Battages de cartes

CHAUB Thomas, LE SCORNET Pierre, ORSINI Nicolas

16/04/2018

Sommaire

Battage par insertion

Définition : 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 $\{1, \ldots, N\}$ la place à laquelle l'insertion à lieu, indépendamment de l'insertion précédente.

Remarque

Insérer à la k-ième place revient à passer de σ à σ' telle que $\sigma' = (k, k-1, \ldots, 2, 1)\sigma$.

Lien avec les chaînes de Markov

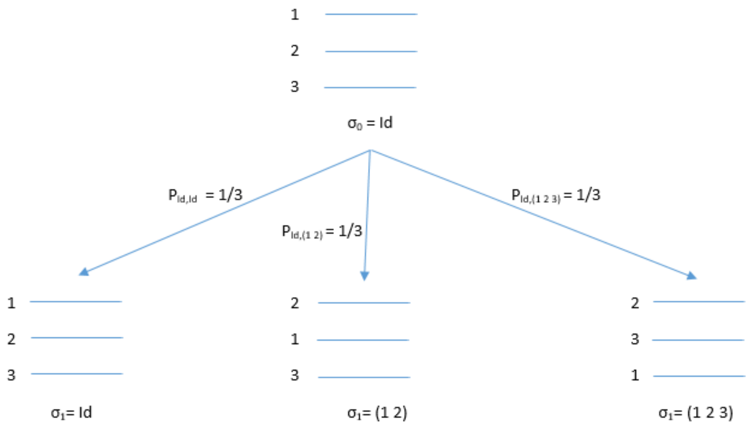
Définition : Chaîne de Markov

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 $i_0,\ i_1, \ldots, i_{n-1},\ i,\ j \in E$:

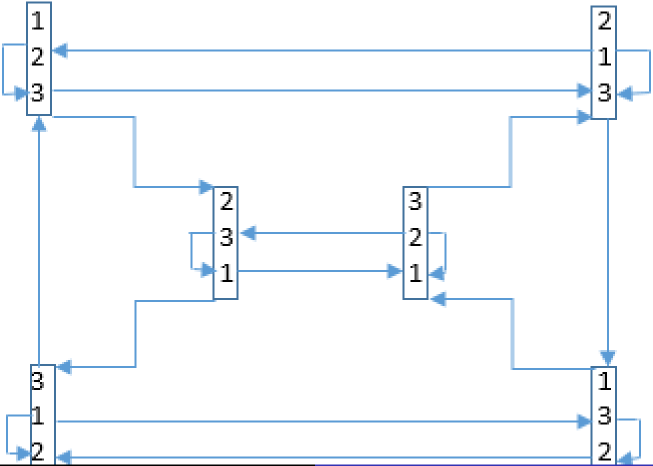
$$\mathbb{P}(X_{n+1} = j \mid X_n = i,\ X_{n-1} = i_{n-1},\ \ldots,\ X_0 = i_0) = \mathbb{P}(X_{n+1} = j \mid X_n = i)$$

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 $(X_n)_{n\in\mathbb{N}}$.

Probabilités de transition



Graphe pour $N = 3$



Convergence en loi de la chaîne de Markov associée

Définition : 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.

Théorème

La chaine de Markov $(X_n)_{n\in\mathbb{N}}$ est irréductible. 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.

Matrice de transition

	Id	(12)	(13)	(23)	(123)	(132)
Id	$1/3$	$1/3$	0	0	$1/3$	0
(12)	$1/3$	$1/3$	0	$1/3$	0	0
(13)	0	$1/3$	$1/3$	0	$1/3$	0
(23)	0	0	$1/3$	$1/3$	0	$1/3$
(123)	0	0	$1/3$	0	$1/3$	$1/3$
(132)	$1/3$	0	0	$1/3$	0	$1/3$

Distance en variation

Définition : 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

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

Proposition

Pour toutes mesures de probabilité μ et ν sur $\mathfrak{S}_N$,

$$d_V(\mu, \nu) = \frac{1}{2} \sum_{\sigma \in \mathfrak{S}_N} |\mu(\{\sigma\}) - \nu(\{\sigma\})|$$

Remontée de la carte C_N

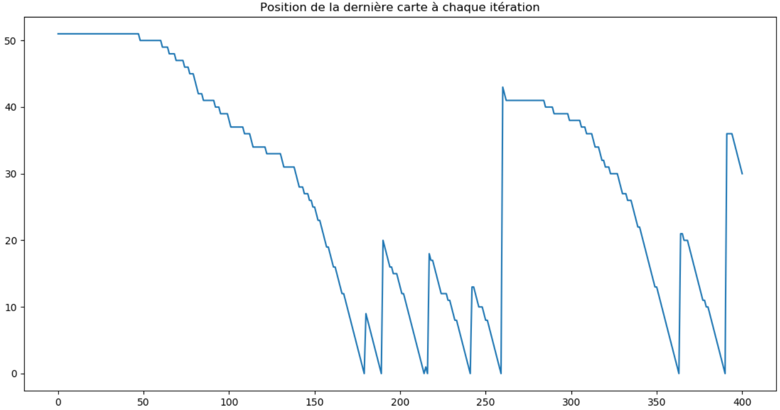
Définition : Temps de remontée de la carte C_N

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

Proposition

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 $\mathfrak{S}_N$.

Estimation qualitative de T



Majoration de $\mathbb{P}(T > n)$

Proposition

Pour tout $n \geq 0$,

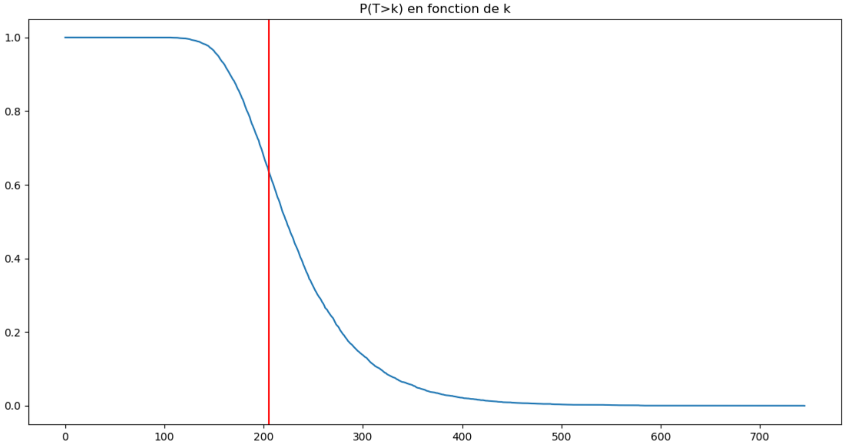
$$d_V(\mu_n, \pi) \leq \mathbb{P}(T > n)$$

Proposition

Pour tout $N \in \mathbb{N} - \{0\}$, pour tout $c > 0$,

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

Représentation de $\mathbb{P}(T > k)$ en fonction de k



Algorithme en $O(N \ln N)$

Théorème : Décroissance exponentielle de $\mathbb{P}(T > m)$

Pour tout entier $m \geq 1$,

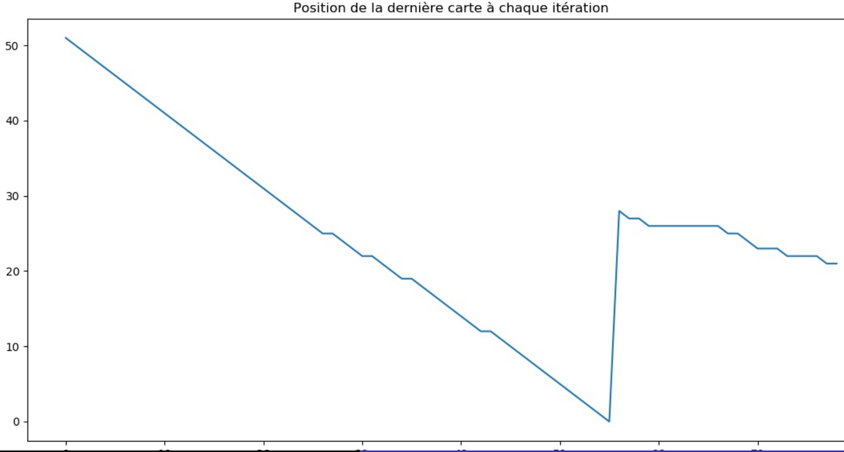
$$\mathbb{P}(T > m) = \mathbb{P}(S > m) \leq N \left(1 - \frac{1}{N}\right)^m \leq N e^{-\frac{m}{N}}$$

Théorème : Minoration asymptotique

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 :

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

Algorithme en $O(N)$



Battage avec coupe et mélange

Définition : Battage américain

- 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 $B\left(N, \frac{1}{2}\right)$.
- 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.

Suites montantes

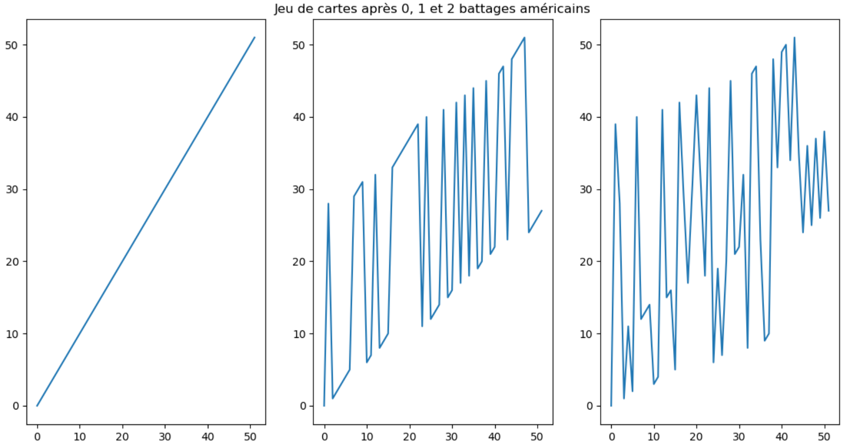
Définition : suite montante

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.

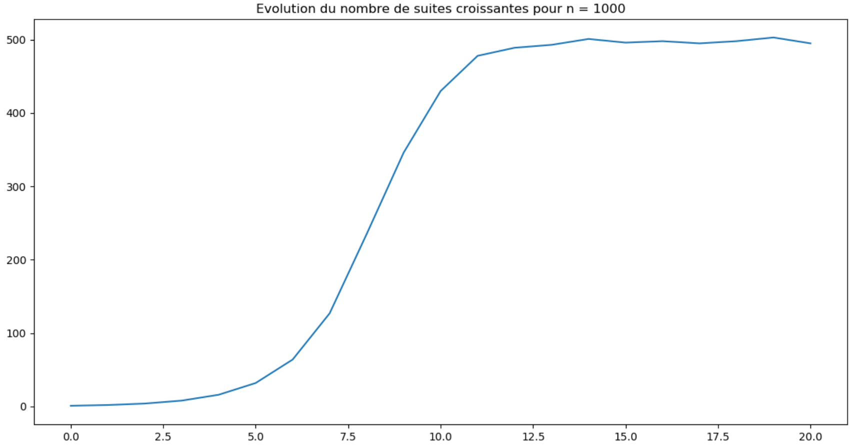
Exemple

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.

Suites montantes après quelques battages



Asymptotique des suites montantes



Algorithme en $O(\log_2(N))$

Théorème : 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 :

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

Théorème : Minoration du nombre de battages

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.

Distances en variation pour $N = 52$

1	2	3	4	5	6	7	8	9
$1,000$	$1,000$	$1,000$	$1,000$	$0,924$	0614	$0,334$	$0,167$	$0,085$

$$\lceil \frac{3}{2} \log_2(52) \rceil = 9$$