$\chi 2$ – test o nezávislosti v kontingenční tabulce

- ***kontingenční tabulka:*** dvourozměrná tabulka, kde alespoň jedna proměnná je slovní
- podstatou testu je porovnání empirických četností s teoretickými četnostmi
- teoretické četnosti = četnosti očekávané v případě nezávislosti (n'_{ij}).

Vzorec pro výpočet teoretických četností:

$$n'_{ij} = \frac{n_{i\bullet} \cdot n_{\bullet j}}{n}$$

Předpoklady testu:

- všechna políčka kontingenční tabulky musí být dostatečně obsazena ($n'_{ij} \geq 5$)
- pokud podmínka není splněna, musíme některé třídy sloučit nebo zvětšit rozsah výběru.

Testovací postup:

1) H_0 *:proměnné x a y jsou nezávislé*

H_1 *: non* H_0

2) ***Testové kritérium:***

$$G = \sum_{i=1}^{r} \sum_{j=1}^{s} \frac{\left(n_{ij} - n'_{ij}\right)^2}{n'_{ij}};$$

Statistika *G* má při platnosti *H₀* rozdělení $\chi^2\left[(r-1)(s-1)\right]$

3) ***Kritický obor:***

$$W \equiv \left\{G;\ G > \chi^2_{1-\alpha}\left[(r-1)(s-1)\right]\right\}$$

4) Závěr testu:

Pokud leží hodnota testového kritéria v kritickém oboru, zamítáme H_0 a přijímáme H_1, tedy prokázali jsme hypotézu H_1 o závislosti proměnných *x* a *y*.

Měření síly (intenzity, těsnosti) závislosti v kontingenční tabulce:

Cramérův koeficient kontingence: $C_{Cr} = \sqrt{\dfrac{G}{n \cdot h}}$; $C_{Cr} \in \langle 0,1 \rangle$

h = min (r-1; s-1)

Pearsonův koeficient kontingence: $C_P = \sqrt{\dfrac{G}{G+n}}$; $C_P \in \langle 0,1)$

$\chi 2$ – test o nezávislosti v asociační tabulce

- ***asociační tabulka:*** dvourozměrná čtyřpolní tabulka
- v podstatě jde o speciální případ kontingenční tabulky
- znaky A a B jsou alternativní
- při zkoumání asociace sledujeme, jak často jevy A a B nastaly či nenastaly současně, a jak často nastal pouze jeden z nich.

Testovací postup:

1) H_0 *:znaky (jevy) A a B jsou nezávislé*

H_1 *: non* H_0

2) *Testové kritérium:*

$$G = n\frac{\left(n_{11}n_{22} - n_{12}n_{21}\right)^2}{n_{\bullet 1}n_{\bullet 2}n_{1\bullet}n_{2\bullet}};$$

Statistika *G* má při platnosti H_0 rozdělení $\chi^2(1)$

3) *Kritický obor:*

$$W \equiv \left\{G;\ G > \chi^2_{1-\alpha}(1)\right\}$$

4) Závěr testu:

Pokud leží hodnota testového kritéria v kritickém oboru, zamítáme H_0 a přijímáme H_1, tedy prokázali jsme hypotézu H_1 o závislosti (asociaci) proměnných A a B.

Měření síly (intenzity, těsnosti) závislosti v asociační tabulce:

Koeficient asociace: $r_{AB} = \dfrac{n_{11}n_{22} - n_{12}n_{21}}{\sqrt{n_{1\bullet}n_{2\bullet}n_{\bullet 1}n_{\bullet 2}}}$; $\qquad r_{AB} \in \langle -1;1 \rangle$

Interpretace koeficientu asociace:

1. znaménko (+/–) udává směr asociace:

$r_{AB} > 0 \Rightarrow$ kladná (přímá) asociace (jevy častěji nastávají či nenastávají společně a méně často nastává jen jeden z nich).

$r_{AB} < 0 \Rightarrow$ záporná (nepřímá) asociace (jevy méně často nastávají či nenastávají společně a častěji nastává jen jeden z nich).

2. $|r_{AB}|$ udává sílu asociace:

$r_{AB} = 0 \qquad \Rightarrow$ úplná nepřímá asociace (kterýkoli z jevů pouze když nenastává jev druhý)

$|r_{AB}| = 1 \quad \Rightarrow$ úplná kladná asociace (jevy nastávají pouze společně)

$|r_{xy}| \to 0 \quad \Rightarrow$ slabá asociace

$|r_{xy}| \to 1 \quad \Rightarrow$ silná asociace