Stat Papers (2011) 52:591–619
DOI 10.1007/s00362-009-0271-3

REGULAR ARTICLE

The generalized inverse Weibull distribution

Felipe R. S. de Gusmão · Edwin M. M. Ortega · Gauss M. Cordeiro

Received: 2 February 2009 / Revised: 21 July 2009 / Published online: 21 August 2009

Abstract The inverse Weibull distribution has the ability to model failure rates which are quite common in reliability and biological studies. A three-parameter generalized inverse Weibull distribution with decreasing and unimodal failure rate is introduced and studied. We provide a comprehensive treatment of the mathematical properties of the new distribution including expressions for the moment generating function and the rth generalized moment. The mixture model of two generalized inverse Weibull distributions is investigated. The identifiability property of the mixture model is demonstrated. For the first time, we propose a location-scale regression model based on the log-generalized inverse Weibull distribution for modeling lifetime data. In addition, we develop some diagnostic tools for sensitivity analysis. Two applications of real data are given to illustrate the potentiality of the proposed regression model.

Keywords Censored data · Data analysis · Inverse Weibull distribution · Maximum likelihood estimation · Moment · Weibull regression model

F. R. S. de Gusmão · G. M. Cordeiro
DEINFO, Universidade Federal Rural de Pernambuco, Recife, Brazil

F. R. S. de Gusmão
e-mail: felipe556@gmail.com

G. M. Cordeiro
e-mail: gausscordeiro@uol.com.br

E. M. M. Ortega (✉)
Departamento de Ciências Exatas, ESALQ, Universidade de São Paulo, Av. Pádua Dias 11, Caixa Postal 9, Piracicaba, SP 13418-900, Brazil
e-mail: edwin@esalq.usp.br

1 Introduction

The inverse Weibull (IW) distribution has received some attention in the literature. Keller and Kamath (1982) study the shapes of the density and failure rate functions for the basic inverse model. The random variable Y has an inverse Weibull distribution if its cumulative distribution function (cdf) takes the form

$$G(t) = \exp\left[-\left(\frac{\alpha}{t}\right)^{\beta}\right], \quad t > 0, \tag{1}$$

where $\alpha > 0$ and $\beta > 0$. The corresponding probability density function (pdf) is

$$g(t) = \beta\alpha^{\beta} t^{-(\beta+1)} \exp\left[-\left(\frac{\alpha}{t}\right)^{\beta}\right]. \tag{2}$$

The IW distribution is also a limiting distribution of the largest order statistics. Drapella (1993) and Mudholkar and Kollia (1994) suggested the names *complementary Weibull* and *reciprocal Weibull* for the distribution (2). The corresponding survival and hazard functions are

$$S_G(t) = 1 - G(t) = 1 - \exp\left[-\left(\frac{\alpha}{t}\right)^{\beta}\right]$$

and

$$h_G(t) = \beta\alpha^{\beta} t^{-(\beta+1)} \exp\left[-\left(\frac{\alpha}{t}\right)^{\beta}\right] \left\{1 - \exp\left[-\left(\frac{\alpha}{t}\right)^{\beta}\right]\right\}^{-1},$$

respectively. The kth moment about zero of (2) is

$$E(T^k) = \alpha^k \Gamma(1 - k\beta^{-1}),$$

where $\Gamma(\cdot)$ is the gamma function.

The rest of the paper is organized as follows. In Sect. 2, we define the generalized inverse Weibull (GIW) distribution. Section 3 provides a general formula for its moments. General expansions for the moments of order statistics of the new distribution are given in Sect. 4. We discuss in Sect. 5 maximum likelihood estimation with censored data. Section 6 provides some properties of a mixture of two GIW distributions. Sections 7 and 8 discuss the log-generalized inverse Weibull distribution and the corresponding regression model, respectively. Some diagnostic measures for the log-generalized inverse Weibull model are proposed in Sect. 9. Section 10 defines two type of residuals. Section 11 illustrates two applications to real data sets including the model fitting and influence analysis. Section 12 ends with some conclusions.

2 The generalized inverse Weibull distribution

Let $G(t)$ be the cdf of the inverse Weibull distribution discussed by Drapella (1993), Mudholkar and Kollia (1994) and Jiang et al. (1999), among others. The cdf of the GIW distribution can be defined by elevating $G(t)$ to the power of $\gamma > 0$, say $F(t) = G(t)^{\gamma} = \exp\left[-\gamma\left(\frac{\alpha}{t}\right)^{\beta}\right]$. Hence, the GIW density function with three parameters $\alpha > 0$, $\beta > 0$ and $\gamma > 0$ is given by

$$f(t) = \gamma\beta\alpha^{\beta}t^{-(\beta+1)}\exp\left[-\gamma\left(\frac{\alpha}{t}\right)^{\beta}\right], \quad t > 0. \tag{3}$$

We can easily prove that (3) is a density function by substituting $u = -\gamma\alpha^{\beta}t^{-\beta}$. The inverse Weibull distribution is a special case of (3) when $\gamma = 1$. If T is a random variable with density (3), we write $T \sim \text{GIW}(\alpha, \beta, \gamma)$. The corresponding survival and hazard functions are

$$S(t) = 1 - F(t) = 1 - \exp\left[-\gamma\left(\frac{\alpha}{t}\right)^{\beta}\right]$$

and

$$h(t) = \gamma\beta\alpha^{\beta}t^{-(\beta+1)}\exp\left[-\gamma\left(\frac{\alpha}{t}\right)^{\beta}\right]\left\{1 - \exp\left[-\gamma\left(\frac{\alpha}{t}\right)^{\beta}\right]\right\}^{-1},$$

respectively.

We can simulate the GIW distribution using the nonlinear equation

$$t = \alpha\left[\frac{-\log(u)}{\gamma}\right]^{-1/\beta}, \tag{4}$$

where u has the uniform $U(0, 1)$ distribution. Plots of the GIW density and survival function for selected parameter values are given in Figs. 1 and 2, respectively.

2.1 Characterizing the failure rate function

The first derivative $h'(t) = \mathrm{d}h(t)/\mathrm{d}t$ to study the hazard function shape is given by

$$h'(t) = h(t)t^{-(\beta+1)}\left\{\gamma\beta\alpha^{\beta}\left\{1 - \exp\left[-\gamma\left(\frac{\alpha}{t}\right)^{\beta}\right]\right\}^{-1} - (\beta+1)t^{\beta}\right\}.$$

We can easily show that $h(t)$ is unimodal with a maximum value at t^*, where t^* satisfies the nonlinear equation

$$\gamma\left(\frac{\alpha}{t^*}\right)\left\{1 - \exp\left[-\gamma\left(\frac{\alpha}{t^*}\right)^{\beta}\right]\right\}^{-1} = 1 + \beta^{-1}.$$

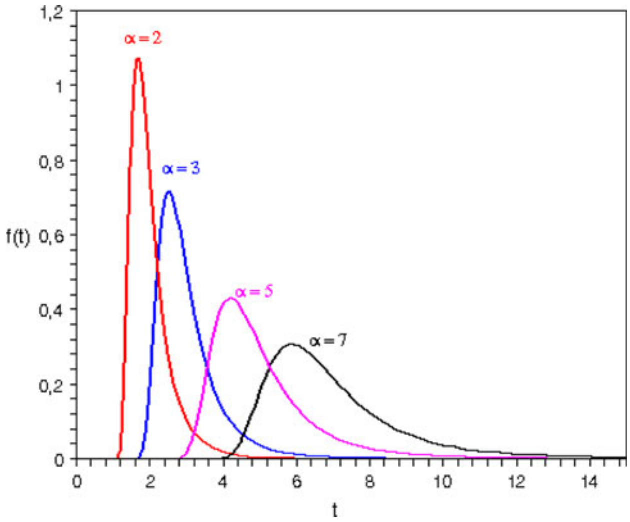

Fig. 1 Plots of the GIW density for some values of α with $\beta = 5$ and $\gamma = 0.5$

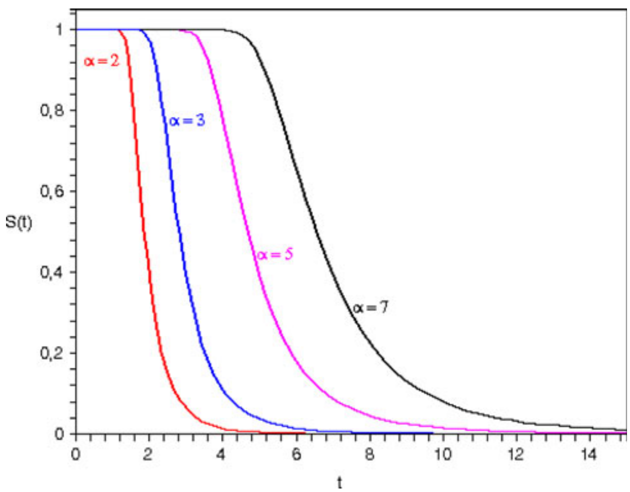

Fig. 2 Plots of the GIW survival functions for some values of α with $\beta = 5$ and $\gamma = 0.5$

Evidently, $\lim_{t\to 0} h(t) = \lim_{t\to\infty} h(t) = 0$.

3 A general formula for the moments

We hardly need to emphasize the necessity and importance of the 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). If the random variable T has the GIW density function (3), the kth moment about zero is

$$E(T^k) = \gamma^{\frac{k}{\beta}} \alpha^k \Gamma(1 - k\beta^{-1}). \tag{5}$$

The moment generating function $M(z)$ of T for $|z| < 1$ is

$$M(z) = E(\mathrm{e}^{zT}) = \sum_{k=0}^{n} \left\{ \frac{z^k}{k!} \gamma^{\frac{k}{\beta}} \alpha^k \Gamma\left(1 - k\beta^{-1}\right) \right\}.$$

Hence, for $|z| < 1$, the cumulant generating function of T is

$$K(z) = \log[M(z)] = \log \left\{ \sum_{k=0}^{n} \left[\frac{z^k}{k!} \gamma^{\frac{k}{\beta}} \alpha^k \Gamma\left(1 - k\beta^{-1}\right) \right] \right\}. \tag{6}$$

The first four cumulants obtained from (6) are

$$\kappa_1 = \gamma^{\frac{1}{\beta}} \alpha \Gamma_1(\beta), \quad \kappa_2 = \gamma^{\frac{2}{\beta}} \alpha^2 \{\Gamma_2(\beta) - \Gamma_1(\beta)^2\},$$
$$\kappa_3 = \gamma^{\frac{3}{\beta}} \alpha^3 \{\Gamma_3(\beta) - 3\Gamma_2(\beta)\Gamma_1(\beta) + 2\Gamma_1(\beta)^3\}$$

and

$$\kappa_4 = \gamma^{\frac{4}{\beta}} \alpha^4 \{\Gamma_4(\beta) - 4\Gamma_3(\beta)\Gamma_1(\beta) - 3\Gamma_1(\beta)^2 + 12\Gamma_2(\beta)\Gamma_1(\beta)^2 - 6\Gamma_1(\beta)^4\},$$

where $\Gamma_k(\beta) = \Gamma(1 - k\beta^{-1})$ for $k = 1, \ldots, 4$. The skewness and kurtosis are then $\rho_3 = \kappa_3/\kappa_2^{\frac{3}{2}}$ and $\rho_4 = \kappa_4/\kappa_2^2$, respectively.

The Shannon entropy of a random variable T with density $f(t)$ is a measure of the uncertainty defined by $E\{-\log[f(T)]\}$. The Shannon entropy for the GIW distribution can be expressed as

$$E\{-\log[f(T)]\} = -\log(\alpha\beta\gamma) + (\beta + 1)\{\log(\alpha) + \beta^{-1}[\log(\gamma) + 0.577216]\} + 1,$$

where 0.577216 is the approximate value for Euler's constant.

4 Moments of order statistics

Order statistics make their appearance in many areas of statistical theory and practice. Let the random variable $T_{r:n}$ be the rth order statistic ($T_{1:n} \leq T_{2:n} \leq \cdots \leq T_{n:n}$) in a sample of size n from the GIW distribution, for $r = 1, \ldots, n$. The pdf of $T_{r:n}$ is given by

$$f_{r:n}(t) = C_{r:n} f(t) F(t)^{r-1} \left[1 - F(t)\right]^{n-r}, \quad t > 0,$$

where $f(t)$ comes from (3), $F(t) = \exp\left[-\gamma \left(\frac{\alpha}{t}\right)^{\beta}\right]$ and $C_{r:n} = n!/[(r-1)!(n-r)!]$.

The kth moment about zero of the rth order statistic is

$$\mu_{r:n}^{(k)} = E(T_{r:n}^k) = C_{r:n} \int_0^\infty f(t) F(t)^{r-1} \left[1 - F(t)\right]^{n-r} \mathrm{d}t.$$

After some algebra, we can obtain

$$\mu_{r:n}^{(k)} = C_{r:n} \alpha^k \gamma^{k/\beta}\, \Gamma[1 - (k/\beta)] \sum_{i=0}^{n-r} \binom{n-r}{i} (-1)^i (r+i)^{\frac{k}{\beta}-1}. \tag{7}$$

An alternative formula for this moment (Barakat and Abdelkader 2004) is

$$E(T_{r:n}^k) = k \sum_{j=n-i+1}^{n} (-1)^{j-n+i-1} \binom{j-1}{n-i} \binom{n}{j} I_j(k), \tag{8}$$

where $I_j(k)$ denotes the integral

$$I_j(k) = \int_0^\infty t^{k-1} [1 - F(t)]^j \,\mathrm{d}t. \tag{9}$$

From Eqs. 8 and 9, we have

$$E(T_{r:n}^k) = k(\gamma\alpha^\beta)^{\frac{k}{\beta}} \Gamma_k(\beta) \sum_{j=n-i+1}^{n} \sum_{l=0}^{j} (-1)^{j-n+i-1} \binom{j-1}{n-i} \binom{n}{j} (-1)^l \binom{j}{l} l^{\frac{k}{\beta}-1}. \tag{10}$$

For $\gamma = 1$, we obtain the moments of order statistics of the IW distribution. Equations 7 and 10 are the main results of this section.

5 Maximum likelihood estimation with censored data

Let T_i be a random variable distributed as (3) with the vector of parameters $\boldsymbol{\theta} = (\alpha, \beta, \gamma)^T$. The data encountered in survival analysis and reliability studies are often censored. A very simple random censoring mechanism that is often realistic is one in which each individual i is assumed to have a lifetime T_i and a censoring time C_i, where T_i and C_i are independent random variables. Suppose that the data consist of n independent observations $t_i = \min(T_i, C_i)$ for $i = 1, \ldots, n$. The distribution of C_i does not depend on any of the unknown parameters of T_i. Parametric inference for such model is usually based on likelihood methods and their asymptotic theory. The censored log-likelihood $l(\boldsymbol{\theta})$ for the model parameters reduces to

$$l(\boldsymbol{\theta}) = r\left[\log(\gamma) + \log(\beta) + \beta\log(\alpha)\right] - (\beta+1)\sum_{i\in F}\log(t_i) - \gamma\alpha^{\beta}\sum_{i\in F} t_i^{-\beta}$$
$$+\sum_{i\in C}\log\left\{1-\exp\left[-\gamma\left(\frac{\alpha}{t_i}\right)^{\beta}\right]\right\},$$

where r is the number of failures and F and C denote the sets of uncensored and censored observations, respectively.

The score functions for the parameters α, β and γ are given by

$$U_\alpha(\boldsymbol{\theta}) = \frac{r\beta}{\alpha} - \gamma\beta\alpha^{\beta-1}\sum_{i\in F} t_i^{-\beta} + \gamma\beta\alpha^{\beta-1}\sum_{i\in C} t_i^{-\beta}\left(\frac{1-u_i}{u_i}\right),$$
$$U_\beta(\boldsymbol{\theta}) = \frac{r}{\beta} + r\log(\alpha) - \sum_{i\in F}\log(t_i) - \gamma\alpha^{\beta}\sum_{i\in F} t_i^{-\beta}\log\left(\frac{\alpha}{t_i}\right)$$
$$+\gamma\alpha^{\beta}\sum_{i\in C} t_i^{-\beta}\log\left(\frac{\alpha}{t_i}\right)\left(\frac{1-u_i}{u_i}\right) \quad \text{and}$$
$$U_\gamma(\boldsymbol{\theta}) = \frac{r}{\gamma} - \alpha^{\beta}\sum_{i\in F} t_i^{-\beta} + \alpha^{\beta}\sum_{i\in C} t_i^{-\beta}\left(\frac{1-u_i}{u_i}\right),$$

where $u_i = 1-\exp\left[-\gamma\left(\frac{\alpha}{t_i}\right)^{\beta}\right]$ is the ith transformed observation.

The MLE $\widehat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}$ is obtained by solving the nonlinear likelihood equations $U_\alpha(\boldsymbol{\theta}) = 0$, $U_\beta(\boldsymbol{\theta}) = 0$ and $U_\gamma(\boldsymbol{\theta}) = 0$ using iterative numerical techniques such as the Newton–Raphson algorithm. We employ the programming matrix language Ox (Doornik 2007).

For interval estimation of (α, β, γ) and hypothesis tests on these parameters, we derive the observed information matrix since the expected information matrix is very complicated and requires numerical integration. The 3×3 observed information matrix $J(\boldsymbol{\theta})$ is

$$J(\boldsymbol{\theta}) = -\begin{pmatrix} \mathbf{L}_{\alpha\alpha} & \mathbf{L}_{\alpha\beta} & \mathbf{L}_{\alpha\gamma} \\ \cdot & \mathbf{L}_{\beta\beta} & \mathbf{L}_{\beta\gamma} \\ \cdot & \cdot & \mathbf{L}_{\gamma\gamma} \end{pmatrix},$$

whose elements are given in Appendix 1. Under conditions that are fulfilled for parameters in the interior of the parameter space but not on the boundary, the asymptotic distribution of

$$\sqrt{n}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \text{ is } N_3(0, I(\boldsymbol{\theta})^{-1}),$$

where $I(\boldsymbol{\theta})$ is the expected information matrix. This asymptotic behavior is valid if $I(\boldsymbol{\theta})$ is replaced by $J(\widehat{\boldsymbol{\theta}})$, i.e., the observed information matrix evaluated at $\widehat{\boldsymbol{\theta}}$. The multivariate normal $N_3(0, J(\widehat{\boldsymbol{\theta}})^{-1})$ distribution can be used to construct approximate confidence regions for the parameters and for the hazard and survival functions. The asymptotic normality is also useful for testing goodness of fit of the GIW distribution and for comparing this distribution with some of its special sub-models using one of the three well-known asymptotically equivalent test statistics—namely, the likelihood ratio (LR) statistic, Wald and Rao score statistics.

6 Mixture of two GIW distributions

Mixture distributions have been considered extensively by many authors. For an excellent survey of estimation techniques, discussion and applications, see Everitt and Hand (1981), Maclachlan and Krishnan (1997) and Maclachlan and Peel (2000). Recently, AL-Hussaini and Sultan (2001), Jiang et al. (2001) and Sultan et al. (2007) reviewed some properties and estimation techniques of finite mixtures for some lifetime models.

The mixture of two generalized inverse Weibull (MGIW) distributions has density given by

$$f(t;\boldsymbol{\theta}) = \sum_{i=1}^{2} p_i f_i(t;\boldsymbol{\theta}_i),$$

where $\sum_{i=1}^{2} p_i = 1$, $\boldsymbol{\theta} = (\boldsymbol{\theta}_1^T, \boldsymbol{\theta}_2^T)^T$, $\boldsymbol{\theta}_1 = (\alpha_1, \beta_1, \gamma_1)^T$, $\boldsymbol{\theta}_2 = (\alpha_2, \beta_2, \gamma_2)^T$ and $f_i(t;\boldsymbol{\theta}_i)$ is the density function of the ith component given by

$$f_i(t;\theta_i) = \gamma_i \beta_i \alpha_i^{\beta_i} t^{-(\beta_i+1)} \exp\left[-\gamma_i \left(\frac{\alpha_i}{t}\right)^{\beta_i}\right], \quad t, \alpha_i, \beta_i, \gamma_i > 0, \quad i = 1, 2.$$

Different shapes of the MGIW density are shown in Figs. 3 and 4. The cdf of the MGIW distribution can be expressed as

$$F(t;\boldsymbol{\theta}) = \sum_{i=1}^{2} p_i F_i(t;\boldsymbol{\theta}_i),$$

where $F_i(t;\theta_i) = \exp\left[-\gamma_i \left(\frac{\alpha_i}{t}\right)^{\beta_i}\right]$ is the cdf of the ith GIW distribution for $i = 1, 2$.

6.1 Properties

Here, we derive some properties of the MGIW distribution by extending the corresponding results for the GIW distribution. The kth moment of the random variable T following the MGIW distribution (5) is given by

$$E(T^k) = \sum_{i=1}^{2} p_i \gamma_i^{\frac{k}{\beta_i}} \alpha_i^k \Gamma(1 - k\beta_i^{-1}).$$

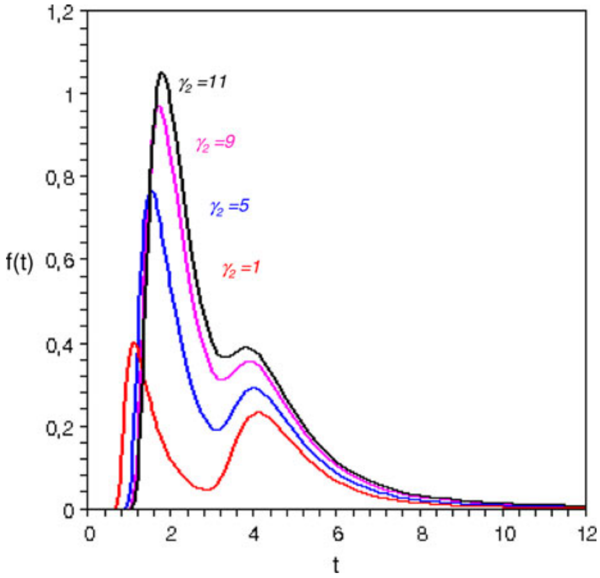

Fig. 3 Plots of the MGIW density for some values of γ_2 with $\alpha_1 = 1, \beta_1 = 2, \gamma_1 = 3, \alpha_2 = 5$ and $\beta_2 = 6$

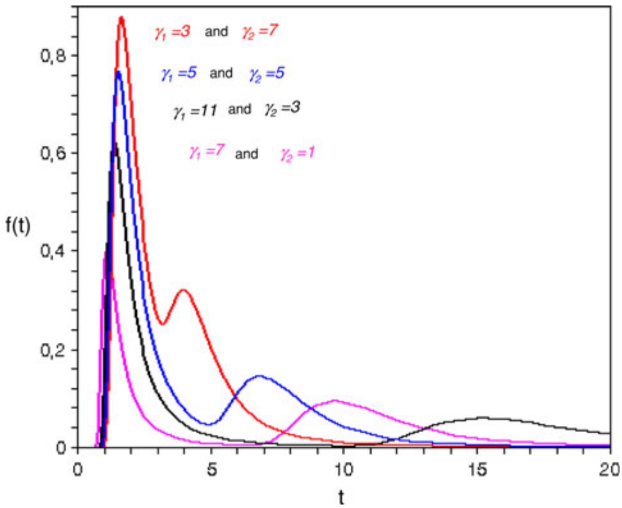

Fig. 4 Plots of the MGIW density for some values of γ_1 and γ_2 with $\alpha_1 = 1, \beta_1 = 2, \alpha_2 = 5$ and $\beta_2 = 6$

The corresponding survival and the hazard functions are

$$S(t) = \sum_{i=1}^{2} p_i \left\{ 1 - \exp\left[-\gamma_i \left(\frac{\alpha_i}{t} \right)^{\beta_i} \right] \right\}$$

and

$$h(t) = \frac{\sum_{i=1}^{2} p_i \gamma_i \beta_i \alpha_i^{\beta_i} t^{-(\beta_i+1)} \exp\left[-\gamma_i \left(\frac{\alpha_i}{t}\right)^{\beta_i}\right]}{\sum_{i=1}^{2} p_i \left\{1 - \exp\left[-\gamma_i \left(\frac{\alpha_i}{t}\right)^{\beta_i}\right]\right\}},$$

respectively. If $\gamma_i = 1$ and $\alpha_i = 1/\delta_i$, the results derived by Sultan et al. (2007) are obtained as a special case.

6.2 Identifiability

Let ϕ be a transform associated with each $F_i \in \Phi$ having the domain of definition D_{ϕ_i} with linear map $M : F_i \to \phi_i$. If there exists a total ordering ($\leqslant$) of Φ such that

(i) $F_1 \leqslant F_2$ $(F_1, F_2 \in \Phi) \Rightarrow D_{\phi_1} \subseteq D_{\phi_2}$ and
(ii) for each $F_1 \in \Phi$, there is some $s_1 \in D_{\phi_1}$, $\phi_1(s) \neq 0$ such that $\lim_{s \to s_1} \phi_2(s)/\phi_1(s) = 0$ for $F_1 < F_2$ $(F_1, F_2 \in \Phi)$,

the class Λ of all finite mixing distributions is identifiable relative to Φ (Chandra 1977).
Using Chandra's approach (assumptions (i) and (ii)) and the results by Sultan et al. (2007), we now prove the following proposition.

Proposition *The class of all finite mixing distributions relative to the GIW distribution is identifiable.*

Proof If T_i is a random variable with the GIW distribution, the sth moment is

$$\phi_i(s) = E(T_i^s) = \gamma_i^{s/\beta_i} \alpha_i^{-s} \Gamma(1 - s\beta_i^{-1}), \quad i = 1, 2.$$

From the cdf of T_i we have

$$F_1 < F_2 \quad \text{when} \quad \beta_1 = \beta_2, \quad \gamma_1 = \gamma_2 \quad \text{and} \quad \alpha_1 < \alpha_2 \tag{11}$$

and

$$F_1 < F_2 \quad \text{when} \quad \alpha_1 = \alpha_2 > \frac{1}{t}, \quad \gamma_1 = \gamma_2 \quad \text{and} \quad \beta_1 < \beta_2. \tag{12}$$

Let $D_{\phi_1}(s) = (-\infty, \beta_1)$, $D_{\phi_2}(s) = (-\infty, \beta_2)$ and $s = \beta_1$. From the facts (11) and (12), we have

$$\lim_{s \to \beta_1} \phi_1(s) = \gamma_1^{\beta_1/\beta_1} \alpha_1^{-\beta_1} \Gamma(1 - \beta_1 \beta_1^{-1}) = \Gamma(0+) = \infty. \tag{13}$$

When $\alpha_1 = \alpha_2 > 1/t$,$\gamma_1 = \gamma_2$ and $\beta_1 < \beta_2$, we obtain

$$\lim_{s \to \beta_1} \phi_2(s) = \gamma_1^{\beta_1/\beta_2} \alpha_1^{-\beta_1} \Gamma(1 - \beta_1 \beta_2^{-1}) > 0. \tag{14}$$

Finally, the results (13) and (14) lead to

$$\lim_{s\to\beta_1} \frac{\phi_2(s)}{\phi_1(s)} = 0,$$

and then the identifiability is proved. □

It is a much more difficult problem to study the identifiability of a mixture of $n \geq 3$ distributions. Lack of identifiability gives no guarantee of convergence to the true values of parameters and therefore usually gives rise to confusing results.

7 The log-generalized inverse Weibull distribution

Let T be a random variable having the GIW density function (3). The random variable $Y = \log(T)$ has a log-generalized inverse Weibull (LGIW) distribution whose density function, parameterized in terms of $\sigma = 1/\beta$ and $\mu = \log(\alpha)$, is given by

$$f(y;\gamma,\sigma,\mu)=\frac{\gamma}{\sigma}\exp\left\{-\left(\frac{y-\mu}{\sigma}\right)-\gamma\exp\left[-\left(\frac{y-\mu}{\sigma}\right)\right]\right\}, \quad -\infty < y < \infty \tag{15}$$

where $\gamma > 0$, $\sigma > 0$ and $-\infty < \mu < \infty$. Four density curves (15) for selected parameter values are given in Fig. 5 to show great flexibility of the new parameter γ.

The corresponding survival function to (15) is

$$S(y) = 1 - \exp\left\{-\gamma\exp\left[-\left(\frac{y-\mu}{\sigma}\right)\right]\right\}.$$

We now define the standardized random variable $Z = (Y-\mu)/\sigma$ with density function

$$\pi(z;\gamma) = \gamma\exp[-z-\gamma\exp(-z)], \quad -\infty < z < \infty \quad \text{and} \quad \gamma > 0. \tag{16}$$

The inverse extreme value distribution follows as a special case when $\gamma = 1$.

The kth ordinary moment of the standardized distribution (16), say $\mu'_k = E(Z^k)$, is

$$\mu'_k = \int_{-\infty}^{\infty} z^k\gamma\exp\left[-z-\gamma\exp(-z)\right]\mathrm{d}z. \tag{17}$$

Substituting $u = \exp(-z)$, we can rewrite Eq. 17 as

$$\mu'_k = (-1)^k\gamma\int_0^{\infty}\left[\log(u)\right]^k\exp(-\gamma u)\mathrm{d}u. \tag{18}$$

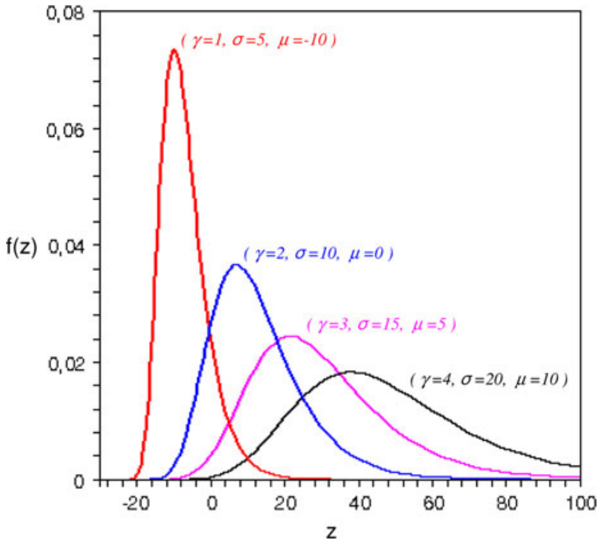

Fig. 5 Plots of the LGIW density for some parameter values

The integral in (18) follows from Prudnikov et al. (1986) and Nadarajah (2006) as

$$\int_0^\infty [\log(u)]^k \exp(-\gamma u) \mathrm{d}u = \left. \frac{\partial^k [\gamma^{-a} \Gamma(a)]}{\partial a^k} \right|_{a=1}.$$

Inserting the last equation in (18), yields the kth moment of Z

$$\mu'_k = (-1)^k \gamma \left. \frac{\partial^k [\gamma^{-a} \Gamma(a)]}{\partial a^k} \right|_{a=1}. \tag{19}$$

The ordinary moments of Y can be obtained via binomial expansion and Eq. 19.

8 The log-generalized inverse Weibull regression model

In many practical applications, the lifetimes t_i are affected by explanatory variables such as the cholesterol level, blood pressure and many others. Let $\mathbf{x}_i = (x_{i1}, \ldots, x_{ip})^T$ be the explanatory variable vector associated with the ith response variable y_i for $i = 1, \ldots, n$. Consider a sample $(y_1, \mathbf{x}_1), \ldots, (y_n, \mathbf{x}_n)$ of n independent observations, where each random response is defined by $y_i = \min\{\log(t_i), \log(c_i)\}$. We assume non-informative censoring and that the observed lifetimes and censoring times are independent.

For the first time, we construct a linear regression model for the response variable y_i based on the LGIW density given by

$$y_i = \mathbf{x}_i^T \boldsymbol{\beta} + \sigma z_i, \quad i = 1, \ldots, n, \tag{20}$$

where the random error z_i follows the distribution (16), $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)^T$, $\sigma > 0$ and $\gamma > 0$ are unknown scalar parameters and $\mathbf{x}_i$ is the explanatory variable vector modeling the location parameter $\mu_i = \mathbf{x}_i^T \boldsymbol{\beta}$. Hence, the location parameter vector $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)^T$ of the LGIW model has a linear structure $\boldsymbol{\mu} = \mathbf{X}\boldsymbol{\beta}$, where $\mathbf{X} = (\mathbf{x}_1, \ldots, \mathbf{x}_n)^T$ is a known model matrix. The log-inverse Weibull (LIW) (or the inverse extreme value) regression model is defined by Eq. 20 with $\gamma = 1$.

Let F and C be the sets of individuals for which y_i is the log-lifetime or log-censoring, respectively. The total log-likelihood function for the model parameters $\boldsymbol{\theta} = (\gamma, \sigma, \boldsymbol{\beta}^T)^T$ can be written from Eqs. 16 and 20 as

$$l(\boldsymbol{\theta}) = r\left[\log(\gamma) - \log(\sigma)\right] - \sum_{i \in F}\left(\frac{y_i - \mathbf{x}_i^T\boldsymbol{\beta}}{\sigma}\right) - \gamma \sum_{i \in F} \exp\left[-\left(\frac{y_i - \mathbf{x}_i^T\boldsymbol{\beta}}{\sigma}\right)\right]$$
$$+ \sum_{i \in C} \log\left[1 - \exp\left\{-\gamma \exp\left[-\left(\frac{y_i - \mathbf{x}_i^T\boldsymbol{\beta}}{\sigma}\right)\right]\right\}\right], \quad (21)$$

where r is the observed number of failures. The MLE $\widehat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}$ can be obtained by maximizing the log-likelihood function (21). We use the matrix programming language Ox (MAXBFGS function) (see Doornik 2007) to compute this estimate. From the fitted model (20), the survival function for y_i can be estimated by

$$S(y_i; \hat{\gamma}, \hat{\sigma}, \widehat{\boldsymbol{\beta}}^T) = 1 - \exp\left\{-\hat{\gamma} \exp\left[-\left(\frac{y_i - \mathbf{x}_i^T\widehat{\boldsymbol{\beta}}}{\hat{\sigma}}\right)\right]\right\}.$$

Under general regularity conditions, the asymptotic distribution of $\sqrt{n}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta})$ is multivariate normal $N_{p+2}(0, K(\boldsymbol{\theta})^{-1})$, where $K(\boldsymbol{\theta})$ is the expected information matrix. The asymptotic covariance matrix $K(\boldsymbol{\theta})^{-1}$ of $\widehat{\boldsymbol{\theta}}$ can be approximated by the inverse of the $(p+2) \times (p+2)$ observed information matrix $J(\boldsymbol{\theta})$ and then the asymptotic inference for the parameter vector $\boldsymbol{\theta}$ can be based on the normal approximation $N_{p+2}(0, J(\boldsymbol{\theta})^{-1})$ for $\widehat{\boldsymbol{\theta}}$. The observed information matrix is

$$J(\boldsymbol{\theta}) = \{-\ddot{L}_{r,s}\} = \begin{pmatrix} -\mathbf{L}_{\gamma\gamma} & -\mathbf{L}_{\gamma\sigma} & -\mathbf{L}_{\gamma\beta_j} \\ \cdot & -\mathbf{L}_{\sigma\sigma} & -\mathbf{L}_{\sigma\beta_j} \\ \cdot & \cdot & -\mathbf{L}_{\beta_j\beta_s} \end{pmatrix},$$

whose elements are given in Appendix 2.

The asymptotic multivariate normal $N_{p+2}(0, J(\boldsymbol{\theta})^{-1})$ distribution can be used to construct approximate confidence regions for some parameters in $\boldsymbol{\theta}$ and for the hazard and survival functions. In fact, an $100(1-\alpha)\%$ asymptotic confidence interval for each parameter θ_r is given by

$$ACI_r = \left(\hat{\theta}_r - z_{\alpha/2}\sqrt{-\widehat{\ddot{L}}^{r,r}}, \hat{\theta}_r + z_{\alpha/2}\sqrt{-\widehat{\ddot{L}}^{r,r}}\right),$$

where $-\widehat{\ddot{L}}^{r,r}$ denotes the rth diagonal element of the inverse of the estimated observed information matrix $J(\widehat{\boldsymbol{\theta}})^{-1}$ and $z_{\alpha/2}$ is the quantile $1-\alpha/2$ of the standard normal distribution. The asymptotic normality is also useful for testing goodness of fit of some sub-models and for comparing some special sub-models using the LR statistic.

The interpretation of the estimated coefficients could be based on the ratio of median times (see Hosmer and Lemeshow 1999) which holds for continuous or categorical explanatory variables. When the explanatory variable is binary (0 or 1), and considering the ratio of median times with $x=1$ in the numerator, if $\hat{\beta}$ is negative (positive), it implies that the individuals with $x=1$ present reduced (increased) median survival time in $[\exp(\hat{\beta})\times 100\%]$ as compared to those individuals in the group with $x=0$, assuming the other explanatory variables fixed.

We are also interested to investigate if the LIW model is a good model to fit the data under investigation. Clearly, the LR statistic can be used to discriminate between the LIW and LGIW models since they are nested models. The hypotheses to be tested in this case are $H_0:\gamma=1$ versus $H_1:\gamma\neq 1$ and the LR statistic reduces to $w=2\{l(\widehat{\boldsymbol{\theta}})-l(\widetilde{\boldsymbol{\theta}})\}$, where $\widetilde{\boldsymbol{\theta}}$ is the MLE of $\boldsymbol{\theta}$ under H_0. The null hypothesis is rejected if $w>\chi^2_{1-\alpha}(1)$, where $\chi^2_{1-\alpha}(1)$ is the quantile of the chi-square distribution with one degree of freedom.

9 Sensitivity analysis

9.1 Global influence

A first tool to perform sensitivity analysis is by means of the global influence starting from case-deletion. Case-deletion is a common approach to study the effect of dropping the ith observation from the data set. The case-deletion model corresponding to model (20) is given by

$$y_l=\mathbf{x}_l^T\boldsymbol{\beta}+\sigma z_l,\quad l=1,\ldots,n,\quad \text{and for}\quad l\neq i. \tag{22}$$

From now on, a quantity with subscript (i) refers to the original quantity with the ith observation deleted. Let $l_{(i)}(\boldsymbol{\theta})$ be the log-likelihood function for $\boldsymbol{\theta}$ under model (22) and $\widehat{\boldsymbol{\theta}}_{(i)}=(\hat{\gamma}_{(i)},\hat{\sigma}_{(i)},\widehat{\boldsymbol{\beta}}^T_{(i)})^T$ be the MLE of $\boldsymbol{\theta}$ obtained by maximizing $l_{(i)}(\boldsymbol{\theta})$. To assess the influence of the ith observation on the unrestricted estimate $\widehat{\boldsymbol{\theta}}=(\hat{\gamma},\hat{\sigma},\widehat{\boldsymbol{\beta}}^T)^T$, the basic idea is to compare the difference between $\widehat{\boldsymbol{\theta}}_{(i)}$ and $\widehat{\boldsymbol{\theta}}$. If deletion of an observation seriously affect the estimates, more attention should be paid to that observation. Hence, if $\widehat{\boldsymbol{\theta}}_{(i)}$ is far from $\widehat{\boldsymbol{\theta}}$, then the ith case is regarded as an influential observation. A first measure of the global influence can be expressed as the standardized norm of $\widehat{\boldsymbol{\theta}}_{(i)}-\widehat{\boldsymbol{\theta}}$ (so-called generalized Cook distance) given by

$$GD_i(\boldsymbol{\theta})=(\widehat{\boldsymbol{\theta}}_{(i)}-\widehat{\boldsymbol{\theta}})^T J(\widehat{\boldsymbol{\theta}})^{-1}(\widehat{\boldsymbol{\theta}}_{(i)}-\widehat{\boldsymbol{\theta}}).$$

Another alternative is to assess the values $GD_i(\boldsymbol{\beta})$ and $GD_i(\gamma,\sigma)$ which reveal the impact of the ith observation on the estimates of $\boldsymbol{\beta}$ and (γ,σ), respectively. Another well-know measure of the difference between $\widehat{\boldsymbol{\theta}}_{(i)}$ and $\widehat{\boldsymbol{\theta}}$ is the likelihood distance

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

Further, we can also compute $\hat{\beta}_j - \hat{\beta}_{j(i)}$ $(j = 1, \ldots, p)$ to obtain the difference between $\widehat{\boldsymbol{\beta}}$ and $\widehat{\boldsymbol{\beta}}_{(i)}$. Alternative global influence measures are possible. We can develop the behavior of a test statistic, such as the Wald test for explanatory variable or censoring effect, under a case deletion scheme.

9.2 Local influence

As a second tool for sensitivity analysis, we now describe the local influence method for the LGIW regression model with censored data. Local influence calculation can be carried out for the model (20). If likelihood displacement $LD(\boldsymbol{\omega}) = 2\{l(\widehat{\boldsymbol{\theta}}) - l(\widehat{\boldsymbol{\theta}}_{\boldsymbol{\omega}})\}$ is used, where $\widehat{\boldsymbol{\theta}}_{\boldsymbol{\omega}}$ is the MLE under the perturbed model, the normal curvature for $\boldsymbol{\theta}$ at the direction $\mathbf{d}$, $\| \mathbf{d} \| = 1$, is given by $C_{\mathbf{d}}(\boldsymbol{\theta}) = 2|\mathbf{d}^T \boldsymbol{\Delta}^T J(\boldsymbol{\theta})^{-1} \boldsymbol{\Delta}\mathbf{d}|$ (see Cook 1986). Here, $\boldsymbol{\Delta}$ is a $(p+2) \times n$ matrix depending on the perturbation scheme, whose elements are given by $\Delta_{ji} = \partial^2 l(\boldsymbol{\theta}|\boldsymbol{\omega})/\partial\theta_j \partial\omega_i$, $i = 1, \ldots, n$ and $j = 1, \ldots, p+2$ evaluated at $\widehat{\boldsymbol{\theta}}$ and $\boldsymbol{\omega}_0$, and $\boldsymbol{\omega}_0$ is the no perturbation vector. For the LGIW model, the elements of $J(\boldsymbol{\theta})$ are given in Appendix 2. We can calculate normal curvatures $C_{\mathbf{d}}(\boldsymbol{\theta})$, $C_{\mathbf{d}}(\gamma)$, $C_{\mathbf{d}}(\sigma)$ and $C_{\mathbf{d}}(\boldsymbol{\beta})$ to perform 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 $\mathbf{B} = \boldsymbol{\Delta}^T J(\boldsymbol{\theta})^{-1} \boldsymbol{\Delta}$ and the index plots of $C_{\mathbf{d}_i}(\boldsymbol{\theta})$, $C_{\mathbf{d}_i}(\gamma)$, $C_{\mathbf{d}_i}(\sigma)$ and $C_{\mathbf{d}_i}(\boldsymbol{\beta})$, so-called the total local influence (see, for example, Lesaffre and Verbeke 1998), where $\mathbf{d}_i$ is an $n{\times}1$ vector of zeros with one at the ith position. Hence, the curvature at direction $\mathbf{d}_i$ has the form $C_i = 2|\boldsymbol{\Delta}_i^T J(\boldsymbol{\theta})^{-1} \boldsymbol{\Delta}_i|$, where $\boldsymbol{\Delta}_i^T$ is the ith row of $\boldsymbol{\Delta}$. It is usual to point out those cases such that

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

9.3 Curvature calculations

Here, under three perturbation schemes, we calculate the matrix

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

from the log-likelihood function (21). Let $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_n)^T$ be the vector of weights.

9.3.1 Case-weight perturbation

In this case, the log-likelihood function has the form

$$l(\boldsymbol{\theta}|\boldsymbol{\omega}) = \left[\log(\gamma) - \log(\sigma)\right]\sum_{i\in F}\omega_i - \sum_{i\in F}\omega_i z_i - \gamma\sum_{i\in F}\omega_i \exp(-z_i) + \sum_{i\in C}\omega_i \log\{1 - \exp[-\gamma\exp(-z_i)]\},$$

where $z_i = (y_i - \mathbf{x}_i^T\boldsymbol{\beta})/\sigma$, $0 \le \omega_i \le 1$ and $\boldsymbol{\omega} = (1,\ldots,1)^T$. From now on, we take $\boldsymbol{\Delta} = (\boldsymbol{\Delta}_1,\ldots,\boldsymbol{\Delta}_{p+2})^T$, $\hat{z}_i = (y_i - \mathbf{x}_i^T\hat{\boldsymbol{\beta}})/\hat{\sigma}$ and $\hat{h}_i = \exp[-\hat{\gamma}\exp(-\hat{z}_i)]$.

The elements of the vector $\boldsymbol{\Delta}_1$ have the form

$$\Delta_{1i} = \begin{cases} \hat{\gamma}^{-1} - \exp(-\hat{z}_i) & \text{if } i \in F \\ \hat{h}_i\exp(-\hat{z}_i)(1-\hat{h}_i)^{-1} & \text{if } i \in C. \end{cases}$$

On the other hand, the elements of the vector $\boldsymbol{\Delta}_2$ are

$$\Delta_{2i} = \begin{cases} \hat{\sigma}^{-1}\{-1 + \hat{z}_i[-1 + \hat{\gamma}\exp(-\hat{z}_i)]\} & \text{if } i \in F \\ \hat{\gamma}\hat{\sigma}^{-1}\hat{z}_i\hat{h}_i\exp(-\hat{z}_i)(1-\hat{h}_i)^{-1} & \text{if } i \in C. \end{cases}$$

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

$$\Delta_{ji} = \begin{cases} \hat{\sigma}^{-1}x_{ij}\left[1 - \hat{\gamma}\exp(\hat{z}_i)\right] & \text{if } i \in F \\ \hat{\gamma}\hat{\sigma}^{-1}x_{ij}\hat{h}_i\exp(-\hat{z}_i)(1-\hat{h}_i)^{-1} & \text{if } i \in C. \end{cases}$$

9.3.2 Response perturbation

We assume that each y_i is perturbed as $y_{iw} = y_i + \omega_i S_y$, where S_y is a scale factor such as the standard deviation of Y and $\omega_i \in \Re$. The perturbed log-likelihood function becomes

$$l(\boldsymbol{\theta}|\boldsymbol{\omega}) = r\left[\log(\gamma) - \log(\sigma)\right] - \sum_{i\in F} z_i^* - \gamma\sum_{i\in F}\exp(-z_i^*) + \sum_{i\in C}\log\left\{1 - \exp[-\gamma\exp(-z_i^*)]\right\},$$

where $z_i^* = \left[(y_i + \omega_i S_y) - \mathbf{x}_i^T\boldsymbol{\beta}\right]/\sigma$. In addition, the elements of the vector $\boldsymbol{\Delta}_1$ take the form

$$\Delta_{1i} = \begin{cases} \hat{\sigma}^{-1}S_y\exp(\hat{z}_i) & \text{if } i \in F \\ \hat{\sigma}^{-1}S_y\hat{h}_i\exp(-\hat{z}_i)\{[\hat{\gamma}\exp(-\hat{z}_i) - 1]\left(1-\hat{h}_i\right) + \hat{h}_i\exp(-\hat{z}_i)\}(1-\hat{h}_i)^{-2} & \text{if } i \in C. \end{cases}$$

On the other hand, the elements of the vector $\mathbf{\Delta}_2$ are

$$\Delta_{2i} = \begin{cases} \hat{\sigma}^{-2} S_y[1 + \hat{\gamma}\hat{z}_i \exp(-\hat{z}_i) - \exp(-\hat{z}_i)] & \text{if } i \in F \\ \hat{\sigma}^{-2}\hat{\gamma} S_y \hat{h}_i \exp(-\hat{z}_i)\{\hat{z}_i \left[\hat{\gamma}\exp(-\hat{z}_i) - 1\right] + 1\}(1-\hat{h}_i)^{-1} & \\ +\hat{\sigma}^{-2}\hat{\gamma} S_y \hat{z}_i \hat{h}_i^2 \exp(-2\hat{z}_i)(1-\hat{h}_i)^{-2} & \text{if } i \in C. \end{cases}$$

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

$$\Delta_{ji} = \begin{cases} \hat{\gamma}\hat{\sigma}^{-2} S_y x_{ij} \exp(-\hat{z}_i) & \text{if } i \in F \\ -\hat{\gamma}\hat{h}_i \exp(-\hat{z}_i)\{-\hat{\sigma}^{-2} S_y x_{ij}[\hat{\gamma}\exp(-\hat{z}_i) - 1]\}(1-\hat{h}_i)^{-1} & \\ +\hat{\gamma}\hat{\sigma}^{-2} S_y x_{ij} \hat{h}_i^2 \exp(-2\hat{z}_i)(1-\hat{h}_i)^{-2} & \text{if } i \in C. \end{cases}$$

9.3.3 Explanatory variable perturbation

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

$$l(\boldsymbol{\theta}|\boldsymbol{\omega}) = r\left[\log(\gamma) - \log(\sigma)\right] - \sum_{i\in F} z_i^{**} - \gamma\sum_{i\in F}\exp(-z_i^{**}) + \sum_{i\in C}\log\left\{1 - \exp[-\gamma\exp(-z_i^{**})]\right\},$$

where $z_i^{**} = (y_i - x_i^{*T}\boldsymbol{\beta})/\sigma$ and $\mathbf{x}_i^{*T}\boldsymbol{\beta} = \beta_1 + \beta_2 x_{i2} + \cdots + \beta_t(x_{it} + \omega_i S_x) + \cdots + \beta_p x_{ip}$.

In addition, the elements of the vector $\mathbf{\Delta}_1$ can be expressed as

$$\Delta_{1i} = \begin{cases} -\hat{\sigma}^{-1} S_x \hat{\beta}_t \exp(-\hat{z}_i) & \text{if } i \in F \\ -\hat{\sigma}^{-1} S_x \hat{\beta}_t \hat{h}_i \exp(-\hat{z}_i)\{[\hat{\gamma}\exp(-\hat{z}_i) - 1](1-\hat{h}_i) + \hat{h}_i\exp(-\hat{z}_i)\}(1-\hat{h}_i)^{-2} & \text{if } i \in C, \end{cases}$$

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

$$\Delta_{2i} = \begin{cases} \hat{\sigma}^{-2} S_x \hat{\beta}_t\{-1 - \hat{\gamma}\hat{z}_i \exp(-\hat{z}_i) + \exp(-\hat{z}_i)\} & \text{if } i \in F \\ -\hat{\gamma}\hat{\sigma}^{-2} S_x \hat{\beta}_t \hat{h}_i \exp(-\hat{z}_i)\{\hat{z}_i[\hat{\gamma}\exp(-\hat{z}_i) - 1] + 1\}(1-\hat{h}_i)^{-1} & \\ -\hat{\gamma}\hat{\sigma}^{-2} S_x \hat{\beta}_t \hat{z}_i \hat{h}_i^2 \exp(-2\hat{z}_i)(1-\hat{h}_i)^{-2} & \text{if } i \in C, \end{cases}$$

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

$$\Delta_{ji} = \begin{cases} -\hat{\gamma}\hat{\sigma}^{-2} S_x \hat{\beta}_t x_{ij} \exp(-\hat{z}_i) & \text{if } i \in F \\ -\hat{\gamma}\hat{\sigma}^{-2} S_x \hat{\beta}_t x_{ij} \hat{h}_i \exp(\hat{z}_i)[\gamma\exp(-\hat{z}_i) + 1](1-\hat{h}_i)^{-1} & \\ -\hat{\gamma}\hat{\sigma}^{-2} S_x \hat{\beta}_t x_{ij} \hat{h}_i^2 \exp(-2\hat{z}_i)(1-\hat{h}_i)^{-2} & \text{if } i \in C, \end{cases}$$

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

$$\Delta_{ti}=\begin{cases}\hat{\sigma}^{-1}S_x\left[1-\hat{\gamma}\hat{\sigma}^{-1}\hat{\beta}_t x_{it}\exp(-\hat{z}_i)-\exp(-\hat{z}_i)\right] & \text{if } i\in F\\ -\hat{\gamma}\hat{\sigma}^{-1}S_x\hat{h}_i\exp(-\hat{z}_i)\{\hat{\sigma}^{-1}\hat{\beta}_t x_{it}[\hat{\gamma}\exp(-\hat{z}_i)-1]-1\}(1-\hat{h}_i)^{-1} & \\ -\hat{\gamma}\hat{\sigma}^{-2}S_x\hat{\beta}_t x_{ij}\hat{h}_i^2\exp(-2\hat{z}_i)(1-\hat{h}_i)^{-2} & \text{if } i\in C.\end{cases}$$

10 Residual analysis

In order to study departures from the error assumption as well as the presence of outliers, we consider the martingale residual proposed by Barlow and Prentice (1988) and some transformations of this residual.

10.1 Martingale residual

This residual was introduced in the counting process (see Fleming and Harrington 1991). For LGIW regression models can be written as

$$r_{M_i}=\begin{cases}1+\log\left\{1-\exp\left[-\hat{\gamma}\exp(-\hat{z}_i)\right]\right\} & \text{if } \quad i\in F\\ \log\left\{1-\exp\left[-\hat{\gamma}\exp(-\hat{z}_i)\right]\right\} & \text{if } \quad i\in C,\end{cases}$$

where $\hat{z}_i=(y_i-\mathbf{x}_i^T\boldsymbol{\beta})/\hat{\sigma}$. The distribution of the residual r_{M_i} is skewed with a maximum value $+1$ and minimum value $-\infty$. Transformations to achieve a more normal shaped form would be more adequate for residual analysis.

10.2 Martingale-type residual

Another possibility is to use a transformation of the martingale residual, based on the deviance residuals for the Cox model with no time-dependent covariates as introduced by Therneau et al. (1990). We use this transformation of the martingale residual in order to have a new residual symmetrically distributed around zero (see Ortega et al. 2008). Hence, a martingale-type residual for the LGIW regression model can be expressed as

$$r_{D_i}=\operatorname{sgn}(r_{M_i})\left\{-2\left[r_{M_i}+\delta_i\log(\delta_i-r_{M_i})\right]\right\}^{1/2}, \tag{23}$$

where r_{M_i} is the martingale residual.

11 Applications

11.1 Vitamin A

We illustrate the proposed model using a data set from a randomized community trial that was designed to evaluate the effect of vitamin A supplementation on diarrheal episodes in 1,207 pre-school children, aged 6–48 months at the baseline, who were

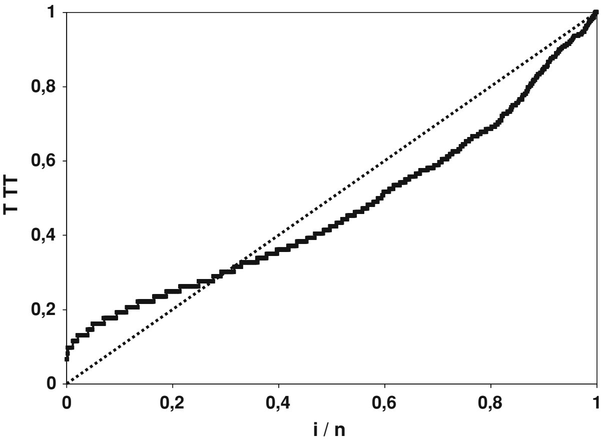

Fig. 6 TTT-plot on vitamin A data

assigned to receive either placebo or vitamin A in a small city of the Northeast of Brazil from December 1990 to December 1991 (see, for example, Barreto et al. 1994).

The vitamin A dosage was 100,000 IU for children younger than 12 months and 200,000 IU for older children, which is the highest dosage guideline established by the World Health Organization (WHO) for the prevention of vitamin A deficiency.

The total time was defined as the time from the first dose of vitamin A until the occurrence of an episode of diarrhea. An episode of diarrhea was defined as a sequence of days with diarrhea and a day with diarrhea was defined when three or more liquid or semi-liquid motions were reported in a 24-h period. The information on the occurrence of diarrhea collected at each visit corresponds to a recall period of 48–72 h. The number of liquid and semi-liquid motions per 24 h was recorded. In many applications, there is a qualitative information about the failure rate function shape which can help in selecting a particular model. In this context, a device called the total time on test (TTT) plot (Aarset 1987) is useful. The TTT plot is obtained by plotting $\mathsf{G}(r/n) = [(\sum_{i=1}^{r} T_{i:n}) + (n-r)T_{r:n}]/(\sum_{i=1}^{n} T_{i:n})$ for $r = 1, \ldots, n$ against r/n (Mudholkar et al. 1996), where $T_{i:n}$, for $i = 1, \ldots, n$, are the order statistics of the sample.

Figure 6 shows the TTT-plot for the current data with the convex and then concave shape, which indicates that the hazard function has a unimodal shape. Hence, the GIW distribution could be an appropriate model for fitting such data. The MLEs (approximate standard errors in parentheses) are: $\hat{\alpha} = 3.3422$ (52.08), $\hat{\beta} = 0.9696$ (0.0245) and $\hat{\gamma} = 8.6448$ (130.66). In order to assess if the model is appropriate, Fig. 7a plots the empirical survival function and the estimated survival function for the GIW distribution. This distribution fits well to the data under analysis. Additionally, the estimated hazard function in Fig. 7b has an upside-down bathtub-shaped curve.

11.2 Multiple myeloma

As a second application, we consider the data set given in Krall et al. (1975) and reported in Lawless (2003, p. 334) connected with survival analysis. This data set has

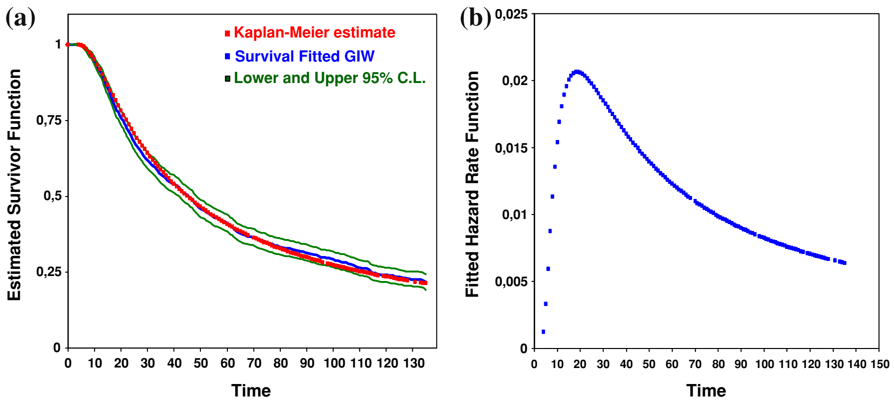

Fig. 7 **a** Estimated survival function and the empirical survival for vitamin A data. **b** Estimated hazard rate function for vitamin A data

recently been analyzed by Jin et al. (2003) and Ghosh and Ghosal (2006) who use a semi-parametric accelerated failure time model. The aim of the study is to relate the logarithm of the survival time (y) for multiple myeloma with a number of prognostic variables for censored data. The data consist of the survival times, in months, of 65 patients who were treated with alkylating agents, of which 48 died during the study and 17 survived. They also include measurements on each patient for the following five predictors: logarithm of a blood urea nitrogen measurement at diagnosis (*urea*) ($x1$); hemoglobin measurement at diagnosis (*hemoglobin*) ($x2$); age at diagnosis (*age*) ($x3$); sex (*sex*) ($x4$, 0 for male and 1 for female) and serum calcium measurement at diagnosis (*serum*) ($x5$).

We fit the LGIW no-intercept regression model

$$y_i = \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4} + \beta_5 x_{i5} + \sigma z_i, \qquad (24)$$

where the errors $z_1, \ldots, z_{65}$ are independent random variables with density function (16).

11.2.1 Maximum likelihood results

Table 1 provides the MLEs of the parameters for the LGIW and LIW regression models fitted to the current data using the subroutine MAXBFGS in Ox. The LR statistic for testing the hypotheses $H_0\colon \gamma = 1$ versus $H_1\colon \gamma \neq 1$, i.e. to compare the LIW and LGIW regression models, is $w = 2\{-76.87 - (-82.69)\} = 11.64$ (P-value <0.01) which yields favorable indications toward to the LGIW model. The LGIW model involves an extra parameter which gives it more flexibility to fit the data, although some parameters may be estimated with less precision, as is the case of β_3 in this example. We note from the fitted LGIW regression model that the age at diagnostic and the sex do not seem to be significant. We expect that the survival time increases as the blood urea nitrogen and the serum calcium measurements at diagnostic decrease, but it increases

Table 1 MLEs of the parameters for the LGIW and LIW regression model fitted to the myeloma data

	LGIW			LIW		
Parameter	Estimate	SE	*P*-value	Estimate	SE	*P*-value
γ	130.25	180.25	–	1	–	–
σ	0.9179	0.1025	–	1.0525	0.1162	–
β_1	–1.5459	0.3605	<0.0001	–1.1777	0.3885	0.0035
β_2	0.1604	0.0456	0.0008	0.2421	0.0502	<0.0001
β_3	0.0006	0.0119	0.9612	0.0309	0.0099	0.0027
β_4	0.3093	0.2604	0.2393	0.8197	0.2659	0.0030
β_5	– 0.1618	0.0607	0.0097	– 0.0605	0.0549	0.2745

Table 2 MLEs of the parameters for the log-Weibull model fitted to the myeloma data

Parameter	Estimate	SE	*P*-value
σ	0.8822	0.0943	–
β_0	4.5642	1.4489	0.0016
β_1	−1.6257	0.5180	0.0017
β_2	0.1181	0.0538	0.0282
β_3	0.0187	0.0140	0.1815
β_4	0.0352	0.2741	0.8979
β_5	−0.1215	0.0883	0.1691

Table 3 Statistics AIC, BIC and CAIC for comparing the LGIW and LIW models

Model	AIC	BIC	CAIC
LGIW	167.8	183.0	169.7
LIW	177.4	190.4	178.8
LW	174.6	189.8	176.6

as the hemoglobin measurement at diagnosis increases. As an alternative analysis, we fit the log-Weibull (LW) regression model (see, for example, Lawless 2003) to these data. The MLEs of the parameters for this model are given in Table 2. Further, we provide in Table 3 the Akaike Information Criterion (AIC), the Bayesian Information Criterion (BIC) and the Consistent Akaike Information Criterion (CAIC) to compare the LGIW, LIW and log-Weibull regression models. The LGIW regression model outperforms the other two models irrespective of the criteria and can be used effectively in the analysis of these data.

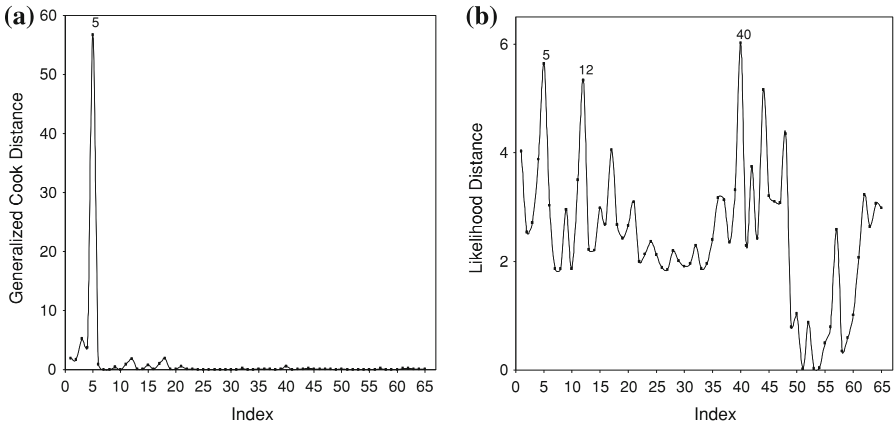

Fig. 8 **a** Index plot of $GD_i(\boldsymbol{\theta})$ (Generalized Cook's distance) on the myeloma data. **b** Index plot of $LD_i(\boldsymbol{\theta})$ (Likelihood distance) on the myeloma data

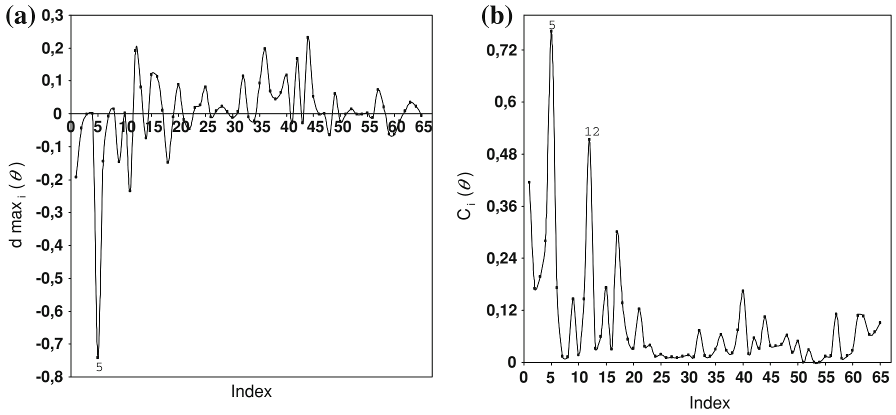

Fig. 9 Index plot for $\boldsymbol{\theta}$ on the myeloma data (case-weight perturbation): **a** $\mathbf{d}_{\max}$ and **b** total local influence

11.2.2 Sensitivity analysis

Global influence analysis. In this section, we compute using Ox the case-deletion measures $GD_i(\boldsymbol{\theta})$ and $LD_i(\boldsymbol{\theta})$ introduced in Sect. 9.1. The influence measure index plots are given in Fig. 8, where we note that the cases ♯5, ♯12 and ♯40 are possible influential observations.

Local and total influence analysis. In this section, we perform an analysis of the local influence for the Golden shiner data set by fitting LGIW regression models.

Case-weight perturbation. We apply the local influence theory developed in Sect. 9.2, where the case-weight perturbation is used and the value $C_{\mathbf{d}_{\max}} = 1.27$ was obtained as a maximum curvature. Figure 9a plots the eigenvector corresponding to $\mathbf{d}_{\max}$, whereas Fig. 9b plots the total influence C_i. The observations ♯5 and ♯12 are very distinguished in relation to the others.

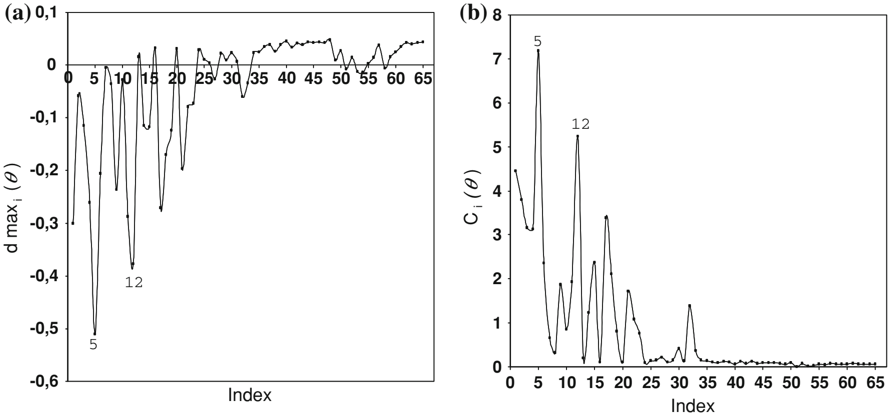

Fig. 10 Index plot for $\boldsymbol{\theta}$ on the myeloma data (response perturbation): **a** $\mathbf{d}_{\max}$ and **b** total local influence

Response variable perturbation. The influence of perturbations on the observed survival times is now analyzed. The value of the maximum curvature was $C_{\mathbf{d}_{\max}} = 14.37$. Figure 10a shows the plot of $\mathbf{d}_{\max}$ versus the observation index, thus indicating that the observations ♯5 and ♯12 are far way from the others. Figure 10b plots the total local influence (C_i) versus the observation index, thus suggesting that the observations ♯5 and ♯12 again stand out.

Explanatory variable perturbation. The perturbation of vectors for each continuous explanatory variable (x_1, x_2, x_3 and x_5) are now investigated. The values for the maximum curvature were $C_{\mathbf{d}_{\max}} = 2.05$, $C_{\mathbf{d}_{\max}} = 1.75$, $C_{\mathbf{d}_{\max}} = 1.05$ and $C_{\mathbf{d}_{\max}} = 1.75$ for x_1, x_2, x_3 and x_5, respectively. Then plots of $\mathbf{d}_{\max}$ and the total local influence C_i against the index of the observations are shown in Figs. 11a–d and 12a–d. We note that the observations ♯5 and ♯12 are very distinguished in relation to the others.

11.2.3 Residual analysis

In order to detect possible outliers in the fitted LGIW regression model, Fig. 13 plots the martingale-type residuals $r_{D\text{mod}_i}$ versus fitted values. We note that almost all observations are within the interval $[-2, 2]$, although a few observations (♯5 and ♯12) appear as possible outliers, thus indicating that the model is well fitted.

11.2.4 Impact of the detected influential observations

From the previous sections we can consider the observations ♯5 and ♯12 as possible influential. The observation ♯5 corresponds to one of the lowest survival times and shows the extent of hemoglobin. The observation ♯12 is also one of the lowest stroke survivors and also shows the extent of the serum calcium. Table 4 gives the relative changes (in percentages) of the parameter estimates defined by $RC_{\boldsymbol{\theta}_j} = [(\widehat{\boldsymbol{\theta}}_j - \widehat{\boldsymbol{\theta}}_j(I))/\widehat{\boldsymbol{\theta}}_j]\times 100$, and the corresponding p-values, where $\widehat{\boldsymbol{\theta}}_j(I)$ denotes the

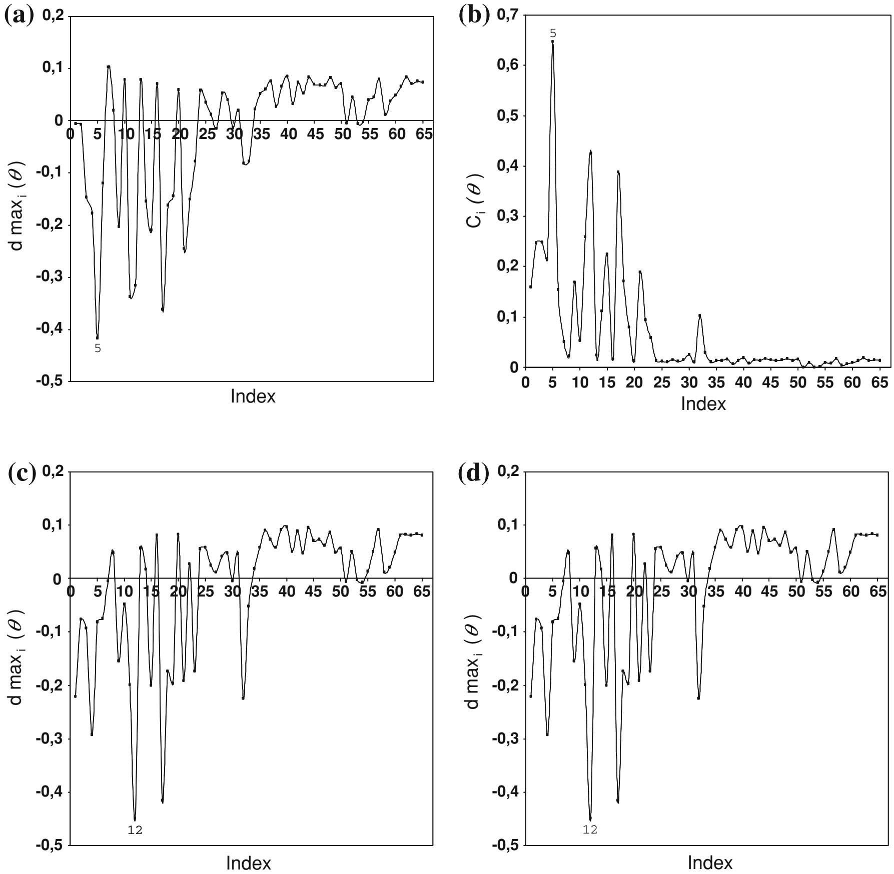

Fig. 11 Index plots for $\boldsymbol{\theta}$ on the myeloma data (explanatory variable perturbation): **a** $\mathbf{d}_{\max}$ (*urea*), **b** total local influence (*urea*), **c** $\mathbf{d}_{\max}$ (*hemoglobin*) and **d** total local influence (*hemoglobin*)

MLE of $\boldsymbol{\theta}_j$ after removing the "set I" of observations in the LGIW model fitted to the current data. We consider the following sets: $I_1 = \{\sharp 5\}$, $I_2 = \{\sharp 12\}$ and $I_3 = \{\sharp 5, \sharp 12\}$.

Table 4 shows that the estimates of the LGIW regression model are not highly sensitive under deletion of the outstanding observations. In general, the significance of the parameter estimates does not change (at the significance level of 5%) after removing the set I. Hence, we do not have inferential changes after removing the observations handed out in the diagnostic plots. The large variations in the parameter estimates occur for the estimates which are not significant, and then they could be removed from the model.

12 Conclusions

We introduce a three parameter lifetime distribution, so-called the generalized inverse Weibull (GIW) distribution, to extend some well-known distributions in the lifetime

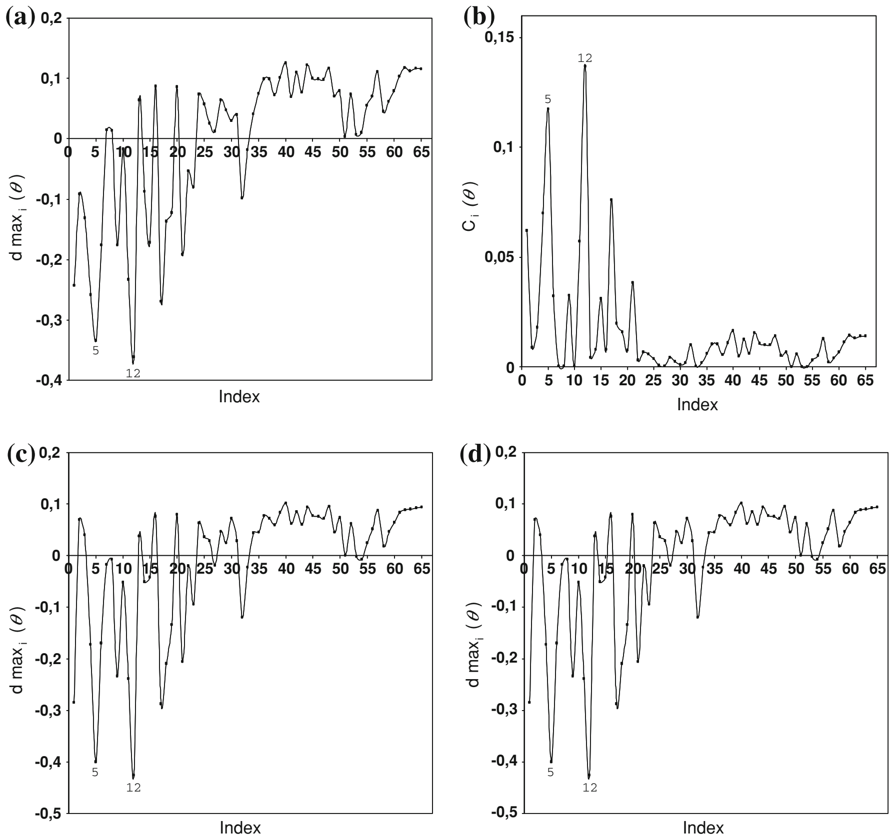

Fig. 12 Index plot for $\boldsymbol{\theta}$ on the myeloma data (explanatory variable perturbation): **a** $\mathbf{d}_{\max}$ (*age*), **b** total local influence (*age*), **c** $\mathbf{d}_{\max}$ (*serum*) and **d** Total local influence (*serum*)

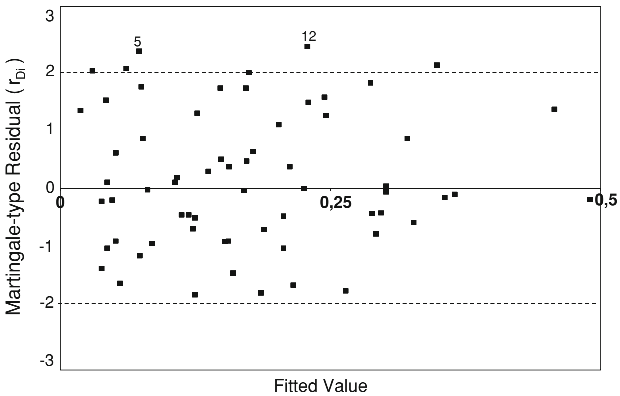

Fig. 13 Plot of the martingale-type residuals (r_{D_i}) versus fitted values for the LGIW model fitted to myeloma data

Table 4 Relative changes [-RC- in %], parameter estimates and their P-values in parentheses for the indicated set

Dropping	$\hat{\gamma}$	$\hat{\sigma}$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$
	–	–	–	–	–	–	–
All observations	130.3	0.918	−1.546	0.160	0.001	0.309	−0.162
	(–)	(–)	(0.000)	(0.000)	(0.961)	(0.239)	(0.010)
	[313]	[3]	[−4]	[30]	[808]	[65]	[−9]
Set I_1	537.7	0.892	−1.603	0.112	−0.004	0.108	−0.177
	(–)	(–)	(0.000)	(0.030)	(0.720)	(0.695)	(0.003)
	[13]	[2]	[3]	[−7]	[−714]	[−48]	[−13]
Set I_2	112.8	0.896	−1.504	0.172	0.005	0.459	−0.182
	(–)	(–)	(0.000)	(0.000)	(0.694)	(0.082)	(0.003)
	[−219]	[5]	[−1]	[20]	[117]	[14]	[−21]
Set I_3	414.9	0.876	−1.555	0.128	0.000	0.265	−0.195
	(–)	(–)	(0.000)	(0.009)	(0.993)	(0.337)	(0.001)

literature. The new distribution is much more flexible than the inverse Weibull distribution and could have increasing, decreasing and unimodal hazard rates. We derive explicit algebraic formula for the rth moment which holds in generality for any parameter values. We discuss maximum likelihood estimation and hypothesis tests for the parameters.

We also propose a log-generalized inverse Weibull (LGIW) regression model for analysis of lifetime censored data. We show how to obtain the maximum likelihood estimates (MLEs) and develop asymptotic tests for the parameters based on the asymptotic distribution of the estimates. We provide applications of influence diagnostics (global, local and total influence) in LGIW regression models to censored data. Further, we discuss the robustness aspects of the MLEs from the fitted LGIW regression model through residual and sensitivity analysis. The practical relevance and applicability of the proposed regression model are demonstrated in two real data analysis.

Appendix 1: The observed information matrix for the GIW model

The elements of the observed information matrix $J(\boldsymbol{\theta})$ for the model parameters (α, β, γ) under censored data are given by

$$\mathbf{L}_{\alpha\alpha}(\boldsymbol{\theta}) = -\frac{r\beta}{\alpha^2} - \gamma\beta(\beta-1)\alpha^{\beta-2}\sum_{i\in F} t_i^{-\beta} + \gamma\beta\alpha^{\beta-2}\sum_{i\in C} t_i^{-\beta}\left(\frac{1-u_i}{u_i}\right)\left[\beta - 1 - \frac{\gamma\beta}{u_i}\left(\frac{\alpha}{t_i}\right)^{\beta}\right];$$

$$\mathbf{L}_{\alpha\beta}(\boldsymbol{\theta}) = \frac{r}{\alpha} - \gamma\alpha^{\beta-1}\sum_{i\in F} t_i^{-\beta}\left[1+\beta\log\left(\frac{\alpha}{t_i}\right)\right] + \gamma\alpha^{\beta-1}\sum_{i\in C} t_i^{-\beta}\left(\frac{1-u_i}{u_i}\right)\left\{1+\beta\log\left(\frac{\alpha}{t_i}\right)\left[1-\frac{\gamma}{u_i}\left(\frac{\alpha}{t_i}\right)^{\beta}\right]\right\};$$

$$\mathbf{L}_{\gamma\beta}(\boldsymbol{\theta}) = -\alpha^{\beta}\sum_{i\in F} t_i^{-\beta}\log\left(\frac{\alpha}{t_i}\right) + \alpha^{\beta}\sum_{i\in C} t_i^{-\beta}\log\left(\frac{\alpha}{t_i}\right)\left(\frac{1-u_i}{u_i}\right)\left[1-\frac{\gamma}{u_i}\left(\frac{\alpha}{t_i}\right)^{\beta}\right];$$

$$\mathbf{L}_{\alpha\gamma}(\boldsymbol{\theta}) = -\beta\alpha^{\beta-1}\sum_{i\in F} t_i^{-\beta} + \beta\alpha^{\beta-1}\sum_{i\in C} t_i^{-\beta}\left(\frac{1-u_i}{u_i}\right)\left[1-\frac{\gamma}{u_i}\left(\frac{\alpha}{t_i}\right)^{\beta}\right];$$

$$\mathbf{L}_{\gamma\gamma}(\boldsymbol{\theta}) = -\frac{r}{\gamma^2} - \alpha^{2\beta}\sum_{i\in C} t_i^{-2\beta}\left(\frac{1-u_i}{u_i^2}\right);$$

and

$$\mathbf{L}_{\beta\beta}(\boldsymbol{\theta}) = -\frac{r}{\beta^2} - \gamma\alpha^{\beta}\sum_{i\in F}\log\left(\frac{\alpha}{t_i}\right)\left[\log(\alpha) - t_i^{-\beta}\log(t_i)\right] + \gamma\alpha^{\beta}\sum_{i\in C}\log\left(\frac{\alpha}{t_i}\right)\left(\frac{1-u_i}{u_i}\right)\left[\log(\alpha) - t_i^{-\beta}\log(t_i) - \frac{t_i^{-\beta}\gamma}{u_i}\left(\frac{\alpha}{t_i}\right)^{\beta}\log\left(\frac{\alpha}{t_i}\right)\right].$$

Here, $u_i = 1-\exp\left[-\gamma\left(\frac{\alpha}{t_i}\right)^{\beta}\right]$.

Appendix 2: The observed information matrix for the LGIW model

We derive the second-order partial derivatives of the log-likelihood function. After some algebraic manipulations, we obtain

$$\mathbf{L}_{\gamma\gamma} = -\frac{r}{\sigma^2} - \sum_{i\in C}\exp(-2z_i)h_i(1-h_i)^{-2};$$

$$\mathbf{L}_{\gamma\sigma} = \sum_{i\in F}\exp(-z_i)(\dot{z}_i)_{\sigma} + \sum_{i\in C} h_i\exp(-z_i)(\dot{z}_i)_{\sigma}[\gamma\exp(-z_i)](1-h_i)^{-1} + \sum_{i\in C} h_i^2\gamma\exp(-2z_i)(\dot{z}_i)_{\sigma}(1-h_i^{-2});$$

$$\begin{aligned}
\mathbf{L}_{\gamma\beta_j} &= \sum_{i\in F} \exp(-z_i)(\dot{z}_i)_{\beta_j} + \sum_{i\in C} h_i \exp(-z_i)(\dot{z}_i)_{\beta_j}[\gamma \exp(-z_i)](1-h_i)^{-1} \\
&\quad + \sum_{i\in C} h_i^2 \gamma \exp(-2z_i)(\dot{z}_i)_{\beta_j}(1-h_i^{-2}); \\
\mathbf{L}_{\sigma\sigma} &= \frac{r}{\sigma^2} + \sum_{i\in F} \left\{-(\ddot{z}_i)_{\sigma\sigma} + \gamma \exp(-z_i)\left[-[(\dot{z}_i)_\sigma]^2 + (\ddot{z}_i)_{\sigma\sigma}\right]\right\} \\
&\quad -\gamma \sum_{i\in C} h_i \exp(-z_i)\left\{[(\dot{z}_i)_\sigma]^2[\gamma \exp(-z_i)-1] + (\ddot{z}_i)_{\sigma\sigma}\right\}(1-h_i)^{-1} \\
&\quad -\gamma \sum_{i\in C} h_i^2 \exp(-2z_i)[(\dot{z}_i)_\sigma]^2(1-h_i)^{-2}; \\
\mathbf{L}_{\sigma\beta_j} &= -\sum_{i\in F} (\ddot{z}_i)_{\beta_j\sigma} + \gamma \sum_{i\in F}\left[-\exp(-z_i)(\dot{z}_i)_{\beta_j}(\dot{z}_i)_\sigma + \exp(-z_i)(\ddot{z}_i)_{\beta_j\sigma}\right] \\
&\quad -\gamma \sum_{i\in C} h_i \exp(-z_i)\left\{(\dot{z}_i)_{\beta_j}(\dot{z}_i)_\sigma[\gamma \exp(-z_i)-1] + (\ddot{z}_i)_{\beta_j\sigma}\right\}(1-h_i)^{-1} \\
&\quad +\gamma \sum_{i\in C} h_i^2 \exp(-2z_i)(\dot{z}_i)_{\beta-j}(\dot{z}_i)_\sigma(1-h_i)^{-2};
\end{aligned}$$

and

$$\begin{aligned}
\mathbf{L}_{\beta_j\beta_s} &= \sum_{i\in F}\left[-\gamma \exp(-z_i)(\dot{z}_i)_{\beta_j}(\dot{z}_i)_{\beta_s}\right] \\
&\quad -\gamma \sum_{i\in C} h_i \exp(-z_i)\left\{(\dot{z}_i)_{\beta_j}(\dot{z}_i)_{\beta_s}[\gamma \exp(-z_i)-1]\right\}(1-h_i)^{-1} \\
&\quad -\gamma \sum_{i\in C} h_i^2 \exp(-2z_i)(\dot{z}_i)_{\beta_j}(\dot{z}_i)_{\beta_s}(1-h_i)^{-2};
\end{aligned}$$

where $h_i = \exp[-\gamma \exp(-z_i)]$, $(\dot{z}_i)_\sigma = -z_i/\sigma$, $(\dot{z}_i)_{\beta_j} = -x_{ij}/\sigma$, $(\dot{z}_i)_{\beta_s} = -x_{is}/\sigma$, $(\ddot{z}_i)_{\sigma\sigma} = 2z_i/\sigma^2$, $(\ddot{z}_i)_{\beta_j\sigma} = x_{ij}/\sigma^2$ and $z_i = (y_i - \mathbf{x}_i^T \boldsymbol{\beta})/\sigma$.

References

Aarset MV (1987) How to identify bathtub hazard rate. IEEE Trans Reliab 36:106–108

AL-Hussaini EK, Sultan KS (2001) Reliability and hazard based on finite mixture models. In: Balakrishnan N, Rao CR (eds) Handbook of statistics, vol 20. Elsevier, Amsterdam pp 139–183

Barakat HM, Abdelkader YH (2004) Computing the moments of order statistics from nonidentical random variables. Stat Methods Appl 13:15–26

Barlow WE, Prentice RL (1988) Residuals for relative risk regression. Biometrika 75:65–74

Barreto ML, Santos LMP, Assis AMO, Araújo MPN, Farenzena GG, Santos PAB, Fiaccone RL (1994) Effect of vitamin A supplementation on diarrhoea and acute lower-respiratory-tract infections in young children in Brazil. Lancet 344:228–231

Chandra S (1977) On the mixtures of probability distributions. Scand J Stat 4:105–112

Cook RD (1986) Assesment of local influence (with discussion). J R Stat Soc B 48:133–169

Doornik J (2007) Ox: an object-oriented matrix programming language. International Thomson Bussiness Press, London

Drapella A (1993) Complementary Weibull distribution: unknown or just forgotten. Qual Reliab Eng Int 9:383–385
Everitt BS, Hand DJ (1981) Finite mixture distributions. Chapman and Hall, London
Fleming TR, Harrington DP (1991) Counting process and survival analysis. Wiley, New York
Ghosh SK, Ghosal S (2006) Semiparametric accelerated failure time models for censored data. In: Upadhyay SK, Singh U, Dey DK (eds) Bayesian statistics and its applications. Anamaya Publishers, New Delhi pp 213–229
Hosmer DW, Lemeshow S (1999) Applied survival analysis. Wiley, New York
Jiang R, Zuo MJ, Li HX (1999) Weibull and Weibull inverse mixture models allowing negative weights. Reliab Eng Syst Saf 66:227–234
Jiang R, Murthy DNP, Ji P (2001) Models involving two inverse Weibull distributions. Reliab Eng Syst Saf 73:73–81
Jin Z, Lin DY, Wei LJ, Ying Z (2003) Rank-based inference for the accelerated failure time model. Biometrika 90:341–353
Keller AZ, Kamath AR (1982) Reliability analysis of CNC machine tools. Reliab Eng 3:449–473
Krall J, Uthoff V, Harley J (1975) A step-up procedure for selecting variables associated with survival. Reliab Eng Syst Saf 73:73–81
Lawless JF (2003) Statistical models and methods for lifetime data. Wiley, New York
Lesaffre E, Verbeke G (1998) Local influence in linear mixed models. Biometrics 54:570–582
Maclachlan GJ, Krishnan T (1997) The EM algorithm and extensions. Wiley, New York
Maclachlan G, Peel D (2000) Finite mixture models. Wiley, New York
Mudholkar GS, Kollia GD (1994) Generalized Weibull family: a structural analysis. Commun Stat Ser A 23:1149–1171
Mudholkar GS, Srivastava DK, Kollia GD (1996) A generalization of the Weibull distribution with application to the analysis of survival data. J Am Stat Assoc 91:1575–1583
Nadarajah S (2006) The exponentiated Gumbel distribution with climate application. Environmetrics 17:13–23
Ortega EMM, Paula GA, Bolfarine H (2008) Deviance residuals in generalized log-gamma regression models with censored observations. J Stat Comput Simul 78:747–764
Prudnikov AP, Brychkov YA, Marichev OI (1986) Integrals and series. Gordon and Breach Science Publishers, New York
Sultan KS, Ismail MA, Al-Moisheer AS (2007) Mixture of two inverse Weibull distributions: properties and estimation. Comput Stat Data Anal 51:5377–5387
Therneau TM, Grambsch PM, Fleming TR (1990) Martingale-based residuals for survival models. Biometrika 77:147–160