Universidade Federal de São Paulo

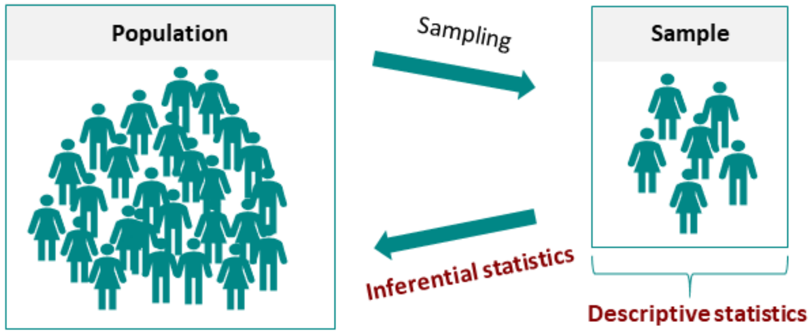
Inferência estatística, distribuição amostral, Teorema do Limite Central, Estimador e Métodos de obtenção de estimadores

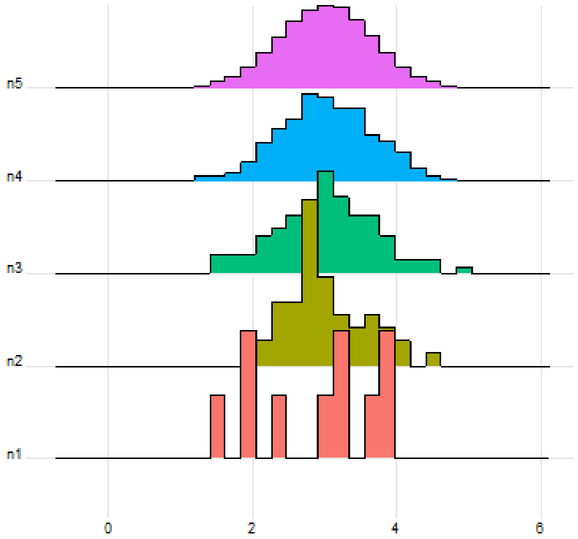
Professor

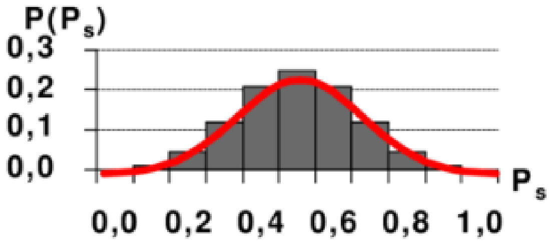
Julio Cezar

AULA DE HOJE

- Inferências a partir de dados amostrais;
- Distribuição amostral;
- Teorema do Limite Central (TLC);
- Estimador;
- Métodos de obtenção de estimadores.





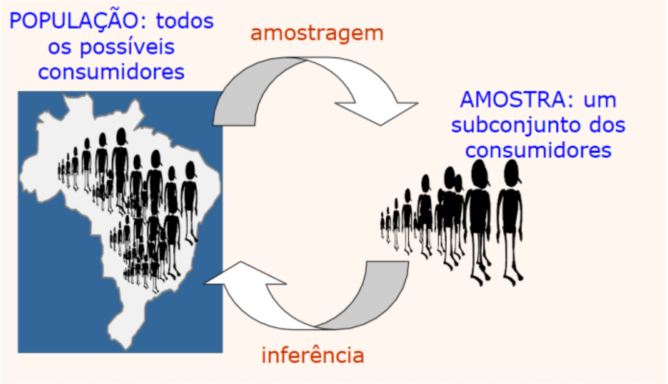


UNIFESP
UNIVERSIDADE FEDERAL DE SÃO PAULO
1933

INFERÊNCIA ESTATÍSTICA

Inferência Estatística

A inferência estatística consiste em fazer afirmações probabilísticas sobre as características do modelo probabilístico, que se supõe representar uma população, a partir dos dados de uma **amostra aleatória (probabilística)** desta mesma população, ou seja, é o processo de se tirar conclusões ou tomar decisões acerca da população com base em uma amostra dessa população.



UNIFESP
UNIVERSIDADE FEDERAL DE SÃO PAULO
1933

INFERÊNCIA ESTATÍSTICA

A inferência pode ser dividida em:

- **Estimação de Parâmetros**: consiste em avaliar uma medida populacional, a partir da informação amostral.

- **Testes de Hipóteses**: consiste em julgar hipóteses sobre populações usando os conhecimentos amostrais.

Conceitos importantes

Na aula anterior vimos o que vem a ser:

- **População**;
- **Censo**;
- **Amostra**.

Também foi mencionado que, à medida que o tamanho de uma amostra aumenta, as informações relativas à amostra vão se tornando cada vez mais verdadeiras.

UNIFESP
Universidade Federal de São Paulo
1933

PARÂMETROS

Parâmetro: As quantidades da população, em geral desconhecidas, sobre as quais temos interesse, são denominadas parâmetros e são representadas por letras gregas θ, μ e σ entre outras.

UNIFESP
UNIVERSIDADE FEDERAL DE SÃO PAULO
1933

PARÂMETROS

Parâmetro: As quantidades da população, em geral desconhecidas, sobre as quais temos interesse, são denominadas parâmetros e são representadas por letras gregas θ, μ e σ entre outras.

Exemplos de Parâmetros:

Parâmetros

Parâmetro: As quantidades da população, em geral desconhecidas, sobre as quais temos interesse, são denominadas parâmetros e são representadas por letras gregas θ, μ e σ entre outras.

Exemplos de Parâmetros:

- μ: a média da população.

UNIFESP
Universidade Federal de São Paulo
1933

Parâmetros

Parâmetro: As quantidades da população, em geral desconhecidas, sobre as quais temos interesse, são denominadas parâmetros e são representadas por letras gregas θ, μ e σ entre outras.

Exemplos de Parâmetros:

- μ: a média da população.
- σ^2: a variância da população.

Parâmetros

Parâmetro: As quantidades da população, em geral desconhecidas, sobre as quais temos interesse, são denominadas parâmetros e são representadas por letras gregas θ, μ e σ entre outras.

Exemplos de Parâmetros:

- μ: a média da população.
- σ^2: a variância da população.
- p: a proporção da população.

UNIFESP
UNIVERSIDADE FEDERAL DE SÃO PAULO
1933

Estatísticas

Obtida uma **amostra aleatória simples**, é possível calcular diversas características desta amostra, como, por exemplo: a média, a mediana, a variância, etc. Qualquer uma destas características é uma função de X_1, X_2,..., X_n e, portanto, o seu valor depende da amostra sorteada. Sendo assim, cada uma dessas características ou funções é também uma variável aleatória (v.a.) . Por exemplo, a média amostral é a v.a. definida por

$$\overline{X} = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

Estatísticas

Estatística: é uma característica da amostra $X_1, X_2, \ldots, X_n$, ou seja, uma estatística T é uma função de $X_1, X_2, \ldots, X_n$ e portanto é uma variável aleatória.

Exemplos de Estatísticas:

- $T(X_1, X_2, \ldots, X_n) = \overline{X} = \sum_{i=1}^{n} \frac{X_i}{n}$: a média da amostra.
- $T(X_1, X_2, \ldots, X_n) = S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X})^2$: a variância da amostra.
- $T(X_1, X_2, \ldots, X_n) = X_{(1)} = \mathrm{Min}(X_1, X_2, \ldots, X_n)$: o menor valor da amostra.

UNIFESP
Universidade Federal de São Paulo
1933

Parâmetros e Estatísticas

	Parâmetro (População)	Estatísticas (Amostras)	Valores Observados
Média	μ	$\bar{X}$	$\bar{x}$
Variância	σ^2	S^2	s^2
Proporção	p	$\hat{p}$	$\hat{p}$

UNIFESP
Universidade Federal de São Paulo
1933

Parâmetros e Estatísticas

População

N

$x_1, x_2, \dots, x_N$

Média populacional

$$\mu = \frac{\sum_{i=1}^{N} x_i}{N}$$

Variância populacional

$$\sigma^2 = \frac{\sum_{i=1}^{N} (x_i - \mu)^2}{N}$$

Desvio padrão populacional

$$\sigma$$

Proporção populacional

$$\rho$$

Parâmetros

$\mu, \sigma^2, \sigma, \rho$

Inferência Estatística

A média amostral é um estimador pontual da média populacional

$$\hat{\mu} = \bar{X}$$

A variância amostral é um estimador pontual da variância populacional

$$\hat{\sigma}^2 = S$$

O desvio padrão amostral é um estimador pontual do desvio padrão populacional

$$\hat{\sigma} = S$$

Amostra

n

$x_1, x_2, \dots, x_n$

Média amostral

$$\hat{\mu} = \bar{X} = \frac{\sum_{i=1}^{n} x_i}{n}$$

Variância amostral

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

Desvio padrão amostral

$$\hat{\sigma} = S$$

Proporção amostral

$$\hat{\rho} = p$$

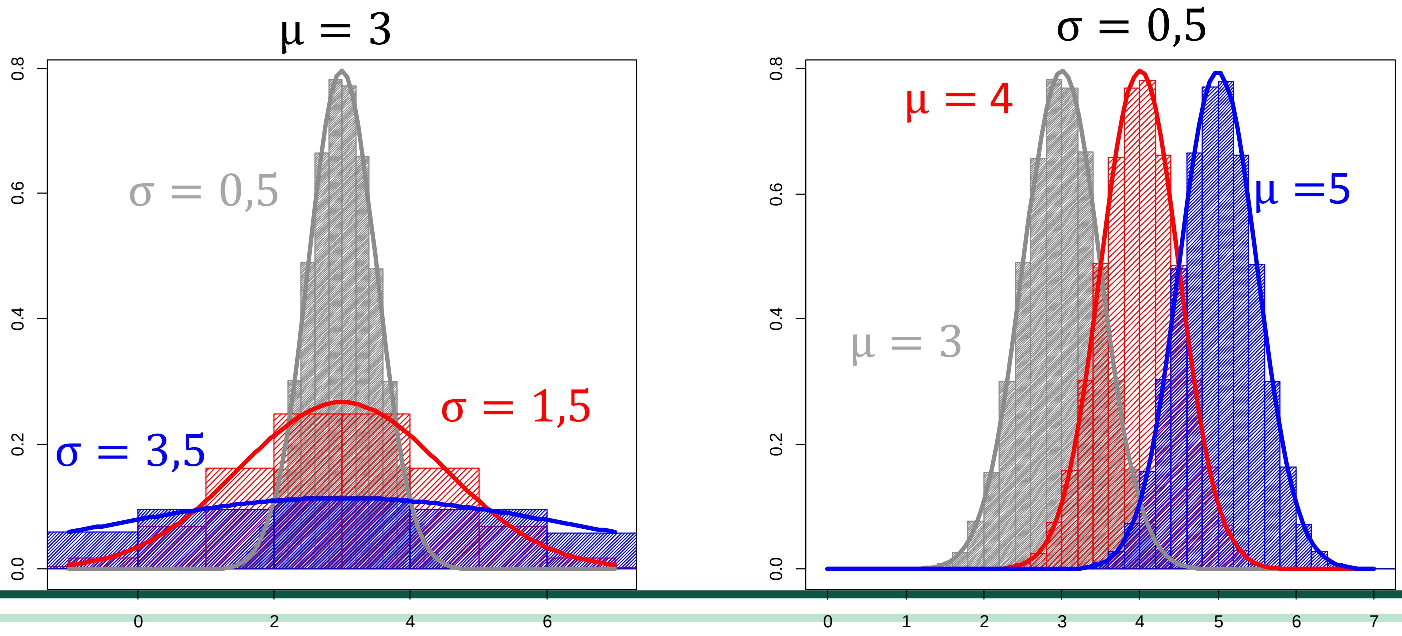
Estimadores

$\hat{\mu} = \bar{X}, \hat{\sigma}^2 = S^2, \hat{\sigma} = S, \hat{\rho} = p$

Distribuição normal

Distribuição Normal: Seja $X_1, \ldots, X_n$ uma amostra de uma variável aleatória X com distribuição normal com média μ é variância σ^2. A função densidade de probabilidade de X é dada por (para $x, \mu \in \mathcal{R}$ e $\sigma^2>0$)

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\left\{-\frac{1}{2\sigma^2}(x-\mu)^2\right\}}$$



Distribuição Amostral

Seja uma população qualquer com um parâmetro θ de interesse, correspondendo a uma estatística T (T é uma v.a., uma vez que depende da amostra sorteada; amostras diferentes fornecerão diferentes valores para T) em uma amostra. Amostras aleatórias são retiradas da população; e, para cada amostra, calcula-se o valor t (não confunda com o t da distribuição t-Student) da estatística T. Os valores de t formam uma nova população que segue uma distribuição de probabilidades, que é chamada de **distribuição amostral de T**.

Distribuição Amostral

Esquematicamente, temos:

1. uma população X, com determinado parâmetro de interesse **θ**;
2. todas as k amostras possíveis são retiradas da população, de acordo com certo procedimento amostra aleatória simples (AAS);
3. para cada amostra, obtém-se o valor X_k da estatística T;
4. os valores t_k formam uma nova população, cuja distribuição recebe o nome de distribuição amostral de T (também possui média, desvio padrão, etc.).

UNIFESP
Universidade Federal de São Paulo
1933

DISTRIBUIÇÃO AMOSTRAL

Definição: Distribuição amostral de uma estatística é a função de probabilidade (ou densidade) que descreve o comportamento probabilístico da estatística em amostras repetidas da mesma população.

UNIFESP
UNIVERSIDADE FEDERAL DE SÃO PAULO
1933

Distribuição Amostral

Definição: Distribuição amostral de uma estatística é a função de probabilidade (ou densidade) que descreve o comportamento probabilístico da estatística em amostras repetidas da mesma população.

Exemplo: A distribuição de probabilidades de $\overline{x}$ é chamada de distribuição amostral da média.

UNIFESP
UNIVERSIDADE FEDERAL DE SÃO PAULO
1933

Distribuição Amostral

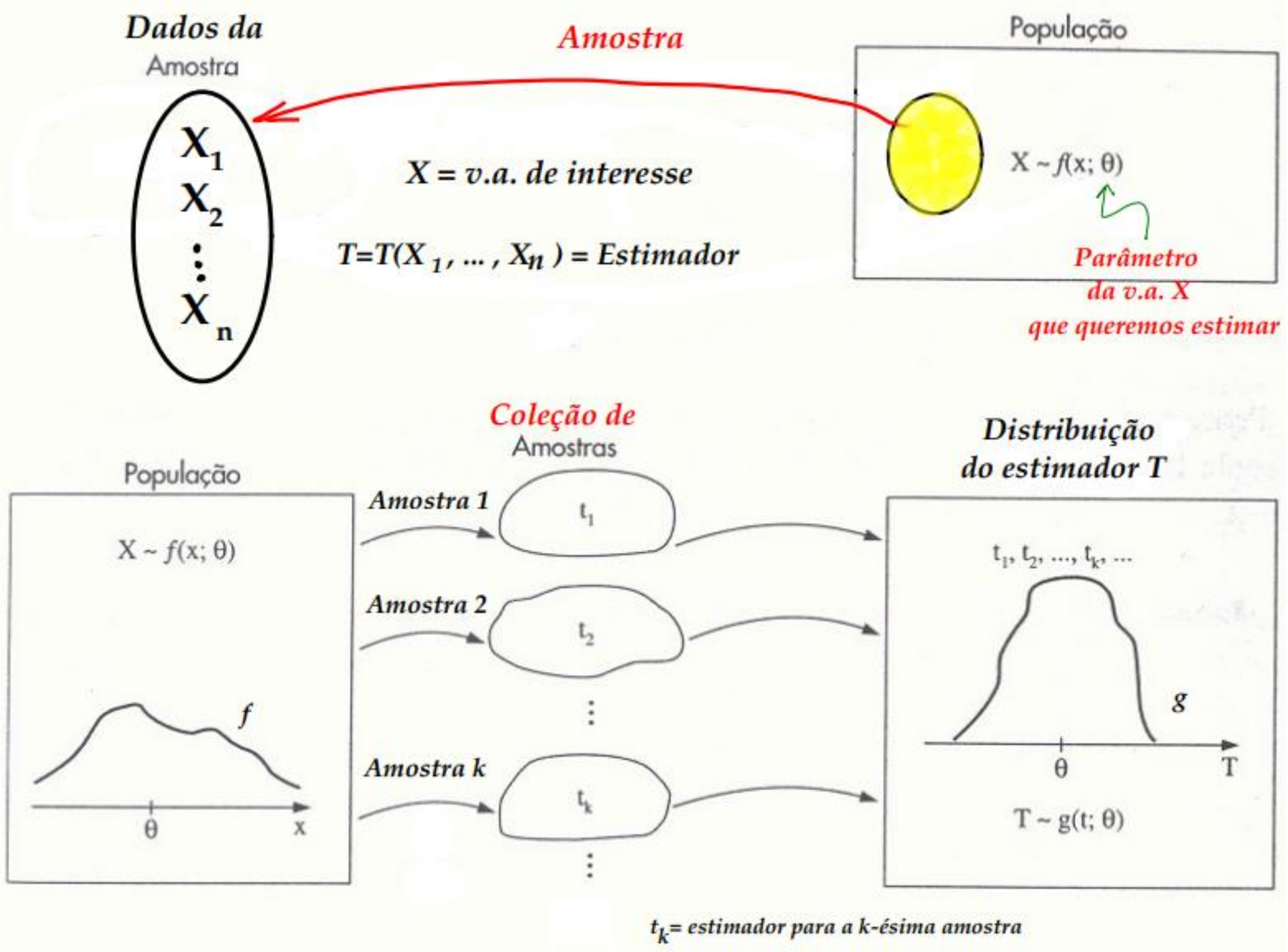
Definição: Distribuição amostral de uma estatística é a função de probabilidade (ou densidade) que descreve o comportamento probabilístico da estatística em amostras repetidas da mesma população.

Exemplo: A distribuição de probabilidades de $\overline{x}$ é chamada de distribuição amostral da média.

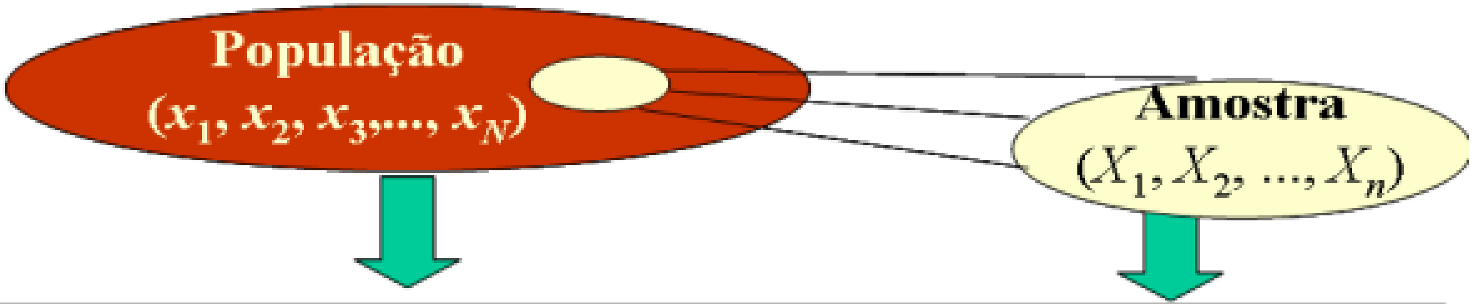
A distribuição amostral de uma estatística depende da distribuição da população, do tamanho da amostra e do método de seleção da amostra.

UNIFESP
UNIVERSIDADE FEDERAL DE SÃO PAULO
1933

Distribuição Amostral



Distribuição Amostral



	Parâmetros	Estatísticas
Proporção	$p = \dfrac{n^\circ \text{ de elementos com o atributo}}{N}$	$\hat{p} = \dfrac{n^\circ \text{ de elementos com o atributo}}{n}$
Média	$\mu = \dfrac{1}{N}\sum_{i=1}^{N} x_i$	$\bar{X} = \dfrac{1}{n}\sum_{i=1}^{n} X_i$
Variância	$\sigma^2 = \dfrac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2$	$S^2 = \dfrac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$

Distribuição Amostral

- A distribuição das **médias** ou **proporções** amostrais representam a população de todas as **médias** ou **proporções** oriundas de uma amostra de tamanho n de uma variável aleatória.

- A convergência em forma de distribuição e dos parâmetros dessa distribuição são elucidadas pela **Lei dos Grandes Números** e pelo **Teorema do Limite Central (TLC)**.

Distribuição Amostral da Média

Um dos procedimentos estatísticos mais comuns é o uso de uma média da amostra $\bar{x}$ para fazer inferências sobre a média da população μ.

Distribuição Amostral da Média

Um dos procedimentos estatísticos mais comuns é o uso de uma média da amostra $\bar{x}$ para fazer inferências sobre a média da população μ.

Definição: A distribuição amostral de $\bar{x}$ é a distribuição de probabilidade de todos os valores possíveis da média da amostra, $\bar{x}$.

Distribuição Amostral da Média

Considere a determinação da distribuição amostral da média $\bar{x}$ da amostra. Suponha que uma amostra aleatória de tamanho n seja retirada de uma população normal com média μ e variância σ^2. Então, temos que

$$\bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n}$$

UNIFESP
Universidade Federal de São Paulo
1933

Distribuição Amostral da Média

Considere a determinação da distribuição amostral da média $\bar{x}$ da amostra. Suponha que uma amostra aleatória de tamanho n seja retirada de uma população normal com média μ e variância σ^2. Então, temos que

$$\bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n}$$

tem uma distribuição normal com média

$$\mu_X = \frac{\mu + \mu + \cdots + \mu}{n} = \mu$$

UNIFESP
UNIVERSIDADE FEDERAL DE SÃO PAULO
1933

Distribuição Amostral da Média

e se uma população for infinita e amostragem for aleatória, ou se a população for finita e a amostragem for com reposição, então a variância da distribuição amostral das médias será dada por

$$\sigma_{X}^{2} = \frac{\sigma^2 + \sigma^2 + \cdots + \sigma^2}{n} = \frac{\sigma^2}{n}$$

Distribuição Amostral da Média

e se uma população for infinita e amostragem for aleatória, ou se a população for finita e a amostragem for com reposição, então a variância da distribuição amostral das médias será dada por

$$\sigma_{\bar{X}}^2 = \frac{\sigma^2 + \sigma^2 + \cdots + \sigma^2}{n} = \frac{\sigma^2}{n}$$

e se a população for de tamanho N, se a amostragem for sem reposição, então a variância será dada com a inclusão do chamado **fator de correção da população finita**:

$$\sigma_{\bar{X}}^2 = \frac{\sigma^2}{n} \cdot \left(\frac{N - n}{N - 1}\right).$$

UNIFESP

Universidade Federal de São Paulo

1933

Distribuição Amostral da Média

Exemplo: Considere uma população em que a variável X pode assumir um dos valores do conjunto {1, 3, 5, 5, 7}. A distribuição de probabilidade de X é

x	1	3	5	7
$P(X = x)$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{2}{5}$	$\frac{1}{5}$

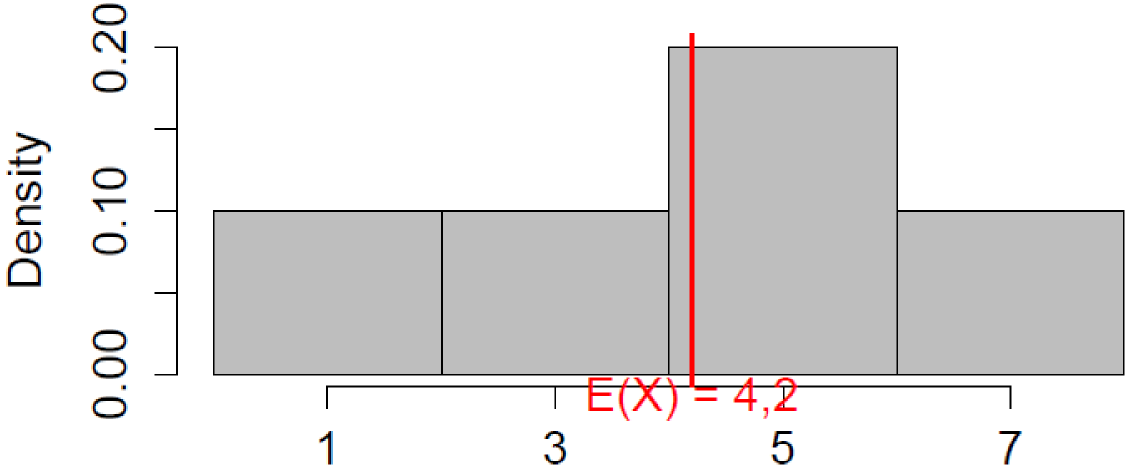
Esperança e Variância:

$$E(X) = \mu_X = 4{,}2$$

$$Var(X) = \sigma_X^2 = 4{,}16$$

UNIFESP
UNIVERSIDADE FEDERAL DE SÃO PAULO
1933

Distribuição Amostral da Média



$n = 1, \text{Var}[X] = 4{,}16$

UNIFESP
Universidade Federal de São Paulo
1933

Distribuição Amostral da Média

Continuação do exemplo: Selecionar todas as amostras aleatórias simples de tamanho 2, $n = 2$, selecionadas ao acaso e com reposição da população X, e encontrar a distribuição do estimador pontual $\overline{X}$, ou seja, encontrar a distribuição da média amostral

$$\overline{X} = \frac{X_1 + X_2}{2}$$

em que

- X_1 é o valor na primeira seleção.
- X_2 é o valor na segunda seleção.

UNIFESP
Universidade Federal de São Paulo
1933

Distribuição Amostral da Média

Amostra (X_1, X_2)	Probabilidade	Média amostral
(1,1)	1/25	1
(1,3)	1/25	2
(1,5)	2/25	3
(1,7)	1/25	4
(3,1)	1/25	2
(3,3)	1/25	3
(3,5)	2/25	4
(3,7)	1/25	5
(5,1)	2/25	3
(5,3)	2/25	4
(5,5)	4/25	5
(5,7)	2/25	6
(7,1)	1/25	4
(7,3)	1/25	5
(7,5)	2/25	6
(7,7)	1/25	7

DISTRIBUIÇÃO AMOSTRAL DA MÉDIA

Distribuição de $\overline{X}$ para $n = 2$

$\bar{x}$	1	2	3	4	5	6	7
$P(\overline{X} = \bar{x})$	$\frac{1}{25}$	$\frac{2}{25}$	$\frac{5}{25}$	$\frac{6}{25}$	$\frac{6}{25}$	$\frac{4}{25}$	$\frac{1}{25}$

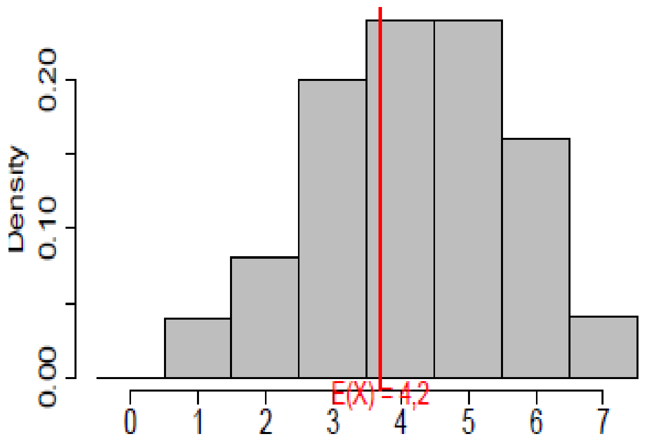
Esperança e Variância de $\overline{X}$ para $n = 2$

$$E[\overline{X}] = \mu_X = 4{,}2$$

$$Var[\overline{X}] = 2{,}08 = \frac{\sigma_X^2}{2}$$

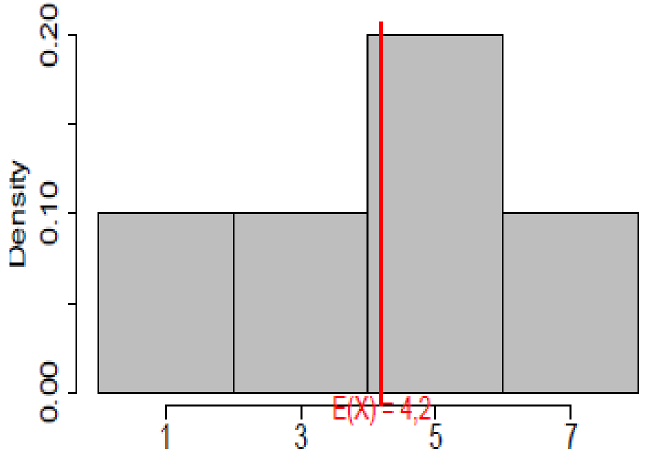
UNIFESP
UNIVERSIDADE FEDERAL DE SÃO PAULO
1933

Distribuição Amostral da Média

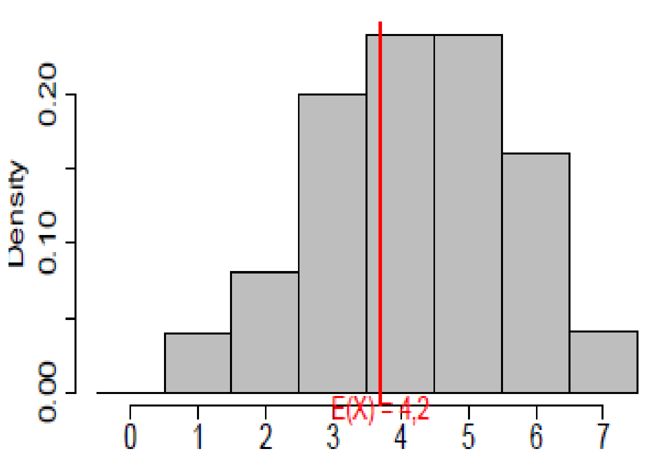


$n = 2, \mathrm{Var}[\overline{\mathrm{X}}] = 2{,}08$

DISTRIBUIÇÃO AMOSTRAL DA MÉDIA



$n = 1, \mathrm{Var}[X] = 4{,}16$



$n = 2, \mathrm{Var}[\overline{X}] = 2{,}08$

UNIFESP
UNIVERSIDADE FEDERAL DE SÃO PAULO
1933

DISTRIBUIÇÃO AMOSTRAL DA MÉDIA

Distribuição de $\overline{\mathrm{X}}$ para $\mathrm{n} = 3$

Observação: Olhar a Tabela 10.6 Distribuição amostral de algumas estatísticas obtidas de amostra de tamanho n = 3 do livro Morettin e Bussab, (2017).

$\bar{x}$	$P(\bar{X} = \bar{x})$
1	1/125
5/3	3/125
7/3	9/125
3	16/125
11/3	24/125
13/3	27/125
5	23/125
17/3	15/125
19/3	6/125
1	1/125

Esperança e Variância de $\overline{\mathrm{X}}$ para $\mathrm{n} = 3$

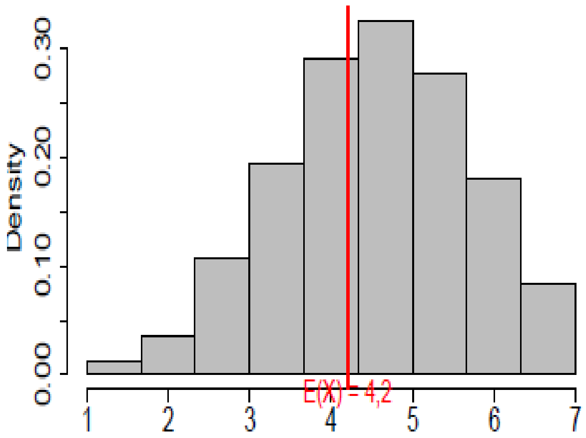
$$\mathrm{E}[\overline{\mathrm{X}}] = \mu_{\mathrm{X}} = 4{,}2 \text{ e } \mathrm{Var}[\overline{\mathrm{X}}] = 1{,}39 = \frac{\sigma_{\mathrm{X}}^{2}}{3}$$

UNIFESP

UNIVERSIDADE FEDERAL DE SÃO PAULO

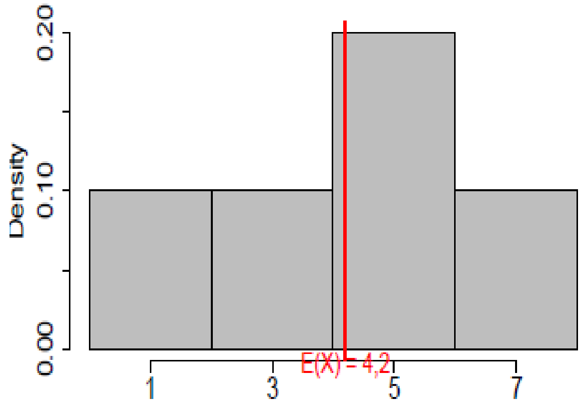
1933

DISTRIBUIÇÃO AMOSTRAL DA MÉDIA

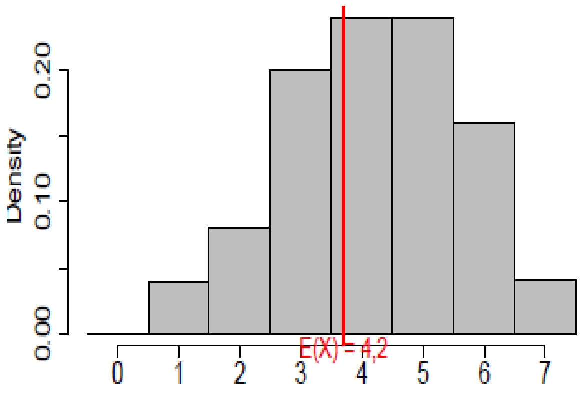


$n = 3, Var[\bar{X}] = 1{,}39$

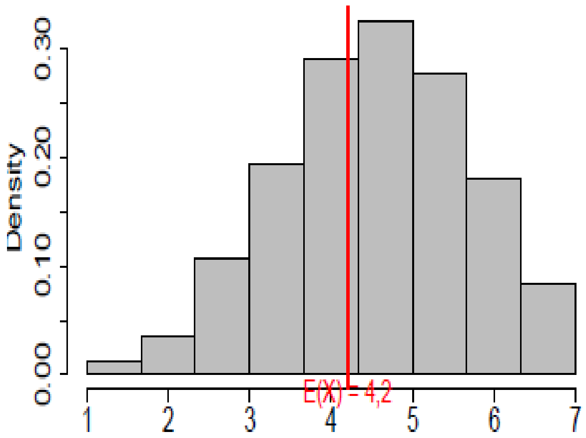
Distribuição Amostral da Média



$n = 1, \mathrm{Var}[X] = 4{,}16$



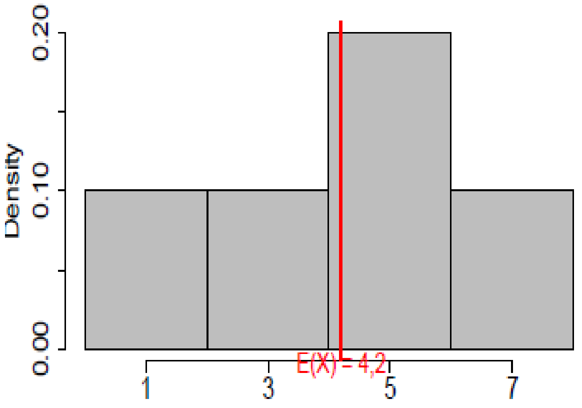
$n = 2, \mathrm{Var}[\overline{X}] = 2{,}08$



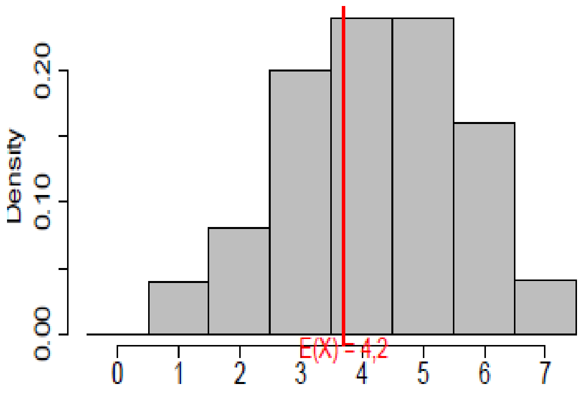
$n = 3, \mathrm{Var}[\overline{X}] = 1{,}39$

UNIFESP
Universidade Federal de São Paulo
1933

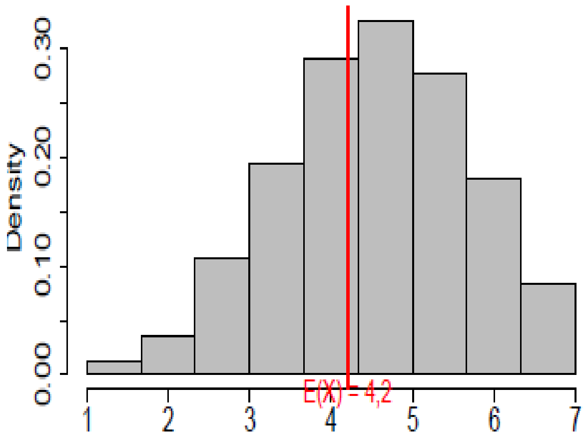
Distribuição Amostral da Média



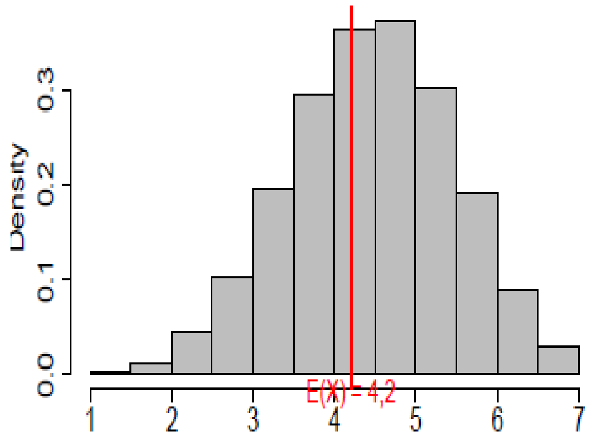
$n = 1, Var[X] = 4{,}16$



$n = 2, Var[\bar{X}] = 2{,}08$



$n = 3, Var[\bar{X}] = 1{,}39$



$n = 4, Var[\bar{X}] = 1{,}04 = \frac{\sigma_X^2}{4}$

UNIFESP
Universidade Federal de São Paulo
1933

DISCUSSÃO DO EXEMPLO

Análise dos Histogramas

- Conforme o tamanho da amostra aumenta, $n \to \infty$, os valores de $\bar{X}$ tendem a concentrar-se cada vez mais em torno de $E[\bar{X}] = 4{,}2$.
- A **variância diminui** a medida que o tamanho da **amostra aumenta**.
- Para n suficientemente grande, a forma da distribuição (histograma) aproxima-se de uma **distribuição simétrica (normal)**.

Distribuição Amostral da Média

Observação: O valor de $\sigma_{\bar{X}}$ (o desvio padrão de $\bar{x}$) é útil para determinar a distância a que a média da amostra pode estar da média populacional. Devido ao papel que $\sigma_{\bar{X}}$ desempenha em calcular possíveis erros de amostragem, $\sigma_{\bar{X}}$ é denominado **erro-padrão da média**.

UNIFESP
Universidade Federal de São Paulo
1933

Distribuição Amostral da Média

Observação: O valor de $\sigma_{\bar{X}}$ (o desvio padrão de $\bar{x}$) é útil para determinar a distância a que a média da amostra pode estar da média populacional. Devido ao papel que $\sigma_{\bar{X}}$ desempenha em calcular possíveis erros de amostragem, $\sigma_{\bar{X}}$ é denominado **erro-padrão da média**.

Se estivermos amostrando de uma população que tenha uma distribuição desconhecida de probabilidades, a distribuição amostral da média da amostra será aproximadamente normal, com média μ e variância σ^2/n, se o tamanho n da amostra for grande. Esse é um dos mais úteis teoremas em Estatística, o chamado **Teorema do Limite Central**.

UNIFESP
Universidade Federal de São Paulo
1933

TEOREMA DO LIMITE CENTRAL

Teorema do Limite Central (TLC)

Para amostras aleatórias $X_1, \ldots, X_n$ iid (se cada variável aleatória tiver a mesma distribuição de probabilidade das outras e todas forem mutuamente independentes), retiradas de uma população com média μ é variância σ^2 finita, a distribuição amostral da média $\overline{X}$ aproxima-se, para n grande $(n \to \infty)$, de uma distribuição normal com média μ é variância ${}^{\sigma^2}/_{n}$,

$$\frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \xrightarrow[n \to \infty]{\text{em distribuição}} Z \sim N(0,1)$$

UNIFESP
UNIVERSIDADE FEDERAL DE SÃO PAULO
1933

TEOREMA DO LIMITE CENTRAL

- A variável aleatória $e = \overline{X} - \mu$ é denominado **erro amostral da média**.
- O desvio padrão $\sigma/\sqrt{n}$ é denominado **erro padrão da média**.

UNIFESP
UNIVERSIDADE FEDERAL DE SÃO PAULO
1933

TEOREMA DO LIMITE CENTRAL

Observação:

- Quanto maior for o tamanho n da amostra, mais a média amostral se aproximará da média da população.
- As propriedades da distribuição amostral asseguram que a média de uma amostra é uma boa estatística para inferir sobre a média da população μ da qual foi extraída.

TEOREMA DO LIMITE CENTRAL

Observação:

- Ao mesmo tempo, o teorema do limite central estabelece que se o tamanho da amostra *n* for suficientemente grande a distribuição da média amostral será normal, qualquer que seja a forma da distribuição da população.

- Portanto, o teorema do limite central permite aplicar a distribuição normal para obter respostas da média de uma amostra de tamanho suficientemente grande retirada de uma população qualquer.

Teorema do limite central

Exemplo 1: Sabe-se que o faturamento diário de um posto de gasolina segue uma distribuição de média 20 mil e desvio padrão 2 mil reais. Qual é a probabilidade de que num período de 60 dias, o faturamento total ultrapasse R$ 1.230.000,00?

UNIFESP
Universidade Federal de São Paulo
1933

Teorema do limite central

Exemplo 2: Uma fábrica de pneus alega que a vida média dos pneus é 30.000 Km, com desvio padrão de 2.000 Km. Tomando-se como verdadeiros estes dados, qual é a probabilidade de uma amostra com 40 pneus apresentar vida média menor que 29.500 Km?

Distribuição Amostral da Proporção

Vamos considerar uma população em que a proporção de elementos portadores de certa característica é p. Logo, podemos definir uma v.a. X, da seguinte maneira:

$$\mathrm{X} = \begin{cases} 1, & \text{se o indivíduo for portador da característica} \\ 0, & \text{se o indivíduo } \textbf{não} \text{ for portador da característica,} \end{cases}$$

tem-se que, $\mathrm{X} \sim \mathrm{Ber}(\mathrm{p})$, logo $\mu = \mathrm{E}[\mathrm{X}] = \mathrm{p}$ e $\sigma^2 = \mathrm{V}[\mathrm{X}] = \mathrm{p}(1 - \mathrm{p})$.

UNIFESP
Universidade Federal de São Paulo
1933

Distribuição Amostral da Proporção

Retirada uma AAS dessa população, e indicando por Y_n o total de indivíduos portadores da característica na amostra, já vimos que

$$Y_n \sim \mathrm{Bin}(n, p)$$

Vamos definir por $\hat{p}$ a proporção de indivíduos portadores da característica na amostra, isto é,

$$\hat{p} = \frac{Y_n}{n}.$$

UNIFESP
Universidade Federal de São Paulo
1933

Distribuição Amostral da Proporção

Então,

$$P(Y_n = k) = P\left(\frac{Y_n}{n} = \frac{k}{n}\right)' = P\left(\hat{p} = \frac{k}{n}\right),$$

ou seja, a distribuição amostral de $\hat{p}$ é obtida da distribuição de Y_n.

Vimos que a distribuição binomial pode ser aproximada pela distribuição normal. Vamos mostrar que a justificativa desse fato está no TLC.

UNIFESP
Universidade Federal de São Paulo
1933

Distribuição Amostral da Proporção

Inicialmente, observe que

$$Y_n = X_1 + X_2 + \cdots + X_n,$$

em que cada X_i tem distribuição de Bernoulli, com média $\mu = p$ e variância $\sigma^2 = p(1-p)$, e são duas a duas independentes. Podemos escrever que

$$Y_n = n\bar{X},$$

mas pelo TLC, $\bar{X}$ terá distribuição aproximadamente normal, com média p e variância $\frac{p(1-p)}{n}$, ou seja,

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

UNIFESP
UNIVERSIDADE FEDERAL DE SÃO PAULO
1933

Distribuição Amostral da Proporção

Logo, a transformação $Y_n = n\overline{X}$ terá a distribuição

$$Y_n \sim N\big(np, np(1-p)\big),$$

Observe que $\overline{X}$, na expressão no slide anterior, é a própria variável $\hat{p}$ e, desse modo, para n grande podemos considerar a distribuição amostral de p como aproximadamente normal:

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

UNIFESP

Universidade Federal de São Paulo

1933

Distribuição Amostral da Proporção

Exemplo 1: Visitantes estrangeiros nos Estados Unidos são cuidadosamente monitorados por organizações de segurança e pela indústria do turismo. Durante o mês de março de 2008 foram monitorados 4,7 milhões de visitantes internacionais nos EUA. Quarenta por cento (40%) de todos os visitantes da Europa Ocidental eram do Reino Unido. Suponha que sejam selecionados aleatoriamente 120 visitantes do mês de março, vindos da Europa Ocidental, e que se determine o número dos visitantes que vieram do Reino Unido.

i) Qual a distribuição da proporção amostral de visitantes do Reino Unido, $\hat{p}$?

ii) Qual a probabilidade de que a proporção amostral seja maior do que 0,50?

UNIFESP

Universidade Federal de São Paulo

1933

Distribuição Amostral da Proporção

Exemplo 2: Suponha que a proporção de peças fora da especificação em um lote é de 40%. Tomada uma amostra de tamanho 30, qual será a probabilidade desta amostra fornecer uma proporção de peças defeituosas menor que 0,50?

ESTIMAÇÃO

Devido a natureza aleatória envolvida num procedimento amostral (AAS), não podemos garantir que repetições de amostras produzam sempre resultados idênticos.

Ao coletarmos uma amostra, não podemos prever antecipadamente seu resultado, pois todas as quantidades associadas à amostra terão caráter aleatório e, portanto, devem receber tratamento probabilístico.

ESTIMAÇÃO

- Mas qualquer estimador (função T) representa bem o parâmetro em estudo?
- Os critérios para escolha do “melhor” estimador para um determinado parâmetro populacional são definidos a partir de “propriedades” desejáveis destes estimadores. As propriedades mais consideradas são:

Estimador Pontual

Estimador Pontual: Um Estimador Pontual é uma estatística construída com a finalidade de estimar um parâmetro de interesse na população. Às vezes denotamos como $\hat{\mu}$, $\hat{\theta}$, $\hat{\sigma}^2$, etc.

Exemplos de Estimador Pontual:

- $\hat{\mu} = \overline{X}$: a média amostral é um estimador da média da população μ.
- $\hat{\sigma}^2 = S^2$: a variância amostral é um estimador da variância populacional σ^2.
- $\hat{p}$: a proporção amostral como estimador da proporção de unidades com a característica de interesse na população p.

UNIFESP
UNIVERSIDADE FEDERAL DE SÃO PAULO
1933

ESTIMADOR

Propriedades dos estimadores

Um estimador é uma v.a. e tem uma distribuição amostral.

Viés: relacionado ao erro **sistemático**.

Precisão: proximidade das estimativas **entre si**. Inversamente relacionada à variância.

Acurácia: proximidade entre as estimativas e **o verdadeiro valor**.

Consistência: proximidade entre as estimativas e o verdadeiro valor quando $\mathbf{n} \rightarrow \infty$.

UNIFESP
UNIVERSIDADE FEDERAL DE SÃO PAULO
1933

ESTIMADOR

Propriedades dos estimadores

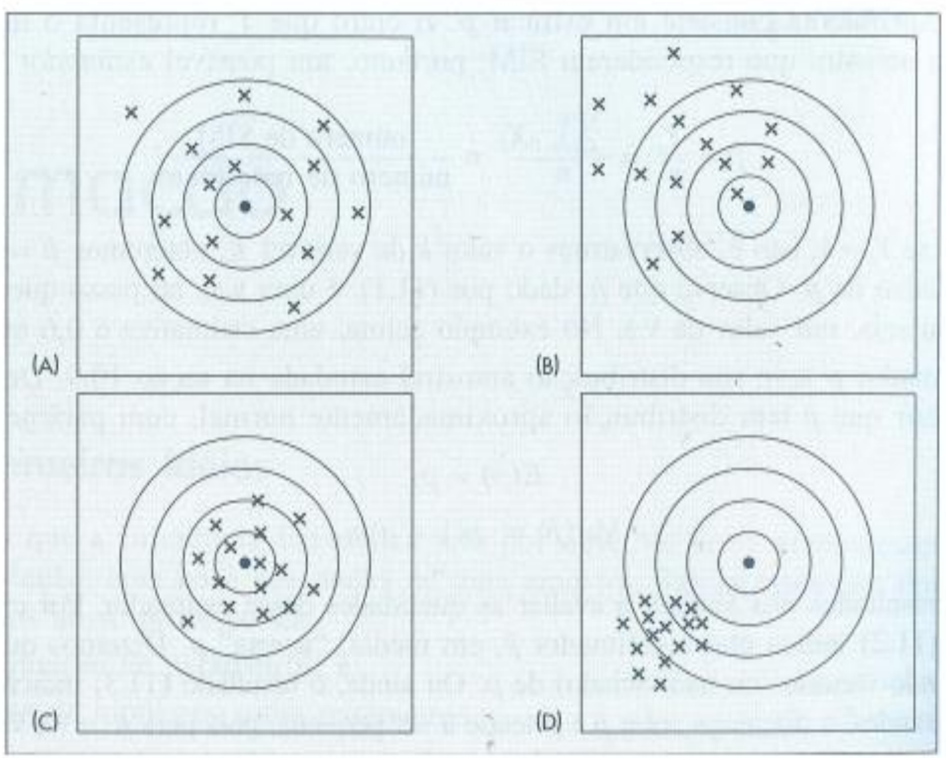
Exemplo: Quatro atiradores

(A) Sem viés, pouco acurado e baixa precisão.

(B) Viesado, pouco acurado e pouca precisão.

(C) Sem viés, acurado e boa precisão.

(D) Viesado, pouco acurado e alta precisão.

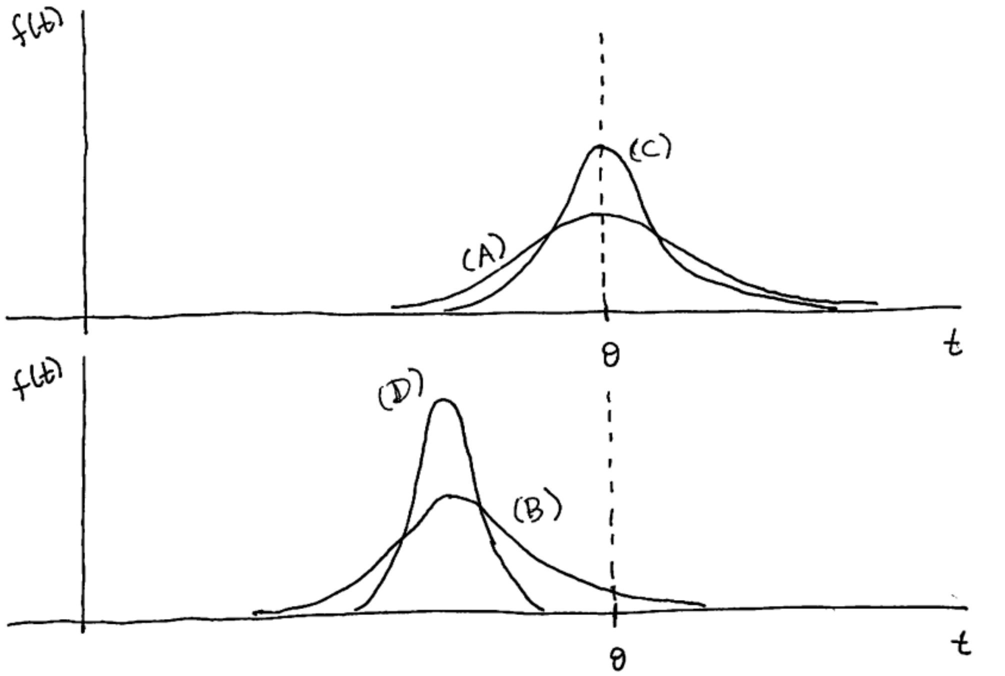


UNIFESP
UNIVERSIDADE FEDERAL DE SÃO PAULO
1933

Propriedades dos estimadores

Exemplo: Quatro atiradores

Representação gráfica de quatro estimadores para θ, todos com distribuição normal e que são semelhantes ao exemplo dos quatro atiradores.



ESTIMADOR

Propriedades dos estimadores

- O **estimador $\hat{\theta}$** é **não viesado** para θ se $E(\hat{\theta}) = \theta$, para todo θ;

Obs: Se o estimador é viesado, a diferença $E(\hat{\theta}) - \theta$ é chamada de **viés** do estimador.

- Se $\hat{\theta}$ é um estimador de θ com base em uma a.a. de tamanho n, dizemos que $\hat{\theta}$ é **consistente** para θ se, para todo $\epsilon > 0$,

$$\lim_{n \to \infty} P\left(\left|\hat{\theta} - \theta\right| \leq \epsilon\right) = 1$$

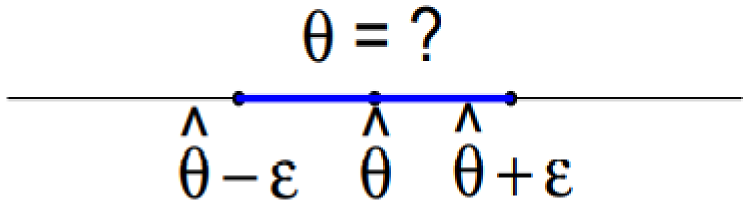
$\hat{\theta} - \theta$ representa o **erro de estimação**.

UNIFESP
UNIVERSIDADE FEDERAL DE SÃO PAULO
1933

ESTIMADOR

Propriedades dos estimadores

Se $\hat{\theta}$ é consistente para θ, então em termos de **probabilidade**, $\hat{\theta}$ pode estar tão **próximo de θ** quanto queiramos (ϵ pequeno), basta tomarmos **n** suficientemente **grande**.



Resultado

$\hat{\theta}$ é um estimador consistente para θ se

$$\lim_{n \to \infty} E(\hat{\theta}) = \theta \quad \text{e} \quad \lim_{n \to \infty} Var(\hat{\theta}) = 0 .$$

ESTIMADOR

Propriedades dos estimadores

$\hat{\theta}_1$ e $\hat{\theta}_2$ são dois estimadores não viesados de θ. Se

$$\mathrm{Var}(\hat{\theta}_1) < \mathrm{Var}(\hat{\theta}_2),$$

então $\hat{\theta}_1$ é **mais eficiente** do que $\hat{\theta}_2$.

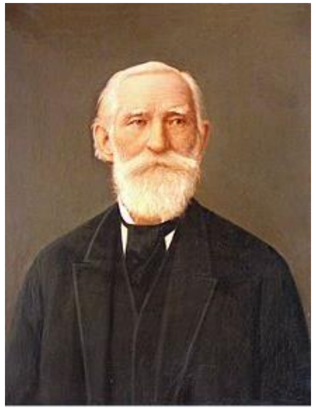
MÉTODOS DE OBTENÇÃO DE ESTIMADORES

Há diversos métodos para obtenção de estimadores, abaixo estão listados apenas dois, sendo:

- **Método dos Momentos (M);**
- **Método da Máxima Verossimilhança (MV).**

MÉTODOS DE OBTENÇÃO DE ESTIMADORES

Método dos momentos



- Método proposto por volta de 1887 por Pafnuty Chebyshev.
- Ideia básica é atribuída a Karl Pearson.
- Método dos momentos: igualar os momentos da população, que são definidos em termos de valores esperados, aos correspondentes momentos da amostra.
- Os momentos da população são funções de parâmetros desconhecidos.
- Solução da(s) equação(ões) são os estimadores dos parâmetros.

Métodos de obtenção de estimadores

Método dos momentos

Seja $(X_1, X_2, \ldots, X_n)$ uma amostra aleatória de uma v.a. X com função densidade f_θ.

Definição: O k-ésimo momento populacional é definido por

$$\mu_k = E(X^k) = \int_{-\infty}^{+\infty} x^k f_\theta(x)dx, \qquad k = 1,2, \ldots$$

Definição: O k-ésimo momento amostral é definido por

$$m_k = \frac{1}{n}\sum_{i=0}^{n} X_i^k, \qquad k = 1,2, \ldots$$

O **estimador de momentos** é a solução da igualdade dos momentos amostrais com os momentos populacionais.

Método dos momentos

Definição: Sendo $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_r)^T$ dizemos que $\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_r$ são estimadores obtidos pelos método dos momentos, se eles forem soluções das equações,

$$m_k = \mu_k, \qquad k = 1, 2, \ldots, r.$$

Métodos de obtenção de estimadores

Estimador de momentos

- Sejam $X_1, X_2, \ldots, X_n$ v.a.'s com fdp ou fp f($\mathbf{x}$; θ) com r parâmetros $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_r)^T$.
- Os estimadores $T_{\theta_1}, \ldots, T_{\theta_r}$ são encontrados igualando os primeiros r momentos populacionais aos primeiros r momentos amostrais.
- Tal procedimento resulta em um conjunto de equações que deve ser resolvido.

$$E(X_1) = m_1$$

$$E(X_2) = m_2$$

$$E(X_3) = m_3$$

UNIFESP
Universidade Federal de São Paulo
1933

MÉTODOS DE OBTENÇÃO DE ESTIMADORES

Recomendações e limitações do método dos momentos

Vantagens

- Concepção simples e intuitiva.
- Fácil de obter (desde que os momentos populacionais estejam disponíveis).
- Em geral, oferece estimadores consistentes.
- Suposições distribucionais não são essenciais.
- Pode ser usado como guia inicial para outros métodos.
- É a base do método dos momentos generalizados.

MÉTODOS DE OBTENÇÃO DE ESTIMADORES

Recomendações e limitações do método dos momentos

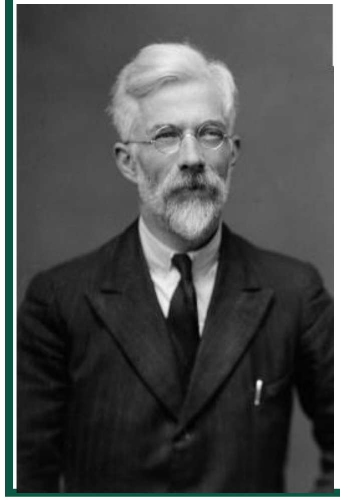
Desvantagens

- Difícil de expressar a incerteza associada às estimativas.
- Difícil de generalizar para modelos e/ou estruturas complexas de dados.
- Em geral, não viés não é garantido.
- Eficiência é difícil de medir e não é garantida mesmo para grandes amostras.
- Pode resultar em estimativas fora do espaço paramétrico.
- Precisa que os momentos populacionais sejam passíveis de calcular.

MÉTODOS DE OBTENÇÃO DE ESTIMADORES

Método da máxima verossimilhança

- Proposto por Ronald Fisher em 1922.
- É o método mais popular em estatística aplicada.
- **Ideia geral:** Encontrar o conjunto de valores para os parâmetros $\boldsymbol{\theta}$ de uma distribuição de probabilidade $f(\mathbf{x}; \theta)$ que maximize a "chance" de observar a amostra de fato observada.



UNIFESP
UNIVERSIDADE FEDERAL DE SÃO PAULO
1933

MÉTODOS DE OBTENÇÃO DE ESTIMADORES

Método da máxima verossimilhança

Sejam dados **x** uma realização de uma v.a. **X** com fd ou fdp $f(\mathbf{x}; \theta)$, a função verossimilhança é

$$L(\theta) = f(\mathbf{x}; \theta)$$

em que $f(\mathbf{x}; \theta)$ é a distribuição conjunta de **X.**

Supondo que as observações são independentes

$$L(\theta) = \prod_{i=1}^{n} f(\mathbf{x}_i; \theta)$$

Notação para enfatizar que a verossimilhança é com o **x** já observado

$$L(\theta|\mathbf{x})$$

UNIFESP

UNIVERSIDADE FEDERAL DE SÃO PAULO

1933

Métodos de obtenção de estimadores

Função de log-verossimilhança e escore

A função de log-verossimilhança é

$$l(\theta; \mathbf{x}) = \ln(L(\theta; \mathbf{x}))$$

No caso de observações independentes, tem-se

$$l(\theta; \mathbf{x}) = \sum_{i=1}^{n} \ln(L(\theta; \mathbf{x_i}))$$

Função escore

Caso de observações independentes

$$U(\theta; \mathbf{x}) = \frac{\partial l(\theta; \mathbf{x})}{\partial \theta_j}, \quad j \in \{1,2, \ldots, p\},$$

UNIFESP
Universidade Federal de São Paulo
1933

MÉTODOS DE OBTENÇÃO DE ESTIMADORES

Estimativa e estimador de máxima verossimilhança

Estimativa de máxima verossimilhança: O valor

$$\hat{\theta} = \hat{\theta}(\mathbf{x})$$

é a estimativa de máxima verossimilhança para θ se $L(\hat{\theta}) \geq L(\theta), \forall\, \theta$.

Estimador de máxima verossimilhança: Se $\hat{\theta}(\mathbf{x})$ é a estimativa de máxima verossimilhança, então

$$\hat{\theta}(\mathbf{X})$$

é o estimador de máxima verossimilhança (EMV).

UNIFESP
UNIVERSIDADE FEDERAL DE SÃO PAULO
1933

MÉTODOS DE OBTENÇÃO DE ESTIMADORES

Exercício: Encontre o estimador de máxima verossimilhança (E.M.V.) de λ do modelo Poisson.

MÉTODOS DE OBTENÇÃO DE ESTIMADORES

Recomendações e limitações do método da máxima verossimilhança

Vantagens

- Concepção intuitiva.
- Propriedades assintóticas desejáveis: não-viés e eficiência.
- Estimadores consistentes.
- Metodologia completa para estimação e inferência (IC e TH).
- É o método de estimação mais popular em estatística.

UNIFESP
UNIVERSIDADE FEDERAL DE SÃO PAULO
1933

Recomendações e limitações do método da máxima verossimilhança

Desvantagens

- Pode ser difícil de obter em termos práticos.
- De forma geral, requer métodos numéricos.
- Suposição explícita de uma distribuição de probabilidade.

MÉTODOS DE OBTENÇÃO DE ESTIMADORES

Comentário finais sobre estimação estatística

- Estimação de parâmetros emprega álgebra, cálculo e métodos numéricos.
- No entanto, os métodos são conceitualmente fáceis de compreender.
  - Momentos: igualar momentos e resolver.
  - Máxima verossimilhança: maximizar a chance de observar a amostra.
- Os estimadores já foram determinados para os principais parâmetros e distribuições .

Métodos de obtenção de estimadores

Comentário finais sobre estimação estatística

- Existem ainda **outros** métodos de estimação.
  - Método de mínimos quadrados.
  - Método da Inferência Bayesiana.
  - Métodos de estimação robustos.
  - Equações de estimação generalizadas.
  - Etc.
- Implementados em *softwares* **estatísticos**.

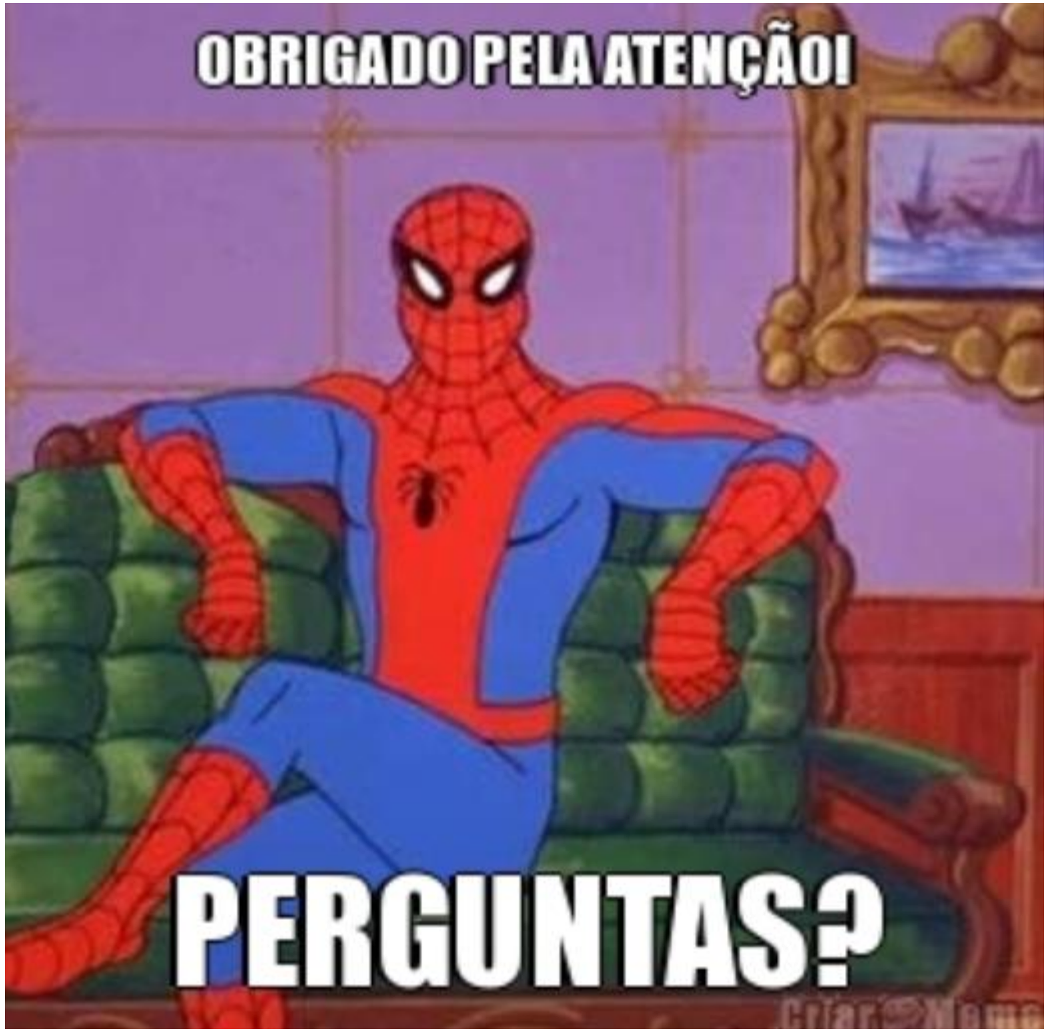
REFERÊNCIAS

MONTGOMERY, D. C.; RUNGER, G. C. Estatística aplicada e probabilidade para engenheiros. 6 ed. Rio de Janeiro: LTC, 2018. 628p.

MORETTIN, P. A.; BUSSAB, W. O. Estatística básica. 9 ed. São Paulo: Saraiva, 2017. 554p.

UNIFESP
UNIVERSIDADE FEDERAL DE SÃO PAULO
1933

CLASS FINISHED



UNIFESP
UNIVERSIDADE FEDERAL DE SÃO PAULO
1933