

Data: 10 de maio de 2021

Duração: 2 horas

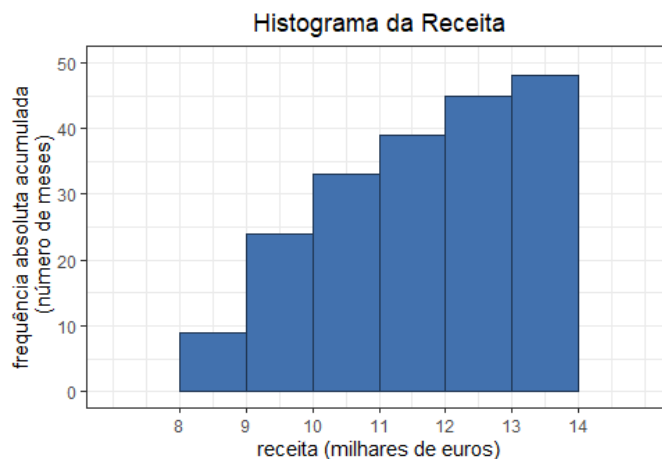
Resolução

1. Os gestores de uma empresa realizaram, ao fim de 4 anos de exercício, um balanço das receitas obtidas em cada mês (em milhares de euros), o que significa que a amostra recolhida tem dimensão $n = 4 \times 12 = 48$ meses.

(a) Variável em estudo: Receita (em milhares de euros)

Tipo variável em estudo: Variável Quantitativa Contínua

Representação gráfica:



- (b) Como "Número de meses (acumulado)" corresponde às frequências absolutas acumuladas (N_i), tem-se a seguinte tabela de frequências:

i	Receita (milhares de euros) Classe - c_i	Representante de classe x'_i	Freq. Absoluta n_i	Freq. Relativa f_i	Freq. Abs. Acumulada N_i	Freq. Rel. Acumulada F_i
1	[8, 9[$\frac{8+9}{2} = 8.5$	9	$\frac{9}{48} = 0.1875$	9	$\frac{9}{48} = 0.1875$
2	[9, 10[$\frac{9+10}{2} = 9.5$	$24 - 9 = 15$	$\frac{15}{48} = 0.3125$	24	$\frac{24}{48} = 0.5$
3	[10, 11[$\frac{10+11}{2} = 10.5$	$33 - 24 = 9$	$\frac{9}{48} = 0.1875$	33	$\frac{33}{48} = 0.6875$
4	[11, 12[$\frac{11+12}{2} = 11.5$	$39 - 33 = 6$	$\frac{6}{48} = 0.125$	39	$\frac{39}{48} = 0.8125$
5	[12, 13[$\frac{12+13}{2} = 12.5$	$45 - 39 = 6$	$\frac{6}{48} = 0.125$	45	$\frac{45}{48} = 0.9375$
6	[13, 14]	$\frac{13+14}{2} = 13.5$	$48 - 45 = 3$	$\frac{3}{48} = 0.0625$	48	$\frac{48}{48} = 1$
			$n = 48$	1		

- (c) Tendo em conta que existem 24 meses cuja receita é superior ou igual a 10 milhares de euros ($n_6 + n_5 + n_4 + n_3 = 3 + 6 + 6 + 9 = 24$) e que 15 meses têm a receita superior ou igual a 11 milhares de euros ($n_6 + n_5 + n_4 = 3 + 6 + 6 = 15$), então a receita mensal em relação ao qual existem 18 meses em que os rendimentos mensais são superiores ou iguais a esse valor é obtida a partir da classe $c_3 = [10, 11[$. Portanto apenas é possível dizer que 18 meses terão a receita superior ou igual a 10.5 milhares de euros (representante da classe).

Ou de forma equivalente:

Como no total tem-se 48 meses, então a receita mensal em relação ao qual existem 18 meses em que os rendimentos mensais são superiores ou iguais a esse valor é obtida a partir da classe onde a frequência relativa acumulada atinge os $48 - 18 = 30$ meses, ou seja, da classe $c_3 = [10, 11[$. Como $F_3 = 33 > 30$, apenas é possível dizer que 18 meses terão a receita superior ou igual a 10.5 milhares de euros (representante da classe).

- (d) Pretende-se determinar $Q_{0.10}$.

Como $F_1 = 0.1875 > 0.10$, ou seja, a primeira frequência relativa acumulada que ultrapassa 0.10 é F_1 , então o valor pretendido é atingido na classe $c_1 = [8, 9[$, logo

$$Q_{0.10} \approx x'_1 = 8.5 \text{ milhares de euros}$$

- (e) Medidas de localização central: média, mediana e moda. Como os dados estão agrupados em classes, os valores das medidas de localização serão calculados com base no representante da classe.

- MÉDIA

usando as frequências absolutas:

$$\bar{x} \approx \frac{9 \times 8.5 + 15 \times 9.5 + 9 \times 10.5 + 6 \times 11.5 + 6 \times 12.5 + 3 \times 13.5}{48} = 10.375 \text{ milhares de euros}$$

Alternativamente poderíamos calcular usando as frequências relativas:

$$\begin{aligned} \bar{x} &\approx 0.1875 \times 8.5 + 0.3125 \times 9.5 + 0.1875 \times 10.5 + 0.125 \times 11.5 + 0.125 \times 12.5 + 0.0625 \times 13.5 = \\ &= 10.375 \text{ milhares de euros} \end{aligned}$$

- MEDIANA

Como $F_2 = 0.50$, então o valor pretendido é o limite superior da classe $c_2 = [9, 10[$

$$\tilde{x} = Q_{0.50} \approx 10 \text{ milhares de euros}$$

- MODA

A classe modal é $c_2 = [9, 10[$ pois é a classe com maior frequência absoluta ($n_2 = 15$)

$$moda \approx x'_2 = 9.5 \text{ milhares de euros}$$

por ser o valor central da classe modal.

Atendendo que

$$moda < mediana < média$$

podemos dizer que a distribuição dos dados é assimétrica positiva (ou enviesada para a direita).

2. Pretende-se investigar se há relação entre a pulsação e a altura de um humano e tem-se uma amostra de dimensão $n = 50$.

(a) Variável independente = X – altura (em centímetros)

Variável dependente = Y – valor da pulsação (em batimentos do coração por minuto)

A reta de regressão: $\hat{y} = a + bx$

em que:

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} = \frac{\sum_{i=1}^{50} x_i}{50} = \frac{8436}{50} = 168.72 \quad \text{e} \quad \bar{y} = \frac{\sum_{i=1}^n y_i}{n} = \frac{\sum_{i=1}^{50} y_i}{50} = \frac{4115}{50} = 82.3$$

$$b = \frac{\sum_{i=1}^n x_i y_i - n \bar{x} \bar{y}}{\sum_{i=1}^n x_i^2 - n \bar{x}^2} = \frac{\sum_{i=1}^{50} x_i y_i - 50 \times \bar{x} \bar{y}}{\sum_{i=1}^{50} x_i^2 - 50 \times \bar{x}^2} = \frac{695171 - 50 \times 168.72 \times 82.3}{1427556 - 50 \times 168.72^2} = 0.21$$

$$a = \bar{y} - b \bar{x} = 82.3 - 0.21 \times 168.72 = 46.87$$

Então a reta de regressão:

$$\hat{y} = 46.87 + 0.21x$$

Previsão:

$$\hat{y} = 46.87 + 0.21 \times 170 = 82.57$$

ou seja, uma pessoa com uma altura de 170 cm terá uma pulsação prevista de 82.57 batimentos do coração por minuto.

- (b) Como não temos os dados não é possível fazer o diagrama de dispersão, então para analisar a correlação linear existente vamos apenas calcular o coeficiente de correlação linear de Pearson:

$$\begin{aligned} r_{XY} &= \frac{\sum_{i=1}^n x_i y_i - n \bar{x} \bar{y}}{\sqrt{\left(\sum_{i=1}^n x_i^2 - n \bar{x}^2\right) \left(\sum_{i=1}^n y_i^2 - n \bar{y}^2\right)}} = \frac{\sum_{i=1}^{50} x_i y_i - 50 \times \bar{x} \bar{y}}{\sqrt{\left(\sum_{i=1}^{50} x_i^2 - 50 \times \bar{x}^2\right) \left(\sum_{i=1}^{50} y_i^2 - 50 \times \bar{y}^2\right)}} = \\ &= \frac{695171 - 50 \times 168.72 \times 82.3}{\sqrt{(1427556 - 50 \times 168.72^2) (342577 - 50 \times 82.3^2)}} = 0.218 \end{aligned}$$

concluimos que existe uma correlação linear positiva fraca entre as variáveis, pois $0 < r_{XY} < 0.5$, pelo que a previsão efetuada é pouco fiável.

3. Seja X a variável aleatória discreta que representa o número de peças defeituosas produzidas por uma máquina em período experimental, o seu domínio é $D_X = \{1, 4, 6, 8\}$.

(a) Pretende-se

$$\begin{aligned} P(X \geq 4 | X \leq 6) &= \frac{P(X \geq 4 \wedge X \leq 6)}{P(X \leq 6)} = \frac{P(4 \leq X \leq 6)}{F(6)} \stackrel{\text{v.a. discreta}}{=} \frac{P(1 < X \leq 6)}{0.8} = \\ &= \frac{F(6) - F(1)}{0.8} = \frac{0.8 - 0.5}{0.8} = 0.375 \end{aligned}$$

(b) A função de probabilidade da variável aleatória X é

x	1	4	6	8
$f(x)$	0.5	0.2	0.1	0.2

pois

$$\begin{aligned}f(1) &= F(1) = 0.5 \\f(4) &= F(4) - F(1) = 0.7 - 0.5 = 0.2 \\f(6) &= F(6) - F(4) = 0.8 - 0.7 = 0.1 \\f(8) &= F(8) - F(6) = 1 - 0.8 = 0.2\end{aligned}$$

então

$$\begin{aligned}E[X] &= 1 \times f(1) + 4 \times f(4) + 6 \times f(6) + 8 \times f(8) = \\&= 1 \times 0.5 + 4 \times 0.2 + 6 \times 0.1 + 8 \times 0.2 = 3.5 \text{ peças defeituosas} \\E[X^2] &= 1^2 \times f(1) + 4^2 \times f(4) + 6^2 \times f(6) + 8^2 \times f(8) = \\&= 1^2 \times 0.5 + 4^2 \times 0.2 + 6^2 \times 0.1 + 8^2 \times 0.2 = 20.1 \\V[X] &= E[X^2] - E^2[X] = 20.1 - 3.5^2 = 7.85 \text{ (peças defeituosas)}^2 \\V[7 - 2X] &= (-2)^2 \times V[X] = 4 \times 7.85 = 31.4\end{aligned}$$

(c) Seja Y a variável aleatória discreta que representa o número de períodos experimentais que resultaram em máquinas devolvidas, de um grupo de 12 períodos experimentais, $Y \sim B(12, 0.5)$ pois

$$\begin{aligned}n &= 12 \text{ períodos experimentais} \\p &= P(X \geq 2) = 1 - P(X < 2) \underset{\text{v.a. discreta}}{=} 1 - P(X \leq 1) = 1 - F(1) = 1 - 0.5 = 0.5\end{aligned}$$

Pretende-se

$$P(Y < 3) \underset{\text{v.a. discreta}}{=} P(Y \leq 2) = F(2) = 0.0193$$

(d) Seja W a variável aleatória contínua que representa o tempo, em minutos, que uma destas máquinas demora a ser reparada

Como 1 hora = 60 minutos, pretende-se

$$\begin{aligned}P(W > 60) &= \int_{60}^{+\infty} f(w)dw = \int_{60}^{80} \left(\frac{1}{40} - \frac{w}{3200} \right) dw + \int_{80}^{+\infty} 0dw = \\&= \left[\frac{1}{40} \times w - \frac{1}{3200} \times \frac{w^2}{2} \right]_{60}^{80} + 0 = \left(\frac{1}{40} \times 80 - \frac{1}{3200} \times \frac{80^2}{2} \right) - \left(\frac{1}{40} \times 60 - \frac{1}{3200} \times \frac{60^2}{2} \right) = \\&= 0.0625\end{aligned}$$

4. Tem-se

X_A a variável aleatória discreta que representa o número de embalagens do modelo A vendidas por dia, tal que $X_A \sim P(4)$ pois $E[X_A] = 4 = \lambda_A$;

X_B a variável aleatória discreta que representa o número de embalagens do modelo B vendidas por dia, tal que $X_B \sim P(9)$ pois $\sqrt{V[X_B]} = 3 \Leftrightarrow V[X_B] = 3^2 = 9 = \lambda_B$;

X_A e X_B são variáveis aleatórias independentes;

A loja está aberta todos os dias da semana, 8 horas por dia.

- (a) Seja $Y = X_A + X_B$ a variável aleatória discreta que representa o número de embalagens de *smartphones* vendidas por dia.

Como X_A e X_B são variáveis aleatórias independentes e ambas com distribuição de Poisson, então pela Aditividade da distribuição de Poisson tem-se $Y \sim P(13)$ pois $\lambda_Y = \lambda_A + \lambda_B = 4 + 9 = 13$

Pretende-se

$$P(Y = 15) = f(15) = 0.0885$$

- (b) Seja W_B a variável aleatória discreta que representa o número de embalagens do modelo B vendidas em 2 horas, $W_B \sim P(2.25)$ pois

$$\begin{aligned} 1 \text{ dia} = 8 \text{ horas} &\mapsto \lambda_B = 9 \\ 2 \text{ horas} &\mapsto \lambda_{W_B} = \frac{2 \times 9}{8} = 2.25 \end{aligned}$$

Pretende-se

$$E[W_B] = \lambda_{W_B} = 2.25 \text{ embalagens do modelo B}$$

- (c) Seja T_A a variável aleatória contínua que representa o tempo, em dias, entre vendas de 2 embalagens do modelo A, recorrendo à relação entre a distribuição de Poisson e a distribuição Exponencial tem-se $T_A \sim \text{Exp}(0.25)$, pois $\theta = \frac{1}{4} = 0.25$. A função de distribuição da variável aleatória T_A é dada por

$$F(t) = \begin{cases} 0 & , \quad t < 0 \\ 1 - e^{-\frac{t}{0.25}} & , \quad t \geq 0 \end{cases}.$$

Pretende-se

$$P(T_A < 3) \underset{\text{v.a. contínua}}{=} F(3) = 1 - e^{-\frac{3}{0.25}} = 1$$

- (d) Seja V a variável aleatória contínua que representa o peso, em gramas, da bateria do *smartphone*, tem-se $V \sim U_{(44,56)}$ pois os limites do intervalo são $a = 44$ e $b = 56$. A função de distribuição da variável aleatória V é dada por

$$F(v) = \begin{cases} 0 & , \quad v < 44 \\ \frac{v-44}{56-44} & , \quad 44 \leq v \leq 56 \\ 1 & , \quad v > 56 \end{cases} \Leftrightarrow F(v) = \begin{cases} 0 & , \quad v < 44 \\ \frac{v-44}{12} & , \quad 44 \leq v \leq 56 \\ 1 & , \quad v > 56 \end{cases}$$

Pretende-se determinar m tal que

$$\begin{aligned} P(V \geq m) = 0.80 &\Leftrightarrow 1 - P(V < m) = 0.80 \Leftrightarrow P(V < m) = 0.20 \underset{\text{v.a. contínua}}{\Leftrightarrow} F(m) = 0.20 \Leftrightarrow \\ &\underset{(*)}{\Leftrightarrow} \frac{m-44}{12} = 0.20 \Leftrightarrow m = 46.4 \text{ gramas} \end{aligned}$$

(*) Só pode ser em $44 \leq m \leq 56$ pois nos outros tem-se:

$$\begin{aligned} \text{se } m < 44, & \quad F(m) = 0 \neq 0.20 \\ \text{se } m > 56, & \quad F(m) = 1 \neq 0.20 \end{aligned}$$