STA 01996 - Probabilidade e Estatística

Lista de exercícios 1 - Estatística Descritiva

Exercício 1

Classifique cada uma das variáveis abaixo:

- (a) Temperatura corporal (°C) e valor da pressão arterial (mmHg) de estudantes após uma bateria de exames físicos.
- (b) Número de alunos com média final abaixo de 5,0.
- (c) Perda de peso de maratonistas na Corrida de São Silvestre, em quilos.
- (d) Intensidade da perda de peso de maratonistas na corrida de São Silvestre (leve, moderada ou forte).

Exercício 2

Quinze pacientes de uma clínica de ortopedia foram avaliados quanto ao número de meses previstos de fisioterapia, se haverá (S) ou não (N) seqüelas após o tratamento e o grau de complexidade da cirurgia realizada: alto(A), médio(M) ou baixo(B). Os dados são apresentados na tabela abaixo:

Paciente	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Fisioterapia	7	8	5	6	4	5	7	7	6	8	6	5	5	4	5
Següelas	S	S	Ν	Ν	Ν	S	S	Ν	Ν	S	S	Ν	S	Ν	Ν

- (a) Classifique cada uma das variáveis.
- **(b)** Divida os pacientes em dois grupos: com (S) e sem (N) seqüelas. Obtenha o boxplot para cada um dos grupos utilizando a técnica tipo II para os quartis. Quais são as suas conclusões?
- (c) Compare o tempo médio com o tempo mediano de fisioterapia em cada um dos grupos. Quais são as suas conclusões?
- (d) Calcule o coeficiente de variação para os dois grupos. Qual dos grupos é mais homogêneo (apresenta menor dispersão) em relação ao tempo médio de fisioterapia?
- **(e)** A clínica deseja oferecer um desconto para 25% dos pacientes com maior previsão de tempo de fisioterapia. Qual é a previsão mínima de tempo de fisioterapia para que o paciente receba o desconto? Que medida descritiva fornece essa informação?

Exercício 3

Um estudo pretende verificar se o problema da desnutrição em adultos medida pelo peso, em quilos, em uma região agrícola (denotada por região A), é maior do que em uma região industrial (região B). Para tanto, uma amostra foi tomada em cada região, fornecendo a tabela de freqüências a seguir:

Regiã	o A	Região B				
Peso	f i	Peso	\mathbf{f}_{i}			
< 40	8	< 60	10			
40 50	25	60 70	34			
50 60	28	70 80	109			
60 70	12	80 90	111			
>= 70	9	>= 90	55			

- (a) Construa, a partir das tabelas, um histograma para cada região. Faça uma suposição conveniente para as faixas não delimitadas.
- **(b)** Com base nos histogramas apresentados em (b), você acha que os dados sugerem que o grau de desnutrição é diferente nas duas regiões?
- (c) Obtenha os quartis relacionados aos dados de cada região. Você acha que essa informação vai de encontro a sua conclusão no item b?

Exercício 4

Os dados abaixo apresentam a distância em (km) entre a residência e o local de trabalho dos funcionários da empresa A.

```
1.8 2.5 0.4 1.9 4.4 2.2 3.5 0.2 0.9 1.4 1.1 1.7 1.2 2.3 1.9 0.8 1.5 1.7 1.4 2.1 3.2 15.1 2.1 1.4 0.5 0.9 1.7 0.5 0.8 3.7 1.4 1.8 2.0 1.1 1.0 0.8
```

- (a) Apresente estes dados em ramo-e-folhas.
- **(b)** Na empresa B, a distância (em Km) até a residência de seus 300 funcionários apresenta as seguintes medidas descritivas:

```
mediana = 2.8 	 1^{\circ} quartil = 1.6, 3^{\circ} quartil = 4.2 mínimo = 0.4 	 máximo = 8.8
```

Quais as principais diferenças entre as empresas A e B, em termos da distância entre a residência e o local de trabalho dos funcionários?

Exercício 5

Um hospital maternidade está planejando a ampliação dos leitos para recémnascidos. Para tal, fez um levantamento dos últimos nascimentos obtendo a informação sobre o número de dias que os bebes permanecem no hospital, antes de terem alta. Os dados, já ordenados, são apresentados a seguir:

- (a) Calcule média, moda e mediana.
- (b) Determine o desvio padrão.
- (c) Você identifica algum valor excepcional dentre os que foram observados? Se sim, remova-o(s) e recalcule os itens (a) e (b). Comente as diferenças encontradas.
- (d) Dentre as medidas de posição calculadas em (a), discuta quais delas seriam mais adequadas para resumir esse conjunto de dados.

Exercício 6

O arquivo TREES do Minitab contém medidas de Diâmetro (em polegadas), Altura (em pés) e Volume (em pés³) de 31 árvores. Estatísticas descritivas destes dados são apresentadas a seguir:

Variable	N	Média	Mediana	Desvio p	oadrão
Diâmetro	31	13.248	12.900	3.138	
Altura	31	76.00	76.00	6.37	
Volume	31	30.17	24.20	16.44	
Variable	Mínir	no Má	áximo	Q1	Q3
Diâmetro	8.30	0 20	0.600	11.000	16.000
Altura	63.0	0 8	7.00	72.00	80.00
Volume	10.2	0 77	' .00	19.10	38.30

- (a) Classifique as variáveis sob estudo;
- **(b)** 50% das árvores apresentam medidas de diâmetro, altura e volume inferiores a quais valores, respectivamente? E se considerarmos 75% das árvores?
- (c) Escolhendo casualmente uma árvore desse conjunto, o que seria mais provável: altura maior ou menor que 72 pés?
- (d) Considerando os valores obtidos da média e mediana, comente a distribuição das medidas de altura destas árvores. Faça o mesmo para as medidas de volume e de diâmetro;
- (e) Calcule os coeficientes de variação. As árvores sob estudo são mais parecidas entre si se levamos em conta seus diâmetros, alturas ou volumes? Justifique.
- **(f)** Suponha que um segundo conjunto de 31 árvores apresentou altura média de 80 pés e variância da altura praticamente nula. Se você vai escolher uma árvore ao acaso de um dos grupos e deseja uma árvore alta (muito acima de 80 pés), de qual grupo você selecionaria?