

Statystyki opisowe

Średnia

$$\bar{x} = \frac{1}{n} \sum_{i=1}^k x_i n_i \quad \bar{x} = \frac{1}{n} \sum_{i=1}^k \dot{x}_i n_i$$

Wariancja

$$s^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \quad s^2 = \frac{1}{n} \sum_{i=1}^n (\dot{x}_i - \bar{x})^2 \cdot n_i$$

Dominanta (modalna)

$$D = x_D + \frac{n_D - n_{D-1}}{(n_D - n_{D-1}) + (n_D - n_{D+1})} \cdot h$$

Mediana i kwartyle

$$\text{Pozycja } Q_1 = \frac{n}{4} \quad \text{Wartość } Q_1 = x_{Q_1} + \frac{\frac{n}{4} - n_{skQ_1-1}}{n_{Q_1}} \cdot h$$

$$\text{Pozycja } M_e = \frac{n}{2} \quad \text{Wartość } M_e = x_{M_e} + \frac{\frac{n}{2} - n_{skM_e-1}}{n_{M_e}} \cdot h$$

$$\text{Pozycja } Q_3 = \frac{3n}{4} \quad \text{Wartość } Q_3 = x_{Q_3} + \frac{\frac{3n}{4} - n_{skQ_3-1}}{n_{Q_3}} \cdot h$$

Odchylenie przeciętne d (dewiata)

$$d = \frac{1}{n} \sum_{i=1}^n |x_i - \bar{x}| \quad d = \frac{1}{n} \sum_{i=1}^n |\dot{x}_i - \bar{x}| \cdot n_i$$

Współczynnik zmienności

$$V_s = \frac{s}{\bar{x}} \cdot 100 \quad \text{lub} \quad V_d = \frac{d}{\bar{x}} \cdot 100$$

Asymetria

$$A_s = \frac{\bar{x} - D}{s} \quad \text{lub} \quad A_d = \frac{\bar{x} - D}{d}$$

Typowy przedział zmienności

$$\bar{x} - s < x_{typ} < \bar{x} + s$$

Estymacja m dla $X \sim N(m; \sigma)$ i nieznanym σ

Jeśli $n \leq 30$

$$\bar{X} - t_{\alpha, n-1} \frac{s}{\sqrt{n-1}} < m < \bar{X} + t_{\alpha, n-1} \frac{s}{\sqrt{n-1}}$$

gdzie $t_{\alpha, n-1}$ z rozkładu T-studenta $P(|T_{n-1}| > t_{\alpha, n-1}) = \alpha$

Jeśli $n > 30$

$$\bar{X} - t_{\alpha} \frac{\sigma}{\sqrt{n}} < m < \bar{X} + t_{\alpha} \frac{\sigma}{\sqrt{n}}$$

gdzie t_{α} z rozkładu Normalnego $\Phi(t_{\alpha}) = \frac{1-\alpha}{2}$

Korelacja i regresja

Współczynnik korelacji Pearsona

$$r_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \cdot \sum_{i=1}^n (y_i - \bar{y})^2}} = \frac{C(X, Y)}{s_x s_y}$$

Prosta regresji $\hat{y} = ax + b$

$$a = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad b = \bar{y} - a \cdot \bar{x}$$

Współczynnik determinacji i zbieżności

$$R^2 = 1 - \varphi^2 \quad \varphi^2 = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

Reszty

$$e_i = y_i - \hat{y}$$

Współczynnik korelacji rang Spearmana

$$r_s = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n(n^2 - 1)}$$

Rozkłady

Dwumianowy (Bernouliego)

$$B(n, p, k) = P(X = k) = C_n^k p^k q^{n-k} = \frac{n!}{k!(n-k)!} p^k q^{n-k}$$

$$E(X) = np$$

$$V(X) = npq$$

Poissona

$$P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}$$

$$E(X) = \lambda$$

$$V(X) = \lambda$$

$$\lambda = np$$

$$e \approx 2,7182$$

Normalny (Gausa)

$$Z = \frac{X - m}{\sigma}$$

$$P(X < a) = P\left(Z < \frac{a - m}{\sigma}\right)$$

$$P(X > b) = P\left(Z > \frac{b - m}{\sigma}\right)$$

$$P(a < X < b) = P\left(\frac{a - m}{\sigma} < Z < \frac{b - m}{\sigma}\right)$$