X2016 – MAP 311

PC 6 – 29 mai 2017 – Fonctions caractéristiques, Monte-Carlo, convergence en loi

Corrigé des questions non abordées en PC

Igor Kortchemski (doublage : Lucas Gerin)– igor.kortchemski@cmap.polytechnique.fr

Corrigé des exercices non traités sur http://www.normalesup.org/~kortchem/MAP311 *un peu après la PC.*

2 Monte Carlo

Exercice 2. **(Des calculs sans calculs)** Soit $f : \mathbb{R} \to \mathbb{R}$ une fonction continue bornée et $\alpha > 0$. Calculer les limites suivantes :

$$\lim_{n\to\infty} \int_{[0,1]^n} f\left(\frac{x_1+\cdots+x_n}{n}\right) \mathrm{d}x_1 \cdots \mathrm{d}x_n \qquad \text{et} \qquad \lim_{n\to\infty} \sum_{k\geq 0} e^{-\alpha n} \frac{(\alpha n)^k}{k!} f\left(\frac{k}{n}\right).$$

Corrigé :

– Soit X une variable aléatoire de Poisson de paramètre αn. Alors

$$\sum_{k\geq 0} e^{-\alpha n} \frac{(\alpha n)^k}{k!} f\left(\frac{k}{n}\right) = \mathbb{E}[f(X)].$$

Or X a la même loi que $V_1 + V_2 + \cdots + V_n$, où $(V_i)_{1\leq i\leq n}$ sont des variables aléatoires indépendantes de même loi de Poisson de paramètre α. Donc

$$\sum_{k\geq 0} e^{-\alpha n} \frac{(\alpha n)^k}{k!} f\left(\frac{k}{n}\right) = \mathbb{E}\left[f\left(\frac{V_1+\cdots+V_n}{n}\right)\right].$$

D'après la loi forte des grands nombres, $\frac{V_1+\cdots+V_n}{n}$ converge presque sûrement vers $\mathbb{E}[V_1] = \alpha$. Donc $f\left(\frac{V_1+\cdots+V_n}{n}\right)$ converge presque sûrement vers $f(\alpha)$. De plus, $|f\left(\frac{V_1+\cdots+V_n}{n}\right)|$ est majorée par $\sup|f|$. Donc

$$\sum_{k\geq 0} e^{-\alpha n} \frac{(\alpha n)^k}{k!} f\left(\frac{k}{n}\right) \underset{n\to\infty}{\longrightarrow} f(\alpha).$$

d'après le théorème de convergence dominée.

□

3 Convergence en loi

Exercice 6. **(Convergence en loi et fonctions caractéristiques)** Soit $\theta > 0$. On considère une suite $(T_n)_{n\geq n_0}$ de variables aléatoires où T_n suit une loi géométrique de paramètre $p_n = \frac{\theta}{n}$. Montrer que la suite $(T_n/n)_{n\geq n_0}$ converge en loi et déterminer sa limite.

Pour des questions, demande d'explications etc., n'hésitez pas à m'envoyer un mail.

Corrigé : On utilise les fonctions caractéristiques. Pour tout $t \in \mathbb{R}$,

$$\phi_{T_n}(t) = \mathbb{E}\left[e^{itT_n}\right] = \sum_{k\geq 1} p_n(1-p_n)^{k-1}e^{itk} = p_n e^{it}\sum_{k\geq 0}\left((1-p_n)e^{it}\right)^k = \frac{p_n e^{it}}{1-(1-p_n)e^{it}} = \frac{p_n}{e^{-it}-(1-p_n)}.$$

Ainsi,

$$\phi_{\frac{T_n}{n}}(t) = \mathbb{E}\left[e^{i\frac{t}{n}T_n}\right] = \frac{\frac{\theta}{n}}{1-i\frac{t}{n}+o(\frac{1}{n})-(1-\frac{\theta}{n})} \quad \underset{n\to\infty}{\longrightarrow} \quad \frac{\theta}{\theta - it}.$$

On reconnaît la fonction caractéristique d'une variable exponentielle de paramètre θ. □

Exercice 7. **(Un exemple)** Pour tout $n \geq 1$, on considère la fonction F_n définie par $F_n(x) = \frac{e^{nx}}{e^{nx}+1}$ pour $x \in \mathbb{R}$. Montrer que, pour tout $n \geq 1$, il existe une variable aléatoire X_n dont F_n est une fonction de répartition. Montrer que la suite $(X_n)_{n\geq 1}$ converge en loi, et identifier la loi limite.

Corrigé : La fonction F_n est croissante, continue, de limite nulle en $-\infty$ et de limite 1 en $+\infty$, donc il existe une variable aléatoire X_n dont F_n est une fonction de répartition. De plus, pour tout $x \in \mathbb{R}$,

$$F_n(x) \quad \underset{n\to\infty}{\longrightarrow} \quad \begin{cases} 0 & \text{si } x < 0 \\ \frac{1}{2} & \text{si } x = 0 \\ 1 & \text{si } x > 0. \end{cases}$$

Ainsi, X_n converge en loi vers 0 car la fonction de répartition de X_n converge vers la fonction de répartition F de la variable aléatoire constante égale à 0 en tout point où F est continue (c'est-à-dire $\mathbb{R}^*$). Alternativement, on voit que pour tout $\epsilon > 0$, on a $\mathbb{P}(|X_n| > \epsilon) \to 0$, donc X_n converge en probabilité et donc en loi vers 0. □

Exercice 8. **(Un autre exemple)** Pour tout $n \geq 1$, on considère la fonction F_n définie par $F_n(x) = 0$ si $x < n$ et $F_n(x) = 1$ si $x \geq n$. Montrer que, pour tout $n \geq 1$, existe une variable aléatoire X_n dont F_n est une fonction de répartition. Est-ce que la suite $(X_n)_{n\geq 1}$ converge en loi ?

Corrigé : La fonction F_n est la fonction de répartition d'une variable aléatoire constante égale à n. Pour tout $x \in \mathbb{R}$, $F_n(x) \to 0$. La suite X_n ne converge donc pas en loi, car la fonction nulle n'est pas égale à une fonction de répartition, sauf en un nombre dénombrable de points. □

5 Plus appliqué (hors PC)

Exercice 10. Cet exercice présente un modèle pour la contamination au mercure.

(1) Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires indépendantes de même loi. On suppose qu'il existe deux réels $\alpha, \lambda > 0$ tels que $\mathbb{P}(X_1 > x) \sim \frac{\alpha}{x^\lambda}$ lorsque $x \to \infty$. Démontrer que la suite $(Z_n)_{n\geq 1}$ définie par $Z_n = n^{-\frac{1}{\lambda}}\max(X_1, \ldots, X_n)$ converge en loi vers une variable aléatoire dont la fonction de répartition est $e^{-\alpha y^{-\lambda}}\mathbb{1}_{y>0}$ (loi de Fréchet de paramètre (α, λ)).

(2) Le mercure, métal lourd, est présent dans peu d'aliments. On le trouve essentiellement dans les produits de la mer. L'Organisation Mondiale de la Santé fixe la dose journalière admissible en mercure à $0.71\mu g$ par jour et par

kilo de poids corporel. Des études statistiques donnent la forme de la queue de distribution empirique de la contamination globale annuelle en gramme de mercure pour un individu de 70 kg : $\mathbb{P}(X > x) = \frac{\alpha}{x^\lambda}$ pour x assez grand avec $\alpha = 3.54 \cdot 10^{-9}$ et $\lambda = 2.58$. Seriez-vous étonné–e qu'au moins une personne soit exposée à ce risque sanitaire en France ? À Palaiseau (Recensement 2013 : 31264 personnes) ? Dans une promotion de cinq cents étudiants ? À partir de quelle valeur de n pouvez-vous affirmer, avec seulement 5% de chances de vous tromper : « Parmi ces n personnes, au moins une a un niveau de mercure trop élevé » ?

Corrigé : (Corrigé d'après le livre *Introduction aux probabilités et aux statistiques* de Jean-François Delmas.)

(1) On utilise les fonctions de répartitions. Pour tout $y \in \mathbb{R}$, on a

$$\mathbb{P}(Z_n \leq y) = \mathbb{P}\left(X_1 \leq n^{1/\lambda} y\right)^n.$$

Si $y \leq 0$, $\mathbb{P}\left(X_1 \leq n^{1/\lambda} y\right) \leq \mathbb{P}(X_1 \leq 0)^n \to 0$ car $\mathbb{P}(X_1 \leq 0) < 1$. Donc $\mathbb{P}(Z_n \leq y) \to 0$. Si $y \geq 0$, on a

$$\mathbb{P}(Z_n \leq y) = \left((1 - \mathbb{P}\left(X_1 > n^{1/\lambda} y\right)\right)^n.$$

Comme $\mathbb{P}\left(X_1 > n^{1/\lambda} y\right) \sim \frac{\alpha y^{-\lambda}}{n}$, un développement limité montre que $\mathbb{P}(Z_n \leq y) \to e^{-\alpha y^{-\lambda}}$. La fonction de répartition de Z_n converge donc en tout point vers celle d'une loi de Fréchet de paramètre (α, λ), d'où le résultat.

(2) On considère n individus et on note X_k le niveau annuel d'exposition au mercure de l'individu k. On suppose les variables aléatoires X_k indépendantes et de même loi que X. Pour un individu de $70kg$, la dose annuelle admissible fixée par l'OMS est $s = 18.14 \cdot 10^{-3} g$. La probabilité qu'un individu au moins ait un niveau de mercure trop élevé est

$$\begin{aligned} p_n = \mathbb{P}(\max(X_1, \ldots, X_n) > s) &= 1 - \mathbb{P}(\max(X_1, \ldots, X_n) \leq s) \\ &= 1 - e^{-n \ln(1 - \alpha s^{-\lambda}} \\ &\simeq 1 - \exp(-\alpha s^{-\lambda} n). \end{aligned}$$

Si $n_0 = 60 \cdot 10^6$ on a $p_{n_0} \simeq 1$, si $n_1 = 31264$ on a $p_{n_1} \simeq 0.97$; si $n_2 = 500$ on a $p_{n_2} \simeq 0.05$. Enfin, $1 - \exp(-\alpha s^{-\lambda} n) \geq 0.95$ si et seulement si $n \geq \frac{\ln(0.05)}{\alpha s^{-\lambda}}$, c'est-à-dire $n \geq 27216$.

□

Exercice 11. On casse un bâton de longueur 1 en n endroits choisis uniformément et indépendamment au hasard. On note L_n la longueur du plus long des $n+1$ bouts obtenus. Quel est le comportement de L_n lorsque $n \to \infty$? Le but de cet exercice est de montrer que $(n+1)L_n - \ln(n+1)$ converge en loi vers une loi de Gumbel (dont la fonction de répartition est $x \mapsto e^{e^{-x}}$).

Pour cela, soit $(X_i)_{1 \leq i \leq n+1}$ des variables aléatoires i.i.d. exponentielles de paramètre 1. On pose, pour $1 \leq i \leq n+1$,

$$S_i = X_1 + \cdots + X_i, \qquad Y_i = \frac{X_i}{S_{n+1}}.$$

(1) Déterminer la loi jointe de $(X_1, \ldots, X_n, S_{n+1})$ et en déduire celle de $(Y_1, \ldots, Y_n)$.

(2) En notant $(\Delta_1, \ldots, \Delta_{n+1})$ la longueur des bouts successifs obtenus, montrer que $(\Delta_1, \ldots, \Delta_n)$ et $(Y_1, \ldots, Y_n)$ ont la même loi (on pourra utiliser le résultat de la question (1) de l'exercice 15 de la PC 4). En déduire que $\max(Y_1, \ldots, Y_{n+1})$ a la même loi que L_n.

P. Bertail. Evaluation des risques d'exposition à un contaminant alimentaire : quelques outils statistiques. www.crest.fr/doctravail/document/2002-41.pdf (2002).

(3) Montrer que pour $x \in \mathbb{R}$, $(x+\ln(n+1))\left(\frac{S_{n+1}}{n+1}-1\right)$ converge en probabilité vers 0.

(4) En déduire le résultat désiré.

Corrigé :

(1) On commence par déterminer la loi jointe $(X_1,\dots,X_n,S_{n+1})$ en utilisant la méthode de la fonction muette Soit $f:\mathbb{R}^{n+1}\to\mathbb{R}$ une fonction continue bornée. On a d'après la formule de transfert

$$\mathbb{E}[f(X_1,\dots,X_n,S_{n+1})] = \int_{]0,\infty[^{n+1}} f(x_1,x_2,\dots,x_1+\dots+x_{n+1})e^{-(x_1+\dots+x_{n+1})}dx_1\cdots d_{x_{n+1}}.$$

En faisant le changement de variable $u_1=x_1,\dots,u_n=x_n, u_{n+1}=x_1+\dots+x_n+1$ de jacobien 1, on obtient

$$\mathbb{E}[f(X_1,\dots,X_n,S_{n+1})] = \int_{]0,\infty[^{n+1}} f(u_1,u_2,\dots,u_{n+1})e^{-u_{n+1}}\mathbb{1}_{u_{n+1}\geq u_1+\dots+u_n}du_1\cdots du_{n+1}$$

ce qui détermine la loi jointe $(X_1,\dots,X_n,S_{n+1})$.

On détermine maintenant la loi jointe $(Y_1,\dots,Y_n,S_{n+1})$. Soit $f:\mathbb{R}^{n+1}\to\mathbb{R}$ une fonction continue bornée. On a d'après la formule de transfert appliquée à $(X_1,\dots,X_n,S_{n+1})$:

$$\mathbb{E}[f(Y_1,\dots,Y_n,S_{n+1})] = \int_{]0,\infty[^{n+1}} f\left(\frac{x_1}{x_{n+1}},\dots\frac{x_n}{x_{n+1}},x_{n+1}\right)\mathbb{1}_{x_{n+1}\geq x_1+\dots+x_n}e^{-x_{n+1}}dx_1\cdots d_{x_{n+1}}.$$

En faisant le changement de variable

$$u_1=\frac{x_1}{x_{n+1}}, u_2=\frac{x_2}{x_{n+1}},\dots,u_n=\frac{x_n}{x_{n+1}}, u_{n+1}=x_{n+1},$$

avec $0\leq u_1+\dots+u_n\leq 1$ et $x_i=u_iu_{n+1}$, de jacobien u_{n+1}^n on obtient

$$\mathbb{E}[f(Y_1,\dots,Y_n,S_{n+1})] = \int_{]0,\infty[^{n+1}} f(u_1,u_2,\dots,u_{n+1})u_{n+1}^n e^{-u_{n+1}}\mathbb{1}_{u_1+\dots+u_n<1}du_1\cdots du_{n+1}.$$

On en déduit que pour toute fonction $f:\mathbb{R}^n\to\mathbb{R}$ continue bornée on a

$$\mathbb{E}[f(Y_1,\dots,Y_n)] = n!\int_{]0,\infty[^n} f(u_1,u_2,\dots,u_n)\mathbb{1}_{u_1+\dots+u_n<1}du_1\cdots du_n, \tag{1}$$

ce qui détermine la loi de $(Y_1,\dots,Y_n)$.

(2) Notons $Z_1,Z_2,\dots,Z_n$ les endroits où le bâton a été cassé, rangés dans l'ordre croissant. L'exercice 15 de la PC 4 donne une densité de $(Z_1,Z_2,\dots,Z_n)$: pour toute fonction $f:\mathbb{R}^n\to\mathbb{R}$ continue bornée, on a

$$\mathbb{E}[f(Z_1,\dots,Z_n)] = n!\int_{\{0\leq x_1<\dots<x_n\leq 1\}} f(x_1,\dots,x_n)dx_1\dots dx_n.$$

On remarque que $\Delta_1=Z_1$ et $\Delta_i=Z_i-Z_{i-1}$ pour $2\leq i\leq n$. On en déduit par changement de variable que

$$\mathbb{E}[f(\Delta_1,\Delta_2,\dots,\Delta_n)] = n!\int_{]0,\infty[^n} f(u_1,u_2,\dots,u_n)\mathbb{1}_{u_1+\dots+u_n<1}du_1\cdots du_n.$$

Compte tenu de (1), on en déduit que $(Y_1,\dots,Y_n)$ et $(\Delta_1,\dots,\Delta_n)$ ont la même loi. Comme $Y_{n+1}=1-Y_1-\dots-Y_n$ et $\Delta_{n+1}=1-\Delta_1-\dots-\Delta_n$, on en conclut que $(Y_1,\dots,Y_{n+1})$ et $(\Delta_1,\dots,\Delta_{n+1})$ ont la même loi. Le fait que $\max(Y_1,\dots,Y_{n+1})$ a la même loi que L_n en découle immédiatement car $L_n=\max(\Delta_1,\dots,\Delta_{n+1})$.

(3) Soit $W_n=(x+\ln(n+1))\left(\frac{X_1+\dots+X_{n+1}}{n+1}-1\right)$. On a $\mathbb{E}[W]=0$ et $\mathrm{Var}(W)\sim\frac{\ln(n)^2}{n}\to 0$. D'après l'inégalité de Bienaymé-Tchebychev, on en déduit que $W_n\to 0$ en probabilité.

(4) Posons $Z_n = \max(X_1, \ldots, X_{n+1}) - \log(n+1)$. D'après la question (2), on a

$$\mathbb{P}((n+1)L_n - \ln(n+1) \le x) = \mathbb{P}(Z_n + W_n \le x).$$

Or un petit calcul montre que $\mathbb{P}(\max(X_1, \ldots, X_{n+1}) - \log(n+1) \le x) \to e^{e^{-x}}$ pour tout $x \in \mathbb{R}$. Donc Z_n converge en loi vers $\mathcal{G}$, où G est la loi de Gumbel de fonction de répartition $x \mapsto e^{e^{-x}}$. Or $W_n \to 0$ en probabilité, donc $Z_n + W_n$ converge en loi vers $\mathcal{G}$ d'après le lemme de Slutsky. Donc $\mathbb{P}(Z_n + W_n \le x) \to e^{e^{-x}}$. On conclut que pour tout $x \in \mathbb{R}$ on a

$$\mathbb{P}((n+1)L_n - \ln(n+1) \le x) \quad \underset{n\to\infty}{\longrightarrow} \quad e^{e^{-x}},$$

ce qui était le résultat voulu.

□

6 Pour aller plus loin (hors PC)

Exercice 12. **(Convergences jointes)** Soient $(X_n)_{n\ge1}$, $(Y_n)_{n\ge1}$ deux suites de variables aléatoires réelles, et X, Y deux variables aléatoires réelles telles que $X_n \to X$ en loi et $Y_n \to Y$ en loi.

(1) On suppose dans cette question que les variables X_n et Y_n sont indépendantes pour tout $n \ge 1$ et que les variables X et Y sont indépendantes. Montrer que $(X_n, Y_n) \to (X, Y)$ en loi.

(2) (**Lemme de Slutsky**) On suppose que $Y = a$ est constante. Montrer que $(X_n, Y_n) \to (X, a)$ en loi.
Indications. On pourra utiliser le fait $Y_n \to Y$ en probabilité (exercice **??***) et écrire*

$$\begin{aligned}|\mathbb{E}[F(X_n,Y_n)] - \mathbb{E}[F(X,a)]| \le |\mathbb{E}[F(X_n,a)] - \mathbb{E}[F(X,a)]| \quad &+ \quad \mathbb{E}\left[|F(X_n,Y_n) - F(X_n,a)|\mathbb{1}_{\{|Y_n-a|\ge\epsilon\}}\right] \\ &+ \quad \mathbb{E}\left[|F(X_n,Y_n) - F(X_n,a)|\mathbb{1}_{\{|Y_n-a|<\epsilon\}}\right].\end{aligned}$$

On admettra également que si Z_n est une suite de variables aléatoires à valeurs dans $\mathbb{R}^k$, alors Z_n converge en loi vers Z si et seulement si pour toute fonction lipschitzienne bornée $f : \mathbb{R}^k \to \mathbb{R}$ on a $\mathbb{E}[f(Z_n)] \to \mathbb{E}[f(Z)]$.

(3) Est-il toujours vrai que $(X_n, Y_n) \to (X, Y)$ en loi ?

Corrigé :

(1) D'après le théorème de Lévy, il suffit de montrer que $\phi_{(X_n,Y_n)}(t,t') \to \phi_{(X,Y)}(t,t')$ pour tout $(t,t') \in \mathbb{R}^2$. Par indépendance (deux fois), on a

$$\phi_{(X_n,Y_n)}(t,t') = \phi_{X_n}(t)\phi_{Y_n}(t') \quad \underset{n\to\infty}{\longrightarrow} \quad \phi_X(t)\phi_Y(t') = \phi_{(X,Y)}(t,t').$$

(2) Il suffit de montrer que $\mathbb{E}[F(X_n,Y_n)] \to \mathbb{E}[F(X,Y)]$ pour une fonction $F : \mathbb{R}^2 \to \mathbb{R}$ lipschitzienne borné. Supposons que $|F(x,y) - F(x',y')| \le L(|x-x'| + |y-y'|)$ pour tous $x, x', y, y' \in \mathbb{R}$. Soit $a \in \mathbb{R}$ tel que $Y = a$ p.s. Pour $\epsilon > 0$ fixé, suivons l'indication en majorant $|\mathbb{E}[F(X_n,Y_n)] - \mathbb{E}[F(X,a)]|$ par

$$\begin{aligned}|\mathbb{E}[F(X_n,a)] - \mathbb{E}[F(X,a)]| \quad &+ \quad \mathbb{E}\left[|F(X_n,Y_n) - F(X_n,a)|\mathbb{1}_{\{|Y_n-a|\ge\epsilon\}}\right] \\ &+ \quad \mathbb{E}\left[|F(X_n,Y_n) - F(X_n,a)|\mathbb{1}_{\{|Y_n-a|<\epsilon\}}\right].\end{aligned}$$

La fonction $x \mapsto F(x,a)$ est continue bornée donc le premier terme de cette somme tend vers 0 (car X_n converge en loi vers X). Le deuxième terme est majoré par $2\sup|F| \cdot \mathbb{P}(|Y_n - a| > \epsilon)$ qui tend vers 0 (car $Y_n \to a$ en probabilité). Pour le dernier terme, on remarque que

$$|F(X_n,Y_n) - F(X_n,a)|\mathbb{1}_{\{|Y_n-a|<\epsilon\}} \le L\epsilon.$$

Ainsi, pour n suffisamment grand,

$$|\mathbb{E}[F(X_n,Y_n)]-\mathbb{E}[F(X,a)]|\leq 3L\epsilon.$$

Le résultat désiré en découle.

(3) Il n'est pas vrai en général que $(X_n,Y_n)\to(X,Y)$ en loi. En effet, considérons les variables aléatoires $X_n = Z = Y_n$ pour tout $n\geq 1$, avec Z gaussienne centrée. La variable Z étant symétrique, on a $X_n\to -Z$ en loi. Si $(X_n,Y_n)\to(-Z,Z)$ en loi, alors $X_n+Y_n\to -Z+Z$ en loi (car la fonction $(x,y)\mapsto x+y$ est continue), c'est à dire $2Z=0$ en loi, ce qui n'est pas.

□

Exercice 13. **(Un calcul sans calcul)** Déterminer la limite de $e^{-n}\sum_{k=0}^{n}\frac{n^k}{k!}$ lorsque $n\to\infty$.
Indication. On pourra utiliser le théorème central limite.

Corrigé : On a

$$e^{-n}\sum_{k=0}^{n}\frac{n^k}{k!}=\mathbb{P}(P_1+\ldots+P_n\leq n),$$

où $P_1,\ldots,P_n$ sont des variables aléatoires de Poisson de paramètre 1 indépendantes. Et

$$\mathbb{P}(P_1+\ldots+P_n\leq n)=\mathbb{P}\left(\frac{P_1+\ldots+P_n-n}{\sqrt{n}}\leq 0\right).$$

- La variance d'une variable aléatoire de loi de Poisson de paramètre λ étant égale à λ, on a d'après le TCL,

$$\frac{P_1+\ldots+P_n-n}{\sqrt{n}}\xrightarrow[n\to\infty]{\text{loi}} N,$$

où N est une variable gaussienne (centrée réduite). Or la fonction de répartition de N est continue donc

$$\lim_{n\to\infty}e^{-n}\sum_{k=0}^{n}\frac{n^k}{k!}=\mathbb{P}(N\leq 0)=\frac{1}{2}.$$

□

Exercice 14. **(Le TCL n'est pas une convergence en probabilité)** Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires réelles indépendantes de même loi. On suppose que $\mathbb{E}[X_1^2]<\infty$, et on note $m=\mathbb{E}[X_1]$, $\sigma^2=\mathrm{Var}(X_1)$ et $Z_n=\frac{1}{\sqrt{n}}\sum_{k=1}^{n}(X_k-m)$.

(1) Rappeler la convergence en loi de la suite $(Z_n)_{n\geq 1}$.

(2) Montrer que la suite $(Z_{2n}-Z_n)_{n\geq 1}$ converge en loi vers une limite qu'on identifiera.
Indication. On pourra écrire $Z_{2n}-Z_n=aZ_n+bZ_n'$ pour $a,b\in\mathbb{R}$ choisis de sorte Z_n et Z_n' soient indépendantes et de même loi.

(3) En déduire que si $\sigma^2>0$ alors la suite $(Z_n)_{n\geq 1}$ ne converge pas en probabilité.

Corrigé :

(1) D'après le TCL, Z_n converge en loi vers σN, où N est une variable aléatoire gaussienne centrée réduite.

(2) On a

$$Z_{2n} - Z_n = \left(\frac{1}{\sqrt{2}} - 1\right) Z_n + \frac{1}{\sqrt{2}} Z_n' \qquad \text{avec} \qquad Z_n' = \frac{1}{\sqrt{n}} \sum_{k=n+1}^{2n} (X_k - m).$$

Comme Z_n' est indépendant de Z_n et a la même loi que Z_n, on en déduit que

$$\begin{aligned}\phi_{Z_{2n}-Z_n}(t) = \phi_{Z_n}\left(\left(\frac{1}{\sqrt{2}} - 1\right)u\right) \cdot \phi_{Z_n}\left(\frac{1}{\sqrt{2}}u\right) \quad &\longrightarrow \quad \phi_{\sigma N}\left(\left(\frac{1}{\sqrt{2}} - 1\right)u\right) \cdot \phi_{\sigma N}\left(\frac{1}{\sqrt{2}}u\right) \\ &= \quad \exp\left(-\frac{u^2}{2}\sigma^2\left(\left(\frac{1}{\sqrt{2}} - 1\right)^2 + \frac{1}{2}\right)\right).\end{aligned}$$

Donc $Z_{2n} - Z_n$ converge en loi vers $\sigma\sqrt{2}\sqrt{1 - \frac{1}{\sqrt{2}}}N$.

(3) Soit $\sigma^2 > 0$ et supposons par l'absurde que Z_n converge en probabilité vers 0. Alors la suite $(Z_{2n} - Z_n)$ converge en probabilité vers 0 (car alors $(Z_n, Z_{2n}) \to (0,0)$ en probabilité, qu'on compose par l'application continue $f(x,y) = x - y$). On déduit de la question précédente que $\sigma = 0$, absurde.

□

Exercice 15. **(Cauchy)** On rappelle qu'une loi de Cauchy a pour densité par rapport à la mesure de Lebesgue $x \mapsto \frac{1}{\pi(1+x^2)}$. On peut démontrer (en utilisant les transformées de Fourier) que sa fonction caractéristique est $\phi(t) = \exp(-|t|)$. Soient $(X_i)_{1 \le i \le n}$ des variables aléatoires indépendantes qui suivent une loi de Cauchy. On pose $S_n = X_1 + \cdots + X_n$. Étudier les convergences en loi et en probabilité des suites suivantes.

$$(1) \quad \left(\frac{S_n}{\sqrt{n}}\right)_{n\ge1} \qquad\qquad (2) \quad \left(\frac{S_n}{n^2}\right)_{n\ge1} \qquad\qquad (3) \quad \left(\frac{S_n}{n}\right)_{n\ge1}.$$

Dans le cas (3), la loi des grands nombres s'applique-t-elle ?

Indication. Pour la troisième suite, on pourra déterminer la loi de $\frac{S_{2n}}{2n} - \frac{S_n}{n}$, raisonner par l'absurde et montrer qu'alors la suite $\frac{S_{2n}}{2n} - \frac{S_n}{n}$ converge en probabilité vers 0.

Corrigé : Rappelons que la fonction caractéristique d'une loi de Cauchy est $\phi(t) = e^{-|t|}$.

(1) Soit $t \in \mathbb{R}$. On a, par indépendance,

$$\phi_{\frac{S_n}{\sqrt{n}}}(t) = \mathbb{E}\left[\prod_{k=1}^{n} e^{i\frac{t}{\sqrt{n}}X_k}\right] = \prod_{k=1}^{n} \mathbb{E}\left[e^{i\frac{t}{\sqrt{n}}X_k}\right] = e^{-|t|\sqrt{n}} \quad \underset{n\to\infty}{\longrightarrow} \quad \begin{cases} 0 & \text{si } t \neq 0 \\ 1 & \text{si } t = 0 \end{cases}$$

La fonction caractéristique de $S_n/\sqrt{n}$ converge donc ponctuellement vers une fonction qui n'est pas continue en 0. Donc $S_n/\sqrt{n}$ ne converge pas en loi (et donc pas en probabilité non plus).

(2) Comme pour (1), on a

$$\phi_{\frac{S_n}{n^2}}(t) = e^{-\frac{|t|}{n}} \quad \underset{n\to\infty}{\longrightarrow} \quad 1.$$

Donc la fonction caractéristique de S_n/n^2 converge en tout point vers celle de la fonction constante égale à 0. Donc S_n/n^2 converge en loi vers 0, et donc aussi en probabilité vers 0 d'après l'exercice **??**.

(3) Pour tout $t \in \mathbb{R}$, on calcule

$$\mathbb{E}\left[e^{it\cdot\frac{X_1+\ldots+X_n}{n}}\right] = \prod_{k=1}^{n} \mathbb{E}\left[e^{i\frac{t}{n}X_k}\right] = \prod_{i=1}^{n} e^{-\frac{|t|}{n}} = e^{-|t|}.$$

Donc $\frac{X_1+\cdots+X_n}{n}$ suit une loi de Cauchy. Donc S_n/n converge en loi vers une loi de Cauchy.

Montrons maintenant que S_n/n ne converge pas en probabilité. On a

$$\frac{S_{2n}}{2n}-\frac{S_n}{n}=\frac{X_{n+1}+\cdots+X_{2n}}{2n}-\frac{X_1+\cdots+X_n}{2n}.$$

On a en utilisant le fait que $\frac{X_{n+1}+\cdots+X_{2n}}{2n}$ et $\frac{X_1+\cdots+X_n}{2n}$ sont indépendants et ont même loi que $S_n/2n$:

$$\phi_{\frac{S_{2n}}{2n}-\frac{S_n}{n}}=\phi_{\frac{S_n}{2n}}(t)^2=e^{-|t|}.$$

Donc $\frac{S_{2n}}{2n}-\frac{S_n}{n}$ suit une loi de Cauchy. Raisonnons maintenant par l'absurde en supposant que S_n/n converge en probabilité vers une limite X. On a alors, pour tout $\epsilon>0$,

$$\left\{\left|\frac{S_{2n}}{2n}-\frac{S_n}{n}\right|\geq\epsilon\right\}\subset\left\{\left|\frac{S_{2n}}{2n}-X\right|\geq\epsilon/2\right\}\cup\left\{\left|\frac{S_n}{n}-X\right|\geq\epsilon/2\right\},$$

et donc

$$\mathbb{P}\left(\left|\frac{S_{2n}}{2n}-\frac{S_n}{n}\right|\geq\epsilon\right)\leq\mathbb{P}\left(\left|\frac{S_{2n}}{2n}-X\right|\geq\epsilon/2\right)+\mathbb{P}\left(\left|\frac{S_n}{n}-X\right|\geq\epsilon/2\right).$$

Par passage à la limite,

$$\mathbb{P}\left(\left|\frac{S_{2n}}{2n}-\frac{S_n}{n}\right|\geq\epsilon\right)\underset{n\to\infty}{\longrightarrow}0.$$

La suite $\frac{S_{2n}}{2n}-\frac{S_n}{n}$ converge donc en probabilité en 0, et donc en loi vers 0, ce qui est absurde car $\frac{S_{2n}}{2n}-\frac{S_n}{n}$ suit une loi de Cauchy.

La loi des grands nombres ne s'applique pas car une loi de Cauchy n'est pas intégrable.

□

Exercice 16. (**Loi faible, non forte**) Soit $(X_n)_{n\geq1}$ une suite de variables aléatoires indépendantes de loi donnée par $\mathbb{P}(X_n=0)=1-\frac{1}{n\ln(n+1)}$ et $\mathbb{P}(X_n=n)=\mathbb{P}(X_n=-n)=\frac{1}{2\ln(n+1)n}$. On pose $Y_n=\frac{X_1+\cdots+X_n}{n}$.

(1) Montrer que Y_n converge en probabilité vers 0. *Indication.* On pourra étudier $\mathbb{E}[Y_n^2]$.

(2) Montrer que presque sûrement, Y_n ne converge pas.

Corrigé :

(1) On a

$$\mathbb{E}[Y_n^2]=\frac{\sum_{k=1}^n\mathbb{E}\left[X_k^2\right]}{n^2}=\frac{1}{n^2}\cdot\sum_{k=1}^n\frac{k}{2\ln(k+1)}.$$

Pour montrer que ceci converge vers 0, on fixe $M>0$ et on écrit

$$n^2\mathbb{E}[Y_n^2]=\sum_{k=1}^M\frac{k}{2\ln(k+1)}+\sum_{k=M}^n\frac{k}{2\ln(k+1)}\leq\sum_{k=1}^M\frac{k}{2\ln(k+1)}+\frac{1}{2\ln(M+1)}\sum_{k=1}^n k.$$

Donc, pour tout $M>0$, pour tout n assez grand $\mathbb{E}[Y_n^2]\leq\frac{1}{\ln(M)}$. Donc $\mathbb{E}[Y_n^2]\to0$. Ainsi, pour tout $\epsilon>0$, $\mathbb{P}(|Y_n|>\epsilon)\leq\frac{1}{\epsilon^2}\mathbb{E}[Y_n^2]\to0$.

(2) On a $\sum_{n\geq2}\mathbb{P}(X_n=n)=\infty$. Donc, d'après le (deuxième) lemme de Borel–Cantelli (les évènements $\{X_n=n\}$ sont indépendants car les variables aléatoires $(X_i)_{i\geq1}$ sont indépendantes), p.s. $X_n=n$ une infinité de fois. Or si Y_n converge, alors $X_n/n\to0$ p.s. (voir l'exercice 17 de la PC 5). Donc p.s. Y_n diverge.

□

Exercice 17. **(Problème des moments)** Soient X et Y deux variables aléatoires réelles bornées. On suppose que $\mathbb{E}[X^n] = \mathbb{E}[Y^n]$ pour tout $n \geq 1$. Montrer que X et Y ont la même loi.

Indication. Utiliser le théorème de Weierstrass sur la densité des polynômes.

Corrigé : Soit $f : \mathbb{R} \to \mathbb{R}$ une fonction continue bornée. Il s'agit de montrer que $\mathbb{E}[f(X)] = \mathbb{E}[f(Y)]$. Soient $a < b$ tels que $\mathbb{P}(a < X < b) = \mathbb{P}(a < Y < b) = 1$. On peut supposer que f est à support compact dans $[a, b]$. Il existe alors une suite $(P_n)_{n\geq 1}$ de polynômes à coefficients réels tels que P_n converge uniformément vers f sur $[a, b]$. En particulier, il existe une constante $C > 0$ telle que $\sup_{[a,b]} |P_n| \leq C$ pour tout $n \geq 1$.

Posons $X_n = P_n(X)$. Alors X_n converge presque sûrement vers $f(X)$ (car la convergence uniforme implique la convergence simple), et $|X_n| \leq C$ pour tout $n \geq 1$. D'après le théorème de convergence dominée, on en déduit que $\mathbb{E}[X_n] = \mathbb{E}[P_n(X)] \to \mathbb{E}[f(X)]$. On montre de même que $\mathbb{E}[P_n(Y)] \to \mathbb{E}[f(Y)]$.

Or, par hypothèse, $\mathbb{E}[P_n(X)] = \mathbb{E}[P_n(Y)]$. Ceci conclut.

□

Exercice 18. **(Sommes aléatoires)** Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires réelles indépendantes de même loi, centrées, de variance finie $\sigma^2 > 0$. On pose $S_n = \sum_{m=1}^{n} X_m$. Soit $(N_k)_{k\geq 1}$ une suite de variables aléatoires à valeurs dans $\mathbb{N}_*$, toutes indépendantes de la suite $(X_n)_{n\geq 1}$. On pose finalement $Z_k = \frac{1}{\sqrt{N_k}} S_{N_k}$. On suppose que $N_k \to \infty$ p.s. lorsque $k \to \infty$. Montrer que Z_k converge en loi vers une variable aléatoire que l'on déterminera.

Corrigé : Posons $Y_n = S_n/\sqrt{n}$ et soit $\mathcal{N}$ une variable aléatoire normale centrée réduite. Nous allons montrer que Z_k converge en loi vers $\mathcal{N}$. Soit $F : \mathbb{R} \to \mathbb{R}_+$ une fonction continue bornée. Fixons $\epsilon > 0$. D'après le théorème central limite, Y_n converge en loi vers $\mathcal{N}$. Il existe donc N tel que pour $n \geq n_0$:

$$|\mathbb{E}[F(Y_n)] - \mathbb{E}[F(\mathcal{N})]| \leq \epsilon.$$

On écrit ensuite :

$$\begin{aligned}
\mathbb{E}[F(Z_k)] &= \mathbb{E}\left[F(Y_{N_k})\right] = \sum_{j=1}^{\infty} \mathbb{E}\left[F(Y_j)\mathbb{1}_{\{N_k=j\}}\right] \\
&= \sum_{j=1}^{\infty} \mathbb{P}(N_k = j)\,\mathbb{E}\left[F(Y_j)\right] \qquad \text{par indépendance de } N_k \text{ et } Y_j
\end{aligned}$$

Ainsi :

$$\begin{aligned}
|\mathbb{E}[F(Z_k)] - \mathbb{E}[F(\mathcal{N})]| &\leq \sum_{j=1}^{\infty} \mathbb{P}(N_k = j)\left|\mathbb{E}\left[F(Y_j)\right] - \mathbb{E}[F(\mathcal{N})]\right| \\
&\leq 2\|F\|_\infty \mathbb{P}(N_k \leq n_0) + \sum_{j\geq n_0} \mathbb{P}(N_k = j)\left|\mathbb{E}\left[F(Y_j)\right] - \mathbb{E}[F(\mathcal{N})]\right| \\
&\leq 2\|F\|_\infty \mathbb{P}(N_k \leq n_0) + \sum_{j\geq n_0} \mathbb{P}(N_k = j)\,\epsilon \\
&\leq 2\|F\|_\infty \mathbb{P}(N_k \leq n_0) + \epsilon.
\end{aligned}$$

Or N_k converge presque sûrement vers 0 et donc en probabilité. Le premier terme de la dernière somme converge donc vers 0. On en déduit que $|\mathbb{E}[F(Z_k)] - \mathbb{E}[F(\mathcal{N})]| \leq 2\epsilon$ pour k suffisamment grand, ce qui conclut. □