MVA - Correction des exercices

Exercice 1 Soit (X,Y) un couple de variables aléatoires réelles. Notons $f_{(X,Y)}$ la densité de la loi jointe de (X,Y). La densité de la loi de Y s'écrit alors $f_Y(y) = \int_{x\in\mathbb{R}} f_{(X,Y)}(x,y)dxdy$ et la densité de la loi de X conditionnelle à Y s'écrit, pour tout y tel que $f_Y(y) \neq 0$, $f_{X/Y}(x,y) = \frac{f_{(X,Y)}(x,y)}{f_Y(y)}$. L'espérance de X conditionnelle à Y est la variable aléatoire $\mathbb{E}(X/Y) = \int x f_{X/Y}(x,Y)dx$.

1. Si la loi jointe de (X,Y) est la loi uniforme sur le triangle $(0,0),(1,0),(0,1)$, alors $f_{(X,Y)} = 2\mathbb{I}_{y+x\leq 1, x\geq 0, y\geq 0}$, donc $f_Y(y) = \int_0^{1-y} 2dx\mathbb{I}_{0\leq y\leq 1} = 2(1-y)\mathbb{I}_{0\leq y\leq 1}$ et $f_{X/Y}(y) = \frac{1}{1-y}\mathbb{I}_{y+x\leq 1, x\geq 0, y\geq 0}$. L'espérance de X conditionnelle à Y est la variable aléatoire $\mathbb{E}(X/Y) = \int_0^{1-Y} \frac{x}{1-Y}\mathbb{I}_{0\leq Y\leq 1}dx = \frac{1-Y}{2}$.
2. Si la loi jointe de (X,Y) est la loi uniforme sur le carré $(0,0),(1,0),(1,1),(0,1)$ alors $f_{(X,Y)} = \mathbb{I}_{0\leq x\leq 1, 0\leq y\leq 1} = \mathbb{I}_{0\leq x\leq 1}\mathbb{I}_{0\leq y\leq 1}$, donc X et Y sont indépendantes. On en déduit que $f_{X/Y}(x,y) = f_X(x)$ et que $\mathbb{E}(X/Y)$ est la v.a. constante égale à $\mathbb{E}(X) = \frac{1}{2}$.

Exercice 2 Un composant électronique est installé à l'instant $t=0$ et tombe en panne à un instant aléatoire $T\geq 0$. On suppose que la loi de la variable aléatoire T a une densité f par rapport à la mesure de Lebesgue et que pour tout $t\geq 0$, $\mathbb{P}(T>t)>0$. On pose $F(t) = 1-\bar{F}(t) = \int_{s=0}^t f(s)ds$ pour tout $t\geq 0$. Si à un instant $t>0$, la panne ne s'est pas encore produite, le risque de panne instantanée se mesure par le **taux de panne** :

$$\lambda(t) = \lim_{u\to 0+} \frac{1}{u}\mathbb{P}(t<T\leq t+u \mid T<t).$$

On dit qu'il y a *usure* de la machine lorsque λ croît et *rodage* lorsque λ décroît.

1. Si $\mathbb{P}(T>s+t \mid T>t) = \mathbb{P}(T>s)$, pour tout $(t,s)\in\mathbb{R}_+^2$, alors la fonction $\bar{F}(t) = \mathbb{P}(T>t)$ vérifie $\bar{F}(t+s) = \bar{F}(t)\bar{F}(s)$ pour tout $(t,s)\in\mathbb{R}_+^2$ car $\mathbb{P}(T>s+t \mid T>t) = \mathbb{P}(T>s+t \text{ et } T>t)/\mathbb{P}(T>t) = \mathbb{P}(T>s+t)/\mathbb{P}(T>t)$. Il existe donc une constante $0<\alpha<1$ telle que $\bar{F}(t) = \alpha^t$. Dans ce cas, $\lambda(t) = \lim_{u\to 0+}\frac{1}{u}\frac{\alpha^t-\alpha^{t+u}}{\alpha^t} = -\ln(\alpha)$.
2. On se place de nouveau dans le cas général.

   $$\lambda(t) = \lim_{u\to 0+}\frac{1}{u}\frac{\bar{F}(t)-\bar{F}(t+u)}{\bar{F}(t)} = \frac{f(t)}{\bar{F}(t)}.$$

   Le **taux de panne intégré** est alors $\Lambda(t) = \int_{s=0}^t \frac{f(s)}{\bar{F}(s)}ds = \ln(\bar{F}(0)) - \ln(\bar{F}(t)) = -\ln(\bar{F}(t))$. Donc $\bar{F}(t) = e^{-\Lambda(t)}$.
3. On suppose ici que T suit une loi Gamma de paramètres $a>0$ et $b>0$, sa densité s'écrit donc : $\forall t>0$,

   $$f(t) = \frac{b^a}{\Gamma(a)}e^{-bt}t^{a-1},$$

   avec $\Gamma(a) = \int_{t=0}^{\infty} e^{-t}t^{a-1}dt$. Par suite, on peut écrire

   $$\lambda(t) = \frac{e^{-bt}t^{a-1}}{\int_t^{+\infty} e^{-bs}s^{a-1}ds} = \frac{1}{\int_t^{+\infty} e^{-b(s-t)}\left(\frac{s}{t}\right)^{a-1}ds} = \frac{1}{\int_0^{+\infty} e^{-bu}\left(1+\frac{u}{t}\right)^{a-1}du}.$$

Cette fonction de t est croissante (usure) si $a > 1$ et décroissante (rodage) si $a < 1$, elle est constante si $a = 1$. Le valeurs de a et b étant fixées, pour t suffisamment grand, on a $|e^{-bu}\left(1+\frac{u}{t}\right)^{a-1}| \leq e^{-\frac{b}{2}u}$ pour tout u positif. On peut donc passer à la limite sous le signe somme (convergence monotone) et lorsque $t \to \infty$, $\lambda(t) \to \frac{1}{\int_0^{+\infty} e^{-bu}du} = b$.

4. Les deux composants étant indépendants, on peut écrire

$$\mathbb{P}[T \leq t] = \mathbb{P}[T_1 \leq t]\mathbb{P}[T_2 \leq t] = \int_0^t \lambda_1 e^{-\lambda_1 s} ds \int_0^t \lambda_2 e^{-\lambda_2 s} ds = (1-e^{-\lambda_1 t})(1-e^{-\lambda_2 t}).$$

La densité de la loi de T est donc $\lambda_1 e^{-\lambda_1 t} + \lambda_2 e^{-\lambda_2 t} - (\lambda_1+\lambda_2)e^{-(\lambda_1+\lambda_2)t}$. Le taux de panne du système est donc $\lambda(t) = \frac{\lambda_1 e^{-\lambda_1 t} + \lambda_2 e^{-\lambda_2 t} - (\lambda_1+\lambda_2)e^{-(\lambda_1+\lambda_2)t}}{e^{-\lambda_1 t} + e^{-\lambda_2 t} - e^{-(\lambda_1+\lambda_2)t}}$. Si $\lambda_1 = \lambda_2$, ce taux devient $\lambda(t) = 2\lambda_1(1 - \frac{1}{2-e^{-\lambda_1 t}})$, qui est une fonction croissante de t, donc il y a usure du système.

Exercice 3 Soient X et Y deux variables aléatoires sur un même espace de probabilité. On suppose X à valeurs dans $\mathbb{N}$ et Y à valeurs dans $\mathbb{R}$ de loi exponentielle de paramètre 1, c'est-à-dire que sa densité est $f_Y(y) = e^{-y}\mathbb{I}_{y \geq 0}$. On suppose aussi que la loi de X conditionnelle à Y est une loi de Poisson de paramètre Y, c'est-à-dire que $\mathbb{P}[X = k/Y] = \frac{Y^k}{k!}e^{-Y}\mathbb{I}_{Y \geq 0}$. La loi de (X, Y) a donc pour "densité"

$$f_{(X,Y)}(k, y) = \frac{y^k}{k!}e^{-2y}\mathbb{I}_{y \geq 0}.$$

Quant à X, pour tout $k \in \mathbb{N}$, $\mathbb{P}[X = k] = \int \frac{y^k}{k!}e^{-2y}\mathbb{I}_{y \geq 0}dy = \frac{1}{2^{k+1}}$. La densité de la loi de Y conditionnelle à l'évènement $X = n$ est donc $f_{(Y/X=n)}(y) = 2^{n+1}\frac{y^n}{n!}e^{-2y}\mathbb{I}_{y \geq 0}$.

Exercice 4 (Vecteur gaussien) Soient $X = (X_1, X_2)$ un vecteur gaussien de dimension 2 dont la densité est donnée par : $\forall (x_1, x_2) \in \mathbb{R}^2$,

$$g(x_1, x_2) = \frac{1}{2\pi\sqrt{1-\rho^2}\sigma_1\sigma_2} \exp\left[-\frac{1}{2(1-\rho^2)}\left(\frac{x_1^2}{\sigma_1^2} - 2\rho\frac{x_1 x_2}{\sigma_1\sigma_2} + \frac{x_2^2}{\sigma_2^2}\right)\right],$$

où $\rho \in]0, 1[$, $(\sigma_1, \sigma_2) \in \mathbb{R}_+^{*2}$. La densité de la loi de X_1 est donnée par

$$\begin{aligned}
f_{X_1}(x_1) &= \int_{x_2 \in \mathbb{R}} g(x_1, x_2)dx_2 = \frac{1}{2\pi\sqrt{1-\rho^2}\sigma_1\sigma_2} \int_{x_2 \in \mathbb{R}} \exp\left[-\frac{1}{2(1-\rho^2)}\left(\frac{x_1^2}{\sigma_1^2} - 2\rho\frac{x_1 x_2}{\sigma_1\sigma_2} + \frac{x_2^2}{\sigma_2^2}\right)\right] dx_2 \\
&= \frac{1}{2\pi\sqrt{1-\rho^2}\sigma_1\sigma_2} \exp\left[-\frac{1}{2(1-\rho^2)}\frac{x_1^2}{\sigma_1^2}(1-\rho^2)\right] \int_{x_2 \in \mathbb{R}} \exp\left[-\frac{1}{2(1-\rho^2)}\left(\frac{x_2}{\sigma_2} - \rho\frac{x_1}{\sigma_1}\right)^2\right] dx_2 \\
&= \frac{1}{2\pi\sqrt{1-\rho^2}\sigma_1\sigma_2} \exp\left[-\frac{1}{2}\frac{x_1^2}{\sigma_1^2}\right] \sqrt{2\pi(1-\rho^2)}\sigma_2 = \frac{1}{\sqrt{2\pi}\sigma_1} \exp\left[-\frac{x_1^2}{2\sigma_1^2}\right],
\end{aligned}$$

X_1 suit donc une loi gaussienne centrée de variance σ_1^2. Par symétrie de g, X_2 suit une loi gaussienne centrée de variance σ_2^2. On en déduit la densité de la loi de X_1 conditionnelle à X_2 :

$$f_{X_1/X_2}(x_1, x_2) = \frac{g(x_1, x_2)}{f_{X_2(x_2)}} = \frac{1}{\sqrt{2\pi(1-\rho^2)}\sigma_1} \exp\left[-\frac{1}{2(1-\rho^2)}\left(\frac{x_1^2}{\sigma_1^2} - 2\rho\frac{x_1 x_2}{\sigma_1\sigma_2} + \frac{x_2^2}{\sigma_2^2}\right)\right] \exp\left[\frac{x_2^2}{2\sigma_2^2}\right].$$

Exercice 5 (Statistiques d'ordre) Soit $X = (X_1, \ldots, X_n)$ un vecteur aléatoire de densité $f(x)$, $x \in \mathbb{R}^n$. Soit σ la permutation de $\{1, \ldots, n\}$ telle que $X_{\sigma(1)} < \ldots < X_{\sigma(n)}$ et soit $R = (R_1, \ldots, R_n)$ le *vecteur rang* des observations : $\forall i \in \{1, \ldots, n\}$, $R_i = \sum_{j=1}^n \mathbb{I}\{X_i \geq X_j\}$. Remarquons que $R_i = \sigma^{-1}(i)$. Autrement dit, $(X_{\sigma(1)}, \ldots X_{\sigma(n)}) = (x_1, \ldots x_n)$ si et seulement si $(X_1, \ldots X_n) = (x_{R_1}, \ldots x_{R_n})$.

La densité de $(X_{\sigma(1)}, \ldots, X_{\sigma(n)})$ est $(x_1, \ldots x_n) \mapsto \sum_{\sigma' \in \Sigma(n)} f(x_{\sigma'(1)}, \ldots x_{\sigma'(n)}) \mathbb{I}_{x_1 < \cdots < x_n}$, où l'on note $\Sigma(n)$ l'ensemble des permutations de $\{1, \ldots n\}$. Par suite,

$$f_{R/(X_{\sigma(1)}, \ldots, X_{\sigma(n)})}(r_1, \ldots r_n, x_1, \ldots x_n) = \frac{f(x_{r_1}, \ldots x_{r_n})}{\sum_{\sigma' \in \Sigma(n)} f(x_{\sigma'(1)}, \ldots x_{\sigma'(n)})} \mathbb{I}_{x_1 < \cdots < x_n}.$$

Si X est un n-échantillon iid d'une loi de densité $f(u)$ sur $\mathbb{R}$, alors

$$f_{R/(X_{\sigma(1)}, \ldots, X_{\sigma(n)})}(r_1, \ldots r_n, x_1, \ldots x_n) = \frac{f(x_{r_1}) \ldots f(x_{r_n})}{\sum_{\sigma' \in \Sigma(n)} f(x_{\sigma'(1)}) \ldots f(x_{\sigma'(n)})} \mathbb{I}_{x_1 < \cdots < x_n} = \frac{1}{n!} \mathbb{I}_{x_1 < \cdots < x_n}.$$

Exercice 6 On rappelle que si Z est une v.a. sur $(\Omega, \mathcal{A}, \mathbb{P})$, alors $\mathbb{E}[Z/\mathcal{B}]$ est une v.a. $\mathcal{B}$-mesurable telle que $\mathbb{E}[(Z - \mathbb{E}[Z/\mathcal{B}])X] = 0$ pour toute v.a. X $\mathcal{B}$-mesurable de carré intégrable. On rappelle également que $\mathbb{E}[\mathbb{E}[Z/\mathcal{B}]] = \mathbb{E}[Z]$ (se déduit de ce qui précède en prenant $X = \mathbb{I}_\Omega$).

Par suite, si Y est une variable aléatoire réelle de carré intégrable sur $(\Omega, \mathcal{A}, \mathbb{P})$, alors,

$$\sigma^2(\mathbb{E}[Y \mid \mathcal{B}]) = \mathbb{E}[(\mathbb{E}[Y \mid \mathcal{B}])^2] - (\mathbb{E}[\mathbb{E}[Y \mid \mathcal{B}]])^2 = \mathbb{E}[(\mathbb{E}[Y \mid \mathcal{B}])^2] - (\mathbb{E}[Y])^2.$$

Par ailleurs, par définition de $\sigma^2(Y|\mathcal{B})$,

$$\mathbb{E}[\sigma^2(Y|\mathcal{B})] = \mathbb{E}[\mathbb{E}[Y^2 \mid \mathcal{B}]] - \mathbb{E}[(\mathbb{E}[Y \mid \mathcal{B}])^2)] = \mathbb{E}[Y^2] - \mathbb{E}[(\mathbb{E}[Y \mid \mathcal{B}])^2)],$$

donc

$$\sigma^2(\mathbb{E}[Y \mid \mathcal{B}]) + \mathbb{E}[\sigma^2(Y|\mathcal{B})] = \mathbb{E}[Y^2] - (\mathbb{E}[Y])^2 = \sigma^2(Y).$$

Si Y est indépendante de la tribu $\mathcal{B}$, alors $\mathbb{E}[Y \mid \mathcal{B}]$ est la v.a constante $\mathbb{E}[Y]$, donc le premier terme de la somme est nul et $\sigma^2(Y|\mathcal{B}) = \sigma^2(Y)$.