Correction de la liste d'exercices n°5

Exercice n°1 :

▶ S'il arrive à 7h30, la loi de son temps d'attente X est uniforme sur [0,10]. Alors

$$\mathbb{E}(X) = 5 \quad var(X) = 25/3$$

▶ S'il a le choix avec la deuxieme ligne, son temps d'attente s'écrit $Z = X \wedge Y$ où X et Y sont indépendantes et uniformes sur [0,10]. On montre que Z a pour densité $f(x) = \frac{2}{10}(1 - \frac{x}{10})$, alors

$$\mathbb{E}(Z) = \frac{10}{3} \quad var(Z) = 50/9$$

▶ S'il arrive à 7h36, la probabilité d'avoir un bus avant 7h40 est

$$\mathbb{P}(X \vee Y > 6) = 1 - \mathbb{P}(X \wedge Y \leq 6) = 1 - \mathbb{P}(X \leq 6 \,\cap\, Y \leq 6) = 1 - \mathbb{P}(X \leq 6)^2 = 1 - \left(\frac{6}{10}\right)^2$$

Exercice n°2 :

▶ Concernant l'existence, c'est la principe de la borne inf...

▶ Soit $x \in \mathbb{R}$, alors par définition $G(F(x)) := \inf\{y \in \mathbb{R}, F(y) \geq \mathbb{F}(x)\}$, comme $F(x) \geq F(x)$!!! l'infimum est au plus égal à x donc $G(F(x)) \leq x$

▶ Soit $p \in [0,1]$ tel que $G(p) \in \mathbb{R}$. Clairement, pour $y > G(p)$ on a $F(y) \geq p$. Soit alors une suite (y_n) convergent en décroissant vers $G(p)$, comme F est continue à droite on a

$$F(G(p)) \geq \lim_n \searrow F(y_n) \geq p$$

▶ Pour le sens $\Rightarrow$ il suffit d'appliquer F et d'utiliser sa croissance. Pour les sens $\Leftarrow$, c'est immédiat vu la définition de la fonction G.

▶ Si $U \sim U_{[0,1]}$ alors pour $x \in \mathbb{R}$, d'après l'équivalence montrée ci dessus

$$P(G \circ U \leq x) = \mathbb{P}(U \leq F(x)) = F(x)$$

La fonction de répartition de $G \circ U$ n'est autre que F.

▶ Si F est bijective, alors G est simplement l'inverse de la fonction F. Par exemple, la variable $X \sim \mathcal{E}(\lambda)$ a pour fonction de répartition $F(x) = 1 - e^{-\lambda x}$ qui est bijective d'inverse $G(y) = F^{-1}(y) = -\frac{1}{\lambda}\log(1 - y)$. Si U est une variable uniforme sur [0,1], la variable $Y = F^{-1}(U)$ suit une loi $\mathcal{E}(\lambda)$.

Exercice n°3 :

► Si l'étudiant part à 7h15, il a 45 minutes pour faire son trajet. S'il souhaite arriver à l'heure, il doit choisir la variable Z telle que $\mathbb{P}(Z \geq 45)$ est minimum. Or

$$\begin{cases} \mathbb{P}(X \geq 45) = \mathbb{P}((X - 35.2)/5 \geq (45 - 35.2)/5) = \mathbb{P}(N \geq (45 - 35.2)/5) \\ \\ \mathbb{P}(Y \geq 45) = \mathbb{P}((Y - 40)/2 \geq (45 - 40)/2) = \mathbb{P}(N \geq (45 - 40)/2) \end{cases}$$

où $N \sim \mathcal{N}(0,1)$. Comme $(45 - 35.2)/5) \leq (45 - 40)/2)$ il faut qu'il choisisse le chemin correspondant à la variable Y.

► Dans le cas où il part à 7h30 on a

$$\begin{cases} \mathbb{P}(X \geq 30) = \mathbb{P}((X - 35.2)/5 \geq (30 - 35.2)/5) = \mathbb{P}(N \geq (30 - 35.2)/5) \\ \\ \mathbb{P}(Y \geq 30) = \mathbb{P}((Y - 40)/2 \geq (30 - 40)/2) = \mathbb{P}(N \geq (30 - 40)/2) \end{cases}$$

où $N \sim \mathcal{N}(0,1)$ Comme $(30 - 40)/2) \leq (30 - 35.2)/5)$, il vaut mieux choisir ici le chemin correspondant à la variable X

Exercice n°4 :

► Soit $Y \sim \mathcal{E}(\lambda)$, alors pour $s,t \in \mathbb{R}^+$,

$$\mathbb{P}(Y > s + t \,|\, Y > s) = \mathbb{P}(Y > s + t)/\mathbb{P}(Y > s) = e^{-\lambda(s+t)}/e^{-\lambda s} = e^{-\lambda t} = \mathbb{P}(Y > t)$$

► Soit X qui vérifie la propriété de vieillissement ci dessus, on introduit la fonction $F : \mathbb{R}^+ \to \mathbb{R}^+$

$$F(s) = -\log(\mathbb{P}(X > s))$$

alors on a

$$F(s + t) = F(s) + F(t)$$

On montre facilement que $F(s) = s \times F(1)$, c'est vrai pour les entiers, les rationnels, puis par continuité sur les réels. On en déduit que $\mathbb{P}(X > s) = e^{-\lambda s}$ avec $\lambda = F(1) > 0$.

► On traite le cas des variables à valeurs dans $\mathbb{N}$. Soit X une variable de loi $\mathcal{G}(p)$, alors $\mathbb{P}(X > n) = (1 - p)^n$ donc pour $m,n \in \mathbb{N}^*$

$$\mathbb{P}(X > n + m | X > n) = \mathbb{P}(X > n + m \text{ et } X > n)/\mathbb{P}(X > n)$$

$$\mathbb{P}(X > n + m | X > n) = \mathbb{P}(X > n + m)/\mathbb{P}(X > n) = (1 - p)^m = \mathbb{P}(X > m)$$

Une variable géométrique vérifie donc la propriété de vieillissement.

► Soit maintenant une variable Y à valeurs dans $\mathbb{N}^*$ qui vérifie la propriété de vieillissement. On définit une fonction $F : \mathbb{N} \to \mathbb{R}^+$ par

$$F(n) = -\log(\mathbb{P}(Y > n))$$

On a bien sûr $F(0) = 0$ et d'après la propriété de vieillissement pour $m,n \in \mathbb{N}^*$:

$$F(n + m) = F(n) + F(m)$$

On déduit facilement que $F(n) = n \times F(1)$ pour tout $n \in \mathbb{N}^*$. La fonction F est à valeurs dans $\mathbb{R}^+$, on peut donc définir $p \in [0,1]$ par $1 - p := e^{-F(1)}$, alors pour $n \in \mathbb{N}^*$

$$\begin{aligned}\mathbb{P}(Y = n) &= 1 - \mathbb{P}(Y \neq n) = 1 - \mathbb{P}(Y > n \quad \text{ou} \quad Y < n) = 1 - \mathbb{P}(Y > n) + \mathbb{P}(Y < n) \\ &= \mathbb{P}(Y \geq n) - \mathbb{P}(Y > n) = \mathbb{P}(Y > n-1) - \mathbb{P}(Y > n) = e^{-F(n-1)} - e^{-F(n)} \\ &= e^{-(n-1)F(1)} - e^{-nF(1)} = e^{-(n-1)F(1)}(1 - e^{-F(1)}) = (1-p)^{n-1} \times p\end{aligned}$$

Y est une donc une variable de loi $\mathcal{G}(p)$.
Les seules variables à valeurs dans $\mathbb{N}^*$ qui vérifie la propriété de vieillissement sont les variables de loi géométriques !

Exercice n°5 :

► Soient U_1, U_2, U_3 trois variables indépendantes de loi uniforme sur $[0,1]$, on notera $U_i \sim U_{[0,1]}$.

$$\mathbb{P}(U_i = U_j) = 1 - \mathbb{P}(U_i \neq U_j) = 1 - (\mathbb{P}(U_i > U_j) + \mathbb{P}(U_j > U_i)) = 1 - 2\mathbb{P}(U_i > U_j)$$

Or

$$\mathbb{P}(U_i > U_j) = \mathbb{E}(1_{U_i > U_j}) = \int_0^1 \int_0^1 1_{x>y} dxdy = 1/2$$

donc $\mathbb{P}(U_i = U_j) = 0$, *i.e* les variables U_1, U_2, U_3 sont *ps* distinctes.

► On ordonne les variables U_1, U_2, U_3 en $U_{(1)} < U_{(2)} < U_{(3)}$. Alors, si f est une fonction continue bornée sur $[0,1]^3$

$$\mathbb{E}(f(U_{(1)},U_{(2)},U_{(3)}) = \int_{[0,1]^3} f(x \wedge y \wedge z, x+y+z-(x \wedge y \wedge z) - (x \vee y \vee z), (x \vee y \vee z))dxdydz$$

On découpe ensuite $[0,1]^3$ (en oubliant les diagonales de mesure nulle d'aprés la première question) en les six parties du type $\{(x,y,z) \in [0,1]^3, x < y < z\}$ (6 parties car il y a 6 façon de permuter x,y et z), alors par symétrie

$$\mathbb{E}(f(U_{(1)},U_{(2)},U_{(3)}) = 6 \int_{\{(x,y,z)\in[0,1]^3, x<y<z\}} f(x,y,z)dxdydz$$

On déduit que $(U_{(1)},U_{(2)},U_{(3)})$ a pour densité

$$6 \times 1_{0<x<y<z<1}$$

► Plus généralement, si $U_1,U_2,...,U_n$ sont n variables indépendantes de loi uniforme sur $[0,a]$, alors le vecteur ordonné $(U_{(1)},U_{(2)},...,U_{(n)})$ a pour densité

$$\frac{n!}{a^n} \times 1_{0<x_1<x_2<...<x_n<a}$$

Pour obtenir les lois marginales, il suffit d'intégrer selon les autres variables. En particulier, la loi de $(U_{(1)},U_{(n)})$ admet pour densité

$$\frac{n!}{a^n} \times \int_{(x_2,x_3,...,x_{n-1})\in[0,a]^{n-2}} 1_{0<x_1<x_2<...<x_n<a} dx_2dx_3...dx_{n-1} = \frac{n!}{a^n} \frac{(x_n - x_1)^{n-2}}{(n-2)!} = n(n-1)\frac{(x_n - x_1)^{n-2}}{a^n}$$

Pour $a = 1$ on trouve que $(U_{(1)},U_{(n)})$ a pour densité $n(n-1)(x_n - x_1)^{n-2}$

► Une autre façon de calculer la densité de $(U_{(1)},U_{(n)})$ consiste à écrire

$$\mathbb{P}(x < U_{(1)} < U_{(n)} < y) = \mathbb{P}(x < U_i < y, i = 1...n) = \mathbb{P}(x < U_1 < y)^n = 1 - (y-x)^n$$

Alors en prenant les dérivées partielles en x et y, on trouve que la densité est $n(n-1)(y-x)^{n-2}$

► Soient $0 < x < y < 1$,

$$\mathbb{P}(U_{(1)} < x < y < U_{(n)}) = 1 - \mathbb{P}(U_{(1)} \geq x \,\cup\, U_{(n)} \leq y)$$

Or

$$\mathbb{P}(U_{(1)} \geq x \,\cup\, U_{(n)} \leq y) = \mathbb{P}(U_{(1)} \geq x) + \mathbb{P}(U_{(n)} \leq y) - \mathbb{P}(x \leq U_{(1)} < U_{(n)} \leq y)$$

Donc

$$\mathbb{P}(U_{(1)} \geq x \,\cup\, U_{(n)} \leq y) = (1-x)^n + y^n - (1-(y-x)^n) \longrightarrow 0 \quad \text{quand} \quad n \to +\infty$$

d'où

$$\lim_{n\to+\infty} \mathbb{P}(U_{(1)} < x < y < U_{(n)}) = 1$$

Exercice n°6 :

► Soient X_1 et X_2 deux variables indépendantes $X_i \sim \mathcal{E}(\lambda_i)$. On ordonne les variables en $X_{(1)} < X_{(2)}$, alors pour $x \geq 0$

$$\mathbb{P}(X_{(1)} > x) = \mathbb{P}(X_1 > x \,\cap\, X_2 > x) = \mathbb{P}(X_1 > x) \times \mathbb{P}(X_2 > x) = e^{-(\lambda_1+\lambda_2)x}$$

Donc $X_{(1)} \sim \mathcal{E}(\lambda_1 + \lambda_2)$.

$$\mathbb{P}(X_{(2)} \leq x) = \mathbb{P}(X_1 \leq x \,\cap\, X_2 \leq x) = \mathbb{P}(X_1 \leq x) \times \mathbb{P}(X_2 \leq x) = (1 - e^{-\lambda_1 x})(1 - e^{-\lambda_2 x})$$

Alors $X_{(2)}$ a pour densité $\lambda_1 e^{-\lambda_1 x} + \lambda_2 e^{-\lambda_2 x} - (\lambda_1 + \lambda_2)e^{-(\lambda_1+\lambda_2)x}$.

► On suppose que $\lambda_1 = \lambda_2$. Soit f une fonction continue bornée, alors

$$\mathbb{E}(f(X_{(1)},X_{(2)} - X_{(1)})) = \int_{\mathbb{R}^+}\int_{\mathbb{R}^+} f(x \wedge y, x \vee y - x \wedge y)\lambda^2 e^{-\lambda(x+y)} dxdy$$

$$= 2\int_{\mathbb{R}^{+2},\, x<y} f(x, y-x)\lambda^2 e^{-\lambda(x+y)} dxdy = 2\int_{\mathbb{R}^+} f(x,v)\lambda^2 e^{-\lambda(v+2x)} dxdv$$

La densité de $(X_{(1)},X_{(2)} - X_{(1)})$ s'écrit comme un produit de densité, on en déduit d'une part que $X_{(1)}$ et $X_{(2)} - X_{(1)}$ sont indépendantes, d'autre part que $X_{(1)} \sim \mathcal{E}(2\lambda)$ et $X_{(2)} - X_{(1)} \sim \mathcal{E}(\lambda)$

Exercice n°7 : Sur les lois $\Gamma(a,b)$

Une variable de loi $\Gamma(a,b)$ est une variable à valeurs dans $\mathbb{R}^+$ dont la densité par rapport à la mesure de Lebesgue s'écrit

$$\frac{b^a}{\Gamma(a)} x^{a-1} e^{-bx}\, 1_{x\geq 0}$$

On remarque qu'une loi $\Gamma(1,\lambda)$ est en fait une loi $\mathcal{E}(\lambda)$. Soient maintenant deux variables indépendantes X et Y de lois respective $\Gamma(n,\lambda)$ et $\Gamma(1,\lambda)$, alors la variable $X + Y$ suit une loi $\Gamma(n+1,\lambda)$. En effet, si f est une fonction continue bornée

$$\mathbb{E}(f(X+Y)) = \int_{\mathbb{R}^{+2}} f(x+y)\frac{\lambda^n}{\Gamma(n)} x^{n-1} e^{-\lambda x} \times \lambda e - \lambda y dxdy = \int_{\mathbb{R}^{+2}} f(x+y)\frac{\lambda^{n+1}}{\Gamma(n)} x^{n-1} e^{-\lambda(x+y)} dxdy$$

$$= \int_{(x,u)\in\mathbb{R}^{+2},u>x} f(u)\frac{\lambda^{n+1}}{\Gamma(n)}x^{n-1}e^{-\lambda u}dxdu = \int_{\mathbb{R}^+} f(u)\frac{\lambda^{n+1}}{\Gamma(n)}e^{-\lambda u}\left(\int_0^u x^{n-1}dx\right)du$$

$$i.e \quad \mathbb{E}(f(X+Y)) = \int_{\mathbb{R}^+} f(u)\frac{\lambda^{n+1}}{\Gamma(n)}e^{-\lambda u}u^n du \quad \text{et} \quad X+Y \sim \Gamma(n+1,\lambda)$$

Calculons la moyenne et la variance d'une variable $X \sim \Gamma(a,b)$

$$\mathbb{E}(X) = \int_{\mathbb{R}^+} x \times \frac{b^a}{\Gamma(a)}x^{a-1}e^{-bx}dx = \frac{b^a}{\Gamma(a)}\int_{\mathbb{R}^+} x^a e^{-bx}dx = \frac{b^a}{\Gamma(a)}\frac{\Gamma(a+1)}{b^{a+1}} = \frac{a}{b}$$

$$\mathbb{E}(X^2) = \int_{\mathbb{R}^+} x^2 \times \frac{b^a}{\Gamma(a)}x^{a-1}e^{-bx}dx = \frac{b^a}{\Gamma(a)}\int_{\mathbb{R}^+} x^{a+1} e^{-bx}dx = \frac{b^a}{\Gamma(a)}\frac{\Gamma(a+2)}{b^{a+2}} = \frac{a(a+1)}{b^2}$$

On en déduit

$$var(X) = \frac{a(a+1)}{b^2} - \left(\frac{a}{b}\right)^2 = \frac{a}{b^2}$$