

Quentin Berger
Francesco Caravenna
Paolo Dai Pra

Introduction aux Probabilités

Quelques exemples (et exercices) supplémentaires

31 janvier 2026

Liste des exemples (ou exercices)

1	Un exercice sur l'approximation de Poisson	1
2	Aiguille de Buffon et théorème de Barbier	3
3	Nombre de cycles d'une permutation aléatoire	7
4	Une caractérisation de la gaussienne : le théorème de Bernstein	9
5	Vecteur aléatoire uniforme dans une boule de grande dimension	12
6	Minimum de variables exponentielles : le lemme des réveils*	15
7	Maximum de variables aléatoires géométriques et log-périodicité	17
8	Probabilité que deux variables géométriques soient premières entre elles	22
9	Problème de survie d'une marche aléatoire	26
9.1	Pour la marche aléatoire simple	26
9.2	Marches aléatoires symétriques à densité : le théorème de Sparre Andersen	29
10	Comportement asymptotique d'un processus de renouvellement	33
10.1	Loi des grands nombres pour le processus de renouvellement	33
10.2	Théorème central limite pour le processus de renouvellement	35
11	Urne de Pólya et loi Beta	37
12	Records successifs d'une suite de variables aléatoires i.i.d.*	42

Préambule

Nous avons rassemblé dans ce document des exemples (ou exercices) qui n'apparaissent dans notre livre. Ils peuvent être vus soit comme des illustrations de notions importantes en probabilités, soit comme des résultats rarement abordés dans un cours standard mais qui mériteraient d'être connus, soit simplement comme des exemples que nous trouvons originaux et/ou amusants. Soulignons qu'il s'agit d'un document de travail : certaines parties sont peut-être incomplètes et il contient sûrement un certain nombre de coquilles. Nous avons néanmoins décidé de rédiger ces quelques notes et de les proposer en ligne afin qu'elles puissent être utiles au plus grand nombre. D'ailleurs, n'hésitez pas à nous contacter si vous avez des commentaires ou des suggestions d'autres problèmes amusants ! En vous souhaitant bonne lecture.

Nous ferons régulièrement référence aux définitions, théorèmes, propositions, équations, etc. de la version publiée de notre livre, en gardant la même numérotation que dans celui-ci ; les équations et figures du document présent seront précédées de la lettre « E » (par exemple, l'équation (E-1.1)).

Ex. 1

Un exercice sur l'approximation de Poisson

Problème 1.1. On attribue à n personnes, de manière aléatoire (et indépendante), un réel choisi de manière uniforme dans $[0, 1]$. Le but du jeu est de trouver la personne qui possède le nombre le plus petit. Cependant, le seul outil que l'on possède est celui de faire un sondage : on pose une question aux n personnes (la même pour tout le monde) à laquelle les seules réponses possibles sont oui ou non. Comment poser la question, et quel est dans ce cas la probabilité de trouver la personne qui possède le nombre le plus petit ?

Solution. On va fixer $x_n \in]0, 1[$ et poser la question suivante :

« Votre nombre est-il plus petit que x_n ? »

Si une seule personne répond « Oui », il s'agit de la personne qui possède le plus petit nombre. L'approximation de Poisson (Section 4.1) nous dit que si $x_n = \frac{\lambda}{n}$, la probabilité qu'il y ait exactement une personne qui réponde *oui* vaut approximativement $P(X = 1)$ où $X \sim \text{Poi}(\lambda)$, c'est-à-dire $\lambda e^{-\lambda}$. On remarque alors que la fonction $\lambda \mapsto \lambda e^{-\lambda}$ est maximale en $\lambda = 1$: on peut donc prendre $x_n = \frac{1}{n}$ et la probabilité qu'il y ait exactement une personne qui réponde *oui* (et donc que l'on trouve la personne qui possède le plus petit nombre) vaut approximativement $e^{-1} \approx 0,36788$.

Mais on peut en fait améliorer un peu cette stratégie. En effet, si k personnes répondent « Oui », on pourra choisir une de ces personnes au hasard et on aura une chance sur k d'avoir trouvé la personne au plus petit nombre. Il reste à choisir proprement le nombre $x = x_n$.

On note $A_j = A_j(x)$ l'événement « la j -ème personne répond *oui* » et $N = N(x) = \sum_{i=1}^n \mathbb{1}_{A_i}$ le nombre de personnes qui répondent *oui* à la question. Ainsi, $N \sim \text{Bin}(n, x)$, car les événements A_j sont indépendants et de probabilité $P(U \leq x) = x$ pour une variable $U \sim U(0, 1)$. Alors, la probabilité de trouver la personne avec le plus petit nombre est

$$p_n(x) := \sum_{k=1}^{\infty} \frac{1}{k} P(N = k).$$

Il reste à optimiser sur le choix de x .

On va supposer que n est très grand et utiliser l'approximation de Poisson : on va choisir $x = \frac{\lambda}{n}$ où $\lambda > 0$ est fixé (à optimiser), de sorte que N soit proche d'une loi $\text{Poi}(\lambda)$. Plus précisément, si $X \sim \text{Poi}(\lambda)$, alors grâce au Théorème 4.8, on obtient

$$\begin{aligned} \left| \sum_{k=1}^{\infty} \frac{1}{k} P(N = k) - \sum_{k=1}^{\infty} \frac{1}{k} P(X = k) \right| &\leq \sum_{k=1}^{\infty} \frac{1}{k} |P(N = k) - P(X = k)| \\ &\leq \sum_{k=1}^n |P(N = k) - P(X = k)| \leq \frac{\lambda^2}{n} \xrightarrow{n \rightarrow \infty} 0. \end{aligned}$$

Ainsi, la probabilité recherchée est proche de

$$\hat{p}(\lambda) := \sum_{k=1}^{\infty} \frac{1}{k} P(X = k) = \sum_{k=1}^n \frac{1}{k} \frac{\lambda^k}{k!} e^{-\lambda}.$$

Il s'avère qu'il n'existe pas de forme fermée pour cette formule : il est possible de montrer qu'il existe un unique maximum en un point $\lambda_0 \in]0, +\infty[$, mais la valeur de λ_0 n'est pas explicite. On peut par contre étudier la dépendance en λ de manière numérique : on trouve que le maximum est $\max \hat{p}(\lambda) \approx 0,51735$, atteint pour $\lambda_0 \approx 1,50286$. On présente dans la Figure E-1.1 le graphe de la fonction $\lambda \mapsto \hat{p}(\lambda)$: d'abord sur $[0, 20]$, puis sur $[1, 2]$ où est atteint le maximum.

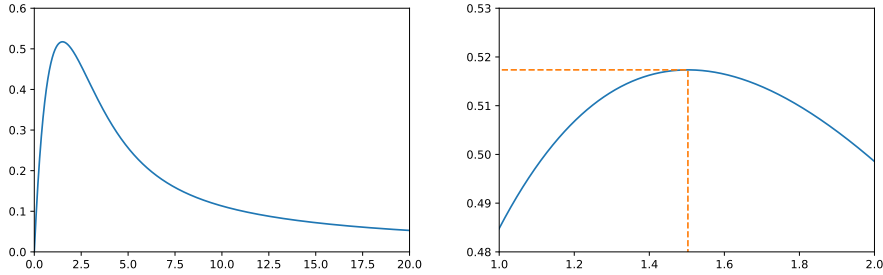


FIG. E-1.1 Graphe de la fonction $\lambda \mapsto \hat{p}(\lambda)$, à deux échelles différentes. Dans le deuxième graphe, on a noté le maximum qui vaut $\approx 0,51735$, atteint pour $\lambda_0 \approx 1,50286$.

En conclusion, il faut poser la question « Votre nombre est-il plus petit que $\frac{\lambda_0}{n}$? », et choisir l'une des personnes au hasard parmi celles qui répondent « oui » (il y en a un nombre aléatoire N de loi proche d'une loi $\text{Poi}(\lambda_0)$). Dans ce cas, la probabilité de trouver la personne qui possède le plus petit nombre est environ 0,51735, soit un peu plus d'une chance sur deux ! \square

Ex. 2

Aiguille de Buffon et théorème de Barbier

Considérons le problème de l'Aiguille de Buffon de l'Exercice 6.47 : on lance une aiguille sur un parquet composé de planches parallèles et on se demande quelle est la probabilité que l'aiguille tombe à cheval sur (au moins) une rainure du parquet. On va maintenant résoudre ce problème d'une manière différente, qui ne requiert quasiment aucun calcul mais repose sur la propriété de linéarité de l'espérance. Il s'agit d'une méthode donnée par Barbier en 1860 pour démontrer un théorème qui porte son nom (voir le Théorème 2.1 ci-dessous) ; la démonstration qu'on donne ici est grandement inspirée de l'exposé « ma preuve préférée » d'Alexandre Gaudillière[‡].

Supposons que les rainures du parquet sont à une distance d les unes des autres et que l'on lance une aiguille de longueur ℓ . Cette fois, plutôt que s'intéresser à l'événement « l'aiguille tombe à cheval sur une rainure », on va étudier la variable aléatoire N_ℓ = « nombre de rainures rencontrées par l'aiguille ». Évidemment, on aura $N_\ell \leq 1$ dans le cas où $\ell < d$, mais on peut avoir $N_\ell \geq 2$ si $\ell \geq d$; de manière générale, N_ℓ est à valeurs dans $\{0, \dots, \lfloor \ell/d \rfloor\}$.

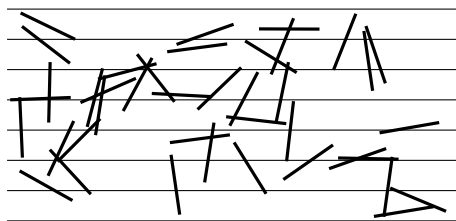
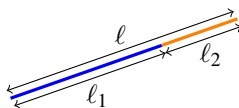


FIG. E-2.1 Illustration du parquet et de la position de 40 aiguilles de longueur $2d$ lancées au hasard.

1ère étape : linéarité de $E(N_\ell)$. La première observation que l'on peut faire est la suivante : si l'on découpe une aiguille de longueur ℓ en deux parties de longueurs respectives ℓ_1, ℓ_2 (avec $\ell_1 + \ell_2 = \ell$), alors on peut écrire $N_\ell = \tilde{N}_{\ell_1} + \tilde{N}_{\ell_2}$, où \tilde{N}_{ℓ_1} est le nombre de rainures rencontrées par la première partie de l'aiguille (de longueur ℓ_1) et \tilde{N}_{ℓ_2} est le nombre de rainures rencontrées par la deuxième partie de l'aiguille (de longueur ℓ_2).



Évidemment, \tilde{N}_{ℓ_1} et \tilde{N}_{ℓ_2} ne sont pas indépendantes (pourquoi ?), mais on peut utiliser la linéarité de l'espérance pour obtenir

$$E(N_\ell) = E(\tilde{N}_{\ell_1}) + E(\tilde{N}_{\ell_2}) = E(N_{\ell_1}) + E(N_{\ell_2}),$$

où on a utilisé le fait que \tilde{N}_{ℓ_1} a la même loi que N_{ℓ_1} (de même pour \tilde{N}_{ℓ_2}).

Ainsi, si l'on pose $g(\ell) := E(N_\ell)$, on a montré que la fonction $g : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ vérifie : pour tous $x, y \geq 0$, $g(x+y) = g(x) + g(y)$. Il s'agit ensuite d'un exercice classique de montrer que

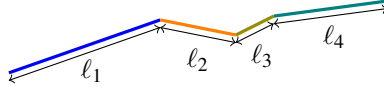
[‡]. Voici le lien : <https://video.math.cnrs.fr/ma-preuve-preferee/>.

la fonction g est linéaire[†] sur \mathbb{R}_+ , et on en conclut donc qu'il existe $\alpha \in \mathbb{R}_+$ tel que

$$\text{pour tout } \ell > 0 \quad E(N_\ell) = \alpha \ell. \quad (\text{E-2.1})$$

Il reste maintenant à déterminer la constante α .

2ème étape : aiguilles polygonales. La deuxième observation est que le point précédent reste vrai si l'on considère une aiguille *polygonale*, constituée de segments de longueurs ℓ_1, \dots, ℓ_k .



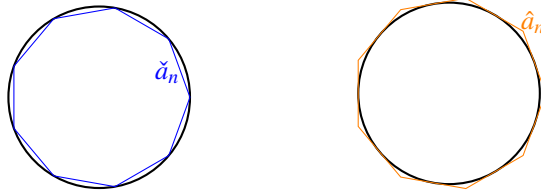
En effet, notons N_a le nombre de rainures rencontrées par une aiguille polygonale a constituée des segments de longueur ℓ_1, \dots, ℓ_k , lancée au hasard sur le parquet. On peut alors écrire $N_a = \tilde{N}_{\ell_1} + \dots + \tilde{N}_{\ell_k}$, où \tilde{N}_{ℓ_i} est le nombre de rainures rencontrées par le i -ème segment (de longueur ℓ_i). Encore une fois, les $(\tilde{N}_{\ell_i})_{1 \leq i \leq k}$ ne sont pas indépendants, mais par linéarité de l'espérance, on obtient encore

$$E(N_a) = E(\tilde{N}_{\ell_1}) + \dots + E(\tilde{N}_{\ell_k}) = \alpha \ell_1 + \dots + \alpha \ell_k = \alpha \ell(a), \quad (\text{E-2.2})$$

où $\ell(a) := \ell_1 + \dots + \ell_k$ est la longueur totale de l'aiguille ; noter que l'on a utilisé (E-2.1) pour obtenir $E(\tilde{N}_{\ell_i}) = \alpha \ell_i$.

3ème étape : approximation du cercle et conclusion. Maintenant, on peut encore généraliser la formule (E-2.2) au cas d'aiguilles *courbes*, au moins dans le cas d'aiguilles que l'on peut approcher par une suite de lignes polygonales.

Prenons l'exemple d'un cercle C_r de rayon r , et notons N_{C_r} le nombre de rainures rencontrées par une aiguille de forme C_r que l'on lance au hasard sur le parquet. Pour tout $n \geq 1$, on peut alors encadrer le cercle C_r par deux polygones à n côtés, le premier \check{a}_n inscrit dans le cercle, le deuxième \hat{a}_n circonscrit au cercle.



Dans ce cas, les nombres de rainures rencontrées par les aiguilles correspondantes vérifient clairement $N_{\check{a}_n} \leq N_{C_r} \leq N_{\hat{a}_n}$, et donc par monotonie de l'espérance,

$$\alpha \ell(\check{a}_n) = E(N_{\check{a}_n}) \leq E(N_{C_r}) \leq E(N_{\hat{a}_n}) = \alpha \ell(\hat{a}_n).$$

Maintenant, en prenant la limite quand $n \rightarrow \infty$ et comme les longueurs des polygones $\ell(\check{a}_n)$, $\ell(\hat{a}_n)$ convergent vers la périmètre du cercle $\ell(C_r) = 2\pi r$, on en conclut que, pour tout $r > 0$

$$E(N_{C_r}) = \alpha \times \ell(C_r) = 2\pi r \alpha.$$

[†]. Donnons ici les étapes : (i) par récurrence, on a $g(n) = ng(1)$ pour tout $n \in \mathbb{N}$; (ii) par récurrence (sur m , à n fixé), on a $g(\frac{n}{m}) = \frac{n}{m}g(1)$ pour tout $\frac{n}{m} \in \mathbb{Q}$; (iii) comme $g : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ est croissante, par approximation on obtient que $g(x) = xg(1)$ pour tout $x \in \mathbb{R}_+$.

Cette identité permet alors de déterminer la valeur de α . En effet, si on considère un cercle de diamètre d (autrement dit de rayon $r = d/2$), où on rappelle que d est la distance entre deux rainures de parquet, alors on se rend compte que ce cercle rencontrera *exactement* 2 rainures, quelle que soit sa position sur le parquet. On a donc $N_{C_{d/2}} = 2$, et

$$2 = E(N_{C_{d/2}}) = \pi d \alpha \implies \alpha = \frac{2}{\pi d}.$$

En conclusion, on a montré que pour toute aiguille A (droite, polygonale ou courbe[†]) de longueur $\ell(A)$, si on note N_A le nombre de rainures rencontrées par l'aiguille A lancée aléatoirement sur le parquet, alors

$$E(N_A) = \frac{2}{\pi d} \ell(A). \quad (\text{E-2.3})$$

Cette relation est aussi connue sous le nom de formule de Crofton dans le domaine de la géométrie intégrale et possède des généralisations en dimension plus grande[‡].

En particulier, si l'on considère une aiguille droite de longueur $\ell < d$, alors N_ℓ vaut soit 0 soit 1, et $E(N_\ell) = P(N_\ell = 1) = \frac{2\ell}{\pi d}$, où l'événement « $N_\ell = 1$ » est exactement l'événement « l'aiguille tombe à cheval sur (au moins) une rainure du parquet ».

Conséquence : le théorème de Barbier. Le même raisonnement s'applique aux *courbes de largeur constante*, qui sont des courbes avec la propriété suivante : pour n'importe quelle paire de droites parallèles tangentes à la courbe en des points opposés (appelées aussi lignes d'appui), la distance entre ces deux droites est constante ; cette distance est alors appelée *diamètre* de la courbe, voir la Figure E-2.2. Le cercle est clairement un exemple de courbe de largeur constante, mais il existe tout un tas de courbes de ce type, par exemple les polygones dit de *Reuleaux*.

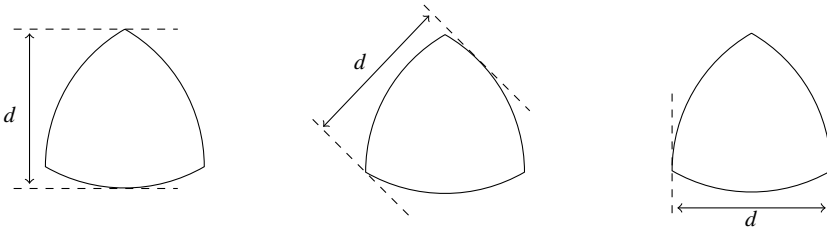


FIG. E-2.2 Une courbe de largeur constante : le triangle de Reuleaux. Des lignes d'appui y sont représentées.

On peut montrer un certain nombre de propriétés de ces courbes : par exemple, il s'agit de courbes strictement convexes. Si on se donne C une courbe de largeur constante égale à d , alors on se rend aussi compte que, de la même manière que pour le cercle, la courbe C rencontrera *exactement* 2 rainures, quelle que soit sa position sur le parquet. Ainsi, grâce à la formule (E-2.3), on obtient $2 = E(N_C) = \frac{2}{\pi d} \ell(C)$, où $\ell(C)$ est la longueur de la courbe C , c'est-à-dire son périmètre. On en déduit alors le théorème suivant, obtenu par Joseph Émile Barbier en 1860.

[†]. Il faut que la courbe soit suffisamment lisse, par exemple une courbe différentiable et régulière, pour être approchable par une ligne polygonale ; une aiguille courbe est parfois appelée « nouille de Buffon ».

[‡]. On peut renvoyer à la vidéo suivante de 3blue1brown : <https://www.youtube.com/watch?v=ltLUadnCyi0>

Théorème 2.1 (de Barbier). *Une courbe de largeur constante de diamètre d a pour périmètre πd .*

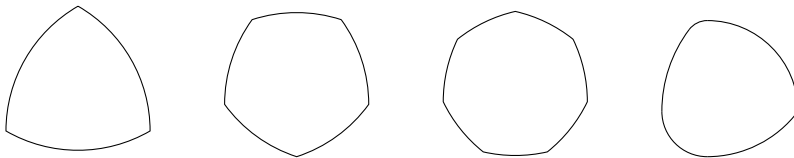


FIG. E-2.3 Quelques autres courbes de largeur constante : un triangle, un pentagone et un heptagone de Reuleaux (dont la pièce de 50 cents britannique est un exemple), et une autre courbe de largeur constante non régulière.

Ex. 3

Nombre de cycles d'une permutation aléatoire

Pour $n \in \mathbb{N}^*$, on note \mathfrak{S}_n le groupe des permutations de $\{1, \dots, n\}$. On choisit une permutation aléatoirement de manière uniforme sur \mathfrak{S}_n , et on note X_n le nombre de cycles disjoints que cette permutation contient. On montre le résultat suivant.

Proposition 3.1. *La variable aléatoire X_n a la même loi que $Y_1 + \dots + Y_n$, où les variables aléatoires (Y_1, \dots, Y_n) sont indépendantes et de lois $Y_i \sim \text{Bern}(\frac{1}{i})$.*

En particulier, on en déduit que $\frac{X_n}{\ln n}$ converge en probabilité vers 1 quand $n \rightarrow \infty$, voir la Section 3.3.6.

Démonstration. Une preuve possible passe par le calcul de la fonction génératrice de X_n . On va montrer que X_n et $Y_1 + \dots + Y_n$ ont la même fonction génératrice.

Étape 1. Combinatoire. Pour $k \in \mathbb{N}$ on pose $\mathcal{C}_{n,k}$ l'ensemble des permutations $\sigma \in \mathfrak{S}_n$ qui ont exactement k cycles à support disjoint et on note $C_{n,k}$ le cardinal de cet ensemble.

Lemme 3.2. *On a $C_{n+1,k} = nC_{n,k} + C_{n,k-1}$ pour tout $n, k \geq 0$.*

Démonstration. On a $C_{n,k} = 0$ si $k > n$ ou si $k < 1$. De plus, $C_{n,n} = 1$ car la seule permutation avec n cycles distincts (n points fixes) est l'identité. Le nombre $C_{n,1}$ de permutations avec un seul cycle vaut $(n-1)!$, car il faut choisir l'image de 1 (il y a $n-1$ possibilités), puis l'image du nombre choisi (il y a $n-2$ possibilités), etc.

Pour obtenir la formule, notons que l'on peut séparer l'ensemble des permutations de longueur $n+1$ avec exactement k cycles en deux sous-ensembles : soit $n+1$ apparaît dans un cycle de longueur au moins 2, soit $n+1$ est un point fixe.

Le premier ensemble est de cardinal $nC_{n,k}$. En effet, étant donné une permutation $\sigma \in \mathfrak{S}_n$ avec exactement k cycles, on peut placer l'indice $n+1$ après n'importe quel indice $i \in \{1, \dots, n\}$ dans la décomposition en cycles de σ : cela donne une permutation $\sigma' \in \mathfrak{S}_{n+1}$ avec exactement k cycles (où $n+1$ apparaît dans un cycle de longueur au moins 2).

Le deuxième ensemble est de cardinal $C_{n,k-1}$. En effet, si on a une permutation $\sigma \in \mathfrak{S}_n$ avec exactement $k-1$ cycles, on peut définir $\sigma' \in \mathfrak{S}_{n+1}$ en définissant $\sigma'(i) = \sigma(i)$ pour $i \in \{1, \dots, n\}$ et $\sigma'(n+1) = n+1$: cela donne une permutation $\sigma' \in \mathfrak{S}_{n+1}$ avec exactement k cycles (où $n+1$ apparaît dans un cycle de longueur 1). \square

Étape 2. Calcul de la fonction génératrice de X_n . Notons que l'on a $\mathbb{P}(X_n = k) = \frac{1}{n!} C_{n,k}$. Ainsi, si on note G_n la fonction génératrice de X_n , on a pour $z \in \mathbb{R}$

$$G_{X_n}(z) = \mathbb{E}[z^{X_n}] = \frac{1}{n!} \sum_{k=0}^n C_{n,k} z^k.$$

Posons $f_n(z) = \sum_{k=0}^n C_{n,k} z^k$ et utilisons la formule de la question précédente. On a

$$\begin{aligned}
f_{n+1}(z) &= \sum_{k=1}^{n+1} C_{n+1,k} z^k = z^{n+1} + n \sum_{k=1}^n C_{n,k} z^k + \sum_{k=1}^n C_{n,k-1} z^k \\
&= z^{n+1} + n \sum_{k=1}^n C_{n,k} z^k + z \sum_{k=0}^{n-1} C_{n,k-1} z^k = (n+z) \sum_{k=1}^n C_{n,k} z^k = (n+z) f_n(z).
\end{aligned}$$

Comme $f_1(z) = C_{1,1}z = z$, on en conclut facilement par récurrence que

$$f_n(z) = \prod_{i=0}^{n-1} (z+i), \quad \forall z \in \mathbb{C}.$$

Donc

$$G_{X_n}(z) = \frac{1}{n!} \prod_{i=0}^{n-1} (z+i).$$

Étape 3. Comparaison avec la fonction génératrice de $Y_1 + \dots + Y_n$. Par indépendance, la fonction génératrice de $Y_1 + \dots + Y_n$ vaut

$$G_{Y_1+\dots+Y_n}(z) = G_{Y_1}(z) \cdots G_{Y_n}(z)$$

où $G_{Y_i}(z) = 1 - \frac{1}{i} + \frac{1}{i}z = \frac{1}{i}(z+i-1)$. On en conclut que

$$G_{Y_1+\dots+Y_n}(z) = \prod_{i=1}^n \frac{1}{i} (z+i-1) = \frac{1}{n!} \prod_{i=0}^{n-1} (z+i) = G_{X_n}(z),$$

ce qui signifie que X_n a la même loi que $Y_1 + \dots + Y_n$. □

Ex. 4

Une caractérisation de la gaussienne : le théorème de Bernstein

Considérons X, Y deux variables aléatoires indépendantes et de même loi, admettant un moment d'ordre deux fini. On suppose pour simplifier que $E(X) = E(Y) = 0$ et $E(X^2) = E(Y^2) = 1$. On va montrer la caractérisation suivante de la loi $N(0, 1)$, due à Bernstein.

Proposition 4.1. *Si X, Y sont deux variables aléatoires indépendantes et de même loi, admettant des moments de tous ordres finis et telles que $E(X) = E(Y) = 0$ et $E(X^2) = E(Y^2) = 1$, alors on a*

$$X, Y \sim N(0, 1) \text{ si et seulement si } X + Y \text{ et } X - Y \text{ sont indépendantes.}$$

À noter que la caractérisation reste valable si on suppose simplement que X, Y ont des moments d'ordre 2 fini (avec des outils qui n'ont pas été introduits ici). Noter aussi que dans tous les cas, on a par bilinéarité et symétrie de la covariance,

$$\text{Cov}(X + Y, X - Y) = \text{Var}(X) + \text{Cov}(Y, X) - \text{Cov}(X, Y) - \text{Var}(Y) = 0,$$

où on a utilisé que $\text{Var}(X) = \text{Var}(Y)$ et $\text{Cov}(X, Y) = \text{Cov}(Y, X)$. Ainsi, si X, Y sont indépendantes (et centrées réduites), même si on a toujours bien $\text{Cov}(X + Y, X - Y) = 0$, le seul cas où $X + Y$ et $X - Y$ sont indépendantes est si $X, Y \sim N(0, 1)$.

Démonstration. Le sens direct est facile. Si $X, Y \sim N(0, 1)$ sont indépendantes, alors (X, Y) est un vecteur gaussien. On déduit que $(X + Y, X - Y)$ est aussi un vecteur gaussien, et comme $\text{Cov}(X + Y, X - Y) = 0$, cela montre que $X + Y$ et $X - Y$ sont indépendantes, voir la Proposition 6.73.

Supposons maintenant que X, Y sont indépendantes (centrées réduites) et que $X + Y$ et $X - Y$ sont indépendantes. Notons $M(t) := M_X(t) = E(e^{tX})$ pour $t \in \mathbb{R}$ la fonction génératrice des moments de X : il suffit de montrer que $M(t) = e^{\frac{1}{2}t^2}$.

a) Si on suppose $M(t) < +\infty$ pour $t \in \mathbb{R}$. Cette hypothèse est plus forte que celle de la proposition, mais la preuve dans ce cas est un peu moins technique[‡]. Rappelons que dans ce cas la fonction $t \mapsto M(t)$ est indéfiniment dérivable et on a notamment $M'(t) = E(Xe^{tX})$ et $M''(t) = E(X^2e^{tX})$.

Calculons de deux manières $E((X - Y)^2 e^{t(X+Y)})$. D'une part, en développant le carré et par linéarité de l'espérance, on obtient

$$\begin{aligned} E((X - Y)^2 e^{t(X+Y)}) &= E(X^2 e^{tX} e^{tY}) - 2E(X e^{tX} Y e^{tY}) + E(Y^2 e^{tX} e^{tY}) \\ &= 2E(X^2 e^{tX}) E(e^{tY}) - 2E(X e^{tX})^2 = 2M''(t)M(t) - 2M'(t)^2, \end{aligned}$$

où on a aussi utilisé le fait que X, Y sont indépendantes et de même loi. D'autre part, en utilisant l'indépendance de $X + Y$ et $X - Y$, on a

[‡]. Et elle s'adapte bien au cas de la *fonction caractéristique* (qui n'a pas été introduite ici), pour donner une preuve générale.

$$\mathbb{E}((X - Y)^2 e^{t(X+Y)}) = \mathbb{E}((X - Y)^2) \mathbb{E}(e^{t(X+Y)}) = 2\mathbb{E}(e^{tX})^2 = 2M(t)^2,$$

où on a utilisé l'indépendance de X, Y , et le fait que $\mathbb{E}(X) = \mathbb{E}(Y) = 0$ et $\mathbb{E}(X^2) = \mathbb{E}(Y^2) = 1$ pour avoir $\mathbb{E}((X - Y)^2) = \mathbb{E}(X^2) - 2\mathbb{E}(XY) + \mathbb{E}(Y^2) = 2$. En combinant les deux égalités, on en conclut que

$$M''(t)M(t) - M'(t)^2 = M(t)^2.$$

Ainsi, si l'on pose $\psi(t) := \log M(t)$ (rappelons que $M(t) > 0$ pour tout $t > 0$), on obtient

$$\psi'(t) = \frac{M'(t)}{M(t)}, \quad \psi''(t) = \frac{M''(t)M(t) - M'(t)^2}{M(t)^2} = 1, \quad \text{pour tout } t \in \mathbb{R}.$$

En notant que $\psi'(0) = \mathbb{E}(X) = 0$, on en déduit que $\psi'(t) = t$ pour tout $t \in \mathbb{R}$. De même, comme $\psi(0) = \log 1 = 0$, on en conclut que $\psi(t) := \log M(t) = \frac{1}{2}t^2$ pour tout $t \in \mathbb{R}$. On a donc montré que $M(t) = e^{\frac{1}{2}t^2}$ pour tout $t \in \mathbb{R}$, ce qui conclut la démonstration.

b) Si on suppose que $\mathbb{E}(X^n) = \mathbb{E}(Y^n) < +\infty$ pour tout $n \in \mathbb{N}$. On va montrer que $\mathbb{E}(X^n) = \mathbb{E}(Z^n)$ pour tout $n \in \mathbb{N}$, où $Z \sim N(0, 1)$: cela montre alors que la série

$$\sum_{n=0}^{\infty} \mathbb{E}(X^n) \frac{t^n}{n!} = \sum_{n=0}^{\infty} \mathbb{E}(Z^n) \frac{t^n}{n!}$$

converge et est égale à $M(t)$, d'où $M(t) = M_Z(t) = e^{\frac{1}{2}t^2}$.

Montrons donc que $\mathbb{E}(X^n) = \mathbb{E}(Z^n)$ pour tout $n \in \mathbb{N}$, par une récurrence forte. C'est évidemment vérifié pour $n = 0, 1, 2$. Soit $n \geq 2$ et supposons que $\mathbb{E}(X^k) = \mathbb{E}(Z^k)$ pour tout $k \in \{0, \dots, n\}$.

Calculons maintenant $\mathbb{E}((X + Y)^{n-1}(X - Y)^2)$ de deux manières. D'une part, en développant $(X + Y)^{n-1}(X - Y)^2$ et en utilisant la linéarité de l'espérance (toutes les espérances sont finies par hypothèse), on obtient

$$\begin{aligned} \mathbb{E}((X + Y)^{n-1}(X - Y)^2) &= \mathbb{E}(X^{n+1} + Y^{n+1}) + \sum_{i,j=0}^n c_{i,j} \mathbb{E}(X^i Y^j) \\ &= 2\mathbb{E}(X^{n+1}) + \sum_{i,j=0}^n c_{i,j} \mathbb{E}(X^i) \mathbb{E}(Y^j), \end{aligned} \tag{E-4.1}$$

où $c_{i,j}$ sont des coefficients donnés (dont la valeur importe peu). De la même manière, si $Z_1, Z_2 \sim N(0, 1)$ sont indépendantes, on a

$$\mathbb{E}((Z_1 + Z_2)^{n-1}(Z_1 - Z_2)^2) = 2\mathbb{E}(Z_1^{n+1}) + \sum_{i,j=0}^n c_{i,j} \mathbb{E}(Z_1^i) \mathbb{E}(Z_2^j).$$

En utilisant l'hypothèse de récurrence, on a $\mathbb{E}(X^i) = \mathbb{E}(Z_1^i)$, $\mathbb{E}(Y^j) = \mathbb{E}(Z_2^j)$ pour tous $i, j \leq n$, donc en faisant la différence des deux dernières identités, on obtient

$$\begin{aligned} 2\mathbb{E}(X^{n+1}) - 2\mathbb{E}(Z_1^{n+1}) &= \mathbb{E}((X + Y)^{n-1}(X - Y)^2) - \mathbb{E}((Z_1 + Z_2)^{n-1}(Z_1 - Z_2)^2) \\ &= \mathbb{E}((X + Y)^{n-1}) \mathbb{E}((X - Y)^2) - \mathbb{E}((Z_1 + Z_2)^{n-1}) \mathbb{E}((Z_1 - Z_2)^2) \\ &= 2\left(\mathbb{E}((X + Y)^{n-1}) - \mathbb{E}((Z_1 + Z_2)^{n-1})\right) \end{aligned}$$

où on a utilisé l'indépendance de $X + Y, X - Y$ (par hypothèse) et de $Z_1 + Z_2, Z_1 - Z_2$ (car Z_1, Z_2 indépendantes de loi $N(0, 1)$). On a aussi utilisé le fait que $E((X - Y)^2) = E((Z_1 - Z_2)^2) = 2$, par un calcul direct (les variables impliquées sont indépendantes, centrées et réduites).

Il reste simplement à voir que $E((X + Y)^{n-1}) = -E((Z_1 + Z_2)^{n-1})$ pour conclure. En effet, en développant $(X + Y)^{n-1}$, en utilisant la linéarité de l'espérance et le fait que X et Y sont indépendantes (de même pour $(Z_1 + Z_2)^{n-1}$), on obtient

$$E((X + Y)^{n-1}) = \sum_{i=1}^n \binom{n-1}{i} E(X^i) E(Y^{n-i}), \quad E((Z_1 + Z_2)^{n-1}) = \sum_{i=1}^n \binom{n-1}{i} E(Z_1^i) E(Z_2^{n-i}).$$

En utilisant l'hypothèse de récurrence, qui donne $E(X^i) = E(Z_1^i)$, $E(Y^j) = E(Z_2^j)$, on obtient que les deux termes sont égaux. On en déduit que $E(X^{n+1}) = E(Z_1^{n+1})$ avec $Z \sim N(0, 1)$, ce qui conclut la récurrence et donc la démonstration de la Proposition 4.1. \square

Ex. 5

Vecteur aléatoire uniforme dans une boule de grande dimension

Soit $Z_n := (X_1, \dots, X_n)$ un vecteur aléatoire de \mathbb{R}^n , de loi uniforme dans la boule de rayon \sqrt{n} , notée $B_n(\sqrt{n}) = \{(x_1, \dots, x_n) \in \mathbb{R}^n, \sum_{i=1}^n x_i^2 \leq n\}$. Autrement dit, Z_n est un vecteur aléatoire de \mathbb{R}^n de densité donnée par

$$f_{Z_n}(x_1, \dots, x_n) = \frac{1}{V_n(\sqrt{n})} \mathbb{1}_{B_n(\sqrt{n})}(x_1, \dots, x_n).$$

Pour tout $k \leq n$, on note $f_{X_1, \dots, X_k}^{(n)}(x_1, \dots, x_k)$ la densité marginale de (X_1, \dots, X_k) dans le vecteur Z_n . On a alors le résultat suivant.

Proposition 5.1. *Pour tout $k \in \mathbb{N}$ fixé, on a pour tous $(x_1, \dots, x_k) \in \mathbb{R}^k$*

$$\lim_{n \rightarrow \infty} f_k^{(n)}(x_1, \dots, x_k) = \prod_{i=1}^k \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x_i^2}.$$

Autrement dit, quand $n \rightarrow \infty$, la densité marginale du vecteur (X_1, \dots, X_k) converge vers celle de k variables $N(0, 1)$ indépendantes.

Commençons par le lemme suivant.

Lemme 5.2. *Soit $V_n(r)$ le volume d'une boule de rayon $r > 0$ en dimension n . Pour tout $n \geq 1$ et tout $r > 0$, on a la formule*

$$V_n(r) = \frac{\pi^{n/2}}{\Gamma(\frac{n}{2} + 1)} r^n.$$

Démonstration. Notons $B_n(r) = \{(x_1, \dots, x_n) \in \mathbb{R}^n, \sum_{i=1}^n x_i^2 \leq r^2\}$ la boule de rayon r dans \mathbb{R}^n . On va démontrer la relation ci-dessus par récurrence. La formule est vérifiée en dimension $n = 1$ car $V_1(r) = 2r$ (noter que $\Gamma(\frac{3}{2}) = \frac{1}{2}\Gamma(\frac{1}{2}) = \frac{1}{2}\sqrt{\pi}$, voir (6.37)) et en dimension $n = 2$ car $V_2(r) = \pi r^2$ (noter que $\Gamma(2) = 1$).

En utilisant le théorème de Fubini–Tonelli, on a la relation de récurrence suivante : pour $n \geq 1$

$$\begin{aligned} V_{n+1}(r) &= \int_{\mathbb{R}^n} \mathbb{1}_{B_{n+1}(r)}(x_1, \dots, x_{n+1}) dx_1 \dots dx_{n+1} \\ &= \int_{-r}^r \left(\int_{\mathbb{R}^n} \mathbb{1}_{\{\sum_{i=1}^n x_i^2 \leq r^2 - x_{n+1}^2\}} dx_1 \dots dx_n \right) dx_{n+1} = \int_{-r}^r V_n(\sqrt{r^2 - x^2}) dx. \end{aligned}$$

Ainsi, en appliquant l'hypothèse de récurrence et par symétrie, on obtient

$$V_{n+1}(r) = \frac{\pi^{n/2}}{\Gamma(\frac{n}{2} + 1)} 2 \int_0^r (r^2 - x^2)^{n/2} dx = \frac{\pi^{n/2}}{\Gamma(\frac{n}{2} + 1)} r^{n+1} \int_0^1 (1 - v)^{n/2} v^{-\frac{1}{2}} dv$$

où on a utilisé le changement de variable $u = \frac{x^2}{r^2}$ pour la deuxième égalité (on laisse les calculs en exercice). Maintenant, la dernière intégrale a déjà été calculée, voir la relation (6.93) : on a

$$\int_0^1 (1-v)^{n/2} v^{-\frac{1}{2}} dv = \beta\left(\frac{n}{2} + 1, \frac{1}{2}\right) = \frac{\Gamma(\frac{n}{2} + 1)\Gamma(\frac{1}{2})}{\Gamma(\frac{n+1}{2} + 1)}.$$

Comme $\Gamma(\frac{1}{2}) = \pi^{1/2}$, on obtient bien la formule $V_{n+1}(r) = \frac{\pi^{(n+1)/2}}{\Gamma(\frac{n+1}{2} + 1)} r^{n+1}$, ce qui conclut la récurrence. \square

Démonstration (de la Proposition 5.1). D'après la Proposition (6.49), la densité marginale est donnée par l'intégrale suivante : pour x_1, \dots, x_k fixés, on a

$$\begin{aligned} f_k^{(n)}(x_1, \dots, x_k) &= \int_{\mathbb{R}^{n-k}} f_{Z_n}(x_1, \dots, x_n) dx_{k+1} \dots dx_n \\ &= \frac{1}{V_n(\sqrt{n})} \mathbb{1}_{\{\sum_{i=1}^k x_i^2 \leq n\}} \int_{\mathbb{R}^{n-k}} \mathbb{1}_{\{\sum_{i=k+1}^n x_i^2 \leq n - \sum_{i=1}^k x_i^2\}} dx_{k+1} \dots dx_n, \end{aligned}$$

de sorte que

$$f_k^{(n)}(x_1, \dots, x_k) = \frac{V_{n-k}(\sqrt{n - \theta_x})}{V_n(\sqrt{n})} \mathbb{1}_{\{\theta_x \leq n\}}, \quad \text{avec } \theta_x := \sum_{i=1}^k x_i^2.$$

On peut maintenant utiliser le Lemme 5.2 : on obtient

$$\begin{aligned} f_k^{(n)}(x_1, \dots, x_k) &= \frac{\Gamma(\frac{n}{2} + 1)}{\pi^{n/2} n^{n/2}} \frac{\pi^{(n-k)/2} (n - \theta_x)^{(n-k)/2}}{\Gamma(\frac{n-k}{2} + 1)} \mathbb{1}_{\{\theta_x \leq n\}} \\ &= \frac{\Gamma(\frac{n}{2} + 1)}{\Gamma(\frac{n-k}{2} + 1)} \frac{1}{n^{k/2}} \frac{1}{\pi^{k/2}} \left(1 - \frac{\theta_x}{n}\right)^{\frac{n}{2} - \frac{k}{2}} \mathbb{1}_{\{\theta_x \leq n\}}. \end{aligned}$$

Il reste maintenant à prendre la limite quand $n \rightarrow \infty$. Pour k fixé et $\theta_x \in \mathbb{R}_+$ fixé, on a

$$\lim_{n \rightarrow \infty} \left(1 - \frac{\theta_x}{n}\right)^{\frac{n}{2}} = e^{-\frac{1}{2}\theta_x}, \quad \lim_{n \rightarrow \infty} \left(1 - \frac{\theta_x}{n}\right)^{-\frac{k}{2}} = 1, \quad \lim_{n \rightarrow \infty} \mathbb{1}_{\{\theta_x \leq n\}} = 1$$

Pour le terme restant, on montrera plus bas (à l'aide des propriétés de la fonction Gamma données dans la Section 6.3.2) que pour $\ell \in \mathbb{N}$,

$$\lim_{\ell \rightarrow \infty} \frac{\Gamma(\frac{1}{2}\ell + 1)}{\Gamma(\frac{1}{2}\ell + \frac{1}{2})\sqrt{\ell}} = \frac{1}{\sqrt{2}}. \quad (\text{E-5.1})$$

Alors, en écrivant le produit télescopique :

$$\frac{1}{n^{k/2}} \frac{\Gamma(\frac{n}{2} + 1)}{\Gamma(\frac{n-k}{2} + 1)} = \frac{1}{n^{k/2}} \prod_{j=0}^{k-1} \frac{\Gamma(\frac{n-j}{2} + 1)}{\Gamma(\frac{n-j}{2} + \frac{1}{2})} = \prod_{j=0}^{k-1} \left(\frac{n-j}{n}\right)^{1/2} \prod_{j=0}^{k-1} \frac{\Gamma(\frac{n-j}{2} + 1)}{\Gamma(\frac{n-j}{2} + \frac{1}{2})\sqrt{n-j}},$$

et pour chaque $j \in \{0, \dots, k-1\}$, en utilisant (E-5.1), on a

$$\lim_{n \rightarrow \infty} \left(\frac{n-j}{n}\right)^{1/2} = 1, \quad \lim_{n \rightarrow \infty} \frac{\Gamma(\frac{n-j}{2} + 1)}{\Gamma(\frac{n-j}{2} + \frac{1}{2})\sqrt{n-j}} = \frac{1}{\sqrt{2}}.$$

Ainsi, pour $k \in \mathbb{N}$ et $(x_1, \dots, x_k) \in \mathbb{R}^k$ fixés, on en déduit que

$$\lim_{n \rightarrow \infty} f_k^{(n)}(x_1, \dots, x_k) = \frac{1}{(\sqrt{2}\pi)^k} e^{-\frac{1}{2}\theta_x} = \frac{1}{(\sqrt{2}\pi)^k} e^{-\frac{1}{2}\sum_{i=1}^k x_i^2},$$

qui est la conclusion désirée.

Pour montrer (E-5.1), on peut utiliser les formules (6.36)-(6.38) de la Section 6.3.2. Notons que l'on peut réécrire la formule (6.37) pour les demi-entiers de la façon suivante : pour $m \in \mathbb{N}$,

$$\Gamma\left(m + \frac{1}{2}\right) = \frac{\sqrt{\pi}}{2^m} \prod_{i=0}^{m-1} (2i+1) = \frac{\sqrt{\pi}}{2^m} \frac{(2m)!}{\prod_{i=1}^m (2i)} = \frac{\sqrt{\pi}}{4^m} \frac{(2m)!}{m!}.$$

Ainsi, en utilisant la formule de Stirling, on obtient

$$\frac{\sqrt{m}\Gamma\left(m + \frac{1}{2}\right)}{\Gamma(m+1)} = \frac{\sqrt{\pi m}}{4^m} \frac{(2m)!}{(m!)^2} \xrightarrow{m \rightarrow \infty} 1, \quad (\text{E-5.2})$$

de manière analogue à (2.13).

On peut maintenant montrer (E-5.1). Dans le cas où ℓ est pair, en considérant l'entier $m := \frac{1}{2}\ell$, on obtient

$$\frac{\Gamma(\frac{1}{2}\ell + 1)}{\Gamma(\frac{1}{2}\ell + \frac{1}{2})\sqrt{\ell}} = \frac{\Gamma(m+1)}{\Gamma(m + \frac{1}{2})\sqrt{2m}} \xrightarrow{m \rightarrow \infty} \frac{1}{\sqrt{2}},$$

en utilisant (E-5.2). Dans le cas où ℓ est impair, en considérant l'entier $m := \frac{\ell}{2} - \frac{1}{2}$, on obtient

$$\frac{\Gamma(\frac{1}{2}\ell + 1)}{\Gamma(\frac{1}{2}\ell + \frac{1}{2})\sqrt{\ell}} = \frac{\Gamma(m + \frac{3}{2})}{\Gamma(m+1)\sqrt{2(m + \frac{1}{2})}} = \frac{1}{\sqrt{2}} \frac{\sqrt{m + \frac{1}{2}}\Gamma(m + \frac{1}{2})}{\Gamma(m+1)} \xrightarrow{m \rightarrow \infty} \frac{1}{\sqrt{2}}$$

où on a aussi utilisé que $\Gamma(\alpha + 1) = \alpha\Gamma(\alpha)$, voir (6.35), puis (E-5.2). Cela conclut la démonstration de (E-5.1), donc de la proposition. \square

Ex. 6

Minimum de variables exponentielles : le lemme des réveils*

Soit $n \geq 1$ un entier naturel fixé et soient X_1, \dots, X_n des variables aléatoires indépendantes. On note

$$R := \min\{X_1, \dots, X_n\},$$

et J le plus petit indice j tel que $X_j = \min\{X_1, \dots, X_n\}$. Autrement dit, s'il existe plusieurs indices j_1, \dots, j_k tels que $X_{j_1} = \dots = X_{j_k} = \min\{X_1, \dots, X_n\}$, on pose $J = j_1$; si le minimum est atteint de manière unique, J est défini de manière univoque comme l'indice pour lequel $X_J = \min\{X_1, \dots, X_n\}$.

On peut interpréter cette situation de la manière suivante : on dispose de n réveils, le k -ème réveil sonnait à une heure aléatoire X_k , indépendamment des autres réveils. Alors, R est le premier instant où un réveil sonne et J est l'indice du réveil qui a sonné. On a alors le résultat suivant.

Proposition 6.1 (Lemme des réveils). *Si les variables aléatoires X_1, \dots, X_n sont indépendantes et de lois exponentielles de paramètres respectifs $\lambda_1 > 0, \dots, \lambda_n > 0$, alors :*

- les variables aléatoires R et J sont indépendantes,
- leurs lois sont données par $R \sim \text{Exp}(\theta)$, $P(J = i) = \frac{\lambda_i}{\theta}$ pour tout $j \in \{1, \dots, n\}$ où on a posé $\theta = \sum_{i=1}^n \lambda_i$.

En particulier, si tous les paramètres λ_i sont égaux ($\lambda_i = \lambda$ pour tout $i \in \{1, \dots, n\}$), alors J est de loi uniforme sur $\{1, \dots, n\}$.

Démonstration. Pour caractériser complètement la loi du vecteur (R, J) , il faut calculer la probabilité $P(R \in A, J \in B)$ pour tous ensembles A, B . Comme R est à valeur dans \mathbb{R}_+ et J est à valeur dans $\{1, \dots, n\}$, on a $P(R \in A, J \in B) = \sum_{j \in B} P(R \in A \cap \mathbb{R}_+, J = j)$, de sorte qu'il suffit de calculer $P(R \in A, J = j)$ pour tout ensemble $A \subset \mathbb{R}_+$ et tout $j \in \{1, \dots, n\}$, car cela caractérise la loi du couple (R, J) à valeurs dans $\mathbb{R}_+ \times \{1, \dots, n\}$.

Tout d'abord, notons que comme les variables aléatoires X_1, \dots, X_n sont à densité, on a $P(X_i \neq X_j \text{ pour tous } i \neq j) = 1$ (voir l'Exercice 6.59). En particulier, le minimum est atteint pour un unique indice j avec probabilité 1. Ainsi, pour tout $j \in \{1, \dots, n\}$ et tout $A \subset \mathbb{R}_+$, on a

$$P(R \in A, J = j) = P(X_j \in A, X_i > X_j \text{ for all } i \neq j).$$

Comme (X_1, \dots, X_n) est un vecteur aléatoire à densité, de densité donnée par le produit des densités f_{X_i} , on obtient

$$\begin{aligned} P(R \leq t, J = j) &= \int_{\mathbb{R}^n} \mathbb{1}_{\{x_j \in A, x_i > x_j \text{ for all } i \neq j\}} \prod_{i=1}^n f_{X_i}(x_i) dx_i \\ &= \int_A \left(\int_{\mathbb{R}^{n-1}} \mathbb{1}_{\{x_i > x_j \text{ for all } i \neq j\}} \prod_{i=1, i \neq j}^n f_{X_i}(x_i) dx_i \right) \lambda_j e^{-\lambda_j x_j} dx_j, \end{aligned}$$

où on a appliqué le théorème de Fubini–Tonelli pour la deuxième identité, ainsi que la forme de la densité de f_{X_j} . Maintenant, pour $x_j \in \mathbb{R}_+$ fixé, l'intégrale intérieure est égale à

$$\int_{\mathbb{R}^{n-1}} \mathbb{1}_{\{x_i > x_j \text{ for all } i \neq j\}} \prod_{i=1, i \neq j}^n f_{X_i}(x_i) dx_i = P(X_i > x_j \text{ for all } i \neq j) = \prod_{i=1, i \neq j}^n P(X_i > x_j),$$

où on a utilisé le fait que les variables $(X_i)_{i \neq j}$ sont indépendantes. Comme $X_i \sim \text{Exp}(\lambda_i)$, on obtient que $P(X_i > x_j) = e^{-\lambda_i x_j}$ pour $x_j \geq 0$, voir (6.44). En combinant avec ce qui précède, on obtient

$$P(R \in A, J = j) = \int_A \lambda_j e^{-\theta x_j} dx_j = \frac{\lambda_j}{\theta} \int_A \theta e^{-\theta x} dx, \quad (\text{E-6.1})$$

où on a posé $\theta = \sum_{i=1}^n \lambda_i$.

Cette identité caractérise complètement la loi du vecteur aléatoire (R, J) . En effet, on peut retrouver la loi marginale de R et J . Tout d'abord, en sommant sur $j \in \{1, \dots, n\}$ et en prenant $A = [0, t]$ pour $t \geq 0$, on a

$$P(R \leq t) = \int_0^t \theta e^{-\theta x} dx = 1 - e^{-\theta t}.$$

On en déduit que $P(R \leq t) = (1 - e^{-\theta t}) \mathbb{1}_{\mathbb{R}_+}(t)$, qui est la fonction de répartition d'une variable aléatoire de loi $\text{Exp}(\theta)$: ainsi, $R \sim \text{Exp}(\theta)$. D'autre part, en prenant $A = \mathbb{R}$ dans (E-6.1), on obtient

$$P(J = j) = \frac{\lambda_j}{\theta} \quad \text{pour tout } j \in \{1, \dots, n\}.$$

Le fait que R et J sont des variables aléatoires indépendantes se voit aussi dans (E-6.1) : on a $P(R \in A, J = j) = P(R \in A) P(J = j)$ pour tous $A \subset \mathbb{R}_+$, $j \in \{1, \dots, n\}$. Le fait que ces probabilités caractérisent la loi suffit pour en déduire que les variables aléatoires R et J sont indépendantes. Il suffit en effet d'étendre cette relation à tous les ensembles A, B , en écrivant comme plus haut

$$P(R \in A, J \in B) = \sum_{j \in B} P(R \in A \cap \mathbb{R}_+, J = j) = \sum_{j \in B} P(R \in A \cap \mathbb{R}_+) P(J = j) = P(R \in A) P(J \in B),$$

ce qui est la définition de l'indépendance de R et J . □

Ex. 7

Maximum de variables aléatoires géométriques et log-périodicité

Soient $(X_i)_{i \geq 1}$ des variables aléatoires indépendantes et de même loi Géom($\frac{1}{2}$). Pour $n \geq 1$, on considère l'événement

$A_n = \ll \text{le maximum } \max\{X_1, \dots, X_n\} \text{ est atteint de manière unique} \gg$,

ou, formulé autrement, $A_n = \ll \text{il existe } i \in \{1, \dots, n\} \text{ tel que } X_i = \max\{X_1, \dots, X_n\} \text{ et } X_j < \max\{X_1, \dots, X_n\} \text{ pour tout } j \neq i \gg$. On a alors le résultat suivant[‡].

Proposition 7.1. *Il existe une fonction continue $f : \mathbb{R} \rightarrow]0, 1[$, périodique de période 1 et non-constante, telle que*

$$\lim_{n \rightarrow \infty} |P(A_n) - f(\log_2 n)| = 0,$$

où $\log_2 x = \frac{\log x}{\log 2}$ désigne le logarithme en base 2.

Plusieurs commentaires s'imposent quant à ce résultat :

- Premièrement, la fonction f étant périodique et continue, elle atteint ses bornes et $0 < \min f(x) < \max f(x) < 1$. Ainsi, la probabilité de l'événement A_n ne tend ni vers 0, ni vers 1 ; on peut en fait le montrer assez facilement, voir les observations préliminaires plus bas.
- Deuxièmement, la probabilité $P(A_n)$ ne converge pas ! On a $P(A_n) \sim f(\log_2 n)$ avec f une fonction qui est périodique non constante (donc oscille) : on dit que $P(A_n)$ est asymptotiquement log₂-périodique.
- La fonction f est explicite, voir le Lemme 7.2 plus bas. Avec un rapide programme Python, on obtient que $\min f(x) \approx 0.72340$ et $\max f(x) \approx 0.72355$ pour tout $x > 0$, voir la Figure E-7.1. Autrement dit, la fonction f est périodique, mais l'amplitude de ses oscillations est de l'ordre de $1,5 \cdot 10^{-5}$...

Observations préliminaires. Avant de montrer le résultat, observons que l'on peut obtenir une borne inférieure et supérieure pour $P(A_n)$ qui permettent de voir facilement que $P(A_n)$ ne tend ni vers 0 ni vers 1. Notons $k_n := \lfloor \log_2 n \rfloor$.

Pour la borne inférieure, notons que

$$P(A_n) \geq P\left(\bigcup_{i=1}^n \{X_i = k_n + 1, X_j \leq k_n \text{ pour } j \neq i\}\right) = \sum_{i=1}^n P(X_i = k_n + 1) P(X_j \leq k_n)^{n-1},$$

où on a utilisé le fait que les événements de la première union sont disjoints et que les variables X_1, \dots, X_n sont indépendantes. Comme $P(X_1 = j) = (\frac{1}{2})^{j-1} \frac{1}{2} = 2^{-j}$ et $P(X_1 > j) = (\frac{1}{2})^j$ pour $j \geq 1$, on obtient

$$P(A_n) \geq n 2^{-(k_n+1)} (1 - 2^{-k_n})^{n-1} \geq \frac{1}{2} \left(1 - \frac{2}{n}\right)^{n-1},$$

[‡]. Inspiré de discussions avec Thomas Duquesne.

où on a utilisé le fait que $\frac{1}{n} \leq 2^{-k_n} \leq \frac{2}{n}$ car $\log_2 n - 1 \leq k_n \leq \log_2 n$. Notons que cette borne inférieure converge, vers $\frac{1}{2}e^{-2} > 0$.

Pour la borne supérieure, on procède de la même manière : on minore $1 - P(A_n) = P(A_n^c)$ en écrivant

$$\begin{aligned} P(A_n^c) &\geq P\left(\bigcup_{1 \leq i < j \leq n} \{X_i = X_j = k_n + 1, X_\ell \leq k_n \text{ pour } \ell \neq i, j\}\right) \\ &= \sum_{1 \leq i < j \leq n} P(X_i = k_n + 1) P(X_j = k_n + 1) P(X_j \leq k_n)^{n-2}. \end{aligned}$$

De même que précédemment, on obtient

$$P(A_n^c) \geq \frac{n(n-1)}{2} 2^{-(k_n+1)} 2^{-(k_n+1)} (1 - 2^{-k_n})^{n-2} \geq \frac{n-1}{8n} \left(1 - \frac{2}{n}\right)^{n-2}.$$

Notons que cette borne inférieure converge, vers $\frac{1}{8}e^{-2} > 0$.

Démonstration de la Proposition 7.1. Commençons par donner une formule explicite pour $P(A_n)$. En effet, en écrivant $A_n = \ll \text{il existe } i \in \{1, \dots, n\} \text{ tel que } X_i > X_j \text{ pour tout } j \neq i \gg$, on obtient

$$\begin{aligned} P(A_n) &= P\left(\bigcup_{i=1}^n \{X_i > X_j \text{ pour tout } j \neq i\}\right) = \sum_{i=1}^n P(X_i > X_j \text{ pour tout } j \neq i) \\ &= n P(X_1 > X_j \text{ pour tout } j \in \{2, \dots, n\}), \end{aligned}$$

où on a utilisé que (X_1, \dots, X_n) a la même loi que $(X_i, X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_n)$ (dans les deux cas, il s'agit de n variables aléatoires indépendantes de loi Géom($\frac{1}{2}$)). Maintenant, en décomposant suivant la valeur de X_1 et en utilisant l'indépendance, on a

$$\begin{aligned} P(X_1 > X_j \text{ pour tout } j \in \{2, \dots, n\}) &= \sum_{k=1}^{\infty} P(X_1 = k + 1, X_j \leq k \text{ pour } j \in \{2, \dots, n\}) \\ &= \sum_{k=1}^{\infty} P(X_1 = k + 1) (1 - P(X_1 > k))^{n-1}. \end{aligned}$$

Comme on a $P(X_1 = j) = (\frac{1}{2})^{j-1} \frac{1}{2} = 2^{-j}$ et $P(X_1 > j) = (\frac{1}{2})^j$ pour $j \geq 1$, on en conclut que

$$P(A_n) = \sum_{k=1}^{\infty} n 2^{-(k+1)} (1 - 2^{-k})^{n-1}. \quad (\text{E-7.1})$$

On introduit maintenant une fonction $g(x)$, dont on donne les propriétés dans le lemme suivant, que l'on démontrera plus bas.

Lemme 7.2. On définit une fonction $g : \mathbb{R}_+^* \rightarrow \mathbb{R}_+^*$, en posant pour $x > 0$,

$$g(x) = \sum_{j \in \mathbb{Z}} x 2^{j-1} e^{-x 2^j} = \sum_{k \in \mathbb{Z}} x 2^{-(k+1)} e^{-x 2^{-k}}. \quad (\text{E-7.2})$$

Alors g est bien définie, elle est continue et non constante, et vérifie $g(x 2^\ell) = g(x)$ pour tous $x > 0$ et $\ell \in \mathbb{Z}$.

Ce lemme montre que la fonction $x \mapsto f(x) := g(2^x)$ est continue, 1-périodique et non-constante ; on a $f(\log_2 n) = g(n)$. On montre alors le résultat suivant :

$$\lim_{n \rightarrow \infty} |P(A_n) - g(n)| = 0, \quad (\text{E-7.3})$$

ce qui conclura la démonstration.

On introduit $\ell_n := \lfloor \frac{2}{3} \log_2 n \rfloor$. En partant de (E-7.1) et de (E-7.2), en appliquant l'inégalité triangulaire, on obtient

$$\begin{aligned} |P(A_n) - g(n)| &\leq \sum_{k=1}^{\ell_n} n 2^{-(k+1)} (1 - 2^{-k})^{n-1} \\ &\quad + \sum_{k > \ell_n} n 2^{-(k+1)} |(1 - 2^{-k})^{n-1} - e^{-n 2^{-k}}| + \sum_{k \in \mathbb{Z}, k \leq \ell_n} n 2^{-(k+1)} e^{-n 2^{-k}}. \end{aligned}$$

Pour le premier terme, en majorant $1 - 2^{-k} \leq 1$ et $1 - 2^{-k} \leq e^{-2^{-k}}$ pour $k \geq 1$, on obtient

$$\sum_{k=1}^{\ell_n} n 2^{-(k+1)} (1 - 2^{-k})^{n-1} \leq \sum_{k=1}^{\ell_n} n 2^{-(k+1)} e^{-n 2^{-k}} \leq n e^{-n 2^{-\ell_n}} \sum_{k=1}^{\infty} 2^{-(k+1)} \leq \frac{1}{2} n e^{-n^{1/3}},$$

où on a utilisé que $\ell_n \leq \frac{2}{3} \log_2 n$ d'où $2^{-\ell_n} \geq n^{-2/3}$. Pour le troisième terme, on utilise la même majoration $1 - 2^{-k} \leq e^{-2^{-k}}$ pour $k \geq 1$, et on obtient, pour $n \geq 2$:

$$\sum_{k \leq 0} n 2^{-(k+1)} e^{-n 2^{-k}} = n \frac{1}{2} \sum_{j \geq 0} 2^{-j} 2^{2j} e^{-n 2^j} \leq \frac{1}{2} n e^{-n} \sum_{j \geq 0} 2^{-(j+1)} = n e^{-n}.$$

Pour l'inégalité, on a utilisé que $2^{2j} e^{-n 2^j} \leq e^{-n}$ pour tout $j \geq 0$, car la fonction $t \mapsto t^2 e^{-nt}$ est décroissante sur $[1, +\infty)$ pour $n \geq 2$ (en effet, sa dérivée vaut $-t(nt - 2)e^{-nt}$). On en conclut que

$$|P(A_n) - g(n)| \leq n e^{-n^{1/3}} + n e^{-n} + \sum_{k > \ell_n} n 2^{-(k+1)} |(1 - 2^{-k})^{n-1} - e^{-n 2^{-k}}|.$$

Il est facile de montrer que, pour t suffisamment petit, on a $1 - t \geq e^{-t-t^2}$. Comme on a la borne $(1 - 2^{-k})^{n-1} \leq e^{-n 2^{-k}}$ valable pour tout $k \geq 1$, on obtient pour $k \geq \ell_n$ et n suffisamment grand,

$$|(1 - 2^{-k})^{n-1} - e^{-n 2^{-k}}| \leq |e^{-(n-1)2^{-k} - (n-1)2^{-2k}} - e^{-n 2^{-k}}| = e^{-n 2^{-k}} |e^{-2^{-k}} e^{-(n-1)2^{-2k}} - 1|.$$

En utilisant maintenant que $1 \geq e^{-2^{-k}} \geq 1 - 2^{-k} \geq 1 - 2^{-\ell_n} \geq 1 - \frac{1}{2} n^{-2/3}$ (on rappelle que l'on a $\ell_n \geq \frac{2}{3} \log_2 n - 1$) et aussi $1 \geq e^{-(n-1)2^{-2k}} \geq 1 - \frac{1}{4} (n-1) n^{-4/3} \geq 1 - \frac{1}{5} n^{-1/3}$ pour $n \geq 5$, on en conclut que

$$0 \leq 1 - e^{-2^{-k}} e^{-(n-1)2^{-2k}} \leq 1 - \left(1 - \frac{1}{2} n^{-1/3}\right) \left(1 - \frac{1}{5} n^{-1/3}\right) \leq \frac{7}{10} n^{-1/3} \leq n^{-1/3}.$$

En combinant toutes les estimées, on obtient

$$|P(A_n) - g(n)| \leq n(e^{-n^{1/3}} + e^{-n}) + n^{-1/3} \sum_{k > \ell_n} n 2^{-(k+1)} e^{-n 2^{-k}} \leq n(e^{-n^{1/3}} + e^{-n}) + n^{1/3} g(n).$$

Comme $g(n)$ est bornée (car $x \mapsto g(2^x)$ est continue 1-périodique), la borne supérieure tend vers 0 quand $n \rightarrow \infty$, ce qui montre (E-7.3) et conclut la démonstration. \square

Démonstration (du Lemme 7.2). Il est facile de voir que $g(x)$ est bien définie, en tant que somme de termes positifs. Elle est finie, car on a

$$\sum_{j \leq 0} x 2^{j-1} e^{-x 2^j} \leq \frac{1}{2} x \sum_{k \geq 0} 2^{-k} < +\infty,$$

$$\sum_{j \geq 1} x 2^{j-1} e^{-x 2^j} = \frac{1}{2} x \sum_{j \geq 1} 2^{-j} 2^{2j} e^{-x 2^j} \leq \frac{1}{2} x c_x \sum_{j \geq 1} 2^{-j} < +\infty,$$

où $c_x = \sup_{j \geq 1} 2^{2j} e^{-x 2^j}$.

Notons de plus que par un simple changement d'indices dans la somme, on a

$$g(2^\ell x) = \sum_{j \in \mathbb{Z}} x 2^{\ell+j-1} e^{-x 2^{\ell+j}} = \sum_{j' \in \mathbb{Z}} x 2^{j'-1} e^{-x 2^{j'}} = g(x).$$

Montrons que $x \mapsto g(x)$ est continue[†]. On écrit $g(x) = \frac{1}{2}x(h_1(x) + h_2(x))$, où on a séparé la somme en termes $j \geq 0$ et $j \leq -1$, c'est-à-dire

$$h_1(x) = \sum_{j \geq 0} 2^j e^{-x 2^j}, \quad h_2(x) = \sum_{k \geq 1} 2^{-k} e^{-x 2^{-k}},$$

et on va montrer que h_1, h_2 sont continues. On se concentre sur h_1 , le même type de raisonnement pouvant être appliqué à h_2 . Pour $x, y \in [a, +\infty[$ avec $x < y$, on a grâce à l'inégalité triangulaire

$$|h_1(x) - h_1(y)| \leq \sum_{j \geq 0} 2^j e^{-x 2^j} |1 - e^{-(y-x) 2^j}| \leq (x-y) \sum_{j \geq 0} 2^{2j} e^{-a 2^j},$$

où on a utilisé $0 \leq 1 - e^{-t} \leq t$ pour $t \geq 0$. On a donc montré que pour tout $a > 0$, $|h_1(x) - h_1(y)| \leq C_a |x - y|$ pour tous $x, y \in [a, +\infty[$, et en particulier $\lim_{y \rightarrow x} |h_1(x) - h_1(y)| = 0$, ce qui montre la continuité de h_1 .

Le fait que g ne soit pas constante n'est pas si facile à voir, mais on présente dans la Figure E-7.1 les graphes des fonctions $g(x)$ sur $[1, 128]$ et $f(x) = g(2^x)$ pour $x \in [0, 7]$. \square

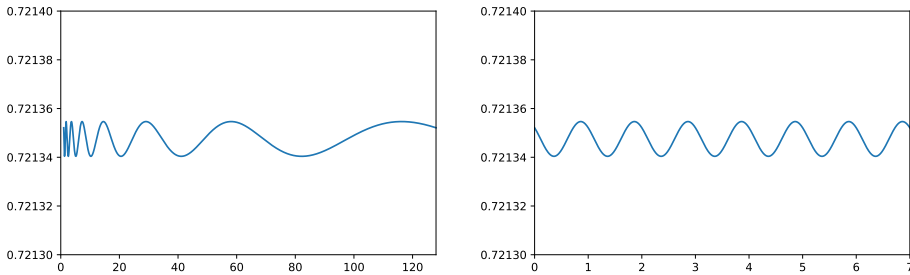


FIG. E-7.1 Graphes des fonctions $x \mapsto g(x)$ pour $x \in [1, 128]$ et $x \mapsto f(x) = g(2^x)$ pour $x \in [0, 7]$. On remarque que f, g sont non-constantes et ont une amplitude d'oscillation de l'ordre de 10^{-5} autour de la valeur ≈ 0.72135 .

Probabilité que le maximum soit atteint par exactement ℓ variables. On peut refaire la même démonstration pour obtenir une estimée de la probabilité que le maximum soit atteint par *exactement* ℓ variables. Notons cet événement $A_n(\ell)$, c'est-à-dire

[†]. On peut en fait montrer que la fonction g est indéfiniment dérivable.

$$A_n(\ell) := \left\{ \text{il existe } I \subset \{1, \dots, n\} \text{ avec } |I| = \ell \text{ tel que :} \right. \\ \left. \forall i \in I X_i = \max\{X_1, \dots, X_n\} \text{ et } \forall j \notin I X_j < \max\{X_1, \dots, X_n\} \right\}.$$

Alors le même calcul que dans (E-7.1) donne la formule

$$P(A_n(\ell)) = \sum_{k=1}^{\infty} \binom{n}{\ell} 2^{-\ell(k+1)} (1 - 2^{-k})^{n-\ell}.$$

On peut alors appliquer le même raisonnement que précédemment pour obtenir le comportement suivant : pour ℓ fixé, quand $n \rightarrow \infty$,

$$P(A_n(\ell)) \sim f_\ell(\log_2 n) = g_\ell(n) \quad \text{avec } g_\ell(x) := \frac{1}{\ell!} \sum_{j \in \mathbb{Z}} x^\ell 2^{(j-1)\ell} e^{-x2^j}.$$

De même que précédemment, f_ℓ est une fonction 1-périodique.

Observation 7.3. Remarquons que par le théorème de Fubini–Tonelli, on a pour $x > 0$

$$\sum_{\ell=1}^{+\infty} g_\ell(x) = \sum_{j \in \mathbb{Z}} \sum_{\ell=1}^{\infty} \frac{1}{\ell!} x^\ell 2^{(j-1)\ell} e^{-x2^j} = \sum_{j \in \mathbb{Z}} (e^{x2^{(j-1)}} - 1) e^{-x2^j} = \sum_{j \in \mathbb{Z}} (e^{-x2^{(j-1)}} - e^{-x2^j}).$$

Notons que l'on ne peut pas écrire que cette somme est pas égale à $\sum_{j \in \mathbb{Z}} e^{-x2^{(j-1)}} - \sum_{j \in \mathbb{Z}} e^{-x2^j}$ car ces deux sommes sont infinies. Mais on voit qu'il s'agit d'une somme télescopique : on a

$$\sum_{j \in \mathbb{Z}} (e^{-x2^{(j-1)}} - e^{-x2^j}) = \lim_{n \rightarrow \infty} \sum_{j=-n}^n (e^{-x2^{(j-1)}} - e^{-x2^j}) = \lim_{n \rightarrow \infty} (e^{-x2^{-n-1}} - e^{-x2^n}) = 1.$$

On en conclut que $\sum_{\ell=1}^{+\infty} g_\ell(x) = 1$. Comme $g_\ell(x) > 0$ pour tout ℓ et $x > 0$, cela montre notamment que g_ℓ est à valeurs dans $]0, 1[$.

Ex. 8

Probabilité que deux variables géométriques soient premières entre elles

On considère X, Y deux variables aléatoires indépendantes, toutes deux de loi géométrique de paramètre $\alpha \in]0, 1[$. On pose $Z_\alpha = \text{pgcd}(X, Y)$ et on note $p_k(\alpha) := P(Z_\alpha = k)$ pour tout $k \in \mathbb{N}^*$. En particulier, $p_1(\alpha)$ est la probabilité que X et Y sont premières entre elles. On montre le résultat suivant.

Proposition 8.1. *Pour tout entier $k \in \mathbb{N}^*$, on a*

$$\lim_{\alpha \downarrow 0} p_k(\alpha) = \lim_{\alpha \downarrow 0} P(Z_\alpha = k) = \frac{c_0}{k^2}, \quad \text{avec } c_0 = \frac{6}{\pi^2}.$$

En particulier, la probabilité que X et Y soient premières entre elles converge vers $\frac{6}{\pi^2}$ quand $\alpha \downarrow 0$.

Noter que le fait de prendre un paramètre $\alpha \downarrow 0$ correspond à considérer une variable aléatoire géométrique de moyenne $1/\alpha \rightarrow +\infty$; la limite $\alpha = 0$ serait de manière informelle « uniforme sur \mathbb{N} ». On s'intéresse donc à la probabilité que deux nombres aléatoires (très grands en moyenne) soient premiers entre eux, dans le même esprit que le Problème 1.27 (les calculs sont d'ailleurs un peu similaires ici, mais pas complètement).

Démonstration. Commençons par un petit résultat préliminaire, intéressant en soi.

Lemme 8.2. *Soit $q_k(\alpha) := P(X \text{ est divisible par } k)$. Alors*

$$q_k(\alpha) = \frac{\alpha(1-\alpha)^{k-1}}{1-(1-\alpha)^k}.$$

Notamment, $\lim_{\alpha \downarrow 0} q_k(\alpha) = \frac{1}{k}$ et $q_k(\alpha) \leq \frac{1}{k}$ pour tout $\alpha \in]0, 1[$. De plus, pour tout $j \in \mathbb{N}^$,*

$$P\left(\frac{1}{k}X = j \mid X \text{ est divisible par } k\right) = (1-\alpha_k)^{j-1}\alpha_k, \quad \text{avec } \alpha_k := 1-(1-\alpha)^k.$$

Autrement dit, conditionnellement à l'événement « X est divisible par k », $\frac{1}{k}X$ suit une loi géométrique de paramètre α_k .

Démonstration. Notons A_k l'événement « X est divisible par k », que l'on peut ré-écrire comme $A_k = \bigcup_{j=1}^{\infty} \{X = kj\}$. On a

$$P(A_k) = \sum_{j=1}^{\infty} P(X = kj) = \sum_{j=1}^{\infty} (1-\alpha)^{kj-1}\alpha = \frac{\alpha}{1-\alpha} \sum_{j=1}^{\infty} ((1-\alpha)^k)^j = \frac{\alpha}{1-\alpha} \frac{(1-\alpha)^k}{1-(1-\alpha)^k},$$

d'où la formule pour $q_k(\alpha)$.

Le fait que $\lim_{\alpha \downarrow 0} q_k(\alpha) = \frac{1}{k}$ vient du fait que $\lim_{\alpha \downarrow 0} \frac{1-(1-\alpha)^k}{\alpha} = k$ (il s'agit de la dérivée en 0 de la fonction $x \mapsto -(1-x)^k$). Pour voir que $q_k(\alpha) \leq \frac{1}{k}$ pour tout $\alpha \in]0, 1[$, on peut

observer que $x \mapsto q_k(x)$ est décroissante (par exemple en calculant sa dérivée[†]) : on en déduit que pour tout $x \in [0, 1]$, $q_k(x) \leq \lim_{x \downarrow 0} q_k(x) = \frac{1}{k}$.

Pour la deuxième formule, on a aisément, en posant $\alpha_k := 1 - (1 - \alpha)^k$ de sorte que $q_k(\alpha) = \frac{1 - \alpha_k}{\alpha_k} \frac{\alpha}{1 - \alpha}$, on obtient

$$\begin{aligned} P\left(\frac{1}{k}X = j \mid X \text{ est divisible par } k\right) &= \frac{1}{P(A_k)} P(X = kj) = \frac{\alpha_k}{1 - \alpha_k} \frac{1 - \alpha}{\alpha} \times (1 - \alpha)^{kj-1} \alpha \\ &= \alpha_k (1 - \alpha_k)^{j-1}, \end{aligned}$$

ce qui donne le résultat voulu. \square

Étape 1. Expression de $P(Z_\alpha = 1)$. Montrons que la probabilité $p_1(\alpha)$ que X et Y soient premiers entre eux converge. On peut utiliser la formule d'inclusion/exclusion de la Proposition 1.23, de la même manière que pour la probabilité que deux entiers pris dans $\{1, \dots, n\}$ soient premiers entre eux, voir le Problème 1.27.

On commence par écrire que $Z_\alpha \neq 1$ si et seulement si X et Y admettent un diviseur commun : ainsi,

$$P(Z_\alpha = 1) = 1 - P(Z_\alpha \neq 1) = 1 - P\left(\bigcup_{i=1}^{\infty} D_{p_i}\right),$$

où p_1, \dots, p_n, \dots désignent les nombres premiers rangés par ordre croissant et où $D_k = \ll k \text{ divise } X \text{ et } Y \gg$. Notons que par indépendance, pour tout $k \geq 1$ on a $P(D_k) = q_k(\alpha)^2$.

On ne peut pas appliquer la formule d'inclusion-exclusion à l'union infinie, mais on peut utiliser le fait que

$$P\left(\bigcup_{i=1}^{\infty} D_{p_i}\right) = \lim_{n \rightarrow \infty} \mathbb{P}\left(\bigcup_{i=1}^n D_{p_i}\right),$$

par continuité croissante des probabilités. On a alors, par la formule d'inclusion-exclusion

$$\begin{aligned} P\left(\bigcup_{i=1}^n D_{p_i}\right) &= \sum_{j=1}^n (-1)^{j+1} \sum_{1 \leq i_1 < \dots < i_j \leq n} P(D_{p_{i_1}} \cap \dots \cap D_{p_{i_j}}) \\ &= \sum_{j=1}^n (-1)^{j+1} \sum_{1 \leq i_1 < \dots < i_j \leq n} P(D_{p_{i_1} \dots p_{i_j}}) = \sum_{j=1}^n (-1)^{j+1} \sum_{1 \leq i_1 < \dots < i_j \leq n} q_{p_{i_1} \dots p_{i_j}}(\alpha)^2. \end{aligned}$$

Notons que l'ensemble des $p_{i_1} \dots p_{i_j}$ est exactement l'ensemble des $d \geq 2$ dont les facteurs premiers sont dans $\{p_1, \dots, p_n\}$ et de multiplicité 1. En utilisant la définition de la fonction de Moebius μ (qui apparaît dans (1.25)), en notant $\mathcal{C}_n = \{\prod_{i=1}^n p_i^{m_i}, (m_1, \dots, m_n) \in \mathbb{N}^n\}$, on obtient

$$1 - \mathbb{P}\left(\bigcup_{i=1}^n D_{p_i}\right) = 1 + \sum_{k \in \mathcal{C}_n \setminus \{1\}} \mu(k) q_k(\alpha)^2 = \sum_{k \in \mathcal{C}_n} \mu(k) q_k(\alpha)^2.$$

On peut alors prendre la limite quand $n \rightarrow \infty$, car la somme converge absolument, en utilisant le fait que $q_k(\alpha)^2 \leq 1/k^2$ et $|\mu(d)| \leq 1$. On a donc montré que

$$p_1(\alpha) = P(Z_\alpha = 1) = \sum_{k=1}^{\infty} \mu(k) q_k(\alpha)^2.$$

[†]. Le calcul donne $q'_k(x) = \frac{(1-x)^{k-2}}{(1-(1-x)^k)^2} (1-kx - (1-x)^k)$, qui est négatif car $(1-x)^k \geq 1-kx$ pour $x \in [0, 1]$ (en observant que $x \mapsto (1-x)^k - 1 + kx$ est nulle en $x = 0$ et de dérivée $k(1 - (1-x)^{k-1}) \geq 0$, donc croissante).

Étape 2. Limite de $P(Z_\alpha = 1)$. On sait que $\lim_{\alpha \downarrow 0} q_k(\alpha)^2 = \frac{1}{k^2}$ et il reste donc à montrer que l'on peut intervertir la limite avec la somme (infinie).

Commençons par observer que $\sum_{k=1}^{\infty} \frac{\mu(k)}{k^2}$ est bien définie car la somme converge absolument. Fixons $\varepsilon > 0$, et $K = K(\varepsilon)$ tel que $\sum_{k>K} \frac{1}{k^2} \leq \varepsilon$. En appliquant l'inégalité triangulaire et en utilisant que $p_1(\alpha) = \sum_{k=1}^{\infty} \mu(k)q_k(\alpha)^2$, on peut écrire

$$\begin{aligned} \left| p_1(\alpha) - \sum_{k=1}^{\infty} \frac{\mu(k)}{k^2} \right| &\leq \left| \sum_{k>K} \mu(k)q_k(\alpha)^2 \right| + \left| \sum_{k=1}^K \mu(k)q_k(\alpha)^2 - \sum_{k=1}^K \frac{\mu(k)}{k^2} \right| + \left| \sum_{k>K} \frac{\mu(k)}{k^2} \right| \\ &\leq 2 \sum_{k>K} \frac{1}{k^2} + \sum_{k=1}^K \left| q_k(\alpha) - \frac{1}{k} \right| \leq 2\varepsilon + \sum_{k=1}^K \left| q_k(\alpha) - \frac{1}{k} \right|, \end{aligned}$$

où on a utilisé que $|\mu(k)| \leq 1$ et $q_k(\alpha)^2 \leq \frac{1}{k^2}$.

Maintenant, comme $\lim_{\alpha \downarrow 0} q_k(\alpha) = \frac{1}{k}$ pour tout $k \in \{1, \dots, K\}$ (avec K fixé), on peut choisir α_0 suffisamment petit (à quel point dépend de ε) pour avoir, pour tout $\alpha \leq \alpha_0$, $|q_k(\alpha) - \frac{1}{k}| \leq \varepsilon/K$. On en conclut que, pour tout $\alpha \leq \alpha_0$ on a $|p_1(\alpha) - \sum_{k=1}^{\infty} \frac{\mu(k)}{k^2}| \leq 3\varepsilon$, ce qui montre que

$$\lim_{\alpha \downarrow 0} p_1(\alpha) = \sum_{k=1}^{\infty} \frac{\mu(k)}{k^2}.$$

On a déjà vu dans le Problème 1.27 que cette somme est égale à $6/\pi^2$, mais on va le redémontrer plus bas.

Étape 3. Expression et limite de $P(Z_\alpha = k)$. Montrons maintenant la relation suivante : On a donc montré que, pour tout $k \in \mathbb{N}^*$,

$$p_k(\alpha) = q_k(\alpha)^2 p_1(\alpha_k), \quad (\text{E-8.1})$$

où $\alpha_k = 1 - (1 - \alpha)^k$ est donné dans le Lemme 8.2. L'idée est d'écrire

$$p_k(\alpha) = P(Z_\alpha = k) = P\left(X \text{ et } Y \text{ sont divisibles par } k \text{ et } \text{pgcd}\left(\frac{1}{k}X, \frac{1}{k}Y\right) = 1\right),$$

et d'utiliser que, conditionnellement au fait que X et Y sont divisibles par k , $\frac{1}{k}X$ et $\frac{1}{k}Y$ sont des variables aléatoires indépendantes de loi Géom(α_k), grâce au Lemme 8.2. Écrivons-le proprement : en décomposant suivant les valeurs de $\frac{1}{k}X$ et $\frac{1}{k}Y$, on a

$$\begin{aligned} p_k(\alpha) &= \sum_{j, j', \text{pgcd}(j, j')=1} P(X \text{ et } Y \text{ sont divisibles par } k \text{ et } \frac{1}{k}X = j, \frac{1}{k}Y = j') \\ &= \sum_{j, j', \text{pgcd}(j, j')=1} P(X \text{ est divisible par } k \text{ et } \frac{1}{k}X = j) P(Y \text{ est divisible par } k \text{ et } \frac{1}{k}Y = j'), \end{aligned}$$

où on a utilisé l'indépendance de X et Y . En notant \hat{X}, \hat{Y} deux variables aléatoires indépendantes de loi Géom(α_k), on obtient grâce au Lemme 8.2

$$\begin{aligned} p_k(\alpha) &= \sum_{j, j', \text{pgcd}(j, j')=1} q_k(\alpha)^2 P(\hat{X} = j) q_k(\alpha) P(\hat{Y} = j') \\ &= q_k(\alpha)^2 \sum_{j, j', \text{pgcd}(j, j')=1} P(\hat{X} = j, \hat{Y} = j') \\ &= q_k(\alpha)^2 P(\text{pgcd}(\hat{X}, \hat{Y}) = 1). \end{aligned}$$

On a donc montré (E-8.1). Maintenant, on sait grâce à l'étape 2 que la limite $\lim_{\alpha \downarrow 0} p_1(\alpha) = c_0$ existe : en prenant la limite $\alpha \downarrow 0$ dans (E-8.1) et en utilisant que $\lim_{\alpha \downarrow 0} q_k(\alpha) = \frac{1}{k}$ et que $\alpha_k \rightarrow 0$ quand $\alpha \rightarrow 0$, on obtient

$$\lim_{\alpha \downarrow 0} p_k(\alpha) = \frac{1}{k^2} c_0, \quad (\text{E-8.2})$$

ce qui est le résultat voulu (une fois que la constante c_0 a été déterminée).

Étape 4. Détermination de la constante $c_0 = \lim_{\alpha \downarrow 0} p_1(\alpha)$. En sommant sur k la relation (E-8.1), on obtient

$$\sum_{k=1}^{\infty} p_k(\alpha) = 1 = \sum_{k=1}^{\infty} q_k(\alpha)^2 p_1(\alpha_k). \quad (\text{E-8.3})$$

Si pouvait échanger la somme et la limite $\alpha \downarrow 0$ dans (E-8.3) on aurait $1 = \sum_{k=1}^{\infty} \frac{c_0}{k^2}$, d'où $c_0 = \left(\sum_{k=1}^{\infty} \frac{1}{k^2} \right)^{-1} = \frac{6}{\pi^2}$. On va démontrer ce résultat à la main.

Fixons $\varepsilon > 0$ et $K = K(\varepsilon)$ tel que $\sum_{k>K} \frac{1}{k^2} \leq \varepsilon$. En reprenant (E-8.3), on obtient grâce à l'inégalité triangulaire

$$\begin{aligned} \left| 1 - \sum_{k=1}^{\infty} \frac{c_0}{k^2} \right| &\leq \sum_{k>K} p_k(\alpha) + \left| \sum_{k=1}^K p_k(\alpha) - \sum_{k=1}^K \frac{c_0}{k^2} \right| + \sum_{k>K} \frac{c_0}{k^2} \\ &\leq 2 \sum_{k>K} \frac{1}{k^2} + \sum_{k=1}^K \left| p_k(\alpha) - \frac{c_0}{k^2} \right| \leq 2\varepsilon + \sum_{k=1}^K \left| p_k(\alpha) - \frac{c_0}{k^2} \right|, \end{aligned}$$

où on a utilisé le fait que $p_k(\alpha) = q_k(\alpha)^2 p_1(\alpha_k) \leq \frac{1}{k^2}$. Cette inégalité étant valable pour tout $\alpha > 0$, on peut prendre la limite $\alpha \downarrow 0$ dans le membre de droite (noter que K est fixé) : on obtient $|1 - \sum_{k=1}^{\infty} \frac{c_0}{k^2}| \leq 2\varepsilon$, et comme $\varepsilon > 0$ est arbitraire, on en déduit que

$$1 = \sum_{k=1}^{\infty} \frac{c_0}{k^2} = c_0 \frac{\pi^2}{6}.$$

Cela permet de conclure que $c_0 = 6/\pi^2$, comme annoncé plus haut. \square

Observation 8.3. On peut adapter la Proposition 8.1 au cas où $X \sim \text{Géom}(\alpha)$ et $Y \sim \text{Géom}(\beta)$, avec $\alpha, \beta \downarrow 0$. \square

Ex. 9

Problème de survie d'une marche aléatoire

Soient $(X_i)_{i \geq 0}$ des variables aléatoires réelles, indépendantes et de même loi. On pose $S_0 = 0$ et pour $k \geq 1$,

$$S_k := \sum_{i=1}^k X_i.$$

La suite $(S_k)_{k \geq 0}$, appelée *marche aléatoire*, permet en pratique de décrire de nombreuses situations, comme le gain d'un joueur lors d'une succession de paris, la position d'un marcheur (ou d'une particule) après k pas effectués de manière aléatoire.

On s'intéresse ici au problème dit *de survie* de la marche aléatoire, c'est-à-dire à la probabilité que $(S_k)_{k \geq 0}$ reste (strictement) positive pendant un temps $n \geq 1$ donné : on cherche à estimer les probabilités

$$P(S_1 \geq 0, \dots, S_n \geq 0) \quad \text{et} \quad P(S_1 > 0, \dots, S_n > 0).$$

Par exemple, si $(S_k)_{k \geq 0}$ représente le gain d'un joueur après k paris, il s'agit de la probabilité que le joueur ne soit jamais endetté au cours des n premières parties.

On va considérer ce problème dans deux cas :

- Le cas de la marche aléatoire simple étudié dans les Sections 2.2 et 4.4, c'est-à-dire si les pas sont à valeur dans $\{+1, -1\}$, avec $P(X_i = +1) = P(X_i = -1) = \frac{1}{2}$;
- Le cas général d'une marche aléatoire *symétrique*, c'est-à-dire où les variables aléatoires $(X_i)_{i \geq 1}$ sont telles que $-X_i$ a la même loi que X_i . Le cas où les variables $(X_i)_{i \geq 1}$ sont à densité s'avère surprenant.

9.1 Pour la marche aléatoire simple

Considérons dans un premier temps le cas de la marche aléatoire simple, c'est-à-dire si les variables aléatoires $(X_i)_{i \geq 1}$ sont i.i.d. de loi

$$P(X_i = +1) = P(X_i = -1) = \frac{1}{2}.$$

On a alors le résultat suivant pour la probabilité de survie. Comme la marche aléatoire simple ne peut revenir en 0 qu'aux instants pairs, on s'intéresse aux probabilités de survie aux instants pairs. En effet, on a par exemple $P(S_1 \geq 0, \dots, S_{2n-1} \geq 0) = P(S_1 \geq 0, \dots, S_{2n} \geq 0)$ ou encore $P(S_1 > 0, \dots, S_{2n} > 0) = P(S_1 > 0, \dots, S_{2n+1} > 0)$ pour tout $n \geq 1$. Comme dans la Section 2.2, on note

$$u_{2n} := \frac{1}{4^n} \binom{2n}{n} = P(S_{2n} = 0), \quad (\text{E-9.1})$$

où la deuxième égalité vient du fait que l'on a $S_{2n} = 0$ si et seulement si il y a eu autant de pas '+1' que de pas '-1' dans les $2n$ premiers pas, voir (2.11).

Théorème 9.1. Soit $(S_k)_{k \geq 0}$ la marche aléatoire simple symétrique définie ci-dessus. Pour tout $n \geq 1$, on a

$$\begin{aligned} P(S_1 \geq 0, \dots, S_{2n} \geq 0) &= u_{2n}, \\ P(S_1 > 0, \dots, S_{2n} > 0) &= \frac{1}{2} u_{2n}. \end{aligned}$$

Comme la marche aléatoire simple $(S_k)_{k \geq 0}$ ne peut pas changer de signe sans passer par 0, il n'y a que deux possibilités pour avoir $S_1 \neq 0, \dots, S_{2n} \neq 0$: soit on a $S_1 > 0, \dots, S_{2n} > 0$, soit on a $S_1 < 0, \dots, S_{2n} < 0$. Ainsi, on en déduit

$$P(S_1 \neq 0, \dots, S_{2n} \neq 0) = P(S_1 > 0, \dots, S_{2n} > 0) + P(S_1 < 0, \dots, S_{2n} < 0) = u_{2n}. \quad (\text{E-9.2})$$

On en déduit le corollaire suivant, concernant l'instant de premier retour en 0 de la marche aléatoire simple.

Corollaire 9.2. Soit $T := \min\{k \geq 1, S_k = 0\}$ l'instant premier retour en 0 de la marche aléatoire simple. Pour tout $n \geq 1$, on a

$$P(T = 2n) = \frac{1}{2n-1} u_{2n}.$$

Démonstration. Commençons par noter que $T > 2n$ si et seulement si on a $S_1 \neq 0, \dots, S_{2n} \neq 0$. Ainsi, d'après le Théorème 9.1, ou plus précisément (E-9.2), on a

$$P(T > 2n) = P(S_1 \neq 0, \dots, S_{2n} \neq 0) = u_{2n}.$$

Notons déjà que cela montre que $P(T = +\infty) = \lim_{n \rightarrow \infty} P(T > 2n) = 0$.

Maintenant, observons que T ne peut prendre que des valeurs paires, car la marche aléatoire ne revient en 0 qu'à des instants pairs. On a donc $T = 2n$ si et seulement si on a $T > 2n-2$ (donc $T > 2n-1$) mais pas $T > 2n$: on obtient

$$P(T = 2n) = P(T > 2n-2) - P(T > 2n) = u_{2(n-1)} - u_{2n}.$$

Notons que $u_{2(n-1)} = \frac{1}{4^{n-1}} \binom{2n-2}{n-1}$ et que $\binom{2n-2}{n-1} = \frac{(2n-2)!}{(n-1)!^2} = \frac{n^2}{2n(2n-1)} \times \frac{(2n-2)!}{(n-1)!^2} = \frac{n}{4n-2} \binom{2n}{n}$.

Ainsi, $u_{2(n-1)} = \frac{4n}{4n-2} u_{2n}$, et

$$P(T = 2n) = u_{2n} \left(\frac{4n}{4n-2} - 1 \right) = \frac{1}{2n-1} u_{2n},$$

ce qui conclut la démonstration. \square

Observation 9.3. On peut aussi déduire du Théorème 9.1 une identité combinatoire intéressante : pour tout $n \geq 1$, on a

$$\sum_{k=0}^n u_{2k} u_{2(n-k)} = 1 \quad \Longleftrightarrow \quad \sum_{k=0}^n \binom{2k}{k} \binom{2(n-k)}{n-k} = 4^n.$$

En effet, en notant $L_{2n} := \max\{0 \leq k \leq 2n, S_k = 0\}$ le dernier instant de retour en 0 de la marche aléatoire simple avant l'instant $2n$, on obtient $\sum_{k=0}^n P(L_{2n} = 2k) = 1$, car L_{2n} est

forcément pair. Mais en écrivant explicitement l'événement $\{L_{2n} = 2k\}$, on a

$$P(L_{2n} = 2k) = P(S_{2k=0}, S_{2k+1} \neq 0, \dots, S_{2n} \neq 0) = P(S_{2k=0}) P(S_1 \neq 0, \dots, S_{2n-2k} \neq 0).$$

Comme $P(S_{2k} = 0) = u_{2k}$ et $P(S_1 \neq 0, \dots, S_{2j} \neq 0) = u_{2j}$ (voir (E-9.2)), cela conclut la démonstration. \square

Démonstration (du Théorème 9.1). On va utiliser une démonstration combinatoire. Comme dans la Section 2.2, notons

$$\tilde{\mathcal{Q}}_{2n} = \{(s_0, \dots, s_{2n}) : s_0 = 0, |s_k - s_{k-1}| = 1 \text{ pour tout } k \in \{1, \dots, 2n\}\}$$

l'ensemble des chemins possibles de longueur $2n$. La loi de la marche aléatoire simple (S_0, \dots, S_{2n}) correspond à la probabilité uniforme sur $\tilde{\mathcal{Q}}_{2n}$.

On considère maintenant les deux sous-ensembles de $\tilde{\mathcal{Q}}_{2n}$ suivant :

$$\begin{aligned} \mathcal{A}_{2n} &:= \{(s_0, \dots, s_{2n}) \in \tilde{\mathcal{Q}}_{2n} : s_{2n} = 0\}, \\ \mathcal{B}_{2n} &:= \{(s_0, \dots, s_{2n}) \in \tilde{\mathcal{Q}}_{2n} : s_k \geq 0 \text{ pour tout } k \in \{1, \dots, 2n\}\}, \end{aligned}$$

de sorte que $P(S_1 \geq 0, \dots, S_{2n} \geq 0) = \frac{1}{|\tilde{\mathcal{Q}}_{2n}|} |\mathcal{B}_{2n}|$ et $P(S_{2n}=0) = \frac{1}{|\tilde{\mathcal{Q}}_{2n}|} |\mathcal{A}_{2n}|$. Il nous reste donc à montrer que $|\mathcal{A}_{2n}| = |\mathcal{B}_{2n}|$.

Pour voir que $|\mathcal{A}_{2n}| = |\mathcal{B}_{2n}|$, on va donner une bijection entre \mathcal{A}_{2n} et \mathcal{B}_{2n} . Il nous suffit en fait donner, pour tout $k \in \{0, \dots, n\}$, une bijection entre $\mathcal{A}_{2n,k}$ et $\mathcal{B}_{2n,k}$, où

$$\begin{aligned} \mathcal{A}_{2n,k} &:= \{(s_0, \dots, s_{2n}) \in \mathcal{A}_{2n} : \min_{0 \leq i \leq 2n} s_i = -k\}, \\ \mathcal{B}_{2n,k} &:= \{(s_0, \dots, s_{2n}) \in \mathcal{B}_{2n} : s_{2n} = 2k\}. \end{aligned}$$

En effet, \mathcal{A}_{2n} est l'union disjointe des $(\mathcal{A}_{2n,k})_{0 \leq k \leq n}$ (noter que si $s_{2n} = 0$, le minimum ne peut pas être inférieur à $-n$) et \mathcal{B}_{2n} est l'union disjointe des $(\mathcal{B}_{2n,k})_{0 \leq k \leq n}$ (noter que s_{2n} est forcément pair, et positif).

Soit $k \in \{0, \dots, n\}$. Une bijection ψ entre $\mathcal{A}_{2n,k}$ et $\mathcal{B}_{2n,k}$ est donnée de la manière suivante. Soit $s = (s_0, \dots, s_{2n})$ un chemin de $\mathcal{A}_{2n,k}$ et soit $m := \min\{i : s_i = -k\}$ le premier instant où le chemin atteint son minimum. On associe alors à s un chemin $\tilde{s} := \psi(s) = (\tilde{s}_0, \dots, \tilde{s}_{2n})$ de la manière suivante :

$$\begin{aligned} \tilde{s}_i &= s_{m+i} - s_m \quad \text{pour } i \in \{0, \dots, 2n-m\} \quad (\text{noter que } \tilde{s}_{2n-m} = k), \\ \tilde{s}_i &= k + (s_m - s_{2n-i}) \quad \text{pour } i \in \{2n-m+1, \dots, 2n\} \quad (\text{noter que } \tilde{s}_{2n} = 2k). \end{aligned}$$

Autrement dit, on décompose s en deux morceaux, $s' = (s_0, \dots, s_m)$ et $s'' = (s_m, \dots, s_{2n})$. Le chemin \tilde{s} est alors construit en deux parties : la première partie $\tilde{s}' = (\tilde{s}_1, \dots, \tilde{s}_\ell)$, avec $\ell = 2n-m$, est identique à s'' (à translation près pour faire démarrer \tilde{s}' en 0) ; la deuxième partie $\tilde{s}'' = (\tilde{s}_\ell, \dots, \tilde{s}_{2n})$ est l'image de s' par une symétrie verticale (à translation près pour placer \tilde{s}'' « à la suite » de \tilde{s}'). Une illustration est donnée dans la Figure E-9.1.

On peut facilement vérifier que l'image de $s \in \mathcal{A}_{2n,k}$ par ψ est un élément de $\mathcal{B}_{2n,k}$: par construction, on a $\tilde{s}_i \geq 0$ pour tout i et $s_{2n} = 2k$. Pour voir qu'il s'agit d'une bijection, il suffit de montrer que ψ possède un inverse : il est décrit de la manière suivante.

Soit $\tilde{s} := \psi(s) = (\tilde{s}_0, \dots, \tilde{s}_{2n})$ un chemin de $\mathcal{B}_{2n,k}$ et soit $\ell = \max\{i \geq 0 : s_i = k\}$ le dernier instant où le chemin \tilde{s} passe par k avant d'atteindre $2k$. L'inverse de ψ est alors construit en associant à \tilde{s} le chemin $s = (s_0, \dots, s_{2n})$ obtenu de la façon suivante :

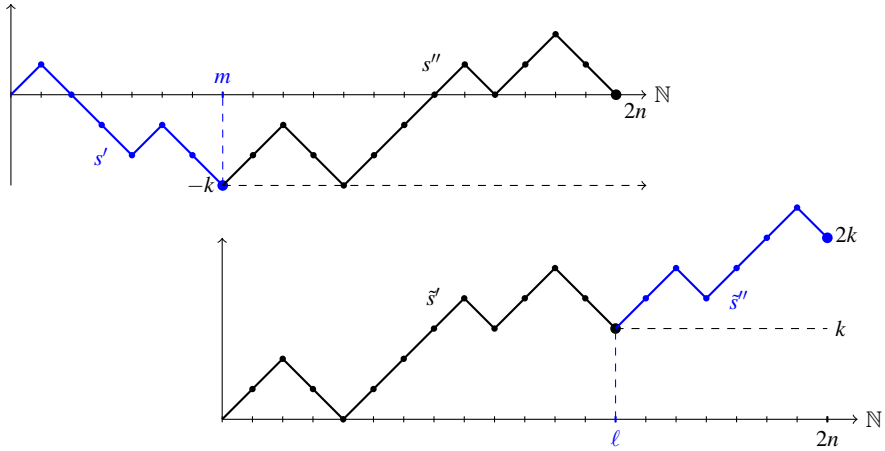


FIG. E-9.1 Illustration de la bijection ψ entre $\mathcal{A}_{2n,k}$ et $\mathcal{B}_{2n,k}$. Dans la figure du haut, on a représenté un chemin $s = (s_0, \dots, s_{2n}) \in \mathcal{A}_{2n,k}$, que l'on a décomposé en deux parties, $s' = (s_0, \dots, s_m)$ et $s'' = (s_m, \dots, s_{2n})$. Dans la figure du bas, on a représenté le chemin $\tilde{s} = \psi(s) \in \mathcal{B}_{2n,k}$, décomposé lui aussi en deux parties, $\tilde{s}' = (\tilde{s}_0, \dots, \tilde{s}_\ell)$ et $\tilde{s}'' = (\tilde{s}_\ell, \dots, \tilde{s}_{2n})$, avec $\ell = 2n - m$. La bijection ψ identifie \tilde{s}' à s'' et \tilde{s}'' au symétrique de s' par rapport à l'axe vertical (dans les deux cas, à translation près).

$$\begin{aligned} s_i &= \tilde{s}_{2n} - \tilde{s}_{2n-i} & \text{pour } i \in \{0, \dots, 2n - \ell\} & \quad (\text{noter que } s_{2n-\ell} = -k), \\ s_i &= -k + \tilde{s}_{i-2n-\ell} & \text{pour } i \in \{\ell + 1, \dots, 2n\} & \quad (\text{noter que } s_{2n} = 0). \end{aligned}$$

Autrement dit, \tilde{s} est composé de deux morceaux $\tilde{s}' = (s_0, \dots, s_\ell)$ et $\tilde{s}'' = (s_\ell, \dots, s_{2n})$. Le chemin s est construit en deux parties : la première partie s' est identique à \tilde{s}'' (à translation près pour faire démarrer s' en 0) ; la deuxième partie s'' est l'image de \tilde{s}' par une symétrie verticale (à translation près pour placer s'' « à la suite » de s'). On renvoie à la Figure E-9.1 pour une illustration. Notons qu'il est ici aussi facile de voir que le chemin s ainsi construit est un élément de $\mathcal{A}_{2n,k}$.

Pour conclure, on a montré que $\psi : \mathcal{A}_{2n,k} \rightarrow \mathcal{B}_{2n,k}$ est une bijection : on en déduit que $|\mathcal{A}_{2n,k}| = |\mathcal{B}_{2n,k}|$ ce qui termine la démonstration. \square

9.2 Marches aléatoires symétriques à densité : le théorème de Sparre Andersen

Considérons maintenant le cas où les variables aléatoires $(X_i)_{i \geq 0}$ sont i.i.d. de loi symétrique, c'est-à-dire que $-X_i$ a la même loi que X_i . Supposons de plus que les $(X_i)_{i \geq 1}$ sont de loi à densité ; le fait que la loi est symétrique s'exprime dans le fait que la densité f_{X_i} est une fonction paire, c'est-à-dire $f_{X_i}(-x) = f_{X_i}(x)$. On a alors le résultat suivant, démontré par Sparre Andersen [†] en 1954.

[†]. Avec une démonstration très différente de ce que nous donnons ci-dessous.

Théorème 9.4 (de Sparre Andersen). *Supposons que les variables aléatoires $(X_i)_{i \geq 1}$ sont i.i.d. de loi symétrique et à densité, et soit $S_k = \sum_{i=1}^k X_i$ la marche aléatoire associée. Alors, pour tout $n \geq 1$, on a*

$$P(S_1 \geq 0, \dots, S_n \geq 0) = P(S_1 > 0, \dots, S_n > 0) = \frac{1}{4^n} \binom{2n}{n}.$$

Observation 9.5. Ce résultat est assez surprenant par plusieurs aspects :

- Aucune hypothèse de moment n'est faite à propos des variables aléatoires $(X_i)_{i \geq 1}$; en particulier, on ne suppose pas que les X_i admettent une espérance ou une variance.
- Le résultat *ne dépend pas* de la loi des variables aléatoires $(X_i)_{i \geq 0}$, pourvu qu'elles soient symétrique et à densité ; on peut par exemple prendre $X_i \sim N(0, 1)$, $X_i \sim U(-1, 1)$, où plus généralement n'importe quel X_i de densité paire.
- La valeur de la probabilité de survie ressemble beaucoup à celle trouvée pour la marche aléatoire simple symétrique dans le Théorème 9.1.

□

Démonstration (du Théorème 9.4). Pour $n \in \mathbb{N}$, on introduit la variable aléatoire

$$K_n := \min \{k \in \{0, \dots, n\} : S_k = \min_{0 \leq j \leq n} S_j\},$$

qui est l'indice où la marche aléatoire atteint son minimum sur $\{0, \dots, n\}$ pour la première fois, voir la Figure E-9.2 pour une illustration.

On a $K_n = 0$ si et seulement si $S_j \geq 0$ pour tout $j \in \{1, \dots, n\}$, de sorte que

$$P(K_n = 0) = P(S_1 \geq 0, \dots, S_n \geq 0).$$

Aussi, $K_n = n$ si et seulement si $S_n < S_j$ pour tout $j \in \{0, \dots, n-1\}$, de sorte que

$$\begin{aligned} P(K_n = n) &= P(S_n < 0, S_n - S_1 < 0, \dots, S_n - S_{n-1} < 0) \\ &= P(X_n < 0, X_n + X_{n-1} < 0, \dots, X_n + \dots + X_1 < 0) \\ &= P(S_1 < 0, S_2 < 0, \dots, S_n < 0), \end{aligned}$$

où on a utilisé pour la dernière égalité que si on pose $\tilde{S}_j = X_n + \dots + X_{n-j+1}$ pour $j \in \{1, \dots, n\}$, alors $(\tilde{S}_j)_{1 \leq j \leq n}$ a la même loi que $(S_j)_{1 \leq j \leq n}$.

Pour $k \in \{1, \dots, n-1\}$, on a $K_n = k$ si et seulement si $S_j > S_k$ pour tout $j \in \{0, \dots, k-1\}$ et $S_j \geq S_k$ pour tout $j \in \{k, \dots, n\}$, voir la Figure E-9.2 pour une illustration. Ainsi, on peut écrire l'égalité d'événements suivante :

$$\begin{aligned} \{K_n = k\} &= \{S_k < 0, S_k - S_1 < 0, \dots, S_k - S_{k-1} < 0\} \cap \{S_{k+1} - S_k \geq 0, \dots, S_n - S_k \geq 0\} \\ &= \{X_k < 0, X_k + X_{k-1} < 0, \dots, X_k + \dots + X_1 < 0\} \cap \{X_{k+1} \geq 0, \dots, X_{k+1} + \dots + X_n \geq 0\}. \end{aligned}$$

En remarquant que le premier événement est généré par les variables aléatoires $(X_j)_{1 \leq j \leq k}$ et le deuxième par les variables aléatoires $(X_j)_{k+1 \leq j \leq n}$, on en déduit qu'ils sont indépendants, grâce à l'indépendance par paquets. Ainsi, la probabilité $P(K_n = k)$ est égale à

$$\begin{aligned} P(X_k < 0, X_k + X_{k-1} < 0, \dots, X_k + \dots + X_1 < 0) P(X_{k+1} \geq 0, \dots, X_{k+1} + \dots + X_n \geq 0) \\ = P(S_1 < 0, \dots, S_k < 0) P(S_1 \geq 0, \dots, S_{n-k} \geq 0), \end{aligned}$$

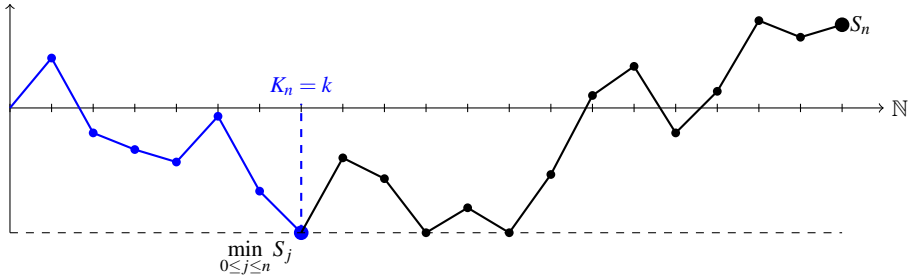


FIG. E-9.2 Illustration de la décomposition de l'événement $\{K_n = k\}$ comme une intersection des deux événements $\{S_k < 0, S_k - S_1 < 0, \dots, S_k - S_{k-1} < 0\}$ et $\{S_{k+1} - S_k \geq 0, \dots, S_n - S_k \geq 0\}$. Ces deux événements sont indépendants car le premier est généré par les variables aléatoires $(X_j)_{1 \leq j \leq k}$ et le deuxième par les variables aléatoires $(X_j)_{k+1 \leq j \leq n}$.

car $(X_k + \dots + X_{k-j+1})_{1 \leq j \leq k}$ a la même loi que $(S_j)_{1 \leq j \leq k}$ et $(X_{k+1} + \dots + X_{k+1+j})_{1 \leq j \leq k}$ a la même loi que $(S_j)_{1 \leq j \leq n-k}$.

Ainsi, en posant, pour $j \geq 1$,

$$p_\ell := P(S_1 \geq 0, \dots, S_\ell \geq 0) \quad \text{et} \quad q_\ell := P(S_1 < 0, \dots, S_\ell < 0)$$

(et par convention $p_0 = q_0 = 1$), on a montré que $P(K_n = k) = q_k p_{n-k}$, pour tout $k \in \{0, \dots, n\}$. En sommant sur $k \in \{0, \dots, n\}$, on obtient

$$\sum_{k=0}^n P(K_n = k) = \sum_{k=0}^n q_k p_{n-k} = 1. \quad (\text{E-9.3})$$

Maintenant, notons que l'on n'a pas (encore) utilisé les hypothèses du théorème : l'identité (E-9.3) est valable pour n'importe quelle marche aléatoire. L'hypothèse de symétrie de la marche aléatoire permet d'obtenir que $(-S_j)_{0 \leq j \leq k}$ a la même loi que $(S_j)_{0 \leq j \leq k}$, de sorte que

$$q_\ell = P(-S_1 > 0, \dots, -S_\ell > 0) = P(S_1 > 0, \dots, S_\ell > 0).$$

L'hypothèse que les variables aléatoires $(X_i)_{i \geq 1}$ possèdent une densité permet d'obtenir que S_ℓ possède une densité, et notamment $P(S_\ell = 0) = 0$, pour tout $\ell \geq 1$. Ainsi, on obtient que

$$p_\ell = P(S_1 \geq 0, \dots, S_\ell \geq 0) = P(S_1 > 0, \dots, S_\ell > 0) = q_\ell,$$

car $P(\exists j \in \{1, \dots, \ell\} S_j = 0) = 0$.

On déduit donc de (E-9.3) que, sous les hypothèses du théorème, pour tout $n \geq 0$

$$\sum_{k=0}^n p_k p_{n-k} = 1. \quad (\text{E-9.4})$$

Notons que cette relation, avec la donnée initial $p_0 = 1$, détermine complètement[†] la suite $(p_n)_{n \geq 0}$.

Remarquons maintenant que, d'après l'Observation 9.3, la suite $(u_{2n})_{n \geq 0}$ définie en (E-9.1) vérifie exactement la même relation (E-9.4), avec aussi $u_0 = 0$. Cela permet donc de conclure

[†]. Par récurrence : en effet, en isolant les termes $k = 0$ et $k = n$ dans la somme, on obtient $2p_n = 1 - \sum_{k=1}^{n-1} p_k p_{n-k}$ et la somme ne contient que des termes p_j avec $j \leq n-1$.

que, pour tout $n \geq 0$,

$$(q_n =) p_n = u_{2n} = \frac{1}{4^n} \binom{2n}{n},$$

comme annoncé. □

Observation 9.6. Si on connaît les séries entières, on peut conclure rapidement à partir de (E-9.4). En effet, on peut calculer la série génératrice suivante de deux manières : pour tout $|x| < 1$,

$$\sum_{n=0}^{\infty} \left(\sum_{k=0}^n p_k p_{n-k} \right) x^n = \sum_{n=0}^{\infty} x^n = \frac{1}{1-x} \quad \text{et} \quad \sum_{n=0}^{\infty} \left(\sum_{k=0}^n p_k p_{n-k} \right) x^n = \left(\sum_{m=0}^{\infty} p_m x^m \right)^2,$$

où pour la deuxième identité on a reconnu un produit de Cauchy de deux séries entières. On en déduit donc que

$$\forall |x| < 1, \quad \sum_{m=0}^{\infty} p_m x^m = \frac{1}{\sqrt{1-x}}.$$

Il reste ensuite à utiliser le développement en série entière de $\frac{1}{\sqrt{1-x}}$ pour obtenir que

$$p_m = \frac{1}{m!} \frac{1}{2} \frac{3}{2} \cdots \frac{2m-1}{2} = \frac{1}{m!} \frac{1}{2^m} \frac{(2m)!}{2 \cdot 4 \cdots 2m} = \frac{1}{4^m} \frac{(2m)!}{(m!)^2},$$

ce qui donne la formule voulue. □

Ex. 10

Comportement asymptotique d'un processus de renouvellement

Soient $(X_i)_{i \geq 1}$ des variables aléatoires indépendantes et de même loi, à valeur dans \mathbb{R}_+ , définies sur le même espace probabilisé (Ω, \mathcal{F}, P) ; on suppose que $P(X_i = 0) \neq 1$ pour éviter des cas dégénérés. On interprète X_i comme des durées de vie de composants d'une chaîne de production : lorsqu'un composant tombe en panne il est immédiatement remplacé par un nouveau composant. Ainsi, pour $k \in \mathbb{N}$, l'instant où le k -ème composant tombe en panne est

$$T_k := X_1 + \cdots + X_k,$$

avec par convention $T_0 = 0$.

On s'intéresse au nombre N_t de composants que l'on aura remplacé à l'instant t , que l'on peut exprimer en fonction de $(T_k)_{k \geq 1}$ de la façon suivante : pour $t \in \mathbb{R}_+$,

$$N_t := \max\{k \in \mathbb{N}, T_k \leq t\}. \quad (\text{E-10.1})$$

On appelle $(N_t)_{t \in \mathbb{R}_+}$ *processus de renouvellement* et on cherche à obtenir des informations sur le comportement en temps long ($t \rightarrow \infty$) de ce processus.

Observation 10.1. Si les variables aléatoires $(X_i)_{i \geq 1}$ sont i.i.d. de loi exponentielle, alors $(N_t)_{t \geq 0}$ est le processus de Poisson étudié dans les Sections 6.6.3 et 6.6.4. \square

10.1 Loi des grands nombres pour le processus de renouvellement

Le premier résultat que l'on obtient est une loi (faible) des grands nombres pour le processus de renouvellement $(N_t)_{t \in \mathbb{R}_+}$.

Théorème 10.2 (Loi faible des grands nombres). Soient $(X_i)_{i \geq 1}$ des variables aléatoires i.i.d., à valeur dans \mathbb{R}_+ , définies sur le même espace probabilisé (Ω, \mathcal{F}, P) . Alors le processus de renouvellement associé $(N_t)_{t \in \mathbb{R}_+}$ défini en (E-10.1) vérifie une loi faible des grands nombres :

$$\forall \varepsilon > 0 : \quad \lim_{t \rightarrow +\infty} P\left(\left|\frac{1}{t}N_t - \frac{1}{E(X_1)}\right| > \varepsilon\right),$$

où par convention $\frac{1}{E(X_1)} = 0$ si $E(X_1) = +\infty$.

Démonstration. Commençons par le cas où $E(X_1) < +\infty$ et notons $\mu := E(X_1)$. On suppose aussi que $\mu > 0$ car sinon on aurait $P(X_1 = 0) = 1$, auquel cas $T_k = 0$ pour tout k et $N_t = +\infty$ pour tout $t > 0$.

Fixons $\varepsilon > 0$ tel que $\varepsilon < 1/\mu$. Comme

$$\left\{\left|\frac{1}{t}N_t - \frac{1}{\mu}\right| > \varepsilon\right\} = \left\{\frac{1}{t}N_t - \frac{1}{\mu} > \varepsilon\right\} \cup \left\{\frac{1}{t}N_t - \frac{1}{\mu} < -\varepsilon\right\},$$

on obtient par sous-additivité

$$P\left(\left|\frac{1}{t}N_t - \frac{1}{\mu}\right| > \varepsilon\right) \leq P\left(\frac{1}{t}N_t > \frac{1}{\mu} + \varepsilon\right) + P\left(\frac{1}{t}N_t < \frac{1}{\mu} - \varepsilon\right), \quad (\text{E-10.2})$$

et on doit donc montrer que les deux probabilités dans le membre de droite tendent vers 0.

Une observation importante est que l'on a l'égalité d'événements $\{N_t \geq k\} = \{T_k \leq t\}$: en effet, $N_t \leq k$ si et seulement si on a dû remplacer les k premiers composants avant l'instant t , c'est-à-dire si et seulement si le k -ème composant est tombé en panne avant l'instant t . Ainsi,

$$P\left(\frac{1}{t}N_t > \frac{1}{\mu} + \varepsilon\right) = P\left(N_t > \frac{t}{\mu} + t\varepsilon\right) = P(N_t \geq k_t^+) = P(T_{k_t^+} \leq t),$$

où on a posé

$$k_t^+ := \left\lfloor \frac{t}{\mu} + t\varepsilon \right\rfloor + 1.$$

Maintenant, on va appliquer la loi des grands nombres (Théorème 7.3) à $T_k = X_1 + \dots + X_k$. On a en effet, pour $t \geq 1$,

$$P(T_{k_t^+} \leq t) = P\left(\frac{1}{k_t^+}T_{k_t^+} - \mu \leq \frac{t}{k_t^+} - \mu\right) \leq P\left(\frac{1}{k_t^+}T_{k_t^+} - \mu \leq -\tilde{\varepsilon}\right),$$

où pour la dernière inégalité, on a utilisé le fait que $k_t^+ \geq \frac{t}{\mu} + t\varepsilon$ et $k_t^+ \leq \frac{t}{\mu} + t\varepsilon + 1 \leq (\frac{1}{\mu} + \varepsilon + 1)t$ pour $t \geq 1$, de sorte que

$$\frac{t}{k_t^+} - \mu = \frac{t - \mu k_t^+}{k_t^+} \leq -\frac{\mu t \varepsilon}{k_t^+} \leq -\frac{\mu \varepsilon}{\frac{1}{\mu} + \varepsilon + 1} =: -\tilde{\varepsilon}.$$

Comme $\tilde{\varepsilon} > 0$ est fixé et que $\lim_{t \rightarrow \infty} k_t^+ = +\infty$, la loi des grands nombres L^1 (Théorème 7.3) montre que $\lim_{t \rightarrow \infty} P(\frac{1}{k_t^+}T_{k_t^+} - \mu \leq -\tilde{\varepsilon}) = 0$, dont on déduit que le premier terme dans le membre de droite de (E-10.2) tend vers 0 quand $t \rightarrow \infty$.

Pour le deuxième terme dans (E-10.2), on utilise la même méthode : on écrit

$$P\left(\frac{1}{t}N_t < \frac{1}{\mu} - \varepsilon\right) = P\left(N_t < \frac{t}{\mu} - t\varepsilon\right) \leq P(N_t < k_t^-) = P(T_{k_t^-} > t),$$

où on a posé

$$k_t^- := \left\lfloor \frac{t}{\mu} - t\varepsilon \right\rfloor + 1$$

et utilisé l'égalité d'événements $\{N_t < k\} = \{T_k > t\}$. On a alors, pour $t \geq 2/\varepsilon$,

$$P(T_{k_t^-} > t) \leq P\left(\frac{1}{k_t^-}T_{k_t^-} - \mu > \frac{t}{k_t^-} - \mu\right) \leq P\left(\frac{1}{k_t^-}T_{k_t^-} - \mu > \tilde{\varepsilon}'\right),$$

où cette fois on a utilisé que $k_t^- \leq \frac{t}{\mu} - t\varepsilon + 1 \leq \frac{t}{\mu} - \frac{1}{2}t\varepsilon$ (d'où aussi $k_t^- \leq \frac{t}{\mu}$) pour $t \geq 2/\varepsilon$, de sorte que

$$\frac{t}{k_t^-} - \mu = \frac{t - \mu k_t^-}{k_t^-} \geq \frac{\mu t \varepsilon}{2k_t^-} \geq \frac{\mu^2 \varepsilon}{2} =: \tilde{\varepsilon}'.$$

Comme précédemment, $\lim_{t \rightarrow \infty} k_t^- = +\infty$, et la loi des grands nombres L^1 (Théorème 7.3) montre que l'on a $\lim_{t \rightarrow \infty} P(\frac{1}{k_t^-}T_{k_t^-} - \mu > \tilde{\varepsilon}') = 0$, dont on déduit que le deuxième terme dans le membre de droite de (E-10.2) tend vers 0 quand $t \rightarrow \infty$. Cela conclut la démonstration dans le cas $E(X_1) < +\infty$.

Il reste à traiter le cas où $E(X_1) = +\infty$. Soit $\varepsilon > 0$, et estimons

$$P\left(\left|\frac{1}{t}N_t\right| > \varepsilon\right) = P(N_t > \varepsilon t) = P(N_t \geq k_t^+) = P(T_{k_t^+} \leq t)$$

où ici $k_t^+ := \lfloor \varepsilon t \rfloor + 1$. On peut alors écrire, pour $t \geq \frac{1}{\varepsilon}$,

$$P(T_{k_t^+} \leq t) = P\left(\frac{1}{k_t^+}T_{k_t^+} \leq \frac{t}{k_t^+}\right) \leq P\left(\frac{1}{k_t^+}T_{k_t^+} \leq \frac{1}{2\varepsilon}\right)$$

où on a utilisé le fait que $k_t^+ \leq \varepsilon t + 1 \leq 2\varepsilon t$. Maintenant, on peut appliquer le résultat de l'Exercice 7.3 à $T_k = X_1 + \dots + X_k$: comme $\lim_{t \rightarrow \infty} k_t^+ = +\infty$, pour $A = (2\varepsilon)^{-1} > 0$ fixé (arbitrairement grand), on obtient que

$$\lim_{t \rightarrow \infty} P\left(\frac{1}{k_t^+}T_{k_t^+} \leq \frac{1}{2\varepsilon}\right) = 0.$$

Cela montre que $\lim_{t \rightarrow \infty} P(T_{k_t^+} \leq t) = 0$, ce qui conclut la démonstration dans le cas $E(X_1) = +\infty$. \square

10.2 Théorème central limite pour le processus de renouvellement

Le deuxième résultat que l'on obtient est un théorème central limite pour le processus $(N_t)_{t \in \mathbb{R}_+}$, sous l'hypothèse un peu plus fortes que les $(X_i)_{i \geq 1}$ admettent un moment d'ordre deux fini. De la même manière que pour le théorème central limite (Théorème 7.17), il s'agit de déterminer à quelle vitesse la quantité $\frac{1}{t}N_t - \frac{1}{\mu}$ tend vers 0. Comme dans (7.20), il faut multiplier la quantité $\frac{1}{t}N_t - \frac{1}{\mu}$ par une suite $c_t \approx \sqrt{t}$, pour obtenir une quantité « qui ne tend ni vers 0 ni vers $+\infty$ » et est une variable aléatoire normale.

Théorème 10.3 (Théorème central limite). Soient $(X_i)_{i \geq 1}$ des variables aléatoires i.i.d., à valeur dans \mathbb{R}_+ , définies sur le même espace probabilisé (Ω, \mathcal{F}, P) et d'espérance et de variance finie

$$\mu := E(X_1) > 0, \quad \sigma^2 := \text{Var}(X_1).$$

Alors le processus de renouvellement associé $(N_t)_{t \in \mathbb{R}_+}$ défini en (E-10.1) vérifie un théorème central limite : quand $t \rightarrow \infty$,

$$Z_t := \frac{\mu^{3/2}\sqrt{t}}{\sigma} \left(\frac{1}{t}N_t - \frac{1}{\mu} \right) \xrightarrow{\mathcal{L}} N(0, 1).$$

Observation 10.4. On a utilisé ci-dessus la notation $Z_t \xrightarrow{\mathcal{L}} N(0, 1)$ de la Section 8.2.3, qui signifie :

$$\lim_{t \rightarrow \infty} P(Z_t \leq x) = P(Z \leq x) \quad \forall x \in \mathbb{R},$$

où Z est une variable aléatoire $N(0, 1)$. \square

Démonstration. On va encore utiliser l'égalité d'événements $\{N_t \geq k\} = \{T_k \leq t\}$, déjà cruciale pour la loi des grands nombres, pour se ramener à une application du théorème central limite (Théorème 7.17) pour $T_k = X_1 + \dots + X_k$. Fixons $x \in \mathbb{R}$ et considérons la probabilité

$$\begin{aligned} P(Z_t > x) &= P\left(\frac{\mu^{3/2}\sqrt{t}}{\sigma}\left(\frac{1}{t}N_t - \frac{1}{\mu}\right) > x\right) = P\left(\frac{1}{t}N_t - \frac{1}{\mu} > \frac{\sigma x}{\mu^{3/2}\sqrt{t}}\right) \\ &= P\left(N_t > \frac{t}{\mu} + \frac{\sigma x}{\mu^{3/2}}\sqrt{t}\right). \end{aligned}$$

Ainsi, en posant

$$k_t := k_t(x) = \left\lfloor \frac{t}{\mu} + \frac{\sigma x}{\mu^{3/2}}\sqrt{t} \right\rfloor + 1$$

(noter que $\lim_{t \rightarrow \infty} k_t = +\infty$), on obtient

$$P(Z_t > x) = P(N_t \geq k_t) = P(T_{k_t} \leq t).$$

On peut alors réécrire l'identité ci-dessus pour se mettre en position d'appliquer le théorème central limite : on a

$$P(Z_t > x) = P\left(\frac{\sqrt{k_t}}{\sigma}\left(\frac{1}{k_t}T_{k_t} - \mu\right) \leq x_t\right), \quad \text{où} \quad x_t := \frac{\sqrt{k_t}}{\sigma}\left(\frac{t}{k_t} - \mu\right) = \frac{t - k_t\mu}{\sigma\sqrt{k_t}}.$$

Remarquons ici que l'on a $\mu k_t \geq t + \frac{\sigma x}{\sqrt{\mu}}\sqrt{t}$ et $\mu k_t \leq t + \frac{\sigma x}{\sqrt{\mu}}\sqrt{t} + 2\mu$: on en déduit que

$$-\frac{x\sqrt{t}}{\sqrt{\mu}\sqrt{k_t}} - \frac{2\mu}{\sigma\sqrt{k_t}} \leq x_t \leq -\frac{x\sqrt{t}}{\sqrt{\mu}\sqrt{k_t}},$$

et comme $k_t \sim \frac{t}{\mu}$ quand $t \rightarrow \infty$, on en conclut que $\lim_{t \rightarrow \infty} x_t = -x$. Ainsi, en utilisant le théorème central limite (Théorème 7.17), ou plus précisément la convergence (7.31) (que l'on peut déduire de la Proposition 7.22), on obtient (rappelons que $\lim_{t \rightarrow \infty} k_t = +\infty$)

$$\lim_{t \rightarrow \infty} P(Z_t > x) = \lim_{t \rightarrow \infty} P\left(\frac{\sqrt{k_t}}{\sigma}\left(\frac{1}{k_t}T_{k_t} - \mu\right) \leq x_t\right) = P(Z \leq -x),$$

où $Z \sim N(0, 1)$. Comme $P(Z \leq -x) = P(Z > x)$, on a donc montré que $\lim_{t \rightarrow \infty} P(Z_t > x) = P(Z > x)$ pour tout $x \in \mathbb{R}$, ce qui conclut la démonstration en passant au complémentaire. \square

Ex. 11

Urne de Pólya et loi Beta

Une urne contient initialement a boules rouges et b boules vertes. Lors du n -ème tour, on tire une boule dans l'urne : on la replace dans l'urne en ajoutant une boule de la même couleur. Peut-être de manière plus précise, si après le $(n-1)$ -ème tour l'urne contient r boules rouges et v boules vertes, alors au n -ème tour : avec probabilité $r/(r+v)$ on tirera une boule rouge et on ajoutera une boule rouge dans l'urne (qui contiendra $r+1$ boules rouges et v boules vertes après le n -ème tour) ; avec probabilité $v/(r+v)$ on tirera une boule verte et on ajoutera une boule verte dans l'urne (qui contiendra r boules rouges et $v+1$ boules vertes après le n -ème tour).

Soit $(X_i)_{i \geq 1}$ la suite de variables aléatoires de Bernoulli définie de la manière suivante : on pose $X_i = 1$ si la boule tirée lors du i -ème tour est rouge et $X_i = 0$ sinon. On s'intéresse au nombre de boules rouges dans l'urne après n tours, que l'on note

$$R_n := a + \sum_{i=1}^n X_i.$$

Le nombre total de boules dans l'urne étant $a+b+n$, on peut aussi étudier la proportion de boules rouges dans l'urne, donnée par $\frac{1}{a+b+n}R_n \in [0, 1]$. On montre le résultat suivant.

Proposition 11.1. *Pour tout $t \in [0, 1]$, on a*

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\frac{1}{a+b+n} R_n \leq t \right) = \mathbb{P}(W \leq t),$$

où W est une variable aléatoire de loi Beta(a, b), introduite dans la Section 6.6.2, voir (6.92).

Observation 11.2. On peut aussi utiliser la notation $\frac{1}{a+b+n}R_n \xrightarrow{\mathcal{L}} \text{Beta}(a, b)$ pour désigner la convergence de la Proposition 11.1, avec des notations analogues à celle de la Section 8.2.3. \square

Démonstration. Commençons par observer que pour tout $n \geq 1$, on peut calculer explicitement la loi de (X_1, \dots, X_n) . Pour tout $n \geq 1$ et tout $(\varepsilon_1, \dots, \varepsilon_n) \in \{0, 1\}^n$, montrons que

$$\mathbb{P} \left((X_1, \dots, X_n) = (\varepsilon_1, \dots, \varepsilon_n) \right) = \frac{\prod_{i=0}^{k-1} (a+i) \prod_{i=0}^{n-k-1} (b+i)}{\prod_{i=0}^{n-1} (a+b+i)}, \quad \text{où } k = \sum_{i=1}^n \varepsilon_i. \quad (\text{E-11.1})$$

On procède par récurrence sur n . On remarque que, en posant $k = \sum_{i=1}^n \varepsilon_i$,

$$\mathbb{P} (X_{n+1} = \varepsilon_{n+1} \mid (X_1, \dots, X_n) = (\varepsilon_1, \dots, \varepsilon_n)) = \begin{cases} \frac{a+k}{a+b+n} & \text{si } \varepsilon_{n+1} = 1 \\ \frac{b+n-k}{a+b+n} & \text{si } \varepsilon_{n+1} = 0 \end{cases}$$

car si l'on sait que $(X_1, \dots, X_n) = (\varepsilon_1, \dots, \varepsilon_n)$, cela signifie qu'il y a $a+k$ boules rouges dans l'urne et $b+n-k$ boules vertes, où $k = \sum_{i=1}^n \varepsilon_i$ est le nombre de fois où l'on a tiré une boule rouge. En posant $k' = \sum_{i=1}^{n+1} \varepsilon_i$, on en déduit par hypothèse de récurrence que

$$P((X_1, \dots, X_{n+1}) = (\varepsilon_1, \dots, \varepsilon_{n+1})) = \frac{\prod_{i=0}^{k-1} (a+i) \prod_{i=0}^{n-k-1} (b+i)}{\prod_{i=0}^{n-1} (a+b+i)} \times \begin{cases} \frac{a+k}{a+b+n} & \text{si } k' = k+1, \\ \frac{b+n-k}{a+b+n} & \text{si } k' = k, \end{cases}$$

c'est-à-dire

$$P((X_1, \dots, X_{n+1}) = (\varepsilon_1, \dots, \varepsilon_{n+1})) = \frac{\prod_{i=0}^{k'-1} (a+i) \prod_{i=0}^{n+1-k'-1} (b+i)}{\prod_{i=0}^n (a+b+i)}.$$

Cela conclut la récurrence et démontre (E-11.1).

Notons que l'on peut réécrire la formule (E-11.1) de la façon suivante :

$$P((X_1, \dots, X_n) = (\varepsilon_1, \dots, \varepsilon_n)) = \frac{(a+b-1)!}{(a-1)!(b-1)!} \frac{(a+k-1)!(b+n-k-1)!}{(a+b+n-1)!}.$$

Comme on a $\Gamma(u) = (u-1)!$ pour $u \in \mathbb{N}^*$, en utilisant la fonction $\beta(u, v) = \frac{\Gamma(u)\Gamma(v)}{\Gamma(u+v)}$ définie dans la Section 8.2.3 (voir (6.42)-(6.43)), on peut le réécrire de manière plus concise sous la forme

$$P((X_1, \dots, X_n) = (\varepsilon_1, \dots, \varepsilon_n)) = \frac{\beta(a+k, b+n-k)}{\beta(a, b)}, \quad \text{où } k = \sum_{i=1}^n \varepsilon_i.$$

Comme il y a exactement $\binom{n}{k}$ n -uplets $(\varepsilon_1, \dots, \varepsilon_n)$ tels que $\sum_{i=1}^n \varepsilon_i = k$, on en déduit que, pour $k \in \{0, \dots, n\}$,

$$P\left(\sum_{i=1}^n X_i = k\right) = \binom{n}{k} \frac{\beta(a+k, b+n-k)}{\beta(a, b)} = \frac{1}{\beta(a, b)} \binom{n}{k} \int_0^1 x^{a+k} (1-x)^{b+n-k} dx, \quad (\text{E-11.2})$$

où on a utilisé le fait que $\beta(u, v) = \int_0^1 x^u (1-x)^v dx$, voir (6.42).

Observation 11.3. On pourrait essayer d'obtenir la formule (E-11.2) de $\sum_{i=1}^n X_i$, par exemple par récurrence sur n , mais c'est en réalité plus compliqué que d'obtenir directement (E-11.1), voir l'Observation 11.4 plus bas. \square

On en déduit que

$$P(R_n = a+k) = P\left(\sum_{i=1}^n X_i = k\right) = \int_0^1 \binom{n}{k} x^k (1-x)^{n-k} f_{a,b}(x) dx,$$

avec $f_{a,b}(x) = \frac{1}{\beta(a,b)} x^a (1-x)^b$ pour $x \in [0, 1]$ la densité d'une loi Beta(a, b), voir (6.92). On peut réinterpréter cette identité de la manière suivante :

$$P(R_n = a+k) = \int_0^1 P(Y_{n,x} = k) f_{a,b}(x) dx$$

où $Y_{n,x}$ est une variable aléatoire de loi Binomiale de paramètre n, x .

Pour $t \in [0, 1]$, posons $k_n := \lfloor (a+b+n)t \rfloor - a$. Alors on a $\lim_{n \rightarrow \infty} \frac{1}{n}k_n = t$ et on peut écrire

$$\begin{aligned} \mathbb{P}\left(\frac{1}{a+b+n}R_n \leq t\right) &= \sum_{k=0}^{k_n} \mathbb{P}(R_n = a+k) = \int_0^1 \sum_{k=0}^{k_n} \mathbb{P}(Y_{n,x} = k) f_{a,b}(x) dx \\ &= \int_0^1 \mathbb{P}(Y_{n,x} \leq k_n) f_{a,b}(x) dx \end{aligned} \quad (\text{E-11.3})$$

où on a utilisé la linéarité de l'espérance (pour une somme finie) dans la deuxième égalité.

Donnons une idée de comment la démonstration se conclut ; on donnera la preuve complète plus bas. Comme $\lim_{n \rightarrow \infty} \frac{1}{n}k_n = t$, la loi faible des grands nombres permet de montrer que [†]

$$\lim_{n \rightarrow \infty} \mathbb{P}\left(\frac{1}{n}Y_{n,x} \leq \frac{1}{n}k_n\right) = \begin{cases} 0 & \text{si } x < t, \\ 1 & \text{si } x > t. \end{cases}$$

Si l'on pouvait échanger la limite et l'intégrale dans l'identité (E-11.3) ci-dessus, on aurait

$$\lim_{n \rightarrow \infty} \mathbb{P}\left(\frac{1}{a+b+n}R_n \leq t\right) = \int_0^t f_{a,b}(x) dx = \mathbb{P}(Z \leq t).$$

Montrons que cette stratégie fonctionne en montrant notamment l'inversion de limite et d'intégrale à la main [‡]. Fixons un $\varepsilon > 0$ et raisonnons par borne supérieure et inférieure. Fixons ensuite $\delta > 0$ (qui dépend de ε) tel que $\int_{t-\delta}^{t+\delta} f_{a,b}(x) dx \leq \varepsilon$.

Pour la borne supérieure, on écrit

$$\mathbb{P}\left(\frac{1}{a+b+n}R_n \leq t\right) \leq \int_0^{t+\delta} f_{a,b}(x) dx + \int_{t+\delta}^1 \mathbb{P}(Y_{n,x} \leq k_n) f_{a,b}(x) dx.$$

Maintenant, la loi faible des grands nombres, plus précisément (7.6), permet de montrer que si $x \geq t + \delta$, pour n suffisamment grand de sorte que $\frac{1}{n}k_n - x \leq -\delta/2$, on a

$$\begin{aligned} \mathbb{P}\left(\frac{1}{n}Y_{n,x} \leq \frac{1}{n}k_n\right) &\leq \mathbb{P}\left(\frac{1}{n}Y_{n,x} - x \leq -\frac{1}{2}\delta\right) \\ &\leq \mathbb{P}\left(\left|\frac{1}{n}Y_{n,x} - x\right| \geq \frac{1}{2}\delta\right) \leq \frac{x(1-x)}{n\frac{1}{4}\delta^2} \leq \frac{1}{n\delta^2}, \end{aligned}$$

en utilisant aussi le fait que $x(1-x) \leq \frac{1}{4}$ pour tout $x \in [0, 1]$. On en déduit que pour n suffisamment grand

$$\int_{t+\delta}^1 \mathbb{P}(Y_{n,x} \leq k_n) f_{a,b}(x) dx \leq \frac{1}{n\delta^2} \int_{t+\delta}^1 f_{a,b}(x) dx \leq \frac{1}{n\delta^2}.$$

On en conclut que

$$\mathbb{P}\left(\frac{1}{a+b+n}R_n \leq t\right) \leq \int_0^t f_{a,b}(x) dx + \varepsilon + \frac{1}{n\delta^2},$$

où on rappelle que δ est tel que $\int_{t-\delta}^{t+\delta} f_{a,b}(x) dx \leq \varepsilon$.

[†]. On le laisse en exercice.

[‡]. Si l'on connaît le théorème de convergence dominée, on peut conclure directement.

Pour la borne inférieure, on écrit

$$\begin{aligned} \mathbf{P}\left(\frac{1}{a+b+n}R_n \leq t\right) &\geq \int_0^{t-\varepsilon} \mathbf{P}(Y_{n,x} \leq k_n) f_{a,b}(x) dx \\ &= \int_0^{t-\delta} f_{a,b}(x) dx - \int_0^{t-\delta} \mathbf{P}(Y_{n,x} > k_n) f_{a,b}(x) dx. \end{aligned}$$

Comme précédemment, la loi faible des grands nombres, plus précisément (7.6), permet de montrer que si $x < t - \delta$, alors pour n suffisamment grand de sorte que $\frac{1}{n}k_n - x \geq \delta/2$, on a

$$\mathbf{P}\left(\frac{1}{n}Y_{n,x} > \frac{1}{n}k_n\right) = \mathbf{P}\left(\frac{1}{n}Y_{n,x} - x > \frac{1}{n}k_n - x\right) \leq \mathbf{P}\left(\left|\frac{1}{n}Y_{n,x} - x\right| \geq \frac{1}{2}\delta\right) \leq \frac{1}{n\delta^2}.$$

On en conclut comme plus haut que

$$\mathbf{P}\left(\frac{1}{a+b+n}R_n \leq t\right) \geq \int_0^t f_{a,b}(x) dx - \varepsilon - \frac{1}{n\delta^2},$$

où on a encore utilisé le fait que $\int_{t-\delta}^{t+\delta} f_{a,b}(x) dx \leq \varepsilon$.

On a donc montré que pour n suffisamment grand (notamment tel que $\frac{1}{n\delta^2} \leq \varepsilon$), on a

$$\left| \mathbf{P}\left(\frac{1}{a+b+n}R_n \leq t\right) - \int_0^t f_{a,b}(x) dx \right| \leq 2\varepsilon,$$

ce qui conclut la démonstration. \square

Commentaire sur l'échangeabilité et le théorème de Finetti ?

Observation 11.4. On pourrait aussi vouloir obtenir directement la loi de $S_n := \sum_{i=1}^n X_i$ par récurrence (mais en fait c'est plus compliqué que (E-11.1), même si je m'y prends peut-être mal), et montrer la formule suivante :

$$\mathbf{P}(S_n = k) = \binom{n}{k} \times \frac{\prod_{i=0}^{k-1} (a+i) \prod_{i=0}^{n-k-1} (b+i)}{\prod_{i=0}^{n-1} (a+b+i)} \quad \text{pour tout } k \in \{0, \dots, n\}, \quad (\text{E-11.4})$$

où par convention $\prod_{i=0}^{k-1} (a+i) = 1$ si $k = 0$ et $\prod_{i=0}^{n-k-1} (b+i) = 1$ si $k = n$. Mais en fait c'est plus compliqué que la formule (E-11.1)...

Noter que la formule (E-11.4) peut aussi se réécrire en

$$\mathbf{P}(S_n = k) = \frac{(a+b-1)!}{(a-1)!(b-1)!} \frac{(a+k-1)!(b+n-k-1)!}{(a+b+n-1)!} \frac{n!}{k!(n-k)!} = \frac{\binom{k+a-1}{k} \binom{n-k+b-1}{n-k}}{\binom{n+a+b-1}{n}}.$$

Pour $n = 1$, on a en effet $S_1 = X_1$, et $\mathbf{P}(X_1 = 1) = \frac{a}{a+b}$, $\mathbf{P}(X_1 = 0) = \frac{b}{a+b}$, car il y a $a+b$ boules dans l'urne, dont a rouges et b vertes. Cela correspond à la formule annoncée.

Ensuite, pour tout $k \in \{0, \dots, n\}$, conditionnellement à $S_n = k$ il y a $a+k$ boules rouges et $b+n-k$ boules vertes dans l'urne, d'où

$$\mathbf{P}(X_{n+1} = 1 \mid S_n = k) = \frac{a+k}{a+b+n}, \quad \mathbf{P}(X_{n+1} = 0 \mid S_n = k) = \frac{b+n-k}{a+b+n}.$$

Ainsi, comme $S_{n+1} = S_n + X_{n+1}$, on obtient :

- Pour $k = 0$, comme $\binom{n}{0} = 1$ pour tout n , on a

$$P(S_{n+1} = 0) = P(S_n = 0) P(X_{n+1} = 0 \mid S_n = 0) = \frac{\prod_{i=0}^{n-1} (b+i)}{\prod_{i=0}^{n-1} (a+b+i)} \frac{b+n}{a+b+n},$$

qui donne la formule voulue (E-11.4) au rang $n+1$ pour $k = 0$.

- De même, pour $k = n+1$, comme $\binom{n}{n} = 1$ pour tout n , on a

$$P(S_{n+1} = n+1) = P(S_n = n) P(X_{n+1} = 1 \mid S_n = n) = \frac{\prod_{i=0}^{n-1} (a+i)}{\prod_{i=0}^{n-1} (a+b+i)} \frac{a+n}{a+b+n},$$

qui donne la formule voulue (E-11.4) au rang $n+1$ pour $k = n+1$.

- Pour $k \in \{1, \dots, n\}$, par la formule des probabilités totales, on a

$$P(S_{n+1} = k) = P(S_n = k-1) P(X_{n+1} = 1 \mid S_n = k-1) + P(S_n = k) P(X_{n+1} = 1 \mid S_n = k-1),$$

et par hypothèse de récurrence

$$\begin{aligned} P(S_n = k-1) P(X_{n+1} = 1 \mid S_n = k-1) &= \binom{n}{k-1} \times \frac{\prod_{i=0}^{k-2} (a+i) \prod_{i=0}^{n-k} (b+i)}{\prod_{i=0}^{n-1} (a+b+i)} \times \frac{a+k-1}{a+b+n} \\ &= \binom{n}{k-1} \times \frac{\prod_{i=0}^{k-1} (a+i) \prod_{i=0}^{n+1-k-1} (b+i)}{\prod_{i=0}^{n+1-1} (a+b+i)}, \end{aligned}$$

et similairement

$$\begin{aligned} P(S_n = k) P(X_{n+1} = 0 \mid S_n = k) &= \binom{n}{k} \times \frac{\prod_{i=0}^{k-1} (a+i) \prod_{i=0}^{n-k-1} (b+i)}{\prod_{i=0}^{n-1} (a+b+i)} \times \frac{b+n-k}{a+b+n} \\ &= \binom{n}{k} \times \frac{\prod_{i=0}^{k-1} (a+i) \prod_{i=0}^{n+1-k-1} (b+i)}{\prod_{i=0}^{n+1-1} (a+b+i)}. \end{aligned}$$

Au final, en sommant les deux dernières expressions, et comme $\binom{n}{k-1} + \binom{n}{k} = \binom{n+1}{k}$, on obtient la formule (E-11.4) au rang $n+1$.

Ex. 12

Records successifs d'une suite de variables aléatoires i.i.d.*

Soient $(X_i)_{i \geq 0}$ des variables aléatoires indépendantes et de même loi, que l'on suppose à densité, de densité commune notée f et de fonction de répartition notée F . Pour $n \geq 1$, on dit qu'un *record* a lieu au n -ème tour si $X_n > X_i$ pour tout $i \in \{0, \dots, n-1\}$, et on note A_n cet événement :

$$A_n := \bigcap_{i=0}^n \{X_n > X_i\}.$$

Lemme 12.1. On a $P(A_n) = \frac{1}{n+1}$ et pour tous $1 \leq k < n$, $P(A_k \cap A_n) = \frac{1}{k+1} \frac{1}{n+1}$. En particulier, $\text{Cov}(\mathbb{1}_{A_k}, \mathbb{1}_{A_n}) = 0$.

Démonstration. Pour calculer $P(A_n)$, on peut utiliser (la méthode de) l'Exercice 6.59, en remarquant que dans les notations de l'exercice, $P(A_n) = P(A_n^{(n)})$, sachant que presque sûrement on a $X_n \neq X_i$ pour tous $0 \leq i \leq n-1$ (voir l'Exercice 6.59). Utilisons ici l'approche directe de l'Exercice 6.60, qui se généralise bien au calcul de $P(A_k \cap A_n)$.

Calcul de $P(A_n)$. On peut réécrire l'événement A_n en fonction des $(X_i)_{0 \leq i \leq n}$: on a $A_n = \{X_0 < X_n, X_1 < X_n, \dots, X_{n-1} < X_n\}$. Ainsi, en écrivant $P(A_n) = E(\mathbb{1}_{A_n})$ et en appliquant la formule de transfert, on a

$$\begin{aligned} P(A_n) &= \int_{\mathbb{R}^{n+1}} f(x_0) \cdots f(x_{n-1}) f(x_n) \mathbb{1}_{\{x_0 < x_n, x_1 < x_n, \dots, x_{n-1} < x_n\}} dx_0 \cdots dx_n \\ &= \int_{-\infty}^{+\infty} f(x_n) \left(\int_{-\infty}^{x_n} \cdots \int_{-\infty}^{x_n} f(x_0) \cdots f(x_{n-1}) dx_0 \cdots dx_{n-1} \right) dx_n, \end{aligned}$$

où on a utilisé le théorème de Fubini–Tonelli pour la deuxième égalité. Maintenant, comme $\int_{-\infty}^x f(t) dt = F(x)$ et que $F'(x) = f(x)$, on obtient

$$P(A_n) = \int_{-\infty}^{+\infty} f(x_n) F(x_n)^n dx_n = \left[\frac{1}{n+1} F(x_n)^{n+1} \right]_{-\infty}^{+\infty} = \frac{1}{n+1},$$

où on a aussi utilisé que $\lim_{t \rightarrow -\infty} F(t) = 0$, $\lim_{t \rightarrow +\infty} F(t) = 1$ pour la dernière identité.

Calcul de $P(A_k \cap A_n)$. On peut maintenant reproduire le calcul, en écrivant

$$A_k \cap A_n = \{X_0 < X_k, X_1 < X_k, \dots, X_{k-1} < X_k, X_{k+1} < X_n, \dots, X_{n-1} < X_n\}.$$

De la même manière que précédemment, en utilisant de nouveau le théorème de Fubini–Tonelli, on obtient

$$\begin{aligned} P(A_k \cap A_n) &= \int_{\mathbb{R}^{n+1}} f(x_0) \cdots f(x_n) \mathbb{1}_{\{x_0 < x_k, \dots, x_{k-1} < x_k, x_{k+1} < x_n, \dots, x_{n-1} < x_n\}} dx_0 \cdots dx_n \\ &= \int_{-\infty}^{+\infty} f(x_n) \left(\int_{-\infty}^{x_n} \cdots \int_{-\infty}^{x_n} f(x_{k+1}) \cdots f(x_{n-1}) dx_{k+1} \cdots dx_{n-1} \right) g_k(x_n) dx_n, \end{aligned} \quad (\text{E-12.1})$$

où, pour éviter d'écrire trop d'intégrales, on a posé

$$g_k(x_n) =: \int_{-\infty}^{x_n} f(x_k) \left(\int_{-\infty}^{x_k} \cdots \int_{-\infty}^{x_k} f(x_0) \cdots f(x_{k-1}) dx_0 \cdots dx_{k-1} \right) dx_k.$$

Remarquons que l'on a $g_k(x) := P(X_0, \dots, X_{k-1} < X_k < x) = P(A_k \cap \{X_k < x\})$, et qu'en reprenant les calculs faits dans le paragraphe précédent on obtient

$$g_k(x) = \int_{-\infty}^x f(x_k) F(x_k)^k dx_k = \frac{1}{k+1} F(x)^{k+1}. \quad (\text{E-12.2})$$

En injectant dans (E-12.1) et en réutilisant le fait que $\int_{-\infty}^x f(u) du = F(x)$, on en déduit que

$$\begin{aligned} P(A_k \cap A_n) &= \int_{-\infty}^{+\infty} f(x_n) F(x_n)^{n-k-1} \frac{1}{k+1} F(x_n)^{k+1} dx_n \\ &= \frac{1}{k+1} \int_{-\infty}^{+\infty} f(x_n) F(x_n)^n dx_n = \frac{1}{k+1} \frac{1}{n+1}, \end{aligned}$$

ce qui est le résultat voulu.

Calcul de la covariance. Le dernier point du lemme vient simplement de l'observation que

$$\text{Cov}(\mathbb{1}_{A_k}, \mathbb{1}_{A_n}) = P(A_k \cap A_n) - P(A_k)P(A_n) = 0,$$

où l'on utilise les calculs précédents qui montrent que $P(A_k \cap A_n) = \frac{1}{k+1} \frac{1}{n+1} = P(A_k)P(A_n)$. \square

On peut déduire du Lemme 12.1 le théorème suivant sur le nombre de records observés jusqu'au n -ème tour, c'est-à-dire

$$R_n := \sum_{i=1}^n \mathbb{1}_{A_i}.$$

En effet, on sait calculer $E[R_n]$ et $\text{Var}(R_n)$:

$$E(R_n) = \sum_{i=1}^n P(A_i) = \sum_{i=1}^n \frac{1}{i+1}$$

qui est la série harmonique décalée de 1. En particulier, $\log(n+1) - 1 \leq E(R_n) \leq \log(n+1)$ et on a $E(R_n) \sim \log n$ quand $n \rightarrow \infty$.

D'autre part, en utilisant la Proposition 3.69-ii, on obtient que comm les covariances $\text{Cov}(\mathbb{1}_{A_k}, \mathbb{1}_{A_n})$ sont nulles,

$$\text{Var}(R_n) = \sum_{i=1}^n \text{Var}(\mathbb{1}_{A_i}) = \sum_{i=1}^n \frac{1}{i+1} \left(1 - \frac{1}{i+1}\right) \leq \sum_{i=1}^n \frac{1}{i+1} \leq \log(n+1).$$

Ainsi, en appliquant l'inégalité de Bienaymé–Tchebychev (Théorème 3.77), de même manière analogue à la Section 3.3.6,

$$P\left(|R_n - E(R_n)| \geq C\sqrt{\log(n+1)}\right) \leq \frac{\text{Var}(R_n)}{C^2 \log(n+1)} \leq \frac{1}{C^2}.$$

On peut donc en déduire que, quand n tend vers l'infini, avec grande probabilité, le nombre de records R_n vaut $\log n$, à une erreur de l'ordre de $O(\sqrt{\log n})$ près.

Observation 12.2. Il s'avère que l'on peut généraliser le calcul fait dans le Lemme 12.1 et obtenir le résultat suivant, suprenant à première vue.

Lemme 12.3. Les événements $(A_n = \{\text{un record est établi au } n^{\text{e}} \text{ tour}\})_{n \geq 1}$ sont indépendants, de probabilité $P(A_n) = \frac{1}{n+1}$.

Ainsi, le nombre de records R_n après n tours suit une loi qui est (très proche de) celle observée dans la Section 3.3.6.

Démonstration (du Lemme 12.3). On doit montrer que pour tous $k_1 < \dots < k_n$ des entiers quelconques

$$P(A_{k_1} \cap \dots \cap A_{k_n}) = P(A_{k_1}) \cdots P(A_{k_n}) = \frac{1}{k_1 + 1} \cdots \frac{1}{k_n + 1}. \quad (\text{E-12.3})$$

De manière analogue à la démonstration du Lemme 12.1, posons, pour $k_1 < \dots < k_n$ $x \in \mathbb{R}$,

$$g_{k_1, \dots, k_n}(x) = P(A_{k_1} \cap \dots \cap A_{k_n}, X_{k_n} \leq x).$$

On va montrer par récurrence que

$$\text{pour tous } k_1 < \dots < k_n \text{ et } x \in \mathbb{R}, \quad g_{k_1, \dots, k_n}(x) = \frac{1}{k_1 + 1} \cdots \frac{1}{k_n + 1} F(x)^n. \quad (\text{E-12.4})$$

Cela montrera (E-12.3), simplement en prenant $x \rightarrow \infty$ dans (E-12.4), car $g_{k_1, \dots, k_n}(x)$ tend vers $P(A_{k_1} \cap \dots \cap A_{k_n})$ et $F(x)$ tend vers 1.

On a déjà fait le calcul de (E-12.4) pour $n = 1$, voir (E-12.2). Il reste donc à montrer l'étape de récurrence. Commençons par observer que l'on peut réécrire

$$g_{k_1, \dots, k_n}(x) = P((X_{k_1}, \dots, X_{k_n}) \in D_{k_1, \dots, k_n}(x)),$$

où $D_{k_1, \dots, k_n}(x)$ est le sous-domaine de \mathbb{R}^{k_n+1} défini par

$$D_{k_1, \dots, k_n}(x) = \{(x_0, \dots, x_{k_n}) \in \mathbb{R}^{k_n+1}; x_0 < x_{k_1}, \dots, x_{k_1-1} < x_{k_1}, x_{k_1+1} < x_{k_2}, \dots, x_{k_2-1} < x_{k_2}, \\ \dots, x_{k_{n-1}-1} < x_{k_{n-1}}, x_{k_{n-1}+1} < x_{k_n}, \dots, x_{k_n-1} < x_{k_n} < x\},$$

de sorte que par la formule de transfert on a

$$g_{k_1, \dots, k_n}(x) = \int_{\mathbb{R}^{k_n+1}} f(x_0) \cdots f(x_{k_n}) \mathbb{1}_{\{(x_0, \dots, x_{k_n}) \in D_{k_1, \dots, k_n}(x)\}} dx_0 \cdots dx_{k_n}. \quad (\text{E-12.5})$$

Soient maintenant $k_1 < \dots < k_{n+1}$ et $x \in \mathbb{R}$. Un point important est que, au vu de sa définition, on peut écrire $D_{k_1, \dots, k_{n+1}}(x)$ de la façon suivante :

$$D_{k_1, \dots, k_{n+1}}(x) = \{(x_0, \dots, x_{k_{n+1}}) \in \mathbb{R}^{k_{n+1}+1}; (x_0, \dots, x_{k_n}) \in D_{k_1, \dots, k_n}(x_{k_{n+1}}), \\ x_{k_{n+1}+1} < x_{k_{n+1}}, \dots, x_{k_{n+1}-1} < x_{k_{n+1}}\}.$$

Ainsi, par la formule de transfert et en appliquant le théorème de Fubini–Tonelli,

$$g_{k_1, \dots, k_{n+1}}(x) = \int_{-\infty}^x f(x_{k_{n+1}}) \left(\int_{-\infty}^{x_{k_{n+1}}} \cdots \int_{-\infty}^{x_{k_{n+1}}} f(x_{k_{n+1}+1}) \cdots f(x_{k_{n+1}-1}) dx_{k_{n+1}+1} \cdots dx_{k_{n+1}-1} \right. \\ \left. \int_{\mathbb{R}^{k_n+1}} f(x_0) \cdots f(x_{k_n}) \mathbb{1}_{\{(x_0, \dots, x_{k_n}) \in D_{k_1, \dots, k_n}(x)\}} dx_0 \cdots dx_{k_n} \right) dx_{k_{n+1}} \\ = \int_{\mathbb{R}} f(x_{k_{n+1}}) F(x_{k_{n+1}})^{k_{n+1}-k_n-1} g(x_{k_{n+1}}) dx_{k_{n+1}},$$

en ayant utilisé encore une fois que $\int_{-\infty}^x f(t) dt = F(x)$, et la formule (E-12.5).

En utilisant l'hypothèse de récurrence (E-12.4), on en déduit que

$$\begin{aligned}
g_{k_1, \dots, k_{n+1}}(x) &= \int_{-\infty}^x f(x_{k_{n+1}}) F(x_{k_{n+1}})^{k_{n+1}-k_n-1} \frac{1}{k_1+1} \cdots \frac{1}{k_n+1} F(x_{k_{n+1}})^{k_n+1} dx_{k_{n+1}} \\
&= \frac{1}{k_1+1} \cdots \frac{1}{k_n+1} \int_{-\infty}^x f(x_{k_{n+1}}) F(x_{k_{n+1}})^{k_{n+1}} dx_{k_{n+1}}.
\end{aligned}$$

Un calcul explicite donne alors que la dernière intégrale vaut

$$\int_{-\infty}^x f(u) F(u)^{k_{n+1}} du = \left[\frac{1}{k_{n+1}+1} F(u) \right]_{-\infty}^x = \frac{1}{k_{n+1}+1} F(x)^{k_{n+1}+1},$$

ce qui montre que (E-12.4) est valable pour $n+1$. Cela conclut la récurrence et la démonstration du lemme. \square