Chapter 13

Discrete-Time Stochastic Epidemic Models and Their Statistical Inference

Raúl Fierro

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/39271

1. Introduction

There is a wide range of models, both stochastic and deterministic, for the spread of an epidemic. Usually, when the population is constituted of a large number of individuals, a deterministic model is useful as a first approximation, and random variations can be neglected. As an alternative, a stochastic model could be more appropriate for describing the epidemic, but it is less tractable and its mathematical analysis is usually possible only when the population size is very small. However, most populations are not large enough to neglect the effect of statistical fluctuations, nor are they small enough to avoid cumbersome mathematical calculations in the stochastic model. In these cases, it uses to be convenient to take into account both types of models and their relationship. The interplay between ordinary differential equations and Markovian counting processes has been widely investigated in the literature. Major references on this subject can be found in [11, 20–22]. Concerning the deterministic epidemic models, those using ordinary differential equations in their formulation have received special attention and a great number of epidemics is modeled by means of Markovian counting processes. For example, some epidemic models known as SIR, SI, SIS, and others derived from these ones, use differential equations and Markovian counting processes in their formulations. Furthermore, stochastic models based on Markovian counting processes and differential equations are mainly used to carry out the statistical analysis of the model parameters. The Mathematical Theory of Infectious Diseases by Bailey [2] represents a classical reference containing a presentation and analysis of these models. However a more recent book by Andersson and Britton [1] entitled Stochastic Epidemic Models and their Statistical Analysis is a more appropriate reference according to the point of view of this chapter. The spread of these epidemics is developed in a closed population, which is divided into three individual compartments, i.e. susceptible, infective and removed cases; different types of transitions can occur among these three groups of individuals. These models include the stochastic and deterministic versions of the Kermack and McKendrick model [19] and the SIS epidemic model, among others. Moreover, a number of variations of these models has been widely studied. Modeling of epidemics by continuous-time Markov chains has a long history; thus, it seems pertinent to cite the works by [4–6, 18, 24, 28].

Also, an epidemic can be modeled by means of discrete-time. This is the case of the classical Reed-Frost model, which is a Markovian discrete-time SIR epidemic model. However, this modeling has two differences with the corresponding one based on counting processes. First, its latent period is assumed constant and equal to the time unit. Secondly, there are no deterministic counterpart based on differential equations as it is the case of an epidemic modeled by means of a Markovian counting process. Another type of population modeling, which is applied to metapopulations, has been introduced by some authors such [9, 10, 27] and other references therein. These researchers derive an approximation that preserves the discrete time structure and reduces the complexity of the models. Probably, these results could be applied to epidemic models and asymptotic inference on the parameters of these models, could be carried out.

This chapter is a compendium of two works by the author, whose references are [12, 13]. A wide class of discrete-time stochastic epidemic models is introduced and analyzed from a statistical point of view. Just as some models based on ordinary differential equations involve a natural alternative through Markovian counting processes, this class includes a counterpart based on differential equations. Unlike those epidemic models where transitions occur at random times, our proposal involves the advantage of being suitable for epidemics that cannot be observed for a long period of time, as in some epidemics where observations are done at previously determined times. This is the main reason for preferentially considering these kind of stochastic models instead of those based on continuous time. It is expected the smaller the periods of time between transitions and the bigger the population, the more similar the stochastic and deterministic models would become. Indeed, one of the main aims of this paper is to prove such a similarity. As a second aim, we are highly interested in carry out statistical analysis on the parameters of the modeling. For this purpose, martingale estimators for the parameters involve in the modeling are derived and their asymptotic normality is proved.

Since the results stated here do not assume a distribution for the process modeling the epidemic, it is not possible to derive a likelihood ratio and hence maximum likelihood estimators cannot be obtained. Even, in many cases when the process representing the model is Markovian, the maximum likelihood estimators cannot be obtained in a closed form, which makes difficult to carry out statistical inference on the parameter of the model. As pointed out in [7], likelihood functions corresponding to epidemic data are often very complicated. In these cases, parameter estimation based on martingale estimators use to be an appropriate alternative to work out this difficulty. This method arises as a natural way of estimation when no distribution in the model is assumed or, when the maximum likelihood estimators cannot be obtained in a closed form.

This chapter is organized as follows. The general form of the model and two preliminary lemmas are introduced in Section 2. Section 3 contains brief definitions of some typical models included in the biomathematical literature. The deterministic counterpart of the general model along with its relationships with it is presented in Section 4. Indeed, the convergence of the stochastic model to the deterministic one and the asymptotic behavior of the corresponding fluctuations are proved. Moreover, in Section 4 a version of the SIS epidemic model is presented and numerical simulations are carried out. The parameter estimators are defined in Section 5, and their asymptotic normality is proved. The General Epidemic Model along with the statistical analysis on the parameters is stated in Section 6.

Results of Section 5 are applied here to some hypothesis tests. Section 7 is devoted to some numerical simulations. Finally, Section 8 contains some conclusions.

2. Modeling and preliminaries

2.1. Discrete-time modeling

The model which we introduce here is defined as follows: a community of individuals divided into three different compartments is considered, namely, susceptible, infective and removed individuals.

Suppose the size population is n, and for each $t \geq 0$, $S^n(t)$, $I^n(t)$ and $R^n(t)$ represent, respectively, the number of susceptible, infective and removed individuals at time t. Since it is assumed the population size is constant, then for each $t \geq 0$, should be $S^n(t) + I^n(t) + R^n(t) = n$. These processes are observed at discrete-time instants which are defined by the sequence $\{t_k^n\}_{k\in\mathbb{N}}$, where for each $k \in \mathbb{N}$, $t_k^n = k\Delta/n$, $(\Delta > 0)$, i.e. each time subinterval has length Δ/n. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. In the sequel, all stochastic processes and random variables are defined on this probability space and for a stochastic process Z, we denote $\Delta Z(t_k^n) = Z(t_k^n) - Z(t_{k-1}^n)$.

Let $M^\top$ denote the transpose of a matrix M and $X^n = (S^n, I^n, R^n)^\top$. Transitions of individuals among the three compartments are determined by m increasing stochastic processes $Z_1^n, \ldots, Z_m^n$, and the number of individuals in each compartment is obtained, for each $t \geq 0$, by means of

$$X^n(t) = X^n(0) + AZ^n(t), \tag{1}$$

where $Z^n(t) = (Z_1^n(t), \ldots, Z_m^n(t))^\top$ and A is a $3 \times m$-incidence matrix.

It is assumed $Z_1^n, \ldots, Z_m^n$ take values in the set of non-negative integer numbers, have right-continuous trajectories and start at zero, i.e. $Z_1^n(0) = \cdots = Z_m^n(0) = 0$. Let $\mathcal{F}_k^n$ be the σ-field $\sigma(Z^n(t_1^n), \ldots, Z^n(t_k^n))$ generated by $Z^n(t_1^n), \ldots, Z^n(t_k^n)$. The stochastic processes $Z_1^n, \ldots, Z_m^n$ increase according to m density dependent transition rates, which are defined by means of m non-negative functions $a_1, \ldots, a_m$, respectively. The domain of these functions is and open set of $\mathbb{R}^3$ containing the 3-simplex $E = \{(u, v, w)^\top \in [0, 1]^3 : u + v + w = 1\}$ and it is assumed the following condition holds:

(C) For each $k \in \mathbb{N}$, $\Delta Z_1^n(t_k^n), \ldots, \Delta Z_m^n(t_k^n)$ are $\mathcal{F}_{k-1}^n-$conditionally independent and satisfy

$$\mathbb{E}(\Delta Z_i^n(t_k^n)|\mathcal{F}_{k-1}^n) = a_i(\chi^n(t_{k-1}^n)), \quad (i \in \{1, \ldots, m\}),$$

where $\chi^n(t) = (\sigma^n(t), \iota^n(t), \rho^n(t))^\top$, $\sigma^n(t) = S^n(t)/n$, $\iota^n(t) = I^n(t)/n$ and $\rho^n(t) = R^n(t)/n$.

A wide variety of stochastic models for epidemics satisfy condition (C). It is important to point out this condition does not determine the law or distribution of χ^n, i.e. there could be two or more processes satisfying this condition, though they have different transition probabilities. This fact enables this condition to be applied to a wide class of models, since in order to verify condition (C), the distribution of the process need not be known. Actually, a stochastic process satisfying condition (C) need not be Markovian. Nevertheless, some Markov chains, having density dependent transition rates, satisfy condition (C) and hence they may be included in our setting.

2.2. Two preliminary lemmas

In what follows, $[a]$ stands for the integer part of a real number a and for each $i = 1, \ldots, m$, we denote $L_i^n(t) = \sum_{k=1}^{[nt]} \xi_k^n(i)\Delta t^n$ and $L^n(t) = (L_1^n(t), \ldots, L_m^n(t))^\top$, $(t \geq 0)$, where $\xi_k^n(i) = \Delta Z_i^n(t_k^n) - a_i(\chi^n(t_{k-1}^n))$ and $\Delta t^n = 1/n$. From Condition (C), it is obtained that, by defining $\mathcal{G}_t^n = \mathcal{F}_{[nt]}^n$, $L^n = \{L^n(t); t \geq 0\}$ is an m-dimensional martingale with respect to $\{\mathcal{G}_t^n; t \geq 0\}$.

Through this chapter, for each $i = 1, \ldots, m$, $v_i^n(t)$ and $\langle L_i^n \rangle$ stand for the random variable

$$v_i^n(t) = \frac{1}{n} \sum_{k=1}^{[nt]} \mathbb{E}(\xi_k^n(i)^2 | \mathcal{F}_{k-1}^n)$$

and the predictable quadratic variation of L_i^n, respectively.

Lemma 2.1. *For each $t \geq 0$, $L_1^n(t), \ldots, L_m^n(t)$ are $\mathcal{G}_{t-}^n$-conditionally independent random variables and, the predictable quadratic variation matrix of L^n is given by*

$$\langle L^n \rangle(t) = \frac{1}{n} \begin{pmatrix} v_1^n(t) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & v_m^n(t) \end{pmatrix}, \quad (t \geq 0).$$

Proof. For $t \geq 0$, the $\mathcal{G}_{t-}^n$-conditional independence of $L_1^n(t), \ldots, L_m^n(t)$ follows from Assumption (C) and it is clear that for each $i = 1, \ldots, m$, the predictable quadratic variation of L_i^n is given by $\langle L_i^n \rangle(t) = \frac{1}{n^2} \sum_{k=1}^{[nt]} \mathbb{E}(\xi_k^n(i)^2 | \mathcal{F}_{k-1}^n)$. Hence $\langle L_i^n \rangle(t) = v_i^n(t)/n$, $(t \geq 0)$, which concludes the proof. □

In the sequel, for each $d \in \mathbb{N}$, $\|\cdot\|$ stands for the Euclidean vector norm in $\mathbb{R}^d$.

Lemma 2.2. *Let $T > 0$ and suppose for each $i = 1, \ldots, m$, $\{\frac{1}{n} v_i^n(T)\}_{n \in \mathbb{N}}$ converges in probability to zero, as n goes to ∞. Then, $\{\sup_{0 \leq t \leq T} \|L^n(t)\|\}_{n \in \mathbb{N}}$ converges in probability to zero.*

For each $T > 0$ and each $i = 1, \ldots, m$, $\{\frac{1}{n} v_i^n(T)\}_{n \in \mathbb{N}}$ converges in probability to zero, as n goes to ∞.

Proof. From Theorem 1 in [26], for any $\epsilon, \eta > 0$ we have

$$\mathbb{P}\Big(\sup_{0 \leq t \leq T} \|L^n(t)\|^2 > \epsilon \Big) \leq \frac{1}{\epsilon} \sum_{i=1}^m \mathbb{E}(\langle L_i^n \rangle(T) \wedge \eta) + \mathbb{P}\Big(\sum_{i=1}^m \langle L_i^n \rangle(T) > \eta \Big)$$

and hence, Lemma 2.1 implies

$$\mathbb{P}\Big(\sup_{0 \leq t \leq T} \|L^n(t)\|^2 > \epsilon \Big) \leq \frac{m\eta}{\epsilon} + \mathbb{P}\Big(\sum_{i=1}^m \frac{1}{n} v_i^n(T) > \eta \Big).$$

By assumption this lemma follows. □

3. Some epidemic models

Most of the typical models involved in the biomathematical literature have a version which belongs to the model class defined in this approach. Some of them are included below.

3.1. The general epidemic model

A deterministic version of this model based on differential equations, was introduced by [19], while the article by [18] was a pioneer in the stochastic version based on counting processes. This model has received the most attention in the literature, and its analysis can be found in [2], [1] and [23]. Two type of transitions are possible for any individual, from susceptible to infected and from infected to removed individuals. Thus under the perspective of this work, the modeling of this epidemic must satisfy (1) with

$$A = \begin{pmatrix} -1 & 0 \\ 1 & -1 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad Z^n = \begin{pmatrix} Z_1^n \\ Z_2^n \end{pmatrix}.$$

Transitions from susceptible to infected and from infected to removed cases are described by Z_1^n and Z_2^n, respectively, and the functions defining their transition rates are given by $a_1(u,v,w) = \beta uv$ and $a_2(u,v,w) = \gamma v$, where β and γ are two parameters denoting the infection and removal rates. This model is also known as SIR model.

3.2. The SIRS Model

This is a slight modification of the preceding model. Besides the transitions determined by Z_1^n and Z_2^n in the SIR model, a transition from removed to susceptible case is allowed and determined by an increasing stochastic process Z_3^n, where its transition rate is defined by $a_3(u,v,w) = \delta w$. In this case, some of the removed cases may become susceptible to be infected again. The incidence matrix defining this model is given by

$$A = \begin{pmatrix} -1 & 0 & 1 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \end{pmatrix}.$$

3.3. The SIS Model

One of the simplest epidemic models is the SIS model, which uses to be suitable for infections resulting from bacteria such as gonorrhea, malaria, etc. In this case, only transitions from susceptible to infected individuals are allowed, as well as transitions from infected to susceptible individuals. According to the approach of this study, two transitions are allowed, and determined by Z_1^n and Z_2^n with transition rates defined by $a_1(u,v,w) = \beta uv$ and $a_2(u,v,w) = \gamma v$, respectively. In this case, the incidence matrix is

$$A = \begin{pmatrix} -1 & 1 \\ 1 & -1 \\ 0 & 0 \end{pmatrix}.$$

Notice the compartment corresponding to removed cases is considered having no individuals.

3.4. The modified SIR model

This is a modification of the general epidemic model and aims to AIDS modeling. As in the general epidemic model, two transitions Z_1^n and Z_2^n define the model with transition rates given by $a_1(u,v,w) = \beta uv/(u+v)$ and $a_2(u,v,w) = \gamma v$, respectively. As before, β and γ correspond to the model parameters. Some references concerning the deterministic version of this model are, for instance, [16, 17], while the stochastic version based on Markovian counting processes was introduced by [4]. The incidence matrix is defined as in the SIR model, i.e.

$$A = \begin{pmatrix} -1 & 0 \\ 1 & -1 \\ 0 & 1 \end{pmatrix}.$$

In general, as much in the SIR as the modified model, the transition rate due to infection is proportional to the susceptible density and the fraction of infected individuals with respect to individuals in circulation. Consequently, in the SIR model this fraction is $v/(u+v+w) = v$, while in an epidemic where to be removed is equivalent to be dead or out of circulation, this fraction is $v/(u+v)$. In modeling AIDS, removed cases are presumed to be so ill with AIDS that they no longer take part in transmission.

4. The deterministic counterpart

In this section, we examine the relationship between the model we are introducing here and an associated ordinary differential equation, which we call its deterministic counterpart.

Let $F(x) = Aa(x)$, where $a(x) = (a_1(x), \ldots, a_m(x))^\top$, $(x \in E)$, and consider the following ordinary differential equation:

$$\frac{\mathrm{d}\chi}{\mathrm{d}t}(t) = F(\chi(t)), \quad \chi(0) = \chi_0, \tag{2}$$

where $\chi_0 = (\sigma_0, \iota_0, \rho_0)^\top \in E$ is the initial condition.

In order to obtain existence and uniqueness of the solution to (2), it will be assumed the following usual Lipchitz condition holds:

(L) $\quad \|F(x) - F(y)\| \leq K\|x - y\|,$

for all $x, y \in E$, where K is a positive constant.

4.1. Comparison between the stochastic and deterministic models

The theorem below stated the consistency of the stochastic model with respect to the deterministic one and we will use it to study the asymptotic behavior of the estimators for the parameters of the model.

Theorem 4.1. *Let χ be the unique solution to (3) and assume conditions (C) and (L) are satisfied. Moreover, suppose the following two conditions hold:*

(4.1.1) *The sequence of initial conditions $\{\chi^n(0)\}_{n\in\mathbb{N}}$ converges in probability to $\chi_0 = (\sigma_0, \iota_0, \rho_0)^\top$, as n goes to ∞.*

(4.1.2) *For each $T > 0$ and each $i = 1, \dots, m$, $\{\frac{1}{n} v_i^n(T)\}_{n\in\mathbb{N}}$ converges in probability to zero, as n goes to ∞.*

Then, $\{\chi^n\}_{n\in\mathbb{N}}$ converges in probability uniformly over compact subsets of $\mathbb{R}_+$ to χ, i.e. for each $T > 0$, $\{\sup_{0\le t\le T} \|\chi^n(t) - \chi(t)\|\}_{n\in\mathbb{N}}$ converges in probability to zero, as n goes to ∞.

Proof. Since, for each $i = 1, \dots, m$, $\triangle Z_i^n(t_k^n) = a_i(\chi^n(t_{k-1}^n)) + \xi_k^n(i)$, we have

$$\begin{aligned} Z_i^n(t) &= n \sum_{k=1}^{[nt]} a_i(\chi^n(t_{k-1}^n))\triangle t^n + n \sum_{k=1}^{[nt]} \xi_k^n(i) \triangle t^n \\ &= n \int_0^t a_i(\chi^n(u))\, \mathrm{d}u - (nt - [nt]) a_i(\chi^n(t)) + n L_i^n(t), \end{aligned}$$

and

$$\chi^n(t) = \chi^n(0) + \int_0^t F(\chi^n(u))\, \mathrm{d}u + A L^n(t) + \frac{(nt - [nt])}{n} F(\chi^n(t)).$$

Hence

$$\chi^n(t) - \chi(t) = \chi^n(0) - \chi(0) + \int_0^t [F(\chi^n(u)) - F(\chi(u))]\, \mathrm{d}u + A L^n(t) - \epsilon^n(t), \tag{3}$$

where $\epsilon^n(t) = \frac{(nt-[nt])}{n} F(\chi^n(t))$.

For any matrix B, let us denote $|||B||| = \sup_{\|x\|=1} \|Bx\|$. Fix $T > 0$ and let

$$g^n(t) = \sup_{0\le s\le t} \|\chi^n(s) - \chi(s)\|, \quad (t \in [0, T]).$$

From (3) and (L), for each $t \ge 0$, we have

$$g^n(t) \le \alpha_n + K \int_0^t g^n(u)\, \mathrm{d}u,$$

where $\alpha_n = g^n(0) + \sup_{0\le t\le T} \|A L^n(t)\| + \mathrm{O}(1/n)$. Since $\{g^n(0)\}_{n\in\mathbb{N}}$ converges in probability to zero and $\sup_{0\le t\le T} \|A L^n(t)\| \le |||A||| \sup_{0\le t\le T} \|L^n(t)\|$, Lemma 2.1 and Gronwall's inequality imply $\{g^n(T)\}_{n\in\mathbb{N}}$ converges in probability to zero. This completes the proof. □

Remark 4.1. *By Chebishev's inequality, Condition (4.1.2) is satisfied whenever the following stronger condition holds: For each $T > 0$ and each $i = 1, \dots, m$,*

$$\lim_{n\to\infty} \mathbb{E}(\langle L^n(i)\rangle(T)) = 0. \tag{4}$$

The following result aims to the problem of finding approximate confident bands for the solution to (2). Before stating it, for each $x \in E$, let $\mathrm{D}(F)(x)$ denote the Jacobian matrix of F at $x = (x_1, x_2, x_3)^\top$, that is,

$$\mathrm{D}(F)(x) = \begin{pmatrix} \frac{\partial F_1}{\partial x_1}(x) & \frac{\partial F_1}{\partial x_2}(x) & \frac{\partial F_1}{\partial x_3}(x) \\ \frac{\partial F_2}{\partial x_1}(x) & \frac{\partial F_2}{\partial x_2}(x) & \frac{\partial F_2}{\partial x_3}(x) \\ \frac{\partial F_3}{\partial x_1}(x) & \frac{\partial F_3}{\partial x_2}(x) & \frac{\partial F_3}{\partial x_3}(x) \end{pmatrix}.$$

In the sequel, $\{Y^n\}_{n\in\mathbb{N}}$ is the sequence defined as $Y^n = \sqrt{n}(\chi^n - \chi)$.

Theorem 4.2. *Let $\{Y^n\}_{n\in\mathbb{N}}$ be the sequence defined as $Y^n = \sqrt{n}(\chi^n - \chi)$ and suppose the following three conditions hold:*

(4.2.1) *For each $i, j = 1, 2, 3$, the partial derivative $\dfrac{\partial F_i}{\partial x_j}(x)$ exists and it is continuous at x in an open set containing E.*

(4.2.2) *For each $\epsilon > 0$ and each $i = 1, \ldots, m$, $\{\dfrac{1}{n}\sum_{k=1}^{n} \mathbb{E}(\xi_k^n(i)^2 \mathrm{I}_{\{|\xi_k^n(i)|>\epsilon\sqrt{n}\}} | \mathcal{F}_{k-1}^n)\}_{n\in\mathbb{N}}$ converges in probability to zero.*

(4.2.3) *$\{Y^n(0)\}_{n\in\mathbb{N}}$ converges in distribution to a three-variate random vector η.*

(4.2.4) *For each $t \geq 0$ and each $i = 1, \ldots, m$, $\{v_i^n(t)\}_{n\in\mathbb{N}}$ converges in probability to $v_i(t)$, where for each $i = 1, \ldots, m$, $v_i : [0, \infty[\to \mathbb{R}$ is an increasing continuous function such that $v_i(0) = 0$.*

Then, $\{Y^n\}_{n\in\mathbb{N}}$ converges in law to the solution Y satisfying the following stochastic differential equation:

$$\mathrm{d}Y(t) = \mathrm{D}(F)(\chi(t))Y(t)\mathrm{d}t + \mathrm{d}M(t), \quad Y(0) = \eta, \tag{5}$$

where M is a continuous martingale with predictable quadratic variation given by the matrix

$$\langle M\rangle(t) = A \cdot \mathrm{Diag}(v_1(t), \ldots, v_m(t)) \cdot A^\top$$

and $\mathrm{Diag}(v_1(t), \ldots, v_m(t))$ stands for the diagonal matrix with entries $v_1(t), \ldots, v_m(t)$ at its diagonal.

Proof. Let $M^n = \sqrt{n}AL^n$. By making use of Corollary 12, Chapter II in [29], (4.2.2) and (4.2.4) imply that $\{\sqrt{n}L^n\}_{n\in\mathbb{N}}$ converges in law to a continuous martingale Q with predictable quadratic variation $\langle Q\rangle$ given by $\langle Q\rangle(t) = \mathrm{Diag}(v_1(t), \ldots, v_m(t))$. Consequently, $\{M^n\}_{n\in\mathbb{N}}$ converges in law to a continuous martingale $M = AQ$ with predictable quadratic variation $\langle M\rangle$ given, for each $t \geq 0$, by $\langle M\rangle(t) = A \cdot \mathrm{Diag}(v_1(t), \ldots, v_m(t)) \cdot A^{\mathrm{T}}$.

From (3), we have

$$Y^n(t) = Y^n(0) + \int_0^t \mathrm{D}(F)(\theta^n(s))Y^n(s)\,\mathrm{d}s + M^n(t) - U^n(t), \tag{6}$$

where $\theta^n(s)$ is between $\chi^n(s)$ and $\chi(s)$, and $U^n(t) = \sqrt{n}\epsilon^n(t) = \frac{(nt-[nt])}{\sqrt{n}}F(\chi^n(t))$.

Put $C_1 = \sup_{x\in E} \|\mathrm{D}(F)(x))\|$ and let $C_2 > 0$ such that $\sup_{t\geq 0} \|U^n(t)\| \leq C_2$. We have

$$\sup_{0\leq u\leq t} \|Y^n(u)\| \leq \|Y^n(0)\| + C_1 \int_0^t \sup_{0\leq u\leq s} \|Y^n(u)\|\,\mathrm{d}s + \sup_{0\leq u\leq t} \|M^n(u)\| + C_2.$$

Consequently, from a standard application of the Gronwall inequality, we obtain

$$\sup_{0\leq u\leq t} \|Y^n(u)\| \leq (\|Y^n(0)\| + \sup_{0\leq u\leq t} \|M^n(u)\| + C_2)\,\mathrm{e}^{C_1 t}. \tag{7}$$

Since $\{Y^n(0)\}_{n\in\mathbb{N}}$ and $\{\sup_{0\leq t\leq 1} \|M^n(t)\|\}_{n\in\mathbb{N}}$ are sequences converging in distribution, Theorem 6.2 in [8] implies

$$\lim_{a\to\infty} \sup_{n\in\mathbb{N}} \mathbb{P}(\|Y^n(0)\| > a) = 0, \tag{8}$$

and, from (18), for each $\epsilon > 0$ and any $t \geq 0$,

$$\lim_{\delta \to 0} \sup_{n \in \mathbb{N}} \mathbb{P}\Big(\sup_{0 \leq u \leq t} \|Y^n(u)\| > \epsilon / \delta \Big) = 0, \tag{9}$$

In order to prove the convergence in law of $\{Y^n\}_{n \in \mathbb{N}}$, fix $T > 0$ and let us define the modulus of continuity $\omega_T^D : \mathrm{D}([0,T], \mathbb{R}^3) \times]0, \infty[\to \mathbb{R}$ as

$$\omega_T^D(x, \delta) = \inf_{\{t_i\}} \max_{0 < i \leq r} \sup_{t_{i-1} \leq s, t < t_i} \|x(s) - x(t)\|,$$

where the infimum extends over the finite sets $\{t_i\}$ of points satisfying

$$\begin{cases} 0 = t_0 < t_1 < \cdots < t_r = T, \\ t_i - t_{i-1} > \delta, \quad i = 1, \ldots, r. \end{cases}$$

Here, $\mathrm{D}([0,T], \mathbb{R}^3)$ stands for the Skorohod space of all functions from $[0,T]$ into $\mathbb{R}^3$, which are right continuous and left-hand limited.

From (6) we have

$$\omega_T^D(Y^n, \delta) \leq \delta C_1 \sup_{0 \leq t \leq T} \|Y^n(t)\| + \omega_T^D(M^n, \delta) + 2C_2 / \sqrt{n}. \tag{10}$$

Since $\{M^n\}_{n \in \mathbb{N}}$ converges in law to M, it follows from Theorem 15.2 by [8] that for each $\epsilon > 0$, $\lim_{\delta \to 0} \sup_{n \in \mathbb{N}} \mathbb{P}(\omega_T^D(M^n, \delta) > \epsilon) = 0$. Hence, from (9) and (10), for each $\epsilon > 0$, we have

$$\lim_{\delta \to 0} \sup_{n \in \mathbb{N}} \mathbb{P}(\omega_T^D(Y^n, \delta) > \epsilon) = 0. \tag{11}$$

Conditions (8) and (11) imply the sequence $\{Y^n\}_{n \in \mathbb{N}}$ satisfies the hypotheses of Theorem 15.2 in [8] and hence, the sequence of probabilities measures $\{\mathbb{P}(Y^n \in \cdot)\}_{n \in \mathbb{N}}$ is tight. This fact, along Theorem 6.1 in [8], imply that for the convergence in law, $\{Y^n\}_{n \in \mathbb{N}}$ is relatively compact. Let Y be a process and $\{Y^{n_k}\}_{k \in \mathbb{N}}$ a subsequence of $\{Y^n\}_{n \in \mathbb{N}}$ such that $\{Y^{n_k}\}_{k \in \mathbb{N}}$ converges in law to Y. Since $\{\sup_{t \geq 0} \|U^n(t)\|\}_{n \in \mathbb{N}}$ converges to zero, it follows from (6), (4.2.1) and Theorem 4.1 that

$$Y(t) = Y(0) + \int_0^t \mathrm{D}(F)(\chi(s)) Y(s)\, \mathrm{d}s + M(t).$$

Moreover, since $\{Y^n(0)\}_{n \in \mathbb{N}}$ converges in distribution to η and $Y(0)$ equals η in distribution, we have Y is a solution to (5). Finally, uniqueness of solutions to (5) implies $\{Y^n\}_{n \in \mathbb{N}}$ converges in distribution to Y, which concludes the proof. □

Remark 4.2. *By Itô's rule, the unique solution to (5) is given by*

$$Y(t) = \Psi(t) \left[\eta + \int_0^t \Psi(s)^{-1}\, \mathrm{d}M(s) \right], \quad 0 \leq t \leq 1,$$

where Ψ *is the unique solution to the matrix differential equation*

$$\Psi'(t) = \mathrm{D}(F)(\chi(t)) \Psi(t), \quad \Psi(0) = \textit{identity matrix}.$$

The stochastic process $\{Y(t); t \geq 0\}$ allows us to give confidence bands for the deterministic model defined by the solution χ to (2). In Section 6 such band is constructed for the SIS epidemic model, where some simulations are carried out.

4.2. Numerical simulations for the SIS epidemic model

In this section, our attention is focused on the SIS epidemic model. As explained in Subsection 3.3, it is assumed that only susceptible and infective individuals are in a closed homogeneously mixing population. Let $\sigma(t)$ and $\iota(t)$ be the densities of susceptible and infective individuals, respectively. The deterministic model is defined by the following system of ordinary differential equations:

$$\begin{aligned}\frac{\mathrm{d}\sigma}{\mathrm{d}t}(t) &= -\beta\sigma(t)\iota(t) + \gamma\iota(t)\\ \frac{\mathrm{d}\iota}{\mathrm{d}t}(t) &= \beta\sigma(t)\iota(t) - \gamma\iota(t).\end{aligned}$$

Since for each $t \geq 0$, $\sigma(t) + \iota(t) = 1$, this model is completely determined by the ordinary differential equation:

$$\frac{\mathrm{d}\iota}{\mathrm{d}t}(t) = \beta(1-\iota(t))\iota(t) - \gamma\iota(t),$$

which, given $\iota(0) = \iota_0 \in]0,1[$, has the unique solution

$$\iota(t) = \begin{cases} \frac{\iota_0}{\iota_0\beta t+1} & \text{if } \beta = \gamma \\ \frac{\iota_0(\beta-\gamma)\,\mathrm{e}^{(\beta-\gamma)t}}{\beta-\gamma+\iota_0\beta(\mathrm{e}^{(\beta-\gamma)t}-1)} & \text{if } \beta \neq \gamma. \end{cases} \tag{12}$$

The relative removal-rate, see for instance [2] and [3], is defined as $\tau = \gamma/\beta$, where γ and β represent the removal and infection rates, respectively. We note that $\iota(t) \to 0$ as $t \to \infty$ if $\tau \geq 1$, while $\iota(t) \to 1-\tau > 0$ as $t \to \infty$ if $\tau < 1$. For this reason, $n(1-\tau)$ is called the endemic level of the process.

Let $\sigma^n(t)$ and $\iota^n(t)$ be the densities of susceptible and infective individuals in the stochastic version of the SIS epidemic model. According to our setting, σ^n and ι^n are defined as

$$\begin{aligned}\sigma^n(t) &= \sigma^n(0) + \tfrac{1}{n}(Z_2^n(t) - Z_1^n(t))\\ \iota^n(t) &= \iota^n(0) + \tfrac{1}{n}(Z_1^n(t) - Z_2^n(t)),\end{aligned}$$

where $\mathbb{E}(\Delta Z_1^n(t_k^n)|\mathcal{F}_{k-1}^n) = \beta\sigma(t_{k-1}^n)\iota(t_{k-1}^n)$ and $\mathbb{E}(\Delta Z_2^n(t_k^n)) = \gamma\iota(t_{k-1}^n)$.

Let us assume $n\iota^n(0) = [n\iota_0]$. From Theorem 4.1, $\{\iota^n\}_{n\in\mathbb{N}}$ converges uniformly in probability to ι, over compact subsets of $\mathbb{R}_+$, whenever $\iota(0) = \iota_0$. It is worth noting $1-\tau$ is an asymptotically stable equilibrium state for the deterministic model and not for the stochastic model. However, it is expected $\iota^n(t)$ is close to this value due to Theorem 4.1, for a large enough n. In Figure 1, the deterministic and stochastic models, for $\beta = 3$, $\gamma = 1$, $\iota_0 = (1-\tau)/2$ and $n = 1{,}000$, are compared. Let $y^n = \sqrt{n}(\iota^n - \iota)$. It follows from Theorem 4.2 that $\{y^n\}_{n\in\mathbb{N}}$ converges in law to y, the solution to the following Langevin equation:

$$\mathrm{d}y(t) = (\beta - \gamma - 2\beta\iota(t))y(t)\,\mathrm{d}t + \sqrt{\beta(1-\iota(t))\iota(t) + \gamma\iota(t)}\,\mathrm{d}W(t), \quad \iota(0) = 0,$$

where W is a one dimensional standard Brownian motion. From Remark 4.2, the solution to this equation is

$$y(t) = \int_0^t b(u)\,\mathrm{e}^{\int_u^t a(s)\,\mathrm{d}s}\,\mathrm{d}W(u), \tag{13}$$

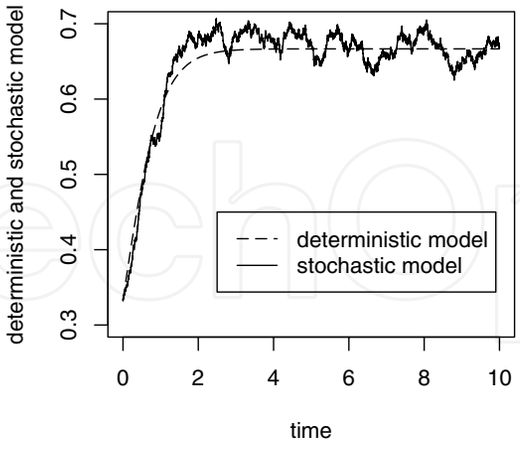

Figure 1. Comparing the deterministic and stochastic models.

where $a(s) = \beta - \gamma - 2\beta y(s)$ and $b(u) = \sqrt{\beta(1 - y(u))y(u) + \gamma y(u)}$. Hence, $y(t)$ has a normal distribution with mean zero and variance

$$\mathrm{Var}(y(t)) = \int_0^t b(u)^2 \, \mathrm{e}^{2\int_u^t a(s)\,\mathrm{d}s} \, \mathrm{d}u.$$

Suppose $\tau < 1$ and the process $n\iota^n$ starts close to the endemic level, i.e. $\iota_0 = 1 - \tau$. Consequently, $\iota(t) = 1 - \tau$ for all $t \geq 0$, and (13) becomes

$$y(t) = \sqrt{2\gamma(1 - \gamma/\beta)} \int_0^t \mathrm{e}^{-(\beta-\gamma)(t-u)} \, \mathrm{d}W(u).$$

In this case, a simple expression for the variance of $y(t)$ can be obtained and a confidence interval for $\iota(t)$ derived. Indeed, for each $t \geq 0$, $y(t)$ has normal distribution with mean zero and variance

$$\mathrm{Var}(y(t)) = \frac{\gamma}{\beta}(1 - \mathrm{e}^{-2(\beta-\gamma)t}).$$

Since for each $t > 0$ $\mathrm{Var}(y(t)) \leq \tau < 1$, $\iota^n(t)$ differs from $\iota(t)$ at most $w_{\alpha/2}\sqrt{\tau/n}$ with an approximate probability equals $1 - \alpha$.

In Figure 2, a simulation of ι^n is carried out for $\beta = 3$, $\gamma = 1$ and $n = 10$, and we note in this case random fluctuations are important to be neglected.

For $\alpha \in]0, 1[$, $[u^-_{\alpha,n}(t), u^+_{\alpha,n}(t)]$ is a confidence interval for $\iota(t)$, with an approximate confidence level $1 - \alpha$, where

$$u^{\pm}_{\alpha,n}(t) = \iota^n(t) \pm w_{\alpha/2}\sqrt{\mathrm{Var}(y(t))/n}, \tag{14}$$

$1 - \Phi(w_{\alpha/2}) = \alpha/2$ and Φ is the cumulative distribution function of a standard normal random variable. The random bounds given by (14) allow to construct (nonuniform) confidence bands for the solution ι given by (12). In turn, by defining $u^{\pm}_{\alpha}(t) = \iota(t) \pm w_{\alpha/2}\sqrt{\mathrm{Var}(y(t))/n}$, we have $\iota^n(t) \in [u^-_{\alpha}(t), u^+_{\alpha}(t)]$ with an approximate probability $1 - \alpha$ for large values of n.

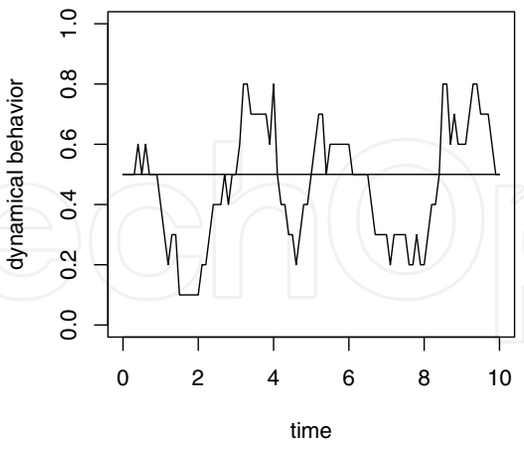

Figure 2. Simulation for $\beta = 3$, $\gamma = 1$ and $n = 10$.

In Figure 3, another simulation of ι^n is carried out for $\beta = 3$ and $\gamma = 1$. However, in order to appreciate the convergence of ι^n to the equilibrium $1 - \tau$ for the deterministic model, a population of size $n = 5,000$ is now considered. The bounds u_α^- and u_α^+ are pictured with dash lines for $\alpha = .05$, and hence, in this case $w_{\alpha/2} = 1.96$.

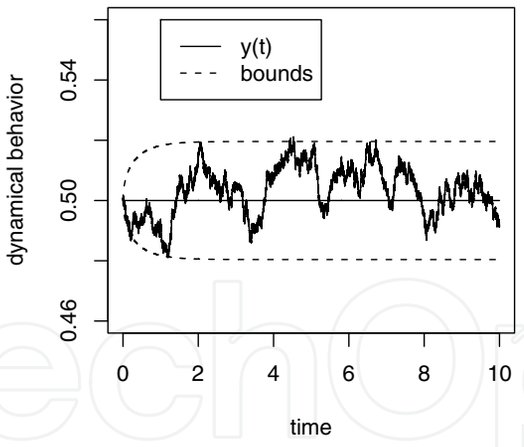

Figure 3. Simulation for $\beta = 3$, $\gamma = 1$ and $n = 5,000$.

Figure 3 confirms $|\iota^n(t) - \iota(t)| \leq C$ with an approximate probability bigger than $1 - \alpha = .95$, where $C = w_{\alpha/2}\sqrt{\tau/n} = .0226$.

5. Estimators and their asymptotic behavior

The main results of this article are stated in this section; however their proofs are deferred to the last section.

5.1. Preliminaries

Let us suppose for each $i = 1, \ldots, m$, a_i splits as $a_i = \beta_i b_i$, where β_i is a parameter of the model and b_i is a non-negative continuous function defined on an open set containing E.

Since for each $i = 1, \ldots, m$, L_i^n is a martingale and

$$Z_i^n(t)/n = \beta_i \sum_{k=1}^{[nt]} b_i(\chi^n(t_{k-1}^n))\Delta t^n + L_i^n(t), \quad (t > 0),$$

by observing the epidemic through the time interval $[0, T]$, $(T > 0)$, we have

$$\widehat{\beta_i^n}(T) = \frac{Z_i^n(T)}{n\sum_{k=1}^{[nT]} b_i(\chi^n(t_{k-1}^n))\Delta t^n} \tag{15}$$

is a martingale estimator of β_i.

For each $T > 0$, let $\widehat{\beta^n}(T) = (\widehat{\beta_1^n}(T), \ldots, \widehat{\beta_m^n}(T))^\top$. Proposition 5.1 below states that $\{\widehat{\beta^n}(T)\}_{n\in\mathbb{N}}$ is a consistent sequence of estimators for $\beta = (\beta_1, \ldots, \beta_m)^\top$.

Proposition 5.1. *Let assume the hypotheses of Theorem 4.1 are satisfied. Then, for each $T > 0$ the sequence $\{\widehat{\beta^n}(T)\}_{n\in\mathbb{N}}$ converges in probability to β.*

Proof. Note that for each $i = 1, \ldots, m$,

$$\widehat{\beta_i^n}(T) = \beta_i + \frac{L_i^n(T)}{\sum_{k=1}^{[nT]} b_i(\chi^n(t_{k-1}^n))\Delta t^n}.$$

From Theorem 4.1, $\{\sum_{k=1}^{[nT]} b_i(\chi^n(t_{k-1}^n))\Delta t^n\}_{n\in\mathbb{N}}$ converges in probability to $\int_0^T b_i(\chi(t))\,\mathrm{d}t$, and from (4.1.1) along with Lemma 2.2 imply $\{L_i^n(T)\}_{n\in\mathbb{N}}$ converges in probability to zero. Therefore, the proof is complete. □

5.2. Asymptotic normality of estimators

In this subsection we state two asymptotic normality results for the martingale estimators for β.

Theorem 5.1. *Let χ be the unique solution to (3) and assume conditions (C) and (L) are satisfied. Moreover, suppose the following three conditions hold:*

(5.1.1) *$\{\chi^n(0)\}_{n\in\mathbb{N}}$ converges in probability to $\chi(0) = (\sigma_0, \iota_0, \rho_0)^\top$, as n goes to ∞.*

(5.1.2) *For each $\epsilon > 0$ and each $i = 1, \ldots, m$, $\{\frac{1}{n}\sum_{k=1}^n \mathbb{E}(\xi_k^n(i)^2 \mathrm{I}_{\{|\xi_k^n(i)|>\epsilon\sqrt{n}\}}|\mathcal{F}_{k-1}^n)\}_{n\in\mathbb{N}}$ converges in probability to zero, as n goes to ∞.*

(5.1.3) *For each $i = 1, \ldots, m$, there exists a continuous increasing function $v_i : [0, \infty[\to \mathbb{R}$ such that $v_i(0) = 0$ and for each $t > 0$, $\{v_i^n(t)\}_{n\in\mathbb{N}}$ converges in probability to $v_i(t)$, as n goes to ∞.*

Then, for each $T > 0$, $\{\sqrt{n}(\widehat{\beta^n}(T) - \beta)\}_{n\in\mathbb{N}}$ converges in distribution, as n goes to ∞, to a normal random vector $N(0, \Sigma)$ having mean zero and variance-covariance matrix $\Sigma = \{\sigma_{ij}(T)\}_{1\leq i,j\leq m}$ satisfying

$$\sigma_{ij}(T) = \begin{cases} v_i(T)/[\int_0^T b_i(\chi(t))\,\mathrm{d}t]^2 & \text{if } i = j \\ 0 & \text{if } i \neq j. \end{cases}$$

Proof. Let $Q^n = (Q_1^n, \ldots, Q_m^n)^\top$, where $Q_i^n = \sqrt{n}L_i^n$, $(i = 1, \ldots, m)$. From Lemma 2.1, for each $t \geq 0$, $Q_1^n(t), \ldots, Q_m^n(t)$ are $\mathcal{G}_{t-}^n$-conditionally independent random variables and for each $i = 1, \ldots, m$, the predictable quadratic variation of the martingale Q_i^n is given by $\langle Q_i^n \rangle = v_i^n$. Condition (5.1.3) indicates for each $i = 1, \ldots, m$, $\{\langle Q_i^n \rangle(t)\}_{n\in\mathbb{N}}$ converges in probability to $v_i(t)$. This fact and Condition (5.1.2) enable to conclude that the hypotheses of Corollary 12 in Chapter II in (Rebolledo, 1979) hold. Consequently, $\{Q^n(T)\}_{n\in\mathbb{N}}$ converges in distribution to a normal random vector $Q(T) = (Q_1(T), \ldots, Q_m(T))^\top$ with mean zero and satisfying

$$\mathbb{E}(Q_i(T)Q_j(T)) = \begin{cases} v_i(T) \text{ if } i = j \\ 0 \quad \text{if } i \neq j. \end{cases}$$

We have $\sqrt{n}(\widehat{\beta^n}(T) - \beta) = D^n Q^n(T)$, where $D^n = Diag(d_1^n, \ldots, d_m^n)$ is the diagonal random matrix with entries $d_1^n, \ldots, d_m^n$ in its diagonal given by

$$d_i^n = \frac{1}{\sum_{k=1}^{[nT]} b_i(\chi^n(t_{k-1}^n))\Delta t^n}.$$

Condition (5.1.3) implies Condition (4.1.2) and consequently, the hypotheses of Theorem 4.1 are satisfied. Thus, for each $i = 1, \ldots, m$, $\{\sum_{k=1}^{[nT]} b_i(\chi^n(t_{k-1}^n))\Delta t^n\}_{n\in\mathbb{N}}$ converges in probability to $\int_0^T b_i(\chi(t))\,\mathrm{d}t$ and Slutzky's theorem (see Theorem 5.1.6 in [25], for instance) enables us to conclude $\{D^n Q^n(T)\}_{n\in\mathbb{N}}$ converges in distribution to a normal random vector with mean zero and variance-covariance matrix $\Sigma = \{\sigma_{ij}\}_{1\leq i,j\leq m}$, which satisfies

$$\sigma_{ij} = \begin{cases} v_i(T)/[\int_0^T b_i(\chi(t))\,\mathrm{d}t]^2 \text{ if } i = j \\ 0 \quad \text{if } i \neq j. \end{cases}$$

This concludes the proof of Theorem 5.1. □

Remark 5.1. *Note that Condition (5.1.2) holds whenever for each $\epsilon > 0$ and each $i = 1, \ldots, m$,*

$$\{\max_{k\leq n} \mathbb{E}(\xi_k^n(i)^2 \mathrm{I}_{\{|\xi_k^n(i)|>\epsilon\sqrt{n}\}})\}_{n\in\mathbb{N}}$$

converges in probability to zero, as n goes to ∞. In particular, this condition is satisfied when the double sequence $\{\xi_k^n(i)^2; 0 \leq k \leq n, n \in \mathbb{N}\}$ is uniformly integrable.

6. The general epidemic model

In order to carry out asymptotical inference for a great number of epidemic models, results stated in Section 5 can be applied. In this subsection, statistical inference for the General Epidemic Model is developed.

6.1. The model

As mentioned in Subsection 3.1, this model contains two parameters β and γ denoting the infection and removal rate, respectively, and it is defined by two increasing integer valued processes, which we denote by A^n and B^n, so that

$$\begin{aligned} S^n(t) &= S^n(0) - A^n(t), \\ I^n(t) &= I^n(0) + A^n(t) - B^n(t), \\ R^n(t) &= R^n(0) + B^n(t). \end{aligned}$$

From (15) and the definition of this model given in Subsection 3.1, the martingale estimators of β and γ in $[0, T]$, are respectively given by

$$\widehat{\beta^n}(T) = \frac{A^n(T)}{\sum_{k=1}^{[nT]} \sigma^n(t_{k-1}^n)\iota^n(t_{k-1}^n)} \quad \text{and} \quad \widehat{\gamma^n}(T) = \frac{B^n(T)}{\sum_{k=1}^{[nT]} \iota^n(t_{k-1}^n)}. \tag{16}$$

In order to verify $\widehat{\beta^n}(T)$ and $\widehat{\gamma^n}(T)$ are martingale estimators, Condition (C) has to be hold. By taking into account some heuristic considerations, which are related to the infection spreading, the distribution of the process can be determined. This fact is sufficient, although not necessary as mentioned previously, to obtain Condition (C). Let β denote the average of effective contacts per time unit between an infected person and any other individual in the population. This constant is known as the contact rate, cf. [15]. Hence, β/n is the average number of effective contacts per time unit per capita of an infected, and it is natural to assume the probability of a susceptible individual to become infective in a time interval $]t_{k-1}^n, t_k^n]$ is $(\beta/n)\iota^n(t_{k-1}^n)$. On the other hand, since the total number of adequate contacts in t_{k-1}^n that may produce an infection in t_k^n equals the susceptible number $S^n(t_{k-1}^n)$, it is assumed that $\Delta A^n(t_k^n)$ and $\Delta B^n(t_k^n)$ conditionally on $\mathcal{F}_{k-1}^n$ have independent Binomial distribution with parameters $(S^n(t_{k-1}^n), (\beta/n)\iota^n(t_{k-1}^n))$ and $(I^n(t_{k-1}^n), \gamma\Delta t_n)$, respectively.

Note that it satisfies

$$\mathbb{E}(\Delta A^n(t_k^n)|\mathcal{F}_{k-1}^n) = a_1(\chi^n(t_{k-1}^n)) \quad \text{and} \quad \mathbb{E}(\Delta B^n(t_k^n)|\mathcal{F}_{k-1}^n) = a_2(\chi^n(t_{k-1}^n)),$$

where $a_1(u, v, w) = \beta uv$ and $a_2(u, v, w) = \gamma v$. Hence, conditions (C) and (L) hold and, $\widehat{\beta^n}(T)$ and $\widehat{\gamma^n}(T)$ are martingales estimators. Also, in the next subsection, we see that this approach satisfies the hypotheses of Theorem 5.1 and hence, asymptotic normality of $\widehat{\beta^n}(T)$ and $\widehat{\gamma^n}(T)$ is obtained.

Regarding the initial state of the epidemic, two natural assumptions can be made and both of them satisfy Condition (5.1.1) in Theorem 5.1. The first consists in assuming $\{\chi^n(0)\}_{n\in\mathbb{N}}$ is a deterministic sequence converging to $\chi(0)$ in E. This assumption is quite reasonable whenever a good knowledge about the initial numbers of susceptible, infected and removed individuals is involved. For instance, this assumption can be done when the population is small and the proportion of individuals belonging to each compartment can be observed and calculated at time zero. The second possible assumption consists in assuming $(S^n(0), I^n(0), R^n(0))^\top$ has multinomial distribution with parameters n, p_1, p_2 and $1 - p_1 - p_2$, $(0 < p_1 + p_2 < 1)$, i.e. for each $(s, i) \in \{0, \ldots, n\} \times \{0, \ldots, n\}$ such that $s + i \leq n$,

$$\mathbb{P}(S^n(0) = s, I^n(0) = i) = \frac{n!}{s!i!(n-s-i)!} p_1^s p_2^i (1 - p_1 - p_2)^{n-s-i}.$$

This second assumption will be held from now on, and although the parameters of this distribution can be estimated by taking a sample at time zero, in the sequel it will be assumed these parameters are known.

Note that the solution $\chi = (\sigma, \iota, \rho)^\top$ to (3) satisfies:

$$\begin{aligned} \frac{\mathrm{d}\sigma}{\mathrm{d}t}(t) &= -\beta\sigma(t)\iota(t) \\ \frac{\mathrm{d}\iota}{\mathrm{d}t}(t) &= \beta\sigma(t)\iota(t) - \gamma\iota(t) \\ \frac{\mathrm{d}\rho}{\mathrm{d}t}(t) &= \gamma\iota(t), \end{aligned} \tag{17}$$

with initial condition $(\sigma(0), \iota(0), \rho(0)) = (p_1, p_2, 1 - p_1 - p_2)$.

6.2. Asymptotic normality

As a consequence of Theorem 5.1, the following proposition is stated, which shows the parameter estimators for the SIR epidemic model are asymptotically normal. In this subsection all notations and facts given on the preceding subsection will be maintained.

Proposition 6.1. *Let $\lambda = (\beta, \gamma)^\top$, $T > 0$ and $\widehat{\lambda^n} = (\widehat{\beta^n}, \widehat{\gamma^n})^\top$. Then, $\{\sqrt{n}(\widehat{\lambda^n}(T) - \lambda)\}_{n\in\mathbb{N}}$ converges in distribution to a normal random bivariate vector with mean zero and variance-covariance matrix*

$$\Sigma(T) = \begin{pmatrix} \beta^2/(\sigma(0) - \sigma(T)) & 0 \\ 0 & \gamma^2/(\rho(T) - \rho(0)) \end{pmatrix}.$$

Proof. From the Kolmogorov Law of Large Numbers, $\{\chi^n(0)\}_{n\in\mathbb{N}}$ converges almost sure to $\chi_0 = (p_1, p_2, 1 - p_1 - p_2)^\top$ and hence, Condition (5.1.1) holds.

Let $\xi_k^n(1) = \Delta A^n(t_k^n) - \beta\sigma^n(t_{k-1}^n)\iota^n(t_{k-1}^n)$ and $\xi_k^n(2) = \Delta B^n(t_k^n) - \gamma\iota^n(t_{k-1}^n)$. In order to verify Condition (5.1.2), it suffices to prove for each $i = 1, \ldots, m$, the double sequence $\{\xi_k^n(i)^2; 0 \leq k \leq n, n \in \mathbb{N}\}$ is uniformly integrable. For this purpose, we note that, for a random variable X with Binomial distribution, the following inequality holds:

$$\mathbb{E}(|X - \mathbb{E}(X)|^3) \leq 8\,\mathbb{E}(X)^3 + 6\,\mathbb{E}(X)^2 + \mathbb{E}(X).$$

Consequently, according to our approach, for each $i = 1, 2$, we have

$$\sup\{\mathbb{E}(|\xi_k^n(i)|^3) : n, k \in \mathbb{N}\} \leq 8\zeta^3 + 6\zeta^2 + \zeta,$$

where

$$\zeta = \begin{cases} \beta \text{ if } i = 1 \\ \gamma \text{ if } i = 2. \end{cases}$$

Hence, $\{\xi_k^n(i)^2\}_{n,k\in\mathbb{N}}$ is uniformly integrable and Condition (5.1.2) holds.

Let

$$v_A^n(t) = \frac{1}{n}\sum_{k=1}^{[nt]} \mathbb{E}([\Delta A^n(t_k^n) - \beta S^n(t_{k-1}^n)\iota^n(t_{k-1}^n)\Delta t_n]^2 | \mathcal{F}_{k-1}^n)$$

and

$$v_B^n(t) = \frac{1}{n}\sum_{k=1}^{[nt]} \mathbb{E}([\Delta B^n(t_k^n) - \gamma I^n(t_{k-1}^n)\Delta t_n]^2 | \mathcal{F}_{k-1}^n).$$

We have,

$$v_A^n(t) = \frac{1}{n}\sum_{k=1}^{[nt]} \beta\sigma^n(t_{k-1}^n)\iota^n(t_{k-1}^n)(1-\beta\iota^n(t_{k-1}^n)\Delta t_n),$$

and

$$v_B^n(t) = \frac{1}{n}\sum_{k=1}^{[nt]} \gamma\iota^n(t_{k-1}^n)(1-\gamma\Delta t_n).$$

Since $0 \leq v_A^n(t) \leq \beta t$ and $0 \leq v_B^n(t) \leq \gamma t$, the sequences $\{v_A^n(t)/n\}_{n\in\mathbb{N}}$ and $\{v_B^n(t)/n\}_{n\in\mathbb{N}}$ converge to zero and hence, Condition (4.1.2) holds. Thus, assumptions of Theorem 4.1 are satisfied and consequently, $\{\chi^n\}_{n\in\mathbb{N}}$ converges to χ uniformly in probability over compact subsets of $\mathbb{R}_+$. This fact along with

$$v_A^n(t) = \int_0^t \beta\sigma^n(u)\iota^n(u)\,\mathrm{d}u + \mathrm{O}(1/n) \quad \text{and} \quad v_B^n(t) = \int_0^t \gamma\iota^n(u)\,\mathrm{d}u + \mathrm{O}(1/n),$$

allow to conclude for each $t \geq 0$, $\{v_A^n(t)\}_{n\in\mathbb{N}}$ and $\{v_B^n(t)\}_{n\in\mathbb{N}}$ converge in probability to

$$v_A(t) = \int_0^t \beta\sigma(u)\iota(u)\,\mathrm{d}u \quad \text{and} \quad v_B(t) = \int_0^t \gamma\iota(u)\,\mathrm{d}u,$$

respectively.

Thus, as a consequence of Theorem 5.1, we have that $\{\sqrt{n}(\widehat{\gamma^n}(T) - \gamma)\}_{n\in\mathbb{N}}$ converges in distribution to a normal random bivariate vector having mean zero and variance-covariance matrix

$$\Sigma(T) = \begin{pmatrix} \beta/\int_0^T \sigma(u)\iota(u)\,\mathrm{d}u & 0 \\ 0 & \gamma/\int_0^T \iota(u)\,\mathrm{d}u \end{pmatrix}.$$

From (6) we have $\sigma(0) - \sigma(T) = \beta\int_0^T \sigma(u)\iota(u)\,\mathrm{d}u$ and $\rho(T) - \rho(0) = \gamma\int_0^T \iota(u)\,\mathrm{d}u$. Hence,

$$\Sigma(T) = \begin{pmatrix} \beta^2/(\sigma(0)-\sigma(T)) & 0 \\ 0 & \gamma^2/(\rho(T)-\rho(0)) \end{pmatrix},$$

and this concludes the proof. □

Since Proposition 6.1 involves a limit distribution depending on the solution of (17), this proposition may be of limited utility. However, Proposition 4.1 and Slutsky's theorem have, as a consequence, the following corollary, which is an alternative to this possible difficulty.

Corollary 6.1. *Let $T > 0$ and*

$$Q_n(T) = \sqrt{n}\begin{pmatrix} (\sigma(0)-\sigma^n(T))^{1/2}(\widehat{\beta^n}(T)-\beta) \\ (\rho^n(T)-\rho(0))^{1/2}(\widehat{\gamma^n}(T)-\gamma) \end{pmatrix}.$$

Then, $\{Q_n(T)\}_{n\in\mathbb{N}}$ converges in distribution to a normal random bivariate vector with mean zero and covariance matrix

$$I = \begin{pmatrix} \beta^2 & 0 \\ 0 & \gamma^2 \end{pmatrix}.$$

6.3. Hypothesis test for the infection rate

In this section, we are interested in proving whether the parameter β of this SIR epidemic model belongs to a subset of the parametric set. This analysis, which is known as hypothesis test, will intend to define whether the infection rate is bigger than a fixed value $\beta_0 > 0$, i.e. the null hypothesis is stated as $H_0 : \beta = \beta_0$ against the one-side alternative $H_1 : \beta > \beta_0$. Since under H_0, $\{\widehat{\beta^n}(T)\}_{n\in\mathbb{N}}$ converges in probability to β_0, H_0 shall be rejected in favor of H_1 when $\widehat{\beta^n}(T)$ is too large, i.e. when $\widehat{\beta^n}(T) > u_n$, being u_n determined by $\mathbb{P}(\widehat{\beta^n}(T) > u_n|H_0) \leq \alpha$. Here, α is the preassigned level of significance which controls the probability of falsely rejecting H_0 when H_0 is true.

By means of Corollary 6.1, it is possible to obtain an approximate value u_n when n is large. For this purpose, $\mathbb{P}(\widehat{\beta^n}(T) > u_n|H_0) \leq \alpha$ is replaced by the following weaker requirement:

$$\mathbb{P}(\widehat{\beta^n}(T) > u_n|H_0) \to \alpha, \quad \text{as} \quad n \to \infty.$$

From the Proposition 6.1, under H_0, $\{\sqrt{n}(\sigma(0) - \sigma(T))^{1/2}(\widehat{\beta^n}(T) - \beta_0)/\beta_0\}_{n\in\mathbb{N}}$ converges in distribution to a random variable having mean zero and variance one. Therefore, with an asymptotic level of significance α,

$$\sqrt{n}(\sigma(0) - \sigma(T))^{1/2}(u_n - \beta_0)/\beta_0 \to t_\alpha \text{ as } n \to \infty,$$

where $1 - \Phi(t_\alpha) = \alpha$. Here, Φ is the cumulative distribution function of a standard normal variable. Hence, $u_n = \beta_0 + \beta_0 t_\alpha/\sqrt{n(\sigma(0) - \sigma(T))} + o(1/n)$. In particular, we can choose $u_n = \beta_0 + \beta_0 t_\alpha/\sqrt{n(\sigma(0) - \sigma(T))}$ and therefore, with an asymptotic level of significance α, a critical region for the test is

$$R_1(n) = \{\sqrt{n}(\sigma(0) - \sigma(T))^{1/2}(\widehat{\beta^n}(T) - \beta_0)/\beta_0 \geq t_\alpha\}.$$

Under H_0, $\sigma(T)$ is known but it can not be explicitly obtained due to the fact that (6) does not admit a closed-form solution. However, we can take advantage of the fact that $\{\sigma^n(T)\}_{n\in\mathbb{N}}$ converges in probability to $\sigma(T)$, and by Slutzky's theorem, it can be obtained that

$$R_2(n) = \{\sqrt{n}(\sigma(0) - \sigma^n(T))^{1/2}(\widehat{\beta^n}(T) - \beta_0)/\beta_0 \geq t_\alpha\}$$

is a critical region for the test with an asymptotic level of significance α.

Let $\Delta > 0$ and let us consider the alternative hypothesis

$$H_1 : \beta = \beta_0 + \Delta/\sqrt{n}.$$

In this case, under H_1 the power of the test $\pi_n(\beta) = \mathbb{P}(R(n)|H_1)$, where $R(n) = R_1(n)$ or $R(n) = R_2(n)$, converges to $\Phi(\frac{(\sigma(0)-\sigma(T))^{1/2}\Delta - t_\alpha\beta_0}{\beta})$.

6.4. Relative removal-rate and basic reproduction number

As pointed out in Subsection 4.2, a fundamental concept resulting from the mathematical theory of the general deterministic epidemic model is the relative removal-rate, see for instance [2] and [3]. This number is defined as $\tau = \gamma/\beta$, where γ and β represent the removal and infection rates, respectively, and it plays a crucial part in determining the probable

occurrence of an epidemic outbreak. Actually, due to the fact that in the deterministic version of this model, we have (see (17))

$$\frac{\mathrm{d}\iota}{\mathrm{d}t}(t) = \beta\sigma(t)\iota(t) - \gamma\iota(t),$$

by assuming $\iota(0) > 0$, one has the derivative of ι at zero is bigger than zero, if and only if, $\tau < \sigma(0)$. Consequently, unless the initial density of susceptible individuals is bigger than the relative removal-rate, no epidemics can start. It is worth pointing out that this parameter is connected with the basic reproduction number, which is defined as the number of cases generated by one infective over the period of infectivity when that infective was introduced into a large population of susceptible individuals. See for instance, [16], or, [14]. Actually, it becomes $R_0 = \sigma(0)/\tau$. Notice $\sigma(0) \approx 1$ when a large population of susceptible individuals is considered. Hence, an epidemic can start with a positive probability, if $R_0 > 1$, and it will die out quickly if $R_0 \leq 1$. The knowledge of the basic reproduction number allows the establishment of vaccination policies when necessary, in order to reduce the number of susceptible individuals in a population to such a level that R_0 is brought below the unity threshold. Moreover, it is useful to carry out the following hypothesis test:

$$\mathrm{H}_0 : R_0 = 1 \quad \text{against} \quad \mathrm{H}_1 : R_0 > 1.$$

A natural critical region for the test is $\{\widehat{R_0^n}(T) > u_n\}$, where $\widehat{R_0^n}(T) = \sigma(0)\widehat{\beta^n}(T)/\widehat{\gamma^n}(T)$ and u_n is a constant which should be determined by $\mathbb{P}(\widehat{R_0^n}(T) > u_n|\mathrm{H}_0) \leq \alpha$, being α a preassigned level of significance controlling the probability of falsely rejecting H_0 when H_0 is true. The following proposition allows us to carry out the mentioned hypothesis test.

Proposition 6.2. *Let Δ be a non-negative real number and $\widehat{U}_n(T) = \sqrt{n}(\widehat{R_0^n}(T) - 1)$. Then, under local alternatives having the form*

$$\mathrm{H}_\Delta : R_0 = 1 + \Delta/\sqrt{n},$$

$\widehat{U_n}(T)$ has as asymptotically normal distribution with mean Δ and variance $v_0(T)^2$, where

$$v_0(T) = \left(\frac{1}{\sigma(0) - \sigma(T)} + \frac{1}{\rho(T) - \rho(0)}\right)^{1/2}.$$

Proof. Under H_Δ we have $\widehat{U_n}(T) = \sqrt{n}(\widehat{R_0^n}(T) - R_0) + \Delta$. Let $f(x, y) = x/y$, $(x, y \in \mathbb{R}, y \neq 0)$. By Taylor's theorem, we have

$$\begin{aligned}\sqrt{n}(\widehat{R_0^n}(T) - R_0) &= \sqrt{n}(f(\sigma(0)\widehat{\beta^n}(T), \widehat{\gamma^n}(T)) - f(\sigma(0)\beta, \gamma)) \\ &= \frac{\sigma(0)}{\gamma}\sqrt{n}(\widehat{\beta^n}(T) - \beta) - \frac{\sigma(0)\beta}{\gamma^2}\sqrt{n}(\widehat{\gamma^n}(T) - \gamma) + E_n,\end{aligned}$$

where

$$E_n = -\frac{\sigma(0)\sqrt{n}}{\gamma_n^2}(\widehat{\beta^n}(T) - \beta)(\widehat{\gamma^n}(T) - \gamma) + \frac{\sigma(0)^2\delta_n\sqrt{n}}{\gamma_n^3}(\widehat{\gamma^n}(T) - \gamma)^2$$

and (δ_n, γ_n) is a random point between $(\sigma(0)\widehat{\beta^n}(T), \widehat{\gamma^n}(T))$ and $(\sigma(0)\beta, \gamma)$.

From Propositions 5.1 and 6.1, $\{E_n\}_{n\in\mathbb{N}}$ converges in probability to zero. Hence, it follows from Proposition 6.1 that $\widehat{U_n}(T)$ is asymptotically normal with mean Δ and variance $v_0^2(T)$. Therefore, this proof is complete. □

The constant u_n can be chosen as $u_n = 1 + t_\alpha v_0(T)/\sqrt{n} + o(1/\sqrt{n})$, where t_α satisfies $1 - \Phi(t_\alpha) = \alpha$. In particular, we can choose $u_n = 1 + t_\alpha v_0(T)/\sqrt{n}$, and therefore, with an asymptotic level of significance α, a critical region for the test is $\{\widehat{U}_n(T)/v_0(T) > t_\alpha\}$. Note also that, under H_Δ, the power of the test $\pi_n(\Delta) = \mathbb{P}(\widehat{U}_n(T)/v_0(T) > t_\alpha | H_\Delta)$ converges to $\pi(\Delta) = \Phi(\Delta/v_0(T) - t_\alpha)$.

Remark 6.1. *An application of the preceding proposition is to calculate the approximate power of the test, with a level of significance α, relative to*

$$H_0 : R_0 = 1 \quad \textit{against} \quad H_1 : R_0 = r,$$

where $r > 1$.

We interpret Δ in H_Δ as $\sqrt{n}(r-1)$ and approximate the power of the test by means of $\pi = \mathbb{P}(\mathcal{N} > t_\alpha | H_\Delta)$, where $\mathcal{N}$ is a normal random variable with mean $\frac{\sqrt{n}(r-1)}{rv_0(T)}$ and variance 1. Consequently, the power of the test can be approximated by

$$\pi = \Phi\left(\frac{\sqrt{n}(r-1)}{rv_0(T)} - t_\alpha\right).$$

7. Numerical simulations for the general epidemic model

In order to carry out asymptotical inference for a great number of epidemic models, results stated in Section 5 can be applied. In this subsection, statistical inference for the General Epidemic Model is developed. In this section we maintain notations used in Section 6.

7.1. Validating the population size

In this subsection, some numerical simulations are carried out in order to validate the appropriate population size under which the results contained in this work are applicable.

Let

$$X_n(T) = \begin{pmatrix} \frac{\sqrt{\sigma(0)-\sigma(T)}}{\beta} & 0 \\ 0 & \frac{\sqrt{\rho(T)-\rho(0)}}{\gamma} \end{pmatrix} \begin{pmatrix} \widehat{\beta^n}(T) \\ \widehat{\gamma^n}(T) \end{pmatrix}.$$

According to Proposition 6.1, the square Euclidean norm of $X_n(T)$, namely $\|X_n(T)\|^2$, has asymptotically $\chi^2(2)$ distribution. The positive part of the real straight line is partitioned in m subintervals, which are determined by $0 = t_0 < t_1 < \cdots < t_{m-1} < \infty$. Let F denote the $\chi^2(2)$ distribution function. For $\beta = 2$, $\gamma = 1$ and $T = 10$, $\|X_n(T)\|^2$ is simulated 1,000 times with $m = 10$, where $t_1, \ldots, t_9$ have been chosen in such a way that $F(t_i) = i/10$, i.e., $t_1 = 0.21$, $t_2 = 0.45$, $t_3 = 0.71$, $t_4 = 1.02$, $t_5 = 1.39$, $t_6 = 1.83$, $t_7 = 2.41$, $t_8 = 3.22$ and $t_9 = 4.61$. In order to $F(t_{10}) = 1$, t_{10} is defined as ∞. For different population sizes ($n = 20, 50, 100, 300, 500, 1000$), percentages obtained into the corresponding subintervals are presented in Table 1. From Proposition 6.1, better approximations correspond to percentages close to 10%. These closeness are measured by means of the standard deviation (SD) of the observations corresponding to the different population sizes, which are indicated at the last column of Table 1 below.

n	$[t_0, t_1[$	$[t_1, t_2[$	$[t_2, t_3[$	$[t_3, t_4[$	$[t_4, t_5[$	$[t_5, t_6[$	$[t_6, t_7[$	$[t_7, t_8[$	$[t_8, t_9[$	$[t_9, t_{10}[$	SD
30	10.7	8.9	8.3	9.0	7.4	7.1	9.1	8.4	9.1	22.0	43.3
50	10.8	7.6	10.0	9.3	7.8	10.0	10.6	9.1	8.3	16.5	25.4
100	11.0	9.6	10.6	9.3	10.3	10.2	8.3	8.1	10.8	11.8	11.8
300	10.0	10.0	9.3	8.6	8.9	10.3	10.0	9.8	10.8	12.3	10.4
500	9.3	10.8	9.2	9.2	9.9	11.0	9.5	10.2	11.3	9.6	7.9
1000	9.4	9.3	10.0	10.3	10.8	10.1	9.8	9.0	9.9	11.4	7.2

Table 1. Percentages of observations of $\|X_n(T)\|^2$ for the indicated population size, $\beta = 2$, $\gamma = 1$ and $T = 10$.

By carrying out five simulations of a random variable $\chi^2(2)$, 1,000 times each, the percentages of values that resulted in the corresponding subintervals for each of the five simulations, with the corresponding SD at the last column, are showed in Table 2 below. From Tables 1 and

N	$[t_0, t_1[$	$[t_1, t_2[$	$[t_2, t_3[$	$[t_3, t_4[$	$[t_4, t_5[$	$[t_5, t_6[$	$[t_6, t_7[$	$[t_7, t_8[$	$[t_8, t_9[$	$[t_9, t_{10}[$	SD
1	10.3	10.3	11.4	9.3	9.7	7.8	8.8	11.2	11.5	9.6	12.0
2	10.4	9.6	9.8	10.4	10.3	11.4	10.5	8.7	10.2	8.4	9.1
3	10.5	11.3	10.3	8.8	9.7	11.0	11.0	9.4	9.4	8.6	9.6
4	9.7	12.3	8.8	9.5	9.7	9.8	8.9	11.7	9.8	9.8	11.2
5	10.6	10.4	10.3	10.1	9.5	8.1	9.9	10.3	9.7	11.1	8.1
										Average SD:	10.0

Table 2. Percentages of $\chi^2(2)$ observations into each subintervals for 5 series of 1,000 trials each.

2, we can conclude that, for a population size over 300, the distribution of $\|X_n(T)\|^2$ is quite approximate to the $\chi^2(2)$ distribution. In Table 3, estimates of β, γ and R_0 have been simulated for different population size. Table 3 corroborates the convergence in probability demands

n	$\widehat{\beta^n}(T)$	$\widehat{\gamma^n}(T)$	$\widehat{R_0^n}(T)$
100	2.87	0.82	1.76
200	1.73	1.17	0.74
500	1.82	0.99	0.93
1,000	1.91	0.98	0.99
2,000	1.95	1.01	0.97
5,000	2.01	0.99	1.01
10,000	2.00	1.00	1.01

Table 3. Estimates for $\widehat{\beta^n}(T)$, $\widehat{\gamma^n}(T)$ and $\widehat{R_0^n}(T)$ for the indicated population size, $\beta = 2$, $\gamma = 1$ and $T = 10$.

bigger population size in order to obtain a suitable approximation of the parameters. We appreciate, appropriate estimates for β, γ and R_0 are obtained for population sizes over $n = 5,000$.

For $\beta = 2$, $\gamma = 1$ and $T = 10$, $X_n(T)$ is simulated 1,000 times again. Let $X_{n,1}(T)$ and $X_{n,2}(T)$ the first and second row of $X_n(T)$, i.e. $X_n(T) = (X_{n,1}(T), X_{n,2}(T))^\top$. In Figure 4 below, histograms of $X_{n,1}(T)$ and $X_{n,2}(T)$ are showed for $n = 300$. It is observed that their corresponding frequencies of simulated values are quite close to the standard normal density. Let $\widehat{U}_n(T)$ and $v_0(T)$ be as in Proposition 6.2 and denote $X_{n,3}(T) = \widehat{U}_n(T)/(R_0 v_0(T))$. By observing in Figure 5 the histogram of $X_{n,3}(10)$, we notice the graphical frequency of its simulated values is also quite close to the standard normal density.

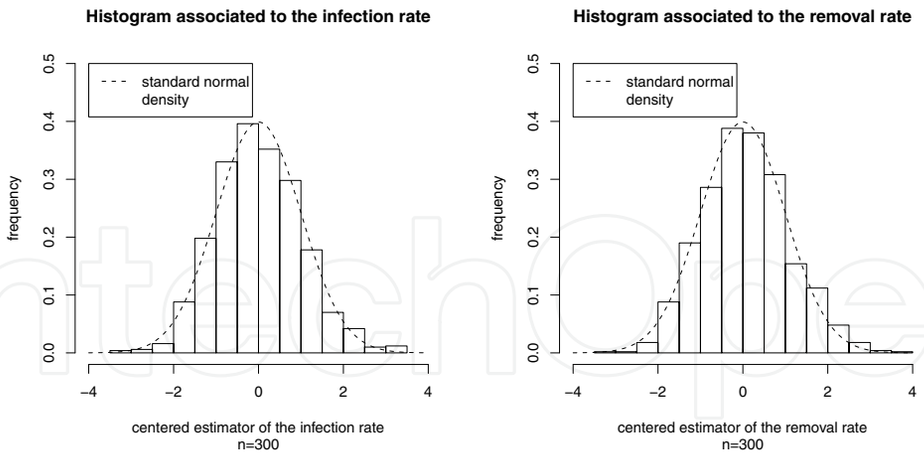

Figure 4. Histograms of $X_{n,1}(10)$ and $X_{n,2}(10)$ for a population size equals 300.

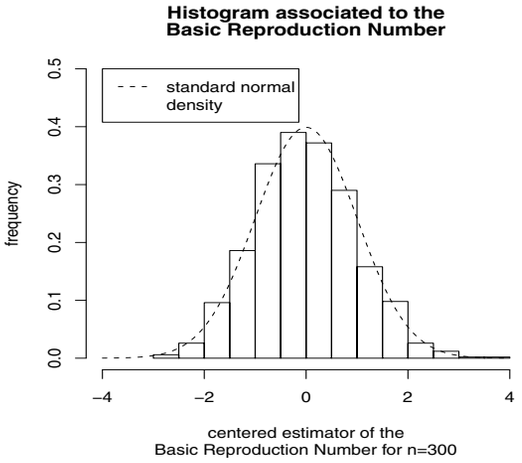

Figure 5. Histograms of $X_{n,3}(10)$ for a population size equals 300.

7.2. The power function of the test for the basic reproduction number

According to Proposition 6.2, the power function $\pi_n(\Delta)$ converges to the asymptotic power function (apf) $\pi(\Delta) = \Phi(\Delta/v_0(T) - t_\alpha)$. The level of significance of the test is assumed to be $\alpha = 0.05$ and consequently, $t_\alpha = 1.644853$. Notice $\pi_n(\Delta)$ can be estimated by $\widehat{\pi}_n(\Delta)$, the frequency H_0 is rejected (rf). Hence, $\widehat{\pi}_n(\Delta)$ is the frequency of times that $\widehat{U}_n(T)/v_0(T) > 1.644853$, under $H_\Delta : R_0 = 1 + \Delta/\sqrt{n}$. By the Law of Large Number, a strong consistent estimator of $\pi_n(\Delta)$ is given by

$$\widehat{\pi}_n(\Delta) = \frac{1}{M}\sum_{k=1}^{M} I_{\{\widehat{U}_{n,k}(T)/v_0(T)>1.644853\}},$$

where $\widehat{U}_{n,1}(T), \ldots, \widehat{U}_{n,M}(T)$ are independent and identically distributed random variables, which have the same distribution as $\widehat{U}_n(T)$ under H_Δ. Even, it can prove that the convergence of $\widehat{\pi}_n$ to π_n, as M goes to ∞, is uniform on compact subsets of $\mathbb{R}_+$, I.e., for each $T > 0$,

$$\lim_{M\to\infty} \sup_{0\le\Delta\le T} |\widehat{\pi}_n(\Delta) - \pi_n(\Delta)| = 0, \quad a.s.$$

In Table 4 below, $M = 1,000$ simulations of $\widehat{\pi}_n(\Delta)$ are carried out for a population size equals 300 and for 150 values of Δ between 0 and 14.9, with step sizes of equal length $\Delta = 0.1$. Since

Δ	.0	.1	.2	.3	.4	.5	.6	.7	.8	.9
0	0.000	0.001	0.001	0.004	0.002	0.005	0.002	0.007	0.008	0.005
1	0.004	0.014	0.004	0.021	0.010	0.015	0.015	0.016	0.018	0.020
2	0.023	0.040	0.032	0.036	0.042	0.047	0.052	0.059	0.063	0.060
3	0.078	0.075	0.104	0.085	0.120	0.105	0.136	0.140	0.143	0.150
4	0.151	0.183	0.190	0.202	0.201	0.210	0.247	0.267	0.289	0.282
5	0.296	0.329	0.343	0.357	0.342	0.376	0.415	0.431	0.427	0.446
6	0.467	0.463	0.509	0.536	0.524	0.558	0.579	0.570	0.604	0.595
7	0.632	0.628	0.655	0.663	0.703	0.706	0.723	0.719	0.730	0.775
8	0.762	0.787	0.765	0.807	0.814	0.798	0.805	0.842	0.845	0.846
9	0.881	0.854	0.874	0.880	0.890	0.900	0.913	0.901	0.913	0.904
10	0.911	0.927	0.934	0.937	0.941	0.941	0.944	0.942	0.958	0.957
11	0.972	0.961	0.969	0.967	0.988	0.973	0.978	0.986	0.988	0.981
12	0.982	0.987	0.982	0.987	0.988	0.989	0.981	0.995	0.990	0.996
13	0.991	0.992	0.993	0.998	0.996	0.998	0.998	0.997	0.994	0.997
14	0.997	1.000	1.000	0.999	0.999	0.999	0.999	1.000	0.998	0.998

Table 4. Frequency of times $\widehat{U}_n(T)/v_0(T) > 1.644853$ for different values of Δ.

for $\Delta = 0$ the power of the test should be close to 0.05, the random function $\widehat{\pi}_n$ cannot be considered a good estimator for π_n, at least for $n = 300$. However, the asymptotic power function π seems to give a better approximation for π_n. The values of $\pi(\Delta)$ are given in Table 5 below for 150 values of Δ between 0 and 14.9, with step sizes of equal length $\Delta = 0.1$. In

Δ	.0	.1	.2	.3	.4	.5	.6	.7	.8	.9
0	0.050	0.053	0.056	0.059	0.063	0.067	0.071	0.075	0.079	0.083
1	0.088	0.093	0.098	0.103	0.108	0.114	0.119	0.125	0.132	0.138
2	0.144	0.151	0.158	0.165	0.173	0.180	0.188	0.196	0.205	0.213
3	0.222	0.230	0.239	0.249	0.258	0.268	0.277	0.287	0.297	0.307
4	0.318	0.328	0.339	0.350	0.361	0.372	0.383	0.394	0.405	0.417
5	0.428	0.440	0.451	0.463	0.475	0.486	0.498	0.510	0.521	0.533
6	0.545	0.556	0.568	0.579	0.591	0.602	0.613	0.625	0.636	0.647
7	0.657	0.668	0.679	0.689	0.699	0.709	0.719	0.729	0.739	0.748
8	0.757	0.766	0.775	0.784	0.792	0.801	0.809	0.817	0.824	0.832
9	0.839	0.846	0.853	0.860	0.866	0.872	0.878	0.884	0.890	0.895
10	0.900	0.905	0.910	0.915	0.919	0.923	0.928	0.931	0.935	0.939
11	0.942	0.946	0.949	0.952	0.955	0.957	0.960	0.962	0.965	0.967
12	0.969	0.971	0.973	0.974	0.976	0.978	0.979	0.981	0.982	0.983
13	0.984	0.985	0.986	0.987	0.988	0.989	0.990	0.991	0.991	0.992
14	0.992	0.993	0.994	0.994	0.994	0.995	0.995	0.996	0.996	0.996

Table 5. Values of $\pi(\Delta)$ for different values of Δ.

Figure 6 below, a graphical comparison between π and $\widehat{\pi}_n$ is given for a population size equals $n = 300$.

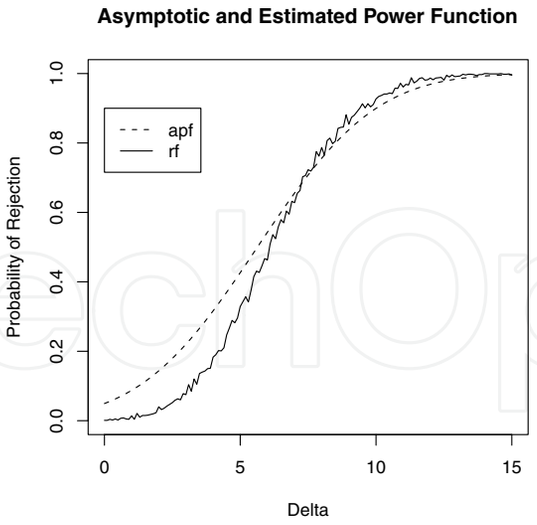

Figure 6. Grafics of $\pi(\Delta)$ and $\widehat{\pi}_n(\Delta)$ for a population size equals $n = 300$.

8. Conclusions

A wide class of discrete-time stochastic epidemic models was introduced and analyzed in this work, the convergence of the stochastic model to the deterministic one as the population increases was proved. Moreover, it was proved the convergence in distribution of a process, depending on the fluctuations between the stochastic and deterministic models, to the solution to a stochastic differential equation. The statistical analysis was focused in proving asymptotic normality of natural martingale estimators, for the parameters of the model, and these results were applied to hypothesis tests for the General Epidemic Model. As a consequence, a hypothesis test for the reproduction number was stated and numerical simulations validated the size populations under which the results are useful.

The stochastic models considered here represent an alternative modeling to those using counting processes and having transitions occurring at random times. However, both are asymptotically consistent with their deterministic counterparts. The advantage of considering stochastic models at discrete times is that it is not necessary to observe the epidemic over a long period of time; however, in order to attain the asymptotical consistency mentioned, our model requires frequent observation when dealing with a large population. Another important conclusion is obtained from Subsection 4.2, where numerical simulations was carried out for the SIS epidemic model. Indeed, the simulation showed large fluctuations of the SIS stochastic epidemic model, for $n = 10$, regarding the deterministic one. However, by simulating the process for $n = 1,000$ it could be appreciated the trajectories of the deterministic and stochastic models were quite closed. Moreover, Theorem 2 provides confidence bounds which give an insight of the fluctuations of the stochastic model regarding the deterministic one. The simulation in Subsection 4.2 for the SIS epidemic model shows the coherence between these bounds and the simulated process (Figure 3). As a conclusion, the deterministic and stochastic model are quite different for small populations, while for large populations both models perform similarly. However, it is worth noting, although the stochastic and deterministic trajectories are similar for large size populations, an attractor state for the deterministic model is not necessarily an attractor for the stochastic model. As shown in Subsection 4.2, this is the case for the SIS epidemic model.

In order to develop statistical inference on the parameters of the model, martingale estimators are proposed. This fact allows to obtain closed form for the estimators. Even, such estimators can be obtained when the distribution of the process governing the epidemic is not completely known.

All the models belong to the class that we are presenting in this chapter assume a stochastic latent period, and consequently, the Reed-Frost model are not included here. As a matter of fact, none of these models could be considered an extension or a particular case of the Reed-Frost model. Moreover, the modeling presented in this chapter need not be Markovian, even though some Markovian epidemic models are included in this setting.

Acknowledgements

This work was partially supported by Instituto de Matemática of the Pontificia Universidad Católica de Valparaíso and by FONDECYT under project 1120879.

Author details

Raúl Fierro
Instituto de Matemática, Pontificia Universidad Católica de Valparaíso, Centro de Investigación y Modelamiento de Fenómenos Aleatorios, Universidad de Valparaíso, Chile

9. References

[1] Andersson, H. & Britton, T. [2000]. *Stochastic Epidemic Models and Their Statistical Analysis. Lecture Notes in Statistics 151*, Springer, New York.
[2] Bailey, N. [1975]. *The Mathematical Theory of Infectious Diseases*, Griffin & Co., New York.
[3] Ball, F. [1983]. The threshold behavior of epidemic models, *Journal of Applied Probability* Vol. 20(No. 2): 227–241.
[4] Ball, F. & O'Neill, P. [1993]. A modification of the general stochastic epidemic motivated by aids modelling, *Journal of Applied Probability* Vol. 25: 39–62.
[5] Barbour, A. [1972]. The principle of the diffusion of arbitrary constants, *Journal of Applied Probability* Vol. 9(No. 3): 519–541.
[6] Barbour, A. [1974]. On a functional central limit theorem for markov population processes, *Advances in Applied Probability* Vol. 6(No. 1): 21–39.
[7] Becker, N. [1993]. Parametric inference for epidemic models, *Mathematical Biosciences* Vol. 117(No. 1-2): 239–251.
[8] Billingsley, P. [1968]. *Convergence of Probability Measures*, Wiley & Sons, New York.
[9] Buckley, F. & Pollet, P. [2009]. Analytical methods for a stochastic mainland-island metapopulation model, *in* R. Anderssen, R. Braddock & L. Newham (eds), *Modelling and Simulation Society of Australia and New Zealand and International Association for Mathematics and Computers in Simulation*, Camberra, Australia, pp. 1767–1773.
[10] Buckley, F. & Pollet, P. [2010]. Limit theorems for discrete-time metapopulation models, *Probability Surveys* Vol. 7: 53–83.
[11] Ethier, S. & Kurtz, N. [1986]. *The Mathematical Theory of Infectious Diseases*, Wiley & Sons, New York.
[12] Fierro, R. [2010]. A class of stochastic epidemic models and its deterministic counterpart, *Journal of the Korean Statistical Society* Vol. 39(No. 4): 397–407.

[13] Fierro, R. [2012]. Asymptotic distribution of martingale estimators for a class of epidemic models, *Statistical Methods and Applications* (to appear).
[14] Heesterbeek, J. & Dietz, K. [1996]. The concept of r_0 in epidemic theory, *Statistica Neerlandica* Vol. 50(No. 1): 89–110.
[15] Hetcote, H. & Van Ark, J. [1987]. Epidemiological models for heterogeneous populations: proportionate mixing, parameter estimation and immunization programs, *Mathematical Biosciences* Vol. 84(No. 1): 85–118.
[16] Jacquez, J. & O'Neill, P. [1991]. Reproduction numbers and thresholds in stochastic epidemic models i. homogeneous populations, *Mathematical Biosciences* Vol. 107(No. 2): 161–186.
[17] Jacquez, J., Simon, C., Koopman, J., Sattenspiel, L. & Perry, T. [1988]. Modelling and analysing hiv transmission: The effect of contact patterns, *Mathematical Biosciences* Vol. 92(No. 2): 119–199.
[18] Kendall, D. [1956]. Deterministic and stochastic epidemic in closed populations, *Proceedings of Third Berkeley Symposium in Mathematical Statistics and Probability* Vol. 4: 149–165.
[19] Kermack, W. & McKendrick, A. [1927]. Contribution to the mathematical theory of epidemics, *Proceedings of the Royal Society of London A* Vol. 115: 700–721.
[20] Kurtz, T. [1970]. Solutions of ordinary differential equations as limits of pure jump markov processes, *Journal of Applied Probability* Vol. 7(No. 1): 49–58.
[21] Kurtz, T. [1971]. Limit theorems for sequences of jump markov processes approximating ordinary differential processes, *Journal of Applied Probability* Vol. 8(No. 2): 344–356.
[22] Kurtz, T. [1981]. *Approximation of Population Processes*, SIAM, Philadelphia.
[23] Lefèvre, C. [1990]. Stochastic epidemic models for s-i-r infectious diseases: a brief survey of the recent general theory, *in* J. Gabriel, C. Lefèvre & P. Picard (eds), *Stochastic Processes in Epidemic Theory. Lecture Notes in Biomathematics 86*, Springer-Verlag, New York, pp. 1–12.
[24] Lefèvre, C. & Picard, P. [1993]. An epidemic model with fatal risk, *Mathematical Biosciences* Vol. 117(No. 1-2): 127–145.
[25] Lehmann, E. [1999]. *Elements of Large-Sample Theory*, Springer-Verlag, New York.
[26] Lenglart, E. [1977]. Relation de domination entre deux processus, *Annales de l'Institut Henri Poincaré (B) Probabilités et Statistiques* Vol. 13(No. 2): 171–179.
[27] McVinish, R. & Pollett, P. [2010]. Limits of large metapopulations with patch dependent extintion probabilities, *Advances in Applied Probability* Vol. 42(No. 4): 1172–1186.
[28] Picard, P. & Lefèvre, C. [1993]. Distribution of the final state and severity of epidemics with fatal risk, *Stochastic Processes and their Applications* Vol. 48(No. 2): 277–294.
[29] Rebolledo, R. [1979]. La méthode des martingales appliquée á l'etude de la convergence en loi de processus, *Mémoires de la Société Mathématique de France* Vol. 62(No. 4): 1–125.