UNIVERSIDADE DE SÃO PAULO

Regression Models for Lifetime Data with Mixture of Normal Distributions

Josmar Mazucheli
Jorge A. Achcar
Robert E. Kass

Nº 59

NOTAS

Instituto de Ciências Matemáticas de São Carlos

Instituto de Ciências Matemáticas e de Computação

ISSN - 0103-2577

Regression Models for Lifetime Data with Mixture of Normal Distributions

Josmar Mazucheli
Jorge A. Achcar
Robert E. Kass

Nº 59

NOTAS DO ICMC
Série Estatística

São Carlos
Jan./2000

Regression Models for Lifetime Data with Mixture of Normal Distributions

Josmar Mazucheli*
Universidade Estadual de Maringá
Department of Statistics, DEs/UEM, Maringá, P.R. - Brazil

Jorge A. Achcar
Universidade de São Paulo - ICMC,
C.P. 668, 13560-970, São Carlos, S.P. - Brazil

Robert E. Kass
Carnegie Mellon University
Department of Statistics, Pittsburgh, P.A. - USA

Abstract

In this paper we present a Bayesian analysis of location-scale regression models assuming standard lifetime distribution and an additional error term with a mixture of normal distributions. Assuming uncensored and censored lifetime data, we use Gibbs with Metropolis-Hastings algorithm to get the posterior quantities of interest. The proposed regression model gives a great flexibility to fit lifetime data.

Key words and phrases: location-scale regression models, lifetime data, mixture of normal distributions, interval-censored data, semiparametric models, MCMC algorithms.

*Josmar Mazucheli is a graduate student at COPPE/PO/UFRJ. Partial financial support from CAPES/Brazil.

1 Introduction

A usual distribution considered for modeling survival time data is the Weibull distribution with density,

$$f(t) = \frac{\beta}{\mu}\left(\frac{t}{\mu}\right)^{\beta-1} \exp\left\{-\left(\frac{t}{\mu}\right)^{\beta}\right\}, \tag{1}$$

where $t > 0$, $\beta > 0$ and $\mu > 0$ are the parameters of shape and scale respectively.

Let us assume a reparametrization of the form,

$$Y = \ln(T) = \alpha + \sigma^* Z, \tag{2}$$

where $\alpha = \ln(\mu)$, $\sigma^* = \beta^{-1}$ and Z is a random variable with a Standard Extreme Value density,

$$f(z) = \exp(z - e^z), \quad -\infty < z < \infty. \tag{3}$$

With a covariate vector $\mathbf{x}$ affecting only the location parameter α but not the scale parameter σ^*, we have a location-scale regression model so $\ln(t)$ has constant variance (see for example, Kalbfleisch and Prentice, 1980), given by

$$Y = \ln(T) = \alpha(\mathbf{x}) + \sigma^* Z. \tag{4}$$

A variety of functional forms for $\alpha(\mathbf{x})$ are often employed, but the most useful form is the log-linear one, for which

$$\alpha(\mathbf{x}) = \mathbf{x}\boldsymbol{\beta}, \tag{5}$$

where $x = (x_1, \ldots, x_k)$ is the $(1 \times k)$ vector of regressor or concomitant variables and $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_k)$ is a $(k \times 1)$ vector of regression coefficients.

Observe that if $\sigma^* = 1$ we have an exponential distribution for the lifetime T and other choices for the distribution for the term error Z also could be considered as a Normal distribution, a Logistic distribution or a Log-Gamma distribution (see for example, Lawless, 1982).

A different form for the location-scale regression model (4) is given by the introduction of an additional term ϵ, that is,

$$Y = \ln(T) = \alpha(\mathbf{x}) + \sigma^* Z + \epsilon, \tag{6}$$

where ϵ has a mixture of Normal distributions with mean zero,

$$\phi(\epsilon) = \sum_{j=1}^{J} p_j \phi_j\left(\epsilon | \mu_j, \sigma_j^2\right), \tag{7}$$

with $p_j > 0$, $\sum_{j=1}^{J} p_j = 1$, $\sum_{j=1}^{J} \mu_j p_j = 0$ and

$$\phi_j\left(\epsilon|\mu_j,\sigma_j^2\right) \propto \sigma_j^{-1}\exp\left\{-\frac{1}{2\sigma_j^2}\left(\epsilon-\mu_j\right)^2\right\}. \tag{8}$$

To avoid identificability problems we assume that the means are ordered, that is $(\mu_1 < \mu_2 < \cdots < \mu_J)$ (see for example, Carrol, Roeder and Wasserman, 1996).

Observe that if J is not fixed, that is, the number of mixtures terms are unknown, we have a semi-parametric mixture model. Some recent works address these subject (see for example Polymenis and Titterington (1998); Stephens (1997,1998) or Richardson and Green (1997)).

The model (6) is a random effect regression model where Z could have any standard parametric model used for modelling lifetime data and can be used as an alternative to nonparametric regression models or semi-parametric models as the proportional hazards regression model proposed by Cox (1972).

Assuming uncensored or censored lifetime data, in this paper we consider a Bayesian analysis of the regression model (6) based on Gibbs Sampling with Metropolis-Hastings algorithms (see for example Smith and Roberts, 1993).

The use of MCMC (Markov Chain Monte Carlo) methods has been recently considered in the literature to analyze data with mixture models (see for example, Diebolt and Robert, 1994; Robert, 1996 or Roeder and Wasserman, 1997).

2 Bayesian Formulation Using MCMC Methods

In model (6) we can see that $(\mathbf{y}|\boldsymbol{\beta},\sigma^*,\mathbf{Z})$ has a mixture of Normal distributions. Considering a prior distribution $\pi(\boldsymbol{\theta})$, $\boldsymbol{\theta}=(\boldsymbol{\beta},\sigma^*,\boldsymbol{\mu},\boldsymbol{\sigma},\mathbf{p},\mathbf{Z})$, $\boldsymbol{\mu}=(\mu_1,\ldots,\mu_J)$, $\boldsymbol{\sigma}=(\sigma_1,\ldots,\sigma_J)$ and $\mathbf{p}=(p_1,\ldots,p_J)$, the joint posterior distribution for $\boldsymbol{\theta}$ is given by,

$$\pi(\boldsymbol{\theta}|\mathbf{y}) \propto \pi(\boldsymbol{\theta})\, L(\boldsymbol{\theta}), \tag{9}$$

where the likelihood function for uncensored data is given by

$$L(\boldsymbol{\theta}) = \prod_{i=1}^{n}\left[\sum_{j=1}^{J} p_j\phi_j\left(\epsilon_i|\mu_j,\sigma_j^2\right)\right], \tag{10}$$

where $\phi_j\left(\epsilon_i|\mu_j,\sigma_j^2\right)$ is the Normal density (8) with $\epsilon_i = y_i - \mathbf{x}_i^t\boldsymbol{\beta} - \boldsymbol{\sigma}^* Z_i$.

For interval-censored lifetime data (see for example, Lindsey and Ryan, 1998) considering n individuals in study, let us assume that the first d out of n failure times are interval censored with the i^{th} individual dying between t_{iL} $(y_{iL}=\ln t_{iL})$ and t_{iU} $(y_{iU}=\ln t_{iU})$. The remaining $n-d$ individuals are right censored at y_{iL}.

In this case, the likelihood function is given by

$$L(\boldsymbol{\theta}) = \prod_{i=1}^{d} \left[G\left(\epsilon_{iU}|\mathbf{x}, \boldsymbol{\theta}\right) - G\left(\epsilon_{iL}|\mathbf{x}, \boldsymbol{\theta}\right)\right] \prod_{i=d+1}^{n} \left[1 - G\left(\epsilon_{iL}|\mathbf{x}, \boldsymbol{\theta}\right)\right], \tag{11}$$

where $G\left(\boldsymbol{\epsilon}|\mathbf{x}_i, \boldsymbol{\theta}\right)$ is the cumulative distribution function for ϵ in (6) given by,

$$G\left(\boldsymbol{\epsilon}|\mathbf{x}, \boldsymbol{\theta}\right) = \sum_{j=1}^{J} p_j G_j\left(\boldsymbol{\epsilon}|\mu_j, \sigma_j^2\right),$$

where G_j is the distribution function for the Normal density $\phi_j, j = 1, 2, \ldots, J$.

The likelihood (11) includes the grouped data where a number of individuals die within the same interval because we can set y_{iL} to be identical for several i; similary for y_{iU}.

To simplify the conditional distributions needed for the Gibbs Sampling algorithm, we introduce latent variables (data augmentation technique; see for example Tanner and Wong, 1987) that allow us to write the likelihood as a product of components model for i.i.d. observations.

This is given by augmenting the original data with two classes of latent variables: one is the truncated random variable W and the other is the index variable denoted by V that convert the mixture model to a model of independent components.

If the i^{th} individual is censored between t_{iL} and t_{iU}, we can generate a latent variable from the truncated density $\frac{f(w_i)}{F(t_{iU}) - F(t_{iL})}$, where $f(.)$ is the density for the survival time T and $F(.)$ is the cumulative distribution.

This can be done by setting

$$w_i = F^{-1}\left\{F\left(\epsilon_{iL}|\mathbf{x}_i, \boldsymbol{\theta}\right) + U\left[F\left(\epsilon_{iU}|\mathbf{x}_i, \boldsymbol{\theta}\right) - F\left(\epsilon_{iL}|\mathbf{x}_i, \boldsymbol{\theta}\right)\right]\right\}, \tag{12}$$

where U has a Uniform$(0, 1)$ distribution and F^{-1} is the inverse function of F.

Similarly, if the i^{th} individual is right-censored at t_{iL}, we can generate a latent variable w_i from the truncated density $\frac{f(w_i)}{1 - F(t_{iL})}$ by setting

$$w_i = F^{-1}\left\{F\left(t_{iL}|\mathbf{x}_i, \boldsymbol{\theta}\right) + U\left[1 - F\left(\epsilon_{iL}|\mathbf{x}_i, \boldsymbol{\theta}\right)\right]\right\}. \tag{13}$$

In this way, the likelihood function is given by

$$L(\boldsymbol{\theta}) = \prod_{i=1}^{n} \left[\sum_{j=1}^{J} p_j \phi_j\left(\epsilon_i|\mu_j, \sigma_j^2\right)\right], \tag{14}$$

where now, $\epsilon_i = w_i - \mathbf{x}_i^t\boldsymbol{\beta} - \sigma^* Z_i$ and $w_i = \ln(t_i)$ for uncensored lifetime.

We can observe that for any other censoring mechanism (left or right censoring with uncensored data; or a combination of both) we get similar latent variable for the censored data.

Assuming only $J = 2$ Normal distributions in the mixture model, the other class of latent variables is given by $\mathbf{v}_i = (v_{i1}, v_{i2}), i = 1, 2, \ldots, n$ where $v_{i1}|\boldsymbol{\theta}, w_i, \mathbf{x}_i \sim b(1, h_{i1})$, a Bernoulli distribution with h_{i1} given by

$$h_{i1} = \frac{p_1 \phi_1 (\epsilon_i|\mu_1, \sigma_1^2)}{\sum_{j=1}^{2} p_j \phi_j \left(\epsilon_i|\mu_j, \sigma_j^2\right)}, \tag{15}$$

that is, $\pi(\mathbf{v}_i) \propto h_{i1}^{v_{i1}} (1 - h_{i1})^{v_{i2}}$ where $v_{i2} = 1$ with probability h_{i1} ($v_{i1} = 0$ with probability $1 - h_{i1}$) and $v_{i1} + v_{i2} = 1$.

Thus,

$$\Pi(\mathbf{v}_1, \ldots, \mathbf{v}_n) \propto \frac{\prod_{i=1}^{n} \prod_{j=1}^{2} \left[p_j \phi_j \left(\epsilon_i|\mu_j, \sigma_j^2\right)\right]^{\mathbf{v}_{ij}}}{\prod_{i=1}^{n} \left[\sum_{j=1}^{2} p_j \phi_j \left(\epsilon_i|\mu_j, \sigma_j^2\right)\right]}. \tag{16}$$

Combining (14) and (16), the joint posterior density for $\boldsymbol{\theta}$ is,

$$\pi(\boldsymbol{\theta}|\mathbf{w}, \mathbf{v}, \mathbf{x}) \propto \pi(\boldsymbol{\theta}) \left\{\prod_{i=1}^{n} \prod_{j=1}^{2} \left[p_j \phi_j \left(\epsilon_i|\mu_j, \sigma_j^2\right)\right]^{\mathbf{v}_{ij}}\right\}. \tag{17}$$

Similar posterior density is obtained when $J > 2$.

To generate samples of the joint posterior distribution (17), use the Gibbs Sampling algorithm. Starting with the initial values $\boldsymbol{\theta}^{(0)} = \left(\theta_1^{(0)}, \ldots, \theta_p^{(0)}\right)$, follow the steps:

i) Generate samples $\mathbf{w}^{(1)} = \left(w_1^{(1)}, \ldots, w_n^{(1)}\right)$ from (12) and (13);

ii) Generate samples $\mathbf{v}^{(1)} = \left(v_1^{(1)}, \ldots, v_n^{(1)}\right)$ where $\mathbf{v}_i^{(1)} = \left(v_{i1}^{(1)}, v_{i2}^{(1)}\right)$ from a Bernoulli distribution $Be(1, h_{i1})$ where h_{i1} is given in (15);

iii) Generate a samples of $\boldsymbol{\theta}$ from the conditional distributions,

$\pi\left(\theta_1|\boldsymbol{\theta}_{-\theta_1}^{(0)}, \mathbf{v}^{(1)}, \mathbf{w}^{(1)}, \mathbf{Z}^{(1)}, \mathbf{x}\right)$,

$\ldots, \pi\left(\theta_J|\boldsymbol{\theta}_{-\theta_J}^{(1)}, \mathbf{v}^{(1)}, \mathbf{w}^{(1)}, \mathbf{Z}^{(1)}, \mathbf{x}\right)$, where $\boldsymbol{\theta}_{-\theta_i} = (\theta_1, \ldots, \theta_{i-1}, \theta_{i+1}, \ldots \theta_J)$.

3 Conditional Distributions for the Gibbs Sampling Algorithm

Consider in the regression model (6), a mixture of two Normal distributions for the error term ϵ and a vector of covariates $\mathbf{x}^t = (x_1, x_2, \ldots, x_k)$ such that $\mathbf{x}^t\boldsymbol{\beta} = \beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k$.

Assuming prior independence let us assume the following prior densities:

i) $\pi(\beta_0, \beta_1, \ldots, \beta_k) \propto$ constant;

ii) $p_1 \sim B(a,b)$, a,b known;

iii) $\sigma_j^2 \sim I\Gamma(c_j,d_j)$, c_j,d_j known, $j=1,2$; (18)

iv) $\sigma^* \sim \Gamma(e,f)$, e,f known;

v) $\mu_j \sim N(g,h^2)$, g,h known;

vi) $Z_i \sim$ Extreme-Value$(0,1)$, $i=1,2,\ldots,n$,

where $B(a,b)$ denotes a Beta distribution with mean $\frac{a}{a+b}$ and variance $\frac{ab}{[(a+b)^2(a+b+1)]}$; $I\Gamma(a,b)$ denotes an Inverse Gamma distribution with mean $\frac{b}{a-1}$ and variance $\frac{b^2}{[(a-1)^2(a-2)]}$; $\Gamma(a,b)$ denotes a Gamma distribution with mean $\frac{a}{b}$ and variance $\frac{a}{b^2}$; $N(\mu,\sigma^2)$ denotes a Normal distribution with mean μ and variance σ^2.

We also could consider a prior distribution for the hiperparameters above.

With the introduction of latent variables $\mathbf{w}$ and $\mathbf{v}$ (see section 2), where $\mathbf{w}=\mathbf{y}$ for the uncensored data, we have from (17) and (18) the conditional distributions needed for the Gibbs-Sampling algorithm given by:

i) $(\beta_0|\boldsymbol{\theta}_{-\beta_0},\mathbf{v},\mathbf{w},\mathbf{Z},\mathbf{x}) \sim N\left\{\frac{\sigma_1^2\sum_{i=1}^n v_{i2}(r_i-\mu_2)+\sigma_2^2\sum_{i=1}^n v_{i1}(r_i-\mu_1)}{n_2\sigma_1^2+n_1\sigma_2^2}, \frac{\sigma_1^2\sigma_2^2}{n_2\sigma_1^2+n_1\sigma_2^2}\right\}$ where $n_1=\sum_{i=1}^n v_{i1}$, $n_2=n-n_1$ and $r_i=w_i-\sum_{s=1}^k \beta_s x_{si}-\sigma^* Z_i$, $i=1,\ldots,n$;

ii) $\left(\beta_{(s)}|\boldsymbol{\theta}_{-\beta_{(s)}},\mathbf{v},\mathbf{w},\mathbf{Z},\mathbf{x}\right) \sim N\left\{\frac{\sigma_1^2\sum_{i=1}^n v_{i2}x_{si}(r_{si}-\mu_2)+\sigma_2^2\sum_{i=1}^n v_{i1}x_{si}(r_{si}-\mu_1)}{a_{2s}\sigma_1^2+a_{1s}\sigma_2^2}, \frac{\sigma_1^2\sigma_2^2}{a_{2s}\sigma_1^2+a_{1s}\sigma_2^2}\right\}$ where $s=1,\ldots,k$, $a_{js}=\sum_{i=1}^n v_{ij}x_{si}^2$, $j=1,2$, $r_{si}=w_i-\beta_0-\sum_{l=1}^k \beta_l x_{li}-\sigma^* Z_i$, and $\beta_{(s)}=(\beta_1,\ldots,\beta_{s-1},\beta_{s+1},\ldots,\beta_k)$;

iii) $(p_1|\boldsymbol{\theta}_{-p_1},\mathbf{v},\mathbf{w},\mathbf{Z},\mathbf{x}) \sim B(a+n_1,b+n_2)$; (19)

iv) $(\mu_1|\boldsymbol{\theta}_{-\mu_1},\mathbf{v},\mathbf{w},\mathbf{Z},\mathbf{x}) \sim N\left(a_1-\frac{p_1b_1\sum_{j=1}^2 p_j a_j}{\sum_{j=1}^2 p_j^2 b_j}, b_1\left(1-\frac{p_1^2 b_1}{\sum_{j=1}^2 p_j^2 b_j}\right)\right)$ where $a_j=\frac{g\sigma_j^2+h^2\sum_{i=1}^n v_{ij}r_i}{\sigma_j^2+n_j h^2}$, $b_j=\frac{h^2\sigma_j^2}{\sigma_j^2+n_j h^2}$ and $r_i=w_i-\beta_0-\sum_{j=1}^k \beta_{ji}-\sigma^* Z_i$ (observe that we are also conditioning on the constraint $\sum_{j=1}^2 \mu_j p_j=0$, see for example Carroll, Roeder and Wasserman (1996));

v) $\left(\sigma_j^2|\boldsymbol{\theta}_{-\sigma_j^2},\mathbf{v},\mathbf{w},\mathbf{Z},\mathbf{x}\right) \sim I\Gamma\left(c_j+\frac{n_j}{2}; d_j+\frac{1}{2}\sum_{i=1}^n v_{ij}^2\left(w_i-\beta_0-\beta_1 x_i+\boldsymbol{\sigma}^* Z_i-\mu_j\right)^2\right)$;

vi) $\pi(\sigma^*|\boldsymbol{\theta}_{-\sigma^*},\mathbf{v},\mathbf{w},\mathbf{x}) \propto \sigma^{*e-1}e^{-f\sigma^*}\psi_1(\boldsymbol{\theta})$ where

$\psi_1(\boldsymbol{\theta})=\exp\left\{-\frac{1}{2}\sum_{i=1}^n\sum_{j=1}^2 v_{ij}\left(\frac{w_i-\beta_0-\sum_{s=1}^k \beta_s x_{si}+\sigma^* Z_i-\mu_j}{\sigma_j}\right)^2\right\}$;

vii) $\pi\left(Z_i|\boldsymbol{\theta}_{-z_i^*},\mathbf{v},\mathbf{w},\mathbf{Z},\mathbf{x}\right) \propto \exp\left(Z_i-e^{Z_i}\right)\psi_2(\boldsymbol{\theta})$ where

$\psi_2(\boldsymbol{\theta})=\exp\left\{-\frac{1}{2}\sum_{j=1}^2 v_{ij}\left(\frac{w_i-\beta_0-\sum_{s=1}^k \beta_s x_{si}+\sigma^* Z_i-\mu_j}{\sigma_j}\right)^2\right\}$ for $i=1,\ldots,n$.

Observe that the variables $\sigma^*, Z_1, \ldots, Z_n$ should be generated using the Metropolis-Hastings algorithm.

4 Some Examples

4.1 An Example With Uncensored Data

In table 1 we have the observed survival times and white blood counts for 17 patients (AG positive group) who died of acute myelogenous leukemia (data set introduced by Feigl and Zelen, 1965). The survival time, t_i, are given in weeks from data of diagnosis. The concomitant variable is the *log(white blood count)* at the time of diagnosis.

Assuming the log-linear model (4) with $\sigma^* = 1$, an exponential regression model for the survival time t_i, given by

$$y_i = \ln t_i = \beta_0 + \beta_1 x_i' + Z_i, \tag{20}$$

where $x_i' = x_i - \bar{x}, i = 1, \ldots, 17$. The maximum likelihood estimates for β_0 and β_1 are given by $\hat{\beta}_0 = 3.934\,(0.2425)$ and $\hat{\beta}_1 = -1.1094\,(0.4138)$.

t_i	WBC	x_i	t_i	WBC	x_i
65	2300	3.36	143	7000	3.85
156	750	2.88	56	9400	3.97
100	4300	3.63	26	32000	4.51
134	2600	3.41	22	35000	4.54
16	6000	3.78	1	100000	5.00
108	10500	4.02	1	100000	5.00
121	10000	4.00	5	52000	4.72
4	17000	4.23	65	100000	5.00
39	5400	3.73			

Table 1: Survival Times and White Blood Counts (AGP Positive Group).

Assuming the log-linear model (6) with a standard Extreme-Value density (3) and a mixture of two Normal distributions for the error and the prior densities with $a = 5$, $b = 7$, $c_1 = 5$, $c_2 = 5$, $d_1 = 0.5$, $d_2 = 0.5, e = 1, f = 1$, $g = 0$ and $h = 1$, we have in table 2, the posterior summaries for the parameters considering two separate Gibbs chains each of which ran for $S = 51000$ iterations. In order to diminish the effect of the starting parameters values we discarded the first 1000 elements of each chain. Convergence of the Gibbs algorithm was observed using diagnostics procedures available in CODA package (see Best, Cowles and Vines 1995). For each parameter we considered every 10^{th} draw and so we finally got a sample size $S = 10000$. The results were generated using Ox package version 2.10 (see Doornik, 1999).

We also have in table 2, the posterior summaries for the parameters considering a standard Normal density, $f(z) = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{z^2}{2}\right]$, and the Logistic density $f(z) = \frac{e^z}{(1+e^z)^2}$, $-\infty < z < \infty$ for Z in the log-linear model (6) with a mixture of two Normal distributions for the error.

Density for Z_i	Parameter	Mean	S.D.	95% Credible Interval
	β_0	3.8656	0.4221	(3.0347; 4.7407)
	β_1	−1.4155	0.6782	(−2.7550; −0.0480)
Standard	σ^*	1.1112	0.3050	(0.6321; 1.8334)
Extreme	μ	−0.0037	0.3327	(−0.6670; 0.6634)
Value	σ_1	0.9051	0.0534	(0.8441; 1.0425)
	σ_2	0.9086	0.0467	(0.8526; 1.0257)
	p	0.4114	0.1420	(0.1549; 0.7023)
	β_0	3.3296	0.4379	(2.4617; 4.1944)
	β_1	−1.8587	0.6687	(−3.2158; −0.5572)
	σ^*	1.2675	0.2785	(0.8274; 1.9030)
Standard	μ	0.0002	0.3350	(−0.6635; 0.6708)
Normal	σ_1	0.9044	0.0526	(0.8444; 1.0415)
	σ_2	0.9082	0.0474	(0.8522; 1.0316)
	p	0.4137	0.1428	(0.1596; 0.7051)
	β_0	3.3262	0.4545	(2.4480; 4.2450)
	β_1	−1.9185	0.7286	(−3.3964; −0.4589)
	σ^*	0.7720	0.1923	(0.4704; 1.2068)
Standard	μ	0.0028	0.3352	(−0.6530; 0.6761)
Logistic	σ_1	0.9040	0.0524	(0.8441; 1.0390)
	σ_2	0.9078	0.0459	(0.8526; 1.0249)
	p	0.4101	0.1414	(0.1565; 0.6931)

Table 2: Posterior Summaries (log-linear model (6)).

In table 3, we have the posterior summaries for $Z_1, \ldots, Z_{17}$ obtained by the same generated Gibbs samples.

In table 4 we have the predicted values $\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i' + \hat{\sigma}^* \hat{Z}_i$, $i = 1, \ldots, 17$ considering the log-linear model (6) with a mixture of Normal distributions where $\hat{\beta}_0, \hat{\beta}_1, \hat{\sigma}^*$ and $\hat{Z}_i$ are the Monte Carlo estimates of the posterior means for $\beta_0, \beta_1, \sigma^*$ and Z_i, based on two combined chains of size $S = 10000$.

As a criterium for comparison of the different distributions for the random effects Z_i, we could consider the sum of squares of the residuals $\hat{\epsilon}_i = y_i - \hat{y}_i$, $i = 1, \ldots, 17$. Observe that $\sum_{i=1}^{17} \hat{\epsilon}_i^2 = 0.1106$ considering the Logistic density for Z_i. Considering a standard Normal density we have $\sum_{i=1}^{17} \hat{\epsilon}_i^2 = 0.1788$ and considering the standard Extreme-Value density for Z_i we have $\sum_{i=1}^{17} \hat{\epsilon}_i^2 = 0.2129$. That is, the log-linear model (6) with a Logistic density for Z_i gives the best fit for the survival data presented in table 1.

$\hat{Z}_i$	Extreme-Value	Normal	Logistic
$\hat{Z}_1$	-0.6336	0.3772	0.6414
$\hat{Z}_2$	-0.4842	-0.3790	-0.7010
$\hat{Z}_3$	0.0463	0.3050	0.4706
$\hat{Z}_4$	0.0214	0.2195	0.3310
$\hat{Z}_5$	-1.3680	-0.8408	-1.4066
$\hat{Z}_6$	0.5447	0.9048	1.4833
$\hat{Z}_7$	0.6032	0.9496	1.5762
$\hat{Z}_8$	-2.0623	-1.2404	-2.0496
$\hat{Z}_9$	-0.6290	-0.2490	-0.4235
$\hat{Z}_{10}$	0.5609	0.8586	1.4119
$\hat{Z}_{11}$	-0.0391	0.3510	0.5640
$\hat{Z}_{12}$	-0.0596	0.5121	0.8649
$\hat{Z}_{13}$	-0.1399	0.4403	0.7533
$\hat{Z}_{14}$	-2.3007	-1.1647	-1.8845
$\hat{Z}_{15}$	-2.3024	-1.1652	-1.8856
$\hat{Z}_{16}$	-1.2130	-0.4090	-0.6135
$\hat{Z}_{17}$	1.0669	1.7901	3.1457

Table 3: Posterior Summaries for $Z_i, i = 1, \ldots, 17$.

Observed	$Z_i \sim$Extreme-Value	$Z_i \sim$Normal	$Z_i \sim$Logistic
4.1744	4.2009	4.2162	4.2396
5.0499	5.0557	5.1185	5.1272
4.6052	4.5716	4.5758	4.5767
4.8978	4.8532	4.8736	4.8881
2.7726	2.7953	2.8546	2.8499
4.6821	4.5766	4.6154	4.6147
4.7958	4.6716	4.7115	4.7271
1.3863	1.3836	1.5074	1.4858
3.6636	3.6812	3.6897	3.6967
4.9628	4.8439	4.8842	4.8974
4.0254	3.9959	4.0028	3.9972
3.2581	3.2200	3.2181	3.2088
3.0910	3.0758	3.0548	3.0479
0.0000	0.0293	0.1729	0.1369
0.0000	0.0274	0.1723	0.1360
1.6094	1.6400	1.6587	1.6629
4.1744	3.7714	3.9182	4.0202

Table 4: Predicted Values $\hat{y}_i, i = 1, \ldots, 1$ (log-linear model (6)).

If we assume the log-linear model (4) for the survival data of table 1 considering standard reference prior for β_0, β_1 and σ^*, we have in table 5 the posterior summaries for the parameters based on two combined chains of size $S = 10000$ generated Gibbs samples using BUGS package (see for example, Gilks et al. (1994); Spiegelhalter et al. (1995a); Spiegelhalter et al. (1995b); Spiegelhalter et. al. (1995c) or Best et al. (1996)).

We also have in table 5 the maximum likelihood estimates for the parameters considering the different densities for the error Z.

			Posterior	Summaries
Density for Z	Parameter	MLE	Mean	S.D.
Extreme-Value	β_0	3.9436	3.9660	0.2927
	β_1	−1.0983	−1.1480	0.4953
	σ^*	0.9785	1.1200	0.2523
Normal	β_0	3.3617	3.3590	0.3188
	β_1	−1.8830	−1.8900	0.5298
	σ^*	1.1579	1.2990	0.2562
Logistic	β_0	3.3469	3.3600	0.3328
	β_1	−1.9299	−1.9280	0.5680
	σ^*	0.6825	0.7594	0.1663

Table 5: Posterior Summaries (log-linear model (4)).

In table 6, we have Bayesian estimators for the predicted values $E\left[y_i|x_i^{'}\right] = \beta_0 + \beta_1 x_i^{'} + \sigma^* E\left(Z_i\right)$, $i = 1, \ldots, 17$. Observe that $E\left(Z_i\right) = -\gamma$ ($\gamma = 0.5722...$ is the Euler constant) for the Extreme-Value density and $E\left(Z_i\right) = 0$ for standard Normal and Logistic densities.

Observe that $\sum_{i=1}^{17} \hat{\epsilon}_i^2 = 26.2052$ ($\hat{\epsilon}_i = y_i - \hat{y}_i$) considering the Extreme-Value density for Z in the log-linear model (4). In the same way $\sum_{i=1}^{17} \hat{\epsilon}_i^2 = 22.7952$ (Normal density for Z) and $\sum_{i=1}^{17} \hat{\epsilon}_i^2 = 22.8075$ (Logistic density for Z).

That is, we have larger values for the residuals considering the log-linear model (4) than using the log-linear model (6) where we consider a mixture of two Normal distributions.

Observe that considering the log-linear model (6) we have better predicted values $\hat{y}_i, i = 1, \ldots, 17$ (see table 4). In summary, the inclusion of the error term ϵ considering a mixture of two Normal distributions gives better fit for the survival data of table 1.

Observed	Z_i ~Extreme-Value	Z_i ~Normal	Z_i ~Logistic
4.1744	4.1624	4.7466	4.7755
5.0499	4.7211	5.6664	5.7138
4.6052	3.8504	4.2331	4.2516
4.8978	4.1013	4.6460	4.6729
2.7726	3.6843	3.9596	3.9727
4.6821	3.4053	3.5003	3.5041
4.7958	3.4296	3.5403	3.5450
1.3863	3.1651	3.1048	3.1007
3.6636	3.7369	4.0461	4.0609
4.9628	3.6075	3.8331	3.8436
4.0254	3.4605	3.5911	3.5968
3.2581	2.8497	2.5856	2.5710
3.0910	2.8051	2.5120	2.4960
0.0000	2.2816	1.6503	1.6170
0.0000	2.2816	1.6503	1.6170
1.6094	2.6077	2.1871	2.1645
4.1744	2.2816	1.6503	1.6170

Table 6: Predicted Values (log-linear model(4)).

For model selection, we also could use the predictive density for t_i (see for example Gelfand 1996) given $t_{(i)} = (t_1, \ldots, t_{i-1}, t_{i+1}, \ldots, t_n)$.

The predictive density for t_i given $t_{(i)}$ is given by

$$\begin{aligned} c_i &= f\left(t_i|t_{(i)}\right) \\ &= \int f\left(t_i|\boldsymbol{\theta}, x_i\right) \pi\left(\boldsymbol{\theta}|\mathbf{t}_{(i)}, x_i\right) d\boldsymbol{\theta}, \end{aligned} \tag{21}$$

where $\pi\left(\boldsymbol{\theta}|\mathbf{t}_{(i)}, x_i\right)$ is the posterior density for $\boldsymbol{\theta}$ given the data $t_{(i)}$, and $f\left(t_i|\boldsymbol{\theta}, x_i\right)$ is the density for the lifetime t_i with covariate x_i, given by

$$f\left(t_i|\boldsymbol{\theta}, x_i\right) = \frac{1}{\sigma^* \boldsymbol{\theta}_{x_i}} \left(\frac{t_i}{\boldsymbol{\theta}_{x_i}}\right)^{\frac{1}{\sigma^*}-1} \exp\left[-\left(\frac{t_i}{\boldsymbol{\theta}_{x_i}}\right)^{\frac{1}{\sigma^*}}\right], \tag{22}$$

where $\boldsymbol{\theta}_{x_i} = \exp\left(\beta_0 + \beta_i x_i'\right)$ considering the location-scale model (4) with a standard Extreme-Value distribution for Z and

$$f\left(t_i|\boldsymbol{\theta}, x_i\right) = \sum_{j=2}^{2} p_j f_j\left(t_i|\boldsymbol{\beta}, Z_i, \mu_j, \sigma_j^2, x_i\right), \tag{23}$$

for the location-scale model (6) with Extreme-Value-Mixture of Normal distributions for the error, here

$$f_j\left(t_i|\beta, Z_i, \mu_j, \sigma_j^2, x_i\right) = \frac{1}{t_i\sqrt{2\pi}\sigma_j} \exp\left[-\frac{1}{2\sigma_j^2}\left(\ln t_i - \beta_0 - \beta_1 x_i' - \sigma^* Z_i - \mu_j\right)^2\right] \tag{24}$$

(a log-normal density).

We can use $c_i = f\left(t_i|t_{(i)}\right)$ in model selection. In this way, we consider plots of c_i versus $i\,(i = 1, \ldots, n)$ for different models; larger values of c_i (in average) indicates the better model. Similarly, we get $c_i\,(i = 1, \ldots, n)$ considering the standard Normal or Logistic densities for Z_i. In figure 1, we have the plots of c_i against i considering both, the location-scale model (4) and the scale model (6).

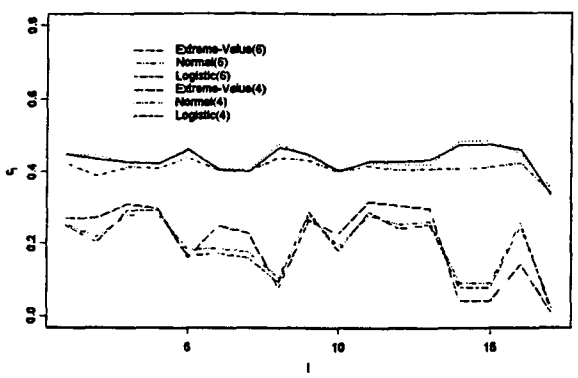

Figure 1: Predictive Densitie.

4.2 An Example With Right Censored Data

In table 7, we have the remission survival times for a group of 21 leukemia patients receiving a certain treatment (Freireich et al., 1963). The concomitant variable is *log(white blood count)*.

t_i	x_i	t_i	x_i	t_i	x_i
6	2.31	22	2.32	19^+	2.05
6	4.06	23	2.57	20^+	2.01
6	3.28	6^+	3.20	25^+	1.78
7	4.43	9^+	2.80	32^+	2.20
10	2.96	10^+	2.70	32^+	2.53
13	2.88	11^+	2.60	34^+	1.47
16	3.60	17^+	2.16	35^+	1.45

Table 7: Remission Survival Times and log(*white blood count*).

From table 7, we observe 9 uncensored observations and 12 right censored observations (observations with +).

Assuming the log-linear model (6) with a standard Extreme Value density (3) and a mixture of two Normal distributions for the error and the prior (18) with $a = 5$, $b = 2$, $c_1 = 0.1$, $c_2 = 0.1$, $d_1 = 0.1$, $d_2 = 0.1$, $e = 15$, $f = 39$, $g = 0.5$, and $h = 50$, we have in table 8, the posterior summaries for the parameters considering two separate Gibbs chains for which ran for $S = 51000$ iterations. For each parameter we consider $1010^{th}, 1020^{th}, \ldots$ samples, that is, we finally got a sample of size $S = 10000$. We also have in table 8, the posterior summaries for the parameters considering a standard Normal density and the Logistic density for Z in the log-linear model (6) with a mixture of two Normal distributions for the error.

In table 9, we have the posterior summaries for $Z_1, \ldots, Z_{21}$ obtained by the generated Gibbs samples.

Density for Z_i	Parameter	Mean	S.D.	95% Credible Interval
	β_0	3.2365	0.2133	(2.9103; 3.6051)
	β_1	−0.8200	0.2410	(−1.2251 − 0.4267)
Standard	σ^*	0.5108	0.0917	(0.3651; 1.8334)
Extreme	μ	−0.8200	0.2410	(−0.1952; 0.1935)
Value	σ_1	0.7050	0.0729	(0.3651; 0.67091)
	σ_2	0.7393	0.1492	(0.6177; 0.8371)
	p	0.6775	0.1434	(0.4111; 0.8775)
	β_0	3.0482	0.3773	(2.5804; 3.7421)
	β_1	−0.8223	0.3408	(−1.3960; −0.3405)
Standard	σ^*	0.4624	0.0984	(0.3228; 0.6385)
Normal	μ	0.1106	0.3749	(−0.3664; 0.7676)
	σ_1	0.8327	0.2458	(0.6233; 1.3288)
	σ_2	0.7950	0.2029	(0.6253; 1.1440)
	p	0.4396	0.1351	(0.2348; 0.6790)
	β_0	3.2157	0.3531	(2.7363; 3.8706)
	β_1	−0.9268	0.3531	(−1.5954; −0.4215)
	σ^*	0.2500	0.0584	(0.1684; 0.3576)
Standard	μ	0.0633	0.2641	(−0.3051; 0.5312)
Logistic	σ_1	0.7592	0.1258	(0.6175; 1.0030)
	σ_2	0.7467	0.1152	(0.6188; 0.9577)
	p	0.4527	0.1256	(0.2588; 0.6713)

Table 8: Posterior Summaries (log-linear model (6)).

$\hat{Z}_i$	Extreme-Value	Normal	Logistic
$\hat{Z}_1$	−3.1216	−5.3087	−2.1113
$\hat{Z}_2$	−0.4899	−0.2528	−0.0814
$\hat{Z}_3$	−1.6219	−2.3908	−1.0520
$\hat{Z}_4$	0.1625	0.9635	0.5262
$\hat{Z}_5$	−1.1735	−1.6951	−0.6629
$\hat{Z}_6$	−0.8256	−1.1194	−0.3406
$\hat{Z}_7$	0.4451	1.2941	0.8227
$\hat{Z}_8$	−0.6839	−1.0206	−0.2141
$\hat{Z}_9$	−0.2651	−0.2270	0.1498
$\hat{Z}_{10}$	−1.5722	0.0642	0.0119
$\hat{Z}_{11}$	−1.6536	0.0642	0.0513
$\hat{Z}_{12}$	−1.7247	0.0920	0.0605
$\hat{Z}_{13}$	−1.2287	0.0813	0.0635
$\hat{Z}_{14}$	−0.9577	0.1201	0.1228
$\hat{Z}_{15}$	−0.7897	0.1506	0.1345
$\hat{Z}_{16}$	−0.8014	0.1247	0.1617
$\hat{Z}_{17}$	−0.6702	0.1769	0.1918
$\hat{Z}_{18}$	0.9763	0.8077	0.6538
$\hat{Z}_{19}$	1.4850	1.3274	0.9374
$\hat{Z}_{20}$	−0.4449	0.2669	0.2440
$\hat{Z}_{21}$	−0.4549	0.2792	0.2438

Table 9: Posterior Summaries for $Z_i, i = 1, \ldots, 21$.

The sum of squares of the residuals, $\hat{\varepsilon}_i = y_i - \hat{y}_i$, considering the uncensored observations is given by $\sum_{i=1}^{21} \hat{\epsilon}_i^2 = 0.0524$, assuming the standard Extreme-Value density for Z_i. Assuming a Normal density for Z_i, we have $\sum_{i=1}^{21} \hat{\epsilon}_i^2 = 5.0851$ and a Logistic density, we have $\sum_{i=1}^{21} \hat{\epsilon}_i^2 = 0.3435$. That is, the best model for the remission times of table 7 is given by the log-linear model (6) with a Extreme-Value density for Z_i.

In table 10, we have the maximum likelihood estimators for β_0, β_1 and σ^* assuming the log-linear model (4) with Z_i assuming the Extreme-Value density, a standard Normal density and a Logistic density (asymptotical standard erros in parenthesis).

Considering only the uncensored data and the log-linear model (4) the sum of squares of the residuals, $\hat{\epsilon}_i = y_i - \hat{y}_i$, where $\hat{y}_i$ are the predicted values based on the maximum likelihood estimators (see table 10), are given by $\sum_{i=1}^{21} \hat{\epsilon}_i^2 = 5.1204$ (standard Extreme-Value), $\sum_{i=1}^{21} \hat{\epsilon}_i^2 = 8.5736$ (standard Normal) and $\sum_{i=1}^{21} \hat{\epsilon}_i^2 = 3.8450$ (standard Logistic).

We observe that considering the linear model (6) with a mixture of two Normal

Parameter	Extreme-Value	Normal	Logistic
β_0	3.2691 (0.1640)	3.0904 (0.1914)	3.0863 (0.1712)
β_1	0.8419(0.1674)	−0.8645 (0.2213)	−0.8380 (0.1930)
σ^*	0.3947 (0.1011)	0.5923 (0.1418)	0.3124 (0.0839)

Table 10: Maximum Likelihood Estimates (log-linear model (4))

distributions for the error, we get better fit for the lifetime data of table 7 (small values for the sum of squares of the residuals).

4.3 An Example With Interval-Censored Data

In table 11 we have medical data from a clinical trial where patients are assessed only at pre-scheduled visits. If the event has not occurred at one visit (at time L) but has by the following visit (at time R), the lifetime T is in the interval $[L, R]$. This data set (introduced by Lindsey and Ryan, 1998) is related to drug resistance (measure using a plaque reduction assay) to zidovudine in patients enrolled in four clinical trials for the treatment of AIDS.

Left	Right	Stage	Left	Right	Stage
0	16	0	0	15	0
15	26	0	3	26	1
1	26	0	4	26	1
17	26	0	1	11	1
13	26	0	13	19	1
0	24	0	0	6	1
6	26	0	0	11	1
0	15	0	6	26	1
14	26	0	0	6	1
12	26	0	2	12	1
13	26	0	1	17	1
12	26	0	0	14	1
12	26	0	0	25	1
0	18	0	2	11	1
0	14	0	0	14	1
0	17	0			

Table 11: AIDS Data Set: Value of Right = 26 Implies Right-Censored Data.

Samples were collected on the patients at a subset of the scheduled visit times dictated by the four protocols. Since the resistance assays were very expensive, there were few assessments on each patient resulting in very wide intervals $[L, R]$, if resistance was seen to have occurred, and a high proportion of right-censored observations.

Associated to the resistance time there are some covariates. One of these covariates is the stage of the disease x (see table 11).

Assuming the log-linear model (6) with a standard Extreme-Value density (3) and a mixture of two Normal distributions for the error and the prior densities (18) with $a = 4$, $b = 9$, $c_1 = 0.1$, $c_2 = 0.1$, $d_1 = 0.1$, $d_2 = 0.1$, $e = 8$, $f = 20$, $g = 0$, and $h = 50$, we have in table 12 the posterior summaries for the parameters considering $S = 10000$ generated Gibbs samples.

Density for Z_i	Parameter	Mean	S.D.	95% Credible Interval
	β_0	2.6731	0.2224	(2.3080; 3.0308)
	β_1	−0.9201	0.4310	(−1.6532; −0.2330)
Standard	σ^*	0.1581	0.0645	(0.0742; 0.2809)
Extreme	μ	−0.0017	0.1191	(−0.1954; 0.1922)
Value	σ_1	0.6446	0.0670	(0.5797; 0.7742)
	σ_2	0.6344	0.0496	(0.5820; 0.7303)
	p	0.3484	0.1089	(0.1899; 0.5442)

Table 12: Posterior Summaries (log-linear model (6)).

In table 13, we have the posterior summaries for $Z_1, \ldots, Z_{31}$ obtained by the same generated Gibbs samples.

$\hat{Z}_1$	0.3212	$\hat{Z}_{17}$	−0.5683
$\hat{Z}_2$	−0.0777	$\hat{Z}_{18}$	−1.0668
$\hat{Z}_3$	0.5767	$\hat{Z}_{19}$	−0.6339
$\hat{Z}_4$	0.0468	$\hat{Z}_{20}$	−0.8771
$\hat{Z}_5$	−0.4189	$\hat{Z}_{21}$	−1.0929
$\hat{Z}_6$	0.1902	$\hat{Z}_{22}$	−0.4597
$\hat{Z}_7$	−0.0754	$\hat{Z}_{23}$	−1.0763
$\hat{Z}_8$	0.0531	$\hat{Z}_{24}$	−0.4355
$\hat{Z}_9$	−0.0627	$\hat{Z}_{25}$	−0.4260
$\hat{Z}_{10}$	−0.0688	$\hat{Z}_{26}$	−0.4338
$\hat{Z}_{11}$	−0.3327	$\hat{Z}_{27}$	−0.4770
$\hat{Z}_{12}$	−0.1639	$\hat{Z}_{28}$	1.1431
$\hat{Z}_{13}$	0.2782	$\hat{Z}_{29}$	−0.4370
$\hat{Z}_{14}$	−0.7402	$\hat{Z}_{30}$	−0.4109
$\hat{Z}_{15}$	−0.4452	$\hat{Z}_{31}$	−0.4446
$\hat{Z}_{16}$	−0.8876		

Table 13: Posterior Summaries for $Z_i, i = 1, \ldots, 31$.

It is interesting to observe that if we assume the log-linear model (4) with a standard Extreme-Value density for Z, the maximum likelihood estimates for

β_0, β_1 and σ^* are given by $\hat{\beta}_0 = 2.8907\,(0.1708)$, $\hat{\beta}_1 = -0.7185\,(0.2711)$ and $\sigma^* = 0.3934\,(0.1389)$.

In table 14, we have the values for the midpoints y_i in each interval-censored observation and the predicted values $y_i, i = 1, \ldots, 31$ considering the log-linear model (4) and (6). For the log-linear model (4) we have the predicted values $\hat{y}_i$ based on the maximum likelihood estimates for β_0, β_1 and σ^*. For the log-linear model (6), we have the predicted values based on the Monte Carlo estimates of the parameters using the $S = 10000$ generated Gibbs samples. We also have in table 14, the residuals, $y_i - \hat{y}_i$, considering de log-linear models (4) and (6). We observe much better fit of model (6) with a mixture of two Normal densities for the AIDS interval-censored data of table 11. We have $\sum_{i=1}^{31} = 21.1211$ for the log-linear model (4) and $\sum_{i=1}^{31} = 11.0088$ for the log-linear model (6).

	Model (4)	Model (6)	Model (4)	Model (6)
y_i	$\hat{y}_i$	$\hat{y}_i$	$\hat{\epsilon}_i$	$\hat{\epsilon}_i$
2.0794	2.8907	2.7238	−0.8113	−0.6444
2.0149	2.8907	2.6608	−0.8758	−0.6459
1.7047	2.8907	2.7642	−1.1860	−1.0595
2.4423	2.8907	1.7603	−0.4484	0.6820
2.5257	2.8907	2.6069	−0.3650	−0.0811
2.3979	2.8907	1.7830	−0.4928	0.6149
1.5041	2.8907	2.6611	=1.3866	=1.1571
1.6094	2.8907	1.7613	−1.2813	−0.1519
1.8718	2.8907	2.6632	−1.0189	−0.7914
1.0986	2.8907	1.7421	−1.7921	−0.6434
2.4849	2.8907	2.6205	−0.4058	−0.1356
1.0986	2.8907	1.7270	−1.7921	−0.6284
2.3026	2.8907	2.7170	−0.5881	−0.4145
1.7047	2.8907	1.6359	−1.1860	0.0688
2.0149	2.8907	2.6027	−0.8758	−0.5878
2.3026	2.8907	1.6126	−0.5881	0.6899
1.7918	2.8907	2.5832	−1.0990	−0.7915
1.0986	2.1722	1.5843	−1.0736	−0.4857
1.9459	2.1722	2.5729	−0.2263	−0.6270
1.6094	2.1722	1.6143	−0.5628	−0.0049
1.8718	2.1722	2.5003	−0.3004	−0.6285
2.0794	2.1722	1.6803	−0.0927	0.3992
1.9459	2.1722	2.5030	−0.2263	−0.5570
1.9459	2.1722	1.6841	−0.2263	0.2618
1.9459	2.1722	2.6057	−0.2263	−0.6598
2.5257	2.1722	1.6844	0.3535	0.8414
2.1972	2.1722	2.5977	0.0250	−0.4005
1.5041	2.1722	1.9336	−0.6681	−0.4295
1.9459	2.1722	2.6040	−0.2263	−0.6581
1.9459	2.1722	1.6880	−0.2263	0.2579
2.1401	2.1722	2.6028	−0.0321	−0.4627

Table 14: Midpoints y_i, Predicted Values $\hat{y}_i$ and Residuals $\hat{\epsilon}_i$.

5 Concluding Remarks

The use of log-linear model (6) with a mixture of two Normal distributions for the error could be a good alternative to analyse censored survival data. In the samples introduced in section (4), we observe better fit of the log-linear model (6) considering a mixture of two Normal distributions in comparison with the usual log-linear model (4).

It is important to point out that the use of MCMC methods to get the posterior summaries of interest does not require sophisticated computational expertise and this approach could be extended to survival data with many covariates and also including more than two Normal distributions for the mixture distribution of the error ϵ in the log-linear model (6).

Acknowledgements

The author Josmar Mazucheli, graduate student at COPPE/PO/UFRJ, gratefully acknowledges partial financial support from CAPES/Brazil and the Department of Statistics DEs/UEM.

References

Best, N. G.; Cowles, M. K.; Vines, S. K. (1995). *CODA: Convergence Diagnosis and Output Analysis Software for Gibbs Sampling Output, Version 0.3.* MRC Biostatistics Unit, Cambridge.

Best, N. G.; Spiegelhalter, D. J.; Thomas, A.; Brayne, C. E. G. (1996). Bayesian Analysis of Realistically Complex Models. Journal of the Royal Statistical Society, A, **159**, 323-342.

Carroll, J. R.; Roeder K.; Wasserman L. (1996). Flexible Parametric Measurement Error Models. Technical Reports #648, Department of Statistics, Carnegie Mellon University. Available from http://lib.stat.cmu.edu/cmu-stats/tr/.

Cox, D. R. (1972). Regression Models and Life Tables (with discussion). Journal of the Royal Statistical Society, B, **24**, 406-423.

Diebolt, J.; Robert, C. P. (1994). Estimation of Finite Mixture Distributions Through Bayesian Sampling. Journal of the Royal Statistical Society, B, **56**, 363-375.

Doornik, J.A. (1999). *Object-Oriented Matrix Programming Using Ox, 3rd* ed. London: Timberlake Consultants Press and Oxford: www.nuff.ox.ac.uk/Users/Doornik.

Feigl, P.; Zelen, M. (1965). Estimation of Exponential Survival Probabilities with Concomitant Information. Biometrics, **21**, 826-838.

Freireich, E. J.; Gehan, E.; Frei, E.; Schroeder, L. R. Wolman, I. J.; Anbari, R.; Burgert, E. O.; Mills, S. D.; Pinkel, D.; Selawry, O. S.; Moon, J. H.; Gendel, B. R.; Spurr, C. L.; Storrs, R.; Haurani, F.; Hoogstraten, B.; Lee, S. The Effect of 6-Mercaptopurine on the Duration of Steroid-Induced Remissions in Acute Leukemia: A model for Evaluation of Other Potentially Used Therapy. Blood, **21**, 699-716.

Gelfand, A. (1996). Model Determination Using Sampling-Based Methods. *In Practical Markov Chain Monte Carlo*, (eds. W. R. Gilks, S. Richardson, D. J. Spiegelhalter). London: Chapman and Hall.

Gilks, W. R.; Thomas, A. and Spiegelhater, D. J. (1994). A language and Program for Complex Bayesian Modelling. The Statistician, **43**, 169-178.

Kalbfleisch, J. D.; Prentice. R. L. (1980). *The Statistical Analysis of Failure Time Data.* New York: Wiley.

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

Lindsey, J. C.; Ryan, L. M. (1998). Tutorial in Biostatistics Methods for Interval-Censored Data. Statistics in Medicine, **17**, 219-238.

Polymenis, A.; Titterington, D. M. (1998). On the Determination of the Number of Components in a Mixture. Statistics & Probability Letter, **38**, 295-298.

Richardson, S.; Green, P. J. (1997). On Bayesian Analysis of Mixtures with an Unknown Number of Components. Journal of the Royal Statistical Society, B, **59**, 731-792.

Robert, C. P. (1996). Mixtures of Distributions: Inference and Estimation. *In Practical Markov Chain Monte Carlo*, (eds. W. R. Gilks, S. Richardson, D. J. Spiegelhalter). London: Chapman and Hall.

Roeder K.; Wasserman L. (1997). Parctical Bayesian Density Estimation Using Mixture of Normals. Journal of the American Statistical Association, 92, 894-902.

Smith, A. F. M.; Roberts, G. (1993). Bayesian Computation via the Gibbs Sampler and Related Markov Chain Monte Carlo Methods (with discussion). Journal of the Royal Statistical Society, B, **55**, 3-23.

Spiegelhalter, D. J.; Thomas, A.; Best, N. G. and Gilks W. R. (1995c). BUGS: Bayesian Inference Using Gibbs Sampling, Version 0.50. MRC Biostatistics Unit Cambridge.

Spiegelhalter, D. J.; Thomas, A. and Best, N. G. (1995b). Computation on Bayesian Graphical Models. In Bayesian Statistics 5. Clarendon Press, Oxford, UK.

Spiegelhalter, D. J,; Best, N. G.; Gilks, W. R. and Inskip, H. (1995a). Hepatits: A Case Study in MCMC Method. *In Practical Markov Chain Monte Carlo*, (eds. W. R. Gilks, S. Richardson, D. J. Spiegelhalter). London: Chapman and Hall.

Stephens, M. (1997). Bayesian Methods for Mixture of Normal Distributions. Ph.D. thesis, University of Oxford.
Available from http://www.stats.ox.ac.uk/~stephens/index.html.

Stephens, M. (1998). Bayesian Analysis of Mixtures with an Unknown Number of Components - An Alternative to Reversible Jump Methods. Submitted to Annals of Statistics.
Available from http://www.stats.ox.ac.uk/~stephens/index.html.

Tanner, M.; Wong, W. (1987). The Calculations of posterior Distributions by Data Augmentation. Journal of the American Statistical Association, **82**, 528-550.

NOTAS DO ICMC

SÉRIE ESTATÍSTICA

058/99 ACHACAR, J. A. ; FORTULAN, V. C. – Meta Analysis: A bayesian approach.

057/99 OLIVEIRA, S.C.; ACHCAR, J.A. – Confiabilidade de redes: um enfoque bayesiano.

056/98 RODRIGUES, J.; CHAVES, J.S. – A note on bayesian exponential regression model with censored data.

055/99 RODRIGUES, J.; SILVEIRA, V.D.R. – Bayesian computation for dichotomous variables with classification errours.

054/98 ACHCAR, J.A.; FORTULAN, V.C. – Relação entre o uso de hormônio e câncer em mulheres: uma aplicação de meta-análise sob um enfoque bayesiano.

053/98 CID, J. E.R.; ACHCAR, J..A. – Bayesian inference for nonhomogeneous poisson processes in software reliability models assuming nonmonotonic intensity functions.

052/98 ANDRADE FILHO, M.G; MIZOI, M.F. – Aplicação de MCMC na estimação de máxima verossimilhança para processos AR(p) e MA(q).

051/98 RODRIGUES, J. – Bayesian analysis for the accelerated life tests with informative prior distributions obtained from fixed stresses.

050/98 ACHCAR, J.A.; ANDRADE, .M.G.; LOIBEL, S. - A bayesian analysis for homogeneous poisson processes with change-points.

049/98 ACHCAR, J.A; PEREIRA, G.A. - Use of mixture of exponential power distributions for interval-censored survival data in presence of covariates.