

PROBABILIDADE E INFERÊNCIA ESTATÍSTICA COM R - MÓDULO 2

Prof^a. Angélica Maria Tortola Ribeiro

October 6, 2020

UTFPR UNIVERSIDADE
QUE TRANSFORMA



MINISTÉRIO DA
EDUCAÇÃO



- **População:** É o conjunto de todos os elementos sob investigação com pelo menos uma característica em comum.
- **Amostra:** É qualquer subconjunto da população.

Exemplos:

- 1 P1: Todos os moradores de Curitiba.
- 2 P2: Todos os estudantes da UTFPR.
- 3 P3: Todos os celulares produzidos por uma fábrica.
- 4 P4: Todos os consumidores de algum produto.

- **População:** É o conjunto de todos os elementos sob investigação com pelo menos uma característica em comum.
- **Amostra:** É qualquer subconjunto da população.

Exemplos:

- 1 P1: Todos os moradores de Curitiba.
- 2 P2: Todos os estudantes da UTFPR.
- 3 P3: Todos os celulares produzidos por uma fábrica.
- 4 P4: Todos os consumidores de algum produto.

- **População:** É o conjunto de todos os elementos sob investigação com pelo menos uma característica em comum.
- **Amostra:** É qualquer subconjunto da população.

Exemplos:

- 1 P1: Todos os moradores de Curitiba.
- 2 P2: Todos os estudantes da UTFPR.
- 3 P3: Todos os celulares produzidos por uma fábrica.
- 4 P4: Todos os consumidores de algum produto.

- **População:** É o conjunto de todos os elementos sob investigação com pelo menos uma característica em comum.
- **Amostra:** É qualquer subconjunto da população.

Exemplos:

- 1 P1: Todos os moradores de Curitiba.
- 2 P2: Todos os estudantes da UTFPR.
- 3 P3: Todos os celulares produzidos por uma fábrica.
- 4 P4: Todos os consumidores de algum produto.

Population and Sample

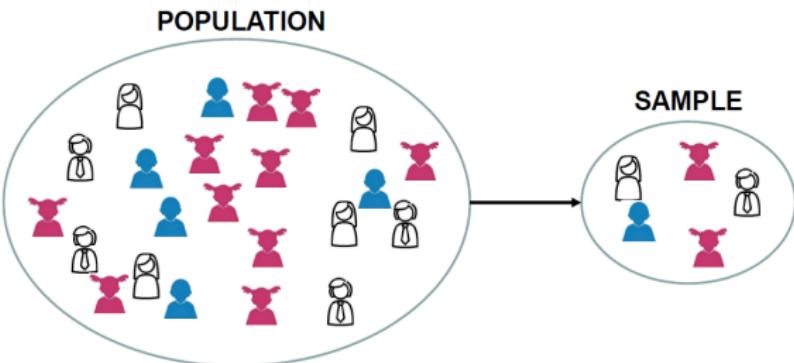


FIGURE: Representação de população e amostra.

■ **Variáveis:** São as características de interesse de uma população.

Exemplos:

1 **P1:** altura, peso, nº de filhos, renda familiar, grau de escolaridade (F, M, S), tem ensino superior (sim ou não), classe social (baixa, média, alta), sexo (masculino, feminino).

2 **P3:** possíveis variáveis são: peso do aparelho, nº de defeitos por aparelho, tem defeito (sim ou não), tamanho do aparelho (pequeno, médio, grande), nº de aparelhos produzidos por hora.

■ **Variáveis:** São as características de interesse de uma população.

Exemplos:

- 1 **P1:** altura, peso, nº de filhos, renda familiar, grau de escolaridade (F, M, S), tem ensino superior (sim ou não), classe social (baixa, média, alta), sexo (masculino, feminino).
- 2 **P3:** possíveis variáveis são: peso do aparelho, nº de defeitos por aparelho, tem defeito (sim ou não), tamanho do aparelho (pequeno, médio, grande), nº de aparelhos produzidos por hora.

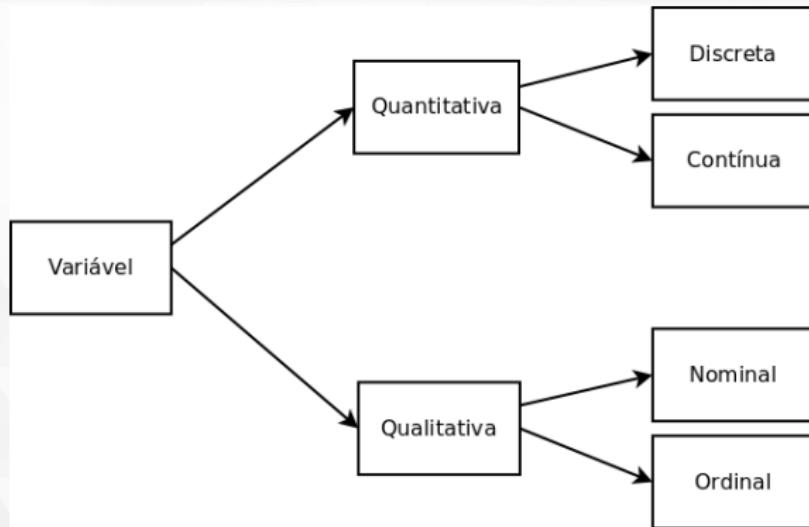


FIGURE: Classificação das variáveis.

TABLE: Classificação das variáveis

Discretas	Contínuas	Ordinais	Nominais
nº de filhos	peso	tamanho	sexo
nº de defeitos	altura	classe social	tem defeito
nº de aparelhos	renda familiar	Grau de escolaridade	tem ensino superior

Exercício: Apresentar 2 exemplos de variáveis para cada uma das classificações da Tabela 1, considerando as populações **P2** e **P4**.

População

Características

In
Estatística

Probabilidade

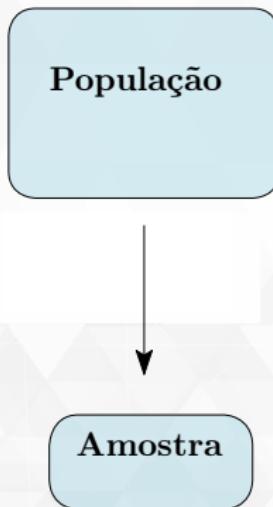
UTFPR UNIVERSIDADE
QUE TRANSFORMA

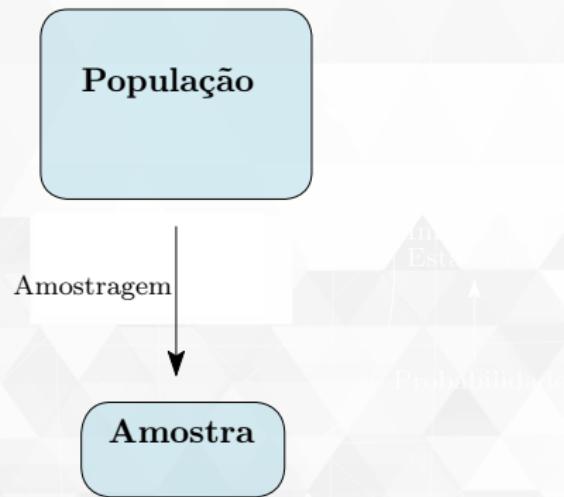
DAEST
DEPARTAMENTO ACADÉMICO DE ESTATÍSTICA

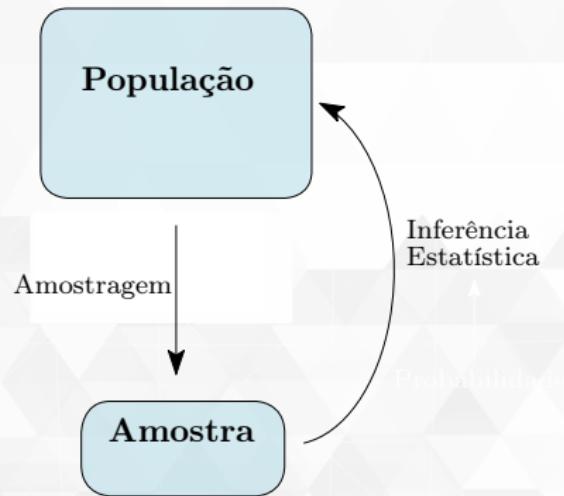
UTP
UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARAÍBA

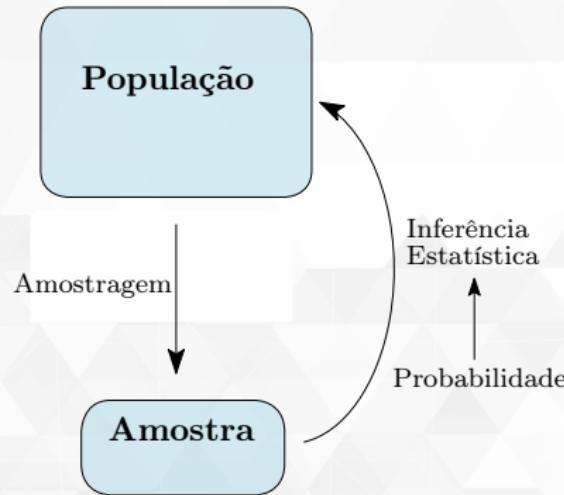
MINISTÉRIO DA
EDUCAÇÃO

PÁTRIA AMADA
BRASIL
GOVERNO FEDERAL









INFERÊNCIA ESTATÍSTICA

Conjunto de técnicas utilizadas para tirar conclusões sobre determinada(s) característica(s) da população, a partir de informações colhidas de uma amostra.

AMOSTRA ALEATÓRIA (A.A.)

Uma a.a. de tamanho n de uma população X , com dada distribuição, é o conjunto de n variáveis aleatórias independentes X_1, X_2, \dots, X_n , cada uma com a mesma distribuição de X .



PARÂMETROS E ESTATÍSTICAS

População

Características

In Estatística

Probabilidade

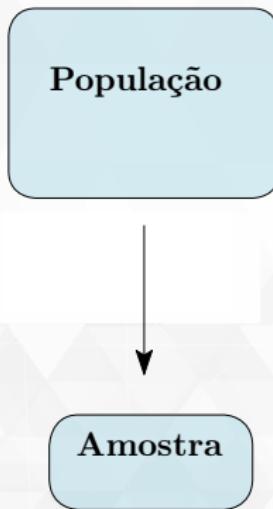
UTFPR UNIVERSIDADE
QUE TRANSFORMA

DAEST
DEPARTAMENTO ACADÉMICO DE ESTATÍSTICA

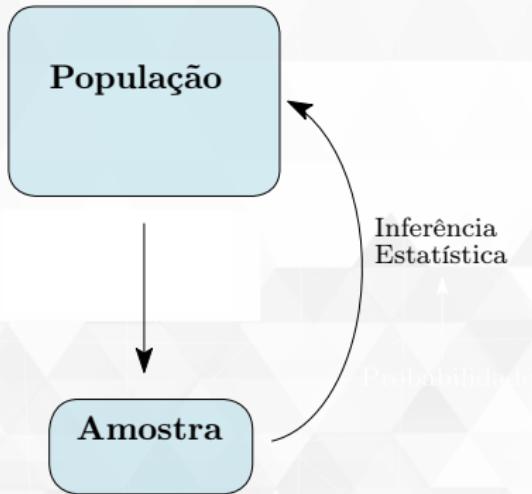
UTP
UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ

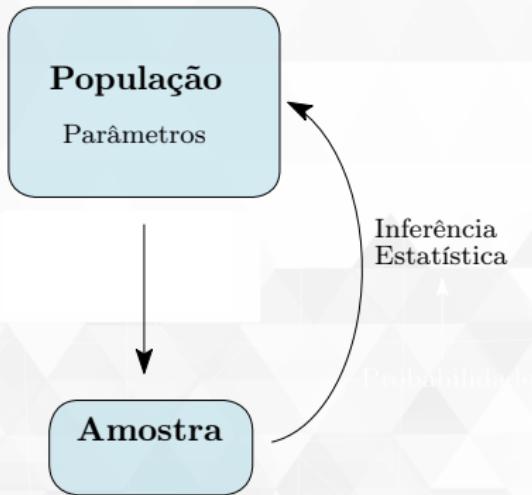
MINISTÉRIO DA
EDUCAÇÃO

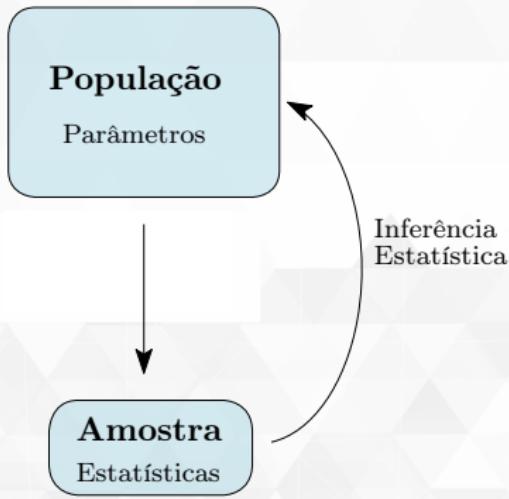
PÁTRIA AMADA
BRASIL
GOVERNO FEDERAL

In
Estad

Probabilidade







PARÂMETROS

São medidas utilizadas para descrever características da população.

ESTATÍSTICAS

São características da amostra, ou seja, uma estatística T é uma função de X_1, X_2, \dots, X_n .

P1: Moradores de Curitiba
Variável X: Altura

Altura

$$\bar{x} = \frac{X_1 + X_2 + \dots + X_n}{n}$$

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

$$S = \sqrt{S^2}$$

P1: Moradores de Curitiba
Variável X: Altura

$$X \sim N(\mu, \sigma^2)$$



AMOSTRA

$$\begin{aligned} S^2 &= \frac{1}{n-1} \sum_{i=1}^{n-1} (X_i - \bar{X})^2 \\ S &= \sqrt{S^2} \end{aligned}$$

estatísticas

UTP
UNIVERSIDADE
QUE TRANSFORMA

DAEST
DEPARTAMENTO ACADÉMICO DE ESTATÍSTICA

UTP
UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ

MINISTÉRIO DA
EDUCAÇÃO

PÁTRIA AMADA
BRASIL
GOVERNO FEDERAL

P1: Moradores de Curitiba
Variável X: Altura

$$X \sim N(\mu, \sigma^2)$$



Parâmetros

ESTIMATIVAS

$$\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n}$$

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

estatísticas

UTP
UNIVERSIDADE
QUE TRANSFORMA

DAEST
DEPARTAMENTO ACADÉMICO DE ESTATÍSTICA

UTP
UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ

MINISTÉRIO DA
EDUCAÇÃO

PÁTRIA AMADA
BRASIL
GOVERNO FEDERAL

P1: Moradores de Curitiba
Variável X: Altura

$$X \sim N(\mu, \sigma^2)$$



Parâmetros

Amostra

P1: Moradores de Curitiba
Variável X: Altura

$$X \sim N(\mu, \sigma^2)$$



Parâmetros

Amostra

$$\begin{aligned}\bar{X} &= \frac{X_1 + X_2 + \dots + X_n}{n} \\ S^2 &= \sum_{i=1}^n \frac{(X_i - \bar{X})^2}{n-1} \\ S &= \sqrt{S^2}\end{aligned}$$

P1: Moradores de Curitiba
Variável X: Altura

$$X \sim N(\mu, \sigma^2)$$



Parâmetros

Amostra

$$\begin{aligned} \bar{X} &= \frac{X_1 + X_2 + \dots + X_n}{n} \\ S^2 &= \sum_{i=1}^n \frac{(X_i - \bar{X})^2}{n-1} \\ S &= \sqrt{S^2} \end{aligned} \rightarrow \text{Estatísticas}$$

P1: Moradores de Curitiba
Variável X: Tem filhos (S, N)

01 - São Francisco

Parametro

estatística



PARÂMETROS E ESTATÍSTICAS

P1: Moradores de Curitiba
Variável X: Tem filhos (S, N)
 $X \sim Bernoulli(p)$

0, 1, ..., n

Parâmetro

Curitiba, PR, 2018. A variável X é uma

estatística

UTP
UNIVERSIDADE
QUE TRANSFORMA

DAEST
DEPARTAMENTO ACADÉMICO DE ESTATÍSTICA

UTP
UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ

MINISTÉRIO DA
EDUCAÇÃO

PÁTRIA AMADA
BRASIL
GOVERNO FEDERAL

P1: Moradores de Curitiba
Variável X: Tem filhos (S, N)
 $X \sim Bernoulli(p)$

$$X = \begin{cases} 0, & \text{se Fracasso} \\ 1, & \text{se Sucesso} \end{cases}$$

Parâmetro

ESTATÍSTICA

P1: Moradores de Curitiba

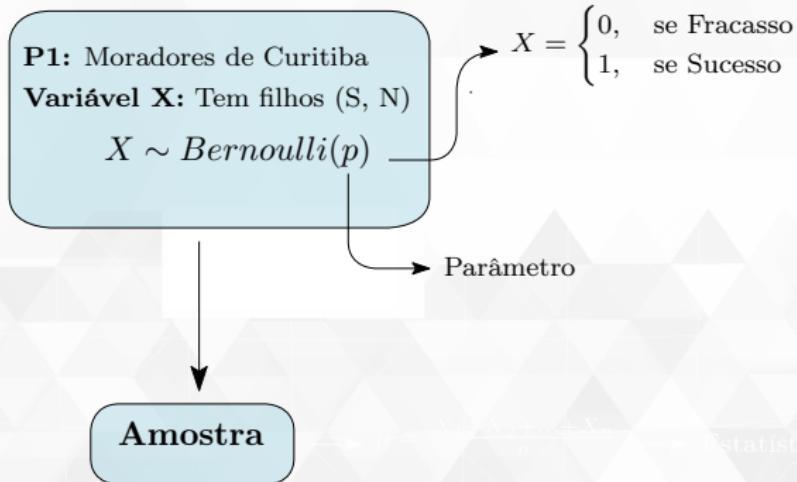
Variável X: Tem filhos (S, N)

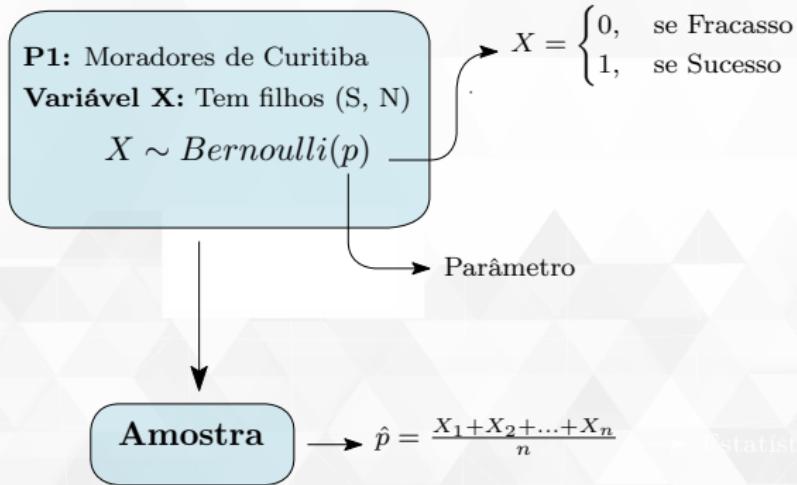
$X \sim Bernoulli(p)$

$$X = \begin{cases} 0, & \text{se Fracasso} \\ 1, & \text{se Sucesso} \end{cases}$$

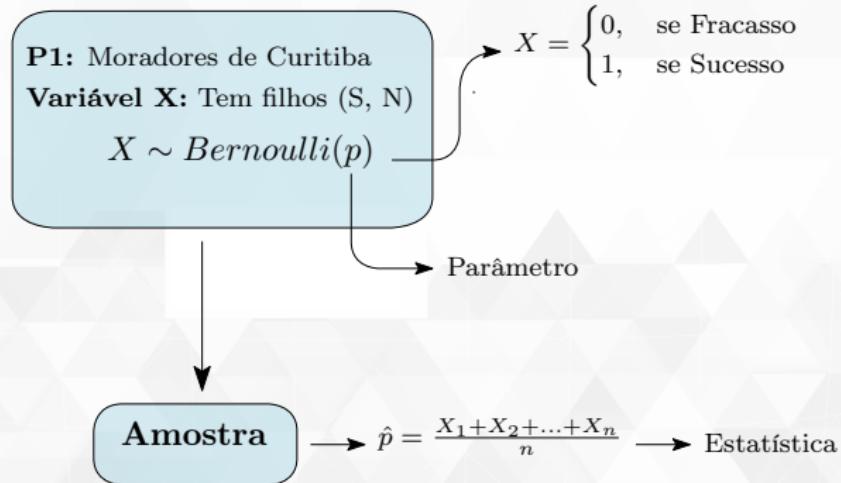
Parâmetro

PARÂMETROS E ESTATÍSTICAS





PARÂMETROS E ESTATÍSTICAS



- **Exemplos de parâmetros:** média populacional, variância populacional, desvio-padrão populacional, proporção populacional;
- **Exemplos de estatísticas:** média amostral, variância amostral, desvio-padrão amostral, proporção amostral;

TABLE: Parâmetros e Estatísticas.

Medidas	Parâmetros	Estatísticas
Média	$\mu = \frac{X_1 + X_2 + \dots + X_N}{N}$	$\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n}$
Variância	$\sigma^2 = \sum_{i=1}^N \frac{(X_i - \mu)^2}{N}$	$S^2 = \sum_{i=1}^n \frac{(X_i - \bar{X})^2}{n-1}$ ou $\hat{\sigma}^2 = \sum_{i=1}^n \frac{(X_i - \bar{X})^2}{n}$
Desvio-padrão	$\sigma = \sqrt{\sigma^2}$	$S = \sqrt{(S^2)}$
Proporção	$p = \frac{X_1 + X_2 + \dots + X_N}{N}$	$\hat{p} = \frac{X_1 + X_2 + \dots + X_n}{n}$



PARÂMETROS E ESTATÍSTICAS

EXEMPLO 1 NO R: PARÂMETROS E ESTATÍSTICAS

UTP
UNIVERSIDADE
QUE TRANSFORMA



DEPARTAMENTO ACADÊMICO DE ESTATÍSTICA

UTP
UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ

MINISTÉRIO DA
EDUCAÇÃO



EXERCÍCIOS NO R:

- 1 Simular 5000 pesos de uma população com distribuição Normal com parâmetros $\mu = 90$ e $\sigma^2 = 9$;
- 2 Construir um histograma para esta variável, considerando as densidades de frequência, e adicionar a curva teórica da Normal;
- 3 Selecionar uma amostra, com reposição de tamanho $n = 100$ desta população;
- 4 Calcular as estatísticas: média, variância e desvio-padrão da amostra.

EXERCÍCIOS NO R:

- 1 Simular 5000 pesos de uma população com distribuição Normal com parâmetros $\mu = 90$ e $\sigma^2 = 9$;
- 2 Construir um histograma para esta variável, considerando as densidades de frequência, e adicionar a curva teórica da Normal;
- 3 Selecionar uma amostra, com reposição de tamanho $n = 100$ desta população;
- 4 Calcular as estatísticas: média, variância e desvio-padrão da amostra.

EXERCÍCIOS NO R:

- 1 Simular 5000 pesos de uma população com distribuição Normal com parâmetros $\mu = 90$ e $\sigma^2 = 9$;
- 2 Construir um histograma para esta variável, considerando as densidades de frequência, e adicionar a curva teórica da Normal;
- 3 Selecionar uma amostra, com reposição de tamanho $n = 100$ desta população;
- Calcular as estatísticas: média, variância e desvio-padrão da amostra.

EXERCÍCIOS NO R:

- 1 Simular 5000 pesos de uma população com distribuição Normal com parâmetros $\mu = 90$ e $\sigma^2 = 9$;
- 2 Construir um histograma para esta variável, considerando as densidades de frequência, e adicionar a curva teórica da Normal;
- 3 Selecionar uma amostra, com reposição de tamanho $n = 100$ desta população;
- 4 Calcular as estatísticas: média, variância e desvio-padrão da amostra.

ESTIMADOR:

São estatísticas utilizadas para estimar os parâmetros. Aos valores observados dos estimadores dá-se o nome de **estimativas**.

TABLE: Parâmetros e Estimadores.

Parâmetros	Estimadores	Propriedades
μ	$\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n}$	Não viciado e consistente
σ^2	$S^2 = \sum_{i=1}^n \frac{(X_i - \bar{X})^2}{n - 1}$	Não viciado e consistente
p	$\hat{p} = \frac{X_1 + X_2 + \dots + X_n}{n}$	Não viciado e consistente



DISTRIBUIÇÃO AMOSTRAL

População

$$X \sim N(\mu, \sigma^2)$$

UTP
UNIVERSIDADE
QUE TRANSFORMA

DAEST
DEPARTAMENTO ACADÉMICO DE ESTATÍSTICA

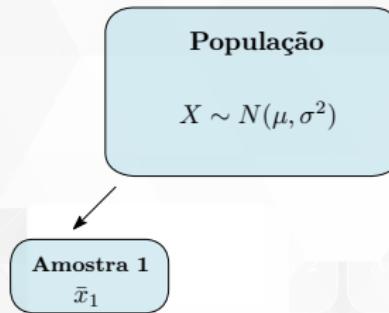
UTP
UNIVERSIDADE TECNOLÓGICA FEDERAL DO PIAUÍ

MINISTÉRIO DA
EDUCAÇÃO

 PÁTRIA AMADA
BRASIL
GOVERNO FEDERAL



DISTRIBUIÇÃO AMOSTRAL



UTP
UNIVERSIDADE
QUE TRANSFORMA

DAEST
DEPARTAMENTO ACADÉMICO DE ESTATÍSTICA

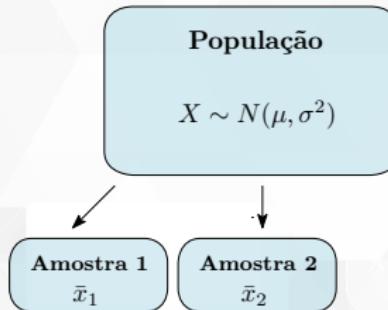
UTFPR
UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ

MINISTÉRIO DA
EDUCAÇÃO

PÁTRIA AMADA
BRASIL
GOVERNO FEDERAL



DISTRIBUIÇÃO AMOSTRAL



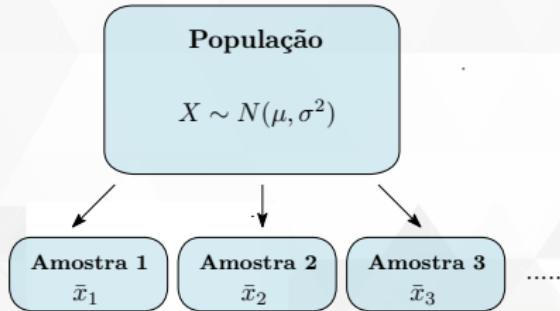
UTP
UNIVERSIDADE
QUE TRANSFORMA

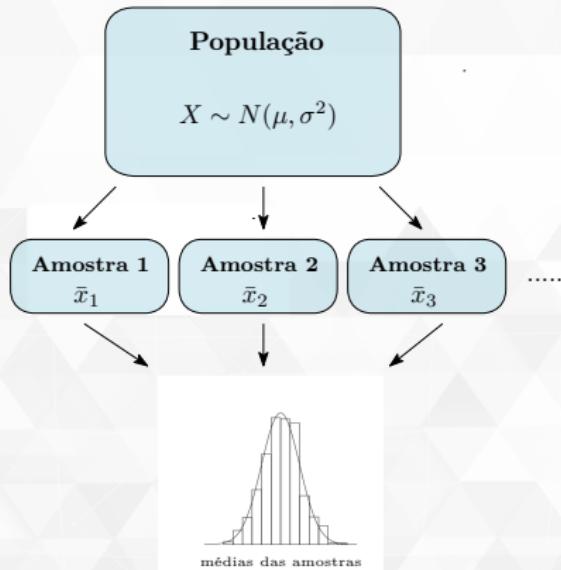
DAEST
DEPARTAMENTO ACADÉMICO DE ESTATÍSTICA

UTFPR
UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ

MINISTÉRIO DA
EDUCAÇÃO

PÁTRIA AMADA
BRASIL
GOVERNO FEDERAL







DISTRIBUIÇÃO AMOSTRAL

DISTRIBUIÇÃO AMOSTRAL DE UM ESTIMADOR:

É o comportamento probabilístico do estimador, isto é, sua distribuição de probabilidade, caso todas as possíveis amostras fossem retiradas.

UTP
UNIVERSIDADE
QUE TRANSFORMA



DAEST
DEPARTAMENTO ACADÊMICO DE ESTATÍSTICA



UTFPR
UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ

MINISTÉRIO DA
EDUCAÇÃO



DISTRIBUIÇÃO AMOSTRAL DA MÉDIA

EXEMPLO 2 NO R: DISTRIBUIÇÃO AMOSTRAL DA MÉDIA: POPULAÇÃO NORMAL.

UTP
UNIVERSIDADE
QUE TRANSFORMA



DAEST
DEPARTAMENTO ACADÊMICO DE ESTATÍSTICA

UTP
UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARAÍBA

MINISTÉRIO DA
EDUCAÇÃO



DISTRIBUIÇÃO AMOSTRAL DA MÉDIA: POPULAÇÃO NORMAL

População

$$X \sim N(\mu, \sigma^2)$$

UTFPR UNIVERSIDADE
QUE TRANSFORMA



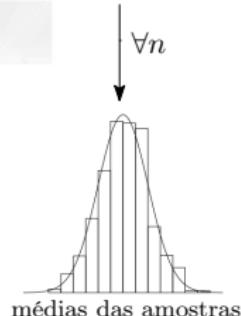
MINISTÉRIO DA
EDUCAÇÃO



DISTRIBUIÇÃO AMOSTRAL DA MÉDIA: POPULAÇÃO NORMAL

População

$$X \sim N(\mu, \sigma^2)$$



UTFPR UNIVERSIDADE
QUE TRANSFORMA

DAEST
DEPARTAMENTO ACADÉMICO DE ESTATÍSTICA

UTP
UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARAÍBA

MINISTÉRIO DA
EDUCAÇÃO

PÁTRIA AMADA
BRASIL
GOVERNO FEDERAL

DISTRIBUIÇÃO AMOSTRAL DA MÉDIA: POPULAÇÃO NORMAL

População

$$X \sim N(\mu, \sigma^2)$$

$\forall n$



$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right), \forall n$$

UTFPR UNIVERSIDADE
QUE TRANSFORMA

DAEST
DEPARTAMENTO ACADÉMICO DE ESTATÍSTICA

UTFPR
UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARAÍBA

MINISTÉRIO DA
EDUCAÇÃO

PÁTRIA AMADA
BRASIL
GOVERNO FEDERAL

DISTRIBUIÇÃO AMOSTRAL DA MÉDIA: POPULAÇÃO NÃO NORMAL

EXEMPLO 3 NO R: DISTRIBUIÇÃO AMOSTRAL DA MÉDIA: POPULAÇÃO NÃO NORMAL.

UTP
UNIVERSIDADE
QUE TRANSFORMA



DAEST
DEPARTAMENTO ACADÊMICO DE ESTATÍSTICA



UTFPR
UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ

MINISTÉRIO DA
EDUCAÇÃO



DISTRIBUIÇÃO AMOSTRAL DA MÉDIA: POPULAÇÃO NÃO NORMAL

População

$$X \sim f(x; \theta) \quad \text{ou}$$
$$X \sim P(X = x; \theta)$$

UTP
UNIVERSIDADE
QUE TRANSFORMA



DAEST
DEPARTAMENTO ACADÊMICO DE ESTATÍSTICA



UTFPR
UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ

MINISTÉRIO DA
EDUCAÇÃO



DISTRIBUIÇÃO AMOSTRAL DA MÉDIA: POPULAÇÃO NÃO NORMAL

População

$$X \sim f(x; \theta) \quad \text{ou}$$

$$X \sim P(X = x; \theta)$$

$$\begin{cases} E(X) \\ Var(X) \end{cases}$$

$$Z = \frac{\bar{X} - E(\bar{X})}{\sqrt{Var(\bar{X})}}$$

UTFPR UNIVERSIDADE
QUE TRANSFORMA



MINISTÉRIO DA
EDUCAÇÃO



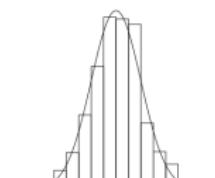
DISTRIBUIÇÃO AMOSTRAL DA MÉDIA: POPULAÇÃO NÃO NORMAL

População

$$X \sim f(x; \theta) \quad \text{ou} \\ X \sim P(X = x; \theta)$$

$$\begin{cases} E(X) \\ Var(X) \end{cases}$$

$n \rightarrow \infty$



médias das amostras

UTFPR
UNIVERSIDADE
QUE TRANSFORMA

DAEST
DEPARTAMENTO ACADÉMICO DE ESTATÍSTICA

UTP
UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ

MINISTÉRIO DA
EDUCAÇÃO

PÁTRIA AMADA
BRASIL
GOVERNO FEDERAL

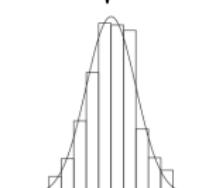
DISTRIBUIÇÃO AMOSTRAL DA MÉDIA: POPULAÇÃO NÃO NORMAL

População

$$X \sim f(x; \theta) \quad \text{ou} \\ X \sim P(X = x; \theta)$$

$$\begin{cases} E(X) \\ Var(X) \end{cases}$$

$n \rightarrow \infty$



médias das amostras

$$\bar{X} \sim N \left(E(X), \frac{Var(X)}{n} \right), n \rightarrow \infty$$

UTFPR UNIVERSIDADE
QUE TRANSFORMA

DAEST
DEPARTAMENTO ACADÉMICO DE ESTATÍSTICA

UTP
UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ

MINISTÉRIO DA
EDUCAÇÃO

PÁTRIA AMADA
BRASIL
GOVERNO FEDERAL



DISTRIBUIÇÃO AMOSTRAL DA PROPORÇÃO

EXEMPLO 4 NO R: DISTRIBUIÇÃO AMOSTRAL DA PROPORÇÃO

UTP
UNIVERSIDADE
QUE TRANSFORMA



MINISTÉRIO DA
EDUCAÇÃO





DISTRIBUIÇÃO AMOSTRAL DA PROPORÇÃO

População

$$X \sim Bernoulli(p)$$

UTP
UNIVERSIDADE
QUE TRANSFORMA

DAEST
DEPARTAMENTO ACADÉMICO DE ESTATÍSTICA

UTP
UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ

MINISTÉRIO DA
EDUCAÇÃO

PÁTRIA AMADA
BRASIL
GOVERNO FEDERAL

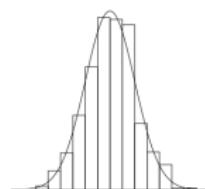


DISTRIBUIÇÃO AMOSTRAL DA PROPORÇÃO

População

$$X \sim Bernoulli(p)$$

$n \rightarrow \infty$



proporções das amostras

UTP
UNIVERSIDADE
QUE TRANSFORMA

DAEST
DEPARTAMENTO ACADÉMICO DE ESTATÍSTICA

UTFPR
UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ

MINISTÉRIO DA
EDUCAÇÃO

PÁTRIA AMADA
BRASIL
GOVERNO FEDERAL

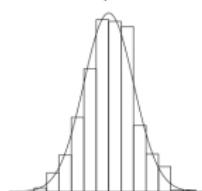


DISTRIBUIÇÃO AMOSTRAL DA PROPORÇÃO

População

$$X \sim Bernoulli(p)$$

$$n \rightarrow \infty$$



proporções das amostras

$$\hat{p} \sim N \left(p, \frac{p(1-p)}{n} \right), n \rightarrow \infty$$

UTP
UNIVERSIDADE
QUE TRANSFORMA

DAEST
DEPARTAMENTO ACADÉMICO DE ESTATÍSTICA

UTP
UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ

MINISTÉRIO DA
EDUCAÇÃO

PÁTRIA AMADA
BRASIL
GOVERNO FEDERAL

EXEMPLO 5 NO R:

- 1 Os salários de funcionários de uma empresa seguem distribuição Normal, com média 1500 e desvio padrão 250. Calcule a probabilidade de que os salários médios das amostras de tamanhos 10 estejam entre 1400 e 1600.



DISTRIBUIÇÕES AMOSTRAIS

$$X \sim N(1500, 250^2)$$

$$\frac{1400 - 1500}{250} = -\frac{100}{250} = -\frac{2}{5}$$

$$P(1400 \leq X \leq 1600)$$

UTP
UNIVERSIDADE
QUE TRANSFORMA

DAEST
DEPARTAMENTO ACADÉMICO DE ESTATÍSTICA

UTP
UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ

MINISTÉRIO DA
EDUCAÇÃO

PÁTRIA AMADA
BRASIL
GOVERNO FEDERAL

$$X \sim N(1500, 250^2)$$



$$\bar{X} \sim N\left(1500, \frac{250^2}{10}\right)$$

$$P(1400 \leq \bar{X} \leq 1600)$$

$$X \sim N(1500, 250^2)$$

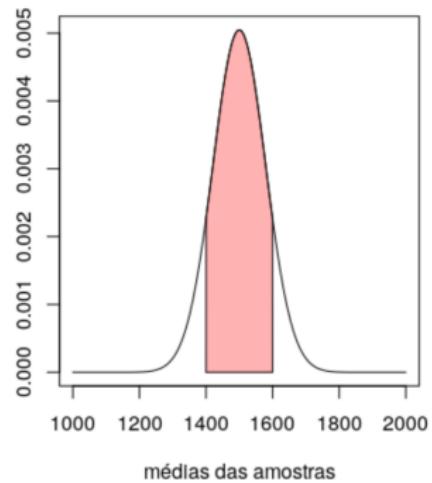


$$\bar{X} \sim N\left(1500, \frac{250^2}{10}\right)$$



$$P(1400 \leq \bar{X} \leq 1600)$$

DISTRIBUIÇÕES AMOSTRAIS



UTP
UNIVERSIDADE
QUE TRANSFORMA



MINISTÉRIO DA
EDUCAÇÃO



DISTRIBUIÇÕES AMOSTRAIS

EXEMPLO 6 NO R:

- 1 Suponha que estamos interessados em estimar a proporção de consumidores de certo produto. Considerando a seleção de uma amostra aleatória de 300 pessoas, calcule a probabilidade de que a proporção amostral de consumidores seja de, no mínimo, 35%, sabendo que a verdadeira proporção é de 40%.

UTFPR UNIVERSIDADE
QUE TRANSFORMA



MINISTÉRIO DA
EDUCAÇÃO





DISTRIBUIÇÕES AMOSTRAIS

$$X \sim Bernoulli(0.4)$$

$$P(p \geq 0.35)$$

UTP
UNIVERSIDADE
QUE TRANSFORMA



MINISTÉRIO DA
EDUCAÇÃO



$$X \sim Bernoulli(0.4)$$



$$\hat{p} \sim N\left(0.4, \frac{0.4(1-0.4)}{300}\right)$$

$$P(\hat{p} \geq 0.35)$$

$$X \sim Bernoulli(0.4)$$



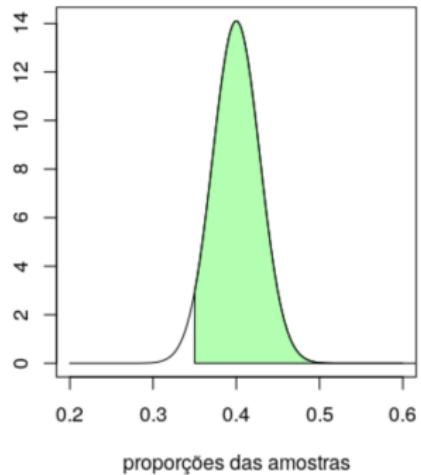
$$\hat{p} \sim N\left(0.4, \frac{0.4(1-0.4)}{300}\right)$$



$$P(\hat{p} \geq 0.35)$$

DISTRIBUIÇÕES AMOSTRAIS

Normal Distribution



UTP
UNIVERSIDADE
QUE TRANSFORMA

DAEST
DEPARTAMENTO ACADÉMICO DE ESTATÍSTICA

UTP
UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ

MINISTÉRIO DA
EDUCAÇÃO

PÁTRIA AMADA
BRASIL
GOVERNO FEDERAL

EXERCÍCIOS NO R:

- 1 Admite-se que as alturas dos estudantes de uma universidade seguem distribuição Normal com média 172.72 cm e desvio padrão 7.62 cm. Calcule a probabilidade de que a altura média amostral se encontre entre 169.67 cm e 173.48 cm, considerando amostras aleatórias de 25 estudantes. **R: 0.6683.**
- 2 Suponha que 60% da população de uma certa cidade seja a favor da criação de um fundo público para fins de criação de áreas de lazer. Se 150 pessoas forem selecionadas aleatoriamente e entrevistadas, qual a probabilidade de que a proporção amostral de pessoas favoráveis seja menor que 0.52? **R: 0.0227**

Tipos de estimação:

- Estimação pontual;
- Estimação por intervalo.



ESTIMAÇÃO PONTUAL

$$\bar{X} \rightarrow \mu$$

$$\hat{p} \rightarrow p$$

$$S^2 \rightarrow \sigma^2$$

UTFPR UNIVERSIDADE
QUE TRANSFORMA



MINISTÉRIO DA
EDUCAÇÃO



$$X \sim N(\mu, \sigma^2)$$

$$\downarrow \quad \forall n$$

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

$$\downarrow$$

$$Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0, 1)$$

$$X \sim N(\mu, \sigma^2)$$



$$\frac{(n-1)S^2}{\sigma^2} \sim \chi_{n-1}^2$$

ESTIMAÇÃO PONTUAL

$$X \sim Bernoulli(p)$$

$$\downarrow \quad n \rightarrow \infty$$

$$\hat{p} \sim N \left(p, \frac{p(1-p)}{n} \right)$$

$$\downarrow$$

$$Z = \frac{\hat{p}-p}{\sqrt{\frac{p(1-p)}{n}}} \sim N(0, 1)$$

UTP
UNIVERSIDADE
QUE TRANSFORMA

DAEST
DEPARTAMENTO ACADÉMICO DE ESTATÍSTICA

UTFPR
UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ

MINISTÉRIO DA
EDUCAÇÃO

PÁTRIA AMADA
BRASIL
GOVERNO FEDERAL

ESTIMAÇÃO POR INTERVALO

- Considera além da estimação pontual, o erro associado à amostra utilizada no processo inferencial;
- Define limites de confiança para um parâmetro, com uma probabilidade (nível de confiança) conhecida;
- Para estimadores com distribuição de probabilidade simétrica, o IC para o parâmetro é simétrico em torno do seu estimador.

ESTIMAÇÃO POR INTERVALO

- De modo geral, seja θ o parâmetro de interesse, então o intervalo de confiança para θ com probabilidade $1 - \alpha$ (nível de confiança) é:

$$IC(\theta, (1 - \alpha)) = (Li; Ls)$$

ou seja,

$$P(Li \leq \theta \leq Ls) = 1 - \alpha$$

- $Li \rightarrow$ limite inferior do intervalo;
- $Ls \rightarrow$ limite superior do intervalo;

UTFPR UNIVERSIDADE
QUE TRANSFORMA



MINISTÉRIO DA
EDUCAÇÃO



INTERVALO DE CONFIANÇA PARA A MÉDIA - σ^2 CONHECIDA

- Supondo $X \sim N(\mu, \sigma^2) \rightarrow \bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$, com σ^2 conhecido.
- E sabendo que $Z = \frac{(\bar{X} - \mu)}{\sigma/\sqrt{n}}$, o IC para μ pode ser obtido por:

$$P\left(-z_{\frac{\alpha}{2}} \leq Z \leq z_{\frac{\alpha}{2}}\right) = 1 - \alpha$$

$$P\left(-z_{\frac{\alpha}{2}} \leq \frac{(\bar{X} - \mu)}{\sigma/\sqrt{n}} \leq z_{\frac{\alpha}{2}}\right) = 1 - \alpha$$

$$P\left(\bar{X} - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha$$

UTFPR UNIVERSIDADE
QUE TRANSFORMA



MINISTÉRIO DA
EDUCAÇÃO



INTERVALO DE CONFIANÇA PARA A MÉDIA - σ^2 CONHECIDA

$$IC(\mu, (1 - \alpha)) = (\bar{X} - e; \bar{X} + e)$$

$$\text{onde } e = z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$$

Intervalo de confiança para a média μ - σ^2 conhecida

Obs: O resultado acima também é válido mesmo que a população não seja Normal, mas a amostra seja suficientemente grande (TLC). O valor $z_{\alpha/2}$ é tal que $P(Z \geq z_{\alpha/2}) = \alpha/2$.

UTFPR
UNIVERSIDADE
QUE TRANSFORMA

DAEST
DEPARTAMENTO ACADÉMICO DE ESTATÍSTICA

UTP
UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ

MINISTÉRIO DA
EDUCAÇÃO

PÁTRIA AMADA
BRASIL
GOVERNO FEDERAL

UI

INTERVALO DE CONFIANÇA PARA A MÉDIA - σ^2 CONHECIDA

- Da relação anterior, pode-se encontrar o tamanho de amostra necessário, fazendo-se

$$n = \left(\frac{z_{\frac{\alpha}{2}} \sigma}{e} \right)^2$$

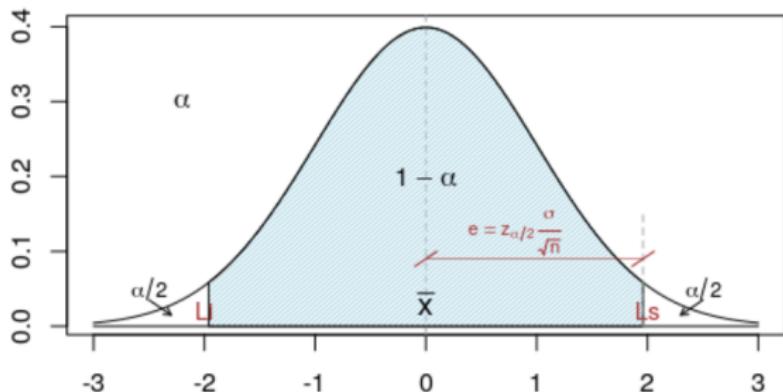
UTFPR UNIVERSIDADE
QUE TRANSFORMA



MINISTÉRIO DA
EDUCAÇÃO



UII INTERVALO DE CONFIANÇA PARA A MÉDIA - σ^2 CONHECIDA



UTFPR UNIVERSIDADE
QUE TRANSFORMA

DAEST
DEPARTAMENTO ACADÉMICO DE ESTATÍSTICA

UTFPR
UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARAÍBA

MINISTÉRIO DA
EDUCAÇÃO

PÁTRIA AMADA
BRASIL
GOVERNO FEDERAL

EXEMPLO 7 NO R:

- 1 Uma empresa deseja estudar seus layouts na Web. Assim, coletaram quanto tempo as pessoas passam em cada uma das página A ou B. Construir um intervalo para o tempo médio de permanencia na página A, com 95% e 99% de confiança, sabendo que os tempos médios de permanência na página A seguem distribuição Normal com desvio-padrão 1.
- 2 No item anterior, qual o tamanho de amostra necessário para se obter um erro de 0.17 com 99% de confiança?

EXERCÍCIOS NO R:

- 1 Construir um IC para o tempo médio de permanencia na página B, com 95% e 99% de confiança, sabendo que os tempos médios de permanencia na página B seguem distribuição Normal com desvio-padrão populacional de 0.5. Qual o tamanho de amostra necessário para se obter um erro de 0.15 com 99% de confiança?

R: $IC(\mu, 0.95) = (1.8134, 2.1712)$, $IC(\mu, 0.99) = (1.7571, 2.2274)$, $n = 74$.

- 2 Considera-se que o tempo de permanência de recém-formados no 1º emprego, em anos, é Normal com variância populacional de 1.96. Para uma amostra de 15 recém-formados, a média obtida foi de 2.7 anos. Construa um IC com 90% de confiança para o tempo médio populacional de permanência de recém-formados no 1º emprego.

R: $IC(\mu, 0.90) = (2.105421, 3.294579)$.



INTERVALO DE CONFIANÇA PARA A MÉDIA - σ^2 DESCONHECIDA

- Supondo $X \sim N(\mu, \sigma^2)$, com σ^2 desconhecido, devemos estimá-lo pelo S , o que introduz uma incerteza maior na estimação. Assim,

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

UTFPR UNIVERSIDADE
QUE TRANSFORMA



MINISTÉRIO DA
EDUCAÇÃO



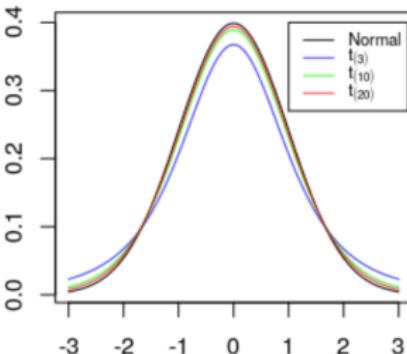
Distribuição t-student com parâmetro ν :

$$E(t) = 0$$

$$\text{Var}(t) = \frac{\nu}{\nu - 2}$$

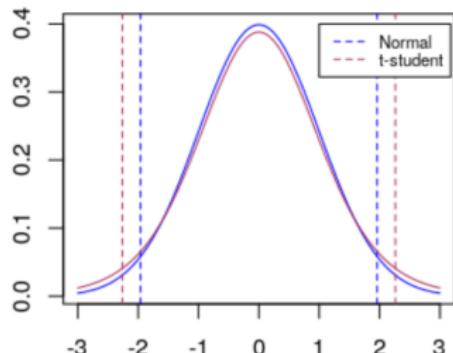
Obs: Quando $\nu \rightarrow \infty$ então

$$t \rightarrow \text{Normal}$$



INTERVALO DE CONFIANÇA PARA A MÉDIA - σ^2 DESCONHECIDA

No R: `dt(x, df)` calcula a densidade da distribuição t com parâmetro df.



Assim, quando σ^2 desconhecida, o IC para μ pode ser obtido por:

$$P(-t_{((n-1), \frac{\alpha}{2})} \leq t \leq t_{((n-1), \frac{\alpha}{2})}) = 1 - \alpha$$

$$P(-t_{((n-1), \frac{\alpha}{2})} \leq \frac{(\bar{X} - \mu)}{S/\sqrt{n}} \leq t_{((n-1), \frac{\alpha}{2})}) = 1 - \alpha$$

$$P(\bar{X} - t_{((n-1), \frac{\alpha}{2})} \frac{S}{\sqrt{n}} \leq \mu \leq \bar{X} + t_{((n-1), \frac{\alpha}{2})} \frac{S}{\sqrt{n}}) = 1 - \alpha$$



INTERVALO DE CONFIANÇA PARA A MÉDIA - σ^2 DESCONHECIDA

$$IC(\mu, (1 - \alpha)) = (\bar{X} - e; \bar{X} + e)$$

onde $e = t_{((n-1), \frac{\alpha}{2})} \frac{S}{\sqrt{n}}$

Intervalo de confiança para a média μ - σ^2 desconhecida

UTFPR
UNIVERSIDADE
QUE TRANSFORMA

DAEST
DEPARTAMENTO ACADÉMICO DE ESTATÍSTICA

UTFPR
UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ

MINISTÉRIO DA
EDUCAÇÃO

PÁTRIA AMADA
BRASIL
GOVERNO FEDERAL

EXEMPLO 8 NO R:

O conjunto de dados income800.csv (disponível na [página do DAEST](#)), consiste de dados de renda anual de candidatos a empréstimos pessoais de uma empresa de empréstimo. Construa um IC com 95% de confiança para o verdadeiro rendimento anual médio.

EXERCÍCIOS NO R:

- 1 Das válvulas fabricadas por uma companhia retira-se uma amostra de 400 válvulas, e obtém-se a vida média de 800 horas e o desvio-padrão de 100 horas. Encontre o intervalo com 99% de confiança para a vida média das válvulas da população. R: $IC(\mu, 0.99) = (787.059, 812.941)$.
- 2 A resistência de determinado material está em estudo e sabe-se que essa variável é normalmente distribuída. Considerando a amostra abaixo, determine o IC para a resistencia média populacional com 90% de confiança. R: $IC(\mu, 0.90) = (5.5021, 6.9422)$.

4.9 7.0 8.1 4.5 5.6 6.8 7.2 5.7 6.2

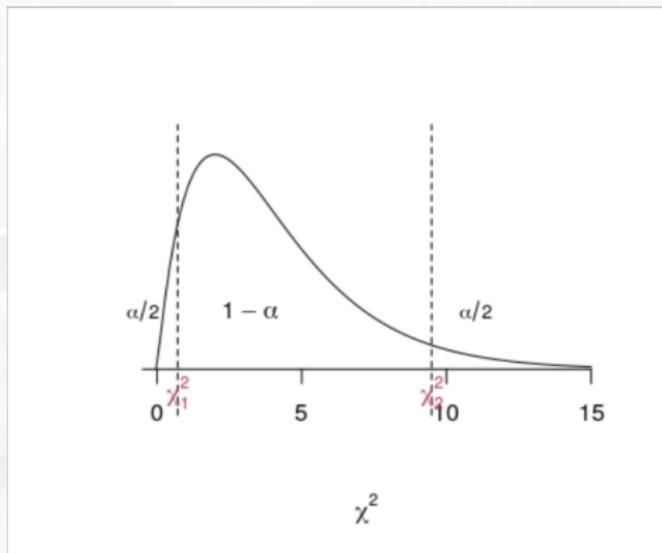
INTERVALO DE CONFIANÇA PARA A VARIÂNCIA

- Supondo $X \sim N(\mu, \sigma^2)$, com média e variância desconhecidos, o IC para a variância populacional é dado por:

$$IC(\sigma^2, (1 - \alpha)) = \left(\frac{(n - 1)S^2}{\chi_2^2}, \frac{(n - 1)S^2}{\chi_1^2} \right)$$

onde, os valores χ_2^2 e χ_1^2 são tais que

$$P(\chi \leq \chi_1^2) = P(\chi \geq \chi_2^2) = \alpha/2$$



UII INTERVALO DE CONFIANÇA PARA A VARIÂNCIA

EXEMPLO 9 NO R:

Sabe-se que as vendas diárias de um produto, em reais, seguem distribuição Normal. Uma amostra de dias de vendas, apresentou os valores:

275 258 242 229 225 134 227 249 259 191

Construa um intervalo com 90% de confiança para a variância populacional.

UTFPR
UNIVERSIDADE
QUE TRANSFORMA



DAEST
DEPARTAMENTO ACADÊMICO DE ESTATÍSTICA

UTP
UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARAÍBA

MINISTÉRIO DA
EDUCAÇÃO



INTERVALO DE CONFIANÇA PARA A PROPORÇÃO

- Sabe-se que se $n \rightarrow \infty$, então $\hat{p} \sim N\left(p, \frac{p(1-p)}{n}\right)$
- Assim, o IC para a proporção fica:

$$IC(p, (1 - \alpha)) = (\hat{p} - e; \hat{p} + e)$$

onde, $e = z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}}$

UTFPR UNIVERSIDADE
QUE TRANSFORMA



MINISTÉRIO DA
EDUCAÇÃO



INTERVALO DE CONFIANÇA PARA A PROPORÇÃO

EXEMPLO 10 NO R:

Considere o conjunto de dados "Melanoma" disponível no pacote "MASS" do R, que consiste de dados amostrais de 205 pacientes com melanoma maligno na Dinamarca. Considerando a variável sexo, encontre o IC para a proporção de pacientes do sexo feminino com 90% de confiança.

UTP
UNIVERSIDADE
QUE TRANSFORMA



MINISTÉRIO DA
EDUCAÇÃO



INTERVALO DE CONFIANÇA PARA A PROPORÇÃO

EXERCÍCIOS NO R:

- 1 Considerando ainda o conjunto de dados "Melanoma" do pacote "MASS", encontre o IC para a proporção de pacientes do sexo masculino com 99% de confiança.

R: $IC(p, 0.99) = (0.2978, 0.4729)$.

- 2 Considerando a variável "ulcer" do conjunto de dados "Melanoma" do pacote "MASS", encontre o IC para a proporção de pacientes com presença de úlcera com 95% de confiança.

R: $IC(p, 0.99) = (0.3710, 0.5069)$.