Processus de branchement

Exercice 1 Générations

Dans une population, chaque adulte mâle (qui a un nom donné) aura un certain nombre d'enfants mâles qui atteindront l'âge adulte. Soit X la variable aléatoire qui représente ce nombre. La probabilité d'avoir $0, 1, 2, 3, \ldots$ est $p_0, p_1, p_2, p_3, \ldots$. S'il y a k telles naissances à la première génération, il y en aura $X_1 + X_2 + \cdots + X_k$ à la deuxième génération. On cherche à étudier la probabilité avec laquelle un nom s'éteint, et ce au bout de combien de générations.

▷ 1. Construire un arbre donnant le nombre de telles naissances à la deuxième génération avec leur probabilité pour $p_0 = 1/2, p_1 = 1/4, p_2 = 1/4$.

▷ 2. Soit d_m la probabilité que le processus s'éteigne à la m^e génération et $d = \lim_{m\to\infty} d_m$. Montrer que $0 \leqslant d \leqslant 1$.

▷ 3. Soit h la fonction génératrice pour les p_i, montrer que $d_m = h(d_{m-1})$. En déduire que déterminer la probabilité qu'un nom s'éteint revient à étudier l'intersection des courbes des fonctions $y = z$ et $y = h(z)$.

▷ 4. Faîtes cette étude. En déduire que si en moyenne, chaque parent a $m \leqslant 1$ *bonnes* naissances, alors $d = 1$ et un nom disparait à coup sûr, sinon, $d < 1$ et un nom disparait avec une probabilité d.

▷ 5. Supposons que le maximum de *bonnes* naissances est 2, i *bonnes* naissances arrivant avec la probabilité p_i. Donner les valeurs possibles de d et le nombre moyen de naissance par adulte mâle m en fonction des p_i. En déduire la condition sur les p_i pour qu'un nom s'éteigne avec une probabilité inférieure à 1.

Solution

▷ 1. On peut proposer le tableau suivant :

Nb de personnes	0	1	2	3	4
Génération 0	0	1	0	0	0
Génération 1	1/2	1/4	1/4	0	0
Génération 2	11/16	1/8	9/64	1/32	1/64

▷ 2. Soit d_m la probabilité que le processus soit éteint à la m^e génération. Comme la suite $(d_m)_{m\in\mathbb{N}}$ est croissante et majorée par 1, elle converge vers une limite comprise entre 0 et 1.

▷ 3. Pour calculer d_m en fonction de d_{m-1} il est plus comode d'ajouter une génération "au début" qu'"à la fin". Considérons l'ancêtre d'une lignée. Il a i enfants mâles avec une probabilité p_i. Chaque enfant voit sa lignée éteinte au bout de $m-1$ générations avec une probabilité d_{m-1}. La lignée de l'encêtre se sera éteinte au bout de m générations si et seulement si la lignée de chaque enfant est éteinte au bout de $m-1$ générations. Donc :

$$d_m = \sum_{i=0}^{\infty} p_i \times {d_{m_1}}^i = h(d_{m-1})$$

où h est la fonction génératrice des p_i. La suite d_m est donc une suite récurrente associée la la fonction h. Sa limite vérifie : $d = h(d)$ par passage à la limite dans l'égalité $d_m = h(d_{m-1})$, et sera même exactement le plus petit point d'intersection des courbes $y = z$ et $y = h(z)$ sur $[0, 1]$.

▷ 4. On cherche à trouver les valeurs comprises entre 0 et 1 pour lesquelles la fonction $z \mapsto h(z) - z$ s'annule. La dérivée de cette fonction vaut $\sum p_i i z^{i-1} - 1$.
Soit $m = \sum p_i i$ le nombre moyen de bonnes naissances par adulte mâle. Si $m \leqslant 1$, la dérivée de $z \mapsto h(z) - z$ reste toujours négative entre 0 et 1, on peut établir le tableau de variation suivant :

z	0		1
$h'(z) - 1$		$-$	
$h(z) - z$	p_0	$\searrow$	0

Le seul point fixe de h est alors 1, et donc $d = 1$. Le nom disparaît alors presque sûrement
Si $m > 1$, la dérivée s'annule en un x compris entre 0 et 1, et le tableau de variations devient :

z	0		x		1
$h'(z) - 1$		$-$		$+$	
$h(z) - z$	p_0	$\searrow$	y	$\nearrow$	0

h admet alors un point fixe inférieur strictement à 1, et donc $d < 1$. Le nom s'éteint alors avec une probabilité d.

▷ 5. Supposons que le maximum de bonnes naissances soit 2. Autrement dit, $\forall i > 2, p_i = 0$. L'on a alors : $h(z) = p_0 + p_1 z + p_2 z^2$, et l'équation $h(z) = z$ devient : $p_2 z^2 + (p_1 - 1)z + p_0 = 0$.
Le discriminant de cette équation du second degré vaut : $(1-p_1)^2 - 4p_0p_2 = (p_2+p_0)^2 - 4p_0p_2 = (p_2-p_0)^2$ car $p_0 + p_1 + p_2 = 1$. Le discriminant est toujours positif - d'ailleurs l'équation admet toujours 1 comme racine - et les racines valent $\frac{1-p_1-|p_0-p_2|}{2p_2}$ et $\frac{1-p_1+|p_0-p_2|}{2p_2}$.

Dans tous les cas, les racines sont p_0/p_2 et 1. Si $p_0 < p_2$, l'on a $d = p_0/p_2 < 1$. Sinon, d vaut 1 et le nom s'éteint à coup sur.
Comme le nombre moyen de naissances m vaut $p_1 + 2p_2$, on a l'équivalence : $m > 1$ ssi $p_2 - p_0 > 0$, et la condition trouvée pour avoir d inférieur à 1 est bien équivalente à celle trouvée dans le cas général.

Exercice 2 Chaîne postale

En 1978, une chaîne postale a consisté à acheter une lettre de 12 (par exemple) noms. Supposons que l'acheteur donne 50$ au vendeur, et 50$ à la personne dont le nom est inscrit en haut de la liste. L'acheteur barre ce nom et ajoute son propre nom à la fin de la liste et la revend.

▷ 1. Si on revend la liste qu'à une seule personne, quelle est notre espérance de gain ?

▷ 2. En imaginant que chaque personne revend la liste à 0, 1 ou 2 personnes avec une certaine probabilité, quelle est notre espérance de gain ?

▷ 3. Quelle consition doit respecter m l'espérance du nombre de lettres que vous avez envoyé pour que ce processus vous soit favorable ?

▷ 4. Supposons maintenant qu'il n'y a aucune limite sur le nombre de personnes à qui on revend la liste. Cependant, supposons que chaque personne connaît un grand nombre N d'acheteurs potentiels mais que la probabilité p de leur vendre la liste est petite. Donner une approximation de p_j la probabilité de vendre la lettre à j personnes.

Solution

▷ 1. Si chaque personne revend la liste à exactement une personne, on est assurés de gagner 50$ à la vente de la liste, et encore 50$ au 12^e descendant, soit 100$ en tout, ce qui rembourse l'investissement.

▷ 2. Posons e_i l'espérance du nombre de personnes reçevant la lettre après i étapes - c'est à dire le nombre de listes avec le nom considéré en i^e position en partant du bas. La première personne considérée envoie la lettre à i personnes, $i = 0, 1$ ou 2, avec une probabilité p_i. Chaque receveur a une espérance de e_{n-1} correspondants au bout de $n - 1$ étapes. Donc l'espérance du nombre de correspondants au bout de n étapes pour la première personne est de

$$e_n = 0 \times p_0 + p_1 \times e_{n-1} + 2 \times p_2 \times e_{n-1} = (p_1 + 2p_2)e_{n-1} = \cdots = (p_1 + 2p_2)^n$$

Car $e_0 = 1$. On peut alors calculer l'espérance E du gain, qui consistera en la somme des espérances des gains à la revente à chaque descendant direct, plus la somme des espérances pour les gains dûs à la 12^e génération, soit :

$$E = 50 \times e_1 + 50 \times e_{12} = 50[(p_1 + 2p_2) + (p_1 + 2p_2)^{12}]$$

▷ 3. La condition pour que l'opération soit intéressante sera donc que le terme $p_1 + 2p_2$ soit supérieur à 1.

▷ 4. Supposons que chaque personne connaisse N acheteurs potentiels, et qu'elle ait une probabilité p de revendre la liste à chacun. Cela peut se modéliser à l'aide d'une distribution binomiale, réussir à revendre la liste étant un succès. Alors, la probabilité p_j qu'une personne donnée arrive à vendre la liste à j personnes est :

$$\binom{N}{j} p^j (1-p)^{N-j}$$

Comme N est grand (il y a plein de gens à qui on peut la revendre) et p est petit (les gens se méfient), on peut approximer ça par une loi de Poisson de paramètre Np.