TD7. Variables indépendantes, produit de convolution

1. Soient X et Y des variables aléatoires indépendantes. On suppose que X suit la loi de Poisson de paramètre $\theta_1 > 0$, et que Y suit la loi de Poisson de paramètre $\theta_2 > 0$.

a) Calculer f_X, la fonction génératrice de X. Quelle est la loi de $X+Y$?

b) Calculer φ_X, la fonction caractéristique de X. Quelle est la loi de $X+Y$?

c) Calculer le produit de convolution de P_X et P_Y. Quelle est la loi de $X+Y$?

Solution de l'exercice 1.

a) Soit $s \in [0, 1]$. Rappelons que $\mathbb{P}(X = n) = e^{-\theta_1} \frac{\theta_1^n}{n!}$, $n \in \mathbb{N}$. Par la formule de transfert (pour fonctions bornées),

$$f_X(s) = \sum_{n=0}^{\infty} s^n \, \mathbb{P}(X = n) = \sum_{n=0}^{\infty} s^n e^{-\theta_1} \frac{\theta_1^n}{n!} = e^{-\theta_1} \, e^{s\theta_1} = e^{-\theta_1(1-s)} .$$

[C'est un calcul que l'on a déjà fait aux TD6.]

Pour déterminer la loi de $X+Y$ qui est une variable aléatoire à valeurs dans $\mathbb{N}$, il suffit de déterminer sa fonction génératrice. Comme $f_Y(s) = e^{-\theta_2(1-s)}$, $s \in [0, 1]$, on a, grâce à l'indépendance,

$$f_{X+Y}(s) = f_X(s) \, f_Y(s) = e^{-(\theta_1+\theta_2)(1-s)}, \qquad s \in [0, 1].$$

Autrement dit, $X+Y$ suit la loi de Poisson de paramètre $\theta_1 + \theta_2$.

b) Par la formule de transfert (pour fonctions bornées), pour $t \in \mathbb{R}$,

$$\varphi_X(t) = \sum_{n \in \mathbb{N}} e^{itn} \, \mathbb{P}(X = n) = \sum_{n \in \mathbb{N}} e^{itn} \, e^{-\theta_1} \frac{\theta_1^n}{n!} = e^{-\theta_1} \, e^{\theta_1 e^{it}} = e^{-\theta_1(1-e^{it})}.$$

De même, $\varphi_Y(t) = e^{-\theta_2(1-e^{it})}$, $t \in \mathbb{R}$. Par indépendance,

$$\varphi_{X+Y}(s) = \varphi_X(s) \, \varphi_Y(s) = e^{-(\theta_1+\theta_2)(1-e^{it})}, \qquad t \in \mathbb{R}.$$

Ainsi, $X+Y$ suit la loi de Poisson de paramètre $\theta_1 + \theta_2$.

c) L'indépendance implique que $P_{X+Y} = P_X * P_Y$ (résultat du cours). Calculons le produit de convolution de P_X et P_Y : pour $n \in \mathbb{N}$,

$$\begin{aligned}
P_{X+Y}(\{n\}) &= \sum_{k=0}^{n} P_X(\{k\})P_Y(\{n-k\}) = \sum_{k=0}^{n} P(X=k)P(Y=n-k) \\
&= \sum_{k=0}^{n} e^{-\theta_1}\frac{\theta_1^k}{k!}e^{-\theta_2}\frac{\theta_2^{n-k}}{(n-k)!} \\
&= e^{-\theta_1-\theta_2}\sum_{k=0}^{n}\frac{\theta_1^k\theta_2^{n-k}}{k!\,(n-k)!} \\
&= \frac{e^{-\theta_1-\theta_2}}{n!}\sum_{k=0}^{n}\binom{n}{k}\theta_1^k\theta_2^{n-k} = \frac{e^{-\theta_1-\theta_2}}{n!}(\theta_1+\theta_2)^n\,.
\end{aligned}$$

On reconnît, de nouveau, que $X+Y$ suit la loi de Poisson de paramètre $\theta_1+\theta_2$.

2. Soit $\theta > 0$ un réel. On effectue n expériences indépendantes, ayant chacune une probabilité $\frac{\theta}{n}$ de réussir. On note X_n le nombre d'expériences ayant réussi.

a) Déterminer la loi de X_n et sa fonction caractéristique φ_{X_n}.

b) Pour tout $t \in \mathbb{R}$, déterminer la limite de $\varphi_{X_n}(t)$ lorsque $n \to \infty$. On note $\varphi(t)$ cette limite. La fonction φ est-elle la fonction caractéristique d'une variable aléatoire ?

Solution de l'exercice 2.

a) X_n est une somme de n variables de Bernoulli indépendantes de même paramètre $\frac{\theta}{n}$, c'est donc une loi binomiale de paramètres n et $\frac{\theta}{n}$. On a vu aux TD6 que pour $t \in \mathbb{R}$,

$$\varphi_{X_n}(t) = \left(1 - \frac{\theta}{n} + \frac{\theta}{n}e^{it}\right)^n.$$

b) En utilisant la limite classique

$$\lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n = e^x, \qquad x \in \mathbb{C},$$

on trouve

$$\lim_{n\to\infty}\varphi_{X_n}(t) = e^{\theta e^{it}-\theta}, \qquad t \in \mathbb{R}.$$

On reconnaît la fonction caractéristique d'une variable aléatoire de Poisson de paramètre θ. [Plus tard, on dira que X_n converge en loi vers une variable aléatoire de Poisson de paramètre θ.]

3. Construire un couple de variables aléatoires X et Y de carré-intégrables tel que $\mathrm{Cov}(X, Y) < 0$.

Solution de l'exercice 3. Il y en a beaucoup. Si $\mathrm{Cov}(X,\, Y) > 0$, alors avec $Z := -Y$, on aura $\mathrm{Cov}(X,\, Y) < 0$. Donc, si $X \in \mathscr{L}^2$ n'est pas une constante (pour assurer que $\mathrm{Var}(X) > 0$; par exemple, si X suit la loi gaussienne $\mathscr{N}(0,\, 1)$) et si $Y := -X$, alors $\mathrm{Cov}(X,\, Y) = -\mathrm{Var}(X) < 0$.

4. Soient $p,\, q \in [0,\, 1]$ des réels. Soient X et Y des variables aléatoires réelles. On suppose que X suit la loi de Bernoulli de paramètre p et que Y la loi de Bernoulli de paramètre q. Montrer que

$$\max\{p+q-1,\, 0\} - pq \le \mathrm{Cov}(X,\, Y) \le \min\{p,\, q\} - pq.$$

Solution de l'exercice 4. On a $\mathrm{Cov}(X,Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = \mathbb{E}[XY] - pq$. Il nous faut donc encadrer $\mathbb{E}[XY]$.

D'une part, $XY \le X$, donc $\mathbb{E}[XY] \le \mathbb{E}[X] = p$. De même, $\mathbb{E}[XY] \le q$. Ainsi, $\mathbb{E}[XY] \le \min(p, q)$.

D'autre part, $\mathbb{E}[XY] = \mathbb{P}(X = 1,\, Y = 1)$. Or pour tous événements A et B, l'égalité $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$ et l'inégalité $\mathbb{P}(A \cup B) \le 1$ entraînent $\mathbb{P}(A \cap B) \ge \mathbb{P}(A) + \mathbb{P}(B) - 1$. Comme de plus $\mathbb{P}(A \cap B) \ge 0$, on a $\mathbb{P}(A \cap B) \ge \max\{\mathbb{P}(A) + \mathbb{P}(B) - 1,\, 0\}$. Avec $A = \{X = 1\}$ et $B = \{Y = 1\}$, on trouve $\mathbb{E}[XY] = \mathbb{P}(A \cap B) \ge \max\{p + q - 1,\, 0\}$. On a ainsi établi l'inégalité voulue.

[Il est possible de prouver que les deux bornes de l'égalité dans cet exercice peuvent être atteintes.]

5. Soient X et Y des variables aléatoires indépendantes suivant la même loi gaussienne $\mathscr{N}(0,\, 1)$. On pose $U := X + Y$ et $V := X - Y$.

a) Déterminer la loi de U, ainsi que celle de V.

b) Montrer que U et V sont indépendantes.

Solution de l'exercice 5.

a) Dans le cours, on a vu que $P_U = \mathscr{N}(0,\, 2)$.

Pour déterminer P_V, on constate que pour tout réel $b \ne 0$, $P_{bY} = \mathscr{N}(0,\, b^2)$: il suffit de remarquer que $\varphi_{bY}(t) = \mathbb{E}(e^{itbY}) = \varphi_Y(bt) = e^{-b^2t^2/2}$, $t \in \mathbb{R}$. En particulier, $-Y$ est une variable aléatoire indépendante de X (résultat du cours), qui suit la loi gaussienne $\mathscr{N}(0,\, 1)$. Donc $P_V = \mathscr{N}(0,\, 2)$.

b) D'après un résultat du cours, il suffit de prouver que $\mathbb{E}(e^{isU}\, e^{itV}) = \mathbb{E}(e^{isU})\, \mathbb{E}(e^{itV})$ pour tous $s,\, t \in \mathbb{R}$.

On écrit $e^{isU}\, e^{itV} = e^{i(s+t)X}\, e^{i(s-t)Y}$. L'indépendance permet de voir que

$$\begin{aligned}\mathbb{E}(e^{isU}\, e^{itV}) &= \mathbb{E}(e^{i(s+t)X})\, \mathbb{E}(e^{i(s-t)Y}) = \varphi_X(s+t)\, \varphi_Y(s-t)\\ &= e^{-(s+t)^2/2}\, e^{-(s-t)^2/2} = e^{-s^2-t^2}\,.\end{aligned}$$

D'autre part, comme $\varphi_U(r) = \varphi_V(r) = e^{-r^2}$, $r \in \mathbb{R}$, on a

$$\mathbb{E}(e^{isU})\, \mathbb{E}(e^{itV}) = \varphi_U(s)\, \varphi_V(t) = e^{-s^2-t^2}\,,$$

ce qui implique l'identité cherchée.

6. Soient X et Y des variables aléatoires indépendantes suivant la même loi gaussienne $\mathscr{N}(0,\, 1)$. On pose $Z := X^2 + Y^2$. Calculer la fonction de répartition de Z. Que peut-on dire de la loi de Z ?

Solution de l'exercice 6. Soit $z \in \mathbb{R}$. Si $z \le 0$, on a $F_Z(z) = 0$. Supposons donc que $z > 0$. D'après un résultat du cours (pour fonctions positives ou bornées),

$$\begin{aligned}F_Z(z) &= \mathbb{P}(Z \le z) = \mathbb{E}(\mathbb{1}_{]-\infty,\, z]}(X^2 + Y^2))\\ &= \int_{\mathbb{R}} \int_{\mathbb{R}} \mathbb{1}_{]-\infty,\, z]}(x^2 + y^2) \frac{1}{(2\pi)^{1/2}} e^{-x^2/2} \frac{1}{(2\pi)^{1/2}} e^{-y^2/2}\, dxdy.\end{aligned}$$

On passe aux coordonnées polaires avec $x = r\cos\theta$ et $y = r\sin\theta$ (pour $r \ge 0$ et $\theta \in [0,\, 2\pi]$), pour voir que

$$F_Z(z) = \int_0^{z^{1/2}} dr \int_0^{2\pi} d\theta\, \frac{1}{2\pi} e^{-r^2/2} r = \int_0^{z^{1/2}} dr\, e^{-r^2/2} r = 1 - e^{-z/2}.$$

[D'après Fubini–Tonelli, on peut intégrer dans un ordre quelconque.] Ainsi, $F_Z(z) = (1 - e^{-z/2})\mathbb{1}_{\mathbb{R}_+}(z)$. On reconnaît la fonction de répartition d'une variable aléatoire suivant la loi exponentielle de paramètre $\frac{1}{2}$. Autrement dit, Z suit la loi exponentielle de paramètre $\frac{1}{2}$.

7. Soient $X_1, \ldots, X_n$ des variables aléatoires indépendantes et identiquement distribuées. On pose $M_n := \max_{1 \le i \le n} X_i$ et $I_n := \min_{1 \le i \le n} X_i$.

a) Déterminer F_{M_n}, la fonction de répartition de M_n, en termes de F_{X_1}.

b) Déterminer F_{I_n}.

c) On suppose que X_1 suit la loi uniforme sur $[0,\, 1]$. Montrer que pour tout réel $x \in \mathbb{R}$, $\mathbb{P}(nI_n \le x)$ converge vers une limite que l'on précisera.

Solution de l'exercice 7.

a) Soit $x \in \mathbb{R}$. Comme $\{M_n \leq x\} = \cap_{i=1}^n \{X_i \leq x\}$, on a, grâce à l'indépendance,

$$F_{M_n}(x) = \prod_{i=1}^n \mathbb{P}(X_i \leq x) = [F_{X_1}(x)]^n.$$

b) Soit $x \in \mathbb{R}$. Cette fois-ci, il convient de considérer $\{I_n > x\}$, qui n'est autre que $\cap_{i=1}^n \{X_i > x\}$. L'indépendance implique alors que

$$F_{I_n}(x) = 1 - \mathbb{P}(I_n > x) = 1 - \prod_{i=1}^n \mathbb{P}(X_i > x) = 1 - [1 - F_{X_1}(x)]^n.$$

c) Si X_1 suit la loi uniforme sur $[0,\, 1]$, alors I_n est à valeurs dans $[0,\, 1]$. Donc $\mathbb{P}(nI_n \leq x) = F_{I_n}(\frac{x}{n}) = 0$ si $x \leq 0$.
Supposons maintenant que $x > 0$. Pour tout n suffisamment grand, $\frac{x}{n} \in [0,\, 1]$, et on a

$$\mathbb{P}(nI_n \leq x) = 1 - [1 - F_{X_1}(\frac{x}{n})]^n = 1 - [1 - \frac{x}{n}]^n .$$

Lorsque $n \to \infty$, l'expression à droite converge vers $1 - e^{-x}$.
En conclusion, $\mathbb{P}(nI_n \leq x) \to (1 - e^{-x})\mathbb{1}_{\mathbb{R}_+}(x)$. [On reconnaît la fonction de répartition d'une variable aléatoire suivant la loi exponentielle de paramètre 1.]

8. Soit X une variable aléatoire réelle suivant la loi uniforme sur $[-1,\, 1]$. On pose $Y := X^2$.

a) Calculer $\mathrm{Cov}(X,\, Y)$.

b) Les variables aléatoires X et Y sont-elles indépendantes ?

Solution de l'exercice 8.

a) Les variables aléatoires X et Y étant bornées, on a $X,\, Y \in \mathscr{L}^2$. La covariance est donc bien définie. On a

$$\mathbb{E}(X) = \int_{-1}^{1} x dx = 0, \qquad \mathbb{E}(X^3) = \int_{-1}^{1} x^3 dx = 0.$$

Donc $\mathrm{Cov}(X,\, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\,\mathbb{E}(Y) = \mathbb{E}(X^3) - \mathbb{E}(X)\,\mathbb{E}(X^2) = 0$.

b) Il est clair que X et Y ne sont pas indépendantes. Par exemple, avec $A := [0,\, \frac{1}{2}]$ et $B := [\frac{1}{2},\, 1]$, on a $\mathbb{P}(X \in A,\, Y \in B) = 0 \neq \mathbb{P}(X \in A)\,\mathbb{P}(Y \in B)$.

9. Soient X et Y des variables aléatoires indépendantes suivant la même loi gaussienne $\mathscr{N}(0,\, 1)$. On pose $Z := \frac{X}{Y}$ (qui est bien définie avec probabilité 1 car $\mathbb{P}(Y \neq 0) = 1$). À l'aide de la fonction de répartition, déterminer la loi de Z.

Solution de l'exercice 9. Soit $z \in \mathbb{R}$. D'après un résultat du cours (pour fonctions positives ou bornées),

$$\begin{aligned} F_Z(z) &= \mathbb{P}(Z \le z) = \mathbb{E}(\mathbb{1}_{]-\infty,\, z]}(\frac{X}{Y})) \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathbb{1}_{]-\infty,\, z]}(\frac{x}{y}) \frac{1}{(2\pi)^{1/2}} e^{-x^2/2} \frac{1}{(2\pi)^{1/2}} e^{-y^2/2}\, dxdy. \end{aligned}$$

Le théorème de Fubini–Tonelli permet d'intégrer dans un ordre quelconque. On fix $y \in \mathbb{R}\backslash\{0\}$, et calcule

$$I(y) := \int_{-\infty}^{\infty} \mathbb{1}_{]-\infty,\, z]}(\frac{x}{y}) \frac{1}{(2\pi)^{1/2}} e^{-x^2/2} \frac{1}{(2\pi)^{1/2}} e^{-y^2/2}\, dx = \frac{1}{2\pi} e^{-y^2/2} \int_{\mathbb{R}} \mathbb{1}_{]-\infty,\, z]}(\frac{x}{y}) e^{-x^2/2} dx.$$

Si $y > 0$, on a $\mathbb{1}_{]-\infty,\, z]}(\frac{x}{y}) = \mathbb{1}_{]-\infty,\, yz]}(x)$, et donc

$$I(y) = \frac{1}{2\pi} e^{-y^2/2} \int_{-\infty}^{yz} e^{-x^2/2} dx.$$

Si $y < 0$, $\mathbb{1}_{]-\infty,\, z]}(\frac{x}{y}) = \mathbb{1}_{[yz,\, \infty[}(x)$, et donc

$$I(y) = \frac{1}{2\pi} e^{-y^2/2} \int_{yz}^{\infty} e^{-x^2/2} dx = \frac{1}{2\pi} e^{-y^2/2} \int_{-\infty}^{-yz} e^{-u^2/2} du,$$

avec un changement de variables $u = -x$. Donc pour tout $y \in \mathbb{R}\backslash\{0\}$,

$$I(y) = \frac{1}{2\pi} e^{-y^2/2} \int_{-\infty}^{|y|z} e^{-x^2/2} dx = \frac{1}{2\pi} e^{-y^2/2} \int_{-\infty}^{z} e^{-y^2u^2/2} |y|\, du,$$

avec un changement de variables $x = |y|u$. Puisque $F_Z(z) = \int_{-\infty}^{\infty} I(y) dy$, il résulte de nouveau du théorème de Fubini–Tonelli que

$$F_Z(z) = \int_{-\infty}^{z} \Big(\int_{-\infty}^{\infty} \frac{1}{2\pi} e^{-y^2/2} e^{-y^2u^2/2} |y| dy \Big) du = \int_{-\infty}^{z} \frac{1}{\pi} \frac{1}{1+u^2}\, du\,.$$

À droite, on reconnaît la valeur de la fonction de répartition au point z d'une variable aléatoire de Cauchy. Comme la fonction de répartition caractérise la loi d'une variable aléatoire réelle, on déduit que Z suit la loi de Cauchy.