

# Técnicas de Amostragem - 2022

## Lista de Exercícios

Kaique Matias de Andrade Roberto

4 de novembro de 2022

Estes são os Exercícios recomendados para a disciplina. Afim de que você possa extrair o maior proveito possível destes exercícios tenha em mente as seguintes observações:

- esta é a **única** lista de exercícios da disciplina toda;
- esta lista **contém** os exercícios que resolveremos em aula;
- as Seções estão nomeadas de acordo com as aulas (por exemplo, na Seção 10 estão os exercícios recomendados para a Aula 10);
- os exercícios que aparecem em aula estão marcados com (A);
- os exercícios com (\*) ou (\*\*) são exercícios que consideramos mais desafiadores.

## 0 Tipos de Variáveis, Estatística Descritiva, Principais Variáveis Aleatórias, Amostragem

**Exercício 0.1.** Qual é a diferença entre variáveis qualitativas e quantitativas?

**Exercício 0.2.** Classificar as variáveis a seguir:

- |  |  |
|--|--|
| a - Faturamento da empresa;                  | e - Distância percorrida em km;                  |
| b - Ranking de desempenho: bom, médio, ruim; | f - Casado: sim ou não;                          |
| c - Número de carros vendidos;               | g - Localização: Barueri ou Santana de Parnaíba; |
| d - Faixa de renda;                          | h - Municípios do Grande ABC.                    |

**Exercício 0.3** (A). Considere a tabela abaixo:

Nome	Idade (anos)	Peso (kg)	Altura (m)
Mariana	48	62	1,60
Luiz	54	84	1,76
Roberta	41	56	1,62
Leonardo	30	82	1,90
Melissa	28	54	1,68
Sandro	50	70	1,72

Calcule a média, variância e desvio-padrão das variáveis Idade, Peso e Altura.

**Exercício 0.4** (Resumo Estatística Descritiva). Faça um Resumo das Seções 2.1-2.4 do Livro [2].

**Exercício 0.5** (Resumo Probabilidade). Faça um Resumo das Seções 4.1-4.7 do Livro [2] ou do Capítulo 5 do Livro [3].

**Exercício 0.6.** Defina o conceito de variável aleatória. Você consegue descrever ao menos três variáveis aleatórias dentro do escopo da sua área de atuação?

**Exercício 0.7** (Resumo Variáveis Aleatórias). Faça um Resumo do Capítulo 5 do Livro [2].

**Exercício 0.8.** Faça uma tabela com a esperança, variância e distribuição acumulada das principais distribuições (vide slides 60 e 61 da Aula-00).

**Exercício 0.9** (Resumo Noções Básicas de Amostragem). Faça um Resumo do Capítulo 1 do Livro [1].

**Exercício 0.10.** Apresente uma questão ligada à sua área de interesse e que poderia ser respondida por um levantamento amostral. Discuta brevemente como realizar esse levantamento tendo como roteiro o Apêndice B do Livro [1].

## 1 Definições e Notações Básicas I: População, Amostra e Planejamento Amostral

**Exercício 1.1.** Para a situação que você apresentou no Exercício 0.10, descreva com as suas palavras a população, o elemento populacional e as característica de interesse.

**Exercício 1.2** (\*). Implemente em Python ou R funções para calcular total, média, variância, covariância, correlação e razão populacionais.

**Exercício 1.3** (A). Seja  $\mathcal{U} = \{1, 2, 3\}$  (vide planilha "aula-01-exemplo"). Considere as amostras

$$\begin{aligned}s_1 &= (1, 2) \\ s_2 &= (2, 1) \\ s_3 &= (1, 1, 3) \\ s_4 &= (3) \\ s_5 &= (2, 2, 1, 3, 2)\end{aligned}$$

Calcule  $f_i(s_j)$  e  $\delta_i(s_j)$ .

**Exercício 1.4** (A). Seja  $\mathcal{U} = \{1, 2, 3\}$  (vide planilha "aula-01-exemplo"). Considere as amostras

$$\begin{aligned}s_1 &= (1, 2) \\ s_2 &= (2, 1) \\ s_3 &= (1, 1, 3) \\ s_4 &= (3) \\ s_5 &= (2, 2, 1, 3, 2)\end{aligned}$$

Calcule  $n(s_j)$  e  $\nu(s_j)$ .

**Exercício 1.5** (A). Seja  $\mathcal{U} = \{1, 2, 3\}$  (vide planilha "aula-01-exemplo"). Descreva  $\mathcal{S}_2(\mathcal{U})$  e  $\mathcal{S}_3(\mathcal{U})$ .

**Exercício 1.6** (\*\*). Mostre que, para uma população  $\mathcal{U} = \{1, \dots, N\}$ , o número de amostras ordenadas de tamanho  $n$ , com reposição, é  $N^n$ , enquanto que, sem reposição, é dado pelo coeficiente binomial  $\binom{N}{n}$ .

**Exercício 1.7** (A). Considere  $\mathcal{U} = \{1, 2, 3\}$  (vide planilha "aula-01-exemplo") e a seguinte regra de sorteio:

- i - sorteia-se com igual probabilidade um elemento de  $\mathcal{U}$ , e anota-se a unidade sorteada;
- ii - este elemento é devolvido à população e sorteia-se um segundo elemento do mesmo modo.

Mostre que este é o mesmo plano amostral do Plano A (slide 45 da Aula-01).

**Exercício 1.8** (A). Seja  $\mathcal{U} = \{1, 2, 3\}$  (vide planilha "aula-01-exemplo"). Considere a regra "Sorteie uma unidade após a outra, repondo a unidade sorteada antes de sortear a seguinte, até o surgimento da unidade 2 ( $i = 2$ ) ou até que 3 unidades tenham sido sorteadas". Mostre que esta regra é equivalente ao Plano Amostral C (slide 47 da Aula-01).

**Exercício 1.9.** Considere  $\mathcal{U} = \{1, 2, 3\}$  (vide planilha "aula-01-exemplo") e a seguinte regra:

- i - sorteia-se um elemento  $\mathcal{U}$  com probabilidade proporcional ao número de trabalhadores;
- ii - sem repor o domicílio selecionado, sorteia-se um segundo também com probabilidade proporcional ao número de trabalhadores.

Mostre que esta regra é equivalente ao Plano Amostral D (slide 48 da Aula-01).

**Exercício 1.10.** Dê exemplos de amostras nos planos AAS, AE, AC, A2E e AS (se quiser, use os dados na planilha "dados-unificados" como inspiração).

## 2 Definições e Notações Básicas II: Estatísticas, Distribuições Amostrais, Estimadores

**Exercício 2.1.** Defina com as suas palavras os conceitos de estatística, distribuição amostral, estimador e estimativa.

**Exercício 2.2** (A). Com os dados/notações da planilha "aula-02-exemplo" calcule as distribuições amostrais da estatística  $r$  segundo os planos amostrais AASc e AASs.

**Exercício 2.3** (A). Com os dados/notações da planilha "aula-02-exemplo" calcule as distribuições amostrais da estatística  $\bar{f}$  segundo os planos amostrais AASc e AASs.

**Exercício 2.4** (A). Considere os dados na planilha "aula-02-exemplo" com dados amostrais

$$D = \begin{pmatrix} F_i \\ T_i \end{pmatrix} = \begin{pmatrix} 12 & 30 & 18 \\ 1 & 3 & 2 \end{pmatrix}, i \in \mathcal{U}$$

e plano amostral AASc. Agora considere as estatísticas  $r$  e  $\bar{f}$ . Calcule

$$E_{AASc}[r], E_{AASc}[\bar{f}], \text{Var}_{AASc}[r], \\ \text{Var}_{AASc}[\bar{f}], \text{Cov}_{AASc}[r, \bar{f}] \text{ e } \text{Corr}_{AASc}[r, \bar{f}].$$

**Exercício 2.5** (A). Considere os dados na planilha "aula-02-exemplo" com plano amostral AASc. Verifique que  $f_2$  e  $f_3$  tem a mesma distribuição que  $f_1$ ,  $\delta_2$  e  $\delta_3$  tem a mesma distribuição que  $\delta_1$ , e

$$E_A[f_1] = 2/3 \text{ e } E_A[\delta_1] = 5/9.$$

**Exercício 2.6** (A). Considere os dados na planilha "aula-02-exemplo" com plano amostral AASc. Verifique que  $r$  é um estimador viesado para  $R$ .

**Exercício 2.7** (A). Considere os dados na planilha "aula-02-exemplo" com plano amostral AASc. Verifique que  $\bar{f}$  é um estimador não-viesado para  $\mu_F$ .

**Exercício 2.8** (\*). Nosso curso não está focado na dedução (demonstração) dos resultados, mas sim na compreensão/uso dos mesmos. Ainda assim, algumas expressões/fórmulas podem ser úteis. Faça um resumo das fórmulas na Seção 2.6 do livro [1].

### 3 Amostragem Aleatória Simples Com Reposição (AASc)

**Exercício 3.1** (A). Seja  $Z \sim N(0, 1)$  a Normal Padrão. Calcule:

- |                    |   |
|--------------------|---|
| a - $P(Z < 1)$ ;   | e - $P(-1 < Z < 1)$ ;                   |
| b - $P(Z < 1.5)$ ; | f - $P(0.5 < Z < 1.7)$ ;                |
| c - $P(Z > 2)$ ;   | g - $P(Z < -0.7 \text{ ou } Z > 0.7)$ ; |
| d - $P(Z > 3)$ ;   | h - $P(Z < -0.4 \text{ ou } Z > 1)$ .   |

**Exercício 3.2** (A). Seja  $Z \sim N(0, 1)$  a Normal Padrão. Calcule:

- a -  $P(-1 < Z < 1)$ ;
- b -  $P(-2 < Z < 2)$ ;
- c -  $P(-3 < Z < 3)$ ;
- d -  $P(-4 < Z < 4)$ .

**Exercício 3.3** (A). Seja  $Z \sim N(8, 36)$  a Normal Padrão. Calcule:

- |                      |   |
|----------------------|---|
| a - $P(X < 0)$ ;     | f - $P(6 \leq X \leq 11)$ ;               |
| b - $P(X \leq 12)$ ; | g - $P(10 \leq X \leq 25)$ ;              |
| c - $P(X \leq 20)$ ; | h - $P(X \leq 3 \text{ ou } X \geq 9)$ ;  |
| d - $P(X \geq 2)$ ;  | i - $P(X \leq 9 \text{ ou } X \geq 20)$ . |
| e - $P(X \geq 5)$ ;  |   |

**Exercício 3.4** (A). Seja  $X \sim N(\mu, \sigma^2)$ . Calcule:

- a -  $P(-\sigma + \mu < Z < \sigma + \mu)$ ;
- b -  $P(-2\sigma + \mu < Z < 2\sigma + \mu)$ ;
- c -  $P(-3\sigma + \mu < Z < 3\sigma + \mu)$ ;
- d -  $P(-4\sigma + \mu < Z < 4\sigma + \mu)$ .

**Exercício 3.5** (A). Considere novamente os dados na planilha "aula-02-exemplo", e considere a variável renda familiar, onde o universo é  $\mathcal{U}$  e o parâmetro populacional é  $\mathbf{D} = (12, 30, 18)$ . Verifique o comportamento das estatísticas  $\bar{y}$  e  $s^2$  com relação as funções paramétricas  $\mu$  e  $\sigma^2$  de  $\mathbf{D}$  para o plano amostral AASc com  $n = 2$ .

**Exercício 3.6.** O que é o intervalo de confiança?

**Exercício 3.7 (A).** Uma máquina enche pacotes de café com uma variância igual a  $100g$ . Ela estava regulada para encher os pacotes com  $500g$ , em média. Agora, ela se desregulou, e queremos saber qual a nova média  $\mu$ . Uma amostra de 25 pacotes apresentou uma média igual a  $485g$ . Vamos construir um intervalo de confiança com 95% de confiança para  $\mu$ .

**Exercício 3.8 (A).** Considere a população de moradores de um condomínio ( $N = 540$ ). Deseja-se estimar a idade média dos condôminos. Com base em pesquisas passadas, pode-se obter a estimativa para  $\sigma^2$  de 463.32. Suponha que será retirada da população uma amostra segundo AASc. Admitindo que a diferença entre a média amostral e a verdadeira média populacional seja, no máximo, de 4 anos, com um nível de confiança de 95%, determine o tamanho da amostra a ser coletada.

**Exercício 3.9 (A).** Suponha que  $p = 30\%$  dos estudantes de uma escola sejam mulheres. Colhemos uma AAS de  $n = 10$  estudantes e calculamos

$$\hat{p} = \text{proporção de mulheres na amostra.}$$

Qual a probabilidade de que  $\hat{p}$  difira de  $p$  em menos de 0,01?

## 4 Exercícios, Amostragem Aleatória Simples Sem Reposição (AASs) I

**Exercício 4.1 (A).** Em uma população com  $N = 6$ , tem-se  $\mathbf{D} = (8, 2, 2, 11, 4, 7)$ . Um plano AASc de tamanho 2 foi adotado.

1. Encontre a distribuição de  $\bar{y}$  e mostre que  $E[\bar{y}] = \mu$ .
2. Encontre  $\text{Var}(\bar{y})$ .
3. Suponha que uma AASc de tamanho  $n = 10$  foi retirada da população com  $\bar{y} = 5.435$  e  $s^2 = 3.6$ . Encontre um intervalo de confiança para  $\mu$  com  $\alpha = 0.02$ .

**Exercício 4.2 (A).** Um plano AASc com  $n = 30$  foi adotado em uma área da cidade contendo 14848 residências. O número de pessoas por residência na amostra observada foi

$$\mathbf{d} = (5, 6, 3, 3, 2, 3, 3, 3, 4, 4, 3, 2, 7, 4, 3, 5, 4, 4, 3, 3, 4, 3, 3, 1, 2, 4, 3, 4, 2, 4).$$

- a - Encontre uma estimativa do número médio de pessoas por residência na população e uma estimativa para a variância da estimativa obtida.
- b - Encontre um intervalo de confiança para  $\mu$ .
- c - Suponha que seja de interesse uma estimativa duas vezes mais precisa que a obtida com a amostra acima. Qual o tamanho da amostra necessário para tal precisão?

**Exercício 4.3 (A).** Considere novamente os dados na planilha "aula-02-exemplo", e considere a variável renda familiar, onde o universo é  $\mathcal{U}$  e o parâmetro populacional é  $\mathbf{D} = (12, 30, 18)$ . Verifique o comportamento das estatísticas  $\bar{y}$  e  $s^2$  com relação as funções paramétricas  $\mu$  e  $\sigma^2$  de  $\mathbf{D}$  para o plano amostral AASs com  $n = 2$ .

## 5 Amostragem Aleatória Simples Sem Reposição (AASs) II

**Exercício 5.1** (A). Uma pesquisa amostral foi conduzida com o objetivo de se estudar o índice de ausência ao trabalho em um determinado tipo de indústria. Uma AAS sem reposição de mil operários de um total de 36 mil é observada com relação ao número de faltas não justificadas em um período de 6 meses. Os resultados obtidos foram:

Faltas:	0	1	2	3	4	5	6	7	8
Trabalhadores:	451	162	187	112	49	21	5	11	2

Para esta amostra tem-se que uma estimativa de  $\mu$  é dada por  $\bar{y} = 1.296$ . Observa-se também que  $s^2 = 2.397$ . Usando a aproximação normal, construa um intervalo de confiança para  $\mu$  com 95% de confiança.

**Exercício 5.2** (A). Considere a população dos operários faltosos do Exercício 5.1. Encontre  $n$  tal que  $B = 0.05$  e  $\gamma = 0.05$ .

**Exercício 5.3** (A). Considere a população dos operários faltosos dos exemplos anteriores. Suponha que até 3 faltas (3 dias) em 6 meses seja considerada aceitável.

a - Construa um intervalo de confiança para  $P$  com nível de confiança de 95%.

b - Calcule o tamanho da amostra com precisão  $B = 0.01$  e nível de confiança de 95%.

**Exercício 5.4.** Refaça o Exercício 4.2 agora considerando que a amostra foi tomada segundo o plano AASs.

**Exercício 5.5.** Faça um resumo com as principais propriedades/características da AASs e AASc.

## 6 Gerando Números Aleatórios

**Exercício 6.1.** O que são os geradores de números pseudo-aleatórios?

**Exercício 6.2.** Como se usa uma tabela de números pseudo-aleatórios para obtenção de uma amostra pseudo-aleatória?

**Exercício 6.3** (A). Há 90 alunos em um período de almoço, e 5 deles serão selecionados aleatoriamente para trabalhar na limpeza toda semana. Cada aluno recebe um número entre 0 e 90. Use uma tabela de números aleatórios para:

a - extrair uma amostra de tamanho 5 segundo o plano AASc;

b - extrair uma amostra de tamanho 5 segundo o plano AASs.

**Exercício 6.4** (A). Uma empresa de cereais está dando um prêmio em cada caixa de cereal e anuncia: "Colecione todos os 6 prêmios!" Cada caixa de cereal tem 1 prêmio, e cada prêmio tem a mesma probabilidade de aparecer em qualquer caixa. Caroline se pergunta quantas caixas leva, em média, para conseguir todos os 6 prêmios. Simule a quantidade de caixas necessárias usando ao menos duas linhas de uma tabela de números aleatórios.

**Exercício 6.5.** Considere os dados na planilha "tecnicas-amostragem-aula-06-exemplos.xlsx" na aba "precos-casas-sp".

a - Use uma tabela de números aleatórios para extrair uma amostra de tamanho 10 para AASc.

b - Calcule o preço médio dessa amostra e um intervalo de confiança para o preço médio da população na planilha.

c - Calcule o erro cometido ao substituir a média populacional pela média da sua amostra.

d - repita os passos (a)-(c) agora com AASs.

**Exercício 6.6.** Considere os dados na planilha "tecnicas-amostragem-aula-06-exemplos.xlsx" na aba "precos-casas-sp".

a - Use o Excel para extrair uma amostra de tamanho 200 para AASc.

b - Calcule o preço médio dessa amostra e um intervalo de confiança para o preço médio da população na planilha.

c - Calcule o erro cometido ao substituir a média populacional pela média da sua amostra.

d - repita os passos (a)-(c) agora com AASs.

Compare os resultados com os resultados obtidos no Exercício 6.5.

## 7 Comentários Pós N1 - Exercícios

Para esta aula não há novos exercícios recomendados.

## 8 Amostragem Estratificada I

**Exercício 8.1** (A). Considere uma pesquisa feita em uma população com  $N = 8$  domicílios, onde são conhecidas as variáveis renda domiciliar ( $Y$ ) e local do domicílio ( $W$ ), com os códigos  $A$  para região alta e  $B$  para região baixa. Tem-se então,

$$\mathcal{U} = \{1, 2, 3, 4, 5, 6, 7, 8\},$$



com

$$\mathbf{D} = \begin{pmatrix} \mathbf{y} \\ \mathbf{w} \end{pmatrix} = \begin{pmatrix} 13 & 17 & 6 & 5 & 10 & 12 & 19 & 6 \\ B & A & B & B & B & A & A & B \end{pmatrix}$$

Analise como se comportam os parâmetros de  $\mathbf{D}$  em função dos parâmetros das subpopulações determinadas pelos estratos A e B (utilize  $AASc(2)$  para avaliar os estimadores).

**Exercício 8.2** (A). Considere uma pesquisa feita em uma população com  $N = 8$  domicílios, onde são conhecidas as variáveis renda domiciliar ( $Y$ ) e local do domicílio ( $W$ ), com os códigos  $A$  para região alta e  $B$  para região baixa. Tem-se então,

$$\mathcal{U} = \{1, 2, 3, 4, 5, 6, 7, 8\},$$

divididas nos estratos

$$\mathbf{D}_1 = (13, 17, 6, 5) \text{ e } \mathbf{D}_2 = (10, 12, 19, 6).$$

Analise como se comportam os parâmetros de  $\mathbf{D}$  em função dos parâmetros das subpopulações determinadas pelos estratos 1 e 2 (utilize  $AASc(2)$  para avaliar os estimadores).

**Exercício 8.3** (A). Considere os dados:

Nome	Categoria	Nota	Nome	Categoria	Nota
Enedina	A	10	Leopoldina	A	3
Machado	A	4	Dandara	C	8
Luiz	A	5	Francisco	A	6
Marilena	B	3	Felipa	A	7
Clarice	B	9	Menininha	C	10
Heitor	B	2	Erenilton	C	9
Camargo	B	8	Vadinho	C	8
Rita	B	10	Jorge	C	3

- Calcule os parâmetros definidos na Seção 3 - Notações da Aula 08.
- Para os estratos A,B,C, realize a alocação  $AL_k$  com uma amostragem AASc com  $n_A = 2$ ,  $n_B = 3$  e  $n_C = 2$ . Para esta alocação, calcule  $\text{Var}_{AL_k}[\bar{y}_{es}]$ .
- Defina uma nova alocação  $AL_e$  de sua preferência e calcule  $\text{Var}_{AL_e}[\bar{y}_{es}]$ .
- Calcule o efeito do planejamento (EPA) entre as alocações  $AL_k$  e  $AL_e$ .

## Referências

- [1] Heleno Bolfarine and Wilton de Oliveira Bussab. *Elementos de amostragem*. Editora Blucher, 2005.
- [2] Luiz Paulo Fávero and Patrícia Belfiore. *Manual de análise de dados: estatística e modelagem multivariada com Excel® , SPSS® e Stata®*. Elsevier Brasil, 2017.
- [3] Pedro A Morettin and Wilton O Bussab. *Estatística básica*. Saraiva Educação, 2010.