POSLOVNA STATISTIKA

Pokazatelji empirijske raspodele:

Mere sredine. Mere odstupanja. Mere simetrije. Mere spljoštenosti.

Mere sredine

Srednje vrednosti (mere centralne tendencije)

- Srednje vrednosti prikazuju sredinu statističke serije.
- Nalaze se između najmanje i najveće vrednosti obeležja
- U zavisnosti od načina definisanja dele se na
  - Izračunate
  - Pozicione

Izračuanate srednje vrednosti

- Izračunate srednje vrednosti se određuju na osnovu svih podataka jedne statističke serije po određenom pravilu.
- U ovu grupu spadaju:
  - Aritmetička sredina
  - Geometrijska sredina
  - Harmonijska sredina
  - Ponderisana sredina

Pozicione srednje vrednosti

- Pozicione srednje vrednosti se određuju na osnovu pozicije podataka u seriji po određenom pravilu.
- Osnovne pozicione vrednosti su
  - Medijana
  - Modus

Aritmetička sredina

- Aritmetička sredina je količnik zbira svih vrednosti obeležja i njihovog broja.
- Obeležava se sa $\bar{x}$
- U nastavku dajemo formule za negrupisane i grupisane podatke

Aritmetička sredina za negrupisane podatke

- Aritmetička sredina za negrupisane podatke data je sa

$$\bar{x} = \frac{x_1 + x_2 + \ldots + x_n}{n} = \frac{\sum_{i=1}^{n} x_i}{n}$$

- Gde je
  - x_i vrednost podataka iz serije
  - n ukupan broj podataka

Aritmetička sredina – primer za negrupisane podatke

- Izračunati prosečan i ukupan broj dana koje je jedan radnik provede na bolovanju, na osnovu podataka o broju dana bolovanja u toku jedne godine. Broj dana: 7, 23, 4, 8, 2, 12, 6,13, 9, 4
- Rešenje:

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{x_1 + x_2 + \ldots + x_n}{n}$$

$$= \frac{7+23+4+8+2+12+6+13+9+4}{10}$$

$$= \frac{88}{10}$$

$$= 8.8$$

- Ukupan broj dana provedenih na bolovanju je 88, a prosečno 8.8 dana na bolovanju.

Aritmetička sredina za grupisane podatke

- Kod grupisanih podataka, za koje je kreirana distribucija frekvencija, prethodna formula za aritmetičku sredinu dobija oblik:

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

- Pri čemu je
  - k ukupan broj klasa,
  - x_i vrednost obeležja i-te klase (kod intervalnih klasa to je sredina intervala),
  - f_i frekvencija i-te klase.

Aritmetička sredina – primer za grupisane podatke (1)

- Na osnovu podataka o broju prodatih mobilnih telefona u jednom danu, odrediti prosečan broj prodatih telefona u danu.

Broj prodatih mobilnih telefona	Frekvencija
8	2
9	4
10	6
11	7
12	5
13	4
14	1
15	1

Aritmetička sredina – primer za grupisane podatke (2)

x	f	fx
8	2	16
9	4	36
10	6	60
11	7	77
12	5	60
13	4	52
14	1	14
15	1	15
Ukupno:	**30**	**330**

- Koristeći formulu za grupisane podatke dobijamo

$$\bar{x} = \frac{\sum_{i=1}^{k} f_i x_i}{\sum_{i=1}^{k} f_i} = \frac{330}{33} = 11$$

- U proseku se dnevno proda 11 telefona.

Aritmetička sredina – primer za grupisane intervalne podatke (1)

- Izračunati prosečno vreme za izradu proizvoda na osnovu podataka iz tabele

Vreme izrade u satima x	Broj radnika f
15.2-15.6	2
15.6-16.0	9
16.0-16.4	12
16.4-16.8	8
16.8-17.2	5
17.2-17.6	4

Aritmetička sredina – primer za grupisane intervalne podatke (2)

Vreme izrade u satima x	Sredina intervala x_{sr}	Broj radnika f	fx_{sr}
15.2-15.6	15.4	2	30.8
15.6-16.0	15.8	9	142.2
16.0-16.4	16.2	12	194.4
16.4-16.8	16.6	8	132.8
16.8-17.2	17	5	85.0
17.2-17.6	17.4	4	69.6
	Ukupno:	40	654.8

$$\bar{x} = \frac{\sum_{i=1}^{k} f_i x_i}{\sum_{i=1}^{k} f_i} = \frac{654.8}{40} = 16.37$$

Aritmetička sredina aritmetičkih sredina

- Kada imamo niz aritmetičkih sredina i potrebno je izračunati njihovu aritmetičku sredinu, **ne može** se koristiti formula koja predstavlja količnik zbira svih sredina i njihovog broja. Mora se uzeti u obzir i obim podataka na osnovu kojih je svaka aritmetička sredina izračunata.
- Aritmetička sredina aritmetičkih sredina je

$$\bar{\bar{x}} = \frac{\sum_{i=1}^{m} n_i \bar{x}_i}{\sum_{i=1}^{m} n_i}$$

- Gde je
  - m ukupan broj aritmetičkih sredina
  - $\bar{x}_i$ aritmetička sredina i-te serije
  - n_i broj podataka i-te serije

Aritmetička sredina aritmetičkih sredina - primer

- Na osnovu podataka o aritmetičkim sredinama plata u tri odeljenja, odrediti platu u celoj firmi.

Odeljenje	Br. radnika	Prosečna plata
Menadžment	10	585250
IT	50	157950
HR	5	81900

$$\overline{\overline{x}} = \frac{\sum_{i=1}^{m} n_i \overline{x}_i}{\sum_{i=1}^{m} n_i} = \frac{10 \cdot 585250 + 50 \cdot 157950 + 5 \cdot 81900}{10 + 50 + 5} = \frac{14159500}{65} = 217838.5$$

- Prosečna plata radnika u firmi je 217838.5

OSOBINE ARITMETIČKE SREDINE

- U statističkoj analizi aritmetička sredina se najčešće koristi zbog:
  - Jednostavnog izračunavanja i interpretacije (jedan broj reprezentuje ceo skup podataka).
  - Svi kvantitativni podaci imaju aritmetičku sredinu i ona je jedinstveno određena.
  - Izračunava se iz svih vrednosti obeležja i izravnava njihove apsolutne razlike.
  - Nalazi se između minimalne i maksimalne vrednosti obeležja.
  - Zbir odstupanja svih vrednosti obeležja od aritmetičke sredine je nula.
  - Zbir kvadrata odstupanja vrednosti obeležja od aritmetičke sredine je minimalan.
- Nedostaci:
  - Na njenu vrednost utiču ekstremne vrednosti
  - Svaki podatak ulazi u obračun, što je problem za velike serije

Geometrijska sredina (1)

- Geometrijska sredina je srednja vrednost koja izjednačava proporcionalne razlike vrednosti obeležja.
- Definisana je samo za pozitivne vrednosti nekog obeležja.
- Označava se sa G i izračunava na sledeći način

$$G = \sqrt[n]{x_1 \cdot x_2 \cdot \ldots \cdot x_n}$$

- Gde je
  - n broj podataka u seriji
  - x_i vrednosti obeležja iz serije

Geometrijska sredina (2)

- Koristi se kada se proučava prosečna promena neke pojave kroz vreme, npr. ako posmatramo procentualno kretanje cene za neku robu ili uslugu, tu aritmetička sredina ne daje zadovoljavajući rezultat.
- Npr. ako je cena narasla za 30% u prvoj godini, zatim opala za 20% u drugoj, u trćoj ponovo narasla za 10%, prosečno kretanje cene biće 4.59% godišnje
- $$\sqrt[3]{\left(1+\frac{30}{100}\right)\cdot\left(1-\frac{10}{100}\right)\cdot\left(1+\frac{20}{100}\right)}-1=0.0459$$

Harmonijska sredina (1)

- Harmonijska sredina je recipročna aritmetička sredina recipročnih vrednosti obeležja.
- Označava se sa H i određuje na sledeći način

$$H = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \ldots + \frac{1}{x_n}}$$

- Gde je
  - n broj podataka u seriji
  - x_i vrednosti obeležja iz serije

Harmonijska sredina (2)

- Koristi se izračunavanje proseka pojave čije su vrednosti obeležja obrnuto proporcionalne veličini pojave.
- Na primer ako želimo da oderimo kupovnu moć dinara na osnovu cena jednog proizvoda u toku 5 godina: 40,80,90,110 i 150. S obzirom da je kupovna moć dinara obrnuto proporcionalna ceni robe tada je prosečna kupovna moć data sa:

$$H = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \frac{1}{x_3} + \frac{1}{x_4} + \frac{1}{x_5}} = \frac{5}{\frac{1}{40} + \frac{1}{80} + \frac{1}{90} + \frac{1}{110} + \frac{1}{150}} = 77.68$$

PONDERISANA SREDINA

- Određuje se po formuli

$$\overline{x_w} = \frac{w_1 x_1 + w_2 x_2 + \ldots + w_n x_n}{w_1 + w_2 + \ldots + w_n} = \frac{\sum_{i=1}^{n} w_i x_i}{\sum_{i=1}^{n} w_i}$$

- Gde je
  - n broj podataka u seriji
  - x_i vrednosti obeležja iz serije
  - w_i težinski faktori (ponderi) odgovarajućih obeležja iz serije

PONDERISANA SREDINA

- Koristi se u situacijama kada nekim vredostima treba da se dodeli veža težina ili uticaj.
- Tako možemo posmatrati situaciju da student polaže ispit iz 3 dela, gde svaki nosi 100 bodova. Na prvom je osvojio 90, na drugom 85, a na trećem 75 bodova.
- S obzirom da je prvi test najlakši nastavnik, on će nositi vrednost 20% ocene, drugi test 35% a treću 45%

$$\overline{x_w} = \frac{20 \cdot 90 + 35 \cdot 85 + 45 \cdot 75}{20 + 35 + +45} = 81.5$$

MEDIJANA

- Medijana je vrednost obeležja koju ima srednji član sređene serije.
- Označava se sa M_e
- Da bi mogla dase odredi prvo podaci u seriji moraju urediti po rastućem redosledu, tj. od najmanjeg do najvećeg.
- Zatim se određuje član koji se nalazi u sredini te serije.
- Način određivanja zavisi od toga da li serija ima paran ili neparan broj podataka.

Medijana za negrupisane podatke

- Posmatramo statističku seriju $x_1, x_2, \ldots, x_n$
- Ako je n neparan broj tada je medijana određena sa

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

- Ako je n paran broj

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

Medijana za negrupisane podatke- primer 1

- Za statističke podatke o broju članova u 7 porodica odrediti medijanu: 2, 3, 2, 5, 4, 4, 3
- 1. korak je sortiranje podataka u rastućem redosledu

x_1	x_2	x_3	x_4	x_5	x_6	x_7
2	2	3	3	4	4	5

- Kako je broj podataka neparan *n*=7, primenjujemo pravilo

$$M_e = x_{\frac{n+1}{2}} = x_{\frac{7+1}{2}} = x_4 = 3$$

Medijana za negrupisane podatke- primer 2

- Za statističke podatke o broju članova u 8 porodica odrediti medijanu:2, 3, 2, 5, 4, 5, 4, 3
- Prvo sortiramo podatke u rastućem redosledu

x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
2	2	3	**3**	**4**	4	5	5

- Sada je *n=8* paran broj

$$M_e = \frac{x_{\frac{n}{2}} + x_{\frac{n}{2}+1}}{2} = \frac{x_{\frac{8}{2}} + x_{\frac{8}{2}+1}}{2} = \frac{x_4 + x_5}{2} = \frac{3+4}{2} = 3.5$$

Medijana za grupisane podatke (1)

- Kada su podaci grupisani u distribuciju frekvencije, klase treba urediti u rastućem redosledu i na osnovu kumulativnih frekvencija odrediti **medijalnu klasu** tj. klasu koja sadrži medijanu.
- Za diskretne vrednosti klasa, može se odrediti tačna vrednost medijane koristeći prethodno definisana pravila.

Medijana za grupisane podatke (2)

- Kada je serija intervalna, ne može se odrediti tačna vrednost medijane. Tu postoje dva pristupa
  - Medijana se računa kao sredina intervala ili
  - se računa po formuli

$$M_e = L_m + \frac{\frac{n}{2} - K_{m-1}}{f_m}\Delta$$

  - Gde je
    - m medijalni interval
    - L_m donja granica medijalnog intervala
    - n broj podataka u seriji
    - K_m-1 kumulativna frekvencija klase koja prethodi medijalnoj
    - f_m frekvencija medijalne klase
    - Δ dužina klasnog intervala

Medijana za grupisane podatke primer 1

Broj prodatih klima uređaja (x_i)	Broj prodavnica (f_i)	Kumulativne frekvencije
16	1	1
18	3	4
20	4	8
22	7	15
24	2	17
Ukupno	**17**	

- Na osnovu rasporeda prodavnica i prema broju prodatih klima uređaja, datog u tabeli odrediti medijanu.
- Datim podacima dodeljujemo kolonu kumulativnih frekvencija.
- Podaci su diskretnog tipa, pa je medijana definisana sa

$$M_e = x_{\frac{n+1}{2}} = x_{\frac{17+1}{2}} = x_9 = 7$$

- Objašnjenje: 9. član se nalazi na prvoj poziciji gde je kumulativna frekvencija veća ili jednaka sa 9.

Medijana za grupisane podatke primer 2

- Na osnovu rasporeda 39 radnika prema izvršenju norme datom u tabeli, odrediti medijanu.

xi	fi	Ki
87 – 92	4	4
92 – 97	5	9
97 – 102	15	24
102 – 107	10	34
107 – 112	5	39
Ukupno	**39**	

- Podacima odmah dodajemo kolonu sa kumulativnim frekvencijama.

- Podataka ima 39, znači da se medijana nalazi na 20. članu serije, ali njega u intervalnoj seriji ne možemo tačno da nađemo.
  - Računamo je kao sredinu medijalnog intervala

$$M_e = \frac{97+102}{2} = 99.5$$

  - Ili koristeći formulu

$$M_e = L_m + \frac{\frac{n}{2} - K_{m-1}}{f_m}\Delta = 97 + \frac{\frac{39}{2} - 9}{15} \cdot 5 = 100.5$$

Modus (1)

- Modus je vrednost obeležja koja se najčešće javlja u seriji.
- Označava se sa M_o
- Predstavlja vrednost klase sa najvećom frekvencijom.
- Ukoliko se svi podaci u seriji javljaju jedanput, modus ne postoji.
- Serije koje imaju jedan modus nazivaju se **unimodalne**,seije sa dva modusa **bimodalne**,sa više modusa **polimodalne**.

Modus (2)

- Za diskretna obeležja modus se lako određuje na osnovu maksimalne frekvencije
- Kod intervalnih klasa, ne može se tačno odrediti pa se kao i u slučaju medijane,
  - može uzeti sredina modalnog intervala ili
  - se može izračunati koristeći formulu

$$M_o = L_m + \frac{f_m - f_{m-1}}{\left(f_m - f_{m-1}\right) + \left(f_m - f_{m+1}\right)} \Delta$$

  - Gde je
    - m modalni interval
    - L_m donja granica modalnog intervala intervala
    - f_m frekvencija modalne klase
    - f_{m-1} frekvencija klase koja prethodi modalnoj
    - f_{m+1} frekvencija klase koja sledi modalnu klasu
    - Δ dužina klasnog intervala

MODUS ZA GRUPISANE PODATKE PRIMER 1

Broj prodatih klima uređaja (x_i)	Broj prodavnica (f_i)
16	1
18	3
20	4
22	7
24	2

- Na osnovu rasporeda prodavnica i prema broju prodatih klima uređaja, datog u tabeli odrediti modus.
- Podatak sa najvećom frekvencijom je 22

$$M_o = 22$$

Modus za grupisane podatke primer 2

- Na osnovu rasporeda 39 radnika prema izvršenju norme datom u tabeli, odrediti modus.

xi	fi
87 – 92	4
92 – 97	5
97 – 102	15
102 – 107	10
107 – 112	5

- Modalni interval je onaj sa frekvencijom 15

- Modus ne možemo tačno odrediti, ali imamo dva pristupa,
  - Određujemo ga kao sredinu modalnog intervala

$$M_o = \frac{97+102}{2} = 99.5$$

  - Ili koristeći formulu

$$M_o = L_m + \frac{f_m - f_{m-1}}{\left(f_m - f_{m-1}\right) + \left(f_m - f_{m+1}\right)}\Delta$$

$$= 97 + \frac{15-5}{(15-5)+(15-10)} \cdot 5$$

$$= 100.3333$$

Mere odstupanja

Mere odstupanja

- Mere odstupanja se nazivaju još i mere varijacije, i mere disperzije.
- Pokazuju kakvo je variranje podataka u seriji. Za podatke sa većim varijabilitetom sredina je manje reprezentativan pokazatelj nego za podatke sa manjim varijabilitetom.
- Na primer neka je u jednoj firmi prosečna plata u finansijama i u IT odeljenju 80000. Međutim plate u finansijama su u rasponu od 70000 do 90000, a plate u IT su u rasponu od 40000 do 120000.
- Iako su prosečne plate jednake, varijabilitet u IT je veći tako da prosečna plata na pravi način ne oslikava stvarnu vrednost plate. Iz tog razloga nam trebaju i mere odstipanja.

Mere odstupanja-klasifikacija

- U zavisnosti od izračunavanja mogu biti:
  - Apsolutne (iskazane su u jedinicama mere obeležja)
    - Interval varijacije
    - Srednje apsolutno odstupanje
    - Varijansa
    - Standardna devijacija
  - Relativne (neimenovani pokazatelji varijabiliteta)
    - Koeficijent varijacije
    - Normalizovano standardno odstupanje

INTERVAL VARIJACIJE

- Interval (rang) varijacije je razlika između maksimalne i minimalne vrednosti obeležja

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

- Gde je
  - x_{max} maksimalna vrednost obeležja
  - x_{min} minimalna vrednost obelažja
- Kada se radi sa neprekidnim obeležjima predstavljenih intervalnim klasama

$$I_v = u_{\max} - l_{\min}$$

- Gde je
  - u_{max} gornja granica intervala sa maksimalnim vrednostima
  - l_{min} donja granica intervala sa minimalnim vrednostima

Srednje apsolutno odstupanje (1)

- Srednje apsolutno odstupanje(AD) je prosečno odstupanje obeležja od srednje vrednosti
- Za negrupisane podatke izračunava se po formuli

$$AD = \frac{\sum_{i=1}^{n} \left| x_i - \bar{x} \right|}{n}$$

- Gde je
  - x_i vrednost obeležja,
  - $\bar{x}$ aritmetička sredina
  - n broj podataka

Srednje apsolutno odstupanje (2)

- Za grupisane podatke sa distribucijama frekvencija izračunava se po formuli

$$AD = \frac{\sum_{i=1}^{k} f_i \left| x_i - \bar{x} \right|}{\sum_{i=1}^{k} f_i}$$

- Gde je
  - x_i vrednost obeležja i-te klase,
  - $\bar{x}$ aritmetička sredina,
  - f_i frekvencija i-te klase,
  - k broj klasa.

Varijansa (1)

- Varijansa ili srednje kvadratno odstupanje je mera varijabiliteta koja se izračunava kao prosek kvadrata odstupanja sredine i vrednosti svakog podatka u seriji.
- Ova mera ima najveću primenu u statistici
- Označava se sa s^2.

Varijansa (2)

- Varijansa za negrupisane podatke se računa po formuli

$$s^2 = \frac{\sum_{i=1}^{n}\left(x_i - \overline{x}\right)^2}{n} = \frac{\sum_{i=1}^{n} x_i^2}{n} - \overline{x}^2$$

- Gde je
  - x_i vrednost obeležja,
  - $\bar{x}$ aritmetička sredina
  - n broj podataka

Varijansa (3)

- Varijansa za grupisane podatke se računa po formuli

$$s^2 = \frac{\sum_{i=1}^{k} f_i \left(x_i - \overline{x}\right)^2}{\sum_{i=1}^{k} f_i} = \frac{\sum_{i=1}^{k} f_i x_i^2}{\sum_{i=1}^{k} f_i} - \overline{x}^2$$

- de je
  - x_i vrednost obeležja i-te klase,
  - $\bar{x}$ aritmetička sredina,
  - f_i frekvencija i-te klase,
  - k broj klasa.

Standardna devijacija

- Varijansa nije pogodna za interpretaciju jer je izražena u kvadratima jedinice u kojima su izraženi podaci.
- Zbog toga se uvodi standardna devijacija koja je kvadratni koren varijanse.

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

- Za negrupisane podatke je $s = \sqrt{\dfrac{\sum_{i=1}^{n}\left(x_i - \overline{x}\right)^2}{n}} = \sqrt{\dfrac{\sum_{i=1}^{n} x_i^2}{n} - \overline{x}^2}$

- Za grupisane podatke $s = \sqrt{\dfrac{\sum_{i=1}^{k} f_i\left(x_i - \overline{x}\right)^2}{\sum_{i=1}^{k} f_i}} = \sqrt{\dfrac{\sum_{i=1}^{k} f_i x_i^2}{\sum_{i=1}^{k} f_i} - \overline{x}^2}$

PRIMER 1

Odrediti varijabilitet za broj dana koje radnik provodi na bolovanju u odnosu na prosečan broj dana bolovanja. Kao meru varijabiliteta koristiti

a) Interval varijacije
b) Srednje apsolutno odstupanje
c) Varijansu
d) Standardnu devijaciju

Broj dana na bolovanju: 7 23 4 8 2 12 6 13 9 4

Rešenje:

x_i	$x_i-\bar{x}$	$\lvert x_i-\bar{x}\rvert$	$(x_i-\bar{x})^2$
7	-1,8	1,8	3,24
23	14,2	14,2	201,64
4	-4,8	4,8	23,04
8	-0,8	0,8	0,64
2	-6,8	6,8	46,24
12	3,2	3,2	10,24
6	-2,8	2,8	7,84
13	4,2	4,2	17,64
9	0,2	0,2	0,04
4	-4,8	4,8	23,04
∑=88		43,6	333,6

a) $I_v = x_{\max} - x_{\min} = 23 - 2 = 21$

b) $\bar{x} = \dfrac{\sum_{i=1}^{n} x_i}{n} = \dfrac{88}{10} = 8,8 \Rightarrow AD = \dfrac{\sum_{i=1}^{n} \lvert x_i - \bar{x}\rvert}{n} = \dfrac{43,6}{10} = 4,36$

c) $s^2 = \dfrac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n} = \dfrac{333,6}{10} = 33,6$

d) $s = \sqrt{s^2} = \sqrt{33,36} = 5,775812 \approx 5,776$

Primer 2

Odrediti varabilitete prodaje TV aparata na osnovu podataka iz tabele. Za meru varijabiliteta koristiti

a) Interval varijacije
b) Srednje apsolutno odstupanje
c)Varijansu
d) Standardnu devijaciju

x_i	f_i
8	2
9	4
10	6
11	7
12	5
13	4
14	1
15	1

PRIMER 2

Rešenje:

a) $I_v = x_{\max} - x_{\min} = 15 - 8 = 7$ Interval varijacije je 7.

x_i	f_i	$f_i \cdot x_i$	$x_i - \overline{x}$	$\lvert x_i - \overline{x} \rvert$	$f_i \lvert x_i - \overline{x} \rvert$	$(x_i - \overline{x})^2$	$f_i (x_i - \overline{x})^2$	x_i^2	$f_i x_i^2$
8	2	16	-3	3	6	9	18	64	128
9	4	36	-2	2	8	4	16	81	324
10	6	60	-1	1	6	1	6	100	600
11	7	77	0	0	0	0	0	121	847
12	5	60	1	1	5	1	5	144	720
13	4	52	2	2	8	4	16	169	676
14	1	14	3	3	3	9	9	196	196
15	1	15	4	4	4	16	16	225	225
Σ	30	330			40		86		3716

$$\overline{x} = \frac{\sum_{i=1}^{n} f_i x_i}{\sum_{i=1}^{n} f_i} = \frac{330}{30} = 11 \Rightarrow AD = \frac{\sum_{i=1}^{n} f_i \lvert x_i - \overline{x} \rvert}{\sum_{i=1}^{n} f_i} = \frac{40}{30} = 1,333$$

Prosečno dnevno variranje prodaje TV aparata,koristeći srednje apsolutno odstupanje je 1,33 komada

1. način

$$s^2 = \frac{\sum_{i=1}^{n} f_i (x_i - \overline{x})^2}{\sum_{i=1}^{n} f_i} = \frac{86}{30} = 2.86666 \approx 2.867$$

2.način

$$s^2 = \frac{\sum_{i=1}^{n} f_i x_i^2}{\sum_{i=1}^{n} f_i} - \overline{x}^2 = \underbrace{\frac{3716}{30}}_{123,867} - \underbrace{11^2}_{121} \approx 2.867$$

Varijansa je 2,867

$s = \sqrt{s^2} = 1.693123347 \approx 1,693$

Variranje prodaje TV aparata, koristeći standardnu devijaciju je 1,639 komada

PRIMER 3

Odrediti varijabilitet izrade proizvoda za podatke iz tabele, kao meru varijabiliteta koristi
a) Interval varijacije
b) Srednje apsolutno odstupanje
c)Varijansu
d) Standardnu devijaciju

Vreme izrade (x_i)	Broj radnika (f_i)
15,2-15,5	3
15,5-15,8	9
15,8-16,1	11
16,1-16,4	8
16,4-16,7	5
16,7-17,0	4

Primer 3

Rešenje:

$I_v = x_{\max} - x_{\min} = 17 - 15,2 = 1,8$

x_i	f_i	x_{si}	$f_i x_{si}$	$x_{si} - \overline{x}$	$\lvert x_{si} - \overline{x}\rvert$	$f_i \lvert x_{si} - \overline{x}\rvert$	$(x_{si} - \overline{x})^2$	$f_i (x_{si} - \overline{x})^2$	x_{si}^2	$f_i x_{si}^2$
15,2-15,5	3	15,35	46,05	-0,7125	0,7125	2,1375	0,5077	1,5231	235,6225	706,8675
15,5-15,8	9	15,65	140,85	-0,4125	0,4125	3,7125	0,1702	1,5318	244,9225	2204,303
15,8-16,1	11	15,95	175,45	-0,1125	0,1125	1,2375	0,0127	0,1397	254,4025	2798,428
16,1-16,4	8	16,25	130,00	0,1875	0,1875	1,5000	0,0352	0,2816	264,0625	2112,5
16,4-16,7	5	16,55	82,75	0,4875	0,4875	2,4375	0,2377	1,1885	273,9025	1369,513
16,7-17	4	16,85	67,40	0,7875	0,7875	3,1500	0,6202	2,4808	283,9225	1135,69
$\sum$	40,00		642,50			14,1750		7,1455		10327,30

$$\overline{x} = \frac{\sum_{i=1}^{n} f_i x_{si}}{\sum_{i=1}^{n} f_i} = \frac{642,50}{40} = 16,0625 \Rightarrow AD = \frac{\sum_{i=1}^{n} f_i \lvert x_{si} - \overline{x}\rvert}{\sum_{i=1}^{n} f_i} = \frac{14,1750}{40} = 0,354375$$

1. način

$$s^2 = \frac{\sum_{i=1}^{n} f_i (x_{si} - \overline{x})^2}{\sum_{i=1}^{n} f_i} = \frac{7,1455}{40} = 0,178638$$

2.način

$$s^2 = \frac{\sum_{i=1}^{n} f_i x_{si}^2}{\sum_{i=1}^{n} f_i} - \overline{x}^2 = \underbrace{\frac{10327,30}{40}}_{258,1825} - \underbrace{16,0625^2}_{258,0039} = 0,1786$$

Varijansa je 0,1786

$\text{s} = \sqrt{\text{s}^2} = 0,422610932 \approx 0,423$

Variranje u odnosu na prosečno vreme izrade je 0,1786

Koeficijent varijacije

- **Koeficijent varijacije**- koristi se za poređenje variajabiliteta u različitim statističkim serijama

$$C_v = \frac{s}{\overline{x}} \cdot 100\%$$

Primer

U jednom preduzeću prosečna plata je 1200 din sa standardnom devijacijom 200, a prosečan radni staž zaposlenih je 17 godina sa varijansom 25. Uporediti varijabilitet plata i radnog staža u ovom preduzeću.

Rešenje: posmatramo dve pojave plata-x i radni staž y

$\overline{x} = 1200;\ s_x = 200$

$\overline{y} = 17;\ s_y^2 = 25 \Rightarrow s_y = 5$

$C_v(x) = \frac{200}{1200} \cdot 100\% = 16,67\%$

$C_v(y) = \frac{5}{17} \cdot 100\% = 29,41\%$

Radni staž ima veći varijabilitet od plata.

NORMALIZOVANO STANDARDNO ODSTUPANJE

- **Normalizovano odstupanje**-koristi se za poređenje odstupanja od aritmetičke sredine

$$z = \frac{x_i - \bar{x}}{s}$$

Primer
Radnik jednog preduzeća ima platu 1800 din, a radnik drugog preduzeća ima platu 2100 din. Uporediti odstupanje plata od proseka preduzeća ako se zna da je u prvom preduzeću prosečna plata 1200din sa standardnom devijacijom 200, a u drugom 1500 din sa standardnom devijacijom 150.
Rešenje:

$x = 1800;\ \bar{x} = 1200;\ s_x = 200$

$y = 2100;\ \bar{y} = 1500; s_y = 150$

$$z(x) = \frac{x - \bar{x}}{s_x} = \frac{1800 - 1200}{200} = \frac{600}{200} = 3$$

$$z(y) = \frac{y - \bar{y}}{s_y} = \frac{2100 - 1500}{150} = \frac{600}{150} = 4$$

Drugi radnik više odstupa od proseka preduzeća (apsolutna odstupanja su ista, pa apsolutno odstupanje nije prava mera

Mere simetrije i spljoštenosti

Mera simetričnsti

- Koeficijent asimetrije

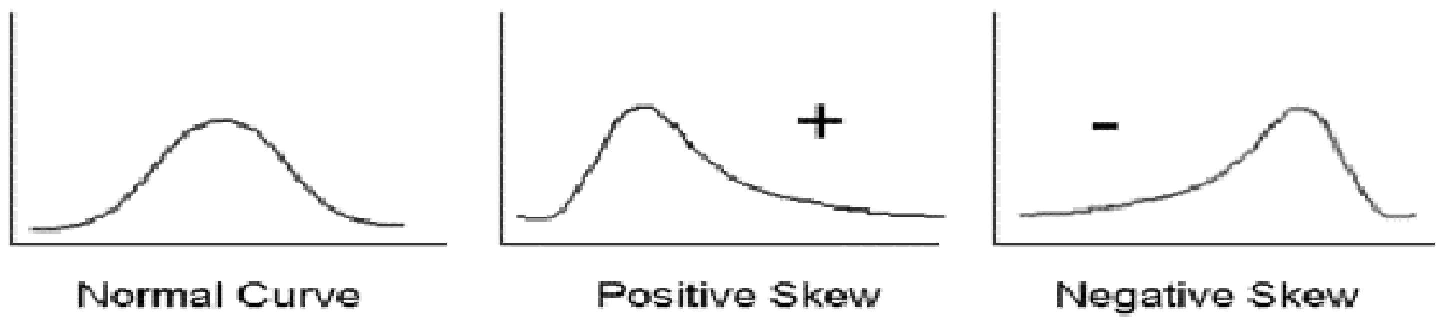
$$\alpha_3 = \frac{m_3}{s^3}$$

- Gde je s standardna devijacija , a m_3 centralni momenat 3. reda

$$m_3 = \frac{\sum_{i=1}^{n} f_i \left(x_i - \overline{x}\right)^3}{\sum_{i=1}^{n} f_i}$$

Simetričnost i asimetričnost

- Distribucija je simetrična ako je $\bar{x} = M_e = M_o$, odnosno ako je $\alpha_3 = 0$
- Distribucija je desno asimetrična ako je $M_o < M_e < \bar{x}$, odnosno ako je $\alpha_3 > 0$
- Distribucija je levo asimetrična ako je $\bar{x} < M_e < M_o$, odnosno ako je $\alpha_3 < 0$



simetrična $\bar{x}$=Me=Mo

desna asimetrija Mo<Me<$\bar{x}$

leva asimetrija $\bar{x}$<Me<Mo

Mera spljoštenosti

- Koeficijent spljoštenosti

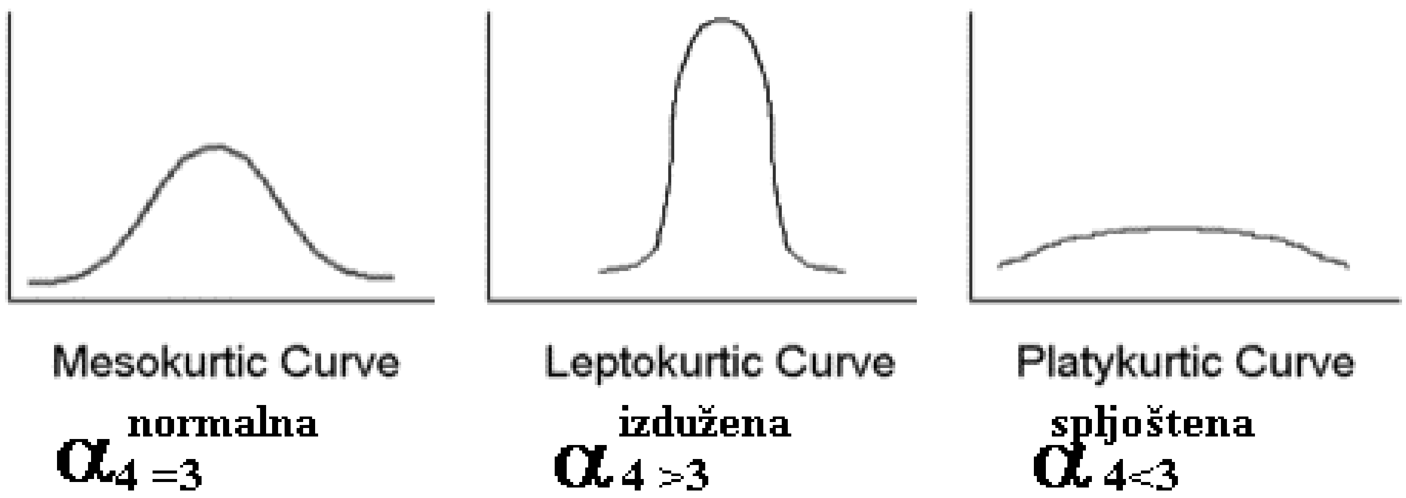
$$\alpha_4 = \frac{m_4}{s^4}$$

- Gde je s standardna devijacija , a m_4 centralni momenat 4. reda

$$m_4 = \frac{\sum_{i=1}^{n} f_i \left(x_i - \bar{x}\right)^4}{\sum_{i=1}^{n} f_i}$$

Spljoštenost

- Distribucija je normalno spljoštena ako je $\alpha_4 = 3$
- Distribucija je izdužena ako je $\alpha_4 > 3$
- Distribucija je levo spljoštena ako je $\alpha_4 < 3$



Primer: simetričnost i spljoštenost

- Ispitati simetričnost i spljoštenost distribucije frekvencija

xi	fi
8	2
9	4
10	6
11	7
12	5
13	4
14	1
15	1

Rešenje

x_i	f_i	$f_i \cdot x_i$	$x_i - \overline{x}$	$(x_i - \overline{x})^2$	$f_i(x_i - \overline{x})^2$	$(x_i - \overline{x})^3$	$f_i(x_i - \overline{x})^3$	$(x_i - \overline{x})^4$	$f_i(x_i - \overline{x})^4$
8	2	16	-3	9	18	-27	-54	81	162
9	4	36	-2	4	16	-8	-32	16	64
10	6	60	-1	1	6	-1	-6	1	6
11	7	77	0	0	0	0	0	0	0
12	5	60	1	1	5	1	5	1	5
13	4	52	2	4	16	8	32	16	64
14	1	14	3	9	9	27	27	81	81
15	1	15	4	16	16	64	64	256	256
$\sum$	30	330			86		36		638

$$\overline{x} = \frac{\sum_{i=1}^{n} f_i x_i}{\sum_{i=1}^{n} f_i} = \frac{330}{30} = 11 \qquad s^2 = \frac{\sum_{i=1}^{n} f_i (x_i - \overline{x})^2}{\sum_{i=1}^{n} f_i} = \frac{86}{30} = 2.86666 \approx 2.867 \qquad \text{s}=\sqrt{\text{s}^2} \approx 1,693$$

$$m_3 = \frac{\sum_{i=1}^{n} f_i (x_i - \overline{x})^3}{\sum_{i=1}^{n} f_i} = \frac{36}{30} = 1,2 \qquad \alpha_3 = \frac{m_3}{s^3} = \frac{1,2}{1,693^3} \approx 0.247 > 0 \text{ desna asimetričnost}$$

$$m_4 = \frac{\sum_{i=1}^{n} f_i (x_i - \overline{x})^4}{\sum_{i=1}^{n} f_i} = \frac{638}{30} \approx 21.267 \qquad \alpha_4 = \frac{m_4}{s^4} = \frac{21,267}{1,693^4} \approx 2,589 < 3$$

vrh distribucije je spljošten u odnosu na normalnu spljoštenost