Mélange d'un jeu de cartes.

On dispose d'un jeu de r cartes initialement dans une pile. On a numéroté les cartes de 1 à r : la carte numéro 1 est celle initialement en haut du tas. On arrange les cartes en prenant celle du dessus et en l'insérant uniformément en k entre la k-ème et la $k+1$-ème, ce qu'on appelle la k-insertion. La 1-insertion ne change rien et la r-insertion met la carte tout en bas. Une configuration du paquet sera donc décrite par une permutation $\sigma \in \mathfrak{S}_r$ avec $\sigma(i)$ qui note la position de la carte initialement en i.

Faire une k-insertion revient à $\sigma \mapsto (k, k-1, \ldots, 1)\sigma$. En effet toutes les cartes avant la k-ème (en dehors de la premiere) sont remontées de 1 cran.

Les insertions sont i.i.d. choisies uniformement sur $C = \{C_1, \ldots, C_r\}$ avec $C_i = (i, i-1, \ldots, 1)$. On peut decrire l'arrangement des cartes par

$$X_{n+1} = \prod_{i=n+1}^{1} \epsilon_i X_0.$$

1. (a) Montrer que $(X_n), n \geq 0$ est une chaine de Markov sur $\mathfrak{S}(r)$, en donner les probabilités de transition.
   (b) Montrer qu'elle est irréductible, apériodique et récurrente. Identifier sa loi invariante.
   (c) Notons T_i les durées entre deux remontées successives de la carte initialement en bas du tas : posons $Y_n = X_n(r)$.
   $$T_k = \min\{n \geq 1, Y_n = r-k\} - (\sum_{i=0}^{k-1} T_i); T_0 = 0.$$
   Simuler les variables T_i pour $r = 3, 4, 5, 10, 20, 52$, puis $T = T_1 + \ldots + T_{r-1}$. On tracera les histogrammes correspondants.
   (d) Quelle est la position de la carte initialement au bas du tas aux temps $T_1, T_1 + T_2, T_1 + T_2 + T_3, \ldots, T_1 + \cdots + T_{r-1}$?
   Calculer $\mathbb{E}[T]$ et vérifier le calcul grâce à la simulation de la question précédente.
   (e) Vérifier par simulation que pour tout $n \geq 0$, X_{T+1+n} suit la loi uniforme sur $\mathfrak{S}_r$.
2. Dans cette partie $r = 3, 4, 5$.
   On note P_n la loi de X_n, et π la loi invariante trouvée en 1.b. Représenter, grâce à une simulation, le graphe de
   $$n \to \frac{1}{2} \sum_{\sigma \in \mathfrak{S}_r} |P_n(\sigma) - \pi(\sigma)|.$$
   Que remarque-t-on ?
3. Suivons à présent les positions successives de l'une des cartes de notre paquet au cours de notre mélange. On note $(Y_n, n \geq 0)$ la suite de ces positions (de sorte que si $Y_0 = i$, on suit la carte qui se trouve initialement en i-ème position).

(a) Montrer que $(Y_n), n \geq 0$ est une chaine de Markov sur $\{1, ..., r\}$, en donner les probabilités de transition.

(b) Montrer qu'elle est irréductible, apériodique et récurrente. Identifier sa loi invariante.

(c) On note $Q_n^{(i)}$ la loi de Y_n lorsque $Y_0 = i$, et $\tilde{\pi}$ la loi invariante. Représenter, grâce à une simulation, les graphes de

$$n \to \frac{1}{2} \sum_{k=1}^{r} |Q_n^{(i)}(k) - \tilde{\pi}(k)|, \quad \text{pour } i = 1, ..., r.$$

Que remarque-t-on ?

(d) Au regard des simulations de la question précédente, au bout d'environ combien d'insertions successives semble-t-on pouvoir affirmer qu'un jeu de 52 cartes est bien mélangé ? Quel lien peut-on faire avec les résultats des questions 1.(c.d.e) ?

Autre méthode de mélange

Comme en IV, on va considérer un modèle de battage de cartes, mais celui-ci est un peu plus proche de la réalité. A nouveau on considère un jeu de r cartes.

Une étape de notre mélange consiste en la démarche suivante :

- On choisit $M \sim \text{Bin}(r, 1/2)$, et on coupe le jeu au niveau de la M-ème carte (i.e. on sépare les M premières cartes des $r - M$ dernières).
- On "re-mélange" nos 2 paquets (on préserve, bien sûr l'ordre relatif dans chacun des 2 sous-paquets) en choisissant uniformément parmi les $\binom{r}{M}$ façons de le faire.

On note $(X_n, n \geq 0)$ la suite de permutations de $\{1, ..., r\}$ obtenue.

Une sous-suite croissante de longueur n d'une permutation $\sigma \in \mathfrak{S}_r$ est un ensemble d'indices $i_1 < ... < i_n$ tels que

$$\sigma(i_1) < \sigma(i_2) < \cdots < \sigma(i_n).$$

Une *plus longue sous-suite croissante* (p.l.s.s.c.) de longueur n est un tel ensemble d'indices maximal au sens de l'inclusion (i.e. si $j \in (i_k, i_{k+1})$, $\sigma(j) \notin (\sigma(i_k), \sigma(i_{k+1}))$).

Par exemple la permutation

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

possède, par exemple, les trois p.l.s.s.c. suivantes :

$$\begin{aligned} &\sigma(1) = 4 < \sigma(2) = 7 < \sigma(4) = 8 < \sigma(5) = 10, \\ &\sigma(3) = 1 < \sigma(6) = 2 < \sigma(8) = 3 < \sigma(10) = 5 \\ &\sigma(7) = 6 < \sigma(9) = 9. \end{aligned}$$

1. (a) Montrer que $(X_n), n \geq 0$ est une chaine de Markov. Montrer que son noyau de transition vérifie $P(\sigma_1, \sigma_2\sigma_1) = P(Id, \sigma_2)$ pour tous $\sigma_1, \sigma_2 \in \mathfrak{S}_r$, et que

$$P(Id, \sigma) = \begin{cases} \frac{r+1}{2^r} \text{ si } \sigma = Id, \\ \frac{1}{2^n} \text{ si } \sigma \text{ possède exactement 2 p.l.s.s.c.,} \\ 0 \text{ sinon.} \end{cases}$$

   (b) Montrer que (X_n) est irréductible, apériodique et récurrente. Identifier sa loi invariante.

   (c) Ecrire un programme qui simule ce mélange de cartes.

2. Soit la procédure suivante : à chaque carte du jeu on attribue l'étiquette $\xi_i \sim \text{Ber}(1/2)$; on replace alors les cartes étiquetées 0 en haut du paquet (en respectant leur ordre relatif).

   On note $(Y_n, n \geq 0)$ la suite de permutations de $\{1, ..., r\}$ obtenue.

(a) Montrer que $(Y_n), n \geq 0$ est une chaine de Markov. Montrer que son noyau de transition Q vérifie, pour tous $\sigma_1, \sigma_2 \in \mathfrak{S}_r$,

$$Q(\sigma_2, \sigma_1) = P(\sigma_1, \sigma_2).$$

Expliquer le lien entre les chaînes X et Y.

(b) Ecrire un programme qui simule la chaîne Y.

3. (a) Dans cette partie uniquement on ne considérera que $r = 3, 4, 5$. On note P_n la loi de X_n, et π la mesure invariante trouvée en 1.b. Représenter le graphe de

$$n \to \frac{1}{2} \sum_{\sigma \in \mathfrak{S}_r} |P_n(\sigma) - \pi(\sigma)|.$$

Que remarque-t-on ?

(b) Même question avec Q_n, loi de Y_n.

4. On admet que

$$\sum_{\sigma \in \mathfrak{S}_r} |P_n(\sigma) - \pi(\sigma)| = \frac{1}{2} \sum_{k=1}^{r} E(r,k) \left| \frac{\binom{2^n + r - k}{k}}{2^{nr}} - \frac{1}{r!} \right|,$$

avec

$$E(1,1) = 1, \quad E(1,k) = 0 \text{ pour tout } k \neq 1,$$
$$E(r,k) = (r - k + 1)E(r-1, k-1) + kE(r-1, k) \text{ pour tous } r \geq 2,\ k \geq 1.$$

(a) Ecrire un programme qui calcule les valeurs des $E(r,k)$, pour $r = 10, 52, 100$, et $k \in \{1, ..., r\}$.

(b) Utiliser la formule admise pour représenter le graphe de

$$n \to \frac{1}{2} \sum_{\sigma \in \mathfrak{S}_{10}} |P_n(\sigma) - \pi(\sigma)|,$$

à nouveau pour $r = 3, 4, 5$ puis pour $r = 10, 52$ et enfin pour $r = 100$. Discuter.