(*public 2008*)

Résumé : Une population bisexuée se reproduit selon un mécanisme qui généralise les processus de Galton–Watson classiques. Sous l'hypothèse que la loi de reproduction inter-générationnelle possède une propriété de sur-additivité, on établit que l'extinction de la population est quasi-certaine selon qu'une quantité intrinsèque au modèle, interprétée comme le taux moyen de croissance espéré, n'excède pas ou dépasse la valeur critique 1. Cette propriété est établie dans le cadre du modèle de Daley. Elle est illustrée par d'autres modèles sur-additifs.

Mots clefs : Processus de branchement, chaîne de Markov, loi des grands nombres, martingale, fonction analytique.

► *Il est rappelé que le jury n'exige pas une compréhension exhaustive du texte. Vous êtes laissé(e) libre d'organiser votre discussion comme vous l'entendez. Des suggestions de développement, largement indépendantes les unes des autres, vous sont proposées en fin de texte. Vous n'êtes pas tenu(e) de les suivre. Il vous est conseillé de mettre en lumière vos connaissances à partir du fil conducteur constitué par le texte. Le jury appréciera que la discussion soit accompagnée d'exemples traités sur ordinateur.*

1. Introduction

Le modèle de Daley s'attache à décrire le mécanisme de reproduction d'une population sexuée composée, à la génération n, de F_n femelles et de M_n mâles. La population de la génération suivante est issue du croisement de Z_n couples. Chaque couple donne naissance, selon la même loi de reproduction G, à X femelles et à Y mâles, indépendamment de l'indice de génération, ainsi que des autres couples. Toutes les variables aléatoires sont définies dans un espace de probabilité commun $(\Omega, \mathscr{F}, \mathbb{P})$.

Définition 1. *Soit une suite $(X_{n,i}, Y_{n,i})_{(n,i)\in\mathbb{N}\times\mathbb{N}^*}$ de variables aléatoires indépendantes et équidistribuées (i.i.d.) selon la loi du couple (X,Y), à valeurs entières. On considère une fonction de croisement $L : \mathbb{R}_+ \times \mathbb{R}_+ \longrightarrow \mathbb{R}_+$ croissante en chaque variable, à valeurs entières, telle que $L(x,y) \le xy$.*
On définit la suite $(Z_n)_{n\in\mathbb{N}}$ par la condition initiale $Z_0 = N$ $(N \ge 1)$ et la formule de récurrence

$$(1) \qquad \begin{cases} (F_{n+1}, M_{n+1}) = \displaystyle\sum_{i=1}^{Z_n} (X_{n,i}, Y_{n,i}), \\ (F_{n+1}, M_{n+1}) = (0,0), \ si\ Z_n = 0, \\ \qquad\qquad Z_n = L(F_n, M_n), \qquad n \ge 1. \end{cases}$$

On rappelle la définition de la fonction génératrice G de la loi du couple (X,Y) :

$$G(z_1,z_2)=\mathbb{E}\left(z_1^X z_2^Y\right).$$

La définition (1) entraîne que pour $|z_1|\leq 1$, $|z_2|\leq 1$ et $j\in\mathbb{N}$

$$\mathbb{E}\left(z_1^{F_{n+1}} z_2^{M_{n+1}}\mid Z_n=j\right)=\left(G(z_1,z_2)\right)^j. \tag{2}$$

Des exemples de lois de reproduction

– La portée d'un couple $T=X+Y$ a pour fonction génératrice $H(z)=\sum_{i\in\mathbb{N}} p_i z^i$. De plus, un nouveau-né est femelle avec probabilité p ($p\in]0,1[$) et mâle avec probabilité $q=1-p$, indépendamment des autres membres de la portée. Dans ce cas,

$$G(z_1,z_2)=H(pz_1+qz_2). \tag{3}$$

– Les nombres de mâles et de femelles d'une portée sont indépendants :

$$G(z_1,z_2)=G_X(z_1)\,G_Y(z_2). \tag{4}$$

Théorème 1. *Sous les hypothèses de la définition 1, la suite $(Z_n)_{n\in\mathbb{N}}$ est une chaîne de Markov homogène, à valeurs dans $\mathbb{N}$. L'état 0 est absorbant.*

Définition 2. *La fonction de croisement L est sur-additive si, quels que soient x_1,x_2,y_1,y_2 réels positifs,*

$$L(x_1+x_2,y_1+y_2)\geq L(x_1,y_1)+L(x_2,y_2).$$

2. Le modèle de Daley

Il s'agit d'un cas particulier de la définition 1. Ici, $L(x,y)=x\min(1,y)$, c'est-à-dire que

$$Z_n=\begin{cases} F_n & \text{si} \quad M_n\geq 1,\\ 0 & \text{si} \quad M_n=0.\end{cases}$$

Ce modèle traduit une situation de promiscuité totale des individus et un pouvoir reproductif illimité des mâles.

On suppose désormais que la loi de reproduction est décrite par le modèle (3). De plus,

$$0<p<1, \qquad p_0+p_1<1 \qquad \text{et} \qquad H(1)=1.$$

Par un argument de chaîne de Markov on établit que tous les états $i\geq 1$ sont transients. Il en découle le

Théorème 2.

$$\mathbb{P}\left(Z_n \xrightarrow[n\to\infty]{} 0\right)+\mathbb{P}\left(Z_n \xrightarrow[n\to\infty]{} \infty\right)=1.$$

L'étude de l'extinction de la population s'articule autour du taux moyen de croissance

$$r_k=\frac{1}{k}\,\mathbb{E}\left(Z_{n+1}|Z_n=k\right), \qquad k\in\mathbb{N}^*. \tag{1}$$

Le prochain résultat est crucial.

Théorème 3 (Fekete)**.** *Si une suite de nombres réels $(\alpha_k)_{k\in\mathbb{N}^*}$ est sous-additive :*

$$\alpha_{k+j} \leq \alpha_k + \alpha_j, \qquad k, j \in \mathbb{N}^*,$$

alors la limite $\lim_{n\to\infty} n^{-1}\alpha_n \left(= \inf\left\{n^{-1}\alpha_n : n \in \mathbb{N}^*\right\}\right)$ *existe. Elle est finie ou égale à* $-\infty$.

Le résultat suivant fournit la condition d'extinction de la population.

Théorème 4. *Quel que soit* $k \in \mathbb{N}^*$

$$q_k = \lim_{n\to\infty} \mathbb{P}(Z_n = 0 | Z_0 = k) = 1$$

si et seulement si

$$\mathsf{r} = \sup_{i\in\mathbb{N}^*} r_i \leq 1.$$

Démonstration succincte. La suite $(kr_k)_{k\in\mathbb{N}^*}$ est sur-additive. D'après le lemme de Fekete, qui précède, $\mathsf{r} = \lim_{n\to\infty} r_n$ existe et l'on a $\mathsf{r} = \sup_{k\in\mathbb{N}^*} r_k$.

Si $\mathsf{r} \leq 1$, alors quel que soit $k \in \mathbb{N}^*$, $r_k = \frac{1}{k}\mathbb{E}(Z_{n+1}|Z_n = k) \leq 1$. La suite $(Z_n)_{n\in\mathbb{N}}$ est une surmartingale positive. La probabilité qu'elle diverge vers l'infini est nulle. La probabilité d'extinction est égale à 1.

On suppose à présent que $\mathsf{r} = pH'(1) > 1$. On pose $g(z) = H(pz+q)$. L'équation

$$g(z) = z$$

possède une racine unique s_0 dans l'intervalle $]0,1[$. De plus, $g(z) < z$ dans l'intervalle $]s_0,1[$. On exclut le cas, lorsque $Z_0 = 1$, où $H(z) = p_0 + p_1 z + p_2 z^2$, avec $p_0 + p_1 + p_2 = 1$. En effet, la chaîne de Markov $(Z_n)_{n\in\mathbb{N}}$ ne visite pas l'état 2 lorsqu'elle est issue de $Z_0 = 1$.
Ce cas particulier exclu, la chaîne de Markov $(Z_n)_{n\in\mathbb{N}}$ possède un état absorbant ($E_0 = \{0\}$) et une classe irréductible d'états transients ($E_1 = \mathbb{N}^*$). Les probabilités q_k sont donc toutes égales à 1 ou toutes strictement inférieures à 1.

La suite $(q_k)_{k\in\mathbb{N}^*}$ est clairement décroissante. En outre,

$$\begin{aligned} 1 - q_k &= \lim_{n\to\infty} \mathbb{P}(Z_n \neq 0 \mid Z_0 = k), \\ &\geq \lim_{n\to\infty} \mathbb{P}(Z_n \geq k+n \mid Z_0 = k), \\ &\geq \lim_{n\to\infty} \prod_{l=1}^{n} \mathbb{P}(Z_l \geq k+l \mid Z_{l-1} = k+l-1). \end{aligned} \tag{2}$$

On a utilisé la propriété de Markov et on établit que

$$\inf_{j\geq k+l} \mathbb{P}\Big(Z_{l+1} \geq k+l+1 \mid Z_l = j\Big) = \mathbb{P}\Big(Z_{l+1} \geq k+l+1 \mid Z_l = k+l\Big).$$

Pour tout entier naturel k, pour tout $s \in]-1,1[$, la probabilité $\mathbb{P}(Z_{n+1} \leq k \mid Z_n = j)$ est le coefficient de s^k dans le développement en série entière, autour de l'origine, de

$$\frac{\mathbb{E}\Big(s^{Z_{n+1}} \mid Z_n = j\Big)}{1-s}.$$

La formule de Cauchy est requise. Elle s'énonce ici

$$\mathbb{P}(Z_{n+1} \leq k | Z_n = k) = \frac{1}{2i\pi} \oint_{\mathscr{C}(O,R)} \frac{g^k(z) - H^k(pz) + H^k(p)}{z^{k+1}(1-z)}\, \mathrm{d}z. \tag{3}$$

L'intégrale est prise le long du lacet $\gamma(t) = R\exp(it), t \in [0;2\pi]$, tracé dans le cercle $\mathscr{C}(O,R)$; le rayon R est choisi de telle façon que $s_0 < R < 1$ et $p\,R + q > p$. Par conséquent, $g(R) = H(pR+q) > H(p)$ et $\rho = g(R)/R < g(s_0)/s_0 = 1$. Le numérateur de l'intégrande (3) est analytique dans le disque unité, à coefficients positifs. Par le principe du maximum,
$\mathbb{P}(Z_{n+1} \leq k \mid Z_n = k) \leq \left(g^k(R) + H^k(p) \right) / \left(R^k(1-R) \right)$. On en déduit que

$$\mathbb{P}(Z_{n+1} \leq k \mid Z_n = k) \leq \frac{2\rho^k}{1-R}. \tag{4}$$

On tire des formules (2) et (4) que $1 - q_k > 0$.

3. Des modèles sur-additifs plus généraux

On revient au modèle général de fonction de croisement L décrite dans la définition 2. On suppose de plus que pour tout entier naturel non nul k

$$\mathbb{P}(Z_1 = k \mid Z_0 = k) < 1.$$

Daley, Hull et Taylor établissent en 1986 que le théorème 4 est encore vérifié. La loi forte des grands nombres permet de décrire le taux moyen de croissance limite r. On admettra ce résultat.

Théorème 5. *Soient deux suites i.i.d.* $(X_n)_{n\in\mathbb{N}}$ *et* $(Y_n)_{n\in\mathbb{N}}$ *indépendantes, de lois parentes respectives la loi de X et la loi de Y. On suppose ces deux lois intégrables. Alors*

$$\frac{1}{k} L\left(\sum_{j=1}^{k} X_j, \sum_{j=1}^{k} Y_j \right) \xrightarrow[k\to\infty]{p.s.} \mathsf{r} = \lim_{n\to\infty} \frac{1}{n} L\Big(n\mathbb{E}(X), n\mathbb{E}(Y) \Big).$$

Terminons par une application du résultat précédent à un exemple de fonction sur-additive. Le modèle est dénommé à couples fidèles et utilise la fonction de croisement $L(x,y) = \min(x,y)$. La population s'éteint quasi-certainement si et seulement si $\mathsf{r} = \min(\mathbb{E}(X), \mathbb{E}(Y)) \leq 1$.

On suppose, par symétrie de L, que $\mathsf{r} = \mathbb{E}(X)$. On trouve que

$$\mathbb{E}(Z_1 | Z_0 = k) = k\mathsf{r} - \mathbb{E}\left(\textstyle\sum_{i=1}^{k} (X_i - Y_i) \right)^+.$$

Lemme 1 (Foster–Tweedie)**.** *Soit* $\mathbf{X} = (X_n)_{n\in\mathbb{N}}$ *une chaîne de Markov à valeurs dans* $\mathbb{R}_+$ *et A une partie borélienne de* $\mathbb{R}_+$. *S'il existe $\varepsilon > 0$ tel que*

$$\forall x \notin A, \qquad \mathbb{E}(X_{n+1} - X_n \mid X_n = x) \leq -\varepsilon,$$

alors le temps de retour dans le borélien A,

$$T_A = \inf\{n \geq 1 : X_n \in A\}$$

vérifie

$$\forall x \notin A, \qquad \mathbb{E}(T_A \mid X_0 = x) \leq \frac{x}{\varepsilon}.$$

Si $\mathsf{r} < 1$, le critère de Foster–Tweedie précédent et, admis, permet d'établir que le singleton $A = \{0\}$ est récurrent positif ; le temps moyen d'extinction est fini.

Lemme 2. *Soient deux variables aléatoires U et V, indépendantes et intégrables. Alors*

$$\mathbb{E}|U+V| \geq \mathbb{E}\Big|U+\mathbb{E}(V)\Big|.$$

Supposons que $\mathsf{r}=1$ et $\mathbb{E}(X)=\mathbb{E}(Y)$. En appliquant le lemme 2 pour $U=X_1-Y_1$, on obtient, pour $k\geq 1$, que

$$\mathbb{E}(Z_1|Z_0=k)=k-\frac{1}{2}\mathbb{E}\left|\sum_{i=1}^{k}(X_i-Y_i)\right|\leq k-\frac{1}{2}\mathbb{E}|X-Y|.$$

En appliquant de nouveau le critère de Foster–Tweedie, on conclut que le temps moyen d'extinction est fini dans le cas où $\mathbb{E}(X)=\mathbb{E}(Y)=1$.

La discussion est moins claire si $\mathsf{r}=1$ et $\mathbb{E}(X)\neq\mathbb{E}(Y)$.

Si $\mathsf{r}>1$, il y a probabilité non nulle que l'espèce se perpétue pour Z_0 suffisamment grand. Plus précisément, il existe une variable aléatoire positive W telle que

$$\frac{Z_n}{\mathsf{r}^n}\xrightarrow[n\to\infty]{p.s.}W.$$

En effet, la suite (Z_n) vérifie la propriété de Markov. La suite $(\mathsf{r}^{-n}Z_n)$ est une surmartingale relativement à sa filtration naturelle. Le résultat est une conséquence de la convergence des surmartingales positives.

Suggestions pour le développement

► *Soulignons qu'il s'agit d'un menu à la carte et que vous pouvez choisir d'étudier certains points, pas tous, pas nécessairement dans l'ordre, et de façon plus ou moins fouillée. Vous pouvez aussi vous poser d'autres questions que celles indiquées plus bas. Il est très vivement souhaité que vos investigations comportent une partie traitée sur ordinateur et, si possible, des représentations graphiques de vos résultats.*

– *Etude de la modélisation et développements mathématiques.* Etablir la relation (2) de la page 2. Vérifier que les deux lois de reproduction des formules (3) et (4), de la page 2, coïncident si H est la fonction génératrice d'une loi de Poisson. Commenter ces modèles. En particulier la sur-additivité de la fonction de croisement L vous semble-t-elle pertinente ? Démontrer le théorème 1. On établira en particulier que dans le cas (3), la matrice de probabilités de transition a pour fonction génératrice

(1) $$\mathbb{E}\left(z^{Z_{n+1}}|Z_n=j\right)=H^j(pz+q)-H^j(pz)+H^j(p),\qquad \text{avec } j\in\mathbb{N},|z|\leq 1.$$

On rappelle que $L(x,y)=x\min(1,y)$. Compléter la démonstration du théorème 4.

– *Etude numérique.* Rédiger une procédure illustrant les probabilités d'extinction du modèle de Daley pour diverses fonctions génératrices H. On pourra comparer les probabilités d'extinction et les taux de croissance moyens (r_k en fonction de k) à ceux du modèle de Galton–Watson (c'est-à-dire $L(x,y)=x$). Etudier le temps moyen d'extinction pour diverses lois de X et Y. Evaluer numériquement, en fonction de l'état initial $Z_0=k$, la probabilité d'extinction q_k dans chacun des trois cas suivants :
  – Le modèle de Galton–Watson ; $L(x,y)=x$.

  - Le modèle de promiscuité totale ; $L(x,y) = x \min(1,y)$.
  - Le modèle des couples fidèles ; $L(x,y) = \min(x,dy)$, avec $d \in \mathbb{N}^*$.
- Quelles conclusions tirez-vous des investigations que *vous* avez *choisi* de mener ?