

# Tópicos de Amostragem: Atividade Avaliativa

Prof. Dr. Fabiano F. T. dos Santos

2024-12-14

## Instruções

- O desenvolvimento desta atividade deve ser realizada de forma individual ou em dupla.
- Deve-se completar o arquivo Rmd enviado na atividade.
- É necessário devolver o arquivo em Rmd e em pdf.
- Valor da atividade: 10 pontos.

**Data disponível: 14/12/2024 .**

**Data máxima para entrega: 14/02/2025**

## Conjunto de dados - Nascidos Vivos em 2024

O conjunto de dados que será utilizado nesta atividade foi obtido no site do OpenDataSUS e consiste da base de informações sobre nascidos vivos 2024 (parcial). Esse conjunto de dados foi extraído no seguinte link <https://opendatasus.saude.gov.br/dataset/sistema-de-informacao-sobre-nascidos-vivos-sinasc> (acesso e download em 04/12/2024), onde também encontramos seu respectivo dicionário de dados. O arquivo Dataframe\_AulaFabiano.csv está disponível na plataforma do curso.

A Atividade Avaliativa consiste nos itens a seguir.

**Item 1 - Considere o plano amostral de Amostragem Aleatória Simples com reposição e a variável IDADEMAE. Com o objetivo de estimar a média das idades de uma população, determine o tamanho necessário da amostra considerando um erro máximo de 1 ano, um nível de confiança de 95% e o conhecimento da variância populacional ( $\sigma^2 = 44,8$ ). Após dimensionar a amostra, selecione-a aleatoriamente, calcule a média amostral e construa um intervalo de confiança para essa média. Por fim, interprete os resultados obtidos.**

O dimensionamento da amostra deverá ser feito utilizando-se a fórmula  $n = \frac{\sigma^2}{(\frac{B}{z_\alpha})^2}$  e o intervalo de confiança deverá ser construído utilizando-se a fórmula  $\left( \bar{x} - z_\alpha \sqrt{\frac{\sigma^2}{n}}; \bar{x} + z_\alpha \sqrt{\frac{\sigma^2}{n}} \right)$ .

**Item 2 - Considere o plano amostral de Amostragem Aleatória Simples sem reposição e a variável IDADEMAE. Com o objetivo de estimar a média das idades de uma população, determine o tamanho necessário da amostra considerando um erro máximo de 1 ano, um nível de confiança de 95% e o conhecimento da variância populacional ( $\sigma^2 = 44,8$ ). Após dimensionar a amostra, selecione-a aleatoriamente, calcule a média amostral e construa um intervalo de confiança para essa média. Por fim, interprete os resultados obtidos.**

O dimensionamento da amostra deverá ser feito utilizando-se a fórmula  $n = \frac{1}{\frac{B^2}{z_\alpha^2 \sigma^2} + \frac{1}{N}}$ , onde  $N$  é o tamanho da população e o intervalo de confiança deverá ser construído utilizando-se a fórmula  $\left(\bar{x} - z_\alpha \sqrt{\left(1 - \frac{n}{N}\right) \frac{\sigma^2}{n}}; \bar{x} + z_\alpha \sqrt{\left(1 - \frac{n}{N}\right) \frac{\sigma^2}{n}}\right)$ .

**Item 3 - Existem diversas maneiras de classificar as pessoas, cada uma com um propósito diferente. Uma das classificações úteis, por exemplo, para questões de marketing, é a classificação em classes sociais. Considerando os critérios de classificação empregados atualmente no Brasil, as categorias podem ser generalizadas da seguinte forma:**

**Classe A: Famílias com renda mensal igual ou superior a R\$ 14.400,00.**

**Classe B: Famílias com renda mensal entre R\$ 7.100,00 e R\$ 14.399,00.**

**Classe C: Famílias com renda mensal entre R\$ 2.600,00 e R\$ 7.099,00.**

**Classe D: Famílias com renda mensal igual ou inferior a R\$ 2.599,00.**

**Suponha que uma determinada população em estudo esteja distribuída nesses estratos da seguinte forma:**

**Classe A: 60 famílias**

**Classe B: 90 famílias**

**Classe C: 120 famílias**

**Classe D: 480 famílias**

**Pretende-se selecionar 100 famílias para analisar o comportamento de consumo dessa população. Responda:**

**1) Quantas unidades amostrais devem ser retiradas de cada classe, se adotarmos o plano de Amostragem Aleatória Estratificada Uniforme?**

**Lembre-se de que o número de elementos do estrato  $h$  que comporá a amostra é dado por  $n_h = \frac{n}{H}$ , onde  $n$  é o tamanho da amostra e  $H$  é a quantidade de estratos.**

**2) Quantas unidades amostrais devem ser retiradas de cada classe, se adotarmos o plano de Amostragem Aleatória Estratificada Proporcional?**

**Lembre-se de que o número de elementos do estrato  $h$  que comporá a amostra é dado por  $n_h = n \cdot \frac{N_h}{N}$ , onde  $n$  é o tamanho da amostra,  $N_h$  é o tamanho de estrato  $h$  e  $N$  é o tamanho da população.**

**Observações:**

- (i) Nos itens 1 e 2 é necessário apresentar o código relativo ao dimensionamento, o código relativo à seleção da amostra, o código relativo ao cálculo da média amostral e código relativo à construção do intervalo de confiança.
- (ii) No item 3, é necessário apresentar os códigos referentes aos dimensionamentos amostrais.

- (iii) Cada item vale 10,0 pontos e a nota final é a média aritmética das notas obtidas nos itens.
- (iv) A entrega da atividade deve ser realizada em Rmd e um arquivo no formato pdf, gerado a partir do RMarkdown.