UNIVERZITET U NOVOM SADU
POLJOPRIVREDNI FAKULTET

STATISTIKA - PRAKTIKUM

(za smer Agroekonomija)

Dr Beba Mutavdžić
Dr Dragana Tekić
Dr Tihomir Novaković

2

STATISTIKA - PRAKTIKUM

STATISTIKA - PRAKTIKUM

(za smer Agroekonomija)

Beba Mutavdžić, Dragana Tekić, Tihomir Novaković

EDICIJA POMOĆNI UDŽBENIK

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

Prof. dr Nedeljko Tica, redovni profesor

Dekan Poljoprivrednog fakulteta u Novom Sadu

Članovi komisije za izdavačku delatnost

dr Branislav Vlahović, redovni profesor, predsednik

dr Ivana Davidov, vanredni profesor, član

dr Dejan Beuković, docent, član

dr Ksenija Mačkić, vanredni profesor, član

CIP – Каталогизација у публикацији

Библиотека Матице српске

ISBN 978-86-7520-597-5

COBISS.SR-ID 1

Autori:

Dr Beba Mutavdžić, vanredni profesor

Dr Dragana Tekić, docent

Dr Tihomir Novaković, docent

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 preštampavanje i fotokopiranje. Sva prava zadržava izdavač.

Štampanje odobrila Komisija za izdavačku delatnost i Naučno-nastavno veće
Poljoprivrednog fakulteta u Novom Sadu.

Tiraž: 20 primeraka

Mesto i godina štampanja: Novi Sad, 2023

PREDGOVOR

Praktikum za predmet „Statistika”, koji se proučava na prvoj godini smera Agroekonomija, 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 deset 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 **–** ***Teorijske distribucije***
- Šesto poglavlje - ***Distribucija sredina uzoraka***
- Sedmo poglavlje **–** ***Ocene na osnovu uzorka***
- Osmo poglavlje **–** ***Testiranje statističkih hipoteza***
- Deveto poglavlje **–** ***Regresiona i korelaciona analiza***
- Deseto poglavlje- ***Indeksni brojevi i analiza vremenskih serija***

Svako poglavlje obuhvata jedan ili više rešenih primera koji će se detaljno rešavati na vežbama, 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 agroekonomskih nauka.

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

Novi Sad, AUTORI

2023. godine

Sadržaj

FORMIRANJE DISTRIBUCIJA FREKVENCIJA

Prekidno obeležje

Primer 1. Broj zaposlenih radnika sa visokom stručnom spremom u trideset poljoprivrednih radnih organizacija dat je u sledećoj seriji:

1	4	5	0	8	7	3	4	9	6	7	3	0	5	1
5	3	4	4	6	5	7	3	2	4	1	2	3	8	6

a) Formirati neintervalnu distribuciju frekvencija;
b) Formirati intervalnu distribuciju frekvencija (i=2);
c) Izračunati relativne frekvencije (strukturu);
d) Formirati kumulativnu distribuciju frekvencija i kumulaciju structure;
e) Podatke grafički predstaviti histogramom i poligonom.

Tačkasti dijagram:

Sređena statistička serija:

Neintervalna serija:

Broj radnika ***(X_i)***	**Broj radnih organizacija** ***(f_i)***
Ukupno	

Intervalna serija:

Broj radnika *(X_i)*	**Broj radnih organizacija** *(f_i)*	**Relativna frekvencija (struktura)**	**Kumulativ**		**Kumulacija strukture**	
			ispod	**iznad**	**Ispod**	**iznad**
Ukupno						

Grafički prikaz (histogram i poligon):

Primer 2. Broj članova domaćinstava u 50 slučajno odabranih seoskih domaćinstava je bio:

3	3	1	2	2	1	6	6	7	8
3	5	3	4	5	6	5	4	5	2
5	4	5	6	3	3	5	5	4	4
7	4	5	6	6	4	5	4	5	3
4	4	6	5	6	5	4	2	5	5

a) Formirati neintervalnu distribuciju frekvencija i grafički je predstaviti;
b) Formirati intervalnu distribuciju frekvencija (i=2);
c) Izračunati relativne frekvencije (strukturu);
d) Formirati kumulativnu distribuciju frekvencija i kumulaciju strukture;
e) Podatke grafički predstaviti histogramom i poligonom.

Tačkasti dijagram:

Sređena statistička serija:

Neintervalna serija:

Broj članova ***(X_i)***	**Broj domaćinstava** ***(f_i)***
Ukupno	

Intervalna serija:

Broj članova ***(X_i)***	**Broj domaćinstava** ***(f_i)***	**Relativna frekvencija (struktura)**	**Kumulativ**		**Kumulacija strukture**	
			ispod	**iznad**	**ispod**	**iznad**
Ukupno						

Grafički prikaz (histogram i poligon):

Neprekidno obeležje

Primer 3. Veličina poseda trideset posmatranih individualnih gazdinstava izražena u hektarima bila je sledeća:

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

a) Formirati dijagram stablo-list;
b) Formirati intervalnu seriju distribucije frekvencija ako je i=2 i grafički je predstaviti primenom histograma i poligona frekvencija;
c) Izračunati relativne frekvencije (strukturu);
d) Formirati kumulativnu distribuciju frekvencija i kumulaciju strukture.

Dijagram stablo-list:

Radna tabela:

Grupni intervali (X_i)	**Frekvencija (f_i)**	**Relativna frekvencija (struktura)**	**Kumulativ**		**Kumulativ strukture**	
			ispod	**iznad**	**ispod**	**iznad**
Ukupno						

Grafički prikaz (histogram i poligon):

Primer 4. Dati su podaci o prosečnoj oceni tokom studija kod 20 slučajno odabranih diplomiranih agroekonomista:

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

a) Formirati dijagram stablo-list.
b) Formirati intervalnu seriju distribucije frekvencija i grafički je predstaviti primenom histograma i poligona frekvencija.
c) Izračunati relativne frekvencije (strukturu).
d) Formirati kumulativnu distribuciju frekvencija i kumulaciju strukture.

Dijagram stablo-list:

Radna tabela:

Grupni intervali ***(X_i)***	**Frekvencija** ***(f_i)***	**Relativna frekvencija (struktura)**	**Kumulativ**		**Kumulativ strukture**	
			ispod	**iznad**	**ispod**	**iznad**
Ukupno						

Grafički prikaz (histogram i poligon):

Zadaci za vežbu:

1. Primenom upitnika na 30 ispitanika zaposlenih u jednom poljoprivrednom preduzeću, dobijeni su podaci o njihovim godinama starosti:

40	42	38	8	35	40	42	41	44	40	43	43	37	39	41
44	38	40	42	37	44	43	39	41	42	39	42	43	41	43

Za navedene podatke formirati neintervalnu distribuciju frekvencija i grafički je predstaviti.

2. Dati su podaci o broju kupaca u toku 1h u 36 prodavnica:

84	49	61	40	83	67	45	66	70	69	80	58
68	60	67	72	73	70	57	63	70	78	52	67
75	61	70	81	76	79	75	76	53	67	58	31

a) Formirati neintervalnu distribuciju frekvencija i grafički je predstaviti;
b) Formirati intervalnu distribuciju frekvencija (i=2);
c) Izračunati relativne frekvencije (strukturu);
d) Formirati kumulativnu distribuciju frekvencija i kumulaciju strukture;
e) Podatke grafički predstaviti histogramom i poligonom.

3. Na osnovu podataka prikazanih dijagramom stablo-list (cifra jedinica je stablo, prva decimala je list), formirati intervalnu distribuciju ako je i=1,5 i grafički je predstaviti.

0 | 3 6 9
1 | 0 1 2 5 5 8 8 9
2 | 1 3 5 6 7
3 | 2 4 6 8
4 | 1 3 5
5 | 6 7
6 | 2

4. Prinos kukuruza (t/ha) kod 20 ispitivanih radnih organizacija je bio:

6,8	4,8	3,1	2,1	7,6	5,8	3,2	5,6	5,6	5,5
4,5	5,7	6,7	5,6	7,8	6,6	4,9	4,6	3,7	3,9

a) Formirati dijagram stablo-list;
b) Formirati intervalnu seriju distribucije frekvencija i grafički je predstaviti primenom histograma i poligona frekvencija;
c) Izračunati relativne frekvencije (strukturu), kumulativnu seriju distribucije frekvencija (kumulativ, kumulaciju) i kumulaciju strukture.

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

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. Dužina radnog staža (u godinama) osam radnika u jednoj fabrici data je u sledećoj seriji:

Dužina radnog staža *(X_i)*	11	10	3	9	4	11	8	7

Utvrditi prosečnu dužinu radnog staža radnika u posmatranoj fabrici.

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

Primer 2. Na osnovu podataka o broju traktora u osamdeset posmatranih radnih organizacija izračunati aritmetičku sredinu (prosečan broj traktora po radnoj organizaciji).

Broj traktora *(X_i)*	**Broj radnih organizacija *(f_i)***	
5	2	
10	3	
12	5	
15	10	
20	25	
24	20	
30	15	
Ukupno	**80**	

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

Primer 3. Utvrditi prosečnu starost radnika u jednoj radnoj organizaciji na osnovu sledeće serije podataka:

Godine starosti	Broj radnika		
25-29	12		
30-34	20		
35-39	24		
40-44	20		
45-49	16		
50-54	8		
Ukupno			

Metod proizvoljnog početka i transformacija obeležja

Primer 1. Primenom metode proizvoljnog početka izračunati aritmetičku sredinu obeležja X: 97 79 81 82 92 95 83.

X_i		
97		
79		
81		
82		
92		
95		
83		
609		

Primer 2. Izračunati aritmetičku sredinu intervalne serije distribucije frekvencija primenom metode proizvoljnog početka i transformacijom obeležja:

Grupni intervali	**Frekvencija**					
2,00-2,49						
2,50-2,99						
3,00-3,49						
3,50-3,99						
4,00-4,49						
4,50-4,99						
Ukupno						

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, odnosno deljiv sa 2 ili ne. Kod intervalnih serija koristi se korigovana formula.

Primer 1. Utvrditi medijalnu vrednost za sledeće serije podataka:

$\boldsymbol{X_i}$	97	78	82	82	95	83	92
Uređena serija							

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

$\boldsymbol{X_i}$	260	260	230	280	290	280
Uređena serija						

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

Primer 2. Utvrditi medijalnu vrednost za sledeću seriju:

Prinos šećerne repe (t/ha) (X_i)	**Površina (ha) (f_i)**	**Kumulativ**
3,9	8	
4,2	12	
4,3	20	
4,5	15	
4,6	5	
Ukupno	**60**	

Primer 3. U sledećoj tabeli dati su podaci o vremenu utrošenom za pružanje određene usluge po klijentu (u minutima) i broj klijenata u jednoj banci. Utvrditi medijalnu vrednost (primenom korigovane formule).

Vreme (min) (X_i)	**Broj klijenata (f_i)**	**Kumulativ**
3,1-5,0	2	
5,1-7,0	4	
7,1-9,0	10	
9,1-11,0	7	
11,1-13,0	4	
13,1-15,0	3	
Ukupno		

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

Modus

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

Primer 1. Utvrditi modus za sledeću seriju podataka:

X_i	22	25	20	25	26	35	30	22	33
Uređena serija									

$M_O{}^1 =$

$M_O{}^2 =$

Primer 2. Utvrditi najčešći broj prodatih proizvoda u toku radnog dana:

Broj prodatih proizvoda *(X_i)*	**Radni dani (f_i)**
5	2
6	5
7	3
8	7

$M_O =$

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

Veličina gazdinstva (ha)	**Frekvencija**
0,51-1,50	16
1,51-2,50	40
2,51-3,50	50
3,51-4,50	30
4,51-5,50	17
5,51-6,50	7
$\sum$	**160**

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

$d_1 =$

$d_2 =$

Geometrijska sredina

Geometrijska sredina (G) se koristi kako bi se izračunala sredina relativnih (procentnih) promena kao što su indeksni brojevi, kamatne stope, stope inflacije.

Primer 1. Stope inflacije za prethodnih 5 godina iznose: 6%, 8%, 5%, 4%, 3%, respektivno. Odrediti kolika je prosečna stopa inflacije za posmatrani period.

Prvi način:

$$G = \sqrt[N]{X_1 \cdot X_2 \cdot \ldots \cdot X_N} =$$

Drugi način:

$$logG = \frac{\sum_{i=1}^{N} logX_i}{N} =$$

$$G = antilog\ (logG) =$$

Primer 2. Dati su podaci o srednjem deviznom kursu (RSD/EUR) u periodu 2003-2016. registrovani 7. marta. Izračunati srednji godišnji relativni porast deviznog kursa i geometrijsku stopu rasta. Predvideti vrednost srednjeg deviznog kursa 7.3.2017.

Godine	**Srednji devizni kurs**	**Relativne promene deviznog kursa (R_t)**
2003	63,63	
2004	69,47	
2005	80,51	
2006	86,96	
2007	81,15	
2008	85,28	
2009	94,56	
2010	99,70	
2011	103,60	
2012	110,59	
2013	111,43	
2014	115,95	
2015	120,86	
2016	123,40	

$$G = \sqrt[n-1]{R_2 \cdot R_3 \cdot R_4 \cdot \ldots \cdot R_t} =$$

$$G = \sqrt[n-1]{\frac{Y_n}{Y_1}} =$$

Stopa promene:

$$p = (G - 1) \cdot 100 =$$

Predviđanje deviznog kursa:

$$Y_{n+1} = Y_n \cdot G =$$

Primer 3. Izračunati geometrijsku sredinu sledeće distribucije frekvencija:

X_i	f_i	$logX_i$	$f_i \cdot logX_i$	f_iX_i
5	34			
7	3			
30	3			
$\sum$	40			

$$logG = \frac{\sum f_i logX_i}{\sum f_i} =$$

$$G = antilog\ (logG) =$$

Harmonijska sredina

Harmonijska sredina se upotrebljava u slučajevima kada su numerička vrednost obeležja i obim pojave u obrnutoj srazmeri.

Primer 1. Sedam radnika proizvodi istu vrstu proizvoda i za jedinicu tog proizvoda utroše sledeće radno vreme (min): 12, 16, 19, 23, 18, 26, 20. Izračunati prosečno radno vreme potrebno za izradu proizvoda.

Vreme (min) *(X_i)*	$\boldsymbol{\frac{1}{X_i}}$
12	
16	
19	
23	
18	
26	
20	
$\sum$	

$$H = \frac{n}{\sum \frac{1}{X_i}} =$$

Primer 2. U jednoj radnoj organizaciji 5 radnika izradi jedan proizvod za 23 minuta, 7 radnika za 28 minuta, 9 radnika za 34 minuta, 4 radnika za 38 minuta i 5 radnika za 43 minuta. Izračunati prosečno vreme izrade tog proizvoda.

Vreme (min.) (*X_i*)	**Broj radnika (*f_i*)**	$\boldsymbol{\frac{f_i}{X_i}}$
23		
28		
34		
38		
43		
$\sum$		

$$H = \frac{\sum f_i}{\sum \frac{f_i}{X_i}} =$$

Kvantili

Kvantili su vrednosti numeričke promenljive koji statističku seriju dele na q jednakih delova. Kvartili su kvantili koji dele statističku seriju na 4 jednaka dela. Kako je u tom slučaju red kvantila q=4 postoje tri kvartila, prvi ili donji (Q_1), drugi kvartil ili medijana (Q_2) i treći (gornji) kvartil (Q_3).
Decili su kvantili koji dele statističku seriju na 10 jednakih delova.
Percentili su kvantili koji dele statističku seriju na 100 jednakih delova.

Primer 1. Izračunati kvartile za sledeću seriju podataka:

X_i	3	5	5	6	6	7	8	10	11	12	12

$$Q_1 = X_{\left[\frac{n}{4}\right]+1} =$$

$$Q_2 = M_e =$$

$$Q_3 = X_{\left[\frac{3n}{4}\right]+1} =$$

Funkcija ceo deo broja *x* koja se označava $[x]$ je najveći ceo broj koji nije veći od *x*.
Na primer, [2,75]=2 ili [8,25]=8.

Primer 2. Izračunati kvartile za sledeću seriju podataka:

X_i	9	10	15	17	20	12	25	28
Sređena serija								

$$Q_1 = \frac{X_{\frac{n}{4}} + X_{\frac{n}{4}+1}}{2} =$$

$$Q_2 = M_e =$$

$$Q_1 = \frac{X_{\frac{3n}{4}} + X_{\frac{3n}{4}+1}}{2} =$$

Primer 3. Izračunati kvartile za sledeću distribuciju frekvencija:

Grupni intervali	f_i	X_i	Kumulativ ispod
0,1-2,0	5		
2,1-4,0	7		
4,1-6,0	10		
6,1-8,0	5		
8,1-10,0	3		
$\sum$	**30**		

$$Q_1 = L + \frac{\frac{n}{4} - F_{Q_1 - 1}}{f_{Q_1}} \cdot i =$$

$$Q_2 = M_e =$$

$$Q_1 = L + \frac{\frac{3n}{4} - F_{Q_3 - 1}}{f_{Q_3}} \cdot i =$$

Zadaci za vežbu:

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

1 5 7 3 9 6 6 5 4 7 6 2

2. Popuniti polja koja nedostaju u tabeli, a zatim izračunati aritmetičku sredinu, modus i medijanu.

Grupni intervali	Frekvencija	Kumulativ ispod
10-20	5	5
20-30	8	13
30-40	12	
40-50		
50-60	7	
60-70	4	50

3. Dati su podaci o dužini staža zaposlenih u jednoj firmi:

Masa teladi (kg)	**Broj teladi**
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

a) Izračunati aritmetičku sredinu (primenom metoda proizvoljnog početka i transformacije obeležja).
b) Izračunati modus i medijanu (primenom korigovanih formula).

4. Na osnovu podataka o prinosu pšenice u periodu od 2012-2016 godine, utvrditi prosečan prinos izračunavanjem geometrijske sredine.

Godine	**Prinos pšenice (t/ha)**
2012	2,80
2013	3,41
2014	2,73
2015	3,47
2016	4,49

5. Izračunati geometrijsku sredinu za sledeću seriju podataka:

Broj krava (X_i)	**Broj gazdinstava (f_i)**
2	10
4	7
5	5
7	4
8	3
10	2
15	2

6. Traktor prelazi razdaljinu od parcele A do parcele D krećući se različitom brzinom. Razdaljinu od 1km od A do B je prešao krećući se brzinom od 4km/h, od B do C brzinom 3km/h i razdaljinu od 1,5 km brzinom od 6km/h. Izračunati prosečnu brzinu kretanja.

7. U jednom reonu na 100 stanovnika dolazi 5km^2 prostora, u drugom 6km^2, u trećem 10km^2 i u četvrtom 15 km^2. Kolika je prosečna gustina naseljenosti u ova četiri regiona?

8. Izračunati kvartile za sledeću seriju podataka:

X_i	28	66	70	47	58	60	67	38	51	68	45	63	56	65	63	66

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. Kod negrupisanih podataka i kod neintervalne serije distribucije frekvencija interval varijacije je razlika maksimalne i minimalne vrednosti obeležja. Kod intervalne distribucije frekvencija interval varijacije predstavlja razliku gornje granice poslednjeg i donje granice prvog grupnog intervala.

Primer 1. U toku jednog meseca osam porodica u uzorku je potrošilo sledeću količinu mleka (lit.):

Potrošnja mleka (lit.) *(X_i)*	10	12	9	15	25	20	28	17
Uređena serija								

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

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 aposlutno odstupanje za seriju podataka koja se odnosi na prinos pšenice (t/ha):

Prinos pšenice (t/ha) (X_i)		
2,5		
2,8		
2,9		
3,1		
3,2		
$\sum$		

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

Primer 3. Izračunati srednje apsolutno odstupanje za sledeću seriju podataka koja se odnosi na prinos pšenice (t/ha):

Prinos pšenice (t/ha) *(X_i)*	**Površina (ha) (f_i)**			
3,0	8			
3,2	12			
3,6	15			
3,7	9			
4,0	6			
$\sum$				

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

Standardna devijacija i varijansa

Standardna devijacija je kvadratni koren 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 i varijansu za sledeću seriju podatka koja se odnosi na nivo padavina tokom jedne nedelje (mm):

Nivo padavina (mm) *(X_i)*			
4			
6			
8			
11			
12			
13			
16			
$\sum$			

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

II način: $\sigma = \sqrt{\frac{\sum_{i=1}^{n} X_i^2 - \frac{\left(\sum_{i=1}^{n} X_i\right)^2}{n}}{n}} =$

Primer 5. Izračunati standardnu devijaciju i varijansu za sledeću seriju podataka koja se odnosi na broj zrna po klasu kod šenice sorte "San pastore":

Broj zrna *(X_i)*	**Broj klasova (f_i)**					
20	2					
25	3					
30	5					
35	17					
40	37					
45	37					
50	48					
55	31					
60	18					
65	2					
$\sum$						

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

II način: $\sigma = \sqrt{\frac{\sum_{i=1}^{k} X_i^2 f_i - \frac{\left(\sum_{i=1}^{k} X_i f_i\right)^2}{\sum_{i=1}^{k} f_i}}{\sum_{i=1}^{k} f_i}} =$

Izračunavanje varijanse metodom proizvoljnog početka i transformacije obeležja

Primer 5. Primenom metoda proizvoljnog početka i transformacije obeležja izračunati varijansu sledeće distribucije frekvencija:

Prinos grožđa (kg/čokotu) *(X_i)*	**Broj čokota (f_i)**							
2,00-2,39	20							
2,40-2,79	35							
2,80-3,19	48							
3,20-3,59	17							
$\sum$								

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. Uporediti varijabilnost dve serije ukoliko su njihove vrednosti sledeće:

$\overline{\mathrm{X}}_1 = 55{,}3$ (kg), $\sigma_1 = 0{,}35$(kg); $\overline{\mathrm{X}}_2 = 37{,}0$ (t), $\sigma_2 = 0{,}25$ (t). 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 distribuciju radnika na osnovu vrednosti dnevnice.

Vrednost dnevnice (\$) (*X_i*)	**Broj radnika (*f_i*)**			
45,1-55,0	5			
55,1-65,0	10			
65,1-75,0	25			
75,1-85,0	35			
85,1-95,0	15			
95,1-105,0	10			
$\sum$	**100**			

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

Zadaci za vežbu:

1. Dati su podaci koji se odnose na prinos šljive po stablu (kg) u jednom voćnjaku.

Prinos (kg)
4
5
6
10
14
17
21

Izračunati apsolutne pokazatelje varijabiliteta.

2. Dati su podaci uzorka od 2.400 individualnih gazdinstava kod kojih je ispitivana površina pod voćnjacima i vinogradima:

Površina (ha)	**Broj gazdinstava**
0,6	140
0,8	240
1,0	500
1,2	700
1,4	160
1,6	455
1,8	110
2,0	95

Izračunati apsolutne pokazatelje varijabiliteta i standardizovano odstupanje za vrednost obeležja X=1,6.

3. Izračunati varijansu metodama proizvoljnog početka i transformacije obeležja na osnovu sledeće serije podataka:

Grupni intervali	**Frekvencija (f_i)**
0,1-2,0	5
2,1-4,0	7
4,1-6,0	10
6,1-8,0	5
8,1-10,0	3

4. Dati su podaci o visini 5 biljaka (cm): 50, 55, 60, 70, 75, 80. Izračunati varijansu datog obeležja. Uporediti varijabilitet datog obeležja i obeležja $Y=2X+2$.

5. Uporediti varijabilnost 2 serije podataka ukoliko su njihove vrednosti sledeće:
$\bar{X}_1 = 60\ (kg)$; $\sigma_1^2 = 24\ (kg)^2$; $\bar{X}_2 = 30\ (cm)$; $\sigma_2 = 5\ (cm)$.

POKAZATELJI OBLIKA DISTRIBUCIJE

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

- Mera asimetričnosti – I Pirsonov koeficijent (β_1)
- Mera spljoštenosti – II Pirsonov koeficijent β_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 ispravnih proizvoda (000 komada) proizvedenih u jednom pogonu utvrditi pokazatelje oblika distribucije i dati odgovarajući komentar.

Broj ispravnih proizvoda (X_i)	$(\boldsymbol{X_i - \bar{X}})$	$(\boldsymbol{X_i - \bar{X}})^{\boldsymbol{2}}$	$(\boldsymbol{X_i - \bar{X}})^{\boldsymbol{3}}$	$(\boldsymbol{X_i - \bar{X}})^{\boldsymbol{4}}$
4				
12				
10				
21				
6				
18				
$\sum$ 71				

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

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

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

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

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

Primer 2. Utvrditi oblik raspodele podataka koji se odnose na broj zastoja po smeni na jednoj proizvodnoj liniji.

Broj zastoja ***(X_i)***	**Broj smena** ***(f_i)***	$(\boldsymbol{X_i} - \overline{\boldsymbol{X}})$	$(\boldsymbol{X_i} - \overline{\boldsymbol{X}})^2$	$\boldsymbol{f_i}(\boldsymbol{X_i} - \overline{\boldsymbol{X}})^2$	$\boldsymbol{f_i}(\boldsymbol{X_i} - \overline{\boldsymbol{X}})^3$	$\boldsymbol{f_i}(\boldsymbol{X_i} - \overline{\boldsymbol{X}})^4$
0	**70**					
1	**50**					
2	**35**					
3	**15**					
4	**10**					
$\sum$						

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

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

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

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

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

Primer 3. Dati su podaci o broju efektivno ostvarenih sati rada u toku meseca i broj radnika. Utvrditi oblik datog rasporeda.

Broj sati ***(X_i)***	**Broj radnika** ***(f_i)***	$(X_i - \bar{X})$	$(X_i - \bar{X})^2$	$f_i(X_i - \bar{X})^2$	$f_i(X_i - \bar{X})^3$	$f_i(X_i - \bar{X})^4$
130-139	**3**					
140-149	**5**					
150-159	**12**					
160-169	**20**					
170-179	**26**					
180-189	**10**					
190-199	**4**					
$\sum$	**80**					

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

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

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

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

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

Primer 4. Za datu seriju podataka utvrditi aritmetičku sredinu, modus i medijanu, a zatim na osnovu izračunatih pokazatelja centralne tendencije opisati oblik date raspodele.

Masa brojlera (kg) ***(X_i)***	**Broj brojlera** ***(f_i)***		
0,7	60		
0,8	120		
1,0	300		
1,3	440		
1,4	300		
1,5	120		
1,7	60		
$\sum$	1400		

Zadaci za vežbu:

1. Na osnovu podataka o broju bodova na testu znanja 100 ispitanika dobijeno je da je prosečan broj bodova 20. Najveći broj ispitanika je imao 25 bodova. Ako je na osnovu dobijenih rezultata izračunato $\sum f(x-\bar{x})^2 = 3.500$ $\sum f(x-\bar{x})^4 = 500.000$ ispitati oblik raspodele broja bodova.

2. Na osnovu distribucije frekvencije veličine poseda (ha) 40 individualnih gazdinstava izračunati pokazatelje asimetričnosti i oblika distribucije.

Veličina poseda (ha)	**Broj gazdinstava**
1,1-3	7
3,1-5	12
5,1-7	5
7,1-9	13
9,1-11	3

3. Na osnovu datih podataka koji se odnose na dužinu staža zaposlenih u jednom preduzeću izračunati pokazatelje centralne tendencije i preko njih ispitati oblik date raspodele.

Godine staža	**Broj zaposlenih**
1-5	20
6-10	35
11-15	50
16-20	10
21-25	10
26-30	5

TEORIJSKE DISTRIBUCIJE

Nasuprot empirijskim distribucijama koje se formiraju na osnovu opserviranih vrednosti numeričkog obeležja, postoje teorijske distribucije koje se mogu očekivati u skladu s iskustvom ili na osnovu nekih pretpostavki. Teorijske distribucije mogu biti prekidne i neprekidne.

Prekidne teorijske distribucije

Binomna distribucija – BD

Binomna distribucija je jedna od najvažnijih prekidnih teorijskih distribucija. U osnovi binomne distribucije su sukcesivni događaji koji imaju dva ishoda. Binomna distribucija je definisana preko Bernulijevog eksperimenta.

Broj "uspeha" u n ponavljanja Bernulijevog eksperimenta je slučajna promenljiva X koja ima binomnu raspodelu. Kako je broj "uspeha" svaki ceo broj u intervalu od 0 do n, vrednosti slučajne promenljive koja ima binomnu raspodelu su X : 0, 1, 2, 3, ...n. Broj modaliteta slučajne promenljive je $n+1$. Verovatnoća $P(X=i)$ za i=0,..., n data je izrazom:

$$p_{(i)} = \binom{n}{i} \cdot p^i q^{n-i}$$

Primer 1. Ako jedna mašina proizvodi 20% defektnih proizvoda odrediti verovatnoću da između četiri slučajno odabrana proizvoda bude: a) 1; b) 0; c) najviše dva defektna.

Primer 2. Šta je verovatnije: da se u pet bacanja homogene kocke dobije jedna šestica ili u šest bacanja da se dobije jedna šestica?

Primer 3. Jaja su pakovana u kutije od 12 komada. Verovatnoća da je jaje razbijeno je 0,35. Izračunati verovatnoću da je u slučajno odabranoj kutiji a) 4 razbijena jajeta; b) najviše 2 razbijena jajeta.

Primer 4. Izračunati verovatnoću da se dobije zbir 9 a) dva puta; b) najmanje dva puta; u 5 bacanja dve kocke.

Primer 5. Prirodna smrtnost komaraca u toku prvih 5 dana iznosi 0.3. Izračunati verovatnoću da u grupi od 3 insekta: a) nijedan neće preživeti 5 dana; b) da će bar dva preživeti 5 dana.

Primer 6. Ako su n=10 i p=0,6 parametri binomne distribucije, izračunati aritmetičku sredinu, varijansu, modus, prvi i drugi Pirsonov koeficijent.

Primer 7. Ispitati saglasnost date empirijske distribucije, koja se odnosi na broj ženskih jedinki (x) u 100 legala od po 5 životinja, sa binomnom distribucijom. Uporediti moduse, varijanse i koeficijente varijacije date empirijske i teorijske binomne distribucije.

Broj ženskih jedinki *(X_i)*	**Frekvencija**	
0	3	
1	16	
2	36	
3	32	
4	11	
5	2	
$\sum$	100	

Poasonova distribucija -PD

Poasonova distribucija je teorijska distribucija koja se odnosi na prekidna obeležja. Vrednost obeležja X su celi nenegativni brojevi 0, 1, 2, 3, ...n,...

Verovatnoće Poasonove distribucije zavise od jednog parametra i to je parametar m. Verovatnoće Poasonove distribucije date su izrazom:

$$p_{(i)} = e^{-m} \cdot \frac{m^i}{i!}$$

Poasonova raspodela je raspodela broja događaja koji se retko javljaju.

Primer 1. U toku jednog sata stigne prosečno dve poruke na određenu internet adresu. Izračunati verovatnoću da u toku jednog sata: a) neće stići nijedna poruka; b) da će stići tačno tri poruke; c) da će stići najviše tri poruke.

Primer 2. U telefonskoj centrali u toku jednog sata bilo je 60 poziva. Izračunati verovatnoću da u toku dva minuta: a) nije bilo ni jednog poziva; b) bilo je najviše tri poziva.

Primer 3. Proizvodi jedne velike serije koja sadrži 0,7% škartova pakuju se u kutije od po 100 proizvoda. Izračunati verovatnoću da je slučajno izabrana kutija : a) bez škartova; b) sa dva ili više škartova.

Primer 4. Poznato je da među ljudima ima 1% levorukih. Neka je slučajna promenljiva X: broj levorukih u grupi od 200 slučajno izabranih ljudi. Odrediti zakon raspodele slučajne promenljive X. Izračunati verovatnoću da se u grupi nalaze dve levoruke osobe.

Primer 5. Broj radnika u jednom preduzeću koji zakasne na posao predstavlja slučajnu promenljivu sa Poasonovom raspodelom čiji je parametar m =1.
Izračunati verovatnoću da su na posao zakasnila:

a) dva radnika;
b) bar dva radnika.

Odrediti varijansu, modus, prvi i drugi Pirsonov koeficijent slučajne promenljive.

Primer 6. Da bi se ocenio rizik pojave određene bolesti odabrano je na slučajan način 1000 uzoraka od po 10 životinja i registrovan je broj obolelih životinja X:

X_i	f_i	
0	500	
1	340	
2	120	
3	30	
≥ 4	10	
$\sum$	**1000**	

a) Ispitati saglasnost empirijske i teorijske Poasonove distribucije.
b) Uporediti moduse empirijske i teorijske distribucije.
c) Uporediti varijanse empirijske i teorijske distribucije.
d) Ispitati simetričnost i spoljoštenost Poasonove distribucije.

Normalna distribucija-ND

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.

Normalna raspodela je neprekidna teorijska distribucija koja zavisi od dva parametra aritmetičke sredine μ i standardne devijacije σ.

Primena Normalne raspodele:

1. $P\,(-a < X < a) = 2F(a)$

2. $P(a < X < b) = F(b) - F(a)$

3. $P\ (-b < X < a) = F(a) + F(b)$

4. $P(-b < X < -a) = F(-b) - F(-a) = F(b) - F(a)$

Primer 1. Neka slučajna promenljiva X ima normalnu raspodelu sa parametrima $\mu= 0$ i standardnom devijacijom $\sigma=1$. Izračunati verovatnoću da je slučajna promenljiva X između -1,5 i 2,5.

Primer 2. Neka slučajna promenljiva X ima normalnu raspodelu sa parametrima $\mu= 0$ i standardnom devijacijom $\sigma=1$. Izračunati verovatnoću da je slučajna promenljiva X između 0,5 i 2.

Primer 3. Neka slučajna promenljiva X ima normalnu raspodelu sa parametrima $\mu= 0$ i standardnom devijacijom $\sigma=1$. Izračunati verovatnoću da je slučajna promenljiva X između -0,5 i -2.

Primer 4. Neka slučajna promenljiva X ima normalnu raspodelu sa parametrima $\mu= 0$ i standardnom devijacijom $\sigma=1$. Izračunati verovatnoću da je slučajna promenljiva X manja od -2.

Primer 5. Data je slučajna promenljiva X koja ima normalnu raspodelu čija je očekivana vrednost 12 i standardna devijacija 2. Izračunati verovatnoće:

a) $P (8 < X < 9)$
b) $P (-2 < X < -1{,}5)$
c) $P (X < 6)$

Primer 6. Neka slučajna promenljiva X ima normalnu raspodelu sa parametrima $\mu = 0$ i standardnom devijacijom $\sigma=1$. Odrediti vrednost a tako da je $P (X < a) = 0{,}85$.

Zadaci za vežbu:

1. Verovatnoća da je neko lice u porodici zaposleno iznosi 0,6. Ako posmatramo 200 tročlanih porodica, izračunati sve moguće verovatnoće i odrediti oblik datog rasporeda. Utvrditi broj porodica u kojima su svi članovi zaposleni.

2. U proizvodnji jednog prehrambenog proizvoda 5% proizvoda ne odgovara standardu. Izračunati verovatnoću da će prost slučajan uzorak od 15 proizvoda sadržati 3 neispravna proizvoda, pretpostavljajući da je izbor uzorka sa ponavljanjem. Koliki je najverovatniji broj neispravnih proizvoda?

3. Slučajna promenljiva X uzima vrednosti 0, 1, 2, 3, 4. Na osnovu realizacije uzorka dobijena je distribucija frekvencija:

X_i	0	1	2	3	4
f_i	67	122	94	38	3

Ispitati saglasnost date empirijske distribucije frekvencija sa teorijskom binomnom distribucijom poređenjem apsolutnih empirijskih i teorijskih frekvencija.

4. U toku jednog sata stigne prosečno dve poruke na određenu internet adresu. Izračunati verovatnoću da u toku jednog sata: a) neće stići nijedna poruka; b) da će stići tačno tri poruke; c) da će stići najviše tri poruke.

5. Za transport kukuruza u jednom danu angažovano je 50 traktora. Raspored traktora prema broju kvarova dat je u tabeli. Izračunati Poasonove verovatnoće, apsolutne frekvencije i Pirsonove koeficijente.

Broj kvarova	0	1	2	3	4
Broj traktora	21	18	7	3	1

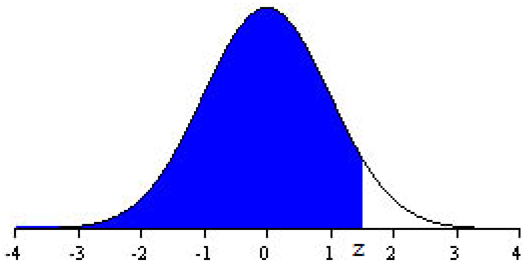
6. Slučajna promenljiva X ima normalnu raspodelu sa aritmetičkom sredinom 0 i varijansom 1. Odrediti konstantu a, tako da je $P(-a<X<a)=0,6528$.

7. Ako slučajna promenljiva ima aritmetičku sredinu 15 i varijansu 9, izračunati verovatnoću da je slučajna promenljiva veća od 9 i manja od 24.

8. Visina biljke soje ima normalnu raspodelu čija je aritmetička sredina 75cm i standardna devijacija 5 cm. Izračunati verovatnoću da će visina slučajno odabrane biljke soje biti veća od 80 cm.

9. Broj posetilaca poljoprivrednog sajma u periodu od 12-14h svakog dana u toku trajanja sajma ima normalan raspored. Prosečan broj posetilaca u navedenom periodu je 350, a varijansa je 400. Kolika je verovatnoća da broj posetilaca bude veći od 310, a manji od 330?

10. Slučajna promenljiva Z ima normalnu raspodelu sa parametrima μ=0 i σ=1.



Odrediti z tako da je verovatnoća da je slučajno promenljiva Z manja od z 0,8531.

11. Nedeljna tražnja jednog proizvoda ima normalnu raspodelu čija je aritmetička sredina 800 pakovanja i varijansa 5625. Izračunati verovatnoće da je nedeljna tražnja:
a) manja od 959 pakovanja;
b) od 650 do 950 pakovanja.

DISTRIBUCIJA SREDINA UZORAKA

Broj prostih slučajnih uzoraka veličine n jedinica koji može da se izabere iz jednog osnovnog skupa veličine N jedinica dat je sledećim izrazom:

- Uzorci sa ponavljanjem: $k = N^n$
- Uzorci bez ponavljanja: $k = \binom{N}{n} = \frac{N!}{n!(N-n)!} = \frac{N(N-1)(N-2)\cdot\ldots\cdot(N-n+1)}{n!}$

Primer 1. Na osnovu osnovnog skupa od pet elemenata formirati sve kombinacije uzoraka od po dve jedinice bez ponavljanja. Uporediti parametre osnovnog skupa i distribucije sredina uzoraka.

X_i	
2	
4	
6	
8	
10	

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 varijavilitet 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_{α}).

Primer 1. Iz osnovnog skupa od 2000 jedinica izabran je prost slučajan uzorak bez ponavljanja od 100 jedinica i izračunata je aritmetička sredina $\bar{X} = 30.$ Odrediti 95% interval poverenja za aritmetičku sredinu osnovnog skupa, ako je varijansa osnovnog skupa $\sigma^2 = 36$.

Primer 2. Za ocenu prosečnog prinosa kukuruza izvršeno je istraživanje na osnovu uzorka od 16 (ha) od ukupno zasejanih 300 (ha). Dobijeni rezultati su dati u tabeli:

Prinos (00kg/ha) (X_i)	**Površina (ha)** (f_i)		
60,1-62	4		
62,1-64	7		
64,1-66	5		
$\sum$			

Oceniti prosečan i ukupn prinos kukuruza sa pouzdanošću od 99 procenata ako je od ranije poznat varijabilitet osnovnog skupa $\sigma^2 = 0{,}14$ (00kg/ha)^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 aritemtičke 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 3. Rezultati uzorka koji se odnose na ispitivanje uticaja jedne vrste pesticida na klijavost semena hibrida kukuruza (%) dati su u tabeli. Odrediti 95% i 99% interval poverenja za prosečanu klijavost semena.

Klijavost semena (%) (X_i)	X_i^2
96	
98	
90	
89	
92	
94	
90	
88	
90	
91	

Primer 4. Oceniti prosečan i ukupan prinos šećerne repe na osnovu prinosa dobijenih sa 20 hektara, u regionu koji ima 480 hektara pod ovom kulturom.

Prinos (vag/ha) ($\boldsymbol{X_i}$)	**Površina** ($\boldsymbol{f_i}$)	$\boldsymbol{X_i f_i}$	$\boldsymbol{X_i^2 f_i}$
4,4	3		
4,6	4		
5,2	7		
5,4	4		
5,9	2		
Σ			

Interval poverenja za ocenu nepoznate proporcije osnovnog skupa

Primer 1. U jednom preduzeću zaposleno je ukupno 5000 radnika, uzet je prost slučajan uzorak od 250 radnika, od tog broja 102 radnika ima lični dohodak manji od 50.000,00 dinara, a preostalih 148 radnika veći. Utvrditi učešće radnika sa ličnim dohotkom manjim od 50.000,00 dinara u posmatranom preduzeću.

Primer 2. Rezultati uzorka vezani za prinos šljive dati su u tabeli. Utvrditi zastupljenost stabala sa prinosom od 10kg u proizvodnji.

Prinos šljive (kg) (X_i)	**Broj stabala** (f_i)
5	100
8	200
10	300
11	200
15	200
Ukupno	

Zadaci za vežbu:

1. Data je potrošnja šećera u 150 slučajno izabranih domaćinstava jednog naselja. Utvrditi prosečnu mesečnu potrošnju šećera po domaćinstvu u ovom naselju, ako je σ^2 u populaciji 0,9, kao i ukupnu mesečnu potrošnju celog naselja ako u njemu ima 3000 domaćinstava.

Potrošnja šećera (kg)	Broj domaćinstava
2	5
3	10
4	20
5	40
6	50
7	25

2. Na osnovu podataka uzorka individualnih gazdinstava kod kojih je posmatrana površina pod voćnjacima i vinogradima utvrditi prosečnu vrednost i razmak poverenja u kome će se kretati površina pod voćnjacima i vinogradima u osnovnom skupu.

Površina pod voćnjacima i vinogradima	Broj gazdinstava
0,6	200
0,8	400
1,0	500
1,2	600
1,4	200
1,6	100

3. Iz osnovnog skupa od 1000 gazdinstava izabran je uzorak bez ponavljanja radi ispitivanja proporcije broja tovnih svinja:

Broj tovnih svinja	Broj gazdinstava
0	4
1	5
2	11
3	37
4	30
5	10
8	3

a) Oceniti proporciju gazdinstava koja nemaju tovne svinje.
b) Oceniti proporciju gazdinstava koja imaju više od 4 tovne svinje.

4. Slučajnim izborom na teritoriji jedne opštine izmeren je prinos kukuruza u mc po ha na 400 parcela od po 1ha. Koliki prosečan i ukupan prinos može da se očekuje u celoj opštini ako je kukuruzom zasejano 80.000ha? Na kolikoj površini može da se očekuje prinos veći od 45 mc/ha?

Prinos (mc/ha)	**Broj parcela**
25,1-35,0	60
35,1-45,0	150
45,1-55,0	100
55,1-65,0	90

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 slučajna uzorka – test značajnosti razlike dve sredine;
3. Upoređivanje više od dve aritmetičke sredine iz više od dva nezavisna slučajna uzorka – metod analize varijanse.

1. Test značajnosti jedne sredine

Primer 1. Rezultati jednog uzorka, n=30, pokazali su da je prosečan prinos graška 3,6 t/ha. Uporediti ovaj prinos sa prinosom standardnih sorti koji iznosi 3,9 t/ha i ima varijansu 1,44 $(t/ha)^2$.

Primer 2. Pri ispitivanju ispravnosti jedne vrste proizvoda (%) u proizvodnom pogonu, dobijeni su sledeći rezultati:

Ispravni proizvodi (%) (X_i)	X_i^2
66	
67	
70	
69	
72	
75	
72	
68	
65	
73	

Da li može da se pretpostavi da je prosečan procenat ispravnih proizvoda 80%?

Primer 3. Raspodela individualnih gazdinstava prema veličini u jednom uzorku data je u tabeli. Utvrditi da li je prosečna veličina gazdinstva na posmatranoj teritoriji 5ha.

Veličina gazdinstva (ha)	**Broj gazdinstava (f_i)**	X_i	$X_i f_i$	$X_i^2 f_i$
0,01-2	15			
2,01-4	30			
4,01-6	35			
6,01-8	13			
8,01-10	7			
Σ				

2. Test značajnosti razlike dve aritmetičke sredine

Primer 1. Cilj eksperimenta je da se utvrdi da li postoji statistički značajna razlika u prosečnom prinosu dve sorte pšenice. Izabrana su dva nezavisna slučajna uzorka. Na osnovu uzorka od 25 (ha) dobijen je prosečni prinos prve sorte 6,2 (t/ha), dok je na osnovu uzorka od 36 (ha) dobijen prosečni prinos druge sorte 4,5 (t/ha). Ako su poznati varijabiliteti posmatranih sorti $\sigma_1^2 = 1$ i σ_2^2=0,75, testirati nultu hipotezu da ne postoji statistički značajna razlika u prosečnom prinosu posmatranih sorti.

Primer 2. Utvrditi da li postoji statistički značajna razlika između prosečnih prinosa dve sorte suncokreta na osnovu podataka iz tabele.

Prinos sorte A (t/ha)	**Prinos sorte B (t/ha)**		
1,8	2,2		
1,9	2,4		
2,0	2,3		
2,2	2,5		
2,3	2,6		
2,2	2,5		
2,1	2,1		
2,0	2,3		

Primer 3. U jednom eksperimentu se ispituje uticaj dva primenjena tretmana na eksperimentalni rezultat. Odabrana su dva nezavisna prosta slučajna uzorka, primenjeni tretmani i dobijeni rezultati dati su u tabeli:

Tretman A		**Tretman B**					
X_1	f_1	X_2	f_2				
1,40	4	1,31	7				
1,38	5	1,37	4				
1,41	11	1,39	15				
1,35	3	1,33	10				
1,42	8	1,35	4				
1,37	9	1,36	5				

Da li može da se prihvati nulta hipoteza da uzorci pripadaju osnovnim skupovima sa jednakim aritmetičkim sredinama?

Zadaci za vežbu:

1. Deklarisana težina pakovanja šećera je 1kg. Na osnovu uzorka od 35 pakovanja čija je prosečna težina 0,96 kg, procenjena je varijansa osnovnog skupa $\sigma^2 = 0{,}09$ kg². Da li pakovanje šećera zadovoljava deklarisanu težinu?

2. Na sedam hektara ostvaren je sledeći prinos zelenog zrna graška: 6,8 6,9 7,1 7,2 7,2 7,4 7,5 (t/ha). Da li je postignuti prosečan prinos na nivou očekivanog prosečnog prinosa od 7,6 (t/ha)?

3. Prinos pšenice (t/ha) sa odabranih 40 ha parcela čija je površina 400 ha dat je u tabeli. Da li se može očekivati da će prosečan prinos pšenice biti na nivou od 3,5 (t/ha)?

Prinos (t/ha)	**Površina (ha)**
4,9	4
4,5	6
4,3	10
3,8	7
3,4	5
3,0	8

4. Dati su podaci o prosečnom trajanju studija (godina) na sedam fakulteta u 2017. i 2018. godini. Da li je uspeh studiranja iskazan prosečnim trajanjem studija u obe godine isti? U kojoj godini dužina studiranja ima veći varijabilitet?

Fakultet	**Vreme studiranja**	
	2017	**2018**
Arhitektura	8,1	6,9
Građevina	8,7	7,6
Defektologija	6,1	6,5
Medicina	7,8	7,0
Poljoprivreda	7,7	6,9
Ekonomija	6,3	6,2
Fakultet za menadžment	4,5	5,0

5. Na osnovu podataka sledeća dva uzorka ustanoviti da li postoji statitički značajna razlika u prosečnoj dnevnoj potrošnji mleka poljoprivrednih i nepoljoprivrednih domaćinstava:

Poljoprivredna domaćinstva		**Nepoljoprivredna domaćinstva**	
Potrošnja (l)	**Broj domaćinstava**	**Potrošnja (l)**	**Broj domaćinstava**
1	4	1	10
2	8	2	8
3	6	3	2
4	2	4	0

3. Analiza varijanse

Primer 1. Na relativno homogenom polju, podeljenom na 20 jednakih parcela, zasejane su 4 sorte pšenice, svaka na 5 parcela. Razmeštaj sorti po parcelama je potpuno slučajan. Rezultati eksperimenta koji se odnose na prinos (00kg/ha) dati su u sledećoj tabeli:

Redni broj	**Sorte**			
	I	**II**	**III**	**IV**
1	32,3	33,3	30,8	29,3
2	34,0	33,0	34,3	26,0
3	34,3	36,3	35,3	29,8
4	35,0	36,8	32,3	28,0
5	36,5	34,5	35,8	28,8

a) Da li postoji statistički značajna razlika u prosečnom prinosu pojedinih sorti?
b) Testirati značajnost razlike parova tretmana primenom: t-testa, testa najmanje značajne razlike (NZR-testa) i višestrukog intervalnog (Dankanovog) testa.

Nulta hipoteza:

Alternativna hipoteza:

Izvori varijacije	**Stepeni slobode**	**Sume kvadrata**	**Sredine suma kvadrata (varijanse)**	**F- odnos**	**F – tablično (r_1=k-1; r_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) –

<u>Sume kvadrata</u>

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

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

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

$$Q_P = Q - Q_T =$$

<u>Sredine suma kvadrata</u>

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

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

t-test

Test najmanje značajne razlike

Dankanov test

Primer 2. Da bi se ispitala značajnost razlike aritmetičkih sredina četiri tretmana, izveden je ogled po planu potpuno slučajnog rasporeda i dobijeni su sledeći rezultati:

Ponavljanja	**Sorte**			
	A	**B**	**C**	**D**
	3	7	3	10
	2	8	2	12
	4	4	1	8
	3	10	2	5
	1	6	4	12
	5		2	10
			3	9
			1	

a) Da li postoji statistički značajna razlika aritmetičkih sredina tretmana?
b) Testirati značajnost razlike parova tretmana primenom t-testa.

Nulta hipoteza:

Alternativna hipoteza:

Izvori varijacije	**Stepeni slobode**	**Sume kvadrata**	**Sredine suma kvadrata (varijanse)**	**F- odnos**	**F – tablično (r_1=k-1; r_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) –

<u>Sume kvadrata</u>

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

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

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

$$Q_P = Q - Q_T =$$

<u>Sredine suma kvadrata</u>

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

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

t-test

Zadaci za vežbu:

1. U cilju ocene kvaliteta jednog proizvoda kod 5 proizvođača na slučajan način je izabrano po 4 proizvoda od svakog proizvođača i njihov kvalitet je ocenjen sa ocenom od 1 do 10. Dobijeni su sledeći rezultati:

Proizvođači				
I	**II**	**III**	**IV**	**V**
10	9	8	7	8
10	10	8	8	9
9	5	5	4	6
8	6	7	4	4

a) Da li postoji statistički značajna razlika u kvalitetu proizvoda u zavisnosti od proizvođača?
b) Primenom NZR testa utvrditi između kojih proizvođača postoji statistički značajna razlika?

2. Menadžer u jednom preduzeću hoće da utvrdi efikasnost četiri programa obuke zaposlenih. Svaki od programa primenjen je na grupi od 8 slučajno odabranih kandidata. Rezultati testa su dati u tabeli:

Programi	**Ponavljanja**							
A	66	74	82	75	73	97	87	78
B	72	51	59	62	74	64	78	63
C	61	60	57	60	81	55	70	71
D	63	61	76	84	58	65	69	80

a) Da li se prosečan broj poena na testu statistički značajno razlikuje u zavisnosti od primenjenog metoda obuke?
b) Primenomm t-testa utvrditi između kojih metoda obuke postoji statistički značajna razlika.

3. Na osnovu podataka koji se odnose na prinos (00kg/ha) tri sorte pšenice, dobijeni su sledeći rezultati:

Sorte		
A	**B**	**C**
3	9	1
7	12	2
7	11	6
6	8	4
2	5	
	9	

a) Da li postoji statistički značajna razlika u prosečnim prinosima sorti A, B i C?
b) Utvrditi između kojih sorti postoji statistički značajna razlika.

Testovi proporcija

Postoje sledeći osnovni testovi za testiranja proporcija:

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

Test značajnosti jedne proporcije

Primer 1. U određenom proizvodnom procesu u toku dana je proizvedeno 100 artikala, od kojih je 5 defektno. Testirati nultu hipotezu da se proporcija defektnih artikala statistički značajno razlikuje od pretpostavljene vrednosti p=0,3.

Test značajnosti razlike dve proporcije

Primer 2. Iz dva naselja na slučajan način je iz prvog odabrano 160 domaćinstava od kojih 116 proizvodi pšenicu, a iz drugog 180 domaćinstava, od kojih 144 proizvodi pšenicu. Da li se može zaključiti da je zastupljenost proizvodnje pšenice ista u oba posmatrana naselja?

Zadaci za vežbu:

1. U uzorku veličine n=100 izniklo je 80 biljaka. Testirati značajnost razlike proporcije izniklih biljaka u uzorku i pretpostavljene proporcije u osnovnom skupu koja iznosi 0,9.

2. Ispitivana su individualna gazdinstva o broju aktivnih poljoprivrednika. Rezultati ispitivanja dati su u tabeli:

Broj aktivnih poljoprivrednika	Broj gazdinstava
0	15
1	35
2	40
3	60
4	30
5	20

Utvrditi proporciju gazdinstava koja imaju po tri aktivna poljoprivrednika. Da li se može pretpostaviti da će u osnovnom skupu proporcija date karakteristike iznositi 0,37?

3. Na osnovu uzorka od 200 proizvoda proizvođala A, nađeno je 12 neispravnih proizvoda. U uzorku od 100 proizvoda proizvođača B nađeno je 6 neispravnih proizvoda. Da li se kvalitet proizvoda ova dva proizvođača, izražen učešćem škarta u proizvodnji, statistički značajno razlikuje?

REGRESIONA I KORELACIONA ANALIZA

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

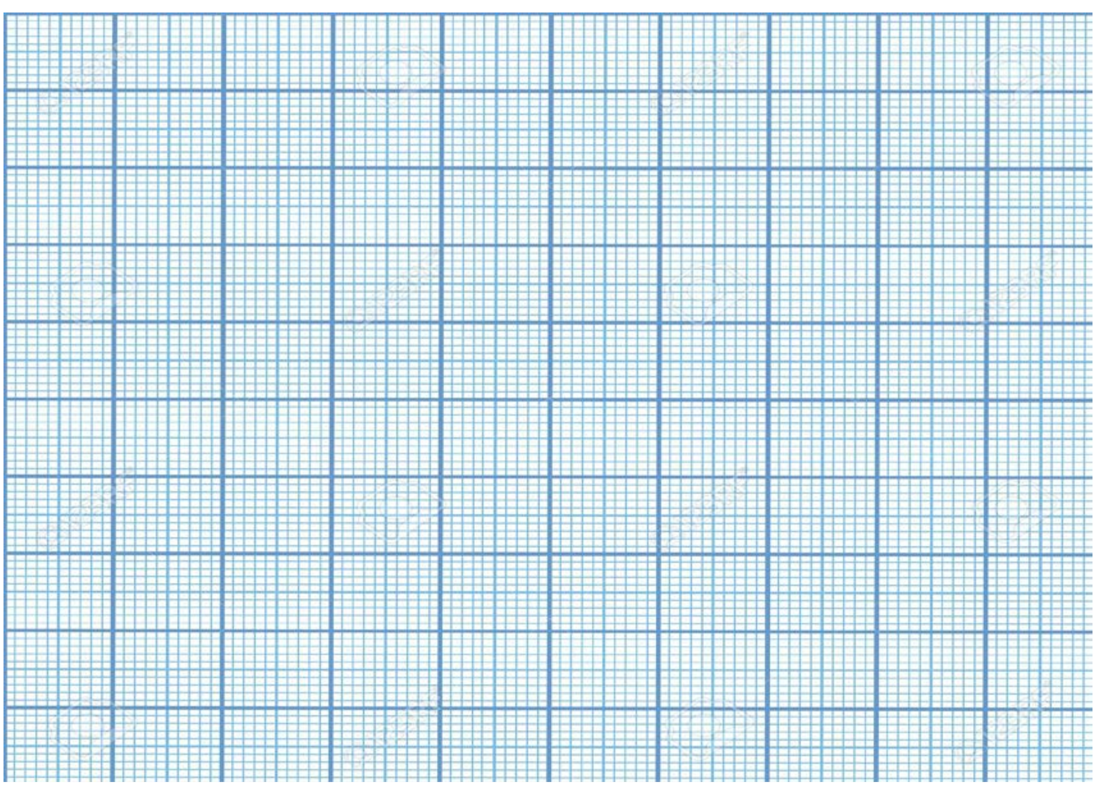
Primer 1. Dati su podaci o ponudi robe (00 komada) i cenama (din/komadu):

Ponuda (Y)	12	14	10	13	17	12	11	15
Cene (X)	5	11	7	8	11	7	6	9

Na osnovu datih parova vrednosti izvesti sledeća računanja:

a) Formirati dijagram rasturanja;
b) Oceniti regresioni model, izračunati ocenjene vrednosti zavisno promenljive i ucrtati liniju regresije na dijagramu rasturanja;
c) Izračunati standardnu grešku regresije;
d) Testirati značajnost regresionog modela primenom analize varijanse;
e) Testirati značajnost koeficijenta pravca regresije. Odrediti interval poverenja za koeficijent pravca regresije, ako je α=0,05;
f) Izračunati kolika vrednost ponude može da se očekuje ako je cena 20 (din/komadu);
g) Izračunati koeficijent korelacije i testirati njegovu značajnost;
h) Izračunati koeficijent determinacije i testirati njegovu značajnost.

Dijagram rasturanja:



Radna tabela:

X	Y								

$$\bar{X} = \frac{\sum X}{n} \qquad \bar{Y} = \frac{\sum Y}{n}$$

$$b = \frac{\sum (X - \bar{X})(Y - \bar{Y})}{\sum (X - \bar{X})^2}$$

$$b = \frac{\sum XY - \frac{\left(\sum X \sum Y\right)}{n}}{\sum X^2 - \frac{\left(\sum X\right)^2}{n}}$$

$$a = \bar{Y} - b\bar{X}$$

$$s_e = \sqrt{\frac{\sum (Y - \hat{Y})^2}{n - 2}}$$

$$Q = \sum (Y - \bar{Y})^2$$

$$Q = \sum Y^2 - \frac{\left(\sum Y\right)^2}{n}$$

$$Q_R = \frac{\left[\sum (X - \bar{X})(Y - \bar{Y})\right]^2}{\sum (X - \bar{X})^2}$$

$$Q_R = \frac{\left[\sum XY - \frac{\left(\sum X \sum Y\right)}{n}\right]^2}{\sum X^2 - \frac{\left(\sum X\right)^2}{n}}$$

$$Q_{VR} = Q - Q_R$$

$$F = \frac{Q_R}{\frac{Q_{VR}}{n-2}}$$

Tabela analize varijanse

Izvori varijacije	Stepeni slobode	Sume kvadrata	Sredine suma kvadrata (varijanse)	F-odnos	F-tablično	
					0,05	0,01
Regresija	1	Q_R	Q_R	Q_R/s_e^2		
Pogreška	*n-2*	Q_{VR}	s_e^2			
Ukupno	*n-1*	Q				

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

$$s_b = \sqrt{\frac{s_e^2}{\sum (X - \bar{X})^2}}$$

$$r = \frac{\sum \left(X - \bar{X} \right)\left(Y - \bar{Y} \right)}{\sqrt{\sum \left(X - \bar{X} \right)^2 \sum \left(Y - \bar{Y} \right)^2}}$$

$$r = \frac{\sum XY - \frac{\left(\sum X \sum Y \right)}{n}}{\sqrt{\left[\sum X^2 - \frac{\left(\sum X \right)^2}{n} \right] \cdot \left[\sum Y^2 - \frac{\left(\sum Y \right)^2}{n} \right]}} =$$

$$t = \frac{r}{s_r} \qquad s_r = \sqrt{\frac{1 - r^2}{n - 2}}$$

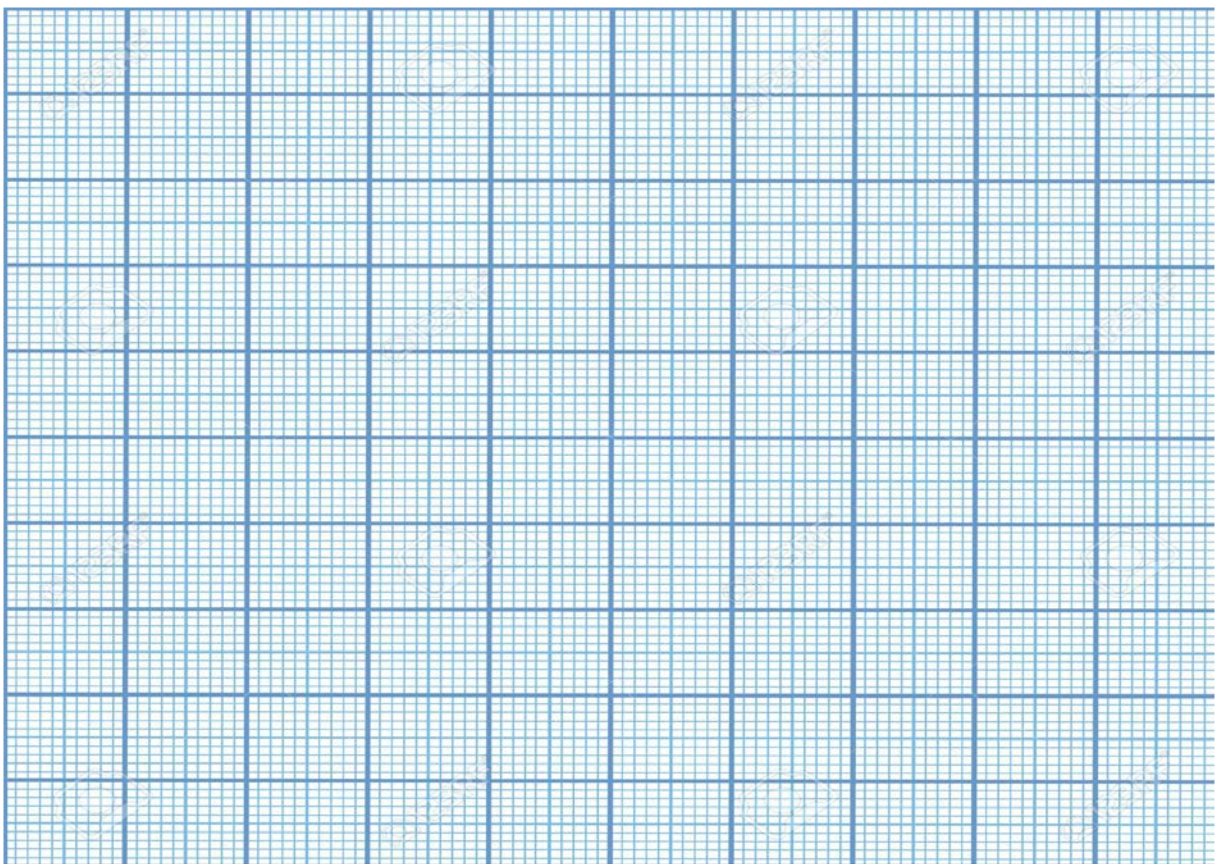
Primer 2. Na 5 parcela istog kvaliteta vrednosti prinosa suncokreta (t/ha) i utrošak mineralnog đubriva (kg) dati su u tabeli:

Prinos suncokreta (t/ha) (Y)	4	2	5	8	10
Količina đubriva (kg) (X)	609	279	709	859	909

Na osnovu datih parova vrednosti izvesti sledeća računanja:

a) Formirati dijagram rasturanja;
b) Oceniti regresioni model, izračunati ocenjene vrednosti zavisno promenljive i ucrtati liniju regresije na dijagramu rasturanja;
c) Izračunati standardnu grešku regresije;
d) Testirati značajnost regresionog modela u celini;
e) Izračunati vrednost koeficijenta korelacije, determinacije i nedetrminacije;
f) Koliki prinos može da se očekuje ukoliko je utrošena količina mineralnog đubriva 750kg?

Dijagram rasturanja:



Radna tabela:

X	**Y**						

Zadaci za vežbu:

1. Dati su podaci o mesečnim zaradama (000din) i broju godina školovanja 19 slučajno odabranih zaposlenih jedne firme:
2.

Broj godina školovanja	6	8	11	12	12	13	14	16	16	21
Mesečna zarada	1	1,5	1	2	4	2,5	5	6	10	8

a) Oceniti linearni regresioni model i grafički ga predstaviti;
b) Kolika je očekivana promena zarade ukoliko se godine školovanja povećaju za jednu godinu? Testirati statističku značajnost koeficijenta pravca regresije;
c) U kom procentu je visina mesečnih zarada objašnjena brojem godina školovanja? Da li je procenat objašnjenosti visine mesečnih zarada brojem godina školovanja statistički značajan?

2. Na osnovu podataka o dužini proizvodnog iskustva radnika i njihovoj produktivnosti:

Godine proizvodnog iskustva	2	4	6	8	10
Produktivnost za 1h rada	6	8	14	12	16

a) Oceniti linearni regresioni model zavisnosti produktivnosti radnika za 1h od godina proizvodnog iskustva i grafički ga predstaviti.
b) Testirati statističku značajnost regresionog modela.
c) Izračunati učešće varijabiliteta objašnjenog linernom regresijom u ukupnom varijabilitetu zavisno promenljive.
d) Izračunati varijansu i standardnu grešku regresije.

3. Na osnovu podataka o padavinama (cm) i prinosu krušaka po stablu (kg/stablu) dobijeni su sledeći zbirovi: n=7, $\sum X$=112, $\sum Y$= 448, $\sum XY$=7254, $\sum X^2$=2084. Oceniti linerni regresioni model i grafički ga predstaviti.

4. Na osnovu slučajnog uzorka od 12 podataka dobijeni su sledeći rezultati: $\sum(X-\bar{X})^2 = 33$, $\sum(X-\bar{X})(Y-\bar{Y}) = -990, \sum(Y-\bar{Y})^2 = 29922$. Izračunati koeficijent proste linearne korelacije. Da li je procenat objašnjenosti zavisno promenljive Y nezavisno promenljivom X statistički značajan? Izračunati standardnu grešku regresije.

INDEKSNI BROJEVI I ANALIZA VREMENSKIH SERIJA

Indeksni brojevi

Indeksi su relativni brojevi koji se dobijaju stavljanjem u odnos dve ili više pojava. Opšti oblik indeksih brojeva dat je sledećim izrazom:

$$I = \frac{tekuća\ vrednost}{bazna\ vrednost} \cdot 100$$

Prema tome da li indeksi izražavaju relativnu promenu jedne ili više pojava razlikuju se individualni i grupni indeksi.

Individualni indeksi

Individualni indeksi izražavaju relativne promene jedne pojave u odnosu na jednu stalnu vrednost (bazni indeksi) ili uvek u odnosu na vrednost iz prethodnog perioda (lančani indeksi).

Bazni indeksi: $I = \frac{Y_i}{Y_B} \cdot 100$

Lančani indeksi: $L = \frac{Y_i}{Y_{i-1}} \cdot 100$

Primer 1. Na osnovu podataka o proizvodnji goveđeg mesa (Y) u periodu od 2012-2017 godine izračunati:

- bazne indekse (I), ako je baza 2012. godina;
- lančane indekse (L);
- pretvoriti bazne indekse 2012=100 u bazne da je 2014=100;
- lančane indekse pretvoriti u bazne, gde je 2012=100;
- lančane indekse pretvoriti u bazne gde je 2014=100.

Godine	**Y**					
2012	362					
2013	357					
2014	371					
2015	352					
2016	339					
2017	359					

Primer 2. Na osnovu podatak o broju zaposlenih u Vojvodini u periodu od 2009-2017. godine izračuanti:

- seriju baznih indeksa, ako je 2013. bazna godina;
- seriju baznih indeksa (2013=100) pretvoriti u seriju baznih indeksa (2015=100);
- seriju baznih indeksa (2015=100) pretvoriti u seriju lančanih indeksa.

Godine	**Broj zaposlenih**			
2009	568			
2010	560			
2011	553			
2012	549			
2013	541			
2014	537			
2015	544			
2016	528			
2017	523			

Grupni indeksi

Grupni indeksi pokazuju relativne promene više pojava koje čine neku celinu. U ekonomskim analizama prisutna su najčešće tri tipa grupnih indeksa:
-indeks fizičkog obima (količina);
-indeks cena i
-indeks vrednosti proizvodnje.

Grupni indeksi cena i fizičkog obima proizvodnje mogu se izračunati primenom četiri tipa formula:
1. Lasperejsova;
2. Pašeova;
3. Kombinovana;
4. Fišerova (idealna).

Primer 1. Na osnovu cena i količina četiri posmatrana proizvoda u 2018 i 2019. godini izračunati grupne indekse cena i količina (fizičkog obima) i indeks vrednosti proizvodnje.

Proizvodi	**2018**		**2019**					
	Cene	Količine	Cene	Količine				
A	10	15	12	19				
B	15	17	17	24				
C	20	23	22	33				
D	25	21	27	30				

1. Lasperejsova:

$$I_{p(L)} = \frac{\sum p_1 q_0}{\sum p_0 q_0} \cdot 100 \qquad I_{q(L)} = \frac{\sum q_1 p_0}{\sum q_0 p_0} \cdot 100$$

2. Pašeova:

$$I_{p(P)} = \frac{\sum p_1 q_1}{\sum p_0 q_1} \cdot 100 \qquad I_{q(P)} = \frac{\sum q_1 p_1}{\sum q_0 p_1} \cdot 100$$

3. Kombinovana :

$$I_{p(K)} = \frac{\sum p_1 q_0 + \sum p_1 q_1}{\sum p_0 q_0 + \sum p_0 q_1} \cdot 100 \qquad I_{q(K)} = \frac{\sum q_1 p_0 + \sum q_1 p_1}{\sum q_0 p_0 + \sum q_0 p_1} \cdot 100$$

4. Fišerova (idealna):

$$I_{p(F)} = \sqrt{\frac{\sum p_1 q_0 \cdot \sum p_1 q_1}{\sum p_0 q_0 \cdot \sum p_0 q_1}} \cdot 100 \qquad I_{p(F)} = \sqrt{\frac{\sum q_1 p_0 \cdot \sum q_1 p_1}{\sum q_0 p_0 \cdot \sum q_0 p_1}} \cdot 100$$

Indeks vrednosti proizvodnje: $I_w = \frac{\sum p_1 q_1}{\sum p_0 q_0} \cdot 100$

Zadaci za vežbu:

1. Dati su podaci o ukupnoj proizvodnji goveđeg mesa u jednom preduzeću (t) u periodu 2009-2015. godina:

Godine	2009	2010	2011	2012	2013	2014	2015
Proizvodnja (t)	146	161	135	117	145	130	129

a) Izračunati bazne indekse (2009=100).
b) Seriju baznih indeksa pretvoriti u lančane indekse.

2. Broj stabala šljive sposobnih za rod u periodu od 2012-2016. godine bio je sledeći:

Godine	2012	2013	2014	2015	2016
Broj stabala	73392	74078	74636	73717	74551

a) Izračunati bazne i lančane indekse.
b) Izračunati bazne indekse ukoliko je baza prosečan broj stabala sposobnih za rod u posmatranom periodu.

3. Dati su podaci o cenama i količinama četiri proizvoda proizvedenih u jednoj organizaciji:

	Cene (000rsd)			**Količine (t)**		
Proizvodi	2017	2018	2019	2017	2018	2019
A	3,35	3,50	3,99	13,1	12,2	14,5
B	0,92	1,07	1,15	6,7	7,6	5,9
C	1,80	1,80	1,94	10,3	10,4	10,6
D	1,13	1,12	1,19	8,5	9,1	9,4

Izračunati grupne indekse cena i fizičkog obima proizvodnje za 2018. i 2019. godinu ukoliko su ponderi iz baznog perioda (2017).

Linearni trend

Uticaj trenda se ispoljava na duži vremenski rok, a ispitivanje trenda treba da ukaže na opštu tendenciju pojave, pri čemu se zanemaruju kratkoročna kolebanja.

Najjednostavnija metoda sagledavanja trenda je grafička metoda. Bolji i precizniji metodi za sagledavanje trenda zasnovani su na izračunavanjima, jedan od tih metoda je metod pokretnih proseka. Pokretni proseci izračunavaju se za tri, četiri, pet ili više godina, na taj način što se svaki član vremenske serije zamenjuje sa odgovarajućim pokretnim prosekom, odnosno aritmetičkom sredinom zbira vrednosti tog podatka i jednog, dva, tri ili više njemu prethodnih i narednih podataka.

Matematička funkcija trenda i stpena ima sledeći oblik: $\hat{Y}_i = a + bX_i$.

Primer 1. Dati su podaci o kretanju proizvodnje goveđeg mesa (000t):

Godine	2010	2011	2012	2013	2014	2015	2016	2017
Proizvodnja	344	341	362	357	371	352	339	359

Na osnovu podataka o kretanju proizvodnje goveđeg mesa (000 t):

a) Naći vrednosti trogodišnjih pokretnih proseka;
b) Izračunati parametre u jednačini linernog trenda;
c) Izračunati ocenjene vrednosti trenda za svaku godinu;
d) Izračunati očekivani obim proizvodnje za 2018. godinu;
e) Izračunati standardnu grešku linernog trenda trenda i grafički predstaviti originalne podatke i liniju trenda.

Radna tabela:

Godine									

Primer 2. Broj zaposlenih u jednom preduzeću u periodu 2011-2019. godine je bio:

Godine	2011	2012	2013	2014	2015	2016	2017	2018	2019
Broj zaposlenih	568	560	553	549	541	537	544	528	523

a) Izračunati pokretne proseke i grafički ih predstaviti;
b) Oceniti model linearnog trenda i grafički ga predstaviti;
c) Koliki broj zaposlenih može da se očekuje 2022. godine?

Radna tabela:

Godine									

Zadaci za vežbu:

1. U tabeli su dati podaci o proizvodnji mleka (mil. litara) u Vojvodini u periodu 2011-2019. godine:

Godine	2011	2012	2013	2014	2015	2016	2017	2018	2019
Proizvodnja	294	292	306	314	348	397	379	380	346

a) Izračunati trogodišnje pokretne proseke i grafički ih predstaviti;
b) Oceniti jednačinu linernog trenda i grafički je predstaviti;
c) Kolika proizvodnja mleka može da se očekuje 2025 godine u Vojvodini?

2. Kretanje broja studenata na jednom fakultetu u periodu 2013-2019. godine, dato je u tabeli:

Godine	2013	2014	2015	2016	2017	2018	2019
Broj studenata	1949	2061	2249	2258	2207	2280	2261

a) Izračunati trogodišnje pokretne proseke i grafički ih predstaviti;
b) Oceniti jednačinu linernog trenda i grafički je predstaviti;
c) Proceniti broj studenata za 2020. godinu.

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
5. Mutavdžić Beba, Novaković T., Tekić Dragana (2020), Praktikum iz statistike (za smer Veterinarska medicina), Poljoprivredni fakultet, Univerzitet u Novom Sadu, Republika Srbija