

## Modèles markoviens en biologie

### Corrigé de l'examen du 20 avril 2007

Barème : un point pour chaque question, soit un total de 35 points, seuillé à 20 pour obtenir la note.

**Exercice 1** (Mort ou vif...). Cet exercice est un clone du premier exercice de l'examen de Rennes qui a eu lieu fin mars. L'énoncé a toutefois subi quelques mutations.

Les taux étant linéaires, il est possible d'établir que si la loi initiale est Poisson-binomiale, alors il en va de même à tout instant. Il est possible d'obtenir une formule explicite pour la loi à tout instant, comme pour la file d'attente M/M/∞.

1. L'interprétation en terme d'horloges exponentielles en compétition indique que lorsque  $X_t = i$ , le temps supplémentaire passé en  $i$  suit la loi  $\mathcal{E}(i(\lambda + \mu))$ , suite à quoi un individu se dédouble avec probabilité  $\lambda/(\lambda + \mu)$  et meurt avec probabilité  $\mu/(\lambda + \mu)$ , d'où la forme du générateur ;
2. L'état 0 est absorbant car  $L(0,0) = 0$  (pas de génération spontanée!). Presque-sûrement, les trajectoires de  $(X_t)_{t \geq 0}$  sont constituées d'une succession de plateaux ouverts à droite et fermés à gauche. Les variations successives de hauteur sont dans  $\{-1, +1\}$ . Si un plateau est de hauteur 0, alors sa longueur est infinie. Les lois des hauteurs et des longueurs correspondent à la description en terme d'horloges exponentielles en compétition. Lorsque  $\lambda = 0$  et  $\mu > 0$ ,  $(X_t)_{t \geq 0}$  est un processus de mort pur, qui est presque-sûrement absorbé par 0. Lorsque  $\mu > 0$  et  $\lambda = 0$ ,  $(X_t)_{t \geq 0}$  est un processus de vie pur, qui tend presque-sûrement vers  $+\infty$ . Dans les deux cas, les états non nuls sont tous transitoires ;
3. Pour tout  $i \in \mathbb{N}$ , on a  $Lf(i) = i\lambda(f(i+1) - f(i)) + i\mu(f(i-1) - f(i)) = 0$ . Par conséquent,  $\mathbf{L}f = 0$  et en notant  $\alpha(t) = P_t f(i)$ , on a  $\alpha'(t) = \partial_t P_t f(i) = P_t Lf(i) = 0$ . Il en découle que  $P_t f(i) = P_0(i)f$ , ou encore  $\mathbb{E}(X_t | X_0 = i) = i$ , ou encore  $\mathbb{E}(X_t) = \mathbb{E}(X_0)$  pour tout  $t \geq 0$ . De même,  $Lg(i) = i\lambda(g(i+1) - g(i)) + i\mu(g(i-1) - g(i)) = 2i\lambda$ . On a donc  $Lg = 2\lambda f$ , et en notant  $\beta(t) = P_t g(i)$ , on obtient  $\beta'(t) = \partial_t P_t g(i) = P_t Lg(i) = 2\lambda P_t f(i) = 2\lambda \alpha(t)$ , d'où  $\beta(t) = 2\lambda t \mathbb{E}(X_0) + \beta(0)$ . Il ne reste plus qu'à utiliser le fait que  $\text{Var}(X_t) = \beta(t) - \mathbb{E}(X_t)^2$  pour obtenir la formule désirée pour la variance ;
4. Si  $L_N$  est le générateur infinitésimal de  $(X^N)_{t \geq 0}$  alors  $NL_N$  est le générateur infinitésimal de  $(X_{Nt}^N)_{t \geq 0}$ . Donc le générateur infinitésimal  $A_N$  de  $(N^{-1}X_{Nt}^N)_{t \geq 0}$  est donné par

$$\forall i \in \mathbb{N}, \quad A_N f(i/N) = iN\lambda_N(f((i+1)/N) - f(i/N)) + iN\mu_N(f((i-1)/N) - f(i/N)).$$

Fixons  $y \in \mathbb{R}_*^+$  et  $f : \mathbb{R}_+ \rightarrow \mathbb{R}$   $\mathcal{C}^2$  et posons  $y_N = N^{-1}\lfloor Ny \rfloor$ . On a alors

$$\begin{aligned} A_N f(y_N) &= N^2 y_N \lambda_N (f(y_N + 1/N) - f(y_N)) + N^2 y_N \mu_N (f(y_N - 1/N) - f(y_N)) \\ &= \lambda_N y_N f''(y_N) + o(1/N). \end{aligned}$$

Pour que  $A_N f(y_N)$  converge vers  $Af(y)$ , il faut et il suffit que  $\lambda_N \rightarrow a$  quand  $N \rightarrow \infty$ . Le processus de diffusion limite  $Y$  est solution de l'équation différentielle stochastique

$$Y_t = Y_0 + \int_0^t \sqrt{2a Y_s} dB_s.$$

5. En prenant l'espérance des deux membres dans l'expression donnée ci-dessus, on obtient  $\mathbb{E}(Y_t) = \mathbb{E}(Y_0)$ . D'autre part, en appliquant la formule d'Itô, on obtient

$$Y_t^2 = y^2 + \int_0^t aY_s ds + \int_0^t 2\sqrt{2a}Y_s^2 dB_s.$$

En espérance, cela conduit à une équation différentielle pour  $t \mapsto \mathbb{V}\text{ar}(Y_t|Y_0 = y)$  dont la solution est précisément  $\mathbb{V}\text{ar}(Y_t|Y_0 = y) = yat$ . Le résultat désiré s'en déduit ;

6. La fonction d'échelle  $S$  est de la forme  $S(x) = x - c$ , d'où  $\mathbb{P}_y(Y_T = r) = (y - l)/(r - l)$ , qui ne dépend pas de  $\lambda$ , et qui tend vers 0 lors  $r \rightarrow +\infty$ .

**Exercice 2** (Développement et déclin d'une population de plantes). Cet exercice est inspiré des processus de Oakes-Hawkes, qui peuvent être vus comme des processus ponctuels autorégressifs, cf. [HO74] et [DVJ03]. Il nécessite une connaissance raisonnable du chapitre sur les processus ponctuels des notes de cours [CM07].

1. Si  $S$  est concentrée autour de 0 alors sa translation  $\delta_x * S$  est concentrée autour de  $x$ . Il est naturel que les plantes filles d'une plante mère en  $x$  soient concentrées autour de  $x$ . D'autre part, la loi de Poisson convient bien à la modélisation des événements discrets rares, ce qui correspond bien à la faible chance de germination des graines ;
2. D'après les hypothèses faites sur les plantes, on a  $H_x = \sum_{k=1}^Z Y_k$  où  $Z$  est une v.a. de loi  $P$  et  $(Y_k)_{k \in \mathbb{N}}$  une suite de v.a. i.i.d. de loi  $\delta_x * S$ , indépendante de  $Z$ . En vertu d'un théorème du cours,  $H_x$  est donc un processus ponctuel de Poisson sur  $\mathbb{R}^2$  de mesure d'intensité  $m(\delta_x * S)$  qui affecte la masse  $mS(A - x)$  à tout ensemble mesurable  $A$  de  $\mathbb{R}^2$ . Cette mesure d'intensité  $m(\delta_x * S)$  est finie car  $S$  est une loi de probabilité. Aussi, l'espace  $\mathbb{R}^2$  contient un nombre d'atomes qui suit la loi de Poisson  $\mathcal{P}(m)$  de moyenne  $m$ . La masse affectée par la mesure d'intensité à l'espace  $\mathbb{R}^2$  est  $m$ . Il est utile de retenir que l'ensemble des atomes d'un processus ponctuel de Poisson de mesure d'intensité finie  $\mu$  sur un espace  $E$  peut être vue comme un échantillon de loi  $\mu/\mu(E)$  dont la taille est aléatoire de loi  $\mathcal{P}(\mu(E))$ , et réciproquement tout échantillon... ;
3. D'après les hypothèses faites sur les plantes, une plante mère située en  $x$  fait des filles selon un processus ponctuel  $H_x$ . Les hypothèses d'indépendance font le reste. Le cas  $|X_n| = 0$  correspond à une population vide à la génération  $n$  (extinction) ;
4. Conditionnellement à  $X_n$ , la variable aléatoire  $X_{n+1}$  est une superposition de processus ponctuels de Poisson indépendants et constitue donc un processus ponctuel de Poisson dont la mesure d'intensité est la somme des mesures d'intensité :  $\sum_{x \in X_n} m(\delta_x * S) = m(X_n * S)$ . Conditionnellement à  $X_n$ , la variable aléatoire  $|X_{n+1}| = \sum_{x \in X_n} |H_x|$  suit une loi de Poisson  $\mathcal{P}(|X_n|m)$ . Par récurrence, on obtient  $\mathbb{E}(|X_{n+1}|) = \mathbb{E}(X_{n+1}(\mathbb{R}^2)) = \Lambda_0(\mathbb{R}^2)m^n$  pour tout entier  $n \geq 0$ . Ainsi, la taille moyenne de la population décroît exponentiellement vite vers 0 si  $m < 1$ , est constante si  $m = 1$ , et explose exponentiellement vite vers  $+\infty$  si  $m > 1$  ;
5. Pour toute fonction  $f : \mathbb{R}^2 \rightarrow \mathbb{R}_+$  mesurable positive, la structure de  $X_{n+1}$  sachant  $X_n$  donne

$$\begin{aligned} \mathbb{E}\left(\exp\left(-\int_{\mathbb{R}^2} f dX_{n+1}\right) \middle| X_n\right) &= \exp\left(-m \int_{\mathbb{R}^2} (1 - e^{-f}) d(X_n * S)\right) \\ &= \exp\left(-m \int_{\mathbb{R}^2} \left(\int_{\mathbb{R}^2} (1 - e^{-f(x+y)}) dS(y)\right) dX_n(x)\right) \\ &= \exp\left(-\int_{\mathbb{R}^2} g(x) dX_n(x)\right). \end{aligned}$$

À présent, si  $X_n$  était un processus ponctuel de Poisson sur  $\mathbb{R}^2$  d'intensité  $\Lambda_n$ , alors

$$\begin{aligned} \mathbb{E}\left(\exp\left(-\int_{\mathbb{R}^2} f dX_{n+1}\right)\right) &= \mathbb{E}\left(\mathbb{E}\left(\exp\left(-\int_{\mathbb{R}^2} f dX_{n+1}\right)\middle|X_n\right)\right) \\ &= \mathbb{E}\left(\exp\left(-\int_{\mathbb{R}^2} g dX_n\right)\right) \\ &= \exp\left(-\int_{\mathbb{R}^2} (1 - e^g) d\Lambda_n\right) \end{aligned}$$

qui n'est pas de la forme  $\exp\left(-\int_{\mathbb{R}^2} (1 - e^f) d\Lambda_{n+1}\right)$  pour une certaine mesure  $\Lambda_{n+1}$  ;

6. Conditionnellement à  $|X_n|$ , la variable aléatoire  $|X_{n+1}|$  est la somme de  $|X_n|$  variables aléatoires i.i.d. de loi  $\mathcal{P}(m)$ . Ceci constitue précisément la définition d'un processus de branchement de Galton-Watson de loi de reproduction  $\mathcal{P}(m)$ . Par conséquent, si  $m < 1$ , alors le processus s'éteint presque-sûrement avec  $\mathbb{P}(|X_n| > 0) < m^n$ . Si  $m = 1$ , alors la population s'éteint également presque-sûrement avec  $\mathbb{P}(|X_n| > 0) \sim 2/(nm^2)$ . En revanche, si  $m > 1$ , alors la probabilité d'extinction est strictement plus petite que 1, et lorsqu'il n'y a pas extinction, la population explose exponentiellement vite et la suite  $(m^{-n}X_n)$  converge en loi sur  $\mathbb{R}_+$  sans atome en 0 ;
7. Le plan  $\mathbb{R}^2$  peut être remplacé par tout espace localement compact à base dénombrable, plus généralement, par tout espace pour lequel la mécanique des processus ponctuels de Poisson est licite. Les résultats ci-dessus resteront identiques. Pour le modèle des plantes, on pourrait remplacer  $\mathbb{R}^2$  par la sphère  $\mathbb{S}^2$ . Notons qu'une restriction de support sur  $\Lambda_0$  et  $S$  ne suffit pas à cantonner la population à une zone donnée. Il faudrait remplacer  $\delta_x * S$  par une loi qui dépend de  $x$  mais ne sort pas d'une zone déterminée à l'avance ;
8. Les plantes ainsi produites sont réparties dans l'espace selon un processus ponctuel de Poisson sur  $\mathbb{R}^2$  de mesure d'intensité  $p\Theta$ , en vertu de la stabilité des processus ponctuels de Poisson par amincissement ;
9. Une manière de prendre en compte ces informations dans le modèle consisterait par exemple à faire dépendre  $P$  de la position de la plante mère  $x$  et de sa génération  $n$ .

**Exercice 3** (Grégarisme et individualisme des oiseaux du pacifique). Cet exercice traite de quelques aspects de la distribution d'Ewens, au moyen d'une chaîne de Markov sur les partitions  $(B_n)$  appelée parfois « processus des restaurants chinois de Pitman », voir par exemple [JKB97], [Kin78], [Ewe72], et le passage sur la formule d'Ewens dans les notes de cours [CM07, ch. 5]. Cependant, la connaissance du chapitre concerné des notes de cours n'est absolument pas requise puisque tout est fait pas à pas.

1. On a  $B \in \mathcal{B}_n$  et donc  $\sum_{b \in B} |b| = |B| = n$  ;
2. Découle de la description du modèle dans l'énoncé et de la question précédente ;
3.  $(B_n)_{n \geq 1}$  est une chaîne de Markov d'espace d'état  $\cup_{n=1}^{\infty} \mathcal{B}_n$  car le mécanisme de transition est sans mémoire : la loi de  $B_{n+1}$  sachant  $B_1, \dots, B_n$  ne dépend que de  $B_n$ . Cependant, le mécanisme de transition dépend du temps  $n$ , et cette chaîne est donc inhomogène ;
4. Le cas limite  $\theta = 0$  empêche toute colonisation d'îlot nouveau et les oiseaux s'installent tous sur l'îlot choisi par le premier arrivant. La probabilité  $\mathbb{P}(B_{n+1} = B' | B_n = B)$  est monotone en  $\theta$ . Pour cette raison,  $\theta$  est interprété comme un coefficient d'individualisme ;
5. Simple raisonnement par récurrence sur  $n$  basé sur l'expression de la loi conditionnelle de  $B_{n+1}$  sachant  $B_n$  ;

6. Fixons  $n$  et  $a_1, \dots, a_n$  avec  $a_1 + 2a_2 + \dots + na_n = n$ . Le nombre de façon de regrouper  $n$  objets en  $a_1$  blocs de taille 1,  $a_2$  blocs de taille 2 etc est donné par le coefficient multinomial  $n! / \prod_{k=1}^n (k!)^{a_k}$  car  $n = 1 + \dots + 1 + 2 + \dots + 2 + \dots + n + \dots + n$  où chaque entier  $k$  est répété  $a_k$  fois. Pour chaque  $1 \leq k \leq n$ , les  $a_k$  blocs de taille  $k$  sont permutables de  $a_k!$  façons différentes, d'où le résultat. Notons que  $a_k \leq \lfloor n/k \rfloor$  pour tout  $1 \leq k \leq n$ . En particulier  $a_n \in \{0, 1\}$  avec  $a_n = 1$  si et seulement si  $a_1 = \dots = a_{n-1} = 0$ ;
7. Tout d'abord,  $|B_n| \leq n$  pour tout  $n \geq 1$  par construction. De plus, pour tout  $B \in \mathcal{B}_n$ , la probabilité  $\mathbb{P}(B_n = B)$  ne dépend que de  $|B|$ . Par conséquent, la sommation sur l'ensemble des partitions  $B$  de taille  $n$  à blocs de tailles  $a_1, \dots, a_n$  prescrites revient à multiplier l'expression de  $\mathbb{P}(B_n = B)$  par le coefficient calculé à la question précédente. Une simple réécriture conduit alors au résultat ;
8. Pour tout  $n \geq 1$ , la loi de  $C_{n+1}$  sachant  $C_1, \dots, C_n$  ne dépend que de  $C_n$  par définition de la suite  $(B_n)_{n \geq 1}$ . Par conséquent,  $(C_n)_{n \geq 1}$  est bien une chaîne de Markov d'espace d'état  $E^n$  où  $E$  désigne l'ensemble des  $n$ -uplets  $(a_1, \dots, a_n)$  d'entiers tels que  $a_1 + 2a_2 + \dots + na_n = n$  ;
9. Sachant  $B_n$ , on a  $|B_{n+1}| - |B_n| \in \{0, 1\}$  et  $|B_{n+1}| = |B_n| + 1$  avec probabilité  $\theta/(\theta + n)$ . Il en découle que  $|B_n| = \xi_1 + \dots + \xi_n$  où  $\xi_1, \dots, \xi_n$  sont des variables aléatoires indépendantes de lois de Bernoulli avec  $\xi \sim \mathcal{B}(\theta/(\theta + \theta))$  pour tout  $1 \leq k \leq n$ . Les formules exprimant les deux premiers moments de  $|B_n|$  découlent de l'indépendance de  $\xi_1, \dots, \xi_n$  et du fait que  $\xi_k$  est de moyenne  $\theta/(\theta + k)$  et de variance  $\theta k/(\theta + k)^2$  pour tout  $1 \leq k \leq n$  ;
10. En vertu de la question précédente, la suite  $(|B_n|)_{n \geq 1}$  est presque-sûrement croissante, et ne fait que des sauts de 1. Elle est cependant de moins en moins croissante en quelque sorte puisque la quantité  $p_n = \mathbb{P}(|B_{n+1}| = |B_n| + 1) = \mathbb{P}(\xi_{n+1} = 1) = \theta/(\theta + n)$  décroît quand  $n$  croît. La suite  $(|B_n|)_{n \geq 1}$  est une chaîne de Markov car elle constitue la suite des sommes cumulées de variables aléatoires indépendantes. Elle n'est cependant pas homogène car ces variables ne sont pas de même loi. L'expression des deux premiers moments donnée par la question précédente entraîne que la moyenne et la variance de  $|B_n|$  sont équivalentes à  $\theta \log(n)$  lorsque  $n$  tend vers  $\infty$  ;
11. L'expression de  $\mathbb{E}(|B_n|)$  suggère d'estimer  $\theta$  avec la solution  $\hat{\theta}$  de l'équation

$$\mathbb{E}(|B_n|) = \sum_{i=0}^{n-1} \frac{\hat{\theta}}{\hat{\theta} + i}.$$

Cette équation en  $\hat{\theta}$  n'est malheureusement pas explicite. Cependant, l'expression de la loi de  $|B_n|$  montre que l'estimateur de  $\theta$  ainsi construit constitue l'estimateur de maximum de vraisemblance, et son caractère non explicite est décevant ;

12. On a

$$\mathbb{E}(|B_{n+1}|) - \mathbb{E}(|B_n|) = \mathbb{E}(\xi_{n+1}) = \frac{\theta}{\theta + n} = \frac{1}{1 + n\theta^{-1}}.$$

Pour  $n \geq 1$  **fixé**, cela fournit, à partir d'un échantillon de taille  $k$  de même loi que  $\xi_{n+1}$ , un estimateur  $\hat{t}_k$  de  $t_n = (1 + n\theta^{-1})^{-1}$ . Notons que l'échantillon de taille  $k$  correspond à  $k$  réalisations indépendantes de l'arrivée des  $n$  premiers oiseaux sur l'île. Le biais et la variance de l'estimateur découlent immédiatement de la moyenne et de la variance de  $\xi_{n+1}$  ;

13. Le cas  $a' = a + 1$  correspond au choix d'un îlot vide (i.e. de taille nulle), le cas  $a' = a - 1$  correspond au choix d'un îlot hébergeant un unique oiseau (i.e. de taille 1), le cas  $a' = a$  correspond au choix d'un îlot de taille plus grande que 2. Les autres cas sont impossibles ! La loi conditionnelle attendue n'est donc qu'une conséquence de celle déjà obtenue pour  $B_n$ . La loi de  $C_{n+1,1}$  sachant  $C_{1,1}, \dots, C_{n,1}$  ne dépend que de  $C_{n,1}$ . La suite  $(C_{n,1})_{n \geq 1}$  est une chaîne de Markov d'espace d'état  $\mathbb{N}$ , inhomogène car la loi conditionnelle dépend de  $n$  ;

14. La formule pour  $\mathbb{E}(C_{n+1,1} | C_{n,1} = a)$  est immédiate. Cela donne  $(n+\theta)m_{n+1} = (\theta+n-1)m_n + \theta$  où  $m_n$  désigne l'espérance de  $C_{n,1}$ . La formule annoncée pour  $m_n$  s'en déduit. Pour la variance, on procède de la même manière avec la formule

$$\text{Var}(X) = \mathbb{E}(\text{Var}(X | Y)) + \text{Var}(\mathbb{E}(X | Y)).$$

Les calculs sont cependant un peu plus lourds que pour l'espérance.

15. Les formules pour les deux premiers moments de  $C_{n,1}$  entraînent

$$\lim_{n \rightarrow \infty} \mathbb{E}(C_{n,1}) = 1 \quad \text{et} \quad \lim_{n \rightarrow \infty} \text{Var}(C_{n,1}) = \theta.$$

16. On a  $C_{n,n} = 1$  si et seulement si  $|B_n| = 1$  (un seul îlot regroupant les  $n$  oiseaux au temps  $n$ ). De plus  $C_{1,1} = 1$  et  $C_{n,n} \in \{0, 1\}$ , pour tout  $n \geq 1$ . On a  $\mathbb{P}(C_{n+1,n+1} = 1 | C_{n,n} = 0) = 0$  car il est impossible de fusionner des îlots non vides. D'autre part,  $\mathbb{P}(C_{n+1,n+1} = 0 | C_{n,n} = 1) = n/(n+\theta)$  par définition de  $(B_n)_{n \geq 1}$ . La suite  $(C_{n,n})_{n \geq 1}$  est presque-sûrement décroissante car  $(|B_n|)_{n \geq 1}$  est presque-sûrement croissante : l'arrivée d'un nouvel oiseau ne peut qu'augmenter le nombre d'îlots occupés. La loi de  $C_{n+1,n+1}$  sachant  $C_{1,1}, \dots, C_{n,n}$  ne dépend que de  $C_{n,n}$  et donc  $(C_{n,n})_{n \geq 1}$  est une chaîne de Markov inhomogène d'espace d'état  $\{0, 1\}$ . En particulier, l'expression des transitions donne la formule requise pour  $\mathbb{P}(C_{n,n} = 1)$ . Cela fait ici encore intervenir une suite de v.a. de Bernoulli indépendantes mais non équidistribuées ;
17. On remarque tout d'abord que comme  $\theta > 0$ , le produit infini  $\prod_{k=1}^{\infty} (1 + \theta n^{-1})$  diverge car la série harmonique diverge, et par conséquent  $\lim_{n \rightarrow \infty} \mathbb{P}(C_{n,n} = 1) = 0$ . D'autre part, la suite d'événements  $(A_n)_{n \geq 1}$  définie par  $A_n = \{C_{n,n} = 0\}$  est croissante car la suite  $(C_{n,n})_{n \geq 1}$  est presque sûrement décroissante. On peut donc écrire

$$\mathbb{P}\left(\lim_{n \rightarrow \infty} C_{n,n} = 0\right) = \mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n \rightarrow \infty} \mathbb{P}(A_n) = 1 - \lim_{n \rightarrow \infty} \mathbb{P}(C_{n,n} = 1) = 1.$$

18. Le modèle utilisé pour l'occupation des îlots par les oiseaux favorise de moins en moins l'apparition d'îlots peu occupés car la probabilité de rejoindre un îlot dépend de sa taille. Ce modèle favorise également le maintien d'îlots peu occupés ! Nous savons que pour  $n$  grand, le nombre d'îlots de taille 1 est en moyenne de l'ordre de  $\theta \log(n)$  (petit par rapport à  $n$ ) tandis que la probabilité qu'un unique îlot hébergeant tous les oiseaux décroît vers 0. La simplicité du modèle et les calculs explicites qu'il permet découlent du fait que la probabilité de rejoindre un îlot est proportionnelle à sa taille. La propriété d'additivité liée à la taille explique pourquoi plusieurs fonctions de  $(B_n)$  liées à la taille sont encore markoviennes. Il est possible de décider qu'à chaque temps  $n$  un couple d'îlots peut inter-agir en transférant un occupant vers l'îlot le moins occupé. Cela constitue une légère modification de la loi conditionnelle de  $B_{n+1}$  sachant  $B_n$ . Cependant, cela brise toutes les formules obtenues ! Ce nouveau modèle défavorise le maintien d'îlots de petite taille et peut même faire dériver la population vers un unique îlot. . .
19. Par récurrence sur  $n$ , on a  $0 \leq P_n \leq 1$  pour tout  $n$  car  $0 \leq W_n \leq 1$  pour tout  $n$ . Par récurrence sur  $n$ , on montre que  $1 - (P_1 + \dots + P_n) = (1 - W_1) \dots (1 - W_n)$  pour tout  $n \geq 1$ . Les variables aléatoires  $(U_n)_{n \geq 1}$  définies par  $U_n = 1 - W_n$  pour tout  $n \geq 1$  sont i.i.d. de loi à densité sur  $[0, 1]$ . En particulier  $\mathbb{E}(U_n) = \mathbb{E}(U_1) \in ]0, 1[$  pour tout  $n \geq 1$ . Or par l'inégalité de Markov  $\mathbb{P}(U_1 \dots U_n > \varepsilon) \leq \mathbb{E}(U_1)^n \varepsilon^{-1}$ , et par conséquent  $\sum_{n=1}^{\infty} \mathbb{P}(U_1 \dots U_n > \varepsilon) < \infty$  pour tout  $\varepsilon > 0$ . En vertu du lemme de Borel-Cantelli, il en découle que  $(U_1 \dots U_n)_{n \geq 1}$  converge presque-sûrement vers 0 ;

## Références

- [CM07] D. CHAFAÏ et F. MALRIEU – « Modèles markoviens en biologie », notes de cours de master mathématiques appliquées, <http://www.lsp.ups-tlse.fr/Chafai/m2r-markbio.html>, 2007.

- [DVJ03] D. J. DALEY et D. VERE-JONES – *An introduction to the theory of point processes. Vol. I*, second éd., Probability and its Applications (New York), Springer-Verlag, New York, 2003, Elementary theory and methods.
- [Ewe72] W. J. EWENS – « The sampling theory of selectively neutral alleles », *Theoret. Population Biology* **3** (1972), p. 87–112; erratum, *ibid.* 3 (1972), 240; erratum, *ibid.* 3 (1972), 376.
- [HO74] A. G. HAWKES et D. OAKES – « A cluster process representation of a self-exciting process », *J. Appl. Probability* **11** (1974), p. 493–503.
- [JKB97] N. L. JOHNSON, S. KOTZ et N. BALAKRISHNAN – *Discrete multivariate distributions*, Wiley Series in Probability and Statistics : Applied Probability and Statistics, John Wiley & Sons Inc., New York, 1997, , A Wiley-Interscience Publication.
- [Kin78] J. F. C. KINGMAN – « Random partitions in population genetics », *Proc. Roy. Soc. London Ser. A* **361** (1978), no. 1704, p. 1–20.