

Some Class

Random Examples

Your Name

Questão 1. Identifique cada uma das variáveis seguintes como qualitativa, quantitativa e como ordinal, nominal ou discreta, contínua.

- (a) A concentração de impurezas em uma amostra de leite, em mg/l.

É uma variável quantitativa, como não é possível colocar os valores em um conjunto enumerável (dado que uma medida poderia apresentar valor 1 ou então 1,0000001) se trata de uma variável quantitativa contínua.

- (b) Tipo de escola de cada candidato ao vestibular da UFC em um determinado ano.

Esta pode ser classificada como uma variável qualitativa. Como temos classificações de escolas como sendo pública ou privadas (sem uma ordem estabelecida entre esses valores), esta é uma variável qualitativa nominal.

- (c) O número de moradores em cada residência de uma cidade.

Se trata se uma variável quantitativa, e como podemos colocar os valores encontrados em um conjunto enumerável se trata de uma variável quantitativa discreta.

- (d) A temperatura de certa região, em determinada época do ano.

A temperatura pode tomar qualquer valor em uma dada escala, desde valores inteiros como 30 graus até valores fracionários 30,6 graus, por exemplo. Portanto, é uma variável quantitativa contínua.

- (e) A produção por hectare de determinado tipo de grão.

A produção por hectare pode tomar qualquer valor em uma dada escala, desde valores inteiros como 30 kg/ha até valores fracionários 30,6 kg/ha, por exemplo. Portanto, é uma variável quantitativa contínua.

Questão 2. Em um estudo sobre contusões causadas durante a prática de esportes, 25 escolas de um estado brasileiro foram selecionadas, ao acaso, entrevistadas. Foram coletados os dados abaixo, sobre o número de contusões classificadas como graves em atletas do sexo masculino para duas modalidades de esporte.

- (a) Construa uma distribuição de frequências para as 50 observações.

	Num. Contusões	Freq
1	1	8
2	2	8
3	3	7
4	4	6
5	5	7
6	6	7
7	7	7

- (b) Construa uma distribuição de frequências para cada modalidade.

Basquete		
	Num. Contusões	Freq
1	1	1
2	2	4
3	3	5
4	4	6
5	5	5
6	6	3
7	7	1

Futebol		
	Num. Contusões	Freq
1	1	7
2	2	4
3	3	2
4	5	2
5	6	4
6	7	6

(c) Represente graficamente cada uma das distribuições.

Para ambas a variável de interesse é uma **variável quantitativa discreta**, Assim a melhor forma de representar graficamente é por meio de um **gráfico de colunas**.

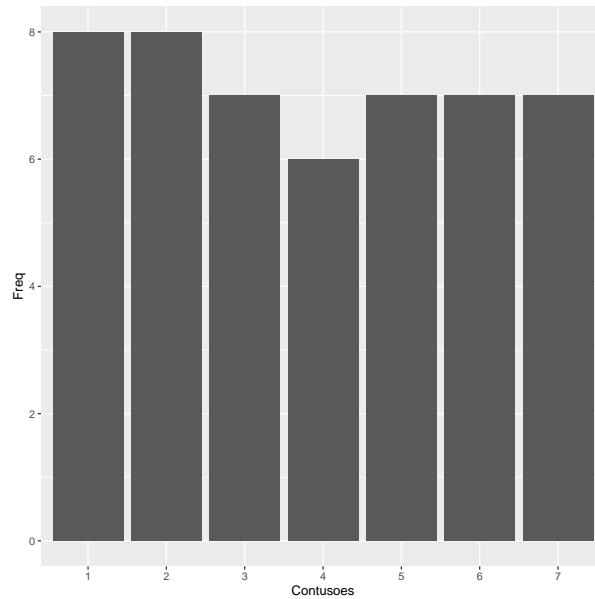


Figure 1: Contusões no Basquete e Futebol

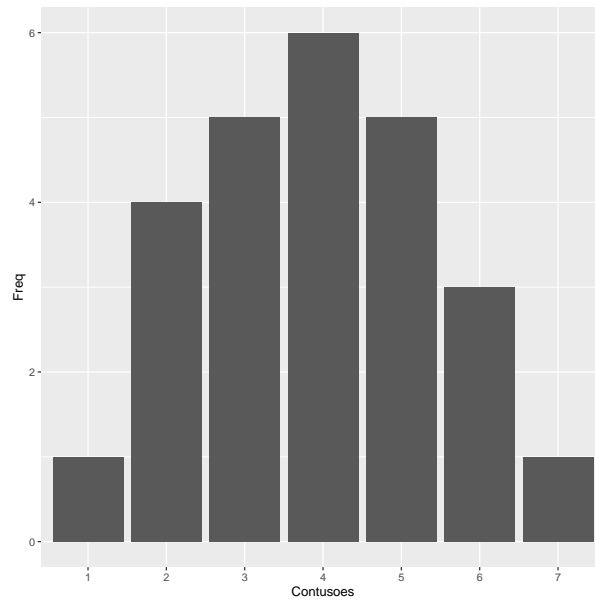


Figure 2: Contusões no Basquete

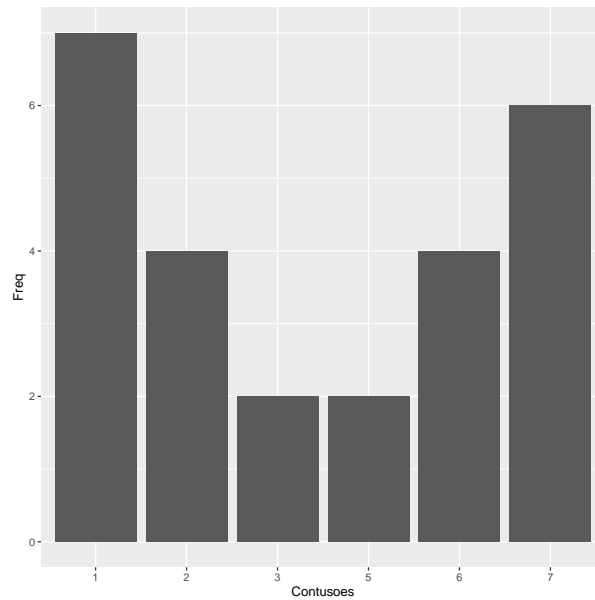


Figure 3: Contusões no Futebol

(d) Comente os resultados.

Percebe-se que a distribuição das contusões no basquete é inferior nos extremos e maior em valores medianos, diferentemente do futebol.

Questão 3. Em uma padaria foi feita uma pesquisa para verificar o consumo de leite e de pão nos primeiros dez dias do mês de janeiro. Foram levantados os seguintes valores diários:

Consumo de leite (litros)	25	26	30	30	28	23	25	29	34	30
Consumo de pão (Kg)	31	40	36	39	39	40	42	38	39	41

(a) Em média, qual foi o produto mais consumido nesses dias?

$$\begin{aligned}\bar{X}_{leite} &= \frac{\sum_{n=1}^{i=10} x_n}{n} \\ &= \frac{25 + 26 + 30 + 30 + 28 + 23 + 25 + 29 + 34 + 30}{10} = 28 \\ \bar{X}_{pao} &= \frac{\sum_{n=1}^{i=10} x_n}{n} \\ &= \frac{31 + 40 + 36 + 39 + 39 + 40 + 42 + 38 + 39 + 41}{10} = 38.5\end{aligned}$$

Assim podemos ver que, em média o consumo de pão foi maior.

(b) Qual dos produtos teve maior variação no consumo, justifique sua resposta.

A variação será dada através do **coeficiente de variação (CV)** tendo a fórmula:

$$\begin{aligned}CV &= \frac{s}{\bar{X}} \times 100; \\ s &= \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n}} \\ \bar{X} &= \frac{\sum_{n=1}^{i=10} x_n}{n}\end{aligned}$$

Assim temos:

$$\begin{aligned}s_{leite} &= 3.265986 \quad \bar{X} = 28 \therefore \\ CV_{leite} &= \frac{3.265986}{28} \times 100 = 11.66424\% \approx 11.66\% \\ s_{pao} &= 3.100179 \quad \bar{X} = 38.5 \therefore \\ CV_{pao} &= \frac{3.100179}{38.5} \times 100 = 8.051948\% \approx 8.05\%\end{aligned}$$

Dessa forma temos um CV maior para o leite, logo o **leite teve maior variação no consumo**.

Questão 4. Os dados abaixo referem-se a dureza de 30 peças de alumínio

- (a) Represente os dados por meio de uma gráfico de ramos-e-folhas.
- (b) Represente os dados por meio de uma distribuição de frequências.

Usando uma distribuição de frequências simples temos, temos:

	Dureza	Freq
1	50.7	1
2	51.1	1
3	52.4	1
4	53	1
5	53.4	1
6	53.5	1
7	54.1	1
8	55.3	1
9	55.7	2
10	59.5	1
11	63.5	1
12	64.3	1
13	67.3	1
14	69.1	1
15	69.5	1
16	70.2	1
17	70.5	1
18	71.4	1
19	72.3	1
20	73	1
21	74.4	1
22	77.8	1
23	78.5	1
24	82.5	1
25	82.7	1
26	84.3	1
27	85.8	1
28	87.5	1
29	95.4	1

Perceba que temos uma grande quantidade de elementos com frequência 1, o que indica que a distribuição é bem dispersa, assim seria mais adequado usar uma distribuição de frequência por intervalos.

$$A_{total} = x_n - x_1 = 95.4 - 50.7 = 44.7$$

Usando a regra da raiz quadrada temos:

$$k = \sqrt{29} = 5.385 \approx 5 \text{ Usando truncamento do valor}$$

Usando a regra de Sturges temos:

$$k = 1 + 3.3 \log_{10}(29) = 1 + 3.3 \times 1.462 = 1 + 4.824 = 5.824 \approx 6$$

Usemos $k = 6$, assim temos:

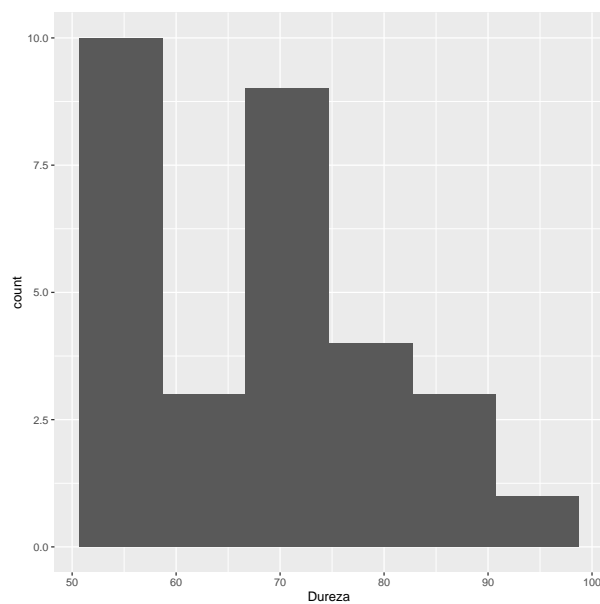
$$h = \frac{A_{total}}{k} = \frac{44.7}{6} = 7.45 \approx 7.4$$

Dessa forma teremos a seguinte distribuição de frequências em intervalos:

	intervals	Freq
1	[49.7,57.1)	10
2	[57.1,64.5)	3
3	[64.5,71.9)	6
4	[71.9,79.3)	5
5	[79.3,86.7)	4
6	[86.7,94.1)	1
7	[94.1,102)	1

- (c) Faça uma representação gráfica para a distribuição de frequências.

Usando a distribuição de frequências por intervalos temos o seguinte histograma:



- (d) Faça o box-plot dos dados. Existem outliers?

Questão 5. Os transdutores de temperatura de um determinado tipo são enviados em lotes de 50. Uma amostra de 60 lotes foi selecionada e o número de transdutores fora das especificações em cada lote foi determinado, resultando nos dados a seguir:

- (a) Construa um diagrama de ramos-e-folhas dos dados. Como ele sugere que a média e mediana serão comparadas?
- (b) Calcule os valores da média (\bar{x}) e da mediana (M_d) amostrais. Dica: $\sum xi = 9638$.

$$\bar{x} = \frac{\sum x_i}{n} = \frac{9638}{26} = 370.6923$$
$$M_d = ?$$

- (c) Em quantos segundos o maior tempo, atualmente 424, pode ser aumentado sem afetar o valor da mediana amostral?
- (d) Quais são os valores de \bar{x} e M_d quando as observações são expressas em minutos?