

HEC-ESCP – Mathématiques II

Mathématiques approfondies

Corrigé du sujet adapté – 2023

Les questions et sous-questions conservent la numérotation de l'énoncé. Les notations $\mathbb{E}$ et $\mathbb{V}$ désignent l'espérance et la variance. La transposée d'une matrice A est notée ${}^tA$.

Questions préliminaires

On note $[t] = \lfloor t \rfloor$ la partie entière de t . Les indices des sommes et de la marche aléatoire sont toujours des entiers.

Corrigé – Questions préliminaires.

1. Approximation d'un réel par sa partie entière.

Pour $n \geq 1$, $nx - 1 < [nx] \leq nx$. En divisant par $n > 0$,

$$x - \frac{1}{n} < \frac{[nx]}{n} \leq x.$$

Les deux bornes tendent vers x ; le théorème d'encadrement donne

$$\lim_{n \rightarrow +\infty} \frac{[nx]}{n} = x \quad (x \in \mathbb{R}).$$

2. Un indice tendant vers l'infini.

Comme $[n/2] \leq n/2$, on a $n - [n/2] \geq n/2$. La minoration tend vers $+\infty$, donc

$$n - [n/2] \xrightarrow{n \rightarrow +\infty} +\infty.$$

Partie I – La loi arcsinus

$\arcsin$ est la réciproque du sinus sur $[-\pi/2, \pi/2]$. On pose $G(x) = 2 \arcsin \sqrt{x}$ sur $[0, 1]$ et $g(x) = 1/\sqrt{x(1-x)}$ sur $]0, 1[$. La loi arcsinus est définie par la densité $f = g/\pi$ sur $]0, 1[$, nulle ailleurs.

Corrigé – Partie I – La loi arcsinus.

3. La fonction arcsinus.

- a) Si $y = \arcsin x$, alors $y \in [-\pi/2, \pi/2]$, $\sin y = x$ et $\cos y \geq 0$. L'identité $\sin^2 y + \cos^2 y = 1$ donne donc

$$\cos(\arcsin x) = \sqrt{1 - x^2} \quad (x \in [-1, 1]).$$

- b) Le sinus est continu et strictement croissant sur $[-\pi/2, \pi/2]$. Le théorème de la bijection assure la continuité de sa réciproque sur $[-1, 1]$. Pour $y \in]-\pi/2, \pi/2[$,

sa dérivée $\cos y$ est strictement positive. Le théorème de dérivation d'une fonction réciproque s'applique sur les intervalles ouverts et donne

$$\arcsin \text{ est continue sur } [-1, 1], \quad (\arcsin)'(x) = \frac{1}{\sqrt{1-x^2}} \quad (-1 < x < 1).$$

4. Étude de G et de g .

- a) La racine carrée est continue sur $[0, 1]$ et à valeurs dans $[0, 1]$; G est donc continue. Sur $]0, 1[$, les règles de dérivation des fonctions d'une variable donnent

$$G'(x) = 2 \frac{1}{\sqrt{1-x}} \frac{1}{2\sqrt{x}}, \quad G'(x) = g(x) = \frac{1}{\sqrt{x(1-x)}}.$$

On a également $G(0) = 0$ et $G(1) = \pi$.

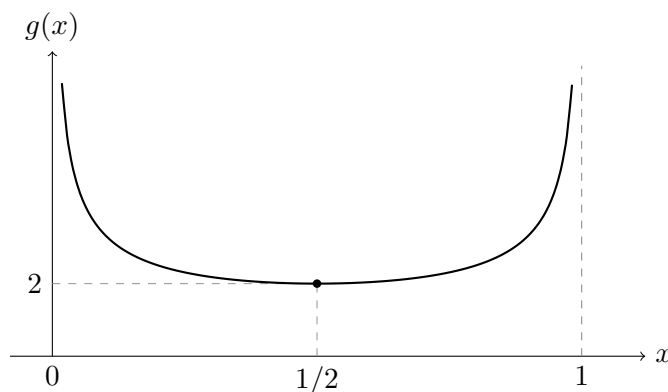
- b) g est positive et de classe $\mathcal{C}^1$ sur $]0, 1[$, avec

$$g'(x) = \frac{2x-1}{2[x(1-x)]^{3/2}}.$$

Elle est strictement décroissante sur $]0, 1/2]$, puis strictement croissante sur $[1/2, 1[$. Les deux limites aux extrémités valent $+\infty$, et $g(1-x) = g(x)$. Ainsi

$$\min_{0 < x < 1} g(x) = g(1/2) = 2.$$

Le graphe est symétrique par rapport à $x = 1/2$ et présente les asymptotes verticales $x = 0$ et $x = 1$.



5. Densité et espérance de la loi arcsinus.

- a) f est positive ou nulle, continue sur $\mathbb{R}$ sauf en 0 et 1. Pour $0 < a < b < 1$,

$$\int_a^b f(t) dt = \frac{G(b) - G(a)}{\pi}.$$

Par continuité de G aux extrémités, l'intégrale impropre converge et

$$\int_{-\infty}^{+\infty} f(t) dt = \frac{G(1) - G(0)}{\pi} = 1; \quad f \text{ est une densité.}$$

- b) Comme $0 \leq X \leq 1$ presque sûrement, $\mathbb{E}(|X|) \leq 1$: son espérance existe. Posons $I = \int_0^1 xf(x) dx = \mathbb{E}(X)$. Le changement $u = 1 - x$ est une bijection de classe $\mathcal{C}^1$ de $]0, 1[$ sur lui-même, de dérivée -1 . Appliqué d'abord sur les segments intérieurs puis aux bornes, il conserve la nature des intégrales impropres. Puisque $f(1 - u) = f(u)$ sur cet intervalle,

$$I = \int_0^1 (1 - u)f(1 - u) du = \int_0^1 f(u) du - \int_0^1 uf(u) du = 1 - I.$$

Ainsi

$$\boxed{\mathbb{E}(X) = \frac{1}{2}.$$

6. Simulation de la loi arcsinus.

- a) La fonction $u \mapsto \sin^2(\pi u/2)$ est strictement croissante de $[0, 1]$ sur $[0, 1]$. Pour $0 < x < 1$,

$$\mathbb{P}(V \leq x) = \mathbb{P}\left(U \leq \frac{2}{\pi} \arcsin \sqrt{x}\right) = \frac{2}{\pi} \arcsin \sqrt{x} = \frac{G(x)}{\pi}.$$

Aux extrémités, $\mathbb{P}(V \leq 0) = 0$ et $\mathbb{P}(V \leq 1) = 1$. Ainsi

$$\boxed{F_V(x) = \begin{cases} 0 & x \leq 0, \\ G(x)/\pi & 0 < x < 1, \\ 1 & x \geq 1. \end{cases}}$$

Cette fonction est continue et dérivable sauf éventuellement en $0, 1$; sa dérivée sur $]0, 1[$ vaut $g(x)/\pi = f(x)$. Par conséquent V suit bien la loi arcsinus.

- b)

```
import numpy as np
import numpy.random as rd
```

```
def arcsinus():
    U = rd.random()
    V = np.sin(np.pi*U/2)**2
    return V
```

La première ligne simule une variable uniforme, la seconde lui applique la transformation de la question précédente.

`arcsinus()` simule la loi de densité f .

7. Encadrement des termes d'une somme.

Si $1 \leq k \leq n/2 - 1$, les deux intervalles $[(k-1)/n, k/n]$ et $[k/n, (k+1)/n]$ se trouvent dans $[0, 1/2]$, où g décroît. Sur le second,

$$\int_{k/n}^{(k+1)/n} g(t) dt \leq \frac{1}{n} g(k/n) = \frac{1}{\sqrt{k(n-k)}}.$$

Sur le premier, l'intégrale est au moins $(1/n)g(k/n)$. Lorsque $k = 1$, cette comparaison se justifie en intégrant d'abord depuis $\varepsilon > 0$ et en faisant $\varepsilon \rightarrow 0$; l'intégrale converge par la question 5. Comme G est une primitive de g ,

$$\boxed{G\left(\frac{k+1}{n}\right) - G\left(\frac{k}{n}\right) \leq \frac{1}{\sqrt{k(n-k)}} \leq G\left(\frac{k}{n}\right) - G\left(\frac{k-1}{n}\right).$$

Si $n/2 + 1 \leq k \leq n - 1$, les deux intervalles sont inclus dans $[1/2, 1]$, où g croît. Les comparaisons sont donc inversées ; en 1, on passe également par une borne intérieure avant la limite :

$$G\left(\frac{k+1}{n}\right) - G\left(\frac{k}{n}\right) \geq \frac{1}{\sqrt{k(n-k)}} \geq G\left(\frac{k}{n}\right) - G\left(\frac{k-1}{n}\right).$$

8. Limite d'une somme tronquée.

Un exemple est

$$a_n = \lfloor \sqrt{n} \rfloor.$$

Cette suite tend vers $+\infty$ et $a_n/n \rightarrow 0$. Si $x > 0$, on a $a_n < [nx]$ pour n assez grand puisque $[nx]/n \rightarrow x$.

Fixons maintenant $0 < x < 1/2$ et une suite quelconque satisfaisant les trois conditions de l'énoncé. Pour n assez grand, tous les indices de a_n à $m_n = [nx]$ vérifient $1 \leq k \leq n/2 - 1$. En sommant les inégalités de la question 7, les différences de G se simplifient :

$$G\left(\frac{m_n+1}{n}\right) - G\left(\frac{a_n}{n}\right) \leq \sum_{k=a_n}^{m_n} \frac{1}{\sqrt{k(n-k)}} \leq G\left(\frac{m_n}{n}\right) - G\left(\frac{a_n-1}{n}\right).$$

Les arguments de G tendent respectivement vers x et 0. Sa continuité et $G(0) = 0$ montrent que les deux bornes tendent vers $G(x)$. Par encadrement,

$$\lim_{n \rightarrow +\infty} \sum_{k=a_n}^{[nx]} \frac{1}{\sqrt{k(n-k)}} = G(x).$$

Partie II – Sommes de variables de valeurs -1 et 1 – Estimation d'un paramètre et dernier passage en zéro

Les variables X_i sont indépendantes de même loi, $\mathbb{P}(X_i = 1) = \theta$ et $\mathbb{P}(X_i = -1) = 1 - \theta$, avec $0 < \theta < 1$. On pose $S_0 = 0$ et $S_n = \sum_{i=1}^n X_i$. Les intervalles d'indices $[0, n]$ ou $[1, n]$ de cette partie sont interprétés comme des ensembles d'entiers, puisque S_k n'est défini qu'aux temps entiers. Ainsi $L_n = \max\{k \in \{0, \dots, n\} : S_{2k} = 0\}$. Cet ensemble est non vide car $S_0 = 0$, et son maximum existe.

Corrigé – Partie II – Sommes de variables de valeurs -1 et 1 .
9. Transformation en variables de Bernoulli.

Le changement $Y_i = (X_i + 1)/2$ transforme les valeurs -1 et 1 en 0 et 1 . Les probabilités correspondantes sont $1 - \theta$ et θ :

$$Y_i \sim \mathcal{B}(1, \theta), \quad \mathbb{E}(Y_i) = \theta, \quad \mathbb{V}(Y_i) = \theta(1 - \theta).$$

Les Y_i sont mutuellement indépendantes, car elles sont des fonctions séparées des X_i indépendantes. Leur somme vérifie

$$\sum_{i=1}^n Y_i = \frac{S_n + n}{2}.$$

La somme de n variables de Bernoulli indépendantes de même paramètre suit donc une loi binomiale :

$$\frac{S_n + n}{2} \sim \mathcal{B}(n, \theta), \quad \mathbb{P}(S_n = 2j - n) = \binom{n}{j} \theta^j (1 - \theta)^{n-j} \quad (0 \leq j \leq n).$$

10. Les variables S_n sont-elles indépendantes ?

Les variables S_1 et S_2 ne sont pas indépendantes. En effet, $S_1 = 1$ signifie $X_1 = 1$, et $S_2 = 2$ signifie $X_1 = X_2 = 1$. L'indépendance des X_i donne

$$\mathbb{P}(S_1 = 1, S_2 = 2) = \theta^2, \quad \mathbb{P}(S_1 = 1)\mathbb{P}(S_2 = 2) = \theta \cdot \theta^2 = \theta^3.$$

Puisque $0 < \theta < 1$, ces deux nombres sont différents. L'indépendance mutuelle des S_n impliquerait celle de cette paire ; donc

$$(S_n)_{n \geq 1} \text{ n'est pas une famille de variables mutuellement indépendantes.}$$

11. Intervalle de confiance asymptotique pour θ .

Posons

$$\hat{\theta}_n = \frac{1}{n} \sum_{i=1}^n Y_i = \frac{1 + \bar{S}_n}{2}, \quad Z_n = \frac{\sqrt{n}(\hat{\theta}_n - \theta)}{\sqrt{\theta(1 - \theta)}}.$$

Les Y_i sont indépendantes de même loi, de variance finie non nulle. Le théorème central limite donne la convergence en loi de Z_n vers une normale centrée réduite : si F_n est sa fonction de répartition,

$$\forall z \in \mathbb{R}, \quad F_n(z) \longrightarrow \Phi(z).$$

On a $\theta(1 - \theta) = 1/4 - (\theta - 1/2)^2 \leq 1/4$, donc $\sqrt{\theta(1 - \theta)} \leq 1/2$. L'intervalle proposé est

$$I_n = \left[\hat{\theta}_n - \frac{t_\alpha}{2\sqrt{n}}, \hat{\theta}_n + \frac{t_\alpha}{2\sqrt{n}} \right].$$

Pour le paramètre θ fixé, l'événement de couverture s'écrit

$$[\theta \in I_n] = [|Z_n| \leq c_\theta], \quad c_\theta = \frac{t_\alpha}{2\sqrt{\theta(1-\theta)}} \geq t_\alpha.$$

Ce seuil est constant en n ; on peut donc utiliser directement la convergence des fonctions de répartition. Pour justifier aussi les bornes incluses, pour chaque $\eta > 0$,

$$F_n(c_\theta) - F_n(-c_\theta) \leq \mathbb{P}(-c_\theta \leq Z_n \leq c_\theta) \leq F_n(c_\theta) - F_n(-c_\theta - \eta).$$

Les expressions aux bornes convergent vers $\Phi(c_\theta) - \Phi(-c_\theta)$ et $\Phi(c_\theta) - \Phi(-c_\theta - \eta)$. La continuité de Φ , puis $\eta \rightarrow 0^+$, donnent

$$\mathbb{P}(\theta \in I_n) \longrightarrow \Phi(c_\theta) - \Phi(-c_\theta) = 2\Phi(c_\theta) - 1.$$

Comme $\Phi(t_\alpha) = 1 - \alpha/2$, il vient

$$\lim_{n \rightarrow +\infty} \mathbb{P}(\theta \in I_n) = 2\Phi\left(\frac{t_\alpha}{2\sqrt{\theta(1-\theta)}}\right) - 1 \geq 1 - \alpha.$$

Ainsi l'intervalle de l'énoncé est un intervalle de confiance asymptotique de niveau *au moins* $1 - \alpha$:

$$I_n = \left[\frac{1}{2} \left(1 - \frac{t_\alpha}{\sqrt{n}} + \bar{S}_n \right), \frac{1}{2} \left(1 + \frac{t_\alpha}{\sqrt{n}} + \bar{S}_n \right) \right].$$

La couverture limite vaut exactement $1 - \alpha$ pour $\theta = 1/2$. Pour $\theta \neq 1/2$, $c_\theta > t_\alpha$, et la stricte croissance de Φ donne une couverture limite strictement supérieure à $1 - \alpha$. La majoration utilisée ne permet donc pas d'affirmer une égalité de couverture asymptotique pour tous les paramètres.

12. Probabilités de passage en zéro.

- a) D'après 9, les valeurs de S_m sont les $2j - m$, donc ont la même parité que m . Pour $m = 2k + 1$ impair, aucune de ces valeurs n'est zéro.

$$\forall k \geq 0, \quad \mathbb{P}(S_{2k+1} = 0) = 0.$$

- b) Pour $k \geq 1$, $S_{2k} = 0$ équivaut à $(S_{2k} + 2k)/2 = k$. La loi binomiale de 9 donne

$$\mathbb{P}(S_{2k} = 0) = \binom{2k}{k} \theta^k (1 - \theta)^k.$$

Pour $k = 0$, $S_0 = 0$ et le membre de droite vaut 1 ; la formule reste donc vraie.

13. Dernier passage à l'instant $2n$.

Puisque n est le plus grand indice admissible dans la définition de L_n , on a $[L_n = n] = [S_{2n} = 0]$. En appliquant 12.b),

$$\mathbb{P}(L_n = n) = \binom{2n}{n} \theta^n (1 - \theta)^n.$$

14. Événement $[L_n = k]$ et probabilité conditionnelle.

Fixons un entier $k \in \{0, \dots, n-1\}$.

- a) L'égalité $L_n = k$ signifie que $S_{2k} = 0$ et qu'aucun des indices pairs strictement postérieurs, jusqu'à $2n$, ne donne zéro. Les indices impairs ne peuvent pas donner zéro, par 12.a). On peut donc inclure tous les indices postérieurs dans l'intersection :

$$[L_n = k] = [S_{2k} = 0] \cap \bigcap_{j=2k+1}^{2n} [S_j \neq 0].$$

Cette identité n'impose aucune condition aux sommes avant $2k$: des passages antérieurs en zéro peuvent avoir eu lieu.

- b) Posons $A = [S_{2k} = 0]$ et, pour $1 \leq j \leq 2(n-k)$,

$$T_j = \sum_{i=2k+1}^{2k+j} X_i.$$

Sur l'événement A , on a $S_{2k+j} = T_j$. Ainsi l'intersection de A avec l'événement étudié coïncide avec $A \cap B$, où $B = \bigcap_{j=1}^{2(n-k)} [T_j \neq 0]$. L'événement A ne dépend que des $2k$ premières variables, tandis que B dépend des variables suivantes. L'indépendance mutuelle des X_i donne donc l'indépendance de A et B . Comme $\mathbb{P}(A) > 0$ par 12.b), la probabilité conditionnelle existe et

$$\mathbb{P}_A \left(\bigcap_{j=2k+1}^{2n} [S_j \neq 0] \right) = \mathbb{P}(B).$$

Le vecteur $(X_{2k+1}, \dots, X_{2n})$ a la même loi que $(X_1, \dots, X_{2(n-k)})$: dans les deux cas les coordonnées sont indépendantes de même loi. Leurs sommes partielles ont donc la même loi.

$$\mathbb{P}_{[S_{2k}=0]} \left(\bigcap_{j=2k+1}^{2n} [S_j \neq 0] \right) = \mathbb{P} \left(\bigcap_{j=1}^{2(n-k)} [S_j \neq 0] \right) = p_{n-k}.$$

Pour $k = 0$, $A = \Omega$ puisque $S_0 = 0$, et le même résultat s'applique sans condition sur des variables antérieures.

15. Loi de L_n en fonction de p_j .

Pour $k < n$, les deux questions précédentes donnent

$$\mathbb{P}(L_n = k) = \mathbb{P}(S_{2k} = 0) \mathbb{P}_{[S_{2k}=0]} \left(\bigcap_{j=2k+1}^{2n} [S_j \neq 0] \right).$$

On remplace les deux facteurs par 12.b) et 14.b). Pour $k = n$, la formule de 13 est identique à celle-ci en adoptant la convention $p_0 = 1$.

$$\forall k \in \{0, \dots, n\}, \quad \mathbb{P}(L_n = k) = \binom{2k}{k} \theta^k (1-\theta)^k p_{n-k}, \quad p_0 = 1.$$

Le cas $k = 0$ donne notamment $\mathbb{P}(L_n = 0) = p_n$.

16. Valeurs visitées et séparation des signes.

À partir de cette question, on suppose $\theta = 1/2$ comme dans l'énoncé.

- a) Fixons une réalisation ω et un entier $r \geq 1$. Notons $s_j = S_j(\omega)$. On a $s_{j+1} - s_j \in \{-1, 1\}$. Au rang $r = 1$, l'ensemble $\{s_1\}$ est un intervalle d'entiers. Supposons $\{s_1, \dots, s_r\} = \{a, a+1, \dots, b\}$. Le nouveau point s_{r+1} est voisin de s_r , qui appartient à cet intervalle. Il appartient donc à l'intervalle existant, ou vaut $a - 1$, ou vaut $b + 1$. Dans les trois cas, l'ensemble augmenté est encore un intervalle d'entiers. Par récurrence,

$$\{S_1(\omega), \dots, S_r(\omega)\} \text{ est un intervalle d'entiers pour tout } r \geq 1.$$

- b) Sur D_n , aucune des valeurs $S_1, \dots, S_{2n}$ n'est nulle. Si une valeur était positive et une autre négative, l'intervalle d'entiers de a) contiendrait 0, contradiction. Toutes les valeurs ont donc le même signe : $D_n = D_n^+ \cup D_n^-$. Les deux événements sont disjoints, d'où

$$D_n = D_n^+ \dot{\cup} D_n^-, \quad \mathbb{P}(D_n) = \mathbb{P}(D_n^+) + \mathbb{P}(D_n^-).$$

17. Décomposition de D_n^+ selon la valeur finale.

Sur D_n^+ , la valeur finale S_{2n} est un entier pair, strictement positif et au plus $2n$. Elle est donc de la forme $2r$ pour un unique entier $r \in \{1, \dots, n\}$. Pour cet r , l'événement $D_n^+ \cap [S_{2n} = 2r]$ est exactement $A_{n,r}$, puisque la condition finale $2r > 0$ complète celles sur $S_1, \dots, S_{2n-1}$. Ces événements sont deux à deux disjoints ; leur réunion est D_n^+ .

$$D_n^+ = \bigcup_{r=1}^n A_{n,r}, \quad \mathbb{P}(D_n^+) = \sum_{r=1}^n a_{n,r}.$$

18. Récurrence pour les probabilités $a_{n,r}$.

On fixe $n \geq 2$. Par convention, $a_{m,q} = 0$ pour $q \leq 0$ ou $q > m$.

- a) Fixons $q, r \geq 1$. Sur $A_{n-1,q}$, toutes les sommes jusqu'au temps $2n-2$ sont strictement positives et $S_{2n-2} = 2q$. Les deux incréments suivants sont indépendants de cet événement, et les quatre couples de valeurs ont chacun une probabilité $1/4$. Les possibilités pour les deux nouvelles sommes sont

(X_{2n-1}, X_{2n})	S_{2n-1}	S_{2n}	valeur de r
$(1, 1)$	$2q + 1$	$2q + 2$	$q + 1$
$(1, -1)$	$2q + 1$	$2q$	q
$(-1, 1)$	$2q - 1$	$2q$	q
$(-1, -1)$	$2q - 1$	$2q - 2$	$q - 1$.

Pour $q \geq 1$, les deux valeurs intermédiaires $2q \pm 1$ sont strictement positives. Dans la dernière ligne, la valeur finale est positive seulement si $q \geq 2$; cela est automatique lorsqu'elle doit valoir $2r$ avec $r \geq 1$ et $q = r + 1$. Il y a donc un seul couple admissible si $q = r - 1$ ou $q = r + 1$, deux si $q = r$, et aucun sinon. L'indépendance avec les premières variables permet de multiplier ces probabilités par $a_{n-1,q}$:

$$\mathbb{P}(A_{n-1,q} \cap A_{n,r}) = \begin{cases} \frac{1}{4} a_{n-1,q}, & q = r - 1 \text{ ou } q = r + 1, \\ \frac{1}{2} a_{n-1,q}, & q = r, \\ 0, & \text{sinon.} \end{cases}$$

Si $q > n - 1$, $A_{n-1,q}$ est vide, et la même formule vaut avec $a_{n-1,q} = 0$.

- b)** Sur $A_{n,r}$, le temps $2n - 2$ est antérieur au temps final et sa somme est donc strictement positive. Cette somme, paire et au plus $2n - 2$, vaut $2q$ pour un unique $q \in \{1, \dots, n - 1\}$. Les conditions antérieures montrent alors que $A_{n-1,q}$ est réalisé. Ainsi

$$A_{n,r} = \bigcup_{q=1}^{n-1} (A_{n-1,q} \cap A_{n,r}).$$

On somme les probabilités obtenues en a). Seuls les indices $q = r - 1, r, r + 1$ contribuent ; la convention de nullité hors de $\{1, \dots, n - 1\}$ traite les extrémités.

$$\boxed{\forall r \geq 1, \quad a_{n,r} = \frac{1}{4}a_{n-1,r-1} + \frac{1}{2}a_{n-1,r} + \frac{1}{4}a_{n-1,r+1}.}$$

Le raisonnement est aussi valable lorsque $r > n$, cas où tous les termes concernés sont nuls.

Troisième partie – Comportement asymptotique

Formule explicite des probabilités de non-retour, simulation du dernier passage par zéro et convergence en loi de L_n/n vers la loi arcsinus.

Corrigé – Troisième partie – Comportement asymptotique.**19. Formule explicite de $a_{n,r}$.**

Pour $n \geq 1$ et $r \geq 0$, posons

$$b_{n,r} = \frac{1}{4^n} \left(\binom{2n-1}{n+r-1} - \binom{2n-1}{n+r} \right),$$

avec la convention sur les coefficients binomiaux de l'énoncé. Pour $r = 0$, la symétrie des coefficients binomiaux donne $\binom{2n-1}{n-1} = \binom{2n-1}{n}$, donc $b_{n,0} = 0 = a_{n,0}$. Pour $r > n$, on a également $b_{n,r} = 0 = a_{n,r}$.

Initialisation. Pour $n = 1$, l'événement $A_{1,1}$ correspond aux deux premiers pas égaux à $+1$, donc $a_{1,1} = 1/4$. La formule donne

$$b_{1,1} = \frac{1}{4} \left(\binom{1}{1} - \binom{1}{2} \right) = \frac{1}{4}.$$

Pour $r > 1$, $a_{1,r} = b_{1,r} = 0$. La formule est donc vraie à l'ordre 1 pour tout $r \geq 1$.

Hérédité. Soit $n \geq 2$, et supposons $a_{n-1,r} = b_{n-1,r}$ pour tout $r \geq 1$. L'égalité est aussi vraie pour $r = 0$, d'après le premier paragraphe. Fixons $r \geq 1$ et posons $m = 2n - 3$ et $j = n + r - 1$. La récurrence de la question 18 donne

$$\begin{aligned} a_{n,r} &= \frac{1}{4} b_{n-1,r-1} + \frac{1}{2} b_{n-1,r} + \frac{1}{4} b_{n-1,r+1} \\ &= \frac{1}{4^n} \left[\binom{m}{j-2} + \binom{m}{j-1} - \binom{m}{j} - \binom{m}{j+1} \right]. \end{aligned}$$

Deux applications de la formule de Pascal montrent que

$$\begin{aligned} \binom{m+2}{j} - \binom{m+2}{j+1} &= \left[\binom{m}{j-2} + 2\binom{m}{j-1} + \binom{m}{j} \right] \\ &\quad - \left[\binom{m}{j-1} + 2\binom{m}{j} + \binom{m}{j+1} \right] \\ &= \binom{m}{j-2} + \binom{m}{j-1} - \binom{m}{j} - \binom{m}{j+1}. \end{aligned}$$

Ces identités restent valables aux bords grâce à la convention des coefficients nuls. Comme $m+2 = 2n-1$, on trouve $a_{n,r} = b_{n,r}$. Le raisonnement s'applique en particulier à $r = 1$, où le terme $a_{n-1,0}$ est nul. Par récurrence,

$$\forall n, r \in \mathbb{N}^*, \quad a_{n,r} = \frac{1}{4^n} \left(\binom{2n-1}{n+r-1} - \binom{2n-1}{n+r} \right).$$

20. Probabilité de ne pas revenir à zéro.

Pour $n \geq 1$, on a $p_n = \mathbb{P}(D_n)$. Changer le signe de chacun des $2n$ premiers pas échange les événements D_n^+ et D_n^- . Comme chaque suite de $2n$ pas a la même probabilité 2^{-2n} ,

leurs probabilités sont égales. Les questions 16 et 17 donnent donc

$$p_n = 2\mathbb{P}(D_n^+) = 2 \sum_{r=1}^n a_{n,r}.$$

En utilisant la question 19, la somme est télescopique :

$$\begin{aligned} p_n &= \frac{2}{4^n} \sum_{r=1}^n \left(\binom{2n-1}{n+r-1} - \binom{2n-1}{n+r} \right) \\ &= \frac{2}{4^n} \binom{2n-1}{n} = \frac{1}{4^n} \binom{2n}{n}. \end{aligned}$$

La dernière égalité découle de $\binom{2n}{n} = 2 \binom{2n-1}{n}$. Pour $n = 0$, la formule donne 1, conformément à la convention $p_0 = 1$. Ainsi

$$\boxed{\forall n \in \mathbb{N}, \quad p_n = \frac{1}{4^n} \binom{2n}{n}.}$$

En combinant ce résultat avec la question 15, on dispose aussi de

$$\boxed{\forall k \in \llbracket 0, n \rrbracket, \quad \mathbb{P}(L_n = k) = p_k p_{n-k}.}$$

21. Simulation et conjecture.

- a) Chaque pas vaut $+1$ ou -1 avec probabilité $1/2$. On actualise la somme des pas, puis l'indice du dernier retour lorsque cette somme s'annule.

```
import numpy.random as rd
```

```
def DernierPassage(n):
    L = 0
    S = 0
    for i in range(1, 2*n+1):
        if rd.random() < 0.5:
            S = S+1
        else:
            S = S-1
        if S == 0:
            L = i/2
    return L
```

La somme de i nombres égaux à $+1$ ou -1 a la même parité que i : un retour à zéro se produit donc uniquement à un instant pair. Le nombre $i/2$, qui est entier lors d'un retour à zéro, est alors exactement l'indice k tel que $i = 2k$. La valeur initiale $L = 0$ prend en compte le passage à zéro à l'instant 0.

$\boxed{\text{DernierPassage}(n) \text{ simule } L_n.}$

- b) L'histogramme est symétrique autour de $1/2$ et présente des valeurs plus élevées près de 0 et de 1, comme la densité de la loi arcsinus étudiée dans la première partie. Il suggère la conjecture

$$\boxed{\frac{L_n}{n} \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} X, \quad X \text{ suit la loi arcsinus.}}$$

Il reste à démontrer cette convergence à partir des fonctions de répartition.

22. Convergence de (C_n) .

- a) Les nombres C_n sont strictement positifs. Un calcul exact à partir des factorielles donne

$$\frac{C_{n+1}}{C_n} = \sqrt{\frac{n+1}{n}} \frac{(2n+2)(2n+1)}{4(n+1)^2} = \sqrt{\frac{n+1}{n}} \frac{2n+1}{2n+2}.$$

En élevant au carré,

$$\left(\frac{C_{n+1}}{C_n}\right)^2 = \frac{(2n+1)^2}{4n(n+1)} = 1 + \frac{1}{4n(n+1)}.$$

Par conséquent,

$$u_{n+1} - u_n = \frac{1}{2} \ln \left(1 + \frac{1}{4n(n+1)} \right).$$

Pour $t \geq 0$, on a $0 \leq \ln(1+t) \leq t$: en effet, la fonction $t - \ln(1+t)$ vaut 0 en 0 et sa dérivée est $t/(1+t) \geq 0$. Ainsi

$$0 < u_{n+1} - u_n \leq \frac{1}{8n(n+1)} = \frac{1}{8} \left(\frac{1}{n} - \frac{1}{n+1} \right).$$

La série majorante est télescopique et converge. Par comparaison,

$$\sum_{n=1}^{+\infty} (u_{n+1} - u_n) \text{ converge.}$$

- b) Pour tout $N \geq 1$,

$$u_N = u_1 + \sum_{n=1}^{N-1} (u_{n+1} - u_n).$$

La convergence de la série entraîne donc celle de u_N vers un réel ℓ . Par continuité de l'exponentielle, $C_N = e^{u_N}$ tend vers $e^\ell > 0$:

$$(C_n) \text{ converge vers une limite strictement positive.}$$

Dans la suite, on utilise la valeur admise de cette limite :

$$\lim_{n \rightarrow +\infty} C_n = c, \quad c = \frac{1}{\sqrt{\pi}}.$$

23. Probabilités cumulées loin de l'origine.

Fixons $x \in]0, 1[$ et une suite (a_n) vérifiant les hypothèses de la question.

- a) Comme L_n est entier, $[L_n/n \leq x] = [L_n \leq [nx]]$. La loi obtenue à la question 20 donne

$$\mathbb{P} \left(\frac{L_n}{n} \leq x \right) = \frac{1}{4^n} \sum_{k=0}^{[nx]} \binom{2k}{k} \binom{2(n-k)}{n-k}.$$

- b) Pour n assez grand, $1 \leq a_n < [nx] < n$. Pour tout entier $k \in \llbracket a_n, [nx] \rrbracket$, les définitions donnent

$$\frac{1}{4^n} \binom{2k}{k} \binom{2(n-k)}{n-k} = \frac{C_k C_{n-k}}{\sqrt{k(n-k)}}.$$

Par définition de m_n et M_n , on a $m_n \leq C_k C_{n-k} \leq M_n$. En multipliant par le nombre positif $1/\sqrt{k(n-k)}$ puis en sommant,

$$\boxed{m_n \sum_{k=a_n}^{\lfloor nx \rfloor} \frac{1}{\sqrt{k(n-k)}} \leq \frac{1}{4^n} \sum_{k=a_n}^{\lfloor nx \rfloor} \binom{2k}{k} \binom{2(n-k)}{n-k} \leq M_n \sum_{k=a_n}^{\lfloor nx \rfloor} \frac{1}{\sqrt{k(n-k)}}.}$$

- c) Soit $\varepsilon > 0$. Puisque $C_k \rightarrow c$, il existe un entier K tel que, pour tout $k \geq K$, $c - \varepsilon < C_k < c + \varepsilon$. Comme $\alpha_n \rightarrow +\infty$, pour n assez grand on a $\alpha_n \geq K$. Tous les indices $k \in \llbracket \alpha_n, \beta_n \rrbracket$ satisfont alors les mêmes bornes. Les minimum et maximum, pris sur cet ensemble fini non vide, vérifient

$$c - \varepsilon \leq \min_{\alpha_n \leq k \leq \beta_n} C_k \leq \max_{\alpha_n \leq k \leq \beta_n} C_k \leq c + \varepsilon.$$

La définition de la convergence donne donc

$$\boxed{\min_{\alpha_n \leq k \leq \beta_n} C_k \longrightarrow c, \quad \max_{\alpha_n \leq k \leq \beta_n} C_k \longrightarrow c.}$$

- d) Appliquons le résultat précédent aux deux ensembles d'indices

$$I_n = \llbracket a_n, \lfloor nx \rfloor \rrbracket, \quad J_n = \llbracket n - \lfloor nx \rfloor, n - a_n \rrbracket.$$

Leurs premiers indices tendent vers $+\infty$, car $a_n \rightarrow +\infty$ et $n - \lfloor nx \rfloor \geq n(1-x) \rightarrow +\infty$. Posons

$$r_n = \min_{k \in I_n} C_k, \quad s_n = \max_{k \in I_n} C_k, \quad r'_n = \min_{j \in J_n} C_j, \quad s'_n = \max_{j \in J_n} C_j.$$

Ces quatre suites tendent vers c . Tous les termes étant positifs,

$$r_n r'_n \leq m_n \leq M_n \leq s_n s'_n.$$

Par encadrement, m_n et M_n tendent vers $c^2 = 1/\pi$. La limite de la somme $\sum_{k=a_n}^{\lfloor nx \rfloor} 1/\sqrt{k(n-k)}$ vaut $G(x)$, d'après le résultat admis à la fin de la première partie. L'encadrement de la question b) entraîne donc

$$\boxed{\frac{1}{4^n} \sum_{k=a_n}^{\lfloor nx \rfloor} \binom{2k}{k} \binom{2(n-k)}{n-k} \longrightarrow \frac{G(x)}{\pi}.}$$

24. Probabilités cumulées près de l'origine.

- a) On peut choisir

$$\boxed{a_n = \lfloor n^{1/4} \rfloor.}$$

En effet, $a_n \rightarrow +\infty$, et $0 \leq a_n/\sqrt{n} \leq n^{-1/4} \rightarrow 0$. Pour $x > 0$ fixé, on a aussi $a_n/n \rightarrow 0$ et $\lfloor nx \rfloor/n \rightarrow x$; donc, pour n assez grand, $1 \leq a_n < \lfloor nx \rfloor$.

- b) La suite convergente $(C_j)_{j \geq 1}$ est bornée : il existe $B > 0$ tel que $C_j \leq B$ pour tout $j \geq 1$. Pour tout $k \geq 0$, $p_k \leq 1$, puisque p_k est une probabilité, y compris $p_0 = 1$. Pour n assez grand et $0 \leq k \leq a_n - 1$, on a $n - k \geq n - a_n + 1 \geq 1$, et donc

$$0 \leq p_k p_{n-k} \leq p_{n-k} = \frac{C_{n-k}}{\sqrt{n-k}} \leq \frac{B}{\sqrt{n-a_n+1}}.$$

En sommant les a_n termes,

$$0 \leq \frac{1}{4^n} \sum_{k=0}^{a_n-1} \binom{2k}{k} \binom{2(n-k)}{n-k} \leq \frac{Ba_n}{\sqrt{n-a_n+1}}.$$

Or

$$\frac{Ba_n}{\sqrt{n-a_n+1}} = B \frac{a_n/\sqrt{n}}{\sqrt{1-a_n/n+1/n}} \rightarrow 0.$$

Par encadrement, pour toute suite (a_n) vérifiant les hypothèses de la question,

$$\boxed{\frac{1}{4^n} \sum_{k=0}^{a_n-1} \binom{2k}{k} \binom{2(n-k)}{n-k} \rightarrow 0.}$$

25. Convergence en loi.

Notons F_n la fonction de répartition de L_n/n . Comme $0 \leq L_n/n \leq 1$, on a $F_n(x) = 0$ pour $x < 0$ et $F_n(x) = 1$ pour $x \geq 1$.

Pour $x \in]0, 1[$, prenons par exemple $a_n = \lfloor n^{1/4} \rfloor$. Cette suite satisfait les hypothèses des questions 23 et 24. En coupant la somme de la question 23 a) à l'indice a_n , on obtient

$$\begin{aligned} F_n(x) &= \frac{1}{4^n} \sum_{k=0}^{a_n-1} \binom{2k}{k} \binom{2(n-k)}{n-k} \\ &\quad + \frac{1}{4^n} \sum_{k=a_n}^{\lfloor nx \rfloor} \binom{2k}{k} \binom{2(n-k)}{n-k} \rightarrow \frac{G(x)}{\pi}. \end{aligned}$$

Au point $x = 0$,

$$F_n(0) = \mathbb{P}(L_n = 0) = p_n = \frac{C_n}{\sqrt{n}} \rightarrow 0.$$

Au point $x = 1$, $F_n(1) = 1$ pour tout n . Ainsi, pour tout $x \in \mathbb{R}$, $F_n(x)$ tend vers

$$F(x) = \begin{cases} 0, & x \leq 0, \\ \frac{2}{\pi} \arcsin(\sqrt{x}), & 0 < x < 1, \\ 1, & x \geq 1. \end{cases}$$

Cette fonction, continue sur $\mathbb{R}$ puisque $G(0) = 0$ et $G(1) = \pi$, est la fonction de répartition de la loi arcsinus étudiée dans la première partie. Par la définition de la convergence en loi,

$$\boxed{\frac{L_n}{n} \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} X, \quad X \text{ suit la loi arcsinus.}}$$