Departamento de Matemática

# Probabilidades e Estatística

## LEAN, LEGM, LEIC-A, LEMat, LMAC, MEAer, MEAmbi, MEBiol, MEBiom, MEC, MEEC, MEFT, MEMec MEQ

1º semestre – 2011/2012 | **1º TESTE** (Época Recurso)
2/02/2012 – **9:00** | Duração: 1 hora e 30 minutos

**Grupo I** — **2.5** + **2.0** + **2.5** + **3.0** = **10.0** valores

**Exercício 1**

(a) • **Quadro de eventos e probabilidades**

| Evento | Probabilidade |
|---|---|
| $L =$ chamada longa | $P(L) = 0.4$ |
| $C =$ chamada curta | $P(C) = 0.6$ |
| $H_0 =$ zero *handoffs* durante a chamada | $P(H_0) = ?$ |
| $H_1 =$ um *handoff* durante a chamada | $P(H_1) = ?$ |
| $H_2 =$ pelo menos dois *handoffs* durante a chamada | $P(H_0) = ?$ |
| $H_0\|L =$ zero *handoffs* dado que a chamada é longa | $P(H_0\|L) = 0.25$ |
| $H_1\|L =$ um *handoff* dado que a chamada é longa | $P(H_1\|L) = 0.25$ |
| $H_2\|L =$ pelo menos dois *handoffs* dado que a chamada é longa | $P(H_2\|L) = 0.5$ |
| $H_0\|C =$ zero *handoffs* dado que a chamada é curta | $P(H_0\|C) = \frac{2}{3}$ |
| $H_1\|C =$ um *handoff* dado que a chamada é curta | $P(H_1\|C) = \frac{1}{6}$ |
| $H_2\|C =$ pelo menos dois *handoffs* dado que a chamada é curta | $P(H_2\|C) = \frac{1}{6}$ |

• **Prob. pedida**

Recorrendo ao teorema da probabilidade total (fazendo uso da partição $\{L, C\}$), tem-se

$$\begin{aligned} P(H_0) &= P(H_0|L)P(L) + P(H_0|C)P(C) \\ &= 0.25 \times 0.4 + \frac{2}{3} \times 0.6 \\ &= 0.5. \end{aligned}$$

(b) • **Prob. pedida**

Fazendo uso do teorema de Bayes, tem-se

$$\begin{aligned} P(L|H_2) &= \frac{P(H_2|L) \times P(L)}{P(H_2)} \\ &= \frac{P(H_2|L) \times P(L)}{P(H_2|L)P(L) + P(H_2|C)P(C)]} \\ &= \frac{0.5 \times 0.4}{0.5 \times 0.4 + \frac{1}{6} \times 0.6} \\ &= \frac{0.2}{0.2 + 01} \\ &= \frac{2}{3}. \end{aligned}$$

**Exercício 2**

(a)
- **V.a.**

  $X_{50}$ = parte inteira do logaritmo do número de bactérias de certo tipo, numa amostra de 50ml de água de um reservatório

- **Distribuição de $X_1$**

  $X_{50} \sim \text{Poisson}(\lambda_{50})$

- **Parâmetro**

$$\begin{aligned} \lambda \quad : \quad & E(X_{50}) = 3.5 \\ & \lambda_{50} = 3.5 \end{aligned}$$

- **Outra v.a.**

  $X_{100}$ = parte inteira do logaritmo do número de bactérias de certo tipo, numa amostra de 100ml de água de um reservatório

- **Distribuição de $X_{100}$**

  Invocando a propriedade reprodutiva da distribuição de Poisson, conclui-se que:

  $X_{100} \sim \text{Poisson}(\lambda_{100} = 2 \times \lambda_{50} = 7)$

- **F.p. de $X_{100}$**

  $P(X_{100} = x) = \frac{e^{-7} \times 7^x}{x!}$, $x = 0, 1, 2, \ldots$

- **Prob. pedida**

$$\begin{aligned} P(X_{100} > 8) &= 1 - P(X_{100} \leq 8) \\ &= 1 - \sum_{x=0}^{8} P(X_{100} = x) \\ &= 1 - F_{Poisson(7)}(8) \\ &\stackrel{tabela}{=} 1 - 0.7291 \\ &= 0.2709. \end{aligned}$$

(b)
- **Nova v.a.**

  $Y$ = número de amostras de 100ml que são recolhidas até ser detectada a primeira amostra nas referidas condições

- **Distribuição de $Y$**

  A v.a. $Y$ corresponde ao número total de provas de Bernoulli i.i.d. realizadas até à ocorrência do primeiro sucesso. Assim,

  $Y \sim \text{Geométrica}(p)$

- **Parâmetro**

  $p = P(X_{100} > 8) \stackrel{(a)}{=} 0.2709$

- **Valor esperado de $Y$**

  $E(Y) \stackrel{form}{=} \frac{1}{p} \simeq 3.691399$

- **Variância de $Y$**

  $V(Y) \stackrel{form}{=} \frac{1-p}{p^2} \simeq 9.935028$

- **Desvio-padrão de $Y$**

  $DP(Y) = +\sqrt{V(Y)} \simeq 3.151988$

| **Grupo II** | **3.0 + 2.0 + 2.5 + 2.5 = 10.0** valores |
|---|---|

**Exercício 1**

(a)
- **V.a.**
  $X =$ diâmetro do tronco de cerejeira (em cm)
- **Distribuição de $X$**
  $X \sim \text{Normal}(\mu, \sigma^2)$
- **Parâmetros**
  $\mu = E(X) = 25cm$
  $\sigma = \sqrt{V(X)} = 10cm$
- **Outra v.a.**
  $Y =$ número de troncos de qualidade inferior, em lote de 1000 troncos
- **Distribuição de $Y$**
  $Y \sim \text{Binomial}(n, p)$
  pois $Y$ representa o número de "sucessos" em 1000 provas de Bernoulli i.i.d.
- **Parâmetros**

$$
\begin{aligned}
n &= 1000 \\
p &= P(\text{tronco de qualidade inferior}) \\
&= P(X < 15) \\
&= P\left[\frac{X - E(X)}{\sqrt{V(X)}} \leq \frac{15 - E(X)}{\sqrt{V(X)}}\right] \\
&= \Phi\left(\frac{15 - 25}{\sqrt{10^2}}\right) \\
&= \Phi(-1) \\
&= 1 - \Phi(1) \\
&\stackrel{tabela}{=} 1 - 0.8413 \\
&= 0.1587.
\end{aligned}
$$

- **F.p. de $Y$**
  $P(Y = y) = \binom{1000}{y} 0.1587^y (1 - 0.1587)^{1000-y}$, $y = 1, 1, \ldots, 1000$
- **Probabilidade pedida — expressão exacta da probabilidade pedida**

$$
\begin{aligned}
P(Y > 500) &= 1 - P(Y \leq 500) \\
&= 1 - \sum_{y=0}^{500} \binom{1000}{y} 0.1587^y (1 - 0.1587)^{1000-y}
\end{aligned}
$$

ou

$$
P(Y > 500) = \sum_{y=501}^{1000} \binom{1000}{y} 0.1587^y (1 - 0.1587)^{1000-y}
$$

- **Probabilidade pedida — valor aproximado** (aproximação da binomial à Normal)
  Dado que
  - $np = 1000 \times 0.1587 = 158.7 > 5$
  - $n(1 - p) = 1000 \times (1 - 0.1587) = 841.3 > 5$

  pode aproximar-se a f.d. da v.a. $Y \sim \text{Binomial}(n = 1000, p = 0.1587)$ pela f.d. da v.a.

$$
Y^* \sim \text{Normal}(E(Y) = np = 158.7, V(Y) = np(1 - p) = 133.514310).
$$

  [Importa notar que esta aproximação (derivada originalmente do teorema de De Moivre-Laplace) coincide com aquela que decorre do Teorema do Limite Central (TLC).]

- **SEM correcção de continuidade**

$$\begin{aligned} P(Y > 500) &= 1 - P(Y \leq 500) \\ &\simeq 1 - P(Y^* \leq 500) \\ &= 1 - P\left[\frac{Y^* - E(Y^*)}{\sqrt{V(Y^*)}} \leq \frac{500 - E(Y^*)}{\sqrt{V(Y^*)}}\right] \\ &= 1 - \Phi\left(\frac{500 - 158.7}{\sqrt{133.514310}}\right) \\ &\simeq 1 - \Phi(29.54) \\ &\simeq 1 - 1.000000 \\ &= 0.000000. \end{aligned}$$

- **Usando máquina de calcular**

  Quem usa máquina não precisa de padronizar e poderá responder

$$\begin{aligned} P(Y > 500) &= 1 - P(Y \leq 500) \\ &\simeq 1 - P(Y^* \leq 500) \\ &= 1 - F_{normal(158.7,133.514310)}(500) \\ &\stackrel{maq.}{=} 1 - 1.00000 \\ &= 0.000000 \end{aligned}$$

- **COM correcção de continuidade**

$$\begin{aligned} P(Y > 500) &= 1 - P(Y \leq 500) \\ &\simeq 1 - P\left(Y^* \leq 500 + \frac{1}{2}\right) \\ &= 1 - P\left[\frac{Y^* - E(Y^*)}{\sqrt{V(Y^*)}} \leq \frac{500 + \frac{1}{2} - E(Y^*)}{\sqrt{V(Y^*)}}\right] \\ &= 1 - \Phi\left(\frac{500 + \frac{1}{2} - 158.7}{\sqrt{133.514310}}\right) \\ &\simeq 1 - \Phi(29.58) \\ &\simeq 1 - 1.000000 \\ &= 0.000000. \end{aligned}$$

• **Probabilidade pedida — valor aproximado** (recurso ao TLC)

- **V.a.**

$$T_i = \begin{cases} 1, & \text{se o tronco } i \text{ é de qualidade inferior} \\ 0, & \text{c.c.} \end{cases}, \; i = 1, \ldots, 1000$$

- **Distribuição de $T_i$**

  $T_i \stackrel{i.i.d.}{\sim} \text{Bernoulli}(p = 0.1587)$, $i = 1, \ldots, 1000$

- **Valor esperado e de variância $T_i$**

  $E(T_i) = p = 0.1587$, $V(T_i) = p(1-p) = 0.133514$

- **Nova v.a.**

  $Y = \sum_{i=1}^{1000} T_i =$ número de troncos de qualidade inferior, em lote de 1000 troncos

- **Valor esperado e variância de $Y$**

  $E(Y) = 1000 \times 0.1587 = 158.7$

  $V(Y) = 1000 \times 0.133514 = 133.514$

- **Distribuição aproximada de $Y$**

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

$$\frac{Y - E(Y)}{\sqrt{V(Y)}} \stackrel{a}{\sim} \text{Normal}(0, 1).$$

– **Valor aproximado da probabilidade pedida**

$$\begin{array}{rcl} P(Y > 500) & = & 1 - P(Y \leq 500) \\ & \simeq & 1 - P(Y* \leq 500) \\ & \overset{TLC}{\simeq} & 1 - \Phi\left(\frac{500 - 158.7}{\sqrt{133.514}}\right) \\ & \simeq & 1 - \Phi(29.54) \\ & =_{tabela} & 1 - 1.000000 \\ & = & 0.000000 \end{array}$$

(b) • **Outra v.a.**

$L$ = lucro por tronco (em euros)

• **Contradomínio de** $L$

$I\!R_L = \{50, 150, 250\}$

• **F.p. de** $L$

$$\begin{array}{rcl} P(L = 50) & = & P(X < 15) \\ & \overset{(a)}{=} & 0.1587 \\ P(L = 150) & = & P(15 \leq X \leq 25) \\ & = & \Phi\left(\frac{25 - 25}{10}\right) - \Phi\left(\frac{15 - 25}{10}\right) \\ & = & \Phi(0) - \Phi(-1) \\ & = & \Phi(0) - [1 - \Phi(1)] \\ & \overset{tabela}{=} & 0.5 - 0.1587 \\ & = & 0.3413 \\ P(L = 250) & = & 1 - P(L = 50) - P(L = 150) \\ & = & 1 - 0.1587 - 0.3413 \\ & = & 0.5 \\ P(L = l) & = & 0,\ l \neq 50, 150, 250 \end{array}$$

• **Valor esperado de** $L$

$$\begin{array}{rcl} E(L) & = & \sum_{l \in I\!R_L} l \times P(L = l) \\ & = & 50 \times 0.1587 + 150 \times 0.3413 + 250 \times 0.5 \\ & = & 184.13. \end{array}$$

**Exercício 2**

(a) • **Par aleatório**

$(X, Y)$

• **Distribuições marginais**

$X \sim \text{Binomial}(n = 1, p = 0.1) \sim \text{Bernoulli}(0.1)$

$Y \sim \text{Hipergeométrica}(N = 5, M = 2, n = 2)$

- **F.p. marginais**

$$
\begin{aligned}
P(X=x) &= 0.1^x(1-0.1)^{1-x} \\
&= \begin{cases} 0.9, & x=0 \\ 0.1, & x=1 \\ 0, & x \neq 0,1 \end{cases} \\
P(Y=y) &= \frac{\binom{2}{y}\binom{5-2}{2-y}}{\binom{5}{2}} \\
&= \frac{\frac{2!}{y!\,(2-y)!}\ \frac{3!}{(2-y)!\,(1+y)!}}{\frac{5!}{2!\,(5-2)!}} \\
&= \begin{cases} \frac{1\times 3}{10} = 0.3, & y=0 \\ \frac{2\times 3}{10} = 0.6, & y=1 \\ \frac{1\times 1}{10} = 0.1, & y=2 \\ 0, & y \neq 0,1,2 \end{cases}
\end{aligned}
$$

- **F.p. conjunta de** $(X,Y)$

Para obtê-la teremos em conta as f.p. marginais de $X$ e de $Y$, que estas f.p. se obtêm à custa da f.p. conjunta de $(X,Y)$ e ainda o facto de $P(X=1|Y=2)=1$. Ora,

$$
\begin{aligned}
P(X=1|Y=2) &= 1 \\
\frac{P(X=1,Y=2)}{P(Y=2)} &= 1 \\
P(X=1,Y=2) &= P(Y=2) \\
&= 0.1.
\end{aligned}
$$

Consequentemente

| $X$ | $Y$ | | | $P(X=x)$ |
|---|---|---|---|---|
| | 0 | 1 | 2 | |
| 0 | $a$ | $b$ | $c$ | 0.9 |
| 1 | $d$ | $e$ | 0.1 | 0.1 |
| $P(Y=y)$ | 0.3 | 0.6 | 0.1 | 1 |

onde

$$
a,b,c,d \in [0,1] \quad : \quad \begin{cases} c+0.1=0.1 \Leftrightarrow c=0 \\ d+e+0.1=0.1 \Leftrightarrow d=e=0 \\ a+d=0.3 \Leftrightarrow a=0.3-d \Leftrightarrow a=0.3 \\ b+e=0.6 \Leftrightarrow b=0.6-e \Leftrightarrow b=0.6. \end{cases}
$$

- **Conclusão**

A f.p. conjunta de $(X,Y)$ é dada por

| $X$ | $Y$ | | | $P(X=x)$ |
|---|---|---|---|---|
| | 0 | 1 | 2 | |
| 0 | 0.3 | 0.6 | 0 | 0.9 |
| 1 | 0 | 0 | 0.1 | 0.1 |
| $P(Y=y)$ | 0.3 | 0.6 | 0.1 | 1 |

(b) - **Correlação entre $X$ e $Y$**

Uma vez que se pretende calcular

$$
\begin{aligned}
corr(X,Y) &= \frac{cov(X,Y)}{\sqrt{V(X)\,V(Y)}} \\
&= \frac{E(XY)-E(X)E(Y)}{\sqrt{V(X)\,V(Y)}},
\end{aligned}
$$

onde

$$\begin{aligned}
E(XY) &= \sum_{x=0}^{1}\sum_{y=0}^{2} xy \times P(X=x, Y=y) \\
&= \sum_{y=0}^{2} y \times P(X=x, Y=y) \\
&= 1 \times 0 + 2 \times 0.1 \\
&= 0.2 \\
E(X) &= E[\text{Bernoulli}(p=0.1)] \\
&= p \\
&= 0.1 \\
E(Y) &= E[\text{Hipergeométrica}(N=5, M=2, n=2)] \\
&= n\frac{M}{N} \\
&= 2 \times \frac{2}{5} \\
&= 0.8 \\
V(X) &= V[\text{Bernoulli}(p=0.1)] \\
&= 0.1 \times (1-0.1) \\
&= 0.09 \\
V(Y) &= E[\text{Hipergeométrica}(N=5, M=2, n=2)] \\
&= n \times \frac{M}{N} \times \frac{N-M}{N} \times \frac{N-n}{N-1} \\
&= 2 \times \frac{2}{5} \times \frac{5-2}{5} \times \frac{5-2}{5-1} \\
&= 0.36
\end{aligned}$$

- **Conclusão**

$$\begin{aligned}
corr(X,Y) &= \frac{E(XY) - E(X)E(Y)}{\sqrt{V(X)V(Y)}} \\
&= \frac{0.2 - 0.1 \times 0.8}{\sqrt{0.09 \times 0.36}} \\
&= \frac{2}{3}
\end{aligned}$$

- **Comentários**

  Visto que $corr(X,Y) = \frac{2}{3} \neq 0$ pode concluir-se que $X$ e $Y$ estão correlacionadas, logo v.a. DEPENDENTES.

  Para além disso, importa notar que $corr(X,Y) > 0$, pelo que as v.a. estão positivamente correlacionadas, pelo que apresentam tendência para variarem no mesmo sentido.

  Estando o valor de $|corr(X,Y)|$ não muito próximo de 1, pode adiantar-se que as v.a. estão moderada e linearmente correlacionadas.