Saloin Lessard
Processus stochastiq

Benaïm el Karoui promenade aléatoire p153
Cottrel, ex de proba p72
3.5

PROCESSUS DE GALTON-WATSON

17min sans faire Etape 2 ni 1er point 1ere prop.

15'06

Contexte : de nombreux phénomènes d'évolution de population peuvent être modélisés en première approximation par un processus de branchement (étude des gènes, survivance d'un nom de famille...).

Prérequis :
- fonction génératrice
- thm des séries (entières) de classe $\mathcal{C}^1$, thm de dérivat° séries entières, rayon de cv
- convexité par le signe de la dérivée seconde. + Rolle

Soit X une v.a. intégrable à valeurs dans $\mathbb{N}$. On note pour $k \in \mathbb{N}$ $p_k = \mathbb{P}(X=k)$ et $m = \mathbb{E}[X] < \infty$. Soit $(X_{i,j})$ iid de loi $\mathbb{P}_X$.

On définit Z_n par $Z_0 = 1$
$$Z_{n+1} = \sum_{i=1}^{Z_n} X_{i,n}$$

On note $\pi_n = \mathbb{P}(Z_n = 0)$ et $\pi_\infty = \mathbb{P}(\exists n \in \mathbb{N}, Z_n = 0)$.

Histoire (ne pas le dire à l'oral)

Des particules sont capables de générer des particules de la m famille. Chaque particule a la proba p_k d'engendrer k particules indépendantes (proba cste au cours des générat°). La particule originale représente la génération 0. Les descendants de la $n$ème générat° forment la $(n+1)$e générat°. Z_n est le nb d'individus à la générat° n. Chaque individu i de la ne génération a un nb $X_{i,n}$ de descendants ($1 \leq i \leq Z_n$), iid ok.
π_n est la proba que la famille soit éteinte à la génération n (mais peut être déjà avant). π_∞ est la proba d'extinction. On étudie cette probabilité.

- Si $p_0 = 0$, $\forall n \in \mathbb{N}^*$ $Z_n \geq 1$ p.s. et donc $\pi_\infty = 0$.
- Si $p_0 = 1$ alors $\forall n \in \mathbb{N}^*$ $Z_n = 0$ p.s et donc $\pi_\infty = 1$.

$\Big\}$ Désormais $p_0 \in]0,1[$.

Etape 1 : Fonction génératrice de X.

Pour $0 \leq t \leq 1$, on définit la fct° génératrice de X : $G(t) = \mathbb{E}[t^X] = \sum_{k=0}^{+\infty} p_k t^k \leq 1$.
et on a $G(1) = 1$.

Prop 1 :
- G est bien définie sur $[0,1]$ et de classe $\mathcal{C}^1$.
- G est (strictement croissante sur $]0,1[$) convexe sur $]0,1[$, croiss sur $[0,1]$
strictement convexe sur $]0,1[\iff p_0 + p_1 < 1$.

G'' en 1 n'existe pas !!

Dém :
- $\times$ $\forall k \in \mathbb{N}$ $t \mapsto p_k t^k \in \mathcal{C}^1([0,1])$.
  $\times$ $\sum_{k \geq 0} p_k 1^k$ converge (vers 1)
  $\times$ $\sum_{k \geq 1} k p_k t^{k-1}$ CN (car X intégrable : $\mathbb{E}[X] < \infty$) donc CU sur $[0,1]$.

  D'où par thm, $\sum_{k \geq 0} p_k t^k$ CU vers G de classe $\mathcal{C}^1$ sur $[0,1]$ (et bien définie)

- La série entière $\sum p_k t^k$ a un rayon de cv ≥ 1 d'où par thm de dérivat° terme à terme d'une série entière.

csqces

$$\forall t \in [0,1[\quad G'(t) = \sum_{k=1}^{+\infty} k p_k t^{k-1}, \quad G''(t) = \sum_{k=2}^{+\infty} k(k-1) p_k t^{k-2}.$$

On a $p_0 < 1$ et $\sum p_k = 1$ d'où $\exists k_0 > 0$ tq $p_{k_0} > 0$, ainsi

$\rightarrow$ $\forall t \in]0,1[$ $G'(t) \geq k_0 p_{k_0} t^{k_0 - 1} > 0 \Rightarrow G$ strict $\nearrow$ sur $]0,1[$ et croiss sur $[0,1]$
$\rightarrow$ $\forall t \in]0,1[$ $G''(t) \geq k_0(k_0-1) p_{k_0} t^{k_0-2} \geq 0 \Rightarrow G$ convexe sur $]0,1[$.
$\rightarrow$ Si $p_0 + p_1 = 1$, $k_0 = 1$ et $G = p_0 + p_1 t$ affine donc pas strict convexe sur $]0,1[$.
Si $p_0 + p_1 < 1$, il existe $k_1 > 1$ tq $p_{k_1} > 0$ et donc $G'' > 0$ sur $]0,1[$. d'où strictement convexe.

On a $m = \mathbb{E}[X] = G'(1)$.

Etape 2 : Fonction génératrice de Z_n, relation de récurrence.

Pour $n \in \mathbb{N}$, on pose G_n la fct° génératrice de Z_n : $\forall t \in [0,1]$ $G_n(t) = \mathbb{E}[t^{Z_n}]$
$= \sum_{k \geq 0} \mathbb{P}(Z_n = k) t^k$

Comme avant, G_n est bien définie sur $[0,1]$.
et remarquons que $G_n(0) = \mathbb{P}(Z_n = 0) = \pi_n$ et $G_n'(1) = \mathbb{E}[Z_n]$.

Lemme 1 : Pour $n \in \mathbb{N}^*$, $i \in \mathbb{N}$, $Z_n \perp\!\!\!\perp X_{i,n}$. (ne pas dem à l'oral)

Dém : $Z_n \in \sigma(Z_{n-1}, X_{i,n-1}, i \in \mathbb{N})$ pour $n \in \mathbb{N}^*$. et $Z_0 = 1$ d'où par récurrence immédiate, $Z_n \in \sigma(X_{i,j}, i \in \mathbb{N}, j \leq n-1)$
Or les $X_{i,j}$ sont indépendantes d'où $Z_n \perp\!\!\!\perp X_{i,n}$ pour $i \in \mathbb{N}$

Prop 2 : Pour $n \in \mathbb{N}^*$, on a $G_n = \overbrace{G \circ \cdots \circ G}^{n \text{ fois}}$ sur $[0,1]$.

Dém : par récurrence sur n. • $G_1(t) = \mathbb{E}[t^{Z_1}] = \mathbb{E}[t^{X_{1,0}}] = \mathbb{E}[t^X] = G(t)$. , $t \in [0,1]$.
• Si $G_n = G \circ \cdots \circ G$. Soit $t \in [0,1]$

$$G_{n+1}(t) = \mathbb{E}[t^{Z_{n+1}}] = \mathbb{E}[t^{\sum_{i=1}^{Z_n} X_{i,n}}] = \mathbb{E}\Big[\prod_{i=1}^{Z_n} t^{X_{i,n}}\Big].$$
$$= \mathbb{E}\Big[\sum_{k=0}^{+\infty} \prod_{i=1}^{k} t^{X_{i,n}} 1_{Z_n = k}\Big] = \sum_{k=0}^{+\infty} \mathbb{E}\Big[\prod_{i=1}^{k} t^{X_{i,n}} 1_{Z_n = k}\Big] \text{ (Fubini-Tonelli)}$$
$$\overset{\perp\!\!\!\perp}{=} \sum_{k=0}^{+\infty} \prod_{i=1}^{k} \mathbb{E}[t^{X_{i,n}}] \mathbb{E}[1_{Z_n = k}] = \sum_{k=0}^{+\infty} \mathbb{P}(Z_n = k) \, \mathbb{E}[t^X]^k \text{ car même loi.}$$
$$= \sum_{k=0}^{+\infty} \mathbb{P}(Z_n = k) G(t)^k = G_n(G(t)) = \underbrace{G \circ \cdots \circ G}_{n+1 \text{ fois}}(t) \text{ par hyp de récurrence.}$$

Ce qui donne $\pi_{n+1} = G(\pi_n)$, $n \in \mathbb{N}$ en évaluant en 0

Etape 3 : Probabilité d'extinction.

Comme $Z_n = 0 \Rightarrow Z_{n+1} = 0$, on a $(\{Z_n = 0\})_{n \in \mathbb{N}}$ croissante. Ainsi,
$\pi_\infty = \mathbb{P}(\bigcup \uparrow (Z_n = 0)) = \lim_{+\infty} \mathbb{P}(Z_n = 0) = \lim_{+\infty} \pi_n$ $\quad \pi_\infty \geq \pi_1 > 0$, $\pi_\infty \leq 1$

Prop 3 : π_∞ est le plus petit point fixe de G sur $[0,1]$

Dém : • Comme $\pi_{n+1} = G(\pi_n)$ pour $n \in \mathbb{N}$ et G est continue sur $[0,1]$, $\pi_\infty = G(\pi_\infty)$.
• Soit $u \in [0,1]$, un pt fixe de G sur $[0,1]$. On va montrer par récurrence que pour tout $n \in \mathbb{N}$ $\pi_n \leq u$.
* $\pi_1 = p_0 = G(0) \leq G(u) = u$.
* si $\pi_n \leq u$ $\quad \pi_{n+1} = G(\pi_n) \leq G(u) = u$.

D'où $\forall n \in \mathbb{N}$ $\pi_n \leq u$, par passage à la limite, $\pi_\infty \leq u$.

Théorème : • Si $m \leq 1$, $\pi_\infty = 1$
• Si $m > 1$ alors π_∞ est l'unique pt fixe de G sur $]0,1[$.

Dém : Comme $G(1) = 1$, le graphe de G coupe la droite $y = x$ sur $[0,1]$ au moins au pt $(1,1)$.
Rappelons qu'on a 2 cas :
* Si $p_0 + p_1 = 1$, G est une droite et comme $G(0) = p_0 \neq 0$., G ne coupe $y = x$ qu'en un unique point. ici $m = p_1 < 1$ écrire la déf. $\quad \pi_\infty = 1$
* sinon G est strictement convexe et $f : x \mapsto G(x) - x$ ~~[illegible]~~ s'annule au plus 2 fois : sur $[0,1]$, si elle s'annulait en 3 points $\neq$ alors par thm de Rolle, il existe $a < b$, $f'(a) = f'(b) = 0$. Or f est convexe donc sa dérivée est croissante donc $f'_{|[a,b]} = 0$ donc f est constante sur $]a,b[$
donc G est affine sur $[a,b]$
[illegible] strictement convexe

Donc dans tous les cas, $f(x) = G(x) - x$ s'annule au ⊕ 2 fois sur $[0,1]$.
On a $G'(0) = p_1$, $G'(1) = m$. On note $f : x \mapsto G(x) - x$.

1er cas : Si $m > 1$:

⚠ strict convexe $\not\Rightarrow f'' > 0$ ex : $x \mapsto x^4$.
mais convexe $\Leftrightarrow f'' \geq 0$

$G'' \geq 0$ (car G est convexe)
d'où $f'' = G'' \geq 0$ donc f' est croissante
de plus $f'(0) = p_1 - 1 < 0$ car $p_0 > 0$.
donc il existe $\alpha \in]0,1[$ tq $f'(\alpha) = 0$.

x	0	π_∞	α		1
f'	$p_1 - 1 < 0$	$-$	0	$+$	$m-1$
f	$p_0 > 0$ ↘	0	↘	↗	0

$f(0) = p_0 > 0$, $f(1) = 0$. d'où $f(\alpha) < 0$.
d'où il existe un point dans l'intervalle $]0, \alpha]$
où f s'annule i.e. G admet un pt fixe.
Or π_∞ est un pt fixe et il y a au ⊕ 2 pts fixes
sur $[0,1]$ d'où c'est π_∞. car π_∞ est le plus petit point fixe de G

2e cas : Si $m \leq 1$:

$G'' \geq 0$ (car G est convexe) d'où $f'' = G'' \geq 0$
d'où f' croissante.
$f'(0) = p_1 - 1 \leq 0$ $f'(1) = m - 1 \leq 0$.
d'où $f' \leq 0$ sur $[0,1]$.
d'où f est décroissante sur $[0,1]$ et
s'annule en 1. Comme f admet au ⊕ 2
annulations, elle ne s'annule qu'en 1, sinon elle s'annulerait sur un intervalle non
réduit à un singleton. Donc $\pi_\infty = 1$.

x	0		1
f'	$p_1 - 1$	$-$	$m-1$
f	$p_0 > 0$	↘	0

Bonus : Espérance de Z_n.

On aurait pu se demander quel est le nb moyen d'individus à la génération n.

Prop : $\mathbb{E}[Z_n] = m^n$.

Dém : 1ère méthode : par récurrence.

- $Z_0 = 1$ donc $\mathbb{E}[Z_0] = 1 = m^0$.
- On peut dériver G_n (m̂ raison que G) et on a pour $s \in [0,1]$.
  $G'_{n+1}(s) = G'(s)\,(G'_n \circ G(s))$
  $G'_{n+1}(1) = \mathbb{E}[X]\,(G'_n(G(1))) = m\,G'_n(1) = m^{n+1}$.

2° méthode pour la récurrence :

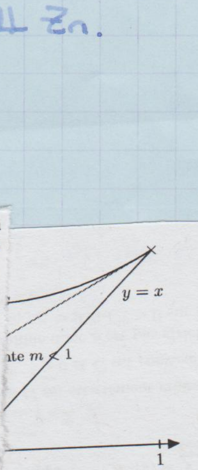
$$\mathbb{E}[Z_{n+1}] = \mathbb{E}\Big[\sum_{i=1}^{Z_n} X_{i,n}\Big] = \mathbb{E}\Big[\mathbb{E}\Big[\sum_{i=1}^{Z_n} X_{i,n} \,\Big|\, Z_n\Big]\Big] = \mathbb{E}\Big[\mathbb{E}\Big[\sum_{i=1}^{+\infty} \mathbb{1}_{i \leq Z_n} X_{i,n} \,\Big|\, Z_n\Big]\Big].$$

$$= \mathbb{E}\Big[\sum_{i=1}^{\infty} \mathbb{1}_{i \leq Z_n}\, \mathbb{E}[X_{i,n} \mid Z_n]\Big]$$ par Fubini Tonelli + $\mathbb{1}_{i \leq Z_n}$ est Z_n-mesurable

$$= \mathbb{E}\Big[\sum_{i=1}^{\infty} \mathbb{1}_{i \leq Z_n}\, \mathbb{E}[X_{i,n}]\Big]$$ car $X_{i,n} \perp\!\!\!\perp Z_n$.

$$= \mathbb{E}\Big[\sum_{i=1}^{Z_n} m\Big] = m\,\mathbb{E}[Z_n] = m^{n+1}.$$

✓ **A l'oral**, on n'explique pas du tout l'histoire : juste brut de maths (ce sera forcément une question du jury). On va vite sur le **1** (6'), on ne fait pas le lemme du **2** (9'27), on fait le **3**.2 géométriquement (13'51). Temps donné en allant hyper vite.
✓ On parle également de processus de branchement.
✓ En fait, la suite (Z_n) est une chaîne de Markov issue de 1, dont l'espace d'états est dénombrable. L'état 0 est absorbant. La chaîne est transciente. La question est ici de savoir si elle "sort " de $\mathbb{N}$ par 0 ou par l'infini.
✓ A l'origine, ce modèle a été introduit par GALTON en 1873 en vue d'étudier la statistique des patronymes dans l'Angleterre victorienne.
♣ Francis GALTON (1822 - 1911) est un homme de science britannique. Il fut anthropologue, explorateur, géographe, inventeur, météorologue, proto-généticien, psychométricien et statisticien. Il est entre autres fondateur de la psychologie différentielle ou comparée. Il a également mis en place de façon systématique la méthode d'identification des individus par empreintes digitales. Il fut anobli en 1909 et reçut la médaille Copley, décernée par la Royal Society.
♣ Henry WATSON (1827 - 1903) est un mathématicien britanique. Il a écrit de nombreux livres sur les mathématiques appliquées à l'électricité et le magnétisme. A ne pas confondre avec George, célèbre pour ses travaux sur les fonctions spéciales.



…URE 2 : Cas $m < 1$

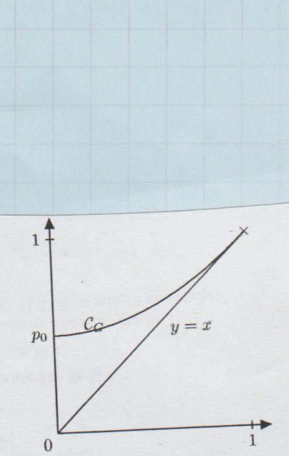


FIGURE 3 : Cas $m = 1$ avec $p_0 + p_1 < 1$