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

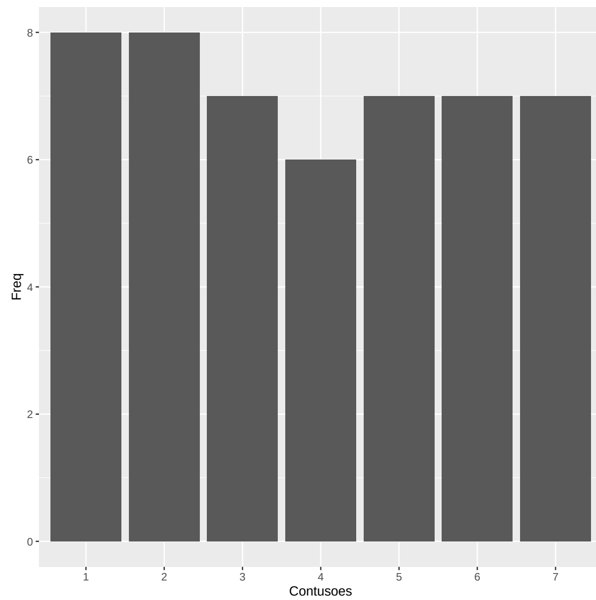


Figure 1: Contusões no Basquete e Futebol

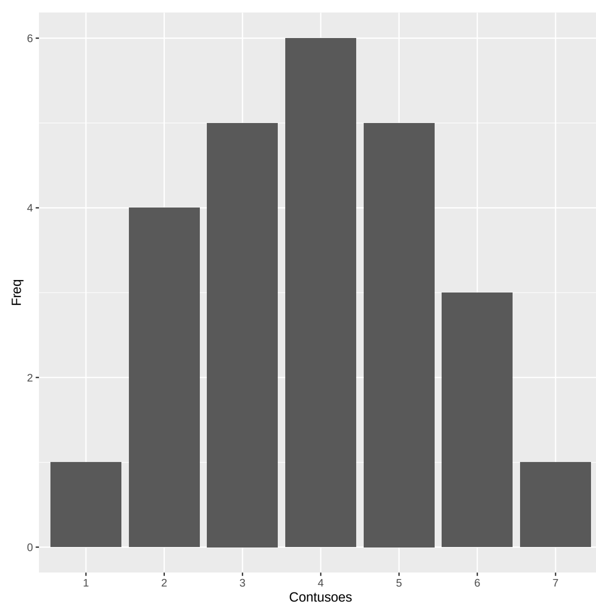


Figure 2: Contusões no Basquete

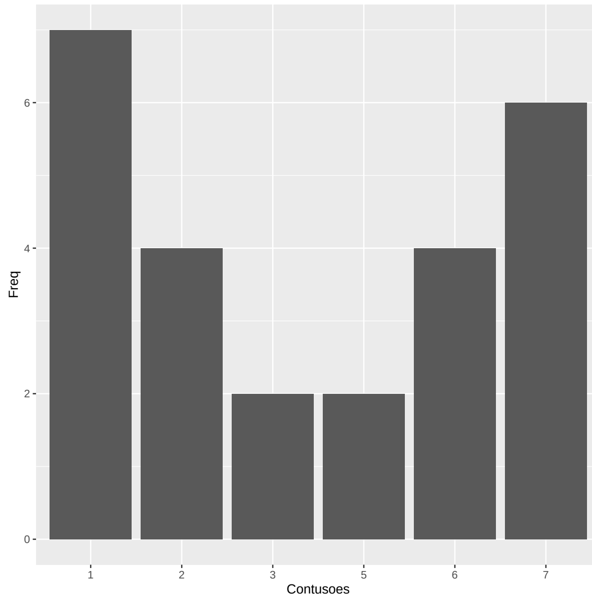


Figure 3: Contusões no Futebol

(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?

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

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

(b) Qual dos produtos teve maior variação no consumo[1] , justifique sua resposta.

A variação será dada através do **coeficiente de variação (CV)** [2] tendo a fórmula:

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

Assim temos:

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

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

[1] Dado a diferença de unidades de medida usadas, a comparação direta não seria adequada, assim utilizamos o **Coeficiente de Variação (CV)**

[2] Dado pela razão entre o desvio padrão s e a média amostral (normalmente expresso em porcentagens)

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.

Gráfico de Ramo-e-Folhas construido para valores inteiros(truncados) dos dados:

5	1 1 2 3 3 4 4
5	5 6 6
6	0 4 4
6	7 9
7	0 0 1 1 2 3 4
7	8 9
8	3 3 4
8	6 8
9	
9	5

(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 para **calcular a quantidade de classes** k temos:

$$k = \sqrt{29} = 5.385 \approx 5$$

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 $k = 6$ [3], assim temos o **cálculo da amplitude do intervalo** h dado por:

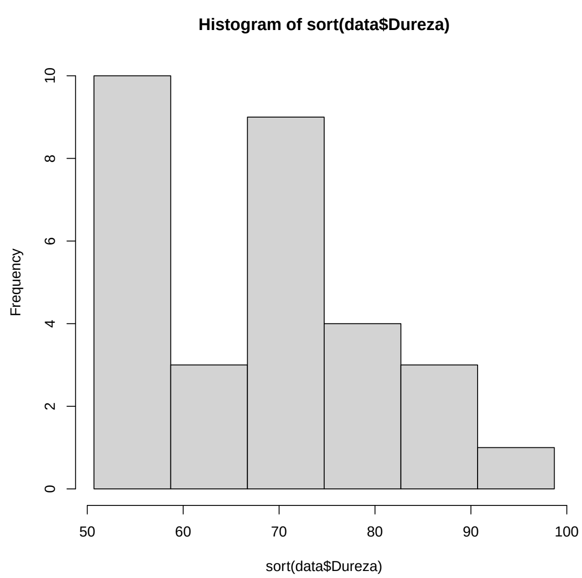
$$h = \frac{A_{total}}{k} = \frac{44.7}{6} = 7.45 \approx 7.4$$

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 [4] dos dados. Existem outliers [5]?

Assim teremos:

$$max = 95.4$$
$$min = 50.7$$

Primeiro quartil $Q_{0.25}$:

$$p_{25} = 0.25 \times (29 + 1) = 0.25 \times 30 = 7.5$$
$$\therefore Q_{0.25} \in [x_7, x_8]$$
$$\therefore Q_{0.25} = 0.5 \times 54.1 + 0.5 \times 55.3 = 54.7$$

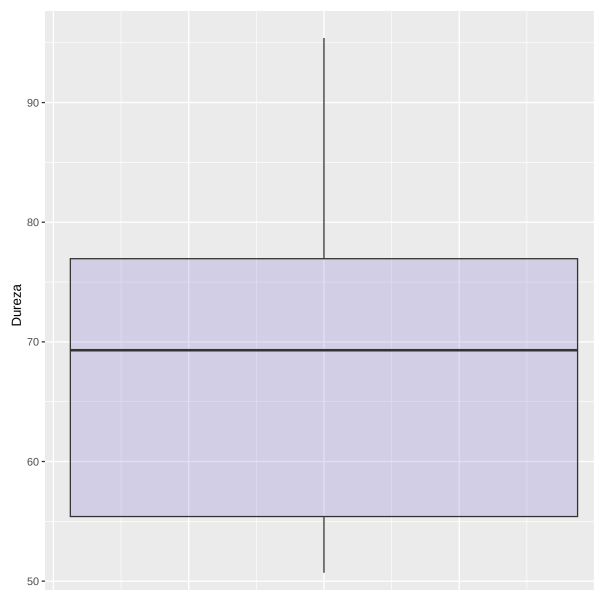
Segundo quartil (mediana) $Q_{0.50} = M_d$:

$$p_{50} = 0.50 \times (29 + 1) = 0.50 \times 30 = 15$$
$$\therefore Q_{0.50} = x_{15} = 69.5$$

Terceiro quartil $Q_{0.75}$:

$$p_{75} = 0.75 \times (29 + 1) = 0.75 \times 30 = 22.5$$
$$\therefore Q_{0.75} \in [x_{22}, x_{23}]$$
$$\therefore Q_{0.75} = 0.5 \times 77.8 + 0.5 \times 78.5 = 78.15$$

Assim temos o box-plot:



Para a análise de **outliers** temos o **limite superior** $LS = Q_3 + (1.5)d_q$ e **limite inferior** $LI = Q_1 - (1.5)d_q$, onde $d_q = Q_3 - Q_1$ ou **distância interquartil**:

$$d_q = 78.15 - 54.7 = 23.45$$
$$LS = 78.15 + (1.5 \times 23.45) = 78.15 + 35.175 = 113.325$$
$$LI = 54.7 - (1.5 \times 23.45) = 54.7 - 35.175 = 19.525$$

Assim, dado os valores encontrados para LS e LI não temos **outliers** na amostra.

[4] Gráfico que relaciona medidas descritivas: máximo, mínimo, mediana (segundo quartil), primeiro e terceiro quartil

[5] Valores discrepantes baseados em um limite superior e inferior

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) Determine as frequências simples e relativas [6] dos valores de x = número de transdutores fora das especificações em um lote.

	fora	Freq	FreqRel
1	0	7.00000000	0.11666667
2	1	12.00000000	0.20000000
3	2	13.00000000	0.21666667
4	3	14.00000000	0.23333333
5	4	6.00000000	0.10000000
6	5	3.00000000	0.05000000
7	6	3.00000000	0.05000000
8	7	1.00000000	0.01666667
9	8	1.00000000	0.01666667
10	Total	60.00000000	1.00000000

(b) Que proporção de lotes na amostra possui no máximo cinco transdutores fora das especificações? Que proporção tem menos de cinco? Que proporção possui no mínimo cinco unidades fora das especificações?

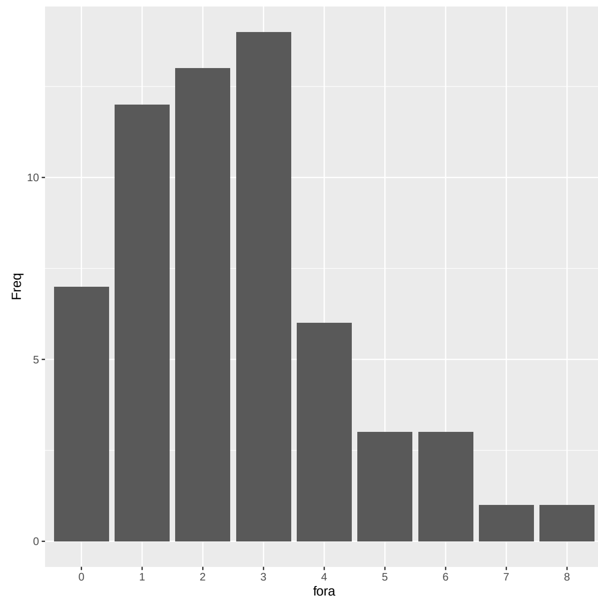
Para aqueles que possuem no máximo cinco transdutores fora temos aqueles que possuem 0, 1, 2, 3, 4 e 5, logo a proporção é de $0.11666667 + 0.20000000 + 0.21666667 + 0.23333333 + 0.10000000 + 0.05000000 = 0.91666667$.

Para aqueles que possuem menos de cinco transdutores fora temos aqueles que possuem 0, 1, 2, 3 e 4, logo a proporção de $0.11666667 + 0.20000000 + 0.21666667 + 0.23333333 + 0.10000000 = 0.86666667$.

Para aqueles que possuem no mínimo cinco transdutores fora temos aqueles que possuem 5, 6, 7 e 8, logo a proporção de $0.10000000 + 0.05000000 + 0.05000000 + 0.01666667 + 0.01666667 = 0.23333333$.

[6] Proporção da frequência em relação a amostra.

(c) Faça a representação gráfica dos dados.



Questão 6. Vinte e seis trabalhadores de plataformas de petróleo offshore participaram de um exercício de fuga simulado, resultando nos dados a seguir (em segundos) para concluir a fuga:

(a) Construa um diagrama de ramos-e-folhas dos dados. Como ele sugere que a média e mediana serão comparadas?

32	5 5
33	4 9
34	
35	6 6 9 9
36	3 4 4 6 9
37	0 3 3 4 5
38	9
39	2 3 4 7
40	2 3
41	
42	4

Assim teremos:

$$\overline{X} = 370.6923$$
$$M_d = 369.5$$
$$\text{ou ainda } P_{50} = 0.5 \times (n+1)$$
$$P_{50} = 0.5 \times 27 = 13.5$$
$$\therefore M_d = \frac{x_{13} + x_{14}}{2} = \frac{369 + 370}{2} = 369.5$$

Quando temos um efeito como esse (mediana próxima da média), temos que a distribuição é simétrica.

(b) Calcule os valores da média ($\overline{x}$) e da mediana (M_d) amostrais. Dica: $\sum xi = 9638$.

Primeiro é necessário fazer a ordenação desses valores, assim teremos:

$$\overline{x} = \frac{\sum x_i}{n} = \frac{9638}{26} = 370.6923$$

Dado a presença de 26 observações, um número par, temos que a mediana será a média dos dois valores centrais:

$$M_d = \frac{x_{13} + x_{14}}{2} = \frac{369 + 370}{2} = 369.5$$

(c) Em quantos segundos o maior tempo, atualmente 424, pode ser aumentado sem afetar o valor da mediana amostral [7]?

$$M_d = \begin{cases} x_{(\frac{n+1}{2})} & \text{se n ímpar} \\ \frac{x_{(\frac{n}{2})} + x_{(\frac{n}{2}+1)}}{2} & \text{se n par} \end{cases}$$

Logo, não afetará a mediana se um valor muito grande for apresentado nos extremos da amostra.

[7] Valor que ocupa posição central na amostra ordenada

(d) Quais são os valores de $\overline{x}$ e M_d quando as observações são expressas em minutos?

Uma das propriedades da média é que ao multiplicar seus valores por uma constante, o resultado também será multiplicado por essa constante, mesmo comportamento observado na mediada, assim:

$$\overline{X} = 370.6923 \times \frac{1}{60} = 6.178$$
$$M_d = 369.5 \times \frac{1}{60} = 6.158$$

Questão 7. As horas semanais dedicadas ao trabalho por uma amostra de 50 profissionais de uma empresa, são apresentadas a seguir:

(a) Complete as informações faltantes, de sorte que a distribuição dos dados seja simétrica.

Horas semanais	Ponto Médio	n_i	$f_i\%$	N_i
$10 \vdash 20$	5	5		
$20 \vdash 30$	10	3		
$30 \vdash 40$	20			
$40 \vdash 50$	10			
$50 \vdash 60$	5			

(b) Qual o percentual de profissionais que trabalham pelo menos 40 horas semanais?

Será 30%.

(c) Construa um histograma com polígono de frequência.

?

Questão 8. Os primeiros quatro desvios em rela¸c˜ao ‘a m´edia de uma amostra de n = 5 tempos de rea¸c˜ao foram 0, 3; 0, 9; 1, 0 e 1, 3. Qual ´e o quinto desvio em rela¸c˜ao ‘a m´edia? Qual a vari^ancia amostral? Forne¸ca uma amostra para qual esses s˜ao os cinco desvios em rela¸c˜ao ‘a m´edia.

Questão 9. foo

(a) bar

Dado $X = \{x_1, x_2, x_3, \dots, x_n\}$

Temos ainda que $\overline{x} = \frac{\sum x_i}{n}$, logo:

$$\sum_{n}^{i=1} (x_i - \overline{x})^2 = \sum_{n}^{i=1} [x_i^2 - 2\overline{x} \sum_{i=1}^{n} x_i + n\overline{x}^2] = \sum_{i=1}^{n} x_i^2 - 2n\overline{x}^2 + n\overline{x}^2 = \sum_{i=1}^{n} x_i^2 - n\overline{x}^2$$

(b) bar2

(c) bar3

A variância será calculada como:

$$s^2 = \frac{1}{n-1} \times \sum_{i=1}^{n} (x_i - \overline{x})2$$

$$s^2 = \frac{1}{16} \times [56.8 - 17(3.34)^2] \approx 0.502 \therefore s = \sqrt{0.502} \approx 0.71$$

Questão 10. ?

Questão 11. Os dados abaixo representam as vendas semanais, em classes de salários mínimos, de vendedores de gêneros alimentícios:

(a) Faça o histograma dos dados.

?

(b) Determine uma média amostral.

A média amostral nesse caso será dada pela média ponderada (sendo o **peso as frequências de cada classe**) dos pontos médios das classes, logo:

$$\overline{x} = \frac{\sum\limits_{i=1}^{k} n_i x_i}{n}, \text{onde } x_i \text{ é o ponto } \textbf{médio da classe}$$
$$= \frac{\sum\limits_{i=1}^{8} n_i x_i}{200} = \frac{(2 \times 32.5) + (10 \times 37.5) + (18 \times 42.5) + (50 \times 47.5) + \cdots + (2 \times 67.5)}{200} =$$
$$\overline{x} = \frac{10240}{200} = 51.2$$

(c) Prove que $\sum_{i=1}^{k} n_i(x_i - \overline{x})^2 = \sum_{i=1}^{k} n_i x_i^2 - n\overline{x}^2$, em que k é o número de classes e n_i a frequência simples da classe i.

Proof. Dado que:

$$\sum_{i=1}^{k} n_i(x_i - \overline{x})^2$$

Podemos expandir, assim:

$$\sum_{i=1}^{k} n_i(x_i^2 - 2x_i\overline{x} + \overline{x}^2) =$$
$$\sum_{i=1}^{k} (n_i x_i^2 - n_i 2x_i\overline{x} + n_i\overline{x}^2)$$
$$\sum_{i=1}^{k} n_i x_i^2 - 2\overline{x}\sum_{i=1}^{k} n_i x_i + \overline{x}^2 \sum_{i=1}^{k} n_i$$

Dado que a **frequência total é a soma das frequências**, ou seja, $\sum_{i=1}^{k} n_i = n$, temos que:

$$\sum_{i=1}^{k} n_i x_i^2 - 2\overline{x}\sum_{i=1}^{k} n_i x_i + \overline{x}^2 n$$

Outro fator é que temos que a **média em distribuições de frequência por classe** é dada por $\overline{x} = \frac{\sum_{i=1}^{k} n_i x_i}{n}$, sendo x_i a média da classe i, n_i a frequência da classe i e n a frequência total, logo:

$$\overline{x} = \frac{\sum_{i=1}^{k} n_i x_i}{n}$$
$$\therefore n\overline{x} = \sum_{i=1}^{k} n_i x_i$$

Assim, temos que:

$$\sum_{i=1}^{k} n_i x_i^2 - 2\overline{x} n \overline{x} + \overline{x}^2 n =$$
$$\sum_{i=1}^{k} n_i x_i^2 - 2\overline{x}^2 n + \overline{x}^2 n =$$
$$\sum_{i=1}^{k} n_i x_i^2 - n\overline{x}^2$$

□

(d) Utilize a informação da parte (c) para determinar o desvio-padrão[8] amostral (s).

Dado que o **desvio-padrão amostral** é dado por $s = \sqrt{\frac{1}{n-1} \times \sum_{i=1}^{k} n_i (x_i - \overline{x})^2}$, temos que:

$$s = \sqrt{\frac{1}{n-1} \times \left[\sum_{i=1}^{k} n_i (x_i - \overline{x})^2 \right]} =$$
$$= \sqrt{\frac{1}{n-1} \times \left[\sum_{i=1}^{k} n_i x_i^2 - n\overline{x}^2 \right]}$$

Assim, com a fórmula atualizada façamos:

$$s = \sqrt{\frac{1}{199} \times \left[\sum_{i=1}^{8} n_i x_i^2 - 200 \times 51.2^2 \right]}$$
$$= \sqrt{\frac{1}{199} \times \left[((2 \times 32.5^2) + (10 \times 37.5^2) + (18 \times 42.5^2) + \cdots + (2 \times 67.5^2)) - 200 \times 51.2^2 \right]} =$$
$$= \sqrt{\frac{1}{199} \times [533050 - 524288]} =$$
$$= \sqrt{44.03} =$$
$$s \approx 6.63$$

[8] Dado a variância, um valor normalmente expresso em casas ao quadrado, utilizamos o desvio-padrão que **faz a transformação dessa unidade (ao quadrado) para a original**

(e) Qual a porcentagem das observações compreendidas entre $\overline{x} - 2s$ e $\overline{x} + 2s$?

Ambos serão interpretados como quantis, assim $Q_1 = \overline{x} - 2s$ e $Q_2 = \overline{x} + 2s$ e assim temos $Q_1 = 37.94$ e $Q_2 = 64.46$.

A partir disso podemos utilizar a **fórmula de quantis em distribuições de frequência por classe**, isto é $Q_p = l_i + \frac{h(p - F_{i-1})}{f_i}$[9], para isso é interessante ter a tabela com as informações da **frequência relativa e relativa acumulada**, assim:

Classe	freqAcum (N)	freqReq (f)	freqRelAcum (F)
1	2	0.01	0.01
2	12	0.05	0.06
3	30	0.09	0.15
4	80	0.25	0.40
5	150	0.35	0.75
6	180	0.15	0.90
7	198	0.09	0.99
8	200	0.01	1.00

Dadas as informações, então temos:

$$\begin{aligned} Q_1 &= l_i + \frac{h(p_1 - F_{i-1})}{f_i} = \\ 37.94 &= 35 + \frac{5(p_1 - 0.01)}{0.05} = \\ \therefore p_1 &= 0.04 \end{aligned}$$

$$\begin{aligned} Q_2 &= l_i + \frac{h(p_2 - F_{i-1})}{f_i} = \\ 64.46 &= 60 + \frac{5(p_2 - 0.90)}{0.09} = \\ \therefore p_2 &= 0.98 \end{aligned}$$

Assim o resultado será $p_2 - p_1 = 0.98 - 0.04 = 0.94$

[9] p sendo a **proporção de observações menores ou iguais ao quantil desejado**

Questão 12. O n´umero de div´orcios em uma determinada cidade, de acordo com a dura¸c˜ao do casa- mento, est´a representado na tabela abaixo.

Questão 13. Foram obtidas observa¸c˜oes sobre a resist^encia ‘a explos˜ao (lb/in2) de soldas de fechamento de bocais de teste e soldas de bocais de can´ısteres de produ¸c˜ao:

Questão 14. As durações (em horas de uso contínuo) de 25 componentes eletrônicos selecionados de um lote de produção são:

Questão 15. A seguir temos a distribuição das estaturas de 100 alunos de uma classe.

(a) A estatura média.

$$
\begin{aligned}
\overline{X} &= \frac{\sum_{i=1}^{n} x_i}{n} ; x_i \text{ sendo o ponto médio de cada classe} \\
&= \frac{1}{100}\Bigg[((1.40+1.50)\times 0.5)+((1.50+1.60)\times 0.5)+((1.60+1.70)\times 0.5) \\
&+((1.70+1.80)\times 0.5)+((1.80+1.90)\times 0.5)+((1.90+2.00)\times 0.5)\Bigg] \\
&= \frac{1}{100}\Bigg[1.45+1.55+1.65+1.75+1.85+1.95\Bigg] = \frac{10.20}{100} = 1.62
\end{aligned}
$$

(b) A estatura modal (utilizando o método de Czuber[10]).

Para calcularmos a moda utilizando o método de Czuber, temos que:

$$M_o = l_i + \frac{h(n_i - n_{i-1})}{(n_i - n_{i-1}) + (n_i - n_{i+1})}$$

Sendo a maior frequência 40, então o intervalo modal[11] será $1.70 \vdash 1.80$, assim teremos:

$$M_o = 1.70 + \frac{0.10(40-30)}{(40-30)+(40-10)} = 1.70 + \frac{1}{40} = 1.73$$

(c) A estatura mediana.

Nesse caso, por se tratar de uma distribuição de frequência com classes, temos que:

$$M_d = l_i + h \times \left(\frac{\frac{\sum f_i}{2} - F_{i-1}}{n_i} \right)$$

Para isso façamos a relação das frequências relativas e relativa acumulada:

Classe	freqAcum	freqReq	freqRelAcum
1	5	0.05	0.05
2	15	0.10	0.15
3	45	0.30	0.45
4	85	0.40	0.85
5	95	0.10	0.95
6	100	0.05	1.00

A partir dessa tabela podemos perceber que nossa classe da mediana será da classe de número 4, ou seja $1.70 \vdash 1.80$, assim temos:

$$
\begin{aligned}
M_d &= 1.70 + 0.10 \times \frac{\frac{5+10+30+40+10+5}{2} - 45}{40} = \\
&= 1.70 + 0.10 \times \frac{50-45}{40} = \\
&= 1.70 + 0.10 \times \frac{5}{40} = 1.70 + 0.0125 = 1.71
\end{aligned}
$$

Questão 16. Em uma granja foi observada a distribuição dos frangos em relação ao peso, que era a seguinte

(a) Qual a média da distribuição?

Por se tratar de uma **distribuição por classe**, temos que a média é dada por:

$$\overline{X} = \frac{\sum_{i=1}^{k} n_i \times x_i}{n}, \text{onde } x_i \text{ é } \textbf{média da classe modal.}$$
$$= \frac{(60 \times 970) + (160 \times 990) + (280 \times 1010) + (260 \times 1030) + (160 \times 1050) + (80 \times 1070)}{1000} =$$
$$\overline{X} = \frac{1020800}{1000} = 1020.8$$

(b) Qual a variância[12] da distribuição?

$$S^2 = \frac{1}{n-1}\left[\sum_{i=1}^{k} n_i(x_i - \overline{x})^2\right], \text{onde } x_i \text{ é } \textbf{média da classe modal.}$$

(c) Queremos dividir os frangos em quatro categorias, em relação ao peso, de modo que:

- os 20% mais leves sejam da categoria D;
- os 30% seguintes sejam da categoria C;
- os 30% seguintes sejam da categoria B;
- os 20% seguintes (ou seja, os 20% mais pesados) sejam da categoria A.

Quais os limites de peso entre as categorias A, B, C e D?

Essa pergunta está relacionada a parte de **quantis**, para facilitar teremos de calcular a frequência relativa e acumulada, o que teremos:

	FreqRel	FreqAcum
1	0.06	0.06
2	0.16	0.22
3	0.28	0.50
4	0.26	0.76
5	0.16	0.92
6	0.08	1.00

Para encontrar os limites de peso vamos utilizar a fórmula de quantis em distribuições com classe, isto é:

(d) O granjeiro decide separar desse lote os animais com peso inferior a dois desvios padrões abaixo da média para receberem ração reforçada, e também separar os ani- mais com peso superior a um e meio desvio padrão acima da média para usá-los como reprodutores. Qual a porcentagem de animais que serão separados em cada caso?

?

[10]Uma aproximação mais refinada, onde o ponto da moda que divide o intervalo em duas partes é proporcional a diferenças entre a frequência da classe modal e as respectivas classes adjacentes

[11]Classe de maior frequência

[12]Motivo pelo qual dividimos por $n-1$.