14 DTD : LGN, TCL

Lois des grands nombres

Exercice 14.1 (Sans calcul)**.** Déterminer les limites suivantes :

1. $\lim_{n\to\infty} \int_{[0,1]^n} f\left(\frac{x_1+\ldots+x_n}{n}\right)\, dx_1 \ldots dx_n$ pour f une fonction continue sur $[0,1]$,

2. $\lim_{n\to\infty} \sum_{k=0}^{n} C_n^k p^k (1-p)^{n-k} f\left(\frac{k}{n}\right)$ pour f une fonction continue sur $[0,1]$ et $p \in [0,1]$,

3. $\lim_{n\to\infty} \sum_{k=0}^{\infty} e^{-\lambda n} \frac{(\lambda n)^k}{k!} f\left(\frac{k}{n}\right)$ pour f une fonction continue et bornée sur $\mathbb{R}_+$ et $\lambda > 0$.

Correction :

1. On a
$$\int_{[0,1]^n} f\left(\frac{x_1+\ldots+x_n}{n}\right)\, dx_1 \ldots dx_n = \mathbb{E}\left(f\left(\frac{U_1+\ldots+U_n}{n}\right)\right)$$
où $U_1, \ldots, U_n$ sont des variables aléatoires indépendantes et uniformes sur $[0,1]$. D'après la loi forte des grands nombres, $n^{-1}(U_1+\ldots+U_n)$ converge p.s. et donc en loi vers $1/2$. Ainsi,
$$\lim_{n\to\infty} \int_{[0,1]^n} dx_1, \ldots dx_n\, f\left(\frac{x_1+\ldots+x_n}{n}\right) = f\left(\frac{1}{2}\right).$$

2. On a
$$\sum_{k=0}^{n} C_n^k p^k (1-p)^{n-k} f\left(\frac{k}{n}\right) = \mathbb{E}\left(f\left(\frac{B_1+\ldots+B_n}{n}\right)\right),$$
où $B_1, \ldots, B_n$ sont des variables aléatoires de Bernoulli de paramètre p indépendantes. D'après la loi forte des grands nombres, $n^{-1}(B_1+\ldots+B_n)$ converge p.s. et donc en loi vers p. Ainsi,
$$\lim_{n\to\infty} \sum_{k=0}^{n} C_n^k p^k (1-p)^{n-k} f\left(\frac{k}{n}\right) = f(p).$$

3. On a
$$\sum_{k=0}^{\infty} e^{-\lambda n} \frac{(\lambda n)^k}{k!} f\left(\frac{k}{n}\right) = \mathbb{E}\left(f\left(\frac{P_1+\ldots+P_n}{n}\right)\right),$$
où $P_1, \ldots, P_n$ sont des variables aléatoires de Poisson de paramètre λ indépendantes. D'après la loi forte des grands nombres, $n^{-1}(P_1+\ldots+P_n)$ converge p.s. et donc en loi vers λ. Ainsi,
$$\lim_{n\to\infty} \sum_{k=0}^{\infty} e^{-\lambda n} \frac{(\lambda n)^k}{k!} f\left(\frac{k}{n}\right) = f(\lambda).$$

Exercice 14.2 (Théorème de Bernstein-Weierstrass)**.** Soit f une fonction continue de $[0,1]$ dans $\mathbb{C}$. Le n-ième polynôme de Bernstein de f est

$$B_n(x) : x \mapsto \sum_{k=0}^{n} \mathrm{C}_n^k x^k (1-x)^{n-k} f\left(\frac{k}{n}\right).$$

1. Soit $S_n(x) := \frac{\mathrm{Bin}(n,x)}{n}$ où $\mathrm{Bin}(n,x)$ est une v.a. de loi binomiale de paramètre n et x. Montrer que $B_n(x) = \mathbb{E}[f(S_n(x))]$.

2. En déduire le Théorème de Berstein-Weierstrass

$$\|B_n - f\|_\infty \longrightarrow 0.$$

Correction :

1. C'est évident.

2. Soit $\varepsilon > 0$, soit $\eta > 0$ le module d'uniforme continuité de f associé à ε. Alors on a

$$\begin{aligned} |f(x) - B_n(x)| &= \mathbb{E}[|f(x) - f(S_n(x))|] \\ &\leq \mathbb{E}[|f(x) - f(S_n(x))|\mathbf{1}_{|S_n(x)-x|\leq\eta}] + \mathbb{E}[|f(x) - f(S_n(x))|\mathbf{1}_{|S_n(x)-x|\geq\eta}] \\ &\leq \varepsilon + 2\|f\|_\infty \mathbb{P}[|S_n(x) - x| \geq \eta] \end{aligned}$$

Pour évaluer $\mathbb{P}[|S_n(x) - x| \geq \eta]$, on utilise l'inégalité de Markov :

$$\mathbb{P}[|S_n(x) - x| \geq \eta] \leq \frac{\mathrm{Var}[|S_n(x)|]}{\eta^2} = \frac{x(1-x)}{n\eta^2} \leq \frac{1}{2n\eta^2}.$$

La majoration ci-dessus est uniforme en x, ce qui permet de conclure.

Exercice 14.3 (Loi faible, non forte)**.** Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires indépendantes de loi

$$\mathbb{P}_{X_n} = \frac{1}{2\ln(n+1)n}(\delta_n + \delta_{-n}) + \left(1 - \frac{1}{n\ln(n+1)}\right)\delta_0.$$

1. Montrer que $Y_n := \frac{X_1 + \ldots + X_n}{n}$ converge en probabilité vers 0.

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

Correction :

1. En montre (en calculant) que

$$\mathbb{E}[(Y_n)^2] \longrightarrow 0.$$

L'inégalité de Markov permet alors de conclure.

2. À l'aide de Borel-Cantelli, on montre que $\mathbb{P}\left\{\frac{X_n}{n \geq 1}, \text{ pour une infinité de } n\right\} = 1$, et l'on conclue comme dans l'exercice 12.2.

Exercice 14.4 (Une LGN avec un goût de TCL)**.** On rappelle que la loi de Cauchy à pour densité par rapport à la mesure de Lebesgue $f(x) = \frac{1}{\pi(1+x^2)}$, et pour fonction caractéristique

$$\phi(\xi) = \exp(-|\xi|).$$

1. Si X et Y sont deux variables de Cauchy indépendantes. Quelles est la loi de $\frac{X+Y}{2}$?
2. Si $X_1, ..., X_n$ sont des vaiid de loi de Cauchy. Quelle est la loi de

$$\frac{X_1 + ... + X_n}{n} ?$$

   Qu'en pensez-vous ?
3. Pour les plus courageux : démontrer à l'aide de la densité de la loi de Cauchy, la forme de la fonction caractéristique.

14.1 Théorème Central Limite

Exercice 14.5. Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires réelles indépendantes et de même loi telle que $\mathbb{E}(X_1) = 0$ et $\mathbb{E}(X_1^2) = 1$. On pose $S_n = X_1 + \ldots + X_n$ pour tout $n \geq 1$.

1. Montrer que pour tout $A > 0$, on a

$$\mathbb{P}\left(\limsup_{n\to\infty} \frac{S_n}{\sqrt{n}} \geq A\right) = 1,$$

   et en déduire que

$$\mathbb{P}\left(\limsup_{n\to\infty} \frac{S_n}{\sqrt{n}} = +\infty\right) = 1.$$

2. Justifier que si $(S_{n_k})_{k\geq 1}$ est une suite extraite de $(S_n)_{n\geq 1}$, alors on a :

$$\mathbb{P}\left(\limsup_{k\to\infty} \frac{S_{n_k}}{\sqrt{n_k}} = +\infty\right) = 1.$$

3. En déduire que la suite $(n^{-1/2}S_n)_{n\geq 1}$ ne converge pas en probabilité.

Correction :

1. D'après le TCL, la suite $(n^{-1/2}S_n)_{n\geq 1}$ converge en loi vers une variable aléatoire N de loi gaussienne centrée réduite. Soit $A > 0$. On a

$$\mathbb{P}\left(\limsup_{n\to\infty} \frac{S_n}{\sqrt{n}} \geq A\right) \geq \mathbb{P}\left(\limsup\left\{\frac{S_n}{\sqrt{n}} > A\right\}\right) \geq \limsup_{n\to\infty} \mathbb{P}\left(\frac{S_n}{\sqrt{n}} > A\right).$$

   La variable N étant à densité, on a

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{S_n}{\sqrt{n}} > A\right) = \mathbb{P}(N > A) > 0,$$

ce qui implique que $\mathbb{P}\left(\limsup_{n\to\infty} \frac{S_n}{\sqrt{n}} \geq A\right) > 0$. Or

$$\left\{\limsup_{n\to\infty} \frac{S_n}{\sqrt{n}} \geq A\right\} \in \mathcal{A}_\infty,$$

où $\mathcal{A}_\infty$ est la tribu asymptotique $\bigcap_{n\geq 1} \sigma(X_n, X_{n+1}, \ldots)$. Ainsi

$$\mathbb{P}\left(\limsup \frac{S_n}{\sqrt{n}} \geq A\right) = 1.$$

En considérant une suite $(A_k)_{k\geq 1} \uparrow +\infty$, on en déduit le résultat.

2. On démontre ce résultat de la même manière que le résultat précédent.

3. Si la suite $(n^{-1/2}S_n)_{n\geq 1}$ converge en probabilité vers une variable aléatoire X, alors on peut extraire une sous-suite $({n_k}^{-1/2}S_{n_k})_{k\geq 1}$ telle que ${n_k}^{-1/2}S_{n_k} \to X$ p.s. Or X a la même loi que N, ce qui contredit le fait que

$$\mathbb{P}\left(\limsup_{k\to\infty} \frac{S_{n_k}}{\sqrt{n_k}} = +\infty\right) = 1.$$

Exercice 14.6 (Formule de Stirling)**.** Question préliminaire : Soit X une variable aléatoire réelle de carré intégrable définie sur $(\Omega, \mathcal{A}, \mathbb{P})$. Montrer que, pour tout $a > 0$, on a :

$$\mathbb{E}(|X - \inf(X, a)|) \leq \left(\mathbb{E}(X^2)\mathbb{P}(X \geq a)\right)^{1/2}.$$

Soit $(X_n, n \geq 1)$ une suite de variables aléatoires indépendantes définies sur $(\Omega, \mathcal{A}, \mathbb{P})$, de même loi de Poisson de paramètre 1. On pose, pour tout entier $n \geq 1$,

$$S_n = \sum_{i=1}^n X_i \ \text{ et } \ Y_n = \frac{S_n - n}{\sqrt{n}}.$$

On note $x^- = \sup(-x, 0)$ pour tout $x \in \mathbb{R}$.

1. Pour tout $n \geq 1$, vérifier que S_n suit la loi de Poisson de paramètre n, calculer $\mathbb{E}(Y_n^2)$ et en déduire que pour tout $a > 0$,

$$\mathbb{P}(Y_n^- \geq a) \leq \frac{1}{a^2}.$$

2. Soit Y une variable aléatoire de loi normale $\mathcal{N}(0, 1)$. Montrer que la suite $(Y_n^-)_{n\geq 1}$ converge en loi vers Y^-.

3. Montrer à l'aide de la question préliminaire que

$$\mathbb{E}(Y_n^-) \underset{n\to\infty}{\longrightarrow} \mathbb{E}(Y^-).$$

4. En déduire la formule de Stirling

$$n! \underset{n\to\infty}{\sim} \left(\frac{n}{e}\right)^n \sqrt{2\pi n}.$$

Correction :

1. On a, d'après l'inégalité de Cauchy-Schwarz,

$$\mathbb{E}(|X - \inf(X,a)|) = \mathbb{E}\left((X-a)\mathbb{1}_{\{X>a\}}\right) \le \mathbb{E}\left(X\mathbb{1}_{\{X>a\}}\right) \le \left(\mathbb{E}(X^2)\mathbb{P}(X>a)\right)^{1/2}.$$

(2) (a) On a

$$\mathbb{E}(Y_n^2) = \frac{1}{n}\sum_{i=1}^{n} \mathrm{Var}(X_n) = 1.$$

On en déduit que

$$\mathbb{P}(Y_n^- \ge a) \le \frac{1}{a^2}\,\mathbb{E}((Y_n^-)^2) \le \frac{1}{a^2}\,\mathbb{E}(Y_n^2) = \frac{1}{a^2}.$$

(2) (b) D'après le TCL, la suite $(Y_n)_{n\ge1}$ converge en loi vers Y. Or la fonction $x \in \mathbb{R} \mapsto x^-$ est continue donc la suite $(Y_n^-)_{n\ge1}$ converge en loi vers Y^-.

(2) (c) Soit $a > 0$. La fonction $x \in \mathbb{R}_+ \mapsto \inf(x,a)$ est continue et bornée donc

$$\mathbb{E}(\inf(Y_n^-, a)) \underset{n\to\infty}{\longrightarrow} \mathbb{E}(\inf(Y^-, a)).$$

Et

$$\begin{aligned}
&\left|\mathbb{E}(Y_n^-) - \mathbb{E}(Y^-)\right| \\
&\quad\le\ \mathbb{E}(|Y_n^- - \inf(Y_n^-, a)|) + \left|\mathbb{E}(\inf(Y_n^-, a) - \inf(Y^-, a))\right| + \mathbb{E}(|\inf(Y^-, a) - Y^-|) \\
&\quad\le\ \left|\mathbb{E}(\inf(Y_n^-, a) - \inf(Y^-, a))\right| + \left(\mathbb{P}(Y_n^- \ge a)\right)^{1/2} + \left(\mathbb{P}(Y^- \ge a)\right)^{1/2},
\end{aligned}$$

d'après la question préliminaire. Or $\mathbb{P}(Y_n^- \ge a) \le a^{-2}$ et

$$\mathbb{P}(Y^- \ge a) \le \frac{\mathbb{E}((Y^-)^2)}{a^2} \le \frac{\mathbb{E}(Y^2)}{a^2} = \frac{1}{a^2}.$$

Ainsi,

$$\limsup_{n\to\infty} \left|\mathbb{E}(Y_n^-) - \mathbb{E}(Y^-)\right| \le \frac{2}{a}.$$

Ceci étant vrai pour tout $a > 0$, on a

$$\limsup_{n\to\infty} \left|\mathbb{E}(Y_n^-) - \mathbb{E}(Y^-)\right| = 0.$$

(2) (d) La variable aléatoire S_n suit une loi de Poisson de paramètre n donc

$$\mathbb{E}(Y_n^-) = \frac{1}{\sqrt{n}}\sum_{k=0}^{n}(n-k)e^{-n}\frac{n^k}{k!} = \frac{e^{-n}}{\sqrt{n}}\left(\sum_{k=0}^{n}\frac{n^{k+1}}{k!} - \sum_{k=0}^{n-1}\frac{n^{k+1}}{k!}\right) = \frac{n^{n+1}}{e^n n!\sqrt{n}} = \left(\frac{n}{e}\right)^n\frac{\sqrt{n}}{n!}.$$

Et

$$\mathbb{E}(Y^-) = \frac{1}{\sqrt{2\pi}}\int_0^{\infty} x e^{-x^2/2}\,dx = \frac{1}{\sqrt{2\pi}}.$$

Donc,

$$\left(\frac{n}{e}\right)^n \frac{\sqrt{n}}{n!} \underset{n\to\infty}{\longrightarrow} \frac{1}{\sqrt{2\pi}},$$

et donc

$$n! \underset{n\to\infty}{\sim} \left(\frac{n}{e}\right)^n \sqrt{2\pi n}.$$

Exercice 14.7 (Sans Calcul)**.** Déterminer la limite suivante :

$$\lim_{n\to\infty} e^{-n} \sum_{k=0}^{n} \frac{n^k}{k!}.$$

Correction : 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,

$$\frac{P_1 + \ldots + P_n - n}{\sqrt{n}} \underset{n\to\infty}{\overset{\text{(loi)}}{\longrightarrow}} N,$$

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 14.8. Soient X une variable aléatoire réelle de loi normale $\mathcal{N}(0,1)$ et ε une variable aléatoire réelle indépendante de X et de loi $\frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1$. On pose $Y = \varepsilon X$.

1. Montrer que Y est de loi normale $\mathcal{N}(0,1)$ et calculer $\mathrm{Cov}(X,Y)$.
2. Les variables X et Y sont-elles indépendantes ? Le vecteur (X,Y) est-il gaussien ?

Correction :

1. La loi $\mathcal{N}(0,1)$ est symétrique donc Y suit la loi $\mathcal{N}(0,1)$. On a

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

2. D'après la question (1), le vecteur (X,Y) est gaussien si et seulement si les variables X et Y sont indépendantes. Or

   $$1 = \mathbb{E}(X^2)\mathbb{E}(Y^2) \neq \mathbb{E}(X^2Y^2) = \mathbb{E}(X^4) = 3.$$

   Donc X et Y ne sont pas indépendantes et (X,Y) n'est pas gaussien.

Exercice 14.9. Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires à valeurs dans $\mathbb{R}^2$ indépendantes et indentiquement distribuées. Étudier le comportement asymptotique de la suite

$$\left(\frac{X_1+\ldots+X_n-n\mathbb{E}(X_1)}{\sqrt{n}}\right)_{n\geq 1}$$

dans les cas suivants :

1. X_1 est de loi $\frac{1}{2}\delta_{(-1,-1)}+\frac{1}{2}\delta_{(1,1)}$,
2. X_1 est de loi $\frac{1}{2}\delta_{(-1,-1)}+\frac{1}{4}\delta_{(1,-1)}+\frac{1}{4}\delta_{(1,1)}$.

Exercice 14.10. Soient $(U_k)_{k\geq 1}$ une suite de variables aléatoires réelles indépendantes et de loi normale $\mathcal{N}(0,\sigma^2)$ et $\theta\in\mathbb{R}$. On définit la suite $(X_k)_{k\geq 1}$ par $X_1=U_1$ et $X_k=\theta U_{k-1}+U_k$. Montrer que pour tout $n\geq 1$, le vecteur $(X_1,\ldots,X_n)$ est un vecteur gaussien non dégénéré dont on précisera la densité, l'espérance et la matrice de covariance.

Correction : Soit A la matrice carré d'ordre n définie par

$$A=\begin{pmatrix} 1 & 0 & 0 & 0 & \ldots \\ \theta & 1 & 0 & 0 & \ldots \\ 0 & \theta & 1 & 0 & \ldots \\ \vdots & \ddots & \ddots & \ddots & \\ & \ldots & 0 & \theta & 1 \end{pmatrix}$$

Alors on voit que $X=AU$ en notant $X=(X_1,\ldots,X_n)$ et $U=(U_1,\ldots,U_n)$. Donc X est un vecteur gaussien. De plus $\det(A)=1$ donc X est non dégénéré. On a $\mathbb{E}(X)=0$ et la matrice de covariance de X est donnée par

$$C_X=A(\sigma^2 I_n)\,{}^tA=\sigma^2 A\,{}^tA=\sigma^2\begin{pmatrix} 1 & \theta & & & \\ \theta & 1+\theta^2 & \theta & & \\ & \ddots & \ddots & \ddots & \\ & & \theta & 1+\theta^2 & \theta \\ & & & \theta & 1+\theta^2 \end{pmatrix}.$$

Enfin la densité de X est

$$f_X(x)=\frac{1}{\sqrt{(2\pi)^n\det(C_X)}}\,e^{-\,{}^txC_X^{-1}x}=\frac{1}{(2\pi\sigma^2)^{n/2}}e^{-\,{}^txC_X^{-1}x},\;\; x\in\mathbb{R}^n. \qquad \square$$

Exercice 14.11. Soit $(X_n,n\geq 0)$ une suite de variables aléatoires réelles définies sur $(\Omega,\mathcal{A},\mathbb{P})$ suivant respectivement une loi gaussienne $\mathcal{N}(m_n,\sigma_n^2)$ avec $m_n\in\mathbb{R}$ et $\sigma_n>0$. Montrer que cette suite converge en loi vers une variable réelle Y si et seulement si les deux suites $(m_n,n\geq 0)$ et $(\sigma_n,n\geq 0)$ convergent, et identifier la loi limite.

Correction : Supposons que $m_n\to m$ et $\sigma_n\to\sigma$. Alors pour tout $t\in\mathbb{R}$,

$$\Phi_{Y_n}(t)=\exp\left(-\frac{\sigma_n^2t^2}{2}+im_nt\right)\underset{n\to\infty}{\longrightarrow}\exp\left(-\frac{\sigma^2t^2}{2}+imt\right)=\Phi_Y(t),$$

où Y est une variable aléatoire de loi gaussienne $\mathcal{N}(m,\sigma)$ (si $\sigma=0$, alors $\mathcal{N}(m,\sigma)=\delta_m$). Ainsi, d'après le théorème de Lévy, $Y_n \to Y$ en loi. Réciproquement, si Y_n converge en loi vers une variable aléatoire Y, alors $|\Phi_{Y_n}(t)|$ converge pour tout $t \in \mathbb{R}$. En particulier pour $t=1$, on a

$$|\Phi_{Y_n}(1)| = \exp\left(-\frac{\sigma_n^2}{2}\right) \in]0,1].$$

On en déduit que σ_n converge vers un certain $\sigma \in \mathbb{R}_+$. En effet, si $\sigma_n \to \infty$ alors $\Phi_Y(t)=0$ pour tout $t \in \mathbb{R}$ ce qui est impossible. Ainsi, $+\infty$ et $-\infty$ ne peuvent être valeurs d'adhérence de $(m_n)_{n\geq 1}$. En effet supposons, quitte à extraire, que $m_n \to \infty$. Alors on a pour $a \in \mathbb{R}$,

$$\begin{aligned}
P(Y<a) &\leq \liminf_{n\to\infty} P(Y_n<a)\\
&= \liminf_{n\to\infty} \int_{-\infty}^{a-m_n} \exp\left(-\frac{x^2}{2\sigma_n^2}\right)\frac{dx}{\sqrt{2\pi\sigma_n^2}}\\
&= \liminf_{n\to\infty} \int_{-\infty}^{\sigma_n^{-1}(a-m_n)} \exp\left(-\frac{x^2}{2}\right)\frac{dx}{\sqrt{2\pi}}\\
&= 0.
\end{aligned}$$

Ainsi $P(Y\in\mathbb{R})=0$, ce qui est absurde. Donc $(m_n)_{n\geq 1}$ est bornée, et a donc une valeur d'adhérence. Soient m, m' deux valeurs d'adhérence de cette suite. Alors par passage à la limite dans la fonction caractéristique et en simplifiant par le module, on a $e^{imt}=e^{im't}$ pour tout $t\in\mathbb{R}$. Les dérivées en 0 sont donc égales, et $m=m'$. Ceci permet de conclure.

14.2 Physionomie

Exercice 14.12. Comment sont obtenues ces images ?
(Les deux premières sont dues à V. Beffara)