*Journal of Statistical Research* *ISSN 0256 - 422 X*
*2008, Vol. 42, No. 1, pp. 85-116*
*Bangladesh*

# THE GENERALIZED LOG–GAMMA MIXTURE MODEL WITH COVARIATES

Edwin M. M. Ortega

*Departamento de Ciencias Exatas, USP Av. Pàdua Dias 11 - Caixa Postal 9*
*13418-900 Piracicaba - São Paulo, Brasil*

*Email: edwin@esalq.usp.br*

Fernanda. B. Rizzato

*ESALQ, Universidade de São Paulo, Piracicaba, Brasil*

*Email: frizato@esalq.usp.br*

Clarice G. B. Demètrio

*ESALQ, Universidade de São Paulo, Piracicaba, Brasil*

*Email: clarice@esalq.usp.br*

SUMMARY

In this paper the generalized log-gamma model is modified for possibility that long-term survivors may be present in the data. The model attempts to separately estimate the effects of covariates on the acceleration/deceleration of the timing of a given event and surviving fraction, that is, the proportion of the population for which the event never occurs. The logistic function is used for the regression model of the surviving fraction. We consider maximum likelihood and Jackknife estimators for the parameters of the model. We derive the appropriate matrices for assessing local influence on the parameter estimates under different perturbation schemes and we also present some ways to perform global influence. Finally, a data set from the medical area is analyzed under the log-gamma generalized mixture model. A residual analysis is performed in order to select an appropriate model.

*Keywords and phrases:* Logistic model; mixture models; log-gamma generalized distribution; local influence; global influence censored data; residual analysis.

*AMS Classification:* Place Classification here. Leave as is, if there is no classification

© Institute of Statistical Research and Training (ISRT), University of Dhaka, Dhaka 1000, Bangladesh.

# 1 Introduction

Models for survival analysis typically assume that every subject in the study population is susceptible to the event under study and will eventually experience such event if the follow-up is sufficiently long. However, there are situations when a fraction of individuals are not expected to experience the event of interest, that is, those individuals are cured or not susceptible. For example, researchers may be interested in analyzing the recurrence of a disease. Many individuals may never experience a recurrence; therefore, a cured fraction of the population exists. Cure rate models have been used to estimate the cured fraction. Cure rate models are survival models which allow for a cured fraction of individuals. These models extend the understanding of time-to-event data by allowing for the formulation of more accurate and informative conclusions. These conclusions are otherwise unobtainable from an analysis which fails to account for a cured or insusceptible fraction of the population. If a cured component is not present, the analysis reduces to standard approaches of survival analysis. Use of cure rate models has been used for modelling time-to-event data for various types of cancers, including breast cancer, non-Hodgkin's lymphoma, leukaemia, prostate cancer and melanoma. Perhaps the most popular type of cure rate models is the mixture model introduced by Berkson and Gage (1958) and Maller and Zhou (1996). In this model, the population is divided into two subpopulation so that an individual either is cured with probability $p$, or has a proper survival function $S(t)$, with probability $1-p$. This gives an improper population survivor function $G(t)$ in the form of mixture, that is,

$$G(t) = p + (1-p)S(t), \quad S(\infty) = 0, \quad G(\infty) = p, \tag{1.1}$$

Common choices for $S(t)$ in (1) are the exponential and Weibull distributions which are particular cases of the family of generalized log-gamma distribution. With those choices we have respectively the exponential mixture model and Weibull. Mixture models involving these distributions have been studied by several authors, including Farewell (1982), Goldman (1984), Greenhouse (1998) and Sy and Taylor (2000). The book by Maller and Zhou (1996) provides a wide range of applications of the long-term survivor mixture model.

Influence diagnostic is an important step in the analysis of a data set as it provides an indication of bad model fitting or of influential observations. However, there are not applications of influence diagnostic to the mixture models. Cook (1986) proposed a diagnostic approach named local influence to assess the effect of small perturbations in the model and/or data on the parameter estimates. Several authors have applied the local influence methodology in more general regression models than the normal regression model (see, for example, Paula 1993, Galea et al., 2002 and Daz-Garca, et al., 2003). Also, some authors have investigated the assessment of local influence in survival analysis models: for instance, Pettit and Bin Daud (1989) have investigated local influence in proportional hazard regression models; Escobar and Meeker (1992) have adapted local influence methods to regression analysis with censoring; Ortega et al. (2003) have considered the problem of assessing local influence in generalized log-gamma regression models with censored observations and Ortega et al. (2006) have derived the curvature calculations under various perturbation schemes in log-Weibull-exponentiated regression models with censored data. In this paper, we consider the generalized log-gamma mixture model with covariates, where the covariates are modelled through the logistic function. In section 2, we present the generalized log-gamma mixture, in addition with the maximum likelihood and Jackknife estimators. The score functions and observed Fisher information matrix are given as well as the process for estimating the regression coefficients and the remaining parameter is discussed. In Section 3 we use several diagnostic measures considering case-deletion and the normal curvatures of local influence are derived under various perturbation schemes in the generalized log-gamma mixture. Section 4, two kinds of deviance-type residuals are proposed. An application with real data, which have not been analyzed from a diagnostic perspective, is discussed in section 5. The last section deals with some concluding remarks.

## 2 The Generalized Log–Gamma Mixture with Covariates

Consider first the uncensored case and the generalized log-gamma model

$$Y = \mu + \sigma Z, \tag{2.1}$$

where $\mu \in \mathbf{R}$ is a location parameter, $\sigma > 0$ is a scale parameter, and $Z$ follows a distribution with probability density function given by

$$f(z;q) = \begin{cases} \frac{|q|}{\Gamma(q^{-2})}(q^{-2})^{q^{-2}}\exp\{q^{-1}z - q^{-2}\exp(qz)\}, & \text{if} \quad q \neq 0 \\ \frac{1}{\sqrt{2\pi}}\exp(-\frac{z^2}{2}), & \text{if} \quad q = 0, \end{cases} \tag{2.2}$$

where $\boldsymbol{\theta} = (\mu, \sigma, q)^T$ and $z = (y-\mu)/\sigma$. The extreme value distribution is a particular case of (2.1), when $q = 1$. With the parameterisation given in (2.1) the maximum likelihood estimation and large-sample methods are less complicated than the parameterisation given in Lawless (2003, Section 5.3). Consider now the censored case with the assumption of uninformative censoring. Let $y$ be either the observed log-lifetime or log-censoring time for the individual. The survivor function $S(y;\boldsymbol{\theta})$ assumes the following forms:

$$S(y;\boldsymbol{\theta}) = \begin{cases} Q\big[q^{-2}, q^{-2}\exp\{q(\frac{y-\mu}{\sigma})\}\big], & \text{if} \quad q > 0 \\ 1 - Q\big[q^{-2}, q^{-2}\exp\{q(\frac{y-\mu}{\sigma})\}\big], & \text{if} \quad q < 0 \\ 1 - \Phi(\frac{y-\mu}{\sigma}), & \text{if} \quad q = 0, \end{cases}$$

with $\Phi(z)$ denoting the standard normal cumulative distribution, while

$$Q(k,a) = \int_a^\infty \frac{x^{k-1}e^{-x}}{\Gamma(k)}dx \tag{2.3}$$

is the incomplete gamma integral and $\Gamma(q)$ the gamma function.

Thus given a sample $y_1, y_2 \ldots, y_n$, where $y_i$ is either the observed log-lifetime or log-censuring time for the ith individual, let a binary random variable $b_i$, $i = 1, \ldots, n$ indicating that the $i$th individual in a population is at risk or not with respect to a certain type of failure, that is, $b_i = 1$ indicates that the $i$th individual will eventually experience a failure event (uncured) and $b_i = 0$ indicates that the individual will never experience such event (cured). For an individual with covariate vector $\mathbf{x}_i$, the proportion of uncured $p$ can be specified to be the logistic link of $\mathbf{x}$ such that the conditional distribution of $b$ is given by

$$Pr\Big(b_i = 1|\mathbf{x}_i\Big) = \frac{1}{1+\exp(\mathbf{x}_i^T\boldsymbol{\beta})} = 1 - p_i$$

where $\boldsymbol{\beta} = (\beta_1, \beta_2, \ldots, \beta_p)^T$ is a vector $p$-dimensional parameters. Observe that the cured probability varies from individual to individual, so that the probability that individual $i$ is cured is modelled by

$$p_i = \frac{\exp(\mathbf{x}_i^T\boldsymbol{\beta})}{1+\exp(\mathbf{x}_i^T\boldsymbol{\beta})}, \tag{2.4}$$

and the logistic link keeps each $p_i$ strictly between 0 and 1. Suppose that $Y_i$s are independent and identically distributed with generalized log-gamma distribution with density function is given by

$$f(y_i;\mu,\sigma,q|Y_i=1) = \begin{cases} \frac{|q|}{\Gamma(q^{-2})\sigma}(q^{-2})^{q^{-2}}\exp\left\{q^{-1}z - exp(qz)\right\} & \text{if} \quad q \neq 0 \\ \frac{1}{\sqrt{2\pi}}\exp\{-\frac{z^2}{2}\}, & \text{if} \quad q = 0. \end{cases} \tag{2.5}$$

where $\mu \in \mathbf{R}$ is the location parameter, $q \in \mathbf{R}$ is the shape parameter, $\sigma > 0$ is the scale parameter and $z = (y-\mu)/\sigma$. Figure 1 presents graphs for the density (2.5).

The maximum likelihood method is used to estimate the parameters. Thus, the contribution of an individual that failed at $y_i$ to the likelihood function is given by

$$\begin{cases} \frac{(1-p_i)|q|}{\Gamma(q^{-2})\sigma}(q^{-2})^{q^{-2}}\exp\left\{q^{-1}z_i - \exp(qz_i)\right\} & \text{if} \quad q \neq 0 \\ \frac{(1-p_i)}{\sqrt{2\pi}\sigma}\exp\{-\frac{z_i^2}{2}\}, & \text{if} \quad q = 0. \end{cases}$$

and the contribution of an individual that is at risk at time $t_i$ is

$$\begin{cases} p_i + (1-p_i)Q[q^{-2}, q^{-2}\exp\{qz_i\}], & \text{if} \quad q > 0 \\ p_i + (1-p_i)\Big(1 - Q[q^{-2}, q^{-2}exp\{qz_i\}]\Big), & \text{if} \quad q < 0 \\ p_i + (1-p_i)\Big[1 - \Phi(z_i)\Big], & \text{if} \quad q = 0, \end{cases}$$

Thus the log-likelihood function corresponding to the parameter vector $\boldsymbol{\theta} = (\mu, \sigma, \beta_1, \beta_2, \ldots, \beta_p)^T$

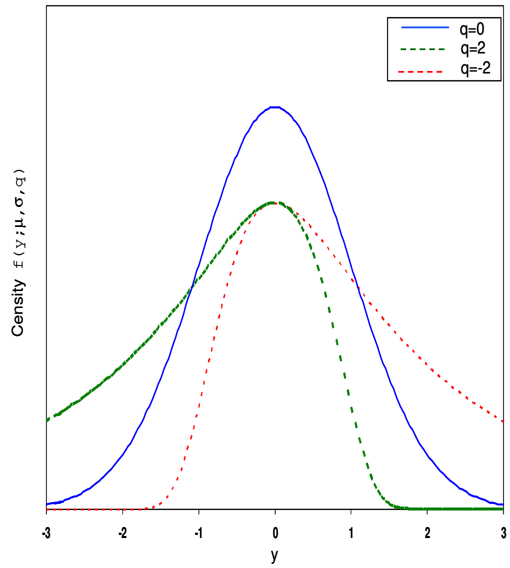


Figure 1: Graphs of the generalized log-gamma distribution for some values of q and by assuming $\mu = 0$ and $\sigma = 1$.

is given by

$$l(\boldsymbol{\theta}) = \begin{cases} \sum\limits_{i \in F} \log\left[\frac{(1-p_i)q}{\Gamma(q^{-2})\sigma}(q^{-2})^{q^{-2}} \exp\left\{q^{-1}z_i - q^{-2}\exp(qz_i)\right\}\right] & \\ + \sum\limits_{i \in C} \log\Big[p_i + (1-p_i)Q[q^{-2}, q^{-2}\exp\{qz_i\}]\Big], & \text{if} \quad q > 0 \\ \sum\limits_{i \in F} \log\left[\frac{(1-p_i)(-q)}{\Gamma(q^{-2})\sigma}(q^{-2})^{q^{-2}} \exp\left\{q^{-1}z_i - q^{-2}\exp(qz_i)\right\}\right] & \\ + \sum\limits_{i \in C} \log\Big[p_i + (1-p_i)\Big\{1 - Q[q^{-2}, q^{-2}\exp\{qz_i\}]\Big\}\Big], & \text{if} \quad q < 0 \\ \sum\limits_{i \in F} \log\left[\frac{(1-p_i)}{\sqrt{2\pi}\sigma}\exp\{-\frac{z_i^2}{2}\}\right] + \sum\limits_{i \in C} log\Big[p_i + (1-p_i)\{1 - \Phi(z_i)\}\Big], & \text{if} \quad q = 0 \end{cases} \tag{2.6}$$

where F and C denote, respectively, that the set of individuals is a log-lifetime or a log-censoring time, $\Phi(.)$ denotes the standard normal cumulative distribution, $Q(.)$ is the incomplete gamma integral, $r$ is the number of uncensored observations (failures), $\mathrm{p}_i =$ is defined by (2.4) and $z_i = \frac{y_i - \mu}{\sigma}$. The maximization of (2.6) follows two steps for obtaining the maximum likelihood estimates of $\boldsymbol{\theta}$. Since, in general, it is reasonable to expect that shape parameter $q$ belongs to interval $[-3, 3]$, we fixed in the first step of the iterative process different $q$ values in this interval. Then, assuming q fixed we find the maximum likelihood estimates $\tilde{\beta}(q)$, $\tilde{\mu}(q)$ and $\tilde{\sigma}(q)$ and the maximized log-likelihood function $l_{\max}(q)$ is determined, using the subroutine MAXBFGS available in Ox (see, for instance, Doornik, 1996). In the second step, the log-likelihood $l_{\max}(q)$ is maximized and then $\hat{q}$ is obtained. The maximum likelihood estimates of $\beta$, $\mu$ and $\sigma$ are, respectively, given by $\hat{\beta} = \tilde{\beta}(q)$, $\hat{\mu} = \tilde{\mu}(q)$ and $\hat{\sigma} = \tilde{\sigma}(q)$.

Covariance estimates for the maximum likelihood estimators $\widehat{\boldsymbol{\theta}}$ can also be obtained using the Hessian matrix. Confidence intervals and hypothesis testing can be conducted by using the large sample distribution of MLE which is a normal distribution with the covariance matrix as the inverse of Fisher information since regularity conditions are satisfied $\left[\widehat{\boldsymbol{\theta}} \sim N_{p+2}\left\{\boldsymbol{\theta}, -\mathbf{I}(\boldsymbol{\theta})^{-1}\right\}\right]$, where the asymptotic covariance matrix is given by $\mathbf{I}^{-1}(\boldsymbol{\theta})$ with $\mathbf{I}(\boldsymbol{\theta}) = -E[\ddot{\mathbf{L}}_{\theta\theta}]$ such that $\ddot{\mathbf{L}}_{\theta\theta} = \left\{\frac{\partial^2 l(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^T}\right\}$.

Since it is not possible to compute the Fisher information matrix $\mathbf{I}(\boldsymbol{\theta})$ due to the censored observations (censoring is random and noninformative), it is possible to use the matrix of second derivatives of the log likelihood, $-\ddot{\mathbf{L}}_{\theta\theta}$, evaluated at MLE $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$, which is consistent. Then

$$\ddot{\mathbf{L}}_{\theta\theta)} = \begin{pmatrix} \mathbf{L}_{\mu\mu} & \mathbf{L}_{\mu\sigma} & \mathbf{L}_{\mu\beta} \\ . & \mathbf{L}_{\sigma\sigma} & \mathbf{L}_{\sigma\beta} \\ . & . & \mathbf{L}_{\beta\beta} \end{pmatrix}$$

with the submatrices given in appendix A.

## 2.1 Jackknife Estimator

The idea of jackknifing is to transform the problem of estimating any population parameter into the problem of estimating a population mean. So, what is done when estimating a mean value is realized in this method but from an unusual point of view. An important work of implementing the jackknife method is given by Lipsitz et al. (1990) that suggest an alternative robust estimator of the covariance matrix based on the jackknife for analyzing data from repeated measures studies. In this paper, we use this method as an alternative to estimate the population parameters.

Suppose that $T_1, T_2, \ldots, T_n$ is a random sample of $n$ values and the sample mean is given by

$$\bar{T} = \sum_{i=1}^{n} \frac{T_i}{n},$$

that is used to estimate the mean of the population.

Now, it is calculated the sample mean with the $lth$ observation missed out,

$$\bar{T}_{-l} = \frac{\sum_{i=1}^{n} T_i - T_l}{n-1}.$$

Then, we obtain

$$T_l = n\hat{T} - (n-1)\bar{T}_{-l}. \tag{2.7}$$

In a general situation, consider that $\theta$ is a parameter estimated by $\hat{E}(T_1, T_2, \ldots, T_n)$ and for ease of notation drop $(T_1, T_2, \ldots, T_n)$. Finally, it is calculated $\hat{E}_{-l}$ that is obtained with the $T_l$ observation missed out. It follows, from (2.7) that pseudo-values can be calculated

$$\hat{E}_l^* = n\hat{E} - (n-1)\hat{E}_{-l}, \quad l = 1, \ldots, n.$$

The average of the pseudo-values is given by

$$\hat{E}^* = \frac{\sum_{l=1}^{n} \hat{E}_l^*}{n}$$

that is the Jackknife estimate of $\theta$.

Manly (1997) suggests that an approximate $100(1-\alpha)\%$ confidence interval for $\theta$ is given by $\hat{E}^* \pm t_{\alpha/2,n-1} s/\sqrt{n}$, where $t_{\alpha/2,n-1}$ is the value that is exceeded with probability $\alpha/2$ for the t distribution with $(n-1)$ degrees of freedom and the Jackknife estimator has the effect of removing bias of order $1/n$.

The Jackknife estimate calculations for the generalized log-gamma mixture with covariates regression model are realized for $\mu$, $\sigma$ and $\beta_j$ $(j = 1, \ldots, p)$ and confidence intervals are calculated separately for each parameter.

# 3 Sensitivity Analysis

## 3.1 Global Influence

The first tool to perform sensitivity analysis as stated before is by means of global influence starting from case-deletion (see, Cook, 1977). Case-deletion is a common approach to study the effect of dropping the *ith* case from the data set. The case-deletion for the model (2.1) is given by

$$Y_l = \mu + \sigma Z_l, \quad l = 1, 2, \ldots, n, \quad l \neq i. \tag{3.1}$$

In the following, a quantity with subscript "(i)" means the original quantity with the *ith* case deleted. For the model (3.1), the log-likelihood function is denoted by $l_{(i)}(\boldsymbol{\theta})$.

Let $\hat{\boldsymbol{\theta}}_{(i)} = (\hat{\mu}_{(i)}, \hat{\sigma}_{(i)}, \hat{\boldsymbol{\beta}}_{(i)}^T)^T$ be the ML estimate of $\boldsymbol{\theta}$ from $l_{(i)}(\boldsymbol{\theta})$. To assess the influence of the *ith* case on the ML estimate $\hat{\boldsymbol{\theta}} = (\hat{\mu}, \hat{\sigma}, \hat{\boldsymbol{\beta}}^T)^T$, the basic idea is to compare the difference between $\hat{\boldsymbol{\theta}}_{(i)}$ and $\hat{\boldsymbol{\theta}}$. If deletion of a case seriously influences the estimates, more attention should be paid to that case. Hence, if $\hat{\boldsymbol{\theta}}_{(i)}$ is far from $\hat{\boldsymbol{\theta}}$, then the case is regarded as an influential observation. A first measure of global influence is defined as the standardized norm of $\hat{\boldsymbol{\theta}}_{(i)} - \hat{\boldsymbol{\theta}}$ (generalized Cook distance)

$$GD_i(\boldsymbol{\theta}) = (\hat{\boldsymbol{\theta}}_{(i)} - \hat{\boldsymbol{\theta}})^T \big\{ -\ddot{\mathbf{L}}(\boldsymbol{\theta}) \big\} (\hat{\boldsymbol{\theta}}_{(i)} - \hat{\boldsymbol{\theta}}).$$

Other alternative is to assess the values $GD_i(\boldsymbol{\beta})$ and $GD_i(\mu, \sigma)$, which reveal the impact of the $ith$ case on the estimates of $\boldsymbol{\beta}$ and $(\mu_1, \sigma)$, respectively. Another popular measure of the difference between $\hat{\boldsymbol{\theta}}_{(i)}$ and $\hat{\boldsymbol{\theta}}$ is the likelihood displacement

$$LD_i(\boldsymbol{\theta}) = 2\Big\{l(\hat{\boldsymbol{\theta}}) - l(\hat{\boldsymbol{\theta}}_{(i)})\Big\}.$$

Besides, we can also compute $\beta_j - \beta_{j(i)} (j = 1, 2, \ldots, p)$ to see the difference between $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\beta}}_{(i)}$. Alternative global influence measures are possible. One could think of the behavior of a test statistics, such as a Wald test for covariates or censoring effect, under a case deletion scheme.

To avoid the direct model estimation for all observations we can use the following one-step approximation to reduce the burden:

$$\hat{\boldsymbol{\theta}}_{(i)} = \hat{\boldsymbol{\theta}} + \ddot{\mathbf{L}}(\hat{\boldsymbol{\theta}})^{-1} \dot{l}_{(i)}(\hat{\boldsymbol{\theta}}),$$

where $\dot{l}_i(\hat{\boldsymbol{\theta}}) = \frac{\partial l_{(i)}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}$ is evaluated at $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$ (see, Cook and Weisberg, 1982).

## 3.2 Local Influence

Local influence calculation can be carried out in model (2.5). If likelihood displacement$LD(\boldsymbol{\omega}) = 2\{l(\hat{\boldsymbol{\theta}}) - l(\hat{\boldsymbol{\theta}}_{\boldsymbol{\omega}})\}$ is used, where $\hat{\boldsymbol{\theta}}_{\boldsymbol{\omega}}$ denotes the MLE under the perturbed model, the normal curvature for $\boldsymbol{\theta}$ at direction $\boldsymbol{\ell}, \| \boldsymbol{\ell} \| = 1$, is given by $C_{\boldsymbol{\ell}}(\boldsymbol{\theta}) = 2|\boldsymbol{\ell}^T \boldsymbol{\Delta}^T \ddot{\mathbf{L}}_{\theta\theta}^{-1} \boldsymbol{\Delta} \boldsymbol{\ell}|$, where $\boldsymbol{\Delta}$ is a $(p+2)n$ matrix that depends on the perturbation scheme and whose elements are given by $\Delta_{ij} = \partial^2 l(\boldsymbol{\theta}|\boldsymbol{\omega}) / \partial\theta_i \partial\omega_j$, $i = 1, 2, \ldots, p+2$ and $j = 1, 2, \ldots, n$ evaluated at $\hat{\boldsymbol{\theta}}$ and $\boldsymbol{\omega}_0$, where $\boldsymbol{\omega}_0$ is the no perturbation vector (see Cook, 1986). For the generalized log-gamma mixture model the elements of $\ddot{\mathbf{L}}_{\hat{\theta}\hat{\theta}}$ are given in appendix A. We can also calculate normal curvatures $C_{\boldsymbol{\ell}}(\boldsymbol{\beta})$,$C_{\boldsymbol{\ell}}(\mu)$ and $C_{\boldsymbol{\ell}}(\sigma)$ to perform various index plots, for instance, the index plot of $l_{max}$, the eigenvector corresponding to $C_{\boldsymbol{\ell}_{max}}$, the largest eigenvalue of the matrix $\mathbf{B} = -\boldsymbol{\Delta}^T \ddot{\mathbf{L}}_{\theta\theta}^{-1} \boldsymbol{\Delta}$ and the index plots of $C_{\boldsymbol{\ell}_i}(\boldsymbol{\beta})$, $C_{\boldsymbol{\ell}_i}(\mu)$ and $C_{\boldsymbol{\ell}_i}(\sigma)$ named total local influence (see, for example, Lesaffre and Verbeke, 1998), where $\boldsymbol{\ell}_i$ denotes an $n1$ vector of zeros with one at the $i-th$ position. Thus, the curvature at direction $\boldsymbol{\ell}_i$ assumes the form $C_i = 2|\boldsymbol{\Delta}_i^T \ddot{\mathbf{L}}_{\theta\theta}^{-1} \boldsymbol{\Delta}_i|$ where $\boldsymbol{\Delta}_i^T$ denotes the $ith$ row of $\boldsymbol{\Delta}$. It is usual to point out those cases such that

$$C_i \geq 2\bar{C}, \quad \bar{C} = \frac{1}{n} \sum_{i=1}^{n} C_i. \tag{3.2}$$

## 3.3 Curvature Calculations

Next, we calculate, for three perturbation schemes, the matrix

$$\boldsymbol{\Delta} = (\boldsymbol{\Delta}_{ji})_{(p+2)\times n} = \left( \frac{\partial^2 l(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial\theta_i \boldsymbol{\omega}_j} \right)_{(p+2)\times n}, \quad j = 1, 2, \ldots, p+2 \quad \text{and} \quad i = 1, 2, \ldots, n, \tag{3.3}$$

considering the model defined in (2.4) and its log-likelihood function given by (2.6).

### 3.3.1 Case–Weights Perturbation

Consider the vector of weights $\boldsymbol{\omega} = (\omega_1, \omega_2, \ldots, \omega_n)^T$.

- $q > 0$

  In this case the log-likelihood function takes the form

$$l(\boldsymbol{\theta}|\boldsymbol{\omega}) = \sum_{i\in F} \omega_i \log\Big[\frac{(1-p_i)q}{\Gamma(q^{-2})\sigma}(q^{-2})^{q^{-2}} \exp\big\{q^{-1}z_i - q^{-2}\exp(qz_i)\big\}\Big] + \sum_{i\in C} \omega_i \log\Big[p_i + (1-p_i)Q[q^{-2}, q^{-2}\exp\{qz_i\}]\Big]. \tag{3.4}$$

- $q < 0$

  Here the weighted log-likelihood function becomes expressed in the form

$$l(\boldsymbol{\theta}|\boldsymbol{\omega}) = \sum_{i\in F} \omega_i \log\Big[\frac{(1-p_i)(-q)}{\Gamma(q^{-2})\sigma}(q^{-2})^{q^{-2}} \exp\big\{q^{-1}z_i - q^{-2}\exp(qz_i)\big\}\Big] + \sum_{i\in C} \omega_i \log\Big[p_i + (1-p_i)\Big\{1 - Q[q^{-2}, q^{-2}\exp\{qz_i\}]\Big\}\Big]. \tag{3.5}$$

- $q = 0$

  In this case the weighted log-likelihood function takes the form

$$l(\boldsymbol{\theta}|\boldsymbol{\omega}) = \sum_{i\in F} \omega_i \log\Big[\frac{(1-p_i)}{\sqrt{2p}\sigma}\exp\{-z_i^2/2\}\Big] + \sum_{i\in C} \omega_i \log\Big[p_i + (1-p_i)\{1-\Phi(z_i)\}\Big], \tag{3.6}$$

  where $0 \le \omega_i \le 1$ and $\boldsymbol{\omega} = (1, \ldots, 1)^T$. The elements of matrix $\boldsymbol{\Delta} = (\boldsymbol{\Delta}_1, \ldots, \boldsymbol{\Delta}_{p+2})^T$ in appendix B.

### 3.3.2 Response Perturbation

We will consider here that each $y_i$ is perturbed as $y_{iw} = y_i + \omega_i S_y$, where $S_y$ is a scale factor that may be the estimated standard deviation of $Y$ and $\omega_i \in \mathbf{R}$.

- $q > 0$

  Here the perturbed log-likelihood function becomes expressed as

$$l(\boldsymbol{\theta}|\boldsymbol{\omega}) = \sum_{i\in F} \log\Big[\frac{(1-p_i)q}{\Gamma(q^{-2})\sigma}(q^{-2})^{q^{-2}} \exp\big\{q^{-1}z_i^* - q^{-2}\exp(qz_i^*)\big\}\Big] + \sum_{i\in C} \log\Big[p_i + (1-p_i)Q[q^{-2}, q^{-2}\exp\{qz_i^*\}]\Big]. \tag{3.7}$$

- $q < 0$

  In this case the perturbed log-likelihood function is expressed as

$$\begin{aligned} l(\boldsymbol{\theta}|\boldsymbol{\omega}) &= \sum_{i \in F} \log \Big[ \frac{(1-p_i)(-q)}{\Gamma(q^{-2})\sigma} (q^{-2})^{q^{-2}} \exp\left\{q^{-1} z_i^* - q^{-2} \exp(q z_i^*)\right\} \Big] \\ &+ \sum_{i \in C} \log \Big[ p_i + (1-p_i) \Big\{ 1 - Q[q^{-2}, q^{-2} \exp\{q z_i^*\}] \Big\} \Big]. \end{aligned} \quad (3.8)$$

- $q = 0$

  In this case the weighted log-likelihood function takes form

$$l(\boldsymbol{\theta}|\boldsymbol{\omega}) = \sum_{i \in F} \log \Big[ \frac{(1-p_i)}{\sqrt{2p}\sigma} \exp\{-\frac{z_i^{2*}}{2}\} \Big] + \sum_{i \in C} \log \Big[ p_i + (1-p_i)\{1 - \Phi(z_i^*)\} \Big], \quad (3.9)$$

where $z_i^* = \frac{(y_i + \omega_i S_y) - \mu}{\sigma}$. The elements of matrix $\boldsymbol{\Delta} = (\boldsymbol{\Delta}_1, \ldots, \boldsymbol{\Delta}_{p+2})^T$ in appendix C.

### 3.3.3 Explanatory Variable Perturbation

Consider now an additive perturbation on a particular continuous explanatory variable, namely $X_t$, by making $x_{it\omega} = x_{it} + \omega_i S_t$, where $S_t$ is a scaled factor, $\omega_i \in \mathbf{R}$. This perturbation scheme leads to the following expressions for the log-likelihood function and for the elements of matrix $\boldsymbol{\Delta}$:

- $q > 0$

  In this case the log-likelihood function takes the form

$$\begin{aligned} l(\boldsymbol{\theta}|\boldsymbol{\omega}) &= \sum_{i \in F} \log \Big[ \frac{(1-p_i^*)q}{\Gamma(q^{-2})\sigma} (q^{-2})^{q^{-2}} \exp\left\{q^{-1} z_i - q^{-2} \exp(q z_i)\right\} \Big] \\ &+ \sum_{i \in C} \log \Big[ p_i^* + (1-p_i^*) Q[q^{-2}, q^{-2} \exp\{q z_i\}] \Big]. \end{aligned} \quad (3.10)$$

- $q < 0$

  In this case the perturbed log-likelihood function is expressed as

$$\begin{aligned} l(\boldsymbol{\theta}|\boldsymbol{\omega}) &= \sum_{i \in F} \log \Big[ \frac{(1-p_i^*)(-q)}{\Gamma(q^{-2})\sigma} (q^{-2})^{q^{-2}} \exp\left\{q^{-1} z_i - q^{-2} \exp(q z_i)\right\} \Big] \\ &+ \sum_{i \in C} \log \Big[ p_i^* + (1-p_i^*) \Big\{ 1 - Q[q^{-2}, q^{-2} \exp\{q z_i\}] \Big\} \Big]. \end{aligned} \quad (3.11)$$

- $q = 0$

  In this case the weighted log-likelihood function takes form

$$l(\boldsymbol{\theta}|\boldsymbol{\omega}) = \sum_{i\in F} \log\Big[\frac{(1-p_i^*)}{\sqrt{2p}\sigma}\exp\{-\frac{z_i^2}{2}\}\Big] + \sum_{i\in C}\log\Big[p_i^* + (1-p_i^*)\{1-\Phi(z_i)\}\Big], \quad (3.12)$$

  where $p_i^* = [1+\exp(-\mathbf{x}_i^{*T}\boldsymbol{\beta})]^{-1}$ and $\mathbf{x}_i^{*T} = \beta_1+\beta_2 x_{i2}+\cdots+\beta_t(x_{it}+\omega_i S_t)+\cdots+\beta_p x_{ip}$. The elements of matrix $\boldsymbol{\Delta} = (\boldsymbol{\Delta}_1,\ldots,\boldsymbol{\Delta}_{p+2})^T$ in appendix D.

# 4 Residual Analysis

In order to study departures from the error assumptions as well as the presence of outlying observations, we will consider two kinds of residuals: deviance component residual (see, for instance, McCullagh & Nelder, 1989) and martingale-type residual (see, for instance, Barlow & Prentice, 1988 and Therneau, Grambsch & Fleming, 1990). More details may be found in Ortega, Bolfarine & Paula (2003).

## 4.1 Martingale–Type and Deviance Component Residual

Therneau, Grambsch and Fleming (1990) introduced the deviance component residual in the counting process area by using basically martingale residuals. For example, the deviance component residual for the Cox model with no time-dependent covariates may be described as

$$r_{D_i} = \mathrm{sgn}(r_{M_i})\Big\{-2\big[r_{M_i} + \delta_i\log(\delta_i - r_{M_i})\big]\Big\}^{1/2}, \quad (4.1)$$

where $r_{M_i}$ is the martingale residual. In the generalized log-gamma mixture model on which we are working, the martingale residual can be described as

- $q > 0$

$$r_{M_i} = \begin{cases} 1+\log\Big\{\hat{p}+(1-\hat{p})Q\big[\hat{q}^{-2},\hat{q}^{-2}\exp\{\hat{q}\hat{z}_i\}\big]\Big\} & \text{if} \quad i\in F \\ \log\Big\{\hat{p}+(1-\hat{p})Q\big[\hat{q}^{-2},\hat{q}^{-2}\exp\{\hat{q}\hat{z}_i\}\big]\Big\} & \text{if} \quad i\in C, \end{cases} \quad (4.2)$$

- $q < 0$

$$r_{M_i} = \begin{cases} 1+\log\Big\{\hat{p}+(1-\hat{p})\Big(1-Q\big[\hat{q}^{-2},\hat{q}^{-2}\exp\{\hat{q}\hat{z}_i\}\big]\Big)\Big\} & \text{if} \quad i\in F \\ \log\Big\{\hat{p}+(1-\hat{p})\Big(1-Q\big[\hat{q}^{-2},\hat{q}^{-2}\exp\{\hat{q}\hat{z}_i\}\big]\Big)\Big\} & \text{if} \quad i\in C, \end{cases} \quad (4.3)$$

- $q = 0$

$$r_{M_i} = \begin{cases} 1 + \log\Big\{\hat{p} + (1-\hat{p})\Big(1 - \Phi\big[\hat{z}_i\big]\Big)\Big\} & \text{if} \quad i \in F \\ \log\Big\{\hat{p} + (1-\hat{p})\Big(1 - \Phi\big[\hat{z}_i\}\big]\Big)\Big\} & \text{if} \quad i \in C, \end{cases} \tag{4.4}$$

where $\hat{z}_i = \frac{y_i - \hat{\mu}}{\hat{\sigma}}$ and $Q(.)$ is the incomplete gamma integral.

More details about counting processes can be found, for instance, in Fleming and Harrington (1994) and Ortega (2001). These authors show that the distribution of the deviance component residual based on the martingale residual has very close asymptotic distribution to the normal distribution.

Therefore, considering q fixed, we have that the deviance component residual for generalized log-gamma mixture model becomes

- $q > 0$

$$r_{D_i} = \begin{cases} sgn(r_{M_i})\sqrt{2}\Big\{-1 - \log\Big[G_1(y_i, \hat{\boldsymbol{\theta}}) \log\big\{G_1(y_i, \hat{\boldsymbol{\theta}})^{-1}\big\}\Big]\Big\}^{\frac{1}{2}} & \text{if} \quad i \in F, \\ sgn(r_{M_i})\sqrt{2}\Big\{\log\big[G_1(y_i, \hat{\boldsymbol{\theta}})^{-1}\big]\Big\}^{\frac{1}{2}} & \text{if} \quad i \in C. \end{cases} \tag{4.5}$$

where $G_1(y_i, \hat{\boldsymbol{\theta}}) = \hat{p} + (1-\hat{p})Q\big[\hat{q}^{-2}, \hat{q}^{-2}\exp\{\hat{q}\hat{z}_i\}\big]$.

- $q < 0$

$$r_{D_i} = \begin{cases} sgn(r_{M_i})\sqrt{2}\Big\{-1 - \log\Big[G_2(y_i, \hat{\boldsymbol{\theta}}) \log\big\{G_2(y_i, \hat{\boldsymbol{\theta}})^{-1}\big\}\Big]\Big\}^{\frac{1}{2}} & \text{if} \quad i \in F, \\ sgn(r_{M_i})\sqrt{2}\Big\{\log\big[G_2(y_i, \hat{\boldsymbol{\theta}})^{-1}\big]\Big\}^{\frac{1}{2}} & \text{if} \quad i \in C. \end{cases} \tag{4.6}$$

where $G_2(y_i, \hat{\boldsymbol{\theta}}) = \hat{p} + (1-\hat{p})\Big\{1 - Q\big[\hat{q}^{-2}, \hat{q}^{-2}\exp\{\hat{q}\hat{z}_i\}\big]\Big\}$.

- $q = 0$

$$r_{D_i} = \begin{cases} sgn(r_{M_i})\sqrt{2}\Big\{-1 - \log\Big[G_3(y_i, \hat{\boldsymbol{\theta}}) \log\big\{G_3(y_i, \hat{\boldsymbol{\theta}})^{-1}\big\}\Big]\Big\}^{\frac{1}{2}} & \text{if} \quad i \in F, \\ sgn(r_{M_i})\sqrt{2}\Big\{\log\big[G_3(y_i, \hat{\boldsymbol{\theta}})^{-1}\big]\Big\}^{\frac{1}{2}} & \text{if} \quad i \in C. \end{cases} \tag{4.7}$$

where $G_3(y_i, \hat{\boldsymbol{\theta}}) = \hat{p} + (1-\hat{p})\Big\{1 - \Phi\big[\hat{z}_i\big]\Big\}$.

# 5 Applications

In this section, the application of the local influence theory to a set of real data on cancer recurrence is discussed. The data are part of an assay on cutaneous melanoma (a type of malignant cancer) for the evaluation of postoperative treatment performance with a high dose of a certain drug (interferon alfa-2b) in order to prevent recurrence. Patients were included in the study from 1991 to 1995, and follow-up was conducted until 1998. The data were collected by Ibrahim et al. (2001); variable T represented the time until the patient's death. The original size of the sample was n = 427 patients , 10 of whom did not present a value for covariable tumor thickness, herein denominated as Breslow. When such cases were removed, a sample of size $n = 417$ patients was considered. The percentage of censored observations was 56%. The following data were associated with each participant, $i = 1, 2, \ldots, n$.

- $t_i$: observed time (in years);
- $\delta_i$: censoring indicator (0=censoring, 1=lifetime observed);
- $x_{i1}$: treatment (0=observation, 1=interferon);
- $x_{i2}$: age (in years);
- $x_{i3}$: nodule (nodule category: to 4);
- $x_{i4}$: sex (0=male, 1=female);
- $x_{i5}$: p.s. (performance status-patient's functional capacity scale as regards his daily activities: 0=fully active, 1=other);
- $x_{i6}$: tumor (tumor thickness in mm.).

The survival function graph, Kaplan-Meier estimate, is presented in Figure 2, from where a significant fraction of survivors can be observed.

## 5.1 Maximum Likelihood and Jackknife Estimation

To obtain the maximum likelihood estimates for the parameters in model (2) we use the subroutine MAXBFGS in Ox, whose results are given in the Table 1. The mean cure fraction estimated was $\hat{p} = 0.4837$ and the only significant variable is x3 (nodule).

In Table 2 we report the Jackknife estimates for the parameters of the generalized log-gamma mixture with covariates.

From Table 2 we may observe that the explanatory variables $x_3$, is significant for the model when the Jackknife estimator is used. Although the estimates from the two methods seem to be very similar. Therefore, since for this sample size ($n = 417$) is expected normality for the Jackknife estimator, one may also expect some symmetric distribution for the MLEs with heavy tails. We will continue the analysis by using the MLEs.

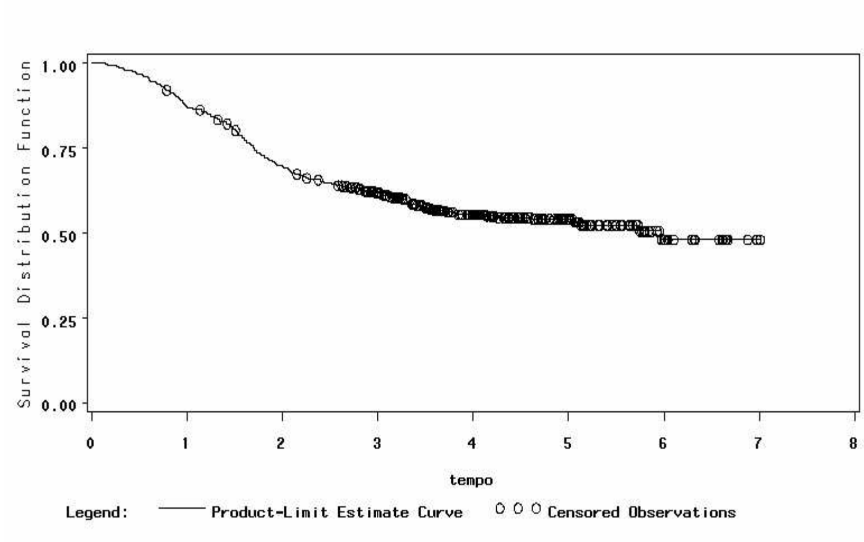


Figure 2: Plot of the Survivor Function.

Table 1: Maximum likelihood estimates from the fit of the generalized log-gamma mixture with covariates

| Parameter | Estimate | SE | p-value | 95% Confidence Interval |
|---|---|---|---|---|
| $\mu$ | 0.6529 | 0.0934 | <0.0001 | (0.4692 ; 0.8366) |
| $\sigma$ | 0.7789 | 0.0648 | <0.0001 | (0.6514 ; 0.9062) |
| $\beta_0$ | 2.4047 | 0.6554 | 0.0003 | (1.1165 ; 3.6929) |
| $\beta_1$ | -0.1590 | 0.2427 | 0.5128 | (-0.6360 ; 0.3181) |
| $\beta_2$ | -0.0158 | 0.0093 | 0.0912 | (-0.0341 ; 0.0025) |
| $\beta_3$ | -0.6016 | 0.1406 | <0.0001 | (-0.8780 ; -0.3253) |
| $\beta_4$ | 0.1999 | 0.2493 | 0.4231 | (-0.2902 ; 0.690) |
| $\beta_5$ | -0.1367 | 0.3661 | 0.7091 | (-0.8564 ; 0.5830) |
| $\beta_6$ | -0.0765 | 0.0481 | 0.1131 | (-0.1712 ; 0.8201) |
| $q$ | 0.2500 | | | |
| $p$ | 0.4837 | | | |

## 5.2 Global and Local Influence Analysis

In this subsection, we use Ox to compute case-deletion measures $GD_i(\boldsymbol{\theta})$ and $LD_i(\boldsymbol{\theta})$ presented in subsection 3.1. The results of such influence measures index plots are displayed

Table 2: Jackknife estimates from the fit of the generalized log-gamma mixture with covariates.

| Parameter | Estimate | SE | 95% Confidence Interval |
|---|---|---|---|
| $\mu$ | 0.6533 | 0.1007 | (0.4553 ; 0.8509) |
| $\sigma$ | 0.8275 | 0.0724 | (0.6852 ; 0.9695) |
| $\beta_0$ | 2.3176 | 0.6987 | (0.9441 ; 3.6885) |
| $\beta_1$ | -0.1845 | 0.2488 | (-0.6736 ; 0.3038) |
| $\beta_2$ | -0.0150 | 0.0094 | (-0.0335 ; 0.0034) |
| $\beta_3$ | -0.5947 | 0.1582 | (-0.9057 ; -0.2843) |
| $\beta_4$ | 0.3046 | 0.2562 | (-0.1990 ; 0.8073) |
| $\beta_5$ | -0.2475 | 0.3781 | (-0.9907 ; 0.4943) |
| $\beta_6$ | -0.0732 | 0.0600 | (-0.1911 ; 0.0445) |

in Figure 3. From this figure we can see that cases ♯47, ♯199, ♯279 and ♯341 are possible influential observations.

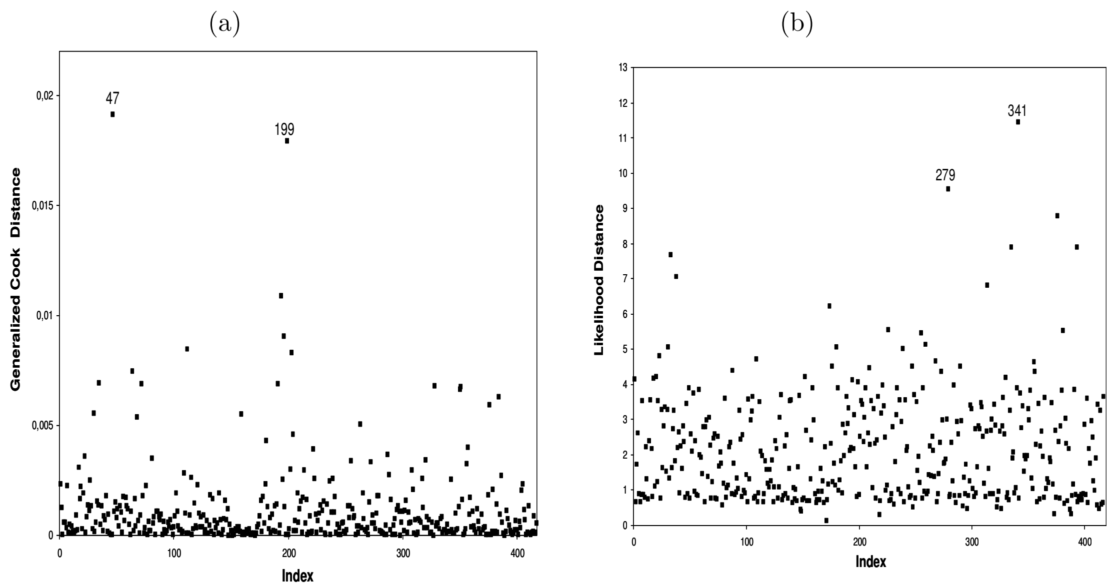


Figure 3: (a) Index plot of $GD_i(\boldsymbol{\theta})$ and (b) Index plot of $LD_i(\boldsymbol{\theta})$ from the fit of model (2) to the generalized log-gamma mixture.

## 5.3 Local Influence Analysis

In this section, we will make an analysis of local influence for the cancer data .

### 5.3.1 Case–Weights Perturbation

By applying the local influence theory developed in sub-section (3.2), where case-weight perturbation is used, value $C_{\boldsymbol{\ell}_{max}} = 1.6566$ was obtained as maximum curvature. In Figure 4(a), the graph of eigenvector corresponding to $C_{\boldsymbol{\ell}_{max}}$ is presented, and total influence $C_i$ is shown in Figure 4(b) and observations 47, 68, 176 and 341 are the most distinguished in relation to the others.

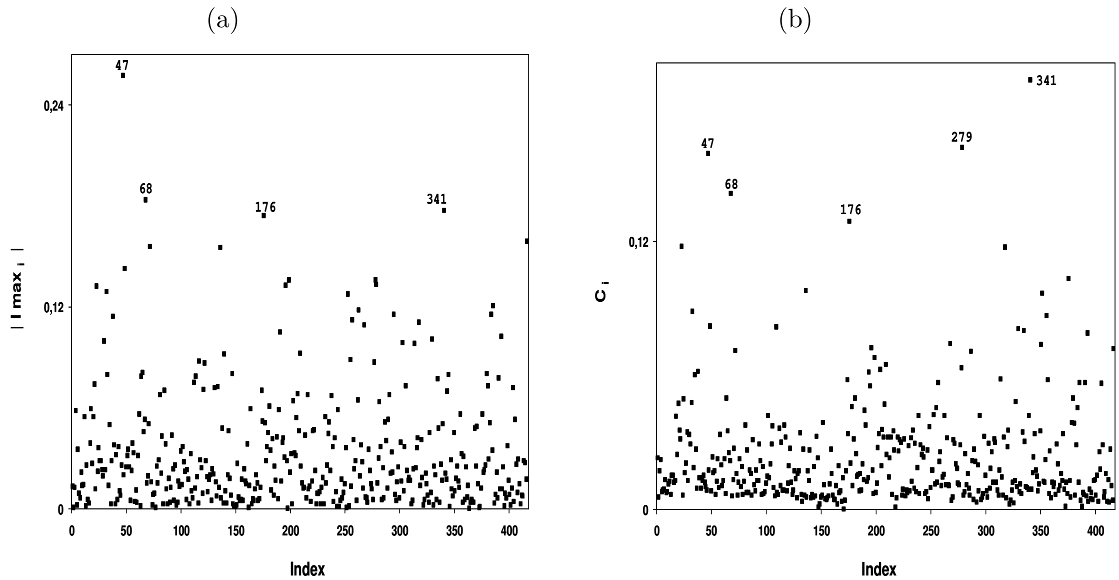


Figure 4: (a)Index plot of $\boldsymbol{\ell}_{max}$ for $\boldsymbol{\theta}$ (case-weights perturbation). (b)Total local influence on the estimates $\boldsymbol{\theta}$ (case-weights perturbation)

### 5.3.2 Influence Using Response Variable Perturbation

Next, the influence of perturbations on the observed survival times will be analyzed. The value for the maximum curvature calculated was $C_{\boldsymbol{\ell}_{max}} = 21.334$. Figure 5(a), containing the graph for $|\boldsymbol{\ell}_{max}|$ versus the observation index, shows that some points were distinguished from the others, among which are points 23, 47 and 176. The same applies to Figure 5(b), which corresponds to total local influence ($C_i$). By analyzing the data associated with these two observations, it is noted that that the highlighted observations refer to patients with shorter non-censored survival times.

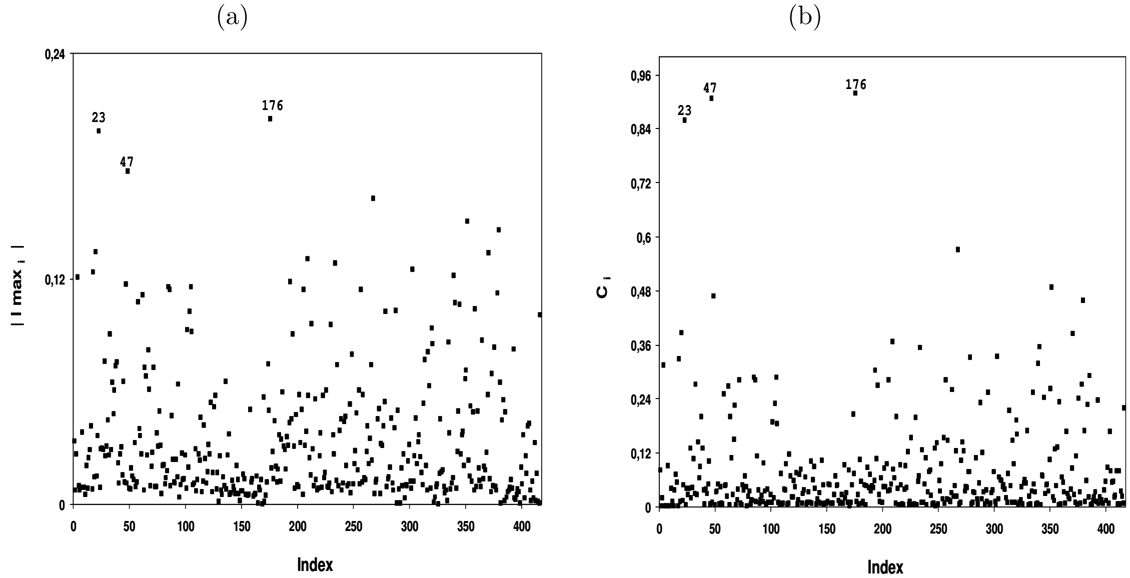


Figure 5: (a)Index plot of $\boldsymbol{\ell}_{max}$ for $\boldsymbol{\theta}$ (response perturbation). (b)Total local influence on the estimates $\boldsymbol{\theta}$ (response perturbation)

### 5.3.3 Influence Using Explanatory Variable Perturbation

The perturbation of vectors for covariables age($x_2$) and tumor($x_6$) is investigated here. For perturbation of covariable age, value $C_{\boldsymbol{\ell}_{max}} = 1.0375$ was obtained as maximum curvature, and for perturbation of covariable tumor, value $C_{\boldsymbol{\ell}_{max}} = 1.5753$ was achieved. The respective graphs of $|\boldsymbol{\ell}_{max}|$ as well as total local influence $C_i$ against the observation index are shown in Figures 6(a), 6(b), 6(c) and 6(d). These four graphs do not present observations with high influence.

## 5.4 Residual Analysis

In order to detect possible outlying observations as well as departures from the assumptions of the generalized log-gamma mixture model, we present in Figure 7 the graphs of $r_{Mi}$ and $r_{Di}$ against the order observations.

By analyzing the residual and martingale deviance graph (Figure 7), a random behavior is observed for the data. A tendency to form two groups is also noted; however, this results from considering the logistic function to introduce covariables. Such problems are also observed in the logistic regression. For further details, refer to Hosmer et al. (1989), McCullagh et al. (1989), among others.

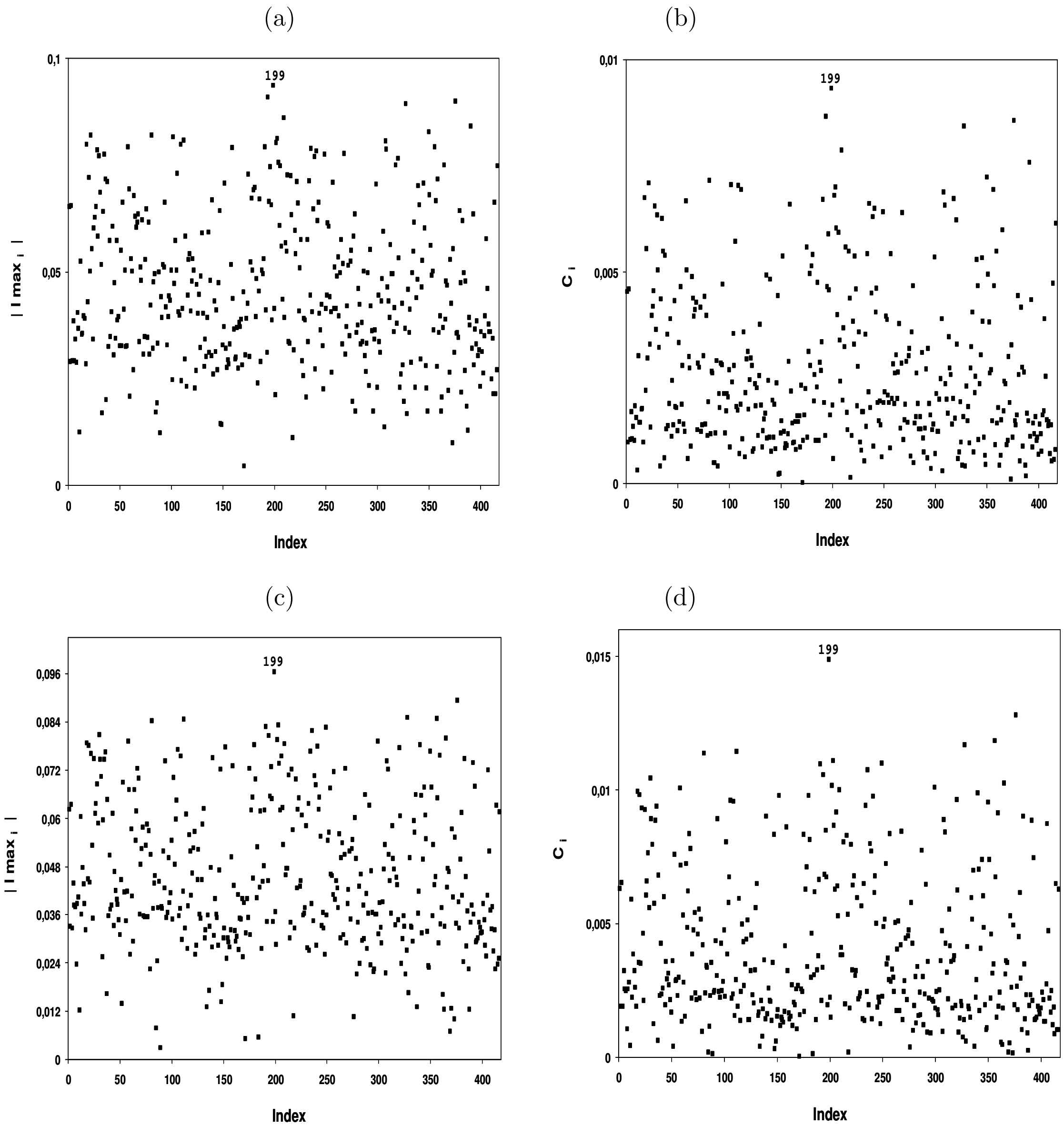


Figure 6: (a)Index plot of $\boldsymbol{\ell}_{max}$ for $\boldsymbol{\theta}$(age explanatory variable perturbation). (b)Total local influence on the estimates $\boldsymbol{\theta}$(age explanatory variable perturbation). (c)Index plot of $\boldsymbol{\ell}_{max}$ for $\boldsymbol{\theta}$( Breslow explanatory variable perturbation). (d)Total local influence on the estimates $\boldsymbol{\theta}$(Breslow explanatory variable perturbation).

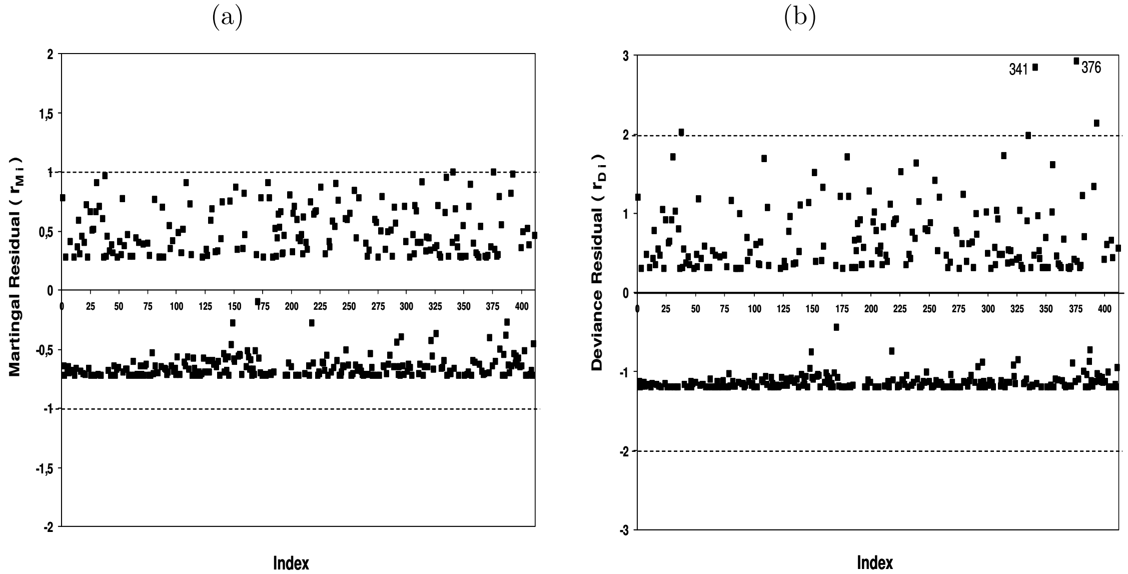


Figure 7: (a) Index plot of the martingale deviance residual $r_{M_i}$. (b) Index plot of the deviance residual $r_{D_i}$.

## 5.5 Impact of the Detected Influential Observations

Therefore, diagnostic analysis (local influence, global influence and residual analysis) detected the following four cases ♯47, ♯176 and ♯341 as potentially influential. In order to reveal the impact of these three observations on the parameter estimates, we refitted the model under some situations. First, we individually eliminated each one of these three cases. In Table 3, we have the relative changes (in percentage) of each parameter estimate, defined by: $RC_{\boldsymbol{\theta}_j} = [(\hat{\boldsymbol{\theta}}_j - \hat{\boldsymbol{\theta}}_j(I))/\hat{\boldsymbol{\theta}}_j]100$, and the corresponding $p$-values, where $\hat{\boldsymbol{\theta}}_j(I)$ denotes the MLE of $\hat{\boldsymbol{\theta}}_j$ after that "set I" of observations has been removed.

From Table 3, we can notice some robust aspects of the maximum likelihood estimates from the generalized log-gamma mixture with covariates. This is an indication that there three observations (♯476/♯176/♯341), deleted in the influential analysis, are masking the importance of the explanatory variable ($x_6$). Since is it not clear to remove these observations from the analysis we will present in the Table 4 the maximum likelihood estimator by removing the most influential ones.

## 5.6 Goodness of Fitting

In order to measure quality of fitting, a Kaplan-Meier survival graph and a survival graph estimated by the generalized log-gamma mixture model with a cure fraction were plotted

Table 3: Relative changes [-**RC**- in %], parameter estimates and their $p$-values in parentheses for the indicated set.

| Dropping | $\hat{\beta}_0$ | $\hat{\beta}_1$ | $\hat{\beta}_2$ | $\hat{\beta}_3$ | $\hat{\beta}_4$ | $\hat{\beta}_5$ | $\hat{\beta}_6$ |
|---|---|---|---|---|---|---|---|
| | - | - | - | - | - | - | - |
| all observations | 2.40 | -0.16 | -0.016 | -0.60 | 0.20 | -0.14 | -0.08 |
| | (<0.0001) | (0.51) | (0.09) | (<0.0001) | (0.42) | (0.71) | (0.11) |
| | [6] | [-13] | [-4] | [-6] | [13] | [-21] | [-21] |
| ♯47 | 2.55 | -0.14 | -0.02 | -0.64 | 0.17 | -0.11 | -0.09 |
| | (<0.0001) | (0.57) | (0.08) | (<0.0001) | (0.49) | (0.77) | (0.06) |
| | [2] | [-5] | [-3] | [-4] | [3] | [-8] | [-6] |
| ♯176 | 2.36 | -0.17 | -0.02 | -0.58 | 0.19 | -0.15 | -0.07 |
| | (<0.0001) | (0.48) | (0.09) | (<0.0001) | (0.43) | (0.68) | (0.11) |
| | [0] | [-12] | [0] | [-2] | [10] | [11] | [-2] |
| ♯341 | 2.40 | -0.18 | -0.02 | -0.59 | 0.22 | -0.15 | -0.07 |
| | (<0.0001) | (0.46) | (0.09) | (<0.0001) | (0.37) | (0.68) | (0.11) |
| | [4] | [-9] | [-1] | [-2] | [17] | [-14] | [-15] |
| ♯47/♯176 | 2.51 | -0.14 | -0.02 | -0.61 | 0.16 | -0.12 | -0.09 |
| | (<0.0001) | (0.55) | (0.08) | (<0.0001) | (0.50) | (0.74) | (0.06) |
| | [4] | [-9] | [-1] | [-2] | [17] | [-14] | [-15] |
| ♯47/♯341 | 2.51 | -0.14 | -0.02 | -0.61 | 0.16 | -0.12 | -0.09 |
| | (<0.0001) | (0.55) | (0.08) | (<0.0001) | (0.50) | (0.74) | (0.06) |
| | [5] | [-2] | [-1] | [0] | [8] | [-4] | [-14] |
| ♯476/♯176/♯341 | 2.51 | -0.16 | -0.02 | -0.60 | 0.18 | -0.13 | -0.09 |
| | (<0.0001) | (0.49) | (0.08) | (<0.0001) | (0.45) | (0.71) | (0.05) |

(see, Figure 8). Good model fitting was observed.

# 6 Concluding Remarks

In this study, the generalized log-gamma regression model was modified in order to include long-term individuals. In the proposal under consideration, log-linear parametric modelling was taken as a basis for survival time. The logit function showed to be an adequate link

Table 4: Maximum likelihood estimates for the generalized log-gamma mixture by removing the most influential observation (♯476/♯176/♯341).

| Parameter | Estimate | SE | p-value |
|---|---|---|---|
| $\mu$ | 0.6191 | 0.0801 | <0.0001 |
| $\sigma$ | 0.7361 | 0.0570 | <0.0001 |
| $\beta_0$ | 2.5142 | 0.6298 | <0.001 |
| $\beta_1$ | -0.1622 | 0.2376 | 0.4950 |
| $\beta_2$ | -0.0160 | 0.0091 | 0.0785 |
| $\beta_3$ | -0.6013 | 0.1312 | <0.0001 |
| $\beta_4$ | 0.1833 | 0.2442 | 0.4529 |
| $\beta_5$ | -0.1311 | 0.3560 | 0.7128 |
| $\beta_6$ | -0.0872 | 0.0453 | 0.0540 |

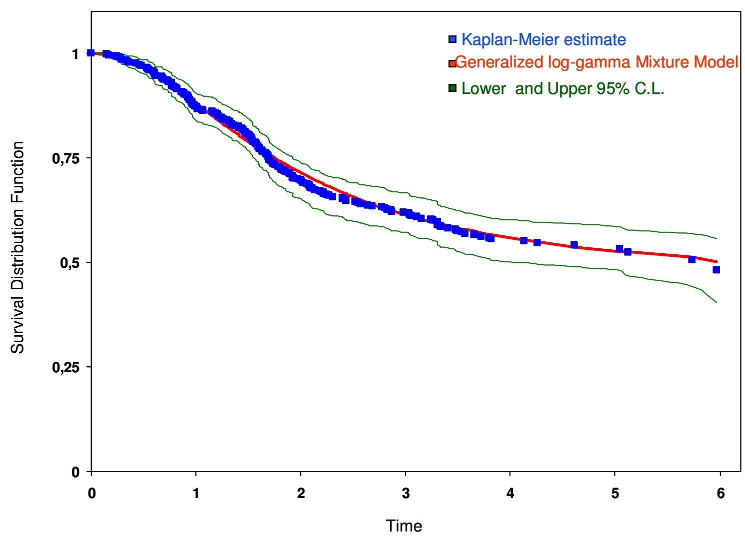


Figure 8: Theoretical survival curve, Kaplan-Meier curve and Upper and Lower 95% confidence limits.

function to include covariables in the proposed model. We used the Quasi-Newton algorithm to obtain the maximum likelihood estimates and were realized asymptotic tests for the parameters based on the asymptotic distribution of the maximum likelihood estimators.

On the other hand, as an alternative analysis, the paper discusses the use of the Jackknife estimator for the generalized log-gamma mixture with covariates. In addition to a study on martingale and deviance residuals in the generalized log-gamma regression model with long-term individuals in order to better evaluate the proposed model. The matrices necessary for application of the technique were obtained by considering various types of perturbation to the data elements and the model. By applying such results to a data set, indication was found which showed that the proposed model is a robust one, since the perturbations to the model and maximum likelihood estimators did not change significantly. The results of applications indicate that the use of the local influence technique as well as the analysis of residuals in the regression model with long-term individuals can be rather useful in the detection of possible influential points. In order to measure goodness of fitting was correct. We also plotted the Kaplan-Meier survival function with that given by proposed model, which indicated good fitting for the model.

## References

[1] Barlow, W. E., and Prentice, R. L.(1988). Residual for relative risk regression. *Biometrika*, **75**, 65–74.

[2] Berkson, J. and Gage, R. P.(1952). Survival curve for cancer patients following treatment. *Journal of the American Statistical Association*, **88**, 1412–1418.

[3] Cook, R. D. (1986). Assessment of local influence (with discussion). *Journal of the Royal Statistical Society*, **48**, 133–169.

[4] Cook, R. D. and Weisberg, S., 1982. *Residuals and Influence in Regression.* New York: Chapman and Hill.

[5] Cook, R. D., Peńa, D. and Weisberg, S. (1988). The Likelihood Displacement: A Unifying Principle for Influence. *Communications in Statistics - Theory and Methods*, **17**, 623–640.

[6] Díaz-Garca, J. A., Galea, M. and Leiva-Sánchez, V. (2004). Influence diagnostics for elliptical multivariate linear regression models. *Communications in statistics - Theory and Methods*, **32**, 625–641.

[7] Doornik, J. (1996). Ox: An Object-Oriented Matrix Programming Language. International Thomson Business Press.

[8] Escobar, L. A. and Meeker, W. Q. (1992). Assessing influence in regression analysis with censored data. *Biometrics*, **48**, 507–528.

[9] Farewell, V. T. (1982). The use of mixture models for the analysis of survival data with long-term survivors. *Biometrics*, **38**, 1041–1046.

[10] Fleming, T. R. and Harrington, D. P. (1991). *Counting Process and Survival Analysis.* Wiley: New York.

[11] Galea, M, Riquelme, M and Paula, G. A. (2002). Diagnostics methods in elliptical linear regression models. *Brazilian Journal of Probability and Statistics*, **14**, 167–184.

[12] Goldman, A. I. (1984). Survivorship analysis when cure is a possibility: A Monte Carlo study. *Statistics in medicine*, **3**, 153–163.

[13] Greenhouse, J. B., and Wolfe, R. A. (1984). A competing risks derivation of a mixture model for the analysis of survival. *Communications in statistics - Theory and Methods*, **13**, 3133–3154.

[14] Hosmer, D. W. and Lemeshow, S. (1989). *Applied Logistic Regression.* John Wiley: New York.

[15] Ibrahim, J. G., Chen, M. H. and Sinha, D. (2001). *Bayesian Survival Analysis*. Springer-Verlag: New York.

[16] Lawless, J. F. (2003). *Statistical Models and Methods for lifetime data*. Wiley: New York.

[17] Le, S. Y., Lu, B. and Song, X. Y. (2006). Assessing local influence for nonlinear structural equation models with ignorable missing data. *Computational Statistics and Data Analysis*, **50** 1356–1377.

[18] Lesaffre, E. and Verbeke, G. (1998). Local influence in linear mixed models. *Biometrics*, **54** 570–582.

[19] Lipsitz, S. R., Laird, N. M. and Harrington, D. P. (1990). Using the Jackknife to estimate the variance of regression estimators from repeated measures studies. *Communications in Statistics - Theory Methods*, **19**, 821–845.

[20] Liu, S. Z. (2000). On local influence for elliptical linear models. *Statist. Papers*, **41**, 211–224.

[21] Maller, R. and Zhou, X. (1996). *Survival Analysis with Long-term Survivors*. Wiley : New York.

[22] Manly, B. F. J., 1997. *Randomization, Bootstrap and Monte Carlo Methods in Biology*, 2nd Edition. Chapman and Hall: London.

[23] McCullagh, P. and Nelder, J. A. (1989). *Generalized Linear Models*, 2nd Edition. Chapman and Hall: London.

[24] Ortega, E. M. M., Cancho, V. G. and Bolfarine, H. (2006). Influence diagnostics in exponentiated-Weibull regression models with censored data. *Statistics and Operations Research Transactions*, **30**, (to appear).

[25] Ortega, E. M. M., Bolfarine, H. and Paula, G. A. (2003). Influence diagnostics in generalized log-gamma regression models. *Computational Statistics and Data Analysis*, **42**, 165–186.

[26] Ortega, E. M. M. (2001). *Influence Analysis and Residual in Generalized Log-gamma Regression Models*. Doctor Thesis, Department of Statistics, University of Sao Paulo, Brasil (in Portuguese).

[27] Paula, G. A. (1993). Assessing local influence in restricted regressions models. *Computational Statistics and Data Analysis*, **16**, 63–79.

[28] Paula, G. A. (1995). Influence residuals in restricted generalized linear models. *J. Statist. Comput. Simulation*, **51**, 63–79.

[29] Pettitt, A. N. and Bin Daud, I. (1989). Case-weight measures of influence for proportional hazards regression. *Applied Statistics*, **38**, 51–67.

[30] Sy, J. P. and Taylor, M. M. G. (2000). Estimation in a proportional hazards cure model. *Biometrics*, **56**, 227–336.

[31] Thomas, W. and Cook, R. D. (1990). Assessing influence on predictions from generalized linear models. *Technometrics*, **32**, 59–65.

[32] Therneau, T. M., Grambsch, P. M., and Fleming, T. R. (1990). Martingale-based residuals for survival models. *Biometrika*, **77**, 147–60.

## Appendix A. Matrix of Second Derivatives $\ddot{\mathbf{L}}(\boldsymbol{\theta})$

Here we derive the necessary formulae to obtain the second order partial derivatives of the log-likelihood function. After some algebraic manipulations, we obtain when $q > 0$

$$\begin{aligned}
\mathbf{L}_{\mu\mu} &= -\frac{1}{q^2\sigma^2}\sum_{i\in F} u_i + \sum_{i\in C}\left\{\frac{(1-p_i)q^2}{\sigma^2\Gamma(q^{-2})}{u_i}^{(q^{-2}+1)}\exp\{-u_i\}\right.\\
&\quad \left.\times\left[\frac{q^{-2}+1-u_i}{h_{i1}} - \frac{(1-p_i){u_i}^{q^{-2}+1}\exp\{-u_i\}}{\Gamma(q^{-2})h_{i1}^2}\right]\right\}\\
\mathbf{L}_{\mu\sigma} &= \frac{r}{q\sigma^2} - \frac{q}{\sigma^2}\sum_{i\in F} u_i - \frac{q^2}{\sigma^2}\sum_{i\in F} u_i z_i + \sum_{i\in C}\left\{\frac{(1-p_i)q}{\sigma^2\Gamma(q^{-2})}{u_i}^{q^{-2}+1}\exp\{-u_i\}\right.\\
&\quad \left.\times\left[\frac{(q^{-2}+1)qz_i - qu_iz_i - 1}{h_{i1}} - \frac{(1-p_i)q{u_i}^{q^{-2}+1}\exp\{-u_i\}z_i}{\Gamma(q^{-2})h_{i1}^2}\right]\right\}
\end{aligned}$$

$$
\begin{aligned}
\mathbf{L}_{\sigma\sigma} &= \frac{r}{\sigma^2} + \frac{2}{q\sigma^2}\sum_{i\in F} z_i - \frac{q^2}{\sigma^2}\sum_{i\in F} z_i^2 u_i - \frac{2q}{\sigma^2}\sum_{i\in F} z_i u_i + \sum_{i\in C}\left\{ \frac{(1-p_i)q{u_i}^{(q^{-2}+1)}}{\sigma^2\Gamma(q^{-2})} \right. \\
&\quad \left. \times \exp\{-u_i\} z_i \left[ \frac{(q^{-2}+1)qz_i - qu_iz_i + 2}{h_{i1}} - \frac{(1-p_i)q{u_i}^{q^{-2}+1} exp\{-u_i\}z_i}{\Gamma(q^{-2})h_{i1}^2} \right] \right\} \\
\mathbf{L}_{\mu\beta} &= \frac{q}{\sigma\Gamma(q^{-2})}\sum_{i\in C} \frac{x_{ij}(u_i)^{q^{-2}+1}\exp\{-u_i\}\exp\{\mathbf{x}_i^T\boldsymbol{\beta}\}}{h_{i1}^2\left[\exp\{\mathbf{x}_i^T\boldsymbol{\beta}\}\right]^2} \\
\mathbf{L}_{\sigma\beta} &= \frac{q}{\sigma\Gamma(q^{-2})}\sum_{i\in C} \frac{x_{ij}z_i(u_i)^{q^{-2}+1}\exp\{-u_i\}\exp\{\mathbf{x}_i^T\boldsymbol{\beta}\}}{h_{i1}^2\left[\exp\{\mathbf{x}_i^T\boldsymbol{\beta}\}\right]^2} \\
\mathbf{L}_{\beta\beta} &= \sum_{i\in F} - \frac{x_{ij}x_{is}\exp\{\mathbf{x}_i^T\boldsymbol{\beta}\}}{(1-p_i)^2\left[1+\exp\{\mathbf{x}_i^T\boldsymbol{\beta}\}\right]^4}\left\{(1-p_i)\left[1+\exp\{\mathbf{x}_i^T\boldsymbol{\beta}\}\right]^2\right. \\
&\quad \left. + \exp\{\mathbf{x}_i^T\boldsymbol{\beta}\} - 2(1-p_i)\exp\{\mathbf{x}_i^T\boldsymbol{\beta}\}\left[1+\exp\{\mathbf{x}_i^T\boldsymbol{\beta}\}\right]\right\} \\
&\quad + \sum_{i\in C} \frac{x_{ij}x_{is}\left[1-Q(q^{-2},u_i)\right]\exp\{\mathbf{x}_i^T\boldsymbol{\beta}\}}{h_{i1}^2\left[1+\exp\{\mathbf{x}_i^T\boldsymbol{\beta}\}\right]^4}\left\{h_{i1}\left[1+\exp\{\mathbf{x}_i^T\boldsymbol{\beta}\}\right]^2\right. \\
&\quad \left. - \exp\{\mathbf{x}_i^T\boldsymbol{\beta}\}\left[1-Q(q^{-2},u_i)\right] - 2h_{i1}\exp\{\mathbf{x}_i^T\boldsymbol{\beta}\}\left[1+\exp\{\mathbf{x}_i^T\boldsymbol{\beta}\}\right]\right\}
\end{aligned}
$$

When $q < 0$

$$
\begin{aligned}
\mathbf{L}_{\mu\mu} &= \frac{1}{q^2\sigma^2}\sum_{i\in F} u_i + \sum_{i\in C}\left\{ \frac{(1-p_i)q^2}{\sigma^2\Gamma(q^{-2})}{u_i}^{q^{-2}+1}\exp\{-u_i\} \right. \\
&\quad \left. \times \left[ \frac{-q^{-2}-1+u_i}{h_{i2}} - \frac{(1-p_i){u_i}^{q^{-2}+1}\exp\{-u_i\}}{\Gamma(q^{-2})h_{i2}^2} \right] \right\} \\
\mathbf{L}_{\mu\sigma} &= \frac{1}{q\sigma^2} - \frac{q}{\sigma^2}\sum_{i\in F} u_i - \frac{q^2}{\sigma^2}\sum_{i\in F} u_i z_i + \sum_{i\in C}\left\{ \frac{(1-p_i)q}{\sigma^2\Gamma(q^{-2})}{u_i}^{q^{-2}+1}\exp\{-u_i\} \right. \\
&\quad \left. \left[ \frac{-(q^{-2}+1)qz_i + qu_iz_i - 1}{h_{i2}} - \frac{(1-p_i)q{u_i}^{q^{-2}+1}\exp\{-u_i\}z_i}{\Gamma(q^{-2})h_{i2}^2} \right] \right\} \\
\mathbf{L}_{\sigma\sigma} &= \frac{r}{\sigma^2} + \frac{2}{q\sigma^2}\sum_{i\in F} z_i - \frac{q^2}{\sigma^2} z_i^2 u_i - \frac{2q}{\sigma^2} z_i u_i + \sum_{i\in C}\left\{ \frac{(1-p_i)qu_i^{(q^{-2}+1)}\exp\{-u_i\}z_i}{\sigma^2\Gamma(q^{-2})} \right. \\
&\quad \left. \left[ \frac{-(q^{-2}+1)qz_i + qu_iz_i - 2}{h_{i2}} - \frac{(1-p_i)q{u_i}^{q^{-2}+1}\exp\{-u_i\}z_i}{\Gamma(q^{-2}){h_{i2}}^2} \right] \right\}
\end{aligned}
$$

$$\begin{aligned}
\mathbf{L}_{\mu\beta} &= \frac{-q}{\sigma\Gamma(q^{-2})}\sum_{i\in C}\frac{x_{ij}(u_i)^{q^{-2}+1}\exp\{-u_i\}\exp\{\mathbf{x}_i^T\boldsymbol{\beta}\}}{h_{i2}^2\big[\exp\{\mathbf{x}_i^T\boldsymbol{\beta}\}\big]^2} \\
\mathbf{L}_{\sigma\beta} &= \frac{-q}{\sigma\Gamma(q^{-2})}\sum_{i\in C}\frac{x_{ij}z_i(u_i)^{q^{-2}+1}\exp\{-u_i\}\exp\{\mathbf{x}_i^T\boldsymbol{\beta}\}}{h_{i2}^2\big[\exp\{\mathbf{x}_i^T\boldsymbol{\beta}\}\big]^2} \\
\mathbf{L}_{\beta\beta} &= \sum_{i\in F} -\frac{x_{ij}x_{is}\exp\{\mathbf{x}_i^T\boldsymbol{\beta}\}}{(1-p_i)^2\big[1+\exp\{\mathbf{x}_i^T\boldsymbol{\beta}\}\big]^4}\Bigg\{(1-p_i)\Big[1+\exp\{\mathbf{x}_i^T\boldsymbol{\beta}\}\Big]^2 \\
&\quad + \exp\{\mathbf{x}_i^T\boldsymbol{\beta}\} - 2(1-p_i)\exp\{\mathbf{x}_i^T\boldsymbol{\beta}\}\Big[1+\exp\{\mathbf{x}_i^T\boldsymbol{\beta}\}\Big]\Bigg\} \\
&\quad + \sum_{i\in C}\frac{x_{ij}x_{is}Q(q^{-2},u_i)\exp\{\mathbf{x}_i^T\boldsymbol{\beta}\}}{h_{i2}^{*2}\Big[1+\exp\{\mathbf{x}_i^T\boldsymbol{\beta}\}\Big]^4}\Bigg\{h_{i2}\Big[1+\exp\{\mathbf{x}_i^T\boldsymbol{\beta}\}\Big]^2 \\
&\quad - \exp\{\mathbf{x}_i^T\boldsymbol{\beta}\}Q(q^{-2},u_i) + 2h_{i2}\exp\{\mathbf{x}_i^T\boldsymbol{\beta}\}\Big[1+\exp\{\mathbf{x}_i^T\boldsymbol{\beta}\}\Big]\Bigg\}
\end{aligned}$$

When $q = 0$

$$\begin{aligned}
\mathbf{L}_{\mu\mu} &= -\frac{r}{\sigma^2} + \frac{1}{\sigma}\sum_{i\in C}(1-p_i)\phi(z_i)h_{i3}^{-2}\Big[z_ih_{i3} - (1-p_i)\phi(z_i)\Big] \\
\mathbf{L}_{\mu\sigma} &= -\frac{2}{\sigma^2}\sum_{i\in F}z_i + \sum_{i\in C}(1-p_i)h_{i3}^{-2}\phi(z_i)\sigma^{-2}\Bigg\{\Big[-2+z_i^2\Big]h_{i3} - (1-p_i)z_i\phi(z_i)\Bigg\} \\
\mathbf{L}_{\sigma\sigma} &= \frac{r}{\sigma^2} - \frac{2}{\sigma}\sum_{i\in F}z_i^2 + \sum_{i\in C}(1-p_i)h_{i3}^{-2}z_i^2\sigma^{-2} \\
&\quad \times\Bigg\{\Bigg[-3z_i^{-1}\phi(z_i) + z_i\sigma^{-1}\exp\Big\{-\frac{z_i^2}{2}\Big\}\Bigg]h_{i3} - \big[\phi(z_i)\big]^2\Bigg\} \\
\mathbf{L}_{\mu\beta} &= -\frac{1}{\sigma}\sum_{i\in C}\frac{x_{ij}p_i\phi(z_i)}{\big[1+\exp\{\mathbf{x}_i^T\boldsymbol{\beta}\}\big]}\Bigg[h_{i3}^{-1} + (1-Pi)h_{i3}^{-2}\Phi(z_i)\Bigg] \\
\mathbf{L}_{\sigma\beta} &= -\frac{1}{\sigma}\sum_{i\in C}\frac{x_{ij}p_iz_i\phi(z_i)}{\big[1+\exp\{\mathbf{x}_i^T\boldsymbol{\beta}\}\big]}\Bigg[h_{i3}^{-1} + (1-Pi)h_{i3}^{-2}\Phi(z_i)\Bigg]
\end{aligned}$$

$$\begin{aligned}\mathbf{L}_{\beta\beta} &= \sum_{i\in F}(1-p_i)^{-2}\Big[1+\exp\{\mathbf{x}_i^T\boldsymbol{\beta}\}\Big]^{-4}\Big[x_{ij}x_{is}\exp\{\mathbf{x}_i^T\boldsymbol{\beta}\}\Big\{(1-p_i)\Big[1+\exp\{\mathbf{x}_i^T\boldsymbol{\beta}\}\Big]^2 \\ &\quad+\exp\{\mathbf{x}_i^T\boldsymbol{\beta}\}-2(1-p_i)\exp\{\mathbf{x}_i^T\boldsymbol{\beta}\}\Big[1+\exp\{\mathbf{x}_i^T\boldsymbol{\beta}\}\Big]\Big\}\Big] \\ &\quad+\sum_{i\in C}\Big[1+\exp\{\mathbf{x}_i^T\boldsymbol{\beta}\}\Big]^{-4}{h_{i3}}^{-2}\Big[\Phi(z_i)x_{ij}x_{is}\exp\{\mathbf{x}_i^T\boldsymbol{\beta}\}\Big\{\Big[1+\exp\{\mathbf{x}_i^T\boldsymbol{\beta}\}\Big]^2h_{i3} \\ &\quad-2\exp\{\mathbf{x}_i^T\boldsymbol{\beta}\}\Big[1+\exp\{\mathbf{x}_i^T\boldsymbol{\beta}\}\Big]h_{i3}+\Phi(z_i)\Big\}\Big]\end{aligned}$$

where

$$h_{i1}=\Big[p_i+(1-p_i)Q(q^{-2},u_i)\Big],\qquad h_{i2}=\Big[p_i+(1-p_i)\Big\{1-Q(q^{-2},u_i)\Big\}\Big],$$
$$h_{i3}=\Big[p_i+(1-p_i)\Big\{1-\Phi(z_i)\Big\}\Big],\qquad u_i=q^{-2}\exp\{qz_i\},\quad z_i=\frac{y_i-\mu}{\sigma}$$
$$p_i=\frac{\exp(\mathbf{x}_i^T\boldsymbol{\beta})}{1+\exp(\mathbf{x}_i^T\boldsymbol{\beta})}$$

$i=1,2,\ldots,n,\ j,s=0,1,2,\ldots,p,\ \Phi(\cdot)$ denoting the standard normal cumulative distribution, $\phi(\cdot)$ standard normal density function and $Q(\cdot)$ is the incomplete gamma integral.

## Appendix B. Case–Weight Perturbation Scheme

Here, we provide the of the elements the matrix $\boldsymbol{\Delta}$ considering the case-weight perturbation scheme.

- $q>0$

  Then the elements of vector $\boldsymbol{\Delta}_1$ take the form

$$\Delta_{1i}=\begin{cases}(\hat{q}\hat{\sigma})^{-1}\big[-1+\exp\{\hat{q}\hat{z}_i\}\big] & \text{if} \quad i\in F\\ (1-\hat{p}_i)\hat{u}_i^{\hat{q}^{-2}}\exp\{-\hat{u}_i\}\hat{q}\Big[\hat{h}_{i1}\hat{\sigma}\Gamma(\hat{q}^{-2})\Big]^{-1} & \text{if} \quad i\in C\end{cases}$$

  On the other hand, the elements of vector $\boldsymbol{\Delta}_2$ can be shown to be given by

$$\Delta_{2i}=\begin{cases}-(\hat{\sigma})^{-1}-(\hat{q}\hat{\sigma})^{-1}\hat{z}_i[1-\exp\{\hat{q}\hat{z}_i\}] & \text{if} \quad i\in F\\ (1-\hat{p}_i)\hat{u}_i^{\hat{q}^{-2}}\exp\{-\hat{u}_i\}\hat{z}_i\hat{q}\Big[\hat{h}_{i1}\hat{\sigma}\Gamma(\hat{q}^{-2})\Big]^{-1} & \text{if} \quad i\in C\end{cases}$$

  The elements of vector $\boldsymbol{\Delta}_j$, for $j=3,\ldots,p+2$, may be expressed as

$$\Delta_{ji}=\begin{cases}-x_{ij}\hat{p}_i(1-\hat{p}_i)^{-1}\big[1+\exp\{\mathbf{x}_i^T\hat{\boldsymbol{\beta}}\}\big]^{-1} & \text{if} \quad i\in F\\ x_{ij}\hat{p}_i\big[1-Q(\hat{q}^{-2},\hat{u}_i)\big]\big[1+\exp(\mathbf{x}_i^T\hat{\boldsymbol{\beta}})\big]^{-1}[\hat{h}_{i1}]^{-1} & \text{if} \quad i\in C\end{cases}$$

- $q < 0$

  Then the elements of vector $\mathbf{\Delta}_1$ take the form

$$\Delta_{1i} = \begin{cases} (\hat{q}\hat{\sigma})^{-1}\big[-1+\exp\{\hat{q}\hat{z}_i\}\big] & \text{if} \quad i \in F \\ (1-\hat{p}_i)\hat{u}_i^{\hat{q}-2}\exp\{-\hat{u}_i\}(-\hat{q})\Big[\hat{h}_{i2}\hat{\sigma}\Gamma(\hat{q}^{-2})\Big]^{-1} & \text{if} \quad i \in C \end{cases}$$

  On the other hand, the elements of vector $\mathbf{\Delta}_2$ can be shown to be given by

$$\Delta_{2i} = \begin{cases} -(\hat{\sigma})^{-1} - (\hat{q}\hat{\sigma})^{-1}\hat{z}_i[1-\exp\{\hat{q}\hat{z}_i\}] & \text{if} \quad i \in F \\ (1-\hat{p}_i)\hat{u}_i^{\hat{q}-2}\exp\{-\hat{u}_i\}\hat{z}_i(-\hat{q})\Big[\hat{h}_{i2}\hat{\sigma}\Gamma(\hat{q}^{-2})\Big]^{-1} & \text{if} \quad i \in C \end{cases}$$

  The elements of vector $\mathbf{\Delta}_j$, for $j = 3, \ldots, p+2$, may be expressed as

$$\Delta_{ji} = \begin{cases} -x_{ij}\hat{p}_i(1-\hat{p}_i)^{-1}\big[1+\exp\{\mathbf{x}_i^T\hat{\boldsymbol{\beta}}\}\big]^{-1} & \text{if} \quad i \in F \\ x_{ij}\hat{p}_i Q(\hat{q}^{-2}, \hat{u}_i)\big[1+\exp(\mathbf{x}_i^T\hat{\boldsymbol{\beta}})\big]^{-1}[\hat{h}_{i2}]^{-1} & \text{if} \quad i \in C \end{cases}$$

- $q = 0$

  Then the elements of vector $\mathbf{\Delta}_1$ take the form

$$\Delta_{1i} = \begin{cases} \hat{\sigma}^{-1}\hat{z}_i & \text{if} \quad i \in F \\ (1-\hat{p}_i)\phi(\hat{z}_i)\hat{\sigma}^{-1}\hat{h}_{i3}^{-1} & \text{if} \quad i \in C \end{cases}$$

  On the other hand, the elements of vector $\mathbf{\Delta}_2$ can be shown to be given by

$$\Delta_{2i} = \begin{cases} \hat{\sigma}^{-1}(-1+\hat{z}_i^2) & \text{if} \quad i \in F \\ (1-\hat{p}_i)\phi(\hat{z}_i)\hat{z}_i\hat{\sigma}^{-1}\hat{h}_{i3}^{-1} & \text{if} \quad i \in C \end{cases}$$

  The elements of vector $\mathbf{\Delta}_j$, for $j = 3, \ldots, p+2$, may be expressed as

$$\Delta_{ji} = \begin{cases} -x_{ij}\hat{p}_i(1-\hat{p}_i)[1+\exp\{\mathbf{x}_i^T\hat{\boldsymbol{\beta}}\}] & \text{if} \quad i \in F \\ x_{ij}\hat{p}_i\Phi(\hat{z}_i)[1+\exp\{\mathbf{x}_i^T\hat{\boldsymbol{\beta}}\}]\hat{h}_{i3}^{-1} & \text{if} \quad i \in C \end{cases}$$

## Appendix C. Response Perturbation Scheme

Here, we provide the of the elements $\mathbf{\Delta}_{ij}$ considering the response variable perturbation scheme.

- $q > 0$ The elements of vector $\mathbf{\Delta}_1$ take form

$$\Delta_{1i} = \begin{cases} \hat{q}^{-1}\hat{\sigma}^{-2}\exp\{\hat{q}\hat{z}_i\}S_y & \text{if} \quad i \in F \\ \frac{(1-\hat{p}_i)\hat{q}S_y}{\hat{\sigma}\Gamma(\hat{q}^{-2})}\left[\frac{\hat{u}_i^{\hat{q}^{-2}}(\hat{q}^{-1}-\hat{u}_i\exp\{-\hat{u}_i\}\hat{q})}{\hat{\sigma}\hat{h}_{i1}} + \frac{\hat{u}_i^{2\hat{q}^{-2}}\hat{q}(1-\hat{p}_i)\exp\{-2\hat{u}_i\}}{\hat{\sigma}\Gamma(\hat{q}^{-2})\hat{h}_{i1}^2}\right] & \text{if} \quad i \in C \end{cases}$$

On the other hand, the elements of vector $\mathbf{\Delta}_2$ can be shown to be given by

$$\Delta_{2i} = \begin{cases} -(\hat{q}\hat{\sigma}^2)^{-1}S_y\big[-1+\exp\{\hat{q}\hat{z}_i\}(1+\hat{z}_i\hat{q})\big] & \text{if} \quad i \in F \\ \frac{(1-\hat{p}_i)\hat{q}S_y}{\hat{\sigma}^2\Gamma(\hat{q}^{-2})}\Bigg\{\frac{\hat{u}_i^{\hat{q}^{-2}}\big[\hat{z}_i(\hat{q}^{-1}-\hat{u}_i\exp\{-\hat{u}_i\}\hat{q})+\exp\{-\hat{u}_i\}\big]}{\hat{h}_{i1}} & \\ \qquad +\frac{\hat{u}_i^{2\hat{q}^{-2}}\exp\{-2\hat{u}_i\}\hat{z}_i(1-\hat{p}_i)\hat{q}}{\Gamma(\hat{q}^{-2})\hat{h}_{i1}^2}\Bigg\} & \text{if} \quad i \in C \end{cases}$$

The elements of vector $\mathbf{\Delta}_j$, for $j = 3, \ldots, p+2$, may be expressed as

$$\Delta_{ji} = \begin{cases} 0 & \text{if} \quad i \in F \\ \frac{x_{ij}\hat{p}_i\hat{u}_i^{\hat{q}^{-2}}\exp\{-\hat{u}_i\}\hat{q}S_y}{\hat{\sigma}\Gamma(\hat{q}^{-2})\left[1+\exp\{\mathbf{x}_i^T\hat{\boldsymbol{\beta}}\}\right]}\left\{\hat{h}_i^{-1}+\hat{h}_i^{-2}(1-\hat{p}_i)\Big[1-Q(\hat{q}^{-2},\hat{u}_i)\Big]\right\} & \text{if} \quad i \in C \end{cases}$$

- $q < 0$

Then the elements of vector $\mathbf{\Delta}_1$ take the form

$$\Delta_{1i} = \begin{cases} \hat{q}^{-1}\hat{\sigma}^{-2}\exp\{\hat{q}\hat{z}_i\}S_y & \text{if} \quad i \in F \\ \frac{(1-\hat{p}_i)(-\hat{q})S_y}{\hat{\sigma}\Gamma(\hat{q}^{-2})}\left[\frac{\hat{u}_i^{\hat{q}^{-2}}(\hat{q}^{-1}-\hat{u}_i\exp\{-\hat{u}_i\}\hat{q})}{\hat{\sigma}\hat{h}_{i2}} - \frac{\hat{u}_i^{2\hat{q}^{-2}}\hat{q}(1-\hat{p}_i)\exp\{-2\hat{u}_i\}}{\hat{\sigma}\Gamma(\hat{q}^{-2})\hat{h}_{i2}^2}\right] & \text{if} \quad i \in C \end{cases}$$

On the other hand, the elements of vector $\mathbf{\Delta}_2$ can be shown to be given by

$$\Delta_{2i} = \begin{cases} -(\hat{q}\hat{\sigma}^2)^{-1}S_y\big[1-\exp\{\hat{q}\hat{z}_i\}(1+\hat{z}_i\hat{q})\big] & \text{if} \quad i \in F \\ \frac{(1-\hat{p}_i)(-\hat{q})S_y}{\hat{\sigma}^2\Gamma(\hat{q}^{-2})}\Bigg\{\frac{\hat{u}_i^{\hat{q}^{-2}}\big[\hat{z}_i(\hat{q}^{-1}-\hat{u}_i\exp\{-\hat{u}_i\}\hat{q})+\exp\{-\hat{u}_i\}\big]}{\hat{h}_{i2}} & \\ \qquad +\frac{\hat{u}_i^{2\hat{q}^{-2}}\exp\{-2\hat{u}_i\}\hat{z}_i(1-\hat{p}_i)\hat{q}}{\Gamma(\hat{q}^{-2})\hat{h}_{i2}^2}\Bigg\} & \text{if} \quad i \in C \end{cases}$$

The elements of vector $\mathbf{\Delta}_j$, for $j = 3, \ldots, p+2$, may be expressed as

$$\Delta_{ji} = \begin{cases} 0 & \text{if} \quad i \in F \\ \frac{x_{ij}\hat{p}_i\hat{u}_i^{\hat{q}^{-2}}\exp\{-\hat{u}_i\}(-\hat{q})S_y}{\hat{\sigma}\Gamma(\hat{q}^{-2})\left[1+\exp\{\mathbf{x}_i^T\hat{\boldsymbol{\beta}}\}\right]}\left\{\hat{h}_i^{-1}+\hat{h}_i^{-2}(1-\hat{p}_i)Q(\hat{q}^{-2},\hat{u}_i)\right\} & \text{if} \quad i \in C \end{cases}$$

- $q = 0$

  Then the elements of vector $\boldsymbol{\Delta}_1$ take the form

$$\Delta_{1i} = \begin{cases} \hat{\sigma}^{-2} S_y & \text{if} \quad i \in F \\ \frac{(1-\hat{p}_i)\phi(\hat{z}_i)S_y}{\hat{\sigma}^2}\Big[-\hat{z}_i\hat{h}_{i3}^{-1} + \phi(\hat{z}_i)(1-\hat{p}_i)\hat{h}_{i3}^{-2}\Big] & \text{if} \quad i \in C \end{cases}$$

  On the other hand, the elements of vector $\boldsymbol{\Delta}_2$ can be shown to be given by

$$\Delta_{2i} = \begin{cases} 2\hat{\sigma}^{-2}\hat{z}_i S_y & \text{if} \quad i \in F \\ \frac{(1-\hat{p}_i)\phi(\hat{z}_i)S_y}{\hat{\sigma}^2}\Big[(1-\hat{z}_i^2)\hat{h}_{i3}^{-1} + \phi(\hat{z}_i)(1-\hat{p}_i)\hat{z}_i\hat{h}_{i3}^{-2}\Big] & \text{if} \quad i \in C \end{cases}$$

  The elements of vector $\boldsymbol{\Delta}_j$, for $j = 3, \ldots, p+2$, may be expressed as

$$\Delta_{ji} = \begin{cases} 0 & \text{if} \quad i \in F \\ \frac{x_{ij}\hat{p}_i\phi(\hat{z}_i)S_y}{\hat{\sigma}\big[1+\exp\{\mathbf{x}_i^T\hat{\boldsymbol{\beta}}\}\big]}\Big[\hat{h}_{i3}^{-1} + \Phi(\hat{z}_i)(1-\hat{p}_i)\hat{h}_{i3}^{-2}\Big] & \text{if} \quad i \in C \end{cases}$$

## Appendix D. Explanatory Variable Perturbation

Here we provide derivations of the of elements $\boldsymbol{\Delta}_{ij}$ considering the explanatory variable perturbation scheme.

- $q > 0$

  The elements of the vector $\boldsymbol{\Delta}_1$ are expressed as

$$\Delta_{1i} = \begin{cases} 0 & \text{if} \quad i \in F \\ \frac{\hat{u}_i^{\hat{q}^{-2}}\exp\{-\hat{u}_i\}(-\hat{q})\hat{p}_i(1-\hat{p}_i)\hat{\beta}_t S_x}{\hat{\sigma}\Gamma(\hat{q}^{-2})}\Big\{\hat{h}_{i1}^{-1} + (1-\hat{p}_i)\Big[1-Q(\hat{q}^{-2},\hat{u}_i)\Big]\hat{h}_{i1}^{-2}\Big\} & \text{if} \quad i \in C, \end{cases}$$

  the elements of vector $\boldsymbol{\Delta}_2$ are expressed as

$$\Delta_{2i} = \begin{cases} 0 & \text{if} \quad i \in F \\ \frac{\hat{u}_i^{\hat{q}^{-2}}\exp\{-\hat{u}_i\}\hat{z}_i(-\hat{q})\hat{p}_i(1-\hat{p}_i)\hat{\beta}_t S_x}{\hat{\sigma}\Gamma(\hat{q}^{-2})}\Big\{\hat{h}_{i1}^{-1} + (1-\hat{p}_i)\Big[1-Q(\hat{q}^{-2},\hat{u}_i)\Big]\hat{h}_{i1}^{-2}\Big\} & \text{if} \quad i \in C, \end{cases}$$

  the elements of vector $\boldsymbol{\Delta}_j$, for $j = 1, \ldots, p$ and $j \neq t$, take the forms

$$\Delta_{ji} = \begin{cases} -x_{ij}\hat{p}_i^2\hat{\beta}_t S_x(1-\hat{p}_i)^{-1}\big[1+\exp\{\mathbf{x}_i^T\hat{\boldsymbol{\beta}}\}\big]^{-1} & \text{if} \quad i \in F \\ -x_{ij}\hat{p}_i^2(1-\hat{p}_i)\hat{\beta}_t S_x\Big[1-Q(\hat{q}^{-2},\hat{u}_i)\Big]^2\Big[1+\exp\{\mathbf{x}_i^T\hat{\boldsymbol{\beta}}\}\Big]^{-1}\hat{h}_{i1}^{-2} & \text{if} \quad i \in C, \end{cases}$$

  the elements of the vector $\boldsymbol{\Delta}_t$ are given by

$$\Delta_{ti} = \begin{cases} -\hat{p}_i S_x\Big[x_{it}\hat{\beta}_t(1-\hat{p}_i)+1\Big] & \text{if} \quad i \in F \\ x_{it}\hat{p}_i(1-\hat{p}_i)\Big[1-Q(\hat{q}^{-2},\hat{u}_i)\Big] & \\ \Big\{\big(1-\hat{\beta}_t\hat{p}_i S_x\big)\hat{h}_{i1}^{-1} - \hat{p}_i(1-\hat{p}_i)\hat{\beta}_t S_x\Big[1-Q(\hat{q}^{-2},\hat{u}_i)\Big]\hat{h}_{i1}^{-2}\Big\} & \text{if} \quad i \in C \end{cases}$$

- $q < 0$

  The elements of vector $\mathbf{\Delta}_1$ are expressed as

$$\Delta_{1i} = \begin{cases} 0 & \text{if} \quad i \in F \\ \frac{\hat{u}_i^{\hat{q}^{-2}}\exp\{-\hat{u}_i\}\hat{q}\hat{p}_i(1-\hat{p}_i)\hat{\beta}_t S_x}{\hat{\sigma}\Gamma(\hat{q}^{-2})}\Big\{\hat{h}_{i2}^{-1} + (1-\hat{p}_i)Q(\hat{q}^{-2},\hat{u}_i)\hat{h}_{i2}^{-2}\Big\} & \text{if} \quad i \in C, \end{cases}$$

  the elements of vector $\mathbf{\Delta}_2$ are expressed as

$$\Delta_{2i} = \begin{cases} 0 & \text{if} \quad i \in F \\ \frac{\hat{u}_i^{\hat{q}^{-2}}\exp\{-\hat{u}_i\}\hat{z}_i\hat{q}\hat{p}_i(1-\hat{p}_i)\hat{\beta}_t S_x}{\hat{\sigma}\Gamma(\hat{q}^{-2})}\Big\{\hat{h}_{i2}^{-1} + (1-\hat{p}_i)Q(\hat{q}^{-2},\hat{u}_i)\hat{h}_{i2}^{-2}\Big\} & \text{if} \quad i \in C, \end{cases}$$

  the elements of vector $\mathbf{\Delta}_j$, for $j = 1, \ldots, p$ and $j \neq t$, take the forms

$$\Delta_{ji} = \begin{cases} -x_{ij}\hat{p}_i^2\hat{\beta}_t S_x(1-\hat{p}_i)^{-1}\big[1+\exp\{\mathbf{x}_i^T\hat{\boldsymbol{\beta}}\}\big]^{-1} & \text{if} \quad i \in F \\ -x_{ij}\hat{p}_i^2(1-\hat{p}_i)\hat{\beta}_t S_x\Big[Q(\hat{q}^{-2},\hat{u}_i)\Big]^2\Big[1+\exp\{\mathbf{x}_i^T\hat{\boldsymbol{\beta}}\}\Big]^{-1}\hat{h}_{i2}^{-2} & \text{if} \quad i \in C, \end{cases}$$

  the elements of vector $\mathbf{\Delta}_t$ are given by

$$\Delta_{ti} = \begin{cases} -\hat{p}_i S_x\Big[x_{it}\hat{\beta}_t(1-\hat{p}_i)+1\Big] & \text{if} \quad i \in F \\ x_{it}\hat{p}_i(1-\hat{p}_i)Q(\hat{q}^{-2},\hat{u}_i)\Big\{\big(1-\hat{\beta}_t\hat{p}_i S_x\big)\hat{h}_{i2}^{-1} & \\ -\hat{p}_i(1-\hat{p}_i)\hat{\beta}_t S_x Q(\hat{q}^{-2},\hat{u}_i)\hat{h}_{i2}^{-2}\Big\} & \text{if} \quad i \in C \end{cases}$$

- $q = 0$

  The elements of vector $\mathbf{\Delta}_1$ are expressed as

$$\Delta_{1i} = \begin{cases} 0 & \text{if} \quad i \in F \\ -\hat{\sigma}^{-1}\hat{p}_i(1-\hat{p}_i)\hat{\beta}_t S_x\phi(\hat{z}_i)\Big[\hat{h}_{i3}^{-1} - (1-\hat{p}_i)\Phi(\hat{z}_i)\hat{h}_{i3}^{-2}\Big] & \text{if} \quad i \in C, \end{cases}$$

  the elements of vector $\mathbf{\Delta}_2$ are expressed as

$$\Delta_{2i} = \begin{cases} 0 & \text{if} \quad i \in F \\ -\hat{\sigma}^{-1}\hat{z}_i\hat{p}_i(1-\hat{p}_i)\hat{\beta}_t S_x\phi(\hat{z}_i)\Big[\hat{h}_{i3}^{-1} - (1-\hat{p}_i)\Phi(\hat{z}_i)\hat{h}_{i3}^{-2}\Big] & \text{if} \quad i \in C, \end{cases}$$

the elements of vector $\mathbf{\Delta}_j$, for $j = 1, \ldots, p$ and $j \neq t$, take the forms

$$\Delta_{ji} = \begin{cases} -x_{ij}\hat{p}_i^2(1-\hat{p}_i)^{-1}\hat{p}_i\hat{\beta}_t S_x & \text{if} \quad i \in F \\ -x_{ij}\hat{p}_i^2(1-\hat{p}_i)\hat{\beta}_t S_x \Big[\Phi(\hat{z}_i)\Big]^2 \Big[1+\exp\{\mathbf{x}_i^T\hat{\boldsymbol{\beta}}\}\Big]^{-1}\hat{h}_{i3}^{-1} & \text{if} \quad i \in C, \end{cases}$$

the elements of vector $\mathbf{\Delta}_t$ are given by

$$\Delta_{ti} = \begin{cases} -\hat{p}_i S_x\big[x_{it}\hat{\beta}_t(1-\hat{p}_i)+1\big] & \text{if} \quad i \in F \\ x_{it}p_i(1-p_i)\Phi(z_i)\Big[\big(1-\beta_t S_x p_i\big)h_{i3}^{-1} - p_i(1-pi)\beta_t S_x \Phi(z_i)h_{i3}^{-2}\Big] & \text{if} \quad i \in C \end{cases}$$

where

$$\widehat{h}_{i1} = \Big[\hat{p}_i + (1-\hat{p}_i)Q(\hat{q}^{-2}, \hat{u}_i)\Big], \qquad \widehat{h}_{i2} = \Big[\hat{p}_i + (1-\hat{p}_i)\Big\{1 - Q(\hat{q}^{-2}, \hat{u}_i)\Big\}\Big],$$
$$\widehat{h}_{i3} = \Big[\hat{p}_i + (1-\hat{p}_i)\Big\{1-\Phi(\hat{z}_i)\Big\}\Big], \qquad \widehat{u}_i = \hat{q}^{-2}exp\{\hat{q}\hat{z}_i\}, \quad \hat{z}_i = \frac{y_i - \hat{\mu}}{\hat{\sigma}}$$
$$\hat{p}_i = \frac{\exp(\mathbf{x}_i^T\hat{\boldsymbol{\beta}})}{1+\exp(\mathbf{x}_i^T\hat{\boldsymbol{\beta}})}$$

$\Phi(\cdot)$ denoting the standard normal cumulative distribution, $\phi(\cdot)$ standard normal density function and $Q(\cdot)$ the incomplete gamma integral.