Nivo poverenja	
$n$	800	$1 - \alpha$	0.95
$\bar{x}$	2.7325		
$s^2$	1.5859		
$s$	1.2593		

- (a) S obzirom da veličina populacije nije poznata, interval poverenja je oblika  $(\bar{x} - \varepsilon, \bar{x} + \varepsilon)$ .

Računanje korekcije		Interval Poverenja	
$z_{0.475}$	1.96	$\bar{x} - \varepsilon$	2.6452
$\varepsilon$	0.0873	$\bar{x} + \varepsilon$	2.8198

Dakle, ako bismo eksperiment, odnosno prebrojavanje mladunaca u leglima, ponavljali "bezbroj" (dovoljno veliki broj) puta, svaki put uzimajući uzorke obima 800, u 95%



slučajeva bi tačna vrednost sredine populacije, koja je nepoznati parametar, bila neka vrednost iz intervala (2.6452, 2.8198).

(b)

Populacija		$n / N$ 40%			
$N$	2000				
Računanje korekcije		Interval Poverenja		Ukupno	
$z_{0.475}$	1.96	$\bar{x} - \kappa\varepsilon$	2.6648	od	5329
$\varepsilon$	0.0873	$\bar{x} + \kappa\varepsilon$	2.8	do	5600
$\kappa$	0.7748				

U slučaju kada je veličina populacije poznata, najpre treba utvrditi da li se korektivni faktor primenjuje ili ne. Budući da je  $800/2000 = 0.4 = 40\% > 4\%$ , neophodno je uključiti ga u izračunavanje. Interval poverenja za sredinu populacije je (2.6648, 2.8), što znači da, ako bismo eksperiment, odnosno prebrojavanje mladunaca u leglima, ponavljali "bezbroj" (dovoljno veliki broj) puta, svaki put uzimajući uzorke obima 800, u 95% slučajeva bi tačna vrednost sredine populacije, koja je nepoznati parametar, bila neka vrednost iz intervala (2.6648, 2.8). Da bismo procenili ukupan broj mladunaca na teritoriji regiona na kojoj ima 2000 legala, potrebno je granice intervala za prosečnu vrednost legala pomnožiti sa 2000. Dakle, zaključujemo da se ukupan broj mladunaca na posmatranoj teritoriji kreće između  $2.6648 \cdot 2000 = 5329$  i  $2.8 \cdot 2000 = 5600$ .

**Vežba 3.2** Prema podacima Zavoda za statistiku, deo uvoza Republike Srbije od januara do oktobra 2016. godine, na slučajnom uzorku, izgledao je:

vrednost u hiljadama \$	broj uvoznika
[0, 30)	100
[30, 60)	400
[60, 90)	150
[90, 120]	350

Sa sigurnošću od 99%, proceniti ukupnu vrednost uvoza u ovom periodu, ako se zna da je broj aktivnih uvoznika bio dvadeset hiljada.

Rešenje:

Interval	$x_i$	$f_i$	$x_i f_i$	$x_i^2 f_i$	Uzorak	
[ 0 , 30 )	15	100	1500	22500	$n$	1000
[ 30 , 60 )	45	400	18000	810000	$\bar{x}$	67.5
[ 60 , 90 )	75	150	11250	843750	$s^2$	978.75
[ 90 , 120 ]	105	350	36750	3858750	$s$	31.285
$\Sigma$		1000	67500	5535000		

Populacija		Nivo poverenja		$n / N$ 5%		
$N$	20000	$1 - \alpha$	0.99			
Računanje korekcije			Interval Poverenja		Ukupno	
$z_{0.495}$	2.58	$\bar{x} - \kappa\varepsilon$	65.0109	od	1300217.61	
$\varepsilon$	2.5537	$\bar{x} + \kappa\varepsilon$	69.9891	do	1399782.39	
$\kappa$	0.9747					

Ako bismo eksperiment, odnosno registrovanje vrednosti dela uvoza Republike Srbije u posmatranom periodu, ponavljali "bezbroj" (dovoljno veliki broj) puta, svaki put uzimajući uzorke obima 1000, sa sigurnošću od 99% odnosno mogućnošću greške do 1%, bi tačna vrednost sredine populacije, to jest prosečna vrednost uvoza aktivnog uvoznika, koja je nepoznati parametar, bila neka vrednost iz intervala (65010.9, 69989.1). Ukupna vrednost uvoza svih aktivnih uvoznika procenjuje se između 1300217610 i 1399782390 dolara.

**Vežba 3.3** Na jednoj studijskoj grupi od 2000 studenata 756 prima stipendije. Od 30 slučajno izabranih studenata, 19 studenata je u toku semestra primilo stipendiju (u hiljadama dinara) i to prema sledećoj raspodeli:

stipendija	broj studenata
[2, 4)	5
[4, 6)	5
[6, 8)	4
[8, 10)	3
[10, 12]	2

Odrediti interval poverenja za prosečnu stipendiju na posmatranoj studijskoj grupi, na nivou poverenja 0.9. U kojim granicama će se kretati ukupna vrednost isplaćenih stipendija ovoj studijskoj grupi sa istim nivoom poverenja?

Rešenje:

Interval	$x_i$	$f_i$	$x_i f_i$	$x_i^2 f_i$	Uzorak	
[ 2 , 4 )	3	5	15	45	$n$	19
[ 4 , 6 )	5	5	25	125	$\bar{x}$	6.1579
[ 6 , 8 )	7	4	28	196	$s^2$	6.8698
[ 8 , 10 )	9	3	27	243	$s$	2.6210
[ 10 , 12 ]	11	2	22	242		
$\Sigma$		19	117	851		

Populacija	Nivo poverenja	$n / N$ 3%
$N$	$1 - \alpha$	
756	0.9	

Računanje korekcije		Interval Poverenja		Ukupno	
$t_{18,0.1}$	1.7341	$\bar{x} - \varepsilon$	5.0866	od	3845.48
$\varepsilon$	1.0713	$\bar{x} + \varepsilon$	7.2292	do	5465.26

Na osnovu ovog uzorka, zaključujemo, sa sigurnošću od 90%, odnosno mogućnošću greške do 10%, da se tačna vrednost sredine populacije, to jest visina prosečne stipendije nalazi u intervalu (5086.6, 7229.2). Ukupna vrednost svih isplata za stipendije se procenjuje između 3845480 i 5465260 dinara.

**Vežba 3.4** Fabrika prodaje svoj proizvod preko mreže od 1000 prodavnica. Radi kontrole prodaje, izabrane su 64 prodavnice i registrovana je mesečna prodaja (izražena u 10000 RSD). Podaci su prikazani u tabeli:

mesečna prodaja	2	3	4	5	6	7
broj prodavnica	5	15	24	11	6	3

Poznato je da raspodela broja prodatih proizvoda u pomenutoj mreži prodavnica ima srednje kvadratno odstupanje 0.9. Odrediti interval poverenja za prosečan broj prodatih proizvoda, na nivou poverenja 0.95. U kojim granicama će se kretati ukupan broj prodatih proizvoda, na istom nivou poverenja?

*Rešenje:* Primetimo da je, za razliku od prethodnih primera, varijansa populacije poznata veličina. Zbog toga, interval poverenja za sredinu populacije nalazimo ovako:

$x_i$	$f_i$	$x_i f_i$	$x_i^2 f_i$												
2	5	10	20	<table border="1"> <thead> <tr> <th colspan="2">Uzorak</th> </tr> </thead> <tbody> <tr> <td><math>n</math></td> <td>64</td> </tr> <tr> <td><math>\bar{x}</math></td> <td>4.1094</td> </tr> <tr> <td><math>s^2</math></td> <td>1.5037</td> </tr> <tr> <td><math>s</math></td> <td>1.2262</td> </tr> </tbody> </table>		Uzorak		$n$	64	$\bar{x}$	4.1094	$s^2$	1.5037	$s$	1.2262
Uzorak															
$n$	64														
$\bar{x}$	4.1094														
$s^2$	1.5037														
$s$	1.2262														
3	15	45	135												
4	24	96	384												
5	11	55	275												
6	6	36	216												
7	3	21	147												
$\Sigma$	64	263	1177												

Populacija		Nivo poverenja		$n / N$	
$N$	1000	$1 - \alpha$	0.95	6.4%	
$\sigma^2$	0.90				
$\sigma$	0.9487				

Računanje korekcije		Interval Poverenja		Ukupno	
$z_{0.475}$	1.96	$\bar{x} - \kappa\varepsilon$	3.8844	od	3884.4
$\varepsilon$	0.2324	$\bar{x} + \kappa\varepsilon$	4.3344	do	4334.4
$\kappa$	0.9680				

Sa sigurnošću od 95%, prosečna vrednost prodatih proizvoda kreće se u intervalu (38844, 43344), a ukupna vrednost prodatih proizvoda svih prodajnih objekata je neka vrednost između 38844000 i 43344000 dinara.

**Vežba 3.5** *Analitičar kompanije SkyLine Inc. dobio je izveštaj o slučajevima izgubljenog prtljaga u poslednjih deset godina. U tabeli su dati podaci o mesečnom broju izgubljenih prtljaga za 20 slučajno izabranih meseci.*

broj izgubljenih prtljaga	5	4	3	2	1
broj meseci	1	3	5	6	5

- (a) U kojim se granicama, na nivou poverenja 90%, kreće ukupan broj izgubljenih prtljaga u proteklih deset godina?
- (b) U kojim se granicama, na nivou poverenja 95%, kreće ukupan broj izgubljenih prtljaga u proteklih deset godina?

Rešenje:

$x_i$	$f_i$	$x_i f_i$	$x_i^2 f_i$	Uzorak		Populacija	
1	5	5	5	$n$	20	$N$	120
2	6	12	24	$\bar{x}$	2.45	$n / N$	17%
3	5	15	45	$s^2$	1.3475		
4	3	12	48	$s$	1.1608		
5	1	5	25				
$\Sigma$	20	49	147				

(a)

Nivo poverenja		Računanje korekcije		Interval Poverenja		Ukupno	
$1 - \alpha$	0.9	$t_{19,0.1}$	1.729	$\bar{x} - \kappa\varepsilon$	2.0279	od	243
		$\varepsilon$	0.4604	$\bar{x} + \kappa\varepsilon$	2.8721	do	345
		$\kappa$	0.9167				

Na nivou poverenja od 90%, ukupan broj izgubljenih prtljaga u proteklih 10 godina kreće se između 243 i 345.

(b)

Nivo poverenja		Računanje korekcije		Interval Poverenja		Ukupno	
$1 - \alpha$	0.95	$t_{19,0.05}$	2.093	$\bar{x} - \kappa\varepsilon$	1.9391	od	233
		$\varepsilon$	0.5574	$\bar{x} + \kappa\varepsilon$	2.9609	do	355
		$\kappa$	0.9167				

Na nivou poverenja od 95%, zaključujemo da se ukupan broj izgubljenih prtljaga u proteklih 10 godina kreće između 233 i 355.

**Vežba 3.6** Ispituje se promet nekretnina u prvom kvartalu 2017. godine u agenciji koja zapošljava 200 agenata prodaje. U tabeli su dati podaci o broju prodatih stanova preko te agencije za slučajno izabrane dane u tom periodu:

broj prodatih stanova	5	7	6	4	3	2	1
broj dana	1	2	4	3	5	6	7

- (a) U kojim se granicama, na nivou poverenja 0.95, kreće prosečna dnevna prodaja stanova preko te agencije u prvom kvartalu 2017?
- (b) U kojim se granicama, na nivou poverenja 0.99, kreće ukupan broj prodatih stanova preko te agencije u prvom kvartalu 2017?

Rešenje: Primitimo da je u ovom zadatku obeležje broj prodatih stanova dnevno.

$x_i$	$f_i$	$x_i f_i$	$x_i^2 f_i$				
1	7	7	7				
2	6	12	24	Uzorak			
3	5	15	45	$n$	28	Populacija	
4	3	12	48	$\bar{x}$	3.1786	$N$	90
5	1	5	25	$s^2$	3.8610	$n / N$	31%
6	4	24	144	$s$	1.9649		
7	2	14	98				
$\Sigma$	28	89	391				

(a)

Nivo poverenja		Računanje korekcije		Interval Poverenja		Ukupno	
$1 - \alpha$	0.95	$t_{27,0.05}$	2.0518	$\bar{x} - \kappa \varepsilon$	2.5310	od	227
		$\varepsilon$	0.7759	$\bar{x} + \kappa \varepsilon$	3.8262	do	345
		$\kappa$	0.8346				

Na nivou poverenja od 95%, prosečna dnevna prodaja nekretnina u prvom kvartalu 2017. godine posredstvom dotične agencije kretala se u intervalu (2.5310, 3.8262).

(b)

Nivo poverenja		Računanje korekcije		Interval Poverenja		Ukupno	
$1 - \alpha$	0.99	$t_{27,0.01}$	2.7707	$\bar{x} - \kappa \varepsilon$	2.3041	od	207
		$\varepsilon$	1.0477	$\bar{x} + \kappa \varepsilon$	4.0531	do	365
		$\kappa$	0.8346				

Ukupan broj prodatih nekretnina posredstvom agencije u prvom kvartalu 2017. godine na nivou poverenja od 99%, se kreće od 207 do 365.

**Vežba 3.7** Rezultati ispitivanja broja sunčanih dana u februaru 2017. godine na uzorku od 25 gradova jednog regiona su:

broj sunčanih dana	[2, 8)	[8, 12)	[12, 18)	[18, 22)	[22, 28]
broj gradova	2	7	10	4	2

U kojim se granicama (uz verovatnoću od 0.95) kretao prosečan broj sunčanih dana u celom regionu? Ako pretpostavimo da je klima takva da su svi meseci slični februaru, koliko se očekuje sunčanih dana u toku 2017. godine u ovom regionu?

Rešenje:

Interval	$x_i$	$f_i$	$x_i f_i$	$x_i^2 f_i$	Uzorak	
[ 2 , 8 )	5	2	10	50	$n$	25
[ 8 , 12 )	10	7	70	700	$\bar{x}$	14.4000
[ 12 , 18 )	15	10	150	2250	$s^2$	26.6400
[ 18 , 22 )	20	4	80	1600	$s$	5.1614
[ 22 , 28 ]	25	2	50	1250		
$\Sigma$		25	360	5850		

Nivo poverenja		Računanje korekcije		Interval Poverenja		Godišnje	
$1 - \alpha$	0.95	$t_{24,0.05}$	2.0639	$\bar{x} - \varepsilon$	12.2255	od	146
		$\varepsilon$	2.1745	$\bar{x} + \varepsilon$	16.5745	do	199

Na nivou poverenja od 95% zaključujemo da se prosečan broj sunčanih dana u celom regionu kretao u intervalu (12.2255, 16.5745). Uz pretpostavku da su, u posmatranom regionu, svi meseci slični februaru, dolazimo do prognoze da se na godišnjem nivou očekuje između  $12.2255 \cdot 12 = 146$  i  $16.5745 \cdot 12 = 199$  sunčanih dana.

**Vežba 3.8** Rezultati ispitivanja mesečne potrošnje hleba u kilogramima na uzorku od 200 domaćinstava iz jednog grada, u kojem je registrovano 5000 domaćinstava, su dati u sledećoj tabeli:

potrošnja hleba	2	3	4	5	6
broj gradova	10	30	60	70	30

U kojim se granicama, sa sigurnošću od 90%, očekuje potrošnja hleba u tom gradu?


Rešenje: Najpre ćemo konstruisati pomoćnu tabelu a zatim odrediti i uzoračke statistike.

$x_i$	$f_i$	$x_i f_i$	$x_i^2 f_i$	Uzorak		Populacija	
2	10	20	40	$n$	200	$N$	5000
3	30	90	270	$\bar{x}$	4.4	$n / N$	4.00%
4	60	240	960	$s^2$	1.14		
5	70	350	1750	$s$	1.0677		
6	30	180	1080				
$\Sigma$	200	880	4100				

Nivo poverenja		Računanje korekcije		Interval Poverenja		Ukupno	
		$z_{0.45}$	1.64	$\bar{x} - \kappa\varepsilon$	4.2780	od	21390
$1 - \alpha$	0.90	$\varepsilon$	0.1245	$\bar{x} + \kappa\varepsilon$	4.5220	do	22610
		$\kappa$	0.9799				


Pod pretpostavkom da eksperiment, odnosno registrovanje mesečne potrošnje hleba u kilogramima, ponavljamo dovoljno veliki broj puta, svaki put uzorkujući podatke na 200 domaćinstava u gradu sa 5000 domaćinstava, zaključujemo da se očekivana ukupna potrošnja hleba kreće između 21390 i 22610 kg.

## Vežbe za samostalni rad

 **Vežba 3.9** U jednom prigradskom naselju na slučajnan način izabrano je nekoliko domaćinstava i popisano je koliko oni dnevno troše mleka. Podaci su sledeći:


broj domaćinstava	5	15	20	11	9
potrošnja mleka (u l)	6	3	2	1	4

U kojim se granicama, na nivou poverenja 0.95, može očekivati prosečna dnevna potrošnja mleka jednog domaćinstva? Ako u tom selu ima 900 domaćinstava, u kojim se granicama kreće ukupna prosečna dnevna potrošnja mleka, na nivou poverenja 0.99?

 **Vežba 3.10** U jednom selu na slučajnan način izabrano je nekoliko pušača i popisano je koliko cigareta dnevno koriste. Podaci su sledeći:

broj cigareta	[1, 3)	[3, 5)	[5, 8)	[8, 13)	[13, 17]
broj pušača	2	7	5	8	5

U kojim se granicama, na nivou poverenja 0.9, može očekivati prosečan broj cigareta koje konzumira pušač u tom selu? Ako selo ima 600 stanovnika, u kojim se granicama kreće ukupna prosečna dnevna potrošnja cigareta, na nivou poverenja 0.95?

 **Vežba 3.11** U tabeli su dati podaci o broju dece koja su se rodila u prethodnih 5 godina, u nekoliko slučajno odabranih porodica. Podaci su sledeći:

broj porodica	17	10	5	2	1
broj dece	2	3	1	4	5

U kojim se granicama, na nivou poverenja 0.95, može očekivati prosečan broj novorođene dece u porodici u poslednjih 5 godina? Ako imamo podatak da u tom kvartu živi 500 porodica, u kojim se granicama kreće ukupan broj dece rođene u poslednjih 5 godina u tom kvartu, na nivou poverenja 0.9?

**Vežba 3.12** Na slučajan način u jednom gradu odabrana je grupa tinejdžera i postavljeno im je pitanje koliko novca troše na zabavu nedeljno. Dobijeni podaci su dati u sledećoj tabeli:

izdaci (u stotinama)	[0, 10)	[10, 15)	[15, 25)	[25, 30)	[30, 50]
broj tinejdžera	5	10	10	20	5

U kojim se granicama, na nivou poverenja 0.95, može očekivati prosečna količina novca koju tinejdžer potroši na zabavu tokom nedelju dana?

**Vežba 3.13** U jednom selu ispitivan je broj vozila po domaćinstvu. Podaci su dati u sledećoj tabeli:

broj vozila	[0, 2)	[2, 4)	[4, 5)	[5, 6)	[6, 7]
broj domaćinstava	5	17	10	2	1

U kojim se granicama, na nivou poverenja 0.95, može očekivati prosečan broj vozila u jednom domaćinstvu? Ako imamo podatak da u tom selu ima 500 domaćinstava, u kojim se granicama nalazi ukupan broj vozila u tom selu, na nivou poverenja 0.99?

**Vežba 3.14** Na jednoj farmi, na slučajan način, izabrano je 100 krava i popisana je dnevna mlečnost u litrama. Podaci su sledeći:

dnevna mlečnost u litrama	[9, 11)	[11, 13)	[13, 15)	[15, 18)	[18, 20]
broj krava	25	40	15	12	8


U kojim se granicama, na nivou poverenja 0.9, može očekivati prosečna mlečnost krava na farmi? Ako imamo podatak da se na toj farmi nalazi 1000 krava, u kojim se granicama kreće ukupna dnevna količina mleka na farmi, na nivou poverenja 0.99?

**Vežba 3.15** Na jednoj studijskoj grupi, na slučajan način, izabrano je 100 studenata i zabeleženo je koliko oni nedeljno troše kredita na mobilnom telefonu. Podaci su sledeći:

ned. potrošnja kredita	[0, 100)	[100, 160)	[160, 260)	[260, 400)	[400, 800]
broj studenata	25	40	15	12	8


U kojim se granicama, na nivou poverenja 0.95, može očekivati prosečna potrošnja kredita za mobilni telefon jednog studenta tokom nedelju dana?




 **Vežba 3.16** U tabeli su dati podaci o broju buketa koje su muškarci iz jednog grada kupili u jednoj cvečari 8. marta. Podaci su sledeći:


broj kupljenih buketa	11	7	5	1	3
broj muškaraca	2	2	5	8	5


U kojim se granicama, na nivou poverenja 0.95, može očekivati prosečan broj buketa koje je muškarac kupio u toj cvečari 8. marta? Ako imamo podatak da je u toj cvečari 1200 muškaraca kupilo cveće 8. marta, u kojim se granicama kreće ukupan broj prodatih buketa tog dana, na nivou poverenja 0.99?

 **Vežba 3.17** Kontroliše se prosečan vek trajanja sijalice kompanije SABA Electric Co. Nakon merenja koje je izvršeno na uzorku obima 100, dobijena je sredina veka trajanja od 930 radnih sati. Izračunati 99% interval poverenja prosečnog veka trajanja sijalice, ako se zna da je koeficijent varijacije 2%.

 **Vežba 3.18** Iz populacije je odabran prost slučajni uzorak sa sredinom 48.9 i standardnom devijacijom 6.7.

- (a) Odrediti 95% interval poverenja ako je obim uzorka 30.
- (b) Odrediti 95% interval poverenja ako je obim uzorka 50.
- (c) Kako povećanje obima uzorka utiče na interval poverenja?

 **Vežba 3.19** Na uzorku od 17 AA baterija u igračkama posmatran je i zabeležen životni vek baterija. Uzoračka sredina trajanja baterije je 3.58 sati, dok je uzoračka standardna devijacija 1.85. Odrediti 99% interval poverenja očekivanog radnog veka baterije.

 **Vežba 3.20** Kompanija Safe Lumber je, za potrebe kontrole i merenja nivoa arsena u podzemnim vodama u okrugu od 1 milje, angažovala nezavisnog kontrolora. Kontrolor, u sklopu svog angažmana, uzorkuje podzemnu vodu na određenim mestima u pomenutom okrugu oko središta kompanije. Odrediti 95% interval poverenja za očekivani nivo arsena u podzemnim vodama, ako se zna da je, na osnovu uzorka od 25 lokaliteta, određen srednji nivo ove toksične supstance od 75 ppb, uz uzoračku standardnu devijaciju od 30 ppb.



## t-test za sredinu populacije

**Cilj:** Uporediti sredinu uzorka i sredinu cele populacije

**Uslov:** Populacija mora imati neprekidnu raspodelu, ako je obim uzorka  $n < 30$

**Statistika:**

$$t = \frac{(\bar{x} - \mu_0)\sqrt{n}}{s}$$

- $n$  - obim uzorka
- $\bar{x}$  - sredina uzorka
- $s$  - standardna devijacija uzorka

### Dvostrani test

**Nulta hipoteza**  $H_0 : \mu = \mu_0$

**Alternativna hipoteza**

$H_A : \mu$  se značajno razlikuje od  $\mu_0$

**Kritična vrednost iz tablice:**

$t_{n-1,\alpha}$

**Zaključak:** Ako je  $|t| > t_{n-1,\alpha}$ , tada se  $H_0$  odbacuje, u korist  $H_A$ .

### Jednostrani test

**Nulta hipoteza**  $H_0 : \mu = \mu_0$

**Alternativna hipoteza**

$H_A : \mu$  je značajno veće od  $\mu_0$   
( $\mu$  je značajno manje od  $\mu_0$ )

**Kritična vrednost iz tablice:**

$t_{n-1,2\alpha}$

**Zaključak:** Ako je  $t > t_{n-1,2\alpha}$   
( $t < -t_{n-1,2\alpha}$ ), tada se  $H_0$  odbacuje u korist  $H_A$ .

**Vežba 4.1** Usled reorganizacije aerodroma Heathrow, analitičar iz službe za kvalitet usluga treba da dostavi izveštaj, na osnovu kojeg će se donositi odluke. Analitičar je prikupio informacije o tačnom vremenu sletanja u odnosu na planirano, za 170 slučajno izabranih letova sa ovog aerodroma u prethodnom kvartalu. Prikupljeni podaci su dati u tabeli:

kašnjenje u minutama	0	10	20	30	40
broj letova	30	60	50	20	10

Upravni odbor kompanije BAA, koja je vlasnik aerodroma, usaglasio se da, ukoliko u prethodnom kvartalu u proseku nema kašnjenja po letu, sa mogućnošću greške do 10%, kompanija neće menjati red letenja. Koji su rezultati izveštaja i kakve su odluke sprovedene?

Rešenje:

$x_i$	$f_i$	$x_i f_i$	$x_i^2 f_i$	Uzorak	
0	30	0	0	$n$	170
10	60	600	6000	$\bar{x}$	15.294
20	50	1000	20000	$s^2$	119.03
30	20	600	18000	$s$	10.910
40	10	400	16000		
$\Sigma$	170	2600	60000		

Prosečno vreme kašnjenja na ovom uzorku je 15.294 minuta. Pitanje je, međutim, da li sa mogućnošću greške do 10% (tj. sa nivoom značajnosti 0.1) možemo odbaciti hipotezu da nema kašnjenja po letu. Dakle, kao nultu hipotezu formulišemo:

$H_0$  : kašnjenje po letu je nula minuta,

a kao alternativnu:

$H_A$  : kašnjenje po letu se značajno razlikuje od nula minuta.

Test statistiku  $t$  računamo na sledeći način:

$$t = \frac{(\bar{x} - \mu_0)\sqrt{n}}{s} = \frac{(15.294 - 0)\sqrt{170}}{10.91} = 18.2776.$$

Zatim je poredimo sa kritičnom vrednošću iz tablice. Skraćeno, to ćemo, ubuduće, zapisivati ovako:

$H_0$ : $\mu = 0$	Sredina populacije		Nivo značajnosti	
$H_A$ : $\mu \neq 0$	$\mu_0$	0	$\alpha$	0.1
	Test statistika		Kritična vrednost	
	$t$	18.2776	$t_{169,0.1}$	1.6539

**Zaključak:** Budući da je  $|t| = 18.2776 > 1.6539 = t_{169,0.1}$ , to se  $H_0$  odbacuje u korist alternativne hipoteze na osnovu datog uzorka. To znači da je u prethodnom kvartalu vreme kašnjenja značajno različito od nula minuta, te zaključujemo da će kompanija odlučiti da menja red letenja.

☞ **Vežba 4.2** Radi kontrole kvaliteta vitaminskog napitka merena je količina  $C$  vitamina u 100 g napitka. Rezultati su prikazani u sledećoj tabeli.

$C$ vitamin u mg	[194, 198)	[198, 200)	[200, 202)	[202, 204)	[204, 208]
broj merenja	2	5	7	4	2

Da li, sa sigurnošću od 90%, možemo tvrditi da je prosečna količina prisutnog  $C$  vitamina u 100 g napitka 200 mg? Da li postoje dokazi da sadržaj vitamina  $C$  u napitku statistički značajno odstupa od predviđene vrednosti od 200 mg po 100 g napitka?

Rešenje:

Interval	$x_i$	$f_i$	$x_i f_i$	$x_i^2 f_i$	Uzorak	
[ 194 , 198 )	196	2	392	76832	$n$	20
[ 198 , 200 )	199	5	995	198005	$\bar{x}$	200.90
[ 200 , 202 )	201	7	1407	282807	$s^2$	6.79
[ 202 , 204 )	203	4	812	164836	$s$	2.6058
[ 204 , 208 ]	206	2	412	84872		
$\Sigma$		20	4018	807352		

$H_0 : \mu = 200$	Sredina populacije	Nivo značajnosti
$H_A : \mu \neq 200$	$\mu_0$ 200	$\alpha$ 0.1

Test statistika	Kritična vrednost
$t$ 1.5446	$t_{19,0.1}$ 1.7291

**Zaključak:** Budući da je  $|t| = 1.5446 < 1.7291 = t_{19,0.1}$ , nemamo dovoljno dokaza protiv  $H_0$  na osnovu datog uzorka, što znači da sa sigurnošću od 90%, odnosno mogućnošću greške od 10%, možemo tvrditi da je u 100 g prosečnog napitka prisutno 200 mg  $C$  vitamina. Drugim rečima, nema dovoljno dokaza da sadržaj  $C$  vitamina u napitku statistički značajno odstupa od predviđene vrednosti od 200 mg u 100 g napitka.

**Vežba 4.3** U tabeli su dati podaci o broju stanara u nekoliko slučajno izabranih stanova u jednom bloku zgrada:

broj stanara	1	7	5	4	3	6	2
broj stanova	5	1	3	6	4	2	7

Da li, na nivou značajnosti 0.01, možemo tvrditi da u posmatranom bloku zgrada u proseku živi 4 stanara po stanu? Šta zaključujemo preko P-vrednosti?

Rešenje:

$x_i$	$f_i$	$x_i f_i$	$x_i^2 f_i$												
1	5	5	5	<table border="1"> <thead> <tr> <th colspan="2">Uzorak</th> </tr> </thead> <tbody> <tr> <td><math>n</math></td> <td>28</td> </tr> <tr> <td><math>\bar{x}</math></td> <td>3.1786</td> </tr> <tr> <td><math>s^2</math></td> <td>2.7895</td> </tr> <tr> <td><math>s</math></td> <td>1.6702</td> </tr> </tbody> </table>		Uzorak		$n$	28	$\bar{x}$	3.1786	$s^2$	2.7895	$s$	1.6702
Uzorak															
$n$	28														
$\bar{x}$	3.1786														
$s^2$	2.7895														
$s$	1.6702														
2	7	14	28												
3	4	12	36												
4	6	24	96												
5	3	15	75												
6	2	12	72												
7	1	7	49												
$\Sigma$	28	89	361												

$H_0 : \mu = 4$	Sredina populacije	Nivo značajnosti
$H_A : \mu \neq 4$	$\mu_0$ 4	$\alpha$ 0.01

Test statistika	Kritična vrednost
$t$ -2.6025	$t_{27,0.01}$ 2.7707

**Zaključak:** Budući da je  $|t| = 2.6025 < 2.7707 = t_{27,0.01}$ , nemamo dovoljno dokaza protiv  $H_0$  na osnovu datog uzorka, što znači da sa sigurnošću od 99%, odnosno mogućnošću greške do 1%, možemo tvrditi da je očekivano da u posmatranom bloku zgrada u proseku živi 4 stanara po stanu. Zaključak preko P-vrednosti je istovetan, budući da je  $P = 0.0148 > 0.01$  (videti tabelu sa interpretacijom na kraju poglavlja).

**Vežba 4.4** Fabrika šećera pakuje šećer u kockama u kutije težine pola kilograma. Da bi se proverilo da li je težina kutija ista, uzima se uzorak od 21 kutije. Pošto je izmerena težina svake kutije, podaci su grupisani u vidu sledeće table:

težina u gramima	[496, 498)	[498, 500)	[500, 504)	[504, 506]
broj kutija	3	8	5	5

Da li, na osnovu ovog uzorka, možemo tvrditi da fabrika šećera poštuje standard da je, sa mogućnošću greške do 1%, prosečna proizvedena kutija šećera težine pola kilograma?

Rešenje:

Interval	$x_i$	$f_i$	$x_i f_i$	$x_i^2 f_i$	Uzorak	
[ 496 , 498 )	497	3	1491	741027	$n$	21
[ 498 , 500 )	499	8	3992	1992008	$\bar{x}$	500.8571
[ 500 , 504 )	502	5	2510	1260020	$s^2$	7.8367
[ 504 , 506 ]	505	5	2525	1275125	$s$	2.7994
$\Sigma$		21	10518	5268180		

$H_0 : \mu = 500$	Sredina populacije	Nivo značajnosti	
$H_A : \mu \neq 500$	$\mu_0$ 500	$\alpha$	0.01

Test statistika		Kritična vrednost	
$t$	1.4303	$t_{20,0.01}$	2.8453

**Zaključak:** Budući da je  $|t| = 1.4303 < 2.8453 = t_{20,0.01}$ , nemamo dovoljno dokaza protiv  $H_0$  na osnovu datog uzorka, što znači da sa sigurnošću od 99%, odnosno mogućnošću greške do 1%, možemo tvrditi da fabrika šećera poštuje pomenuti standard.

**Vežba 4.5** U okviru ankete o održavanju nastave za nekoliko slučajno izabranih studenata, zabeležen je broj profesora kod kojih odlaze na konsultacije:

broj studenata	1	2	3	4	5	6
broj profesora	1	4	6	5	2	3

Da li, na osnovu ovog uzorka, možemo tvrditi, sa mogućnošću greške do 8%, da student prosečno posećuje konsultacije kod 3 profesora?

Rešenje:

$x_i$	$f_i$	$x_i f_i$	$x_i^2 f_i$	Uzorak	
1	1	1	1	$n$	21
2	5	10	20	$\bar{x}$	3.5714
3	6	18	54	$s^2$	2.2449
4	2	8	32	$s$	1.4983
5	4	20	100		
6	3	18	108		
$\Sigma$	21	75	315		

$H_0 : \mu = 3$	Sredina populacije	Nivo značajnosti	
$H_A : \mu \neq 3$	$\mu_0$ 3	$\alpha$	0.08

Test statistika		Kritična vrednost	
$t$	1.7477	$t_{20,0.08}$	1.8443

**Zaključak:** Budući da je  $|t| = 1.7477 < 1.8443 = t_{20,0.08}$ , nemamo dovoljno dokaza protiv  $H_0$  na osnovu datog uzorka, što znači da sa sigurnošću od 92%, odnosno mogućnošću greške do 8%, možemo tvrditi da student, na osnovu posmatranog uzorka, u proseku posećuje konsultacije kod 3 profesora.

**Vežba 4.6** U okviru ankete o održavanju nastave za nekoliko slučajno izabranih profesora zabeležen je broj studenata koji su dolazili na konsultacije u toku nastave:

broj studenata	1	2	3	4	5	6
broj profesora	1	4	6	5	2	3

- a) Da li, na osnovu ovog uzorka, možemo tvrditi, sa mogućnošću greške do 5%, da konsultacije profesora u proseku posećuje 3 studenta?
- b) Da li, na osnovu ovog uzorka, možemo tvrditi, sa mogućnošću greške do 5%, da konsultacije profesora u proseku posećuje više od 3 studenta?

Rešenje:

$x_i$	$f_i$	$x_i f_i$	$x_i^2 f_i$	Uzorak	
1	1	1	1	$n$	21
2	4	8	16	$\bar{x}$	3.5714
3	6	18	54	$s^2$	1.9592
4	5	20	80	$s$	1.3997
5	2	10	50		
6	3	18	108		
$\Sigma$	21	75	309		

$H_0 : \mu = 3$	Sredina populacije	Nivo značajnosti
$H_A : \mu \neq 3$	$\mu_0$ 3	$\alpha$ 0.05

Test statistika		Kritična vrednost	
$t$	1.8708	$t_{20,0.05}$	2.086

**Zaključak:** Budući da je  $|t| = 1.8708 < 2.086 = t_{20,0.05}$ , nemamo dovoljno dokaza da odbacimo  $H_0$ , što znači da sa sigurnošću od 95%, možemo tvrditi da broj studenata koji posećuju konsultacije u proseku iznosi 3.



- b) Da bismo odgovorili na ovo pitanje, korišćićemo jednostrani test. Pomoćna tabela ostaje ista, a ostatak izračunavanja sada izgleda ovako:

$H_0 : \mu = 3$	Sredina populacije		Nivo značajnosti	
$H_A : \mu > 3$	$\mu_0$	3	$\alpha$	0.05

Test statistika		Kritična vrednost	
$t$	1.8708	$t_{20,0.1}$	1.725

**Zaključak:** Budući da je  $|t| = 1.8708 > 1.725 = t_{20,0.1}$ , to se  $H_0$  odbacuje u korist  $H_A$ , što znači da sa sigurnošću od 95%, možemo tvrditi da konsultacije profesora u proseku posećuje više od 3 studenta.

**Vežba 4.7** Usled reorganizacije aerodroma Heathrow, analitičar iz službe za kvalitet usluga treba da dostavi izveštaj na osnovu koga će se donositi odluke. Analitičar je prikupio informacije o broju izgubljenih prtljaga na nekoliko slučajno izabranih letova sa ovog aerodroma u prethodnom kvartalu. Prikupljeni podaci su dati u tabeli:

broj izgubljenih prtljaga	0	1	2	3	4
broj letova	3	6	5	2	1

Upravni odbor kompanije BAA, koja je vlasnik aerodroma, usaglasio se da ukoliko u prethodnom kvartalu u proseku nema izgubljenih prtljaga po letu, sa mogućnošću greške do 5%, kompanija neće uvoditi restriktivne mere zaposlenima aerodroma. Koji su rezultati izveštaja i kakve su odluke sprovedene?

Rešenje:

$x_i$	$f_i$	$x_i f_i$	$x_i^2 f_i$	Uzorak	
0	3	0	0	$n$	17
1	6	6	6	$\bar{x}$	1.5294
2	5	10	20	$s^2$	1.1903
3	2	6	18	$s$	1.0910
4	1	4	16		
$\Sigma$	17	26	60		

$H_0 : \mu = 0$	Sredina populacije		Nivo značajnosti	
$H_A : \mu > 0$	$\mu_0$	0	$\alpha$	0.05
Test statistika		Kritična vrednost		P vrednost
$t$	5.7799	$t_{16,0.1}$	1.7459	$P$ 0.0000141

**Zaključak:** Budući da je  $t = 5.7799 > 1.7459 = t_{16,0.1}$ , to se  $H_0$  odbacuje u korist alternativne hipoteze na osnovu datog uzorka, što znači da je u prethodnom kvartalu broj izgubljenih prtljaga statistički značajno veći od 0, te zaključujemo da će kompanija uvesti restriktivne mere prema zaposlenima.

**Napomena:** Čitalac će se vrlo često u praksi susresti sa pristupom statističkom zaključivanju upotrebom takozvane  $P$ -vrednosti. Što je ova veličina manja, to su dokazi protiv  $H_0$  jači. U tom smislu, kad god je neophodno okarakterisati jačinu dokaza, to činimo na sledeći način:

SLUČAJ	INTERPRETACIJA
$P < 0.01$	veoma jak dokaz protiv $H_0$
$0.01 \leq P < 0.05$	umeren dokaz protiv $H_0$
$0.05 \leq P < 0.1$	nagoveštaj protiv $H_0$
$0.1 \leq P$	malo ili bez realnih dokaza protiv $H_0$

Dakle, u svetlu interpretacije statističkog zaključka upotrebom  $P$ -vrednosti, u prethodnom zadatku možemo konstatovati da su dokazi protiv  $H_0$  *veoma jaki*.



## Vežbe za samostalni rad

**Vežba 4.8** Laborant jedne klinike ispitivao je broj leukocita u serumu 9 bolesnika (izražen u  $10^9/l$ ) i prikupio sledeće podatke:

9.8, 13.5, 10, 11.8, 14.1, 12.6, 13.8, 8.5, 9.5.


Da li, na osnovu datog uzorka, postoje dokazi, sa mogućnošću greške do 5%, da je broj prisutnih leukocita kod obolele osobe iz posmatrane kategorije 9.1?

**Vežba 4.9** Profesor želi da zna da li grupa koji sluša uvod u statistiku ima dobro predznanje iz matematike. Na slučajan način odabrao je 6 studenata iz grupe i testirao njihova znanja iz oblasti matematike. Profesor nastoji da grupa ostvari dobar rezultat na testu, i to oko 70 poena. Rezultati koje su odabrani studenti ostvarili su sledeći: 62, 92, 75, 68, 83, 95. Da li profesor, sa sigurnošću od 90%, može očekivati da će srednja vrednost uspeha preostalih studenata na testu biti oko 70?

**Vežba 4.10** Trener Male lige u bejzbolu želi da sazna da li osvajanje poena njegovog tima odgovara nacionalnom učinku ostalih timova. Na nacionalnom nivou, prosečan broj osvojenih poena po utakmici iznosi 5.7. Trener je, na slučajan način, odabrao 5 utakmica u kojima je njegov tim ostvario sledeće poene:

5, 9, 4, 11, 8.


Da li se rezultat njegovog tima značajno razlikuje od rezultata na nacionalnom nivou? Pretpostaviti da je  $\alpha=0.05$ .

 **Vežba 4.11** Uprava jedne škole je zainteresovana da sazna da li je došlo do promena čitačkih sposobnosti učenika u odnosu na isto obeležje pre 30 godina. Za potrebe testiranja, na slučajan način, odabrano je 30 učenika. Cilj testiranja je utvrditi da li postoje statistički značajne razlike čitačkih sposobnosti danas i proseka od 78 (pocena na standardizovanom testu) ostvarenog pre 30 godina. Rezultati sadašnjeg testiranja su:

63, 58, 76, 95, 84, 68, 90, 45, 75, 38, 52, 90, 100, 45, 76, 80, 45, 68, 64, 50, 59, 63, 75, 72, 68, 92, 64, 35, 87, 86.

Da li se, sa sigurnošću od 95%, može izvesti zaključak da su se čitačke sposobnosti za poslednjih 30 godina:

- a) promenile,      b) povećale,      c) smanjile?

 **Vežba 4.12** Istraživačka grupa pri Univerzitetu je zainteresovana za proveru rezultata do kojih je došla jedna druga istraživačka grupa prilikom ispitivanja uticaja hemijskog preparata pod nazivom Formula Z-protect, na suzbijanje alfida na biljkama. Minuli rezultati pokazali su da je, nakon tretmana hemijskim preparatom, srednja vrednost koncentracije alfida na biljkama 52. Novim istraživanjem obuhvaćene su 24 biljke, na kojima su izmerene sledeće koncentracije alfida:

45, 58, 72, 52, 56, 49, 48, 67, 71, 52, 62, 56, 38, 39, 54, 52, 61, 42, 45, 46, 51, 34, 58, 68.

Ispitati da li su rezultati novog istraživanja u skladu sa zaključcima do kojih je došla prethodna istraživačka grupa, sa sigurnošću od 99%.



## Analiza varijanse (ANOVA)

**Cilj:** Uporediti sredine  $m$  uzoraka iz  $m$  populacija

**Uslovi:** Populacije moraju imati neprekidnu raspodelu, varijanse u svim populacijama moraju biti jednake, uzorci se biraju nezavisno jedan od drugog.

**Statistika:**

Izvor varijabiliteta	Suma kvadrata	Stepeni slobode	Varijanse između i unutar	F-statistika
između uzoraka	$Q_i$	$m - 1$	$Q_i / (m - 1)$	$\frac{Q_i / (m - 1)}{Q_u / (m(n - 1))}$
unutar uzoraka	$Q_u$	$m(n - 1)$	$Q_u / (m(n - 1))$	
ukupno	$Q$			


- $m$  - broj uzoraka
- $n$  - obim svakog uzorka
- $Q_u$  - suma kvadrata odstupanja podataka od odgovarajućih sredina unutar uzorka
- $Q$  - suma kvadrata odstupanja podataka od ukupne sredine
- $Q_i = Q - Q_u$

**Nulta hipoteza**  $H_0 : \mu_1 = \mu_2 = \dots = \mu_m$

**Alternativna hipoteza**  $H_A$  : Bar dve od ovih sredina se značajno razlikuju

**Kritična vrednost iz tablice:**  $F_{m-1, m(n-1), \alpha}$

**Zaključak:**  $F > F_{m-1, m(n-1), \alpha} \implies H_0$  se odbacuje

 **Vežba 5.1** Na 24 slučajno izabrana kandidata sprovedeno je istraživanje o uticaju zvuka na koncentraciju pri učenju. Kandidati su podeljeni u tri jednake grupe: oni koji uče uz pozadinski zvuk konstantne jačine, oni koji uče uz periodične slučajne promene jačine zvuka i oni koji uče u tišini. Nakon testiranja, koje se sastoji od 10 pitanja iz materijala koji su učili, ispitanici su ostvarili sledeće rezultate:

grupa	broj tačnih odgovora							
konstantan zvuk	7	4	6	8	6	6	2	9
promenljiv zvuk	5	5	3	4	4	7	2	2
bez zvuka	2	4	7	1	2	1	5	5

Da li se, na nivou značajnosti do 5%, na osnovu ovog uzorka, može tvrditi da promene u jačini zvuka utiču značajno na koncentraciju pri učenju?

*Rešenje:* Nezavisno promenljiva, odnosno faktor uticaja, je intenzitet zvuka, a zavisno promenljiva je koncentracija pri učenju, izražena brojem tačnih odgovora. S obzirom da nulta hipoteza uvek upućuje na *status quo*, formulisaćemo je u sledećem obliku,

$H_0$  : Intenzitet zvuka ne utiče značajno na koncentraciju pri učenju.

Samim tim, alternativna hipoteza glasi

$H_A$  : Intenzitet zvuka značajno utiče na koncentraciju pri učenju.

Najpre računamo sredine uzoraka i ukupnu sredinu,

	Sredine uzoraka								Zbir	Sredina
U1	7	4	6	8	6	6	2	9	48	6
U2	5	5	3	4	4	7	2	2	32	4
U3	2	4	7	1	2	1	5	5	27	3.375
Ukupna sredina										4.4583

a zatim određujemo sume kvadrata odstupanja od ukupne sredine ( $Q$ ) kao i sume kvadrata odstupanja unutar uzoraka, ( $Q_u$ ), dok sume kvadrata odstupanja između uzoraka ( $Q_i$ ) računamo kao:

$$Q_i = Q - Q_u.$$

	Kvadrati odstupanja od ukupne sredine								Zbir
U1	6.4601	0.2101	2.3767	12.5434	2.3767	2.3767	6.0434	20.6267	53.0139
U2	0.2934	0.2934	2.1267	0.2101	0.2101	6.4601	6.0434	6.0434	21.6806
U3	6.0434	0.2101	6.4601	11.9601	6.0434	11.9601	0.2934	0.2934	43.2639
$Q$									117.9583


Kvadrati odstupanja unutar uzorka									Zbir
U1	1	4	0	4	0	0	16	9	34
U2	1	1	1	0	0	9	4	4	20
U3	1.8906	0.3906	13.141	5.6406	1.8906	5.6406	2.6406	2.6406	33.875
$Q_u$									87.875

$m$	3
$n$	8
Nivo značajnosti	0.05
Kritična vrednost	3.4668

Konačno, računamo ANOVA tabelu:

Izvor varijabiliteta	Suma kvadrata	Stepeni slobode	Varijanse	F statistika
Između uzoraka	30.0833	2	15.0417	3.5946
Unutar uzoraka	87.8750	21	4.1845	
Ukupno	117.9583			

**Zaključak:** S obzirom da je  $F = 3.5946 > 3.4668 = F_{2,21,0.05}$ , na osnovu datog uzorka zaključujemo da se na nivou značajnosti do 5%  $H_0$  odbacuje u korist  $H_A$ , odnosno da sa dozvoljenom greškom zaključivanja do 5% možemo tvrditi da promene u jačini zvuka utiču na koncentraciju pri učenju.

 **Vežba 5.2** Radnik fabrike koja pakuje kavijar u teglice je, radi kontrole kvaliteta, na slučajan način izabrao po 6 teglica kavijara sa svakog od 5 postrojenja za pakovanje. Neto sadržaj kavijara je dat u sledećoj tabeli:

postrojenje	sadržaj teglice u gramima					
1	103	101	99	98	97	105
2	101	100	103	99	102	101
3	102	99	98	100	100	101
4	100	102	103	101	100	100
5	101	97	98	99	99	100

Da li se, na nivou značajnosti do 5%, na osnovu ovog uzorka, može tvrditi da sva postrojenja proizvode teglice kavijara iste prosečne mase?

*Rešenje:* S obzirom da je nezavisno promenljiva, odnosno faktor uticaja, postrojenje na kojem se vrši pakovanje kavijara u teglice, a zavisno promenljiva sadržaj (izražen u gramima) koji se tom prilikom napuni, nulta hipoteza glasi

$H_0$  : Izbor postrojenja ne utiče na prosečnu masu teglice,

dok je alternativna

$H_A$  : Izbor postrojenja utiče na prosečnu masu teglice.

	Sredine uzoraka						Zbir	Sredina
U1	103	101	99	98	97	105	603	100.5
U2	101	100	103	99	102	101	606	101
U3	102	99	98	100	100	101	600	100
U4	100	102	103	101	100	100	606	101
U5	101	97	98	99	99	100	594	99
							Ukupna sredina	100.3

	Kvadrati odstupanja od ukupne sredine						Zbir	
U1	7.29	0.49	1.69	5.29	10.89	22.09	47.74	
U2	0.49	0.09	7.29	1.69	2.89	0.49	12.94	
U3	2.89	1.69	5.29	0.09	0.09	0.49	10.54	
U4	0.09	2.89	7.29	0.49	0.09	0.09	10.94	
U5	0.49	10.89	5.29	1.69	1.69	0.09	20.14	
							$Q$	102.3

	Kvadrati odstupanja unutar uzorka						Zbir	
U1	6.25	0.25	2.25	6.25	12.25	20.25	47.5	
U2	0	1	4	4	1	0	10	
U3	4	1	4	0	0	1	10	
U4	1	1	4	0	1	1	8	
U5	4	4	1	0	0	1	10	
							$Q_u$	85.5

$m$	5
$n$	6
Nivo značajnosti	0.05
Kritična vrednost	2.7587

Izvor varijabiliteta	Suma kvadrata	Stepeni slobode	Varijanse	F statistika
Između uzoraka	16.8	4	4.2	1.2281
Unutar uzoraka	85.5	25	3.42	
Ukupno	102.3			



**Zaključak:** Kako je  $F = 1.2281 < 2.7587 = F_{4,25,0.05}$ , na osnovu datog uzorka i sa mogućnošću greške do 5%, zaključujemo da nemamo dovoljno dokaza protiv  $H_0$ , odnosno da postrojenja proizvode teglice kavijara približno iste prosečne mase.

**Vežba 5.3** Radi testiranja novog sistema obuke zaposlenih, jedna firma je na slučajan način izabrala dvadeset ispitanika, koji su podeljeni u četiri grupe po pet. Prva grupa je primila obuku tipa A, druga tipa B, treća tipa C a četvrta tipa D. Zatim je, tokom probnog perioda od mesec dana, merena efikasnost ispitanika u obavljanju radnih zadataka. Rezultati testiranja su dati u tabeli:

tip obuke	efikasnost u %				
A	87	79	85	92	77
B	72	71	80	73	84
C	92	85	87	78	83
D	70	72	70	88	75

Da li se, na nivou značajnosti do 1%, na osnovu ovog uzorka, može tvrditi da tip obuke ne utiče na efikasnost obavljanja radnih zadataka? Da li se do istog zaključka dolazi i na nivou značajnosti do 5%?

Rešenje:

Shodno tekstu zadatka, zaključujemo da je nezavisno promenljiva tip obuke, a zavisno promenljiva efikasnost u % registrovana nakon što je obuka sprovedena. S tim u vezi, nulta hipoteza glasi

$H_0$  : Tip obuke ne utiče na efikasnost,

a alternativna je da

$H_A$  : Tip obuke utiče na efikasnost.

	Sredine uzoraka					Zbir	Sredina
U1	87	79	85	92	77	420	84
U2	72	71	80	73	84	380	76
U3	92	85	87	78	83	425	85
U4	70	72	70	88	75	375	75
Ukupna sredina							80

	Kvadrati odstupanja od ukupne sredine					Zbir
U1	49	1	25	144	9	228
U2	64	81	0	49	16	210
U3	144	25	49	4	9	231
U4	100	64	100	64	25	353
$Q$						1022

Kvadrati odstupanja unutar uzorka						Zbir
U1	9	25	1	64	49	148
U2	16	25	16	9	64	130
U3	49	0	4	49	4	106
U4	25	9	25	169	0	228
$Q_u$						612


$m$	4
$n$	5
Nivo značajnosti	0.01
Kritična vrednost	5.2922

$m$	4
$n$	5
Nivo značajnosti	0.05
Kritična vrednost	3.2389

Izvor varijabiliteta	Suma kvadrata	Stepeni slobode	Varijanse	F statistika
Između uzoraka	410	3	136.67	3.5730
Unutar uzoraka	612	16	38.25	
Ukupno	1022			

**Zaključak:** Pošto je  $F = 3.5730 < 5.2922 = F_{3,16,0.01}$ , na osnovu datog uzorka i na nivou značajnosti do 1%, ne možemo zaključiti da tip obuke utiče na efikasnost radnika u obavljanju radnih zadataka. Međutim, ako dozvolimo mogućnost greške do 5%, budući da je  $F = 3.5730 > 3.2389 = F_{3,16,0.05}$ , tada možemo zaključiti da tip obuke utiče na efikasnost (odbacujemo nultu hipotezu).

## Vežbe za samostalni rad

 **Vežba 5.4** Da bi se ispitalo uticaj hrane na porast mase pilića (PMP u %), uzete su tri hranljive smese: A, B i C. Pilići su podeljeni u tri grupe i tokom mesec dana hranjeni su odgovarajućom smesom. Podaci dobijeni nakon isteka mesec dana dati su u tabeli. Da li se, na nivou značajnosti do 0.01, može tvrditi da vrsta hrane utiče na porast mase pilića?

Vrsta hrane	A	B	C
P	13.7	12.2	13.9
	14.2	13.2	14.3
M	14.1	12.1	13.6
P	13.5	12.8	14.1
	13.2	13.2	13.4

**Vežba 5.5** Na 15 slučajno izabranih kandidata sprovedeno je istraživanje o uticaju kardio vežbi na potrošnju kalorija. Kandidati su podeljeni u tri jednake grupe: oni koji trče, oni koji plivaju i oni koji voze bicikl. Dobijeni rezultati su prikazani u tabeli. Da li se, na nivou značajnosti do 5%, može tvrditi da promene u načinu vežbanja značajno utiču na gubitak kalorija?

Tip vežbanja	gubitak kalorija nakon 1h				
trčanje	800	850	750	820	900
plivanje	720	800	820	740	830
bicikl	850	870	900	920	890

**Vežba 5.6** Farmaceutska kompanija vrši testiranje tri vrste analgetika za tretman glavobolje. Za potrebe istraživanja, na slučajan način izabrano je 27 ispitanika i oni su podeljeni u tri grupe kojima su dodeljeni lekovi tipa A, B odnosno C. Ispitanici su upućeni da uzmu lek u slučaju narednog napada migrene, te da kvantifikuju subjektivnu procenu intenziteta bola nakon određenog vremena na skali od 1 do 10. Podaci do kojih se došlo prezentovani su tabelom. Utvrditi da li, sa dozvoljenom greškom do 1%, možemo tvrditi da izbor vrste leka daje različite rezultate u tretmanu glavobolje.

lek	intenzitet bola								
A	4	5	4	3	2	4	3	4	4
B	6	8	4	5	4	6	5	8	6
C	6	7	6	6	7	5	6	5	5

**Vežba 5.7** U okviru istraživanja uticaja bakterije *Dehalococcoides ethogenes* na razgradnju toksičnog otpada koji sadrži hlor, u in vitro uslovima posmatran je uticaj temperature na proces ekstrakcije hlora. Eksperiment je sproveden po 10 puta na temperaturama od 15, 20 i 25 °C i zabeležen je procenat ekstrahovanog hlora. Utvrditi da li, sa dozvoljenom greškom do 1%, možemo tvrditi da se proces ekstrakcije hlora odvija različito na različitim temperaturama?

t u °C	količina ekstrahovanog hlora u %									
15	45	35	41	33	32	43	44	45	43	46
20	56	57	45	43	51	55	54	48	51	49
25	31	34	41	37	32	31	28	27	31	33

**Vežba 5.8** Ispituje se primena posebne metode sticanja manuelnih veština na grupi muških osoba starosti 30-35 godina. Ispitanici su podeljeni prema nivou obrazovanja. Nakon sprovedenog testiranja veština, zabeležen je procenat uspešnosti za svakog ispitanika. Utvrditi da li, sa dozvoljenom greškom do 1%, možemo tvrditi da pomenuta metoda rada daje različite rezultate u zavisnosti od obrazovanja ispitanika?

nivo obrazovanja	uspešnost na testiranju u %							
srednje	75	62	77	58	78	55	45	80
više	95	86	77	65	66	68	70	56
visoko	75	65	55	80	82	65	45	65

**Vežba 5.9** Sprovodi se istraživanje sa ciljem da se ispita da li postoje razlike u prosečnom dnevnom unosu Ca kod odraslih osoba različitih stanja kostiju. Ispitanici od 60 godina starosti, izabrani na slučajan način pomoću bolničkog registra, pozvani su da pristupe istraživanju i podeljeni su u 3 kategorije, i to one koji imaju kosti normalne gustine (NGK), one koji imaju osteopeniju (OA) i one kojima je konstatovana osteoporozna (OS). Za svakog ispitanika ponaosob registrovan je dnevni unos Ca (DUCa), putem hrane i suplemenata, a podaci su predstavljeni tabelom u nastavku. Da li se, na nivou značajnosti do 0.05, može tvrditi da stanje kostiju diktira količinu dnevnog unosa pomenutog minerala?

stanje kostiju	NGK	OA	OS
D	1200	1000	890
	1000	1100	650
U	980	700	1100
Ca	900	800	900
	750	500	400

**Vežba 5.10** Ispituje se primena posebne metode obrazovanja odraslih u oblasti novih "zelenih" tehnologija u drvenoj industriji. Ispitanici su podeljeni prema starosti u tri grupe. Nakon sprovedenog testiranja, zabeležen je procenat uspešnosti za svakog ispitanika. Utvrditi da li, sa dozvoljenom greškom do 5%, možemo zaključiti da pomenuta metoda obrazovanja daje jednake rezultate bez obzira na godište ispitanika?

starosna grupa	uspešnost na testiranju u %							
grupa 30-35	75	62	77	58	89	60	45	80
grupa 35-40	90	86	43	53	66	42	70	55
grupa 40-45	70	67	45	88	92	65	45	78


## Linearna korelacija

**Pearson-ov koeficijent:**  $r_{xy} = \frac{s_{xy}}{s_x s_y}$

- $n$  - veličina uzorka
- $x_i$  - vrednosti prvog obeležja
- $y_i$  - vrednosti drugog obeležja
- $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$  - sredina prvog obeležja
- $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$  - sredina drugog obeležja
- $s_x^2 = \frac{1}{n} \sum_{i=1}^n x_i^2 - \bar{x}^2$  - varijansa prvog obeležja
- $s_y^2 = \frac{1}{n} \sum_{i=1}^n y_i^2 - \bar{y}^2$  - varijansa drugog obeležja
- $s_{xy} = \frac{1}{n} \sum_{i=1}^n x_i y_i - \bar{x} \bar{y}$  - kovarijansa

INTERVAL KOJEM PRIPADA VREDNOST $ r $	INTERPRETACIJA
[0.00, 0.40)	slaba korelacija
[0.40, 0.75)	umerena korelacija
[0.75, 0.85)	dobra korelacija
[0.85, 1.00]	odlična korelacija

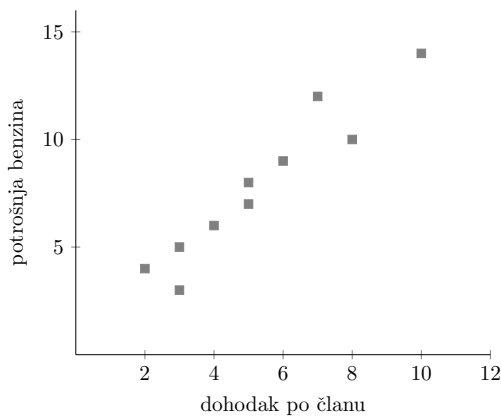
**Koeficijent determinacije:**  $r_{xy}^2$

 **Vežba 6.1** Prikupljeni podaci o prosečnom dohotku po članu domaćinstva i potrošnji benzina za nekoliko slučajno izabranih domaćinstava u jednom gradu dati su u sledećoj tabeli:

dohodak po čl. domać. u 10000 RSD	2	3	3	4	5	5	6	7	8	10
potrošnja benzina u 10l	4	3	5	6	7	8	9	12	10	14

Prikazati podatke u vidu dijagrama rasipanja, a zatim odrediti koeficijent korelacije i koeficijent determinacije.

**Rešenje:** Najpre ćemo podatke predstaviti grafički, u formi *dijagrama rasipanja*, a zatim na osnovu tabele za pomoćna izračunavanja odrediti statistike za oba obeležja, kao i kovarijansu, koeficijent korelacije i koeficijent determinacije.



$i$	$x_i$	$y_i$	$x_i^2$	$y_i^2$	$x_i y_i$
1	2	4	4	16	8
2	3	3	9	9	9
3	3	5	9	25	15
4	4	6	16	36	24
5	5	7	25	49	35
6	5	8	25	64	40
7	6	9	36	81	54
8	7	12	49	144	84
9	8	10	64	100	80
10	10	14	100	196	140
$\Sigma$	53	78	337	720	489

Statistike za prvo obeležje		Statistike za drugo obeležje	
$\bar{x}$	5.3	$\bar{y}$	7.8
$s_x^2$	5.61	$s_y^2$	11.16
$s_x$	2.3685	$s_y$	3.3407

Kovarijansa	
$s_{xy}$	7.56
Koeficijent korelacije	
$r_{xy}$	0.9555
Koeficijent determinacije	
$r_{xy}^2$	0.9129

**Zaključak:** S obzirom da je koeficijent korelacije pozitivan, zaključujemo da je linearna veza između dohotka i potrošnje benzina direktna, što znači da povećanje dohotka implicira i povećanu potrošnju benzina. Dodatno, s obzirom da je koeficijent korelacije 0.9555, zaključujemo da je u pitanju odlična korelacija.

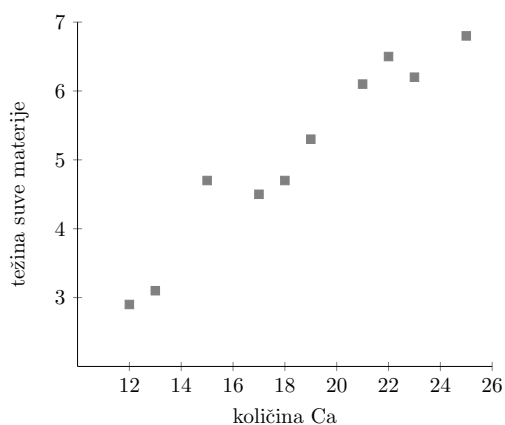
Koeficijent determinacije iznosi 91.29%, što znači da se 91.29% varijacija jedne promenljive objašnjava, odnosno pripisuje varijacijama druge promenljive.

**Vežba 6.2** Ispituje se težina suve materije paradajza uzgajanog na tlu sa različitim količinom kalcijuma. Podaci su dati u tabeli:

Ca u tlu ( $mg/cm^3$ )	12	13	15	17	18	19	21	22	23	25
težina suve materije u g	2.9	3.1	4.7	4.5	4.7	5.3	6.1	6.5	6.2	6.8

Nacrtati dijagram rasipanja i ispitati korelaciju između sadržaja Ca u tlu i težine suve materije u plodu paradajza.

**Rešenje:** Kao i u prethodnom primeru, najpre ćemo podatke prikazati grafički, a potom pristupiti izračunavanjima.



$i$	$x_i$	$y_i$	$x_i^2$	$y_i^2$	$x_i y_i$
1	12	2.9	144	8.41	34.8
2	13	3.1	169	9.61	40.3
3	15	4.7	225	22.09	70.5
4	17	4.5	289	20.25	76.5
5	18	4.7	324	22.09	84.6
6	19	5.3	361	28.09	100.7
7	21	6.1	441	37.21	128.1
8	22	6.5	484	42.25	143
9	23	6.2	529	38.44	142.6
10	25	6.8	625	46.24	170
$\Sigma$	185	50.8	3591	274.68	991.1

Statistike za prvo obeležje		Statistike za drugo obeležje	
$\bar{x}$	18.5	$\bar{y}$	5.08
$s_x^2$	16.85	$s_y^2$	1.6616
$s_x$	4.1049	$s_y$	1.2890

Kovarijansa	
$s_{xy}$	5.13
Koefficient korelacije	
$r_{xy}$	0.9695
Koefficient determinacije	
$r_{xy}^2$	0.9400

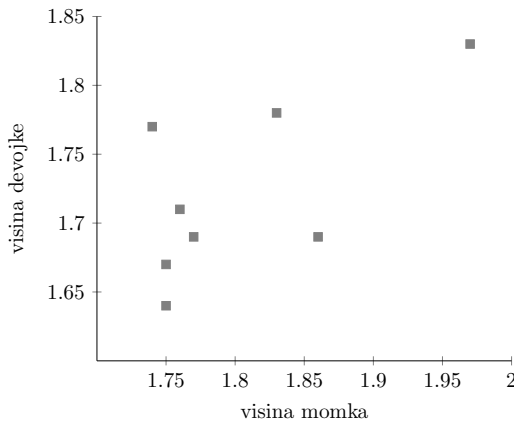
**Zaključak:** Linearna povezanost između količine Ca i težine suve materije paradajza je direktna i odlična, budući da je koeficijent korelacije 0.9695. Koeficijent determinacije je 94%.

**Vežba 6.3** U popularnom ženskom časopisu objavljen je članak o tome koliko je devojka važna visina muškarca. Za jedan od pokazatelja izabrana je korelacija između visine momka i visine devojke za nekoliko nasumično odabranih parova:

visina momka u m	1.76	1.75	1.74	1.86	1.77	1.83	1.97	1.75
visina devojke u m	1.71	1.67	1.77	1.69	1.69	1.78	1.83	1.64

Nacrtati dijagram rasipanja, odrediti koeficijent korelacije između visine momka i visine devojke i koeficijent determinacije.

Rešenje:




$i$	$x_i$	$y_i$	$x_i^2$	$y_i^2$	$x_i y_i$
1	1.76	1.71	3.1	2.92	3.01
2	1.75	1.67	3.06	2.79	2.92
3	1.74	1.77	3.03	3.13	3.08
4	1.86	1.69	3.46	2.86	3.14
5	1.77	1.69	3.13	2.86	2.99
6	1.83	1.78	3.35	3.17	3.26
7	1.97	1.83	3.88	3.35	3.6
8	1.75	1.64	3.06	2.69	2.87
$\Sigma$	14.43	13.78	26.07	23.67	24.87

Statistike za prvo obeležje		Statistike za drugo obeležje	
$\bar{x}$	1.8038	$\bar{y}$	1.7225
$s_x^2$	0.0055	$s_y^2$	0.0036
$s_x$	0.0745	$s_y$	0.0602

Kovarijansa	
$s_{xy}$	0.0029
Koeficijent korelacije	
$r_{xy}$	0.6466
Koeficijent determinacije	
$r_{xy}^2$	0.4181

**Zaključak:** Na osnovu prikupljenih podataka možemo zaključiti da je linearna veza između visine momka i visine devojke direktna, ali da je umerena. Koeficijent determinacije iznosi 41.81%.

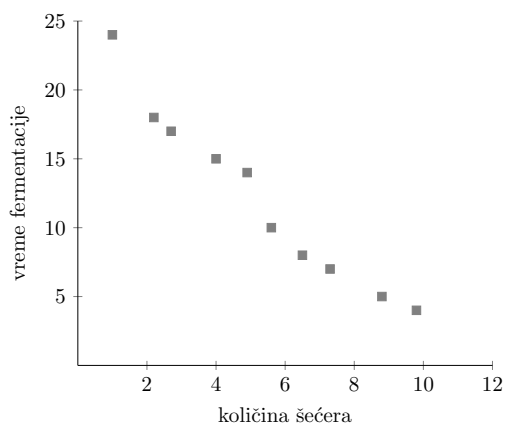


 **Vežba 6.4** Pri proizvodnji multivitaminskog soka pretpostavlja se da brzina fermentacije zavisi od količine dodatog šećera. Na slučajnom uzorku dobijeni su podaci:

kol. dodatog šećera u kg	1.0	2.2	2.7	4.0	4.9	5.6	6.5	7.3	8.8	9.8
vreme ferm. u danima	24	18	17	15	14	10	8	7	5	4

Nacrtati dijagram rasipanja i ispitati korelaciju između količine dodatog šećera i brzine fermentacije.

Rešenje:



$i$	$x_i$	$y_i$	$x_i^2$	$y_i^2$	$x_i y_i$
1	1	24	1	576	24
2	2.2	18	4.84	324	39.6
3	2.7	17	7.29	289	45.9
4	4	15	16	225	60
5	4.9	14	24.01	196	68.6
6	5.6	10	31.36	100	56
7	6.5	8	42.25	64	52
8	7.3	7	53.29	49	51.1
9	8.8	5	77.44	25	44
10	9.8	4	96.04	16	39.2
$\Sigma$	52.8	122	353.52	1864	480.4

Statistike za prvo obeležje		Statistike za drugo obeležje	
$\bar{x}$	5.28	$\bar{y}$	12.2
$s_x^2$	7.4736	$s_y^2$	37.56
$s_x$	2.7338	$s_y$	6.1286

Kovarijansa	
$s_{xy}$	-16.376
Koeficijent korelacije	
$r_{xy}$	-0.9774
Koeficijent determinacije	
$r_{xy}^2$	0.9553

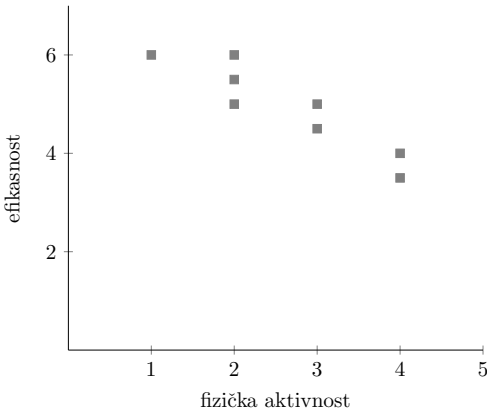
**Zaključak:** Budući da je koeficijent korelacije negativan, linearna veza između količine dodatog šećera i vremena fermentacije je obrnuta, što znači da veća količina šećera podrazumeva kraće vreme fermentacije. Apsolutna vrednost koeficijenta korelacije, koja iznosi 0.9774, govori da je linearna korelacija ova dva obeležja odlična.

**Vežba 6.5** U popularnom muškom časopisu objavljuje se članak o tome kako nivo fizičkih aktivnosti utiče na radnu sposobnost. Za jedan od pokazatelja izabrana je korelacija između broja sati fizičkog vežbanja dnevno i efikasnosti na poslu, izraženoj kroz vreme potrebno da se obave radni zadaci. Podaci prikupljeni na slučajnom uzorku od nekoliko muškaraca starosne dobi 30-35 godina dati su u tabeli:

fizičke aktivnosti u h	1	2	2	3	3	2	4	4
efikasnost u h	6	5.5	5	5	4.5	6	4	3.5

Nacrtati dijagram rasipanja, odrediti koeficijent korelacije između broja sati fizičkog vežbanja i efikasnosti na poslu i koeficijent determinacije.

Rešenje:




$i$	$x_i$	$y_i$	$x_i^2$	$y_i^2$	$x_i y_i$
1	1	6	1	36	6
2	2	5.5	4	30.25	11
3	2	5	4	25	10
4	3	5	9	25	15
5	3	4.5	9	20.25	13.5
6	2	6	4	36	12
7	4	4	16	16	16
8	4	3.5	16	12.25	14
$\Sigma$	21	39.5	63	200.75	97.5

Statistike za prvo obeležje		Statistike za drugo obeležje	
$\bar{x}$	2.6250	$\bar{y}$	4.9375
$s_x^2$	0.9844	$s_y^2$	0.7148
$s_x$	0.9922	$s_y$	0.8455

Kovarijansa	
$s_{xy}$	-0.7734
Koeficijent korelacije	
$r_{xy}$	-0.9220
Koeficijent determinacije	
$r_{xy}^2$	0.8501


**Zaključak:** Koeficijent korelacije između broja sati provedenih u bavljenju fizičkom aktivnošću i efikasnosti u obavljanju radnih zadataka iznosi -0.922, što znači da osobama koje više vremena provedu vežbajući treba manje vremena za kompletiranje radnih obaveza, odnosno da vežbanje ima pozitivan uticaj na efikasnost. Apsolutna vrednost koeficijenta korelacije je 0.922 i ona sugeriše da je linearna veza između pomenuta dva obeležja odlična.

## Vežbe za samostalni rad

 **Vežba 6.6** Ispituje se uticaj duvana na životni vek ljudi. Na osnovu podataka iz 2013. godine imamo sledeći slučajni uzorak:


broj cigareta dnevno	10	17	5	20	35	10
životni vek	60	62	79	59	55	65

Odrediti koeficijent korelacije i protumačiti vezu između broja cigareta konzumiranih na dnevnom nivou, sa jedne, i životnog veka, sa druge strane.

 **Vežba 6.7** Godišnja dobit jedne kompanije i ulaganja u reklamu te kompanije dati su u tabeli:

dobit kompanije (u milionima RSD)	12	14	20	26	50
ulaganja u reklamu (u milionima RSD)	2	4	5	6	10


Ispitati korelaciju.

 **Vežba 6.8** Među dečacima uzrasta 10 godina ispituje se da li dužina skoka iz mesta zavisi od mase pojedinca. Ako su na osnovu slučajnog uzorka dobijeni su sledeći podaci, ispitati korelaciju.

telesna masa (kg)	42	50	35	40	62
dužina skoka (m)	1.35	1.2	1.7	1.7	1.1

 **Vežba 6.9** Ispituje se da li broj sklopljenih brakova utiče na broj novorođene dece. Ispitati korelaciju, na osnovu podataka za 5 slučajno odabranih meseci:

broj novorođenčadi	48	40	54	53	80
broj sklopljenih brakova	35	20	30	42	50

 **Vežba 6.10** Među studentima koji polažu Matematiku sa statistikom, ispituje se da li druženje utiče na uspešnost na kolokvijumu. Anketirano je pet studenata. U tabeli su dati podaci o vremenu koje student provede dnevno družeći se (u minutima) i broju bodova osvojenih na kolokvijumu:

broj bodova	10	50	28	43	35
druženje	240	60	150	85	40

Ispitati korelaciju.



## Linearna regresija

Prva regresiona prava:

$$\hat{y} = ax + b, \quad a = \frac{s_{xy}}{s_x^2}, \quad b = \bar{y} - a\bar{x}$$

Druga regresiona prava:

$$\hat{x} = cy + d, \quad c = \frac{s_{xy}}{s_y^2}, \quad d = \bar{x} - c\bar{y}$$

**Vežba 7.1** Prikupljeni podaci o prosečnom dohotku po članu domaćinstva i potrošnji benzina, za nekoliko slučajno izabranih domaćinstava u jednom gradu, dati su u sledećoj tabeli:

dohodak po čl. dom. u 10000 RSD	2	3	3	4	5	5	6	7	8	10
potrošnja benzina u 10l	4	3	5	6	7	8	9	12	10	14

Odrediti prvu regresionu pravu i prikazati je na dijagramu rasipanja. Zatim, odrediti očekivanu potrošnju benzina domaćinstva čiji članovi u proseku primaju 55000 RSD.

Rešenje: Označimo sa  $x$  dohodak po članu, a sa  $y$  potrošnju benzina. Pomoćna tabela je data u rešenju zadatka 6.1, stoga navodimo samo izračunate statistike.

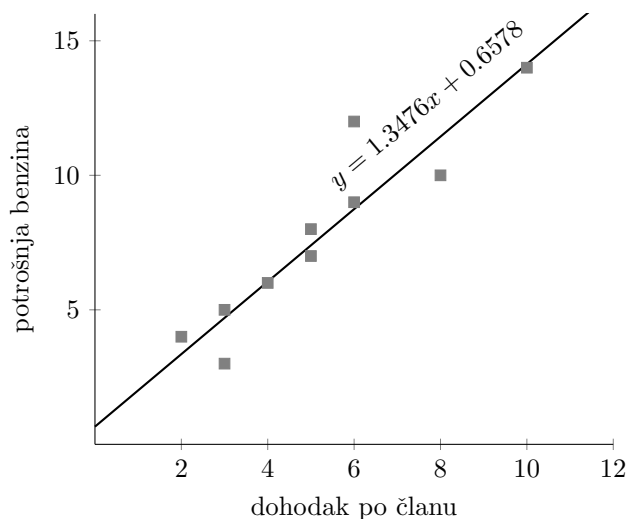
Statistike za $x$		Statistike za $y$			
$\bar{x}$	5.3	$\bar{y}$	7.8	Kovarijansa	Koeficijent korelacije
$s_x^2$	5.61	$s_y^2$	11.16	$s_{xy}$	7.56
$s_x$	2.3685	$s_y$	3.3407	$r_{xy}$	0.9555

Dalje, koeficijenti prve regresione prave su:

Koeficijenti prve reg. prave	
$a$	1.3476
$b$	0.6578

Dakle, jednačina prve regresione prave je


$$\hat{y} = 1.3476x + 0.6578.$$



Da bismo prognozirali potrošnju benzina domaćinstva čiji članovi u proseku primaju 55000 RSD, koristićemo upravo prvu regresionu pravu, i to tako što ćemo za vrednost  $x$  uzeti  $x = 5.5$ :

Prognoza	
$x$	5.5
$\hat{y}$	8.0695

**Zaključak:** Dakle, za domaćinstvo koje ima prosečni mesečni prihod od 55000 RSD očekuje se potrošnja od oko 8 litara benzina.

 **Vežba 7.2** Ispituje se težina suve materije paradajza uzgajanog na tlu sa različitom količinom kalcijuma. Podaci su dati u tabeli:

<i>Ca u tlu (<math>mg/cm^3</math>)</i>	12	13	15	17	18	19	21	22	23	25
<i>težina suve materije u g</i>	2.9	3.1	4.7	4.5	4.7	5.3	6.1	6.5	6.2	6.8

Odrediti neophodan sadržaj kalcijuma u zemljištu na kome se očekuje plod paradajza čija će težina suve materije biti 5 grama.

*Rešenje:* Označimo sa  $x$  količinu  $Ca$  u tlu, a sa  $y$  težinu suve materije. S obzirom da nas interesuje procena za  $x$  pri datoj vrednosti za  $y$ , koristićemo drugu regresionu pravu.

Statistike za $x$		Statistike za $y$		Kovarijansa		Koeficijent korelacije	
$\bar{x}$	18.5	$\bar{y}$	5.08	$s_{xy}$	5.13	$r_{xy}$	0.9695
$s_x^2$	16.85	$s_y^2$	1.6616				
$s_x$	4.1049	$s_y$	1.2890				

Koeficijenti druge reg. prave	
$c$	3.0874
$d$	2.8161


Jednačina druge regresione prave je

$$\hat{x} = 3.0874y + 2.8161.$$

Odatle, uzimajući za  $y = 5$  i uvrštavajući ovu vrednost u jednačinu druge regresione prave, dobijamo procenu neophodnog sadržaja  $Ca$  u tlu.

Prognoza	
$y$	5
$\hat{x}$	18.2530

**Zaključak:** U zemljištu je neophodno da bude  $18,25 \text{ mg/cm}^3$   $Ca$ , kako bi se obezbedila količina od 5 g suve materije paradajza.

 **Vežba 7.3** U popularnom ženskom časopisu objavljuje se članak o tome koliko je devojka važna visina muškarca. Nekoliko slučajno izabranih parova je anketirano i dobijeni su sledeći rezultati:

visina momka u m	1.76	1.75	1.74	1.86	1.77	1.83	1.97	1.75
visina devojke u m	1.71	1.67	1.77	1.69	1.69	1.78	1.83	1.64

Kakve je rezultate o očekivanoj visini momka za devojku visoku 175 cm objavio časopis?

*Rešenje:* Neka je sa  $x$  označena visina momka, a sa  $y$  visina devojke. S obzirom na postavljeno pitanje, zaključujemo da treba odrediti jednačinu druge regresione prave.

Statistike za $x$		Statistike za $y$		Kovarijansa		Koefficijent korelacije	
$\bar{x}$	1.8038	$\bar{y}$	1.7225	$s_{xy}$	0.0029	$r_{xy}$	0.6466
$s_x^2$	0.0055	$s_y^2$	0.0036				
$s_x$	0.0745	$s_y$	0.0602				

Koefficijenti druge reg. prave	
$c$	0.8092
$d$	0.4100

Jednačina druge regresione prave glasi

$$\hat{x} = 0.8092y + 0.41.$$

Da bismo procenili očekivanu visinu momka čija je devojka visoka 175 cm, uvrstićemo u jednačinu regresione prave vrednost  $y = 175$ .

Prognoza	
$y$	1.75
$\hat{x}$	1.8260

**Zaključak:** Očekivana visina momka, čija je devojka visoka 175 cm, je 183 cm.



**Vežba 7.4** Pri proizvodnji multivitaminskog soka pretpostavlja se da brzina fermentacije zavisi od količine dodatog šećera. Na slučajnom uzorku dobijeni su podaci:

kol. dodatog šećera u kg	1.0	2.2	2.7	4.0	4.9	5.6	6.5	7.3	8.8	9.8
vreme ferm. u danima	24	18	17	15	14	10	8	7	5	4

Odrediti prvu regresionu pravu i prikazati je na dijagramu rasipanja. Zatim, odrediti očekivano vreme fermentacije soka u koji je dodato 5 kg šećera.

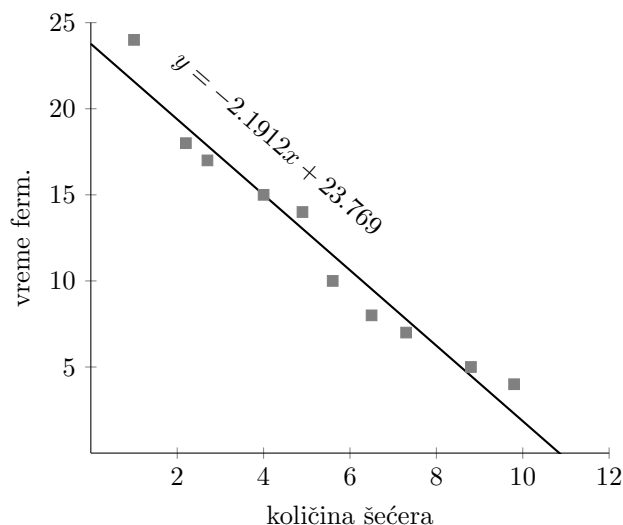
Rešenje: Označimo sa  $x$  količinu dodatog šećera, a sa  $y$  vreme fermentacije.

Statistike za $x$		Statistike za $y$		Kovarijansa		Koeficijent korelacije	
$\bar{x}$	5.28	$\bar{y}$	12.2	$s_{xy}$	-16.38	$r_{xy}$	-0.9776
$s_x^2$	7.4736	$s_y^2$	37.56				
$s_x$	2.7338	$s_y$	6.1286				

Koeficijenti prve reg. prave	
$a$	-2.1912
$b$	23.7694

Jednačina prve regresione prave glasi


$$\hat{y} = -2.1912x + 23.7694.$$



Prognoza	
$x$	5
$\hat{y}$	12.8135

Da bismo procenili vreme fermentacije u slučaju dodatka 5 kg šećera, uvrstićemo u jednačinu regresione prave vrednost  $x = 5$ .

**Zaključak:** Očekivano vreme fermentacije soka u koji je dodato 5 kg šećera iznosi približno 13 dana.

 **Vežba 7.5** U popularnom muškom časopisu objavljuje se članak o tome kako nivo fizičkih aktivnosti utiče na radnu sposobnost. Za jedan od pokazatelja izabrana je korelacija između broja sati fizičkog vežbanja dnevno i efikasnosti na poslu, izraženoj kroz vreme potrebno da se obave radni zadaci. Podaci prikupljeni na slučajnom uzorku od nekoliko muškaraca starosti 30-35 godina dati su u tabeli:

fizičke aktivnosti u h	1	2	2	3	3	2	4	4
efikasnost u h	6	5.5	5	5	4.5	6	4	3.5

Kakve je rezultate objavio časopis o tome kolika je očekivana efikasnost muškarca koji ne radi fizičke vežbe i o tome koliko treba vežbati da očekivana efikasnost iznosi 2h?

*Rešenje:* Neka je sa  $x$  označeno vreme koje muškarac provede baveći se fizičkom aktivnošću, a sa  $y$  njegova efikasnost u obavljanju radnih zadataka.

Statistike za $x$		Statistike za $y$		Kovarijansa		Koeficijent korelacije	
$\bar{x}$	2.6250	$\bar{y}$	4.9375	$s_{xy}$	-0.7734	$r_{xy}$	-0.9219
$s_x^2$	0.9844	$s_y^2$	0.7148				
$s_x$	0.9922	$s_y$	0.8455				

Koeficijenti prve reg. prave	
$a$	-0.7857
$b$	7.0000

Jednačina prve regresione prave je

$$\hat{y} = -0.7857x + 7.$$

Prognoza	
$x$	0
$\hat{y}$	7.0000

**Zaključak:** Očekivano vreme za obavljanje radnih zadataka prosečnog muškarca koji ne vežba iznosi 7 sati.

Koeficijenti druge reg. prave	
$c$	-1.0820
$d$	7.9672

Jednačina druge regresione prave je

$$\hat{x} = -1.0820y + 7.9672.$$

Prognoza	
$y$	2
$\hat{x}$	5.8032

**Zaključak:** Očekivan broj sati fizičkog vežbanja potrebnog da prosečan muškarac obavi radne zadatke za 2 sata iznosi oko šest sati.

☞ **Vežba 7.6** U jednom preduzeću, tokom 6 meseci, ispituje se veza između broja kišnih dana u mesecu i broja izostanaka sa posla. Podaci su sledeći:

broj kišnih dana u mesecu	16	11	10	8	7	2
broj izostanaka u mesecu	17	13	13	11	10	5

Ispitati korelaciju između broja kišnih dana u mesecu i broja izostanaka sa posla. Na osnovu ovog uzorka pomoću linearne regresije odrediti očekivani broj izostanaka sa posla ako je čitav mesec sunčan. Nacrtati dijagram rasipanja i prvu regresionu pravu.

*Rešenje:* Označimo sa  $x$  broj kišnih dana u mesecu, a sa  $y$  broj izostanaka sa posla.

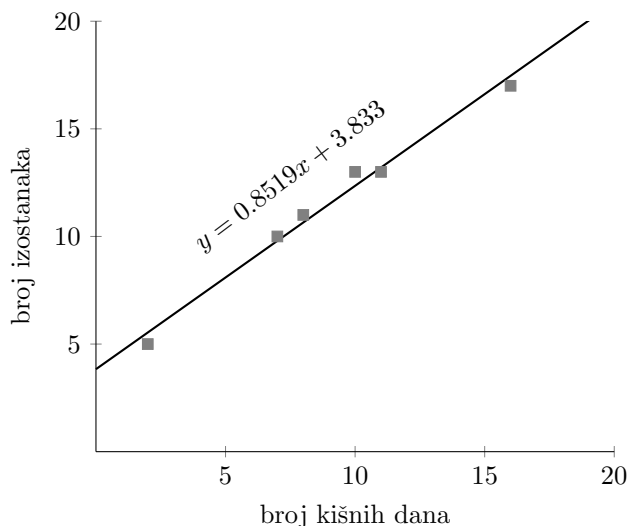
Statistike za $x$		Statistike za $y$		Kovarijansa		Koeficijent korelacije	
$\bar{x}$	9.0000	$\bar{y}$	11.5000	$s_{xy}$	15.3333	$r_{xy}$	0.99
$s_x^2$	18.0000	$s_y^2$	13.2500				
$s_x$	4.2426	$s_y$	3.6401				

Dakle, koeficijent korelacije je 0.99, što govori da je veza između broja dana provedenih na poslu i broja kišnih dana direktna, te da je linearna korelacija odlična.

Ako je čitav mesec sunčan, to znači da je broj kišnih dana  $x = 0$ . Za prognozu koristimo prvu regresionu pravu:

Koeficijenti prve reg. prave		Prognoza	
$a$	0.8519	$x$	0
$b$	3.8333	$\hat{y}$	3.8333

**Zaključak:** Tokom meseca sa svim sunčanim danima očekivani broj izostanaka je približno 4.



**Vežba 7.7** U popularnom muškom časopisu objavljuje se članak o tome koliko dnevna količina unetih zasićenih masti utiče na kondiciju. Za jedan od pokazatelja izabrana je korelacija između dnevne količine unetih zasićenih masti u gramima i vremena potrebnog ispitaniku da sprintom istrči 100 m u sekundama. Dobijeni su sledeći rezultati:

količina zasićenih masti u g	50	55	60	65	70	75	80
vreme u s	12.5	12.9	13	14	14.8	15	15

Ispitati korelaciju, a zatim na osnovu ovog uzorka pomoću linearne regresije odrediti očekivano prolazno vreme ispitanika, ako je dnevna količina unetih zasićenih masti 100 g. Koliko je potrebno da ispitanik dnevno unese zasićenih masti da bi trčao 100 m za 10 s?

**Rešenje:** Ako označimo sa  $x$  količinu zasićenih masti u g, a sa  $y$  vreme u s, tada dobijamo sledeće statistike:

Statistike za $x$		Statistike za $y$	
$\bar{x}$	65	$\bar{y}$	13.8857
$s_x^2$	100	$s_y^2$	1.0012
$s_x$	10	$s_y$	1.0006

Kovarijansa		Koeficijent korelacije	
$s_{xy}$	9.6429	$r_{xy}$	0.9637

Budući da je koeficijent korelacije 0.9637, to znači da je linearna veza između količine unetih zasićenih masti i vremena koje je neophodno da osoba istrči 100 m direktna i odlična.


Koeficijenti prve reg. prave		Prognoza	
$a$	0.0964	$x$	100
$b$	7.6179	$\hat{y}$	17.2607

**Zaključak:** Ako je dnevna količina unetih zasićenih masti 100 g, tada se očekuje prolazno vreme ispitanika oko 17 s.

Koeficijenti druge reg. prave		Prognoza	
$c$	9.6311	$y$	10
$d$	-68.7342	$\hat{x}$	27.5764


**Zaključak:** Očekuje se da ispitanik treba dnevno da unese oko 27 g zasićenih masti, da bi distancu od 100 m istrčao za 10 s.

## Vežbe za samostalni rad

 **Vežba 7.8** Istraživač je, usled porasta broja Internet provajdera, odlučio da ispita da li postoji veza između mesečne cene usluge provajdera (u stotinama RSD), i stepena zadovoljstva kupaca pruženom uslugom (na skali od 1 do 10, pri čemu 1 predstavlja totalno nezadovoljstvo, a 10 apsolutno zadovoljstvo klijenta). Podaci su prikupljeni i predstavljeni u tabeli:

<i>cena</i>	11	18	17	15	9	5	12	19	22	25
<i>zadovoljstvo</i>	6	8	10	4	9	6	3	5	2	10

- (a) Koja je procena zadovoljstva klijenta čiji mesečni izdaci za pomenutu uslugu iznose 1350 RSD?
- (b) Odrediti procenu cene usluge mesečne naknade klijenta čije je zadovoljstvo dato sa 7.5.

 **Vežba 7.9** Posmatra se telesna masa deteta spram njegovih godina života. Podaci su prikazani u tabeli:

<i>godine</i>	2	3	5	7	8
<i>masa</i>	14	20	32	42	44

- (a) Odrediti jednačinu regresione prave koja opisuje zavisnost godina života od telesne mase.
- (b) Na osnovu datih podataka, dati prognozu telesne mase šestogodišnjaka.

**Vežba 7.10** U sklopu evaluacije rada i sposobnosti zaposlenih, svaki od 6 službenika odradio je test sposobnosti, sa namerom da se, na osnovu rezultata testa, utvrdi veza između uspeha na testu i ostvarene prodaje u prvom narednom mesecu angažmana (u stotinama dolara). Podaci su objedinjeni i dati tabelarno:

uspeh na testu ( $x$ )	25	42	33	54	29	36
ostvarena prodaja ( $y$ )	42	72	50	90	45	48

- (a) Izračunati koeficijent korelacije i dati tumačenje rezultata.
- (b) Odrediti jednačinu regresione prave i dati prognozu očekivane prodaje službenika koji na testu osvoji 47 poena.

**Vežba 7.11** Rezultati ispitivanja o broju sati provedenih u snu i onih provedenih gledajući televiziju, koje je obuhvatilo 50 ispitanika, prikazani su u tabeli:

$h$ u snu ( $x$ )	6	7	8	9	10
$h$ ispred televizije ( $y$ )	4	3	3	2	1
apsolutne frekvencije ( $f_i$ )	3	16	20	10	1

- (a) Izračunati koeficijent korelacije.
- (b) Odrediti jednačinu regresione prave.
- (c) Dati prognozu očekivanog broja sati koje osoba provede gledajući televiziju, ako u snu provede 8 sati.

**Vežba 7.12** Pretpostavlja se da fluktuacije nivoa norepinefrina (NE) prate fluktuacije bipolarnog afektivnog poremećaja (manično - depresivne bolesti). Stoga, nivo NE tokom stanja depresije opada, dok je tokom maničnih stanja nivo NE u porastu. Naučnici su, sa namerom da ispituju pomenutu vezu, merili nivo NE tako što su registrovali koncentraciju 3-metoksi-4-hidroksifenilglikol metabolita (MHPG) u mikrogramima na svaka 24 sata. Stepni depresije su takođe zabeleženi, na skali od 0 do 30. Dobijeni podaci su prikazani u tabeli:

MHPG	980	1209	1403	1950	1814	1280	1073	1066	880	776
rasp.	22	26	8	10	5	19	26	12	23	28

- (a) Odrediti koeficijent korelacije. Tumačiti dobijeni rezultat sa stanovišta pretpostavke o postojanju veze između nivoa MHPG i raspoloženja pacijenta.
- (b) Koji procenat odstupanja je objašnjen postojanjem veze između promenljivih?
- (c) Dati prognozu raspoloženja pacijenta kojem je izmeren MHPG na nivou od 1100. Koja je prognoza u slučaju kada je MHPG 950, a koja ako je 700?







## Primeri kolokvijuma

Visine (u cm) 10 bambusovih biljaka koje se rasle tokom meseca juna na lokalitetima sa određenim brojem sunčanih dana su date tabelom:

visina [cm]	49	45	46	48	45	44	43	49	51	53
osunčanost [dani]	17	15	12	22	10	8	10	19	15	21

- a) Šta se može zaključiti na osnovu koeficijenta korelacije?
- b) Dati prognozu očekivanog broja sunčanih dana u junu ako je visina bambusove biljke iznosila pola metra.
- 

Trener DC dresira psa Lolu, a trening se sastoji od nedeljnog istrčavanja (u (km/h)\*10, radi jednostavnosti i izbegavanja decimalnog zapisa), a po završetku treninga trener DC očitava registrovani broj Lolinih otkucaja srca (u bpm). Nakon 7 treninga, podatke je prikazala tabelarno:

brzina [(km/h)*10]	11	16	31	40	50	60	69
broj otkucaja srca [bpm]	61	69	78	80	85	87	90

- a) Šta se može zaključiti na osnovu koeficijenta korelacije?
- b) Dati prognozu očekivane Loline brzine u (km/h)\*10 ako njen trenutni broj otkucaja srca iznosi 99 bpm.
- 

Sprovodi se istraživanje sa ciljem ispitivanja veze između broja nedelja koje je novorođenče provelo u gestacionom periodu (vreme provedeno u majčinoj utrobi) i telesne mase nakon porođaja. Slučajno izabrani podaci prikazani su tabelom u nastavku:

gestacioni period (nedelja)	34	29	40	36	37	41	38
telesna masa (kg)	1.9	1.4	2.8	3.1	2.7	3.6	2.7

- a) Izračunati koeficijent korelacije. Šta se može zaključiti na osnovu istog? Odrediti i koeficijent determinacije.
- b) Kolika bi bila telesna masa novorođenčeta koje je u gestacionoj fazi provelo 30 nedelja?
- 

Posmatra se broj parkova u gradovima. U jednom regionu u kom se broj gradova procenjuje na 40, prikupljeni su podaci o broju parkova i dati sledećom tabelom:

broj parkova	3	4	7	5	9	8	2
broj gradova	6	4	4	3	1	2	7

- a) Odrediti interval poverenja za srednju vrednost broja parkova po gradu na nivou poverenja od 95% kao i procenu ukupnog broja parkova u posmatranom regionu.
- b) Da li, sa mogućnošću greške od 5%, možemo tvrditi da je broj parkova u prosečnom gradu posmatranog regiona jednak 5?
- 

Laborant sprovodi istraživanje sa ciljem ispitivanja veze između indeksa telesne mase sa jedne i ukupnog holesterola sa druge strane. Slučajno izabrani podaci prikazani su tabelom u nastavku:

indeks telesne mase	21	22	23	25	27	31	32
ukupni holesterol	160	190	165	200	220	210	260

- a) Izračunati koeficijent korelacije. Šta se može zaključiti na osnovu istog? Odrediti i koeficijent determinacije.
- b) Koliki bi bio ukupni holesterol osobe čiji indeks telesne mase iznosi 29?
- 

Sa ciljem ispitivanja uticaja i štetnog dejstva stresa na zdravlje zaposlenih u kompaniji čiji radni zadaci nose sa sobom dosta neizvesnosti, lekar je na slučajan način izabrao 7 radnika starosne dobi od 35 do 55 godina i zabeležio nivoe sistolnog krvnog pritiska (SBP) za svakog od ispitanika ponaosob. Podaci su dati tabelom u nastavku:

starosna dob	37	39	44	47	48	51	52
SBP	133	143	151	143	158	146	168

- a) Izračunati koeficijent korelacije. Šta se može zaključiti na osnovu istog? Odrediti i koeficijent determinacije.
- b) Kolika bi bila procena starosne dobi radnika kojem je izmeren SBP od 160?
- 

Prikupljeni podaci o broju ugostiteljskih objekata (UO) u pešačkoj zoni za svaki od gradova jednog regiona, u kojem je broj gradova 40, prezentovani su sledećom tabelom:

broj UO	7	4	3	5	2	8	9
broj gradova	4	4	6	3	7	2	1

- a) Odrediti interval poverenja za srednju vrednost broja ugostiteljskih objekata po gradu na nivou poverenja od 90% kao i procenu ukupnog broja ove vrste objekata u posmatranom regionu.

- b) Da li, sa mogućnošću greške od 10%, možemo tvrditi da je broj ugostiteljskih objekata u prosečnom gradu posmatranog regiona jednak 6?
- 

Poznata prodavnica sladoleda u Rimu odlučila je da ispita zavisnost temperature (merene u °C) i prihoda od dnevne prodaje ove letnje poslastice. Podaci za vremenski okvir od nedelju dana dati su sledećom tabelom:

temperatura (°C)	17	22	23	19	15	11	14
prihod od prodaje (€)	408	445	544	412	332	185	215

- a) Izračunati koeficijent korelacije. Šta se može zaključiti na osnovu istog? Odrediti i koeficijent determinacije.
- b) Koliki bi bio očekivani prihod dana kojeg je izmerena temperatura od 29 °C?
- 

Prikupljeni podaci o broju ugostiteljskih objekata (UO) u pešačkoj zoni za svaki od gradova jednog regiona, u kojem je broj gradova 40, prezentovani su sledećom tabelom:

broj gradova	4	4	6	3	7	2	1
broj UO	7	4	3	5	2	8	9

- a) Odrediti interval poverenja za srednju vrednost broja ugostiteljskih objekata po gradu na nivou poverenja od 99% kao i procenu ukupnog broja ove vrste objekata u posmatranom regionu.
- b) Da li, sa mogućnošću greške od 1%, možemo tvrditi da je broj ugostiteljskih objekata u prosečnom gradu posmatranog regiona jednak 6?
- 

Prikupljeni podaci o broju kasa u objektima trgovinskog lanca za svaki od gradova jednog regiona, u kojem je broj gradova 40, prezentovani su sledećom tabelom:

broj gradova	4	4	6	3	7	2	1
broj kasa	7	4	3	5	2	8	9

- a) Odrediti interval poverenja za srednju vrednost broja kasa u trgovinskom objektu po gradu na nivou poverenja od 90% kao i procenu ukupnog broja kasa u posmatranom regionu.
- b) Da li, sa mogućnošću greške od 10%, možemo tvrditi da je broj kasa u prosečnom gradu posmatranog regiona jednak 6?

Analitičar zaposlen u marketing timu uspješne kompanije je, za potrebe izveštavanja Upravnog odbora, sa ciljem povećanja prodaje novog proizvoda, odlučio da ispita poznatu vezu između količine sredstava uloženi u reklamu i promociju proizvoda sa jedne i uspešnosti prodaje sa druge strane. Istorijski podaci za 7 slučajno izabranih meseci iz prethodne poslovne godine u vezi sa uložnim sredstvima u promociju kao i uspešnosti prodaje dati su tabelom u nastavku.

promocija [u hiljadama €]	5.5	5.9	6.2	6.4	6.5	6.7	6.8
prodaja [u hiljadama €]	100	115	117	121	120	117	123

- a) Odrediti koeficijent korelacije  $r_{xy}$ . Šta je analitičar zaključio na osnovu istog?
- b) Dati prognozu očekivane prodaje [u hiljadama €] za naredni mesec ako je Upravni odbor kompanije doneo odluku da za potrebe promocije uloži 5700 €.
- 

Analitičar zaposlen u odeljenju za istraživanje i praćenje tržišta u uspešnoj brokerskoj kući je, za potrebe izveštavanja Upravnog odbora ove institucije, prikupio podatke o dnevnom prometu akcija (broju transakcija tipa kupovina-prodaja) u sklopu berzanskog poslovanja kao i broju dana tokom kojih je promet realizovan. Podaci su dati tabelom u nastavku. (*Primeru radi, promet od 0 do 1.000 transakcija registrovan je u 2 dana u posmatranom periodu*).

promet (u hiljadama)	[0, 1)	[1, 2)	[2, 3)	[3, 4)	[4, 5)	[5, 6)	[6, 7)
broj dana	2	5	3	4	6	2	3

- a) Odrediti interval poverenja za srednju vrednost prometa (izraženog u hiljadama) u tom periodu na nivou poverenja od 99%.
- b) Da li, sa mogućnošću greške od 1%, možemo tvrditi da je prosečan dnevni promet u posmatranom periodu bio jednak 5 hiljada?
- 

Sa ciljem povećanja broja novih potencijalnih klijenata, analitičar konsalting kuće odlučio je da ispita vezu između količine sredstava uloženi u reklamne aktivnosti sa jedne i uspešnosti poslovanja (merene brojem klijenata) sa druge strane. Tabelom u nastavku prikazani su istorijski podaci za 7 slučajno izabranih meseci iz prethodne poslovne godine koji prate pomenuta dva obeležja.

broj novih klijenata	6	7	7	9	11	12	14
reklama [u hiljadama €]	3	5	6	8	11	13	15

- a) Odrediti koeficijent korelacije  $r_{xy}$ . Šta možemo zaključiti na osnovu istog?
- b) Ako je Upravni odbor ove kompanije odlučio da narednog meseca privuče 17 novih klijenata, koliko je novčanih sredstava (u hiljadama €) neophodno da izdvoji?
- 

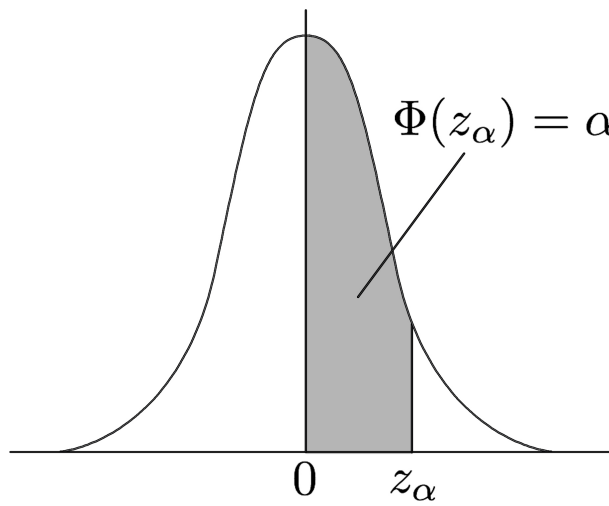
Start-up programerska kuća je, za potrebe evaluacije uspešnosti dosadašnjeg poslovanja, prikupila podatke o broju prodatih aplikacija namenjenih mobilnim platformama kao i broju dana tokom kojih su te prodaje realizovane. Podaci su dati tabelom u nastavku.

br. prodatih app	[0, 10)	[10, 30)	[30, 40)	[40, 60)	[60, 70)	[70, 90)	[90, 100]
broj dana	2	5	3	4	6	2	3

- a) Odrediti interval poverenja za srednju vrednost broja prodatih aplikacija u posmatranom periodu na nivou poverenja od 95%.
- b) Da li, sa mogućnošću greške od 5%, možemo tvrditi da je prosečan broj prodatih aplikacija na dnevnom nivou u posmatranom periodu bio jednak 35?

*A*

*z*-tablica

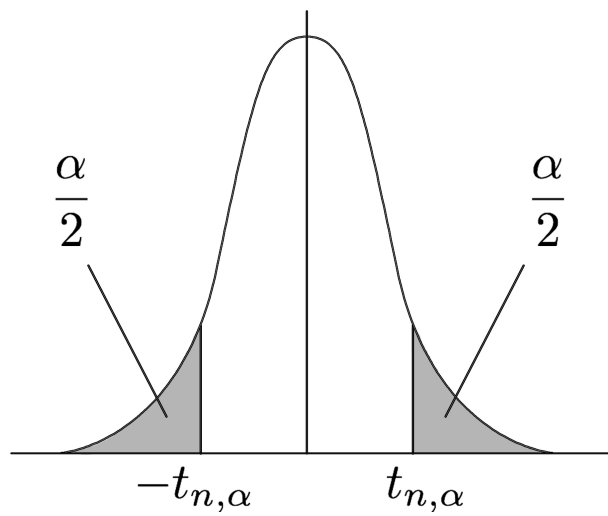


$z$	$\Phi(z)$	$z$	$\Phi(z)$	$z$	$\Phi(z)$	$z$	$\Phi(z)$	$z$	$\Phi(z)$
0.00	0.0000	0.26	0.1026	0.52	0.1985	0.78	0.2823	1.04	0.3508
0.01	0.0040	0.27	0.1064	0.53	0.2019	0.79	0.2852	1.05	0.3531
0.02	0.0080	0.28	0.1103	0.54	0.2054	0.80	0.2881	1.06	0.3554
0.03	0.0120	0.29	0.1141	0.55	0.2088	0.81	0.2910	1.07	0.3577
0.04	0.0160	0.30	0.1179	0.56	0.2123	0.82	0.2939	1.08	0.3599
0.05	0.0199	0.31	0.1217	0.57	0.2157	0.83	0.2967	1.09	0.3621
0.06	0.0239	0.32	0.1255	0.58	0.2190	0.84	0.2995	1.10	0.3643
0.07	0.0279	0.33	0.1293	0.59	0.2224	0.85	0.3023	1.11	0.3665
0.08	0.0319	0.34	0.1331	0.60	0.2257	0.86	0.3051	1.12	0.3686
0.09	0.0359	0.35	0.1368	0.61	0.2291	0.87	0.3078	1.13	0.3708
0.10	0.0398	0.36	0.1406	0.62	0.2324	0.88	0.3106	1.14	0.3729
0.11	0.0438	0.37	0.1443	0.63	0.2357	0.89	0.3133	1.15	0.3749
0.12	0.0478	0.38	0.1480	0.64	0.2389	0.90	0.3159	1.16	0.3770
0.13	0.0517	0.39	0.1517	0.65	0.2422	0.91	0.3186	1.17	0.3790
0.14	0.0557	0.40	0.1554	0.66	0.2454	0.92	0.3212	1.18	0.3810
0.15	0.0596	0.41	0.1591	0.67	0.2486	0.93	0.3238	1.19	0.3830
0.16	0.0636	0.42	0.1628	0.68	0.2517	0.94	0.3264	1.20	0.3849
0.17	0.0675	0.43	0.1664	0.69	0.2549	0.95	0.3289	1.21	0.3869
0.18	0.0714	0.44	0.1700	0.70	0.2580	0.96	0.3315	1.22	0.3888
0.19	0.0753	0.45	0.1736	0.71	0.2611	0.97	0.3340	1.23	0.3907
0.20	0.0793	0.46	0.1772	0.72	0.2642	0.98	0.3365	1.24	0.3925
0.21	0.0832	0.47	0.1808	0.73	0.2673	0.99	0.3389	1.25	0.3944
0.22	0.0871	0.48	0.1844	0.74	0.2704	1.00	0.3413	1.26	0.3962
0.23	0.0910	0.49	0.1879	0.75	0.2734	1.01	0.3438	1.27	0.3980
0.24	0.0948	0.50	0.1915	0.76	0.2764	1.02	0.3461	1.28	0.3997
0.25	0.0987	0.51	0.1950	0.77	0.2794	1.03	0.3485	1.29	0.4015



$z$	$\Phi(z)$	$z$	$\Phi(z)$	$z$	$\Phi(z)$	$z$	$\Phi(z)$	$z$	$\Phi(z)$	$z$	$\Phi(z)$
1.56	0.4406	1.82	0.4656	2.08	0.4812	2.34	0.4904	2.60	0.4953	2.86	0.4979
1.57	0.4418	1.83	0.4664	2.09	0.4817	2.35	0.4906	2.61	0.4955	2.87	0.4979
1.58	0.4429	1.84	0.4671	2.10	0.4821	2.36	0.4909	2.62	0.4956	2.88	0.4980
1.59	0.4441	1.85	0.4678	2.11	0.4826	2.37	0.4911	2.63	0.4957	2.89	0.4981
1.60	0.4452	1.86	0.4686	2.12	0.4830	2.38	0.4913	2.64	0.4959	2.90	0.4981
1.61	0.4463	1.87	0.4693	2.13	0.4834	2.39	0.4916	2.65	0.4960	2.91	0.4982
1.62	0.4474	1.88	0.4699	2.14	0.4838	2.40	0.4918	2.66	0.4961	2.92	0.4982
1.63	0.4484	1.89	0.4706	2.15	0.4842	2.41	0.4920	2.67	0.4962	2.93	0.4983
1.64	0.4495	1.90	0.4713	2.16	0.4846	2.42	0.4922	2.68	0.4963	2.94	0.4984
1.65	0.4505	1.91	0.4719	2.17	0.4850	2.43	0.4925	2.69	0.4964	2.95	0.4984
1.66	0.4515	1.92	0.4726	2.18	0.4854	2.44	0.4927	2.70	0.4965	2.96	0.4985
1.67	0.4525	1.93	0.4732	2.19	0.4857	2.45	0.4929	2.71	0.4966	2.97	0.4985
1.68	0.4535	1.94	0.4738	2.20	0.4861	2.46	0.4931	2.72	0.4967	2.98	0.4986
1.69	0.4545	1.95	0.4744	2.21	0.4864	2.47	0.4932	2.73	0.4968	2.99	0.4986
1.70	0.4554	1.96	0.4750	2.22	0.4868	2.48	0.4934	2.74	0.4969	3.00	0.4987
1.71	0.4564	1.97	0.4756	2.23	0.4871	2.49	0.4936	2.75	0.4970	3.01	0.4987
1.72	0.4573	1.98	0.4761	2.24	0.4875	2.50	0.4938	2.76	0.4971	3.02	0.4987
1.73	0.4582	1.99	0.4767	2.25	0.4878	2.51	0.4940	2.77	0.4972	3.03	0.4988
1.74	0.4591	2.00	0.4772	2.26	0.4881	2.52	0.4941	2.78	0.4973	3.04	0.4988
1.75	0.4599	2.01	0.4778	2.27	0.4884	2.53	0.4943	2.79	0.4974	3.05	0.4989
1.76	0.4608	2.02	0.4783	2.28	0.4887	2.54	0.4945	2.80	0.4974	3.06	0.4989
1.77	0.4616	2.03	0.4788	2.29	0.4890	2.55	0.4946	2.81	0.4975	3.07	0.4989
1.78	0.4625	2.04	0.4793	2.30	0.4893	2.56	0.4948	2.82	0.4976	3.08	0.4990
1.79	0.4633	2.05	0.4798	2.31	0.4896	2.57	0.4949	2.83	0.4977	3.09	0.4990
1.80	0.4641	2.06	0.4803	2.32	0.4898	2.58	0.4951	2.84	0.4977	3.10	0.4990
1.81	0.4649	2.07	0.4808	2.33	0.4901	2.59	0.4952	2.85	0.4978	3.11	0.4991

$z$	$\Phi(z)$	$z$	$\Phi(z)$	$z$	$\Phi(z)$	$z$	$\Phi(z)$
3.12	0.4991	3.38	0.4996	3.64	0.4999	3.90	0.5000
3.13	0.4991	3.39	0.4997	3.65	0.4999	3.91	0.5000
3.14	0.4992	3.40	0.4997	3.66	0.4999	3.92	0.5000
3.15	0.4992	3.41	0.4997	3.67	0.4999	3.93	0.5000
3.16	0.4992	3.42	0.4997	3.68	0.4999	3.94	0.5000
3.17	0.4992	3.43	0.4997	3.69	0.4999	3.95	0.5000
3.18	0.4993	3.44	0.4997	3.70	0.4999	3.96	0.5000
3.19	0.4993	3.45	0.4997	3.71	0.4999	3.97	0.5000
3.20	0.4993	3.46	0.4997	3.72	0.4999	3.98	0.5000
3.21	0.4993	3.47	0.4997	3.73	0.4999	3.99	0.5000
3.22	0.4994	3.48	0.4997	3.74	0.4999		
3.23	0.4994	3.49	0.4998	3.75	0.4999		
3.24	0.4994	3.50	0.4998	3.76	0.4999		
3.25	0.4994	3.51	0.4998	3.77	0.4999		
3.26	0.4994	3.52	0.4998	3.78	0.4999		
3.27	0.4995	3.53	0.4998	3.79	0.4999		
3.28	0.4995	3.54	0.4998	3.80	0.4999		
3.29	0.4995	3.55	0.4998	3.81	0.4999		
3.30	0.4995	3.56	0.4998	3.82	0.4999		
3.31	0.4995	3.57	0.4998	3.83	0.4999		
3.32	0.4995	3.58	0.4998	3.84	0.4999		
3.33	0.4996	3.59	0.4998	3.85	0.4999		
3.34	0.4996	3.60	0.4998	3.86	0.4999		
3.35	0.4996	3.61	0.4998	3.87	0.4999		
3.36	0.4996	3.62	0.4999	3.88	0.4999		
3.37	0.4996	3.63	0.4999	3.89	0.4999		



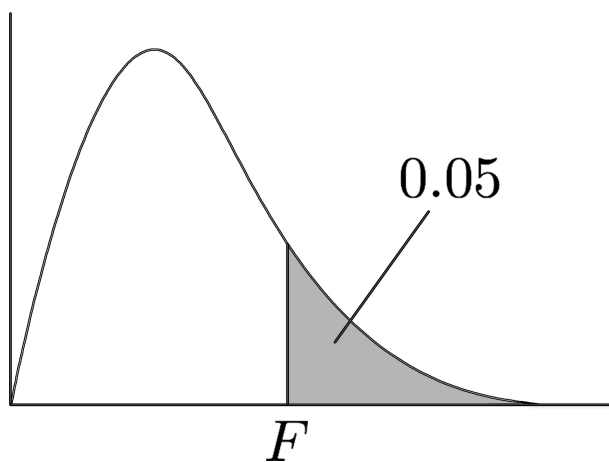
$n \setminus \alpha$	0.500	0.400	0.300	0.200	0.100	0.050	0.025	0.010	0.005	0.001
1	1.000	1.376	1.963	3.078	6.314	12.706	25.452	63.656	127.321	636.578
2	0.816	1.061	1.386	1.886	2.920	4.303	6.205	9.925	14.089	31.600
3	0.765	0.978	1.250	1.638	2.353	3.182	4.177	5.841	7.453	12.924
4	0.741	0.941	1.190	1.533	2.132	2.776	3.495	4.604	5.598	8.610
5	0.727	0.920	1.156	1.476	2.015	2.571	3.163	4.032	4.773	6.869
6	0.718	0.906	1.134	1.440	1.943	2.447	2.969	3.707	4.317	5.959
7	0.711	0.896	1.119	1.415	1.895	2.365	2.841	3.499	4.029	5.408
8	0.706	0.889	1.108	1.397	1.860	2.306	2.752	3.355	3.833	5.041
9	0.703	0.883	1.100	1.383	1.833	2.262	2.685	3.250	3.690	4.781
10	0.700	0.879	1.093	1.372	1.812	2.228	2.634	3.169	3.581	4.587
11	0.697	0.876	1.088	1.363	1.796	2.201	2.593	3.106	3.497	4.437
12	0.695	0.873	1.083	1.356	1.782	2.179	2.560	3.055	3.428	4.318
13	0.694	0.870	1.079	1.350	1.771	2.160	2.533	3.012	3.372	4.221
14	0.692	0.868	1.076	1.345	1.761	2.145	2.510	2.977	3.326	4.140
15	0.691	0.866	1.074	1.341	1.753	2.131	2.490	2.947	3.286	4.073
16	0.690	0.865	1.071	1.337	1.746	2.120	2.473	2.921	3.252	4.015
17	0.689	0.863	1.069	1.333	1.740	2.110	2.458	2.898	3.222	3.965
18	0.688	0.862	1.067	1.330	1.734	2.101	2.445	2.878	3.197	3.922
19	0.688	0.861	1.066	1.328	1.729	2.093	2.433	2.861	3.174	3.883
20	0.687	0.860	1.064	1.325	1.725	2.086	2.423	2.845	3.153	3.850

$n \setminus \alpha$	0.500	0.400	0.300	0.200	0.100	0.050	0.025	0.010	0.005	0.001
21	0.686	0.859	1.063	1.323	1.721	2.080	2.414	2.831	3.135	3.819
22	0.686	0.858	1.061	1.321	1.717	2.074	2.405	2.819	3.119	3.792
23	0.685	0.858	1.060	1.319	1.714	2.069	2.398	2.807	3.104	3.768
24	0.685	0.857	1.059	1.318	1.711	2.064	2.391	2.797	3.091	3.745
25	0.684	0.856	1.058	1.316	1.708	2.060	2.385	2.787	3.078	3.725
26	0.684	0.856	1.058	1.315	1.706	2.056	2.379	2.779	3.067	3.707
27	0.684	0.855	1.057	1.314	1.703	2.052	2.373	2.771	3.057	3.689
28	0.683	0.855	1.056	1.313	1.701	2.048	2.368	2.763	3.047	3.674
29	0.683	0.854	1.055	1.311	1.699	2.045	2.364	2.756	3.038	3.660
30	0.683	0.854	1.055	1.310	1.697	2.042	2.360	2.750	3.030	3.646
31	0.682	0.853	1.054	1.309	1.696	2.040	2.356	2.744	3.022	3.633
32	0.682	0.853	1.054	1.309	1.694	2.037	2.352	2.738	3.015	3.622
33	0.682	0.853	1.053	1.308	1.692	2.035	2.348	2.733	3.008	3.611
34	0.682	0.852	1.052	1.307	1.691	2.032	2.345	2.728	3.002	3.601
35	0.682	0.852	1.052	1.306	1.690	2.030	2.342	2.724	2.996	3.591
36	0.681	0.852	1.052	1.306	1.688	2.028	2.339	2.719	2.990	3.582
37	0.681	0.851	1.051	1.305	1.687	2.026	2.336	2.715	2.985	3.574
38	0.681	0.851	1.051	1.304	1.686	2.024	2.334	2.712	2.980	3.566
39	0.681	0.851	1.050	1.304	1.685	2.023	2.331	2.708	2.976	3.558
40	0.681	0.851	1.050	1.303	1.684	2.021	2.329	2.704	2.971	3.551

$n \setminus \alpha$	0.500	0.400	0.300	0.200	0.100	0.050	0.025	0.010	0.005	0.001
41	0.681	0.850	1.050	1.303	1.683	2.020	2.327	2.701	2.967	3.544
42	0.680	0.850	1.049	1.302	1.682	2.018	2.325	2.698	2.963	3.538
43	0.680	0.850	1.049	1.302	1.681	2.017	2.323	2.695	2.959	3.532
44	0.680	0.850	1.049	1.301	1.680	2.015	2.321	2.692	2.956	3.526
45	0.680	0.850	1.049	1.301	1.679	2.014	2.319	2.690	2.952	3.520
46	0.680	0.850	1.048	1.300	1.679	2.013	2.317	2.687	2.949	3.515
47	0.680	0.849	1.048	1.300	1.678	2.012	2.315	2.685	2.946	3.510
48	0.680	0.849	1.048	1.299	1.677	2.011	2.314	2.682	2.943	3.505
49	0.680	0.849	1.048	1.299	1.677	2.010	2.312	2.680	2.940	3.500
50	0.679	0.849	1.047	1.299	1.676	2.009	2.311	2.678	2.937	3.496
55	0.679	0.848	1.046	1.297	1.673	2.004	2.304	2.668	2.925	3.476
60	0.679	0.848	1.045	1.296	1.671	2.000	2.299	2.660	2.915	3.460
70	0.678	0.847	1.044	1.294	1.667	1.994	2.291	2.648	2.899	3.435
80	0.678	0.846	1.043	1.292	1.664	1.990	2.284	2.639	2.887	3.416
90	0.677	0.846	1.042	1.291	1.662	1.987	2.280	2.632	2.878	3.402
100	0.677	0.845	1.042	1.290	1.660	1.984	2.276	2.626	2.871	3.390
120	0.677	0.845	1.041	1.289	1.658	1.980	2.270	2.617	2.860	3.373

C

*F*-tablica,  $\alpha = 0.05$

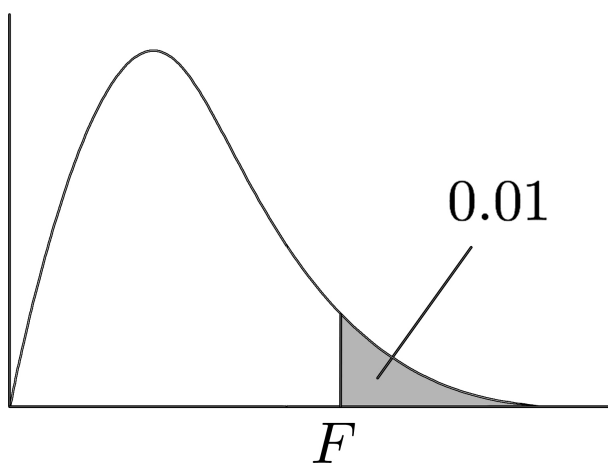


$n_1 \setminus n_2$	1	2	3	4	5	6	7	8	12	24
1	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	243.91	249.05
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.41	19.45
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.74	8.64
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	5.91	5.77
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.68	4.53
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.00	3.84
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.57	3.41
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.28	3.12
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.07	2.90
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	2.91	2.74
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.79	2.61
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.69	2.51
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.60	2.42
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.53	2.35
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.48	2.29
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.42	2.24
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.38	2.19
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.34	2.15
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.31	2.11
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.28	2.08
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.00	1.79
80	3.96	3.11	2.72	2.49	2.33	2.21	2.13	2.06	1.88	1.65



$\mathcal{D}$

$F$ -tablica,  $\alpha = 0.01$



$n_1 \setminus n_2$	1	2	3	4	5	6	7	8	12	24
1	4052.18	4999.50	5403.35	5624.58	5763.65	5858.99	5928.36	5981.07	6106.32	6234.63
2	98.50	99.00	99.17	99.25	99.30	99.33	99.36	99.37	99.42	99.46
3	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.05	26.60
4	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.37	13.93
5	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	9.89	9.47
6	13.75	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.72	7.31
7	12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.47	6.07
8	11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.67	5.28
9	10.56	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.11	4.73
10	10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.71	4.33
11	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.40	4.02
12	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.16	3.78
13	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	3.96	3.59
14	8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14	3.80	3.43
15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.67	3.29
16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.55	3.18
17	8.40	6.11	5.18	4.67	4.34	4.10	3.93	3.79	3.46	3.08
18	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.37	3.00
19	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.30	2.92
20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.23	2.86
40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.66	2.29
80	6.96	4.88	4.04	3.56	3.26	3.04	2.87	2.74	2.42	2.03

# Literatura

- [1] Agresti, A.: *Statistics. The Art and Science of Learning from Data*, Pearson Prentice Hall, New Jersey, 2006.
- [2] Anderson, D.R., Sweeney, D.J., Williams, T.A.: *Essentials of Statistics for Business and Economics*, 8th Edition, South-Western/Thomson Learning, 2002.
- [3] Arsham, H.: *Statistical Thinking for Managerial Decisions*, <http://home.ubalt.edu/ntsbarsh/Business-stat/opre504.html>
- [4] Chase, W., Brown, F.: *General Statistics*, 2nd Edition, John Wiley and Sons, New York, 1992.
- [5] Cvetković, Lj.: *Matematika sa statistikom*, PMF Institut za matematiku, Novi Sad, 1996.
- [6] Cvetković, Lj., Lozanov Crvenković, Z.: *Verovatnoća i statistika*, Zbirka zadataka, Futura publikacije, Novi Sad, 2002.
- [7] Cvetković, Lj.: *Poslovna statistika*, Futura publikacije, Novi Sad, 2006.
- [8] Easton, V.J., McColl, J.H.: *Statistics Glossary*, <http://www.cas.lancs.ac.uk/glossary/v1.1/main.html>
- [9] Johnson, R. A., Bhattacharya, G. K.: *Statistics : Principles and Methods*, 2nd Edition, John Wiley and Sons, New York, 1992.
- [10] Stojković, M.: *Statistika*, Ekonomski fakultet, Subotica, 2001.
- [11] Sullivan III, M.: *Statistics - Informed Decisions Using Data*, 2nd Edition, Pearson Prentice Hall, New Jersey, 2004.