DM DEPARTAMENTO DE MATEMÁTICA TÉCNICO LISBOA

Probabilidades e Estatística

LEAN + LEE + LEGI + LEMat + MEAer + MEAmbi + MEBiol + MEBiom + MEEC + MEMec + MEQ

1º semestre – 2013/2014
16/11/2013 – 9:00

Duração: 90 minutos **1º teste A**

Justifique convenientemente todas as respostas!

Grupo I	10 valores

1. Sabe-se que 2% das mensagens recebidas na caixa de correio electrónico de um utilizador vêm contaminadas com algum tipo de vírus informático. O programa anti-vírus instalado no computador do utilizador é tal que:

– a probabilidade de detectar a existência de vírus quando uma mensagem tem vírus é 0.99;

– a probabilidade de não detectar a existência de vírus quando uma mensagem não tem vírus é 0.9995.

(a) Qual é a probabilidade de o programa anti-vírus detectar a existência de vírus numa mensagem recebida pelo utilizador? (2.5)

Sejam V =“uma mensagem recebida está contaminada por algum vírus informático” e D =“o programa anti-vírus detecta a existência de vírus numa mensagem”. Tem-se $P(V) = 0.02$, $P(D \mid V) = 0.99$ e $P(\bar{D} \mid \bar{V}) = 0.9995$.
$P(D) = P(D \mid V)P(V) + P(D \mid \bar{V})P(\bar{V}) = P(D \mid V)P(V) + (1 - P(\bar{D} \mid \bar{V}))(1 - P(V)) = 0.0203$.

(b) Calcule a probabilidade de que uma mensagem tenha vírus sabendo que o programa anti-vírus assinalou essa presença. (2.0)

$P(V \mid D) = \frac{P(D\cap V)}{P(D)} = \frac{P(D|V)P(V)}{P(D)} = 0.9758$.

2. A chegada de autocarros a uma certa paragem é bem modelada por um processo de Poisson com valor esperado de 8 autocarros por hora.

(a) Calcule a probabilidade de num período de um quarto de hora chegarem pelo menos 3 autocarros à paragem. (1.0)

Seja $X(t)$ =“número de autocarros que chegam à paragem num período de t horas”. Tem-se que $X(t) \sim Poi(\lambda t)$ com $E[X(1)] = \lambda = 8$.
Uma vez que $X(0.25) \sim Poi(2)$, tem-se $P(X(0.25) \geq 3) = 1 - P(X(0.25) < 3) = 1 - P(X(0.25) \leq 2) = 1 - F_{X(0.25)}(2) = 1 - 0.6767 = 0.3233$.

(b) Considerando um conjunto de 5 períodos de um quarto de hora disjuntos, determine a probabilidade de haver no máximo um período em que chegam pelo menos 3 autocarros. (2.5)

Seja Y =“número de períodos em que chegam pelo menos 3 autocarros num conjunto de 5 períodos de um quarto de hora disjuntos”. Como Y representa o número de sucessos (chegada de pelo menos 3 autocarros) em 5 repetições independentes (períodos disjuntos num processo de Poisson) de uma prova de Bernoulli então $Y \sim Bi(5, p)$, com $p = P(X(0.25) \geq 3) = 0.3233$.
$P(Y \leq 1) = F_Y(1) = 0.4809$.

(c) Qual é a probabilidade de o tempo de espera entre dois autocarros consecutivos ser inferior a um quarto de hora? (2.0)

Seja T =“tempo que decorre entre dois autocarros consecutivos, em horas”. Tem-se $T \sim Exp(8)$ e $P(T < 1/4) = \int_0^{1/4} 8e^{-8t}\,dt = 1 - e^{-2} \approx 0.8647$.
Alternativa: $P(T < 1/4) = 1 - P(X(0.25) = 0) = 1 - F_{X(0.25)}(0)$.

Grupo II	10 valores

1. Uma empresa está a construir uma nova máquina para encher sacos de cimento e considera que o peso de cada saco enchido pela máquina, em Kg, é uma variável aleatória, X, com distribuição uniforme contínua num intervalo $[a,b]$.

(a) A cimenteira que encomendou a máquina pretende encher sacos com um peso médio de 50 Kg e requer ainda que 90% dos sacos tenham um peso superior a 48 Kg. Determine os valores das constantes a e b que satisfazem estas especificações. (2.0)

$E[X]=50 \iff \frac{a+b}{2}=50$ e $P(X>48)=0.9 \iff \int_{48}^{b}\frac{1}{b-a}\,dx=0.9 \iff \frac{b-48}{b-a}=0.9$.

$$\begin{cases} a+b=100 \\ 0.9a+0.1b=48 \end{cases} \iff \begin{cases} a=47.5 \\ b=52.5 \end{cases}$$

(b) Os sacos de cimento irão ser comercializados em lotes de 100 unidades. Calcule aproximadamente o valor da probabilidade de o peso de um lote exceder 5050 Kg, admitindo que os pesos de diferentes sacos são variáveis aleatórias independentes. (3.0)

Nota: se não resolveu a alínea (a) considere que $Var[X]=25/12$.

Sejam $X_i=$"peso do iº saco", $i=1,\ldots,100$ e $T=\sum_{i=1}^{100}X_i=$"peso de um lote de 100 sacos".
Temos que $E[T]=E[\sum_{i=1}^{100}X_i]=\sum_{i=1}^{100}E[X_i]=100E[X]=5000$ e $Var[T]=Var[\sum_{i=1}^{100}X_i]=\sum_{i=1}^{100}Var[X_i]=100Var[X]=\frac{2500}{12}$, uma vez que as variáveis X_i são independentes e identicamente distribuídas a X. Nestas condições, o teorema do limite central garante que

$$\frac{T-E[T]}{\sqrt{Var[T]}} \overset{a}{\sim} N(0,1).$$

$$P(T>5050)=P\left(\frac{T-5000}{\sqrt{2500/12}}>\frac{5050-5000}{\sqrt{2500/12}}\right)\approx 1-\Phi(\sqrt{12})=2.7\times10^{-4}.$$

2. Sejam X e Y duas variáveis aleatórias discretas que representam, respectivamente, o número de erros de sintaxe e de *input/output* num programa elaborado por um perito informático. De um estudo prévio, sabe--se que a função de probabilidade conjunta deste par aleatório é dada por:

$X\backslash Y$	0	1	2
0	0.2	0.3	0.1
1	0.1	0.2	0.1

(a) Determine o valor esperado de X condicionalmente a $Y=1$. (1.5)

$f_{X|Y=1}(x)=\frac{f_{X,Y}(x,1)}{f_Y(1)}=\begin{cases}0.6, & x=0\\ 0.4, & x=1\end{cases} \iff X|Y=1\sim Ber(0.4)$,
uma vez que $f_Y(1)=\sum_{x=0}^{1}f_{X,Y}(x,1)=0.5$. $E[X|Y=1]=\sum_{x=0}^{1}xf_{X|Y=1}(x)=0.4$.

(b) Calcule o coeficiente de correlação entre X e Y e comente o valor obtido. (3.5)

$Corr[X,Y]=\frac{Cov[X,Y]}{\sqrt{Var[X]Var[Y]}}=\frac{0.04}{\sqrt{0.24\times0.49}}\approx 0.1166.$

$f_X(x)=\sum_{y=0}^{2}f_{X,Y}(x,y)=\begin{cases}0.6, & x=0\\ 0.4, & x=1\end{cases} \iff X\sim Ber(0.4)$, $E[X]=0.4$ e $Var[X]=0.24$.

$f_Y(y)=\sum_{x=0}^{1}f_{X,Y}(x,y)=\begin{cases}0.3, & y=0\\ 0.5, & y=1\\ 0.2, & y=2\end{cases}$, $E[Y]=\sum_{y=0}^{2}yf_Y(y)=0.9$, $E[Y^2]=\sum_{y=0}^{2}y^2f_Y(y)=1.3$

$Var[Y]=E[Y^2]-(E[Y])^2=0.49$

$E[XY]=\sum_{x=0}^{1}\sum_{y=0}^{2}xyf_{X,Y}(x,y)=0.4$ e $Cov[X,Y]=E[XY]-E[X]E[Y]=0.04$.

Como $Corr[X,Y]>0$, existe uma associação linear positiva entre X e Y. No entanto, essa associação é fraca uma vez que o valor de $Corr[X,Y]$ é pequeno.