Estatística e Probabilidade - Avaliação Presencial 01 - 2022/02 - Gabarito

Prof. Hugo Carvalho

08/11/2022

Questão 1:

a) *(0,5)* Note que $f(x) \geq 0$, para todo x, se $C \geq 0$. Para encontrar o valor exato de tal constante, usemos o fato de que uma função densidade de probabilidade deve integrar 1, sobre todo o seu domínio:

$$\begin{aligned} 1 &= \int_{-\infty}^{+\infty} C \sin(x)\ dx \\ &= C \int_0^{\pi} \sin(x)\ dx \\ &= -C\cos(x)|_{x=0}^{x=\pi} \\ &= 2C \Rightarrow C = \frac{1}{2}. \end{aligned}$$

b) *(0,5)* Para encontrar a média de X, podemos argumentar que, dentro do domínio $[0, 2\pi]$, a função $\sin(x)$ é simétrica em torno de $x = \pi/2$, ou seja, $\sin(\pi/2 + x) = \sin(\pi/2 - x)$, para $0 \leq x \leq \pi/2$, de modo que tenhamos $\mathbb{E}[X] = \pi/2$. Uma outra forma de concluir este fato é simplesmente realizando a cont, através de uma integral por partes:

$$\begin{aligned} \mathbb{E}[X] &= \int_{-\infty}^{+\infty} xf(x)\ dx \\ &= \int_0^{\pi} x\frac{1}{2}\sin(x)\ dx \\ &= \frac{1}{2}\left[-x\cos(x)|_{x=0}^{x=\pi} + \int_0^{\pi}\cos(x)\ dx\right] \\ &= \frac{1}{2}\left[\pi + \sin(x)|_{x=0}^{x=\pi}\right] \\ &= \frac{\pi}{2}. \end{aligned}$$

c) *(1,0)* Para calcular a variância de X, usamos a relação $\mathbb{V}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$, pois o item b) já nos dá que o segundo termo é igual a $\pi^2/4$. Calculemos, portanto, o primeiro termo, através de uma integral por partes:

$$\begin{aligned} \mathbb{E}[X^2] &= \int_{-\infty}^{+\infty} x^2 f(x)\ dx \\ &= \int_0^{\pi} x^2\frac{1}{2}\sin(x)\ dx \\ &= \frac{1}{2}\left[-x^2\cos(x)|_{x=0}^{x=\pi} + \int_0^{\pi} 2x\cos(x)\ dx\right] \\ &= \frac{1}{2}\left[\pi^2 + 2\left[x\sin(x)|_{x=0}^{x=\pi} - \int_0^{\pi}\sin(x)\ dx\right]\right] \\ &= \frac{1}{2}\left[\pi^2 + 2\cos(x)|_{x=0}^{x=\pi}\right] \\ &= \frac{1}{2}\left[\pi^2 - 4\right]. \end{aligned}$$

Dessa forma, temos que

$$\begin{aligned} \mathbb{V}(x) &= \mathbb{E}[X^2] - \mathbb{E}[X]^2 \\ &= \left[\frac{\pi^2}{2} - 2\right] - \frac{\pi^2}{4} \\ &= \frac{\pi^2}{4} - 2. \end{aligned}$$

d) *(1,0)* Como a função densidade de probabilidade de X toma seus valores positivos entre 0 e π, podemos concluir que $F(x) = 0$ se $x \leq 0$ (pois não há probabilidade a ser acumulada antes de 0), e $F(x) = 1$ se $x \geq \pi$ (pois toda probabilidade já foi acumulada entre 0 e π. Calculemos a expressão de tal função para $0 \leq x \leq \pi$:

$$\begin{aligned} F(x) &= \mathbb{P}(X \leq x) \\ &= \int_{-\infty}^{x} f(t)\ dt \\ &= \int_{0}^{x} \frac{1}{2}\sin(t)\ dt \\ &= -\frac{1}{2}\cos(t)\Big|_{t=0}^{t=x} \\ &= \frac{1}{2}(1-\cos(x)),\ \text{para } 0 \leq x \leq \pi. \end{aligned}$$

Questão 2:

a) *(1,0)* Podemos começar preenchendo o valor de $p_{X,Y}(2,2) = p_X(2)p_Y(2) = 0{,}20 \times 0{,}30 = 0{,}06$. Usando o fato de que $p_{X,Y}(1,1) = 0{,}35$, podemos "chutar" que $p_X(1)$ é 0,5 ou 0,7 e que $p_Y(1)$ é 0,7 ou 0,5, respectivamente. Porém, note que $p_Y(1)$ não pode ser igual a 0,7, pois isso implicaria que $p_Y(3) = 0$, o que é uma contradição com termos $p_{X,Y}(1,3) = 0{,}14$. Dessa forma, concluímos que $p_X(1) = 0{,}7$ e que $p_Y(1) = 0{,}5$. Utilizando o fato de que as marginais devem somar 1, temos também que $p_X(3) = 0{,}1$ e que $p_Y(3) = 0{,}2$. Finalmente, com as marginais completas, usamos a independência para preencher toda a tabela, visto que teremos $p_{X,Y}(x,y) = p_X(x)p_Y(y)$, para todos x e y. A tabela final ficará conforme abaixo (em **negrito** estão os valores preenchidos):

	Y			
X	1	2	3	$p_X(x)$
1	0,35	**0,21**	0,14	**0,7**
2	**0,1**	**0,06**	**0,04**	0,20
3	0,05	**0,03**	**0,02**	**0,1**
$p_Y(y)$	**0,5**	0,30	**0,2**	

b) *(0,5)* A largura média e o comprimento médio são, respectivamente, os valores de $\mathbb{E}[Y]$ e $\mathbb{E}[X]$. Façamos os cálculos:

$$\begin{aligned} \mathbb{E}[Y] &= \sum_y y p_Y(y) \\ &= 1 \times 0{,}5 + 2 \times 0{,}3 + 3 \times 0{,}2 = 1{,}7 \text{ km.} \end{aligned}$$

$$\begin{aligned} \mathbb{E}[X] &= \sum_x x p_X(x) \\ &= 1 \times 0{,}7 + 2 \times 0{,}2 + 3 \times 0{,}1 = 1{,}4 \text{ km.} \end{aligned}$$

c) *(1,0)* Em função de X e Y, podemos expressar a área como $A = XY$ e o perímetro como $B = 2X + 2Y$. Façamos então o cálculo que se pede:

$$\begin{aligned} \mathbb{P}(A=4|B=8) &= \frac{\mathbb{P}(A=4; B=8)}{\mathbb{P}(B=8)} \\ &= \frac{\mathbb{P}(X=2;Y=2)}{\mathbb{P}(X=2;Y=2)+\mathbb{P}(X=1;Y=3)+\mathbb{P}(X=3;Y=1)} \\ &= \frac{0{,}06}{0{,}06+0{,}14+0{,}05} \\ &= 0{,}24 \end{aligned}$$

Questão 3:

a) *(0,5)* Conforme apresentado em sala de aula e por dados do problema, denotando por X a quantidade de celulares com defeito a cada lote, temos que $X \sim \text{Poi}(1)$. Dessa forma, a probabilidade desejada é dada por

$$\mathbb{P}(X = 0) = e^{-1}\frac{1^0}{0!} = e^{-1}.$$

b) *(1,0)* Tendo um modelo probabilístico para X, a probabilidade desejada é dada por

$$\begin{aligned}\mathbb{P}(X \geq 5) &= 1 - \mathbb{P}(X \leq 4) \\ &= 1 - \sum_{x=0}^{4} \frac{e^{-1}}{x!}.\end{aligned}$$

c) *(1,0)* Denotando por p a probabilidade de que um dado lote tenha ao menos cinco celulares defeituosos, denote por Y a quantidade de lotes com pelo menos cinco celulares defeituosos, ao longo de um ano de produção. Temos então que $Y \sim \text{Bin}(300, p)$. A probabilidade desejada é dada então por:

$$\begin{aligned}\mathbb{P}(Y \geq 1) &= 1 - \mathbb{P}(Y = 0) \\ &= 1 - \binom{300}{0} p^0 (1-p)^{300} \\ &= 1 - (1-p)^{300} \\ &= 1 - \left[\sum_{x=0}^{4} \frac{e^{-1}}{x!}\right]^{300}.\end{aligned}$$

d) *(1,0)* O valor que está entre colchetes ao final do item c) parece ser "pequeno", pois as parcelas da soma são compostas por e^{-1}, um número menor que 1, dividido por quantias que vão até $4! = 24$. Tomando a potência 300 de tal número, espera-se obter um valor ainda menor, de forma que espera-se que a probabilidade $\mathbb{P}(Y \geq 1)$ seja um valor "grande", ou seja, perto de 1.

Questão 4: *(1,0)* Tal igualdade é chamada a propriedade de perda de memória da distribuição exponencial. Segue uma possível interpretação para essa fórmula: assumindo que X representa o tempo de vida útil de determinado produto, a probabilidade de que o produto dure pelo menos mais s unidades de tempo sabendo que ele já durou por ao menos t unidades de tempo é a mesma probabilidade de um produto novo durar ao menos s unidades de tempo. Ou seja, é como se, a cada instante, o produto "esquecesse" o tempo que já tem de uso, e sempre se comportasse como um produto novo. Pensando, por exemplo, em um produto mecânico que sofre desgaste de uso ao longo do tempo, essa modelagem pode se mostrar muito inacurada, à longo prazo. Mas caso o produto em questão demore a sentir os efeitos de um desgaste de uso, talvez esse modelo seja razoável, em determinada janela curta de tempo.

Questão 5: Questão bônus!

a) *(1,0)* A expressão para $\mathbb{P}(X = x)$ é dada por $p(1-p)^x$. Pensando que p é a probabilidade de obter sucesso em determinado experimento, tal expressão nos dá exatamente a probabilidade de ocorrência de um sucesso e x fracassos. Como não há um "termo de correção", como na distribuição Binomial, para contabilizar as diferentes ordenações possíveis, qualquer ordenação é abarcada em tal probabilidade, inclusive a que desejamos: primeiro observamos x fracassos e em seguida um sucesso.

b) *(1,0)* Fazendo o que se pede, temos que:

$$
\begin{aligned}
0 &= \frac{d}{dp}1 \\
&= \frac{d}{dp}\sum_{x=0}^{\infty} p(1-p)^x \\
&= \sum_{x=0}^{\infty} \frac{d}{dp} p(1-p)^x \\
&= \sum_{x=0}^{\infty} (1-p)^x - px(1-p)^{x-1} \\
&= \sum_{x=0}^{\infty} (1-p)^x - \sum_{x=0}^{\infty} px(1-p)^{x-1} \\
&= \frac{1}{p} - \sum_{x=1}^{\infty} px(1-p)^{x-1} \\
&= \frac{1}{p} - \sum_{y=0}^{\infty} p(y+1)(1-p)^y \\
&= \frac{1}{p} - \sum_{y=0}^{\infty} py(1-p)^y - \sum_{y=0}^{\infty} p(1-p)^y \\
&= \frac{1}{p} - \mathbb{E}[X] - 1.
\end{aligned}
$$

Rearranjando os termos, temos que:

$$
\begin{aligned}
\mathbb{E}[X] &= \frac{1}{p} - 1 \\
&= \frac{1-p}{p}.
\end{aligned}
$$

c) *(1,0)* Um jogo de Mega-Sena consiste na escolha de seis dezenas dentre um total de sessenta. No dia do sorteio, seis dezenas são sorteadas, e o jogador que acertou as seis, leva o prêmio. Dessa forma, a probabilidade de se vencer um jogo de Mega-Sena é de $p = 1/\binom{60}{6} \approx 50 \times 10^6$. Dessa forma, o tempo médio até ganhar pela primeira vez na Mega-Sena é dado por $(1-p)/p \approx 50 \times 10^6$. Ou seja, um jogador deverá fazer em média tal quantidade de jogos para ganhar pela primeira vez. Assumindo que um ano tem 52 semanas, com dois jogos por semana temos 104 jogos por ano, de modo que o tempo médio até ganhar pela primeira vez é de aproximadamente 481.383 anos. Sabendo que cada jogo custa R\$ 4,50, o custo médio desembolsado nessa empreitada é de R\$ 4,50 $\times (1-p)/p =$ R\$ 225.287.365,50. Para fins de comparação, o prêmio médio pago na Mega-Sena é de R\$ 50.000.000,00, e a Mega-Sena da virada, sorteada no último dia do ano, têm pago acima de R\$ 300.000.000,00 nos últimos anos.