Troisième Année Licence M.I.A.S.H.S. 2023 – 2024

Statistique 2

Contrôle Continu 1, Mars 2024

Examen de 1h30. Tout document ou calculatrice est interdit.

Exercice 1 (Sur 9 points)

On s'intéresse à la durée de vie T d'une ampoule électrique. Après observation du comportement d'un grand nombre d'ampoules, on modélise T (en milliers d'heures) de la façon suivante: T est une variable aléatoire définie sur $(\Omega, \mathcal{A}, \mathbb{P})$, dont la fonction de répartition F_T est telle que pour $x \in \mathbb{R}$:

$$F_T(x) = \frac{1}{10} \mathbb{I}_{[0,+\infty[}(x) + \frac{9}{10} \left(1 - \frac{1}{(1+x^a)^b}\right) \mathbb{I}_{]0,+\infty[}(x),$$

où $a > 0$ et $b \in \mathbf{R}$.

1. Donner une condition sur (a, b) pour que F_T soit bien une fonction de répartition (justifier rigoureusement qu'elle en est une!) **(1.5pts)**. Quelle est la probabilité qu'une ampoule vive exactement 1000 heures **(0.5pts)**? Strictement moins de 1000 heures **(0.5pts)**? Ait une durée de vie nulle **(0.5pts)**?
2. Soit $\mathbb{P}_T$ la mesure de probabilité associée à F_T. Décomposer, en justifiant, $\mathbb{P}_T$ sous la forme d'une somme pondérée d'une masse de Dirac et d'une mesure de probabilité μ absolument continue par rapport à la mesure de Lebesgue λ sur $\mathbb{R}$ dont on donnera la densité **(2pts)**. Pour $I = [\alpha, \beta]$ intervalle de $\mathbf{R}_+$, donner l'expression de $\mu(I)$ **(0.5pts)**.
3. Donner l'expression générale de $\mathbb{E}[T]$ (on ne cherchera pas à simplifier l'intégrale) **(0.5pts)**. Quelle relation doivent vérifier a et b pour que $\mathbb{E}[T] < \infty$ **(1pt)**?
4. Dans le cas où $a = 2$ et $b = 1$, calculer $\mathbb{E}[T]$ **(1pt)** et $\mathrm{var}(T)$ **(1pt)**.

Proof.

1. Pour que F_T soit bien une fonction de répartition il faut que F_T soit croissante, avec limite 0 en $-\infty$ et limite 1 en $+\infty$. Cela impose que $b > 0$. Par ailleurs, F_T est continue en tout point de $\mathbf{R}$ sauf en 0, mais elle a une limite à gauche qui est 0 et elle est continue à droite en 0 (avec valeur 1/10): c'est bien une fonction de répartition.
On a $\mathbb{P}(T = 1) = F_T(1) - \lim_{t \to 1^-} F_T(t)$. Or la fonction F_T est continue en 1. Donc $\mathbb{P}(T = 1) = 0$.
On a $\mathbb{P}(T < 1) = F_T(1) - \mathbb{P}(T = 1) = 1/10 + 9/10(1 - 1/2) = 11/20$.
Enfin, $\mathbb{P}(T = 0) = F_T(0) = 1/10$.

2. On veut que $\mathbb{P}_T\big(]-\infty, x]\big) = F_T(x)$. Si on écrit $\mathbb{P}_T(B) = \frac{1}{10}\,\delta_{\{0\}}(B) + \frac{9}{10} \int_B f_T(x)\, d\lambda(x)$ avec $\frac{9}{10} f_T(x) = F_T'(x)$ pour presque tout $x \in \mathbf{R}$ (en fait pour $x \in \mathbf{R}^*$), soit $f_T(x) = \frac{ab\,x^{a-1}}{(1+x^a)^{b+1}} \mathbb{I}_{x>0}$. On vérifie ainsi que pour $x < 0$, alors $\mathbb{P}_T\big(]-\infty, x]\big) = 0$, pour $x = 0$ alors $\mathbb{P}_T\big(]-\infty, x]\big) = 1/10$ et pour $x > 0$, $\mathbb{P}_T\big(]-\infty, x]\big) = 1/10 + 9/10 \int_0^x f_T(t)\, dt = 1/10 + 9/10\big(1 - (1+x^a)^b\big)$. D'où
$$\mathbb{P}_T(B) = \frac{1}{10}\,\delta_{\{0\}}(B) + \frac{9}{10} \int_B \frac{ab\,x^{a-1}}{(1+x^a)^{b+1}} \mathbb{I}_{x>0}\, d\lambda(x) = \frac{1}{10}\,\delta_{\{0\}}(B) + \frac{9}{10}\,\mu(B).$$
Pour $I = [\alpha, \beta]$ avec $0 \le \alpha \le \beta$, il est clair que comme μ est absolument continue par rapport à la mesure de Lebesgue sur $\mathbf{R}$, alors $\mu(I) = \frac{10}{9}\big(F_T(\beta) - F_T(\alpha)\big)$.

3. On a T fonction mesurable positive. On peut donc calculer $\mathbb{E}[T] = \dfrac{1}{10} \times 0 + \dfrac{9\,ab}{10} \displaystyle\int_0^\infty \frac{x^{1+(a-1)}}{(1+x^a)^{b+1}}\, dx = \dfrac{9\,ab}{10} \displaystyle\int_0^\infty \frac{x^a}{(1+x^a)^{b+1}}\, dx$.
Il s'agit d'une intégrale généralisée. La fonction est équivalente à $x^{a-(b+1)a}$ en $+\infty$. Pour que l'intégrale converge en $+\infty$ il faut que $a - (b+1)a < -1$, soit $ab > 1$.

4. Dans ce cas, on a alors $\mathbb{E}[T] = \dfrac{9}{5} \displaystyle\int_0^\infty \frac{x^2}{(1+x^2)^2}\, dx$. Par intégration par parties, on obtient:
$$\mathbb{E}[T] = \frac{9}{5}\left(\Big[-x\,\frac{1}{2(1+x^2)}\Big]_0^\infty + 2\int_0^\infty \frac{1}{1+x^2}\, dx\right) = \frac{9}{5}\,\pi.$$
Par ailleurs, $\mathbb{E}[T^2] = +\infty$ car cela revient à calculer $\int_0^\infty \frac{x^3}{(1+x^2)^2}\, dx = \infty$ (équivalent en $1/x$ en $+\infty$). Donc $\mathrm{var}(T) = +\infty$.

□

Exercice 2 (Sur 8 points)

Soit le vecteur aléatoire (X,Y) défini sur $(\Omega,\mathcal{A},\mathbb{P})$ et on suppose que la loi de (X,Y) est absolument continue par rapport à la mesure de Lebesgue sur $\mathbf{R}^2$ avec pour densité $f_{(X,Y)}$. On note $M=\max(X,Y)$.

1. Montrer que M est une variable aléatoire définie sur $(\Omega,\mathcal{A},\mathbb{P})$ **(1pt)**.
2. Dans le cas où X et Y sont indépendantes, rappeler d'abord les expressions des densités f_X et f_Y de X et de Y **(0.5pts)**. Démontrer que M est une v.a. de loi absolument continue par rapport à la mesure de Lebesgue sur $\mathbf{R}$ **(1pt)**. Préciser sa densité f_M en fonction de f_X et de f_Y **(1pt)**, puis en déduire que pour presque tout $m\in\mathbf{R}$,
$$\textbf{(1pt)}\qquad f_M(m)=\int_{-\infty}^{m}\left(f_{(X,Y)}(z,m)+f_{(X,Y)}(m,z)\right)dz. \tag{1}$$
3. On ne suppose plus que X et Y sont indépendantes. Montrer que $\mathbb{P}(M\le m)=\mathbb{P}(X\le Y\le m)+\mathbb{P}(Y\le X\le m)$ pour tout $m\in\mathbf{R}$ **(1pt)**. En déduire que M est une v.a. de loi absolument continue par rapport à la mesure de Lebesgue sur $\mathbf{R}$ et préciser sa densité en fonction de $f_{(X,Y)}$ **(2.5pts)**.

Proof.

1. X et Y sont des fonctions mesurables de $(\Omega,\mathcal{A})$ dans $(\mathbf{R},\mathcal{B}(\mathbf{R}))$. La fonction $(x,y)\in\mathbf{R}^2\mapsto\max(x,y)$ est une fonction continue, donc mesurable de $(\mathbf{R}^2,\mathcal{B}(\mathbf{R}^2))$ dans $(\mathbf{R},\mathcal{B}(\mathbf{R}))$. Par composition des fonctions mesurables, M est donc mesurable de $(\Omega,\mathcal{A})$ dans $(\mathbf{R},\mathcal{B}(\mathbf{R}))$: c'est bien une variable aléatoire définie sur $(\Omega,\mathcal{A},\mathbb{P})$.
2. Pour tout $(x,y)\in\mathbf{R}^2$, on sait que $f_X(x)=\int_{-\infty}^{\infty}f_{(X,Y)}(x,y)\,dy$ et $f_Y(y)=\int_{-\infty}^{\infty}f_{(X,Y)}(x,y)\,dx$.
Soit $m\in\mathbf{R}$. Alors la fonction de répartition F_M de M en m est définie par:
$$F_M(m)=\mathbb{P}\big(M\le m\big)=\mathbb{P}\big(X\le m\cap Y\le m\big)=\mathbb{P}\big(X\le m\big)\,\mathbb{P}\big(Y\le m\big),$$
car X et Y sont indépendantes. Cette fonction est donc continue car X et Y sont des v.a. "continues" donc leurs fonctions de répartition sont continues. Par ailleurs, leurs fonctions de répartition sont aussi dérivables presque partout sur $\mathbf{R}$, donc le produit de ces fonctions est aussi dérivables presque partout sur $\mathbf{R}$. On en déduit que M est une v.a. de loi absolument continue par rapport à la mesure de Lebesgue sur $\mathbf{R}$.
Grâce au calcul précédent, on en déduit que pour presque tout m:
$$f_M(m)=\frac{\partial}{\partial m}F_M(m)=f_X(m)\int_{-\infty}^{m}f_Y(y)\,dy+f_Y(m)\int_{-\infty}^{m}f_X(x)\,dx.$$
Il est donc également possible d'écrire cette densité f_M en fonction de $f_{(X,Y)}$ puisque $f_X(m)\,f_Y(y)=f_{(X,Y)}(m,y)$ et $f_Y(m)\,f_X(x)=f_{(X,Y)}(x,m)$:
$$f_M(m)=\frac{\partial}{\partial m}F_M(m)=\int_{-\infty}^{m}f_{(X,Y)}(m,y)\,dy+\int_{-\infty}^{m}f_{(X,Y)}(x,m)\,dx=\int_{-\infty}^{m}\Big(f_{(X,Y)}(m,z)+f_{(X,Y)}(z,m)\Big)\,dz.$$
3. On a $\mathbb{P}(M\le m)=\mathbb{P}(X<Y\le m)+\mathbb{P}(Y<X\le m)+\mathbb{P}(X=Y\le m)$ par la formule des probabilités totales. Mais l'ensemble $\{(x,y)\in[-\infty,m]^2, x=y\}$ est une demi-droite. Et comme la mesure de Lebesgue sur $\mathbf{R}^2$ d'une droite ou demi-droite vaut 0, et que le vecteur (X,Y) a une loi absolument continue par rapport à cette mesure de Lebesgue sur $\mathbf{R}^2$, alors $\mathbb{P}(X=Y\le m)=0$. D'où la formule demandée car pour la même raison $\mathbb{P}(X<Y\le m)=\mathbb{P}(X\le Y\le m)$.
Grâce à cette formule, on peut écrire que pour tout $m\in\mathbf{R}$,
$$\begin{aligned}F_M(m)&=\mathbb{P}(X\le Y\le m)+\mathbb{P}(Y\le X\le m)\\&=\int_{-\infty}^{m}\Big(\int_{-\infty}^{y}f_{(X,Y)}(x,y)\,dx\Big)\,dy+\int_{-\infty}^{m}\Big(\int_{-\infty}^{x}f_{(X,Y)}(x,y)\,dy\Big)\,dx\\&=\int_{-\infty}^{m}\Big(\int_{-\infty}^{u}\big(f_{(X,Y)}(z,u)+f_{(X,Y)}(u,z)\big)\,dz\Big)\,du.\end{aligned}$$
D'où le fait que $F_M(m)$ s'écrit bien sous la forme $\int_{-\infty}^{m}f_M(u)\,du$ avec f_M une fonction positive mesurable, donc M est bien une v.a. "continue" et sa densité est également celle donnée par (1).

$\square$

Exercice 3 (Sur 6 points)

Soit $(X_i)_{i\in\mathbf{N}}$ une suite de variables aléatoires indépendantes toutes définies sur le même espace de probabilité $(\Omega,\mathcal{A},\mathbb{P})$ et toutes de même loi de Bernoulli $\mathcal{B}(p)$ avec $0<p\le 1$. Soit N une variable aléatoire définie également sur $(\Omega,\mathcal{A},\mathbb{P})$, indépendante de $(X_i)_{i\in\mathbf{N}}$, et suivant une loi de Poisson de paramètre $\lambda>0$. Soit
$$Z=\sum_{i=1}^{N}X_i\quad\text{avec la convention}\quad\sum_{i=1}^{0}X_i=0.$$

1. Montrer que Z est une variable aléatoire définie sur $(\Omega,\mathcal{A},\mathbb{P})$ **(1.5pts)** et déterminer son ensemble de valeurs possibles **(0.5pts)**.

2. Pour $(n,k) \in \mathbf{N}^2$, déterminer $\mathbb{P}(Z = k \cap N = n)$ en traitant à part le cas $n = 0$ **(2pts)**. En déduire que Z suit une loi de Poisson de paramètre $p\,\lambda$ **(2pts)**.

Proof. 1. On peut écrire que $Z = \sum_{n=0}^{\infty} \mathbb{I}_{N=n} \sum_{i=1}^{n} X_i$. Pour tout $n \in \mathbf{N}$, $\omega \in \Omega \mapsto \mathbb{I}_{N(\omega)=n} = \mathbb{I}_{\{\omega \in N^{-1}(\{n\})\}}$ est une fonction mesurable positive car $\omega \in \Omega \mapsto N(\omega)$ est mesurable. De plus $\omega \in \Omega \mapsto \sum_{i=1}^{n} X_i(\omega)$ est une fonction mesurable positive comme somme de fonction mesurable positive. Ainsi $\omega \in \Omega \mapsto \mathbb{I}_{N(\omega)=n} \sum_{i=1}^{n} X_i(\omega)$ est une fonction mesurable positive, et il en est de même pour Z: c'est donc bien une variable aléatoire.
Si $\lambda > 0$ alors N prend ses valeurs dans $\mathbf{N}$. Et si $p > 0$ alors Z prend ses valeurs dans N comme somme d'entiers non tous nuls.

2. Pour $n = 0$, alors il est clair que $\mathbb{P}\big(Z = k \cap N = n\big) = 0$ si $k \geq 1$ et $\mathbb{P}\big(Z = 0 \cap N = 0\big) = \mathbb{P}\big(N = 0\big) = e^{-\lambda}$.
Par ailleurs, pour $(n,k) \in \mathbf{N}^* \times \mathbf{N}$, on a

$$\begin{aligned}
\mathbb{P}\big(Z = k \cap N = n\big) &= 0 \quad \text{si } n < k \text{ car } Z \leq N \text{ puisque } X_i \leq 1. \\
&= \mathbb{P}\Big(\sum_{i=1}^{n} X_i = k \bigcap N = n\Big) \quad \text{si } n \geq k \\
&= \mathbb{P}\Big(\sum_{i=1}^{n} X_i = k\Big)\, \mathbb{P}\Big(N = n\Big) \quad \text{par l'indépendance entre } N \text{ et les } X_i \\
&= \binom{k}{n}\, p^k (1-p)^{n-k}\, e^{-\lambda}\, \frac{\lambda^n}{n!} \quad \text{car alors loi binomiale,}
\end{aligned}$$

et cette formule est également valable pour $n = k = 0$.
Pour $k \in \mathbf{N}$, on peut utiliser la formule des probabilités totales et on obtient:

$$\begin{aligned}
\mathbb{P}\big(Z = k\big) &= \sum_{n=k}^{\infty} \mathbb{P}\big(Z = k \cap N = n\big) \\
&= \sum_{n=k}^{\infty} \binom{k}{n}\, p^k (1-p)^{n-k}\, e^{-\lambda}\, \frac{\lambda^n}{n!} \\
&= e^{-\lambda}\, \frac{1}{k!}\, p^k\, \lambda^k \sum_{n=k}^{\infty} (1-p)^{n-k}\, \frac{\lambda^{n-k}}{(n-k)!} \\
&= e^{-\lambda}\, \frac{1}{k!}\, (\lambda\, p)^k \sum_{n'=0}^{\infty} (1-p)^{n'}\, \frac{\lambda^{n'}}{n'!} \quad \text{après changement de variable } n' = n - k \\
&= e^{-\lambda}\, \frac{1}{k!}\, (\lambda\, p)^k\, e^{(1-p)\lambda} = e^{-p\,\lambda}\, \frac{(\lambda\, p)^k}{k!},
\end{aligned}$$

qui est la probabilité d'une loi de Poisson de paramètre $\lambda\, p$.

$\square$