A Note on Estimation of Multi-Sigmoidal Gompertz Functions with Random Noise

Patricia Román-Román [1,2,*,†] ![ORCID], **Juan José Serrano-Pérez** [1,†] ![ORCID] **and Francisco Torres-Ruiz** [1,2,†] ![ORCID]

1   Departamento de Estadística e Investigación Operativa, Facultad de Ciencias, Universidad de Granada, Avenida Fuente Nueva s/n, Granada 18071, Spain; jjserra@ugr.es (J.J.S.-P.); fdeasis@ugr.es (F.T.-R.)
2   Instituto de Matemáticas de la Universidad de Granada (IEMath-GR), Calle Ventanilla 11, 18001 Granada, Spain
*   Correspondence: proman@ugr.es; Tel.: +34-958-240-491
†   These authors contributed equally to this work.

Abstract: The behaviour of many dynamic real phenomena shows different phases, with each one following a sigmoidal type pattern. This requires studying sigmoidal curves with more than one inflection point. In this work, a diffusion process is introduced whose mean function is a curve of this type, concretely a transformation of the well-known Gompertz model after introducing in its expression a polynomial term. The maximum likelihood estimation of the parameters of the model is studied, and various criteria are provided for the selection of the degree of the polynomial when real situations are addressed. Finally, some simulated examples are presented.

Keywords: multi-sigmoidal growth curves; diffusion processes; maximum likelihood estimation; model selection

1. Introduction

Growth curves with sigmoidal behaviour are widely used for data analysis across several fields of application. Given the variety of approaches employed in the profuse literature dealing with this subject, several mathematical models have been proposed for their study.

In general terms, a sigmoidal function is a function defined in the real line, bounded and differentiable with positive derivative. Its graph has a typical *S* shape showing a slow growth at the beginning, followed by a fast (exponential) growth that slows down gradually until it reaches an equilibrium value (usually named *carrying capacity* or level of saturation). Often, the sigmoid function refers to the particular case of the logistic function $f(t) = k/(1 + be^{-rt})$, although there is a great variety of curves with these characteristics. One of them is the Gompertz function, which is used in the modeling of systems that are saturated for large values of t, and whose most general expression is $f(t) = a \exp(-b \exp(-ct))$.

New sigmoidal curves have been introduced over the years, and their application has extended to several new fields. Among them we may mention hyperbolic curves (Eby et al. [1]), sigmoidal hyperbolic functions of the first and second kind (Menon et al. [2]), or the beta growth function (Yin et al. [3]). Regarding their application to new fields, researchers have looked into the diffusion of innovations (Giovanis and Skiadas [4]); the calculation of oil production peaks (Gallagher [5]); predicting changes in language (Yokohama and Sanada [6]); analyzing fatigue and fractures in materials and structures (Paolino and Cavatorta [7]); etc.

Traditionally, most of the sigmoidal growth models cited above have arisen from the solution of ordinary differential equations. In that sense, they are deterministic and do not include other information than that provided by the variable under study. In order to incorporate these influences, the idea of growth in a random environment emerged (see Ricciardi, [8] and references therein).

Thus, the so-called dynamic growth models appeared, and among them diffusion processes. Some of these diffusion models emerge as solutions to a stochastic differential equation after modifying a deterministic one by introducing in it a term of white noise. Other diffusion processes are constructed in such a way that their mean function is a certain sigmoidal growth curve. Usually, those of the first type retain the name of the corresponding deterministic equation of origin. For instance, Schurz [9], collect a wide variety of logistic diffusion processes. However, for most of them the stochastic differential equation does not have an explicit solution. For this reason, Román and Torres [10] constructed a logistic-type stochastic differential equation in the second sense (its mean is a logistic function). The same situation has presented itself in the context of other growth curves and related diffusion processes. Such is the case of the Gompertz process, which was introduced by [11]. In such process, the upper limit of the curve is independent of the initial value of the population under study, something not always verified in real situations. For this reason, Gutiérrez et al. [12] introduced a new Gompertz-type process in which the *carrying capacity* of the system depends on the initial state. This line of action has also been applied to the Bertalanffy curve (see Quiming et al. [13] and Román et al. [14]), the Hubbert curve (Luz-Sant'Ana et al. [15]), and, more recently, to the hyperbolastic curve of type I (Barrera et al. [16]).

Dynamic models are used in the fields in which the deterministic case has proved to be useful in fitting sigmoidal behaviour patterns to observed data. Researchers have developed several methods of estimation for these dynamic models. As far as maximum likelihood estimation methods are concerned, we can differentiate between those that take as a starting point the stochastic differential equation related to the model (usually known as continuous sampling methods) from those who build the likelihood function from the transition density functions of the process (discrete sampling methods). Alternatively, some authors have dealt with inference from a Bayesian perspective (Tang and Heron [17]).

An interesting aspect in this type of processes is the possibility of introducing, into their infinitesimal moments, time functions that allow us to regulate the evolution of the variable under study. Given that the functional form of such functions is not known, several strategies have been devised for their estimation. Some work carried out along this line includes studies by Albano et al. [18,19] and Román et al. [20] which centered on modifications of the Gompertz process.

There are multiple real situations in which the maximum level of growth is reached after successive stages, in each of which there is a deceleration followed by an explosion of the exponential type. For this reason, the use of sigmoidal curves with more than one inflection point is a good approach. A typical example of this behaviour is observed in the growth of various fruit species, such as stone fruits (Álvarez and Boché [21]). Cairns et al. [22] used double-sigmoidal models to study fatigue profiles in mouse muscles, while Amorim et al. [23] detected this type of behaviour in the different phases in which the fungus Ustilago Scitaminea Sydow infects the sugarcane and produces its characteristic smut.

The way that this multi-sigmoidal behaviour is modeled is far from unique. For example, Roper [24] used hyperbolic functions to study the transition between various temperature states in certain geological zones. Other authors have achieved this goal by including terms that define inflection points and additional parameters (Lipovetsky [25]). However, these models have not addressed the incorporation of external information to the variable under study, similarly to the dynamic models already mentioned. In this paper we address this problem by introducing a diffusion process whose mean obeys a pattern of multi-sigmoidal behaviour. In particular, we will deal with the case of Gompertz growth with multiple inflection points, following the idea mentioned in [23] for the case of the generalized monomolecular and Gompertz curves.

The rest of the paper is organized as follows: in Section 2 the multi-sigmoidal Gompertz curve is introduced by including a polynomial in the usual expression of the curve. In Section 3, the Gompertz multi-sigmoidal diffusion process is defined. To this end, the lognormal diffusion process with exogenous factors is considered, since it allows us to model behavioral patterns that verify the properties exhibited by the curve. The estimation of the process, which is performed by maximum

likelihood using discrete sampling, is the subject of Section 4. The matter of obtaining initial solutions to solve the resulting system of equations deserves special attention. Other important aspect is to determine the degree of polynomial that should be considered since, in general, this aspect will be unknown in real applications. To this end, some criteria are considered in that section. Finally, in Section 5, some simulation examples are considered.

2. Multi-Sigmoidal Gompertz Curve

Let $Q_{\beta}(t) = \sum_{l=1}^{p} \beta_l t^l$ be a p-degree polynomial , where $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)^T$ ($p > 1$) denotes a real parametric vector with positive leading coefficient β_p. We define the multi-sigmoidal Gompertz function as

$$f_{\boldsymbol{\theta}}(t) = k \exp\left(-\alpha\, e^{-Q_{\beta}(t)}\right), \qquad t \geq t_0 \geq 0, \qquad \alpha, k > 0, \qquad \boldsymbol{\theta} = (\alpha, \boldsymbol{\beta}^T)^T. \tag{1}$$

Denoting $P_{\beta}(t) = \frac{dQ_{\beta}(t)}{dt}$, curve (1) satisfies the ordinary linear differential equation

$$\frac{df_{\boldsymbol{\theta}}(t)}{dt} = f_{\boldsymbol{\theta}}(t) h_{\boldsymbol{\theta}}(t), \tag{2}$$

where

$$h_{\boldsymbol{\theta}}(t) = \alpha\, P_{\beta}(t)\, e^{-Q_{\beta}(t)}. \tag{3}$$

Taking into account that $\ln f_{\boldsymbol{\theta}}(t) = \ln k - \alpha\, e^{-Q_{\beta}(t)}$, the above equation can be expressed as

$$\frac{df_{\boldsymbol{\theta}}(t)}{dt} = f_{\boldsymbol{\theta}}(t)\left(\ln k - \ln f_{\boldsymbol{\theta}}(t)\right) P_{\beta}(t). \tag{4}$$

However, the resolution of both equations, with initial condition $f_{\boldsymbol{\theta}}(t_0) = f_0 > 0$, leads to two expressions of curve (1). Indeed, the solution of (2) is

$$f_{\boldsymbol{\theta}}(t) = f_0 \exp\left(-\alpha\, \left(e^{-Q_{\beta}(t)} - e^{-Q_{\beta}(t_0)}\right)\right), \tag{5}$$

while the solution of (4) is

$$f_{\boldsymbol{\theta}}(t) = \exp\left(\ln k\left(1 - e^{-Q_{\beta}(t) - Q_{\beta}(t_0)}\right) + \ln f_0\, e^{-Q_{\beta}(t) - Q_{\beta}(t_0)}\right). \tag{6}$$

Equation (4) is a generalization of the classical gompertzian differential equation, giving rise to the curve (6), used by authors like Ricciardi et al. [26] in the case $Q_{\beta}(t) = \beta t$. On the other hand, Equation (2) is a linear differential equation of the Malthusian type whose solution generalizes the expression of the Gompertz curve used by authors such as Laird [27] and Gutiérrez et al. [12].

The main difference between (5) and (6) lies in their limit value, which in the first case is $k(\boldsymbol{\theta}) = f_0 \exp\left(\alpha\, e^{-Q_{\beta}(t_0)}\right)$, and k in the second. This may lead to the choice of either expression depending on the knowledge available about the influence of initial value f_0 on the limit value. This is the case when the phenomenon under study shows Gompertz-type growth and several sample paths are available, each with a common growth pattern but with different initial values and a different limit value (for example, the particular weight of each individual of the same species). In the rest of the paper, we will consider the situation in which the carrying capacity of the system modeled by the curve depends on the initial value of the population under study. So, (5) can be expressed as $f_{\boldsymbol{\theta}}(t) = k(\boldsymbol{\theta}) g_{\boldsymbol{\theta}}(t)$ where $g_{\boldsymbol{\theta}}(t) = \exp\left(-\alpha\, e^{-Q_{\beta}(t)}\right)$, verifying $\lim_{t \to \infty} g_{\boldsymbol{\theta}}(t) = 1$.

Focusing on the general expression given by (1), from (2) it follows that the growth intervals of the curve depend on the roots of the equation $P_\beta(t) = 0$. As for the inflection points, the candidates will be the solutions of $\frac{d^2 f_\theta(t)}{dt^2} = 0$, which results in solving the equation

$$\frac{dP_\beta(t)}{dt} = P_\beta^2(t) \left(1 - \alpha \, e^{-Q_\beta(t)}\right).$$

(7)

Figure 1 shows some possible situations for various choices of the polynomial $Q_\beta(t)$. In each case, the Gompertz curve is represented together with its first and second derivatives. In particular, figure (a) represents the case in which $P_\beta(t)$ has no roots and the curve is strictly increasing and presenting two inflection points, as in case (b), although in the latter case the curve presents both decreasing and increasing intervals. Finally, figure (c) shows an example with three inflection points.
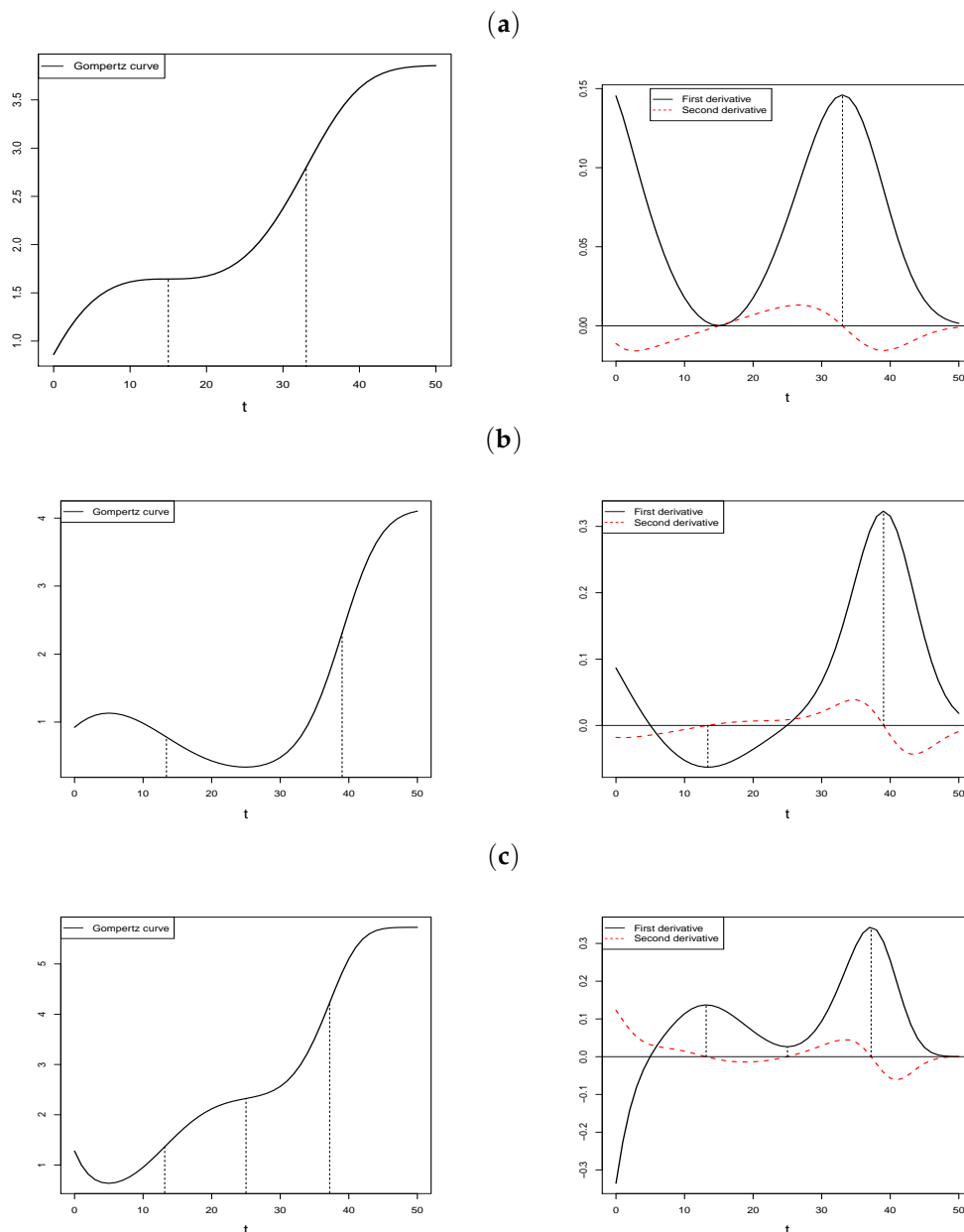

Figure 1. Strictly increasing Gompertz curve with two inflection points (**a**). Decreasing and increasing curves with two (**b**) and three (**c**) inflection points.

3. Multi-Sigmoidal Gompertz Diffusion Process

Assuming that (7) has at least one solution where $f_{\boldsymbol{\theta}}$ presents an inflection point, this function is found in the conditions listed in Román and Torres [28], which ensures that the growth phenomenon represented by the curve can be modeled by a non-homogeneous lognormal diffusion process whose mean function is such a function. In Román et al. [29], a general study of this process is carried out, including the distribution and main characteristics as well as aspects related to inference. Following the notation used in that paper, we define the multi-sigmoidal Gompertz process as a diffusion process $\{X(t); t \in I\}$ that takes values in $(0, +\infty)$ and with infinitesimal moments

$$A_1(x, t) = h_{\boldsymbol{\theta}}(t)x$$
$$A_2(x) = \sigma^2 x^2, \qquad \sigma > 0$$

where $I = [t_0, +\infty)$ is a real interval ($t_0 \geq 0$) and $\Theta \subseteq \mathbb{R}^{p+1}$ is an open set such that $\boldsymbol{\theta} = (\alpha, \boldsymbol{\beta}^T)^T \in \Theta$, where $h_{\boldsymbol{\theta}}$ is given by (3). This process is determined from the stochastic differential equation

$$dX(t) = h_{\boldsymbol{\theta}}(t)X(t)dt + \sigma X(t)dW(t), \qquad X(t_0) = X_0,$$

where $W(t)$ is a Wiener process (Brownian motion), independent of the initial condition $X_0 = X(t_0)$, $t \geq t_0$. The solution to this equation can be expressed as

$$X(t) = X_0 \exp\left(H_{\boldsymbol{\xi}}(t_0, t) + \sigma(W(t) - W(t_0))\right), \qquad t \geq t_0 \tag{8}$$

where

$$H_{\boldsymbol{\xi}}(s, t) = \int_s^t h_{\boldsymbol{\theta}}(u)du - \frac{\sigma^2}{2}(t - s) = -\alpha\left(e^{-Q_{\boldsymbol{\beta}}(t)} - e^{-Q_{\boldsymbol{\beta}}(s)}\right) - \frac{\sigma^2}{2}(t - s), \qquad s < t, \boldsymbol{\xi} = (\boldsymbol{\theta}^T, \sigma^2)^T.$$

Regarding the distribution of the process, if X_0 is distributed according to a lognormal distribution $\Lambda_1\left[\mu_0; \sigma_0^2\right]$, or X_0 is a degenerate variable (i.e., $P[X_0 = x_0] = 1$), the finite dimensional distributions of the process are lognormal. Thus, $\forall n \in \mathbb{N}$ and $t_1 < \cdots < t_n$, vector $(X(t_1), \ldots, X(t_n))^T$ is distributed according to an n-dimensional lognormal distribution $\Lambda_n[\boldsymbol{\varepsilon}, \boldsymbol{\Sigma}]$, where the components of the vector $\boldsymbol{\varepsilon}$ are $\varepsilon_i = \mu_0 + H_{\boldsymbol{\xi}}(t_0, t_i)$, $i = 1, \ldots, n$, being $\sigma_{ij} = \sigma_0^2 + \sigma^2(\min(t_i, t_j) - t_0)$, $i, j = 1, \ldots, n$, those of the matrix $\boldsymbol{\Sigma}$.

From the two-dimensional distributions $(X(s), X(t))^T$, $s < t$, the transitions of the process can be obtained, which are also lognormal; concretely,

$$X(t) \mid X(s) = y \rightsquigarrow \Lambda_1\left(\ln y + H_{\boldsymbol{\xi}}(s, t), \sigma^2(t - s)\right), \qquad s < t.$$

Once the distribution of the process has been established, different characteristics associated with it can be calculated, including the mean and conditioned mean functions, whose expressions are

$$m(t) = E[X(t)] = E[X_0]\exp\left(-\alpha\left(e^{-Q_{\boldsymbol{\beta}}(t)} - e^{-Q_{\boldsymbol{\beta}}(t_0)}\right)\right)$$

and

$$m(t|t_0) = E[X(t)|X(t_0) = x_0] = x_0\exp\left(-\alpha\left(e^{-Q_{\boldsymbol{\beta}}(t)} - e^{-Q_{\boldsymbol{\beta}}(t_0)}\right)\right)$$

respectively, being able to verify that both functions are of the type introduced in the previous section. The expression that adopts other characteristics can be consulted in Román et al. [29].

4. Maximum Likelihood Estimation

In this section we will deal with the estimation of the parameters of the process by maximum likelihood. In Román and Torres [29], a general treatment of this question is carried out for the

non-homogeneous lognormal process, of which the process dealt with in this paper is a particular case. Below we summarize the main results obtained in the aforementioned paper, adapting them to the actual case.

The starting point is a discrete sampling of the process based on d sample paths observed at time instants t_{ij}, $(i = 1, \ldots, d, j = 1, \ldots, n_i)$. Note that the observation time instants do not have to be the same in each trajectory, although we will suppose that $t_{i1} = t_0$, $i = 1, \ldots, d$. Denote by $\mathbf{X} = (\mathbf{X}_1^T | \cdots | \mathbf{X}_d^T)^T$, where $\mathbf{X}_i^T$ is the vector that contains the variables of the i-th sample-path, that is $\mathbf{X}_i = (X(t_{i1}), \ldots, X(t_{i,n_i}))^T$, $i = 1, \ldots, d$.

Assuming that the distribution of $X(t_0)$ is lognormal $\Lambda_1(\mu_1, \sigma_1^2)$, and taking into account the transitions of the process, the probability density function of $\mathbf{X}$ is given by

$$f_{\mathbf{X}}(\mathbf{x}) = \prod_{i=1}^d \frac{\exp\left(-\frac{[\ln x_{i1} - \mu_1]^2}{2\sigma_1^2}\right)}{x_{i1}\sigma_1\sqrt{2\pi}} \prod_{j=1}^{n_i-1} \frac{\exp\left(-\frac{\left[\ln\left(x_{i,j+1}/x_{ij}\right) - m_\xi^{i,j+1,j}\right]^2}{2\sigma^2 \Delta_i^{j+1,j}}\right)}{x_{ij}\sigma\sqrt{2\pi\Delta_i^{j+1,j}}},$$

where $m_\xi^{i,j+1,j}$ and $\Delta_i^{j+1,j}$ are given by

$$m_\xi^{i,m,n} = H_\xi(t_{in}, t_{im}) = -\alpha\,\phi_{i,m,n}^{\beta,0} - \frac{\sigma^2}{2}\Delta_i^{m,n}$$

$$\Delta_i^{m,n} = t_{im} - t_{in}$$

with

$$\phi_{i,m,n}^{\beta,l} = t_{im}^l\,e^{-Q_\beta(t_{im})} - t_{in}^l\,e^{-Q_\beta(t_{in})}, \qquad l = 0, 1, \ldots; \qquad i = 1, \ldots, d; \qquad m, n \in \{1, \ldots, n_{i-1}\}, m > n.$$

Next, we consider the change of variables

$$V_{0i} = X_{i1}, \quad i = 1, \ldots, d$$

$$V_{ij} = (\Delta_i^{j+1,j})^{-1/2} \ln \frac{X_{i,j+1}}{X_{ij}}, \quad i = 1, \ldots, d; j = 1, \ldots, n_i - 1$$

that transforms vector $\mathbf{X}$ in $\mathbf{V} = \left[\mathbf{V}_0^T | \mathbf{V}_1^T | \cdots | \mathbf{V}_d^T\right]^T = \left[\mathbf{V}_0^T | \mathbf{V}_{(1)}^T\right]^T$, whit probability density function

$$f_{\mathbf{V}}(\mathbf{v}) = \frac{\exp\left(-\frac{1}{2\sigma_1^2}(\ln \mathbf{v}_0 - \mu_1 \mathbf{1}_d)^T (\ln \mathbf{v}_0 - \mu_1 \mathbf{1}_d)\right)}{\prod_{i=1}^d v_{0i}\left(2\pi\sigma_1^2\right)^{\frac{d}{2}}} \frac{\exp\left(-\frac{1}{2\sigma^2}\left(\mathbf{v}_{(1)} - \gamma^\xi\right)^T \left(\mathbf{v}_{(1)} - \gamma^\xi\right)\right)}{(2\pi\sigma^2)^{\frac{n}{2}}} \tag{9}$$

where $\ln \mathbf{v}_0 = (\ln v_{01}, \ldots, \ln v_{0d})^T$, $n = \sum_{i=1}^d (n_i - 1)$, $\mathbf{1}_d = (1, \ldots, 1)_{d \times 1}^T$, and γ^ξ is an n-dimensional vector with components $\gamma_{ij}^\xi = (\Delta_i^{j+1,j})^{-1/2} m_\xi^{i,j,j+1}$, $i = 1, \ldots, d; j = 1, \ldots, n_i - 1$.

For a fixed value $\mathbf{v}$, expression (9) provides the likelihood function, whose logarithm is

$$L_{\mathbf{v}}(\eta, \xi) = -\frac{(n+d)\ln(2\pi)}{2} - \frac{d\ln\sigma_1^2}{2} - \sum_{i=1}^d \ln v_{0i} - \frac{\sum_{i=1}^d [\ln v_{0i} - \mu_1]^2}{2\sigma_1^2} - \frac{n\ln\sigma^2}{2} - \frac{Z_1 + \Phi_\xi - 2\Gamma_\xi}{2\sigma^2}$$

where $\eta = (\mu_1, \sigma_1^2)^T$ is the vector that contains the parameters of the initial distribution, being

$$Z_1 = \sum_{i=1}^{d} \sum_{j=1}^{n_i-1} v_{ij}^2, \qquad \Phi_\xi = \sum_{i=1}^{d} \sum_{j=1}^{n_i-1} \frac{\left(m_\xi^{i,j+1,j}\right)^2}{\Delta_i^{j+1,j}}, \qquad \Gamma_\xi = \sum_{i=1}^{d} \sum_{j=1}^{n_i-1} \frac{v_{ij} m_\xi^{i,j+1,j}}{(\Delta_i^{j+1,j})^{1/2}}.$$

Assuming that η and ξ are functionally independent, the estimate of η leads to

$$\widehat{\mu}_1 = \frac{1}{d} \sum_{i=1}^{d} \ln v_{0i} \text{ and } \widehat{\sigma}_1^2 = \frac{1}{d} \sum_{i=1}^{d} (\ln v_{0i} - \widehat{\mu}_1)^2,$$

while that of ξ is obtained (see [29] for details) from the system of equations

$$\Psi_\theta - \Omega_\xi = 0 \tag{10}$$

$$Z_1 + \Phi_\xi - 2\Gamma_\xi - \sigma^2 Z_2 + \sigma^2 Y_\xi = n\sigma^2 \tag{11}$$

where

$$\Omega_\xi = \frac{1}{2} \frac{\partial \Phi_\xi}{\partial \theta^T}, \qquad \Psi_\theta = \frac{1}{2} \frac{\partial \Gamma_\xi}{\partial \theta^T}, \qquad Y_\xi = -\frac{\partial \Phi_\xi}{\partial \sigma^2}, \qquad Z_2 = -2 \frac{\partial \Gamma_\xi}{\partial \sigma^2}.$$

Taking into account that

$$\frac{\partial m_\xi^{i,j+1,j}}{\partial \theta^T} = \left(-\phi_{i,j+1,j}^{\beta,0}, \alpha \phi_{i,j+1,j}^{\beta,1}, \alpha \phi_{i,j+1,j}^{\beta,p} \right),$$

and the previous expressions of Ω_ξ and Ψ_θ, the subsystem of equations (10) remains in the form

$$X_l^\beta + \alpha Y_l^\beta + \frac{\sigma^2}{2} W_l^\beta = 0, \qquad l = 0, 1, \ldots, p \tag{12}$$

where, for $l = 0, 1, \ldots, p$ one has

$$X_l^\beta = \sum_{i=1}^{d} \sum_{j=1}^{n_i-1} \frac{v_{ij} \phi_{i,j+1,j}^{\beta,l}}{(\Delta_i^{j+1,j})^{1/2}}, \qquad Y_l^\beta = \sum_{i=1}^{d} \sum_{j=1}^{n_i-1} \frac{\phi_{i,j+1,j}^{\beta,0} \phi_{i,j+1,j}^{\beta,l}}{\Delta_i^{j+1,j}}, \qquad W_l^\beta = \sum_{i=1}^{d} \phi_{i,n_i,1}^{\beta,l}.$$

On the other hand, and since

$$\Phi_\xi = \alpha^2 Y_0^\beta + \frac{\sigma^4}{4} Z_3 + \alpha \sigma^2 W_0^\beta, \qquad \Gamma_\xi = -\alpha X_0^\beta - \frac{\sigma^2}{2} Z_2, \qquad Y_\xi = -\alpha W_0^\beta - \frac{\sigma^2}{2} Z_3, \qquad Z_3 = \sum_{i=1}^{d} \Delta_i^{n_i,1}$$

Equation (11) transforms into

$$\sigma^2 \left[n + \sigma^2 Z_3/4 \right] - \alpha \left[2X_0^\beta + \alpha Y_0^\beta \right] - Z_1 = 0. \tag{13}$$

4.1. System of Equations and Numerical Computations

The system of Equations (12) and (13) can not be solved explicitly, and it is therefore necessary to use numerical methods, such as Newton-Raphson, for which an initial solution is required. Next, we present a strategy to achieve this, based on the information provided by the sample data.

As a matter of fact, and taking into account that the mean function of the process is a Gompertz multi-sigmoidal curve, as well as the expression (1), it follows

$$\ln \ln \frac{k}{f_\theta(t)} = \ln \alpha - Q_\beta(t).$$

Noting m_i the values of the mean of the sample paths at t_i, we propose to fit, by linear regression, a polynomial taking as data the pairs of values $(t_i, \ln(\ln(k/m_i)))$. The estimated coefficients will provide the initial values for α and β. Regarding σ, its initial estimation is based on the fact that for a lognormal distribution $\Lambda_1[\eta, \delta]$, the quotient between the arithmetic mean and the geometric one provides an estimation of δ; concretely $\widehat{\delta} = 2\ln(E[X]/E_G[X])$. Applying this result to the distribution of $X(t)$ we obtain, for each t_i, an estimate of $\sigma^2(t_i - t_0)$; that is, $\sigma_i^2 = 2\ln(m_i/m_i^g)$, $i = 1, \ldots$, where m_i^g are the values of the geometric sample mean. Finally, the initial value of σ is calculated by performing a simple linear regression of the σ_i values against t_i.

In this procedure there are several questions that must be taken into account:

- The value of k, in general, will not be known. Therefore, we suggest taking as an approximation the last value of the mean. However, it is possible that in real cases, and due to the fluctuations of the process, there could be values m_i verifying $k \leq m_i$, so transformation $\ln(\ln(k/m_i))$ would not be determined. In such cases, usually not many in practical cases, those points must be removed from the regression analysis.

- Since, generally, the degree of the polynomial will not be known a priori, it is necessary to have some mechanism that will allow for its selection. To this end we propose a forward procedure, introducing polynomials in a consecutive way in the model. Each time a polynomial is introduced, a measure of the adjustment made is calculated and compared with the previous ones. If the adjustment is improved, the procedure continues; otherwise it stops. However, and even in this case, it is convenient to perform one more iteration due to the parity of the polynomial.

- Regarding the measures that can be used to evaluate the adjustment, we propose the following:

  - The absolute relative errors between the sample mean of the process and the fitted mean for each estimated model

  $$RAE_j = \frac{1}{N}\sum_{i=1}^{N}\frac{|m_i - \widehat{E}[X^{(j)}(t_i)]|}{m_i}, j = 1, 2, \ldots$$

  - The resistor-average distance (see Johnson and Sinanovic [30]). This is a distance based on the Kullback–Leibler divergence, which will be used for calculating the distance between the sample distribution (available from the data) and that obtained from each estimated model. The expression for this measure is

  $$D_{RA}(f_s||f_i) = \frac{D_{KL}(f_s||f_i) \cdot D_{KL}(f_i||f_s)}{D_{KL}(f_s||f_i) + D_{KL}(f_i||f_s)},$$

  where $D_{KL}(f_s||f_i)$ denotes the Kullback–Leibler divergence between the sample distribution (f_s) and that for the i-th estimated model (f_i). Its expression is given by

  $$D_{KL}(f_s||f_i) = \frac{1}{2}\left[\ln\left(\frac{\widehat{\sigma}^2(t_i - t_0)}{\sigma_i^2}\right) + \frac{\sigma_i^2}{\widehat{\sigma}^2(t_i - t_0)} + \frac{\left(\ln m_i^g - \ln\widehat{E[X_0]} - H_{\widehat{\xi}}(t_0, t_i)\right)^2}{\widehat{\sigma}^2(t_i - t_0)} - 1\right].$$

  In practice, this is the expression of the distance that should be used since the theoretical model will not be known in real applications. However, for simulation studies the distance between theoretical and estimated models could be considered. In this case, the previous expression would be slightly modified.

  - The Akaike information criterion (AIC) and the Bayes information criterion (BIC).

4.2. About t_0

Another interesting aspect to consider is that of the time instants, especially when they take high values. This can mainly affect the obtention of the initial values for the parameters since a polynomial regression has been proposed. One option is to apply orthogonal polynomials, as it is usual when considering this type of regression.

One alternative is to consider a new diffusion process $\{Y(t); t \geq 0\}$ obtained from $\{X(t); t \geq t_0\}$ by considering a shift of length t_0 in time, that is $Y(t) = X(t + t_0)$, so that the original data can be considered as observations of the new process with an initial instant equal to zero. Let's see how this change affects the infinitesimal moments of the processes:

In general, let $\{X(t); t \geq t_0\}$ be the original process and $\{Y(t); t \geq 0\}$ verifying $X(t) = Y(t - t_0)$. Denote by $A_m^X(x,t)$ and $A_m^Y(x,t)$ their respective m-th order infinitesimal moments. Taking into account the definition of infinitesimal moment of order m,

$$A_m^X(x,t) = \lim_{h \to 0} \frac{1}{h} E\left[(X(t+h) - X(t))^m | X(t) = x\right],$$

it verifies that $A_m^X(x,t) = A_m^Y(x, t - t_0)$. Obviously, the strategy of considering this translation over time will be useful when the resulting process is of the same type as the original.

In the case of the multi-sigmoidal Gompertz process, let us consider $\{Y(t); t \geq 0\}$ with infinitesimal moments

$$A_1^Y(x,t) = \eta \widetilde{P}_\gamma(t)\, e^{-\widetilde{Q}_\gamma(t)}, \qquad A_2^Y(x) = \sigma^2\, x^2$$

where $\gamma = (\gamma_1, \ldots, \gamma_p)^T$.

Taking into account that

$$\widetilde{Q}_\gamma(t - t_0) = \sum_{l=1}^{p} \gamma_l\, (t - t_0)^l = \beta_0 + \sum_{m=1}^{p} \beta_m\, t^m = \beta_0 + Q_\beta(t),$$

where $\gamma = (\gamma_1, \ldots, \gamma_m)$, being

$$\beta_0 = \sum_{j=1}^{p} \gamma_j(-t_0)^j, \qquad \beta_m = \sum_{j=m}^{p} \binom{j}{m}(-t_0)^{j-m}, m = 1, \ldots, p,$$

then the corresponding infinitesimal moments for the process $\{X(t); t \geq t_0\}$ given by $X(t) = Y(t - t_0)$ are

$$A_1^X(x,t) = A_1^Y(x, t - t_0) = \eta \widetilde{P}_\gamma(t - t_0)\, e^{-\widetilde{Q}_\gamma(t-t_0)} = \alpha\, P_\beta(t) e^{-Q_\beta(t)}, \qquad A_2^X(x) = A_2^Y(x)$$

with $\alpha = \eta\, e^{-\beta_0}$ and P_β the derivative of polynomial Q_β.

Note that $X(t)$ is also a multi-sigmoidal Gompertz diffusion process whose infinitesimal moments differ from those of $Y(t)$ in the reparametrization occurred in η and γ. The same happens for the finite dimensional distributions, transition distributions and main characteristics of the process. In particular, $E[X(t)] = E[Y(t - t_0)]$.

5. Simulations

In this section, some simulation examples will be carried out with the aim of illustrating the developments previously established, focusing on the strategies related to the estimation of the parameters of the model as well as the selection of the model that best fits the data. All the simulations were performed according to the following common pattern: 25 sample paths were simulated, each one obtained from expression (8), which relates the Gompertz process under consideration and the Wiener process. All of them contain the same number of data (501), being $(i-1) \cdot 0.1, i = 1, \dots, 501$ the observation time instants. For simplicity we have chosen a degenerate initial distribution $(P[X(0) = 5] = 1)$. After obtaining each trajectory, we chose 51 values starting from the first one and using a step equal to 1. Hence, a sample of 51 data was obtained for each sample path.

With regard to the processes chosen for the simulation, two have been selected that correspond to situations in which there are two inflection points. The former presents a strictly increasing mean, while in the second an initial decrease is observed.

5.1. The Case of Increasing Mean

As a first example we have selected a multi-sigmoidal Gompertz diffusion process for which the degree of the polynomial included in the infinitesimal mean is $p = 3$, being $\boldsymbol{\beta} = (0.1225, -0.0075; 0.00017)^T$. The value of α is $\alpha = e^{-1}$, while two values of σ have been considered (concretely $\sigma = 0.01, 0.05$) to verify the effect of increasing the infinitesimal variance in the estimation process. Figure 2 shows the 25 simulated sample paths for each value of σ.
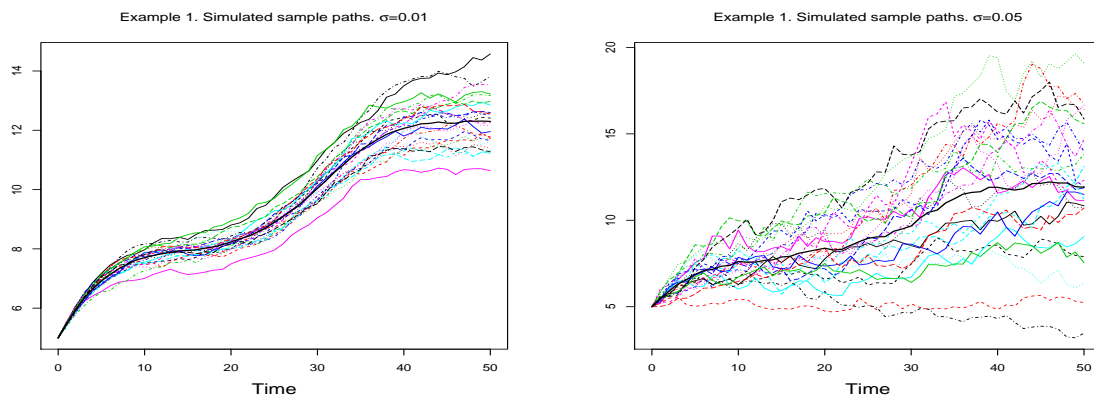

Figure 2. Example 1. Simulated sample-paths. Black lines represent the sample mean.

In order to find the process that best fits the data, for each sigma value, multi-sigmoidal Gompertz processes including polynomials from grade 2 to 5 have been considered successively. Table 1 includes, for each model, the initial values of the parameters as well as their definitive estimates. The initial values have been obtained following the procedure described above. Note that the initial σ value is common for all cases since it is calculated directly from the sample data.

Table 1. Example 1. Estimates of the parameters of the models for each value of p.

Degree		α	σ	β_1	β_2	β_3	β_4	β_5
				$\sigma = 0.01$				
2	Initial	0.43264	0.00951	−0.08412	0.00405			
	Final	0.44732	0.01632	0.19938	0.00726			
3	Initial	0.82721	0.00951	0.09462	−0.00576	0.00014		
	Final	0.90536	0.00975	0.12100	−0.00738	0.00016		
4	Initial	1.01846	0.00951	0.19476	−0.01580	0.00048	−0.0000037	
	Final	0.52895	0.01221	0.33611	−0.02349	0.00043	0.0000019	
5	Initial	0.80607	0.00951	0.01539	0.01266	−0.00119	0.000037	−0.00000037
	Final	0.80607	0.00951	0.01539	0.01266	−0.00119	0.000037	−0.00000037
				$\sigma = 0.05$				
2	Initial	0.40172	0.05147	−0.09713	0.00468			
	Final	−0.04547	0.05450	−0.98848	0.28151			
3	Initial	0.89617	0.05147	0.14054	−0.00929	0.00021		
	Final	0.87641	0.04832	0.11912	−0.00698	0.00015		
4	Initial	0.93752	0.05147	0.16394	−0.01181	0.00030	−0.0000010	
	Final	0.95883	0.05139	0.15846	−0.01143	0.00030	−0.0000016	
5	Initial	0.76767	0.05147	−0.00203	0.01645	−0.00147	0.0000458	−0.00000043
	Final	0.76767	0.05147	−0.00203	0.01645	−0.00147	0.0000458	−0.00000043

Figure 3 displays the theoretical and the sample mean functions, together with those corresponding to each estimated model. This figure suggests considering the model with $p = 3$ as the optimum, although this must be endorsed by numerical measures of goodness of fit.
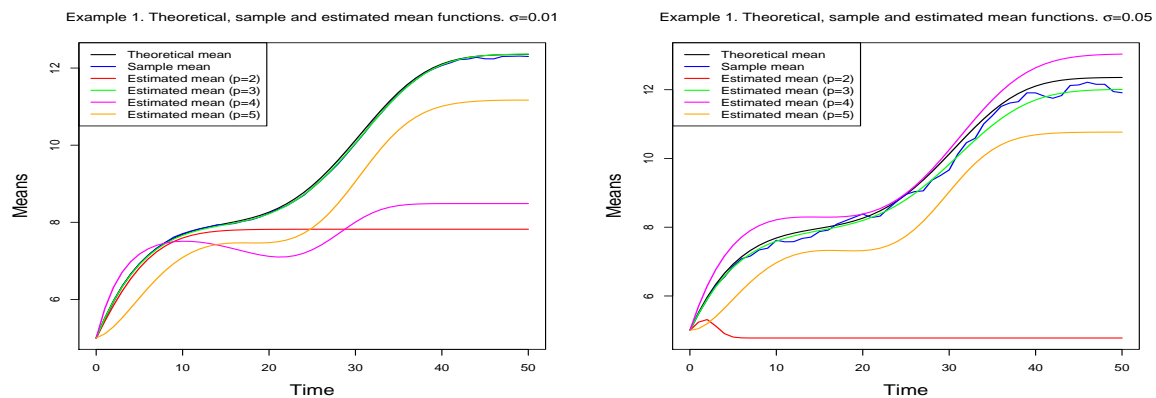

Figure 3. Example 1. Theoretical, sample and estimated mean functions.

Table 2 summarizes the measures that have been used (RAE, AIC and BIC). For all of them, it can be observed how from $p = 3$ the goodness of fit can not be improved. Therefore, the third-degree model has been chosen as the optimal one.

Table 2. Example 1. Measures for choosing the estimated model.

		Degree			
		2	**3**	**4**	**5**
	RAE	0.168	0.002	0.171	0.092
$\sigma = 0.01$	AIC	-9034.903	-10314.497	-8952.280	-7870.896
	BIC	-9014.380	-10288.842	-8921.495	-7834.980
	RAE	0.459	0.011	0.056	0.099
$\sigma = 0.05$	AIC	-5967.582	-6315.135	-6275.419	-6188.681
	BIC	-5947.058	-6289.480	-6244.633	-6152.765

The use of the resistor-average distance also leads us to this conclusion. For each value of p, and for each value of t, the distance between the estimated one-dimensional distribution and the corresponding theoretical and sample distributions has been calculated. This provides, for each degree, two functions whose graphs are shown in Figure 4, showing how odd-grade models seem to be preferable. With the idea of obtaining a globalizing measure that allows selecting the best model, Table 3 shows, for each of them, the means and medians of the values of the distances. These two measures confirm that the model with $p = 3$ is the one that should be selected as optimal. It should be noted that in practical applications, as the theoretical model is not available, the distance to be considered is that which takes the sample distribution as a reference. However, in this first example we have included the two possibilities.
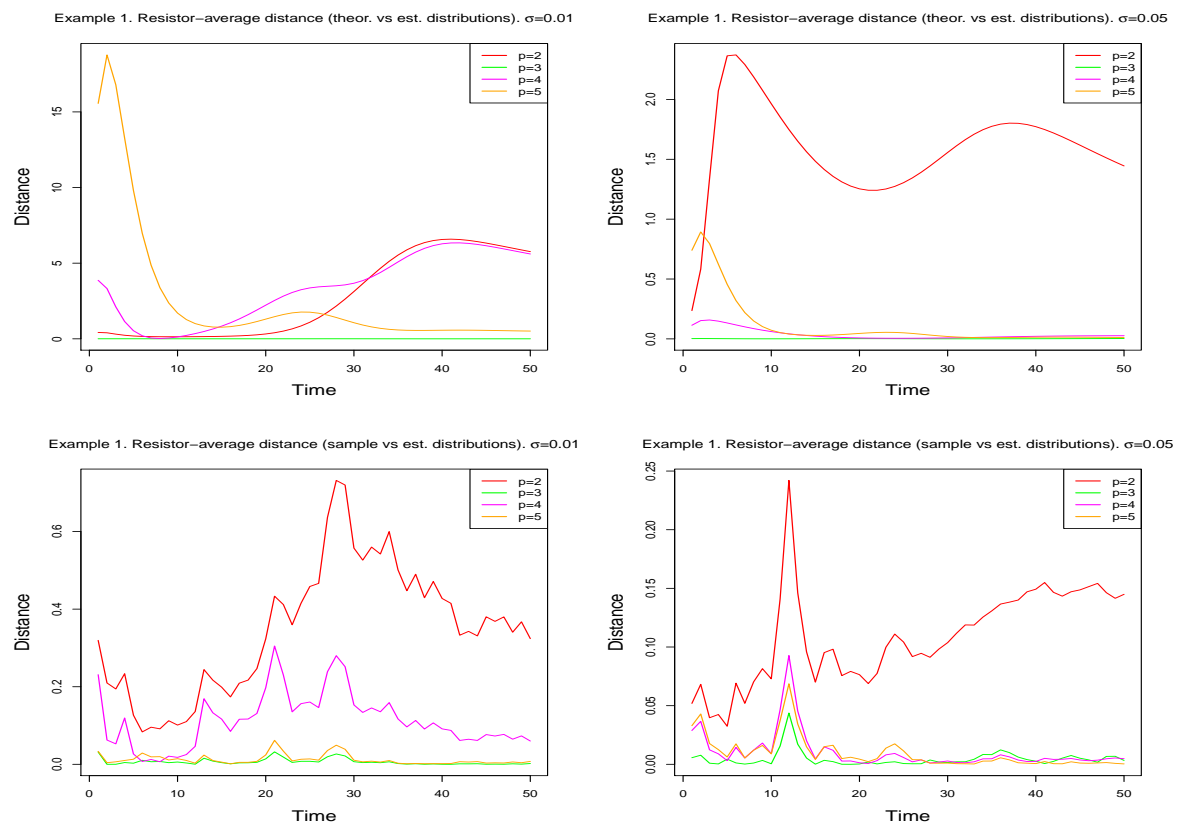

Figure 4. Example 1. Resistor-average distances between the theoretical and estimated models (**up**) and between the sample and estimated models (**down**).

Table 3. Example 1. Means and medians of the resistor-average distances.

		Degree							
		2		**3**		**4**		**5**	
	Distr.	**Mean**	**Median**	**Mean**	**Median**	**Mean**	**Median**	**Mean**	**Median**
$\sigma = 0.01$	Theoretical	2.71439	1.24435	0.00210	0.00208	3.33965	3.45329	2.60928	0.95538
	Sample	0.34820	0.35109	0.00641	0.00421	0.11279	0.10989	0.01249	0.00759
$\sigma = 0.05$	Theoretical	1.58825	1.59751	0.00148	0.00139	0.03698	0.02188	0.10850	0.03090
	Sample	0.10873	0.10398	0.00464	0.00257	0.01023	0.00473	0.00938	0.00408

Finally, Figure 5 shows, for each σ value, the first and second derivatives of the theoretical, sample and estimated mean functions for the selected model. In order to obtain the derivatives of the sample mean function, a smoothing of the function has previously been carried out using polynomial local regression. A good fit between these functions can be observed, which is corroborated by Table 4, which contains the theoretical, sample, and estimated values of the inflection time instants.
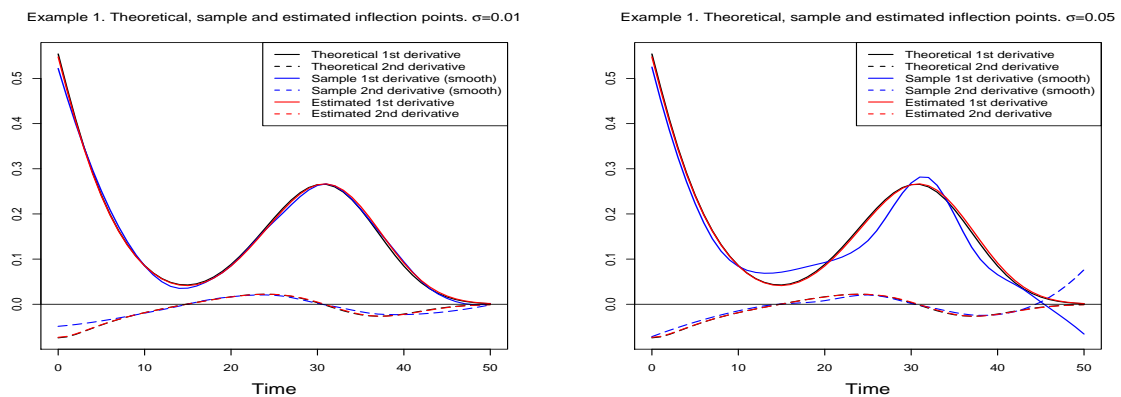

Figure 5. Example 1. Theoretical, sample, and estimated inflection time instants.

Table 4. Example 1. Sample and estimated inflection time instants for the model chosen. Theoretical values are in parentheses.

		Inflection Time Instants	
		t_1 **(14.787)**	t_2 **(30.589)**
$\sigma = 0.01$	Sample	14.922	30.522
	Estimated	14.869	30.834
$\sigma = 0.05$	Sample	14.595	30.394
	Estimated	14.789	30.553

5.2. The Case of Mean Decreasing, Then Increasing

This example illustrates a case in which the data present an initial decrease and then grow up to the value of the upper bound. As in the previous case, we have selected a model with $p = 3$, being $\boldsymbol{\beta} = (0.0626, -0.009, 0.0002)^T$, $\alpha = e^{-1}$ and $\sigma = 0.025$. Figure 6 shows the simulated sample paths.
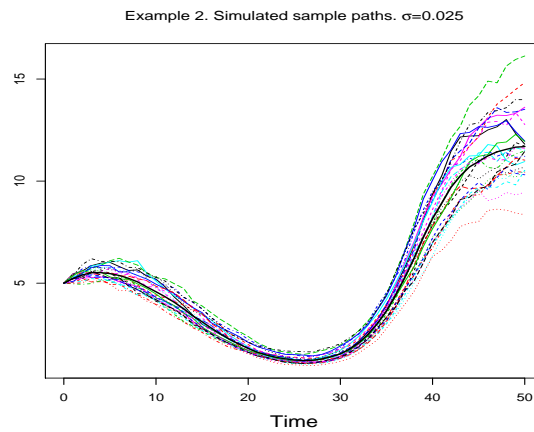
Figure 6. Example 2. Simulated sample-paths. The black line represents the sample mean.

Following the same methodology developed in the previous example, once again the procedure stops when considering the polynomial of degree 5. Figure 7 shows the estimated means together with the theoretical and the sample ones.
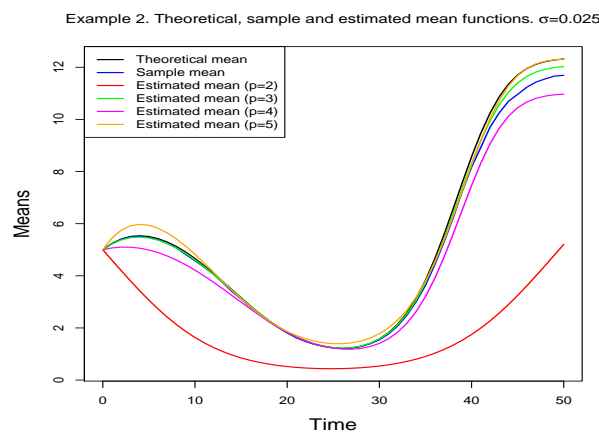

Figure 7. Example 2. Theoretical, sample, and estimated mean functions.

Table 5 contains the values of goodness-of-fit measurements, from which it follows that the model containing the third-degree polynomial must be chosen.

Table 5. Example 2. Measures for choosing the estimated model.

	Degree			
	2	**3**	**4**	**5**
RAE	0.168	0.002	0.171	0.092
AIC	-9034.903	-10314.497	-8952.280	-7870.896
BIC	-9014.380	-10288.842	-8921.495	-7834.980

Regarding resistor-average distances, in this example we have only considered those calculated between the estimated models and the sample distribution. Figure 8 and Table 6 contain the results obtained, confirming the previous choice of the model.
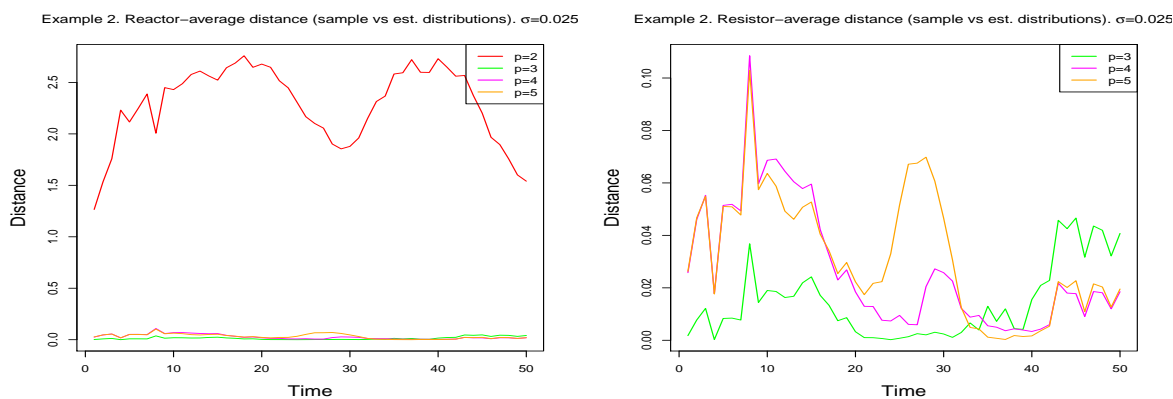
Figure 8. Example 2. Resistor-average distances between the sample and estimated models. The figure on the right enlarges the previous one for a better view.

Table 6. Example 2. Means and medians of the resistor-average distances between estimated and sample unidimensional distributions.

Degree							
2		**3**		**4**		**5**	
Mean	Median	Mean	Median	Mean	Median	Mean	Median
3.60322	3.39936	0.00288	0.00270	0.23962	0.13691	0.15147	0.01742

Table 7 shows the theoretical values of the inflection points (in parentheses) together with the sample and the estimated ones. In each case, the values have been obtained by numerically solving Equation (7) for the theoretical, estimated, and sample mean functions. In the latter case, a natural cubic spline has been previously adjusted to the sample mean. Figure 9 shows the situation graphically. It is readily apparent how the estimation is optimal and provides estimated values very close to the theoretical and sample values.

Table 7. Example 2. Sample and estimated inflection time instants for the model chosen. Theoretical are in parentheses.

	Inflection Time Instants	
	t_1 (**13.888**)	t_2 (**38.403**)
Sample	13.884	38.926
Estimated	13.903	38.474

The results of the simulations carried out demonstrate the suitability of the procedures developed to adjust data that follow a Gompertz multi-sigmoidal pattern. The method of obtaining initial solutions for the resolution of the system of likelihood equations provides optimal values for this purpose from the information provided by the data. The procedures introduced for calculating the degree of the polynomial operate according to two types of criteria: one based on the adjustment of the data to the model and another based on the existing proximity between the sample and estimated distributions of the process. It can be seen, in the two simulations carried out, how both types of criteria coincide in the conclusions drawn from their application.
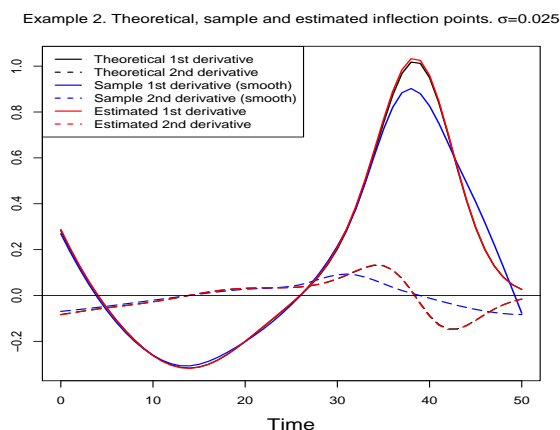
Figure 9. Example 2. Theoretical, sample and estimated inflection time instants.

6. Conclusions

A wide variety of curves (logistic, Gompertz, Bertalanffy, Richards, among others) have been used in order to describe sigmoidal growth patterns in many scientific fields. All these curves have a point of inflection that is always at a fixed proportion of its asymptotic value. Nevertheless, there are real situations in which several growth phases appear, each representing a sigmoidal pattern.

Starting from a modification of the classic Gompertz curve (by including a polynomial function in it), this paper introduces a diffusion process whose mean function is a curve of such characteristics, which allows us to model situations showing this type of behaviour. The process introduced, following the methodology given in Román and Torres [28], is a particular case of the lognormal process with exogenous factors, to which we apply the findings of Román et al. [29]. In particular, the results relative to the estimation of the process by maximum likelihood are adapted, providing strategies that provide the initial solutions required for solving the system of equations by which the parameters are estimated.

One of the main problems arising from the use of this model in real situations is that of determining the degree of the polynomial. For this purpose, the use of several criteria has been suggested. These criteria are based on the goodness of fit (relative absolute error and Akaike and BIC criteria), as well as on the measurement of the distance between the estimated and sample distributions (resistor-average distance, based on the Kullback–Leibler divergence).

Author Contributions: The three authors have participated equally in the development of this work, either in the theoretical developments or in the applied aspects. The paper was also written and reviewed cooperatively.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Eby, W.M.; Tabatabai, M.A.; Bursac, Z. Hyperbolastic modeling of tumor growth with a combined treatment of iodoacetate and dimenthylsuphoxide. *BMC Cancer* **2010**, *10*, 509. [CrossRef] [PubMed]
2. Menon, A.; Mehrotra, K.; Mohan, C.K.; Ranka, S. Characterization of a class of sigmoid functions with applications to rural networks. *Neural Netw.* **1996**, *9*, 819–835. [CrossRef]
3. Yin, X.; Goudriaan, J.; Lantinga, E.A.; Vos, J.; Spiertz, H.J. A flexible sigmoid function of determinate growth. *Ann. Bot.-Lond.* **2003**, *91*, 361–371. [CrossRef] [PubMed]
4. Giovanis, A.P.; Skiadas, C.H. A new modeling approach investigating the diffusion speed of mobile telecomunication services in EU-15. *Comput. Econ.* **2007**, *29*, 97–106. [CrossRef]

5.　Gallagher, B. Peak oil analyzed with a logistic function and idealized Hubbert curve. *Energ. Policy* **2011**, *39*, 709–802. [CrossRef]

6.　Yokohama, S.; Sanada, H. Logistic regression model for predicting language change. In *Issues in Quantitative Linguistics*; Köhler, R., Ed.; RAM-Verlag: Lüdenscheid, Germany, 2009; pp. 176–192, ISBN 978-3-9802659-9-7.

7.　Paolino, D.D.; Cavatorta, M.P. Sigmoidal crack growth rate: Statistical modelling and applicactions. *Fatig. Fract. Eng. Mater. Struct*. **2012**, *36*, 316–326. [CrossRef]

8.　Ricciardi, L.M. On the conjecture concerning population growth in random environment. *Biol. Cybern*. **1979**, *32*, 95–99. [CrossRef]

9.　Schurz, H. Modeling, analysis and discretization of stochastic logistic equations. *Int. J. Numer. Anal. Mod.* **2007**, *4*, 178–197.

10.　Román-Román, P.; Torres-Ruiz, F. Modelling logistic growth by a new diffusion process: Application to biological systems. *Biosystems* **2012**, *110*, 9–21. [CrossRef]

11.　Capocelli, R.M.; Ricciardi, L.M. Growth with regulation in random environment. *Kybernetik* **1974**, *15*, 147–157. [CrossRef]

12.　Gutiérrez, R.; Román, P.; Romero, D.; Serrano, J.J.; Torres, F. A new gompertz-type diffusion process with application to random growth. *Math. Biosci.* **2007**, *208*, 147–165. [CrossRef]

13.　Qiming, L.; Pitchford, J.W. Stochastic Von Bertalanffy models, with applications to fish recruitment. *J. Theor. Biol*. **2007**, *244*, 640–655. [CrossRef]

14.　Román-Román, P.; Romero, D.; Torres-Ruiz, F. A diffusion process to model generalized von Bertalanffy growth patterns: Fitting to real data. *J. Theor. Biol.* **2010**, *263*, 59–69. [CrossRef]

15.　Luz-Sant'Ana, I.; Román-Román, P.; Torres-Ruiz, F. Modeling oil production and its peak by means of a stochastic diffusion process based on the Hubbert curve. *Energy* **2017**, *133*, 455–470. [CrossRef]

16.　Barrera, A.; Román-Román, P.; Torres-Ruiz, F. A hyperbolastic type-I diffusion process: Parameter estimation bymeans of the firefly algorithm. *Biosystems* **2018**, *163*, 11–22. [CrossRef]

17.　Tang, S.; Heron, E. Bayesian inference for a stochastic logistic model with switching points. *Ecol. Model*. **2008**, *219*, 153–169. [CrossRef]

18.　Albano, G.; Giorno, V.; Román-Román, P.; Torres-Ruiz, F. Inferring the effect of therapy on tumors showing stochastic Gompertzian growth. *J. Theor. Biol*. **2011**, *276*, 67–77. [CrossRef]

19.　Albano, G.; Giorno, V.; Román-Román, P.; Román-Román, S.; Torres-Ruiz, F. Estimating and determining the effect of a therapy on tumor dymamics by means of a modified Gompertz diffusion process. *J. Theor. Biol.* **2015**, *364*, 206–219. [CrossRef]

20.　Román-Román, P.; Román-Román, S.; Serrano-Pérez, J.J.; Torres-Ruiz, F. Modeling tumor growth in the presence of a therapy with an effect on rate growth and variability by means of a modified Gompertz diffusion process. *J. Theor. Biol.* **2016**, *407*, 1–17. [CrossRef]

21.　Álvarez, O.; Boché, S. Modelos matemáticos para describir crecimiento doble-sigmoideos en frutos de un nectarín tardío (cv. Sun Grand). *AgroSur* **1999**, *27*, 21–28.

22.　Cairns, S.P.; Robinson, D.M.; Loiselle, D.S. Double-sigmoid model for fitting fatigue profiles in mouse fast- and slow-twitch muscle. *Exp. Physiol*. **2008**, *93*, 851–862. [CrossRef]

23.　Amorim, L.; Bergamin Filho, A.; Hau, B. Analysis of progress curves of sugarcane smut on different cultivars using functions of double sigmoid pattern. *Phytopathology* **1993**, *83*, 933–936. [CrossRef]

24.　Roper, L.D. Using Sigmoid and Double-Sigmoid Functions for Earth-StatesTransitions. 2000. Available online: http://www.roperld.com/Science/DoubleSigmoid.pdf (accessed on 10 June 2019).

25.　Lipovetsky, S. Double logistic curve in regression modeling. *J. Appl. Stat*. **2010**, *37*, 1785–1793. [CrossRef]

26.　Ricciardi, L.M.; Sacerdote, L.; Sato, S. Diffusion approximation and first-passage-time problem for a model neuron II. Outline of a computation Method. *Math. Biosci.* **1983**, *64*, 29–44. [CrossRef]

27.　Laird, A.K. Dynamic of tumour growth: Comparison of growth rates and extrapolation of growth curve to one cell. *Br. J. Cancer* **1965**, *19*, 278–291. [CrossRef]

28.　Román-Román, P.; Torres-Ruiz, F. The nonhomogeneous lognormal diffusion process as a general process to model particular types of growth patterns. In *Lecture Notes of Seminario Interdisciplinare di Matematica, Vol XII*; Università degli Studi della Basilicata: Potenza, Italy, 2015; pp. 201–219.

29. Román-Román, P.; Román-Román, S.; Serrano-Pérez, J.J.; Torres-Ruiz, F. Some Notes about inference for the lognormal diffusion process with exogenous factors. *Mathematics* **2018**, *6*, 85. [CrossRef]

30. Johnson, D.H.; Sinanovic, S. Symmetrizing the Kullback-Leibler distance. 2001. Available online: http://www.ece.rice.edu/∼dhj/resistor.pdf (accessed on 10 June 2019)