

LISTA DE EXERCÍCIOS 1

Componente Curricular: Estatística I

1. Defina com suas próprias palavras o que é estatística descritiva e estatística inferencial.
2. Dê um exemplo para cada tipo de variável: discreta, contínua, ordinal e nominal (de preferência diferente do exemplificado em sala de aula).
3. Uma pessoa que ordena uma lista de cidades segundo a limpeza das ruas, do mais limpo para o mais sujo, está operando no nível de medida:
 - a. nominal.
 - b. ordinal.
4. Determine se os valores descritos abaixo são de um conjunto de dados discreto ou contínuo.
 - a. Um estudante de estatística obtém dados amostrais e encontra que o consumo de lixo é em média de 1 kg por pessoa por dia.
 - b. Quando 19.218 amostras de galões de água foram testadas, verificou-se que 10.322 delas estavam impróprias para o consumo humano.
 - c. Em uma pesquisa com 1059 adultos, verificou-se que 39% deles não sabiam explicar a função das matas ciliares.
5. Construa uma apresentação de ramo-e-folhas para os dados médios de umidade relativa do ar (%) coletados durante 16 dias do mês de maio na cidade de Cuiabá, e faça um comentário sobre a distribuição dos dados.

57	68	70	65	82	83	72	75
80	64	76	72	75	78	58	85

6. Os dados que seguem são (a) pesos de lixo (classificado como plástico, unidade: gramas) descartados em uma semana por domicílio e (b) pesos de lixo (classificado como papel, unidade: gramas) descartados em uma semana por domicílio.

a.	605	606	405	406	602	503	503	404	400	509
	208	208	509	509	100	607	609	304	507	400
	408	408	501	501	502	100	203	307	609	507

- b.** 1500 2184 1050 2500 1058 1895 2192 1805 1878 1823 1854 1596
2163 2186 2145 1893 1500 1560 2184 1075 2515 1891 2520

Construa uma apresentação de ramo-e-folhas para cada conjunto de dados e compare explicando eventuais diferenças.

7. Determine qual nível de mensuração é mais apropriado, se nominal ou ordinal.

a. Quando se faz uma pesquisa para encontrar o número de pessoas que realizam separação do lixo seco e úmido em suas residências.

b. Quando um professor atribui notas A, B, C e D nas avaliações aplicadas.

8. Identifique qual tipo de amostragem é usado nos casos a seguir:

a. Um repórter de noticiário de uma rede de TV analisa a opinião das pessoas sobre a disponibilidade dos recursos hídricos entrevistando pessoas que passam em frente ao seu estúdio.

b. Um pesquisador da Universidade Federal de Rondônia examina a qualidade da água de todos os poços residenciais de cinco bairros selecionados aleatoriamente em Ji-Paraná.

c. Um engenheiro de controle de qualidade seleciona cada centésima fonte de computador que passa em uma esteira transportadora.

9. Identifique qual tipo de amostragem deveria ser utilizada na situação em que um pesquisador quer verificar a qualidade da água em um bairro onde se sabe que tem apenas poços artesianos, e o mesmo coleta a amostra a cada cinco casas:

a. Sistemática.

b. Estratificada.

c. Conveniência.

d. Aleatória simples.

e. Conglomerado.

10. Identifique qual(is) tipo(s) de amostragem é usado nos casos abaixo:

a. Um pesquisador seleciona aleatoriamente 10 quarteirões em Ji-Paraná e pergunta a todos os adultos residentes nos mesmos se fazem ou não a separação dos resíduos seco e úmido de suas residências.

b. Um estudante de estatística obtém dados sobre altura/peso entrevistando membros da família.

c. Um economista está estudando o efeito da educação sobre o salário e realiza uma pesquisa com 150 trabalhadores selecionados aleatoriamente de cada uma das seguintes categorias: menos do que Ensino Médio; Ensino Médio; mais do que Ensino Médio.

d. Motivado pelo fato de um estudante ter morrido por excesso de bebida, a Faculdade de Medicina da USP fez um estudo do hábito de bebida dos estudantes, selecionando aleatoriamente 10 classes diferentes e entrevistando todos os estudantes em cada uma dessas classes.

- e. O Comissário de Jurados de um distrito obtém uma lista de 5.763 proprietários de carros e obtém um conjunto de jurados selecionando cada centésimo nome na lista.
- f. Em uma pesquisa do IBAMA, os sujeitos da entrevista foram selecionados usando-se um computador para gerar aleatoriamente números de telefones, que eram então discados.

11. Complete a tabela abaixo:

i	Classes	Frequência	Fr. Relativa	Fr. Percentual	Fr. Acumulada
1	0 - 8	4			
2	8 - 16	10			
3	16 - 24	14			
4	24 - 32	9			
5	32 - 40	3			
$\Sigma = 40$					

12. Com base nos dados do exercício 5 construa uma tabela de distribuição de frequência completa (use uma amplitude da classe de 6).

- a. Analise os dados e novamente faça um comentário sobre a distribuição dos mesmos.
- b. Compare as novas análises com as obtidas no exercício 5 e relate se foi possível observar mais detalhes ou não, uma vez que agora os dados estão separados por classes.

13. Com base nos dados do exercício 6 construa uma tabela de distribuição de frequência completa (use uma amplitude da classe de 100 gramas no item a e de 250 no b).

- a. Analise os dados e novamente explique eventuais diferenças.
- b. Compare as novas análises com as obtidas no exercício 6 e relate se foi possível observar mais detalhes ou não ao comparar os dois conjuntos de dados, uma vez que agora os dados estão separados por classes.

14. Um estudo realizado com uma amostra de 50 estudantes universitários para verificar o uso de tecnologia doméstica coletou informação sobre a quantidade de horas que cada estudante fazia uso do computador pessoal durante uma semana.

4,1 1,5 10,4 5,9 3,4 5,7 1,6 6,1 3,0 3,7
 3,1 4,8 2,0 14,8 5,4 4,2 3,9 4,1 11,1 3,5
 4,1 4,1 8,8 5,6 4,3 3,3 7,1 10,3 6,2 7,6
 10,8 2,8 9,5 12,9 12,1 0,7 4,0 9,2 4,4 5,7
 7,2 6,1 5,7 5,9 4,7 3,9 3,7 3,1 6,1 3,1

Sintetize os dados construindo:

- a. Uma distribuição de frequência (use uma amplitude da classe de 3 horas).
- b. Uma distribuição de frequência relativa.
- c. Um histograma e um polígono de frequência.
- d. Uma ogiva.
- e. Comente o que os dados indicam sobre o uso do computador pessoal.

15. Com base na tabela de distribuição de frequência feita no exercício 12, encontre:

- a. a amplitude total;
- b. a frequência da segunda classe;
- c. o limite superior da terceira classe;
- d. o ponto médio da quarta classe;
- e. a frequência relativa da quinta classe.

16. Com base na tabela de distribuição de frequência feita no exercício 13, encontre para os dois conjuntos de dados:

- a. a amplitude amostral;
- b. a frequência da quinta classe;
- c. o limite inferior da segunda classe;
- d. o ponto médio da primeira classe;
- e. a frequência percentual da sexta classe.

17. Os seguintes dados são medidas de viscosidade para um produto químico observado de hora em hora (leia para baixo e então, da esquerda para a direita).

47,9	48,4	48,8	48,9	48,6	48,5	43,2	43,0	43,0	43,6
47,9	48,1	48,1	48,6	48,0	48,1	43,0	42,9	42,8	42,2
48,6	48,0	48,3	48,0	47,9	48,0	43,5	43,6	43,1	43,5
48,0	48,6	48,2	47,5	48,3	48,3	43,1	43,3	43,2	43,0

- a. Sintetize os dados construindo uma distribuição de frequência.
- b. Construa o gráfico que achares mais apropriado para representar os dados.
- c. As especificações de viscosidade desse produto é de 48 ± 2 . Que conclusões você pode tirar sobre o desempenho desse processo?

18. A seguir temos uma distribuição de frequência das áreas de 400 lotes que serão utilizados na construção de uma indústria alimentícia no Estado de Rondônia:

Áreas (m ²)	300	400	500	600	700	800	900	1.000	1.100	1.200
Número de lotes	14	46	58	76	68	62	48	22	6	

Com referência a esses dados, determine:

- a. Amplitude total.
- b. O limite superior da quinta classe.
- c. O limite inferior da oitava classe.
- d. O ponto médio da sétima classe.
- e. A amplitude de intervalo da segunda classe.
- f. A frequência da quarta classe.
- g. A frequência relativa da sexta classe.
- h. A frequência acumulada da quinta classe.
- i. O número de lotes cuja área não atinge 700 m².
- j. O número de lotes cuja área atinge e ultrapassa 800 m².

- k. A percentagem dos lotes cuja área não atinge 600 m^2 .
- l. A percentagem dos lotes cuja área seja maior ou igual a 900 m^2 .
- m. A percentagem dos lotes cuja área é de 500 m^2 , no mínimo, mas inferior a 1.000 m^2 .
- n. A classe do 72° lote.
- o. Até que classe estão incluídos 60% dos lotes.

19. Com base no exercício 5 construa um gráfico de linhas e uma ogiva, e comente mediante análise gráfica o que se pode observar relacionado a esses dados.

20. Com base no exercício 6 construa um histograma e uma ogiva para cada conjunto de dados, e comente mediante análise gráfica o que os dados indicam sobre eventuais diferenças na quantidade de lixo coletado.

21. Um questionário para verificar a opinião dos rondonienses quanto à construção da usina hidrelétrica no rio Madeira forneceu as seguintes respostas: 62 a favor, 48 contra e 30 sem opinião.

- a. Na construção de um gráfico de pizza, quantos graus haveria na seção que mostra as respostas afirmativas?
- b. Quantos graus haveria na seção que mostra as respostas negativas?
- c. Construa um gráfico de pizza.
- d. Construa um gráfico de barras.

22. Represente a tabela abaixo usando o gráfico em barras:

Produção brasileira de petróleo bruto	
Anos	Quantidade (1.000 m³)
1991	36.180,4
1992	36.410,5
1993	37.164,3

Fonte: Petrobrás

23. Represente a tabela abaixo usando o gráfico em setores:

Área terrestre Brasil	
Regiões	Relativa (%)
Norte	45,25
Nordeste	18,28
Sudeste	10,85
Sul	6,76
Centro-Oeste	18,86
Total	100,00

Fonte: IBGE

24. Para facilitar um projeto de ampliação da rede de esgoto de certa região de uma cidade, as autoridades tomaram uma amostra de tamanho 50 dos 270 quarteirões que compõem a região, e foram encontrados os seguintes números de casas por quarteirão:

73	83	66	85	78	97	61	74	61	75
89	92	76	69	30	54	80	53	51	51
45	36	55	42	52	59	55	33	62	32
67	18	57	21	13	75	30	65	44	45
71	39	68	41	29	45	46	58	25	55

- Use intervalo igual a 5 e construa um histograma e um polígono de frequência.
- Use a regra de Sturges para encontrar o número de classes sugerido para esse conjunto de dados e construa um novo histograma.
- Faça um comentário sobre as possíveis diferenças entre os dois histogramas.

25. A moda, média e mediana são conhecidas como medidas de tendência central porque:

- tendem para o centro de uma distribuição, onde está situada a maioria dos escores.
- tendem a ser centrais para nosso entendimento de Estatística.
- tendem a localizar-se no ponto médio de um intervalo de classe.
- todas as alternativas anteriores.

26. Que medida de tendência central representa o ponto de maior frequência em uma distribuição?

- Moda.
- Mediana.
- Média.

27. Que medida de tendência central é considerada o ponto de equilíbrio de uma distribuição?

- Moda.
- Mediana.
- Média.

28. Uma distribuição de rendas é fortemente assimétrica. Que medida de tendência central você utilizaria para caracterizar a renda?

- Moda.
- Mediana.
- Média.

29. Os seguintes dados são medidas de intensidade solar direta (watts.m^{-2}), em dias diferentes, em uma localização no sul do Brasil. Calcule as medidas de tendência central e comente sobre os dados.

562	869	708	775	775	704	809	856	655	806	878	909
918	558	768	870	918	940	946	661	820	898	935	952
957	693	835	905	939	955	960	498	653	730	753	

30. Considere os dados de intensidade solar do exercício prévio. Com esses dados construa um gráfico de linhas. Indique a média no gráfico. Dê uma interpretação prática da média da amostra.

31. O pH de uma solução é medido oito vezes por uma operadora usando o mesmo instrumento. Ela obtém os seguintes dados: 7,15; 7,20; 7,18; 7,19; 7,21; 7,20; 7,16 e 7,18.

- Calcule a média da amostra
- Calcule a mediana.

32. O transporte público e o automóvel são dois meios que um empregado pode usar para ir ao trabalho diariamente. Amostras de tempo para cada meio estão registradas a seguir. Os tempos estão em minutos.

Transporte público:	28	29	32	37	33	25	29	32	41	34
Automóvel:	29	31	33	32	34	30	31	32	35	33

- Calcule o tempo médio da amostra de cada meio de trabalho e a amplitude amostral.
- Com base nos resultados, que meio de transporte deve ser preferido? Explique.
- Calcule o percentil 60 para cada meio de transporte e interprete.
- Calcule o primeiro quartil para cada meio de transporte e interprete.

33. O conteúdo médio de material sólido em suspensão na água do rio R costuma ser de 205 mg/L. Uma coleta recente, em 9 pontos desse curso d'água, forneceu os dados a seguir. Calcule todas as medidas de tendência central e comente o resultado.

210	242	226	268	251	206	218	215	207
-----	-----	-----	-----	-----	-----	-----	-----	-----

34. Consideremos a distribuição salarial de 16 engenheiros ambientais que ocupam o mesmo cargo em certa empresa que possui filiais em várias partes do Brasil:

Salários (R\$)	3.200,00	3.350,00	3.700,00	4.000,00	4.800,00
Número de funcionários	3	5	4	3	1

- Calcule o salário médio, o salário mediano, o salário modal e comente os resultados.
- Calcule o primeiro e o terceiro quartis e interprete.
- Calcule o percentil 40 e interprete.

35. Com base nos dados do exercício 5 calcule a média, a mediana, a moda e comente os resultados.

36. Com base no exercício 6 faça o que se pede para cada conjunto de dados:

- a. Calcule a média, a mediana, a moda e comente os resultados
- b. Calcule o segundo quartil e interprete.
- c. Calcule o percentil 60 e interprete.

37. Num teste automobilístico de distância e de consumo de gasolina, 13 automóveis foram testados por 300 quilômetros em estrada, nas mesmas condições de direção na cidade e no campo. Os seguintes dados foram registrados para o desempenho quilômetros por litro.

Cidade: 16,2 16,7 15,9 14,4 13,2 15,3 16,8 16,0 16,1 15,3 15,2 15,3 16,2
Campo: 19,4 20,6 18,3 18,6 19,2 17,4 17,2 18,6 19,0 21,1 19,4 18,5 18,7

Use a média, a mediana e a moda para fazer um relatório sobre a diferença no desempenho para a condução na cidade e no campo.

38. Os dados abaixo representam a quantidade de copos plásticos que são utilizados em 20 dias úteis consecutivos em uma empresa com 900 funcionários :

440 677 481 690 707 514 671 488 483 554
611 638 572 514 623 664 631 570 484 612

- a. Determine a média, a mediana e a moda desse conjunto de valores.
- b. Arredonde os 20 valores para a dezena mais próxima e calcule novamente a média, a mediana e a moda.
- c. Arredonde os 20 valores para a centena mais próxima e calcule novamente a média, a mediana e a moda.
- d. Enuncie uma conclusão sobre o efeito do arredondamento da média, da mediana e da moda.

39. Decidiu-se investigar a distribuição dos profissionais com nível universitário em duas regiões, A e B. As informações pertinentes foram obtidas e encontram-se no quadro abaixo expressas em salários mínimos. Faça uma descrição das principais diferenças observadas.

Região	Média	Mediana	Moda	Q ₁	Q ₃
A	20,00	20,32	20,15	17,32	22,68
B	20,00	18,00	17,00	16,00	24,00

40. Os registros de uma grande universidade mostram que 22 alunos do último ano do curso de Engenharia consultaram os seguintes números de livros durante o ano acadêmico:

62	73	40	72	79	88	35	51	48	42	75
65	69	82	50	66	103	68	54	38	52	72

- a. Determine a mediana.
b. Ache os quartis Q_1 e Q_3 e os interprete.

41. Os dados a seguir referem-se ao grau de conforto (valores mais altos, mais conforto) no uso de dois tipos de pipetas de laboratório. O primeiro tipo foi experimentado por 5 pessoas e o segundo por outra 5, totalizando 10 usuários. Com que tipo de pipeta os usuários sentem-se mais confortáveis?

Tipo A:	5	1	3	4	2
Tipo B:	5	8	9	7	6

42. Os seguintes dados são as temperaturas ($^{\circ}\text{C}$), em dias consecutivos, do efluente na descarga de uma unidade de tratamento de esgoto.

43	47	51	48	52	50	46	49
45	52	46	51	44	49	46	51
49	45	44	50	48	50	49	50

- a. Calcule a média e a mediana da amostra.
b. Calcule o desvio padrão.
c. Construa um diagrama em caixa e comente sobre a informação nesse diagrama.
d. Encontre os percentis 5 e 95 da temperatura.

43. A concentração de uma solução é medida seis vezes por uma operadora que usa o mesmo instrumento. Ela obtém os seguintes dados: 63,2; 67,1; 65,8; 64,0; 65,1; 65,3 (g/L).

a. Calcule a média da amostra. Suponha que o valor desejado para essa solução tenha sido especificado em 65 g/L. Você acha que o valor médio calculado aqui é suficientemente próximo do valor alvo, para que se possa afirmar que a solução tenha atingido o alvo? Justifique sua resposta.

b. Calcule o desvio padrão.

c. Suponha que, para medir a concentração, a operadora tenha de calibrar o aparelho e usar um material reagente. Quais as maiores fontes de variabilidade que você imagina para esse experimento? Por que é desejável ter uma pequena variância dessas medidas?

44. Certa bióloga está medindo o dano ambiental em uma reserva biológica. Em seis locais dessa reserva, determinou um escore de dano (em graus), obtendo os dados a seguir.

Local:	L1	L2	L3	L4	L5	L6
Dano:	2	5	1	0	3	4

Calcule a média, a mediana e o desvio padrão e comente o resultado.

45. Com base no exercício 5 faça o que se pede:

- a. Calcule o desvio padrão e o coeficiente de variação.
- b. Determine o tipo de assimetria, o coeficiente de curtose e sua classificação.

46. Com base no exercício 6 faça o que se pede para cada conjunto de dados:

- a. Calcule o desvio padrão e o coeficiente de variação.
- b. Determine o tipo de assimetria, o coeficiente de curtose e sua classificação.

47. Considere os dados de intensidade solar do exercício 29. Com esses dados calcule o desvio padrão e faça novas considerações sobre os conjuntos de dados. Calcule o primeiro e o terceiro quartis e interprete e determine o tipo de assimetria.

48. Considere os dados de pH do exercício 31. Com esses dados calcule o desvio padrão e liste quais são as maiores fontes de variabilidade nesse experimento. Determine também o tipo de assimetria, o coeficiente de curtose e sua classificação.

49. Abaixo estão listadas as idades de motociclistas quando foram vítimas fatais de acidentes de trânsito.

17	40	14	31	16	30
25	27	23	34	21	28
38	20	18	37	42	25

- a. Calcule a média, o desvio-padrão e o coeficiente de variação e verifique se os resultados confirmam a crença comum de que tais fatalidades são ocasionadas por uma grande proporção de motoqueiros mais jovens? Explique.
- b. Calcule as medidas de assimetria e de curtose, classifique e interprete os resultados.

50. A *Folha de São Paulo* regularmente publica o índice da qualidade do ar para várias cidades do Brasil. Uma amostra de valores do índice da qualidade do ar para São Paulo forneceu os seguintes dados: 28, 42, 58, 48, 45, 55, 60, 49 e 50.

- a. Calcule a amplitude;
- b. Calcule a variância e o desvio-padrão da amostra;
- c. Uma amostra de leituras do índice da qualidade do ar para o Rio de Janeiro forneceu uma média de amostra de 48,5, uma variância de 136 e um desvio-padrão de 11,66. Que comparações você pode fazer entre a qualidade do ar em São Paulo e no Rio de Janeiro com base nessas estatísticas descritivas?

51. Em uma empresa A, a média dos salários é 10.000 unidades e o 3º quartil é 5.000.

- a. Se você se apresentasse como candidato a funcionário nessa firma e se o seu salário fosse escolhido ao acaso entre todos os possíveis salários, o que seria mais provável: ganhar mais ou menos que 5.000 unidades?
- b. Suponha que na companhia B a média dos salários seja 7.000 unidades, a variância praticamente zero e o salário também seja escolhido ao acaso. Em qual companhia você se

apresentaria para procurar o emprego?

52. Dão-se a seguir os tempos (em segundos) de reação a um alarme de incêndio, após a liberação de fumaça de uma fonte fixa:

12 9 11 7 9 14 6 10

Calcule a amplitude, a variância, e o desvio padrão e comente os resultados.

53. O dano ecológico devido ao despejo de substâncias produzidas por certa fábrica foi medido em quatro pontos de um curso d'água: antes da saída do efluente da fábrica (ponto A), na saída do efluente (B) e em dois outros pontos situados após o local B (C e D). Valores maiores indicam danos ecológicos maiores.

Índices de dano ecológico.

Ponto A	Ponto B	Ponto C	Ponto D
1	5	3	3
2	4	0	3
	5	2	2
			2

Considere que todos os conjuntos de dados dessa lista têm distribuição dos dados aproximadamente normal.

**Penso noventa e nove vezes e nada descubro;
deixo de pensar, mergulho em profundo silêncio –
e eis que a verdade se me revela.**

Albert Einstein (1879-1955), físico alemão.