Cortés, J.; El-Labany, S.; Navarro-Quiles, A.; Selim, MM.; Slama, H. (2020). A comprehensive probabilistic analysis of SIR-type epidemiological models based on full randomized Discrete-Time Markov Chain formulation with applications. Mathematical Methods in the Applied Sciences. 43(14):8204-8222. https://doi.org/10.1002/mma.6482

Received 02 September 2019; | Revised XX; | Accepted XX

DOI: xxx/xxxx

ARTICLE TYPE

A comprehensive probabilistic analysis of SIR-type epidemiological models based on full randomized Discrete-Time Markov Chain formulation with applications

J.-C. Cortés[1] | S. K. El-Labany[2] | A. Navarro-Quiles*[1] | M. M. Selim[2] | H. Slama[2]

[1]Instituto Universitario de Matemática Multidisciplinar, Universitat Politècnica de València, Camino de Vera s/n, 46022 Valencia, España

[2]Physics Department, Faculty Science, Damietta University, New Damietta City 34517, Egypts

Correspondence
*A. Navarro-Quiles, Instituto Universitario de Matemática Multidisciplinar, Universitat Politècnica de València, Camino de Vera s/n, Edificio 8G, acceso C, 2^a planta, 46022 Valencia, Spain . Email: annaqui@doctor.upv.es

Present Address
Instituto Universitario de Matemática Multidisciplinar, Universitat Politècnica de València, Camino de Vera s/n, Edificio 8G, acceso C, 2^a planta, 46022 Valencia, Spain

Summary

This paper provides a comprehensive probabilistic analysis of a full randomization of SIR-type epidemiological models based on Discrete-Time Markov Chain formulation. The randomization is performed by assuming that all input data (initial conditions, the contagion and recovering rates involved in the transition matrix) are random variables instead of deterministic constants. In a first part of the paper, we determine explicit expressions for the so called first probability density function of each subpopulation identified as the corresponding states of the Markov chain (susceptible, infected and recovered) in terms of the probability density function of each input random variable. Afterwards, we obtain the probability density functions of the times until a given proportion of the population remains susceptible, infected and recovered, respectively. The theoretical contribution of this paper is completed by computing explicit expressions of important randomized epidemiological quantities, namely, the basic reproduction number, the effective reproduction number and the herd immunity threshold. All this theoretical information is derived under very general assumptions and taking extensive advantage of the Random Variable Transformation method via appropriate transformations for each one of the aforementioned target probability density functions. The second part of the paper is devoted to apply our theoretical findings to study, from a probabilistic standpoint, the dynamics of the pandemic influenza using real data from Egypt. The probabilistic study of pandemic influenza in Egypt is complemented by providing valuable information which is seldom displayed in epidemiological models.

KEYWORDS:
Randomized Discrete-Time Markov Chains, SIR epidemiological model, First probability density function, Random Variable Transformation method, Simulations

1 | INTRODUCTION AND MOTIVATION

Markov chains have demonstrated to be useful tools to model important problems appearing in medicine. In particular, they have been successfully applied to model the dynamics of epidemics, i.e., the rapid spread (transmission) of an infectious disease to a

[0]**Abbreviations:** 1-PDF, first probability density function; PDF, probability density function; RVT, random variable transformation; DTMC, discrete-time Markov chain

large number of individuals in a given population. Since the transmission of a disease is usually subject to complex factors whose nature is uncertain, and whose current state depends upon the previous levels of the disease, the application of probabilistic models, such as Markovian chains, is a natural and powerful approach. One important type of Markovian epidemiological models to describe the dynamics of a disease is the so-called SIR model formulated via Discrete-Time Markov Chains (DTMC)[1, ch. 2]. In the DTMC SIR epidemic model time is assumed to be discrete, $t = 01, 2, \ldots$, while random variables, defining the different states of the Markov chain, are discrete. Specifically, in the SIR model individuals of the population are categorized into one of the following three groups: Susceptible (S), Infected (I) and Recovered (R). The SIR model is appropriate when the following assumptions hold: the population is fixed, i.e., remains constant over the time; the only way an individual can leave the susceptible group is to become infected ($S \to I$) and, in the classical SIR model, the probability of being infected is independent of age, sex, social status, etc.; the only way an individual can leave the infected subpopulation is to recover from the disease ($I \to R$); once an individual has recovered, he/she develops permanent immunity. So, the class R represents individuals that are permanently immune. In addition, there is no inherited immunity, i.e., the whole population is susceptible to be infected. In the SIR model, homogeneous mixing hypothesis is also assumed, i.e., each individual has the same interactions with one another to the same degree[2]. Some diseases whose dynamics have been described using the SIR model include measles, mumps, chickenpox, smallpox, influenza, etc.[2,3,4,5,6]. In Figure 1 , we show the flow diagram associated to this Markovian epidemiological model, where $\alpha > 0$ denotes the infectious rate, i.e., the transition rate to pass from S to I, and $\beta > 0$ is the transition rate from I to R.

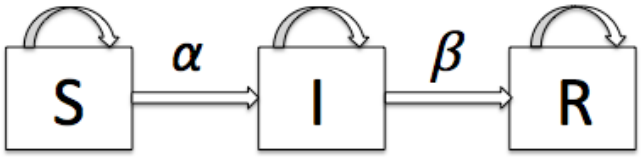


FIGURE 1 Flow diagram to SIR epidemiological model.

Given an initial vector $(s_0, i_0, r_0)^\top$, the SIR Markovian model can be formulated by

$$\begin{pmatrix} s_{n+1} \\ i_{n+1} \\ r_{n+1} \end{pmatrix} = T \begin{pmatrix} s_n \\ i_n \\ r_n \end{pmatrix}, \qquad n = 0, 1, 2, \ldots, \tag{1}$$

where s_n, i_n and r_n are the proportion of susceptible, infected and recovered subpopulations in cycle n, respectively. According to the classical homogeneous Markovian chain theory, we assume that $s_n + i_n + r_n = 1$, that is all the states form a closed system at every cycle n. This means that any individual can neither leave nor join the system, hence keeping the total population size constant over the time. Then, the initial cohort can be taken as $(s_0, i_0, 1 - s_0 - i_0)^\top$. Without loss of generality, we can assume that the transition matrix T is given by

$$T = \begin{pmatrix} \mathrm{e}^{-\alpha} & 0 & 0 \\ 1 - \mathrm{e}^{-\alpha} & \mathrm{e}^{-\beta} & 0 \\ 0 & 1 - \mathrm{e}^{-\beta} & 1 \end{pmatrix}, \tag{2}$$

since $\mathrm{e}^{-\alpha}$ and $\mathrm{e}^{-\beta}$ lie in the interval $]0, 1[$ when $\alpha, \beta \in]0, +\infty[$, so each entry of transition matrix T represents a probability and the sum of each column is 1. The way that this matrix has been represented in expression (2) can be interpreted from another view point using the Poisson distribution. It is well-known that the Poisson distribution represents the probability of a given number of events occurring in a fixed interval of time, whenever these events occur with a constant rate and independently of the time since the last event happened. The probability of observing k events in an interval is given by

$$\mathbb{P}[k; \lambda] = \mathrm{e}^{-\lambda} \frac{\lambda^k}{k!},$$

where $\lambda > 0$ is the average number of events per interval. In our case, taking $k = 0$ and $\lambda = \alpha$, the probability of not being infected is $\mathrm{e}^{-\alpha}$, which corresponds to the element $(1, 1)$ of transition matrix T. The element $(3, 1)$ represents the probability of the transition $S \to R$, which is zero. Then, the element $(2, 1)$ is given by $1 - \mathrm{e}^{-\alpha}$, since the sum of the columns must be 1.

Analogously for the transition rate β and for the elements in second column of matrix $\boldsymbol{T}$. Finally, if an individual reaches the R state it can not leave it since R is an absorbent state and the probability of staying in it is 1.

On the one hand, an aspect that confers classical homogeneous Markovian chains certain rigidity, in dealing with epidemiological models, is the fact that probabilities included in the transition matrix are assumed to be constants. As we have previously explained, in our context these probabilities represent the transmission rates of being infected (α) and recovered (β). Obviously, these parameters depend upon complex factors involving certain degree of randomness, such as age, sex, social status, genetic, weather, etc. This is a key aspect that, as previously indicated, in the classical homogeneous Markovian chains approach is neglected. On the other hand, the initial proportions, s_0 (susceptibles) and i_0 (infected), of the whole population are rarely known in an exact (deterministic) way but via surveys, so involving sampling errors. These facts aim us at treating both the transmission rates, α, β, and the initial conditions, s_0 and i_0, as random variables rather than deterministic constants.

Hereinafter, we will therefore assume that $s_0(\omega)$, $i_0(\omega)$, $\alpha(\omega)$ and $\beta(\omega)$, $\omega \in \Omega$, are independent and absolutely continuous random variables defined on a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$, being f_{s_0}, f_{i_0}, f_α and f_β their probability density functions (PDFs), respectively. The hypothesis of independence between the input data is realistic from a practical standpoint and it also facilitates computations and notational aspects throughout subsequent analysis, anyway it is worth pointing out that our approach can be entirely developed working with a joint PDF $f_{s_0,i_0,\alpha,\beta}$ instead. Given a random variable, say $x(\omega)$, henceforth its domain will be denoted by $\mathcal{D}(x(\omega))$.

To conduct our probabilistic study, we will take extensive advantage of the Random Variable Transformation (RVT) technique (see Theorem 1 later). Using this key result, we will be able to compute explicit expressions to the first probability density function (1-PDF) of the solution of the randomized DTMC SIR-type model, i.e., of the stochastic processes describing the dynamics of susceptibles, infected and recovery subpopulations. We will also be able to determine the PDFs of important random variables modelling relevant information of the disease. This is a relatively new approach that provides richer information about the dynamics of SIR-type epidemiological model. The RVT technique has been successfully applied to deal with both discrete and continuous models. In the former context, RVT method has been applied to perform a comprehensive probabilistic analysis of binary Markov chains with applications to model the spread of a new technology[7] and to model stroke disease using real data[8], for instance. In the latter framework, the RVT has been served to study relevant epidemiological continuos models formulated via random differential equations, namely, the SI and SIS-type models[9,10]. Recently, some of the authors have proposed a probabilistic analysis of the SIR-type continuous model[11]. In a sense, our study can be regarded as a complement of that contribution[11] since now we deal with its discrete counterpart. Besides, we now apply the randomized DTMC SIR-type model to study the dynamics of a disease using real data rather than just performing simulations. Finally, we stress that the RVT technique has also been applied to study other important problems in Biology and Engineering,[12,13,14,15,16,17,18,19].

The main goals of this paper are twofold: First to perform a comprehensive analysis of the randomized SIR Markovian model (1)–(2) by computing important probabilistic distributions associated to the dynamics of epidemics. Second, we will apply our theoretical findings to model the dynamics of pandemic influenza using real-world data from Egypt. To achieve these objectives, the paper is organized as follows: In Section 2, we will determine the 1-PDFs of each subpopulations. In Section 3, we will calculate the PDFs of the time until a certain proportion of the population remains susceptible, infected and recovered. Section 4 is devoted to complete the previous analysis by calculating the PDF of some key randomized epidemiological quantities as the basic reproduction number (R_0), the effective reproduction number (R_e) and the so-called herd immunity threshold (HIT). Section 5 is addressed to carry out some simulations and to model pandemic influenza using data from Egypt. To conduct this real application, we assign plausible parametric probabilistic distributions to every input data and then we determine the values of the parameters so that the mean of each subpopulation of the randomized SIR model, computed via its corresponding 1-PDF, adjusts the available real data using a mean square error as goodness-of-fit measure. Once the parameters of the probabilistic distributions allocated for each input random variable have been computed, we construct punctual (via the mean) and probabilistic (via confidence intervals) predictions for susceptible, infected and recovered subpopulations. Specifically, we construct the PDFs of each subpopulation, the PDFs of the times until a given proportion of the population remains in each state, as well as the PDFs of R_0, R_e and HIT. Finally, detailed interpretations of these key epidemiological quantities, in the context of the pandemic influenza in Egypt, are provided to stress the usefulness of computing them. Conclusions are outlined in Section 6.

2 | FIRST PROBABILITY DENSITY FUNCTION OF EACH SUBPOPULATION

This section is devoted to provide explicit expressions for the 1-PDF of susceptibles, infected and recovered subpopulations of the randomized Markovian chain SIR model (1)–(2). The key tool to calculate these 1-PDFs is the RVT method, which is stated in the following theorem. Observe that this result allows us to calculate the PDF, $f_{\mathbf{y}}(\mathbf{y})$, of an absolutely continuous random vector, $\mathbf{y}(\omega)$, which results from mapping another absolutely continuous random vector, $\mathbf{x}(\omega)$, whose the PDF, $f_{\mathbf{x}}(\mathbf{x})$, is known.

Theorem 1 (Multidimensional RVT method[20])**.** Let $\mathbf{x}(\omega) = (x_1(\omega), \dots, x_m(\omega))$ and $\mathbf{y}(\omega) = (y_1(\omega), \dots, y_m(\omega))$ be two m-dimensional absolutely continuous random vectors defined on a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let $\mathbf{r} : \mathbb{R}^m \to \mathbb{R}^m$ be a one-to-one deterministic transformation of $\mathbf{x}(\omega)$ into $\mathbf{y}(\omega)$, i.e., $\mathbf{y}(\omega) = \mathbf{r}(\mathbf{x}(\omega))$, for each $\omega \in \Omega$. Assume that $\mathbf{r}$ is continuous in $\mathbf{x}(\omega)$ and has continuous partial derivatives with respect to $\mathbf{x}(\omega)$, for each $\omega \in \Omega$. Then, if $f_{\mathbf{x}}(\mathbf{x})$ denotes the probability density function of the absolutely continuous random vector $\mathbf{x}(\omega)$, and $\mathbf{s} = \mathbf{r}^{-1} = (s_1(y_1, \dots, y_m), \dots, s_m(y_1, \dots, y_m))$ denotes the inverse of $\mathbf{r} = (r_1(x_1, \dots, x_m), \dots, r_m(x_1, \dots, x_m))$, the probability density function of the absolutely continuous random vector $\mathbf{y}(\omega)$ is given by

$$f_{\mathbf{y}}(\mathbf{y}) = f_{\mathbf{x}}(\mathbf{s}(\mathbf{y}))\,|J|\,, \tag{3}$$

where $|J|$, which is assumed to be different from zero, denotes the absolute value of the Jacobian defined by the determinant

$$J = \det \begin{pmatrix} \dfrac{\partial s_1(y_1, \dots, y_m)}{\partial y_1} & \cdots & \dfrac{\partial s_m(y_1, \dots, y_m)}{\partial y_1} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial s_1(y_1, \dots, y_m)}{\partial y_m} & \cdots & \dfrac{\partial s_m(y_1, \dots, y_m)}{\partial y_m} \end{pmatrix}.$$

The following technical lemma is a direct consequence of Theorem 1 and it will be required later to calculate the 1-PDFs of each subpopulation of the randomized Markovian SIR model (1)–(2).

Lemma 1. Let $x(\omega)$ be a positive absolutely continuous random variable with probability density function, $f_x(x)$. Then, the probability density function of the absolutely continuous random variable $y(\omega) = 1 - \mathrm{e}^{-x(\omega)}$ is given by

$$f_y(y) = f_x(-\ln(1-y))\frac{1}{1-y}. \tag{4}$$

Proof. For each $\omega \in \Omega$, let us define the following transformation $y = r(x) = 1 - \mathrm{e}^{-x}$, $x > 0$, whose inverse mapping $s :]0, 1[\to]0, +\infty[$ and Jacobian J are given by

$$x = s(y) = -\ln(1-y), \qquad J = 1/(1-y),$$

respectively. Then applying Theorem 1, one obtains the PDF of $y(\omega)$ given in (4). □

Observe that the solution stochastic process of the randomized Markovian model (1)–(2) is

$$\begin{pmatrix} s_n(\omega) \\ i_n(\omega) \\ r_n(\omega) \end{pmatrix} = (T(\omega))^n \begin{pmatrix} s_0(\omega) \\ i_0(\omega) \\ 1 - s_0(\omega) - i_0(\omega) \end{pmatrix}, \quad T(\omega) = \begin{pmatrix} \mathrm{e}^{-\alpha(\omega)} & 0 & 0 \\ 1 - \mathrm{e}^{-\alpha(\omega)} & \mathrm{e}^{-\beta(\omega)} & 0 \\ 0 & 1 - \mathrm{e}^{-\beta(\omega)} & 1 \end{pmatrix}. \tag{5}$$

For the sake of simplicity, we will denote the probabilities in matrix $T(\omega)$ as $k(\omega) := 1 - \mathrm{e}^{-\alpha(\omega)}$ and $p(\omega) := 1 - \mathrm{e}^{-\beta(\omega)}$. Then carrying out computations, the following explicit expressions for each subpopulation are straightforwardly obtained,

$$\begin{aligned} s_n(\omega) &= (1 - k(\omega))^n s_0(\omega), \\ i_n(\omega) &= (1 - p(\omega))^n i_0(\omega) + \frac{((1 - k(\omega))^n - (1 - p(\omega))^n)\, k(\omega) s_0(\omega)}{p(\omega) - k(\omega)}, \\ r_n(\omega) &= 1 - (1 - p(\omega))^n i_0(\omega) + \frac{((1 - p(\omega))^n k(\omega) - (1 - k(\omega))^n p(\omega))\, s_0(\omega)}{p(\omega) - k(\omega)}. \end{aligned} \tag{6}$$

1-PDF, $f_s(s,n)$, of susceptibles subpopulation

Let n be a fixed cycle, and, for each $\omega \in \Omega$, define the deterministic mapping $\mathbf{r} : \mathbb{R}^2 \to \mathbb{R}^2$

$$\begin{aligned} x_1 &= r_1(s_0,k) = (1-k)^n s_0, \\ x_2 &= r_2(s_0,k) = k. \end{aligned}$$

The inverse mapping $\mathbf{s} : \mathbb{R}^2 \to \mathbb{R}^2$ of r is

$$\begin{aligned} s_0 &= s_1(x_1,x_2) = x_1(1-x_2)^{-n}, \\ k &= s_2(x_1,x_2) = x_2, \end{aligned}$$

being its Jacobian $J = (1-x_2)^{-n}$. Then applying Theorem 1, the PDF of the absolutely continuous random vector $\mathbf{x}(\omega) = (x_1(\omega), x_2(\omega))$ is

$$f_{\mathbf{x}}(x_1,x_2) = f_{s_0,k}(x_1(1-x_2)^{-n}, x_2)(1-x_2)^{-n}.$$

Notice that $1 - x_2(\omega) > 0$, since $x_2(\omega) = k(\omega) = 1 - \mathrm{e}^{-\alpha(\omega)}$, being $\alpha(\omega) > 0$, for each $\omega \in \Omega$, so we do not need to add the absolute value in the Jacobian term as indicated in the general formula (3).

As $s_0(\omega)$ and $\alpha(\omega)$ are assumed to be independent random variables, $s_0(\omega)$ and $k(\omega)$ are independent too[21, Th. 3, p.9]. Then, $f_{s_0,k}(s_0,k) = f_{s_0}(s_0)f_k(k)$. Marginalizing with respect to $x_2(\omega) = k(\omega)$ and taking n arbitrary, we obtain the 1-PDF of the susceptible subpopulation, $x_1(\omega) = s_n(\omega)$

$$f_s(s,n) = \int\limits_{\mathcal{D}(k(\omega))} f_{s_0}(s(1-k)^{-n}) f_k(k)(1-k)^{-n}\,\mathrm{d}k, \tag{7}$$

where, as it was indicated previously, $\mathcal{D}(k(\omega))$ denotes the domain of random variable $k(\omega)$. Now, we express this representation of $f_s(s,n)$ in terms of the data. In this regard, recall that we are assuming that the PDF, $f_\alpha(\alpha)$, of the absolutely continuous input random variable $\alpha(\omega)$ is given and that $k(\omega) = 1 - \mathrm{e}^{-\alpha(\omega)}$. So, applying Lemma 1, we can express the PDF $f_k(k)$, appearing in (7), in terms of $f_\alpha(\alpha)$. This yields

$$f_s(s,n) = \int\limits_{\mathcal{D}(1-\mathrm{e}^{-\alpha(\omega)})} f_{s_0}(s(1-k)^{-n}) f_\alpha(-\ln(1-k))(1-k)^{-(n+1)}\,\mathrm{d}k. \tag{8}$$

1-PDF, $f_i(i,n)$, of infected subpopulation

Following a similar reasoning to the one exhibited previously for susceptibles, but using another appropriate mapping $\mathbf{r}$, it can be shown, by applying the Theorem 1, that the 1-PDF, $f_i(i,n)$, of the infected subpopulation is given by

$$\begin{aligned} f_i(i,n) = &\int\limits_{\mathcal{D}(1-\mathrm{e}^{-\beta(\omega)})} \int\limits_{\mathcal{D}(1-\mathrm{e}^{-\alpha(\omega)})} \int\limits_{\mathcal{D}(i_0(\omega))} f_{s_0}\left(\frac{(i-i_0(1-p)^n)(-k+p)}{((1-k)^n-(1-p)^n)\,k}\right) f_{i_0}(i_0) f_\alpha(-\ln(1-k)) \\ &\times f_\beta(-\ln(1-p)) \left|\frac{(1-k)^{-1}(1-p)^{-1}(-k+p)}{((1-k)^n-(1-p)^n)\,k}\right| \mathrm{d}i_0\,\mathrm{d}k\,\mathrm{d}p. \end{aligned} \tag{9}$$

For the sake of clarity in the presentation, the technical details are reported in the Appendix.

1-PDF, $f_r(r,n)$, of recovered subpopulation

The explicit expression to the 1-PDF, $f_r(r,n)$, of the recovered subpopulation is given by

$$\begin{aligned} f_r(r,n) = &\int\limits_{\mathcal{D}(1-\mathrm{e}^{\beta(\omega)})} \int\limits_{\mathcal{D}(1-\mathrm{e}^{\alpha(\omega)})} \int\limits_{\mathcal{D}(i_0(\omega))} f_{s_0}\left(\frac{(-1+r+i_0(1-p)^n)(-k+p)}{(1-p)^n k-(1-k)^n p}\right) f_{i_0}(i_0) \\ &\times f_\alpha(-\ln(1-k)) f_\beta(-\ln(1-p)) \left|\frac{(1-k)^{-1}(1-p)^{-1}(-k+p)}{(1-p)^n k-(1-k)^n p}\right| \mathrm{d}i_0\,\mathrm{d}k\,\mathrm{d}p. \end{aligned} \tag{10}$$

In Appendix, we show the mathematical development to obtain this expression by applying Theorem 1 and choosing an appropriate mapping $\mathbf{r}$.

3 | PDF OF THE TIME UNTIL A GIVEN PROPORTION OF THE POPULATION REMAINS IN EACH STATE

From an applied point of view it is very helpful to know when the percentage of susceptibles, infected or recovered in the population will attain a specific level. Since these times will obviously depend on random initial subpopulations, $s_0(\omega)$ and $i_0(\omega)$, and on random parameters, $\alpha(\omega)$ and $\beta(\omega)$, the aforementioned times will be random themselves. In this section, we determine the PDFs of the times until a given proportion of the population remain susceptibles, infected and recovered. Bearing in mind this goal, let us fix a proportion of population remaining susceptible, infected or recovered, ρ_s, ρ_i and ρ_r, respectively. In the following we will obtain the PDF of these times in each one of the three states by applying the RVT method stated in Theorem 1.

PDF, $f_{n_s}(n)$, of random time for susceptibles

Let us denote by $n_s(\omega)$ the time until a given proportion of the population, $\rho_s \in (0,1)$, remains susceptible. Then, isolating $n_s(\omega)$ from the solution of susceptible subpopulation, given in (6), and taking into account that $k(\omega) = 1 - \mathrm{e}^{-\alpha(\omega)}$, we obtain

$$n_s(\omega) = \frac{1}{\alpha(\omega)} \ln\left(\frac{s_0(\omega)}{\rho_s}\right).$$

Now, we apply Theorem 1 to the mapping $\mathbf{r} : \mathbb{R}^2 \to \mathbb{R}^2$, whose components are defined by

$$\begin{aligned} x_1 &= r_1(s_0, \alpha) = \frac{1}{\alpha} \ln\left(\frac{s_0}{\rho_s}\right), \\ x_2 &= r_2(s_0, \alpha) = \alpha. \end{aligned}$$

The inverse mapping $\mathbf{s} : \mathbb{R}^2 \to \mathbb{R}^2$ of $\mathbf{r}$ is

$$\begin{aligned} s_0 &= s_1(x_1, x_2) = \mathrm{e}^{x_1 x_2}\, \rho_s, \\ \alpha &= s_2(x_1, x_2) = x_2, \end{aligned}$$

being the absolute value of the Jacobian of this transformation: $|J| = \rho_s x_2\, \mathrm{e}^{x_1 x_2}$. Then, applying Theorem 1, and considering $\rho_s \in (0,1)$, the joint PDF of the absolutely continuous random vector $\mathbf{x}(\omega) = (x_1(\omega), x_2(\omega))$ is given by

$$f_{\mathbf{x}}(x_1, x_2) = \rho_s x_2 f_{s_0,\alpha}(\rho_s\, \mathrm{e}^{x_1 x_2}, x_2)\, \mathrm{e}^{x_1 x_2}.$$

Therefore, the PDF of the absolutely continuous random variable $n_s(\omega)$ is the marginal PDF of $f_{\mathbf{x}}(x_1, x_2)$ with respect to $x_2(\omega) = \alpha(\omega)$ and, as $s_0(\omega)$ and $\alpha(\omega)$ are assumed to be independent random variables, one obtains

$$f_{n_s}(n; \rho_s) = \rho_s \int_{D(\alpha(\omega))} \alpha f_{s_0}\left(\rho_s\, \mathrm{e}^{n\alpha}\right) f_\alpha(\alpha)\, \mathrm{e}^{n\alpha}\, \mathrm{d}\alpha. \tag{11}$$

PDF, $f_{n_i}(n)$, of random time for infected

To compute the PDF of the time $n_i(\omega)$ until a given proportion of the population, $\rho_i \in (0,1)$, remains infected, first observe that the first equation in (6) can be expressed as $\rho_s = (1 - k(\omega))^n s_0(\omega)$ and, taking into account that $k(\omega) = 1 - \mathrm{e}^{-\alpha(\omega)}$ and $p(\omega) = 1 - \mathrm{e}^{-\beta(\omega)}$, then the second equation in (6) writes

$$\rho_i = \mathrm{e}^{-n\beta(\omega)}\, i_0(\omega) + \frac{\left(\rho_s - \mathrm{e}^{-\beta(\omega)n}\, s_0(\omega)\right)\left(1 - \mathrm{e}^{-\alpha(\omega)}\right)}{\mathrm{e}^{-\alpha(\omega)} - \mathrm{e}^{-\beta(\omega)}}.$$

Now, isolating $n_i(\omega)$ one obtains

$$n_i(\omega) = \frac{1}{\beta(\omega)} \ln\left(\frac{\mathrm{e}^{\alpha(\omega)}\, i_0(\omega) + \mathrm{e}^{\alpha(\omega)+\beta(\omega)}\, s_0(\omega) - \mathrm{e}^{\beta(\omega)}(i_0(\omega) + s_0(\omega))}{\mathrm{e}^{\alpha(\omega)}\, \rho_i + \mathrm{e}^{\alpha(\omega)+\beta(\omega)}\, \rho_s - \mathrm{e}^{\beta(\omega)}(\rho_i + \rho_s)}\right).$$

We apply Theorem 1 (see details in Appendix) to obtain the following expression to the PDF, $f_{n_i}(n)$, of the time until a given proportion of the population remains infected, $n_i(\omega)$,

$$f_{n_i}(n;\rho_s,\rho_i) = \int\limits_{D(i_0(\omega))}\int\limits_{D(\alpha(\omega))}\int\limits_{D(\beta(\omega))} f_{s_0}\left(\frac{\mathrm{e}^{-\beta}}{-1+\mathrm{e}^{\alpha}}\left(\mathrm{e}^{\beta n}\left(-\mathrm{e}^{\beta}\left(\rho_i+\rho_s\right)+\mathrm{e}^{\alpha}\left(\rho_i+\mathrm{e}^{\beta}\,\rho_s\right)\right)+\left(-\mathrm{e}^{\alpha}+\mathrm{e}^{\beta}\right)i_0\right)\right) \times f_{i_0}(i_0)f_{\alpha}(\alpha)f_{\beta}(\beta)\left|\frac{\mathrm{e}^{(-1+n)\beta}\,\beta}{-1+\mathrm{e}^{\alpha}}\left(-\mathrm{e}^{\beta}\left(\rho_i+\rho_s\right)+\mathrm{e}^{\alpha}\left(\rho_i+\mathrm{e}^{\beta}\,\rho_s\right)\right)\right|\mathrm{d}\beta\,\mathrm{d}\alpha\,\mathrm{d}i_0. \tag{12}$$

PDF, $f_{n_r}(n)$, of random time for recovered

As before, we obtain $n_r(\omega)$ in terms of the fixed proportions ρ_s and ρ_r from the third equation of expression (6) as

$$\rho_r = 1-\mathrm{e}^{-\beta n}\,i_0(\omega)+\frac{\mathrm{e}^{-\beta n}(1-\mathrm{e}^{-\alpha})s_0-\rho_s(1-\mathrm{e}^{-\beta})}{\mathrm{e}^{-\alpha}-\mathrm{e}^{-\beta}}.$$

Then, isolating $n_r(\omega)$, one gets

$$n_r(\omega)=\frac{1}{\beta(\omega)}\ln\left(\frac{\mathrm{e}^{\alpha}\,i_0+\mathrm{e}^{\alpha+\beta}\,s_0-\mathrm{e}^{\beta}(i_0+s_0)}{\mathrm{e}^{\alpha}(1-\rho_r-\rho_s)-\mathrm{e}^{\beta}(1-\rho_r-\mathrm{e}^{\alpha}\,\rho_s)}\right).$$

Now, the PDF , $f_{n_r}(n)$, of the time until a given proportion of the population remains recovered, $n_r(\omega)$, can be calculated by applying Theorem 1 (see details in the Appendix)

$$f_{n_r}(n;\rho_s,\rho_r) = \int\limits_{D(i_0(\omega))}\int\limits_{D(\alpha(\omega))}\int\limits_{D(\beta(\omega))} f_{s_0}\left(\frac{\mathrm{e}^{-\beta}}{-1+\mathrm{e}^{\alpha}}\left(\mathrm{e}^{\beta n}\left(\mathrm{e}^{\beta}\left(-1+\rho_r\right)+\mathrm{e}^{\alpha+\beta}\,\rho_s+\mathrm{e}^{\alpha}\left(1-\rho_r-\rho_s\right)\right)+\left(\mathrm{e}^{\beta}-\mathrm{e}^{\alpha}\right)i_0\right)\right) \times f_{i_0}(i_0)f_{\alpha}(\alpha)f_{\beta}(\beta)\left|\frac{\mathrm{e}^{(-1+n)\beta}\,\beta}{-1+\mathrm{e}^{\alpha}}\left(\mathrm{e}^{\beta}\left(-1+\rho_r\right)+\mathrm{e}^{\alpha+\beta}\,\rho_s+\mathrm{e}^{\alpha}\left(1-\rho_r-\rho_s\right)\right)\right|\mathrm{d}\beta\,\mathrm{d}\alpha\,\mathrm{d}i_0. \tag{13}$$

4 | PDF OF SOME KEY RANDOMIZED QUANTITIES IN EPIDEMIC THEORY

This section is addressed to determine the expression of the PDF of some key randomized quantities in Epidemiology: the basic reproduction number, R_0, the effective reproduction number, R_{e} and the herd immunity threshold, HIT. These quantities, and other reproduction quantities related to them, are useful in designing control strategies[22].

Basic reproduction number, R_0: This quantity is used to measure the transmission of a disease. From R_0 we can determine whether the disease dies out or it spreads out. Then, we have information about if we are dealing with a pandemic. A pandemic occurs if the number or infected individuals increases more than the recovered. The epidemiological definition of R_0 is the average number of secondary cases produced by one infected individual introduced into a population of susceptible individuals[23]. Therefore, taking into account Figure 1 , R_0 can defined as

$$R_0=\frac{\alpha}{\beta},$$

being α the rate of infection and β the rate of recovery. Thus, the disease spreads out when $R_0>1$ ($\alpha>\beta$) and it will die out if $R_0<1$ ($\alpha<\beta$) when the time passes. Observe that these conditions are very intuitive. As we have motivated in the introduction section, in this paper we assume that parameters α and β are absolutely continuous random variables. Then, $R_0(\omega)=\frac{\alpha(\omega)}{\beta(\omega)}$ is also a random variable. To discuss, from a probabilistic standpoint, how the epidemic will evolve in the long-run, we will compute the PDF of R_0 taking advantage of the the RVT method stated in Theorem 1. To this end, we define the deterministic mapping $\mathbf{r}:\mathbb{R}^2\to\mathbb{R}^2$

$$\begin{aligned} x_1 &= r_1(\alpha,\beta)=\alpha/\beta,\\ x_2 &= r_2(\alpha,\beta)=\beta. \end{aligned}$$

The inverse mapping $\mathbf{s}:\mathbb{R}^2\to\mathbb{R}^2$ of $\mathbf{r}$ is

$$\begin{aligned} \alpha &= s_1(x_1,x_2)=x_1x_2,\\ \beta &= s_2(x_1,x_2)=x_2. \end{aligned}$$

The absolute value of its Jacobian is $|J| = x_2$ since $\beta > 0$. Then applying Theorem 1, the PDF of the absolutely continuous random vector $\mathbf{x}(\omega) = (x_1(\omega), x_2(\omega))$ is

$$f_{\mathbf{x}}(x_1, x_2) = f_{\alpha,\beta}\left(x_1 x_2, x_2\right) x_2.$$

Then, marginalizing with respect to the random variable $x_2(\omega) = \beta(\omega)$ and assuming that $\alpha(\omega)$ and $\beta(\omega)$ are independent random variables, the PDF of $R_0(\omega)$ is obtained

$$f_{R_0}(r) = \int\limits_{D(\beta(\omega))} f_\alpha(r\beta) f_\beta(\beta)\beta \, \mathrm{d}\beta. \tag{14}$$

As it was pointed before, in the deterministic theory if $R_0 > 1$ we are dealing with an epidemic (the disease spreads out). In the probabilistic scenario, we can calculate the probability that this event occurs

$$\mathbb{P}_0 = \mathbb{P}\left[\{\omega \in \Omega \ : \ R_0(\omega) > 1\}\right] = \int\limits_1^\infty f_{R_0}(r)\mathrm{d}r, \tag{15}$$

where $f_{R_0}(r)$ is given by (14).

Effective reproduction number, R_e: It is rare to find a disease where the total population is susceptible. There exists always a percentage of the population which is immune to the disease. For example, immunity may be the result of a previous vaccine. The effective reproduction number (also termed replacement number) considers this circumstance. It is the average number of secondary cases per infectious case in a population made up of both susceptible and non-susceptible hosts, and it is defined by $R_\mathrm{e} = R_0 x$, being R_0 the basic reproduction number and x the fraction of the host population that is susceptible. In this case, if $R_\mathrm{e} > 1$ the number of cases will increase, such as at the start of an epidemic. If $R_\mathrm{e} = 1$ the disease is endemic, that is, the disease usually affects a particular region. The number of cases decreases when $R_\mathrm{e} < 1$, and then the disease tends to disappear.

In this paper we consider that the percentage of susceptible population is not necessarily one. Then, there is a percentage of the initial population that is infected or recovered. This fact allows us to contemplate the possibility of immunity. Immune persons fit perfectly in the initial recovered subpopulation. Thus, in this case x is s_0, being

$$R_\mathrm{e} = R_0 s_0 = \frac{\alpha s_0}{\beta}. \tag{16}$$

As in the case of the basic reproduction number, we consider that parameters s_0, α and β are absolutely continuous random variables. Then, $R_\mathrm{e}(\omega) = \frac{\alpha(\omega)s_0(\omega)}{\beta(\omega)}$ and we apply the RVT to determine the following expression for its PDF

$$f_{R_\mathrm{e}}(r) = \int\limits_{D(\beta(\omega))} \int\limits_{D(\alpha(\omega))} f_{s_0}\left(\frac{r\beta}{\alpha}\right) f_\alpha(\alpha) f_\beta(\beta) \frac{\beta}{\alpha} \, \mathrm{d}\alpha \, \mathrm{d}\beta.$$

This allows us to compute the following probability

$$\mathbb{P}_\mathrm{e} = \mathbb{P}\left[\{\omega \in \Omega \ : \ R_\mathrm{e}(\omega) > 1\}\right] = \int\limits_1^\infty f_{R_e}(r)\mathrm{d}r, \tag{17}$$

that provides key probabilistic information about the likelihood that the number of infected cases increases.

Herd immunity threshold, HIT: This is an important measure used in infectious disease control and immunisation and eradication programmes. Herd immunity happens if a significant percentage of the population is immune, for example they have been vaccinated. The HIT is defined from the basic reproduction number

$$\mathrm{HIT} = 1 - \frac{1}{R_0} = 1 - \frac{\beta}{\alpha}.$$

It represents the percentage of the population that needs to be immune in order for an infectious disease to become stable in case of no stability, i.e., when $\alpha > \beta$ (so assuring that, indeed $\mathrm{HIT} \in (0, 1)$). Now, we consider that α and β are absolutely continuous random variables, then $\mathrm{HIT}(\omega) = 1 - \frac{\beta(\omega)}{\alpha(\omega)}$ is an absolutely continuous random variable and by applying the RVT technique it can be seen that its PDF is given by

$$f_{\mathrm{HIT}}(h) = \int\limits_{D(\alpha(\omega))} f_\beta(\alpha(1-h)) f_\alpha(\alpha) \, \alpha \, \mathrm{d}\alpha.$$

5 | APPLYING THE RANDOMIZED DISCRETE-TIME MARKOV CHAIN SIR MODEL TO REAL DATA

In this section we apply the theoretical results previously established to illustrate how the randomized Markov chain SIR model given by (1)–(2) can be applied to model the pandemic influenza in Egypt. We take advantage of the real data collected in Figure 2 , which have been excerpted from reference[24]. In Figure 2 , we have plotted the percentages of Susceptibles (s_n), Infected (i_n) and Recovered (r_n) from pandemic influenza in Egypt corresponding to $n = 0, 2, 4, 6, \ldots, 44$ time instants (days) during the year 2006,[24].

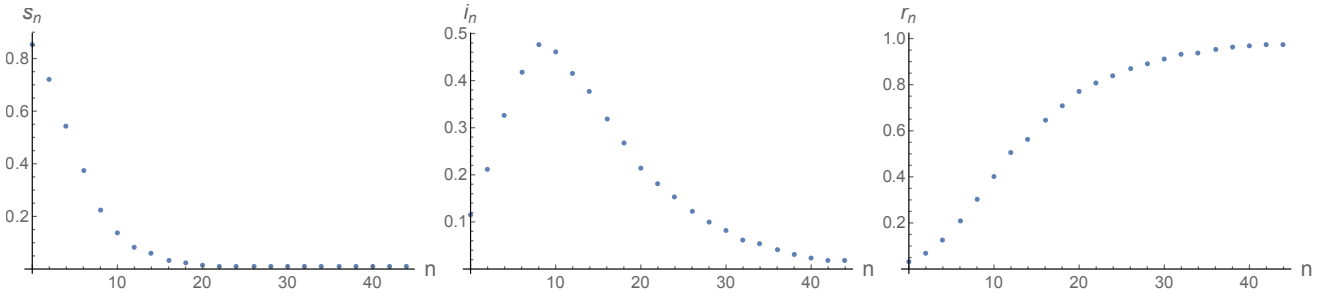


FIGURE 2 Percentage of Susceptibles (Left), Infected (Center) and Recovered (Right) in 23 days during the year 2006 for pandemic influenza in Egypt. Source:[24].

First, we must choose appropriate PDFs for the absolutely continuous random inputs $s_0(\omega)$, $i_0(\omega)$, $\alpha(\omega)$ and $\beta(\omega)$. As it has been said in our previous theoretical development, these random variables are assumed to be independent random variables. Notice that the random variable $s_0(\omega)$ represents a proportion, thus its values lies between 0 and 1. Therefore, we will assume that $s_0(\omega)$ has a Uniform distribution on the interval $(s_{0,1}, s_{0,2})$, where $0 < s_{0,1} < s_{0,2} < 1$ will be determined later. Based on similar arguments, the initial condition i_0 is assumed to have Uniform distribution on an interval $(i_{0,1}, i_{0,2})$ with $0 < i_{0,1} < i_{0,2} < 1$. On the other hand, $\alpha(\omega)$ represents the infection rate, which by definition, is the number of infections over the number of those in a population that are at risk of infection (susceptible). Thus, this value lies also in the unit interval $(0, 1)$. In this case, we choose a Beta distribution with positive parameters $a_1, b_1 > 0$, i.e., $\alpha(\omega) \sim \text{Be}(a_1; b_1)$. Since the Beta distribution depends on two parameters, with this choice we allow for greater flexibility in fitting the model. Based upon the same arguments, since $\beta(\omega)$ also represents a rate, then, we choose $\beta(\omega) \sim \text{Be}(a_2; b_2)$, being $a_2, b_2 > 0$. In order to determine the parameters $s_{0,1}$, $s_{0,2}$, $i_{0,1}$, $i_{0,2}$, a_1, b_1, a_2 and b_2, that best fit the randomized Markov chain SIR model (1)–(2) to data, we will minimize the mean square error, represented in (18) by the error function $e(s_{0,1}, s_{0,2}, i_{0,1}, i_{0,2}, a_1, b_1, a_2, b_2)$ (see (18)), between data given in Figure 2 and the expectation of each one of the states: susceptible, $s_n(\omega) := s_n(\omega; s_{0,1}, s_{0,2}, a_1, b_1)$; infected $i_n(\omega) := i_n(\omega; s_{0,1}, s_{0,2}, i_{0,1}, i_{0,2}, a_1, b_1, a_2, b_2)$, and recovered $r_n(\omega) := r_n(\omega; s_{0,1}, s_{0,2}, i_{0,1}, i_{0,2}, a_1, b_1, a_2, b_2)$, evaluated at the time instants (days) $n \in \mathcal{T} = \{0, 2, 4, 6, \ldots, 44\}$. Summarizing, this corresponds to the following optimization programme

$$\min_{\substack{0 < s_{0,1}, s_{0,2} < 1 \\ 0 < i_{0,1}, i_{0,2} < 1 \\ a_1, b_1, a_2, b_2 > 0}} e(s_{0,1}, s_{0,2}, i_{0,1}, i_{0,2}, a_1, b_1, a_2, b_2) = \sum_{n \in \mathcal{T}} \left(s_n - \mathbb{E}\left[s_n(\omega)\right]\right)^2 + \left(i_n - \mathbb{E}\left[i_n(\omega)\right]\right)^2 + \left(r_n - \mathbb{E}\left[r_n(\omega)\right]\right)^2. \tag{18}$$

Observe that the expectation for each subpopulation is calculated as follows

$$\begin{aligned}
\mathbb{E}\left[s_n(\omega)\right] &= \mathbb{E}\left[s_n(\omega; s_{0,1}, s_{0,2}, a_1, b_1)\right] = \int_0^1 s f_s(s, n)\, \mathrm{d}s, \quad n \in \mathcal{T}, \\
\mathbb{E}\left[i_n(\omega)\right] &= \mathbb{E}\left[i_n(\omega; s_{0,1}, s_{0,2}, i_{0,1}, i_{0,2}, a_1, b_1, a_2, b_2)\right] = \int_0^1 i f_i(i, n)\, \mathrm{d}i, \quad n \in \mathcal{T}, \\
\mathbb{E}\left[r_n(\omega)\right] &= \mathbb{E}\left[r_n(\omega; s_{0,1}, s_{0,2}, i_{0,1}, i_{0,2}, a_1, b_1, a_2, b_2)\right] = \int_0^1 r f_r(r, n)\, \mathrm{d}r, \quad n \in \mathcal{T},
\end{aligned} \tag{19}$$

where $f_s(s, n)$, $f_i(i, n)$ and $f_r(r, n)$ are given by (8), (9) and (10), respectively.

To optimize programme (18), we have applied the Nelder-Mead algorithm carrying out computations by Mathematica® software. We obtain the following values

$$\begin{array}{rlrl} s_{0,1} &= 0.908819, & s_{0,2} &= 0.918742, \\ i_{0,1} &= 0.046341, & i_{0,2} &= 0.054682, \\ a_1 &= 206.273096, & b_1 &= 1052.315715, \\ a_2 &= 31.252079, & b_2 &= 242.792953. \end{array} \tag{20}$$

Firstly, with these values for the parametric distributions, we check out that the integral of the 1-PDF for each subpopulation, given in expressions (8), (9) and (10), is one at every time instant n, i.e.,

$$\int_0^1 f_s(s,n)\,\mathrm{d}s = 1, \quad \int_0^1 f_i(i,n)\,\mathrm{d}i = 1 \quad \int_0^1 f_r(r,n)\,\mathrm{d}r = 1, \quad n \in \mathcal{T}.$$

Apart from the expectations of each subpopulation, given by (19), the calculation of the 1-PDFs $f_s(s,n)$, $f_i(i,n)$ and $f_r(r,n)$ allows us to compute their respective variances. For example, for the subpopulation of susceptibles, we use

$$\mathbb{V}\left[s_n(\omega)\right] = \int_0^1 (s - \mathbb{E}\left[s_n(\omega)\right])^2 f_s(s,n)\,\mathrm{d}s, \quad n \in \mathcal{T}. \tag{21}$$

Besides the punctual predictions for each subpopulations obtained by means of the above expectations, we can construct probabilistic predictions via confidence intervals for any $(1-\gamma)\times 100\%$ confidence level. For example for the susceptible subpopulation, fixed $\gamma \in (0,1)$, for each $\hat{n} \in \mathcal{T}$ fixed, we can calculate $s^1_{\hat{n}}$ and $s^2_{\hat{n}}$ such that

$$\int_0^{s^1_{\hat{n}}} f_s(s,\hat{n})\,\mathrm{d}s = \frac{\gamma}{2} = \int_{s^2_{\hat{n}}}^1 f_s(s,\hat{n})\,\mathrm{d}s\,, \tag{22}$$

where

$$1-\gamma = \mathbb{P}\left(\left\{\omega \in \Omega : s_{\hat{n}}(\omega) \in \left[s^1_{\hat{n}}, s^2_{\hat{n}}\right]\right\}\right) = \int_{s^1_{\hat{n}}}^{s^2_{\hat{n}}} f_s(s,\hat{n})\,\mathrm{d}s\,. \tag{23}$$

Usually $\gamma = 0.05$ is taken so that 95% confidence intervals are built.

Notice that analogous expressions to (21)–(23) can be given for the percentage of infected and recovered subpopulations by changing $f_s(s,n)$ by $f_i(i,n)$, and $f_r(r,n)$, respectively.

Furthermore, it is important to point out that the knowledge of the 1-PDF of each subpopulation permits the computation of key information as the probability that, for $\hat{n}$ arbitrary but fixed, the percentage of susceptibles (infected/recovered) lies within a specific interval of interest, say, $[s^1_{\hat{n}}, s^2_{\hat{n}}] = [\hat{s}^1, \hat{s}^2]$

$$\mathbb{P}\left(\left\{\omega \in \Omega : s_{\hat{n}}(\omega) \in \left[\hat{s}^1, \hat{s}^2\right]\right\}\right) = \int_{\hat{s}^1}^{\hat{s}^2} f_s(s,\hat{n})\,\mathrm{d}s\,. \tag{24}$$

In Figure 3 the 1-PDFs of susceptible, infected and recovered subpopulations, given in formulas (8)–(10), are represented for the optimal values given in (20). Although we have data at the time instants (days) $n \in \{0,2,4,\ldots,44\}$, these probabilistic functions have been plotted only for the time periods $n \in \{0,2,4,\ldots,32\}$, since as we can observe in Figure 3 the 1-PDFs tend to stabilize its behaviour from this last time instant. In addition, $f_s(s;n)$ starts to be leptokurtic at $n = 32$, which makes its graphical representation difficult. From Figure 3 we observe that the number of susceptible vanishes over time while the recovered subpopulation tends to be the total population. This circumstance is in agreement with the fact that R is an absorbent state (see Figure 1). With respect to the infected subpopulation, at first it increases but approximately from the 8th day it decreases to zero.

In Figure 4 we validate the model constructing the mean and confidence intervals. We show the graphical fitting performed by means of the 1-PDFs for each subpopulation. In this plot, we have represented the real data (points), the mean (solid line) and the confidence interval (dashed lines) constructed via the expectation plus/minus 1.96 standard deviations at each day where data is available[21]. We observe that the data lies within the confidence intervals, hence the model is capable of capturing the variability of the real data and the model can be validated at 95% confidence level. Furthermore, in Figure 4 the dynamic

behaviour of three populations can be observed and it is in agreement with the time evolution of the PDFs of each subpopulation described previously in Figure 3 .

Figure 5 shows the PDF, $f_{n_s}(n;\rho_s)$, of the time n_s until a given proportion, ρ_s, of susceptible subpopulation remains susceptible. This graphical representation has been performed for different values of parameter $\rho_s \in \{0.1, 0.2, \dots, 0.9\}$. From this PDF, we can calculate the expectation of random variable n_s given a fixed value ρ_s,

$$\mathbb{E}[n_s] = \int_0^\infty n f_{n_s}(n; 0.4) = 5.06116,$$

where $f_{n_s}(n;\rho_s)$ is given by (11). Then, approximately the fifth day represents the average time until the 40% of the susceptible subpopulation will remain at the same state. It can be also seen in Figure 5 . It is worth noting that this result is in agreement with the graphical representation of the 1-PDF, $f_s(s;n)$, shown in Figure 3 (top). In Table 1 , we have computed the expectation for the following values of $\rho_s \in \{0.1, 0.2, \dots, 0.9\}$. The figures collected in this table admit a similar interpretation as the one previously specified.

ρ_s	0.1	0.2	0.3	0.4	
$\mathbb{E}[n_s]$	13.5542	9.30767	6.82362	5.06116	
ρ_s	0.5	0.6	0.7	0.8	0.9
$\mathbb{E}[n_s]$	3.69409	2.57711	1.63272	0.819997	0.165793

TABLE 1 Expectation of the time, $n_s(\omega)$, until a given proportion, ρ_s, of the susceptible subpopulation remains susceptible for different values of $\rho_s \in \{0.1, 0.2, \dots, 0.9\}$.

We can look at expressions (12) and (13) that each PDF, $f_{n_i}(n;\rho_s,\rho_i)$ and $f_{n_r}(n;\rho_s,\rho_r)$, depends on three parameters, the time n and two proportions, $\{\rho_s,\rho_i\}$ and $\{\rho_s,\rho_r\}$, respectively. Therefore, now we will first plot $f_{n_i}(n;\rho_s,\rho_i)$ for fixed proportions of susceptible, ρ_s and of infected, ρ_i, and secondly, we will plot $f_{n_r}(n;\rho_s,\rho_r)$ fixing the proportions $\{\rho_s,\rho_r\}$. In Figure 6 both functions are plotted for particular values of the aforementioned proportions. Specifically, on the left, we show the PDF of the time $n_i(\omega)$ until a given proportion, $\rho_i = 0.3$, of infected remains infected knowing that the proportion of susceptibles is $\rho_s = 0.15$, while on the right, we show the PDF of the time $n_r(\omega)$, until a given proportion, $\rho_r = 0.8$, of recovered remains recovered knowing that $\rho_s = 0.05$. It must pointed out that these plots have been performed on the time intervals $[9, 20]$ and $[14, 30]$, respectively, where it can be checked that the corresponding integrals have value 1, thus representing the PDFs. For illustrative purposes only, for example in this second case, we can compute the mean or expectation

$$\mathbb{E}[n_r] = \int_0^\infty n f_{n_r}(n; 0.05, 0.8) = 20.6126.$$

This means that approximately the 21st day represents the average time until the 80% of the recovered subpopulation remains recovered when 5% of the subpopulation remains susceptible.

Finally, we represent the distribution of the epidemiological randomized quantities described in Section 4. In Figure 7 , we show the PDFs, f_{R_0}, f_{R_e} and f_{HIT}, of the basic reproduction number (left), the effective reproduction number (center) and the herd immunity threshold (right), respectively. Since random variable $s_0(\omega)$ has a PDF concentrated about the value $0.91 \approx 1$ (recall that $s_0(\omega) \sim \text{U}(0.908819, 0.9187742)$), then according to (16) the PDFs f_{R_0} and f_{R_e} are very similar (see left and center panels in Figure 7). We can calculate the probabilities $\mathbb{P}_0$ and $\mathbb{P}_e$, given in expressions (15) and (17), for each key random variable $R_0(\omega)$ and $R_e(\omega)$, respectively. We obtain $\mathbb{P}_0 = 0.984957$ and $\mathbb{P}_e = 0.946996$, respectively. Then, there exists a high likelihood that the epidemic evolves. With regard to plot of PDF f_{HIT} (see Figure 7 , right panel), it is worth nothing that a piece of its domain is negative because there exists a (small) likelihood that the event $\{\omega \in \Omega : \alpha(\omega) < \beta(\omega)\}$ happens, then $\text{HIT}(\omega) = 1 - \frac{\beta(\omega)}{\alpha(\omega)}$ could take negative values with a small probability.

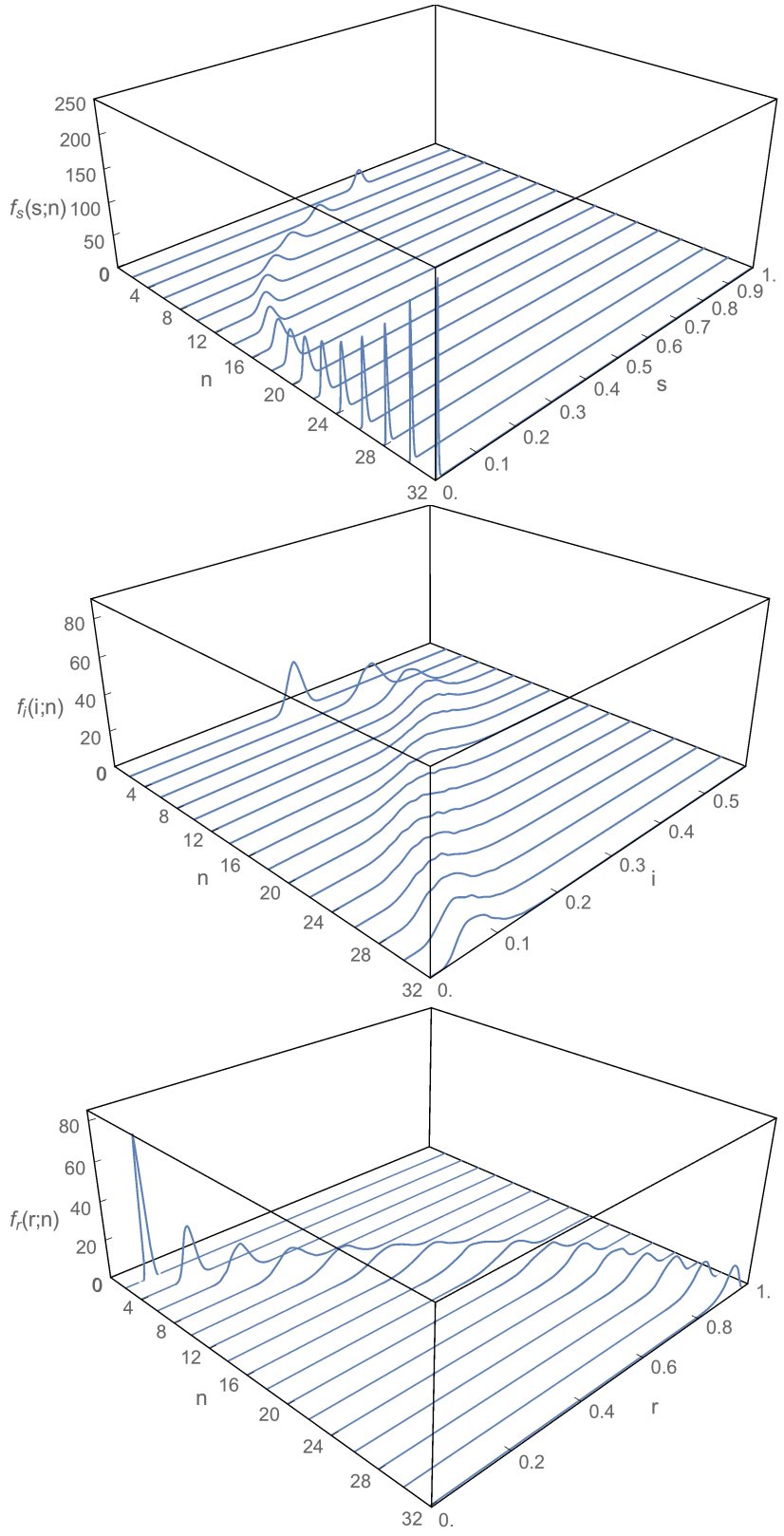


FIGURE 3 1-PDFs of susceptible (top), infected (center) and recovered (bottom) subpopulations for the periods $n \in \{0, 2, \ldots, 32\}$ describing the evolution of pandemic influenza in Egypt. Notice that these graphical results are in full agreement with Figure 2 .

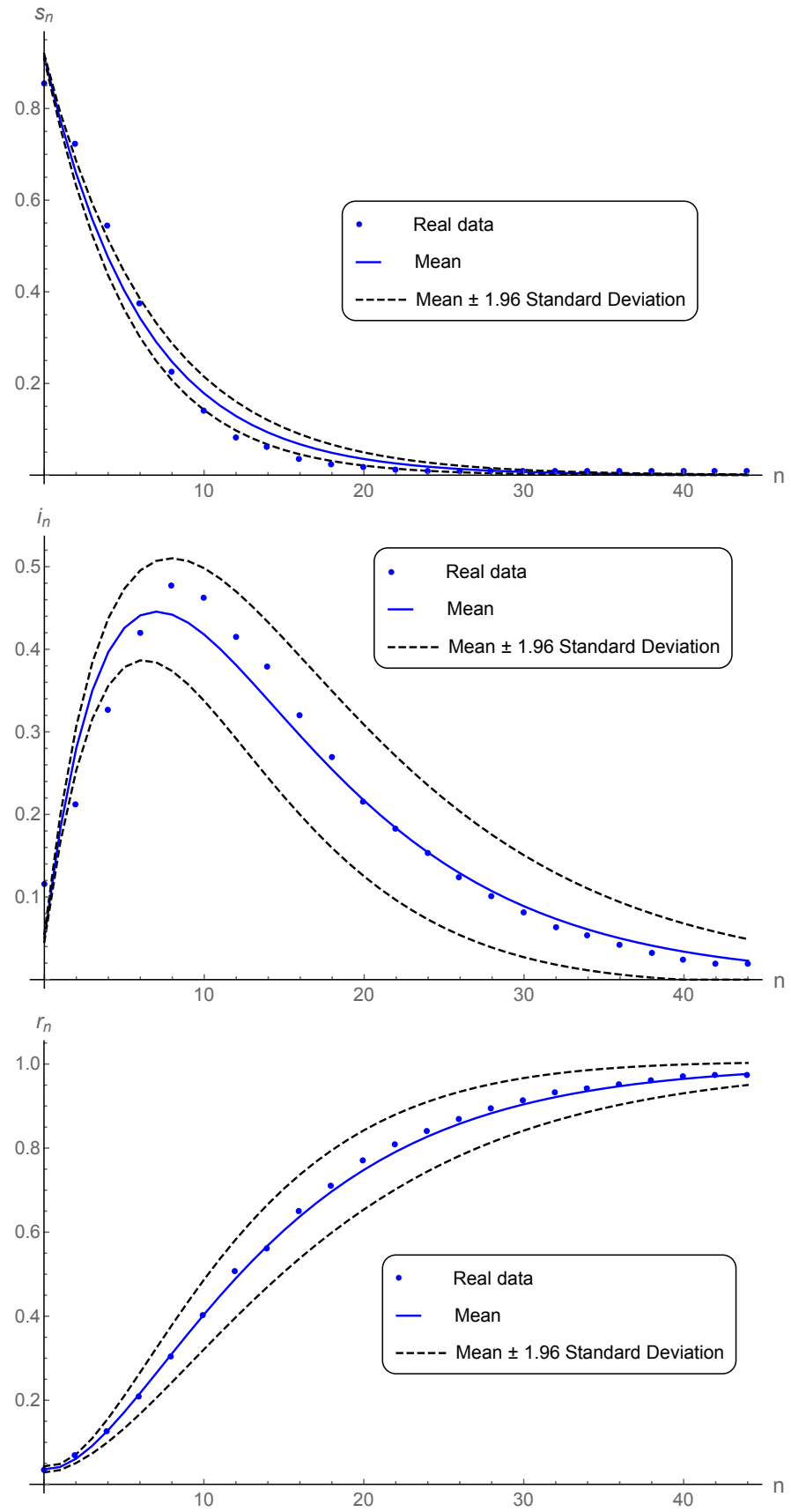


FIGURE 4 Expectation (solid line) and 95% confidence intervals (dashed lines) in the time interval $n \in [0, 40]$ for susceptible (top), infected (center) and recovered (bottom) subpopulations. These plots describe the probabilistic evolution of pandemic influenza in Egypt. Notice that these graphical results are in full agreement with Figure 2 .

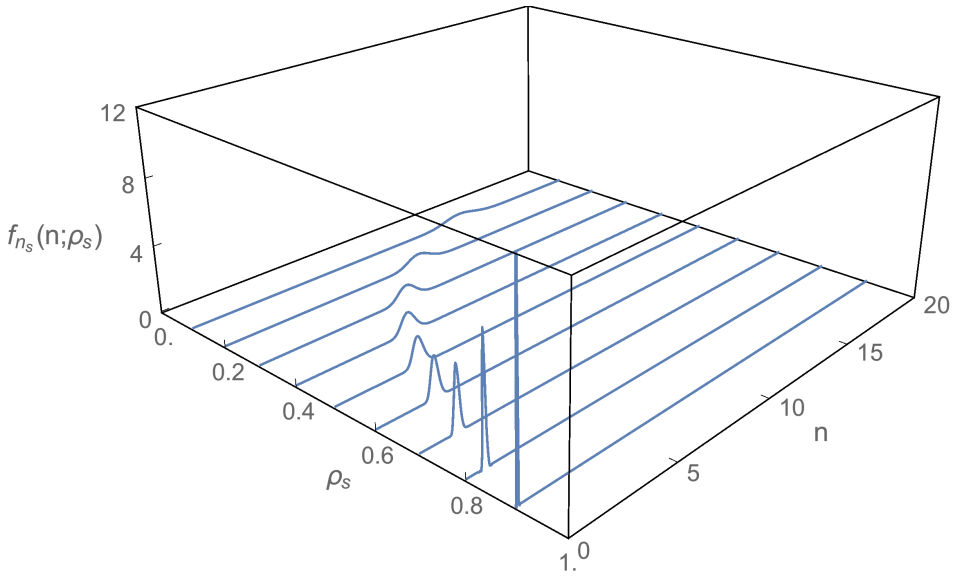


FIGURE 5 PDFs of time, $n_s(\omega)$, until a given proportion, ρ_s, of the susceptible subpopulation remains susceptible for the following values of $\rho_s \in \{0.1, 0.2, \ldots, 0.9\}$.

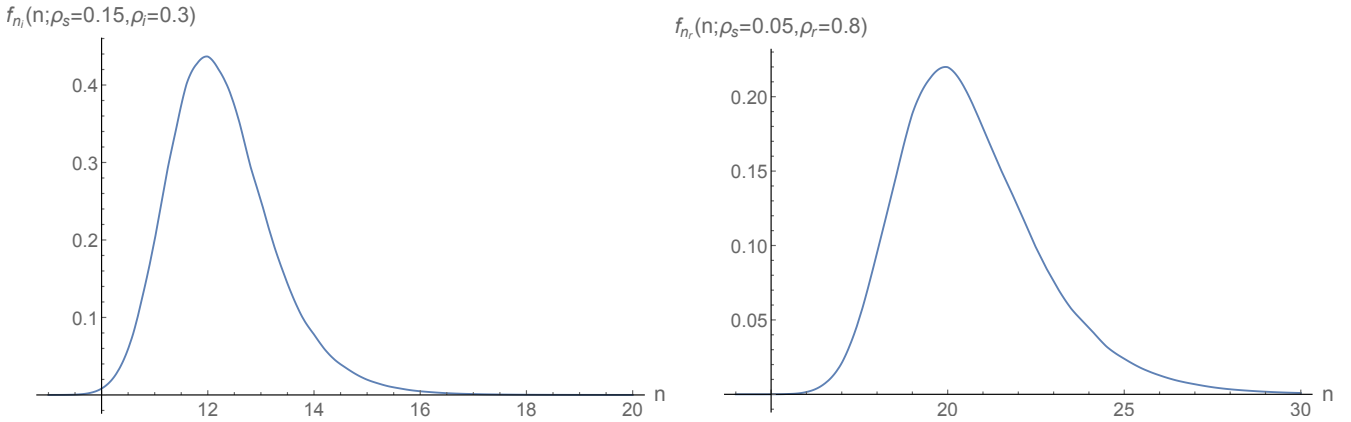


FIGURE 6 Left: PDF of time $n_i(\omega)$ until a given proportion, $\rho_i = 0.3$, of the infected subpopulation remains infected knowing that a proportion, $\rho_s = 0.15$, is susceptible. Rigth: PDF of time $n_r(\omega)$ until a given proportion, $\rho_r = 0.8$, of the recovered subpopulation remains recovered knowing that a proportion, $\rho_s = 0.05$, is susceptible. Both plots correspond to data of pandemic influenza in Egypt represented in Figure 2 .

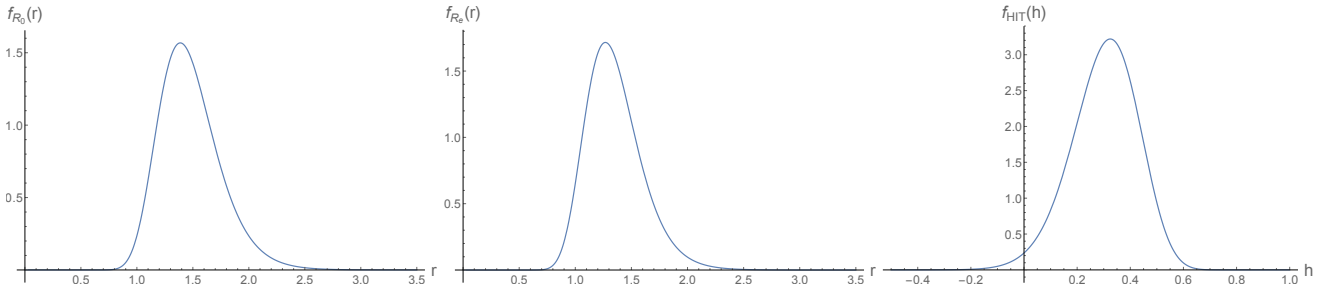


FIGURE 7 PDFs of the basic reproduction number (left), the effective reproduction number (center) and the herd immunity threshold (right). These plots correspond to data of pandemic influenza in Egypt represented in Figure 2 .

6 | CONCLUSIONS

In this paper we have performed a comprehensive study of a randomized Discrete-Time Markov Chain (DTMC) describing SIR-type epidemiological models under very general assumptions. The main novelty of this approach consists of considering that all input data (initial conditions and, contagion and recovering rates) are absolutely continuous random variables instead of deterministic values, and then determining explicit formulas for important probabilistic information associated to the solution stochastic process. Specifically, we have determined the first probability density function (1-PDF) of each subpopulation or state (susceptible, infected and recovered). In this manner, all the one-dimensional moments of the solution, including the mean and the variance, as well as the probability that the solution lies in a set of particular interest can be straightforwardly computed via integration of the PDF. Additional key probabilistic information has been determined as the PDF until a given proportion of the population remains in each state. The probabilistic analysis has been completed by determining explicit expressions for important epidemiological quantities that play a key role in studying the long-run behaviour of a disease. Our approach allows more flexibility and credibility than treating model parameters as deterministic values since real data often involves uncertainty because ignorance and lack of knowledge about the factors determining diseases. This fact has been illustrated by means of the probabilistic study conducted for the pandemic influenza in Egypt. Our approach can also be useful to provide a comprehensive analysis of another epidemiological models via their full randomization and then computing the PDF of the solution stochastic process instead of just computing the mean and the variance as it is usually done instead. Finally, we want to underline that we plan to consider the full randomization of deterministic inhomogeneous DTMC SIR model, that is, when the probabilities in the transition matrix depend on time, so dealing with the scenario where these probabilities are treated as stochastic processes rather than random variables. In this case, we hope to be able to capture uncertainties embedded in the dynamics of the contagion and recovery rates over the time.

ACKNOWLEDGMENTS

This work has been supported by the Spanish Ministerio de Economía y Competitividad grant MTM2017–89664–P. Ana Navarro Quiles acknowledges the funding received from Generalitat Valenciana through a postdoctoral contract (APOSTD/2019/128).

Conflict of interest

The authors declare no potential conflict of interests.

How to cite this article: J.-C. Cortés, S. K. El-Labany, A. Navarro-Quiles, M. M. Selim, and H. Slama (2019), A comprehensive probabilistic analysis of SIR-type epidemiological models based on full randomized Discrete-Time Markov Chain formulation with applications, *Mathematical methods in applied sciences*, *XX*.

APPENDIX

Calculation of the 1-PDF, $f_i(i,n)$, of infected subpopulation

Let n be a fixed cycle, and for each $\omega \in \Omega$, define the mapping $\mathbf{r} : \mathbb{R}^4 \to \mathbb{R}^4$

$$\begin{aligned} x_1 &= r_1(s_0, i_0, k, p) = (1-p)^n i_0 + \frac{((1-k)^n - (1-p)^n)\, k s_0}{p-k}, \\ x_2 &= r_2(s_0, i_0, k, p) = i_0, \\ x_3 &= r_3(s_0, i_0, k, p) = k, \\ x_4 &= r_4(s_0, i_0, k, p) = p. \end{aligned}$$

The inverse mapping $\mathbf{s} : \mathbb{R}^4 \to \mathbb{R}^4$ of $\mathbf{r}$ is

$$\begin{aligned} s_0 &= s_1(x_1, x_2, x_3, x_4) = \frac{(x_1 - x_2(1-x_4)^n)(-x_3 + x_4)}{\left((1-x_3)^n - (1-x_4)^n\right) x_3}, \\ i_0 &= s_2(x_1, x_2, x_3, x_4) = x_2, \\ k &= s_3(x_1, x_2, x_3, x_4) = x_3, \\ p &= s_4(x_1, x_2, x_3, x_4) = x_4, \end{aligned}$$

being its Jacobian

$$J = \frac{-x_3 + x_4}{\left((1-x_3)^n - (1-x_4)^n\right) x_3}.$$

Then applying Theorem 1, the PDF of the random vector $\mathbf{x}(\omega) = (x_1(\omega), x_2(\omega), x_3(\omega), x_4(\omega))$ is

$$f_{\mathbf{x}}(x_1, x_2, x_3, x_4) = f_{s_0, i_0, k, p}\left(\frac{(x_1 - x_2(1-x_4)^n)(-x_3 + x_4)}{\left((1-x_3)^n - (1-x_4)^n\right) x_3}, x_2, x_3, x_4\right) \left|\frac{-x_3 + x_4}{\left((1-x_3)^n - (1-x_4)^n\right) x_3}\right|.$$

As by hypothesis, $s_0(\omega)$, $i_0(\omega)$, $\alpha(\omega)$ and $\beta(\omega)$ are independent random variables, then $s_0(\omega)$, $i_0(\omega)$, $k(\omega)$ and $p(\omega)$ too[21, Th. 3, p.9]. Then, $f_{s_0, i_0, k, p}(s_0, i_0, k, p) = f_{s_0}(s_0) f_{i_0}(i_0) f_k(k) f_p(p)$. Now, marginalizing with respect to $x_2(\omega) = i_0(\omega)$, $x_3(\omega) = k(\omega)$ and $x_4(\omega) = p(\omega)$ and taking n arbitrary, we obtain the 1-PDF of the infected subpopulation $x_1(\omega) = i_n(\omega)$

$$\begin{aligned} f_i(i, n) = \int_{D(p(\omega))} \int_{D(k(\omega))} \int_{D(i_0(\omega))} & f_{s_0}\left(\frac{(i - i_0(1-p)^n)(-k + p)}{((1-k)^n - (1-p)^n)\, k}\right) f_{i_0}(i_0) f_k(k) f_p(p) \\ & \times \left|\frac{-k + p}{((1-k)^n - (1-p)^n)\, k}\right| \mathrm{d}i_0 \, \mathrm{d}k \, \mathrm{d}p. \end{aligned} \tag{1}$$

Recall that $k(\omega) = 1 - \mathrm{e}^{-\alpha(\omega)}$ and $p(\omega) = 1 - \mathrm{e}^{-\beta(\omega)}$, and that we are assuming that the PDFs $f_\alpha(\alpha)$ and $f_\beta(\beta)$ are known. Then, Lemma 1 allows us to express the PDF of $k(\omega)$ and $p(\omega)$ in terms of the PDFs of $\alpha(\omega)$ and $\beta(\omega)$, respectively. Therefore, applying twice Lemma 1, first to $x(\omega) = \alpha(\omega)$, and second to $y(\omega) = p(\omega)$, this yields expression (9) for $f_i(i, n)$.

Calculation of the 1-PDF, $f_r(r, n)$, of recovered subpopulation

Let n be a fixed cycle, we define the mapping $\mathbf{r} : \mathbb{R}^4 \to \mathbb{R}^4$

$$\begin{aligned} x_1 &= r_1(s_0, i_0, k, p) = 1 - (1-p)^n i_0 + \frac{((1-p)^n k - (1-k)^n p)\, s_0}{p - k}, \\ x_2 &= r_2(s_0, i_0, k, p) = i_0, \\ x_3 &= r_3(s_0, i_0, k, p) = k, \\ x_4 &= r_4(s_0, i_0, k, p) = p. \end{aligned}$$

The inverse mapping $\mathbf{s} : \mathbb{R}^4 \to \mathbb{R}^4$ of $\mathbf{r}$ is

$$\begin{aligned} s_0 &= s_1(x_1, x_2, x_3, x_4) = \frac{(-1 + x_1 + x_2(1-x_4)^n)(-x_3 + x_4)}{x_3(1-x_4)^n - (1-x_3)^n x_4}, \\ i_0 &= s_2(x_1, x_2, x_3, x_4) = x_2, \\ k &= s_3(x_1, x_2, x_3, x_4) = x_3, \\ p &= s_4(x_1, x_2, x_3, x_4) = x_4, \end{aligned}$$

being its Jacobian

$$J = \frac{-x_3 + x_4}{x_3(1-x_4)^n - (1-x_3)^n x_4}.$$

Then applying Theorem 1, the PDF of the random vector $\mathbf{x}(\omega) = (x_1(\omega), x_2(\omega), x_3(\omega), x_4(\omega))$ is

$$f_{\mathbf{x}}(x_1, x_2, x_3, x_4) = f_{s_0, i_0, k, p}\left(\frac{(-1 + x_1 + x_2(1-x_4)^n)(-x_3 + x_4)}{x_3(1-x_4)^n - (1-x_3)^n x_4}, x_2, x_3, x_4\right) \left|\frac{-x_3 + x_4}{x_3(1-x_4)^n - (1-x_3)^n x_4}\right|.$$

We know that $f_{s_0,i_0,k,p}(s_0,k,p) = f_{s_0}(s_0)f_{i_0}(i_0)f_k(k)f_p(p)$, then marginalizing with respect to $i_0(\omega)$, $k(\omega)$ and $p(\omega)$, and taking n arbitrary, the 1-PDF of the recovered subpopulation is straightforwardly obtained

$$f_r(r,n) = \int\limits_{D(p(\omega))}\int\limits_{D(k(\omega))}\int\limits_{D(i_0(\omega))} f_{s_0}\left(\frac{(-1+r+i_0(1-p)^n)(-k+p)}{(1-p)^n k-(1-k)^n p}\right) f_{i_0}(i_0)f_k(k)f_p(p) \times\left|\frac{-k+p}{(1-p)^n k-(1-k)^n p}\right| \mathrm{d}i_0\,\mathrm{d}k\,\mathrm{d}p. \tag{2}$$

Finally, we apply twice Lemma 1 to express (2) in terms of the data. This yields expression (10) for $f_r(r,n)$.

Calculation of the PDF, $f_{n_i}(n)$, of the time until a given proportion of the population remains infected

Let $\rho_s, \rho_i \in (0,1)$ fixed proportions of susceptible and infected subpopulations, respectively. Notice that $0 < \rho_s + \rho_i < 1$. For each $\omega \in \Omega$ we define the mapping $\mathbf{r} : \mathbb{R}^4 \to \mathbb{R}^4$

$$\begin{aligned} x_1 &= r_1(s_0,i_0,\alpha,\beta) = \frac{1}{\beta}\ln\left(\frac{\mathrm{e}^{\alpha}\, i_0 + \mathrm{e}^{\alpha+\beta}\, s_0 - \mathrm{e}^{\beta}(i_0+s_0)}{\mathrm{e}^{\alpha}\,\rho_i + \mathrm{e}^{\alpha+\beta}\,\rho_s - \mathrm{e}^{\beta}(\rho_i+\rho_s)}\right)\\ x_2 &= r_2(s_0,i_0,\alpha,\beta) = i_0,\\ x_3 &= r_3(s_0,i_0,\alpha,\beta) = \alpha,\\ x_4 &= r_4(s_0,i_0,\alpha,\beta) = \beta. \end{aligned}$$

The inverse mapping $\mathbf{s} : \mathbb{R}^4 \to \mathbb{R}^4$ of $\mathbf{r}$ is

$$\begin{aligned} s_0 &= s_1(x_1,x_2,x_3,x_4) = \frac{\mathrm{e}^{-x_4}}{-1+\mathrm{e}^{x_3}}\left(\mathrm{e}^{x_1x_4}\left(-\mathrm{e}^{x_4}\left(\rho_i+\rho_s\right)+\mathrm{e}^{x_3}\left(\rho_i+\mathrm{e}^{x_4}\,\rho_s\right)\right)+\left(-\mathrm{e}^{x_3}+\mathrm{e}^{x_4}\right)x_2\right),\\ i_0 &= s_2(x_1,x_2,x_3,x_4) = x_2,\\ \alpha &= s_3(x_1,x_2,x_3,x_4) = x_3,\\ \beta &= s_4(x_1,x_2,x_3,x_4) = x_4, \end{aligned}$$

being its Jacobian

$$J = \frac{\mathrm{e}^{(-1+x_1)x_4}\,x_4}{-1+\mathrm{e}^{x_3}}\left(-\mathrm{e}^{x_4}\left(\rho_i+\rho_s\right)+\mathrm{e}^{x_3}\left(\rho_i+\mathrm{e}^{x_4}\,\rho_s\right)\right).$$

Then applying Theorem 1, the PDF of the random vector $\mathbf{x}(\omega) = (x_1(\omega), x_2(\omega), x_3(\omega), x_4(\omega))$ is

$$f_{\mathbf{x}}(x_1,x_2,x_3,x_4) = f_{s_0,i_0,\alpha,\beta}\left(\frac{\mathrm{e}^{-x_4}}{-1+\mathrm{e}^{x_3}}\left(\mathrm{e}^{x_1x_4}\left(-\mathrm{e}^{x_4}\left(\rho_i+\rho_s\right)+\mathrm{e}^{x_3}\left(\rho_i+\mathrm{e}^{x_4}\,\rho_s\right)\right)+\left(-\mathrm{e}^{x_3}+\mathrm{e}^{x_4}\right)x_2\right), x_2, x_3, x_4\right) \times\left|\frac{\mathrm{e}^{(-1+x_1)x_4}\,x_4}{-1+\mathrm{e}^{x_3}}\left(-\mathrm{e}^{x_4}\left(\rho_i+\rho_s\right)+\mathrm{e}^{x_3}\left(\rho_i+\mathrm{e}^{x_4}\,\rho_s\right)\right)\right|.$$

We know that $f_{s_0,i_0,\alpha,\beta}(s_0,\alpha,\beta) = f_{s_0}(s_0)f_{i_0}(i_0)f_{\alpha}(\alpha)f_{\beta}(\beta)$, then marginalizing with respect to $i_0(\omega)$, $\alpha(\omega)$ and $\beta(\omega)$, the PDF, $f_{n_i}(n;\rho_s,\rho_i)$, of the time until a given proportion, ρ_i, remains infected is straightforwardly obtained by expression (12).

Calculation of the PDF, $f_{n_r}(n)$, of the time until a given proportion of the population remains recovered

Let $\rho_s, \rho_r \in (0,1)$ fixed proportions of infected and recovered subpopulations, respectively. Notice that $0 < \rho_s + \rho_r < 1$. For each $\omega \in \Omega$ we define the mapping $\mathbf{r} : \mathbb{R}^4 \to \mathbb{R}^4$

$$x_1 = r_1(s_0,i_0,\alpha,\beta) = \frac{1}{\beta(\omega)}\ln\left(\frac{\mathrm{e}^{\alpha}\, i_0 + \mathrm{e}^{\alpha+\beta}\, s_0 - \mathrm{e}^{\beta}(i_0+s_0)}{\mathrm{e}^{\alpha}(1-\rho_r-\rho_s) - \mathrm{e}^{\beta}(1-\rho_r-\mathrm{e}^{\alpha}\,\rho_s)}\right).$$

$$\begin{aligned} x_2 &= r_2(s_0,i_0,\alpha,\beta) = i_0\\ x_3 &= r_3(s_0,i_0,\alpha,\beta) = \alpha\\ x_4 &= r_4(s_0,i_0,\alpha,\beta) = \beta \end{aligned}$$

The inverse mapping $\mathbf{s} : \mathbb{R}^4 \to \mathbb{R}^4$ of $\mathbf{r}$ is

$$\begin{aligned} s_0 &= s_1(x_1, x_2, x_3, x_4) = \frac{\mathrm{e}^{-x_4}}{-1+\mathrm{e}^{x_3}} \left(\mathrm{e}^{x_1 x_4}\left(\mathrm{e}^{x_4}\left(-1+\rho_r\right) + \mathrm{e}^{x_3+x_4}\,\rho_s + \mathrm{e}^{x_3}\left(1-\rho_r-\rho_s\right)\right) + \left(-\mathrm{e}^{x_3}+\mathrm{e}^{x_4}\right) x_2\right), \\ i_0 &= s_2(x_1, x_2, x_3, x_4) = x_2, \\ \alpha &= s_3(x_1, x_2, x_3, x_4) = x_3, \\ \beta &= s_4(x_1, x_2, x_3, x_4) = x_4, \end{aligned}$$

being its Jacobian

$$J = \frac{\mathrm{e}^{(-1+x_1)x_4}\, x_4}{-1+\mathrm{e}^{x_3}} \left(\mathrm{e}^{x_4}\left(-1+\rho_r\right) + \mathrm{e}^{x_3+x_4}\,\rho_s + \mathrm{e}^{x_3}\left(1-\rho_r-\rho_s\right)\right).$$

Then applying Theorem 1, the PDF of the random vector $\mathbf{x}(\omega) = (x_1(\omega), x_2(\omega), x_3(\omega), x_4(\omega))$ is

$$f_{\mathbf{x}}(x_1, x_2, x_3, x_4) = f_{s_0, i_0, \alpha, \beta}\left(\frac{\mathrm{e}^{-x_4}}{-1+\mathrm{e}^{x_3}} \left(\mathrm{e}^{x_1 x_4}\left(\mathrm{e}^{x_4}\left(-1+\rho_r\right) + \mathrm{e}^{x_3+x_4}\,\rho_s + \mathrm{e}^{x_3}\left(1-\rho_r-\rho_s\right)\right) + \left(-\mathrm{e}^{x_3}+\mathrm{e}^{x_4}\right) x_2\right), x_2, x_3, x_4\right)$$

$$\times \left| \frac{\mathrm{e}^{(-1+x_1)x_4}\, x_4}{-1+\mathrm{e}^{x_3}} \left(\mathrm{e}^{x_4}\left(-1+\rho_r\right) + \mathrm{e}^{x_3+x_4}\,\rho_s + \mathrm{e}^{x_3}\left(1-\rho_r-\rho_s\right)\right) \right|.$$

We know that $f_{s_0, i_0, \alpha, \beta}(s_0, \alpha, \beta) = f_{s_0}(s_0) f_{i_0}(i_0) f_\alpha(\alpha) f_\beta(\beta)$, then marginalizing with respect to $i_0(\omega)$, $\alpha(\omega)$ and $\beta(\omega)$, the PDF, $f_{n_r}(n; \rho_s, \rho_r)$, of the time until a given proportion, ρ_r, remains recovered is straightforwardly obtained by expression (13).

References

1. Allen Linda J.S.. *An Introduction to Stochastic Processes with Applications to Biology*. New York: CRC Press; 2010.
2. Hethcote H. W.. The mathematics of infectious diseases. *SIAM Review.* 2000;42(4):599–653.
3. Brauer F., Castillo-Chávez C.. *Mathematical Models in Population Biology and Epidemiology*. New York: Springer; 2001.
4. Bailey N. T. J.. *The Mathematical Theory of Infectious Diseases*. New York: Hafner Press; 2nd ed.1975.
5. Anderson R. M.. *Population Dynamics of Infectious Diseases*. London: Chapman & Hall; 1982.
6. Anderson R. M., May R. M.. *Infectious Diseases of Humans*. London: Oxford University; 1992.
7. Cortés J. C., Navarro-Quiles A., Romero J. V., Roselló M. D.. Some results about randomized binary Markov chains: theory, computing and applications. *International Journal of Computer Mathematics.* ;.
8. Cortés J. C., Navarro-Quiles A., Romero J. V., Roselló M. D.. Randomizing the parameters of a Markov chain to model the stroke disease. *Journal of Computational and Applied Mathematics.* 2017;324:225–240.
9. Casabán M. C., Cortés J. C., Romero J. V., Roselló M. D.. Probabilistic solution of random SI-type epidemiological models using the Random Variable Transformation technique. *Communications in Nonlinear Science and Numerical Simulation.* 2015;24(1–3):86–97.
10. Casabán M. C., Cortés J. C., Navarro-Quiles A., Romero J. V., Roselló M. D., Villanueva R. J.. A comprehensive probabilistic solution of random SIS-type epidemiological models using the Random Variable Transformation technique. *Communications in Nonlinear Science and Numerical Simulation.* 2016;32:199–210.
11. Slama H., Hussein A., El-Bedwhey N. A., Selim M. M.. An approximate probabilistic solution of a random SIR-type epidemiological model using RVT technique. *Applied Mathemtics and Computation.* 2019;361:144–156.
12. Hussein A., Selim M. M.. Solution of the stochastic radiative transfer equation with Rayleigh scattering using RVT technique. *Applied Mathematics and Computation.* 2012;218(13):7193-7203.
13. Santos L. T., Dorini F. A., Cunha M. C. C.. The probability density function to the random linear transport equation. *Applied Mathematics and Computation.* 2010;216(5):1524–1530.

14. Hussein A., Selim M. M.. Solution of the stochastic transport equation of neutral particles with anisotropic scattering using RVT technique. *Applied Mathematics and Computation.* 2009;213(1):250–261.

15. Hussein A., Selim M. M.. Solution of the stochastic radiative transfer equation with Rayleigh scattering using RVT technique. *Applied Mathematics and Computation.* 2012;218(13):7193–7203.

16. Slama H., El-Bedwhey N. A., El-Depsy A., Selim M. M.. Solution of finite Milne problem in stochastic media with RVT technique. *European Physics Journal Plus.* 2017;132:505.

17. Hussein A., Selim M. M.. On the generalization of probabilistic transformation method. *Applied Mathematics and Computation.* 2007;190(15):1284–1289.

18. Kegan B., Webster West R.. Modeling the simple epidemic with deterministic differential equations and random initial conditions. *Mathematical Biosciences.* 2005;195(5):179–193.

19. Dorini F. A., Cecconello M. S., Dorini L. B.. On the logistic equation subject to uncertainties in the environmental carrying capacity and initial population density. *Communications in Nonlinear Science and Numerical Simulation.* 2016;33:160–173.

20. Soong T. T.. *Random Differential Equations in Science and Engineering*. New York: Academic Press; 1973.

21. Grimmett G. R., Stirzaker D. R.. *Probability and Random Processes*. Oxford: Clarendon Press; 2000.

22. Van Den Driessche P.. Reproduction numbers of infectious disease models. *Infectious Disease Modelling.* 2017;2:288–303.

23. Heffernan J. M., Smith R. J., Wahl L. M.. Perspectives on the basic reproductive ratio. *Journal of the Royal Society Interface.* 2005;2:281–293.

24. Khalil K. M., Abdel-Aziz M., Nazmy T. T., Salem A.-B. M.. An agent-based modeling for pandemic influenza in Egypt. *Handbook on Decision Making.* 2012;33:205–218.

25. Øksendal B.. *Stochastic Differential Equations: An Introduction with Applications*. New York: Springer; 6th ed.2010.

26. Allen E.. *Modeling with Itô Stochastic Differential Equations*. New York: Springer; 2007.

27. Nelsen R. B.. *An Introduction to Copulas*. New York: Springer; 1999.

28. Villafuerte L., Braumann C.A., Cortés J. C., Jódar L.. Random differential operational calculus: Theory and Applications. *Computers and Mathematics with Applications.* 2010;59(1):115–125.

29. Cortés J. C., Villafuerte L., Jódar L.. Numerical solution of random differential initial value problems: Multistep methods. *Mathematical Methods in the Applied Science.* 2011;34(1):63–75.

30. Ghanem R. G., Spanos P. D.. *Stochastic Finite Elements: A Spectral Approach*. New York: Courier Dover Publ.; 2003.

31. Zhang J., Yu K. F.. What's the relative risk? A method of correcting the odds ratio in cohort studies of common outcomes. *Journal of American Medical Association (JAMA).* 1998;290(19):1690–1691.

32. McNutt L. A., Wu C., Xue X., Hafner J. P.. Estimating the Relative Risk in Cohort Studies and Clinical Trials of Common Outcomes. *American Journal of Epidemiology.* 2003;157(10):940–943.

33. Calbo G., Cortés J. C., Jódar L.. Random Hermite differential equations: Mean square power series solutions and statistical properties. *Applied Mathematics and Computation.* 2011;218(7):3654–3666.

34. Kroese D. P., Taimre T., Botev Z. I.. *Handbok of Monte Carlo Methods*. New Jersey: Wiley; 2011.

35. El-Tawil M. A.. The approximate solutions of some stochastic differential equations using transformations. *Applied Mathematics and Computation.* 2005;164(1):167–178.

36. Casabán M. C., Cortés J. C., Romero J. V., Roselló M. D.. Determining the first probability density function of linear random initial value problems by the Random Variable Transformation (RVT) technique: A comprehensive study. *Abstract and Applied Analysis.* 2014;2014–ID248512:1–25.

37. Reid W. T.. *Riccati Differential Equations*. New York: Academic Press; 1972.

38. Ascher U. M., Mattheij R. M., Rusell R. D.. *Numerical Solution of Boundary Value Problems for Ordinary Differential Equations*. New Jersey: Englewood Cliffs: Prentice Hall; 1988.

39. Cortés J. C., Jódar L., Company R., Villafuerte L.. Solving Riccati time-dependent models with random quadratic coefficient. *Applied Mathematics Letters.* 2011;24(12):2193–2196.

40. Rogers E.M.. *Diffusion of Innovations*. New York: Free Press; 2003.

41. Calbo G., Cortés J. C., Jódar L.. Random matrix difference models arising in long-term medical drug strategies. *Applied Mathematics and Computation.* 2010;217(5):2149–2161.

42. Casabán M. C., Cortés J. C., Romero J. V., Roselló M. D.. Probabilistic solution of random homogeneous linear second-order difference equations. *Applied Mathematics Letters.* 2014;34(2):27–32.

43. Zhang D., Ntoko A.. Mathematical Model of Technology Diffusion in Developing Countries Computational Methods in Decision-Making. *Mathematical Biosciences.* 2005;195(5):179–193.

44. Abramowitz M., Stegun I. A. (editors). *Handbook of Mathematical Functions, with Formulas, Graphs, and Mathematical Tables*. New York: Dover; 1972.

45. El-Tawil M. A., El-Tahan W., Hussein A.. Using FEM-RVT technique for solving a randomly excited ordinary differential equation with a random operator. *Applied Mathematics and Computation.* 2007;187(2):856–867.

46. Hussein A., Selim M. M.. A developed solution of the stochastic Milne problem using probabilistic transformations. *Applied Mathematics and Computation.* 2009;216(10):2910–2919.

47. Casella G., Berger R.L.. *Statistical Inference*. New York: Brooks/Cole; 2002.

48. Nelder J. A., Mead R.. A simplex method for function minimization. *Computer Journal.* 1964;7(1):308–313.

49. Enders W.. *Applied Econometric Times Series*. New York: John Wiley & Sons; 1995.

50. Gard T. C.. *Introduction to Stochastic Differential Equations*. New York: Marcel Dekker; 1988.

51. He B.. Computational Structural Engineeringch. Building Structures Vibration Differential Equations under Random Excitation, :787–793. The Netherlands: Springer 2009.

52. Kadry S.. Probabilistic solution of rational difference system with random parameters. *ISRN Applied Mathematics.* 2012;2012:Article ID 290186, 6 pages.

53. Kloeden P. E., Platen E.. *Numerical Solution of Stochastic Differential Equations* Applications of Mathematics: Stochastic Modelling and Applied Probability, vol. 23: . New York: Springer; 3th ed.1999.

54. Papoulis A., Pillai S. U.. *Probability, Random Variables and Stochastic Processes*. New York: McGraw-Hill; 4th ed.2002.

55. Anderson R. M., May R. M.. *Infectious Diseases of Humans*. London: Oxford University; 1992.

56. Simonoff J. S.. *Smoothing Methods in Statistics*. New York: Springer-Verlag; 1996.

57. Yang Z., Xu D.. Mean square exponential stability of impulsive stochastic difference equations. *Applied Mathematics Letters.* 2007;20(8):938–945.

58. CMT , ed.*Anual Report of the Spanish National Committee of Telecommunications 2010–2012*. Madrid: CMT; 2013.

59. Briggs A., Sculpher M.. An Introduction to Markov Modelling for Economic Evaluation. *PharmacoEconomics.* 1998;13(4):397–409.

60. Sonnenberg F.A., Beck R.. Markov models in medical decision making: a practical guide. *Medical Decision Making.* 1993;13(4):322–338.

61. Mar J., Antoñanzas F., Pradas R., Arrospide A.. Los modelos de Markov probabilísticos en la evaluación económica de tecnologías sanitarias: una guía práctica. *Gaceta Sanitaria.* 2010;24(3):209–214.

62. Mar J., Sainz-Ezkerra M., Miranda-Serrano E.. Calculation of prevalence with Markov models: Budget impact analysis of thrombolysis for stroke. *Medical Decision Making.* 2008;28(4):481–490.

63. Hazen G., Li Z.. Cohort decomposition for Markov cost-effectiveness models. *Medical Decision Making.* 2011;31(1):19–34.

64. Sacristán J. A., Oliva J., Del Llano J., Prieto L., Pinto J. L.. ¿Qué es una tecnología sanitaria eficiente en España?. *Gaceta Sanitaria.* 2002;4(16):334–343.

65. Pinto-Prades J. L., Puig-Junoy J., Ortún-Rubio V.. Análisis coste-utilidad. *Atención Primaria.* 2001;27:569–573.

66. Rubio-Terrés C., Echevarría A. Modelos de Markov: Una herramienta útil para el análisis farmacoeconómico. *Pharmacoeconomics.* 2006;3:71–78.

67. Badia X., Bueno H., González-Juanatey J. R., Valentín V., Rubio M.. Análisis de la relación coste-efectividad a corto y largo plazo de clopidogrel añadido a terapia estándar en pacientes con síndrome coronario agudo en España. *Rev. Esp. Cardiol.*. 2005;58:1385–1395.

68. Antoñanzas F., Brenes F., Molero J. M., et al. Cost-Effectiveness of the combination therapy of dutasteride and Tamsulosin in the treatment of benign Prostatic Hyperlasia in Spain. *Actas Urol. Esp.*. 2011;35(2):65–71.

69. Prieto L., Sacristán J. A., Antoñanzas F., Rubio-Terrés C., Pinto J. L., Rovira J.. Análisis coste-efectividad en la evaluación económica de intervenciones sanitarias. *Med. Clin.*. 2004;122(13):505–510.

70. Karlsson G., Johannesson M.. The decision rules of cost-effectiveness analysis. *PharmacoEconomics.* 1996;9(2):113–120.

71. Cooper B., Lipsitch M.. The analysis of hospital infection data using hidden Markov models. *Biostatistics.* 2004;5:223âĂŞ-237.

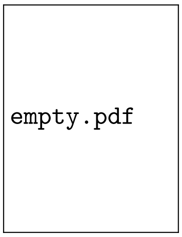
72. Mandel M., Betensky R. A.. Estimating time-to-event from longitudinal ordinal data using random-effects Markov models: application to multiple sclerosis progression. *Biostatistics.* 2008;9:750âĂŞ-764.

73. Korostil I. A., Peters G. W., Cornebis J., Regan R. G.. Adaptive Markov chain Monte Carlo forward projection for statistical analysis in epidemic modelling of human papillomavirus. *Statistics in Medicine.* 2012;32(11):1917-âĂŞ1953.

74. Kapetanakis V., Matthews F. E., Hout A.. A semi-Markov model for stroke with piecewise-constant hazards in the presence of left, right and interval censoring. *Statistics in Medicine.* 2013;32(4):697-âĂŞ713.

75. Pan S. L., Wu H. M., Yen A. M. F., Chen T. H. H.. A Markov regression random-effects model for remission of functional disability in patients following a first stroke: A Bayesian approach. *Statistics in Medicine.* 2007;26(29):5335-âĂŞ5353.

AUTHOR BIOGRAPHY



Author Name. This is sample author biography text this is sample author biography text.