DM DEPARTAMENTO DE MATEMÁTICA TÉCNICO LISBOA	**Probabilidades e Estatística**	$2^{\underline{o}}$ semestre – 2015/2016
	LEAN, LEE, LEGI, LEMat, LETI, LMAC, MEAmb, MEAer, MEBiol, MEBiom, MEEC, MEFT, MEQ	30/04/2016 – **11:00**

Duração: 90 minutos **$1^{\underline{o}}$ Teste B**

Justifique convenientemente todas as respostas!

Grupo I 10 valores

1. Vários estudos têm mostrado que o consumo excessivo de tabaco pode provocar doenças nos pulmões e na bexiga e que estas doenças se manifestam independentemente uma da outra, tanto em indivíduos fumadores com em não fumadores. Num grupo de interesse, verificou-se que 20% dos indivíduos são fumadores, 40% dos indivíduos fumadores são doentes pulmonares, 25% dos indivíduos não fumadores são doentes pulmonares e 76% dos indivíduos fumadores não são doentes da bexiga. Tendo sido escolhido, ao acaso, um indivíduo do referido grupo:

(a) Calcule a probabilidade de o indivíduo selecionado ter doença pulmonar. (2.0)

- **Quadro de acontecimentos e probabilidades**

Evento	**Probabilidade**
$DP = \{$indivíduo tem doença pulmonar$\}$	$P(DP) = ?$
$DB = \{$indivíduo tem doença na bexiga$\}$	$P(DB) = ?$
$F = \{$indivíduo é fumador$\}$	$P(F) = 0.2$
	$P(DP \mid F) = 0.4$
	$P(DP \mid \overline{F}) = 0.25$
	$P(\overline{DB} \mid F) = 0.76$

- **Probabilidade pedida**

$$\begin{aligned} P(DP) &= P(DP \mid F) \times P(F) + P(DP \mid \overline{F}) \times P(\overline{F}) \qquad \text{(lei da probabilidade total)} \\ &= P(DP \mid F) \times P(F) + P(DP \mid \overline{F}) \times [1 - P(F)] \\ &= 0.4 \times 0.2 + 0.25 \times (1 - 0.2) \\ &= 0.28. \end{aligned}$$

(b) Dado que o indivíduo selecionado é fumador, qual é a probabilidade de ter doença pulmonar ou da bexiga? (2.0)

- **Probabilidade pedida:** $P[(DP \cup DB) \mid F]$
Tirando partido do facto de os acontecimentos DP e DB serem condicionalmente independentes dado F, i.e.,

$$P[(DP \cap DB) \mid F] = P(DP \mid F) \times P(DB \mid F),$$

temos

$$\begin{aligned} P[(DP \cup DB) \mid F] &= P(DP \mid F) + P(DB \mid F) - P[(DP \cap DB) \mid F] \\ &= P(DP \mid F) + P(DB \mid F) - P(DP \mid F) \times P(DB \mid F) \\ &= P(DP \mid F) + [1 - P(\overline{DB} \mid F)] - P(DP \mid F) \times [1 - P(\overline{DB} \mid F)] \\ &= 0.4 + (1 - 0.76) - 0.4 \times (1 - 0.76) \\ &= 0.544. \end{aligned}$$

2. Paragens não programadas de uma máquina industrial de corte e gravação a laser em uso contínuo ocorrem segundo um processo de Poisson com taxa mensal igual a 3.

(a) Qual é a probabilidade de terem ocorrido menos de 4 paragens não programadas da máquina num mês, sabendo que nesse mesmo mês ocorreu pelo menos uma dessas paragens? (2.0)

- **Variável aleatória de interesse**
  $X =$ número de paragens não programadas da máquina num mês

- **Distribuição de X**
  Dado que lidamos com um processo de Poisson com taxa igual a 3 paragens por mês, temos $X \sim \text{Poisson}(3)$.

- **F.p. de X**
  $P(X=x) = \frac{e^{-3}3^x}{x!}, \quad x = 0,1,2,\ldots$

- **Prob. pedida**

$$
\begin{aligned}
P(X<4 \mid X \geq 1) &= \frac{P(X<4, X\geq 1)}{P(X\geq 1)} \\
&= \frac{P(1 \leq X < 4)}{1-P(X<1)} \\
&= \frac{P(0 < X \leq 3)}{1-P(X=0)} \\
&= \frac{F_X(3)-F_X(0)}{1-F_X(0)} \\
&\stackrel{tabela/calc.}{=} \frac{0.6472-0.0498}{1-0.0498} \\
&\simeq 0.6287.
\end{aligned}
$$

[Alternativamente,

$$
\begin{aligned}
P(X<4 \mid X \geq 1) &= \ldots \\
&= \frac{\sum_{x=1}^{3} P(X=x)}{1-P(X=0)} \\
&= \frac{3\,e^{-3} + \frac{9\,e^{-3}}{2} + \frac{27\,e^{-3}}{6}}{1-e^{-3}} \\
&= \frac{12\,e^{-3}}{1-e^{-3}} \\
&\simeq 0.6287.]
\end{aligned}
$$

(b) Obtenha a mediana do número de paragens não programadas da máquina num mês. (2.0)

- **Mediana de X**
  Represente-se a mediana de X por $me(X)$. Então

$$
\begin{aligned}
me(X) \quad : \quad & \frac{1}{2} \leq F_X[me(X)] \leq \frac{1}{2} + P[X = me(X)] && (1)\\
& \frac{1}{2} \leq F_X[me(X)] \leq \frac{1}{2} + [F_X[me(X)] - F_X[me(X)^-]] \\
& F_X[me(X)^-] \leq \frac{1}{2} \leq F_X[me(X)]. && (2)
\end{aligned}
$$

Ora, tirando partido da definição de mediana em (1) e de

$$
\begin{aligned}
\frac{1}{2} \leq F_X(3) \stackrel{(a)}{=} 0.6472 \quad &\leq \frac{1}{2} + P(X=3) = \frac{1}{2} + [F_X(3) - F_X(2)] \\
&= \frac{1}{2} + (0.6472 - 0.4232) = 0.7240,
\end{aligned}
$$

concluímos que $me(X) = 3$.

[Em alternativa, notemos que $F_X(3) = P(X \leq 3) \stackrel{(a)}{=} 0.6472 \geq \frac{1}{2}$; mais, de uma nova consulta da tabela da f.d. da Poisson tem-se $F_X(2) = F_X(3^-) = 0.4232 \leq \frac{1}{2}$. Logo o resultado (2), leva-nos a concluir que $me(X) = 3$.]

(c) Obtenha a probabilidade de o intervalo entre duas paragens não programadas consecutivas da máquina não exceder um mês. (2.0)

- **Variável aleatória de interesse**
  T = intervalo (em meses) entre duas paragens não programadas consecutivas da máquina
- **Distribuição de T**
  Uma vez que lidamos com um processo de Poisson com taxa igual a 3 paragens por mês, temos $T \sim \text{Exponencial}(\lambda = 3)$.
- **F.d.p. de T**
  $$f_T(t) = \begin{cases} 0, & t < 0 \\ 3 \times e^{-3t}, & t \geq 0 \end{cases}$$
- **Prob. pedida**
  $$\begin{aligned} P(T \leq 1) &= \int_0^1 3 \times e^{-3t} dt \\ &= -e^{-3t}\Big|_0^3 \\ &= 1 - e^{-3} \\ &\simeq 0.9502. \end{aligned}$$

Resolução alternativa

- **Variável aleatória de interesse**
  T = intervalo (em meses) entre duas paragens não programadas consecutivas da máquina
- **Prob. pedida**
  $$\begin{aligned} P(T \leq 1) &= 1 - P(T > 1) \\ &= 1 - P(X = 0) \\ &= 1 - \frac{e^{-3} \times 3^0}{0!} \\ &= 1 - e^{-3} \\ &\simeq 0.9502. \end{aligned}$$

Grupo II	10 valores

1. O controlo de qualidade de azulejos fabricados artesanalmente numa empresa consiste na inspeção de 10 azulejos, escolhidos ao acaso e sem reposição, de cada lote de 100 azulejos produzidos.

(a) Admitindo que existem 3 azulejos não conformes num desses lotes de azulejos, qual é a probabilidade de serem detetados quanto muito 2 azulejos não conformes na respetiva inspeção? (2.0)

- **Variável aleatória de interesse**
  X = número de azulejos não conformes numa amostra de 10 azulejos seleccionados ao acaso e SEM reposição de um lote com 100, dos quais 3 são não conformes
- **Distribuição de X**
  $X \sim \text{Hipergeométrica}(N, M, n)$
  com:
  $N = 100$ (azulejos no lote);
  $M = 3$ (azulejos não conformes no lote);
  $n = 10$ (azulejos seleccionados ao acaso e SEM reposição).
- **F.p. de X**
  $$\begin{aligned} P(X = x) &\overset{form.}{=} \begin{cases} \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}, & x = \max\{0, n-(N-M)\}, \ldots, \min\{n, M\} \\ 0, & \text{caso contrário} \end{cases} \\ &= \begin{cases} \frac{\binom{3}{x}\binom{100-3}{10-x}}{\binom{100}{10}}, & x = 0, 1, 2, 3 \\ 0, & \text{caso contrário} \end{cases} \end{aligned}$$

- **Prob. pedida**

$$
\begin{aligned}
P(X \leq 2) &= 1 - P(X > 2) \\
&= 1 - P(X = 3) \\
&= 1 - \frac{\binom{3}{3}\binom{100-3}{10-3}}{\binom{100}{10}} \\
&= 1 - \frac{\binom{97}{7}}{\binom{100}{10}} \\
&= 1 - \frac{\frac{97!}{7!90!}}{\frac{100!}{10!90!}} \\
&= 1 - \frac{\frac{97!}{7!}}{\frac{100\times 99\times 98\times 97!}{10\times 9\times 8\times 7!}} \\
&= 1 - \frac{10 \times 9 \times 8}{100 \times 99 \times 98} \\
&= \frac{2693}{2695} \\
&\simeq 0.999256.
\end{aligned}
$$

(b) Supondo que a percentagem de azulejos não conformes é de 3%, obtenha um valor aproximado da probabilidade de em 50 lotes o controlo de qualidade detetar mais de 25 azulejos não conformes. (3.0)

- **Variável aleatória de interesse**
  X_i = número de azulejos não conformes na amostra i, $\quad i = 1, \ldots, n^\star$
  $n^\star = 50$

- **Distribuição, valor esperado e variância comuns**
  $X_i \overset{i.i.d.}{\sim} X, \quad i = 1, \ldots, n^\star$
  $X \sim \text{Hipergeométrica}(N, M, n)$, com $(N, M, n) = (100, 3, 10)$
  $E(X_i) = E(X) = \mu \overset{form.}{=} n\,\frac{M}{N} = 10 \times \frac{3}{100} = 0.3, \quad i = 1, \ldots, n^\star$
  $V(X_i) = V(X) = \sigma^2 \overset{form.}{=} n\,\frac{M}{N}\,\frac{N-M}{N}\,\frac{N-n}{N-1} = 10 \times \frac{3}{100} \times \frac{97}{100} \times \frac{90}{99} = \frac{291}{1100} = 0.26(45), \quad i = 1, \ldots, n^\star$

- **Nova v.a.**
  $S_{n^\star} = \sum_{i=1}^{n^\star} X_i$ = número de azulejos não conformes nas $n^\star$ amostras

- **Valor esperado e variância de $S_{n^\star}$**
  $E(S_{n^\star}) = E\left(\sum_{i=1}^{n^\star} X_i\right) = \sum_{i=1}^{n^\star} E(X_i) \overset{X_i \sim X}{=} n^\star\, E(X) = n^\star\, \mu = 50 \times 0.3 = 15$

  $V(S_{n^\star}) = V\left(\sum_{i=1}^{n^\star} X_i\right) \overset{X_i\, indep.}{=} \sum_{i=1}^{n^\star} V(X_i) \overset{X_i \sim X}{=} n^\star\, V(X) = n^\star\, \sigma^2 = 50 \times \frac{291}{1100} = \frac{291}{22} \simeq 13.2273$

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

$$
\frac{S_{n^\star} - E(S_{n^\star})}{\sqrt{V(S_{n^\star})}} = \frac{S_{n^\star} - n^\star\, \mu}{\sqrt{n^\star}\, \sigma} \overset{a}{\sim} \text{Normal}(0,1).
$$

- **Valor aproximado da prob. pedida**

$$
\begin{aligned}
P(S_{n^\star} > 25) &= 1 - P(S_{n^\star} \leq 25) \\
&= 1 - P\left(\frac{S_{n^\star} - n^\star\, \mu}{\sqrt{n^\star}\, \sigma} \leq \frac{25 - n^\star\, \mu}{\sqrt{n^\star}\, \sigma}\right) \\
&\overset{TLC}{\simeq} 1 - \Phi\left(\frac{25-15}{\sqrt{13.2273}}\right) \\
&\simeq 1 - \Phi(2.75) \\
&\overset{tabela/calc.}{\simeq} 1 - 0.9970 \\
&= 0.0030.
\end{aligned}
$$

2. Num pequeno laboratório químico trabalham dois engenheiros e dois estagiários. Suponha que as variáveis aleatórias X (resp. Y) indicam o número de engenheiros (resp. estagiários) envolvidos no desenvolvimento de um projecto selecionado ao acaso. Admita que a função de probabilidade conjunta de (X,Y) é dada pela tabela seguinte:

	Y		
X	0	1	2
0	0	0.1	0.1
1	0.1	0.3	0
2	0.05	0.2	0.15

(a) Averigúe se $E(X \mid Y=1) = E(X)$. Que pode concluir relativamente à independência entre as variáveis aleatórias X e Y? (2.0)

- **Par aleatório** (X,Y)
  X = nº de engenheiros envolvidos no desenvolvimento do projecto selecionado
  Y = nº de estagiários envolvidos no desenvolvimento de um projecto selecionado

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

	Y			
X	0	1	2	$P(X=x)$
0	0	0.1	0.1	0.2
1	0.1	0.3	0	0.4
2	0.05	0.2	0.15	0.4
$P(Y=y)$	0.15	0.6	0.25	1

- **F.p. de** $X \mid Y=1$

$$\begin{aligned} P(X=x \mid Y=1) &= \frac{P(X=x,Y=1)}{P(Y=1)} \\ &= \begin{cases} \frac{0.1}{0.6}=\frac{1}{6}, & x=0 \\ \frac{0.3}{0.6}=\frac{1}{2}, & x=1 \\ \frac{0.2}{0.6}=\frac{1}{3}, & x=2 \\ 0, & \text{restantes valores de } x \end{cases} \end{aligned}$$

- **Valor esperado de** $X \mid Y=1$

$$\begin{aligned} E(X \mid Y=1) &= \sum_{x=0}^{2} x \times P(X=x \mid Y=1) \\ &= 0\times\frac{1}{6}+1\times\frac{1}{2}+2\times\frac{1}{3} \\ &= \frac{7}{6} \\ &= 1.1(6) \end{aligned}$$

- **Valor esperado de** X

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

- **Comentário**
  Note-se que caso X e Y fossem v.a. independentes então $(X \mid Y=1)$ e X possuíriam a mesma distribuição e, consequentemente, $E(X \mid Y=1)=E(X)$. Ora,

$$\begin{aligned} E(X \mid Y=1) &= 1.1(6) \\ &\neq \\ E(X) &= 1.2, \end{aligned}$$

  logo X e Y são v.a. DEPENDENTES.

(b) Obtenha o valor esperado e a variância da variável aleatória $(X+Y)$, que representa o número total de engenheiros e estagiários envolvidos no desenvolvimento de um projecto selecionado ao acaso. (3.0)

- **Valor esperado pedido**

$$\begin{aligned} E(X+Y) &= E(X)+E(Y) \\ &\overset{(a)}{=} 1.2+\sum_{y=0}^{2} y\times P(Y=y) \\ &\overset{(a)}{=} 1.2+(0\times 0.15+1\times 0.6+2\times 0.25) \\ &= 1.2+1.1 \\ &= 2.3 \end{aligned}$$

- **Variância pedida**

Uma vez que se pretende calcular

$$\begin{aligned} V(X+Y) &= V(X)+V(Y)+2\times cov(X,Y) \\ &= V(X)+V(Y)+2\times[E(XY)-E(X)\times E(Y)], \end{aligned}$$

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

- **Valor esperado e variância de X**

$$\begin{aligned} E(X) &\overset{(a)}{=} 1.2 \\ V(X) &= E(X^2)-E^2(X) \\ &= \sum_{x=0}^{2} x^2\times P(X=x)-1.2^2 \\ &= (0^2\times 0.2+1^2\times 0.4+2^2\times 0.4)-1.2^2 \\ &= 2-1.44 \\ &= 0.56 \end{aligned}$$

- **Valor esperado e variância de Y**

$$\begin{aligned} E(Y) &= 1.1 \\ V(Y) &= E(Y^2)-E^2(Y) \\ &= \sum_{y=0}^{2} y^2\times P(Y=y)-1.2^2 \\ &= (0^2\times 0.15+1^2\times 0.6+2^2\times 0.25)-1.1^2 \\ &= 1.6-1.21 \\ &= 0.39 \end{aligned}$$

- **Valor esperado de XY**

$$\begin{aligned} E(XY) &= \sum_{x=0}^{2}\sum_{y=0}^{2} xy\times P(X=x,Y=y) \\ &= 1\times 1\times 0.3+2\times 1\times 0.2+2\times 2\times 0.15 \\ &= 1.3 \end{aligned}$$

- **Covariância**

$$\begin{aligned} cov(X,Y) &= E(XY)-E(X)\times E(Y) \\ &= 1.3-1.2\times 1.1 \\ &= -0.02 \end{aligned}$$

- **Variância pedida (cont.)**

$$\begin{aligned} V(X+Y) &= V(X)+V(Y)+2\times cov(X,Y) \\ &= 0.56+0.39+2\times(-0.02) \\ &= 0.91. \end{aligned}$$