Perído: Segundo trimestre 2020

Curso: Estatística para Ciência de Dados

Tutores: Caio, Guilherme, Isis, Marina, Matheus e Tobias Criado em: 25 de Abril de 2020

### Questão 1

Um inspetor de qualidade extrai uma amostra de 10 tubos (produzidos de forma independente um do outro) armazenados em um depósito onde, de acordo com os padrões de produção, espera-se um total de 20% de tubos defeituosos. Quando se encontram quatro ou mais tubos defeituosos, o processo de produção é interrompido para revisão. Qual é a probabilidade que isto aconteça desnecessariamente (ou seja, se a probabilidade de fabricação de um tubo defeituoso está dentro do esperado)?

Alternativas:

- (a) 0.88
- (b) 0,12
- (c) 0,0016
- (d) 0,26
- (e) 0.03

Solução: Alternativa b.

Tem-se que:

$$n = 10$$

$$p = 0, 2$$

$$X_i = \begin{cases} 1, & \text{se o tubo \'e defeituoso,} \\ 0, & \text{se o tubo n\~ao \'e defeituoso.} \end{cases}$$

Assim, pode-se definir a variável aleatória Y (número de tubos defeituosos na amostra de tamanho n=10) da seguinte forma:

$$Y = \sum_{i=1}^{n} X_i \sim Binomial(n = 10, p = 0, 2)$$

Desta forma, o que se quer calcular é:

$$P(Y \ge 4) = \sum_{y=4}^{10} {10 \choose y} 0, 2^{y} (1 - 0, 2)^{10 - y}.$$

Note que pode-se calcular de forma equivalente:

$$\begin{split} P(Y \ge 4) &= 1 - P(Y \le 3) = 1 - [P(Y = 0) + P(Y = 1) + P(Y = 2) + P(Y = 3)] \\ &= 1 - [0, 11 + 0, 26 + 0, 31 + 0, 20] \\ &= 1 - 0, 88 = 0, 12. \end{split}$$

# Questão 2

Uma variável aleatória X tem uma distribuição binomial e uma outra variável aleatória Y tem distribuição de Poisson. As duas têm médias iguais a 3. É possível determinar qual variável aleatória tem maior variância?

Alternativas:

- (a) Não, precisamos saber o número de ensaios de Bernoulli, n, para X.
- (b) Não, precisamos saber o valor da taxa de ocorrência por unidade de tempo,  $\lambda$ , para Y.
- (c) Sim, Y tem uma variância maior.
- (d) Sim, X tem uma variância maior.
- (e) Não, precisamos saber a probabilidade de sucesso, p, para X.

Solução: Alternativa c.

Tem-se que:

$$X \sim Binomial(n,p) \begin{cases} \mathbb{E}(X) &= np \\ \mathbb{V}(X) &= np(1-p) \end{cases} \qquad Y \sim Poisson(\lambda) \begin{cases} \mathbb{E}(Y) &= \lambda \\ \mathbb{V}(Y) &= \lambda \end{cases}$$

Comparando-se as variâncias, tem-se que

$$\mathbb{V}(X) \underset{np = 3}{=} 3 * \underbrace{(1-p)}_{\text{probabilidade}^*} < \mathbb{E}(X) = \mathbb{E}(Y) = \mathbb{V}(Y).$$

\*É um valor entre zero e um.

### Questão 3

Um banco faz operações via Internet e supõe um modelo exponencial com média igual a 4/k para o tempo de conexão (em minutos, sendo k=1 se o cliente for pessoa física e k=2 se o cliente for pessoa jurídica. A porcentagem de pessoas físicas utilizando este sistema é 20%. Qual é a alternativa INCORRETA? Escolha uma:

Alternativas:

- (a) O número de operações realizadas em um intervalo de uma hora possui distribuição de Poisson.
- (b) A probabilidade de uma pessoa jurídica passar mais de dois minutos conectada é 0, 37.
- (c) A probabilidade do tempo de conexão ser maior do que 4 minutos dado que já durou 2 minutos é diferente da probabilidade de conexão ser maior do que 2 minutos.
- (d) A probabilidade de um cliente qualquer passar mais de dois minutos conectado é 0, 42.
- (e) A probabilidade de uma pessoa física passar mais de dois minutos conectada é 0,61.

Solução: Alternativa c.

Pode-se definir, segundo o problema:

$$X \sim Exponencial\left(\lambda = \frac{k}{4}\right) \text{: tempo de conexão}$$
 
$$\mathbb{E}(X) = \frac{4}{k}, \begin{cases} k = 1, & \text{se pessoa física,} \\ k = 2, & \text{se pessoa jurídica.} \end{cases}$$
 
$$Y \sim Bernoulli(p = 0, 2), \begin{cases} Y = 1, & \text{se pessoa física,} \\ Y = 0, & \text{se pessoa jurídica.} \end{cases}$$

Tem-se que:

- Em um Processo de Poisson, o número de eventos tem distribuição de Poisson e o intervalo entre os eventos tem distribuição Exponencial.
- Como a distribuição Exponencial tem falta de memória, P(X > 4|X > 2) = P(X > 2).

• 
$$P(X > 4 \cap Y = 1 | X > 2) + P(X > 4 \cap Y = 0 | X > 2) =$$
  
=  $P(X > 4 | X > 2, Y = 1) P(Y = 1) + P(X > 4 | X > 2, Y = 0) P(Y = 0) =$   
=  $P(X > 2 | Y = 1) P(Y = 1) + P(X > 2 | Y = 0) P(Y = 0) =$   
=  $P(X > 2 | Y = 1) * 0, 2 + P(X > 2 | Y = 0) * 0, 8 =$   
=  $e^{-\frac{1}{4}*2} * 0, 2 + e^{-\frac{1}{2}*2} * 0, 8 =$   
=  $0, 61 * 0, 2 + 0, 37 * 0, 8 =$   
=  $0, 12 + 0, 30 = 0, 42$ .

### Questão 4

O tempo de vida de um certo componente (designado por A) é uma variável aleatória tendo distribuição normal com média 7000h e desvio padrão 900h. Um concorrente desenvolveu um componente mais simples (designado por B), afirmando que seu tempo de vida também segue uma distribuição normal, mas com média de 6500h e desvio padrão de 1400h. Componentes são adquiridos em grandes lotes e quanto maior for a proporção de componentes com tempo de vida superior a 9000h, melhor é o lote. Sob condições descritas, qual afirmação é verdadeira? Escolha uma:

#### Alternativas:

- (a) Não é possível afirmar qual dos dois componentes é melhor porque ambos têm probabilidade de durar mais de 9000h.
- (b) O componente B é melhor porque tem maior probabilidade de durar mais de 9000h do que o componente A.
- (c) O componente A é melhor porque tem menor variabilidade.
- (d) Não é possível afirmar qual dos dois componentes é melhor porque não sabemos o número de componentes dos lotes.
- (e) Nenhuma das anteriores.

### Solução: Alternativa b.

Tem-se que:

 $A \sim Normal(7000, 900)$  $B \sim Normal(6500, 1400).$  Assim, fazendo a padronização, tem-se:

$$P(A > 9000) \approx P\left(Z_A > \frac{9000 - 7000}{900}\right) = P(Z_A > 2, 2) = 0,013$$
  
 $P(B > 9000) \approx P\left(Z_B > \frac{9000 - 6500}{1400}\right) = P(Z_B > 1,79) = 0,037,$ 

isto é, a probabilidade de que o componente B tenha um tempo de vida superior a 9000h é maior que a mesma probabilidade para o componente A.

# Questão 5

Uma fábrica de tubos resolveu instalar uma norma de qualidade para melhorar seu processo de produção e minimizar as perdas por fabricação de tubos defeituosos. Após esse processo ter sido implantado, espera-se que a proporção de tubos fabricados com defeito caia para 10%. Para verificar tal fato, um inspetor extraiu 1000 tubos (produzidos de forma independente um do outro) e observou o número de defeituosos. Qual a probabilidade <u>aproximada</u> de ele encontrar mais de 110 defeituosos, se a probabilidade de fabricar um item defeituoso realmente caiu para 10%?

Alternativas:

- (a) 0, 10
- (b) 0,135
- (c) 0,146
- (d) 0,854
- (e) 0,456

### Solução: Alternativa c.

Pode-se definir a variável aleatória X (número de tubos defeituosos na amostra de tamanho n=1000) da seguinte forma:

$$X \sim Binomial(n = 1000, p = 0, 1).$$

Note que n é grande. Assim, usando o Teorema Central do Limite (TCL),

$$X \approx Normal(np = 100, np(1 - p) = 90).$$

Fazendo a padronização, tem-se:

$$P(X > 110) \approx P\left(Z\frac{110 - 100}{\sqrt{90}}\right) = 0,146.$$