ESTATÍSTICA I - 2º Ano/Economia, 1º semestre, EN 05. 01.2021
2 horas (20 valores)

Nome:____ **Nº:**_________

Espaço reservado para classificações

1a.(10)	3a.(10)	4a.(15)	6a.(15)	T:
1b.(10)	3b.(10)	4b.(10)	6b.(10)	
2a.(15)	3c.(10)	4c.(10)	7. (10)	
2b.(15)	3d.(15)	5. (20)	8.(15)	

Atenção: todas as questões devem ser devidamente formalizadas e justificadas.

1. Seja (X, Y) uma variável aleatória bidimensional com função probabilidade conjunta dada por:

$y \setminus x$	0	1	2
0	0.12	0.22	0.06
1	0.21	0.10	0.03
2	0.07	0.06	0.01
3	0.07	0.02	0.03

a. Calcule $P(X < Y)$. Obtenha também $E(X + Y|Y = 1)$

$$P(X < Y) = f(0,1) + f(0,2) + f(0,3) + f(1,2) + f(1,3) + f(2,3)$$
$$= 0.21 + 0.07 + 0.07 + 0.06 + 0.02 + 0.03 = 0.46$$

$E(X + Y|Y = 1) = E(X|Y = 1) + E(Y|Y = 1) = X + 1 = Z$ já que

$E(Y|Y = 1) = 1$ e que

$$E(X|Y = 1) = \sum_{x=0}^{2} x\, f_{X|Y=1}(x) = 0 \times \frac{0.21}{0.34} + 1 \times \frac{0.10}{0.34} + 2 \times \frac{0.03}{0.34} = 0.4706$$

b. Obtenha a **função de distribuição** de X .

Começa por obter-se a função probabilidade de X , dada por $f_X(x) = \begin{cases} 0.47 & x = 0 \\ 0.40 & x = 1 \\ 0.13 & x = 2 \end{cases}$ para se determinar a

função de distribuição $F_X(x) = \begin{cases} 0 & x < 0 \\ 0.47 & 0 \le x < 1 \\ 0.87 & 1 \le x < 2 \\ 1 & x \ge 2 \end{cases}$

2. Seja X uma variável aleatória com função densidade dada por $f_X(x) = 2\,(x - 1),\ 1 < x < 2$.

a. Obtenha **a função de distribuição** de X e calcule $P(X < 1.5|X > 1.2)$.

$$F_X(x) = \int_{-\infty}^{x} f_X(u)du = \begin{cases} 0 & x < 1 \\ \int_1^x 2\,(u-1)\,du & 1 \le x < 2 \\ 1 & x \ge 2 \end{cases} = \begin{cases} 0 & x < 1 \\ x^2 - 2x + 1 & 1 \le x < 2 \\ 1 & x \ge 2 \end{cases}$$

Já que $\int_1^x 2\,(u-1)\,du = (u^2 - 2u]_1^x = (x^2 - 2x) - (1 - 2) = x^2 - 2x + 1$

$$P(X < 1.5|X > 1.2) = \frac{P(1.2<X<1.5)}{P(X>1.2)} = \frac{F_X(1.5)-F_X(1.2)}{1-F_X(1.2)} = \frac{0.21}{0.96} = 0.219$$

b. Seja Y outra variável aleatória cuja função densidade marginal, condicionada por $X = x$, é dada por $f_{Y|X=x}(y) = \frac{1}{x-1}$, $1 < y < x$ com $1 < x < 2$. Obtenha a **função de densidade marginal** de Y e calcule $P(X > 1.5, Y > 1.5)$.

$f_{X,Y}(x,y) = f_X(x)\ f_{Y|X=x}(y) = 2\,(x-1)\,\frac{1}{x-1} = 2 \qquad 1 < y < x\,, 1 < x < 2$

Logo

$f_Y(y) = \int_{-\infty}^{+\infty} f_{X,Y}(x,y)\,dx = \int_y^2 2\,dx = 2\,(2-y) \qquad 1 < y < 2$

$P(X > 1.5, Y > 1.5) = \int_{1.5}^2 \int_{1.5}^x 2\,dy\,dx = \int_{1.5}^2 (2y]_{1.5}^x\,dx = \int_{1.5}^2 (2x-3)\,dx = \left(x^2 - 3x\right]_{1.5}^2 = 0.25$

3. O número de livros vendidos num site de *online* segue um processo de Poisson com taxa média de 4 por hora.

a. Qual a probabilidade de se venderem mais de 9 livros num período de hora e meia?

X_1 – Nº de livros vendidos numa hora $\qquad X_1 \sim Po(4)$

X_2 – Nº de livros vendidos numa hora e meia $\qquad X_2 \sim Po(6)$

$P(X_2 > 9) = 1 - P(X_2 \leq 9) = 1 - 0.9161 = 0.0839$

b. Sabendo que entre as 10 e as 11 se venderam 6 livros qual a probabilidade de se venderem 5 na meia-hora seguinte?

De acordo com as regras do processo de Poisson, os acontecimentos em intervalos de tempo disjuntos são independentes, logo basta calcular a probabilidade de se venderem 5 livros em meia hora

X_3 – Nº de livros vendidos em meia hora $\qquad X_3 \sim Po(2)$

$P(X_3 = 5) = 0.0361$

c. Qual a variância do tempo necessário para vender 2 livros?

De acordo com as regras do processo de Poisson, o tempo entre acontecimentos consecutivos é dado por uma exponencial e portanto o tempo necessário para vender 2 livros, seja T, será dado por uma gama. Assim, $T \sim G(2,4)$ estando o tempo medido em horas.

Logo $var(T) = \frac{2}{4^2} = 0.125$

d. O período da manhã (entre as 8 e o meio-dia) for dividido em 8 meias horas. Qual a probabilidade de não se venderem livros apenas num desses períodos.

Y – Nº de períodos de meia-hora (dos 8 possíveis) em que não se vendem livros

$Y \sim B(8,p)$ com $p = P(X_3 = 0) = 0.1353$

$P(Y = 1) = \binom{8}{1}\ 0.1353\ \times (1 - 0.1353)^7$ =0.3912

Ou arredondar *p* para 0.15 e ir à tabela → 0.3847

Ou ir à tabela e calcular a média com p=0.10 e p=0.15 → (0.3826+0.3847)/2=0.38365

4. O peso de um pacotinho de açúcar é uma variável aleatória com distribuição Normal de média 6 gramas e variância 1. Admite-se que o enchimento dos pacotes é feito de forma independente.

a. Os saquinhos são embalados em caixas de 10 unidades. Qual a percentagem de caixas com peso inferior a 58 gramas?

X_i – Peso (em gramas) do pacotinho i $\qquad X_i \sim n(6,1)$
$Y = \sum_{i=1}^{10} X_i$ – peso de uma caixa de 10 unidades $\qquad Y \sim n(60,10)$
$P(Y < 58) = \Phi\left(\frac{58-60}{\sqrt{10}}\right) = \Phi(-0.63) = 1 - \Phi(0.63) = 1 - 0.7357 = 0.2643$
Temos então 26.43% de caixas com peso inferior a 58 gramas

b. Em cada caixa com 10 saquinhos qual a probabilidade de haver mais de 3 saquinhos com peso inferior a 5 gramas.

Seja p a probabilidade de um saquinho pesar menos de 5 gramas e N o número de saquinhos (em 10) com peso inferior a 5 gramas. $\qquad N \sim B(10; p)$
$p = P(X < 5) = \Phi\left(\frac{5-6}{1}\right) = \Phi(-1) = 1 - \Phi(1) = 1 - 0.8413 = 0.1587$
$P(N > 3) = 1 - P(N \leq 3) = 0.06$ ou 0.05 recorrendo à tabela com $p = 0.15$

c. Calcule o limite inferior de gramas de açúcar que se encontra em 80% dos saquinhos.

Procura-se o valor k tal que $P(X \geq k) = 0.8$.
Assim $0.8 = P(X \geq k) = P\left(Z \geq \frac{k-6}{1}\right)$ sendo $Z \sim n(0; 1)$. Logo $k - 6 = -0.842$ e portanto $k = 5.158$

5. Seja X uma variável aleatória contínua com distribuição Uniforme entre 0 e 4. Da população X recolheu-se uma amostra aleatória de dimensão 75. Qual a probabilidade de a média da amostra diferir da média da população por um valor superior a 0.2?

O objetivo é calcular $P(|\bar{X} - \mu| > 0.2)$
Vai-se aplicar o TLC com $\mu = 2$ e $\sigma^2 = 16/12$ (média e variância de uma uniforme (0,4)). Pelo TLC sabe-se que
$Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} = \frac{\bar{X}-\mu}{0.1333} \dot{\sim} n(0,1)$
Assim
$$P(|\bar{X} - \mu| > 0.2) = 1 - P(|\bar{X} - \mu| \leq 0.2) = 1 - P(-0.2 \leq \bar{X} - \mu \leq 0.2) = 1 - P(-1.5 \leq Z \leq 1.5)$$
$$= 1 - \left(\Phi(1.5) - \Phi(-1.5)\right) = 1 - \left(\Phi(1.5) - \left(1 - \Phi(1.5)\right)\right) = 2 \times 0.0668 = 0.1336$$

6. A principal causa do denominado "cheiro a rolha" em garrafas de vinho é a contaminação do vedante de cortiça por um composto químico designado abreviadamente por TCA. Alguns estudos estimam em 3% a percentagem de garrafas de vinho afetadas por esta contaminação. No entanto, a capacidade de detetar a sua presença utilizando apenas o aparelho olfativo varia de indivíduo para indivíduo.
Um reputado crítico de vinhos afirma que deteta 97% das garrafas contaminadas e que, se uma garrafa não estiver contaminada, a probabilidade de ele a classificar erradamente é de 0.01. Nestas condições,

a. O crítico classificou uma garrafa como tendo "cheiro a rolha". Qual a probabilidade de ele se ter enganado?

Sejam os acontecimentos A "A garrafa está contaminada" e B "O critico diz que a garrafa está contaminada". Procura-se $P(\bar{A}|B)$, isto é a garrafa não estava contaminada muito embora o crítico dissesse o contrário. Notando que $\{A, \bar{A}\}$ forma uma partição do espaço de resultados, aplica-se a fórmula de Bayes,

$$P(\bar{A}|B) = \frac{P(\bar{A} \cap B)}{P(B)} = \frac{P(B|\bar{A}) \times P(\bar{A})}{P(B|A) \times P(A) + P(B|\bar{A}) \times P(\bar{A})} = \frac{0.01 \times 0.97}{0.97 \times 0.03 + 0.01 \times 0.97} = 0.25$$

b. Ao receber um lote com milhares de garrafas, um importador contrata os serviços deste crítico para avaliar o vinho de 5 garrafas selecionadas casualmente desse lote. Se pelo menos 2 garrafas forem classificadas pelo crítico como tendo "cheiro a rolha", todo o lote é rejeitado. Qual a probabilidade de o importador rejeitar o lote? (Nota: caso não tenha resolvido a alínea a) pode assumir que a resposta era 0.05, que não é a resposta correta)

Seja X o número de garrafas rejeitadas (em 5), logo $X \sim B(5, p)$ em que p é a probabilidade de o crítico dar a garrafa como má. Da alínea anterior sabe-se que $p = P(B) = 0.97 \times 0.03 + 0.01 \times 0.97 = 0.0388$.

$$P(X \geq 2) = 1 - P(X \leq 1) = 1 - \left(\binom{5}{0} 0.0388^0 \times 0.9612^5 + \binom{5}{1} 0.0388^1 \times 0.9612^4\right)$$
$$= 1 - (0.9612^5 + 5 \times 0.0388^1 \times 0.9612^4) = 1 - 0.9861 = 0.0139$$

Pela tabela com $p = 0.05$ vem uma probabilidade de 0.0226

7. Prove que dois acontecimentos incompatíveis, ambos de probabilidade positiva, definidos no mesmo espaço de resultados não podem ser independentes.

   Se dois acontecimentos, A e B são incompatíveis, $A \cap B = \emptyset$ logo , $P(A \cap B) = 0$
   Se dois acontecimentos, A e B são independentes, $P(A \cap B) = P(A) \times P(B) > 0$ já que $P(A) > 0$ e $P(B) > 0$
   Logo se os acontecimentos são incompatíveis, não podem ser independentes já que $P(A \cap B)$ não pode ser simultaneamente positivo e nulo.

8. Sejam três variáveis aleatórias independentes, $X \sim n(0, 1)$, $Y \sim n(0, 3)$ e $Z \sim \chi^2(5)$. Calcule $P\left(\frac{3\,X^2 + Y^2}{2\,Z} > 9.01\right)$.

$$\left.\begin{array}{c} X \sim n(0,1) \to X^2 \sim \chi^2_{(1)} \\ Y \sim n(0,3) \to \dfrac{Y^2}{3} \sim \chi^2_{(1)} \\ X^2 \ e \ Y^2 \ indep. \end{array}\right\} \to U = X^2 + \frac{Y^2}{3} = \frac{3\,X^2 + Y^2}{3} \sim \chi^2_{(2)}$$

E portanto $F = \frac{U/2}{Z/5} \sim F(2;\ 5)$.
Ora $F = \frac{U/2}{Z/5} = \frac{\frac{3\,X^2+Y^2}{6}}{\frac{Z}{5}} = \frac{5\left(3\,X^2+Y^2\right)}{6\,Z} = \frac{5}{3}\,\frac{\left(3\,X^2+Y^2\right)}{2\,Z}$.
Assim $P\left(\frac{3\,X^2+Y^2}{2\,Z} > 9.01\right) = P\left(F > \frac{5}{3} \times 9.01\right) = P(F > 15.01) < 0.01$ (valor de 13.27 na tabela)