Teste de Student "t"

O *teste de Student*, ou simplesmente *teste t* é o método mais utilizado para se avaliarem as diferenças entre as médias de dois grupos.

Por exemplo, o teste t pode ser usado para 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”

Podemos montar uma experiência aos pares e efetuar o teste para o mesmo grupo de pessoas em duas situações diferentes.

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.

Teste “t” para observações independentes

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

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

Calculam-se:

a) A média de cada grupo; indica-se:

$\bar{x}_1$: **Média do grupo 1**

$\bar{x}_2$: **Média do grupo 2**

Teste “t” para observações independentes

b) A variância de cada grupo; indica-se:

s_1^2 : **variância do grupo 1**

s_2^2 : **variância do grupo 2**

c) A variância ponderada, dada pela fórmula:

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

Teste “t” para observações independentes

d) O valor de t, definido por

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

e) Graus de liberdade

$$n_1 + n_2 - 2$$

Teste “t” para observações independentes

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.

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

Exemplo: Para verificar se duas dietas para emagrecer são igualmente eficientes, um médico separou, ao acaso, um conjunto de pacientes em dois grupos. Cada um seguiu a dieta designada para o seu grupo.

Perda de peso, em quilogramas, segundo a dieta.

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

Teste "t" para observações independentes

1º) α=5%

a) A média de cada grupo:

$$\bar{x}_1 = \frac{12+8+15+13+10+12+14+11+12+13}{10} = \frac{120}{10} = 12$$

$$\bar{x}_2 = \frac{15+19+15+12+13+16+15}{7} = \frac{105}{7} = 15$$

Teste "t" para observações independentes

b) A variância de cada grupo:

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

$$s_1^2 = \frac{1476 - \frac{120^2}{10}}{9} = \frac{36}{9} = 4$$

$$s_2^2 = \frac{1605 - \frac{105^2}{7}}{6} = \frac{30}{6} = 5$$

c) A variância ponderada, dada pela fórmula:

$$s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} = \frac{9 \times 4 + 6 \times 5}{10 + 7 - 2} = 4{,}4$$

Teste "t" para observações independentes

d) O valor de t, definido por

$$t = \frac{\bar{x}_2 - \bar{x}_1}{\sqrt{s^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{15-12}{\sqrt{4{,}4 \times \left(\frac{1}{10} + \frac{1}{7}\right)}} = 2{,}902$$

e) Graus de liberdade

$$n_1 + n_2 - 2 = 10 + 7 - 2 = 15$$

Teste "t" para observações independentes

Tabela A.6

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
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 independentes

t_c =2,13 valor tabelado.

t = 2,902 valor calculado.

$t > t_c$ conclui-se que, em média, 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;**
- **Efeito de um tratamento: o dentista aplica o tratamento de um lado da arcada e deixa o outro sem tratamento;**
- **Tratamentos onde se observam os mesmos indivíduos duas vezes, antes e depois do tratamento: pressão arterial.**

Teste “t” para observações pareadas

Para testar o efeito de um tratamento, quando as observações são pareadas, aplica-se o teste t.

Temos que calcular:

a) A diferença entre as unidades de cada um dos *n* pares

$$d = x_2 - x_1$$

b) A média das diferenças

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

Teste "t" para observações pareadas

c) A variância das diferenças

$$s^2 = \frac{\sum d^2 - \frac{\left(\sum d\right)^2}{n}}{n-1}$$

d) O valor de t

$$t = \frac{\bar{d}}{\sqrt{\frac{s^2}{n}}}$$

e) 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

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

Dieta Antes	Dieta Depois	Diferenças
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

Teste “t” para observações pareadas

b) A média das diferenças: $\bar{d} = \frac{-30}{9} = -3{,}333$

c) A variância das diferenças:

$$s^2 = \frac{\sum d^2 - \frac{\left(\sum d\right)^2}{n}}{n-1} = \frac{300 - \frac{(-30)^2}{9}}{8} = \frac{200}{8} = 25$$

d) O valor de t

$$t = \frac{\bar{d}}{\sqrt{\frac{s^2}{n}}} = \frac{-3{,}333}{\sqrt{\frac{25}{9}}} = \frac{-3{,}333}{1{,}666} = -2{,}0$$

e) Grau de liberdade: n-1=9-1=8

Teste “t” para observações independentes

Tabela A.6

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
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

t_c =3,36 valor tabelado.

t = 2,00 valor absoluto calculado.

$t < t_c$ conclui-se que, o tratamento não tem efeito significante, ao nível de 1%. Em termos práticos, o experimento não provou que dieta emagrece.

Teste “t” para observações independentes quando as variâncias não são iguais

Como se estabelece que as variâncias das populações são iguais?

Para testar a hipótese de que as variâncias das duas populações são iguais, aplica-se o teste F.

1º estabelecer o nível de significância, em seguida calcula-se:

Teste “t” para observações independentes quando as variâncias não são iguais

a) A variância de cada grupo; indica-se:

s_1^2 : **variância do grupo 1**

s_2^2 : **variância do grupo 2**

b) O valor de F, dado pela razão entre a maior e a menor variância.

$$Se\ s_1^2 > s_2^2\ o\ valor\ de\ F = \frac{s_1^2}{s_2^2}$$

Está associado a n_1-1(numerador) e n_2-1 (denominador) graus de liberdade.

Teste “t” para observações independentes quando as variâncias não são iguais

Está associado a n_1-1(numerador) e n_2-1 (denominador) graus de liberdade.

Precisamos procurar o valor de F na tabela, com nível de significância igual a metade do nível de significância estabelecido, suponha 5%.

Valores de F para α = 2,5%, segundo o número de graus de liberdade do numerador e do denominador

Nº de g. l. do denominador	Número de graus de liberdade do numerador								
	1	2	3	4	5	6	7	8	9
1	648,0	800,0	864,0	900,0	922,0	937,0	948,0	957,0	963,0
2	38,5	39,0	39,2	39,2	39,3	39,3	39,4	39,4	39,4
3	17,4	16,0	15,4	15,1	14,9	14,7	14,6	14,5	14,5
4	12,2	10,6	9,98	9,60	9,36	9,20	9,07	8,98	8,90
5	10,0	8,43	7,76	7,39	7,15	6,98	6,85	6,76	6,68
6	8,81	7,26	6,60	6,23	5,99	5,82	5,70	5,60	5,52
7	8,07	6,54	5,89	5,52	5,29	5,12	4,99	4,90	4,82
8	7,57	6,06	5,42	5,05	4,82	4,65	4,53	4,43	4,36
9	7,21	5,71	5,08	4,72	4,48	4,32	4,20	4,10	4,03

Teste “t” para observações independentes quando as variâncias não são iguais

Calculam-se:

a) A média de cada grupo; indica-se:

$\bar{x}_1$: **Média do grupo 1**

$\bar{x}_2$: **Média do grupo 2**

Teste "t" para observações independentes quando as variâncias não são iguais

b) A variância de cada grupo; indica-se:

s_1^2 : **variância do grupo 1**

s_2^2 : **variância do grupo 2**

c) O valor de t, definido por

$$t = \frac{\bar{x}_2 - \bar{x}_1}{\sqrt{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)}}$$

Teste "t" para observações independentes quando as variâncias não são iguais

d) Graus de liberdade

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

Teste “t” para observações independentes quando as variâncias não são iguais

Feitos os cálculos, é preciso procurar o valor de t na tabela, ao nível de significância estabelecido.

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

Teste "t" para observações independentes quando as variâncias não são iguais

Teste a hipótese de que recém nascidos de ambos os sexos têm, em média, a mesma estatura. Teste essa hipótese, ao nível de significância de 5%.

Sexo	n	$\bar{x}$	s^2
Masculino	1442	49,29	5,76
Feminino	1361	48,54	6,30

Teste "t" para observações independentes quando as variâncias não são iguais

Testamos se as variâncias são iguais

$$F = \frac{6,30}{5,76} = 1,09$$

1360(numerador) e 1441(denominador)

Nº de g. l. do denominador	Número de graus de liberdade do numerador									
	10	12	15	20	24	30	40	60	120	∞
1	969	977	985	993	997	1000	1010	1010	1010	1020
2	39,4	39,4	39,4	39,4	39,5	39,5	39,5	39,5	39,5	39,5
3	14,4	14,3	14,3	14,2	14,1	14,1	14,0	14,0	13,9	13,9
4	8,84	8,75	8,66	8,56	8,51	8,46	8,41	8,36	8,31	8,26
5	6,62	6,52	6,43	6,33	6,28	6,23	6,18	6,12	6,07	6,02
6	5,46	5,37	5,27	5,17	5,12	5,07	5,01	4,96	4,90	4,85
7	4,76	4,67	4,57	4,47	4,42	4,36	4,31	4,25	4,20	4,14
8	4,30	4,20	4,10	4,00	3,95	3,89	3,84	3,78	3,73	3,67
9	3,96	3,87	3,77	3,67	3,61	3,56	3,51	3,45	3,39	3,33
10	3,72	3,62	3,52	3,42	3,37	3,31	3,26	3,20	3,14	3,08
11	3,53	3,43	3,33	3,23	3,17	3,12	3,06	3,00	2,94	2,88
12	3,37	3,28	3,18	3,07	3,02	2,96	2,91	2,85	2,79	2,72
13	3,25	3,15	3,05	2,95	2,89	2,84	2,78	2,72	2,66	2,60
14	3,15	3,05	2,95	2,84	2,79	2,73	2,67	2,61	2,55	2,49
15	3,06	2,96	2,86	2,76	2,70	2,64	2,59	2,52	2,46	2,40
16	2,99	2,89	2,79	2,68	2,63	2,57	2,51	2,45	2,38	2,32
17	2,92	2,82	2,72	2,62	2,56	2,50	2,44	2,38	2,32	2,25
18	2,87	2,77	2,67	2,56	2,50	2,44	2,38	2,32	2,26	2,19
19	2,82	2,72	2,62	2,51	2,45	2,39	2,33	2,27	2,20	2,13
20	2,77	2,68	2,57	2,46	2,41	2,35	2,29	2,22	2,16	2,09
21	2,73	2,64	2,53	2,42	2,37	2,31	2,25	2,18	2,11	2,04
22	2,70	2,60	2,50	2,39	2,33	2,27	2,21	2,14	2,08	2,00
23	2,67	2,57	2,47	2,36	2,30	2,24	2,18	2,11	2,04	1,97
24	2,64	2,54	2,44	2,33	2,27	2,21	2,15	2,08	2,01	1,94
25	2,61	2,51	2,41	2,30	2,24	2,18	2,12	2,05	1,98	1,91
26	2,59	2,49	2,39	2,28	2,22	2,16	2,09	2,03	1,95	1,88
27	2,57	2,47	2,36	2,25	2,19	2,13	2,07	2,00	1,93	1,85
28	2,55	2,45	2,34	2,23	2,17	2,11	2,05	1,98	1,91	1,83
29	2,53	2,43	2,32	2,21	2,15	2,09	2,03	1,96	1,89	1,81
30	2,51	2,41	2,31	2,20	2,14	2,07	2,01	1,94	1,87	1,79
40	2,39	2,29	2,18	2,07	2,01	1,94	1,88	1,80	1,72	1,64
60	2,27	2,17	2,06	1,94	1,88	1,82	1,74	1,67	1,58	1,48
120	2,16	2,05	1,94	1,82	1,76	1,69	1,61	1,53	1,43	1,31
∞	2,05	1,94	1,83	1,71	1,64	1,57	1,48	1,39	1,27	1,00

Teste “t” para observações independentes quando as variâncias não são iguais

O valor tabelado é igual a 1 menor que o valor calculado, logo rejeita-se a hipótese de que as variâncias são iguais, ao nível de 2,5%.

$$t = \frac{\bar{x}_2 - \bar{x}_1}{\sqrt{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)}}$$

$$t = \frac{49{,}29 - 48{,}54}{\sqrt{\frac{5{,}76}{1442} + \frac{6{,}30}{1361}}} = 8{,}076$$

Teste "t" para observações independentes quando as variâncias não são iguais

Que está associado aos graus de liberdade

$$g = \frac{\left(\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}\right)^2}{\dfrac{\left(\dfrac{s_1^2}{n_1}\right)^2}{n_1 - 1} + \dfrac{\left(\dfrac{s_2^2}{n_2}\right)^2}{n_2 - 1}} = \frac{\left(\dfrac{5,76}{1442} + \dfrac{6,30}{1361}\right)^2}{\dfrac{\left(\dfrac{5,76}{1442}\right)^2}{1441} + \dfrac{\left(\dfrac{6,30}{1361}\right)^2}{1360}} = 2772$$

Tabela A.6

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
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
21	1,72	2,08	2,83
22	1,72	2,07	2,82
23	1,71	2,07	2,81
24	1,71	2,06	2,80
25	1,71	2,06	2,79
26	1,71	2,06	2,78
27	1,70	2,05	2,77
28	1,70	2,05	2,76
29	1,70	2,04	2,76
30	1,70	2,04	2,75
40	1,68	2,02	2,70
60	1,67	2,00	2,66
120	1,66	1,98	2,62
∞	1,64	1,96	2,58

Teste “t” para observações independentes quando as variâncias não são iguais

Como o valor calculado de t é maior que o valor tabelado, logo rejeita-se a hipótese de que recém-nascidos de ambos os sexos têm, em média, a mesma estatura, ao nível de 5%.

Em termos práticos, os meninos nascem com estatura maior do que as meninas.

Teste “t” para o coeficiente de correlação

O teste t pode ser usado para testar a hipótese de que o coeficiente de correlação entre duas variáveis é igual a zero, contra a hipótese de que é diferente de zero.

Para aplicar o teste t, usa-se a fórmula

$$t = \frac{r}{\sqrt{1-r^2}}\sqrt{n-2}$$

Graus de liberdade n-2

Teste "t" para o coeficiente de correlação

Exemplo: Suponha que o coeficiente de correlação entre duas variáveis é -0,775 e n=14

Para aplicar o teste t, usa-se a fórmula

$$t = \frac{r}{\sqrt{1-r^2}}\sqrt{n-2} = \frac{-0,775}{\sqrt{1-0,601}}\sqrt{14-2} = -4,25$$

Graus de liberdade n-2=14-2=12 α=5%

Ao nível de significância de 5% o valor tabelado é 2,18, logo a correlação entre as variáveis é significante ao nível de 5%.

Intervalos de Confiança

Intervalos de Confiança

Imagine uma amostra casual simples de n elementos. A média dos dados dessa amostra constitui uma *estimativa* da média da população, de onde essa amostra proveio.

Para indicar a precisão dessa estimativa, calcula-se o intervalo de confiança para a média.

Intervalos de Confiança

Um população é constituída pelos valores 14, 20, 26.

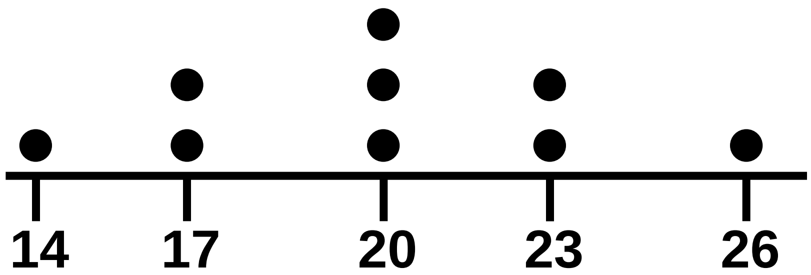
$$\mu = \frac{14 + 20 + 26}{3} = 20$$

Vamos considerar todas as amostra de dois elementos que podem ser retirados dessa população, com reposição.

Intervalos de Confiança

Médias das amostras de dois elementos obtidos da população constituída pelos números 14, 20 e 26.

Amostra	Média
14 e14	14
14 e 20	17
14 e 26	20
20 e 14	17
20 e 20	20
20 e 26	23
26 e 14	20
26 e 20	23
26 e 26	26



Intervalos de Confiança

Vamos medir a dispersão das médias das amostras em torno da média da população.

$$\sigma_{\bar{x}}^2 = \frac{\sum_{i=1}^{r}\left(\overline{x_i} - \mu\right)^2}{r}$$

logo para o exemplo temos.

r número de amostras que podem ser obtidas da população

$$\sigma_{\bar{x}}^2 = \frac{(14-20)^2 + (17-20)^2 + \cdots + (26-20)^2}{9} = \frac{108}{9} = 12$$

Intervalos de Confiança

É impossível calcular a variância da média pela fórmula dada, pois, o pesquisador dispõe de uma única amostra, para estimar a média da população.

Foi demonstrado que a estimativa da variância da média é dada pela fórmula:

$$s_{\bar{x}}^2 = \frac{s^2}{n}$$

Intervalos de Confiança

Médias das amostras de dois elementos obtidos da população constituída pelos números 14, 20 e 26.

Amostra	Média	Variância	Variância da Média
14 e14	14	0	0
14 e 20	17	18	9
14 e 26	20	72	36
20 e 14	17	18	9
20 e 20	20	0	0
20 e 26	23	18	9
26 e 14	20	72	36
26 e 20	23	18	9
26 e 26	26	0	0
Média	20	24	12

Intervalos de Confiança

Por definição, erro padrão da média é a raiz quadrada com sinal positivo da variância da média

$$s_{\bar{x}} = \frac{s}{\sqrt{n}}$$

Seja X uma variável aleatória com distribuição normal de média μ e variância σ². Com base numa amostra aleatória de n elemento podemos obter as estimativas $\bar{x}$ e s^2 **de μ e variância σ² respectivamente.**

$$\bar{x} - t\frac{s}{\sqrt{n}} < \mu < \bar{x} + t\frac{s}{\sqrt{n}}$$

O valor de t é tabelado com n-1 graus de liberdade e α.

Intervalos de Confiança

Exemplo:X é v. a. que representa a taxa de colesterol no plasma sangüíneo. Uma amostra casual simples de n=25 indivíduos, foram obtidos a média 198mg/100ml e desvio padrão de 30 mg/100ml. Seja α=10% 24 graus de liberdade t=1,71.

$$\bar{x} - t\frac{s}{\sqrt{n}} < \mu < \bar{x} + t\frac{s}{\sqrt{n}} \Rightarrow 198 - 1{,}71\frac{30}{\sqrt{25}} < \mu < 198 + 1{,}71\frac{30}{\sqrt{25}}$$

$$187{,}74 < \mu < 208{,}26$$

Interpretação do intervalo de confiança: quando são obtidas muitas amostras de n elementos de uma mesma população e se determina, para cada amostra, um intervalo de confiança, (100-α)% desses intervalos contém a média da população.

Intervalos de Confiança

(100-α)% é denominado nível de confiança, ou seja os intervalos são de confiança (100-α)%

Na área biológica é comum apresentar os valores de

$\bar{x}$ e $s_{\bar{x}}$ **escritos na forma** $\bar{x} \pm s_{\bar{x}}$

Este intervalo pode ser visto como um intervalo de confiança, mas com nível de confiança indeterminado. Isto por que neste caso t=1, e o valor de t depende dos graus de liberdade.