χ^2-coefficient

Measure association between two nominal variables

$$\chi^2(X,Y) = \sum^{r} \sum^{c} \frac{\left(n_{ij} - \frac{n_{i\cdot}\, n_{\cdot j}}{n}\right)^2}{\frac{n_{i\cdot}\, n_{\cdot j}}{n}} =$$

$$= \sum^{r} \sum^{c} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

Properties:

- $\chi^2(X,Y) \approx 0$ corresponds to the independence (no association)

χ^2-coef.: Chuproff correlation

Properties

- $0 \leq T(X,Y) \leq 1$

Correlation ratio

numerical r.v. and nominal one

numeric

$$\bar{y}_{i} = \frac{1}{n_{i\bullet}} \sum_{j=1}^{I} y_j^* n_{ij} \qquad i=1,\dots,k$$

$$\hat{\eta}(Y|X) = \frac{\sqrt{\frac{1}{n}\sum^{k} (\bar{y}_i - \bar{y})^2 n_{i\bullet}}}{S_Y}$$

Kendall's tau

between columns of ranked data

$$\text{Kendall's } \tau = \frac{C - D}{C + D}$$

C - # concordant pairs

D - # disconcordant pairs