Battages de cartes

CHAUB Thomas, LE SCORNET Pierre, ORSINI Nicolas

16/04/2018

1 Position du problème et démarche suivie

Le but de ce travail est de comprendre un modèle mathématique simple de battage de cartes. En particulier, en se donnant une définition précise de ce que peut être un mélange "satisfaisant" du paquet par rapport à sa configuration initiale, l'étude s'intéresse à approcher le plus finement possible le nombre optimal de battages nécessaires pour parvenir à un tel mélange.
De manière générale, si on se donne N cartes disposées en un paquet, on peut associer à chaque carte un nombre compris entre 1 et N, et ainsi associer à la configuration initiale du paquet une permutation de $\mathfrak{S}_N$. Un mélange du paquet aboutit dans ce cas à une permutation σ' de $\mathfrak{S}_N$.

— Nous présentons un modèle de battage très simple, dit par insertion, dans lequel les permutations associées sont étudiées après le mouvement d'une seule carte à l'intérieur du paquet. Nous montrons les propriétés de convergence du modèle vers un mélange convenable dans un sens à préciser.

— Ce cadre nous invite à nous intéresser au mouvement de la dernière carte de la configuration initiale au cours du temps, que nous illustrons par une simulation informatique.

— Les estimations numériques conduisent ensuite à un minorant précis du nombre de battages à effectuer.

— Enfin, les limites du modèle sont explorées, et d'autres types de battages sont introduits afin d'améliorer le minorant exhibé.

La référence principale de ce document est un sujet de l'agrégation externe de mathématiques [1].
Les simulations informatiques sont réalisées à l'aide de PYTHON.

2 Modèle du battage par insertion

2.1 Présentation et exemple

Commençons par présenter le modèle de battage par insertion. Plaçons-nous dans le cadre théorique ci-dessus et numérotons les N cartes de C_1 à C_N. La permutation σ associée à la configuration initiale est celle qui pour tout $i \in \{1, \dots, N\}$ associe C_i.
Un mélange peut être décomposé en plusieurs échanges successifs de position entre deux cartes. Ici, nous étudions le type le plus simple d'échange possible.

Définition 1 (**Insertion à la k-ième place**). *Soit $k \in \{1, \dots, N\}$. On appelle insertion à la k-ième place l'opération consistant à prendre la première carte du paquet et à la placer entre la k-ième et la (k+1)-ième carte.*

Remarquons tout d'abord qu'insérer à la k-ième place revient à composer à gauche par une permutation circulaire. En effet, la permutation σ' associée à la configuration finale du paquet est donnée par $\sigma' = (k, k-1, \dots, 2, 1)\sigma$. En particulier, une insertion à la première place ne permet pas de changer de configuration, alors qu'une insertion à la N-ième place amène la première carte sous le paquet. Dans ce cas, toutes les cartes sauf celle située initialement en première position "remontent" (leur position est incrémentée). Nous verrons comment ce phénomène est lié à une minoration du nombre de battages nécessaires.
Nous pouvons alors définir rigoureusement le battage par insertion.

Définition 2 (**Battage par insertion**). *Le battage par insertion d'un jeu de N cartes consiste à effectuer une suite d'insertions aléatoires, en choisissant à chaque étape au hasard uniformément dans $\{1, \dots, N\}$ la place à laquelle l'insertion à lieu, indépendamment de l'insertion précédente.*

Ceci revient à étudier une suite de variables aléatoires $(X_n)_{n\in\mathbb{N}}$ à valeurs dans $\mathfrak{S}_N$, et la loi de probabilité de chacune de ces variables.
Notons que l'indépendance de l'insertion étudiée par rapport à la précédente impose que la loi de probabilité de X_{k+1} ne dépende que de celle de X_k, pour tout $k \in \{1, \dots, N\}$. Il s'agit d'un type particulier de processus stochastique.

Définition 3 (**Chaîne de Markov (homogène)**). *Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires à valeurs dans un ensemble dénombrable E. $(X_n)_{n\in\mathbb{N}}$ est appelée chaîne de Markov si pour tout $n > 0$ et tous $i_0,\ i_1, \dots, i_{n-1},\ i,\ j \in E$:*

$$p_{ij} = \mathbb{P}(X_{n+1} = j \mid X_n = i,\ X_{n-1} = i_{n-1},\ \dots,\ X_0 = i_0) = \mathbb{P}(X_{n+1} = j \mid X_n = i)$$

Dans ce cas, E est appelé ***espace d'états*** *et les probabilités $\mathbb{P}(X_{n+1} = j \mid X_n = i)$ sont appelées* ***probabilités de transition*** *de $(X_n)_{n\in\mathbb{N}}$.*
La chaîne de Markov est dite ***homogène*** *si les probabilités de transition sont indépendantes de n.*

Le modèle de battage étudié fait apparaître une chaîne de Markov $(X_n)_{n\in\mathbb{N}}$ dont l'espace d'états est $\mathfrak{S}_N$ et dont les probabilités de transition sont définies par :

$$\forall\ \sigma, \sigma' \in \mathfrak{S}_N,\ \mathbb{P}(X_{n+1} = \sigma' \mid X_n = \sigma) = \begin{cases} \frac{1}{N} \text{ s'il existe } k \in \{1, \dots, N\} \text{ tel que } \sigma' = (k, k-1, \dots, 2, 1)\sigma \\ 0 \text{ sinon} \end{cases}$$

Il s'agit alors d'une chaîne de Markov homogène.

Exemple 1 (**Cas N=3**). *Voici une illustration simple dans le cas de $\mathfrak{S}_3$.*

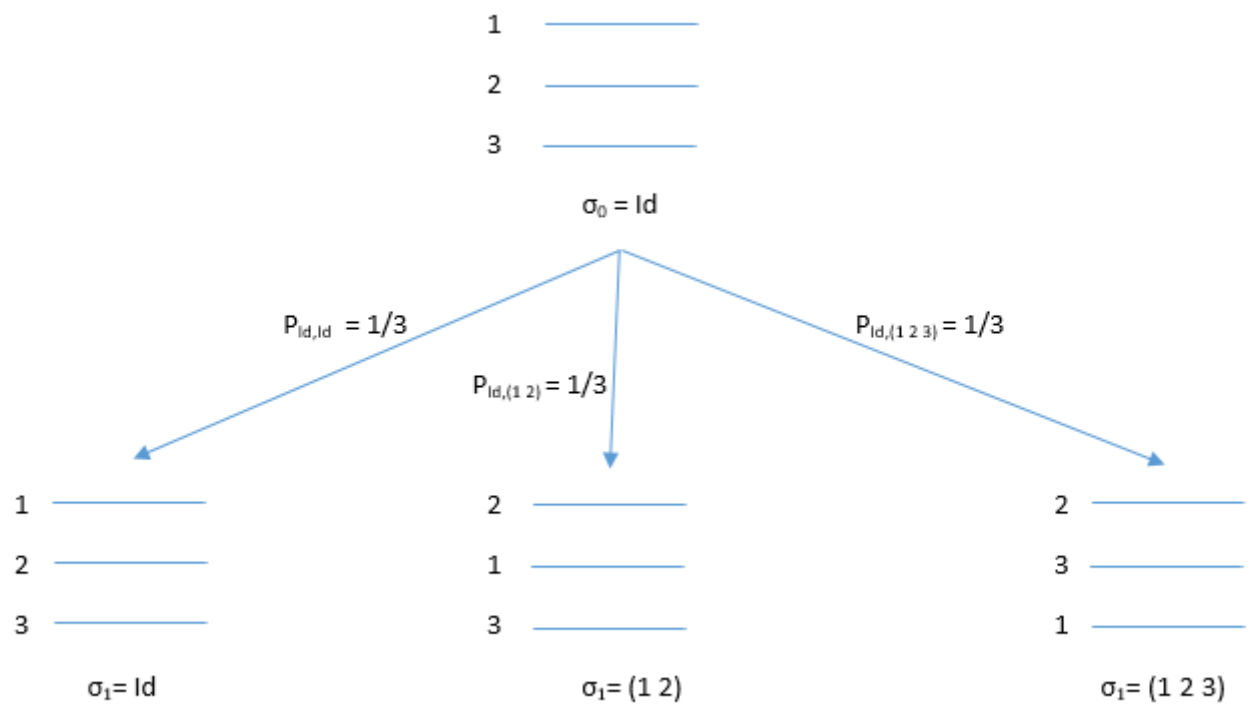


FIGURE 1 – **Probabilités de transition du battage par insertion d'un jeu de 3 cartes pour une configuration initiale égale à l'identité**

2.2 Convergence de la chaîne de Markov associée

A partir de maintenant, E est supposé **fini** (comme l'est $\mathfrak{S}_N$). En pratique, les définitions suivantes peuvent être généralisées, en adaptant notamment les règles de calcul matriciel [2]

Définition 4 (**Matrice de transition**). *La matrice de transition d'une chaîne de Markov est la matrice des probabilités de transition* $P = (p_{ij})_{1 \leq i,j \leq \mathrm{Card}(E)}$.

Une matrice de transition P est alors stochastique, par définition d'une mesure de probabilité : 1 est valeur propre de P.

Donnons quelques définitions relatives aux chaînes de Markov utiles pour la suite. On se fixe ici $(X_n)_{n\in\mathbb{N}}$ une chaîne de Markov.

Définition 5 (**Communication**). *Soit* $i, j \in E$. *On dit que* i *et* j *communiquent s'il existe* $M, M' > 0$ *tels que* $\mathbb{P}(X_M = j \mid X_0 = i) > 0$ *et* $\mathbb{P}(X_{M'} = i \mid X_0 = j) > 0$.

La relation de communication est clairement une relation d'équivalence, ce qui amène à partitionner E en classes de communication.

Définition 6 (**Irréductibilité**). *Une chaîne de Markov et sa matrice de transition sont dites irréductibles s'il n'existe qu'une seule classe de communication.*

Autrement dit, dans une chaîne de Markov irréductible, on peut toujours passer d'un état à un autre par un chemin de probabilité non nulle.

Définition 7 (**Apériodicité**). *Soit* $(X_n)_{n\in\mathbb{N}}$ *une chaîne de Markov irréductible de probabilités de transition* p_{ij}. *Notons* $p_{ii}^{(n)}$ *le coefficient en position* (i,i) *de* P^n.
La quantité $d = \mathrm{pgcd}(n \geq 1 \mid p_{ii}^{(n)} > 0)$ *ne dépend pas de* $i \in E$ *et est appelée la période de la chaîne.*
Si $d = 1$, $(X_n)_{n\in\mathbb{N}}$ *est dite apériodique.*

De manière informelle, une chaîne de Markov apériodique ne contient pas de "cycles".

Concluons cette séquence introductive par des définitions relatives à la convergence des chaînes de Markov.

Définition 8 (**Probabilité invariante**). *Soit une chaîne de Markov de matrice de transition* P. *Un vecteur non nul de coordonnées non négatives est une mesure de probabilité invariante de* P *si c'est un vecteur propre de* ${}^tP$ *pour la valeur propre* 1 *tel que la somme des coordonnées soit égale à* 1.

Ce vecteur est lui-même stochastique, ce qui explique le nom de mesure (il s'agit d'une mesure de probabilité sur $(E,\ P(E))$).
Voici un lemme qui sera utile concernant l'existence et l'unicité d'une mesure de probabilité invariante.

Lemme 1. *Si une chaîne de Markov* $(X_n)_{n\in\mathbb{N}}$ *irréductible admet une probabilité invariante* π, *alors celle-ci est unique et, pour tout* $i \in E$, $\pi(i) > 0$.
De plus, pour toute distribution initiale,

$$\lim_{n\to+\infty} \mathbb{P}(X_n = i) = \pi(i)$$

Définition 9 (**Loi uniforme**). *La loi uniforme* π *est l'application*

$$\begin{cases} P(\mathfrak{S}_N) & \longrightarrow & [0,1] \\ A & \mapsto & \dfrac{\mathrm{Card}(A)}{N!} \end{cases}$$

Il s'agit d'une loi de probabilité sur $(\mathfrak{S}_N, P(\mathfrak{S}_N))$.

Ces définitions permettent d'énoncer les propriétés de la chaîne de Markov donnée par le battage par insertion.

Théorème 1 (Caractéristiques de la chaîne de Markov $(X_n)_{n\in\mathbb{N}}$)**.** *La chaine de Markov $(X_n)_{n\in\mathbb{N}}$ est irréductible et apériodique. Elle possède une unique mesure de probabilité invariante sur $\mathfrak{S}_N$, qui est la mesure uniforme π et vers laquelle elle converge en loi.*

Démonstration. Procédons en trois étapes.

— Montrons que $(X_n)_{n\in\mathbb{N}}$ est irréductible. Pour cela, considérons σ et σ' deux permutations et montrons qu'il existe $(m,n)\in\mathbb{N}^2$ tels que $\mathbb{P}(X_n=\sigma' \mid X_0=\sigma)$ et $\mathbb{P}(X_m=\sigma \mid X_0=\sigma')$ soient strictement positifs.
Les permutations $(2,1)$ et $(N,N-1,\dots,2,1)$ engendrent $\mathfrak{S}_N$.

Il existe alors $p\in\mathbb{N}-\{0\}$ et $(\alpha_i)_{i\in[|1,p|]}\in\{(2,1),(N,N-1,\dots,2,1)\}^p$ tels que $\sigma'\sigma^{-1}=\alpha_1\circ\ldots\circ\alpha_p$.

$$\mathbb{P}(X_p=\sigma' \mid X_0=\sigma)=\frac{\mathbb{P}(X_p=\sigma'\bigcap X_0=\sigma)}{\mathbb{P}(X_0=\sigma)}\geq\frac{\mathbb{P}(X_p=\sigma'\bigcap\ X_{p-1}=\alpha_2\circ\ldots\circ\alpha_p\sigma\bigcap\ldots\bigcap X_0=\sigma)}{\mathbb{P}(X_0=\sigma)}$$

par croissance de la mesure de probabilité $\mathbb{P}$.
Posons a égal au membre de droite de l'inégalité. La propriété de Markov permet d'obtenir

$$a=\mathbb{P}(X_p=\sigma' \mid X_{p-1}=\alpha_2\circ\ldots\circ\alpha_p\sigma)\,\mathbb{P}(X_{p-1}=\alpha_2\circ\ldots\circ\alpha_p\sigma \mid X_{p-2}=\alpha_3\circ\ldots\circ\alpha_p\sigma)\ldots\mathbb{P}(X_1=\alpha_p\sigma \mid X_0=\sigma)=\left(\frac{1}{N}\right)^p$$

Alors, $\mathbb{P}(X_p=\sigma \mid X_0=\sigma)\geq\left(\frac{1}{N}\right)^p>0$
Par symétrie des rôles de σ et σ', p convient pour n et m.
Ceci montre que tous les états de $\mathfrak{S}_N$ communiquent, et donc la chaîne de Markov $(X_n)_{n\in\mathbb{N}}$ est irréductible.

— Montrons que $(X_n)_{n\in\mathbb{N}}$ est apériodique. Posons, pour tout $\sigma\in\mathfrak{S}_N$, $M_\sigma=\{n\geq 1,\ p_{\sigma\sigma}^{(n)}>0\}$.
La période de σ est définie par $\mathrm{pgcd}(M_\sigma)$. La chaîne étant irréductible, il y a une seule classe de communication, qui est de période 1 car $p_{\sigma\sigma}^{(n)}>0$. Alors, la chaîne $(X_n)_{n\in\mathbb{N}}$, est apériodique.

— Montrons que la mesure uniforme π est l'unique mesure de probabilité invariante par $(X_n)_{n\in\mathbb{N}}$ sur $\mathfrak{S}_N$, et qu'il y a convergence en loi. Posons, pour tout $\tau\in\mathfrak{S}_N$, $S=\{\sigma\in\mathfrak{S}_N,\ p_{\sigma\tau}>0\}$.
Par définition de la matrice de transition de $(X_n)_{n\in\mathbb{N}}$ et de $p_{\sigma\tau}$, $S=\{\tau,\ (2,1)\tau,\ \ldots\ (N,N_1,\dots,2,1)\tau\}$.
$\mathfrak{S}_N$ étant un groupe, Card $S=N$, ce qui montre que $\sum_{\sigma\in\mathfrak{S}_N}p_{\sigma\tau}=1$. Alors, ${}^tP$ est aussi une matrice de transition.

Or si $\pi=\frac{1}{N!}(1,\dots,1)$ est la mesure uniforme sur $\mathfrak{S}_N$, π est un vecteur propre à gauche de ${}^tP$, car ${}^t\pi=P\,{}^t\pi$.
$(X_n)_{n\in\mathbb{N}}$ est alors une chaîne de Markov homogène admettant une probabilité invariante π, et le lemme 1 montre l'unicité de la probabilité invariante. De plus, pour toute distribution initiale

$$\lim_{n\to+\infty}P(X_n=\sigma)=\pi(\sigma)=\frac{1}{N!}$$

Cette égalité montre la convergence étroite des mesures de probabilité $\mathbb{P}_{X_n}$ vers π, équivalant à la convergence en loi de $(X_n)_{n\in\mathbb{N}}$ vers π.

□

2.3 Formalisation d'un mélange "satisfaisant"

Le principal enjeu de l'étude est la détermination d'un nombre de battages nécessaires pour que le paquet soit convenablement mélangé.
Nous allons utiliser la convergence en loi donnée par le théorème 1 pour établir une notion de mélange "convenable". Etant données deux mesures de probabilité sur $\mathfrak{S}_N$, il s'agit de mesurer le plus grand écart de mesure sur une partie de $\mathfrak{S}_N$.

Définition 10 (Distance en variation). *Soient μ et ν deux mesures de probabilité sur $\mathfrak{S}_N$. La distance en variation entre μ et ν est le réel $d_V(\mu,\nu) \in [0,1]$ défini par*

$$d_V(\mu,\nu) = \max_{A \subset P(\mathfrak{S}_N)} |\mu(A) - \nu(A)|$$

Cette définition est assez intuitive : en particulier, il est clair qu'il s'agit bien d'une distance sur l'ensemble des mesures de probabilité définies sur $\mathfrak{S}_N$.
L'intérêt majeur de cette définition est son bon comportement lors d'un passage à la limite. En effet, si on note, pour tout $n \in \mathbb{N}$, μ_n la loi de probabilité de X_n, on a le résultat :

$$\lim_{n \to +\infty} d_V(\mu_n, \pi) = 0$$

Cela découle immédiatement de la proposition suivante.

Proposition 1. *Pour toutes mesures de probabilité μ et ν sur $\mathfrak{S}_N$, $d_V(\mu,\nu) = \frac{1}{2} \sum_{\sigma \in \mathfrak{S}_N} |\mu(\{\sigma\}) - \nu(\{\sigma\})|$*

Démonstration. Soient μ et ν deux mesures de probabilité sur $\mathfrak{S}_N$.
$\sum_{\sigma \in \mathfrak{S}_N} \mu(\{\sigma\}) = \sum_{\sigma \in \mathfrak{S}_N} \nu(\{\sigma\}) = 1 \Rightarrow \sum_{\sigma \in \mathfrak{S}_N} \mu(\{\sigma\}) - \nu(\{\sigma\}) = 0$

Or, par définition des parties positive et négative :

$$\begin{aligned} \sum_{\sigma \in \mathfrak{S}_N} (\mu(\{\sigma\}) - \nu(\{\sigma\}))^+ &= \sum_{\mu(\{\sigma\}) \geq \nu(\{\sigma\})} (\mu(\{\sigma\}) - \nu(\{\sigma\}))^+ + \sum_{\mu(\{\sigma\}) \leq \nu(\{\sigma\})} (\mu(\{\sigma\}) - \nu(\{\sigma\}))^+ \\ &= \sum_{\mu(\{\sigma\}) \geq \nu(\{\sigma\})} (\mu(\{\sigma\}) - \nu(\{\sigma\})) \end{aligned}$$

et, de même, $\sum_{\sigma \in \mathfrak{S}_N} (\mu(\{\sigma\}) - \nu(\{\sigma\}))^- = - \sum_{\mu(\{\sigma\}) \leq \nu(\{\sigma\})} (\mu(\{\sigma\}) - \nu(\{\sigma\}))$

Donc, $\sum_{\sigma \in \mathfrak{S}_N} (\mu(\{\sigma\}) - \nu(\{\sigma\}))^+ - \sum_{\sigma \in \mathfrak{S}_N} (\mu(\{\sigma\}) - \nu(\{\sigma\}))^- = \sum_{\sigma \in \mathfrak{S}_N} \mu(\{\sigma\}) - \nu(\{\sigma\}) = 0$

Or, $d_V(\mu,\nu) = \max \left[\sum_{\sigma \in \mathfrak{S}_N} (\mu(\{\sigma\}) - \nu(\{\sigma\}))^+, \sum_{\sigma \in \mathfrak{S}_N} (\mu(\{\sigma\}) - \nu(\{\sigma\}))^- \right]$, donc l'égalité précédente donne

$$\sum_{\sigma \in \mathfrak{S}_N} |\mu(\{\sigma\}) - \nu(\{\sigma\})| = \sum_{\sigma \in \mathfrak{S}_N} (\mu(\{\sigma\}) - \nu(\{\sigma\}))^+ + \sum_{\sigma \in \mathfrak{S}_N} (\mu(\{\sigma\}) - \nu(\{\sigma\}))^- = 2 d_V(\mu,\nu)$$

ce qui termine la démonstration. □

Nous avons ici décrit le battage par insertion en terme de convergence d'une suite de distances entre mesures de probabilité.
A présent, pour apporter une réponse quantitative au problème posé, nous pouvons examiner la vitesse de convergence de cette suite.

3 Temps de remontée d'une carte

3.1 Propriétés du temps de remontée

Dans toute la suite, nous noterons, pour toute suite de variables aléatoires $(X_n)_{n\in\mathbb{N}}$ et pour tout $n \in \mathbb{N}$, $X^{(n)} = \sum_{k=1}^{n} X_k$ (somme partielle d'ordre n).

Supposons à présent que la configuration initiale soit $X_0 = Id$ (toutes les cartes sont rangées dans l'ordre), et intéressons-nous au mouvement de la carte C_N.
L'insertion de la première carte à chaque étape étant aléatoire, on peut définir un temps T_1 (éventuellement infini) pendant lequel la carte C_N reste au fond du paquet. Par définition, à $n = T_1$, C_N se trouve en position $N-1$.
Définissons alors par récurrence un temps T_k tel que à $n = T^{(k)}$, la carte C_N passe de la position $N-k+1$ à la position $N-k$. En notant $X_n(N)$ l'image de N par la permutation X_n, i. e. la position à l'instant n de la carte C_N, on peut obtenir la formule de récurrence suivante :

$$\begin{cases} T_0 = 0 \\ \forall\ k \geq 1,\ T_k = \min\{n \mid X_n(N) = N-k\} - T^{(k-1)} \end{cases}$$

Posons enfin $T = T^{(N)}$, le plus petit instant auquel la carte C_N occupe la première place du paquet (remontée complète).

Exemple 2. *Voici une illustration de la définition de temps de remontée.*

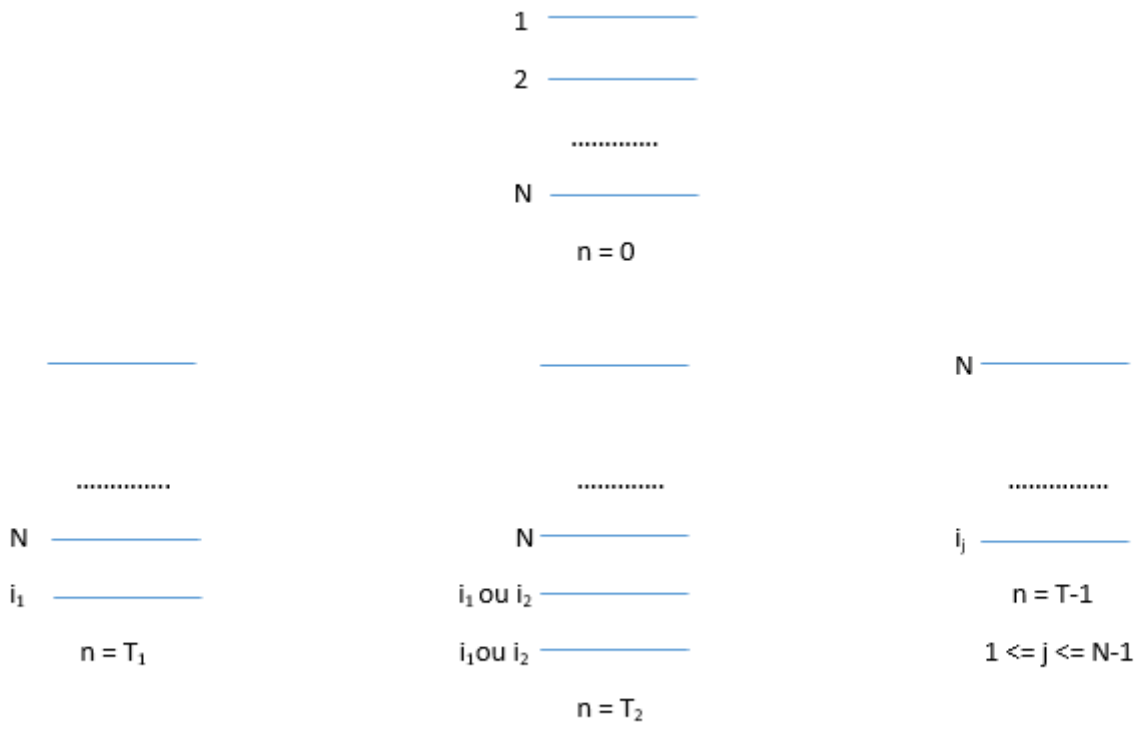


FIGURE 2 – **Remontée de la carte C_N, initialement dernière carte du paquet, et première à l'instant $n = T-1$**

Ces définitions permettent d'établir la proposition suivante.

Proposition 2. *La permutation aléatoire X_T est indépendante de T et de loi uniforme sur $\mathfrak{S}_N$. Plus généralement, pour tout entier $m \geq 0$, la permutation aléatoire X_{T+m} est indépendante de T et de loi uniforme sur $\mathfrak{S}_N$.*

Démonstration. Montrons d'abord que les variables considérées suivent une loi uniforme, et déduisons-en l'indépendance avec T.

— Pour établir que X_T est de loi uniforme sur $\mathfrak{S}_N$, montrons par récurrence sur i la propriété H_i :
Quitte à changer de numérotation des cartes, la loi de probabilité de la position des i cartes en dessous de la carte N au temps $T^{(i)}$ est la loi uniforme sur $\mathfrak{S}_i$.

— **Initialisation :** Au temps T_1, il y a une seule carte en dessous de la carte C_N, donc le résultat est vrai.

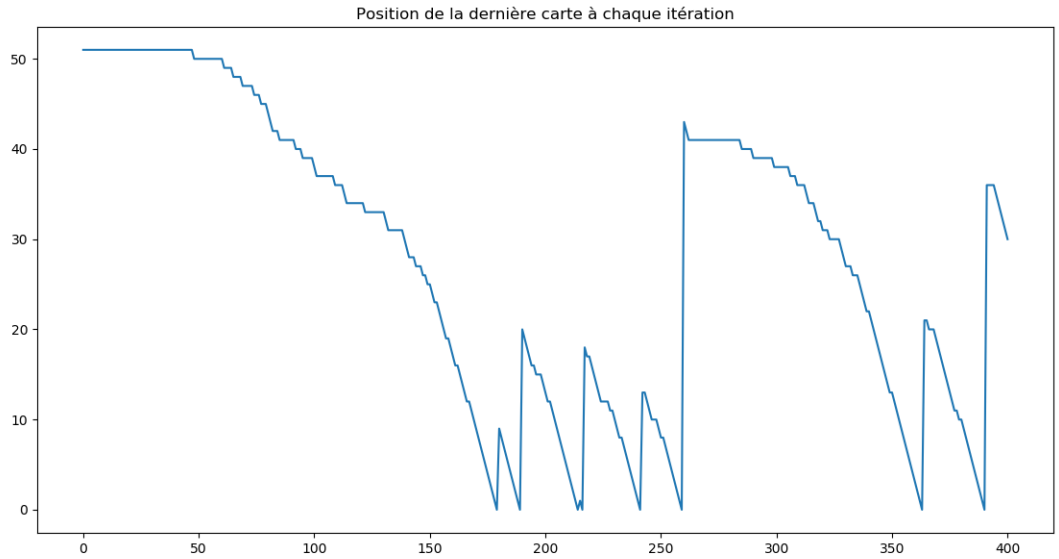
— **Hérédité :** Supposons le résultat vérifié au rang i et considérons le temps $T' = T^{(i+1)} - 1$. L'ordre des cartes aux positions supérieures aux égales à i aux temps $T^{(i)}$ et T' étant identique, l'hypothèse de récurrence montre, sous réserve de changement de numérotation, que toute permutation de $\mathfrak{S}_i$ est atteinte dans l'ensemble de ces i cartes probabilité $\frac{1}{i!}$.
Soit $\tau \in \mathfrak{S}_i$. Au temps $T^{(i+1)}$, les seules permutations pouvant être atteintes par $X_{T^{(i+1)}}$ sont celles dont la restriction à $\mathfrak{S}_i$ est égale à τ après réindexation (i.e. pour σ' une telle permutation, on considère sa restriction à $\mathfrak{S}_i$ et pour tout $k \leq i$, on retranche 1 à $\sigma(k)$ si $\sigma(k) > \sigma(i+1)$, et on impose $\sigma(k) = k$ si $\{k, i+1\}$ est une orbite de σ). Il y a $i+1$ telles permutations, suivant le choix de l'image de $i+1$, qui a pour loi de probabilité la probabilité uniforme sur $[|1, i+1|]$.
La loi uniforme sur $\mathfrak{S}_i$ permet alors d'atteindre $i! \times (i+1) = (i+1)!$ permutations, soit $\mathfrak{S}_{i+1}$ tout entier, avec la loi uniforme sur $\mathfrak{S}_{i+1}$.

— **Conclusion :** X_T est de loi uniforme sur $\mathfrak{S}_N$.

— Pour tous $k,\ m \in \mathbb{N}$ et tout $\sigma \in \mathfrak{S}_N$, $\mathbb{P}(X_{T+m} = \sigma | T = k) = \mathbb{P}(X_{T+m} = \sigma)$, donc les variables X_{T+m} et T sont indépendantes. Soit $m > 0$. $X_{T+m} = X_T({}^tP)^m = X_T$, car X_T suit la loi uniforme, probabilité invariante par la matrice de transition P d'après le théorème 1. Donc, X_{T+m} a même loi que X_T : elle est de loi uniforme sur $\mathfrak{S}_N$ et indépendante de T.

□

Enfin, une première simulation informatique nous renseigne sur le comportement de la position de la carte C_N au cours du temps, pour $N = 52$.



3.2 Estimations quantitatives

L'inconvénient de cette description est de ne pas permettre d'arrêter le battage au temps T sans connaître à chaque instant la position de la carte C_N. Il nous faut alors procéder par estimations de la distance de variation entre μ_n et π.
Il est possible de montrer le résultat suivant.

Proposition 3.

$$\forall\ n \geq 0,\ d_V(\mu_n, \pi) \leq \mathbb{P}(T > n)$$

La démonstration consiste simplement à majorer, pour tout $A \in P(\mathfrak{S}_N)$, $\mu_n(A)$ par $\pi(A) + \mathbb{P}(T > n)$, en utilisant le résultat d'indépendance fourni par la proposition 2.
Ce résultat est très utile car il permet de ramener l'étude d'une suite de réels peu aisément quantifiables à l'étude d'une loi de probabilité (celle de T) construite pour être une loi classique.
En effet, pour tout $i \in \{1, \ldots, N-1\}$, la variable T_i est définie comme le premier instant de succès d'une expérience de Bernoulli, si on définit le succès comme étant l'événement "passer de la position $N-i-1$ à la position i". Elle suit alors une loi géométrique de paramètre $\dfrac{i}{N}$.
Ces renseignements permettent d'aboutir à l'estimation suivante.

Proposition 4. *Pour tout $N \in \mathbb{N} - \{0\}$, pour tout $c > 0$,*

$$\mathbb{P}(T > N \ln N + cN) \leq \frac{K}{c^2},\ \text{avec}\ K = \sum_{k=1}^{+\infty} \frac{1}{k^2}$$

Démonstration. Soient $N \in \mathbb{N} - \{0\}$ et $c > 0$. D'après l'inégalité de Bienaymé-Tchebychev,

$$\mathbb{P}(|T - \mathbb{E}(T)| > cN) \leq \frac{V(T)}{N^2 c^2}$$

— Calculons tout d'abord l'espérance de T.
T est une somme de N variables aléatoires indépendantes T_i de loi géométrique de paramètre $\dfrac{i}{N}$, donc d'espérance $\mathbb{E}(T_i) = \dfrac{N}{i}$. Alors

$$\mathbb{E}(T) = \sum_{i=1}^{N} \mathbb{E}(T_i) = N \sum_{i=1}^{N} \frac{1}{i} \underset{+\infty}{=} N(\ln N + \gamma) + \circ(N) \underset{+\infty}{\simeq} N \ln N$$

avec γ la constante d'Euler.
Le calcul montre l'inégalité $\mathbb{E}(T) < N \ln N$ pour tout $N \in \mathbb{N} - \{0\}$ (somme de termes positifs). Alors, par positivité de T,

$$T - \mathbb{E}(T) > T - N \ln N \ \Rightarrow \mathbb{P}(T - N \ln N > cN) \leq \mathbb{P}(T - \mathbb{E}(T) > cN) \leq \mathbb{P}(|T - \mathbb{E}(T)| > cN)$$

.
— Calculons à présent la variance de T.
T est une somme de N variables aléatoires indépendantes T_i de loi géométrique de paramètre $\dfrac{i}{N}$, donc de variance $V(T_i) = \dfrac{1 - \frac{i}{N}}{\left(\frac{i}{N}\right)^2} = N\dfrac{N-i}{i^2}$. Alors

$$\frac{V(T)}{N^2} = \sum_{i=1}^{N} \frac{V(T_i)}{N^2} = \sum_{i=1}^{N} \frac{N-i}{Ni^2} = \sum_{i=1}^{N} \frac{1}{i^2} - \frac{1}{N} \sum_{i=1}^{N} \frac{1}{i} \underset{+\infty}{=} \sum_{i=1}^{N} \frac{1}{i^2} - \frac{\ln N}{N} - \frac{\gamma}{N} + \circ(\frac{1}{N})$$

En posant $K = \displaystyle\sum_{i=1}^{+\infty} \frac{1}{i^2}$ la somme d'une série convergente, on obtient $\dfrac{V(T)}{N^2} \underset{+\infty}{\simeq} K$
L'inégalité précédente montre alors

$$\mathbb{P}(T - N \ln N > cN) \leq \frac{K}{c^2}$$

□

Nous pouvons alors essayer de vérifier par une simulation informatique cette estimation quantitative. L'axe des ordonnées est gradué en échelle semi-logarithmique.

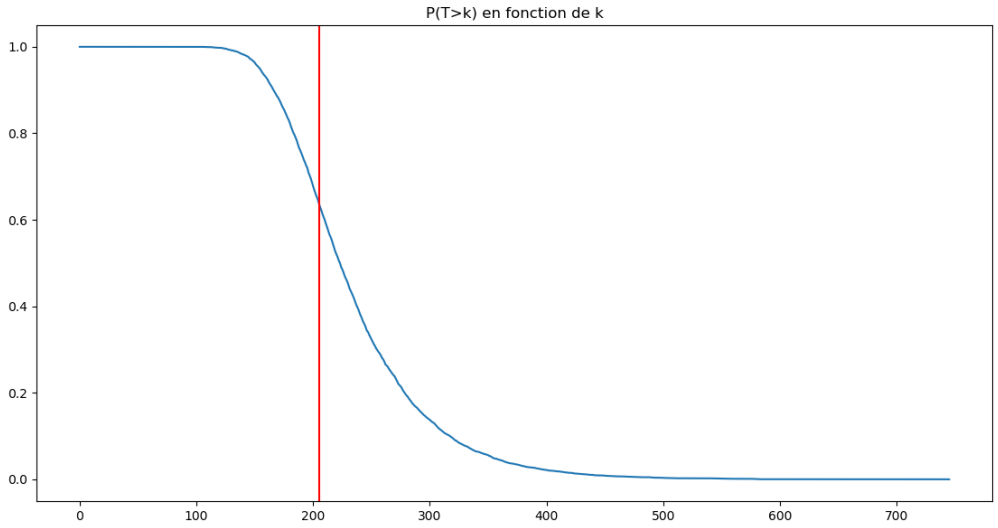


FIGURE 3 – **Représentation de $\mathbb{P}(T > n)$ en fonction de n pour $N = 52$.**

Nous constatons que la décroissance semble être plus rapide qu'une décroissance géométrique pour n assez grand En particulier, on observe un saut autour de $N \ln N$ (qui vaut environ 205 pour $N = 52$).
Le résultat suivant améliore l'estimation précédente pour rendre compte d'un comportement asymptotique exponentiel de la distance en variation.

Théorème 2. *Pour tout réel positif c,*

$$d_V(\mu_{N \ln N + cN}, \pi) \leq e^{-c}$$

Pour démontrer ce théorème, il est nécessaire d'interpréter différemment la loi de T.

3.3 Problème du collectionneur

Etablissons un lien avec un autre problème probabiliste, appelé "problème du collectionneur", en justifiant le rapprochement avec le problème qui nous intéresse. Imaginons un philatéliste qui reçoit une lettre affranchie avec un timbre choisi au hasard uniformément parmi les N timbres en vigueur.
On remarque que le fait d'obtenir les premiers timbres est très rapide, ce qui est bien plus difficile pour les derniers timbres. Or, dans la modélisation précédente, c'est l'inverse qui se produit : la carte C_N remonte d'autant plus facilement que sa position est élevée.
A présent, traduisons mathématiquement cette analogie. Définissons la suite de variables aléatoires $(S_i)_{i \in [|1,N|]}$ qui modélise la suite des premiers instants auxquels le collectionneur complète sa collection par un nouveau timbre. Notons enfin $S = S^{(N)}$.
La proposition suivante établit le lien avec le problème qui nous occupe, c'est-à-dire la loi de T.

Proposition 5. *Les variables S_i sont indépendantes et pour tout $i \in [|1, n|]$, S_i suit la loi géométrique de paramètre $\dfrac{N-i+1}{N}$. Ainsi, S a la même loi que T.*

Démonstration. Soit $i \in [|1, n|]$. Entre les instants S_{i-1} et S_i, sur le total de N timbres, $N-i+1$ timbres ne sont pas déjà présents dans la collection. De plus, chaque timbre a la même probabilité d'être reçu. A chaque réception de timbre, la probabilité que le timbre reçu soit nouveau est alors égal à $\frac{N-i+1}{N}$.
Ainsi, S_i est le premier instant de réception d'un nouveau timbre dans une succession d'épreuves de Bernoulli indépendantes de paramètre $\frac{N-i+1}{N}$, ce qui caractérise la loi de S_i : elle est géométrique de paramètre $\frac{N-i+1}{N}$.
Ce raisonnement prouve au passage que $(S_i)_{1\leq i\leq N}$ est une suite de variables aléatoires mutuellement indépendantes, car pour tout $1 \leq i \leq N$, le temps S_i ne dépende que du nombre de timbres reçus après le temps $S^{(i-1)}$.
De plus, pour tout $1 \leq k \leq N-1$, T_k et S_{N_k} ont même loi, ainsi que T_N et S_1 qui suivent la loi constante égale à 1.
Alors, S à la même loi que T. □

A présent, remarquons que cette description de la loi de T permet de montrer plus aisément le résultat suivant.

Proposition 6. *Pour tout entier $m \geq 1$,*

$$\mathbb{P}(T > m) = \mathbb{P}(S > m) \leq N\left(1 - \frac{1}{N}\right)^m \leq Ne^{-\frac{m}{N}}$$

Démonstration. Soit $m \geq 1$. Numérotons les timbres de 1 à N, et, pour j entre 1 et N, appelons B_j^m l'événement suivant : "le jour m, le collectionneur n'a toujours pas reçu de lettre affranchie avec le timbre numéro j". On a, par sous-additivité :

$$\mathbb{P}(S > m) = \mathbb{P}\left(\bigcup_{j=1}^{N} B_j^m\right) \leq \sum_{j=1}^{N} \mathbb{P}(B_j^m)$$

Les tirages étant indépendants, la famille d'événements $(B_j^p)_{1\leq p\leq m}$ est une famille d'événements mutuellement indépendants. De plus, la probabilité que le timbre reçu ne soit pas le timbre j est égale à $1 - \frac{1}{N}$ (loi uniforme).
On en déduit $\mathbb{P}(B_j) = (1 - \frac{1}{N})^m$, pour tout j. Ceci démontre la première inégalité.
La seconde inégalité est une application de la convexité de l'exponentielle. □

Convenons que, pour tout entier $m \geq 1$ et tout $m \leq p < m-1$, $\mathbb{P}(T > p) = \mathbb{P}(T > m)$. La proposition 3 permet de majorer la distance en variation par une probabilité $\mathbb{P}(T > m)$. Le résultat précédent appliqué à $p = N \ln N + cN$ permet de déduire immédiatement le théorème 2, ce qui était notre objectif.
Néanmoins, une estimation du nombre de battages nécessaires pour parvenir à un jeu convenablement mélangé n'est pas fournie par cette étude.

4 Minoration du nombre de battages

La simulation numérique montre que le temps de remontée de la carte C_N a une forte probabilité d'être proche de $N \ln N$. En particulier, $\mathbb{P}(T > N)$ reste proche de 1 pour de petites valeurs de n. Formalisons ce phénomène.

Le but est de montrer que, pour un temps n légèrement inférieur à $N \ln N$, le paquet est mal mélangé. Considérons alors une partie A_j de $\mathfrak{S}_N$ telle que la probabilité de l'événement $X_n \in A_j$ soit minorée par une probabilité proche de 1.
Pour j entier entre 2 et N, définissons

$$A_j = \{\sigma \in \mathfrak{S}_N \mid \forall\, k \in [|1, j-1|],\ \sigma(N-j+k) < \sigma(N-j+k+1)\}$$

A_j est alors l'ensemble des configurations du paquet qui n'ont pas mélangé les j dernières cartes initiales. Par définition de la distance en variation, on obtient les inégalités :

$$d_V(\mu_n, \pi) \geq |\mathbb{P}(X_n \in A_j) - \pi(A_j)| \geq \mathbb{P}(X_n \in A_j) - \frac{1}{j!}$$

Elles traduisent le fait que si $\mathbb{P}(X_n \in A_j)$ est proche de 1, la distance de variation est trop importante avec la loi uniforme pour pouvoir décider que le paquet est convenablement mélangé, et ce d'autant plus que j est grand.

Pour minorer $\mathbb{P}(X_n \in A_j)$, observons le mouvement de la carte C_{N-j+1} au cours du temps, de la même manière que dans la section précédente pour la carte C_N. On peut alors minorer $\sup_{n\in\mathbb{N}}\{X_n \in A_j\}$ par R_j défini comme le premier instant auquel la carte C_{N-j+1} se retrouve au sommet du paquet, c'est-à-dire $R_j = \min_{n\in\mathbb{N}}\{X_n(N-j+1) = 1\}$.
Nous avons effectivement généralisé l'étude précédente, puisque, dans la section précédente, $T = R_1 + 1$, avec les nouvelles définitions introduites dans cette partie.
Ce raisonnement général permet de montrer la proposition suivante.

Proposition 7. *R_j a la même loi que $\sum_{k=j}^{N-1} T_k$. De plus, pour tous n et j, $\mathbb{P}(X_n \in A_j) \geq \mathbb{P}(R_j \geq n)$.*

Démonstration. La définition de R_j conduit à l'équivalence

$$X_0(N-j+1) = N-j+1 \iff X_{T^{(j-1)}}(N) = N-j+1$$

Le temps que met la carte C_N à remonter le paquet est alors le temps $R_j = T - T^{(j-1)}$, donc les deux variables ont même loi.
Supposons $n \leq R_j$. La carte C_j n'est pas en première position, ce qui implique que seules les $N-j$ premières cartes ont pu changer de position. De plus, leur insertion ne modifie pas l'ordre relatif des $N_j + 1$ dernières cartes. Alors, par définition de A_j, $X_n \in A_j$.
Par croissance de la mesure de probabilité, $\mathbb{P}(X_n \in A_j) \geq \mathbb{P}(R_j \geq n)$. □

A présent, nous pouvons établir le théorème suivant.

Théorème 3 (Minoration asymptotique du nombre de battages par $N \ln N$)**.** *Soit $(c_N)_{N\geq 1}$ une suite de réels positifs de limite infinie et telle que, pour tout $N \geq 1$, $C_N N \leq N \ln N$. Alors :*

$$\lim_{N\to+\infty} d_V(\mu_{N\ln N - c_N N}, \pi) = 1$$

Démonstration. Soit j entre 2 et N fixé. La proposition 7 montre que R_j a même loi qu'une somme de variables aléatoires indépendantes dont l'espérance et la variance ont été estimées précédemment. En effet :

$$\mathbb{E}(R_j) = E(T) - N\sum_{k=1}^{j-1}\frac{1}{k} \underset{+\infty}{\simeq} E(T) \underset{+\infty}{\simeq} N\ln N$$

et

$$V(R_j) \underset{+\infty}{\simeq} N^2 \sum_{k=j}^{N}\frac{1}{k^2} \underset{+\infty}{\simeq} C(j)N^2$$

avec $C(j)$ une constante ne dépendant que de j.
L'inégalité de Bienaymé-Tchebychev donne alors :

$$\mathbb{P}(R_j \leq N \ln N - c_N N) \leq \mathbb{P}(|R_j - N \ln N| \geq c_N N) \leq \frac{C(j)N^2}{c_N^2 N^2} = \frac{C(j)}{c_N^2}$$

par le même argument de positivité que dans la section 3.
Par passage à la limite lorsque N tend vers $+\infty$:

$$\lim_{N \to +\infty} \mathbb{P}(R_j \geq N \ln N - c_N N) = 1$$

La majoration de la proposition 7 montre alors :

$$\lim_{N \to +\infty} \mathbb{P}(X_{N \ln N - c_N N} \in A_j) = 1$$

Il s'ensuit, en utilisant la minoration sur la distance en variation obtenue en début de section :

$$\lim_{N \to +\infty} d_V(\mu_{N \ln N - c_N N}, \pi) = 1 - \frac{1}{j!}$$

Dans cette limite $N \to +\infty$, on peut également faire tendre j vers $+\infty$ puisque l'inégalité est prouvée pour tout j entre 2 et N. Elle décrit alors ce qui se passe pour le mouvement d'une carte qui se situe "en bas" du paquet initialement. On obtient alors le résultat voulu. □

Exemple 3 (Nombre de battages pour un jeu de 52 cartes)**.** *En estimant qu'une distance en variation de* $0,2$ *est acceptable, pour* $N = 52$*, on obtient avec la section 3 qu'il suffit de mélanger le jeu* 290 ***fois****.*
De plus, le travail de cette section apporte un renseignement supplémentaire : il est nécessaire d'effectuer un nombre de battages proche pour obtenir un jeu convenablement mélangé : en d'autres termes, il n'y a pas d'amélioration substantielle possible du résultat.

Le nombre très élevé de battages obtenu nous amène alors à réfléchir à d'autres modèles de battages et de modélisation d'un jeu mélangé de manière satisfaisante.

5 Vers d'autres modèles de mélanges

5.1 Autres modèles de battage par insertion

On s'intéresse à différentes manières possibles d'améliorer le battage par insertion.

Tout d'abord, un des inconvénients du battage par insertion est qu'il autorise une carte à ne pas être mélangée, en étant repositionnée à la première place du paquet à l'étape où elle doit être insérée. On peut alors penser, à l'étape n, à imposer l'insertion de la première carte du paquet à une position choisie uniformément entre $N-n+1$ et N. Ainsi, on insère toujours en dessous de la carte C_N, et on force la carte C_N à se positionner en tête du paquet mélangé en seulement N battages. On est ainsi sûr d'obtenir un mélange parfait pour la distance en variation avec la loi uniforme au bout d'un nombre fini d'étapes N. Ceci est plus intéressant que l'estimation $N\ln N$ précédente, asymptotique de surcroît.

Cependant, un tel battage implique la connaissance de la position de la carte C_N à la fin du mélange. Or une contrainte naturelle que l'on souhaite imposer à un mélange convenable est l'absence de certitude sur la position de chaque carte.
Pour conjuguer les avantages des deux variantes, on peut envisager le battage suivant.

Définition 11 (**Battage par insertion, variante**). *Soient $N \geq 3$ cartes.*

- *On effectue tout d'abord $\lfloor \frac{N}{2} \rfloor$ battages en insérant, à l'étape k, la première carte du paquet au hasard, uniformément entre les positions $N-k+1$ et N.*
- *On effectue ensuite $2\lfloor \frac{N}{2} \rfloor$ battages de la manière suivante : pour tout k entre 1 et $\lfloor \frac{N}{2} \rfloor$, on insère, à l'étape $2k-1$, la première carte du paquet au hasard, uniformément entre les positions $N-k+1$ et N, et on insère, à l'étape $2k$ la première carte du paquet au hasard, uniformément dans le jeu, sans contrainte de position, tel qu'a été présenté le battage par insertion en début de travail.*

Ce modèle de battage présente l'avantage d'obtenir sûrement un mélange convenable du paquet au bout de $3\lfloor \frac{N}{2} \rfloor$ opérations, tout en conservant inconnue la position de chaque carte, en particulier celle de la carte C_N. Plus précisément, le raisonnement par récurrence établi en section 3 sur la remontée de la carte C_N permet de démontrer le résultat suivant.

Proposition 8 (**Terminaison du battage**). *Il existe k entre $\lfloor \frac{N}{2} \rfloor$ et $3\lfloor \frac{N}{2} \rfloor - 1$ tel que, à l'étape k, la carte C_N se retrouve en première position.*

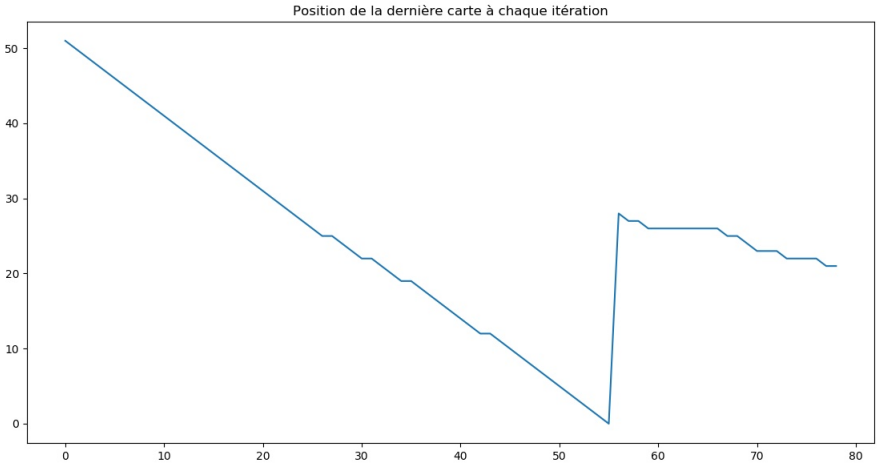
Démonstration. Il existe deux types d'insertion.

- Si, à l'étape k, on insère la première carte entre les positions $N-k+1$ et N, la position de la carte C_N est incrémentée de 1.
- Si on insère la première carte à une position quelconque, la position de la carte C_N peut être incrémentée de 1 ou rester inchangée.

Il y a donc au moins $2\lfloor \frac{N}{2} \rfloor$ étapes pour lesquelles la carte C_N "remonte", ce qui prouve le résultat. En effet, à l'étape $3\lfloor \frac{N}{2} \rfloor$, la première carte est toujours insérée à une place quelconque dans le jeu.

□

La simulation suivante illustre cette proposition dans le cas $N = 52$ (ici, k est environ égal à 55).



Néanmoins, les différentes méthodes employées jusqu'ici ne sont ni efficaces ni réalistes. De plus, les améliorations effectuées ne permettent pas de réduire le nombre de battages requis de telle sorte à pouvoir appliquer en pratique ces modèles pour trier un jeu de cartes. Ainsi, pour un jeu de 52 cartes, il faudrait 78 battages pour aboutir à un jeu mélangé.
Ces nouveaux battages introduisent même la contrainte supplémentaire de compter les positions acceptables d'insertion à chaque étape.
Par ailleurs, les précédents modèles ne font pas appel à une technique de tri couramment utilisée, qui est la coupe du jeu. Elle est même effectuée le plus souvent avant le mélange.
Tentons alors d'introduire une formalisation de la coupe d'un jeu de cartes.

5.2 Introduction de la coupe et battage dit "américain"

Un battage est composé d'une coupe et d'un mélange [3]. Il est toujours associé à une permutation de $\mathfrak{S}_N$. On se place dans le cas où la permutation représentant à la configuration initiale est l'identité.
Nous noterons m le nombre de battages, X_m les variables aléatoires modélisant la permutation associée au mélange au battage m et μ_m leur loi de probabilité.

Définition 12 (**Battage américain**). *Soient N cartes.*

- *Coupons le paquet ainsi formé en deux paquets numérotés 1 et 2, de J et $N-J$ cartes respectivement, avec J choisi entre 0 et N en suivant une loi binomiale* $\mathrm{B}\left(N, \frac{1}{2}\right)$.
- *Après avoir coupé, on mélange le jeu de cartes de la manière suivante.*
  Soient Y_1 et Y_2 les variables aléatoires comptant le nombre de cartes dans les paquets 1 et 2 respectivement à chaque étape. On reconstitue le jeu en choisissant, à chaque étape, la carte placée en dernière position dans le paquet 1 avec probabilité $\frac{Y_1}{Y_1 Y_2}$, et dans le paquet 2 sinon.

Pour la coupe, la loi binomiale permet d'obtenir :

$$\forall\, k \in [|0, N|],\ \mathbb{P}(j = k) = \frac{1}{2^N} \frac{N!}{k!(N-k)!}$$

En effet, la paramètre binomial $p = \frac{1}{2}$ respecte l'intuition du joueur de cartes qui coupe plus naturellement "au milieu" du jeu. Par ailleurs, il s'agit de la probabilité d'obtenir une partie à k éléments dans un ensemble à N éléments, si toutes les parties sont équiprobables.
Le paramètre $\frac{Y_1}{Y_1 + Y_2}$ traduit le fait qu'une carte est choisie avec plus grande probabilité dans un paquet plus gros.

L'efficacité de ce modèle de battage par rapport aux précédents réside dans l'exploitation d'une propriété des permutations que nous introduisons ici.

Définition 13 (**Suite montante**). *Soit σ une permutation de $\mathfrak{S}_N$. Une suite montante dans σ est une sous-suite maximale de l'ensemble $\{\sigma(i), i \in [|1, N|]\}$ constituée de nombres consécutifs.*

On dispose d'un algorithme permettant de trouver les suites montantes d'une permutation, donné par la démonstration de la proposition suivante.

Proposition 9. *Toute permutation se décompose de manière unique en une juxtaposition de suites montantes.*

Pour trouver les suites montantes d'une permutation, on procède comme ceci : on commence par repérer la carte 1. Si la carte 2 est avant 1, alors (1) est une suite montante, sinon on cherche 3. Si 3 est avant 2, alors (1,2), est une suite montante, sinon on cherche 4, et ainsi de suite jusqu'à épuisement du paquet.

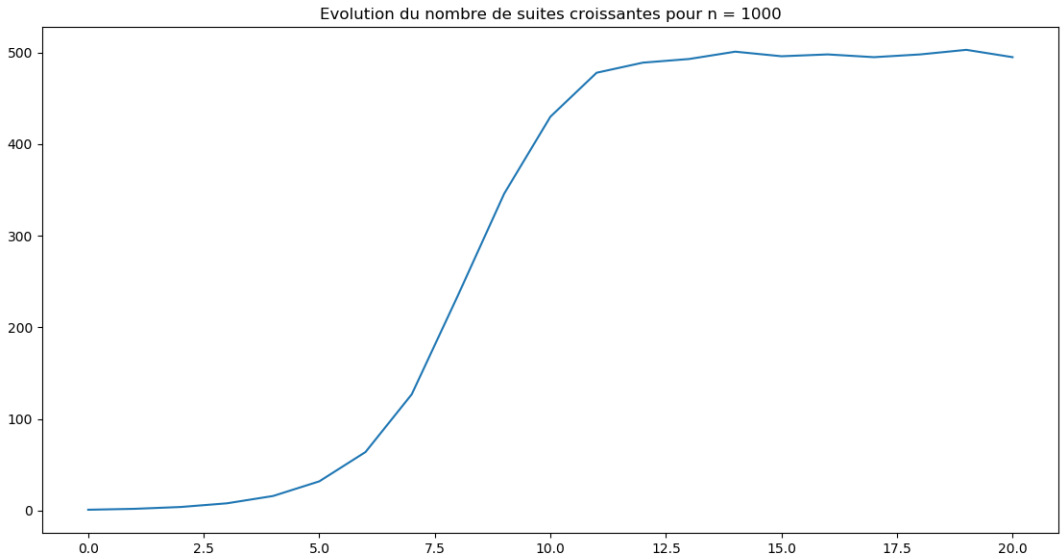
Exemple 4. *La permutation*

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 \\ 7 & 3 & 2 & 12 & 1 & 4 & 5 & 8 & 11 & 6 & 10 & 9 \end{pmatrix}$$

possède les suites montantes (1), (2), (3 4 5 6), (7 8 9), (10), (11) *et* (12).

Exemple 5. *Lors d'une coupe d'un paquet de cartes dans la configuration identité au niveau de la carte C_k, deux suites montantes sont crées si $2 \leq k \leq N-1$: $(1\ 2\ \ldots\ k)$ et $(k\ k+1\ \ldots\ N)$. Si $k = 1$ ou $k = N-1$, il n'y a qu'une seule suite montante.*

Ici, nous voyons apparaître l'idée que le nombre de suites montantes croît presque sûrement avec le nombre de coupes. On considère empiriquement que le paquet est bien mélangé si le nombre de suites montantes de la permutation associée est proche de $\lfloor \frac{N}{2} \rfloor$. Ce raisonnement grossier et très qualitatif est néanmoins corroboré par la simulation (nous avons choisi $N = 1000$).
De plus, nous voyons se dessiner un ordre de grandeur du nombre de battages nécessaire (aux alentours de 10 battages). Il donne l'intuition que ce modèle est bien plus performant que les modèles précédents.



Un lemme combinatoire permet de démontrer le théorème principal sur la probabilité d'une permutation.

Lemme 2. *Le nombre d'arrangements de q éléments dans p compartiments est égal à C_{p+q-1}^{p-1}.*

On s'autorise, dans le lemme précédent, à placer plusieurs objets dans un même compartiment. Ceci correspond à la possibilité de pouvoir couper plusieurs fois à la même position et à obtenir un paquet vide, dans la démonstration du théorème suivant.

Théorème 4 (Probabilité d'une configuration). *Soit $m \geq 1$ et $\sigma \in \mathfrak{S}_N$. Notons r le nombre de suites montantes dans σ.*
La probabilité que le paquet soit dans la configuration σ après m battages est égale à :

$$\mathbb{P}(X_m = \sigma) = \frac{1}{2^{Nm}} \mathrm{C}_{2^m+N-r}^{N}$$

Démonstration. Distinguons deux cas.

— **Cas $m = 1$:** L'exemple 4 montre que les seules valeurs possibles de r sont $r = 1$ et $r = 2$.
Soit j la position de coupe : le paquet est divisé en deux petits paquets, numérotés 1 et 2, de taille respective j et $N - j$.
On choisit ensuite successivement chaque carte dans l'un des deux paquets après n opérations. Notons, pour k entre 1 et N et pour i_k égal à 1 ou 2, x_k le nombre de cartes restant dans le i_k-ième paquet à la k-ième étape. La probabilité de ce mélange associé à σ' est égale à :

$$\mathbb{P}(X_m = \sigma') = \prod_{k=1}^{N} \frac{x_k}{N-k-1}$$

En remarquant ensuite que pour tout mélange, il existe une étape telle qu'il reste k cartes dans le paquet 1, et une étape (éventuellement distincte) telle qu'il reste $N - k$ cartes dans le paquet 2, on en déduit que tous les mélanges σ sont équiprobables, de probabilité :

$$\mathbb{P}(X_m = \sigma) = \frac{1}{2^N}$$

Enfin, observons que, dans ce cas simple, les deux suites montantes obtenues en coupant caractérisent l'entier j si $r = 2$. Si $r = 1$, alors la permutation obtenue est l'identité, et on peut l'obtenir exactement une fois pour chaque valeur de j, pour $0 \leq j \leq N$. Alors, $\mathbb{P}(X_m = Id) = \dfrac{N+1}{2^N} = \dfrac{1}{2^N}\mathrm{C}_{N+2}^{N}$.
Le théorème est donc vérifié.

— **Cas général :** Donnons les grandes idées de la preuve.

— On peut se ramener à une situation plus simple dans laquelle on effectue toutes les coupes avant les mélanges. En effet, une coupe correspond à séparer chaque paquet obtenu en deux petits paquets à une position choisie selon la loi binomiale comme précédemment de sorte à obtenir, après m coupes 2^m paquets (éventuellement vides), séparés par $2^m - 1$ opérations. En notant $(j_k)_{k\in[|1,2^m-1|]}$ la suite des opérations, on montre par récurrence sur m la propriété suivante :

$$\mathbb{P}((j_k)_{k\in[|1,2^m-1|]}) = \frac{n!}{2^{mN} \prod_{i=1}^{2^m} (j_{i-1} - j_i)!}$$

Un raisonnement analogue au cas $m = 1$ permet alors de conclure que toutes les manières de réaliser un mélange de 2^m paquets sont équiprobables, de probabilité $\dfrac{1}{2^{mN}}$

— On dénombre les manières d'obtenir une permutation $\sigma \in \mathfrak{S}_N$. La proposition 9 montre que cela revient à se donner les r suites montantes de σ.
Les $2^m - 1$ positions de chaque coupe d'un paquet sont choisies, et $r - 1$ coupes sont déterminées par les derniers éléments des $r - 1$ premières suites montantes.

Les $2^m - r$ dernières coupes peuvent être réalisées en $N+1$ positions quelconques.
Le lemme 2 montre que le nombre de possibilités est égal à $\mathrm{C}^N_{2^m+N-r}$.

— On conclut en multipliant ce nombre de manières d'obtenir une permutation donnée par le facteur $\dfrac{1}{2^{Nm}}$ trouvé.

□

Terminons l'étude en remarquant que le comportement asymptotique de la probabilité d'obtenir une permutation conduit à une minoration du nombre de battages successifs pour bien mélanger le jeu. Notons que la définition d'une distance pour mesurer effectivement si un mélange est correct est inchangée : nous gardons la même définition de distance en variation.

Théorème 5 (Estimation du nombre de battages américains nécessaires). *Au sens de la distance de variation entre les mesures de probabilité des variables modélisant le battage et la loi uniforme, il faut* $\lceil \frac{3}{2}\log_2(n) \rceil$ *battages pour mélanger un jeu de cartes.*

Démonstration. Là encore, donnons une ébauche de démonstration, en en présentant les idées.
La suite des variables aléatoires X_m converge presque sûrement vers une variable aléatoire suivant la loi uniforme sur $\mathfrak{S}_N$, car, pour tout $\sigma \in \mathfrak{S}_N$ et tout nombre r de suites montantes de σ, on a :

$$\mathbb{P}(X_m = \sigma) = \frac{1}{2^{Nm}}\mathrm{C}^N_{2^m+N-r} \underset{m\to+\infty}{\longrightarrow} \frac{1}{N!}$$

Or, la convergence presque sûre entraîne la convergence en loi.
En notant, pour tout $r \leq N$, $x_{n,r}$ le nombre de permutations contenant r suites montantes, le théorème précédent permet de montrer que, pour tout $m \geq 1$:

$$d_V(\mu_m, \pi) = \frac{1}{2}\sum_{r=1}^{N} |x_{N,r}\frac{1}{2^{Nm}}\mathrm{C}^N_{2^m+N-r} - \frac{1}{N!}|$$

En calculant de manière approchée les nombres $x_{n,r}$, on peut aboutir à une expression intégrale de $d_V(\mu_m, \pi)$ pour les valeurs du type $m = \frac{3}{2}\lceil \log_2(n) \rceil + h$.
Cette expression a l'avantage de fournir des valeurs numériques proches de 1 pour $h >> 0$ et proches de 0 pour $h << 0$. □

La démonstration montre que l'estimation numérique ainsi établie est fiable et optimale (les deux caractères sont donnés chacun par une variation de la quantité h). En particulier, ce résultat est à mettre en relation avec la minoration obtenue pour le battage par insertion simple en section 4, qui est elle aussi optimale.
L'avantage considérable de ce modèle est sa rapidité d'exécution, comme le montre l'application au cas du jeu à 52 cartes, le plus couramment utilisé.

Exemple 6. *Voici les valeurs de la distance en variation entre* μ_m *et* π *pour les neuf premières valeurs de* m *pour* $N = 52$ *:*

$d_V(\mu_1,\pi)$	$d_V(\mu_2,\pi)$	$d_V(\mu_3,\pi)$	$d_V(\mu_4,\pi)$	$d_V(\mu_5,\pi)$	$d_V(\mu_6,\pi)$	$d_V(\mu_7,\pi)$	$d_V(\mu_8,\pi)$	$d_V(\mu_9,\pi)$
1,000	1,000	1,000	1,000	0,924	0614	0,334	0,167	0,085

On a $m = \frac{3}{2}\lceil \log_2(52) \rceil = 9$*, ce que confirment les valeurs numériques : la distance en variation entre* μ_9 *et* π *peut raisonnablement être considérée comme faible.*
Si nous nous en tenons à l'estimation faite pour le battage par insertion, pour laquelle nous avons considéré une distance de variation égale à $0,2$ *comme acceptable, nous pouvons même convenir que* 8 ***battages suffisent****.*

Références

[1] Agrégation externe de mathématiques. Epreuve de modélisation, option a : Probabilités et statistiques.

[2] *Initiation aux Probabilités et aux chaînes de Markov*, chapter 8. 2009.

[3] D. Bayer and P. Diaconis. Trailing the dovetail shuffle to its lair. *Ann. Appl. Prob.*, 1992.