Formulário de Estatística

Medidas descritivas

$$\mu = \frac{\sum_i y_i}{n} \qquad \sigma^2 = \frac{\sum_i (y_i-\mu)^2}{n} = \frac{\sum_i y_i^2 - n\overline{y}^2}{n}$$

$$\overline{y} = \frac{\sum_i y_i}{n} \qquad S^2 = \frac{\sum_i (y_i-\overline{y})^2}{n-1} = \frac{\sum_i y_i^2 - n\overline{y}^2}{n-1} \qquad CV = 100\frac{S}{\overline{y}}$$

$$\chi^2 = \sum_i \frac{(o_i-e_i)^2}{e_i} \qquad C = \sqrt{\frac{\chi^2}{\chi^2+n}} \qquad C' = \sqrt{\frac{\chi^2/n}{(r-1)(s-1)}} \qquad C'' = \frac{C}{\sqrt{(t-1)/t}}$$

$$r = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i-\overline{x}}{s_x}\right)\left(\frac{y_i-\overline{y}}{s_y}\right) = \frac{\sum_i x_i y_i - n\overline{xy}}{\sqrt{(\sum_i x_i^2 - n\overline{x}^2)(\sum_i y_i^2 - n\overline{y}^2)}} \qquad Y = g(X) \to f_Y(y) = f_X(g^{-1}(y))\left|\frac{\mathrm{d}g^{-1}(y)}{\mathrm{d}y}\right|$$

Probabilidades

$$\mu_Y = E[Y] = \sum_i y_i P(Y=y_i) \qquad \sigma_Y^2 = Var[Y] = \sum_i (y_i-\mu_y)^2 P(Y=y_i)$$

$$\mu_Y = E[Y] = \int y f_Y(y)dy \qquad \sigma_Y^2 = Var[Y] = \int (y-\mu_y)^2 f_Y(y)$$

Distribuições de Probabilidade

Distribuição	Fç de Probabilidade/Densidade	Domínio	E[Y]	Var[Y]
$Y \sim \mathrm{U}(k)$	$\frac{1}{k}$	$y \in \{1,2,\ldots,k\}$	$\frac{y_{(1)}+y_{(n)}}{2}$	$\frac{k^2-1}{12}$
$Y \sim \mathrm{B}(n,p)$	$\binom{n}{y}p^y(1-p)^{n-y}$	$y \in \{0,1,2,\ldots,n\}$	np	$np(1-p)$
$Y \sim \mathrm{HG}(m,n,k)$	$\frac{\binom{m}{y}\binom{n}{k-y}}{\binom{m+n}{k}}$	$y \in \{\max(0,k-n),\ldots,\ldots,\min(k,m)\}$	$kp = k\frac{m}{m+n}$	$kp(1-p)\frac{m+n-k}{m+n-1}$
$Y \sim \mathrm{P}(\lambda)$	$\frac{e^{-\lambda}\lambda^y}{y!}$	$y \in \{0,1,2,\ldots\}$	λ	λ
$Y \sim \mathrm{G}(p)$	$(1-p)^y p$	$y \in \{0,1,2,\ldots\}$	$\frac{1-p}{p}$	$\frac{(1-p)}{p^2}$
$Y \sim \mathrm{BN}(r,p)$	$\binom{y+r-1}{r-1}(1-p)^y p^r$	$y \in \{0,1,2,\ldots\}$	$r\frac{(1-p)}{p}$	$r\frac{(1-p)}{p^2}$
$Y \sim \mathrm{U}[a,b]$	$\frac{1}{b-a}$	$a \le y \le b$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
$Y \sim \mathrm{N}(\mu,\sigma^2)$	$\frac{1}{\sqrt{2\pi\sigma^2}}\exp\{-\frac{1}{2\sigma^2}(y-\mu)^2\}$	$y \in (-\infty,\infty)$	μ	σ^2
$Y \sim \mathrm{LN}(\mu,\sigma^2)$	$\frac{1}{y\sqrt{2\pi\sigma^2}}\exp\{-\frac{1}{2\sigma^2}(\ln(y)-\mu)^2\}$	$y \in (0,\infty)$	$\exp\{\mu+\sigma^2/2\}$	$\exp\{2\mu+\sigma^2\}(\exp\{\sigma^2\}-1)$
$Y \sim \mathrm{Exp}(\lambda)$	$\lambda\exp(-\lambda y)$	$y \ge 0$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
$Y \sim \mathrm{Erlang}(\lambda,r)$	$\lambda^r y^{r-1}\exp(-\lambda y)/(r-1)!$	$y \ge 0$; $r = 1,2,\ldots$	$\frac{r}{\lambda}$	$\frac{r}{\lambda^2}$
$Y \sim \mathrm{G}(\alpha,\beta)$	$\frac{1}{\Gamma(\alpha)\,\beta^\alpha}\ y^{\alpha-1}\ \exp\{-y/\beta\}$	$y \ge 0$	$\alpha\beta$	$\alpha\beta^2$
$Y \sim \mathrm{Weibull}(\alpha,\beta)$	$\frac{\alpha}{\beta}\ (y/\beta)^{\alpha-1}\exp\{-(y/\beta)^\alpha\}$	$y \ge 0$	$\beta\,\Gamma(1+\frac{1}{\alpha})$	$\beta^2\left[\Gamma(1+\frac{2}{\alpha})-\Gamma^2(1+\frac{1}{\alpha})\right]$
$Y \sim \mathrm{Beta}(\alpha,\beta)$	$\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\ y^{\alpha-1}(1-y)^{\beta-1}$	$0 \le y \le 1$	$\frac{\alpha}{\alpha+\beta}$	$\frac{\alpha\,\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$

$$Z = (Y-\mu)/\sigma \qquad \Gamma(\alpha) = \int_0^\infty t^{\alpha-1}\ \exp\{-t\}\ dt \qquad \Gamma(\alpha) = (\alpha-1)! \qquad \Gamma(\alpha+1) = \alpha\,\Gamma(\alpha) \qquad \Gamma(1/2) = \sqrt{\pi}$$

Distribuições Amostrais, Intervalos de Confiança e Testes de Hipóteses

Estimador	Intervalo de Confiança	Estatística de Teste	Obs.
$\overline{y} \sim N(\mu, \frac{\sigma^2}{n})$	$\overline{y} \pm z_t \frac{\sigma}{\sqrt{n}}$	$z_c = \frac{\overline{y} - \mu_0}{\sqrt{\sigma^2/n}}$	
$\overline{y} \sim t_{n-1}(\mu, \frac{S^2}{n})$	$\overline{y} \pm t_t \frac{S}{\sqrt{n}}$	$t_c = \frac{\overline{y} - \mu_0}{\sqrt{S^2/n}}$	$\nu = n - 1$
$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)}$	$\left(\frac{(n-1)S^2}{\chi^2_{sup}}, \frac{(n-1)S^2}{\chi^2_{inf}} \right)$	$\chi^2_c = \frac{(n-1)S^2}{\sigma_0^2}$	$\nu = n - 1$
$\hat{p} \sim N(p, \frac{p(1-p)}{n})$	$\hat{p} \pm z_t \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$	$z_c = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$	$\hat{p} \pm z_t \sqrt{\frac{1}{4n}}$
$\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \sim F(n_1 - 1, n_2 - 1)$	$\left(\frac{1}{F_{\alpha/2, n_1-1, n_2-1}} \frac{S_1^2}{S_2^2}, F_{\alpha/2, n_2-1, n_1-1} \frac{S_1^2}{S_2^2} \right)$	$F_c = \frac{S_1^2}{S_2^2}$	$F_{1-\alpha/2, \nu_2, \nu_1} = \frac{1}{F_{\alpha/2, \nu_1, \nu_2}}$
$\overline{y}_1 - \overline{y}_2 \sim N\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)$	$(\overline{y}_1 - \overline{y}_2) \pm z_{\alpha/2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$	$z_c = \frac{(\overline{y}_1 - \overline{y}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$	
$\overline{y}_1 - \overline{y}_2 \sim t_\nu \left(\mu_1 - \mu_2, \frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)$	$(\overline{y}_1 - \overline{y}_2) \pm t_{\alpha/2, \nu} \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$	$t_c = \frac{(\overline{y}_1 - \overline{y}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$	$gl = \nu = \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{(S_1^2/n_1)^2}{n_1-1} + \frac{(S_2^2/n_2)^2}{n_2-1}}$
$\overline{y}_1 - \overline{y}_2 \sim t_\nu \left(\mu_1 - \mu_2, S_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2}\right)\right)$	$(\overline{y}_1 - \overline{y}_2) \pm t_{\alpha/2, \nu} \sqrt{S_p^2 (\frac{1}{n_1} + \frac{1}{n_2})}$	$t_c = \frac{(\overline{y}_1 - \overline{y}_2) - (\mu_1 - \mu_2)}{\sqrt{S_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$	$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$ $gl = \nu = n_1 + n_2 - 2$
$\overline{d} \sim t_{n-1}(\mu_d, \frac{S_d^2}{n})$	$\overline{d} \pm t_t \frac{S_d}{\sqrt{n}}$	$t_c = \frac{\overline{d} - d_0}{\sqrt{S_d^2/n_d}}$	$\nu = n_d - 1$
$\hat{p}_1 - \hat{p}_2 \sim N\left(p_1 - p_2\ , \frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}\right)$	$\hat{p}_1 - \hat{p}_2 \pm z_{\alpha/2} \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$	$z_c = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\overline{p}(1-\overline{p}) \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$	$\overline{p} = \frac{y_1 + y_2}{n_1 + n_2}$
$\frac{1}{2} \ln \left(\frac{1+r}{1-r}\right) \sim N\left(\frac{1}{2} \ln \left(\frac{1+\rho}{1-\rho}\right), \frac{1}{n-3}\right)$	$\left(\tanh\left(\text{arctanh}(r) \pm \frac{z_{\alpha/2}}{\sqrt{n-3}}\right)\right)$	$z_c = (\arctan(r) - \arctan(\rho))\sqrt{n-3}$	
$\frac{r\sqrt{n-2}}{\sqrt{1-r^2}} \sim t_{n-2}$		$t_c = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$	apenas para testar $\rho = 0$
		$\chi^2_c = \sum_i \frac{(o_i - e_i)^2}{e_i}$	$\nu = k - 1$(aderência) $\nu = (L-1)(C-1)$(independência)

$y_i = \beta_0 + \beta_1 x_i + e_i$	$e_i \sim N(0, \sigma^2)$	$\hat{\beta}_0 = \overline{y} - \hat{\beta}_1 \overline{x}$	$\hat{\beta}_1 = \frac{\sum_i x_i y_i - n\overline{x}\overline{y}}{\sum_i x_i^2 - n\overline{x}^2}$
$SQreg = \hat{\beta}_1^2 \sum_i (x_i - \overline{x})^2$	$SQtot = \sum_i (y_i - \overline{y})^2$	$S_e^2 = \frac{\sum_i (y_i - \hat{y}_i)^2}{n-2}$	$R^2 = \frac{SQreg}{SQtot}$