CE-003: Estatística II - Turma: K/O, 1^a Prova (17/10/2016)

GRR: ______________ **Nome:** ____ **Turma:** ________

1. A probabilidade de haver algum acidente considerado grave em um dia, em um trecho de uma rodovia é de 0,04 se não chove e de 0,12 se chove. Sabe-se que, no período considerado, chove em 30% dos dias.

   (a) Se em um determinado dia não houve nenhum acidente, qual a probabilidade que não tenha chovido?

   (b) qual a probabilidade de que, chovendo ou não, haja acidente?

   Solução:
   Eventos e probabilidades informadas:

$$
\begin{aligned}
A &: \text{ocorre acidente} \quad \overline{A} : \text{não ocorre acidente} \\
C &: \text{chove} \quad \overline{C} : \text{não chove} \\
P[A|\overline{C}] &= 0,04 \longrightarrow P[\overline{A}|\overline{C}] = 1 - 0,04 = 0,96 \\
P[A|C] &= 0,12 \longrightarrow P[\overline{A}|C] = 1 - 0,12 = 0,88 \\
P[\overline{C}] &= 0,30 \longrightarrow P[\overline{C}] = 1 - 0,30 = 0,70
\end{aligned}
$$

   Probabilidades pedidas:

   (a) $P[\overline{C}|\overline{A}] = \frac{P[\overline{C}\cap\overline{A}]}{P[\overline{A}]} = \frac{P[\overline{C}\cap\overline{A}]}{P[C\cap\overline{A}]+P[\overline{C}\cap\overline{A}]} = \frac{P[\overline{C}]\cdot P[\overline{A}|\overline{C}]}{P[\overline{C}]\cdot P[\overline{A}|\overline{C}]+P[C]\cdot P[\overline{A}|C]} = \frac{0,70\cdot 0,96}{0.70\cdot 0,96+0,30\cdot 0,88} = 0,718$

   (b) $P[A] = P[A\cap C] + P[A\cap\overline{C}] = P[C]\cdot P[A|C] + P[\overline{C}]\cdot P[A|\overline{C}] = 0,30\cdot 0,12 + 0,70\cdot 0,04 = 0,064$

---

2. Ainda no contexto da questão anterior:

   (a) qual distribuição poderia ser usada para descrever o tempo entre acidentes graves?

   (b) qual a probabilidade de se passarem 10 dias sem acidentes graves?

   (c) qual o tempo médio entre acidentes graves?

   (d) se não houve acidentes por um período de 5 dias consecutivos, qual a probabilidade de haver um acidente nos próximos 10 dias?

   Solução:
   Supondo que os acidentes são eventos independentes pode-se considerar que ocorrem segundo um Processo de Poisson. Desta forma, a probabilidade de ocorrer acidente em um dia $P(A) = 0,064$ corresponde ao número médio de acidentes por dia.
   Pode-se então usar a relação entre a distribuição Poisson (do número de acidentes em um dia) e a exponencial (do intervalo de tempo entre acidentes).

   (a)

$$
\begin{aligned}
X &: \text{Numero de acidentes por dia} \\
X &\sim \mathrm{P}(\lambda = 0,064) \\
T &: \text{Tempo (em dias) entre acidentes} \\
T &\sim \mathrm{Exp}(\lambda = 0,064) \\
f(t) &= \lambda \exp\{-\lambda t\} = \exp\{-0,064t\} \\
F(t) &= 1 - \exp\{-\lambda t\} = 1 - \exp\{-0,064t\}
\end{aligned}
$$

   (b) $P[T > 10] = 1 - F(10) = 0,527$

   (c) $\mathrm{E}[T] = \frac{1}{\lambda} = 1/0,064 = 15,62$ dias

   (d) $P[T < 15|T > 5] = P[T < 10] = 0,473$

---

3. Seja a função de densidade de probabilidade dada por $f(x) = Cx^2 I_{[0,4]}(x)$. Obtenha:

   (a) o valor de C,
   (b) $P[X > 0,5]$,
   (c) $P[X > 0,7|X > 0,5]$,
   (d) $E(X)$,
   (e) o terceiro quartil.

   Solução:

$$f(x) = Cx^2 I_{[0,4]}(x) \longrightarrow F(x) = \int_0^x f(x)\mathrm{d}x = \frac{C}{3}x^3 I_{[0,4]}(x)$$

   (a) $\int_0^4 f(x)\mathrm{dx} = 1 \longrightarrow \mathrm{C} = 3/64$
   (b) $P[X > 0,5] = 1 - F(0,5) = 0,998$
   (c) $P[X > 0,7|X > 0,5] = \frac{1-F(0,7)}{1-F(0,5)} = 0,997$
   (d) $E(X) = \int_0^4 x \cdot f(x)\mathrm{dx} = \ldots = 3$
   (e) $q_3 : \int_0^{q_3} f(x)\mathrm{d}x = 0,75 \longrightarrow q_3 = (64 \cdot 0,75)^{1/3} = 3,63$

---

4. Um grande número de alunos faz uma prova visando avaliar a sua formação e a partir da qual é atribuído a cada um deles um *escore de proficiência* (EP). Os resultados mostram que os EP's podem ser descritos por uma distribuição normal de média 300 e variância 400.

   (a) Qual a percentagem de alunos com escores acima de 320?
   (b) Qual a percentagem de pessoas com escores entre 250 e 350?
   (c) Qual a percentagem de pessoas com escores que não se afastem da média mais do que 50?
   (d) Qual valor deve ter 15% dos escores abaixo dele?
   (e) Deseja-se dividir os escores em três grupos (alto, médio e baixo) com a mesma proporção de alunos. Quais os valores de corte que dividem as categorias/grupos?
   (f) Mantendo-se a mesma média, quanto deveria ser o desvio padrão para que se tenha não mais que 10% dos escores abaixo de 270?
   (g) Mantida a variância de 400 quanto deveria ser o escore médio para que se tenha não mais que 10% dos escores abaixo de 270?

$$X : \text{escores no exame e proficiência}$$
$$X \sim N(400, 45^2)$$

   (a) $P[X > 320] = P[Z > \frac{320-300}{20}] = P[Z > 1] = 0,1587 \longrightarrow 15,87\%$
   (b) $P[250 < X < 350] = P[\frac{250-300}{20} < Z < \frac{350-300}{20}] = P[-2,5 < Z < 2,5] = 0,9876 \longrightarrow 98,76\%$
   (c) $P[|X - 300| < 50] = P[250 < X < 350] = P[\frac{250-300}{20} < Z < \frac{250-300}{20}] = P[2,5 < Z < 2,5] = 0,9876 \longrightarrow 98,76\%$
   (d) $P[X < a] = 0,15 \longrightarrow z = -1,04 = \frac{a-300}{20} \longrightarrow a = 279,3$
   (e)

$$P[X < x_1] = 1/3 \longrightarrow z = -0,431 = \frac{x_1 - 300}{20} \longrightarrow x_1 = 291,4$$
$$P[X < x_2] = 2/3 \longrightarrow z = 0,431 = \frac{x_2 - 300}{20} \longrightarrow x_2 = 308,6$$

   (f) $P[X < 270|\mu = 300, \sigma] = 0,10 \longrightarrow z = -1,28 = \frac{270-300}{\sigma} \longrightarrow \sigma = 23,4$
   (g) $P[X < 270|\mu, \sigma = 20] = 0,10 \longrightarrow z = -1,28 = \frac{270-\mu}{20} \longrightarrow \mu = 295,6$

   Solução computacional com o programa *R*:

```
> (ita <- round(pnorm(320, 300, 20, low=F),dig=4))
```

```
[1] 0,1587

> (itb <- round(diff(pnorm(c(250, 350), 300, 20)),dig=4))

[1] 0,9876

> (itc <- round(diff(pnorm(c(250, 350), 300, 20)),dig=4))

[1] 0,9876

> (itd <- round(qnorm(0.15, 300, 20), dig=1))

[1] 279,3

> (ite <- round(qnorm(c(1/3,2/3), 300, 20), dig=1))

[1] 291,4 308,6

> (itf <- round((270-300)/qnorm(0.10), dig=1))

[1] 23,4

> (itg <- round(270-qnorm(0.10)*20, dig=1))

[1] 295,6
```

---

5. Assume-se que o tempo entre acessos a um blog tem uma distribuição com média de 1,5 segundos. Assumindo alguma distribuição responda os itens a seguir.
   (a) Qual a probabilidade de haver duas conexões com intervalo inferior a 1,5 segundos?
   (b) Qual a probabilidade de se passarem 5 segundos sem conexão alguma?
   (c) Tendo havido uma conexão, qual a probabilidade da próxima conexão ocorrer entre 0,5 e 2,5 segundos?
   (d) Se já se passou 1 segundo sem conexão, qual a probabilidade de se passar mais 0,5 segundos adicionais sem conexão?
   (e) Qual a probabilidade do intervalo entre conexões não superar 3,5 segundos se já se passaram 2 segundos sem conexão?

Solução:
Não se especificou a distribuição e vamos assumir a distribuição exponencial considerando: (i) que devem ser valores positivos, (ii) pela possibilidade de cálculos com as informações fornecidas.

$$X : \text{intervalo de tempo entre conexões (segundos)}$$
$$X \sim \text{Exp}(\lambda = 1/1,5 = 2/3)$$
$$f(x) = \frac{2}{3}\mathrm{e}^{-2x/3} \;\; \mathrm{I}_{(0,\infty)}(x) \qquad F(x) = 1 - \mathrm{e}^{-2x/3}$$

(a) $P[X < 1,5] = \int_0^{1,5} f(x)\mathrm{d}x = F(1,5) = 0,63$
(b) $P[X > 5] = \int_5^{\infty} f(x)\mathrm{d}x = 1 - F(5) = 0,036$
(c) $P[X < 0,5] = \int_{0,5}^{2,5} f(x)\mathrm{d}x = F(2,5) - F(0,5) = 0,53$
(d) $P[X > 1,5|X > 1] = \frac{\int_{1,5}^{\infty} f(x)\mathrm{d}x}{\int_1^{\infty} f(x)\mathrm{d}x} = {}^1P[X > 0,5] = 1 - F(0,5) = 0,72$
(e) $P[X < 3,5|X > 2] = \frac{\int_2^{3,5} f(x)\mathrm{d}x}{\int_2^{3,5} f(x)\mathrm{d}x} = \frac{F(3,5)-F(2)}{1-F(2)} = {}^1P[X < 1,5] = F(1,5) = 0,63$

---

6. (Bônus) Um professor preparou 40 versões diferentes de uma lista de exercícios. As listas são atribuídas ao acaso para estudantes sorteando-se para cada um um número de 1 a 40 que identifica a lista. Se um grupo de três colegas decide fazer as listas juntos, que a probabilidade de que ao menos dois deles recebam a mesma versão?

[1] propriedade de falta de memória da exponencial

Solução:

$$\begin{aligned}
\text{Espaço Amostral: } & S \text{ todas possíveis atribuições de 40 listas para 3 estudantes} \\
n(S) &= 40 \cdot 40 \cdot 40 \\
\text{Evento: } & E \text{ coincidência de lista em ao menos 2 estudantes} \\
& \overline{E} \text{ sem coincidência de listas} \\
n(\overline{E}) &= 40 \cdot 39 \cdot 38 \\
P[E] &= 1 - P[\overline{E}] = 1 - \frac{n(\overline{E})}{n(S)} = 1 - \frac{40 \cdot 39 \cdot 38}{40^3} = 0,0737.
\end{aligned}$$

Nota: este exercício é semelhante ao problema da *coincidência de aniversários* em um grupo de pessoas.
