Modélisation de l’évolution d’une épidémie

Mohammed El Azhar

Numéro du candidat: 33855

Année:2021

Thème: Enjeux sociétaux

El Azhar Mohammed , Numéro candidat: 33855

Introduction

↓

Modèle déterministe

↓

Problématique

↓

Modèle stochastique

↓

Résultats

↓

Annexe

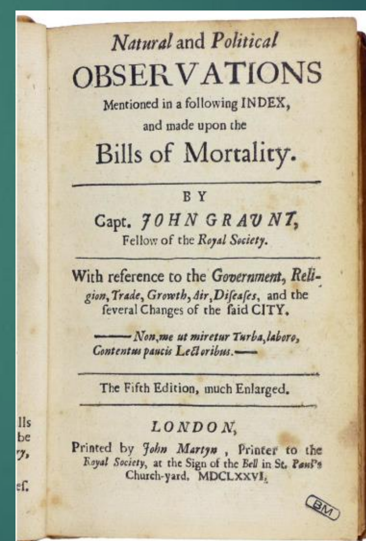
Introduction

El Azhar Mohammed , Numéro candidat: 33855

Histoire:

L'épidémiologie mathématique aurait vu le jour avec John Graunt (1620-1674)

Captain John Graunt



Natural and *Political*
OBSERVATIONS
Mentioned in a following INDEX,
and made upon the
Bills of Mortality.

BY
Capt. *JOHN GRAUNT*,
Fellow of the *Royal Society*.

With reference to the *Government*, *Religion*, *Trade*, *Growth*, *Air*, *Diseases*, and the several Changes of the said CITY.

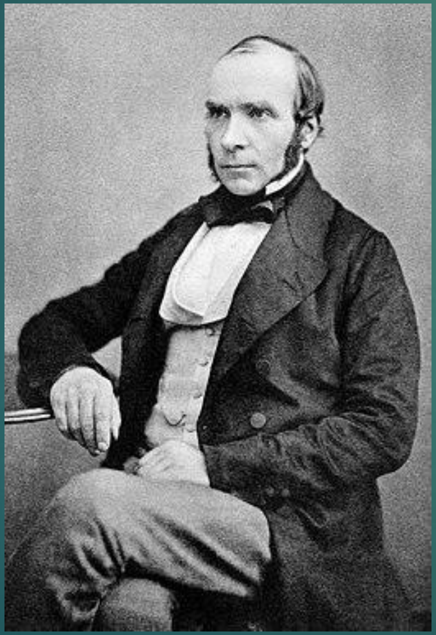
——*Non, me ut miretur Turba, laboro, Contentus paucis Lectoribus.*——

The Fifth Edition, much Enlarged.

LONDON,
Printed by *John Martyn*, Printer to the *Royal Society*, at the Sign of the *Bell* in St. *Paul's* Church-yard. MDCLXXVI.

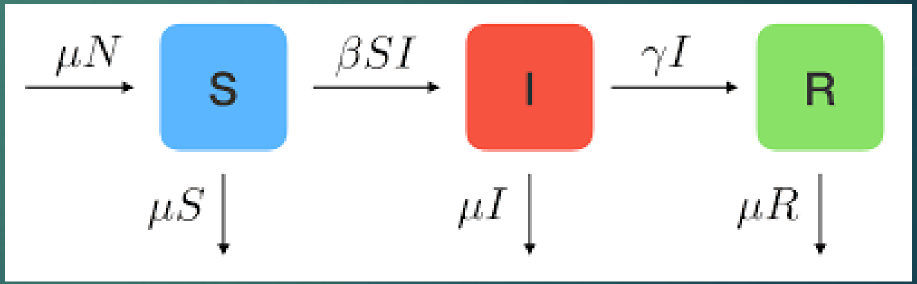
18eme siècle: Daniel Bernoulli élabora ce que l'on penserait être le premier modèle d'épidémiologie mathématique

John Snow étudie la mortalité par Cholera dans la ville de Londres entre 1849 et 1854



John Snow’s London Cholera map

Début 20eme siècle: A. G. Mckendrick et W. O. Kermack proposèrent un modèle fondamental dans la théorie d'épidémiologie mathématique: Modèle SIR



Quels sont les piliers de cette modélisation ?

Modélisation d'évolution d'une épidémie

- Description: équations + paramétrage
- Résolution
- Prévision

Modèles compartimentaux

Définition:

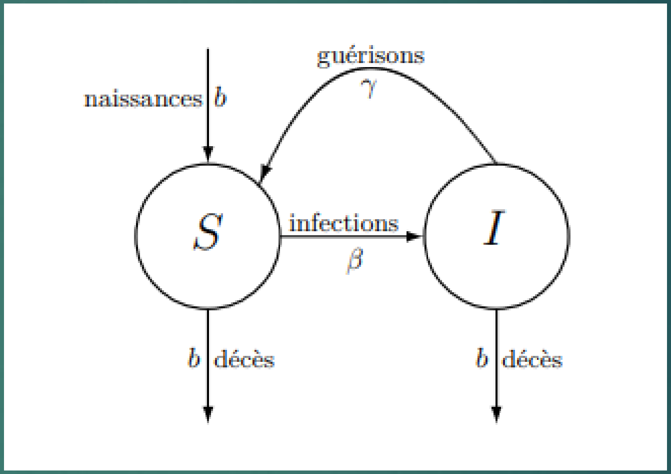
Il s'agit de segmenter la population proie à une épidémie en différentes catégories:

- Susceptibles: personnes non encore touchées
- Infectieux
- Retirés
- ..

Notations:

- N=S+I : la taille de la population (considéré 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.

Modèle SIS Déterministe



Système d'équations différentielles:

$$\begin{cases} S' = -\dfrac{\beta}{N} S I + (b+\gamma) I \\[2ex] I' = \dfrac{\beta}{N} S I - (b+\gamma) I \end{cases}$$

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 :

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

Solutions de l'équation différentielle non linéaire:

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

Nombre de reproduction de base:

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 :

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

À t=0 : S_0~N donc l'équation devient :

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

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

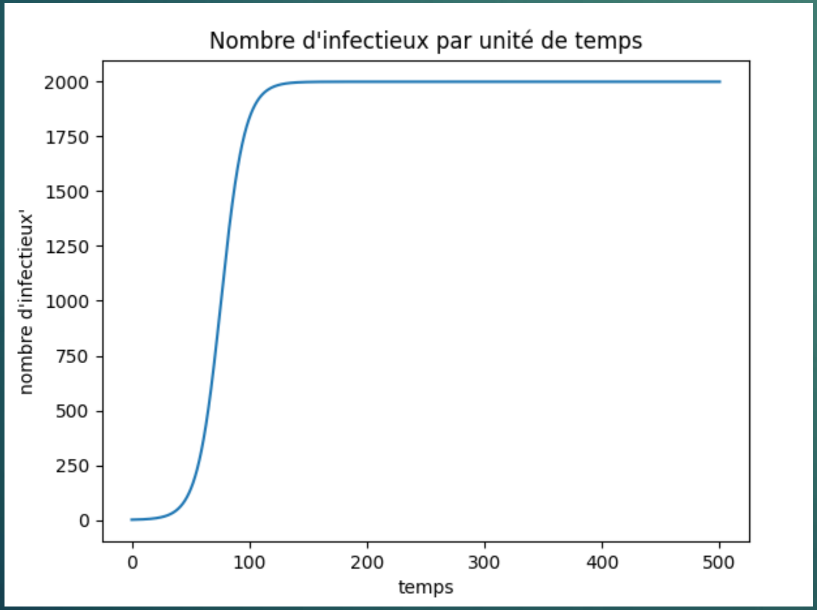
Si $R_0 > 1$, alors $I'|_{t=0} > 0$. Le nombre d'infectieux croît au début.

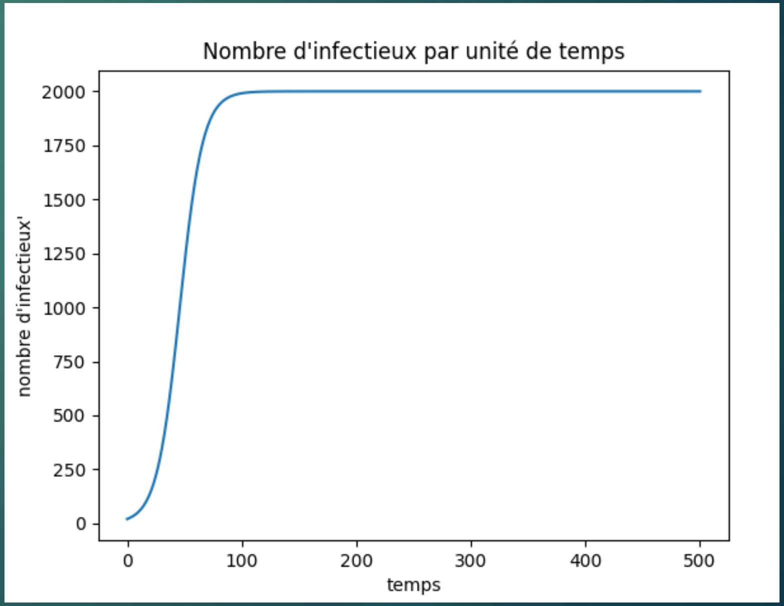
Si $R_0 < 1$, alors $I'|_{t=0} < 0$. Le nombre d'infectieux décroît

Conclusion

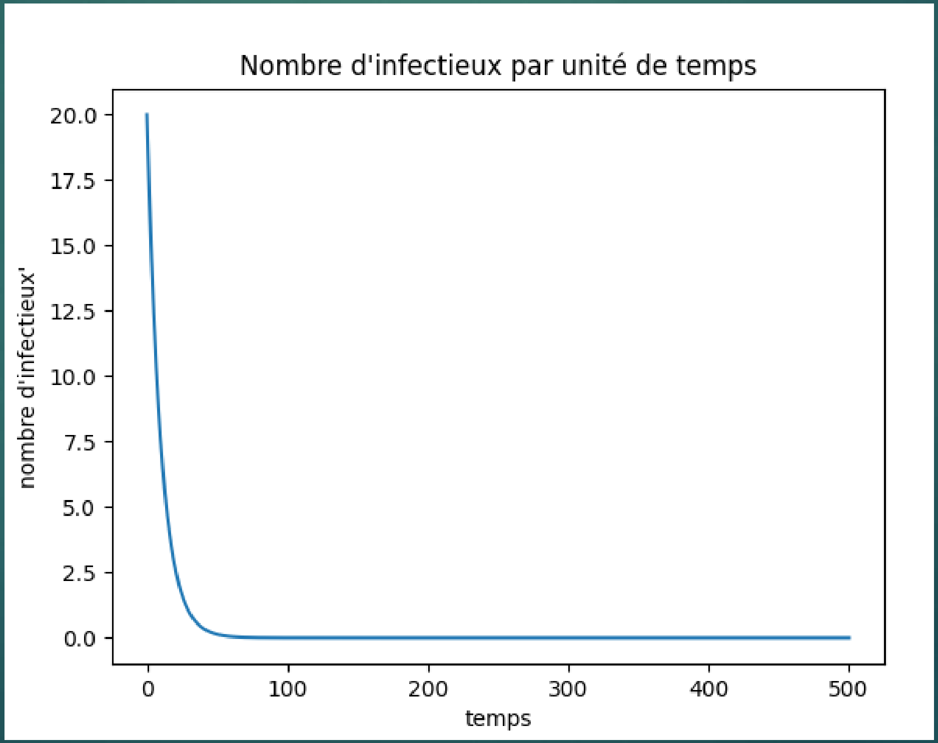
Ainsi le contrôle d'une épidémie est lié directement au contrôle du paramètre R_0. D'où l'intérêt de la distanciation sociale(contrôle de β), les paramètre b et γ, eux ne peuvent pas être contrôlés directement

• 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 :

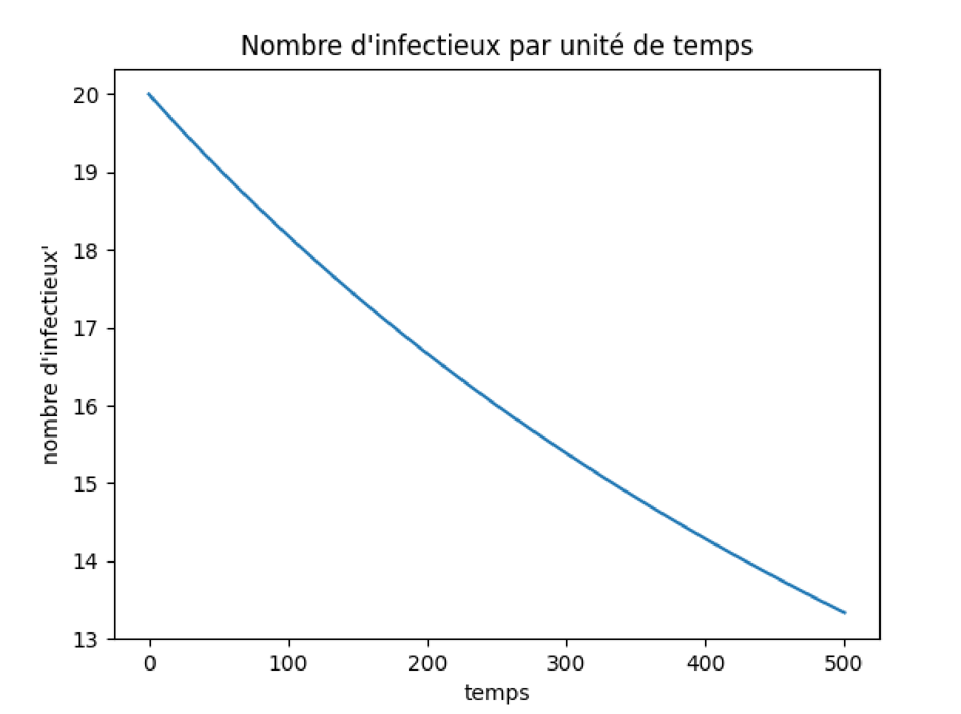




- On fixe : $N = 10^5, c = -0.1, \beta = 0.5$ (Donc $R_0 < 1$)



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



Contrôler l'épidémie $\Leftrightarrow$ Rendre R0<1

On peut se demander dans ce cas(lorsque R0<1) combien va durer l'épidémie ?

Problématique

Comment estimer le temps moyen d'absorption d'une épidémie?

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.

Modèle SIS stochastique

Notations adoptées:

- S(t) et I(t) deux variables aléatoires discrètes
- t $\in \{0,\ \Delta t,\ 2\Delta t,\ \dots\}$
- S(t), I(t) $\in \{0, 1, 2, \dots, N\}$
- $p_i(t) = \mathbb{P}(I(t) = i)$
- I_n=I(nΔt)
- $p_{ji} = \mathbb{P}(I_{n+1} = i | I_n = j)$

CMTD: Chaîne de Markov à temps discret

- *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 :

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

Hypothèse:

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

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

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

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

On allège les notations :

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

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 :

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

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

Distribution de T0:

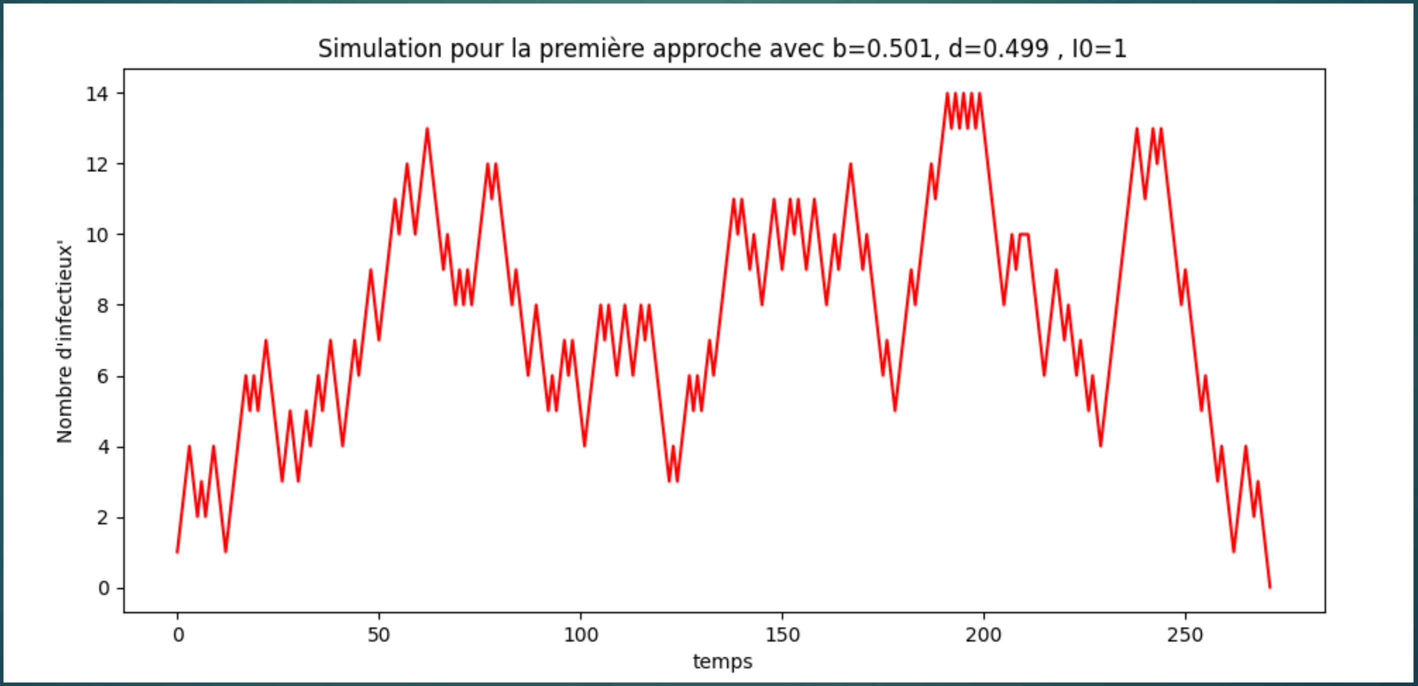
$$\mathbb{P}(T_0 = 2n) = 0 \ \forall n \in \mathbb{N}$$

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

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

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

On obtient la simulation suivante en implémentant notre modèle sur Python:



La fonction *tau1premiereapproche* 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

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

Les résultats théoriques et pratiques sont proches, donc on n'a pas d'erreur dans le raisonnement

Critique

- Non réaliste
- Ne décrit pas le phénomène fidèlement

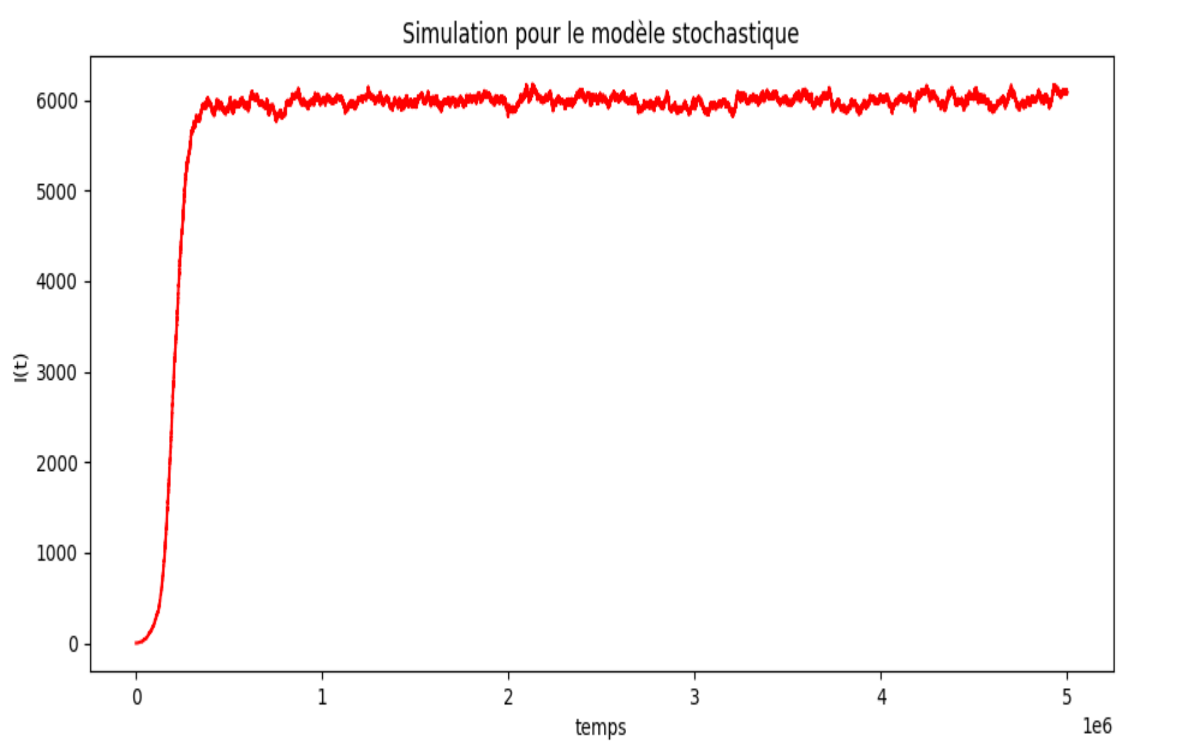
Conclusion

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

Seconde approche : modèle complet

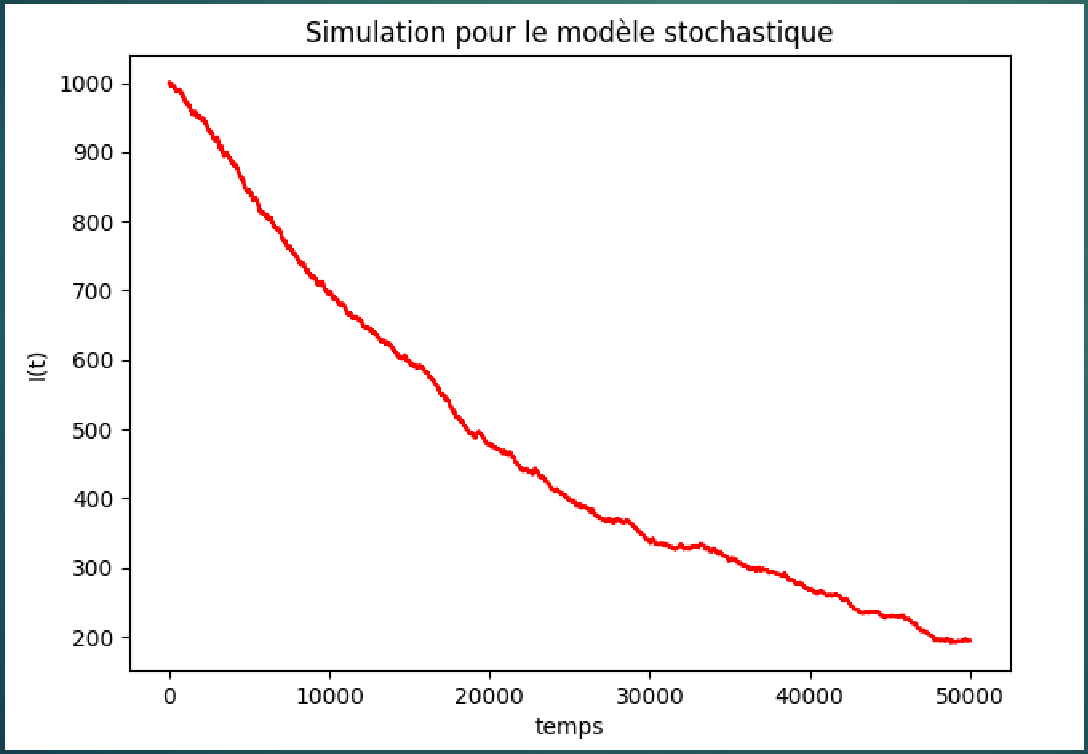
On peut calculer l'espérance de la variable T_0 dans le modèle SIS que nous avions décrit tout au début (sans les hypothèses simplificatrices). Cependant cette manoeuvre dépasse le cadre du programme. On se contentera donc d'estimer cette espérance numériquement plus tard à l'aide de l'outil informatique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

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



Cette courbe ressemble sensiblement à celle trouvée par le modèle déterministe

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



Le même profil déterministe est obtenu à nouveau : l'épidémie tend à disparaître quand R0 <1

La fonction *tau1deuxiemeapproche* permet de calculer la moyenne du temps d'absorption sur une base de 10^2 ou 10^3 simulations (Par de soucis de complexité croissante dans le temps, on ne peut pas faire plus de simulations comme dans la première approche). Les résultats obtenus sont :

```
>>> tau1_deuxiemeapproche(100)
207737.79

>>> tau1_deuxiemeapproche(100)
206702.74

>>> tau1_deuxiemeapproche(100)
217823.08

>>> tau1_deuxiemeapproche(100)
213616.73

>>> tau1_deuxiemeapproche(100)
209881.08

>>> tau1_deuxiemeapproche(1000)
209832.716

>>> tau1_deuxiemeapproche(1000)
210659.658

>>> tau1_deuxiemeapproche(1000)
210828.352
```

La matrice de transition utilisée est :

$$\text{matricetransition}(\text{N}=10^4, \beta = 0.5, b = 0.4, \gamma = 0.4, \Delta t = 0.0001)$$

Merci pour votre attention

Annexe

La solution de l'équation différentielle en I(t):

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

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

Distribution de T_0

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

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

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

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→i+1) et n transitions vers le bas(i→i-1), chaque excursion a la probabilité $(bd)^n$

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

Calcul de l'espérance de T_0 :

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

$$\begin{aligned} G_{T_0}(x) &= \sum_{n=0}^{\infty} \mathbb{P}(T_0 = n)x^n \\ &= \sum_{n=0}^{\infty} \mathbb{P}(T_0 = 2n+1)x^{2n+1} \\ &= \sum_{n=0}^{\infty} b^n d^{n+1} C_n x^{2n+1} \\ &= \sum_{n=0}^{\infty} b^n d^{n+1} \frac{\binom{2n}{n}}{n+1} x^{2n+1} \\ &= dx \sum_{n=0}^{\infty} \frac{\binom{2n}{n}}{n+1} (bdx^2)^n \end{aligned}$$

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

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 :

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

D'où:

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

Annexe: Code python des différents algorithmes utilisés

Modèle déterministe

Implémentation de I(t) pour I0=20

```
import numpy as np
pip install matplotlib
import matplotlib.pyplot as plt
import numpy.random as rnd
from numpy import exp
pip install scipy

beta=0.5
c=0
N=10000
I0=20
def I(t):
        if beta*I0==c/N:
                return c*N/beta
        elif c==0:
                return I0/(1+I0*beta*t/N)
        else:
                return c/(beta/N -(beta /N - c/I0)*exp(-c*t))


tt=np.linspace(0,500,10000)
y=[I(t) for t in tt]
plt.plot(tt,y)
plt.title("Nombre d'infectieux par unité de temps")
plt.xlabel('temps')
plt.ylabel("nombre d'infectieux'")
plt.show()
```

Modèle stochastique

Première approche: Simulation de marche aléatoire

```
def premiereapproche(n):
        x = 0
        y = 5
        xposition = [0]
        yposition = [1]
        for i in range (1,n+1):
                step = np.random.uniform(0,1)
                if step < 0.6:
                        x += 1
                        y += 1
                if step > 0.6:
                        x += 1
                        y += -1
                if y==0:
                        break
                xposition.append(x)
                yposition.append(y)
        return [xposition,yposition]
Randwalk = premiereapproche(50000)
plt.plot(Randwalk[0],Randwalk[1],'r-', label = "premiereapproche")
plt.title("Simulation pour la première approche")
plt.xlabel("temps")
plt.ylabel("Nombre d'infectieux'")
plt.show()
```

Calcul de la moyenne expérimentale du temps moyen d'absorption:

```
def tau1premiereapproche(n):
        m=0
        l=0
        for i in range(n):
                if premiereapproche(100000)[1][-1]==0:
                        m+=1
                        l+=premiereapproche(100000)[0][-1]
        return l/m
```

Implémentation de la matrice de transition du processus Markovien du modèle SIS

```
def matricetransition(N,beta,b,gamma,dt):
        M=np.zeros((N+1,N+1))
        for i in range(1,N):
                M[i+1][i]=beta*i*(N-i)*dt/N
                M[i-1][i]=(b+gamma)*i*dt
                M[i][i]=1-beta*i*(N-i)*dt/N - (b+gamma)*i*dt
        M[0][0]=1
        M[N-1][N]=(b+gamma)*N*dt
        M[N][N]=1-(b+gamma)*N*dt
        return M
```

Deuxième approche: Simulation utilisant le modèle complet .

```
N=10000
def deuxiemeapproche(n):
        x=0
        y=1
        xposition=[0]
        yposition=[1]
        for i in range(n):
                step=np.random.uniform(0,1)
                if step<M[y-1][y]:
                        x+=1
                        y+=-1
                if M[y-1][y]<=step<=M[y-1][y]+M[y][y]:
                        x+=1
                        y+=0
                if step>M[y-1][y]+M[y][y]:
                        x+=1
                        y+=1
                xposition.append(x)
                yposition.append(y)
                if y in [0,N]:
                        break
        return [xposition,yposition]

Randwalk = deuxiemeapproche(50000)
plt.plot(Randwalk[0],Randwalk[1],'r-', label = "Deuxiemeapproche")
plt.title("Simulation pour le modèle stochastique")
plt.xlabel("temps")
plt.ylabel("I(t)")
plt.show()
```

Calcul experimental de la moyenne du temps d'absorption:

```
def tau1deuxiemeapproche(n):
    m=0
    l=0
    for i in range(n):#on evalue sur n test
        if deuxiemeapproche(500000)[1][-1]==0:
            m+=1
            l+=deuxiemeapproche(500000)[0][-1]
    return l/m
```