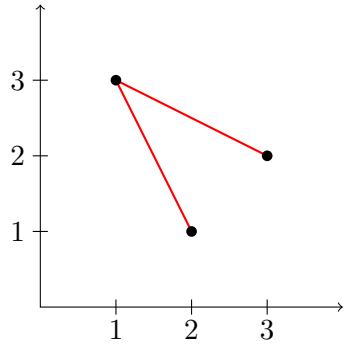
Nombre d'inversions d'une permutation aléatoire

mots-clés : permutations aléatoires, séries génératrices, grandes déviations.

Dans tout le texte, $\mathfrak{S}(N)$ désigne l'ensemble des permutations de taille N, c'est-à-dire les bijections $\sigma : [1, N] \to [1, N]$. On rappelle que $\operatorname{card} \mathfrak{S}(N) = N! = 1 \times 2 \times 3 \times \cdots \times N$. Une *inversion* d'une permutation $\sigma \in \mathfrak{S}(N)$ est un couple (k, l) d'entiers dans $[1, N]$ tels que $k < l$ et $\sigma(k) > \sigma(l)$. Par exemple, la permutation $\sigma = 312$ de taille 3 qui envoie 1 sur 3, 2 sur 1 et 3 sur 2 a deux inversions, à savoir les couples $(1, 2)$ et $(1, 3)$. Si l'on représente une permutation par son graphe, alors les inversions correspondent aux segments décroissants que l'on peut dessiner entre deux points du graphe :



Si $\sigma \in \mathfrak{S}(N)$, notons $I(\sigma)$ le nombre d'inversions de σ. Ainsi, $I(312) = 2$. L'objectif de ce texte est d'étudier la variable aléatoire $I_N = I(\sigma_N)$, où σ_N est une permutation aléatoire choisie suivant la distribution uniforme sur $\mathfrak{S}(N)$:

$$\forall \sigma \in \mathfrak{S}(N), \quad \mathbb{P}[\sigma_N = \sigma] = \frac{1}{N!}.$$

En particulier, on donnera un équivalent asymptotique de $\mathbb{P}[I_N \geq xN^2]$ pour certaines valeurs de x, et pour N tendant vers $+\infty$.

1 Construction de permutations uniformes

Si $\sigma \in \mathfrak{S}(N)$, alors ses inversions forment une partie de l'ensemble des paires d'éléments de $[1, N]$, donc

$$0 \leq I(\sigma) \leq \binom{N}{2}. \tag{1}$$

Pour comprendre la distribution de I_N, il est utile d'avoir une construction explicite d'une permutation aléatoire σ_N de loi uniforme sur $\mathfrak{S}(N)$, et qui se comporte bien vis-à-vis des nombres d'inversions :

Théorème 1. *Supposons donnée une permutation $\sigma' \in \mathfrak{S}(N-1)$ de loi uniforme sur $\mathfrak{S}(N-1)$, et un entier k choisi uniformément dans $[1, N]$ et indépendant de σ'. On définit alors $\sigma \in \mathfrak{S}(N)$ par la formule suivante :*

$$\begin{aligned} \sigma(N) &= k; \\ \sigma(i) &= \sigma'(i) && \text{si } 1 \leq \sigma'(i) < k; \\ \sigma(j) &= \sigma'(j) + 1 && \text{si } k \leq \sigma'(j) \leq N-1. \end{aligned}$$

1. *La permutation σ est de loi uniforme sur $\mathfrak{S}(N)$.*

2. *Le nombre d'inversions de σ est $I(\sigma) = I(\sigma') + N - k$.*

Démonstration. Si l'on montre que l'association

$$\begin{aligned} \Psi : \mathfrak{S}(N-1) \times [\![1, N]\!] &\to \mathfrak{S}(N) \\ (\sigma', k) &\mapsto \sigma \end{aligned}$$

est une bijection, alors comme la loi du couple (σ', k) est uniforme sur l'ensemble produit $\mathfrak{S}(N-1) \times [\![1, N]\!]$ qui est de taille $N!$, la loi de σ sera bien la loi uniforme sur $\mathfrak{S}(N)$. Or, étant donnée $\sigma \in \mathfrak{S}(N)$, on peut récupérer le couple (σ', k) en posant $k = \sigma(N)$, puis en renumérotant les lettres du mot $\sigma(1)\sigma(2)\cdots\sigma(N-1)$: toutes les lettres strictement inférieures à k sont conservées, et toutes les lettres strictement supérieures à k sont diminuées d'une unité. Par exemple, si $N = 9$ et $\sigma = 357926814$, alors $k = 4$ et σ' est obtenue en renumérotant le préfixe 35792681 en 34682571. On a donc bien inversé l'application Ψ.

Comparons maintenant les deux ensembles d'inversions $\mathrm{Inv}(\sigma')$ et $\mathrm{Inv}(\sigma)$. Le point clef est que toute inversion (i, j) de σ' avec $1 \leq i < j \leq N-1$ est encore une inversion de σ. De plus, tous les couples (i, N) avec $\sigma'(i) \geq k$ sont des inversions de σ, car $\sigma(i) = \sigma'(i) + 1 > k = \sigma(N)$. Il y a $N - k$ entiers $i \in [\![1, N-1]\!]$ tels que $k \leq \sigma'(i) \leq N-1$, d'où la formule pour le nombre d'inversions de σ. □

Si des entiers k_n sont choisis indépendamment et uniformément dans $[\![1, n]\!]$ pour tout $n \in [\![1, N]\!]$, alors on peut ainsi construire récursivement une permutation σ de loi uniforme sur $\mathfrak{S}(N)$, dont le nombre d'inversions est

$$\sum_{n=1}^{N} (n - k_n) = \frac{N(N+1)}{2} - \sum_{k=1}^{n} k_n.$$

On a donc :

Corollaire 2. *La loi de I_N est celle d'une somme $\sum_{n=1}^{N} U_n$, où les U_n sont indépendants, et où chaque variable U_n suit une loi uniforme sur l'ensemble d'entiers $\{0, 1, \ldots, n-1\}$.*

2 Moyenne, variance et série génératrice

Le corollaire 2 implique immédiatement les formules suivantes pour la moyenne et la variance de I_N :

$$\begin{aligned} \mathbb{E}[I_N] &= \sum_{n=1}^{N} \mathbb{E}[U_n] = \sum_{n=1}^{N} \frac{n-1}{2} = \frac{N(N-1)}{4}; \\ \mathrm{var}(I_N) &= \sum_{n=1}^{N} \mathrm{var}(U_n) = \sum_{n=1}^{N} \frac{n^2-1}{12} = \frac{N(N-1)(2N+5)}{72}. \end{aligned}$$

Il est aussi facile de voir que I_N et $\binom{N}{2} - I_N$ ont même loi. Pour mieux comprendre la distribution de I_N, on peut utiliser la série génératrice :

$$\mathbb{E}[z^{I_N}] = \prod_{n=1}^{N} \left(\frac{1 + z + \cdots + z^{n-1}}{n} \right). \tag{2}$$

Théorème 3. *Lorsque N tend vers l'infini, $A_N = \frac{1}{N^{3/2}} (I_N - \mathbb{E}[I_N])$ tend en loi vers une gaussienne centrée de variance $\frac{1}{36}$.*

Démonstration. La Formule (2) est valable pour tout nombre complexe z, donc on peut l'employer pour calculer les transformées de Fourier :

$$\begin{aligned}
\mathbb{E}[\mathrm{e}^{\mathrm{i}\xi I_N}] &= \prod_{n=1}^{N} \frac{1+\mathrm{e}^{\mathrm{i}\xi}+\cdots+\mathrm{e}^{\mathrm{i}(n-1)\xi}}{n};\\
\mathbb{E}[\mathrm{e}^{\mathrm{i}\xi A_N}] &= \left(\prod_{n=1}^{N} \frac{1+\mathrm{e}^{\frac{\mathrm{i}\xi}{N^{3/2}}}+\cdots+\mathrm{e}^{\mathrm{i}\frac{(n-1)\xi}{N^{3/2}}}}{n}\right) \mathrm{e}^{-\frac{N-1}{4N^{1/2}}\,\mathrm{i}\xi}\\
&= \left(\prod_{n=1}^{N} \frac{1-\mathrm{e}^{n\frac{\mathrm{i}\xi}{N^{3/2}}}}{n(1-\mathrm{e}^{\frac{\mathrm{i}\xi}{N^{3/2}}})}\,\mathrm{e}^{-\frac{(n-1)\mathrm{i}\xi}{2N^{3/2}}}\right) = \prod_{n=1}^{N} \frac{\sin(\frac{n\xi}{2N^{3/2}})}{n\sin(\frac{\xi}{2N^{3/2}})}.
\end{aligned}$$

Si ξ est fixé dans $\mathbb{R}\setminus\{0\}$, alors le développement de Taylor jusqu'à l'ordre $O(\frac{1}{N^2})$ donne :

$$\mathbb{E}[\mathrm{e}^{\mathrm{i}\xi A_N}] = \prod_{n=1}^{N} \exp\left(-\frac{(n^2-1)\,\xi^2}{24N^3}+O\left(\frac{1}{N^2}\right)\right) = \exp\left(-\frac{\xi^2}{72}+O\left(\frac{1}{N}\right)\right), \tag{3}$$

les constantes dans les $O(\cdot)$ pouvant dépendre de $|\xi|$. La convergence point par point des transformées de Fourier implique ensuite la convergence en loi. □

Le théorème 3 indique que pour N grand, le nombre d'inversions I_N d'une permutation aléatoire uniforme σ_N est proche de $\mathbb{E}[I_N] \simeq \frac{N^2}{4}$ avec grande probabilité, et avec des fluctuations gaussiennes d'ordre $N^{3/2}$. En particulier, si $x \in (\frac{1}{4}, \frac{1}{2})$, alors la probabilité $\mathbb{P}[I_N \geq xN^2]$ d'avoir une déviation à la moyenne d'ordre $(x-\frac{1}{4})\,N^2$ doit tendre vers 0. La section suivante donne une estimation précise de ceci.

3 Déviations d'ordre N^2

Dans toute cette section, ε est un paramètre dans $(0,1)$, et on cherche un équivalent asymptotique de la probabilité $\mathbb{P}[I_N \geq (1+\varepsilon)\mathbb{E}[I_N]]$. Posons $\phi(t) = \log\left(\frac{\mathrm{e}^t-1}{t}\right)$; en prolongeant cette fonction par continuité en $\phi(0)=0$, on obtient une fonction infiniment dérivable sur $\mathbb{R}$. En remplaçant $\mathrm{i}\xi$ par un paramètre réel $t\in\mathbb{R}$ dans les formules précédentes, on obtient :

$$\log \mathbb{E}\left[\mathrm{e}^{t\frac{I_N}{N}}\right] = \sum_{n=1}^{N} \log\left(\frac{\mathrm{e}^{\frac{tn}{N}}-1}{n(\mathrm{e}^{\frac{t}{N}}-1)}\right) = \left(\sum_{n=1}^{N}\phi\left(\frac{nt}{N}\right)\right) - N\phi\left(\frac{t}{N}\right).$$

Pour toute fonction f de classe $\mathcal{C}^2$ sur $[0,1]$, la somme de Riemann $\sum_{n=1}^{N} f(\frac{n}{N})$ vérifie la formule d'approximation d'Euler–Maclaurin :

$$\sum_{n=1}^{N} f\left(\frac{n}{N}\right) = N\int_0^1 f(x)\,dx + \frac{f(1)-f(0)}{2} + O\left(\frac{\|f'\|_\infty}{N}\right).$$

Ceci implique :

Proposition 4. *Soit $\psi(t) = \phi(t) - \frac{t}{2} = \log\left(\frac{\sinh\frac{t}{2}}{\frac{t}{2}}\right)$. La fonction ψ est paire et infiniment dérivable sur $\mathbb{R}$, et*

$$\log \mathbb{E}\left[\mathrm{e}^{t\,\frac{I_N-\mathbb{E}[I_N]}{N}}\right] = N\int_0^1 \psi(xt)\,dx + \frac{\psi(t)}{2} + O\left(\frac{1}{N}\right)$$

avec une constante universelle (indépendante de t) dans le $O(\cdot)$.

L'inégalité de Chernov permet d'en déduire une bonne estimée de $\mathbb{P}[I_N \geq (1+\varepsilon)\mathbb{E}[I_N]]$. Remarquons que les fonctions $\psi(t)$ et $\Lambda(t) = \int_0^1 \psi(xt)\,dx$ sont convexes, avec $\lim_{t\to\infty} \Lambda'(t) = \frac{1}{4}$. En particulier, pour tout $\varepsilon \in (0,1)$, il existe un unique paramètre $h \in (0,+\infty)$ tel que $\Lambda'(h) = \frac{\varepsilon}{4}$. De plus, h vérifie :

$$\forall u \in \mathbb{R}_+,\ \ \Lambda(u) - \frac{u\varepsilon}{4} \geq \Lambda(h) - \frac{h\varepsilon}{4}.$$

Alors,

$$\log \mathbb{P}[I_N \geq (1+\varepsilon)\mathbb{E}[I_N]] \leq N\left(\Lambda(h) - \frac{h\varepsilon}{4}\right) + \frac{\psi(h)}{2} + h\varepsilon\frac{1}{4} + O\left(\frac{1}{N}\right). \tag{4}$$

Comme $\psi(h) + \frac{h\varepsilon}{2} < \psi(h) + \frac{h}{2} = \phi(h)$, on en déduit :

Théorème 5. *Si $\varepsilon \in (0,1)$ et h est l'unique paramètre positif tel que $\Lambda'(h) = \frac{\varepsilon}{4}$, alors pour N assez grand,*

$$\mathbb{P}[I_N \geq (1+\varepsilon)\mathbb{E}[I_N]] \leq \exp\left(-N\left(\frac{h\varepsilon}{4} - \Lambda(h)\right) + \frac{\phi(h)}{2}\right).$$

La fonction $\varepsilon \mapsto \Lambda^*(\varepsilon) = \frac{h\varepsilon}{4} - \Lambda(h)$ est une bijection croissante convexe de $(0,1)$ vers $(0,+\infty)$.

Ainsi, la probabilité d'avoir une déviation plus grande que $\varepsilon\,\mathbb{E}[I_N] \simeq \frac{\varepsilon N^2}{4}$ décroît au moins comme $\mathrm{e}^{-N\,\Lambda^*(\varepsilon)}$. Des techniques plus avancées de calcul permettraient de trouver un vrai équivalent de la probabilité lorsque N tend vers l'infini : ainsi, avec les mêmes notations qu'avant, on peut montrer que :

$$\mathbb{P}[I_N \geq (1+\varepsilon)\mathbb{E}[I_N]] = \frac{1}{h\sqrt{2\pi N \Lambda''(h)}}\exp\left(-N\,\Lambda^*(\varepsilon) + \frac{\psi(h)}{2}\right)(1+o(1)). \tag{5}$$

Questions

1. L'inégalité (1) est-elle optimale ? On pourra chercher des permutations de taille N avec $I(\sigma) = 0$ ou $I(\sigma) = \binom{N}{2}$.

2. Démontrer complètement la seconde partie du Théorème 1 : en particulier, expliquer pourquoi on a $\mathrm{Inv}(\sigma') \subset \mathrm{Inv}(\sigma)$.

3. Écrire un algorithme qui, étant donné un entier $N \geq 1$, construit les N premiers états d'une chaîne de Markov $(\sigma_k, i_k)_{1\leq k\leq N}$ d'espace d'états

$$\mathfrak{X} = \bigsqcup_{n=0}^{\infty} (\mathfrak{S}(n) \times \mathbb{N})$$

   avec les propriétés suivantes :
   - $i_k = I(\sigma_k)$ pour tout $k \in [\![1, N]\!]$;
   - la distribution de σ_k est uniforme sur $\mathfrak{S}(k)$ pour tout $k \in [\![1, N]\!]$.

4. Illustrer le Théorème 3, par exemple en dessinant la fonction de répartition empirique de la variable aléatoire I_N pour $N = 100$. Commenter cette illustration.

5. Démontrer la formule (2) à partir du Corollaire 2.

6. Partant de la formule $\mathbb{E}[\mathrm{e}^{\mathrm{i}\xi A_N}] = \prod_{n=1}^{N} \frac{\sin(\frac{n\xi}{2N^{3/2}})}{n\sin(\frac{\xi}{2N^{3/2}})}$, démontrer l'estimée (3). Pourquoi implique-t-elle

$$A_N \rightharpoonup_{N\to\infty} \mathcal{N}\left(0, \frac{1}{36}\right) \quad \text{(convergence en loi)?}$$

7. Dessiner la fonction
$$\psi(t) = \log\left(\frac{\sinh\frac{t}{2}}{\frac{t}{2}}\right)$$
(attention au fait que c'est un sinus hyperbolique au numérateur). Pourquoi son prolongement par continuité ($\psi(0) = 0$) est-il de classe $\mathcal{C}^\infty$ sur $\mathbb{R}$? Montrer aussi que la fonction ψ est paire, convexe, strictement croissante sur $\mathbb{R}_+$, avec $\lim_{t\to\infty} \psi'(t) = \frac{1}{2}$.

8. Montrer que la fonction $\Lambda(t)$ est elle aussi $\mathcal{C}^\infty$, paire, convexe, strictement croissante sur $\mathbb{R}_+$, et avec $\lim_{t\to\infty} \Lambda'(t) = \frac{1}{4}$. Dessiner (à la main) l'aspect de cette fonction.

9. En combinant la Proposition 4 et l'inégalité de Chernoff, démontrer l'inégalité (4). Pourquoi la définition de h par la formule $\Lambda'(h) = \varepsilon$ donne-t-elle une inégalité optimale ?

10. On pose $P(t) = 4\,\Lambda'(t)$; c'est une bijection croissante de $(0, +\infty)$ vers $(0, 1)$. Pour $\varepsilon \in (0, 1)$, exprimer $(\Lambda^*)'(\varepsilon)$ en fonction de $P^{-1}(\varepsilon)$. En déduire que Λ^* est une bijection croissante convexe de $(0, 1)$ vers $(0, +\infty)$. L'estimée (5) vous semble-t-elle aisée à mettre en évidence par des simulations ?