

Professor: Fábio Kravetz

Leia atentamente as instruções:

- Responder a todas as questões apresentadas abaixo. O objetivo da atividade em questão é a consolidação da compreensão sobre o conteúdo.
- As soluções devem ser realizadas com a linguagem R.

1. Uma agência de marketing está planejando duas pesquisas distintas para avaliar a aceitação de um novo produto. Para garantir a validade estatística, ambos os projetos devem seguir os mesmos critérios de rigor:
 - Nível de Confiança: 95%;
 - Erro máximo: 5%
 - Proporção estimada: Como não há dados prévios, deve-se usar o valor mais conservador de 50% para maximizar o tamanho da amostra e garantir precisão.

Utilizando a linguagem R, implemente um script que:

- a) Calcule e exiba o tamanho da amostra necessário para a **pesquisa com a população finita** de 1.000 clientes.
 - b) Calcule e exiba o tamanho da amostra necessário para a **pesquisa com a população infinita**. c) (Opcional) Compare os resultados e explique por que são diferentes.
2. Uma empresa com 1000 funcionários deseja realizar uma pesquisa interna para medir o nível de satisfação com o ambiente de trabalho. A equipe de RH necessita de um plano de amostragem que garanta resultados confiáveis para executar ações de melhoria, caso estas se façam necessárias.

Deste modo, foram definidos os seguintes parâmetros estatísticos para a pesquisa:

- Nível de confiança: 95%;
- Erro máximo: 5%;
- Proporção estimada: Como não há dados prévios, deve-se usar o valor mais conservador de 50% para maximizar o tamanho da amostra e garantir precisão.

Você foi encarregado de usar a linguagem R para preparar o processo de seleção dos funcionários que participarão da pesquisa. Os funcionários são identificados por números de 1 a 1000.

Desenvolva um script em R que:

1. **Calcule o número de funcionários (n)** que precisam ser entrevistados, com base nos parâmetros fornecidos e na fórmula de cálculo para populações finitas.
 2. **Gere uma primeira lista de funcionários** utilizando a técnica de **amostragem aleatória simples**. Exiba a lista dos IDs selecionados e a contagem total.
 3. **Gere uma segunda lista de funcionários** utilizando a técnica de **amostragem sistemática**. Exiba a lista dos IDs selecionados e a contagem total para esta segunda abordagem.
3. Uma equipe de pesquisadores educacionais da UFPE, em Recife, está planejando dois estudos distintos para entender melhor a experiência dos estudantes. Ressalta-se que será utilizada a linguagem R para selecionar as amostras.
 - 1) O primeiro estudo visa analisar os hábitos de estudo dos alunos. A população de **1.000 estudantes** de um determinado setor está dividida em três grandes áreas de conhecimento (estratos). A composição é a seguinte:
 - **Estrato A (Humanas):** 500 alunos
 - **Estrato B (Exatas):** 300 alunos
 - **Estrato C (Biológicas):** 200 alunos

Os pesquisadores querem garantir que a amostra final de **100 alunos** seja representativa, mantendo as mesmas proporções de cada área de conhecimento.

- Desenvolva um script em R que calcule o **tamanho da amostra proporcional** a ser coletada de cada estrato, arredondando o valor para cima.
- Para garantir que os resultados sejam reproduzíveis, utilize a semente aleatória `set.seed(123)`.
- Para cada estrato, gere uma **amostra aleatória simples** de IDs de estudantes (considere os IDs de 1 até o total de alunos do estrato).
- Exiba as listas de IDs selecionados para cada um dos três estratos.

- 2) O segundo estudo tem como objetivo avaliar a qualidade da infraestrutura de pesquisa. A universidade possui **10 centros de pesquisa (conglomerados)**, cada um representando uma unidade primária de amostragem.

Por questões logísticas, a equipe só poderá visitar e avaliar a infraestrutura de **3 centros de pesquisa**, que devem ser escolhidos aleatoriamente.

- Utilizando R, realize uma seleção aleatória dos 3 centros de pesquisa que serão incluídos no estudo.
- Para garantir a reprodutibilidade da seleção, utilize a semente aleatória `set.seed(123)`.
- Exiba os IDs dos centros de pesquisa (conglomerados) selecionados.

4. Em uma população de 10.000 indivíduos, deseja-se estimar a proporção de pessoas que usam um determinado serviço. Sabendo que a proporção esperada é de 40%, com um erro máximo de 5% e nível de confiança de 95%, calcule o tamanho da amostra necessária e selecione a amostra aleatória simples.

5. Uma cidade com 15.000 habitantes é dividida em três regiões: região A (60% da população), região B (25%) e região C (15%). Deseja-se realizar uma pesquisa com margem de erro de 3% e nível de confiança de 95%. Determine o tamanho da amostra total e a amostra proporcional de cada estrato, e selecione as amostras.

6. Uma empresa possui um banco de dados com 1.000 clientes e deseja realizar uma pesquisa de satisfação. A empresa precisa selecionar uma amostra de clientes para serem entrevistados, de forma sistemática. Para garantir uma margem de erro de 5% e um nível de confiança de 95%, considere que a proporção esperada de clientes satisfeitos é de 50%.