

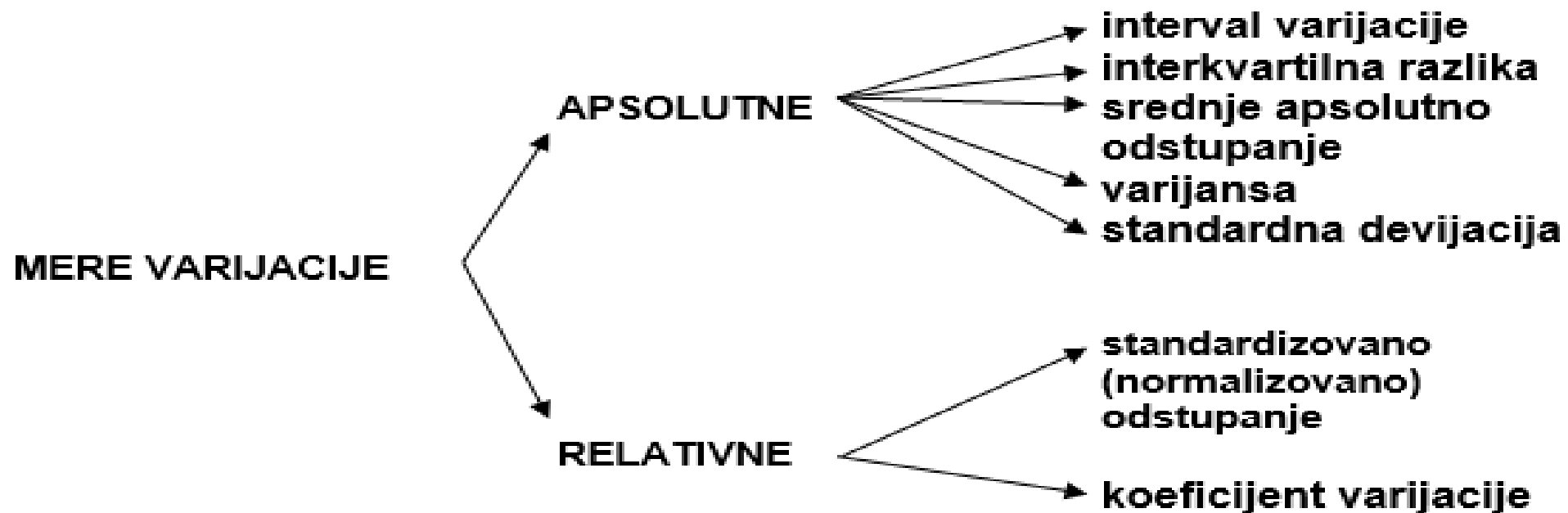


STATISTIKA

POKAZATELJI STATISTIČKIH SERIJA – MERE VARIJACIJA (DISPERZIJE)

POKAZATELJI STATISTIČKIH SERIJA – MERE VARIJACIJA ILI MERE DISPERZIJE

- Može se desiti da dve ili više statističkih serija imaju istu srednju vrednost i ako im se elementi međusobno jako razlikuju. Zato za statistički opis serije pored srednje vrednosti treba koristiti i odgovarajuću meru disperzije, tj. varijacije.



POKAZATELJI STATISTIČKIH SERIJA – APSOLUTNE MERE VARIJACIJA (DISPERZIJE)

- Interval varijacije je razlika ekstremnih vrednosti obeležja u seriji:

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

- Kod intervalnih distribucija frekvencija, interval varijacije je razlika između gornje granice poslednje grupe i donje granice prve grupe.
- Interval varijacije se lako izračunava, međutim, njegova primena je ograničena jer se zasniva samo na dve vrednosti obeležja.
- Interkvartilna razlika je razlika između vrednosti obeležja na mestu trećeg i prvog kvartila, tj.

$$I_q = Q_3 - Q_1.$$

POKAZATELJI STATISTIČKIH SERIJA – APSOLUTNE MERE VARIJACIJA (DISPERZIJE)

- Prvi kvartil (Q_1) je vrednost obeležja na mestu $1/4$ sredene serije, tj.

$$Q_1 = x_{\frac{n+1}{4}}$$

- a treći kvartil je vrednost obeležja na mestu $3/4$ sredene serije, odnosno:

$$Q_3 = x_{3\frac{n+1}{4}}$$

- Za intervalne serije kvartili se određuju pomoću sledećih formula:

$$Q_1 = g_1 + \frac{\frac{\sum_{i=1}^k f_i}{4} - K_{i-1}}{f_{q_1}} \cdot d$$

$$Q_3 = g_1 + \frac{3\frac{\sum_{i=1}^k f_i}{4} - K_{i-1}}{f_{q_3}} \cdot d$$

$g_{1(3)}$ - donja granica grupnog intervala u kojem se nalazi mesto prvog (trećeg) kvartila;

f_{q_1} - frekvencija grupnog intervala u kojem se nalazi mesto prvog kvartila;

f_{q_3} - frekvencija grupnog intervala u kojem se nalazi mesto trećeg kvartila;

K_{i-1} - kumulativna frekvencija koja prethodi toj klasi i

d - dužina klasnog intervala.

POKAZATELJI STATISTIČKIH SERIJA – APSOLUTNE MERE VARIJACIJA (DISPERZIJE)

- Srednje absolutno odstupanje se izračunava tako što se zbir absolutnih odstupanja svih vrednosti obeležja od aritmetičke sredine podeli brojem članova serije.

- Za negrupisane podatke formula je:

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

- Za grupisane podatke formula je:

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

- Sredina kvadrata odstupanja vrednosti obeležja od aritmetičke sredine naziva se varijansa.

POKAZATELJI STATISTIČKIH SERIJA – APSOLUTNE MERE VARIJACIJA (DISPERZIJE)

- Za negrupisane podatke formula je:

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n} \quad \text{ili} \quad s^2 = \frac{\sum_{i=1}^n x_i^2}{n} - \bar{x}^2$$

- Za grupisane podatke formula je:

$$s^2 = \frac{\sum_{i=1}^k f_i (x_i - \bar{x})^2}{\sum_{i=1}^k f_i} \quad \text{ili} \quad s^2 = \frac{\sum_{i=1}^k f_i x_i^2}{\sum_{i=1}^k f_i} - \bar{x}^2$$

POKAZATELJI STATISTIČKIH SERIJA – APSOLUTNE MERE VARIJACIJA (DISPERZIJE)

- Standardna devijacija predstavlja kvadratni koren iz varijanse.
- Za negrupisane podatke formula je:

$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}} \quad \text{ili} \quad s = \sqrt{\frac{\sum_{i=1}^n x_i^2}{n} - \bar{x}^2}$$

- Za grupisane podatke formula je:

$$s = \sqrt{\frac{\sum_{i=1}^k f_i (x_i - \bar{x})^2}{\sum_{i=1}^k f_i}} \quad \text{ili} \quad s = \sqrt{\frac{\sum_{i=1}^k f_i x_i^2}{\sum_{i=1}^k f_i} - \bar{x}^2}$$

POKAZATELJI STATISTIČKIH SERIJA – RELATIVNE MERE VARIJACIJA (DISPERZIJE)

- Poređenje varijabiliteta različitih serija na osnovu apsolutnih mera varijacije moguće je samo u slučaju kada se odnose na isto obeležje i kada imaju isti prosečan nivo obeležja. Takvi slučajevi su retki, pa je potrebno varijabilitet iskazati nekom neimenovanom merom.

- Koeficijent varijacije je procentualni odnos standardne devijacije i aritmetičke sredine:

$$C_v = \frac{S}{\bar{x}} \cdot 100.$$

- S obzirom da se iskazuje u procentima koeficijent varijacije ne zavisi od jedinice mere obeležja i zato se može koristiti za poređenje varijabiliteta raznoimenih obeležja.

POKAZATELJI STATISTIČKIH SERIJA – RELATIVNE MERE VARIJACIJA (DISPERZIJE)

- *Standardizovano* ili, kako se još naziva, *normalizovano odstupanje* daje informaciju za koliko je standardnih devijacija posmatrana vrednost obeležja udaljena od aritmetičke sredine:

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

- Na osnovu predznaka standardizovanog odstupanja se vidi da li je vrednost obeležja manja ili veća od aritmetičke sredine.

POKAZATELJI STATISTIČKIH SERIJA – RELATIVNE MERE VARIJACIJA (DISPERZIJE)

- Izračunate pojedinačne standardizovane vrednosti obeležja formiraju novu takozvanu normalizovanu ili standardizovanu distribuciju.
- Aritmetička sredina distribucije standardizovanog obeležja je jednaka nuli ($\bar{z}=0$), a standardna devijacija iznosi jedan ($s_z=1$). Grafički prikazane distribucije sa standardizovanim obeležjem sve imaju aritmetičku sredinu u koordinatnom početku i jednak raspon.
- Standardizovano obeležje je pogodno za upoređivanje varijabiliteta pojedinačnih podataka jedne serije, kao i distribucija frekvencija različitog oblika i za različite vrste obeležja.

POKAZATELJI STATISTIČKIH SERIJA – MERE VARIJACIJA (DISPERZIJE)

- **Primer 1:** Evidentirana je sledeća visina grebena kod 7 grla crvene danske rase:

Visina grebena (cm): 127 131 135 130 125 128 134

Izračunati sve apsolutne i relativne pokazatelje disperzije.

Rešenje:

$$I_v = x_{\max} - x_{\min} = 135 - 125 = 10$$

$$I_Q = Q_3 - Q_1 = 134 - 127 = 7$$

$$Q_1 = x_{\frac{n+1}{4}} = x_{\frac{7+1}{4}} = x_2 = 127$$

$$Q_3 = x_{3\frac{n+1}{4}} = x_{3\frac{7+1}{4}} = x_6 = 134$$

POKAZATELJI STATISTIČKIH SERIJA – MERE VARIJACIJA (DISPERZIJE)

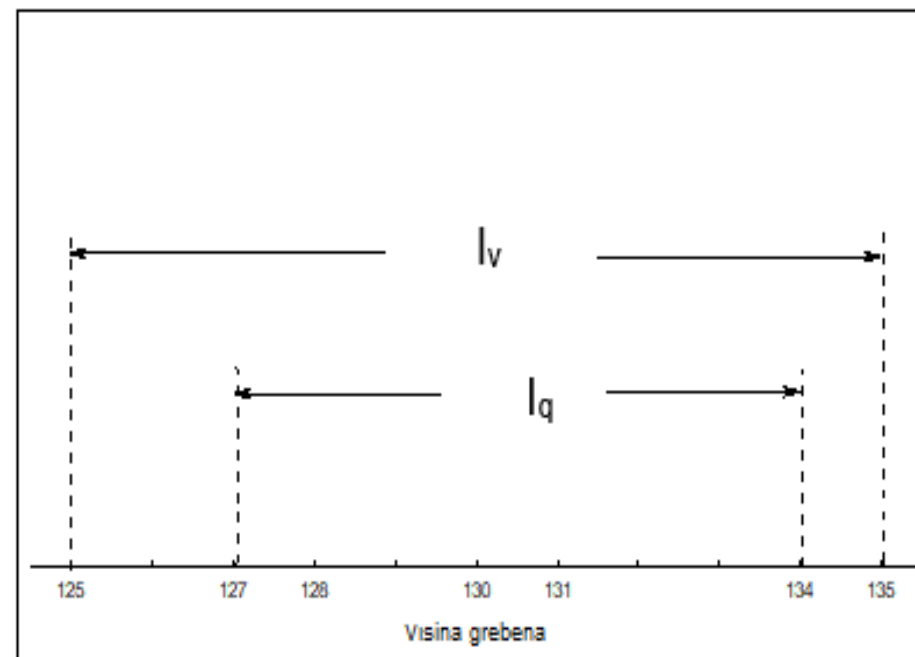
$$SO = \frac{\sum_{i=1}^n |x_i - \bar{x}|}{n} = \frac{20}{7} = 2,8571$$

$$\bar{x} = \frac{\sum x_i}{n} = \frac{910}{7} = 130$$

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n} = \frac{80}{7} = 11,4286$$

odnosno

$$s^2 = \frac{\sum_{i=1}^n x_i^2}{n} - \bar{x}^2 = \frac{118380}{7} - 130^2 = 11,4286$$



Slika 3.13: Odnos intervala varijacije i interkvartilne razlike

POKAZATELJI STATISTIČKIH SERIJA – MERE VARIJACIJA (DISPERZIJE)

Tabela 3.11: Radna tabela

| Visina grebena (cm) x_i | $x_i - \bar{x}$ | $ x_i - \bar{x} $ | $(x_i - \bar{x})^2$ | x_i^2 |
|------------------------------|-----------------|-------------------|---------------------|---------------|
| 125 | -5 | 5 | 25 | 15625 |
| 127 | -3 | 3 | 9 | 16129 |
| 128 | -2 | 2 | 4 | 16384 |
| 130 | 0 | 0 | 0 | 16900 |
| 131 | 1 | 1 | 1 | 17161 |
| 134 | 4 | 4 | 16 | 17956 |
| 135 | 5 | 5 | 25 | 18225 |
| 910 | 0 | 20 | 80 | 118380 |

POKAZATELJI STATISTIČKIH SERIJA – MERE VARIJACIJA (DISPERZIJE)

$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}} = \sqrt{\frac{80}{7}} = \sqrt{11,4286} = 3,3806$$

ili

$$s = \sqrt{\frac{\sum_{i=1}^n x_i^2}{n} - \bar{x}^2} = \sqrt{\frac{118380}{7} - 130^2} = \sqrt{11,4286} = 3,3806$$

$$C_v = \frac{s}{\bar{x}} \cdot 100 = \frac{3,3806}{130} \cdot 100 = 2,60\%$$

POKAZATELJI STATISTIČKIH SERIJA – MERE VARIJACIJA (DISPERZIJE)

$$z_1 = \frac{x_1 - \bar{x}}{s} = \frac{125 - 130}{3,3806} = -1,4790$$

$$z_5 = \frac{x_5 - \bar{x}}{s} = \frac{131 - 130}{3,3806} = 0,2958$$

$$z_2 = \frac{x_2 - \bar{x}}{s} = \frac{127 - 130}{3,3806} = -0,8874$$

$$z_6 = \frac{x_6 - \bar{x}}{s} = \frac{134 - 130}{3,3806} = 1,1832$$

$$z_3 = \frac{x_3 - \bar{x}}{s} = \frac{128 - 130}{3,3806} = -0,5916$$

$$z_7 = \frac{x_7 - \bar{x}}{s} = \frac{135 - 130}{3,3806} = 1,4790$$

$$z_4 = \frac{x_4 - \bar{x}}{s} = \frac{130 - 130}{3,3806} = 0,0000$$

POKAZATELJI STATISTIČKIH SERIJA – MERE VARIJACIJA (DISPERZIJE)

- **Primer 2:** Prosečan prinos kukuruza po hektaru je 10 tona sa varijansom 2,5 t². Prosečan broj biljaka kukuruza po hektaru je 60000 sa standardnom devijacijom 2000. Koje obeležje više varira? (Koje obeležje ima veću disperziju?)

$$c_{v_x} = \frac{s_x}{\bar{x}} \cdot 100 = \frac{\sqrt{2,5}}{10} \cdot 100 = 15,81\%$$

$$c_{v_y} = \frac{s_y}{\bar{y}} \cdot 100 = \frac{2000}{60000} \cdot 100 = 3,33\%$$

POKAZATELJI STATISTIČKIH SERIJA – MERE VARIJACIJA (DISPERZIJE)

- **Primer 3:** Prosečna masa ploda jedne sorte kajsija je 77g sa standardnom devijacijom 3g. Kod jednog ploda ove sorte kajsija izmerena je masa od 72g. Prosečna masa ploda druge sorte kajsija je 60g, a srednje kvadratno odstupanje $4g^2$. Jedan plod iz druge sorte imao je težinu od 65g. Koji plod je bolji prema masi u odnosu na prosečnu masu sorte?

$$\begin{array}{ll} \bar{x}_1 = 77 & \bar{x}_2 = 60 \\ s_1 = 3 & s_2^2 = 4 \\ x_{1_i} = 72 & x_{2_i} = 65 \end{array}$$

$$z_1 = \frac{x_1 - \bar{x}_1}{s_1} = \frac{72 - 77}{3} = -1,67 \quad z_2 = \frac{x_{2_i} - \bar{x}_2}{s_2} = \frac{65 - 60}{2} = 2,5$$

- U odnosu na masu pripadajuće sorte bolji je plod iz druge sorte, jer ima masu veću za 2,5 standardnih devijacija od proseka .