Formulário de Estatística

Medidas descritivas

$$\mu = \frac{\sum_i x_i}{n} \qquad \sigma^2 = \frac{\sum_i (x_i - \mu)^2}{n} = \frac{\sum_i x_i^2 - n\bar{x}^2}{n}$$

$$\bar{x} = \frac{\sum_i x_i}{n} \qquad S^2 = \frac{\sum_i (x_i - \bar{x})^2}{n-1} = \frac{\sum_i x_i^2 - n\bar{x}^2}{n-1}$$

$$CV = 100\frac{S}{\bar{x}}$$

$$\chi^2 = \sum_i \frac{(o_i - e_i)^2}{e_i} \qquad C = \sqrt{\frac{\chi^2}{\chi^2 + n}} \qquad C^* = \frac{C}{\sqrt{(t-1)/t}}$$

$$r = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i - \bar{x}}{s_x}\right)\left(\frac{y_i - \bar{y}}{s_y}\right) = \frac{\sum_i x_i y_i - n\bar{x}\bar{y}}{\sqrt{(\sum_i x_i^2 - n\bar{x}^2)(\sum_i y_i^2 - n\bar{y}^2)}}$$

Probabilidades

$$\mu_X = E[X] = \sum_i x_i P(X = x_i) \qquad \sigma_X^2 = Var[X] = \sum_i (x_i - \mu_x)^2 P(X = x_i)$$

$$\mu_X = E[X] = \int x f_X(x) dx \qquad \sigma_X^2 = Var[X] = \int (x - \mu_x)^2 f_X(x)$$

Distribuições de Probabilidade

Distribuição	Densidade	Domínio	E[X]	Var[X]
$X \sim U(k)$	$\frac{1}{k}$	$x = 1, 2, \ldots, k$	$\frac{min(X)+max(x)}{2}$	$\sqrt{\frac{(max(X)-min(X)+1)^2-1}{12}}$
$X \sim B(n,p)$	$\binom{n}{x}p^x(1-p)^{n-x}$	$x = 0, 1, 2, \ldots, n$	np	$np(1-p)$
$X \sim HG(N,K,n)$	$\frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}}$	$x = 0, 1, 2, \ldots, min(K,n)$	np	$np(1-p)\frac{N-n}{N-1}$
$X \sim P(\lambda)$	$\frac{e^{-\lambda}\lambda^x}{x!}$	$x = 0, 1, 2, \ldots$	λ	λ
$X \sim G(p)$	$(1-p)^{x-1}p$	$x = 1, 2, \ldots$	$\frac{1}{p}$	$\frac{(1-p)}{p^2}$
$X \sim BN(r,p)$	$\binom{x-1}{r-1}(1-p)^{x-r}p^r$	$x = r, r+1, r+2, \ldots$	$\frac{r}{p}$	$\frac{r(1-p)}{p^2}$
$X \sim U[a,b]$	$\frac{1}{b-a}$	$a \le x \le b$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
$X \sim N(\mu, \sigma^2)$	$\frac{1}{\sqrt{2\pi\sigma^2}}\exp\{-\frac{1}{2\sigma^2}(x-\mu)^2\}$	$x \in (-\infty, \infty)$	μ	σ^2
$X \sim Exp(\lambda)$	$\lambda \exp(-\lambda x)$	$x \ge 0$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
$X \sim G(\alpha, \beta)$	$\frac{1}{\Gamma(\alpha)\beta^\alpha}\, x^{\alpha-1}\, \exp\{-x/\beta\}$	$x \ge 0$	$\alpha\beta$	$\alpha\beta^2$
$X \sim Wei(\gamma, \beta)$	$\frac{\gamma}{\beta}\, x^{\gamma-1}\exp\{-x^\gamma/\beta\}$	$x \ge 0$	$\beta^{1/\gamma}\Gamma(1+\frac{1}{\gamma})$	$\beta^{2/\gamma}\left[\Gamma(1+\frac{2}{\gamma}) - \Gamma^2(1+\frac{1}{\gamma})\right]$
$X \sim Beta(\alpha, \beta)$	$\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\, x^{\alpha-1}(1-x)^{\beta-1}$	$0 \le x \le 1$	$\frac{\alpha}{\alpha+\beta}$	$\frac{\alpha\,\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$

$$Z = (X - \mu)/\sigma^2 \qquad \Gamma(\alpha) = \int_0^\infty x^{\alpha-1}\, \exp\{-x\}\, dx \qquad \Gamma(\alpha) = (\alpha-1)! \qquad \Gamma(\alpha+1) = \alpha\,\Gamma(\alpha)$$

Distribuições Amostrais, Intervalos de Confiança e Testes de Hipóteses

Estimador	Intervalo de Confiança	Estatística de Teste	Obs.
$\bar{x} \sim N(\mu, \frac{\sigma^2}{n})$	$\bar{x} \pm z_t \frac{\sigma}{\sqrt{n}}$	$z_c = \frac{\bar{x} - \mu_0}{\sqrt{\sigma^2/n}}$	
$\bar{x} \sim t_{n-1}(\mu, \frac{S^2}{n})$	$\bar{x} \pm t_t \frac{S}{\sqrt{n}}$	$t_c = \frac{\bar{x} - \mu_0}{\sqrt{S^2/n}}$	$\nu = n - 1$
$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)}$	$\left(\frac{(n-1)S^2}{\chi^2_{sup}}, \frac{(n-1)S^2}{\chi^2_{inf}}\right)$	$\chi^2_c = \frac{(n-1)S^2}{\sigma_0^2}$	$\nu = n - 1$
$\hat{p} \sim N(p, \frac{p(1-p)}{n})$	$\hat{p} \pm z_t \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$	$z_c = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$	$\hat{p} \pm z_t \sqrt{\frac{1}{4n}}$
$\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \sim F(n_1 - 1, n_2 - 1)$	$\left(\frac{1}{F_{\alpha/2,n_1-1,n_2-1}} \frac{S_1^2}{S_2^2}, F_{\alpha/2,n_2-1,n_1-1} \frac{S_1^2}{S_2^2}\right)$	$F_c = \frac{S_1^2}{S_2^2}$	$F_{1-\alpha/2,\nu_2,\nu_1} = \frac{1}{F_{\alpha/2,\nu_1,\nu_2}}$
$\bar{x}_1 - \bar{x}_2 \sim N\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)$	$(\bar{x}_1 - \bar{x}_2) \pm z_{\alpha/2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$	$z_c = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$	
$\bar{x}_1 - \bar{x}_2 \sim t_\nu\left(\mu_1 - \mu_2, \frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)$	$(\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2,\nu} \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$	$t_c = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$	$gl = \nu = \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{(S_1^2/n_1)^2}{n_1-1} + \frac{(S_2^2/n_2)^2}{n_2-1}}$
$\bar{x}_1 - \bar{x}_2 \sim t_\nu\left(\mu_1 - \mu_2, S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)\right)$	$(\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2,\nu} \sqrt{S_p^2(\frac{1}{n_1} + \frac{1}{n_2})}$	$t_c = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$	$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$ $gl = \nu = n_1 + n_2 - 2$
$\bar{d} \sim t_{n-1}(\mu_d, \frac{S_d^2}{n})$	$\bar{d} \pm t_t \frac{S_d}{\sqrt{n}}$	$t_c = \frac{\bar{d} - d_0}{\sqrt{S_d^2/n_d}}$	$\nu = n_d - 1$
$\hat{p_1} - \hat{p_2} \sim N\left(p_1 - p_2\,,\ \frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}\right)$	$\hat{p_1} - \hat{p_2} \pm z_{\alpha/2} \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$	$z_c = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}}$	
		$\chi^2_c = \sum_i \frac{(o_i - e_i)^2}{e_i}$	$\nu = (L - 1)(C - 1)$ $\nu = k - 1$

Regressão linear simples

$$y_i = \beta_0 + \beta_1 x_i + e_i \qquad e_i \sim N(0, \sigma^2) \qquad \hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} \qquad \hat{\beta}_1 = \frac{\sum_i x_i y_i - n\bar{x}\bar{y}}{\sum_i x_i^2 - n\bar{x}^2}$$

$$SQreg = \hat{\beta}_1^2 \sum_i (x_i - \bar{x})^2 \qquad SQtot = \sum_i (y_i - \bar{y})^2 \qquad S_e^2 = \frac{\sum_i (y_i - \hat{y}_i)^2}{n - 2} \qquad R^2 = \frac{SQreg}{SQtot}$$