SORT 32 (2) July-December 2008, 115-140

Assessing influence in survival data with a cure fraction and covariates

Edwin M. M. Ortega[1], Vicente G. Cancho[2] and Victor Hugo Lachos[3]

[1]*ESALQ-USP,* [2]*ICMC-USP* *and* [3]*IMECC-UNICAMP*

Abstract

Diagnostic methods have been an important tool in regression analysis to detect anomalies, such as departures from error assumptions and the presence of outliers and influential observations with the fitted models. Assuming censored data, we considered a classical analysis and Bayesian analysis assuming no informative priors for the parameters of the model with a cure fraction. A Bayesian approach was considered by using Markov Chain Monte Carlo Methods with Metropolis-Hasting algorithms steps to obtain the posterior summaries of interest. Some influence methods, such as the local influence, total local influence of an individual, local influence on predictions and generalized leverage were derived, analyzed and discussed in survival data with a cure fraction and covariates. The relevance of the approach was illustrated with a real data set, where it is shown that, by removing the most influential observations, the decision about which model best fits the data is changed.

MSC: 62F15, 62H10, 62J20, 62N01, 62N02

Keywords: Cure fraction, Bayesian inference, local influence, generalized leverage, survival data.

1 Introduction

Models for survival analysis typically assume that every subject in a population is susceptible to the event under study and will eventually experience it if follow-up is sufficiently long. However, there are situations where a fraction of individuals are

[1] ESALQ, University of São Paulo. Departamento de Ciências Exatas, USP Av. Pádua Dias 11 - Caixa Postal 9, 13418-900 Piracicaba - São Paulo - Brazil. e-mail: edwin@esalq.usp.br

[2] ICMC, University of São Paulo, São Carlos, São Paulo, Brazil. E-mail: garibay@icmc.usp.br

[3] IMECC, University of Campinas, Campinas, Brazil. E-mail: vlachos@ime.unicamp.br

Received: March 2007
Accepted: January 2008

not expected to experience the event of interest, that is, those individuals are cured or insusceptible. 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 utilized to estimate the cured fraction.

Cure rate models are survival models which allow for a fraction of cured 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.

Cure rate models have been used for modeling time-to-event data for various types of cancers, including breast cancer, non-Hodgkins lymphoma, leukemia, prostate cancer and melanoma. Perhaps the most popular type of cure rate models is the mixture model introduced by Berkson and Gage (1958). In this model, the population is divided into two subpopulations 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,

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

A common choice of the $S(t)$ in (1) is exponential and the Weibull distribution. With those choices, we have respectively an exponential model with a cured fraction and a Weibull model with a cured fraction. This mixture model has been studied by several authors, including Farrell (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. We considered a classical analysis for model Weibull with a cured fraction and covariates. The inferential part was carried out using the asymptotic distribution of the maximum likelihood estimators, which in situations when the sample is small, may present difficult results to be justified. As an alternative for classical analysis, we explored the use of techniques of the Markov Chain Monte Carlo (MCMC) method to develop a Bayesian inference for the Weibull model with a cure fraction.

The development of influence diagnostics is an important step in the analysis of a data set as it provides us with an indication of bad model fitting or of influential observations. However, there are no applications of influence diagnostics to survival data with a cured fraction and covariates. 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., 2000 and Dias, et al., 2003). Also, some authors have investigated the assessment of local influence in survival analysis models:

for instance, Pettit and Bin Daud (1989) investigate local influence in proportional hazard regression models; Escobar and Meeker (1992) adapt local influence methods to regression analysis with censoring, Ortega et al. (2003) considered the problem of assessing local influence in generalized log-gamma regression models with censored observations, Ortega et al. (2006) derived curvature calculations under various perturbation schemes in exponentiated-Weibull regression models with censored data and Fachini et al. (2007) adapt local influence methods to polyhazard models under the presence of covariates.

In this article, we present diagnostic methods based on local influence and residual analysis in survival data with a cure fraction and covariates, where the covariates are modeled through p via a binomial regression model. In section 2, we present the Weibull model with a cured fraction and covariates and discuss the process of estimation for the parameters in the model. Section 3 deals with a Bayesian analysis using MCMC methodology under informative priors. In Section 4, 5 and 6, we discuss the local influence method, local influence on predictions and generalized leverage. Likelihood displacement is used to evaluate the influence of observations on the maximum likelihood estimators. Section 7 presents the results of an analysis with a real data set and residual analysis.

2 The Weibull model with a cure fraction and covariates

Let a binary random variable Y_i, $i = 1, \ldots, n$ indicate that the ith individual in a population is at risk or not with respect to a certain type of failure, that is, $Y_i = 1$ indicates that the ith individual will eventually experience a failure event (uncured) and $Y_i = 0$ 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 a logistic link of $\mathbf{x}$ such that the conditional distribution of Y is given by

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

where $\boldsymbol{\beta} = (\beta_1, \beta_2, \ldots, \beta_p)^T$ is a vector p-dimensional parameter. Note that the cure probability varies from individual to individual so that the probability that individual i is cured is modeled by $p_i = \frac{\exp(\mathbf{x}_i^T\boldsymbol{\beta})}{1 + \exp(\mathbf{x}_i^T\boldsymbol{\beta})}$. The logistic link keeps each p_i strictly between 0 and 1.

Letting T_i be the ith time of occurrence of the failure event and considering that T_i´s are independent and identically distributed with the Weibull distribution, the density function is given by

$$f(t; \alpha, \lambda \,|\, Y_i = 1) = \alpha t^{\alpha-1} \exp\{\lambda - t^{\alpha} e^{\lambda}\}, \tag{2}$$

where $\alpha > 0$ is a shape parameter and $\lambda \in R$ is a scale parameter. Thus, the contribution of an individual that failed at t_i to the likelihood function is given by $(1 - p_i)\alpha t_i^{\alpha-1}\exp\{\lambda - t_i^{\alpha} e^{\lambda}\}$, and the contribution of an individual that is at risk at time t_i is $p_i + (1 - p_i)\exp\{-t_i^{\alpha} e^{\lambda}\}$.

Given a sample $t_1, \ldots, t_n$, where we observed $t_i = \min(T_i, C_i)$ where T_i is the lifetime for the ith individual and C_i is the censoring time for the ith individual. In this case the log-likelihood function corresponding to the parameter vector $\boldsymbol{\theta} = (\alpha, \lambda, \boldsymbol{\beta}^T)^T$ is given by

$$\begin{aligned} l(\boldsymbol{\theta}) \quad \propto \quad & \mathrm{r}\log(\alpha) + \mathrm{r}\lambda + \sum_{i \in F} \log(1 - p_i) + (\alpha - 1) \sum_{i \in F} \log(t_i) - \\ & \exp\{\lambda\} \sum_{i \in F} t_i^{\alpha} + \sum_{i \in C} \log\Big[p_i + (1 - p_i)\exp\{-t_i^{\alpha} \mathrm{e}^{\lambda}\}\Big], \end{aligned} \tag{3}$$

where r is the number of uncensored observations (failures), F denotes the set of uncensored observations, C denotes the set of censored observations. Maximum likelihood estimates for parameter vector $\boldsymbol{\theta}$ can be obtained by maximizing the likelihood function, while Bayesian estimation is discussed. In this paper, software Ox (MAXBFGS subroutine) (see Doornik, 1996) was used to compute maximum likelihood estimates (MLE). Covariance estimates for maximum likelihood estimators $\widehat{\boldsymbol{\theta}}$ can also be obtained by 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 as long as regularity conditions are satisfied. More specifically, the asymptotic covariance matrix is given by $\mathbf{I}^{-1}(\boldsymbol{\theta})$ with $\mathbf{I}(\boldsymbol{\theta}) = -E[\ddot{\mathbf{L}}(\boldsymbol{\theta})]$ such that $\ddot{\mathbf{L}}(\boldsymbol{\theta}) = \left\{\frac{\partial^2 l(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}\,\partial\boldsymbol{\theta}^T}\right\}$.

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), the matrix of second derivatives of the log likelihood, $-\ddot{\mathbf{L}}(\boldsymbol{\theta})$, evaluated at MLE $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$, which is consistent, can be used. Then

$$\ddot{\mathbf{L}}(\boldsymbol{\theta}) = \begin{pmatrix} \mathbf{L}_{\alpha\alpha} & \mathbf{L}_{\alpha\lambda} & \mathbf{L}_{\alpha\beta} \\ . & \mathbf{L}_{\lambda\lambda} & \mathbf{L}_{\lambda\beta} \\ . & . & \mathbf{L}_{\beta\beta} \end{pmatrix}$$

with the submatrices in appendix A.

3 A Bayesian analysis using MCMC

In this section, we consider a Bayesian approach to the MCMC methodology for approximating the posterior distribution for quantities of interest in survival data with a cure fraction and covariates. As seen in the previous section, likelihood based inference in small samples can be somewhat misleading. Thus, Bayesian inference may play an important role in such cases. Since the derivation of exact posterior densities is not feasible for the Weibull model with a cure fraction and covariates, we make use of the MCMC methodology to obtain approximation for such densities. We consider the joint prior density for $\boldsymbol{\theta} = (\alpha, \lambda, \boldsymbol{\beta}^T)^T$ of the form

$$\pi(\boldsymbol{\theta}) = \prod_{i=1}^{p} \left(\phi(\beta_i \,|\, \mu_{\beta_i}, \sigma^2_{\beta_i}))\right) \phi(\lambda \,|\, \mu_\lambda, \sigma_\lambda) \Gamma(\alpha \,|\, a, b), \tag{4}$$

where $\phi(\cdot \,|\, \mu, \sigma^2)$ denotes the probability density function of the Normal distribution with mean μ and variance σ^2 and $\Gamma(\cdot \,|\, a, b)$ denoting the Gamma distribution with shape parameter $a > 0$ and scale $b > 0$. Here all the hyperparameters are specified.

Combining likelihood function $L(\boldsymbol{\theta}) \propto \exp\{l(\boldsymbol{\theta})\}$ and prior to specification (4), the joint posterior distribution for $\boldsymbol{\theta}$ is given by

$$\pi(\boldsymbol{\theta} \,|\, D) \propto T^{\alpha-1} \alpha^{r+a-1} \exp\left\{ -b\alpha - \frac{\lambda^2}{2\sigma^2_\lambda} + r\lambda - \frac{1}{2} \sum_{j=1}^{p} \frac{\beta_j^2}{\sigma^2_{\beta_j}} - \right. \tag{5}$$

$$\left. e^{\lambda} \sum_{i \in F} t_i^{\alpha} + \sum_{i \in F} \log(1 - p_i) + \sum_{i \in C} \log\Big[p_i + (1 - p_i) \exp\{-t_i^{\alpha} e^{\lambda}\} \Big] \right\},$$

where r is the number of uncensored observations, $T = \prod_{i \in F} t_i$, $i = 1, 2, \ldots, n$ and D denotes the observed data.

To implement the MCMC methodology, we consider Gibbs within the Metropolis-Hasting sampler, which requires the derivation of the complete set of conditional posterior distributions. After some algebraic manipulations, it follows that the conditional posterior densities are given by

$$\pi(\alpha \,|\, \boldsymbol{\beta}, \lambda, D) \propto T^{\alpha-1} \alpha^{r+a-1} \exp\left\{ -b\alpha - e^{\lambda} \sum_{i \in F} t_i^{\alpha} + \sum_{i \in C} \log\Big[p_i + (1 - p_i) \exp\{-t_i^{\alpha} e^{\lambda}\} \Big] \right\}$$

$$\pi(\lambda \,|\, \alpha, \boldsymbol{\beta}, D) \propto \exp\left\{ -\frac{(\lambda - \mu_\lambda)^2}{2\sigma^2_\lambda} + r\lambda - e^{\lambda} \sum_{i \in F} t_i^{\alpha} + \sum_{i \in C} \log\Big[p_i + (1 - p_i) \exp\{-t_i^{\alpha} e^{\lambda}\} \Big] \right\}$$

$$\pi(\boldsymbol{\beta} \,|\, \alpha, \lambda, D) \propto \left\{ -\frac{1}{2} \sum_{j=1}^{p} \frac{(\beta_j - \mu_{\beta_j})^2}{\sigma^2_{\beta_j}} + \sum_{i \in F} \log(1 - p_i) + \right.$$

$$\left. \sum_{i \in C} \log\Big[p_i + (1 - p_i) \exp\{-t_i^{\alpha} e^{\lambda}\} \Big] \right\}$$

Since the conditional posteriors do not present standard forms, the use of the Metropolis-Hasting sampler is required.

4 Influence diagnostics

4.1 Local influence

Let $l(\boldsymbol{\theta})$ denote the log-likelihood function from the postulated model, where $\boldsymbol{\theta} = (\alpha, \lambda, \boldsymbol{\beta}^T)^T$, and let $\boldsymbol{\omega}$ be an $n \times 1$ vector of perturbations restricted to some open subset $\boldsymbol{\Omega} \subset \mathbb{R}^n$. The perturbations are made in the log-likelihood function. We will assume, in particular, the case-weights perturbation scheme such that the log-likelihood function takes the form

$$l(\boldsymbol{\theta} \,|\, \boldsymbol{\omega}) = \sum_{i \in F} \omega_i \log\left[(1 - p_i)\alpha t_i^{\alpha-1} \exp\{\lambda - t_i^{\alpha} e^{\lambda}\}\right] + \sum_{i \in C} \omega_i \log\left[p_i + (1 - p_i)\exp\{-t_i^{\alpha} e^{\lambda}\}\right],$$

where $0 \leq \omega_i \leq 1$ and $\boldsymbol{\omega}_0 = (1, 1, \ldots, 1)^T$ is the vector of no perturbation. Note that $l(\boldsymbol{\theta} \,|\, \boldsymbol{\omega}_0) = l(\boldsymbol{\theta})$. To assess the influence of the perturbations in the maximum likelihood estimate $\hat{\boldsymbol{\theta}}$, we consider the likelihood displacement

$$LD(\boldsymbol{\omega}) = 2\{l(\hat{\boldsymbol{\theta}}) - l(\hat{\boldsymbol{\theta}}_{\omega})\},$$

where $\hat{\boldsymbol{\theta}}_{\omega}$ denotes the maximum likelihood estimate under the model $l(\boldsymbol{\theta} \,|\, \boldsymbol{\omega})$.

The idea of local influence (Cook, 1986) is concerned with characterizing the behavior of $LD(\boldsymbol{\omega})$ around $\boldsymbol{\omega}_0$. The procedure consists in selecting a unit direction $\mathbf{d}$, $\| \mathbf{d} \| = 1$, and then considering the plot of $LD(\boldsymbol{\omega}_0 + a\mathbf{d})$ against a, where $a \in \boldsymbol{R}$. This plot is called lifted line. Note that, since $LD(\boldsymbol{\omega}_0) = 0$, $LD(\boldsymbol{\omega}_0 + a\mathbf{d})$ has a local minimum at $a = 0$. Each lifted line can be characterized by considering the normal curvature $C_{\mathbf{d}}(\boldsymbol{\theta})$ around $a = 0$. This curvature is interpreted as the inverse radius of the best fitting circle at $a = 0$. The suggestion is to consider direction $\mathbf{d}_{max}$ corresponding to the largest curvature $C_{\mathbf{d}_{max}}(\boldsymbol{\theta})$. The index plot of $\mathbf{d}_{max}$ may reveal those observations which, under small perturbations, exercise notable influence on $LD(\boldsymbol{\omega})$. Cook(1986) showed that the normal curvature at direction $\mathbf{d}$ takes the form $C_{\mathbf{d}}(\boldsymbol{\theta}) = 2\,|\,\mathbf{d}^T \boldsymbol{\Delta}^T (\ddot{\mathbf{L}})^{-1} \boldsymbol{\Delta}\mathbf{d}\,|$ where $-\ddot{\mathbf{L}}$ is the observed Fisher information matrix for the postulated model ($\boldsymbol{\omega} = \boldsymbol{\omega}_0$) and $\boldsymbol{\Delta}$ is the $(p+1) \times n$ matrix with elements $\Delta_{ji} = \partial^2 L(\boldsymbol{\theta} \,|\, \boldsymbol{\omega}) / \partial\theta_j \, \partial\omega_i$, evaluated at $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$ and $\boldsymbol{\omega} = \boldsymbol{\omega}_0$, $j = 1, \ldots, p + 2$ and $i = 1, \ldots, n$. Then, $C_{\mathbf{d}_{max}}$ is the largest eigenvalue of the matrix $\mathbf{B} = \boldsymbol{\Delta}^T (\ddot{\mathbf{L}})^{-1} \boldsymbol{\Delta}$, and $\mathbf{d}_{max}$ is the corresponding eigenvector. The index plot of $\mathbf{d}_{max}$ for the matrix $\boldsymbol{\Delta}^T (\ddot{\mathbf{L}})^{-1} \boldsymbol{\Delta}$ can show how to perturb the log-likelihood function to obtain larger changes in the estimate of $\boldsymbol{\theta}$.

Another procedure is the total local curvature corresponding to the *ith* element, which follows by taking $\mathbf{d}_i$ or an $n \times 1$ vector of zeros with one at the *ith* position. Thus, the curvature at the direction $\mathbf{d}_i$ assumes the form

$$C_i = 2\,|\,\boldsymbol{\Delta}_i^T \ddot{\mathbf{L}}(\boldsymbol{\theta})^{-1} \boldsymbol{\Delta}_i\,| \qquad (6)$$

where $\mathbf{\Delta}_i^T$ denotes the *ith* row of $\mathbf{\Delta}$. This is named total local influence (see, for example, Lesaffre and Verbeke, 1998). It is suggested looking at the index plot of C_i.

We find, after some algebraic manipulation, the following expressions for the weighted log-likelihood function and for the elements of the matrix $\mathbf{\Delta}$:

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

$$\begin{aligned} l(\boldsymbol{\theta}\,|\,\boldsymbol{\omega}) \;=\; & \Big[\log(\alpha) + \lambda\Big] \sum_{i\,\epsilon\,F} \omega_i + \sum_{i\,\epsilon\,F} \omega_i \log(1 - p_i) + (\alpha - 1) \sum_{i\,\epsilon\,F} \omega_i \log(t_i) \\ & - \exp\{\lambda\} \sum_{i\,\epsilon\,F} \omega_i t_i^{\alpha} + \sum_{i\,\epsilon\,C} \omega_i \log\Big[p_i + (1 - p_i)\exp\{-t_i^{\alpha} e^{\lambda}\}\Big] \end{aligned} \tag{7}$$

Let us denote $\mathbf{\Delta} = (\mathbf{\Delta}_1, \ldots, \mathbf{\Delta}_{p+2})^T$.

Then the elements of vector $\mathbf{\Delta}$ is given in appendix B.

However, if the interest is only in vector $\boldsymbol{\beta}$, the normal curvature in direction $\mathbf{d}$ is given by $C_{\mathbf{d}}(\boldsymbol{\beta}) = 2\,|\,\mathbf{d}^T\mathbf{\Delta}^T(\ddot{\mathbf{L}}^{-1} - \mathbf{B}_{22})\mathbf{\Delta}\mathbf{d}\,|$ (see Cook, 1986), where

$$\mathbf{B}_{22} = \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \ddot{\mathbf{L}}_{22}^{-1} \end{pmatrix}$$

with $\ddot{\mathbf{L}}_{22}$ denoting the submatrix of $\ddot{\mathbf{L}}$ obtained according to partition

$$\ddot{\mathbf{L}}(\boldsymbol{\theta}) = \begin{pmatrix} \mathbf{L}_{11} & \mathbf{L}_{12} \\ \mathbf{L}_{21} & \mathbf{L}_{22} \end{pmatrix}$$

The index plot of the largest eigenvector of $\mathbf{\Delta}^T(\ddot{\mathbf{L}}^{-1} - \mathbf{B}_{22})\mathbf{\Delta}$ can reveal those observations to be most influential on $\hat{\boldsymbol{\beta}}$.

4.2 Local influence on predictions

Let $\mathbf{z}$ be a $p \times 1$ vector of values of the explanatory variables, for which we do not necessarily have an observed response. Then, the prediction at $\mathbf{z}$ is $\hat{\mu}(\mathbf{z}) = \sum_{j=1}^{p} z_j\hat{\beta}_j$. Analogously, the point prediction at $\mathbf{z}$ based on the perturbed model becomes $\hat{\mu}(\mathbf{z}, \boldsymbol{\omega}) = \sum_{j=1}^{p} z_j\hat{\beta}_{j\omega}$, where $\hat{\boldsymbol{\beta}}_{\omega} = (\hat{\beta}_{1\omega}, \ldots, \hat{\beta}_{p\omega})^T$ denotes the maximum likelihood estimate from the perturbed model. Thomas and Cook (1990) have investigated the effect of small perturbations in predictions at some particular point $\mathbf{z}$ in continuous generalized linear models and by assuming ϕ known or estimated separately from $\hat{\beta}$. ϕ^{-1} is defined as a dispersion parameter. For more details, see McCullagh and Nelder (1989). They defined three objective functions based on different residuals. Because the diagnostic calculations were identical for the proposed functions, they concentrated the application of the methodology on the objective function $f(\mathbf{z}, \boldsymbol{\omega}) = \{\hat{\mu}(\mathbf{z}) - \hat{\mu}(\mathbf{z}, \boldsymbol{\omega})\}^2$.

Similarly, we will concentrate our study on investigating the normal curvature of the surface formed by vector $\boldsymbol{\omega}$ and function $f(\mathbf{z},\boldsymbol{\omega})$, around $\boldsymbol{\omega}_0$. The normal curvature at unit direction $\mathbf{d}$ takes, in this case, the form $C_{\mathbf{d}}(\mathbf{z}) = 2 \mid \mathbf{d}^T\ddot{\mathbf{f}}\mathbf{d} \mid$, where $\ddot{\mathbf{f}} = \partial^2 f/\partial\boldsymbol{\omega}\,\partial\boldsymbol{\omega}^T$ is evaluated at $\boldsymbol{\omega}_0$ and $\hat{\boldsymbol{\beta}}$. From Thomas and Cook (1990) one has that

$$\ddot{\mathbf{f}} = \boldsymbol{\Delta}^T(\ddot{\mathbf{L}}_{\beta\beta}^{-1}\mathbf{z}\mathbf{z}^T\ddot{\mathbf{L}}_{\beta\beta}^{-1})\boldsymbol{\Delta},$$

where $\boldsymbol{\Delta} = \partial^2 l(\boldsymbol{\theta} \mid \boldsymbol{\omega})/\partial\boldsymbol{\beta}\,\partial\boldsymbol{\omega}^T$. Consequently

$$\mathbf{d}_{max}(\mathbf{z}) \propto -\boldsymbol{\Delta}^T\ddot{\mathbf{L}}_{\beta\beta}^{-1}\mathbf{z}.$$

In the sequence we discuss the calculation of $\mathbf{d}_{max}(\mathbf{z})$ under additive perturbations for the response and for each continuous explanatory variable.

4.2.1 Response perturbation

Consider regression model (3) by assuming now that each t_i is perturbed as $t_i \rightarrow t_i + (S_t)\omega_i = t_i^*$, $i = 1,\ldots,n$, where (S_t) is a scale factor that can be the estimated standard deviation of T and $w_i \epsilon \mathbb{R}$. Below we give the expressions for the log-likelihood function

Here the perturbed log-likelihood function is expressed as

$$\begin{aligned} l(\boldsymbol{\theta}|\boldsymbol{\omega}) &= r\log(\alpha) + r\lambda + \sum_{i\,\epsilon\,F}\log(1-p_i) + (\alpha-1)\sum_{i\,\epsilon\,F}\log(t_i^*) - \\ &\quad \exp\{\lambda\}\sum_{i\,\epsilon\,F} t_i^{*\alpha} + \sum_{i\,\epsilon\,C}\log\left[p_i + (1-p_i)\exp\{-t_i^{*\alpha}e^{\lambda}\}\right] \end{aligned} \quad (8)$$

where $t_i^* = t_i + (S_t)\omega_i$.

Matrix $\boldsymbol{\Delta} = (\Delta_1, \Delta_2, \ldots, \Delta_{p+2})^T$ is given in appendix C.

Vector $\mathbf{d}_{max}(\mathbf{z})$ is constructed by taking $\mathbf{z} = \mathbf{x}_i$, which corresponds to the $n \times 1$ vector

$$\mathbf{d}_{max}(\mathbf{x}_i) \propto -\boldsymbol{\Delta}^T\ddot{\mathbf{L}}_{\beta\beta}^{-1}\mathbf{x}_i. \quad (9)$$

A large value for the ith component of (15), $\mathbf{d}_{max_i}(\mathbf{x}_i)$, indicates that the ith observation should have substantial local influence on $\hat{y}_i$. Then, the suggestion is to take the index plot of the $n \times 1$ vector $(\mathbf{d}_{max_1}(\mathbf{x}_1), \ldots, \mathbf{d}_{max_n}(\mathbf{x}_n))^T$ in order to identify those observations with high influence on its own fitted value.

4.2.2 Explanatory variable perturbation

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

The perturbed log-likelihood function is, in this case, expressed as

$$\begin{aligned} l(\boldsymbol{\theta}\,|\,\boldsymbol{\omega}) \;&=\; r\log(\alpha) + r\lambda + \sum_{i\,\epsilon\,F}\log(1-p_i^*) + (\alpha-1)\sum_{i\,\epsilon\,F}\log(t_i) - \\ &\exp\{\lambda\}\sum_{i\,\epsilon\,F} t_i^{\alpha} + \sum_{i\,\epsilon\,C}\log\left[p_i^* + (1-p_i^*)\exp\{-t_i^{\alpha}e^{\lambda}\}\right] \end{aligned} \quad (10)$$

where $p_i^* = \dfrac{\exp\{\mathbf{x}_i^{*T}\boldsymbol{\beta}\}}{1+\exp\{\mathbf{x}_i^{*T}\boldsymbol{\beta}\}}$ and $\mathbf{x}_i^{*T}\boldsymbol{\beta} = \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_t(x_{it} + \omega_i S_x) + \cdots + \beta_p x_{ip}$.

Matrix $\boldsymbol{\Delta} = (\Delta_1, \Delta_2, \ldots, \Delta_{p+2})^T$ is given in appendix D.

Similarly to the response perturbation case, the suggestion here is to evaluate the largest curvature at $\mathbf{z} = \mathbf{x}_i$, which leads to

$$C_{max}(\mathbf{x}_i) = 2\,|\,\mathbf{d}_{max}^T\,\ddot{\mathbf{f}}\mathbf{d}_{max}\,|,$$

and consequently

$$\mathbf{d}_{max}(\mathbf{x}_i) \propto -\boldsymbol{\Delta}^T\ddot{\mathbf{L}}_{\beta\beta}^{-1}\mathbf{x}_i.$$

To see for which observed values of X_t the prediction is most sensitive under small changes in X_t, we can perform the plot of $C_{max}(\mathbf{x}_i)$ against x_{it}. The index plot of the $n \times 1$ vector $(\ell_{max_1}(\mathbf{x}_1), \ldots, \ell_{max_n}(\mathbf{x}_n))^T$ can indicate those observations for which a small perturbation in the value of X_t leads to a substantial change in the prediction.

4.3 Generalized leverage

Let $l(\boldsymbol{\theta})$ denote the log-likelihood function from the postulated model in equation (10), $\widehat{\boldsymbol{\theta}}$ the MLE of $\boldsymbol{\theta}$ and $\boldsymbol{\mu}$ the expectation of $\mathbf{T}$, then, $\widehat{\mathbf{t}} = \boldsymbol{\mu}(\widehat{\boldsymbol{\theta}})$ will be the predicted response vector.

The main idea behind the concept of leverage (see, for instance, Cook and Weisberg, 1982; Wei et al., 1998) is that of evaluating the influence of t_i on its own predicted value. This influence may well be represented by derivative $\frac{\partial\widehat{t_i}}{\partial ti}$ that equals h_{ii} is the i-th principal diagonal element of the projection matrix $\mathbf{H} = \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T$ and $\mathbf{X}$ is the model matrix. Extensions to more general regression models have been given, for instance, by St. Laurent and Cook (1992), and Wei, et al. (1998) and Paula (1999), when $\boldsymbol{\theta}$ is restricted with inequalities. Hence, it follows from Wei et al.(1998) that the nxn matrix $(\frac{\partial\widehat{\mathbf{t}}}{\partial\mathbf{t}})$ of generalized leverage can be expressed as:

$$\mathbf{GL}(\widehat{\boldsymbol{\theta}}) = \left\{\mathbf{D}_{\boldsymbol{\theta}}\left[\text{-}\ddot{\mathbf{L}}(\boldsymbol{\theta})\right]^{-1}\ddot{\mathbf{L}}_{\boldsymbol{\theta}\mathbf{t}}\right\} \quad (11)$$

evaluated at $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$.

Matrix $\mathbf{D}_{\boldsymbol{\theta}} = \left(\mathbf{D}_{\alpha}, \mathbf{D}_{\lambda}, \mathbf{D}_{\boldsymbol{\beta}}\right)$ is given in appendix E.

5 Residual analysis

In order to study departures from the error assumption as the well as the presence of outliers, we will first consider the martingale residual proposed by Barlow and Prentice (1988) (see also Therneau et al., 1990). This residual was introduced in counting processes and can be written for the Weibull model with a cure fraction and covariates as

$$r_{M_i} = \begin{cases} 1 + \log\left[\hat{p}_i + (1 - \hat{p}_i)\exp\{-t_i^{\hat{\alpha}} e^{\hat{\lambda}}\}\right], & \text{if} \quad i \in F; \\ \log\left[\hat{p}_i + (1 - \hat{p}_i)\exp\{-t_i^{\hat{\alpha}} e^{\hat{\lambda}}\}\right], & \text{if} \quad i \in C. \end{cases} \tag{12}$$

Due to the skewness distributional form of r_{M_i}, it has maximum value +1 and minimum value $-\infty$, and transformations to achieve a more normal shaped form would be more appropriate for residual analysis. Another possibility is to use the deviance residual (see, for instance, McCullagh and Nelder, 1989, Section 2.4), which has been largely applied in generalized linear models (GLMs). Various authors have investigated the use of deviance residuals in GLMs (see, for instance, Williams, 1987; Hinkley et al., 1991; Paula 1995; Ortega et al., 2007) as well as in other regression models (see, for example, Farhrmeir and Tutz, 1994). In the Weibull model with a cure fraction and covariates, the modified residual deviance is expressed here as

$$r_{D_i} = \begin{cases} \operatorname{sgn}(r_{M_i})\Big[-2 - 2\log\Big\{\left[\hat{p}_i + (1 - \hat{p}_i)\exp\{-t_i^{\hat{\alpha}} e^{\hat{\lambda}}\}\right]\times \\ \log\left[\hat{p}_i + (1 - \hat{p}_i)\exp\{-t_i^{\hat{\alpha}} e^{\hat{\lambda}}\}\right]^{-1}\Big\}\Big]^{1/2}, & \text{if} \quad i \in F; \\ \operatorname{sgn}(r_{M_i})\Big\{-2\log\left[\hat{p}_i + (1 - \hat{p}_i)\exp\{-t_i^{\hat{\alpha}} e^{\hat{\lambda}}\}\right]\Big\}^{1/2}, & \text{if} \quad i \in C, \end{cases} \tag{13}$$

where r_{M_i} is the residual martingale corresponding to the Weibull model with a cure fraction and covariates.

6 Application

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

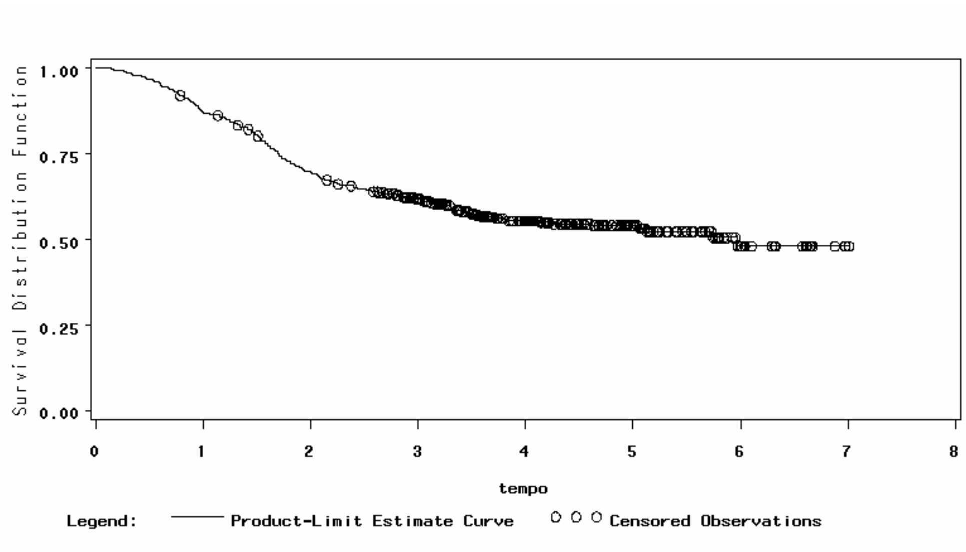

Figure 1: *Plot of the Survivor Function.*

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);
- δ_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}: Breslow (tumor thickness in mm).

The survival function graph, Kaplan-Meier estimate, is presented in Figure 1, from where a fraction of survivors can be observed.

6.1 Maximum likelihood results

To obtain the maximum likelihood estimates for the parameters in the Weibull model we use the subroutine MAXBFGS in Ox, whose results are given in the Table following.

The mean cure fraction estimated was $\hat{p} = 0.5162$.

In Table 1, it is estimated that the only significant variable is x3(nodule). Also, the information criteria based on the decision theory which penalize models with a large

Table 1: *Maximum likelihood estimates for the complete data set of the Weibull model with a cure fraction and covariates*

Parameter	Estimate	SE	p-value
α	1.6104	0.1066	—
λ	−1.2877	0.1217	—
β_0	2.2656	0.5811	< 0.0001
β_1	−0.1603	0.2247	0.4756
β_2	−0.0142	0.0086	0.0977
β_3	−0.5392	0.1142	< 0.0001
β_4	0.2019	0.2315	0.3832
β_5	−0.1509	0.3352	0.6527
β_6	−0.0599	0.0391	0.1253
Statistics	Value	Statistics	Value
AIC	1045.578	BIC	1081.876

number of parameters were used. The used criteria are based on the AIC statistics (Akaike Information Criterion) and BIC (Bayesian Information Criterion) (see Table 1).

6.2 Bayesian analysis

We consider now a Bayesian analysis for the data considering the following independent prior (4) with values of the hyperparameters given for $a = b = 0,1$, $\mu_\lambda = \mu_{\beta_j} = 0$ and $\sigma^2_\lambda = \sigma^2_{\beta_j} = 100$, $j = 0, 1, \ldots, 6$. Considering those prior densities we generated two parallel independent runs of the Gibbs sampler chain with size 40,000 for each parameter, discarding the first 5,000 iterations. To eliminate the effect of the initial values and to avoid correlation problems, we considered a spacing of size 10, obtaining a sample of size 3,500 from each chain. To monitor the convergence of the Gibbs

Table 2: *Bayesian estimates. Posterior summary results of fitting the Weibull model with a cure fraction and covariates to the data set.*

Parameters	Mean	SD	95% credible interval	$\hat{R}$
α	1.5760	0.1123	(1.353 ; 1.793)	1.017
λ	−1.3020	0.1227	(−1.544 ; −1.071)	1.000
β_0	2.2870	0.5962	(1.164 ; 3.508)	1.002
β_1	−0.1506	0.2325	(−0.603 ; 0.299)	1.001
β_2	−0.0136	0.0086	(−0.031 ; 0.002)	1.001
β_3	−0.5700	0.1268	(−0.826 ; −0.339)	1.005
β_4	0.2095	0.2377	(−0.259 ; 0.674)	1.072
β_5	−0.1508	0.3446	(−0.839 ; 0.509)	1.001
β_6	−0.0681	0.0432	(−0.159 ; 0.009)	1.011

samples, we used the between and within sequence information, following the approach developed in Gelman and Rubin (1992) to obtain the potential scale reduction, $\hat{R}$. In all cases, these values were close to one, indicating the convergence of the chain. In Table 2 we report posterior summaries for the parameters of the Weibull, mixture model and, in Figure 2, we have the approximate marginal posterior densities considering 7,000 Gibbs samples.

In Table 2, we observe that only the covariate nodule (x_3) presents significant effect on lifetime. It is interesting to note that the Bayesian analysis is very similar to the classical analysis. The computational implementation of the algorithm was developed in the software package R jointly with package R2Winbug (see Gelman, 2004), and the programs can be requested from the authors.

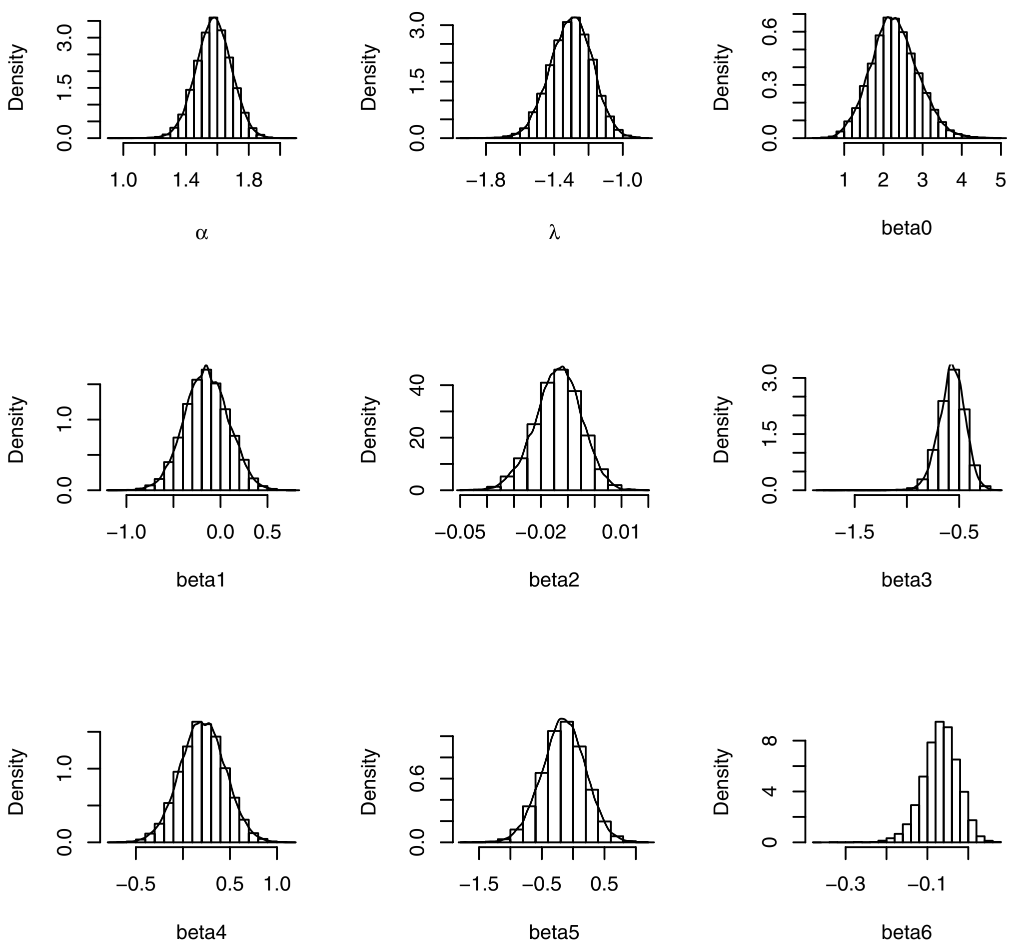

Figure 2: Approximate marginal posterior densities for parameters of the Weibull model with a cure fraction and covariates.

(a) (b)

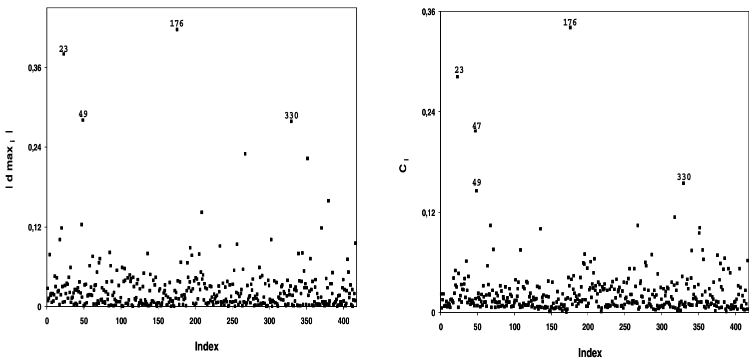

Figure 3: *(a) Index plot of* $\mathbf{d}_{max}$ *for* $\boldsymbol{\theta}$ *(case-weights perturbation). (b) Total local influence on the estimates* $\boldsymbol{\theta}$ *(case-weights perturbation)*

6.3 Local influence analysis

In this section, we will make an analysis of local influence for the data set given in Ibrahim et. al. (2001), using a cure fraction in the Weibull model.

6.3.1 Case-weights perturbation

By applying the local influence theory developed in Section 3, where case-weight perturbation is used, value $C_{\mathbf{d}_{max}} = 1.5820$ was obtained as maximum curvature. In Figure 3(a), the graph of the eigenvector corresponding to $C_{\mathbf{d}_{max}}$ is presented, and total influence C_i is shown in Figure 3(b). Observations ♯23 and ♯176 are the most distinguished in relation to the others.

6.3.2 Prediction of influence using the response variable perturbation

Next, the influence of perturbations on the observed survival times will be analyzed. The value for the maximum curvature calculated was $C_{\mathbf{d}_{max}} = 11.21$. Figure 4 (a), containing the graph for $|\mathbf{d}_{max}|$ versus the observation index, shows that some points were distinguished from the others, among which are points ♯279 and ♯341. The same applies to Figure 4(b), which corresponds to total local influence (C_i). By analyzing the data associated with these two observations, it is noted that the highlighted observations refer to patients with shorter non-censored survival times.

6.3.3 Prediction of influence using the explanatory variable perturbation

The perturbation of vectors for covariates age (x_2) and Breslow (x_6) are investigated here. For perturbation of covariable age, value $C_{\mathbf{d}_{max}} = 1.0374$ was obtained as maximum curvature, and for the perturbation of covariable Breslow, value $C_{\mathbf{d}_{max}} = 1.2864$ was achieved. The respective graphs of $|\mathbf{d}_{max}|$ as well as total local influence C_i against the

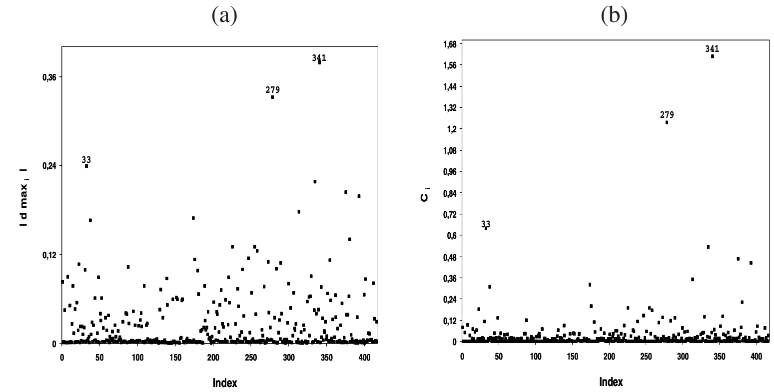

Figure 4: *(a) Index plot of $\mathbf{d}_{max}$ for $\boldsymbol{\theta}$ (response perturbation). (b) Total local influence on the estimates $\boldsymbol{\theta}$ (response perturbation)*

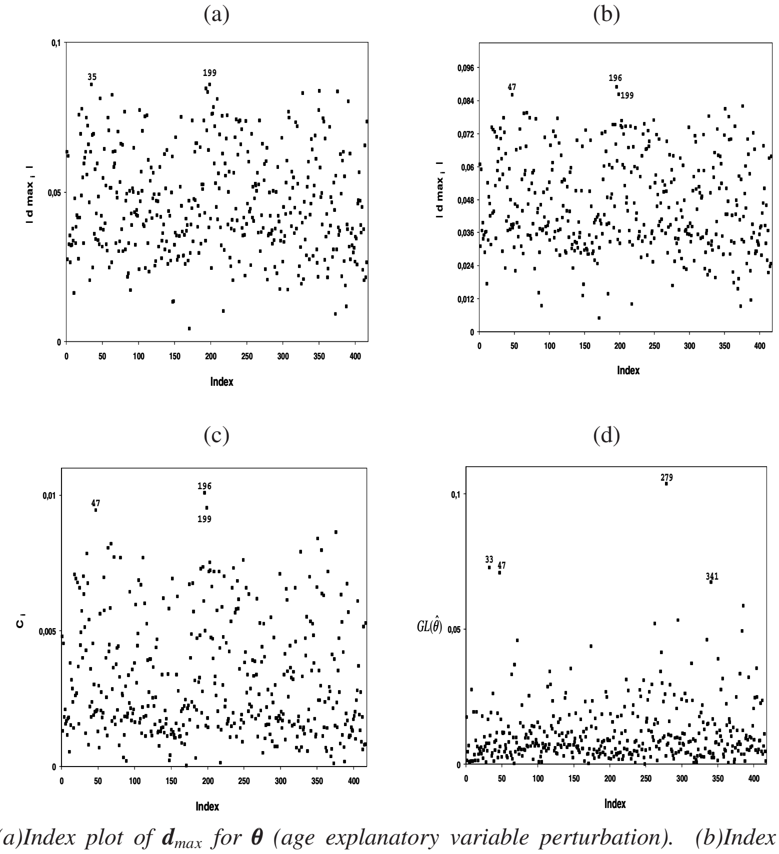

Figure 5: *(a)Index plot of $\mathbf{d}_{max}$ for $\boldsymbol{\theta}$ (age explanatory variable perturbation). (b)Index plot of $\mathbf{d}_{max}$ for $\boldsymbol{\theta}$ (Breslow explanatory variable perturbation). (c)Total local influence on the estimates $\boldsymbol{\theta}$ (Breslow explanatory variable perturbation). (d) Generalized leverage for $\boldsymbol{\theta}$*

observation index are shown in Figures 5(a), 5(b) and 5(c). In these three graphs, we can see no influential observation.

6.3.4 Generalized leverage analysis

Figure 5(d) exhibits the index plot of $GL(\boldsymbol{\theta})$, using the model given in equation (12). The generalized leverage graph presented in Figure 5(d) confirms the tendencies observed under local and total influence methods. Observations with large and small values for **t** tend to have a high influence on these own-fitted values. We note outstanding influence observations ♯33, ♯279 and ♯341. The graph for $GL(\boldsymbol{\theta})$ is very similar to the one given in Figure 4(a).

6.4 Residual analysis

In order to detect possible outlying observations as well as departures from the assumptions generalized log-gamma regression models with a cure fraction, we present, in Figure 6(a) and 6(b), the graphs of r_{M_i} and r_{D_i} against the order observations.

By analyzing the martingale residual and modified deviance residual graph, 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. (2003), Mccullagh et al. (1989), among others.

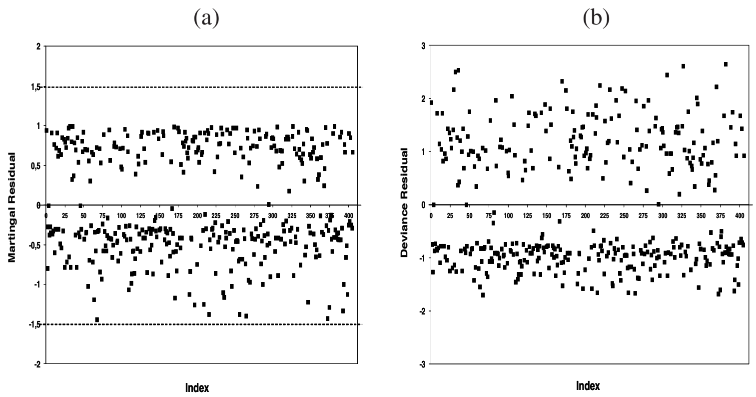

Figure 6: *(a) Index plot of the martingal residual r_{M_i}. (b) Index plot of the modified deviance residual r_{D_i}.*

6.5 Impact of the detected influential observations

Therefore, diagnostic analysis (local influence, local influence on predictions, generalized leverage and residual analysis) detected the following four cases ♯23, ♯176, ♯279 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 4, we have

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

Propping	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	$\hat{\beta}_6$
	–	–	–	–	–	–	–
all observations	2.27	−0.16	−0.01	−0.54	0.20	−0.15	−0.06
	(< 0.01)	(0.48)	(0.10)	(< 0.01)	(0.38)	(0.65)	(0.13)
	[1]	[−5]	[−3]	[−1]	[3]	[−3]	[−1]
♯23	2.23	−0.15	−0.01	−0.53	0.20	−0.15	−0.06
	(< 0.01)	(0.50)	(0.11)	(< 0.01)	(0.40)	(0.64)	(0.12)
	[1]	[−4]	[−2]	[−2]	[4]	[−7]	[−4]
♯176	2.25	−0.17	−0.01	−0.53	0.19	−0.16	−0.06
	(< 0.01)	(0.45)	(0.10)	(< 0.01)	(0.40)	(0.63)	(0.13)
	[1]	[−2]	[−2]	[−1]	[2]	[−14]	[0]
♯279	2.24	−0.16	−0.01	−0.53	0.20	−0.13	−0.06
	(< 0.01)	(0.47)	(0.11)	(< 0.01)	(0.39)	(0.70)	(0.13)
	[1]	[−11]	[−2]	[−1]	[9]	[−8]	[−2]
♯341	2.29	−0.18	−0.01	−0.54	0.22	−0.16	−0.06
	(< 0.01)	(0.43)	(0.09)	(< 0.01)	(0.34)	(0.63)	(0.12)
	[0]	[−3]	[−1]	[−4]	[10]	[−11]	[−5]
♯23/♯176	2.26	−0.16	−0.01	−0.52	0.18	−0.17	−0.06
	(< 0.01)	(0.48)	(0.09)	(< 0.01)	(0.42)	(0.61)	(0.13)
	[1]	[−5]	[−1]	[−3]	[9]	[−9]	[−2]
♯23/♯279	2.25	−0.15	−0.01	−0.52	0.18	−0.14	−0.06
	(< 0.01)	(0.49)	(0.10)	(< 0.01)	(0.42)	(0.68)	(0.12)
	[1]	[−3]	[−3]	[−2]	[3]	[−12]	[0]
♯23/♯341	2.30	−0.17	−0.01	−0.53	0.21	−0.17	−0.06
	(< 0.01)	(0.46)	(0.08)	(< 0.01)	(0.37)	(0.61)	(0.12)
	[2]	[−6]	[−5]	[−3]	[6]	[−7]	[−4]
♯176/♯279	2.22	−0.17	−0.01	−0.52	0.19	−0.14	−0.06
	(< 0.01)	(0.44)	(0.11)	(< 0.01)	(0.41)	(0.67)	(0.13)
	[0]	[−15]	[0]	[−3]	[5]	[−14]	[−2]
♯176/♯341	2.27	−0.18	−0.01	−0.52	0.21	−0.17	−0.06
	(< 0.01)	(0.41)	(0.09)	(< 0.01)	(0.35)	(0.60)	(0.12)
	[0]	[−12]	[−1]	[−2]	[7]	[−6]	[−1]
♯279/♯341	2.26	−0.18	−0.01	−0.53	0.22	−0.14	−0.06
	(< 0.01)	(0.42)	(0.10)	(< 0.01)	(0.35)	(0.67)	(0.12)

the relative changes (in percentage) of each parameter estimate, defined by: $RC_{\boldsymbol{\theta}_j} = [(\hat{\boldsymbol{\theta}}_j - \hat{\boldsymbol{\theta}}_j(I))/\hat{\boldsymbol{\theta}}_j] \times 100$, and the corresponding p-values, where $\hat{\boldsymbol{\theta}}_j(I)$ denotes the MLE of $\boldsymbol{\theta}_j$ after that "set I" of observations has been removed.

*Table 4: Continuation Relative changes [-**RC**- in %], parameter estimates and their p-values in parentheses for the indicated set.*

Propping	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	$\hat{\beta}_6$
	[1]	[−1]	[−3]	[−5]	[13]	[−3]	[−6]
♯23/♯176/♯279	2.24	−0.16	−0.01	−0.51	0.18	−0.15	−0.06
	(< 0.01)	(0.47)	(0.10)	(< 0.01)	(0.44)	(0.65)	(0.13)
	[0]	[−5]	[0]	[−3]	[0]	[−2]	[−1]
♯23/♯279/♯341	2.27	−0.17	−0.01	−0.52	0.20	−0.15	−0.06
	(< 0.01)	(0.45)	(0.09)	(< 0.01)	(0.38)	(0.66)	(0.12)
	[1]	[−17]	[−3]	[−4]	[3]	[−1]	[−2]
♯176/♯279/♯341	2.24	−0.19	−0.01	−0.52	0.21	−0.15	−0.06
	(< 0.01)	(0.40)	(0.10)	(< 0.01)	(0.36)	(0.65)	(0.12)
	[0]	[−10]	[−2]	[−5]	[4]	[−5]	[−4]
♯23/♯176/♯279/♯341	2.26	−0.18	−0.01	−0.51	0.19	−0.16	−0.06
	(< 0.01)	(0.42)	(0.10)	(< 0.01)	(0.39)	(0.63)	(0.12)

From Tables 3 and 4 we can notice some robust aspects of the maximum likelihood estimates from the Weibull model with a cure fraction and covariates. In general, the significance of the parameter estimates does not change after removing set I at the level of 5 %. A significant change was found when observations 23 and 341 were removed, from which it was noted that covariate age was significant if an 8% level were taken into account. Therefore, we did not encounter inferential change after removing the observations given in the diagnostic graphs.

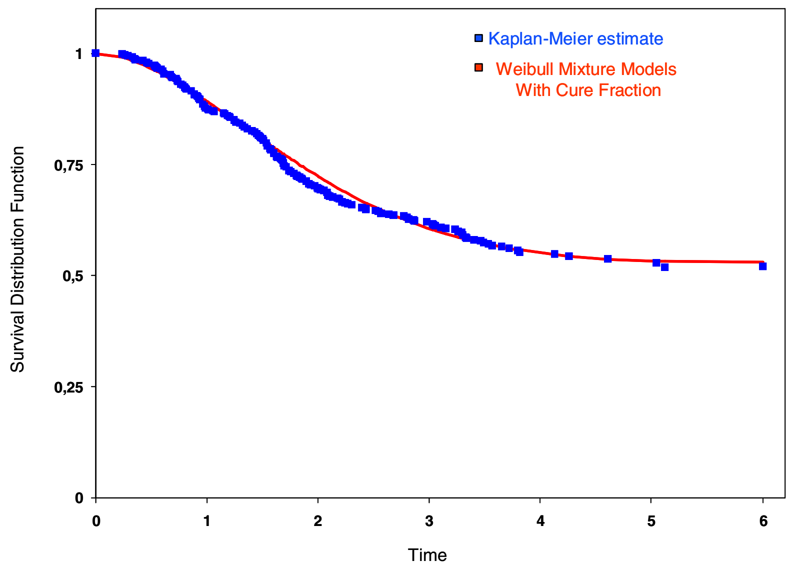

Figure 7: *Theoretical survival curve and Kaplan-Meier curve*

6.6 Quality of fitting

In order to measure quality of fitting, a Kaplan-Meier survival graph and a survival graph estimated by the Weibull model with a cure fraction were plotted. Good model fitting was observed.

7 Concluding Remarks

The local influence theory (Cook (1986) and Thomas and Cook (1990)), that of generalized leverage proposed by Wei et al. (1998) and a study based on martingale and deviance residual in a survival model with a cure fraction were discussed in this study by using two estimation approaches: the maximum likelihood and the Bayesian approaches. The matrices necessary for application of the technique were obtained by taking into account various types of perturbations to the data elements and to the models. By applying such results to a data set, indication was found of which observations or set of observations would sensitively influence the analysis results. This fact is illustrated in Application (Section 7). By means of a real data set, it was observed that, for some perturbation schemes, the presence of certain observations could considerably change the levels of significance of certain variables. The results of the applications indicate that the local influence technique as well as that of generalized leverage in models with a cure fraction can be rather useful in the detection of possibly influential points by admitting two types of estimation methods: maximum likelihood and Bayesian. In order to measure quality of fitting, martingale and deviance residuals were used, which showed that the model fitting was correct. The Kaplan-Meier survival function was also plotted with the survival function for the proposed model, indicating good model fitting.

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

Here we derive the necessary formulas to obtain the second order partial derivatives of the log-likelihood function. After some algebraic manipulations, we obtain

$$\begin{aligned}
\mathbf{L}_{\alpha\alpha} &= -\frac{\mathrm{r}}{\alpha^2} - \exp\{-\lambda\} \sum_{i \in F} t_i^{\alpha} \Big[\log(t_i)\Big]^2 \\
&\quad - \sum_{i \in C} \frac{(1-p_i)\Big[\log(t_i)\Big]^2 \Big[-\log(\mathrm{h}_i)\Big]\mathrm{h}_i \Big[p_i\{1+\log(h_i)\} + (1-p_i)h_i\Big]}{\Big[p_i + (1-p_i)h_i\Big]^2}. \\
\mathbf{L}_{\alpha\lambda} &= -\exp\{\lambda\} \sum_{i \in F} t_i^{\alpha} \log(t_i) \\
&\quad + \sum_{i \in C} \frac{(1-p_i)\Big[\log(t_i)\Big]\Big[\log(h_i)\Big] h_i \Big[p_i\{1+\log(h_i)\} + (1-p_i)h_i\Big]}{\Big[p_i + (1-p_i)h_i\Big]^2}.
\end{aligned}$$

$$\mathbf{L}_{\alpha\beta} = -\sum_{i \epsilon C} \frac{(x_{ij})p_i\Big[\log(t_i)\Big]\Big[\log(h_i)\Big]h_i}{\Big[1+\exp\{\mathbf{x}_i^T\boldsymbol{\beta}\}\Big]\Big[p_i+(1-pi)h_i\Big]^2}.$$

$$\begin{aligned}\mathbf{L}_{\lambda\lambda} &= -\exp\{\lambda\}\sum_{i \epsilon F} t_i^{\alpha} \\ &+ \sum_{i \epsilon C} \frac{(1-p_i)h_i\log(h_i)\Big[1+\log(h_i)+(1-p_i)h_i\{-\log(h_i)\}\Big]}{\Big[p_i+(1-pi)h_i\Big]^2}.\end{aligned}$$

$$\mathbf{L}_{\lambda\beta} = -\sum_{i \epsilon C} \frac{(x_{ij})p_i\Big[-\log(h_i)\Big]h_i}{\Big[1+\exp\{\mathbf{x}_i^T\boldsymbol{\beta}\}\Big]\Big[p_i+(1-pi)h_i\Big]^2}.$$

$$\begin{aligned}\mathbf{L}_{\beta\beta} &= \sum_{i \epsilon F} \frac{-(x_{ij}^2)p_i\Big[1+\exp\{\mathbf{x}_i^T\boldsymbol{\beta}\}(p_i-1)\Big]}{(1-pi)^2\Big[1+\exp\{\mathbf{x}_i^T\boldsymbol{\beta}\}\Big]^2} \\ &+ \sum_{i \epsilon C} \frac{(x_{ij}^2)p_i\Big[1-h_i\Big]\Big\{\Big[1-\exp\{\mathbf{x}_i^T\boldsymbol{\beta}\}\Big]\Big[p_i+(1-pi)h_i\Big]-p_i\Big[1-h_i\Big]\Big\}}{\Big[1+\exp\{\mathbf{x}_i^T\boldsymbol{\beta}\}\Big]^2\Big[p_i+(1-p_i)h_i\Big]^2}.\end{aligned}$$

where $h_i = \exp\{-t_i^{\alpha}e^{\lambda}\}$, $p_i = \dfrac{\exp(\mathbf{x}_i^T\boldsymbol{\beta})}{1+\exp(\mathbf{x}_i^T\boldsymbol{\beta})}$, $i = 1, 2, \ldots, n$ and $j = 1, 2, \ldots, p$.

Appendix B: Local influence: Case-weight perturbation $\ddot{\mathbf{L}}(\boldsymbol{\gamma})$

Here, we provide the derivatives of the elements considering the case-weight perturbation scheme. Then the elements of vector $\boldsymbol{\Delta}_1$ take the form

$$\Delta_{1i} = \begin{cases} \dfrac{1}{\hat{\alpha}} + \log(t_i)[1+\log(\hat{h}_i)], & \text{if} \quad i \epsilon F; \\[2ex] \dfrac{(1-\hat{p}_i)[\log(\hat{h}_i)][\log(t_i)]\hat{h}_i}{[\hat{p}i+(1-\hat{p}_i)\hat{h}_i]}, & \text{if} \quad i \epsilon C. \end{cases}$$

The elements of vector $\boldsymbol{\Delta}_2$ take the form

$$\Delta_{2i} = \begin{cases} 1+\log(\hat{h}_i), & \text{if} \quad i \epsilon F; \\[2ex] \dfrac{(1-\hat{p}_i)[\log(\hat{h}_i)]\hat{h}_i}{[\hat{p}i+(1-\hat{p}_i)\hat{h}_i]}, & \text{if} \quad i \epsilon C. \end{cases}$$

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

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

where

$$\hat{h}_i = \exp\{-t_i^{\hat{\alpha}}e^{\hat{\lambda}}\} \quad \hat{p}_i = \frac{\exp(\mathbf{x}_i^T\hat{\boldsymbol{\beta}})}{1+\exp(\mathbf{x}_i^T\hat{\boldsymbol{\beta}})}.$$

Appendix C: Local influence on predictions: Response perturbation

Here we provide the derivatives of elements $\boldsymbol{\Delta}_{ij}$ of matrix $\boldsymbol{\Delta}$ considering the response variables perturbation scheme. The elements of vector $\boldsymbol{\Delta}_1$ take the form

$$\Delta_{1i} = \begin{cases} \dfrac{(S_t)}{t_i} + (S_t)\log(\hat{h}_i)(t_i)^{-1}[(\hat{\alpha})\log(t_i)+1], & \text{if} \quad i \in F; \\ (1-\hat{p}_i)(S_t)\log(\hat{h}_i)(t_i)^{-1}\hat{h}_i\left\{\dfrac{(\hat{\alpha})\log(t_i)[1+\log(\hat{h}_i)]+1}{\hat{p}_i+(1-\hat{p}_i)\hat{h}_i} - \right. & \\ \left.\dfrac{(\hat{\alpha})\log(\hat{h}_i)(1-\hat{p}_i)\log(t_i)\hat{h}_i}{[\hat{p}_i+(1-\hat{p}_i)\hat{h}_i]^2}\right\}, & \text{if} \quad i \in C. \end{cases}$$

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

$$\Delta_{2i} = \begin{cases} (S_t)(\hat{\alpha})\log(\hat{h}_i)(t_i)^{-1}, & \text{if} \quad i \in F; \\ (S_t)(1-\hat{p}_i)(\hat{\alpha})\log(\hat{h}_i)(t_i)^{-1}\left\{\dfrac{1+\log(\hat{h}_i)}{\hat{p}_i[\hat{h}_i-1]+1} - \dfrac{(1-\hat{p}_i)\log(\hat{h}_i)}{[\hat{p}_i(\hat{h}_i-1)+1]^2}\right\}, & \text{if} \quad i \in C. \end{cases}$$

and the elements of the vector $\boldsymbol{\Delta}_j$, $j = 3, \ldots, p+2$ are expressed as

$$\Delta_{ji} = \begin{cases} 0, & \text{if} \quad i \in F; \\ -(x_{ij})(S_t)(\hat{p}_i)(\hat{\alpha})\log(\hat{h}_i)(t_i)^{-1}\hat{h}_i & \\ \left\{\dfrac{1}{[1+\exp\{\mathbf{x}_i^T\hat{\boldsymbol{\beta}}\}][\hat{p}i+(1-\hat{p}i)\hat{h}_i]} + \dfrac{(1-\hat{p}_i)(1-\hat{h}_i)}{[1+\exp\{\mathbf{x}_i^T\hat{\boldsymbol{\beta}}\}][\hat{p}_i+(1-\hat{p}_i)\hat{h}_i]^2}\right\}, & \text{if} \quad i \in C. \end{cases}$$

where

$$\hat{h}_i = \exp\{-t_i^{\hat{\alpha}}e^{\hat{\lambda}}\} \quad \hat{p}_i = \frac{\exp(\mathbf{x}_i^T\hat{\boldsymbol{\beta}})}{1+\exp(\mathbf{x}_i^T\hat{\boldsymbol{\beta}})}.$$

Appendix D: Local influence on predictions: Explanatory variable perturbation

In this appendix we provide the derivatives of elements $\boldsymbol{\Delta}_{ij}$ of matrix $\boldsymbol{\Delta}$ considering the explanatory variables perturbation scheme. The elements of vector $\boldsymbol{\Delta}_1$ are expressed as

$$\Delta_{1i} = \begin{cases} 0, & \text{if} \quad i \in F; \\ -\dfrac{\hat{\beta}_t(S_x)(\hat{p}_i)\log(\hat{h}_i)\log(t_i)}{[1+\exp\{\mathbf{x}_i^T\hat{\boldsymbol{\beta}}\}]}\left\{\dfrac{\hat{h}_i}{\hat{p}_i+(1-\hat{p}_i)\hat{h}_i} + \dfrac{(1-\hat{p}i)\hat{h}_i(1-\hat{h}_i)}{[\hat{p}_i+(1-\hat{p}_i)\hat{h}_i]^2}\right\}, & \text{if} \quad i \in C. \end{cases}$$

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

$$\Delta_{2i} = \begin{cases} 0, & \text{if} \quad i \in F; \\ -\dfrac{(\beta_t)(S_x)\log(\hat{h}_i)(\hat{h}_i)(\hat{p}_i)}{[1+\exp\{\mathbf{x}_i^T\hat{\boldsymbol{\beta}}\}][\hat{p}_i+(1-\hat{p}_i)\hat{h}_i]^2}, & \text{if} \quad i \in C. \end{cases}$$

the elements of vector $\boldsymbol{\Delta}_j$, for $j = 3, \ldots, p+2$ and $j \neq t$, take the forms

$$\Delta_{ji} = \begin{cases} -\dfrac{x_{ij}(\hat{\beta}_t)(S_x)(\hat{p}_i)}{[1+\exp\{\mathbf{x}_i^T\hat{\boldsymbol{\beta}}\}]}, & \text{if} \quad i \in F; \\ -x_{ij}(\hat{p}_i)(S_x)(\hat{\beta}_t)(1-\hat{h}_i)\left\{\dfrac{\hat{p}_i(1-\hat{h}_i)}{[1+\exp\{\mathbf{x}_i^T\boldsymbol{\beta}\}]^2[\hat{p}_i+(1-\hat{p}i)\hat{h}_i]^2} \right. & \\ \left. -\dfrac{[1-\exp\{x_i^T\hat{\boldsymbol{\beta}}\}]}{[\hat{p}_i+(1-\hat{p}_i)\hat{h}_i][1+\exp\{\mathbf{x}_i^T\hat{\boldsymbol{\beta}}\}]^2}\right\}, & \text{if} \quad i \in C. \end{cases}$$

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

$$\Delta_{ti} = \begin{cases} -(S_x)(\hat{p}_i)\left[1+\dfrac{x_{it}\hat{\beta}_t}{[1+\exp\{\mathbf{x}_i^T\hat{\boldsymbol{\beta}}\}]}\right], & \text{se} \quad i \in F; \\ \dfrac{(S_x)(\hat{p}_i)(1-\hat{p}_i)^2(1-\hat{h}_i)}{[\hat{p}_i+(1-\hat{p}_i)\hat{h}_i]}\left\{1+\hat{\beta}_t x_{it}[1-\exp\{\mathbf{x}_i^T\hat{\boldsymbol{\beta}}\}]\right\} & \\ -\dfrac{(x_{it})(S_x)\hat{\beta}_t(\hat{p}_i)(1-\hat{p}_i)^2[1-\hat{h}_i]^2}{[\hat{p}_i+(1-\hat{p}i)\hat{h}_i]^2}, & \text{se} \quad i \in C. \end{cases}$$

where

$$\hat{h}_i = \exp\{-t_i^{\hat{\alpha}}e^{\hat{\lambda}}\} \quad \hat{p}_i = \frac{\exp(\mathbf{x}_i^T\hat{\boldsymbol{\beta}})}{1+\exp(\mathbf{x}_i^T\hat{\boldsymbol{\beta}})}.$$

Appendix E: Generalized leverage

In this appendix we provide the derivatives of elements $\mathbf{D}_\alpha, \mathbf{D}_\lambda, \mathbf{D}_{\boldsymbol{\beta}}$, of matrix $\mathbf{D}_{\boldsymbol{\theta}}$ considering generalized leverage.

The elements of vector $\mathbf{D}_{\boldsymbol{\theta}}$ are expressed as

$$\mathbf{D}_\alpha = (1-\hat{p}_i)(\hat{\alpha})^{-2}\exp\Big\{-\frac{\hat{\lambda}}{\hat{\alpha}}\Big\}\Big\{\log\Big(\frac{\hat{\alpha}+1}{\hat{\alpha}}\Big)+\Big(\frac{\hat{\alpha}+1}{\hat{\alpha}}\Big)\Big\}.$$

$$\mathbf{D}_\lambda = (1-\hat{p}_i)(\hat{\alpha}^{-1})\Big(-\exp\Big\{-\frac{\hat{\lambda}}{\hat{\alpha}}\Big\}\Big)\Big\{\log\Big(\frac{\hat{\alpha}+1}{\hat{\alpha}}\Big)\Big\}.$$

$$\mathbf{D}_{\beta_j} = (x_{ij})(\hat{p}_i)\Big[1+\exp\{\mathbf{x}_i^T\hat{\boldsymbol{\beta}}\}\Big]^{-1}\exp\Big\{-\frac{\hat{\lambda}}{\hat{\alpha}}\Big\}\log\Big(\frac{\hat{\alpha}+1}{\hat{\alpha}}\Big).$$

where

$$\ddot{\mathbf{L}}_{\boldsymbol{\theta}\mathbf{t}} = \frac{\partial^2 l(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}\,\partial\mathbf{t}^T} = \begin{pmatrix}\ddot{\mathbf{L}}_{\alpha t_i}\\ \ddot{\mathbf{L}}_{\lambda t_i}\\ \ddot{\mathbf{L}}_{\beta_j t_i}\end{pmatrix}$$

with

$$\ddot{\mathbf{L}}_{\alpha t_i} = \begin{cases} t_i^{-1}-\exp\{\hat{\lambda}\}t_i^{\hat{\alpha}-1}[\hat{\alpha}\log(t_i)+1], & \forall i: i\in F;\\ -\hat{g}_i^{-2}(1-\hat{p}_i)\exp\{\hat{\lambda}\}\hat{h}_i t_i^{\hat{\alpha}-1}\log(t_i) & \\ \Big\{\hat{g}_i(-\exp\{\hat{\lambda}\}\hat{\alpha}t_i^{\hat{\alpha}}+\hat{\alpha}+[\log(t_i)]^{-1})- & \\ (1-\hat{p}_i)\hat{h}_i\exp\{\hat{\lambda}\}\hat{\alpha}t_i^{\hat{\alpha}}\Big\}, & \forall i: i\in C.\end{cases}$$

$$\ddot{\mathbf{L}}_{\lambda t_i} = \begin{cases} -\hat{\alpha}t_i^{\hat{\alpha}-1}\exp\{\hat{\lambda}\}, & \forall i: i\in F;\\ \hat{g}_i^{-2}(1-\hat{p}_i)\exp\{\hat{\lambda}\}\hat{h}_i\hat{\alpha}t_i^{\hat{\alpha}-1}\Big[\hat{g}_i\hat{\alpha}t_i^{\hat{\alpha}}+(1-\hat{p}_i)\exp\{\hat{\lambda}\}t_i^{\hat{\alpha}}\hat{h}_i\Big], & \forall i: i\in C.\end{cases}$$

$$\ddot{\mathbf{L}}_{\beta_j t_i} = \begin{cases} 0, & \forall i: i\in F;\\ \hat{g}_i^{-2}\hat{p}_i x_{ij}[1+\exp\{\mathbf{x}_i^T\hat{\boldsymbol{\beta}}\}]^{-1}\hat{h}_i\exp\{\hat{\lambda}\}\hat{\alpha}t_i^{\hat{\alpha}-1}\times\Big\{\hat{g}_i-(1-\hat{p}_i)[1-\hat{h}_i]\Big\}, & \forall i: i\in C.\end{cases}$$

where

$$\hat{h}_i = \exp\Big\{-t_i^{\hat{\alpha}}\exp\{\hat{\lambda}\}\Big\},\quad \hat{g}_i = \hat{p}_i+(1-\hat{p}_i)\hat{h}_i \quad\text{and}\quad \hat{p}_i = \frac{\exp\{\mathbf{x}_i^T\hat{\boldsymbol{\beta}}\}}{1+\exp\{\mathbf{x}_i^T\hat{\boldsymbol{\beta}}\}}.$$

References

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

Beckman, R. J., Nachtsheim, C. J. and Cook, R. D. (1987). Diagnostics for mixed-model analysis of variance. *Technometrics*, 29, 413-426.

Bolfarine, H. and Cancho, V. (2001). Modelling the presence of immunes by using the exponentiated-Weibull model. *Journal of Applied Statistics*, 28, 659-671.

Cancho, V., Bolfarine, H. and Achcar, J. A. (1999). A Bayesian analysis for the Exponentiated-Weibull distribution. *Journal Applied Statistical Science*, 8, 227-242.

Chatterjee, S. and Hadi, A. S. (1988). *Sensitivity Analysis in Linear Regression.* New York: John Wiley.

Cox, D. R. and Snell, E. J. (1968). A general definition of residuals. *Journal of the Royal Statistical Society B*, 30, 248-275.

Collet, D. (1994). *Modelling Survival Data in Medical Research.* Chapman and Hall: London.

Cook, R. D. (1977). Detection of influential observations in linear regression. *Technometrics* , 19 15-18.

Cook, R. D. (1986). Assesment of local influence (with discussion). *Journal of the Royal Statistical Society*, 48 2, 133-169.

Cook, R. D. and Weisberg, S. (1982). *Residuals and Influence in Regression.* New York: Chapman and Hill.

Davison, A. C. and Gigli, A. (1989). Deviance residuals and normal scores plots. *Biometrika*, 76, 211-221.

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

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

Fachini, J. B., Ortega, E. M. M. and Louzada-Neto, F. (2007). Influence diagnostics for polyhazard models in the presence of covariates. Statistical Methods and Applications. In Press.

Fahrmeir, L. and Tutz, G. (1994). *Multivariate Statistical Modelling Based on Generalized Linear Models.* Springer-Verlag: New York.

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

Fung, W. K. and Kwan, C. W. (1997). A note on local influence based on normal curvature. *Journal of the Royal Statistical Society B*, 59, 839-843.

Galea, M., Paula, G. A. and Bolfarine, H. (1997). Local influence in elliptical linear regression models. *The Statistician*, 46, 71-79.

Gelman, A. (2004). Running WingBUGS from R, Available for download at (http:carn.us.r-project.org/src/contrib/Descriptions/R2WinBUGS.html).

Gu, H. and Fung, W. K. (1998). Assessing local influence in canonical analysis. *Annals of the Institute of Statistical Mathematics*, 50, 755-772.

Hinkley, D. V., Reid, N. and Snell, E. J.(1991). *Statistical Theory and Modelling-In honour of Sir David Cox*. Chapman & Hall, London.

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

Kim, M. G. (1995). Local influence in multivariate regression. *Communications in Statistics: Theory and Methods*, 20, 1271-1278.

Kwan, C. W. and Fung, W. K. (1998). Assesing local influence for specific restricted likelihood: Applications to factor analysis. *Psychometrika*, 63, 35-46.

Lawrence, A. J. (1988). Regression transformation diagnostics using local influence. *Journal of the American Statistical Association*, 83, 1067-1072.

Lawless, J. F. (1982). *Statistical Models and Methods for Lifetime Data*. Wiley: New York.

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

Liu, S. Z. (2000). On local influence for elliptical linear models. *Statistical Papers*, 41, 211-224.

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

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

Mudholkar, G. S., Srivastava, D. K. and Friemer, M. (1995). The exponentiated Weibull family: A reanalysis of the bus-motor-failure data. *Technometrics*, 37, 436-445.

Nelson, W. B. (1990). *Accelerated Testing; Statistical Models, Test Plans and Data Analysis*. New York: John Wiley.

O'Hara, R. J., Lawless, J. F. and Carter, E. M. (1992). Diagnostics for a cumulative multinomial generalized linear model with application to grouped toxicological mortality data. *Journal of the American Statistical Association*, 87, 1059-1069.

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.

Ortega, E. M. M., Cancho, V. G., Bolfarine, H. (2006). Influence diagnostics in exponentiated -Weibull regression models with censored data. *Statistics and Operation Reserch Transactions*, 30, 171-192.

Ortega, E. M. M., Paula, G. A. and Bolfarine, H. (2007). Deviance Residuals in Generalized Log-Gamma Regression Models with Censored Observations. *Journal of Statistical Computation and Simulation*, 77. In press.

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

Paula, G. A. (1995). Influence residuals in restricted generalized linear models. *Journal of Statistical Computation and Simulation*, 51, 63-79.

Paula, G. A. (1999). Leverage in inequality constrained regression models. *The Statistician*, 48, 529-538.

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

Prentice, R. L. (1974). A log-gamma model and its maximum likelihood estimation. *Biometrica*, 61, 539-544.

Stacy, E. W. (1962). A generalization of the gamma distribution. *Ann. Math. Stat.*, 33, 1187-1192.

St. Laurent, R. T. and Cook, R. D. (1992). Leverage and superleverage in nonlinear regression. *Journal of the American Statistical Association*, 87, 985-990.

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

Tsai, C. and Wu, X. (1992). *Transformation-model diagnostics*. *Technometrics*, 34, 197-202.

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

Williams, D. A. (1987). Generalized linear model diagnostic using the deviance and single case deletion. *Applied Statistics*, 36, 181-191.

Wei, B., Hu, Y. Q. and Fung, W. K. (1998). Generalized leverage and its applications. *Scandinavian Journal of statistics*, 25, 25-37.