

Probabilidades e Estatística

LEAN, LEE, LEGI, LEGM, LEIC-A, LEIC-T, LEMat, LERC, LMAC,
MEAer, MEAmbi, MEBiol, MEBiom, MEEC, MEFT, MEMec, MEQ

2º semestre – 2011/2012

25/06/2012 – 9:00

1º TESTE (Época de Recurso)

Duração: 1 hora e 30 minutos

Justifique convenientemente **todas as respostas!**

Grupo I

10 valores

1. Numa dada experiência aleatória, sejam A e B dois acontecimentos independentes, tais que $P(A) = P(B) = 1/2$. Calcule $P[A|(A \cup B)]$. (2.0)

• **Eventos**

$$A, B : \begin{cases} P(A) = P(B) = \frac{1}{2} \\ A \perp\!\!\!\perp B \Leftrightarrow P(A \cap B) = P(A) \times P(B) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4} \end{cases}$$

• **Probabilidade pedida**

$$\begin{aligned} P[A|(A \cup B)] &= \frac{P[A \cap (A \cup B)]}{P(A \cup B)} \\ &= \frac{P(A)}{P(A \cup B)} \\ &= \frac{P(A)}{P(A) + P(B) - P(A \cap B)} \\ &= \frac{\frac{1}{2}}{\frac{1}{2} + \frac{1}{2} - \frac{1}{4}} \\ &= \frac{2}{3}. \end{aligned}$$

2. Um jardineiro efectua uma sementeira de um determinado número de sementes calibradas de uma espécie de plantas. Por experiência, o jardineiro sabe que cada semente não germina com probabilidade 0.2, independentemente do que acontece com as restantes sementes.

- (a) Se o jardineiro usar 20 sementes, qual é a probabilidade de menos de 4 não germinarem? (2.5)

• **V.a.**

X = número de sementes não germinadas, em 20 usadas

• **Distribuição de X**

Supondo que as sementes germinam de forma independente, a v.a. X corresponde ao número de sucessos em n provas de Bernoulli i.i.d., pelo que $X \sim \text{Binomial}(n, p)$

• **Parâmetros**

$$n = 20$$

$$p = P(\text{semente não germinar}) = 0.2$$

• **Prob. pedida**

$$\begin{aligned} P(X < 4) &= P(X \leq 3) \\ &= F_{\text{Bin}(20,0.2)}(3) \\ &\stackrel{\text{tabela}}{=} 0.4114. \end{aligned}$$

- **Alternativa — Prob. pedida**

$$\begin{aligned}
 P(X < 4) &= P(X \leq 3) \\
 &= \sum_{x=0}^3 \binom{20}{x} 0.2^x (1 - 0.2)^{20-x} \\
 &= 0.4114.
 \end{aligned} \tag{0.3}$$

(b) Qual é o menor número de sementes que o jardineiro deve semear para que, com probabilidade superior a 50%, pelo menos 3 sementes não germinem? (1.5)

- **V.a.**

X_n = número de sementes não germinadas em n usadas

- **Distribuição de X_n**

$X_n \sim \text{Binomial}(n, 0.2)$

- **Obtenção do parâmetro n**

$$\begin{aligned}
 n &: P(X_n \geq 3) > 0.5 \\
 &1 - P(X_n \leq 2) > 0.5 \\
 &F_{\text{Bin}(n,0.2)}(2) < 1 - 0.5 = 0.5.
 \end{aligned}$$

Ora, ao pesquisar a tabela da função de distribuição da binomial, em particular a coluna correspondente a $p = 0.2$, conclui-se que: $F_{\text{Bin}(n,0.2)}(2)$ é uma função monótona decrescente de n ; $F_{\text{Binomial}(13,0.2)}(2) = 0.5017$; $F_{\text{Bin}(14,0.2)}(2) = 0.4481$ e

$$F_{\text{Bin}(n,0.2)}(2) < 0.5, \text{ para } n \geq 14.$$

- **Conclusão**

O valor mínimo de n que satisfaz a condição $P(X_n \geq 3) > 0.5$ é $n^* = 14$.

3. A distribuição de probabilidade do montante (em euros) associado a um tipo de transacção efectuado via cartão de crédito por clientes individuais de um banco é função do tipo de vínculo – forte ou fraco – do cliente ao banco. Para clientes com vínculo forte ao banco essa distribuição é exponencial com valor médio de 150 euros, enquanto que para clientes com vínculo fraco ao banco a mesma é exponencial com valor médio de 75 euros. Considere que 50% das transacções do tipo referido são realizadas por clientes com vínculo forte ao banco.

(a) Sabendo que um cliente com vínculo forte ao banco realizou uma transacção do tipo referido de valor superior a 90 euros, qual é a probabilidade do montante dessa transacção exceder 180 euros. (2.0)

- **V.a.**

X = montante da transacção

- **Outra v.a.**

Considerando o evento F = cliente com vínculo forte ao banco, pode definir-se a v.a.:

$X | F$ = montante da transacção, sabendo que foi efectuada por cliente com vínculo forte

- **Distribuição de $X | F$**

$X | F \sim \text{Exponencial}(\lambda_F)$

- **Parâmetro**

$$\begin{aligned}
 \lambda_F &: E(X | F) = 150 \\
 &\frac{1}{\lambda_F} = 150 \\
 &\lambda_F = \frac{1}{150}
 \end{aligned}$$

- **F.d.p. de $X | F$**

$$f_{X|F}(x) = \frac{1}{150} e^{-\frac{x}{150}}, x \geq 0$$

- **Probabilidade pedida**

Atendendo a que a f.d. de $X | F$ é igual a

$$\begin{aligned}
 F_{X|F}(x) &= P(X \leq x | F) \\
 &= \int_{-\infty}^x f_{X|F}(t) dt \\
 &= \int_0^x \frac{1}{150} e^{-\frac{t}{150}} dt \\
 &= \left(-e^{-\frac{x}{150}}\right)\Big|_0^x \\
 &= 1 - e^{-\frac{x}{150}}, x \geq 0,
 \end{aligned}$$

tem-se

$$\begin{aligned}
 P(X > 180 | X > 90, F) &= \frac{P(X > 180, X > 90 | F)}{P(X > 90 | F)} \\
 &= \frac{P(X > 180 | F)}{P(X > 90 | F)} \\
 &= \frac{1 - F_{X|F}(180)}{1 - F_{X|F}(90)} \\
 &= \frac{1 - \left(1 - e^{-\frac{180}{150}}\right)}{1 - \left(1 - e^{-\frac{90}{150}}\right)} \\
 &= e^{-\frac{180-90}{150}} \\
 &= e^{-\frac{9}{15}} \\
 &= e^{-0.6} \\
 &\simeq 0.548812.
 \end{aligned}$$

- **Alternativa 1 — Probabilidade pedida**

$$\begin{aligned}
 P(X > 180 | X > 90, F) &= \frac{P(X > 180, X > 90 | F)}{P(X > 90 | F)} \\
 &= \frac{P(X > 180 | F)}{P(X > 90 | F)} \\
 &= \frac{\int_{180}^{+\infty} \frac{1}{150} e^{-\frac{t}{150}} dt}{\int_{90}^{+\infty} \frac{1}{150} e^{-\frac{t}{150}} dt} \\
 &= \frac{\left(-e^{-\frac{x}{150}}\right)\Big|_{180}^{+\infty}}{\left(-e^{-\frac{x}{150}}\right)\Big|_{90}^{+\infty}} \\
 &= \frac{e^{-\frac{180}{150}}}{e^{-\frac{90}{150}}} \\
 &= e^{-\frac{9}{15}} \\
 &= e^{-0.6} \\
 &\simeq 0.548812.
 \end{aligned}$$

- **Alternativa 2 — Probabilidade pedida**

Invocando a falta de memória da distribuição Exponencial tem-se:

$$\begin{aligned}
 P(X > 180 | X > 90, F) &= P(X > 180 - 90 | F) \\
 &= P(X > 90 | F) \\
 &= \int_{90}^{+\infty} \frac{1}{150} e^{-\frac{t}{150}} dt \\
 &= \left(-e^{-\frac{x}{150}}\right)\Big|_{90}^{+\infty} \\
 &= e^{-\frac{9}{15}} \\
 &= e^{-0.6} \\
 &\simeq 0.548812.
 \end{aligned}$$

- (b) Sabendo que um cliente realizou uma transacção do tipo mencionado de valor superior a 90 euros, qual é a probabilidade de esse cliente ter vínculo forte ao banco? (2.0)

• **Outra v.a.**

Considerando o evento \bar{F} = cliente com vínculo fraco, pode definir-se outra v.a.:

$X | \bar{F}$ = montante da transacção, sabendo que foi efectuada por cliente com vínculo fraco

• **Distribuição de $X | \bar{F}$**

$X | \bar{F} \sim \text{Exponencial}(\lambda_{\bar{F}})$

• **Parâmetro**

$$\lambda_{\bar{F}} : E(X | \bar{F}) = 75$$

$$\frac{1}{\lambda_{\bar{F}}} = 75$$

$$\lambda_{\bar{F}} = \frac{1}{75}$$

• **Probabilidade pedida**

Aplicando o teorema de Bayes, tirando partido das duas distribuições condicionais ao vínculo e do facto de

$$P(X > 90 | \bar{F}) = \dots = e^{-\frac{90}{75}} = e^{-1.2}$$

$$P(F) = P(\bar{F}) = 0.5,$$

obtém-se:

$$\begin{aligned} P(F | X > 90) &= \frac{P(X > 90 | F) \times P(F)}{P(X > 90)} \\ &= \frac{P(X > 90 | F) \times P(F)}{P(X > 90 | F) \times P(F) + P(X > 90 | \bar{F}) \times P(\bar{F})} \\ &= \frac{e^{-0.6} \times 0.5}{e^{-0.6} \times 0.5 + e^{-1.2} \times 0.5} \\ &= \frac{e^{-0.6}}{e^{-0.6} + e^{-1.2}} \\ &\simeq 0.645656. \end{aligned}$$

Grupo II

10 valores

1. Seja X (respectivamente, Y) o tempo em minutos entre chegadas sucessivas de mensagens electrónicas a um servidor A (respectivamente, B). Considere que (X, Y) é um par aleatório contínuo com função de densidade de probabilidade conjunta

$$f_{X,Y}(x, y) = \begin{cases} \frac{1}{2} e^{-\left(\frac{x}{2} + y\right)}, & x, y > 0 \\ 0, & \text{caso contrário.} \end{cases}$$

- (a) Determine o valor esperado do tempo que decorre entre chegadas sucessivas de mensagens electrónicas ao servidor B . (2.5)

• **Par aleatório**

X = tempo em minutos entre chegadas sucessivas de mensagens electr. a um servidor A

Y = tempo em minutos entre chegadas sucessivas de mensagens electr. a um servidor B

• **F.d.p. conjunta de (X, Y)**

$$f_{X,Y}(x, y) = \begin{cases} \frac{1}{2} e^{-\left(\frac{x}{2} + y\right)}, & x, y > 0 \\ 0, & \text{caso contrário.} \end{cases}$$

- **F.d.p. marginal de Y**

$$\begin{aligned}
 f_Y(y) &= \int_{-\infty}^{+\infty} f_{X,Y}(x,y) dx \\
 &= \int_0^{+\infty} \frac{1}{2} e^{-(\frac{x}{2}+y)} dx \\
 &= e^{-y} \times \int_0^{+\infty} \frac{1}{2} e^{-\frac{x}{2}} dx \\
 &= e^{-y} \times \int_0^{+\infty} f_{Exp(1/2)}(x) dx \quad \text{ou} \quad e^{-y} \times (-e^{-\frac{x}{2}})|_0^{+\infty} \\
 &= e^{-y}, y > 0
 \end{aligned}$$

- **Valor esperado de Y**

A consulta do formulário permite-nos concluir que $Y \sim \text{Exponencial}(\lambda = 1)$, logo $E(Y) \stackrel{\text{form.}}{=} \frac{1}{\lambda} = 1$

- **Alternativa 1 — Valor esperado de Y**

$$\begin{aligned}
 E(Y) &= \int_{-\infty}^{+\infty} y f_Y(y) dy \\
 &= \int_0^{+\infty} y e^{-y} dy \\
 &= (y e^{-y})|_0^{+\infty} + \int_0^{+\infty} e^{-y} dy \\
 &= 0 + (-e^{-y})|_0^{+\infty} \\
 &= 1
 \end{aligned}$$

(0.2)

- **Alternativa 2 — F.d.p. marginal de Y e valor esperado de Y**

$$\begin{aligned}
 E(Y) &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} y f_{X,Y}(x,y) dx dy \\
 &= \int_0^{+\infty} \int_0^{+\infty} y \times \frac{1}{2} e^{-(\frac{x}{2}+y)} dx dy \\
 &= \int_0^{+\infty} y e^{-y} \left(\int_0^{+\infty} \frac{1}{2} e^{-\frac{x}{2}} dx \right) dy \\
 &= \int_0^{+\infty} y e^{-y} (-e^{-\frac{x}{2}})|_0^{+\infty} dy \\
 &= \int_0^{+\infty} y e^{-y} dy \\
 &= (-y e^{-y})|_0^{+\infty} + \int_0^{+\infty} e^{-y} dy \\
 &= 0 + (-e^{-y})|_0^{+\infty} \\
 &= 1
 \end{aligned}$$

(b) Calcule $P(X < Y)$.

(2.0)

- **Probabilidade pedida**

$$\begin{aligned}
 P(X < Y) &= P((X,Y) \in \{(x,y) \in \mathbb{R}^2 : x < y\}) \\
 &= \int_{-\infty}^{+\infty} \int_{-\infty}^y f_{X,Y}(x,y) dx dy \\
 &= \int_0^{+\infty} \int_0^y \frac{1}{2} e^{-(\frac{x}{2}+y)} dx dy \\
 &= \int_0^{+\infty} e^{-y} \left(\int_0^y \frac{1}{2} e^{-\frac{x}{2}} dx \right) dy \\
 &= \int_0^{+\infty} e^{-y} (-e^{-\frac{x}{2}})|_0^y dy
 \end{aligned}$$

$$\begin{aligned}
&= \int_0^{+\infty} e^{-y} \left(1 - e^{-\frac{y}{2}}\right) dy \\
&= \left(-e^{-y} + \frac{2}{3} e^{-\frac{3y}{2}}\right) \Big|_0^{+\infty} \\
&= 1 - \frac{2}{3} \\
&= \frac{1}{3}.
\end{aligned}$$

2. O número de horas de sol descoberto, numa cidade do litoral, num dia aleatório do mês de Julho é uma variável aleatória X de valor esperado 9 horas e desvio padrão 2 horas. Admita-se que os números de horas de sol descoberto, nessa cidade, nos 31 dias do mês de Julho são variáveis aleatórias independentes.

(a) Calcule um valor aproximado para a probabilidade de essa cidade ter no total mais de 300 horas de sol descoberto no mês de Julho. (3.0)

• **V.a.**

X_i = no. de horas de sol descoberto no i – ésimio dia de Julho nessa cidade, $i = 1, \dots, 31$

• **Distribuição, valor esperado e variância de X_i**

$X_i \stackrel{i.i.d.}{\sim} X, i = 1, \dots, 31$

$E(X_i) = E(X) = \mu = 9$

$V(X_i) = V(X) = \sigma^2 = 2^2$

• **Nova v.a.**

$S = \sum_{i=1}^{31} X_i$ = no. total de horas de sol descoberto durante o mês de Julho nessa cidade

• **Valor esperado e variância de S**

$E(S) = E\left(\sum_{i=1}^{31} X_i\right) = \sum_{i=1}^{31} E(X_i) \stackrel{X_i \sim X}{=} 31 \times E(X) = 31 \times 9 = 279$

$V(S) = V\left(\sum_{i=1}^{31} X_i\right) \stackrel{X_i \text{ indep.}}{=} \sum_{i=1}^{31} V(X_i) \stackrel{X_i \sim X}{=} 31 \times V(X) = 31 \times 2^2 = 124$

• **Distribuição aproximada de S**

Pelo Teorema do Limite Central (TLC) pode escrever-se

$$\frac{S - E(S)}{\sqrt{V(S)}} = \frac{\sum_{i=1}^n X_i - n\mu}{\sqrt{n\sigma^2}} \stackrel{a}{\sim} \text{Normal}(0, 1).$$

• **Prob. pedida — valor aproximado**

$$\begin{aligned}
P(S > 300) &= 1 - P\left[\frac{S - E(S)}{\sqrt{V(S)}} \leq \frac{300 - E(S)}{\sqrt{V(S)}}\right] \\
&= 1 - P\left[\frac{S - E(S)}{\sqrt{V(S)}} \leq \frac{300 - 279}{\sqrt{124}}\right] \\
&\stackrel{TLC}{\simeq} 1 - \Phi\left(\frac{300 - 279}{\sqrt{124}}\right) \\
&\simeq 1 - \Phi(1.89) \\
&\stackrel{tabela}{=} 1 - 0.9706 \\
&= 0.0294.
\end{aligned}$$

(b) A família A passa férias nessa cidade nos primeiros 14 dias do mês Julho e a família B nos últimos 6 dias do mesmo mês. Admitindo que X tem distribuição normal, calcule a probabilidade de a família A usufruir de pelo menos o dobro do número de horas de sol descoberto usufruídas pela família B durante as respectivas férias nessa cidade. (2.5)

• **V.a.**

X_i = no. de horas de sol descoberto no i – ésimio dia de Julho nessa cidade, $i = 1, \dots, 31$

• **Distribuição, valor esperado e variância de X_i**

$X_i \stackrel{i.i.d.}{\sim} X \sim \text{Normal}(\mu, \sigma^2), i = 1, \dots, 31$

$$E(X_i) = E(X) = \mu = 9$$

$$V(X_i) = V(X) = \sigma^2 = 2^2$$

- **Novas v.a.**

$$S_A = \sum_{i=1}^{14} X_i = \text{no. de horas de sol descoberto durante as férias da família } A$$

$$S_B = \sum_{i=26}^{31} X_i = \text{no. de horas de sol descoberto durante as férias da família } B$$

- **Distribuições de S_A e S_B**

Tratando-se S_A e S_B de somas de v.a. independentes com distribuição normal, tem-se:

$$S_A \sim \text{Normal}(E(S_A), V(S_A));$$

$$S_B \sim \text{Normal}(E(S_B), V(S_B)).$$

- **Parâmetros de S_A e S_B**

$$E(S_A) = E\left(\sum_{i=1}^{14} X_i\right) = 14 \times 9 = 126$$

$$V(S_A) = V\left(\sum_{i=1}^{14} X_i\right) = 14 \times 2^2 = 56$$

$$E(S_B) = E\left(\sum_{i=26}^{31} X_i\right) = 6 \times 9 = 54$$

$$V(S_B) = V\left(\sum_{i=26}^{31} X_i\right) = 6 \times 2^2 = 24$$

- **Probabilidade pedida**

Importa notar que a probabilidade pedida é igual a

$$P(S_A \geq 2 \times S_B) = P(S_A - 2 \times S_B \geq 0) = 1 - P(S_A - 2 \times S_B < 0).$$

É, pois, conveniente lidar com a v.a. $S_A - 2 \times S_B$, que não passa de uma combinação linear de duas v.a. independentes com distribuição normal logo também normalmente distribuída:

$$S_A - 2 \times S_B \sim \text{normal}(E(S_A - 2 \times S_B), V(S_A - 2 \times S_B)), \text{ onde}$$

$$E(S_A - 2 \times S_B) = E(S_A) - 2 \times E(S_B) = 126 - 2 \times 54 = 18$$

$$V(S_A - 2 \times S_B) = V(S_A) + 2^2 \times V(S_B) = 56 + 2^2 \times 24 = 152.$$

Conclui-se que

$$\begin{aligned} P(S_A \geq 2 \times S_B) &= 1 - P(S_A - 2 \times S_B < 0) \\ &= 1 - \Phi\left[\frac{0 - E(S_A - 2 \times S_B)}{\sqrt{V(S_A - 2 \times S_B)}}\right] \\ &= 1 - \Phi\left(\frac{0 - 18}{\sqrt{152}}\right) \\ &\simeq 1 - \Phi(-1.46) \\ &= \Phi(1.46) \\ &\stackrel{\text{tabela}}{=} 0.9279. \end{aligned}$$

Alternativamente, usando a máquina de calcular:

$$\begin{aligned} P(S_A \geq 2 \times S_B) &= 1 - P(S_A - 2 \times S_B < 0) \\ &= 1 - F_{N(18,152)}(0) \\ &\simeq 0.927854 \end{aligned}$$

- **Alternativa — Probabilidade pedida**

Importa notar que a probabilidade pedida é igual a

$$\begin{aligned} P(S_A \geq 2 \times S_B) &= P\left(\sum_{i=1}^{14} X_i - 2 \times \sum_{i=26}^{31} X_i \geq 0\right) \\ &= 1 - P\left(\sum_{i=1}^{14} X_i - 2 \times \sum_{i=26}^{31} X_i < 0\right). \end{aligned}$$

É, pois, conveniente lidar com a v.a. $T = \sum_{i=1}^{14} X_i - 2 \times \sum_{i=26}^{31} X_i$, que não passa de uma combinação linear de v.a. independentes com distribuição normal logo também normalmente distribuída:

$$T \sim \text{normal}(E(T), V(T)), \text{ onde}$$

$$E(T) = \sum_{i=1}^{14} E(X_i) - 2 \times \sum_{i=26}^{31} E(X_i) = 14\mu - 2 \times 6\mu = 18$$

$$V(T) = \sum_{i=1}^{14} V(X_i) + 2^2 \times \sum_{i=26}^{31} V(X_i) = 14\sigma^2 + 2^2 \times 6\sigma^2 = 152.$$

Conclui-se que

$$\begin{aligned}P(S_A \geq 2 \times S_B) &= 1 - P(T < 0) \\&= 1 - \Phi \left[\frac{0 - E(T)}{\sqrt{V(T)}} \right] \\&= 1 - \Phi \left(\frac{0 - 18}{\sqrt{152}} \right) \\&\simeq 1 - \Phi(-1.46) \\&= \Phi(1.46) \\&\stackrel{tabela}{=} 0.9279.\end{aligned}$$

Alternativamente, usando a máquina de calcular:

$$\begin{aligned}P(S_A \geq 2 \times S_B) &= 1 - P(T < 0) \\&= 1 - F_{N(18,152)}(0) \\&\simeq 0.927854\end{aligned}$$