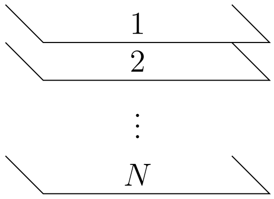
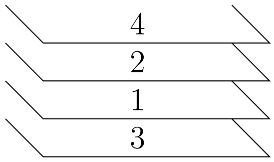
BATTAGES DE CARTES

On considère N cartes numérotées de 1 à N, placées initialement dans l'ordre, avec la carte 1 en haut du paquet, et la carte N en bas :



On note $\mathfrak{S}(N)$ l'ensemble des bijections de $[\![1, N]\!]$ dans $[\![1, N]\!]$ (*permutations* de taille N), et on identifie un ordre de cartes dans un paquet de N cartes avec un élément de $\mathfrak{S}(N)$. Par exemple, si $N = 4$ et si l'on regarde le paquet de cartes

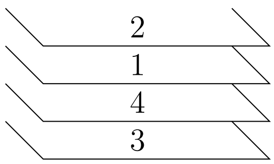


alors on peut le décrire par la bijection $\sigma \in \mathfrak{S}(4)$ qui envoie 1 sur 4, 2 sur 2, 3 sur 1 et 4 sur 3. On notera de façon abrégée cette bijection $\sigma = 4213$. D'autre part, on rappelle que $\mathfrak{S}(N)$ est un groupe pour l'opération $\circ$ de composition des bijections.

(1) Quel est le cardinal de $\mathfrak{S}(N)$? En déduire le nombre de paquets ordonnés de cartes possibles.

1. Battage par insertion

On se propose d'étudier la chaîne de Markov d'espace d'états $X = \mathfrak{S}(N)$, et dont les transitions peuvent être décrites comme suit. À chaque étape n, on tire au hasard un entier $k_n \in [\![1, N]\!]$ uniformément, et on insère la carte tout en haut du paquet à la k_n-ième position, en comptant les positions du haut vers le bas. Par exemple, si $\sigma_{n-1} = 4213$ est l'ordre du paquet au temps $n - 1$ et si l'on tire l'entier $k_n = 3$, alors le nouveau paquet de cartes σ_n est :



On suppose que $\sigma_0 = 1234 \ldots N$, et on note $(\sigma_n)_{n \in \mathbb{N}}$ la suite de permutations aléatoires ainsi obtenues, les entiers k_n étant supposés indépendants. Pour $k \in [\![1, N]\!]$, on note C_k la bijection cyclique qui envoie 1 sur 2, 2 sur 3, 3 sur 4, *etc.*, k sur 1, et laisse stable les autres entiers :

$$C_k = 23 \ldots k1(k + 1)(k + 2) \ldots N.$$

(2) Montrer que la suite de permutations $(\sigma_n)_{n\in\mathbb{N}}$ vérifie l'équation de récurrence
$$\sigma_n = \sigma_{n-1} \circ C_{k_n}.$$
En déduire que $(\sigma_n)_{n\in\mathbb{N}}$ est une chaîne de Markov sur l'espace $X = \mathfrak{S}(N)$, de matrice de transition :
$$P(\sigma,\tau) = \begin{cases} \frac{1}{N} & \text{si } \tau = \sigma \circ C_k \text{ pour un certain } k \in [\![1,N]\!], \\ 0 & \text{sinon.} \end{cases}$$

(3) Si $\sigma \in \mathfrak{S}(N)$, montrer qu'il existe $\tau \in \mathfrak{S}(N)$ avec $\tau(N) = N$, et un certain entier $k \in [\![1,N]\!]$, tels que
$$\tau = \sigma \circ (C_N)^k.$$
Montrer que la chaîne de Markov $(\sigma_n)_{n\in\mathbb{N}}$ est irréductible.

(4) On note $\pi_n(\sigma) = \mathbb{P}[\sigma_n = \sigma]$ la loi de la permutation aléatoire σ_n. Dessiner le graphe de la chaîne de Markov lorsque $N = 3$. Calculer dans ce cas π_1, π_2 et π_3, les premières marginales de la chaîne de Markov.

(5) Pour tout $\sigma \in \mathfrak{S}(N)$, que vaut $P(\sigma,\sigma)$? En déduire que la chaîne de Markov $(\sigma_n)_{n\in\mathbb{N}}$ est apériodique.

(6) Montrer que l'unique mesure invariante π telle que $\pi = \pi P$ est la mesure uniforme donnée par $\pi(\sigma) = \frac{1}{\mathrm{card}\,\mathfrak{S}(N)}$.

(7) Énoncer le théorème de convergence vérifié par la loi π_n de σ_n.

2. Remontée des cartes et borne sur la distance en variation totale

Dans cette section, on relie deux quantités calculées à partir de la chaîne de Markov :

— la position de la carte N dans le paquet ;

— et la *distance en variation totale* entre π_n et π :
$$d_{\mathrm{TV}}(\pi_n,\pi) = \sup_{A\subset\mathfrak{S}(N)} |\pi_n(A) - \pi(A)|$$
qui tend vers 0 et mesure la vitesse de la convergence en loi $\pi_n \to \pi$.

On note $x_n = \sigma_n^{-1}(N)$; par hypothèse, $x_0 = N$.

(8) Montrer que la distance en variation totale est donnée par la formule
$$d_{\mathrm{TV}}(\pi_n,\pi) = \frac{1}{2} \sum_{\sigma\in\mathfrak{S}(N)} |\pi_n(\sigma) - \pi(\sigma)|.$$
On pourra identifier un événement A qui maximise la différence $|\pi_n(A) - \pi(A)|$.

(9) Supposons $x_{n-1} = k > 1$. Montrer que
$$\mathbb{P}[x_n = k | x_{n-1} = k] = \frac{k-1}{N} \qquad ; \qquad \mathbb{P}[x_n = k-1 | x_{n-1} = k] = \frac{N+1-k}{N}$$
de sorte qu'au temps n, la carte N reste au même endroit, ou remonte d'une position dans le paquet si elle n'était pas en haut du paquet au temps $n-1$.

(10) Montrer que $(x_n)_{n\in\mathbb{N}}$ est une chaîne de Markov, et donner son espace d'états et sa matrice de transition. Décrire ses classes d'états.

(11) Pour $k \geq 1$, on pose $T_k = \inf\{n \geq 1,\ X_n = N - k\}$. Notons que T_{N-1} est le premier temps tel que la carte N soit remontée tout en haut du paquet. Montrer que les variables aléatoires $T_1, T_2 - T_1, \ldots, T_{N-1} - T_{N-2}$ sont indépendantes, et que chaque $U_k = T_k - T_{k-1}$ est une variable géométrique de paramètre $p = \frac{k}{N}$:

$$\forall j \geq 1,\ \mathbb{P}[T_k - T_{k-1} = j] = \left(\frac{N-k}{N}\right)^{j-1} \frac{k}{N}.$$

(12) Calculer $\mathbb{E}[T_k - T_{k-1}]$ et $\mathrm{var}(T_k - T_{k-1})$ pour tout $k \in [\![1, N-1]\!]$. En déduire que si $T = T_{N-1} + 1$, alors

$$\mathbb{E}[T] = N \log N + O(N) \qquad ; \qquad \mathrm{var}(T) = O(N^2).$$

On peut remarquer que, à tout temps $n \in [\![T_k, T_{k+1} - 1]\!]$, l'ordre des k cartes placées sous la carte N dans le paquet a pour loi la distribution uniforme sur l'ensemble des permutations de ces cartes. En effet, cette propriété est invariante par insertion aléatoire uniforme d'une nouvelle carte, et c'est précisément ce qui arrive aux temps T_1, T_2, *etc.* Par conséquent, la loi de l'ordre des $N-1$ cartes $1, \ldots, N-1$ placées sous N au temps T_{N-1} est la loi uniforme sur $\mathfrak{S}(N-1)$, et la loi de $\sigma_{T=T_{N-1}+1}$ est la loi uniforme sur $\mathfrak{S}(N)$:

$$\forall \sigma \in \mathfrak{S}(N),\ \mathbb{P}[\sigma_T = \sigma] = \frac{1}{\mathrm{card}\, \mathfrak{S}(N)}.$$

On peut alors deviner que, si l'on prend un temps déterministe n plus grand que la valeur de l'espérance du temps aléatoire T, on obtiendra avec σ_n une permutation de loi proche de la loi uniforme.

(13) Montrer que pour toute partie $A \subset \mathfrak{S}(N)$ et tout $k \leq n$, on a

$$\mathbb{P}[\sigma_n \in A \text{ et } T = k] = \mathbb{P}[T = k]\, \pi(A).$$

En déduire, en conditionnant par rapport à la valeur de T, que

$$\pi_n(A) = \mathbb{P}[X_n \in A] \leq \pi(A) + \mathbb{P}[T > n].$$

(14) En considérant le complémentaire de A, montrer que

$$\pi(A) \leq \pi_n(A) + \mathbb{P}[T > n],$$

puis que $d_{\mathrm{TV}}(\pi_n, \pi) \leq \mathbb{P}[T > n]$.

(15) Si $n = N \log N + cN$, montrer avec l'inégalité de Bienaymé–Chebyshev que

$$d_{\mathrm{TV}}(\pi_n, \pi) \leq \frac{K}{c^2}$$

pour une certaine constante positive K, et tout $c > 0$ suffisamment grand. Commenter.

On peut montrer qu'inversement, si $n = N \log N - cN$ avec $c > 0$, alors $d_{\mathrm{TV}}(\pi_n, \pi)$ reste assez grand, de sorte qu'une coupure s'opère au voisinage du temps $n = N \log N$.

Battage de cartes – Corrigé

(1) Le nombre de permutations de taille N est $N! = 1 \times 2 \times 3 \times \cdots \times N$.

(2) Supposons que le paquet de cartes au temps $n-1$ soit dans l'ordre suivant :

$$\begin{array}{c} \sigma_{n-1}(1) \\ \sigma_{n-1}(2) \\ \vdots \\ \sigma_{n-1}(k_n) \\ \vdots \\ \sigma_{n-1}(N) \end{array}$$

Si l'insertion au temps n se fait en position k_n, alors σ_n est égal à

$$\begin{array}{c} \sigma_{n-1}(2) \\ \sigma_{n-1}(3) \\ \vdots \\ \sigma_{n-1}(1) \\ \vdots \\ \sigma_{n-1}(N) \end{array} = \begin{array}{c} \sigma_{n-1}(C_{k_n}(1)) \\ \sigma_{n-1}(C_{k_n}(2)) \\ \vdots \\ \sigma_{n-1}(C_{k_n}(k_n)) \\ \vdots \\ \sigma_{n-1}(C_{k_n}(N)) \end{array}$$

On a donc bien $\sigma_n = \sigma_{n-1} \circ C_{k_n}$. On a ainsi écrit σ_n comme un système dynamique aléatoire $\sigma_n = f(\sigma_{n-1}, k_n)$, avec les k_n variables i.i.d. et $f : X \times [\![1, N]\!] \to X$; et ceci implique que $(\sigma_n)_{n\in\mathbb{N}}$ est une chaîne de Markov. Plus précisément, calculons $\mathbb{P}[\sigma_n = \tau \,|\, \sigma_{n-1} = \sigma]$. D'après ce qui précède, cette probabilité est nulle si τ ne s'écrit pas sous la forme $\sigma \circ c$ avec c cycle de la forme $(1, 2, \ldots, k)$. De plus, si $c = C_k = (1, 2, \ldots, k)$, alors

$$\mathbb{P}[\sigma_n = \sigma \circ C_k \,|\, \sigma_{n-1} = \sigma] = \mathbb{P}[k_n = k] = \frac{1}{N},$$

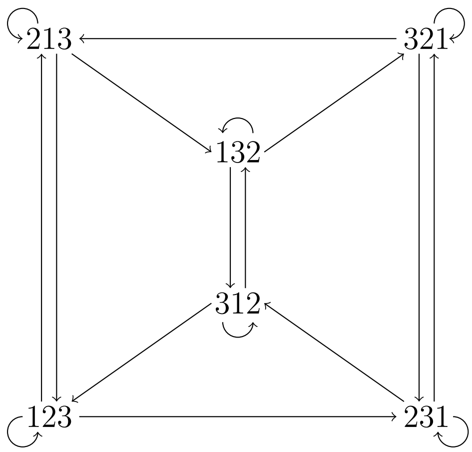
car k_n est choisi uniformément dans $[\![1, N]\!]$. Ceci donne la forme de la matrice de transition P.

(3) Soit $\sigma \in \mathfrak{S}(N)$, et $k = \sigma^{-1}(N)$, l'image réciproque de N par la bijection σ. L'image de N par $(C_N)^k$ est k, donc,

$$\left(\sigma \circ (C_N)^k\right)(N) = \sigma(k) = N.$$

Comme les probabilités de transition $\sigma \to \sigma C_N \to \sigma(C_N)^2 \to \cdots \to \sigma(C_N)^k$ sont non nulles, σ communique avec $\tau = \sigma(C_N)^k$, que l'on peut voir comme un élément de $\mathfrak{S}(N-1)$ puisque $\tau(N) = N$. Notons que respectivement, $\sigma = \tau(C_N)^{N-k}$, donc τ communique aussi avec σ. On montre alors par récurrence sur N que toute permutation $\sigma \in \mathfrak{S}(N)$ communique avec la permutation identité id_N, et réciproquement. Supposons le résultat vrai jusqu'au rang $N-1$. Alors, si $\sigma \in \mathfrak{S}(N)$, σ communique avec $\tau \in \mathfrak{S}(N-1)$, et réciproquement ; et τ communique dans les deux sens avec id_N. Par transitivité, on a donc bien σ qui communique avec id_N, et réciproquement. Ceci est vrai pour toute permutation $\sigma \in \mathfrak{S}(N)$, et implique donc que la chaîne de Markov est irréductible.

(4) Lorsque $N = 3$, la chaîne de Markov a pour graphe :



chaque flèche portant la probabilité $\frac{1}{3}$. Les marginales sont données par les vecteurs suivants :

$$\pi_1 = \left(\frac{1}{3}, 0, \frac{1}{3}, \frac{1}{3}, 0, 0\right) ;$$
$$\pi_2 = \left(\frac{2}{9}, \frac{1}{9}, \frac{2}{9}, \frac{2}{9}, \frac{1}{9}, \frac{1}{9}\right) ;$$
$$\pi_3 = \left(\frac{5}{27}, \frac{4}{27}, \frac{5}{27}, \frac{5}{27}, \frac{4}{27}, \frac{4}{27}\right) ;$$

où l'on fait figurer les probabilités des permutations dans cet ordre : 123, 132, 213, 231, 312, 321.

(5) On a $P(\sigma, \sigma) = \frac{1}{N}$, cette probabilité correspondant au choix du cycle aléatoire C_1, qui est la permutation identité. Comme cette probabilité est non nulle, la période de la chaîne de Markov (pgcd des entiers k tels que $P^k(\sigma, \sigma) > 0$) est égale à 1, *i.e.*, la chaîne est apériodique.

(6) La chaîne étant irréductible, il existe une unique mesure invariante. Vérifions que c'est la mesure uniforme $\pi(\sigma) = \frac{1}{N!}$. On a :

$$(\pi P)(\sigma) = \sum_{\tau \in \mathfrak{S}(N)} \pi(\tau)\, P(\tau, \sigma) = \frac{1}{N!} \sum_{\tau \in \mathfrak{S}(N)} P(\tau, \sigma).$$

Or, l'ensemble des permutations $\tau \in \mathfrak{S}(N)$ telles que $P(\tau, \sigma) \neq 0$ est égal à $\{\sigma \circ (C_k)^{-1},\ \ k \in [\![1, N]\!]\}$. Cet ensemble est de taille N, et chacun de ses éléments donne une probabilité de transition $\frac{1}{N}$, donc :

$$(\pi P)(\sigma) = \frac{1}{N!}\left(N \times \frac{1}{N}\right) = \frac{1}{N!} = \pi(\sigma).$$

Ainsi, la mesure uniforme est bien la loi invariante.

(7) On a une chaîne de Markov irréductible apériodique sur un ensemble fini, donc également récurrente positive. Par conséquent, le théorème de convergence vers la

loi stationnaire s'applique, et pour toute permutation $\sigma \in \mathfrak{S}(N)$,

$$\lim_{n\to\infty} \pi_n(\sigma) = \pi(\sigma) = \frac{1}{N!}.$$

(8) Soit B un événement dans $\mathfrak{S}(N)$, qu'on sépare en deux parties B_+ et B_- :

$$B_+ = \{\sigma \in B \,|\, \pi_n(\sigma) \geq \pi(\sigma)\} \qquad ; \qquad B_- = \{\sigma \in B \,|\, \pi_n(\sigma) < \pi(\sigma)\}.$$

On peut alors écrire :

$$\pi_n(B) - \pi(B) = (\pi_n(B_+) - \pi(B_+)) - (\pi(B_-) - \pi_n(B_-)),$$

qui est la différence de deux quantités positives. On a donc toujours :

$$\begin{aligned} -|\pi_n(B_-) - \pi(B_-)| &= -(\pi(B_-) - \pi_n(B_-)) \\ &\leq \pi_n(B) - \pi(B) \\ &\leq \pi_n(B_+) - \pi(B_+) = |\pi_n(B_+) - \pi(B_+)|. \end{aligned}$$

Par conséquent, $|\pi_n(B) - \pi(B)|$ est toujours inférieure à $|\pi_n(B_-) - \pi(B_-)|$ et à $|\pi_n(B_+) - \pi(B_+)|$. Une partie A qui maximise $|\pi_n(A) - \pi(A)|$ peut donc toujours être supposée constituée entièrement d'éléments tels que $\pi_n(\sigma) \geq \pi(\sigma)$, ou d'éléments tels que $\pi_n(\sigma) < \pi(\sigma)$. Appelons positifs ou négatifs de tels éléments, et positive ou négative une partie entièrement constituée d'éléments positifs ou négatifs. Si B est une partie positive, alors la quantité $\pi_n(B) - \pi(B)$ croît si l'on rajoute à B d'autres éléments positifs. De même, si B est une partie négative, alors la quantité $\pi(B) - \pi_n(B)$ croît si l'on rajoute à B d'autres éléments négatifs. Ainsi, on a démontré que $|\pi_n(A) - \pi(A)|$ est maximisée soit par

$$A_+ = \{\sigma \in \mathfrak{S}(N) \,|\, \pi_n(\sigma) \geq \pi(\sigma)\},$$

soit par

$$A_- = \{\sigma \in \mathfrak{S}(N) \,|\, \pi_n(\sigma) < \pi(\sigma)\}.$$

En réalité, ces deux ensembles donnent la même différence $|\pi_n(A) - \pi(A)|$, car A_- est le complémentaire de A_+, et

$$\pi_n(A_+) - \pi(A_+) = (1 - \pi_n(A_-)) - (1 - \pi(A_-)) = -(\pi_n(A_-) - \pi(A_-)).$$

On conclut finalement que

$$\begin{aligned} d_{\mathrm{TV}}(\pi_n, \pi) &= |\pi_n(A_+) - \pi(A_+)| = |\pi_n(A_-) - \pi(A_-)| \\ &= \frac{|\pi_n(A_+) - \pi(A_+)| + |\pi_n(A_-) - \pi(A_-)|}{2} \\ &= \frac{1}{2}\left(\sum_{\sigma\in A_+} |\pi_n(\sigma) - \pi(\sigma)| + \sum_{\sigma\in A_-} |\pi_n(\sigma) - \pi(\sigma)|\right) \\ &= \frac{1}{2}\sum_{\sigma\in\mathfrak{S}(N)} |\pi_n(\sigma) - \pi(\sigma)|. \end{aligned}$$

(9) Si $x_{n-1} = k$ avec $k > 1$, au temps n, la carte en haut du paquet est insérée soit au dessus de la carte N, soit en dessous. Il y a $k-1$ positions possibles au-dessus

de N, chacune avec probabilité $\frac{1}{N}$. Pour chacune de ces insertions, le nombre de cartes au-dessus de N ne change pas par l'insertion, donc

$$\mathbb{P}[x_n = k | x_{n-1} = k] = (k-1) \times \frac{1}{N} = \frac{k-1}{N}.$$

Pour les $N-(k-1)$ autres positions possibles, qui sont en dessous de la carte N, les insertions correspondantes font remonter d'une unité la carte N dans le paquet, donc

$$\mathbb{P}[x_n = k-1 | x_{n-1} = k] = (N-(k-1)) \times \frac{1}{N} = \frac{N-k+1}{N}.$$

(10) Justifions d'abord le fait que $(x_n)_{n\in\mathbb{N}}$ est une chaîne de Markov. On remarque que si $x_{n-1} = k > 1$, alors $x_n = k-1$ si $k_n \geq k$, et $x_n = k$ si $k_n < k$. Supposons maintenant $k = 1$; alors, la carte N est en haut du paquet, et au rang suivant, elle est réinsérée à une position $k_n \in [\![1, N]\!]$, donc $x_n = k_n$ dans ce cas. On a donc $x_n = g(x_{n-1}, k_n)$ avec la fonction

$$g(k,k') = \begin{cases} k-1 & \text{si } k' \geq k > 1, \\ k & \text{si } k > \max(1,k'), \\ k' & \text{si } k = 1. \end{cases}$$

Ceci prouve que $(x_n)_{n\in\mathbb{N}}$ est une chaîne de Markov. On a calculé dans la question précédente les probabilités de transition $Q(k,l)$ de la chaîne $(x_n)_{n\in\mathbb{N}}$ avec $k > 1$, et le raisonnement ci-dessus donne :

$$Q(1,l) = \frac{1}{N} \quad \text{pour tout } l \in [\![1, N]\!].$$

La matrice de transition de $(x_n)_{n\in\mathbb{N}}$ est donc :

$$Q = \begin{pmatrix} \frac{1}{N} & \frac{1}{N} & \frac{1}{N} & \cdots & \frac{1}{N} \\ \frac{N-1}{N} & \frac{1}{N} & 0 & \cdots & 0 \\ 0 & \frac{N-2}{N} & \frac{2}{N} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \frac{1}{N} & \frac{N-1}{N} \end{pmatrix}$$

L'espace des états de $(x_n)_{n\in\mathbb{N}}$ est l'intervalle $[\![1, N]\!]$, et tous les états communiquent entre eux (on peut en un nombre fini d'étapes remonter à la position $k = 1$, puis redescendre à l'étape suivante à n'importe quelle position k). La chaîne est donc irréductible, et elle est récurrente positive puisque d'espace fini.

(11) Chaque temps T_k est un temps d'atteinte d'un état par la chaîne de Markov $(x_n)_{n\in\mathbb{N}}$, donc un temps d'arrêt. De plus, la chaîne est récurrente positive, donc elle visite presque sûrement tous les états ; les temps T_k sont donc p.s. finis. Notons que chaque différence $U_k = T_k - T_{k-1}$ est égale au temps de visite de l'état $N-k+1$ par $(x_n)_{n\in\mathbb{N}}$ avant d'atteindre $N-k$. En particulier, U_k est une fonction mesurable de la chaîne décalée $(x_{n+T_{k-1}})_{n\in\mathbb{N}}$, qui est indépendante de la tribu $\mathcal{F}_{T_{k-1}}$ par la propriété de Markov forte. Les variables $T_1 \leq T_2 \leq \cdots \leq T_{k-1}$ étant $\mathcal{F}_{T_{k-1}}$-mesurables, on en déduit que U_k est indépendant de $(T_1, T_2, \ldots, T_{k-1})$, et donc de $(U_1, \ldots, U_{k-1})$. Par récurrence sur k, on montre donc que $U_1, \ldots, U_{N-1}$ sont des variables indépendantes.

Calculons maintenant la loi de U_k. On note $\mathbb{P}_x$ la loi de la chaîne de Markov sur $[\![1, N]\!]$ de matrice Q, et d'état initial x. Rappelons d'autre part que T_k est le temps d'atteinte de l'état $N-k$. Pour tout $r \geq 2$ et tout $k \in [\![1, N-1]\!]$, on a

$$\begin{aligned}
\mathbb{P}_N[U_k \geq r] &= \mathbb{P}_{N-k+1}[T_k \geq r] \\
&= \mathbb{P}_{N-k+1}[T_k \geq r \text{ et } x_1 = N-k+1] \\
&= \mathbb{P}_{N-k+1}[x_1 = N-k+1]\; \mathbb{P}_{N-k+1}[T_k \geq r \,|\, x_1 = N-k+1] \\
&= \frac{N-k}{N}\; \mathbb{P}_{N-k+1}[T_k \geq r-1],
\end{aligned}$$

en utilisant la propriété de Markov pour la chaîne décalée $(x_{n+1})_{n\in\mathbb{N}}$, qui a loi $\mathbb{P}_{N-k+1}$ conditionnellement à $x_1 = N-k+1$. Sous $\mathbb{P}_0$, la variable U_k a donc une loi géométrique de paramètre $p = \frac{k}{N}$.

(12) L'espérance d'une variable géométrique de paramètre p est $\frac{1}{p}$, et sa variance est $\frac{1}{p^2} - \frac{1}{p}$. On a donc

$$\mathbb{E}[U_k] = \frac{N}{k} \qquad ; \qquad \mathrm{var}(U_k) = \frac{N^2}{k^2} - \frac{N}{k}.$$

Comme les variables U_k sont indépendantes sous $\mathbb{P}_N$, on en déduit que

$$\mathbb{E}[T] = 1 + \sum_{k=1}^{N-1} \frac{N}{k} = \sum_{k=1}^{N} \frac{N}{k} = N\log N + N\gamma + o(N) = N\log N + O(N),$$

car $\sum_{k=1}^{N} \frac{1}{k} = \log N + \gamma + o(1)$ avec $\gamma = 0.57\ldots$ De même,

$$\begin{aligned}
\mathrm{var}(T) &= \sum_{k=1}^{N-1} \frac{N^2}{k^2} - \frac{N}{k} = \sum_{k=1}^{N} \frac{N^2}{k^2} - \frac{N}{k} \\
&= \frac{\pi^2 N^2}{6} - O(N) - O(N\log N) = \frac{\pi^2 N^2}{6} - O(N\log N)
\end{aligned}$$

car $\sum_{k=1}^{N} \frac{1}{k^2} = \frac{\pi^2}{6} - O(\frac{1}{N})$.

(13) Par la propriété de Markov forte, $(\sigma_{T+m})_{m\in\mathbb{N}}$ est une chaîne de Markov de matrice de transition P, et de loi initiale la loi de σ_T, qui est la loi uniforme. De plus, elle est indépendante de T. On peut alors écrire :

$$\begin{aligned}
\mathbb{P}[\sigma_n \in A \text{ et } T = k] &= \mathbb{P}[\sigma_{T+(n-k)} \in A \text{ et } T = k] \\
&= \mathbb{P}[\sigma_{T+(n-k)} \in A]\, \mathbb{P}[T = k].
\end{aligned}$$

Or, la loi uniforme est invariante par P, donc toutes les marginales de $(\sigma_{T+m})_{m\in\mathbb{N}}$ sont toutes égales à la loi uniforme π. Par conséquent,

$$\mathbb{P}[\sigma_n \in A \text{ et } T = k] = \pi(A)\, \mathbb{P}[T = k].$$

On peut alors borner $\pi_n(A)$ comme suit :

$$\begin{aligned}\pi_n(A) &= \sum_{k=0}^{\infty} \mathbb{P}[\sigma_n \in A \text{ et } T = k] \\ &= \left(\sum_{k=0}^{n} \mathbb{P}[\sigma_n \in A \text{ et } T = k]\right) + \mathbb{P}[\sigma_n \in A \text{ et } T > n] \\ &= \left(\sum_{k=0}^{n} \pi(A)\,\mathbb{P}[T = k]\right) + \mathbb{P}[\sigma_n \in A \text{ et } T > n] \\ &\leq \pi(A) + \mathbb{P}[T > n].\end{aligned}$$

(14) Réciproquement, on a

$$\pi(A) = 1 - \pi(A^{\mathrm{c}}) \leq 1 - \pi_n(A^{\mathrm{c}}) + \mathbb{P}[T > n] = \pi_n(A) + \mathbb{P}[T > n]$$

en appliquant l'inégalité précédente à $A^{\mathrm{c}} = \{\sigma \in \mathfrak{S}(N) \,|\, \sigma \notin A\}$. Par conséquent, $|\pi_n(A) - \pi(A)| \leq \mathbb{P}[T > n]$, et comme ceci est vrai pour toute partie $A \subset \mathfrak{S}(N)$,

$$d_{\mathrm{TV}}(\pi_n, \pi) = \sup_{A \subset \mathfrak{S}(N)} |\pi_n(A) - \pi(A)| \leq \mathbb{P}[T > n].$$

(15) Par l'inégalité de Bienaymé Chebyshev,

$$\mathbb{P}[T > N \log N + cN] = \mathbb{P}[T - \mathbb{E}[T] > (c - O(1))(N)] \leq \frac{\mathrm{var}(T)}{(c - O(1))^2 N^2}.$$

Or, $\mathrm{var}(T) \leq KN^2$ pour une certaine constante K (on peut prendre $K = \frac{\pi^2}{6}$), donc

$$\mathbb{P}[T > N \log N + cN] \leq \frac{K}{(c - O(1))^2},$$

qui est plus petit que $\frac{2K}{c^2}$ pour c assez grand. On a alors montré que

$$d_{\mathrm{TV}}(\pi_n, \pi) \leq \mathbb{P}[T > n] = O\left(\frac{1}{c^2}\right)$$

si $n = N \log N + cN$. Donc, la distance en variation totale entre π_n et π devient petite à partir de $n \sim N \log N$ (en particulier, $d_{\mathrm{TV}}(\pi_n, \pi) \to 0$, et le théorème de convergence vers la loi stationnaire est redémontré).