
Soutien - Variables aléatoires discrètes

Extrait du sujet de Modélisation 2024

Dans ce sujet, on propose d'étudier différents modèles d'évolution de population, et d'étudier les conditions de son extinction. Il est composé de deux problèmes indépendants.

Le **Problème A** propose d'étudier des modèles déterministes qui mettent en valeur une condition d'extinction.

Le **Problème B** propose d'étudier le modèle probabiliste de Galton-Watson. Il est composé de deux parties ; la deuxième est dédiée à l'étude algorithmique du modèle.

Les candidats (et même les candidates) peuvent admettre le résultat d'une question ou d'une sous-question pour passer aux questions suivantes, à condition de le mentionner explicitement.

Une annexe dans laquelle certaines commandes Python sont rappelées est jointe à la fin du sujet. **Pour les questions d'informatique, on considérera que les importations de modules nécessaires ont été préalablement faites.**

Problème B. Le modèle probabiliste de Galton-Watson

I- Le modèle de Galton-Watson, exemples.

Un modèle de croissance probabiliste pour une espèce est le modèle de Galton-Watson. On considère une population dont on va décrire l'évolution génération par génération. On appelle Z_n la variable aléatoire qui compte le nombre d'individus à la génération n et on considère que :

- Les générations ne se superposent pas,
- Chaque individu a un nombre aléatoire de descendants : le nombre de descendants d'un individu est une variable aléatoire. Les variables aléatoires pour chacun sont indépendantes et de même loi.

On s'intéresse aux conditions sous lesquelles on a extinction ou survie de l'espèce. On dit que la lignée est éteinte à la génération n si $Z_n = 0$ et on souhaite étudier la suite de terme général $P(Z_n = 0)$.

Formellement, le modèle est donné par :

$$Z_0 = 1 \text{ et } \forall n \in \mathbb{N}, Z_{n+1} = \sum_{i=1}^{Z_n} X_{n,i}$$

où les variables aléatoires $(X_{n,i})_{(n,i) \in \mathbb{N} \times \mathbb{N}^*}$ sont à valeurs dans $\mathbb{N}$ et sont **indépendantes** et **de même loi**. $X_{n,i}$ est le nombre de descendants de l'individu numéro i de la génération n . On notera Y une autre variable aléatoire qui suit la même loi que les variables $X_{n,i}$.

Par exemple, si $Z_n = 12$ alors $Z_{n+1} = X_{n,1} + \dots + X_{n,12}$. Z_{n+1} est la somme du nombre de descendants de chacun des 12 individus de la génération n .

On remarquera que comme $Z_0 = 1, Z_1 = X_{0,1}$ qui est le nombre de descendants de l'unique individu de la génération 0. Ainsi Z_1 et Y suivent la même loi.

5. Que se passe-t-il si toutes les variables $X_{n,i}$ sont constantes égales à $q \in \mathbb{N}$?
6. Dans cette question, on suppose que le nombre de descendants de chaque individu suit une loi de Bernoulli de paramètre $p \in]0, 1[$.

Démontrer que pour tout $n \in \mathbb{N}^*$, $Z_n = 0$ ou $Z_n = 1$ et que :

$$P(Z_n = 1) = p^n$$

En déduire la valeur de $\lim_{n \rightarrow +\infty} P(Z_n = 0)$.

7. On définit la suite (u_n) par $u_n = P(Z_n = 0)$. Justifier que (u_n) est croissante puis qu'elle converge.

Dans la suite du sujet, on appelle la limite de (u_n) la **probabilité d'extinction de la lignée**.

8. **Étude complète dans un cas simple.** Dans cette question uniquement, la loi de reproduction est la suivante : chaque individu a une probabilité $p \in]0, 1[$ de donner deux descendants, par exemple en se divisant, et $1 - p$ de disparaître sans descendant.

- (a) Donner l'ensemble des valeurs prises par Z_1 ainsi que sa loi de probabilité. Calculer l'espérance $E(Z_1)$ et la variance $V(Z_1)$.
- (b) Démontrer que pour tout $n \in \mathbb{N}$,

$$P(Z_{n+1} = 0) = (1-p)P_{[Z_1=0]}(Z_{n+1} = 0) + pP_{[Z_1=2]}(Z_{n+1} = 0)$$

- (c) Justifier, avec une phrase, que $P_{[Z_1=2]}(Z_{n+1} = 0) = u_n^2$ puis démontrer que pour tout $n \in \mathbb{N}$,

$$u_{n+1} = (1-p) + pu_n^2$$

- (d) En déduire que les deux limites possibles de (u_n) sont 1 et $\frac{1-p}{p}$.
- (e) Démontrer que si $p \leq \frac{1}{2}$, la probabilité d'extinction vaut 1.
- (f) Si $p > \frac{1}{2}$, démontrer que pour tout $n \in \mathbb{N}$,

$$u_n \leq \frac{1-p}{p} < 1$$

En déduire la valeur de la probabilité d'extinction.

- (g) Tracer la probabilité d'extinction en fonction de $E(Z_1)$. Commenter le tracé obtenu.

9. Dans cette question uniquement, on suppose que la loi de reproduction est donnée par :

$$\forall k \in \mathbb{N}, P(Y = k) = p^k(1-p)$$

pour une certaine valeur $p \in]0, 1[$ fixée.

- (a) On admet que pour tout $k \in \mathbb{N}$, $P_{[Z_1=k]}(Z_{n+1} = 0) = u_n^k$. En utilisant un système complet d'événements associé à Z_1 , démontrer que :

$$\forall n \in \mathbb{N}, u_{n+1} = \frac{1-p}{1-pu_n}$$

- (b) On admet que (u_n) converge vers la plus petite des solutions de l'équation dont l'inconnue est ℓ :

$$\ell = \frac{1-p}{1-p\ell}$$

Déterminer la probabilité d'extinction en fonction de p .

- (c) Reconnaître la loi de $Y + 1$. En déduire que les propositions suivantes sont équivalentes :

- i. la probabilité d'extinction vaut 1.
- ii. $E(Y) \leq 1$.

Commenter.

II- Modélisation informatique

Dans cette partie, on s'intéresse à l'implémentation informatique du processus de GaltonWatson. Notre objectif est de faire des conjectures sur le comportement de la population dans le cas où la loi de reproduction est plus complexe : on considèrera dans cette partie que les variables aléatoires $X_{n,i}$ suivent la loi de Poisson de paramètre $\lambda > 0$.

On rappelle que la commande `rd.poisson(x)` simule une variable aléatoire qui suit une loi de Poisson de paramètre x .

La fonction suivante simule l'évolution d'une population et stocke le nombre d'individus à chaque génération dans une liste.

```
def galton_watson(lambda_, n):
    population = np.zeros(n + 1)
    population[0] = 1
    Z = 1
    for i in range(1, n+1):
        descendants = 0
        for j in range(Z):
            descendants += rd.poisson(lambda_)
        population[i] = descendants
        Z = descendants
        if descendants == 0:
            return population
    return population
```

11. (a) À quoi correspondent les deux arguments de la fonction `galton_watson` ?
(b) À quoi servent les lignes suivantes ?

```
        if descendants == 0:
            return population

    population = np.zeros(n + 1)

    descendants = 0
    for j in range(Z):
        descendants += rd.poisson(lambda_)
```

- (c) Pourquoi peut-on remplacer les lignes 6 à 8 par :

```
        descendants = rd.poisson(Z*lambda_)
```

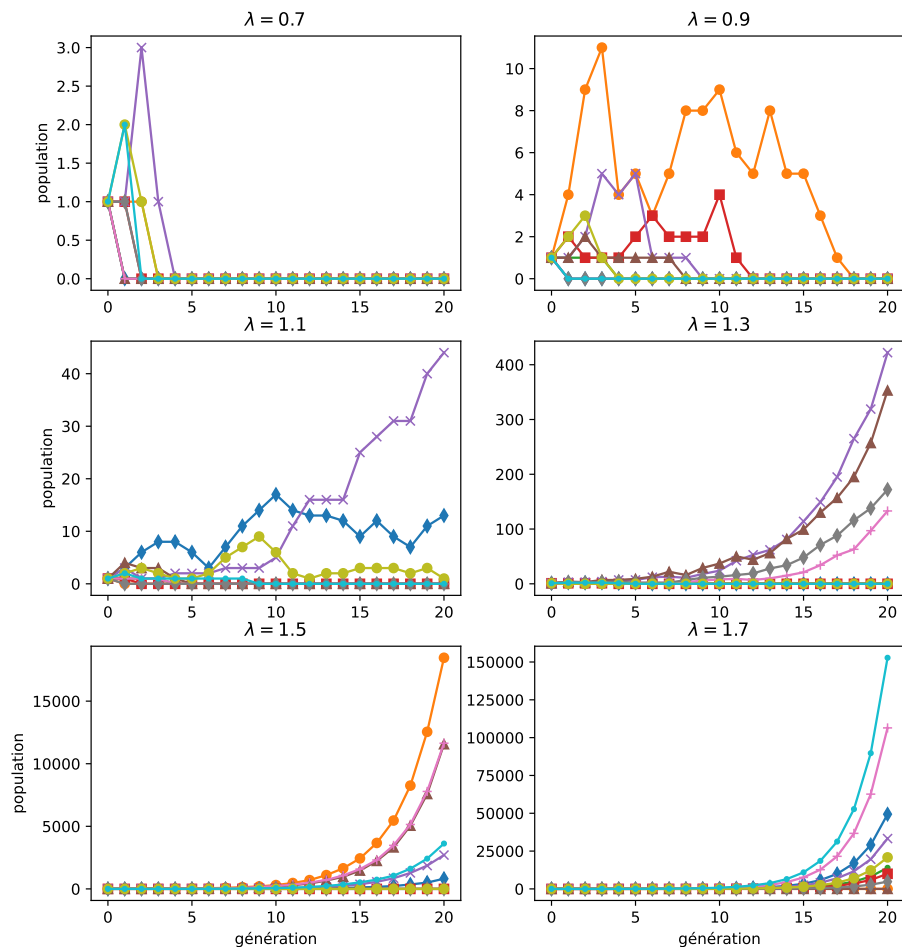
12. On souhaite réaliser :

- des simulations pour différentes valeurs de λ ,
- pour chacun des choix, plusieurs simulations.

- (a) Compléter le programme suivant pour qu'il réalise 10 simulations pour $\lambda = 0,7$ et 20 générations et les trace sur un même graphique.

```
lambda_ = 0.7
for k in ## LIGNE À COMPLÉTER ##
    plt.plot(## LIGNE À COMPLÉTER ##)
plt.show()
```

- (b) Les simulations donnent les résultats suivants pour λ variant entre 0,7 et 1,7 avec un pas de 0,2 . Quelles conjectures peut-on faire quant à la probabilité d'extinction de l'espèce ?



Les valeurs de la population sont déterminées pour des nombres entiers de générations. Des lignes ont cependant été tracées entre les valeurs de la population pour les générations successives, afin de faciliter la visualisation de l'évolution d'une population au cours des générations.

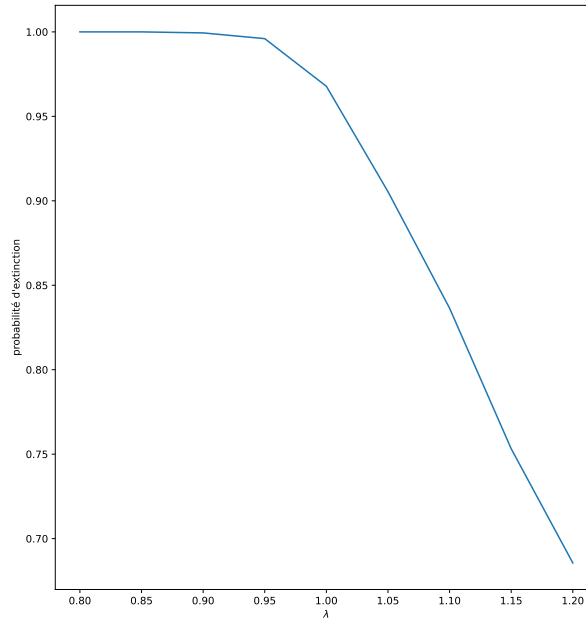
13. Dans cette question, on s'intéresse à la probabilité d'extinction.

- (a) Comment modifier la fonction `galton_watson` pour qu'elle renvoie 1 si la lignée est éteinte et 0 si la lignée n'est pas éteinte ?

Pour la suite on appelle la fonction ainsi modifiée : `galton_watson_2`.

- (b) Écrire une fonction `extinction`, qui prend en entrée un paramètre `lambda_` et qui, à partir de 5000 simulations de Galton-Watson, renvoie une approximation de la probabilité d'extinction. (On s'arrêtera à 60 générations).

La simulation pour différentes valeurs de λ donne le graphique suivant. (On a pris λ entre 0,8 et 1,2 avec un pas de 0,05).



Annexe Python

- Dans le module `matplotlib.pyplot` importé sous l'alias `plt` :
`plt.plot(X, Y)` prend en entrée deux vecteurs ou deux listes de même taille, et réalise le tracé des points d'abscisses prises dans `X` et d'ordonnées prises dans `Y`. Si on donne un seul argument à `plt.plot`, cela trace juste la suite des termes de `X`.
On utilise `plt.show()` pour afficher le tracé.
- Dans le module `numpy` importé sous l'alias `np` :
`np.zeros(n)` crée une matrice unidimensionnelle de n coefficients tous nuls.
- Dans le module `numpy.random` importé sous l'alias `rd` :
`rd.poisson(x)` simule une variable aléatoire suivant une loi de Poisson de paramètre x .