

Étude comparative de la propagation épidémique : Du modèle déterministe SIR à l'approche stochastique par chaînes de Markov

YAKOUBI Yanis

Sous la direction de : Léon Matar TINE
TIPE – L2 Mathématiques

Table des matières

Remerciements et motivations	3
1 Introduction et cadre théorique	4
2 Partie I : L'approche déterministe et les lois universelles de propagation	5
2.1 Mise en équation : Hypothèses biologiques et modèle SIR	5
2.2 Validation du modèle : Conservation et positivité	6
2.3 Le seuil épidémique R_0 et la gestion du pic infectieux	8
2.4 Visualisation des trajectoires : L'épidémie dans l'espace des phases	11
2.5 Prévission à long terme : L'immunité collective naturelle	13
2.6 Justification rigoureuse : Stabilité du point d'équilibre	15
2.7 Le modèle SIR ouvert : Démographie et équilibre endémique	16
2.8 Synthèse : Portée et limites des modèles déterministes	18
3 Partie II : L'approche stochastique par les chaînes de Markov	20
3.1 Limites du continu et nécessité d'une approche discrète	20
3.2 Définitions mathématiques : Chaînes de Markov à temps continu	20
3.3 Le cadre mathématique : Processus de saut de Markov	21
3.4 L'équation maîtresse (Équation de Kolmogorov forward)	22
3.5 Outil mathématique : Le processus de Galton-Watson	23
3.6 Le résultat fondamental : La probabilité d'extinction précoce	25
3.7 Simulation exacte : L'algorithme de Gillespie	27
3.8 Application aux petits foyers infectieux : La distribution bimodale	28
3.9 Synthèse de l'approche stochastique	30
4 Partie III : La confrontation des modèles et la loi des grands nombres	31
5 Conclusion et perspectives	34
Références	35
Annexe	36
A Scripts Python des simulations	36

Remerciements et motivations

Je tiens tout d'abord à adresser mes plus sincères remerciements à mon encadrant, Monsieur Léon Matar Tine, pour avoir accepté de superviser ce Travail d'Initiative Personnelle Encadré (TIPE). Ses précieux conseils, sa grande rigueur mathématique et sa bienveillance ont été d'une aide inestimable pour structurer ma pensée, corriger mes incertitudes et mener à bien ce travail. Je tiens également à remercier Monsieur Guillaume Aubrun et Monsieur Laurent Bétermin pour m'avoir permis de réaliser ce travail tout en suivant l'UE Probabilités discrètes et statistiques descriptives en parallèle.

Le choix de ce sujet est né d'un intérêt profond pour la modélisation mathématique et sa capacité à capturer la complexité du monde réel. L'attrait principal de cette thématique résidait, pour moi, dans la confrontation conceptuelle de deux branches mathématiques fondamentalement différentes : le déterminisme strict de l'analyse (via les équations différentielles) et la nature imprévisible des probabilités.

Au-delà de l'aspect purement théorique, ce travail représentait une occasion unique de lier la réflexion mathématique à l'expérimentation numérique. Il m'a permis de développer une réelle appétence pour la simulation algorithmique et l'étude des processus stochastiques. Comprendre comment la loi des grands nombres parvient à faire émerger un ordre collectif implacable à partir d'un chaos individuel absolu a été une découverte fascinante.

1 Introduction et cadre théorique

Contexte et problématique

La prévision de l'évolution des maladies infectieuses constitue un enjeu majeur de santé publique, comme l'ont rappelé avec vigueur les crises sanitaires récentes (Ébola, COVID-19). Si la biologie décrit les mécanismes cellulaires de l'infection, ce sont les mathématiques qui permettent de comprendre la dynamique de propagation à l'échelle d'une population.

Historiquement, l'approche dominante, initiée par Kermack et McKendrick en 1927, est **déterministe**. Elle repose sur des systèmes d'équations différentielles considérant les grandes populations comme des fluides continus. Ce modèle est efficace pour les grandes populations (Loi des grands nombres), mais il présente une faiblesse conceptuelle majeure : il ignore le caractère intrinsèquement aléatoire des interactions individuelles.

Cependant, au début d'une épidémie ou dans un petit foyer (une école, un service hospitalier), le nombre d'individus est faible. Le hasard joue alors un rôle critique : une épidémie potentiellement grave peut s'éteindre spontanément si le "patient zéro" guérit avant de transmettre le pathogène.

Ce mémoire se propose d'étudier cette dichotomie. **Dans quelle mesure l'approche déterministe classique échoue-t-elle à décrire la réalité des petits foyers infectieux, et comment l'approche stochastique par chaînes de Markov permet-elle de quantifier les phénomènes d'extinction aléatoire ?**

Pour répondre à cette problématique, nous définirons d'abord le cadre mathématique du modèle compartimental **SIR (Sains - Infectés - Rétablis)**, avant de confronter ses résultats théoriques à une modélisation probabiliste et à des simulations numériques.

2 Partie I : L'approche déterministe et les lois universelles de propagation

Dans cette partie, nous analysons le comportement global des solutions du système SIR. L'objectif est double : valider la cohérence biologique du modèle et fournir aux décideurs des outils prédictifs (pic épidémique, taux de vaccination nécessaire).

2.1 Mise en équation : Hypothèses biologiques et modèle SIR

Pour modéliser la propagation, nous divisons la population totale N (supposée constante) en trois compartiments distincts.

Définition 1 (Compartiments épidémiologiques). *Soit $t \in \mathbb{R}^+$ le temps. On note :*

- $S(t)$: le nombre d'individus **Sains** (*Susceptibles*) à l'instant t .
- $I(t)$: le nombre d'individus **Infectés** (*et infectieux*) à l'instant t .
- $R(t)$: le nombre d'individus **Rétablis** (*ou Retirés, immunisés ou décédés*) à l'instant t .

Par hypothèse de population fermée, nous avons la relation de conservation suivante :

$$\forall t \geq 0, \quad S(t) + I(t) + R(t) = N$$

Dans cette approche, on considère que S, I, R sont des fonctions dérivables du temps. La dynamique est régie par la "loi d'action de masse" : le nombre de nouvelles infections est proportionnel au nombre de rencontres entre sains et infectés.

Définition 2 (Système SIR de Kermack-McKendrick). *L'évolution de l'épidémie est décrite par le système d'équations différentielles non-linéaires autonome suivant :*

$$\begin{cases} S'(t) = -\frac{\beta}{N}S(t)I(t) \\ I'(t) = \frac{\beta}{N}S(t)I(t) - \gamma I(t) \\ R'(t) = \gamma I(t) \end{cases} \quad (1)$$

où

- $\beta > 0$ est le **taux de contact efficace**.
- $\gamma > 0$ est le **taux de guérison** (*inverse de la durée moyenne d'infection*).

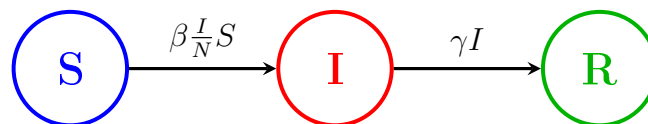


FIGURE 1 – **Schéma du modèle SIR classique.** Les individus passent de l'état Sain à Infecté avec un taux de transmission β , puis guérissent avec un taux γ .

2.2 Validation du modèle : Conservation et positivité

Pour que notre modèle soit physiquement et biologiquement cohérent avec l'hypothèse d'une population fermée, il est nécessaire de vérifier que nos équations garantissent deux propriétés fondamentales : la conservation de la population totale au cours du temps, et la positivité stricte ou nullité des effectifs de chaque compartiment.

Proposition 1 (Conservation de la population).

La population totale $N(t) = S(t) + I(t) + R(t)$ est un invariant du système. Elle est constante pour tout $t \geq 0$.

Démonstration. Calculons la dérivée de la population totale en sommant les trois équations du système différentiel :

$$N'(t) = S'(t) + I'(t) + R'(t)$$

En remplaçant par les expressions du système (1) :

$$N'(t) = \left(-\frac{\beta}{N} S(t) I(t) \right) + \left(\frac{\beta}{N} S(t) I(t) - \gamma I(t) \right) + (\gamma I(t))$$

On observe que tous les termes s'annulent :

$$N'(t) = 0$$

La dérivée est nulle sur $\mathbb{R}^+$, donc la fonction $N(t)$ est constante. On notera $N(0) = N$. $\square$

Proposition 2 (Positivité des solutions). *Si les conditions initiales S_0, I_0, R_0 sont positives, alors les fonctions $S(t)$, $I(t)$ et $R(t)$ restent positives pour tout $t \geq 0$.*

Démonstration. 1. Positivité des individus sains $S(t)$:

Partons de la première équation du système :

$$S'(t) = -\frac{\beta}{N} S(t) I(t)$$

En transposant tous les termes dans le même membre, l'équation se réécrit :

$$S'(t) + \frac{\beta}{N} I(t) S(t) = 0$$

Multiplions cette équation par la fonction $E(t) = \exp\left(\int_0^t \frac{\beta}{N} I(u) du\right)$. Par composition des dérivées, la dérivée de $E(t)$ est :

$$E'(t) = \frac{\beta}{N} I(t) \exp\left(\int_0^t \frac{\beta}{N} I(u) du\right) = \frac{\beta}{N} I(t) E(t)$$

L'équation multipliée par $E(t)$ devient donc :

$$S'(t)E(t) + S(t)\frac{\beta}{N}I(t)E(t) = 0$$

Soit, en remplaçant le second terme par la dérivée $E'(t)$:

$$S'(t)E(t) + S(t)E'(t) = 0$$

On reconnaît la dérivée d'un produit $(u \cdot v)' = u'v + uv'$:

$$\frac{d}{dt} [S(t)E(t)] = 0$$

La dérivée de la fonction $t \mapsto S(t)E(t)$ étant nulle, cette fonction est constante sur $\mathbb{R}^+$. On l'évalue en $t = 0$:

$$S(t)E(t) = S(0)E(0)$$

Puisque $E(0) = \exp(0) = 1$, on a $S(t)E(t) = S_0$. En isolant $S(t)$, on obtient l'expression explicite :

$$S(t) = S_0 \exp \left(- \int_0^t \frac{\beta}{N} I(u) du \right)$$

La fonction exponentielle est à valeurs strictement positives sur $\mathbb{R}$. Par hypothèse, le nombre initial de sains est strictement positif ($S_0 > 0$). Le produit des deux est donc strictement positif.

Conclusion : Pour tout $t \geq 0$, $S(t) > 0$.

2. Positivité des individus infectés $I(t)$:

Un raisonnement similaire s'applique. Partons de la deuxième équation :

$$I'(t) = \left(\frac{\beta S(t)}{N} - \gamma \right) I(t)$$

On la réécrit sous la forme :

$$I'(t) - \left(\frac{\beta S(t)}{N} - \gamma \right) I(t) = 0$$

Cette fois, on multiplie par le facteur intégrant $F(t) = \exp \left(- \int_0^t \left(\frac{\beta S(u)}{N} - \gamma \right) du \right)$. La dérivée de ce facteur est $F'(t) = - \left(\frac{\beta S(t)}{N} - \gamma \right) F(t)$. L'équation multipliée par $F(t)$ donne :

$$I'(t)F(t) - I(t) \left(\frac{\beta S(t)}{N} - \gamma \right) F(t) = 0$$

Ce qui équivaut à la dérivée d'un produit :

$$I'(t)F(t) + I(t)F'(t) = 0 \implies \frac{d}{dt} [I(t)F(t)] = 0$$

En intégrant de la même manière et sachant que $F(0) = \exp(0) = 1$, on trouve $I(t)F(t) = I_0$, soit :

$$I(t) = I_0 \exp \left(\int_0^t \left(\frac{\beta S(u)}{N} - \gamma \right) du \right)$$

Puisque l'exponentielle est toujours strictement positive et que $I_0 > 0$ (il y a au moins un malade au départ), on en déduit que le produit est strictement positif.

Conclusion : Pour tout $t \geq 0$, $I(t) > 0$.

3. Pour les rétablis $R(t)$: On a $R'(t) = \gamma I(t)$. Comme $\gamma > 0$ et que nous venons de montrer que $I(t) > 0$, alors $R'(t) > 0$. $R(t)$ est une fonction strictement croissante. Comme $R(0) \geq 0$, alors $R(t) \geq 0$ pour tout t . $\square$

Définition 3 (Espace des phases). *Compte tenu des propositions précédentes, l'étude du système se restreint au domaine compact triangulaire Ω dans le plan (S, I) :*

$$\Omega = \{(S, I) \in \mathbb{R}_+^2 \mid S + I \leq N\}$$

2.3 Le seuil épidémique R_0 et la gestion du pic infectieux

Nous cherchons ici à caractériser les variations du nombre d'infectés $I(t)$. Le concept central de l'épidémiologie mathématique est le seuil de déclenchement.

On définit le taux de reproduction de base R_0 comme le nombre moyen de cas secondaires produits par un unique individu infecté dans une population entièrement saine.

Définition 4 (Taux de reproduction de base R_0). *Dans le cas spécifique de notre modèle SIR, le taux de reproduction de base est donné par :*

$$R_0 = \frac{\beta}{\gamma}$$

Justification de la formule :

Un individu infectieux transmet la maladie avec un taux de contact efficace β . De plus, puisque les individus guérissent au taux γ , la durée moyenne de la période de contagion est donnée par $\frac{1}{\gamma}$. Le nombre total de personnes qu'un malade va infecter correspond donc au produit de sa "vitesse de transmission" par sa "durée de contagiosité" :

$$R_0 = \beta \times \frac{1}{\gamma} = \frac{\beta}{\gamma}$$

Illustration par des exemples réels :

Pour bien mesurer l'impact de ce nombre de reproduction, voici quelques estimations classiques. Ces exemples illustrent comment la valeur de R_0 dicte l'urgence de la réponse sanitaire :

- **Grippe saisonnière** ($R_0 \approx 1,3$) : La propagation est modérée et gérable par le système de santé sans mesures restrictives majeures.
- **COVID-19 (souche initiale $R_0 \approx 3$)** : Chaque malade contamine en moyenne 3 personnes. Cette valeur entraîne une croissance exponentielle très rapide, justifiant historiquement la mise en place de confinements.
- **Rougeole** ($R_0 > 12$) : C'est l'une des maladies virales les plus contagieuses connues. Un tel R_0 implique qu'une épidémie flambera instantanément à moins qu'une immense majorité de la population ne soit vaccinée.

Lemme 1 (Condition de croissance). *Le signe de la dérivée $I'(t)$ dépend exclusivement de la position de $S(t)$ par rapport à un seuil critique $S_c = \frac{N}{R_0}$.*

Démonstration. Partons de la deuxième équation du système SIR :

$$I'(t) = \frac{\beta}{N} S(t) I(t) - \gamma I(t)$$

Factorisons par $\gamma I(t)$ (qui est strictement positif pour $t > 0$) :

$$I'(t) = \gamma I(t) \left(\frac{\beta}{\gamma N} S(t) - 1 \right)$$

On reconnaît le terme $R_0 = \beta/\gamma$. L'équation devient :

$$I'(t) = \gamma I(t) \left(\frac{R_0}{N} S(t) - 1 \right)$$

Le signe de $I'(t)$ est donc identique au signe de $\frac{R_0}{N} S(t) - 1$. On en déduit que :

$$I'(t) > 0 \iff \frac{R_0}{N} S(t) > 1 \iff S(t) > \frac{N}{R_0}$$

□

Ce lemme décrit la mécanique interne du système à tout instant $t \geq 0$. Son application immédiate aux conditions initiales permet d'énoncer le théorème du seuil.

Théorème 1 (Théorème du seuil). *Considérons une introduction infinitésimale d'infectés dans une population saine ($S(0) \approx N$).*

- Si $R_0 \leq 1$, alors $I'(0) \leq 0$: l'épidémie ne démarre pas.
- Si $R_0 > 1$, alors $I'(0) > 0$: il y a une flambée épidémique.

Démonstration. C'est une conséquence immédiate du lemme précédent appliqué à l'instant initial. À $t = 0$, on a $S(0) \approx N$. La condition de croissance initiale $I'(0) > 0$ est équivalente à :

$$S(0) > \frac{N}{R_0} \iff N > \frac{N}{R_0} \iff 1 > \frac{1}{R_0} \iff R_0 > 1$$

□

Application politique : L'objectif des politiques de santé publique est d'agir sur les deux paramètres de R_0 pour le faire passer en dessous du seuil critique de 1. D'une part, les mesures "barrières" (masques, confinement) visent à réduire artificiellement le taux de contact β . D'autre part, le développement de traitements médicaux ou les politiques de dépistage et d'isolement rapide permettent d'augmenter le taux de retrait γ (en réduisant la durée effective de la période contagieuse). Si l'on parvient à combiner ces facteurs pour que le R_0 effectif soit strictement inférieur à 1, alors $I'(t)$ devient négatif et l'épidémie régresse, même si la population n'est pas encore immunisée.

Théorème 2 (Existence et valeur du pic épidémique). *Supposons $R_0 > 1$ et $S_0 > \frac{N}{R_0}$. La fonction $I(t)$ n'est pas monotone : elle est d'abord croissante, atteint un maximum unique (le pic) à l'instant t_{pic} tel que $S(t_{pic}) = \frac{N}{R_0}$, puis décroît et tend vers 0 lorsque $t \rightarrow +\infty$.*

Démonstration. La démonstration s'articule en quatre points rigoureux.

(1) Étude du signe de la dérivée

Factorisons la dérivée $I'(t)$ par $\gamma I(t)$ (strictement positif sur l'intervalle d'existence) :

$$I'(t) = \gamma I(t) \left(\frac{R_0}{N} S(t) - 1 \right)$$

Le signe de $I'(t)$ dépend exclusivement de la position de $S(t)$ par rapport au seuil critique $S_c = N/R_0$.

(2) Application du Théorème des Valeurs Intermédiaires (TVI)

Considérons la fonction $S(t)$ sur $[0, +\infty[$:

- Elle est **continue** et **strictement décroissante** (car $S' < 0$).
- À $t = 0$, $S(0) > S_c$ (hypothèse de départ).
- À $t \rightarrow +\infty$, $S_\infty < S_c$ (car S décroît et ne peut pas rester coincée au-dessus du seuil, sinon I exploserait, ce qui est impossible dans une population finie).

Pour visualiser les conditions d'application du théorème, dressons un tableau de comparaison des valeurs aux bornes :

Temps t	0	t_{pic}	$+\infty$
Variations de $S(t)$	S_0	$S_c = \frac{N}{R_0}$	S_∞

The diagram illustrates the variation of the susceptible population S(t) over time. It shows a decreasing curve starting at S0 at t=0, passing through the critical threshold Sc = N/R0 at time t_pic, and approaching S_infinity as t goes to infinity. A vertical dashed line marks the point (t_pic, Sc) on the curve.

D'après le **TVI** pour les fonctions monotones, il existe un **unique** instant t_{pic} tel que $S(t_{pic}) = S_c = N/R_0$. C'est l'unique instant où la dérivée s'annule : $I'(t_{pic}) = 0$.

(3) Tableau de variations de l'épidémie

Nous en déduisons les variations de $I(t)$:

Temps t	0	t_{pic}	$+\infty$
Signe de $I'(t)$	+	0	-
Variations de $I(t)$	I_0	I_{maz}	0

The diagram illustrates the variation of the infected population I(t) over time. It shows an increasing curve starting at I0 at t=0, reaching a maximum value Imaz at time t_pic, and then decreasing towards 0 as t goes to infinity. A vertical dashed line marks the point (t_pic, Imaz) on the curve.

(4) Preuve de la convergence vers 0 (par l'absurde)

Pour $t > t_{pic}$, la fonction $I(t)$ est décroissante et minorée par 0. D'après le **théorème de la limite monotone**, elle converge vers une limite réelle positive $I_\infty \geq 0$.

Supposons par l'absurde que cette limite soit strictement positive ($I_\infty > 0$). Alors, asymptotiquement, la vitesse de décroissance des sains deviendrait constante :

$$\lim_{t \rightarrow +\infty} S'(t) = \lim_{t \rightarrow +\infty} \left(-\frac{\beta}{N} S(t) I(t) \right) = -\frac{\beta}{N} S_\infty I_\infty = -K < 0, \quad \text{avec } K = \frac{\beta}{N} S_\infty I_\infty > 0$$

Si la dérivée tend vers une constante négative $-K$, cela signifie géométriquement que la courbe $S(t)$ finit par se comporter comme une droite de pente négative ($S(t) \approx -Kt$). Elle plongerait inévitablement vers $-\infty$, ce qui est impossible car une population doit

rester positive ($S(t) \geq 0$).

Conclusion : L'hypothèse est fausse. On conclut donc par l'absurde :

$$\lim_{t \rightarrow +\infty} I(t) = 0$$

□

Enjeu hospitalier "aplatir la courbe" : Le pic épidémique est le moment critique. Si la valeur maximale $I(t_{pic})$ dépasse la capacité d'accueil des hôpitaux (nombre de lits), le système de santé sature. Les stratégies de "freinage" visent non pas à annuler l'épidémie, mais à étaler ce pic dans le temps pour qu'il reste sous le seuil de saturation hospitalière.

2.4 Visualisation des trajectoires : L'épidémie dans l'espace des phases

Au lieu d'étudier l'évolution temporelle, il est pertinent d'analyser la relation directe entre les infectés I et les sains S . Cela revient à éliminer la variable temps t .

Proposition 3 (Équation des trajectoires). *Dans le plan de phase (S, I) , les trajectoires solutions suivent la courbe d'équation :*

$$I(S) = -S + \frac{N}{R_0} \ln(S) + K \quad (2)$$

où K est une constante déterminée par les conditions initiales.

Démonstration. Pour obtenir la relation entre I et S , nous divisons membre à membre les équations différentielles :

$$\frac{dI}{dS} = \frac{dI/dt}{dS/dt} = \frac{\frac{\beta}{N}SI - \gamma I}{-\frac{\beta}{N}SI}$$

Pour $I \neq 0$, on peut simplifier par I :

$$\frac{dI}{dS} = \frac{I(\frac{\beta S}{N} - \gamma)}{I(-\frac{\beta S}{N})} = \frac{\frac{\beta S}{N} - \gamma}{-\frac{\beta S}{N}}$$

Séparons la fraction :

$$\frac{dI}{dS} = \frac{\frac{\beta S}{N}}{-\frac{\beta S}{N}} + \frac{-\gamma}{-\frac{\beta S}{N}} = -1 + \frac{\gamma N}{\beta S}$$

On reconnaît l'inverse de $R_0 = \beta/\gamma$:

$$\frac{dI}{dS} = -1 + \frac{N}{R_0 S}$$

Il s'agit d'une équation différentielle élémentaire. Intégrons par rapport à la variable S :

$$I(S) = \int \left(-1 + \frac{N}{R_0 S} \right) dS$$

$$I(S) = -S + \frac{N}{R_0} \ln(S) + K$$

Où K est une constante d'intégration fixée par l'état initial (S_0, I_0) .

□

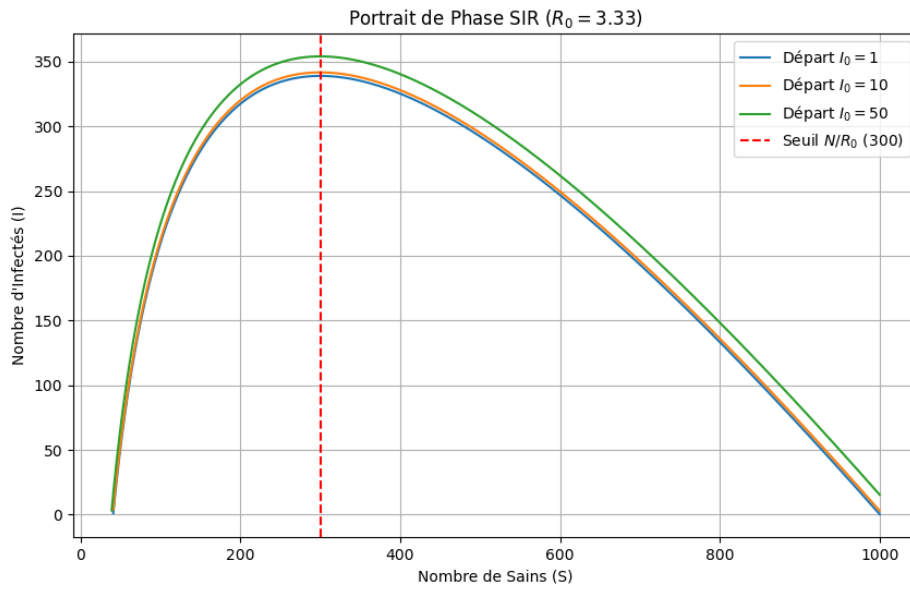


FIGURE 2 – **Portrait de Phase théorique (I en fonction de S).**

Paramètres de simulation : Population $N = 1000$, taux de contact $\beta = 0.5$, taux de guérison $\gamma = 0.15$ (soit $R_0 \approx 3,33$).

Lecture : Ce graphique se lit de la droite vers la gauche, car le nombre de sains $S(t)$ diminue au cours du temps. Les différentes courbes de couleur (bleue, orange, verte) correspondent à des introductions d'infectés initiaux variés ($I_0 = 1, 10, 50$).

Conclusion majeure : Bien que l'intensité maximale de l'épidémie (hauteur du pic) varie légèrement selon le point de départ, le pic se produit systématiquement à la même abscisse critique $S_c = N/R_0 \approx 300$ (ligne rouge). Cela démontre que le seuil d'immunité collective est une propriété structurelle du système, indépendante des conditions initiales.

2.5 Pr vision   long terme : L'immunit  collective naturelle

Une question fondamentale est de savoir combien d'individus auront  t  infect s   la fin de l' pid mie (quand $t \rightarrow +\infty$).

Proposition 4 (Taille finale de l' pid mie). *L' pid mie s' teint toujours ($\lim_{t \rightarrow \infty} I(t) = 0$), mais ne touche pas toute la population ($S_\infty > 0$). La valeur finale S_∞ est l'unique solution dans $]0, N/R_0[$ de l' quation transcendante :*

$$\ln\left(\frac{S_0}{S_\infty}\right) = R_0\left(1 - \frac{S_\infty}{N}\right) \quad (3)$$

(En supposant $I_0 \approx 0$ et $S_0 \approx N$).

D monstration. Partons de la premi re  quation du syst me : $S' = -\frac{\beta}{N}SI$. Divisons par S (strictement positif) pour faire appara tre une d riv e logarithmique :

$$\frac{S'(t)}{S(t)} = -\frac{\beta}{N}I(t)$$

Int grons cette relation sur la dur e totale de l' pid mie, c'est- -dire entre $t = 0$ et $t = +\infty$:

$$\int_0^{+\infty} \frac{S'(t)}{S(t)} dt = -\frac{\beta}{N} \int_0^{+\infty} I(t) dt$$

Membre de gauche : C'est une primitive usuelle :

$$\int_0^{+\infty} \frac{S'(t)}{S(t)} dt = [\ln S(t)]_0^{+\infty} = \ln(S_\infty) - \ln(S_0)$$

Membre de droite : Pour calculer l'int grale de $I(t)$, utilisons la troisi me  quation du syst me : $R'(t) = \gamma I(t)$, ce qui implique $I(t) = \frac{1}{\gamma} R'(t)$.

$$-\frac{\beta}{N} \int_0^{+\infty} I(t) dt = -\frac{\beta}{N} \int_0^{+\infty} \frac{1}{\gamma} R'(t) dt = -\frac{\beta}{N\gamma} [R(t)]_0^{+\infty}$$

On reconna t $R_0 = \beta/\gamma$. De plus, on sait que $R(0) = 0$ et que $R_\infty = N - S_\infty$ (car $I_\infty = 0$).

$$-\frac{\beta}{N} \int_0^{+\infty} I(t) dt = -\frac{R_0}{N} (R_\infty - 0) = -\frac{R_0}{N} (N - S_\infty) = -R_0 \left(1 - \frac{S_\infty}{N}\right)$$

Conclusion : En  galisant les deux membres :

$$\ln(S_\infty) - \ln(S_0) = -R_0 \left(1 - \frac{S_\infty}{N}\right)$$

En multipliant par -1 pour inverser le logarithme ($\ln(A) - \ln(B) = -\ln(B/A)$), on obtient la formule annonc e :

$$\ln\left(\frac{S_0}{S_\infty}\right) = R_0 \left(1 - \frac{S_\infty}{N}\right)$$

□

Interprétation (Calcul de la taille finale de l'épidémie) :

Cette formule est implicite (on ne peut pas isoler S_∞). Posons p la proportion totale de la population **ayant été infectée au cours de toute l'épidémie** (c'est ce que l'on appelle la *taille finale de l'épidémie*).

Puisqu'à la fin de l'épidémie il n'y a plus de malades ($I_\infty = 0$), tous les individus ayant contracté le virus se trouvent dans le compartiment des rétablis.

On a donc $p = \frac{R_\infty}{N} = \frac{N-S_\infty}{N}$. Partons de l'équation (3) et appliquons l'hypothèse selon laquelle l'immense majorité de la population est saine au déclenchement de l'épidémie ($S_0 \approx N$). L'équation s'écrit :

$$\ln\left(\frac{N}{S_\infty}\right) = R_0\left(1 - \frac{S_\infty}{N}\right)$$

Exprimons les termes de cette équation en fonction de p . Par définition, on a $p = \frac{N-S_\infty}{N}$, ce qui peut se réécrire :

$$p = 1 - \frac{S_\infty}{N}$$

On en déduit immédiatement que le membre de droite contient p . Pour le membre de gauche, on remarque que $\frac{S_\infty}{N} = 1 - p$, d'où, par passage à l'inverse, $\frac{N}{S_\infty} = \frac{1}{1-p}$.

En substituant ces expressions dans notre équation initiale, on obtient la suite d'équivalences suivante :

$$\begin{aligned}\ln\left(\frac{N}{S_\infty}\right) = R_0\left(1 - \frac{S_\infty}{N}\right) &\iff \ln\left(\frac{1}{1-p}\right) = R_0p \\ &\iff -\ln(1-p) = R_0p \quad (\text{car } \ln(1/x) = -\ln(x)) \\ &\iff \ln(1-p) = -R_0p \\ &\iff e^{\ln(1-p)} = e^{-R_0p} \quad (\text{par composition avec l'exponentielle}) \\ &\iff 1-p = e^{-R_0p}\end{aligned}$$

Cette relation permet de calculer numériquement l'ampleur totale de la crise sanitaire :

- Pour $R_0 = 1.5$ (ex : Grippe), on trouve $p \approx 58\%$.
- Pour $R_0 = 3$ (ex : COVID-19 initial), on trouve $p \approx 94\%$.
- *Remarque* : La solution p de cette équation est une fonction strictement croissante de R_0 . Pour un virus extrêmement contagieux comme la rougeole ($R_0 \approx 15$), la proportion de personnes touchées dépassera 99,99% si aucune vaccination n'est mise en place.

Cela montre que dans un modèle classique, l'épidémie s'arrête **naturellement avant** d'avoir touché 100% de la population. Lorsque la proportion d'immunisés est suffisante, le virus ne trouve plus assez de cibles (sains) pour se maintenir.

2.6 Justification rigoureuse : Stabilité du point d'équilibre

Pour valider rigoureusement le seuil R_0 , nous étudions la stabilité des points d'équilibre du système linéarisé.

Points d'équilibre : Les points où $S' = I' = R' = 0$. Le point principal est l'équilibre "sans maladie" (Disease Free Equilibrium - DFE) : $E_0 = (N, 0)$.

Théorème 3 (Stabilité du DFE). *Le point d'équilibre E_0 est :*

- **Localement asymptotiquement stable** si $R_0 < 1$.
- **Instable** si $R_0 > 1$.

Démonstration. Considérons le sous-système réduit (S, I) . La matrice Jacobienne J est définie par les dérivées partielles :

$$J(S, I) = \begin{pmatrix} \frac{\partial S'}{\partial S} & \frac{\partial S'}{\partial I} \\ \frac{\partial I'}{\partial S} & \frac{\partial I'}{\partial I} \end{pmatrix}$$

Calculons les termes :

$$\begin{aligned} \text{— } \frac{\partial S'}{\partial S} &= -\frac{\beta I}{N} \text{ et } \frac{\partial S'}{\partial I} = -\frac{\beta S}{N} \\ \text{— } \frac{\partial I'}{\partial S} &= \frac{\beta I}{N} \text{ et } \frac{\partial I'}{\partial I} = \frac{\beta S}{N} - \gamma \end{aligned}$$

La matrice Jacobienne générale est donc :

$$J(S, I) = \begin{pmatrix} -\frac{\beta I}{N} & -\frac{\beta S}{N} \\ \frac{\beta I}{N} & \frac{\beta S}{N} - \gamma \end{pmatrix}$$

Évaluons cette matrice au point d'équilibre $E_0 = (N, 0)$ (en remplaçant S par N et I par 0) :

$$J(E_0) = \begin{pmatrix} 0 & -\beta \\ 0 & \beta - \gamma \end{pmatrix}$$

C'est une matrice triangulaire supérieure. Ses valeurs propres sont donc ses coefficients diagonaux :

$$\lambda_1 = 0 \quad \text{et} \quad \lambda_2 = \beta - \gamma$$

La stabilité du système dépend du signe de la partie réelle de λ_2 .

$$\lambda_2 < 0 \iff \beta - \gamma < 0 \iff \frac{\beta}{\gamma} < 1 \iff R_0 < 1$$

Ainsi, si $R_0 < 1$, toute petite perturbation (introduction d'un malade) s'amortit (retour à l'équilibre). Si $R_0 > 1$, la perturbation croît exponentiellement (instabilité). $\square$

2.7 Le modèle SIR ouvert : Démographie et équilibre endémique

Le modèle SIR classique, étudié jusqu'ici, suppose une population fermée. Par conséquent, la maladie finit toujours par disparaître par épuisement des cibles. Pour comprendre comment un virus peut s'installer durablement dans une population (comme la grippe ou le rhume), il faut introduire un renouvellement démographique.

Définition 5 (Système SIR ouvert). *Considérons un flux d'entrée de nouveaux individus (naissances) et un flux de sortie (décès naturel) dictés par un même taux $\mu > 0$, maintenant la population totale N constante. Le système devient :*

$$\begin{cases} S'(t) = \mu N - \frac{\beta}{N} S(t) I(t) - \mu S(t) \\ I'(t) = \frac{\beta}{N} S(t) I(t) - \gamma I(t) - \mu I(t) \\ R'(t) = \gamma I(t) - \mu R(t) \end{cases} \quad (4)$$

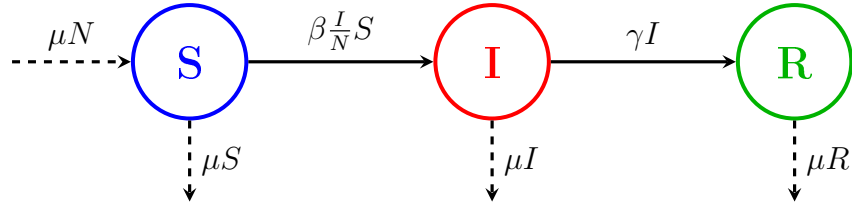


FIGURE 3 – **Schéma compartimental du modèle SIR ouvert.** Les flèches pleines représentent la dynamique infectieuse classique, tandis que les flèches en pointillés modélisent le flux démographique (ici avec un taux de renouvellement constant μ , c'est à dire un taux de natalité et de mortalité égal).

L'ajout de la mortalité naturelle pour les infectés (μI) réduit la durée moyenne pendant laquelle un individu est contagieux. Le nouveau taux de reproduction de base s'écrit alors :

$$R_0 = \frac{\beta}{\gamma + \mu}$$

Recherche des points d'équilibre :

Un point d'équilibre est atteint lorsque les dérivées s'annulent ($S' = I' = R' = 0$). L'équation des infectés donne :

$$I \left(\frac{\beta}{N} S - (\gamma + \mu) \right) = 0$$

Ce produit nul conduit à deux situations biologiquement très différentes :

1) L'équilibre trivial (DFE - Disease Free Equilibrium) :

Si $I = 0$, l'équation de S' donne $\mu N - \mu S = 0 \implies S = N$. On retrouve l'équilibre sans maladie $E_0 = (N, 0)$.

2) L'équilibre endémique (EE) :

Si $I \neq 0$, alors le second facteur s'annule :

$$\frac{\beta}{N} S_{ee} - (\gamma + \mu) = 0 \implies S_{ee} = \frac{\gamma + \mu}{\beta} N$$

En remarquant que $\frac{\gamma+\mu}{\beta} = \frac{1}{R_0}$, on obtient l'abscisse de l'équilibre endémique :

$$S_{ee} = \frac{N}{R_0}$$

En réinjectant S_{ee} dans l'équation de $S' = 0$, on trouve le nombre de malades à l'équilibre :

$$\mu N - \frac{\beta}{N} \left(\frac{N}{R_0} \right) I_{ee} - \mu \left(\frac{N}{R_0} \right) = 0 \implies I_{ee} = \frac{\mu N}{\beta} (R_0 - 1)$$

Proposition 5 (Condition d'existence de l'endémie). *Pour que l'équilibre endémique ait un sens biologique, il faut un nombre de malades strictement positif ($I_{ee} > 0$). Puisque $\frac{\mu N}{\beta} > 0$, on retrouve l'importance capitale du taux de reproduction : **l'équilibre endémique n'existe que si $R_0 > 1$.***

Contrairement au modèle fermé où les trajectoires s'écrasent sur l'axe des abscisses ($I \rightarrow 0$), l'ajout d'un renouvellement démographique ($\mu > 0$) permet à la maladie de se maintenir.

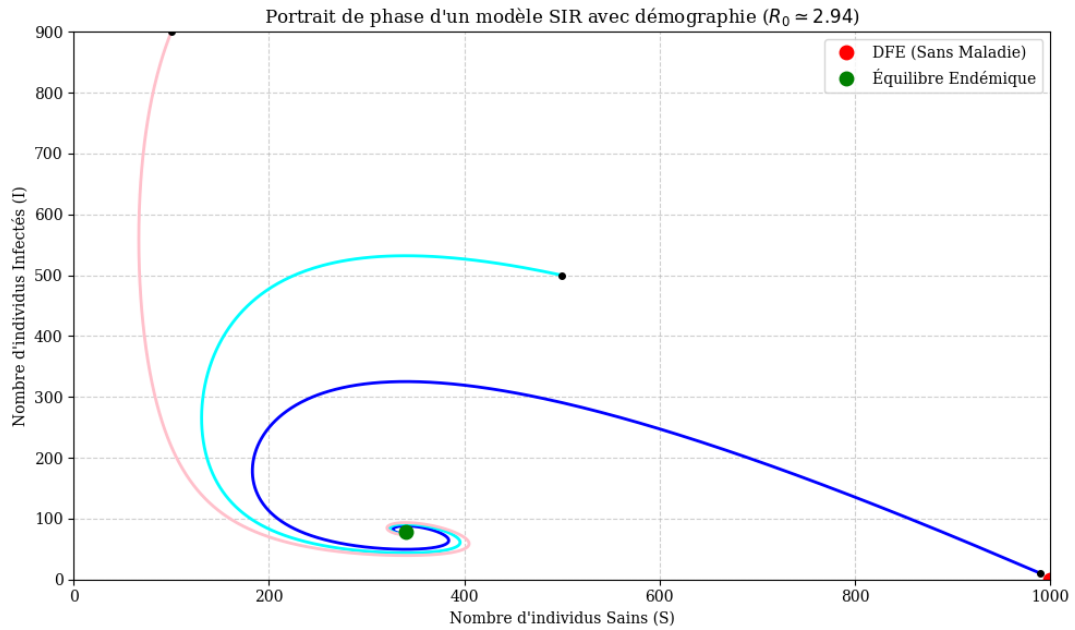


FIGURE 4 – **Portrait de phase d'un modèle SIR avec démographie (ouvert).**

Ce graphique illustre l'évolution du nombre d'infectés (I) en fonction du nombre d'individus sains (S). Les trois trajectoires représentent des scénarios initiaux radicalement différents :

- **Courbe bleue** : Le début d'une épidémie standard ($S_0 = 990$, $I_0 = 10$).
- **Courbe cyan** : Une situation où la moitié de la population est déjà infectée ($S_0 = 500$, $I_0 = 500$).
- **Courbe rose** : Un scénario extrême où la quasi-totalité de la population est initialement malade ($S_0 = 100$, $I_0 = 900$).

Le système présente un comportement remarquable : indépendamment du point de départ, toutes les trajectoires s'enroulent en spirale (traduisant des vagues épidémiques

successives et amorties) pour converger vers un attracteur unique : **l'équilibre endémique (point vert)**. Cela démontre géométriquement que la gravité de l'état final d'une maladie endémique ne dépend pas du nombre de patients zéros, mais uniquement des caractéristiques intrinsèques du virus et de la démographie (les paramètres β , γ et μ). Par ailleurs, l'équilibre sans maladie (DFE, point rouge) est ici un point instable que le système fuit naturellement puisque $R_0 > 1$.

2.8 Synthèse : Portée et limites des modèles déterministes

L'étude mathématique du modèle SIR nous a permis d'établir les lois fondamentales qui régissent une épidémie en population fermée. Trois résultats majeurs se dégagent de cette première approche :

1. **La dichotomie du seuil R_0** : Nous avons démontré que le déclenchement d'une épidémie n'est pas automatique. Il dépend strictement de la condition $R_0 > 1$ (instabilité du point d'équilibre). Ce nombre sans dimension est donc la clé de voûte de la prévision sanitaire.
2. **L'auto-limitation du processus** : Contrairement à une croissance exponentielle infinie, l'épidémie finit toujours par s'éteindre d'elle-même ($\lim_{t \rightarrow \infty} I(t) = 0$). Nous avons prouvé que cette extinction survient alors qu'il reste encore des individus sains ($S_\infty > 0$), illustrant le principe d'immunité collective naturelle.
3. **La prévisibilité du pic** : L'analyse du portrait de phase a montré que le maximum d'infectés est atteint précisément lorsque le réservoir de sains franchit le seuil critique $S_c = N/R_0$.

Bien que robuste mathématiquement, ce modèle repose sur l'hypothèse forte d'un *mélange homogène* (chaque individu a la même probabilité de rencontrer n'importe qui) et de paramètres β et γ constants dans le temps. Dans la réalité, les interactions sont complexes et la dynamique peut être modifiée par plusieurs facteurs biologiques ou démographiques qu'il est possible d'intégrer en enrichissant le modèle :

- **La période de latence (Modèle SEIR)** : Le modèle SIR suppose qu'un individu devient contagieux à l'instant même de son infection. En réalité, il existe souvent une période d'incubation. L'ajout d'un compartiment E (Exposés) permet de retarder le pic épidémique et de lisser la dynamique.

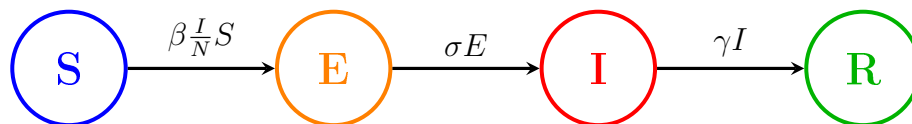


FIGURE 5 – **Schéma du modèle SEIR.** L'ajout du compartiment E (Exposés) modélise la période d'incubation. Les individus sont infectés mais pas encore contagieux (taux de passage σ).

- **La perte d'immunité (Modèle SIRS) :** Nous avons supposé une immunité définitive. Si l'immunité s'estompe avec le temps, les rétablis R redeviennent des cibles potentielles S , ce qui peut provoquer des vagues épidémiques successives.

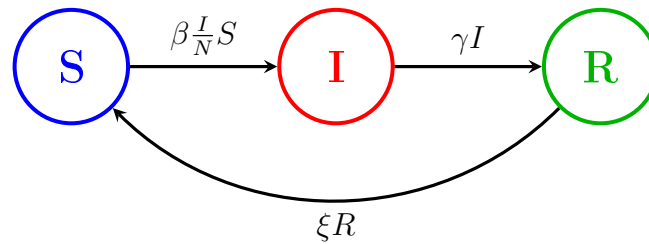


FIGURE 6 – **Schéma du modèle SIRS.** Une immunité temporaire est modélisée par un retour des individus Rétablis vers le compartiment des Sains à un taux ξ .

- **La combinaison des deux (Modèle SEIRS) :** Pour de nombreux virus respiratoires (comme le SARS-CoV-2 ou la grippe), on observe à la fois une phase d'incubation et une immunité temporaire. Le modèle SEIRS fusionne ces deux aspects et décrit avec une grande fidélité les cycles d'infections récurrentes, modélisant ainsi l'apparition de vagues saisonnières.

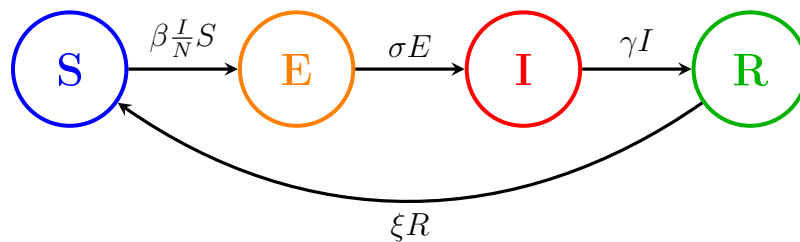


FIGURE 7 – **Schéma du modèle SEIRS.** Il combine la période de latence (σ) et la perte d'immunité (ξ).

L'approche déterministe est très riche, mais elle atteint une de ses limites majeures lorsqu'on étudie les petits foyers infectieux : elle ignore l'extinction aléatoire due au hasard des premières rencontres. L'étude de ce phénomène, par le prisme des probabilités, fera l'objet de la partie suivante.

3 Partie II : L'approche stochastique par les chaînes de Markov

3.1 Limites du continu et nécessité d'une approche discrète

Dans la première partie, le modèle déterministe a supposé que la population était suffisamment grande pour être assimilée à un fluide continu. Les variables $S(t)$, $I(t)$ et $R(t)$ étaient traitées comme des fonctions réelles continues et dérivables.

Cependant, la réalité biologique est fondamentalement discrète (on ne peut pas avoir 0,5 individu infecté) et aléatoire. Au tout début d'une épidémie, lorsqu'il n'y a qu'un seul "patient zéro" ($I_0 = 1$), deux scénarios s'affrontent :

- Soit le patient rencontre une personne saine et transmet le virus.
- Soit le patient guérit (ou s'isole) avant d'avoir pu contaminer qui que ce soit.

Dans le second cas, l'épidémie s'éteint instantanément par pur hasard. Le modèle déterministe, régi par $I'(0) > 0$ dès lors que $R_0 > 1$, est incapable de capturer cette probabilité d'extinction prématurée. Pour pallier ce défaut et étudier les petits foyers d'infection, nous devons abandonner les équations différentielles au profit d'une modélisation probabiliste. Le choix spécifique des **chaînes de Markov à temps continu** s'impose ici naturellement face à d'autres processus stochastiques. En effet, la dynamique épidémique locale vérifie intrinsèquement la propriété biologique "d'absence de mémoire" : la probabilité qu'une nouvelle infection ou qu'une guérison se produise à l'instant suivant dépend uniquement de l'état présent du système (le nombre exact d'individus sains et infectés à l'instant t), et absolument pas de l'historique ou du chemin parcouru par l'épidémie pour y arriver. Les chaînes de Markov constituent l'outil mathématique parfait pour traduire cette réalité.

3.2 Définitions mathématiques : Chaînes de Markov à temps continu

Pour modéliser l'aspect aléatoire des rencontres et des guérisons, nous utilisons le formalisme des processus stochastiques. Contrairement aux chaînes de Markov à temps discret (où les événements se produisent à des étapes $n = 0, 1, 2, \dots$), l'épidémie évolue en temps continu.

Définition 6 (Processus stochastique et espace d'états). *On appelle chaîne de Markov à temps continu un processus stochastique $(X(t))_{t \geq 0}$ prenant ses valeurs dans un espace d'états E discret (fini ou dénombrable). À chaque instant t , la variable aléatoire $X(t)$ représente l'état du système.*

La caractéristique fondamentale de ce processus est l'absence de mémoire, signifiant que le futur du système ne dépend que de son état présent, et non du chemin parcouru pour y arriver.

Définition 7 (Propriété de Markov en temps continu). *Le processus $(X(t))_{t \geq 0}$ vérifie la propriété de Markov si, pour tout nombre entier n , pour toute suite d'instants $0 \leq t_1 < t_2 < \dots < t_n < t_{n+1}$ et pour tous états $x_1, x_2, \dots, x_{n+1} \in E$, on a :*

$$\mathbb{P}\left(X(t_{n+1}) = x_{n+1} \mid X(t_1) = x_1, \dots, X(t_n) = x_n\right) = \mathbb{P}\left(X(t_{n+1}) = x_{n+1} \mid X(t_n) = x_n\right)$$

On supposera de plus que le processus est **homogène**, c'est-à-dire que cette probabilité de transition ne dépend que du délai $\Delta t = t_{n+1} - t_n$ et non de l'instant de départ t_n .

Plutôt que de manipuler des matrices de transition globales comme en temps discret, la dynamique en temps continu est caractérisée par des taux de saut instantanés entre les états.

Définition 8 (Taux de transition et générateur infinitésimal). *Pour deux états distincts i et j de E , on définit le taux de transition $q_{i,j}$ (ou propension de saut) de l'état i vers l'état j par la limite :*

$$q_{i,j} = \lim_{\Delta t \rightarrow 0^+} \frac{\mathbb{P}(X(t + \Delta t) = j \mid X(t) = i)}{\Delta t}$$

Ce qui se traduit par le développement asymptotique suivant pour un Δt très petit :

$$\mathbb{P}(X(t + \Delta t) = j \mid X(t) = i) = q_{i,j}\Delta t + o(\Delta t)$$

Cette définition infinitésimale permet de déduire une propriété cruciale pour la simulation informatique du processus : le temps passé dans un état donné avant de "sauter" vers un autre état est une variable aléatoire qui suit une loi bien précise.

Proposition 6 (Temps de séjour et loi exponentielle). *Soit le système dans l'état i à l'instant t . Notons T_i le temps de séjour (le temps qu'il reste dans cet état avant le prochain événement). Grâce à la propriété d'absence de mémoire de la chaîne de Markov, la variable aléatoire T_i suit nécessairement une **loi exponentielle** de paramètre q_i , où q_i est la somme de tous les taux de sortie possibles depuis l'état i :*

$$q_i = \sum_{j \neq i} q_{i,j}$$

L'espérance du temps passé dans l'état i est donc $\mathbb{E}[T_i] = \frac{1}{q_i}$.

Ce socle théorique étant posé, nous pouvons maintenant l'appliquer rigoureusement à la dynamique épidémique du modèle SIR.

3.3 Le cadre mathématique : Processus de saut de Markov

Définition 9 (Espace d'état discret). *À tout instant $t \geq 0$, l'état du système n'est plus un point dans un espace continu, mais un vecteur de variables aléatoires à valeurs entières :*

$$X(t) = (S(t), I(t)) \in \mathbb{N}^2$$

Puisque la population totale reste constante ($N = S(t) + I(t) + R(t)$), la connaissance des variables $S(t)$ et $I(t)$ suffit pour caractériser entièrement le système.

L'évolution du système n'est plus décrite par des trajectoires lisses, mais par une suite de "sauts" aléatoires. L'hypothèse fondamentale de notre modèle est la **propriété de Markov** : l'évolution future du système dépend uniquement de son état présent, et non de son passé (le processus est dit "sans mémoire").

Définition 10 (Probabilités de transition). *Considérons le système dans un état (s, i) à l'instant t . Pendant un intervalle de temps infinitésimal Δt , l'évolution du système présente trois issues possibles et mutuellement exclusives : soit la survenue de l'un des deux événements épidémiologiques (une infection ou une guérison), soit l'absence de tout événement.*

1. Une nouvelle infection : *L'état passe de (s, i) à $(s - 1, i + 1)$.*

La probabilité que cet événement se produise dans l'intervalle $[t, t + \Delta t[$ est proportionnelle au nombre de rencontres possibles entre les s sains et les i infectés :

$$\mathbb{P}(X(t + \Delta t) = (s - 1, i + 1) \mid X(t) = (s, i)) = \frac{\beta}{N} si \Delta t + o(\Delta t)$$

2. Une guérison : *L'état passe de (s, i) à $(s, i - 1)$.*

La probabilité qu'un des i individus infectés guérisse est proportionnelle au nombre de malades :

$$\mathbb{P}(X(t + \Delta t) = (s, i - 1) \mid X(t) = (s, i)) = \gamma i \Delta t + o(\Delta t)$$

3. Aucun événement : *L'état reste (s, i) .*

La probabilité que rien ne se passe est le complémentaire à 1 des deux événements précédents :

$$\mathbb{P}(X(t + \Delta t) = (s, i) \mid X(t) = (s, i)) = 1 - \left(\frac{\beta}{N} si + \gamma i \right) \Delta t + o(\Delta t)$$

Contrairement au modèle déterministe qui impose une vitesse de variation stricte, ces probabilités (souvent appelées "propensions") indiquent seulement la *tendance* du système à évoluer. C'est cette dimension probabiliste qui va nous permettre de simuler l'aléatoire inhérent aux petites populations.

3.4 L'équation maîtresse (Équation de Kolmogorov forward)

Puisque la variable $X(t) = (S(t), I(t))$ est aléatoire, nous ne pouvons plus suivre sa trajectoire exacte. L'objectif analytique est alors de déterminer la distribution de probabilité du système à tout instant t .

Notons $p_{(s,i)}(t) = \mathbb{P}(S(t) = s, I(t) = i)$ la probabilité que le système contienne exactement s sains et i infectés à l'instant t . En utilisant la formule des probabilités totales sur l'intervalle $[t, t + \Delta t]$, on peut déduire une équation différentielle appelée **Équation maîtresse** (ou équation de Kolmogorov *forward*) :

$$\begin{aligned} \frac{dp_{(s,i)}(t)}{dt} = & \underbrace{\frac{\beta}{N}(s+1)(i-1)p_{(s+1,i-1)}(t)}_{\text{Transition : une infection vient de se produire}} \\ & + \underbrace{\gamma(i+1)p_{(s,i+1)}(t)}_{\text{Transition : une guérison vient de se produire}} \\ & - \underbrace{\left(\frac{\beta}{N} si + \gamma i \right) p_{(s,i)}(t)}_{\text{Transition : le système quitte l'état } (s,i)} \end{aligned} \quad (5)$$

Cette équation différentielle linéaire décrit la dynamique des probabilités. Cependant, l'espace des états possède $\frac{(N+1)(N+2)}{2}$ éléments. Pour $N = 1000$, cela représente plus de 500 000 équations différentielles couplées ! Sa résolution analytique exacte est donc impossible en pratique, ce qui justifie le recours aux simulations numériques ou à des approximations asymptotiques.

3.5 Outil mathématique : Le processus de Galton-Watson

Pour justifier rigoureusement les calculs liés à l'extinction du virus lors des premiers instants de l'épidémie, il est nécessaire d'introduire le modèle fondamental de la famille des processus de branchement : le processus de Galton-Watson.

Définition 11 (Processus de Galton-Watson). *Un processus de Galton-Watson modélise l'évolution de la taille d'une population au fil des générations (à temps discret). On note Z_n le nombre d'individus à la génération n . À chaque génération, chaque individu i donne naissance à un nombre aléatoire de "descendants" (ou de nouvelles infections) X_i , de manière mutuellement indépendante et selon une même loi de probabilité. Ainsi, la taille de la génération $n + 1$ s'écrit :*

$$Z_{n+1} = \sum_{i=1}^{Z_n} X_i$$

L'étude de l'extinction de la population (c'est-à-dire l'événement $\{\exists n \in \mathbb{N}, Z_n = 0\}$) se fait en conditionnant par le nombre de descendants du tout premier individu.

Théorème 4 (Probabilité d'extinction). *Soit q la probabilité d'extinction d'une lignée issue d'un unique "patient zéro" ($Z_0 = 1$). Si l'on note $p_k = \mathbb{P}(X_1 = k)$ la probabilité qu'un individu ait exactement k descendants directs, alors, par la formule des probabilités totales, q vérifie l'équation :*

$$q = \sum_{k=0}^{+\infty} p_k q^k$$

*La théorie des processus de branchement démontre que la probabilité d'extinction exacte de la lignée correspond à la **plus petite racine dans l'intervalle** $[0, 1]$ de cette équation polynomiale.*

Remarque 1 (Justification de la plus petite racine). *Le fait que la probabilité d'extinction soit la plus petite racine de l'équation $q = f(q)$ (où $f(x) = \sum p_k x^k$) provient de l'étude temporelle du processus. Si l'on note $q_n = \mathbb{P}(Z_n = 0)$ la probabilité que l'extinction ait eu lieu à la génération n , on a $q_0 = 0$ (le patient zéro est présent). La suite des probabilités d'extinction vérifie la relation de récurrence $q_{n+1} = f(q_n)$. La fonction f étant croissante sur $[0, 1]$ et $q_0 \leq q_1$, la suite (q_n) est croissante. Étant majorée par 1, elle converge vers une limite q vérifiant $q = f(q)$. Puisque la suite démarre de 0 et croît, elle converge nécessairement vers le premier point fixe rencontré, c'est-à-dire la plus petite racine positive de l'équation.*

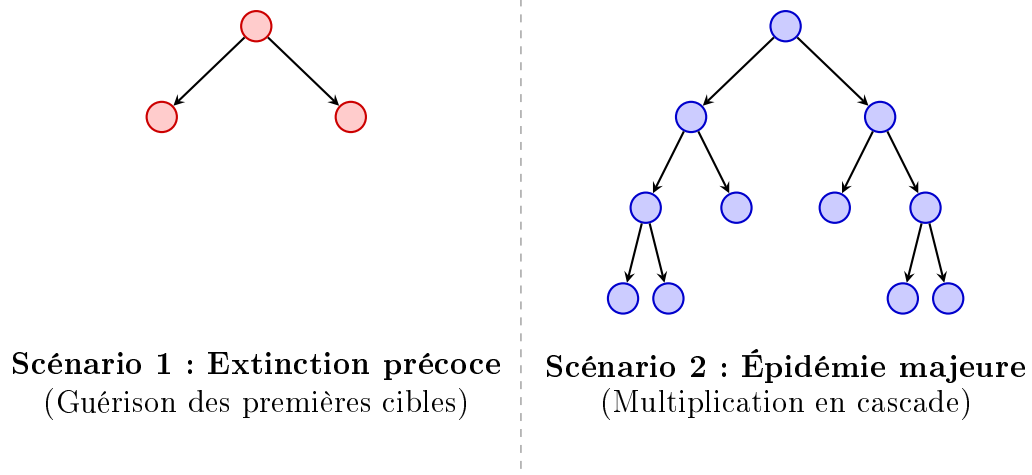


FIGURE 8 – **Visualisation du processus de Galton-Watson ($1 \rightarrow 0$ ou 2 descendants).**

Le processus stochastique génère deux devenirs possibles. À gauche, la chaîne de transmission s'interrompt rapidement par un concours de circonstances aléatoires (les deux premiers malades guérissent avant de transmettre). À droite, les infections s'enchaînent plus vite que les guérisons, provoquant une croissance exponentielle incontrôlable de l'arbre épidémique.

C'est ce théorème puissant que nous allons adapter à notre dynamique en temps continu pour modéliser le devenir du foyer épidémique initial.

3.6 Le résultat fondamental : La probabilité d'extinction précoce

L'intérêt majeur du modèle stochastique est de pouvoir calculer la probabilité qu'une épidémie s'éteigne d'elle-même au tout début, malgré un $R_0 > 1$.

Au début de l'épidémie, si la population totale N est très grande par rapport aux quelques premiers cas, on peut faire l'approximation $S(t) \approx N$. Les taux de transition se simplifient considérablement :

- Taux d'infection : $q_{(s,i) \rightarrow (s-1,i+1)} = \frac{\beta}{N} si \approx \beta i$
- Taux de guérison : $q_{(s,i) \rightarrow (s,i-1)} = \gamma i$

Sous cette approximation, la dynamique des infectés $I(t)$ devient indépendante des sains et se modélise comme un processus de naissance et de mort. Chaque individu infecté "vit" un temps exponentiel de paramètre γ et infecte une nouvelle cible à un taux β .

Théorème 5 (Probabilité d'extinction). *Soit un foyer infectieux initialement composé de I_0 malades dans une grande population saine. Si $R_0 > 1$, l'épidémie peut s'éteindre de manière aléatoire avant d'avoir pu créer une flambée, avec une probabilité exacte :*

$$\mathbb{P}_{ext} = \left(\frac{1}{R_0} \right)^{I_0}$$

Démonstration. Considérons l'introduction d'un unique "patient zéro" ($I_0 = 1$). Notons $q \in [0, 1]$ la probabilité que la lignée d'infections issue de cet individu s'éteigne définitivement.

Bien que notre modèle soit à temps continu, la loi exponentielle garantit que les événements se produisent strictement un par un (la probabilité de deux événements simultanés est nulle). Ainsi, l'observation du système au moment exact du premier événement se ramène formellement à une bifurcation discrète à deux issues.

À cet instant initial, puisqu'il y a exactement un seul malade ($I_0 = 1$), les taux de transition valent respectivement β pour une nouvelle infection et γ pour une guérison. D'après la Proposition 6, le taux de sortie global de cet état est la somme de ces propensions, soit $\beta + \gamma$. Le temps d'attente avant le tout premier événement suit donc une loi exponentielle de paramètre $\beta + \gamma$, la probabilité que cet événement soit une guérison est $\frac{\gamma}{\beta + \gamma}$, et la probabilité que ce soit une nouvelle infection est $\frac{\beta}{\beta + \gamma}$.

À partir de ce premier "saut", le patient zéro agit formellement comme l'ancêtre d'un processus de Galton-Watson discret, avec deux nombres de "descendants" (cibles infectées présentes dans le système) possibles :

1. **Il guérit avant de transmettre la maladie (0 descendant) :** La probabilité d'avoir $k = 0$ enfant est $p_0 = \frac{\gamma}{\beta + \gamma}$.
2. **Il transmet la maladie avant de guérir (2 descendants dans le système) :** Il y a maintenant lui-même (qui est toujours malade) et la nouvelle cible. Le système se sépare en **2** lignées indépendantes. La probabilité d'avoir $k = 2$ enfants est $p_2 = \frac{\beta}{\beta + \gamma}$.

D'après le théorème du processus de Galton-Watson, la probabilité d'extinction q vérifie l'équation $q = p_0 q^0 + p_2 q^2$, soit :

$$q = \frac{\gamma}{\beta + \gamma} \times 1 + \frac{\beta}{\beta + \gamma} \times q^2$$

Multiplions par $(\beta + \gamma)$ et réorganisons :

$$\beta q^2 - (\beta + \gamma)q + \gamma = 0$$

Cette équation du second degré admet comme racine évidente $q_1 = 1$. Le produit des racines valant $\frac{\gamma}{\beta}$, la seconde racine est $q_2 = \frac{\gamma}{\beta}$.

D'après la théorie de Galton-Watson, la probabilité d'extinction est la plus petite racine dans $[0, 1]$. Or, $\frac{\gamma}{\beta} = \frac{1}{R_0}$. Ainsi, si $R_0 > 1$, la probabilité d'extinction issue d'un patient est bien $q = \frac{1}{R_0}$.

Par indépendance, si l'on commence avec I_0 patients zéros, l'épidémie s'éteint si et seulement si toutes les I_0 chaînes s'éteignent, ce qui a une probabilité de :

$$\mathbb{P}_{\text{ext}} = q^{I_0} = \left(\frac{1}{R_0} \right)^{I_0}$$

□

Remarque 2 (Sur le nombre de "descendants"). *Il peut paraître contre-intuitif que le patient zéro ne puisse avoir que 0 ou 2 "descendants" dans notre équation (p_0 et p_2), alors qu'en réalité, un malade (si $R_0 = 3$) infectera en moyenne 3 personnes au cours de sa période contagieuse.*

*Cette apparente contradiction s'explique par la manière dont on relie le temps continu à la notion de "génération". Dans notre processus de Markov, les probabilités de transition ne permettent qu'une seule contamination à la fois (il est impossible d'avoir des infections simultanées ou des "jumeaux"). L'équation conditionne donc la dynamique sur ce **tout premier événement** générationnel.*

*Lorsqu'une infection survient, le système passe de 1 à 2 malades (le nouveau contaminé **et** le patient zéro qui est toujours contagieux). Dans le formalisme de Galton-Watson, on considère que l'état initial a donné naissance à 2 branches indépendantes. Si le patient zéro infecte plus tard une autre personne, cela correspondra simplement à un nouveau branchement sur l'une de ces lignées. La possibilité d'avoir un grand nombre de cas finaux est donc mathématiquement incluse dans le développement en cascade du terme q^2 .*

Conséquence sanitaire majeure :

Même avec un pathogène extrêmement contagieux (par exemple $R_0 = 3$), si l'épidémie démarre avec 1 seul individu malade, il y a $\frac{1}{3} \approx 33\%$ de chances que le foyer s'éteigne naturellement et n'entraîne aucune pandémie. Cela montre la limite fondamentale du modèle déterministe qui, dans ce cas précis, aurait prédit avec certitude l'explosion des cas.

3.7 Simulation exacte : L'algorithme de Gillespie

Comme nous l'avons souligné à la section 3.4, l'explosion combinatoire de l'équation maîtresse rend toute résolution analytique globale impossible dès que la population dépasse quelques dizaines d'individus. Pour contourner cette impasse analytique et observer la dynamique stochastique exacte sans recourir à des approximations, nous devons utiliser des méthodes de Monte-Carlo. L'algorithme de choix pour simuler ces processus de Markov à temps continu est l'**Algorithme de Gillespie**. Plutôt que de faire avancer le temps par petits pas fixes Δt , cet algorithme détermine précisément "quand" a lieu le prochain événement et "lequel" se produit.

Définition 12 (Principe de l'algorithme de Gillespie). *À un instant t , le système est dans l'état (S, I) . On calcule les "propensions" (ou taux d'occurrence) des deux seuls événements possibles :*

- Taux d'une nouvelle infection : $R_{inf} = \frac{\beta}{N}SI$
- Taux d'une guérison : $R_{gue} = \gamma I$

Le taux global d'activité du système est la somme de ces propensions : $R_{tot} = R_{inf} + R_{gue}$.

L'algorithme repose sur les propriétés des lois exponentielles :

1. **Le temps d'attente** τ avant le prochain événement (quel qu'il soit) suit une loi exponentielle de paramètre R_{tot} .
2. **La nature de l'événement** est déterminée par un tirage aléatoire pondéré par les propensions relatives. La probabilité que ce soit une infection est $\frac{R_{inf}}{R_{tot}}$, et une guérison $\frac{R_{gue}}{R_{tot}}$.

Déroulement de la boucle algorithmique :

Étape 1 : Initialisation. On fixe le temps $t = 0$ et l'état initial (S_0, I_0, R_0) .

Étape 2 : Calcul des taux. On évalue R_{inf} , R_{gue} et R_{tot} . Si $R_{tot} = 0$ (c'est-à-dire $I = 0$), l'épidémie est terminée, la boucle s'arrête.

Étape 3 : Tirages aléatoires. On tire deux nombres pseudo-aléatoires uniformes $u_1, u_2 \in]0, 1[$.

— On calcule le temps écoulé : $\tau = \frac{1}{R_{tot}} \ln \left(\frac{1}{u_1} \right)$.

— On détermine l'événement : si $u_2 < \frac{R_{inf}}{R_{tot}}$, c'est une infection. Sinon, c'est une guérison.

Étape 4 : Mise à jour. On avance l'horloge globale ($t \leftarrow t + \tau$). On met à jour la population : $(S \leftarrow S-1, I \leftarrow I+1)$ pour une infection, ou $(I \leftarrow I-1, R \leftarrow R+1)$ pour une guérison.

Étape 5 : On retourne à l'Étape 2.

Cet algorithme va nous permettre de générer des trajectoires "exactes" de notre modèle stochastique afin de les confronter aux solutions lisses des équations différentielles dans la partie suivante.

3.8 Application aux petits foyers infectieux : La distribution bimodale

Le calcul précédent de la probabilité d'extinction a été réalisé sous l'hypothèse d'une population suffisamment grande pour considérer $S(t) \approx N$. Cependant, la véritable puissance de l'approche stochastique se révèle lors de l'étude de foyers restreints (une classe de 30 élèves, un service hospitalier, un EHPAD), où la population N est petite.

Dans ce contexte, le modèle déterministe continu perd tout son sens. Si l'on résout le système d'équations différentielles pour $N = 30$ et $R_0 = 2.5$, le modèle déterministe prédit de manière univoque que la maladie va se propager et toucher environ 89% de la classe (soit 26 à 27 élèves). Il n'offre aucune autre alternative.

L'approche probabiliste par chaînes de Markov montre une réalité beaucoup plus nuancée et conforme aux observations de terrain. Si l'on simule un grand nombre de fois la même introduction d'un "patient zéro" dans cette classe de 30 élèves, la taille finale de l'épidémie (le nombre total de personnes ayant été infectées) n'est plus une valeur fixe, mais une **variable aléatoire suivant une distribution bimodale** (en forme de "U").

- **Le premier mode (L'extinction précoce)** : Dans une proportion significative des cas (environ $\frac{1}{R_0}$ comme démontré en section 3.6), le patient zéro et les toutes premières cibles guérissent avant d'avoir pu créer une véritable chaîne de transmission. L'épidémie s'éteint avec un bilan très faible (1 à 3 malades au total).
- **Le second mode (L'épidémie majeure)** : Si la chaîne de transmission survit à ces premiers instants critiques, le virus se propage dans ce petit espace confiné. Le bilan final rejoint alors l'estimation du modèle déterministe, touchant la grande majorité du groupe.

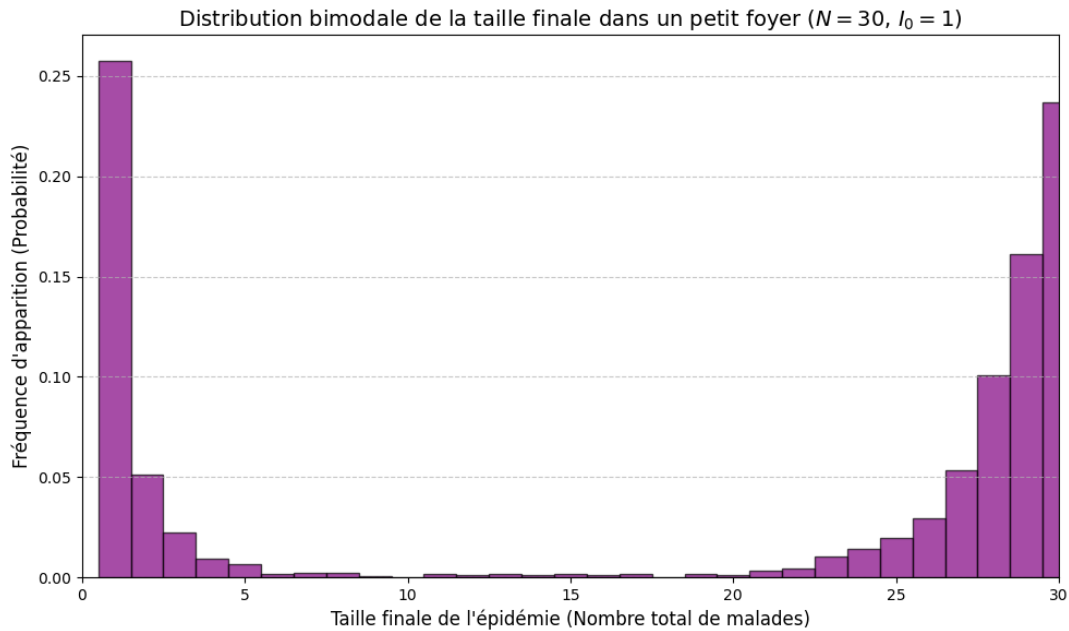


FIGURE 9 – Distribution bimodale de la taille finale dans un petit foyer ($N = 30$, $I_0 = 1$, $R_0 \approx 3,33$).

L’histogramme **Figure 9** illustre la répartition des tailles finales sur 2000 simulations. On y observe distinctement les deux modes caractéristiques du modèle stochastique :

- **Le pic d’extinction précoce** (à gauche) : environ 26 % des simulations s’arrêtent dès les premiers cas. Cet écart avec la valeur théorique asymptotique ($1/R_0 \approx 30\%$) illustre la limite du modèle de Galton-Watson. En effet, ce théorème repose sur l’hypothèse d’une population de taille infinie, garantissant un taux de transmission constant. Or, dans un petit foyer ($N = 30$), chaque nouvelle contamination réduit significativement le réservoir de cibles restantes. Cet épuisement rapide des individus sains modifie les probabilités de transmission à chaque étape, expliquant ainsi que la dynamique exacte du petit foyer s’écarte de la formule asymptotique pure.
- **Le pic d’épidémie majeure** (à droite) : le reste des trajectoires converge vers une infection massive du groupe, rejoignant le modèle déterministe.

Illustration :

Afin de mieux visualiser la dynamique temporelle sous-jacente à cette distribution bimodale, il est pertinent de confronter, grâce à l’algorithme de Gillespie, deux simulations de l’évolution du nombre de malades $I(t)$ dont les allures sont diamétralement opposées. Deux foyers stochastiques distincts, générés à partir de paramètres initiaux strictement identiques ($N = 30$, $I_0 = 1$), sont présentés ci-dessous.

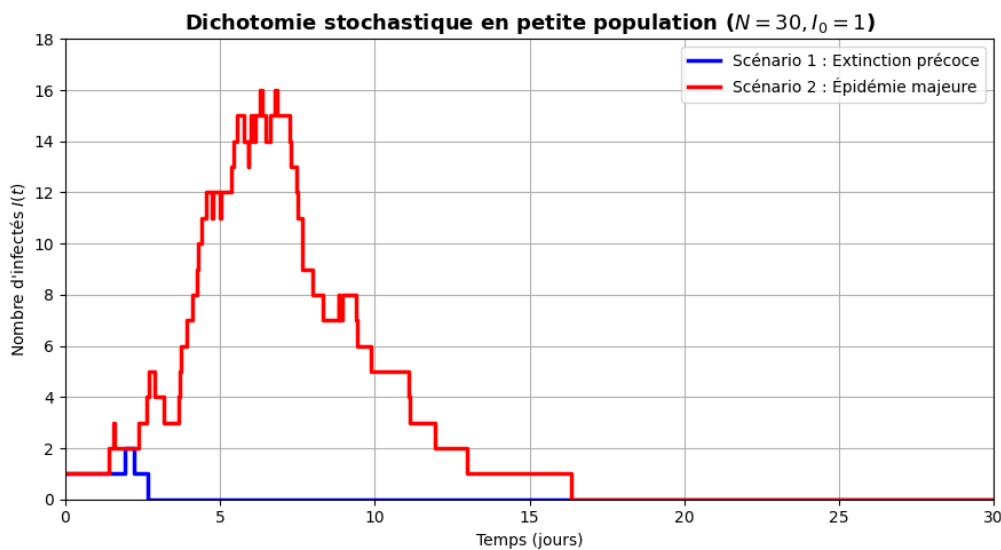


FIGURE 10 – **Dichotomie stochastique en petite population.** Ces deux trajectoires correspondent aux premières simulations générées par l’algorithme satisfaisant des critères d’amplitude prédéfinis afin d’illustrer ces deux scénarios extrêmes. La courbe bleue illustre le premier mode de l’histogramme (extinction précoce après quelques guérisons). La courbe rouge illustre le second mode (épidémie majeure s’emballant jusqu’au pic).

L'observation directe de ces trajectoires (**Figure 10**) met en évidence la criticité des premiers jours d'infection. La courbe bleue illustre un avortement rapide de la chaîne de transmission, dicté par le hasard des premières guérisons. Ce scénario correspond au pic gauche de l'histogramme (dont la probabilité d'occurrence approche $1/R_0$). À l'inverse, la courbe rouge représente une situation où le virus survit aux premières fluctuations aléatoires, déclenchant une flambée épidémique qui correspond au pic droit de l'histogramme.

Conclusion sanitaire :

Dans un petit foyer, la propagation n'est pas une fatalité moyenne, c'est un phénomène du "tout ou rien". Cette propriété probabiliste fondamentale justifie les mesures de gestion de crise locales (comme la fermeture immédiate d'une classe de maternelle dès l'apparition d'un cas de COVID-19) : l'objectif est de forcer mathématiquement le système à basculer dans le premier mode (l'extinction précoce) avant que la loi des grands nombres ne reprenne le dessus.

3.9 Synthèse de l'approche stochastique

En conclusion de cette partie, le passage d'une vision fluide et déterministe à une modélisation discrète et stochastique a radicalement transformé notre compréhension de l'amorce épidémique. Là où le système d'équations différentielles prédisait une explosion systématique dès lors que $R_0 > 1$, l'approche par chaînes de Markov a révélé l'existence d'une probabilité d'extinction spontanée, quantifiée par la relation $\mathbb{P}_{\text{ext}} = (1/R_0)^{I_0}$.

L'étude des processus de branchement a permis de démontrer que, dans les milieux restreints tels que les écoles ou les services hospitaliers, la dynamique n'est pas régie par une fatalité déterministe mais par un phénomène bimodal du "tout ou rien". Nous avons vu que l'équation maîtresse, bien que rigoureuse, se heurte à une impasse analytique face à l'explosion combinatoire des états, justifiant ainsi le recours à l'algorithme de Gillespie pour simuler l'aléatoire inhérent au vivant.

Toutefois, une question fondamentale demeure : comment ces deux visions, si divergentes dans leurs conclusions pour les petits groupes, peuvent-elles décrire la même réalité biologique ? L'enjeu de la partie suivante sera de confronter ces modèles afin de mettre en évidence la convergence des trajectoires stochastiques vers les courbes déterministes sous l'effet de la loi des grands nombres.

4 Partie III : La confrontation des modèles et la loi des grands nombres

Dynamique comparative et influence de la taille de la population :

Le modèle déterministe (Partie I) et le modèle stochastique (Partie II) semblent à première vue mener à des conclusions divergentes. Là où les équations différentielles imposent une trajectoire unique et une flambée épidémique systématique dès lors que $R_0 > 1$, le modèle de Markov met en évidence une distribution bimodale où l'extinction précoce reste possible par pur hasard.

Pour comprendre le lien profond entre ces deux visions, il est nécessaire d'étudier l'impact du paramètre structurel majeur : la taille de la population totale N . À l'aide de l'algorithme de Gillespie, nous avons simulé la dynamique temporelle de la proportion d'infectés $I(t)/N$ pour différentes échelles de population ($N = 30$, $N = 300$, et $N = 3000$). Les paramètres biologiques sont constants ($\beta = 0.5, \gamma = 0.15$, soit $R_0 \approx 3,33$) et la proportion initiale d'infectés est fixée à environ 3%.

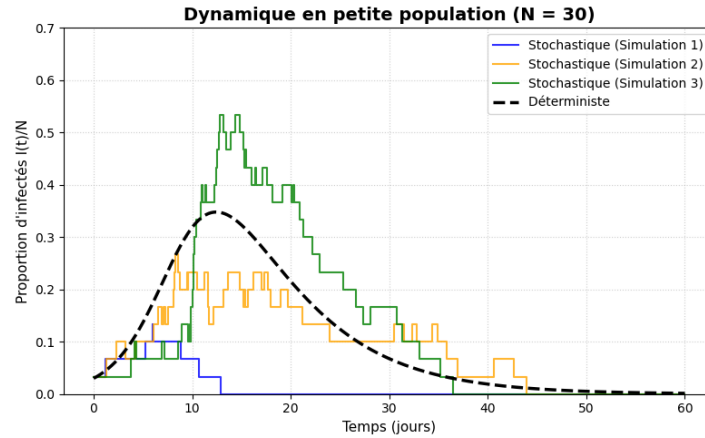


FIGURE 11 – **Dynamique en très petite population ($N = 30$).** Évolution de la proportion d'infectés $I(t)/N$ en fonction du temps. La courbe en pointillés noirs représente la solution du modèle déterministe continu. Les courbes de couleur (bleue, orange, verte) correspondent à trois réalisations stochastiques distinctes générées par l'algorithme de Gillespie. Ces trois trajectoires ont été délibérément sélectionnées parmi un grand nombre de tirages afin d'illustrer la diversité des scénarios extrêmes possibles à cette échelle.

Le modèle stochastique (courbes en escalier) révèle une extrême variabilité par rapport à la prévision déterministe (en pointillés noirs). La simulation 1 (bleue) illustre une extinction précoce, la simulation 2 (orange) une flambée retardée, tandis que la simulation 3 (verte) affiche un dépassement massif du pic théorique induit par les fluctuations aléatoires.

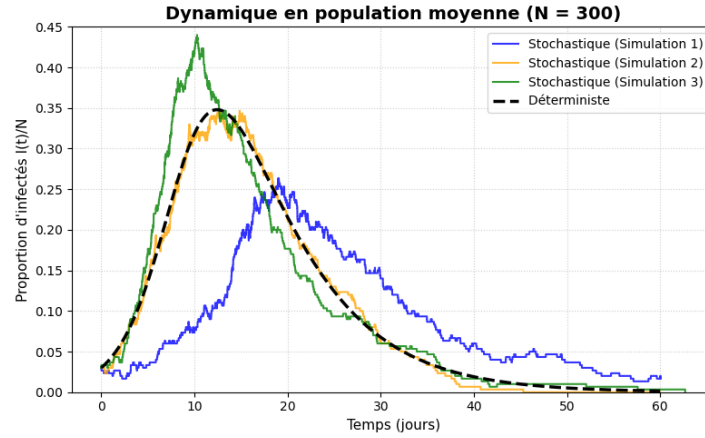


FIGURE 12 – **Dynamique en population moyenne ($N = 300$)**. Évolution de la proportion d'infectés $I(t)/N$ en fonction du temps. La courbe en pointillés noirs représente la prévision du modèle déterministe. Les courbes de couleur (bleue, orange, verte) représentent trois tirages stochastiques successifs obtenus par l'algorithme de Gillespie, sans sélection préalable.

Les trajectoires aléatoires commencent à se resserrer autour de la tendance déterministe. La forme en "cloche" émerge, bien que des décalages temporels majeurs (courbe bleue) et des variations locales restent bien visibles.

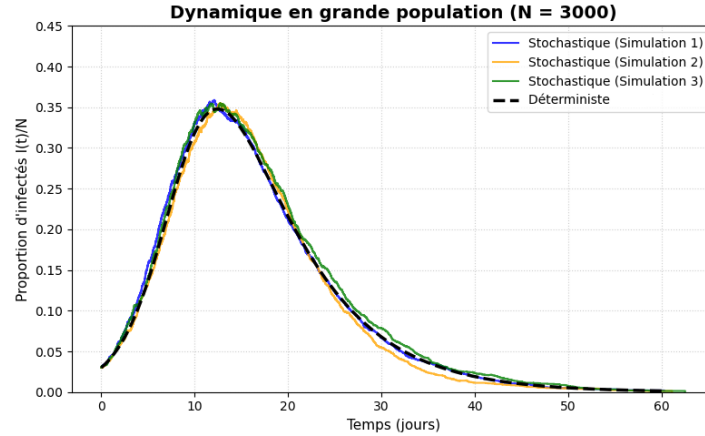


FIGURE 13 – **Dynamique en grande population ($N = 3000$)**. Évolution de la proportion d'infectés $I(t)/N$ en fonction du temps. La courbe en pointillés noirs indique la solution déterministe, superposée à trois trajectoires stochastiques consécutives (courbes de couleur) générées aléatoirement par l'algorithme de Gillespie.

Les fluctuations liées au hasard des rencontres se compensent. Les différentes réalisations stochastiques se superposent presque parfaitement et deviennent mathématiquement indiscernables de la solution fluide continue.

L'analyse de ces trois figures permet d'observer la transition mathématique fascinante entre le monde microscopique (aléatoire) et le monde macroscopique (déterministe) :

- **À très faible effectif ($N = 30$, Figure 11) :** Le modèle continu s'avère incapable de décrire la réalité du petit foyer. Les trajectoires vivent leur propre dynamique, dictée par le hasard des toutes premières rencontres. On y observe visuellement l'extinction précoce (courbe bleue), une épidémie erratique et retardée (courbe orange), ou au contraire une flambée soudaine dépassant de très loin le pic théorique (courbe verte). À cette échelle, la courbe noire déterministe n'a aucune réalité biologique.
- **À effectif modéré ($N = 300$, Figure 12) :** L'allure générale de l'épidémie en "cloche" commence à émerger du chaos. Les sauts discrets deviennent plus petits proportionnellement à la taille totale du groupe. Toutefois, les fluctuations aléatoires restent suffisantes pour modifier l'intensité de la crise ou décaler le pic épidémique de plusieurs jours (comme l'illustre le retard important de la courbe bleue).
- **En grande population ($N = 3000$, Figure 13) :** Le hasard individuel s'efface totalement au profit du comportement collectif. Les marches d'escalier microscopiques deviennent infinitésimales à l'échelle macroscopique. Les trajectoires probabilistes épousent alors parfaitement la courbe lisse des équations différentielles de Kermack-McKendrick.

Ce phénomène illustre l'action implacable de la **loi des grands nombres** appliquée aux processus de sauts : à mesure que le nombre d'individus et d'interactions augmente, les fluctuations aléatoires se compensent et s'annulent. La moyenne probabiliste dicte alors une trajectoire unique et inéluctable, justifiant ainsi l'utilisation du modèle déterministe à l'échelle d'un pays.

5 Conclusion et perspectives

Au terme de cette étude comparative, la modélisation de la propagation épidémique nous a permis de confronter deux approches mathématiques majeures.

D'une part, l'approche déterministe classique par systèmes d'équations différentielles offre un cadre analytique puissant pour comprendre les grandes épidémies. Elle permet d'établir des invariants structurels universels, comme le taux de reproduction de base R_0 , et de définir des seuils critiques exploitables par les autorités sanitaires, comme le niveau d'immunité collective requis pour éteindre une épidémie. Sa force réside dans sa simplicité de calcul et sa capacité de prédiction globale à l'échelle macroscopique.

D'autre part, l'approche stochastique par les chaînes de Markov à temps continu lève les limites conceptuelles du continu en redonnant au hasard et à l'individualité leur place légitime. En s'appuyant sur les processus de branchement de Galton-Watson et les simulations exactes de l'algorithme de Gillespie, nous avons démontré que dans les populations restreintes, la trajectoire moyenne n'a plus de sens. L'amorce d'une épidémie obéit à une logique du "tout ou rien", où une extinction spontanée reste fortement probable.

La confrontation finale de ces deux modèles réconcilie ces deux visions. Elle met en lumière la beauté des probabilités appliquées : l'ordre parfait et prévisible du modèle déterministe macroscopique émerge de manière rigoureuse et inéluctable du chaos stochastique microscopique individuel sous l'effet de la loi des grands nombres.

Perspectives et prolongements analytiques :

Bien que la convergence du modèle stochastique vers le modèle déterministe ait été clairement mise en évidence par les simulations numériques, une démonstration analytique rigoureuse de ce phénomène constituerait le prolongement naturel de ce travail. Une telle étude pourrait être menée en s'appuyant sur les travaux de Thomas Kurtz [T. Kurtz, Solutions of ordinary differential equations as limits of pure jump Markov processes, Journal of Applied Probability, 1970]. Le théorème de Kurtz sur la limite des processus de saut de Markov établit formellement que, sous des conditions de régularité des taux de transition, la trajectoire stochastique normalisée par la taille de la population (ici la proportion $\frac{I(t)}{N}$) tend en probabilité, et uniformément sur tout intervalle de temps fini, vers la solution du système d'équations différentielles déterministe lorsque $N \rightarrow +\infty$. Ce cadre théorique permettrait de confirmer analytiquement les observations numériques de cette étude, en démontrant rigoureusement la transition entre la modélisation probabiliste en effectif fini et la limite continue.

Références

- [1] W.O. Kermack et A.G. McKendrick. *A Contribution to the Mathematical Theory of Epidemics*. Proceedings of the Royal Society of London, A, 115(772), 1927. (<http://acdc2007.free.fr/kermack1927.pdf>).
- [2] D.T. Gillespie. *Exact Stochastic Simulation of Coupled Chemical Reactions*. The Journal of Physical Chemistry, 81(25), 1977. (<https://www.cmor-faculty.rice.edu/~cox/gillespie.pdf>).
- [3] N. Bacaër. *Mathématiques et épidémies*. Éditions Cassini. (Ouvrage culturel et mathématique de référence).
- [4] L.J.S. Allen. *An Introduction to Stochastic Processes with Applications to Biology*. Pearson Education. (<https://sistemas.fciencias.unam.mx/~silo/Cursos/coronavirus/Allen.pdf>).
- [5] H. Andersson et T. Britton. *Stochastic epidemic models and their statistical analysis*. 2000. (<https://staff.math.su.se/tom.britton/newmain.pdf>).
- [6] K.B. Athreya et P.E. Ney. *Branching Processes*. Springer Verlag, 1972. (<https://fr.scribd.com/document/849438472/Springer-Athreya-Ney-Branching-Processes>).
- [7] G. Sallet. *Mathematical Epidemiology*. Université de Lorraine, 2018. (<https://hal.science/hal-04688889>).
- [8] Wikipédia, l'encyclopédie libre. *Modèles compartimentaux en épidémiologie*. (https://fr.wikipedia.org/wiki/Mod%C3%A8les_compartimentaux_en_%C3%89pid%C3%89miologie).
- [9] Wikipedia, l'encyclopédie libre. *Nombre de reproduction de base*. (https://fr.wikipedia.org/wiki/Nombre_de_reproduction_de_base).
- [10] Wikipedia, the free encyclopedia. *Continuous-time Markov chain* (https://en.wikipedia.org/wiki/Continuous-time_Markov_chain)
- [11] Wikipedia, l'encyclopédie libre. *Processus de Galton-Watson*. (https://fr.wikipedia.org/wiki/Processus_de_Galton-Watson).
- [12] Wikipedia, the free encyclopedia. *Gillespie algorithm*. (https://en.wikipedia.org/wiki/Gillespie_algorithm)
- [13] INRIA (Institut national de recherche en sciences et technologies du numérique). *Vidéo sur la modélisation épidémique*. (<https://www.youtube.com/watch?v=LLjQxWVJwdk>).
- [14] Institut Pasteur. *Notes pédagogiques sur le R_0 et la dynamique des épidémies*. (<https://modelisation-covid19.pasteur.fr>).

Annexe

A Scripts Python des simulations

Ce document regroupe l'ensemble des scripts Python utilisés pour générer les figures et modéliser les dynamiques épidémiques (modèles déterministes et stochastiques) présentées dans le rapport. Pour des raisons de concision, les codes ayant servi à générer des figures similaires avec des paramètres différents (comme l'évolution de la population N) sont regroupés.

Portrait de phase du modèle SIR fermé

```
1 import numpy as np
2 import matplotlib.pyplot as plt
3
4 # Paramètres du modèle SIR (fermé)
5 N = 1000
6 beta = 0.5          # Taux de contact efficace
7 gamma = 0.15        # Taux de guérison
8 R0 = beta / gamma
9
10 def trajectoire_phase(S, I0, S0, N, R0):
11     K = I0 + S0 - (N / R0) * np.log(S0)
12     return -S + (N / R0) * np.log(S) + K
13
14 S_valeurs = np.linspace(1, N, 2000)
15
16 infections_initiales = [1, 10, 50]
17 couleurs = ['blue', 'orange', 'green']
18
19 plt.figure(figsize=(10, 6))
20
21 for i, I0 in enumerate(infections_initiales):
22     S0 = N - I0
23     I_valeurs = trajectoire_phase(S_valeurs, I0, S0, N, R0)
24
25     # un nombre d'infectés négatif n'a pas de sens
26     # On crée une variable pour ne tracer que la partie positive de la
    courbe
27     valide = I_valeurs >= 0
28
29     plt.plot(S_valeurs[valide], I_valeurs[valide],
30             color=couleurs[i], label=f'Départ $I_0 = {I0}$')
31
32 seuil = N / R0
33 plt.axvline(x=seuil, color='r', linestyle='--', label=f'Seuil $N/R_0$ ({
    int(seuil)})')
34 plt.title(f"Portrait de phase SIR ($R_0 = {R0:.2f}$)")
35 plt.xlabel("Nombre d'individus Sains ($S$)")
36 plt.ylabel("Nombre d'individus Infectés ($I$)")
37 plt.legend()
38 plt.grid(True)
39 plt.show()
```

Portrait de phase du modèle SIR ouvert (avec démographie)

```
1 import numpy as np
2 import matplotlib.pyplot as plt
3 from scipy.integrate import odeint
4
5 # Paramètres du modèle SIR (ouvert)
6 N = 1000
7 beta = 0.5          # Taux de contact
8 gamma = 0.15        # Taux de guérison
9 mu = 0.02           # Taux de renouvellement démographique (natalité =
    mortalité)
10
11 R0_ouvert = beta / (gamma + mu)
12
13 def sir_ouvert(y, t, N, beta, gamma, mu):
14     S, I = y
15     # Flux entrant (naissances) et sortants (infections, guérisons, décès)
16     dSdt = mu * N - (beta * S * I / N) - mu * S
17     dIdt = (beta * S * I / N) - (gamma + mu) * I
18     return [dSdt, dIdt]
19
20 temps = np.linspace(0, 300, 5000)
21
22 # On teste trois situations de départ différentes (S, I)
23 conditions_initiales = [[990, 10], [500, 500], [100, 900]]
24
25 trajectoires = [odeint(sir_ouvert, ci, temps, args=(N, beta, gamma, mu))
    for ci in conditions_initiales]
26
27 # L'équilibre sans maladie : Disease Free Equilibrium (DFE)
28 S_dfe, I_dfe = N, 0
29
30 # Équilibre endémique (EE)
31 S_ee = N / R0_ouvert
32 I_ee = (mu * N / beta) * (R0_ouvert - 1)
33
34 plt.figure(figsize=(10, 6))
35
36 couleurs = ['blue', 'cyan', 'pink']
37
38 for i, sol in enumerate(trajectoires):
39     plt.plot(sol[:, 0], sol[:, 1], color=couleurs[i])
40     plt.plot(sol[0, 0], sol[0, 1], 'ko', markersize=4)
41 plt.plot(S_dfe, I_dfe, 'ro', markersize=8, label='DFE (Sans Maladie)')
42 plt.plot(S_ee, I_ee, 'go', markersize=8, label='Équilibre Endémique')
43 plt.title(f"Portrait de phase d'un modèle SIR avec démographie ($R_0 = {
    R0_ouvert:.2f}$)")
44 plt.xlabel("Nombre d'individus Sains (S)")
45 plt.ylabel("Nombre d'individus Infectés (I)")
46 plt.legend()
47 plt.grid(True)
48 plt.show()
```

Distribution bimodale de la taille finale

```
1 import numpy as np
2 import matplotlib.pyplot as plt
3
4 # Paramètres du petit foyer
5 N = 30
6 I0 = 1
7 S0 = 29
8 beta = 1.0      # Taux de contact élevé (milieu fermé)
9 gamma = 0.3     # Taux de guérison
10 R0 = beta/gamma # R0 = 3.33
11
12 n_simulations = 2000
13 tailles_finales = []
14
15 #Algorithme de Gillespie (sans tau) : le temps d'attente n'est pas
    utilisé car inutile pour calculer la taille finale.
16 for _ in range(n_simulations):
17     S = S0
18     I = I0
19
20     while I > 0:
21         r_inf = beta * S * I / N
22         r_gue = gamma * I
23         r_tot = r_inf + r_gue
24
25         u = np.random.uniform(0, 1)
26         if u < r_inf / r_tot:
27             S -= 1
28             I += 1
29         else:
30             I -= 1
31
32     tailles_finales.append(N - S)
33
34 plt.figure(figsize=(10, 6))
35
36 bins = np.arange(0.5, N + 1.5, 1)
37 poids = np.ones_like(tailles_finales) / 2000
38
39 plt.hist(tailles_finales, bins=bins, weights=poids, color='purple',
    edgecolor='black', alpha=0.7)
40 plt.title(f"Distribution bimodale de la taille finale dans un petit
    foyer ($N={N}$, $I_0={I0}$)", fontsize=14)
41 plt.xlabel("Taille finale de l'épidémie (Nombre total de malades)",
    fontsize=12)
42 plt.ylabel("Fréquence d'apparition (Probabilité)", fontsize=12)
43 plt.grid(axis='y', linestyle='--', alpha=0.7)
44 plt.xlim(0, N)
45 plt.tight_layout()
46 plt.show()
```

Dichotomie stochastique en petite population

```
1 import numpy as np
2 import matplotlib.pyplot as plt
3
4 # Paramètres du petit foyer
5 N = 30
6 I0 = 1
7 S0 = 29
8 beta = 1.0          # Taux de contact
9 gamma = 0.3         # Taux de guérison
10 t_max = 30
11
12 def gillespie_I(S0, I0, beta, gamma, N, t_max):
13     t, S, I = [0], [S0], [I0]
14     while t[-1] < t_max and I[-1] > 0:
15         r_inf = beta * S[-1] * I[-1] / N
16         r_gue = gamma * I[-1]
17         r_tot = r_inf + r_gue
18         if r_tot == 0: break
19         t.append(t[-1] + np.random.exponential(1.0 / r_tot))
20         u = np.random.uniform(0, 1)
21         if u < r_inf / r_tot:
22             S.append(S[-1] - 1); I.append(I[-1] + 1)
23         else:
24             S.append(S[-1]); I.append(I[-1] - 1)
25     if t[-1] < t_max:
26         t.append(t_max)
27         I.append(I[-1])
28     return np.array(t), np.array(I)
29
30 traj_extinction = None
31 traj_maj = None
32
33 # On cherche une trajectoire qui s'éteint vite (max I < 3) et une qui
34   flambe (max I > 10)
35 while traj_extinction is None or traj_maj is None:
36     t, I = gillespie_I(S0, I0, beta, gamma, N, t_max)
37     if max(I) <= 2 and traj_extinction is None:
38         traj_extinction = (t, I)
39     elif max(I) >= 10 and traj_maj is None:
40         traj_maj = (t, I)
41
42 plt.figure(figsize=(9, 5))
43
44 t_ext, I_ext = traj_extinction
45 plt.step(t_ext, I_ext, color='blue', linewidth=2.5, where='post', label=
46   "Scénario 1 : Extinction précoce")
47
48 t_maj, I_maj = traj_maj
49 plt.step(t_maj, I_maj, color='red', linewidth=2.5, where='post', label="
50   Scénario 2 : Épidémie majeure")
51
52 plt.title(f"Dichotomie stochastique en petite population ($N=30, I_0=1$)
53   ", fontsize=13, fontweight='bold')
54 plt.xlabel("Temps (jours)")
55 plt.ylabel("Nombre d'infectés $I(t)$")
56 plt.xlim(0, t_max)
57 plt.ylim(0, 18)
```

```

53 plt.yticks(np.arange(0, 19, 2))
54 plt.grid(True)
55 plt.legend()
56 plt.tight_layout()
57 plt.show()

```

Convergence stochastique vers déterministe

Le script suivant a été utilisé pour simuler la convergence du modèle stochastique vers le modèle déterministe. Pour générer les figures correspondantes aux populations $N = 30$, $N = 300$ et $N = 3000$, la variable N a été modifiée.

```

1 import numpy as np
2 import matplotlib.pyplot as plt
3 from scipy.integrate import odeint
4
5 N = 3000 # variable à modifier
6
7 def deriv(y, t, beta, gamma):
8     return [-beta * y[0] * y[1], beta * y[0] * y[1] - gamma * y[1]]
9
10 def gillespie(S0, I0, beta, gamma, N, t_max):
11     t, S, I = [0], [S0], [I0]
12     while t[-1] < t_max and I[-1] > 0:
13         rate_inf, rate_rec = beta * S[-1] * I[-1] / N, gamma * I[-1]
14         rate_tot = rate_inf + rate_rec
15         if rate_tot == 0: break
16         t.append(t[-1] + np.random.exponential(1.0 / rate_tot))
17         if np.random.uniform(0, 1) < rate_inf / rate_tot:
18             S.append(S[-1] - 1); I.append(I[-1] + 1)
19         else:
20             S.append(S[-1]); I.append(I[-1] - 1)
21     if t[-1] < t_max:
22         t.append(t_max); S.append(S[-1]); I.append(I[-1])
23     return np.array(t), np.array(I)/N
24
25 prop_I0 = 0.03
26 I0 = max(1, int(N * prop_I0))
27 S0 = N - I0
28 beta, gamma, t_max = 0.5, 0.15, 60
29
30 t_det = np.linspace(0, t_max, 500)
31 sol_det = odeint(deriv, [1.0 - prop_I0, prop_I0], t_det, args=(beta,
32     gamma))
33
34 plt.figure(figsize=(8, 5))
35 colors = ['blue', 'orange', 'green']
36
37 for j in range(3):
38     t_stoch, I_stoch = gillespie(S0, I0, beta, gamma, N, t_max)
39     plt.step(t_stoch, I_stoch, color=colors[j], alpha=0.8, where='post',
40         label=f"Stochastique (Simulation {j+1})", linewidth=1.5)
41
42 plt.plot(t_det, sol_det[:, 1], 'k--', linewidth=2.5, label="Déterministe")
43
44 plt.title(f"Dynamique en grande population (N = {N})", fontsize=14,
45     fontweight='bold')
46 plt.xlabel("Temps (jours)", fontsize=11)

```



```
44 plt.ylabel("Proportion d'infectés  $I(t)/N$ ", fontsize=11)
45 plt.ylim(0, 0.45)
46 plt.grid(True, linestyle=':', alpha=0.6)
47 plt.legend()
48 plt.tight_layout()
49 plt.show()
```