



Stochastic Dynamics and Extinction Time in SIR Epidemiological Models

Rachid El Chaal^{1*}, Said Bouchefra², Moulay Othman Aboutafail¹

¹ Engineering Sciences Laboratory, Data Analysis, Mathematical Modeling, and Optimization Team, Department of Computer Science, Logistics and Mathematics, Ibn Tofail University National School of Applied Sciences ENSA, 14 000 Kenitra, Morocco

² Laboratory of Biology and Health, Team of Nutritional Sciences, Food and Health, Faculty of Sciences, University Ibn Tofail, FSK-Kenitra, 14 000 Kenitra, Morocco

* Correspondence: Rachid El Chaal (rachid.elchaal@uit.ac.ma)

Received: 11-12-2023

Revised: 12-19-2023

Accepted: 12-25-2023

Citation: R. El Chaal, S. Bouchefra, and M. O. Aboutafail, “Stochastic dynamics and extinction time in SIR epidemiological models,” *Acadlore Trans. Appl Math. Stat.*, vol. 1, no. 3, pp. 181–202, 2023. <https://doi.org/10.56578/atams010305>.



© 2023 by the authors. Published by Acadlore Publishing Services Limited, Hong Kong. This article is available for free download and can be reused and cited, provided that the original published version is credited, under the CC BY 4.0 license.

Abstract: In the realm of epidemiological modeling, the intricacies of epidemic dynamics are elucidated through the lens of compartmental models, with the SIR (Susceptible-Infectious-Recovered) and its variant, the SIS (Susceptible-Infectious-Susceptible) model, being pivotal. This investigation delves into both deterministic and stochastic frameworks, casting the SIR model as a continuous-time Markov chain (CTMC) in stochastic settings. Such an approach facilitates simulations via Gillespie’s algorithm and integration of stochastic differential equations. The latter are formulated through a bivariate Fokker-Planck equation, originating from the continuous limit of the master equation. A focal point of this study is the distribution of extinction time, specifically, the duration until recovery in a population with an initial count of infected individuals. This distribution adheres to a Gumbel distribution, viewed through the prism of a birth and death process. The stochastic analysis reveals several insights: firstly, the SIR model as a CTMC encapsulates random fluctuations in epidemic dynamics. Secondly, stochastic simulation methods, either through Gillespie’s algorithm or stochastic differential equations, offer a robust exploration of disease spread variability. Thirdly, the precision of modeling is enhanced by the incorporation of a bivariate Fokker-Planck equation. Fourthly, understanding the Gumbel distribution of extinction time is crucial for gauging recovery periods. Lastly, the non-linear nature of the SIR model, when analyzed stochastically, enriches the comprehension of epidemic dynamics. These findings bear significant implications for epidemic mitigation and recovery strategies, informing healthcare resource planning, vaccine deployment optimization, implementation of social distancing measures, public communication strategies, and swift responses to epidemic resurgences.

Keywords: Epidemiological modeling; SIR model; Stochastic analysis; Extinction time distribution; Gumbel distribution

1 Introduction

In December 2019, Dr. Zhang Jixian notified Chinese authorities about the emergence of a novel coronavirus, SARS-CoV-2 [1], which has infected 180 individuals in Wuhan, China. Europe witnessed a rapid escalation of the crisis as the year came to a close, with the first instances of the virus being reported in France, Italy, and Germany between January 24 and 29, 2020. The events that transpired led to the ongoing worldwide battle against the COVID-19 epidemic.

The significance of the SIR model in the context of COVID-19 resides in its capacity to offer a simplified yet potent depiction of disease transmission throughout a community. The model specifically focuses on the susceptible-infectious-recovered compartments, aiding in the enhanced comprehension and estimation of the epidemic’s progression. Here are few key factors that emphasize this significance:

(1) Projection of Epidemic Trends: The SIR model allows for projecting epidemic trends based on parameters such as the transmission rate and recovery rate. These projections can assist in anticipating the future spread of the disease and implementing preventive measures.

(2) Estimation of Key Parameters: By fitting the model to real-world data, it becomes possible to estimate key parameters such as the basic reproduction number (\mathcal{R}_0), providing crucial information to assess the transmission dynamics of the virus.

(3) Analysis of Intervention Scenarios: By modifying model parameters, the effectiveness of different interventions, such as social distancing, vaccination, or other control measures, can be evaluated, contributing to the development of pandemic control strategies.

(4) Identification of Critical Points: The SIR model can help identify critical points where swift interventions are necessary to prevent uncontrolled virus spread, contributing to proactive crisis management.

(5) Understanding Disease Dynamics: By breaking down the population into compartments, the model offers an intuitive understanding of different phases of the disease, from initial spread to recovery, facilitating the analysis of epidemiological dynamics.

However, this global crisis unexpectedly fueled a newfound interest in numbers. Access to pandemic statistics empowered various individuals, ranging from self-proclaimed statisticians to seasoned analysts, to construct graphs and curves. Media coverage introduced terms such as exponential growth, flattening the curve, and parameter \mathcal{R}_0 , which, when taken out of context, often seemed ambiguous or specific to epidemiology.

Motivated by this context, our work delves into the dynamics of virus spread, presenting a model less complex than SARS-CoV-2 but providing an introduction to epidemiology and population behavior modeling during a viral epidemic. Epidemiological compartmental models, despite simplifying reality, offer an elegant framework to modulate the system's complexity, yielding qualitative and elementary insights [2].

The mathematical study of epidemic dynamics experienced a resurgence during the SARS-CoV-2 pandemic. This revival is crucial, not only for understanding and managing epidemics but also for informing political decisions at the authorities' level and influencing global public health affairs amid a growing population [3].

Several questions propel this study. How can we comprehend epidemic dynamics? What determines an epidemic's persistence? How can we model interactions within a population? What will be the extent of the affected population? Epidemiology thrives on complexity, as individual interactions are inherently local and nonlinear. With a country's average population of 10^7 individuals, considering all micro- and macroscopic variables influencing virus spread becomes impractical and illusory.

This is where the stochastic components come into play when modeling interactions between individuals. We will return to this point later.

Specifically, the aim here is to study the SIR model, which belongs to the class of compartmental models in epidemiology, from both a deterministic and a stochastic point of view. Within the framework of the stochastic study, we will study this model as being a continuous-time Markovian jump process, which we will be able to simulate using the Gillespie algorithm and from which we will be able to obtain the master equation. Then, we will take the continuous limit of this master equation in order to obtain a bivariate Fokker-Planck equation. Finally, we will study the distribution of the epidemic's extinction time, a central issue in epidemiology [4].

The study objectives can be formulated as follows:

(1) Explore Epidemiological Dynamics: Gain an in-depth understanding of the dynamics of disease spread, with a focus on transitions between the susceptible, infectious, and recovered compartments of the SIR model.

(2) Evaluate the Adaptability of the SIR Model: Examine the SIR model's capacity to adapt to various epidemiological scenarios, including stochastic variability, to determine its robustness and applicability in real-world situations.

(3) Analyze stochastic aspects: deepen the stochastic analysis of the SIR model, highlighting the advantages of this approach over deterministic models and exploring how random fluctuations may influence outcomes.

(4) Understand the Extinction Time Distribution: Scrutinize the extinction time distribution in the context of the stochastic SIR model, emphasizing a comprehensive understanding of this distribution and its impact on the recovery duration of infected individuals.

(5) Contribute to the Optimization of Control Strategies: Provide practical insights and perspectives to refine disease control strategies by leveraging the results of the SIR model analysis, with a particular focus on implications for the management and prevention of epidemics.

The study's main goals are to improve our understanding of how diseases spread using the SIR model, see how well it can be used in different situations, look into random factors, figure out how long it takes for a disease to die out, and finally, make disease control strategies better.

2 Unraveling Epidemiological Dynamics: Exploring Compartmental Models in Public Health

A concise chronicle. To conduct an analysis of an epidemic, it is necessary to have daily census data. This is the reason why the field of epidemiology is relatively new. The initial instances of mortality investigations conducted by D. Bernoulli can be dated back to 1766, during which his objective was to exert an impact on public policy about a vaccine technique (or its contemporary counterpart) in order to mitigate the propagation of smallpox [5].

It was not until the beginning of the 20th century that more systematic work was produced to model the spread of diseases and viruses. Among these pioneers were R. Ross, W. H. Hamer, W. O. Kermack, and A. G. McKendrick, whose work in epidemiology was considerable between 1900 and 1935 [6]. Ross won the Nobel Prize for Medicine in 1902 for studying and demonstrating the dynamics of the propagation of the malaria parasite between mosquitoes and humans. He showed that below a certain critical threshold, the mosquito population is no longer able to carry the parasite and thus produce an epidemic. Hamer is best known for studying and modeling the recurrent nature of measles epidemics, and in particular for showing that an epidemic must fortiori depend on the number of susceptible and infected individuals.

This was followed, between 1927 and 1933, by articles by Kermack and McKendrick [7], which laid the foundations for compartmental models. The first compartmental diagrams appeared in their second article in 1932.

Principle: The idea is to segment the population into a certain number of compartments, each containing a certain number of individuals. A given individual can only be located in one and only one compartment. It can, however, be transferred to another bin according to a certain transfer rate or transition rate (which may depend on the individual chosen or the size of the population, for example). Its state is then modified. Depending on the model, the total population may or may not be retained. The latter case takes into account the death or birth of individuals.

An epidemiological compartmental model is composed of two essential components: compartments and rules. The rules dictate the circumstances under which an individual might transition from one compartment to another. The selection of compartments is unrestricted and contingent upon the subject of investigation. Table 1 provides a comprehensive list of the compartments that are frequently utilized. It is important to mention that, due to clear reasons, a minimum of compartments S and I are necessary.

The greater the number of compartments, the more the system will illustrate complex interactions and be able to highlight fairly detailed behavior. Such systems are generally, if not always, non-analytical. On the other hand, a small number of compartments is often synonymous with a simple model, which most of the time omits details and remains qualitative [8].

Once the compartments and rules have been chosen, a reaction scheme can be associated with the model. Often complementary to the reaction diagram, it is also possible to construct a compartmental diagram, which in itself summarizes the entire model. Figure 1, for example, illustrates the compartmental diagram of an SIRD model. In such a model, a healthy individual who comes into contact with an infected individual becomes infected himself. An infected individual can either die or recover.

Table 1. The different compartments commonly used and their definitions

Compartment	Type of Individual
S	Sains
I	Infected
E	Exposed
D	Deceased
R	Immune
M	Maternal immunity
C	Asymptomatic carrier
Q	In quarantine

Note: Compartment E refers to the latency period between exposure and infection

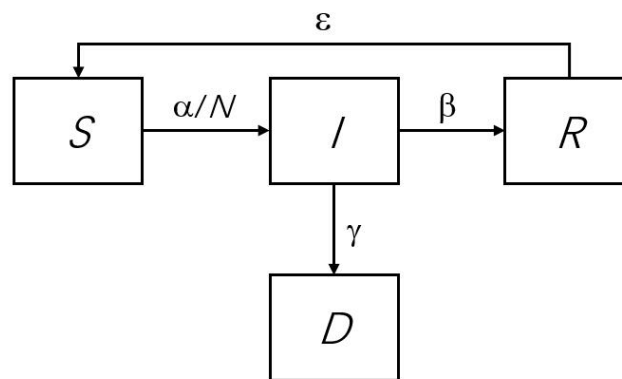


Figure 1. SIRD model in the form of compartments

In practice. Compartmental models can be interpreted deterministically or stochastically. This flexibility increases their scope. For each of the deterministic models, however complex, it is possible to define a quantity, called the basic reproduction number, that defines the threshold between epidemic growth and decline [9].

In particular, compartmental models are used to support political decisions: What will be the effects of a vaccine? A quarantine. As models, the results remain qualitative and cannot take into account unpredictable factors (environmental or economic catastrophe, civil war, etc.). However, they have already proved their worth for the spread of tuberculosis, AIDS, influenza, Ebola, and malaria [9]. The global SARS-CoV-2 coronavirus pandemic has given rise to a revival of compartmental models of all kinds, from the simplest to the most elaborate.

There has been a steady improvement in the SIR (Susceptible-Infectious-Recovered) model in epidemiology in order to better show how complicated epidemiological dynamics are. The significance of this evolution lies in the improvement of model accuracy to better anticipate, understand, and control the spread of diseases. Here is a perspective on the evolution of the SIR model and its significance:

(1) Origins of the SIR Model: The SIR model was initially developed to describe infectious disease outbreaks by considering three population compartments—susceptible, infectious, and recovered. It provided a simple yet powerful conceptual framework for understanding disease transmission within a population.

(2) Incorporation of Stochasticity: The recent evolution of the SIR model has included a more profound consideration of stochasticity, acknowledging that epidemiological phenomena may involve random elements. Stochastic SIR models better represent natural fluctuations in the spread of diseases.

(3) Spatial and Temporal Models: Variants of the SIR model have been developed to incorporate spatial and temporal elements, accounting for the mobility of individuals and geographical variations in disease transmission. This has strengthened the models' ability to predict spread on regional and national scales.

(4) Integration of More Complex Parameters: Advanced versions of the SIR model integrate more complex parameters, such as variable transmission rates, differentiated incubation periods, and specific subpopulations. This allows for a more realistic modeling of the diversity of epidemiological situations.

(5) Adaptation to Emerging Diseases: The evolution of the SIR model has been guided by the necessity to adapt to emerging diseases, such as COVID-19. Researchers continually adjust the models to make them relevant in specific epidemiological contexts.

The evolution of the SIR model underscores its constant adaptation to meet the growing challenges posed by the understanding and management of infectious diseases. This evolution enhances the relevance and applicability of the model in various epidemiological contexts, reinforcing its value as an essential tool for modeling and predicting epidemics.

The transition from deterministic models to stochastic models in epidemiology represents a fundamental shift in how we approach the modeling of infectious diseases. This transition is motivated by several factors and offers significant advantages. Here is a discussion on this transition and the underlying motivation for using stochastic models:

(1) Incorporation of Uncertainty: Deterministic models consider the spread of diseases as a predictable and deterministic process. However, in reality, many epidemiological variables involve an element of uncertainty. Stochastic models incorporate this uncertainty, allowing for a better representation of the random fluctuations observed in the transmission of diseases.

(2) Handling Small Populations: In small populations where random effects can have a significant impact, stochastic models are more suitable. They take into account the probability of disease extinction, which may not be considered in deterministic models.

(3) Exploration of Random Fluctuations: Stochastic models enable the exploration of inherent random fluctuations in epidemiological dynamics. This includes variations in the number of infections, the timing of epidemic events, and other factors that are challenging to predict deterministically.

(4) Precision in Complex Scenarios: In complex epidemiological situations, such as epidemics with variable social interactions, individual behavior changes, or public health interventions, stochastic models offer increased precision by considering inherent variability.

(5) Adaptation to Real Data: Real epidemiological data often exhibit unexpected variations. Stochastic models allow for the adjustment of parameters to better align with observed data, enhancing their ability to faithfully represent reality.

The adoption of stochastic models in epidemiology is a response to the necessity of accurately representing the uncertainty and variability inherent in the transmission of illnesses. This methodology offers a pragmatic and flexible viewpoint, empowering researchers and decision-makers to make well-informed choices in a wide range of intricate and varied epidemiological scenarios.

3 Unraveling the SIR Deterministic Model: Insights into Epidemic Dynamics

The model we propose to study is one of the simplest, with the SIS. It is defined as follows. Consider a population of N individuals, made up of three types of individuals: susceptible individuals (S) infected individuals (I) and permanently immunized individuals (R). On the principle of compartmental models, the population is divided into three groups as shown in Figure 2.

The number of individuals in each of these groups is obviously not constant. An individual in the S group may, for example, move to the I group, but not directly to the R . The compartments accessible to each individual can be summarized by writing the model in the form of compartments as shown in Figure 3. This defines the rules of the model.

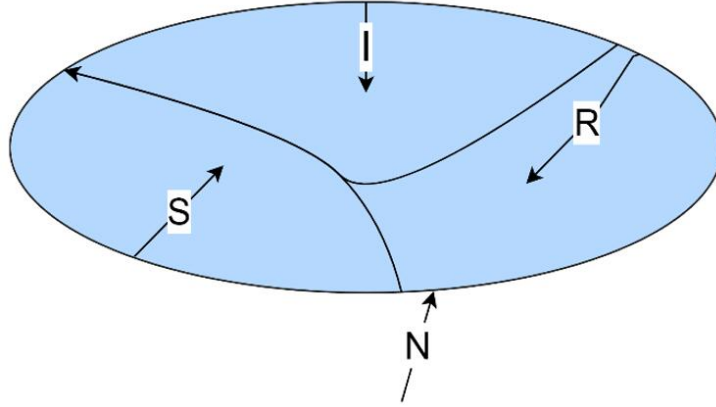


Figure 2. Division of the population of N individuals into three groups: S , I and R

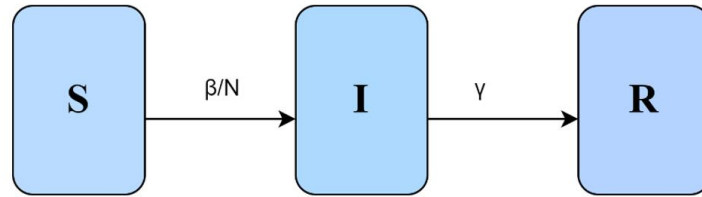


Figure 3. SIR model in compartment form

Let's define each individual's role precisely.

Susceptible individual: an individual who has no immunity to the infectious agent and therefore becomes infected if exposed.

Infected individual: an individual who can transmit the infection to susceptible individuals if they come into contact.

Immunized individual: an individual who can no longer contract the infection and therefore no longer infects other individuals. This individual is passive.

This is known as the epidemic SIR model [10]. Another version is the endemic SIR model, which deals with a disease/infection affecting a population on a permanent basis. The essential difference between the two lies in the time scales. If we note τ the characteristic time of the disease/infection (from contraction to recovery) and T the average lifespan of individuals in the population, then for an epidemic SIR model, $\tau \ll T$ whereas for an endemic SIR model, $\tau \sim T$. In the latter case, the evolution of the various population groups over the long term (i.e. of the order of T) is taken into account. It then becomes necessary to take birth and death rates into account [11]. Another difference is that the epidemic model is a closed system, $S + I + R = N = c^{ste}$ whereas the endemic model is open, $N = N(t)$. In the following, we will focus on the epidemic SIR model.

The pattern of reactions between individuals is as follows [12]:



with β transmission rate and τ recovery rate. Note that in the first reaction, the transmission rate is denoted β/N to take account of the interaction between two individuals from different groups. Note that the asymptotic state is necessarily trivial and corresponds to the case where $I = 0, S(\infty) + R(\infty) = N^\dagger$. Let's now derive the kinetic equations associated with this reaction scheme. Eq. (1a) gives,

$$\begin{aligned}\frac{dS}{dt} &= -\frac{\beta}{N}SI \\ \frac{dI}{dt} &= \frac{\beta}{N}SI\end{aligned}$$

and Eq. (1b) gives,

$$\begin{aligned}\frac{dI}{dt} &= -\gamma I \\ \frac{dR}{dt} &= \gamma I\end{aligned}$$

Summing up these contributions, we obtain the kinetic equations of the SIR model [13],

$$\frac{dS}{dt} = -\frac{\beta}{N}SI \quad (2a)$$

$$\frac{dI}{dt} = \frac{\beta}{N}SI - \gamma I \quad (2b)$$

$$\frac{dR}{dt} = \gamma I \quad (2c)$$

Let's define the following variables,

$$\begin{cases} x \stackrel{\text{def}}{=} S/N \\ y \stackrel{\text{def}}{=} I/N \\ z \stackrel{\text{def}}{=} R/N \end{cases}$$

Eqs. (2a)-(2c) then take the following form,

$$\frac{dx}{dt} = -\beta xy \quad (3a)$$

$$\frac{dy}{dt} = \beta xy - \gamma y \quad (3b)$$

$$\frac{dz}{dt} = \gamma y \quad (3c)$$

The dynamics of this system of differential equations can therefore be studied on the 2-simplex defined by $\{(x, y, z) \in \mathbb{R}_+^3 \mid x + y + z = 1\}$. In this sense, it will suffice to study the dynamics implied by Eqs. (3a) and (3b) on the projection \mathcal{P} of the 2-simplex on the plane (x, y) .

The transition from a 3-dimensional system to the 2-simplex and the subsequent projection involves several mathematical steps. Here's a detailed explanation:

1. Three-Dimensional System:

Initially, the system is described in a 3-dimensional space with variables (x, y, z) , representing different compartments of the population.

2. 2-Simplex Definition:

The 2-simplex is a geometric object defined by the constraint $x + y + z = 1$. Geometrically, it forms a triangular region in 3D space.

3. Normalization:

The constraint $x + y + z = 1$ is a normalization condition often used in epidemiological models. It ensures that the sum of proportions of the compartments (S, I, R) is constant and equal to 1, representing the entire population.

4. Parametrization of the 2-Simplex:

The 2-simplex can be parametrized as $(x, y, z) = (x, y, 1 - x - y)$, where x and y are parameters that vary within certain bounds (e.g., $x, y \geq 0$ and $x + y \leq 1$).

5. Projection onto (x, y) Plane (P):

To simplify the analysis, a projection onto the (x, y) plane is performed. This involves discarding the z -coordinate. The projection is typically denoted as $\{(x, y, 1 - x - y)\}$, representing points on the 2-simplex projected onto the (x, y) plane.

6. Significance of 2-Simplex and Projection:

The 2-simplex and its projection are significant because they reduce the dimensionality of the system, making it more manageable for analysis. The constraint $x + y + z = 1$ ensures that all points lie within a feasible region representing valid proportions of the population.

Let's define \mathcal{R}_0 the number of basic reproductions. Suppose that at the initial time $t = 0$ the whole population is healthy: $S(0) = N$. Now let's introduce a single infected individual into this same population.

The parameter $\mathcal{R}_0 \cdot x_0$ is the product of two essential terms in epidemiological modeling: \mathcal{R}_0 (the basic reproduction number) and x_0 (the initial proportion of the susceptible population). Its interpretation and control are crucial for understanding and managing an epidemic. Here is a more detailed discussion on these aspects:

1. Basic Reproduction Number (\mathcal{R}_0):

\mathcal{R}_0 represents the average number of secondary cases resulting from a single infection in a fully susceptible population. The higher \mathcal{R}_0 is, the more potential for rapid spread of the epidemic.

2. Initial Proportion of the Susceptible Population (x_0):

x_0 is the initial proportion of the population that is susceptible to infection. This depends on various factors such as control measures, pre-existing immunity, and individual behaviors.

3. Interpretation of $\mathcal{R}_0 \cdot x_0$:

$\mathcal{R}_0 \cdot x_0$ represents the expected number of secondary cases resulting from a single infection in a partially susceptible population. Its importance lies in its connection to the potential intensity of disease spread at the beginning of the epidemic.

4. Influence on the Actual Epidemic:

Control of \mathcal{R}_0 : Measures such as social distancing, vaccination, and hygiene practices can reduce \mathcal{R}_0 , thereby limiting the spread.

Influence of x_0 : Awareness campaigns, quarantine measures, and vaccination can influence the initial proportion of the susceptible population.

5. Control Strategies:

Reducing $\mathcal{R}_0 \cdot x_0$: By acting on both \mathcal{R}_0 and x_0 , health authorities can reduce the number of new infections, flatten the epidemic curve, and avoid healthcare system overload.

Understanding and controlling $\mathcal{R}_0 \cdot x_0$ are key elements for effective epidemic management. This requires targeted interventions to reduce transmission and increase immunity in the population.

This individual will infect the N other individuals at a rate $\frac{\beta}{N}S = \beta$ over a period of $1/\gamma \cdot \mathcal{R}_0 = \beta/\gamma$ healthy individuals will therefore be infected.

Eq. (3b) can then be written as [14],

$$\frac{dy}{dt} = \gamma (\mathcal{R}_0 x - 1) y$$

As $y \geq 0$ the epidemic will end when y cancels out. Note that $\frac{dy}{dt} < 0$. i.e., $y(t)$ decreases, as soon as $\mathcal{R}_0 x < 1$. But since $0 \leq x \leq 1$ this is equivalent to having $\mathcal{R}_0 < 1$. Note also that $\frac{dx}{dt} < 0 \forall t$ so that $x(t)$ decreases $\forall t$. Thus $y(t)$ can increase when $\mathcal{R}_0 x > 1$ reaches a maximum when $x(t) = 1/\mathcal{R}_0$ and then decreases again. The same applies to $z(t)$ which only increases.

The behavior of $x(t)$, $y(t)$ and $z(t)$ is summarized by Propositions 1 and 2.

Proposition 1. Let x , y and z solutions of the system (3a)-(3c). Then $\lim_{t \rightarrow \infty} y(t) = 0$, $\lim_{t \rightarrow \infty} x(t) = x_\infty$ and $\lim_{t \rightarrow \infty} z(t) = 1 - x_\infty$.

Demonstration:

Let $(x_0, y_0, 0)$ be any initial condition in \mathcal{P} such that $x_0 + y_0 = 1$. Let's start by showing that $x(t)$ has a limit for $t \rightarrow \infty$. We have that,

$$\frac{dx}{dt} = -\beta xy \geq -\beta x (x_0 + y_0)$$

because $y \leq x_0 + y_0 = 1$. i.e., $1 \geq x(t) \geq x_0 e^{-\beta(x_0+y_0)t} > 0$. $x(t)$ has a limit for $t \rightarrow \infty$. Note this limit x_∞ . Let us now show that $y(t)$ also has a limit.

$$\frac{dy}{dt} = \beta xy - \gamma y \geq -\gamma y$$

and therefore $1 \geq y(t) \geq y_0 e^{-\gamma t} > 0$. $y(t)$ therefore admits a limit which we note y_∞ . Necessarily, $z(t)$ also admits a limit which we note $z_\infty = 1 - x_\infty - y_\infty$.

Using that $\lim_{t \rightarrow \infty} z(t) = z_\infty$,

$$\lim_{t \rightarrow \infty} \frac{dz}{dt} = 0 = \gamma \lim_{t \rightarrow \infty} y(t)$$

and therefore, $\lim_{t \rightarrow \infty} y(t) = 0$ because $\gamma > 0$.

To assess the quantities x_∞ and z_∞ let's divide (3b) by (3a).

$$\frac{dy}{dx} = -1 + \frac{1}{\mathcal{R}_0 x}$$

In other words,

$$y = y_0 + x_0 - \frac{1}{\mathcal{R}_0} \ln x_0 - x + \frac{1}{\mathcal{R}_0} \ln x \quad (4)$$

with $x(t=0) = x_0$ and $y(t=0) = y_0$.

Using that $\lim_{t \rightarrow \infty} y(t) = 0$ we can obtain the asymptotic values x_∞ and $z_\infty = 1 - x_\infty$. Indeed, taking the limit $t \rightarrow \infty$ of Eq. (4), we obtain that x_∞ is a solution of the equation,

$$x_\infty - \frac{1}{\mathcal{R}_0} \ln x_\infty = y_0 + x_0 - \frac{1}{\mathcal{R}_0} \ln x_0$$

which has no closed-form solution.

In the sense of this proposition, the end of the epidemic is therefore asymptotic and does not correspond to the case where $y(t) = 0$. We'll come back to this later.

Proposition 2. Let x and y solution of (3a) and (3b) in \mathcal{P} .

If $\mathcal{R}_0 x_0 \leq 1$ then y immediately decreases towards 0 for $t \rightarrow \infty$.

If $\mathcal{R}_0 x_0 > 1$ then y grows until it reaches a maximum y_{\max} and then decreases.

Demonstration:

In fact, there's nothing to demonstrate. As previously mentioned, y grows as long as $\mathcal{R}_0 x > 1$, so in particular for $\mathcal{R}_0 x_0 > 1$.

In addition, $y(t) = y_{\max}$ for t such that $x(t) = 1/\mathcal{R}_0$.

Using Eq. (4), we have that,

$$y_{\max} = y_0 + x_0 - \frac{1}{\mathcal{R}_0} \ln x_0 - \frac{1}{\mathcal{R}_0} - \frac{1}{\mathcal{R}_0} \ln \mathcal{R}_0 = y_0 + x_0 - \frac{1}{\mathcal{R}_0} [1 + \ln x_0 \mathcal{R}_0]$$

Given \mathcal{R}_0 knowledge of x_0 allows us to predict whether the epidemic will continue or not. According to Proposition 2, if $\mathcal{R}_0 x_0 \leq 1$ there is no epidemic, as the number of infected individuals decreases towards zero. On the other hand, if $\mathcal{R}_0 x_0 > 1$ the epidemic becomes maximal in $x = 1/\mathcal{R}_0$. The parameter $\mathcal{R}_0 x_0$ is central to the control and dynamics of the model. SIR. As such, the parameter \mathcal{R}_0 parameter alone is not relevant to this model. Note that the number $\mathcal{R} = \mathcal{R}_0 x_0$ is identified in the literature as an effective reproduction number [15].

Figures 4a and 4b illustrate the phase portrait corresponding to the system (3a)-(3b) on \mathcal{P} for $\mathcal{R}_0 > 1$ and $\mathcal{R}_0 < 1$. The nullcline $x = \frac{1}{\mathcal{R}_0}$ is also shown. For $\mathcal{R}_0 < 1$ the nullcline no longer intersects \mathcal{P} . Figures 4c and 4d show the evolution of x, y and z as a function of t for $\mathcal{R}_0 x_0 > 1$ and $\mathcal{R}_0 x_0 \leq 1$.

In short, parameter \mathcal{R}_0 is a bifurcation parameter for the phase portrait of the (3a)-(3b) system. Parameter $\mathcal{R}_0 x_0$ is a bifurcation parameter for the evolution of $y(t)$.

The stability of the $(1, 0)$ The stability of the disease-free state [5] is also a question of interest in the epidemiological context. To this end, we carry out a stability analysis of the differential system (3a)-(3b). This system admits a fixed-point line given by the set of points $(x_*, 0)$. In particular, the point $(1, 0)$ is a fixed point of the system. The Jacobian matrix of the system evaluated in $(1, 0)$ is given by,

$$DJ_{(1,0)} = \begin{pmatrix} 0 & -\beta \\ 0 & \beta - \gamma \end{pmatrix}$$

and has eigenvalues 0 and $\beta - \gamma = \gamma(\mathcal{R}_0 - 1)$. The latter is the eigenvalue of interest and is associated with the eigenvector,

$$\begin{pmatrix} -\frac{\beta}{\beta-\gamma} \\ 1 \end{pmatrix} = \begin{pmatrix} -\frac{\mathcal{R}_0}{\mathcal{R}_0-1} \\ 1 \end{pmatrix}$$

Thus, if $\mathcal{R}_0 > 1$ the eigenvalue is positive and the eigenvector is oriented towards the x decreasing and y increasing. The state $(1, 0)$ is in this sense unstable and the epidemic is imminent. This is shown in Figure 4a. Conversely, if $\mathcal{R}_0 < 1$ the eigenvalue is negative, the state $(1, 0)$ is stable and there is no epidemic [16].

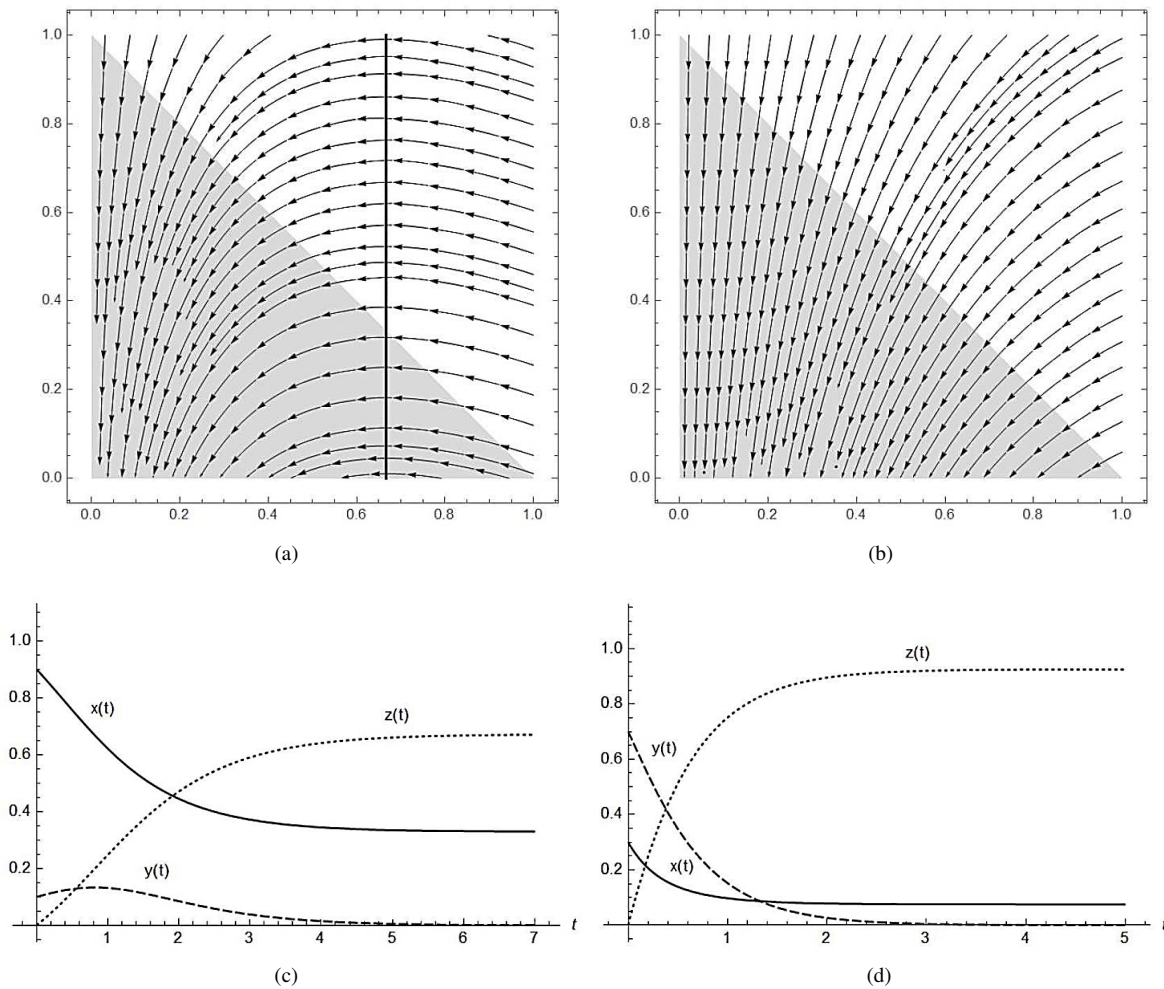


Figure 4. (a) Phase portrait for $\mathcal{R}_0 = \frac{3}{2} > 1, \gamma = 2$, (b) Phase portrait for $\mathcal{R}_0 = \frac{2}{3} < 1, \gamma = 3$, (c) Evolution of x, y and z for $\mathcal{R}_0 x_0 = \frac{3}{2} > 1, \gamma = 2$, (d) Evolution of x, y and z for $\mathcal{R}_0 x_0 = \frac{3}{2} < 1, \gamma = 2$

The stability analysis of the disease-free state in epidemiological models has significant practical implications for public health strategies and interventions. Here is a more detailed discussion of how the stability of this state translates into public health actions:

(1) Early Detection and Rapid Response:

A stable disease-free state suggests that, under certain conditions, the disease may not establish itself in the population. Early detection of outbreaks becomes crucial to prevent the transition to an endemic state. Public health strategies should focus on robust surveillance systems and rapid response mechanisms.

(2) Effective Vaccination Campaigns:

Stability analysis provides insights into the conditions for maintaining a disease-free state. This information is valuable for designing effective vaccination campaigns. Vaccination efforts can be targeted to enhance stability by ensuring a sufficiently high level of immunity in the population.

(3) Optimization of Resource Allocation:

Stability analysis helps understand the factors influencing disease dynamics. This knowledge enables the

optimization of resources, such as healthcare facilities, staff, and medical supplies. Resource allocation can be adjusted based on the stability conditions of the disease-free state.

(4) Adaptation of Public Health Measures:

Public health interventions, such as quarantine measures, social distancing, and hygiene practices, can be adapted based on stability conditions. For diseases with a stable disease-free state, interventions can focus on preventing reintroduction from external sources.

These practical implications highlight the importance of stability analysis in guiding evidence-based public health strategies and interventions to effectively manage and control infectious diseases.

4 SIR Stochastic Model of Epidemics: From Motivation to Gillespie Algorithm

The work is contributive in the field of epidemiological modeling, particularly through the use of stochastic approaches such as the Gillespie model and the Fokker-Planck equation to study the dynamics of infectious diseases. These approaches allow for the consideration of uncertainty and random fluctuations, providing a realistic perspective on the spread of diseases.

Specifically, the use of the Gillespie model to simulate stochastic jump processes in the context of epidemics appears to be a significant contribution. Additionally, the analysis of the distribution of extinction time provides valuable insights into the dynamics of infectious diseases and may have important implications for public health.

Regarding improvements over existing models, the stochastic approach adopted here enables the capture of random variations in disease spread, which can be crucial in real-world situations. Furthermore, the transition to a Fokker-Planck equation in the continuous limit offers an analytical perspective on system behavior, which can be advantageous compared to purely numerical approaches.

It would also be relevant to highlight how these new approaches can be applied to real epidemiological data and how they compare to more traditional deterministic models in terms of accuracy and predictive ability.

The work brings forth original ideas and extends the scope of epidemiological modeling by introducing stochastic elements and in-depth analytical analysis.

Complex systems are made up of a multitude of elements interacting according to the laws of the microscopic and macroscopic worlds. Therefore, it is totally unrealistic to model all these interactions. Predictions of the behavior of such systems are severely limited. The notion of fluctuation also comes into play: the evolution of a given observable cannot be rigorously monotonic, i.e., smooth. The large number of interactions necessarily diverges the evolution of the observable from that predicted by the equivalent deterministic model. It is in this sense that stochastic calculus takes on its full importance in the study of complex systems. In short, the emergence of a random dimension is based on the impossibility of knowing the behavior of a system down to the smallest detail. This lack of information can be modeled as noise with a given, postulated probability distribution [17].

This provides an initial motivation for the stochastic approach to epidemiological models. Moreover, stochastic models are important when the number of affected individuals is small compared with the population size or when transmission, infection, births, and deaths are highly variable. Other global variations, such as environmental ones, could also be considered [18].

Yes, the stochastic SIR model and the Gillespie algorithm are likely to have a substantial impact on the understanding of epidemic spread. Here are some reasons why these approaches are important:

(1) Incorporation of Uncertainty and Fluctuations: Unlike deterministic models, the stochastic SIR model takes into account the inherent uncertainty in epidemiological processes. It models random fluctuations in disease transmission, which is crucial in real-world scenarios where individual interactions cannot always be predicted with certainty.

(2) Realistic Simulation of Jump Processes: The Gillespie algorithm provides a realistic simulation approach for stochastic jump processes associated with disease spread. It captures discrete events such as contacts between individuals, infections, and recoveries, which can have a significant impact on epidemic dynamics.

(3) Analysis of Temporal Distributions: The use of stochastic models allows the analysis of temporal distributions, such as the distribution of the extinction time of the infection. This provides valuable insights into the temporal variability of epidemics and can help better understand factors influencing the duration of their spread.

(4) Consideration of Small Populations: Stochastic models, like the Gillespie model, are particularly useful when the number of affected individuals is small compared to the total population size. They can better represent situations where rare events have a significant impact.

(5) Complementary Approach to Deterministic Models: While deterministic models are useful for understanding general trends, stochastic models provide a complementary perspective by taking into account stochastic aspects and individual variations.

The stochastic SIR model and the Gillespie algorithm help us understand how epidemics spread better and more realistically by taking into account the unknown and inherent differences in how people interact and how epidemics happen.

The article establishes a connection between the mathematical development of the stochastic SIR model, the Gillespie algorithm, and the practical implications in the context of infectious disease spread. Here's how this work enhances our understanding of infectious disease propagation:

(1) Precision in Modeling: The stochastic SIR model captures random fluctuations and uncertainties inherent in individual interactions and epidemiological events. This enhanced precision contributes to a more realistic modeling of the spread processes.

(2) Realistic Simulation of Epidemic Processes: The Gillespie algorithm can be used to simulate stochastic jump processes, which gives a realistic picture of discrete epidemiological events like people coming into contact with each other, getting sick, and getting better. This enables an in-depth understanding of epidemic dynamics.

(3) Analysis of Temporal Distribution: Analyzing the distribution of infection extinction times provides crucial insights into the temporal variability of epidemics. Understanding the duration of spread and factors influencing extinction time contributes to better epidemic management.

(4) Study of Small Populations: The stochastic model is particularly relevant when the number of affected individuals is small compared to the total population size. This extends the model's applicability to realistic situations where rare events can have a significant impact.

(5) Integration of Stochastic and Deterministic Analysis: The stochastic approach complements deterministic models by offering a perspective that considers individual variations and uncertainties. This integration allows for a more holistic understanding of disease spread.

(6) Implications for Public Health: By better understanding epidemic dynamics through the stochastic model, health authorities can make more informed decisions regarding control strategies, resource allocation, and preventive interventions.

Thus, the work enriches our understanding by providing advanced mathematical tools while demonstrating how these tools can be applied to draw meaningful conclusions about infectious disease spread in the real world.

Nor is it more natural to define a probability of infection between two individuals rather than a mechanism by which infection does or does not occur? Stochastic calculus then offers the possibility of assigning uncertainties to the mean values of observables or to estimates [19].

Markov jump process: The variables $S(t)$, $I(t)$ and $R(t)$ are now considered as discrete random variables that can take an integer between 0 and N . The state space is therefore \mathbb{N}^3 . As always, $S(t) + I(t) + R(t) = N$ stochastic dynamics must preserve the number of individuals in the population. In this sense, the study can be limited to the variables $S(t)$ and $I(t)$.

The SIR model is now considered as a Markov jump process precisely defined as [20],

$$(S = m, I = n) \longrightarrow (m - 1, n + 1) \text{ avec taux } \frac{\beta}{N}mn \longrightarrow (m, n - 1) \text{ avec taux } \gamma n$$

An immediate consequence is that extinction is no longer asymptotic and corresponds to $I(t) = 0$. This Markov process is assumed to be continuous-time. If the process were discrete-time, we would have had to consider a jump corresponding to the absence of a transition [21].

The time evolution of the probability distribution $P(S(t) = m, I(t) = n) \stackrel{\text{def}}{=} P_t(m, n)$ is given by the master equation (or Kolmogorov equation). To obtain the expression of the master equation, we need to look at the jumps that lead the system to the state (m, n) and those that lead away from it.

$$\begin{aligned} (m + 1, n - 1) &\rightarrow (m, n) \text{ avec taux } \frac{\beta}{N}(m + 1)(n - 1) \\ (m, n + 1) &\rightarrow (m, n) \text{ avec taux } \gamma(n + 1) \end{aligned}$$

The master equation is written as the sum of gain and loss terms. The gain terms are the positive contributions to the master equation associated with transitions to state (m, n) . Loss terms, on the other hand, are the negative contributions associated with transitions from state (m, n) . In this sense, we have that,

$$\begin{aligned} \partial_t P_t(m, n) &= \frac{\beta}{N} [(m + 1)(n - 1)P_t(m + 1, n - 1) - mnP_t(m, n)] \\ &+ \gamma [(n + 1)P_t(m, n + 1) - nP_t(m, n)] \end{aligned} \quad (5)$$

This equation is the basis of all stochastic analysis. Given an initial condition ($S_0 = m_0, I_0 = n_0$) if we can solve the master equation, we obtain the propagator $P_t(m, n | m_0, n_0) \forall t$. Although it is relatively simple to obtain the master equation of a given jump process, solving the equation is generally non-trivial, if not impossible as in this case.

The question now is to understand how to simulate such a Markov jump process so as to be able to obtain numerical results.

Gillespie algorithm. The Gillespie algorithm, introduced by D. T. Gillespie in 1977, is a simple algorithm for simulating any jump process [22].

Consider a scheme of p reactions with respective reaction rates α_j for $j = 1, \dots, p$. Assume that this reaction scheme involves q entities. In the case of the SIRmodel, we would have $p = 2$ and $q = 3$. Let h_j be the number of possible combinations of reactants in the reaction j . For example, if the reaction j is of the form $X + Y \rightarrow$ something, then $h_j = XY$. If it is of the form $X + X + X \rightarrow$ something, then $h_j = \binom{X}{3}$. The algorithm first requires the maximum process time t_{\max} of the process, the reaction rates α_j and the initial number of entities $X_{i,0}$. Gillespie's algorithm is shown below.

We'd now like to justify this algorithm. To do this, let's assume that at time t the system is in state (X_1, \dots, X_q) . Two questions arise: when will the next reaction take place? What will be the next reaction to take place? To this end, let's note $P(\tau, \mu)d\tau$ the probability that the next reaction will take place within the time interval $[t + \tau, t + \tau + d\tau]$ and be the reaction μ with $0 \leq \tau < \infty$ and $\mu = 0, \dots, p$. We have that,

- $\alpha_\mu dt \stackrel{\text{def}}{=} \text{probability that a particular combination of reaction reactants } \mu \text{ gives rise to the reaction } \mu \text{ between } t \text{ and } t + dt.$
- $h_\mu \alpha_\mu dt \stackrel{\text{def}}{=} \text{probability of the reaction } \mu \text{ between } t \text{ and } t + dt \text{ knowing that at time } t \text{ the system was in state } (X_1, \dots, X_q).$

Gillespie algorithm [23]:

input: $t_{\max}, X_{0,i} (i = 1, \dots, q)$ and $\alpha_j (j = 1, \dots, p)$
 procedure GILLESPIE ($t_{\max}, X_{0,i}, \alpha_j$)
 $X_i \leftarrow X_{i,0} (i = 1, \dots, q)$
 $t \leftarrow 0$
 while $t < t_{\max}$ do
 for $k \leftarrow 1$ to p do
 $a_k \leftarrow h_k \alpha_k$
 $a_0 \leftarrow \sum_{k=1}^p a_k$
 $r_1 \leftarrow U([0, 1])$
 $r_2 \leftarrow U([0, 1])$
 $\tau \leftarrow \frac{1}{a_0} \ln \frac{1}{r_1}$
 $\sum_{k=1}^{\mu-1} a_k < r_2 a_0 \leq \sum_{k=1}^{\mu} a_k$ Choose μ verifying condition
 $t \leftarrow t + \tau$
 $f(\mu)$ f modifies the value of X_i according the reaction scheme
 return $X_i (i = 1, \dots, q)$

It is then possible to calculate $P(\tau, \mu)$ as $P(\tau, \mu)d\tau = P_0(\tau)a_\mu d\tau$ with $P_0(\tau)$ the probability of no reaction in the interval $[t, t + \tau]$ from state (X_1, \dots, X_q) . Note that $a_\mu \stackrel{\text{de}}{=} h_\mu \alpha_\mu$. Then,

$$P_0(\tau + d\tau) = P_0(\tau) \left[1 - \sum_{k=1}^p a_k d\tau \right]$$

Indeed, the term in square brackets is precisely the probability that there will be no reaction between $t + \tau$ and $t + \tau + d\tau$. In other words,

$$\frac{dP_0(\tau)}{d\tau} = - \sum_{k=1}^p a_k P_0(\tau) = -a_0 P_0(\tau)$$

Or else,

$$P(\tau, \mu) = a_\mu e^{-a_0 \tau} \quad (6)$$

Note that so far we haven't made any approximations, so the expression for the joint distribution $P(\tau, \mu)$ is, for the moment, exact.

Now we just need to understand how to sample a pair of random variables (τ, μ) according to distribution (6). In theory, this could be done exactly, provided we have an ideal uniform random number generator (URNG). In practice, computers can only sample psuedo-uniform random numbers (GNPAU). Let's generate two GNPAUs r_1 and r_2 such that,

$$r_1, r_2 \sim U([0, 1])$$

Writing that $r_1 = e^{-a_0\tau}$ we obtain that $\tau = \frac{1}{a_0} \ln \frac{1}{r_1}$. Next, we need to write an expression that can give us the value of μ for a certain frame of r_2 . Starting from the relation $\frac{1}{a_0} \sum_{k=1}^p a_k$ we can choose μ verifying,

$$\sum_{k=1}^{\mu-1} a_k < r_2 a_0 \leq \sum_{k=1}^{\mu} a_k$$

Gillespie's algorithm is now proven.

Figure 5 shows a numerical implementation of Gillespie's algorithm for the (1a) -(1b) jump process.

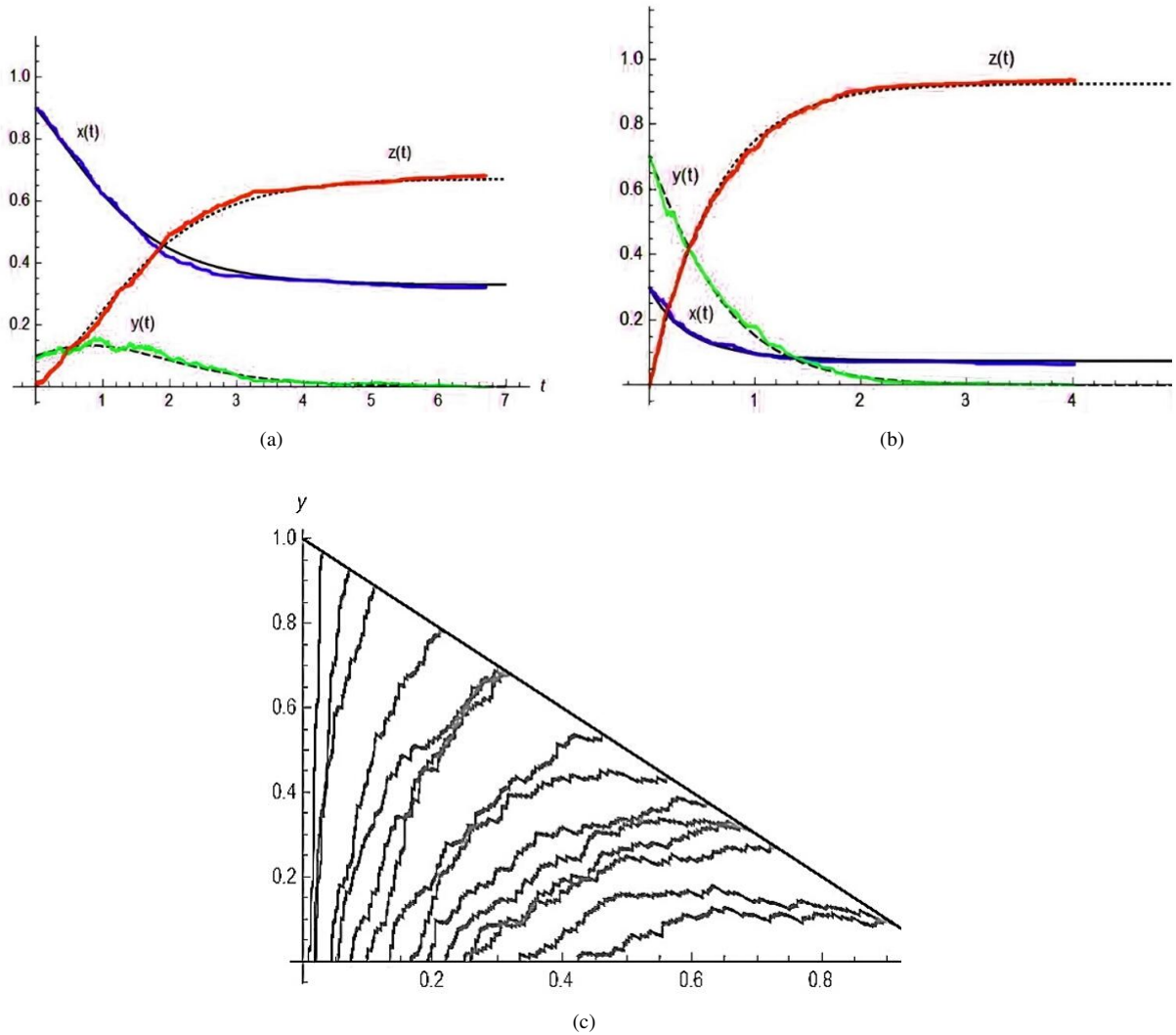


Figure 5. (a) Gillespie algorithm implemented for the evolution of x, y and z on a population of 500 individuals $\mathcal{R}_0 x_0 = \frac{3}{2} \frac{9}{10} > 1, \gamma = 2$. (b) Gillespie algorithm implemented for the evolution of x, y and z on a population of 500 individuals. $\mathcal{R}_0 x_0 = \frac{3}{2} \frac{3}{10} > 1, \gamma = 2$. (c) Gillespie algorithm implemented for the evolution of y as a function of x on a population of 500 individuals. The initial conditions have been randomly chosen on the right-hand side, $x_0 + y_0 = 1$

Continuous limit and Fokker-Planck equation [24]. The idea here is to look at the master equation in the continuous limit in order to obtain the Fokker-Planck equation associated with the process. Recall that the Fokker-Planck equation is the equation that must be satisfied by the probability density of a system described by a continuous-time Markov process. As a result, it will be possible to associate a stochastic differential equation in the sense of Itô (or possibly in the sense of Stratonovich) [25] with the Markov jump process [20] and integrate its trajectories to obtain stochastic trajectories.

Let's put it this way,

$$\begin{cases} x = m/N \\ y = n/N \\ \varepsilon = 1/N \end{cases}$$

The continuous limit corresponding to $N \rightarrow \infty \Leftrightarrow \varepsilon \rightarrow 0$. The Taylor series expansion of the master equation is called the Kramers-Moyal expansion. Developing with respect to the small parameter ε to order 2, we get,

$$\begin{aligned} P_t(m, n+1) &= P_t(xN, yN+1) = P_t\left(xN, N\left(y + \frac{1}{N}\right)\right) \\ &= \mathcal{P}(x, y + \varepsilon, t) = \mathcal{P}(x, y, t) + \varepsilon \partial_x \mathcal{P}(x, y + \varepsilon, t) \\ &\quad + \frac{\varepsilon^2}{2} \partial_x^2 \mathcal{P}(x, y + \varepsilon, t) + o(\varepsilon^3) \end{aligned}$$

where, we have defined $\mathcal{P}(x, y, t) \stackrel{\text{def}}{=} P_t(xN, yN)$. The same applies to $P_t(m+1, n-1)$,

$$\begin{aligned} P_t(m+1, n-1) &= \mathcal{P}(x, y, t) + \varepsilon \partial_x \mathcal{P}(x, y + \varepsilon, t) - \varepsilon \partial_y \mathcal{P}(x, y + \varepsilon, t) \\ &\quad + \frac{\varepsilon^2}{2} \partial_x^2 \mathcal{P}(x, y + \varepsilon, t) + \frac{\varepsilon^2}{2} \partial_y^2 \mathcal{P}(x, y + \varepsilon, t) - \varepsilon^2 \partial_x \partial_y \mathcal{P}(x, y, t) + o(\varepsilon^3) \end{aligned}$$

Now we need to replace the developments of $P_t(m, n+1)$ and $P_t(m+1, n-1)$ in the master Eq. (5). In the following, we will implicitly note $\mathcal{P} \stackrel{\text{def}}{=} \mathcal{P}(x, y, t)$. Restricting ourselves to terms of order ε^2 the calculation yields,

$$\begin{aligned} \partial_t \mathcal{P} &= \beta \left\{ xy (\partial_x - \partial_y) + \frac{\varepsilon}{2} xy (\partial_x^2 + \partial_y^2 - 2\partial_x \partial_y) \right. \\ &\quad \left. - (x - y) - \varepsilon(x - y) (\partial_x - \partial_y) - \varepsilon \right\} \mathcal{P} + \gamma \left\{ y \partial_y + \frac{\varepsilon}{2} y \partial_y^2 + 1 + \varepsilon \partial_y \right\} \mathcal{P} \end{aligned}$$

where, ε has been highlighted in the first term to simplify with N . Next, we need to group terms of the same order into ε .

Terms in ε^0 ,

$$\beta \{ xy (\partial_x - \partial_y) - (x - y) \} \mathcal{P} + \gamma \{ y \partial_y + 1 \} \mathcal{P} = -\partial_x \{ -\beta xy \} \mathcal{P} - \partial_y \{ \beta xy - \gamma y \} \mathcal{P}$$

In square brackets, we recognize the expressions (3a) and (3b) for $\frac{dx}{dt}$ and $\frac{dy}{dt}$. The operator ∂ operator applies not only to the terms in square brackets, but also to \mathcal{P} .

Terms in ε^1 ,

$$\begin{aligned} &\beta \left\{ \frac{1}{2} xy (\partial_x^2 + \partial_y^2 - 2\partial_x \partial_y) - (x - y) (\partial_x - \partial_y) - 1 \right\} \mathcal{P} + \gamma \left\{ \frac{1}{2} y \partial_y^2 + \partial_y \right\} \mathcal{P} \\ &= \varepsilon \beta \left\{ \partial_x^2 \frac{xy}{2} + \partial_y^2 \frac{xy}{2} - 2\partial_x \partial_y xy \right\} \mathcal{P} + \varepsilon \gamma \left\{ \partial_y^2 \frac{y}{2} \right\} \mathcal{P} \end{aligned}$$

Now let's define the vector $\mu(x, y)$ and the diffusion matrix $\mathbf{D}(x, y)$ of size 2×2 such that,

$$\begin{aligned} \mu(x, y) &= \begin{pmatrix} -\beta xy \\ \beta xy - \gamma y \end{pmatrix} \\ \mathbf{D}(x, y) &= \frac{\varepsilon}{2} \begin{pmatrix} \beta xy & -\beta xy \\ -\beta xy & \beta xy + \gamma y \end{pmatrix} \end{aligned}$$

We have shown that the time evolution of the probability distribution $\mathcal{P}(x, y, t)$ verifies the Fokker-Planck equation,

$$\begin{aligned} \partial_t \mathcal{P}(x, y, t) &= \partial_x \{ -\mu_1(x, y) + \partial_x \mathbf{D}_{11}(x, y) + \partial_y \mathbf{D}_{12}(x, y) \} \mathcal{P}(x, y, t) \\ &\quad + \partial_y \{ -\mu_2(x, y) + \partial_x \mathbf{D}_{21}(x, y) + \partial_y \mathbf{D}_{22}(x, y) \} \mathcal{P}(x, y, t) \end{aligned}$$

which is a bivariate Fokker-Planck equation.

Finally, it is possible to associate a stochastic differential equation in the sense of Itô with this Fokker-Planck equation. To do this, recall that the generic Fokker-Planck equation,

$$\partial_t \mathcal{P}(x) = \partial_x \left\{ -f(x, t) + \frac{1}{2} \partial_x g(x, t)^2 \right\} \mathcal{P}(x, t)$$

which is associated with the following stochastic differential equation in the sense of Itô,

$$dx = f(x, t)dt + g(x, t)dW$$

with $dW(t)$ an infinitesimal realization of the Wiener process. In the bivariate case, this result transposes naturally to the following stochastic differential equation [26],

$$d \begin{pmatrix} x \\ y \end{pmatrix} = \mu(x, y)dt + \sigma(x, y)d \begin{pmatrix} W_1(t) \\ W_2(t) \end{pmatrix} \quad (7)$$

with $\sigma(x, y)$ such that $D(x, y) = \frac{1}{2}\sigma(x, y)\sigma(x, y)^\top$. The expression for $\sigma(x, y)$ as it is too long.

Note that it is the diffusive terms that separate the stochastic trajectory from the deterministic one. These terms are proportional to the inverse of the system size, so the smaller the system, the greater the fluctuations around the deterministic trajectory. Figures 6a and 6b illustrate three stochastic trajectories obtained by integrating (7) in the Itô sense for the same initial condition.

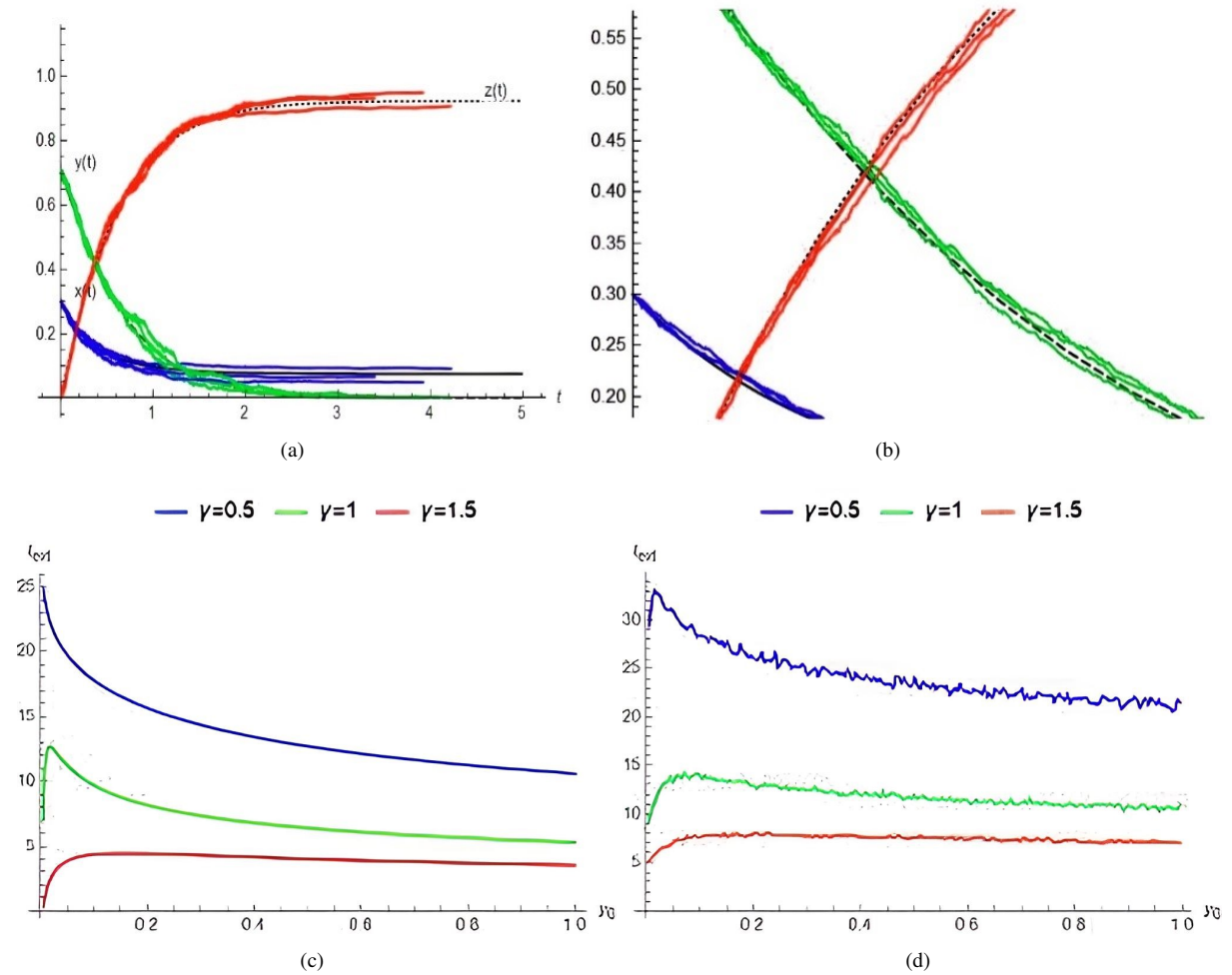


Figure 6. (a) Evolution of x, y and z obtained by integrating Itô's stochastic differential equation on a population of 200 individuals. $\mathcal{R}_0 x_0 = \frac{3}{2} \frac{3}{10} > 1, \gamma = 2$. (b) Zooming in on a certain region to highlight the three different integrations. (c) Evolution of t_{ext} as a function of y_0 and γ calculated from a population of 200 individuals using the deterministic model. (d) Evolution of t_{ext} as a function of y_0 and γ calculated from a population of 200 individuals using the stochastic model

Distribution of extinction time. We would now like to consider the following question: given a number Y_0 (or concentration y_0) of infected individuals, what is the probability that the extinction time lies within the interval $[t_{\text{ext}}, t_{\text{ext}} + dt]$? In other words, what is the average time that will have elapsed between the moment when infected individuals are introduced into the healthy population and the moment when the population no longer contains any infected individuals?

Since the SIR model is non-linear, the analysis of the extinction process from a stochastic point of view is nontrivial. In the following [4], let's note t_{ext} this extinction time. We'll also set $\beta = 1$ this being of little importance, given that the bifurcation parameter is $\mathcal{R}_0 x_0 = \frac{1}{\gamma} (1 - y_0)$.

Note that t_{ext} must depend on the initial condition. To gain some intuition, let's start by studying the extinction time as a function of the initial condition y_0 and parameter γ in the light of the deterministic model. To do this, simply integrate the system of coupled differential Eq. (3a) and (3b) until $y(t) < \frac{1}{N}$ for a population of N individuals. t_{ext} is then such that $y(t_{\text{ext}}) = \frac{1}{N}$ because by Proposition 1, y is a decreasing function of t which ensures that for $t > t_{\text{ext}}$, $y(t) < \frac{1}{N}$ corresponding to no more infected individuals.

The Figure 6c shows the evolution of extinction time t_{ext} as a function of y_0 and γ . We have set $N = 200$ although the result does not depend qualitatively on N .

These three curves seem to illustrate an antagonistic behavior. Let's take a look at each case. For $\gamma = 1.5$, $\mathcal{R}_0 x_0 = \frac{2}{3} (1 - y_0)$. the threshold $\mathcal{R}_0 x_0 = 1$ is reached for a value $y_0 = -0.5$ which is not possible here.

In this case, all the initial conditions correspond to a situation without epidemics. In this sense, it's quite intuitive to understand that the smaller the number of infected individuals introduced into the population, the shorter the extinction time.

When $\gamma = 0.5$, the situation becomes epidemic for $y_0 < 0.5$. Then, when y_0 decreases towards zero, the maximum reached by y , y_{max} increases, which explains why the extinction time increases. On the other hand, when $y_0 > 0.5$ and y_0 increases towards $y_0 = 1$, the extinction time decreases, which is the opposite of the behavior in the case where $\gamma = 1.5$.

Finally, when $\gamma = 1$, the variation in extinction time is surprising. It increases until it reaches a maximum, then decreases. The maximum corresponds to a very small number of infected individuals, around 0.015 in this case. We'll come back to this case a little later.

Analysis of the extinction time from deterministic equations allows us to introduce, albeit roughly, the behavior of the mean value of the extinction time within the framework of the stochastic process. The mean extinction time can be calculated as a function of y_0 and γ in the same way as for the deterministic model. The result is shown in Figure 6d. To do this, for each value of y_0 (in steps of 0.005) and γ , 1000 jump processes were simulated using Gillespie's algorithm, and the extinction times obtained and averaged.

Our next question is whether it is possible to express the extinction time distribution analytically.

Several methods have been proposed in the literature to study the mean extinction time and its distribution. These include the method of Allen [27] and the method of Daley and Gani [28]. Lefèvre [29] provides a list of other methods in use.

Barbour's theorem: In the following, we would like to examine in greater detail a result obtained by Barbour [30]. This result concerns the distribution of the extinction time when the fraction of infected individuals at the initial time is not too small compared with the fraction of healthy individuals.

Before introducing the Theorem in question, we need to discuss the Gumbel distribution. The Gumbel distribution is a special case of the generalized extremum law, whose cumulative function is formally written as,

$$F(x, \mu, \sigma) = e^{-e^{\frac{\mu-x}{\sigma}}}$$

with $-\infty < x < \infty$ and whose probability distribution is,

$$X \sim \text{Gumbel}(\mu, \sigma) \Leftrightarrow p(X = x; \mu, \sigma) = \frac{e^{-z} z}{\sigma}$$

with $z \stackrel{\text{def}}{=} e^{\frac{\mu-x}{\sigma}}$. In the following, only the standard Gumbel distribution ($\mu = 0$ and $\sigma = 1$) will be of interest.

As a distributor, $\int_{-\infty}^{\infty} dx e^{-e^{-x}-x} = 1$. In addition, $\mathbb{E}[X = x] = \int_{-\infty}^{\infty} dx x e^{-e^{-x}-x} = \gamma \simeq 0.577 \dots$ is the EulerMascheroni constant. Figure 7 illustrates the standard Gumbel distribution.

Let's take a population of size $X_0 + Y_0 = N + I(N)$, i.e., $x_0 + y_0 = 1 + h$ with $h \stackrel{\text{def}}{=} \frac{I}{N} = c^{\text{ste}} \forall N$ and independent of N . Let's note T_N the extinction time. Barbour's Theorem is then written as,

Theorem. Let $W \sim \text{Gumbel}(0, 1)$. Then, for $N \rightarrow \infty$, T_N converges in distribution to,

$$T_N \rightarrow \frac{1}{(\gamma - x_{\infty})} (W + \ln N + k)$$

with k such that,

$$k = \lim_{m \rightarrow \infty} \left\{ -\ln m + (\gamma - x_{\infty}) J \left[x_{\infty} \left(1 + m^{-1} (\gamma - x_{\infty})^{-1} \right), 1 \right] + \ln \left(1 - \frac{x_{\infty}}{\gamma} \right) \right\}$$

and $J[a, b]$ such that,

$$J[a, b] = \int_a^b \frac{du}{u(1 + h + \gamma \ln u - u)}$$

Let's compare the distribution T_N given by the Theorem with a histogram generated by Gillespie's algorithm.

Figure 8 illustrates the distribution T_{5000} obtained from the Theorem, onto which we have superimposed a histogram obtained from several 10^3 runs of Gillespie's algorithm.

The distribution T_N overlaps very reasonably with the histogram.

When the number of infected individuals introduced into the population is very small compared to the number of healthy individuals, the above Theorem is no longer valid. Indeed, for Y_0 typically of the order of 1, a significant proportion of the stochastic trajectories cancel out immediately (giving a very small extinction time), while another proportion approaches $Y = 0$ a first time at the initial time and a second time at the extinction time. This marginal case will not be dealt with here and is illustrated in Figure 9.

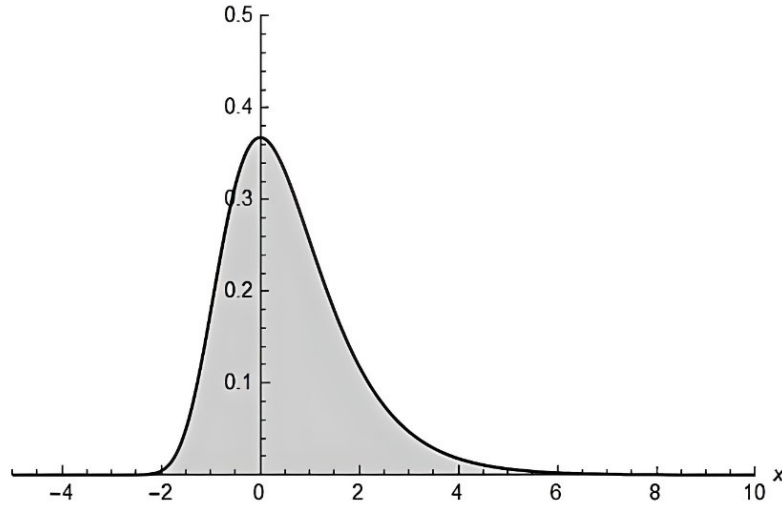


Figure 7. Standard Gumbel distribution

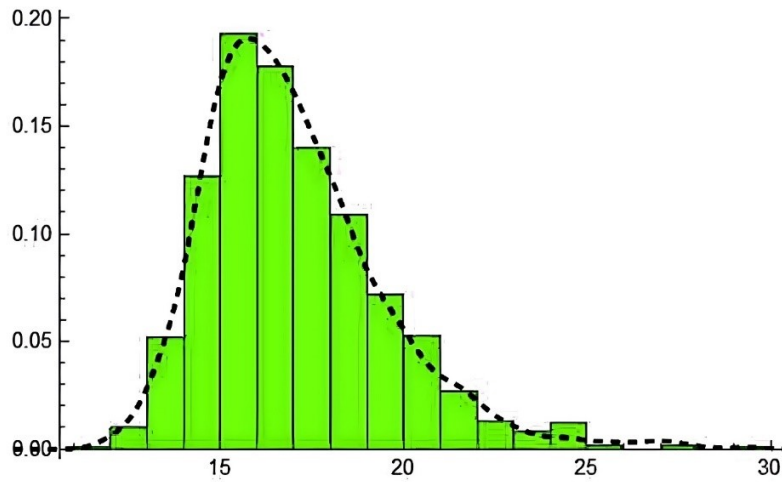


Figure 8. Dotted line: distribution T_{5000} calculated from the Theorem. In green: histogram obtained by 10^3 runs of Gillespie's algorithm. $h = 0.3$ and $\gamma = 0.8$

Proof of this Theorem is based on a comparison of the leap process leading to extinction with a birth-and-death process. We will not demonstrate this Theorem here. However, we would like to discuss the emergence of the Gumbel distribution from a birth-and-death jump process. There are two processes that modify the number of infected individuals in the population. On the one hand, process (1a) increases the number of infected individuals by one, $I \rightarrow I + 1$ and secondly, process (1b), which decreases the number of infected individuals by one unit

$I \rightarrow I - 1$ each with a corresponding reaction rate. This binary vision of the process leading to extinction suggests an analogy with a process of birth and death.

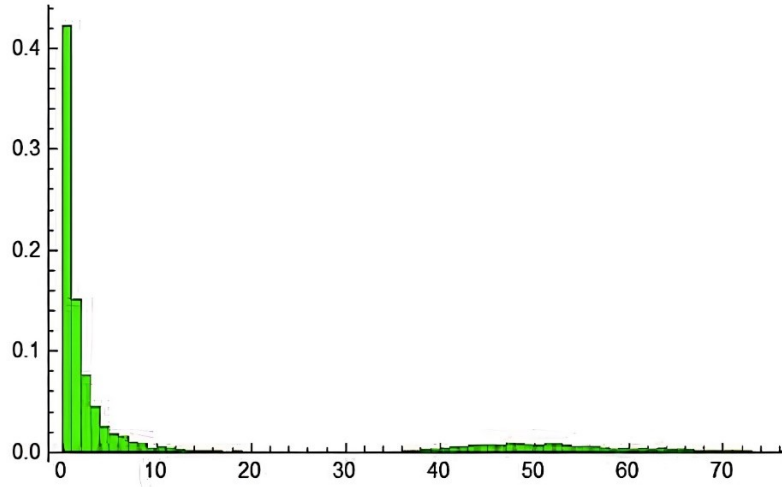


Figure 9. Histogram for $X_0 = 5000$ and $Y_0 = 1$ for $\gamma = 0.8$. A significant proportion of the trajectories undergo almost immediate extinction, while another proportion evolve in the vicinity of $Y = 0$ without being extinguished

In other words, a birth and death process with respective reaction rates of a and b . In the following, we assume $a < b$. The jump process is then as follows,

$$\begin{aligned} n &\rightarrow n + 1 \text{ avec taux } an \\ n &\rightarrow n - 1 \text{ avec taux } bn \\ n - 1 &\rightarrow n \text{ avec taux } a(n - 1) \\ n + 1 &\rightarrow n \text{ avec taux } b(n + 1) \end{aligned}$$

from which we can immediately deduce the expression of the master equation,

$$\partial_t P_t(n) = a(n - 1)P_t(n - 1) + b(n + 1)P_t(n + 1) - (a + b)nP_t(n)$$

with n which represents the number of infected individuals and therefore cannot be negative. The master equation is therefore problematic in $n = 0$ and the edge condition should be chosen [18],

$$\begin{aligned} \partial_t P_t(0) &= bP_t(1) \\ P_t(n) &= 0, \forall n < 0 \end{aligned}$$

Such a master equation is often studied using the generating function $G = G(z, t)$ such as $G(z, t) = \sum_{n=0}^{\infty} P_t(n)z^n$. In particular,

$$\begin{aligned} \frac{\partial G}{\partial t} &= \sum_{n=0}^{\infty} \partial_t P_t(n)z^n = \sum_{n=0}^{\infty} \{a(n - 1)P_t(n - 1) + b(n + 1)P_t(n + 1) \\ &\quad - (a + b)nP_t(n)\} z^n = \sum_{n=0}^{\infty} \{az^2 - (a + b)z + b\} nP_t(n)z^{n-1} \\ \frac{\partial G}{\partial t} &= \Phi(z) \frac{\partial G}{\partial z} \end{aligned}$$

with $\Phi(z) \stackrel{\text{def}}{=} az^2 - (a + b)z + b$ and where edge conditions are used. In order to solve this first-order partial differential equation, we look for a path $z = z(s), t = t(s)$ along which Eq. (8) is verified [31]. The differential of $G(z(s), t(s)) \stackrel{\text{def}}{=} G(s)$ is written as,

$$\frac{dG(s)}{ds} = \frac{\partial G}{\partial z} \frac{dz}{ds} + \frac{\partial G}{\partial t} \frac{dt}{ds}$$

From this it follows that,

$$\begin{aligned}\frac{dz}{ds} &= \Phi(z) \\ \frac{dt}{ds} &= -1\end{aligned}$$

The general solution to this first equation is written as,

$$z(t) = \frac{e^{-bt+ac_1} - be^{-at+bc_1}}{e^{-bt+ac_1} - ae^{-at+bc_1}}$$

with $c_1 \in \mathbb{R}$ and where we have chosen $t = -s$ as the solution to the second equation. Let's express $e^{c_1(b-a)} \stackrel{\text{def}}{=} c_2$ in terms of z ,

$$\frac{z-1}{az-b} e^{(a-b)t} = c_2$$

By virtue of (8), we have that $\frac{dG}{ds} = 0 \Leftrightarrow G(s) = f(c_2)$. It is possible to find an expression for f by imposing that at the initial time $t = 0$ there are exactly n_0 individuals. Then $P_{t=0}(n) = \delta_{n,n_0}$ and $G(z, t = 0) = \sum_{n=0}^{\infty} P_{t=0}(n) z^n = z^{n_0}$. f verifies,

$$G(z, 0) = f\left(\frac{z-1}{az-b}\right) = z^{n_0}$$

that is,

$$f(x) = \left[\frac{bx-1}{ax-1} \right]^{n_0}$$

So,

$$G(z, t) = \left\{ \frac{b(z-1)e^{(a-b)t} - az + b}{a(z-1)e^{(a-b)t} - az + b} \right\}^{n_0}$$

note that $G(1, t) = \sum_{n=0}^{\infty} P_t(n) = 1$.

What's the point of all this reasoning? The probability that the extinction time t_{ext} is less than or equal to t is given by,

$$P(t_{\text{ext}} \leq t) = P_t(n = 0) = G(z = 0, t)$$

and therefore,

$$P(t_{\text{ext}} = t) = \frac{d}{dt} G(z = 0, t) = n_0 \frac{(a-b)^2}{b^2} \frac{(1 - e^{(a-b)t})^{n_0-1}}{(1 - \frac{a}{b} e^{(a-b)t})^{n_0+1}} e^{(a-b)t}$$

The question is to understand where the Gumbel distribution emerges from this expression.

Heuristically, two regimes can be identified. During the first regime, a set of stochastic trajectories originating from the same initial condition roughly follow the deterministic trajectory. In the second regime, stochastic effects become more important, and the time.

Heuristically, it is appropriate to identify two regimes. During the first regime, a set of stochastic trajectories from the same initial condition closely follow the deterministic trajectory. During the second regime, stochastic effects become important and the extinction time becomes fully random. Figure 10 illustrates this phenomenon.

Let n_* denote the separation between the two regimes. If $n_0 \gg n_*$, then $t_{\text{ext}} = t \gg 1$, so that $e^{(a-b)t} \ll 1$, meaning that,

$$\left(1 - e^{(a-b)t}\right)^{(n_0-1)} \approx 1 - (n_0 - 1) e^{(a-b)t} \approx \exp\left(-n_0 e^{(a-b)t}\right)$$

In summary, it has been shown that,

$$P(t_{\text{ext}} = t) \propto \exp\left[\left(\frac{a}{b} - 1\right) n_0 e^{(a-b)t}\right]$$

which leading to the Gumbel distribution of the extinction time.

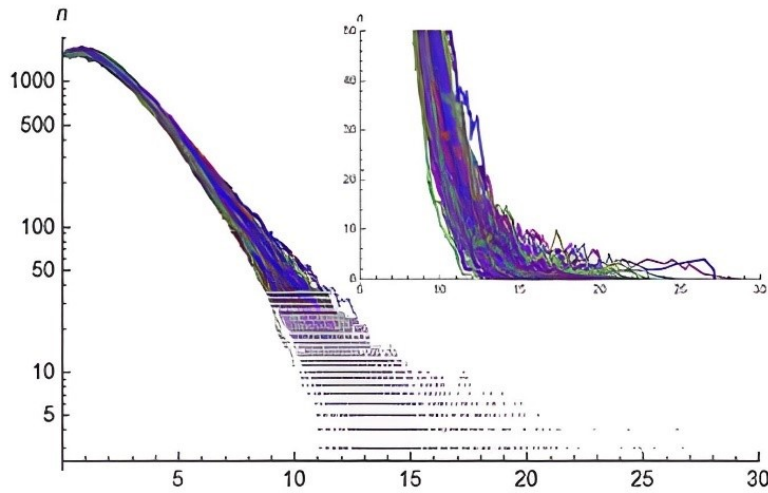


Figure 10. Illustration of the two regimes of evolution n of in terms of t . A set of trajectories (200 in total) were generated from the same initial condition. In the first regime, the set of trajectories is grouped around the deterministic trajectory. In the second, the trajectories disperse to give random extinction times

5 Conclusions

We first talked about the history, general ideas, and real-world uses of compartmental models in epidemiology. Then we suggested a study of the SIR model that looked at both deterministic and stochastic factors. The analysis of the deterministic dynamics was summarized in Propositions 1 and 2. The deterministic dynamics of the SIR model with stochastic introductions have been summarized in Propositions 1 and 2. Proposition 1 establishes that, asymptotically, the number of infected individuals converges to a stable state, while Proposition 2 details the growth or decline behavior of the epidemic based on the reproduction parameter \mathcal{R}_0 . These findings underscore the critical threshold role of \mathcal{R}_0 in the emergence and control of epidemics, enhancing the understanding of the model dynamics.

After motivating the stochastic analysis, we considered the SIR model as a continuous-time Markov jump process, and we determined its master equation. Two methods were proposed to simulate stochastic trajectories. First, using the Gillespie algorithm, which we demonstrated, and second, using stochastic differential equations in the sense of Itô obtained from the bivariate Fokker-Planck equation. The latter was obtained through a continuous limit of the master equation. Finally, we investigated the distribution of extinction times, motivated by deterministic analysis. Firstly, a theorem obtained by A. D. Barbour was introduced. Secondly, a development explaining the emergence of a Gumbel distribution for the extinction time was proposed.

The study of the stochastic SIR model, the Gillespie algorithm, and the Fokker-Planck equation in the context of epidemics provides a powerful approach to understanding the dynamics of infectious diseases. These models accurately capture the uncertainty and fluctuations inherent in individual interactions and epidemiological events, offering a realistic perspective on the spread of diseases.

The Gillespie algorithm is used to implement the stochastic SIR model. This model simulates stochastic jump processes and gives a realistic picture of discrete epidemiological events. The analysis of temporal distributions, especially the extinction time of the infection, yields valuable insights into the temporal variability of epidemics.

Transitioning to the Fokker-Planck equation in the continuous limit provides an analytical description of the system's behavior. When you add diffusion terms to the Fokker-Planck equation, deterministic paths take on a random component. This shows how important randomness is, especially in small populations.

Analyzing the distribution of extinction times, whether through stochastic simulations or the Fokker-Planck equation, allows for quantifying the uncertainty surrounding the moment when an infectious disease disappears from a population. This deeper understanding of epidemiological dynamics contributes to more precise predictions and more effective public health strategies.

By integrating stochastic elements, this approach expands our understanding of epidemics and provides advanced mathematical tools for interpreting real epidemiological data. Thus, this study makes a significant contribution to epidemiological modeling, paving the way for better-informed public health decisions and more targeted preventive interventions.

In conclusion, this paper has provided valuable insights into the dynamics of infectious diseases by employing both deterministic and stochastic analyses. The deterministic dynamics were effectively summarized in Propositions 1 and 2, which highlighted key features of the system's behavior under deterministic conditions. Proposition 1 likely encapsulates the stability conditions or equilibrium states, while Proposition 2 may shed light on the critical points

associated with disease spread or control.

The pivotal transition from deterministic to stochastic analysis marked a significant shift in the study's focus. The stochastic approach, particularly through the Gillespie algorithm and the Fokker-Planck equation, allowed for a more realistic modeling of infectious disease spread. By considering random fluctuations and uncertainties inherent in individual interactions, the stochastic analysis captured the intricacies of epidemics that deterministic models might overlook.

While the deterministic analysis provides a foundational understanding, the stochastic perspective offers a more comprehensive and applicable framework, considering the inherent variability in real-world epidemiological processes. The introduction of noise, as represented by the stochastic components, enriches the model and better aligns with the unpredictability of infectious disease dynamics.

However, it is crucial to acknowledge the assumptions made in the model, such as the assumption of well-mixed populations and homogeneity in disease transmission. These assumptions may limit the applicability of the model to specific scenarios. Future research could focus on refining the model to incorporate spatial heterogeneity, non-homogeneous mixing patterns, or other realistic complexities.

In terms of future research, an exploration of how environmental factors influence disease spread within the stochastic framework could be insightful. Additionally, investigating the impact of interventions, such as vaccination or quarantine measures, within the stochastic SIR model would contribute to practical applications in public health. Addressing these aspects could enhance the model's predictive ability and provide more nuanced guidance for disease control strategies.

In summary, this study bridges the gap between deterministic and stochastic approaches in epidemiological modeling, offering a nuanced understanding of infectious disease dynamics. Acknowledging the model's strengths and limitations, future research avenues should focus on refining the model's realism and exploring its applications to diverse epidemiological scenarios.

Author Contributions

Conceptualization, R.E.C.; methodology, R.E.C and S.B.; software, R.E.C.; validation, R.E.C and S.B.; formal analysis, R.E.C and S.B; investigation, R.E.C and S.B; resources, R.E.C and S.B; data curation.; writing—review and editing, R.E.C and S.B.

Data Availability

Not applicable.

Conflicts of Interest

The authors declare no conflict of interest.

References

- [1] F. D. Juengling, A. Maldonado, F. Wuest, and T. H. Schindler, "The role of nuclear medicine for COVID-19: Time to act now," *J. Nucl. Med.*, vol. 61, no. 6, pp. 781–782, 2020. <https://doi.org/10.2967/jnumed.120.246611>
- [2] R. Ikram, A. Khan, M. Zahri, A. Saeed, M. Yavuz, and P. Kumam, "Extinction and stationary distribution of a stochastic COVID-19 epidemic model with time-delay," *Comput. Biol. Med.*, vol. 141, p. 105115, 2022. <https://doi.org/10.1016/j.compbimed.2021.105115>
- [3] "How epidemiology has shaped the COVID pandemic," *Nature*, vol. 589, no. 7843, pp. 491–492, 2021. <https://doi.org/10.1038/d41586-021-00183-z>
- [4] H. C. Tuckwell and R. J. Williams, "Some properties of a simple stochastic epidemic model of SIR type," *Math. Biosci.*, vol. 208, no. 1, pp. 76–97, 2007. <https://doi.org/10.1016/j.mbs.2006.09.018>
- [5] M. Gayer, D. Legros, P. Formenty, and M. A. Connolly, "Conflict and emerging infectious diseases," *Emerg. Infect. Dis.*, vol. 13, no. 11, pp. 1625–1631, 2007. <https://doi.org/10.3201/eid1311.061093>
- [6] K. Mamis and M. Farazmand, "Stochastic compartmental models of the COVID-19 pandemic must have temporally correlated uncertainties," *Proc. R. Soc. A.*, vol. 479, no. 2269, 2023. <https://doi.org/10.1098/rspa.2022.0568>
- [7] W. O. Kermack and A. G. McKendrick, "Contributions to the mathematical theory of epidemics—II. The problem of endemicity," *Bull. Math. Biol.*, vol. 53, no. 1-2, pp. 57–87, 1991. [https://doi.org/10.1016/s0092-8240\(05\)80041-2](https://doi.org/10.1016/s0092-8240(05)80041-2)
- [8] M. Martcheva, "Introduction to epidemic modeling," in *Texts in Applied Mathematics*. Springer, Boston, MA, USA, 2015, pp. 9–31. https://doi.org/10.1007/978-1-4899-7612-3_2

- [9] P. E. Greenwood and L. F. Gordillo, "Stochastic epidemic modeling," in *Mathematical and Statistical Estimation Approaches in Epidemiology*. Springer, Dordrecht, Netherlands, 2009, pp. 31–52. https://doi.org/10.1007/978-90-481-2313-1_2
- [10] M. S. Bartlett and N. T. J. Bailey, "The mathematical theory of epidemics," *J. R. Stat. Soc. Ser. A*, vol. 121, no. 2, p. 236, 1958. <https://doi.org/10.2307/2343364>
- [11] M. Martcheva, "Introduction to epidemic modeling," in *Texts in Applied Mathematics*. Springer, Boston, MA, USA, 2015, pp. 9–31. https://doi.org/10.1007/978-1-4899-7612-3_2
- [12] J. Gillis, "The mathematical theory of epidemics," *Interdiscip. Sci. Rev.*, vol. 4, no. 4, pp. 306–314, 1979. <https://doi.org/10.1179/isr.1979.4.4.306>
- [13] V. Isham, "Stochastic models for epidemics," in *Celebrating Statistics*. Oxford University Press, Oxford, UK, 2005, pp. 27–54. <https://doi.org/10.1093/acprof:oso/9780198566540.003.0002>
- [14] L. J. S. Allen, "An introduction to stochastic epidemic models," in *Mathematical Epidemiology*. Springer, Berlin, Heidelberg, 2008, pp. 81–130. https://doi.org/10.1007/978-3-540-78911-6_3
- [15] H. W. Hethcote, "The mathematics of infectious diseases," *SIAM Rev.*, vol. 42, no. 4, pp. 599–653, 2000. <https://doi.org/10.1137/s0036144500371907>
- [16] J. Reddingius, "Notes on the mathematical theory of epidemics," *Acta Biotheor.*, vol. 20, no. 3-4, pp. 125–157, 1971. <https://doi.org/10.1007/bf01556687>
- [17] P. J. White and M. C. Enright, "Mathematical models in infectious disease epidemiology," in *Infectious Diseases*, vol. 1. Elsevier, 2010, pp. 70–75. <https://doi.org/10.1016/b978-0-323-04579-7.00005-8>
- [18] L. J. S. Allen, "A primer on stochastic epidemic models: Formulation, numerical simulation, and analysis," *Infect. Dis. Model.*, vol. 2, no. 2, pp. 128–142, 2017. <https://doi.org/10.1016/j.idm.2017.03.001>
- [19] H. Andersson and T. Britton, *Stochastic Epidemic Models and Their Statistical Analysis*. Springer, New York, NY, USA, 2000.
- [20] M. Bladt and M. Sørensen, "Statistical inference for discretely observed Markov jump processes," *J. R. Stat. Soc. Ser. B Stat. Methodol.*, vol. 67, no. 3, pp. 395–410, 2005. <https://doi.org/10.1111/j.1467-9868.2005.00508.x>
- [21] G. L. Werneck, "Epidemiology and the Covid-19 pandemic: Opportunities to review trajectories and plan for the future," *Interface Comun. Saúde, Educ.*, vol. 27, pp. 1–16, 2023. <https://doi.org/10.1590/interface.220486>
- [22] D. T. Gillespie, "Exact stochastic simulation of coupled chemical reactions," *J. Phys. Chem.*, vol. 81, no. 25, pp. 2340–2361, 1977. <https://doi.org/10.1021/j100540a008>
- [23] N. Masuda and L. E. C. Rocha, "A Gillespie algorithm for non-Markovian stochastic processes," *SIAM Rev.*, vol. 60, no. 1, pp. 95–115, 2018. <https://doi.org/10.1137/16m1055876>
- [24] M. Schienbein and H. Gruler, "Langevin equation, Fokker-Planck equation and cell migration," *Bull. Math. Biol.*, vol. 55, no. 3, pp. 585–608, 1993. [https://doi.org/10.1016/s0092-8240\(05\)80241-1](https://doi.org/10.1016/s0092-8240(05)80241-1)
- [25] G. I. Zmievskaya, T. A. Averina, and A. L. Bondareva, "Numerical solution of stochastic differential equations in the sense of Stratonovich in an amorphization crystal lattice model," *Appl. Numer. Math.*, vol. 93, pp. 15–29, 2015. <https://doi.org/10.1016/j.apnum.2014.05.006>
- [26] C. Gardiner, *Stochastic Methods: A Handbook for the Natural and Social Sciences*. Springer, Berlin, Heidelberg, 2009.
- [27] L. J. S. Allen, *An Introduction to Stochastic Processes with Applications to Biology, Second Edition*. CRC Press, USA, 2010.
- [28] D. J. Daley and J. Gani, *Epidemic Modelling*. Cambridge University Press, USA, 1984.
- [29] C. Lefèvre, "Stochastic epidemic models for S-I-R infectious diseases: A brief survey of the recent general theory," in *Stochastic Processes in Epidemic Theory*. Springer, Berlin, Heidelberg, 1990, pp. 1–12. https://doi.org/10.1007/978-3-662-10067-7_1
- [30] A. D. Barbour, "The duration of the closed stochastic epidemic," *Biometrika*, vol. 62, no. 2, pp. 477–482, 1975. <https://doi.org/10.1093/biomet/62.2.477>
- [31] B. S. Khatri, "A simple stochastic theory of extinction shows rapid elimination of a Sars-like pandemic," *medRxiv*, 2020. <https://doi.org/10.1101/2020.08.10.20171454>