DM DEPARTAMENTO DE MATEMÁTICA TÉCNICO LISBOA

Probabilidades e Estatística

LEIC-A, LEIC-T, LEGM, MA, MEMec

2º semestre – 2015/2016
30/04/2016 – **9:00**

Duração: 90 minutos **1º Teste A**

Justifique convenientemente todas as respostas!

Grupo I	10 valores

1. Uma biblioteca universitária é frequentada por três categorias de utentes: alunos, docentes e funcionários. Da consulta dos registos concluiu-se que: 50% dos utentes são alunos, 30% são docentes e 20% são funcionários. Apurou-se também que: 25% dos utentes são alunos que requisitaram livros; 6% são docentes que requisitaram livros; e 4% são funcionários que requisitaram livros. Tendo sido selecionado ao acaso um utente da biblioteca:

(a) Calcule a probabilidade de o utente selecionado ter requisitado livros da biblioteca. (1.5)

- **Quadro de acontecimentos e probabilidades**

Evento	Probabilidade
$A = \{$utente é aluno$\}$	$P(A) = 0.5$
$D = \{$utente é docente$\}$	$P(D) = 0.3$
$F = \{$utente é funcionário$\}$	$P(F) = 1 - P(A) - P(D) = 0.2$
$R = \{$utente requisitou livros$\}$	$P(R) = ?$
$R \cap A = \{$utente requisitou livros e é aluno$\}$	$P(R \cap A) = 0.25$
$R \cap D = \{$utente requisitou livros e é docente$\}$	$P(R \cap D) = 0.06$
$R \cap F = \{$utente requisitou livros e é funcionário$\}$	$P(R \cap F) = 0.04$

- **Acontecimento**
  $R = \{$utente requisitou livros$\}$

- **Probabilidade pedida**
  Tirando partido do facto de os acontecimentos A, D e F constituírem uma partição do espaço de resultados Ω, podemos escrever:

$$\begin{aligned} P(R) &= P[(R \cap A) \cup (R \cap D) \cup (R \cap F)] \\ &= P(R \cap A) + P(R \cap D) + P(R \cap F) \\ &= 0.25 + 0.06 + 0.04 \\ &= 0.35. \end{aligned}$$

(b) Sabendo que o utente escolhido requisitou livros, determine a probabilidade de ser um aluno. (1.0)

- **Probabilidade pedida**

$$\begin{aligned} P(A \mid R) &= \frac{P(A \cap R)}{P(R)} \qquad \text{(def. de probabilidade condicionada)} \\ &\overset{(a)}{=} \frac{0.25}{0.35} \\ &= \frac{5}{7}. \end{aligned}$$

(c) São $\{$utente requisitou livros$\}$ e $\{$utente é um aluno$\}$ acontecimentos independentes? (1.5)

- **Averiguação de independência**
  Uma vez que os acontecimentos R e A se dizem independentes se e só se $P(R\cap A) = P(R)\times P(A)$ e que
  $$\begin{aligned} P(R\cap A) &= 0.25 \\ &\neq \\ P(R)\times P(A) &= 0.35\times 0.5 \\ &= 0.175, \end{aligned}$$
  pode afirmar-se que R e A não são acontecimentos independentes.
  [Em alternativa, recordemos que, caso A e R sejam acontecimentos independentes (ambos com probabilidades não nulas), $P(A\mid R) = P(A)$. Ora,
  $$\begin{aligned} P(A\mid R) &= \frac{5}{7} \\ &\neq \\ P(A) &= 0.5, \end{aligned}$$
  pelo que pode concluir-se que A e R não são acontecimentos independentes.]

2. Em 1938, o físico Frank Benford propôs a seguinte função de probabilidade para a variável aleatória X que representa o primeiro algarismo de um número inteiro genuíno escrito em base decimal:

x	1	2	3	4	5	6	7	8	9
$P(X=x)$	0.301	0.176	0.125	0.097	0.079	0.067	0.058	0.051	0.046

Constatou-se também que o primeiro algarismo de um número inteiro fraudulento, doravante representado pela variável aleatória Y, possui distribuição uniforme no conjunto $\{1,2,\dots,9\}$.

(a) Obtenha as modas de X e Y. (2.0)

- **Variável aleatória X**
  X = primeiro algarismo de um número inteiro escrito em base decimal
- **Moda de X**
  Represente-se a moda de X por $mo(X)$. Então
  $$mo(X):\quad P[X = mo(X)] = \max_{x\in\{1,2,\dots,9\}} P(X=x).$$
  Atendendo a que $P(X=1) = 0.301$ é superior a qualquer dos restantes valores da f.p. de X, temos $mo(X) = 1$.
- **Variável aleatória Y**
  Y = primeiro algarismo de um número inteiro FRAUDULENTO em base decimal
- **Distribuição de Y**
  $Y \sim$ uniforme discreta$(\{1,2,\dots,9\})$
- **F.p. de Y**
  $$P(Y=y) = \begin{cases} \frac{1}{9}, & y = 1,2,\dots,9 \\ 0, & \text{caso contrário.} \end{cases}$$
- **Moda de Y**
  Neste caso temos
  $$mo(Y):\quad P[Y = mo(Y)] = \max_{y\in\{1,2,\dots,9\}} P(Y=y).$$
  Dado que a f.p. de Y é constante em $\{1,2,\dots,9\}$, a moda não é única. Com efeito, todos os valores possíveis desta v.a. são valores modais, i.e., $mo(Y) = 1,2,\dots,9$.

(b) Determine e compare $P(X>3)$ e $P(Y>3)$. (2.0)

- **Probabilidades pedidas**

$$\begin{aligned} P(X>3) &= 1-P(X\leq 3) \\ &= 1-\sum_{x=1}^{3} P(X=x) \\ &= 1-(0.301+0.176+0.125) \\ &\simeq 1-0.602 \\ &= 0.398 \\ P(Y>3) &= \sum_{y=4}^{9} P(Y=y) \\ &= \sum_{y=4}^{9} \frac{1}{9} \\ &= \frac{9-4+1}{9} \\ &= \frac{2}{3} \\ &\simeq 0.6667 \end{aligned}$$

- **Comentário**

  Em caso de fraude o primeiro algarismo será superior a 3 mais frequentemente que na ausência de fraude. [Poderemos tirar partido deste e de outros factos para emitir alertas de fraude.]

(c) Ao examinar um registo fiscal selecionado ao acaso, uma inspectora suspeita de fraude com probabilidade 0.1. Calcule o valor esperado do número de registos examinados até que a inspectora suspeite de fraude pela primeira vez. Indique as suas hipóteses de trabalho. (2.0)

- **Variável aleatória de interesse**

  R = registos examinados até que a inspectora suspeite de fraude pela primeira vez

- **Hipóteses de trabalho**

  Admitiremos que:
  - a inspectora examina os registos de modo independente;
  - a probabilidade de suspeita de fraude se mantém constante ao longo dos exames aos registos.

- **Distribuição de R**

  $R \sim$ geométrica(p)

  com

  $p = P(\text{suspeita de fraude}) = 0.1$

- **Valor esperado de R**

$$\begin{aligned} E(R) &\stackrel{form.}{=} \frac{1}{p} \\ &= 10 \text{ registos.} \end{aligned}$$

Grupo II	10 valores

1. Admita que a função de distribuição da velocidade de impacto (X, em milhas náuticas por hora) de ondas em cascos de navios em determinada região do globo é dada por:

$$F_X(x) = \begin{cases} 0, & x<0 \\ 1-e^{-x^2}, & x\geq 0. \end{cases}$$

(a) Obtenha a probabilidade de a velocidade de impacto de uma onda pertencer ao intervalo [1, 1.5]. (1.0)

- **Probabilidade pedida**

$$\begin{aligned} P(1 \le X \le 1.5) &= P(X \le 1.5) - P(X \le 1) \\ &= F_X(1.5) - F_X(1) \\ &= \left(1 - e^{-1.5^2}\right) - \left(1 - e^{-1^2}\right) \\ &= e^{-1} - e^{-2.25} \\ &\simeq 0.2625. \end{aligned}$$

(b) Sabendo que $E(X) = \frac{\sqrt{\pi}}{2}$ e $V(X) = 1 - \frac{\pi}{4}$, obtenha um valor aproximado para a probabilidade de a velocidade média de impacto de 50 ondas ser superior a 1. (3.0)

- **V.a.**
  X_i = velocidade de impacto da onda i, $\quad i = 1, \dots, n$
  $n = 50$

- **Distribuição, valor esperado e variância comuns**
  $X_i \stackrel{i.i.d.}{\sim} X, \quad i = 1, \dots, n$
  $E(X_i) = E(X) = \mu = \frac{\sqrt{\pi}}{2}, \quad i = 1, \dots, n$
  $V(X_i) = V(X) = \sigma^2 = 1 - \frac{\pi}{4}, \quad i = 1, \dots, n$

- **Nova v.a.**
  $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ = velocidade média de impacto de n ondas

- **Valor esperado e variância de $\bar{X}$**
  $E(\bar{X}) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) \stackrel{X_i \sim X}{=} \frac{1}{n} \times nE(X) = E(X) = \mu$
  $V(\bar{X}) = V\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) \stackrel{X_i\ indep.}{=} \frac{1}{n^2} \times \sum_{i=1}^{n} V(X_i) \stackrel{X_i \sim X}{=} \frac{1}{n^2} \times nV(X) = \frac{V(X)}{n} = \frac{\sigma^2}{n}$

- **Distribuição aproximada de $\bar{X}$**
  Pelo teorema do limite central (TLC) pode escrever-se

$$\frac{\bar{X} - E(\bar{X})}{\sqrt{V(\bar{X})}} = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \stackrel{a}{\sim} \text{Normal}(0,1).$$

- **Valor aproximado da probabilidade pedida**
  Atendendo a que $\mu = \frac{\sqrt{\pi}}{2} \simeq 0.8862$ e $\frac{\sigma^2}{n} = \frac{1-\frac{\pi}{4}}{50} \simeq 4.292 \times 10^{-3}$, temos

$$\begin{aligned} P(\bar{X} > 1) &= 1 - P(\bar{X} \le 1) \\ &= 1 - P\left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \le \frac{1 - \mu}{\sigma/\sqrt{n}}\right) \\ &\stackrel{TLC}{\simeq} 1 - \Phi\left(\frac{1 - 0.8862}{\sqrt{4.292 \times 10^{-3}}}\right) \\ &\simeq 1 - \Phi(1.74) \\ &\stackrel{tabela/calc.}{\simeq} 1 - 0.9591 \\ &= 0.0409. \end{aligned}$$

2. Sejam X e Y as variáveis aleatórias que representam, respetivamente, o número de dias consecutivos em que há equipamento parado e o número de trabalhadores dispensados numa fábrica de componentes eletrónicas, quando há necessidade de proceder à manutenção ou reparação de alguma máquina. Admita que a função de probabilidade conjunta de (X, Y) é dada por:

	Y		
X	0	1	2
1	0.3	0.1	0
2	0.1	0.2	0.05
3	0.05	0.1	0.1

(a) Determine o valor da função de distribuição conjunta no ponto $(2, 1)$. (1.0)

- **Par aleatório** (X,Y)
  X = número de dias consecutivos em que há equipamento parado
  Y = número de trabalhadores dispensados
- **F.p. conjunta**
  $P(X=x,Y=y)$ (ver enunciado).
- **Probabilidade pedida**

$$\begin{aligned} F_{X,Y}(2,1) &= P(X\leq 2, Y\leq 1) \\ &= \sum_{x=1}^{2}\sum_{y=0}^{1} P(X=x,Y=y) \\ &= 0.3+0.1+0.1+0.2 \\ &= 0.7. \end{aligned}$$

(b) Calcule o valor esperado e a variância da variável aleatória $Y \mid X=3$. (2.5)

- **V.a.**
  $Y \mid X=3$
- **F.p. de** $Y \mid X=3$
  Atendendo a que

$$\begin{aligned} P(X=3) &= \sum_{x=1}^{3} P(X=3,Y=y) \\ &= 0.05+0.1+0.1 \\ &= 0.25, \end{aligned}$$

temos

$$\begin{aligned} P(Y=y \mid X=3) &= \frac{P(X=3,Y=y)}{P(X=3)} \\ &= \begin{cases} \frac{0.05}{0.25}=0.2, & y=0 \\ \frac{0.1}{0.25}=0.4, & y=1,2 \\ 0, & \text{restantes valores de } x \end{cases} \end{aligned}$$

- **Valor esperado de** $Y \mid X=3$

$$\begin{aligned} E(Y \mid X=3) &= \sum_{y=0}^{2} y\times P(Y=y \mid X=3) \\ &= 0\times 0.2+1\times 0.4+2\times 0.4 \\ &= 1.2 \end{aligned}$$

- **Variância de** $Y \mid X=3$

$$\begin{aligned} V(Y \mid X=3) &= E(Y^2 \mid X=3)-E^2(Y \mid X=3) \\ &= \left[\sum_{y=0}^{2} y^2\times P(Y=y \mid X=3)\right]-1.2^2 \\ &= (0^2\times 0.2+1^2\times 0.4+2^2\times 0.4)-1.2^2 \\ &= 2-1.44 \\ &= 0.56. \end{aligned}$$

(c) Obtenha o valor da covariância entre X e Y. Que conclui? (2.5)

- **Covariância entre** X **e** Y
  Uma vez que se pretende calcular

$$cov(X,Y)=E(XY)-E(X)\times E(Y),$$

serão necessários alguns cálculos auxiliares que envolverão as f.p. conjunta de (X,Y) e marginais de X e Y.

- **F.p. conjunta e f.p. marginais**
  $P(X=x, Y=y)$, $P(X=x)=\sum_{y=0}^{2} P(X=x, Y=y)$ e $P(Y=y)=\sum_{x=1}^{3} P(X=x, Y=y)$ encontram-se sumariadas na tabela seguinte:

	Y			
X	0	1	2	$P(X=x)$
1	0.3	0.1	0	0.4
2	0.1	0.2	0.05	0.35
3	0.05	0.1	0.1	0.25
$P(Y=y)$	0.45	0.4	0.15	1

- **Valor esperado de XY**

$$\begin{aligned} E(XY) &= \sum_{x=1}^{3}\sum_{y=0}^{2} xy \times P(X=x, Y=y) \\ &= 1\times 1\times 0.1 + 2\times 1\times 0.2 + 2\times 2\times 0.05 + 3\times 1\times 0.1 + 3\times 2\times 0.1 \\ &= 1.6 \end{aligned}$$

- **Valor esperado de X**

$$\begin{aligned} E(X) &= \sum_{x=1}^{3} x \times P(X=x) \\ &= 1\times 0.4 + 2\times 0.35 + 3\times 0.25 \\ &= 1.85 \end{aligned}$$

- **Valor esperado de Y**

$$\begin{aligned} E(Y) &= \sum_{y=0}^{2} y \times P(Y=y) \\ &= 0\times 0.45 + 1\times 0.4 + 2\times 0.15 \\ &= 0.7 \end{aligned}$$

- **Covariância entre X e Y (cont.)**

$$\begin{aligned} cov(X,Y) &= E(XY) - E(X)E(Y) \\ &= 1.6 - 1.85\times 0.7 \\ &= 0.305 \end{aligned}$$

- **Conclusão**
  - Visto que $cov(X,Y) \neq 0$ podemos concluir que X e Y são v.a. DEPENDENTES.
  - Dado que $cov(X,Y) > 0$ podemos adiantar que X e Y tenderão a variar no mesmo sentido relativamente aos respectivos valores esperados.