Épreuve de Mathématiques 8

Durée 4 h

L'usage des calculatrices est interdit.

La présentation, la lisibilité, l'orthographe, la qualité de la rédaction et la précision des raisonnements entreront pour une part importante dans l'appréciation des copies. En particulier, les résultats non justifiés ne seront pas pris en compte.
Les candidats sont invités à encadrer les résultats de leurs calculs.

Exercice 1

Soit n un entier naturel non nul. Une entreprise dispose d'un lot du n feuilles originales qu'elle a numérotées $l,\ 2,\ \cdots,\ n$. Elle photocopie ces n feuilles originales et souhaite que chaque original soit agrafé avec sa copie. L'entreprise programme le photocopieur afin que chaque original soit agrafé avec sa copie. Cependant . suite à un défaut informatique, la photocopieuse a mélangé les originaux et les copies. L'entreprise décide donc de placer les n originaux et les n copies dans une boite. Une personne est alors chargée du travail suivant : elle pioche simultanément et au hasard 2 feuilles dans la boite. S'il s'agit d'un original et de sa copie, elle les agrafe et les sort de la boite. Sinon, elle repose les deux feuilles dans la, boite et elle recommence.
On modélise l'expérience par un espace probabilité $(\Omega, \mathcal{B}, \mathrm{P})$. Soit T_n la variable aléatoire égale au nombre de pioches qui sont nécessaires pour vider la boite lorsque celle-ci contient n originaux et n copies (soit $2n$ feuilles).
On considère l'événement A_n : « à l'issue de la première pioche, les deux feuilles piochées ne sont pas agrafées » et a_n sa probabilité c'est-à-dire que $a_n = \mathrm{P}(A_n)$.

1) Calculer a_n.

2) Étude de T_2. On suppose dans cette question que $n = 2$, c'est-à-dire que la boite contient deux originaux et deux copies.

- **a)** Montrer que pour tout entier $k \geqslant 2$: $\mathrm{P}(T_2 = k) = (1 - a_2)(a_2)^{k-2}$.
- **b)** Justifier que la variable $S_2 = T_2 - 1$ suit une loi géométrique dont on précisera le paramètre. En déduire l'espérance et la variance de T_2 en fonction de a_2

3) Étude de T_3. On suppose dans cette question que $n = 3$, c'est-à-dire que la boite, contient trois originaux et trois copies.

- **a)** Calculer $\mathrm{P}(T_3 = 2)$ puis $\mathrm{P}(T_3 = 3)$ en fonction de a_2 et a_3
- **b)** A l'aide du système complet d'événements $\left(A_3, \overline{A_3}\right)$ démontrer pour tout $k \geqslant 2$ que :

$$\mathrm{P}(T_3 = k+1) = (1 - a_3)\,\mathrm{P}(T_2 = k) + a_3\,\mathrm{P}(T_3 = k)$$

- **c)** Montrer que :

$$k \geqslant 2, \quad \mathrm{P}(T_3 = k) = \frac{(1 - a_2)(1 - a_3)}{a_3 - a_2}\left[(a_3)^{k-2} - (a_2)^{k-2}\right].$$

- **d)** Calculer $\displaystyle\sum_{k=2}^{+\infty} \mathrm{P}(T_3 = k)$.
- **e)** Prouver que la variable aléatoire $T_3 - 1$ admet une espérance et calculer $E(T_3 - 1)$. Donner la valeur de $E(T_3)$ en fonction de a_2 et a_3.

f) Établir que la variable aléatoire $T_3\,(T_3-1)$ admet une espérance et donner sa valeur en fonction de a_2 et a_3.
En déduire que T_3 admet une variance.

Exercice 2

On désigne par n un entier naturel supérieur ou égal à 2. On note p un réel de $]0;1[$ et on pose $q=1-p$.
On dispose d'une pièce donnant "Pile" avec la probabilité p et "Face" avec la probabilité q.
On lance cette pièce et on arrête les lancers dans l'une des deux situations suivantes :
• Soit si l'on a obtenu "Pile" .
• Soit si l'on a obtenu n fois "Face".
Pour tout entier naturel k non nul, on note P_k (respectivement F_k l'événement « on obtient "Pile" (respectivement "Face") au k^{e} lancer ».
On note T_n le nombre de lancers effectués, X_n le nombre de "Pile" obtenus et enfin Y_n le nombre de "Face" obtenus.
On admet que T_n , X_n et Y_n sont des variables aléatoires toutes les trois définies sur un espace probabilisé $(\Omega;\mathscr{A};P)$ que l'on ne cherchera pas à préciser.

1) Loi de T_n .

a) Pour tout k de $[\![1;n-1]\!]$, déterminer, en distinguant le cas $k=1$, la probabilité $P(T_n=k)$.

b) Déterminer $P\left(T_n=n\right)$.

c) Vérifier que $\displaystyle\sum_{k=1}^{n} P\left(T_n=k\right)=1$.

d) Établir que T_n possède une espérance et vérifier que $E\left(T_n\right)=\dfrac{1-q^n}{1-q}$.

2) Loi de X_n .

a) Donner la loi de X_n .

b) Vérifier que $E(X_n)=1-q^n$.

3) Loi de Y_n .

a) Déterminer, pour tout k de $[\![0;n-1]\!]$, la probabilité $P(Y_n=k)$.

b) Déterminer $P(Y_n=n)$.

c) Écrire une égalité liant les variables aléatoires T_n , X_n et Y_n , puis en déduire $E\left(Y_n\right)$.

Exercice 3

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 $\mathcal{S}_N$ l'ensemble des permutations possibles pour ce paquet de N cartes et on rappelle que $\text{card}(\mathcal{S}_N)=N!$
On se place dans un espace probabilisé $(\Omega,\mathscr{A},\mathbf{P})$ avec $\Omega=\mathcal{S}_N$, $\mathscr{A}=\mathscr{P}(\mathcal{S}_N)$ l'ensemble des parties de $\mathcal{S}_N$ et $\mathbf{P}$ 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 toute permutation σ de $\mathcal{S}_N$ la probabilité que le tas de cartes se trouve dans la configuration σ vaut $1/N!$

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 $\{1, \cdots, N\}$ la place à laquelle l'insertion a lieu, indépendamment des insertions précédentes.
Les instants successifs d'insertions seront notées $1, 2, \ldots, n, \ldots$, 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 $T_1(\omega) = 3$, $T_2(\omega) = 5$ et $T_3(\omega) = 6$ 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 ?

2) Loi de Δ_1.
Déterminer pour tout entier $n \in \mathbb{N}$, $\mathbf{P}(\Delta_1 > n)$ et reconnaître la loi de Δ_1.

3) Soit $i \in [\![2, N-1]\!]$. Loi de Δ_i.

a) Établir que pour tout entier $n \in \mathbb{N}$, on a $\mathbf{P}(\Delta_i > n) = \left(\dfrac{N-i}{N}\right)^n$. En déduire que Δ_i suit une loi usuelle que l'on précisera.

b) En déduire $E(\Delta_i) = \dfrac{N}{i}$ et $V(\Delta_i) = N\dfrac{N-i}{i^2}$.

4) Loi de T_2. Soit $n \geqslant 2$.

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

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

5) 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 ?

6) À l'instant T_3, la carte C_N est située en position $N-3$ 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) ?

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

i) on obtienne $(a_1, a_2, a_3) = (N-2, N-1, N)$?

ii) on obtienne $(a_1, a_2, a_3) = (N-2, N, N-1)$?

7) Justifier la phrase suivante :

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

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 C_N bien sûr !*

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

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

$$\forall n \geqslant 1 \quad H_n = \sum_{k=1}^{n} \frac{1}{k} \text{ et } u_n = H_n - \ln(n)$$

8) Espérance et variance de T

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

9) Étude de la suite (u_n)

a) Montrer que pour tout entier $k \geqslant 1$, on a $\frac{1}{k+1} \leqslant \int_k^{k+1} \frac{\mathrm{d}t}{t} \leqslant \frac{1}{k}$.

b) En déduire successivement :

i) la décroissance de la suite (u_n),

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

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

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

b) Quelle est la nature de la suite $\left(\dfrac{V(T)}{N^2}\right)_{N\in\mathbb{N}^*}$? *(on prendra garde au fait que $V(T)$ dépend de N).*
Justifier qu'il existe une constante α, strictement positive, telle que

$$V(T) \underset{N\to\infty}{\sim} \alpha N^2 \text{ et } V(T) \leqslant \alpha N^2$$

11) Écart à la moyenne
On rappelle l'inégalité de Bienaymé-Chebychev valable pour une variable aléatoire X admettant une espérance et une variance :

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

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

a) Justifier que $\forall \omega \in \Omega, \quad |T(\omega) - N\ln(N)| \leqslant |T(\omega) - E(T)| + N$.
Comparer par une inclusion les événements suivants

$$\Big(|T - N\ln(N)| \geqslant cN\Big) \text{ et } \Big(|T - E(T)| \geqslant N(c-1)\Big)$$

b) Démontrer que

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

où α a été définie à la question 10b.
Le nombre N étant fixé, que vaut $\lim\limits_{c\to+\infty} \mathbf{P}\Big(|T - N\ln N| \geqslant cN\Big)$?

12) Démontrer aussi que pour tout $\varepsilon > 0$:

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

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,* $52\ln(52) \simeq 205$.

Partie 3 - Distance variationnelle à la loi uniforme

Notations :

• On note π l'équiprobabilité sur $\mathcal{S}_N$, c'est-à-dire l'application de $\mathscr{P}(\mathcal{S}_N)$ dans $[0,1]$ telle que :

$$\forall A \subset \mathcal{S}_N \quad \pi(A) = \frac{\text{card}(A)}{N!}; \text{ en particulier, } \forall \sigma \in \mathcal{S}_N \quad \pi(\{\sigma\}) = \frac{1}{N!}$$

• On note également μ_n la probabilité sur $\mathcal{S}_N$ définie comme suit :
pour chaque configuration σ de $\mathcal{S}_N$, $\mu_n(\{\sigma\})$ 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 $\mathcal{S}_N$, $\mu_n(A) = \sum\limits_{\sigma\in A} \mu_n(\{\sigma\})$.
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 :

$$d(\mu_n, \pi) = \max\Big\{ |\mu_n(A) - \pi(A)|, \ A \subset \mathcal{S}_N \Big\}$$

13) Soient A une partie de $\mathcal{S}_N$, $n \in \mathbb{N}^*$ et E_n l'événement : "à l'instant n le paquet de cartes se trouve dans une configuration qui appartient à la partie A."

a) Expliquer, en utilisant la question 7, l'égalité suivante : Tr $\mathbf{P}_{(T\leqslant n)}(E_n) = \pi(A)$.
En déduire $\mathbf{P}(E_n \cap (T \leqslant n)) = \pi(A)\mathbf{P}(T \leqslant n)$.

b) Établir que $\mathbf{P}(E_n \cap (T > n)) \leqslant \mathbf{P}(T > n)$.

c) Montrer que

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

14) Soit A une partie de $\mathcal{S}_N$ 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 $\mu_n(A)$ et $\pi(A)$.

b) Déduire des questions précédentes la majoration :

$$|\mu_n(A) - \pi(A)| \leqslant \mathbf{P}(T > n)$$

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

Partie 4- Une majoration de $\mathbf{P}(T > n)$

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 $j \in [\![1, N]\!]$, 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 $(S_k)_{k\in[\![1,N]\!]}$ sont indépendantes.

16) Déterminer la loi de S_1.

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

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 $\mathbf{P}(T > n)$.

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

a) Exprimer l'événement $(S > m)$ à l'aide des événements $B_1^m,\ B_2^m,\ \ldots,\ B_N^m$.

b) Que vaut $\mathbf{P}(B_j^m)$ pour tout entier $j \in [\![1, N]\!]$?

c) On rappelle que pour tout entier $n \geqslant 2$ et pour toute famille d'événements $A_1, \ldots, A_n$, on a l'inégalité : $\mathbf{P}(\bigcup\limits_{i=1}^{n} A_i) \leqslant \sum\limits_{i=1}^{n} \mathbf{P}(A_i)$ En déduire $\mathbf{P}(S > m) \leqslant N\left(1 - \dfrac{1}{N}\right)^m$.

20) **a)** Montrer que $\ln(1+x) \leqslant x$ pour tout $x \in]-1, +\infty[$.

b) Déduire des résultats précédents la majoration

$$\forall m \in \mathbb{N},\ \mathbf{P}(T > m) \leqslant Ne^{-\frac{m}{N}}$$

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

b) *Application numérique.* On estime qu'une distance en variation à la loi uniforme de $0,2$ est acceptable.
Avec un jeu de 32 cartes, combien de *battages par insertions* doit-on faire pour considérer le paquet mélangé de façon acceptable ?

FIN DE L'ÉPREUVE