Probabilités

Devoir 3

A rendre fin avril

Problème : **Urne aléatoire** Ce problème est divisé en deux parties indépendantes.

Première partie : Urne de Polya A l'instant 1, une urne contient a boules blanches et $b = N_0 - a$ boules rouges. A chaque étape, on tire une boule uniformément dans l'urne, et on la remplace par deux boules de la même couleur que celle tirée. Pour $n \in \mathbb{N}$, on note Y_n le nombre de boules blanches dans l'urne, et

$$X_n = \frac{Y_n}{N_0 + n - 1}$$

la proportion de boules blanches. Soit $\mathcal{F}_n = \sigma(X_0, \cdots X_n)$.

(1) Calculer $\mathbb{P}(Y_{n+1} = Y_n | \mathcal{F}_n)$, et montrer que (X_n) est une martingale qui converge p.s. vers une variable aléatoire U. Montrer que pour tout $k \in \mathbb{N}$, on a

$$\lim_{n \to +\infty} \mathbb{E}(X_n^k) = \mathbb{E}(U^k).$$

(2) On suppose dans cette question uniquement que $a = b = 1$. Montrer par récurrence que Y_n suit la loi uniforme sur $\{1, \dots n\}$, en déduire la loi de U.

(3) On fixe $k \geq 1$, et on pose, pour tout $n \in \mathbb{N}$

$$Z_n = \frac{Y_n(Y_n + 1) \cdots (Y_n + k - 1)}{(n + N_0 - 1)(n + N_0) \cdots (n + N_0 + k - 2)}$$

Montrer que (Z_n) est une $(\mathcal{F}_n)$-martingale, et en déduire la valeur de $\mathbb{E}(U^k)$.

(4) Montrer que la fonction caractéristique d'une variable aléatoire réelle bornée est développable en séries entières, en déduire qu'on a caractérisé la loi de U.

(5) On rappelle que la loi β de paramètre a et b admet pour densité par rapport à la mesure de Lebesgue

$$\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} u^{a-1}(1-u)^{b-1} \mathbf{1}_{\{u \in [0,1]\}}.$$

Calculer les moments entiers d'une variable aléatoire de loi β de paramètres a et b.

(6) En déduire la loi de U.

Solution. 1. D'après l'énoncé, la probabilité de tirer une boule blanche conditionnellement à $\mathcal{F}_n$ est égale à la proportion de boules blanches X_n, par conséquent,

$$\mathbb{P}(Y_{n+1} = Y_n | \mathcal{F}_n) = 1 - X_n.$$

Par conséquent, on a

$$\mathbb{E}(Y_{n+1} | \mathcal{F}_n) = Y_n + \mathbb{P}(Y_{n+1} = Y_n + 1 | \mathcal{F}_n) = Y_n + X_n = (N_0 + n) X_n,$$

donc X est une martingale, bornée par 0 et 1, qui converge donc p.s. et dans L^p pour tout p. En particulier, pour tout $k \in \mathbb{N}$, on a

$$\lim_{n \to +\infty} \mathbb{E}(X_n^k) = \mathbb{E}(U^k).$$

2. Y_1 suit bien la loi uniforme sur 1. On suppose la propriété vraie au temps n, on a alors

$$\begin{aligned}\mathbb{P}(Y_{n+1} = k) = \mathbb{E}(\mathbb{P}(Y_{n+1} = k | \mathcal{F}_n)) &= \mathbb{E}(\mathbf{1}_{\{Y_n = k-1\}} X_n + \mathbf{1}_{\{Y_n = k\}} (1 - X_n)) \\ &= \frac{1}{n} \left[\frac{k-1}{n+1} + \frac{n+1-k}{n+1} \right]\end{aligned}$$

ce qui conclut à la récurrence. Or (X_n) converge p.s. donc en loi vers U, et pour toute fonction continue bornée f, on a

$$\mathbb{E}(f(X_n)) = \frac{1}{n}\sum_{k=1}^{n} f(k/n) \to \int_0^1 f(s)ds,$$

ce qui montre que U est uniforme sur $[0,1]$.

3. On observe que

$$\begin{aligned}
\mathbb{E}(Z_{n+1}|\mathcal{F}_n) &= \mathbb{E}\left[\left.\frac{Y_{n+1}(Y_{n+1}+1)\cdots(Y_{n+1}+k-1)}{(n+N_0)\cdots(n+N_0+k-1)}\right|\mathcal{F}_n\right] \\
&= \frac{Y_n(Y_n+1)\cdots(Y_n+k-1)}{(n+N_0)\cdots(n+N_0+k-1)} \times (1-X_n) + \frac{(Y_n+1)\cdots(Y_n+k)}{(n+N_0)\cdots(n+N_0+k-1)} \times X_n \\
&= \frac{Y_n(Y_n+1)\cdots(Y_n+k-1)}{(n+N_0-1)\cdots(n+N_0+k-1)} = Z_n.
\end{aligned}$$

Par conséquent, (Z_n) est une martingale, à valeurs dans $[0,1]$, donc converge p.s. et dans L^1 vers Z_∞. D'autre part, quand $n \to +\infty$, on a $Z_n \sim X_n^k$, donc $X_n^k \to Z_\infty$ p.s. d'où $Z_\infty = U^k$. On en déduit

$$\mathbb{E}(U^k) = \mathbb{E}(Z_0) = \frac{a(a+1)\cdots(a+k-1)}{(N_0)(N_0+1)\cdots(N_0+k-1)}$$

4. On utilise le théorème de Fubini. On observe que d'une part

$$\mathbb{E}\left(\sum_{k=0}^{+\infty}\left|\frac{(i\xi X)^k}{k!}\right|\right) \le \sum_{k=0}^{+}\infty\frac{(\xi M)^k}{k!} < +\infty,$$

et d'autre part

$$\begin{aligned}
\mathbb{E}\left(e^{i\xi X}\right) &= \mathbb{E}\left(\sum_{k=0}^{+\infty}\frac{(i\xi X)^k}{k!}\right) \\
&= \sum_{k=0}^{+\infty}(i\xi)^k\mathbb{E}\left(X^k\right),
\end{aligned}$$

ce qui conclut. Or on connaît tous les moments entiers de U, qui est une variable aléatoire bornée. On connaît donc sa transformée de Laplace, donc sa loi.

5. On rappelle que

$$\int_0^1 u^{a-1}(1-u)^{b-1}du = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}.$$

Par conséquent, on a

$$\begin{aligned}
\mathbb{E}(\beta(a,b)^k) &= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\int_0^1 u^{a+k-1}(1-u)^{b-1}du \\
&= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\frac{\Gamma(a+k)\Gamma(b)}{\Gamma(a+b+k)} \\
&= \frac{a(a+1)\cdots(a+k-1)}{(a+b)(a+b+1)\cdots(a+b+k-1)}
\end{aligned}$$

par propriétés de la fonction Γ –en particulier $\Gamma(x+1) = x\Gamma(x)$.

6. Coïncidence, si on calcule

$$\lim_{n\to+\infty}\mathbb{E}(X_n^k) = \frac{a(a+1)\cdots(a+k-1)}{(a+b)(a+b+1)\cdots(a+b+k-1)},$$

grâce à la question 4, on en déduit que U suit une loi $\beta(a,b)$, car les transformées de Fourier sont égales. $\square$

Deuxième partie : Urne triangulaire On considère un processus légèrement modifié. A l'instant 1, il y a 1 boule blanche et 0 boules rouges. A chaque instant n, on tire uniformément au hasard une boule dans l'urne. Si cette boule est blanche, on la replace accompagnée d'une boule blanche et d'une boule rouge, si la boule est rouge on la replace avec deux nouvelles boules rouges. On note encore Y_n le nombre de boules blanches, et $\mathcal{F}_n = \sigma(Y_1, \dots Y_n)$.

(1) Déterminer le nombre a_n de boules dans l'urne à l'instant n, et montrer que $\frac{Y_n}{\prod_{j=1}^n \left(1+\frac{1}{a_j}\right)}$ est une martingale.

(2) En déduire que $\frac{Y_n}{\sqrt{n}}$ converge p.s. vers une variable aléatoire $U \in L^1$.

(3) Déterminer c_n tel que $\frac{Y_n(Y_n+1)}{c_n}$ est une martingale, et montrer que

$$\lim_{n\to+\infty} \frac{Y_n(Y_n+1)}{c_n} = U^2.$$

(4) Proposer une manière d'identifier la loi de U.

Solution. 1. On observe qu'on ajoute à chaque étape deux boule dans l'urne. On a donc $a_{n+1} = a_n + 2$ et $a_1 = 1$, par conséquent, $a_n = 2n - 1$. La probabilité de tirer une boule blanche a l'étape n est donc de $\frac{Y_n}{a_n}$, on a

$$\mathbb{E}\{()\, Y_{n+1}|\mathcal{F}_n) = Y_n + 1 \times \frac{Y_n}{a_n} = Y_n(1 + \tfrac{1}{a_n}).$$

Par conséquent, $Z_n = \frac{Y_n}{\prod_{j=1}^n \left(1+\frac{1}{a_j}\right)}$ est une martingale.

2. Z_n est une martingale positive, qui converge donc p.s. vers une variable aléatoire Z qui est intégrable (par Fatou). De plus, on a

$$\log \prod_{j=1}^n \left(1 + \frac{1}{a_j}\right) = \sum_{j=1}^n \log(1 + \tfrac{1}{2j-1}) = \frac{1}{2}\log n + C + o(1),$$

par conséquent, $\frac{Y_n}{\sqrt{n}}$ converge p.s. vers $U = e^C Z$.

3. On calcule maintenant

$$\begin{aligned}
\mathbb{E}\{[\}\, Y_{n+1}(Y_{n+1}+1)|\mathcal{F}_n] &= Y_n(Y_n+1) \times (1 - \tfrac{Y_n}{a_n}) + (Y_n+1)(Y_n+2) \times \tfrac{Y_n}{a_n} \\
&= Y_n(Y_n+1) \times \left[1 - \tfrac{Y_n}{a_n} + \tfrac{Y_n+2}{a_n}\right] \\
&= Y_n(Y_n+1)\left[1 + \frac{2}{2n-1}\right].
\end{aligned}$$

Par conséquent, $c_n = \prod_{j=1}^n (1 + \frac{2}{2j-1}) \sim Cn$. De plus, on a

$$\frac{Y_n(Y_n+1)}{c_n} \sim \frac{Y_n^2}{Cn} \sim \frac{U^2}{C}.$$

4. Grâce à la question 3. on sait que $\mathbb{E}\{()\, Y_n^2)$ est bornée, donc on a convergence, p.s. et dans L^2 de Y_n. On pourrait de la même façon calculer les moments d'ordre p de la variable aléatoire limite U. Si ces moments ont une décroissance assez rapide, le même argument que précédemment permettra d'identifier la transformée de Fourier de la loi – qui pour la petite histoire, a pour densité

$$\sum_{k=0}^{+\infty} \frac{(-1)^k x^k}{\Gamma(\frac{-k}{2}) k!}.$$

$\square$

Exercice : On considère $X_0 = 0$, et $(X_n)_{n\geq 1}$ une suite de variables aléatoires réelles i.i.d. de loi normale centrée réduite. On introduit

$$Y_i = \frac{X_i - X_{i-1}}{i}, i \geq 1.$$

On note $\mathcal{F}_n = \sigma(X_i, i \leq n)$, et on pose

$$S_n = \sum_{i=1}^{n} Y_i.$$

(1) Montrer que $\mathbb{E}(S_{n+1}|\mathcal{F}_n) = S_n - \frac{X_n}{n+1}$.
(2) En déduire que $\bar{S}_n = S_n - \frac{X_n}{n+1}$ est une $\mathcal{F}_n$-martingale.
(3) Trouver une suite de réels positifs (u_j) telle que pour tout $n \geq 1$, on a

$$\bar{S}_n = \sum_{i=1}^{n} u_i X_i$$

en déduire une nouvelle démonstration de 2.
(4) Calculer $\mathbb{E}(|X_1|)$, et montre que pour tout $n \in \mathbb{N}$

$$\mathbb{E}([|\bar{S}_n|] \leq \sqrt{\frac{2}{\pi}} \sum_{j=1}^{n} \frac{1}{j(j+1)}.$$

(5) En déduire que lorsque $n \to +\infty$, $\bar{S}_n$ converge presque sûrement vers une variable aléatoire $\bar{S}_\infty$. Montrer alors que S_n converge également vers $\bar{S}_\infty$ presque sûrement.
(6) En utilisant la question 4, montrer que $\bar{S}_\infty$ est une variable aléatoire gaussienne centrée dont on déterminera la variance.

Solution. 1. On a

$$\mathbb{E}(S_{n+1}|\mathcal{F}_n) = S_n + \mathbb{E}(Y_{n+1}|\mathcal{F}_n) = S_n - \frac{X_n}{n+1}.$$

2. On a

$$\mathbb{E}(\bar{S}_{n+1}|\mathcal{F}_n) = S_n - \frac{X_n}{n+1} + \mathbb{E}\left(\frac{X_{n+1}}{n+2}\middle| \mathcal{F}_n\right) = \bar{S}_n.$$

3. On observe que

$$\bar{S}_n = \sum_{k=1}^{n} Y_k - \frac{X_n}{n+1} = \sum_{k=1}^{n} \frac{X_k - X_{k-1}}{k} - \frac{X_n}{n+1} = \sum_{j=1}^{n} X_j \left(\frac{1}{j} - \frac{1}{j+1}\right),$$

donc $u_j = \frac{1}{j(j+1)}$. On en redéduit immédiatement que $\bar{S}_n$ est une martingale, comme somme de n variables aléatoires indépendantes centrées.
4. On a

$$\mathbb{E}(|X_1|) = \frac{2}{\sqrt{2\pi}} \int_0^{+\infty} xe^{-\frac{x^2}{2}}\,dx = \sqrt{\frac{2}{\pi}}.$$

Dès lors,

$$\mathbb{E}(|\bar{S}_n|) \leq \sum_{k=1}^{n} \mathbb{E}(|u_i X_i|) \leq \sqrt{\frac{2}{\pi}} \sum_{j=1}^{n} \frac{1}{j(j+1)} = \sqrt{\frac{2}{\pi}} \left(1 - \frac{1}{n+1}\right).$$

5. La suite $\bar{S}$ est une martingale bornée dans L^1, qui converge donc p.s. vers une variable aléatoire $\bar{S}_\infty$. De plus, comme $\mathbb{E}\{()\,|X_1|) < +\infty$, par Borel-Cantelli, on a $\frac{X_n}{n} \to 0$ p.s. donc S converge également vers $\bar{S}_\infty$.
6. On observe que $\bar{S}_n$ est une variable aléatoire de loi gaussienne centrée de variance $\sigma_n^2 = \sum_{j=1}^{n} \frac{1}{(j(j+1))^2}$. Or

$$\frac{1}{(j(j+1))^2} = \left(\frac{1}{j} - \frac{1}{j+1}\right)^2 = \frac{1}{j^2} + \frac{1}{(j+1)^2} + \frac{2}{j} - \frac{2}{j+1}.$$

Par conséquent, on a

$$\lim_{n\to+\infty} \sigma_n^2 = \left(2\frac{\pi^2}{6} - 1\right) + 2 = \frac{\pi^2}{3} + 1.$$

Dès lors, on a

$$\lim_{n\to+\infty} \mathbb{E}\left[e^{i\xi\bar{S}_n}\right] = e^{-\frac{\xi^2(\frac{\pi^2}{3}+1)}{2}},$$

donc $\bar{S}_n$ converge en loi vers une variable aléatoire gaussienne centrée de variance $\frac{\pi^2}{3}+1$. On en conclut donc $\bar{S}_\infty \sim \mathcal{N}(0, \frac{\pi^2}{3}+1)$. $\square$