Correction du TD 1

Exercice 1

1. On dit que $(X_n)_{n\geq 1}$ converge vers X (l'espace sous-jacent étant $(\Omega, \mathcal{F}, \mathbb{P})$) :
   — en probabilité ($X_n \xrightarrow{\mathbb{P}} X$), si $\forall \varepsilon > 0$ on a $\mathbb{P}(|X_n - X| \geq \varepsilon) \underset{n\to\infty}{\to} 0$.
   — en norme $\mathbb{L}^2$ ($X_n \xrightarrow{\mathbb{L}^2} X$), si $\mathbb{E}(|X_n - X|^2) \underset{n\to\infty}{\to} 0$.
   — en loi ($X_n \rightsquigarrow X$ ou $X_n \xrightarrow{\mathcal{L}} X$), si $F_n(x) \underset{n\to\infty}{\to} F(x)$, pour tout x tel que F est continue en x ; F_n désignant la fonction de répartition (f.d.r.) de la v.a. X_n et F désignant celle de la v.a. X. (Il existe d'autres caractérisations de la convergence en loi.)
   On peut également utiliser la définition suivante : $X_n \xrightarrow{\mathcal{L}} X$ si pour toute fonction continue bornée φ de $\mathbb{R}$ dans $\mathbb{R}$, $\lim_n \mathbb{E}\left[\varphi(X_n)\right] = \mathbb{E}\left[\varphi(X)\right]$.
   — presque sûrement ($X_n \xrightarrow{p.s.} X$), si $\exists \Omega_0 \in \mathcal{F}$, t.q. $\mathbb{P}(\Omega_0) = 1$ et $\forall \omega \in \Omega_0$, on a $X_n(\omega) \underset{n\to\infty}{\to} X(\omega)$.
2. En utilisant le corollaire de l'inégalité de Markov avec la fonction carrée (croissante et positive sur $\mathbb{R}_+$), on obtient pour tout $\varepsilon > 0$
$$\mathbb{P}(|X_n - X| \geq \varepsilon) \leq \frac{\mathbb{E}|X_n - X|^2}{\varepsilon^2} \underset{n\to+\infty}{\longrightarrow} 0$$
   qui tend vers 0 par convergence L^2. On en déduit la convergence en probabilité.
3. Si $(X_n)_n$ converge en probabilité vers a alors $(X_n)_n$ converge en loi vers a, c'est-à-dire que pour tout fonction f continue bornée on a $\mathbb{E}f(X_n) \underset{n+\infty}{\to} \mathbb{E}f(a) = f(a)$.
   En particulier, pour toute fonction f continue bornée, h étant continue, $f \circ h$ est continue bornée. Pour toute fonction f continue bornée, on a donc $\mathbb{E}\left[f(h(X_n))\right] \underset{n+\infty}{\to} \mathbb{E}\left[f(h(a))\right] = f(h(a))$, ce qui implique que $h(X_n)$ converge en loi vers $h(a)$.
   De plus, comme la convergence en loi vers une constante implique la convergence en probabilité vers cette constante, nous obtenons finalement que $(h(X_n))_n$ converge en probabilité vers $h(a)$.
   En fait, les convergences p.s., en probabilité et en loi « passent » toutes à la continuité (avec une limite pas forcément constante p.s. qui plus est) : ce résultat est connu en anglais sous le nom de Continuous Mapping Theorem.
4. (a) On a
$$\mathbb{E}(X_1^2) = \int_{\mathbb{R}} x^2 f(x)dx = \int_{\mathbb{R}_+} 2x^3 \exp(-x^2)dx$$
$$= \int_{\mathbb{R}_+} u \exp(-u)du = 1,$$
   en posant $u = x^2$.
   (b) On vient de montrer que $X_1^2 \in \mathbb{L}^1$, la suite (X_n^2) vérifie donc la loi forte des grands nombres (LFGN), c'est-à-dire,
$$V_n = \frac{1}{n}\sum_{i=1}^n X_i^2 \xrightarrow{p.s.} \mathbb{E}(X_1^2) = 1.$$

(c) La fonction $x \mapsto 1/x$ est continue sur $\mathbb{R}_+^*$ et $\mathbb{E}(X_1^2) = 1 \neq 0$. Par le théorème de continuité et la question précédente, on a donc

$$W_n = n / \sum_{i=1}^{n} X_i^2 \overset{p.s.}{\to} \frac{1}{\mathbb{E}(X_1^2)} = 1.$$

5. — $X_n = 1/n$ est une suite déterministe qui tend vers 0 lorsque n tend vers l'infini. Elle converge donc p.s. et donc aussi en probabilité et en loi. Elle converge également dans L^2 ;
   — $X_n = (-1)^n$ est aussi déterministe, mais ne converge pas en loi car pour f continue bornée avec $f(x) = x$ pour $|x| \leq 1$, $\mathbb{E}f(X_n) = (-1)^n$ ne converge pas. Ainsi, il n'y a pas convergence en probabilité, L^2 ou p.s. ;
   — $X_n = \mathbb{1}_{A_n}$ avec $\mathbb{P}(A_n) \to 0$. La convergence L^2 vers 0 se vérifie facilement : $\mathbb{E}X_n^2 = \mathbb{P}(A_n) \to 0$ (ce qui implique aussi la convergence en probabilité et en loi).
   En revanche, on n'a pas nécessairement la convergence p.s. Contre-exemple classique de la "bosse" qui se balade dans $[0, 1]$ avec un support qui rétrécit. On construit les ensembles A_n de la façon suivante :
     — 1ère étape : $A_1 = [0, 1]$,
     — 2ème étape : on découpe A_1 en 2 et on fait glisser les ensembles $A_2 = [0, 1/2]$, $A_3 = [1/2, 1]$,
     — 3ème étape : on recommence $A_4 = [0, 1/4]$, $A_5 = [1/4, 1/2]$, $A_6 = [1/2, 3/4]$, $A_7 = [3/4, 1]$, etc.

   avec pour espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P}) = ([0, 1], \mathcal{B}([0, 1]), \mathrm{Leb}[0, 1])$. Ainsi on a bien $\mathbb{P}(A_n) = \mathbb{E}(\mathbb{1}_{A_n}) = 1/2^{\lfloor \ln_2(n) \rfloor} \to 0$, mais $\forall \omega \in [0, 1]$, $\mathbb{1}_{A_n}(\omega)$ n'admet pas de limite. Il ne peut donc pas y avoir la convergence p.s.
   — $X_n = Z_n \mathbb{1}_{B_n}$ avec $Z_n \rightsquigarrow Z$ et $\mathbb{P}(B_n) \to 1$.
   (X_n) converge en loi vers Z. En effet, Z_n converge en loi vers Z et $\mathbb{1}_{B_n}$ converge en probabilité vers 1 :
   $$\mathbb{P}(\mathbb{1}_{B_n} \neq 1) = 1 - \mathbb{P}(B_n) \to 0$$
   soit, pour tout $1 > \varepsilon > 0$
   $$\mathbb{P}(|\mathbb{1}_{B_n} - 1| \geq \varepsilon) = \mathbb{P}(\mathbb{1}_{B_n} \neq 1) \to 0$$
   Ainsi, par le lemme de Slutsky, X_n converge en loi vers Z.
   En revanche, la convergence en probabilité ne peut pas être assurée (Z_n peut ne pas converger en probabilité). En effet, prenons $Z \sim \mathcal{N}(0, 1)$, $Z_n = (-1)^n Z$ et $B_n = \Omega$ pour tout $n \geq 1$. On a alors, pour tout $n \geq 1$, $X_n = Z_n$ et puisque $Z_n \sim \mathcal{N}(0, 1)$, $X_n \overset{\mathcal{L}}{\to} Z$. Cependant, X_n ne converge pas en probabilité : supposons qu'il existe une variable aléatoire réelle Y telle que $X_n \overset{\mathbb{P}}{\to} Y$, alors $Z_{2n} \overset{\mathbb{P}}{\to} Y$ et $Z_{2n+1} \overset{\mathbb{P}}{\to} Y$, autrement dit, $Z \overset{\mathbb{P}}{\to} Y$ et $Z \overset{\mathbb{P}}{\to} -Y$. Par unicité de la limite en probabilité (résultat admis), $Y = -Y$ *p.s.*, i.e. $Y = 0$ *p.s.*. Il vient donc $Z = 0$ *p.s.*, ce qui est absurde. Ainsi, X_n ne converge pas en probabilité, et il n'y a pas non plus convergence p.s. ni L^2.

6. La suite (X_n) converge p.s. vers X

   Il faut retenir le corollaire du lemme de Borel-Cantelli énoncé dans le cours : si pour tout $\varepsilon > 0$,
   $$\sum_{n=1}^{\infty} \mathbb{P}(|X_n - X| \geq \varepsilon) < \infty$$
   alors (X_n) converge p.s. vers X.
   La preuve n'est pas à savoir.

   Preuve à partir du lemme de Borel-Cantelli. Soit $k \geq 1$. Posons $A_n^{(k)}$ l'événement $\{|X_n - X| \geq 1/k\}$. Par hypothèse, nous avons $\sum_n \mathbb{P}(A_n^{(k)}) < +\infty$. Par application du théorème de Borel-Cantelli, nous en déduisons que

$$\mathbb{P}(\limsup_n \{|X_n - X| \geq 1/k\}) = \mathbb{P}\left(\cap_n \cup_{m \geq n} A_n^{(k)}\right) = 0.$$

En considérant l'événement complémentaire, on en déduit que

$$\mathbb{P}(\liminf_n \{|X_n - X| \leq 1/k\}) = \mathbb{P}\left(\cup_n \cap_{m \geq n} \overline{A_n^{(k)}}\right) = 1,$$

c'est-à-dire que

$$\exists \Omega_k, \quad \text{t.q.} \quad \mathbb{P}(\Omega_k) = 1, \quad \forall \omega \in \Omega_k, \quad \exists n_k \in \mathbb{N}, \quad \forall m \geq n_k, \quad |X_m(\omega) - X(\omega)| \leq 1/k,$$

et ceci est vrai quel que soit $k \geq 1$. On pose $\Omega_0 = \cap_k \Omega_k$, et on note que $\mathbb{P}(\Omega_0) = 1$ (par passage au complémentaire et en utilisant une *union bound*). Notons également que pour tout $\varepsilon > 0$, il existe $k \in \mathbb{N}$ tel que $\varepsilon > 1/k$. On en déduit finalement que

$$\forall \omega \in \Omega_0, \quad \forall \varepsilon > 0, \quad \exists N(\omega, \varepsilon), \quad \forall n \geq N(\varepsilon, \omega), \quad |X_n(\omega) - X(\omega)| \leq \varepsilon.$$

Ceci signifie que la suite de variables aléatoires $(X_n)_{n \geq 1}$ converge p.s. vers X. □

Exercice 2

1. Pour déterminer la loi de la v.a. M_n/n, on calcule sa fonction de répartition (c'est un réflexe à avoir quand on veut calculer la loi d'un min ou d'un max). Pour $x \in \mathbb{R}$, on a

$$\mathbb{P}(M_n/n \leq x) = \mathbb{P}(M_n \leq nx) = \mathbb{P}(\forall i = 1, \ldots, n, X_i \leq nx) = (\mathbb{P}(X_1 \leq nx))^n,$$

   la dernière égalité étant vraie car les X_i sont i.i.d.. De plus,

$$\mathbb{P}(X_1 \leq nx) = \int_{-\infty}^{nx} \frac{1}{\pi(1+u^2)} du = \frac{1}{\pi}\big[\arctan(u)\big]_{-\infty}^{nx} = \frac{1}{\pi}\arctan(nx) + \frac{1}{2}.$$

   Ainsi,

$$\mathbb{P}(M_n/n \leq x) = \left(\frac{1}{\pi}\arctan(nx) + \frac{1}{2}\right)^n.$$

2. On étudie la convergence simple de la fonction de répartition de M_n/n. D'après la question précédente, pour tout $x \in \mathbb{R}$, on a

$$\mathbb{P}(M_n/n \leq x) = \exp\left(n \ln\left(\frac{1}{\pi}\arctan(nx) + \frac{1}{2}\right)\right).$$

   Lorsque $x < 0$ ou $x = 0$, il est aisé de voir que $\mathbb{P}(M_n/n \leq x)$ tend vers 0. Pour déterminer la limite de $\mathbb{P}(M_n/n \leq x)$ quand $x > 0$, il suffit d'utiliser le rappel : $(1/\pi)\arctan(nx) + 1/2 = 1 - (1/\pi)\arctan(1/(nx))$. Ainsi

$$\mathbb{P}(M_n/n \leq x) = \exp\left(n \ln\left(1 - \frac{1}{\pi}\arctan(1/(nx))\right)\right).$$

   Comme $\frac{1}{\pi}\arctan(1/(nx)) \underset{n\to\infty}{\longrightarrow} 0$,

$$n \ln\left(1 - \frac{1}{\pi}\arctan(1/(nx))\right) \sim -\frac{n}{\pi}\arctan(1/(nx)) \sim \frac{n}{\pi}\frac{1}{(nx)} \to -\frac{1}{\pi x}.$$

Remarquons que nous n'avons pas composé les équivalences à gauche, chose interdite en générale ! Nous avons finalement pour tout $x \in \mathbb{R}$,

$$\mathbb{P}\left(M_n/n \leq x\right) \underset{n\to\infty}{\to} F^{\infty}(x),$$

avec

$$F^{\infty} : x \mapsto \begin{cases} 0 & \text{pour} \quad x \leq 0 \\ e^{-\frac{1}{\pi x}} & \text{pour} \quad x > 0 \end{cases}.$$

On vérifie que F^{∞} est une fonction de répartition (continue à droite, croissante, limite nulle en $-\infty$, limite 1 en $+\infty$). On en déduit que la suite de suite de v.a. M_n/n converge en loi vers la loi dont la fonction de répartition est F^{∞}. Pour la culture : cette loi limite est une loi de Fréchet.

3. Montrons à présent que cette loi admet une densité par rapport à la mesure de Lebesgue. Attention, ceci n'est pas automatique, car il y a des lois sur $\mathbb{R}$ qui ne sont pas absolument continues par rapport à la mesure de Lebesgue ! Par exemple, la fonction de répartition $F(x) = \mathbb{1}_{[0,+\infty[}(x)$ (mesure Dirac en 0) n'admet pas de densité par rapport à la mesure de Lebesgue. Une condition suffisante pour qu'une fonction de répartition admette une densité par rapport à la mesure de Lebesgue est qu'elle soit continue et C^1 par morceaux [1]. Alors la densité coïncide avec la dérivée de la fonction de répartition (p.p.).

   Après ce petit rappel, reprenons l'exercice et montrons que la loi associée à F^{∞} admet une densité. La fonction F^{∞} est continue et C^1 par morceaux, la loi associée à F^{∞} admet donc une densité f^{∞} par rapport à la mesure de Lebesgue, égale p.p. à la dérivée de F^{∞}. La densité f^{∞} est ainsi donnée par :

$$f^{\infty} : x \mapsto \begin{cases} 0 & \text{pour} \quad x \leq 0 \\ \frac{1}{\pi x^2} \exp\left(-\frac{1}{\pi x}\right) & \text{pour} \quad x > 0 \end{cases}.$$

Exercice 3

1. Comme f est positive, il suffit de vérifier que $\int_{\mathbb{R}} f(x)\,dx = 1$. Ceci s'obtient en faisant le changement de variable $y = \theta x$:

$$\int_{\mathbb{R}} f(x)\,dx = \int_0^{\infty} \frac{(\theta x)^{p-1}}{\Gamma(p)}\, e^{-\theta x}\,\theta dx = \frac{1}{\Gamma(p)} \int_0^{\infty} y^{p-1} e^{-y} dy = 1.$$

2. De la même façon que plus haut, on a

$$\mathbb{E}(X^k) = \int_{\mathbb{R}} x^k f(x)\,dx = \frac{1}{\theta^k \Gamma(p)} \int_0^{\infty} (\theta x)^{k+p-1}\, e^{-\theta x}\,\theta dx = \frac{\Gamma(k+p)}{\theta^k \Gamma(p)}.$$

   Ainsi, on a

$$\begin{aligned} \mathbb{E}(X) &= \Gamma(p+1)/(\theta\Gamma(p)) = p/\theta \\ \mathbb{E}(X^2) &= \Gamma(p+2)/(\theta^2\Gamma(p)) = p(p+1)/\theta^2 \\ \mathrm{Var}(X) &= \mathbb{E}(X^2) - (\mathbb{E}(X))^2 = \theta^{-2}(p(p+1) - p^2) = p/\theta^2. \end{aligned}$$

1. Pour rappel, une fonction $f : \mathbb{R} \to \mathbb{R}$ est dite C^1 par morceaux si pour tout segment de $\mathbb{R}$, il existe une partition finie composée d'intervalles, telle que sur chaque cellule A (de la forme $[a,b],]a,b], [a,b[$ ou $]a,b[$), la restriction de f à l'intérieur de A (c'est-à-dire à $]a,b[$) est prolongeable en une fonction C^1 sur A. Autrement dit, il est nécessaire est suffisant que f soit C^1 sur $]a,b[$ et que f et f' admettent des limites finies en a^+ et en a^-.

3. Pour toute fonction Ψ borélienne bornée,

$$\begin{aligned}\mathbb{E}(\Psi(Y^2)) &= \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}\Psi(y^2)e^{-y^2/2}dy \\ &= \frac{1}{\sqrt{2\pi}}\int_0^{\infty}\Psi(y^2)y^{-1}e^{-y^2/2}2ydy \\ &= \frac{1}{\sqrt{2\pi}}\int_0^{\infty}\Psi(u)u^{-1/2}e^{-u/2}du,\end{aligned}$$

en posant $u = y^2$. Comme $1/\sqrt{2\pi} = 1/(\sqrt{2}\Gamma(1/2))$, on a $Y^2 \sim \gamma(1/2, 1/2)$. Par ailleurs, on sait que $Y^2 \sim \chi^2(1)$ par définition.

4. Pour toute fonction Ψ borélienne bornée,

$$\begin{aligned}\mathbb{E}(\Psi(X/a)) &= \frac{\theta^p}{\Gamma(p)}\int_0^{\infty}\Psi(x/a)x^{p-1}\, e^{-\theta x}\, dx \\ &= \frac{(a\theta)^p}{\Gamma(p)}\int_0^{\infty}\Psi(x/a)(x/a)^{p-1}\, e^{-a\theta x/a}\, dx/a \\ &= \frac{(a\theta)^p}{\Gamma(p)}\int_0^{\infty}\Psi(y)y^{p-1}\, e^{-a\theta y}\, dy\end{aligned}$$

d'où le résultat.

5. Pour déterminer la loi de $X+Y$, on calcule sa fonction caractéristique $\Phi_{X+Y}(t) = \mathbb{E}[e^{it(X+Y)}]$. Grâce à l'indépendance, cette fonction est égale au produit des fonctions caractéristiques de X et de Y :

$$\Phi_{X+Y}(t) = \Phi_X(t)\,\Phi_Y(t) = \frac{1}{(1-i\theta^{-1}t)^{p_1}(1-i\theta^{-1}t)^{p_2}} = \frac{1}{(1-i\theta^{-1}t)^{p_1+p_2}}.$$

Ainsi, Φ_{X+Y} est la fonction caractéristique de la loi $\gamma(p_1+p_2, \theta)$.

6. D'après la question précédente (et par une récurrence triviale), $S_n \sim \gamma(n, \theta)$. En utilisant la question 2, nous avons donc $\mathbb{E}(S_n) = n/\theta$, $\mathrm{Var}(S_n) = n/\theta^2$.

7. En utilisant la question 3, les X_i^2 sont i.i.d. de loi $\gamma(1/2, 1/2)$. Ainsi, $S'_n \sim \gamma(n/2, 1/2)$. Comme nous savons par ailleurs que $S'_n \sim \chi^2(n)$ (par définition), nous avons $\gamma(n/2, 1/2) = \chi^2(n)$. En utilisant la question 2, on retrouve que $\mathbb{E}S'_n = n$ et $\mathrm{Var}\, S'_n = 2n$.

Exercice 4

1. $\forall\lambda \in \mathbb{R}$, on a :

$$\mathbb{E}\left(\exp(\lambda(X-\theta))\right) = \sum_{k\geq 0}\exp(\lambda(k-\theta))\mathbb{P}(X=k) = \sum_{k\geq 0}\exp(\lambda(k-\theta))e^{-\theta}\frac{\theta^k}{k!}$$

$$= \quad e^{-\theta(\lambda+1)}\sum_{k\geq 0}\frac{\left(e^{\lambda}\theta\right)^k}{k!} = e^{-\theta(\lambda+1)}e^{e^{\lambda}\theta} = \exp(\theta\phi(\lambda)).$$

2. Pour tout $x > 0$ et $\lambda \geq 0$, on a, par croissance de $x \mapsto e^{\lambda x}$ et par l'inégalité de Markov,

$$\begin{aligned}\mathbb{P}\big(X-\theta \geq x\big) &\leq \mathbb{P}\big(\exp(\lambda(X-\theta)) \geq \exp(\lambda x)\big) \\ &\leq \exp(-\lambda x)\mathbb{E}\left(\exp(\lambda(X-\theta))\right) \\ &= \exp(-\lambda x)\exp(\theta\phi(\lambda)),\end{aligned}$$

en utilisant la question 1.

3. Soit $f : \lambda \mapsto \lambda x - \phi(\lambda)$ définie sur $[0, +\infty[$, pour un $x > 0$ fixé. On a pour tout $\lambda \geq 0$,

$$f'(\lambda) > 0 \Leftrightarrow (x+1) - e^{\lambda} > 0 \Leftrightarrow \lambda < \ln(x+1).$$

On en déduit que f est strictement croissante sur $[0, \ln(x+1)]$ et strictement décroissante sur $[\ln(x+1), +\infty[$. Ainsi, elle atteint son maximum en l'unique point $\lambda_0 = \ln(x+1)$. Par suite,

$$\sup_{\lambda > 0}(\lambda x - \phi(\lambda)) = x\ln(x+1) - e^{\ln(x+1)} + \ln(x+1) + 1 = h(x).$$

4. Ainsi, en remarquant que l'inégalité trouvée à la question 2. est vraie pour tout $\lambda \geq 0$,

$$\begin{aligned}\mathbb{P}\big(X - \theta \geq x\big) &\leq \inf_{\lambda \geq 0}\left\{\exp\left(-\theta(\lambda\frac{x}{\theta} - \phi(\lambda))\right)\right\} \\ &= \exp\left(-\theta\sup_{\lambda \geq 0}\left\{\lambda\frac{x}{\theta} - \phi(\lambda)\right\}\right) \\ &= \exp\left(-\theta h(x/\theta)\right),\end{aligned}$$

en utilisant cette fois la question 3.

5. L'inégalité de Bienaymé-Tchebychev s'écrit, pour $X \sim \mathcal{P}(\theta)$,

$$\mathbb{P}\big(X - \theta \geq x\big) \leq \mathbb{P}\big(|X - \theta| \geq x\big) = \mathbb{P}\big(|X - \mathbb{E}X| \geq x\big) \leq \mathrm{Var}(X)/x^2 = \theta/x^2.$$

Ainsi, lorsque $\theta = 2$ et $x = 10 - \theta = 8$, il s'agit de comparer la borne de Bennett $e^{-2h(8/2)} = e^{-2h(4)} \simeq 0.0003$ avec celle de Bienaymé-Tchebychev $\theta/x^2 = 2/8^2 = 1/32 \simeq 0.031$. Encore une fois l'inégalité de Bienaymé-Tchebychev est moins fine que l'inégalité établie dans l'exercice. Cette inégalité de Bennett est vérifiée par les sommes de variables aléatoires qui peuvent n'être bornées que d'un côté : soient $X_1, \ldots, X_n$ des variables aléatoires indépendantes avec variance finie et telles que $X_i \leq b$ presque sûrement, pour $b \geq 0$ et pour tout $i \leq n$. Alors, quelque soit $x > 0$, on a

$$\mathbb{P}\left(\sum_{i=1}^{n}(X_i - \mathbb{E}X_i) \geq x\right) \leq \exp\left(-\frac{v}{b^2}h\left(\frac{bx}{v}\right)\right),$$

où $v = \sum_{i=1}^{n}\mathbb{E}[X_i^2]$.

Exercice 5

1. Pour tout $\lambda \geq 0$, la fonction $x \mapsto \exp(\lambda x)$ est convexe sur $\mathbb{R}$, on a donc l'inégalité suivante : pour tout $x \in [-M, M]$,

$$\exp(\lambda x) = \exp\left(\frac{M-x}{2M}(-\lambda M) + \frac{M+x}{2M}(\lambda M)\right) \leq \frac{M-x}{2M}\exp(-\lambda M) + \frac{M+x}{2M}\exp(\lambda M).$$

2. Pour tout $i \in \mathbb{N}$, on obtient par passage à l'espérance :

$$\begin{aligned}\mathbb{E}(\exp(\lambda X_i)) &\leq \mathbb{E}\left(\frac{M - X_i}{2M}\right)\exp(-\lambda M) + \mathbb{E}\left(\frac{M + X_i}{2M}\right)\exp(\lambda M) \\ &= \frac{\exp(-\lambda M)}{2} + \frac{\exp(\lambda M)}{2} \\ &= \cosh(\lambda M) \\ &\leq \exp(\lambda^2 M^2/2).\end{aligned}$$

La dernière inégalité est vraie car

$$\cosh(u) = \frac{1}{2}\sum_{k\geq 0}\frac{u^k}{k!} + \frac{1}{2}\sum_{k\geq 0}\frac{(-u)^k}{k!} = \sum_{k\geq 0}\frac{u^{2k}}{(2k)!}$$
$$\exp(u^2/2) = \sum_{k\geq 0}\frac{u^{2k}}{2^k k!}$$

et $2^k k!$ est le produit des facteurs pairs (seulement) de $(2k)!$.

3. Pour tout $\lambda \geq 0$ et $x > 0$, on a par l'inégalité de Markov :

$$\begin{aligned}\mathbb{P}\left(\frac{S_n}{n} \geq x\right) &= \mathbb{P}\left(\exp(\lambda S_n) \geq \exp(\lambda n x)\right)\\ &\leq \exp(-\lambda n x)\mathbb{E}\left(\exp(\lambda S_n)\right)\\ &= \exp(-\lambda n x)\prod_{i=1}^{n}\mathbb{E}\exp(\lambda X_i) \qquad ((X_i)_i \text{ indépendants mais pas i.i.d.})\\ &\leq \exp\left(n\left(\frac{\lambda^2 M^2}{2} - \lambda x\right)\right),\end{aligned}$$

en utilisant la question précédente.

4. Comme l'inégalité précédente est vraie pour tout $\lambda \geq 0$, on a pour tout $x > 0$,

$$\begin{aligned}\mathbb{P}\left(\frac{S_n}{n} \geq x\right) &\leq \inf_{\lambda>0}\exp\left(n\left(\frac{\lambda^2 M^2}{2} - \lambda x\right)\right) = \exp\left(n\inf_{\lambda>0}\left(\frac{\lambda^2 M^2}{2} - \lambda x\right)\right)\\ &= \exp\left(n\left(\frac{x^2}{2M^2} - \frac{x^2}{M^2}\right)\right) = \exp\left(-\frac{nx^2}{2M^2}\right),\end{aligned}$$

car l'infimum est atteint pour $\lambda = x/M^2$. On obtient la deuxième inégalité en remarquant que :

$$\mathbb{P}\left(\left|\frac{S_n}{n}\right| \geq x\right) \leq \mathbb{P}\left(\frac{S_n}{n} \geq x\right) + \mathbb{P}\left(-\frac{S_n}{n} \geq x\right),$$

où la première inégalité vient du fait que l'événement $\{\frac{S_n}{n} = x\}$ se retrouve dans les deux événements de droite, qui ne sont donc pas disjoints. Comme $-S_n$ est la somme des v.a. $-X_k$ qui restent centrées et dans $[-M, M]$, l'inégalité de Hoeffding précédente s'applique encore et on obtient le résultat voulu.

5. Soit $U_0, U_1, U_2, \ldots$ une suite de v.a. i.i.d. uniformes sur $[0, 2]$. Les v.a. $X_k = U_k - 1$ sont centrées et dans $[-1, 1]$, on peut donc appliquer l'inégalité de Hoeffding et on obtient :

$$\mathbb{P}\left(\left|\frac{\sum_{k=1}^n X_k}{n}\right| \geq x\right) = \mathbb{P}\left(\left|\frac{\sum_{k=1}^n U_k}{n} - 1\right| \geq x\right) \leq 2\exp(-nx^2/2) =: \Delta_H.$$

L'inégalité de Bienaymé-Tchebychev appliquée à la v.a. $\sum_{k=1}^n U_k/n$ (on calcule $\mathbb{E}(U_1) = 1$ et $\mathbb{V}(U_1) = 1/3$) s'écrit

$$\mathbb{P}\left(\left|\frac{\sum_{k=1}^n U_k}{n} - 1\right| \geq x\right) \leq \frac{\mathbb{V}\big(\sum_{k=1}^n U_k/n\big)}{x^2} = \frac{1}{3nx^2} =: \Delta_{BT}$$

Les bornes Δ_H et Δ_{BT} se comparent de la façon suivante :
— pour $n = 100$ et $x = 0.1$, on a $nx^2 = 1$ donc $\Delta_H \simeq 1.21$ et $\Delta_{BT} = 1/3$;
— pour $n = 100$ et $x = 0.5$, on a $nx^2 = 25$ donc $\Delta_H \simeq 7.46\,10^{-6}$ et $\Delta_{BT} = 1/75 \simeq 0.013$;
— pour $n = 1000$ et $x = 0.1$, on a $nx^2 = 10$ donc $\Delta_H \simeq 0.01$ et $\Delta_{BT} = 1/30 \simeq 0.03$.

La borne de Hoeffding est mauvaise pour les petites valeurs de nx^2 mais devient beaucoup plus fine que celle de Bienaymé-Tchebychev pour de plus grandes valeurs de nx^2 donc de x ou n. Par contre, la borne de Hoeffding utilise des hypothèses plus contraignantes (v.a. bornées).

Exercice 6

1. Pour toute fonction Ψ borélienne bornée, on a

$$\begin{aligned}\mathbb{E}(\Psi(X_1^2)) &= \int_0^\infty \Psi(x^2) 2xe^{-x^2} dx \\ &= \int_0^\infty \Psi(u) e^{-u} du,\end{aligned}$$

en posant $u = x^2$ ($du = 2xdx$). Ainsi, $X_1^2 \sim \mathcal{E}(1)$ (loi exponentielle de paramètre 1). D'après l'exercice 3, comme les X_i^2 sont i.i.d. de loi $\mathcal{E}(1) = \gamma(1,1)$ nous avons $\sum_{i=1}^n X_i^2 \sim \gamma(n,1)$ puis $V_n = n^{-1} \sum_{i=1}^n X_i^2 \sim \gamma(n,n)$.

2. On applique le TCL aux variables aléatoires X_i^2 i.i.d. et de carré intégrable pour obtenir

$$\sqrt{n}(V_n - \mathbb{E}(X_1^2)) \rightsquigarrow \mathcal{N}\left(0, \mathrm{Var}(X_1^2)\right).$$

Comme $X_1^2 \sim \mathcal{E}(1) = \gamma(1,1)$, on calcule aisément $\mathbb{E}(X_1^2) = 1$ et $\mathrm{Var}(X_1^2) = 1$. Ainsi, en utilisant la caractérisation de la convergence en loi avec les fonctions de répartition, on a que pour tout $t \in \mathbb{R}$,

$$\mathbb{P}(\sqrt{n}(V_n - \mathbb{E}V_n) \geq t) \underset{n\to\infty}{\longrightarrow} 1 - \Phi(t).$$

Par ailleurs, comme $1 - \Phi(x) = \Phi(-x)$ par symétrie de la loi $\mathcal{N}(0,1)$, on a bien le résultat escompté.

3. (a) Comme $V_n \sim \gamma(n,n)$, V_n n'est pas bornée (p.s.) et on ne peut pas utiliser l'inégalité de Hoeffding. Cependant, V_n possède des moments exponentiels (comme on le verra plus bas), donc on cherche à utiliser la méthode de Chernoff, en espérant que les calculs soient faisables.

   (b) Comme $V_n \sim \gamma(n,n)$, on a

$$\begin{aligned}\mathbb{E}\left(\exp(\lambda V_n)\right) &= \frac{n^n}{\Gamma(n)} \int_0^\infty e^{\lambda x} x^{n-1} e^{-nx} dx \\ &= \frac{n^n}{(n-\lambda)^n} \frac{(n-\lambda)^n}{\Gamma(n)} \int_0^\infty x^{n-1} e^{-(n-\lambda)x} dx \\ &= \frac{1}{(1-\lambda/n)^n} \times 1,\end{aligned}$$

   en reconnaissant la loi $\gamma(n, n-\lambda)$. Par suite,

$$\mathbb{E}\left(\exp(\lambda(V_n - 1))\right) = \exp\left(-\lambda - n\ln(1-\lambda/n)\right),$$

   d'où $\phi(\lambda) = -\lambda - n\ln(1-\lambda/n)$.

   (c) Il s'agit de maximiser la fonction $\lambda \mapsto \lambda(x+1) + n\ln(1-\lambda/n)$ sur $[0,+\infty[$ pour un $x > 0$ fixé. La dérivée en λ est positive si et seulement si

$$x + 1 - n/(n-\lambda) \geq 0$$

   c'est-à-dire $\lambda \leq nx/(x+1)$. Donc le maximum est atteint en $\lambda_0 = nx/(x+1)$. On vérifie bien que $\lambda_0 < n$. De plus, la fonction vaut en λ_0

$$h(x) = \lambda_0(x+1) + n\ln(1-\lambda_0/(n)) = xn - n\ln(x+1).$$

(d) D'après la méthode de Chernoff (refaire la preuve dans notre cas pour vous en convaincre !), on a pour tout $t > 0$,

$$\mathbb{P}\left(V_n - 1 \geq t/\sqrt{n}\right) \leq \exp\left(-h(t/\sqrt{n})\right),$$

avec

$$\begin{aligned} h(t/\sqrt{n}) &= t\sqrt{n} - n\ln(t/\sqrt{n} + 1) \\ &= n\psi(t/\sqrt{n}), \end{aligned}$$

ce qui prouve finalement que

$$\mathbb{P}\left(V_n - 1 \geq t/\sqrt{n}\right) \leq \exp\left(-n\psi(t/\sqrt{n})\right).$$

Lorsque n tend vers l'infini, on a $-n\psi(t/\sqrt{n}) = -t^2/2 + o(1)$. Donc la borne est équivalente à $e^{-t^2/2}$, ce qui est toujours plus grand que $\Phi(-t)$ (tout en étant souvent proche), ce qui semble indiquer que l'inégalité plus haut est (légèrement) moins bonne que celle obtenue dans la question 2. Cependant, il est important de noter que, en contrepartie, l'inégalité obtenue plus haut est valable pour tout $n \geq 1$ alors que le résultat de la question 2 n'est valable que lorsque n tend vers l'infini.

Finalement, cet exemple montre surtout que la méthode de Chernoff est assez précise lorsque n tend vers l'infini.

Exercice 7

1. La LFGN appliquée aux variables aléatoires X_i i.i.d. et intégrables donne

$$\overline{X}_n \overset{p.s.}{\to} \mathbb{E}(X_1) = m.$$

De plus, nous avons la décomposition (à savoir !)

$$\sigma_n^2 = n^{-1}\sum_{j=1}^n X_j^2 - (\overline{X}_n)^2.$$

La LFGN appliquée aux variables aléatoires X_i^2 i.i.d. et intégrables (car $\mathbb{E}[X_1^2] = \operatorname{Var} X_1 + \mathbb{E}[X_1]^2 = \sigma^2 + m^2 < \infty$) donne

$$n^{-1}\sum_{j=1}^n X_j^2 \overset{p.s.}{\to} \mathbb{E}(X_1^2) = \sigma^2 + m^2.$$

Au final, nous avons

$$\sigma_n^2 \overset{p.s.}{\to} \sigma^2.$$

Remarque : attention ! On ne peut pas appliquer la LFGN directement aux $X_i - \overline{X}_n$ car ces variables ne sont pas indépendantes !

2. Les v.a. X_i sont i.i.d. dans $\mathbb{L}^2$, on peut donc leur appliquer le TCL :

$$\sqrt{n}\,\frac{\overline{X}_n - m}{\sigma} \rightsquigarrow \mathcal{N}(0,1).$$

De plus, d'après la question précédente, $\sigma_n^2 \xrightarrow{\text{p.s.}} \sigma^2$. Montrons qu'alors $\frac{\sigma}{\sqrt{\sigma_n^2}} \overset{p.s.}{\to} 1$ (quand bien même $\frac{\sigma}{\sqrt{\sigma_n^2}}$ peut prendre la valeur ∞). Cela n'est pas direct car σ_n^2 peut s'annuler avec probabilité non-nulle.

Première option Soit $\varphi : x \in \mathbb{R} \mapsto \begin{cases} \frac{\sigma}{\sqrt{x}} & \text{si } x > 0 \\ \infty & \text{sinon.} \end{cases}$. Cette fonction est C^0 sur $\mathbb{R}_+$ mais elle est surtout continue en σ^2 (puisque $\sigma^2 > 0$), qui est la limite p.s. de σ_n^2. Ainsi, par le théorème de continuité, $\varphi(\sigma_n^2) = \frac{\sigma}{\sqrt{\sigma_n^2}} \overset{p.s.}{\to} 1$.

Deuxième option (hors programme) Puisque, $\sigma_n^2 \overset{\text{p.s.}}{\longrightarrow} \sigma^2$, il existe donc un événement Ω_0 de mesure 1, tel que $\forall \omega \in \Omega_0, \sigma_n^2(\omega) \underset{n\to\infty}{\longrightarrow} \sigma^2 > 0$. Ainsi, il existe un rang $n_0(\omega)$ tel que $\forall n \geq n_0(\omega), \sigma_n^2(\omega) > 0$ et $\forall n \geq n_0(\omega)$, il est licite de composer par la fonction continue $x \in \mathbb{R}_+^* \mapsto \frac{\sigma}{\sqrt{x}}$, ce qui assure que $\frac{\sigma}{\sqrt{\sigma_n^2(\omega)}} \underset{n\to\infty}{\longrightarrow} 1$. Enfin, puisque cela est vrai pour tout $\omega \in \Omega_0$ et que $\mathbb{P}(\Omega_0) = 1$, $\frac{\sigma}{\sqrt{\sigma_n^2}} \overset{p.s.}{\to} 1$.

Ces deux convergences (en loi et $p.s.$) et le lemme de Slutsky impliquent que

$$\sqrt{n}\,\frac{\overline{X}_n - m}{\sqrt{\sigma_n^2}} = \sqrt{n}\,\frac{\overline{X}_n - m}{\sigma}\,\frac{\sigma}{\sqrt{\sigma_n^2}} \rightsquigarrow \mathcal{N}(0,1).$$

Exercice 8

1. Soit $Z_1 = -\ln(X_1)$ (bien défini car $\mathbb{P}(X_1 = 0) = 0$), alors Z_1 suit une loi exponentielle de paramètre 1 : en effet, pour tout $u \in \mathbb{R}$,

$$\mathbb{P}(Z_1 \leq u) = \mathbb{P}(X_1 \geq e^{-u}) = (1 - e^{-u})\mathbb{1}_{[0,+\infty[}(u).$$

2. Comme $\ln Y_n = \overline{Z}_n$, où $\overline{Z}_n = n^{-1}\sum_{i=1}^n Z_i$, $\mathbb{E}Z_1 = 1$, $\operatorname{Var} Z_1 = 1$, le théorème de la limite centrale appliqué aux variables $\{Z_i\}_i$ i.i.d. vérifiant $\mathbb{E}Z_1^2 < \infty$ donne

$$\sqrt{n}(\overline{Z}_n - 1) \rightsquigarrow \mathcal{N}(0,1).$$

La méthode Delta appliquée avec la fonction $g(u) = e^u$ qui est dérivable en 1 avec $g'(1) = e$ donne alors

$$\sqrt{n}(Y_n - e) \rightsquigarrow \mathcal{N}(0, e^2).$$

Exercice 9

1. Par la LFGN (X_i i.i.d et intégrables), $\overline{X}_n$ converge p.s. vers $\mathbb{E}(X_1) = 0$. De même, le TCL (X_i i.i.d et de carré intégrable) donne trivialement $\sqrt{n}\overline{X}_n \rightsquigarrow \mathcal{N}(0,1)$.
2. En se souvenant que $\cos' = -\sin$ et par la Delta méthode (cos dérivable en 0), on obtient $\sqrt{n}(\cos(\overline{X}_n) - 1) \rightsquigarrow 0$. Ici la limite est dégénérée car p.s. constante. On en déduit uniquement que la vitesse de convergence de $\cos(\overline{X}_n)$ vers 1 est plus rapide que $1/\sqrt{n}$. Il faut donc pousser le développement à l'ordre suivant pour pouvoir être plus précis sur vitesse et limite.
3. Le développement limité de cos en 0 à l'ordre 2 est :

$$\begin{aligned} \cos x &= \cos 0 + \cos'(0)x + \cos''(0)x^2/2 + o(x^2) \\ &= 1 - x^2/2 + o(x^2). \end{aligned}$$

Autrement dit, il existe une fonction r tendant vers 0 en 0 telle que :

$$\cos x = 1 - \frac{x^2}{2}(1 + r(x)).$$

4. D'après la question précédente, nous avons :

$$\cos(\overline{X}_n) - 1 = -\frac{(\overline{X}_n)^2}{2}(1 + r(\overline{X}_n)),$$

d'où

$$-2n(\cos(\overline{X}_n) - 1) = (\sqrt{n}\overline{X}_n)^2(1 + r(\overline{X}_n)).$$

Il reste maintenant à étudier la convergence du terme de droite avec comme outil le théorème de Slutsky. Tout d'abord, la fonction r est prolongeable par continuité en 0 de sorte que $r(0) = 0$. Or $\overline{X}_n$ converge p.s. (et a fortiori en probabilité) vers 0, donc par le théorème de continuité, $r(\overline{X}_n) \xrightarrow{\mathbb{P}} 0$. De même, $\sqrt{n}\overline{X}_n \rightsquigarrow \mathcal{N}(0,1)$, donc par continuité $(\sqrt{n}\overline{X}_n)^2 \rightsquigarrow Y^2$ avec $Y \sim \mathcal{N}(0,1)$. Autrement dit, $(\sqrt{n}\overline{X}_n)^2 \rightsquigarrow \chi^2(1)$. Finalement, le théorème de Slutsky donne :

$$-2n(\cos(\overline{X}_n) - 1) \rightsquigarrow \chi^2(1).$$

Bilan : la suite de v.a. $(\cos(\overline{X}_n))$ converge à vitesse $1/n$ vers sa limite, c'est-à-dire bien plus rapidement que pour le TCL classique.

Exercice 10

1. Bien que X_n se comporte comme une somme de variables aléatoires de Bernoulli, il est impossible d'utiliser la loi des grands nombres pour montrer que X_n/n converge vers θ, car X_n s'écrit comme une somme de Bernoulli en loi et non p.s.
2. Pour démontrer la convergence p.s., on va utiliser le lemme de Borel-Cantelli, autrement dit, on cherche à montrer que, pour tout $\varepsilon > 0$,

$$\sum_{n\geq 1} \mathbb{P}\left(|X_n/n - \theta| \geq \varepsilon\right) < \infty.$$

Or, puisqu'une variable de Bernoulli est bornée par 0 et 1, l'inégalité de Hoeffding (la version du cours) donne (pour tout $c > 0$) :

$$\mathbb{P}\left(|X_n/n - \theta| \geq c\right) \leq 2\exp(-2c^2 n).$$

Ainsi,

$$\sum_{n\geq 1} \mathbb{P}\left(|X_n/n - \theta| \geq c\right) \leq 2\sum_{n\geq 1} \exp(-2c^2 n) < \infty,$$

donc par le lemme de Borel-Cantelli, X_n/n converge p.s. vers θ.

3. Soit $\varphi : x \in \mathbb{R}_+ \mapsto \begin{cases} \ln(x) & \text{si } x > 0 \\ 1 & \text{sinon.} \end{cases}$. Alors $Y_n = \varphi(X_n/n)$ et φ est continue en θ (puisque $\theta > 0$), qui est la limite p.s. de X_n/n. Ainsi, puisque $X_n/n \xrightarrow{\text{p.s.}} \theta$, le théorème de continuité assure que $Y_n = \varphi(X_n/n) \xrightarrow{\text{p.s.}} \varphi(\theta) = \ln(\theta)$.
4.

Première méthode D'après ce qui précède, $Y_n = \varphi(X_n/n)$ avec φ dérivable en θ et de dérivée $\varphi'(\theta) = 1/\theta > 0$. Or, d'après le TCL,

$$\sqrt{n}(X_n/n - \theta) \rightsquigarrow \mathcal{N}(0, \theta(1-\theta)).$$

Donc, par application de la méthode Delta avec la fonction φ,

$$\sqrt{n}(\varphi(X_n/n) - \varphi(\theta)) = \sqrt{n}(Y_n - \ln(\theta)) \rightsquigarrow \mathcal{N}(0, 1/\theta - 1).$$

Deuxième méthode On remarque tout d'abord que

$$\sqrt{n}(Y_n - \ln(\theta)) = \sqrt{n}(\ln(X_n/n) - \ln(\theta))\mathbb{1}_{X_n \geq 1} + \sqrt{n}(1 - \ln(\theta))\mathbb{1}_{X_n = 0}.$$

Montrons que le terme de droite converge vers 0 en probabilité. Pour tout $\epsilon > 0$,

$$\mathbb{P}(|\sqrt{n}(1 - \ln(\theta))\mathbb{1}_{X_n=0}| \geq \epsilon) \leq \mathbb{P}(X_n = 0) = (1-\theta)^n \underset{n\to\infty}{\longrightarrow} 0.$$

Ainsi $\sqrt{n}(1 - \ln(\theta))\mathbb{1}_{X_n=0} \xrightarrow{\mathbb{P}} 0$.
Concernant le terme de gauche, le TCL donne

$$\sqrt{n}(X_n/n - \theta) \rightsquigarrow \mathcal{N}(0, \theta(1-\theta)).$$

De plus, par la méthode Delta appliquée avec ln (dérivable en θ), nous obtenons :

$$\sqrt{n}(\ln(X_n/n) - \ln(\theta)) \rightsquigarrow \mathcal{N}(0, 1/\theta - 1).$$

Ensuite, par application du théorème de Slutsky en tenant compte que $\mathbb{1}_{X_n \geq 1}$ tend vers 1 en probabilité, nous obtenons :

$$\sqrt{n}(\ln(X_n/n) - \ln(\theta))\mathbb{1}_{Y_n = \ln(X_n/n)} \rightsquigarrow \mathcal{N}(0, 1/\theta - 1).$$

Finalement, puisque

$$\sqrt{n}(Y_n - \ln(\theta)) = \sqrt{n}(\ln(X_n/n) - \ln(\theta))\mathbb{1}_{X_n \geq 1} + \sqrt{n}(1 - \ln(\theta))\mathbb{1}_{X_n = 0},$$

on obtient (en appliquant de nouveau le théorème de Slutsky) :

$$\sqrt{n}(Y_n - \ln(\theta)) \rightsquigarrow \mathcal{N}(0, 1/\theta - 1).$$