

ME 430 A - Técnicas de amostragem
Segundo semestre de 2021
Lista de Exercícios III

OBS: As resoluções de todas as questões relativas à análise de dados devem sempre incluir análises descritivas e inferenciais apropriadas, EPA (todos os pertinentes), estimativas pontuais, intervalares, testes de hipótese (quando pertinentes). Se nada for dito, a escolha do(s) estimador(es), plano(s) amostral(is), níveis de confiança/significância, tamanho(s) amostral(is), tipo(s) de alocação(ões), precisão/erro da estimativa, etc, fica à cargo do(s) aluno(s). Considere que as unidades a serem selecionadas, independentemente do PA, tem a mesma probabilidade de serem selecionadas. Apresente a solução em forma de relatório, sempre com as devidas: justificativas, comentários e conclusões.

1. Resolva os exercícios deixados em sala.
2. Seja $Y = \sum_{i=1}^N X_i$, em que $X_i \sim \text{Bernoulli}(p)$, independentes entre si e $N \sim \text{Poisson}(\lambda)$, independente de todos os X_i 's. Calcule a esperança e a variância de Y . Sugestão: use esperanças e variâncias condicionais.
3. Considere o seguinte conjunto de variáveis $Y_i = \alpha + \beta x_i + \xi_i, i = 1, 2, \dots, n$, em que ξ_i são variáveis aleatórias independentes e identicamente distribuídas (como uma certa média e uma certa variância). Determine α e β que minimiza $\sum_{i=1}^n (Y_i - \alpha - \beta x_i)^2$.
4. Do livro Bolfarine & Bussab (2005):
 - Capítulo 5, páginas 138, 139, 140 e 141 : exercícios 5.1, 5.3, 5.4, 5.5, 5.7, 5.11, 5.12, 5.13, 5.15.
5. Do livro Silva et al. (2021):
 - Capítulo 6, seção 6.8, exercício: 6.1, 6.2, 6.3.
6. Considere o arquivo “Rendimentos.csv”, relativo aos dados do livro Silva et al. (2021) (disponíveis no site do curso). Considere que ele representa uma população de interesse, que consiste nos rendimentos de domicílios (renda total dos moradores que trabalham) de determinada cidade. Estime a renda média por domicílio.
7. Considere o arquivo “Aves.csv”, relativo aos dados do livro Silva et al. (2021) (disponíveis no site do curso). Considere que ele representa uma população de interesse, que consiste nos tipos de aves que existem em uma determinada granja. Responda aos itens abaixo:
 - a) Estime a proporção de: frangos, gansos, marrecos e patos.
 - b) Proponha, baseado na abordagem probabilística, um teste assintótico para testar se as proporções do item a) são iguais, contra que elas são diferentes, justificando, adequadamente, seus desenvolvimentos e comentários.
8. Considere o arquivo “Empresas.csv”, relativo aos dados do livro Silva et al. (2021) (disponíveis no site do curso). Considere que ele representa uma população de empresas de interesse, das quais temos o total de funcionários e a renda média desses funcionários, por empresa. Responda aos itens abaixo:

- a) Estime o total de funcionários das empresas.
 - b) Estime a média salarial paga pelas empresas.
 - c) Estime a média salarial dos funcionários das empresas.
9. Com relação as Questões 6, 7 e 8, compare os resultados obtidos, com exatamente um (para cada questão) dentro da abordagem clássica (ou seja, escolha, para cada questão, exatamente um modelo probabilístico), justificando, adequadamente, seus desenvolvimentos, resultados e comentários. Suas conclusões eram esperadas?
10. Considere o conjunto de dados “tutorial” disponível no pacote do R “R2MLwiN” (veja a descrição das variáveis no respectivo help (?tutorial). Considere as variáveis “normexam”, “sex” (respostas), “school” (estratos) e “standlrt” (explicativa). Assuma que esses dados correspondes à uma população dividida nos mencionados estratos. Considerando, uma PA = AAS_s (ou seja, ignore os estratos), responda aos itens abaixo:
- a) Estime a média populacional da variável “normexam”, usando o estimador usual.
 - b) Estime a proporção populacional de estudantes do sexo feminino.
 - c) Usando o estimador razão, estime a média populacional da variável “normexam”.
 - d) Usando o estimador regressão, estime a média populacional da variável “normexam”.
 - e) Compare os resultados dos itens c) e d).
11. Repita o item 10 considerando um PA: AE_2 , comparando com os resultados obtidos na Questão 10.
12. Considere agora, que na Questão 10, as escolas são as unidades amostrais de interesse e que o objetivo é estimar o total de alunos da população. Para tal fim, utilize (e compare) os estimadores: expansão, razão e regressão (para estes dois últimos use a mesma covariável, de sua escolha, dentro do banco de dados).