Université de Tours — **M1 Mathématiques**
Année 2019-2020 — **Probabilités 1**

Contrôle continu 2
Durée de l'épreuve : 4 heures

Le sujet est long et il est normal de ne pas traiter toutes les questions dans le temps imparti. Les exercices sont essentiellement indépendants. On pourra toutefois utiliser les résultats de l'exercice 3 dans les suivants. Vous pouvez admettre une question et traiter les suivantes. Les réponses devront être étayées par des arguments construits. Bon travail !

Exercice 1

Doudou le hamster paresseux ne connaît que trois endroits dans sa cage : les copeaux où il dort, la mangeoire où il mange et la roue où il fait de l'exercice. Ses journées sont assez semblables les unes aux autres, et son activité se représente aisément par une chaîne de Markov. Toutes les minutes, il peut soit changer d'activité, soit continuer celle qu'il était en train de faire. L'appellation *processus sans mémoire* n'est pas du tout exagérée pour parler de Doudou.

- Quand il dort, il a 9 chances sur 10 de ne pas se réveiller la minute suivante.
- Quand il se réveille, il y a 1 chance sur 2 qu'il aille manger, et 1 chance sur 2 qu'il aille faire de l'exercice.
- Le repas ne dure qu'une minute, après il part faire autre chose.
- Après avoir mangé, il y a 3 chances sur 10 qu'il aille faire un peu de sport dans sa roue, mais surtout 7 chances sur 10 qu'il retourne dormir.
- Courir est fatigant : il a donc 8 chances sur 10 de retourner dormir au bout d'une minute, sinon il continue de courir, en oubliant qu'il est déjà un peu fatigué.

Combien d'heures Doudou passe-t-il en moyenne à dormir par jour ?

Exercice 2

On considère la chaîne de Markov $(X_n)_{n\geq 0}$ sur $E = \{1,2,3,4,5\}$ de matrice de transition P donnée par

$$P = \begin{pmatrix} 1/2 & 1/4 & 1/4 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 3/8 & 5/8 & 0 & 0 \\ 0 & 0 & 1/3 & 1/3 & 1/3 \\ 0 & 0 & 1/2 & 1/4 & 1/4 \end{pmatrix}.$$

1) Dessiner le graphe de la chaîne de Markov X.
2) Déterminer les classes de communication de P.
3) On note $T_3 = \inf\{n \geq 0 \ : \ X_n = 3\}$. Déterminer $\mathbb{E}_4(T_3)$ et $\mathbb{E}_5(T_3)$.
4) Déterminer toutes les mesures de probabilité invariantes de X.
5) La suite $(Z_n)_{n\geq 1}$ définie par

$$\forall n \geq 1, \quad Z_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$

a-t-elle une limite quand $n \to \infty$? Si oui, quelle est-elle ?

Exercice 3

Soit E un espace d'états dénombrable et $\mathcal{P}(E)$ l'ensemble des mesures de probabilités sur E. On rappelle la définition de la distance en variation totale sur $\mathcal{P}(E)$:

$$\forall (\mu, \tilde{\mu}) \in \mathcal{P}(E)^2, \quad d_{\mathrm{VT}}(\mu, \tilde{\mu}) = \frac{1}{2} \sup_{f : E \to [-1,1]} \left(\int f d\mu - \int f d\tilde{\mu} \right).$$

1) Démontrer que pour toute fonction $f : E \to [-1, 1]$,

$$\int f d\mu - \int f d\tilde{\mu} \leq \sum_{x \in E} |\mu(x) - \tilde{\mu}(x)|.$$

2) En déduire que

$$d_{\mathrm{VT}}(\mu, \tilde{\mu}) = \frac{1}{2} \sum_{x \in E} |\mu(x) - \tilde{\mu}(x)|.$$

3) Démontrer que

$$d_{\mathrm{VT}}(\mu, \tilde{\mu}) \leq \inf_{(X, \tilde{X})} \mathbb{P}(X \neq \tilde{X})$$

où l'infimum porte sur tous les couples de variables aléatoires sur $E \times E$ de lois marginales μ et $\tilde{\mu}$.

Remarque. L'inégalité ci-dessus est en fait une égalité mais il n'est pas demandé de le démontrer ici.

Exercice 4

Un professeur consciencieux mais imparfait (qui ne l'est pas ?) rédige un poly pour son cours sur les chaînes de Markov. Il le remanie régulièrement et tente d'en corriger les coquilles. À chaque étape, il corrigera donc certaines erreurs (d'autres échapperont à sa relecture) mais est également susceptible d'en ajouter. Lors d'une relecture, chaque coquille est corrigée avec une probabilité donnée, indépendamment des autres et un nombre aléatoire de coquilles est ajouté.

Soit $p \in]0, 1[$ et $\lambda > 0$. On note X_n le nombre de coquilles après n relectures. On introduit les variables aléatoires suivantes :

- X_0 de loi μ_0 sur $\mathbb{N}$,
- $(Y_{n,k})_{n,k \geq 1}$ de loi de Bernoulli de paramètre p,
- $(Z_n)_{n \geq 1}$ de loi de Poisson de paramètre λ :

$$\forall l \in \mathbb{N}, \quad \mathbb{P}(Z_n = l) = e^{-\lambda} \frac{\lambda^l}{l!}.$$

On suppose que toutes ces variables aléatoires sont indépendantes. On pose alors, pour tout $n \in \mathbb{N}$,

$$X_{n+1} = \sum_{k=1}^{X_n} Y_{n+1,k} + Z_{n+1}.$$

Si T est une variable aléatoire sur $\mathbb{N}$, on note G_T sa fonction génératrice définie par

$$\forall s \in [-1, 1], \quad G_T(s) = \mathbb{E}\big(s^T\big) = \sum_{n=0}^{\infty} \mathbb{P}(T = n) s^n.$$

1) Démontrer que la fonction génératrice d'une variable aléatoire à valeurs dans $\mathbb{N}$ est bien définie sur $[-1,1]$. Justifier l'affirmation suivante : deux variables aléatoires à valeurs dans $\mathbb{N}$ ont même loi si et seulement si elles ont même fonction génératrice.
2) Démontrer que si X et Y sont deux variables aléatoires indépendantes à valeurs dans $\mathbb{N}$ alors $G_{X+Y} = G_X G_Y$.
3) Déterminer les fonctions génératrices de variables aléatoires de loi binomiale $\mathcal{B}(n,p)$ et de loi de Poisson $\mathcal{P}(\lambda)$.
4) Faire le lien entre la situation concrète et le modèle mathématique. En particulier, que représente le paramètre p ? On admet dans la suite que $(X_n)_{n\in\mathbb{N}}$ est une chaîne de Markov. On ne cherchera pas à expliciter son noyau de transition.
5) La chaîne de Markov $(X_n)_{n\in\mathbb{N}}$ est-elle irréductible ? apériodique ?
6) Démontrer que si $X_0 = x$ alors

$$\forall s \in [-1,1], \quad G_{X_1}(s) = e^{\lambda(s-1)}(ps+1-p)^x.$$

7) Établir la relation de récurrence

$$\forall s \in [-1,1], \quad G_{X_{n+1}}(s) = e^{\lambda(s-1)} G_{X_n}(ps+1-p).$$

En déduire l'expression de G_{X_n} si $X_0 = x$.
8) À l'aide de la question 6, montrer que si X_0 suit la loi de Poisson de paramètre α alors X_1 suit la loi de Poisson de paramètre $p\alpha + \lambda$.
9) En déduire que la chaîne $(X_n)_{n\geq 0}$ admet une loi de Poisson dont précisera le paramètre pour loi invariante. En a-t-elle d'autres ?
10) Soit $x, \tilde{x} \in \mathbb{N}$ tels que $x \leq \tilde{x}$. Notons μ_n (resp. $\tilde{\mu}_n$) la loi de X_n lorsque $X_0 = x$ (resp. $X_0 = \tilde{x}$). Grâce aux questions 2 et 7, établir que $\tilde{\mu}_n$ est la loi de la somme de deux variables aléatoires indépendantes de lois respectives μ_n et $\mathcal{B}(\tilde{x}-x, p^n)$.
11) En déduire que

$$d_{\mathrm{VT}}(\mu_n, \tilde{\mu}_n) \leq \mathbb{P}(D_n \neq 0)$$

où D_n est une variable aléatoire de loi $\mathcal{B}(\tilde{x}-x, p)$.
12) Soit $m \in \mathbb{N}^*$. En étudiant la fonction $a \in [0,1] \mapsto (1-a)^m + am$, établir que

$$1-(1-a)^m \leq ma.$$

Conclure que, pour tout $n \in \mathbb{N}$,

$$d_{\mathrm{VT}}(\mu_n, \tilde{\mu}_n) \leq (\tilde{x}-x)p^n.$$

Exercice 5

Soit $d \in \mathbb{N}$ et $E = \{0,1\}^d$. Si $x \in E$, la coordonnée $i \in \{1,2,\ldots,d\}$ de x est notée $x(i)$. On dit que deux éléments x et y de E sont voisins (on note $x \sim y$) s'ils diffèrent d'une coordonnée exactement. Soit $U = (U_n)_{n\geq 1}$ et $B = (B_n)_{n\geq 1}$ des variables aléatoires indépendantes telles que, pour $n \geq 1$, U_n et B_n soient de lois uniformes sur $\{1,2,\ldots,d\}$ et $\{0,1\}$ respectivement. Soit X_0 une variable aléatoire à valeurs dans E indépendante de (U,B). Pour $n \in \mathbb{N}$, on définit X_{n+1} en posant

$$X_{n+1}(i) = \begin{cases} B_{n+1} & \text{si } i = U_{n+1}, \\ X_n(i) & \text{sinon.} \end{cases}$$

1) Montrer que X est une chaîne de Markov sur E de matrice de transition P donnée par

$$\forall (x,y) \in E^2, \quad P(x,y) = \begin{cases} \frac{1}{2} & \text{si } y = x, \\ \frac{1}{2d} & \text{si } y \sim x, \\ 0 & \text{sinon.} \end{cases}$$

2) La chaîne est-elle irréductible ? récurrente ? apériodique ?

3) La matrice P possède-t-elle une unique mesure de probabilité invariante ? Si oui, la déterminer.

4) Soit T la variable aléatoire à valeurs dans $\mathbb{N}$ définie par

$$T = \inf \{n \geq 1 \ : \ \{U_1, U_2, \ldots, U_n\} = \{1, 2, \ldots, d\}\}.$$

a) Démontrer que, pour tout $n \in \mathbb{N}$,

$$\{T > n\} \subset \bigcup_{i=1}^{d} E_{n,i} \quad \text{où} \quad E_{n,i} = \{U_1 \neq i, \ldots, U_n \neq i\},$$

puis que, pour $1 \leq i \leq d$,

$$\mathbb{P}(E_{n,i}) = \left(1 - \frac{1}{d}\right)^n.$$

b) En déduire que, pour tout $n \in \mathbb{N}$,

$$\mathbb{P}(T > n) \leq d e^{-n/d}.$$

c) À quel autre modèle classique cette variable aléatoire T fait-elle penser ?

5) Soit $\tilde{X}_0$ une variable aléatoire à valeurs dans E indépendante des suites U et B (mais pas nécessairement de X_0). On définit la suite $(\tilde{X}_n)_n$ de manière analogue à $(X_n)_n$ en remplaçant X_0 par $\tilde{X}_0$. Démontrer que

$$\forall n \geq T, \quad X_n = \tilde{X}_n.$$

En déduire une majoration de la distance en variation totale entre μP^n et $\tilde{\mu} P^n$ valable pour toutes mesures de probabilités μ et $\tilde{\mu}$ sur E.

Exercice 6

Soit

$$K = \begin{pmatrix} 1 & 1 & -1 \\ 1 & 2 & 0 \\ -1 & 0 & 3 \end{pmatrix}.$$

1) Montrer qu'il existe un vecteur gaussien (X, Y, Z) centré de covariance K. Sa loi admet-elle une densité par rapport à la mesure de Lebesgue sur $\mathbb{R}^3$?

2) Déterminer la loi de $U = X + 2Y - Z$.

3) Établir que $V = Y^2/2 + Z^2/3$ suit la loi $\chi^2(2)$.

4) Montrer que $(X + Y, Y - Z)$ est un vecteur gaussien dont on déterminera l'espérance et la matrice de covariance et une densité de sa loi par rapport à la mesure de Lebesgue sur $\mathbb{R}^2$.