TELECOM et Management SudParis
Institut TELECOM
CONTROLE DE CONNAISSANCES

Module MAT3002 : Probabilités (deuxième partiel)
Durée : 1,5h
Sans document et sans calculatrice (une feuille manuscrite) A4 est autorisée
Date : 27 Février 2009
Coordonnateur : DELMAS Jean Pierre

Nom : Prénom : Année (1,2,3)

Justifier chaque réponse de façon concise. Répondre sur le questionnaire même dans les limites qui vous sont proposées. Il est conseillé de faire un rapide brouillon sur une feuille à part avant de reporter votre réponse. Les quatres exercices sont indépendants.

1] **Génération d'une suite de variables aléatoires**

On désire générer par des moyens informatiques une suite X_n de variables aléatoires indépendantes de loi de probabilité de Rayleigh (c'est à dire de densité de probabilité $f_X(x) = \frac{x}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}} \mathbb{1}_{[0,\infty[}(x)$), à partir d'un générateur de suite U_n de variables aléatoires indépendantes de loi de Bernoulli à valeurs $\{0, 1\}$ équiprobables. Proposer au moins une méthode.

2] **Matrice de covariance**

On considère une variable aléatoire discrète réelle à deux dimensions $\mathbf{X} = (X_1, X_2)$ qui prend ses valeurs dans $\mathbb{R}^2$ dont la probabilité est définie par les probabilités 1/5 aux points (0,0) (1,0), (0,1), (-1,0) et (0,-1).
2.1] Quelle est la matrice de covariance $\mathbf{C}_X$ de cette variable aléatoire discrète bidimensionelle $\mathbf{X}$?
2.2] Les variables aléatoires X_1 et X_2 sont elles non corrélées, indépendantes ?

3] **Théorème de transfert**

Soit a une constante réelle positive et X une variable aléatoire de densité de probabilité $f_X(x)$.
3.1] Donner l'expression de $\mathrm{E}(Y)$ pour $Y = \mathrm{Max}(X, a)$, en fonction de a et $f_X(x)$. (On pourra utiliser avec profit le théorème de transfert que l'on précisera).
3.2] Calculer $\mathrm{E}(Y)$ dans le cas particulier où $f_X(x) = \lambda e^{-\lambda x} \mathbb{1}_{[0,\infty}[(x)$ Que remarque-t-on lorsque a tend vers 0 ?

4] **Calcul d'espérance mathématique, variance, coefficient de corrélation et d'espérance conditionnelle**

Soit (X, Y) une variable aléatoire bidimensionnelle de densité de probabilité :

$$f_{X,Y}(x, y) = 2\ \mathbb{1}_{(x\geq 0)\cap(y\geq 0)\cap(x+y\leq 1)}$$

4.1] Déterminer les densités de probabilités marginales $f_X(x)$ et $f_Y(y)$. En déduire les expressions de $\mathrm{E}(X)$, $\mathrm{E}(Y)$, $\mathrm{Var}(X)$, $\mathrm{Var}(Y)$.
4.2] Calculer le coefficient de corrélation entre X et Y et recalculer $\mathrm{E}(X)$, $\mathrm{E}(Y)$, $\mathrm{Var}(X)$, $\mathrm{Var}(Y)$ à l'aide du théorème de transfert que l'on précisera.
4.3] Déterminer les densités de probabilité conditionnelles $f_{X/Y=y}(x)$ et $f_{Y/X=x}(y)$. Les variables aléatoires X et Y sont elles indépendantes ?

4.4] Calculer $\mathrm{E}[X/Y=y]$ et en particulier $\mathrm{E}[X/Y=\frac{1}{2}]$. Quelle est la meilleure approximation $\widehat{X}$ de X comme fonction $g(.)$ de Y au sens de $\min_g \mathrm{E}([g(Y)-X]^2)$. Interprétation des résultats des questions 4.3] et 4.4] ?

Correction très succincte du contrôle de connaissance de MAT3002 deuxième partiel

1] **Génération d'une suite de variables aléatoires**

A partir de la suite U_n, on peut générer une suite de variables aléatoires V_n indépendantes de loi uniforme sur [0,1] en regroupant K valeurs successives de U_n.

$$V_n = \sum_{k=1}^{K} U_{nK+k} 2^{-k}$$

où K dépend du processeur disponible. On réalise en réalité une approximation car V_n est une variable aléatoire discrète à 2^K valeurs équiprobables régulièrement réparties dans $\{0, \frac{1}{2^K}, ..., \frac{k}{2^K}, 1-\frac{1}{2^K}\}$. Puis à l'aide de la méthode de la fonction de répartition inverse, puisque :

$$F_X(x) = \int_{-\infty}^{x} f_X(t)dt = (1-e^{-\frac{x^2}{2\sigma^2}})\mathbb{1}_{[0,\infty[}(x),$$

on obtient la génération suivante :

$$X_n = \sigma\sqrt{-2\ln(1-V_n)}$$

où plus simplement $X_n = \sigma\sqrt{-2\ln(V_n)}$, car les variables aléatoires $1-V_n$ et V_n ont même loi de probabilité uniforme sur [0,1].

2] **Matrice de covariance**

2.1] $\mathrm{E}(X_1) = \sum_{k=1}^{5} x_{1,k}\frac{1}{5} = \frac{1}{5} - \frac{1}{5} = 0$ et $\mathrm{E}(X_2) = \sum_{k=1}^{5} x_{2,k}\frac{1}{5} = \frac{1}{5} - \frac{1}{5} = 0$,
puis
$\mathrm{E}(X_1X_2) = \sum_{k=1}^{5} x_{1,k}x_{2,k}\frac{1}{5} = 0$, $\mathrm{E}(X_1^2) = \sum_{k=1}^{5} x_{1,k}^2\frac{1}{5} = \frac{1}{5} + \frac{1}{5} = \frac{2}{5}$ et $\mathrm{E}(X_2^2) = \sum_{k=1}^{5} x_{2,k}^2\frac{1}{5} = \frac{1}{5} + \frac{1}{5} = \frac{2}{5}$,
d'où
$\mathrm{Cov}(X_1, X_2) = \mathrm{E}(X_1X_2) - \mathrm{E}(X_1)\mathrm{E}(X_2) = 0$. Par suite :

$$\mathbf{C}_X = \frac{1}{5}\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}.$$

2.2] Les variables aléatoires X_1 et X_2 sont non corrélées car $\mathrm{Cov}(X_1, X_2) = 0$. Elles ne sont pas indépendantes car $P[(X_1=0)\cap(X_2=0)] \overset{\text{def}}{=} P[(X_1, X_2) = (0,0)] = \frac{1}{5} \neq P[X_1=0]P[X_2=0] = \frac{3}{5}\frac{3}{5}$.

3] **Théorème de transfert**

3.1] D'après le théorème de transfert,

$$\begin{aligned} \mathrm{E}(Y) &= \int_{-\infty}^{+\infty} \mathrm{Max}(X,a) f_X(x)dx \\ &= \int_{-\infty}^{a} a f_X(x)dx + \int_{a}^{+\infty} x f_X(x)dx = aF_X(a) + \int_{a}^{+\infty} x f_X(x)dx. \end{aligned}$$

3.2] Dans le cas particulier où $f_X(x) = \lambda e^{-\lambda x}\mathbb{1}_{[0,\infty[}(x)$, $F_X(x) = (1-e^{-\lambda x})\mathbb{1}_{[0,\infty[}(x)$. Alors :

$$\mathrm{E}(Y) = a(1-e^{-\lambda a}) + \int_{a}^{+\infty} x\lambda e^{-\lambda x}dx = a(1-e^{-\lambda a}) + ae^{-\lambda a} + \frac{1}{\lambda}e^{-\lambda a} = a + \frac{1}{\lambda}e^{-\lambda a}.$$

Lorsque a tend vers 0, $\mathrm{E}(Y)$ tend vers $\frac{1}{\lambda} = \mathrm{E}(X)$. Ce qui est assez intuitif car dans ce cas la variable aléatoire Y converge en loi vers la variable aléatoire X.

4] **Calcul d'espérance mathématique, variance, coefficient de corrélation et d'espérance conditionnelle**

4.1] $f_X(x) = \int_{-\infty}^{+\infty} f_{X,Y}(x,y)dy = \int_0^{1-x} 2dy = 2(1-x)\mathbb{1}_{[0,1]}(x)$.

De même :

$f_Y(y) = 2(1-y)\mathbb{1}_{[0,1]}(y)$.

Par suite

$\mathrm{E}(X) = \int_{-\infty}^{+\infty} x f_X(x)dx = \int_0^1 2x(1-x)dx = \frac{1}{3}$ et $\mathrm{E}(X^2) = \int_{-\infty}^{+\infty} x^2 f_X(x)dx = \int_0^1 2x^2(1-x)dx = \frac{1}{6}$.

avec de même $\mathrm{E}(Y) = \frac{1}{6}$.

Par suite $\mathrm{Var}(X) = \mathrm{E}(X^2) - [\mathrm{E}(X)]^2 = \frac{1}{6} - \frac{1}{3^2} = \frac{1}{18}$, de même $\mathrm{Var}(Y) = \frac{1}{18}$.

4.2] Soit Δ, le domaine du plan défini par $\Delta \stackrel{\text{def}}{=} \{(x \geq 0) \cap (y \geq 0) \cap (x+y \leq 1)\}$. Alors en utlisant le théorème de transfert qui reporte le calcul d'espérance mathématique sur la variable aléatoire (X,Y), on obtient :

$$\mathrm{E}(X) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x f_{X,Y}(x,y)dxdy = \int\int_{\Delta} 2xdxdy = \int_0^1 2x\left[\int_0^{1-x} dy\right]dx = \frac{1}{3}$$

avec de même $\mathrm{E}(Y) = \frac{1}{3}$.

$$\mathrm{E}(X^2) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x^2 f_{X,Y}(x,y)dxdy = \int\int_{\Delta} 2x^2dxdy = \int_0^1 2x^2\left[\int_0^{1-x} dy\right]dx = \frac{1}{6}$$

Le coefficient de corrélation entre X et Y est $\rho_{X,Y} \stackrel{\text{def}}{=} \frac{\mathrm{Cov}(X,Y)}{\sigma_X\sigma_Y}$ avec $\mathrm{Cov}(X,Y) = \mathrm{E}(XY) - \mathrm{E}(X)\mathrm{E}(Y)$ où :

$$\mathrm{E}(XY) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xy f_{X,Y}(x,y)dxdy = \int\int_{\Delta} 2xydxdy = \int_0^1 2x\left[\int_0^{1-x} ydy\right]dx = \frac{1}{12}.$$

Par suite :

$$\mathrm{Cov}(X,Y) = \frac{1}{12} - \left(\frac{1}{3}\right)^2 = -\frac{1}{36} \quad \text{et} \quad \rho_{X,Y} = -\frac{1}{36}/\frac{1}{18} = -\frac{1}{2}.$$

4.3] $f_{X/Y=y}(x) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \frac{1}{1-y}\mathbb{1}_{[0,1-y]}(x)$ pour $y \in [0,1[$, c'est à dire une loi de probabilité uniforme sur $[0,1-y]$.

De même

$f_{Y/X=x}(y) = \frac{1}{1-x}\mathbb{1}_{[0,1-x]}(y)$ pour $x \in [0,1[$.

Puisque $f_{X/Y=y}(x) \neq f_X(x)$, les variables aléatoires X et Y ne sont pas indépendantes.

4.4]

$$\mathrm{E}[X/Y=y] = \int_{-\infty}^{+\infty} x f_{X/Y=y}(x)dx = \frac{1-y}{2}$$

et en particulier $\mathrm{E}[X/Y=\frac{1}{2}] = \frac{1}{4}$.

Puisque d'après le théorème de transfert

$$\begin{aligned}\mathrm{E}([X-g(Y)]^2) &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}[x-g(y)]^2 f_{X,Y}(x,y)dxdy \\ &= \int_{-\infty}^{+\infty}\left[\int_{-\infty}^{+\infty}[x-g(y)]^2\frac{f_{X,Y}(x,y)}{f_Y(y)}dx\right]f_Y(y)dy \quad \text{pour} \quad f_Y(y) \neq 0\end{aligned}$$

et que $\min_a \mathrm{E}([X-a]^2)$ a pour solution $a = \mathrm{E}(X) = \int_{-\infty}^{+\infty} x f_X(x)dx$, la "meilleure" fonction $g(.)$ sera donc donnée par $g(y) = \int_{-\infty}^{+\infty} x f_{X/Y=y}(x)dx$ avec $f_{X/Y=y}(x) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$.

Les résultats 4.3] et 4.4] peuvent se déduire de la remarque que la loi de probabilité de (X,Y) est de loi uniforme sur Δ.