Elementos de Probabilidade e Estatística | 21037

Proposta de resolução e critérios de correção do E-fólio A. Em determinadas questões é possível existir mais do que uma versão de resolução. Apresentamos apenas uma versão. Outras versões e interpretações, desde que corretas, foram corrigidas com os mesmos critérios da presente resolução.

Questão 1

Fez-se um inquérito aos youtubers portugueses sobre o número de vezes que, ao longo do mês de janeiro, fizeram vídeos em direto. Com os dados obtidos construiu-se o seguinte quadro:

Nº de vídeos em direto	0	1	2	3	4	6	10	15	Total
Nº de youtubers	320	200	109	A	30	30	B	23	C
Percentagem de youtubers	40	D	E	F	G	H	3	I	J

1.1 Determine os valores de $A, B, C, D, E, F, G, H, I$ e J.

Resolução:

- Claramente $J = 100\%$ pois representa o total das percentagens.
- Tendo em conta que a 320 youtubers corresponde 40%, então a C youtubers corresponderá $\frac{320\times100}{40}$ youtubers, ou seja, $C = 800$. Pelo mesmo raciocínio obtém-se o valor $B = 24$.
- O raciocínio anterior aplica-se também para o cálculo de D, E, G, H e I. Assim, $D = 25\%$, $E = 13.63\%$, $G = H = 3.75\%$ e $I = 2.87\%$.
- O total de youtubers corresponde a 800, logo $320 + 200 + 109 + A + 30 + 30 + 24 + 23 = 800$, donde $A = 64$.
- Finalmente, sabendo A calcula-se F como anteriormente obtendo-se $F = 8\%$.

O quadro acima fica com o aspeto:

Nº de vídeos em direto	0	1	2	3	4	6	10	15	Total
Nº de youtubers	320	200	109	64	30	30	24	23	800
Percentagem de youtubers	40	25	13.63	8	3.75	3.75	3	2.87	100

`Cotação: por cada valor calculado` 0.01. `Total:` $0.01\times10 = 0.1$.

1.2 Calcule a média, a moda e o desvio-padrão do número de youtubers que fizeram vídeos em direto. Estude a assimetria da distribuição de frequências.

Resolução:

Vamos recorrer à seguinte tabela para ajudar nos cálculos:

x_i	f_i	$x_i f_i$	$x_i^2 f_i$
0	320	0	0
1	200	200	200
2	109	218	436
3	64	192	576
4	30	120	480
6	30	180	1080
10	24	240	2400
15	23	345	5175
Total	800	1495	10347

Nesta tabela x_i representa o número de vídeos em direto e f_i representa a frequência absoluta.

- Cálculo da média:

$$\overline{x} = \frac{1}{N}\sum_{i=1}^{N} x_i f_i = \frac{1}{800} \times 1495 \approx 1.87.$$

- Cálculo da moda: $m_o = 0$, pois representa o valor com maior frequência.
- Cálculo do desvio-padrão:

$$s = \sqrt{\frac{1}{N-1}\left(\sum_{i=1}^{k} f_i x_i^2 - N\overline{x}^2\right)} = \sqrt{\frac{1}{800-1}\left(10347 - 800 \times 1.87^2\right)} \approx 3.07.$$

- Estudo da assimetria: o grau de assimetria de Pearson é o mais adequado visto os cálculos anteriores fornecerem toda a informação necessária para o seu cálculo. Assim,

$$g = \frac{\overline{x} - m_o}{s} = \frac{1.87 - 0}{3.07} \approx 0.61.$$

  Uma vez que $g > 0$ a distribuição de frequências apresenta uma assimetria positiva.

`Cotação: por cada medida calculada` 0.05. `Total:` $0.05 \times 4 = 0.2$.

Questão 2
Mostre que $-1 \leq r \leq 1$, onde r representa o coeficiente de correlação de Pearson.

Exemplo de resolução. A cotação para outras resoluções levará em conta a justificação de todos os passos, igualdades e equivalências apresentadas bem como a obrigatoriedade da definição dos conceitos utilizados, por exemplo, "produto interno".

(Passo 1:) Considere $t \in \mathbb{R}$ e observe que $\displaystyle\sum_{i=1}^{N}\left((y_i-\overline{y})-t(x_i-\overline{x})\right)^2 \geq 0$. Esta desigualdade é equivalente a

$$t^2\sum_{i=1}^{N}(x_i-\overline{x})^2-2t\sum_{i=1}^{N}(x_i-\overline{x})(y_i-\overline{y})+\sum_{i=1}^{N}(y_i-\overline{y})^2 \geq 0.$$

(Passo 2:) Tendo em conta que esta inequação é do 2º grau em t, vem que o seu discriminante[1] é não positivo, ou seja,

$$\left(2\sum_{i=1}^{N}(x_i-\overline{x})(y_i-\overline{y})\right)^2-4\sum_{i=1}^{N}(x_i-\overline{x})^2\sum_{i=1}^{N}(y_i-\overline{y})^2 \leq 0.$$

(Passo 3:) Dividindo a última inequação por $4\displaystyle\sum_{i=1}^{N}(x_i-\overline{x})^2\sum_{i=1}^{N}(y_i-\overline{y})^2$ resulta em

$$\frac{\left(2\sum_{i=1}^{N}(x_i-\overline{x})(y_i-\overline{y})\right)^2-4\sum_{i=1}^{N}(x_i-\overline{x})^2\sum_{i=1}^{N}(y_i-\overline{y})^2}{4\sum_{i=1}^{N}(x_i-\overline{x})^2\sum_{i=1}^{N}(y_i-\overline{y})^2} \leq 0,$$

(Passo 4:) ou seja,

$$\frac{\left(2\sum_{i=1}^{N}(x_i-\overline{x})(y_i-\overline{y})\right)^2}{4\sum_{i=1}^{N}(x_i-\overline{x})^2\sum_{i=1}^{N}(y_i-\overline{y})^2}-1 \leq 0 \iff r^2 \leq 1,$$

(Passo 5:) donde se conclui que $-1 \leq r \leq 1$.

[1] Recorde que na equação $ax^2+bx+c=0$ tem-se que o discriminante é b^2-4ac.

`Cotação: Passos 1, 3, 4 e 5: 0.1 cada. Passo 2: 0.3 . Total:` $0.1 \times 4 + 0.3 = 0.7$.

Questão 3
Um estudo realizado na empresa EPE, onde o número de colaboradores do género masculino a trabalhar é o dobro do número de colaboradores do género feminino a trabalhar, revelou que as mulheres estão ausentes do trabalho com uma probabilidade igual a 0.05, enquanto os homens estão ausentes do trabalho com uma probabilidade igual a 0.02. Considere os eventos:

$$M = \text{“colaborador homem”}, \ F = \text{“colaborador mulher”},$$
$$A = \text{“estar ausente do trabalho”}.$$

Determine os valores de $P(M)$, $P(F)$, $P(A)$, $P(M|A)$ e $P(F|A)$.

Resolução:

- $P(M) = 2P(F)$ e $P(M) + P(F) = 1$. Logo, $P(M) = 2/3$ e $P(F) = 1/3$.
- $P(A) = P(A|M)P(M) + P(A|F)P(F) = 0.02 \times \dfrac{2}{3} + 0.05 \times \dfrac{1}{3} = 0.03$, pois $P(A|M) = 0.02$ e $P(A|F) = 0.05$ pelo enunciado.
- $P(M|A) = \dfrac{P(A \cap M)}{P(A)} = \dfrac{P(M) \times P(A|M)}{P(A)} = \dfrac{\frac{2}{3} \times 0.02}{0.03} = 0.45.$
- $P(F|A) = \dfrac{P(A \cap F)}{P(A)} = \dfrac{P(F) \times P(A|F)}{P(A)} = \dfrac{\frac{1}{3} \times 0.05}{0.03} = 0.55.$

`Cotação: cálculo de` $P(M)$ `e` $P(F)$`: 0.05 cada.`
`Cotação: cálculo de` $P(A)$, $P(M|A)$ `e` $P(F|A)$`: 0.3 cada.`
`Cotação total :` $0.05 \times 2 + 0.3 \times 3 = 1.0$.

Questão 4
Numa empresa de chocolates, o departamento de controlo de qualidade é constituído por 4 homens e 4 mulheres. O CEO da empresa quer nomear 3 elementos desse departamento para representarem a empresa no estrangeiro e, para ser imparcial, resolveu selecioná-los aleatoriamente.
Seja X = “número de mulheres nessa amostra aleatória”.

Escreva a função distribuição da variável aleatória X e calcule $P(X > 2)$ e $P(0 \leqslant X < 1.5)$

Resolução:

Sejam $H =$ "colaborador Homem", $M =$ "colaborador Mulher" e
$X =$ Número de Mulheres na amostra aleatória.

- $P(X=0) = P((HHH)) = \frac{4}{8} \times \frac{3}{7} \times \frac{2}{6} = \frac{1}{14}$.
- $P(X=1) = P((MHH) \cup (HMH) \cup (HHM)) = \left(\frac{4}{8} \times \frac{4}{7} \times \frac{3}{6}\right) + \left(\frac{4}{8} \times \frac{4}{7} \times \frac{3}{6}\right) + \left(\frac{4}{8} \times \frac{3}{7} \times \frac{4}{6}\right) = \frac{6}{14}$.
- $P(X=2) = P((MMH) \cup (HMM) \cup (MHM)) = \left(\frac{4}{8} \times \frac{3}{7} \times \frac{4}{6}\right) + \left(\frac{4}{8} \times \frac{4}{7} \times \frac{3}{6}\right) + \left(\frac{4}{8} \times \frac{4}{7} \times \frac{3}{6}\right) = \frac{6}{14}$.
- $P(X=3) = P((MMM)) = \frac{4}{8} \times \frac{3}{7} \times \frac{2}{6} = \frac{1}{14}$.

Assim, a função distribuição da variável aleatória X é dada por

$$F(x) = \begin{cases} 0 & \text{para } x < 0 \\ 1/14 & \text{para } 0 \leq x < 1 \\ 7/14 & \text{para } 1 \leq x < 2 \\ 13/14 & \text{para } 2 \leq x < 3 \\ 1 & \text{for } x \geq 3. \end{cases}$$

$$P(X > 2) = 1 - P(X \leq 2) = 1 - F(2) = 1 - 13/14 = 1/14$$
$$P(0 \leq X < 1.5) = F(1.5^-) - F(0^-) = 7/14 - 0 = 7/14.$$

`Cotação:  Cálculo de` $P(X = i)$ `: 0.1 cada (`$i = 0, 1, 2, 3.$`).`
`Cotação:  Escrita de` $F(x)$`:  0.2`
`Cotação:  Cálculo das duas probabilidades:  0.2 cada`
`Cotação total :` $0.1 \times 4 + 0.2 + 0.2 \times 2 = 1.0.$

Questão 5
Considere a variável aleatória X e a sua função massa de probabilidade dada por

$$f(x) = \frac{1}{8} \cdot \frac{3!}{x!(3-x)!},$$

para $x = 0, 1, 2, 3$. Determine a função geradora de momento de X e aproveite este resultado para determinar os momentos de ordem 1 e 2 de X.

Resolução:

$$M_X(t) = \mathbb{E}\left[e^{tX}\right] = \frac{1}{8}\sum_{x=0}^{3} e^{tx}\frac{3!}{x!(3-x)!} = \frac{1}{8}\left(1+3e^t+3e^{2t}+e^{3t}\right) = \frac{1}{8}\left(1+e^t\right)^3.$$

$$\mathbb{E}\left[X\right] = M'_X(t)\Big|_{t=0} = \frac{3e^t}{8}\left(1+e^t\right)^2\Big|_{t=0} = \frac{3}{2}.$$

$$\mathbb{E}\left[X^2\right] = M''_X(t)\Big|_{t=0} = \frac{3e^{2t}}{4}\left(1+e^t\right) + \frac{3e^t}{8}\left(1+e^t\right)^2\Big|_{t=0} = 3.$$

```
Cotação:  Cálculo da fgm:  0.6
Cotação:  Cálculo do momento de ordem 1:  0.2
Cotação:  Cálculo do momento de ordem 2:  0.2
Cotação total :  0.6+0.2+0.2=1.0.
```

FIM