

# Testes t para comparação de médias de dois grupos independentes

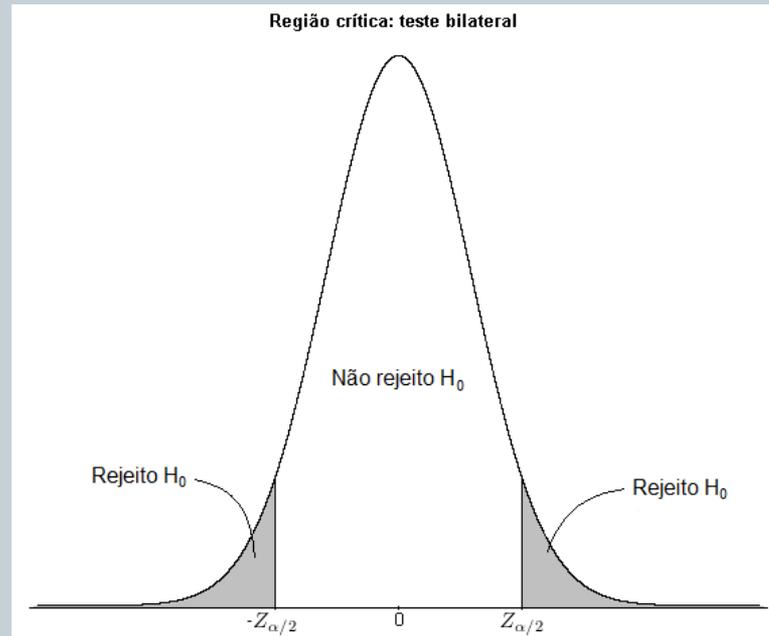


**Acadêmicas do curso de Zootecnia -  
Aline Cristina Berbet Lopes  
Amanda da Cruz Leinoski  
Larissa Ceccon**

# Teste de Hipóteses



- Os testes de hipótese constituem uma forma de inferência estatística. Hipóteses são afirmações sobre parâmetros populacionais e são testadas para ver se são consideradas verdadeiras ou não.



# Teste $t$ de Student



- 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
- Este trabalho tem como objetivo explicar e exemplificar a comparação de duas amostras independentes.

# Por que utilizar a comparação de duas amostras independentes?



- Este teste se aplica a planos amostrais onde se deseja comparar dois grupos independentes . Esses grupos podem ter sido formados de duas maneiras diferentes:
- a) Extraiu-se uma amostra da população A e outra amostra da população B;
- b) Indivíduos da mesma população foram alocados aleatoriamente a um dos dois tratamentos em estudo.

# Quando utilizar?

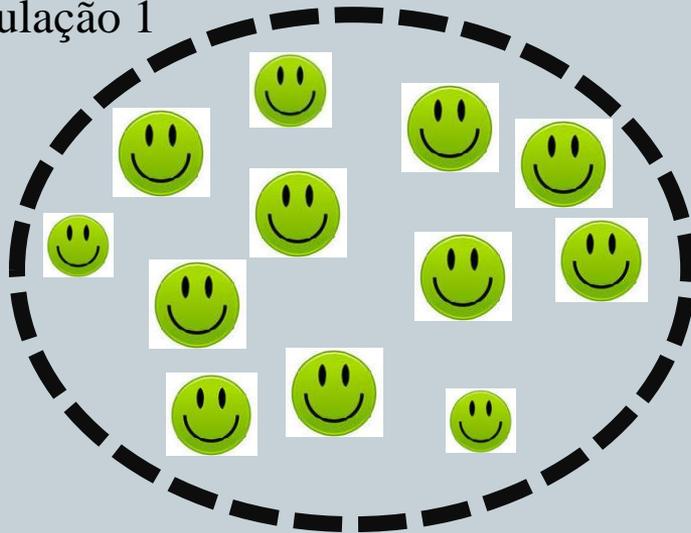


- Quando o objetivo é comparar duas populações quanto a uma variável quantitativa, é muito comum que os pesquisadores não conheçam os parâmetros de nenhuma delas, isto é, sejam desconhecidas as médias ( $\mu$ ) e também os desvios padrão ( $\sigma$ ) populacionais;
- Dessa forma, muitos estudos biológicos são realizados com duas amostras independentes de indivíduos, denominadas grupo experimental e grupo controle, respectivamente.

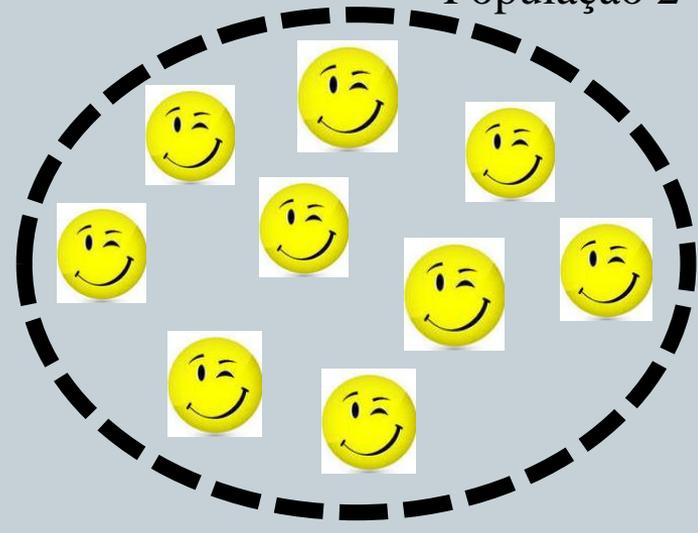


- Suponha duas populações distintas, compostas por um número grande de indivíduos.

População 1



População 2





- Seja  $X$  uma variável aleatória (contínua) de interesse...

Na população  
1, a média de  
 $X$  é  $\mu_1$

Na população  
2, a média de  
 $X$  é  $\mu_2$

- O objetivo seria comparar as médias populacionais  $\mu_1$  e  $\mu_2$ .
- O teste de hipótese seria:  $H_0: \mu_1 = \mu_2$

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

# Exemplo 1



- Um estudo objetivou analisar a associação entre diversas variáveis com a síndrome metabólica (SM) em indivíduos de origem japonesa, com mais de 30 anos de idade, residentes em um município do interior de São Paulo.

❖ População 1:  
indivíduos com  
SM

❖ População 2:  
indivíduos sem  
SM

# Pergunta-se...



- Indivíduos com SM e sem SM possuem, em média, valores iguais para a pressão arterial sistólica (PAS)?
  - ➔  $\mu_1$  é a média da PAS na população 1;
  - ➔  $\mu_2$  é a média da PAS na população 2.
- Testando as hipóteses:  $H_0: \mu_1 = \mu_2$   
 $H_1: \mu_1 \neq \mu_2$
- Lembrando que  $\sigma_1$  é o desvio padrão da PAS na população 1, e  $\sigma_2$  é o desvio padrão da PAS na população 2;
- Pressupondo que  $\sigma = \sigma_1 = \sigma_2$ .

# Exemplo 1



- De cada uma dessas populações retira-se uma amostra de tamanho  $n$ .

❖ População 1:  
amostra de  
tamanho  $n_1$

❖ População 2:  
amostra de  
tamanho  $n_2$

# Exemplo 1



- Se a variável de interesse segue uma distribuição próxima de uma curva normal em ambas populações,

$$T = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{S_p^2 \left( \frac{1}{n_1} + \frac{1}{n_2} \right)}}$$

- Segue uma distribuição t Student com  $n_1 + n_2 - 2$  graus de liberdade.

# Exemplo 1



- Se a hipótese  $H_0$  é verdadeira  $T_0 = \frac{(\bar{x}_1 - \bar{x}_2)}{\sqrt{S_p^2 \left( \frac{1}{n_1} + \frac{1}{n_2} \right)}}$
- Portanto, testamos  $H_0$  encontrando um valor  $t_0$  para  $T_0$ , com base em amostras de tamanho  $n_1$  e  $n_2$  retiradas das respectivas populações  $t_0 = \frac{(\bar{x}_1 - \bar{x}_2)}{\sqrt{S_p^2 \left( \frac{1}{n_1} + \frac{1}{n_2} \right)}}$

# Exemplo 1



- Mas o que seria  $S_p^2$ ? Estamos pressupondo que a variável de interesse tem o desvio padrão igual nas duas populações.
- Entretanto, isso não significa que os desvios padrão  $S_1$  e  $S_2$  que encontraremos nas amostras serão iguais.
- Assim temos que  $S_p^2$  é uma estimativa da variância populacional  $\sigma^2$ .
- $S_p^2$  é uma média entre as variâncias amostrais  $S_1^2$  e  $S_2^2$ , ponderada pelos respectivos graus de liberdade  $n_1 - 1$  e  $n_2 - 1$ .

# Exemplo 1



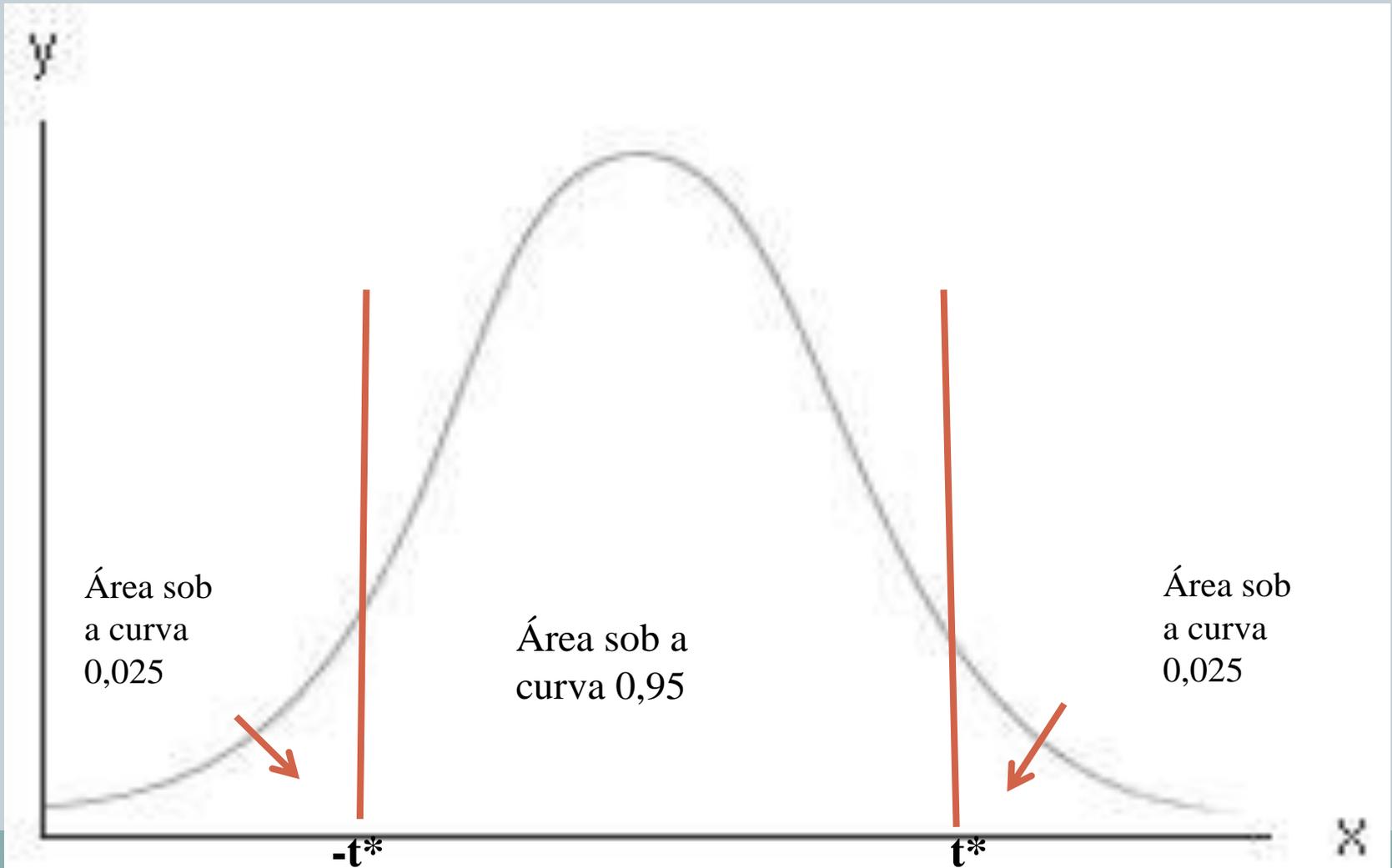
- Então teremos  $S_p^2 = \frac{(n_1-1)S_p^2 + (n_2-1)S_p^2}{n_1+n_2-2}$ .
- Lembrando que se  $H_0$  é verdadeira,  $t_0$  é um valor de uma variável aleatória que segue uma distribuição t Student com  $n_1 + n_2 - 2$  graus de liberdade.
- $t_0 = \frac{(\bar{x}_1 - \bar{x}_2)}{\sqrt{S_p^2 \left( \frac{1}{n_1} + \frac{1}{n_2} \right)}}$  em que  $S_p^2 = \frac{(n_1-1)S_p^2 + (n_2-1)S_p^2}{n_1+n_2-2}$

# Exemplo 1



- Seja o tamanho da amostra da população 1 = 52;  
população 2 = 50.
- Seja um nível de significância  $\alpha = 0,05$ ;
- $1 - \alpha = 0,95$ .
- O número de graus de liberdade é  $52 + 50 - 2 = 100$ .

# Exemplo 1



# Na tabela...

g!	Probabilidade p							
	0,60	0,70	0,75	0,80	0,85	0,90	0,925	0,95
1	1,376	1,963	2,414	3,078	4,165	6,314	8,449	12,706
2	1,061	1,386	1,604	1,886	2,282	2,920	3,443	4,303
3	0,978	1,250	1,423	1,638	1,924	2,353	2,681	3,182
4	0,941	1,190	1,344	1,533	1,778	2,132	2,392	2,776
5	0,920	1,156	1,301	1,476	1,699	2,015	2,242	2,571
6	0,906	1,134	1,273	1,440	1,650	1,943	2,151	2,447
7	0,896	1,119	1,254	1,415	1,617	1,895	2,090	2,365
8	0,889	1,108	1,240	1,397	1,592	1,860	2,046	2,306
9	0,883	1,100	1,230	1,383	1,574	1,833	2,013	2,262
10	0,879	1,093	1,221	1,372	1,559	1,812	1,987	2,228
11	0,876	1,088	1,214	1,363	1,548	1,796	1,966	2,201
12	0,873	1,083	1,209	1,356	1,538	1,782	1,949	2,179
13	0,870	1,079	1,204	1,350	1,530	1,771	1,935	2,160
14	0,868	1,076	1,200	1,345	1,523	1,761	1,923	2,145
15	0,866	1,074	1,197	1,341	1,517	1,753	1,913	2,131
16	0,865	1,071	1,194	1,337	1,512	1,746	1,904	2,120
17	0,863	1,069	1,191	1,333	1,508	1,740	1,897	2,110
18	0,862	1,067	1,189	1,330	1,504	1,734	1,890	2,101
19	0,861	1,066	1,187	1,328	1,500	1,729	1,884	2,093
20	0,860	1,064	1,185	1,325	1,497	1,725	1,878	2,086
21	0,859	1,063	1,183	1,323	1,494	1,721	1,873	2,080
...	...	...	...	...	...	...	...	...
99	0,845	1,042	1,157	1,290	1,451	1,660	1,799	1,984
100	0,845	1,042	1,157	1,290	1,451	1,660	1,799	1,984
110	0,845	1,041	1,156	1,289	1,450	1,659	1,797	1,982

# Exemplo 1



- Não se rejeita  $H_0$  se encontrar um valor de  $t_0 > -1,984$

$$t_0 < 1,984$$

$$n_1 = 52$$

$$n_2 = 50$$

$$\bar{x}_1 = 142,1 \text{ mmHg}$$

$$\bar{x}_2 = 121,6 \text{ mmHg}$$

$$S_1 = 23,0 \text{ mmHg}$$

$$S_2 = 21,3 \text{ mmHg}$$

$$S_p^2 = \frac{(52-1)23,0^2 + (50-1)21,3^2}{52+50-2} = 492,0981$$

$$t_0 = \frac{142,1 - 121,6}{\sqrt{492,0981 \left( \frac{1}{52} + \frac{1}{50} \right)}} \cong 4,67$$

# Exemplo 1



- Como  $t_0 > t^*$ , rejeita-se  $H_0$  para um nível de significância de 0,05.
- Assim temos evidências de que na população de indivíduos de origem japonesa, com mais de 30 anos e residentes no município do interior de São Paulo, com SM e sem SM possuem médias diferentes de PAS.

# Exemplo 2



- Num estudo comparativo do tempo médio de adaptação, uma amostra aleatória, de 50 homens e 50 mulheres de um grande complexo industrial, produziu os seguintes resultados:

	Homens	Mulheres
Média	3,2 anos	3,7 anos
Desvio padrão	0,8 anos	0,9 anos

- Que conclusões você poderia tirar para a população dessa indústria? Quais suposições você deve fazer?

# Exemplo 2



- Queremos determinar se há diferença entre o tempo de adaptação de homens e mulheres. Devemos supor que o tempo de adaptação tem distribuição Normal. A amostra foi colhida de maneira independente. As variâncias populacionais, ainda que desconhecidas, sejam as mesmas.
- Queremos testar a hipótese que as médias são iguais, isto é,  $H_0: \mu_H = \mu_M$ , ou equivalentemente,  $H_0: \mu_H - \mu_M = 0$ . Note que as suposições acima podem (ou melhor, devem) todas ser verificadas através de testes de hipótese específicos.

# Exemplo 2



- O teste T com variâncias iguais mas desconhecidas é baseado na seguinte estatística:

$$T = \frac{\bar{X}_H - \bar{X}_M}{S_p \sqrt{\frac{1}{n_H} + \frac{1}{n_M}}} \sim t_{(n_H + n_M - 2)}$$

- Onde  $S_p$ , o desvio padrão comum é dado por:

$$S_p^2 = \frac{(n_H - 1)S_H^2 + (n_M - 1)S_M^2}{n_H + n_M - 2}$$

- No problema apresentado,  $sp = 0.8514$

# Exemplo 2



- A estatística observada foi:

$$t_0 = \frac{3.2 - 3.7}{0.8514 \sqrt{\frac{1}{50} + \frac{1}{50}}} = -2.9363$$

- Note que a região crítica agora é dada por:

$$RC(0.05) = \{[T < -1.984] \cup [T > 1.984]\}$$

onde  $q = -1.984$  é o ponto tal que  $P([T < q]) = 0.025$ , etc.

- E como  $-2.9363 \in RC$ , rejeitamos a hipótese nula. Ou seja, há evidência em favor da diferença entre o tempo médio de adaptação dos homens e das mulheres.

# Referências



- Comparação entre duas amostras independentes. Disponível em PDF:  
[http://docente.lages.ifsc.edu.br/liciana.garcia/MaterialDidatico/Aulas%20pdf%200Biotec/Aula%206\\_Comparacao%20entre%20duas%20amostras%20indep.pdf](http://docente.lages.ifsc.edu.br/liciana.garcia/MaterialDidatico/Aulas%20pdf%200Biotec/Aula%206_Comparacao%20entre%20duas%20amostras%20indep.pdf)
- Exemplo de exercício: níveis séricos de frutossamina. Disponível em:  
<http://leg.ufpr.br/~shimakur/CE055/node87.html>
- Teste de hipóteses - Comparação entre duas médias populacionais - Amostras independentes. Disponível em:  
[https://www.youtube.com/watch?v=ZsAd2QND\\_wc](https://www.youtube.com/watch?v=ZsAd2QND_wc)
- Teste  $t$  de Student. Disponível em PDF:  
<http://sistemas.eferp.usp.br/myron/arquivos/2540410/e8fc3b72347400901a2750cb214bf4e0.pdf>
- Estatística Não-Paramétrica. Disponível em PDF:  
[http://people.ufpr.br/~prbg/public\\_html/ce050/apostcap4a.PDF](http://people.ufpr.br/~prbg/public_html/ce050/apostcap4a.PDF)

# Referências



- Teste de Hipótese. Disponível em:  
[http://stat2.med.up.pt/cursop/print\\_script.php3?capitulo=testet&numero=3&titulo=Testes%20de%20Hip%F3tese%20Teste%20T](http://stat2.med.up.pt/cursop/print_script.php3?capitulo=testet&numero=3&titulo=Testes%20de%20Hip%F3tese%20Teste%20T)
- Bioestatística. Disponível em:  
[http://www.pucrs.br/famat/helio/Caderno\\_Bio\\_puc](http://www.pucrs.br/famat/helio/Caderno_Bio_puc)
- Instituto de Matemática, Estatística e Computação Científica – Aula p12.  
Disponível em PDF: [http://www.ime.unicamp.br/~hildete/Aula\\_p12.pdf](http://www.ime.unicamp.br/~hildete/Aula_p12.pdf)



obrigada !!

