

## Unidade IV - Inferência estatística

- 4.1. Introdução e histórico
- 4.2. Conceitos fundamentais
- 4.3. Distribuições amostrais e Teorema central do limite
- 4.4. Estimação de parâmetros
- 4.5. Testes de hipóteses**
- 4.6. Quebras das pressuposições no processo de inferência
- 4.7. Testes de qui-quadrado

### Teoria dos testes de hipóteses

## Testes de hipóteses

- ⇒ Outro problema a ser resolvido pela Inferência Estatística é o de **testar uma hipótese**.
- ⇒ Feita uma afirmação a respeito de uma população (parâmetro) desejamos **saber se os resultados experimentais (amostra) contrariam tal afirmação**.
- ⇒ Muitas vezes, essa afirmação sobre a população é derivada de teorias desenvolvidas no campo do conhecimento por meio da observação e do raciocínio.
- ⇒ A adequação dessa teoria ao universo real pode ser verificada ou refutada pela amostra.
- ⇒ O objetivo do teste estatístico de hipótese é então fornecer uma metodologia que permita verificar se os dados amostrais trazem evidências que apoiam ou não a hipótese (estatística) formulada.

## Teste de hipóteses

Procedimento estatístico onde se busca verificar uma hipótese a respeito da **população**, no sentido de aceitá-la ou rejeitá-la, a partir de **dados amostrais**, tendo por base a teoria das probabilidades.

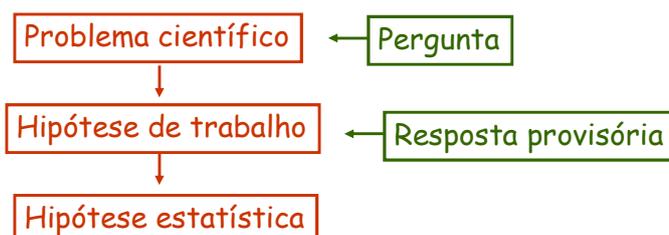
### Algoritmo para construção de um teste de hipóteses

1. Definir as hipóteses estatísticas.
2. Fixar a taxa de erro aceitável.
3. Escolher a estatística para testar a hipótese e verificar as pressuposições para o seu uso.
4. Usar as observações da amostra para calcular o valor da estatística do teste.
5. Decidir sobre a hipótese testada e concluir.

### 1. Hipóteses estatísticas

## Hipóteses estatísticas

A hipótese estatística é uma suposição feita a respeito de um ou mais parâmetros ( $\mu$ ,  $\sigma^2$ ,  $\pi$ , etc.).



Profa. Clause Piana

7

Existem dois tipos básicos de hipóteses estatísticas:

- ⇒ **Hipótese de nulidade ( $H_0$ )**: é a hipótese que está sob verificação. Esta hipótese supõe a igualdade dos parâmetros que estão sendo comparados.
- ⇒ **Hipótese alternativa ( $H_A$ )**: é a hipótese que será considerada caso a hipótese de nulidade seja rejeitada. Esta hipótese supõe que os parâmetros comparados são diferentes.

Situações comuns em testes de hipóteses a respeito de  $\mu$

- ▣ Comparação de uma média ( $\mu$ ) com um valor padrão ( $\mu_0$ )
- ▣ Comparação entre duas médias ( $\mu_1$  e  $\mu_2$ )

Profa. Clause Piana

8

- ▣ Comparação de uma média ( $\mu$ ) com um valor padrão ( $\mu_0$ )



**Exemplo 1:** Para verificar se uma nova droga é eficaz no tratamento da pressão alta, a pressão média de um grupo de pacientes submetidos a esta droga (amostra) é comparada com um valor que é considerado normal (valor padrão).

Profa. Clause Piana

9

### Exemplo 2:

Um engenheiro de desenvolvimento de um fabricante de pneus está investigando a vida do pneu em relação a um novo componente de borracha. Ele fabricou 40 pneus e testou-os até o fim da vida em um teste na estrada. A média e o desvio padrão da amostra são 61.492 e 6.085 km, respectivamente.

O engenheiro acredita que a vida média desse novo pneu supera 60.000 km.

Teste a hipótese de que a verdadeira vida média do pneu não difere de 60.000 km e conclua a respeito da suposição do engenheiro.

Profa. Clause Piana

10

□ Comparação de uma média ( $\mu$ ) com um valor padrão ( $\mu_0$ )

$$H_0 : \mu = \mu_0 \text{ ou } \mu - \mu_0 = 0$$

$$H_A : \mu \neq \mu_0 \text{ ou } \mu - \mu_0 \neq 0$$

$$\mu > \mu_0 \text{ ou } \mu - \mu_0 > 0$$

$$\mu < \mu_0 \text{ ou } \mu - \mu_0 < 0$$

Bilateral

Unilateral direita

Unilateral esquerda

Escolher uma das três

□ Comparação entre duas médias ( $\mu_1$  e  $\mu_2$ )

População 1

População 2



Amostra 1

Amostra 2



Estimativa de  $\mu_1$



Estimativa de  $\mu_2$

Comparação das estimativas

**Exemplo 1:** Para verificar, entre dois métodos de ensino, qual é o mais eficiente, comparamos os desempenhos médios de dois grupos de alunos (duas amostras), cada um submetido a um método diferente.

**Exemplo 2:**

Um pesquisador da área de computação está investigando a utilidade de duas diferentes linguagens de programação (A e B) na melhoria das tarefas computacionais.

Trinta programadores experientes, familiarizados com ambas as linguagens, foram divididos aleatoriamente em dois grupos. Cada grupo codificou uma função padrão em uma das linguagens e os tempos de codificação da função (em minutos) foram registrados.

As medidas de cada grupo são apresentadas no quadro abaixo.

	n	Média	Variância
Linguagem A	15	17,5	2,31
Linguagem B	15	20,5	3,02

Utilizando um teste de hipótese, verifique se, em média, o tempo de codificação da função padrão difere entre as linguagens.

▣ Comparação entre duas médias ( $\mu_1$  e  $\mu_2$ )

$$H_0 : \mu_1 = \mu_2 \text{ ou } \mu_1 - \mu_2 = 0$$

$$H_A : \mu_1 \neq \mu_2 \text{ ou } \mu_1 - \mu_2 \neq 0$$

$$\mu_1 > \mu_2 \text{ ou } \mu_1 - \mu_2 > 0$$

$$\mu_1 < \mu_2 \text{ ou } \mu_1 - \mu_2 < 0$$

Bilateral

Unilateral direita

Unilateral esquerda

Escolher uma das três

**Exemplo 1:**

**Problema científico:** Ração a base de soja promove maior ganho de peso em bovinos do que ração a base de milho?

**População 1** - bovinos tratados com farelo de soja

**População 2** - bovinos tratados farelo de milho

**Variável em estudo:** X - medida do ganho de peso (kg)

$E(X_1) = \mu_1$  = ganho de peso médio com farelo de soja

$E(X_2) = \mu_2$  = ganho de peso médio com farelo de milho

**Hipótese de trabalho:** As duas rações têm efeitos diferentes sobre o ganho de peso de bovinos.

Profa. Clause Piana

15

**Hipóteses estatísticas:**

⇒ **Hipótese de nulidade:** o ganho de peso médio de bovinos tratados com farelo de soja não difere do ganho de peso médio de bovinos tratados com farelo de milho

⇒ **Hipótese de alternativa:** o ganho de peso médio de bovinos tratados farelo de soja difere do ganho de peso médio de bovinos tratados com farelo de milho

$$\begin{cases} H_0 : \mu_1 = \mu_2 \text{ ou } \mu_1 - \mu_2 = 0 \\ H_A : \mu_1 \neq \mu_2 \text{ ou } \mu_1 - \mu_2 \neq 0 \end{cases} \leftarrow \text{Bilateral}$$

Quando não temos motivos suficientes para supor que uma das médias será maior que a outra, formulamos uma **hipótese alternativa bilateral** (mais genérica).

Quando a hipótese alternativa é bilateral, dizemos que o **teste de hipóteses é bilateral**.

**Exemplo 2:**

**Problema científico:** Um novo inseticida é eficaz no combate à lagarta da soja?

**População 1** - lavouras de soja com aplicação do inseticida

**População 2** - lavouras de soja sem aplicação do inseticida

**Variável em estudo:**  $X$  - número de insetos mortos

$E(X_1) = \mu_1$  = número médio de insetos da população 1 (Com)

$E(X_2) = \mu_2$  = número médio de insetos da população 2 (Sem)

**Hipótese de trabalho:** O novo inseticida é eficaz no combate à lagarta da soja

Profa. Clause Piana

17

**Hipóteses estatísticas:**

⇒ **Hipótese de nulidade:** a média de insetos mortos na lavoura com inseticida não difere da média de insetos mortos na lavoura sem inseticida

⇒ **Hipótese alternativa:** a média de insetos mortos na lavoura com inseticida é maior que a média de insetos mortos na lavoura sem inseticida

$$\begin{cases} H_0 : \mu_1 = \mu_2 \text{ ou } \mu_1 - \mu_2 = 0 \\ H_A : \mu_1 > \mu_2 \text{ ou } \mu_1 - \mu_2 > 0 \end{cases} \leftarrow \text{Unilateral direita}$$

Quando temos motivos suficientes para supor que uma das médias será maior que a outra, podemos formular uma **hipótese alternativa unilateral** (mais específica).

Quando a hipótese alternativa é unilateral, dizemos que o **teste de hipóteses é unilateral**.

**Objetivo:** verificar a hipótese

Podemos verificar a hipótese de duas formas:

- ⇒ avaliar as populações inteiras (todos os bovinos tratados com as duas rações ou todas as lavouras de soja) e comparar suas médias
- ⇒ avaliar amostras retiradas das populações e utilizar um teste estatístico que compare as médias das amostras

Devemos considerar:

- ⇒ seria impossível avaliar todos os bovinos tratados com as duas rações ou todas as lavouras de soja
- ⇒ o processo de amostragem pode fornecer precisão suficiente

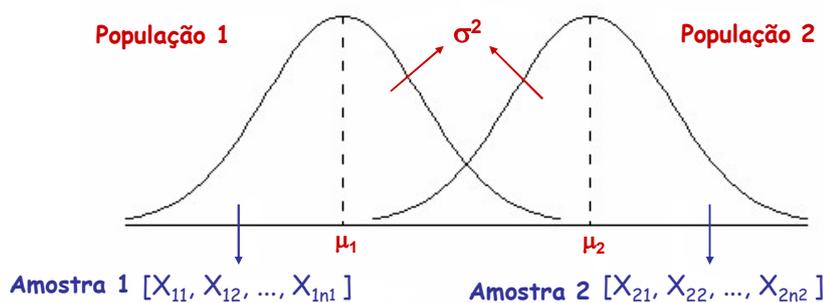
Será muito mais econômico e menos trabalhoso utilizar amostras das populações.

Profa. Clause Piana

19

**Pressuposições:**

- A variável em estudo tem distribuição normal
- As variâncias das populações são iguais
- As amostras retiradas das populações são independentes



Para determinar (dimensionar) o tamanho da amostra, devemos definir:

- ◆ o tamanho do erro que estamos dispostos a admitir
- ◆ a variabilidade aproximada das populações ( $\sigma^2$ )
- ◆ a magnitude da diferença a ser detectada

## 2. Erros de conclusão

### Erros de conclusão

Como a hipótese sob verificação é  $H_0$ , dois tipos de erro estão associados à decisão a respeito dela:

- ♦ **Erro Tipo I:** rejeitar  $H_0$  quando esta é verdadeira  
 $\alpha = P(\text{erro tipo I}) \rightarrow$  probabilidade de cometer o erro tipo I
- ♦ **Erro Tipo II:** não rejeitar  $H_0$  quando esta é falsa  
 $\beta = P(\text{erro tipo II}) \rightarrow$  probabilidade de cometer o erro tipo II

Como consequência:

$1 - \alpha$  é a probabilidade de não cometer o erro tipo I  
(probabilidade de não rejeitar  $H_0$  verdadeira)

$1 - \beta$  é a probabilidade de não cometer o erro tipo II  
(probabilidade de rejeitar  $H_0$  falsa)

A probabilidade  $1 - \beta$  é denominada **poder do teste**.

Réu	Decisão do juiz	
	Não condenar	Condenar
Inocente	Acerto	Erro 1
Culpado	Erro 2	Acerto

$H_0$	Decisão	
	Não rejeitar	Rejeitar
Verdadeira	Acerto	Erro Tipo I
Falsa	Erro Tipo II	Acerto

**Erro Tipo I:** Declarar diferença quando ela não existe

**Erro Tipo II:** Não declarar diferença quando ela existe

### Importante!!!

- ⇒ As duas taxas de erro  $\alpha$  e  $\beta$  estão relacionadas negativamente, de modo que a redução de  $\alpha$  implica no aumento de  $\beta$  e vice-versa.
- ⇒ Para que os testes de hipótese tenham validade, é necessário que sejam delineados de modo a minimizar os erros de conclusão.
- ⇒ O único meio de reduzir ambos os tipos de erro é aumentando o tamanho da amostra, o que nem sempre é viável.
- ⇒ Em geral, a preocupação está voltada para o erro tipo I, pois na maioria dos casos ele é considerado o mais grave.
- ⇒ A probabilidade de ocorrência do erro tipo I ( $\alpha$ ) é chamada de nível de significância do teste.

### 3. Estatística do teste

#### Estatística do teste

Para testar hipóteses a respeito do parâmetro  $\mu$ , utilizamos a variável aleatória  $T$  que tem distribuição **t de Student**.

$$T = \frac{\hat{\theta} - \theta}{S(\hat{\theta})} \sim t(v)$$

onde:

$\theta$  é o parâmetro estimado

$\hat{\theta}$  é estimador do parâmetro

$S(\hat{\theta})$  é o estimador do erro padrão do  $\hat{\theta}$

$v$  é o número de graus de liberdade

### 1. Comparação de uma média ( $\mu$ ) com um valor padrão ( $\mu_0$ )

**Pressuposição:** A variável em estudo tem distribuição normal

**Hipótese sob verificação:**  $H_0 : \mu = \mu_0$

**Estatística do teste:**  $T = \frac{\hat{\theta} - \theta}{S(\hat{\theta})} \sim t(v)$

onde:

$$\theta = \mu = \mu_0$$

$$\hat{\theta} = \bar{X}$$

$$S(\hat{\theta}) = S(\bar{X}) = \frac{S}{\sqrt{n}}$$

$$v = n - 1$$

$$T = \frac{\bar{X} - \mu_0}{\frac{S}{\sqrt{n}}}$$

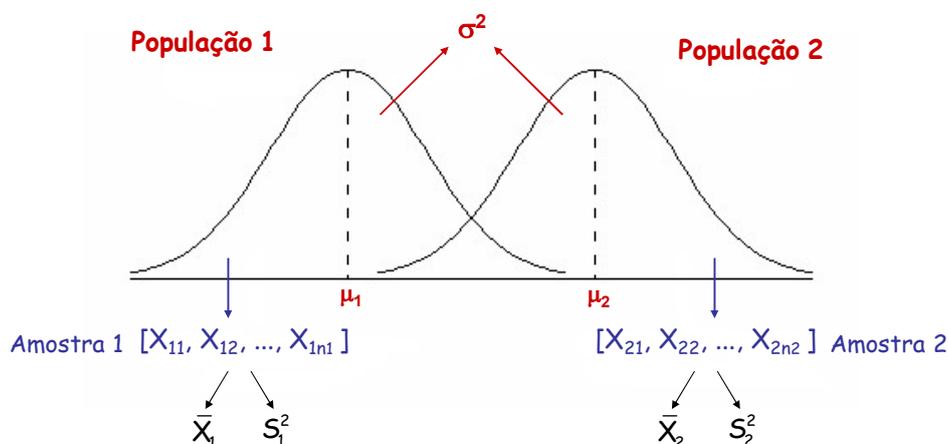
Valor que deve ser calculado na amostra

**Estatística do teste**

### 2. Comparação entre duas médias ( $\mu_1$ e $\mu_2$ )

**Pressuposições para o uso da estatística T:**

- A variável em estudo tem distribuição normal.
- As variâncias das populações são iguais.
- As amostras retiradas das populações são independentes.



## 2. Comparação entre duas médias ( $\mu_1$ e $\mu_2$ )

### Pressuposições:

A variável em estudo tem distribuição normal

As variâncias das populações são iguais

As amostras retiradas das populações são independentes

Hipótese sob verificação:  $H_0 : \mu_1 - \mu_2 = 0$

Estatística do teste:  $T = \frac{\hat{\theta} - \theta}{S(\hat{\theta})} \sim t(v)$

Valor que deve ser calculado na amostra

onde:

$$\theta = \mu_1 - \mu_2 = 0$$

$$\hat{\theta} = \bar{X}_1 - \bar{X}_2$$

$$S(\hat{\theta}) = S(\bar{X}_1 - \bar{X}_2) = \sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right) S^2}$$

$$v = (n_1 - 1) + (n_2 - 1)$$

$$T = \frac{\hat{\theta} - 0}{S(\hat{\theta})} = \frac{\hat{\theta}}{S(\hat{\theta})}$$

$$T = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right) S^2}}$$

Estatística do teste

## 1. Comparação de uma média ( $\mu$ ) com um valor padrão ( $\mu_0$ )

Pressuposição: A variável em estudo tem distribuição normal

Hipótese sob verificação:  $H_0 : \mu = \mu_0$

Estatística do teste:  $T = \frac{\bar{X} - \mu_0}{\frac{S}{\sqrt{n}}} \sim t(v)$

## 2. Comparação entre duas médias ( $\mu_1$ e $\mu_2$ )

### Pressuposições:

A variável em estudo tem distribuição normal

As variâncias das populações são iguais

As amostras retiradas das populações são independentes

Hipótese sob verificação:  $H_0 : \mu_1 - \mu_2 = 0$

Estatística do teste:  $T = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right) S^2}} \sim t(v)$

Prof. Clause Piana

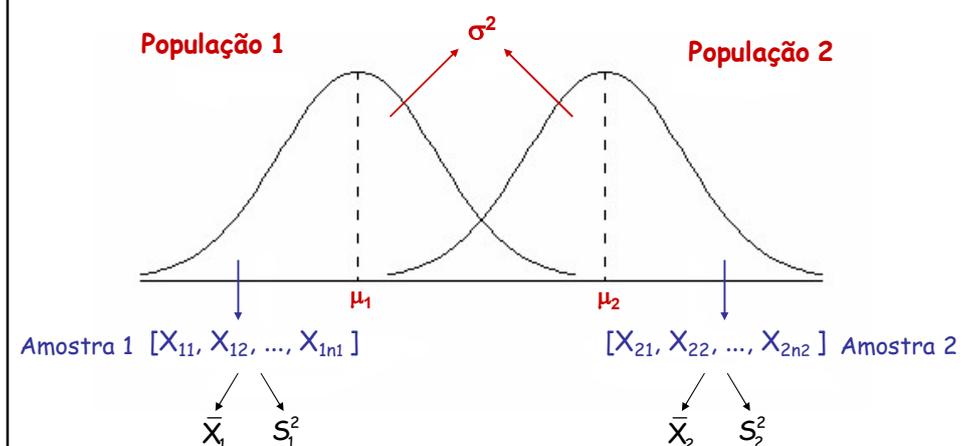
30

#### 4. Critério de decisão

Comparação de duas médias  $\begin{cases} H_0 : \mu_1 - \mu_2 = 0 \\ H_A : \mu_1 - \mu_2 \neq 0 \end{cases}$

Possíveis conclusões para  $H_0$

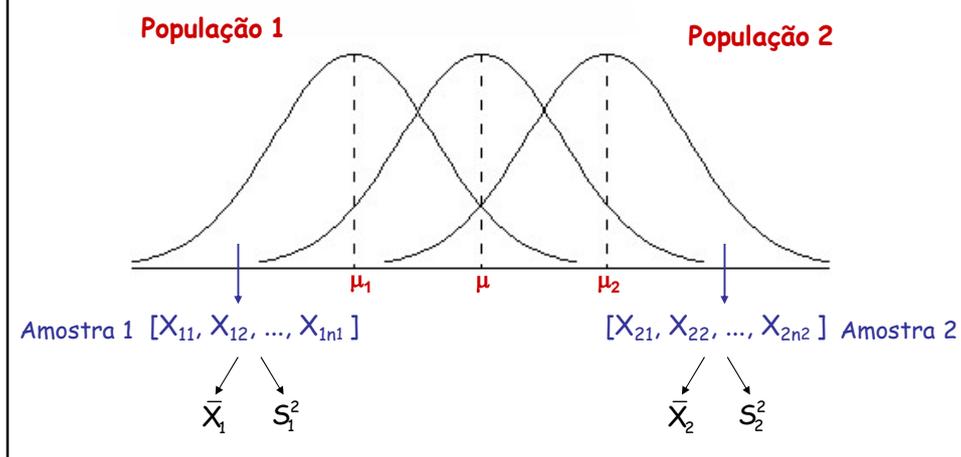
1.  $H_0$  é falsa: temos duas populações com médias diferentes



2.  $H_0$  é verdadeira: as médias das duas populações não diferem

$$H_0 : \mu_1 = \mu_2 = \mu$$

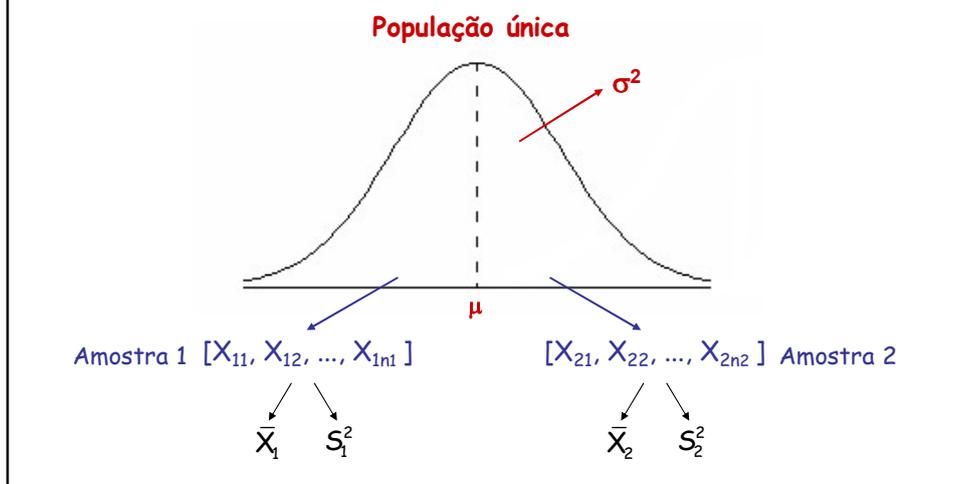
Portanto, em vez de duas populações, temos apenas uma

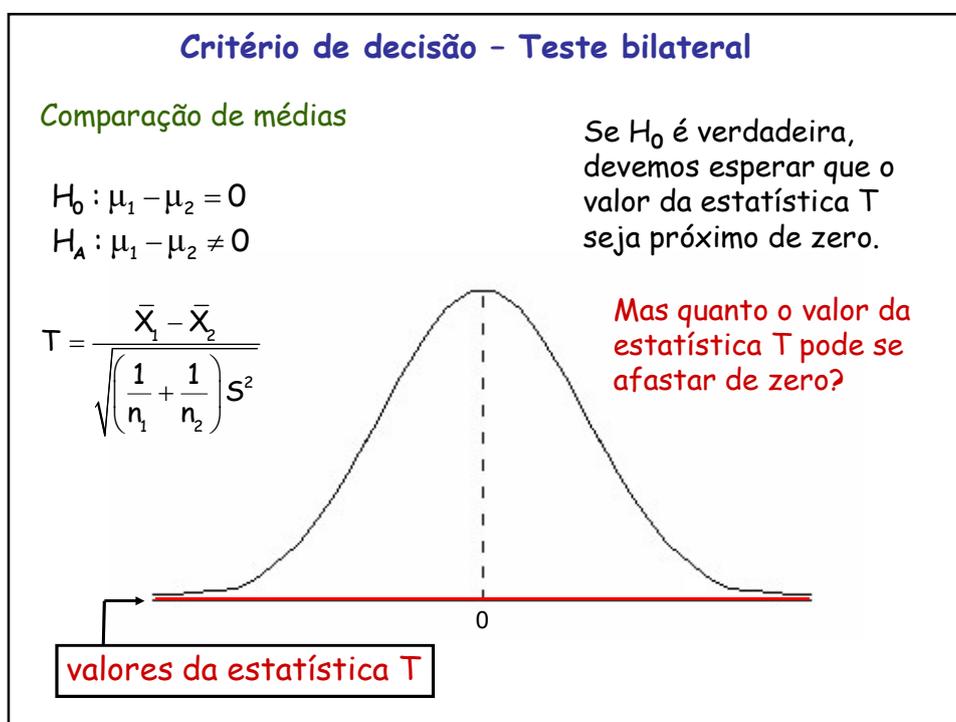
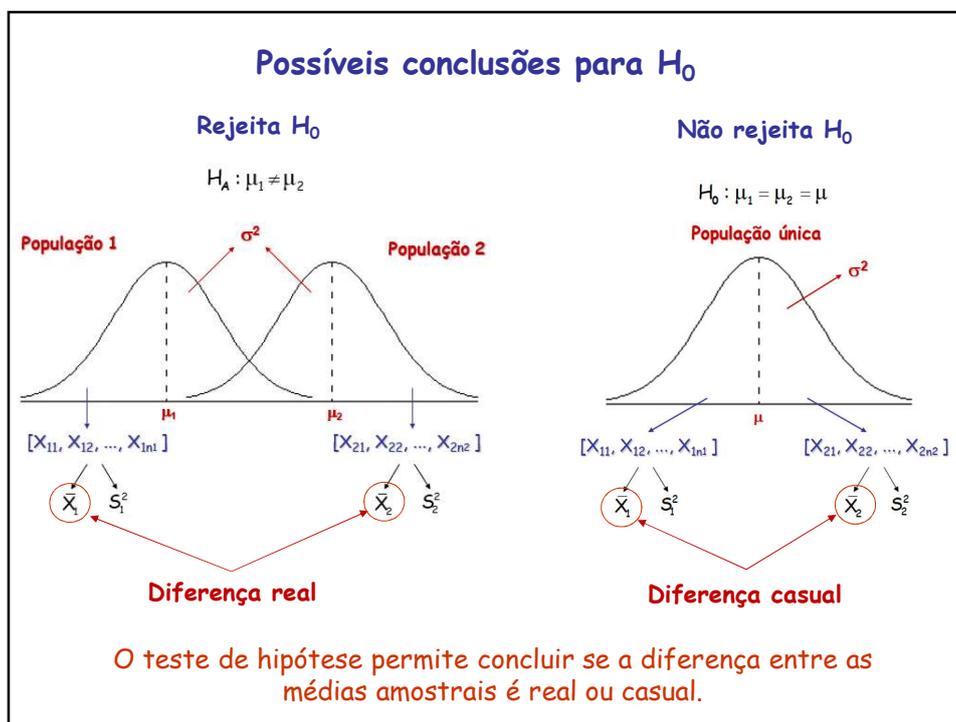


2.  $H_0$  é verdadeira: as médias das duas populações não diferem

$$H_0 : \mu_1 = \mu_2 = \mu$$

Portanto, em vez de duas populações, temos apenas uma





### Comparação de médias

$$H_0 : \mu_1 - \mu_2 = 0$$

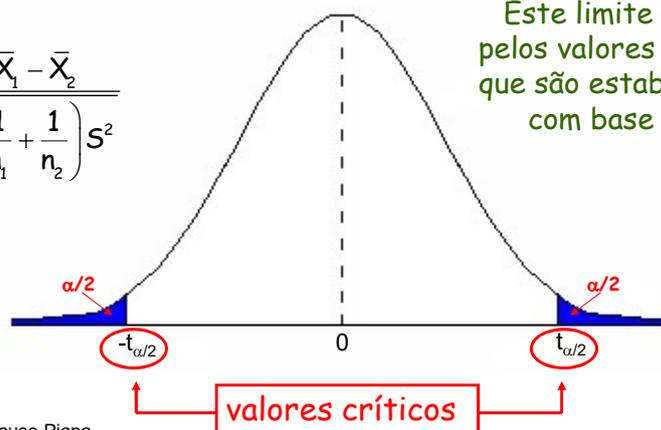
$$H_A : \mu_1 - \mu_2 \neq 0$$

$$T = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right) S^2}}$$

Se  $H_0$  é verdadeira, devemos esperar que o valor da estatística T seja próximo de zero.

Mas quanto o valor da estatística T pode se afastar de zero?

Este limite é dado pelos valores críticos, que são estabelecidos com base no  $\alpha$ .



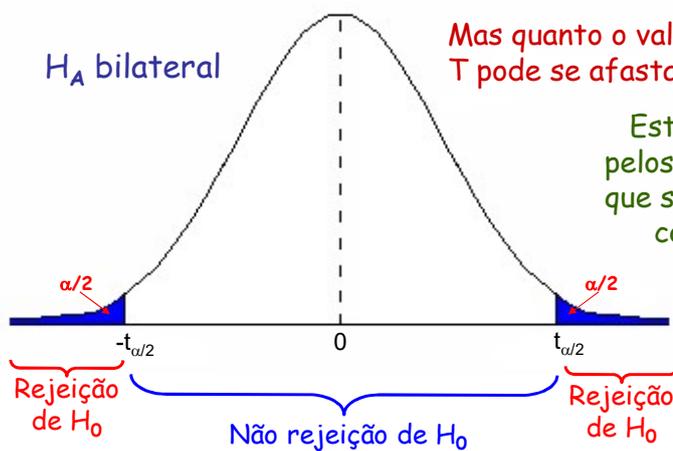
Profa. Clause Piana

37

### $H_A$ bilateral

Mas quanto o valor da estatística T pode se afastar de zero?

Este limite é dado pelos valores críticos, que são estabelecidos com base no  $\alpha$ .



⇒ Se o valor  $t$  (calculado na amostra) estiver entre  $-t_{\alpha/2}$  e  $t_{\alpha/2}$ , não temos motivos para rejeitar  $H_0$

⇒ Se o valor  $t$  (calculado na amostra) for menor que  $-t_{\alpha/2}$  ou maior que  $t_{\alpha/2}$ , rejeitamos  $H_0$

### Como tomar a decisão a respeito de $H_0$ ?

Para decidir comparamos o valor da estatística T (calculado) com o valor crítico (tabelado):

⇒ Rejeitamos  $H_0$ , ao nível  $\alpha$ , se o valor da estatística, em módulo, for maior que o valor crítico:

$$|t| > t_{\alpha/2} \Rightarrow t \text{ é atípico}$$

⇒ Não temos motivos suficientes para rejeitar  $H_0$ , ao nível  $\alpha$ , se o valor da estatística, em módulo, for menor que o valor crítico:

$$|t| \leq t_{\alpha/2} \Rightarrow t \text{ é típico}$$

Profa. Clause Piana

39

### Critério de decisão - Teste unilateral

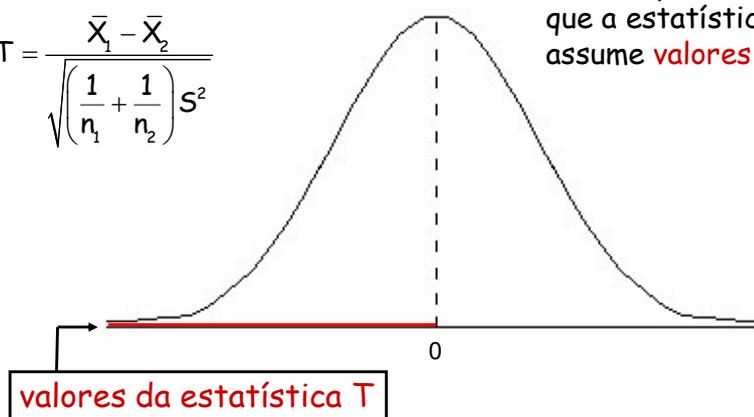
#### Comparação de médias

$$H_0 : \mu_1 - \mu_2 = 0$$

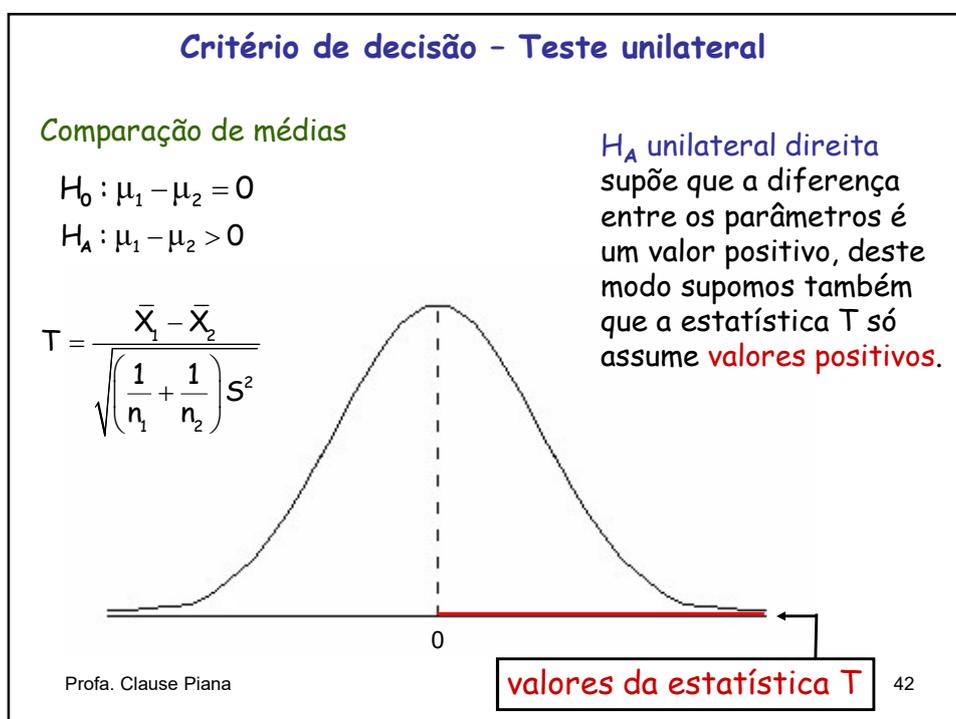
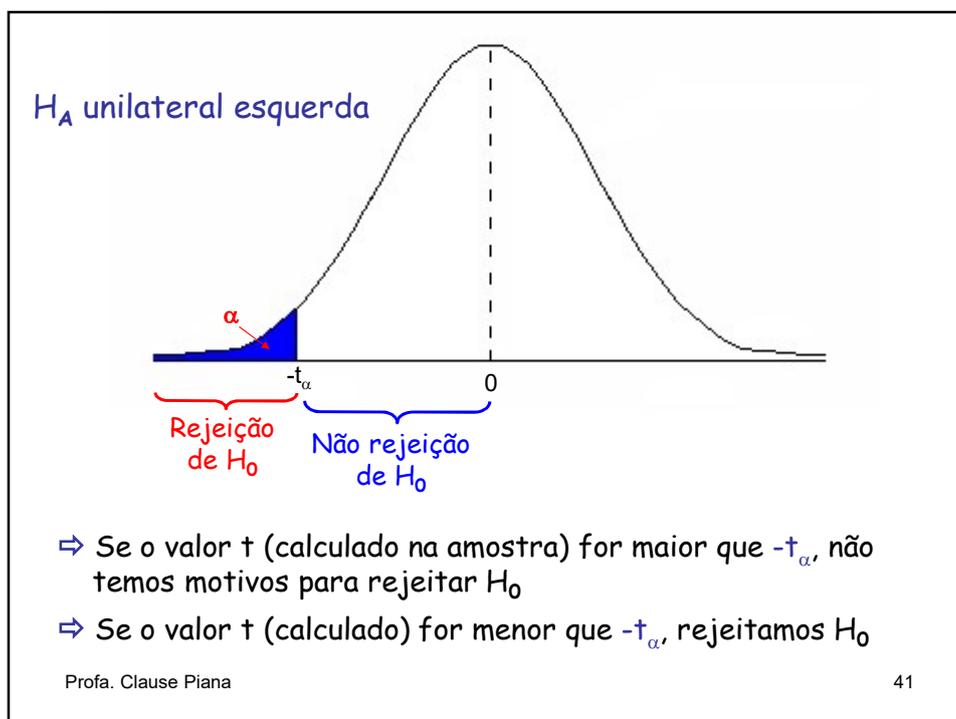
$$H_A : \mu_1 - \mu_2 < 0$$

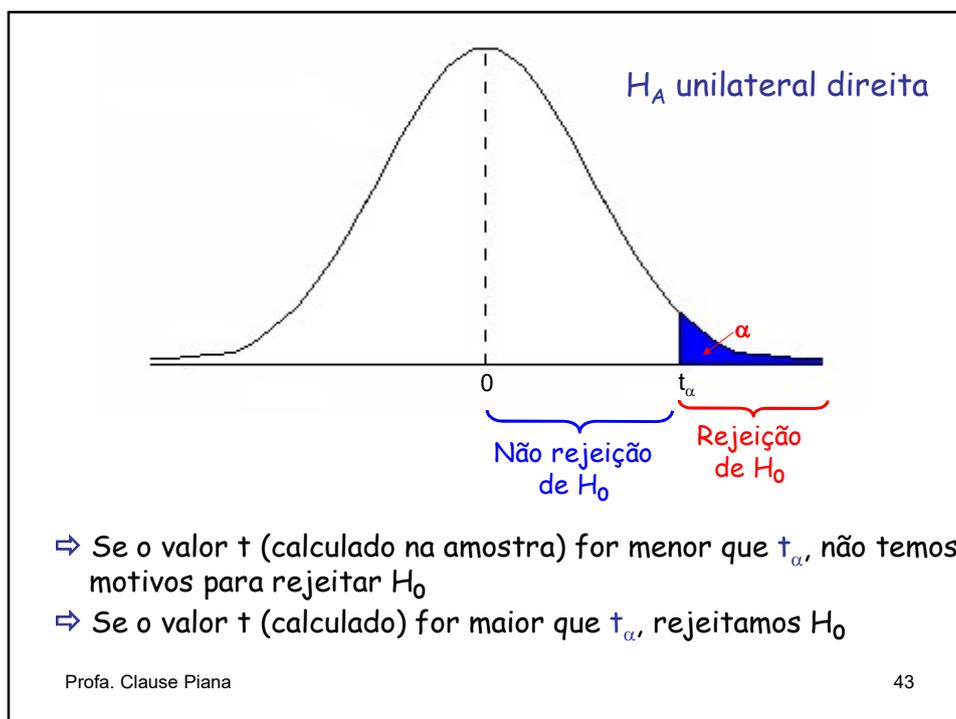
$$T = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right) S^2}}$$

$H_A$  unilateral esquerda supõe que a diferença entre os parâmetros é um valor negativo, deste modo supomos também que a estatística T só assume **valores negativos**.



valores da estatística T





### Como tomar a decisão a respeito de $H_0$ ?

Para decidir comparamos o valor da estatística  $T$  (calculado) com o valor crítico (tabelado):

- ⇒ Rejeitamos  $H_0$ , ao nível  $\alpha$ , se o valor da estatística, em módulo, for maior que o valor crítico:

$$|t| > t_\alpha \Rightarrow t \text{ é atípico}$$

- ⇒ Não temos motivos suficientes para rejeitar  $H_0$ , ao nível  $\alpha$ , se o valor da estatística, em módulo, for menor que o valor crítico:

$$|t| \leq t_\alpha \Rightarrow t \text{ é típico}$$

Graus de Liberdade (v)	Limites bilaterais: $P( t  > t_{\alpha/2})$							
	Nível de Significância ( $\alpha$ )							
	0,50	0,20	0,10	0,05	0,025	0,02	0,01	0,005
1	1,000	3,078	6,314	12,706	25,542	31,821	63,657	127,320
2	0,916	1,895	2,920	4,303	6,205	6,965	9,925	14,089
3	0,715	1,638	2,353	3,183	4,177	4,541	5,841	7,453
4	0,741	1,533	2,132	2,776	3,495	3,747	4,604	5,598
5	0,727	1,476	2,015	2,571	3,163	3,365	4,032	4,773
6	0,718	1,440	1,943	2,447	2,969	3,143	3,707	4,317
7	0,711	1,415	1,895	2,365	2,841	2,998	3,500	4,029
8	0,706	1,397	1,860	2,306	2,752	2,896	3,355	3,833
9	0,703	1,383	1,833	2,262	2,685	2,821	3,250	3,690
10	0,700	1,372	1,813	2,228	2,634	2,764	3,169	3,581
11	0,697	1,363	1,796	2,201	2,503	2,718	3,106	3,497
12	0,695	1,356	1,782	2,179	2,560	2,681	3,055	3,428
13	0,694	1,350	1,771	2,160	2,533	2,650	3,012	3,373
14	0,692	1,345	1,761	2,145	2,510	2,624	2,977	3,326
15	0,691	1,341	1,753	2,132	2,490	2,602	2,947	3,286
16	0,690	1,337	1,746	2,120	2,473	2,583	2,921	3,252
17	0,689	1,333	1,740	2,110	2,458	2,567	2,898	3,223
18	0,688	1,330	1,734	2,101	2,445	2,552	2,878	3,197
19	0,688	1,328	1,729	2,093	2,433	2,539	2,861	3,174
20	0,687	1,325	1,725	2,086	2,423	2,528	2,845	3,153
21	0,686	1,323	1,721	2,080	2,414	2,518	2,831	3,135
22	0,686	1,321	1,717	2,074	2,406	2,508	2,819	3,119
23	0,685	1,319	1,714	2,069	2,398	2,500	2,807	3,104
24	0,685	1,318	1,711	2,064	2,391	2,492	2,797	3,091
25	0,684	1,316	1,708	2,060	2,385	2,485	2,787	3,078
26	0,684	1,315	1,706	2,056	2,379	2,479	2,779	3,067
27	0,684	1,314	1,703	2,052	2,373	2,473	2,771	3,057
28	0,683	1,313	1,701	2,048	2,369	2,467	2,763	3,047
29	0,683	1,311	1,699	2,045	2,364	2,462	2,756	3,038
30	0,683	1,310	1,697	2,042	2,360	2,457	2,750	3,030
40	0,681	1,303	1,684	2,021	2,329	2,423	2,705	2,971
60	0,679	1,296	1,671	2,000	2,299	2,390	2,660	2,915
120	0,677	1,289	1,658	1,980	2,270	2,358	2,617	2,860
...	0,674	1,282	1,645	1,960	2,241	2,326	2,576	2,807
Graus de Liberdade (v)	Limites unilaterais: $P( t  > t_{\alpha})$							
	Nível de Significância ( $\alpha$ )							
	0,25	0,10	0,05	0,025	0,0125	0,01	0,005	0,0025

## Critério de decisão

A regra de decisão a respeito de  $H_0$  pode ser estabelecida com base num **valor crítico**:

⇒ Se o teste for **bilateral** (hipótese alternativa bilateral), o valor crítico será:

$t_{\alpha/2(v)}$ : valor da estatística T que delimita a área  $\alpha/2$ , para  $v$  graus de liberdade (Tabela: limites bilaterais)

⇒ Se o teste for **unilateral** (hipótese alternativa unilateral), o valor crítico será:

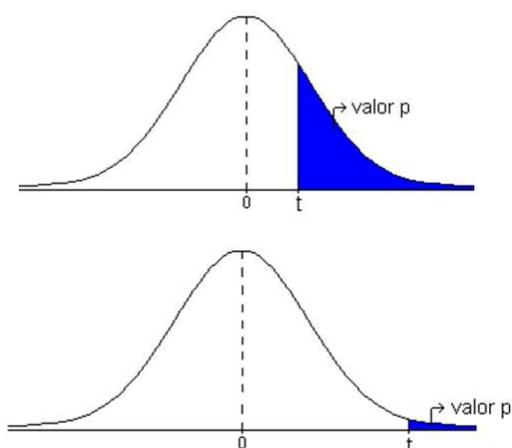
$t_{\alpha(v)}$ : valor da estatística T que delimita a área  $\alpha$ , para  $v$  graus de liberdade (Tabela: limites unilaterais)

### Nível de significância do teste

- ⇒ A probabilidade  $\alpha$  de se cometer um erro tipo I é um valor arbitrário e recebe o nome de nível de significância do teste
- ⇒ O resultado da amostra é tanto mais significativo para rejeitar  $H_0$  quanto menor for esse nível  $\alpha$ . Ou seja, quanto menor for  $\alpha$ , menor é a probabilidade de se obter uma amostra com estatística pertencente a região crítica, sendo pouco verossímil a obtenção de uma amostra da população para a qual  $H_0$  seja verdadeira.
- ⇒ A probabilidade do erro tipo II, por outro lado, na maioria dos casos, não pode ser calculada, pois a hipótese alternativa usualmente especifica um conjunto de valores para o parâmetro. Esta probabilidade  $\beta$  só pode ser calculada se for especificado um valor alternativo para  $\mu$ .
- ⇒ Usualmente, o valor  $\alpha$  é fixado em 5%, 1% e 0,1%. A fixação do valor de  $\alpha$  envolve uma questionável arbitrariedade. Neste sentido, há um modo alternativo de se proceder, que será exposto a seguir.

### Outro critério de decisão

**Valor p:** probabilidade de que seja obtido um valor  $t$  maior que aquele **calculado na amostra**, dado que  $H_0$  é verdadeira



### Outro critério de decisão

**Valor p:** probabilidade de que seja obtido um valor de T maior do que aquele calculado na amostra, dado que  $H_0$  é verdadeira

#### Como tomar a decisão a respeito de $H_0$ ?

- ⇒ Se o **valor p** for maior que  $\alpha$ : **não rejeitamos** a hipótese nula pois **t é típico** ou está em uma região de alta probabilidade
- ⇒ Se o **valor p** for menor que  $\alpha$ : **rejeitamos** a hipótese nula pois **t é atípico** ou está em uma região de baixa probabilidade

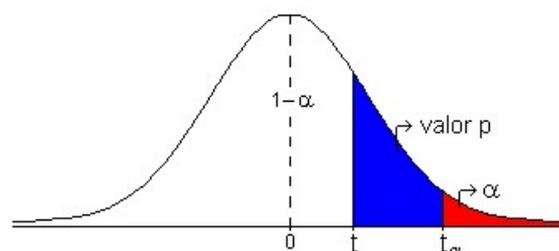
Profa. Clause Piana

49

Se  $p > \alpha \Rightarrow t$  é típico



Não rejeitamos a hipótese de nulidade



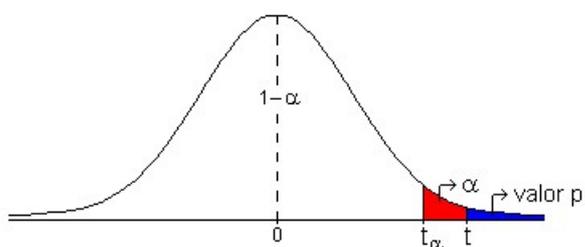
Profa. Clause Piana

50

Se  $p < \alpha \Rightarrow t$  é atípico



Rejeitamos a hipótese de nulidade



Profa. Clause Piana

51

### Decisão sobre $H_0$

**Critério 1:** comparar o módulo de  $t$  calculado com o  $t_\alpha$  (tabelado)

$|t| > t_\alpha \rightarrow$  Rejeição de  $H_0 \rightarrow t$  grande é evidência contra  $H_0$

$|t| \leq t_\alpha \rightarrow$  Não rejeição de  $H_0 \rightarrow t$  pequeno é evidência a favor de  $H_0$

**Critério 2:** comparar o valor  $p$  com o  $\alpha$

$p < \alpha \rightarrow$  Rejeição de  $H_0 \rightarrow$  valor  $p$  pequeno é evidência contra  $H_0$

$p \geq \alpha \rightarrow$  Não rejeição de  $H_0 \rightarrow$  valor  $p$  grande é evidência a favor de  $H_0$

### Poder do teste

⇒ Vimos que  $1-\beta$  é a probabilidade de não cometer o erro tipo II, ou seja, é a probabilidade de rejeitar  $H_0$  quando ela é falsa.

⇒ Esta probabilidade  $1-\beta$  é denominada **poder do teste**.

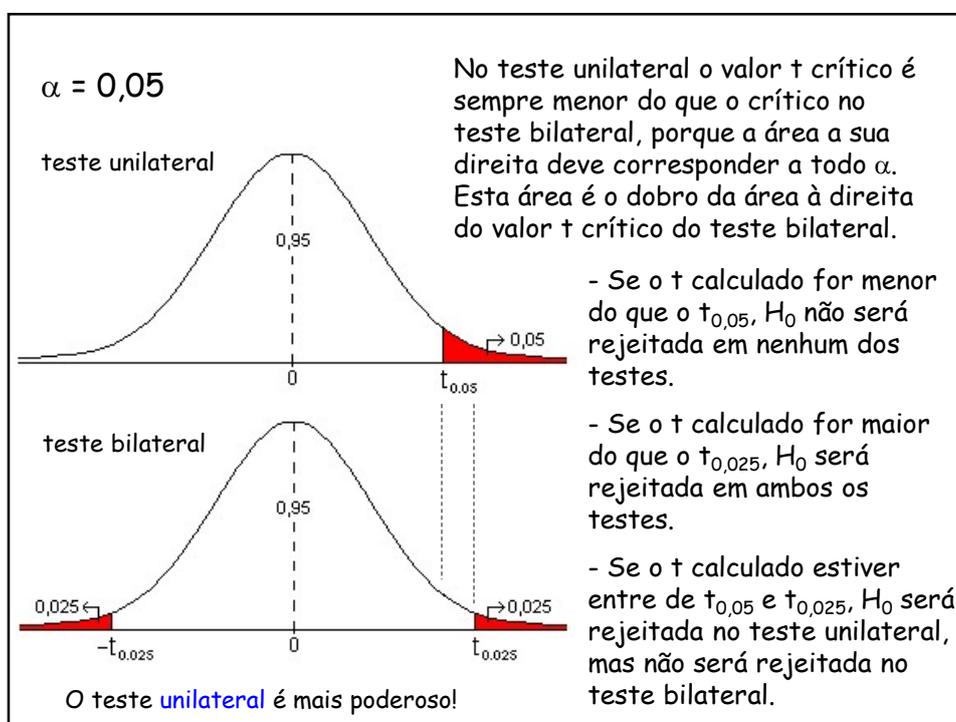
⇒ Assim, quanto maior a probabilidade de um teste rejeitar  $H_0$  mais poderoso ele é, pois tem maior probabilidade de detectar a diferença entre as médias.

Qual dos testes é mais poderoso: **unilateral** ou **bilateral**?

O mais poderoso é aquele que, para um mesmo valor  $\alpha$ , tem maior probabilidade rejeitar  $H_0$ .

Profa. Clause Piana

53



### Considerações finais

- ⇒ Os intervalos de confiança e os testes de hipóteses bilaterais são procedimentos estatísticos relacionados.
- ⇒ Se forem utilizados para analisar os mesmos dados, ao mesmo nível de significância, devem conduzir aos mesmos resultados.
- ⇒ O intervalo de confiança para uma média está relacionado com o teste de hipóteses que compara uma média com um padrão.  
 $H_0$  não rejeitada  $\Leftrightarrow$  valor padrão está coberto pelo intervalo
- ⇒ O intervalo de confiança para a diferença entre duas médias está relacionado com o teste de hipóteses que compara duas médias.  
 $H_0$  não rejeitada  $\Leftrightarrow$  zero está coberto pelo intervalo

### Intervalo de confiança para uma média ( $\mu$ )

$$IC (\mu; 1-\alpha): \bar{X} \pm t_{\alpha/2} \frac{S}{\sqrt{n}}$$

### Estatística T para comparação de média ( $\mu$ ) com valor padrão ( $\mu_0$ )

$$T = \frac{\bar{X} - \mu_0}{\frac{S}{\sqrt{n}}} \quad \text{Valor crítico: } t_{\alpha/2}$$

- ⇒ Se no teste de hipóteses, ao nível de 1% de significância, rejeitamos  $H_0$ , significa que a diferença entre a média e o valor padrão não é zero, ou seja, média e valor padrão são diferentes.
- ⇒ Construindo o intervalo de confiança para  $\mu$ , ao nível de 99%, devemos esperar que o valor padrão ( $\mu_0$ ) esteja fora do intervalo. Caso contrário, os resultados seriam contraditórios.

Intervalo de confiança para a diferença entre duas médias ( $\mu_1 - \mu_2$ )

$$\text{IC} (\mu_1 - \mu_2; 1-\alpha): \bar{X}_1 - \bar{X}_2 \pm t_{\alpha/2} \sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right) S^2}$$

Estatística T para a comparação entre duas médias ( $\mu_1$  e  $\mu_2$ )

$$T = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right) S^2}} \quad \text{Valor crítico: } t_{\alpha/2}$$

⇒ Se no teste de hipóteses, ao nível de 5% de significância, não rejeitamos  $H_0$ , significa que a diferença entre as duas médias é zero, ou seja, as médias são iguais.

⇒ Construindo o intervalo de confiança, ao nível de 95%, para a diferença entre as médias ( $\mu_1 - \mu_2$ ), devemos esperar que o valor zero seja coberto pelo intervalo.

Exemplos resolvidos

**Exemplo 1.**

O fabricante de uma certa marca de aparelhos eletrônicos informou que a potência média dos seus aparelhos é de 27 microwatts. O gerente de uma loja que vende os aparelhos utiliza uma amostra de 15 aparelhos para checar se a informação do fabricante é verdadeira. Os valores (em microwatts) obtidos para a amostra foram os seguintes: 26,7; 25,8; 24,0; 24,9; 26,4; 25,9; 24,4; 21,7; 24,1; 25,9; 27,3; 26,9; 27,3; 24,8 e 23,6.

Utilize um teste de hipótese, ao nível de 5% de significância, e verifique qual foi a conclusão do gerente.

**Resolução:** **Variável em estudo:** X = potência de aparelhos eletrônicos (microwatts)

**Pressuposição:** A variável em estudo tem distribuição normal.

**Hipóteses**  $\left\{ \begin{array}{l} H_0 : \mu = \mu_0 = 27 \\ H_A : \mu \neq \mu_0 = 27 \end{array} \right.$  **Taxa de erro tipo I:**  $\alpha = 0,05$

**Cálculo da estatística do teste**

$$\begin{array}{l} \bar{x} = 25,31 \quad n = 15 \\ s = 1,589 \end{array} \quad T = \frac{\bar{X} - \mu_0}{\frac{S}{\sqrt{n}}} \rightarrow t = \frac{25,31 - 27}{\frac{1,589}{\sqrt{15}}} = -4,12$$

**Decisão:**  $|t| = -4,12 > t_{\alpha/2(14)} = 2,145 \rightarrow$  Rejeitamos  $H_0$

**Conclusão:** Concluímos, ao nível de 5% de significância, que a verdadeira potência média de aparelhos eletrônicos desta marca difere de 27 microwatts. Portanto, a informação do fabricante não é verdadeira.

**Exemplo 2.**

Um pesquisador da área de computação está investigando a utilidade de duas diferentes linguagens de programação (A e B) na melhoria das tarefas computacionais.

Trinta programadores experientes, familiarizados com ambas as linguagens, foram divididos aleatoriamente em dois grupos. Cada grupo codificou uma função padrão em uma das linguagens e os tempos de codificação da função (em minutos) foram registrados.

As medidas de cada grupo são apresentadas no quadro abaixo.

	n	Média	Variância
Linguagem A	15	17,5	2,31
Linguagem B	15	20,5	3,02

Utilizando um teste de hipótese, verifique se, em média, o tempo de codificação da função padrão difere entre as linguagens.

**Resolução:** Variável em estudo:  $X$  = tempo de codificação da função (minutos)

- 1. Pressuposições:**
1. A variável em estudo tem distribuição normal.
  2. As variâncias das populações são homogêneas.
  3. As amostras retiradas das populações são independentes.

**2. Hipóteses estatísticas:**

$$\begin{cases} H_0 : \mu_1 = \mu_2 \\ H_A : \mu_1 \neq \mu_2 \end{cases}$$

**3. Taxa de erro tipo I:**  $\alpha = 0,05$

**4. Estatística do teste:**

$$t = \frac{17,5 - 20,5}{\sqrt{\left(\frac{1}{15} + \frac{1}{15}\right) 2,67}} = \frac{-3}{0,5967} = -5,028$$

**5. Decisão e conclusão:**

$$\alpha = 0,05$$

$$v = (15 - 1) + (15 - 1) = 28$$

$$t_{\alpha/2(28)} = 2,048$$

$$|-5,028| > t_{\alpha/2(28)} = 2,048$$

Sendo o  $t$  calculado maior que o crítico, rejeitamos  $H_0$

Concluímos, ao nível de 5% de significância, que o tempo médio de programadores que usam a linguagem A difere do tempo médio de programadores que usam a linguagem B. Portanto, a linguagem A, em média, é mais eficiente que a linguagem B na execução das tarefas computacionais.

## Bibliografia

BUSSAB, W.O.; MORETTIN, P.A. **Estatística Básica**. São Paulo: Saraiva. 2006. 526p.

FERREIRA, D.F. **Estatística Básica**. Lavras: Editora UFLA, 2005, 664p.

MLODINOW, L. **O andar do bêbado. Como o acaso determina nossas vidas**. Rio de Janeiro: Editora Zahar, 2009, 264p.

SILVEIRA JÚNIOR, P., MACHADO, A.A., ZONTA, E.P., SILVA, J.B. da Curso de Estatística v.1, Pelotas: Universidade Federal de Pelotas, 1989. 135p.

Sistema Galileu de Educação Estatística. Disponível em: <http://www.galileu.esalq.usp.br/topico.html>