

Estatística Avançada - 2022

Lista de Exercícios

Kaique Matias de Andrade Roberto

31 de outubro 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 Fundamentos: Tipos de Variáveis, Estatística Descritiva, 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 [1].

Exercício 0.5 (Resumo Amostragem). Faça um Resumo das Seções 6.1-6.4 do Livro [1].

Exercício 0.6 (A). Inspirado nos dados da planilha motocicletas.xlsx, realize as seguintes tarefas:

- a - Crie uma base de dados (uma planilha por exemplo) com ao menos 15 entradas, 3 variáveis quantitativas e 3 variáveis qualitativas.
- b - Realize a análise descritiva da sua base de dados.

1 Variáveis Aleatórias e Distribuição de Probabilidades I

Exercício 1.1 (Resumo Probabilidade). Faça um Resumo das Seções 4.1-4.7 do Livro [1] ou do Capítulo 5 do Livro [2].

Exercício 1.2. Defina o conceito de variável aleatória. Você consegue descrever ao menos três variáveis aleatórias dentro do escopo da sua área de atuação?

Exercício 1.3 (A). Considere o experimento de lançar uma certa moeda e observar se ocorre cara (k) ou coroa (c). Descreva o comportamento da variável

$$N = \{\text{Número de caras em dois lançamentos dessa moeda}\}.$$

Exercício 1.4 (A). Considere duas extrações, sem reposição, de uma urna contendo duas bolas brancas e três bolas vermelhas. Descreva o comportamento da variável

$$X = \{\text{Número de bolas vermelhas obtidas nas duas extrações}\}.$$

Exercício 1.5 (A). Considere duas extrações, com reposição, de uma urna contendo duas bolas brancas e três bolas vermelhas. Descreva o comportamento da variável

$$Y = \{\text{Número de bolas vermelhas obtidas nas duas extrações}\}.$$

Exercício 1.6 (A). Para as variáveis N, X, Y calcule:

- a - a distribuição de probabilidades;
- b - a distribuição acumulada;
- c - a esperança;
- d - a variância e o desvio-padrão.

Exercício 1.7 (A). Um dado não viciado é lançado, de modo que a variável aleatória X que representa o valor da face voltada para cima. Temos que X assume distribuição discreta de probabilidades. Calcule $E[X]$ e $\text{Var}[X]$.

Exercício 1.8 (A). A final do Masterchef Brasil ocorrerá entre os participantes G (de Greice Kelly) e M (de Michael Jaquisson). A variável aleatória X representa o vencedor. Sabe-se que a probabilidade de G ser vencedora é 0,60. Determine a distribuição de X , além da esperança e variância de X .

Exercício 1.9 (A). Determinada peça é produzida em uma linha de produção. A probabilidade de que a peça não tenha defeitos é de 99%. Se forem produzidas 30 peças, qual a probabilidade de que pelo menos 28 delas esteja em boas condições? Determine também a média e a variância da variável aleatória associada.

Exercício 1.10 (A). Seja X a variável aleatória que denota o número de clientes que chegam a um banco e suponha que X siga uma distribuição Poisson. Verifica-se que, em média, chegam 12 clientes por minuto. Calcule:

- a - probabilidade de chegada de 10 clientes no próximo minuto;
- b - probabilidade de chegada de 40 clientes nos próximos 5 minutos;
- c - média e variância de X .

Exercício 1.11. Uma moeda viciada tem probabilidade de cara igual a 0.4. Para dois lançamentos independentes dessa moeda, estude o comportamento da variável

$$X = \text{Número de caras}$$

e faça um gráfico da sua função de distribuição.

Exercício 1.12. Quais são as principais diferenças entre os modelos Bernoulli, Binomial e Poisson?

Exercício 1.13 (*). Uma moeda viciada tem probabilidade de cara igual a 0.3. Para três lançamentos independentes dessa moeda, estude o comportamento da variável

$$Y = \text{Número de coroas}$$

e faça um gráfico da sua função de distribuição.

Exercício 1.14 (A).

- a - Adapte as fórmulas da planilha "estatistica-avancada-aula-01-binomial-poisson.xlsx" para plotar (desenhar) as distribuições $b(2^k, 1/2^{k+1})$, $k = 1, 2, 3, 4$ ¹.
- b - Em seguida, plote o gráfico da distribuição Poisson(1/2).
- c - Use estes gráficos para dar uma justificativa empírica da convergência das binomiais $b(2^k, 1/2^{k+1})$ para a poisson Poisson(1/2).

Exercício 1.15 (**). Seja $X \sim \text{Bern}(p)$. Mostre que

$$E[X] = p \text{ e } \text{Var}(X) = p(1 - p).$$

Exercício 1.16 (**). Seja $X \sim b(n, p)$. Mostre que

$$E[X] = np \text{ e } \text{Var}(X) = np(1 - p).$$

2 Variáveis Aleatórias e Distribuição de Probabilidades II

Exercício 2.1 (A). Aponte as diferenças entre uma variável aleatória discreta e uma variável aleatória contínua.

Exercício 2.2 (A). Descreva o que é a "falta de memória" da Distribuição Geométrica.

Exercício 2.3 (A). Uma empresa fabrica determinado componente eletrônico, de modo que, ao final do processo, cada componente é testado, um a um. Suponha que a probabilidade de um componente eletrônico estar defeituoso seja de 0,05. Determine a probabilidade de que o primeiro defeito seja encontrado no oitavo componente testado. Calcule também o valor esperado e a variância da variável aleatória.

Exercício 2.4 (A). Uma urna contém 15 bolas, das quais 5 delas são vermelhas. São escolhidas 7 bolas ao acaso, sem reposição. Determine:

- a - A probabilidade de que exatamente duas bolas vermelhas sejam sorteadas.
- b - A probabilidade de que pelo menos duas bolas vermelhas sejam sorteadas.

¹Isto é, você deve plotar as binomiais $b(2, 0.25)$, $b(4, 0.125)$ e assim por diante...

c - O número esperado de bolas vermelhas sorteadas.

d - A variância do número de bolas vermelhas sorteadas.

Exercício 2.5 (A). A variável aleatória X representa o tempo de utilização dos caixas eletrônicos de um banco (em minutos) e segue uma distribuição uniforme no intervalo $[1, 5]$. Determine:

a - $P(X < 2)$;

d - $E(X)$;

b - $P(X > 3)$;

c - $P(3 < X < 4)$;

e - $\text{Var}(X)$.

Exercício 2.6 (A). Uma indústria fabrica lâmpadas especiais que ficam em operação continuamente. A empresa oferece a seus clientes a garantia de reposição, caso a lâmpada dure menos de 50 horas. A vida útil dessas lâmpadas é modelada através da distribuição Exponencial com parâmetro $1/8000$. Determine a proporção de trocas por defeito de fabricação.

Exercício 2.7. Faça um quadro "mnemônico" com a descrição para as distribuições de probabilidades que usamos neste curso.

3 A Distribuição Normal

Exercício 3.1. Por quê usamos tabelas para calcular probabilidades associadas à distribuição Normal?

Exercício 3.2 (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 3.3 (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 3.4 (A). Seja $Z \sim N(8, 36)$ a Normal Padrão. Calcule:

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

Exercício 3.5 (A). Os depósitos no Banco de Palmares durante o mês de Setembro são distribuídos normalmente, com média R\$10000,00 e desvio-padrão R\$1500,00. Um depósito é selecionado ao acaso dentre todos os referentes ao mês em questão. Encontre a probabilidade de que o depósito seja:

- a - R\$100000,00 ou menos;
- b - pelo menos R\$10000,00;
- c - um valor entre R\$12000,00 e R\$15000,00;
- d - maior que R\$20000,00.

Exercício 3.6 (A). Seja $X \sim N(\mu, \sigma^2)$ a Normal Padrão. 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 - $P(-4\sigma + \mu < Z < 4\sigma + \mu)$.

4 Fundamentos da Inferência Estatística: A Amostra

Exercício 4.1 (A). Quais são os objetivos da Inferência Estatística?

Exercício 4.2 (A). Dê exemplos de População e Amostra (de preferência, baseados na sua área de atuação).

Exercício 4.3 (A). Qual é a diferença entre parâmetro e estatística?

Exercício 4.4 (A). O que é Amostragem Aleatória Simples?

Exercício 4.5 (A). Descreva com as suas palavras o conceito de Distribuição Amostral.

Exercício 4.6 (A). Considere os dados no Slide 92 da Aula-04:

Unidade	Nome do Chefe	Sexo	Idade	Fumante	Renda Bruta Familiar	Número de Trabalhadores
1	Ada	0	20	0	12	1
2	Beto	1	30	1	30	3
3	Ema	0	40	1	18	2

Com o parâmetro

$$\theta = \text{Idade},$$

e plano amostral AAS com amostras de tamanho 2, calcule:

a - distribuição amostral da estatística \bar{X} (média amostral);

b - distribuição amostral da estatística S^2 (variância amostral).

Exercício 4.7 (A). Qual é a importância do Teorema do Limite Central?

5 Fundamentos da Inferência Estatística: Estimação

Exercício 5.1 (A). Qual é a diferença entre estimador e estimativa?

Exercício 5.2 (A). Dê exemplos de estimadores não-viesados.

Exercício 5.3 (A). Como devemos interpretar os intervalos de confiança?

Exercício 5.4 (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 μ . Uma amostra de 25 pacotes apresentou uma média igual a 485g. Construa um intervalo de confiança com 95% de confiança para μ .

Exercício 5.5 (A). Numa pesquisa de mercado, $n = 400$ pessoas foram entrevistadas sobre determinado produto, e 60% delas preferiram a marca A. Construa um intervalo de confiança com 95% de confiança para p com $\gamma = 0.95$.

Exercício 5.6 (A). Suponha que em $n = 400$ provas obtemos $k = 80$ sucessos. Obtenha um intervalo de confiança para p com $\gamma = 0.90$.

6 Testes de Hipóteses I

Exercício 6.1. Explique com as suas palavras o que são os Erro do Tipo I e Tipo II.

Exercício 6.2. Explique com as suas palavras o que é a hipótese nula e o que é a hipótese alternativa.

Exercício 6.3. Quais tipos de hipóteses alternativas estamos considerando e por quê?

Exercício 6.4. Explique com as suas palavras o que é a região crítica.

Exercício 6.5 (A). Uma máquina automática para encher pacotes de café enche-os segundo uma distribuição normal, com média μ e variância sempre igual a $400g^2$. A máquina foi regulada para $\mu = 500g$. Desejamos, periodicamente, colher uma amostra de 16 pacotes e verificar se a produção está sob controle, isto é, se $\mu = 500g$ ou não. Se uma dessas amostras apresentasse uma média $\bar{x} = 492g$, você pararia ou não a produção para regular a máquina?

7 Testes de Hipóteses II

Exercício 7.1 (A). A associação dos proprietários de indústrias metalúrgicas está muito preocupada com o tempo perdido com acidentes de trabalho, cuja média, nos últimos tempos, tem sido da ordem de 60 horas/homem por ano e desvio padrão de 20 horas/homem. Tentou-se um programa de prevenção de acidentes, após o qual foi tomada uma amostra de nove indústrias e medido o número de horas/homens perdidas por acidente, que foi de 50 horas. Você diria, no nível de 5%, que há evidência de melhoria?

Exercício 7.2 (A). O salário médio dos empregados das indústrias siderúrgicas de um país é de 2,5 salários mínimos, com um desvio padrão de 0,5 salários mínimos. Uma indústria é escolhida ao acaso e desta é escolhida uma amostra de 49 empregados, resultando um salário médio de 2,3 salários mínimos. Podemos afirmar que esta indústria paga salários inferiores à média nacional, com o nível de 5%?

Exercício 7.3 (A). Uma estação de televisão afirma que 60% dos televisores estavam ligados no seu programa especial da última segunda-feira. Uma rede competidora deseja contestar essa afirmação e decide usar uma amostra de 200 famílias para um teste. Qual deve ser o procedimento adotado para avaliar a veracidade da afirmação da estação?

Exercício 7.4 (A). O consumidor de um certo produto acusou o fabricante, dizendo que mais de 20% das unidades fabricadas apresentam defeito. Para confirmar sua acusação, ele usou uma amostra de tamanho 50, onde 27% das peças eram defeituosas. Mostre como o fabricante poderia refutar a acusação. Utilize um nível de significância de 10%.

Exercício 7.5 (A). Um fabricante garante que 90% dos equipamentos que fornece a uma fábrica estão de acordo com as especificações exigidas. O exame de uma amostra de 200 peças desse equipamento revelou 25 defeituosas. Teste a afirmativa do fabricante, nos níveis de 5% e 1%.

8 Testes de Hipóteses III

Exercício 8.1 (A). Uma companhia de serviços de ônibus intermunicipais planejou uma nova rota para servir vários locais situados entre duas cidades importantes. Um estudo

preliminar afirma que a duração das viagens pode ser considerada uma v.a. normal, com média igual a 300 minutos e desvio padrão 30 minutos. As dez primeiras viagens realizadas nessa rota apresentaram média igual a 314 minutos. Esse resultado comprova ou não o tempo médio determinado nos estudos preliminares?

Exercício 8.2 (A). Suponha que queiramos testar $H_0 : \mu = 50$ contra $H_1 : \mu > 50$, onde μ é a média de uma normal $N(\mu, 900)$. Extraída uma amostra de $n = 36$ elementos da população, obtemos $\bar{x} = 52$. Calcule o p -valor α do teste.

Exercício 8.3 (A). Os novos operários de uma empresa são treinados a operarem uma máquina, cujo tempo X (em horas) de aprendizado é anotado. Observou-se que X segue de perto a distribuição $N(25, 100)$. Uma nova técnica de ensino, que deve melhorar o tempo de aprendizado, foi testada em 16 novos empregados, o quais apresentaram 20,5 horas como tempo médio de aprendizado. Usando o p -valor, você diria que a nova técnica é melhor que a anterior?

Referências

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