Függetlenségvizsgálat

Legyen (ξ_i, η_i) n elemű véletlen minta a (ξ, η) valószínűségi vektorváltozóra. Nullhipotézis H_0 : a ξ és η valószínűségi változók függetlenek. Soroljuk a mintát rs csoportba a következőképpen. Az osztályhatárok legyenek $a_1 < a_2 < \cdots < a_{r-1}$ illetve $b_1 < b_2 < \cdots < b_{s-1}$. Az $A_1 = \{\xi < a_1\}$, $A_i = \{a_{i-1} \le \xi < a_i\}$ $(i = 2, \ldots, r-1)$, $A_r = \{\xi \ge a_{r-1}\}$ illetve a $B_1 = \{\eta < b_1\}$, $B_i = \{b_{i-1} \le \eta < b_i\}$ $(i = 2, \ldots, s-1)$, $B_s = \{\eta \ge b_{s-1}\}$ események teljes eseményrendszert alkotnak. A nullhipotézisünk $H_0 : P(A_i \cap B_j) = P(A_i)P(B_j)$ minden i, j-re alakot ölt. Legyen k_{ij} a $(\xi, \eta) \in A_i \times B_j$ esemény gyakorisága. Az adatokat (k_{ij}) matrixba rendezve, az utolsó sor és oszlop után az adott sor f_{i*} illetve az adott oszlop f_{*j} összegét feltüntetve kapjuk a **kontingenciatáblázat**ot.

A tiszta függetlenségvizsgálat esetén a ξ, η valószínűségi változók elméleti eloszlása ismert, azaz ismertek a $P(A_i) = p_i$ és a $P(B_j) = q_j$ valószínűségek. Ekkor a nullhipotézisünk $H_0 : P(A_i \cap B_j) = p_i q_j$ lesz. A

$$K^2 = \sum_{i=1}^{r} \sum_{j=1}^{s} \frac{(k_{ij} - np_i q_j)^2}{np_i q_j}$$

próbastatisztika $n \to \infty$ esetén az $(rs-1)$-szabadságfokú χ^2-eloszláshoz tart, alkalmazhatjuk a χ^2-próbát a szokott módon.

A becsléses függetlenségvizsgálat esetén a ξ, η valószínűségi változók elméleti eloszlásai nem ismertek, azokat az f_{i*}/n illetve f_{*j}/n relatív gyakoriságokkal becsüljük. A

$$K^2 = \sum_{i=1}^{r} \sum_{j=1}^{s} \frac{(k_{ij} - f_{i*} f_{*j}/n)^2}{f_{i*} f_{*j}/n}$$

próbastatisztika $n \to \infty$ esetén az $rs - 1 - (r + s - 2) = (r-1)(s-1)$ szabadságfokú χ^2-eloszláshoz tart (mivel $p_1 + \cdots + p_r = 1$, ezért $r - 1$ a becsült p_i paraméterek száma, és a q_j-kre pedig $s - 1$). A χ^2-próba további menete a megszokott.