Robust Estimation in Gompertz Diffusion Model of Tumor Growth

Jaya PN Bishwal*

Department of Mathematics and Statistics, University of North Carolina at Charlotte, USA

*Corresponding author: Jaya PN Bishwal, Department of Mathematics and Statistics, University of North Carolina at Charlotte, 376 Fretwell Building, 9201 University City Blvd, Charlotte, NC 28223, USA.

Submission: January 27, 2018; Published: May 10, 2018

Abstract

Stochastic Gompertz diffusion model describes the in vivo tumor growth. The drift parameter describes the intrinsic growth rate (mitosis rate) of the tumor. The paper introduces some new approximate minimum contrast estimators of the tumor growth acceleration parameter in the Gompertz diffusion model based on discretely sampled data which are robust and studies their asymptotic distributional properties with precise rates of convergence.

Keywords: Itô stochastic differential equation; Gompertz diffusion process; Black-Karasinski model; Discrete observations; Approximate minimum contrast estimators; Robustness; efficiency; Berry-Esseen bound

Model and Estimators

The Gompertz diffusion process has been used in tumor growth modeling. Ferrante et al. [1]. Lo [2] considered a Gompertz diffusion model in which the size of the tumor cells is bounded and used Lie-algebraic method to derive the exact analytical solution of the functional Fokker-Planck equation obeyed by the density function of the size of the tumor. Giorno et al. [3] proposed a non-homogeneous time dependent Gompertz diffusion process with jumps to describe the evolution of a solid tumor subject to an intermittent therapeutic program. Moummou et al. [4] obtained explicit expressions for the maximum likelihood estimators with discrete sampling from the Gompertz diffusion model by using functional optimization orthogonal projections. However, the statistical properties of the model were not studied.

Ferrante et al. [1] studied maximum likelihood estimation of natural growth parameters of tumor for such models. However, they did not study distributional properties of the estimators. The knowledge of the distribution of the estimator may be applied to evaluate the distribution of other important growing parameters used to access tumor treatment modalities. We study distributional properties of approximate minimum contrast estimators of the unknown parameters in the model from discrete data with precise rates of convergence which are robust and efficient.

Let \( \Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0} \) be a stochastic basis on which is defined the Gompertz diffusion process \( \{X_t\}_{t \geq 0} \) satisfying the Itô stochastic differential equation

\[
dX_t = \left( \alpha X_t - \beta X_t \ln X_t \right) dt + \sigma X_t dW_t, \quad t \geq 0, \quad X_0 = x_0
\]

(1.1)

Where \( \{W_t\}_{t \geq 0} \) is a standard Brownian motion with the filtration \( \{\mathcal{F}_t\}_{t \geq 0} \) and \( \alpha > 0; \beta > 0; \sigma > 0 \) are the unknown parameters to be estimated on the basis of discrete observations of the process \( \{X_t\}_{t \geq 0} \). We assume equi-spaced sampling for simplicity. We assume two types of high frequency data: 1) \( T \to \infty, n \to \infty, \frac{T}{\sqrt{n}} \to 0 \) 2) \( T \to \infty, n \to \infty, \frac{T}{\sqrt{n}} \to 0 \).

Here \( X_t \) is the tumor volume which is measured at discrete time, \( \alpha \) is the intrinsic growth rate of the tumor, \( \beta \) is the tumor growth acceleration factor, and \( \sigma \) is the diffusion coefficient. Other parameters are the plateau of the model \( x_\alpha = \exp \left( \frac{\alpha T}{\sigma^2} \right) \) tumor growth decay, and the first time the growth curve of the model reaches \( x_\alpha \). We assume that the growth deceleration factor \( \beta \) does not change, while the variability of environmental conditions induces fluctuations in the intrinsic growth rate (mitosis rate) \( \alpha \).

In finance literature, this model is known as Black-Karasinski model which is a geometric mean reverting Vasicek model used for modeling term structure of interest rates. However, the statistical properties of the model were not studied.

Let the continuous realization be \( \{X_t\}_{0 \leq t \leq T} \) denoted by \( x^* \). Denote \( \Theta = (\alpha, \beta, \sigma) \). Let \( P_{\theta} \) be the measure generated on the space \( (C_0, \mathcal{B}_0) \) of continuous functions on \([0, T]\) with the associated Borel \( \sigma \) algebra \( \mathcal{B}_0 \) generated under the supremum norm by the process \( X^* \) and let \( \nu_* \) be the standard Wiener measure. It is well known that when \( \Theta \) is the true value of the parameter \( \nu_* \) is absolutely continuous with respect to \( \nu_* \) and the Radon-Nikodym derivative (likelihood) of \( \nu_* \) with respect to \( \nu_* \) based on the data \( x^* \) is given by

\[
\frac{dP_{\theta}}{d\nu_*} = \exp \left( \frac{-1}{2} \sigma^2 \theta T \right) \left( \frac{\sigma^2}{\pi} \right)^{1/2} \int \exp \left( \frac{-1}{2} \sigma^2 \theta \right) \left( \frac{\sigma^2}{\pi} \right)^{1/2} d\nu_*
\]

(2.1)

Where \( \theta = (\alpha, \beta, \sigma) \) and \( P_{\theta} \) is the probability measure induced by the diffusion process \( X^* \) with parameter \( \theta \) and \( \nu_* \) is the Lebesgue measure.
Consider the log-likelihood function, which is given by
\[ l_t(\theta) = \frac{d^2}{dx^2} \left( x_0^t \right) = -\frac{1}{\sigma^2} \left( (\alpha - \beta \ln x_0) x_0^t dx_0 - \frac{1}{2 \sigma^2} (\alpha - \beta \ln x_0)^2 dt \right) \tag{1:2} \]

A solution of the estimating equation \( \gamma(\theta) = 0 \) provides the conditional maximum likelihood estimators (MLEs)
\[ a_t = \frac{1}{T} \left\{ \int x_0^t dx_0 - \beta_t \int \ln x_0 dt \right\} \tag{1:4} \]
\[ \hat{\beta}_t = \frac{\sum_{i=1}^{n} (X_i^n - X_i^n)^2}{\sum_{i=1}^{n} X_i^n (t_i - t_{i-1})} \tag{1:6} \]

As an alternative to maximum likelihood method and to obtain robust estimators with higher efficiency we use contrast functions. Suppose \( \alpha \) and \( \sigma \) are known, for simplicity let \( \gamma = \frac{1}{2}, \sigma = 1, x_0 = 1 \) and our aim is to estimate the tumor growth acceleration parameter \( \beta \). Using Itô formula, the score function can be written as
\[ \gamma(\beta) = -\frac{1}{2} \int T \left( \frac{1}{\sigma^2} \right) \left( \frac{1}{\sigma^2} \right) dt \tag{1:7} \]

Using a contrast function which is related to the negative derivative of the log-likelihood function, we consider the estimating function
\[ M_T(\beta) = \frac{T}{2} \left( \int \ln x_0 dt \right) \tag{1:8} \]

Then the minimum contrast estimate (MCE) of \( \beta \) which is the solution of \( M_T(\beta) = 0 \) is given by
\[ \hat{\beta}_T = \frac{T}{2} \left( \int \ln x_0 dt \right) = \frac{T}{2T} \tag{1:9} \]

Where \( T = \int a_T \). Hence \( M_T(\beta) = \hat{\beta}_T - \frac{T}{2} \). Sometimes we will denote this by just \( M_T \). We find several discrete approximations of the MCE. Define a weighted approximation of \( L_1 \):
\[ K_n,T := \frac{\sum_{i=1}^{n} \ln x_{i-1} \ln x_{i-1} + \sum_{i=1}^{n} (1 - \omega_i) \ln x_{i-1} \ln x_{i-1}}{\sum_{i=1}^{n} \ln x_{i-1}} \tag{1:10} \]

Where \( w_i \geq 0 \) is a weight function. Denote the forward and backward approximations of \( L_1 \):
\[ I_{n,T} := \frac{\sum_{i=1}^{n} \ln x_{i-1}}{n} \tag{1:11} \]
\[ J_{n,T} := \frac{\sum_{i=1}^{n} \ln x_{i-1}}{n} \tag{1:12} \]

General weighted AMCE is defined as
\[ \hat{\beta}_{n,F} := \left( \frac{2}{T} \right) K_{n,T}^{-1} \tag{1:13} \]

With \( \omega_i = 1 \) in (1.10), we obtain the forward AMCE as
\[ \hat{\beta}_{n,F} := \left( \frac{2}{T} \right) I_{n,F}^{-1} \tag{1:14} \]

With \( \omega_i = 0 \) in (1.10), we obtain the backward AMCE as
\[ \hat{\beta}_{n,B} := \left( \frac{2}{T} \right) J_{n,B}^{-1} \tag{1:15} \]

With \( w_i(1) = 0.5 \) in (1.10), the simple symmetric AMCE is defined as
\[ \hat{\beta}_{n,S} := \left( \frac{2}{T} \right) \left( I_{n,F} + J_{n,B} \right)^{-1} = \left( \frac{2}{T} \sum \ln^2 x_{i-1} + 0.5(\ln^2 x_{i-1} + \ln^2 x_{i-1}) \right)^{-1} \tag{1:16} \]

Define the weighted symmetric estimators: With the weight function
\[ \omega_i = \left\{ \begin{array}{cl}
0 & \text{if } i = 1 \\\text{or } i = n+1 \\
1 & \text{if } i = n+1 \\
\end{array} \right. \tag{1:17} \]

the weighted symmetric AMCE is defined as
\[ \hat{\beta}_{n,W} := \left( \frac{2}{T} \right) \left( I_{n,F} + J_{n,B} \right)^{-1} \tag{1:18} \]

The AMCE has several good properties. The AMCE is simpler to calculate, in the sense that it does not involve simulation of a stochastic integral unlike AMLE. Hence AMCE is a more practical estimator. This is robust since M-estimator is reduced to the AMCE. The AMCE is efficient, Tanaka [5]. Tanaka [5] calculated the asymptotic relative efficiency of the minimum contrast estimator with respect to least squares estimator (LSE) and showed that MCE is asymptotically efficient while LSE is inefficient. We study the distributional properties of the AMCE. We obtain the rate of weak convergence to normal distribution of the AMCE using different normings. We also obtain stochastic bound on the difference of the AMCE. The AMCE is efficient, Tanaka [5]. Tanaka [5] calculated the asymptotic relative efficiency of the minimum contrast estimator with respect to least squares estimator (LSE) and showed that MCE is asymptotically efficient while LSE is inefficient. We study the distributional properties of the AMCE. We obtain the rate of weak convergence to normal distribution of the AMCE using different normings. We also obtain stochastic bound on the difference of the AMCE.
**Theorem 2.1**

Denote $b_{n,T} := O(\max(T^{2/3}(\log \log T)^{2/3} n^{-1/3}(\log \log T)^{-1}))$

(a) $\sup_{x \in R} \mathcal{P}\left(\frac{T}{2\beta} \sqrt{2} \left(\hat{\beta}_{n,T}-\beta\right) \leq x\right) - \mathcal{O}(x) = O(b_{n,T}^{-1})$

(b) $\sup_{x \in R} \mathcal{P}\left(\frac{1}{\sqrt{2}} \sqrt{\frac{T}{2\beta}} \left(\hat{\beta}_{n,T}-\beta\right) \leq x\right) - \mathcal{O}(x) = O(b_{n,T}^{-1})$

(c) $\sup_{x \in R} \mathcal{P}\left(\left(\hat{\beta}_{n,T}-\beta\right) - \frac{\mathcal{E}_{n,T} \mathcal{L}_{n,T}}{2} \leq x\right) - \mathcal{O}(x) = O(b_{n,T}^{-1})$

Proof: (a) Observe that $\left(\frac{T}{2\beta}\right)^{2/3} \left(\hat{\beta}_{n,T}-\beta\right) = \frac{\mathcal{E}_{n,T}^{2/3} \mathcal{L}_{n,T}^{2/3}}{2}$ (2.1)

Thus, we have $\mathcal{I}$ Hence

\[
\left(\frac{T}{2\beta}\right)^{2/3} \left(\hat{\beta}_{n,T}-\beta\right) = \frac{\mathcal{E}_{n,T}^{2/3} \mathcal{L}_{n,T}^{2/3}}{2}
\]

(2.2)

Further,

\[
P\left(\left(\frac{2\beta}{T}\right) \left|\mathcal{I}_{n,T} \lambda - 1\right| > \epsilon\right) = \mathcal{P}\left(\left(\frac{2\beta}{T}\right) \left|\mathcal{I}_{n,T} \lambda - 1 + \mathcal{I}_{n}\right| > \epsilon\right)
\]

(2.6)

We have from (2.3),

\[
U_1 \leq C T^{-1/2} + \mathcal{C}_{18} \mathcal{T}^{2/3} + C \mathcal{T}^{2/3} n^{-1/2} + \mathcal{E}
\]

Further,

\[
U_2 = \mathcal{P}\left(\left(\frac{2\beta}{T}\right) \left|\mathcal{I}_{n,T} \lambda - 1\right| > \epsilon\right) = \mathcal{P}\left(\left(\frac{2\beta}{T}\right) \left|\mathcal{I}_{n,T} \lambda - 1 + \mathcal{I}_{n}\right| > \epsilon\right)
\]

\[
\leq \mathcal{P}\left(\left(\frac{2\beta}{T}\right) \left|\mathcal{I}_{n,T} \lambda - 1\right| > \epsilon\right) + \mathcal{P}\left(\left(\frac{2\beta}{T}\right) \left|\mathcal{I}_{n,T} \lambda - 1\right| > \delta\right)
\]

(2.7)
\begin{equation}
\delta = C e^{-C e^2} + \delta_i, \quad \text{where } \delta_i > 0.
\end{equation}

(2.8)

Here, the bound for the first term in the right hand side of (2.7) comes from Lemma 2.1(c) and that for the second term is obtained from (2.3) [2]. Now, using the bounds (2.7) and (2.8) in (2.6) with \( \delta = CT \sqrt{T \log T} \), we obtain that the terms in (2.6) are of the order \( O(1) \). Hence, upon choosing \( nT^4 \rightarrow \infty \), we obtain the part (b) of Theorem follows.

Then
\begin{equation}
(2 \beta_{nT,F} - \beta)^{1/2} = (2 \beta)^{1/2} \left( 1 - \frac{\beta - \beta_{nT,F}}{\beta} \right)^{1/2} = (2 \beta)^{1/2} \left[ 1 + \frac{\beta - \beta_{nT,F}}{\beta} \right] + O(d^2)
\end{equation}

Further, we obtain
\begin{equation}
P(G_T^c) = P \left( \beta_{nT,F} - \beta > C \sqrt{T \log T} \right) \leq C \sqrt{T \log T} \left( 2 \beta \right)^{1/2}
\end{equation}

Choosing \( \epsilon = \frac{T^{1/2}}{n^2} \), the theorem follows.

The following theorem gives stochastic bound on the error of approximation of the continuous MCE by AMCEs.

(3.1)

(a) \( \beta_{nT,F} - \beta = O_p \left( \frac{T^{1/2}}{n^2} \right) \), (b) \( \beta_{nT,F} - \beta = O_p \left( \frac{T^{1/2}}{n^2} \right) \).

Choosing \( \epsilon = \frac{T^{1/2}}{n^2} \), the theorem follows.

\begin{thebibliography}{9}


How to cite this article: Jaya PN Bishwal. Robust Estimation in Gompertz Diffusion Model of Tumor Growth. Open Acc Biostat Bioinform. 1(4). OABB.000521.
