ON AN EPIDEMIC MODEL GIVEN BY A STOCHASTIC DIFFERENTIAL EQUATION

Zarife ZARARSIZ

June 2009

MSI
Växjö University
SE-351 95 VÄXJÖ

Report 09017
ISSN 1650-2647
ISRN VXU/MSI/MA/C/--09017/--SE
Zarife ZARARSIZ

ON AN EPIDEMIC MODEL GIVEN BY A STOCHASTIC DIFFERENTIAL EQUATION

Bachelor’s thesis

Mathematics

2009
Abstract

We investigate a certain epidemics model, with and without noise. Some parameter analysis is performed together with computer simulations. The model was presented in Iacus (2008).

Key-words: Epidemics models, stochastic differential equations, Lotka-Volterra equations, scale measure, speed measure, limiting distributions, reflecting boundaries, killed processes.

Acknowledgments

We would like to express our gratitude to all those gave us the possibility to complete our study at the Växjö University. Thanks to our teachers, Roger Pettersson for providing us all the knowledge, that we needed to understand the Matlab, Latex, Statistics. Thanks to Växjö University for having the open doors for international students and gave us the opportunity to expand our knowledge through the world. Thanks to Växjö University to offer us information sources (library, database and so on.). Each of us for making effort to get the project effective and productive...

Contents

1 Discussion on deterministic and stochastic models 1
2 Population and epidemics modelling 1
3 A classical epidemics model 2
   3.1 Classical deterministic epidemics model 2
   3.2 Classical epidemics model, here with noise 4
4 A model for the fraction of infected people 6
   4.1 A deterministic model 6
   4.2 Stochastic epidemics fraction model 10
A Appendix 18
   A.1 Classical epidemics models 18
   A.2 Code for fraction models 19
References 20
1 Discussion on deterministic and stochastic models

A deterministic model is a mathematical model in which outcomes are precisely determined through known relationships among states and events, without any room for random variation. In such models, a given input will always produce the same output, such as in a known chemical reaction. In comparison, in stochastic models the outcome is random according to probability distributions. Often a purpose of a model is to make predictions and try "What If?" scenarios. You can change the inputs and recalculate the model and you'll get a new answer. You might even want to plot a graph of the future value vs. time. In economics applications you may calculate under a fixed interest rate, but what do you do if the interest rate is allowed to change? For this simple situation, you might only care to know a worst/best case scenario, where you calculate the future value based upon the lowest and highest interest rates that you might expect. Another example, the planets move around the sun according to Newton's laws and their position can be predicted with great accuracy well into the future. Even in this situation, a totally deterministic relationship is unlikely due to unpredictable factors - for example, a previously unknown comet moving through the solar system could perturb somewhat some planets.

A stochastic model is a mathematical model involving randomness: you don’t know future outcome values but you describe the outcomes in terms of probability. Exact prediction is not possible, but it may be possible to predict within a known confidence interval - or predict the probability that a particular value will be observed at a particular time.

2 Population and epidemics modelling

The mathematical theory of epidemiology [1] contains many types of models, both deterministic and stochastic which have been adapted and handled also outside population dynamics modelling. There are different ways to model epidemics for example the classical model described in the next Section.

Why stochastic modelling? From [6], pp 1-4: "Recent interest in population dynamics has polarized to an undesirable extent, though there are hopeful signs that this situation may soon improve." ... " Whilst conversely, some mathematically oriented biologists have gained considerable mileage out of developing supposedly plausible deterministic models with interesting mathematical features. As long as no one questions whether these features relate to biological reality or pure imagination, then not only will this approach remain unchallenged but it will also thrive on its own increasing aura of biological respectability. The tragedy is that too few researchers realize that both deterministic and stochastic models have important roles to play in the analysis of any particular system. Slavish obedience to one specific approach can lead to disaster. Provided that population numbers never become too small then a deterministic model may enable sufficient biological understanding to be gained about the system; if at any population numbers do become small then a stochastic analysis is vital. So pursuing both approaches simultaneously ensures that we do not become trapped either by deterministic fantasy"... "As an example, suppose that all members of a population develop independently from each

\[ X \]  \[ \omega \]  \[ C(\mathbb{R}) \]  \[ \omega(t) \]

\[ X \] is a function \( \Omega \ni \omega \mapsto X(\omega) \in \mathbb{R} \). The \( \omega \) represents a random event. Often, for convenient reasons, one writes \( X(\omega) \) instead of \( X \). The random variable should be measurable in a certain way so that we can give probabilities for events of the type \( \{X \leq x\} \), where \( x \) is a fixed value in \( \mathbb{R} \). A stochastic process \( \{X_t : t \geq 0\} \) is a set of random variables: for each \( t \), we have a random variable \( X_t \). A for example real-valued continuous stochastic process \( X = \{X_t : t \geq 0\} \) can also be seen as a function \( \{C(\mathbb{R}) \ni \omega \mapsto X(\omega) \in C(\mathbb{R}) : X_t(\omega) = \omega(t)\} \), that means the outcome of the process \( \{X_t : t \geq 0\} \) is a random continuous function \( \omega(t) \).
other, and reproduce at rate $\lambda$ and die at rate $\mu$. Then the deterministic number of individuals alive at time $t$, starting from an initial population of size $N(0)$ at time 0, is given by

$$N(t) = N(0) \exp(\lambda - \mu)t.$$ 

If births predominate over deaths then this result tells us that the population size will explode exponentially fast, whilst if deaths predominate then extinction is inevitable. For large $N(0)$ and realistically small $t$ this is indeed an excellent description, but let us reflect for a moment on what happens if $N(0)$ is small. Indeed, suppose that $N(0) = 1$ and $\lambda = 2\mu$ so that births are twice as likely to occur as deaths. Then results

$$N(t) = N(0) \exp(\lambda - \mu)t$$

predicts exponential growth with $N(t) = \exp(\mu t)$. But the first event to occur may be a death, with probability $\mu/(\mu + \lambda) = 1/3$, and this results in the population immediately becoming extinct. Thus the probability of ultimate extinction is at least $\frac{1}{3}$ which is in direct contradiction to the deterministic prediction. The situation becomes even more contradictory when $\lambda = \mu$. Then $N(t)$ remains absolutely constant at $N(t) = 1$ in spite of the fact that the actual (i.e. stochastic) process specifically involves birth and death. Far more sinister is the popular desire to infer the nature of population development from a single experimental run." Ask people "to predict how a population develops through time when they are presented with a single sequence of observations and their response suddenly changes. An apparently hard lesson to learn is that different realizations from the same process can vary enormously." "The modelling of epidemics exposes the problem that deterministic prediction can be substantially different from what is expected under a stochastic model (i.e. the average of a large number of simulated realizations). For example, the expected rate at which new infectives accrue is of considerable interest to public health officers in deciding the severity and likely length of duration of an epidemic. Yet deterministic and stochastic assumptions generate substantially different values for this rate and so the two approaches might easily lead to different conclusions. Full conditions for agreement between deterministic and expected stochastic solutions are at present unknown."

3  A classical epidemics model

There is a classical deterministic epidemics model similar to the famous Lotka-Volterra equations. In this section we present this model briefly. We also present a stochastic version which may be more realistic including uncertainties and probabilities.

3.1  Classical deterministic epidemics model

First we consider the Lotka-Volterra equations, describing not epidemics but the numbers of preys and predator, namely

$$\begin{align*}
\frac{dN_1}{dt} &= N_1(r_1 - b_1N_2) \\
\frac{dN_2}{dt} &= N_2(-r_2 + b_2N_1),
\end{align*}$$

and make the following correspondence, n.o. (number of)

\begin{align*}
N_1(t) &\quad (\text{n.o. preys}) \rightarrow x(t) \quad (\text{n.o. susceptible to a disease}) \\
N_2(t) &\quad (\text{n.o. predators}) \rightarrow y(t) \quad (\text{n.o. individuals infected by a disease}).
\end{align*}
Then on taking $r_1 = 0$, $b_1 = b_2 = \beta$, and $r_2 = \gamma$, (3.1) may be written in the epidemic form

$$\begin{align*}
\frac{dx}{dt} &= -\beta xy \\
\frac{dy}{dt} &= \beta xy - \gamma y.
\end{align*}$$

(3.2)

In brief, $\beta$ is the infection rate, $xy$ is the number of possible contact-pairs between susceptibles and infectives, and $\gamma$ is the death (or removal) rate of infectives. The fundamental difference between these two sets of equations lies in their interpretation. In (3.1) the death of a susceptible automatically gives rise to the birth of an infective, since it involves a transfer of state for the same individual, whilst in (3.2) preys do not become predators but merely act as food for them. Though equations (3.1) and (3.2) are naïve descriptors of epidemic development, they do generate useful qualitative predictions about possible modes of behaviour. From [6], pp324-325: "the importance of epidemiology cannot be overstated. To quote Bailey’s figures" (see for example [1]), "in fourteenth century Europe there were 25 million deaths out of a population of 100 million from the Black death alone, with whole towns and villages being virtually annihilated; the Azhects lost half their population of $3\frac{1}{2}$ million from smallpox; whilst around 20 million people died in the world pandemic of influenza in 1919. Today, the over-riding concern is the spread of AIDS, whilst vast numbers of people remain affected by less media-consious diseases such as malaria (350 million living in endemic areas), filariasis (250 million) and hookworm disease (450 million). The total burden of human misery and suffering that results from communicable disease is immense, and any understanding that modelling techniques can bring to alleviate this terrible state of affairs is truly important. Light can be shed on the life-cycle of the parasite, the transmission and spread of an infectious disease, the nature of threshold population densities above which an epidemic can flare up, and methods for optimal immunization and control. In all of these matters mathematical and statistical studies have an essential role to play. Indeed, bearing in mind the great variety of infectious diseases that currently affect our planet, and how relatively few theoretical inroads have been made into understanding most of them, there is clearly a vast field of worthwhile and directly applicable research just waiting to take off for those with the interest and imagination to embark upon it."

Figure 3.1 illustrates (3.2) and Figure 3.2 describes the fraction of infected individuals from the same equation.

![Figure 3.1](image)

Figure 3.1: Modelling of number of infected and non-infected individuals (3.2) with $\beta = 1$, $\gamma = 0.2$. Start values for the infected and non-infected are 1 and 2, respectively.
Figure 3.2: Modelling of the number of infected and non-infected individuals (3.2) with $\beta = 1$, $\gamma = 0.2$. Instead of the numbers, the fraction of infected individuals are plotted. Start values for the infected and non-infected are 1 and 2, respectively, i.e. the start value for the fraction of infected individuals is $1/3$.

3.2 Classical epidemics model, here with noise

A stochastic version of (3.2) is

$$
\frac{dx}{dt} = \begin{bmatrix}
-\beta xy \\
\beta xy
\end{bmatrix} dt + \begin{bmatrix}
dw_1 \\
dw_2
\end{bmatrix}
$$

(3.3)

where $w_1$ and $w_2$ are two independent Brownian motions$^2$ and $C$ satisfies $CC^T = B$ for

$$
B = \begin{bmatrix}
\beta xy & -\beta xy \\
-\beta xy & \beta xy - \gamma y
\end{bmatrix}
$$

see [3] for similar model, but there aspects of a stochastic Lotka Volterra equation was investigated.

Figure 3.3 illustrates (3.3) and Figure 3.4 describes the fraction of infected individuals from the same equation.

---

$^2$Heuristically: at an infinitesimal time interval $[t, t + dt]$, $w_1$ has moved a step $dw_1(t) = \sqrt{dt}N(0,1)$: in other words $w_1(t + dt) - w_1(t) = \sqrt{dt}N(0,1)$. Similarly with $w_2$. A more precise definition of a Brownian motion $w = \{w(t) : t \geq 0\}$ is that it is a continuous stochastic process starting at zero with independent increments $w(t + h) - w(t) \sim \sqrt{h}N(0,1)$-distributed, $h > 0$. 

Figure 3.3: Stochastic modelling of number of infected and non-infected individuals (3.3) with $\beta = 1$, $\gamma = 0.2$. Start values for the number of infected and non-infected are 1 and 2, respectively.

Figure 3.4: Stochastic modelling of number of infected and non-infected individuals (3.3) with $\beta = 1$, $\gamma = 0.2$. Here the fraction infected individuals. Start values for the infected and non-infected are 1 and 2, respectively, i.e. the start value for the fraction of infected individuals is 1/3.
4 A model for the fraction of infected people

Here we discuss a model where the fraction of infected individuals are directly modelled, [2]. First we discuss a model described by an ordinary differential equation. Then we discuss a stochastic variant.

4.1 A deterministic model

We investigate a simple epidemic model presented in Iacus [2]. Let \( x(t) \) be the fraction of a population that has an infectious disease at time \( t \). If the disease does not confer immunity, then one model dealing directly with the fraction of infected individuals is that the fraction of infected individuals are given by the deterministic ordinary differential equation

\[
\frac{dx}{dt} = ax(1-x)dt - bxdt + c(1-x)dt,
\]

where \( a > 0 \) is the rate of person-to-person transmission, \( b \geq 0 \) is the rate of recovery (or dying or removal), and \( c \geq 0 \) is the rate of transmission from an external source. We assume throughout that \( a > 0, b, c \geq 0 \) (if \( a = 0 \) then (4.1) is hardly a relevant epidemics model).

A simple parameter choice is \( b = c = 0 \). Let us discuss this case first. We have

\[
\frac{dx}{dt} = ax(1-x)dt.
\]

Recall that \( x \) is the fraction of infected individuals. The product \( x(1-x) \) describes the number of meetings between infected and non-infected individuals. The constant \( a > 0 \) describes the intensity of those meetings. Note that if \( x(0) \in (0,1) \) then the derivative \( x' \) of \( x \) is positive so \( x \) is increasing with limit 1 since \( x' = 0 \) at \( x = 1 \). Everyone at the end will be infected.

We can of course write the equation (4.1) as

\[
\frac{dx}{dt} = \mu(x)dt, \quad \mu(x) = ax(1-x)dt - bx + c(1-x)
\]

Without solving the equation we can get a qualitative behaviour of a solution of (4.2). For example if \( \mu(x) > 0 \) then \( x' = \mu(x) > 0 \), i.e. \( t \mapsto x(t) \) increases. If \( \mu(x) < 0 \) then \( x' = \mu(x) < 0 \), i.e. \( t \mapsto x(t) \) decreases. If \( \mu(x) = 0 \) then \( x' = \mu(x) = 0 \), i.e. \( t \mapsto x(t) \) is neither increasing nor decreasing, i.e. constant. So now we investigate the sign of \( \mu \). For that it is simplest to first look at its zeros, i.e. when \( \mu(x) = 0 \).

\[
\mu(x) = ax(1-x) - bx + c(1-x)
\]

\[
= -ax^2 + (a-b-c)x + c
\]

\[
= -a \left( x^2 - \frac{2(a-b-c)x}{2a} - \frac{c}{a} \right)
\]

\[
= -a \left( x^2 - 2dx - \frac{c}{a} \right), \quad d = \frac{a-b-c}{2a}
\]

\[
= -a(x-x_-)(x-x_+),
\]

where

\[
x_\pm = d \pm \sqrt{d^2 + \frac{c}{a}}
\]

We remark that \( \mu(x) < 0 \) if \( x < x_- \) or \( x > x_+ \) and \( \mu(x) > 0 \) if \( x_- < x < x_+ \). That means that a solution to (4.2) is decreasing if \( x \) is outside the interval \( (x_- , x_+) \) and increasing if \( x \in (x_- , x_+) \). Clearly

\[
x_- = d - \sqrt{d^2 + \frac{c}{a}} \leq 0,
\]
and

\[ x_+ = d + \sqrt{d^2 + \frac{c}{a}} \geq 0 \]

with equality iff (if and only if) \( d \leq 0 \) and \( c = 0 \), i.e. \( a \leq b \) and \( c = 0 \). Note

\[
d^2 + \frac{c}{a} = \left( \frac{a - b - c}{2a} \right)^2 + \frac{4ac}{4a^2} \]

\[
= (a^2 + b^2 + c^2 - 2ab - 2ac + 2bc + 4ac) \frac{1}{4a^2} \]

\[
= (a^2 + b^2 + c^2 - 2ab + 2ac + 2bc) \frac{1}{4a^2} \]

\[
= (a^2 + b^2 + c^2 + 2ab + 2ac + 2bc - 4ab) \frac{1}{4a^2} \]

\[
= ((a + b + c)^2 - 4ab) \frac{1}{4a^2} \]

\[
\leq \frac{1}{4a^2} (a + b + c)^2 \quad \text{(equality iff } b = 0) \]

where we in the iff statement used that we assumed \( a > 0 \). Hence

\[
x_+ = d + \sqrt{d^2 + \frac{c}{a}} = \frac{a - b - c}{2a} + \sqrt{\frac{1}{4a^2} ((a + b + c)^2 - 4ab)} \]

\[
\leq \frac{a - b - c}{2a} + \frac{1}{2a} |a + b + c| \]

\[
= \frac{a - b - c + (a + b + c)}{2a} = 1 \]

with equality iff \( b = 0 \). We thus conclude that \( 0 < x_+ < 1 \) iff \( b > 0 \) and the complement of \( \{a \leq b \text{ and } c = 0\} \), i.e. \( \{a > b \text{ or } c > 0\} \) is true. By the above discussion we obtain the following result.

**Result 4.1.** If \( a > 0, b, c \geq 0 \) and \( x_0 \in (0,1) \), then

\[
\lim_{t \to \infty} x_t = x_+ = d + \sqrt{d^2 + c/a}, \quad d = (a - b - c)/(2a). \]

Furthermore, \( 0 \leq x_+ \leq 1 \) and

\[
x_+ = 1 \iff b = 0 \]

\[
x_+ = 0 \iff a \leq b \text{ and } c = 0 \]

\[
0 < x_+ < 1 \iff a > b > 0 \text{ or } b, c > 0. \]

Result 4.1 was not presented in Iacus [2]. Result 4.1 means that in the long run all individuals will be infected iff (if and only if) \( b = 0 \). The disease will die out iff \( a < b \) and \( c = 0 \). The disease will live on forever among infected and non-infected individuals fraction iff \( a > b > 0 \) or \( b, c > 0 \). For illustration of Result 4.1, see Figures 4.1–4.5.
Figure 4.1: Solutions of (4.1) with different start values, $a = 1, b = 0, c = 0$. $b = 0$ makes the upper root $x_+$ of $\mu(x) = 0$ equal to one. That means $x' = \mu(x) > 0$ for $x \in (0, 1)$. Convergence to $x_+ = 1$.

Figure 4.2: Solutions of (4.1) with different start values, $a = 1, b = 0, c = 1$. $b = 0$ makes $x' = \mu(x) > 0$ for $x \in (0, 1)$, convergence to $x_+ = 1$. Here $c > 0$, which makes the convergence quicker than in 4.1 where $c$ was zero.

Figure 4.3: Solutions of (4.1) with different start values, $a = 1, b = 3, c = 0$. $a < b$ and $c = 0$ makes the upper root $x_+$ of $\mu(x) = 0$ equal to zero. That makes $x' = \mu(x) < 0$ for $x \in (0, 1)$. Convergence to $x_+ = 0$. 
Figure 4.4: Solutions of (4.1) with different start values, $a = 1$, $b = .1$, $c = 0$. $a > b > 0$ or $b, c > 0$. Convergence to $x_+ = 0.9$.

Figure 4.5: Solutions of (4.1) with different start values, $a = 1$, $b = .5$, $c = 0$. $a \geq b > 0$ or $\{ b > 0 \text{ and } c > 0 \}$ makes the upper root $x_+$ of $\mu(x) = 0$ be in the open interval $(0, 1)$, here $x_+ = 0.5$. That makes $x' = \mu(x) > 0$ for $x \in (0, 0.5)$ and $x' = \mu(x) < 0$ for $x \in (0.9, 1)$. Convergence to $x_+ = 0.5$. Here $b$ is larger than in 4.4, which makes the equilibrium point $x_+$ smaller.
4.2 Stochastic epidemics fraction model

In [2] there is suggested a stochastic version of (4.1):

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t, \quad (4.4)$$

where

$$\mu(x) = ax(1 - x) - bx + c(1 - x),$$

and

$$\sigma^2(x) = 2\varepsilon x(1 - x), \quad \varepsilon > 0$$

and $W$ is a Brownian motion.

Throughout in this subsection we let $\varepsilon > 0$. It is not clear from [2] that there exists a solution to (4.4) given initial condition $X_0$.

Let us discuss existence of a solution to (4.4) and some of its properties. For this we need some terminology. For simplification assume $X_0 = x_0$, a nonrandom point in the open interval $(0, 1)$. Following Proposition 5.22 and Definition 5.20 in [4], if $x_0 \in (1/n, 1 - 1/n)$, there exists at least what is called a weak solution up to a stopping time $T_n = \inf\{t \geq 0 : X_t \not\in (1/n, 1 - 1/n)\}$ for any $n \geq 2$. Let $T = \lim_{n \to \infty} T_n$. It means

$$T = \inf\{t \geq 0 : X_t = 0 \text{ or } X_t = 1\}.$$

First note that

$$\int_{x-\delta}^{x+\delta} \sigma^{-2}(y)(1 + |\mu(y)|)dy < \infty \quad (4.5)$$

for any $x \in (0, 1)$ and $\delta$ such that $(x - \delta, x + \delta) \in (0, 1)$. We also have

$$\sigma^2(x) > 0 \text{ for any } x \in (0, 1). \quad (4.6)$$

For a fixed $x^*$ in $(0, 1)$, let

$$s(x) = \exp\{-2\int_{x^*}^{x} \frac{\mu(y)}{\sigma^2(y)}dy\}$$

and

$$m(x) = \frac{1}{\sigma^2(x)s(x)}.$$

$s(x)$ is known as a so called scale measure (or scale function). $m(x)$ is known as a so called speed measure (or speed function).

Thanks to (4.5) and (4.6) we can apply a result in [4] (Proposition 5.22) here adapted to (4.4).

**Result 4.2.** Let $T = \inf\{t \geq 0 : X_t = 0 \text{ or } X_t = 1\}$.

(a) If

$$\int_{x^*}^{x} s(x)dx = \int_{x^*}^{1} s(x)dx = \infty$$

then

$$P(T = \infty) = P(\inf_{0 \leq t < \infty} X_t = 0) = P(\sup_{0 \leq t < \infty} X_t = 1) = 1.$$ 

In particular, $\{X_t : t \geq 0\}$ is recurrent: for any $x \in (0, 1)$:

$$P(X_t = x \text{ for some } t \in [0, \infty)) = 1.$$
(b) If
\[ \int_0^{x^*} s(x)dx < \infty = \int_{x^*}^1 s(x)dx \]
then
\[ P(\lim_{t\uparrow T} X_t = 0) = P(\sup_{0 \leq t < T} X_t < 1) = 1. \]

(c) If
\[ \int_0^{x^*} s(x)dx = \infty > \int_{x^*}^1 s(x)dx \]
then
\[ P(\inf_{0 \leq t < T} X_t > 0) = P(\lim_{t\uparrow T} X_t = 1) = 1. \]

(d) If
\[ \int_0^{x^*} s(x)dx < \infty \text{ and } \int_{x^*}^1 s(x)dx = \infty \]
then
\[ P(\lim_{t\uparrow T} X_t = 0) = 1 - P(\lim_{t\uparrow T} X_t = 1) = \frac{\int_0^1 s(x)dx}{\int_0^{x^*} s(x)dx}, \]

where \( X_0 = x_0 \).

In case (a) the process never touches the boundary.

We have
\[
\begin{align*}
s(x) &= \exp\left(-2 \int_{x^*}^x \frac{ay(1-y) - by + c(1-y)}{2\varepsilon y(1-y)} dy\right) \\
&= \exp\left(-\frac{1}{\varepsilon} \int_{x^*}^x a - \frac{b}{1-y} + \frac{c}{y} dy\right) \\
&= \exp\left(-\frac{a}{\varepsilon} (x - x^*) - \frac{b}{\varepsilon} (\ln(1-x) - \ln(1-x^*)) - \frac{c}{\varepsilon} (\ln x - \ln x^*)\right) \\
&= \exp\left(-\frac{a}{\varepsilon} (x - x^*)\right) \cdot \frac{1-x}{1-x^*} \cdot \left(\frac{x}{x^*}\right)^{-\frac{s}{\varepsilon}}
\end{align*}
\]

From this we easily see that
\[ \int_0^{x^*} s(x)dx \begin{cases} < \infty & \text{if } c/\varepsilon < 1 \\
= \infty & \text{if } c/\varepsilon \geq 1 \end{cases} \]
and
\[ \int_{x^*}^1 s(x)dx \begin{cases} < \infty & \text{if } b/\varepsilon < 1 \\
= \infty & \text{if } c/\varepsilon \geq 1 \end{cases} \]

To really see that, note that for example that for \( x \) close to zero, the critical factor of \( s(x) \) is \( x^{-c/\varepsilon} \) and \( \int_0^\delta x^{-c/\varepsilon} dx \) is finite for small \( \delta > 0 \) if and only if \( c/\varepsilon < 1 \).

Hence we obtain by result 4.2

Result 4.3. Let \( T = \inf\{t \geq 0 : X_t = 0 \text{ or } X_t = 1\} \).

(a) If \( \varepsilon \leq c \) and \( \varepsilon \leq b \), then
\[ P(T = \infty) = P(\inf_{0 \leq t < \infty} X_t = 0) = P(\sup_{0 \leq t < \infty} X_t = 1) = 1. \]
In particular, \( \{X_t : t \geq 0\} \) is recurrent: for any \( x \in (0, 1) \),

\[
P(X_t = x \text{ for some } t \in [0, \infty)) = 1.
\]

(b) If \( \varepsilon > c \) and \( \varepsilon \leq b \), then

\[
P(\lim_{t \uparrow T} X_t = 0) = P(\sup_{0 \leq t < T} X_t < 1) = 1.
\]

(c) If \( \varepsilon \leq c \) and \( \varepsilon > b \), then

\[
P(\inf_{0 \leq t < T} X_t > 0) = P(\lim_{t \uparrow T} X_t = 1) = 1.
\]

(d) If \( \varepsilon > c \) and \( \varepsilon > b \), then

\[
P(\lim_{t \uparrow T} X_t = 0) = 1 - P(\lim_{t \uparrow T} X_t = 1) = \frac{\int_0^1 s(x)dx}{\int_0^1 s(x)dx} - \frac{\int_0^1 \exp\left(-\frac{a}{\varepsilon}x\right)(1-x)^{-b/\varepsilon} x^{-c/\varepsilon} dx}{\int_0^1 \exp\left(-\frac{a}{\varepsilon}x\right)(1-x)^{-b/\varepsilon} x^{-c/\varepsilon} dx},
\]

For the case the (a) there is a limiting distribution if \( \int_0^1 m(x)dx < \infty \) according to the following fact, [4], Exercise 5.40. See also [2].

**Fact 4.4.** If

\[
\int_0^{x^*} s(x)dx = \int_0^1 s(x)dx = \infty
\]

and \( \int_0^1 m(x)dx < \infty \) then \( X_t \) converges to a random variable as \( t \to \infty \) with the distribution

\[
\pi(x) = \frac{m(x)}{\int_0^1 m(y)dy}.
\]

In our situation,

\[
m(x) = \frac{1}{2\varepsilon x(1-x)} \exp\left(\frac{1}{\varepsilon} \left( \int_{x^*}^{x} a - \frac{b}{1-u} + \frac{c}{u} du \right) \right)
\]

\[
= \frac{1}{2\varepsilon x(1-x)} \exp\left(-\frac{a}{\varepsilon}(x-x^*)\left(\frac{1-x}{1-x^*}\right) - \frac{b}{\varepsilon} \left(\frac{x}{x^*}\right)^{-\frac{b}{\varepsilon}}\right)
\]

\[
= c^* \exp\left(\frac{a}{\varepsilon}x\right)(1-x)^{-\frac{b}{\varepsilon}} x_{-1}^{-1}
\]

where \( c^* = c(a, b, c, \varepsilon) \) is a coefficient depending on \( a, b, c, x^* \) but not on \( x \). We see easily that \( \int_0^1 m(x)dx < \infty \) if and only if \( \frac{c^*}{\varepsilon} > -1 > -1 \) and \( \frac{b}{\varepsilon} > -1 > -1 \), i.e.

\[
\int_0^1 m(x)dx < \infty \iff c > 0 \text{ and } b > 0
\]

We obtain from Result 4.3 and Fact 4.4

**Result 4.5.** If \( \varepsilon \leq c \) and \( \varepsilon \leq b \) then there exists a limiting density function

\[
\pi(x) = c \exp\left(\frac{a}{\varepsilon}x\right)(1-x)^{-\frac{b}{\varepsilon}} x^{-1}, \quad (4.7)
\]

where \( 1/c = \int_0^1 \exp\left(\frac{a}{\varepsilon}x\right)(1-x)^{-\frac{b}{\varepsilon}} x^{-1} dx < \infty \).
Remark 4.6. Note that in Results 4.2 and 4.3 the process is killed at the hitting time $T$ of $\{0\}$ or $\{1\}$. By [2], even if we are not in case (a) there exists a limiting distribution for the non-killed process, which has the form (4.7). In fact, after the hitting time the boundaries are instantaneously reflecting. More precise, for case (b) in Result 4.3, $\{0\}$ is an instanteously reflecting boundary (and $\{1\}$ never reached), in case (c) $\{1\}$ is an instantenously reflecting boundary (and $\{0\}$ never reached), in case (d) both $\{0\}$ and $\{1\}$ are instanteously reflecting boundaries. Possibly, what happens after the hitting time is not that relevant for describing epidemics.

In Figures 4.6 - 4.9 are simulated some selected paths of (4.4). Due to Result 4.3, it seems appropriate to apply a 'killed' Euler-Maruyama scheme, that means a scheme such that when the Euler method go outside $[0,1]$, it is stopped. For a partition $0 = t_0 < t_1 < \ldots < t_n = T$, $X_i = X_{t_i}$, $\Delta_i = t_{i+1} - t_i$, the scheme is simply $X_0 = x_0$, and, given $X_i$, it works simply as follows: if $0 < X_i + \mu (X_i) \Delta_i + \sigma (X_i) \sqrt{\Delta_i} N(0,1) < 1$ then $X_{i+1} = X_i + \mu (X_i) \Delta_i + \sigma (X_i) \sqrt{\Delta_i} N(0,1)$. If $X_i + \mu (X_i) \Delta_i + \sigma (X_i) \sqrt{\Delta_i} N(0,1) \geq 1$, then $X_{i+1} = X_{i+2} = \ldots = 1$. If $X_i + \mu (X_i) \Delta_i + \sigma (X_i) \sqrt{\Delta_i} N(0,1) \leq 0$, then $X_{i+1} = X_{i+2} = \ldots = 0$.

One interesting question is: are there more appropriate one-step numerical method than this 'killed' scheme?

Figure 4.6: Case (a), $\varepsilon \leq c$ and $\varepsilon \leq b$. Here $a = 1$, $b = 1$, $c = 1$, $\varepsilon = 0.5$. $X_t$ never hits the boundaries.

Figure 4.7: Case (b), $\varepsilon > c$ and $\varepsilon \leq b$. Here $a = 1$, $b = 2$, $c = 1$, $\varepsilon = 1.5$. The first hit boundary is 0.
Figure 4.8: Case (c), $\varepsilon \leq c$ and $\varepsilon > b$. Here $a = 1$, $b = 1$, $c = 2$, $\varepsilon = 1.5$. The first hit boundary is 1.

Figure 4.9: Case (d), $\varepsilon \geq b$, $c$. $a = 1$, $b = 0.2$, $c = 0.1$, $\varepsilon = 0.6$. It first hits 0 with a certain probability strictly between 0 and 1. Some of the paths actually hits 0 first, the other ones hits 1 first.
4.2.1 Bimodality

From [2], it is claimed that "this distribution may be bimodal depending on the number of real solutions to \( d \pm \sqrt{d^2 - \frac{c - \varepsilon}{a}} \), where \( d = (a - b - c + 2\varepsilon)/(2a) \). When there are two positive real solutions, the smaller is the epidemic threshold and the larger is the size of the epidemic. This is a model of stochastic threshold when \( 0 < c < \varepsilon \). The epidemic is unlikely to happen if the model (the number of infected people in the population) stays below the lowest mode. When the model is above it the epidemic size is around the second mode. When \( c > \varepsilon \) (high infection rate from external sources), the epidemic is guaranteed."

The last statement on \( c > \varepsilon \) appears to be doubtful in the light of Result 4.3.

Let us investigate the bimodality. Bimodality of a density means just that there are points \( x_- \) and \( x_+ \) for which \( \pi'(x) = d\pi(x)/dx = 0 \). One way to investigate \( \pi' \) is to note that

\[
(\pi\sigma^2)' = 2\mu \pi
\]
i.e.

\[
\pi' = \frac{(2\mu - (\sigma^2)')\pi}{\sigma^2}.
\]

where

\[
2\alpha(1-x) - 2bx + 2c(1-x) - 2\varepsilon(1-x) + 2\varepsilon x = -2ax^2 + 2(a - b - c + 2\varepsilon)x + 2(c - \varepsilon) = -2a(x - x_-)(x - x_+)
\]

where

\[
x_\pm = d \pm \sqrt{d^2 - \frac{c - \varepsilon}{a}}, \quad d = \frac{a - b - c + 2\varepsilon}{2a} \quad (4.8)
\]

For convenience we may write

\[
x_\pm = d \pm \sqrt{d^2 - \left(\xi - \gamma\right)}, \quad d = \frac{1 - \beta - \gamma + 2\xi}{2}
\]

where \( \beta = \frac{b}{2}, \gamma = \frac{c}{a}, \xi = \frac{\varepsilon}{a} \). By elementary calculations we see that \( x_\pm \) are real if and only if

\[
(\xi - \frac{\beta + \gamma}{2})^2 \geq \frac{1}{4} + \frac{1}{2}(\beta - \gamma)
\]
i.e.

\[
(\varepsilon - \frac{\beta + c}{2})^2 \geq \frac{1}{2}((b - c) - \frac{a}{2}) \quad (4.9)
\]

It is interesting to note that given \( b, c \) and \( \varepsilon \) we get a simple condition for \( a \), namely

\[
a \geq 2(b - c) - 4(\varepsilon - \frac{b + c}{2})^2.
\]

If \( a, b, \) and \( c \) are given, we get a condition for \( \varepsilon \):

**Result 4.7.** Sufficient and necessary conditions for \( x_\pm \) given by (4.8) to be real:

If \( b - c \leq \frac{a}{2} \) then \( \varepsilon > 0 \).

If \( b - c > \frac{a}{2} \) then \( |\varepsilon - \frac{b + c}{2}| \geq \frac{1}{2}\sqrt{(b - c) - \frac{a}{2}} \).

If \( x_- \) and \( x_+ \) are real then they are both positive iff (if and only if)

\[
x_- > 0 \iff d - \sqrt{d^2 - (\xi - \gamma)} > 0 \iff d > \sqrt{d^2 - (\xi - \gamma)} \iff d > 0 \text{ and } \xi > \gamma
\]

\[
\iff \frac{1 - \beta - \gamma + 2\xi}{2} > 0 \text{ and } \xi > \eta \iff \frac{a - b - c + 2\varepsilon}{2a} > 0 \text{ and } \varepsilon > c \iff a - b - c + 2\varepsilon > 0 \text{ and } \varepsilon > c \iff \varepsilon > (b + c - a)/2 \text{ and } \varepsilon > c.
\]
If \( x_\pm \) real then \( x_- \leq x_+ < 1 \) iff \( x_+ < 1 \). Now,

\[
x_+ = d + \sqrt{d^2 - (\xi - \gamma)} = \frac{1 - \beta - \gamma + 2\xi}{2} + \sqrt{\left(\frac{1 - \beta - \gamma + 2\xi}{2}\right)^2 - (\xi - \gamma)}
\]

\[
= \frac{1 - \beta - \gamma + 2\xi}{2} + \frac{1}{2} \sqrt{(1 - \beta - \gamma + 2\xi)^2 - 4(\xi - \gamma)}
\]

\[
= \frac{1 - (\beta + \gamma - 2\xi)}{2} + \frac{1}{2} \sqrt{(1 - (\beta + \gamma - 2\xi))^2 - 4(\xi - \gamma))}
\]

\[
= \frac{1 - (\beta + \gamma - 2\xi)}{2} + \frac{1}{2} \sqrt{(1 + (\beta + \gamma - 2\xi)^2 - 4(\xi - \gamma) - 4(\beta - \gamma - 2\xi))}
\]

\[
= \frac{1 - (\beta + \gamma - 2\xi)}{2} + \frac{1}{2} \sqrt{(1 + (\beta + \gamma - 2\xi)^2 + 4(\xi - \beta))}
\]

If \( x_+ \) real, and \( \beta + \gamma - 2\xi \leq -1 \) then

\[
x_+ \geq \frac{1}{2}(1 - (\beta + \gamma - 2\xi)) \geq 1
\]

If \( \beta + \gamma - 2\xi > -1 \) and \( \xi < \beta \) then

\[
x_+ < \frac{1}{2}(1 - (\beta + \gamma - 2\xi)) + \frac{1}{2}|1 + (\beta + \gamma - 2\xi)|
\]

\[
= 12(1 - (\beta + \gamma - 2\xi)) + \frac{1}{2}(1 + (\beta + \gamma - 2\xi)) = 1
\]

If \( \beta + \gamma - 2\xi > -1 \) and \( \xi = \beta \) then \( x_+ = 1. \)
If \( \beta + \gamma - 2\xi > -1 \) and \( \xi > \beta \) then

\[
x_+ \geq \frac{1}{2}(1 - (\beta + \gamma - 2\xi)) + \frac{1}{2}|1 + (\beta + \gamma - 2\xi)| = 1.
\]

Hence, if \( x_+ \) real then \( x_+ < 1 \) iff \( \beta + \gamma - 2\xi > -1 \) and \( \xi < \beta \) i.e. \( a + b + c - 2\varepsilon > 0 \) and \( \varepsilon < b \) i.e. \( \varepsilon < (a + b + c)/2 \) and \( \varepsilon < b \).

We summarize in the following result:

**Result 4.8.** For the limiting distribution \( \pi, \pi' = 0 \) has roots \( x_\pm \) given by (4.8) which may be complex. The roots are real iff the conditions in 4.7 are satisfied.

The roots are positive iff furthermore \( \varepsilon > (b + c - a)/2 \) and \( \varepsilon > c \). The roots are smaller than one iff furthermore \( \varepsilon < (a + b + c)/2 \) and \( \varepsilon < b \).

Concerning bimodality we get:

**Result 4.9.** The limiting distribution \( \pi \) is bimodal iff \( a, b, c \) and \( \varepsilon \) are such that the conditions in 4.7 are satisfied and

\[
\max\left(\frac{b + c - a}{2}, c\right) < \varepsilon < \min\left(\frac{a + b + c}{2}, b\right)
\]

Note that if the inequalities in (4.9) are satisfied, then we are in the case (b).
In Figure 4.9 there is one simulated case with bimodality.
Figure 4.10: Another case of (b), $\epsilon > c$ and $\epsilon \leq b$, for which the limiting distribution $\pi$ is bimodal. The parameters are chosen so that we get bimodality. Here $a = .356$, $b = 0.2$, $c = 0.1$, $\epsilon = 0.15$. Here the process is not killed at the first hitting time.

Figure 4.11: The limiting distribution $\pi$ for the process in Figure 4.10. Note its bimodality, i.e. there are two points for which $\pi$ is flat.
A Appendix

In this Appendix the codes for producing the figures are presented, ordered by the sub-section for which the corresponding models appear.

A.1 Classical epidemics models

A.1.1 Classical deterministic epidemics model

```matlab
% clear, epidemics
x0=2;
y0=1;
N=1;
n=1000;
dt=1/(n-1);
T=50;
t=linspace (0,T,n);
x=x0*ones (1,N);
y=y0*ones (1,N);

beta=1;
gamma=.2;
r=0;
dW=sqrt (dt)*randn (2,n);
for i=1:n-1
    x1=x (i,:); x2=y (i,:);
    A=[x1*(r-beta*x2);x2*(-gamma+beta*x1)];
    B=[r*x1+beta*x1*x2, -beta*x1*x2;-beta*x1*x2, 
        beta*x1*x2+gamma*x2];
    C=zeros (2,2);% sqrtm (B); %C zeros makes it deterministic
    dW=sqrt (dt)*randn (2,1);
    x (i+1,:)=(1,:)*A+C (1,:) *dW;
    y (i+1,:)=(, :)*A+C (2,:) *dW;
end
plot (t,x,t,y)
display ('press button'), pause
plot (t,y./(x+y))
```

A.1.2 Classical epidemics model, here with noise

```matlab
% clear, epidemics
x0=2;
y0=1;
N=1;
n=1000;
dt=1/(n-1);
T=50;
t=linspace (0,T,n);
x=x0*ones (1,N);
y=y0*ones (1,N);

beta=1;
gamma=.2;
r=0;
dW=sqrt (dt)*randn (2,n);
for i=1:n-1
    x1=x (i,:); x2=y (i,:);
    A=[x1*(r-beta*x2);x2*(-gamma+beta*x1)];
    B=[r*x1+beta*x1*x2, -beta*x1*x2;-beta*x1*x2, 
        beta*x1*x2+gamma*x2];
    C=zeros (2,2);% sqrtm (B); %C zeros makes it deterministic
    dW=sqrt (dt)*randn (2,1);
    x (i+1,:)=(1,:)*A+C (1,:) *dW;
    y (i+1,:)=(, :)*A+C (2,:) *dW;
end
plot (t,x,t,y)
display ('press button'), pause
plot (t,y./(x+y))
```
\[ r=0; \]
\[ dW = \sqrt(dt) \cdot \text{randn}(2,n); \]
\[ \text{for } i=1:n-1 \]
\[ x1=x(i,:); \ x2=y(i,:); \]
\[ A=[x1*(r-beta*x2); x2*(-gamma+beta*x1)]; \]
\[ B=[r*x1+beta*x1*x2, -beta*x1*x2; -beta*x1*x2, beta*x1*x2+gamma*x2]; \]
\[ C=\text{sqrtm}(B); % \text{C is now stochastic} \]
\[ dW = \sqrt(dt) \cdot \text{randn}(2,1); \]
\[ x(i+1,:) = x1 + A(1,:)*dt + C(1,:)*dW; \]
\[ y(i+1,:) = x2 + A(2,:)*dt + C(2,:)*dW; \]
\[ \text{end} \]
\[ \text{plot}(t,x,t,y) \]
\[ \text{display ('press button')}, \text{pause} \]
\[ \text{plot}(t,y./(x+y)) \]

**A.2 Code for fraction models**

**A.2.1 Code for the deterministic fraction models**

clear, clf

\[ a=1; b=0; c=0; \]
\[ epsilon=0; \]
\[ T=20; \]
\[ d=(a-b-c+2*epsilon)/(2*a); \]
\[ xm=d-sqrt(d^2-(epsilon-c)/a); \]
\[ xp=d+sqrt(d^2-(epsilon-c)/a); \]
\[ [xm \ xp] \]
\[ xdot=@(t,x)a*x.*(1-x)-b*x+c*(1-x); \]
\[ x0=linspace(0,1,10); \]
\[ \text{for } i=1:length(x0), \]
\[ [ti \ xi]=\text{ode23}(xdot,[0 \ T],x0(i)); \]
\[ \text{plot}(ti,xi), \text{hold on} \]
\[ \text{end} \]
\[ \text{hold off} \]

**A.2.2 Code for the stochastic fraction model**

% clear
% clf
\[ c=.1;b=.2; \]
\[ epsilon=c+.5*(b-c); \]
\[ areal=b-c+2*sqrt(abs(epsilon^2-(b-c)^2)); \]
\[ aint=abs(b+c-2*eps); \]
\[ [areal \ aint] \]
\[ a=max(areal,aint)*1.1; \]
\[ [a \ b \ c \ epsilon] \]
\[ d=(a-b-c+2*epsilon)/(2*a); \]
\[ [d-sqrt(d^2-(epsilon-c)/a) \ d+sqrt(d^2-(epsilon-c)/a)] \]
\[ mu=@(x)a*x.*(1-x)-b*x+c*(1-x); \]
\[ sigma=@(x)sqrt(2*epsilon*max(x.*(1-x),0)); \]
N=50;
x0=linspace (0,1,N);
n=2000;
T=10;
t=linspace (0,T,n);
dt=T/(n-1);
x=x0;
mu=@(x)a*x.*(1-x)-b*x+c*(1-x);
sigma=@(x)sqrt (2*epsilon*max (x.*(1-x),0));
dW=sqrt (dt)*randn (n,1);
for i=1:n-1
    x (i+1,:)=x (i,:)+mu (x (i,:))*dt+sigma (x (i,:)).*dW (i,:);
end
y=[x;ones (1,N)];
ind=zeros (1,N);
for j=1:N
    ind (j)=find (((y (:,j)>=1) | (y (:,j)<=0),1,'first');
end
z=zeros (n,N); %z killed Euler
for j=1:N
    z (:,j)=[y (1:ind (j)-1,j);(y (ind (j),j)>=1)*ones (n+1-ind(j),1)];
end
%subplot (211), plot (t,x)
%subplot (211), plot (t,z)
%histogramfn (z (n,:))
%subplot (211), plot (x0,x (n,:),'.',x0,z (n,:),'.')
ox=0.5;
prepi=@(x)1./(x.*(1-x)).*exp (1/epsilon*(a*(x-xo)+b*(log (1-x)-log...
   (1-xo)))+c*(log (x)-log (xo)));
M=1/quad (prepi,0.01,0.99);
pidensity=@(x)M*prepi (x);
xp=linspace (0.001,0.999);
%subplot (212), plot (xp,pidensity (xp))

References
   simuleringsstudie ur matematiskt och datalogiskt perspektiv, Student thesis, Växjö
   University.