

# Chaines de Markov et épidémiologie : propagation d'une épidémie dans une population

3I005

21 novembre 2022

L'objectif de ce projet est de manipuler des chaînes de Markov pour étudier la propagation d'une épidémie dans une population. Votre rendu sera un notebook, il est attendu que les codes soient **commentés** et les résultats **interprétés**. Les packages `numpy`, `random` et `matplotlib` sont conseillés.

Nous allons étudier des populations constituées de 3 types d'individus. Chaque individu est dans un des 3 **états** suivant : sain  $S$ , infecté  $I$  ou guéri  $R$ .

## 1 Apprentissage des paramètres d'un modèle à partir de données

Dans notre modèle, nous allons considérer qu'à chaque temps :

- chaque individu sain peut rester sain ou devenir infecté
- chaque individu infecté peut rester infecté ou devenir guéri
- chaque individu guéri reste guéri

et que la probabilité de passer d'un état à l'autre **ne dépend que de l'état précédent**.

Nous disposons d'une séquence d'observations et nous souhaitons apprendre les paramètres de la chaîne de Markov permettant de modéliser le processus sous-jacent qui a généré la séquence. Nous avons suivi un individu pendant 10 jours, afin de déterminer à chaque temps dans quel état se trouvait l'individu. Nous avons obtenu la séquence d'observation suivante :  $S, S, S, I, I, I, I, I, I, R$ .

1. À partir de cette séquence d'observation, estimez les probabilités de transition entre les états et dressez la matrice de probabilité de transitions.

Nous avons ensuite suivi une population de 5000 individus, pendant 200 jours. Pour lire les données utiliser `np.loadtxt('data_exo1.txt')`. Les individus sains sont notés 0, les infectés 1 et les guéris 2.

1. Lire des données
2. Estimez les probabilités de transition entre les états et dressez la matrice de probabilité de transitions (indice pour vérifier vos résultats : la première ligne de la matrice est  $[0.9308, 0.0691, 0.]$ )

## 2 Description du premier modèle

La probabilité pour un individu d'être dans un de ces 3 états au temps  $t$ , ne dépend que de l'état dans lequel il est au temps  $t - 1$ .

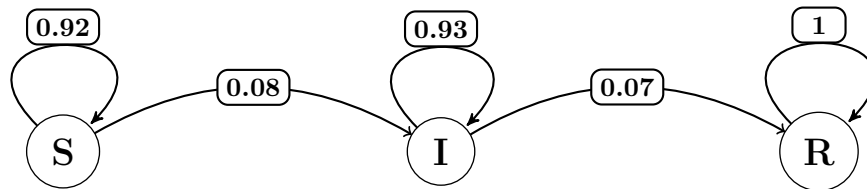


FIGURE 1 – Diagramme représentant les 3 états du modèle et les probabilités de transition

Un individu dans l'état sain a une probabilité de 0.92 de rester sain et une probabilité de 0.08 de devenir infecté. Si l'individu est infecté, il peut le rester avec une probabilité de 0.93 et être guéri avec une probabilité de 0.07. S'il est dans l'état guéri, il reste dans cet état avec une probabilité de 1.

1. A partir du graphe de transition, créez la matrice de transition  $A$ , la matrice contenant les probabilités de transition entre les différents états. Créez une fonction permettant de vérifier qu'une matrice est stochastique.

Au temps  $t = 0$ , un individu a une probabilité de 0.9 d'être sain et 0.1 d'être infecté.

2. Créez  $\pi_0$  la distribution de probabilité initiale.

Nous allons étudier l'évolution du nombre d'individu sains, infectés et guéris au cours du temps, dans un premier temps nous étudierons la distribution théorique, puis la distribution observée sur des simulations.

## Distribution $\pi_t$

On note  $\pi_t$ , le vecteur de probabilité au temps  $t$ .

1. En utilisant  $\pi_0$  et  $A$ , donnez la probabilité pour un individu d'être sain, infecté ou guéri au temps  $t = 1$  (faire d'abord le calcul à la main).
2. Donnez la probabilité pour un individu d'être sain, infecté ou guéri au temps  $t = 2$  (faire d'abord le calcul à la main).
3. De même pour chaque temps  $t$  entre 1 et 200, calculez la distribution théorique des effectifs dans chaque état (Rappel  $\pi_{t+1} = \pi_t A$ )
4. Représentez graphiquement la probabilité d'être dans chaque état en fonction du temps. (+ décrivez un peu ce que vous observez)

## Tirage aléatoire des états

Vous allez générer une séquence de taille  $T$  en utilisant cette chaîne de Markov. Pour générer une séquence aléatoire, choisissez un état initial au hasard (en utilisant  $\pi_0$ ) ; puis choisissez les états suivants en suivant les probabilités de transition (= la matrice de transition  $A$ ).

Vous pouvez prendre  $T=150$ .

## Modélisation d'une population

Vous avez généré une séquence d'état pour un individu. Maintenant vous allez générer un ensemble de séquences pour une population de 200 individus.

1. A chaque temps  $t$ , comptez le nombre d'individus sains, infectés et guéris dans la population et affichez l'évolution du nombre d'individus dans les trois états en fonction du temps.
2. Affichez le pourcentage d'individus sains infectés et guéris en fonction du temps.
3. Quand  $t$  est grand, quelle est la proportion d'individus sains, infectés et guéris ?
4. Refaites les questions précédentes avec des populations de tailles différentes, 5 individus et 5000 individus par exemple.

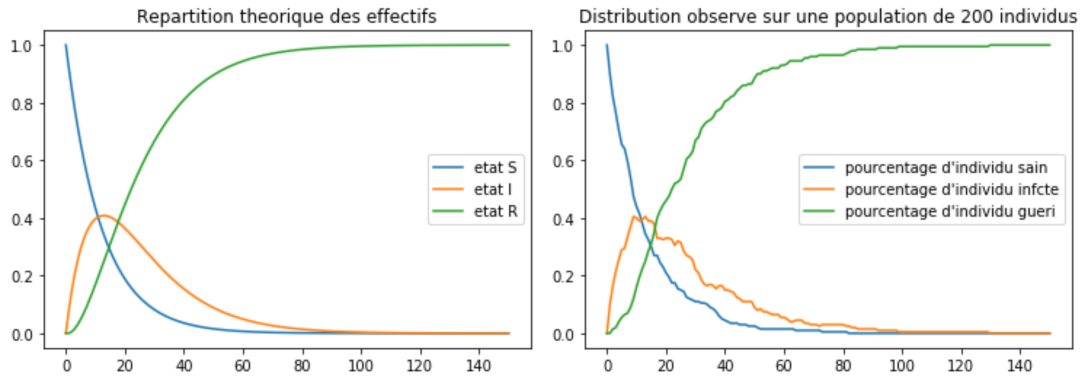


FIGURE 2 – Pourcentage théorique et observé d’individus dans chaque état en fonction du temps

### Longueur de l’infection

1. À partir des simulations, estimez la longueur moyenne d’une séquence de  $I$ .
2. Afficher la distribution observée de la longueur d’infection (peut-être que pour représenter une distribution ça serait une bonne idée de faire un histogramme).
3. Calculez la longueur théorique d’une séquence de  $I$  (vous pourrez utiliser l’espérance de la loi géométrique en justifiant).
4. Comparer la longueur estimée et la longueur théorique.
5. Affichez la distribution théorique de la longueur d’infection.

### Petites modifications autour de ce premier modèle

Vous pouvez maintenant modifier le modèle pour étudier différents cas de figure, par exemple en faisant varier

1. la taille de la population
2. la distribution de probabilité initiale ;
3. les probabilités de transition.

### 3 Description du second modèle

Nous allons maintenant considérer un second modèle, les individus guéris peuvent redevenir sains avec une probabilité de 0.02. Ils-elles peuvent perdre leur immunité face à la maladie.

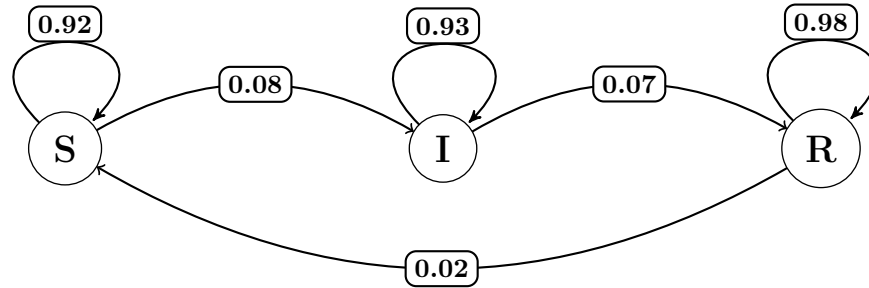


FIGURE 3 – Diagramme représentant les 3 états du modèles et les probabilités de transition

#### Analyse du modèle

1. Ce processus peut-il être modélisé par une chaîne de Markov ?
2. Donnez la nouvelle matrice de transition
3. Quels est la nature des états de cette chaîne de Markov ? Est-elle périodique ? Est-elle irréductible ?
4. Calculez la matrice  $A \times A$ . A quoi correspond-elle ? Est elle stochastique ? Même question pour  $A^3$  et  $A^4$ .
5. Réalisez les nouvelles simulations. Comment la population évolue-t-elle ?
6. Refaites les simulations avec une autre distribution de probabilité initiale (par exemple si au temps  $t = 0$ , nous avons 90% d'infectés et 10% de sains). Explorez d'autres initialisations et comme pour chaque question **commentez** vos observations.
7. Calculez la distribution de probabilité stationnaire et comparez ce résultat avec les simulations pour  $t$  assez grand.

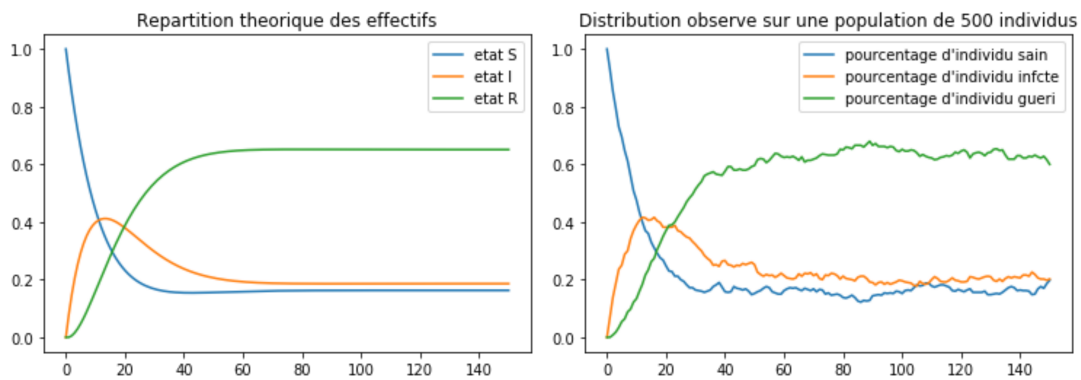


FIGURE 4 – Pourcentage théorique et observe d'individus dans chaque état en fonction du temps

#### Longueur de l'immunité

On peut se demander combien de temps un individu qui a été malade, reste protégé contre la maladie.

1. À partir des simulations, estimez la longueur moyenne d'une séquence de  $R$
2. Calculez la longueur théorique d'une séquence de  $R$
3. Affichez la distribution théorique et la distribution observée de la longueur de l'immunité.

## Modifier le modèle

1. Comment l'épidémie évolue-t-elle si vous modifiez la probabilité pour un individu sain de devenir infecté ? Quelle est la nouvelle distribution à l'équilibre ?
2. Même question si vous changez la probabilité pour un individu guéri de redevenir sain.

## 4 Confinement

On peut imaginer que si des mesures de distanciation sociale sont mises en place, la probabilité de devenir infecté devient nulle.

1. Nous allons alterner entre les périodes de non distanciation et de distanciation.
  1. En période de non-confinement, nous utilisons la matrice de transition de l'exercice précédent ;
  2. En période de confinement, la probabilité de devenir infecté pour un individu sain devient nulle
- Commencez les simulations avec la matrice de transition de l'exercice précédent. On peut considérer qu'au temps initial tous les individus sont sains.
- Quand il y a 25% d'individus infectés dans la population, nous passons en période de distanciation
- Le nombre d'individus infectés va décroître. Quand il y a moins de 10% d'infectés, le confinement est levé.
2. Faites les simulations pour une population assez grande, représentez l'évolution du nombre d'individus à chaque temps  $t$  (vous devriez voir des "vagues"), et notez les temps de confinement et de deconfinement
3. Combien de confinements/deconfinements sont nécessaires ?

## 5 Discussion

1. Quelles remarques critiques pouvez faire sur les modèles utilisés ?
2. Optionnel : Proposez des améliorations et implémentez-les si possible.