21037 : e-fólio A- proposta de resolução

1. Os montantes de depósito a prazo, em unidades codificadas (UC), correspondem a uma variável quantitativa contínua, e estão organizados em classes com a mesma amplitude.

O número total de observações (N) calcula-se com base na frequência absoluta(n_i) de cada classe:

$$N = n_1 + n_2 + \cdots + n_{10} = 1860, i = \{1, 2, \ldots 10\}$$

Nota: como todas as classes têm a mesma amplitude (h_i) não é necessário calcular as frequências corrigidas.

As frequências absolutas acumuladas (N_i) calculam-se adicionando as frequências absolutas até à classe i inclusive. Exemplo: $N_7 = n_1 + n_2 + \cdots + n_7 = N_6 + n_7 = 1280$.

Para cada classe i calculam-se as frequências relativas: $f_i = \frac{n_i}{N}$. Exemplo: $f_4 = \frac{n_4}{1280} = \frac{160}{1280} = 0{,}08602$.

As frequências relativas acumuladas (F_i) calculam-se adicionando as frequências relativas até à classe i inclusive. Exemplo: $F_3 = f_1 + f_2 + f_3 = F_2 + f_3 = 0{,}14517$.

Tabela de frequências absolutas e relativas, simples e acumuladas:

i	Classes	x_i'	n_i	N_i	f_i	F_i	$x_i' \times f_i$
1	[0, 50[	25	74	74	0,03979	0,03979	0,994624
2	[50, 100[	75	90	164	0,04839	0,08818	3,629032
3	[100, 150[	125	106	270	0,05699	0,14517	7,123656
4	[150, 200[	175	160	430	0,08602	0,23119	15,05376
5	[200, 250[	225	198	628	0,10645	0,33764	23,95161
6	[250, 300[	275	252	880	0,13548	0,47312	37,25806
7	[300, 350[	325	400	1280	0,21505	0,68817	69,89247
8	[350, 400[	375	380	1660	0,20430	0,89247	76,6129
9	[400, 450[	425	140	1800	0,07527	0,96774	31,98925
10	[450, 500[	475	60	1860	0,03226	1,00000	15,32258
		Total	1860		1,00000		281,828

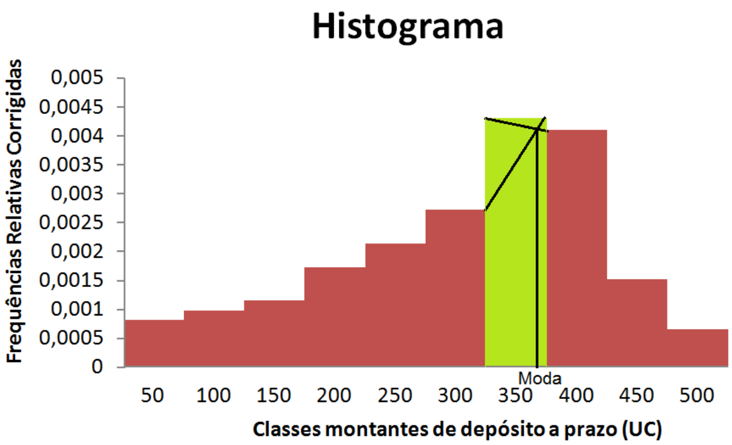
x_i' representa o centro da classe, ou seja, na classe [a, b[o centro da classe calcula-se $x_i' = \frac{a_i + b_i}{2}$. Exemplo: $x_9' = \frac{a_9 + b_9}{2} = \frac{400+450}{2} = 425$.

1.1

Representação da medida de localização central moda num histograma:

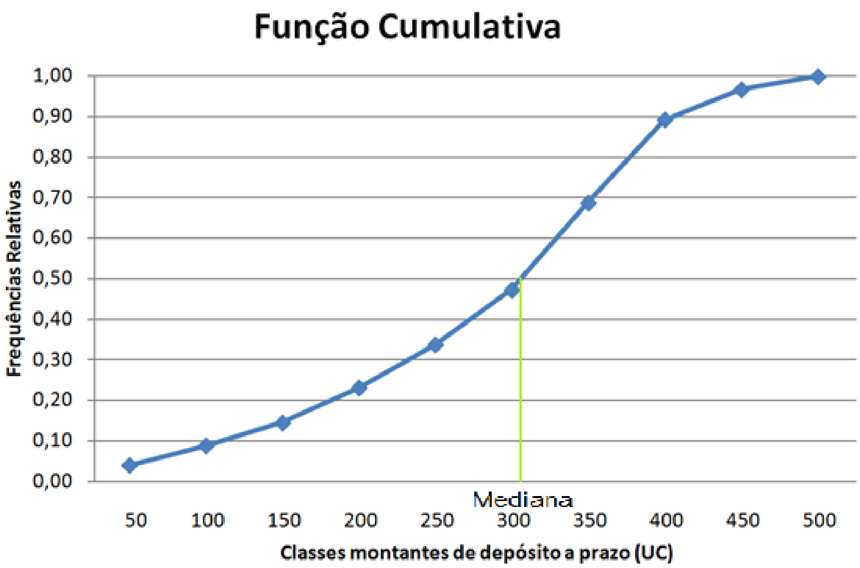
Consideram-se as frequências relativas corrigidas para que a área representada seja igual à unidade.

A verde está representada a classe modal.



Representação da medida de localização central mediana num gráfico de função cumulativa.

A mediana está representada a verde, é o valor em que 50% das observações são inferiores ou iguais e 50% das observações são superiores ou iguais.



1.2

Média:

$$\bar{x} = \sum_{i=1}^{k} (f_i x_i') = \frac{\sum_{i=1}^{k} (n_i x_i')}{n} = \sum_{i=1}^{10} (f_i x_i') = 281{,}828$$

Média – é a medida de localização mais usual, é o valor para o qual a soma dos desvios é zero; a média é pouco resistente, ou seja é sensível a valores muito altos ou muito baixos (outliers), a média é uma boa medida de localização quando a distribuição é simétrica e não se verificam outliers, pois a média é calculada com base em todos os dados.

A mediana encontra-se na 7ª classe com o limite inferior (a_i) igual a 300, uma amplitude igual a 50 e um frequência relativa igual a 0,21505; a frequência relativa acumulada da classe anterior é igual a 0,47505. Aplicando a fórmula, conclui-se que a mediana é igual a 306,25.

$$Me = a_i + \frac{0{,}5 - f_{i-1}}{fi} \times (b_i - a_i) = 300 + \frac{0{,}5 - 0{,}47312}{0{,}21505} \times (300 - 350) = 306{,}25$$

Mediana – é o valor que ocupa o lugar central da distribuição, quando considerado o número total de dados ordenados por ordem crescente, tendo como vantagem a sua resistência a valores extremos, é a medida de localização que melhor representa um conjunto de dados quando a amplitude é muito grande.

A moda, neste caso, situa-se na 7ª classe com o limite inferior (a_i) igual a 300, uma amplitude igual a 50 e as frequências relativas das classes anterior e posterior são 0,13548 e 0,20430 respetivamente; aplicando a fórmula, conclui-se que a moda é igual a 330,06.

$$Mo = a_i + \frac{f_{i+1}}{f_{i-1} + f_i} \times (b_i - a_i) = 300 + \frac{0{,}20430}{0{,}13548 + 0{,}20430} \times 50 = 330{,}06$$

Moda – é o valor mais frequente de uma distribuição, aquele que aparece mais vezes, tendo como vantagens a identificação de preferências ou maiorias na variável em estudo e a sua não influencia por valores extremos; a desvantagem é que nem sempre existe, há distribuições que são amodais e a sua falta de eficiência quando as frequências absolutas de diferentes valores (ou classes) são muito próximas.

1.3

Todos os valores aumentam 100% assim a média sofre igual alteração e a variância sofre a alteração elevada ao quadrado. Seja Y=kX e n o número de classes, com i=1,...,n.

$$\bar{y} = \sum_{i=1}^{n} \left(f_i (k x_i')\right) = k \sum_{i=1}^{n} (f_i x_i') = k\bar{x}$$

$$s_y^2 = \sum_{i=1}^{n} (f_i (k x_i')^2) - (k\bar{x})^2 = \sum_{i=1}^{n} \left(f_i k^2 x_i'^2\right) - k^2 \bar{x}^2 = k^2 \left(\sum_{i=1}^{n} \left(f_i x_i'^2\right) - \bar{x}^2 \right) = k^2 s_x^2$$

Assim sendo, para o exercício em causa em que o aumento é de 100%, significa que a média aumenta para o dobro e a variância para o quádruplo.

Seria então de considerar a v.a Y, com Y=2X, sendo $\bar{y} = 2\overline{x}$ e $s_y^2 = 4 s_x^2$, ou seja, concretizando para a média, dados os resultados obtidos em 1.2, teríamos $\bar{y} = 563{,}656$.

1.4

Como a todos os valores é adicionada uma constante, a média sofre igual alteração enquanto a variabilidade dos dados permanece igual, ou seja a variância é a mesma. Seja Z=X+c

$$\bar{z} = \sum_{i=1}^{n}\left(f_i(c + x_i')\right) = \sum_{i=1}^{n}(cf_i) + \sum_{i=1}^{n}(f_i x_i') = c\sum_{i=1}^{n}(f_i) + \bar{x} = c \times 1 + \bar{x} = c + \bar{x}$$

Usando para a média o valor obtido em 1.2 teríamos,

$$\bar{z} = 50 + 281{,}828 = 331{,}828$$

Quanto à variância, verifica-se que se mantém:

$$s_Z^2 = \sum_{i=1}^{n}(f_i(c + x_i')^2) - (c + \bar{x})^2 = \sum_{i=1}^{n}\left(f_i\left(c^2 + 2cx_i' + {x_i'}^2\right)\right) - (c^2 + 2c\bar{x} + \bar{x}^2) =$$

$$= \sum_{i=1}^{n}\left(f_i c^2 + 2cf_i x_i' + f_i {x_i'}^2\right) - c^2 - 2c\bar{x} - \bar{x}^2 = c^2\sum_{i=1}^{n} f_i + 2c\sum_{i=1}^{n} f_i\, x_i' + \sum_{i=1}^{n} f_i {x_i'}^2 - c^2 - 2c\bar{x} - \bar{x}^2$$

$$= c^2\sum_{i=1}^{n} f_i - c^2 + 2c\sum_{i=1}^{n} f_i\, x_i' - 2c\bar{x} + \sum_{i=1}^{n} f_i {x_i'}^2 - \bar{x}^2 = \sum_{i=1}^{n} f_i {x_i'}^2 - \bar{x}^2 = s_x^2$$

2.1

Acontecimentos

M - Marcar

$\bar{M}$ - Não marcar

E - Rematar para a esquerda

F - Rematar para a frente

D - Rematar para a direita

$$P(M) = 0{,}90 \Leftrightarrow P(\bar{M}) = 1 - 0{,}90 = 0{,}10$$

$$P(\bar{M}|E) = 0{,}25 = \frac{P(\bar{M} \cap E)}{P(E)} \Leftrightarrow P(M|E) = 0{,}75 = \frac{P(M \cap E)}{P(E)}$$

$$P(M|F) = 0{,}95 = \frac{P(M \cap F)}{P(F)}$$

$$P(M|D) = 0{,}90 = \frac{P(M \cap D)}{P(D)}$$

Sabe-se ainda que $2P(D) = P(F)$

Probabilidade de marcar:

$$\begin{aligned} P(M) = 0{,}90 &= 0{,}75P(E) + 0{,}95P(F) + 0{,}90P(D) \\ &= 0{,}75P(E) + 0{,}95 \times 2P(D) + 0{,}90P(D) = 0{,}75P(E) + 2{,}8P(D) \end{aligned}$$

$$P(\bar{M}) = 0{,}10 = 0{,}25P(E) + 0{,}05P(F) + 0{,}10P(D)$$
$$= 0{,}25P(E) + 0{,}05 \times 2P(D) + 0{,}10P(D) = 0{,}25P(E) + 0{,}2P(D)$$

$$\begin{cases} 0{,}75P(E) + 2{,}8P(D) = 0{,}90 \\ 0{,}25P(E) + 0{,}2P(D) = 0{,}10 \\ P(E) + P(F) + P(D) = 1{,}00 \end{cases} \Leftrightarrow \begin{cases} P(E) = 0,(18) \\ P(D) = 0,(27) \\ P(F) = 2 \times 0,(27) = 0{,}5455 \end{cases}$$

A probabilidade de rematar à esquerda é 0,1818.

Sabendo que não marcou golo qual a probabilidade de ter rematado em frente:

$$P(M|F) = 0{,}95 = \frac{P(M \cap F)}{P(F)}$$

$$\Leftrightarrow P(\bar{M}|F) = 0{,}05 = \frac{P(\bar{M} \cap F)}{P(F)} \Leftrightarrow P(\bar{M} \cap F) = 0{,}05 \times 0{,}5455$$

$$P(F|\bar{M}) = \frac{P(F \cap \bar{M})}{P(\bar{M})} = \frac{0{,}05 \times 0{,}5455}{0{,}10} = 0{,}27275$$

3.

$$P(B) = P(B_1)P(B|B_1) + P(B_2)P(B|B_2) + P(B_3)P(B|B_3) + P(B_4)P(B|B_4)$$
$$= \frac{4}{22} \times 0 + \frac{5}{22} \times \frac{1}{22} + \frac{6}{22} \times \frac{4}{22} + \frac{7}{22} \times \frac{9}{22} = \frac{92}{484} = \frac{23}{121} = 0{,}19$$

$$P(B_i|B) = \frac{P(B_i)P(B|B_i)}{P(B)}$$

$$P(B_i|B) = \frac{\left(\frac{i+3}{22}\right)\left(\frac{(i-1)^2}{22}\right)}{P(B)} = \frac{\left(\frac{i+3}{22}\right)\left(\frac{i^2-2i+1}{22}\right)}{P(B)} = \frac{\frac{i^3-2i^2+i+3i^2-6i+3}{484}}{P(B)}$$
$$= \frac{i^3-2i^2+i+3i^2-6i+3}{484P(B)} = \frac{i^3+i^2-5i+3}{484P(B)}, i = \{1, 2, 3, 4\}$$

$$P(B_1|B) = \frac{i^3+i^2-5i+3}{484\frac{23}{121}} = \frac{1+1-5+3}{92} = 0$$

$$P(B_2|B) = \frac{i^3+i^2-5i+3}{484\frac{23}{121}} = \frac{8+4-10+3}{92} = \frac{5}{92}$$

$$P(B_3|B) = \frac{i^3+i^2-5i+3}{484\frac{23}{121}} = \frac{27+9-15+3}{92} = \frac{24}{92}$$

$$P(B_4|B) = \frac{i^3+i^2-5i+3}{22\frac{23}{121}} = \frac{64+16-20+3}{92} = \frac{63}{92}$$