

**APLIKOVANÁ STATISTIKA**

**Přehled základních vzorců**

**Obor: VSRR**

# I. Testy neparametrické

## 1. Testy dobré shody

### $\chi^2$ – test dobré shody

$$\chi^2 = \sum_{j=1}^k \frac{(n_j - np_j)^2}{np_j}$$

kde:  $n_j$  ..... empirické (skutečné) četnosti v j-té třídě

$np_j$  ... teoretické četnosti v j-té třídě ( $j = 1, 2, \dots, k$ )

$k$  ..... počet tříd (skupin)

Testovací kritérium má za předpokladu, že provádíme dostatečně velký výběr, přibližně  $\chi^2$ -rozdělení pro  $f = (k - c - 1)$  stupně volnosti ( $k$  je počet intervalů a  $c$  je počet parametrů, které nejsou hypotézou  $H_0$  specifikovány).

### Kolmogorov - Smirnovův test

$$D = \frac{1}{n} \max |N_j - H_j|$$

kde:  $N_j$  ..... kumulativní empirické četnosti

$H_j$  ..... kumulativní teoretické četnosti

$n$  ..... rozsah sledovaného souboru

Tabulka kritických hodnot  $D_\alpha$  je sestavena pouze pro  $n \leq 40$ . Pro výběry větších rozsahů se kritické hodnoty určí podle vztahů:

$$D_{0,05} = \frac{1,36}{\sqrt{n}}$$

$$D_{0,01} = \frac{1,63}{\sqrt{n}}$$

## 2. Klasické neparametrické testy (pořadové)

### Dvouvýběrový Wilcoxonův test

$$U = \min(U_x, U_y), \text{ kde } U_x = m \cdot n + \frac{m \cdot (m+1)}{2} - T_x$$

$$U_y = m \cdot n + \frac{n \cdot (n+1)}{2} - T_y$$

$U_{\alpha(m, n)}$  – kritická hodnota pro  $m$  a  $n$  – rozsahy souborů a  $\alpha$  – hladinu významnosti

### Wilcoxonův test

$$W = \min(W^-, W^+)$$

$W_{\alpha(n)}$  – kritická hodnota pro  $\alpha$  a  $n$ , kde  $n$  je počet nenulových diferencí

### Znaménkový test

$$Z = \min(Z^+, Z^-)$$

$Z_{\alpha(n)}$  – kritická hodnota, kde  $n$  je počet nenulových diferencí

### Kruskalův - Wallisův test

$$KW = \frac{12}{n(n+1)} \sum_{i=1}^k \frac{T_i^2}{n_i} - 3(n+1), \text{ kde } T_i \text{ je součet pořadových čísel pro } i\text{-tý výběr}$$

Testové kritérium se řídí  $\chi^2$ -rozdělením o  $(k - 1)$  stupních volnosti (kde  $k$  je počet úrovní třídícího znaku).

## 3. Ostatní neparametrické testy

### Dixonův test

$$Q_1 = \frac{x_2 - x_1}{x_n - x_1} \quad Q_n = \frac{x_n - x_{n-1}}{x_n - x_1}$$

$Q_{\alpha(n)}$  – tabulková hodnota pro Dixonův test, hladinu významnosti  $\alpha$  a  $n$  (kde  $n$  je rozsah souboru)

## II. Analýza závislosti kvantitativních znaků

### 1. Jednoduchá korelace a regrese

funkce lineární (přímka)

$$y' = a + b \cdot x_i$$

Soustava normálních rovnic

$$n a_{yx} + b_{yx} \cdot \sum_{i=1}^n x_i = \sum_{i=1}^n y_i$$

$$a_{yx} \sum_{i=1}^n x_i + b_{yx} \cdot \sum_{i=1}^n x_i^2 = \sum_{i=1}^n x_i y_i$$

Výpočtové vzorce

$b_{yx} = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{n \sum x_i^2 - (\sum x_i)^2}$	$b_{xy} = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{n \sum y_i^2 - (\sum y_i)^2}$
$b_{yx} = \frac{\text{cov } xy}{s_x^2}$	$b_{xy} = \frac{\text{cov } xy}{s_y^2}$
$b_{yx} = r \cdot \frac{s_y}{s_x}$	$b_{xy} = r \cdot \frac{s_x}{s_y}$
$a_{yx} = \bar{y} - b_{yx} \cdot \bar{x}$	$a_{xy} = \bar{x} - b_{xy} \cdot \bar{y}$
$\text{cov } xy = \bar{x}\bar{y} - \bar{x} \cdot \bar{y}$	

**Síla závislosti – korelační koeficient**

$$r_{xy} = r_{yx} = \frac{\text{cov } xy}{s_x \cdot s_y}$$

$$r_{yx} = r_{xy} = \pm \sqrt{b_{yx} \cdot b_{xy}}$$

$$r_{yx} = r_{xy} = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{\sqrt{[n \sum x_i^2 - (\sum x_i)^2] \cdot [n \sum y_i^2 - (\sum y_i)^2]}}$$

**Síla závislosti – Spearmanův koeficient pořadové korelace**

$$r_s = 1 - \frac{6 \cdot \sum d_i^2}{n \cdot (n^2 - 1)},$$

kde  $d_i$  = rozdíly pořadí  $R_x$  a  $R_y$

## 2. Vícenásobná lineární regrese a korelace

$$\text{Rovina: } y' = a + b_{y x_1 \cdot x_2} \cdot x_1 + b_{y x_2 \cdot x_1} \cdot x_2$$

Soustava normálních rovnic pro určení parametrů  $a$ ,  $b_1$ ,  $b_2$  funkce

$$y' = a + b_1 \cdot x_1 + b_2 \cdot x_2$$

$$n a + b_1 \sum x_1 + b_2 \sum x_2 = \sum y_i$$

$$a \sum x_1 + b_1 \sum x_1^2 + b_2 \sum x_1 x_2 = \sum y_i x_1$$

$$a \sum x_2 + b_1 \sum x_1 x_2 + b_2 \sum x_2^2 = \sum y_i x_2$$

$$b_{y x_1 \cdot x_2} = \frac{b_{y x_1} - b_{y x_2} \cdot b_{x_2 x_1}}{1 - b_{x_2 x_1} \cdot b_{x_1 x_2}}$$

$$b_{y x_1 \cdot x_2} = \frac{s_y}{s_{x_1}} \cdot \frac{r_{y x_1} - r_{x_1 x_2} \cdot r_{y x_2}}{1 - r_{x_1 x_2}^2}$$

$$b_{y x_2 \cdot x_1} = \frac{b_{y x_2} - b_{y x_1} \cdot b_{x_1 x_2}}{1 - b_{x_2 x_1} \cdot b_{x_1 x_2}}$$

$$b_{y x_2 \cdot x_1} = \frac{s_y}{s_{x_2}} \cdot \frac{r_{y x_2} - r_{x_1 x_2} \cdot r_{y x_1}}{1 - r_{x_1 x_2}^2}$$

$$a = \bar{y} - b_{y x_1 \cdot x_2} \cdot \bar{x}_1 - b_{y x_2 \cdot x_1} \cdot \bar{x}_2$$

**Síla závislosti**

a) úplný korelační koeficient

$$R_{y \cdot x_1 x_2} = \sqrt{\frac{r_{y x_1}^2 + r_{y x_2}^2 - 2 \cdot r_{x_1 x_2} \cdot r_{y x_1} \cdot r_{y x_2}}{1 - r_{x_1 x_2}^2}}$$

b) parciální korelační koeficienty

$$r_{y x_1 \cdot x_2} = \frac{r_{y x_1} - r_{y x_2} \cdot r_{x_1 x_2}}{\sqrt{(1 - r_{y x_2}^2)(1 - r_{x_1 x_2}^2)}}$$

$$r_{y x_2 \cdot x_1} = \frac{r_{y x_2} - r_{y x_1} \cdot r_{x_1 x_2}}{\sqrt{(1 - r_{y x_1}^2)(1 - r_{x_1 x_2}^2)}}$$

**$\beta$  – koeficienty**

$$\beta_{y x_1 x_2} = b_{y x_1 \cdot x_2} \cdot \frac{s_{x_1}}{s_y} = \frac{r_{y x_1} - r_{y x_2} \cdot r_{x_1 x_2}}{1 - r_{x_1 x_2}^2}$$

$$\beta_{y x_2 x_1} = b_{y x_2 \cdot x_1} \cdot \frac{s_{x_2}}{s_y} = \frac{r_{y x_2} - r_{y x_1} \cdot r_{x_1 x_2}}{1 - r_{x_1 x_2}^2}$$

### III. Vybrané ukazatele demografického vývoje

#### Index feminity ( $I^F$ )

$$I^F = \frac{P_F}{P_M}, \quad \text{kde } P_F \text{ je počet žen, } P_M \text{ je počet mužů}$$

#### Index maskulinity ( $I^M$ )

$$I^M = \frac{P_M}{P_F}$$

#### Index stáří (IS)

$$IS = \frac{\text{III. biologická generace}}{\text{I. biologická generace}} \quad \text{dle Sauvyho}$$

$$IS = \frac{\text{III. generace 65+}}{\text{I. generace do 14 let}} \quad \text{častěji (uvádí ČSÚ)}$$

#### Index hospodářského zatížení (IHZ)

$$IHZ = \frac{\text{I.e.g.} + \text{II.e.g.} + \text{III.e.g.}}{\text{II.e.g.}}$$

#### Vážený index hospodářského zatížení

$$IHZ_{0,7;1;0,7} = \frac{0,7 \cdot \text{I.e.g.} + \text{II.e.g.} + 0,7 \cdot \text{III.e.g.}}{\text{II.e.g.}}$$

#### Index závislosti mladých (IZM)

$$IZM = \frac{\text{I.e.g.}}{\text{II.e.g.}}$$

#### Index závislosti starých (IZS)

$$IZS = \frac{\text{III.e.g.}}{\text{II.e.g.}}$$