Branchement

TD6-MAPI3

2016-2017

1 La poule et les poussins

Une poule pond N oeufs, N suit une loi de Poisson de paramètre λ. Chaque oeuf éclôt selon une loi de Bernoulli de paramètre p, indépendamment des autres oeufs. On appelle K le nombre de poussins.

1. Calculer $\mathbb{P}[K|N]$, puis $\mathbb{E}[K|N]$ et enfin $\mathbb{E}[K]$.
2. Calculer $\mathbb{P}[N|K]$ puis $\mathbb{E}[N|K]$.

Théorème. *Soit $X_1, X_2, \dots$ une suite de variables aléatoires iid à valeurs dans $\mathbb{N}$ et de fonction génératrice G_X. Soit N une variable aléatoire à valeurs dans $\mathbb{N}$, indépendante des X_i et de fonction génératrice G_N. Alors si $S_N = X_1 + \cdots + X_N$, la fonction génératrice de S_N est*

$$G_{S_N}(s) = G_N(G_X(s)).$$

3. En déduire la loi de K.
4. Écrire une fonction qui simule le processus de ponte d'oeuf ainsi que la naissance ou non des poussins.
5. Vérifier par un histogramme l'adéquation de la loi empirique à la loi théorique de K.

2 Processus de Galton-Watson : probabilité d'extinction

On étudie la transmission du nom X porté à l'origine par un seul homme. Cet homme forme la génération 0. Les descendants mâles directs de la n-ième génération forment la $(n+1)$-ième génération et la probabilité p_k qu'un homme ait k fils ($k = 0, 1, 2, \dots$) est constante au cours des générations. On suppose $0 < p_0 < 1$. Soit Z_n le nombre d'hommes portant le nom X à la n-ième génération.

On pose $x_n = \mathbb{P}[Z_n = 0]$ la probabilité d'extinction du nom à la n-ième génération et $G(s) = \sum_{k=0}^{\infty} p_k s^k$ la fonction génératrice du nombre de fils d'un homme.

1. En utilisant le théorème de l'exercice précédent, donner une relation de récurrence permettant de calculer la fonction génératrice G_n de Z_n.
2. Montrer que G est monotone et convexe sur $[0, 1]$. Discuter la valeur de $\lim_{n\to\infty} x_n$ en fonction de $m = \sum_{k=0}^{\infty} kp_k$. Conclure sur le problème de l'extinction du nom X.

3 Processus de Galton-Watson et forme autorégressive

On considère un processus de Galton-Watson, ou processus de branchement,

$$X_{n+1} = \sum_{k=1}^{X_n} Y_{n,k}$$

avec $X_0 = 1$ et $(Y_{k,n})$ est une famille de variables aléatoires indépendantes qui suivent une même loi de moyenne m et de variance σ^2.

Montrer que le processus peut s'écrire sous la forme autorégressive $X_{n+1} = mX_n + \epsilon_{n+1}$ et calculer $\mathbb{E}[\epsilon_{n+1}|X_n]$ et $\mathbb{E}[\epsilon_{n+1}^2|X_n]$.

4 Processus de Galton-Watson : cas sous-critique, critique et sur-critique

On considère un processus de Galton-Watson, ou processus de branchement, $X_{n+1} = \sum_{k=1}^{X_n} Y_{n,k}$ avec $X_0 = 1$ et $(Y_{k,n})$ est une famille de variables aléatoires indépendantes qui suivent une même loi de moyenne m et de variance $\sigma^2 > 0$.

Théorème. *On note q la probabilité d'extinction de la population.*

- *Si $m < 1$, on est dans le* ***cas sous-critique*** *: $q = 1$ donc la population va s'éteindre presque sûrement.*
- *Si $m = 1$, on est dans le* ***cas critique*** *: $q = 1$ donc la population va également s'éteindre presque sûrement. On a aussi $n\mathbb{P}[X_n > 0] \to 2/\sigma^2$ et*
$$\frac{1}{n}\mathbb{E}[X_n | X_n > 0] \to \frac{\sigma^2}{2}.$$
- *Si $m > 1$, on est dans le* ***cas sur-critique*** *: q est l'unique point fixe < 1 de la fonction génératrice associée à la loi de $Y_{n,k}$. De plus (X_n/m^n) converge p.s. et en poyenne quadratique vers une variable aléatoire finie L avec $\mathbb{E}[L] = 1$, $Var[L] = \sigma^2/(m^2 - m)$ et $\mathbb{P}[L = 0] = q$.*

1. Simuler un processus de Galton-Watson pour lequel chaque individu admet 0, 1 ou 2 descendants (avec le choix de p_0 et p_1 en entrée). Calculer l'espérance m et la probabilité d'extinction q.
2. Dans le cas où $m < 1$, représenter une évolution assez longue de la population. Pour 100 trajectoires, calculer la moyenne empirique de X_n pour chaque n et vérifier que cette moyenne divisée par m^n est proche de 1.
3. Dans le cas où $m = 1$, représenter une évolution de la population. Pour 100 trajectoires pour lesquelles $X_n > 0$, calculer la moyenne empirique de X_n sachant $X_n > 0$ avec $n = 10, 20, 50$. Vérifier que ctte moyenne est proche de $n\sigma^2/2$.
4. Dans le cas où $m > 1$, représenter une évolution de la population puis représenter plusieurs trajectoires de X_n/m^n.

5 Interprétation de code

On considère un processus de Galton-Watson, ou processus de branchement, $X_{n+1} = \sum_{k=1}^{X_n} Y_{n,k}$ avec $X_0 = 1$ et $(Y_{k,n})$ est une suite de variables aléatoires indépendantes qui suivent une loi géométrique de paramètre p.

1. Expliquez ce que le code suivant est censé faire.

```
N = input('Entrer le nombre d'itérations : ');
NR = input('Entrer le nombre de réalisations : ');
p = input('Entrer le paramètre de la loi géométrique : ');
X = zeros(NR, N + 1); X(1 : NR, 1) = 1;
for i = 1 : NR,
  for j = 1 : N,
   X(i, j+1) = sum(geornd(p, [1, X(i, j)])); %geornd crée un vecteur de taille X(i,j) de variables de loi geometrique
    if X(i, j + 1) == 0; break end; %Sortie de boucle si extinction population
  end;
end
P = sum(X == 0)/NR; %Vecteur des proportions de 0 dans chaque colonne de X
plot([0 : N], P) %Sortie graphique
```

2. Faites tournez le programme (dans le langage que vous voulez) pour $N = NR = 100$ et $p = 0.47$, puis $0.48, 0.49$ et 0.5 et comparez avec la probabilité d'extinction théorique $p/(1 - p)$.