

Teste t para dados emparelhados

(Leitura complementar ao [capítulo 5](#))

Sumário:

[Dados emparelhados](#)

[Duas amostras pertencem à mesma população?](#)

[Estimação do tamanho amostral](#)

[Menor diferença detectável](#)

[Sensibilidade do teste](#)

[Teste bicaudal](#)

[Teste monocaudal](#)

[Teste t](#)

Dados emparelhados

É muito frequente necessitarmos *comparar médias de amostras diferentes*.

É importante notar que se forem utilizados *dados emparelhados* a comparação será mais perfeita, já que outros fatores poderão influenciar o resultado em quaisquer duas amostras, tais como:

- variabilidade genética,
- extração social diferente,
- higidez de cada amostra.

Todas as amostras deveriam ser planejadas para serem coletadas aos pares, ou seja cada amostra deveria ter sempre o sua amostra controle.

Para evitar críticas, toma-se uma amostra composta por pares de medidas tiradas de indivíduos que, em cada par, mostrem a maior semelhança possível, ou seja:

- no mesmo indivíduo
- em gêmeos monozigóticos
- em gêmeos dizigóticos
- em irmãos do mesmo sexo
- em irmãos de sexo diferente

Exemplos

- Tempo de coagulação do sangue de mulheres homozigotas para um tipo de hemofilia e em heterozigotas normais.

- O número de cristas dermopapilares na falange distal dos dedos das mãos de mongolóides, com o de indivíduos normais

- Pressão arterial de pacientes que apresentam uma determinada doença, em repouso e após um certo exercício físico.

- Peso de pacientes no pré e no pós-operatório.

Teste t

O parâmetro usado para descrever a distribuição t é

$$\text{número de graus de liberdade} = \text{tamanho da amostra (n)} - 1$$

Quando o tamanho da amostra cresce, a distribuição t se aproxima da normal.

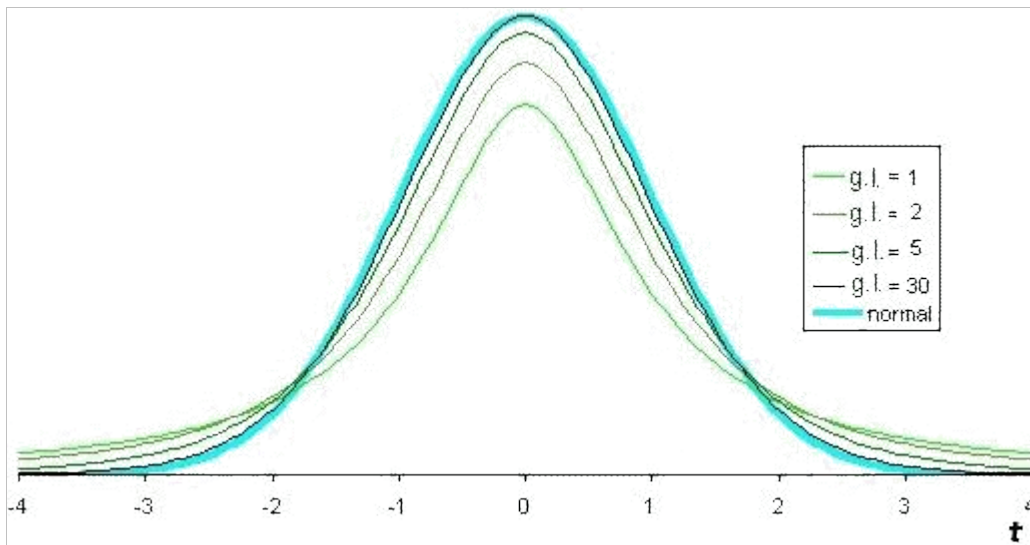


Figura: Tamanho amostral, distribuições t e normal

As amostras podem ser chamadas de 1 e 2 e as medidas dos indivíduos de x_1 e x_2 . Ao se aplicar o teste, quer-se saber se os valores de x_1 são, em média, iguais aos valores de x_2 , ou seja, o objetivo do teste é verificar se a média das diferenças intrapar difere significativamente ou não de zero.

Há *dois* tipos de teste:

- *bicaudal* (ou bilateral): utilizado quando interessam os resultados de ambos os lados da curva e
- *monocaudal* (ou unilateral): usado quando são importantes os resultados de apenas um lado da curva.

Ao realizar o teste, chega-se a 2 *valores* de t:

1. O t crítico (t_c) encontrado na [tabela de t](#) e lido

- na linha correta de GL (em que GL = número de pares -1)
- na coluna de 5%, se o teste for bicaudal (alfa do teste = 5%) e
- na coluna de 10%, se o teste for monocaudal, para manter o alfa do teste em 5%, já que a tabela é bicaudal.

Essa é a diferença entre os dois tipos de teste:

Teste	na tabela de t ler o valor de t_c na linha	na coluna
Bicaudal	correspondente a n-1 graus de liberdade	5%
Monocaudal	idem	10%

2. E calcula-se o t, usando as seguintes fórmulas :

$$\text{desvio} = d = x_1 - x_2$$

$$d \text{ médio} = \underline{d} = \Sigma d / n$$

$$\text{erro da média} = s_d = \sqrt{(\Sigma d^2 - [(\Sigma d)^2 / n] / n \cdot (n-1))}$$

$$t = d \text{ médio} / \text{erro da média} = \underline{d} / S_d$$

Estabelece-se as hipóteses:

H. Nula: A média dos desvios é igual a zero. $H_0: \mu_d = 0$

H. Alternativa: A média dos desvios é diferente de zero. (Teste bicaudal) $H_a: \mu_d \neq 0$ **ou**

H. Alternativa: A média dos desvios é maior que zero. (Teste monocaudal 1) $H_a: \mu_d > 0$ **ou**

H. Alternativa: A média dos desvios é menor que zero. (Teste monocaudal 2) $H_a: \mu_d < 0$

A análise é feita com base nesses *critérios*:

- Teste bicaudal:

$t < -t_c$	$-t_c$	$-t_c < t < +t_c$	$+t_c$	$t > +t_c$
rejeita-se a hipótese nula		aceita-se a hipótese nula		rejeita-se a hipótese nula

- Teste monocaudal :

Se $t < t_c$ aceita-se a hipótese nula.	t_c	Se $t > t_c$ rejeita-se a hipótese nula
---	-------	---

Teste bicaudal

Exemplo: Foi verificado o diâmetro da reação de Mitsuda, em milímetros, provocada por duas partidas de lepromina (A e B) em uma amostra de 10 indivíduos. (O braço direito será sempre inoculado com 0,1 ml de lepromina da partida A. A leitura será sempre feita pelo mesmo operador). Foram obtidos os seguintes resultados. As médias das duas partidas são diferentes?

Indivíduo	Lepromina A (x_1)	Lepromina B (x_2)	$d = x_1 - x_2$	d^2
1	4	5	-1	1
2	8	7	1	1
3	4	4	0	0
4	6	7	-1	1
5	5	4	1	1
6	6	8	-2	4
7	7	7	0	0
8	7	6	1	1
9	3	5	-2	4
10	5	4	1	1
Total	55	57	-2	14

Quais são as hipóteses? (ou seja, porque é um teste bicaudal?)

H. Nula: as leprominas têm o mesmo efeito, ou seja, a média dos desvios é igual a zero.

$$H_0: \mu_d = 0$$

H. Alternativa: as leprominas não tem o mesmo efeito, isto é, a média dos desvios difere de zero.

$$H_a: \mu_d \neq 0$$

Portanto,

$$d \text{ médio} = \underline{d} = \Sigma d / n = -2/10 = -0,20 \text{ mm}$$

$$\text{erro da média} = S_d = \sqrt{(\Sigma d^2 - [(\Sigma d)^2 / n]) / n(n-1)} =$$

$$S_d = \sqrt{(14 - 2^2 / 10) / 10 \cdot 9} = 0,39$$

O t calculado (t) e o t crítico (t_c)

O valor calculado para t deve ser comparado com o valor tabelado de t crítico

$$t = d \text{ médio} / S_d = -0,2 / 0,39 = -0,513 \text{ com } 0,60 < P < 0,70.$$

É importante perceber que o valor de t crítico (t_c) depende da hipótese alternativa:

Como H_a = μ_d ≠ 0 t é bicaudal Interessam valores de t > 0 e < que 0

Como a hipótese alternativa é bicaudal, procura-se o valor de t_c na fileira correspondente a n- 1 graus de liberdade.

Encontra-se t_c = 2,262, com uma probabilidade de 0,60 < P < 0,70.

Como o valor obtido de t (-0,513) encontra-se entre - 2,262 e + 2,262 aceita-se a hipótese nula, ou seja, admite-se que as reações das pessoas às diferentes partidas de lepromina não diferem entre si.

Para facilitar os cálculos utilize uma planilha especial:

Teste t entre médias - exemplos.

Copie a planilha comprimida em formato xls ou em ods

<http://www.cultura.ufpa.br/dicas/biome/biozip/testet.zip>

Outra resolução

Pode-se usar a fórmula do item Curva Normal para o *Intervalo de confiança* e calcular o intervalo de confiança de 95% no qual se situa a médias das diferenças intrapar:

$$d \text{ médio} - (t_c \times S_d) \leq \mu_d \leq d \text{ médio} + (t_c \times S_d)$$

$$-0,2 - (2,262 \times 0,39) \leq \mu_d \leq -0,2 + (2,262 \times 0,39)$$

$$-1,082 \leq \mu_d \leq 0,682$$

Constata-se que a média dos desvios (μ_d) pode assumir o valor zero (já que zero está entre

esses limites). Assim, pode-se aceitar a hipótese nula.

Teste monocaudal

Exemplo: Abaixo encontram-se as médias obtidas pelos gêmeos de uma amostra de 20 pares de gêmeos monozigóticos submetidos a duas técnicas de ensino (N = nova e A = antiga).

- sorteou-se em cada par de MZ qual será submetido à técnica A e qual será submetido à B.
- ao final do experimento fez-se uma série de provas, cujas médias serão comparadas.

Pergunta-se: técnica nova é melhor que a antiga?

Gêmeos	Técnica N (x_1)	Técnica A (x_2)	$d = x_1 - x_2$	d^2
1	54	55	-1	1
2	57	54	3	9
3	60	57	3	9
4	70	60	10	100
5	55	50	5	25
6	62	62	0	0
7	53	48	5	25
8	70	61	9	81
9	62	55	7	49
10	56	55	1	1
11	61	52	9	81
12	65	70	-5	25
13	54	57	-3	9
14	59	51	8	64
15	60	58	2	4
16	68	61	7	49
17	65	60	5	25
18	70	66	4	16
19	63	54	9	81
20	58	56	2	4
Total	1222	1142	80	658

É importante notar que a pergunta refere-se a *média melhor*. Obviamente refere-se a maior. Portanto,

H. Nula: as duas técnicas tem o mesmo efeito, ou seja, a média dos desvios não difere de zero

$$H_0 = \mu_d = 0$$

H. Alternativa: a técnica nova N é melhor que a antiga A, ou seja, a média dos desvios é maior que zero

$$H_a = \mu_d > 0$$

Portanto:

$$d \text{ médio} = \bar{d} = \Sigma d / n = 4$$

$$\text{erro da média} = S_d = \sqrt{(\Sigma d^2 - [(\Sigma d)^2 / n]) / n(n-1)}$$

$$S_d = \sqrt{[658 - (80^2 / 20)] / 20 \times 19} = 0,94$$

$$t_{19} = 4 / 0,94 = 4,257 \text{ com } P < 0,001$$

Como a hipótese alternativa é monocaudal e a tabela de t é bicaudal, procura-se o valor do t

crítico na fileira correspondente a n-1 graus de liberdade, e na coluna com o *dobro do grau de significância* desejado (10%). Nesse caso, $t_{19} = 1,729$.

Como o valor de t_{19} (4,257) é maior que t_c (1,729) rejeita-se a hipótese nula, e os resultados obtidos após a técnica nova diferem dos antigos. A técnica nova é melhor que a antiga já que ocasiona médias maiores.

Outra resolução

Pode-se usar a fórmula do item Curva Normal (*intervalo de confiança*) e calcular o intervalo de confiança de 95% no qual se situa a médias das diferenças intrapar:

$$\underline{d} - (t_c \times s_d) \leq \mu_d \leq + (t_c \times s_d)$$

$t_c = 2,093$ (t lido na coluna de 5% e na linha de 19 graus de liberdade, pois, nesse caso interessam as 2 caudas da distribuição).

Substituindo os valores:

$$\begin{aligned} &= 4 - (2,093 \times 0,94) \leq \mu_d \leq 4 + (2,093 \times 0,94) \\ &= 2,033 \leq \mu_d \leq 5,967 \end{aligned}$$

Deve-se notar que os dois limites fiduciais são positivos, não incluindo o valor zero. Portanto, constata-se que a média μ_d não pode assumir o valor zero (já que zero está fora desses limites). Pode-se rejeitar a hipótese nula.

Menor diferença detectável

Quando se tem dados emparelhados, o produto do erro da média das diferenças intrapar (s_d) pelo valor de t_c , ao nível de 5%, indicará a menor diferença detectável = δ . Portanto:

$$\text{Menor diferença detectável} = \delta = s_d \cdot t_c$$

No caso do exemplo anterior, dos pares de gêmeos submetidos a técnicas de avaliação diferentes, teríamos:

$$\delta = s_d \cdot t_c = 0,94 \times 2,093 = 1,97$$

Ou seja, o número de pares usado é suficiente para detectar uma diferença de no mínimo 1,97 pontos nas médias das provas realizadas por esses gêmeos com o nível de significância de 5%.

Sensibilidade do teste

Se tomarmos a soma dos valores dos 20 pares de gêmeos e tirarmos a média e fizermos a razão da *menor diferença detectável* (δ) por essa média, teremos:

$$x \text{ médio} = \bar{x} = (1222 + 1142) / 40 = 59,10$$

$$\text{sensibilidade do teste} = 1,97 / 59,10 = 0,0333 = 3,33\%$$

Como 3,33% é menor que o nível de significância (5%) pode-se dizer que o teste tem boa sensibilidade

Já, no caso do primeiro exemplo, o exame das 2 partidas de lepromina, teríamos:

$$\delta = s_d \cdot t_c = 0,39 \times 2,262 = 0,88 \text{ mm}$$

$$x \text{ médio} = (55 + 57) / 20 = 5,60,$$

portanto sensibilidade do teste = $0,88 / 5,60 = 0,1571 = 15,71\%$

O valor obtido é muito maior que 5%. Assim a sensibilidade do teste não é tão boa.

Mas um fator deve ser considerado: o valor 0,88 é menor que 1 mm, que é a precisão do teste de Mitsuda, portanto o experimento continua válido.

Estimação do tamanho amostral

Se não estivermos satisfeitos com a [sensibilidade](#) de um experimento com dados emparelhados é possível estimar o número de pares necessário para se obter a precisão desejada.

Por exemplo, com relação aos dados sobre leprominas obtivemos $d = 0,88$. E se quisermos que esse valor passe a ser de 0,80?

$$t_c = d / s_d$$

$$t_c = 0,80 / 0,39 = 2,051$$

Na [tabela de t](#) procura-se o valor mais próximo. Acha-se 2,052, que corresponde a 27 *gl*, o que define que o *tamanho amostral* deve ser de 28 pares de indivíduos.

Como só 10 pares foram coletados deve-se, portanto, investigar mais 18 pares para conseguir detectar uma diferença mínima de 0,80 entre eles.

Duas amostras pertencem à mesma população?

Ou seja, elas são homogêneas? Seus elementos podem ser adicionados? Uma análise conjunta pode ser feita?

Como já foi dito antes, ao estudar uma variável com [distribuição normal](#) em duas ou mais amostras, em geral é necessário saber se uma amostra difere significativamente das outras, ou seja, se elas podem ser consideradas como extraídas da mesma população.

Como a distribuição normal é determinada pela [média](#) e [desvio padrão](#) (ou variância) é óbvio que se as médias e variâncias de 2 ou + amostras não diferirem significativamente pode-se aceitar que elas foram extraídas da mesma população.

Para responder essas perguntas, sabe-se que devem ser comparadas as variâncias e as médias das amostras.

Assim, duas amostras pertencem à mesma população se:

1. as suas médias não diferem significativamente ($\bar{x}_1 = \bar{x}_2$) e
2. as suas variâncias não diferem significativamente ($s^2_1 = s^2_2$)

Portanto, para se verificar se esses valores obtidos na amostra são estatisticamente iguais deve-se estabelecer se há desvio significativo entre suas variâncias e, também, entre suas médias. Evidentemente, em ambas as comparações, não pode haver desvios significativos.

Comparação entre as 2 variâncias

Um teste que pode ser feito para comparar variâncias é o Teste F, proposto por Snedecor, em 1934, em homenagem a Fisher, em que:

F = razão entre a maior variância e a menor variância, ou seja:

$$F = s^2_{\text{maior}} / s^2_{\text{menor}}$$

Consulta-se a [tabela de F](#), ao nível de 2,5%, para encontrar F_c .

Usa-se o seguinte critério:

$F < F_c$	F_c	$F > F_c$
As variâncias <i>não diferem</i> significativamente		As variâncias <i>diferem</i> significativamente

Se as variâncias forem diferentes, com certeza as amostras são heterogêneas e não é necessário continuar o teste.

Mas, se as variâncias não forem estatisticamente diferentes, continua-se:

Comparação entre as 2 médias

Pode ser utilizado o seguinte teste t:

$$t = (\bar{x}_1 - \bar{x}_2) / \text{raiz} ((s^2_1 / n_1) + (s^2_2 / n_2))$$

em que \bar{x}_1 = média maior e \bar{x}_2 = média menor

Exemplo 1

Foram medidas as distâncias interpupilares em 2 amostras, uma masculina (1) outra feminina (2).

Que se pode dizer sobre essas distâncias nos dois sexos?

<i>Amostra 1</i>	<i>Amostra 2</i>
$n_1 = 100$	$n_2 = 131$
$\bar{x}_1 = 62,02$	$\bar{x}_2 = 59,20$
$s^2_1 = 8,00$	$s^2_2 = 7,61$

1. Comparação entre as 2 variâncias

$$F_{(99,130)} = 8,00 / 7,61 = 1,05$$

$$F_c(99,130) = 1,27$$

Como $F < F_c$, as variâncias não são diferentes. Portanto, passa-se à:

2. Comparação entre as 2 médias

Estabelece-se as hipóteses:

Hipótese Nula: $\mu_a - \mu_b = 0$ ou seja $H_0: \mu_d = 0$

H. Alternativa: $\mu_a - \mu_b \neq 0$ ou seja $H_a: \mu_d \neq 0$ (teste bicaudal)

Encontrando t_c

$$n_1 + n_2 = 100 + 131 = 231. \text{ Portanto, } gl = 231 - 1 - 1 = 229$$

Procurando t_c na coluna de 5% e na fileira de infinito, encontra-se: $t_c = 1,96$

Calculando t

$$\text{Calcula-se } t_{(229)} = (\bar{x}_1 - \bar{x}_2) / \text{raiz} ((s^2_1 / n_1) + (s^2_2 / n_2))$$

$$t = (62,02 - 59,20) / \text{raiz} (8,00 / 100 + 7,61 / 131) = 7,589, \text{ com } P < 0,001$$

Como $t > t_c$ rejeita-se H_0 e aceita-se H_a .

Assim, conclui-se que as médias das 2 amostras mostram diferenças significativas. Ou seja, a distância interpupilar dos homens (amostra 1) é maior que a das mulheres (amostra 2), já que sua média é maior.

Entretanto, nota-se que se outras hipóteses poderiam ter sido estabelecidas:

Hipótese Nula: $\mu_1 - \mu_2 = 0$ ou seja $H_0: \mu_d = 0$

H. Alternativa: $\mu_1 - \mu_2 > 0$ ou seja $H_a: \mu_d > 0$ (teste monocaudal)

Nesse caso, o t_c seria lido na coluna de 10% e na fileira de infinito, $t_c = 1,645$

E o t calculado, evidentemente, continua a ser 7,589.

Como $t > t_c$ rejeita-se H_0 e aceita-se H_a . Ou seja, também se concluiria que as médias amostrais são estatisticamente diferentes.

Portanto, retornando à questão original, ou seja, o que se pode dizer sobre as distâncias interpupilares, quando se compara uma amostra masculina e outra feminina, constata-se que sua variância não diferem mas que suas médias dão diferentes estatisticamente.

Assim, pode-se afirmar que as amostras *não são homogêneas* e que a distância interpupilar mostra uma distribuição significativamente diferente nos dois sexos. (Portanto, é uma característica sexual secundária?)

Para facilitar os cálculos utilize uma planilha especial:

Duas amostras pertencem à mesma população?

Copie a planilha comprimida em formato xls ou em ods

<http://www.cultura.ufpa.br/dicas/biome/biozip/amo-pop.zip>

Exemplo 2

Em 2 amostras foi analisada a mesma variável, com distribuição normal e obteve-se:

Amostra 1	Amostra 2
$n_1 = 10$	$n_2 = 20$
$s^2_1 = 8,10$	$s^2_2 = 2,02$
$gl_1 = 9$	$gl_2 = 19$

$$F_{c(9,19)} = 2,88$$

$$F_{(9,19)} = 8,10 / 2,02 = 4,01$$

Como F obtido ($F_{(9,19)} = 4,01$) é maior que F_c ($F_{c(9,19)} = 2,88$) conclui-se que as variâncias são diferentes.

Nesse caso, como as variâncias já são diferentes, não é preciso continuar a análise e admite-se que as amostras não pertencem à mesma população, ou seja, que não são homogêneas.

Casos especiais:

Em casos raros é necessário fazer a comparação entre 2 médias de amostras com variâncias diferentes. O método é o mesmo, mas deve ser usada uma fórmula especial para o cálculo de t (Cochrane Cox, 1950).

Para amostras que apresentam distribuição de Poisson também há um modo especial de calcular.

Este "site", destinado prioritariamente aos alunos de Fátima Conti, pretende auxiliar quem esteja começando a se interessar por Bioestatística, computadores e programas, estando em permanente construção. Sugestões e comentários são bem vindos. Agradeço antecipadamente.

Endereço dessa página:

HTML: <http://www.cultura.ufpa.br/dicas/biome/biomed.htm>

PDF: <http://www.cultura.ufpa.br/dicas/pdf/biomed.pdf>

Última alteração: 13 out 2009 (Solicito conferir datas. Pode haver atualização só em HTML)