

Lista IC, tamanho de amostra e TH

1. Considere a amostra abaixo e construa um intervalo de confiança para a média populacional. Considere um nível de confiança de 99%.

17	23	22
19	23	23
21	18	20
13	17	16

Como não temos o desvio padrão populacional, devemos calcular o Intervalo de Confiança com base no desvio padrão amostral.

O Intervalo de Confiança é: $\left[\bar{x} - t_{n-1, \alpha/2} \cdot \frac{S}{\sqrt{n}}; \bar{x} + t_{n-1, \alpha/2} \cdot \frac{S}{\sqrt{n}} \right]$

Probabilidade do I.C. conter o valor da média populacional = $1 - \alpha = 99\% \Rightarrow \alpha = 1\%$

$$S = \text{desvio padrão da amostra} = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}}$$

\bar{x} = média da amostra = 19,33

$$\Rightarrow S = 3,23$$

n = número de elementos da amostra = 12

$$t_{n-1, \alpha/2} = t_{11, 0,5\%} = 3,106$$

Assim:

$$\text{I.C.} = \left[19,33 - 3,106 \cdot \frac{3,23}{\sqrt{12}}; 19,33 + 3,106 \cdot \frac{3,23}{\sqrt{12}} \right] = [19,33 - 2,90; 19,33 + 2,90]$$

$$\text{I.C.} = [16,43; 22,23]$$

$$\text{I.C.} = 19,33 \pm 2,90$$

2. Um estagiário deseja construir um intervalo de confiança para a resistência média de um certo tipo de parafuso. Este IC deve ter nível de confiança de pelo menos 99% e sua amplitude total não deve ser superior a 0,5 Kgf/mm². Para determinar o tamanho da amostra, ele analisou alguns parafusos (amostra piloto) e obteve os seguintes dados

$$N = 20$$

$$\text{Média Amostral} = 56,5 \text{ Kgf/mm}^2$$

$$\text{Desvio Padrão Amostral} = 1,5 \text{ Kgf/mm}^2$$

Para determinarmos o tamanho da amostra, utilizamos os dados da amostra piloto por meio da seguinte fórmula:

$$n = \left(\frac{t_{n'-1, \alpha/2} \cdot S}{e_0} \right)^2$$

n' = quantidade de elementos da amostra piloto = 20

Probabilidade do I.C. conter o valor da média populacional = $1 - \alpha = 99\% \Rightarrow \alpha = 1\%$

$$t_{n'-1, \alpha/2} = t_{19, 0,5\%} = 2,861$$

$$S = \text{desvio padrão amostral} = 1,5 \text{ kgf/mm}^2$$

e_0 = semi amplitude do Intervalo de Confiança $\leq 0,5 \text{ kgf/mm}^2$. Analisaremos o caso limite em que e_0 vale $0,5 \text{ kgf/mm}^2$.

$$\Rightarrow n = \left(\frac{2,861 \cdot 1,5}{0,5} \right)^2 = 73,67 \approx 74 \text{ elementos}$$

3. Sua empresa está pensando em anunciar seus produtos em uma estação de rádio. Uma pesquisa sobre audiência com 600 ouvintes mostrou que esta estação era a preferida de 48 deles. Construa um intervalo com 95% de confiança para a audiência da estação.

O Intervalo de Confiança é: $\left[p' - z_{\alpha/2} \cdot \sqrt{\frac{p'(1-p')}{n}}; p' + z_{\alpha/2} \cdot \sqrt{\frac{p'(1-p')}{n}} \right]$

p' = frequência relativa amostral = $48/600 = 0,08$

Probabilidade do I.C. conter o valor da frequência relativa populacional = $1 - \alpha = 95\%$
 $\Rightarrow \alpha = 5\%$

$z_{\alpha/2} = z_{2,5\%} =$ valor para o qual $P(0 \leq Z \leq z_{2,5\%}) = 0,5 - \frac{\alpha}{2} \Rightarrow$

$P(0 \leq Z \leq z_{2,5\%}) = 0,5 - \frac{0,05}{2} = 0,475 \Rightarrow z_{2,5\%} = 1,96$

n = tamanho da amostra = 600

Assim:

$$\text{I.C.} = \left[0,08 - 1,96 \cdot \sqrt{\frac{0,08(1-0,08)}{600}}; 0,08 + 1,96 \cdot \sqrt{\frac{0,08(1-0,08)}{600}} \right]$$

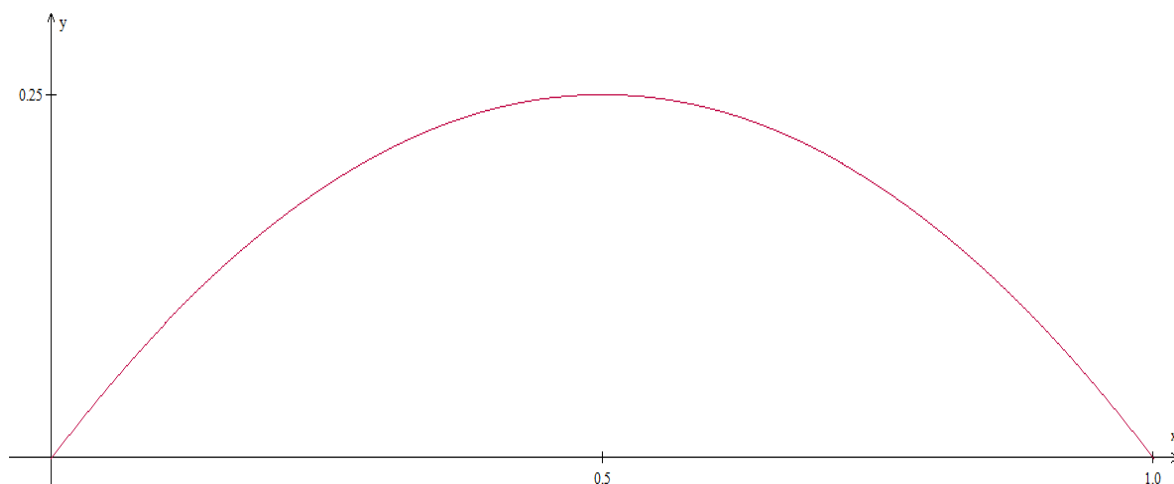
$$\text{I.C.} = [0,08 - 0,022; 0,08 + 0,022] = [0,058; 0,102]$$

4. Sabe-se que o percentual de pessoas que comprariam um determinado produto está entre 15% e 25%. Contudo, o gerente de produto gostaria de construir um intervalo de confiança ao nível de 99% para a parcela da população que compraria o produto cuja semi-amplitude não fosse superior a 2%. Qual deveria ser o tamanho da amostra?

Para descobrirmos o tamanho da amostra, devemos usar a seguinte fórmula:

$$n = \left(\frac{z_{\alpha/2}}{e_0} \right)^2 p(1-p)$$

O problema é que desconhecemos p (proporção populacional) e p' (proporção amostral). Porém podemos analisar o fator $p(1-p)$:



No eixo x temos p e no eixo y temos $p(1-p)$. É uma parábola com valor máximo igual a 0,25. Como sabemos que a proporção p de pessoas que comprariam um determinado produto está entre 15% e 25%, podemos analisar o caso limite em que p vale 0,25. Este é o caso no qual o fator $p(1-p)$ atinge o valor máximo para p entre 0,15 e 0,25. Desse modo:

$$n = \left(\frac{z_{\alpha/2}}{e_0} \right)^2 0,25(1 - 0,25) = \left(\frac{z_{\alpha/2}}{e_0} \right)^2 \cdot 0,1875$$

Probabilidade do I.C. conter o valor da frequência relativa populacional = $1 - \alpha = 99\% \Rightarrow \alpha = 1\%$

$z_{\alpha/2} = z_{0,5\%} =$ valor para o qual $P(0 \leq Z \leq z_{0,5\%}) = 0,5 - \frac{\alpha}{2} \Rightarrow$

$P(0 \leq Z \leq z_{0,5\%}) = 0,5 - \frac{0,01}{2} = 0,495 \Rightarrow z_{0,5\%} \approx 2,58$

e_0 = semi amplitude do Intervalo de Confiança $\leq 2\% = 0,02$. Analisaremos o caso limite em que e_0 vale 0,02.

$$\Rightarrow n = \left(\frac{2,58}{0,02} \right)^2 \cdot 0,1875 = 3120,2 \approx 3121 \text{ elementos.}$$

5. Um engenheiro está avaliando a precisão de uma máquina que produz esferas para rolamentos. Ele colheu uma amostra cujos dados estão abaixo, e analisou o diâmetro das esferas. Construa um intervalo de confiança ao nível de 90% para variância populacional.

$N = 20$

Média Amostral = 32 mm

Desvio Padrão Amostral = 1,2 mm

O Intervalo de Confiança é: $\left[\frac{(n-1) \cdot S^2}{X^2_{n-1, \alpha/2}}; \frac{(n-1) \cdot S^2}{X^2_{n-1, 1-\alpha/2}} \right]$

Probabilidade do I.C. conter o valor da variância populacional = $1 - \alpha = 90\% \Rightarrow \alpha = 10\%$

$S = \text{desvio padrão amostral} = 1,2 \text{ mm} \Rightarrow S^2 = 1,44 \text{ mm}^2$

$X^2_{n-1, \alpha/2} = X^2_{19, 5\%} = 30,144$

$X^2_{n-1, 1-\alpha/2} = X^2_{19, 95\%} = 10,117$

Assim:

$\text{I.C.} = \left[\frac{19 \cdot 1,44}{30,144}; \frac{19 \cdot 1,44}{10,117} \right] = [0,908; 2,704]$

6. Um técnico de basquete está analisando a possibilidade de contratar um novo ala para sua equipe. Ele recebeu a informação que este jogador tem uma média de pontos por partida superior a 16 e uma média de assistências por partida superior a 8. Para verificar esta informação ele analisou os dados das últimas 10 partidas (ver tabela abaixo) com nível de significância de 5%. As afirmações sobre pontuação e assistência são aceitáveis?

Partida	Pontos	Assistências
1	18	11
2	19	12
3	13	14
4	16	17
5	23	8
6	16	15
7	13	15
8	19	13
9	13	13
10	19	13

Para verificarmos se a afirmação sobre a pontuação é aceitável, devemos testar a seguinte hipótese:

$$H_0: \mu = 16$$

$$H_1: \mu > 16$$

Como o desvio padrão populacional é desconhecido, rejeita-se H_0 se $t_{n-1} > t_{n-1, \alpha}$.

$\alpha = \text{nível de significância} = 5\%$

$n = \text{tamanho da amostra} = 10$

$$t_{n-1} = t_9 = \frac{\bar{x} - \mu_0}{S/\sqrt{n}}$$

$\bar{x} = \text{média da amostra} = 16,9$

$\mu_0 = \text{valor a ser testado} = 16$

$$S = \text{desvio padrão da amostra} = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}} = 3,31$$

$$\Rightarrow t_9 = \frac{16,9 - 16}{3,31/\sqrt{10}} = 0,860$$

$$t_{n-1, \alpha} = t_{9, 5\%} = 1,833$$

Como $t_9 < t_{9,5\%}$, aceita-se H_0 , ou seja, a média de pontos não é superior a 16 ao nível de significância de 5%.

Para verificarmos se a afirmação sobre as assistências é aceitável, devemos testar a seguinte hipótese:

$$H_0: \mu = 8$$

$$H_1: \mu > 8$$

Como o desvio padrão populacional é desconhecido, rejeita-se H_0 se $t_{n-1} > t_{n-1,\alpha}$.

α = nível de significância = 5%

n = tamanho da amostra = 10

μ_0 = valor a ser testado = 8

$$t_{n-1} = t_9 = \frac{\bar{x} - \mu_0}{S/\sqrt{n}}$$

\bar{x} = média da amostra = 13,1

$$S = \text{desvio padrão da amostra} = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}} = 2,47$$

$$\Rightarrow t_9 = \frac{13,1-8}{2,47/\sqrt{10}} = 6,529$$

$$t_{n-1,\alpha} = t_{9,5\%} = 1,833$$

Como $t_9 > t_{9,5\%}$, rejeita-se H_0 , ou seja, a média de assistências é superior a 8 ao nível de significância de 5%.

7. Um gerente recebeu a informação de que um funcionário estava com um desempenho médio abaixo do aceitável. O mínimo aceitável para a produção média é de 500 unidades por dia. Na última semana (5 dias úteis) a produção média do funcionário foi de 475 unidade com desvio padrão de 10. O desempenho está baixo? Use um nível de confiança de 1%.

Para verificarmos se o desempenho está baixo, devemos testar a seguinte hipótese:

$$H_0: \mu = 500$$

$$H_1: \mu < 500$$

Como o desvio padrão populacional é desconhecido, rejeita-se H_0 se $t_{n-1} < -t_{n-1,\alpha}$.

α = nível de confiança = 1%

n = tamanho da amostra = 5

$$t_{n-1} = t_4 = \frac{\bar{x} - \mu_0}{S/\sqrt{n}}$$

\bar{x} = média da amostra = 475

μ_0 = valor a ser testado = 500

S = desvio padrão da amostra = 10

$$\Rightarrow t_4 = \frac{475-500}{10/\sqrt{5}} = -5,590$$

$$t_{n-1,\alpha} = t_{4,1\%} = 3,747$$

Como $t_4 < -t_{4,1\%}$, rejeita-se H_0 , ou seja, o desempenho está baixo (produção média abaixo de 500 unidades) ao nível de confiança de 1%.

8. Uma empresa fez uma campanha de comunicação e deseja agora avaliar o desempenho desta iniciativa. Nos critérios atuais, uma campanha é considerada de sucesso se conseguir atingir pelo menos 30% do seu público alvo. Em uma amostra de 40 indivíduos, 20 afirmaram que viram as peças publicitárias. Ao afirmar que a campanha é um sucesso, qual a probabilidade de estar cometendo um erro?

A empresa deseja avaliar se sua campanha é um sucesso, ou seja, se atinge pelo menos 30% do seu público alvo. Assim, ela deve realizar o seguinte teste:

$$H_0: p \leq 0,3$$

$$H_1: p > 0,3$$

ou, de modo análogo:

$$H_0: p = 0,3$$

$$H_1: p > 0,3$$

Ambos são iguais se levarmos em consideração que no segundo caso estamos eliminando a possibilidade de $p < 0,3$.

Queremos saber qual a probabilidade da empresa estar cometendo um erro ao afirmar que a campanha é um sucesso. Em outras palavras, queremos saber qual é a probabilidade de rejeitar H_0 tal que H_0 é verdadeira. Essa probabilidade é exatamente a definição do nível de significância (α).

$$P(\text{rejeitar } H_0 / H_0 \text{ é verdadeira}) = \alpha$$

Assim, queremos descobrir qual é o valor de α .

Observe que podemos supor que $np \geq 5$ e $n(1 - p) \geq 5$. H_0 será rejeitada se $z > z_\alpha$.

$$z = \frac{p' - p_0}{\sqrt{p_0(1 - p_0)/n}}$$

$$p' = \text{proporção observada na amostra} = 20/40 = 0,5$$

$$p_0 = \text{valor que queremos testar} = 0,3$$

$$\Rightarrow z = \frac{0,5 - 0,3}{\sqrt{0,3(1 - 0,3)/40}} = 2,76$$

No caso limite, $z = z_\alpha$ (é o caso onde α será menor). Assim, $z_\alpha = 2,76$. Procurando na tabela, encontramos $P(0 \leq Z \leq z_\alpha) = 0,4971 = 0,5 - \alpha \Rightarrow \alpha = 0,0029 = 0,29\%$.

9. Um comprador deve escolher entre dois fornecedores de parafusos. A escolha do fornecedor será baseada na resistência dos parafusos. Uma amostra de cada fabricante foi analisada e os resultados obtidos estão na tabela abaixo.

	Amostra	
	1	2
Tamanho	8	6
Média	56,38	64,5
Desvio padrão	0,74	3,15

- O fornecedor 1 é melhor por que a média de sua resistência é maior.
- O fornecedor 2 é melhor por que a média de sua resistência é maior.
- O fornecedor 1 é melhor por que a variância de sua resistência é menor.
- O fornecedor 2 é melhor por que a variância de sua resistência é menor.
- O fornecedor 1 é melhor por que a variância de sua resistência é maior.

Já que a escolha do comprador será baseada na resistência dos parafusos, devemos verificar qual fornecedor entrega a maior média de resistência. Assim, testaremos as hipóteses:

$$H_0: \mu_1 = \mu_2$$

$$H_1: \mu_1 < \mu_2 \text{ ou } \mu_2 > \mu_1$$

Perceba que os dados não são emparelhados (ou seja, as amostras não estão relacionadas por algum critério de modo que haja uma influência marcante sobre ambas). Além disso, não conhecemos os desvios padrão populacionais e não sabemos se são iguais. Assim, utilizaremos o método de Aspin-Welch, já que as amostras não são grandes: H_0 será rejeitada se $t_v < -t_{v,\alpha}$.

$$t_v = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{S_1^2/n_1 + S_2^2/n_2}} = \frac{56,38 - 64,5}{\sqrt{0,74^2/8 + 3,15^2/6}} = -6,187$$

Para $t_{v,\alpha}$, o número de graus de liberdade será dado por:

$$\nu = \frac{(w_1 + w_2)^2}{w_1^2/(n_1 + 1) + w_2^2/(n_2 + 1)} - 2, \text{ onde } w_1 = \frac{S_1^2}{n_1} \text{ e } w_2 = \frac{S_2^2}{n_2}$$

$$w_1 = \frac{0,74^2}{8} = 0,0925; w_2 = \frac{3,15^2}{6} = 1,65375$$

$$\Rightarrow \nu = \frac{(0,0925 + 1,65375)^2}{0,0925^2/(8 + 1) + 1,65375^2/(6 + 1)} - 2 = 5,59 \Rightarrow 5 < \nu < 6$$

Observando a tabela de distribuições de t de Student, podemos perceber que para $\nu = 5$ e $\nu = 6$, todos os graus de significância indicam que $t_\nu < -t_{\nu,\alpha}$. Assim, H_0 será rejeitada, e $\mu_2 > \mu_1$. O fornecedor 2 é melhor porque a média de sua resistência é maior.

10. Com os dados da amostra ao lado é possível afirmar, ao nível de significância de 5%, que a variância populacional é superior a 12,4?

220	218	202
210	206	208
206	207	218

Para verificarmos se a variância populacional é superior a 12,4, devemos testar a seguinte hipótese:

$$H_0: \sigma^2 = 12,4$$

$$H_1: \sigma^2 > 12,4$$

Rejeita-se H_0 se $X_{n-1}^2 > X_{n-1,\alpha}^2$.

α = nível de significância = 5%

n = tamanho da amostra = 9

$$X_{n-1}^2 = X_8^2 = \frac{(n-1) \cdot S^2}{\sigma_0^2}$$

$$S^2 = \text{variância da amostra} = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}$$

\bar{x} = média da amostra = 210,55

$$\Rightarrow S^2 = 41,77$$

σ_0^2 = valor a ser testado = 12,4

$$\Rightarrow X_8^2 = \frac{8 \cdot 41,77}{12,4} = 26,948$$

$$X_{n-1,\alpha}^2 = X_{8,5\%}^2 = 15,507$$

Como $X_8^2 > X_{8,5\%}^2$, rejeita-se H_0 , ou seja, a variância populacional é superior a 12,4 ao nível de significância de 5%.

11. Um dispositivo para redução do consumo de combustível está sendo testado em veículos. Para cada veículo, foi medido o consumo com e sem o dispositivo durante uma semana de uso. Os dados estão abaixo. Ao nível de 5% e significância, é possível afirmar que o dispositivo reduz o consumo de combustível?

Veículo	Consumo médio (Km/l)	
	Sem o dispositivo	Com o dispositivo
1	8	17
2	15	19
3	11	15
4	7	17
5	11	16
6	12	17
7	22	16
8	14	14
9	11	14

10	15	14
Média	12,60	15,90
Desvio Padrão	4,25	1,66

Devemos verificar se a média de consumo com o dispositivo é menor do que sem o dispositivo. Perceba que os dados são emparelhados, ou seja, os resultados das amostras estão relacionados dois a dois por algum critério de modo que haja uma influência marcante entre os pares, mas que essa influência seja igual em cada par. Nesse caso, a influência marcante é a característica de cada veículo. Porém, ela deve ser igual em cada valor de um determinado par, e se tomarmos a diferença entre cada par, essa influência tenderá a desaparecer, sobrando somente os efeitos da inserção do dispositivo. Assim, faremos a seguinte comparação:

$$H_0: \mu_d = \Delta = 0$$

$$H_1: \mu_d > \Delta = 0$$

onde $\mu_d = \mu_2 - \mu_1$ é a diferença das médias com o dispositivo e sem o dispositivo. Fazendo a diferença entre as amostras, temos uma nova tabela:

Veículo	Diferença d_i
1	9
2	4
3	4
4	10
5	5
6	5
7	-6
8	0
9	3
10	-1
Média \bar{d}	3,30
Desvio Padrão S_d	4,715

\bar{d} = média das diferenças = 3,30

$$S_d = \text{desvio padrão das diferenças} = \sqrt{\frac{\sum_{i=1}^n (d_i - \bar{d})^2}{n-1}} = 4,715$$

H_0 será rejeitada se $t_{n-1} > t_{n-1,\alpha}$.

α = nível de significância = 5%

$$t_{n-1} = t_9 = \frac{\bar{d} - \Delta}{S_d/\sqrt{n}} = \frac{3,3 - 0}{4,715/\sqrt{10}} = 2,213$$

$$t_{n-1,\alpha} = t_{9,5\%} = 1,833$$

Como $t_9 > t_{9,5\%}$, H_0 será rejeitada, e $\mu_d > 0$. Ou seja, o consumo é menor com o dispositivo.

Respostas

Q1

N 12
media 19,3
DP 3,2

GL 11
1-alfa 99,0%
alfa 1,0%
t 3,11

IC: 19,3 +/- 2,9

Q2

N	20
media	56,5
DP	1,5
GL	19
1-alfa	99,0%
alfa	1,0%
t	2,86
e	0,5
N amostra	74

Q3

N	600
f	48
p	8,00%
1-alfa	95,0%
alfa	5,0%
z	1,96
IC	8,0% +/- 2,2%

Q4

1-alfa	99,0%
alfa	1,0%
z	2,58
e	2,0%
p max	25,0%
N	3.121

Q5

N	20
media	32
DP	1,2
Var	1,44
1-alfa	90,0%
alfa	10,0%
$\chi^2_{\alpha/2}$	30,14
$\chi^2_{1-\alpha/2}$	10,12
$\sigma^2_{\text{Mínimo}}$	0,76
$\sigma^2_{\text{Máximo}}$	2,25

Q6

Pontos	Não rejeitar Ho
Média	16,9
DP	3,3
Var	11
Alfa	5%
N	10
GL	9
t _{Calculado}	0,86
t _{Crítico}	1,83

Assistências - Rejeitar Ho

Média	13,1
DP	2,5
Var	6,1
Alfa	5%
N	10
GL	9
t _{Calculado}	6,45
t _{Crítico}	1,83

Q7

Valor	500
Média	475
DP	10
Var	100
alfa	1%
N	5
GL	4
t _{Calculado}	-5,59
t _{Crítico}	-3,75
Está baixo sim	

Q8

N	40
f	20
p	50,0%
Valor	30,0%
Z _{Calculado}	2,53
Significância	0,57%

Q9 – Item B

Q10

Valor	12,4
Média	210,56
Variância	41,78
N	9
GL	8
alfa	5%

$\chi^2_{\text{Calculado}}$	26,95
$\chi^2_{\text{Crítico}}$	15,51

Rejeitar Ho

Q11

Veículo	Sem	Com	Dif
1	8	17	9
2	15	19	4
3	11	15	4
4	7	17	10
5	11	16	5
6	12	17	5
7	22	16	-6
8	14	14	0
9	11	14	3
10	15	14	-1

Média	3,3
DP	4,72
alfa	5,0%
N	10
GL	9
t _{Calculado}	2,21
t _{Crítico}	1,83

Rejeitar Ho