

Considerem que o arquivo com os dados usados na Lista de Exercícios 1, referente aos alunos de EAD da TYU seja a POPULAÇÃO da sua pesquisa. Desta população vocês irão retirar amostras aleatórias para resolver os problemas a seguir. O relatório com as respostas deverá ser entregue IMPRESSO com todas as tabelas, medidas e gráficos julgados convenientes, produzidos através do R, Excel ou qualquer outro aplicativo. As amostras coletadas devem ser enviadas por e-mail ao professor, devidamente identificadas com os nomes dos alunos e nomes das variáveis.

Primeira parte – Distribuição Amostral da Média

1) Usando o suplemento do Excel disponibilizado pelo professor proceda a retirada de 200 amostras aleatórias da variável **Renda** (primeiramente excluam as linhas com dados perdidos), com os seguintes tamanhos: 4, 8, 16, 30 e 100 elementos. Com base nos resultados encontrados responda os itens a seguir:

- a) Pelo teorema central do limite admite-se que o valor esperado da média amostral seja a média populacional que se pretende estimar, e eles serão tão mais próximos à medida que aumenta o tamanho da amostra. Isso é confirmado pelos resultados obtidos nas amostras? JUSTIFIQUE.
- b) Além do que foi dito acima, admite-se que o desvio padrão das médias amostrais será igual ao desvio padrão populacional dividido pela raiz quadrada do tamanho da amostra, e eles serão tão mais próximos à medida que aumenta o tamanho da amostra. As amostras retiradas confirmam essa afirmação? JUSTIFIQUE.
- c) Outra característica muito importante do teorema central do limite é que a distribuição amostral da média será cada vez mais próxima de uma distribuição normal, à medida que o tamanho da amostra aumenta, independentemente da forma da distribuição da variável na população. Sabe-se que a variável **Renda** não tem distribuição normal na população. Os resultados das amostras corroboram esta afirmação (usem os gráficos apropriados). JUSTIFIQUE.

Segunda parte – Intervalos de Confiança para Média

2) Há uma grande preocupação em estimar o comportamento da variável **Idade** dos alunos, de maneira a caracterizar melhor seu o perfil. Alguém da universidade sugeriu que vocês retirassem uma amostra de 20 alunos e registrassem os valores de Idade. Com base nos resultados encontrados respondam os itens a seguir (lembre-se de excluir inicialmente as linhas com dados perdidos e de que o tamanho da população é **conhecido**).

- a) Encontrem o intervalo de 95% de confiança para a média populacional da Idade dos alunos. Interpretem o resultado.
- b) Se quiséssemos encontrar um intervalo de 99% de confiança para a média populacional da Idade dos alunos, com uma precisão de 2 anos, a amostra coletada seria suficiente? JUSTIFIQUE.
- c) Você concorda com o plano de amostragem usado (que considerou a população homogênea)? JUSTIFIQUE.

3) A estimação da média populacional da variável **Renda** também é muito importante. Retirem uma amostra de 20 alunos e registrem seus valores. Com base nos resultados encontrados respondam os itens a seguir (lembre-se de excluir inicialmente as linhas com dados perdidos e de que o tamanho da população é **conhecido**).

- a) Encontre o intervalo de 95% de confiança para a média populacional de Renda. Interpretem o resultado.
- b) Repitam o item a para 99% de confiança. O intervalo ficou muito “largo”? O que poderia ser feito para torná-lo mais preciso? JUSTIFIQUEM.
- c) Qual deveria ser o tamanho mínimo de amostra para determinar um intervalo de 99% de

confiança e precisão de 1,5 salários mínimos para a Renda mensal dos alunos?

Terceira parte – Intervalos de Confiança para Proporção

4) Vamos avaliar a forma de **Pagamento** usada pelos alunos para quitar suas mensalidades. Retirem uma amostra aleatória de 200 elementos (novamente, primeiramente excluam as linhas com dados perdidos). Registrem a proporção amostral de alunos que usam incentivos federais e respondam os itens a seguir (lembre-se de que o tamanho da população é **conhecido**):

- Encontrem o intervalo de 95% de confiança para a proporção populacional de alunos que usam incentivos federais. Interpretem o resultado.
- Suponha que haja interesse em obter um intervalo de 95% para a proporção populacional de alunos que usam incentivos federais com uma precisão de 2%. A amostra coletada é suficiente? JUSTIFIQUE.
- Qual deveria ser o tamanho mínimo de amostra para estimar com 95% de confiança e precisão de 2% a proporção populacional de alunos que usam incentivos federais, se vocês não pudessem ter tirado a amostra piloto?

5) Outra grande preocupação é a avaliação da **Opinião** geral dos alunos sobre a EAD da TYU. Retirem uma amostra aleatória de 200 elementos (novamente, primeiramente excluam as linhas com dados perdidos). Registrem a proporção amostral de alunos com opiniões “Insatisfeito” e “Muito insatisfeito” (opiniões negativas) e respondam os itens a seguir (lembre-se de que o tamanho da população é **conhecido**):

- Encontrem o intervalo de 95% de confiança para a proporção populacional de alunos com opiniões negativas sobre a EAD da TYU. Interpretem o resultado.
- Suponha que haja interesse em obter um intervalo de 95% para a proporção populacional de alunos com opiniões negativas sobre a EAD da TYU, com uma precisão de 2%. A amostra coletada é suficiente? JUSTIFIQUE.
- Qual deveria ser o tamanho mínimo de amostra para estimar com 95% de confiança e precisão de 2% a proporção populacional de alunos com opiniões negativas sobre a EAD da TYU, se vocês não pudessem ter tirado a amostra piloto?

6) Imaginem que a TYU considera alunos “alvo” aqueles clientes que tem **Renda** familiar mensal de *mais de* 2,5 salários mínimos. Recodifiquem a variável Renda em uma variável qualitativa com dois valores (primeiramente retirem as linhas com dados perdidos): alunos-não alvo e alunos-alvo. Retirem uma amostra aleatória de 250 elementos (posicione a primeira célula do intervalo de população na coluna onde está a variável recodificada). Registrem a proporção amostral de alunos-alvo e respondam os itens a seguir (lembre-se de que o tamanho da população é **conhecido**):

- Encontrem o intervalo de 95% de confiança para a proporção populacional de alunos-alvo na variável Renda recodificada. Interpretem o resultado.
- Suponha que haja interesse em obter um intervalo de 95% para a proporção populacional de alunos-alvo, com uma precisão de 2%. A amostra coletada é suficiente?
- Qual deveria ser o tamanho mínimo de amostra para estimar com 95% de confiança e precisão de 2% a proporção populacional de alunos-alvo, se vocês não pudessem ter tirado a amostra piloto?
- Se vocês pudessem escolher, o que iriam preferir: estimar o intervalo de confiança da variável Renda diretamente ou após a recodificação em variável qualitativa? JUSTIFIQUEM.