

POKAZATELJI VARIJABILITETA

U pokazatelje varijacije spadaju:

- Razmak (interval) varijacije – I
- Interkvartilna razlika - IQR
- Srednje apsolutno odstupanje – SO
- Standardna devijacija - σ
- Varijansa - σ^2
- Koeficijent varijacije – V
- Standardizovano (normalizovano) odstupanje - Z

Interval (razmak) varijacije I

Predstavlja razliku ekstremnih vrednosti obeležja u nekoj seriji. Kod negrupisanih podataka interval varijacije je razlika maksimalne i minimalne vrednosti obeležja.

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

Kod intervalne distribucije frekvencija interval varijacije predstavlja razliku gornje granice poslednjeg i donje granice prvog grupnog intervala.

Primer 1. Na osnovu podataka o prinosu kukuruza (t/ha) na 10 pacela izračunati interval varijacije.

Prinos (X_i)
4,0
4,2
4,5
4,6
4,8
5,3
5,4
5,4
5,7
6,2

Interkvartilna razlika – IQR

Slično kao interval varijacije, interkvartilna razlika predstavlja razliku trećeg i prvog kvartila.

$$IQR = Q_3 - Q_1$$

Primer 1. Na osnovu podataka iz prethodnog primera zračunati interkvartilnu razliku.

$$Q_1 = X_{\frac{n}{4}+1} = X_{\frac{10}{4}+1} = X_{3,5} = X_3 = 4,5$$

$$Q_3 = X_{\frac{3n}{4}+1} = X_{\frac{3 \cdot 10}{4}+1} = X_{8,5} = X_8 = 5,4$$

$$IQR = Q_3 - Q_1 = 5,4 - 4,5 = 0,9$$

Srednje apsolutno odstupanje SO

Srednje apsolutno odstupanje je količnik zbira odstupanja individualnih vrednosti obeležja od njihovog proseka i njihovog broja. Srednje apsolutno odstupanje se utvrđuje bez obzira na predznak odstupanja.

Srednje apsolutno odstupanje kod negrupisanih podataka:

$$SO = \frac{\sum |X_i - \bar{X}|}{N}$$

Primer 1. Izračunati SO za sledeću seriju podataka:

Prinos pšenice X_i	$ X_i - \bar{X} $
2.5	0.4
2.8	0.1
2.9	0
3.1	0.2
3.2	0.3
Σ 14,5	1

$$SO = \frac{\sum |X_i - \bar{X}|}{N}$$

$$\bar{X} = \frac{\sum X_i}{N} = \frac{14,5}{5} = 2,9t / ha$$

$$SO = \frac{1,0}{5} = 0,2t / ha$$

Srednje apsolutno odstupanje kod grupisanih podataka:

$$So = \frac{\sum f_i |X_i - \bar{X}|}{\sum f_i}$$

$$SO = \frac{\sum f_i |X_i - \bar{X}|}{N}$$

Primer 2. Izračunati srednje apsolutno odstupanje iz sledećih podataka.

Prinos pšenice (t/ha) X_i	Površina (ha) f_i	$f_i X_i$	$ X_i - \bar{X} $	$f_i X_i - \bar{X} $
3	8	24	0.474	3.792
3.2	12	38.4	0.274	3.288
3.6	15	54	0.126	1.89
3.7	9	33.3	0.226	2.034
4	6	24	0.526	3.154
Σ	50	173.7		14.16

$$SO = \frac{\sum f_i |X_i - \bar{X}|}{N}$$

$$\bar{X} = \frac{\sum f_i X_i}{N} = \frac{173,7}{50} = 3,474t / ha$$

$$SO = \frac{14,16}{50} = 0,283t / ha$$

Standardna devijacija (σ)

Standardna devijacija je kvadratni koren iz sredine kvadrata odstupanja vrednosti obeležja od aritmetičke sredine. Standardna devijacija je najviše upotrebljivan pokazatelj disperzije.

Za negrupisane podatke standardna devijacija se može izračunati na dva načina:

1. na bazi odstupanja vrednosti obeležja od proseka

$$\sigma = \sqrt{\frac{\sum (X_i - \bar{X})^2}{N}}$$

2. izračunavanje direktno iz podataka

$$\sigma = \sqrt{\frac{\sum X_i^2 - \frac{(\sum X_i)^2}{N}}{N}}$$

Primer1. Izračunati standardnu devijaciju na osnovu podataka o nivou padavina u toku jedne nedelje.

Prvi način:

Nivo padavina (mm) X_i	$(X_i - \bar{X})$	$(X_i - \bar{X})^2$
4	-6	36
6	-4	16
8	-2	4
11	1	1
12	2	4
13	3	9
16	6	36
Σ 70	0	106

$$\sigma = \sqrt{\frac{\Sigma(X_i - \bar{X})^2}{N}}$$

$$\bar{X} = \frac{\Sigma X_i}{N} = \frac{70}{7} = 10mm$$

$$\sigma = \sqrt{\frac{106}{7}} = \sqrt{15,143} = 3,89mm$$

Drugi način:

Nivo padavina (mm) X_i	X_i^2
4	16
6	36
8	64
11	121
12	144
13	169
16	256
Σ 70	806

$$\sigma = \sqrt{\frac{\Sigma X_i^2 - \frac{(\Sigma X_i)^2}{N}}{N}}$$

$$\sigma = \sqrt{\frac{806 - \frac{(70)^2}{7}}{7}} = \sqrt{\frac{806 - 700}{7}} = \sqrt{15,143} = 3,89mm$$

Za grupisane podatke (kod distribucije frekvencija) standardna devijacija se može izračunati na tri načina:

1. na bazi odstupanja vrednosti obeležja od proseka

$$\sigma = \sqrt{\frac{\sum f_i (X_i - \bar{X})^2}{\sum f_i}}$$

2. izračunavanje direktno iz podataka

$$\sigma = \sqrt{\frac{\sum f_i X_i^2 - \frac{(\sum f_i X_i)^2}{N}}{N}}$$

3. skraćeno izračunavanje

$$\sigma = i \cdot \sqrt{\frac{\sum f_i d_i^2 - \frac{(\sum f_i d_i)^2}{N}}{N}}$$

i – dužina grupnog razmaka

f_i – frekvencija datog obeležja

d_i – transformisana vrednost obeležja

Primer 2. Na osnovu podataka o broju zrna po klasu kod pšenice sorte "san pastore" izračunati standardnu devijaciju.

Prvi način:

Broj zrna X_i	Broj klasova f_i	$f_i X_i$	$(X_i - \bar{X})$	$(X_i - \bar{X})^2$	$f_i (X_i - \bar{X})^2$
20	2	40	-26.6	707.56	1415.12
25	3	75	-21.6	466.56	1399.68
30	5	150	-16.6	275.56	1377.8
35	17	595	-11.6	134.56	2287.52
40	37	1480	-6.6	43.56	1611.72
45	37	1665	-1.6	2.56	94.72
50	48	2400	3.4	11.56	554.88
55	31	1705	8.4	70.56	2187.36
60	18	1080	13.4	179.56	3232.08
65	2	130	18.4	338.56	677.12
Σ	200	9320			14838

$$\sigma = \sqrt{\frac{\sum f_i (X_i - \bar{X})^2}{\sum f_i}}$$

$$\bar{X} = \frac{\sum f_i X_i}{\sum f_i} = \frac{9320}{200} = 46,6 \text{ zrna}$$

$$\sigma = \sqrt{\frac{14838}{200}} = 8,61 \text{ zrna}$$

Drugi način:

Broj zrna X_i	Broj klasova f_i	$f_i X_i$	$f_i X_i^2$
20	2	40	800
25	3	75	1875
30	5	150	4500
35	17	595	20825
40	37	1480	59200
45	37	1665	74925
50	48	2400	120000
55	31	1705	93775
60	18	1080	64800
65	2	130	8450
Σ	200	9320	449150

$$\sigma = \sqrt{\frac{\sum f_i X_i^2 - \frac{(\sum f_i X_i)^2}{N}}{N}}$$

$$\sigma = \sqrt{\frac{449150 - \frac{(9320)^2}{200}}{200}} = 8,61 \text{ zrna}$$

$$i = 5$$

Treći način:

Broj zrna X_i	Broj klasova f_i	d_i	$f_i d_i$	$f_i d_i^2$
20	2	-6	-12	72
25	3	-5	-15	75
30	5	-4	-20	80
35	17	-3	-51	153
40	37	-2	-74	148
45	37	-1	-37	37
50	48	0	0	0
55	31	1	31	31
60	18	2	36	72
65	2	3	6	18
Σ	200		-136	686

$$\sigma = i \cdot \sqrt{\frac{\sum f_i d_i^2 - \frac{(\sum f_i d_i)^2}{N}}{N}}$$

$$\sigma = 5 \cdot \sqrt{\frac{686 - \frac{(-136)^2}{200}}{200}} = 8,61 \text{ zrna}$$

Kvadrat standardne devijacije predstavlja varijansu. *Varijansa* (σ^2) takođe može da se izračuna na više načina.

Izračunavanje varijanse kod negrupisanih podataka:

1. na bazi odstupanja obeležja od proseka

$$\sigma^2 = \frac{\sum(X_i - \bar{X})^2}{N}$$

2. izračunavanje direktno iz podataka

$$\sigma^2 = \frac{\sum X_i^2 - \frac{(\sum X_i)^2}{N}}{N}$$

Kod distribucije frekvencija varijansa se može izračunati na tri načina:

1. na bazi odstupanja obeležja od proseka

$$\sigma^2 = \frac{\sum f_i (X_i - \bar{X})^2}{\sum f_i}$$

2. izračunavanje direktno iz podataka

$$\sigma^2 = \frac{\sum f_i X_i^2 - \frac{(\sum f_i X_i)^2}{N}}{N}$$

3. skraćeno izračunavanje

$$\sigma^2 = i^2 \cdot \frac{\sum f_i d_i^2 - \frac{(\sum f_i d_i)^2}{N}}{N}$$

Koeficijent varijacije - V

Koeficijent varijacije - je relativni pokazatelj disperzije. Koeficijent varijacije pokazuje koliko procentualno iznosi standardna devijacija u odnosu na aritmetičku sredinu.

$$V = \frac{\sigma}{\bar{X}} \cdot 100(\%)$$

Primer 1. U jednom prasilištu 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 kg². Uporediti varijabilitet telesne mase i varijabilitet broja prasica. Obrazložiti.

$$V_1 = \frac{\sigma_1}{\bar{X}_1} \times 100 (\%) = \frac{2}{12} \times 100 (\%) = 16,67\%$$

$$V_2 = \frac{\sigma_2}{\bar{X}_2} \times 100 (\%) = \frac{\sqrt{2,56}}{8} \times 100 (\%) = 20\%$$

Varijabilitet telesne mase je veći u odnosu na varijabilitet broja prasića ($V_2 > V_1$).

Standardizovano (normalizovano) odstupanje - Z

Standardizovano odstupanje je takođe relativni pokazatelj disperzije obeležja.

Standardizovano odstupanje je mera udaljenosti pojedinih vrednosti obeležja od aritmetičke sredine iskazana u odnosu na standardnu devijaciju.

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

Standardizovano odstupanje može biti izračunato kao pozitivna ili negativna vrednost. Za razliku od ostalih pokazatelja varijacije, standardizovano odstupanje pokazuje varijabilitet pojedinačnih vrednosti obeležja.

Primer1. Na osnovu podataka o broju zrna po klasu iz prethdnog primera utvrditi vrednost standardizovanog odstupanja za vrednost obeležja $X_i = 50$.

$$X_i = 50$$

$$\bar{X} = 46,6$$

$$\sigma = 8,61$$

$$Z_i = \frac{50 - 46,6}{8,61} = 0,395 \Rightarrow 39,5\%$$

Osobine varijanse

- Varijansa je pokazatelj varijacije izražen istim jedinicama mere kao i posmatrano obeležje, ali se uz izračunatu vrednost varijanse nikada ne stavlja ju jedinice mere.

- Ako su sve vrednosti obeležja u nekoj seriji međusobno jednake varijansa i standardna devijacija su jednake nuli.

- Ako svim vrednostima obeležja u nekoj seriji dodamo ili oduzmemo konstantu varijansa novih vrednosti obeležja se ne menja.

$$X_i \pm c \Rightarrow X_i'$$

$$\sigma^2 = \sigma'^2$$

- Ako sve vrednosti obeležja u nekoj seriji pomnožimo konstantom, varijansa novih vrednosti obeležja biće jednaka proizvodu kvadrata konstante i prethodno izračunate varijanse.

$$X_i \cdot c \Rightarrow X'_i$$

$$\sigma'^2 = \sigma^2 \cdot c^2$$