Probabilidade e Estatística – Prova da 2^a área – Semestre 2020/1

Importante:

- resolva a prova em folhas em branco ou pautadas que não tenham sido utilizadas anteriormente
- resolva a prova de preferência à caneta ou lápis/lapiseira com ponta grossa, para facilitar a visualização da mesma
- anote o nome e o número de matrícula em todas as folhas
- tire fotos das folha, de forma que seu conteúdo seja legível. Atente para boa iluminação e contraste
- junte as fotos em um arquivo em formato PDF, utilizando algum *software* específico para tal fim (coloquei algumas sugestões no Moodle da disciplina)
- envie a prova via Moodle até o horário agendado!
- **em todas as questões, explicite seu raciocínio, identifique as equações e as suposições utilizadas. Este procedimento faz parte da avaliação!**

---

1. (2 pontos) A função probabilidade conjunta de duas variáveis X e Y é dada por $f(x, y) = c(2x + y)$, onde x e y assumem todos os valores **inteiros** de forma que $(0 \le x \le 2)$ e $(0 \le y \le 3)$ e $f(x, y) = 0$ em caso contrário.

a) Determine a constante c

b) Verifique se as variáveis X e Y são independentes;

c) Calcule $P(1 \le Y \le 2 \,|\, X = 1)$

Resposta

Temos que as variáveis são discretas (pois os valores são inteiros, conforme enunciado), portanto:

a) Temos as seguintes contas:

$$\begin{aligned}
\sum_{x,y} f(x,y) &= \sum_{x,y} c(2x+y) \\
&= c\sum_{x}\left[2x\sum_{y} 1 + \sum_{y} y\right] \\
&= c\sum_{x}\left[2x(4)+6\right] \\
&= c\left[8\sum_{x} x + 6\sum_{x} 1\right] \\
&= c(8(3)+6(3)) \\
&= c(24+18) \\
&= 42c
\end{aligned}$$

Por normalização temos que $\sum_{x,y} f(x,y) = 1$, logo $c = 1/42$.

b) Vamos calcular as distribuições marginais:

$$\begin{aligned} f_1(x) &= \sum_y f(x,y) \\ &= \sum_y c(2x+y) \\ &= c\Big[2x\sum_y 1 + \sum_y y\Big] \\ &= c[2x(4)+6] \\ &= \frac{1}{42}(8x+6) \end{aligned}$$

Além disso:

$$\begin{aligned} f_2(y) &= \sum_x f(x,y) \\ &= \sum_x c(2x+y) \\ &= c\Big[2\sum_x x + y\sum_x 1\Big] \\ &= c[2(3)+y(3)] \\ &= \frac{1}{42}(6+3y) \end{aligned}$$

Das equações acima fica evidente que $f(x,y) \neq f_1(x) \times f_2(y)$ e, portanto, as variáveis são dependentes.

c) Podemos usar aqui:

$$\begin{aligned} f(y|x) &= \frac{f(x,y)}{f_1(x)} \\ &= \frac{c(2x+y)}{c(8x+6)} \\ &= \frac{(2x+y)}{(8x+6)} \end{aligned}$$

Assim:

$$\begin{aligned} f(Y=1|X=1) &= \frac{2(1)+1}{[8(1)+6]} \\ &= \frac{3}{14} \end{aligned}$$

e

$$\begin{aligned} f(Y=2|X=1) &= \frac{2(1)+2}{[8(1)+6]} \\ &= \frac{4}{14} \end{aligned}$$

Logo $P(1 \leq Y \leq 2 \,|\, X=1) = P(Y=1|\, X=1) + P(Y=2\,|\, X=1) = 7/14 = 0,5$.

2. (2 pontos) Um vendedor de sorvete ganha R$ 50/dia, em média, quando é dia de sol. Caso chova, ele ganha R$ 5/dia, mesmo sem trabalhar. Sabe-se também que, indiferentemente do fato de ter sol ou chuva, ele pode ganhar R$ 30/dia nos dias em que trabalhar como pintor.

a) Se às 19h00 o metereologista diz que temos 60% de probabilidade de chuva para o dia seguinte, deverá ele decidir por vender sorvete ou optar por pintura, visando a maior probabilidade possível de lucro no dia seguinte? (esta decisão é tomada ainda na noite anterior, ainda antes de se saber se choverá ou não de fato).

b) Qual deveria ser a probabilidade mínima de chover para que ele decida, no dia anterior, não vender sorvete no dia posterior?

Resposta

a) Neste caso, o valor recebido Y é uma função de X, que é a variável referente à chuva (por simplificação X_1 é quando não chove e X_2 é quando chove). O valor médio em caso de ser vender sorvete μ_S:

$$\begin{aligned}\mu_S &= \mathrm{E}\left[Y\right]\\ &= \sum_x (Y \times X)\\ &= Y_1 \times \Pr\left(X_1\right) + Y_2 \times \Pr\left(X_2\right)\\ &= 50 \times 0,4 + 5 \times 0,6\\ &= 20 + 3 = 23\end{aligned}$$

No caso do sorveteiro não vender sorvete em determinado dia, e decidir pintar, ele recebe os R$30,00 da pintura e ainda assim ele vai receber os R$ 5,00 da sorveteria pois o enunciado fala que ele recebe este valor mesmo sem trabalhar vendendo sorvete. Assim o valor médio recebido no caso de pintar μ_P é:

$$\begin{aligned}\mu_P &= \mathrm{E}\left[Y\right]\\ &= \sum_x (Y \times X)\\ &= Y_1 \times \Pr\left(X_1\right) + Y_2 \times \Pr\left(X_2\right)\\ &= 35 \times 0,4 + 35 \times 0,6\\ &= 35\end{aligned}$$

Assim, compensa mais pintar do que vender sorvete, no caso específico do dia de amanhã.

b) Neste caso as duas médias deveriam ser iguais, ou seja $\mu_S = \mu_P$. Se p é a probabilidade de **chover**, então:

$$\begin{aligned}\mu_P &= \mu_S\\ 35 &= 50(1-p) + 5 \times p\\ &= 50 - 50p + 5p \quad (\div 5)\\ 7 &= 10 - 10p + p\end{aligned}$$

que resulta em $p = 3/9 = 1/3\%$. Ou seja, para qualquer probabilidade igual ou maior de chuva do que $1/3$ ($\sim 33,33...\%$), pintar se torna a melhor opção para este sorveteiro, do ponto de vista de probabilidade de ganhos no dia seguinte.

3. (2 pontos) Um produto tem custo médio de R$ 10,00 e desvio-padrão de R$ 0,80. Calcular o preço de venda médio, bem como seu desvio-padrão, de forma que o lucro médio seja de R$ 4,00 e seu desvio-padrão de R$ 1 ,00. Lembre-se que o lucro de um produto é o valor de venda, descontado do valor de custo.

Resposta

Temos que $L = V - C$, onde L é o valor do lucro líquido, V é o valor da venda e C é o valor do custo. Dos dados do problema, e das propriedades da média, temos $\mathrm{E}[L] = \mathrm{E}[V] - \mathrm{E}[C]$, logo

$$\begin{aligned} \mathrm{E}[V] &= \mathrm{E}[L] + \mathrm{E}[C] \\ &= 4 + 10 \\ &= 14 \end{aligned}$$

Para o desvio-padrão temos $\sigma_L^2 = \sigma_V^2 + \sigma_C^2$, logo

$$\begin{aligned} \sigma_V^2 &= \sigma_L^2 - \sigma_C^2 \\ &= 1^2 - 0,8^2 \\ &= 1 - 0,64 = 0,36 \end{aligned}$$

Logo $\sigma_V = \sqrt{\sigma^2} = 0,6$.

4. (2 pontos) A probabilidade de um sapato apresentar defeito de fabricação, em determinada fábrica, é de 3%. Para que um par de sapatos seja rejeitado pelo controle de qualidade, basta que apenas um dos pés, direito ou esquerdo, apresente defeito. Em um conjunto de de 40.000 pares, qual o valor esperado e o desvio-padrão do número de pares rejeitados? Argumente sobre a motivação para a distribuição de probabilidade escolhida e a motivação do(s) parâmetro(s) escolhido(s) para tal distribuição.

Resposta

Podemos usar a distribuição binomial. Para este caso temos $n = 40000$. Precisamos saber o valor da probabilidade p de termos um par rejeitado. Esta probabilidade é igual à probabilidade de termos o sapato direito com defeito ou o sapato esquerdo com defeito ou os dois sapatos com defeito. Por outro lado, p também é igual à probabilidade de termos 1 menos a probabilidade de nenhum sapato ter defeito, logo:

$$p = 1 - (0,97)^2 = 0,0591$$

Como estamos usando uma distribuição binomial, a média é dada por $\mu = np = 40000 \times 0,0591 = 2364$. A variância é dada por: $\sigma^2 = np(1-p) = 40000 \times 0,0591 \times 0,9409 = 2224,2876$. Logo $\sigma = \sqrt{\sigma^2} = 47,1623536308$.
Por outro lado podemos também utilizar Poisson. Neste caso temos $\lambda = np = 2364$ e $\sigma = \sqrt{\sigma^2} = \sqrt{\lambda} = \sqrt{2364} = 48,6209831246$.

5. (2 pontos) Por experiências do passado em determinada empresa, verificou-se que o tempo médio gasto por um candidato a determinada vaga, em um teste específico, é aproximadamente normal (Gaussiano) com média de 58 minutos e desvio-padrão de 12 minutos.

a) Que porcentagem de candidatos levará menos de 50 minutos para concluir o teste?

b) Que porcentagem não terminará o teste se o tempo máximo concedido é de 85 minutos?

c) Se 50 candidatos fazem o teste, quantos podemos esperar que o terminem no primeiros 40 minutos?

Resposta

Temos:

a) basta usar a distribuição normal padrão:

$$\begin{aligned}
\Pr\left(X<50\right) &= \Pr\left(Z<\frac{50-58}{12}\right)\\
&= \Phi\left(\frac{-8}{12}\right)\\
&= \Phi\left(\frac{-2}{3}\right)\\
&= 0,252492537546923 \quad (\text{ou } 0.25142889509531)
\end{aligned}$$

b) novamente:

$$\begin{aligned}
\Pr\left(X>85\right) &= \Pr\left(Z>\frac{85-58}{12}\right)\\
&= 1-\Phi\left(2,25\right)\\
&= 1-0,987775527344955\\
&= 0.0122244726550447
\end{aligned}$$

c) Neste caso:

$$\begin{aligned}
\Pr\left(X<40\right) &= \Pr\left(Z<\frac{40-58}{12}\right)\\
&= \Phi\left(\frac{-18}{12}\right)\\
&= \Phi\left(\frac{-3}{2}\right)\\
&= 0.0668072012688581
\end{aligned}$$

Isto significa que 6,680...% dos candidatos terminam a prova antes dos 40 minutos. Assim, como temos 50 candidatos, $50 \times 0.0668072012688581 = 3.3403600634429 \sim 3$ candidatos terminam a prova antes de 40 minutos.