

FORMULÁRIO

Estatística Descritiva

Regra de Struges: n° de classes $\cong 1 + \log_{10} n / \log_{10} 2$

Média aritmética: $\bar{x} = \frac{1}{n} \sum_{i=1}^p n_i x_i = \sum_{i=1}^p f_i x_i$

Moda: (dados agrupados) Fórmula de King: $Mo = a_j + \frac{n_{j+1}}{n_{j-1} + n_{j+1}} (a_{j+1} - a_j)$

Fórmula de Czuber: $Mo = a_j + \frac{n_j - n_{j-1}}{(n_j - n_{j-1}) + (n_j - n_{j+1})} (a_{j+1} - a_j)$

Variância: $\sigma^2 = \frac{1}{n} \sum_{i=1}^p n_i (x_i - \bar{x})^2 = \sum_{i=1}^p f_i (x_i - \bar{x})^2 = \frac{1}{n} \left(\sum_{i=1}^p n_i x_i^2 - n \bar{x}^2 \right) = \sum_{i=1}^p f_i x_i^2 - \bar{x}^2$

ou, $s^2 = \frac{1}{n-1} \sum_{i=1}^p n_i (x_i - \bar{x})^2 = \frac{1}{n-1} \left(\sum_{i=1}^p n_i x_i^2 - n \bar{x}^2 \right)$

Coefficiente de dispersão: $cd = \frac{s}{\bar{x}}$ ou $\frac{\sigma}{\bar{x}}$

Coefficiente de variação: $cv = cd \times 100\%$

Momento simples de ordem r ou momento ordinário de ordem r : $m'_k = \sum_{i=1}^p f_i x_i^k = \frac{1}{n} \sum_{i=1}^p n_i x_i^k$

Momento centrado de ordem r : $m_k = \sum_{i=1}^p f_i (x_i - \bar{x})^k = \frac{1}{n} \sum_{i=1}^p n_i (x_i - \bar{x})^k$

Coefficiente de assimetria: $g_1 = \frac{m_3}{\sqrt{m_2^3}}$

Coefficiente de achatamento ou curtose: $g_2 = \frac{m_4}{m_2^2}$

Probabilidades

Probabilidade condicional ou condicionada de A dado B: $P(A|B) = \frac{P(A \cap B)}{P(B)}$, se $P(B) > 0$

Variáveis Aleatórias

Média, valor esperado ou esperança matemática:

– $E(X) \equiv \mu_X = \sum_i x_i f_X(x_i)$, e $E(\phi(X)) \equiv \mu_{\phi(X)} = \sum_i \phi(x_i) f_X(x_i)$, se X é discreta com função de probabilidade f_X e tomando valores em $\{x_1, x_2, \dots\}$;

– $E(X) \equiv \mu_X = \int_{-\infty}^{+\infty} x f_X(x) dx$, e $E(\phi(X)) \equiv \mu_{\phi(X)} = \int_{-\infty}^{+\infty} \phi(x) f_X(x) dx$, se X é (absolutamente) contínua com função densidade de probabilidade f_X .

Variância:

- $\text{Var}(X) \equiv \sigma_x^2 = \sum_i (x_i - \mu_x)^2 f_x(x_i) = E((X - \mu_x)^2)$, se X é discreta com função de probabilidade f_x e tomando valores em $\{x_1, x_2, \dots\}$;
- $\text{Var}(X) \equiv \sigma_x^2 = \int_{-\infty}^{+\infty} (x - \mu_x)^2 f_x(x) dx = E((X - \mu_x)^2)$, se X é (absolutamente) contínua com função densidade de probabilidade f_x .

Tem-se: $\text{Var}(X) = E(X^2) - [E(X)]^2$

Desvio padrão: $\sigma_x = \sqrt{\text{Var}(X)}$

Distribuições de Probabilidade

Se X tem distribuição de Bernoulli de parâmetro p :

$$f_x(x) = \begin{cases} p^x (1-p)^{1-x} & \text{se } x = 0 \vee x = 1 \\ 0 & \text{caso contrário} \end{cases}, \quad E(X)=p \quad \text{e} \quad \text{Var}(X)=p(1-p)$$

Se X tem distribuição de binomial de parâmetros n e p :

$$f_x(x) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x} & \text{se } x = 0, 1, 2, \dots \\ 0 & \text{caso contrário} \end{cases}, \quad E(X)=np \quad \text{e} \quad \text{Var}(X)=np(1-p)$$

Se X tem distribuição hipergeométrica de parâmetros M , k e n :

$$f_x(x) = \begin{cases} \frac{\binom{k}{x} \binom{M-k}{n-x}}{\binom{M}{n}} & \text{se } x = \max\{0, n - (M - k)\}, \dots, \min\{n, k\} \\ 0 & \text{caso contrário} \end{cases},$$

$$E(X) = n \times \frac{k}{M} \quad \text{e} \quad \text{Var}(X) = n \times \frac{k}{M} \times \frac{M-k}{M} \times \frac{M-n}{M-1}$$

Se X tem distribuição de Poisson de parâmetro μ :

$$f_x(x) = \begin{cases} \frac{e^{-\mu} \mu^x}{x!} & \text{se } x = 0, 1, 2, \dots \\ 0 & \text{caso contrário} \end{cases}, \quad E(X)=\mu \quad \text{e} \quad \text{Var}(X)=\mu$$

Regressão e correlação linear simples

Estimativas dos coeficientes do modelo: $b_1 = \frac{\sum_{i=1}^n x_i y_i - n \bar{x} \bar{y}}{\sum_{i=1}^n x_i^2 - n \bar{x}^2}$; $b_0 = \bar{y} - b_1 \bar{x}$

Coeficiente de determinação: $r^2 = \frac{SSR}{SST} = \frac{SST - SSE}{SST} = 1 - \frac{SSE}{SST} = \frac{b_0 \sum_{i=1}^n y_i + b_1 \sum_{i=1}^n y_i x_i - n \bar{y}^2}{\sum_{i=1}^n y_i^2 - n \bar{y}^2}$