DEPARTAMENTO DE MATEMÁTICA
MÉTODOS ESTATÍSTICOS
2.º SEMESTRE - 2022/2023
1.º Teste (recuperação)

Data: 11 de julho de 2023 Duração: 2 horas

Resolução

O teste foi resolvido recorrendo ao software R: ver script_1TesteRecup_ME_22_23.

1. (a) População: todos os cereais de pequeno almoço americanos

Unidade estatística: cereal de pequeno almoço americano

Variável estatística: calories
Dados estatísticos: 70, 120,...
Classificação: Quantitativa contínua (*)

Variável estatística: carbo
Dados estatísticos: 5, 8,...
Classificação: Quantitativa contínua (*)

Variável estatística: shelf
Dados estatísticos: 1, 2, 3 (códigos)
Classificação: Qualitativa ordinal

Variável estatística: rating
Dados estatísticos: 68.4, 34,...
Classificação: Quantitativa contínua

(*) foi recolhida como Quantitativa discreta.

(b) Como o gráfico refere-se a <u>todos</u> os cereais de pequeno almoço americanos, então a informação é sobre a População. A marca que mais cereais de pequeno almoço comercializa tem 60 tipos de cereais americanos, logo como a

$$\text{Moda} = K$$

pois apresenta a maior percentagem (30%), então

$$\text{Dimensão da População} = \frac{60 \times 100}{30} = 200 \text{ cereais}$$

Como o ficheiro só tem os cereais de pequeno almoço americanos utilizados no estudo (comercializados pelo hipermercado), então a informação é sobre a amostra recolhida, logo

$$\text{Dimensão da Amostra} = n = 77 \text{ cereais}$$

(c) Tabela de frequências:

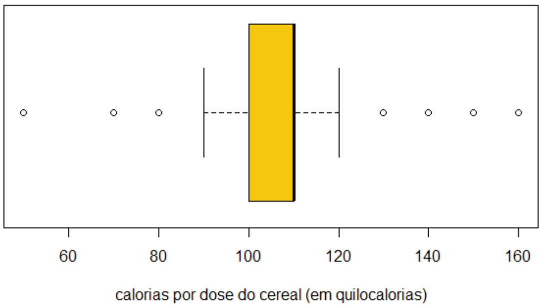
i	Prateleira x_i	Freq. Absoluta n_i	Freq. Relativa f_i	Freq. Absoluta Acum. N_i	Freq. Relativa Acum. F_i
1	mais baixa	20	0.25974	20	0.25974
2	do meio	21	0.27273	41	0.53247
3	mais alta	36	0.46753	77	1
		$n = 77$	1		

Como os dados são qualitativos, a única medida adequada é a Moda:

$$\text{moda} = \text{prateleira mais alta}$$

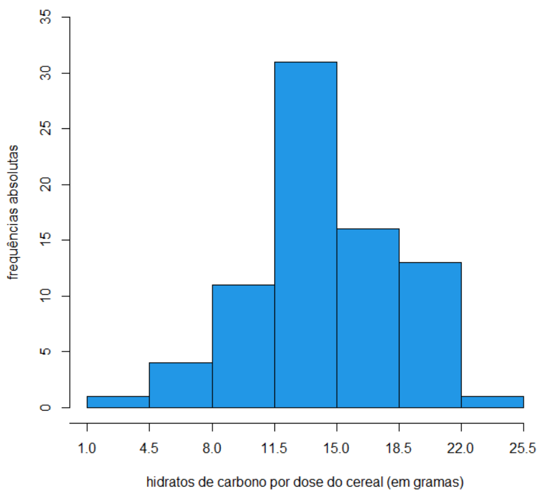
(pois $n_3 = 36$ é a maior frequência absoluta).

(d) Diagrama de extremos e quartis das calorias por dose de cereal com indicação dos *outliers* (moderados e severos)



No total há 14 dados considerados *outliers*, 6 pelo limite inferior e 8 pelo limite superior. Dos 14 dados considerados *outliers*, 6 podem ser classificados como *outliers* severos (3 pelo limite inferior e 3 pelo limite superior) e 8 podem ser classificados como *outliers* moderados (3 pelo limite inferior e 5 pelo limite superior).

(e) Histograma



O histograma aparenta que os dados se distribuem de forma simétrica, com uma ligeira assimetria negativa (ou enviesada para a esquerda).

(f) Como os dados têm unidades de medidas diferentes, a única medida de dispersão adequada é o coeficiente de variação:

$$CV_{\text{calories}} = 18.2294\% \qquad \text{e} \qquad CV_{\text{carbo}} = 28.6722\%$$

Como $CV_{\text{calories}} < CV_{\text{carbo}}$, os dados das calorias por dose de cereal apresentam menor dispersão do que os dados dos hidratos de carbono por dose de cereal.

2. Seja X a variável aleatória discreta, que representa a procura diária de um dado modelo de *smartphone* na loja I-TEL.

(a)

$$\begin{aligned} P\left(X < 3 \,|\, X \geq 1\right) &= \frac{P(X<3 \wedge X \geq 1)}{P(X \geq 1)} = \frac{P(1 \leq X < 3)}{1 - P(X<1)} = \frac{P(0 < X \leq 2)}{1 - P(X \leq 0)} \\ &= \frac{F(2) - F(0)}{1 - F(0)} = \frac{\frac{7}{8} - \frac{1}{4}}{1 - \frac{1}{4}} = \frac{5}{6} \end{aligned}$$

(b) Para determinar a variância é necessário calcular a função probabilidade. Temos que:

$f(0) = F(0) = \frac{1}{4} = 0.25$
$f(1) = F(1) - F(0) = \frac{3}{8} - \frac{1}{4} = \frac{1}{8} = 0.125$
$f(2) = F(2) - F(1) = \frac{7}{8} - \frac{3}{8} = \frac{1}{2} = 0.5$
$f(3) = F(3) - F(2) = 1 - \frac{7}{8} = \frac{1}{8} = 0.125$

Assim:

x	0	1	2	3
$f(x)$	0.25	0.125	0.5	0.125

Pretende-se calcular $V[X] = E\left[X^2\right] - E^2[X]$
onde

$$E[X] = 0 \times 0.25 + 1 \times 0.25 + 2 \times 0.5 + 3 \times 0.125 = 1.5$$

e

$$E\left[X^2\right] = 0^2 \times 0.25 + 1^2 \times 0.25 + 2^2 \times 0.5 + 3^2 \times 0.125 = 3.25$$

Logo

$$V[X] = 3.25 - (1.5)^2 = 1$$

(c) Seja Y = remuneração mensal (em unidades monetárias) dos vendedores da I-TEL, uma variável aleatória contínua. Como f é uma função densidade de probabilidade, então $f(y) \geq 0, \forall y$ e $\int_{-\infty}^{+\infty} f(y)\,dy = 1$.
Portanto

$$\int_{-\infty}^{+\infty} f(y)\,dy = 1 \Leftrightarrow \int_{10}^{20} \frac{y^2}{7000}dy + \int_{10}^{20} k dy = 1 \Leftrightarrow \frac{1}{7000}\left[\frac{y^3}{3}\right]_{10}^{20} + k\,[y]_{20}^{30} = 1 \Leftrightarrow$$

$$\Leftrightarrow \frac{1}{7000}\left(\frac{20^3}{3} - \frac{10^3}{3}\right) + k\,(30-20) = 1 \Leftrightarrow \frac{1}{3} + 10k = 1 \Leftrightarrow k = \frac{1}{15}$$

e para $f(y) \geq 0, \forall y$ é necessário que $k \geq 0$, e tem-se $k = \frac{1}{15} \geq 0$.

Logo $k = \frac{1}{15}$ é o valor pretendido.

(d) Seja $B-$ tempo entre entradas consecutivas de clientes na loja , em minutos,
então, $B \sim Exp\,(15)$ porque $\theta = E[B] = 15$

i.

$$P\,(B \geq 40 + 10 \,|\, B \geq 10) \underset{(*)}{=} P\,(B \geq 40) = 1 - P(B < 40) \underset{\text{v.a. cont.}}{=} 0.0695$$

$(*)$ falta de memória da distribuição exponencial

ii. Considere a variável aleatória discreta $C-$ número de clientes que entram na loja, em 3 horas. Pela relação entre a distribuição Exponencial e a distribuição de Poisson tem-se

$$C \sim P\left(\lambda = \frac{t}{\theta} = \frac{3 \times 60}{15} = 12\right)$$

$$P\,(C \geq 10) = 1 - P(C < 10) \underset{\text{v.a. discreta}}{=} 1 - P(C \leq 9) = 0.7576$$

(e) Seja $L-$ tempo de produção de um lote, em dias, onde

$$L \sim N\,(50, 5)$$

i. Pretende-se determinar k (prazo de entrega), tal que:

$$P\,(L \leq k) = 0.95 \quad \Leftrightarrow F(k) = 0.95 \Leftrightarrow k = F^{-1}(0.95) \Leftrightarrow k = 58.224$$

Pelo que, para garantir que em 95% das vezes o prazo é cumprido, este terá de ser cerca de 58 dias.

ii. Seja V- número de lotes em 10, que demoram mais de 60 dias em produção.
Por outro lado,

$$p = P(\text{lote demorar mais de 60 dias}) = P(L > 60) = 1 - P(L \leq 60) = 0.0227 \text{ e } n = 10.$$

Assim $\quad V \sim B\,(10, 0.0227)$ e

$$P\,(V = 3) = 0.0012$$

(f) Seja W = distância (em quilómetros) que a empresa transportadora percorre, numa hora, para entrega das encomendas de *smartphones*. Como $W \sim U_{(30,40)}$ então

$$V[W] = \frac{(40-30)^2}{12} = \frac{25}{3}$$

$$V[T] = V[-3W+2] = (-3)^2 \times V[W] = 9 \times \frac{25}{3} = 75$$

e

$$\sigma_T = \sqrt{V[T]} = \sqrt{75} = 8.6603$$