

MATURITNÉ PRÍKLADY Z MATEMATIKY

MATURITNÝ OKRUH 20: ŠTATISTIKA

1. príklad (220/Pr. 2)

Zadanie: Dokážte, že rozptyl $s^2 = \frac{1}{n} \sum_{i=1}^n x_i^2 - (\bar{x})^2$.

Dôkaz (priamy):

Podľa definície $s^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$. Teda

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

ČBTD.

2. príklad (222/1)

Zadanie: Nech y je znak, ktorý dostaneme zo znaku x pomocou rovnosti $y = 2x + 3$. Dokážte, že $\bar{y} = 2\bar{x} + 3$, $\text{mod}(y) = 2 \cdot \text{mod}(x) + 3$, $\text{med}(y) = 2 \cdot \text{med}(x) + 3$, $s(y) = 2 \cdot s(x)$, $r_{xy} = 1$.

Dôkaz (priamy):

Nech x_1, x_2, \dots, x_n sú hodnoty štatistického súboru ($x_1 \leq x_2 \leq \dots \leq x_n$) a nech platí $\forall i \in \{1, 2, \dots, n\}$; $y_i = 2x_i + 3$. Potom

$$1. \quad \bar{y} = \frac{1}{n} \sum_{i=1}^n y_i = \frac{1}{n} \sum_{i=1}^n (2x_i + 3) = \frac{1}{n} \sum_{i=1}^n 2x_i + \frac{1}{n} \sum_{i=1}^n 3 = 2 \cdot \frac{1}{n} \sum_{i=1}^n x_i + \frac{3n}{n} = \underline{\underline{2\bar{x} + 3}} \quad \text{ČBTD.}$$

2. $\text{mod}(y)$:

Je to najčastejšie sa vyskytujúca hodnota znaku y . Keďže znaky x a y spolu priamo súvisia, musí sa rovnať najčastejšie sa vyskytujúcej hodnote znaku x vynásobenej dvomi a sčítanej s tromi, čiže $\text{mod}(y) = 2 \cdot \text{mod}(x) + 3$ ČBTD.

3. $\text{med}(y)$:

Keďže funkcia, ktorá priraduje hodnotu znaku y je prostá, tak stredná hodnota z postupnosti x_1, x_2, \dots, x_n priamo určí aj prostrednú hodnotu postupnosti y_1, y_2, \dots, y_n , čiže $\text{med}(y) = 2 \cdot \text{med}(x) + 3$ ČBTD.

$$4. \quad s_y = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2} = \sqrt{\frac{1}{n} \sum_{i=1}^n (2x_i - 3 - (2\bar{x} + 3))^2} = \sqrt{4 \cdot \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2} = 2 \cdot \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2} = \underline{\underline{2s_x}}$$

ČBTD.

$$5. \quad r_{xy} = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{s_x s_y} = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(2x_i + 3 - (2\bar{x} + 3))}{s_x \cdot 2s_x} = \frac{2 \cdot \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(x_i - \bar{x})}{s_x \cdot 2s_x} = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2}{s_x^2} = \frac{s_x^2}{s_x^2} = \underline{\underline{1}} \quad \text{ČBTD.}$$

MATURITNÝ OKRUH 20: ŠTATISTIKA

3. príklad (222/2)

Zadanie: Dokážte, že mimo intervalu $(\bar{x} - 3s; \bar{x} + 3s)$ sa nachádza najviac devätina členov postupnosti $x_1, x_2, x_3, \dots, x_n$.

Dôkaz (priamy):

Nech x_1, x_2, \dots, x_n sú hodnoty štatistického súboru, pričom $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ a $s = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2}$. Nech množina A je množinou všetkých takých indexov $i \in \{1, 2, \dots, n\}$, pre ktoré platí, že hodnota x_i je mimo intervalu $(\bar{x} - 3s; \bar{x} + 3s)$. Potom platí:

$$\forall i \in A; |x_i - \bar{x}| \geq 3s \Rightarrow (x_i - \bar{x})^2 \geq 9s^2 \Rightarrow \left(\left(\sum_{\text{cez } \forall i \in A} (x_i - \bar{x})^2 \geq \sum_{\text{cez } \forall i \in A} 9s^2 \right) \wedge \left(\sum_{i=1}^n (x_i - \bar{x})^2 = ns^2 \geq \sum_{\text{cez } \forall i \in A} (x_i - \bar{x})^2 \right) \right) \Rightarrow$$

$$\Rightarrow ns^2 \geq |A| \cdot 9s^2 \Rightarrow \underline{\underline{|A| \leq \frac{n}{9}}}$$

Keďže $|A| \leq \frac{n}{9}$, tak platí aj, že najviac devätina členov postupnosti $x_1, x_2, x_3, \dots, x_n$ sa nachádza mimo intervalu $(\bar{x} - 3s; \bar{x} + 3s)$ ČBTD.

4. príklad (222/4)

Zadanie: Dokážte, že pre koeficient korelácie r_{xy} platí $|r_{xy}| \leq 1$.

Dôkaz (priamy):

Nech x_1, x_2, \dots, x_n sú hodnoty štatistického súboru pre znak x , pričom $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$, $s_x = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2}$.

Nech y_1, y_2, \dots, y_n sú hodnoty štatistického súboru pre znak y , pričom $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$, $s_y = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2}$.

Zároveň uvažujme vektory $\vec{x} = [x_1 - \bar{x}, x_2 - \bar{x}, \dots, x_n - \bar{x}]$ a $\vec{y} = [y_1 - \bar{y}, y_2 - \bar{y}, \dots, y_n - \bar{y}]$. Pre koeficient korelácie potom podľa definície platí:

$$r_{xy} = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{s_x \cdot s_y} = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2}} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} =$$

$$= \frac{\vec{x} \cdot \vec{y}}{|\vec{x}| \cdot |\vec{y}|} = \cos(\angle \vec{x}, \vec{y}) \in (-1, 1) \Rightarrow \underline{\underline{|r_{xy}| \leq 1}} \quad \text{ČBTD.}$$

MATURITNÉ PRÍKLADY Z MATEMATIKY

MATURITNÝ OKRUH 20: ŠTATISTIKA

5. príklad (nie je v MK)

Zadanie: Daná je tabuľka absolútnych početností hodnôt známok žiakov z predmetov slovenský jazyk a matematika. Zistite vzájomnú závislosť týchto známok u žiakov.

M \ S _j	1	2	3
1	6	8	1
2	4	7	1
3	-	-	1

Riešenie:

Najprv si trochu doplníme tabuľku:

M \ S _j	1	2	3	n_{iM}
1	6	8	1	15
2	4	7	1	12
3	-	-	1	1
n_{iS}	10	15	3	$n = 28$

Nech známka z matematiky je znak x a známka zo slovenčiny je znak y . Vypočítajme teda koeficient korelácie znakov x a y .

Matematika:

$$\bar{x} = \frac{1}{n} \cdot \sum_{i=1}^k x_i^* \cdot n_{iM} = \frac{1}{28} \cdot (1 \cdot 15 + 2 \cdot 12 + 3 \cdot 1) = 1,5$$

$$s_x = \sqrt{\frac{1}{n} \cdot \sum_{i=1}^k (x_i^* - \bar{x})^2 \cdot n_{iM}} = \sqrt{\frac{1}{28} \cdot ((1-1,5)^2 \cdot 15 + (2-1,5)^2 \cdot 12 + (3-1,5)^2 \cdot 1)} = 0,57$$

Slovenčina:

$$\bar{y} = \frac{1}{n} \cdot \sum_{i=1}^k y_i^* \cdot n_{iS} = \frac{1}{28} \cdot (1 \cdot 10 + 2 \cdot 15 + 3 \cdot 3) = 1,75$$

$$s_y = \sqrt{\frac{1}{n} \cdot \sum_{i=1}^k (y_i^* - \bar{y})^2 \cdot n_{iS}} = \sqrt{\frac{1}{28} \cdot ((1-1,75)^2 \cdot 10 + (2-1,75)^2 \cdot 15 + (3-1,75)^2 \cdot 3)} = 0,63$$

$$r_{xy} = \frac{k}{s_x \cdot s_y} \cdot \frac{1}{n} \cdot \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) = \frac{1}{0,57 \cdot 0,63} =$$

$$\frac{6 \cdot (1-1,5) \cdot (1-1,75) + 8 \cdot (1-1,5) \cdot (2-1,75) + 1 \cdot (1-1,5) \cdot (3-1,75) + 4 \cdot (2-1,5) \cdot (1-1,75) + 7 \cdot (2-1,5) \cdot (2-1,75) + 1 \cdot (2-1,5) \cdot (3-1,75) + 1 \cdot (3-1,5) \cdot (3-1,75)}{28 \cdot 0,57 \cdot 0,63} = 0,249$$

Keďže platí $|r_{xy}| \leq 0,3$, tak lineárna závislosť medzi známkami z matematiky a slovenského jazyka je slabá.