Pour le jeudi 4 février 2021

Devoir de Mathématiques numéro 5

Exercice 1

Une urne contient au départ 1 boule noire et 1 boule blanche. On effectue une suite de tirages qui consiste à tirer une boule de l'urne, regarder sa couleur, et la remettre dans l'urne en ajoutant une boule de la même couleur avant le tirage suivant. On cherche à déterminer l'évolution de la proportion de boules noires dans l'urne.
Pour tout $n \in \mathbb{N}$, on note X_n le nombre de boules noires à l'issue du n-ième tirage. En particulier $X_0 = 1$.
Pour tout $n \in \mathbb{N}^*$, on note A_n l'événement « tirer une boule noire lors du n-ième tirage ».

1) Quel est le nombre de boules lors du n-ième tirage ?

2) Déterminer les lois de X_0, X_1 et X_2.

3) Déterminer par récurrence la loi de X_n.

4) À l'aide de la loi des $(X_k)_{k\in\mathbb{N}}$, déterminer la probabilité de l'évènement A_n.

Exercice 2

Soit X_1 une variable aléatoire définie sur un espace probabilisé $(\Omega, \mathscr{A}, P)$ qui suit une loi de Poisson de paramètre $\lambda > 0$.

On définit sur $(\Omega, \mathscr{A}, P)$ la variable aléatoire X_2 par :

$$\forall \omega \in \Omega, \quad X_2(\omega) = \begin{cases} 0 \ \text{ si } \ X_1(\omega) = 0 \ \text{ ou si } \ X_1(\omega) \text{ est impair} \\ \dfrac{X_1(\omega)}{2} \quad \text{si } \ X_1(\omega) \text{ est pair et non nul} \end{cases}$$

1) Déterminer la loi de la variable aléatoire X_2.

2) Calculer l'espérance de la variable aléatoire X_2 et (5/2) l'exprimer à l'aide de fonctions usuelles.

Exercice 3 (Chaîne de Markov)

Chaque soir, pour dîner, vous avez 3 possibilités : vous faire des pâtes chez vous (A), aller au self (B), ou commander une pizza (C).

i. Si vous avez mangé chez vous la veille, la probabilité d'aller au self le lendemain est de 1/2, celle de commander une pizza 1/3, et donc celle de rester chez vous 1/6.

ii. Si vous êtes allé au self la veille, la probabilité de manger chez vous est 1/2 et celle de commander une pizza 1/4.

iii. Si vous avez commandé une pizza la veille, la probabilité de rester chez vous est de 1/3 et d'aller au self 1/3.

Pour tout $n \in \mathbb{N}$, on note X_n le type de repas pris le n-ième jour.
On suppose que le premier repas est pris au self.

1) Pour tout $n \in \mathbb{N}^*$, donner $X_n(\Omega)$.

2) Donner la loi de X_0.

3) Quelle est la probabilité $P(X_{n+1} = A | X_n = B)$? (probabilité de $X_{n+1} = A$ sachant $X_n = B$).

4) Traduire les conditions **i** à **iii** sous forme de probabilités conditionnelles.

5) Déterminer la probabilité de manger une pizza le jour $n = 1$, le jour $n = 2$.

6) Soit $n \in \mathbb{N}$. On pose $V_n = \begin{pmatrix} P(X_n = A) \\ P(X_n = B) \\ P(X_n = C) \end{pmatrix}$. Montrer qu'il existe $M \in \mathscr{M}_3(\mathbb{R})$ que l'on déterminera telle que

$$V_{n+1} = MV_n$$

7) Déterminer V_n en fonction de M, de n et de V_0.

Exercice 4 (Bonus : battre un paquet de cartes)

Rappel 1
Soit $N \in \mathbb{N}^*$. Une permutation de $[\![1, N]\!]$ est une bijection $\sigma : [\![1, N]\!] \to [\![1, N]\!]$ (cf. TP de cryptographie). Une permutation σ peut être identifiée à la liste $[\sigma(1), \ldots, \sigma(N)]$. Par exemple, la permutation $\sigma(1) = 2$, $\sigma(2) = 3$ et $\sigma(3) = 1$ de $[\![1, 3]\!]$ sera notée $[2, 3, 1]$.
L'ensemble des permutations de $[\![1, N]\!]$ sera noté $\mathcal{S}_N$. Card $\mathcal{S}_N = N!$

Introduction : 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 $\boxed{1}$ à $\boxed{N}$ et disposées en un paquet sur une table. Un joueur bat les cartes et repose le paquet sur la table. Le résultat du mélange, appelée configuration du paquet, est une permutation de ces N cartes (*cf.* exemple plus bas).
L'univers de cet exercice sera donc l'ensemble Ω des suites de configurations $(\sigma_0, \sigma_1, \sigma_2, ...)$ où $\sigma_n \in \mathcal{S}_N$ est la configuration du paquet de cartes à l'instant n.

Vocabulaire : Un paquet de N cartes est identifié à l'ensemble $[\![1, N]\!]$: les cartes sont numérotées, et on appelle une carte par son numéro. La position de la carte $\boxed{i}$ dans le paquet est notée C_i. Une carte située au sommet de la pile est dite *en position 1*, celle qui se trouve immédiatement en dessous est dite *en position 2*, etc. Ainsi une carte $\boxed{i}$ située en position $C_i = N$ désigne la carte située en bas de la pile.
Pour toute variable aléatoire X on notera $E(X)$ et $V(X)$ l'espérance et la variance de X lorsqu'elles existent.

État initial : Partons d'un tas de cartes rangées initialement dans l'ordre suivant :

Pour tout i élément de $[\![1, N]\!]$, la carte $\boxed{i}$ se trouve en position $C_i = i$.

Ainsi, à l'instant initial, la carte $\boxed{1}$ se trouve sur le dessus du paquet alors que $\boxed{N}$ se trouve donc tout en dessous du paquet.

But : On considère qu'un paquet est *convenablement mélangé* à l'instant n 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 $1/N!$

Étapes : 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 $[\![1, N]\!]$ la place à laquelle l'insertion a lieu, <u>indépendamment</u> des insertions précédentes.
Les instants successifs d'insertions seront notées $1, 2, \ldots, n, \ldots$ et l'instant initial est $n = 0$.

Notations : Nous notons :

- Pour $n \in \mathbb{N}^*$, X_n est la place où l'on insère la carte $\boxed{i}$ prise sur le dessus du paquet, à l'instant n.
- 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 $\boxed{N}$ se trouve remontée de la position $C_N = N$ à la position $C_N = N-1$,
- T_2 le premier instant où la carte $\boxed{N}$ se trouve remontée en position $C_N = N-2$,
- et plus généralement, pour i dans $[\![1, N-1]\!]$, T_i le premier instant où la carte $\boxed{N}$ atteint la position $C_N = N-i$.
- On posera également $\Delta_1 = T_1$ et $\forall\ i \in [\![2, N-1]\!]$, $\Delta_i = T_i - T_{i-1}$.
- Enfin, on notera $T = T_{N-1} + 1$.

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 X_n d'insertions.
La colonne n figure la configuration du paquet à l'instant n.

	instant n	0	1	2	3	4	5	6	7
	insertion en place X_n		3	2	4	1	3	4	2
Configuration	position 1	$\boxed{1}$	$\boxed{2}$	$\boxed{3}$	$\boxed{2}$	$\boxed{2}$	$\boxed{1}$	$\boxed{4}$	$\boxed{2}$
du	position 2	$\boxed{2}$	$\boxed{3}$	$\boxed{2}$	$\boxed{1}$	$\boxed{1}$	$\boxed{4}$	$\boxed{2}$	$\boxed{4}$
paquet	position 3	$\boxed{3}$	$\boxed{1}$	$\boxed{1}$	$\boxed{4}$	$\boxed{4}$	$\boxed{2}$	$\boxed{3}$	$\boxed{3}$
	position 4	$\boxed{4}$	$\boxed{4}$	$\boxed{4}$	$\boxed{3}$	$\boxed{3}$	$\boxed{3}$	$\boxed{1}$	$\boxed{1}$

Partie 1 (Mise en place)

1) Soit $n \in \mathbb{N}^*$. Donner $X_n(\Omega)$ et la loi de X_n.

2) Exemple : Dans l'exemple ω ci-dessus, donner les valeurs de
- $X_n(\omega)$ pour tout $n \in [\![1, 7]\!]$;
- $T_i(\omega)$ pour $i \in [\![1, N-1]\!]$;
- $\Delta_i(\omega)$ pour $i \in [\![1, N-1]\!]$;
- $T(\omega)$.

3) 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 ?

4) Loi de $\Delta_1 = T_1$.
Déterminer, pour tout $n \in \mathbb{N}$, $\mathbb{P}(\Delta_1 = n)$ et reconnaître la loi de Δ_1. Commencer par préciser $\Delta_1(\Omega)$.
Indication : *On pensera à décrire l'événement* $(\Delta_1 = n)$ *à l'aide des* X_k.

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

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

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

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

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

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

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

7) À l'instant T_2, la carte $\boxed{N}$ 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 ?

8) À l'instant T_3, la carte $\boxed{N}$ 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) <u>Quelques exemples.</u> 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 $(a_1,a_2,a_3)=(N-2,N,N-1)$?

9) 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 $\boxed{N}$ bien sûr !*

Partie 2 (Estimation du nombre d'insertions pour bien mélanger les cartes)
Notations : On introduit la suite $(H_n)_{n\geqslant 1}$ et $(u_n)_{n\in\mathbb{N}^*}$ définie par :

$$\forall n\geqslant 1 \qquad H_n=\sum_{k=1}^{n}\frac{1}{k}$$

On admet (exercice classique d'analyse, comparaison série-intégrale) qu'il existe $\gamma\in[0,1]$ appelée constante d'Euler telle que

$$H_n=\ln n+\gamma+o(1)$$

1) <u>Espérance et variance de T</u>

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

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

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

3) <u>Écart à la moyenne</u>
<u>Indication</u> : *Pensez à l'inégalité de Bienaymé-Chebychev.*
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

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

où α a été définie à la question 2b.

Le nombre N étant fixé, que vaut $\lim_{c\to+\infty} \mathbb{P}\Big(|T - N\ln N| \geqslant cN\Big)$?

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

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

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