Probabilités et statistiques
M1 - 2009-2010 - Examen

Exercice 1 : On pose

$$K = \begin{bmatrix} 1 & 3 & -1 \\ 3 & 9 & -3 \\ -1 & -3 & 2 \end{bmatrix}$$

et on note V le vecteur $(1,1,1)$.

1. Montrer que K est une matrice symétrique positive. On note dans la suite (X,Y,Z) un vecteur aléatoire gaussien d'espérance V et de matrice de covariance K.
2. Quelle est la loi du vecteur aléatoire $(X+Y, Y-Z)$?
3. Montrer qu'il existe un unique réel λ tel que $X+\lambda Y$ et X soient indépendants.
4. En déduire $E(Y|X)$.

Exercice 2 : Soient X, Y, Z trois variables aléatoires indépendantes de loi uniforme sur $[-1,1]$. On pose $A = (X+Y+Z)^2$. Calculer $E(A|X)$.

Exercice 3 : Soit (X,Y) un vecteur aléatoire de $\mathbb{R}^2$ dont la loi admet pour densité la fonction définie par

$$(x,y) \mapsto 1_{[0,1]}(x) 1_{[0,1]}(y) c(x+y)$$

où c est une constante.

1. Que vaut la constante c ?
2. Donner la densité de la loi de X.
3. Donner la famille des lois conditionnelles de Y sachant X.
4. Calculer $E(Y|X)$.
5. Calculer $E(XY^2|X)$.

Exercice 4 : Soit $(T_n)_{n\geq 1}$ une suite de variables aléatoires indépendantes de loi exponentielle de moyenne $m > 0$ [1]. L'objectif de cet exercice est d'étudier une méthode d'estimation de m. On introduit pour cela pour tout entier n et tout réel $m' > 0$ les variables aléatoires

$$M_n = \frac{T_1 + \cdots + T_n}{n}$$

et

$$R_n^{m'} = \frac{2nM_n}{m'} = \frac{2(T_1 + \cdots + T_n)}{m'}.$$

[1] On rappelle qu'une variable aléatoire suit une loi exponentielle de moyenne $m > 0$ si sa loi admet pour densité la fonction définie par

$$x \mapsto 1_{\mathbb{R}_+}(x) \frac{1}{m} \exp(-x/m).$$

L'espérance d'une telle variable aléatoire est m.

Enfin, pour tout n, on note a_n et b_n les réels tels que $P(C_n \leq a_n) = P(C_n > b_n) = 0.025$ où C_n suit une loi du chi-deux à $2n$ degrés de liberté [2].

1. Calculer $E(M_n)$ pour tout entier $n \geq 1$.
2. Montrer que M_n converge presque sûrement vers m.
3. Montrer que $2T_1/m$ suit une loi exponentielle de moyenne 2.
4. Montrer que la loi exponentielle de moyenne 2 est aussi la loi du chi-deux à 2 degrés de libertés.
5. En déduire que, pour tout n, R_n^m suit une loi du chi-deux à $2n$ degrés de liberté.
6. En déduire, pour tout n, $P(R_n^m \in [a_n, b_n]) = 0.95$.
7. Soit $\varepsilon > 0$. Montrer que $P(R_n^m/n \in [2-\varepsilon, 2+\varepsilon])$ tend vers 1 quand n tend vers l'infini.
8. Déduire de la question précèdente que, pour tout $\varepsilon > 0$, on a $[a_n, b_n] \subset [2n - n\varepsilon, 2n + n\varepsilon]$ pour n assez grand.
9. Soit $m' \neq m$. Déduire des deux questions précédentes que $P(R_n^{m'} \in [a_n, b_n])$ tend vers 0 quand n tend vers l'infini.
10. On se donne $(t_1, \ldots, t_n)$ un échantillon que l'on suppose issu d'une suite de variables aléatoires indépendantes de loi exponentielle de moyenne m inconnue. On se donne un réel $m' > 0$. Comment construire un test statistique pour déterminer si l'hypothèse $m = m'$ est plausible ?
11. L'une des questions précédentes permet de montrer que $b_n - a_n$ est négligeable devant n. Peut-on être plus précis ?

[2] On rappelle que la loi du chi-deux à d degrés de liberté est la loi de la somme des carrés de d variables aléatoires indépendantes de loi gaussienne centrée réduite. On rappelle également que la loi gaussienne centrée réduite admet pour densité la fonction définie par

$$x \mapsto \frac{1}{\sqrt{2\pi}} \exp(-x^2/2).$$