Test niezależności chi-kwadrat (χ^2)

H_0 : zmienne są niezależne
H_1 : zmienne są zależne

$$\chi_e^2 = \sum_{i=1}^{r}\sum_{j=1}^{k}\frac{(n_{ij} - \hat{n}_{ij})^2}{\hat{n}_{ij}} \qquad \text{gdzie:} \quad \hat{n}_{ij} = \frac{n_{i.} \cdot n_{.j}}{n}$$

r – liczba wierszy w tablicy wielodzielczej
k – liczba kolumn w tablicy wielodzielczej
n – liczebność badanej próby
n_{ij} – liczba jednostek na przecięciu i-tego wiersza i j-tej kolumny w tablicy korelacyjnej (liczebność empiryczna)
$\hat{n}_{ij}$ – liczebność teoretyczna

$n_{i.} = \sum_j n_{ij}$ – liczebność brzegowa liczona dla i-tego wiersza po wszystkich kolumnach

$n_{.j} = \sum_i n_{ij}$ – liczebność brzegowa liczona dla j-tej kolumny po wszystkich wierszach

Jeżeli:

- $\chi_e^2 < \chi^2_{\alpha;\ df=(r-1)(k-1)}$ to na poziomie istotności α nie ma podstaw do odrzucenia hipotezy zerowej
- $\chi_e^2 \geq \chi^2_{\alpha;\ df=(r-1)(k-1)}$ to na poziomie istotności α hipotezę zerową należy odrzucić na rzecz hipotezy alternatywnej

---

Współczynnik kontyngencji

$$C_{xy} = \sqrt{\frac{\chi^2}{\chi^2 + n}}$$

$$C_{\max} = \frac{\sqrt{\frac{r-1}{r}} + \sqrt{\frac{k-1}{k}}}{2}$$

r – liczba wierszy w tabeli wielodzielczej
k – liczba kolumn w tabeli wielodzielczej

$C_{skor} = C_{xy} / C_{max}$