

# 3I005, Chaînes de Markov : propagation d'une épidémie dans une population

maureen.muscat@upmc.fr

20 novembre 2023

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 fortement 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  $R$ .

Nous considérons 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 (pour le premier exercice)

L'information utile pour la prédiction du futur est entièrement contenue dans l'état présent du processus (propriété de Markov).

## 1 Description du premier modèle

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

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.

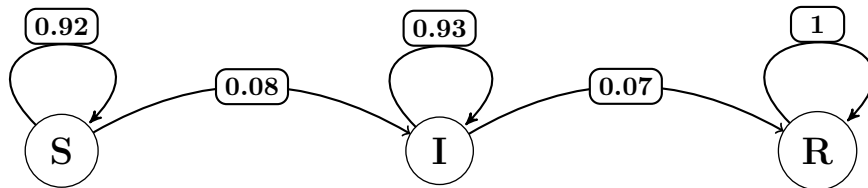


FIGURE 1 – Diagramme représentant les 3 états du modèles et les probabilités de transition

1. Combien d'états y a t'il ? Quel est la taille du vecteur de probabilité initiale, quelle est la taille de la matrice de transition ?
2. Créez  $\pi_0$  le vecteur de probabilité initiale.
3. Créez la matrice de transition  $A$ , la matrice contenant les probabilités de transition entre les différents états.
4. Créez une fonction permettant de vérifier qu'une matrice est stochastique et vérifier que  $A$  est stochastique.

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.

## 1.1 Distribution théorique

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)

## 1.2 Distribution observée

### 1.2.1 Tirage aléatoire des états

Créer une fonction permettant de générer une séquence aléatoire de taille  $T$  avec cette chaîne de Markov. (Indice : tirer au hasard le premier état en utilisant  $\pi_0$  ; puis tirer au hasard les états suivants en suivant les probabilités de transition).

Vous pouvez prendre  $T = 150$ .

### 1.2.2 Modélisation d'une population

Vous avez généré une séquence aléatoire représentant un individu, maintenant vous allez générer une population de 200 individus.

Dans cette population :

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.

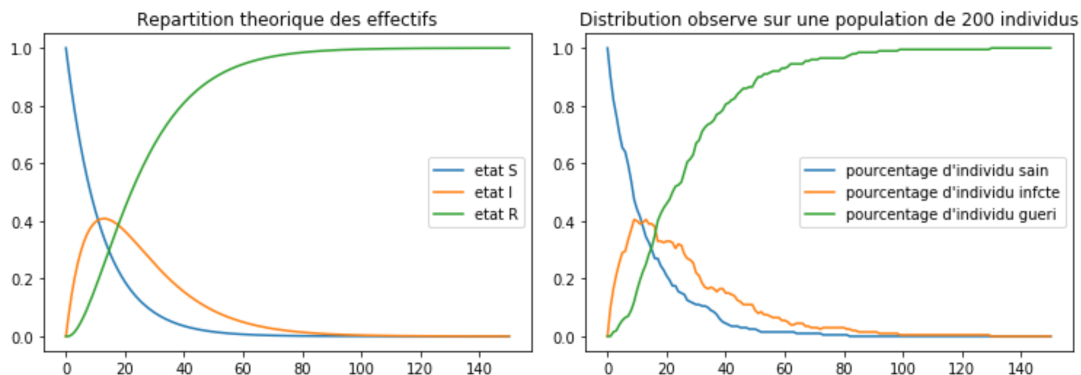


FIGURE 2 – Pourcentage théorique et observé d'individus dans chaque état en fonction du temps

### 1.3 Longueur de l'infection

1. À partir des simulations, estimez la longueur moyenne d'une séquence de  $I$ .
2. Calculez la longueur théorique d'une séquence de  $I$  (vous pourrez utiliser l'espérance de la loi géométrique **en justifiant**).
3. Comparer la longueur moyenne observée et la longueur théorique.
4. Afficher la distribution observée de la longueur d'infection (faire un histogramme).
5. Affichez et superposez la distribution théorique de la longueur d'infection et la distribution observée.

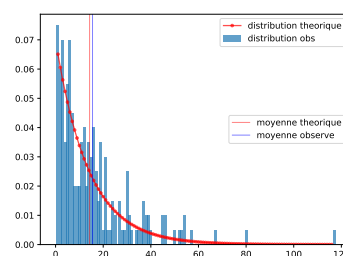
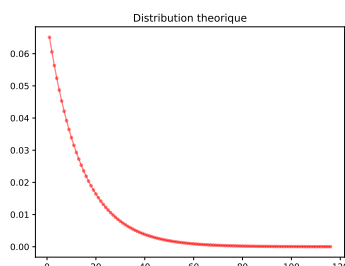
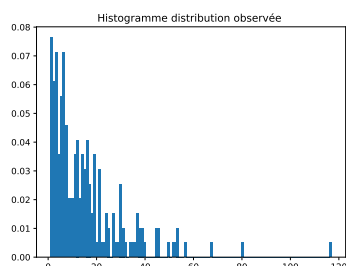


FIGURE 3 – Histogramme, distribution obs

FIGURE 4 – Loi geometrique, distribution th

FIGURE 5 – en superposant tout ca et en rajoutant les moyennes

### 1.4 Apprentissage des paramètres du modèle à partir de données

Vous allez maintenant utiliser les séquences que vous avez généré, pour estimer les paramètres du modèle.

1. A partir d'une séquence d'observation (donc un individu), estimez de la matrice de probabilité de transitions.
2. Même question avec 10 individus, 100 individus, *ect.*
3. Combien de séquences sont nécessaires pour approximer correctement le modèle.

### 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.

## 2 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.

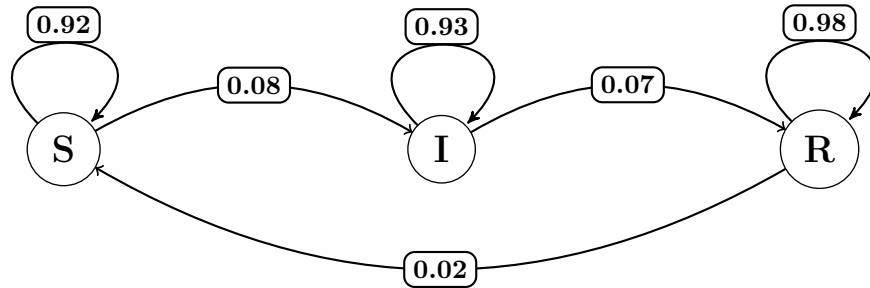


FIGURE 6 – Diagramme représentant les 3 états du modèles et les probabilités de transition

### 2.1 Analyse du modèle

1. Ce processus peut-il être modélisé par une chaîne de Markov ?
2. Donnez la nouvelle matrice de transition
3. Quelle est la nature des états de cette chaîne de Markov ? Est-elle périodique ? Est-elle irréductible ?
4. Calculez la matrice  $A^2 = A \times A$ . A quoi correspond-elle ? Est-elle stochastique ? Même question pour  $A^3$ ,  $A^4$ , etc (attention).
5. Réalisez les nouvelles simulations. Comment la population évolue-t-elle ?
6. Refaites les simulations avec d'autres distributions de proba initiales 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.

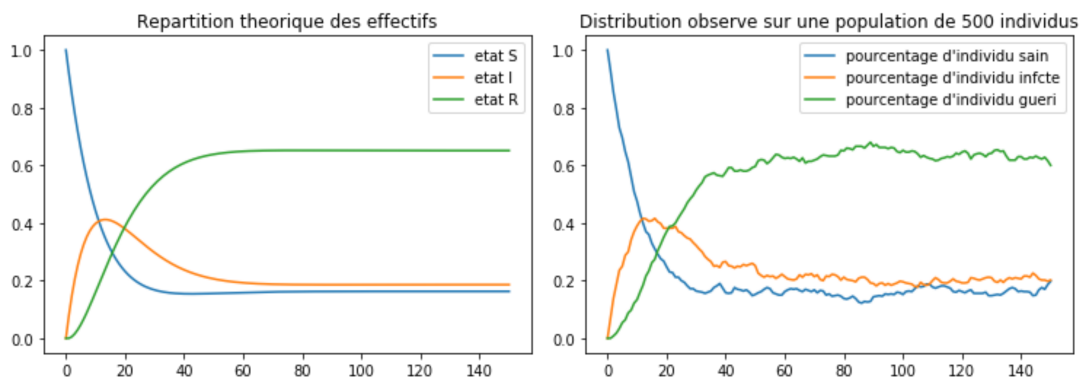


FIGURE 7 – Pourcentage théorique et observe d'individus dans chaque état en fonction du temps

### 2.2 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.

## 2.3 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.

### 3 Vaccination

Nous supposons qu'une partie de la population a été vaccinée. Voici la nouvelle matrice des probabilités de transitions chez les individus vaccinés.

```
MatriceVaccin = np.array([[0.9 , 0.01, 0.09], [0. , 0.8 , 0.2 ], [0.02, 0. , 0.98]])
```

### 3.1 Questions rapides

1. Qu'est ce que la vaccination a changé? (qu'est ce qui change par rapport au précédent modèle? )
2. Représenter graphiquement la distribution théorique
3. Représenter graphiquement et comparer les distributions stationnaires du modèle 2 et du modèle vaccin.

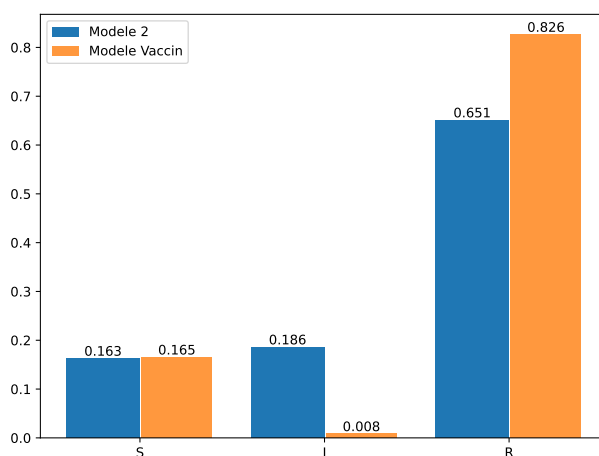


FIGURE 8 – Proba stationnaires

### 3.2 Classification

[illegible]

1. Évaluez la probabilité d'obtenir cette séquence sachant le modèle 2?
2. Évaluez la probabilité d'obtenir cette séquence sachant le modèle vaccin.
3. Selon vous est-ce-que l'individu a été vacciné? (A justifier bien sur, comme toutes les questions :) )

## 4 Confinement

Imaginons qu'il n'y ait pas de vaccins disponibles, dans ce cas 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 2 ;
  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.