

## Semana 7

### Introdução à Amostragem e Inferência Estatística

#### Links Fixos

##### Livro PDF Teórico:

[https://edisciplinas.usp.br/pluginfile.php/4445638/mod\\_resource/content/1/Book\\_EstatBas%20-%20Morettin%20%20Bussab.pdf](https://edisciplinas.usp.br/pluginfile.php/4445638/mod_resource/content/1/Book_EstatBas%20-%20Morettin%20%20Bussab.pdf)

##### Site educativo:

<https://seeing-theory.brown.edu/basic-probability>

#### Capítulo 10

Livro Estatística Básica - Morettin e Bussab

#### Introdução a amostragem - explicação sobre o contexto da amostragem e os tipos de amostra

[https://drive.google.com/a/pyfound.org/file/d/18ccmlA3v7C-Jh\\_TSElWat7r3qPkaFEJa/view?usp=drivesdk](https://drive.google.com/a/pyfound.org/file/d/18ccmlA3v7C-Jh_TSElWat7r3qPkaFEJa/view?usp=drivesdk)

#### Artigo Sobre Viés na Seleção de Amostra

Health Indicators: Eliminating bias from convenience sampling estimators

Bethany L. Hedt and Marcello Pagano

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3951147/>

#### Sobre viés em algoritmo e como ajustar isso

<https://www.ere.net/dont-embed-your-bias-in-your-algorithm/>

#### Exemplo dos casos de contratação com viés de Etnia e Idade

<https://www.linkedin.com/pulse/fix-labor-market-discrimination-bas-van-de-haterd/>

#### Sobre o algoritmo com viés de contratação sexista da Amazon

<https://www.ere.net/what-we-should-take-away-from-amazons-embedding-bias-in-its-algorithm-and-stopping-the-project/>

<https://www.theverge.com/2018/10/10/17958784/ai-recruiting-tool-bias-amazon-report>

## Parte prática

Usando os dados do Prodam vamos estimar a média salarial com a abordagem de

- Amostra Aleatória Simples
- Amostra Aleatória Estratificada
- Amostra Aleatória por Conglomerado

### Amostra Aleatória Simples

Para criar uma amostra aleatória estratificada, há quatro etapas:

1) definindo a população;

Em nosso exemplo, a população é dos funcionários do Prodam de 2018. A população é expressa em N. Como estamos interessados em todos esses funcionários, podemos dizer que nossa população é de X. Se estivéssemos interessados apenas em funcionárias do Prodam, por exemplo, excluiríamos todos os homens na criação de nosso quadro de amostragem.

2) listando a população;

Precisamos identificar todos os funcionários.

3) escolhendo o tamanho da amostra;

Informações mais detalhadas sobre esse cálculo em

<https://drive.google.com/file/d/1prfw8WCNZcarUZF4fsIb5zphaxpHZoIF/view?usp=sharing>

A fórmula para cálculo do tamanho da amostra para uma estimativa confiável da MÉDIA POPULACIONAL ( $\mu$ ) é dada por:

$$n = \left( \frac{Z_{\alpha/2} \cdot \sigma}{E} \right)^2$$

Equação 1

Onde:

- |                |   |  |
|----------------|---|--|
| n              | = | Número de indivíduos na amostra  |
| $Z_{\alpha/2}$ | = | Valor crítico que corresponde ao grau de confiança desejado.   |
| $\sigma$       | = | Desvio-padrão populacional da variável estudada (no exemplo, RENDA).   |
| E              | = | Margem de erro ou ERRO MÁXIMO DE ESTIMATIVA. Identifica a diferença máxima entre a MÉDIA AMOSTRAL ( $\bar{X}$ ) e a verdadeira MÉDIA |

Precisamos definir qual Grau de Confiança adotaremos para encontrar o tamanho da amostra.

Tabela 1 – Valores críticos associados ao grau de confiança na amostra

Grau de Confiança	$\alpha$	Valor Crítico $Z_{\alpha/2}$
90%	0,10	1,645
95%	0,05	1,96
99%	0,01	2,575

4) selecionando a amostra.

### Amostra Aleatória Estratificada

Para criar uma amostra aleatória estratificada, há sete etapas:

1) definindo a população;

2) escolha da estratificação relevante;

Se quiséssemos analisar as diferenças entre funcionários do sexo masculino e feminino, isso significaria escolher o sexo como estratificação, mas poderia envolver da mesma forma a escolha de funcionários de diferentes departamentos ou outras variáveis.

Para os fins deste exemplo, qual variável usaremos como estratos?

3) listando a população;

Precisamos identificar todos os funcionários.

4) listar a população de acordo com a estratificação escolhida;

Assim como nas técnicas de amostragem aleatória simples e de amostragem aleatória sistemática, precisamos atribuir um número consecutivo de 1 a NK a cada um dos funcionários em cada estrato. Como resultado, terminaríamos com duas listas, uma detalhando todos os funcionários do sexo masculino e outra detalhando todas as funcionárias.

5) escolhendo o tamanho da amostra;

Vamos imaginar que escolhemos um tamanho de amostra de 100 funcionários. A amostra é expressa como  $n$ . Esse número foi escolhido porque reflete o limite de nosso orçamento e o tempo que temos para distribuir nosso questionário aos funcionários, por exemplo. No entanto, também poderíamos ter determinado o tamanho da amostra necessário usando um cálculo de tamanho da amostra, que é uma ferramenta estatística particularmente útil.

6) calcular uma estratificação proporcional;

Imagine que dos funcionários, 60% deles são do sexo feminino e 40% do sexo masculino. Precisamos garantir que o número de unidades selecionadas para a amostra de cada estrato seja proporcional ao número de homens e mulheres na população. Para conseguir isso, primeiro multiplicamos o tamanho da amostra desejado ( $n$ ) pela proporção de

unidades em cada estrato. Portanto, para calcular o número de mulheres necessário em nossa amostra, multiplicamos 100 por 0,60 (ou seja,  $0,60 = 60\%$  da população de funcionárias), o que nos dá um total de 60 mulheres. Se fizermos o mesmo com funcionários do sexo masculino, obteremos 40 homens (ou seja,  $40\%$  são do sexo masculino, onde  $100 \times 0,40 = 40$ ). Isso significa que precisamos selecionar 60 funcionários do sexo feminino e 40 do sexo masculino para a nossa amostra de 100 funcionários.

7) usando uma amostra aleatória simples para selecionar sua amostra.

Agora que escolhemos para nossa amostra 40 funcionários do sexo masculino e 60 do sexo feminino, ainda precisamos selecionar esses funcionários da população listada. Fazemos isso usando amostragem aleatória simples.