Distribuição amostral

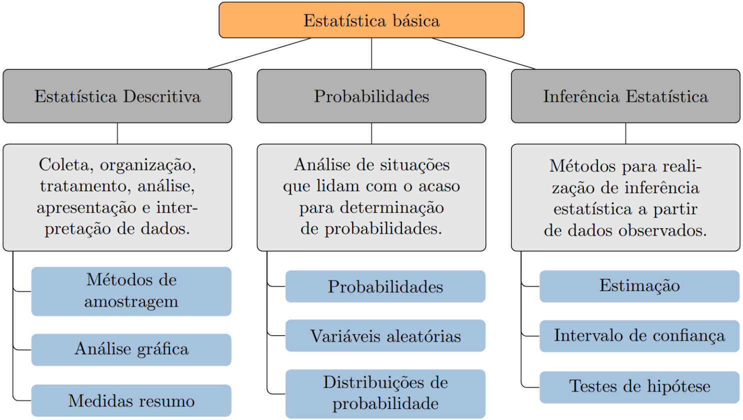
Parte 1

Prof.: Eduardo Vargas Ferreira

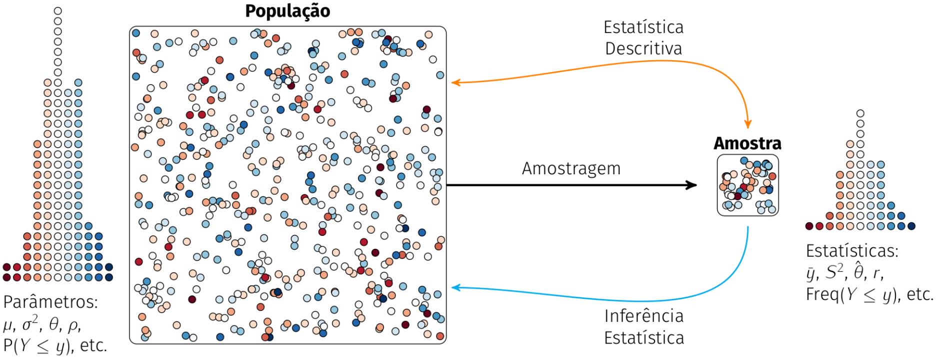
UFPR
UNIVERSIDADE FEDERAL DO PARANÁ

DEST
Departamento de Estatística

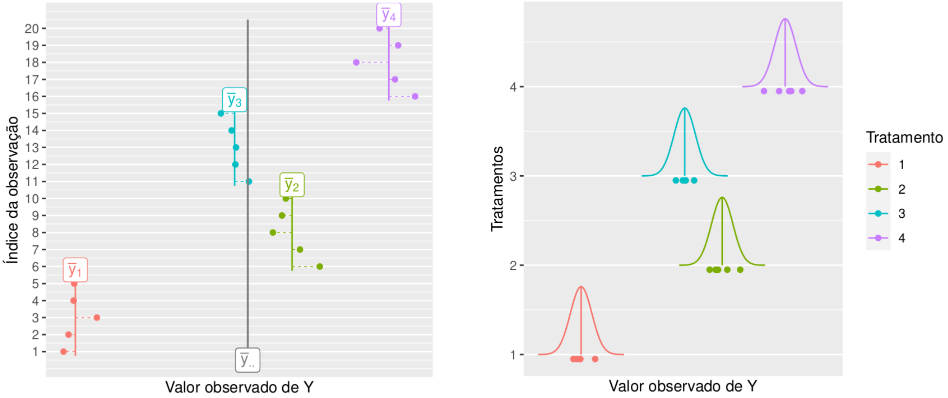
Resumo


Estatística básica
Estatística Descritiva
Probabilidades
Inferência Estatística
Coleta, organização, tratamento, análise, apresentação e interpretação de dados.
Análise de situações que lidam com o acaso para determinação de probabilidades.
Métodos para realização de inferência estatística a partir de dados observados.
Métodos de amostragem
Análise gráfica
Medidas resumo
Probabilidades
Variáveis aleatórias
Distribuições de probabilidade
Estimação
Intervalo de confiança
Testes de hipótese

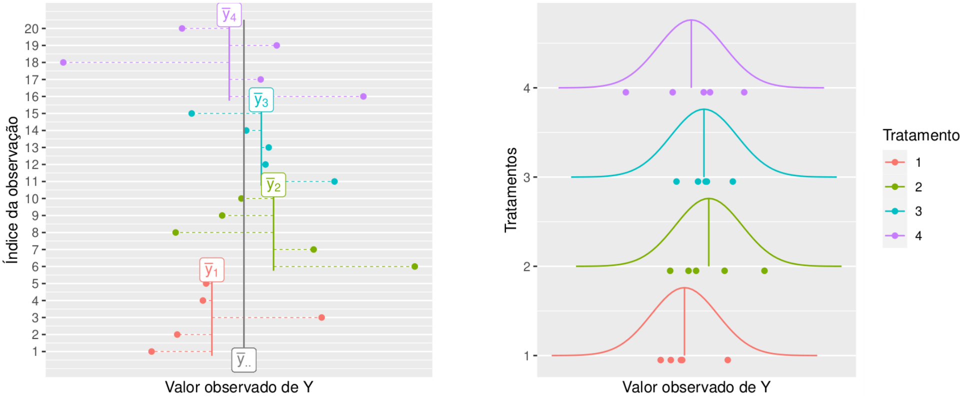
O que vimos até aqui



Exemplo: diferença entre os tratamentos



Exemplo: diferença entre os tratamentos



Parâmetros e Estatísticas

Definição: um **parâmetro** é uma medida usada para descrever uma característica da população.

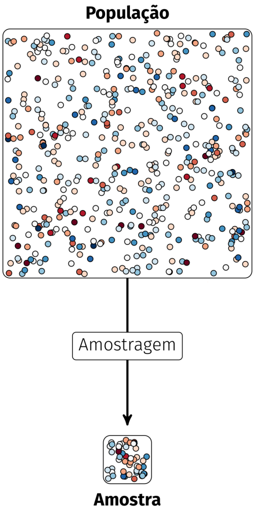
Denominação	População
Média	μ
Variância	σ^2
Proporção	p
Mediana	Q_2
Intervalo inter-quartil	$d_Q = Q_3 - Q_1$
Função de densidade	$f(x)$
Função de distribuição	$F(x)$

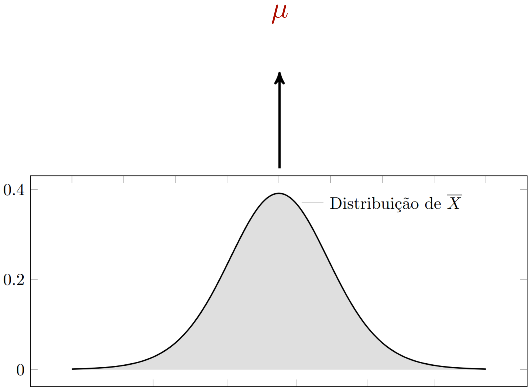
Parâmetros e Estatísticas

Definição: Uma **estatística** é uma característica da amostra.

Denominação	População	Amostra
Média	μ	$\overline{X} = \sum \frac{X_i}{n}$
Variância	σ^2	$S^2 = \sum \frac{\left(X_i - \overline{X}\right)^2}{(n-1)}$
Proporção	p	$\hat{p}$
Mediana	Q_2	q_2
Intervalo inter-quartil	$d_Q = Q_3 - Q_1$	$d_q = q_3 - q_1$
Função de densidade	$f(x)$	histograma
Função de distribuição	$F(x)$	$F_e(x)$

Parâmetros e Estatísticas





$\bar{x_5} = 4.7 \quad \bar{x_2} = 5.1 \quad \bar{x_1} = 5.6 \quad \bar{x_3} = 6.1 \quad \bar{x_4} = 6.8$

Distribuição amostral da média

Distribuição amostral da média

Definição: Seja X uma v.a. com média μ e variância σ^2, e seja $(X_1, \ldots, X_n)$ uma AAS de X. Então,

$$\mathbb{E}(\bar{X}) = \mu$$

$$\mathbb{V}ar(\bar{X}) = \frac{\sigma^2}{n}$$

$$\begin{aligned}
\mathbb{E}(\bar{X}) &= \mathbb{E}\left(\frac{X_1 + X_2 + \ldots + X_n}{n}\right) \\
&= \frac{\mathbb{E}(X_1) + \mathbb{E}(X_2) + \ldots + \mathbb{E}(X_n)}{n} \\
&= \frac{n \times \mathbb{E}(X_1)}{n} \\
&= \mu.
\end{aligned}$$

$$\begin{aligned}
\mathbb{V}ar(\bar{X}) &= \mathbb{V}ar\left(\frac{X_1 + X_2 + \ldots + X_n}{n}\right) \\
&= \frac{\mathbb{V}ar(X_1) + \mathbb{V}ar(X_2) + \ldots + \mathbb{V}ar(X_n)}{n^2} \\
&= \frac{n \times \mathbb{V}ar(X_1)}{n^2} \\
&= \frac{\sigma^2}{n}.
\end{aligned}$$

Exemplo: capacidade máxima no elevador

- A capacidade máxima de um elevador é de 500 kg. Se a distribuição dos pesos dos usuários é $N(70, 100)$, qual a probabilidade de que 7 pessoas ultrapassem este limite?

X_i = o peso do indivíduo i.

$$
\begin{aligned}
P\left(\sum_{i=1}^{7} X_i > 500\right) &= P\left(\frac{\sum_{i=1}^{7} X_i}{7} > \frac{500}{7}\right) \\
&= P\left(\bar{X} > 71,42\right) \quad \text{como} \quad \bar{X} \sim N\left(70, \frac{100}{7}\right) \\
&= P\left(\frac{\bar{X} - 70}{\sqrt{100/7}} > \frac{71,42 - 70}{\sqrt{100/7}}\right) \\
&= P\left(Z > 0.37\right) = 0.35
\end{aligned}
$$

Exemplo: salário de pilotos

- O salário anual médio dos pilotos de avião pode ser modelado por uma distribuição $N(41979, 5000^2)$. Suponha que uma AAS de 50 pilotos seja selecionada.

X_i = salário anual do i-ésimo piloto.

1. Qual a probabilidade da média amostral não diferir da média populacional em até $R\$1000,00$?

$$
\begin{aligned}
P\left(\,|\bar{X}-\mu| < 1000\,\right) &= P\left(\frac{-1000}{5000/\sqrt{50}} < \frac{\bar{X}-\mu}{5000/\sqrt{50}} < \frac{1000}{5000/\sqrt{50}}\right) \\
&= P\left(-1.41 < Z < 1.41\right) \\
&\approx 0.84
\end{aligned}
$$

Exemplo: salário de pilotos

- O salário anual médio dos pilotos de avião pode ser modelado por uma distribuição $N(41979, 5000^2)$. Suponha que uma AAS de 50 pilotos seja selecionada.

X_i = salário anual do i-ésimo piloto.

2. Como a probabilidade do item anterior seria alterada caso a amostra fosse de tamanho 100?

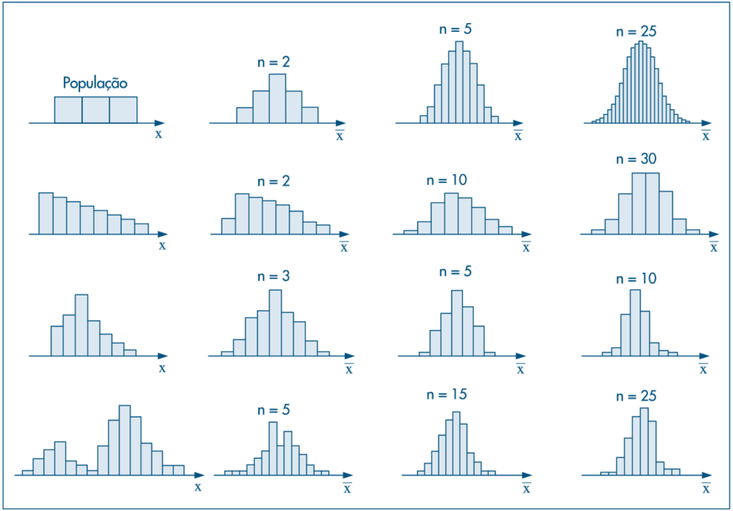
$$
\begin{aligned}
P\left(|\bar{X} - \mu| < 1000\right) &= P\left(\frac{-1000}{5000/\sqrt{100}} < \frac{\bar{X} - \mu}{5000/\sqrt{100}} < \frac{1000}{5000/\sqrt{100}}\right) \\
&= P\left(-2 < Z < 2\right) \\
&\approx 0.95
\end{aligned}
$$

Vimos exemplos com distribuição normal

X_i = o peso do indivíduo i.

X_i = salário anual do i-ésimo piloto.

Caso geral



População
x
n = 2
$\bar{x}$
n = 5
$\bar{x}$
n = 25
$\bar{x}$
x
n = 2
$\bar{x}$
n = 10
$\bar{x}$
n = 30
$\bar{x}$
x
n = 3
$\bar{x}$
n = 5
$\bar{x}$
n = 10
$\bar{x}$
x
n = 5
$\bar{x}$
n = 15
$\bar{x}$
n = 25
$\bar{x}$

Teorema Central do Limite

Distribuição amostral da média

- Seja $X_1, X_2, \ldots, X_n$ uma sequência de variáveis aleatórias independentes e identicamente distribuídas, cada uma com $\mathbb{E}(X_i) = \mu$ e $\mathbb{V}ar(X_i) = \sigma^2$, então:

$$\mathbb{E}(\bar{X}) = \mu$$

$$\mathbb{V}ar(\bar{X}) = \frac{\sigma^2}{n}$$

$$\begin{aligned}
\mathbb{E}(\bar{X}) &= \mathbb{E}\left(\frac{X_1 + X_2 + \ldots + X_n}{n}\right) \\
&= \frac{\mathbb{E}(X_1) + \mathbb{E}(X_2) + \ldots + \mathbb{E}(X_n)}{n} \\
&= \mu.
\end{aligned}$$

$$\begin{aligned}
\mathbb{V}ar(\bar{X}) &= \mathbb{V}ar\left(\frac{X_1 + X_2 + \ldots + X_n}{n}\right) \\
&= \frac{\mathbb{V}ar(X_1) + \mathbb{V}ar(X_2) + \ldots + \mathbb{V}ar(X_n)}{n^2} \\
&= \frac{\sigma^2}{n}.
\end{aligned}$$

Distribuição amostral da média

- Seja $X_1, X_2, \ldots, X_n$ uma sequência de variáveis aleatórias independentes e identicamente distribuídas, cada uma com $\mathbb{E}(X_i) = \mu$ e $\mathbb{V}ar(X_i) = \sigma^2$, então:

$$\mathbb{E}(\bar{X}) = \mu$$

$$\mathbb{V}ar(\bar{X}) = \frac{\sigma^2}{n}$$

- O **Teorema Central do Limite** nos diz que

$$\frac{\bar{X} - \mathbb{E}\left[\bar{X}\right]}{\sqrt{\mathbb{V}ar\left[\bar{X}\right]}} = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \xrightarrow{D} N(0,1), \quad \text{quando} \quad n \to \infty$$

Distribuição amostral da média

- Seja $X_1, X_2, \ldots, X_n$ uma sequência de variáveis aleatórias independentes e identicamente distribuídas, cada uma com $\mathbb{E}(X_i) = \mu$ e $\mathbb{V}ar(X_i) = \sigma^2$, então:

$$\mathbb{E}(\bar{X}) = \mu$$

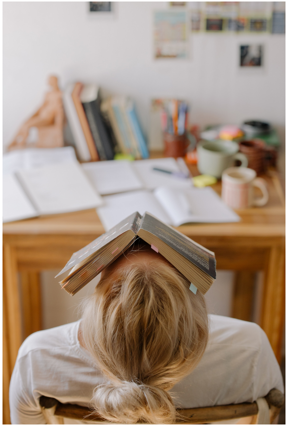
$$\mathbb{V}ar(\bar{X}) = \frac{\sigma^2}{n}$$

- O **Teorema Central do Limite** nos diz que

$$\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \xrightarrow{D} N(0,1), \quad \text{quando} \quad n \to \infty$$

Exemplo: notas dos alunos

- Suponha que X_i, $i = 1, \ldots, 10$, sejam as notas dos alunos de uma classe de Estatística, tal que $E(X_i) = 5$ e $Var(X_i) = 16$ **(conhecido)**. Calcule uma aproximação para



$$P\left(\bar{X} > 7\right)$$

Sabemos que $E\left(\bar{X}\right) = 5$ e $Var\left(\bar{X}\right) = 16/10$. Então, pelo TCL,

$$\begin{aligned} P\left(\bar{X} > 7\right) &= P\left(\frac{\bar{X} - 5}{\sqrt{\frac{16}{10}}} > \frac{7 - 5}{\sqrt{\frac{16}{10}}}\right) \\ &= P\left(Z > 1.58\right) \\ &\approx 0.057. \end{aligned}$$

Referências

- Bussab, WO; Morettin, PA. Estatística Básica. São Paulo: Editora Saraiva, 2006 (5ª Edição).
- Magalhães, MN; Lima, ACP. Noções de Probabilidade e Estatística. São Paulo: EDUSP, 2008.