

Intervalo de Confiança para a diferença de proporções ($p_1 - p_2$)

$$\hat{p}_1 - \hat{p}_2 \pm z \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

$$P(-z \leq Z \leq z) = \gamma \text{ coeficiente de confiança}$$
$$Z \sim N(0,1)$$

Exemplo

Uma companhia de seguros analisou a frequência com que duas amostras de 1000 homens e 1000 mulheres utilizaram o hospital. Deseja-se testar se a proporção de utilização do hospital é a mesma para homens e mulheres

	Usaram Hospital	Não usaram	
Homens	100	900	1000
Mulheres	150	850	1000

Parâmetros:

p_1 - proporção de mulheres que utilizam o hospital
 p_2 - proporção de homens que utilizam o hospital

$$H_0: p_1 = p_2$$

$$\hat{p}_1 = \frac{150}{1000} = 0,15$$

$$\hat{p}_2 = \frac{100}{1000} = 0,10$$

$$H_a: p_1 \neq p_2$$

$$\hat{V} = \hat{p}(1-\hat{p}) \left(\frac{1}{n_1} + \frac{1}{n_2} \right)$$

$$\hat{p} = \frac{100+150}{2000} = 0,125$$

$$\hat{V} = 0,125(1-0,125) \left(\frac{1}{1000} + \frac{1}{1000} \right) = 0,00022$$

Estatística do Teste:
$$Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{V}}}$$

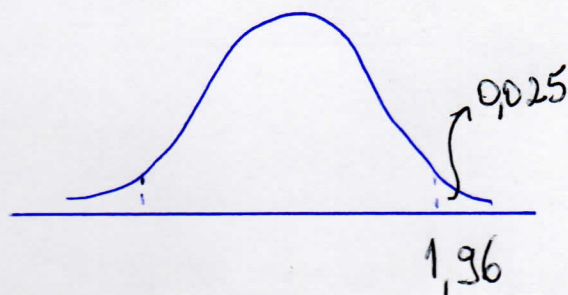
Região Crítica
$$Z \leq -z_c \quad \text{ou} \quad Z \geq z_c$$

$$P(Z \geq z_c) = \frac{\alpha}{2} \quad Z \sim N(0,1)$$

No caso
$$Z = \frac{0,15 - 0,10}{\sqrt{0,00022}} = \frac{0,05}{0,0148} = 3,38$$

$$\alpha = 0,05 \quad P(Z \geq z_c) = 0,025$$

RC:
$$Z \geq 1,96 \quad \text{ou} \quad Z \leq -1,96$$



$$Z_{obs} = 3,38 \in RC \Rightarrow \text{Rejeita-se } H_0.$$

Os dados sugerem desigualdade das proporções de utilizações do hospital para homens e mulheres.

Quando H_0 é rejeitada, é comum a construção³ de um intervalo de confiança para o parâmetro. No caso, o parâmetro é $p_1 - p_2$ e o IC é:

$$\hat{p}_1 - \hat{p}_2 \pm z \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}, \quad P(-z \leq Z \leq z) = \gamma$$

$$\gamma = 0,95 \Rightarrow z = 1,96$$

$$\text{IC: } 0,15 - 0,10 \pm 1,96 \sqrt{\frac{0,15 \cdot 0,85}{1000} + \frac{0,10 \cdot 0,90}{1000}}$$

$$0,05 \pm 1,96 \cdot 0,0147$$

$$[0,0212; 0,0788]$$

$0 \notin \text{IC} \Rightarrow$ Rejeita-se H_0 ao nível $\alpha = 1 - \gamma = 0,05$.

Esta equivalência entre testes de hipóteses e intervalos de confiança funciona apenas para testes bicaudais.

4
Teste de Hipóteses para a média de uma distribuição normal com variância desconhecida

$X \sim N(\mu, \sigma^2)$ μ e σ^2 desconhecidos

$$\hat{\mu} = \bar{X} \quad \hat{\sigma}^2 = S^2$$

$$S^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1} = \frac{\sum_{i=1}^n X_i^2 - n\bar{X}^2}{n-1}$$

Resultado

Se $X \sim N(\mu, \sigma^2)$ e X_1, X_2, \dots, X_n é uma amostra aleatória dessa distribuição, a variável aleatória

$$\frac{\bar{X} - \mu}{S/\sqrt{n}}$$

tem distribuição t de student com $n-1$ graus de liberdade. Tabela Apêndice A, pag. 372, Magalhães e Lima

Notação: $\frac{\bar{X} - \mu}{S/\sqrt{n}} \sim t_{n-1}$

Teste bicaudal para μ

$X \sim N(\mu, \sigma^2)$ μ e σ^2 desconhecidos

$$H_0: \mu = \mu_0$$

$$H_a: \mu \neq \mu_0$$

Estatística de teste:

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

Região Crítica:

Fixado um nível de significância α e tomada uma amostra aleatória de n observações

RC: $T > t_c$ ou $T < -t_c$ onde t_c é tal que

$$P(T > t_c) = \frac{\alpha}{2} \quad T \sim t\text{-student com } n-1 \text{ gl}$$

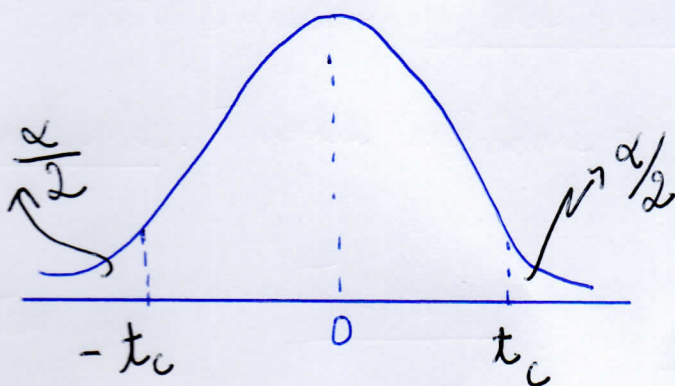


Tabela da distribuição t-student
 Apêndice A, pág 372

Distribuição t-Student : Valores t_c tais que $P(-t_c \leq t \leq t_c) = 1 - p$

	p->90%	80%	70%	60%	50%	40%	30%	20%	10%	5%	4%	2%	1%	0,2%	0,1%
1	0,158	0,325	0,510	0,727	1,000	1,376	1,963	3,078	6,314	12,706	15,894	31,821	63,656	318,289	636,578
2	0,142	0,289	0,445	0,617	0,816	1,061	1,386	1,886	2,920	4,303	4,849	6,965	9,925	22,328	31,600
3	0,137	0,277	0,424	0,584	0,765	0,978	1,250	1,638	2,353	3,182	3,482	4,541	5,841	10,214	12,924
4	0,134	0,271	0,414	0,569	0,741	0,941	1,190	1,533	2,132	2,776	2,999	3,747	4,604	7,173	8,610
5	0,132	0,267	0,408	0,559	0,727	0,920	1,156	1,476	2,015	2,571	2,757	3,365	4,032	5,894	6,869
6	0,131	0,265	0,404	0,553	0,718	0,906	1,134	1,440	1,943	2,447	2,612	3,143	3,707	5,208	5,959
7	0,130	0,263	0,402	0,549	0,711	0,896	1,119	1,415	1,895	2,365	2,517	2,998	3,499	4,785	5,408
8	0,130	0,262	0,399	0,546	0,706	0,889	1,108	1,397	1,860	2,306	2,449	2,896	3,355	4,501	5,041
9	0,129	0,261	0,398	0,543	0,703	0,883	1,100	1,383	1,833	2,262	2,398	2,821	3,250	4,297	4,781
10	0,129	0,260	0,397	0,542	0,700	0,879	1,093	1,372	1,812	2,228	2,359	2,764	3,169	4,144	4,587
11	0,129	0,260	0,396	0,540	0,697	0,876	1,088	1,363	1,796	2,201	2,328	2,718	3,106	4,025	4,437
12	0,128	0,259	0,395	0,539	0,695	0,873	1,083	1,356	1,782	2,179	2,303	2,681	3,055	3,930	4,318
13	0,128	0,259	0,394	0,538	0,694	0,870	1,079	1,350	1,771	2,160	2,282	2,650	3,012	3,852	4,221
14	0,128	0,258	0,393	0,537	0,692	0,868	1,076	1,345	1,761	2,145	2,264	2,624	2,977	3,787	4,140
15	0,128	0,258	0,393	0,536	0,691	0,866	1,074	1,341	1,753	2,131	2,249	2,602	2,947	3,733	4,073
16	0,128	0,258	0,392	0,535	0,690	0,865	1,071	1,337	1,746	2,120	2,235	2,583	2,921	3,686	4,015
17	0,128	0,257	0,392	0,534	0,689	0,863	1,069	1,333	1,740	2,110	2,224	2,567	2,898	3,646	3,965
18	0,127	0,257	0,392	0,534	0,688	0,862	1,067	1,330	1,734	2,101	2,214	2,552	2,878	3,610	3,922
19	0,127	0,257	0,391	0,533	0,688	0,861	1,066	1,328	1,729	2,093	2,205	2,539	2,861	3,579	3,883
20	0,127	0,257	0,391	0,533	0,687	0,860	1,064	1,325	1,725	2,086	2,197	2,528	2,845	3,552	3,850
21	0,127	0,257	0,391	0,532	0,686	0,859	1,063	1,323	1,721	2,080	2,189	2,518	2,831	3,527	3,819
22	0,127	0,256	0,390	0,532	0,686	0,858	1,061	1,321	1,717	2,074	2,183	2,508	2,819	3,505	3,792
23	0,127	0,256	0,390	0,532	0,685	0,858	1,060	1,319	1,714	2,069	2,177	2,500	2,807	3,485	3,768
24	0,127	0,256	0,390	0,531	0,685	0,857	1,059	1,318	1,711	2,064	2,172	2,492	2,797	3,467	3,745
25	0,127	0,256	0,390	0,531	0,684	0,856	1,058	1,316	1,708	2,060	2,167	2,485	2,787	3,450	3,725
26	0,127	0,256	0,390	0,531	0,684	0,856	1,058	1,315	1,706	2,056	2,162	2,479	2,779	3,435	3,707
27	0,127	0,256	0,389	0,531	0,684	0,855	1,057	1,314	1,703	2,052	2,158	2,473	2,771	3,421	3,689
28	0,127	0,256	0,389	0,530	0,683	0,855	1,056	1,313	1,701	2,048	2,154	2,467	2,763	3,408	3,674
29	0,127	0,256	0,389	0,530	0,683	0,854	1,055	1,311	1,699	2,045	2,150	2,462	2,756	3,396	3,660
30	0,127	0,256	0,389	0,530	0,683	0,854	1,055	1,310	1,697	2,042	2,147	2,457	2,750	3,385	3,646
35	0,127	0,255	0,388	0,529	0,682	0,852	1,052	1,306	1,690	2,030	2,133	2,438	2,724	3,340	3,591
40	0,126	0,255	0,388	0,529	0,681	0,851	1,050	1,303	1,684	2,021	2,123	2,423	2,704	3,307	3,551
50	0,126	0,255	0,388	0,528	0,679	0,849	1,047	1,299	1,676	2,009	2,109	2,403	2,678	3,261	3,496
60	0,126	0,254	0,387	0,527	0,679	0,848	1,045	1,296	1,671	2,000	2,099	2,390	2,660	3,232	3,460
120	0,126	0,254	0,386	0,526	0,677	0,845	1,041	1,289	1,658	1,980	2,076	2,358	2,617	3,160	3,373
∞	0,126	0,253	0,385	0,524	0,675	0,842	1,036	1,282	1,645	1,960	2,054	2,327	2,576	3,091	3,291

Ex:

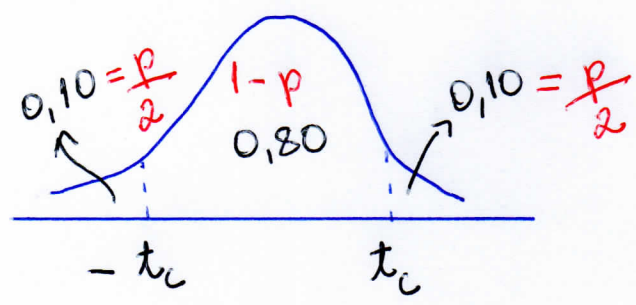
20 gl t_c | $P(-t_c \leq T \leq t_c) = 0,90 = 1 - p \Rightarrow p = 0,10$

p 10%

$t_c = 1,725$

gl 20 --- 1,725

22 gl t_c | $P(T > t_c) = 0,10$



$P(-t_c < T < t_c) = 0,8 \Rightarrow p = 0,2$

p 20%

$t_c = 1,321$

gl . 22 --- 1,321

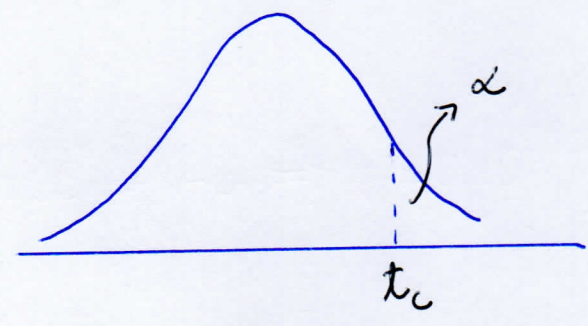
t_c é obtido a partir da tabela para $n-1$ g.l.

Para $gl > 120$ usar a linha ∞ que é equivalente a consultar a tabela da distribuição normal.

Teste unicaudal à direita

$H_0: \mu = \mu_0$

$H_a: \mu > \mu_0$



Estatística de Teste

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

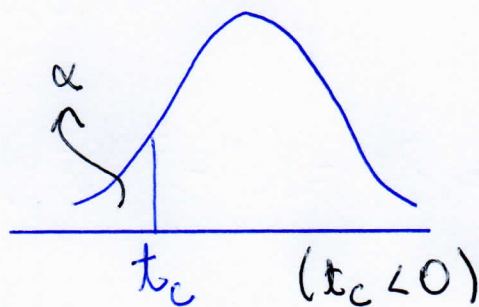
RC: $T > t_c$ t_c | $P(T > t_c) = \alpha$

$T \sim t_{n-1}$

Teste unicaudal à esquerda

$$H_0: \mu = \mu_0$$

$$H_a: \mu < \mu_0$$



Estatística de Teste

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

$$RC: T < t_c \quad t_c \mid P(T < t_c) = \alpha$$

$$T \sim t_{n-1}$$

Exemplo

Deseja-se investigar se uma certa moléstia altera o consumo renal médio de oxigênio. Para indivíduos sãos, esse consumo tem distribuição normal com média $12 \text{ cm}^3/\text{min}$. Os valores medidos para cinco pacientes com a moléstia foram:

14,4 12,9 15,0 13,7 13,5.

Qual seria a conclusão ao nível de significância de 1%?

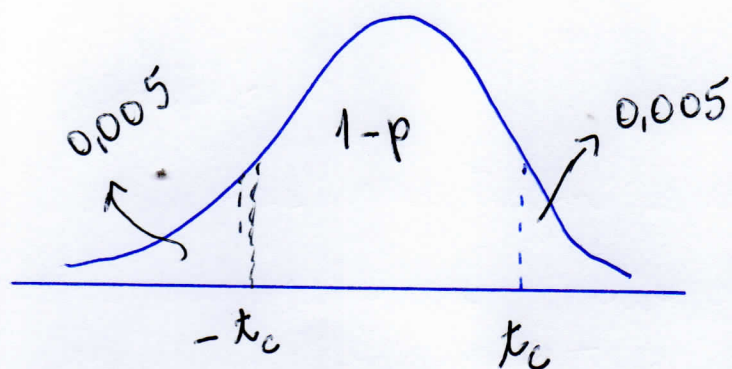
X - consumo renal de oxigênio de portadores da moléstia

$$X \sim N(\mu, \sigma^2) \quad \mu \text{ e } \sigma^2 \text{ desconhecidos}$$

$H_0: \mu = 12$ (A moléstia não altera o consumo renal médio de oxigênio)

$H_a: \mu \neq 12$ (Portadores da moléstia têm a média alterada)

RC: $T > t_c$ ou $T < -t_c$ $\alpha = 0,01$ $n = 5$ 4 gl



Basta procurar na tabela a soma das caudas

gl	p
4	1% 4,604

RC: $-4,604$ $4,604$

O diagrama mostra uma linha horizontal com hachuras nas extremidades à esquerda de $-4,604$ e à direita de $4,604$, indicando a região de rejeição.

Efetuada os cálculos: $\bar{x} = 13,9$ $s^2 = 0,67$

$$T_{obs} = \frac{13,9 - 12}{\sqrt{0,67}/\sqrt{5}} = 5,19 \in RC$$

Rejeita-se H_0 . Ao nível de significância de 0,01, conclui-se que a moléstia altera o consumo renal médio de oxigênio.

Exemplo

Um fabricante afirma que a quantidade média de nicotina em seus cigarros não ultrapassa 30mg. Uma amostra de 25 cigarros forneceu quantidade média de nicotina de 31,5 mg e desvio padrão de 3mg. Ao nível de significância de 5%, podemos concordar com a afirmação do fabricante?

X - quantidade de nicotina por cigarro

$X \sim N(\mu, \sigma^2)$ μ e σ^2 desconhecidos

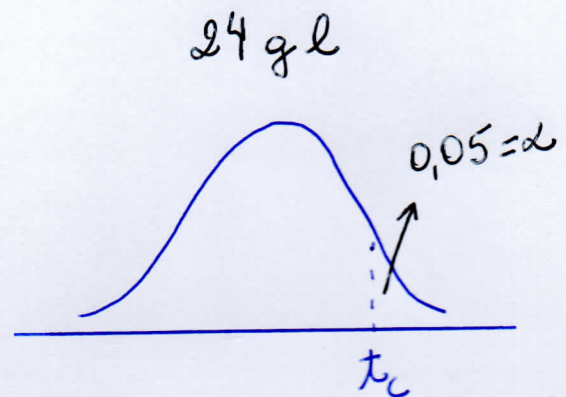
$$H_0: \mu = 30$$

$$H_a: \mu > 30$$

$$RC: T = \frac{\bar{X} - \mu_0}{s/\sqrt{n}} > t_c$$

$$n = 25 \quad \bar{x} = 31,5 \quad \alpha = 0,05$$

$$s = 3$$



$$g/l \quad P = 10\% \quad t_c = 1,711$$

$$\vdots \quad \vdots$$

$$24 - \dots - 1,711$$

$$RC: T > 1,711$$

$$T_{obs} = \frac{31,5 - 30}{3/5} = 2,5 \in RC$$

Rejeitamos H_0 . Ao nível de significância de 0,05, os dados sugerem que a quantidade média de nicotina é superior a 30 g.

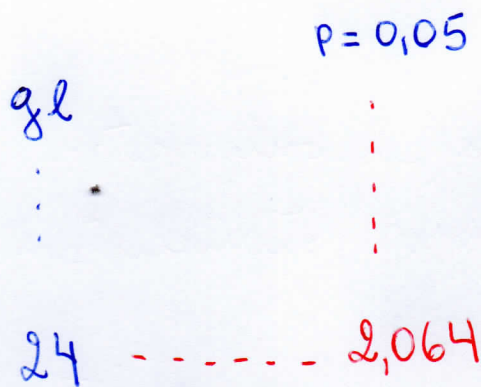
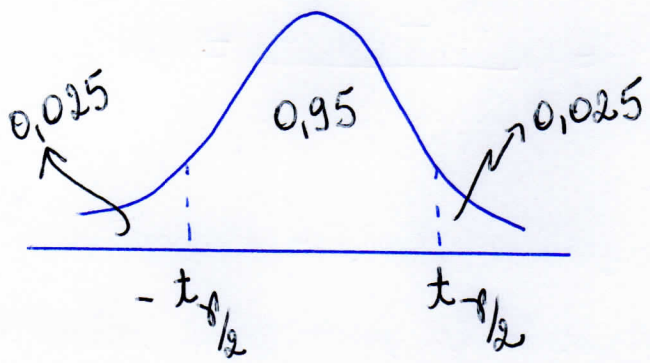
Intervalos de Confiança para a média da distribuição normal com variância desconhecida

$$\left[\bar{X} - t_{\gamma/2} \frac{S}{\sqrt{n}} ; \bar{X} + t_{\gamma/2} \frac{S}{\sqrt{n}} \right]$$

$$t_{\gamma/2} \text{ tal que } P(-t_{\gamma/2} \leq t \leq t_{\gamma/2}) = \gamma$$

γ coeficiente de confiança do intervalo

No exemplo, para $\gamma = 0,95$



$$IC: 31,5 \pm 2,064 \cdot \frac{3}{5}$$

$$[30,27; 32,73]$$