



CE009- Introdução à Estatística

Lista 2

Exercício 1

Classifique cada uma das variáveis abaixo em qualitativa (nominal / ordinal) ou quantitativa (discreta / contínua).

- (a) Ocorrência de hipertensão pré-natal em grávidas, com mais de 35 anos, em que as possíveis respostas são *sim* ou *não*.
- (b) Intenção de voto em que as possíveis respostas são os nomes dos candidatos.
- (c) Perda de peso (em gramas) de maratonistas na Corrida de São Silvestre.
- (d) Grau de cansaço (leve, moderado ou forte) dos maratonistas na Corrida de São Silvestre.
- (e) Nível de satisfação da população com relação ao trabalho do prefeito (valores de 0 a 5, com 0 indicando totalmente insatisfeito e 5 totalmente satisfeito).

Para ver a resposta clique aqui: 1

Exercício 2

Um grupo de estudantes foi submetido a um teste de matemática resultando em:

Nota	[0; 2)	[2; 4)	[4; 6)	[6; 8)	[8; 10)
Frequência	14	28	27	11	4

Se a nota mínima para aprovação é 5, qual será a porcentagem de aprovados? Foi necessária alguma suposição adicional?

Para ver a resposta clique aqui: 2

Exercício 3

A pulsação (em batimentos por minuto) de 10 estudantes no início de uma prova de Estatística foram as seguintes: 80, 91, 84, 86, 93, 88, 80, 89, 85, 86. Calcule a média e a variância desse conjunto de dados.

Para ver a resposta clique aqui: [3](#)

Exercício 4

Responda certo ou errado, justificando:

- (a) Suponha duas amostras de uma mesma população, sendo uma de tamanho 100 e outra de tamanho 200. Então, a amostra de tamanho maior é mais representativa da população.
- (b) Duas variáveis com *box-plot* iguais não podem ter valores diferentes.

Para ver a resposta clique aqui: [4](#)

Exercício 5

Discuta quais medidas de posição seriam mais adequadas para os conjuntos de dados abaixo. Comente suas escolhas.

- (a) Estão disponíveis dados mensais sobre a incidência de envenenamento por picada de cobra. Deseja-se planejar a compra mensal de antídoto.
- (b) O número diário de usuários, entre 17 e 19 horas, de determinada linha de ônibus foi anotado. Pretende-se utilizar essa informação para dimensionar a frota em circulação.
- (c) Um fabricante de baterias deseja divulgar a durabilidade do seu produto. Para tanto, coletou informação sobre a duração de 100 de suas baterias.
- (d) Num voo internacional, uma companhia serve dois tipos de pratos de jantar: peixe ou frango. Um banco de dados contém os pedidos feitos nos últimos 200 voos. Pretende-se planejar o número de cada tipo a ser colocado à disposição dos passageiros.

Para ver a resposta clique aqui: [5](#)

Exercício 6

Uma pesquisa com usuários de transporte coletivo na cidade de São Paulo avaliou, entre os passageiros, o número de meios de transporte diferentes (entre ônibus, metrô e trem) que eles utilizam diariamente. O resultado foi o seguinte:

2, 3, 2, 1, 2, 1, 2, 1, 2, 3, 1, 1, 1, 2, 2

3, 1, 1, 1, 1, 2, 1, 1, 2, 2, 1, 2, 1, 2, 3

- (a) Organize os dados em uma tabela de frequência.
- (b) Admitindo que essa amostra represente bem o comportamento do usuário paulistano, você acha que a porcentagem dos usuários que utilizam mais de um tipo de transporte é grande?

Para ver a resposta clique aqui: 6

Exercício 7

Um novo medicamento para cicatrização está sendo testado e um experimento foi feito para estudar o tempo (em dias) de completo fechamento dos cortes provenientes da cirurgia. A amostra de trinta cobaias forneceu os seguintes valores:

17, 16, 15, 17, 14, 17, 16, 16, 17, 15, 18, 14, 17, 15, 15

14, 15, 16, 17, 18, 18, 17, 15, 16, 14, 18, 18, 16, 15, 14

- (a) Organize uma tabela de frequência.
- (b) Que porcentagem das observações estão abaixo de 16 dias?
- (c) Classifique como *rápida* as cicatrizações iguais ou inferiores a 15 dias e como *lenta* as demais. Indique as porcentagens para cada classificação.

Para ver a resposta clique aqui: 7

Exercício 8

Um grupo de pedagogos estuda a influência da troca de escolas no desempenho de alunos. Durante o levantamento, foi anotado o número de escolas cursadas pelos alunos participantes.

Nº de escolas cursadas	1	2	3	4	5
Frequência	46	57	21	15	4

- (a) Qual é a porcentagem dos alunos que cursaram mais de uma escola?
- (b) Classifique os alunos em dois grupos segundo a rotatividade:
- *Alta*: para alunos com mais de 2 escolas;
 - *Baixa* para os demais.

Obtenha a tabela de frequência dessa nova variável.

Para ver a resposta clique aqui: 8

Exercício 9

Os alunos de Educação Física foram submetidos a um teste de resistência, avaliando quantos quilômetros conseguiam correr sem parar. Os dados estão apresentados a seguir:

Faixas	[0; 4)	[4; 8)	[8; 12)	[12; 16)	[16; 20)
Frequência	438	205	125	22	9

- (a) Qual é a variável em estudo?
- (b) Construa o histograma e o *box-plot*.

Para ver a resposta clique aqui: 9

Exercício 10

Vinte e cinco residências foram sorteadas e visitadas por um entrevistador que, entre outras questões, perguntou sobre o número de televisores. Os dados foram os seguintes:

2, 2, 2, 3, 1, 2, 1, 1, 1, 1, 0, 1, 2, 2, 2, 2, 3, 1, 1, 3, 1, 2, 1, 0, 2.

Organize os dados numa tabela de frequências e determine as diversas medidas de posição.

Para ver a resposta clique aqui: 10

Exercício 11

Você está indeciso em comprar uma televisão. Diante disso, decide avaliar algumas estatísticas fornecidas pelo fabricante sobre a duração dos aparelhos (em horas).

Marca da TV	GA	FB	HW
Média	8000	8200	8000
Mediana	8000	9000	7000
Desvio Padrão	600	1500	2500

Com qual marca você ficaria? Justifique.

Para ver a resposta clique aqui: 11

Exercício 12

Foram anotados os níveis de colesterol (em mg/100ml) de trinta pacientes. As medidas se referem a homens entre 40 e 60 anos de idade que foram à clínica fazer um *check up*.

Paciente	1	2	3	4	5	6	7	8	9	10
Colesterol	160	160	161	163	167	170	172	172	173	177
Paciente	11	12	13	14	15	16	17	18	19	20
Colesterol	178	181	181	182	185	186	194	197	199	203
Paciente	21	22	23	24	25	26	27	28	29	30
Colesterol	203	205	206	206	208	209	211	214	218	225

- (a) Calcule a média, a moda e a variância a partir da tabela de dados brutos.
- (b) Organize uma tabela de frequência com faixas de tamanho de 10 a partir de 160.
- (c) Refaça o item (a) usando a tabela de frequência obtida em (b).
- (d) Comente as diferenças encontradas entre os valores das medidas calculadas em (a) e (c).

Para ver a resposta clique aqui: 12

Exercício 13

Certa cidade com uma população de 150000 habitantes possui 3 jornais: I, II e III. As proporções de moradores que leem esses jornais são as seguintes:

$$\begin{array}{lll} \text{I} = 15\%; & \text{I e II} = 8\%; & \text{I, II e III} = 1\%; \\ \text{II} = 24\%; & \text{I e III} = 2\%; & \\ \text{III} = 5\%; & \text{II e III} = 4\%; & \end{array}$$

1. Determine o número de pessoas que leem apenas um jornal. Justifique sua resposta.
2. Quantas pessoas leem pelo menos dois jornais? Justifique sua resposta.

Para ver a resposta clique aqui: [13](#)

Exercício 14

Com dados do último censo, a assistente social de um Centro de Saúde constatou que 20% das famílias têm um filho, 30% têm dois filhos, 35% têm três, e as restantes se dividem igualmente entre quatro, cinco ou seis filhos.

1. Faça a distribuição de frequência do número de filhos.
2. Encontre a média do número de filhos.

Para ver a resposta clique aqui: [14](#)

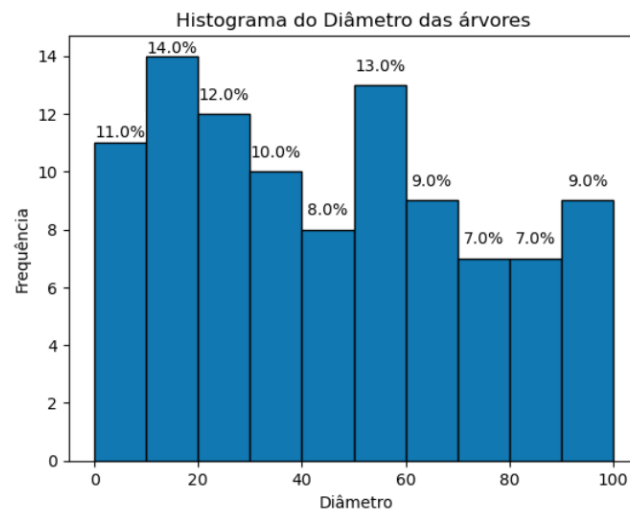
Exercício 15

Em um estudo sobre hábitos religiosos de pessoas que vivem em uma cidade, 500 moradores foram questionadas sobre frequentar à igreja regularmente. Desses, 240 eram homens, dos quais 40 disseram que frequentam um culto regularmente, enquanto entre as mulheres, 80 disseram que frequentam um culto regularmente. Construa uma tabela de frequência que contemple as variáveis sexo e hábito de frequentar a igreja.

Para ver a resposta clique aqui: [15](#)

Exercício 16

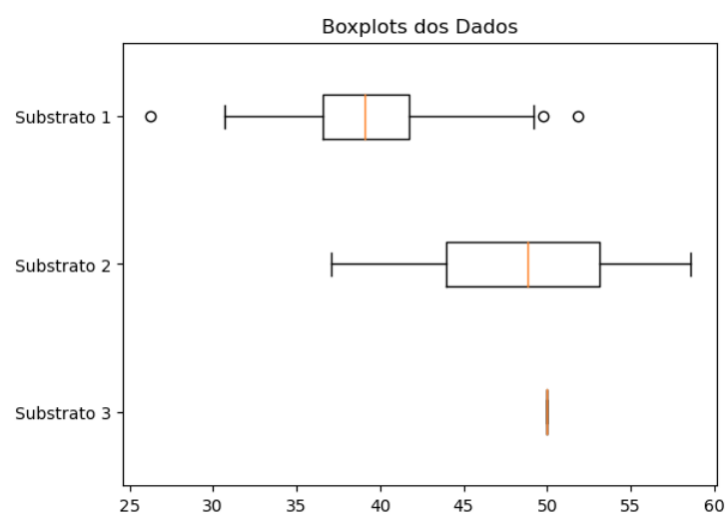
Um grupo está conduzindo um estudo sobre a distribuição do diâmetro das árvores em determinada área. O histograma abaixo fornece o resultado desta pesquisa. Encontre a mediana desses dados.



Para ver a resposta clique aqui: [16](#)

Exercício 17

Imagine que se esteja analisando o crescimento de uma espécie de planta em três diferentes condições de substrato. Os resultados abaixo referem-se à altura (em cm) de 50 unidades em cada um dos tratamentos. Analise cada um dos box-plots.



Para ver a resposta clique aqui: [17](#)

Respostas

Resposta do exercício 1

- (a) Qualitativa nominal.
- (b) Qualitativa nominal.
- (c) Quantitativa contínua.
- (d) Qualitativa ordinal.
- (e) Qualitativa ordinal.

Resposta do exercício 2

A porcentagem de aprovados é 34%. Neste caso, assumiu-se que os indivíduos com nota entre $[4; 6)$ estavam igualmente distribuídos no intervalo.

Resposta do exercício 3

Utilizando as expressões para dados não agrupados, temos: $\bar{x}_{obs} = 86.20$; $var_{obs} = 16.36$.

Resposta do exercício 4

- (a) Errado. A representatividade da população não necessariamente depende do tamanho da amostra. A forma de amostragem, por exemplo, é mais importante.
- (b) Errado. Para a construção do *box-plot* utilizamos medidas entre quantis, mas os dados que as geram não necessariamente são iguais.

Resposta do exercício 5

Este exercício é para possibilitar discussão sobre as várias formas de resumir a informação e, portanto, não existe uma resposta correta.

- (a) Considerando o valor mediano, ele deve ser suficiente em cerca de 50% dos meses. Se houver meses com um grande número de acidentes por razões sazonais, como colheitas ou enchentes, a média será afetada. É importante verificar se a variabilidade dos dados é grande, pois isso pode influenciar a escolha da medida de posição mais adequada.

- (b) Para dimensionar a frota de ônibus, a decisão pode ser tomada considerando o ponto de vista do usuário ou da empresa. No primeiro caso, deseja-se mais espaço e menos superlotação, então a mediana pode ser uma boa escolha, pois reflete um valor mais típico, menos afetado por picos extremos. No ponto de vista da empresa, a média diária de passageiros pode ser usada para calcular a frota necessária, dividindo-se pelo número de assentos ou pela capacidade total dos ônibus.
- (c) A duração média da bateria é provavelmente a melhor medida a ser considerada, pois representa o desempenho esperado para um consumidor comum. No entanto, se a moda indicar uma longa duração, pode ser interessante utilizá-la para fins publicitários, destacando um valor típico que seja atraente ao público.
- (d) Pode-se definir uma variável quantitativa indicando a proporção de passageiros que pedem frango (ou peixe) em cada voo. A moda dessa variável, multiplicada pelo total de passageiros, poderia ser usada para determinar a quantidade de pratos a ser disponibilizada. No entanto, se houver grande variabilidade nos pedidos, pode ser prudente considerar um intervalo de segurança para evitar escassez de um dos pratos.

Resposta do exercício 6

(a)

	Número de transportes			Total
	1	2	3	
Frequência	14	12	4	30
Frequência relativa	0.47	0.40	0.13	1

- (b) Como mais da metade dos usuários depende de mais de um meio de transporte, pode-se dizer que a porcentagem dos usuários que utilizam múltiplos meios de transporte é considerável. Isso pode indicar a necessidade de integração eficiente entre os diferentes modais na cidade de São Paulo.

Resposta do exercício 7

(a)

	Cicatrização					Total
	14	15	16	17	18	
Frequência	5	7	6	7	5	30
Frequência relativa	0.167	0.233	0.2	0.233	0.167	1

(b) 40%.

(c) Cicatrização rápida = 40% e lenta = 60%.

Resposta do exercício 8

(a) 67%

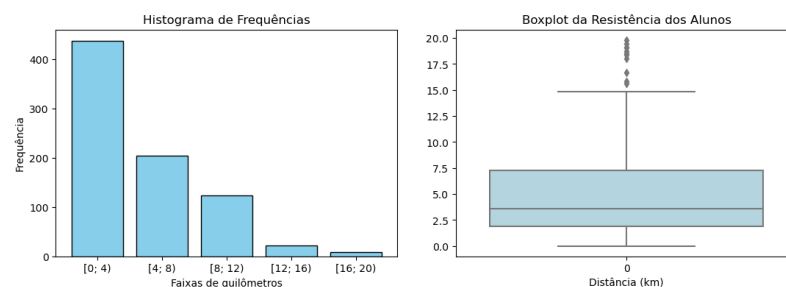
(b)

Rotatividade	Baixa	Alta
Frequência	103	40

Resposta do exercício 9

(a) Quilômetros percorridos sem parar.

(b) Para o *box-plot* use mínimo = 0, $Q_1 = 1.82$; $Q_2 = 3.65$; $Q_3 = 7.13$; máximo = 20.



Resposta do exercício 10

Nº de TVs	0	1	2	3	total
Frequência	2	10	10	3	25

- $\bar{x}_{obs} = 1.56$;
- $md_{obs} = 13^o$ termo em ordem crescente = 2;
- $mo_{obs} = 1$ ou 2.

Resposta do exercício 11

As médias são similares. A mediana de FB é mais alta, o que é um fator positivo. Por outro lado, HW tem a menor mediana e, portanto, essa marca deve ser desconsiderada. Notemos que o desvio padrão de FB é duas vezes e meia maior do que o de GA. Como GA tem mediana não muito baixa e pouca variabilidade, parece ser a melhor opção. Portanto, recomendamos comprar a marca GA.

Resposta do exercício 12

- (a)
- $\bar{x}_{obs} = 188.87$;
 - $mo_{obs} = 160, 172, 181, 203$ ou 206
 - $var_{obs} = 357.09$.

(b)

Colesterol	[160; 170)	[170; 180)	[180; 190)	[190; 200)	[200; 210)	[210; 220)	[220; 230)
Frequência	5	6	5	3	7	3	1

(c) Com a tabela de frequência no item anterior e usando o ponto médio da faixa obtemos:

- $\bar{x}_{obs} = 189.67$;
- $mo_{obs} = 205$.
- $var_{obs} = 311.56$.

- (d) Os valores obtidos no item (a) são mais precisos. Entretanto, para um grande número de dados, o cálculo fica muito trabalhoso para ser feito sem o auxílio de um computador. Notamos que, mesmo a variável sendo quantitativa discreta, a organização em faixas é conveniente e os valores encontrados nas duas formas não são muito diferentes.

Resposta do exercício 13

- (a) O número de pessoas que leem apenas um jornal é dado por:

$$\begin{aligned}|I \text{ apenas}| &= |I| - (|I \cap II| + |I \cap III| - |I \cap II \cap III|) \\&= 15\% - (8\% + 2\% - 1\%) \\&= 6\%\end{aligned}$$

$$\begin{aligned}|II \text{ apenas}| &= |II| - (|I \cap II| + |II \cap III| - |I \cap II \cap III|) \\&= 24\% - (8\% + 4\% - 1\%) \\&= 13\%\end{aligned}$$

$$\begin{aligned}|III \text{ apenas}| &= |III| - (|I \cap III| + |II \cap III| - |I \cap II \cap III|) \\&= 5\% - (2\% + 4\% - 1\%) \\&= 0\%\end{aligned}$$

Em números absolutos:

$$\begin{aligned}|I \text{ apenas}| &: 6\% \times 150000 = 9000 \\|II \text{ apenas}| &: 13\% \times 150000 = 19500 \\|III \text{ apenas}| &: 0\% \times 150000 = 0\end{aligned}$$

- (b) Número de pessoas que leem pelo menos dois jornais:

$$|I \cap II| + |I \cap III| + |II \cap III| - 2|I \cap II \cap III|$$

Substituindo os valores:

$$8\% + 2\% + 4\% - 2 \times 1\% = 14\% - 2\% = 12\%$$

Em números absolutos:

$$12\% \times 150000 = 18000$$

Resposta do exercício 14

(a) A distribuição de frequência do número de filhos é por:

Número de Filhos	Proporção (%)
1	20
2	30
3	35
4	5
5	5
6	5

(b) Agora, calculando a média do número de filhos:

$$\begin{aligned}\bar{x} &= (1 \times 0.20) + (2 \times 0.30) + (3 \times 0.35) + (4 \times 0.05) + (5 \times 0.05) + (6 \times 0.05) \\ &= 2.6\end{aligned}$$

Resposta do exercício 15

Sexo	Frequenta Igreja	Não Frequenta	Total
Homens	40	200	240
Mulheres	80	180	260
Total	120	380	500

Resposta do exercício 16

Os dados extraídos do histograma:

Intervalo (Diâmetro)	Frequência (f)	Frequência Acumulada (F)
0 - 10	11	11
10 - 20	14	25
20 - 30	12	37
30 - 40	10	47
40 - 50	8	55
50 - 60	13	68
60 - 70	9	77
70 - 80	7	84
80 - 90	7	91
90 - 100	9	100

Assim, dos 8% referente ao intervalo $[40 - 50[$, vamos fazer uma regra de três com 3%.

$$\begin{aligned}Md &= 40 + \left(\frac{50 - 47}{8} \right) \times 10 \\ &= 43.75\end{aligned}$$

Resposta do exercício 17

A análise dos **boxplots** apresentados permite identificar algumas características do crescimento da planta em cada uma das três condições de substrato:

Substrato 1

- A mediana está próxima de **40 cm**, indicando que metade das plantas cresceu abaixo desse valor e metade acima.
- A dispersão dos dados é relativamente baixa, sugerindo uma variação pequena nas alturas das plantas.

- Há **outliers** abaixo e acima dos limites, o que indica algumas plantas que cresceram significativamente menos ou mais do que a maioria.

Substrato 2

- A mediana está próxima de **50 cm**, superior à do Substrato 1, sugerindo um crescimento maior.
- Porém, a dispersão dos dados é maior do que no Substrato 1, evidenciado pelo comprimento da caixa e dos “bigodes”.

Substrato 3

- A mediana está próxima de **50 cm**, semelhante ao Substrato 2.
- O boxplot indica **uma variação muito baixa** nos valores da altura, pois a caixa e os “bigodes” são muito pequenos, o que sugere que a maioria das plantas cresceu quase na mesma altura.