Teste de Student “t”

O teste t é muito utilizado para comparar duas amostras (grupos). Por exemplo, pode-se testar o efeito provocado por uma determinada droga.

- **Grupo tratamento** – pacientes que receberam a droga;
- **Grupo controle** – pacientes que receberam o placebo.

Teste de Student “t”

É um teste de hipótese que usa conceitos estatísticos para rejeitar ou não uma hipótese nula quando a estatística de teste (t) segue uma distribuição t de Student.

Teste t pode ser conduzido para:

- Comparar uma amostra com uma população;
- Comparar duas amostras pareadas;
- Comparar duas amostras independentes

Teste de Student “t”

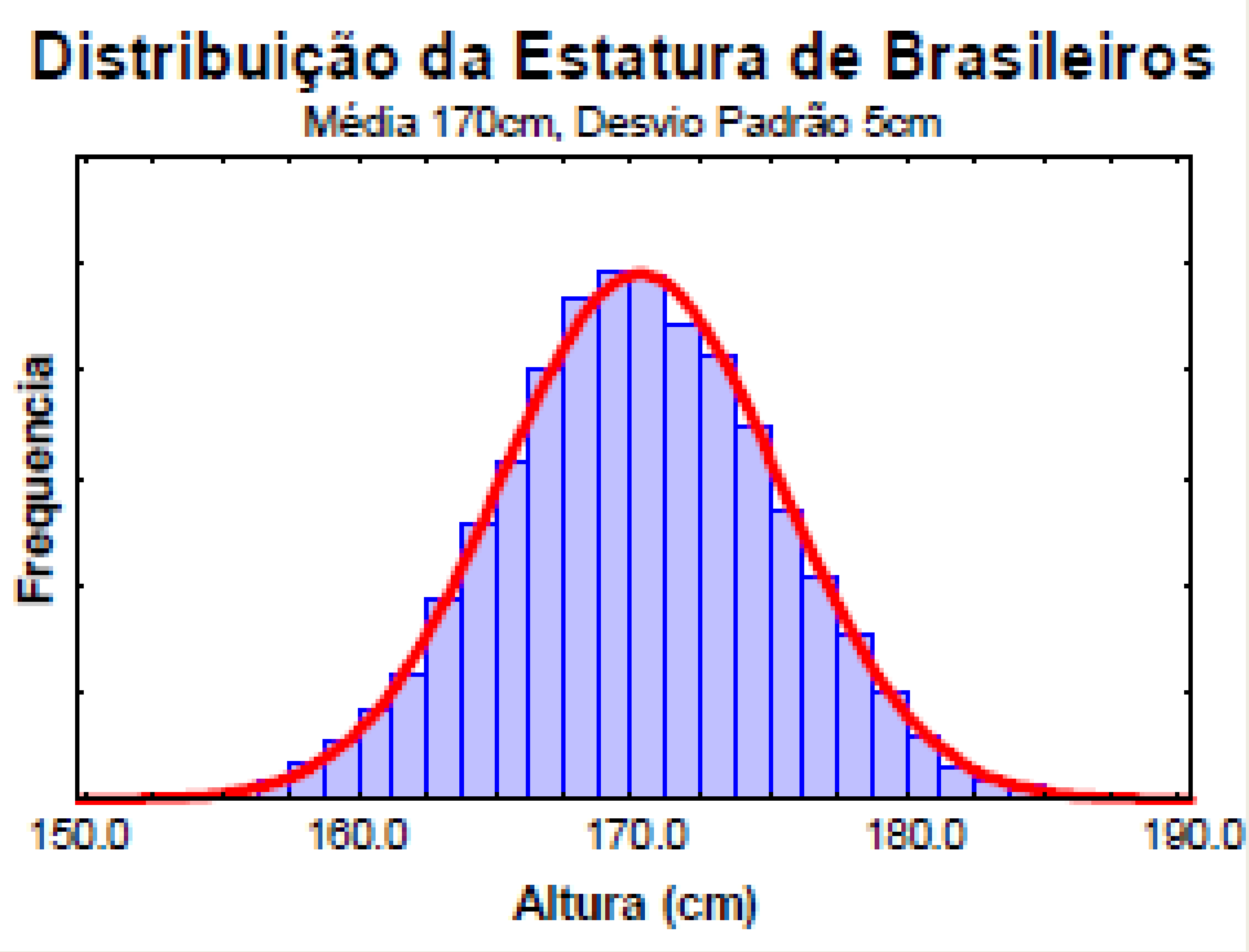
O teste t pode ser usado mesmo que as amostras sejam pequenas (n=10) desde que seja admitido que as populações que deram origem às amostras tenham **distribuição normal** e variabilidades não significativamente diferentes.

O que é DISTRIBUIÇÃO NORMAL

Em probabilidade e estatística, a **distribuição normal** é uma das distribuições de probabilidade mais utilizadas para modelar fenômenos naturais. Isso se deve ao fato de que um grande número de fenômenos naturais apresenta sua distribuição de probabilidade tão proximamente normal, que a ela pode ser com sucesso referida, e, portanto, com adequado acerto por ela representada como se normal fosse.

A distribuição normal também é chamada distribuição gaussiana, distribuição de Gauss ou distribuição de Laplace–Gauss, em referência aos matemáticos, físicos e astrônomos francês Pierre–Simon Laplace (1749 – 1827) e alemão Carl Friedrich Gauss (1777 – 1855).

DISTRIBUIÇÃO NORMAL



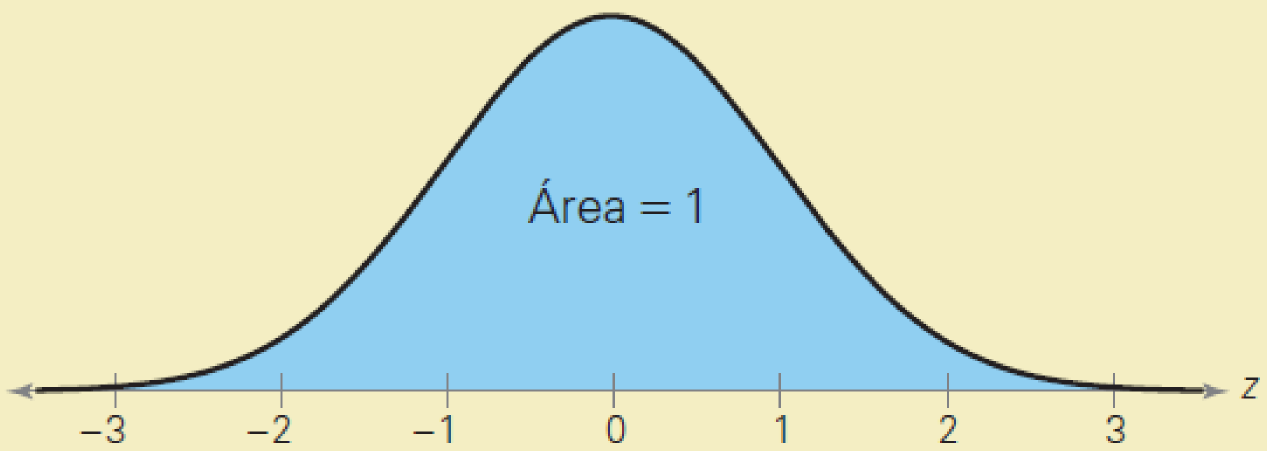
Fonte: Rangel, T. (UFG)

DISTRIBUIÇÃO NORMAL X DISTRIBUIÇÃO T

Definição

A **distribuição normal padrão** é uma distribuição normal com média 0 e desvio padrão 1. A área total sob sua curva normal é 1 (veja a Figura A.2).

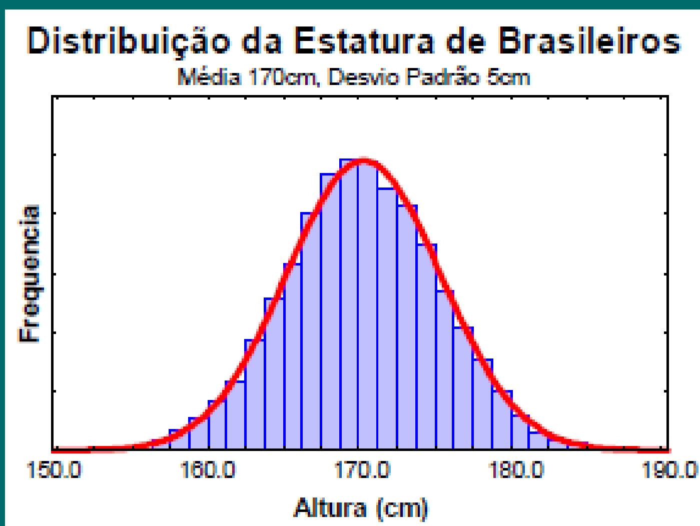
Figura A.2 Distribuição normal padrão.

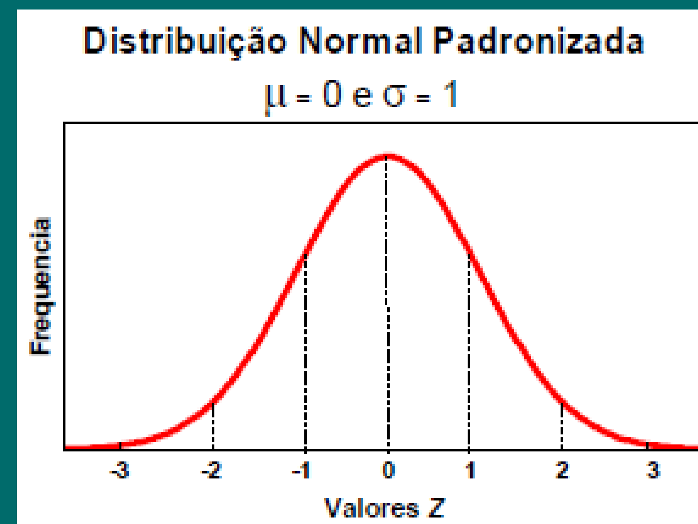


Larson e Farber (Estatística Aplicada)

DISTRIBUIÇÃO NORMAL

$$z = \frac{x - \mu}{\sigma}$$





$$x = (z \times \sigma) + \mu$$

Fonte: Rangel, T. (UFG)

O que é DISTRIBUIÇÃO T

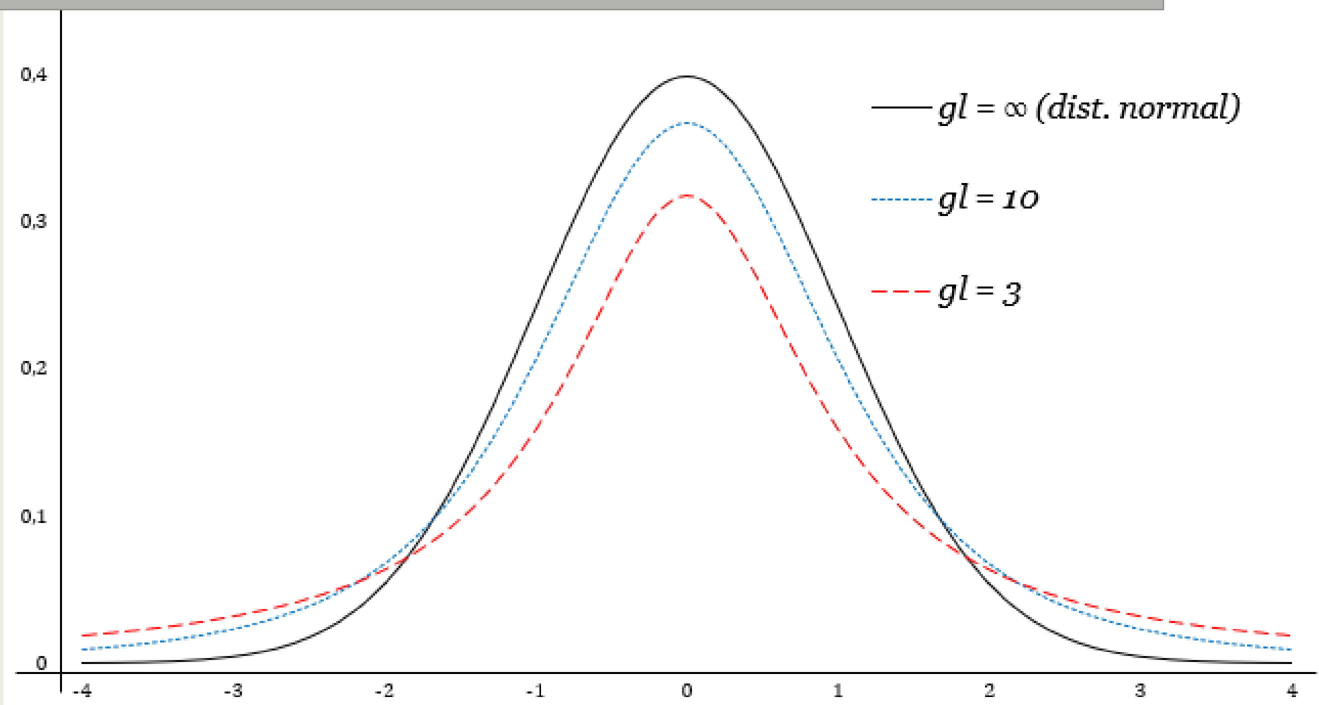
A distribuição t de Student é uma distribuição de probabilidade, publicada por William Sealy Gosset sob o pseudônimo Student no início do século XX, a partir de trabalhos na cervejaria Guinnes na Irlanda.

A distribuição t de Student aparece naturalmente no problema de se determinar a média de uma população (que segue a distribuição normal) a partir de uma amostra. Neste problema, não se sabe qual é a média ou o desvio padrão da população, mas ela deve ser normal.

DISTRIBUIÇÃO T

TEOREMA DO LIMITE CENTRAL

Quando o tamanho da amostra aumenta, a distribuição amostral da sua média aproxima-se cada vez mais de uma distribuição normal



Teste “t” para observações independentes

Procedimento: a variável em análise tem distribuição normal ou aproximadamente normal?

Se a resposta for afirmativa pode-se aplicar o teste t, para comparar as médias.

Com o objetivo de comparar duas médias amostrais advindas de populações independentes, as Hipóteses do teste t seriam:

$$\boldsymbol{H_0: \bar{x}_1 = \bar{x}_2}$$

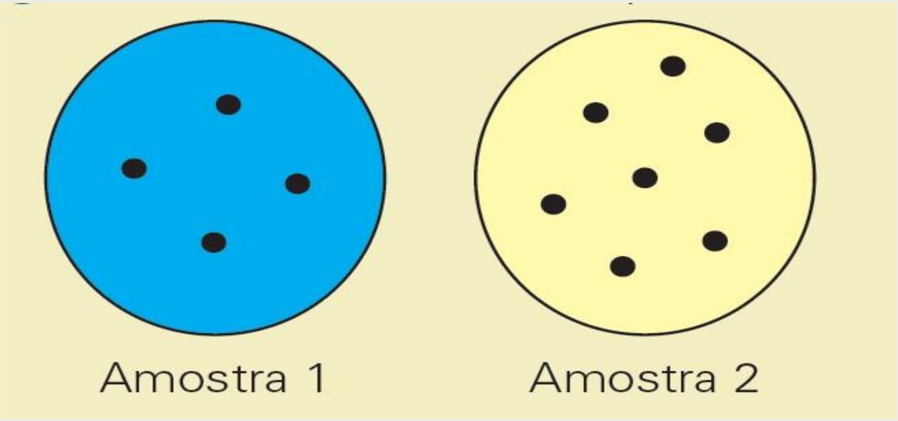
$$\boldsymbol{H_1: \bar{x}_1 \neq \bar{x}_2}$$

Teste "t" para observações independentes

Calculam-se:

1) A média de cada grupo

2) A variância de cada grupo



Média do grupo 1 $(\bar{x}_1)$ Média do grupo 2 $(\bar{x}_2)$

Variância do grupo 1 (s_1^2) Variância do grupo 2 (s_2^2)

Teste “t” para observações independentes

3) A variância Ponderada dos Grupos

4) O Valor de t_{calc}

5) Os Graus de Liberdade dos grupos

6) Obter valor tabelado de t (t_{tab}) e compará-lo com valor de t calculado (t_{calc})

Teste “t” para observações independentes

Valores de *t*, segundo os graus de liberdade e o valor de α

Graus de Liberdade	α 10%	5%	1%
1	6,31	12,71	63,66
2	2,92	4,30	9,92
3	2,35	3,18	5,84
4	2,13	2,78	4,60
5	2,02	2,57	4,03

Teste “t” para observações independentes

7) Fazer a interpretação

Toda vez que o valor calculado de t, em valor absoluto for igual ou maior do que o tabelado, conclui-se que as médias não são iguais, ao nível de significância estabelecido.

8) Concluir o trabalho conforme seus objetivos

EXEMPLO

Objetivo:

- Verificar se duas dietas para emagrecer são igualmente eficientes.

Método:

- Separar ao acaso um conjunto de pacientes em dois grupos.
- Cada um deve seguir a dieta designada para o seu grupo.

Resultados:

Tabela: Perda de Peso (Kg) de duas dietas

Dieta 1	Dieta 2
12	15
8	19
15	15
13	12
10	13
12	16
14	15
11	
12	
13	

Calcular:

1. Média dos grupos:

$$\bar{x} = \frac{\sum x_i}{n}$$

2. Variância dos grupos (s^2)

$$s^2 = \frac{\sum (x_i - \bar{x})^2}{n-1}$$

ou

$$s^2 = \frac{\sum x_i^2 - \frac{(\sum x_i)^2}{n}}{n-1}$$

Resultados:

Tabela: Perda de Peso (Kg) de duas dietas

Dieta 1	Dieta 2
12	15
8	19
15	15
13	12
10	13
12	16
14	15
11	
12	
13	

Calcular:

3. Variância ponderada $\left(s_p^2\right)$

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{\underbrace{n_1 + n_2 - 2}_{\text{Graus de Liberdade (G.L.)}}}$$

4. Valor de ***t*** de student

$$t = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

Teste “t” para observações independentes

1) A média de cada grupo:

$$\bar{x} = \frac{\sum x_i}{n}$$

$$\bar{\boldsymbol{x}}_{\mathbf{1}} = \frac{120}{10} = \mathbf{12} \qquad \bar{\boldsymbol{x}}_{\mathbf{2}} = \frac{105}{7} = \mathbf{15}$$

2) A variância de cada grupo:

$$s^2 = \frac{\sum x_i^2 - \frac{(\sum x_i)^2}{n}}{n-1}$$

$$\boldsymbol{s}_{\mathbf{1}}^{\mathbf{2}} = \frac{1476 - \frac{120^2}{10}}{9} = \mathbf{4} \qquad \boldsymbol{s}_{\mathbf{2}}^{\mathbf{2}} = \frac{1605 - \frac{105^2}{7}}{6} = \mathbf{5}$$

Teste “t” para observações independentes

3) A variância ponderada, dada pela fórmula:

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

$$\boldsymbol{s_p^2} = \frac{(9)4 + (6)5}{10 + 7 - 2} = \mathbf{4,4}$$

4) O valor de t, calculado por:

$$t = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$\boldsymbol{t} = \frac{|12 - 15|}{\sqrt{\left(4{,}4\left(\frac{1}{10} + \frac{1}{7}\right)\right)}} = \mathbf{2,902}$$

5) Valor de t tabelado – Tabela de Student (5%)

Graus de Liberdade (GL) = $n_1 + n_2 - 2 = 10 + 7 - 2 =$ **15**

Valores de *t*, segundo os graus de liberdade e o valor de α

Graus de liberdade	10%	α 5%	1%
1	6,31	12,71	63,66
2	2,92	4,30	9,92
3	2,35	3,18	5,84
4	2,13	2,78	4,60
5	2,02	2,57	4,03
6	1,94	2,45	3,71
7	1,90	2,36	3,50
8	1,86	2,31	3,36
9	1,83	2,26	3,25
10	1,81	2,23	3,17
11	1,80	2,20	3,11
12	1,78	2,18	3,06
13	1,77	2,16	3,01
14	1,76	2,14	2,98
15	1,75	2,13	2,95
16	1,75	2,12	2,92

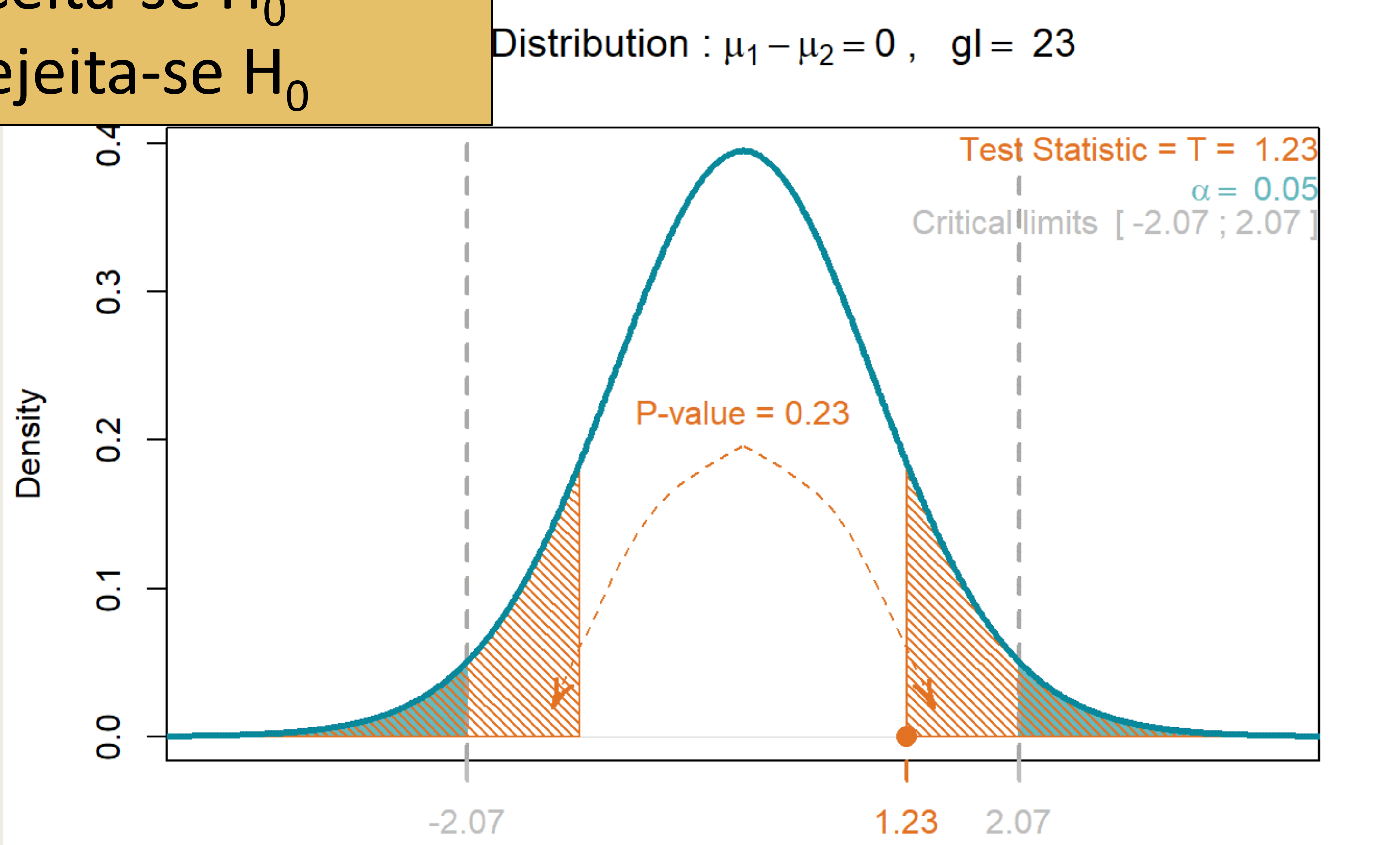
Como decidir se aceita ou rejeita H_0?

Regra de decisão

Quando:

$t_{calc} < t_{tab} \rightarrow$ aceita-se H_0

$t_{calc} > t_{tab} \rightarrow$ rejeita-se H_0



Neste exemplo o valor de $t_{calc.}$ é menor que t_{tab}: **Aceita-se H_0**

Teste "t" para observações independentes

6) Comparar os valores de t

t_c = 2,902 valor calculado

t_{tab} = 2,13 valor tabelado (α=5%)

Regra de decisão

Quando:

$t_c < t_{tab} \rightarrow$ aceita-se H_0

$t_c > t_{tab} \rightarrow$ rejeita-se H_0

7) Interpretar o resultado

$t_c > t_{ab}$, rejeita-se H_0 e aceita-se H_1, as médias dos grupos são diferentes.

Teste “t” para observações independentes

6) Comparar os valores de t

t_c = 2,902 valor calculado

t_{tab} = 2,13 valor tabelado (α=5%)

Regra de decisão

Quando:

$t_c < t_{tab} \rightarrow$ aceita-se H_0

$t_c > t_{tab} \rightarrow$ rejeita-se H_0

7) Interpretar o resultado

$t_c > t_{ab}$, rejeita-se H_0 e aceita-se H_1, as médias dos grupos são diferentes.

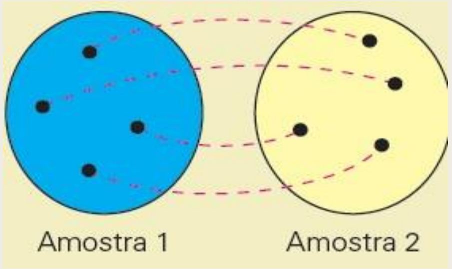
Teste “t” para observações independentes

8) Conclusão da pesquisa

As perdas de pesos de pacientes submetidos aos dois tipos de dieta são diferentes. Em termos práticos, a perda de peso é maior quando os pacientes são submetidos a dieta 2.

Teste “t” para observações pareadas

Para estudar o efeito de um tratamento, muitas vezes comparam-se pares de indivíduos.



Exemplo:

- Psicologia: compara pares de gêmeos.
- Tratamentos onde se observam os mesmos indivíduos duas vezes, antes e depois do tratamento: pressão arterial.
- A altura de árvores medidas com dois instrumentos diferentes.

Teste “t” para observações pareadas

Hipótese para o teste t dependente

$$\boldsymbol{H_0}: \bar{\boldsymbol{d}} = \mathbf{0}$$

$$\boldsymbol{H_1}: \bar{\boldsymbol{d}} \neq \mathbf{0}; \bar{\boldsymbol{d}} < \mathbf{0}; \bar{\boldsymbol{d}} > \mathbf{0}$$

Teste “t” para observações pareadas

1) A diferença das duas medidas por observação

2) A média das diferenças $\left(\bar{d}\right)$

3) A variância das diferenças

4) O valor de t_c

$$t_c = \frac{\bar{d}}{\sqrt{\frac{s_d^2}{n}}}$$

5) Graus de liberdade: *n-1*

Teste "t" para observações pareadas

Exemplo: São dados os pesos de 9 pessoas, antes e depois da dieta para emagrecimento.

Dieta	
Antes	**Depois**
77	80
62	58
61	61
80	76
90	79
72	69
86	90
59	51
88	81

Teste “t” para observações pareadas

Dieta		Diferenças	d
Antes	Depois		
77	80	80 - 77 =	3
62	58	58 - 62 =	- 4
61	61	61 - 61 =	0
80	76	76 - 80 =	- 4
90	79	79 - 90 =	- 11
72	69	69 - 72 =	- 3
86	90	90 - 86 =	4
59	51	51 - 59 =	- 8
88	81	81 - 88 =	- 7
		Soma ($\sum$d)	**- 30**

Teste "t" para observações pareadas

2) A média das diferenças:

$$\bar{d} = \frac{-30}{9} = -3{,}333$$

$$\bar{d} = \frac{\sum d_i}{n}$$

3) A variância das diferenças:

$$s_d^2 = \frac{300 - \frac{(-30)^2}{9}}{8} = 25$$

$$s_d^2 = \frac{\sum d_i^2 - \frac{(\sum d_i)^2}{n}}{n-1}$$

4) O valor de t_c:

$$t_c = \frac{-3{,}333}{\sqrt{\left(\frac{25}{9}\right)}} = -2{,}0$$

$$t_c = \frac{\bar{d}}{\sqrt{\frac{s_d^2}{n}}}$$

5) Valor de t tabelado – Tabela de Student (1%)

Graus de Liberdade (GL) = $n_1 - 1 = 9 - 1 = $ **8**

Tabela A.6

Valores de *t*, segundo os graus de liberdade e o valor de α

Graus de liberdade	α 10%	α 5%	α 1%
1	6,31	12,71	63,66
2	2,92	4,30	9,92
3	2,35	3,18	5,84
4	2,13	2,78	4,60
5	2,02	2,57	4,03
6	1,94	2,45	3,71
7	1,90	2,36	3,50
8	1,86	2,31	3,36
9	1,83	2,26	3,25
10	1,81	2,23	3,17
11	1,80	2,20	3,11
12	1,78	2,18	3,06
13	1,77	2,16	3,01
14	1,76	2,14	2,98
15	1,75	2,13	2,95
16	1,75	2,12	2,92
17	1,74	2,11	2,90
18	1,73	2,10	2,88
19	1,73	2,09	2,86
20	1,73	2,09	2,84

Teste “t” para observações pareadas

Regra de decisão
Quando:
$t_c < t_{tab} \rightarrow$ aceita-se H_0
$t_c > t_{tab} \rightarrow$ rejeita-se H_0

6) Comparar os valores de t

$t_c = 2{,}00$ (valor calculado)

$t_{tab} = 3{,}36$ (valor tabelado)

7) Interpretar o resultado

$t_c < t_{tab}$; Aceita-se H_0, ou seja, não há diferença entre os valores medidos nas duas ocasiões (diferença igual a zero).

Teste “t” para observações pareadas

8) Conclusão da pesquisa

As perdas de pesos dos pacientes submetidos à dieta não foi significativa. Assim, o experimento não provou que a dieta emagrece.