Intervalo de confiança para a média ($\boldsymbol{\mu}$)

(i) **Variância ($\boldsymbol{\sigma^2}$) é conhecida:** I.C.(μ; 100γ%) $= \bar{x} \pm z_c\, \sigma/\sqrt{n}$

onde $\gamma = P(-z_c \leq Z \leq z_c)$ e $Z \sim \mathbb{N}(0;1)$.

Tamanho da amostra: $n = (z_c\, \sigma/\varepsilon_{\bar{x}})^2$, onde $\varepsilon_{\bar{x}} = |\bar{x} - \mu|$ é a margem de erro.

(ii) **Variância ($\boldsymbol{\sigma^2}$) é desconhecida:** I.C.(μ; 100γ%) $= \bar{x} \pm t_c\, s/\sqrt{n}$

em que $s = \sqrt{\frac{1}{(n-1)} \sum_{i=1}^{n} (x_i - \bar{x})^2}$, $\gamma = P(-t_c \leq T \leq t_c)$ e $T \sim t_{(n-1)}$

Intervalo de confiança para a proporção ($\boldsymbol{p}$)

(i) **Grandes amostras:** I.C.(p; 100γ%) $= \hat{p} \pm z_c \sqrt{\hat{p}(1-\hat{p})/n}$, onde $\gamma = P(-z_c \leq Z \leq z_c)$ e $Z \sim \mathbb{N}(0;1)$.

Tamanho da amostra: $n = \hat{p}(1-\hat{p})\, \left[z_c/\varepsilon_{\hat{p}}\right]^2$, onde $\varepsilon_{\hat{p}}$ é a margem de erro e $\hat{p}$ é uma estimativa conhecida.

(ii) **Intervalo conservativo:** I.C.(p; 100γ%) $= \hat{p} \pm z_c \sqrt{(0{,}25)/n}$

Tamanho da amostra: $n = (1/4)\left[z_c/\varepsilon_{\hat{p}}\right]^2$, onde $\varepsilon_{\hat{p}}$ é a margem de erro.

Intervalo de confiança para a variância ($\boldsymbol{\sigma^2}$)

$$\text{I.C.}(\sigma^2, 100\gamma\%) = \left[\frac{(n-1)s^2}{q_2}; \frac{(n-1)s^2}{q_1}\right], \text{ onde } \gamma = P(q_1 \leq Q \leq q_2) \text{ e } Q \leq \chi^2_{(n-1)}$$

Intervalo de confiança para a diferença de médias de populações independentes ($\boldsymbol{\mu_1 - \mu_2}$)

(i) **Variâncias desconhecidas, mas iguais.**

$$\text{I.C.}(\mu_1 - \mu_2; 100\gamma\%) = (\bar{x}_1 - \bar{x}_2) \pm t_c \sqrt{s^2_{comum}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

em que $s^2_{comum} = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{(n_1+n_2-2)}$, $\gamma = P(-t_c \leq T \leq t_c)$ e $T \sim t_{(n_1+n_2-2)}$

(ii) **Variâncias desconhecidas e diferentes.**

$$\text{I.C.}(\mu_1 - \mu_2; 100\gamma\%) = (\bar{x}_1 - \bar{x}_2) \pm t_c^* \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

em que $\gamma = P(-t_c^* \leq T^* \leq t_c^*)$ e $T^* \sim t_{(v)}$ e $v = \dfrac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1-1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2-1}}$ (Fórmula de Sattertwaite)

Teste de hipóteses para a média (μ)

$H_0: \mu = \mu_0$
$H_a: \mu < \mu_0$ (ou $H_a: \mu \neq \mu_0$ ou $H_a: \mu > \mu_0$)

$(\boldsymbol{i})$ **Variância** $(\boldsymbol{\sigma^2})$ **é conhecida** Estatística do teste: $Z = \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} \sim \mathbb{N}(0;1)$

$(\boldsymbol{ii})$ **Variância** $(\boldsymbol{\sigma^2})$ **é desconhecida** Estatística do teste: $T = \frac{\bar{x}-\mu_0}{s/\sqrt{n}} \sim t_{(n-1)}$

Teste de hipótese para a proporção $(\boldsymbol{p})$ $(\boldsymbol{n} \to \infty)$

$H_0: p = p_0$
$H_a: p < p_0$ (ou $H_a: p \neq p_0$ ou $H_a: p > p_0$)

Estatística do teste: $Z = \frac{\hat{p}-p_0}{\sqrt{p_0(1-p_0)/n}} \sim \mathbb{N}(0;1)$

Teste de hipótese para a variância $(\boldsymbol{\sigma^2})$

$H_0: \sigma^2 = \sigma_0^2$
$H_a: \sigma^2 < \sigma_0^2$ (ou $H_a: \sigma^2 \neq \sigma_0^2$ ou $H_a: \sigma^2 > \sigma_0^2$)

Estatística do teste: $Q = \frac{(n-1)s^2}{\sigma_0^2} \sim \chi^2_{(n-1)}$

Comparação das variâncias de duas populações (σ_1^2 e σ_2^2)

$H_0: \sigma_1^2 = \sigma_2^2 = \sigma^2$
$H_a: \sigma_1^2 < \sigma_2^2$ (ou $H_a: \sigma_1^2 \neq \sigma_2^2$ ou $H_a: \sigma_1^2 > \sigma_2^2$)

Estatística do teste: $F = {s^2_{maior}}/{s^2_{menor}} \sim F(n_1 - 1, n_2 - 1)$, em que n_1 e n_2 são os tamanhos das amostras que têm as variâncias $maior$ e $menor$, respectivamente.

Comparação de médias de duas populações normais independentes

$H_0: \mu_1 - \mu_2 = d$
$H_a: \mu_1 - \mu_2 < d$ (ou $H_a: \mu_1 - \mu_2 \neq d$ ou $H_a: \mu_1 - \mu_2 > d$)

$(\boldsymbol{i})$ **Variâncias desconhecidas, mas iguais.**

Estatística do teste: $T = \frac{(\bar{x}_1-\bar{x}_2)-d}{\sqrt{s^2_{comum}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} \sim t_{(n_1+n_2-2)}$ em que $s^2_{comum} = \frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{(n_1+n_2-2)}$

$(\boldsymbol{ii})$ **Variâncias desconhecidas e diferentes.**

Estatística do teste: $T^* = \frac{(\bar{x}_1-\bar{x}_2)-d}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}} \sim t_{(v)}$ em que o número de graus de liberdade (v) é calculado pela fórmula de Sattertwaite.

Comparação de médias de duas populações normais pareadas

$H_0: \mu_d = \mu_0$
$H_a: \mu_d < \mu_0$ (ou $H_a: \mu_d \neq \mu_0$ ou $H_a: \mu_d > \mu_0$)

Estatística do teste: $T = \frac{\bar{D} - \mu_0}{s_d/\sqrt{n}} \sim t_{(n-1)}$

onde $\mu_d = \mu_X - \mu_Y$, $D_i = X_i - Y_i$, $\bar{D} = \frac{1}{n}\sum_{i=1}^{n} D_i$ e $s_d = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(D_i - \bar{D})^2}$

Coeficiente de correlação linear de Pearson

Estimador: $r(X,Y) = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}}$

i*) Teste de Independência das variáveis normais *X* e *Y

$H_0: \rho(X,Y) = 0$ (as variáveis X e Y são independentes)
$H_a: \rho(X,Y) \neq 0$ (as variáveis X e Y não são independentes)

Estatística do teste: $T = \frac{r(X,Y)\sqrt{n-2}}{\sqrt{1 - r^2(X,Y)}} \sim t_{(n-2)}$

***i*) Teste para $\boldsymbol{\rho(X,Y) = \rho_0}$, para $\boldsymbol{\rho_0 \neq 0}$**

$H_0: \rho(X,Y) = \rho_0$
$H_a: \rho(X,Y) < \rho_0$ (ou $H_a: \rho(X,Y) \neq \rho_0$, ou $H_a: \rho(X,Y) > \rho_0$)

Estatística do teste: $Z = \frac{z - \mu_z}{\sigma_z} \sim \mathbb{N}(0;1)$

onde $z = 0{,}5\ ln\left[\frac{1 + r(X,Y)}{1 - r(X,Y)}\right]$, $\mu_z = 0{,}5\ ln\left(\frac{1+\rho_0}{1-\rho_0}\right)$ e $\sigma_z = 1\Big/\sqrt{(n-3)}$

Regressão linear simples

Modelo: $y_i = a + bx_i + \varepsilon_i$, para $i = 1, 2, \ldots, n$.

Modelo ajustado (ou estimado): $\hat{y}_i = \hat{a} + \hat{b}x_i$ **Resíduo do modelo:** $\hat{\varepsilon}_i = y_i - \hat{y}_i$

Onde:

$$\hat{b} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}} = \frac{\sum_{i=1}^{n} x_i y_i - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right)}{\sqrt{\sum_{i=1}^{n} x_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2}} \qquad \hat{a} = \bar{y} - \hat{b}\bar{x}$$

Coeficiente de determinação: $R^2 = r^2(X,Y) = (\hat{b})^2 \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$

Variância residual: $s_{y/x}^2 = \frac{1}{(n-2)}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2$

Regressão linear simples – **Intervalos de confiança**

$$\text{I.C.}(a, 100\gamma\%) = \hat{a} \pm t_c \sqrt{s^2_{y/x}\left(\frac{1}{n} + \frac{\bar{x}^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}\right)}, \text{ onde } \gamma = \mathrm{P}(-t_c \leq T \leq t_c) \text{ e } T \sim t_{(n-2)}$$

$$\text{I.C.}(b, 100\gamma\%) = \hat{b} \pm t_c \sqrt{\frac{s^2_{y/x}}{\sum_{i=1}^{n}(x_i - \bar{x})^2}}, \text{ onde } \gamma = \mathrm{P}(-t_c \leq T \leq t_c) \text{ e } T \sim t_{(n-2)}$$

$$\text{I.C.}(y_p, 100\gamma\%) = \hat{y}_p \pm t_c \sqrt{s^2_{y/x}\left(\frac{1}{n} + \frac{(x_p - \bar{x})^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}\right)}, \text{ onde } \hat{y}_p = \hat{a} + \hat{b}x_p.$$

Regressão linear simples – **Testes de hipóteses**

Teste para o intercepto $\boldsymbol{a}$

$H_0: a = a_0$
$H_a: a < a_0$ (ou $H_a: a \neq a_0$ ou $H_a: a > a_0$)

Estatística do teste: $T = \dfrac{\hat{a} - a_0}{\sqrt{s^2_{y/x}\left(\frac{1}{n} + \frac{\bar{x}^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}\right)}} \sim t_{(n-2)}$

Teste para o coeficiente angular $\boldsymbol{b}$

$H_0: b = b_0$
$H_a: b < b_0$ (ou $H_a: b \neq b_0$ ou $H_a: b > b_0$)

Estatística do teste: $T = \dfrac{\hat{b} - b_0}{\sqrt{\frac{s^2_{y/x}}{\sum_{i=1}^{n}(x_i - \bar{x})^2}}} \sim t_{(n-2)}$