Statistics
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/gsta20

On estimation and diagnostics analysis in log-generalized gamma regression model for interval-censored data

Elizabeth M. Hashimoto [a], Edwin M.M. Ortega [a], Vicente G. Cancho [b] & Gauss M. Cordeiro [c]

[a] Departmento de Ciências Exatas, Universidade de São Paulo, Avenue Pádua Dias 11, Caixa Postal 9, 13418-900, Piracicaba, São Paulo, Brazil

[b] Departamento de Matemática Aplicada e Estatística, Universidade de São Paulo, Brazil

[c] Departamento de Estatística, Universidade Federal de Pernambuco, Brazil

Available online: 17 Aug 2011

To cite this article: Elizabeth M. Hashimoto, Edwin M.M. Ortega, Vicente G. Cancho & Gauss M. Cordeiro (2011): On estimation and diagnostics analysis in log-generalized gamma regression model for interval-censored data, Statistics, DOI:10.1080/02331888.2011.605888

To link to this article: http://dx.doi.org/10.1080/02331888.2011.605888

iFirst

Statistics, iFirst, 2011, 1–20

On estimation and diagnostics analysis in log-generalized gamma regression model for interval-censored data

Elizabeth M. Hashimoto[a], Edwin M.M. Ortega[a]*, Vicente G. Cancho[b] and Gauss M. Cordeiro[c]

[a]*Departmento de Ciências Exatas, Universidade de São Paulo, Avenue Pádua Dias 11, Caixa Postal 9, 13418-900 Piracicaba, São Paulo, Brazil;* [b]*Departamento de Matemática Aplicada e Estatística, Universidade de São Paulo, Brazil;* [c]*Departamento de Estatística, Universidade Federal de Pernambuco, Brazil*

(*Received 27 November 2009; final version received 13 July 2011*)

The interval-censored survival data appear very frequently, where the event of interest is not observed exactly but it is only known to occur within some time interval. In this paper, we propose a location-scale regression model based on the log-generalized gamma distribution for modelling interval-censored data. We shall be concerned only with parametric forms. The proposed model for interval-censored data represents a parametric family of models that has, as special submodels, other regression models which are broadly used in lifetime data analysis. Assuming interval-censored data, we consider a frequentist analysis, a Jackknife estimator and a non-parametric bootstrap for the model parameters. We derive the appropriate matrices for assessing local influence on the parameter estimates under different perturbation schemes and present some techniques to perform global influence.

Keywords: log-generalized gamma regression; generalized gamma distribution; interval-censored data; maximum likelihood; regression model; sensitivity analysis

1. Introduction

In several studies, survival response can be interval-censored such that the event of interest is not observed exactly, but it is only known to occur within some time intervals that may overlap and vary in length. The literature presents many applications of survival models for interval-censored data by taking the Weibull family of distributions [1]. This family is very suitable in situations where the failure rate function is constant or monotone. However, it is not suitable in situations where the failure rate function presents a bathtub or a unimodal shape. To cope with these situations, several distributions were derived from the Weibull distribution to exhibit bathtub-shaped or unimodal failure rate functions, one of which is the generalized gamma (GG) distribution [2].

In some situations, the times of the events of interest T may only be known to have occurred within an interval of time, say $[U, V]$, where $U \leq T \leq V$. This can occur in a clinical trial, for

*Corresponding author. Email: edwin@esalq.usp.br

ISSN 0233-1888 print/ISSN 1029-4910 online

DOI: 10.1080/02331888.2011.605888
http://www.informaworld.com

example, when patients are assessed only at pre-scheduled visits. If the event has not occurred at one visit (at time U) but has by the following visit (at time V), T is known only to be within the interval $[U, V]$. These are known as interval-censored data. Note that exactly observed, right- and left-censored data are special cases of interval-censored data, with $U = V$ for exactly observed data, $V = \infty$ for right-censored and $U = 0$ for left-censored observations.

In this paper, we examine some statistical inference aspects for modelling interval-censored data based on the log-GG (LGG) regression model. The inferential aspects were carried out using the asymptotic distribution of the maximum-likelihood estimators (MLEs), where the normality is more difficult to be justified when the sample size is small. As an alternative to the frequentist analysis, we explore the use of the Jackknife estimator for the LGG regression model for interval-censored data. A punctual and an interval estimation methodology, based on bootstrap re-sampling methods, are also proposed.

After modelling, it is important to check the model assumptions and conduct sensitivity studies to detect possible influential or extreme observations that can cause distortions in the results from the analysis. In this paper, we discuss the influence diagnostics based on case deletion [3], in which the influence of the ith observation on the parameter estimates is studied by removing the case from the analysis. We propose diagnostic measures based on case deletion for the LGG regression models for interval-censored data in order to determine which subjects might be influential in the analysis. This methodology has been applied in various statistical models. See, for instance [4–6].

Nevertheless, when case deletion is used, all information from a single subject is deleted at once and, therefore, it is hard to say whether that subject has some influence on a specific aspect of the model. A solution for the earlier problem can be found in the local influence approach, where we discuss how the results from the analysis are changed under small perturbations in the model or data. Cook [7] proposed a general framework to detect the influence of observations which indicates how sensitive is the analysis when small perturbations are provoked on the data and the model. Some authors have investigated the assessment of local influence in survival analysis models. For instance, Pettitt and Bin Daud [8] investigated the local influence in proportional hazard regression models, Escobar and Meeker [9] adapted local influence methods to regression analysis with censoring, Ortega *et al.* [10] considered the problem of assessing the local influence in LGG regression models with censored observations. Further, Magnus and Vasnev [11] confronted sensitivity analysis with diagnostic testing with applications in econometrics and Xie and Wei [12] developed the application of influence diagnostics in censored generalized Poisson regression models based on the case-deletion method and the local influence analysis. More recently, Fachini *et al.* [13] adapted local influence methods to poly-hazard models under the presence of explanatory variables, Carrasco *et al.* [14] investigated the influence diagnostics in log-modified Weibull regression models with censored data, Silva *et al.* [15] performed the global and local influence methods in log-Burr XII regression models with censored data and Ortega *et al.* [16] derived curvature calculations under various perturbation schemes in regression models with the cure fraction. Here, we propose a similar methodology to detect influential subjects as the one in the log-exponentiated Weibull regression model for interval-censored data [17].

The paper is organized as follows. In Section 2, we present the LGG regression model for interval-censored data in addition to maximum-likelihood estimation, the Jackknife estimator and bootstrap re-sampling methods. The score functions and the observed information matrix are derived, and the process for estimating the regression coefficients and the remaining parameters is discussed. In Section 3, we perform a simulation study for the proposed model. In Section 4, we adopt some diagnostic measures considering the case deletion and the normal curvatures of local influence under various perturbation schemes in the LGG regression model with interval-censored data. In Section 5, a real data set is analysed to show the usefulness of the techniques described. Finally, in Section 6, we offer some concluding remarks.

2. The LGG regression models for interval-censored data

From now on, let T be a random variable having the GG probability density function (pdf) with parameters $(\alpha, \tau, k)^\top$ given by

$$f(t;\alpha,\tau,k) = \frac{\tau}{\alpha\Gamma(k)}\left(\frac{t}{\alpha}\right)^{\tau k-1}\exp\left[-\left(\frac{t}{\alpha}\right)^{\tau}\right], \quad t,\alpha,\tau,k>0, \tag{1}$$

where $\Gamma(\cdot)$ is the gamma function, α and τ are the shape parameters and k is a scale parameter. The survival function reduces to $S(t;\alpha,\tau,k) = 1 - Q[k,(t/\alpha)^\tau]$, where $Q(k,x) = \Gamma(k)^{-1}\int_0^x u^{k-1}\,\mathrm{e}^{-u}\,\mathrm{d}u$ is the incomplete gamma integral. The GG family is a very flexible model since it includes several well-known distributions as submodels [18]. The main submodels of the GG distribution are: the exponential ($k = \tau = 1$), the gamma ($\tau = 1$) and the Weibull ($k = 1$) distributions. The log-normal distribution is also obtained as a limiting distribution when $k \to \infty$. If $\tau = 2$, we obtain the generalized normal (GN) distribution, say GN$(2k,\alpha)$. The GN distribution is itself a flexible family that includes the half-normal ($k = 0.5$), Rayleigh ($k = 1$), Maxwell–Boltzmann ($k = \frac{3}{2}$) and Chi ($k = \nu/2, \nu = 1,2,\ldots$) distributions. The GG hazard function is simply given by $h(t;\alpha,\tau,k) = f(t;\alpha,\tau,k)/S(t;\alpha,\tau,k)$. The great flexibility of this model to fit lifetime data is due to the different forms that the hazard function can take, that is: (i) if $\tau > 1$ and $k = 1$, the hazard function is monotonically increasing; (ii) if $\tau < 1$ and $k = 1$, the hazard function is monotonically decreasing; (iii) if $1 < \tau < 1/k$ and $k < 1$, the hazard function is bathtub-shaped and (iv) if $1/k < \tau < 1$ and $k > 1$, we have a unimodal hazard function.

Applications of the GG distribution in reliability and survival studies were investigated by Lawless [1]. Cox *et al.* [19] proposed a parametric survival analysis and the taxonomy of hazard functions for the GG distribution, Almpanidis and Kotropoulos [20] presented a text-independent automatic phone segmentation algorithm based on the GG distribution, Nadarajah [21] analysed some incorrect references with respect to the use of this distribution in electrical and electronics engineering, Ortega *et al.* [22] proposed deviance residuals in generalized log-gamma (LG) regression models with censored observations and Gomes *et al.* [23] developed a study on the parameter estimation of the GG distribution. Recently, Ortega *et al.* [16] developed generalized LG regression models with the cure fraction.

The random variable $Y = \log(T)$ defined from the GG random variable T has an LGG density function, parametrized in terms of $\mu = \log(\alpha) + \tau^{-1}\log(\lambda^{-2})$, $\sigma = (\tau\sqrt{k})^{-1}$ and $\lambda = (\sqrt{k})^{-1}$, given by

$$f(y;\lambda,\sigma,\mu) = \begin{cases} \dfrac{|\lambda|}{\sigma\Gamma(\lambda^{-2})}(\lambda^{-2})^{\lambda^{-2}}\exp\left\{\lambda^{-2}\left[\left(\dfrac{y-\mu}{\sigma}\right)\lambda - \exp\left\{\left(\dfrac{y-\mu}{\sigma}\right)\lambda\right\}\right]\right\} & \text{if } \lambda \neq 0, \\ \dfrac{1}{\sqrt{2\pi\sigma^2}}\exp\left[-\dfrac{1}{2}\left(\dfrac{y-\mu}{\sigma}\right)^2\right] & \text{if } \lambda = 0, \end{cases} \tag{2}$$

where $-\infty < y < \infty$, $-\infty < \lambda < \infty$ is the shape parameter, $\sigma > 0$ is the scale parameter and $-\infty < \mu < \infty$ is the location parameter. The corresponding survival function can be expressed as

$$S(y;\lambda,\sigma,\mu)=\begin{cases} Q\left\{\lambda^{-2},\lambda^{-2}\exp\left[\lambda\left(\dfrac{y-\mu}{\sigma}\right)\right]\right\} & \text{if } \lambda>0,\\ 1-Q\left\{\lambda^{-2},\lambda^{-2}\exp\left[\lambda\left(\dfrac{y-\mu}{\sigma}\right)\right]\right\} & \text{if } \lambda<0,\\ 1-\Phi\left(\dfrac{y-\mu}{\sigma}\right) & \text{if } \lambda=0,\end{cases}$$

where $\Phi(\cdot)$ is the standard normal cumulative distribution and $Q(\cdot,\cdot)$ was defined before. Plots of the density function (2) for selected parameter values are given in Figure 1.

The standardized random variable $Z=(Y-\mu)/\sigma$ has a density function given by

$$f(z;\lambda,\sigma,\mu)=\begin{cases} \dfrac{|\lambda|(\lambda^{-2})^{\lambda^{-2}}}{\Gamma(\lambda^{-2})}\exp[\lambda^{-1}z-\lambda^{-2}\exp(\lambda z)] & \text{if } \lambda\neq 0,\\ \dfrac{1}{\sqrt{2\pi}}\exp\left(-\frac{z^2}{2}\right) & \text{if } \lambda=0.\end{cases} \tag{3}$$

The extreme value of standard distribution corresponds to the particular choice $\lambda=1$.

We hardly need to emphasize the necessity and importance of moments in any statistical analysis, especially in applied work. Some of the most important features and characteristics of a distribution can be studied through moments (e.g. tendency, dispersion, skewness and kurtosis). Now, consider the following theorem:

THEOREM 1 *If $Y\sim \mathrm{LGG}(\lambda,\sigma,\mu)$, the rth moment of Y for $\lambda>0$ is given by*

$$\mu_r'=E(Y^r)=\frac{1}{\Gamma(w)}\sum_{j=0}^{r}\binom{r}{j}[2\sigma\log(w^{-1/2})^{w^{1/2}}+\mu]^{r-j}\sigma^j w^{j/2}\Gamma^{(j)}(w),$$

where $w=\lambda^{-2}$, $\Gamma^{(j)}(w)=\partial^j\Gamma(w)/\partial w^j$.

Proof The rth moment of the LGG distribution is

$$\mu_r'=\int_{-\infty}^{\infty}\frac{y^r\lambda}{\sigma\Gamma(\lambda^{-2})}(\lambda^{-2})^{\lambda^{-2}}\exp\left\{\lambda^{-2}\left[\left(\frac{y-\mu}{\sigma}\right)\lambda-\exp\left\{\left(\frac{y-\mu}{\sigma}\right)\lambda\right\}\right]\right\}\,\mathrm{d}y.$$

Setting $x=\lambda^{-2}\exp\{\lambda[(y-\mu)/\sigma]\}$, μ_r' can be reduced to

$$\mu_r'=\frac{1}{\Gamma(\lambda^{-2})}\int_0^{\infty}\left\{\frac{\sigma}{\lambda}[\log(x)+2\log(\lambda)]+\mu\right\}^r x^{\lambda^{-2}-1}\,\mathrm{e}^{-x}\,\mathrm{d}x. \tag{4}$$

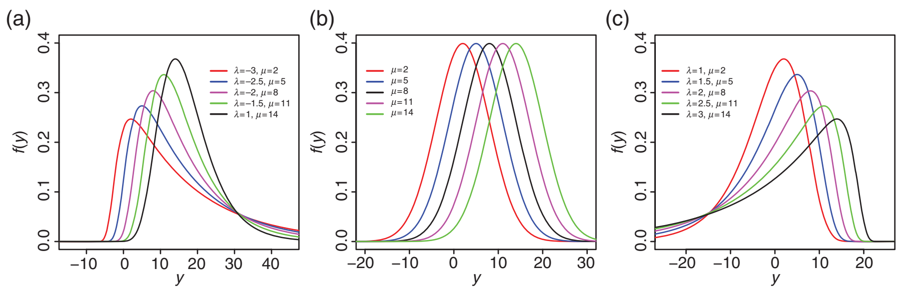

Figure 1. Plots of the LGG density function. (a) For $\lambda<0$ and $\sigma=6$, (b) for $\lambda=0$ and $\sigma=6$ and (c) for $\lambda>0$ and $\sigma=6$.

Applying the binomial expansion in Equation (4) and after some algebra, we obtain

$$\mu_r' = \frac{1}{\Gamma(\lambda^{-2})} \sum_{j=0}^{r} \binom{r}{j} [2\sigma \log(\lambda^{1/\lambda}) + \mu]^{r-j} \left(\frac{\sigma}{\lambda}\right)^j \int_0^\infty x^{\lambda^{-2}-1} \mathrm{e}^{-x} [\log(x)]^j \,\mathrm{d}x. \qquad (5)$$

The integral in Equation (5) can be expressed as

$$\int_0^\infty x^{\lambda^{-2}-1} \mathrm{e}^{-x} [\log(x)]^j \,\mathrm{d}x = \frac{\partial^j \Gamma(\lambda^{-2})}{\partial (\lambda^{-2})^j}.$$

Substituting the previous result in Equation (5) and setting $w = \lambda^{-2}$, we have

$$\mu_r' = \frac{1}{\Gamma(w)} \sum_{j=0}^{r} \binom{r}{j} [2\sigma \log(w^{-1/2})^{w^{1/2}} + \mu]^{r-j} \sigma^j w^{j/2} \Gamma^{(j)}(w). \qquad \blacksquare$$

Hence, for $\lambda > 0$, $E(Y) = \mu + \sigma\lambda^{-1}[\psi(\lambda^{-2}) - \log(\lambda^{-2})]$ and $\mathrm{Var}(Y) = \lambda^{-2}\sigma^2\psi'(\lambda^{-2})$, where $\psi(\cdot)$ and $\psi'(\cdot)$ denote the digamma and trigamma functions, respectively. Following similar steps of Theorem 1, the rth moment of the LGG distribution for $\lambda < 0$ can also be determined.

In many practical applications, the lifetimes are affected by explanatory variables such as the cholesterol level, blood pressure and many others. Let $\mathbf{x} = (x_1, \ldots, x_p)^\top$ be the explanatory variable vector associated with the response variable y. Based on the LGG density function, we can construct a linear regression model linking the response variable y_i and the explanatory variable vector $\mathbf{x}_i$ as follows:

$$y_i = \mathbf{x}_i^\top \boldsymbol{\beta} + \sigma z_i, \quad i = 1, \ldots, n, \qquad (6)$$

where the random error z_i has the density function (3), $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)^\mathrm{T}$, $\sigma > 0$ and $\lambda > 0$ are unknown parameters and $\mathbf{x}_i^\top = (x_{i1}, \ldots, x_{ip})$ is the explanatory variable vector modelling the location parameter μ_i. Hence, the location parameter vector $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)^\top$ of the LGG regression model can be expressed as a linear model $\boldsymbol{\mu} = \mathbf{X}\boldsymbol{\beta}$, where $\mathbf{X} = (\mathbf{x}_1, \ldots, \mathbf{x}_n)^\top$ is a known model matrix.

From the log-linear model (6), the survival function of $Y_i|\mathbf{x}$ can take three different forms:

$$S(y_i|\mathbf{x}) = \begin{cases} Q\left\{\lambda^{-2}, \lambda^{-2}\exp\left[\lambda\left(\dfrac{y_i - \mathbf{x}_i^\top\boldsymbol{\beta}}{\sigma}\right)\right]\right\} & \text{if } \lambda > 0, \\ 1 - Q\left\{\lambda^{-2}, \lambda^{-2}\exp\left[\lambda\left(\dfrac{y_i - \mathbf{x}_i^\top\boldsymbol{\beta}}{\sigma}\right)\right]\right\} & \text{if } \lambda < 0, \\ 1 - \Phi\left(\dfrac{y_i - \mathbf{x}_i^\top\boldsymbol{\beta}}{\sigma}\right) & \text{if } \lambda = 0. \end{cases}$$

For interval-censored data, the observed data consist of an interval $(\log(u_i), \log(v_i))$ for each individual, where such intervals are known to include $y_i = \log(t_i)$ with probability 1, i.e. $P[\log(u_i) \le y_i \le \log(v_i)] = 1$ and if $\log(v_i) = \infty$, then it is a right-censored time for y_i. This model will be referred to as the LGG regression model for interval-censored data. It is an extension of an accelerated failure time model using the GG distribution for interval-censored data.

Setting $\lambda = 1$ in model (6), we obtain the log-Weibull (LW) (or extreme value) regression model for interval-censored data. Further, if in addition, $\sigma = 1$, model (6) reduces to the log-exponential (LE) regression model for interval-censored data. If $\lambda = -1$, we obtain the log-reciprocal Weibull (LRW) regression model for interval-censored data. Finally, if $\tau = 1$, we have the LG regression models for interval-censored data.

2.1. *Maximum-likelihood estimation*

Given a set of interval-censored observations and explanatory variables $(\log(u_1), \log(v_1), \mathbf{x}_1), \ldots,$ $(\log(u_n), \log(v_n), \mathbf{x}_n)$ of n observations, where $(\log(u_i), \log(v_i))$ is the observed data and $\mathbf{x}_i$ the explanatory variable vector, the full log-likelihood function for the parameter vector $\boldsymbol{\theta} = (\lambda, \sigma, \boldsymbol{\beta}^\top)^\top$ can be expressed as

$$l(\boldsymbol{\theta}) = \sum_{i \in F} l_1(\lambda, zu_i, zv_i) + \sum_{i \in C} l_2(\lambda, zu_i), \tag{7}$$

where

$$l_1(\lambda, zu_i, zv_i) = \begin{cases} \log\{Q\{\lambda^{-2}, \lambda^{-2}\exp[\lambda(zu_i)]\} - Q\{\lambda^{-2}, \lambda^{-2}\exp[\lambda(zv_i)]\}\} & \text{if } \lambda > 0, \\ \log\{Q\{\lambda^{-2}, \lambda^{-2}\exp[\lambda(zv_i)]\} - Q\{\lambda^{-2}, \lambda^{-2}\exp[\lambda(zu_i)]\}\} & \text{if } \lambda < 0, \\ \log[\Phi(zv_i) - \Phi(zu_i)] & \text{if } \lambda = 0 \end{cases}$$

and

$$l_2(\lambda, zu_i, zv_i) = \begin{cases} \log\{Q\{\lambda^{-2}, \lambda^{-2}\exp[\lambda(zu_i)]\}\} & \text{if } \lambda > 0, \\ \log\{1 - Q\{\lambda^{-2}, \lambda^{-2}\exp[\lambda(zu_i)]\}\} & \text{if } \lambda < 0, \\ \log[1 - \Phi(zu_i)] & \text{if } \lambda = 0, \end{cases}$$

where F denotes the set of individuals with interval censoring, that is, $y_i \in (\log(u_i), \log(v_i)]$, C denotes the set of individuals with direct censoring, that is, $y_i \in (\log(u_i), +\infty)$, $zu_i = [\log(u_i) - \mathbf{x}_i^\top \boldsymbol{\beta}]/\sigma$ and $zv_i = [\log(v_i) - \mathbf{x}_i^\top \boldsymbol{\beta}]/\sigma$. The maximization of Equation (7) follows the same two steps for obtaining the MLE of $\boldsymbol{\theta}$ under the uncensored case. In general, it is reasonable to consider that the shape parameter λ is in the interval $[-3, 3]$. We fixed, in the first step of the iterative process, different q values in this interval. Then, we obtain the MLEs $\tilde{\sigma}(\lambda)$ and $\tilde{\boldsymbol{\beta}}(\lambda)$, and the maximized log-likelihood function $L_{\max}(\lambda)$ is then determined. We use, in this step, the matrix programming language Ox, subroutine MAXBFGS (see, for instance, [24]). In the second step, the log-likelihood $L_{\max}(\lambda)$ is maximized, and then $\hat{\lambda}$ is obtained. The MLEs of σ and $\boldsymbol{\beta}$ are $\hat{\sigma} = \tilde{\sigma}(\hat{\lambda})$ and $\hat{\boldsymbol{\beta}} = \tilde{\boldsymbol{\beta}}(\hat{\lambda})$, respectively.

The procedures discussed in this work are developed by assuming q fixed. The estimate of the covariance matrix of the MLEs $\hat{\boldsymbol{\theta}}$ can also be defined by the Hessian matrix. Under standard regularity conditions [25], confidence intervals (CIs) and hypothesis tests can be conducted by using the large sample multivariate normal distribution of the MLEs, where the covariance matrix is given by the inverse of the information matrix. More specifically, the asymptotic covariance matrix of $\hat{\boldsymbol{\theta}}$ is simply $\mathbf{I}(\boldsymbol{\theta})^{-1}$, where $\mathbf{I}(\boldsymbol{\theta}) = E[\ddot{\mathbf{L}}(\boldsymbol{\theta})]$ and $\ddot{\mathbf{L}}(\boldsymbol{\theta}) = -\partial^2 l(\boldsymbol{\theta})/\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^{\mathrm{T}}$.

We cannot compute the expected information matrix $\mathbf{I}(\boldsymbol{\theta})$ due to censored observations (censoring is random and noninformative), but it is possible to use minus the matrix of second derivatives of the log-likelihood, $\ddot{\mathbf{L}}(\boldsymbol{\theta})$, evaluated at the MLE $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$, which is a consistent estimate of $\mathbf{I}(\boldsymbol{\theta})$. The asymptotic normal approximation for $\hat{\boldsymbol{\theta}}$ may be expressed as $\hat{\boldsymbol{\theta}} \sim N_{(p+2)}\{\boldsymbol{\theta}, \ddot{\mathbf{L}}(\boldsymbol{\theta})^{-1}\}$, where $\ddot{\mathbf{L}}(\boldsymbol{\theta})$ is the $(p+2) \times (p+2)$ observed information matrix given by

$$\ddot{\mathbf{L}}(\boldsymbol{\theta}) = \begin{pmatrix} \mathbf{L}_{\lambda\lambda} & \mathbf{L}_{\lambda\sigma} & \mathbf{L}_{\lambda\beta_j} \\ \cdot & \mathbf{L}_{\sigma\sigma} & \mathbf{L}_{\sigma\beta_j} \\ \cdot & \cdot & \mathbf{L}_{\beta_j\beta_s} \end{pmatrix},$$

whose submatrices are determined in Appendix 1.

2.2. *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 performed in this method, but from an unusual point of view. An important work of implementing the Jackknife method is given by Lipsitz *et al.* [26], who suggest an alternative robust estimator of the covariance matrix based on the jackknife for analysing data from repeated measures studies. Here, we use this method as an alternative to estimate the population parameters.

Suppose that $T_1, \ldots, T_n$ is a random sample of n values and $\bar{T} = \sum_{i=1}^{n} T_i/n$ is the sample mean used to estimate the mean of the population. The sample mean calculated with the lth observation missing out is

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

for which

$$T_l = n\bar{T} - (n-1)\bar{T}_{-l}. \tag{8}$$

In a general situation, suppose that θ is a parameter estimated by $\hat{E}(T_1, \ldots, T_n)$, and for ease of notation, we drop $(T_1, \ldots, T_n)$. Thus, $\hat{E}_{-l}$ is calculated when the observation T_l is missed out. It follows, from Equation (8), that the pseudo-values can be determined by

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

The average of the pseudo-values is the Jackknife estimate of θ given by

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

Manly [27] suggested that an approximate $100(1-\alpha)\%$ CI for θ is given by $\hat{E}^* \pm t_{\alpha/2,n-1}\, s/\sqrt{n}$, where s is the standard deviation of the pseudo-values, $t_{\alpha/2,n-1}$ is the upper $(1-\alpha/2)$ point of the t-distribution with $n-1$ degrees of freedom, which has the effect of removing the bias of order n^{-1}.

The Jackknife estimate calculations for the LGG regression model for interval-censored data are performed for λ, σ and β_j $(j = 1, \ldots, p)$ and CIs are calculated separately for each parameter.

2.3. *Bootstrap re-sampling method*

The bootstrap re-sampling method, proposed by Efron [28], considers that the observed sample represents the population. From the information obtained from such sample, B bootstrap samples of similar size to that of the observed sample are generated, from which it is possible to estimate various characteristics of the population, such as the mean, variance, percentiles and so on.

According to the literature, the re-sampling method may be non-parametric or parametric. In this study, the non-parametric bootstrap method is addressed for which the distribution function F can be estimated by the empirical distribution $\hat{F}$.

Let $\mathbf{T} = (T_1, \ldots, T_n)$ be an observed random sample and $\hat{F}$ the empirical distribution of $\mathbf{T}$. Thus, a bootstrap sample $\mathbf{T}^*$ is constructed by re-sampling with the replacement of n elements from the sample $\mathbf{T}$. From the B generate bootstrap samples, $T_1^*, \ldots, T_B^*$, the bootstrap replication of the parameter of interest for the bth sample is given by

$$\hat{\boldsymbol{\theta}}_b^* = s(T_b^*),$$

that is, the value of $\hat{\boldsymbol{\theta}}$ for sample T_b^* with $b = 1, \ldots, B$.

The bootstrap estimator of the standard error [29] is the standard deviation of these bootstrap samples. It is denoted by $\widehat{EP}_B$ and can be obtained by the following result:

$$\widehat{EP}_B = \left[\frac{1}{(B-1)}\sum_{b=1}^{B}(\hat{\theta}_b^* - \bar{\theta}_B)^2\right]^{1/2},$$

where $\bar{\theta}_B = (1/B)\sum_{b=1}^{B}\hat{\theta}_b^*$. Note that B is the number of generate bootstrap samples. According to [29], assuming $B \geq 200$, it is generally sufficient to present good bootstrap estimators. However, to achieve greater accuracy, a reasonably high B-value should be considered. We describe the bias corrected and accelerated (BCa) method for constructing approximated CIs based on the bootstrap re-sampling method. For further details on bootstrap intervals, see [29–31].

2.3.1. *BCa bootstrap interval*

The bootstrap interval based on the BCa method assumes that the percentiles used in delimiting the bootstrap CIs depend on the corrections to tendency $\hat{a}$ and acceleration $\hat{z}_0$. The bias correction value $\hat{z}_0$ is generated based on the proportion of estimations of bootstrap samples that are smaller than the original estimate $\hat{\boldsymbol{\theta}}$. Equation for $\hat{z}_0$ is given by

$$\hat{z}_0 = \Phi^{-1}\left(\frac{\sharp(\hat{\boldsymbol{\theta}}_b^* < \hat{\boldsymbol{\theta}})}{B}\right), \quad b = 1, \ldots, B.$$

Here, $\Phi^{-1}(\cdot)$ is the inverse of the standard normal cumulative distribution, B the number of generated bootstrap samples, $\hat{\boldsymbol{\theta}}$ the MLE of the observed sample and $\hat{\boldsymbol{\theta}}_b^*$ the MLE of the bth bootstrap sample.

Let $\hat{\boldsymbol{\theta}}_{(i)}$ be the MLE of the sample without the ith observation. Then, $\hat{a}$ is given by

$$\hat{a} = \frac{\sum_{i=1}^{n}[\hat{\boldsymbol{\theta}}_{(\cdot)} - \hat{\boldsymbol{\theta}}_{(i)}]^3}{6\{\sum_{i=1}^{n}[\hat{\boldsymbol{\theta}}_{(\cdot)} - \hat{\boldsymbol{\theta}}_{(i)}]^2\}^{3/2}}.$$

Note that $\hat{\boldsymbol{\theta}}_{(\cdot)} = \sum_{i=1}^{n}\hat{\boldsymbol{\theta}}_{(i)}/n$ and n is the sample size.

Hence, the BCa bootstrap interval of coverage $100(1-2\alpha)\%$ can be reduced to

$$[\hat{\boldsymbol{\theta}}^*_{(B\alpha_1)}, \hat{\boldsymbol{\theta}}^*_{(B\alpha_2)}],$$

where

$$\alpha_1 = \Phi\left\{\hat{z}_0 + \frac{\hat{z}_0 + \Phi^{-1}(\alpha)}{1 - \hat{a}[\hat{z}_0 + \Phi^{-1}(\alpha)]}\right\} \quad \text{and} \quad \alpha_2 = \Phi\left\{\hat{z}_0 + \frac{\hat{z}_0 + \Phi^{-1}(1-\alpha)}{1 - \hat{a}[\hat{z}_0 + \Phi^{-1}(1-\alpha)]}\right\}.$$

The quantities α_1 and α_2 are simple corrections to the bootstrap percentiles [29].

3. Simulation study

We conducted a Monte Carlo simulation study to assess on the finite sample behaviour of the MLEs of σ, β_0 and β_1. All results were obtained from 1000 Monte Carlo replications. The simulations were carried out using the matrix programming language Ox [24]. In each replication, we generated event times of the GG distribution with parameters $\lambda = 1$ and 3. In our simulations we have one binary covariate with values drawn from a Bernoulli distribution with parameter 0.5.

Table 1. Averages of the MLEs and their standard errors (SD), CP, MCI and square RMSE for the parameters of the LGG regression model for interval-censored data.

λ	n	Parameter	Mean	RMSE	CP	MCI	SD
1	50	σ	0.463	0.152	0.937	0.572	0.147
		β_0	−1.054	0.456	0.941	1.658	0.453
		β_1	0.762	0.710	0.929	2.490	0.707
	100	σ	0.481	0.105	0.939	0.402	0.104
		β_0	−1.030	0.313	0.946	1.177	0.312
		β_1	0.746	0.463	0.931	1.744	0.461
	300	σ	0.494	0.059	0.949	0.231	0.059
		β_0	−1.010	0.178	0.945	0.679	0.176
		β_1	0.755	0.260	0.946	1.002	0.255
3	50	σ	0.469	0.141	0.949	0.551	0.137
		β_0	−0.961	0.398	0.951	1.474	0.396
		β_1	0.712	0.600	0.935	2.190	0.600
	100	σ	0.479	0.101	0.942	0.387	0.099
		β_0	−0.961	0.276	0.940	1.044	0.273
		β_1	0.684	0.396	0.944	1.539	0.396
	300	σ	0.491	0.057	0.946	0.223	0.056
		β_0	−0.970	0.154	0.950	0.604	0.151
		β_1	0.717	0.223	0.956	0.889	0.222

The censoring times were sampled from the uniform distribution on the interval $(0, \phi)$, where ϕ was set in order to control the proportion of censored observations. In this study the proportion of censored observations was on average approximately equal to 20%. The intervals of each occurrence time were calculated following the same method proposed by Hashimoto *et al.* [17]. The BFGS method (see, for example, [32]) has been used by the authors to maximize the log-likelihood. The true parameter values for the data-generating processes are: $\tau = 0.5$, $\beta_0 = -1.0$ and $\beta_1 = 0.7$. We consider sample sizes equal to $50, 100$ and 300. For each configuration, we conducted 1000 replicates and then we averaged the estimates of the parameters and obtained the standard errors (SD), mean of the CI (MCI), coverage probability (CP) of the 95% CI and the square root of the mean square error (RMSE). The figures in Table 1 show that:

- the biases and RMSEs of the MLEs of τ, β_0 and β_1 decay towards zero as the sample size increases, as expected;
- future research should be conducted to obtain bias corrections for these estimators, thus reducing their systematic errors in finite samples;
- the empirical coverage probabilities are close to the nominal coverage level when the sample size increases;
- the MCIs decrease when the sample size increases;
- the MLEs were consistent.

4. Sensitivity analysis

4.1. *Global influence*

The first tool to perform sensitivity analysis, as previously stated, is by means of global influence by starting from case deletion [3]. Case deletion is a common approach to study the effect of dropping the ith case from the data set. Case deletion for model (6) is given by

$$Y_l = \mathbf{x}_l^{\mathrm{T}} \boldsymbol{\beta} + \sigma z_l, \quad l = 1, \ldots, n, \quad l \neq i. \tag{9}$$

In the following, a quantity with subscript 'i' means the original quantity with the ith observation deleted. For model (9), the log-likelihood function is denoted by $l_{(i)}(\boldsymbol{\theta})$.

Let $\hat{\boldsymbol{\theta}}_{(i)} = (\hat{\lambda}_{(i)}, \hat{\sigma}_{(i)}, \hat{\boldsymbol{\beta}}_{(i)}^{\top})^{\top}$ be the MLE of $\boldsymbol{\theta}$ obtained by maximizing $l_{(i)}(\boldsymbol{\theta})$. To assess the influence of the ith observation on the MLE $\hat{\boldsymbol{\theta}} = (\hat{\lambda}, \hat{\sigma}, \hat{\boldsymbol{\beta}}^{\top})^{\top}$, the basic idea is to compare the difference between $\hat{\boldsymbol{\theta}}_{(i)}$ and $\hat{\boldsymbol{\theta}}$. If deletion of an observation seriously influences the estimates, more attention should be paid to that observation. Hence, if $\hat{\boldsymbol{\theta}}_{(i)}$ is far from $\hat{\boldsymbol{\theta}}$, then this 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)

$$\mathrm{GD}_i(\hat{\boldsymbol{\theta}}) = (\hat{\boldsymbol{\theta}}_{(i)} - \hat{\boldsymbol{\theta}})^{\mathrm{T}}\{\ddot{\mathbf{L}}(\hat{\boldsymbol{\theta}})\}(\hat{\boldsymbol{\theta}}_{(i)} - \hat{\boldsymbol{\theta}}).$$

Another alternative is to assess values $\mathrm{GD}_i(\boldsymbol{\beta})$ and $\mathrm{GD}_i(\sigma)$, which reveal the impact of the ith observation on the estimates of $\boldsymbol{\beta}$ and σ, respectively. Another popular measure of the difference between $\hat{\boldsymbol{\theta}}_{(i)}$ and $\hat{\boldsymbol{\theta}}$ is the likelihood displacement

$$\mathrm{LD}_i(\hat{\boldsymbol{\theta}}) = 2\{l(\hat{\boldsymbol{\theta}}) - l(\hat{\boldsymbol{\theta}}_{(i)})\}.$$

Further, we can also compute $\beta_j - \beta_{j(i)} (j = 1, 2, \ldots, p)$ to calculate the difference between $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\beta}}_{(i)}$. Alternative global influence measures are possible. One could think of the behaviour of a test statistic, such as the Wald test for explanatory variable 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

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

where $\dot{l}_{(i)}(\hat{\boldsymbol{\theta}}) = \partial l_{(i)}(\boldsymbol{\theta})/\partial\boldsymbol{\theta}$ is evaluated at $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$ [33].

4.2. *Local influence*

Another approach is suggested by Cook [7] giving weights to the observations instead of removing them. Local influence calculation can be carried out for model (6). If likelihood displacement $\mathrm{LD}(\boldsymbol{\omega}) = 2\{l(\hat{\boldsymbol{\theta}}) - l(\hat{\boldsymbol{\theta}}_{\boldsymbol{\omega}})\}$ is used, where $\hat{\boldsymbol{\theta}}_{\boldsymbol{\omega}}$ denotes the MLE under the perturbed model, the normal curvature for $\boldsymbol{\theta}$ at direction $\mathbf{d}$, $\|\mathbf{d}\| = 1$, is given by $C_{\mathbf{d}}(\boldsymbol{\theta}) = 2|\mathbf{d}^{\mathrm{T}}\boldsymbol{\Delta}^{\mathrm{T}}[\ddot{\mathbf{L}}(\boldsymbol{\theta})]^{-1}\boldsymbol{\Delta}\mathbf{d}|$, where $\boldsymbol{\Delta}$ is a $(p+2)n$ matrix that depends on the perturbation scheme, whose elements are given by $\Delta_{vi} = \partial^2 l(\boldsymbol{\theta}|\boldsymbol{\omega})/\partial\theta_v\partial\omega_i$, for $i = 1, 2, \ldots, n$ and $v = 1, 2, \ldots, p+2$, evaluated at $\hat{\boldsymbol{\theta}}$ and $\boldsymbol{\omega}_0$ and $\boldsymbol{\omega}_0$ is the no perturbation vector. The elements of $\ddot{\mathbf{L}}(\boldsymbol{\theta})$ in the LGG regression model with interval-censored data are given in Appendix 1. We can also calculate normal curvatures $C_{\mathbf{d}}(\lambda)$, $C_{\mathbf{d}}(\sigma)$ and $C_{\mathbf{d}}(\boldsymbol{\beta})$ to construct various index plots, for instance, the index plot of $\mathbf{d}_{\max}$, the eigenvector corresponding to $C_{\mathbf{d}_{\max}}$, the largest eigenvalue of the matrix $\mathbf{B} = -\boldsymbol{\Delta}^{\mathrm{T}}[\ddot{\mathbf{L}}(\boldsymbol{\theta})]^{-1}\boldsymbol{\Delta}$ and the index plots of $C_{\mathbf{d}_i}(\lambda)$, $C_{\mathbf{d}_i}(\sigma)$ and $C_{\mathbf{d}_i}(\boldsymbol{\beta})$, called the total local influence (see, for example, [34]), where $\mathbf{d}_i$ denotes an $n \times 1$ vector of zeros with one at the ith position. Thus, the curvature at direction $\mathbf{d}_i$ takes the form $C_i = 2|\boldsymbol{\Delta}_i^{\mathrm{T}}[\ddot{\mathbf{L}}(\boldsymbol{\theta})]^{-1}\boldsymbol{\Delta}_i|$, where $\boldsymbol{\Delta}_i^{\mathrm{T}}$ denotes the ith row of $\boldsymbol{\Delta}$. It is usual to point out those cases such that $C_i \geq 2\bar{C}$, where $\bar{C} = (1/n)\sum_{i=1}^{n} C_i$.

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

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

for model (6) and its associated log-likelihood function (7). Consider the vector of weights $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_n)^{\mathrm{T}}$.

- *Case-weight perturbation*: In this case, the log-likelihood function takes the form $l(\boldsymbol{\theta}|\boldsymbol{\omega}) = \sum_{i \in F} \omega_i l_1(\lambda, zu_i, zv_i) + \sum_{i \in C} \omega_i l_2(\lambda, zu_i)$, where $0 \le \omega_i \le 1$ and $\boldsymbol{\omega}_0 = (1, \dots, 1)^\top$ and $l_1(\cdot)$ and $l_2(\cdot)$ are defined in Equation (7). The matrix $\boldsymbol{\Delta} = (\boldsymbol{\Delta}_\lambda, \boldsymbol{\Delta}_\sigma, \boldsymbol{\Delta}_{\boldsymbol{\beta}})^\top$ is determined numerically.
- *Response perturbation* ($\log(u_i)$): Here, we consider that each u_i is perturbed as $u_{iw} = u_i + \omega_i S_u$, where S_u is a scale factor that may be equal to the estimated standard deviation of U, $\omega_i \in \mathbf{R}$. Here, the perturbed log-likelihood function can be expressed as $l(\boldsymbol{\theta}|\boldsymbol{\omega}) = \sum_{i \in F} l_1(\lambda, zu_i^*, zv_i) + \sum_{i \in C} l_2(\lambda, zu_i^*)$, where $zu_i^* = [\log(u_i^*) - \mathbf{x}_i^\top \hat{\boldsymbol{\beta}}]/\hat{\sigma}$, $u_i^* = [\log(u_i) + \omega_i S_u]$, $\boldsymbol{\omega}_0 = (0, \dots, 0)^{\mathrm{T}}$ and $l_1(\cdot)$ and $l_2(\cdot)$ are defined in Equation (7). The matrix $\boldsymbol{\Delta} = (\boldsymbol{\Delta}_\lambda, \boldsymbol{\Delta}_\sigma, \boldsymbol{\Delta}_{\boldsymbol{\beta}})^\top$ is determined numerically.
- *Response perturbation* ($\log(v_i)$): Here, we consider that each v_i is perturbed as $v_{iw} = v_i + \omega_i S_v$, where S_v is a scale factor that may be equal to the estimated standard deviation of V, $\omega_i \in \mathbf{R}$. Then, the perturbed log-likelihood function can be expressed as $l(\boldsymbol{\theta}|\boldsymbol{\omega}) = \sum_{i \in F} l_1(\lambda, zu_i, zv_i^*) + \sum_{i \in C} l_2(\lambda, zu_i)$, where $zv_i^* = [\log(v_i^*) - \mathbf{x}_i^\top \hat{\boldsymbol{\beta}}]/\hat{\sigma}$, $v_i^* = [\log(v_i) + \omega_i S_v]$, $\boldsymbol{\omega}_0 = (0, \dots, 0)^\top$ and $l_1(\cdot)$ and $l_2(\cdot)$ are defined in Equation (7). The matrix $\boldsymbol{\Delta} = (\boldsymbol{\Delta}_\lambda, \boldsymbol{\Delta}_\sigma, \boldsymbol{\Delta}_{\boldsymbol{\beta}})^\top$ is obtained numerically.
- *Simultaneous response perturbation* ($\log(u_i), \log(v_i)$): We consider that each of the u_i and v_i is perturbed as $u_{iw} = u_i + \omega_i S_u$, $v_{iw} = v_i + \omega_i S_v$, respectively, where S_u and S_v are scale factors that may be equal to the estimated standard deviation of U and V, $\omega_i \in \mathbf{R}$. So, the perturbed log-likelihood function can be expressed as $l(\boldsymbol{\theta}|\boldsymbol{\omega}) = \sum_{i \in F} l_1(\lambda, zu_i^*, zv_i^*) + \sum_{i \in C} l_2(\lambda, zu_i^*)$, where $zu_i^* = [\log(u_i^*) - \mathbf{x}_i^\top \hat{\boldsymbol{\beta}}]/\hat{\sigma}$, $\log(u_i^*) = (\log(u_i) + \omega_i S_u)$, $zv_i^* = [\log(v_i^*) - \mathbf{x}_i^\top \hat{\boldsymbol{\beta}}]/\hat{\sigma}$, $\log(v_i^*) = [\log(v_i) + \omega_i S_v]$, $\boldsymbol{\omega}_0 = (0, \dots, 0)^\top$ and $l_1(\cdot)$ and $l_2(\cdot)$ are defined in Equation (7). The matrix $\boldsymbol{\Delta} = (\boldsymbol{\Delta}_\lambda, \boldsymbol{\Delta}_\sigma, \boldsymbol{\Delta}_{\boldsymbol{\beta}})^\top$ is determined numerically.
- *Explanatory variable perturbation*: Now, consider an additive perturbation on a particular continuous explanatory variable, namely X_t, by taking $x_{it\omega} = x_{it} + \omega_i S_x$, where S_x is a scale factor, $\omega_i \in \mathbf{R}$. This perturbation scheme leads to the following expression for the log-likelihood function: $l(\boldsymbol{\theta}|\boldsymbol{\omega}) = \sum_{i \in F} l_1(\lambda, zu_i^{**}, zv_i^{**}) + \sum_{i \in C} l_2(\lambda, zu_i^{**})$, where $zu_i^{**} = [\log(u_i) - \mathbf{x}_i^{*\top} \boldsymbol{\beta}]/\sigma$, $zv_i^{**} = [\log(v_i) - \mathbf{x}_i^{*\top} \boldsymbol{\beta}]/\sigma$, $\mathbf{x}_i^{*\top} \boldsymbol{\beta} = \beta_1 + \beta_2 x_{i2} + \cdots + \beta_t(x_{it} + \omega_i S_x) + \cdots + \beta_p x_{ip}$, $\boldsymbol{\omega}_0 = (0, \dots, 0)^\top$ and $l_1(\cdot)$ and $l_2(\cdot)$ are defined in Equation (7). The matrix $\boldsymbol{\Delta} = (\boldsymbol{\Delta}_\lambda, \boldsymbol{\Delta}_\sigma, \boldsymbol{\Delta}_{\boldsymbol{\beta}})^\top$ is obtained numerically.

All the programmes for estimating and for calculating the diagnostic measures are available from the authors upon request.

5. Application

We consider a data set on the AIDS cohort study of haemophiliacs discussed in [35,36]. This study consists of individuals with Type A or B haemophilia, who were at risk for HIV infection through the contaminated blood factor they received for their treatment. The subjects were classified into two groups, lightly and heavily treated groups, according to the amount of blood they received and age indicators that indicate whether the age of a subject was below 20 at his or her HIV infection. There were 257 individuals in the original study. In the following, we will focus on 256 subjects who were known to be infected by HIV before the end of the study. Note that, in the original data set, most of the observations for AIDS diagnosis are exact or right-censored and the remainder are interval-censored with only two time points included in each interval. One objective of the study is to test and estimate the possible difference of the AIDS incubation distributions between the two groups. The covariates considered in the models are: x_{i1}, age ($0 =$ above 20, $1 =$ below 20) and x_{i2}, treated group ($0 =$ lightly, $1 =$ heavily).

5.1. *Estimation*

5.1.1. *Maximum-likelihood and Jackknife estimation*

The MLEs of the parameters in the LGG regression model for interval-censored data are calculated using the subroutine MAXBFGS in Ox, whose results are listed in Table 2. Additionally, in Table 2, we report the Jackknife estimates. To obtain MLEs of parameters of the LGG regression model, we fix different values for λ. We choose the value the maximizes the likelihood function over several values of $\lambda \in (-0.5, 5)$, obtaining $\lambda = 1.1$ (see, Figure 2). Hence this value is assumed as the MLE of λ.

We can observe that the explanatory variable x_2 is significant (at 5%) for the log-survival time. Note that the estimates in Table 2 from the two methods seem very similar. From this table, we note that $\hat{\beta}_1$ is close to zero thus indicating that the values of x_1 do not explain the estimated survival probability. On the other hand, the negative value of $\hat{\beta}_2$ indicates that the survival probabilities for the patients in the lightly treated group ($x_2 = 0$) are greater than those of the probabilities for patients in the heavily treated group ($x_2 = 1$).

5.1.2. *Bootstrap re-sampling method*

We considered $B = 3000$ bootstrap samples of the LGG regression model with interval censoring. Using the bootstrap method described in Section 2.3, we obtain the estimated bootstrap and the BCa CIs given in Table 3.

The estimates from the three methods are quite close. The MLEs seem more conservative (the standard errors are smaller). Then, since normality for the Jackknife estimator is expected for this

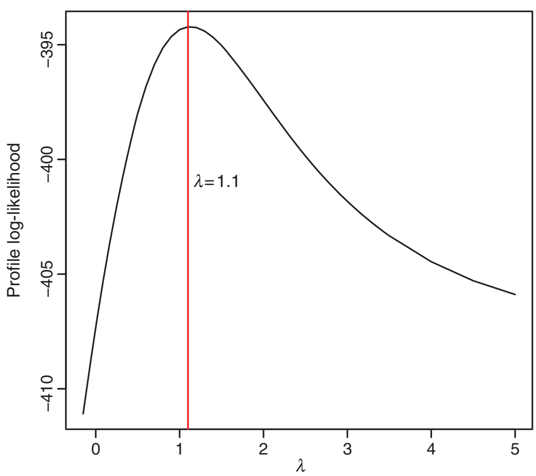

Figure 2. Profile log-likelihood.

Table 2. Maximum-likelihood and Jackknife estimates for the parameters of the LGG regression model for interval-censored fitted to haemophilia data.

	MLEs				Jackknife estimates		
$\boldsymbol{\theta}$	Estimate	SE	p-Value	95% CI	Estimate	SE	95% CI
σ	0.247	0.018	–	(0.211; 0.283)	0.247	0.020	(0.286; 0.208)
β_0	2.774	0.031	0.000	(2.714; 2.835)	2.775	0.033	(2.839; 2.711)
β_1	0.047	0.040	0.247	(−0.032; 0.125)	0.047	0.044	(−0.039; 0.133)
β_2	−0.217	0.039	<0.001	(−0.293; −0.141)	−0.216	0.039	(−0.294; −0.138)

Table 3. Bootstrap estimate and CIs based on the non-parametric bootstrap re-sampling method for haemophilia data.

	Bootstrap estimates		
$\boldsymbol{\theta}$	Estimate	SE	95% CI (BCA)
σ	0.246	0.020	(0.221; 0.285)
β_0	2.774	0.023	(2.737; 2.813)
β_1	0.047	0.045	(−0.026; 0.122)
β_2	−0.216	0.035	(−0.275; −0.161)

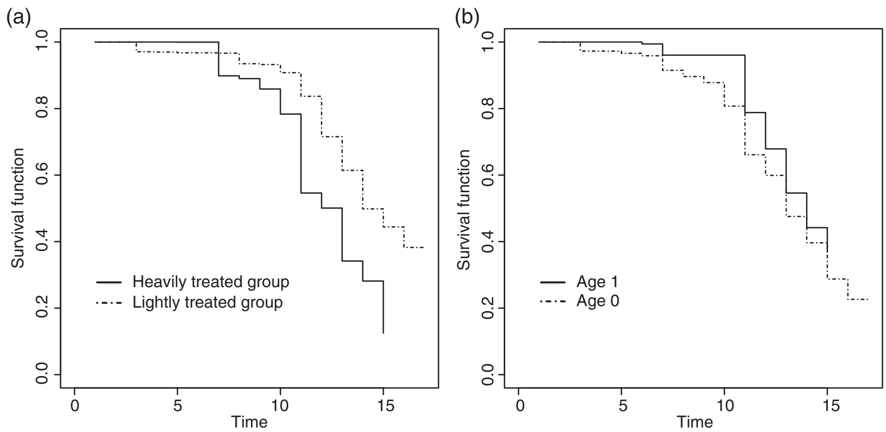

Figure 3. Product-limit estimator for the survival function for haemophilia data. (a) For heavily and lightly treated group and (b) for age.

sample size ($n = 256$), we can consider that the MLEs should follow a symmetric distribution with heavy tails. We will continue the analysis using the MLEs and considering the LGG regression models.

For the sake of illustration, Figure 3 shows the estimations by product-limit estimator [37] for the survival function by treatment (heavily and lightly treated groups) and age separately. The algorithm of the product-limit estimator for interval-censored data used in this analysis was described by Colosimo and Giolo [38]. There is no evidence for the difference of the empirical survival functions for the two groups of ages. Such difference is more accentuated between heavily and lightly treated groups. The plots in Figure 3 show that the curves do not have significative differences at the beginning of the study. However, for individuals with times event up to 8, we note that the survival function decreases considerably for the patients in the heavily treated group. Further, there is no difference for the survival times for individuals with age less than and greater than 20 years old.

Therefore, based on these estimates, we obtain the final regression equation

$$y_i = \beta_0 + \beta_2 x_{i2}, \quad i = 1, \ldots, 256. \tag{10}$$

The MLEs for the parameters in the final model (10) are listed in Table 4.

5.2. *Sensitivity analysis*

In this section, we use Ox to compute the case-deletion and local influence measures for the haemophilia data set using the LGG regression model (10).

Table 4. MLEs from the fit of the LGG regression model for interval-censored fitted to the haemophilia data.

$\boldsymbol{\theta}$	Estimate	SE	p-Value	95% CI
σ	0.248	0.018	–	(0.212; 0.284)
β_0	2.793	0.027	0.000	(2.739; 2.846)
β_2	−0.222	0.039	<0.001	(−0.298; −0.146)

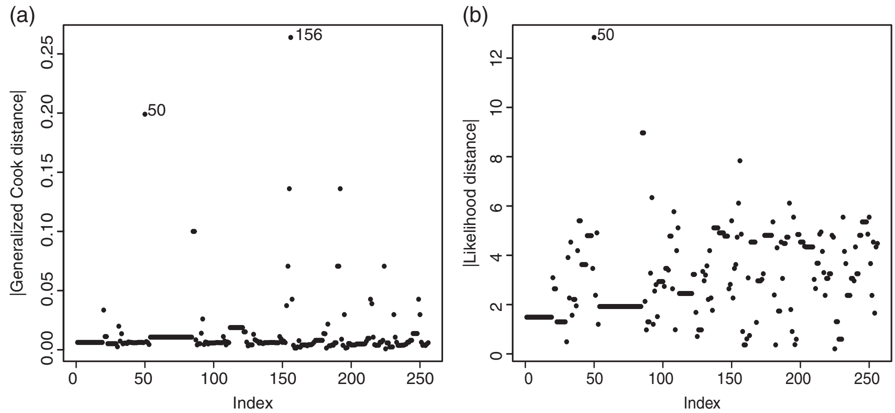

Figure 4. (a) Index plot of $\mathrm{GD}_i(\boldsymbol{\theta})$ (generalized Cook's distance) and (b) index plot of $\mathrm{LD}_i(\boldsymbol{\theta})$ (likelihood distance).

The case-deletion measures $\mathrm{GD}_i(\boldsymbol{\theta})$ and $\mathrm{LD}_i(\boldsymbol{\theta})$ are presented in Section 4.1. The results of such influence measure index plots are displayed in Figure 4. From this figure, we note that the cases ♯50 and ♯156 are possible influential observations.

5.2.1. *Case-weight perturbation*

By applying the local influence theory discussed in Section 4.2, where case-weight perturbation is used, the value $C_{\mathbf{d}_{\max}} = 1.25$ was obtained as a maximum curvature. In Figure 5(a), the plot of the eigenvector corresponding to $\mathbf{d}_{\max}$ is presented, and the total influence C_i is shown in Figure 5(b). The observations ♯50 and ♯156 are very distinguished in relation to the others.

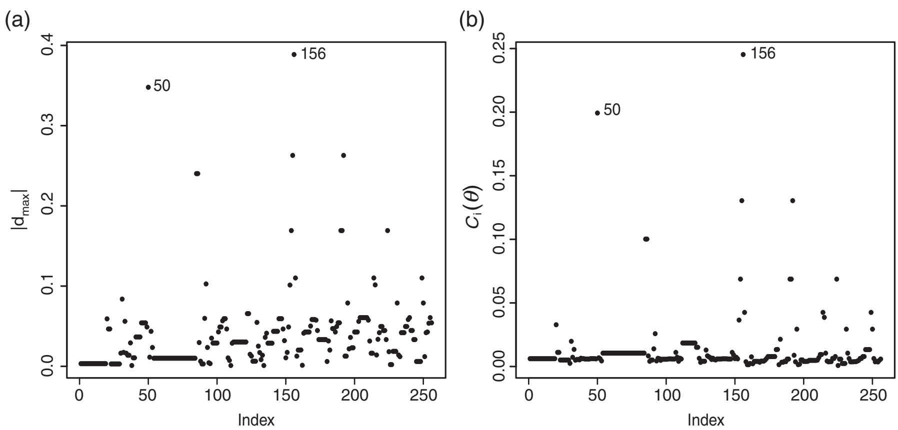

Figure 5. (a) Index plot of $\mathbf{d}_{\max}$ for $\boldsymbol{\theta}$ (case-weight perturbation) and (b) total local influence for $\boldsymbol{\theta}$ (case-weight perturbation) from the fit of the LGG model to the haemophilia data.

5.2.2. *Response variable perturbation*

Next, the influence of perturbations on the observed survival times will be analysed (simultaneous response perturbation $(\log(U_i), \log(V_i))$). The value for the maximum curvature was $C_{\mathbf{d}_{\max}} = 1.54$. In Figure 6(a), we plot $\mathbf{d}_{\max}$ versus the observation index which shows that the observation ♯156 is more salient in relation to the others. Figure 6(b) provides plots for the total local influence (C_i), where again the observation ♯156 stands out.

5.3. *Impact of the detected influential observations*

Hence, the diagnostic analysis (global influence and local influence) detected the observations ♯50 and ♯156 as potentially influential. The observation ♯156 presents a high lifetime value (right-censored). Moreover, the observation ♯50 represents the lowest interval-censored data. To reveal the impact of these two observations on the parameter estimates, we re-fitted the model under some situations. First, we individually eliminated each one of these two cases. Next, we removed the totality of potentially influential observations from the set 'A' (original data set).

In Table 5, we give the relative changes (in percentage) of each parameter estimate, defined by $\mathbf{RC}_{\theta_j} = [(\hat{\theta}_j - \hat{\theta}_{j(I)})/\hat{\theta}_j] \times 100$, parameter estimates and the corresponding p-values, where

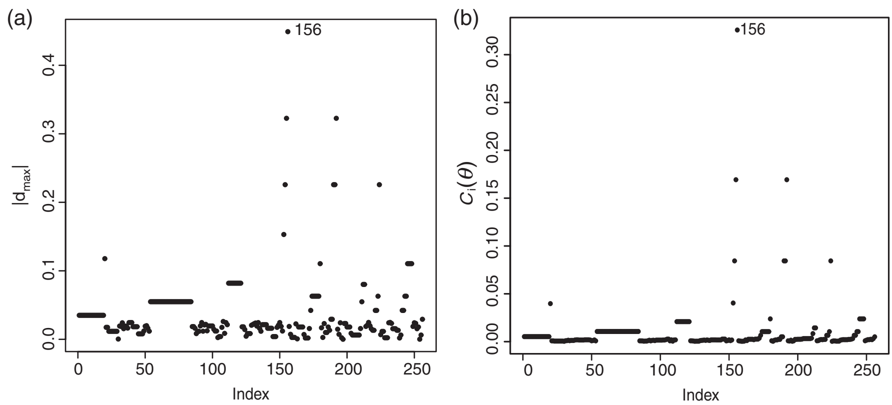

Figure 6. (a) Index plot of $\mathbf{d}_{\max}$ for $\boldsymbol{\theta}$ (simultaneous response perturbation) and (b) total local influence for $\boldsymbol{\theta}$ (simultaneous response perturbation) from the model fitting to the haemophilia data.

Table 5. Relative changes [-**RC**- in %], estimates and the corresponding p-values in parentheses for the regression coefficients to explain the log-survival time.

Set{I}	σ	β_0	β_2
None	–	–	–
	0.248	2.793	−0.222
	(–)	(0.000)	(<0.001)
Set I_1	[3]	[0]	[1]
	0.240	2.793	−0.220
	(–)	(0.000)	(<0.001)
Set I_2	[17]	[2]	[−37]
	0.207	2.739	−0.304
	(–)	(0.000)	(<0.001)
Set I_3	[5]	[0]	[−3]
	0.235	2.792	−0.228
	(–)	(0.000)	(<0.001)

$\hat{\theta}_{j(I)}$ denotes the MLE of θ_j after the set 'I' of observations being removed. Table 5 provides the following sets: $I_1 = \{\sharp 50\}$, $I_2 = \{\sharp 156\}$ and $I_3 = \{\sharp 50, \sharp 156\}$. From Table 5, we note that the MLEs from the LGG regression model for interval-censored are highly robust under deletion of the outstanding observations. In general, the significance of the parameter estimates does not change (at the level of 1%) after removing the set I. Therefore, we do not have inferential changes after removing the observations handed out in the diagnostic plots.

6. Concluding remarks

In this paper, we introduce and study the LGG regression model for interval-censored data. We adopt the quasi-Newton algorithm to obtain the MLEs and use likelihood ratio tests for testing the model parameters. On the other hand, as an alternative analysis, we discuss the use of the Jackknife estimator and non-parametric bootstrap for the LGG regression model for interval-censored data. The required matrices for the application of the techniques were obtained by taking into account some usual perturbation in the model/data. By applying the procedures in a data set from the medical area, we can assess the sensitivity aspects of the MLEs under the perturbation schemes and check the goodness of fit of the postulated model. Although the diagnostic plots detected some possible influential observations, their deletion does not cause inferential changes in the results.

Acknowledgements

We would like to thank the reviewers for the constructive comments. This study was supported by FAPESP Grant 2010/04496-2 and CNPq, Brazil.

References

[1] J.F. Lawless, *Statistical Models and Methods for Lifetime Data*, Wiley, New York, 2003.
[2] E.W. Stacy, *A generalization of the gamma distribution*, Ann. Math. Statist. 33 (1962), pp. 1187–1192.
[3] R.D. Cook, *Detection of influential observations in linear regression*, Technometrics 19 (1977), pp. 15–18.
[4] R. Christensen, L. Pearson, and W. Johnson, *Case-deletion diagnostics for mixed models*, Technometrics 34 (1992), pp. 38–45.
[5] A.C. Davison and C.L. Tsai, *Regression model diagnostics*, Int. Statist. Rev. 60 (1992), pp. 337–355.
[6] F. Xie and B.C. Wei, *Diagnostics analysis for log-Birnbaum–Saunders regression models*, Comput. Statist. Data Anal. 51 (2007), pp. 4692–4706.
[7] R.D. Cook, *Assessment of local influence* (*with discussion*), J. Roy. Statist. Soc. Ser. B 48 (1986), pp. 133–169.
[8] A.N. Pettitt and I. Bin Daud, *Case-weight measures of influence for proportional hazards regression*, Appl. Statist. 38 (1989), pp. 51–67.
[9] L.A. Escobar and W.Q. Meeker, *Assessing influence in regression analysis with censored data*, Biometrics 48 (1992), pp. 507–528.
[10] E.M.M. Ortega, H. Bolfarine, and G.A. Paula, *Influence diagnostics in generalized log-gamma regression models*, Comput. Statist. Data Anal. 42 (2003), pp. 165–186.
[11] J.R. Magnus and A.L. Vasnev, *Local sensitivity and diagnostic tests*, Econometrics J. 10 (2007), pp. 166–192.
[12] F.C. Xie and B.C. Wei, *Diagnostics analysis in censored generalized Poisson regression model*, J. Statist. Comput. Simul. 77 (2007), pp. 695–708.
[13] J.B. Fachini, E.M.M. Ortega, and F. Louzada-Neto, *Influence diagnostics for poly-hazard models in the presence of covariates*, Statist. Methods Appl. 17 (2008), pp. 413–433.
[14] J.M.F. Carrasco, E.M.M. Ortega, and G.A. Paula, *Log-modified Weibull regression models with censored data: Sensitivity and residual analysis*, Comput. Statist. Data Anal. 52 (2008), pp. 4021–4029.
[15] G.O. Silva, E.M.M. Ortega, V.C. Garibay, and M.L. Barreto, *Log-Burr XII regression models with censored data*, Comput. Statist. Data Anal. 52 (2008), pp. 3820–3842.
[16] E.M.M. Ortega, V.G. Cancho, and G.A. Paula, *Generalized log-gamma regression models with cure fraction*, Lifetime Data Anal. 15 (2009), pp. 79–106.
[17] E.M. Hashimoto, E.M.M. Ortega, V.G. Cancho, and G.M. Cordeiro, *The log-exponentiated Weibull regression model for interval-censored data*, Comput. Statist. Data Anal. 54 (2010), pp. 1017–1035.
[18] N.L. Johnson, S. Kotz, and N. Balakrishnan, *Continuous Univariate Distributions*, Wiley, New York, 1994.
[19] C. Cox, H. Chu, M.F. Schneider, and A. Muñz, *Parametric survival analysis and taxonomy of hazard functions for the generalized gamma distribution*, Statist. Med. 26 (2007), pp. 4352–4374.

[20] G. Almpanidis and C. Kotropoulos, *Phonemic segmentation using the generalized gamma distribution and small sample Bayesian information criterion*, Speech Commun. 50 (2008), pp. 38–55.
[21] S. Nadarajah, *On the use of the generalized gamma distribution*, Int. J. Electron. 95 (2008), pp. 1029–1032.
[22] E.M.M. Ortega, G.A. Paula, and H. Bolfarine, *Deviance residuals in generalized log-gamma regression models with censored observations*, J. Statist. Comput. Simul. 78 (2008), pp. 747–764.
[23] O. Gomes, C. Combes, and A. Dussauchoy, *Parameter estimation of the generalized gamma distribution*, Math. Comput. Simul. 79 (2008), pp. 955–963.
[24] J. Doornik, *Ox: Object-Oriented Matrix Programming Using Ox*, 5th ed., Timberlake Consultants Ltd, London, 2007.
[25] P.K. Sen and J.M. Singer, *Large Sample Methods in Statistics: An Introduction with Applications*, Chapman and Hall, New York, 1993.
[26] S.R. Lipsitz, N.M. Laird, and D.P. Harrington, *Using the Jackknife to estimate the variance of regression estimators from repeated measures studies*, Comm. Statist. Theory Methods 19 (1990), pp. 821–845.
[27] B.F.J. Manly, *Randomization, Bootstrap and Monte Carlo Methods in Biology*, 2nd ed., Chapman and Hall, London, 1997.
[28] B. Efron, *Bootstrap methods: another look at the jackknife*, Ann. Statist. 7 (1979), pp. 1–26.
[29] B. Efron and R.J. Tibshirani, *An Introduction to the Bootstrap*, Chapman & Hall, New York, 1993.
[30] T.J. DiCiccio and B. Efron, *Bootstrap confidence intervals*, Statist. Sci. 11 (1996), pp. 189–228.
[31] A.C. Davison and D.V. Hinkley, *Bootstrap Methods and their Application*, Cambridge University Press, New York, 1997.
[32] W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery, *Numerical Recipes in C: The Art of Scientific Computing*, 2nd ed., Cambridge University Press, Cambridge, 1992.
[33] R.D. Cook and S. Weisberg, *Residuals and Influence in Regression*, Chapman & Hall, New York, 1982.
[34] E. Lesaffre and G. Verbeke, *Local influence in linear mixed models*, Biometrics 54 (1998), pp. 570–582.
[35] M.Y. Kim, V.G. DeGruttola, and S.W. Lagakos, *Analyzing doubly censored data with covariates, with application to AIDS*, Biometrics 49 (1993), pp. 13–22.
[36] J. Sun, Q. Liao, and M. Pagano, *Regression analysis of doubly censored failure time data with applications to AIDS studies*, Biometrics 55 (1999), pp. 909–914.
[37] W. Turnbull, *The empirical distribution function with arbitrarily grouped censored and truncated data*, J. R. Statist. Soc. Ser. B 38 (1976), pp. 290–295.
[38] E.A. Colosimo and S.R. Giolo, *Análise de sobrevivência aplicada*, Edgard Blücher, São Paulo, 2006.

Appendix 1: Matrix of second derivatives $\ddot{\mathbf{L}}(\boldsymbol{\theta})$

Here, we provide the elements of the observed information matrix. After some algebraic manipulations, we can obtain the following:

For $\lambda > 0$:

$$\mathbf{L}_{\lambda\lambda} = \sum_{i\in F}\left\{\frac{[\ddot{g}(\lambda, zu_i)]_{\lambda\lambda} - [\ddot{g}(\lambda, zv_i)]_{\lambda\lambda}}{g(\lambda, zu_i) - g(\lambda, zv_i)} - \left[\frac{[\dot{g}(\lambda, zu_i)]_{\lambda} - [\dot{g}(\lambda, zv_i)]_{\lambda}}{g(\lambda, zu_i) - g(\lambda, zv_i)}\right]^2\right\}$$

$$+ \sum_{i\in C}\left\{\frac{[\ddot{g}(\lambda, zu_i)]_{\lambda\lambda}}{g(\lambda, zu_i)} - \left[\frac{[\dot{g}(\lambda, zu_i)]_{\lambda}}{g(\lambda, zu_i)}\right]^2\right\},$$

$$\mathbf{L}_{\lambda\sigma} = \sum_{i\in F}\left\{\frac{[\ddot{g}(\lambda, zu_i)]_{\lambda\sigma} - [\ddot{g}(\lambda, zv_i)]_{\lambda\sigma}}{g(\lambda, zu_i) - g(\lambda, zv_i)} - \frac{\{[\dot{g}(\lambda, zu_i)]_{\lambda} - [\dot{g}(\lambda, zv_i)]_{\lambda}\}\{[\dot{g}(\lambda, zu_i)]_{\sigma} - [\dot{g}(\lambda, zv_i)]_{\sigma}\}}{[g(\lambda, zu_i) - g(\lambda, zv_i)]^2}\right\}$$

$$+ \sum_{i\in C}\left\{\frac{[\ddot{g}(\lambda, zu_i)]_{\lambda\sigma}}{g(\lambda, zu_i)} - \frac{[\dot{g}(\lambda, zu_i)]_{\lambda}[\dot{g}(\lambda, zu_i)]_{\sigma}}{[g(\lambda, zu_i)]^2}\right\},$$

$$\mathbf{L}_{\lambda\beta_j} = \sum_{i\in F}\left\{\frac{[\ddot{g}(\lambda, zu_i)]_{\lambda\beta_j} - [\ddot{g}(\lambda, zv_i)]_{\lambda\beta_j}}{g(\lambda, zu_i) - g(\lambda, zv_i)} - \frac{\{[\dot{g}(\lambda, zu_i)]_{\lambda} - [\dot{g}(\lambda, zv_i)]_{\lambda}\}\{[\dot{g}(\lambda, zu_i)]_{\beta_j} - [\dot{g}(\lambda, zv_i)]_{\beta_j}\}}{[g(\lambda, zu_i) - g(\lambda, zv_i)]^2}\right\}$$

$$+ \sum_{i\in C}\left\{\frac{[\ddot{g}(\lambda, zu_i)]_{\lambda\beta_j}}{g(\lambda, zu_i)} - \frac{[\dot{g}(\lambda, zu_i)]_{\lambda}[\dot{g}(\lambda, zu_i)]_{\beta_j}}{[g(\lambda, zu_i)]^2}\right\},$$

$$\mathbf{L}_{\sigma\sigma} = \sum_{i\in F}\left\{\frac{[\ddot{g}(\lambda, zu_i)]_{\sigma\sigma} - [\ddot{g}(\lambda, zv_i)]_{\sigma\sigma}}{g(\lambda, zu_i) - g(\lambda, zv_i)} - \left[\frac{[\dot{g}(\lambda, zu_i)]_{\sigma} - [\dot{g}(\lambda, zv_i)]_{\sigma}}{g(\lambda, zu_i) - g(\lambda, zv_i)}\right]^2\right\}$$

$$+ \sum_{i\in C}\left\{\frac{[\ddot{g}(\lambda, zu_i)]_{\sigma\sigma}}{g(\lambda, zu_i)} - \left[\frac{[\dot{g}(\lambda, zu_i)]_{\sigma}}{g(\lambda, zu_i)}\right]^2\right\},$$

$$\mathbf{L}_{\sigma\beta_j} = \sum_{i\in F}\left\{\frac{[\ddot{g}(\lambda,zu_i)]_{\sigma\beta_j} - [\ddot{g}(\lambda,zv_i)]_{\sigma\beta_j}}{g(\lambda,zu_i)-g(\lambda,zv_i)} - \frac{\{[\dot{g}(\lambda,zu_i)]_{\sigma} - [\dot{g}(\lambda,zv_i)]_{\sigma}\}\{[\dot{g}(\lambda,zu_i)]_{\beta_j} - [\dot{g}(\lambda,zv_i)]_{\beta_j}\}}{[g(\lambda,zu_i)-g(\lambda,zv_i)]^2}\right\}$$
$$+\sum_{i\in C}\left\{\frac{[\ddot{g}(\lambda,zu_i)]_{\sigma\beta_j}}{g(\lambda,zu_i)} - \frac{[\dot{g}(\lambda,zu_i)]_{\sigma}[\dot{g}(\lambda,zu_i)]_{\beta_j}}{[g(\lambda,zu_i)]^2}\right\},$$

$$\mathbf{L}_{\beta_j\beta_s} = \sum_{i\in F}\left\{\frac{[\ddot{g}(\lambda,zu_i)]_{\beta_j\beta_s} - [\ddot{g}(\lambda,zv_i)]_{\beta_j\beta_s}}{g(\lambda,zu_i)-g(\lambda,zv_i)} - \left[\frac{[\dot{g}(\lambda,zu_i)]_{\beta_j} - [\dot{g}(\lambda,zv_i)]_{\beta_j}}{g(\lambda,zu_i)-g(\lambda,zv_i)}\right]^2\right\}$$
$$+\sum_{i\in C}\left\{\frac{[\ddot{g}(\lambda,zu_i)]_{\beta_j\beta_s}}{g(\lambda,zu_i)} - \left[\frac{[\dot{g}(\lambda,zu_i)]_{\beta_j}}{g(\lambda,zu_i)}\right]^2\right\}.$$

For $\lambda < 0$:

$$\mathbf{L}_{\lambda\lambda} = \sum_{i\in F}\left\{\frac{[\ddot{g}(\lambda,zv_i)]_{\lambda\lambda} - [\ddot{g}(\lambda,zu_i)]_{\lambda\lambda}}{g(\lambda,zv_i)-g(\lambda,zu_i)} - \left[\frac{[\dot{g}(\lambda,zv_i)]_{\lambda} - [\dot{g}(\lambda,zu_i)]_{\lambda}}{g(\lambda,zv_i)-g(\lambda,zu_i)}\right]^2\right\}$$
$$-\sum_{i\in C}\left\{\frac{[\ddot{g}(\lambda,zu_i)]_{\lambda\lambda}}{1-g(\lambda,zu_i)} - \left[\frac{[\dot{g}(\lambda,zu_i)]_{\lambda}}{1-g(\lambda,zu_i)}\right]^2\right\},$$

$$\mathbf{L}_{\lambda\sigma} = \sum_{i\in F}\left\{\frac{[\ddot{g}(\lambda,zv_i)]_{\lambda\sigma} - [\ddot{g}(\lambda,zu_i)]_{\lambda\sigma}}{g(\lambda,zv_i)-g(\lambda,zu_i)} - \frac{\{[\dot{g}(\lambda,zv_i)]_{\lambda} - [\dot{g}(\lambda,zu_i)]_{\lambda}\}\{[\dot{g}(\lambda,zv_i)]_{\sigma} - [\dot{g}(\lambda,zu_i)]_{\sigma}\}}{[g(\lambda,zv_i)-g(\lambda,zu_i)]^2}\right\}$$
$$-\sum_{i\in C}\left\{\frac{[\ddot{g}(\lambda,zu_i)]_{\lambda\sigma}}{1-g(\lambda,zu_i)} - \frac{[\dot{g}(\lambda,zu_i)]_{\lambda}[\dot{g}(\lambda,zu_i)]_{\sigma}}{[1-g(\lambda,zu_i)]^2}\right\},$$

$$\mathbf{L}_{\lambda\beta_j} = \sum_{i\in F}\left\{\frac{[\ddot{g}(\lambda,zv_i)]_{\lambda\beta_j} - [\ddot{g}(\lambda,zu_i)]_{\lambda\beta_j}}{g(\lambda,zv_i)-g(\lambda,zu_i)} - \frac{\{[\dot{g}(\lambda,zv_i)]_{\lambda} - [\dot{g}(\lambda,zu_i)]_{\lambda}\}\{[\dot{g}(\lambda,zv_i)]_{\beta_j} - [\dot{g}(\lambda,zu_i)]_{\beta_j}\}}{[g(\lambda,zv_i)-g(\lambda,zu_i)]^2}\right\}$$
$$-\sum_{i\in C}\left\{\frac{[\ddot{g}(\lambda,zu_i)]_{\lambda\beta_j}}{1-g(\lambda,zu_i)} - \frac{[\dot{g}(\lambda,zu_i)]_{\lambda}[\dot{g}(\lambda,zu_i)]_{\beta_j}}{[1-g(\lambda,zu_i)]^2}\right\},$$

$$\mathbf{L}_{\sigma\sigma} = \sum_{i\in F}\left\{\frac{[\ddot{g}(\lambda,zv_i)]_{\sigma\sigma} - [\ddot{g}(\lambda,zu_i)]_{\sigma\sigma}}{g(\lambda,zv_i)-g(\lambda,zu_i)} - \left[\frac{[\dot{g}(\lambda,zv_i)]_{\sigma} - [\dot{g}(\lambda,zu_i)]_{\sigma}}{g(\lambda,zv_i)-g(\lambda,zu_i)}\right]^2\right\}$$
$$-\sum_{i\in C}\left\{\frac{[\ddot{g}(\lambda,zu_i)]_{\sigma\sigma}}{1-g(\lambda,zu_i)} - \left[\frac{[\dot{g}(\lambda,zu_i)]_{\sigma}}{1-g(\lambda,zu_i)}\right]^2\right\},$$

$$\mathbf{L}_{\sigma\beta_j} = \sum_{i\in F}\left\{\frac{[\ddot{g}(\lambda,zv_i)]_{\sigma\beta_j} - [\ddot{g}(\lambda,zu_i)]_{\sigma\beta_j}}{g(\lambda,zv_i)-g(\lambda,zu_i)} - \frac{\{[\dot{g}(\lambda,zv_i)]_{\sigma} - [\dot{g}(\lambda,zu_i)]_{\sigma}\}\{[\dot{g}(\lambda,zv_i)]_{\beta_j} - [\dot{g}(\lambda,zu_i)]_{\beta_j}\}}{[g(\lambda,zv_i)-g(\lambda,zu_i)]^2}\right\}$$
$$-\sum_{i\in C}\left\{\frac{[\ddot{g}(\lambda,zu_i)]_{\sigma\beta_j}}{1-g(\lambda,zu_i)} - \frac{[\dot{g}(\lambda,zu_i)]_{\sigma}[\dot{g}(\lambda,zu_i)]_{\beta_j}}{[1-g(\lambda,zu_i)]^2}\right\},$$

$$\mathbf{L}_{\beta_j\beta_s} = \sum_{i\in F}\left\{\frac{[\ddot{g}(\lambda,zv_i)]_{\beta_j\beta_s} - [\ddot{g}(\lambda,zu_i)]_{\beta_j\beta_s}}{g(\lambda,zv_i)-g(\lambda,zu_i)} - \left[\frac{[\dot{g}(\lambda,zv_i)]_{\beta_j} - [\dot{g}(\lambda,zu_i)]_{\beta_j}}{g(\lambda,zv_i)-g(\lambda,zu_i)}\right]^2\right\}$$
$$-\sum_{i\in C}\left\{\frac{[\ddot{g}(\lambda,zu_i)]_{\beta_j\beta_s}}{1-g(\lambda,zu_i)} - \left[\frac{[\dot{g}(\lambda,zu_i)]_{\beta_j}}{1-g(\lambda,zu_i)}\right]^2\right\}.$$

For $\lambda = 0$:

$$\mathbf{L}_{\lambda\lambda} = 0, \quad \mathbf{L}_{\lambda\sigma} = 0, \quad \mathbf{L}_{\lambda\beta_j} = 0,$$

$$\mathbf{L}_{\sigma\sigma} = \sum_{i\in F}\left\{\frac{[\ddot{\Phi}(zv_i)]_{\sigma\sigma} - [\ddot{\Phi}(zu_i)]_{\sigma\sigma}}{\Phi(zv_i)-\Phi(zu_i)} - \left[\frac{[\dot{\Phi}(zv_i)]_{\sigma} - [\dot{\Phi}(zu_i)]_{\sigma}}{\Phi(zv_i)-\Phi(zu_i)}\right]^2\right\}$$
$$-\sum_{i\in C}\left\{\frac{[\ddot{\Phi}(zu_i)]_{\sigma\sigma}}{1-\Phi(zu_i)} - \left[\frac{[\dot{\Phi}(zu_i)]_{\sigma}}{1-\Phi(zu_i)}\right]^2\right\},$$

$$\mathbf{L}_{\sigma\beta_j} = \sum_{i\in F}\left\{\frac{[\ddot{\Phi}(zv_i)]_{\sigma\beta_j} - [\ddot{\Phi}(zu_i)]_{\sigma\beta_j}}{\Phi(zv_i)-\Phi(zu_i)} - \frac{\{[\dot{\Phi}(zv_i)]_{\sigma} - [\dot{\Phi}(zu_i)]_{\sigma}\}\{[\dot{\Phi}(zv_i)]_{\beta_j} - [\dot{\Phi}(zu_i)]_{\beta_j}\}}{[\Phi(zv_i)-\Phi(zu_i)]^2}\right\}$$
$$- \sum_{i\in C}\left\{\frac{[\ddot{\Phi}(zu_i)]_{\sigma\beta_j}}{1-\Phi(zu_i)} - \frac{[\dot{\Phi}(zu_i)]_{\sigma}[\dot{\Phi}(zu_i)]_{\beta_j}}{[1-\Phi(zu_i)]^2}\right\},$$
$$\mathbf{L}_{\beta_j\beta_s} = \sum_{i\in F}\left\{\frac{[\ddot{\Phi}(zv_i)]_{\beta_j\beta_s} - [\ddot{\Phi}(zu_i)]_{\beta_j\beta_s}}{\Phi(zv_i)-\Phi(zu_i)} - \left[\frac{[\dot{\Phi}(zv_i)]_{\beta_j} - [\dot{\Phi}(zu_i)]_{\beta_j}}{\Phi(zv_i)-\Phi(zu_i)}\right]^2\right\}$$
$$- \sum_{i\in C}\left\{\frac{[\ddot{\Phi}(zu_i)]_{\beta_j\beta_s}}{1-\Phi(zu_i)} - \left[\frac{[\dot{\Phi}(zu_i)]_{\beta_j}}{1-\Phi(zu_i)}\right]^2\right\}.$$

Note that

$$zu_i = \frac{[\log(u_i) - \mathbf{x}_i^{\top}\boldsymbol{\beta}]}{\sigma},\quad zv_i = \frac{[\log(v_i) - \mathbf{x}_i^{\top}\boldsymbol{\beta}]}{\sigma},\quad g(\lambda, zu_i) = Q[\lambda^{-2}, hu_i],$$
$$g(\lambda, zv_i) = Q[\lambda^{-2}, hv_i],\quad hu_i = \lambda^{-2}\exp(\lambda zu_i),\quad hv_i = \lambda^{-2}\exp(\lambda zv_i),$$
$$a_{ui}(\lambda) = \int_0^{hu_i} w^{\lambda^{-2}-1}\exp(-w)\,\mathrm{d}w,\quad a_{vi}(\lambda) = \int_0^{hv_i} w^{\lambda^{-2}-1}\exp(-w)\,\mathrm{d}w,$$
$$[\dot{g}(\lambda, zu_i)]_{\lambda} = \frac{\{[\dot{a}_{ui}(\lambda)]_{\lambda} - \psi(\lambda^{-2})a_{ui}(\lambda)\}}{\Gamma(\lambda^{-2})},\quad [\dot{g}(\lambda, zv_i)]_{\lambda} = \frac{\{[\dot{a}_{vi}(\lambda)]_{\lambda} - \psi(\lambda^{-2})a_{vi}(\lambda)\}}{\Gamma(\lambda^{-2})},$$
$$[\dot{a}_{ui}(\lambda)]_{\lambda} = \exp(-hu_i)(hu_i)^{\lambda^{-2}}(zu_i - 2\lambda^{-1}) - 2\lambda^{-6}\sum_{k=0}^{\infty}\frac{(-1)^k}{k!}J(hu_i, \lambda^{-2}-1+k, 1),$$
$$[\dot{a}_{vi}(\lambda)]_{\lambda} = \exp(-hv_i)(hv_i)^{\lambda^{-2}}(zv_i - 2\lambda^{-1}) - 2\lambda^{-6}\sum_{k=0}^{\infty}\frac{(-1)^k}{k!}J(hv_i, \lambda^{-2}-1+k, 1),$$
$$J(hu_i, \lambda^{-2}-1+k, 1) = \int_0^{hu_i} w^{\lambda^{-2}-1+k}\log(w)\,\mathrm{d}w,\quad J(hv_i, \lambda^{-2}-1+k, 1) = \int_0^{hv_i} w^{\lambda^{-2}-1+k}\log(w)\,\mathrm{d}w,$$
$$[\dot{g}(\lambda, zu_i)]_{\sigma} = -\left[\frac{(\lambda\sigma^{-1}zu_i)}{\Gamma(\lambda^{-2})}\right]\exp(-hu_i)(hu_i)^{\lambda^{-2}},\quad [\dot{g}(\lambda, zv_i)]_{\sigma} = -\left[\frac{(\lambda\sigma^{-1}zv_i)}{\Gamma(\lambda^{-2})}\right]\exp(-hv_i)(hv_i)^{\lambda^{-2}},$$
$$[\dot{g}(\lambda, zu_i)]_{\beta_j} = -\left[\frac{(x_{ij}\lambda\sigma^{-1})}{\Gamma(\lambda^{-2})}\right]\exp(-hu_i)(hu_i)^{\lambda^{-2}},\quad [\dot{g}(\lambda, zv_i)]_{\beta_j} = -\left[\frac{(x_{ij}\lambda\sigma^{-1})}{\Gamma(\lambda^{-2})}\right]\exp(-hv_i)(hv_i)^{\lambda^{-2}},$$
$$[\dot{\Phi}(zu_i)]_{\sigma} = -\frac{[f(zu_i)zu_i]}{\sigma},\quad [\dot{\Phi}(zv_i)]_{\sigma} = -\frac{[f(zv_i)zv_i]}{\sigma},$$
$$[\dot{\Phi}(zu_i)]_{\beta_j} = -\frac{[x_{ij}f(zu_i)]}{\sigma},\quad [\dot{\Phi}(zv_i)]_{\beta_j} = -\frac{[x_{ij}f(zv_i)]}{\sigma},$$
$$[\ddot{g}(\lambda, zu_i)]_{\lambda\lambda} = \frac{\{[\ddot{a}_{ui}(\lambda)]_{\lambda\lambda} - \psi'(\lambda^{-2})a_{ui}(\lambda) - 2\psi(\lambda^{-2})[\dot{a}_{ui}(\lambda)]_{\lambda} + [\psi(\lambda^{-2})]^2 a_{ui}(\lambda)\}}{\Gamma(\lambda^{-2})},$$
$$[\ddot{g}(\lambda, zv_i)]_{\lambda\lambda} = \frac{\{[\ddot{a}_{vi}(\lambda)]_{\lambda\lambda} - \psi'(\lambda^{-2})a_{vi}(\lambda) - 2\psi(\lambda^{-2})[\dot{a}_{vi}(\lambda)]_{\lambda} + [\psi(\lambda^{-2})]^2 a_{vi}(\lambda)\}}{\Gamma(\lambda^{-2})},$$
$$[\ddot{a}_{ui}(\lambda)]_{\lambda\lambda} = q(\lambda, zu_i) + 12\lambda^{-7}\sum_{k=0}^{\infty}\frac{(-1)^k}{k!}J(hu_i, \lambda^{-2}-1+k, 1) - 2\lambda^{-6}\sum_{k=0}^{\infty}\frac{(-1)^k}{k!}J(hu_i, \lambda^{-2}-1+k, 2),$$
$$[\ddot{a}_{vi}(\lambda)]_{\lambda\lambda} = q(\lambda, zv_i) + 12\lambda^{-7}\sum_{k=0}^{\infty}\frac{(-1)^k}{k!}J(hv_i, \lambda^{-2}-1+k, 1) - 2\lambda^{-6}\sum_{k=0}^{\infty}\frac{(-1)^k}{k!}J(hv_i, \lambda^{-2}-1+k, 2),$$
$$q(\lambda, zu_i) = \exp(-hu_i)(hu_i)^{\lambda^{-2}}\{2\lambda^{-2} - hu_i(zu_i - 2\lambda^{-1})^2 + \lambda^{-2}(zu_i - 2\lambda^{-1})[zu_i - 2\lambda^{-1}(\log(hu_i)+1)]\},$$
$$q(\lambda, zv_i) = \exp(-hv_i)(hv_i)^{\lambda^{-2}}\{2\lambda^{-2} - hv_i(zv_i - 2\lambda^{-1})^2 + \lambda^{-2}(zv_i - 2\lambda^{-1})[zv_i - 2\lambda^{-1}(\log(hv_i)+1)]\},$$
$$J(hu_i, \lambda^{-2}-1+k, 2) = \int_0^{hu_i} w^{\lambda^{-2}-1+k}[\log(w)]^2\,\mathrm{d}w,\quad J(hv_i, \lambda^{-2}-1+k, 2) = \int_0^{hv_i} w^{\lambda^{-2}-1+k}[\log(w)]^2\,\mathrm{d}w,$$

$$[\ddot{g}(\lambda, zu_i)]_{\lambda\sigma} = \frac{\exp(-hu_i)(hu_i)^{\lambda^{-2}}}{\sigma\Gamma(\lambda^{-2})}\{zu_i[1-\lambda hu_i(2\lambda^{-1}-zu_i)-\lambda^{-1}[zu_i-2\lambda^{-1}(\log(hu_i)+1)]] + \lambda zu_i\psi(\lambda^{-2})\},$$

$$[\ddot{g}(\lambda, zv_i)]_{\lambda\sigma} = \frac{\exp(-hv_i)(hv_i)^{\lambda^{-2}}}{\sigma\Gamma(\lambda^{-2})}\{zv_i[1-\lambda hv_i(2\lambda^{-1}-zv_i)-\lambda^{-1}[zv_i-2\lambda^{-1}(\log(hv_i)+1)]] + \lambda zv_i\psi(\lambda^{-2})\},$$

$$[\ddot{g}(\lambda, zu_i)]_{\lambda\beta_j} = -\frac{x_{ij}\exp(-hu_i)(hu_i)^{\lambda^{-2}}}{\sigma\Gamma(\lambda^{-2})}\{1+\lambda^{-1}[hu_i(2\lambda^{-1}-zu_i)+zu_i-2\lambda^{-1}(\log(hu_i)+1)]-\lambda\psi(\lambda^{-2})\},$$

$$[\ddot{g}(\lambda, zv_i)]_{\lambda\beta_j} = -\frac{x_{ij}\exp(-hv_i)(hv_i)^{\lambda^{-2}}}{\sigma\Gamma(\lambda^{-2})}\{1+\lambda^{-1}[hv_i(2\lambda^{-1}-zv_i)+zv_i-2\lambda^{-1}(\log(hv_i)+1)]-\lambda\psi(\lambda^{-2})\},$$

$$[\ddot{g}(\lambda, zu_i)]_{\sigma\sigma} = \left[\frac{(\sigma^{-2}zu_i)}{\Gamma(\lambda^2)}\right]\exp(-hu_i)(hu_i)^{\lambda^{-2}}[zu_i(1-\lambda hu_i)+2\lambda],$$

$$[\ddot{g}(\lambda, zv_i)]_{\sigma\sigma} = \left[\frac{(\sigma^{-2}zv_i)}{\Gamma(\lambda^2)}\right]\exp(-hv_i)(hv_i)^{\lambda^{-2}}[zv_i(1-\lambda hv_i)+2\lambda],$$

$$[\ddot{g}(\lambda, zu_i)]_{\sigma\beta_j} = \left[\frac{(x_{ij}\sigma^{-2}\lambda}{\Gamma(\lambda^2)}\right]\exp(-hu_i)(hu_i)^{\lambda^{-2}}[1+zu_i\lambda(hu_i-\lambda^{-2})],$$

$$[\ddot{g}(\lambda, zv_i)]_{\sigma\beta_j} = \left[\frac{(x_{ij}\sigma^{-2}\lambda}{\Gamma(\lambda^2)}\right]\exp(-hv_i)(hv_i)^{\lambda^{-2}}[1+zv_i\lambda(hv_i-\lambda^{-2})],$$

$$[\ddot{g}(\lambda, zu_i)]_{\beta_j\beta_s} = \left[\frac{(x_{ij}x_{is}\sigma^{-2}\lambda)}{\Gamma(\lambda^2)}\right]\exp(-hu_i)(hu_i)^{\lambda^{-2}}[\lambda(\lambda^{-2}-hu_i)],$$

$$[\ddot{g}(\lambda, zv_i)]_{\beta_j\beta_s} = \left[\frac{(x_{ij}x_{is}\sigma^{-2}\lambda)}{\Gamma(\lambda^2)}\right]\exp(-hv_i)(hv_i)^{\lambda^{-2}}[\lambda(\lambda^{-2}-hv_i)],$$

$$[\ddot{\Phi}(zu_i)]_{\sigma\sigma} = \left(\frac{zu_i}{\sigma^2}\right)[f'(zu_i)zu_i+2f(zu_i)], \quad [\ddot{\Phi}(zv_i)]_{\sigma\sigma} = \left(\frac{zv_i}{\sigma^2}\right)[f'(zv_i)zv_i+2f(zv_i)],$$

$$[\ddot{\Phi}(zu_i)]_{\sigma\beta_j} = \left(\frac{x_{ij}}{\sigma^2}\right)[f'(zu_i)zu_i+f(zu_i)], \quad [\ddot{\Phi}(zv_i)]_{\sigma\beta_j} = \left(\frac{x_{ij}}{\sigma^2}\right)[f'(zv_i)zv_i+f(zv_i)],$$

$$[\ddot{\Phi}(zu_i)]_{\beta_j\beta_s} = \left[\frac{(x_{ij}x_{is})}{\sigma^2}\right]f'(zu_i), \quad [\ddot{\Phi}(zv_i)]_{\beta_j\beta_s} = \left[\frac{(x_{ij}x_{is})}{\sigma^2}\right]f'(zv_i).$$