STATYSTYKA
dr inż Krzysztof Bryś

Wykład 7

Dwuwymiarowa zmienna losowa typu skokowego. Współczynnik korelacji. Prosta regresji.

Niech X, Y będą jednowymiarowymi zmiennymi losowymi. Parę (X,Y) nazywamy **dwuwymiarową zmienną losową** a X i Y jej współrzędnymi.
Jeżeli X i Y są typu skokowego, to (X,Y) jest **dwuwymiarową zmienną losową typu skokowego**.

Rozkład prawdopodobieństwa dwuwymiarowej zmiennej losowej typu skokowego: funkcja $P : W_X \times W_Y \rightarrow < 0, 1 >$ taka, że :

$$1) \text{ dla każdego } x_i \in W_X, y_j \in W_Y, \quad P(X = x_i, Y = y_j) = p_{ij} > 0,$$

$$2) \sum_i \sum_j p_{ij} = 1.$$

Rozkład brzegowy zmiennej losowej X:

$$p_{i.} = P(X = x_i) = \sum_j p_{ij} \quad \text{dla każdego } x_i \in W_X$$

Rozkład brzegowy zmiennej losowej Y:

$$p_{.j} = P(Y = y_j) = \sum_i p_{ij} \quad \text{dla każdego } y_j \in W_Y$$

Zmienne losowe X i Y są **niezależne** $\Leftrightarrow$ dla każdego $A \subseteq R, B \subseteq R$, zdarzenia $X \in A$ oraz $Y \in B$ są niezależne.
Zmienne losowe X i Y typu skokowego są **niezależne** $\Leftrightarrow$ dla każdego $x_i \in W_X, y_j \in W_Y$, $p_{ij} = p_{i.} \cdot p_{.j}$

Kowariancja zmiennych losowych X i Y typu skokowego:

$$cov(X,Y) = E(X \cdot Y) - E(X) \cdot E(Y)$$

gdzie

$$E(X \cdot Y) = \sum_i \sum_j x_i \cdot y_j \cdot p_{ij}$$

Kowariancja określa siłę i kierunek zależności liniowej (korelacji) mię dzy zmiennymi losowymi X i Y.

Współczynnik korelacji zmiennych losowych X i Y:

$$\rho_{(X,Y)} = \frac{cov(X,Y)}{D(X) \cdot D(Y)} = \frac{E(X \cdot Y) - E(X) \cdot E(Y)}{D(X) \cdot D(Y)}$$

Współczynnik korelacji określa siłę i kierunek zależności liniowej (korelacji) mię dzy zmiennymi losowymi X i Y. Przyjmuje wartości z przedziału $< -1; 1 >$.

Zmienne losowe X i Y są **nieskorelowane** $\Leftrightarrow \rho_{(X,Y)} = 0$.

Prosta regresji drugiego rodzaju zmiennej losowej Y względem zmiennej losowej X: prosta o równaniu $y = a \cdot x + b$, której współczynniki są tak dobrane, że średnie odchylenie kwadratowe zmiennej losowej Y od zmiennej $aX + b$ jest minimalne.
Można wykazać, że:

$$a = \frac{cov(X,Y)}{D^2(X)} = \rho_{(X,Y)} \cdot \frac{D(Y)}{D(X)}$$

$$b = E(Y) - aE(X)$$