



UNIVERZITET U NOVOM SADU  
POLJOPRIVREDNI FAKULTET

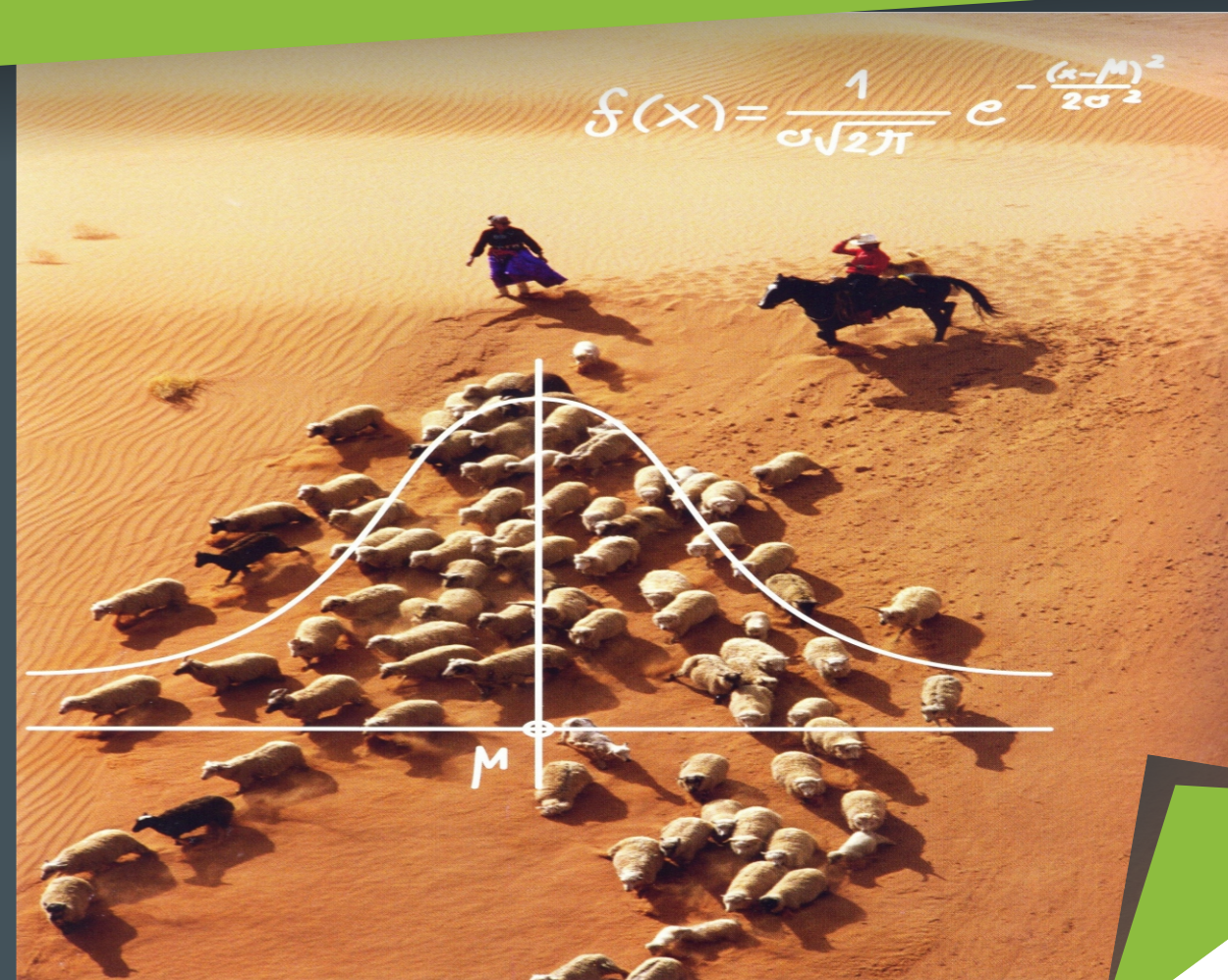
PRAKTIKUM IZ STATISTIKE



# PRAKTIKUM IZ STATISTIKE

Prof. dr Beba Mutavdžić  
MMgt Tihomir Novaković  
Mast.aekon. Dragana Tekić

Prof. dr Beba Mutavdžić  
MMgt Tihomir Novaković  
Mast.aekon. Dragana Tekić



# **PRAKTIKUM IZ STATISTIKE**

**(za smer Veterinarska medicina)**

Beba Mutavdžić, Tihomir Novaković, Dragana Tekić

## **EDICIJA PRAKTIKUM**

### **Osnivač i izdavač edicije**

Univerzitet u Novom Sadu

Poljoprivredni fakultet

Trg Dositeja Obradovića br.8, Novi Sad

### **Godina osnivanja**

1954.

### **Glavni i odgovorni urednik edicije**

Dr Nedeljko Tica, redovni profesor

Dekan Poljoprivrednog fakulteta u Novom Sadu

### **Članovi komisije za izdavačku delatnost**

Dr Ljiljana Nešić, redovni profesor, predsednik

Dr Zorica Srđević, redovni profesor, član

Dr Nada Plavša, vanredni profesor, član

Dr Branislav Vlahović, redovni profesor, član

CIP –

ISBN: 978-86-7520-504-3

**Autori:**

Dr Beba Mutavdžić

MMgt Tihomir Novaković

MSc Dragana Tekić

**Glavni i odgovorni urednik edicije**

Dr Nedeljko Tica, redovni profesor

Dekan Poljoprivrednog fakulteta u Novom Sadu

**Recenzenti**

Dr Otilija Sedlak, redovni profesor

Ekonomski fakultet Subotica, Univerzitet u Novom Sadu

Mr Emilija Nikolić-Đorić

Poljoprivredni fakultet Novi Sad, Univerzitet u Novom Sadu

**Izdavač**

Poljoprivredni fakultet Novi Sad, Univerzitet u Novom Sadu

Zabranjeno preštampanje i fotokopiranje. Sva prava zadržava izdavač.

Štampa:

Štampanje odobrila Komisija za izdavačku delatnost i Naučno-nastavno veće

Poljoprivrednog fakulteta u Novom Sadu.

Tiraž: 20 primeraka

Mesto i godina štampanja: Novi Sad, 2020



## SADRŽAJ

PREDGOVOR .....	1
FORMIRANJE DISTRIBUCIJA FREKVENCIJA .....	2
POKAZATELJI CENTRALNE TENDENCIJE .....	8
POKAZATELJI VARIJABILITETA .....	15
POKAZATELJI OBLIKA DISTRIBUCIJE .....	22
NORMALNA RASPODELA .....	28
OCENE NA OSNOVU UZORKA .....	33
TESTIRANJE STATISTIČKIH HIPOTEZA .....	39
REGRESIONA I KORELACIONA ANALIZA .....	55
LITERATURA .....	62

## PREDGOVOR

Praktikum za predmet „Statistika”, koji se proučava na prvoj godini smjera Veterinarska medicina Poljoprivrednog fakulteta, Univerziteta u Novom Sadu obuhvata materiju koja je u skladu sa aktuelnim akreditovanim programom za navedeni predmet i navedeni smer. Praktikum se sastoji iz osam poglavlja, koja čine strukturnu i sadržajnu celinu:

- Prvo poglavlje - *Formiranje distribucije frekvencija*
- Drugo poglavlje - *Pokazatelji centralne tendencije*
- Treće poglavlje – *Pokazatelji varijabiliteta*
- Četvrto poglavlje – *Pokazatelji oblika distribucije*
- Peto poglavlje – *Normalna raspodela*
- Šesto poglavlje – *Ocene na osnovu uzorka*
- Sedmo poglavlje – *Testiranje statističkih hipoteza*
- Osmo poglavlje – *Regresiona i korelaciona analiza*

Svako poglavlje obuhvata jedan ili više radnih zadataka, kao i primere za vežbanje. Formule u primerima date su samo kao radni obrasci.

Autori se nadaju da će ovaj praktikum olakšati studentima upotrebu osnovnih statističkih metoda u rešavanju problema koji su u domenu poljoprivrednih i bioloških nauka, odnosno konkretno problema iz oblasti veterinarske medicine. Ideja autora je da se studenti upoznaju sa deskriptivnim metodama, kao i metodama analize rezultata ogleđa.

Zahvaljujemo se svima koji su na direktan ili indirektan način pomogli izradu ove knjige, a naročito recenzentima: prof. dr Otiliji Sedlak i Mr Emiliji Nikolić-Đorić na korisnim sugestijama.

Novi Sad,  
2020. godine

AUTORI



## FORMIRANJE DISTRIBUCIJA FREKVENCIJA

### *Prekidno obeležje*

**Primer 1.** Na osnovu broja svinja kod 24 poljoprivredna gazdinstva formirati neintervalnu i intervalnu distribuciju frekvencija:

12 5 6 7 5 8 10 9 5 12 9 8  
8 7 7 10 11 8 10 10 11 9 12 8

Tačkasti dijagram:

Neintervalna serija:

Broj svinja ( $X_i$ )	Broj gazdinstava ( $f_i$ )
<b>Ukupno</b>	

Intervalna serija:


### *Neprekidno obeležje*

**Primer 2.** Dnevna mlečnost (lit.) 20 ispitivanih krava je data u sledećoj seriji:

10,1 12,0 17,6 14,5 18,4 12,7 16,2 15,8 16,4 13,0  
19,9 12,1 13,5 15,1 17,9 13,3 17,0 19,3 16,5 14,5

- Formirati intervalnu distribuciju frekvencija ( $i=2$ );
- izračunati relativne frekvencije (strukturu);
- formirati kumulativnu distribuciju frekvencija
- formirati kumulaciju strukture
- podatke grafički predstaviti histogramom i poligonom

Dijagram stablo-list:

Radna tabela:

Dnevna mlečnost (lit.) ( $X_i$ )	Broj krava ( $f_i$ )	Relativna frekvencija (struktura)	Kumulativ		Kumulativ strukture	
			ispod	iznad	ispod	iznad
<b>Ukupno</b>						

Grafički prikaz (histogram i poligon):

### Zadaci za vežbu:

1. Na osnovu podataka o broju teladi kod 30 poljoprivrednih gazdinstava formirati neintervalnu distribuciju frekvencija i grafički je predstaviti primenom poligona:

7	8	10	8	11	10	10	12	5	10	8	7	11	11	6
8	10	7	11	5	10	10	11	12	8	12	11	9	8	5

2. Dužina butne kosti (cm) merena je kod 40 pasa. Dobijene su sledeće vrednosti:

17,6	14,1	13,9	13,1	16,8	14,7	13,2	15,7	13,5	15,4
14,3	13,9	13,2	13,6	15,3	13,2	16,6	14,7	18,1	18,4
15,6	15,5	15,4	14,7	17,6	14,4	19,8	13,6	14,7	14,7
13,9	17,8	18,2	13,8	13,1	13,5	13,6	14,9	14,7	14,1

Formirati dijagram stablo list i na osnovu njega intervalnu seriju distribucije frekvencija ako je  $i=0,5$ . Utvrditi strukturu raspodele, kumulativ i kumulativ strukture. Dati komentar za kumulativne vrednosti najfrekventnije vrednosti obeležja. Grafički predstaviti podatke primenom histograma.

3. Dati su podaci o plućnoj ventilaciji (lit/min) 25 odraslih ovaca:

8,3	8,0	9,9	6,1	5,5
10,3	6,5	7,6	7,6	7,6
6,9	10,3	7,8	7,3	8,9
10,1	7,6	9,1	8,3	4,8
10,2	6,5	9,1	7,0	11,9

Formirati dijagram stablo list i intervalnu distribuciju frekvencija ako je dužina grupnog interval 1,5 (lit/min). Podatke grafički predstaviti.





## POKAZATELJI CENTRALNE TENDENCIJE

Pokazatelji centralne tendencije predstavljaju vrednosti koje kvantifikuju tendenciju podataka u seriji prema njihovom "centru", odnosno sredini.

U pokazatelje centralne tendencije ubrajaju se:

- Aritmetička sredina
- Geometrijska sredina
- Harmonijska sredina
- Medijana
- Modus
- Kvantili

### *Aritmetička sredina*

Aritmetička sredina predstavlja izračunatu prosečnu vrednost posmatrane promenljive i računa se kao količnik zbira vrednosti obeležja i ukupnog broja podataka. Razlikuje se izračunavanje proste (za negrupisane podatke) i ponderisane (za grupisane podatke) aritmetičke sredine.

**Primer 1.** Izračunati prosečnu vrednost dnevnog prirasta svinja (kg) na osnovu sledeće serije podataka:

<b>Prirast (kg) (<math>X_i</math>)</b>	0,80	0,74	0,83	0,82	0,66	0,78
--	------	------	------	------	------	------

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n} =$$

**Primer 2.** Utvrditi prosečanu dnevnu mlečnost (lit.) 20 ispitivanih krava na osnovu sledeće serije podataka:

<b>Dnevna mlečnost (lit.)</b> <b>(<math>X_i</math>)</b>	<b>Broj krava</b> <b>(<math>f_i</math>)</b>	
11	2	
13	4	
15	6	
17	5	
19	3	
<b>Ukupno</b>	<b>20</b>	

$$\bar{X} = \frac{\sum_{i=1}^k X_i f_i}{\sum_{i=1}^k f_i} =$$

**Primer 3.** Utvrditi prosečan broj teladi po radnoj organizaciji na osnovu sledeće serije podataka:

Broj teladi	Poljoprivredne radne organizacije		
30-39	3		
40-49	4		
50-59	4		
60-69	6		
70-79	3		
80-89	3		
90-99	2		
<b>Ukupno</b>			

### *Medijana*

Medijana je pozicioni pokazatelj centralne tendencije i predstavlja vrednost obeležja koja uređenu statističku seriju deli na dva jednaka dela. Za neintervalne statističke serije potrebno je prethodno utvrditi da li je ukupan broj podataka paran ili neparan broj. Kod intervalnih serija koristi se korigovana formula.

**Primer 1.** Utvrditi medijalnu vrednost za sledeće serije podataka:

<i>X<sub>i</sub></i>	1230	1320	1200	1500	1600
<b>Uređena serija</b>					

$$M_e = X_{\frac{n}{2}+1} =$$

<i>X<sub>i</sub></i>	2110	1900	2170	1980	2100	2000
<b>Uređena serija</b>						

$$M_e = \frac{X_{\frac{n}{2}} + X_{\frac{n}{2}+1}}{2} =$$



**Primer 2.** Utvrditi medijalnu vrednost za sledeću seriju:

Masa jagnjadi ( $X_i$ )	Broj jagnjadi ( $f_i$ )	Kumulativ
3,9	8	
4,2	12	
4,3	20	
4,5	15	
4,6	5	
<b>Ukupno</b>	<b>60</b>	

**Primer 3.** Utvrditi medijalnu vrednost (primenom korigovane formule) za sledeću seriju:

Dužina trupa ( $X_i$ )	Broj goveda ( $f_i$ )	Kumulativ
148,1-152	10	
152,1-156	19	
156,1-160	23	
160,1-164	34	
164,1-168	34	
168,1-172	28	
172,1-176	13	
176,1-180	3	
<b>Ukupno</b>	<b>164</b>	

$$M_e = L + \left( \frac{\frac{n}{2} - F_{M_e-1}}{f_{M_e}} \right) \times i =$$

### Modus

Modus je najučestalija vrednost obeležja u statističkoj seriji. Prema broju modusa statističke serije mogu biti unimodalne, bimodalne ili mogu imati tri modusa.

**Primer 1.** Utvrditi modus za sledeću seriju podataka:

<b><math>X_i</math></b>	75	45	50	70	55	65	60	65	70
<b>Uređena serija</b>									

$$M_0^1 =$$

$$M_0^2 =$$

**Primer 2.** Utvrditi najčešći broj larvi skočibuba po 1 m<sup>2</sup> na posmatranom lokalitetu:

<b>Prosečan broj larvi po 1 m<sup>2</sup> (<math>X_i</math>)</b>	<b>Broj parcela (<math>f_i</math>)</b>
4,5	10
7,5	24
10,5	36
13,5	90

$$M_0 =$$

**Primer 3.** Utvrditi modus (primenom korigovane formule) za sledeću seriju podataka:

<b>Prinos kukuruza (t/ha)</b>	<b>Površina (ha)</b>
4,01-5,00	3
5,01-6,00	3
6,01-7,00	4
7,01-8,00	5
8,01-9,00	6
9,01-10,00	3
$\Sigma$	<b>24</b>

$$M_0 = L + \left( \frac{d_1}{d_1 + d_2} \right) \times i =$$

$$d_1 =$$

$$d_2 =$$

**Zadaci za vežbu:**

1. Odrediti aritmetičku sredinu, modus i medijanu za sledeće vrednosti numeričkog obeležja:

2 3 4 13 1 6 6 7 6 8 11

2. Na osnovu podataka uzorka na jednoj farmi koja ima 250 krmača utvrditi aritmetičku sredinu, modus i medijanu.

<b>Broj prasadi</b>	<b>Broj krmača</b>
5	1
6	0
7	2
8	3
9	3
10	10
11	8
12	5
13	3
14	2

3. Za podatke koji se odnose na masu teladi (kg) na jednoj farmi izračunati aritmetičku sredinu, modus i medijanu.

<b>Masa teladi (kg)</b>	<b>Broj teladi</b>
60,1-70,0	2
70,1-80,0	6
80,1-90,0	11
90,1-100,0	17
100,1-110,0	26
110,1-120,0	19
120,1-130,0	12
130,1-140,0	6
140,1-150,0	1





## POKAZATELJI VARIJABILITETA

Pokazatelji varijabiliteta predstavljaju meru disperzije vrednosti obeležja u okviru jedne statističke serije. Razlikuju se apsolutni i relativni pokazatelji varijabiliteta.

Apsolutni pokazatelji varijabiliteta su:

- Interval varijacije
- Srednje apsolutno odstupanje
- Standardna devijacija
- Varijansa

Relativni pokazatelji varijabiliteta su:

- Koeficijent varijacije
- Standardizovano (normalizovano) odstupanje

### *Interval varijacije*

Interval varijacije predstavlja razliku između ekstremnih vrednosti obeležja u nekoj statističkoj seriji.

**Primer 1.** Na osnovu podataka iz uzorka ustanovljene su sledeće težine proizvoda (kg):

<b>Težina proizvoda (kg) (<math>X_i</math>)</b>	4,8	4,9	5,9	5,2	4,9	5,0	5,4	5,3	5,1
<b>Uređena serija</b>									

$$I = X_{max} - X_{min} =$$

### *Srednje apsolutno odstupanje*

Srednje apsolutno odstupanje predstavlja količnik zbira apsolutnih vrednosti odstupanja pojedinačnih vrednosti obeležja od njihovog proseka i ukupnog broja podataka. Izražava se u jedinicama mere obeležja.

**Primer 2.** Izračunati srednje apsolutno odstupanje za seriju podataka koja se odnosi na dužinu polutki svinja (cm):

<b>Dužina polutki (cm) (<math>X_i</math>)</b>		
80		
89		
93		
100		
105		
112		
115		
$\Sigma$		

$$SO = \frac{\sum_{i=1}^n |X_i - \bar{X}|}{n} =$$

**Primer 3.** Izračunati srednje apsolutno odstupanje za sledeću seriju podataka koja se odnosi na mlečnost krava:

Mlečnost po kravi (lit.) ( $X_i$ )	Broj krava ( $f_i$ )			
8	40			
10	8			
12	25			
16	5			
20	25			
$\Sigma$				

$$SO = \frac{\sum_{i=1}^k f_i |X_i - \bar{X}|}{\sum_{i=1}^k f_i} =$$

### *Standardna devijacija i varijansa*

Standardna devijacija je kvadratni koren iz sredine kvadrata odstupanja vrednosti obeležja od aritmetičke sredine. Vrednost standardne devijacije pokazuje koliko su udaljene grupisane vrednosti obeležja od aritmetičke sredine.

**Primer 4.** Izračunati standardnu devijaciju za sledeću seriju podataka koja se odnosi na masu prasića (kg):

Masa prasića (kg) ( $X_i$ )			
19			
20			
22			
23			
25			
26			
28			
$\Sigma$			

$$I \text{ na\u0107in: } \sigma = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n}} =$$

$$II \text{ na\u0107in: } \sigma = \sqrt{\frac{\sum_{i=1}^n X_i^2 - \frac{(\sum_{i=1}^n X_i)^2}{n}}{n}} =$$

**Primer 5.** Izra\u010dunati standardnu devijaciju za slede\u0107u seriju podataka koja se odnosi na potro\u0161nju mleka:

Potro\u0161nja mleka (lit.) ( $X_i$ )	Broj doma\u0107instava ( $f_i$ )					
12,0-19,9	12					
20,0-27,9	23					
28,0-35,9	85					
36,0-43,9	55					
44,0-51,9	25					
$\Sigma$						

$$I \text{ na\u0107in: } \sigma = \sqrt{\frac{\sum_{i=1}^k f_i (X_i - \bar{X})^2}{\sum_{i=1}^k f_i}} =$$

$$II \text{ na\u0107in: } \sigma = \sqrt{\frac{\sum_{i=1}^k X_i^2 f_i - \frac{(\sum_{i=1}^k X_i f_i)^2}{\sum_{i=1}^k f_i}}{\sum_{i=1}^k f_i}} =$$



### **Koeficijent varijacije**

Koeficijent varijacije je relativni pokazatelj varijabiliteta koji svoju primenu pronalazi prilikom poređenja varijabiliteta pojava koje su izražene u različitim jedinicama mere. Izražava se u procentima

**Primer 6.** U jednom pasilištu prosečan broj prasadi po leglu je 12 komada sa standardnom devijacijom 2 komada. Prosečna telesna masa prasadi iznosi 8 kg dok je varijansa 2,56 kg<sup>2</sup>. Prosečna telesna masa ima \_\_\_\_\_ varijabilitet u poređenju sa prosečnim brojem prasadi po leglu. Obrazložiti.

$$V_1 = \frac{\sigma_1}{\bar{X}_1} \times 100 (\%) =$$

$$V_2 = \frac{\sigma_2}{\bar{X}_2} \times 100 (\%) =$$

### **Standardizovano (normalizovano) odstupanje**

Standardizovano odstupanje je mera udaljenosti pojedinih vrednosti obeležja od aritmetičke sredine iskazana u odnosu na standardnu devijaciju. Standardizovano odstupanje je takođe relativni pokazatelj disperzije obeležja i nema jedinicu mere.

**Primer 7.** Utvrditi standardizovano odstupanje najučestalije vrednosti obeležja za seriju podataka koja se odnosi na dnevnu mlečnost (lit.) 20 ispitivanih krava.

<b>Dnevna mlečnost (lit.) (<math>X_i</math>)</b>	<b>Broj krava (<math>f_i</math>)</b>			
10,1-12,0	2			
12,1-14,0	4			
14,1-16,0	6			
16,1-18,0	5			
18,1-20,0	3			
$\Sigma$	<b>20</b>			

$$Z_i = \frac{X_i - \bar{X}}{\sigma} =$$

**Zadaci za vežbu:**

1. Dati su podaci koji se odnose na prirast pilića starosti 3 nedelje.

<b>Prirast (g)</b>
392
402
385
440
415
397
400

Izračunati apsolutne pokazatelje varijabiliteta.

2. Dati su podaci koji se odnose na sadržaj glikoze u prednjoj komori oka i koncentraciji u serumu krvi 30 zdravih pasa:

<b>Koncentracija glikoze (%)</b>	<b>Broj pasa</b>
70-79	8
80-89	15
90-99	4
100-109	3

Izračunati apsolutne pokazatelje varijabiliteta i standardizovano odstupanje za vrednost obeležja  $X=100$ .

3. Date su vrednosti numeričkog obeležja X: 10, 30, 50, 70, 90. Varijansa datog obeležja je  $\sigma_x^2 = 800$ . Uporediti varijabilitet obeležja X i obeležja  $Y=2X+3$ .

4. Uporediti varijabilnost 2 serije podataka ukoliko su njihove vrednosti sledeće:

$$\bar{X}_1 = 70 (cm); \sigma_1^2 = 25 (cm)^2; \bar{X}_2 = 35 (g); \sigma_2 = 4 (g).$$





## POKAZATELJI OBLIKA DISTRIBUCIJE

Pokazatelji oblika distribucije sagledavaju dve karakteristike, asimetričnost i spljoštenost. Najčešće korišćeni pokazatelji ovih karakteristika distribucije su:

- Koeficijent asimetričnosti – I Pirsonov koeficijent ( $\beta_1$ )
- Koeficijent spljoštenosti – II Pirsonov koeficijent ( $\beta_2$ ).

Za izračunavanje ovih koeficijenata potrebno je prvo da se izračunaju centralni momenti. Pod centralnim momentom  $k$ -tog reda podrazumeva se sredina sume odstupanja vrednosti obeležja od aritmetičke sredine stepenovana na  $k$ -ti stepen.

**Primer 1.** Na osnovu podataka o broju tovne junadi (000 komada) na jednoj teritoriji utvrditi pokazatelje oblika distribucije i dati odgovarajući komentar.

Broj junadi ( $X_i$ )	$(X_i - \bar{X})$	$(X_i - \bar{X})^2$	$(X_i - \bar{X})^3$	$(X_i - \bar{X})^4$
3				
5				
7				
10				
15				
20				
$\Sigma$ 60				

$$\beta_1 = \frac{\mu_3}{\mu_2} =$$

$$\beta_2 = \frac{\mu_4}{\mu_2^2} =$$

$$\mu_2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n} =$$

$$\mu_3 = \frac{\sum_{i=1}^n (X_i - \bar{X})^3}{n} =$$

$$\mu_4 = \frac{\sum_{i=1}^n (X_i - \bar{X})^4}{n} =$$

**Primer 2.** Utvrditi oblik raspodele podataka koji se odnose na mesečnu potrošnju svinjskog mesa u 100 slučajno odabranih gazdinstava na teritoriji Grada Novog Sada.

Potrošnja mesa ( $X_i$ )	Broj domaćinstava ( $f_i$ )	$(X_i - \bar{X})$	$(X_i - \bar{X})^2$	$f_i(X_i - \bar{X})^2$	$f_i(X_i - \bar{X})^3$	$f_i(X_i - \bar{X})^4$
2,2	10					
2,8	15					
3,0	40					
3,3	20					
3,8	10					
4,0	5					
$\Sigma$	100					

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} =$$

$$\beta_2 = \frac{\mu_4}{\mu_2^2} =$$

$$\mu_2 = \frac{\sum_{i=1}^n f_i (X_i - \bar{X})^2}{\sum_{i=1}^n f_i} =$$

$$\mu_3 = \frac{\sum_{i=1}^n f_i (X_i - \bar{X})^3}{\sum_{i=1}^n f_i} =$$

$$\mu_4 = \frac{\sum_{i=1}^n f_i (X_i - \bar{X})^4}{\sum_{i=1}^n f_i} =$$

**Primer 3.** Za datu seriju podataka utvrditi aritmetičku sredinu, modus i medijanu, a zatim na osnovu izračunatih pokazatelja centralne tendencije opisati oblik date raspodele.

<b>Masa jagnjadi (kg)</b> <i>(<math>X_i</math>)</i>	<b>Broj jagnjadi</b> <i>(<math>f_i</math>)</i>	<b>Kumulativ</b>
3,9	8	8
4,2	12	20
4,3	20	40
4,5	15	55
4,6	5	60
$\Sigma$	60	

**Zadaci za vežbu:**

1. Na osnovu podataka o distribuciji nove smeše za ishranu koka nosilja po maloprodajnim objektima, dobijeni su sledeći rezultati:  $\sigma^2 = 12,25$ ,  $\mu_3 = -5,27$  i  $\mu_4 = 291,7$ . Ispitati asimetriju i oblik ove empirijske distribucije frekvencija?

2. Na osnovu N=100 vrednosti obeležja formirana je intervalna serija distribucije frekvencija i izračunate su sledeće sume:

$$\sum f(x - \bar{x})^2 = 688, \sum f(x - \bar{x})^3 = -144, \sum f(x - \bar{x})^4 = 12544.$$

Izračunati vrednosti prvog i drugog Pirsonovog koeficijenta i uporediti ih sa vrednostima normalne raspodele.

3. Na osnovu datih podataka koji se odnose na mesečnu potrošnju mesa po domaćinstvu izračunati pokazatelje centralne tendencije i preko njih ispitati oblik date raspodele.

<b>Mesečna potrošnja (kg)</b>	<b>Broj domaćinstava</b>
2,2	10
2,8	40
3,0	15
3,3	20
3,8	10
4,0	5







## NORMALNA RASPODELA

Normalna raspodela je najvažniji model teorijske distribucije verovatnoće. Značaj ovog oblika distribucije u statističkoj teoriji i statističkim istraživanjima se ogleda u tome što se mnoge empirijske pojave modeliraju normalnom distribucijom. Parametarska statistika je zasnovana na pretpostavci da osnovni skup kome pripada uzorak ima normalnu distribuciju.

Primena Normalne raspodele:

1.  $p(-a < x < a) = 2F(a)$

2.  $p(b < x < a) = F(a) - F(b)$

3.  $p(-b < x < a) = F(a) + F(b)$

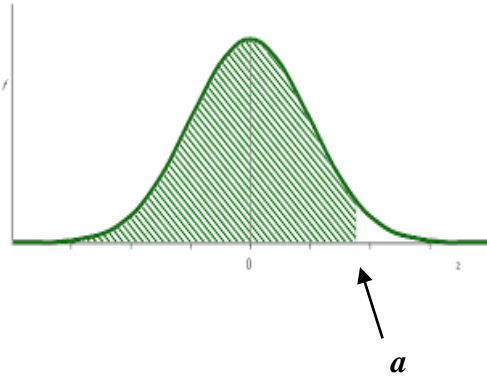
4.  $p(-b < x < -a) = F(-b) - F(-a) = F(b) - F(a)$

**Primer 1.** Kod 250 junadi prosečan dnevni prirast iznosi 0.8 kg sa varijansom  $0.0625\text{kg}^2$ .

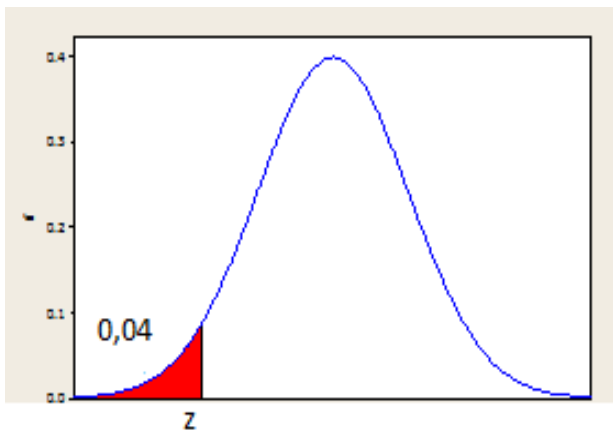
- a) Izračunati verovatnoću da je prosečan dnevni prirast junadi 0.6-0.8kg.
- b) Izračunati očekivan broj junadi sa prosečnim dnevnim prirastom 0.6-0.8kg.
- c) Koliki je očekivan broj junadi sa prosečnim prirastom većim od 1.05kg?

### Zadaci za vežbu:

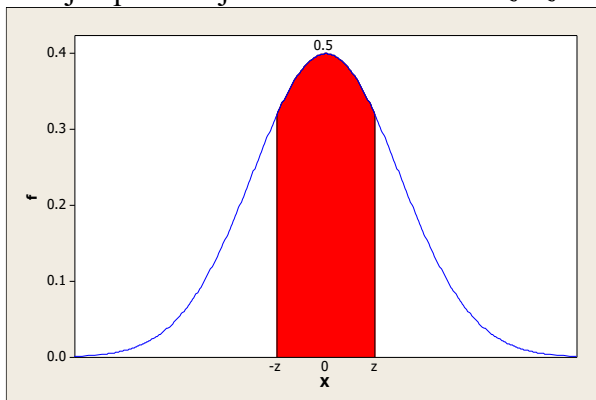
1. Slučajno promenljiva  $Z$  ima normalnu raspodelu sa parametrima  $\mu=0$  i  $\sigma=1$ . Odrediti  $a$  ukoliko osenčena površina zauzima 97% površine ispod krive.



2. Odrediti vrednost  $z$  za standardizovanu normalnu krivu tako da je površina na levom kraju raspodele 0,04.



3. Odrediti vrednost  $z$  za standardizovanu normalnu krivu gde verovatnoća da se vrednost slučajne promenljive  $X$  nalazi između  $-z$  i  $z$  iznosi 0,5.



4. Pretpostavljajući da je sadržaj gvožđa u serumu psa slučajna promenljiva koja ima normalnu raspodelu sa parametrima  $\mu = 25(\mu\text{mol/l})$  i varijansom  $\sigma^2 = 4(\mu\text{mol/l})$ .

a) Izračunati verovatnoću da je sadržaj gvožđa u serumu psa 21-29 ( $\mu\text{mol/l}$ ).

b) Izračunati verovatnoću da je sadržaj gvožđa u serumu 19-31 ( $\mu\text{mol/l}$ ).

## OCENE NA OSNOVU UZORKA

U praktičnom radu, u svrhu donošenja zaključaka o karakteristikama osnovnog skupa, uzima se jedan ili više uzoraka dovoljne veličine na osnovu kojeg se ocenjuju, odnosno procenjuju nepoznati parametri osnovnog skupa.

U slučaju da je poznata veličina osnovnog skupa moguće je izračunati i total osnovnog skupa.

### **Interval poverenja za ocenu nepoznate sredine osnovnog skupa**

Intervali poverenja za ocenu nepoznate aritmetičke sredine osnovnog skupa na osnovu uzorka se razlikuju prema tome da li je poznat varijabilitet osnovnog skupa ili nije.

### ***Poznat varijabilitet osnovnog skupa***

Ukoliko je poznat varijabilitet osnovnog skupa u intervalu poverenja figuriraju tačkasta ocena aritmetičke sredine iz uzorka ( $\bar{x}$ ), prava standardna greška aritmetičke sredine ( $\sigma_{\bar{x}}$ ) i kritične vrednosti iz tablice Normalne raspodele ( $Z_{\alpha}$ ).

**Primer 1.** U rejonu od 10.000 domaćinstava izabran je uzorak čiji su rezultati dati sledećim vrednostima:

Mesečna potrošnja goveđeg mesa (kg) ( $X_i$ )	Broj domaćinstava ( $f_i$ )	$X_i f_i$
1	10	
1,6	15	
1,8	40	
2,1	20	
2,6	10	
2,8	5	
Ukupno		

Oceniti prosečnu mesečnu i ukupnu potrošnju goveđeg mesa sa pouzdanošću od 95 procenata ako je od ranije poznat varijabilitet osnovnog skupa  $\sigma^2 = 3,4 \text{ kg}^2$ .



### *Nije poznat varijabilitet osnovnog skupa*

S druge strane, ukoliko varijabilitet osnovnog skupa nije poznat, pored tačkaste ocene aritmetičke sredine iz uzorka u intervalu poverenja figuriraju još i ocenjena standardna greška aritmetičke sredine ( $S_{\bar{x}}$ ), kao i kritične vrednosti iz tablice Normalne raspodele (za velike uzorke) odnosno kritične vrednosti iz tablice Studentove raspodele (za male uzorke) ( $t_{\alpha;n-1}$ ).

**Primer 1.** Na osnovu podataka o mlečnosti 57 istočno-frizijskih krava na jednoj farmi u toku 305 dana laktacije oceniti prosečnu količinu dobijenog mleka po kravi za posmatrani period. Formirati 95% i 99% interval poverenja.

Količina mleka (lit.) ( $X_i$ )	Broj krava ( $f_i$ )	$X_i f_i$	$X_i^2 f_i$
3.600	3		
3.800	4		
4.000	7		
4.200	10		
4.400	15		
4.600	9		
4.800	4		
5.000	3		
5.200	2		
$\Sigma$			

**Primer 2.** Primenom prostog slučajnog uzorka bez ponavljanja odabrano je 5 od 200 pilića koji su hranjeni određenim koncentratom. Prirast pilića u periodu od 3 nedelje je dat u tabeli. Odrediti 95% i 99% interval poverenja za prosečan i ukupan prirast.

<b>Prirast (g) (<math>X_i</math>)</b>	<b><math>X_i^2</math></b>
391	
401	
405	
418	
432	

### Interval poverenja za ocenu nepoznate proporcije osnovnog skupa

**Primer 1.** Dati su podaci o potrošnji svinjskog mesa. Sa pouzdanošću od 95 i 99% oceniti proporciju domaćinstava sa potrošnjom manjom od 2,5 kg. Sa istom pouzdanošću oceniti ukupan broj domaćinstava sa potrošnjom manjom od 2,5 kg. ukoliko je poznato da je broj domaćinstava u osnovnom skupu 1.000.

Mesečna potrošnja ( $X_i$ )	Broj domaćinstava ( $f_i$ )
2,2	10
2,8	15
3,0	40
3,3	20
3,8	10
4,0	5
Ukupno	

**Zadaci za vežbu:**

1. Date su vrednosti prekidnog numeričkog obeležja koje se odnose na uzorak veličine  $n=20$ :

87      109      79      80      96      95      90      92      96      98  
 101      91      78      112      94      98      94      107      81      96

Ako je poznat varijabilitet osnovnog skupa  $\sigma = 4,4$ , odrediti granice 98% interval poverenja za aritmetičku sredinu osnovnog skupa.

2. Iz osnovnog skupa od 1000 gazdinstava izabran je uzorak radi ispitivanja broja tovnih svinja. Odrediti 95% i 99% interval poverenja za prosečan i ukupan broj svinja u osnovnom skupu, ako je poznato da je primenjen prost slučajan uzorak bez ponavljanja.

<b>Broj tovnih svinja (<math>X_i</math>)</b>	0	1	2	3	4	5	8
<b>Broj gazdinstava (<math>f_i</math>)</b>	4	5	11	37	30	10	3

3. Metodom slučajnog izbora na teritoriji jedne opštine izabran je prost slučajan uzorak od 25 parcela od po jednog hektara zasejanih kukuruzom na kojima su dobijeni sledeći prinosi (t/ha):

<b>Prinos (t/ha)(<math>X_i</math>)</b>	2,51-3,50	3,51-4,50	4,51-5,50	5,51-6,50
<b>Broj parcela (<math>f_i</math>)</b>	3	10	8	4

Odrediti granice 95% i 99% interval poverenja u kojima se može očekivati da se nalazi prosečan prinos po hektaru u celoj opštini.

4. U uzorku od 100 pasa koji su dobili prvu vakcinu obolelo je 7 pasa. Odrediti 95% i 99% interval poverenja za proporciju obolelih pasa. Koliki se broj obolelih pasa može očekivati ako je poznato da osnovni skup broji 500 pasa.



## TESTIRANJE STATISTIČKIH HIPOTEZA

Pod hipotezom se podrazumeva naučna pretpostavka zasnovana na poznatim činjenicama radi izvođenja nekog zaključka. Predmet statističkog testiranja mogu biti različiti parametri, a najčešće su to aritmetička sredina i proporcija.

### Testovi aritmetičkih sredina

Postoje sledeći osnovni testovi za testiranja aritmetičkih sredina:

1. Upoređivanje aritmetičke sredine uzorka sa aritmetičkom sredinom osnovnog skupa ili sa nekom hipotetičkom vrednošću – test značajnosti jedne sredine;
2. Upoređivanje dve aritmetičke sredine iz dva nezavisna uzorka – test značajnosti razlike dve sredine;
3. Upoređivanje više od dve sredine iz više od dva uzorka – metod analize varijanse.

#### *1. Test značajnosti jedne sredine*

**Primer 1.** Rezultati uzorka,  $n=70$ , pokazali su da je prosečan prirast prasića starosti do 6 nedelja 395 g/dan. Od ranije je poznato da je prosečan prirast prasića starosti do 6 nedelja 430 g/dan sa standardnom devijacijom 30 g/dan. Da li je postignuti prinos na nivou očekivanog.

**Primer 2.** Dati su podaci o prosečnom utrošku hrane 25 pilića starosti od dva meseca izabranih primenom prostog slučajnog uzorka bez ponavljanja od ukupnog broja od 500 pilića koliko ih ima na farmi:

<b>Utrošak hrane po piletu (g)</b>	196	198	200	202
<b>Broj pilića</b>	3	10	7	5

Da li može da se prihvati nulta hipoteza da je prosečan utrošak hrane u osnovnom skupu  $\mu=199(g)$ ?

## 2. Test značajnosti razlike dve sredine

**Primer 1.** Na osnovu oglada u kome se ispituje razlika u prirastu (g) dve grupe životinja kod kojih su primenjena dva načina ishrane A i B dobijeni su rezultati koji se odnose na veličine prostih slučajnih nezavisnih uzoraka, aritmetičke sredine i standardne devijacije osnovnih skupova koje su od ranije poznate. Testirati nultu hipotezu da je prosečan prirast kod obe grupe isti nezavisno od primenjenog načina ishrane.

<b>Ishrana A</b>	<b>Ishrana B</b>
$n_1=15$	$n_2=12$
$\bar{x}_1 = 68,4$	$\bar{x}_2 = 83,42$
$\sigma_1 = 16,47$	$\sigma_2 = 17,63$



**Primer 2.** Pri ispitivanju uticaja dve vrste hraniva u ishrani junadi postavljen je ogled sa dve grupe od po deset grla. Rezultati eksperimenta se odnose na prosečan dnevni prirast junadi (kg) što je predstavljeno u sledećoj tabeli:

<b>Uzorak 1</b>	<b>Uzorak 2</b>
1,40	1,31
1,38	1,30
1,41	1,33
1,35	1,37
1,42	1,39
1,37	1,36
1,42	1,35
1,40	1,32
1,43	1,31
1,45	1,35

Testirati hipotezu o jednakosti sredina dva nezavisna uzorka.

**Zadaci za vežbu:**

1. U štali sa 10 krava ostvarena je dnevna mlečnost (lit.): 3, 4, 5, 7, 8, 9, 7, 10, 9, 10. Ako je varijansa osnovnog skupa poznata i iznosi 6, testirati  $H_0: \bar{x} = \mu$  protiv alternativne hipoteze  $H_1: \bar{x} \neq \mu$ , ako važi da je  $\mu=7,8$  (lit.).

2. Na osnovu podataka datih u tabeli o inkubaciji tetanusa (danima) i broju obolelih životinja testirati nultu hipotezu da je prosečna inkubacija 16 dana.

Inkubacija tetanusa	Broj obolelih životinja
5-8	5
9-12	5
13-16	4
17-20	3
21-24	2
25-28	1

3. Prosečna dnevna potrošnja svežeg mleka, u uzorku od 30 poljoprivrednih domaćinstava iznosi 2,1 litar, sa standardnom devijacijom osnovnog skupa 0,9. U uzorku od 20 radničkih domaćinstava prosečna dnevna potrošnja svežeg mleka je 1,5 litara sa varijansom osnovnog skupa 0,49. Da li se statistički značajno razlikuje prosečna dnevna potrošnja svežeg mleka poljoprivrednih i radničkih domaćinstava?

4. Dati su podaci o broju dana potrebnih za dresiranje pasa ukoliko se pri treniranju primenjuje nagrada (I grupa) ili kazna (II grupa).

<b>A</b>	43	41	48	44	51	48	47	35	
<b>B</b>	42	47	57	53	74	59	65	54	46

Testirati nultu hipotezu da ne postoji statistički značajna razlika u prosečnom vremenu potrebnom za dresiranje pasa u zavisnosti od načina dresiranja.

5. Da bi se utvrdilo da li postoji statistički značajna razlika u brzini delovanja dva leka A i B protiv bola, istraživač je primenio lekove na dva uzorka slučajno odabranih pacijenata. Rezultati testa su dati u sledećoj tabeli. Prosečno vreme i sume kvadrata odstupanja aritmetičke sredine vremena od početka delovanja leka u uzorcima (minutima) dati su u tabeli:

Lek	Veličina uzorka	Prosečno vreme do početka delovanja leka	Sume kvadrata odstupanja od aritmetičke sredine
A	19	40	$\sum(X_1 - \bar{X}_1)^2 = 1815$
B	12	50	$\sum(X_2 - \bar{X}_2)^2 = 729$

Da li postoji statistički značajna razlika u prosečnom vremenu do početka delovanja lekova A i B?





### 3. Analiza varijanse

**Primer 1.** U jednom poljoprivrednom preduzeću organizovana je ishrana svinja na tri različita načina. Da bi se ispitalo da li način ishrane značajno utiče na tov svinja iz svake grupe, formiran je slučajni uzorak sa po 8 svinja. Ustanovljen je sledeći prirast u težini (kg) za dve nedelje:

Redni broj	Prirast		
	I	II	III
1	15	10	20
2	13	25	22
3	16	19	14
4	20	18	17
5	24	16	13
6	18	20	18
7	18	22	15
8	23	14	10

- Da li postoji statistički značajna razlika u prosečnom prirastu svinja u zavisnosti od načina ishrane?
- Primenom t-testa i testa najmanje značajne razlike uporediti aritmetičke sredine tretmana.

Nulta hipoteza:

Alternativna hipoteza:

Izvori varijacije	Stepeni slobode	Sume kvadrata	Sredine suma kvadrata (varijanse)	F- odnos	F – tablično ( $\alpha_1=k-1; \alpha_2=N-k$ )	
					0,05	0,01
Tretmani	k-1=	$Q_T=$	$S_T^2=$	$S_T^2/S_P^2=$		
Pogreška	N-k=	$Q_P=$	$S_P^2=$			
Total	N-1=	$Q=$				

k (broj tretmana) –

n (veličina uzorka) –

N (ukupan broj podataka) –

Sume kvadrata

$$Q = \sum_{i=1}^k \sum_{j=1}^{n_i} X_{ij}^2 - C =$$

C – je korektivni faktor

$$C = \frac{\left( \sum_{i=1}^k \sum_{j=1}^{n_i} X_{ij} \right)^2}{N} =$$

$$Q_T = \frac{\sum_{i=1}^k \left( \sum_{j=1}^{n_i} X_{ij} \right)^2}{n} - C =$$

$$Q_P = Q - Q_T =$$

Sredine suma kvadrata

$$S_T^2 = \frac{Q_T}{k-1} =$$

$$S_P^2 = \frac{Q_P}{N-k} =$$

t-test

Test najmanje značajne razlike

**Zadaci za vežbu:**

1. Da bi se utvrdilo da li prosečan dnevni prirast zavisi od vrste žitarica primenjene u ishrani, izveden je eksperiment u kome je svakoj grupi od 7 miševa dat obrok jedne od četiri vrste žitarica. Prirast (g) dat je u sledećoj tabeli:

<b>T R E T M A N I</b>			
<b>A</b>	<b>B</b>	<b>C</b>	<b>D</b>
9	5	2	3
7	4	1	8
8	6	1	5
8	4	2	9
7	5	2	2
8	7	3	7
8	3	2	8

- Testirati nultu hipotezu da se prosečni prinos statistički značajno ne razlikuje u zavisnosti od načina ishrane;
- Primenom testa najmanje značajne razlike uporediti aritmetičke sredine parova tretmana.

2. U eksperimentu organizovanom po planu potpuno slučajnog rasporeda ispituje se efikasnost tri hemijska preparata. Intezitet iritantnosti je označen 0-10 i izmeren je na 8 eksperimentalnih životinja:

<b>Preparati</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>
<b>I</b>	6	9	6	5	7	5	6	6
<b>II</b>	5	9	9	8	8	7	7	7
<b>III</b>	3	4	3	6	8	5	5	6

- Testirati nultu hipotezu da je efikasnost sva tri preparata jednaka.
- Testirati značajnost razlike parova tretmana primenom NZR testa.







## **Testovi proporcija**

Postoje sledeći osnovni testovi za testiranja proporcija:

1. kada upoređujemo proporciju iz uzorka sa pretpostavljenom proporcijom osnovnog skupa ili sa nekom teorijskom vrednošću – test značajnosti jedne proporcije
2. kada upoređujemo dve proporcije iz dva nezavisna uzorka- test značajnosti razlike dve proporcije.

### ***Test značajnosti jedne proporcije***

**Primer 1.** U uzorku od 164 grla jedne rase bilo je 52 goveda sa dužinom trupa ispod 160 cm. Može li se prihvatiti nulta hipoteza da u osnovnom skupu proporcija goveda sa dužinom trupa ispod 160 cm iznosi 0,4?

### ***Test značajnosti razlike dve proporcije***

**Primer 2.** U uzorcima od po 110 grla goveda dve rase, obolela grla učestvuju sa 6% i 13%. Utvrditi da li je otpornost dve rase goveda prema ispitivanoj bolesti ista.

**Zadaci za vežbu:**

1. U uzorku od 100 jedinki vrste *Tribolium castaneum* rastvorom aktivne supstance hlorprifos metila posle 24 časa je registrovan mortalitet 73 jedinki. Očekivana proporcija smrtnosti je 0,9. Može li se smatrati da razlika konstatovane i očekivane smrtnosti nije statistički značajna?

2. Ispitivanjem prostog slučajnog uzorka od 200 životinja utvrđeno je da je 4 životinje obolelo. Proporcija obolelih životinja u uzorku je \_\_\_\_\_, dok je ocena za standardnu grešku uzoračke proporcije \_\_\_\_\_. Nulta hipoteza da je uzoračka proporcija statistički značajna se sa rizikom od 5% \_\_\_\_\_.

3. Ispituje se efikasnost dve vakcine A i B. Prva vakcina je primenjena na uzorku od 32 životinje i bila je uspešna u 24 slučaja. Druga vakcina je primenjena na uzorku od 26 životinja i bila je efikasna u 18 slučajeva. Da li postoji statistički značajna razlika u efikasnosti vakcina?



## REGRESIONA I KORELACIONA ANALIZA

Regresiona i korelaciona analiza se primenjuje s namerom da se opiše, predvidi i kontroliše zavisna promenljiva u odnosu na jednu ili više nezavisno promenljivih, kao i da se utvrdi postojanje i intenzitet veze između zavisne i nezavisnih promenljivih.

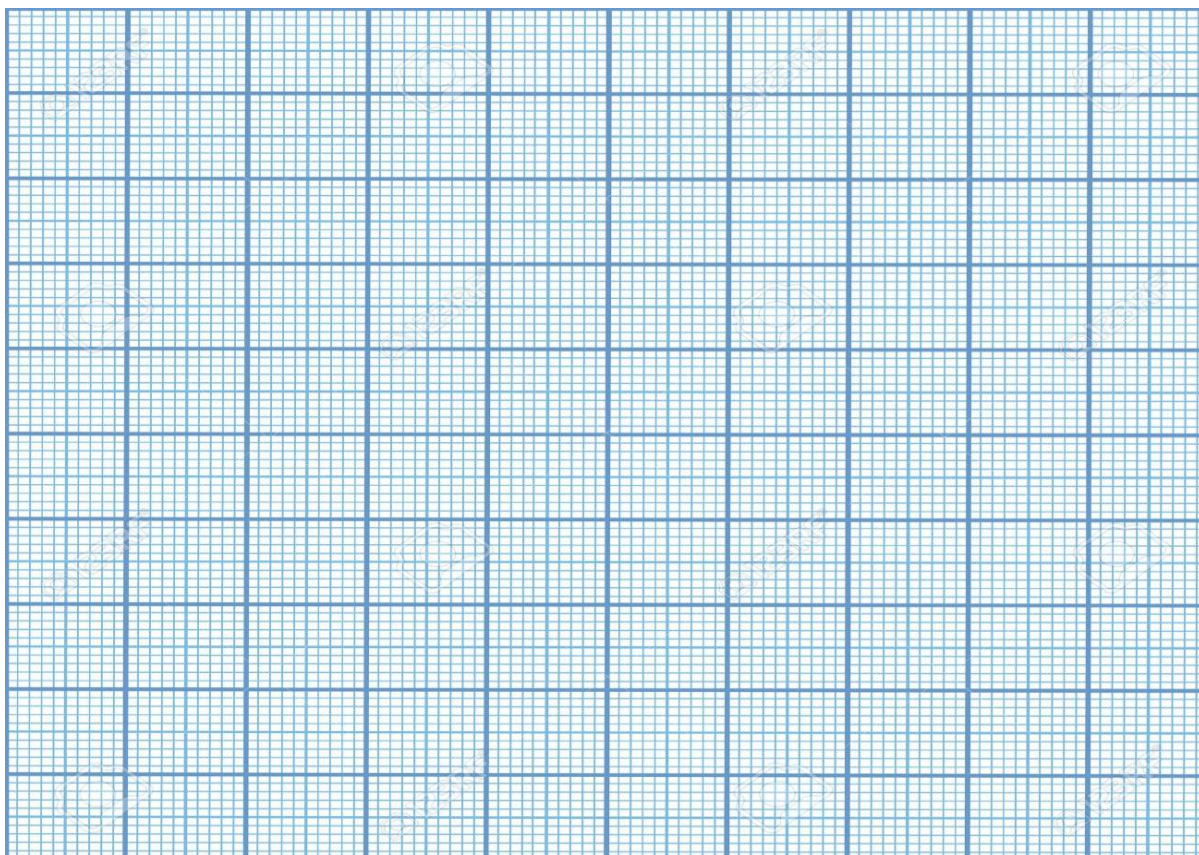
**Primer 1.** Dati su podaci o mlečnosti koze (kg/24h) u zavisnosti od dana laktacije:

<b>Mlečnost (Y)</b>	1,84	1,75	1,50	1,62	1,51	1,25	1,26	1,15
<b>Dani Laktacije (X)</b>	14	24	34	44	54	64	74	84

Na osnovu datih parova vrednosti izvesti sledeća računanja:

- Formirati dijagram rasturanja;
- Oceniti regresioni model, izračunati ocenjene vrednosti zavisno promenljive i ucrtati liniju regresije na dijagramu rasturanja;
- Izračunati standardnu grešku regresije;
- Izračunati vrednost koeficijenta korelacije, determinacije i nedetrminacije;
- Na osnovu jednačine regresije oceniti mlečnost 67. dana laktacije.

Dijagram rasturanja:



Radna tabela:

<b>X</b>	<b>Y</b>							





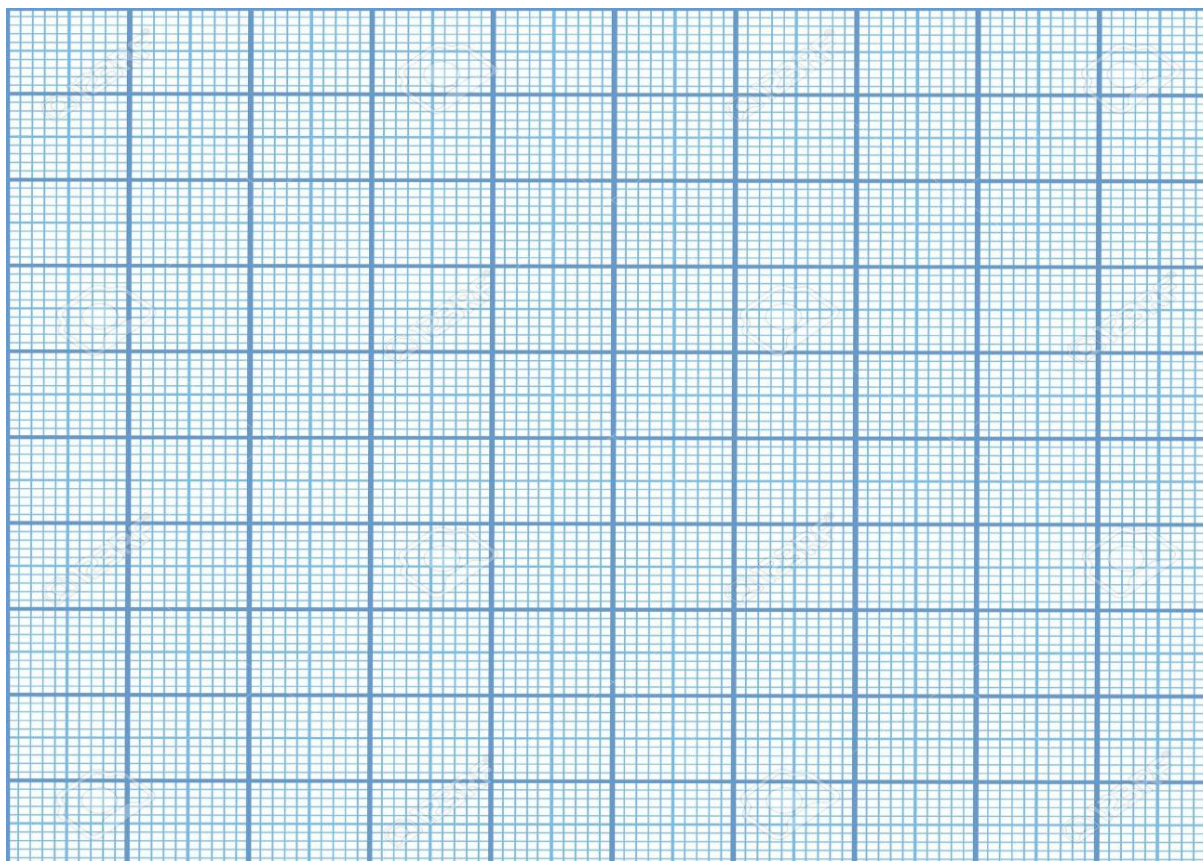
**Primer 2.** Rezultati ispitivanja uticaja vitamina holina, dodatog smeši hrane, na prirast kod ispitivanih prasadi dati su sledećim vrednostima:

<b>Prosečan dnevni prirast (kg) (Y)</b>	420	405	387	377
<b>Dodatak holina (mg/kg hrane) (X)</b>	0	250	500	1000

Na osnovu datih parova vrednosti izvesti sledeća računanja:

- f) Formirati dijagram rasturanja;
- g) Oceniti regresioni model, izračunati ocenjene vrednosti zavisno promenljive i ucrtati liniju regresije na dijagramu rasturanja;
- h) Izračunati standardnu grešku regresije;
- i) Izračunati vrednost koeficijenta korelacije, determinacije i nedetrminacije;
- j) Na osnovu jednačine regresije oceniti prosečan dnevni prirast za dodatak holina od 750 mg/kg hrane.

Dijagram rasturanja:



Radna tabela:

<b>X</b>	<b>Y</b>						

### Zadaci za vežbu:

1. Dati su podaci o dozi leka i vremenu izlečenja:

<b>Doza leka</b>	0	1	2	3	4	5	6
<b>Vreme izlečenja</b>	15,4	12,2	13,7	9,2	9,9	6,1	4,1

- Oceniti parametre linearnog regresionog modela i objasniti njihovo značenje;
- Izračunati koeficijent determinacije;
- izračunati standardnu grešku regresije.

2. Na osnovu podataka o težini 10 teladi pri rođenju i njihovom prirastu izračunato je:  $\sum X=410$ ,  $\sum Y=304$ ,  $\sum XY=12665$ ,  $\sum X^2=16974$ ,  $\sum Y^2=9508$ . Izračunati koeficijente korelacije i determinacije i objasniti njihovo značenje. Da li telad koja imaju veću težinu pri rođenju imaju i veći očekivani prirast?

3. Na osnovu podataka o proizvodnji mleka (000 lit.) i količini hrane (t) 10 krava izračunate su sledeće sume:  $\sum X=1530$ ,  $\sum Y=748$ ,  $\sum XY=116210$ ,  $\sum X^2=248700$ . Oceniti linearni regresioni model i objasniti značenje regresionog koeficijenta.



## LITERATURA

1. Čobanović Katarina (2003), Primeri za vežbanje iz statistike, Poljoprivredni fakultet, Univerzitet u Novom Sadu, Republika Srbija
2. Lozanov-Crvenković Zagorka (2002), Statistika, Prirodno-matematički fakultet, Univerzitet u Novom Sadu, Republika Srbija
3. Man S.P. (2017), Uvod u statistiku, deveto izdanje, John Wiley & Sons
4. Mutavdžić Beba, Nikolić-Đorić Emilija (2018), Statistika (za smer Veterinarska medicina), Poljoprivredni fakultet, Univerzitet u Novom Sadu, Republika Srbija