Correction de l'examen du 6 Mai 2024 (durée 2h)

Exercice 1. — Questions de cours

1. Rappeler l'énoncé de la loi forte des grands nombres.
2. Rappeler l'énoncé du TCL (théorème central limite).
3. Est-ce qu'un processus de Galton-Watson dont la mesure de reproduction μ est donnée par $\mu = \frac{2}{3}\delta_0 + \frac{1}{6}\delta_2 + \frac{1}{6}\delta_4$ a une chance de survivre indéfiniment ? Justifier.

Correction. 3. Si $X \sim \mu$, alors $\mathbb{E}\big[X\big] = \ldots = 1$. Etant donné que $\mu \neq \delta_1$, le théorème vu en cours sur Galton-Watson implique que la généalogie $(Z_n)_{n\geq 0}$ s'éteint presque sûrement (i.e. il existe presque surement un $N \geq 1$ tel que $Z_N = 0$).

Exercice 2. —

Soit $(X_n)_{n\geq 1}$ une suite de v.a. i.i.d vérifiant $\mathbb{P}\big[X_n = 1\big] = \mathbb{P}\big[X_n = 0\big] = 1/2$. Pour $n \geq 1$, on pose

$$Y_n := X_n + X_{n+1}$$

1. Les v.a $(Y_n)_{n\geq 1}$ sont-elles identiquement distribuées ? Justifier.
2. Pour $n \geq 1$, déterminer la loi, l'espérance et la variance de Y_n.
3. Les v.a. $(Y_n)_{n\geq 1}$ sont-elles indépendantes ? Justifier.
4. Les v.a. $(Y_{2n})_{n\geq 1}$ sont-elles indépendantes ? Justifier.
5. Montrer que la suite $\big(\frac{Y_1+\ldots+Y_n}{n}\big)_{n\geq 1}$ converge presque sûrement vers une v.a. à préciser.

Correction.
1. Chaque v.a. Y_n est une fonction $\phi(X_n, X_{n+1})$ ou le vecteur (X_n, X_{n+1}) est identiquement distribué, donc les Y_n sont identiquement distribués.
2. La loi de Y_n est la mesure de proba supportée sur $\{0, 1, 2\}$. Elle est donnée par $\mu = \frac{1}{4}\delta_0 + \frac{1}{2}\delta_1 + \frac{1}{4}\delta 2$. En effet, $\mathbb{P}\big[Y_1 = 0\big] = \mathbb{P}\big[X_1 = X_2 = 0\big] = \mathbb{P}\big[X_1 = 0\big]\mathbb{P}\big[X_2 = 0\big]$ (par indépendance!) et est donc égal a 1/4. Même justification pour $\mathbb{P}\big[Y_1 = 2\big]$ et on conclut pour la dernière valeur possible par exemple en utilisant que la somme des probas vaut 1.
La moyenne $\mathbb{E}\big[Y_n\big]$ est bien définie car Y_n est une v.a. positive. Elle est donnée par

$$\mathbb{E}\big[Y_n\big] = \tfrac{1}{2} + \tfrac{1}{4} * 2 = 1.$$

Sa variance est

$$\sigma^2 = \mathrm{Var}\big[Y_n\big] = \mathbb{E}\big[(Y_n - 1)^2\big] = \tfrac{1}{4}1^2 + \tfrac{1}{4}1^2 = \tfrac{1}{2}\,.$$

3. **non.** En effet, on a $\mathbb{P}\big[Y_n = Y_{n+1} = 0\big] = \mathbb{P}\big[X_n = X_{n+1} = X_{n+2} = 0\big] = \frac{1}{8} \neq \mathbb{P}\big[Y_n = 0\big]\mathbb{P}\big[Y_{n+1} = 0\big]$.
4. **oui.** En effet, pour tout $n \geq 1$, $Y_{2n} = \phi(X_{2n}, X_{2n+1})$ (où $\phi(x, y) := x + y$) et les vecteurs $Z_n := (X_{2n}, X_{2n+1})$ forment une suite i.i.d de vecteurs aléatoires.
5. Raisonnons pour n pair, on a dans ce cas,

$$\frac{Y_1 + \ldots + Y_n}{n} = \frac{Y_1 + Y_3 + \ldots + Y_{n-1}}{n} + \frac{Y_2 + Y_4 + \ldots + Y_n}{n}$$

On peut appliquer la loi des grands nombres à chacune des deux sommes (dont les termes sont i.i.d d'apres le même raisonnement que dans la question 4). Ces deux sommes comptent chacune $\frac{n}{2}$ termes. La loi des grands nombres s'applique a chacune des deux sommes car les Y_{2n} (et Y_{2n+1}) sont non seulement i.i.d mais aussi L^1 (question 2.). La première somme converge p.s. vers $\frac{1}{2} * \mathbb{E}\big[Y_1\big]$.

(Le facteur $\frac{1}{2}$ provient du fait qu'on doit diviser par le nombre de termes, i.e. $(n \pm 1)/2$ plutot que n). La réunion de deux ensembles de mesure nulle est encore de mesure nulle, on obtient donc que la somme de ces deux termes tend p.s. vers $\frac{1}{2} + \frac{1}{2} = 1$.
Le cas n impair s'obtient en s'arrêtant au dernier indice pair et en constatant que le dernier terme impair contribue au plus $2/n \to 0$ (de manière déterministe).

Exercice 3. — Loi de Poisson

On rappelle que si X suit une loi de Poisson de paramètre $\lambda \geq 0$ (on notera $X \sim \mathcal{P}(\lambda)$) alors $\mathbb{P}\big[X = 0\big] = e^{-\lambda}$ et pour tout $k \in \mathbb{N} \setminus \{0\}$, $\mathbb{P}\big[X = k\big] = e^{-\lambda}\frac{\lambda^k}{k!}$.

1. Calculer la fonction caractéristique d'une loi de Poisson $\mathcal{P}(\lambda)$
2. En déduire que si X et Y sont indépendantes et respectivement de loi $\mathcal{P}(\lambda)$ et $\mathcal{P}(\mu)$ ($\lambda \geq 0, \mu \geq 0$), alors $Z = X + Y$ suit une loi de Poisson $\mathcal{P}(\lambda + \mu)$.
3. En utilisant la question précédente, montrer que si $\lambda_1 \leq \lambda_2$, et si $X_1 \sim \mathcal{P}(\lambda_1), X_2 \sim \mathcal{P}(\lambda_2)$, alors pour tout $n \in \mathbb{N}$,
$$\mathbb{P}\big[X_1 \geq n\big] \leq \mathbb{P}\big[X_2 \geq n\big]$$

Correction.
1. Pour tout $\xi \in \mathbb{R}$ on a (N.B. l'espérance ci-dessous est bien définie car $|e^{i\xi X}| \leq 1$).

$$\Phi_X(\xi) := \mathbb{E}\big[e^{i\xi X}\big] = \sum_{k\geq 0} e^{-\lambda}\frac{\lambda^k e^{i\xi k}}{k!} = e^{-\lambda + \lambda e^{i\xi}}$$

2. Vu que X et Y sont indépendantes, on a pour tout $\xi \in \mathbb{R}$,

$$\Phi_{X+Y}(\xi) = \Phi_X(\xi)\Phi_X(\xi) = e^{-(\lambda+\mu)+(\lambda+\mu)e^{i\xi}}$$

On reconnaît alors la fonction caractéristique d'une loi de Poisson. Par un théorème du cours (injectivité de $\mu \mapsto \hat{\mu}$ sur l'espace des mesures de probas), on en déduit ainsi que $X + Y \sim \mathcal{P}(\lambda + \mu)$.
3. Ecrivons $\lambda_2 := \lambda_1 + \delta\lambda$ avec $\delta\lambda \geq 0$. Soient X et Y deux v.a. de poissons **indépendantes** et de paramètres respectifs λ_1 et $\delta\lambda$. On a donc par 2. X_1 a même loi que X et X_2 a même loi que $X + Y$. On en déduit donc

$$\mathbb{P}\big[X_2 \geq n\big] = \mathbb{P}\big[X + Y \geq n\big] \geq \mathbb{P}\big[X \geq n\big] = \mathbb{P}\big[X_1 \geq n\big]$$

Exercice 4. — Une propriété qui caractérise la Gaussienne

Soient X, Y deux variables aléatoires i.i.d L^2 ayant la propriété que $\frac{X+Y}{\sqrt{2}}$ a même loi que X. On pose $\sigma^2 = \mathrm{Var}(X)$.

1. Vérifier que la loi Gaussien $\mathcal{N}(0, \sigma^2)$ a la propriété requise. Le but de l'exercice est de montrer que c'est la seule loi ayant cette propriété.
2. Montrer que $\mathbb{E}\big[X\big] = 0$.
3. Montrer que pour tout $n \in \mathbb{N}$, si $X_1, \ldots, X_{2^n}$ sont des variables aléatoires i.i.d de même loi que X alors, la variable aléatoire
$$\frac{1}{2^{n/2}}(X_1 + \ldots + X_{2^n})$$
a même loi que X.
4. Conclure.

Correction.
1. Il fallait rappeler une propriété de la Gaussienne vue en cours.
2. Important, on a le droit de considérer $\mathbb{E}\big[(X+Y)/\sqrt{2}\big]$ car on a supposé X (et Y) L^2. La somme de ces deux variables aléatoires est donc L^2 (on a pas besoin de l'indépendance ici). Et $L^2 \subset L^1$ (sur un espace de probas). On peut donc considérer cette espérance et par linéarité plus le fait que X et Y ont même loi, on a

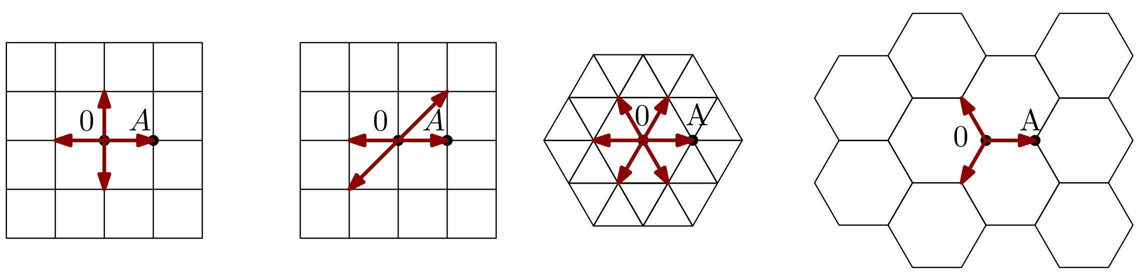
$$\mathbb{E}\big[\frac{X+Y}{\sqrt{2}}\big] = \sqrt{2}\mathbb{E}\big[X\big]$$

mais c'est aussi égal à $\mathbb{E}\big[X\big]$ par hypothèse. On a donc $\sqrt{2}\mathbb{E}\big[X\big] = \mathbb{E}\big[X\big]$ ce qui implique $\mathbb{E}\big[X\big] = 0$.
3. et 4. C'est une application du TCL. Tous les ingrédients de l'énoncé du TCL sont bien réunis : indépendance plus le fait d'être dans L^2, on a donc

$$\frac{1}{2^{n/2}}(X_1 + \ldots + X_{2^n}) \xrightarrow{(loi)} \mathcal{N}(0, \sigma^2)$$

Par ailleurs, via une récurrence facile en partant de l'hypothèse (l'identité en loi $\frac{X+Y}{\sqrt{2}} \overset{(loi)}{=} X$), on obtient que la loi du terme de gauche est constante et est égale pour tout n à celle de X. On peut maintenant conclure de plein de manières différentes, par exemple en observant que si X_n est une suite de v.a. de même loi et qui converge en loi vers X, alors la loi de X a la même que celle des X_n, en effet pour tout $\phi \in \mathcal{C}_b(\mathbb{R})$, on a $\mathbb{E}\big[\phi(X_n)\big] = \mathbb{E}\big[\phi(X_1)\big] \to \mathbb{E}\big[\phi(X)\big]$ et donc $\mathbb{E}\big[\phi(X_n)\big] = \mathbb{E}\big[\phi(X)\big]$ pour toute telle fonction test ϕ ce qui suffit à caractériser la loi de X.

Exercice 5. — Marches aléatoires sur des réseaux du plan



marche a) marche b) marche c) marche d)

On considère 4 marches aléatoires dans $\mathbb{R}^2$ qui partent de l'origine $0 = (0,0) \in \mathbb{R}^2$. Pour les 3 premières, les marches a), b) et c), elles sont de la forme $S_n = X_1 + \ldots + X_n$ où les $(X_i)_{i\geq 1}$ sont des variables aléatoires i.i.d dans $\mathbb{R}^2$ dont la loi est la **mesure uniforme** sur les vecteurs issus de l'origine et dessinés sur le graphique ci-dessus. Pour fixer l'échelle, le point A correspond à chaque fois au point $(1,0)$. Par exemple, la marche aléatoire à gauche (le cas a)) correspond à la marche aléatoire dont les incréments i.i.d X_i ont loi $\mu_a = \frac{1}{4}(\delta_{(1,0)} + \delta_{(0,1)} + \delta_{(-1,0)} + \delta_{(0,-1)})$. La dernière marche (d) est expliquée dans la dernière question.

1. Rappeler quelle est la définition d'un vecteur Gaussien $Y = (Y_1, \ldots, Y_k)$.
2. Dans l'exemple de gauche a), vers quoi converge le vecteur aléatoire $\frac{1}{\sqrt{n}}S_n$?
3. Traiter l'exemple b) et la limite correspondante pour $\frac{1}{\sqrt{n}}S_n$. On commencera par expliciter la loi des incréments μ_b puis la matrice de covariance de $X_1 \sim \mu_b$.
4. Traiter l'exemple c) et la limite correspondante pour $\frac{1}{\sqrt{n}}S_n$. (Cette marche est appelée marche aléatoire sur le *réseau triangulaire*).
5. La dernière marche est une marche au hasard sur le *réseau hexagonal*. S_0 part de l'origine 0 et choisit l'un des trois sommets les plus proches (voir figure d)) avec proba $1/3, 1/3, 1/3$.

Ce nouveau sommet est S_1. Pour le deuxième pas de la marche, on choisit à nouveau l'un des 3 sommets les plus proches uniformément au hasard, ce qui nous conduit à S_2 et ainsi de suite. Vérifier que $\mathbb{P}\big[S_2 = 0\big] = \frac{1}{3}$. Montrer que pour tout $n \geq 1$, on peut écrire $S_{2n} = X_1 + Y_1 + X_2 + Y_2 + \ldots + X_n + Y_n$ où les X_i et Y_i sont des suites de variables i.i.d de loi différente. En déduire une limite pour le vecteur aléatoire $S_n/\sqrt{n}$.

Correction.
1. Question de cours. Il fallait rappeler l'une des définitions du cours ici (soit le fait que toutes les combinaisons linéaires sont Gaussiennes, soit ...)
2. Dans cet exemple, les incréments sont des vecteurs i.i.d dont toutes les composantes (la composante x et la composante y) sont dans L^2. Par ailleurs, ces vecteurs i.i.d sont de moyenne $0 = (0, 0)$ et leur matrice de covariance est donnée par

$$\mathrm{Cov}_{\mu_a} = A := \begin{bmatrix} 1/2 & 0 \\ 0 & 1/2 \end{bmatrix}$$

On peut donc appliquer le TCL vectoriel pour obtenir que $S_n/\sqrt{n}$ converge **en loi** vers un vecteur Gaussien **centré** de matrice de covariance A.
3. Même argumentaire, i.e. L^2, indépendance, moyenne nulle, et cette fois la matrice de covariance des incréments est donnée par

$$\mathrm{Cov}_{\mu_b} = B := \begin{bmatrix} 1 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}$$

En effet, si $X_1 = (X, Y) \in \mathbb{R}^2$. On a $\mathbb{E}\big[X^2\big] = 1$, $\mathbb{E}\big[Y^2\big] = \frac{1}{2}$ et $\mathbb{E}\big[XY\big] = \frac{1}{4} * 1^2 + \frac{1}{4} * (-1)^2 = \frac{1}{2}$. On peut donc appliquer le TCL vectoriel pour obtenir que $S_n/\sqrt{n}$ converge **en loi** vers un vecteur Gaussien **centré** de matrice de covariance B.
4. Même argumentaire, i.e. L^2, indépendance, moyenne nulle, et cette fois la matrice de covariance des incréments est donnée par

$$\mathrm{Cov}_{\mu_c} = C := \begin{bmatrix} 1/2 & 0 \\ 0 & 1/2 \end{bmatrix}$$

En effet, on a, si $X_1 = (X, Y)$, $\mathbb{E}\big[X^2\big] = 2/6 * 1^2 + 4/6 * (1/2)^2 = 1/2$. A partir de la, on peut soit calculer $\mathbb{E}\big[Y^2\big]$ puis $\mathbb{E}\big[XY\big]$ (méthode la plus sûre mais plus fatiguante) soit raisonner géométriquement en observant que le vecteur Gaussien limite se doit d'être invariant par $e^{i\pi/3}$, cela force la matrice de covariance à être multiple de l'identité et évite tout autre calcul. On peut donc appliquer le TCL vectoriel pour obtenir que $S_n/\sqrt{n}$ converge **en loi** vers un vecteur Gaussien **centré** de matrice de covariance C.
5. On raisonne comme dans l'exercice 2. I.e. en différentiant les pas pairs des pas impairs. En prime, à la différence de l'exercice 2. on considère ici le TCL plutot que la loi des grands nombres. Les pas impairs sont centrés, L^2, i.i.d et leur matrice de covariance est donnée par

$$\mathrm{Cov}_{\mu_{d-Impair}} = D_{Impair} := \begin{bmatrix} 1/2 & 0 \\ 0 & 1/2 \end{bmatrix}$$

En effet, on a $\mathbb{E}\big[X^2\big] = \frac{1}{3} * 1^2 + \frac{2}{3} * (1/2)^2 = \frac{1}{2}$. On peut ensuite calculer ou raisonner géométriquement (invariance par $e^{2i\pi/3}$ cette fois) pour obtenir la matrice ci-dessus. La matrice de covariance D_{Pair} est la même (a nouveau soit par calcul soit géométriquement).
Les termes pair et impair convergent chacun en loi vers $1/\sqrt{2}$ fois un vecteur Gaussien centré de matrice de covariance $D_{Impair} = D_{Pair}$. De même que dans l'exercice 4. La somme de deux vecteurs Gaussiens centrés, de même loi et **indépendants** (les incréments pairs et impairs sont indépendants) divisé par $\sqrt{2}$ est (en loi) le même vecteur Gaussien. On obtient donc que $S_n/\sqrt{n}$ converge **en loi** vers un vecteur Gaussien **centré** de matrice de covariance $D = D_{Impair}$.