TD 10 – Indépendance – Corrigé

Exercice à préparer du TD 9

Exercice 0. Soit X une variable aléatoire réelle.

1. On suppose que X admet un moment d'ordre $n \in \mathbb{N}^*$.

   (a) Montrer que pour tout $t \in \mathbb{R}$:

$$\phi_X(t) = \sum_{k=0}^{n-1} \frac{(it)^k}{k!} \mathbb{E}\left[X^k\right] + \frac{(it)^n}{(n-1)!} \mathbb{E}\left[X^n \int_0^1 (1-u)^{n-1} \exp(ituX) du\right].$$

   (b) Montrer que

$$\phi_X(t) = \sum_{k=0}^{n} \frac{(it)^k}{k!} \mathbb{E}\left[X^k\right] + \frac{(it)^n}{n!} \epsilon_n(t),$$

   où $|\epsilon_n(t)| \leq 2\mathbb{E}[|X|^n]$ et $\lim_{t \to 0} \epsilon_n(t) = 0$.

2. On suppose que X admet des moments de tous ordres et que

$$\limsup_{n \to \infty} \frac{\|X\|_n}{n} = \frac{1}{R} < \infty.$$

   Ici, $\|X\|_n = \mathbb{E}[|X|^n]^{1/n}$. Montrer qu'alors ϕ_X est développable en série entière au voisinage de tout réel, le rayon de convergence étant $\geq R/e$. En déduire que :

$$\forall t \in \left]-\frac{R}{e}, \frac{R}{e}\right[, \qquad \phi_X(t) = \sum_{k=0}^{\infty} \frac{(it)^k}{k!} \mathbb{E}\left[X^k\right].$$

Corrigé :

1. (a) On applique la formule de Taylor avec reste intégral à $y \mapsto \exp(iy)$:

$$\exp(iy) = \sum_{k=0}^{n-1} \frac{(iy)^k}{k!} + \frac{(iy)^n}{(n-1)!} \int_0^1 (1-u)^{n-1} \exp(iuy) du.$$

   Le résultat s'ensuit en prenant $y = X$ et en prenant l'espérance (qui est linéaire).

Pour des questions, demande de précisions ou explications, n'hésitez pas à m'envoyer un mail à igor.kortchemski@ens.fr , ou bien à venir me voir au bureau V4.

(b) En remarquant que $\int_0^1 (1-u)^{n-1} du = 1/n$, on a :

$$\phi_X(t) = \sum_{k=0}^{n} \frac{(it)^k}{k!} \mathbb{E}\left[X^k\right] + \frac{(it)^n}{(n-1)!} \mathbb{E}\left[X^n \int_0^1 (1-u)^{n-1} (\exp(ituX)-1) du\right].$$

Ainsi,

$$\epsilon_n(t) = n\mathbb{E}\left[X^n \int_0^1 (1-u)^{n-1} (\exp(ituX)-1) du\right],$$

et en appliquant le théorème de Fubini, on obtient la majoration

$$|\epsilon_n(t)| \leq 2n\mathbb{E}\left[|X|^n\right] \int_0^1 (1-u)^{n-1} du = 2\mathbb{E}\left[|X|^n\right].$$

De plus, pour $u \in [0,1]$ on a

$$\left| n \int_0^1 (1-u)^{n-1} (\exp(ituX)-1) \right| \leq 2|X|^n (1-u)^{n-1},$$

majoration indépendante de t par une application $\lambda_{[0,1]} \otimes \mathbb{P}$ intégrable. Le théorème de convergence dominée garantit que $\lim_{t\to 0} \epsilon_n(t) = 0$.

2. Soit $t_0 \in \mathbb{R}$. Puisque X a des moments de tout ordre, d'après le TD précédent, elle est C^∞ et possède donc un développement de Taylor à tout ordre $n \geq 1$:

$$\phi_X(t) = \phi_X(t_0) + \sum_{k=1}^{n} \frac{(t-t_0)^k}{k!} \phi_X^{(k)}(t_0) + R_n(t, t_0),$$

où le reste est défini par

$$R_n(t_0, t) = \int_{t_0}^{t} \frac{(t-u)^n}{n!} \phi_X^{(n+1)}(u) du.$$

Il s'agit donc de démontrer que celui-ci tend vers 0. On a vu au TD précédent que $\phi_X^{(n+1)}(u) = i^{n+1}\mathbb{E}\left[X^{n+1}\exp(iuX)\right]$. Ainsi :

$$|R_n(t_0, t)| \leq \frac{(|t-t_0| \|X\|_{n+1})^{n+1}}{(n+1)!}.$$

L'hypothèse sur $\|X\|_n$ et la formule de Stirling permettent aisément de voir que cette quantité tend vers 0 pour $|t-t_0| < R/e$.

On a vu en TD que si X, Y sont deux variables aléatoires réelles telles que $\limsup_{n\to\infty} \frac{\|X\|_n}{n} < \infty$, $\limsup_{n\to\infty} \frac{\|Y\|_n}{n} < \infty$ et $\mathbb{E}[X^n] = \mathbb{E}[Y^n]$ pour tout $n \geq 1$, alors $\phi_X(t) = \phi_Y(t)$ pour tout $t \in \mathbb{R}$, et donc X et Y n'ont pas même loi (en revanche, ce n'est pas vrai en général, cf exercice 9 du TD 9).
□

0 – Petite question

Soient X et Y deux variables aléatoires réelles indépendantes. Soit $F : \mathbb{R}\times\mathbb{R} \to \mathbb{R}_+$ une fonction borélienne. Montrer que $\mathbb{E}[F(X,Y)] = \mathbb{E}[g(Y)]$, où $g : \mathbb{R} \to \mathbb{R}$ est la fonction définie par $g(y) = \mathbb{E}[F(X,y)]$ pour $y \in \mathbb{R}$.

Corrigé :

On écrit :

$$\begin{aligned}
\mathbb{E}[F(X,Y)] &= \int_{\mathbb{R}\times R} F(x,y) P_{(X,Y)}(dxdy) = \int_{\mathbb{R}\times R} F(x,y) P_X(dx) \otimes P_Y(dy) \\
&= \int_{\mathbb{R}} P_Y(dy) \left(\int_{\mathbb{R}} F(x,y) P_X(dx) \right) = \int_{\mathbb{R}} P_Y(dy) g(y) = \mathbb{E}[g(Y)].
\end{aligned}$$

□

1 – Variables aléatoires indépendantes

Exercice 1. Soient X et Y deux variables aléatoires gaussiennes (centrées réduites) indépendantes. Montrer que les variables aléatoires $\frac{X+Y}{\sqrt{2}}$ et $\frac{X-Y}{\sqrt{2}}$ soient indépendantes.

Corrigé : Soient $F, G : \mathbb{R} \to \mathbb{R}_+$ deux fonctions boréliennes. En remarquant que $(x,y) \mapsto (\frac{x+y}{\sqrt{2}}, \frac{x-y}{\sqrt{2}})$ est un C^1 difféomorphisme de $\mathbb{R}\times\mathbb{R}$ sur $\mathbb{R}\times\mathbb{R}$ de jacobien -1, la formule du changement de variable donne

$$\begin{aligned}
\mathbb{E}\left[F\left(\frac{X+Y}{\sqrt{2}}\right)G\left(\frac{X-Y}{\sqrt{2}}\right)\right] &= \int_{\mathbb{R}\times\mathbb{R}} F\left(\frac{x+y}{\sqrt{2}}\right)G\left(\frac{x-y}{\sqrt{2}}\right)\frac{1}{2\pi}e^{-\frac{x^2+y^2}{2}}dxdy \\
&= \int_{\mathbb{R}\times\mathbb{R}} F\left(\frac{x+y}{\sqrt{2}}\right)G\left(\frac{x-y}{\sqrt{2}}\right)\frac{1}{2\pi}e^{-\frac{\left(\frac{x+y}{\sqrt{2}}\right)^2+\left(\frac{x-y}{\sqrt{2}}\right)^2}{2}}dxdy \\
&= \int_{\mathbb{R}\times\mathbb{R}} F(x)G(y)\frac{1}{2\pi}e^{-\frac{x^2+y^2}{2}}dxdy \\
&= \left(\int_{\mathbb{R}} F(x)\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx\right)\left(\int_{\mathbb{R}} G(y)\frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}dy\right).
\end{aligned}$$

On en déduit que $\frac{X+Y}{\sqrt{2}}$ et $\frac{X-Y}{\sqrt{2}}$ soient indépendantes, et sont toutes les deux gaussiennes centrées réduites.

Exercice 2. On définit sur $(\Omega, \mathcal{A}, \mathbb{P})$ des variables aléatoires $U_1, \dots, U_n$ indépendantes et de loi uniforme sur $\{1, 2, \dots, p\}$.

1. Trouver la loi de $M_n = \max_{1\le k\le n} U_k$.
2. Montrer que
$$\frac{\mathbb{E}[M_n]}{p} \underset{p\to\infty}{\longrightarrow} \frac{n}{n+1}.$$

Corrigé :

1. Pour tout $k \in \{1, \dots, p\}$, on a,
$$\mathbb{P}(M_n \le k) = \mathbb{P}(U_1 \le k, \dots, U_n \le k) = \mathbb{P}(U_1 \le k)\dots\mathbb{P}(U_n \le k) = \mathbb{P}(U_1 \le k)^n = \left(\frac{k}{p}\right)^n.$$
On en déduit la loi de M_n :
$$\mathbb{P}(M_n = k) = \mathbb{P}(M_n \le k) - \mathbb{P}(M_n \le k-1) = \frac{k^n - (k-1)^n}{p^n}.$$
2. On a,
$$\mathbb{E}(M_n) = \sum_{k=1}^{p} \mathbb{P}(M_n \ge k),$$
et donc
$$\frac{E(M_n)}{p} = \frac{1}{p}\sum_{k=1}^{p}\left(1 - \left(\frac{k-1}{p}\right)^n\right) = \frac{1}{p}\sum_{k=0}^{p-1}\left(1 - \left(\frac{k}{p}\right)^n\right).$$
On reconnaît une somme de Riemann de la fonction $x \mapsto 1 - x^n$, dont l'intégrale entre 0 et 1 vaut $n/(n+1)$. D'où le résultat.

□

Exercice 3. **(Formule de compensation.)** Soient $(X_n)_{n\geq 1}$ une suite de variables aléatoires réelles indépendantes et de loi μ et N une variable aléatoire à valeurs dans $\mathbb{N}$ indépendante de la suite $(X_n)_{n\geq 1}$.

1. On suppose que μ est la loi de Bernoulli de paramètre $p \in]0,1[$ c'est-à-dire que
$$\mu = p\delta_1 + (1-p)\delta_0,$$
et que N suit la loi de Poisson de paramètre λ c'est-à-dire que
$$\mathbb{P}(N = k) = e^{-\lambda}\frac{\lambda^k}{k!}.$$
On pose
$$P = \sum_{i=1}^{N} X_i \quad \text{et} \quad F = N - P = \sum_{i=1}^{N}(1 - X_i),$$
avec $P = F = 0$ sur $\{N = 0\}$. Les variables aléatoires P et F représentent respectivement le nombre de piles et de faces dans un jeu de pile ou face de paramètre p à N lancers.
   (a) Déterminer la loi du couple (P,N).
   (b) En déduire les lois de P et F et montrer que P et F sont indépendantes.
2. On ne fait plus d'hypothèse sur les lois. Soit $f : \mathbb{R} \to \mathbb{R}_+$ une fonction mesurable. Montrer que
$$\mathbb{E}\left(\sum_{i=1}^{N} f(X_i)\right) = \mathbb{E}(N)\int_{\mathbb{R}} f(x)\,\mu(dx),$$
avec $\sum_{i=1}^{N} f(X_i) = 0$ sur $\{N = 0\}$.

Corrigé :

(1)(a) On a $\mathbb{P}(P = 0, N = 0) = \mathbb{P}(N = 0) = e^{-\lambda}$, et pour $n \geq 1$ et $0 \leq k \leq n$,
$$\begin{aligned}
\mathbb{P}(P = k, N = n) &= \mathbb{P}\left(N = n, \sum_{i=1}^{n} X_i = k\right) \\
&= \mathbb{P}(N = n)\mathbb{P}\left(\sum_{i=1}^{n} X_i = k\right) \\
&= e^{-\lambda}\frac{\lambda^n}{n!}\binom{n}{k}p^k(1-p)^{n-k} \\
&= e^{-\lambda}\frac{(\lambda p)^k}{k!}\frac{(\lambda(1-p))^{n-k}}{(n-k)!}
\end{aligned}$$

(1)(b) On a pour $k, l \geq 0$,
$$\mathbb{P}(P = k, F = l) = \mathbb{P}(P = k, N = k + l) = \left(e^{-\lambda p}\frac{(\lambda p)^k}{k!}\right)\left(e^{-\lambda(1-p)}\frac{(\lambda(1-p))^l}{l!}\right).$$
Donc les variables aléatoires P et F sont indépendantes et de lois respectives les lois de Poisson de paramètres λp et $\lambda(1-p)$.

(2) On a

$$\begin{aligned}
\mathbb{E}\left(\sum_{i=1}^{N} f(X_i)\right) &= \mathbb{E}\left(\sum_{n\geq 1} \mathbb{1}_{\{N=n\}} \sum_{i=1}^{n} f(X_i)\right) \\
&= \sum_{n\geq 1} \mathbb{E}\left(\mathbb{1}_{\{N=n\}} \sum_{i=1}^{n} f(X_i)\right) \\
&= \sum_{n\geq 1} \mathbb{P}(N=n)\mathbb{E}\left(\sum_{i=1}^{n} f(X_i)\right) \\
&= \sum_{n\geq 1} \mathbb{P}(N=n) n \mathbb{E}(f(X_1)) \\
&= \mathbb{E}(N) \int_{\mathbb{R}} f(x)\mu(dx).
\end{aligned}$$

□

2 – Indépendance et fonctions caractéristiques

Exercice 4. Soient X et Y deux variables aléatoires bornées. Démontrer que X et Y sont indépendantes si, et seulement si :

$$\forall k, l \in \mathbb{N}, \qquad \mathbb{E}\left[X^k Y^l\right] = \mathbb{E}\left[X^k\right]\mathbb{E}\left[Y^l\right]. \tag{1}$$

Indication. Lire le titre de cette partie.

Corrigé :

L'implication est facile, car si X et Y sont indépendantes, alors X^k et Y^k le sont également. Réciproquement, supposons que (1) est vérifiée. La fonction caractéristique de (X, Y) est donnée en $(u, v) \in \mathbb{R}^2$ par :

$$\phi_{(X,Y)}(u,v) = \mathbb{E}[\exp(iuX)\exp(iuY)] = \mathbb{E}\left[\left(\sum_{k=0}^{\infty} \frac{(iu)^k X^k}{k!}\right)\left(\sum_{l=0}^{\infty} \frac{(iv)^l Y^l}{l!}\right)\right]$$

Soit C un majorant de $|X|$ et $|Y|$, de sorte que

$$\sum_{k,l\geq 0} \frac{|u|^k |v|^k C^{k+l}}{k!l!} = \exp(|u|C)\exp(|v|C) < \infty.$$

On peut ainsi appliquer le théorème de convergence dominée :

$$\phi_{(X,Y)}(u,v) = \sum_{k,l\geq 0} i^{k+l} \frac{u^k v^l}{k!l!} \mathbb{E}\left[X^k Y^l\right].$$

D'après l'hypothèse, on a donc

$$\phi_{(X,Y)}(u,v) = \sum_{k,l\geq 0} i^{k+l} \frac{u^k v^l}{k!l!} \mathbb{E}\left[X^k\right]\mathbb{E}\left[Y^l\right] = \left(\sum_{k=0}^{\infty} \frac{(iu)^k \mathbb{E}\left[X^k\right]}{k!}\right)\left(\sum_{l=0}^{\infty} \frac{(iv)^l \mathbb{E}\left[Y^l\right]}{l!}\right).$$

En appliquant le théorème de convergence dominée de la même manière, il en découle

$$\phi_{(X,Y)}(u,v) = \mathbb{E}\left[\sum_{k=0}^{\infty} \frac{(iu)^k X^k}{k!}\right]\mathbb{E}\left[\sum_{l=0}^{\infty} \frac{(iv)^l Y^l}{l!}\right] = \phi_X(u) \cdot \phi_X(v),$$

d'où le résultat.

4 – Compléments (hors TD)

Exercice 5.

1. Mathias a deux enfants dont une fille. Quelle est la probabilité que l'autre enfant soit un garçon ?
2. Mathilde a deux enfants. Le plus jeune est une fille. Quelle est la probabilité que l'aîné soit un garçon ?

Corrigé :
Une famille de deux enfants peut se représenter par (a_1, a_2) où a_i est f (fille) ou g (garçon) suivant que le i-ième enfant est une fille ou un garçon. Tous les couples (a_i, a_j) sont équiprobables. Posons donc $\Omega = \{(f,g),(f,f),(g,g),(g,f)\}$.

1. On sait qu'un enfant est une fille on cherche donc $\mathbb{P}(E|A)$ avec $E = \{(f,g),(g,f),(g,g)\}$ et $A = \{(f,g),(f,f),(g,f)\}$. Ainsi, $\mathbb{P}(E|A) = 2/3$.
2. On cherche désormais $P(F|B)$ où $F = \{(g,g),(g,f)\}$ et $B = \{(f,f),(g,f)\}$. On trouve alors $\mathbb{P}(F|B) = 1/2$. $\square$

Exercice 6. (**Processus de Poisson**) Soit $(X_n, n \geq 1)$ une suite de variables aléatoires indépendantes définies sur $(\Omega, \mathcal{A}, \mathbb{P})$, de même loi exponentielle de paramètre 1. On pose $T_0 = 0$ et pour tout $n \geq 1$,

$$T_n = X_1 + \ldots + X_n.$$

Pour tout $t \geq 0$, on pose

$$N_t = \max\{n \geq 0 : T_n \leq t\}.$$

1. Soit $n \geq 1$. Calculer la loi du n-uplet $(T_1, \ldots, T_n)$.
2. En déduire la loi de N_t pour tout $t > 0$.
3. Pour $n \geq 1$ et $t > 0$, on définit sur Ω une nouvelle mesure de probabilité $\mathbb{Q}^{n,t}$ par la formule

$$\mathbb{Q}^{n,t}(A) = \frac{\mathbb{P}(A \cap \{N_t = n\})}{\mathbb{P}(N_t = n)}, \ A \in \mathcal{A}.$$

   Calculer la loi du n-uplet $(T_1, \ldots, T_n)$ sous la mesure de probabilité $\mathbb{Q}^{n,t}$.

Corrigé :

1. Soit $f : \mathbb{R}_+^n \to \mathbb{R}_+$, une fonction mesurable. On a

$$\mathbb{E}(f(T_1, \ldots, T_n)) = \int_{\mathbb{R}_+^n} f(x_1, x_1 + x_2, \ldots, x_1 + x_2 + \ldots + x_n) e^{-(x_1+x_2+\ldots+x_n)} \, dx_1 \ldots dx_n.$$

   Or $\phi : (x_1, \ldots, x_n) \in (R_+^*)^n \mapsto (x_1, x_1+x_2, \ldots, x_1+x_2+\ldots+x_n) \in \{0 < t_1 < \ldots < t_n\}$ est un $\mathcal{C}^1$-difféomorphisme de jacobien égal à 1 donc d'après la formule du changement de variables,

$$\mathbb{E}(f(T_1, \ldots, T_n)) = \int_{\mathbb{R}_+^n} f(t_1, t_2, \ldots, t_n) e^{-t_n} \mathbb{1}_{\{0<t_1<\ldots<t_n\}} \, dt_1 \ldots dt_n,$$

   ce qui signfie que la loi de $(T_1, \ldots, T_n)$ a pour densité $e^{-t_n} \mathbb{1}_{\{0<t_1<\ldots<t_n\}}$ par rapport à la mesure de Lebesgue sur $\mathbb{R}^n$.

2. Soit $n \in \mathbb{N}$. On a $\{N_t = n\} = \{T_n \leq t, T_{n+1} > t\}$ car la suite $(T_n)_{n\geq 0}$ est p.s. croissante et donc $\mathbb{P}(N_t = n) = \mathbb{P}(T_n \leq t, T_{n+1} > t)$. On en déduit d'après la question (1) que pour $n \geq 1$,

$$\begin{aligned}
\mathbb{P}(N_t = n) &= \int_{\mathbb{R}_+^{n+1}} e^{-t_{n+1}} \mathbb{1}_{\{0<t_1<\ldots<t_n<t_{n+1}\}} \mathbb{1}_{\{t_n \leq t\}} \mathbb{1}_{\{t_{n+1}>t\}} \, dt_1 \ldots dt_n dt_{n+1} \\
&= \int_{\mathbb{R}_+^n} \mathbb{1}_{\{0<t_1<\ldots<t_n \leq t\}} \, dt_1 \ldots dt_n \int_t^\infty e^{-t_{n+1}} \\
&= \frac{t^n}{n!} e^{-t},
\end{aligned}$$

où la deuxième égalité est une conséquence du théorème de Fubini-Tonelli. Et l'on a

$$\mathbb{P}(N_t = 0) = \mathbb{P}(T_1 > t) = \mathbb{P}(X_1 > t) = \int_t^\infty e^{-x} \, dx = e^{-t}.$$

On voit que N_t suit la loi de Poisson de paramètre t.

3. Soit $f : \mathbb{R}_+^n \to \mathbb{R}_+$, une fonction mesurable. On a

$$\begin{aligned}
&\mathbb{E}(f(T_1, \ldots, T_n)\mathbb{1}_{\{N_t=n\}}) \\
&\quad= \mathbb{E}(f(T_1, \ldots, T_n)\mathbb{1}_{\{T_n \leq t, T_{n+1}>t\}}) \\
&\quad= \int_{\mathbb{R}_+^{n+1}} f(t_1, \ldots, t_n)\mathbb{1}_{\{t_n \leq t, t_{n+1}>t\}}) e^{-t_{n+1}} \mathbb{1}_{\{0<t_1<\ldots<t_n<t_{n+1}\}} \, dt_1 \ldots dt_n dt_{n+1} \\
&\quad= \int_{\mathbb{R}_+^n} f(t_1, \ldots, t_n)\mathbb{1}_{\{0<t_1<\ldots<t_n<t\}} \, dt_1 \ldots dt_n \int_t^\infty e^{-t_{n+1}} dt_{n+1} dt_{n+1} \\
&\quad= e^{-t} \int_{\mathbb{R}_+^n} f(t_1, \ldots, t_n)\mathbb{1}_{\{0<t_1<\ldots<t_n<t\}} \, dt_1 \ldots dt_n,
\end{aligned}$$

où la troisième égalité est une conséquence du théorème de Fubini-Tonelli. Donc

$$\begin{aligned}
\mathbb{E}^{\mathbb{Q}^{n,t}}(f(T_1, \ldots, T_n)) &= \frac{\mathbb{E}(f(T_1, \ldots, T_n)\mathbb{1}_{\{N_t=n\}})}{\mathbb{P}(N_t = n)} \\
&= \frac{n!}{t^n} \int_{\mathbb{R}_+^n} f(t_1, \ldots, t_n)\mathbb{1}_{\{0<t_1<\ldots<t_n<t\}} \, dt_1 \ldots dt_n,
\end{aligned}$$

ce qui signfie que la loi de $(T_1, \ldots, T_n)$ sous $\mathbb{Q}^{n,t}$ a pour densité $n! \, t^{-n} \, \mathbb{1}_{\{0<t_1<\ldots<t_n<t\}}$ par rapport à la mesure de Lebesgue sur $\mathbb{R}^n$. □

Exercice 7. (✯✯) Soient X et Y deux variables aléatoires réelles indépendantes et telles que les variables aléatoires $X + Y$ et $X - Y$ soient indépendantes. Montrer que les deux variables X et Y sont deux variables aléatoires gaussiennes.

Grandes étapes de la solution : en utilisant une équation fonctionnelle vérifiée par les fonctions caractéristiques, trouver le module de la fonction caractéristique de $X + Y$, puis son argument.

Fin