Concours Blanc n°1 – Maths 2
ESSEC2, 2011
Corrigé

Mélange de jeux de cartes

Une question que se pose un joueur de cartes est de savoir combien de fois il est nécessaire de battre les cartes pour que le paquet soit convenablement mélangé. Ce problème décrit un procédé très élémentaire pour mélanger les cartes et propose de répondre alors à cette question.
Considérons un jeu de N cartes numérotées de C_1 à C_N et disposées en un paquet sur une table. Un joueur bat les cartes et repose le paquet sur la table. Le résultat du mélange est une permutation de ces N cartes.

Notations et Rappel :
On note S_N l'ensemble des ordres possibles pour ce paquet de N cartes et on rappelle que $\mathrm{Card}(S_N) = N!$.
On se place dans un espace probabilisé $(\Omega, \mathscr{A}, P)$ avec $\Omega = S_N$, $\mathscr{A} = \mathscr{P}(S_N)$ l'ensemble des parties de S_N et P l'équiprobabilité sur Ω. Pour toute variable aléatoire X on notera $E(X)$ et $V(X)$ l'espérance et la variance de X lorsqu'elles existent.

On considère qu'un paquet est *convenablement mélangé* lorsque toutes les permutations sont équiprobables, c'est-à-dire lorsque pour tout ordre de S_N, la probabilité que le tas de cartes se trouve dans cet ordre vaut $\frac{1}{N!}$.

Vocabulaire et notations :
Une carte située au sommet de la pile est dite *en position n°1*, celle qui se trouve immédiatement en dessous est dite *en position n°2*, etc. Ainsi une carte située *en position n°N* désigne la carte située en bas de la pile.
On prendra garde à bien distinguer la position d'une carte dans le paquet du numéro qu'elle porte.

Partons d'un tas de cartes rangées initialement dans l'ordre suivant : pour tout i élément de $[\![1, N]\!]$, la carte C_i se trouve en position i. Ainsi, à l'instant initial, la carte C_1 se trouve sur le dessus du paquet alors que C_N se trouve donc tout en dessous du paquet.

Pour k élément de $[\![1, N]\!]$, on appelle *insertion à la k-ième place* l'opération qui consiste à prendre la carte située au-dessus du paquet et à l'insérer entre la k-ième et la $(k+1)$-ième place. Une insertion à la première place ne change pas l'ordre des cartes. Une insertion à la N-ième place consiste à faire glisser la carte située au-dessus du paquet pour la mettre sous le paquet.

Le *battage par insertions* du jeu de cartes consiste à effectuer une suite d'insertions aléatoires, en choisissant, à chaque instant, au hasard uniformément dans $\{1, \dots, N\}$ la place à laquelle l'insertion a lieu, indépendamment des insertions précédentes. Les instants successifs d'insertions seront notées $1, 2, \dots, n, \dots$; l'instant initial est $n = 0$.

Notations. Nous notons :

- **T_1 le premier instant où la carte située sur le dessus du paquet est glissée en dernière position, c'est-à-dire le premier instant où la carte C_N se trouve remontée de la position N à la position $N-1$,**
- **T_2 le premier instant où la carte C_N se trouve remontée en position $N-2$,**
- **et plus généralement, pour i dans $[\![1, N-1]\!]$, T_i le premier instant où la carte C_N atteint la position $N-i$.**
- **On posera également $\Delta_1 = T_1$ et $\forall i \in [\![2, N-1]\!]$, $\Delta_i = T_i - T_{i-1}$.**
- **Enfin, on notera $T = T_{N-1} + 1$.**

On admet que les conditions de l'expérience permettent de faire l'hypothèse que les variables aléatoires $(\Delta_i)_{i\in[\![1,N-1]\!]}$ sont indépendantes.

Description d'un exemple. Dans le tableau ci-dessous, nous décrivons les résultats d'une expérience faite sur un paquet de $N = 4$ cartes. La première ligne du tableau indique les instants n, la deuxième ligne indique les positions d'insertions , et dans la dernière ligne figure la configuration du paquet à l'instant n.

	instant n	0	1	2	3	4	5	6	7
	insertion en place $k = ...$		3	2	4	1	3	4	2
Configuration du paquet	position 1	C_1	C_2	C_3	C_2	C_2	C_1	C_4	C_2
	position 2	C_2	C_3	C_2	C_1	C_1	C_4	C_2	C_4
	position 3	C_3	C_1	C_1	C_4	C_4	C_2	C_3	C_3
	position 4	C_4	C_4	C_4	C_3	C_3	C_3	C_1	C_1

Pour cette expérience, on a les résultats $T_1(\omega) = 3$, $T_2(\omega) = 5$, $T_3(\omega) = 6$ et $T(\omega) = 7$.

Partie 1 - Description et premiers résultats

1. **Justifier que $\forall\, i \in [\![2,N-1]\!]$, $T_i = \Delta_1 + \Delta_2 + \cdots + \Delta_i$.**
   Que représente l'intervalle de temps Δ_i ?

   D'après les définitions données :

$$\forall\, i \in [\![2,N-1]\!],\ \sum_{k=1}^{i} \Delta_k = \Delta_1 + \sum_{k=2}^{i} \left(T_k - T_{k-1}\right) = \Delta_1 + (T_i - T_1) = T_i \quad \text{avec } T_1 = \Delta_1$$

   Les Δ sont les intervalles de temps entre deux mouvements de la carte C_N ; plus précisément Δ_i est l'intervalle de temps entre l'arrivée de la carte C_N en position $N - i + 1$ et son arrivée en $N - i$.

2. **Lois des variables Δ_i.**

   (a) **Justifier que $\Delta_1 \hookrightarrow \mathcal{G}\left(\frac{1}{N}\right)$; puis que, pour tout $i \in [\![1,N-1]\!]$, $\Delta_i \hookrightarrow \mathcal{G}\left(\frac{i}{N}\right)$.**

   Δ_1 est le temps au bout duquel la carte C_N passera en $(N-1)$-ème position : ceci aura lieu au premier moment où l'insertion choisie sera une insertion à la N-ième place.
   Toutes les insertions étant équiprobables, et les choix d'insertion étant indépendants les uns des autres, on est dans une situation de loi géométrique : on répète de manière indépendante une expérience de Bernoulli de proba de succès $\frac{1}{N}$ jusqu'à obtenir un succès.
   Ainsi : $\Delta_1 \hookrightarrow \mathcal{G}\left(\frac{1}{N}\right)$.

   Le raisonnement est similaire pour $i \in [\![2,N]\!]$: par définition Δ_i est l'intervalle de temps nécessaire pour que la carte C_N passe de la position $N - i + 1$ à $N - i$; ceci arrive à la première insertion en position $N, N-1, N-2, \ldots$ ou $N - i + 1$. Il y a i telles insertions ; la proba que l'une d'entre elles soit choisie est donc égale à $\frac{i}{N}$.
   Par les mêmes arguments que précédemment, $\Delta_i \hookrightarrow \mathcal{G}\left(\frac{i}{N}\right)$.

   (b) **En déduire, pour tout $i \in [\![1,N-1]\!]$, $E(\Delta_i) = \frac{N}{i}$ et $V(\Delta_i) = N\frac{N-i}{i^2}$.**

   Ce sont les formules de cours : si $X \hookrightarrow \mathcal{G}(p)$, alors $E(X) = \frac{1}{p}$ et $V(X) = \frac{(1-p)}{p^2}$. Avec $p = \frac{i}{N}$, on trouve :

$$E(\Delta_i) = \frac{1}{i/N} = \frac{N}{i} \quad \text{et} \quad V(\Delta_i) = \frac{1 - i/N}{i^2/N^2} = \frac{N^2}{i^2}\frac{N-i}{N} = N\frac{N-i}{i^2}$$

3. **Loi de T_2 . Soit $n \geqslant 2$.**

(a) **Démontrer que** $P(T_2 = n) = \sum_{k=1}^{n-1} P(\Delta_2 = n-k)P(\Delta_1 = k)$**.**

On a vu en question 1 que $T_2 = \Delta_1 + \Delta_2$. Comme les Δ sont à valeurs dans $\mathbb{N}^*$, on a

$$(T_2 = n) = \bigsqcup_{k=1}^{n-1} \big((\Delta_1 = k) \cap (\Delta_2 = n-k)\big)$$

et donc, en utilisant l'indépendance des Δ_i (donnée par l'énoncé) :

$$\begin{aligned} P(T_2 = n) &= \sum_{k=1}^{n-1} P\big((\Delta_1 = k) \cap (\Delta_2 = n-k)\big) \\ &= \sum_{k=1}^{n-1} P\big(\Delta_1 = k\big) P(\Delta_2 = n-k) \end{aligned}$$

(b) **Justifier que** $\sum_{k=1}^{n-1} \left(\frac{1-1/N}{1-2/N}\right)^k = N\left(1-\frac{1}{N}\right)\left[\left(\frac{1-1/N}{1-2/N}\right)^{n-1} - 1\right]$**.**

C'est une somme géométrique, de raison $\frac{1-1/N}{1-2/N} \neq 1$. On trouve avec la formule habituelle, et un peu d'abnégation :

$$\begin{aligned} \sum_{k=1}^{n-1} \left(\frac{1-1/N}{1-2/N}\right)^k &= \frac{1-1/N}{1-2/N} \frac{1-\left(\frac{1-1/N}{1-2/N}\right)^{n-1}}{1-\left(\frac{1-1/N}{1-2/N}\right)} \\ &= \frac{(1-1/N)\left(1-\left(\frac{1-1/N}{1-2/N}\right)^{n-1}\right)}{(1-2/N)-(1-1/N)} \\ &= \frac{(1-1/N)\left(1-\left(\frac{1-1/N}{1-2/N}\right)^{n-1}\right)}{-1/N} \\ \sum_{k=1}^{n-1} \left(\frac{1-1/N}{1-2/N}\right)^k &= N\left(1-\frac{1}{N}\right)\left[\left(\frac{1-1/N}{1-2/N}\right)^{n-1} - 1\right] \end{aligned}$$

Rq : il faut aussi supposer $N \geqslant 3$ pour que le dénominateur ne s'annule pas, ce que l'énoncé ne dit pas clairement...

(c) **En déduire que l'on a :** $P(T_2 = n) = \frac{2}{N}\left[\left(1-\frac{1}{N}\right)^{n-1} - \left(1-\frac{2}{N}\right)^{n-1}\right]$**.**

Avec $\Delta_1 \hookrightarrow \mathscr{G}\left(\frac{1}{N}\right)$ et $\Delta_2 \hookrightarrow \mathscr{G}\left(\frac{2}{N}\right)$, on trouve à l'aide des deux questions précédentes :

$$\begin{aligned} P(T_2 = n) &= \sum_{k=1}^{n-1} \frac{1}{N}\left(1-\frac{1}{N}\right)^{k-1} \times \frac{2}{N}\left(1-\frac{2}{N}\right)^{n-k-1} \\ &= \frac{2}{N^2}\left(1-\frac{2}{N}\right)^{n-1}\left(1-\frac{1}{N}\right)^{-1} \sum_{k=1}^{n-1}\left(\frac{1-1/N}{1-2/N}\right)^k \\ &= \frac{2}{N^2}\left(1-\frac{2}{N}\right)^{n-1}\left(1-\frac{1}{N}\right)^{-1} N\left(1-\frac{1}{N}\right)\left[\left(\frac{1-1/N}{1-2/N}\right)^{n-1} - 1\right] \\ P(T_2 = n) &= \frac{2}{N}\left[\left(1-\frac{1}{N}\right)^{n-1} - \left(1-\frac{2}{N}\right)^{n-1}\right] \end{aligned}$$

4. **À l'instant** T_2**, la carte** C_N **est située en position** $N-2$ **et deux cartes se trouvent sous elle qui ont été insérées aux instants** T_1 **et** T_2**.**
Que valent alors les probabilités, qu'à l'instant T_2 **:**

   (a) la carte insérée à l'instant T_1 **soit en place** $N-1$ **et celle insérée à l'instant** T_2 **en place** N **?**

   (b) la carte insérée à l'instant T_2 **soit en place** $N-1$ **et celle insérée à l'instant** T_1 **en place** N **?**

On note C_a la carte insérée en bas du paquet à l'instant T_1 : le paquet est alors de la forme $\begin{array}{c}\vdots\\ \mathrm{C}_{\mathrm{N}-1}\\ \mathrm{C}_{\mathrm{N}}\\ \mathrm{C}_a\end{array}$

À l'instant T_2, la carte C_{N} remonte en position $\mathrm{N}-2$ ssi la position d'insertion tirée vaut N (auquel cas on aura un paquet $\begin{array}{c}\vdots\\ \mathrm{C}_{\mathrm{N}}\\ \mathrm{C}_a\\ \mathrm{C}_b\end{array}$ – donc la première alternative donnée par la question) ou $\mathrm{N}-1$ (ce qui donnera le paquet $\begin{array}{c}\vdots\\ \mathrm{C}_{\mathrm{N}}\\ \mathrm{C}_b\\ \mathrm{C}_a\end{array}$ – donc la seconde alternative). Il s'agit alors de probas conditionnelles : la proba de la première alternative vaut $\mathrm{P}_{k\in\{\mathrm{N}-1,\mathrm{N}\}}(k=\mathrm{N})$, où k est la position tirée pour l'insertion de C_b ; et la proba de la seconde est $\mathrm{P}_{k\in\{\mathrm{N}-1,\mathrm{N}\}}(k=\mathrm{N}-1)$.

Par équiprobabilité, ces deux probas valent $\dfrac{1}{2}$.

5. **À l'instant T_3, la carte C_{N} est située en position $\mathrm{N}-3$ et trois cartes, insérées aux instants T_1, T_2 et T_3, se trouvent sous elle. On note alors, pour $i\in\{1,2,3\}$, a_i la position de la carte ayant été insérée à l'instant T_i.**

   (a) **Combien y a-t-il de résultats possibles pour le triplet (a_1,a_2,a_3) ?**

   On généralise la discussion précédente : si C_a, C_b, C_c sont les 3 cartes insérées en dessous de C_{N} aux temps $\mathrm{T}_{1,2,3}$, les 6 ordres sont possibles :

   $\begin{array}{c}\vdots\\ \mathrm{C}_{\mathrm{N}}\\ \mathrm{C}_a\\ \mathrm{C}_b\\ \mathrm{C}_c\end{array}$ (3 insertions successives au bas du paquet), $\begin{array}{c}\vdots\\ \mathrm{C}_{\mathrm{N}}\\ \mathrm{C}_b\\ \mathrm{C}_a\\ \mathrm{C}_c\end{array}$ (insertions en N, puis $\mathrm{N}-1$, puis N), etc.

   (b) **Quelques exemples. Donner les probabilités qu'à l'instant T_3 :**

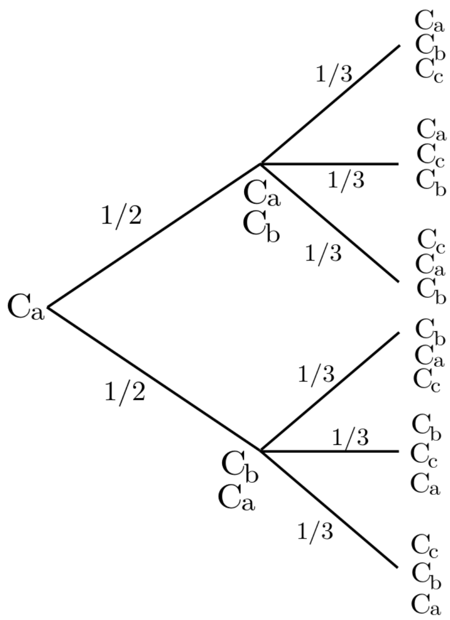
   i. on obtienne $(a_1,a_2,a_3)=(\mathrm{N}-2,\mathrm{N}-1,\mathrm{N})$?

   ii. on obtienne $(a_1,a_2,a_3)=(\mathrm{N}-2,\mathrm{N},\mathrm{N}-1)$?

   On discute encore probas conditionnelles à l'aide d'un arbre.

   - À la première insertion, C_a arrive en bas du paquet ;
   - à la seconde insertion, C_b arrive en bas du paquet avec proba $1/2$, et au-dessus de C_a avec proba $1/2$;
   - à la troisième insertion, C_c arrive en position $\mathrm{N},\mathrm{N}-1,\mathrm{N}-2$ avec des probas $1/3$

   Les deux probas demandées correspondent à deux chemins sur le graphe, donc sont de probas $\dfrac{1}{2}\times\dfrac{1}{3}$ par probas composées : (i) donne le paquet $\begin{array}{c}\vdots\\ \mathrm{C}_{\mathrm{N}}\\ \mathrm{C}_a\\ \mathrm{C}_b\\ \mathrm{C}_c\end{array}$ et (ii) donne $\begin{array}{c}\vdots\\ \mathrm{C}_{\mathrm{N}}\\ \mathrm{C}_a\\ \mathrm{C}_c\\ \mathrm{C}_b\end{array}$



6. **Justifier la phrase suivante :**

«À partir de l'instant T, toutes les configurations du jeu de cartes sont équiprobables.»

L'instant T est celui où on insère la carte C_N dans le paquet. Avant cet instant, on a inséré les $N-1$ cartes précédentes en-dessous de C_N : on obtient $(N-1)!$ ordres possibles, chacun de proba $\frac{1}{(N-1)!}$ par généralisation de ce qui précède. On rajoute alors un niveau à notre arbre (qui était de profondeur $N-1$, et devient maintenant de profondeur N) pour les N possibilités de l'insertion de C_N ; les branches rajoutées sont pondérées par une proba $\frac{1}{N}$.

On obtient au final les N! ordres possibles du paquet, avec une proba $\frac{1}{(N-1)!} \times \frac{1}{N} = \frac{1}{N!}$: à l'instant T, tous les ordres du paquet sont alors équiprobables.

NB : à ce stade-là vous devriez raisonnablement avoir récupéré tous les points de la question...

L'énoncé va toutefois plus loin et demande : à partir de l'instant T.
On peut procéder à une récurrence pour conclure. À l'instant T, notre arbre comporte N! branches équiprobables.
On continue la procédure en insérant la carte du dessus du paquet dans une des N positions tirée au hasard. On rajoute alors un niveau à l'arbre, qui compte maintenant $N \times N!$ sommets, et chaque sommet est obtenu avec proba $\frac{1}{N!} \times \frac{1}{N}$. Comme il n'y a que N! configurations du paquet, chaque ordre apparaît plusieurs fois : justifions que chaque ordre apparaît N fois à l'aide d'un exemple. Sur un jeu de 3 cartes, le paquet $\begin{matrix} C_2 \\ C_1 \\ C_3 \end{matrix}$ peut sortir : après insertion en 1ère position en partant de $\begin{matrix} C_2 \\ C_1 \\ C_3 \end{matrix}$; après insertion en 2ème position en partant de $\begin{matrix} C_1 \\ C_2 \\ C_3 \end{matrix}$; après insertion en 3ème position en partant de $\begin{matrix} C_3 \\ C_2 \\ C_1 \end{matrix}$.
Pour revenir à N quelconque, un ordre étant donné, il existe donc N configuration précédentes menant à cet ordre, après insertion de la carte du dessus dans une des N positions possibles. Chaque ordre apparaît donc N fois, chaque fois avec une proba $\frac{1}{N!} \times \frac{1}{N}$; donc finalement avec une proba $\frac{1}{N!}$. Une fois que les ordres sont devenus équiprobables (à l'instant T), il el restent donc à chaque étape, ce qui permet de

conclure.

On retiendra que si on arrête le battage des cartes par insertion exactement à l'instant T***, on a un paquet*** **convenablement mélangé.** ***Cependant le temps*** T ***étant aléatoire, il n'est pas possible d'arrêter de battre les cartes à cet instant précis, à moins de marquer la carte*** $\mathrm{C_N}$ ***bien sûr!***

Partie 2 - Estimation du nombre d'insertions pour bien mélanger les cartes

Notations : on introduit les suites $(\mathrm{H}_n)_{n\geqslant 1}$ **et** $(u_n)_{n\in\mathbb{N}^*}$ **définies par :**

$$\forall n \geqslant 1,\ \mathrm{H}_n = \sum_{k=1}^{n} \frac{1}{k} \text{ et } u_n = \mathrm{H}_n - \ln(n)$$

7. **Espérance et variance de** T

Justifier que $\mathrm{E(T)} = \mathrm{N\,H_N}$ **et que** $\mathrm{V(T)} = \mathrm{N}^2\left(\sum_{k=1}^{\mathrm{N}} \frac{1}{k^2}\right) - \mathrm{N\,H_N}$**.**

On a vu dans la question 1 que $\mathrm{T}_{n-1} = \Delta_1 + \Delta_2 + \cdots + \Delta_{\mathrm{N}-1}$; d'où

$$\mathrm{T} = 1 + \mathrm{T}_{\mathrm{N}-1} = 1 + \sum_{i=1}^{\mathrm{N}-1} \Delta_i$$

Par linéarité de l'espérance, et en utilisant $\Delta_i \hookrightarrow \mathscr{G}\left(\frac{i}{\mathrm{N}}\right)$ (question 2a), on a

$$\mathrm{E(T)} = 1 + \sum_{i=1}^{\mathrm{N}-1} \frac{\mathrm{N}}{i} = \mathrm{N} \times \frac{1}{\mathrm{N}} + \mathrm{N}\sum_{i=1}^{\mathrm{N}-1} \frac{1}{i} = \mathrm{NH_N}$$

et par indépendance des Δ_i, et leur variance vue en 2b :

$$\begin{aligned}\mathrm{V(T)} = \mathrm{V}\left(1 + \sum_{i=1}^{\mathrm{N}-1} \Delta_i\right) = \sum_{i=1}^{\mathrm{N}} \mathrm{V}(\Delta_i) = \mathrm{N}\sum_{i=1}^{\mathrm{N}-1} \frac{\mathrm{N}-i}{i^2} &= \mathrm{N}^2\sum_{i=1}^{\mathrm{N}-1} \frac{1}{i^2} - \mathrm{N}\sum_{i=1}^{\mathrm{N}-1} \frac{1}{i} \\ &= \mathrm{N}^2\left(\sum_{i=1}^{\mathrm{N}} \frac{1}{i^2} - \frac{1}{\mathrm{N}^2}\right) - \mathrm{N}\left(\sum_{i=1}^{\mathrm{N}} \frac{1}{i} - \frac{1}{\mathrm{N}}\right) \\ &= \mathrm{N}^2\sum_{i=1}^{\mathrm{N}} \frac{1}{i^2} - \mathrm{N}\sum_{i=1}^{\mathrm{N}} \frac{1}{i} \\ &= \mathrm{N}^2\left(\sum_{k=1}^{\mathrm{N}} \frac{1}{k^2}\right) - \mathrm{N\,H_N}\end{aligned}$$

8. **Étude de la suite** (u_n)

   (a) **Montrer :** $\forall\, x > -1, \ln(1+x) \leqslant x$.
   Vue maintes fois en cours / Ds précédents.

   (b) **En déduire :** $\forall\, k \in \mathbb{N}^*,\ \frac{1}{k+1} \leqslant \ln(k+1) - \ln(k) \leqslant \frac{1}{k}$.
   On pourra, pour l'une des deux inégalités, poser $x = -\frac{1}{k+1}$**.**
   Soit $k \in \mathbb{N}^*$. On a alors $\frac{1}{k} > 0 > -1$, donc la question précédente donne $\ln\left(1 + \frac{1}{k}\right) \leqslant \frac{1}{k}$, ce qui s'écrit aussi : $\ln\left(\frac{k+1}{k}\right) \leqslant \frac{1}{k}$, ou encore

$$\ln(k+1) - \ln(k) \leqslant \frac{1}{k}$$

   On a également $k+1 \geqslant 2$, donc $\frac{1}{k+1} \leqslant \frac{1}{2}$, donc $-\frac{1}{k+1} \geqslant -\frac{1}{2} > -1$; la question précédente donne cette fois : $\ln\left(1 - \frac{1}{k+1}\right) \leqslant -\frac{1}{k+1}$, d'où $\ln\left(\frac{k}{k+1}\right) \leqslant -\frac{1}{k+1}$, puis $\ln(k) - \ln(k+1) \leqslant -\frac{1}{k+1}$; en multipliant par (-1) on a bien :

$$\ln(k+1) - \ln(k) \geqslant \frac{1}{k+1}$$

(c) **En déduire successivement :**

i. la décroissance de la suite (u_n)**,**

ii. l'encadrement : $\forall n \in \mathbb{N}^*$**,** $\ln(n+1) \leqslant \mathrm{H}_n \leqslant \ln(n)+1$**.**

Pour tout $n \in \mathbb{N}^*$:

$$u_{n+1} - u_n = \mathrm{H}_{n+1} - \mathrm{H}_n - \ln(n+1) + \ln(n) = \frac{1}{n+1} - \ln(n+1) + \ln(n)$$

L'inégalité de gauche de l'encadrement de la question précédente ($n \geqslant 1$ donc on peut s'en servir) donne alors immédiatement :

$$\forall\, n \in \mathbb{N}^*,\ u_{n+1} - u_n \leqslant 0$$

et la suite (u_n) est bien décroissante.

On a aussi : $\forall\, k \in \mathbb{N}^*,\ \dfrac{1}{k} \geqslant \ln(k+1) - \ln(k)$; donc en sommant pour k de 1 à n :

$$\mathrm{H}_n = \sum_{k=1}^{n} \frac{1}{k} \geqslant \sum_{k=1}^{n} \big(\ln(k+1) - \ln(k)\big) = \ln(n+1) - \ln(1) = \ln(n+1)$$

(par somme téléscopique).

Enfin, on écrit, pour $n \geqslant 2$:

$$\mathrm{H}_n = 1 + \sum_{k=2}^{n} \frac{1}{k} = 1 + \sum_{k=1}^{n-1} \frac{1}{k+1} \leqslant 1 + \sum_{k=1}^{n-1} \big(\ln(k+1) - \ln(k)\big) = 1 + \ln(n) - \ln(1) = 1 + \ln(n)$$

inégalité encore vraie pour $n = 1$ car $\mathrm{H}_1 = 1$ et $1 + \ln(1) = 1$.

(d) **Déduire de ce qui précède que la suite** (u_n) **est convergente et que sa limite, notée** γ **appartient à** $[0,1]$**.**

Comme (u_n) est décroissante, il suffit de montrer qu'elle est minorée. Or :

$$\forall\, n \in \mathbb{N}^*,\ \mathrm{H}_n \geqslant \ln(n+1) \Rightarrow u_n \geqslant \ln(n+1) - \ln(n) \geqslant 0 \quad \text{croissance de ln}$$

donc (u_n) est minorée par 0 : elle converge vers $\gamma \in \mathbb{R}$.
De plus on a montré que $\forall\, n \in \mathbb{N}^*,\ 0 \leqslant u_n \leqslant 1$: en passant à la limite on obtient $0 \leqslant \gamma \leqslant 1$.

9. (a) **Établir que** $\mathrm{E(T)} \underset{\mathrm{N}\to+\infty}{\sim} \mathrm{N}\ln(\mathrm{N})$ **et** $\mathrm{E(T)} = \mathrm{N}\ln(\mathrm{N}) + \mathrm{N}\gamma + o(\mathrm{N})$.

Cela revient à établir l'équivalent de la somme harmonique.
On a vu que $\mathrm{H_N} = \ln(\mathrm{N}) + u_\mathrm{N}$, avec $u_\mathrm{N} \to \gamma \in \mathbb{R}$. Alors $\dfrac{\mathrm{H_N}}{\ln(\mathrm{N})} = 1 + \dfrac{u_\mathrm{N}}{\ln(\mathrm{N})} \to 1$ pour $\mathrm{N} \to +\infty$, de sorte que $\mathrm{H_N} \sim_{\mathrm{N}\to+\infty} \ln(\mathrm{N})$.
Comme $\mathrm{E(T)} = \mathrm{NH_N}$, on en déduit $\mathrm{E(T)} \underset{\mathrm{N}\to+\infty}{\sim} \mathrm{N}\ln(\mathrm{N})$.

De plus comme $(u_\mathrm{N}) \to \gamma \Leftrightarrow u_\mathrm{N} = \gamma + o(1)$, a aussi

$$\mathrm{E(N)} = \mathrm{NH_N} = \mathrm{N}(\ln(\mathrm{N}) + u_\mathrm{N}) = \mathrm{N}(\ln(\mathrm{N}) + \gamma + o(1)) = \mathrm{N}\ln(\mathrm{N}) + \mathrm{N}\gamma + o(\mathrm{N})$$

(b) **Quelle est la nature de la suite** $\left(\dfrac{\mathrm{V(T)}}{\mathrm{N}^2}\right)_{\mathrm{N}\in\mathbb{N}^*}$ **? (on prendra garde au fait que** $\mathrm{V(T)}$ **dépend de** N**).**
Déterminer une constante α **strictement positive (à écrire sous forme d'une somme infinie), telle que** $\mathrm{V(T)} \underset{\mathrm{N}\to+\infty}{\sim} \alpha\mathrm{N}^2$**. Montrer qu'alors** $\mathrm{V(T)} \leqslant \alpha\mathrm{N}^2$**.**

D'après 7, on a $\dfrac{\mathrm{V(T)}}{\mathrm{N}^2} = \displaystyle\sum_{k=1}^{\mathrm{N}} \frac{1}{k^2} - \frac{\mathrm{H_N}}{\mathrm{N}}$. On sait que $\sum 1/n^2$ converge (Riemann de paramètre $2 > 1$) ; et que $\dfrac{\mathrm{H_N}}{\mathrm{N}} \underset{\mathrm{N}\to+\infty}{\sim} \dfrac{\ln(\mathrm{N})}{\mathrm{N}}$ donc tend vers 0 pour $\mathrm{N} \to +\infty$ par croissances comparées.

On en déduit

$$\lim_{\mathrm{N}\to+\infty} \frac{\mathrm{V(T)}}{\mathrm{N}^2} = \sum_{k=1}^{+\infty} \frac{1}{k^2} = \alpha > 0$$

ce qui donne directement $V(T) \underset{N\to+\infty}{\sim} \alpha N^2$.

De plus, pour tout $N \geqslant 1$,

$$V(T) = N^2 \sum_{k=1}^{N} \frac{1}{k^2} - N H_N \leqslant N^2 \sum_{k=1}^{N} \frac{1}{k^2} \leqslant N^2 \sum_{k=1}^{+\infty} \frac{1}{k^2} = \alpha N^2$$

car $H_N \geqslant 0$ et tous les $1/k^2$ sont > 0.

10. **Écart à la moyenne**
On donne l'inégalité de *Bienaymé-Chebychev* valable pour une variable aléatoire X admettant une espérance et une variance :

$$\forall \varepsilon > 0,\ P\big(|X - E(X)| \geqslant \varepsilon\big) \leqslant \frac{V(X)}{\varepsilon^2}$$

Soit N fixé et une constante c strictement plus grande que 1.

(a) **Justifier que** $\forall \omega \in \Omega,\ |T(\omega) - N\ln(N)| \leqslant |T(\omega) - E(T)| + N$ **(on remarquera que** $N\ln(N) = E(T) - Nu_N$**).**
Montrer l'inclusion entre événements :

$$\big(|T - N\ln(N)| \geqslant cN\big) \subset \big(|T - E(T)| \geqslant N(c-1)\big)$$

On utilise $E(T) = NH_N = N\ln(N) + Nu_N$. $\forall\, \omega \in \Omega$, on peut écrire :

$$|T(\omega) - N\ln(N)| = |T(\omega) - E(T) + Nu_N| \leqslant |T(\omega) - E(T)| + N|u_N| \leqslant |T(\omega) - E(T)| + N$$

(avec inégalité triangulaire et $|u_N| \leqslant 1$, car $u_N \in [0,1]$).
Soit ω tel que $|T(\omega) - N\ln(N)| \geqslant cN$; alors on a aussi, en utilisant l'inégalité précédente :

$$|T(\omega) - E(T)| \geqslant |T(\omega) - N\ln(N)| - N \geqslant cN - N = (c-1)N$$

ce qui montre que :

$$\big(|T - N\ln(N)| \geqslant cN\big) \subset \big(|T - E(T)| \geqslant N(c-1)\big)$$

(b) **Démontrer que :**

$$P\big(|T - N\ln(N)| \geqslant cN\big) \leqslant \frac{\alpha}{(c-1)^2}$$

où α a été définie à la question 9b.
Le nombre N étant fixé, que vaut $\lim_{c\to+\infty} P\big(|T - N\ln(N)| \geqslant cN\big)$ **?**
L'inclusion précédente, puis BT avec $\varepsilon = N(c-1)$, donnent :

$$P\big(|T - N\ln(N)| \geqslant cN\big) \leqslant P\big(|T - E(T)| \geqslant N(c-1)\big) \leqslant \frac{V(T)}{N^2(c-1)^2}$$

Or $V(T) \leqslant \alpha N^2$ d'après 9b ; on obtient bien

$$P\big(|T - N\ln(N)| \geqslant cN\big) \leqslant \frac{\alpha}{(c-1)^2}$$

et, en minorant par 0 (une proba est positive) et en faisant tendre $c \to +\infty$, on obtient par gendarmes que

$$\lim_{c\to+\infty} P\big(|T - N\ln(N)| \geqslant cN\big) = 0$$

11. **Démontrer aussi que pour tout $\varepsilon > 0$:**

$$\lim_{N\to+\infty} P\big(|T - N\ln(N)| \geqslant \varepsilon N\ln(N)\big) = 0$$

On peut reprendre ce qui précède avec $c = \varepsilon\ln(N)$ (qui sera > 1 pour N assez grand) : on trouve

$$P\big(|T - N\ln(N)| \geqslant \varepsilon N\ln(N)\big) \leqslant \frac{\alpha}{(\varepsilon\ln(N) - 1)^2}$$

et on a alors, toujours par gendarmes :

$$\lim_{N\to+\infty} P\big(|T-N\ln(N)| \geqslant \varepsilon N\ln(N)\big)=0$$

On peut traduire ces résultats en disant que l'événement : «T **s'écarte de** $N\ln(N)$ **de manière significative»** ***est un événement asymptotiquement rare.***
Pour information, pour un paquet de 32 cartes, on donne $32\ln(32)\simeq 110$***, et pour un paquet de 52 cartes,*** $52\ln(52)\simeq 205$**.**

12. **Simulation informatique**

On veut implémenter cette méthode de mélange en Python. On suppose les imports usuels effectués :

```
import numpy as np
import numpy.random as rd
```

On considère un paquet de 32 cartes ; on représente le paquet par une liste de 32 éléments, qui sont les numéros des cartes. Le premier élément de la liste est la carte du dessus ; le dernier est l'élément du bas de la pile. Initialement, le paquet sera donc représenté par la liste `[1,2,3,...,31,32]`**.**
Une insertion dans le paquet modifie cette liste ; si on reprend les insertions exposées dans le tableau de début de problème, pour un jeu de 4 cartes, la liste représentant le paquet vaudra successivement :

`[1,2,3,4]`, `[2,3,1,4]`, `[3,2,1,4]`, `[2,1,4,3]`, `[2,1,4,3]`, `[1,4,2,3]`, `[4,2,3,1]`, `[2,4,3,1]`

(a) **Modélisation d'une insertion.**
Compléter le programme suivant pour qu'il effectue une insertion de la carte du dessus à la position `k`**.**
(Attention : si le paquet est représenté par une liste `L`**, alors la carte « à la position** `k` **» est** `L[k-1]`**).**

Il faut décaler les cartes d'indice $0,1,2,...,k-1$ vers la gauche ; donc faire les substitutions « `L[0]` reçoit `L[0]` » , « `L[1]` reçoit `L[2]` » ,...« `L[k-2]` reçoit `L[k-1]` » ; puis insérer la carte du dessus en position `L[k-1]`.

```
def insertion(paquet,k):
    # k dans [1,n] : position où la nouvelle carte est insérée
    paquet_final=paquet
    carte_dessus=paquet_final[0]
    if k>1 :
        for i in range(k-1):
            paquet_final[i] = paquet_final[i+1]
        paquet_final[k-1]=carte_dessus
    return paquet_final
```

(b) **On rappelle que la comparaison « différent de » se note** `!=`**.**
Compléter le programme suivant pour qu'il effectue un mélange suivant le protocole du problème, et renvoie le paquet mélangé, et le temps T nécessaire à ce mélange.
La liste `[1,2,...,N]` s'obtient par `list(range(1,N+1))`.
Ensuite on tire au hasard (équiprobable) une position d'insertion dans $[\![1,N]\!]$ puis on insère la carte du dessus à cette position, à l'aide de la fonction précédente.
On fait ça tant que la carte du dessus n'est pas la dernière.
On fait ensuite une dernière insertion aléatoire pour obtenir l'état mélangé définitif.
Au cours du programme, le compteur `t` compte le nombre d'insertions effectuées.

```
def temps_melange(N): # N : taille du paquet
    paquet=list(range(1,N+1))
    t=0
    while paquet[0]!=N:
        indice=rd.randint(1,N+1)
        paquet=insertion(paquet,indice)
        t+=1
    indice=rd.randint(1,N+1)
    paquet=insertion(paquet,indice)
    t+=1
    return paquet,t
```

On souhaite tester le caractère uniforme de ce mélange (c'est-à-dire que tous les ordres possibles du paquet apparaissent de manière équiprobable).
On se place dans le cas d'un paquet de 4 cartes.

(c) **Le script suivant :**

```
n=0
for k in range(100000):
    if temps_melange(4)[0] == [1,2,3,4]:
        n=n+1
print(n/100000)
print(1/24)
```

produit l'affichage :

```
0.04153
0.041666666666666664
```

Qu'en pensez-vous ?
Dans ce code, on effectue 100 000 mélanges du paquet `[1,2,3,4]` en suivant le protocole indiqué, et `n` compte le nombre de fois où, à l'issue de ce mélange, on retrouve l'ordre initial.
`n/100000` est donc la fréquence d'apparition de l'ordre initial à l'issue du mélange.
Le problème affirme que chaque ordre est équiprobable à l'issue du mélange. Pour un paquet de 4 cartes, il y a $4! = 24$ ordres possibles, qui doivent dons tous sortir avec une proba $1/24$. Avec la simulation effectuée, on remarque bien que la fréquence d'apparition de l'ordre initial est $\simeq 1/24$, ce qui va dans le sens des calculs effectués.

Partie 3 - Distance variationnelle à la loi uniforme

Notations :

- **On note π l'équiprobabilité sur S_N, c'est-à-dire l'application de $\mathscr{P}(S_N)$ dans $[0,1]$ telle que :**
$$\forall A \subset S_N \quad \pi(A) = \frac{\text{Card}(A)}{N!} \textbf{; en particulier, } \forall \sigma \in S_N,\ \pi(\{\sigma\}) = \frac{1}{N!}$$
- **On note également μ_n la probabilité sur S_N définie comme suit :**
pour chaque configuration σ de S_N, $\mu_n(\{\sigma\})$ désigne la probabilité qu'à l'instant n le tas de cartes se trouve dans la configuration σ.
On a alors pour pour toute partie A de S_N, $\mu_n(A) = \sum_{\sigma \in A} \mu_n(\sigma)$.
On peut mesurer la qualité du mélange à un instant donné n en estimant l'écart entre μ_n et π. Une *distance* d entre ces probabilités est définie de la manière suivante :
$$d(\mu_n, \pi) = \max\left\{\left|\mu_n(A) - \pi(A)\right|,\ A \subset S_N\right\}$$

13. **Soient A une partie de S_N, $n \in \mathbb{N}^*$, et E_n l'événement : «à l'instant n le paquet de cartes se trouve dans une configuration qui appartient à la partie A.» On a donc $\mu_n(A) = P(E_n)$.**

(a) **Expliquer, en utilisant la question ??, l'égalité suivante : $P_{T \leqslant n}(E_n) = \pi(A)$.**
En déduire $P(E_n \cap (T \leqslant n)) = \pi(A)P(T \leqslant n)$.
Si $T \leqslant n$, le nombre d'étapes n est supérieur au rang T à partir duquel, d'après la question 6, toutes les configurations du paquet sont équiprobables. Si A est un ensemble de telles configurations, la définition de E_n donne $P_{T \leqslant n}(E_n) = \dfrac{\text{Card}(A)}{N!} = \pi(A)$.
Par définition d'une probabilité conditionnelle, on en déduit que $\dfrac{P(E_n \cap (T \leqslant n))}{P(T \leqslant n)} = \pi(A)$; ce qui donne bien l'égalité voulue.

(b) **Établir que $P(E_n \cap (T > n)) \leqslant P(T > n)$.**
On a l'inclusion entre événements : $(E_n \cap (T > n)) \subset (T > n)$; ce qui donne directement le résultat par propriété des probabilités.

(c) **Montrer que :**

$$\mu_n(A) \leqslant \pi(A) + P(T > n)$$

D'après l'énoncé on a $\mu_n(A) = P(E_n)$. On peut alors appliquer les probas totales au SCE $\{(T \leqslant n), (T > n)\}$ et en utilisant les questions précédentes :

$$\begin{aligned} P(E_n) &= P(E_n \cap (T > n)) + P(E_n \cap (T \leqslant n)) \\ &\leqslant P(T > n) + \pi(A)P(T \leqslant n) \\ P(E_n) &\leqslant P(T > n) + \pi(A) \qquad \text{en utilisant } P(T \leqslant n) \leqslant 1 \end{aligned}$$

14. **Soit A une partie de** S_N **et** $n \in \mathbb{N}^*$. **On note** $\overline{A}$ **l'événement contraire de** A.

(a) **Exprimer** $\mu_n(\overline{A}) - \pi(\overline{A})$ **en fonction de** $\mu_n(A) - \pi(A)$.
μ_n et π étant des probabilités, on a $\mu_n(\overline{A}) = 1 - \mu_n(A)$, et $\pi(\overline{A}) = 1 - \pi(A)$; d'où

$$\mu_n(\overline{A}) - \pi(\overline{A}) = \pi(A) - \mu_n(A)$$

(b) **Déduire des questions précédentes la majoration :**

$$\left|\mu_n(A) - \pi(A)\right| \leqslant P(T > n)$$

Le résultat de la question 13c peut s'appliquer aux événements A et $\overline{A}$: on obtient

$$\mu_n(A) - \pi(A) \leqslant P(T > n) \quad \text{et} \quad \mu_n(\overline{A}) - \pi(\overline{A}) \leqslant P(T > n)$$

ce qui donne encore

$$\mu_n(A) - \pi(A) \leqslant P(T > n) \quad \text{et} \quad \pi(A) - \mu_n(A) \leqslant P(T > n)$$

Si $x \leqslant y$ et $-x \leqslant y$, on peut conclure $|x| \leqslant y$; ce qui permet de conclure.

15. **Montrer que** $\forall\ n \in \mathbb{N}^*,\ 0 \leqslant d(\mu_n, \pi) \leqslant P(T > n)$. **Déterminer la limite** $\lim\limits_{n\to+\infty} d(\mu_n, \pi)$.

D'après l'énoncé, $d(\mu_n, \pi)$ est le max des $|\mu_n(A) - \pi(A)|$ sur tous les événements A. D'après la question précédente, ces quantités sont toutes majorées par $P(T > n)$; donc leur max l'est également. On a bien $d(\mu_n, \pi) \leqslant P(T > n)$. Enfin on a aussi $0 \leqslant d(\mu_n, \pi)$ car c'est un max de réels positifs.

Pour calculer la limite on cherche à appliquer le théorème des gendarmes : montrons que $P(T > n) \to 0$ pour $n \to +\infty$.

On sait que $P(T > n) = \sum_{k=n+1}^{+\infty} P(T = k)$. La série $\sum P(T = k)$ est convergente (de somme 1 car $T(\Omega) = \mathbb{N}$), donc ses restes partiels tendent vers 0 : ce sont exactement les $P(T > n)$.

On a donc montré : $\lim\limits_{n\to+\infty} P(T > n) = 0$: par théorème des gendarmes on en déduit

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

Partie 4- Une majoration de $P(T > n)$

Dans cette partie, nous nous intéressons provisoirement à un collectionneur de timbres. Celui-ci reçoit chaque jour une lettre affranchie avec un timbre choisi au hasard uniformément parmi les N timbres en vigueur. On étudie ici le nombre de jours que doit attendre le collectionneur pour posséder la collection complète des N timbres. Le jour 0 il n'a aucun timbre.
On note alors :

- **pour tout entier** $k \in [\![1, N]\!]$, S_k **le nombre aléatoire de jours que doit attendre le collectionneur pour que le nombre de timbres différents qu'il possède passe de** $k - 1$ **à** k**,**

- $S = S_1 + S_2 + \cdots + S_N$ **, soit la variable aléatoire correspondant au nombre de jours à attendre pour posséder la collection complète des** N **timbres,**

- **en supposant les N timbres en vigueur numérotés de 1 à N, pour tout $j \in [\![1,\mathrm{N}]\!]$, B_j^m l'événement «le jour m, le collectionneur n'a toujours pas reçu de lettre affranchie avec le timbre numéro j.»**

On admet que les variables aléatoires $(\mathrm{S}_k)_{k\in[\![1,\mathrm{N}]\!]}$ sont indépendantes.

16. **Déterminer la loi de S_1.**

    S_1 est le temps attendu par le collectionneur pour recevoir son premier timbre : cela arrive au premier jour. On a donc S_1 constante, égale à 1.

17. **Déterminer pour tout entier $k \in [\![2,\mathrm{N}]\!]$ la loi de la variable S_k.**

    S_k est le temps nécessaire, possédant déjà k timbres distincts, d'en recevoir un k-ième distinct de ceux possédés. Il s'agit donc de tirer des timbres parmi les N disponibles, de manière équiprobable, jusqu'à obtnir un qui ne soit pas dans les $k-1$ déjà possédés. On reconnaît une expérience menant à une loi géométrique, de paramètre la proba de succès, égale à $\dfrac{\mathrm{N}-(k-1)}{\mathrm{N}}$ par équiprobabilité des tirages.

    $$\mathrm{S}_k \hookrightarrow \mathscr{G}\left(\frac{\mathrm{N}-k+1}{\mathrm{N}}\right)$$

18. **En déduire que la variable S suit la même loi de probabilité que la variable T étudiée dans les parties précédentes.**
    Ce résultat sera utilisé pour estimer la quantité $\mathrm{P}(\mathrm{T}>n)$.

    On a donc $\mathrm{S} = 1+\mathrm{S}_2+\cdots+\mathrm{S}_\mathrm{N}$ (avec $\mathrm{S}_1 = 1$), les S_i indépendantes, et $\mathrm{S}_2 \hookrightarrow \mathscr{G}\left(\frac{\mathrm{N}-1}{\mathrm{N}}\right)$, $\mathrm{S}_3 \hookrightarrow \mathscr{G}\left(\frac{\mathrm{N}-2}{\mathrm{N}}\right)$, ..., $\mathrm{S}_\mathrm{N} \hookrightarrow \mathscr{G}\left(\frac{1}{\mathrm{N}}\right)$ d'une part ;

    $\mathrm{T} = \Delta_1+\Delta_2+\cdots+\Delta_{\mathrm{N}-1}+1$, avec les Δ_i indépendantes, et $\Delta_1 \hookrightarrow \mathscr{G}\left(\frac{1}{\mathrm{N}}\right)$, $\Delta_2 \hookrightarrow \mathscr{G}\left(\frac{2}{\mathrm{N}}\right)$, ... , $\Delta_{\mathrm{N}-1} \hookrightarrow \mathscr{G}\left(\frac{\mathrm{N}-1}{\mathrm{N}}\right)$ d'autre part ;

    S et T sont donc obtenues par la même construction, donc suivant la même loi.

19. **Soit $m \in \mathbb{N}^*$.**

    (a) **Exprimer l'événement $(\mathrm{S}>m)$ à l'aide des événements $\mathrm{B}_1^m, \mathrm{B}_2^m, \ldots, \mathrm{B}_\mathrm{N}^m$.**
    $(\mathrm{S}>m)$ a lieu ssi le collectionneur n'a pas complété sa collection au temps m : c'est-à-dire qu'il existe au moins un timbre qu'il n'a pas encore reçu au temps m. Ceci donne

    $$(\mathrm{S}>m) = \bigcup_{j=1}^{\mathrm{N}} \mathrm{B}_j^m$$

    (b) **Que vaut $\mathrm{P}(\mathrm{B}_j^m)$ pour tout entier $j \in [\![1,\mathrm{N}]\!]$?**
    B_j^m est réalisé ssi les m premiers tirages ont donné un autre timbre que le timbre j ; par indépendance des tirages successifs, on a

    $$\forall\, j \in [\![1,\mathrm{N}]\!],\ \mathrm{P}(\mathrm{B}_j^m) = \left(\frac{\mathrm{N}-1}{\mathrm{N}}\right)^m$$

    (c) **On rappelle que pour tout entier $n \geqslant 2$ et pour toute famille d'événements $\mathrm{A}_1, \ldots, \mathrm{A}_n$, on a l'inégalité : $\mathrm{P}\left(\bigcup_{i=1}^{n} \mathrm{A}_i\right) \leqslant \sum_{i=1}^{n} \mathrm{P}(\mathrm{A}_i)$. En déduire $\mathrm{P}(\mathrm{S}>m) \leqslant \mathrm{N}\left(1-\frac{1}{\mathrm{N}}\right)^m$.**
    D'après les deux questions qui précèdent et la propriété rappelée :

    $$\mathrm{P}(\mathrm{S}>m) = \mathrm{P}\left(\bigcup_{j=1}^{\mathrm{N}} \mathrm{B}_j^m\right) \leqslant \sum_{j=1}^{\mathrm{N}} \mathrm{P}(\mathrm{B}_j^m) = \sum_{j=1}^{\mathrm{N}} \left(\frac{\mathrm{N}-1}{\mathrm{N}}\right)^m = \mathrm{N}\left(1-\frac{1}{\mathrm{N}}\right)^m$$

20. **Déduire des résultats précédents et de la question (8a) la majoration :**

    $$\forall m \in \mathbb{N},\ \mathrm{P}(\mathrm{T}>m) \leqslant \mathrm{N}e^{-m/\mathrm{N}}$$

$N \geqslant 2$ donc $-\dfrac{1}{N} > -1$, d'où $\ln\left(1-\dfrac{1}{N}\right) \leqslant -\dfrac{1}{N}$ avec (8a). On en déduit par croissance de exp : $1-\dfrac{1}{N} \leqslant e^{-1/N}$ et donc $\left(1-\dfrac{1}{N}\right)^m \leqslant e^{-m/N}$ d'où avec la majoration précédente :

$$P(S > m) \leqslant N e^{-m/N}$$

S et T suivant la même loi on a donc bien :

$$\forall m \in \mathbb{N},\ P(T > m) \leqslant N e^{-m/N}$$

21. **On reprend les notations introduites dans la partie précédente.**

 (a) **Soit $c > 0$ fixé. Montrer que pour n entier supérieur ou égal à $N \ln N + cN$ on a : $d(\mu_n, \pi) \leqslant e^{-c}$.**
 On a vu en partie 3 : $\forall\ n \in \mathbb{N}^*,\ 0 \leqslant d(\mu_n, \pi) \leqslant P(T > n)$; avec la majoration de la question précédente il vient $0 \leqslant d(\mu_n, \pi) \leqslant N e^{-n/N}$. Si on choisit $n \geqslant N\ln(N) + cN$, la borne supérieure donne :

$$N e^{-n/N} \leqslant N e^{-\ln(N)-c} = e^{-c}$$

 et on trouve bien

$$0 \leqslant d(\mu_n, \pi) \leqslant e^{-c}$$

 (b) ***Application numérique.* On estime qu'une distance en variation à la loi uniforme de $\dfrac{1}{5}$ est acceptable.**
 Avec un jeu de 32 cartes, combien de *battages par insertions* doit-on faire pour considérer le paquet mélangé de façon acceptable ? (on donne $\ln 5 \simeq 1,6$ et $\ln(32) \simeq 3,4$).
 Pour obtenir $d(\mu_n, \pi) \leqslant e^{-c}$, il suffit de choisir c tel que $e^{-c} \leqslant \dfrac{1}{5} \Leftrightarrow c \geqslant \ln(5)$. Ceci donne $n \geqslant N \ln N + cN = 32\ln(32) + 32\ln(5) \simeq 32(3,4+1,6) = 160$ avec les valeurs fournies.
 On peut donc considérer le paquet comme bien mélangé au bout de 160 insertions.