Estatística Aplicada

Inferência estatística com desvio padrão populacional desconhecido

Estimativa de proporção

Capítulo 10

Itens 10.4 a 10.5

Casos – população infinita

- Quando o desvio-padrão populacional (σ) é desconhecido, adota-se o desvio-padrão amostral (s).
- Para determinar os valores críticos do Intervalo de Confiança:
  - Amostras grandes $(\boldsymbol{n > 30})$: usamos a tabela de distribuição normal para os valores de z
  - Para amostras pequenas $(\boldsymbol{n \leq 30})$: usamos a tabela de distribuição t de Student (William S. Gosset, 1876-1937)

$$I_C = m \pm z \cdot \frac{s}{\sqrt{n}} \qquad n > 30$$

$$I_C = m \pm t \cdot \frac{s}{\sqrt{n}} \qquad n \leq 30$$

- Obs.: Quando o desvio padrão da população é conhecido, usamos SEMPRE o valor z.

Como funciona a tabela *t-Student*

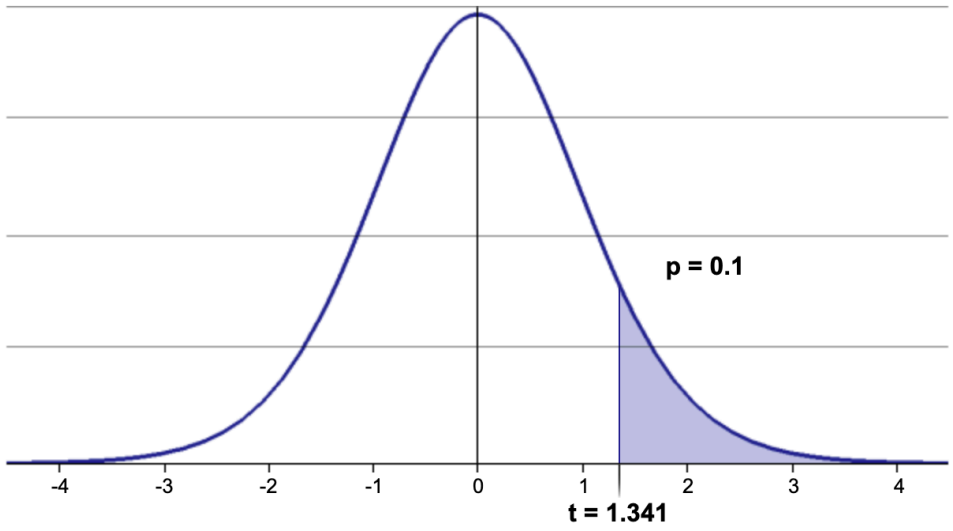
- A tabela *t-Student* utiliza a área que fica na cauda direita e os Graus de Liberdade. O valor de t fica no centro da tabela.
- Por exemplo: qual o valor de *t* acima do qual ficam 10% dos valores de uma amostra de 16 elementos.

Tabela 2 Distribuição de Student: coeficiente *t*

(Área no extremo da cauda)

Graus de liberdade	0,10 (10%)	0,05 (5%)	0,025 (2,5%)	0,01 (1%)	0,005 (0,5%)
1	3,078	6,314	12,706	31,821	63,657
2	1,886	2,920	4,303	6,965	9,925
3	1,638	2,353	3,182	4,541	5,841
4	1,533	2,132	2,776	3,747	4,604
5	1,476	2,015	2,571	3,365	4,032
6	1,440	1,943	2,447	3,143	3,707
7	1,415	1,895	2,365	2,998	3,499
8	1,397	1,860	2,306	2,896	3,355
9	1,383	1,833	2,262	2,821	3,250
10	1,371	1,812	2,228	2,764	3,169
11	1,363	1,796	2,201	2,718	3,106
12	1,356	1,782	2,179	2,681	3,055
13	1,350	1,771	2,160	2,650	3,012
14	1,345	1,761	2,145	2,624	2,977
15	1,341	1,753	2,131	2,602	2,947
16	1,337	1,746	2,120	2,583	2,921
17	1,333	1,740	2,110	2,567	2,898

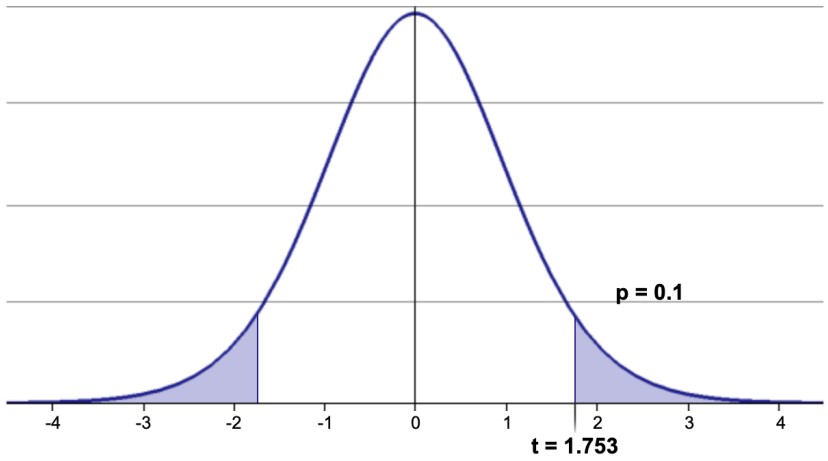
$$Gl = n - 1$$
$$Gl = 16 - 1 = 15$$



Como funciona a tabela t

- Para determinação do intervalo de confiança, porém, necessitamos os valores superior e inferior.
- Por exemplo, quais os valores críticos de *t* para um intervalo de 90% de confiança, em uma amostra de 16 elementos.
- Como a distribuição é simétrica, basta entrar na tabela com a metade do valor da porcentagem

$$p = \frac{100\% - 90\%}{2} = 5\% = 0{,}05$$



Graus de liberdade	0,10 (10%)	0,05 (5%)	0,025 (2,5%)
1	3,078	6,314	12,706
2	1,886	2,920	4,303
3	1,638	2,353	3,182
4	1,533	2,132	2,776
5	1,476	2,015	2,571
6	1,440	1,943	2,447
7	1,415	1,895	2,365
8	1,397	1,860	2,306
9	1,383	1,833	2,262
10	1,371	1,812	2,228
11	1,363	1,796	2,201
12	1,356	1,782	2,179
13	1,350	1,771	2,160
14	1,345	1,761	2,145
15	1,341	1,753	2,131
16	1,337	1,746	2,120

Tabela Completa

	Área na cauda superior								
gl	0,25	0,10	0,05	0,025	0,01	0,005	0,0025	0,001	0,0005
1	1,000	3,078	6,314	12,71	31,82	63,66	127,3	318,3	636,6
2	0,816	1,886	2,920	4,303	6,965	9,925	14,09	22,33	31,60
3	0,765	1,638	2,353	3,182	4,541	5,841	7,453	10,21	12,92
4	0,741	1,533	2,132	2,776	3,747	4,604	5,598	7,173	8,610
5	0,727	1,476	2,015	2,571	3,365	4,032	4,773	5,894	6,869
6	0,718	1,440	1,943	2,447	3,143	3,707	4,317	5,208	5,959
7	0,711	1,415	1,895	2,365	2,998	3,499	4,029	4,785	5,408
8	0,706	1,397	1,860	2,306	2,896	3,355	3,833	4,501	5,041
9	0,703	1,383	1,833	2,262	2,821	3,250	3,690	4,297	4,781
10	0,700	1,372	1,812	2,228	2,764	3,169	3,581	4,144	4,587
11	0,697	1,363	1,796	2,201	2,718	3,106	3,497	4,025	4,437
12	0,695	1,356	1,782	2,179	2,681	3,055	3,428	3,930	4,318
13	0,694	1,350	1,771	2,160	2,650	3,012	3,372	3,852	4,221
14	0,692	1,345	1,761	2,145	2,624	2,977	3,326	3,787	4,140
15	0,691	1,341	1,753	2,131	2,602	2,947	3,286	3,733	4,073
16	0,690	1,337	1,746	2,120	2,583	2,921	3,252	3,686	4,015
17	0,689	1,333	1,740	2,110	2,567	2,898	3,222	3,646	3,965
18	0,688	1,330	1,734	2,101	2,552	2,878	3,197	3,610	3,922
19	0,688	1,328	1,729	2,093	2,539	2,861	3,174	3,579	3,883
20	0,687	1,325	1,725	2,086	2,528	2,845	3,153	3,552	3,850
21	0,686	1,323	1,721	2,080	2,518	2,831	3,135	3,527	3,819
22	0,686	1,321	1,717	2,074	2,508	2,819	3,119	3,505	3,792
23	0,685	1,319	1,714	2,069	2,500	2,807	3,104	3,485	3,768
24	0,685	1,318	1,711	2,064	2,492	2,797	3,091	3,467	3,745
25	0,684	1,316	1,708	2,060	2,485	2,787	3,078	3,450	3,725
26	0,684	1,315	1,706	2,056	2,479	2,779	3,067	3,435	3,707
27	0,684	1,314	1,703	2,052	2,473	2,771	3,057	3,421	3,689
28	0,683	1,313	1,701	2,048	2,467	2,763	3,047	3,408	3,674
29	0,683	1,311	1,699	2,045	2,462	2,756	3,038	3,396	3,660
30	0,683	1,310	1,697	2,042	2,457	2,750	3,030	3,385	3,646
35	0,682	1,306	1,690	2,030	2,438	2,724	2,996	3,340	3,591
40	0,681	1,303	1,684	2,021	2,423	2,704	2,971	3,307	3,551
45	0,680	1,301	1,679	2,014	2,412	2,690	2,952	3,281	3,520
50	0,679	1,299	1,676	2,009	2,403	2,678	2,937	3,261	3,496
z	0,674	1,282	1,645	1,960	2,326	2,576	2,807	3,090	3,291

Exercícios

- Encontre os valores críticos de t para os seguintes casos

1) $n = 25$ e 90% de confiança.

2) $n = 8$ e 95% de confiança.

3) $n = 19$ e 80% de confiança.

4) $n = 13$ e 99,5% de confiança.

5) $n = 16$ e 99% de confiança.

6) $n = 10$ e 98% de confiança.

Exercícios

- Encontre os valores críticos de t para os seguintes casos

1) $n = 25$ e 90% de confiança.

$gl = 25 - 1 = 24$

$p = \frac{100 - 90}{2} = 5\% = 0{,}05$

$t = \pm 1{,}711$

2) $n = 8$ e 95% de confiança.

3) $n = 19$ e 80% de confiança.

4) $n = 13$ e 99,5% de confiança.

5) $n = 16$ e 99% de confiança.

6) $n = 10$ e 98% de confiança.

Exemplo

Duas amostragens, a primeira com 12 medidas e a segunda com 40 medidas, foram feitas para determinar um intervalo de 95% de confiança para a tensão média de ruptura de uma produção de tecido de algodão. Ambas amostras apresentaram uma tensão média de ruptura 7,38 unid. e um desvio padrão de 1,24 unid. Determine o intervalo de confiança nos dois casos. Compare e analise os resultados.

a) $n = 12$ ($n \leq 30$, usar "t")

$s = 1{,}24$ kg (amostra)

$\bar{m} = 7{,}38$ kg

$p = \frac{100-95}{2} = \frac{5}{2} = 2{,}5 \Rightarrow 0{,}025$

$gl = 12-1 = 11$

$t = \pm 2{,}201 \Rightarrow IC = 7{,}38 \pm 2{,}201 \cdot \frac{1{,}24}{\sqrt{12}} = 7{,}38 \pm 0{,}788$ kg

b) $n = 40$ ($n > 30$, usar "z")

$\alpha = 0{,}95 \Rightarrow \frac{\alpha}{2} = 0{,}475$

$z_{\frac{\alpha}{2}} = \pm 1{,}96 \Rightarrow IC = 7{,}38 \pm 1{,}96 \cdot \frac{1{,}24}{\sqrt{40}} = 7{,}38 \pm 0{,}384$ kg

Populações finitas

- Quando a população é finita e o tamanho da amostra é superior a 5% da população, devemos usar o fator de correção.
- Assim, as fórmulas ficam:

$$I_C = m \pm z \cdot \frac{s}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N-1}}$$

$$n > 30 \; e \; superior \; a \; 5\% \; da \; população$$

$$I_C = m \pm t \cdot \frac{s}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N-1}}$$

$$n \leq 30 \; e \; superior \; a \; 5\% \; da \; população$$

Resumo

$$Ic = m \pm z \text{ ou } t \cdot \frac{\sigma \text{ ou } s}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N-1}}$$

$z \text{ ou } t \Rightarrow$ usar t se $\begin{cases} \sigma \text{ for desconhecido e} \\ n \leq 30 \end{cases}$

$\sigma \text{ ou } s \Rightarrow$ usar o que for fornecido no problema

$\sqrt{\frac{N-n}{N-1}} \Rightarrow$ usar se $\begin{cases} N \text{ for conhecido e} \\ \frac{n}{N} > 5\%. \end{cases}$

Estimativa de uma proporção populacional

- Podemos desejar saber uma proporção populacional a partir da proporção de uma média.

- Por exemplo: em uma amostra de 120 lajotas de cerâmica, 15 delas apresentaram imperfeições que inviabilizaram sua utilização. Determinar a estimativa intervalar de 95% de confiança para a proporção de lajotas com defeito em toda a produção.

Estimativa de uma proporção (população infinita)

- Para determinar a proporção de uma população a partir de uma proporção identificada em uma amostra, usamos a seguinte relação
- Estimativa pontual
  - É a própria porporção da amostra $P = p$

- Estimativa intervalar
  - É dada pela fórmula

$$P = p \pm z \cdot \sqrt{\frac{p \cdot q}{n}}$$

O fator $E = z \cdot \sqrt{\frac{p \cdot q}{n}}$ é chamada de Erro de Estimativa ou margem de erro

Obs.: para a estimativa intervalar, utilizamos sempre o coeficiente z

  - Onde

$$p \Rightarrow proporção\ de\ x\ sucessos, p = \frac{x}{n}$$

$$q \Rightarrow proporção\ de\ fracassos, q = 1 - p$$

Exemplo

Em uma amostra de 120 lajotas de cerâmica, 15 delas apresentaram imperfeições que inviabilizaram sua utilização. Determinar:

a) A estimativa pontual da proporção populacional
b) O intervalo de 95% de confiança para a proporção de lajotas com defeito em toda a produção.
c) O erro de estimativa da proporção populacional

Exemplo

Uma pesquisa feita com 1850 eleitores revelou que 1134 votariam em certo candidato.

a) Construir um intervalo de 90% de confiança para a verdadeira proporção dos eleitores que pretendem votar nesse candidato
b) Determine a margem de erro dessa estimativa

Estimativa de uma proporção (população finita)

- Valem as mesmas considerações feitas na estimativa da média, sendo que para proporção, teremos

$$P = p \pm z \cdot \sqrt{\frac{p \cdot q}{n}} \cdot \sqrt{\frac{N - n}{N - 1}}$$

$$E = z \cdot \sqrt{\frac{p \cdot q}{n}} \cdot \sqrt{\frac{N - n}{N - 1}}$$

Relembrando, a população é considerada finita quando a amostra é superior a 5% da população

Exemplo

Uma indústria possui um total de N máquinas. O engenheiro mecânico selecionou uma amostra aleatória de n máquinas e constatou que 34 delas tiveram algum tipo de problema durante o mês anterior.

a) Construir um intervalo de 95% de confiança para a proporção de todas as máquinas com defeito, sendo N= 360 e n=100. Identifique a margem de erro dessa estimativa.
b) Construir um intervalo de 98% de confiança para a proporção de todas as máquinas com defeito, sendo agora N=1500 e n=60. Identifique a margem de erro dessa estimativa.

Determinação do tamanho da amostra

- Para populações infinitas

$$n = \frac{z^2 \cdot p \cdot q}{E^2}$$

Obs.: quando o valor de p é totalmente desconhecido, usamos $p = 0{,}5$

- Para populações finitas

Para o dimensionamento da amostra, sempre usamos *z*.

$$n = \frac{z^2 \cdot p \cdot q \cdot N}{(N-1) \cdot E^2 + z^2 \cdot p \cdot q}$$

Exemplo

Um componente eletrônico é produzido em uma empresa e um engenheiro de produção deseja dimensiona a amostra necessária para determinar a proporção de componentes defeituosos, com erro máximo de 3,2% e confiança de 98%, se:

a) Em um estudo anterior, a proporção de defeituosos foi de 4,8%.
b) Não se tem nenhuma informação sobre a proporção de defeituosos.
c) Refazer o item ”a” com 90% de confiança
d) Refazer o item “b” com 90% de confiança
e) Analise os resultados

Exercício

Um engenheiro eletricista de uma fábrica de aparelhos eletrônicos deseja estimar a proporção populacional dos aparelhos de som que apresentam algum tipo de defeito durante o período de garantia. Utilizando uma confiança de 99% e um erro de estimativa de 4%, determinar a quantidade mínima de aparelhos desse tipo que devem ser testados, se:

a) Em um estudo anterior, constatou-se que 12,5% dos aparelhos apresentam defeitos durante o período de garantia
b) Não se tem qualquer informação que possa sugerir o valor de p.

Exercícios

- Estimativa de proporção e tamanho da amostra
  - Página 208, 1 a 11