DEPARTAMENTO DE MATEMÁTICA
ESTATÍSTICA PARA A SAÚDE
1.º Semestre - 2023/2024
1.º Teste (recuperação)

Data: 14 de fevereiro de 2024 — Duração: 2 horas

- Parte do teste foi resolvido recorrendo ao software R: ver script_1teste_recup_ES_23_24.R.

Resolução

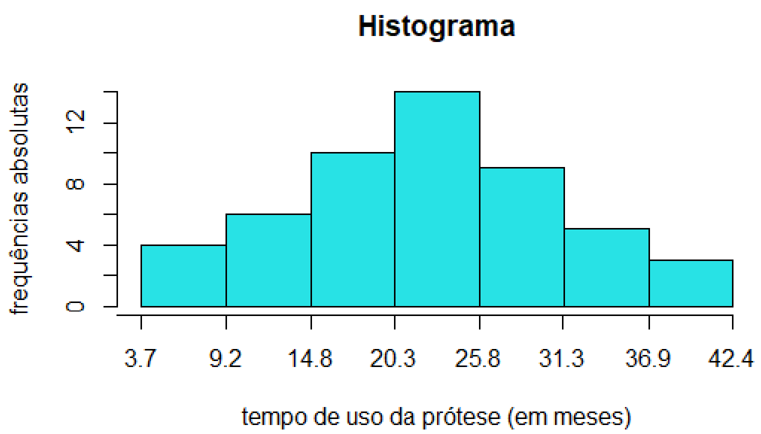
1. (a) Amostra: Alguns pacientes amputados que utilizam próteses de membros inferiores
Dimensão da Amostra: $n = 51$ pacientes
Unidade estatística: paciente amputado que utiliza próteses de membros inferiores
Variável estatística: idade. Classificação: Quantitativa contínua (mas foi recolhida como quantitativa discreta).
Variável estatística: protese. Classificação: Qualitativa nominal.
Variável estatística: tempo. Classificação: Quantitativa contínua.
Variável estatística: satisfacao. Classificação: Qualitativa ordinal.

(b) Tabela de frequências da variável Grau de satisfação:

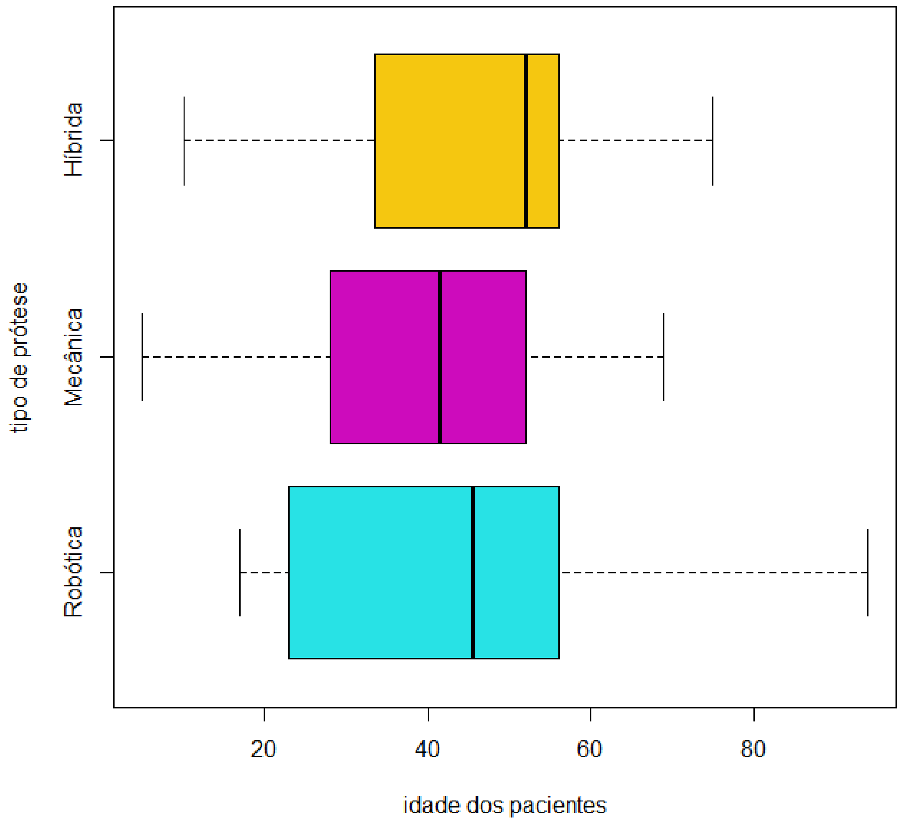
i	Grau de Satisfação x_i	Frequência Absoluta n_i	Frequência Relativa f_i	Frequência Absoluta Acumulada N_i	Frequência Relativa Acumulada F_i
1	1	6	0.1176	6	0.1176
2	2	9	0.1765	15	0.2941
3	3	7	0.1373	22	0.4314
4	4	15	0.2941	37	0.7255
5	5	14	0.2745	51	1
		$n = 51$	1		

(c) Pretende-se determinar em que valor da satisfação foi acumulada no máximo 30% das observações, pelas frequências relativas acumuladas verifica-se que foi em $F_3 = 0.4314 > 0.30$ (primeira frequência relativa acumulada que ultrapassa 0.30), ou seja, para um valor de satisfação até 3 (pois $x_3 = 3$).
Outra possibilidade de resposta é considerar a definição de quantil, ou seja, pretende-se o quantil 0.30: $Q_{0.30} = 3$.

(d) Histograma com classes fechadas à esquerda e abertas à direita definidas com base na regra de Sturges:



(e) Diagrama de extremos e quartis da idade dos pacientes por tipo de prótese:



A distribuição das idades dos pacientes apresentam diferenças quando separadas por tipo de prótese. As idades nas próteses robóticas apresentam uma clara assimetria à direita, tem os pacientes mais idosos, tendo 94 anos o mais idoso, e apresenta a maior amplitude interquartil, 29.5 anos. As idades nas próteses mecânicas têm os pacientes mais novos, tendo 5 anos o mais novo, e a "caixa com bigodes"é a que se aproxima mais da simetria. As idades nas próteses híbridas é a que apresenta a mediana mais elevada, 52 anos, e tem a menor distância entre o terceiro quartil e a mediana, 4 anos (ou seja, 25% das idades encontram-se entre os 52 anos e 56 anos). Não há idades consideradas "outliers".

(f) Como os dados têm unidades de medida diferentes, é necessário recorrer ao coeficiente de variação para comparar a dispersão:

$$CV_{idade} = 43.9\% \qquad CV_{tempo} = 37.7\%$$

A idade dos pacientes apresenta maior dispersão do que o tempo de uso das próteses, pois $CV_{idade} > CV_{tempo}$.

2. Seja X = número de produtos produzidos, por semana, uma variável aleatória discreta com $D_x = \{1, 3, 4, , 7\}$. Tem-se $E[X] = 5$ produtos por semana.

(a) Como $f(x)$ é função de probabilidade, sabe-se:

- $f(x) \geq 0, \forall x$, logo $f(1) = a \geq 0$ e $f(7) = b \geq 0$
- $\sum_x f(x) = 1$, logo

$$f(1) + f(3) + f(4) + f(5) + f(7) = 1 \Leftrightarrow a + 0.1 + 0.2 + 0.2 + b = 1 \Leftrightarrow a + b = 0.5$$

Como $E[X] = 5$ tem-se

$$1\times f(1)+3\times f(3)+4\times f(4)+5\times f(5)+7\times f(7) = 5 \Leftrightarrow 1\times a+3\times 0.1+4\times 0.2+5\times 0.2+7\times b = 5 \Leftrightarrow a+7b = 2.9$$

Juntando todos os resultados tem-se:

$$\begin{cases} a + b = 0.5 \\ a + 7b = 2.9 \\ a \geq 0 \wedge b \geq 0 \end{cases} \Leftrightarrow \begin{cases} a = 0.5 - b \\ 0.5 - b + 7b = 2.9 \\ a \geq 0 \wedge b \geq 0 \end{cases} \Leftrightarrow \begin{cases} a = 0.1 \\ b = 0.4 \\ 0.1 \geq 0 \wedge 0.4 \geq 0 \end{cases} \Leftrightarrow \begin{cases} a = 0.1 \\ b = 0.4 \end{cases}$$

(b) Com $a = 0.1$ e $b = 0.4$ tem-se

x	1	3	4	5	7
$f(x)$	0.1	0.1	0.2	0.2	0.4

$$V\left[\frac{1-2X}{3}\right] = V\left[\frac{1}{3} - \frac{2}{3}X\right] = \left(\frac{2}{3}\right)^2 \times V[X] = \frac{4}{9} \times \left(E\left[X^2\right] - E^2\left[X\right]\right)$$

Como $E\left[X\right] = 5$, falta calcular $E\left[X^2\right]$. Assim

$$E\left[X^2\right] = 1^2 \times f(1) + 3^2 \times f(3) + 4^2 \times f(4) + 5^2 \times f(5) + 7^2 \times f(7) =$$

$$= 1^2 \times 0.1 + 3^2 \times 0.1 + 4^2 \times 0.2 + 5^2 \times 0.2 + 7^2 \times 0.4 = 28.8$$

logo

$$V\left[\frac{1-2X}{3}\right] = \frac{4}{9} \times \left(28.8 - 5^2\right) = 1.6889$$

(c) Seja Y = duração, em centenas de horas, de um componente eletrónico, uma variável aleatória contínua. Sabe-se que 250 horas = 2.5 centenas de hora.

Seja W a variável aleatória discreta:

W = número de componentes, de um grupo de 10000 componentes, que possuem uma duração superior a 250 horas. $W \sim B\left(10000, 0.16\right)$ pois

$n = 10000$ componentes

$$p = P\left(\text{Sucesso}\right) = P\left(Y > 2.5\right) = 1 - P\left(Y \leq 2.5\right) = 1 - \int_{-\infty}^{2.5} f(y)dy =$$

$$= 1 - \left(\int_{-\infty}^{1} 0dy + \int_{1}^{2.5} (2y^{-3})dy\right) = 1 - \left(0 + \left[2 \times \frac{y^{-2}}{-2}\right]_1^{2.5}\right) =$$

$$= 1 - \left(2 \times \frac{2.5^{-2}}{-2} - 2 \times \frac{1^{-2}}{-2}\right) = 0.16$$

então

$$E\left[W\right] = n \times p = 10000 \times 0.16 = 1600 \text{ componentes}$$

3. Considere a variável aleatória discreta X = procura mensal do medicamento, $X \sim P(2)$, pois $E[X] = 2$ medicamentos por mês. Sabe-se que o *stock* mensal é de 5 medicamentos.

   (a) $P\left(X > 5\right) = 1 - P\left(X \leq 5\right) = 1 - F(5) = 0.0166$

   (b) Pretende-se determinar $s > 0$ tal que

$$P\left(X > s\right) < 0.01 \Leftrightarrow 1 - P\left(X \leq s\right) < 0.01 \Leftrightarrow F\left(s\right) > 0.99 \Leftrightarrow s \geq 6 \text{ medicamentos}$$

4. Qualquer pacotinho é considerado defeituoso se a sua quantidade de açúcar for inferior a 5 gramas ou superior a 8 gramas.

   (a) Considere a variável aleatória contínua X = quantidade, em gramas, de açúcar contida num pacotinho embalado por M_1, $X \sim N(7, 2)$ pois $\mu = E[X] = 7$ gramas e $\sigma = \sqrt{V[X]} = 2$ gramas.

      i. Pretende-se

$$P\left(4 \leq X \leq 8 | X \geq 5\right) = \frac{P\left(4 \leq X \leq 8 \wedge X \geq 5\right)}{P\left(X \geq 5\right)} = \frac{P\left(5 \leq X \leq 8\right)}{1 - P\left(X < 5\right)} \underset{\text{v.a. contínua}}{=}$$

$$= \frac{F(8) - F(5)}{1 - F(5)} = \frac{0.6915 - 0.1587}{1 - 0.1587} = 0.6333$$

ii. Seja W a variável aleatória discreta:

W = número de pacotinhos de açúcar, de uma caixa embalada por M_1, serem defeituosos.

$W \sim B\,(25, 0.4672)$ pois

$n = 25$ pacotinhos de açúcar numa caixa

$p = P\,(\text{Sucesso}) = P\,(\text{pacotinho defeituoso}) = P\,(X < 5 \vee X > 8) = 1 - P\,(5 \leq X \leq 8) \underset{\text{v.a. contínua}}{=}$

$= 1 - (F(8) - F(5)) = 1 - (0.6915 - 0.1587) = 0.4672$

então

$$P\,(W \geq 10) = 1 - P\,(W < 10) \underset{\text{v.a. discreta}}{=} 1 - P\,(W \leq 9) = 1 - F\,(9) = 1 - 0.1916 = 0.8084$$

(b) Seja Y = quantidade, em gramas, de açúcar contida num pacotinho embalado por M_2, tem-se $Y \sim N(\mu_2, 1.1)$ pois $\sigma = \sqrt{V[Y]} = 1.1$ gramas. Tem-se

$$Y \sim N(\mu_2, 1.1) \Leftrightarrow Z = \frac{Y - \mu_2}{1.1} \sim N(0,1)$$

Pretende-se determinar μ_2 tal que

$$P\,(Y > 6) = 0.55 \Leftrightarrow 1 - P\,(Y \leq 6) = 0.55 \Leftrightarrow P\,(Y \leq 6) = 0.45 \Leftrightarrow$$
$$\Leftrightarrow P\left(Z \leq \frac{6 - \mu_2}{1.1}\right) = 0.45 \Leftrightarrow \Phi\left(\frac{6 - \mu_2}{1.1}\right) = 0.45 \Leftrightarrow \frac{6 - \mu_2}{1.1} = z_{0.45} \Leftrightarrow$$
$$\Leftrightarrow \frac{6 - \mu_2}{1.1} = -0.1257 \Leftrightarrow \mu_2 = 6.1382 \text{ gramas}$$

Fim do Teste