

# SME0816 - Planejamento de Experimentos I

Comparações simples

---

Profa. Cibele Russo

(Referências: Montgomery (2012), Notas de aula de Roseli Leandro; Clarice Demétrio; Marinho Andrade)

- **Unidade experimental:** é a unidade física ou biológica para conduzir o experimento, também denominada de **parcela**. Ex.: uma pessoa, uma peça, uma planta, um animal, etc.
- **Variável:** é a condição ou característica medida ou observada no experimento.
- **Tratamento:** é o elemento que está em teste no experimento.
  - Ex.: um método de ensino, um fertilizante, uma droga terapêutica, um tipo de ração, uma liga metálica, etc.

# Tipos de Tratamentos

- **Qualitativos:** são aqueles que possuem naturezas intrinsecamente diferentes.
  - Ex.: comparar dois tipos de drogas terapêuticas para dor de cabeça (Paracetamol versus Dipirona Sódica),
  - Ex.: comparar três tipos de fertilizantes (diferentes marcas), etc.
  - Ex.: comparar quatro tipos de ligas metálicas usada na fabricação de uma peça
- **Quantitativos:** são aqueles que se distinguem pela quantidade (dose) que está sendo utilizada no experimento;
  - percentual de determinado composto em uma substância ou material; etc.

## **Grupo controle:**

- Em um estudo em que se deseja estudar o efeito de um ou mais tratamentos em relação a situação sem nenhum tratamento, o grupo controle é o grupo de unidades experimentais que não recebe tratamento.

**Comparar a média de dois níveis  $i = 1, 2$  de um fator  $\mu$ .**

**Modelo estatístico:**

$$Y_{ij} = \mu_i + \varepsilon_{ij}$$

em que

- $Y_{ij}$ , a variável resposta na  $j$ -ésima unidade no nível  $i$ .
- $\mu_i$ , média (populacional) do fator no nível  $i$ ;  $i = 1, 2$ .
- $\varepsilon_{ij}$ , erro aleatório no nível  $i$  da unidade  $j$ ,  $j = 1, \dots, n_i$ .

Comparar a média de dois níveis  $i = 1, 2$  de um fator  $\mu$ .

Modelo estatístico:

$$Y_{ij} = \mu_i + \varepsilon_{ij}$$

**Conjecturas sobre o modelo**

- Podemos considerar  $\varepsilon_{ij} \sim N(0, \sigma^2)$ ?
- Se  $\varepsilon_{ij} \sim N(0, \sigma_i^2)$  então  $Y_{ij} \sim N(\mu_i, \sigma_i^2)$
- As variâncias  $\sigma_i^2$ ,  $i = 1, 2$ , são conhecidas ou desconhecidas?
- As variâncias  $\sigma_i^2$ ,  $i = 1, 2$ , são iguais ou diferentes?

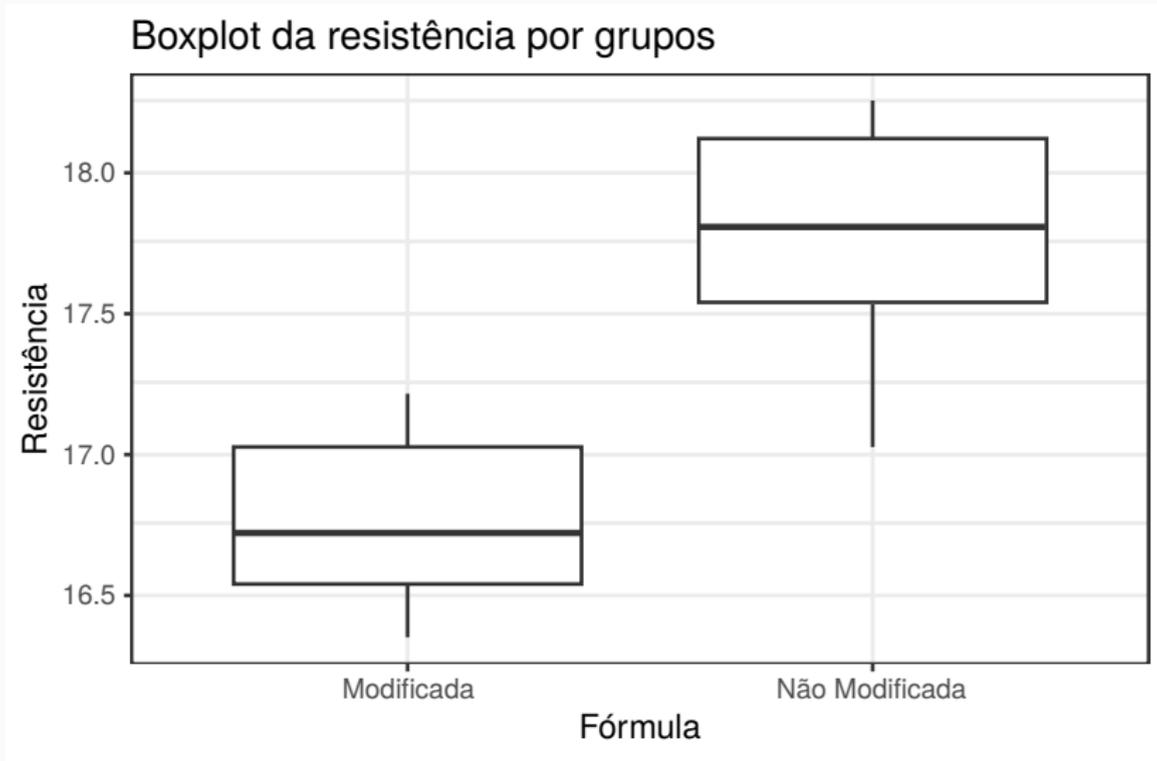
## Exemplo 1

A resistência à tração da argamassa de cimento Portland é uma característica importante do produto. Um engenheiro está interessado em comparar a resistência de uma fórmula modificada, à qual se acrescentou emulsão de polímero de látex durante a mistura, contra a resistência da argamassa feita com a fórmula não modificada. O experimentador preparou 10 amostras de cada tipo e mediu a resistência à tração:

- Modificada: 16.85, 16.40, 17.21, 16.35, 16.52, 17.04, 16.96, 17.15, 16.59, 16.57
- Não modificada: 17.50, 17.63, 18.25, 18.00, 17.86, 17.75, 18.22, 17.02, 17.08, 18.15

Formule um modelo para avaliar a comparação das médias das resistências nos dois grupos.

# Exemplo 1



## Modelo estatístico:

$$Y_{ij} = \mu_i + \varepsilon_{ij}$$

- $Y_{ij}$ , a variável resposta no nível  $i$  da unidade  $j$ .
- $\mu_i$ , média do fator no nível  $i$ ;  $i = 1, 2$ .
- $\varepsilon_{ij}$ , erro aleatório no nível  $i$  da unidade  $j$ ;  $j = 1, \dots, n$ .

### Suposição do modelo:

- A componente aleatória tem média zero e variância  $\sigma_i^2$ .
- A componente aleatória é não correlacionada.
- A componente aleatória tem distribuição normal.

Portanto,  $\varepsilon_{ij} \stackrel{ind}{\sim} N(0, \sigma_i^2)$ , para  $i = 1, 2$ .

O que implica que  $Y_{ij} \stackrel{ind}{\sim} N(\mu_i, \sigma_i^2)$ , para  $i = 1, 2$ .

Sejam  $Y_{11}, \dots, Y_{1n_1}$  as  $n_1$  observações (resistências) com primeiro nível do fator (argamassa modificada) e  $Y_{21}, \dots, Y_{2n_2}$  as  $n_2$  observações no segundo nível do fator (argamassa não modificada).

- $Y_{ij} = \mu_i + \epsilon_{ij}$ ,  $i = 1, 2$ ,  $j = 1, \dots, n_i$
- $Y_{ij}$ : é a  $j$ -ésima observação do nível  $i$  do fator
- $\mu_i$ : média (populacional) da variável resposta no nível  $i$  do fator
- $\epsilon_{ij}$ : v. a. associada a  $ij$  - observação (componente aleatória)

## Populações independentes com distribuição normal

- Amostra da população 1:

$$Y_{11}, \dots, Y_{1n_1} \stackrel{i.i.d.}{\sim} N(\mu_1, \sigma_1^2) \Rightarrow \bar{Y}_1 \sim N\left(\mu_1, \frac{\sigma_1^2}{n_1}\right)$$

- Amostra da população 2:

$$Y_{21}, \dots, Y_{2n_2} \stackrel{i.i.d.}{\sim} N(\mu_2, \sigma_2^2) \Rightarrow \bar{Y}_2 \sim N\left(\mu_2, \frac{\sigma_2^2}{n_2}\right)$$

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

Obs. Se a distribuição de  $Y_1$  e  $Y_2$  não for normal, os resultados são válidos aproximadamente, usando o Teorema do Limite Central.

## Testes de hipóteses sobre $\mu_1 - \mu_2$

As duas populações são independentes.

(i) Formulação das hipóteses:

- $H_0 : \mu_1 - \mu_2 = \Delta_0$  contra  $H_1 : \mu_1 - \mu_2 < \Delta_0$  (à esquerda)
- $H_0 : \mu_1 - \mu_2 = \Delta_0$  contra  $H_1 : \mu_1 - \mu_2 > \Delta_0$  (à direita)
- $H_0 : \mu_1 - \mu_2 = \Delta_0$  contra  $H_1 : \mu_1 - \mu_2 \neq \Delta_0$  (bilateral)

onde  $\Delta_0$  é uma constante conhecida (valor de teste).

$\Delta_0 = 0$  corresponde à igualdade das duas médias.

## Tipos de erros

**Erro tipo I:** Rejeitar  $H_0$  quando  $H_0$  é verdadeira.

**Erro tipo II:** Não rejeitar  $H_0$  quando  $H_0$  é falsa.

**Quadro resumo:**

	Decisão	
Situação real	Rejeito $H_0$	Não rejeito $H_0$
$H_0$ é Verdadeira	Erro Tipo I	Acerto
$H_0$ é Falsa	Acerto	Erro Tipo II

## Nível de significância e poder do teste:

$$\begin{aligned}P(\text{Erro tipo I}) &= \alpha \quad (\text{nível de significância}) \\ &= P(\text{Rejeitar } H_0 \mid H_0 \text{ verdadeira})\end{aligned}$$

$$\begin{aligned}P(\text{Erro tipo II}) &= \beta = P(\text{Não rejeitar } H_0 \mid H_0 \text{ falsa}) \\ &= P(\text{Não rejeitar } H_0 \mid H_1 \text{ verdadeira})\end{aligned}$$

$$1 - \beta = P(\text{Rejeitar } H_0 \mid H_0 \text{ é falsa}): \text{ poder do teste}$$

## Testes de hipóteses sobre $\mu_1 - \mu_2$ :

(ii) Estatística de teste

(a)  $\sigma_1^2$  e  $\sigma_2^2$  **conhecidas**:

$$Z = \frac{\bar{Y}_1 - \bar{Y}_2 - \Delta_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0, 1) \text{ sob } H_0$$

## Testes de hipóteses sobre $\mu_1 - \mu_2$ :

(ii) Estatística de teste

(b)  $\sigma_1 = \sigma_2 = \sigma$  **desconhecida**:

$$T = \frac{\bar{Y}_1 - \bar{Y}_2 - \Delta_0}{\sqrt{\frac{S_p^2}{n_1} + \frac{S_p^2}{n_2}}} \underset{\text{sob } H_0}{\sim} t_{n_1+n_2-2}$$

onde  $S_p^2$  é a **variância combinada (pooled variance)**:

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{(n_1 - 1) + (n_2 - 1)}$$

e  $S_i^2$  é a variância amostral para a amostra  $i = 1, 2$ :

$$S_i^2 = \frac{\sum_{j=1}^{n_i} (Y_{ij} - \bar{Y}_i)^2}{n_i - 1}, \quad i = 1, 2.$$

## Testes de hipóteses sobre $\mu_1 - \mu_2$ :

(ii) Estatística de teste

(c)  $\sigma_1 \neq \sigma_2$ , ambas desconhecidas:

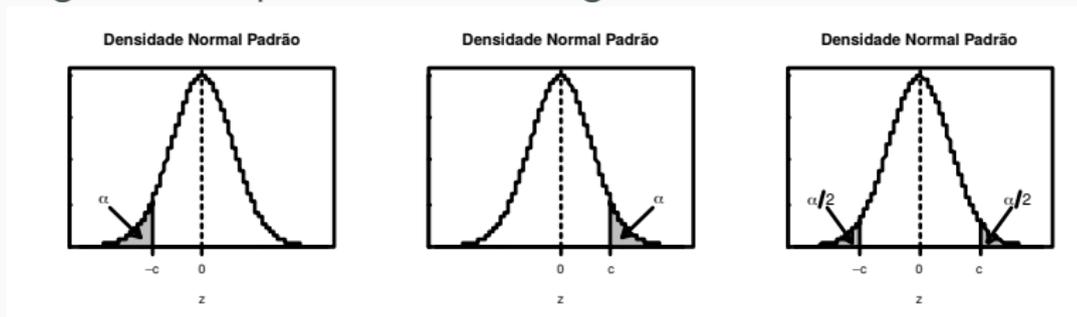
$$T = \frac{\bar{Y}_1 - \bar{Y}_2 - \Delta_0}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \underset{\text{sob } H_0}{\sim} t_g \text{ aproximadamente}$$

em que

$$g = \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{(S_1^2/n_1)^2}{(n_1-1)} + \frac{(S_2^2/n_2)^2}{(n_2-1)}}$$

## Testes de hipóteses sobre $\mu_1 - \mu_2$

(iii) Região crítica para um nível de significância  $\alpha$  escolhido:



$$H_1 : \mu_1 - \mu_2 < \Delta_0$$

$$R_C(Z) = \{Z < -c\}$$

$$R_C(T) = \{T < -c\}$$

$$H_1 : \mu_1 - \mu_2 > \Delta_0$$

$$R_C(Z) = \{Z > c\}$$

$$R_C(T) = \{T > c\}$$

$$H_1 : \mu_1 - \mu_2 \neq \Delta_0$$

$$R_C(Z) = \{|Z| > c\}$$

$$R_C(T) = \{|T| > c\}$$

(iv) Se  $Z \in R_C$  ou  $T \in R_C$ , rejeita-se  $H_0$ ; caso contrário, não se rejeita  $H_0$ .

Obs. Nas regiões críticas com  $Z$  e  $T$ , o valor de  $c$  não é o mesmo.

## IC para $\mu_1 - \mu_2$

1. Estimador pontual para  $\mu_1 - \mu_2$ :  $\bar{Y}_1 - \bar{Y}_2$ . Podemos provar que um intervalo de confiança (IC) de  $100(1 - \alpha)\%$  para  $\mu_1 - \mu_2$  é dado por

$$IC = [L; U] = [\bar{Y}_1 - \bar{Y}_2 - E; \bar{Y}_1 - \bar{Y}_2 + E],$$

sendo que  $E$  é o erro máximo do IC.

2.  $\sigma_1^2$  e  $\sigma_2^2$  conhecidas:  $E = z_{\alpha/2} \times \frac{\sigma}{\sqrt{n}}$ .
3.  $\sigma_1 = \sigma_2 = \sigma$  desconhecida:  
$$E = t_{\alpha/2, n+m-2} \times S_p \times \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}.$$
4.  $\sigma_1 \neq \sigma_2$ , ambas desconhecidas:  $E = t_{\alpha/2, g} \times \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$ .

## IC para $\mu_1 - \mu_2$ e Testes de Hipóteses

O teste da hipótese

$$H_0 : \mu_1 - \mu_2 = \Delta_0 \text{ contra}$$

$$H_1 : \mu_1 - \mu_2 \neq \Delta_0$$

a um nível de significância  $\alpha$  pode ser efetuado utilizando um intervalo de confiança (IC) com coeficiente de confiança igual a  $1 - \alpha$ .

Construímos o IC de  $100(1 - \alpha)\%$  para  $\mu_1 - \mu_2$ , dado por

$$IC = [L; U] = [\bar{Y}_1 - \bar{Y}_2 - E; \bar{Y}_1 - \bar{Y}_2 + E],$$

sendo que no cálculo do erro máximo ( $E$ ) utilizamos anteriormente.

Se  $\Delta_0 \notin IC$ , rejeitamos  $H_0$ ; caso contrário, não rejeitamos  $H_0$ .

## Exemplo 2

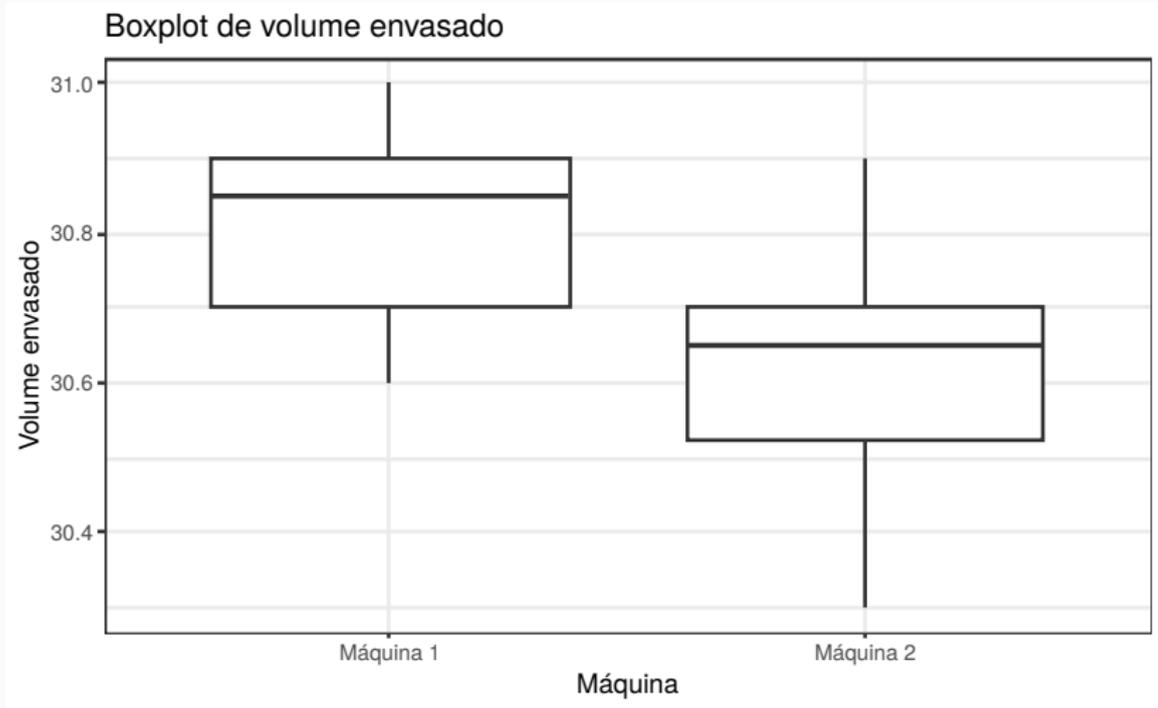
Duas máquinas são utilizadas para envasar um líquido em frascos de plástico. Com o objetivo de verificar se há diferença entre os volumes médios envasados, duas amostras de 12 e 10 frascos foram selecionadas. Os volumes (em ml) foram medidos resultando nos seguintes valores:

- Máquina 1: 30.9, 30.9, 30.8, 30.7, 30.9, 30.6, 30.8, 30.9, 30.7, 30.9, 30.7, 31.0;
- Máquina 2: 30.8, 30.9, 30.7, 30.5, 30.5, 30.6, 30.7, 30.3, 30.6, 30.7.

Utilizando os dados coletados, qual o resultado da verificação?

Adote  $\alpha = 5\%$ .

## Exemplo 2



Exercício: Avalie a normalidade das amostras.

## Exemplo 2

**Solução.** O problema envolve a comparação de duas médias de variáveis contínuas. Definimos  $Y_1$  e  $Y_2$  como sendo os volumes envasados pelas máquinas 1 e 2, tais que  $E(Y_1) = \mu_1$ ,  $\text{var}(Y_1) = \sigma_1^2$ ,  $E(Y_2) = \mu_2$  e  $\text{var}(Y_2) = \sigma_2^2$ .

**Hipóteses:**

$$H_0 : \mu_1 = \mu_2 \text{ contra}$$

$$H_1 : \mu_1 \neq \mu_2 \text{ (ou seja, } \Delta_0 = 0)$$

## Exemplo 2

- Estatística de teste (variâncias diferentes e desconhecidas):

$$T = \frac{\bar{Y}_1 - \bar{Y}_2 - 0}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \quad \text{sob } H_0 \sim t_g, \text{ aproximadamente,}$$

em que

$$g = \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{(S_1^2/n_1)^2}{(n_1-1)} + \frac{(S_2^2/n_2)^2}{(n_2-1)}}$$

## Exemplo 2

- Utilizando os dados coletados, calculamos:

$$\bar{Y}_1 = \frac{\sum_{j=1}^{12} Y_{1j}}{12} = \frac{369,8}{12} = 30,82 \text{ ml},$$

$$\bar{Y}_2 = \frac{\sum_{j=1}^{10} Y_{2j}}{10} = \frac{306,3}{10} = 30,63 \text{ ml},$$

$$S_1^2 = \frac{\sum_{j=1}^{12} (Y_{1j} - \bar{Y}_1)^2}{12 - 1} = \frac{\sum_{j=1}^{12} (Y_{1j} - 30,82)^2}{0,0142 \text{ ml}^2},$$

$$S_2^2 = \frac{\sum_{j=1}^{10} (Y_{2j} - \bar{Y}_2)^2}{10 - 1} = \frac{\sum_{j=1}^{10} (Y_{2j} - 30,63)^2}{0,0290 \text{ ml}^2},$$

$$g = \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{(S_1^2/n_1)^2}{(n_1-1)} + \frac{(S_2^2/n_2)^2}{(n_2-1)}} = 16.$$

## Exemplo 2

- A região crítica para  $\alpha = 0,05$  é obtida consultando a tabela da distribuição t-Student com  $g = 16$  g.l.:

$$R_C = \{|T| > 2,120\}.$$

- Calculamos:

$$T = \frac{\bar{X} - \bar{Y} - 0}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} = \frac{30,82 - 30,63}{\sqrt{\frac{0,0142}{12} + \frac{0,0290}{10}}} = 2,972.$$

- Como  $|T| = 2,972 \in R_C$ , rejeitamos  $H_0$ .
- **Conclusão:** De acordo com os dados coletados e com um nível de significância de 5%, verificamos que há diferença entre os volumes médios envasados pelas duas máquinas.

## Nível descritivo do teste (valor-p)

No exemplo 2 a região crítica é da forma  $Rc = |T| > c$ , sendo que  $T$  tem distribuição t-Student com 16 g.l., sob  $H_0$  verdadeira.

Com os dados coletados calculamos  $|T| = 2,972$ . Se adotarmos  $c = 2,972$  obtemos  $Rc = \{|T| > 2,972\}$  e o nível de significância do teste correspondente é  $\alpha = P(|T| > 2,972) = 0,0090$ .

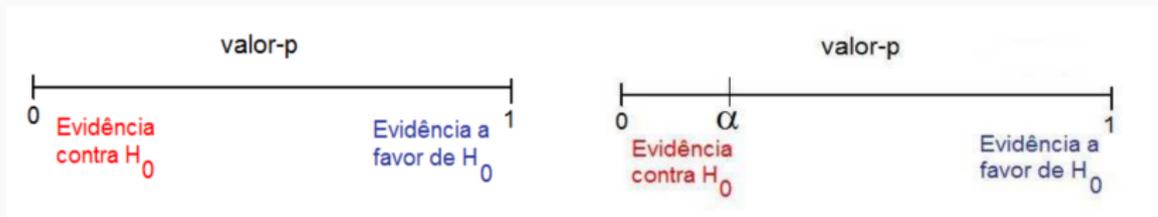
Uma vez que para qualquer escolha tal que  $0 < c < 2,972$ , rejeitaremos  $H_0$ , concluimos que testes com  $\alpha > 0,0090$  levam à rejeição de  $H_0$  (se  $c < 2,972$ , então  $\alpha > 0,0090$ ).

$\alpha = 0,0090$  é chamado de probabilidade de significância, nível descritivo, valor-p (p-value) ou p.

## Nível descritivo do teste (valor-p)

- Como o nível de significância é a probabilidade de um erro tipo I (rejeição de  $H_0$  verdadeira), quanto menor for valor-p, mais fortemente rejeitamos  $H_0$ .
- Quanto menor for valor-p, mais evidência contra  $H_0$  (e vice-versa).
- Caso a variância fosse conhecida e  $Rc = \{|Z| > c\}$ , sendo que  $Z$  tem distribuição  $N(0, 1)$ , se  $H_0$  for verdadeira. Com os dados coletados calculamos  $|Z| = 1,90$ . Neste caso,  $\text{valor-p} = P(|Z| > 1,90) = 2 \times P(Z < -1,90) = 2 \times 0,0287 = 0,0574$  (mudança na conclusão do teste).

# Nível descritivo do teste (valor-p)



## Nível descritivo do teste (valor-p)

- (...) Experimental postural force directions clustered near both maxima; there were no medial postural forces near the absolute minimum. However, the medians of the anterior and posterior postural force direction histograms in the right hindlimb were rotated counter-clockwise from the FFS maxima ( $p < 0.05$ ; Wilcoxon signed-rank test).. *Journal of Biomechanics* 40(10), 2254-2260, 2007.
- The Biomechanics Concept Inventory (BCI) was given to 53 kinesiology majors before and after an introductory biomechanics class together with a survey of student behaviours to determine factors that assisted in learning. Analysis of scores from 49 students showed significant ( $p < 0.001$ ).

## Exemplo 3

Em um teste de dureza uma esfera de aço é pressionada contra a superfície de um bloco de material a uma carga padrão. Mede-se o diâmetro (em mm) da cavidade produzida, que está relacionado à dureza do material da superfície.

Na realização do teste duas esferas (A e B) estão disponíveis. Suspeita-se que a esfera A gera cavidades com diâmetro médio com diferença superior a 0,2 mm em relação à esfera B. As duas esferas foram utilizadas em 10 blocos ( $n = 10$ ) obtendo-se os dados a seguir:

Bloco	Esfera A	Esfera B	Diferença
1	7.5	5.2	2.3
2	4.6	4.1	0.5
3	5.7	4.3	1.4
4	4.3	4.7	-0.4
5	5.8	3.2	2.6
6	3.2	4.9	-1.7
7	6.1	5.2	0.9
8	5.6	4.4	1.2
9	3.4	5.7	-2.3
10	6.5	6.0	0.5

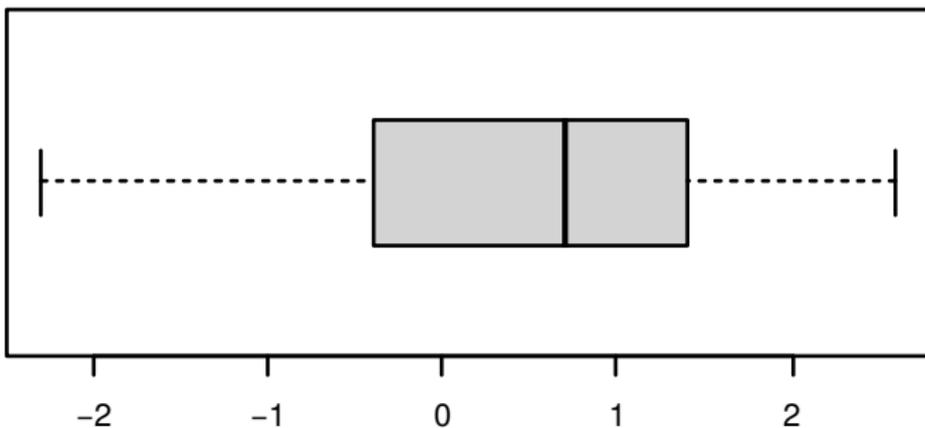
O que os dados permitem concluir sobre a suspeita formulada?

Adote  $\alpha = 5\%$ .

## Exemplo 3

Análise exploratória

**Boxplot da diferença**



## Modelo estatístico

O modelo estatístico é representado por:

$$Y_{ij} = \mu_i + \beta_j + \epsilon_{ij}$$

Onde:

- $Y_{ij}$ : diâmetro do  $j$  bloco na esfera  $i$
- $\mu_i$ : diâmetro médio da esfera  $i$
- $\beta_j$ : efeito do  $j$ -ésimo bloco
- $\epsilon_{ij}$ : erro aleatório associado a  $ij$  - observação (v.a.)

Suposição do modelo:

- A componente aleatória tem média zero e variância  $\sigma_i^2$

## Populações dependentes com distribuição normal

- Pares:  $(Y_{11}, Y_{21}), (Y_{12}, Y_{22}), \dots, (Y_{1n}, Y_{2n})$ .
- Diferença:  $D_j = Y_{1j} - Y_{2j}$  com  
 $\mu_D = E(Y_{1j} - Y_{2j}) = E(Y_{1j}) - E(Y_{2j}) = \mu_1 - \mu_2$  e  
 $var(D_j) = \sigma_D^2$ .
- Calculamos  $D_j = Y_{1j} - Y_{2j}$ :

$$\bar{D} = \sum_{j=1}^n \frac{D_j}{n} = \bar{Y}_1 - \bar{Y}_2 \text{ e } s_D = \sqrt{\frac{\sum_{j=1}^n (D_j - \bar{D})^2}{(n-1)}}$$

- Resultado  $\bar{D} \sim N\left(\mu_D, \frac{\sigma_D^2}{n}\right)$ .
- Obs.: Se a distribuição de  $\bar{D}$  não for normal, o resultado é válido aproximadamente.

(i) **Formulação das hipóteses:**

$$H_0 : \mu_D = \Delta_0$$

$$H_1 : \mu_D < \Delta_0 \quad (\text{À esquerda})$$

$$H_1 : \mu_D > \Delta_0 \quad (\text{À direita})$$

$$H_1 : \mu_D \neq \Delta_0 \quad (\text{Bilateral})$$

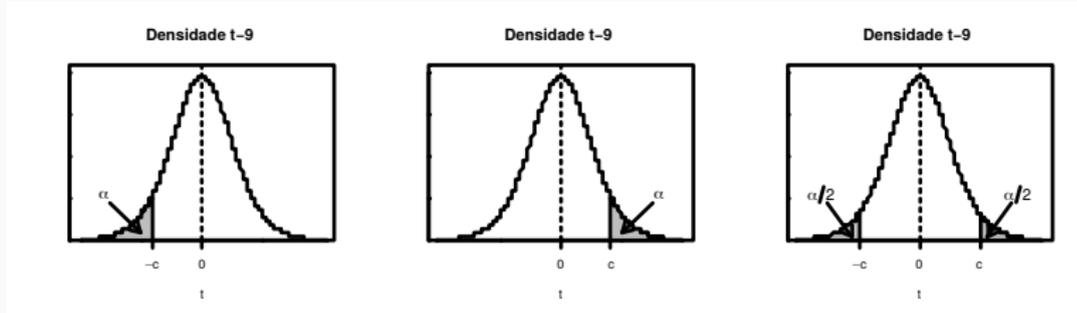
sendo que  $\Delta_0$  é uma constante conhecida (valor de teste).  $\Delta_0 = 0$  corresponde à igualdade das duas médias ( $\mu_D = \mu_1 - \mu_2$ ).

## Testes de hipóteses sobre $\mu_1 - \mu_2$

(ii) Estatística de teste:

$$T = \frac{n(\bar{D} - \Delta_0)}{s_D} \stackrel{H_0}{\sim} t_{n-1}$$

(iii) Região crítica para um nível de significância  $\alpha$  escolhido:



$$\begin{array}{lll} H_1 : \mu_1 - \mu_2 < \Delta_0 & H_1 : \mu_1 - \mu_2 > \Delta_0 & H_1 : \mu_1 - \mu_2 \neq \Delta_0 \\ R_C(T) = \{T < -c\} & R_C(T) = \{T > c\} & R_C(T) = \{|T| > c\} \end{array}$$

(iv) Se  $T \in R_C$ , rejeita-se  $H_0$ ; caso contrário, não se rejeita  $H_0$ .

## Intervalo de Confiança para $\mu_1 - \mu_2$

**Estimador pontual para  $\mu_D = \mu_1 - \mu_2$ :**

$$\bar{D} = \bar{Y}_1 - \bar{Y}_2$$

Podemos provar que um intervalo de confiança (IC) de  $100(1 - \alpha)\%$  para  $\mu_1 - \mu_2$  é dado por:

$$IC = [L, U] = [\bar{D} - E, \bar{D} + E]$$

onde  $E$  é o erro máximo do IC:

$$E = t_{\alpha/2, n-1} \times \frac{s_D}{\sqrt{n}}$$

## Intervalo de Confiança para $\mu_1 - \mu_2$ e Testes de Hipóteses

O teste da hipótese

$$H_0 : \mu_D = \Delta_0 \text{ contra}$$

$$H_1 : \mu_D \neq \Delta_0$$

a um nível de significância  $\alpha$  pode ser efetuado utilizando um IC com coeficiente de confiança igual a  $1 - \alpha$ .

Construímos o IC de  $100(1 - \alpha)\%$  para  $\mu_D = \mu_1 - \mu_2$ , dado por

$$[L, U] = [\bar{D} - E, \bar{D} + E]$$

sendo que no cálculo do erro máximo ( $E$ ) utilizamos a lâmina 11.

Se  $\Delta_0 \notin IC$ , rejeitamos  $H_0$ ; caso contrário, não rejeitamos  $H_0$ .

### Exemplo 3

**Solução.** Problema envolve duas médias de variáveis contínuas. Definimos  $Y_1$  e  $Y_2$  como sendo os diâmetros das crateras produzidas pelas esferas A e B, tais que  $E(Y_1) = \mu_1$  e  $E(Y_2) = \mu_2$ . Como os dados são pareados, utilizamos  $\bar{D} = \bar{Y}_1 - \bar{Y}_2$  com  $\mu_D = E(\bar{Y}_1 - \bar{Y}_2) = \mu_1 - \mu_2$  e  $var(D) = \sigma_D^2$ .

**Hipóteses:**  $H_0 : \mu_D = 0,2$  contra  $H_1 : \mu_D > 0,2$  (ou seja,  $\Delta_0 = 0,2$ ).

**Estatística de teste:**

$$T = \frac{n(\bar{D} - \Delta_0)}{s_D}$$

A região crítica para  $\alpha = 0,05$  é obtida consultando a tabela da distribuição t-Student com 9 g.l. ( $= n - 1$ ) e  $p = 10\%$ :

$$Rc = \{T > 1,833\}$$

## Exemplo 3

Calculamos:

$$\bar{D} = \frac{\sum_{j=1}^n D_j}{n} = \frac{5,0}{10} = 0,5 \text{ mm}$$

$$S^2 = \frac{\sum_{j=1}^n (D_j - \bar{D})^2}{n - 1} = \frac{\sum_{j=1}^n (D_j - 0,5)^2}{9} = 2,51 \text{ mm}^2$$

$$T = \frac{n(\bar{D} - \Delta_0)}{s_D} = \frac{10(0,5 - 0,2)}{\sqrt{2,51}} = 0,599$$

Como  $T = 0,599 \notin R_c$ , não rejeitamos  $H_0$ .

## Exemplo 3

**Conclusão:** De acordo com os dados coletados e com um nível de significância de 5%, não se confirma a suspeita de que a esfera A gera cavidades com diâmetro médio superior a 0,2 mm em relação à esfera B.

Obs. A região crítica é da forma  $R_c = \{T > c\}$ , sendo que  $T$  tem distribuição t-Student com 9 g.l., se  $H_0$  for verdadeira. Neste caso, valor-p =  $P(T > 0,599) = 0,282$ . Não rejeitamos  $H_0$ , pois valor-p  $\geq \alpha$ .

## Exemplo 3

Calculamos:

$$\bar{D} = \frac{\sum_{j=1}^n D_j}{n} = \frac{5,0}{10} = 0,5 \text{ mm}$$

$$S^2 = \frac{\sum_{j=1}^n (D_j - \bar{D})^2}{n - 1} = \frac{\sum_{j=1}^n (D_j - 0,5)^2}{9} = 2,51 \text{ mm}^2$$

$$T = \frac{n(\bar{D} - \Delta_0)}{s_D} = \frac{10(0,5 - 0,2)}{\sqrt{2,51}} = 0,599.$$

Como  $T = 0,599 \notin R_c$ , não rejeitamos  $H_0$ .

**Conclusão:** De acordo com os dados coletados e com um nível de significância de 5%, não se confirma a suspeita de que a esfera A gera cavidades com diâmetro médio superior a 0,2 mm em relação à esfera B.

Obs. A região crítica é da forma  $R_c = \{T > c\}$ , sendo que  $T$  tem distribuição t-Student com 9 g.l., se  $H_0$  for verdadeira. Neste caso, valor-p =  $P(T > 0,599) = 0,282$ . Não rejeitamos  $H_0$ , pois valor-p  $\geq \alpha$ .

## Exemplo 4

Estudos anteriores indicam que a vida (em horas) de um termopar produzido em uma indústria é uma variável aleatória com distribuição aproximadamente normal. Um grande comprador suspeita que o tempo de vida é inferior a 560 h. Em uma amostra aleatória de 15 termopares adquiridos foram medidos os tempos de vida (em h) 553, 552, 567, 579, 550, 541, 537, 553, 552, 546, 538, 553, 581, 539 e 529. O que os dados permitem concluir sobre a suspeita do comprador? Adote  $\alpha = 5\%$ .

## Exemplo 4

### Solução

- Problema envolve uma população com distribuição normal.
- Definimos  $X$  como sendo o tempo de vida (em horas) de um termopar, com  $E(X) = \mu$  e  $\text{var}(X) = \sigma^2$ .
- Pelo enunciado,  $X \sim N(\mu, \sigma^2)$ ,  $\sigma^2$  desconhecida.
- Hipóteses:
  - $H_0 : \mu = 560$  contra  $H_1 : \mu < 560$  (ou seja,  $\mu_0 = 560$ ).
- Estatística de teste:
  - $$T = \frac{n(\bar{X} - \mu_0)}{s}$$

## Exemplo 4

A região crítica para  $\alpha = 0,05$  é obtida consultando a tabela da distribuição t-Student com 14 g.l. ( $n - 1$ ) e  $p = 10\%$ :

$$R_c = \{T < -1,761\}.$$

Calculamos:

$$\bar{X} = \frac{\sum_{j=1}^n X_j}{n} = \frac{8270}{15} = 551,3 \text{ h}$$

$$S^2 = \frac{\sum_{j=1}^n (X_j - \bar{X})^2}{n - 1} = \frac{\sum_{j=1}^n (X_j - 551,3)^2}{14} = 14,8 \text{ h}^2$$

$$T = \frac{n(\bar{X} - \mu_0)}{s} = \frac{15(551,3 - 560)}{14,8} = -2,266.$$

Como  $T = -2,266 \in R_c$ , rejeitamos  $H_0$ .

## Exemplo 4

**Conclusão:** De acordo com os dados coletados e com um nível de significância de 5%, concluímos que a vida média dos termopares é inferior a 560 h.

Obs. A região crítica é da forma  $R_c = \{T < -c\}$ , sendo que  $T$  tem distribuição t-Student com 14 g.l., se  $H_0$  for verdadeira. Calculamos  $T = -2,266$ . Neste caso,  $\text{valor-p} = P(T < -2,266) = 0,0199$ .  
Rejeitamos  $H_0$ , pois  $\text{valor-p} < \alpha$ .

## Determinação do tamanho das amostras

No planejamento de experimentos para comparar dois tratamentos, surge a questão de qual deve ser o número  $n$  de ensaios para cada tratamento. A determinação de  $n$  e a probabilidade do erro tipo II estão relacionados.

Suponha que estamos testando as hipóteses:

$$H_0 : \mu_1 = \mu_2$$

$$H_1 : \mu_1 \neq \mu_2$$

E que as médias não são iguais, de modo que  $\delta = \mu_1 - \mu_2$ . Já que  $H_0 : \mu_1 = \mu_2$  não é verdadeira, o problema está em decidir erradamente não rejeitar  $H_0$  (erro tipo II).

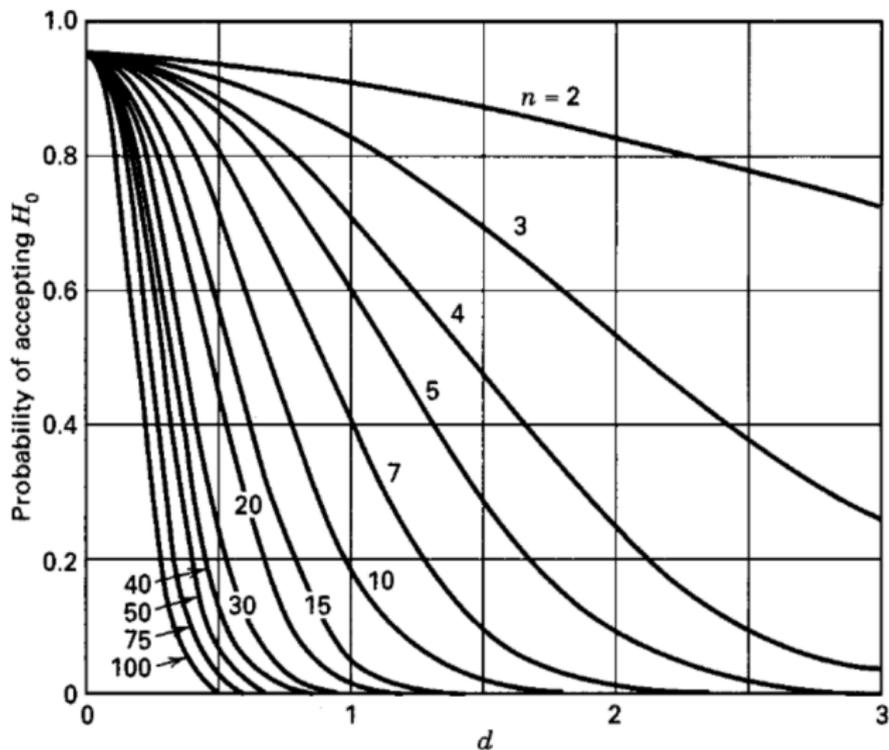
Erro tipo II: não rejeitar  $H_0$  quando  $H_0$  é falsa.

## Determinação do tamanho das amostras

A probabilidade do erro tipo II depende da diferença verdadeira que há entre as médias  $\delta$ . O gráfico de  $\beta$  e  $\delta$  no caso de um verdadeiro valor particular do tamanho da amostra, se conhece como curva característica operativa (CO) para o teste.

- A probabilidade do erro tipo II,  $\beta$ , é uma função de  $n$ . Geralmente, para um valor dado de  $\delta$ ,  $\beta$  diminui à medida que  $n$  se incrementa.
- Isto significa que uma diferença específica é mais fácil de ser detectada em amostras grandes.

## Determinação do tamanho das amostras



**Figure 2-12** Operating characteristic curves for the two-sided  $t$ -test with  $\alpha = 0.05$ . (Reproduced with permission from "Operating Characteristics for the Common Statistical Tests of Significance," C. L. Ferris, F. E. Grubbs, and C. L. Weaver. *Annals of Mathematical Statistics*, June 1946.)

## Curva Característica da Operação (CO) para Teste de Hipóteses

A Figura mostra Curva Característica de Operação para as hipóteses

$$H_0 : \mu_1 = \mu_2$$

$$H_1 : \mu_1 \neq \mu_2,$$

no caso  $\sigma_1 = \sigma_2 = \sigma$  desconhecida e para um nível de significância  $\alpha = 0.05$  com  $n_1 = n_2 = n$ .

O eixo horizontal da Figura é  $\frac{\mu_1 - \mu_2}{2\sigma}$ .

O tamanho amostral usado na construção destas curvas é  $n^* = 2n - 1$ .

## Observações da Curva CO

Das curvas CO observamos:

- Maior a diferença das médias, menor é a probabilidade do erro tipo II dado o tamanho da amostra e  $\alpha$ . Isto é, que para um valor específico de tamanho da amostra e um nível  $\alpha$ , o teste detecta mais facilmente diferenças grandes que pequenas.
- À medida que o tamanho de amostra aumenta, a probabilidade de erro tipo II diminui quando o nível de diferença das médias e o valor de  $\alpha$  estão dados. Isto significa que para detectar uma diferença específica  $\delta$ , o teste pode ser mais eficaz se o tamanho da amostra for incrementado.

## Exemplo 5: Detecção de Diferença Mínima

Considere o problema do cimento Portland apresentado anteriormente. Suponha que se deseja detectar com alta probabilidade uma diferença mínima de 0,5 unidades no nível médio dos dados. Supondo  $\delta = \mu_1 - \mu_2 = 0,5$  é um valor crítico na diferença do nível médio que deseja detectar. Então:

$$\begin{aligned}\mu_1 - \mu_2 &= 0,5 \\ d &= \frac{0,5}{2\sigma} \\ &= \frac{0,5}{2 \times 0,25} \\ &= 1\end{aligned}$$

Suponha que estudos anteriores mostraram que o desvio padrão da resistência excede 0,25. Daí consideramos  $\sigma = 0,25$ , o que implica que  $d = 1$ .

## Determinação do Tamanho da Amostra

Se deseja que a probabilidade de rejeitar  $H_0$  quando  $H_0 : \mu_1 - \mu_2 = 0,5$  seja 0,95 (poder do teste), o que implica que  $\beta = 0,05$ . A partir da Figura com  $d = 1$  e  $\beta = 0,05$ , tem-se que  $n^* = 16$  (aproximadamente). Mas  $n^* = 2n - 1 = 16$ , então temos que  $n = 8,5 \approx 9$ . Ou seja, deve-se usar tamanhos amostrais no mínimo  $n_1 = n_2 = n = 9$ .

## Testes de Hipóteses sobre Igualdade de Variâncias

Seja  $Y_{11}, \dots, Y_{1n_1}$  uma amostra aleatória  $n_1$  observações de uma população normal com média  $\mu_1$  e variância  $\sigma_1^2$ , e  $Y_{21}, \dots, Y_{2n_2}$  uma amostra aleatória  $n_2$  observações de uma população normal com média  $\mu_1$  e variância  $\sigma_2^2$ .

(i) Formulação das hipóteses:

$$H_0 : \sigma_1^2 = \sigma_2^2$$

$$H_1 : \sigma_1^2 \neq \sigma_2^2$$

(ii) Estatística de teste:

$$F = \frac{S_1^2}{S_2^2} \stackrel{\text{Sob } H_0}{\sim} F(n_1 - 1, n_2 - 1)$$

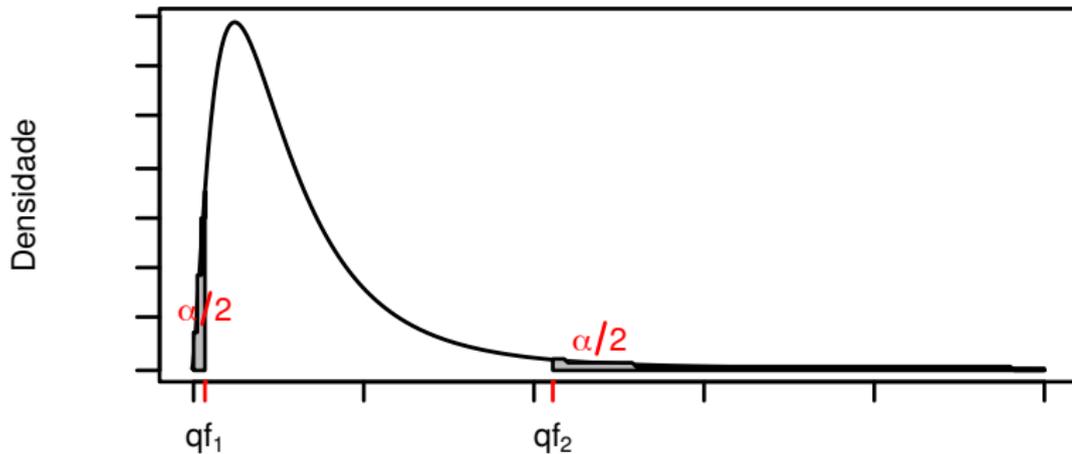
(iii) A região crítica, para  $\alpha$  fixado como mostrado na figura.

$$RC = \{F \leq f_{1-\alpha/2, n-1, m-1} \text{ ou } F \geq f_{\alpha/2, n-1, m-1}\}$$

(iv) Se a  $F_{\text{obs}} \in RC$ , rejeita-se  $H_0$ ; caso contrário, não se rejeita  $H_0$ .

# Testes de Hipóteses sobre Igualdade de Variâncias

Densidade da Distribuição F



## Testes de Hipóteses sobre Igualdade de Variâncias

Exemplo: A tabela seguinte sintetiza medidas de corrente elétrica (em mA) obtidas por dois engenheiros na análise de placas de circuito integrado.

- I: 35, 35, 37, 33, 31, 33
- II: 32, 34, 34, 31, 32

Há evidência para se afirmar que as medidas possuem a mesma homogeneidade?

- Engenheiro I: Média = 34,00, Mediana = 34, Desvio Padrão = 2,098
- Engenheiro II: Média = 32,60, Mediana = 32, Desvio Padrão = 1,342

## Testes de Hipóteses sobre Igualdade de Variâncias

Sejam:

- $X$ : Medição do Eng. I,  $X \sim N(\mu, \sigma^2)$
- $Y$ : Medição do Eng. II,  $Y \sim N(\mu, \sigma^2)$

Nosso interesse é provar as seguintes hipóteses:

$$H_0 : \sigma_1^2 = \sigma_2^2$$

$$H_1 : \sigma_1^2 \neq \sigma_2^2$$

(ii) A estatística de teste:

$$F = \frac{S_1^2}{S_2^2} \sim F(n - 1, m - 1)$$

onde  $n = 6$  e  $m = 5$ .

## Testes de Hipóteses sobre Igualdade de Variâncias

(iii) Os valores críticos para  $\alpha = 0,10$  são:

$$f_{0,05,5,4} = 6,26$$

$$f_{0,95,4,5} = 0,193$$

A região crítica é dada por:

$$Rc = \{F \leq 0,193 \text{ ou } F \geq 6,26\}$$

(iv) A estatística calculada é:

$$F_{\text{cal}} = \frac{2,098^2}{1.342^2} = 2,444$$

Como  $F_{\text{cal}} \notin Rc$ , então não rejeitamos  $H_0$ .