

FORMULÁRIO MATEMÁTICA II

Indicadores para dados univariados

amostra observada x_1, x_2, \dots, x_n

amostra ordenada $x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}$

$$\text{mediana } \tilde{x} = \begin{cases} x_{(\frac{n+1}{2})} & n \text{ ímpar} \\ \frac{x_{(n/2)} + x_{(n/2+1)}}{2} & n \text{ par} \end{cases}$$

Quantil de ordem θ ($0 < \theta < 1$)

$$Q_\theta = \begin{cases} \frac{1}{2} (x_{(n\theta)} + x_{(n\theta+1)}) & \text{se } n\theta \text{ inteiro} \\ x_{(\lceil n\theta \rceil + 1)} & \text{se } n\theta \text{ não inteiro} \end{cases}$$

$\lceil n\theta \rceil$ a parte inteira de $n\theta$

barreira inferior $BI = Q_1 - 1.5(Q_3 - Q_1)$

barreira superior $BS = Q_3 + 1.5(Q_3 - Q_1)$

$$s_x^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1} = \frac{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2}{n(n-1)}$$

se $x_i^* = a + bx_i$ então $\bar{x}^* = a + b\bar{x}$, $s_{x^*}^2 = b^2 s_x^2$

$$CV = \frac{s}{\bar{x}} \times 100\%$$

regra de Sturges (divisão em classes)

$$\text{número de classes próximo de } \left[1 + \frac{\ln n}{\ln 2} \right]$$

Indicadores para dados agrupados em c classes

n_i , x'_i frequência absoluta e ponto médio da classe i

$$\bar{x}' = \frac{1}{n} \sum_{i=1}^c n_i x'_i = \frac{\sum_{i=1}^c f_i x'_i}{\sum_{i=1}^c f_i}$$

$$s_x'^2 = \frac{\sum_{i=1}^c (x'_i - \bar{x}')^2 n_i}{n} = \frac{\sum_{i=1}^c x_i'^2 n_i}{n} - \bar{x}'^2$$

Seja k a primeira classe tal que $F_k \geq \theta$

$$Q'_\theta = x_k^{\min} + (x_k^{\max} - x_k^{\min}) \frac{\theta - F_{k-1}}{f_k}$$

Seja k a classe com maior frequência

$$\text{moda}' = x_k^{\min} + h \frac{f_{k+1}}{f_{k-1} + f_{k+1}}$$

Indicadores para dados bivariados

$(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$

$$\text{cov}(x, y) = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{n-1}$$

$$= \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n(n-1)}$$

$$|\text{cov}(x, y)| \leq s_x s_y$$

$$r = r_{xy} = \frac{\text{cov}(x, y)}{s_x s_y} \text{ se } s_x \neq 0, s_y \neq 0$$

se $x_i^* = a + bx_i$ e $y_i^* = c + dy_i$

$$\text{cov}(x^*, y^*) = bd \text{cov}(x, y)$$

$$r_{x^* y^*} = \begin{cases} r_{xy} & \text{se } bd > 0 \\ -r_{xy} & \text{se } bd < 0 \end{cases}$$

Regressão linear simples

$$\hat{y}_i = b_0 + b_1 x_i$$

$$y_i = b_0 + b_1 x_i + e_i$$

$$\begin{cases} b_1 = \frac{\text{cov}(x, y)}{s_x^2} = r \frac{s_y}{s_x}, & s_x \neq 0 \\ b_0 = \bar{y} - b_1 \bar{x} \end{cases}$$

Probabilidade de acontecimentos

$$P(A - B) = P(A \cap \bar{B}) = P(A) - P(A \cap B)$$

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$$

$$P(\bar{A}|B) = 1 - P(A|B), \text{ se } P(B) \neq 0$$

Parâmetros (de funções) de uma v.a. X

Quantil de probabilidade p , χ_p , de uma v.a. X é o menor valor x tal que $F(x) \geq p$ (se X contínua $F(\chi_p) = p$)

$$E[\varphi(X)] = \begin{cases} \sum_i \varphi(x_i) p_i, & \text{se } X \text{ discreta} \\ \int_{-\infty}^{+\infty} \varphi(x) f(x) dx, & \text{se } X \text{ contínua} \end{cases}$$

$$\sigma^2 = \text{Var}(X) = E[(X - \mu)^2] = E[X^2] - \mu^2 \quad (\mu = E[X])$$

$$\text{Var}(a + bX) = b^2 \text{Var}(X)$$

Pares aleatórios (X, Y) discretos com f. massa de probabilidade conjunta $P[X = x_i, Y = y_j] = p_{ij}$

$$p_{i \cdot} = \sum_j p_{ij} \quad p_{\cdot j} = \sum_i p_{ij}$$

$$P[X = x_i | Y = y_j] = \frac{p_{ij}}{p_{\cdot j}}$$

Parâmetros de funções de um par aleatório (X, Y)

$$\sigma_{X, Y} = \text{Cov}(X, Y) = E[(X - \mu_X)(Y - \mu_Y)]$$

$$= E[XY] - E[X]E[Y] \quad (\mu_X = E[X], \mu_Y = E[Y])$$

$$\text{Var}(X \pm Y) = \text{Var}(X) + \text{Var}(Y) \pm 2\text{Cov}(X, Y)$$

$$\text{Cov}(a + bX, c + dY) = bd \text{Cov}(X, Y)$$

$$\rho = \rho_{X, Y} = \frac{\sigma_{X, Y}}{\sigma_X \sigma_Y}, \quad \sigma_X \neq 0, \sigma_Y \neq 0$$

$$\rho_{a+bX, c+dY} = \rho_{X, Y} \text{ se } bd > 0$$

Distribuição binomial

$X \cap B(n, p) \Leftrightarrow n - X \cap B(n, q)$, com $q = 1 - p$

$$P[X = x] = \binom{n}{x} p^x q^{n-x}, \quad x = 0, 1, \dots, n$$

$$E[X] = np \quad \text{Var}[X] = npq$$

Distribuição Poisson

$X \cap P(\lambda)$

$$P[X = x] = \frac{e^{-\lambda} \lambda^x}{x!}, \quad x = 0, 1, 2, \dots$$

$$E[X] = \lambda \quad \text{Var}[X] = \lambda$$

Distribuição uniforme contínua

$$X \cap U(a, b)$$

$$f(x) = \begin{cases} \frac{1}{b-a} & a < x < b \\ 0 & \text{o.v.} \end{cases}$$

$$E[X] = \frac{a+b}{2}, \quad Var[X] = \frac{(b-a)^2}{12}$$

Distribuição normal

$$X \cap N(\mu, \sigma)$$

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]$$

$$E[X] = \mu \quad Var[X] = \sigma^2$$

Teorema Limite Central

Sejam X_1, \dots, X_n v.a.'s independentes e identicamente distribuídas com valor médio μ e variância σ^2 (finita),

$S_n = \sum_{i=1}^n X_i$ e $\bar{X}_n = \frac{S_n}{n}$, então

$$\frac{S_n - n\mu}{\sigma\sqrt{n}} \sim \mathcal{N}(0, 1) \quad \text{e} \quad \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0, 1)$$

Aproximações das distribuições

$X \cap B(n, p)$, $np > 5$ e $nq > 5 \Rightarrow X \sim \mathcal{N}(\mu, \sigma)$, com

$$\mu = np \text{ e } \sigma = \sqrt{npq}$$

$X \cap P(\lambda)$ e $\lambda > 12 \Rightarrow X \sim \mathcal{N}(\mu, \sigma)$, com $\mu = \lambda$ e $\sigma = \sqrt{\lambda}$

Estimação

T estimador de um parâmetro θ

centrado se $E[T] = \theta$

$$EQM = E[(T - \theta)^2]$$

Fórmula geral de interpolação

$$z(x_0) = \frac{\sum_{i=1}^m w_i z(x_i)}{\sum_{i=1}^m w_i}$$

Variograma

$$\gamma[Z(x_i), Z(x_i + \mathbf{h})] = \frac{1}{2} Var[Z(x_i) - Z(x_i + \mathbf{h})] = \gamma(\mathbf{h})$$

$$\gamma^*(\mathbf{h}) = \frac{\sum_{i=1}^{n(\mathbf{h})} [z(x_i) - z(x_i + \mathbf{h})]^2}{2n(\mathbf{h})}$$

$$\gamma(\mathbf{h}) = C(0) - C(\mathbf{h})$$

Modelo esférico

$$\gamma(h) = \begin{cases} C \left[1.5 \frac{h}{\phi} - 0.5 \left(\frac{h}{\phi} \right)^3 \right], & 0 \leq h \leq \phi \\ C, & h > \phi \end{cases}$$

Modelo exponencial

$$\gamma(h) = C (1 - e^{-h/\phi}), \quad h \geq 0$$

modelo Gaussiano

$$\gamma(h) = C \left(1 - e^{-(h/\phi)^2} \right), \quad h \geq 0$$

Krigagem normal

$$Z^*(x_0) = \sum_{i=1}^n \lambda_i Z(x_i)$$

$$Var[Z^*(x_0) - Z(x_0)] = \sum_{i=1}^n \lambda_i \gamma(x_i, x_0) + \alpha$$

com $(\lambda_1, \dots, \lambda_n, \alpha)$ solução do sistema linear

$$\begin{bmatrix} \gamma_{11} & \dots & \gamma_{1n} & 1 \\ \dots & \dots & \dots & \dots \\ \gamma_{n1} & \dots & \gamma_{nn} & 1 \\ 1 & \dots & 1 & 0 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \dots \\ \lambda_n \\ \alpha \end{bmatrix} = \begin{bmatrix} \gamma_{10} \\ \dots \\ \gamma_{n0} \\ 1 \end{bmatrix}$$

Expressões úteis

Combinações de n elementos k a k , $n, k \in \mathbb{N}_0$, $k \leq n$

$$C_k^n = \binom{n}{k} = \frac{n!}{k!(n-k)!}$$

Algumas (regras de) primitivas

$$P(xe^{-x}) = -e^{-x}(x+1) + C$$

$$P(fg) = Fg - P(Fg'), \text{ com } F = Pf$$

$$P(f'f^\alpha) = \frac{f^{\alpha+1}}{\alpha+1} + C, \text{ com } \alpha \in \mathbb{R} \setminus \{-1\}$$

$$P(f'e^f) = e^f + C$$

$$P\frac{f'}{f} = \log|f| + C$$