

# TER : Calcul des indices de Sobol pour un modèle épidémiologique

Mièze Perrine & Leclercq David  
Sous la direction de : TRAN Viet Chi.

8 mai 2017

# Table des matières

|   |    |
|---|----|
| Introduction  | 2  |
| Chaînes de Markov : en temps discret et en temps continu                  | 3  |
| 1 Processus stochastique et chaîne de Markov en temps discret             | 4  |
| Application au modèle épidémiologique : le modèle SIR et indices de Sobol | 18 |
| 2 Introduction au modèle SIR  | 19 |
| 3 Analyse de sensibilité : indices de Sobol                               | 32 |
| 3.1 Généralités . . . . .   | 32 |
| 3.2 Application au modèle SIR . . . . .                                   | 34 |
| Bibliographie   | 37 |

# Introduction

Dans le domaine de l'épidémiologie, pour l'étude de l'évolution d'un virus dans une population, on peut considérer des modèles compartimentaux où les individus sont partitionnés en plusieurs catégories; par exemple : les individus susceptibles, infectés et rémis. Ce modèle, que nous étudierons par la suite, s'appelle le modèle SIR.

Le but de ce travail de recherche est d'étudier l'influence des taux d'infection et de rémission sur la propagation de l'épidémie. Pour cela, nous étudierons les indices de Sobol appliqué au modèle SIR. Ces indices permettent de prévoir les conséquences sur l'évolution d'une maladie selon la variation des paramètres du modèle afin de mettre en place des stratégies pour prévenir la maladie (vaccination, campagne de prévention...).

L'évolution d'une maladie dépend de plusieurs paramètres (taux d'infection, taux de rémission, lieu, âge, vaccination...). Dans notre cas, nous étudierons le modèle avec deux paramètres  $\lambda$  et  $\mu$  représentant les taux d'infection et de rémission.

Dans une première partie, nous énoncerons les notions théoriques probabilistes et statistiques utiles pour la compréhension du sujet. Le modèle SIR étant un processus continu, nous définirons ce que sont les processus stochastiques, processus de Markov, chaîne de Markov et processus de saut. Nous étudierons le processus de Poisson, qui est une chaîne de Markov en temps continu.

Dans une seconde partie, nous introduirons le modèle SIR sur population de taille finie puis dans le cas limite d'une grande population. Nous analyserons le comportement de chaque individu de la population durant une certaine durée. Nous remarquerons que le modèle peut se représenter par un processus de Poisson et que son évolution s'écrit sous la forme d'un système d'équations différentielles.

Nous ferons l'analyse de sensibilité du modèle. Pour cela, nous utiliserons une décomposition de la variance permettant de calculer les indices de Sobol.

## Première partie

# Chaînes de Markov : en temps discret et en temps continu

# Chapitre 1

## Processus stochastique et chaîne de Markov en temps discret

L'étude d'un modèle épidémiologique nécessite un cadre mathématique rigoureux. Il est donc important de définir plusieurs notions indispensables à sa compréhension. Définissons tout d'abord ce qu'est un processus stochastique.

**Définition 1** (Processus stochastique).

*Il y a deux façons d'appréhender les processus stochastiques :*

1. *Soit  $\mathbb{I} = \mathbb{N}$  ou  $\mathbb{R}_+$ . Un processus stochastique est une famille de variables aléatoires indexées par  $\mathbb{I}$ . On le note  $(X_t, t \in \mathbb{I})$ .*
2. *On peut aussi voir un processus stochastique comme une variable à valeurs dans un espace de trajectoires.*
  - *Dans le cas discret, un processus stochastique est une variable aléatoire à valeurs dans  $\mathbb{R}^{\mathbb{N}}$ .*
  - *Dans le cas continu, un processus stochastique est une variable aléatoire dans  $C^0$  ou continue à droite et limitée à gauche.*

**Remarque 1.** *Un exemple type de processus stochastique est la chaîne de Markov, que nous allons définir ci-dessous.*

### I. Chaîne de Markov à temps discret

**Définition 2** (Chaîne de Markov à temps discret). *Un processus de Markov est un processus stochastique qui vérifie la propriété de Markov : le comportement aléatoire dans le futur du processus conditionnellement au passé ne dépend du passé que par l'information donnée par son état présent.*

Autrement dit :

$$\forall n \geq 0, \forall (i_1, \dots, i_{n-1}, i, j) \in \mathbb{R}^{n+1}$$

$$\mathbb{P}(X_{n+1} = j \mid X_1 = i_1, \dots, X_{n-1} = i_{n-1}, X_n = i) = \mathbb{P}(X_{n+1} = j \mid X_n = i)$$

**Exemple 1.** Une chaîne de Markov permet de modéliser la cotisation d'assurance automobile versée par un assuré : la cotisation versée à l'année  $n + 1$  dépend de la cotisation versée à l'année  $n$  et des accidents commis lors cette même année.

Ici, on note  $X_i$  la variable représentant la cotisation d'assurance à la  $i^e$  année. Le comportement de  $X_{i+1}$  ne dépend pas des  $X_{i-k}$ , pour  $k \in \{1, \dots, i - 1\}$  dès lors qu'on connaît  $X_i$ .

Il existe plusieurs cas particuliers de chaîne de Markov. Nous allons définir l'une d'entre elles :

**Définition 3** (Chaîne de Markov homogène). On dit qu'une chaîne de Markov  $(X_1, \dots, X_n, \dots)$  est homogène si :

$$\forall n \in \mathbb{N}, \forall (i, j) \in \mathbb{R}^2, \\ \mathbb{P}(X_{n+1} = j \mid X_n = i) \text{ ne dépend pas de } n.$$

**Conséquence 1.**

$$\forall n \in \mathbb{N}^*, \forall (i, j) \in \mathbb{R}^2, \\ \mathbb{P}(X_{n+1} = j \mid X_n = i) = \mathbb{P}(X_n = j \mid X_{n-1} = i).$$

**Exemple 2** (Exemple de la marche aléatoire sur  $\mathbb{Z}$ ). Soient  $(X_n)_{n \in \mathbb{N}}$  une suite de variables aléatoires indépendantes et de même loi, à valeurs dans  $\{-1, 1\}$ .

Soit  $p \in [0, 1]$  la probabilité d'avancer ( $X_i = 1, i \in \mathbb{N}$ ) et  $1 - p$  la probabilité de reculer ( $X_i = -1, i \in \mathbb{N}$ ).

Ce modèle est une chaîne de Markov homogène. En effet :

$$\mathbb{P}(X_{i+2} = 1 \mid X_{i+1} = -1) = \mathbb{P}(X_{i+1} = 1 \mid X_i = -1), \quad \forall i \in \mathbb{N}.$$

## II. Processus de saut et chaîne de Markov en temps continu

### 1. Processus de saut

Nous allons maintenant définir la notion de processus de saut qui englobe les chaînes de Markov en temps continu. Ceci nous permettra de faire intervenir le processus de Poisson dans la prochaine partie.

**Définition 4** (Processus de saut). *Un processus de saut  $(X_t, t \geq 0)$  est un processus dont les trajectoires sont presque-sûrement continues à droite, limitées à gauche et constantes entre des instants de sauts.*

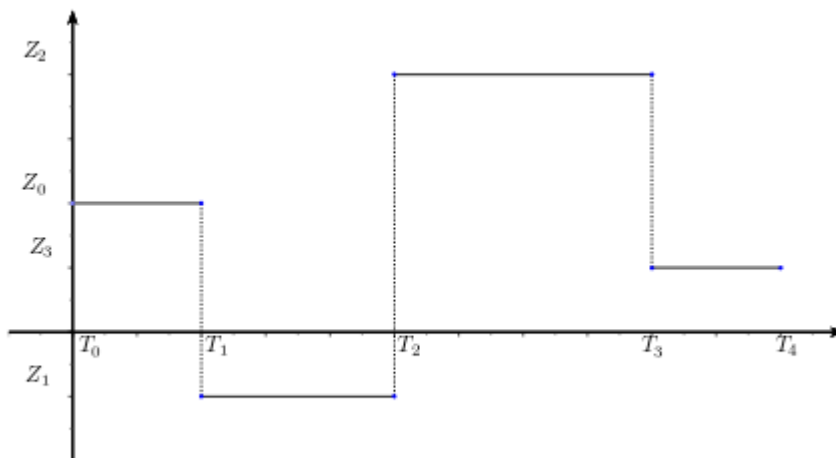


FIG. 1.1 : Exemple de processus de saut

**Notation 1** (Notations de la Figure 1.1). *Soit  $T_n \in \mathbb{R} \cup \{+\infty\}$  le  $n$ -ième instant de saut.*

*Soit  $Z_n = X_{T_n} \in E$  la position du processus à l'instant  $T_n$ .*

*La donnée du processus  $(X_t, t \geq 0)$  est équivalente à la donnée de la suite de variables aléatoires  $(T_n, Z_n)$ ,  $n \in \mathbb{N}$ .*

*De plus,  $T_0 = 0$  et les instants de saut  $T_n \in \mathbb{R} \cup \{+\infty\}$  forment une suite croissante telle que :*

$$T_n(\omega) < T_{n+1}(\omega) \text{ si } T_n(\omega) < +\infty.$$

## 2. Chaîne de Markov en temps continu

Pour l'étude du modèle épidémiologique, nous avons besoin de la notion de chaîne de Markov en temps continu.

**Définition 5** (Chaîne de Markov en temps continu).

1. Un processus de saut  $(X_t, t \geq 0)$  à valeurs dans  $\mathbb{E}$  ensemble dénombrable est un processus markovien de saut, appelé également chaîne de Markov en temps continu, si :

$\forall 0 < s < t$ , la loi conditionnelle de la variable aléatoire  $X_t$  sachant  $(X_u, u \leq s)$  ne dépend que de  $X_s$ .

Autrement dit :

$$\forall n \in \mathbb{N}, \quad \forall 0 \leq t_0 < t_1 < \dots < t_n < s, \quad \forall i_0, i_1, \dots, i_n, i, j \in \mathbb{E} :$$

$$\mathbb{P}(X_t = j \mid X_{t_0} = i_0, \dots, X_{t_n} = i_n, X_s = i) = \mathbb{P}(X_t = j \mid X_s = i).$$

2. Soient  $0 < s < t$  et  $i, j \in \mathbb{E}$  ensemble dénombrable.

On dit que ce processus markovien est homogène si la probabilité

$\mathbb{P}(X_t = j \mid X_s = i)$  ne dépend de  $s$  et  $t$  que par la différence  $t - s$ .

**Notation 2.** Soient  $0 < s < t$  et  $i, j \in \mathbb{E}$ .

On note  $\mathbb{P}(X_t = j \mid X_s = i) = P_{i,j}(t-s)$  pour un processus de saut Markovien homogène.

**Définition 6** (Noyau de transition). Soient  $(\mathbb{X}, \epsilon)$  et  $(\mathbb{Y}, \tau)$  deux ensembles mesurables. Soit  $x \in \mathbb{X}$  et  $dy \in \mathbb{Y}$ .

$(P(x, dy))$  est un noyau de transition si :

- $x \rightarrow P(x, dy)$  est  $\epsilon$ -mesurable.
- $\forall x, P(x, dy)$  est une probabilité sur  $\mathbb{Y}$

Nous pouvons alors introduire la notion de matrice markovienne (opérateur markovien).

**Définition 7** (Matrice markovienne). Une matrice markovienne, notée  $P(t) = (P_{i,j}(t))_{i,j \in \mathbb{E}}$ ,  $\forall t > 0$ , est un noyau de transition tel que :

$$\forall i, j \in \mathbb{N}, \forall t > 0,$$

$$P_{i,j}(t) \geq 0,$$

$$\sum_{j \in \mathbb{N}} P_{i,j}(t) = 1$$

La matrice définie par les transitions de la notation 2 est une matrice markovienne.



**Proposition 1.** *Les lois fini-dimensionnelles d'un processus markovien de saut homogène ne dépendent que de sa loi initiale  $\mu$  et de sa matrice de transition  $(P(t))_{t \in \mathbb{R}^+}$ . Pour tout  $n \in \mathbb{N}$ , pour tout  $0 < t_1 < \dots < t_n$ , la loi du vecteur aléatoire  $(X_0, X_{t_1}, \dots, X_{t_n})$  est donnée par :*

$$\begin{aligned} & \forall i_0, i_1, \dots, i_n \in \mathbb{N}, \\ & \mathbb{P}(X_0 = i_0, X_{t_1} = i_1, \dots, X_{t_n} = i_n) \\ & = \mu_{i_0} P_{i_0, i_1}(t_1) P_{i_1, i_2}(t_2 - t_1) \times \dots \times P_{i_{n-1}, i_n}(t_n - t_{n-1}). \end{aligned}$$

*Démonstration.* Rappelons la formule des probabilités conditionnelles pour deux variables. Soient  $X$  et  $Y$  deux événements aléatoires :

$$\mathbb{P}(X | Y) = \frac{\mathbb{P}(X, Y)}{\mathbb{P}(Y)}.$$

On l'applique ici pour  $n + 1$  variables.

$$\begin{aligned} & \mathbb{P}(X_0 = i_0, X_{t_1} = i_1, \dots, X_{t_n} = i_n) \\ & = \mathbb{P}(X_0 = i_0) \mathbb{P}(X_{t_1} = i_1 | X_0 = i_0) \\ & \quad \mathbb{P}(X_{t_2} = i_2 | X_0 = i_0, X_{t_1} = i_1) \times \dots \times \\ & \quad \mathbb{P}(X_{t_n} = i_n | X_0 = i_0, \dots, X_{t_{n-1}} = i_{n-1}) \\ & = \mu_{i_0} P_{i_0, i_1}(t_1) \times \dots \times P_{i_{n-1}, i_n}(t_n - t_{n-1}). \end{aligned}$$

Avec  $\mu_{i_0}$  la loi initiale définie comme  $\mathbb{P}(X_0 = i_0)$ . □

**Proposition 2** (Equation de Chapman-Kolmogorov).

*Les matrices  $(P(t))_{t \in \mathbb{R}^+}$  vérifient :*

- $P(0) = Id$
- *l'équation de Chapman Kolmogorov, aussi appelée relation de semi-groupe :*

$$P(t + s) = P(t)P(s) = P(s)P(t)$$

*Démonstration.* Soient  $i, k \in \mathbb{N}$ . On a :

•

$$P_{i,k}(0) = \mathbb{P}(X_0 = k | X_0 = i) = \begin{cases} 0 & \text{si } k \neq i \\ 1 & \text{si } k = i \end{cases}$$

- Pour  $s, t > 0$  :

$$\begin{aligned}
P_{i,k}(t+s) &= \mathbb{P}(X_{t+s} = k \mid X_0 = i) \\
&= \sum_{j \in \mathbb{N}} \mathbb{P}(X_t = j, X_{t+s} = k \mid X_0 = i) \\
&= \sum_{j \in \mathbb{N}} \mathbb{P}(X_{t+s} = k \mid X_t = j) \mathbb{P}(X_t = j \mid X_0 = i) \\
&= \sum_{j \in \mathbb{N}} P_{i,j}(t) P_{j,k}(s).
\end{aligned}$$

D'où  $P(t+s) = P(t)P(s)$ . □

**Définition 8** (Matrice génératrice). *On appelle matrice génératrice  $A$  du processus markovien de saut de semi-groupe  $(P(t))_{t \in \mathbb{R}^+}$  la matrice des dérivées en 0 de la matrice de transition :*

$$\forall i, j \in \mathbb{E}, \quad A(i, j) = \frac{\partial}{\partial t} P_{i,j}(0) = \lim_{t \rightarrow 0} \frac{P_{i,j}(t) - P_{i,j}(0)}{t}.$$

**Conséquence 2.** *Pour  $f : \mathbb{N} \rightarrow \mathbb{R}^+, \forall i \in \mathbb{N}$*

$$\begin{aligned}
\mathcal{A}f(i) &= \sum_{j \in \mathbb{N}} \mathcal{A}(i, j) f(j) \\
&= \sum_{j \in \mathbb{N}} \frac{\partial}{\partial t} P_{i,j}(0) f(j) \\
&= \frac{\partial}{\partial t} \sum_{j \in \mathbb{N}} P_{i,j}(0) f(j) \\
&= \frac{\partial}{\partial t} P_0 f(i).
\end{aligned}$$

Où  $P_0$  est la matrice de transition évaluée en 0.

Donc,  $\frac{\partial}{\partial t} P(t) = \mathcal{A}P(t) = P(t)\mathcal{A}$ .

En effet, on a :  $P(t+h) = P(t)P(h)$  par l'équation de Chapman Kolmogorov.

$$\begin{aligned}
\frac{P(t+h) - P(t)}{h} &= \frac{P(t)P(h) - P(t)}{h} \\
&= P(t) \frac{P(h) - I}{h}.
\end{aligned}$$

De même :

$$\begin{aligned}
\frac{P(t+h) - P(t)}{h} &= \frac{P(h)P(t) - P(t)}{h} \\
&= \frac{P(h) - I}{h} P(t).
\end{aligned}$$

Or  $P(0) = I$  et  $\lim_{h \rightarrow 0} \frac{P(h) - I}{h} = \mathcal{A}$ . On obtient bien le résultat souhaité.

**Remarque 2.** Une matrice génératrice  $A$  vérifie :

- $\forall i \neq j, A(i, j) \geq 0$
- $\forall i, \sum_j A(i, j) = 0$

Le processus de Poisson est un cas particulier du processus de saut. Il nécessite quelques notions que nous allons d'abord définir.

**Définition 9** (Processus ponctuel).

Un processus ponctuel sur  $\mathbb{E}$  se décrit comme la donnée d'une suite presque sûrement croissante de temps aléatoires  $0 < T_1 < T_2 < \dots < T_n < \dots$ , définis sur un espace de probabilité  $(\Omega, \mathcal{F}, \mathbb{P})$  à valeurs dans  $\mathbb{E}$  et vérifiant  $T_n \rightarrow +\infty$  presque sûrement, quand  $n \rightarrow +\infty$ .

On peut le représenter comme :

$$\mathcal{N}(dt) = \sum_{i \geq 1} \delta_{T_i}.$$

On l'associe en général à la fonction aléatoire de comptage.

**Définition 10** (Fonction aléatoire de comptage). La fonction aléatoire de comptage  $(\mathcal{N}_t)_{t \geq 0}$  associée au processus ponctuel  $\{T_n, n \in \mathbb{N}\}$  est définie par :

$$\mathcal{N}_t = \mathcal{N}([0, t]) = \sup\{n, T_n \leq t\} = \sum_{i \geq 1} \mathbf{1}_{T_i \leq t}.$$

Nous pouvons maintenant définir un processus de Poisson.

**Définition 11** (Processus de Poisson). Le processus ponctuel  $\{\mathcal{N}_t, t \geq 0\}$  est appelé processus de Poisson de mesure d'intensité  $\lambda(dx)$  si :

- Soit  $A$  un ensemble mesurable. Alors  $\mathcal{N}(A) \sim \mathcal{P}(\lambda(A))$ .
- Soient  $m \in \mathbb{N}$ ,  $A_1, A_2, \dots, A_m$  des ensembles mesurables disjoints deux à deux. Alors :  $\mathcal{N}(A_i) \perp \mathcal{N}(A_j)$ ,  $i \neq j$

Cette définition est équivalente à la définition suivante, lorsque  $\lambda(dx) = \lambda \cdot dx$ .

**Définition 12.** Le processus ponctuel  $\{\mathcal{N}_t, t \geq 0\}$  à valeurs dans  $\mathbb{R}^+$  est appelé processus de Poisson de mesure d'intensité  $\lambda dx$  où  $\lambda \in \mathbb{R}^+$  et  $dx$  est la mesure de Lebesgue sur  $\mathbb{R}^+$  si :

Pour  $0 \leq s < t$ , la loi de  $\mathcal{N}_t - \mathcal{N}_s$  ne dépend de  $s$  et  $t$  que par la différence  $t - s$  : la loi de  $\mathcal{N}_t - \mathcal{N}_s$  est égale à la loi de  $\mathcal{N}_{t-s}$  et  $\mathcal{N}_t - \mathcal{N}_s \perp \sigma(\mathcal{N}_u, u \geq s)$ .

**Définition 13.** Le paramètre  $\lambda$  est appelé intensité du processus de Poisson. Il est égal au nombre moyen d'événements qui se produisent pendant une unité de temps, puisque :

$$\mathbb{E}(\mathcal{N}_{t+1} - \mathcal{N}_t) = \lambda.$$

**Remarque 3.** On a, par définition de la loi de Poisson :

$$\mathbb{P}(\mathcal{N}_t = k) = e^{-\lambda t} \frac{(\lambda t)^k}{k!}$$

**Remarque 4.** Remarquons ces quelques égalités qui seront utiles dans la preuve.

Soit  $n \in \mathbb{N}$  et  $t, s \in \mathbb{R}^+$ . Alors :

$$\begin{aligned} \{T_n > t\} &= \{\mathcal{N}_t < n\} \\ \{\mathcal{N}_t \geq n\} &= \{T_n \leq t\} \\ \{\mathcal{N}_t = n\} &= \{T_n \leq t < T_{n+1}\} \\ \{\mathcal{N}_t \geq n > N_s\} &= \{s < T_n \leq t\} \end{aligned}$$

On rappelle la définition de la fonction génératrice.

**Définition 14.** Soit  $X$  une variable aléatoire discrète, soit  $u \in [0, 1]$ . On note  $g(u)$  la fonction génératrice de  $X$  et on la définit comme :

$$g(u) = \mathbb{E}(u^X) = \sum_{k \geq 0} \mathbb{P}(X = k) u^k.$$

**Remarque 5** (Fonction génératrice d'une loi de Poisson). Soit  $X \sim \mathcal{P}(\lambda)$ . On note  $g_X$  la fonction génératrice de  $X$  :

$$\begin{aligned} g_X(u) &= \mathbb{E}(u^X) \\ &= \sum_{k=0}^{+\infty} u^k e^{-\lambda} \frac{\lambda^k}{k!} \\ &= e^{-\lambda} \sum_{k=0}^{+\infty} \frac{(\lambda u)^k}{k!} \\ &= e^{-\lambda} e^{\lambda u} = e^{-\lambda(1-u)}. \end{aligned}$$

*Démonstration.* ( $\Rightarrow$ )

$[s, t]$  et  $[0, s[$  sont deux ensembles disjoints.

D'après la définition 11,  $\mathcal{N}_{t-s}$  et  $\mathcal{N}_s$  sont indépendants.

$$\mathcal{N}([0, t]) = \mathcal{N}_t \sim \mathcal{P}(\lambda t),$$

$$\mathcal{N}([0, s]) = \mathcal{N}_s \sim \mathcal{P}(\lambda s),$$

$$\mathcal{N}([s, t]) = \mathcal{N}_{t-s} \sim \mathcal{P}(\lambda(t-s))$$

Montrons que  $\mathcal{N}_t - \mathcal{N}_s \stackrel{\text{loi}}{=} \mathcal{N}_{t-s}$

$$\begin{aligned}
& \mathbb{P}(\mathcal{N}_{t-s} + \mathcal{N}_s = k) \\
&= \sum_{j=0}^k \mathbb{P}(\mathcal{N}_{t-s} = j, \mathcal{N}_s = k - j) \\
&= \sum_{j=0}^k e^{-\lambda(t-s)} \frac{\lambda^j (t-s)^j}{j!} e^{-\lambda s} \frac{\lambda^{k-j} s^{k-j}}{(k-j)!} \\
&= e^{-\lambda t} \lambda^k \sum_{j=0}^k \frac{(t-s)^j s^{k-j}}{j!(k-j)!} \\
&= \frac{e^{-\lambda t} \lambda^k}{k!} \sum_{j=0}^k \frac{k!}{j!(k-j)!} (t-s)^j s^{k-j} \\
&= \frac{e^{-\lambda t} \lambda^k}{k!}.
\end{aligned}$$

Ainsi la loi de  $\mathcal{N}_t = \mathcal{N}_{t-s} + \mathcal{N}_s$  est la même que celle de  $\mathcal{N}_{t-s} + \mathcal{N}_s$ . Donc on peut identifier la loi de  $\mathcal{N}_{t-s}$  à la loi de  $\mathcal{N}_t - \mathcal{N}_s$ .

( $\Leftarrow$ )

$0 \leq s < t$ , avec  $s, t \in \mathbb{R}$

Par hypothèse,  $\mathcal{N}_t - \mathcal{N}_s$  a la même loi que  $\mathcal{N}_{t-s}$ . Vérifions que  $\mathcal{N}_t \sim \mathcal{P}(\lambda t)$ .

Nous remarquons tout d'abord que  $[s, t]$  et  $[0, s[$  sont deux ensembles disjoints. Par la définition 12, on a :

$$\mathcal{N}_t - \mathcal{N}_s \perp\!\!\!\perp \sigma(\mathcal{N}_u, u \geq s).$$

Ainsi,  $\mathcal{N}_{t-s}$  et  $\mathcal{N}_s$  sont indépendants.

Notons  $g_t(u)$  la fonction génératrice de  $\mathcal{N}_t$  appliquée à  $u$ . Vérifions que :

$$g_t(u) = \exp(-(\lambda t(1 - u)))$$

On a :

$$\begin{aligned}
g_t(u) &= g_{t-s+s}(u) \\
&= \mathbb{E}(u^{\mathcal{N}_{t-s} + \mathcal{N}_s}) \\
&= \mathbb{E}(u^{\mathcal{N}_{t-s}}) \mathbb{E}(u^{\mathcal{N}_s}) \\
&= \mathbb{E}(u^{\mathcal{N}_{t-s}}) \mathbb{E}(u^{\mathcal{N}_s}) \\
&= g_{t-s}(u) g_s(u).
\end{aligned}$$

Cas où  $t \in \mathbb{N}^*$  :

$$g_t(u) = g_1(u)g_{t-1}(u) = g_1(u)^t$$

Cas où  $t \in \mathbb{Q}^*$ . Soit  $t = \frac{p}{q}$ .

$$g_t(u) = g_{\frac{p}{q}}(u) = g_{\frac{1}{q}}(u)^p$$

$$\text{Or : } g_1(u) = g_{\frac{1}{q}}(u)^q = \prod_{k=1}^q g_{\frac{1}{q}}(u)$$

$$g_{\frac{1}{q}}(u) = g_1(u)^{\frac{1}{q}}$$

$$\text{Donc : } g_t(u) = g_1(u)^{\frac{p}{q}} = g_1(u)^t.$$

Si  $t \in \mathbb{R}^+$ , on déduit par densité de  $\mathbb{Q}$  dans  $\mathbb{R}$  que :  $g_t(u) = g_1(u)^t$ .

On remarque également que :  $g_1(u) \leq 1, \forall u \in [0, 1]$ .

On a :  $\mathcal{N}_1 \geq 0 \Rightarrow u^{\mathcal{N}_1} \leq 1, \forall u \in [0, 1] \Rightarrow \mathbb{E}(u^{\mathcal{N}_1}) \leq 1$

Donc  $g_1(u) \leq 1$ .

Etudions la fonction  $g_t(u)$  avec  $u \in [0, 1]$ .

- On remarque la décroissance en  $t$  de  $g_t(u)$  par décroissance de la fonction  $u^t$ .
- De plus,

$$g_t(u) = \sum_{k \geq 0} \mathbb{P}(\mathcal{N}_t = k) u^k \geq \mathbb{P}(\mathcal{N}_t = 0) = \mathbb{P}(T_1 > t) \rightarrow 1 \text{ quand } t \rightarrow 0$$

Donc  $g_t$  a une forme exponentielle. On note :  $g_t(u) = e^{-\lambda(u)t}$   
avec  $\lambda(u) > 0$  fonction de  $u$  ( $\lambda(u) = -\ln g_1(u)$ ).

Il est également facile de voir que  $g_1(u) \neq 0$  en constatant que :

$$g_1(u) = \sqrt[t]{g_t(u)} \text{ et } g_t(u) \geq \mathbb{P}(T_1 > t) \rightarrow 1 \text{ quand } t \rightarrow 0$$

De plus, on a :

$$g_1(u) \leq 1 \Rightarrow g_t(u) = (g_1(u))^t \rightarrow 0 \text{ quand } t \rightarrow +\infty$$

Montrons que  $\lambda(u)$  est de la forme  $\lambda(0)(1 - u)$ .

$$\begin{aligned}
 g_t(u) &= e^{-\lambda(u)t} \\
 -\lambda(u)t &= \ln(g_t(u)) = \ln(1 + g_t(u) - 1) \\
 \lambda(u) &= \frac{-1}{t} \ln(1 + g_t(u) - 1) \quad \forall t. \\
 &= \lim_{t \rightarrow 0} \frac{1}{t} (1 - g_t(u)) \\
 &= \lim_{t \rightarrow 0} \frac{1}{t} \left[ \sum_{k \geq 1} \mathbb{P}(\mathcal{N}_t = k)(1 - u^k) \right].
 \end{aligned}$$

On remarque :

$$\begin{aligned}
 0 &\leq \frac{1}{t} (1 - g_t(u)) - \frac{1}{t} \mathbb{P}(\mathcal{N}_t = 1)(1 - u) \\
 &= \frac{1}{t} \left[ \sum_{k \geq 1} \mathbb{P}(\mathcal{N}_t = k)(1 - u^k) - \mathbb{P}(\mathcal{N}_t = 1)(1 - u) \right] \\
 &= \frac{1}{t} \sum_{k \geq 2} \mathbb{P}(\mathcal{N}_t = k)(1 - u^k) \\
 &\leq \frac{1}{t} \mathbb{P}(\mathcal{N}_t \geq 2).
 \end{aligned}$$

Supposons que :

$$\begin{aligned}
 \frac{1}{t} \mathbb{P}(\mathcal{N}_t \geq 2) &\rightarrow 0 \quad \text{quand } t \rightarrow 0. \\
 0 &\leq \frac{1}{t} (1 - g_t(u)) - \frac{1}{t} \mathbb{P}(\mathcal{N}_t = 1)(1 - u) \rightarrow 0 \quad \text{quand } t \rightarrow 0. \\
 \lim_{t \rightarrow 0} \frac{1}{t} (1 - g_t(u)) &= \lim_{t \rightarrow 0} \frac{1}{t} \mathbb{P}(\mathcal{N}_t = 1)(1 - u) = \lambda(u).
 \end{aligned}$$

On pose  $\lambda(0) = \lim_{t \rightarrow 0} \frac{1}{t} \mathbb{P}(\mathcal{N}_t = 1)$ .

$$\lambda(u) = \lambda(0)(1 - u) \Rightarrow g_t(u) = e^{-\lambda(0)t(1-u)}$$

On obtient bien la fonction génératrice d'une variable aléatoire suivant une loi de Poisson  $\mathcal{P}(\lambda(0)t)$ .

□

**Exemple 3.** Voici quelques exemples où on peut utiliser les processus de Poisson pour la modélisation :

- Appel téléphonique à un standard
- Arrivée des clients à un guichet
- Modèle SIR en épidémiologie (nous reviendrons sur ce modèle par la suite)

On peut en déduire la loi du premier temps de saut.

**Corollaire 1.** La loi du premier temps de saut  $T_1$  est une loi exponentielle de paramètre  $\lambda$ . De même, pour tout  $s > 0$ , la loi du premier évènement après  $s$ , soit  $T_{N_s+1} - s$ , est une loi exponentielle de paramètre  $\lambda$ .

*Démonstration.* On remarque que :

$$\{T_{N_s+1} - s > t\} = \{T_{N_s+1} > t + s\} = \{\mathcal{N}_{t+s} = \mathcal{N}_s\}$$

- Pour  $t > 0$ , on a :

$$\mathbb{P}(T_1 > t) = \mathbb{P}(N_t = 0) = e^{-\lambda t}$$

- De même, on a :

$$\mathbb{P}(T_{N_s+1} - s > t) = \mathbb{P}(N_{s+t} - N_s = 0) = \mathbb{P}(N_t = 0).$$

□

**Proposition 3.** Le processus de Poisson d'intensité  $\lambda$  est un processus de Markov. Sa matrice de transition est donnée par :

$$P_{i,j}(t) = e^{-\lambda t} \frac{(\lambda t)^{j-i}}{(j-i)!} \quad \text{si } j \geq i.$$

Si  $i > j$ , alors  $P_{i,j}(t) = 0$ .



*Démonstration.* On rappelle que :

$$\mathbb{P}(\mathcal{N}_t = j \mid \mathcal{N}_s = i) = P_{i,j}(t - s)$$

$$\begin{aligned} P_{i,j}(t) &= \mathbb{P}(\mathcal{N}_{t+s} = j \mid \mathcal{N}_s = i) \\ &= \frac{\mathbb{P}(\mathcal{N}_{t+s} - \mathcal{N}_s = j - i, \mathcal{N}_s = i)}{\mathbb{P}(\mathcal{N}_s = i)} \\ &= \frac{\mathbb{P}(\mathcal{N}_t = j - i) \mathbb{P}(\mathcal{N}_s = i)}{\mathbb{P}(\mathcal{N}_s = i)} \\ &= \mathbb{P}(\mathcal{N}_t = j - i) \\ &= e^{-\lambda t} \frac{(\lambda t)^{j-i}}{(j-i)!}. \end{aligned}$$

Par définition de  $\mathcal{N}_t$ , on a obligatoirement  $j \geq i$ . Sinon, la probabilité est nulle.  $\square$

Nous pouvons donc construire la matrice génératrice du processus de Poisson.

**Proposition 4.** *La matrice génératrice d'un processus de Poisson d'intensité  $\lambda$ , notée  $\mathcal{A} \in \mathcal{M}_{\mathbb{N},\mathbb{N}}(\mathbb{R})$  est de la forme :*

- $\mathcal{A}_{i,j} = 0$ , si  $j \neq i + 1$
- $\mathcal{A}_{i,i+1} = \lambda$
- $\mathcal{A}_{i,i} = -\lambda$

*Démonstration.* Calculons tout d'abord la dérivée de  $P_{i,j}(t)$  quand  $j \geq i$ .

$$P'_{i,j}(t) = \frac{\lambda^{j-i}}{(j-i)!} e^{-\lambda t} (-\lambda t^{j-i} + (j-i)t^{j-i-1})$$

Pour obtenir la matrice génératrice, il faut évaluer la dérivée en 0 :

- Dans le cas  $j \neq i + 1$  et  $j \neq i$

On a clairement :  $\mathcal{A}_{i,j} = 0$ .

- Dans le cas  $i + 1 = j$   
 $P'_{i,i+1}(t) = \lambda e^{-\lambda t} (-\lambda t + 1)$   
Donc  $\mathcal{A}_{i,i+1} = \lambda$ .
- Dans le cas  $i = j$   
 $P'_{i,i}(t) = e^{-\lambda t} (-\lambda)$   
Donc  $\mathcal{A}_{i,i} = -\lambda$ .

$\square$

**Conséquence 3.** Dans le cas du processus de Poisson d'intensité  $\lambda$ , on a pour  $f : \mathbb{E} \rightarrow \mathbb{R}^+$  :

$$\forall i \in \mathbb{E}, \quad \mathcal{A}f(i) = \lambda[f(i+1) - f(i)]$$

**Définition 15.** Etant donné un processus de Poisson  $(\mathcal{N}_t, t \geq 0)$ , une variable aléatoire  $S$  à valeurs dans  $\mathbb{R}^+ \cup \{+\infty\}$  est un  $\mathcal{F}_t^{\mathcal{N}}$  temps d'arrêt si :

$$\forall t \in \mathbb{R}^+, \quad \{S \leq t\} \in \mathcal{F}_t^{\mathcal{N}}$$

**Remarque 6.** Pour tout  $n$ , le temps de saut  $T_n$  est un  $(\mathcal{F}_t^{\mathcal{N}})$  temps d'arrêt. En effet,  $\{T_n \leq t\} = \{\mathcal{N}_t \geq n\}$ .

**Proposition 5.** Soit  $(\mathcal{N}_t, t \geq 0)$  un processus de Poisson de paramètre  $\lambda$  et  $S$  un temps d'arrêt pour  $(\mathcal{N}_t)_t$ . Sur l'événement  $\{S < +\infty\}$ , on pose pour  $t \geq 0$  :

$$\mathcal{N}_t^S = \mathcal{N}_{S+t} - \mathcal{N}_S$$

Conditionnellement à  $\{S < +\infty\}$ , le processus  $(\mathcal{N}_t^S, t \geq 0)$  est un processus de Poisson d'intensité  $\lambda$ , indépendant de la tribu engendrée par la trajectoire de  $\mathcal{N}$  jusqu'à  $S$ .

*Démonstration.*

- Dans le cas constant  $S = s$ , le résultat est vérifié par définition.
- Supposons  $S$  un temps d'arrêt non constant pour  $(\mathcal{N}_t)_t$ , et supposons que  $S$  prenne ses valeurs dans une suite croissante de réels positifs  $(s_j)_j$ .  
Soit  $\mathcal{F}_t = \sigma(\mathcal{N}_s, s \leq t)$  une tribu. Nous pouvons alors définir la tribu d'arrêt  $\mathcal{F}_S = \{A \in \mathcal{F}_\infty \mid A \cap \{S \leq t\} \in \mathcal{F}_t\}$  (où  $\mathcal{F}_\infty = \bigcup_{s \in \mathbb{N}} \mathcal{F}_s$ ).  
Nous avons donc :  $\mathcal{N}_t^s, \forall t > 0$  indépendants de  $\mathcal{F}_S$  et la loi de  $\mathcal{N}_t^s$  est la même que celle de  $\mathcal{N}_t$ .  
Soient  $A \in \mathcal{F}_S, 0 < t_1 < \dots < t_k \in \mathbb{R}, (s_i)_{i \in \mathbb{N}} \in \mathbb{R}$  et  $n_1, \dots, n_k \in \mathbb{N}$ .

$$\begin{aligned} & \mathbb{P}\left(A \bigcap_{i=1}^k \{\mathcal{N}_{t_i}^S - \mathcal{N}_{t_{i-1}}^S = n_i\}\right) \\ &= \sum_{j \in \mathbb{N}} \mathbb{P}(\{S = s_j\} \cap A \bigcap_{i=1}^k \{\mathcal{N}_{s_j+t_i} - \mathcal{N}_{s_j+t_{i-1}} + \mathcal{N}_{s_j} - \mathcal{N}_{s_j} = n_i\}) \\ &= \sum_{j \in \mathbb{N}} \mathbb{P}(\{S = s_j\} \cap A) \prod_{i=1}^k \mathbb{P}(\mathcal{N}_{s_j+t_i} - \mathcal{N}_{s_j+t_{i-1}} = n_i) \\ &= \mathbb{P}(A) \prod_{i=1}^k \mathbb{P}(\mathcal{N}_{t_i-t_{i-1}} = n_i). \end{aligned}$$

Nous retrouvons donc un processus de Poisson indépendant de la tribu d'arrêt. □

## Deuxième partie

### Application au modèle épidémiologique : le modèle SIR et indices de Sobol

# Chapitre 2

## Introduction au modèle SIR

Pour étudier l'évolution d'une maladie en épidémiologie, il peut être utile d'étudier l'évolution de sous-populations dans le temps. On utilise alors le modèle SIR.

### I. Le modèle SIR pour une population finie de taille $N$

#### 1. Modélisation

**Définition 16** (Modèle SIR). *Un modèle compartimental permet de diviser une population supposée constante en plusieurs catégories. Pour le modèle SIR, on a 3 catégories : les individus susceptibles, infectés et rémis.*

Pour étudier l'évolution des sous-populations au cours du temps, on définit une chaîne de Markov en temps continu à valeurs dans le sous-espace de  $\mathbb{E} = \{0, \dots, N\}^3$  tel que  $S_t + I_t + R_t = N$ ,  $\forall t \geq 0$ .

Au début du modèle, on a :

- $S_0 > 0$ , le nombre d'individus susceptibles
- $I_0 > 0$  le nombre d'individus infectés
- $R_0 = 0$  le nombre d'individus rémis.

Soit  $t \in \mathbb{R}^+$ . On note  $S_t, I_t, R_t$  l'évolution des sous-populations au cours du temps. On note  $\lambda$  le taux d'infection et  $\mu$  le taux de rémission.

Chaque couple "Infecté-susceptible" transmet la maladie au taux  $\lambda$ . Ainsi, le taux global d'infection est  $\lambda S_t I_t$  au temps  $t$ . La durée de transmission de la maladie sur une arête suit une exponentielle  $\mathcal{E}(\lambda)$ . (i.e : sur toutes les arêtes, la durée de transmission suit une exponentielle  $\mathcal{E}(\lambda S_t I_t)$ .)

$\mu$  est le taux de rémission individuel. Le taux global de rémission est  $\mu I_t$ . La durée de rémission sur une arête suit une loi exponentielle  $\mathcal{E}(\mu)$ . Sur toutes les arêtes, la durée de rémission suit une exponentielle  $\mathcal{E}(\mu I_t)$ . On a donc :

- $(S_t, I_t, R_t) \rightarrow (S_t - 1, I_t + 1, R_t)$  dans le cas d'une infection au taux  $\lambda S_t I_t$ .
- $(S_t, I_t, R_t) \rightarrow (S_t, I_t - 1, R_t + 1)$  dans le cas d'une rémission au taux  $\mu I_t$ .

Il y a donc deux transitions possibles du temps  $t$  au temps  $t + 1$ . Ces transitions sont représentées avec le schéma ci-dessous.



FIG. 2.1 : Représentation des transitions entre chaque état

On remarque que les individus susceptibles peuvent devenir infectés au taux  $\lambda S_t I_t$  et que les individus infectés peuvent être rémis avec un taux  $\mu I_t$ . Cependant un individu rémis ne peut plus être à nouveau infecté dans ce modèle.

**Lemme 1.** Soient  $X \sim \mathcal{E}(\lambda)$  et  $Y \sim \mathcal{E}(\mu)$  deux variables indépendantes. Alors on a :

- $\min(X, Y) \sim \mathcal{E}(\lambda + \mu)$ .
- $\mathbb{P}(X < Y) = \frac{\lambda}{\mu + \lambda}$

*Démonstration.* Soit  $X \sim \mathcal{E}(\lambda)$  et  $Y \sim \mathcal{E}(\mu)$  deux variables indépendantes. On a alors :

$$\min(X, Y) \sim \mathcal{E}(\lambda + \mu)$$

En effet :

$$\begin{aligned}
 \mathbb{P}(\min(X, Y) > t) &= \mathbb{P}(\{X > t\} \cap \{Y > t\}) \\
 &= \mathbb{P}(X > t)\mathbb{P}(Y > t) \\
 &= e^{-\lambda t} e^{-\mu t} \\
 &= e^{-(\lambda + \mu)t}.
 \end{aligned}$$

On cherche  $\mathbb{P}(X < Y)$  :

$$\begin{aligned}
\mathbb{P}(X < Y) &= \int_0^{+\infty} \int_0^y \lambda e^{-\lambda x} \mu e^{-\mu y} dx dy \\
&= \int_0^{+\infty} \mu e^{-\mu y} (1 - e^{-\lambda y}) dy \\
&= \int_0^{+\infty} \mu e^{-\mu y} - \mu e^{-(\mu+\lambda)y} dy \\
&= \left[ -e^{-\mu y} + \frac{\mu}{\mu + \lambda} e^{-(\lambda+\mu)y} \right]_0^{+\infty} \\
&= 1 - \frac{\mu}{\mu + \lambda} \\
&= \frac{\lambda}{\mu + \lambda}.
\end{aligned}$$

□

**Conséquence 4.** Soit  $t \geq 0$ . Dans notre cas, on note  $X = \text{''durée d'infection''} \sim \mathcal{E}(\lambda S_t I_t)$  et  $Y = \text{''durée de rémission''} \sim \mathcal{E}(\mu I_t)$ .

On sait aussi qu'entre deux sauts  $S_t$  et  $I_t$  sont constants.

On note  $\min(X, Y)$  le premier temps d'infection ou rémission. Ainsi,  $\min(X, Y) \sim \mathcal{E}(\lambda S_t I_t + \mu I_t)$ .

On a également :

$$\mathbb{P}(\text{''le 1<sup>er</sup> temps est une infection''}) = \frac{\lambda S_t I_t}{\mu I_t + \lambda S_t I_t}$$

On représente le modèle SIR par deux processus de Poisson :

- un processus de Poisson d'intensité  $\lambda N^2$  pour modéliser le passage de l'état susceptible à l'état infecté.
- un processus de Poisson d'intensité  $\mu N$  pour modéliser le passage de l'état infecté à l'état rémis.

On suppose donc que le temps de transition entre deux états suit une loi exponentielle de paramètre  $\lambda S_t I_t$  en ce qui concerne une infection ou de paramètre  $\mu I_t$  en ce qui concerne une rémission.

On peut alors construire la matrice génératrice du modèle. On la note  $\mathcal{A}$  et on la construit comme suit :

$$\forall f : \{0, \dots, N\}^3 \rightarrow \mathbb{R}^+, \quad \forall (s, i, r) \in \{0, \dots, N\}^3,$$

$$\mathcal{A}f(s, i, r) = \lambda s i [f(s-1, i+1, r) - f(s, i, r)] + \mu i [f(s, i-1, r+1) - f(s, i, r)]$$

**Rappel 1.**  $\frac{\partial}{\partial t} P(t) = \mathcal{A}P(t) = P(t)\mathcal{A}$

Soient  $f : \mathbb{E} \rightarrow \mathbb{R}$ ,  $i \in \mathbb{E}$ . Alors :

$$\frac{\partial}{\partial t} P_t f(i) = P_t \mathcal{A} f(i)$$

**Remarque 7.** Soit  $(X_t)_{t \geq 0}$  une chaîne de Markov en temps continu avec matrice génératrice  $\mathcal{A}$  et matrice de transition  $P_t, t \geq 0$ . On a :

$$\begin{aligned} \frac{\partial}{\partial t} (\mathbb{E}_i f(X_t)) &= P_t \mathcal{A} f(i) = \mathbb{E}_i (\mathcal{A} f(X_t)) \\ \mathbb{E}_i f(X_t) - f(i) &= \int_0^t \mathbb{E}_i (\mathcal{A} f(X_s)) ds. \end{aligned}$$

Appliqué à notre modèle, on a :

$$\begin{aligned} \mathbb{E}_{(S_0, I_0, 0)} (f(S_t, I_t, R_t)) - f(S_0, I_0, 0) &= \int_0^t \mathbb{E}_{(S_0, I_0, 0)} \left[ \lambda S_s I_s (f(S_s - 1, I_s + 1, R_s) - f(S_s, I_s, R_s)) \right. \\ &\quad \left. + \mu I_s (f(S_s, I_s - 1, R_s + 1) - f(S_s, I_s, R_s)) \right] ds. \end{aligned}$$

On peut étudier trois fonctions particulières :

- En posant  $f(s, i, r) = i$  :

$$\mathbb{E}(I_t) = I_0 + \mathbb{E} \left[ \int_0^t (\lambda S_s I_s - \mu I_s) ds \right]$$

- En posant  $f(s, i, r) = s$  :

$$\mathbb{E}(S_t) = S_0 + \mathbb{E} \left[ \int_0^t (-\lambda S_s I_s) ds \right]$$

- En posant  $f(s, i, r) = r$  :

$$\mathbb{E}(R_t) = \mathbb{E} \left[ \int_0^t (\mu I_s) ds \right]$$

## II. Etude du modèle SIR sur une très grande population (limite EDO)

### 1. Convergence de $S_t$ , $I_t$ et $R_t$

On considère  $N \rightarrow +\infty$  et on pose :

$$S_t^{(N)} = \frac{S_t}{N}, \quad I_t^{(N)} = \frac{I_t}{N}, \quad R_t^{(N)} = \frac{R_t}{N}, \quad \lambda = \frac{\lambda'}{N}$$

avec  $\lambda' > 0$ .

En reprenant les notations de la partie précédente :

$$\begin{aligned} & \mathbb{E} \left[ f(S_t^{(N)}, I_t^{(N)}, R_t^{(N)}) \right] - f(S_0^{(N)}, I_0^{(N)}, 0) \\ &= \int_0^t \mathbb{E} \left( \lambda S_s^{(N)} I_s^{(N)} N^2 \left[ f(S_s^{(N)} - \frac{1}{N}, I_s^{(N)} + \frac{1}{N}, R_s^{(N)}) - f(S_s^{(N)}, I_s^{(N)}, R_s^{(N)}) \right] \right. \\ & \quad \left. + \mu N I_s^{(N)} \left[ f(S_s, I_s - \frac{1}{N}, R_s + \frac{1}{N}) - f(S_s, I_s, R_s) \right] \right) ds. \end{aligned}$$

Ainsi, en reprenant  $f(s, i, r) = s$  :

$$\begin{aligned} \mathbb{E}(S_t^{(N)}) &= \frac{S_0}{N} - \mathbb{E} \left( \int_0^t (\lambda S_s^{(N)} I_s^{(N)} N) ds \right) \\ \mathbb{E}(S_t^{(N)2}) &= \frac{S_0^2}{N^2} - 2 \int_0^t \lambda' S_s^{(N)2} I_s^{(N)} ds + \frac{1}{N} \int_0^t \lambda' I_s^{(N)} S_s^{(N)} ds \end{aligned}$$

**Rappel 2.** Soit  $t \geq 0$  et  $(X_t)_{t \geq 0}$  une chaîne de Markov en temps continu. Alors pour toute fonction  $f$  :

$$\mathbb{E}(f(X_t)) = f(X_0) + \mathbb{E} \left( \int_0^t \mathcal{A}f(X_s) ds \right)$$

**Proposition 6.** Soit  $t \geq 0$ ,  $f$  une fonction quelconque et  $(X_t)_{t \geq 0}$  une chaîne de Markov en temps continu.

$$M_t^f := f(X_t) - f(X_0) - \int_0^t \mathcal{A}f(X_s) ds \tag{2.1}$$

est une martingale à temps continu.

**Conséquence 5.** Par définition de  $M_t^f$ , on a  $\mathbb{E}(M_t^f) = 0$ .



*Démonstration.* Montrons que  $\mathbb{E}(M_t^f | \mathcal{F}_s) = M_s^f$  pour  $t \geq s$  et  $\mathcal{F}_s = \sigma(X_u, u \leq s)$ .

$$\begin{aligned}
\mathbb{E}(M_t^f | \mathcal{F}_s) &= \mathbb{E}\left(f(X_t) - f(X_0) - \int_0^t \mathcal{A}f(X_u)du \mid \mathcal{F}_s\right) \\
&= \mathbb{E}(f(X_t)) - f(X_0) - \int_0^s \mathcal{A}f(X_u)du \\
&\quad - \int_s^t \mathbb{E}\left(\mathcal{A}f(X_u) \mid \mathcal{F}_s\right)du \\
&= f(X_0) + \int_0^t \mathbb{E}\left(\mathcal{A}f(X_u) \mid \mathcal{F}_s\right)du - f(X_0) \\
&\quad - \int_0^s \mathcal{A}f(X_u)du - \int_s^t \mathbb{E}\left(\mathcal{A}f(X_u) \mid \mathcal{F}_s\right)du \\
&= \int_0^s \mathbb{E}\left(\mathcal{A}f(X_u) \mid \mathcal{F}_s\right)du - \int_0^s \mathcal{A}f(X_u)du \\
&= \mathbb{E}(f(X_s) \mid \mathcal{F}_s) - f(X_0) - \int_0^s \mathcal{A}f(X_u)du \\
&= M_s^f.
\end{aligned}$$

Donc  $M_t^f$  est une martingale à temps continu. □

- Pour  $f(x) = x$ , on a  $M_t^x = X_t - X_0 - \int_0^t \mathcal{A}Id(X_s)ds$ .
- Pour  $f(x) = x^2$ , on a  $M_t^{x^2} = X_t^2 - X_0^2 - \int_0^t \mathcal{A}f(X_s)ds$ .

D'autre part, par la formule de Itô, on a :

$$X_t^2 = X_0^2 + \int_0^t 2X_s \mathcal{A}Id(X_s)ds + \frac{1}{2} \langle M^x \rangle_t + N_t^x \quad (2.2)$$

où  $N_t^x$  est une martingale et  $\langle M^x \rangle$  est la variation quadratique de  $\langle N^x \rangle$  telle que  $\mathbb{E}(\langle M^x \rangle_t) = Var(M_t^x)$

En réappliquant à notre modèle, on obtient :

- Pour  $f(s, i, r) = s$ , on a  $\mathcal{A}S^{(N)} = -\lambda' S^{(N)} I^{(N)}$ .
- Pour  $f(s, i, r) = s^2$ , on a  $\mathcal{A}S^{(N)2} = \lambda' S^{(N)} I^{(N)} \left(\frac{1}{N} - 2S^{(N)}\right)$

Ainsi, par (2.1) et (2.2), on a deux expressions pour  $(S_t^{(N)})^2$  :

$$\begin{aligned} S_t^{(N)2} &= S_0^{(N)2} - \int_0^t 2\lambda' S_s^{(N)2} I_s^{(N)} ds + \frac{\langle M^x \rangle_t}{2} + N_t^x \\ &= S_0^{(N)2} - \int_0^t 2\lambda' S_s^{(N)2} I_s^{(N)} ds + \frac{1}{N} \int_0^t \lambda' S_s^{(N)} I_s^{(N)} ds + P_t^x. \end{aligned}$$

où  $P_t^x$  et  $N_t^x$  sont des martingales d'espérance nulle.

On en déduit :

$$\frac{1}{2} \mathbb{E}(\langle M^x \rangle_t) = \frac{1}{N} \mathbb{E} \left( \int_0^t \lambda' S_s^{(N)} I_s^{(N)} ds \right)$$

Donc  $Var(M_t^x) = \mathcal{O}(\frac{1}{N})$ .

Rappelons que  $S_t^{(N)} = S_0^{(N)} - \int_0^t \lambda' S_s^{(N)} I_s^{(N)} ds + M_t^x$ . Ainsi, si  $S_0^{(N)} \rightarrow s_0$  quand  $N \rightarrow +\infty$ , on a convergence simple en t :

$$S_t^{(N)} \rightarrow s_t = s_0 - \int_0^t \lambda' s_u i_u du \quad (2.3)$$

De la même manière,

$$I_t^{(N)} \rightarrow i_t = i_0 + \int_0^t (\lambda' s_u i_u - \mu i_u) du$$

$$R_t^{(N)} \rightarrow r_t = \int_0^t \mu i_u du$$

On peut également montrer la convergence uniforme de  $S_t$  vers  $s_t$ . En effet, on a :

$$\begin{aligned} |S_t^{(N)} - s_t| &= |S_0^{(N)} - \int_0^t \lambda' S_s^{(N)} I_s^{(N)} ds + M_t^x - s_0 + \int_0^t \lambda' s_u i_u du| \\ &\leq |S_0^{(N)} - s_0| + \int_0^t \lambda' |S_u^{(N)} I_u^{(N)} - s_u i_u| du + |M_t^x| \\ &\leq |S_0^{(N)} - s_0| + \int_0^t \lambda' |S_u^{(N)} - s_u| I_u^{(N)} du + \int_0^t \lambda' |I_u^{(N)} - i_u| s_u du + |M_t^x|. \end{aligned}$$

On sait que  $|S_0^{(N)} - s_0| \xrightarrow{N \rightarrow +\infty} 0$  et que  $I_u^{(N)} \leq 1$  et  $s_u \leq 1$ . On a alors :

$$|S_t^{(N)} - s_t| \leq |S_0^{(N)} - s_0| + \int_0^t \lambda' |S_u^{(N)} - s_u| du + \int_0^t \lambda' |I_u^{(N)} - i_u| du + |M_t^x|.$$

De plus, on sait par l'inégalité de Doob que :

$$\mathbb{E} \left( \sup_{t \leq T} |M_t^x| \right) \leq C \mathbb{E} \left( (M_t^x)^2 \right).$$

Ainsi :

$$\sup_{t \leq T} |S_t^{(N)} - s_t| \leq |S_0^{(N)} - s_0| + \int_0^t \lambda' \sup_{v \leq u} |S_v^{(N)} - s_v| du + \int_0^t \lambda' \sup_{v \leq u} |I_v^{(N)} - i_v| du + \sup_{t \leq T} |M_t^x|.$$

En passant à l'espérance, on obtient donc :

$$\begin{aligned} \mathbb{E} \left( \sup_{t \leq T} |S_t^{(N)} - s_t| \right) &\leq \mathbb{E} (|S_0^{(N)} - s_0|) + \int_0^t \lambda' \mathbb{E} \left( \sup_{v \leq u} |S_v^{(N)} - s_v| \right) du \\ &\quad + \int_0^t \lambda' \mathbb{E} \left( \sup_{v \leq u} |I_v^{(N)} - i_v| \right) du + \mathbb{E} \left( \sup_{t \leq T} |M_t^x| \right) \\ &\leq C_1 + \int_0^t \lambda' \mathbb{E} \left[ \sup_{v \leq u} (|S_v^{(N)} - s_v| + |I_v^{(N)} - i_v|) \right] du + C_2 \mathbb{E}(M_t^{x2}). \end{aligned}$$

Or, on a vu précédemment que  $\text{Var}(M_t^x) = \mathbb{E}(M_t^{x2}) = \mathcal{O}(\frac{1}{N}) \xrightarrow{N \rightarrow +\infty} 0$ .

On obtient alors :

$$\mathbb{E} \left( \sup_{t \leq T} |S_t^{(N)} - s_t| \right) \leq C(N, T) + \int_0^t \lambda' \mathbb{E} \left( \sup_{v \leq u} |I_v^{(N)} - i_v| + |S_v^{(N)} - s_v| \right) du.$$

De même, pour  $I_t$ , on obtient :

$$\begin{aligned} \mathbb{E} \left( \sup_{t \leq T} |I_t^{(N)} - i_t| \right) &\leq \tilde{C}(N, T) + \int_0^t \lambda' \mathbb{E} \left( \sup_{v \leq u} |I_v^{(N)} - i_v| + |S_v^{(N)} - s_v| \right) du \\ &\quad + \int_0^t \mu \mathbb{E} \left( \sup_{v \leq u} |I_s^{(N)} - i_s| \right) du. \end{aligned}$$

**Rappel 3** (Inégalité de Gronwall).

$$f(t) \leq c + \int_0^t c' f(s) ds \Rightarrow f(t) \leq ce^{c't}$$

**Rappel 4** (Inégalité de Gronwall). Soit  $f$  une fonction à valeurs positives, si on a :

$$f(t) \leq c + \int_0^t c' f(s) ds$$

où  $c$  et  $c'$  sont des constantes positives, alors :

$$f(t) \leq ce^{c't}$$

On a :

$$\begin{aligned} \mathbb{E} \left[ \sup_{t \leq T} \left( |S_t^{(N)} - s_t| + |I_t^{(N)} - i_t| \right) \right] &\leq C(N, T) + \tilde{C}(N, T) \\ &\quad + \int_0^t (\lambda' + \mu) \mathbb{E} \left[ \sup_{v \leq u} \left( |S_v^{(N)} - s_v| + |I_v^{(N)} - i_v| \right) \right] du. \end{aligned}$$

On peut appliquer l'inégalité de Gronwall.

$$\mathbb{E} \left[ \sup_{t \leq T} \left( |S_t^{(N)} - s_t| + |I_t^{(N)} - i_t| \right) \right] \leq \left( C(N, T) + \tilde{C}(N, T) \right) e^{T(\lambda + \mu)}.$$

Or on sait que  $C(N, T) + \tilde{C}(N, T) \xrightarrow{N \rightarrow +\infty} 0$ .

On a donc convergence uniforme de  $S_t^{(N)}$  vers  $s_t$  et  $I_t^{(N)}$  vers  $i_t$ , pour tout  $t \geq 0$ .

## 2. Résolution par système d'équations différentielles

**Remarque 8.**  $s_t$  définie en (2.3) est de classe  $\mathcal{C}^\infty$ .

Grâce à la convergence uniforme de  $S_t^{(N)}$ ,  $I_t^{(N)}$  et  $R_t^{(N)}$ , on obtient ce système d'équations différentielles :

$$\begin{cases} s_t' = -\lambda s_t i_t \\ i_t' = \lambda s_t i_t - \mu i_t \\ r_t' = \mu i_t \end{cases} \quad (2.4)$$

On sait, d'après Cauchy-Lipschitz, qu'il existe une unique solution. En effet,  $(s_t, i_t) \rightarrow \lambda s_t i_t$  est lipschitzienne. De même,  $i_t \rightarrow \mu i_t$  est également lipschitzienne.

**Remarque 9.** Les valeurs de  $s, i$  et de  $r$  sont comprises entre 0 et 1. De plus, la croissance de  $i$  est inférieure ou égale à la décroissance de  $s$ . En effet, si le contraire se produit,  $s$  va tendre trop rapidement vers 0 et  $i$  va tendre trop rapidement vers 1, et on ne peut donc rien observer.

Grâce au logiciel R, on peut représenter la solution du système d'équations différentielles au cours du temps. On pose des solutions initiales :

- $S_0 = 500$
- $I_0 = 50$
- $R_0 = 0$

De plus, on pose le taux d'infection  $\lambda = \frac{0.2}{S_0 + I_0}$  et le taux de rémission  $\mu = 0.01$ .

On calcule les solutions  $S_t$ ,  $I_t$  et  $R_t$  par les formules page 22. Ensuite, on calcule la solution du système d'équations différentielles à partir des taux  $\lambda$  et  $\mu$ .

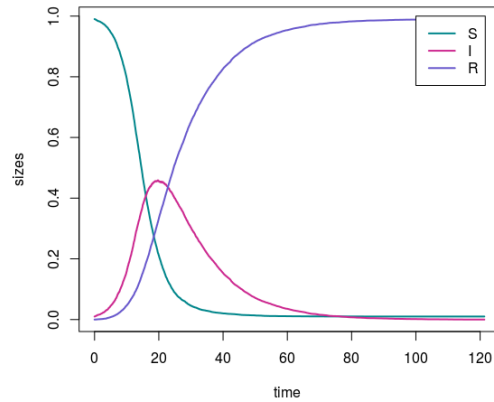


FIG. 2.2 : Représentation de l'évolution de chaque sous-population

On remarque sur le graphique que le nombre d'individus susceptible diminue au cours du temps. Dans un premier temps, le nombre d'individus infectés augmente rapidement et le nombre d'individus rémis croît lentement. Dans un second temps, le nombre d'individus infectés décroît aussi rapidement que la croissance du nombre d'individus rémis.

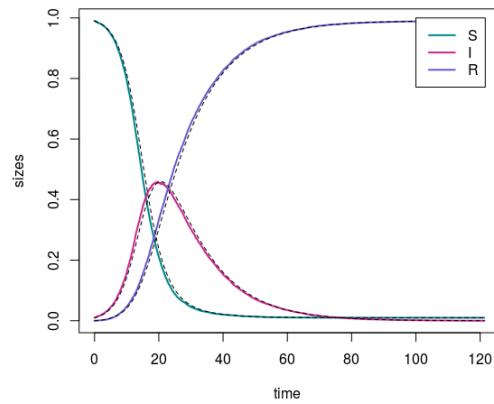


FIG. 2.3 : Représentation de la solution du système d'équations différentielles

On remarque que la solution du système d'équations différentielles approche très bien l'évolution de la maladie sur la population.

### III. Algorithmes de simulation

Pour simuler le modèle SIR sur R, il faut tout d'abord construire l'échantillon de  $\lambda$  et  $\mu$  :

```
echant<-function(n,a,b){
  lambda=runif(n,0,a)
  mu=runif(n,0,b)
  A=matrix(0,nrow=n,ncol=2)
  A[,1]=lambda
  A[,2]=mu
  return(A)
}
```

Cette fonction prend en paramètres un entier  $n$  et deux réels  $a$  et  $b$ . Elle renvoie une matrice de taille  $n \times 2$ . La première colonne (l'échantillon de  $\lambda$ ) est un échantillon de taille  $n$  d'une loi *Unif* ( $[0, a]$ ). La seconde colonne (l'échantillon de  $\mu$ ) est un échantillon de taille  $n$  d'une loi *Unif* ( $[0, b]$ ).

```
simul <- fonction(n,S0, I0, T,a,b){
  N<- I0+S0
  A <- echant(n,a,b)
  B=matrix(0,nrow=n,ncol=3)

  for (i in 1 :n){
    dates<- array(data=c(0),dim=c(T))
    I<- array(data=c(0),dim=c(T))
    S<- array(data=c(0),dim=c(T))
    R<- array(data=c(0),dim=c(T))
    I[1]=I0
    S[1]=S0

    t<-2
    while((t<T) && (I[t-1]>0))
      {
        dates[t]<-dates[t-1]+rexp(1,A[i,1]*S[t-1]*I[t-1]+A[i,2]*I[t-1])
        u<-rbinom(1,1, A[i,1]*S[t-1]*I[t-1]/(A[i,1]*S[t-1]*I[t-1]+A[i,2]*I[t-1]))
        if(u==1)
          {
            I[t]<-I[t-1]+1
            S[t]<-S[t-1]-1
            R[t]<-R[t-1]
          }
        else
```

```

    {
      S[t]<-S[t-1]
      I[t]<-I[t-1]-1
      R[t]<-R[t-1]+1
    }
    t<-t+1
  }

pas<-0.1
datesedp<-seq(0,dates[t-1],pas)
Sedp<-array(data=c(0),dim=c(dates[t-1]/pas+1))
Iedp<-array(data=c(0),dim=c(dates[t-1]/pas+1))
Redp<-array(data=c(0),dim=c(dates[t-1]/pas+1))

Sedp[1]=S0/N
Iedp[1]=I0/N
lambda2<- A[i,1]*(S0+I0)

for(tt in 2 :(dates[t-1]/pas+1))
{
  Sedp[tt]=Sedp[tt-1]-lambda2*Sedp[tt-1]*Iedp[tt-1]*pas
  Iedp[tt]=Iedp[tt-1]+lambda2*Sedp[tt-1]*Iedp[tt-1]*pas-A[i,2]*Iedp[tt-1]*pas
  Redp[tt]=Redp[tt-1]+A[i,2]*Iedp[tt-1]*pas
}
B[i,1]=Sedp[(dates[t-1]/pas+1)]
B[i,2]=Iedp[(dates[t-1]/pas+1)]
B[i,3]=Redp[(dates[t-1]/pas+1)]
}
return(B)
}

```

La fonction simul prend en paramètres le nombre de simulations  $n$ , le nombre d'individus susceptibles à l'état initial  $S_0$ , le nombre d'individus infectés à l'état initial  $I_0$ , le temps total pour la simulation  $T$ , le taux de  $\lambda$  a et le taux de  $\mu$  b.

Elle renvoie une matrice taille  $n \times 3$ . La première colonne représente pour chaque simulation le pourcentage d'individus susceptibles au temps T. La deuxième colonne représente le pourcentage d'individus infectés au temps T. La dernière représente le pourcentage d'individus rémis au temps T.

La population totale est représentée par  $N = I_0 + S_0$ . On simule un échantillon de  $\lambda$  et  $\mu$  avec la fonction echant.

On construit tout d'abord une simulation aléatoire dépendant d'une variable aléatoire  $u \sim \text{Bernoulli}(\mathcal{B}(1, \frac{\lambda SI}{\lambda SI + \mu I}))$ . Si  $u = 1$ , il y a infection. Sinon, il y a rémission. Ensuite, on construit la solution du système différentiel (2.4). On prend les conditions initiales  $S_{edp} = \frac{S_0}{N}$  et  $I_{edp} = \frac{I_0}{N}$  et on pose  $\lambda_{edp} = \frac{\lambda}{N}$ . On résout le système par différences finies. On range le dernier terme de  $S_{edp}$ ,  $I_{edp}$  et  $R_{edp}$  dans une matrice de taille  $n \times 3$ . On obtient ainsi la simulation du modèle SIR pour  $n$  observations.

```

    result<-function(n,S0, I0, T,a,b){
  Y <- 1 :n
  B<-simul(n,S0,I0,T,a,b)
  for(i in 1 :n){
    Y[i]<-B[i,2]+B[i,3]
  }
  return(Y)
}

```

La fonction `result` prend en paramètres le nombre de simulations  $n$ , le nombre d'individus susceptibles à l'instant initial  $S_0$ , le nombre d'individus infectés à l'instant initial  $I_0$ , le temps total pour la simulation  $T$  et les taux  $a$  et  $b$  pour  $\lambda$  et  $\mu$ .

Elle renvoie un vecteur  $Y$  de taille  $n$  correspondant au nombre d'individus tombés malade au cours de la simulation (on compte les individus infectés et rémis).



# Chapitre 3

## Analyse de sensibilité : indices de Sobol

On va maintenant étudier si le modèle précédent est bien fidèle à la réalité. Pour cela, on va étudier l'analyse de sensibilité du modèle par les indices de Sobol.

### 3.1 Généralités

Pour définir les indices de Sobol, on va d'abord se restreindre à un échantillon  $X = (X_1, \dots, X_n) \in \mathbb{R}^n$  indépendants. Soit  $f$  une fonction analytique  $\mathbb{R}^n \rightarrow \mathbb{R}$ . On note la variable  $Y = f(X) \in \mathbb{R}$ .

On veut étudier l'importance de la variable  $X_i$  sur la variance de  $Y$ . Plus la variable  $X_i$  sera importante, plus  $Var(\mathbb{E}(Y | X_i))$  sera grande. On prend alors sa forme normalisée pour étudier la sensibilité de la variable  $X_i$  sur  $Y$ .

**Définition 17** (Indice de sensibilité de premier ordre). *On note  $S_i$  la part de variance de  $Y$  due à la variable  $X_i$  :*

$$S_i = \frac{Var(\mathbb{E}(Y | X_i))}{Var(Y)}$$

On peut également étudier la sensibilité de plusieurs variables sur la variance  $Y$ . Pour cela, on utilise la décomposition de la fonction  $f$ .

$$\begin{aligned} Y &= f(X_1, \dots, X_n) \\ &= f_0 + \sum_{i=1}^n f_i(X_i) + \sum_{1 \leq i < j \leq n} f_{ij}(X_i, X_j) + \dots + f_{1, \dots, n}(X_1, \dots, X_n) \end{aligned}$$

où

$$\begin{aligned}
 f_0 &= \mathbb{E}(Y) \\
 f_i(X_i) &= \mathbb{E}(Y \mid X_i) - \mathbb{E}(Y) \\
 f_{i,j}(X_i, X_j) &= \mathbb{E}(Y \mid X_i, X_j) - \mathbb{E}(Y \mid X_i) - \mathbb{E}(Y \mid X_j) + \mathbb{E}(Y) \\
 &\dots
 \end{aligned}$$

Pour obtenir des indices de sensibilité pour plusieurs variables (d'ordre supérieur à 1), on utilise la décomposition de la variance.

**Théorème 1** (Décomposition de la variance).

$$V = \sum_{i=1}^n V_i + \sum_{1 \leq i < j \leq n} V_{i,j} + \dots + V_{1,\dots,n}$$

où

$$\left\{ \begin{array}{l}
 V_i = \text{Var}(\mathbb{E}(Y \mid X_i)) \\
 V_{i,j} = \text{Var}(\mathbb{E}(Y \mid X_i, X_j)) - V_i - V_j \\
 \dots \\
 V_{1,\dots,n} = V - \sum_{i=1}^n V_i - \sum_{1 \leq i < j \leq n} V_{i,j} - \dots - \sum_{1 \leq i_1 < i_2 < \dots < i_{n-1} \leq n} V_{i_1,\dots,i_{n-1}}
 \end{array} \right.$$

*Démonstration.* En remplaçant les  $V_i, V_{i,j}, \dots, V_{1,\dots,n}$  dans la formule, par simplification, on retrouve  $V$ .  $\square$

A partir de cette décomposition, on peut définir les indices de sensibilité d'ordre supérieur à 1.

En reprenant les notations du théorème, on note l'indice de sensibilité d'ordre 2 comme :

$$\begin{aligned}
 S_{i,j} &= \frac{V_{i,j}}{V} \\
 &= \frac{\text{Var}(\mathbb{E}(Y \mid X_i, X_j)) - \text{Var}(\mathbb{E}(Y \mid X_i)) - \text{Var}(\mathbb{E}(Y \mid X_j))}{V}.
 \end{aligned}$$

et un indice de sensibilité d'ordre k comme :

$$S_{1,\dots,k} = \frac{V_{1,\dots,k}}{V}.$$

**Conséquence 6.** Par le théorème précédent, on a la somme des indices de sensibilité égale à 1.

On peut facilement interpréter ces indices de sensibilité. En effet, plus l'indice  $S_{1,\dots,k}$  sera proche de 1, plus les variables  $X_1, \dots, X_k$  seront importantes.

On peut également étudier l'indice de sensibilité total à la variable  $X_i$ . Celui-ci représente la somme des indices de sensibilité faisant intervenir la variable  $X_i$ .

On le note :

$$S_{T_i} = \sum_{k \neq i} S_k$$

où  $\#i$  représente tous les ensembles d'indices contenant l'indice  $i$ .

## 3.2 Application au modèle SIR

Appliqué au modèle SIR, on a  $X = (X_1, X_2) = (\lambda, \mu)$  indépendants. Pour le calcul des indices de Sobol, nous allons utiliser l'estimateur de Nadaraya-Watson.

**Définition 18** (Estimateur de Nadaraya-Watson). *Soient  $(X_1, Y_1), \dots, (X_n, Y_n)$  un échantillon de même loi que  $(X, Y)$ .*

*Soient  $K : \mathbb{R} \rightarrow \mathbb{R}^+$  tel que  $\int_{\mathbb{R}} K(u)du = 1$  et  $h$  un pas. On estime  $f(x) = \mathbb{E}(Y | X = x)$  par :*

$$\hat{f}_n(x) = \frac{\sum_{i=1}^n Y_i K\left(\frac{X_i - x}{h}\right)}{\sum_{i=1}^n K\left(\frac{X_i - x}{h}\right)}$$

Dans notre cas, nous avons choisi  $Y$  comme défini dans la fonction result et  $K(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ .

Nous avons simulé les estimateurs de Nadaraya-Watson pour une et deux variables :

```

est <- fonction(x,y,h){
n<-length(x)
D<-as.matrix(dist(x))
K<-dnorm(D,0,h)
E<-1 :n
for(i in 1 :n){
  E[i]<-sum(y*K[,i])/sum(K[,i])
}
return(E)
}

est2<-fonction(x1,x2,y,h1,h2){
n <- length(x1)
D1<-as.matrix(dist(x1))
D2<-as.matrix(dist(x2))
K1<-dnorm(D1,0,h1)
K2<-dnorm(D2,0,h2)
E<-1 :n
for(i in 1 :n){

```

```

    E[i]<- sum(y*K1[,i]*K2[,i])/sum(K1[,i]*K2[,i])
  }
  return(E)
}

```

Ainsi, on peut définir les indices de Sobol pour le modèle :

```

  sobol<-function(n,x,y,h){
E<-est(x,y,h)
S<-var(E)/var(y)
return(S)
}

  sobol2<-function(n,x1,x2,y,h1,h2){
E<-est2(x1, x2,y,h1, h2)
E1<-est(x1,y,h1)
E2<-est(x2,y,h2)
S<-(var(E)-var(E1)-var(E2))/var(y)
return(S)
}

```

On remarque que l'indice de Sobol dominant parmi les trois calculés est l'indice correspondant à  $\mu$ . Ainsi,  $\mu$  est la variable qui influe le plus sur la variance de  $Y$ .

**Remarque 10.** *On peut choisir une autre loi pour  $\lambda$  et  $\mu$ . Dans notre cas, nous avons choisi une loi uniforme pour  $\lambda$  et  $\mu$ .*

# Conclusion

Au cours de ce travail de recherche, on a étudié le modèle SIR et l'influence des taux d'infection et de rémission sur l'évolution de la maladie. On a remarqué que selon le choix des  $\lambda$  et  $\mu$ , nos résultats de simulation étaient différents. En effet, en prenant pour  $\lambda$  une loi uniforme  $Unif([0, a])$  et  $\mu$  une loi  $Unif([0, b])$  où  $a \ll b$  (par exemple :  $a = 0.000000001$  et  $b = 0.001$ ), on obtient peu d'infections. A l'inverse, en prenant  $a \gg b$  (par exemple : ( $a = 0.01$  et  $b = 0.000000001$ ), on n'a plus d'individus susceptibles.

Cependant, ce modèle n'est pas le plus complet. Nous aurions pu ajouter d'autres paramètres comme le taux de décès, de naissances, de vaccinations... Le modèle aurait alors été plus réaliste : il y aurait une variation de la taille de population au cours du temps et moins d'infections.

Il existe d'autres modèles connus en épidémiologie. Le modèle SIS (susceptible, infecté, susceptible) permet de modéliser l'évolution d'une maladie dans le cas où chaque individu peut tomber malade plusieurs fois. Le modèle SEIR (susceptible, exposé, infecté, rémis) est un peu plus précis que le modèle SIR. Il permet de distinguer le moment où un individu passe de l'état sain au début d'infection...

D'autres modèles peuvent être créés ou complétés : on peut ajouter d'autres compartiments de populations (D pour décédé, Q pour quarantaine, M pour immunisé à la naissance). Selon les modèles, plus le nombre de compartiments est élevé, plus la modélisation est proche de la réalité et complexe à réaliser.

La santé publique est importante dans notre société. Il est donc important de prévoir l'évolution des virus tel que la grippe H5N1 (grippe aviaire). Le modèle SIR peut suffire dans la prédiction ; le gouvernement pourra alors adapter ses moyens pour lutter contre la maladie.

# Bibliographie

1. LALEUF Jean-Claude. *Processus et intégrales stochastiques, cours et exercices*, Ellipses, collection références sciences, 2014, 470 pages.
2. LESSARD Sabin. *Processus stochastiques, cours et exercices corrigés*, Ellipses, collection références sciences, 2014, 261 pages.
3. CASTELLAN Gwenaëlle, COUSIEN Anthony, TRAN Viet Chi. *Non-parametric adaptive estimation of order 1 Sobol indices in stochastic models, with an application to Epidemiology*, 2015, <hal-01249333v2>
4. MELEARD Sylvie. *Modèles aléatoires en Ecologie et Evolution, Processus de population en temps continu*, Springer, 2016, 267 pages.
5. JACQUES Julien. *Pratique de l'analyse de sensibilité : comment évaluer l'impact des entrées aléatoires sur la sortie d'un modèle mathématique*, Pub. IRMA Lille Vol. 71-III, 2011.