

IFCE- Departamento de Física e Matemática
Disciplina: Probabilidade e Estatística - 2024.1
Prof: Valberto Feitosa

01) Em cada população abaixo, classifique as variáveis descritas em qualitativas nominal ou ordinal, e quantitativas discretas ou contínuas:

a. Funcionários de uma indústria em Fortaleza.

Salário: _____
Tempo (anos) na empresa: _____
Estado civil: _____
Grau de instrução: _____
Nº de funcionários que recebem adicional de salubridade: _____

b. Alunos do Centro de Tecnologia da UFC, 2023.1

Nº de alunos matriculados: _____
Classificação do aluno no ENEM: _____
Disciplina cursada pelo aluno nesse semestre: _____
Renda familiar: _____
Número de irmãos: _____

c. Computadores ligados à Internet no Ceará em 2021.

Custo das instalações: _____
Nº de usuários ao mês: _____
Marca dos computadores: _____
Ordem de inscrição na rede: _____

02) Foi retirada uma amostra de galões de tinta produzida por uma fábrica, e registradas as seguintes medidas para o tempo de secagem, em horas, de certa marca de tinta látex:

3,4 ; 2,5 ; 4,8 ; 2,9 ; 3,6 ; 2,8 ; 3,3 ; 5,6 ; 3,7 ; 2,8 ; 4,4 ; 4,0 ; 5,2 ; 3,0 ; 4,8.

- Qual é o tamanho da amostra acima?
- Qual a variável (dado) que está sendo investigado? E qual é o tipo de dado em estudo?
- Determine a média, a mediana e a moda para essa amostra.
- Determine o mínimo, máximo, a variância, desvio-padrão e Coeficiente de variação (%).

03) A quantidade de radiação recebida em uma estufa tem importante papel na determinação da taxa de fotossíntese. As observações a seguir, sobre radiação solar, foram lidas de um gráfico no artigo "*Radiation Components over Bare and Planted Soils in a Greenhouse*" (Solar Energy, 1990, p. 1011-1016):

6,3 ; 6,4 ; 7,7 ; 8,4 ; 8,5 ; 8,8 ; 8,9 ; 9,0 ; 9,1 ; 10,0 ; 10,1 ; 10,2 ; 10,6 ; 10,6 ;
10,7 ; 10,7 ; 10,8 ; 10,9 ; 11,1 ; 11,2 ; 11,2 ; 11,4 ; 11,9 ; 11,9 ; 12,2 ; 13,1

Use os métodos discutidos nas aulas para descrever e resumir os dados.

04) Estamos interessados em estudar a idade dos 12.325 funcionários da Cia. Distribuidora de Leite Teco, e isso será feito por meio de uma amostra. Para determinar que tamanho deverá ter essa amostra, foi colhida uma amostra-piloto. As idades (ano) observadas foram: 42, 35, 27, 21, 55, 18, 27, 30, 21, 24.

- Determine as medidas descritivas dos dados que você conhece.
- Qual dessas medidas você acredita que será a mais importante para julgar o tamanho final da amostra? Por quê?

05) Para facilitar um projeto de ampliação da rede de esgoto de uma 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:

2	2	3	10	13	14	15	15	16	16	18	18	20	21	22	59	61
26	27	29	29	30	32	36	42	44	45	45	46	48	52	58	61	61
6	66	68	75	78	80	89	90	92	97	22	23	24	25	25	65	

- Determine uma medida de posição central, para representar o número de casas/quarteirão dessa região, e uma medida de dispersão que você acha mais adequada para representar esses dados.
- Uma outra região apresentou um $CV(\%)=47,3\%$ para o número médio de casas/quarteirão. Em qual das duas regiões observa-se uma homogeneidade maior de resultados? Justifique.

06) Foram selecionadas amostras de três tipos de corda e o limite de fadiga (MPa) foi determinado para cada amostra, resultando os dados a seguir.

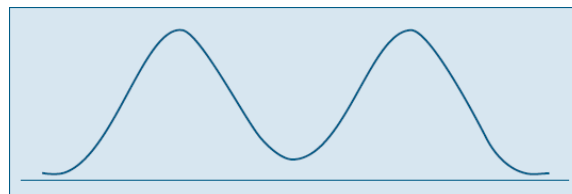
Tipo 1: 350, 350, 350, 358, 370, 370, 370, 371, 371, 372, 372, 384, 391, 391, 392

Tipo 2: 350, 354, 359, 363, 365, 368, 369, 371, 373, 374, 376, 380, 383, 388, 392

Tipo 3: 350, 361, 362, 364, 364, 365, 366, 371, 377, 377, 377, 379, 380, 380, 392

- Calcule as medidas descritiva, apresentadas nas aulas, para representar e resumir as amostras acima.
- Qual tipo de corda apresentou resultado mais homogêneo? Justifique.
- Construa um *boxplot* comparativo e comente as semelhanças e diferenças.
- Construa um gráfico de pontos (*dotplot*) comparativo (um *dotplot* para cada amostra com uma escala comum). Comente as semelhanças e diferenças.
- O *boxplot* comparativo da parte (a) fornece uma avaliação informativa das semelhanças e diferenças? Explique seu raciocínio.

07) Suponha que a variável de interesse tenha a distribuição como na figura abaixo.



Você acha que a média é uma boa medida de posição? E a mediana? Justifique.

08) Um órgão do governo do estado está interessado em determinar padrões sobre o investimento em educação, por habitante, realizado pelas prefeituras em diferentes cidades de uma região. De um levantamento de dez cidades, foram obtidos os valores (codificados) da tabela abaixo:

Cidade	A	B	C	D	E	F	G	H	I	J
Investimento(\$)	20	16	14	8	19	15	14	16	19	18

Nesse caso, será considerado como *investimento básico* a *média final* das observações, calculada da seguinte maneira:

- Obtém-se uma média inicial geral.
- Eliminam-se do conjunto aquelas observações que forem superiores à média inicial mais duas vezes o desvio padrão, ou inferiores à média inicial menos duas vezes o desvio padrão.
- Calcula-se a média final com o novo conjunto de observações.

Qual o investimento básico que você daria como resposta: o que considera a média das observações ou essa proposta acima?

Observação: O procedimento do item 2 tem a finalidade de eliminar do conjunto a cidade cujo investimento é muito diferente dos demais.

09) Quinze amostras de ar de certa região foram obtidas e para cada uma delas foi determinada a concentração de monóxido de carbono. Os resultados (em ppm) foram: 9,3; 10,7; 8,5; 9,6; 12,2; 15,6; 9,2; 10,5; 9,0; 13,2; 11,0; 8,8; 13,7; 12,1; 9,8.

Calcule as medidas descritivas usuais para esse conjunto, além dos Quartis, Decil 7 e Centil 40, interpretando o que cada valor significa. Podemos supor que esses dados são heterogêneos? Justifique sua resposta.

10) O elevado consumo de energia durante o exercício continua após o fim do treino. Como as calorias queimadas após o exercício contribuem para a perda de peso e têm outras consequências, é importante entender esse processo. O estudo “*Effect of Weight Training Exercise and Treadmill Exercise on Post-Exercise Oxygen Consumption*” (Medicine and Science in Sports and Exercise, 1998, p. 518-522) relatou os dados a seguir de um estudo em que o consumo de oxigênio (litros) foi medido continuamente por 30 minutos para cada um dos 15 indivíduos após um exercício de levantamento de peso (X) e após um exercício em esteira rolante (Y).

Indivíduo	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
(x)	14,6	14,4	19,5	24,3	16,3	22,1	23,0	18,7	19,0	17,0	19,1	19,6	23,2	18,5	15,9
(y)	11,3	5,3	9,1	15,2	10,1	19,6	20,8	10,3	10,3	2,6	16,6	22,4	23,6	12,6	4,4

- Calcule as medidas descritivas de X e Y.
- Construa um *boxplot* comparativo (um *boxplot* para cada amostra com uma escala comum) das observações de levantamento de peso e de esteira rolante e comente as características encontradas.
- Como os dados estão na forma de pares (x, y), com as medidas x e y da mesma variável em duas condições diferentes, é natural focar as diferenças entre os pares: $d1=x_1-y_1$, ..., $d_n=x_n-y_n$. Construa um *boxplot* das diferenças das amostras. O que ele sugere?

11) Considere os dados a seguir sobre tempo de conserto ativo (horas) para uma amostra de $n=46$ rádios de comunicação aérea:

0,2 0,3 0,5 0,5 0,5 0,6 0,6 0,7 0,7 0,7 0,8 0,8 0,8 1,0 1,0 1,0 1,0 1,1 1,3 1,5 1,5
 1,5 1,5 2,0 2,0 2,2 2,5 2,7 3,0 3,0 3,3 3,3 4,0 4,0 4,5 4,7 5,0 5,4 5,4 7,0 7,5 8,8
 9,0 10,3 22,0 24,5

- Calcule as medidas descritivas para resumir o conjunto.
- Construa uma distribuição de frequências com seis intervalos de classe, que tenham a mesma largura.
- Calcule as medidas descritivas considerando a distribuição de frequência em (b).
- Compare os resultados das medidas calculadas em (a) e em (c). Comente.

12) Em uma pesquisa realizada com 100 famílias, levantaram-se as seguintes informações:

Número de filhos	10	11	12	13	4	5	Mais que 5
Frequência de famílias	17	20	28	19	7	4	5

- Qual a mediana do número de filhos?
- E a moda?
- Que problemas você enfrentaria para calcular a média? Faça alguma suposição e encontre-a.

13) Estudando-se o consumo diário de leite, verificou-se que, em certa região, 25% das famílias consomem até um litro, 25% consomem entre um e dois litros, 25% consomem entre dois e três litros e o restante consome entre três e cinco litros. Para a variável em estudo:

- Escreva as informações acima na forma de uma tabela de frequências.
- Construa o histograma.
- Calcule a média e a mediana.
- Calcule a variância e o desvio padrão.
- Qual o valor do 1º quartil?

14) A MB Indústria e Comércio, desejando melhorar o nível de seus funcionários em cargos de chefia, organizou um curso experimental e indicou 25 funcionários para a primeira turma. Os dados referentes à seção a que pertencem, notas e graus obtidos no curso estão na tabela a seguir. Como havia dúvidas quanto à adoção de um único critério de avaliação, cada instrutor adotou seu próprio sistema de aferição. Usando dados daquela tabela, responda às questões:

Func.	Seção(*)	Administr.	Direito	Redação	Estatíst.	Inglês	Metodologia	Política	Economia
1	P	8,0	9,0	8,6	9,0	B	A	9,0	8,5
2	P	8,0	9,0	7,0	9,0	B	C	6,5	8,0
3	P	8,0	9,0	8,0	8,0	D	B	9,0	8,5
4	P	6,0	9,0	8,6	8,0	D	C	6,0	8,5
5	P	8,0	9,0	8,0	9,0	A	A	6,5	9,0
6	P	8,0	9,0	8,5	10,0	B	A	6,5	9,5
7	P	8,0	9,0	8,2	8,0	D	C	9,0	7,0
8	T	10,0	9,0	7,5	8,0	B	C	6,0	8,5
9	T	8,0	9,0	9,4	9,0	B	B	10,0	8,0
10	T	10,0	9,0	7,9	8,0	B	C	9,0	7,5
11	T	8,0	9,0	8,6	10,0	C	B	10,0	8,5
12	T	8,0	9,0	8,3	7,0	D	B	6,5	8,0
13	T	6,0	9,0	7,0	7,0	B	C	6,0	8,5
14	T	10,0	9,0	8,6	9,0	A	B	10,0	7,5
15	V	8,0	9,0	8,6	9,0	C	B	10,0	7,0
16	V	8,0	9,0	9,5	7,0	A	A	9,0	7,5
17	V	8,0	9,0	6,3	8,0	D	C	10,0	7,5
18	V	6,0	9,0	7,6	9,0	C	C	6,0	8,5
19	V	6,0	9,0	6,8	4,0	D	C	6,0	9,5
20	V	6,0	9,0	7,5	7,0	C	B	6,0	8,5
21	V	8,0	9,0	7,7	7,0	D	B	6,5	8,0
22	V	6,0	9,0	8,7	8,0	C	A	6,0	9,0
23	V	8,0	9,0	7,3	10,0	C	C	9,0	7,0
24	V	8,0	9,0	8,5	9,0	A	A	6,5	9,0
25	V	8,0	9,0	7,0	9,0	B	A	9,0	8,5

Nota: (*) (P = departamento pessoal, T = seção técnica e V = seção de vendas)

- Após observar atentamente cada variável, e com o intuito de resumi-las, como você identificaria (qualitativa ordinal ou nominal e quantitativa discreta ou contínua) cada uma das 9 variáveis listadas?
- Compare e indique as diferenças existentes entre as distribuições das variáveis Direito, Política e Estatística.
- Calcule as medidas descritivas para cada prova. Interprete os resultados.
- Construa o histograma para as notas da variável Redação.
- Construa a distribuição de frequências da variável Metodologia e faça um gráfico para indicar essa distribuição.
- Como é o aproveitamento dos funcionários na disciplina Estatística, segundo a seção a que eles pertencem?
- Faça uma representação gráfica da variável Seção com cada uma das demais variáveis qualitativas.
- Represente na forma gráfica, através de um gráfico de dispersão as notas obtidas nos cursos de estatística e Economia, identificando no gráfico, qual é a seção de cada funcionário.
- Na figura obtida em (h) identifique, no gráfico, os funcionários que apresentaram o melhor e o pior rendimento nos dois cursos.

15) Mostre que:

$$a) \sum_{i=1}^n (X_i - \bar{X}) = 0$$

$$b) \sum_{i=1}^n (X_i - \bar{X})^2 = \sum_{i=1}^n X_i^2 - n\bar{X}^2$$

$$c) \sum_{i=1}^k n_i (X_i - \bar{X})^2 = \sum_{i=1}^k n_i X_i^2 - n\bar{X}^2$$