Revista Colombiana de Estadística
Junio 2011, volumen 34, no. 1, pp. 111 a 131

A Bayesian Analysis in the Presence of Covariates for Multivariate Survival Data: An example of Application

Análisis bayesiano en presencia de covariables para datos de sobrevivencia multivariados: un ejemplo de aplicación

Carlos Aparecido Santos[1,a], Jorge Alberto Achcar[2,b]

[1]Departamento de Estatística, Centro de Ciências Exatas, UEM - Universidade Estadual de Maringá, Maringá-PR, Brasil

[2]Departamento de Medicina Social, FMRP - Faculdade de Medicina de Ribeirão Preto, USP - Universidade de São Paulo, Ribeirão Preto-SP, Brasil

Abstract

In this paper, we introduce a Bayesian analysis for survival multivariate data in the presence of a covariate vector and censored observations. Different "frailties" or latent variables are considered to capture the correlation among the survival times for the same individual. We assume Weibull or generalized Gamma distributions considering right censored lifetime data. We develop the Bayesian analysis using Markov Chain Monte Carlo (MCMC) methods.

***Key words*:** Bayesian methods, Bivariate distribution, MCMC methods, Survival distribution, Weibull distribution.

Resumen

En este artículo, se introduce un análisis bayesiano para datos multivariados de sobrevivencia en presencia de un vector de covariables y observaciones censuradas. Diferentes "fragilidades" o variables latentes son consideradas para capturar la correlación entre los tiempos de sobrevivencia para un mismo individuo. Asumimos distribuciones Weibull o Gamma generalizadas considerando datos de tiempo de vida a derecha. Desarrollamos el análisis bayesiano usando métodos Markov Chain Monte Carlo (MCMC).

***Palabras clave*:** distribución bivariada, distribución de sobrevivencia, distribución Weibull, métodos bayesianos, métodos MCMC.

[a]Adjoint professor. E-mail: casantos@uem.br
[b]Professor. E-mail: jorge.achcar@pq.cnpq.br

1. Introduction

Different parametric regression models are introduced in the literature to analyse lifetime data in the presence of censored data (see for example, Lawless 1982). A popular semi-parametric regression model to analyse survival data was introduced by Cox (1972) assuming proportional hazards (see also, Cox & Oakes 1984). In these models, the survival times are independent, that is, the individuals are not related to each other.

In many practical situations, especially in medical studies, to have dependent survival times is common, when the individuals are related to each other (same family, repeated measurements in the same individual or two or more measuremenst in the same patient).

As an example, we could consider a survival data set introduced by McGilchrist & Aisbett (1991) related to kidney infection where the recurrence of infection of 38 kidney patients, using portable dialysis machines, is recorded. Infections may occur at the location of insertion of the catheter. The time recorded, called infection time, is either the survival time (in days) of the patient until an infection occurred and the catheter had to be removed, or the censored time, where the catheter was removed by others reasons. The catheter is reinserted after some time and the second infection time is again observed or censored (data set in Table 1).

Different survival multivariate models are introduced in the literature to analyse dependent lifetime data in the presence of a covariate vector and censored observations.

To capture the correlation among two or more survival times, we could consider the introduction of "frailties" or latent variables (see for example, Clayton & Cuzick (1985), Oakes (1986, 1989) and Shih & Louis (1992)), assuming proportional hazard models.

Clayton (1991) uses a Levy process (Kalbfleisch 1978) as a nonparametric Bayesian model for the baseline hazard, applied to continuous data, that is, data with no ties.

In this paper, we assume parametric regression models for dependent survival data in the presence of censored observations considering the special Weibull distribution, a popular lifetime model and the Generalized Gamma distribution, a supermodel that generalizes some common models used for lifetime data as the Weibull, the Gamma, and the log-normal distributions.

Different "frailties" are assumed to model the dependent structure of the data, under the Bayesian paradigm.

For a Bayesian analysis of the proposed models, we use MCMC (Markov Chain Monte Carlo) methods to obtain posterior summaries of interest (see for example, Gelfand & Smith (1990) and Chib & Greenberg (1995)).

The paper is organized as follows: in Section 2, we introduce a Weibull regression model for multivariate survival data; in Section 3, we introduce a Bayesian analysis; in Section 4, we consider the use of a generalized Gamma distribution for

multivariate survival data; in Section 5 we present an analysis for the recurrence times introduced in Table 1.

TABLE 1: Recurrence times of infections in 38 kidney patients.

Patient	First time	Second time	Censoring first time	Censoring second time	Sex
1	8	16	1	1	1
2	23	13	1	0	2
3	22	28	1	1	1
4	447	318	1	1	2
5	30	12	1	1	1
6	24	245	1	1	2
7	7	9	1	1	1
8	511	30	1	1	2
9	53	196	1	1	2
10	15	154	1	1	1
11	7	333	1	1	2
12	141	8	1	0	2
13	96	38	1	1	2
14	149	70	0	0	2
15	536	25	1	0	2
16	17	4	1	0	1
17	185	117	1	1	2
18	292	114	1	1	2
19	22	159	0	0	2
20	15	108	1	0	2
21	152	562	1	1	1
22	402	24	1	0	2
23	13	66	1	1	2
24	39	46	1	0	2
25	12	40	1	1	1
26	113	201	0	1	2
27	132	156	1	1	2
28	34	30	1	1	2
29	2	25	1	1	1
30	130	26	1	1	2
31	27	58	1	1	2
32	5	43	0	1	2
33	152	30	1	1	2
34	190	5	1	0	2
35	119	8	1	1	2
36	54	16	0	0	2
37	6	78	0	1	2
38	63	8	1	0	1

(Censoring (0); infection ocurrence (1); male (1); female (2))

2. A Weibull Regression Model for Multivariate Survival Data

Let T_{ji} be a random variable denoting the survival time of the i^{th} individual ($i = 1, 2, \ldots, n$) in the j^{th} repeated measurement for the same individual ($j = 1, 2, \ldots, k$) with a Weibull (1951) distribution with density,

$$f(t_{ji} \mid \nu_j, \lambda_j(i)) = \nu_j \lambda_j(i) t_{ji}^{\nu_j - 1} \exp\{-\lambda_j(i) t_{ji}^{\nu_j}\} \tag{1}$$

where $t_{ji} > 0$; $\nu_j > 0$ is the shape parameter and $\lambda_j(i)$ is the scale parameter.

To capture the correlation among the repeated measures $T_{1i}, T_{2i}, \ldots, T_{ki}$ for the same individual, we introduce a "frailty" or latent variable W_i, $i = 1, 2, \ldots, n$ with a normal distribution, that is,

$$W_i \overset{\text{iid}}{\sim} N(0, \sigma_w^2) \tag{2}$$

In the presence of a covariate vector $\mathbf{x}_i = (x_{i1}, x_{i2}, \ldots, x_{ip})'$ and the latent variable W_i, we assume the regression model in (1), given by

$$\lambda_j(i) = \exp\{w_i + \boldsymbol{\beta}_j' \mathbf{x}_i\} \tag{3}$$

where $\boldsymbol{\beta}_j = (\beta_{j1}, \beta_{j2}, \ldots, \beta_{jp})$ is the vector of regression parameters, $j = 1, 2, \ldots, k$.

The hazard function is given by

$$h_j(t_{ji} \mid \mathbf{x}_i, w_i) = \nu_j t_{ji}^{\nu_j - 1} \exp\{w_i + \boldsymbol{\beta}_j' \mathbf{x}_i\} \tag{4}$$

The survival function for a given t_{ji} is

$$S(t_{ji} \mid \mathbf{x}_i, w_i) = \exp\{-t_{ji}^{\nu_j} e^{w_i + \boldsymbol{\beta}_j' \mathbf{x}_i}\} \tag{5}$$

for $i = 1, 2, \ldots, n$ and $j = 1, 2, \ldots, k$.

Let us denote the model defined by (1)-(5) as "model 1".

From (4), we observe that we can have constant, decreasing or increasing hazards, assuming, respectively, $\nu_j = 1$, $\nu_j < 1$ or $\nu_j > 1$.

The conditional mean and variance for T_{ji} given $\mathbf{x}_i$ and w_i, are given, respectively, by

$$E(T_{ji} \mid \mathbf{x}_i, w_i) = \frac{\Gamma(1 + 1/\nu_j)}{\exp\left\{\frac{1}{\nu_j}\left(w_i + \boldsymbol{\beta}_j' \mathbf{x}_i\right)\right\}}$$

and

$$\text{Var}(T_{ji} \mid \mathbf{x}_i, w_i) = \frac{1}{\exp\left\{\frac{2}{\nu_j}\left(w_i + \boldsymbol{\beta}_j' \mathbf{x}_i\right)\right\}} \left\{\Gamma\left(1 + \frac{2}{\nu_j}\right) - \Gamma^2\left(1 + \frac{1}{\nu_j}\right)\right\}$$

for $i = 1, 2, \ldots, n$; $j = 1, 2, \ldots, k$.

The unconditional mean for T_{ji} is obtained from the result, $E(T_{ji} \mid \mathbf{x}_i) = E[E(T_{ji} \mid \mathbf{x}_i, w_i)]$, that is,

$$E(T_{ji} \mid \mathbf{x}_i) = \frac{\Gamma(1 + 1/\nu_j)}{\exp\left(\frac{\boldsymbol{\beta}_j' \mathbf{x}_i}{\nu_j}\right)} E\left\{e^{-W_i/\nu_j}\right\}$$

Observe that, since $W_i \sim N(0, \sigma_w^2)$, we have

$$g(W_i) = e^{-W_i/\nu_j} \overset{\text{a}}{\sim} N\left\{g(0); [g'(0)]^2 \sigma_w^2\right\}$$

("delta method"), that is,

$$e^{-W_i/\nu_j} \overset{\text{a}}{\sim} N\left[1; \frac{\sigma_w^2}{\nu_j^2}\right]$$

Thus, the unconditional mean for T_{ji} given $\mathbf{x}_i$ is,

$$E(T_{ji} \mid \mathbf{x}_i) = \frac{\Gamma(1+1/\nu_j)}{\exp\left(\frac{\boldsymbol{\beta}_j'\mathbf{x}_i}{\nu_j}\right)} \tag{6}$$

for $i = 1, 2, \ldots, n$; $j = 1, 2, \ldots, k$.

The unconditional variance for T_{ji} is obtained from $\text{Var}(T_{ji} \mid \mathbf{x}_i) = \text{Var}\{E(T_{ji} \mid \mathbf{x}_i, \mathbf{w}_i)\} + E\{\text{Var}(T_{ji} \mid \mathbf{x}_i, w_i)\}$, that is,

$$\begin{aligned}\text{Var}(T_{ji} \mid \mathbf{x}_i) = {} & \frac{\Gamma^2(1+1/\nu_j)}{\exp\left(\frac{2\beta_j'\mathbf{x_i}}{\nu_j}\right)}\text{Var}(e^{-W_i/\nu_j}) \\ & + \frac{\left[\Gamma(1+2/\nu_j) - \Gamma^2(1+1/\nu_j)\right]}{\exp\left(\frac{2\boldsymbol{\beta}_j'\mathbf{x}_i}{\nu_j}\right)}E(e^{-2W_i/\nu_j})\end{aligned}$$

Also using the "delta method", we observe that $g(W_i) = e^{-2W_i/\nu_j} \overset{\text{a}}{\sim} N\left[1; \frac{4\sigma_w^2}{\nu_j^2}\right]$, that is,

$$\begin{aligned}\text{Var}(T_{ji} \mid \mathbf{x}_i) = {} & \frac{\sigma_w^2\Gamma^2(1+1/\nu_j)}{\nu_j^2\exp\left(\frac{2\boldsymbol{\beta}_j'\mathbf{x}_i}{\nu_j}\right)} \\ & + \frac{1}{\exp\left(\frac{2\boldsymbol{\beta}_j'\mathbf{x}_i}{\nu_j}\right)}\left\{\Gamma(1+2/\nu_j) - \Gamma^2(1+1/\nu_j)\right\}\end{aligned} \tag{7}$$

Observe that not considering the presence of the "frailty" W_i, the variance for T_{ji}, given, $\mathbf{x}_i$ is

$$\text{Var}(T_{ji} \mid \mathbf{x}_i) = \frac{\Gamma(1+2/\nu_j) - \Gamma^2(1+1/\nu_j)}{\exp\left(\frac{2\boldsymbol{\beta}_j'\mathbf{x}_i}{\nu_j}\right)} \tag{8}$$

From (7) and (8), we observe that the extra-Weibull variability in the presence of the "frailty" W_i with normal distribution (2) is given by

$$\frac{\sigma_w^2\Gamma^2(1+1/\nu_j)}{\nu_j^2\exp\left(\frac{2\boldsymbol{\beta}_j'\mathbf{x}_i}{\nu_j}\right)}$$

for $j = 1, 2, \ldots, k$; $i = 1, 2, \ldots, n$.

A different model could be considered replacing (3) by

$$\lambda_j(i) = w_i e^{\boldsymbol{\beta}'\mathbf{x}_i} \tag{9}$$

with

$$W_i \overset{\text{iid}}{\sim} \text{Gamma}(\phi^{-1}, \phi^{-1}) \tag{10}$$

From (9), observe that $E(W_i) = 1$ and $\text{Var}(W_i) = 0$.

Let us assume the model defined by (1) and (9) as "model 2".

From "model 2", the conditional mean and variance for T_{ji} given $\mathbf{x}_i$ and w_i, are given, respectively, by,

$$E(T_{ji} \mid \mathbf{x}_i, w_i) = \frac{\Gamma(1 + 1/\nu_j)}{w_i^{1/\nu_j} e^{\boldsymbol{\beta}_j' \mathbf{x}_i/\nu_j}} \tag{11}$$

and

$$\text{Var}(T_{ji} \mid \mathbf{x}_i, w_i) = \frac{\Gamma(1 + 2/\nu_j) - \Gamma^2(1 + 1/\nu_j)}{w_i^{2/\nu_j} e^{2\boldsymbol{\beta}_j' \mathbf{x}_i/\nu_j}}$$

for $i = 1, 2, \ldots, n$; $j = 1, 2, \ldots, k$.

Following the same arguments used in the determination of the unconditional mean and variance for T_{ji} assuming "model 1", and observing that the "frailty" W_i has a Gamma (ϕ^{-1}, ϕ^{-1}) distribution, the unconditional mean for T_{ji} assuming "model 2" is (see section 6) given by

$$E(T_{ji} \mid \mathbf{x}_i) = \frac{\Gamma(1 + 1/\nu_j)(\phi^{-1})^{1/\nu_j} \Gamma(\phi^{-1} - \nu_j^{-1})}{\exp\{\boldsymbol{\beta}_j' \mathbf{x}_i/\nu_j\} \Gamma(\phi^{-1})}$$

for $\phi^{-1} > \nu_j^{-1}$, $i = 1, 2, \ldots, n$; $j = 1, 2, \ldots, k$.

The unconditional variance for T_{ji} (see Section 7) is given by,

$$\text{Var}(T_{ji} \mid \mathbf{x}_i) = \frac{(\phi^{-1})^{2/\nu_j}}{\exp(2\boldsymbol{\beta}_j' \mathbf{x}_i/\nu_j)} \times \left\{ \frac{\Gamma(1 + 2/\nu_j)\Gamma(\phi^{-1} - 2/\nu_j)}{\Gamma(\phi^{-1})} - \left[\frac{\Gamma(1 + 1/\nu_j)\Gamma(\phi^{-1} - 1/\nu_j)}{\Gamma(\phi^{-1})} \right]^2 \right\}$$

for $i = 1, 2, \ldots, n$; $j = 1, 2, \ldots, k$.

Generalization of "model 1" and "model 2" could be considered assuming that the covariate vector $\mathbf{x}_i$ also affect the shape parameter ν_j, that is, assuming the regression model $\nu_j(i) = \exp\{\boldsymbol{\alpha}_j' \mathbf{x}_i\}$, where $\boldsymbol{\alpha}_\mathbf{j} = (\alpha_{j1}, \alpha_{j2}, \ldots, \alpha_{jp})$ is another vector of regression parameters, $j = 1, 2, \ldots, k$. Let us denote these models as "model 3" and "model 4", respectively.

3. A Bayesian Analysis

Assuming lifetime in the presence of censored observations and a covariate vector $\mathbf{x} = (x_1, x_2, \ldots, x_p)'$, let us define an indicator variable for censoring or not censoring observations, by

$$\delta_{ji} = \begin{cases} 1 & \text{for observed lifetime} \\ 0 & \text{for censored observation} \end{cases} \tag{12}$$

Assuming "model 1" defined by (1), (2) and (3), the likelihood function is given by

$$f(\mathbf{t} \mid \mathbf{x}, \boldsymbol{\beta}, \boldsymbol{\nu}, \mathbf{w}) = \prod_{i=1}^{n} \prod_{j=1}^{k} \left[f(t_{ji} \mid \mathbf{x}_i, w_i)\right]^{\delta_{ji}} \left[S(t_{ji} \mid \mathbf{x}_i, w_i)\right]^{1-\delta_{ji}}$$

where $S(t_{ji} \mid \mathbf{x}_i, w_i)$ is the survival function defined by (5), $\boldsymbol{\beta} = (\beta_1, \beta_2, \ldots, \beta_k)$, $\boldsymbol{\beta}_j = (\beta_{j1}, \beta_{j2}, \ldots, \beta_{jp})$, $j = 1, 2, \ldots, k$; $\boldsymbol{\nu} = (\nu_1, \nu_2, \ldots, \nu_k)$, $\mathbf{w} = (w_1, w_2, \ldots, w_n)$, $\mathbf{x}_i = (x_{i1}, x_{i2}, \ldots, x_{ip})$, $i = 1, 2, \ldots, n$.

That is,

$$f(\mathbf{t} \mid \mathbf{x}, \boldsymbol{\beta}, \boldsymbol{\nu}, \mathbf{w}) = \prod_{i=1}^{n} \prod_{j=1}^{k} \nu_j^{\delta_{ji}} t_{ji}^{\delta_{ji}(\nu_j - 1)} \exp\left\{\delta_{ji}[w_i + \boldsymbol{\beta}_j' \mathbf{x}_i]\right\} \exp\{-t_{ji}^{\nu_j} e^{w_i + \boldsymbol{\beta}_j' \mathbf{x}_i}\} \tag{13}$$

Assuming "model 2", the likelihood function is (from (12) and (9)) given by

$$f(\mathbf{t} \mid \mathbf{x}, \boldsymbol{\beta}, \boldsymbol{\nu}, \mathbf{w}) = \prod_{i=1}^{n} \prod_{j=1}^{k} \nu_j^{\delta_{ji}} t_{ji}^{\delta_{ji}(\nu_j - 1)} w_i^{\delta_{ji}} e^{\delta_{ji} \boldsymbol{\beta}' \mathbf{x}_i} \exp\{-t_{ji}^{\nu_j} w_i e^{\boldsymbol{\beta}' \mathbf{x}_i}\}$$

For a hierarchical Bayesian analysis of "model 1", we assume in the first stage, the following prior distributions for the parameters:

$$\begin{aligned} \nu_j &\sim \text{Gamma}(a_j, b_j) \\ \beta_{jl} &\sim \text{N}(0; c_{jl}^2) \end{aligned} \tag{14}$$

where $j = 1, 2, \ldots, k$; $l = 1, 2, \ldots, p$; a_j, b_j, c_{jl} are known hyperparameters and $\text{Gamma}(a, b)$ denotes a gamma distribution with mean a/b and variance a/b^2.

In the second stage of the hierarchical Bayesian analysis, we assume a gamma prior distribution for σ_w^2, that is,

$$\sigma_w^2 \sim \text{Gamma}(d, e) \tag{15}$$

where d and e are known hyperparameters.

We further assume independence among the parameters.

Combining (2), (13), (14) and (15), we get the joint posterior distribution for $\mathbf{w}, \boldsymbol{\nu}, \boldsymbol{\beta}$ and σ_w^2, given by

$$\begin{aligned} \pi(\boldsymbol{\nu}, \boldsymbol{\beta}, \mathbf{w}, \sigma_w^2 \mid \mathbf{x}, \mathbf{t}) \propto & \left\{\prod_{i=1}^{n} \exp\left(-\frac{w_i^2}{2\sigma_w^2}\right)\right\} \left\{\prod_{j=1}^{k} \prod_{l=1}^{p} \exp\left(-\frac{\beta_{jl}^2}{2c_j^2}\right)\right\} \\ & \times (\sigma_w^2)^{d-1} \exp(-e\sigma_w^2) \left(\prod_{j=1}^{k} \nu_j^{a_j - 1} e^{-b_j \nu_j}\right) \\ & \times \prod_{i=1}^{n} \prod_{j=1}^{k} \nu_j^{\delta_{ji}} t_{ji}^{\delta_{ji}(\nu_j - 1)} \exp\{\delta_{ji}(w_i + \boldsymbol{\beta}_j' \mathbf{x}_i)\} \\ & \times \exp\{-t_{ji}^{\nu_j} e^{w_i + \boldsymbol{\beta}_j' \mathbf{x}_i}\} \end{aligned} \tag{16}$$

To get the posterior summaries of interest, we simulate samples of the joint posterior distribution (16) using MCMC methods as the popular Gibbs sampling algorithm (see for example, Gelfand & Smith 1990) or the Metropolis-Hastings algorithm (see for example, Chib & Greenberg 1995).

A great simplification in the simulation of the samples for the joint posterior distribution is given by the WinBugs software (Spiegelhalter, Thomas, Best & Lunn 2003), which requires only the specification of the joint distribution for the data and the prior distributions for the parameters.

Assuming "model 2", we consider the same priors (14) for ν_j and β_{jl}, and an uniform prior distribution for ϕ, that is,

$$\phi \sim \mathrm{U}(0, f) \tag{17}$$

where $U(a, b)$ denotes an uniform distribution in the interval (a, b) and f is a known hyperparameter.

The joint posterior distribution for $\mathbf{w}, \boldsymbol{\nu}, \boldsymbol{\beta}$ and ϕ is given by

$$\begin{aligned}\pi(\boldsymbol{\beta}, \boldsymbol{\beta}, \mathbf{w}, \phi \mid \mathbf{x}, \mathbf{t}) \propto & \left\{\prod_{i=1}^{n} w_i^{\phi^{-1}-1} e^{-\phi^{-1} w_i}\right\} \phi^{f-1} e^{-g\phi} \\ & \times \left\{\prod_{j=1}^{k}\prod_{l=1}^{p} \exp\left(-\frac{\beta_{jl}^2}{2c_j^2}\right)\right\} \left\{\prod_{j=1}^{k} \nu_j^{a_j-1} e^{-b_j\nu_j}\right\} \\ & \times \prod_{i=1}^{n}\prod_{j=1}^{k} \nu_j^{\delta_{ji}} t_{ji}^{\delta_{ji}(\nu_j-1)} w_i^{\delta_{ji}} \exp\{\delta_{ji}\boldsymbol{\beta}_j'\mathbf{x}_i\} \exp\{-t_{ji}^{\nu_j} w_i e^{\boldsymbol{\beta}_j'\mathbf{x}_i}\}\end{aligned}$$

4. Use of a Generalized Gamma Distribution for Multivariate Survival Data

In this section, we assume that the lifetime T_{ji} has a generalized gamma distribution with density

$$f(t_{ji} \mid \nu_j, \mu_j(i), \theta_j) = \frac{\theta_j}{\Gamma(\nu_j)} [\mu_j(i)]^{\theta_j\nu_j} t_{ji}^{\theta_j\nu_j - 1} \exp\left\{-[\mu_j(i) t_{ji}]^{\theta_j}\right\} \tag{18}$$

where $t_{ji} > 0$, $i = 1, 2, \ldots, n$; $j = 1, 2, \ldots, k$; $\theta_j > 0$; $\nu_j > 0$ and $\mu_j(i) > 0$.

The generalized gamma distribution is a fairly flexible family of distributions that includes as special cases the exponential ($\theta_j = \nu_j = 1$), Weibull ($\nu_j = 1$) and gamma ($\theta_j = 1$) distributions. The log-normal distribution also arises as a limiting form of (18), that is, the generalized gamma model includes as special cases all of the most commonly used lifetime distributions. This makes it useful for discriminating among these other models.

The survival function for a given value of T_{ji} is given by

$$S(t_{ji} \mid \nu_j, \mu_j(i), \theta_j) = P(T_{ji} > t_{ji}) = \frac{\theta_j}{\Gamma(\nu_j)} [\mu_j(i)]^{\theta_j\nu_j} \int_{t_{ji}}^{\infty} z^{\theta_j\nu_j - 1} e^{-[\mu_j(i) z]^{\theta_j}} dz$$

To capture the correlation among the repeated measures $T_{1i}, T_{2i}, \ldots, T_{ki}$ for the same individual, we introduce "frailties" W_i, $i = 1, 2, \ldots, n$. In the presence of a covariate vector $\mathbf{x}_i = (x_{i1}, x_{i2}, \ldots, x_{ip})'$, we assume the regression models

$$\mu_j(i) = \exp\{w_i + \boldsymbol{\beta}_j' \mathbf{x}_i\}$$

where W_i has a normal distribution (2) $i = 1, 2, \ldots, n$; $j = 1, 2, \ldots, k$, denoted as "model 5", or,

$$\mu_j(i) = w_i e^{\boldsymbol{\beta}_j' \mathbf{x}_i}$$

where W_i has a gamma distribution (10) denoted as "model 6".

Assuming "model 5", we consider the following prior distributions in a first stage of a hierarchical Bayesian analysis:

$$\begin{aligned} \nu_j &\sim \text{Gamma}(a_j, b_j); \\ \theta_j &\sim \text{Gamma}(c_j, d_j); \\ \beta_{jl} &\sim \text{N}(0, e_{jl}^2); \end{aligned} \tag{19}$$

where $j = 1, 2, \ldots, k$; $l = 1, 2, \ldots, p$; a_j, b_j, c_j, d_j and e_{jl} are known hyperparameters. In a second stage of the hierarchical Bayesian analysis, let us assume a gamma prior (15) for σ_w^2.

Assuming "model 6", we consider the same priors (19) for ν_j, θ_j and β_{jl}, and a gamma prior (17) for ϕ.

To develop a Bayesian analysis for the generalized gamma distribution of multivariate survival data in the presence of covariates and censored observations, we need informative prior distributions to get convergence for the Gibbs sampling algorithm. Observe that using the generalized gamma distribution usually we have great difficulties to get classical inferences of interest (see for example, Stacy & Mihram (1965), Parr & Webster (1965) and Hager & Bain (1970)).

Samples of the joint posterior distribution for the parameters of "model 3" or "model 4" are obtained using MCMC methods.

5. Model Selection

Different model selection methods could be used to choose the most adequate model to analyse multivariate survival data in the presence of covariates and censored observartions. As a special situation, we could use the generalized gamma distribution (see Section 4). In this way, if credible intervals for the parameters ν_j, $j = 1, 2, \ldots, k$ include the value one, this is an indication that the use of Weibull distribution in the presence of "frailties" gives good fit for the survival data.

We also could consider the Deviance Information Criterion (DIC), which is a criterion specifically useful for selection models under the Bayesian approach where samples of the posterior distribution for the parameters of the model are obtained using MCMC methods.

The deviance is defined by

$$D(\theta) = -2\log L(\theta) + c$$

where θ is a vector of unknown parameters of the model, $L(\theta)$ is the likelihood function of the model and c is a constant that does not need to be known when the comparison between models is made.

The DIC criterion defined by Spiegelhalter, Best, Carlin & Van der Linde (2002) is given by,

$$DIC = D(\widehat{\theta}) + 2n_D$$

where $D(\widehat{\theta})$ is the deviance evaluated at the posterior mean $\widehat{\theta} = E(\theta \mid data)$ and n_D is the effective number of parameters of the model given by $n_D = \overline{D} - D(\widehat{\theta})$, where $\overline{D} = E(D(\theta) \mid data)$ is the posterior deviance measuring the quality of the data fit for the model. Smaller values of DIC indicates better models. Note that these values could be negative.

6. Some Results About Gamma Distribution

Let W_i be a random variable with a Gamma(a, b) distribution, with density

$$f(w_i \mid a, b) = \frac{b^a}{\Gamma(a)} w_i^{a-1} e^{-bw_i} \tag{20}$$

where $w_i > 0$, $a > 0$, $b > 0$, $i = 1, 2, \ldots, n$.

From (20) we observe that

$$\int_0^\infty w_i^{a-1} e^{-bw_i} dw_i = \frac{\Gamma(a)}{b^a} \tag{21}$$

Also observe that

$$E(w_i^{-k}) = \int_0^\infty w_i^{-k} \frac{b^a}{\Gamma(a)} w_i^{a-1} e^{-bw_i} dw_i = \frac{b^a}{\Gamma(a)} \int_0^\infty w_i^{(a-k)-1} e^{-bw_i} dw_i$$

From (21), we have:

$$E(w_i^{-k}) = \frac{b^k \Gamma(a-k)}{\Gamma(a)}$$

for $a > k$.

Assuming $a = b = \phi^{-1}$, we have:

i) With $k = 1/\nu_j$,

$$E(w_i^{-1/\nu_j}) = \frac{(\phi^{-1})^{\nu_j^{-1}} \Gamma(\phi^{-1} - \nu_j^{-1})}{\Gamma(\phi^{-1})} \tag{22}$$

for $i = 1, 2, \ldots, n$; $j = 1, 2, \ldots, k$;

ii) With $k = 2/\nu_j$,

$$E(w_i^{-2/\nu_j}) = \frac{(\phi^{-1})^{2/\nu_j}\Gamma(\phi^{-1} - 2/\nu_j)}{\Gamma(\phi^{-1})} \tag{23}$$

for $i = 1, 2, \ldots, n$; $j = 1, 2, \ldots, k$.

From (10), we have:

$$E(T_{ji} \mid \mathbf{x}_i) = E[E(T_{ji} \mid \mathbf{x}_i, w_i)] = \frac{\Gamma(1 + 1/\nu_j)}{\exp\{\boldsymbol{\beta}_j'\mathbf{x}_i/\nu_j\}} E(W_i^{-1/\nu_j})$$

Thus, from (22), we find the unconditional mean for T_{ji}, given by

$$E(T_{ji} \mid \mathbf{x}_i) = \frac{\Gamma(1 + 1/\nu_j)(\phi^{-1})^{1/\nu_j}\Gamma(\phi^{-1} - 1/\nu_j)}{\Gamma(\phi^{-1})\exp\{\boldsymbol{\beta}_j'\mathbf{x}_i/\nu_j\}} \tag{24}$$

From (10) and (11) and using the result $\mathrm{Var}(T_{ji} \mid \mathbf{x}_i) = E\{\mathrm{Var}(T_{ji} \mid \mathbf{x}_i, w_i)\} + \mathrm{Var}\{E(T_{ji} \mid \mathbf{x}_i, w_i)\}$, we have:

$$\mathrm{Var}(T_{ji} \mid \mathbf{x}_i) = \frac{[\Gamma(1 + 2/\nu_j) - \Gamma^2(1 + 1/\nu_j)]}{\exp\{2\boldsymbol{\beta}_j'\mathbf{x}_i\}} E(W_i^{-2/\nu_j}) + \tag{25}$$

$$+ \frac{\Gamma^2(1 + 1/\nu_j)}{\exp\{2\boldsymbol{\beta}_j'\mathbf{x}_i/\nu_j\}} \mathrm{Var}(W_i^{-1/\nu_j}) \tag{26}$$

Observe that $\mathrm{Var}(W_i^{-1/\nu_j}) = E(W_i^{-2/\nu_j}) - [E(W_i^{-1/\nu_j})]^2$, that is, from (22) and (23),

$$\mathrm{Var}(W_i^{-1/\nu_j}) = \frac{(\phi^{-1})^{2/\nu_j}\Gamma(\phi^{-1} - 2/\nu_j)}{\Gamma(\phi^{-1})} - \frac{(\phi^{-1})^{2/\nu_j}\Gamma^2(\phi^{-1} - 1/\nu_j)}{\Gamma^2(\phi^{-1})}$$

That is, from (23) and (24), we find the unconditional variance for T_{ji} given by

$$\mathrm{Var}(T_{ji} \mid \mathbf{x}_i) = \frac{(\phi^{-1})^{2/\nu_j}}{\exp(2\boldsymbol{\beta}_j'\mathbf{x}_i/\nu_j)} \times$$
$$\times \left\{ \frac{\Gamma(1 + 2/\nu_j)\Gamma(\phi^{-1} - 2/\nu_j)}{\Gamma(\phi^{-1})} - \left[\frac{\Gamma(1 + 1/\nu_j)\Gamma(\phi^{-1} - 1/\nu_j)}{\Gamma(\phi^{-1})} \right]^2 \right\}$$

7. Analysis of the Recurrence Times of Infections for Kidney Patients

To analyse the recurrence times of infections (see Table 1), let us assume a Weibull regression model ("model 1") in the presence of a "frailty" W_i with a normal distribution (2).

In this case, we have only a covariate x_i (sex; $x_i = 1$ for male; $x_i = 0$ for female) and $k = 2$ recurrence times.

From (3), we have the regression model

$$\lambda_j(i) = \exp\{\beta_{1j} + \beta_{2j}x_i + w_i\}$$

$i = 1, 2, \ldots, 38$; $j = 1, 2$.

For a Bayesian analysis of "model 1", let us assume the prior distributions (14) and (15) with $a_1 = b_1 = a_2 = b_2 = 1$; $c_{11} = c_{21} = c_{12} = c_{22} = 10$ and $d = e = 0.1$.

Using the WinBugs software (Spiegelhalter et al. 2003), we discarded the first 5000 simulated Gibbs samples ("burn-in-sample") to eliminate the effect of the initial values for the parameters of the model. Choosing every 20^{th} simulated Gibbs sample, we obtained a final sample of size 2000 to get the posterior summaries of interest (see Table 2). Convergence of the Gibbs sampling algorithm was monitored using existing methods as time series plots for the simulated samples and Gelman & Rubin (1992) indexes. This simulation procedure also was employed for the other models considered in this section.

In Table 2, we also have the Monte Carlo estimate for the posterior mean of the median survival time in each recurrence time. Observe that the median survival time not including the covariate x_i is given by $\text{Med}_j = [(\log 2)e^{-\beta_{1j}}]^{1/\nu_j}$, $j = 1, 2$.

Table 2: Posterior summaries ("model 1").

Parameter	Mean	S.D.	95% Credible Interval
β_{21}	1.9820	0.5694	(0.8885; 3.1550)
β_{22}	0.7594	0.5570	(−0.3279; 1.8510)
β_{11}	−5.5490	0.8571	(−7.3400; −3.9880)
β_{12}	−5.6110	0.8805	(−7.4620; −3.9940)
med 1	120.40	31.780	(69.350; 194.50)
med 2	106.60	29.520	(62.720; 173.10)
ν_1	1.0880	0.1610	(0.8012; 1.4270)
ν_2	1.1310	0.1742	(0.8113; 1.5100)
$1/\sigma_w^2$	3.1350	3.1060	(0.7188; 11.270)

In Figure 1, we have the time series plots for the simulated Gibbs samples under "model 1". From these plots, we observe convergence of the algorithm in all cases.

Assuming "model 2", that is, defined by the Weibull density (1) where $\lambda_j(i)$ is given by (9), we have

$$\lambda_j(i) = w_i \exp\{\beta_{1j} + \beta_{2j}x_i\}$$

$i = 1, 2, \ldots, 38$; $j = 1, 2$. Let us assume the prior distributions (14) and (17) with $a_1 = b_1 = a_2 = b_2 = 1$; $c_{11} = c_{21} = c_{12} = c_{22} = 10$ and $f = 5$.

Following the same simulation steps considered in the generation of samples for the joint posterior distribution of the parameters of "model 1", we have, in Table 3, the posterior summaries of interest assuming the final Gibbs sample of size 2000.

In Figure 2, we have plots for the simulated Gibbs samples under "model 2".

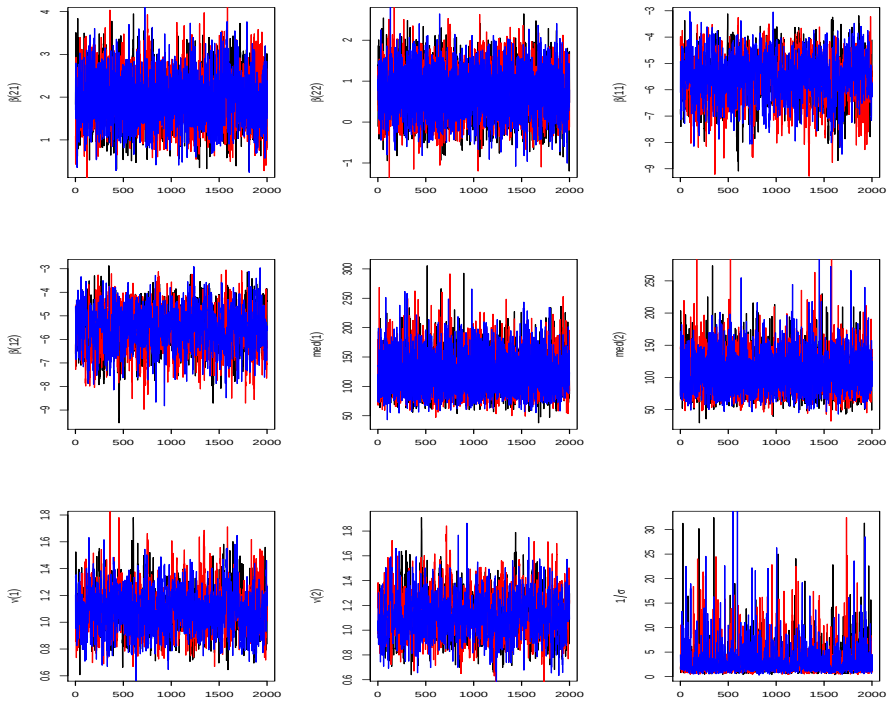

FIGURE 1: Simulated Gibbs samples ("model 1").

TABLE 3: Posterior summaries ("model 2").

Parameter	Mean	S.D.	95% Credible Interval
β_{21}	2.2370	0.6192	$(1.0770; 3.5080)$
β_{22}	1.0180	0.6234	$(-0.1979; 2.2800)$
β_{11}	-5.6340	0.8506	$(-7.3790; -4.1350)$
β_{12}	-5.6690	0.8794	$(-7.5240; -4.0420)$
med 1	98.590	26.260	$(54.860; 157.90)$
med 2	87.340	23.140	$(50.320; 141.90)$
ν_1	1.1560	0.1719	$(0.8508; 1.5130)$
ν_2	1.1950	0.1863	$(0.8604; 1.5900)$
$1/\phi$	2.4670	2.6360	$(0.7685; 7.7670)$

From the results of Tables 2 and 3, we observe similar results considering "model 1" and "model 2". We observe that in both models, we have a significative effect of sex for the first recurrence time, since zero is not included in the 95% credible interval for β_{21}; in the same way, we observe that sex does not have a significative effect in the second recurrence time, since zero is included in the 95% credible interval for β_{22}.

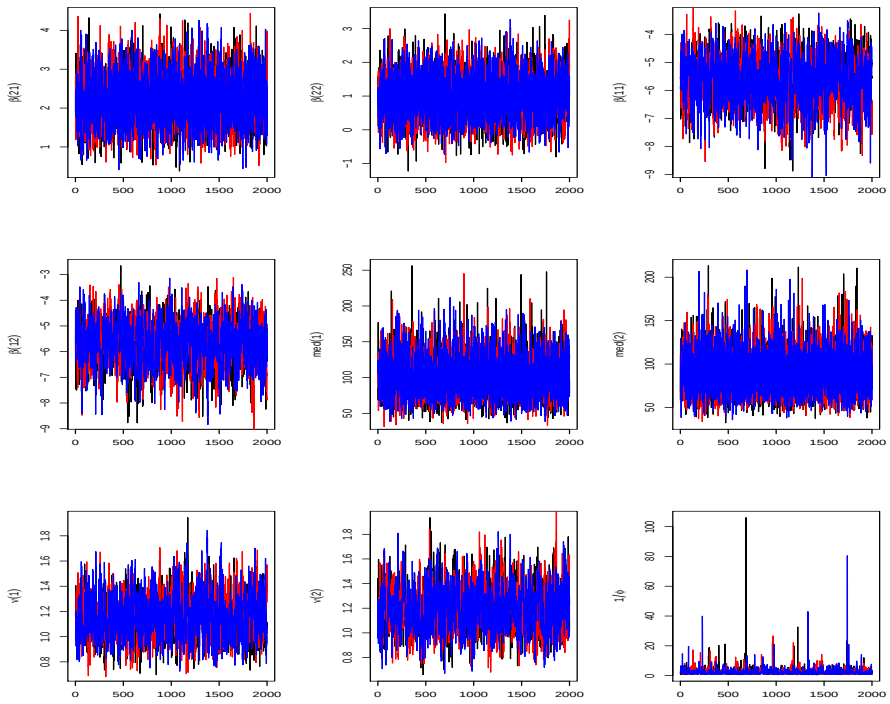

FIGURE 2: Simulated Gibbs samples ("model 2").

A Monte Carlo estimate for DIC (see Section 5), based on the 2000 simulated Gibbs samples considering "model 1", is given by DIC = 667.07. Considering "model 2", we have DIC = 662.86. That is, since we have a small decreasing in the value of DIC assuming "model 2", we could conclude that "model 2" is better fitted by the recurrence times of infection for kidney patients. To point out that other discrimination methods also could be used to decide by the best model is important.

A further modification could be assumed for "model 1" and "model 2", introducing the effect of covariate sex (x_i) in the shape parameter ν_j, $j = 1, 2$.

In this way, we assume for "model 1" and "model 2" the regression model for the shape parameter given by

$$\nu_j(i) = \exp\{\alpha_{1j} + \alpha_{2j} x_i\}$$

$i = 1, 2, \ldots, 38$; $j = 1, 2$.

Let us denote these models as "model 3" and "model 4".

For "model 3" and "model 4", we assume informative normal prior distributions for β_{1j} and β_{2j} considering means close to the obtained posterior means for β_{1j} and β_{2j} assuming "model 1" and "model 2", respectively. We also assume normal priors for α_{1j} and α_{2j}, $j = 1, 2$, considering small variances.

In Table 4, we have the posterior summaries obtained from 2000 simulated Gibbs samples for the joint posterior distributions of interest.

Table 4: Posterior summaries ("model 3" and "model 4").

Model	Parameter	Mean	S.D.	95% Credible Interval
"model 3"	β_{21}	2.0050	0.2932	$(1.4410; 2.5880)$
DIC = 660.87	β_{22}	1.9470	0.3007	$(1.3610; 2.5370)$
	β_{11}	-5.9650	0.2970	$(-6.5460; -5.3840)$
	β_{12}	-6.0650	0.2937	$(-6.6580; -5.4970)$
	α_{21}	0.0603	0.1186	$(-0.1818; 0.2850)$
	α_{22}	-0.1855	0.1180	$(-0.4252; 0.0356)$
	α_{11}	0.1395	0.0662	$(0.0064; 0.2670)$
	α_{12}	0.1899	0.0697	$(0.0452; 0.3159)$
	$1/\sigma_w^2$	1.7700	0.8131	$(0.6983; 3.8310)$
"model 4"	β_{21}	2.0170	0.3044	$(1.4210; 2.5960)$
DIC = 658.06	β_{22}	1.9510	0.3013	$(1.3610; 2.5440)$
	β_{11}	-5.9410	0.2938	$(-6.5280; -5.3700)$
	β_{12}	-6.0510	0.2921	$(-6.6460; -5.4680)$
	α_{21}	0.1142	0.1242	$(-0.1362; 0.3476)$
	α_{22}	-0.1263	0.1281	$(-0.3848; 0.1203)$
	α_{11}	0.1772	0.0696	$(0.0430; 0.3143)$
	α_{12}	0.2267	0.0691	$(0.0838; 0.3593)$
	$1/\phi$	2.1050	1.9220	$(0.7696; 5.5800)$

In Figures 3 and 4, we have plots for the simulated Gibbs samples considering "model 3" and "model 4", respectively.

From the results in Table 4, we observe that "model 3" and "model 4" give similar inferences. We observe that the covariate x_i (sex) does not have a significative effect on the shape parameter of the Weibull distribution for the recurrences times, since zero is included in the 95% credible intervals for α_{21} and α_{22} assuming both models. We also observe that "model 4" gives a smaller value for DIC (658.06) when compared to models 1, 2 and 3.

Another way, to check if the Weibull regression model is well fitted by the data, is to assume a generalized gamma distribution.

Considering "model 5" with a generalized gamma density (18) with regression model,

$$\mu_j(i) = \exp\{\beta_{1j} + \beta_{2j} x_i + w_i\}$$

where the "frailty" W_i has a normal distribution (2), let us assume the priors (19) and (15) for the parameters of the model with hyperparameter values $a_1 = a_2 = b_1 = b_2 = c_1 = c_2 = d_1 = d_2 = 1$ and normal distributions for β_{11}, β_{12}, β_{21} and β_{22} with variance equals to one.

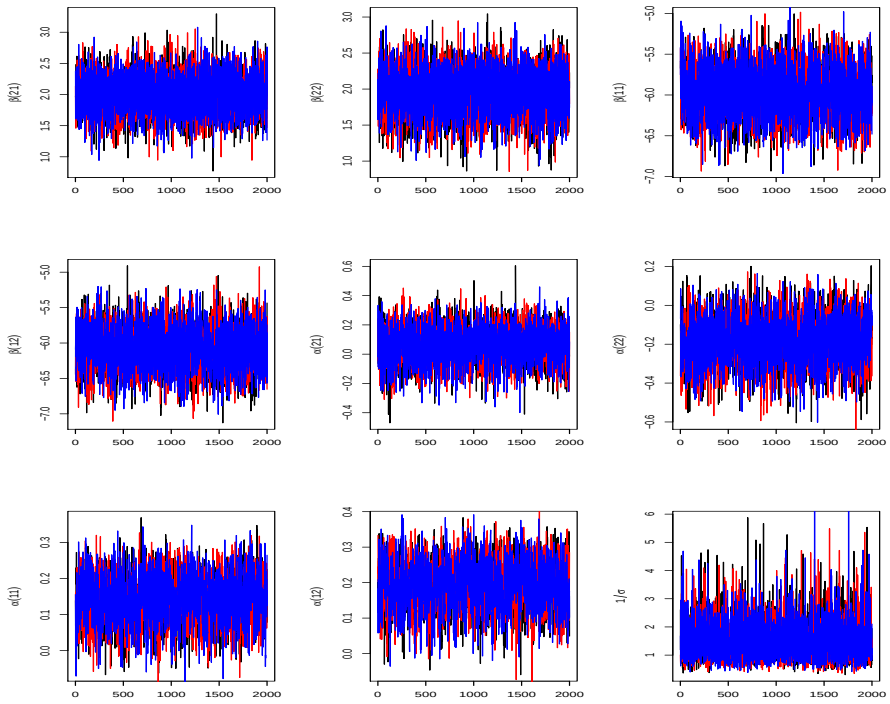

FIGURE 3: Simulated Gibbs samples ("model 3").

Assuming "model 6", with a generalized gamma density (18), and a regression model,

$$\mu_j(i) = w_i \exp\{\beta_{1j} + \beta_{2j} x_i\}$$

where the "frailty" W_i has a gamma distribution (10), let us assume the same prior distributions considered for "model 5", in the first stage of the hierarchical Bayesian analysis and a Gamma$(1, 1)$ prior for the parameter ϕ.

In Table 5, we have the posterior summaries of interest considering "model 5" and "model 6".

From the results of Table 5, we observe that assuming "model 5" or "model 6", the 95% credible intervals for ν_1 and ν_2 include the value one, that is, an indicator that the Weibull models in the presence of "frailties" give good fit for the multivariate survival data introduced in table 1.

8. Discussion and Concluding Remarks

Longitudinal survival data is common in many studies as in medicine or in engineering. Usually, we have repeated measures for the same patient or unit. In these studies, the presence of covariates and censoring data is common . The use of

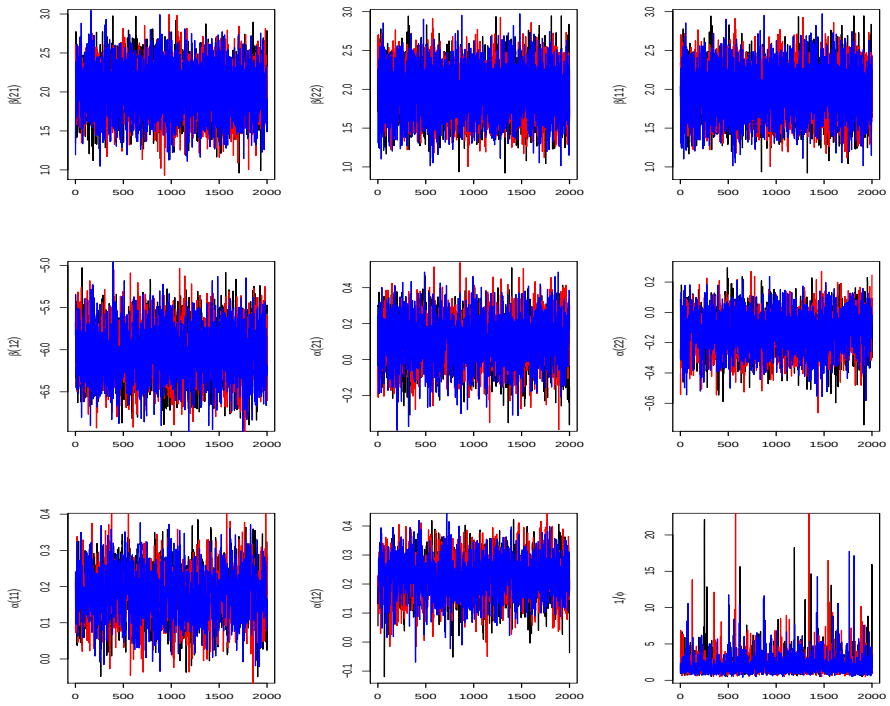

FIGURE 4: Simulated Gibbs samples ("model 4").

Bayesian hierarchical models with "frailties" or latent variables assuming different structures is a powerful way to get the inferences of interest.

Observe that considering independent survival times assuming Weibull distribution (1) and regression model (3) to analyse the survival data introduced in Table 1, we have the value of DIC given by 678.82 considering non-informative priors for the parameters of the model and the same Gibbs algorithm steps assumed for the other proposed models. That is, since DIC is larger assuming independent Weibull models, we have a great indication of the presence of a correlation structure for the survival data of Table 1.

In Table 6, we have the posterior summaries assuming independent Weibull models.

Since we have only a covariate x_i (sex; $x_i = 1$ for male and $x_i = 0$ for female), we can compare the obtained means and variances assuming independent Weibull distributions and "model 1" in the presence of a "frailty". Observe that for "model 1" we use the approximate formulas (6) and (7) for the unconditional means and variances for the survival times (see Table 7). In Table 7, we also have the sample means and sample variances for each combination sex versus response assuming only the uncensored data.

Table 5: Posterior summaries ("model 5" and "model 6").

Model	Parameter	Mean	S.D.	95% Credible Interval
"model 5"	β_{21}	1.8270	0.4312	$(0.9458; 2.6560)$
	β_{22}	0.8333	0.4353	$(-0.0450; 1.6790)$
	β_{11}	-5.1650	0.6007	$(-6.1090; -3.7340)$
	β_{12}	-4.7660	0.5579	$(-5.7620; -3.5470)$
	θ_1	1.4920	0.7905	$(0.6186; 3.7110)$
	θ_2	1.3930	0.6994	$(0.6453; 3.3830)$
	ν_1	0.9544	0.5290	$(0.2398; 2.2420)$
	ν_2	1.2580	0.5967	$(0.3553; 2.6830)$
	$1/\sigma_w^2$	1.9060	0.7860	$(0.8513; 3.9330)$
"model 6"	β_{21}	1.9010	0.4240	$(1.0380; 2.6750)$
	β_{22}	0.9467	0.4284	$(0.1055; 1.7830)$
	β_{11}	-4.8950	0.6403	$(-5.8470; -3.3000)$
	β_{12}	-4.6930	0.5428	$(-5.6470; -3.4300)$
	θ_1	1.3900	0.7163	$(0.6086; 3.3030)$
	θ_2	1.4630	0.6724	$(0.6761; 3.2120)$
	ν_1	1.0330	0.5737	$(0.2642; 2.4950)$
	ν_2	1.1450	0.5517	$(0.3419; 2.5530)$
	$1/\phi$	2.8710	2.1210	$(1.1220; 7.3010)$

Table 6: Posterior summaries (independent Weibull model).

Parameter	Mean	S.D.	95% Credible Interval
β_{21}	1.5540	0.4271	$(0.6873; 2.3750)$
β_{22}	0.2536	0.4351	$(-0.6191; 1.1000)$
β_{11}	-4.8710	0.7078	$(-6.3190; -3.5540)$
β_{12}	-4.8880	0.7298	$(-6.4030; -3.5710)$
med 1	123.80	31.640	$(71.730; 193.80)$
med 2	104.10	28.470	$(61.170; 171.00)$
ν_1	0.9388	0.1238	$(0.7112; 1.1960)$
ν_2	0.9792	0.1392	$(0.7279; 1.2720)$

Table 7: Means and variances ("model 1" and independent Weibull distributions).

	data without censoring		independent Weibull		"model 1"	
	sample mean	sample var	mean	var	unc mean	unc var
$(x = 1)$, resp 1	32.8	2052.09	34.36	1338.18	25.68	565.55
$(x = 0)$, resp 1	162.2	28358.6	186.95	39619.1	86.22	14307.6
$(x = 1)$, resp 2	105.8	36214.1	117.29	14512.9	69.76	3836.03
$(x = 0)$, resp 2	115.9	10609.0	148.36	23243.7	136.51	14658.7

(resp=response; unc=unconditional; male $(x = 1)$; female $(x = 0)$)
var = variance

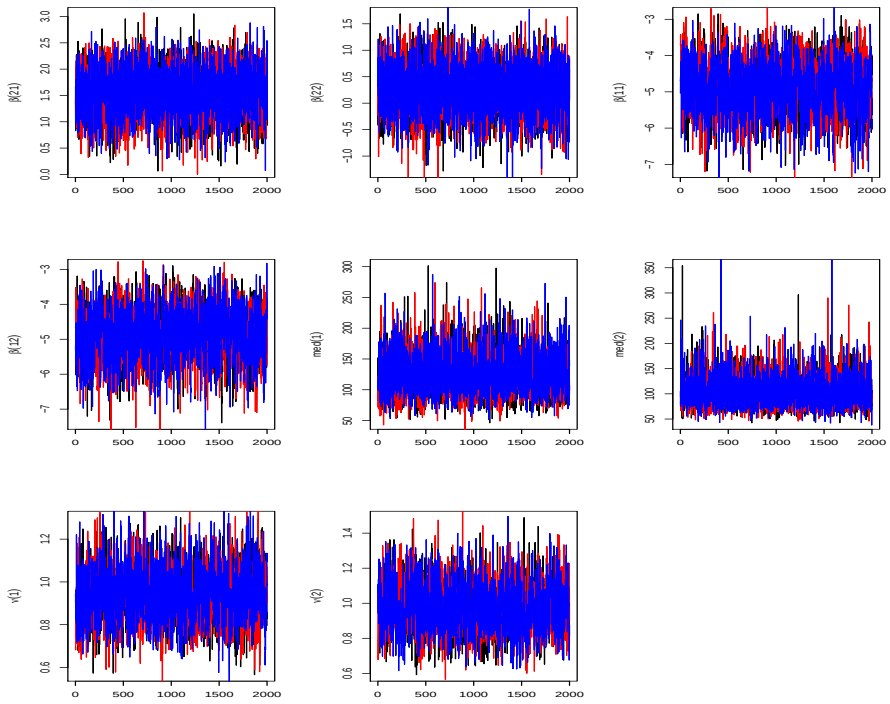

FIGURE 5: Simulated Gibbs samples independent Weibull model.

From the results of Table 7, we observe that the variances of the survival times have a great influence of the presence of the "frailty". Also to point out that these differences could be affected by the sample sizes for each class sex x response is important.

Acknowledgments

The authors would like to thank the editor and referees for their helpful comments.

[Recibido: julio de 2009 — Aceptado: enero de 2011]

References

Chib, S. & Greenberg, E. (1995), 'Understanding the Metropolis-Hastings algorithm', *The American Statistician* (49), 327–335.

Clayton, D. (1991), 'A Monte Carlo Method for Bayesian inference in frailty models', *Biometrics* (47), 467–485.

Clayton, D. & Cuzick, J. (1985), 'Multivariate generalizations of the proportional hazards model', *Journal of the Royal Statistical Society* **A**(148), 82–117.

Cox, D. R. (1972), 'Regression models and life tables', *Journal of the Royal Statistical Society* **B**(34), 187–220.

Cox, D. R. & Oakes, D. (1984), *Analysis of Survival Data*, Chapman & Hall, London.

Gelfand, A. E. & Smith, A. F. M. (1990), 'Sampling based approaches to calculating marginal densities', *Journal of the American Statistical Association* (85), 398–409.

Gelman, A. & Rubin, D. B. (1992), 'Inference from iterative simulation using multiple sequences (with discussion)', *Statistical Science* **7**(4), 457–472.

Hager, H. W. & Bain, L. J. (1970), 'Inferential procedures for the generalized gamma distribution', *Journal of the American Statistical Association* (65), 1601–1609.

Kalbfleisch, J. D. (1978), 'Nonparametric Bayesian analysis of survival time data', *Journal of the Royal Statistical Society* **B**(40), 214–221.

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

McGilchrist, C. A. & Aisbett, C. W. (1991), 'Regression with frailty in survival analysis', *Biometrics* (47), 461–466.

Oakes, D. (1986), 'Semiparametric inference in a model for association in bivariate survival data', *Biometrika* **73**, 353–361.

Oakes, D. (1989), 'Bivariate survival models induced by frailties', *Journal of the American Statistical Association* **84**, 487–493.

Parr, V. B. & Webster, J. T. (1965), 'A method for discriminating between failure density functions used in reliability predictions', *Technometrics* (7), 1–10.

Shih, J. A. & Louis, T. A. (1992), Models and analysis for multivariate failure time data, Technical report, Division of Biostatistics, University of Minnesota.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P. & Van der Linde, A. (2002), 'Bayesian measures of model complexity and fit (with discussion)', *Journal of the Royal Statistical Society* **B**(64), 583–639.

Spiegelhalter, D. J., Thomas, A., Best, N. G. & Lunn, D. (2003), *WinBugs version 1.4 user manual*, Institute of Public Health and Department of Epidemiology & Public Health, London.
*http://www.mrc-bsu.com.ac.uk/bugs

Stacy, E. W. & Mihram, G. A. (1965), 'Parameter estimation for a generalized gamma distribution', *Technometrics* (7), 349–358.

Weibull, W. (1951), 'A statistical distribution function of wide applicability', *Journal of Applied Mechanics* **18**(3), 292–297.