Cirrigé du Final - SY01/A16, durée 2h, 4 exercices, 2 pages

Indication : Un rappel est donné en fin de sujet.

Exercice 1. (Dans cet exercice les questions sont indépendantes)

1. Soient X et Y deux variables aléatoires indépendantes et de même loi Uniforme sur $[-1, 1]$. On définit la variable aléatoire $Z = (X+Y)/2$, déterminer sa fonction de répartition, puis en déduire sa densité.

   Corrigé :

$$F_Z(t) = \begin{cases} 0, & \text{si } t < -1, \\ * & \text{si } t \in [-1,\ 1], \\ 1 & \text{si } t > 1. \end{cases}$$

Ainsi pour $t \in [-1,\ 1]$ on a

$$F_Z(t) = P(X+Y \leq 2t) \Leftrightarrow F_Z(t) = F_{X+Y}(2t). \tag{1}$$

Or,

$$F_{X+Y}(t) = \int_R F_Y(t-x) f_X(x)dx = \int_R \frac{t-x+1}{2} 1_{\{-1 \leq t-x \leq 1\}} \frac{1}{2} 1_{\{-1 \leq x \leq 1\}} dx$$

$$= \frac{1}{4} \int_R (t-x+1) 1_{\{-1+t \leq x \leq 1+t\}} 1_{\{-1 \leq x \leq 1\}} dx = \begin{cases} 0, & \text{si } t < -2, \\ \frac{1}{4}\int_{-1}^{t+1}(t-x+1)dx = \frac{t}{2} + \frac{t^2}{8} + \frac{1}{2}, & \text{si } t \in [-2,\ 0], \\ \frac{1}{2} + \frac{1}{4}\int_{t-1}^{1}(t-x+1)dx = \frac{t}{2} - \frac{t^2}{8} + \frac{1}{2}, & \text{si } t \in]0,\ 2], \\ 1, & \text{si } t > 2. \end{cases}$$

Il s'ensuit que

$$F_{X+Y}(2t) = \begin{cases} 0, & \text{si } t < -1, \\ t + \frac{t^2}{2} + \frac{1}{2}, & \text{si } t \in [-1,\ 0], \\ t - \frac{t^2}{2} + \frac{1}{2}, & \text{si } t \in]0,\ 1], \\ 1, & \text{si } t > 1. \end{cases}$$

Il suffit de dériver pour trouver la densité.
On aurait pu aussi faire : de (1), on en déduit que pour $t \in [-1, 1]$, on a $f_Z(t) = 2f_{X+Y}(2t)$, soit $f_Z(t) = 2(f_X * f_Y)(2t)$. Ainsi

$$f_z(x) = \begin{cases} 2 \times \frac{1}{4}(2 - |2x|), & \text{si } |2x| \leq 2, \\ 0 & \text{sinon.} \end{cases}$$

Soit finalement

$$f_z(x) = \begin{cases} (1 - |x|), & \text{si } |x| \leq 1, \\ 0 & \text{sinon.} \end{cases}$$

2. Soit $a > 0$ et $\lambda > 0$, on considère la loi $\gamma(a, \lambda)$ de densité

$$f(x) = \frac{\lambda^a}{\Gamma(a)} e^{-\lambda x} x^{a-1} 1_{\{x \in \mathbb{R}^+\}}.$$

   Calculer l'espérance de la variable aléatoire X de loi $\gamma(a, \lambda)$.

   Corrigé :

$$E[X] = \int_{\mathbb{R}^+} x \frac{\lambda^a}{\Gamma(a)} e^{-\lambda x} x^{a-1} dx = \int_{\mathbb{R}^+} \frac{\lambda^a u^a}{\lambda \Gamma(a)} e^{-u} du = \frac{\Gamma(a+1)}{\lambda \Gamma(a)} = \frac{a\Gamma(a)}{\lambda\Gamma(a)} = \frac{a}{\lambda},$$

la deuxième égalité est obtenue par le changement de variable $u = \lambda x$.

3. Soit
$$f(x,y)=\left\{\begin{array}{ll} C, & \text{si } \max(|x|,|y|)\leq 2,\\ 0 & \text{sinon.}\end{array}\right.$$
Déterminer C, pour que $f(x,y)$ soit la densité de probabilité d'un couple de variables aléatoires.
Corrigé : Nécessairement $C>0$ et
$$f(x,y)=C1_{\{\max(|x|,|y|)\leq 2\}}=C1_{[0,2]^2}(|x|,|y|)=C1_{[-2,2]^2}(x,y),$$
Ainsi
$$\int_{\mathbb{R}}\int_{\mathbb{R}} f(x,y)dxdy=C\int_{-2}^{2}\int_{-2}^{2}dxdy=16C,$$
d'où $C=1/16$.
4. Soient X et Y deux variables aléatoires indépendantes et de même loi $\mathcal{P}(1/2)$. Soit $S=X+Y$. Pour $n\in\mathbb{N}$, déterminer la loi conditionnelle de X sachant $\{S=n\}$, ainsi que l'espérance conditionnelle de X sachant $\{S=n\}$ (notée $E[X|S=n]$).
Corrigé : On sait d'après le cours que la loi de S est la loi $\mathcal{P}(1/2+1/2)$. Soit $k\in\{0,1,\ldots,n\}$, on a
$$P(X=k|X+Y=n)=\frac{P(X=k,Y=n-k)}{P(X+Y=n)}=\frac{P(X=k)P(Y=n-k)}{P(X+Y=n)}$$
$$=\frac{\frac{e^{-1/2}(\frac{1}{2})^k}{k!}\frac{e^{-1/2}(\frac{1}{2})^{n-k}}{(n-k)!}}{e^{-1}\frac{1^n}{n!}}=C_n^k\left(\frac{1}{2}\right)^k\left(\frac{1}{2}\right)^{n-k}.$$
La loi conditionnelle de X sachant $\{S=n\}$ est la loi $B\left(n,\frac{1}{2}\right)$, ainsi $E[X|S=n]=\frac{n}{2}$.
5. Soient X et Y deux variables aléatoires indépendantes et de même loi exponentielle de paramètre 1, notée $\epsilon(1)$. Calculer l'espérance conditionnelle $E[\min\{X,Y\}|X]$.
Corrigé : On doit calculer d'abord pour $x>0$, $E[\min\{x,Y\}|X=x]$.
$$E[\min\{x,Y\}|X=x]=E[\min\{x,Y\}]=\int_0^{+\infty}\min\{x,y\}e^{-y}dy=\int_0^x ye^{-y}dy+\int_x^{+\infty}xe^{-y}dy$$
on fait une intégration par partie de la première intégrale d'où
$$E[\min\{x,Y\}|X=x]=-xe^{-x}+\int_0^x e^{-y}dy+xe^{-x}=-xe^{-x}+1-e^{-x}+xe^{-x}=1-e^{-x}.$$
Ainsi, $E[\min\{X,Y\}|X]=1-e^{-X}$.

Exercice 2. A l'approche des élections, un institut de sondage contacte successivement des individus. Le modèle est le suivant : les appels sont indépendants et chaque individu répond qu'il va voter pour le candidat A avec probabilité p_A (et pour le candidat B avec probabilité $p_B=1-p_A$). Le but est d'estimer le paramètre p_A du modèle.

1. Soit $N_A(n)$ le nombre de réponses en faveur du candidat A collectées en n appels. Que peut-on dire de la convergence $p.s.$ de la suite $\frac{N_A(n)}{n}$? (Une réponse non justifiée clairement ne sera pas recevable).
Corrigé : Pour $k\in\{1,\ldots,n\}$ on pose
$$X_k=\left\{\begin{array}{ll} 1, & \text{si le k-ième individu vote pour A,}\\ 0 & \text{sinon.}\end{array}\right.$$
Par l'énoncé les v.a. X_k sont indépendantes et de même loi de Bernoulli de paramètre p_A. Ainsi $N_A(n)=\sum_{i=1}^n X_i$ et donc d'après la loi des grands nombres
$$\frac{N_A(n)}{n}=\frac{\sum_{i=1}^n X_i}{n}\xrightarrow{p.s.}p_A.$$

2. L'application de l'inégalité de Bienaymé-Tchebichev à la variable $\frac{N_A(n)}{n}$ permet-elle de retrouver une version plus faible du résultat précédent ?
Corrigé : On a $E\left[\frac{N_A(n)}{n}\right] = p_A$ et $V\left[\frac{N_A(n)}{n}\right] = \frac{Var(X_1)}{n} = \frac{p_A(1-p_A)}{n}$, l'inégalité de Bienaymé-Tchebichev appliquée à la variable aléatoire $\frac{N_A(n)}{n}$ donne, pour tout $\epsilon > 0$

$$P\left(|\frac{N_A(n)}{n} - p_A| \geq \epsilon\right) \leq \frac{p_A(1-p_A)}{n\epsilon^2} \longrightarrow 0, \text{ lorsque } n \to \infty,$$

donc on a la convergence en probabilité de $\frac{N_A(n)}{n}$ vers p_A.

Exercice 3. Soient X et Y deux variables aléatoires indépendantes et de même loi $N(0,1)$.

1. On pose $U = X$ et $V = X^2 + Y^2$. Déterminer la densité du couple (U, V).
Corrigé : Les v.a. étant indépendantes la densité de (X, Y) est donnée par

$$f_{(X,Y)}(x,y) = \frac{1}{2\pi}e^{-1/2(x^2+y^2)}, (x,y) \in \mathbb{R}^2.$$

Remarque : Pour avoir la bijection des domaines on doit choisir $y \geq 0$ et $x \in \mathbb{R}$ ou $y \leq 0$ et $x \in \mathbb{R}$.
Le jacobien de la transformation étant : $|\frac{1}{2\sqrt{v-u^2}}|$, nous avons

$$g_{(U,V)}(u,v) = \frac{1}{2\pi}e^{-v/2}\frac{1}{2\sqrt{v-u^2}}1_{\{v>u^2\}}.$$

2. En déduire la densité de V. Retrouve-t-on une loi usuelle ?
Corrigé : Pour $v > 0$, on déduit que la densité de V est donnée par

$$g_V(v) = \frac{1}{4\pi}e^{-v/2}\int_{-\sqrt{v}}^{\sqrt{v}}\frac{du}{\sqrt{v-u^2}}$$

or, nous avons que

$$\frac{1}{4}\int_0^\infty e^{-v/2}dv = \frac{1}{2}$$

(en multipliant à gauche et à droite par 2) on a donc que $g_V(v) = \frac{1}{2}e^{-v/2}1_{\{v\geq 0\}}$, soit V suit la loi $\varepsilon(1/2)$.

3. En déduire la densité conditionnelle de U sachant V.
Corrigé : Pour $v > 0$

$$g_{U|V}(u|v) = \frac{1_{\{|u|<\sqrt{v}\}}}{\pi\sqrt{v-u^2}},$$

et $g_{U|V}(u|v) = 0$ si $v \leq 0$.

4. Calculer l'espérance conditionnelle de $|U|$ (c.a.d. valeur absolue de la variable aléatoire U) sachant V, notée $E[|U||V]$.
Corrigé :

$$E[|U||V = v] = \int_R |u|g_{U|V}(u|v)du = \frac{2}{\pi}\int_0^{\sqrt{v}}\frac{u}{\sqrt{v-u^2}}du = \frac{2\sqrt{v}}{\pi}.$$

Ainsi, on en déduit que $E[|U||V] = \frac{2\sqrt{V}}{\pi}$.

5. On pose $X = R\cos\Theta$ et $Y = R\sin\Theta$, $R > 0$ et $\Theta \in]0, 2\pi]$. Déterminer la densité du couple (R, Θ). Les v.a. R et Θ sont-elles indépendantes ?
Corrigé :

$$g_{(R,\Theta)}(r,\theta) = re^{-r^2/2}1_{]0,\infty]}(r)\frac{1_{]0,2\pi]}(\theta)}{2\pi}.$$

Les v.a. R et Θ sont donc indépendantes de densités respectives $g_R(r) = re^{-r^2/2}1_{]0,\infty]}(r)$ et $g_\Theta(\theta) = \frac{1_{]0,2\pi]}(\theta)}{2\pi}$.

6. A l'aide de la question précédente, retrouver l'espérance conditionnelle de $|U|$ sachant V, notée $E[|U||V]$.
Corrigé : Pour $v > 0$ on a

$$E[|U||V = v] = \sqrt{v}E[|\cos\Theta|] = \frac{4}{2\pi}\sqrt{v}\int_0^{\pi/2}\cos\theta d\theta = \frac{2\sqrt{v}}{\pi}.$$

On en déduit $E[|U||V] = \frac{2\sqrt{V}}{\pi}$.

Exercice 4. Soient X et Y deux variables aléatoires indépendantes, X suit la loi $N(\mu_1, \sigma_1^2)$, et Y suit la loi $N(\mu_2, \sigma_2^2)$.

1. En utilisant la fonction génératrice des moments, déduire la loi de la variable aléatoire $X - Y$.
Corrigé $\phi_{X-Y}(t) = E[e^{t(X-Y)}] = \phi_X(t)\phi_Y(-t)$ par indépendance. D'où $\phi_{X-Y}(t) = e^{\sigma_1^2 t^2/2+\mu_1 t}e^{\sigma_2^2 t^2/2-t\mu_2}$, d'où $\phi_{X-Y}(t) = e^{(\sigma_1^2+\sigma_1^2)t^2/2+(\mu_1-\mu_2)}$ qui est la fonction génératrice de moments d'une variable aléatoire de loi $N(\mu_1 - \mu_2, \sigma_1^2 + \sigma_2^2)$.
2. Vérifiez rapidement la valeur des paramètres de la loi de $X - Y$, obtenus à la question précédente.
Corrigé : $E[X - Y] = \mu_1 - \mu_2$ et $Var(X - Y) = Var(X) + Var(Y) = \sigma_1^2 + \sigma_2^2$.
3. Application : Des ingénieurs civils pensent que W, le poids (en tonnes) qu'une travée de pont peut supporter sans subir de dommage au niveau de sa structure, suit la loi normale, de moyenne 400 et d'écart type 40. Supposons que le poids (également en tonnes) d'un véhicule est une variable aléatoire de moyenne 3 et d'écart type 0.3.
Notons X_i la masse du i-ème véhicule, et S_n la somme des masses de n parmi ceux-ci.
   a. Exprimer la condition de rupture pour n véhicules sur la travée.
   Corrigé : Il y a rupture si $S_n > W$ ce qui est équivalent à $S_n - W > 0$.
   b. Pour n grand, quelle loi peut-on considérer pour la variable aléatoire S_n ? justifier clairement votre réponse.
   Corrigé : La variable aléatoire S_n est la somme des masses de n véhicules donc $S_n = X_1 + \cdots + X_n$, les variables X_i sont considérés indépendantes, admettant une moyenne et une variance finies. Pour n grand la loi de S_n (par le TCL) peut être considérée comme étant la loi normale $N(3n, (0,3)^2 n)$.
   c. Combien de véhicules devraient se trouver sur la travée pour que la probabilité de rupture soit supérieure à 0.1 ? On ne demande que d'exprimer l'inégalité qu'il faudrait résoudre pour trouver explicitement la valeur de n minimale.
   Indication : la lecture de la table donne $\phi_U(1.28) = 0.9$, pour U variable aléatoire de loi Normale centrée et réduite.
   Corrigé : On cherche donc n (minimal) tel que $P(S_n - W > 0) \geq 0.1$. Or, d'après les questions précédentes ceci revient au calcul de $P(Z > 0) \geq 0.1$ où Z est une variable aléatoire normale $N(3n - 400, (0.3)^2 n + 40^2)$, soit
   $P\left(\frac{Z-(3n-400)}{\sqrt{(0.3)^2 n+40^2}} > \frac{400-3n}{\sqrt{(0.3)^2 n+40^2}}\right) \geq 0.1$. Or, $\phi_U(1.28) = 0.9$, nous cherchons n tel que $\frac{400-3n}{\sqrt{(0.3)^2 n+40^2}} \leq 1.28$.

Rappels :

a. $\frac{d}{dx}(\arcsin(x)) = \frac{1}{\sqrt{1-x^2}}$, pour $x \in]-1, 1[$.

b. $\Gamma(a) = \int_0^{+\infty} x^{a-1}e^{-x}dx,\ a \in \mathbb{R}_+^*$.