CE-009: Introdução à Estatística - Turma: A, 1^a Prova (05/05/2017)

GRR: ______________ **Nome:** ____ **Turma:** _______

Para as questões (1) a (5) a seguir considere que há três tipos de sementes de milho, **A**, **B** e **C**, com taxas de germinação de 80, 90 e 95%, respectivamente.

Notação e dados:

$$
\begin{aligned}
&G : \text{a semente germina} \quad \overline{G} : \text{a semente não germina}\\
&A : \text{a semente é do tipo A}\\
&B : \text{a semente é do tipo B}\\
&C : \text{a semente é do tipo C}\\
P[G|A] = 0,80 \quad & P[G|B] = 0,90 \quad P[G|A] = 0,95\\
P[\overline{G}|A] = 0,20 \quad & P[\overline{G}|B] = 0,10 \quad P[\overline{G}|A] = 0,05
\end{aligned}
$$

1. Tomando-se, aleatoriamente, uma semente de cada tipo, qual a probabilidade de que alguma delas germine?

   Solução:
   $P[(G|A) \cup (G|B) \cup (G|C)] = 1 - P[(\overline{G}|A) \cup (\overline{G}|B) \cup (\overline{G}|C)] \overset{ind.}{=} 1 - \{P[(\overline{G}|A)] \cdot P[(\overline{G}|B)] \cdot P[\overline{G}|C)] = 1 - 0,20 \cdot 0,10 \cdot 0,05 = 0,999$

---

2. Um lote é feito com 20% de sementes do tipo **A**, 30% de **B** e 50% de **C**. Toma-se uma semente ao acaso deste lote e verifica-se que ela não germina. Determine de qual tipo a semente tem maior chance de ser. (Dica: Calcule as probabilidades da semente ser de cada um dos tipos).

   Solução:

   $P[A] = 0,20 \quad P[B] = 0,30 \quad P[C] = 0,50$

   Pelo Teorema de Bayes:

$$P[A|\overline{G}] = \frac{P[\overline{G}|A]P[A]}{P[\overline{G}|A]P[A] + P[\overline{G}|B]P[B] + P[\overline{G}|C]P[C]} = \frac{0,20 \cdot 0,20}{0,20 \cdot 0,20 + 0,10 \cdot 0,30 + 0,05 \cdot 0,50} = 0,421$$

$$P[B|\overline{G}] = \frac{P[\overline{G}|B]P[B]}{P[\overline{G}|A]P[A] + P[\overline{G}|B]P[B] + P[\overline{G}|C]P[C]} = \frac{0,10 \cdot 0,30}{0,20 \cdot 0,20 + 0,10 \cdot 0,30 + 0,05 \cdot 0,50} = 0,316$$

$$P[B|\overline{G}] = \frac{P[\overline{G}|C]P[C]}{P[\overline{G}|A]P[A] + P[\overline{G}|B]P[B] + P[\overline{G}|C]P[C]} = \frac{0,05 \cdot 0,50}{0,20 \cdot 0,20 + 0,10 \cdot 0,30 + 0,05 \cdot 0,50} = 0,263$$

---

3. São tomadas 6 sementes do tipo **A** para um teste. Qual a probabilidade de que três ou mais não germinem?

Solução:

$$
\begin{aligned}
X &: \text{número de sementes que } \textbf{não} \text{ germinam dentre seis testadas}\\
x &\in \{0, 1, 2, \ldots 6\}\\
X &\sim \mathrm{B}(n = 6, p = 0,2)\\
P[X \geq 3] &= P[X = 3] + P[X = 4] + P[X = 5] + P[X = 6]\\
&= \binom{6}{3} 0,20^3 0,80^3 + \binom{6}{4} 0,20^4 0,80^2 + \binom{6}{5} 0,20^5 0,80^1 + \binom{6}{6} 0,20^6 0,80^0 = 0,09888
\end{aligned}
$$

Solução computacional:

```
> pbinom(2, size=6, prob=0.2, lower=FALSE)
```

```
[1] 0,09888
```

---

4. Sementes do tipo **B** são tomadas testadas uma a uma até que seja encontrada a primeira que não germine. Qual a probabilidade de que seja necessário tomar quatro ou mais sementes?

Solução:

$$
\begin{aligned}
X &: \text{número de sementes testadas que germinam que seja encontrada a primeira que } \textbf{não} \text{ germinam}\\
x &\in \{0, 1, 2, \ldots\}\\
X &\sim \mathrm{G}(n = 6, p = 0,2)\\
P[X \geq 3] &= 1 - P[X \leq 2] = 1 - \{P[X = 0] + P[X = 1] + P[X = 2]\}\\
&= 1 - \{0,90^0 0,10 + 0,90^1 0,10 + 0,90^2 0,10\} = 1 - 0,10 \cdot \{0,90^0 + 0,90^1 + 0,90^2\} = 0,729
\end{aligned}
$$

OBS: note a definiçãode X como o "numero de falhas até o primeiro sucesso" (e não o número de tentativas...)

Solução computacional:

```
> pgeom(2, prob=0.1, lower=FALSE)
```

```
[1] 0,729
```

---

5. Foram misturadas em uma amostra 30 sementes do tipo **B** e 10 do tipo **C**. Serão tomadas ao acaso, oito sementes da mistura para testes. Qual a probabilidade de que se tenha ao menos duas delas do tipo **C**.

Solução:

$$
\begin{aligned}
X &: \text{número de sementes do tipo } \mathbf{C}\\
x &\in \{0, 1, 2, \ldots 8\}\\
X &\sim \mathrm{HG}(N = 40, K = 10, n = 8)\\
P[X \geq 2] &= 1 - P[X \leq 1] = 1 - \{P[X = 0] + P[X = 1]\} = 1 - \left\{ \frac{\binom{10}{0}\binom{30}{8}}{\binom{40}{8}} + \frac{\binom{10}{1}\binom{30}{7}}{\binom{40}{8}} \right\} = 0,6592
\end{aligned}
$$

Solução computacional:

```
> phyper(1, m=10, n=30, k=8, lower=FALSE)
```

```
[1] 0,6592
```

---

6. Estudos mostram que no plantio do tipo **B** ocorre, em média, 1,4 falhas por 10 metros de plantio.
   (a) Qual a probabilidade de ocorrer no máximo duas falhas em uma determinada linha de plantio de 10 metros?
   (b) Qual a probabilidade de ocorrer cinco ou mais falhas em 20 metros de plantio?

Solução:

(a)

$$\begin{aligned}
X_1 &: \text{número de falhas em 10 metros de plantio}\\
x &\in \{0, 1, 2, \ldots\}\\
X &\sim \mathrm{P}(\lambda = 1,4)\\
P[X \leq 2] = P[X=0] + P[X=1] + P[X=2] &= \frac{\mathrm{e}^{-1,4} 1,4^0}{0!} + \frac{\mathrm{e}^{-1,4} 1,4^1}{1!} + \frac{\mathrm{e}^{-1,4} 1,4^2}{2!}\\
&= \mathrm{e}^{-1,4}\left(\frac{1,4^0}{0!} + \frac{1,4^1}{1!} + \frac{1,4^2}{2!}\right) = \mathrm{e}^{-1,4}\left(1 + 1,4 + \frac{1,4^2}{2}\right) = 0,8335
\end{aligned}$$

(b)

$$\begin{aligned}
X_2 &: \text{número de falhas em 20 metros de plantio}\\
x &\in \{0, 1, 2, \ldots\}\\
X &\sim \mathrm{P}(\lambda = 2,8)\\
P[X \geq 5] &= 1 - \{P[X=0] + P[X=1] + P[X=2] + P[X=3] + P[X=4]\}\\
&= 1 - \left\{\frac{\mathrm{e}^{-2,8} 2,8^0}{0!} + \frac{\mathrm{e}^{-2,8} 2,8^1}{1!} + \frac{\mathrm{e}^{-2,8} 2,8^2}{2!} + \frac{\mathrm{e}^{-2,8} 2,8^3}{3!} + \frac{\mathrm{e}^{-2,8} 2,8^4}{4!}\right\}\\
&= \mathrm{e}^{-2,8}\left(\frac{2,8^0}{0!} + \frac{2,8^1}{1!} + \frac{2,8^2}{2!} + \frac{2,8^3}{3!} + \frac{2,8^4}{4!}\right) = \mathrm{e}^{-2,8}\left(1 + 2,8 + \frac{2,8^2}{2} + \frac{2,8^3}{6} + \frac{2,8^4}{24}\right) = 0,152
\end{aligned}$$

Soluções computacionais:

```
> ppois(2, lambda=1.4)
```

```
[1] 0,8335
```

```
> ppois(4, lambda=2.8, lower=FALSE)
```

```
[1] 0,1523
```

---

7. Seja a função de densidade de probabilidade dada por $f(x) = K(10 - 2x)I_{[0,5]}(x)$.
   (a) Determine o valor de K para que $f(x)$ seja uma função de densidade de probabilidade.
   (b) Calcule $P[X > 4]$
   (c) Calcule $P[X \leq 4]$

(d) Calcule $P[2 < X < 3]$

(e) Calcule $P[X > 4|X > 2]$

(f) Calcule os quartis da distribuição.

(g) A média da distribuição (E[X]) é: (a) $E[X] < 2,5$, (b) $E[X] = 2,5$ ou (c) $E[X] > 2,5$? Justifique.

Solução:

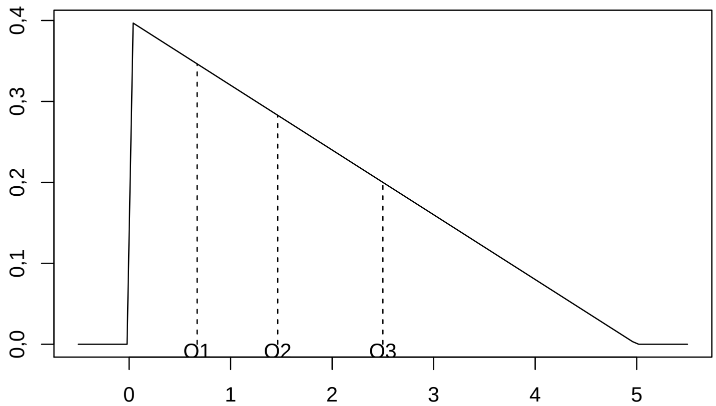

Figura 1: Função de densidade de probabilidade de $f(x) = 0.04(10 - 2x)I_{[0,5]}(x)$ e seus quartis.

Os cálculos para responder os items solicitados podem ser feitos por (i) integração ou (ii) por área de figuras planas sob o gráfico. Vamos aqui fazer por área de figuras. Na solução computacional utilizamos integração.

(a)

$$\int_0^5 f(x)\mathrm{d}x = 1$$
$$\frac{B \cdot h}{2} = \frac{5 \cdot 10K}{2} = 1$$
$$K = 0,04 = 1/25$$

(b) $P[X > 4] = \int_4^5 f(x)\mathrm{d}x = \frac{B \cdot h}{2} = \frac{(5-4) \cdot f(4)}{2} = \frac{(5-4) \cdot 0,04}{2} = 0,04$

(c) $P[X \leq 4] = 1 - P[X > 4] = 1 - 0,04 = 0,96$

(d) $P[2 < X < 3] = \frac{(B+b) \cdot h}{2} = \frac{(f(2)+f(3)) \cdot (3-2)}{2} = \frac{(0,24+0,16) \cdot 1}{2} = 0,20$

(e) $P[X > 4|X > 2] = \frac{P[X>4 \cap X>2]}{P[X>2]} = \frac{P[X>4]}{P[X>2]} = \frac{((5-4)f(4))/2}{((5-2)f(2))/2} = \frac{0,08}{(3 \cdot 0,24)} = 0,111$

(f)

$$q_1 : \int_0^{q_1} f(x)\mathrm{d}x = 0,25 \longrightarrow \int_{q_1}^5 f(x)\mathrm{d}x = 0,75 \longrightarrow \frac{(5-q_1)f(q_1)}{2} = 0,75 \longrightarrow q1 = 0,67$$
$$q_2 : \int_0^{q_2} f(x)\mathrm{d}x = 0,50 \longrightarrow \int_{q_2}^5 f(x)\mathrm{d}x = 0,50 \longrightarrow \frac{(5-q_2)f(q_2)}{2} = 50 \longrightarrow q2 = 1,46$$
$$q_3 : \int_0^{q_3} f(x)\mathrm{d}x = 0,75 \longrightarrow \int_{q_3}^5 f(x)\mathrm{d}x = 0,75 \longrightarrow \frac{(5-q_3)f(q_3)}{2} = 0,75 \longrightarrow q3 = 2,5$$

(g) A média é $E[X] < 2,5$ pois a distribuição é assimétrica com maior massa de probabilidade abaixo deste valor.

Resoluções computacionais:

```
> fxK <- function(x){ifelse(x>0 & x<5, (10-2*x), 0)}
> K <- 1/integrate(fxK, 0, 5)$value
> require(MASS)
> fractions(1/integrate(fxK, 0, 5)$value)

[1] 1/25

> fx = function(x) ifelse(x>0 & x<=5, K*(10-2*x), 0)
> Fx = function(x) {x[x<0] <- 0;x[x>5] <- 5; K*(10*x-x^2)}
> integrate(fx, 0, 5)$value

[1] 1

> (I1 <- integrate(fx, 4, 5)$value)

[1] 0,04

> 1 - Fx(4)

[1] 0,04

> (I2 <- integrate(fx, 0, 4)$value)

[1] 0,96

> Fx(4)

[1] 0,96

> (I3 <- integrate(fx, 2, 3)$value)

[1] 0,2

> Fx(3) - Fx(2)

[1] 0,2

> (I4 <- integrate(fx, 4, 5)$value/integrate(fx, 2, 5)$value)

[1] 0,1111

> (1-Fx(4))/(1-Fx(2))

[1] 0,1111

> qfx <- function(p)
+     optimize(function(x) (Fx(x) - p)^2, interval = c(0, 5))[[1]]
> (q1 <- qfx(0.25))

[1] 0,6699

> integrate(fx, 0, q1)$value

[1] 0,25
```

```
> (q2 <- qfx(0.5))
```

```
[1] 1,464
```

```
> integrate(fx, 0, q2)$value
```

```
[1] 0,5
```

```
> (q3 <- qfx(0.75))
```

```
[1] 2,5
```

```
> integrate(fx, 0, q3)$value
```

```
[1] 0,75
```

---

8. O peso de um certo tipo de semente de soja (expresso como peso de 1000 sementes) possui média de 170 g e desvio padrão de 12 g. Tomando-se um lote de 1000 sementes calcule:

   (a) a probabilidade de que o peso esteja acima 150 g,
   (b) a probabilidade de que o peso esteja entre 170 e 180 g,
   (c) a probabilidade de que o peso esteja entre 165 e 180 g,
   (d) a probabilidade de que o peso esteja acima 190 g,
   (e) o valor cuja probabilidade de estar acima dele seja de 0,15.
   (f) Supondo o mesmo desvio padrão, qual deveria ser o peso médio tal que a probabilidade do lote estar abaixo de 150 g fosse no máximo de 0,5%?

Solução:

$$X : \text{peso de 1000 sementes}$$
$$X \sim N(400, 45^2)$$

(a) $P[X > 150] = P[Z > \frac{150-170}{12}] = P[Z > -1,667] = 0,9522$
(b) $P[170 < X < 180] = P[\frac{170-170}{12} < Z < \frac{180-170}{12}] = P[0 < Z < 0,833] = 0,2977$
(c) $P[165 < X < 180] = P[\frac{165-170}{12} < Z < \frac{180-170}{12}] = P[-0,417 < Z < 0,833] = 0,4592$
(d) $P[X > 190] = P[Z > \frac{190-170}{12}] = P[Z > 1,667] = 0,0478$
(e) $P[X > a] = 0,15 \longrightarrow z = 1,04 = \frac{a-170}{12} \longrightarrow a = 182,4$
(f) $P[X < 150|\mu, \sigma = 12] = 0,005 \longrightarrow z = -2,58 = \frac{150-\mu}{12} \longrightarrow \mu = 180,9$

Solução computacional com o programa R:

```
> (ita <- round(pnorm(150, 170, 12, low=F),dig=4))
```

```
[1] 0,9522
```

```
> (itb <- round(diff(pnorm(c(170, 180), 170, 12)),dig=4))
```

```
[1] 0,2977
```

```
> (itc <- round(diff(pnorm(c(165, 180), 170, 12)),dig=4))

[1] 0,4592

> (itd <- round(pnorm(190, 170, 12, low=F),dig=4))

[1] 0,0478

> (ite <- round(qnorm(0.85, 170, 12), dig=1))

[1] 182,4

> (itf <- round(150-qnorm(0.005)*12, dig=1))

[1] 180,9
```