

VARIAČNÉ HODNOTY (MIERY VARIABILITY)

- Dva štatistické súbory popisujúce rovnaký jav sa môžu líšiť nielen svojimi strednými hodnotami, ale tiež spôsobom, akým sú usporiadané konkrétne hodnoty znaku okolo týchto stredných hodnôt.
- Skúmané hodnoty znaku môžu byť okolo strednej hodnoty usporiadané veľmi voľne (t.j. rozptýlene) až veľmi tesne (t.j. hodnoty sú rozmiestnené v bezprostrednom susedstve so strednou hodnotou).

VARIABILITA – PREMENLIVOSŤ - odlišnosti skúmaného znaku vnútri daného súboru, ktoré vyplývajú z viacerých príčin, ktoré môžu mať charakter vecný, časový a priestorový.

- Určením variability rozširujeme informácie o skúmaných štatistických súboroch, pričom vo všeobecnosti platí, že malá variabilita sledovaného znaku svedčí o vyrovnanosti jednotiek súboru a veľká variabilita hovorí o heterogenite jednotiek z hľadiska skúmaného znaku.

- Ak chceme charakterizovať variabilitu, tak aby čo najpresnejšie odrážala variácie celého skúmaného radu hodnôt, je potrebné vychádzať z odchýlok jednotlivých hodnôt znaku **od zvolenej konštanty**. Ako konštanta sa používa niektorá zo stredných hodnôt, t.j. aritmetický priemer, medián, alebo módu

- Miery variability sú **odchýlky** od niektorej **strednej hodnoty** alebo odchýlky **medzi jednotlivými hodnotami**.

- Miery variability sú obdobou chýb merania, vyplývajúcich z nemožnosti zistiť akúkoľvek premennú veličinu s absolútnou presnosťou.

MIERY VARIABILITY ZÍSKANÉ VÝPOČTOM ZO VŠETKÝCH HODNÔT RADU

- Najdôležitejšie miery variability vychádzajú z veľkosti všetkých hodnôt súboru.

1. ABSOLÚTNE MIERY VARIABILITY

- umožňujú charakterizovať variabilitu v štatistickom súbore

ODCHÝLKA

OD ARITMETICKÉHO PRIEMERU d_i

- najčastejšie používaná miera variability
- súčet jednotlivých odchýlok hodnôt od aritmetického priemeru je rovný nule, preto sa tieto odchýlky vyjadrujú v tvare **absolútnej hodnoty**: $d_i = |x_i - \bar{x}|$

♦ podobne je možné vyjadriť ODCHÝLKU OD MEDIÁNU: $d_i = |x_i - \tilde{x}|$

♦ ODCHÝLKA OD MÓDU: $d_i = |x_i - \hat{x}|$

PRIEMERNÁ ODCHÝLKA

OD ARITMETICKÉHO PRIEMERU

- **jednoduchá forma:**

- je to vlastne jednoduchý aritmetický priemer jednotlivých absolútnych odchýlok od aritmetického priemeru

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

- **vážená forma:**

$$\bar{d}_{\bar{x}} = \frac{\sum_{i=1}^k |x_i - \bar{x}| \cdot n_i}{\sum_{i=1}^k n_i} = \frac{\sum_{i=1}^k |d_i| \cdot n_i}{\sum_{i=1}^k n_i}$$

- ak hodnotíme opakované merania na tom istom objekte, priemerná odchýlka od aritmetického priemeru sa označuje ako **ABSOLÚTNA STREDNÁ CHYBA**

- ♦ podobne sa vyjadruje PRIEMERNÁ ODCHÝLKA OD MEDIÁNU $\bar{d}_{\tilde{x}}$
- ♦ a aj PRIEMERNÁ ODCHÝLKA OD MODU $\bar{d}_{\hat{x}}$

STREDNÁ DIFERENCIÁCIA d_{dif} alebo D

- špecifická forma priemernej odchýlky, ktorá sa niekedy používa na presné meranie variability
- používa sa pri veľmi heterogénnych súboroch, kedy nemá zmysel triediť variačný rad na rad rozloženia početností
- vyjadruje priemer rozdielov každej hodnoty od každej (t.j. vzájomné rozdiely hodnoty x_n od všetkých ďalších až po x_{n-1} , x_{n-1} od všetkých ďalších až po x_{n-2} , ... až po x_2-x_1 :

$$d_{dif} = \frac{\sum_{i=1}^n \sum_{j=1}^n |x_i - x_j|}{n \cdot (n-1)}$$

počet všetkých vzájomných rozdielov n hodnôt znaku je $n \cdot (n-1)$, pretože každá z hodnôt sa líši od ostatných v $(n-1)$ prípadoch

ROZPTYL, VARIANCIA $S_x^2, \sigma^2, \text{var } x$

(PRIEMERNÁ ŠTVORCOVÁ ODCHÝLKA, PRIEMERNÁ KVADRATICÁ ODCHÝLKA)

- táto veľmi často používaná miera, charakterizuje vzájomnú odlišnosť jednotlivých hodnôt znaku a zároveň aj varianciu jednotlivých hodnôt znaku od priemeru
- je definovaný ako aritmetický priemer štvorcov odchýlok jednotlivých hodnôt znaku od ich aritmetického priemeru

- **jednoduchá forma:** $s_x^2 = \text{var } x = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}$

- pri jeho výpočte nedochádza ku vzájomnej kompenzácii kladných a záporných odchýlok od aritmetického priemeru, pretože sa počíta so štvorcem odchýlok, ktoré majú vždy kladnú hodnotu

- **vážená forma:** $s_x^2 = \text{var } x = \frac{\sum_{i=1}^n (x_i - \bar{x})^2 \cdot n_i}{\sum_{i=1}^n n_i}$

SMERODAJNÁ (ŠTANDARDNÁ) ODCHÝLKA s_x, σ

- kladná druhá odmocnina rozptylu
- stredné hodnoty nie sú umocnené na druhú a rozptyl je v podobe štvorca, preto sa veľmi často používa jeho druhá odmocnina

- **jednoduchá forma:** $s_x = \sqrt{\text{var } x} = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}}$

- **vážená forma:** $s_x = \sqrt{\text{var } x} = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2 \cdot n_i}{\sum_{i=1}^n n_i}}$

2. RELATÍVNE (POMERNÉ) MIERY VARIABILITY

- umožňujú porovnávať variabilitu v rôznych štatistických súboroch
- zvyčajne charakterizujú variabilitu vzhľadom na aritmetický priemer
- najčastejšie používané relatívne miery variability sú relatívna priemerná odchýlka a variačný koeficient

RELATÍVNA PRIEMERNÁ ODCHÝLKA

- je to podiel absolútnej priemernej odchýlky a tej strednej hodnoty, od ktorej boli počítané odchýlky
- najčastejšie sa používa **RELATÍVNA PRIEMERNÁ ODCHÝLKA OD ARITMETICKÉHO PRIEMERU**
- podiel priemernej odchýlky a aritmetického priemeru

$$\delta_x = \overline{d'_x} = \frac{\overline{d_x}}{\bar{x}}$$

- vyjadruje sa v %
- ak hodnotíme opakované merania na tom istom objekte relatívna priemerná odchýlka od aritmetického priemeru sa označuje aj ako **RELATÍVNA (POMERNÁ) STREDNÁ CHYBA**

VARIAČNÝ KOEFICIENT

- podiel smerodajnej odchýlky a aritmetického priemeru daného znaku:

$$V_x = \frac{S_x}{\bar{x}}$$

- často sa vyjadruje v %
- definoval významný anglický štatistik Karl Pearson (1858-1936)

MIERY VARIABILITY ZÍSKANÉ VÝPOČTOM IBA Z NIEKTORÝCH HODNÔT RADU

VARIAČNÉ ROZPÄTIE R_x , alebo v_x

- najjednoduchšia a ľahko vypočítateľná, no zároveň najmenej presná miera variability
- vyjadruje rozdiel najvyššej a najnižšej hodnoty sledovaného znaku:

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

- je to len približná miera variability, pretože je vypočítaná z extrémnych hodnôt a nevystihuje skutočnú variabilitu vnútri medzi extrémami (aj dve úplne odlišné rady, ak majú náhodou zhodné extrémne hodnoty môžu mať rovnaké variačné rozpätie)

KVANTILOVÉ ODCHÝLKY

- sú lepšou mierou variability, pretože nie sú založené len na náhodných extrémnych hodnotách
- vychádzajú z kladných odchýlok susedných kvantilov

KVANTILOVÉ ROZPÄTIE:

- rozdiel medzi posledným a prvým kvantilom
- pr. **KVARTILOVÉ ROZPÄTIE:** $R_K = \tilde{x}_{75} - \tilde{x}_{25}$
alebo $Q_r = x_{Q3} - x_{Q1}$

- pr. **DECILOVÉ ROZPÄTIE:** $R_D = \tilde{x}_{90} - \tilde{x}_{10}$

KVANTILOVÁ ODCHÝLKA:

- aritmetický priemer dvoch susedných kvantilov
- najčastejšie sa používa **KVANTILOVÁ ODCHÝLKA:** S_K alebo Q_n

$$S_K = \frac{(\tilde{x}_{75} - \tilde{x}) + (\tilde{x} - \tilde{x}_{25})}{2} = \frac{\tilde{x}_{75} - \tilde{x}_{25}}{2}$$

- pr. **DECILOVÁ ODCHÝLKA:**

$$S_D = \frac{(\tilde{x}_{90} - \tilde{x}_{80}) + (\tilde{x}_{80} - \tilde{x}_{70}) + \dots + (\tilde{x}_{20} - \tilde{x}_{10})}{8} = \frac{(\tilde{x}_{90} - \tilde{x}_{10})}{8}$$