Modelling logistic growth by a new diffusion process: Application to biological systems

Patricia Román-Román, Francisco Torres-Ruiz *

Departamento de Estadística e Investigación Operativa, Universidad de Granada, Spain

ABSTRACT

The present paper introduces a new diffusion process for the purpose of modelling logistic-type behaviour patterns. Unlike other processes in the same context, this one verifies that its mean function is a logistic curve. In addition, its transition density can be found explicitly, which allows to analyse inference from the discrete sampling of trajectories. The main features of the process will be analysed and the maximum likelihood estimation of parameters will be carried out through discrete sampling. Regarding the numerical problems found to solve the likelihood equations, several strategies are developed for obtaining initial solutions for the usual numerical procedures. Such strategies are compared by means of a simulation example. Also, another simulation study is carried out in order to compare the estimation in this process to that developed by means of continuous sampling in the logistic diffusion model considered by Giovanis and Skiadas (1999). Finally an example is given for the growth of a microorganism culture. This example illustrates the predictive possibilities of the new process, as well as its ability to study time variables formulated as first-passage-times.

© 2012 Elsevier Ireland Ltd. All rights reserved.

1. Introduction

For a long time the study of phenomena associated to growth curves has drawn the attention of researchers. A wide range of curves have been applied to a variety of scientific areas. Out of them, the logistic curve has deserved special attention and has been the genesis of many of the curves that appeared later (see Birch, 1999; Tsoularis and Wallace, 2002).

The logistic curve was introduced by Verhulst in the XIX century, for the purpose of studying population growth. In the 1920s the interest was rekindled and many pieces of research have focused on it since then, due to the fact that it is an excellent model for the development and evolution of many growth phenomena, even beyond those that gave it rise.

There are indeed many phenomena in nature exhibiting a fast (virtually exponential) initial growth that then slow down after a certain point (where the curve reaches an inflection point) until a point of equilibrium or saturation of the system (*carrying capacity*).

The origins of the logistic model can be found in Ecology, where it served the purpose of explaining population growth, but it has also been used in Demography, and in Biology and Medicine for the analysis of the growth of bacteria, tumours, and several species of animals and plants (see Tsoularis and Wallace, 2002 and references therein). In the last years, its applications have been extended, and the curve has even been employed in Economy for the illustration of how innovation spreads (Giovanis and Skiadas, 2007); Gallagher (2011) uses the curve for the analysis of the global oil production and to locate production peak; Yokohama and Sanada (2009) use it to predict linguistic changes and variations; Modis (2011) regards it as a competition model and applies it to the process of granting Nobel prizes. Gámez et al. (2009) consider a logistic growth in order to describe the interaction between the populations of healthy and irradiated cells of an organ under the effect of a constant radiation. Several variations are currently under study, as in the case of the double logistic curve (see Lipovetsky, 2010).

As is well known, the deterministic logistic model is defined in terms of the differential equation

$$\frac{dx(t)}{dt} = \alpha x(t) - \beta x^2(t)$$
$$x(t_0) = x_0,$$

(1)

where α, β are positive constants. Constant α defines the growth rate, whereas the term $-\beta x^2$ serves to inhibit or retard this rate. In this sense, β is usually smaller than α. When the population $x = x(t)$ is small, this term has little effect on the value of $x' = dx(t)/dt$ and so the population starts off with almost exponential growth. However, as x increases, the inhibiting term eventually slows the rate of growth dramatically.

* Corresponding author at: Departamento de Estadística e Investigación Operativa, Facultad de Ciencias, Universidad de Granada, Avda. Fuentenueva s/n, 18071 Granada, Spain. Tel.: +34 958241000x20056; fax: +34 958243267.
  E-mail address: fdeasis@ugr.es (F. Torres-Ruiz).

Despite the model (1) having been widely used, it does not account for possible random influences affecting the system under study. As stated by Li et al. (2011), "*the noises are ubiquitous in both nature and human society; for instance environmental fluctuations, a lack of precision of measurements. In any case, one has to deal with effects of randomness on the model.*" It is therefore necessary to modify the deterministic model in order to introduce the random element, and employ stochastic differential equations (SDE).

Several alternatives have been proposed for the stochastic modification of the equation (1). Most of these add to the differential equation a random term made of the variable under study plus a white noise:

$$dX(t) = (\alpha X(t) - \beta X^2(t)) dt + f(X(t)) dW(t), \qquad (2)$$

where $W(t)$ represents the standard Wiener process.

The first to propose such formulation was Feller in 1939, who considered $f(x) = c\sqrt{x}$, $c > 0$. Nevertheless, the Kolmogorov backward and forward equations associated to the diffusion process solution of (2) have never been solved. Tuckwell and Koziol (1987) show a summary of some diffusion processes generated by modifying the previous function, some of which are linked to specific applications like Demography (Artzrouni and Reneke, 1990), or energy consumption (Giovanis and Skiadas, 1999). Also, Schurz (2007) considers a more general version for the stochastic differential equation associated to logistic growth, and some versions of such equations have been recently introduced, which include delays (see Rupsys, 2008).

It must be noted that several of the works published in this field find the following obstacle: although the SDE may have a solution (sometimes not an explicit one), the resulting diffusion process is difficult to handle because the associated Kolmogorov equations have not been solved. This precludes the use of the transition density functions, and also to address some problems like the first-passage-time densities through time-variable boundaries (see Gutiérrez et al. (1997)). Even more so, in the few cases where the transition density has been calculated (Tuckwell, 1974), it has been impossible to find an explicit expression for the expected value of the solution for the SDE. For this reason, many of the works cited above deal with such processes from an overly theoretical perspective, analysing the existence of a solution for the SDE, dealing with the stability of the solution (see Sun and Wang, 2008), or simulating the probability distribution (see Hu and Wang, 2011). For the cases where an application has been made to some real phenomenon, the estimation procedures have been based on derivations of the SDE, either by using continuous sampling or through the discretisation of the differential equation. Giovanis and Skiadas (1999) considered the maximum likelihood estimation from the continuous sampling of paths, and even Bayesian procedures have been used, as in Tang and Heron (2008).

In this paper we propose a new diffusion process associated with the logistic curve that does not present the previously described obstacles. For this process, whose mean function will be a logistic curve, it is possible to explicitly determine the transition density function, which allows to tackle questions like the estimation of parameters through discrete sampling of paths after finding the likelihood from the transition density functions and the initial distribution. Such improvement enables the researcher to calculate the first-passage-time density, which requires an explicit form of such densities.

The present paper is structured as follows. Section 2 introduces a new reparameterisation of the logistic curve with significant differences to the ones usually employed, since it verifies that the upper bound is dependent on the initial value. This can be particularly useful in real-life situations in which information is available on the growth of specific individuals, each exhibiting a logistic behaviour, and with a limit value dependent on the initial one, and it is

necessary to develop a common model (as, for example, to determine the maximum growth each individual will reach). From such curve, in Section 3 a new diffusion process is developed with a mean function (which can be explicitly calculated) that fits the curve. Also, the distribution of the process will be found, which makes it possible to find the transition densities and the main characteristics of the process. Section 4 deals with the maximum likelihood estimation of the parameters of the process by means of discrete sampling of the paths. The complexity of likelihood equations imposes the use of numerical procedures which require, in turn, initial solutions. In order to find these, three alternatives are proposed, and a simulation is carried out to determine the most appropriate one according to the variability of the process and the sample size. Also, this section includes a simulation example in order to compare the estimation in this process to that developed by means of continuous sampling in the logistic diffusion model considered by Giovanis and Skiadas (1999), a process which also models logistic behaviour patterns, but for which inference is dealt with by means of continuous sampling since there was no explicit form for the transition densities. The results show how the process herein proposed provides a better estimation as the variability increases and the sample size decreases. Finally, in Section 5, an application to study the growth of microorganisms in a culture is presented in order to show how the new logistic process can be applied to practical matters.

2. A reformulation of the logistic curve

Before presenting the new reparameterisation of the logistic curve, let us summarise some interesting aspects of the most commonly used expression. This expression is found as the solution of the differential equation (1) with results

$$x(t) = \frac{\alpha/\beta}{1 + (((\alpha/\beta)/x_0) - 1)e^{-\alpha(t-t_0)}}, \quad t \geq t_0. \qquad (3)$$

This curve verifies that:

- It is increasing if and only if $\alpha > \beta x_0$.
- $\lim_{t \to \infty} x(t) = \alpha/\beta$.
- The curve exhibits an inflection point if and only if $\alpha > \beta$, in which case it occurs at

$$t_I = t_0 + \frac{1}{\alpha} \ln \left(\frac{\alpha/\beta}{x_0} - 1 \right).$$

Note that the limit value of the logistic curve is independent of x_0. Nevertheless, there may be practical situations in which this must not be taken for granted, as when a growth phenomenon is being studied and there are data from several individuals with a common logistic behaviour but different initial and limit values (as the weight or length of several individuals of the same animal species). For this reason, it would be useful to find a new expression of the curve which makes it possible to find a limit value that is dependent on the initial value.

To this end, we will start with the most general expression of the curve (3), namely

$$f(t) = \frac{a}{1 + be^{-ct}}, \quad a, b, c > 0.$$

If we assume that $f(t_0) = x_0 > 0$, then $a = x_0(1 + b e^{-ct_0})$, from where the new expression for the logistic curve is deduced:

$$f(t) = x_0 \frac{1 + b e^{-ct_0}}{1 + b e^{-ct}}, \quad t \geq t_0; b, c > 0, \qquad (4)$$