# UNIVERSIDADE DE SÃO PAULO

## REGRESSION MODELS FOR BIVARIATE SURVIVAL DATA: A BAYESIAN APPROACH

JORGE A. ACHCAR
ROSELI A. LEANDRO

Nº 34

## NOTAS

*Instituto de Ciências Matemáticas de São Carlos*

Instituto de Ciências Matemáticas de São Carlos

ISSN - 0103-2577

# REGRESSION MODELS FOR BIVARIATE SURVIVAL DATA: A BAYESIAN APPROACH

**JORGE A. ACHCAR**
**ROSELI A. LEANDRO**

**Nº 34**

NOTAS DO ICMSC
Série Estatística

São Carlos
Dez./1996

# Resumo

Em muitas aplicações podemos ter algumas variáveis explanatórias afetando o tempo de vida de um sistema com dois componentes $X$ e $Y$. Neste artigo apresentamos inferências Bayesianas para modelos de regressão considerando dados de sobrevivência bivariados e a distribuição ACBVE (ver BLOCK and BASU, 1974) usando algoritmos Metropolis-com-Gibbs Sampling.

**Palavras-chaves**: distribuição exponencial bivariada, modelo de regressão, Gibbs Sampling, algoritmo de Metropolis-Hastings, Confiabilidade.

# Regression Models For Bivariate Survival Data: A Bayesian Approach

Jorge A. Achcar * Roseli A. Leandro †

University of São Paulo, Brazil

**Abstract**

Usually in problems of lifetime testing, we may have some explanatory or independent variables affecting the lifetimes $X$ and $Y$ of a two-component system. In this paper, we present Bayesian inferences for regression models considering bivariate survival data and the ACBVE distribution (see BLOCK and BASU, 1974) and using Metropolis-within-Gibbs algorithms.

**Key words**: Bivariate exponential distribution, Regression Model, Gibbs Sampling, Metropolis-Hastings algorithm, Reliability.

# 1 Introduction

In many applications of life testing, we usually have two lifetimes $X$ and $Y$ associated to each unit. Among the different existing bivariate lifetime models to be used in these applications (see for example, FREUND, 1961; MARSHALL and OLKIN, 1967; SARKAR, 1987, BLOCK and BASU, 1974) one model has been very well explored in the literature: the BLOCK and BASU (1974) exponential distribution.

* ICMSC, Caixa Postal 668, 13560-970, São Carlos, S.P., Brazil.

† ESALQ, Caixa Postal 11, 13418-900, Piracicaba,S.P., Brazil.

The bivariate exponential distribution of BLOCK and BASU (ACBVE) with parameters $\lambda_1$, $\lambda_2$ and $\lambda_3$, for the lifetimes $X$ and $Y$ has a joint density function given by

$$f(x,y)=\begin{cases} f_1(x,y)=\dfrac{\lambda\lambda_1\lambda_{23}}{\lambda_{12}}exp\{-\lambda_1 x-\lambda_{23}y\} & \text{if } x<y \\ \\ f_2(x,y)=\dfrac{\lambda\lambda_2\lambda_{13}}{\lambda_{12}}exp\{-\lambda_{13}x-\lambda_2 y\} & \text{if } x\geq y \end{cases} \tag{1}$$

where

$$\lambda_{12}=\lambda_1+\lambda_2,\ \lambda_{13}=\lambda_1+\lambda_3,\ \lambda_{23}=\lambda_2+\lambda_3\ and\ \lambda=\lambda_1+\lambda_2+\lambda_3.$$

The joint generating function for the ACBVE distribution is given by

$$m(s,t)=E(e^{sX+tY})=\frac{\lambda}{\lambda_{12}(\lambda-t-s)}\left\{\frac{\lambda_1\lambda_{23}}{\lambda_{23}-t}+\frac{\lambda_2\lambda_{13}}{\lambda_{13}-s}\right\} \tag{2}$$

From (2), we get moments of interest for $X$ e $Y$; thus, the means and variances for $X$ and $Y$ are given by,

$$\begin{aligned} E(X) &= \frac{1}{\lambda_{13}}+\frac{\lambda_2\lambda_3}{\lambda\lambda_{12}\lambda_{13}} \\ E(Y) &= \frac{1}{\lambda_{23}}+\frac{\lambda_1\lambda_3}{\lambda\lambda_{12}\lambda_{23}} \qquad (3) \\ \sigma_X^2=var(X) &= \frac{1}{\lambda_{13}^2}+\frac{\lambda_2\lambda_3(2\lambda_1\lambda+\lambda_2\lambda_3)}{\lambda^2\lambda_{12}^2\lambda_{13}^2} \end{aligned}$$

$$\sigma_Y^2 = var(Y) \quad = \quad \frac{1}{\lambda_{23}^2} + \frac{\lambda_1\lambda_3(2\lambda_2\lambda + \lambda_1\lambda_3)}{\lambda^2\lambda_{12}^2\lambda_{23}^2}$$

The correlation coefficient for $X$ and $Y$ is given by

$$\rho_{XY} = \frac{\lambda_3[(\lambda_1^2 + \lambda_2^2)\lambda + \lambda_1\lambda_2\lambda_3]}{\phi_1\phi_2} \tag{4}$$

where

$$\phi_1 = \quad [\lambda_{12}^2\lambda_{13}^2 + \lambda_2(\lambda_2 + 2\lambda_1)\lambda^2]^{1/2}$$

$$\phi_2 = \quad [\lambda_{12}^2\lambda_{23}^2 + \lambda_1(\lambda_1 + 2\lambda_2)\lambda^2]^{1/2}$$

Observe that $0 \leq \rho_{XY} \leq 1$ and $\rho_{XY} = 0$ only for the trivial cases $\lambda_3 = 0$ or $\lambda_1 = \lambda_2 = 0$.

Usually, researchers consider the use of standard asymptotic results based on the normality of the maximum likelihood estimators to get inferences for the parameters of the ACBVE distribution (1), but in general these asymtotic results can be very poor for small or moderate sample sizes.

A Bayesian analysis of the ACBVE model (1) is introduced by ACHCAR and SANTANDER (1993) using non-informative prior densities for the parameters and Laplace's method of approximation for integrals (see for example, TIERNEY and KADANE, 1986) to get the posterior summaries of interest.

ACHCAR and LEANDRO (1996) present Bayesian inferences for the ACBVE distribution (1) using Metropolis-within-Gibbs algorithms (see for example, CHIB and GREENBERG, 1995). They also consider the use of Gibbs sampling to develop Bayesian inference for accelerated life tests assuming a power rule model and the ACBVE distribution.

In this paper, we present Bayesian inferences for regression models with the ACBVE distribution (1) using Metropolis-within-Gibbs algorithms.

## 2 A Regression Model for Bivariate Lifetime Data with the ACBVE Distribution

Usually, in many applications we can have $K$ explanatory or independent variables affecting the lifetimes $X$ and $Y$ of a two-component system. In this case, let us assume that the two-component lifetimes $X$ and $Y$ has a ACBVE distribution (1) and consider a regression model given by

$$
\begin{aligned}
\lambda_{1j} &= c_1 \exp(\beta_1 V_{1j} + \beta_2 V_{2j} + \ldots + \beta_K V_{Kj}) \\
\lambda_{2j} &= c_2 \exp(\beta_1 V_{1j} + \beta_2 V_{2j} + \ldots + \beta_K V_{Kj}) \qquad (5) \\
\lambda_{3j} &= c_3 \exp(\beta_1 V_{1j} + \beta_2 V_{2j} + \ldots + \beta_K V_{Kj})
\end{aligned}
$$

where $j = 1, \cdots, J$ are $J$ different levels of the explanatory variables $V_1, \cdots, V_K$; $c_1, c_2, c_3, \beta_1, \ldots, \beta_k$ are regression parameters.

Assume that for the $j^{th}$ fixed level of the explanatory vector $(V_{1j}, \cdots, V_{Kj})$, $j = 1, \cdots, J$, we have $n_j$ replicates of bivariate lifetimes $(X_{1j}, Y_{1j})$, $\cdots$, $(X_{n_j j}, Y_{n_j j})$. From (1) and (5), the likelihood function for $c_1, c_2, c_3, \beta_1, \cdots, \beta_K$ is given by

$$L_j(c_1, c_2, c_3, \beta_1, \beta_2, \ldots, \beta_K) = \prod_{i=1}^{n_j} f_1^{\delta_{ij}}(X_{ij}, Y_{ij}) f_2^{1-\delta_{ij}}(X_{ij}, Y_{ij}) \tag{6}$$

where $\delta_{ij} = 1$ if $X_{ij} < Y_{ij}$ and $\delta_{ij} = 0$ if $X_{ij} \geq Y_{ij}$, and

$$f_1(X_{ij}, Y_{ij}) = \frac{c_1 c_{123} c_{23}}{c_{12}} \exp\{2 \sum_{k=1}^{K} \beta_k V_{kj}\} \exp\{-(c_1 X_{ij} + c_{23} Y_{ij}) \exp\{\textstyle\sum_{k=1}^{K} \beta_k V_{kj}\}\} \quad \text{if} \quad X_{ij} < Y_{ij} \tag{7}$$

$$f_2(X_{ij}, Y_{ij}) = \frac{c_2 c_{123} c_{13}}{c_{12}} \exp\{2 \sum_{k=1}^{K} \beta_k V_{kj}\} \exp\{-(c_{13} X_{ij} + c_2 Y_{ij}) \exp\{\textstyle\sum_{k=1}^{K} \beta_k V_{kj}\}\} \quad \text{if} \quad X_{ij} \geq Y_{ij}$$

where $c_{12} = c_1 + c_2$, $c_{13} = c_1 + c_3$, $c_{23} = c_2 + c_3$ and $c_{123} = c_1 + c_2 + c_3$.

That is,

$$L_j(c_1, c_2, c_3, \beta_1, \beta_2, \ldots, \beta_K) = \frac{c_1^{r_j} c_{123}^{n_j} c_{23}^{r_j} c_2^{n_j - r_j} c_{13}^{n_j - r_j}}{c_{12}^{n_j}} \exp\{2 n_j \sum_{k=1}^{K} \beta_k V_{kj}\} \exp\{-(c_1 n_j \bar{X}_j + c_2 n_j \bar{Y}_j + c_3 R_j) \exp\{\textstyle\sum_{k=1}^{K} \beta_k V_{kj}\}\} \tag{8}$$

where $n_j \bar{X}_j = \sum_{i=1}^{n_j} X_{ij}$; $n_j \bar{Y}_j = \sum_{i=1}^{n_j} Y_{ij}$; $R_j = \sum_{i=1}^{n_j} [\delta_{ij} Y_{ij} + (1 - \delta_{ij}) X_{ij}]$

Assuming that the data obtained for the $J$ levels of the explanatory vector are independent, the likelihood function for $c_1, c_2, c_3, \beta_1, \cdots, \beta_K$ is given by,

$$L(c_1, c_2, c_3, \beta_1, \beta_2, \ldots, \beta_K) = \prod_{i=1}^{J} L_j(c_1, c_2, c_3, \beta_1, \beta_2, \ldots, \beta_K) \tag{9}$$

That is,

$$L(c_1, c_2, c_3, \beta_1, \beta_2, \ldots, \beta_K) = \frac{c_1^r c_{123}^n c_{23}^r c_2^{n-r} c_{13}^{n-r}}{c_{12}^n}$$

$$\exp(Z(\boldsymbol{\theta})) \exp\{-(c_1 S_X(\boldsymbol{\theta}) + c_2 S_Y(\boldsymbol{\theta}) + c_3 T(\boldsymbol{\theta})\} \tag{10}$$

where

(i) $r = \sum_{j=1}^{J} r_j, \quad n = \sum_{j=1}^{J} n_j$

(ii) $Z(\boldsymbol{\theta}) = 2n \sum_{k=1}^{K} \beta_k V_{kj}$,

(iii) $S_X(\boldsymbol{\theta}) = \sum_{j=1}^{J} n_j \bar{X}_j \exp\{\sum_{k=1}^{K} \beta_k V_{kj}\}$,

(iv) $S_Y(\boldsymbol{\theta}) = \sum_{j=1}^{J} n_j \bar{Y}_j \exp\{\sum_{k=1}^{K} \beta_k V_{kj}\}$,

(v) $T(\boldsymbol{\theta}) = \sum_{j=1}^{J} R_j \exp\{\sum_{k=1}^{K} \beta_k V_{kj}\}$.

Considering the introduction of a latent variable $N_1$ representing the number of observations such that $X_{i1} < Y_{i1}$, we assume the following prior densities for $N_1$, $c_1$, $c_2$, $c_3$, $\beta_k$, $k = 1, \cdots, K$:

$$N_1 \sim b(n_1, c_1/c_{12})$$

$$c_1 \sim \Gamma(a_1, b_1), \quad a_1 \textit{ and } b_1 \quad \text{Known}$$

$$c_2 \sim \Gamma(a_2, b_2), \quad a_2 \textit{ and } b_2 \quad \text{Known}$$

$$c_3 \sim \Gamma(a_3, b_3), \quad a_3 \textit{ and } b_3 \quad \text{Known}$$

$$\beta_k \sim N(\mu_{0k}, \sigma_{0k}^2), \quad \mu_{0k} \text{ and } \sigma_{0k}^2 \quad \text{Known for } k = 1, \ldots, K. \tag{11}$$

where $b(n_1, c_1/c_{12})$ denotes a binomial distribution with mean $n_1 c_1/c_{12}$, $c_1/c_{12} = P(X_{i1} < Y_{i1})$; $\Gamma(a_i, b_i)$ denotes a gamma distribution with mean $a_i/b_i$ and variance $a_i/b_i^2$, and $N(\mu_{0k}, \sigma_{0k}^2)$ denotes a normal distribution with mean $\mu_{0k}$ and variance $\sigma_{0k}^2$. We further assume independence among the parameters.

The joint posterior density is,

$$\begin{aligned}
&\pi(N_1, c_1, c_2, c_3, \beta_1, \ldots, \beta_K \mid \mathcal{D}) \propto \\
&\binom{n_1}{N_1} \left(\frac{c_1}{c_{12}}\right)^{N_1} \left(\frac{c_2}{c_{12}}\right)^{n_1 - N_1} \\
&\frac{c_1^{r+a_1-1} c_2^{n-r+a_2-1} c_3^{a_3-1} c_{23}^{r} c_{13}^{n-r} c_{123}^{n}}{c_{12}^{n}} \\
&\prod_{i=1}^{K} \exp\left\{-\frac{1}{2\sigma_{0i}^2}(\beta_i - \mu_{0i})^2\right\} \\
&\exp(Z(\boldsymbol{\theta})) \exp\{-[b_1 + S_X(\boldsymbol{\theta})]c_1 - [b_2 + S_Y(\boldsymbol{\theta})]c_2 - [b_3 + T(\boldsymbol{\theta})]c_3\}
\end{aligned} \tag{12}$$

where $\mathcal{D}$ denote the data set.

The marginal conditional densities for the Gibbs algorithm are given by,

$$N_1 \mid c_1, c_2, c_3, \beta_1, \ldots, \beta_K, \mathcal{D} \sim b(n_1, \frac{c_1}{c_{12}})$$

$$\pi(c_1 \mid N_1, c_2, c_3, \beta_1, \ldots, \beta_K, \mathcal{D}) \propto$$

$$\frac{c_1^{N_1+r+a_1-1} c_{13}^{n-r} c_{123}^{n}}{c_{12}^{n+n_1}} \exp\{-[b_1 + S_X(\boldsymbol{\theta})]c_1\} \tag{13}$$

$$\pi(c_2 \mid N_1, c_1, c_3, \beta_1, \ldots, \beta_K, \mathcal{D}) \propto$$

$$\frac{c_2^{n_1-N_1+n-r+a_2-1} c_{23}^{r} c_{123}^{n}}{c_{12}^{n+n_1}} \exp\{-[b_2 + S_Y(\boldsymbol{\theta})]c_2\}$$

$$\pi(c_3 \mid N_1, c_1, c_2, \beta_1, \ldots, \beta_K, \mathcal{D}) \propto$$

$$c_3^{a_3-1} c_{23}^{r} c_{13}^{n-r} c_{123}^{n} \exp\{-[b_3 + T(\boldsymbol{\theta})]c_3\}$$

$$\pi(\beta_i \mid N_1, c_1, c_2, c_3, \beta_1, \ldots, \beta_{i-1}, \beta_{i+1}, \ldots, \beta_K, \mathcal{D}) \propto$$

$$\exp(Z(\boldsymbol{\theta})) \exp\{-\frac{1}{2\sigma_{0i}^2}(\beta_i - \mu_{0i})^2\}$$

$$\exp\{-c_1 S_X(\boldsymbol{\theta}) - c_2 S_Y(\boldsymbol{\theta}) - c_3 T(\boldsymbol{\theta})\}$$

where $i = 1, \cdots, K$.

Observe that, we need to use the Metropolis-Hastings algorithm to generate the variables $c_1, c_2, c_3$, $\beta_1$, $\cdots$, $\beta_k$, $k = 1, \ldots, K$

# 3 A numerical Ilustration

In table (1), we have 90 bivariate observations $(X, Y)$ generated

Table 1: Generated bivariate life time data

| $X_{i1}$ | $Y_{i1}$ | $X_{i2}$ | $Y_{i2}$ | $X_{i3}$ | $Y_{i3}$ |
|---|---|---|---|---|---|
| $V_{11}=-10$, | $V_{21}=10$ | $V_{12}=0$, | $V_{22}=10$ | $V_{13}=30$, | $V_{23}=10$ |
| 57.7766 | 50.6024 | 102.9000 | 95.7546 | 44.7246 | 37.5118 |
| 30.4980 | 23.1823 | 102.2590 | 95.1141 | 1.8376 | 11.3359 |
| 15.4739 | 7.8367 | 31.0630 | 23.7533 | 13.2085 | 5.4753 |
| 79.2714 | 72.1200 | 17.9920 | 26.5806 | 35.7469 | 28.4910 |
| 26.0849 | 18.7102 | 12.0230 | 4.2296 | 56.6761 | 49.4997 |
| 31.3256 | 39.7426 | 16.7710 | 9.1805 | 62.4791 | 55.3127 |
| 12.7805 | 21.5171 | 9.2560 | 1.2895 | 21.3563 | 13.8923 |
| 24.6601 | 17.2619 | 16.2570 | 24.8867 | 9.4380 | 1.4846 |
| 60.5910 | 53.4218 | 47.7180 | 40.5175 | 10.5767 | 2.6998 |
| 9.7454 | 1.8138 | 5.6490 | 14.7689 | 25.5447 | 34.0128 |
| $X_{i4}$ | $Y_{i4}$ | $X_{i5}$ | $Y_{i5}$ | $X_{i6}$ | $Y_{i6}$ |
| $V_{14}=-10$, | $V_{24}=30$ | $V_{15}=0$, | $V_{25}=30$ | $V_{16}=-10$, | $V_{26}=30$ |
| 3.8839 | 13.1613 | 25.5541 | 18.1710 | 14.2563 | 6.5701 |
| 27.2439 | 19.8866 | 27.5553 | 20.2024 | 9.3858 | 1.4287 |
| 0.6602 | 10.3194 | 48.0861 | 40.8864 | 11.0007 | 3.1498 |
| 13.5954 | 22.3037 | 6.8127 | 15.8480 | 24.0162 | 16.6065 |
| 3.7031 | 12.9936 | 8.2092 | 0.1604 | 11.0740 | 3.2276 |
| 14.7652 | 7.1002 | 11.3058 | 3.4729 | 14.2036 | 6.5152 |
| 34.6080 | 27.3328 | 12.7174 | 4.9604 | 14.7183 | 7.0514 |
| 21.4295 | 29.9532 | 23.2775 | 15.8538 | 10.7445 | 2.8781 |
| 9.2034 | 1.2329 | 21.8704 | 14.4178 | 30.5760 | 38.9983 |
| 4.9006 | 14.0814 | 4.7548 | 13.9481 | 8.6363 | 0.6224 |
| $X_{i7}$ | $Y_{i7}$ | $X_{i8}$ | $Y_{i8}$ | $X_{i9}$ | $Y_{i9}$ |
| $V_{17}=-10$, | $V_{27}=50$ | $V_{18}=0$, | $V_{28}=50$ | $V_{19}=-10$, | $V_{29}=50$ |
| 9.2096 | 1.23963 | 2.0914 | 11.5584 | 12.2095 | 20.9673 |
| 11.4386 | 3.61338 | 9.6654 | 1.7281 | 8.5057 | 0.4814 |
| 8.6549 | 0.64245 | 11.1968 | 3.3576 | 21.4919 | 14.0309 |
| 10.2216 | 2.32202 | 14.2027 | 6.5143 | 10.1247 | 2.2187 |
| 9.2695 | 1.30394 | 13.1331 | 5.3963 | 12.6162 | 4.8541 |
| 10.2223 | 2.32271 | 8.0696 | 0.0091 | 8.1856 | 0.1349 |
| 14.1025 | 6.40972 | 9.9314 | 2.0125 | 9.4584 | 1.5065 |
| 8.4783 | 0.45175 | 12.8646 | 5.1148 | 10.6680 | 2.7968 |
| 10.9192 | 3.06351 | 11.7797 | 3.9736 | 8.7267 | 0.7199 |
| 14.3369 | 6.65412 | 8.7661 | 0.7624 | 8.3439 | 0.3064 |

from a ACBVE distribution (1) and the regression model (5) with two explanatory variables $V_1$ and $V_2$, assuming $\beta_1 = 0.01$, $\beta_2 = 0.05$,$c_1 = 0.002$, $c_2 = 0.006$ and $c_3 = 0.009$

Considering from (11) the priors $N_1 \sim b(n_1, c_1/c_{12})$, $c_1 \sim \Gamma(a_1, b_1)$, $c_2 \sim \Gamma(a_2, b_2)$, $c_3 \sim \Gamma(a_3, b_3)$ and $\beta_k \sim N(\mu_{0k}, \sigma_{0k})$, $n_1, a_1, b_1, a_2, b_2, a_3, b_3$, $\mu_{0k}$, $\sigma_{0k}$ are suitable values, we generated 10 separate Gibbs chains each of which ran for 2000 iterations. For each parameter, we considered the $1010^{th}$, $1020^{th}$, $\cdots$, $2000^{th}$ iteration, which for 10 chains yields a sample of size 1000. The convergence of the Gibbs samplers was monitored using the GELMAN and RUBIN (1992) method that uses the analysis of variance technique to determine if further iterations are needed. In table (2), we have the obtained posterior summaries for the parameters $c_1$, $c_2$, $c_3$, $\beta_1$ and $\beta_2$.

Table 2: Posterior summaries for the regression model with ACBVE distribution

| | mean | median | s.d | 95% credible interval |
|---|---|---|---|---|
| $c_1$ | 0.00203 | 0.00203 | 0.00005 | ( 0.00193; 0.00215) |
| $c_2$ | 0.00602 | 0.00602 | 0.00005 | ( 0.00590; 0.00612) |
| $c_3$ | 0.00904 | 0.00904 | 0.00005 | ( 0.00893; 0.00914) |
| $\beta_1$ | 0.00986 | 0.01161 | 0.02506 | (-0.03636; 0.05754) |
| $\beta_2$ | 0.05115 | 0.04911 | 0.02380 | ( 0.00693; 0.098960) |

In figure 1, we have the approximate marginal posterior densities for $c_1$, $c_2$, $c_3$, $\beta_1$ and $\beta_2$ considering the 1000 Gibbs samples.

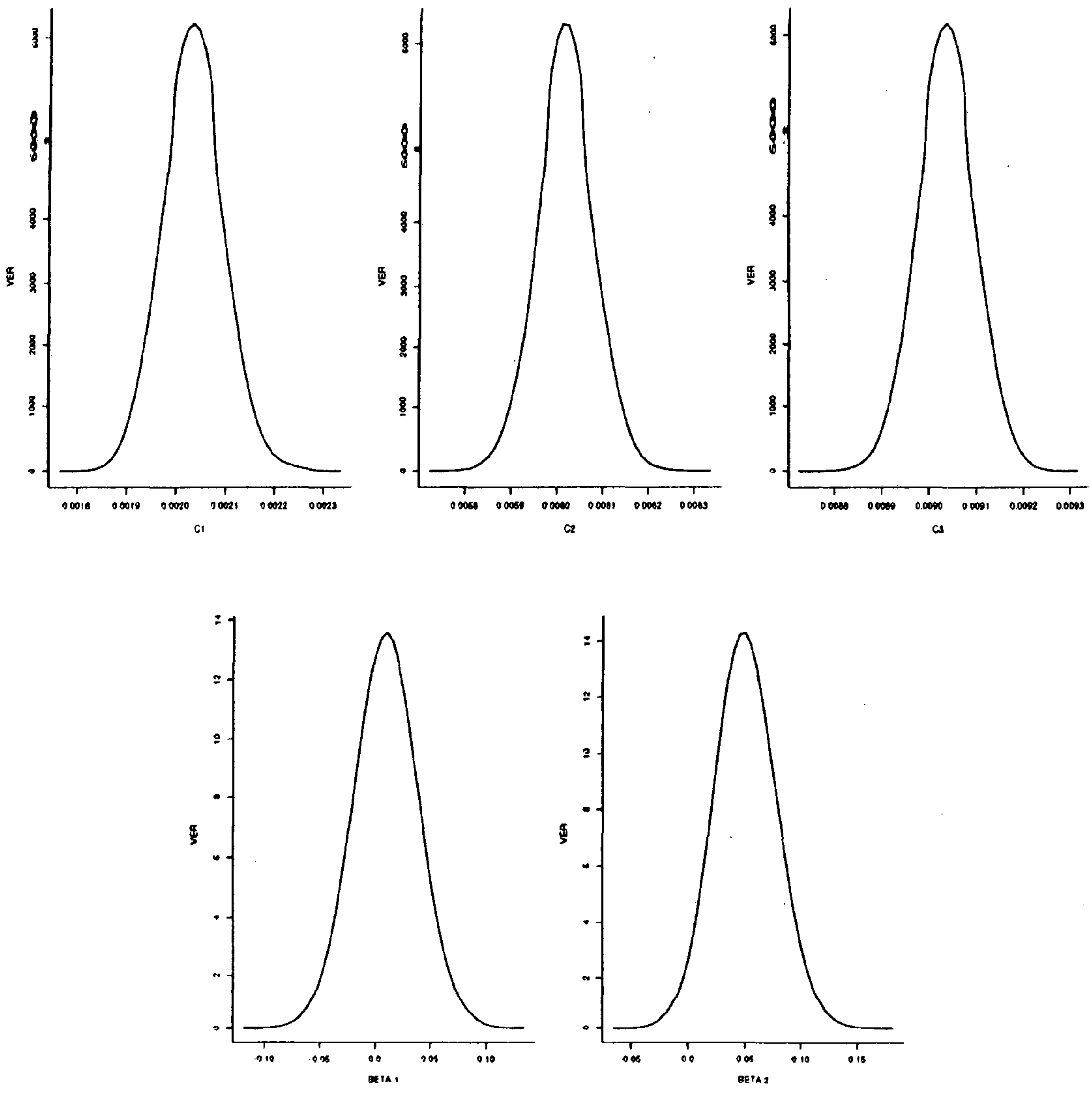

Figure 1: Approximate marginal posterior densities for para $c_1$, $c_2$, $c_3$, $\beta_1$, $\beta_2$

# 4 Concluding remarks

The use of Metropolis-within-Gibbs algorithms is a suitable way to obtain posterior summaries of interest for regression models for bivariate survival data. Other different regression models could be considered as alternative to the regressive model given in (6). We also could consider other prior distribution for the parameters of the model. The use of standard asymptotic results based on the asymptotical normality of the maximum likelihood estimators could be not appropriate for regression models with bivariate lifetime data, especially for small or moderate sample sizes (see for example, LAWLESS, 1982).

# Appendix

## The Gibbs Sampling and Metropolis-Hastings Algorithms

The Gibbs sampler is a Markov chain Monte Carlo (MCMC) technique for generating random variables from a distribution without calculating the density itself (see for example, GELFAND and SMITH, 1990).

Given a collection of $k$ random variables $\boldsymbol{\theta} = (\theta_1, \theta_2, ..., \theta_k)$, we want to generate a random sample from their joint distribution $p(\theta_1, \theta_2, ..., \theta_k)$ or $p(\boldsymbol{\theta})$. By supressing the dependence on the data, we need the complete conditional distributions $p(\theta \mid \boldsymbol{\theta}_{(s)})$, $s = 1, \cdots, k$ where $(\boldsymbol{\theta}_{(s)})$ denotes the random vector $k-1$ random variables with the $s^{th}$ random variable being deleted. Given the initial values of $\theta_1^0$, $\theta_2^0$, $\cdots$, $\theta_k^0$, the Gibbs sampling algorithm proceeds as follows: generate a value $\theta_1^1$ from the conditional density $p(\theta_1 \mid \theta_2^0, \theta_3^0, \cdots, \theta_k^0)$. Similarly, generate a value $\theta_2^1$ from the conditional density $p(\theta_2 \mid \theta_1^1, \theta_3^0, \cdots, \theta_k^0)$ and continue up to the value $\theta_k^1$ from the conditional density $p(\theta_k \mid \theta_1^1, \theta_2^1, \cdots, \theta_{k-1}^1)$.

Then, replace the initial values with the new realization $\boldsymbol{\theta}^1$ of $\boldsymbol{\theta}$, and iterate the above process $t$ times, producing $\boldsymbol{\theta}^{(t)}$. GEMAN and GEMAN, (1984) showed that the $k$-triple produced at the $t^{th}$ iteration of the sampling scheme, $(\theta_1^t, \theta_2^t, \cdots, \theta_k^t)$, converges in distribution to a random variate from $p(\theta_1, \theta_2, \cdots, \theta_k)$ if $t$ is sufficiently large. Further, $\theta_i^k$ can be regarded as a simulated observation from $p(\theta_i)$, the marginal distribution of $\theta_i$.

Replicating the above process $B$ times, we obtain $B$ many $k$-tuples $\{\theta_{1g}^t, \theta_{2g}^t, \cdots, \theta_{kg}^t, g = 1, 2, \cdots, B\}$. Upon convergence of the Gibbs sampler, any characteristic of the marginal density $p(\theta_i)$ can be obtained. In particular, if $p(\theta_s \mid \boldsymbol{\theta}_s)$ is available in closed form, then

$$\hat{\theta}_s = \frac{1}{B}\sum_{g=1}^{B} p(\theta_s \mid \boldsymbol{\theta}_{(s)g})$$

$s = 1, 2, \cdots, k.$

The Gibbs sampler involves drawing random samples from all full conditional densities of $p(\boldsymbol{\theta})$. When the conditional densities are not easily identified, such as in cases without conjugate priors, we can employ the Metropolis-Hastings algorithm.

Suppose we desire to sample a variate from a nonregular density $p(\theta_i \mid \boldsymbol{\theta}_{(i)})$. Observe that $p(\theta_i \mid \boldsymbol{\theta}_{(i)})$ only needs to be known up a constant and we denote $p(\theta)$ as the target density, supressing the subscript and the conditional variables for brevity. Let us define a transition kernel $q(\theta, X)$ which maps $\theta$ to $X$. If $\theta$ is a real variable which ranges in $(-\infty, \infty)$, we can construct $q$ such that $X \longleftarrow \theta + \sigma Z$, with $Z$ being the standard normal random variable and $\sigma^2$ reflecting the conditional variance of $\theta$ in $p(\theta)$. If $\theta$ is bounded with range $(a, b)$, we can use a transformation to map $(a, b)$ to $(-\infty, \infty)$, then use the transition kernel $q$ and apply the Metropolis algorithm to the density of the transformed variable. Then the Metropolis algorithm proceeds as follows:

( i) start with any point $\theta^{(0)}$, and stage indicator $j = 0$;

( ii) generate a point X according to the transition kernel $q(\theta^{(j)}, X)$;

(iii) update $\theta^{(j)}$ to $\theta^{(j+1)} = X$ with probability $p = min\{1, p(X)/p(\theta^{(j)})\}$; stay at $\theta^{(j)}$ with probability $1 - p$;

( iv) repeat (ii) and (iii) by increasing the stage indicator until the process reaches a stationary distribution.

# References

ACHCAR, J. A; LEANDRO, R. A. (1996). Use of Markov chain Monte Carlo methods in a bayesian analysis of the Block and Basu bivariate exponential distribution. **Notas do ICMSC**. Série Estatística, 032/96.

ACHCAR, J. A; SANTANDER, L. A. M.(1993). Use of approximate Bayesian methods for the Block and Basu bivariate exponential distribution, **Journal of the Italian Statistical Society**, 3, 233-250.

BLOCK, H.; BASU, A. P. (1974). A continuous bivariate exponential extension, **Journal of the American Statistical Association**, 69, 1031-1037.

CHIB, S.; GREENBERG, E. (1995). Understanding the Metropolis-Hastings algoritm, **American Statistician**, 49, 4, 327-335.

FREUND, J. E. (1961) . A bivariate extension of the exponential distribution, **Journal of the American statistical Association**, 56, 971-977.

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. E.; RUBIN, D. (1992). Inference from iterative simulation using multiple sequences, **Statistical Science**, 7, 457-472.

GEMAN, S.; GEMAN, D. (1984). Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, 6, 721-741.

LAWLESS, J. F. (1982).Statistical models and methods for lifetime data. New York: John Wiley.

MARSHALL, A. W; OLKIN, I. (1967). A multivariate exponential distribution, **Journal of the American Statistical Association**, 62, 30-44.

SARKAR, S. K. (1987). A continuous bivariate exponential distribution, **Journal of the American Statistical Association**, 82, 667-675.

TIERNEY, L.; KADANE, J. B. (1986). Accurate approximation for posterior moments and marginal, **American Statistical Association**, 81, 82-86.