Master 1 Mathématiques, MAT414 Université Joseph Fourier
2009-2010

Feuille d'exercices 2 : encore des espérances conditionnelles...

Exercice 1. *Couple de variables aléatoires discrètes.* Soit $\lambda > 0$, $\mu > 0$ et $0 < \nu \leq 1$. Soit (X, Y) un couple de variables aléatoires entières positives telles que

$$P_{(X,Y)}(n,m) = c_{\lambda,\mu,\nu} \frac{\lambda^n \mu^m \nu^{nm}}{n!\, m!} \quad \text{pour tout } (n,m) \in \mathbb{N}^2,$$

où $c_{\lambda,\mu,\nu}$ est une constante de normalisation.

1. Vérifier que $\displaystyle\sum_{(n,m)\in\mathbb{N}^2} \frac{\lambda^n \mu^m \nu^{nm}}{n!\, m!} < \infty$ (la constante $c_{\lambda,\mu,\nu}$ est alors l'inverse de cette somme).
2. Donner les lois marginales de X et Y. Trouver une condition nécessaire et suffisante pour que X et Y soient indépendantes.
3. Déterminer $\mathbb{E}[X|Y]$.

Exercice 2. *Baisse de 100% du pouvoir d'achat au Casino.* Un joueur possède une fortune $X_n > 0$ à l'instant n, avec $X_0 = x_0$ déterministe. A chaque instant $n \in \mathbb{N}^\star$ il joue à pile ou face une mise qu'il double s'il gagne et perd dans le cas contraire. Le joueur décide d'arrêter de jouer quand il possède une somme s fixée d'avance (s déterministe, $s \geq x_0$) ou quand il est ruiné (instant N pour lequel $X_N = 0$). Sa stratégie est de miser à chaque coup toute sa fortune X_n si $X_n \leq s/2$ et de miser $s - X_n$ si $X_n > s/2$.

1. Le jeu s'arrête t'il avec probabilité 1 au bout d'un nombre fini de coups ? Justifier votre réponse.
2. Montrer que pour tout $(k,n) \in \mathbb{N}^2$, $k \geq n$, on a $\mathbb{E}[X_k|X_1, X_2, \cdots, X_n] = X_n$.
3. En déduire la probabilité que le joueur termine le jeu en étant ruiné.

Exercice 3. *Indépendance conditionnelle.* Deux tribus $\mathcal{G}_1$ et $\mathcal{G}_2$ sont dites conditionnellement indépendantes par rapport à la tribu $\mathcal{G}$ si

$$\mathbb{E}[X_1 X_2|\mathcal{G}] = \mathbb{E}[X_1|\mathcal{G}]\, \mathbb{E}[X_2|\mathcal{G}]$$

pour toute variable aléatoire X_1 (respectivement X_2) positive $\mathcal{G}_1$-mesurable (resp. $\mathcal{G}_2$-mesurable).

1. Montrer que $\mathcal{G}_1$ et $\mathcal{G}_2$ sont conditionnellement indépendantes par rapport $\mathcal{G}$ si et seulement si $\mathbb{E}[X_2|\sigma(\mathcal{G}, \mathcal{G}_1)] = \mathbb{E}[X_2|\mathcal{G}]$ pour toute variable aléatoire X_2 positive $\mathcal{G}_2$-mesurable.
2. Montrer que $\mathcal{G}_1$ et $\mathcal{G}_2$ peuvent être indépendantes sans être conditionnellement indépendantes par rapport à une troisième tribu $\mathcal{G}$.
3. Soit (X_0, X_1, X_2) un triplet de variables aléatoires indépendantes. Montrer que $\sigma(X_0+X_1+X_2)$ et $\sigma(X_0)$ sont conditionnellement indépendantes par rapport à $\sigma(X_0 + X_1)$.

Exercice 4. *Maximum de variables aléatoires i.i.d.* Soit $(X_k)_{1\leq k\leq n}$ des variables aléatoires réelles indépendantes de même loi admettant une densité $\rho(x)$ par rapport à la mesure de Lebesgue.

1. Calculer la loi de la variable aléatoire $M = \max\{X_1, \cdots, X_n\}$.
2. Déterminer la loi conditionnelle de X_k sachant M.

Exercice 5. *Régression linéaire.* Soit $X = {}^t(X_1, \cdots, X_n)$ un vecteur aléatoire dans $\mathbb{R}^n$ et Y une variable aléatoire réelle. On suppose que $X_1, \cdots, X_n$ et Y sont de carré intégrable. On veut approcher Y par une fonction affine des X_k de la forme $A \cdot X + b$, où $A \in \mathbb{R}^n$ et $b \in \mathbb{R}$ sont déterministes. Dans la suite, on considère A comme un vecteur ligne et X comme un vecteur colonne, de sorte que AX coïncide avec le produit scalaire canonique $A \cdot X$ de A et de X dans $\mathbb{R}^n$. On cherche A et b de manière à minimiser la distance au carré $\mathbb{E}[(Y - AX - b)^2]$.

1. On considère d'abord le cas $n = 1$. Montrer que la meilleure approximation $\widehat{Y} = AX + b$ de Y au sens de la distance dans $L^2(\Omega, \mathcal{F}, \mathbb{P})$ ci-dessus vaut :
$$\widehat{Y} = \frac{\mathrm{cov}(X, Y)}{\mathrm{var}(X)} \big(X - \mathbb{E}[X]\big) + \mathbb{E}[Y] \, .$$

2. On étudie à présent le cas $n \geq 2$. Soit $\widetilde{X} = X - \mathbb{E}[X]$ et $\widetilde{Y} = Y - \mathbb{E}[Y]$. On note respectivement $\Gamma_X = \mathbb{E}[\widetilde{X}\, {}^t\widetilde{X}]$ et $\Gamma_{Y,X} = \mathbb{E}[\widetilde{Y}\, {}^t\widetilde{X}])$ la matrice de covariance de X (matrice $n \times n$) et la matrice d'intercovariance de Y et X (vecteur ligne). La matrice de covariance du vecteur aléatoire ${}^t(X_1, \cdots, X_n, Y)$ vaut alors
$$\Gamma = \begin{pmatrix} \Gamma_X & {}^t\Gamma_{Y,X} \\ \Gamma_{Y,X} & \mathrm{var}(Y) \end{pmatrix} .$$
   (a) Montrer que
$$\begin{aligned} \mathbb{E}[(Y - AX - b)^2] &= \big(A\Gamma_X^{1/2} - \Gamma_{Y,X}\Gamma_X^{-1/2}\big)\, {}^t\big(A\Gamma_X^{1/2} - \Gamma_{Y,X}\Gamma_X^{-1/2}\big) \\ &\quad + \mathrm{var}(Y) - \Gamma_{Y,X}\Gamma_X^{-1}\, {}^t\Gamma_{Y,X} + \big(\mathbb{E}[Y] - A\mathbb{E}[X] - b\big)^2 . \end{aligned}$$
   (b) En déduire que la meilleure approximation de Y par une fonction affine des X_k vaut :
$$\widehat{Y} = \Gamma_{Y,X}\Gamma_X^{-1}\big(X - \mathbb{E}[X]\big) + \mathbb{E}[Y] \, .$$

3. On suppose que $(X_1, \cdots, X_n, Y)$ est un vecteur aléatoire gaussien. Montrer que
$$\widehat{Y} = \mathbb{E}[Y | X_1, \cdots, X_n] \, .$$

Exercice 6. *Processus de Poisson.* Un matériau radioactif émet des particules à des temps aléatoires $T_0 = 0 \leq T_1 \leq T_2 \leq \cdots \leq T_{k-1} \leq T_k \leq \cdots$. On suppose que les temps d'attente $\Delta_k = T_k - T_{k-1}$ entre deux émissions successives sont indépendants entre eux et identiquement distribués selon une loi exponentielle de paramètre λ. Pour tout $t > 0$, soit
$$N_t = \sum_{k=1}^{\infty} 1_{\{T_k \leq t\}}$$
la variable aléatoire comptant le nombre d'émissions dans l'intervalle de temps $]0, t]$.

1. Montrer que N_t suit une loi de Poisson de paramètre λt.

2. Déterminer les densités des lois des variables aléatoires T_k, $k \in \mathbb{N}^\star$.

3. On suppose que les particules émises sont détectées à partir du temps $s > 0$. Pour tout $k \in \mathbb{N}^\star$, soit $T_k^{(s)}$ le k-ième temps de détection compté à partir du temps s et $\Delta_k^{(s)} = T_k^{(s)} - T_{k-1}^{(s)}$ (avec la convention $T_0^{(s)} = s$), c'est-à-dire,
$$T_k^{(s)} = T_{N_s + k} \qquad \text{et} \qquad \Delta_k^{(s)} = \begin{cases} T_{N_s+1} - s & \text{si } k = 1 \\ \Delta_{N_s+k} & \text{si } k \geq 2. \end{cases}$$

(a) Calculer $\mathbb{P}[\Delta_1^{(s)} \geq \delta | N_s = n]$ et $\mathbb{P}[\Delta_1^{(s)} \geq \delta]$ pour tout $\delta > 0$ et $n \in \mathbb{N}$. Conclusion ?

(b) Montrer que $(\Delta_k^{(s)})_{k \in \mathbb{N}^\star}$ est une suite de variables aléatoires indépendantes identiquement distribuées, de même loi exponentielle que celle des Δ_k.
A-t-on indépendance de $(\Delta_k^{(s)})_{k \geq 1}$ et de N_s ?

(c) Soit $t \geq s > 0$. Donner la loi de $N_{]s,t]} = N_t - N_s$ et montrer que $N_{]s,t]}$ et N_s sont indépendantes.