Processus stochastiques – Feuille d'exercices 3

Processus de Galton-Watson et martingales

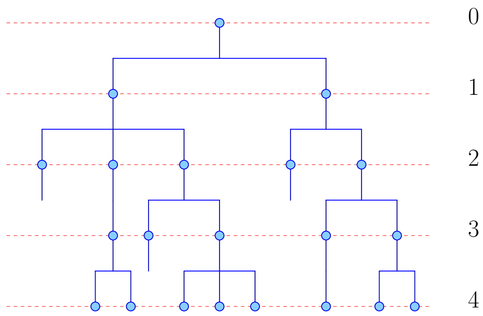
1 Processus de Galton-Watson

À l'époque victorienne, certaines personnes ont craint la disparition des noms des familles aristocratiques. Sir Francis Galton posa originellement la question de déterminer la probabilité d'un tel événement dans le *Educational Times* de 1873, et le Révérend Henry William Watson répondit avec une solution. Ensemble, ils écrivirent alors, en 1874, un article intitulé « On the probability of extinction of families ». Leur modèle suppose (cela étant considéré comme allant de soi à l'époque de Galton, et étant encore le cas le plus courant dans la plupart des pays) que le nom de famille est transmis à tous les enfants mâles par leur père. Il suppose également que le nombre de fils d'un individu est une variable aléatoire à valeurs dans $\mathbb{N}$, et que le nombre de fils d'hommes différents sont des variables aléatoires indépendantes de même loi.

Plus généralement, supposons qu'une population évolue par générations, et notons Z_n le nombre d'individus de la $n^{\text{ème}}$ génération. Chaque membre de la $n^{\text{ème}}$ génération donne naissance à une famille, éventuellement vide, de la génération suivante ; la taille de la famille est une variable aléatoire. On fait les hypothèses suivantes :

— les tailles de chaque famille forment une collection de variables aléatoires indépendantes ;
— les tailles des familles suivent toutes la même loi.

Sous ces hypothèses, le processus est bien défini dès que la taille de la population initiale Z_0 est donnée ; on supposera ici que $Z_0 = 1$. Ce modèle peut également représenter la croissance d'une population de cellules, celle de neutrons dans un réacteur, la propagation d'une maladie dans une population, etc.



On s'intéresse à la suite aléatoire $Z_0, Z_1, Z_2, \ldots$ des tailles des générations successives. Ce processus peut être défini de la façon suivante : on considère une collection $(X_k^n)_{n\geq 0, k\geq 1}$ de variables aléatoires i.i.d. à valeurs dans $\mathbb{N}$; X_k^n représente le nombre d'enfants du $k^{\text{ème}}$ individu de la $n^{\text{ème}}$ génération (si celui-ci existe). On note G la fonction génératrice commune à ces variables aléatoires (encodant donc la loi du nombre de fils d'un individu). On peut alors poser $Z_0 = 1$ et, pour $n \geq 1$,

$$Z_n = X_1^{n-1} + X_2^{n-1} + \cdots + X_{Z_{n-1}}^{n-1}\,, \tag{1}$$

le nombre d'individus de la $n^{\text{ème}}$ génération étant égal au nombre total de fils des individus de la $(n-1)^{\text{ème}}$ génération. On notera $G_n(s) = \mathbb{E}(s^{Z_n})$ la fonction génératrice de Z_n. Observez qu'en particulier $G_1 = G$, puisque $Z_0 = 1$.

Question 1.1. Pour tout $n \geq 1$, montrer que

$$G_n = G^{\circ n} \equiv \underbrace{G \circ G \circ \cdots \circ G}_{n \text{ fois}}\,.$$

Les moments de la variable aléatoire Z_n peuvent facilement s'exprimer en termes des moments de la variable aléatoire Z_1 décrivant la taille d'une famille typique.

Question 1.2. Soit $\mu = \mathbb{E}(Z_1)$ et $\sigma^2 = \mathrm{Var}(Z_1)$. Montrer que

$$\mathbb{E}(Z_n) = \mu^n,$$
$$\mathrm{Var}(Z_n) = \begin{cases} n\sigma^2 & \text{si } \mu = 1 \\ \sigma^2(\mu^n - 1)\mu^{n-1}(\mu - 1)^{-1} & \text{si } \mu \neq 1. \end{cases}$$

Une question particulièrement intéressante concerne le destin de la population : va-t-elle s'éteindre après un temps fini, ou au contraire, toutes les générations auront-elles une taille strictement positive ? Cette question peut être reformulée sous la forme suivante : la limite $\lim_{n\to\infty} \mathbb{P}(Z_n = 0)$ est-elle égale à 1 (extinction inéluctable) ou strictement inférieure à 1 (survie possible) ? (Observez que s'il est possible qu'un individu n'ait pas de descendance, alors la probabilité d'extinction est toujours strictement positive.) Le destin de la population est étroitement lié à la taille moyenne des familles.

Question 1.3. Pourquoi la limite $\lim_{n\to\infty} \mathbb{P}(Z_n = 0)$ existe-t-elle ?

Question 1.4. Montrer que la taille totale de l'arbre $\mathcal{T}$ de Galton-Watson a une espérance finie si et seulement si $\mu < 1$, auquel cas :

$$\mathbb{E}(|\mathcal{T}|) = \frac{1}{1-\mu} ,$$

et la population s'éteint nécessairement.

Question 1.5. Soit $\mu = \mathbb{E}(Z_1)$, la taille moyenne d'une famille. Montrer que la probabilité d'extinction

$$q = \lim_{n\to\infty} \mathbb{P}(Z_n = 0)$$

est donnée par la plus petite racine positive de l'équation $s = G(s)$. En particulier, $q = 1$ si $\mu < 1$ et $q < 1$ si $\mu > 1$. Lorsque $\mu = 1$, on a $q = 1$ sauf si $\mathbb{P}(Z_1 = 1) = 1$.

Question 1.6. Montrer que G est convexe et finie sur $[0, 1]$ et faire un dessin faisant apparaître la probabilité d'extinction q.

Question 1.7. *Autosimilarité.*
Soit $X = (X_n)_{n\geq 0}$ un processus à valeurs dans $\mathbb{N}$. On note $X^{(i)}$, $i \geq 1$ une suite de copies i.i.d de X, indépendante de X. On définit un processus X' par :

$$X'_0 = 1$$
$$X'_n = \textstyle\sum_{i=1}^{X_1} X^{(i)}_{n-1} \quad n \geq 1 .$$

Montrer que X' a même loi que X si et seulement si X est un processus de Galton-Watson de loi de reproduction la loi de X_1.

Question 1.8. On suppose que la probabilité d'extinction q est strictement positive. Montrer que conditionnellement à son extinction, Z est encore un processus de Galton-Watson, de loi de reproduction ayant pour fonction génératrice $s \mapsto G(qs)/q$.

Question 1.9. On suppose maintenant que $\mu \leq 1$ (cas critique et sous-critique). On note $h(s) = \mathbb{E}(s^{|\mathcal{T}|})$ la fonction génératrice de la taille de l'arbre généalogique.

1. Montrer, en conditionnant par rapport à la taille Z_1 de la génération à l'instant 1, que pour tout $s \in [0,1]$, $h(s)$ vérifie :
$$h(s) = sG(h(s)) \ .$$
2. Montrer que pour tout $s \in [0,1[$, l'équation $x = sG(x)$ admet une unique solution dans $[0,1[$.
3. Déterminer $h(s)$ dans le cas où la loi du nombre de descendants d'un individu est une loi géométrique de paramètre p sur $\mathbb{N}$: $\mathbb{P}(X_1^1 = k) = p(1-p)^k$ pour tout $k \geq 0$. Expliciter la loi de $|\mathcal{T}|$ dans le cas $p = 1/2$.

2 Sous l'angle des martingales

Question 2.1.

1. On suppose désormais que $\mu > 0$, et on note $X_n = Z_n/\mu^n$. Montrer que X_n est une martingale, et en déduire que X_n converge presque sûrement vers une variable aléatoire que l'on notera X.
2. Que vaut X si $\mu \leq 1$?
   (a) Cas $\mu < 1$: En utilisant le point (1), montrer que $\mathbb{P}(Z_n = 0) \to 1$ lorsque $n \to \infty$. En déduire que $Z_n \to 0$ presque sûrement et dans $\mathcal{L}^1$.
   (b) Cas $\mu = 1$: Montrer que X_n ne converge pas vers X dans $\mathcal{L}^1$.

Question 2.2. *Décomposition de Doob*
Montrer que toute sous-martingale $(Y_n)_n$ peut être écrite d'une manière unique comme $Y_n = M_n + A_n$, où M_n est une martingale et A_n est un processus prévisible croissant tel que $A_0 = 0$.

Indications :

1. Supposer d'abord que la décomposition existe et calculer $\exp(Y_n|\mathcal{F}_{n-1})$. En déduire deux relations $A_n = f(A_{n-1}, Y_{n-1}, \mathbb{E}(Y_n|\mathcal{F}_{n-1}))$ et $M_n = g(A_n, Y_n)$ qui définissent A_n et M_n univoquement (par récurrence).
2. Montrer que $(A_n)_n$ défini ainsi est croissant ($A_{n-1} \leq A_n$) et prévisible (A_n est $\mathcal{F}_{n-1}$-mesurable).
3. Montrer que $(M_n)_n$ défini ainsi est une martingale.

Question 2.3. *Application au processus de Galton-Watson*

1. On note toujours $X_n = Z_n/\mu^n$. Montrer que $(X_n^2)_n$ est une sous-martingale.
2. Dans la question précédente on a vu que X_n^2 peut être écrite d'une manière unique comme $X_n^2 = M_n + A_n$, où M_n est une martingale et A_n est un processus prévisible croissant. Lorsque $(X_n)_n$ est une martingale, on appelle ce processus prévisible le crochet de X_n, et on le note $\langle X \rangle_n$, autrement dit :
$$X_n^2 = M_n + \langle X \rangle_n. \qquad (*)$$
Montrer que
$$\langle X \rangle_n = \sigma^2 \sum_{m=1}^{n} \frac{X_{m-1}}{\mu^{m+1}}$$
et calculer $\mathbb{E}(\lim_{n\to\infty} \langle X \rangle_n)$.

Indications : En conditionnant $(*)$ par rapport à $\mathcal{F}_{n-1}$, montrer (par induction) que

$$\langle X \rangle_n = \sum_{m=1}^{n} \mathbb{E}((X_m - X_{m-1})^2 | \mathcal{F}_{m-1}).$$

Montrer ensuite que $Z_n - \mu Z_{n-1} = \eta_{n,1} + ... + \eta_{n,Z_{n-1}}$ avec $\eta_{n,i} := \xi_{n,i} - \mu$, et faire le calcul.

3. Si $\mu > 1$, montrer que $\langle X \rangle_\infty := \lim_{n\to\infty} \langle X \rangle_n < \infty$. En déduire que (X_n) est bornée dans $\mathcal{L}^2$.

Question 2.4. 1. Montrer que $\mathbb{P}(X = 0) = G(\mathbb{P}(X = 0))$, et en déduire que $\mathbb{P}(X = 0) \in \{q, 1\}$.

2. Dans le cas $\mu > 1$,
   (a) Montrer que $\mathbb{P}(X > 0) > 0$.
   Indication : Montrer que (X_n) est uniformément intégrable, et donc que $\mathbb{E}(X) = 1$.
   (b) Montrer que $X > 0$ sur l'événement $\{Z_n > 0 \ \forall n\}$, i.e. si la population survit, elle croît exponentiellement vite.
   (c) On suppose que Z_1 est de carré intégrable. Montrer que $\mathbb{P}(W = 0) = q$. Interpréter.