Inteligência Artificial
IAD-001: Análise Estatística de Dados
Lista de Exercícios, Aula 05: Derivadas e Otimização
Larissa Driemeier • Ciclo 2025-1

Nomes:

---

1. O clássico "problema da Linda"(Tversky & Kahneman, 1983, p. 297): "Linda tem 31 anos, é solteira, extrovertida e muito inteligente. Ela é formada em filosofia. Como estudante, ela era bastante preocupada com questões de discriminação e justiça social, participava de manifestações e de demonstrações antinucleares."Baseando-se no texto, qual das seguintes alternativas é a mais provável:

   a) Linda é caixa de banco

   b) Linda é caixa de banco e ativista do movimento feminista

   c) Alternativas a e b têm a mesma probabilidade de ocorrer

   d) É impossível saber o que é mais provável com os dados fornecidos

   Solução. A opção correta é a alternativa *a)*. □

2. Considere um dado com seis faces (quatro faces VERDES e duas AZUIS). O dado será jogado 20 vezes e as sequências de verdes V e azuis A serão registradas. Supondo que você ganhe R$100 se a sequência que você escolher aparecer em sucessivas jogadas do dado, marque a sequência na qual você prefere apostar:

   a) $VAVAAAA$ *b)* $VAAAAAA$ *c)* $AVAAAA$ *d)* Todas as opções têm a mesma probabilidade

   Solução. A opção correta é a alternativa *c)*. □

3. Verifique se as frases são Verddeiras ou Falsas.

   a) Um jornal americano realizou, nos EUA, uma pesquisa com 800 pessoas divorciadas perguntando se elas desejavam se casar novamente. Foi encontrado que 58 % dos entrevistados não desejam um novo casamento. Os 800 divorciados constituem uma população.

   b) A Estatística Descritiva descreve quantitativamente ou resume características de uma coleção de informações.

   c) A Estatística Indutiva está relacionada com a análise e interpretação dos dados e previsões futuras.

   d) A Estatística Descritiva está baseada na teoria das probabilidades.

   a) FVVF *b)* VVFF *c)* FVFV *d)* VFVF

   Solução. A opção correta é a alternativa *a)*. □

4. Verifique em cada um dos itens abaixo e marque as variáveis como: qualitativa nominal (N), qualitativa ordinal (O), quantitativa discreta (D) e quantitativa contínua (C):

N O C D Cor dos olhos de uma população;
N O C D Temperatura de uma certa região, durante um certo período do ano;
N O C D Nível educacional;
N O C D Vida média de válvulas de descarga hidráulica;
N O C D Concentração de impurezas em uma amostra de leite, em mg por litro;
N O C D Tipo sanguíneo de uma população;
N O C D Religião de um indivíduo;
N O C D Tempo de voo entre duas cidades;
N O C D Procedência de cada candidato ao vestibular da USP em certo ano;
N O C D Quantidade de estudantes de uma disciplina;
N O C D Diâmetro de uma bola de futebol;
N O C D Número de parafusos refugados em determinado lote;
N O C D Tempo de reação de um indivíduo após submetido a certo estímulo;
N O C D Classificação de um café (bom, regular, ruim...);
N O C D Estado civil dos alunos de Estatística Básica deste ano;
N O C D Sexo de uma criança;
N O C D Quantidade de cômodos de uma residência;
N O C D Consumo de bebida em BH (bebe muito, bebe pouco, não bebe).

Solução. As opções corretas são: N C O C C N N C N D C D C O N N D O □

5. Diversas características são consideradas para o cálculo dos preços de venda e de aluguel de imóveis comerciais e residenciais. Ordene as variáveis abaixo na sequência: (1) Quantitativa Contínua, (2) Quantitativa Discreta, (3) Qualitativa Ordinal, (4) Qualitativa Nominal.

1 2 3 4 Região da cidade que se encontra o imóvel
1 2 3 4 Área do imóvel
1 2 3 4 Estado de conservação do imóvel
1 2 3 4 Número de quartos do imóvel

Solução. A sequência é: 4321 □

6. Em uma empresa, o salário médio dos estagiários é de R$500,00. Os salários médios pagos aos estagiários nível I e II são R$520,00 e R$420,00, respectivamente. Pode-se dizer, então, que:

a) O número de estagiários nível I é o dobro do número de estagiários nível II .
b) O número de estagiários nível I é o triplo do número de estagiários nível II.
c) O número de estagiários nível I é o quádruplo do número de estagiários nível II.
d) O número de estagiários nível II é o triplo do número de estagiários nível I.

e) O número de número de estagiários nível II é o quádruplo do número de estagiários nível I.

Dica:

$$\bar{x} = \frac{\bar{x}_1 n_1 + \bar{x}_2 n_2}{n_1 + n_2}$$

Solução. A opção correta é a alternativa *c)*. □

7. Que medida de tendência central é mais adequada para se analisar cada um dos casos abaixo? Use somente uma vez cada definição: média, mediana, moda. Justifique.

a) Uma grande construtora publica os preços de venda de seus apartamentos disponíveis na grande São Paulo.

b) Um estudante faz quatro provas em uma disciplina de cálculo. Suas notas são $8,8; 7,5; 9,5$ e $10,0$.

c) A tabela abaixo mostra a distribuição de frequência para os corredores da São Silvestre em 2024, considerando duas categorias, masculino e feminino.

Sexo	Frequência
Masculino	22.875
Feminino	14625

Solução. A sequência correta é: mediana, média, moda. □

8. Considere a idade de cada membro do grupo de avós e netos que brincam em uma praça,

1 1 1 2 2 2 3 3 3 3 3 50 52 55 55 57 58 59 59 60 61 65

Qual medida melhor representa a média de idade dos frequentadores do parque?

Solução.

A tabela abaixo resumo uma análise das medidas de tendência central para as idades dos frequentadores do parque.

Medida	**Cálculo**	**Valor**	**Observação**
Média	$\frac{\sum \text{Idades}}{22}$	$\approx 29{,}8$	Não reflete o perfil dos frequentadores do parque.
Mediana	Valor central da lista ordenada	26.5	Não reflete o perfil dos frequentadores do parque.
Moda	Valor mais frequente	3	Representa apenas as crianças, ignorando adultos/idosos.

Da tabela acima, conclui-se que, para determinar a melhor medida de tendência central que representa a *média de idade* dos frequentadores do parque, vamos analisar os dados divididos em *três grupos*:

Crianças (1 a 3 anos): idades: $1, 1, 1, 2, 2, 2, 3, 3, 3, 3, 3$. São 11 pessoas (50 % do total), resultando em uma média de $\approx 2{,}18$ anos;

Adultos (50 a 59 anos): idades: $50, 52, 55, 55, 57, 58, 59, 59$. São 8 pessoas (36.4 % do total), resultando em uma média de $\approx$ 56,75 anos;

Idosos (60+ anos): idades: $60, 61, 65$. São 3 pessoas (13.6 % do total), resultando em uma média de $\approx$ 62,0 anos.

Dependendo do contexto e objetivo da pesquisa, agrupar os frequentadores em dois grandes grupos: crianças e adultos. □

9. Se o salário médio de todos os empregados homens da companhia A excede o salário médio de todos os empregados homens da companhia B, e se o salário médio de todas as empregadas mulheres da companhia A excede o salário médio de todas as empregadas mulheres da companhia B, decorre daí que o salário médio da totalidade dos empregados da companhia A excede o salário médio da totalidade dos empregados da companhia B? Explique sua resposta.

Solução.

A resposta é não, pois depende do número de Homens e Mulheres em cada empresa. Por exemplo, a Empresa A: 4 homens ganham 10 reais e 1000 mulheres ganham 2 reais. $\bar{x}_A = 2040/1004$; Empresa B: 1000 homens ganham 9 reais e 4 mulheres ganham 1 reais. $\bar{x}_B = 9004/1004$. □

10. Durante as três semanas antes do Natal, 12 pessoas fizeram compras, em média, em 5,75 lojas de roupas. É possível que ao menos sete delas tenham feito compras em pelo menos 10 lojas?

Solução. O tamanho da amostra é dado por: $n = 5,75 \times 12 = 69$. Portanto, o número total de lojas em que as 12 pessoas fizeram compras é 69. Porém, se 7 fizerem compras em 10, tem-se já 70 lojas. Portanto, não é possível que ao menos sete delas tenham feito compras em pelo menos 10 lojas. □

11. Ache os quartis dos seguintes conjuntos de dados:

$$A = \{3 \quad 4 \quad 6 \quad 8 \quad 11 \quad 14 \quad 16 \quad 17 \quad 20 \quad 21 \quad 23 \quad 24\}$$
$$B = \{3 \quad 7 \quad 8 \quad 5 \quad 12 \quad 14 \quad 21 \quad 13 \quad 18\}$$
$$C = \{19 \quad 26 \quad 25 \quad 37 \quad 32 \quad 28 \quad 22 \quad 23 \quad 29 \quad 34 \quad 39 \quad 31\}$$

Solução. Para os conjuntos de dados fornecidos, os quartis são:

Conjunto $A = \{3, 4, 6, 8, 11, 14, 16, 17, 20, 21, 23, 24\}$

$$\begin{aligned} Q_1 &= \frac{6+8}{2} = 7, \\ Q_2 \text{ (Mediana)} &= \frac{14+16}{2} = 15, \\ Q_3 &= \frac{20+21}{2} = 20{,}5. \end{aligned}$$

Conjunto $B = \{3, 5, 7, 8, 12, 13, 14, 18, 21\}$

$$\begin{aligned} Q_1 &= \frac{5+7}{2} = 6, \\ Q_2 \text{ (Mediana)} &= 12, \\ Q_3 &= \frac{14+18}{2} = 16. \end{aligned}$$

Conjunto $C = \{19, 22, 23, 25, 26, 28, 29, 31, 32, 34, 37, 39\}$

$$Q_1 = \frac{23+25}{2} = 24,$$
$$Q_2 \text{ (Mediana)} = \frac{28+29}{2} = 28{,}5,$$
$$Q_3 = \frac{32+34}{2} = 33.$$

Resumo em tabela:

Conjunto	Q_1	Q_2 (Mediana)	Q_3
A	7	15	20,5
B	6	12	16
C	24	28.5	33

□

12. Construa o diagrama de caixa dos 40 dados a seguir

59	60	61	62	62	63	63	64
64	64	65	65	65	65	65	65
65	65	65	66	66	67	67	68
68	69	70	70	70	70	70	71
71	72	72	73	74	74	75	77

Solução.

- **Tamanho do conjunto:** $n = 40$ (par).
- **Mediana (Q_2):**

$$Q_2 = \frac{20^\circ \text{ valor} + 21^\circ \text{ valor}}{2} = \frac{65+66}{2} = 65{,}5$$

- **Primeiro Quartil (Q_1):**

$$\text{Primeira metade: } 59,\ 60,\ \ldots,\ 66 \quad (20 \text{ valores})$$
$$Q_1 = \frac{10^\circ \text{ valor} + 11^\circ \text{ valor}}{2} = \frac{64+65}{2} = 64{,}5$$

- **Terceiro Quartil (Q_3):**

$$\text{Segunda metade: } 66,\ 67,\ \ldots,\ 77 \quad (20 \text{ valores})$$
$$Q_3 = \frac{10^\circ \text{ valor} + 11^\circ \text{ valor}}{2} = \frac{70+70}{2} = 70{,}0$$

Quartil	**Cálculo**	**Valor**
Q_1	$\frac{64+65}{2}$	64,5
Q_2	$\frac{66+66}{2}$	66
Q_3	$\frac{70+70}{2}$	70,0

Cálculo da IQR:

$$\text{IQR} = Q_3 - Q_1 = 70,0 - 64,5 = \boxed{5,5}$$

Portanto, o valor mínimo é:

$$Min = Q2 - 1,5IQR = 66 - 1,5 \times 5,5 = 57,75$$

e o máximo:

$$Max = Q2 + 1,5IQR = 66 + 1,5 \times 5,5 = 74,25$$

□

13. Calcule a covariância e correlação entre:

$$x : \{2,1,\ 2,5,\ 3,6,\ 4,0\}$$
$$y : \{8,\ 10,\ 12,\ 14\}$$

Solução.

Calcula-se, inicialmente, as médias dos dois conjuntos de dados:

$$\bar{x} = \frac{2,1 + 2,5 + 3,6 + 4,0}{4} = \frac{12,2}{4} = 3,05$$
$$\bar{y} = \frac{8 + 10 + 12 + 14}{4} = \frac{44}{4} = 11$$

O cálculo da covariância é dado por

$$\text{Cov}(x, y) = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{n - 1}$$

x_i	y_i	$x_i - \bar{x}$	$y_i - \bar{y}$	$(x_i - \bar{x})(y_i - \bar{y})$
2.1	8	$-0,95$	-3	2,85
2.5	10	$-0,55$	-1	0,55
3.6	12	0,55	1	0,55
4.0	14	0,95	3	2,85

$$\text{Cov}(x, y) = \frac{2,85 + 0,55 + 0,55 + 2,85}{3} = \frac{6,8}{3} \approx 2,267$$

A correlação é dada por:

$$r = \frac{\text{Cov}(x, y)}{\sigma_x \cdot \sigma_y}$$

Portanto, é necessário calcular o desvio padrão de cada conjunto de dados, de modo que:

$$\sigma_x = \sqrt{\frac{\sum(x_i - \bar{x})^2}{n - 1}} = \sqrt{\frac{0,9025 + 0,3025 + 0,3025 + 0,9025}{3}} \approx 0,896$$
$$\sigma_y = \sqrt{\frac{\sum(y_i - \bar{y})^2}{n - 1}} = \sqrt{\frac{9 + 1 + 1 + 9}{3}} \approx 2,582$$

Portanto:

$$r = \frac{2,267}{0,896 \cdot 2,582} \approx \frac{2,267}{2,315} \approx 0,979$$

Dessa forma,

$$\boxed{\text{Covariância} \approx 2,267 \quad \text{e} \quad \text{Correlação} \approx 0,979}$$

□