

Rapport de TIPE

Modélisation des épidémies

El Azhar Mohammed,identifiant:33855 ,
2020-2021,Enjeux sociétaux

8 juin 2021

1 Introduction à l'épidémiologie mathématique :

L'épidémiologie mathématique est une approche visant à prévoir le comportement d'une épidémie au sein d'une population compte tenu de certains paramètres propre à la population, elle fournit un cadre de raisonnement adapté, manipulable et réfutable pour l'étude de la propagation des maladies infectieuses, là où l'intrication des différents processus (physiopathologie, transmission, politiques sanitaires) est un obstacle à l'intuition et à l'argumentation qualitative. En outre, elle permet d'apporter des éléments de réponse à des problèmes pour lesquels les approches empiriques courantes sont limitées (par le temps, l'espace, l'éthique), sinon impossibles. De plus, elle permet de conclure quant à l'efficacité de certaines solutions;(quarantaine, vaccinations..)

L'épidémiologie mathématique aurait vu le jour avec John Graunt (1620-1674) qui fut le premier à recenser des données de maladies et les analyser. Pendant le 18ème siècle durant la variole, Daniel Bernoulli élaborait ce que l'on penserait être le premier modèle d'épidémiologie mathématique. Il fut suivi par les travaux de John Snow qui étudia la mortalité par Cholera dans la ville de Londres entre 1849 et 1854. Puis au 20ème siècle A. G. Mckendrick et W. O. Kermack proposèrent un modèle fondamental dans la théorie d'épidémiologie mathématique qui sera ensuite inscrit dans ce que l'on appelle aujourd'hui les modèles compartimentaux.

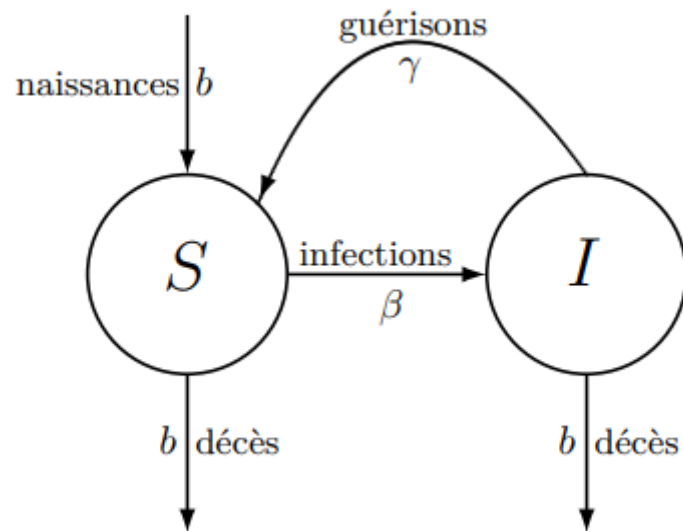
Les principales notations adoptées, remarques et définitions nécessaires :

- $N=S+I$: la taille de la population (considérée constante)
- S : les individus susceptibles d'être infectés
- I : les individus infectieux

- $b > 0$: le taux de naissance=taux de létalité
- $\gamma > 0$: le taux de guérison.
- $\beta > 0$: le taux de contact.

2 Modèle SIS déterministe :

le modèle est décrit par le schéma suivant :



On en déduit le système d'équations différentielles suivant :

$$\begin{cases} S' = -\frac{\beta}{N}SI + (b + \gamma)I \\ I' = \frac{\beta}{N}SI - (b + \gamma)I \end{cases}$$

Comme $N=S+I$, il suffit d'étudier une seule variable : I le nombre d'infectieux par exemple.

On suppose $I(t)$ initialisé à I_0 et on note : $c=\beta - (b + \gamma)$

On en obtient ainsi cette équation différentielle :

$$I'(t) = cI(t) - \frac{\beta}{N}I^2$$

2.1 Solution de l'équation :

L'ensemble de solution de l'équation différentielle est donné par :

1. Si $I_0 = \frac{c}{\beta}N$ alors $\forall t \geq 0, I(t) = I_0$
2. Si $c = 0$, alors $\forall t \geq 0, I(t) = \frac{I_0}{1 + \beta \frac{I_0}{N}t}$
3. Sinon : alors $\forall t \geq 0, I(t) = \frac{\beta \frac{c}{N} - \frac{c}{I_0}}{\frac{\beta}{N} - (\frac{\beta}{N} - \frac{c}{I_0})e^{-ct}}$

2.2 Preuve :

On traite les 3 cas :

1. La solution constante $I(t) = \frac{c}{\beta}N$ est solution du problème de cauchy. On conclut par unicité
2. Si $c=0$, l'équation devient $I'(t) = -\frac{\beta}{N}I^2$. Donc $\frac{d}{dt}(\frac{1}{I(t)}) = \frac{\beta}{N}$. Ensuite on effectue le changement de variable $u(t) = \frac{1}{I(t)}$ pour déduire que $I(t) = \frac{I_0}{1 + \beta \frac{I_0}{N}t}$
3. En divisant les termes de l'équation par $I(t)(c - \frac{\beta}{N}I(t))$ on obtient :

$$1 = \frac{I'(t)}{I(t)(c - \frac{\beta}{N}I(t))} = \frac{I'(t)}{cI(t)} - \frac{\frac{\beta}{N}I'(t)}{c(\frac{\beta}{N}I(t) - c)}$$

Donc : $\ln(\frac{I(t)}{I_0}) - \ln(\frac{\frac{\beta}{N}I(t) - c}{\frac{\beta}{N}I_0(t) - c}) = ct$ On en déduit enfin , par des opérations

élémentaires en introduisant l'exponentielle que : $I(t) = \frac{\beta \frac{c}{N} - \frac{c}{I_0}}{\frac{\beta}{N} - (\frac{\beta}{N} - \frac{c}{I_0})e^{-ct}}$

2.3 Conclusion : Introduction du Paramètre R_0

On entend souvent parler de ce fameux paramètre R_0 , ici on présente sa signification en particulier pour le modèle SIS. Revenons à l'équation :

$$I'|_{t=0} = \frac{I_0}{b + \gamma}(\frac{\beta S_0}{N(b + \gamma)} - 1)$$

À $t=0$: $S_0 \sim N$ donc l'équation devient :

$$I'|_{t=0} = \frac{I_0}{b + \gamma}(\frac{\beta}{(b + \gamma)} - 1)$$

On note $R_0 = \frac{\beta}{(b + \gamma)}$. Deux cas se présentent :

Si $R_0 > 1$, alors $I'|_{t=0} > 0$. Le nombre d'infectieux croît au début.

Si $R_0 < 1$, alors $I'|_{t=0} < 0$. Le nombre d'infectieux décroît

Ainsi le contrôle d'une épidémie est lié directement au contrôle du paramètre R_0 . D'où l'intérêt de la distanciation sociale(contrôle de β), les paramètre b et γ , eux ne peuvent pas être contrôlés directement

3 Modèle SIS stochastique :

Une version stochastique du modèle SIS précédemment présenté serait utile pour avoir une idée sur le temps d'extinction d'une épidémie. L'on serait amené à calculer des espérances de variables aléatoires.

3.1 Formalisation : chaîne de Markov discrète dans le temps

Les principales notations adoptées, remarques et définitions nécessaires :

- $S(t)$ et $I(t)$ deux variables aléatoires discrètes
- $t \in \{0, \Delta t, 2\Delta t, \dots\}$
- $S(t), I(t) \in \{0, 1, 2, \dots, N\}$
- $I_n = I(n\Delta t)$
- $p_{ji} = \mathbb{P}(I_{n+1} = j | I_n = i)$
- *Une chaîne de Markov* est un processus stochastique à temps discret vérifiant la propriété de Markov : le présent ne dépend que de l'instant qui le précède. c'est à dire :

$$\mathbb{P}(I(t) = i | I(0), \dots, I(t - \Delta t)) = \mathbb{P}(I(t) = i | I(t - \Delta t))$$

Hypothèses du modèle :

- On prend Δt suffisamment petit pour supposer que les seules transitions possibles entre t et $t + \Delta t$ soient (si $I(t) = i$) :

$$i \rightarrow i+1 \text{ ou } i \rightarrow i \text{ ou } i \rightarrow i-1$$

Compte tenu de la définition du modèle SIS , sa version stochastique serait résumée de la manière suivante :

$$p_{ji} = \begin{cases} \frac{\beta i(N-i)\Delta t}{N} & \text{si } j = i+1 \\ (b+\gamma)i\Delta t & \text{si } j = i-1 \\ 1 - (\frac{\beta i(N-i)\Delta t}{N} + (b+\gamma)i\Delta t) & \text{si } j = i \\ 0 & \text{si } j \notin \{i+1, i, i-1\} \end{cases}$$

On allège les notations :

$$p_{ji} = \begin{cases} b_i & \text{si } j = i+1 \\ d_i & \text{si } j = i-1 \\ 1 - (b_i + d_i) & \text{si } j = i \\ 0 & \text{si } j \notin \{i+1, i, i-1\} \end{cases}$$

4 Problématique : Comment estimer la moyenne du temps d'absorption d'une épidémie ?

4.1 Première approche ; Méthode simplifiée :

Notations et hypothèses du modèle :

- $I_1 = 1$
- $b_i = b \forall i \in \mathbb{N}$, (b est une constante à ne pas confondre avec le taux de létalité)
- $d_i = d \forall i \in \mathbb{N}$
- $b+d=1$
- Soit par ailleurs la variable aléatoire $(X_n)_{n \in \mathbb{N}}$ définie par :

$$X_i = \begin{cases} X_1 = I_1 & \text{si } i=1 \\ X_i = I_i - I_{i-1} & \forall i, i \geq 2 \end{cases}$$

On a donc $\mathbb{P}(X_i = 1) = b$ et $\mathbb{P}(X_i = -1) = d$ et $I_n = \sum_{i=1}^n X_i$

Notre processus stochastique est donc une marche aléatoire dans \mathbb{N} . Notre objectif est de calculer le temps de retour en 0.

- On note enfin $T_0 = \inf\{n \in \mathbb{N} / I_n = 0\}$ la variable aléatoire qui retourne l'entier n correspondant au premier retour en 0

4.1.1 Distribution de T_0

Notre variable aléatoire suit la loi suivante :

$$\mathbb{P}(T_0 = 2n) = 0 \quad \forall n \in \mathbb{N}$$

$$\mathbb{P}(T_0 = 2n + 1) = b^n d^{n+1} C_n \quad \forall n \in \mathbb{N} \text{ où } C_n = \frac{\binom{2n}{n}}{n+1}$$

Preuve :

Comme notre marche est issue de 1, elle occupe alors des états pairs aux temps impairs, et inversement des états impairs aux temps pairs. Donc T_0 est soit impair soit infini. Ainsi $\mathbb{P}(T_0 = 2n) = 0 \quad \forall n \in \mathbb{N}$

Si $T_0 = 2n+1$, la marche jusqu'à $2n$ est une excursion dont tous les états sont supérieurs à un, de taille $2n$. Le nombre d'excursions de taille $2n$ est donné par le nombre de Catalan :

$$C_n = \frac{\binom{2n}{n}}{n+1} = \frac{(2n)!}{(n+1)!n!}$$

Par ailleurs, comme au total, pendant une excursion, on effectue n transition vers le haut ($i \rightarrow i+1$) et n transitions vers le bas ($i \rightarrow i-1$), chaque excursion a la probabilité $(bd)^n$

D'où : $\mathbb{P}(T_0 = 2n + 1) = b^n d^{n+1} C_n$

4.1.2 Calcul de l'espérance de T_0 :

Soit $\tau_1 = \mathbb{E}(T_0)$ l'espérance de la chaîne issue de 1. Alors :

$$\tau_1 = \frac{2d}{\sqrt{1-4bd}(1+\sqrt{1-4bd})}$$

Preuve :

Soit G_{T_0} la fonction génératrice de T_0 , alors : $\mathbb{E}(T_0) = G'_{T_0}(1)$. D'autre part :

$$G_{T_0}(x) = \sum_{n=0}^{\infty} \mathbb{P}(T_0 = n)x^n \quad (1)$$

$$= \sum_{n=0}^{\infty} \mathbb{P}(T_0 = 2n + 1)x^{2n+1} \quad (2)$$

$$= \sum_{n=0}^{\infty} b^n d^{n+1} C_n x^{2n+1} \quad (3)$$

$$= \sum_{n=0}^{\infty} b^n d^{n+1} \frac{\binom{2n}{n}}{n+1} x^{2n+1} \quad (4)$$

$$= dx \sum_{n=0}^{\infty} \frac{\binom{2n}{n}}{n+1} (bdx^2)^n \quad (5)$$

Or $\sum_{n=0}^{\infty} \frac{\binom{2n}{n}}{n+1} x^n = \frac{2}{1 + \sqrt{1-4x}} \frac{2xd}{1 + \sqrt{1-4bdx^2}}$ par développement en série entière au voisinage de 0. Donc $G_{T_0}(x) = \frac{2xd}{1 + \sqrt{1-4bdx^2}}$

Ainsi par dérivation : $G'_{T_0}(x) = 2d \left(\frac{1 + \sqrt{1-4bdx^2} - \frac{-8bdx^2}{2\sqrt{1-4bdx^2}}}{(1 + \sqrt{1-4bdx^2})^2} \right)$ D'où :

$$\mathbb{E}(T_0) = 2d \left(\frac{1 + \sqrt{1-4bd} + \frac{4bd}{\sqrt{1-4bd}}}{(1 + \sqrt{1-4bd})^2} \right) \quad (6)$$

$$= 2d \left(\frac{\sqrt{1-4bd} + 1 - 4bd + 4bd}{\sqrt{1-4bd}(1 + \sqrt{1-4bd})^2} \right) \quad (7)$$

$$= \frac{2d}{\sqrt{1-4bd}(1 + \sqrt{1-4bd})} \quad (8)$$

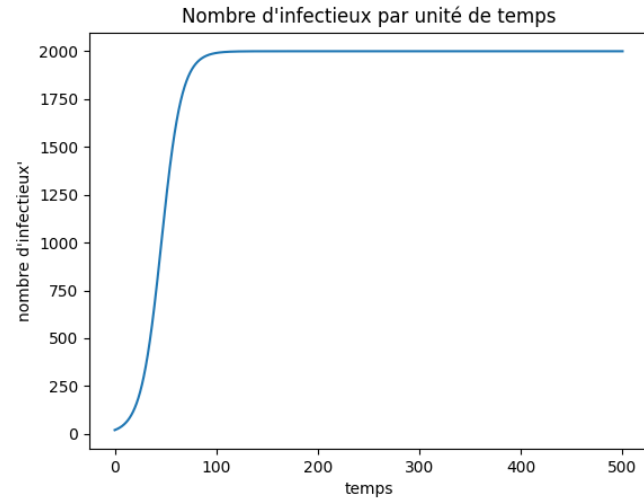
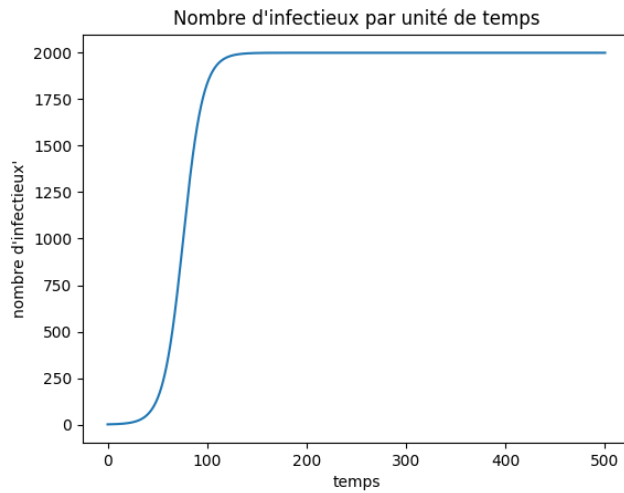
4.2 Seconde approche : modèle complet

On peut calculer l'espérance de la variable T_0 dans le modèle SIS que nous avons décrit tout au début (sans les hypothèses simplificatrices). Cependant cette manoeuvre dépasse le cadre du programme. On se contentera donc d'estimer cette espérance numériquement plus tard à l'aide de l'outil informatique Python

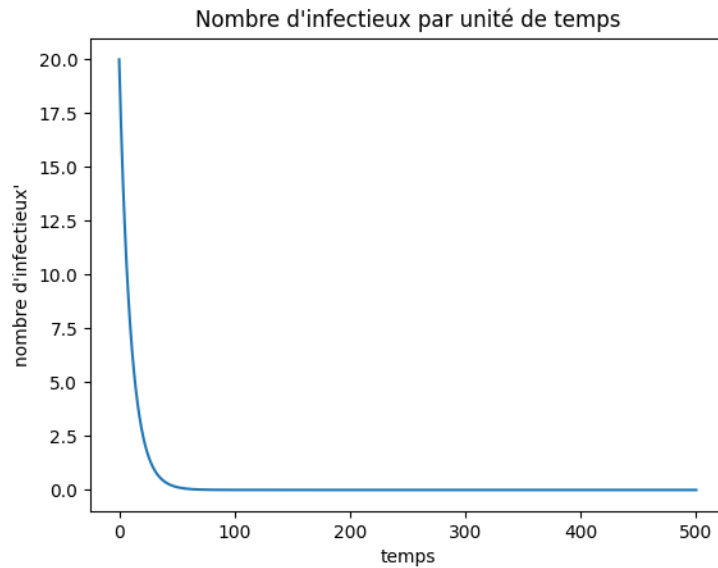
A Simulations :

A.1 Représentation du modèle déterministe :

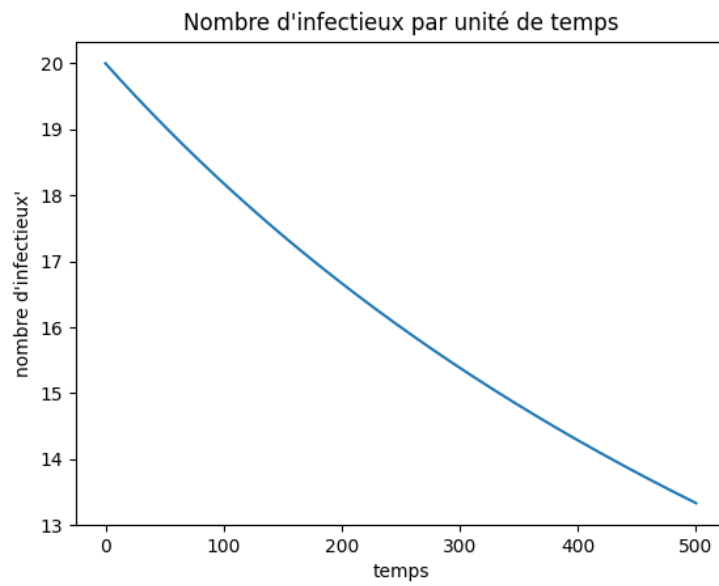
• On fixe : $N = 10^5$, $c = 0.1$, $\beta = 0.5$ (Donc $1 < R_0$). Les deux figures suivantes montrent respectivement le résultat pour $I_0 = 1$ et $I_0 = 20$. On remarque que le régime permanent est atteint dans les deux cas :



• On fixe : $N = 10^5$, $c = -0.1$, $\beta = 0.5$ (Donc $R_0 < 1$)

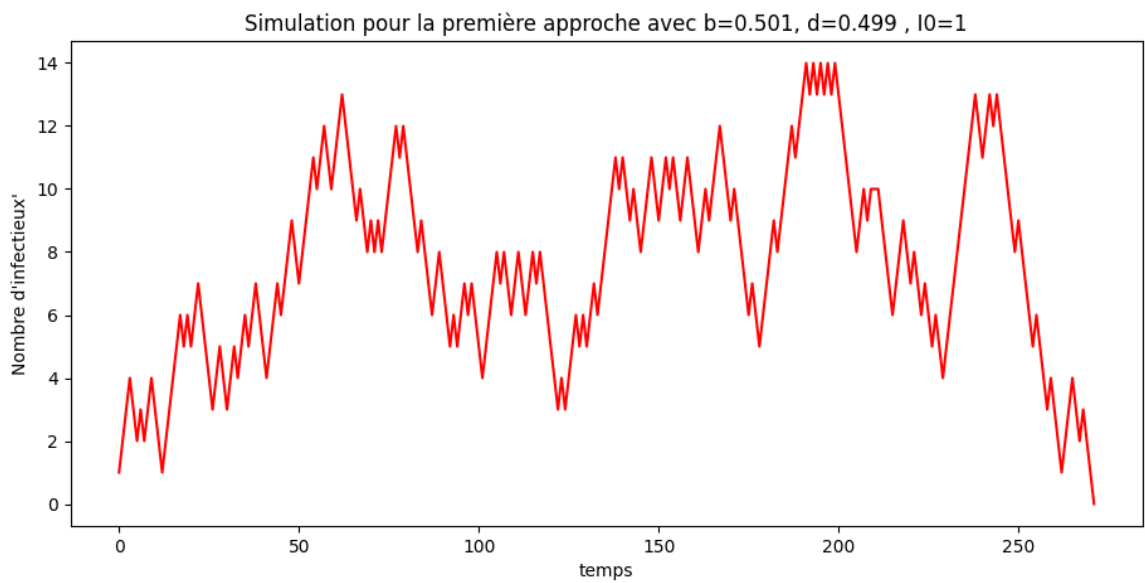


• On fixe : $N = 10^5$, $c = 0$, $\beta = 0.5$ (Donc $R_0 = 1$)



A.2 Représentation du modèle stochastique :

A.2.1 Première approche :



La fonction *taulpremiereapproche* décrite en annexe donne la moyenne du temps d'absorption pour 10^5 simulations :

```

(<module>)>>> taul_premiereapproche(100000)
535.4449343621606

(<module>)>>> taul_premiereapproche(100000)
575.3980358753382

(<module>)>>> taul_premiereapproche(100000)
521.9933874361286

(<module>)>>> taul_premiereapproche(100000)
535.6264166081637

(<module>)>>> taul_premiereapproche(100000)
591.7752380952381

(<module>)>>> taul_premiereapproche(100000)
524.2414795509222

(<module>)>>> taul_premiereapproche(100000)
501.32598456759195

```

La formule qu'on a trouvé dans le 4.1.2 donne :

```

(<module>)>>> 2*0.499/sqrt(1-4*0.501*0.499)*(1+sqrt(1-4*0.501*0.499))
499.9979999997504

```

On est sûr ainsi que la formule qu'on a donné est correct

Critiques du modèle :

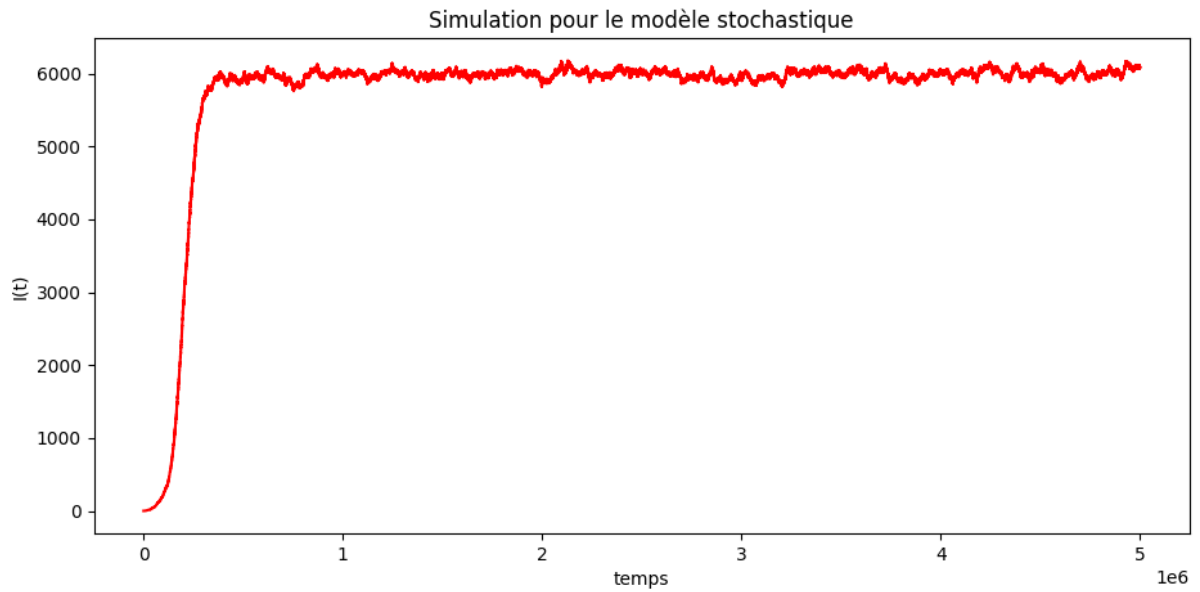
- Non réaliste car la probabilité d'attraper une maladie ne peut être constante puisqu'elle est liée directement au nombre d'infectieux à un instant donné. Or cette première approche ne prend même pas en considération la taille de la population.
- Le modèle SIS décrit une épidémie pour laquelle on ne développe pas d'immunité (comme certaines maladies transmissibles sexuellement 'STD'), ainsi on peut prévoir que le taux de contact peut être faible mais le nombre d'infectieux croira éventuellement. Ceci est mieux décrit par le modèle complet.

Conclusion :

Réduire notre problème à celui d'une marche aléatoire simple est donc incapable de décrire fidèlement le phénomène étudié, c'est donc à rejeter.

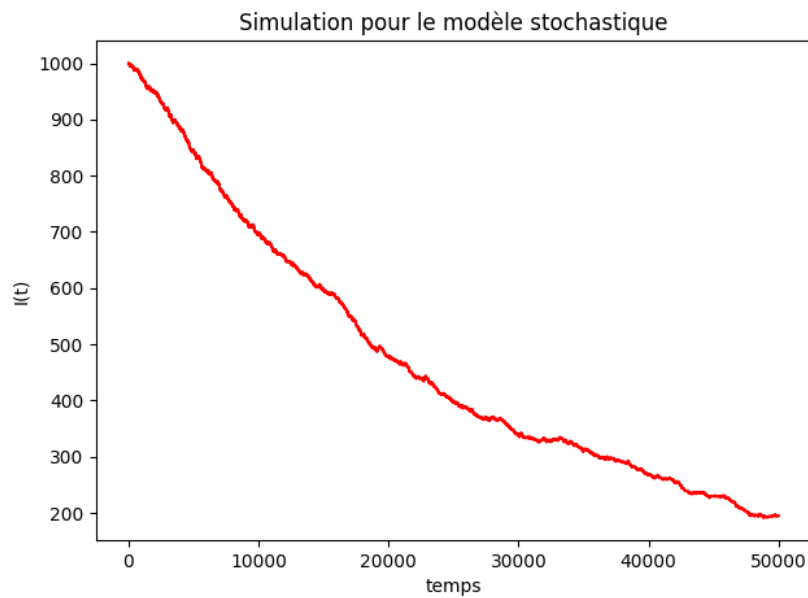
A.2.2 Deuxième approche :

- Une simulation faite pour le modèle complet avec : $N = 10^4$ $\beta = 0.5$, $b=\gamma=0.1$, $\Delta t = 10^{-4}$. Donc $R_0 > 1$



Cette simulation ressemble sensiblement à celle obtenue par la version déterministe du modèle : le 'régime permanent' est atteint ici également.

- Une autre simulation pour, $N = 10^4$, avec $I_0 = 10^3$, $\beta = 0.5$, $b=\gamma=0.4$ et $\Delta t = 10^{-4}$. Donc $R_0 < 1$



Le même profil du modèle déterministe est obtenu à nouveau : avec $R_0 < 1$, l'épidémie tend à disparaître.

La fonction *tauldeuxiemeapproche* permet de calculer la moyenne du temps d'absorption sur une base de 10^2 ou 10^3 simulations (Par soucis de complexité croissante dans le temps, on ne peut pas faire plus de simulations comme dans la première approche). Les résultats obtenus sont :

```
>>> taul_deuxiemeapproche(100)
207737.79

>>> taul_deuxiemeapproche(100)
206702.74

>>> taul_deuxiemeapproche(100)
217823.08

>>> taul_deuxiemeapproche(100)
213616.73

>>> taul_deuxiemeapproche(100)
209881.08

>>> taul_deuxiemeapproche(1000)
209832.716

>>> taul_deuxiemeapproche(1000)
210659.658

>>> taul_deuxiemeapproche(1000)
210828.352
```

La matrice de transition utilisée est :

`matricetransition(N=104, $\beta = 0.5$, $b = 0.4$, $\gamma = 0.4$, $\Delta t = 0.0001$)`

FIN