Resolução de Exercícios

Profª Julienne Borges

PUC Minas Virtual

Estimação de Parâmetros

Estimação Pontual

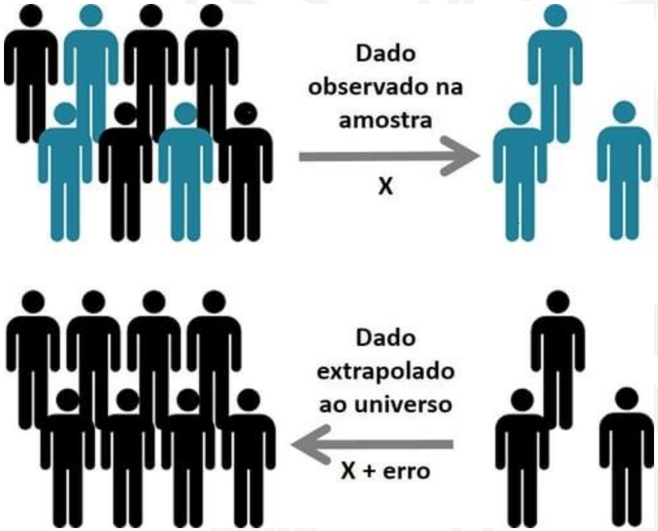
Parâmetro populacional	Estimador
μ (Média)	$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$
σ (Desvio padrão)	$s = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}}$
p (Proporção)	$\hat{p} = \frac{x}{n} = \frac{n^{\text{º}}\ de\ itens\ na\ amostra\ com\ a\ característica\ de\ interesse}{tamanho\ da\ amostra}$

Estimação intervalar ou Intervalo de confiança

Parâmetro	Estimativa pontual	Estimativa intervalar
μ	$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$	$\boldsymbol{\bar{x} \pm E}$
p	$\hat{p} = \frac{x}{n}$	$\boldsymbol{\hat{p} \pm E}$

Margem de erro

Quando utilizamos dados amostrais para estimar uma média populacional μ, a ***margem de erro***, denotada por ***E***, é a diferença máxima provável (com probabilidade 1 - α) entre a média amostral observada ($\bar{x}$) e a verdadeira média populacional μ. A margem de erro é chamada também de erro máximo da estimativa.



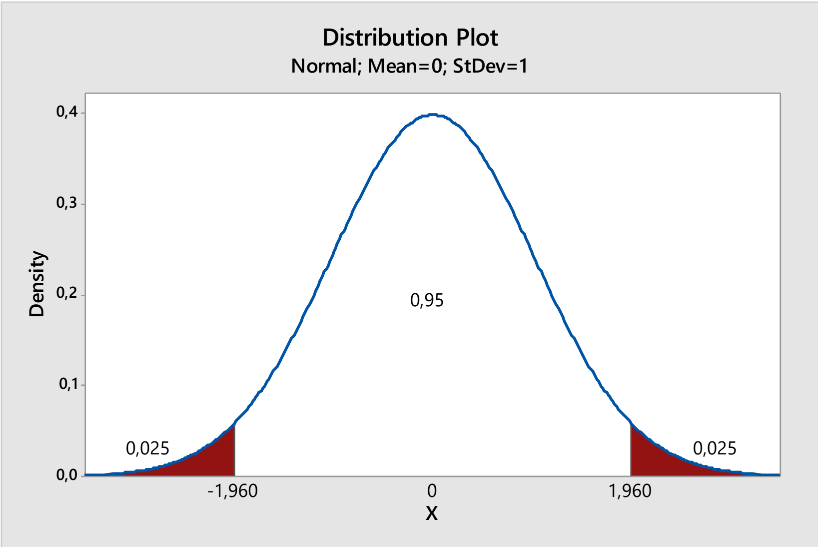
Fonte: https://www.netquest.com/blog/br/blog/br/amostragem-porque-funciona

Nível de confiança

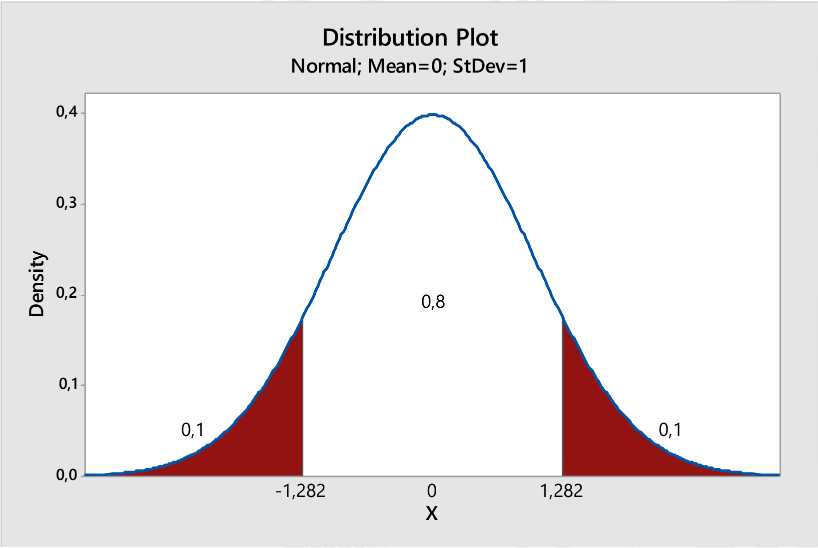
O ***grau de confiança*** é a probabilidade 1 - α de o intervalo de confiança conter o verdadeiro valor do parâmetro populacional.

Um intervalo de 95% de confiança indica que **a probabilidade do intervalo de confiança calculado conter o verdadeiro valor do parâmetro populacional é de 0,95**.

Valores críticos $\left(z_{\frac{\alpha}{2}}\right)$ vs Nível de confiança



Nível de confiança de 95%



Nível de confiança de 80%

Utilizando o Microsoft Excel

Intervalo de confiança para uma média populacional

σ é conhecido

$n =$	28	digite o tamanho da amostra
$\bar{x} =$	16,14	digite a média amostral
$\sigma =$	4,58	digite o desvio-padrão populacional (sigma)
NC =	0,95	digite o nível de confiança do IC
zc =	1,960	valor crítico da tabela Z
E =	1,696424568	margem de erro
IC (μ) =	14,44	intervalo de confiança
	17,83642457	

IC_Z_Média | IC_t_Média | IC_Z_Proporção

PUC Minas Virtual

Utilizando o R – IC para uma média

Cálculo de intervalo de confiança para uma média com σ desconhecido
Dados brutos (amostra de valores está disponível)

```
t.test(amostra1, conf.level=0.95)
```

`amostra1` → Vetor contendo a amostra da qual se construir um intervalo de confiança para a média populacional
`conf.level` → Grau de confiança do intervalo.

Utilizando o R – IC para uma média

Cálculo de intervalo de confiança para uma média com σ conhecido

Dados resumidos

```
#Intervalo de confiança para uma média
#Sigma É conhecido
#Fornecer os dados do problema
n <- 25
xbar <- 300
sigma <- 18.5
NC<-0.95
 p<-NC+((1-NC)/2)
 #valor critico
z<- qnorm(p, mean = 0, sd = 1, lower.tail = TRUE, log.p = 
FALSE)
#margem de erro
erro <- z*sigma/sqrt(n)
 #limite inferior do intervalo
li <- xbar-erro
 #limite superior do intervalo
ls <- xbar+erro
 #impressão dos resultados
#Valor crítico (z)
print(z)
#Margem de erro do intervalo de confiança
calculado (E)
print(erro)
#Intervalo de confiança [li ; ls]
print(li)
print(ls)
```

Utilizando o R – IC para uma média

Cálculo de intervalo de confiança para uma média com σ NÃO É conhecido

Dados resumidos

```
#Intervalo de confiança para uma média
#Sigma NÃO é conhecido
#Fornecer os dados do problema
n <- 25
xbar <- 300
s <- 18.5
NC<-0.95
p<-NC+((1-NC)/2)
#valor critico
t<- qt(p, n-1, lower.tail = TRUE, log.p = FALSE)
#margem de erro
erro <- t*s/sqrt(n)
#limite inferior do intervalo
li <- xbar-erro
#limite superior do intervalo
ls <- xbar+erro
#impressão dos resultados
#Valor crítico (t)
print(t)
#Margem de erro do intervalo de confiança calculado (E)
print(erro)
#Intervalo de confiança [li ; ls]
print(li)
print(ls)
```

Utilizando o R – IC para uma proporção

Cálculo de intervalo de confiança para uma proporção

```
prop.test(x, n, conf.level=0.95, correct=FALSE)
```

x → número de sucessos
n → tamanho da amostra
conf.level → Grau de confiança do intervalo.
correct → FALSE não utiliza a correção de continuidade de Yates ou TRUE utiliza a correção de continuidade de Yates

Observação importante! Essa função utiliza a distribuição Qui-quadrado para a construção do intervalo de confiança, ou seja, obtém valores ligeiramente diferentes do que vimos na disciplina.

Utilizando o R – IC para uma proporção

Cálculo de intervalo de confiança para uma proporção

```
binom.test(x, n, conf.level=0.95)
```

```
x → número de sucessos
n → tamanho da amostra
conf.level → Grau de confiança do intervalo.
```

A função *binom.test* utiliza a distribuição exata, a distribuição binomial, para construção do intervalo de confiança para uma proporção.

https://www.est.ufmg.br/~monitoria/Material/ApostilaR/InferenciaProporcoes.html

Utilizando o R – IC para uma proporção

Cálculo de intervalo de confiança para uma proporção

```
#Intervalo de confiança para uma proporção
#Fornecer os dados do problema
n <- 230
x <- 145
p_amostral <- x/n
NC<-0.92
p<-NC+((1-NC)/2)
#valor critico
z<- qnorm(p, mean = 0, sd = 1, lower.tail = TRUE, log.p
= FALSE)
#margem de erro
erro <- z*(sqrt((p_amostral*(1-p_amostral))/n))
```

```
#limite inferior do intervalo
li <- p_amostral-erro
#limite superior do intervalo
ls <- p_amostral+erro
#impressão dos resultados
#Valor crítico (z)
print(z)
#Margem de erro do intervalo de confiança
calculado (E)
print(erro)
#Intervalo de confiança [li ; ls]
print(li)
print(ls)
```

Exercício 1

Os resíduos industriais jogados nos rios, muitas vezes, absorvem o oxigênio necessário a respiração de peixes e de outras formas de vida aquática. Seis amostras de água retiradas de um rio revelaram os índices: 4,9; 5,1; 4,9; 5,0; 5,0 e 4,7 ppm de oxigênio dissolvido. Construir o intervalo com 95% de confiança para a verdadeira média de oxigênio dissolvido, em ppm e interprete.

Exercício 1

Os resíduos industriais jogados nos rios, muitas vezes, absorvem o oxigênio necessário a respiração de peixes e de outras formas de vida aquática. Seis amostras de água retiradas de um rio revelaram os índices: 4,9; 5,1; 4,9; 5,0; 5,0 e 4,7 ppm de oxigênio dissolvido. Construir o intervalo com 95% de confiança para a verdadeira média de oxigênio dissolvido, em ppm e interprete.

Parâmetro: quantidade média de oxigênio dissolvido (em ppm)

↓

σ NÃO é conhecido

↓

Tabela t-Student

↓

Dados brutos

PUC Minas Virtual

Exercício 1

Utilizando a função t.test do R, temos:

```
oxigenio<-c(4.9, 5.1, 4.9, 5.0, 5.0, 4.7)
t.test(oxigenio, conf.level=0.95)
```

Pode-se dizer com 95% de confiança que a verdadeira média de oxigênio dissolvido no rio estudado pode variar de 4,79 ppm a 5,08 ppm.

```
One Sample t-test

data:  oxigenio
t = 88.447, df = 5, p-value =
3.502e-09
alternative hypothesis: true mean
is not equal to 0
95 percent confidence interval:
 4.789953 5.076714
sample estimates:
mean of x
 4.933333
```

Exercício 2

Um engenheiro de computação está desenvolvendo um novo sistema para executar certo programa e pretende determinar o tempo médio necessário para a sua realização. Para tanto, executou esse programa por 20 vezes obtendo um tempo médio de 352 segundos e um desvio padrão de 8 segundos. Construa um intervalo de 99% de confiança para o verdadeiro tempo médio para execução desse programa.

Parâmetro: Tempo médio para execução desse programa (em segundos).

↓

σ NÃO é conhecido

↓

Tabela t-Student

↓

Dados resumidos

n=20; $\bar{x}$=352 ; s=8

Exercício 2

Parâmetro: Tempo médio para execução desse programa (em segundos).

↓

σ NÃO é conhecido

↓

Tabela t-Student

↓

Dados resumidos

n=20; $\bar{x}$=352 ; s=8

Intervalo de confiança para uma média populacional

σ NÃO é conhecido

	A	B	C	D
1		**Intervalo de confiança para uma média populacional**		
2			**σ NÃO é conhecido**	
3				
4		n =	20	digite o tamanho da amostra
5		$\bar{x}$ =	352,00	digite a média amostral
6		s =	8,00	digite o desvio-padrão amostral
7		NC =	0,99	digite o nível de confiança do IC
8		GL =	**19**	graus de liberdade
9		tc =	**2,861**	valor crítico da tabela t
10		E =	**5,117795407**	margem de erro
11		IC (μ) =	**346,8822046**	intervalo de confiança
12			**357,1177954**	
13				

IC_Z_Média | **IC_t_Média** | IC_Z_Proporção | +

Exercício 2

```
#Intervalo de confiança para uma média
#Sigma NÃO é conhecido
#Fornecer os dados do problema
n <- 20
xbar <- 352
s <- 8
NC<-0.99
p<-NC+((1-NC)/2)
#valor critico
t<- qt(p, n-1, lower.tail = TRUE, log.p = FALSE)
#margem de erro
erro <- t*s/sqrt(n)
#limite inferior do intervalo
li <- xbar-erro
#limite superior do intervalo
ls <- xbar+erro
#impressão dos resultados
#Valor crítico (t)
print(t)
#Margem de erro do intervalo de confiança calculado (E)
print(erro)
#Intervalo de confiança [li ; ls]
print(li)
print(ls)
```

Resultado obtido pelo R

```
> #impressão dos resultados
> #Valor crítico (t)
> print(t)
[1] 2.860935
> #Margem de erro do intervalo de confiança calculado (E)
> print(erro)
[1] 5.117795
> #Intervalo de confiança [li ; ls]
> print(li)
[1] 346.8822
> print(ls)
[1] 357.1178
```

	A	B	C	D
1		**Intervalo de confiança para uma média populacional**		
2		**σ NÃO é conhecido**		
3				
4		n =	20	digite o tamanho da amostra
5		$\bar{x}$ =	352,00	digite a média amostral
6		s =	8,00	digite o desvio-padrão amostral
7		NC =	0,99	digite o nível de confiança do IC
8		GL =	**19**	graus de liberdade
9		tc =	**2,861**	valor crítico da tabela t
10		E =	**5,117795407**	margem de erro
11		IC (μ) =	**346,8822046**	intervalo de confiança
12			**357,1177954**	
13				

IC_Z_Média | **IC_t_Média** | IC_Z_Proporção | +

Pode-se dizer com 99% de confiança que o verdadeiro tempo médio para execução desse programa varia de 346,88 a 357,12 segundos.

Exercício 3

O sistema gerenciador de um hotel acompanha a porcentagem de reservas feitas pela internet para ajustar sua propaganda e o sistema de reservas pela web. Tais dados são binários: ou a reserva é feita pela internet ou não é. Na semana passada uma amostra aleatória de 25 reservas foi selecionada mostrando que 15 haviam sido realizadas pela web. Encontre e interprete um intervalo de confiança de 85% para a verdadeira proporção de reservas realizadas pela internet.

Exercício 3

Parâmetro: Proporção de reservas para o hotel realizadas pela internet.

↓

Dados resumidos:

n=25; x=15; $\hat{p}$=?

O sistema gerenciador de um hotel acompanha a porcentagem de reservas feitas pela internet para ajustar sua propaganda e o sistema de reservas pela web. Tais dados são binários: ou a reserva é feita pela internet ou não é. Na semana passada uma amostra aleatória de 25 reservas foi selecionada mostrando que 15 haviam sido realizadas pela web. Encontre e interprete um intervalo de confiança de 85% para a verdadeira proporção de reservas realizadas pela internet.

Exercício 3

Parâmetro: Proporção de reservas para o hotel realizadas pela internet.

↓

Dados resumidos:

n=25; x=15; $\hat{p}$=?

	A	B	C	D
1		**Intervalo de confiança para uma proporção populacional**		
2				
3		n =	25	digite o tamanho da amostra
4		$\hat{p}$ =	0,60	digite a proporção de sucessos na amostra [0 ; 1]
5		NC =	0,85	digite o nível de confiança do IC
6		zc =	**1,440**	valor crítico da tabela Z
7		E =	**0,141044703**	margem de erro
8		IC (p) =	**0,458955297**	intervalo de confiança
9			**0,741044703**	
10				
11				
12				
13				

IC_Z_Média | IC_t_Média | **IC_Z_Proporção** | +

Exercício 3

```
#Intervalo de confiança para uma proporção
#Fornecer os dados do problema
n <- 25
x <- 15
p_amostral <- x/n
NC<-0.85
p<-NC+((1-NC)/2)
#valor critico
z<- qnorm(p, mean = 0, sd = 1, lower.tail = TRUE,
log.p = FALSE)
#margem de erro
erro <- z*(sqrt((p_amostral*(1-p_amostral))/n))
#limite inferior do intervalo
li <- p_amostral-erro
#limite superior do intervalo
ls <- p_amostral+erro
#impressão dos resultados
#Valor crítico (z)
print(z)
#Margem de erro do intervalo de confiança
calculado (E)
print(erro)
#Intervalo de confiança [li ; ls]
print(li)
print(ls)
```

Resultado obtido pelo R

```
> #impressão dos resultados
> #Valor crítico (z)
> print(z)
[1] 1.439531
> #Margem de erro do intervalo de confiança calculado (E)
> print(erro)
[1] 0.1410447
> #Intervalo de confiança [li ; ls]
> print(li)
[1] 0.4589553
> print(ls)
[1] 0.7410447
```

Intervalo de confiança para uma proporção populacional

n =	25	digite o tamanho da amostra
$\hat{p}$ =	0,60	digite a proporção de sucessos na amostra [0 ; 1]
NC =	0,85	digite o nível de confiança do IC
zc =	1,440	valor crítico da tabela Z
E =	0,141044703	margem de erro
IC (p) =	0,458955297	intervalo de confiança
	0,741044703	

IC_Z_Média | IC_t_Média | IC_Z_Proporção

Pode-se dizer com 85% de confiança que a verdadeira proporção de reservas do hotel realizadas pela internet varia de 45,9% a 74,1%.

Aumentando o tamanho amostral, n=25 → n=200, mantendo o $\hat{p}$=0,6 e o nível de 85% de confiança.

	A	B	C	D
1		**Intervalo de confiança para uma proporção populacional**		
2				
3		n =	200	digite o tamanho da amostra
4		$\hat{p}$ =	0,60	digite a proporção de sucessos na amostra [0 ; 1]
5		NC =	0,85	digite o nível de confiança do IC
6		zc =	**1,440**	valor crítico da tabela Z
7		E =	**0,049866833**	margem de erro
8		IC (p) =	**0,550133167**	intervalo de confiança
9			**0,649866833**	
10				
11				
12				
13				

IC_Z_Média | IC_t_Média | **IC_Z_Proporção** | +

n=25; $\hat{p}$=0,6
E=0,1410
I.C.=[0,4589 ; 0,7410]
A=0,2821

↓

n=200; $\hat{p}$=0,6
E=0,049
I.C.=[0,5501 ; 0,6499]
A=0,0998

Mantendo o tamanho amostral n=25, mantendo o $\hat{p}$=0,6 e aumentando o nível de confiança de 85% para 95%.

	A	B	C	D
1		Intervalo de confiança para uma proporção populacional		
2				
3		n =	25	digite o tamanho da amostra
4		$\hat{p}$ =	0,60	digite a proporção de sucessos na amostra [0 ; 1]
5		NC =	0,95	digite o nível de confiança do IC
6		zc =	1,960	valor crítico da tabela Z
7		E =	0,192036467	margem de erro
8		IC (p) =	0,407963533	intervalo de confiança
9			0,792036467	
10				
11				
12				
13				

IC_Z_Média | IC_t_Média | **IC_Z_Proporção** | +

n=25; $\hat{p}$=0,6;
85% de confiança
E-0,1410
I.C.=[0,4589 ; 0,7410]
A=0,2821

↓

n=25; $\hat{p}$=0,6;
95% de confiança
E=0,1920
I.C.=[0,408 ; 0,792]
A=0,384

Ponderações importantes

Aumentar o tamanho (amplitude) do intervalo de confiança significa reduzir a precisão da estimativa por intervalo. Para aumentar a precisão da estimativa temos que reduzir o tamanho (amplitude) do intervalo.

Só podemos fazer isto através de três maneiras:
1) reduzir o grau de confiança do intervalo;
2) aumentar o tamanho n da amostra;
3) reduzir a variância da população.
Como a variância da população geralmente é um dado do problema, temos apenas as duas primeiras opções.

Ponderações importantes

Se não podemos alterar o tamanho da amostra, quando aumentamos a precisão do intervalo somos obrigados a reduzir o seu grau de confiança e quando diminuímos a precisão automaticamente aumentamos o seu grau de confiança.

Na verdade só existe uma maneira de aumentarmos simultaneamente a precisão e confiança do intervalo: aumentarmos o tamanho da amostra!

PUC Minas
Virtual