

TESTE T PARA POPULAÇÕES INDEPENDENTES

ISABELA GOMES DA SILVA
ISABELLA NAOMI FURUIE
MARIA JÚLIA JORGE MAURO

DISTRIBUIÇÃO T DE STUDENT

Distribuição para uma amostra n retirada de uma população com distribuição normal padrão.

Na teoria utilizamos esta fórmula para, por exemplo, calcular o IC, porém na prática não sabemos o desvio padrão da população (σ).

σ - é um parâmetro, o desvio padrão da população total e tipicamente não temos esse valor.

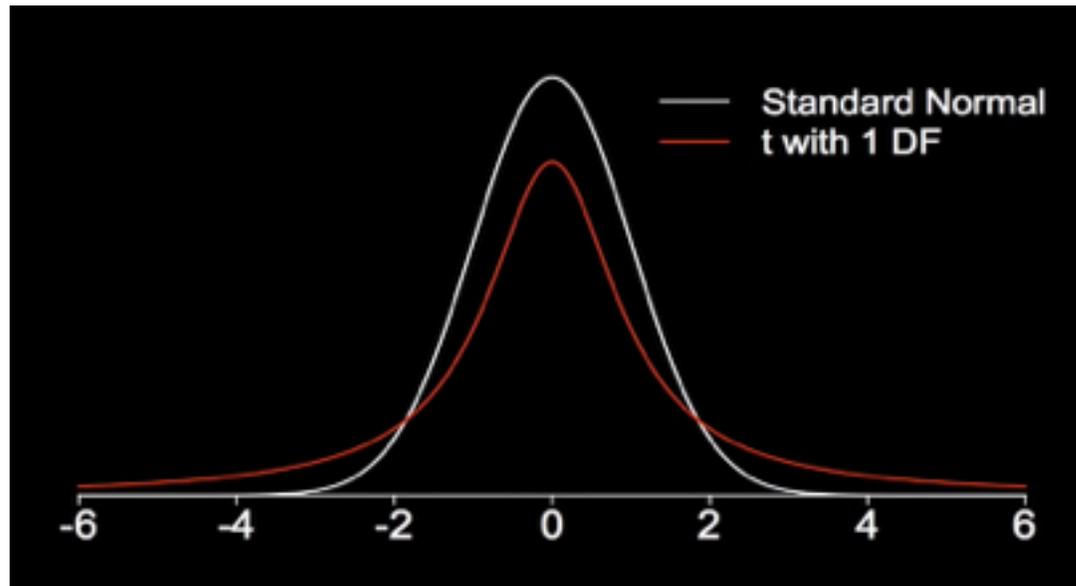
s - desvio padrão da amostra

Quando a amostra é proveniente de uma população de distribuição normal padrão, ela terá distribuição t com $n-1$ graus de liberdade

$$Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}}$$

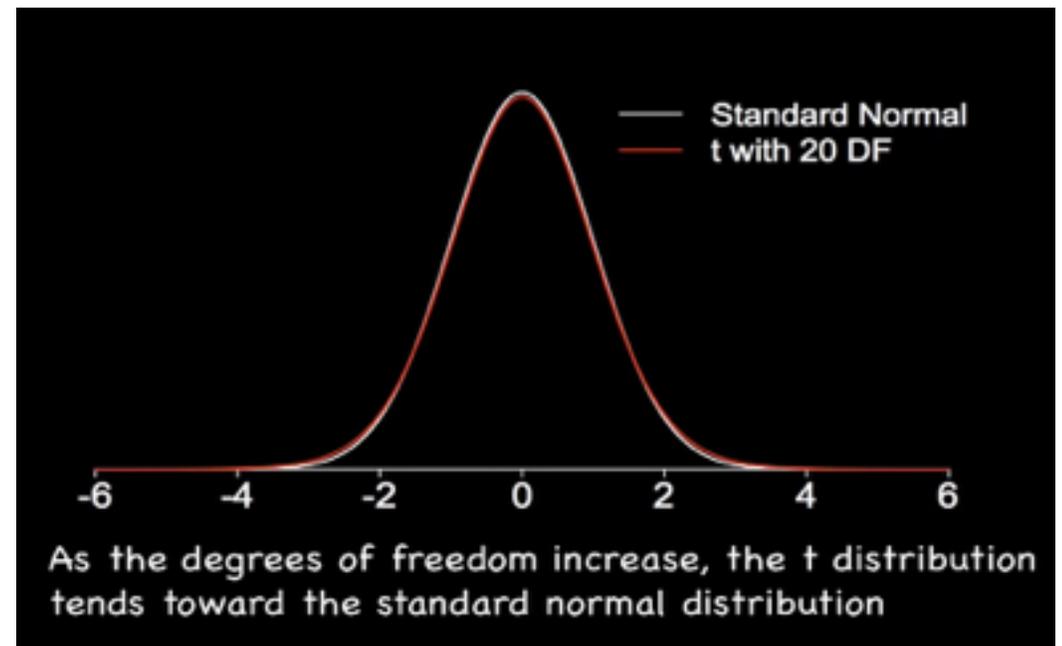
$$t = \frac{\bar{X} - \mu}{s / \sqrt{n}}$$

DISTRIBUIÇÃO T DE STUDENT



A fórmula é parecida com a Normal Z, porém estamos usando uma estatística (s) para avaliar um parâmetro. A variância nesse caso é maior.

Quanto maior o grau de liberdade, mais semelhante a curva será à distribuição normal padrão



DISTRIBUIÇÃO T DE STUDENT

Constructing a 95% confidence interval
(for the population mean)

If σ is known:

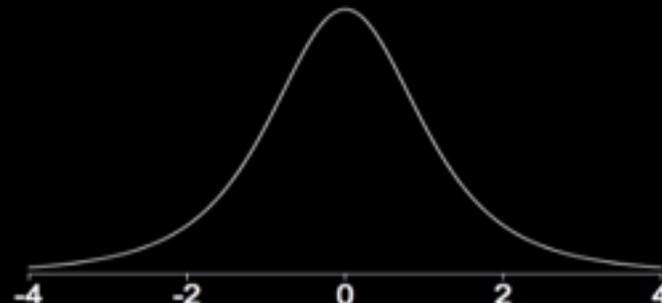
$$\bar{X} \pm 1.96 \times \frac{\sigma}{\sqrt{n}}$$



$$Z_{0,025} = 1,96$$

If σ is not known:

$$\bar{X} \pm ? \times \frac{s}{\sqrt{n}}$$



$$T_{0,025} > Z_{0,025}$$

DISTRIBUIÇÃO T DE STUDENT

$t_{.025}$ values (for 95% confidence intervals)

n	df	$t_{.025}$
6	5	2.571
11	10	2.228
31	30	2.042
51	50	2.009
101	100	1.984
∞	∞	1.960

COMPARANDO A MÉDIA DE DOIS GRUPOS INDEPENDENTES

1. Definir a hipótese
2. Definir alfa
3. Calcular grau de liberdade
4. Definir a regra da decisão
5. Calcular o teste estatístico
6. Definir os resultados
7. Conclusão

EXEMPLO 1

Um professor de estatística quer comparar os resultados da prova de duas de suas turmas para ver se o desempenho das duas foi diferente. Turma A tem 25 alunos com uma nota média de 70 e desvio padrão de 15. Turma B tem 20 alunos com uma nota média de 74 e desvio padrão de 25. Utilizando $\alpha = 0,05$, essas duas turmas tiveram desempenhos diferentes nessa prova?

1 - Definir a Hipótese

$$H_0; \mu_{classA} = \mu_{classB}$$

$$H_1; \mu_{classA} \neq \mu_{classB}$$

2- Definir o alfa

$$\alpha = 0.05$$

EXEMPLO 1

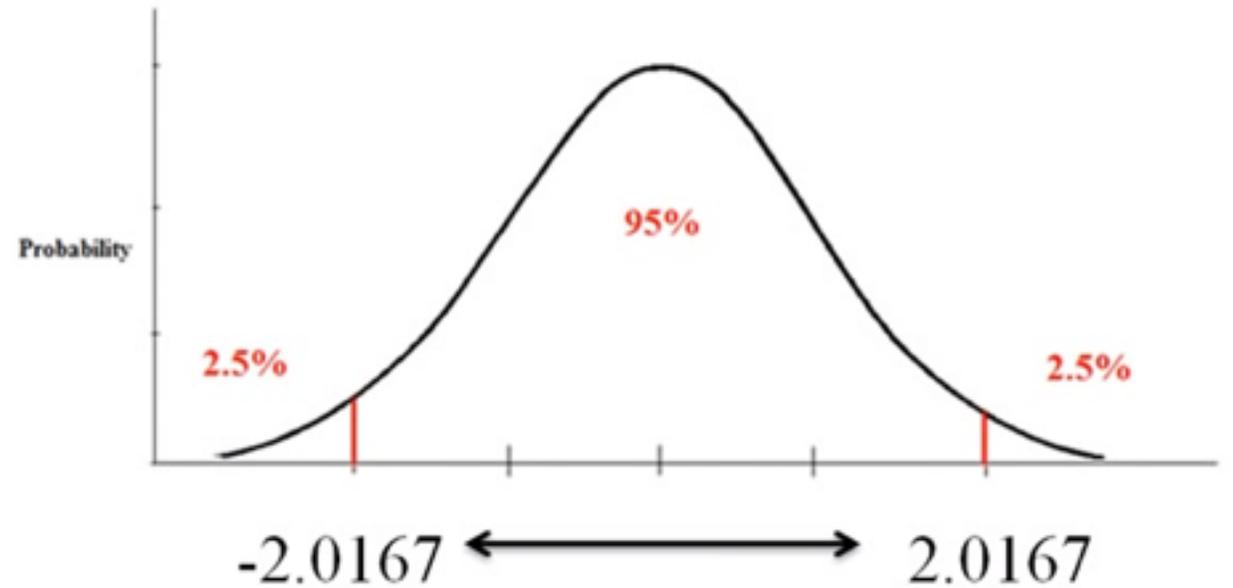
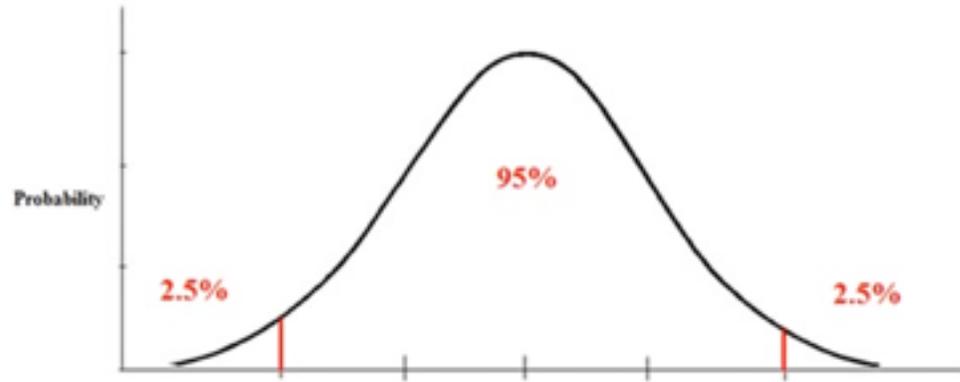
3- Calcular Grau de Liberdade

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

$$df = (25 - 1) + (20 - 1) = 43$$

EXEMPLO 1

4- Definir a regra de decisão



t Table

Critical Values						Critical Values					
df	0.1	0.05	0.025	0.01	0.005	df	0.1	0.05	0.025	0.01	0.005
1-tailed						1-tailed					
2-tailed						2-tailed					
1	3.0777	6.3138	12.7062	31.8205	63.6567	32	1.3086	1.6939	2.0369	2.4487	2.7385
2	1.8856	2.9200	4.3027	6.9646	9.9248	33	1.3077	1.6924	2.0345	2.4448	2.7333
3	1.6377	2.3534	3.1824	4.5407	5.8409	34	1.3070	1.6909	2.0322	2.4411	2.7284
4	1.5332	2.1318	2.7764	3.7469	4.6041	35	1.3062	1.6896	2.0301	2.4377	2.7238
5	1.4759	2.0150	2.5706	3.3649	4.0321	36	1.3055	1.6883	2.0281	2.4345	2.7195
6	1.4398	1.9432	2.4469	3.1427	3.7074	37	1.3049	1.6871	2.0262	2.4314	2.7154
7	1.4149	1.8946	2.3646	2.9980	3.4995	38	1.3042	1.6860	2.0244	2.4286	2.7116
8	1.3968	1.8595	2.3060	2.8965	3.3554	39	1.3036	1.6849	2.0227	2.4258	2.7079
9	1.3830	1.8331	2.2622	2.8214	3.2498	40	1.3031	1.6839	2.0211	2.4233	2.7045
10	1.3722	1.8125	2.2281	2.7638	3.1693	41	1.3025	1.6829	2.0195	2.4208	2.7012
11	1.3634	1.7959	2.2010	2.7181	3.1058	42	1.3020	1.6820	2.0181	2.4185	2.6981
12	1.3562	1.7823	2.1788	2.6810	3.0545	43	1.3016	1.6811	2.0167	2.4163	2.6951
13	1.3502	1.7709	2.1604	2.6503	3.0123	44	1.3011	1.6802	2.0154	2.4141	2.6923
14	1.3450	1.7613	2.1448	2.6245	2.9768	45	1.3006	1.6794	2.0141	2.4121	2.6896
15	1.3406	1.7531	2.1314	2.6025	2.9467	46	1.3002	1.6787	2.0129	2.4102	2.6870
16	1.3368	1.7459	2.1199	2.5835	2.9208	47	1.2998	1.6779	2.0117	2.4083	2.6846
17	1.3334	1.7396	2.1098	2.5669	2.8982	48	1.2994	1.6772	2.0106	2.4066	2.6822
18	1.3304	1.7341	2.1009	2.5524	2.8784	49	1.2991	1.6766	2.0096	2.4049	2.6800
19	1.3277	1.7291	2.0930	2.5395	2.8609	50	1.2987	1.6759	2.0086	2.4033	2.6778
20	1.3253	1.7247	2.0860	2.5280	2.8453	60	1.2958	1.6706	2.0003	2.3901	2.6603
21	1.3232	1.7207	2.0796	2.5176	2.8314	70	1.2938	1.6669	1.9944	2.3808	2.6479
22	1.3212	1.7171	2.0739	2.5083	2.8188	80	1.2922	1.6641	1.9901	2.3739	2.6387
23	1.3195	1.7139	2.0687	2.4999	2.8073	90	1.2910	1.6620	1.9867	2.3685	2.6316
24	1.3178	1.7109	2.0639	2.4922	2.7969	100	1.2901	1.6602	1.9840	2.3642	2.6259
25	1.3163	1.7081	2.0595	2.4851	2.7874	110	1.2893	1.6588	1.9818	2.3607	2.6213
26	1.3150	1.7056	2.0555	2.4786	2.7787	120	1.2886	1.6577	1.9799	2.3578	2.6174
27	1.3137	1.7033	2.0518	2.4727	2.7707	130	1.2881	1.6567	1.9784	2.3554	2.6142
28	1.3125	1.7011	2.0484	2.4671	2.7633	140	1.2876	1.6558	1.9771	2.3533	2.6114
29	1.3114	1.6991	2.0452	2.4620	2.7564	150	1.2872	1.6551	1.9759	2.3515	2.6090
30	1.3104	1.6973	2.0423	2.4573	2.7500	∞	1.2816	1.6449	1.9600	2.3263	2.5758
31	1.3095	1.6955	2.0395	2.4528	2.7440						

EXEMPLO 1

5- Teste Estatístico

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

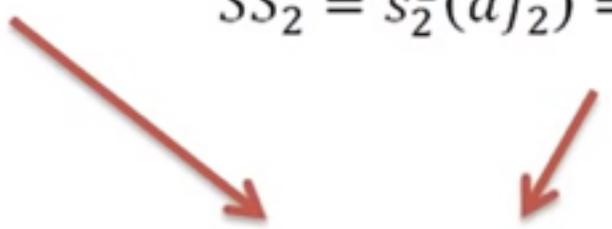
$$s_p^2 = \frac{SS_1 + SS_2}{df_1 + df_2}$$

$$df_1 = n_1 - 1 = 25 - 1 = 24$$

$$df_2 = n_2 - 1 = 20 - 1 = 19$$

$$SS_1 = s_1^2(df_1) = (15^2)(24) = 5400$$

$$SS_2 = s_2^2(df_2) = (25^2)(19) = 11875$$


$$s_p^2 = \frac{5400 + 11875}{24 + 19} = 401.74$$

EXEMPLO 1

5- Teste Estatístico

$$t = \frac{(70 - 74)}{\sqrt{\frac{401.74}{25} + \frac{401.74}{20}}} = \frac{-4}{\sqrt{36.16}} = -0.67$$

EXEMPLO 1

6- Resultados

$$t = -0.67$$

7- Conclusão

Como definimos a regra de decisão como: rejeitar a hipótese nula se t for maior que 2.0167 ou menor que -2.0167, e -0.67 não preenche esses critérios, não podemos rejeitar a hipótese nula de que a média da turma A é igual à da turma B.

Não há diferença estatisticamente significativa entre as médias da turma A e turma B, $t = -0.67, p > 0,05$

EXEMPLO 2

Dois grupos de ratos foram alimentados com dietas com alto e baixo conteúdo de proteínas. O quadro abaixo fornece o ganho de peso destes ratos entre os dias 28 e 84 de vida:

Conteúdo de proteína	Ganho de peso											
Alto	123	134	146	104	119	124	161	107	83	113	129	97
Baixo				70	118	101	85	107	132	94		

EXEMPLO 2

Um grupo de pesquisadores deseja testar a hipótese de que o alto conteúdo de proteína leva a maior aumento de peso

PERGUNTAS:

A) ao nível de significância de 1%, há evidência estatística de que a dieta com alto conteúdo de proteína aumenta o ganho de peso?

Em outras palavras...

Podemos rejeitar a hipótese H_0 : média do peso 1 = média do peso 2?

EXEMPLO 2

PASSO 1 – calcular o valor de T

$$H_0: \mu_1 = \mu_2$$

$$H_a: \mu_1 > \mu_2$$

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

$$\frac{\bar{x}_1 - \bar{x}_2}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} > t_{n_1 + n_2 - 2; 1 - \frac{\alpha}{2}}$$

EXEMPLO 2

PASSO 1

$$\bar{x}_1 = 120 \quad s = 21,39 \quad n_1 = 12$$

$$\bar{x}_2 = 101 \quad s = 20,62 \quad n_2 = 7$$

$$T_{obs} = \frac{120 - 101}{21,12 \sqrt{\frac{1}{12} + \frac{1}{7}}} = 1,89$$

$$s_{comb} = \sqrt{\frac{(12 - 1)(21,29)^2 + (7 - 1)(20,62)^2}{12 + 7 - 2}} = 21,12$$

EXEMPLO 2

PASSO 2 – achar o valor de t para n_1+n_2-2 ; $1-\alpha/2$ na tabela de t de student, para o nível de significancia de 1%

$$N_1+n_2-2= 12+7-2= 17$$

$$T= 2,567 \text{ (aproximadamente } 2,57)$$

PASSO 3 – CONCLUSÃO

$$T < 2,57$$

Portanto, não podemos descartar a hipótese de que os pesos médios entre os grupos de ratos se equívalem.

<i>gl</i>	0,600	0,750	0,900	1,050	1,200	1,350	1,500	1,650	1,800
1	0,325	1,000	3,078	6,214	12,700	22,700	31,821	63,657	636,619
2	0,289	0,816	1,886	2,920	4,303	6,965	9,925	31,598	
3	0,277	0,765	1,638	2,353	3,182	4,541	5,841	12,924	
4	0,271	0,741	1,533	2,132	2,776	3,747	4,604	8,610	
5	0,267	0,727	1,476	2,015	2,571	3,365	4,032	6,869	
6	0,265	0,718	1,440	1,943	2,447	3,143	3,707	5,959	
7	0,263	0,711	1,415	1,895	2,365	2,998	3,499	5,408	
8	0,262	0,706	1,397	1,860	2,306	2,896	3,355	5,041	
9	0,261	0,703	1,383	1,833	2,262	2,821	3,250	4,781	
10	0,260	0,700	1,372	1,812	2,228	2,764	3,169	4,587	
11	0,260	0,697	1,363	1,796	2,201	2,718	3,106	4,437	
12	0,259	0,695	1,356	1,782	2,179	2,681	3,055	4,318	
13	0,259	0,694	1,350	1,771	2,160	2,650	3,012	4,221	
14	0,258	0,692	1,345	1,761	2,145	2,624	2,977	4,140	
15	0,258	0,691	1,341	1,753	2,131	2,602	2,947	4,073	
16	0,258	0,690	1,337	1,746	2,120	2,583	2,921	4,015	
17	0,257	0,689	1,333	1,740	2,110	2,567	2,898	3,965	
18	0,257	0,688	1,330	1,734	2,101	2,552	2,878	3,922	
19	0,257	0,688	1,328	1,729	2,093	2,539	2,861	3,883	
20	0,257	0,687	1,325	1,725	2,086	2,528	2,845	3,850	
21	0,257	0,686	1,323	1,721	2,080	2,518	2,831	3,819	
22	0,256	0,686	1,321	1,717	2,074	2,508	2,819	3,792	
23	0,256	0,685	1,319	1,714	2,069	2,500	2,807	3,768	
24	0,256	0,685	1,318	1,711	2,064	2,492	2,797	3,745	
25	0,256	0,684	1,316	1,708	2,060	2,485	2,787	3,725	
26	0,256	0,684	1,315	1,706	2,056	2,479	2,779	3,707	
27	0,256	0,684	1,314	1,703	2,052	2,473	2,771	3,689	
28	0,256	0,683	1,313	1,701	2,048	2,467	2,763	3,674	
29	0,256	0,683	1,311	1,699	2,045	2,462	2,756	3,660	
30	0,256	0,683	1,310	1,697	2,042	2,457	2,750	3,646	
40	0,255	0,681	1,303	1,684	2,021	2,423	2,704	3,551	
60	0,254	0,679	1,296	1,671	2,000	2,390	2,660	3,460	
120	0,254	0,677	1,289	1,658	1,980	2,358	2,617	3,373	
∞	0,253	0,674	1,282	1,645	1,960	2,326	2,576	3,291	

EXEMPLO 2

Calculando o valor de P

Na linha 17 da tabela, 1,89 fica entre 1,74 e 2,11 (P entre 0,05 e 0,025%)

Se $P < \alpha$ = descarta a hipótese H_0

Como o valor de P fica num intervalo maior que 0,01 $>$ não se descarta a hipótese H_0

	$p(t_p \leq T)$							
gl	0,600	0,750	0,900	0,950	0,975	0,990	0,995	0,9995
1	0,325	1,000	3,078	6,314	12,706	31,821	63,657	636,619
2	0,289	0,816	1,886	2,920	4,303	6,965	9,925	31,598
3	0,277	0,765	1,638	2,353	3,182	4,541	5,841	12,924
4	0,271	0,741	1,533	2,132	2,776	3,747	4,604	8,610
5	0,267	0,727	1,476	2,015	2,571	3,365	4,032	6,869
6	0,265	0,718	1,440	1,943	2,447	3,143	3,707	5,959
7	0,263	0,711	1,415	1,895	2,365	2,998	3,499	5,408
8	0,262	0,706	1,397	1,860	2,306	2,896	3,355	5,041
9	0,261	0,703	1,383	1,833	2,262	2,821	3,250	4,781
10	0,260	0,700	1,372	1,812	2,228	2,764	3,169	4,587
11	0,260	0,697	1,363	1,796	2,201	2,718	3,106	4,437
12	0,259	0,695	1,356	1,782	2,179	2,681	3,055	4,318
13	0,259	0,694	1,350	1,771	2,160	2,650	3,012	4,221
14	0,258	0,692	1,345	1,761	2,145	2,624	2,977	4,140
15	0,258	0,691	1,341	1,753	2,131	2,602	2,947	4,073
16	0,258	0,690	1,337	1,746	2,120	2,583	2,921	4,015
17	0,257	0,689	1,333	1,740	2,110	2,567	2,898	3,965
18	0,257	0,688	1,330	1,734	2,101	2,552	2,878	3,922
19	0,257	0,688	1,328	1,729	2,093	2,539	2,861	3,883
20	0,257	0,687	1,325	1,725	2,086	2,528	2,845	3,850
21	0,257	0,686	1,323	1,721	2,080	2,518	2,831	3,819
22	0,256	0,686	1,321	1,717	2,074	2,508	2,819	3,792
23	0,256	0,685	1,319	1,714	2,069	2,500	2,807	3,768
24	0,256	0,685	1,318	1,711	2,064	2,492	2,797	3,745
25	0,256	0,684	1,316	1,708	2,060	2,485	2,787	3,725
26	0,256	0,684	1,315	1,706	2,056	2,479	2,779	3,707
27	0,256	0,684	1,314	1,703	2,052	2,473	2,771	3,689
28	0,256	0,683	1,313	1,701	2,048	2,467	2,763	3,674
29	0,256	0,683	1,311	1,699	2,045	2,462	2,756	3,660
30	0,256	0,683	1,310	1,697	2,042	2,457	2,750	3,646
40	0,255	0,681	1,303	1,684	2,021	2,423	2,704	3,551
60	0,254	0,679	1,296	1,671	2,000	2,390	2,660	3,460
120	0,254	0,677	1,289	1,658	1,980	2,358	2,617	3,373
∞	0,253	0,674	1,282	1,645	1,960	2,326	2,576	3,291

EXEMPLO 3 COMPARAÇÃO DA TIANEPTINA COM PLÁCEBO

- ensaio clínico aleatorizado, duplo-cego
- Tianeptina: antidepressivo tricíclico
- administrar a droga a dois grupos de pacientes -> quantificar a depressão através da escala de Montgomery-Asberg (MADRS)

Grupo	Escores								
Placebo	6	33	21	26	10	29	33	29	
	37	15	2	21	7	26	13		
Tianeptina	10	8	17	4	17	14	9	4	
	21	3	7	10	29	13	14	2	

$$n_1 = 15 \quad \bar{x}_1 = 20,53 \quad s_1 = 11,09$$

$$n_2 = 16 \quad \bar{x}_2 = 11,37 \quad s_2 = 7,26$$

OBJETIVO: Testar se uma variável difere entre dois grupos independentes de sujeitos

$$H_0 : \mu_1 = \mu_2$$

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

EXEMPLO 3 COMPARAÇÃO DA TIANEPTINA COM PLÁCEBO

$$\begin{array}{lll} n_1 = 15 & \bar{x}_1 = 20,53 & s_1 = 11,09 \\ n_2 = 16 & \bar{x}_2 = 11,37 & s_2 = 7,26 \end{array}$$

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

$$s_p = \sqrt{\frac{14(11,09)^2 + 15(7,26)^2}{15 + 16 - 2}} = 9,31$$

Como a suposição é que as variâncias dos dois grupos são iguais, podemos estimar σ como uma média ponderada, com pesos proporcionais aos tamanhos dos grupos

$$T = \left| \frac{\bar{x}_1 - \bar{x}_2}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \right| = \left| \frac{20,53 - 11,37}{9,31 \sqrt{\frac{1}{15} + \frac{1}{16}}} \right| = 2,74$$

EXEMPLO 3 COMPARAÇÃO DA TIANEPTINA COM PLÁCEBO

t crítico → $t_{n_1+n_2-2; 1-\frac{\alpha}{2}}$

- percentil de ordem $1 - \alpha/2$ da distribuição t com $n_1 + n_2 - 2$ graus de liberdade

$n_1 = 15$	$\bar{x}_1 = 20,53$	$s_1 = 11,09$
$n_2 = 16$	$\bar{x}_2 = 11,37$	$s_2 = 7,26$

- Graus de liberdade:
 $n_1 + n_2 - 2 = 15 + 16 - 2 = 29$
- Percentil:
 $\alpha = 0.05 \rightarrow 1 - \alpha/2 = 1 - 0.05 = 0.975$
- Procurar na tabela o percentil

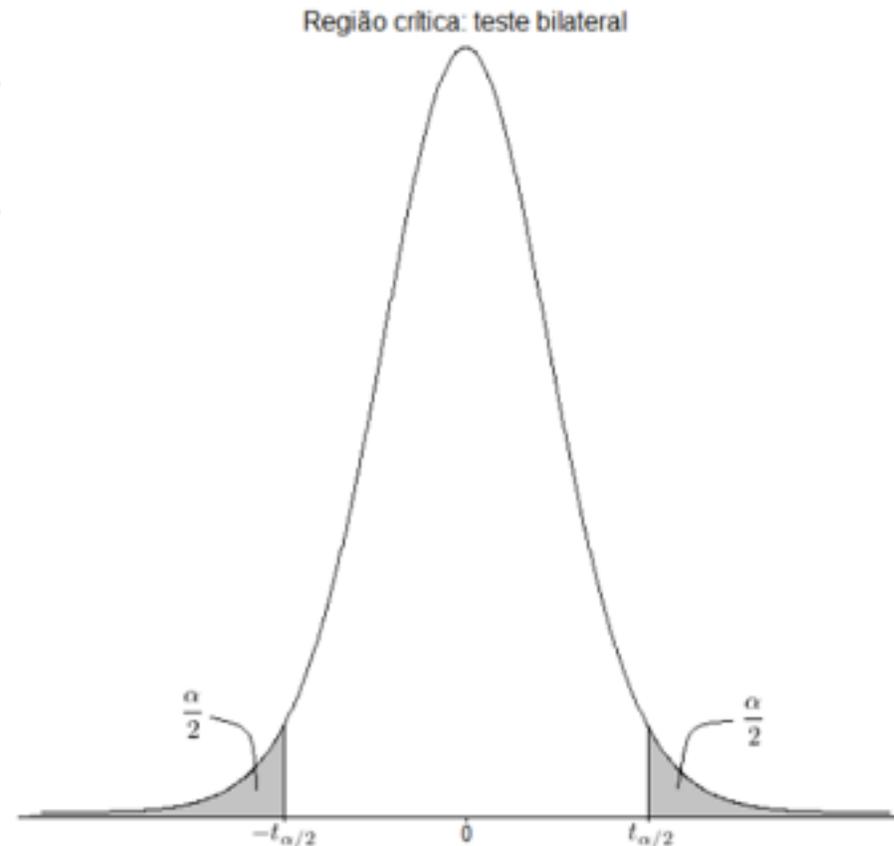


TABELA COM VALORES DE T $P_{97.5} = 2.2381 = 97.5\%$ 2391 T

Graus de liberdade ($n-1$)	P_{90}	P_{95}	$P_{97.5}$	P_{99}	$P_{99.5}$
1	3,0780	6,3138	12,7060	31,8210	63,6570
2	1,8860	2,9200	4,3027	6,9650	9,9248
3	1,6380	2,3534	3,1825	4,5410	5,8409
4	1,5330	2,1318	2,7764	3,7470	4,6041
5	1,4760	2,0150	2,5706	3,3650	4,0321
6	1,4400	1,9432	2,4469	3,1430	3,7074
7	1,4150	1,8946	2,3846	2,9980	3,4995
8	1,3970	1,8596	2,3360	2,8960	3,3554
9	1,3830	1,8331	2,2822	2,8210	3,2498
10	1,3720	1,8125	2,2281	2,7640	3,1693
11	1,3630	1,7959	2,2010	2,7180	3,1058
12	1,3560	1,7823	2,1788	2,6810	3,0545
13	1,3500	1,7709	2,1604	2,6500	3,0123
14	1,3450	1,7613	2,1448	2,6240	2,9768
15	1,3410	1,7530	2,1315	2,6020	2,9467
16	1,3370	1,7459	2,1190	2,5830	2,9208
17	1,3330	1,7398	2,1098	2,5670	2,8982
18	1,3300	1,7341	2,1009	2,5520	2,8784
19	1,3280	1,7291	2,0930	2,5390	2,8609
20	1,3250	1,7247	2,0860	2,5280	2,8453
21	1,3230	1,7207	2,0796	2,5180	2,8314
22	1,3210	1,7171	2,0739	2,5080	2,8188
23	1,3190	1,7139	2,0687	2,5000	2,8073
24	1,3180	1,7109	2,0639	2,4920	2,7969
25	1,3160	1,7081	2,0595	2,4850	2,7874
26	1,3150	1,7056	2,0555	2,4790	2,7787
27	1,3140	1,7033	2,0518	2,4730	2,7707
28	1,3130	1,7011	2,0484	2,4670	2,7633
29	1,3110	1,6991	<u>2,0452</u>	2,4620	2,7564
30	1,3100	1,6973	2,0423	2,4570	2,7500

EXEMPLO 3 COMPARAÇÃO DA TIANEPTINA COM PLÁCEBO

$$\frac{|\bar{x}_1 - \bar{x}_2|}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} > t_{n_1+n_2-2; 1-\frac{\alpha}{2}}$$

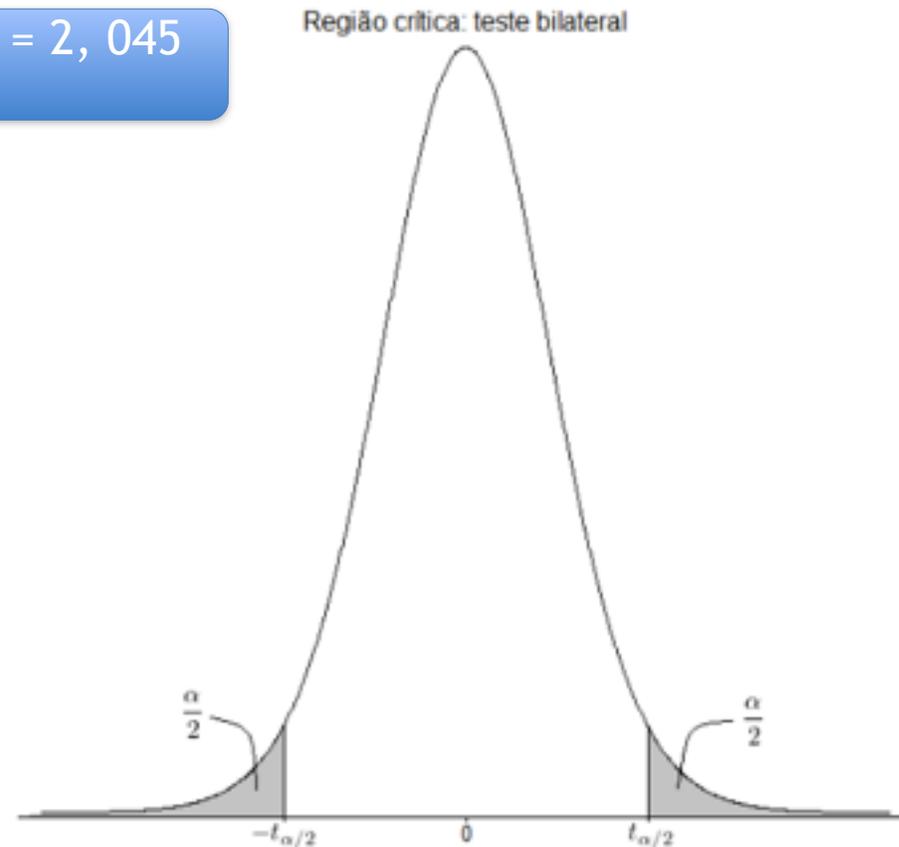
t calculado = 2,74

×

t 29;0,975 = 2,045

2,74 > 2,045

Rejeição da igualdade entre os dois grupos no nível de 5%



OBRIGADA!