

FORMULÁRIO DE ESTATÍSTICA II

• VALOR ESPERADO, MOMENTOS E PARÂMETROS

- $\text{Var}(X) = E(X - \mu)^2 = E(X^2) - \mu^2$
- $\text{Cov}(X, Y) = E\{(X - \mu_x)(Y - \mu_y)\} = E(XY) - E(X)E(Y)$; $\rho_{x,y} = \frac{\text{Cov}(X, Y)}{\sigma_x \sigma_y}$
- $E(aX + bY) = aE(X) + bE(Y)$; $\text{Var}(aX + bY) = a^2 \text{Var}(X) + b^2 \text{Var}(Y) + 2ab \text{Cov}(X, Y)$ com a, b constantes

• DISTRIBUIÇÕES TEÓRICAS

• UNIFORME (DISCRETA)

- Caso $x = 1, 2, \dots, n$: $f(x) = \frac{1}{n}$; $E(X) = \frac{n+1}{2}$; $\text{Var}(X) = \frac{n^2 - 1}{12}$
- Caso $x = 0, 1, 2, \dots, m$: $f(x) = \frac{1}{m+1}$; $E(X) = \frac{m}{2}$; $\text{Var}(X) = \frac{m(m+2)}{12}$

• BERNOULLI $X \sim B(1, \theta)$

$$f(x|\theta) = \theta^x (1-\theta)^{1-x}, \quad x = 0, 1 \quad (0 < \theta < 1); \quad E(X) = \theta; \quad \text{Var}(X) = \theta(1-\theta)$$

• BINOMIAL $X \sim B(n, \theta)$

$$f(x|\theta) = \binom{n}{x} \theta^x (1-\theta)^{n-x}, \quad x = 0, 1, 2, \dots, n, \quad (0 < \theta < 1);$$

$$E(X) = n\theta; \quad \text{Var}(X) = n\theta(1-\theta); \quad \mathfrak{Z}(\theta) = \frac{n}{\theta(1-\theta)}, \quad n \text{ conhecido}$$

Propriedades:

$$- X \sim B(n, \theta) \Leftrightarrow (n - X) \sim B(n, 1 - \theta)$$

$$- X_i \sim B(n_i, \theta), \text{ independentes } (i = 1, 2, \dots, k) \Rightarrow \sum_{i=1}^k X_i \sim B(n, \theta), \quad n = \sum_{i=1}^k n_i$$

• POISSON $X \sim \text{Po}(\lambda)$

$$f(x|\lambda) = \frac{e^{-\lambda} \lambda^x}{x!}, \quad x = 0, 1, 2, \dots, \quad (\lambda > 0); \quad E(X) = \lambda; \quad \text{Var}(X) = \lambda; \quad \mathfrak{Z}(\lambda) = \frac{1}{\lambda}$$

Propriedades:

$$- X_i \sim \text{Po}(\lambda_i) \text{ independentes } (i = 1, 2, \dots, k) \Rightarrow X = \sum_{i=1}^k X_i \sim \text{Po}\left(\sum_{i=1}^k \lambda_i\right)$$

$$- X \sim B(n, \theta) \text{ com } n \text{ grande e } \theta \text{ pequeno, então } X \sim \text{Po}(n\theta)$$

• UNIFORME (CONTÍNUA) $X \sim U(\alpha, \beta)$

$$f(x|\alpha, \beta) = \frac{1}{\beta - \alpha} \quad \alpha < x < \beta; \quad E(X) = \frac{\alpha + \beta}{2}; \quad \text{Var}(X) = \frac{(\beta - \alpha)^2}{12}$$

• NORMAL $X \sim N(\mu, \sigma^2)$

$$f(x|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}(x - \mu)^2\right\}, \quad -\infty < x < +\infty, \quad -\infty < \mu < +\infty, \quad 0 < \sigma < +\infty;$$

$$E(X) = \mu; \quad \text{Var}(X) = \sigma^2; \quad \mathfrak{Z}(\mu) = \frac{1}{\sigma^2} \quad \sigma^2 \text{ conhecido}; \quad \mathfrak{Z}(\sigma^2) = \frac{1}{2\sigma^4} \quad \mu \text{ conhecido}$$

Propriedades:

$$- \text{Normal estandardizada } Z = \frac{X - \mu}{\sigma} \sim N(0, 1); \quad \phi(z) = \phi(-z); \quad \Phi(z) = 1 - \Phi(-z)$$

$$- X_i \sim N(\mu, \sigma^2) \text{ independentes } (i = 1, 2, \dots, n) \Rightarrow Y = \sum_{i=1}^n X_i \sim N(n\mu, n\sigma^2); \quad \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

$$- X_i \sim N(\mu_i, \sigma_i^2) \text{ independentes } (i = 1, 2, \dots, k) \Rightarrow Y = \sum_{i=1}^k \alpha_i X_i \sim N(\mu_Y, \sigma_Y^2); \quad \mu_Y = \sum_{i=1}^k \alpha_i \mu_i; \quad \sigma_Y^2 = \sum_{i=1}^k \alpha_i^2 \sigma_i^2$$

- **EXPONENCIAL** $X \sim \text{Ex}(\lambda)$; $X \sim \text{Ex}(\lambda) \Leftrightarrow X \sim G(1, \lambda)$

$$f(x|\lambda) = \lambda e^{-\lambda x} \quad x > 0, \lambda > 0; \quad F(x|\lambda) = 1 - e^{-\lambda x}; \quad E(X) = \frac{1}{\lambda}; \quad \text{Var}(X) = \frac{1}{\lambda^2}; \quad \mathfrak{I}(\lambda) = \frac{1}{\lambda^2}$$

Propriedades:

$$- X_i \sim \text{Ex}(\lambda) \text{ independentes } (i = 1, 2, \dots, k) \Rightarrow \sum_{i=1}^k X_i \sim G(k; \lambda) \text{ e } \min_i X_i \sim \text{Ex}(k\lambda)$$

- **GAMA** $X \sim G(\alpha, \lambda)$

$$\text{Função gama: } \Gamma(\alpha) = \int_0^{+\infty} e^{-x} x^{\alpha-1} dx \quad (\alpha > 0) \text{ com } \Gamma(\alpha) = (\alpha-1)\Gamma(\alpha-1), \alpha > 1; \quad \Gamma(n) = (n-1)!; \quad \Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$$

$$f(x|\alpha, \lambda) = \frac{\lambda^\alpha e^{-\lambda x} x^{\alpha-1}}{\Gamma(\alpha)}, \quad x > 0, \alpha, \lambda > 0; \quad E(X) = \frac{\alpha}{\lambda}; \quad \text{Var}(X) = \frac{\alpha}{\lambda^2}; \quad \mathfrak{I}(\lambda) = \frac{\alpha}{\lambda^2} \quad \alpha \text{ conhecido}$$

Propriedades:

$$- X_i \sim G(\alpha_i, \lambda), (i = 1, 2, \dots, k) \text{ independentes} \Rightarrow \sum_{i=1}^k X_i \sim G(\alpha, \lambda); \quad \alpha = \sum_{i=1}^k \alpha_i$$

$$- X \sim G(\alpha, \lambda) \Rightarrow cX \sim G\left(\alpha, \frac{\lambda}{c}\right) \quad c > 0 \text{ constante}$$

- **QUI-QUADRADO** $X \sim \chi^2(n)$

$$f(x|n) = \frac{e^{-\frac{x}{2}} x^{\frac{n}{2}-1}}{2^{\frac{n}{2}} \Gamma\left(\frac{n}{2}\right)}, \quad x > 0, \quad n > 0 \text{ inteiro}; \quad E(X) = n; \quad \text{Var}(X) = 2n;$$

Propriedades:

$$- X \sim \chi^2(n) \Leftrightarrow X \sim G\left(\frac{n}{2}, \frac{1}{2}\right)$$

$$- \sqrt{2\chi^2(n)} - \sqrt{2n-1} \overset{a}{\sim} N(0,1)$$

$$- X \sim G(n; \lambda) \Leftrightarrow 2\lambda X \sim \chi^2(2n)$$

$$- X_i \sim \chi^2(n_i), \text{ independentes } (i = 1, 2, \dots, k) \Rightarrow \sum_{i=1}^k X_i \sim \chi^2(n), \quad n = \sum_{i=1}^k n_i$$

$$- X_i \sim N(0,1) \text{ independentes } (i = 1, 2, \dots, k) \Rightarrow \sum_{i=1}^n X_i^2 \sim \chi^2(n); \quad X \sim N(0,1) \Rightarrow X^2 \sim \chi^2(1)$$

- **t-“STUDENT”** $T \sim t(n)$

$$T = \frac{U}{\sqrt{V/n}} \sim t(n) \text{ com } U \sim N(0,1) \text{ e } V \sim \chi^2(n) \text{ independentes}; \quad E(T) = 0; \quad \text{Var}(T) = \frac{n}{n-2} \quad (n > 2)$$

- **F-SNEDCOR** $X \sim F(m, n)$

$$F = \frac{U/m}{V/n} \sim F(m, n) \text{ com } U \sim \chi^2(m), \quad V \sim \chi^2(n) \text{ independentes}$$

$$E(X) = \frac{n}{n-2} \quad (n > 2); \quad \text{Var}(X) = \frac{2n^2(m+n-2)}{m(n-2)^2(n-4)} \quad (n > 4)$$

$$\text{Propriedades:} \quad - X \sim F(m, n) \Rightarrow \frac{1}{X} \sim F(n, m) \quad - T \sim t(n) \Rightarrow T^2 \sim F(1, n)$$

• TEOREMA DO LIMITE CENTRAL E COROLÁRIOS

$$X_i \text{ iid, com } E(X_i) = \mu \text{ e } \text{Var}(X_i) = \sigma^2 \Rightarrow \frac{\sum_{i=1}^n X_i - n\mu}{\sqrt{n} \sigma} = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \overset{a}{\sim} N(0,1)$$

$$\text{Corolário: } X_i \sim B(1; \theta), \text{ independentes então } \frac{\sum_{i=1}^n X_i - n\theta}{\sqrt{n\theta(1-\theta)}} \overset{a}{\sim} N(0,1)$$

$$\text{Correcção de continuidade: } P(a \leq X \leq b) \approx \Phi\left(\frac{b + \frac{1}{2} - n\theta}{\sqrt{n\theta(1-\theta)}}\right) - \Phi\left(\frac{a - \frac{1}{2} - n\theta}{\sqrt{n\theta(1-\theta)}}\right) \text{ com } a \text{ e } b \text{ inteiros}$$

$$\text{Corolário: } X \sim \text{Po}(\lambda) \Rightarrow \frac{X - \lambda}{\sqrt{\lambda}} \overset{a}{\sim} N(0,1), \text{ quando } \lambda \rightarrow +\infty$$

Correcção de continuidade: $P(a \leq X \leq b) \approx \Phi\left(\frac{b + \frac{1}{2} - \lambda}{\sqrt{\lambda}}\right) - \Phi\left(\frac{a - \frac{1}{2} - \lambda}{\sqrt{\lambda}}\right)$

• **AMOSTRAGEM. DISTRIBUIÇÕES POR AMOSTRAGEM**

• **MÉDIA E VARIÂNCIA AMOSTRAIS**

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i ; S^2 = \frac{1}{n} \sum_{i=1}^n X_i^2 - \bar{X}^2 ; S'^2 = \frac{n}{n-1} S^2 ; E(\bar{X}) = \mu ; \text{Var}(\bar{X}) = \frac{\sigma^2}{n} ; E(S^2) = \frac{n-1}{n} \sigma^2 ; E(S'^2) = \sigma^2$$

• **DISTRIBUIÇÕES POR AMOSTRAGEM**

POPULAÇÕES NORMAIS

Média	$\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$	$\frac{\bar{X} - \mu}{S'/\sqrt{n}} \sim t(n-1)$
Diferença de médias	Variâncias conhecidas $\frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}} \sim N(0,1)$	
	Variâncias desconhecidas mas iguais $T = \frac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{\sqrt{\frac{1}{m} + \frac{1}{n}} \sqrt{\frac{(m-1)S_1'^2 + (n-1)S_2'^2}{m+n-2}}} \sim t(m+n-2)$	
Variância	$\frac{nS^2}{\sigma^2} = \frac{(n-1)S'^2}{\sigma^2} \sim \chi^2(n-1)$	
Relação de variâncias	$\frac{S_1'^2}{S_2'^2} \frac{\sigma_2^2}{\sigma_1^2} \sim F(m-1, n-1)$ ou $\frac{S_2'^2}{S_1'^2} \frac{\sigma_1^2}{\sigma_2^2} \sim F(n-1, m-1)$	
Amostras emparelhadas	$\frac{\bar{Z} - \mu_X - \mu_Y}{S'_Z/\sqrt{n}} \sim t(n-1),$ (X_i, Y_i) - amostra emparelhada, $Z_i = X_i - Y_i ; \bar{Z} = \bar{X} - \bar{Y}$	

GRANDES AMOSTRAS: CASO GERAL

Média	$\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$	$\frac{\bar{X} - \mu}{S'/\sqrt{n}} \sim N(0,1)$ ou $\frac{\bar{X} - \mu}{S/\sqrt{n}} \sim N(0,1)$
Diferença de médias	$\frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}} \sim N(0,1)$	$\frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1'^2}{m} + \frac{S_2'^2}{n}}} \sim N(0,1)$ ou $\frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{m} + \frac{S_2^2}{n}}} \sim N(0,1)$

GRANDES AMOSTRAS: POPULAÇÃO DE BERNOULLI

Proporção	$\frac{\bar{X} - \theta}{\sqrt{\frac{\theta(1-\theta)}{n}}} \stackrel{a}{\sim} N(0,1)$	$\frac{\bar{X} - \theta}{\sqrt{\frac{\bar{X}(1-\bar{X})}{n}}} \stackrel{a}{\sim} N(0,1)$
Diferença de proporções	$\frac{\bar{X}_1 - \bar{X}_2 - (\theta_1 - \theta_2)}{\sqrt{\frac{\theta_1(1-\theta_1)}{m} + \frac{\theta_2(1-\theta_2)}{n}}} \stackrel{a}{\sim} N(0,1)$	$\frac{\bar{X}_1 - \bar{X}_2 - (\theta_1 - \theta_2)}{\sqrt{\frac{\bar{X}_1(1-\bar{X}_1)}{m} + \frac{\bar{X}_2(1-\bar{X}_2)}{n}}} \stackrel{a}{\sim} N(0,1)$
Igualdade de proporções	$\frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\left(\frac{1}{m} + \frac{1}{n}\right)\hat{\theta}(1-\hat{\theta})}} \stackrel{a}{\sim} N(0,1)$ com $\hat{\theta} = \frac{m\bar{X}_1 + n\bar{X}_2}{m+n}$	

GRANDES AMOSTRAS: POPULAÇÃO DE POISSON

Média	$\frac{\bar{X} - \lambda}{\sqrt{\frac{\lambda}{n}}} \stackrel{a}{\sim} N(0,1)$	$\frac{\bar{X} - \lambda}{\sqrt{\frac{\bar{X}}{n}}} \stackrel{a}{\sim} N(0,1)$
Diferença de médias	$\frac{\bar{X}_1 - \bar{X}_2 - (\lambda_1 - \lambda_2)}{\sqrt{\frac{\lambda_1}{m} + \frac{\lambda_2}{n}}} \stackrel{a}{\sim} N(0,1)$	$\frac{\bar{X}_1 - \bar{X}_2 - (\lambda_1 - \lambda_2)}{\sqrt{\frac{\bar{X}_1}{m} + \frac{\bar{X}_2}{n}}} \stackrel{a}{\sim} N(0,1)$
Igualdade de médias	$\frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\left(\frac{1}{m} + \frac{1}{n}\right)\bar{X}}} \stackrel{a}{\sim} N(0,1)$ com $\bar{X} = \frac{m\bar{X}_1 + n\bar{X}_2}{m+n}$	

- **ESTATÍSTICA-TESTE DO χ^2**

- **TESTE DE AJUSTAMENTO**

$$Q = \sum_{j=1}^m \frac{(N_j - fe_j)^2}{fe_j} \stackrel{a}{\sim} \chi^2(m-1); \quad fe_j = np_{\circ j} - \text{frequência esperada da classe } j;$$

Com estimação de k parâmetros para obter as estimativas $\hat{p}_{\circ j}$: $Q \stackrel{a}{\sim} \chi^2(m-k-1)$

- **TESTE DE INDEPENDÊNCIA:**

$$Q = \sum_{i=1}^r \sum_{j=1}^s \frac{(N_{ij} - fe_{ij})^2}{fe_{ij}} \stackrel{a}{\sim} \chi^2((r-1)(s-1)); \quad fe_{ij} = \frac{N_{i\circ} N_{\circ j}}{n} - \text{frequência esperada da classe } ij$$