Estatística: Aplicação ao Sensoriamento Remoto

SER 204 - ANO 2024

Intervalo de Confiança
(Extra)

Camilo Daleles Rennó
camilo.renno@inpe.br
http://www.dpi.inpe.br/~camilo/estatistica/

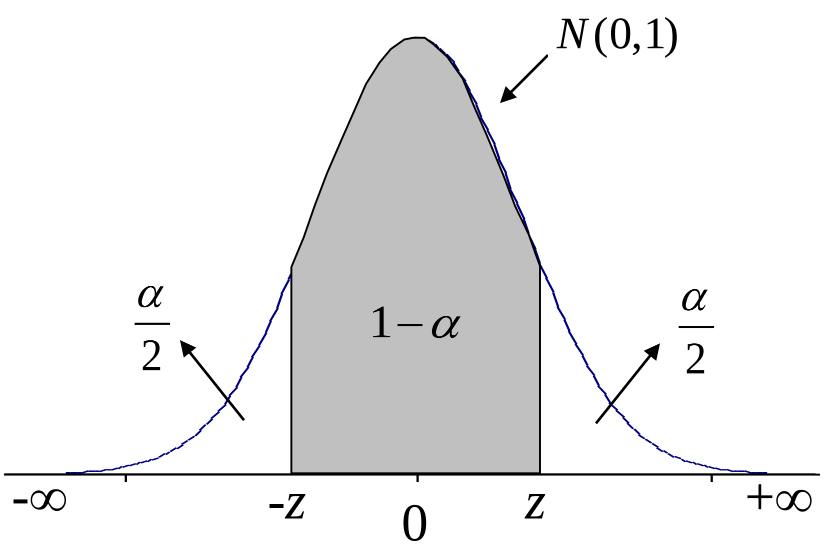
Intervalo de Confiança para μ_1 - μ_2

$X_1 \sim N(\mu_1, \sigma_1^2)$ $X_2 \sim N(\mu_2, \sigma_2^2)$ μ_i desconhecidas, mas σ_i^2 conhecidas

$\bar{X}_1 \sim N(\mu_1, \frac{\sigma_1^2}{n_1})$ $\bar{X}_2 \sim N(\mu_2, \frac{\sigma_2^2}{n_2})$ $\bar{X}_1 - \bar{X}_2 \sim N(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2})$

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1)$$

$$P\left(-z < \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} < z\right) = 1 - \alpha$$



$P(-z < Z < z) = 1 - \alpha$

$$\boxed{P\left(\bar{X}_1 - \bar{X}_2 - z\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} < \mu_1 - \mu_2 < \bar{X}_1 - \bar{X}_2 + z\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\right) = 1 - \alpha}$$

IC para μ_1 - μ_2

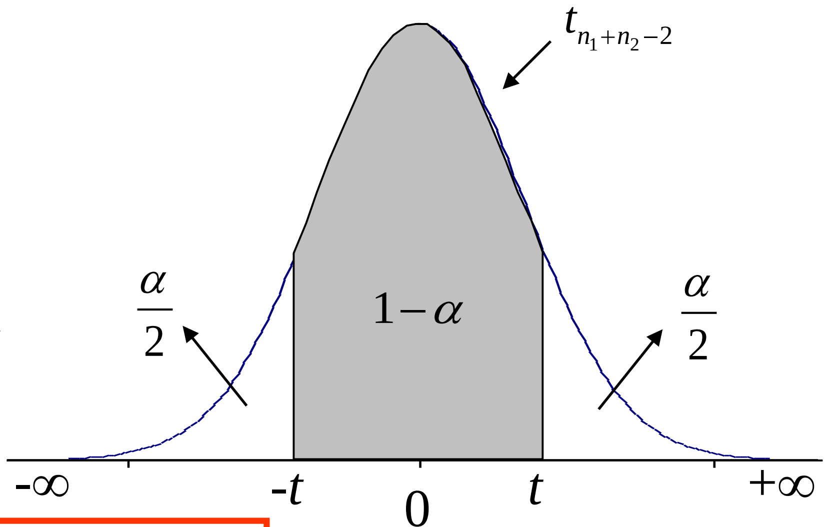
Intervalo de Confiança para μ_1 - μ_2

$X_1 \sim N(\mu_1, \sigma_1^2)$ $\quad X_2 \sim N(\mu_2, \sigma_2^2)$ $\quad \mu_i$ e σ_i^2 desconhecidas

$$\frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1) \qquad \frac{(n_1-1)s_1^2}{\sigma_1^2} + \frac{(n_2-1)s_2^2}{\sigma_2^2} \sim \chi^2_{n_1+n_2-2}$$

$$\frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \sim t_{n_1+n_2-2}$$

$$P\left(-t < \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} < t\right) = 1-\alpha$$



$$\boxed{P\left(\bar{X}_1 - \bar{X}_2 - t\sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}} < \mu_1 - \mu_2 < \cdots\right) = 1-\alpha}$$

$$P(-t < T < t) = 1-\alpha$$

IC para μ_1 - μ_2 (atenção: $\sigma_1^2 = \sigma_2^2$ $\quad t$ homocedástico)

Intervalo de Confiança para μ_1 - μ_2

$X_1 \sim N(\mu_1, \sigma_1^2)$ $\quad X_2 \sim N(\mu_2, \sigma_2^2)$ $\quad \mu_i$ e σ_i^2 desconhecidas

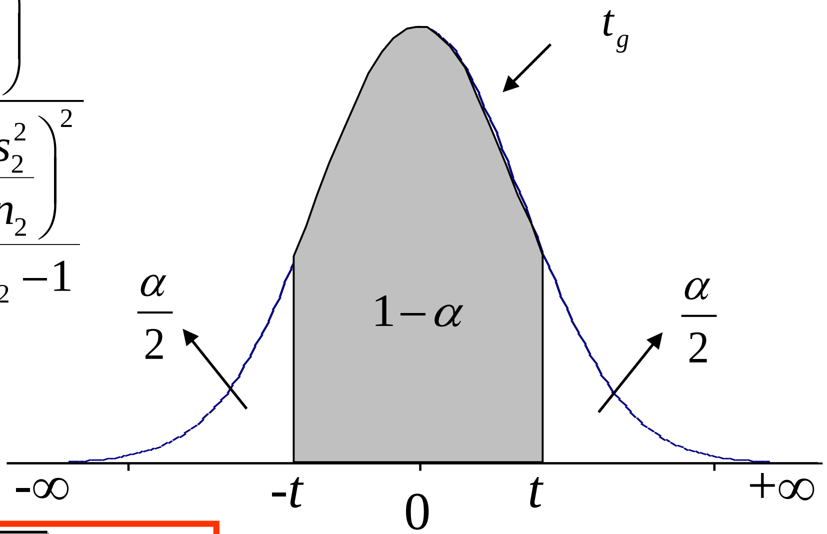
$$\frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1)$$

(considerando $\sigma_1^2 \neq \sigma_2^2$)

$$\frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \sim t_g$$

$$g \approx \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2 - 1}}$$

$$P(-t < \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} < t) = 1 - \alpha$$



$$\boxed{P(\bar{X}_1 - \bar{X}_2 - t\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} < \mu_1 - \mu_2 < \bar{X}_1 - \bar{X}_2 + t\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}) = 1 - \alpha}$$

$$P(-t < T < t) = 1 - \alpha$$

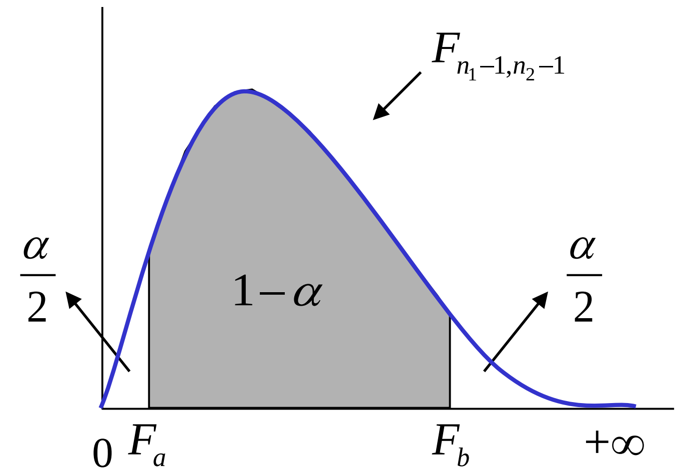
IC para μ_1 - μ_2 (atenção: $\sigma_1^2 \neq \sigma_2^2$ $\quad t$ heterocedástico)

Intervalo de Confiança para σ_1^2 / σ_2^2

$X_1 \sim N(\mu_1, \sigma_1^2)$ $X_2 \sim N(\mu_2, \sigma_2^2)$ μ_i e σ_i^2 desconhecidas

$$\frac{s_1^2 \sigma_2^2}{s_2^2 \sigma_1^2} \sim F_{n_1-1, n_2-1}$$

F



$$P(F_a < \frac{s_1^2 \sigma_2^2}{s_2^2 \sigma_1^2} < F_b) = 1-\alpha$$

$$P(F_a < F < F_b) = 1-\alpha$$

$$P(\frac{1}{F_b} < \frac{s_2^2 \sigma_1^2}{s_1^2 \sigma_2^2} < \frac{1}{F_a}) = 1-\alpha$$

$$\boxed{P(\frac{s_1^2}{s_2^2}\frac{1}{F_b} < \frac{\sigma_1^2}{\sigma_2^2} < \frac{s_1^2}{s_2^2}\frac{1}{F_a}) = 1-\alpha}$$

IC para σ_1^2 / σ_2^2

OBS: por exemplo, se $1 - \alpha = 95\%$

$$F_a \rightarrow F_{0,975;n_1-1;n_2-1} \quad \Rightarrow \quad \frac{1}{F_a} \rightarrow F_{0,025;n_2-1;n_1-1}$$

$$F_b \rightarrow F_{0,025;n_1-1;n_2-1}$$

IC para μ_1 - μ_2 e σ_1^2 / σ_2^2

Exemplo: duas v.a. quaisquer têm distribuições desconhecidas com médias e variâncias também desconhecidas. Retira-se uma amostra de cada população e calcula-se a média e a variância para cada amostra. Construa um IC de 95% para a razão entre variâncias e para a diferença entre médias supondo que

$$n_1 = 26 \quad \bar{X}_1 = 10,3 \quad s_1^2 = 2,34$$

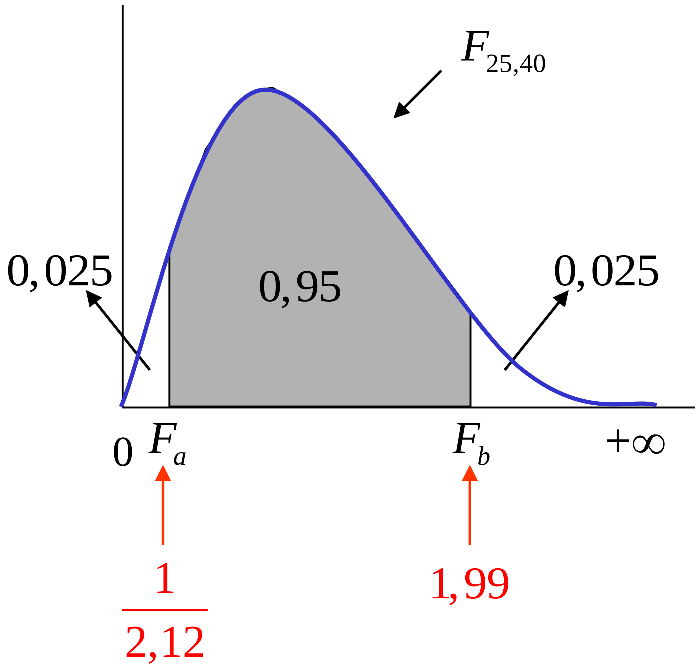
$$n_2 = 41 \quad \bar{X}_2 = 15,7 \quad s_2^2 = 1,91$$

IC para σ_1^2 / σ_2^2

$$P\left(\frac{s_1^2}{s_2^2}\frac{1}{F_b} < \frac{\sigma_1^2}{\sigma_2^2} < \frac{s_1^2}{s_2^2}\frac{1}{F_a}\right) = 0,95$$

$$P\left(\frac{2,34}{1,91}\frac{1}{1,99} < \frac{\sigma_1^2}{\sigma_2^2} < \frac{2,34}{1,91}2,12\right) = 0,95$$

$$\boxed{P\left(0,615 < \frac{\sigma_1^2}{\sigma_2^2} < 2,597\right) = 0,95}$$



As variâncias podem ser iguais?

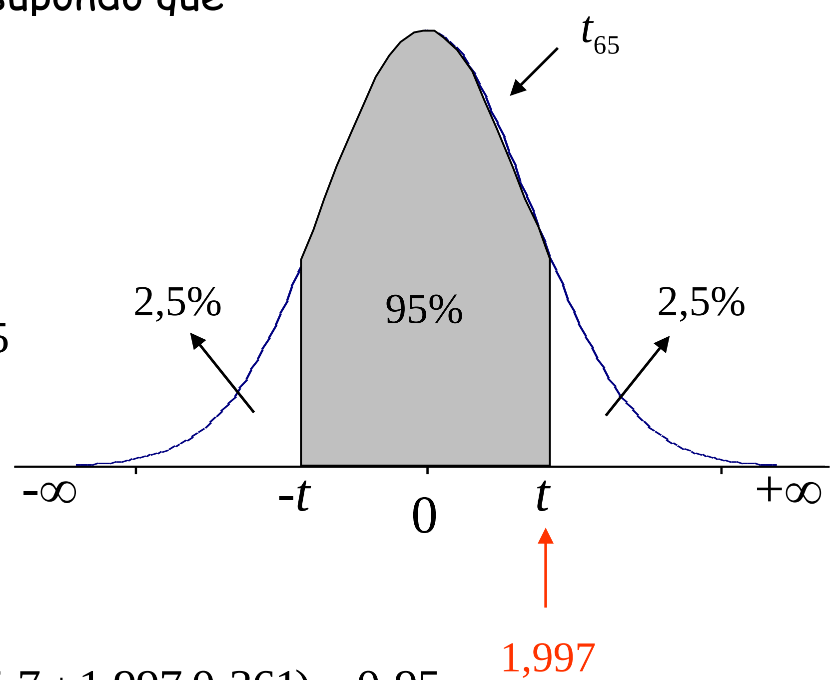
R: não há razão para discordar disso. $\Rightarrow$ pode-se fazer o IC para μ_1 - μ_2 (homocedástico)

IC para μ_1 - μ_2 e σ_1^2/σ_2^2

Exemplo: duas v.a. quaisquer têm distribuições desconhecidas com médias e variâncias também desconhecidas. Retira-se uma amostra de cada população e calcula-se a média e a variância para cada amostra. Construa um IC de 95% para a razão entre variâncias e para a diferença entre médias supondo que

$$n_1 = 26 \quad \bar{X}_1 = 10{,}3 \quad s_1^2 = 2{,}34$$

$$n_2 = 41 \quad \bar{X}_2 = 15{,}7 \quad s_2^2 = 1{,}91$$



IC para μ_1 - μ_2

$$P(\bar{X}_1 - \bar{X}_2 - tK < \mu_1 - \mu_2 < \bar{X}_1 - \bar{X}_2 + tK) = 0{,}95$$

$$K = \sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$$

$$K = \sqrt{\frac{25\;2{,}34 + 40\;1{,}91}{65}}\sqrt{\frac{1}{26}+\frac{1}{41}} = 0{,}361$$

$$P(10{,}3 - 15{,}7 - 1{,}997\;0{,}361 < \mu_1 - \mu_2 < 10{,}3 - 15{,}7 + 1{,}997\;0{,}361) = 0{,}95$$

$$\boxed{P(-6{,}121 < \mu_1 - \mu_2 < -4{,}679) = 0{,}95}$$

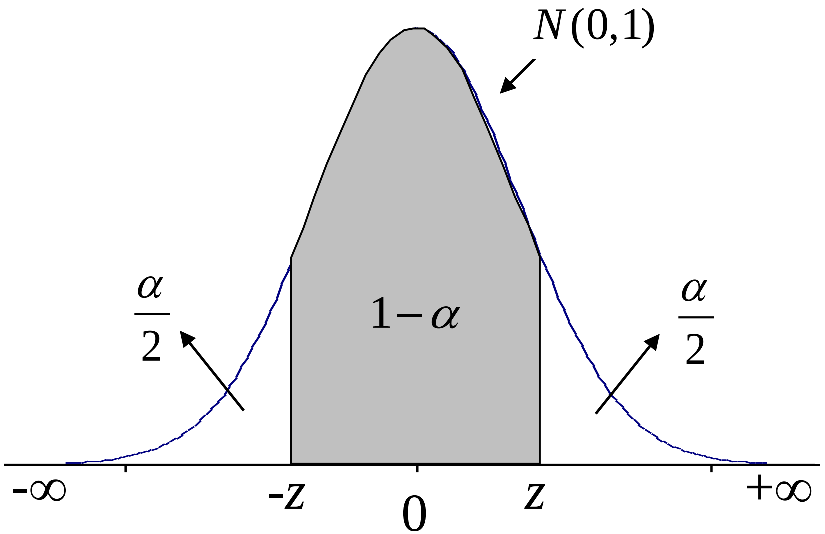
$\mu_1 = \mu_2$? $\quad \Rightarrow \mu_1 < \mu_2$

Intervalo de Confiança para $p_1 - p_2$

$$\hat{p}_1 \sim N(p_1, \frac{p_1 q_1}{n_1}) \qquad \hat{p}_2 \sim N(p_2, \frac{p_2 q_2}{n_2}) \qquad \hat{p}_1 - \hat{p}_2 \sim N(p_1 - p_2, \frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2})$$

$$\frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}} \sim N(0,1)$$

$$P(-z < \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}} < z) = 1 - \alpha$$



$$P(-z < Z < z) = 1 - \alpha$$

$$\boxed{P(\hat{p}_1 - \hat{p}_2 - z\sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}} < p_1 - p_2 < \hat{p}_1 - \hat{p}_2 + z\sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}}) = 1 - \alpha}$$

IC para $p_1 - p_2$

Intervalos de Confiança (Resumo)

para μ $\begin{cases} N(0,1) & \text{se } \sigma^2 \text{ é conhecida} \\ t_{n-1} & \text{se } \sigma^2 \text{ é desconhecida} \end{cases}$

para σ^2 $\left\{ \chi^2_{n-1} \right.$

para μ_1 - μ_2 $\begin{cases} N(0,1) & \text{se } \sigma_1^2 \text{ e } \sigma_2^2 \text{ são conhecidas} \\ t_{n_1+n_2-2} & \text{se } \sigma_1^2 \text{ e } \sigma_2^2 \text{ são desconhecidas, mas } \sigma_1^2 = \sigma_2^2 \\ t_g & \text{se } \sigma_1^2 \text{ e } \sigma_2^2 \text{ são desconhecidas, mas } \sigma_1^2 \neq \sigma_2^2 \end{cases}$

para $\dfrac{\sigma_1^2}{\sigma_2^2}$ $\left\{ F_{n_1-1,n_2-1} \right.$

para p $\left\{ N(0,1) \right.$

para $p_1 - p_2$ $\left\{ N(0,1) \right.$