UNIVERSIDADE DE SÃO PAULO

CONSTANT HAZARD FUNCTION MODELS WITH A CHANGE-POINT: A BAYESIAN ANALYSIS USING MARKOV CHAIN MONTE CARLO METHODS

JORGE ALBERTO ACHCAR
SELENE LOIBEL

Nº 37

NOTAS

Instituto de Ciências Matemáticas de São Carlos

Instituto de Ciências Matemáticas de São Carlos

ISSN - 0103-2577

CONSTANT HAZARD FUNCTION MODELS WITH A CHANGE-POINT: A BAYESIAN ANALYSIS USING MARKOV CHAIN MONTE CARLO METHODS

JORGE ALBERTO ACHCAR
SELENE LOIBEL

Nº 37

NOTAS DO ICMSC
Série Estatística

São Carlos
Fev./1997

RESUMO

Algoritmos de Metropolis com etapas de Gibbs sao propostas para desenvolver uma analise Bayesiana em modelos de funcao de risco constante com ponto de mudanca considerando densidades a priori diferentes para os parametros e dados de sobrevivencia censurados.Tambem apresentamos algumas generalicoes para a comparacao de dois tratamentos.A metodologia e ilustrada com alguns exemplos.

CONSTANT HAZARD FUNCTION MODELS WITH A CHANGE-POINT: A BAYESIAN ANALYSIS USING MARKOV CHAIN MONTE CARLO METHODS

Jorge Alberto Achcar
Selene Loibel
Universidade de São Paulo
ICMSC, C.P. 668, São Carlos, S.P., Brazil

ABSTRACT

Metropolis algorithms along with Gibbs steps are proposed to perform a Bayesian analysis for change-point constant hazard function models considering different prior densities for the parameters and censored survival data. We also present some generalizations for the comparison of two treatments. The methodology is illustrated with some examples.

Key words: Constant hazard, change-point, Gibbs sampling, Metropolis algorithm.

1. Introduction

A frequently recurring question posed by medical researchers concerns a test of a constant failure rate against the alternative for a failure rate involving a single change-point. Let T be the survival time of a patient, the hazard function is given by

$$\lambda_\tau(t) = \begin{cases} \lambda & if\ \ t < \tau \\ \rho\lambda & if\ \ t \geq \tau \end{cases} \tag{1}$$

where λ, ρ, $\tau > 0$

Inferences based on standard asymptotic likelihood results are not possible with this model. Matthews and Farewell (1982) consider the problem of testing the hypothesis ρ=1 based on the likelihood ratio test statistic and use simulations to find the distribution of the statistic of interest. In another paper, Matthews, Farewell and Pyke (1985), present asymptotic score statistics processes to test for constant hazard ρ=1 against a change point alternative $\rho \neq 1$. Considering λ known or λ unknown, they show that the asymptotic significance level for tests based on the maximal score statistics involve the solution to a first passage time problem of an Ornstein-Uhlenbeck process. This solution is based on asymptotic results, assuming τ known. A Bayesian analysis for model (1) is introduced by Achcar and Bolfarine (1989), considering censored data and non-informative prior densities for the parameters of model (1).

In this paper, we present Bayesian inferences for the model (1) using different prior densities and Metropolis-with-Gibbs algorithms (see for example, Gelfand and Smith, 1990 ; Chib and Greenberg, 1995 or Smith and Roberts, 1993).

2. The likelihood function for λ, ρ and τ

Let $T_1^0, T_2^0, \ldots, T_n^0$ be the true survival times of n individuals considered as a random sample of size n and let $C_1, C_2, \ldots., C_n$ be the fixed censoring times associated with each individual (type I censoring). The observed data are given by $T_i = min(T_i^0, C_i)$. We define an indicator variable δ_i such that $\delta_i = 1$ if $T_i = T_i^0$ (failure) and $\delta_i = 0$ if $T_i < T_i^0$(censoring). Associated with the true survival times, we consider the model with hazard function (1). The density and survival function are, respectively

$$f_\tau(t) = \begin{cases} \lambda e^{-\lambda t} & if \;\; t < \tau \\ \rho\lambda e^{-\lambda\tau - \rho\lambda(t-\tau)} & if \;\; t \geq \tau \end{cases} \tag{2}$$

and

$$S_\tau(t) = \begin{cases} e^{-\lambda t} & if \;\; t < \tau \\ e^{-\lambda\tau - \rho\lambda(t-\tau)} & if \;\; t \geq \tau \end{cases} \tag{3}$$

Considering $\epsilon_i = 1$ if $T_i < \tau$ and $\epsilon_i = 0$ if $T_i \geq \tau$, the likelihood function is

$$L(\lambda, \rho, \tau) = \prod_{i=1}^{n} \left\{ \left(\lambda e^{-\lambda t_i}\right)^{\epsilon_i} \left[\rho\lambda exp\{-\lambda\tau - \rho\lambda(t_i - \tau)\}\right]^{1-\epsilon_i} \right\}^{\delta_i} .$$
$$\prod_{i=1}^{n} \left\{ \left(e^{-\lambda t_i}\right)^{\epsilon_i} \left[exp\{-\lambda\tau - \rho\lambda(t_i - \tau)\}\right]^{1-\epsilon_i} \right\}^{1-\delta_i}$$

That is,

$$L(\lambda, \rho, \tau) = \lambda^{d_3} \rho^{d_3 - d_1(\tau)} exp\left\{ -\lambda\left[S_1(\tau) + \rho S_2(\tau)\right]\right\} \tag{4}$$

where,

$$d_1(\tau) = \sum_{i=1}^{n} \delta_i \epsilon_i \quad , \quad d_2(\tau) = \sum_{i=1}^{n} \epsilon_i \;, \quad d_3 = \sum_{i=1}^{n} \delta_i$$

$$w_1(\tau) = \sum_{i=1}^{n} \delta_i \epsilon_i t_i + \sum_{i=1}^{n} (1 - \delta_i)\epsilon_i t_i \quad ,$$
$$w_2(\tau) = \sum_{i=1}^{n} \delta_i (1 - \epsilon_i) t_i + \sum_{i=1}^{n} (1 - \delta_i)(1 - \epsilon_i) t_i \quad ,$$
$$w_3(\tau) = \sum_{i=1}^{n} \delta_i (1 - \epsilon_i) + \sum_{i=1}^{n} (1 - \delta_i)(1 - \epsilon_i) = \sum_{i=1}^{n} (1 - \epsilon_i) = n - d_2(\tau)$$

$$S_1(\tau) = w_1(\tau) + \tau w_3(\tau) \quad ,$$

$S_2(\tau) = w_2(\tau) - \tau w_3(\tau)$.

3. Bayesian inference for model (1)

Let us assume τ taking discrete values $\tau_i = t_i$, with prior probabilities $\pi_0(\tau_i = t_i)$, $i = 1, 2, ..., n$, where n is the sample size. The prior density for λ, ρ and τ_i is,

$$\pi(\lambda, \rho, \tau_i) = \pi(\lambda, \rho \backslash \tau_i = t_i)\pi_0(\tau_i = t_i) \tag{5}$$

Given $\tau_i = t_i$, assuming approximate independence between the parameters λ and ρ, a non-informative prior density for λ and ρ (see for example, Box and Tiao, 1973) is given by

$$\pi(\lambda, \rho \backslash \tau_i = t_i) \propto \frac{1}{\lambda\rho} \tag{6}$$

where λ, $\rho > 0$.

Assuming an uniform prior density $\pi_0(\tau_i = t_i) = \frac{1}{n}$, the joint posterior density for λ, ρ and τ is given by

$$\pi(\lambda, \rho, \tau \backslash \mathcal{D}) \propto \lambda^{d_3 - 1}\rho^{d_3 - d_1(\tau) - 1} exp\Big\{ - \lambda\Big[S_1(\tau) + \rho S_2(\tau)\Big]\Big\} \tag{7}$$

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

The conditional posterior densities for the Gibbs algorithm are given by

$i)$ $\lambda \backslash \rho, \tau, \mathcal{D} \sim \Gamma[d_3, S_1(\tau) + \rho S_2(\tau)]$

$ii)$ $\rho \backslash \lambda, \tau, \mathcal{D} \sim \Gamma[d_3 - d_1(\tau), \lambda S_2(\tau)]$ (8)

$iii)$ $\pi(\tau \backslash \lambda, \rho, \mathcal{D}) \propto \rho^{-d_1(\tau)} exp\{ - \lambda[S_1(\tau) + \rho S_2(\tau)]\}$

where $\Gamma(a, b)$ denotes a Gamma distribution with mean $\frac{a}{b}$ and variance $\frac{a}{b^2}$.

Observe that, we need to use the Metropolis-Hastings algorithm to generate the variable τ. We could monitor the convergence of the Gibbs samples using the Gelman and Rubin (1992) method that uses the analysis of variance technique to determine if further iterations are needed.

We also could consider other prior densities for parameters of model (1). Assuming $\tau = \tau^*$ known, consider the following prior densities for λ and ρ,

$i)$ $\lambda \sim \Gamma(a_1, b_1)$, a_1 and b_1 known,

(9)

$ii)$ $\rho \sim \Gamma(a_2, b_2)$, a_2 and b_2 known.

With this choice of prior density for λ and ρ, and assuming independence between the parameters, the joint posterior density for λ and ρ is given by,

$$\pi(\lambda, \rho \backslash \tau^*, \mathcal{D}) \propto \lambda^{d_3+a_1-1} \rho^{d_3-d_1(\tau^*)+a_2-1} \qquad (10)$$

$$exp\{ -\lambda[b_1 + S_1(\tau^*) + \rho S_2(\tau^*)] - \rho b_2\}$$

The conditional posterior densities for Gibbs algorithm are given by

$$i)\ \lambda \backslash \rho, \tau^*, \mathcal{D} \sim \Gamma[\, d_3 + a_1,\ b_1 + S_1(\tau^*) + \rho S_2(\tau^*)\,]$$

$$ii)\ \rho \backslash \lambda, \tau^*, \mathcal{D} \sim \Gamma[\, d_3 - d_1(\tau^*) + a_2,\ b_2 + \lambda S_2(\tau^*)\,] \qquad (11)$$

4. Comparison for two treatments

In the comparison of two samples, let $T_{11}, T_{12}, \ldots, T_{1n_1}$ be a random sample of size n_1 from a treatment 1 and $T_{21}, T_{22}, \ldots, T_{2n_2}$ be a random sample of size n_2 from a treatment 2.

Assume that the hazard function is given by

$$\lambda_i(t) = \begin{cases} \lambda & if\ \ t_{ij} < \tau_i \\ \rho\lambda e^{\beta x_i} & if\ \ t_{ij} \geq \tau_i \end{cases} \qquad (12)$$

where $i = 1, 2$; $j = 1, 2, \ldots, n_i$; $x_i = 0$ for treatment 1 ($i = 1$) and $x_i = 1$ for treatment 2 ($i = 2$).

As it was considered for model (1), define an indicator variable $\delta_{ij} = 1$ if $T_{ij} = T_{ij}^0$ (failure) and $\delta_{ij} = 0$ if $T_{ij} < T_{ij}^0$ (censoring); $i = 1, 2$; $j = 1, 2, \ldots, n_i$. Assuming $\epsilon_{ij} = 1$ if $T_{ij} < \tau_i$ and $\epsilon_{ij} = 0$ if $T_{ij} \geq \tau_i$, the likelihood function is given by

$$L(\lambda, \rho, \beta, \tau_1, \tau_2) = \prod_{j=1}^{n_1} \left\{ \left(\lambda e^{-\lambda t_{1j}}\right)^{\epsilon_{1j}} \left[\rho\lambda exp\{ -\lambda\tau_1 - \rho\lambda(t_{1j} - \tau_1)\}\right]^{1-\epsilon_{1j}} \right\}^{\delta_{1j}}$$

$$\prod_{j=1}^{n_1} \left\{ \left(e^{-\lambda t_{1j}}\right)^{\epsilon_{1j}} \left[exp\{ -\lambda\tau_1 - \rho\lambda(t_{1j} - \tau_1)\}\right]^{1-\epsilon_{1j}} \right\}^{1-\delta_{1j}} \qquad (13)$$

$$\prod_{j=1}^{n_2} \left\{ \left(\lambda e^{-\lambda t_{2j}}\right)^{\epsilon_{2j}} \left[\rho\lambda e^{\beta} exp\{ -\lambda\tau_2 - \rho\lambda e^{\beta}(t_{2j} - \tau_2)\}\right]^{1-\epsilon_{2j}} \right\}^{\delta_{2j}}$$

$$\prod_{j=1}^{n_2} \left\{ \left(e^{-\lambda t_{2j}}\right)^{\epsilon_{2j}} \left[exp\{ -\lambda\tau_2 - \rho\lambda e^{\beta}(t_{2j} - \tau_2)\}\right]^{1-\epsilon_{2j}} \right\}^{1-\delta_{2j}}$$

That is,

$$L(\lambda,\rho,\beta,\tau_1,\tau_2) = \lambda^{d_3}\rho^{d_3-d_1(\tau_1,\tau_2)}e^{\{\beta(d_3^{(2)}-d_1^{(2)})-\lambda[S_1(\tau_1,\tau_2)+\rho S_2(\tau_1,\tau_2,\beta)]\}} \quad (14)$$

where,

$$d_1^{(i)}(\tau_i) = \sum_{j=1}^{n_i}\delta_{ij}\epsilon_{ij}, \qquad d_1(\tau_1,\tau_2) = d_1^{(1)}(\tau_1) + d_1^{(2)}(\tau_2)$$

$$d_2^{(i)}(\tau_i) = \sum_{j=1}^{n_i}\epsilon_{ij},$$

$$d_3^{(i)} = \sum_{j=1}^{n_i}\delta_{ij}, \qquad d_3 = d_3^{(1)} + d_3^{(2)}$$

$$w_1^{(i)}(\tau_i) = \sum_{j=1}^{n_1}\delta_{ij}\epsilon_{ij}t_{ij} + \sum_{j=1}^{n_1}(1-\delta_{ij})\epsilon_{ij}t_{ij}$$

$$w_2^{(i)}(\tau_i) = \sum_{j=1}^{n_2}\delta_{ij}(1-\epsilon_{ij})t_{ij} + \sum_{j=1}^{n_2}(1-\delta_{ij})(1-\epsilon_{ij})t_{ij}$$

$i = 1, 2$

$$S_1(\tau_1,\tau_2) = w_1^{(1)}(\tau_1) + w_1^{(2)}(\tau_2) + \tau_1(n_1 - d_2^{(1)}) + \tau_2(n_2 - d_2^{(2)})$$

$$S_2(\tau_1,\tau_2,\beta) = w_2^{(1)}(\tau_2) - \tau_1(n_1 - d_2^{(1)}) + e^{\beta}[w_2^{(2)}(\tau_2) + \tau_2(n_2 - d_2^{(2)})]$$

Assuming independent samples with τ_1 and τ_2 known, consider a non-informative prior density given by

$$\pi(\lambda,\rho,\beta) \propto \frac{1}{\lambda\rho} \quad (15)$$

where λ, $\rho > 0$ and $-\infty<\beta<\infty$.

With this choice of prior, the joint posterior density for λ, ρ and β is given by

$$\pi(\lambda,\rho,\beta\backslash\tau_1^*,\tau_2^*,\mathcal{D}) \propto \lambda^{d_3-1}\ \rho^{d_3-d_1(\tau_1^*,\tau_2^*)-1} \quad (16)$$

$$exp\{\beta(d_3^{(2)} - d_1^{(2)}) - \lambda[S_1(\tau_1^*,\tau_2^*) + \rho S_2(\tau_1^*,\tau_2^*,\beta)]\}$$

The conditional posterior densities for the Gibbs algorithm are given by

$i)$ $\lambda\backslash\rho,\beta,\tau_1^*,\tau_2^*,\mathcal{D} \sim \Gamma[d_3,\ S_1(\tau_1^*,\tau_2^*) + \rho S_2(\tau_1^*,\tau_2^*,\beta)]$

$ii)$ $\rho\backslash\lambda,\beta,\tau_1^*,\tau_2^*,\mathcal{D} \sim \Gamma[d_3 - d_1(\tau_1^*,\tau_2^*), \lambda S_2(\tau_1^*,\tau_2^*,\beta)]$ (17)

$iii)\quad \phi\backslash\rho,\lambda,\tau_1^*,\tau_2^*,\mathcal{D} \sim \Gamma\Big[\, d_3^{(2)} - d_1^{(2)}\, ,\ \lambda\rho\,[w_2^{(2)} - \tau_2(n_2 - d_2^{(2)})]\Big]$

where $\phi = e^{\beta}$.

Also assuming τ_1 and τ_2 known, consider now the following prior densities for λ, ρ and β:

$i)\ \lambda \sim \Gamma(a_1, b_1)$, a_1 and b_1 known,

$ii)\ \rho \sim \Gamma(a_2, b_2)$, a_2 and b_2 known, (18)

$iii)\ \beta \sim N[\mu_0, \sigma_0^2]$, μ_0 and σ_0^2 known,

where $N(\mu_0, \sigma_0^2)$ denotes a normal distribution with mean μ_0 and variance σ_0^2.

Assuming independence among the parameters, the joint posterior density for λ,ρ and β is given by

$$\pi(\lambda,\rho,\beta\backslash\tau_1^*,\tau_2^*,\mathcal{D}) \propto \lambda^{d_3+a_1-1}\ \rho^{d_3-d_1(\tau_1^*,\tau_2^*)+a_2-1} \tag{19}$$

$$exp\Big\{-\tfrac{1}{2\sigma_0^2}(\beta-\mu_0)^2\Big\}exp\Big\{\beta(d_3^{(2)} - d_1^{(2)}(\tau_2^*)\Big\}$$

$$exp\Big\{-\lambda[b_1 + S_1(\tau_1^*,\tau_2^*) + \rho S_2(\tau_1^*,\tau_2^*,\beta)]\Big\}\, exp\{-\rho b_2\}$$

The conditional posterior densities required for the Gibbs algorithm are given by

$i)\ \lambda\backslash\rho,\beta,\mathcal{D} \sim \Gamma\,[\,d_3{+}a_1\,,\ b_1 + S_1(\tau_1^*,\tau_2^*) + \rho S_2(\tau_1^*,\tau_2^*,\beta)\,]$

$ii)\ \rho\backslash\lambda,\beta,\mathcal{D} \sim \Gamma\,[\,d_3 - d_1(\tau_1^*,\tau_2^*) + a_2,\ b_2 + S_2(\tau_1^*,\tau_2^*,\beta)\,]$ (20)

$$iii)\ \pi(\beta\backslash\lambda,\rho,\mathcal{D}) \propto\ exp\Big\{-\tfrac{1}{2\sigma_0^2}(\beta-\mu_0)^2\Big\}exp\Big\{\beta(d_3^{(2)} - d_1^{(2)}(\tau_2^*)\Big\}\ exp\Big\{-\lambda\rho\, S_2(\tau_1^*,\tau_2^*,\beta)\Big\}$$

Observe that, we need to use the Metropolis-Hastings algorithm to generate the variable β.

5. Some examples

5.1) An example with one sample.

In table I, we have the remission times of 84 patients with acute nonlymphoblastic leukemia (data in Matthews and Farewell, 1982). We have n=84 and $d_3 = 51$(the number of uncensored observations).

Non-censored observations (51 patients)

24	82	111	152	197	249	270	304	487	534	1160
46	89	117	166	209	254	273	332	510	608	
57	90	128	171	223	258	284	341	516	642	
64	90	143	186	230	264	294	393	518	697	
65	90	148	191	247	269	304	395	518	955	

Censored observations (33 patients)

68	182	182	182	182	182	182	182	182	1310	1908
119	182	182	182	182	182	182	182	182	1538	1966
182	182	182	182	182	182	182	182	583	1634	2057

Table I - Survival times in days.

Considering a non-informative prior density (see (5) and (6)) with an uniform prior density for $\tau_i = t_i$, $i = 1, 2, ..., n$, the conditional posterior densities (8) are given by,

$$i)\ \lambda \backslash \rho, \tau, \mathcal{D} \sim \Gamma[51,\ w_1(\tau) + \rho\, w_2(\tau) + \tau(1-\rho)(84 - d_2(\tau))]$$

$$ii)\ \rho \backslash \lambda, \tau, \mathcal{D} \sim \Gamma[3,\ w_2(\tau) + \lambda\tau(84 - d_2(\tau))] \qquad (21)$$

$$iii)\ \pi(\tau \backslash \lambda, \rho, \mathcal{D}) \propto \rho^{-d_1(\tau)} exp\{ -\lambda[S_1(\tau) + \rho S_2(\tau)]\}$$

We generated 5 separate Gibbs chains each of which ran for 2000 iterations, and we monitored the convergence of the Gibbs samples using the Gelman and Rubin (1992) method that uses the analysis of variance technique to determine if further iterations are needed.

For each parameter, we considered the $515^{th}, 530^{th}, ..., 2000^{th}$ iterations, which for 5 chains yields a sample os size 500.

In table II, we have the obtained posterior summaries for the parameter λ, ρ and τ, and in figure 1 , we have the approximate marginal posterior densities considering the 500 Gibbs samples. It is interesting to observe that the maximum likelihood estimators for λ, ρ and τ are given by $\hat{\lambda}$ = 0.00204, $\hat{\rho}$ =0.21046 and $\hat{\tau}$ = 697.

	Mean	Mode	S.D.	95% Credible Interval
τ	752.0304	738.3420	57.5157	(628.1459 , 855.0770)
λ	0.0020	0.0020	0.0003	(0.0016 , 0.0026)
ρ	0.1602	0.0961	0.1077	(0.0149 , 0.4175)

Table II - Posterior summaries (non-informative prior)

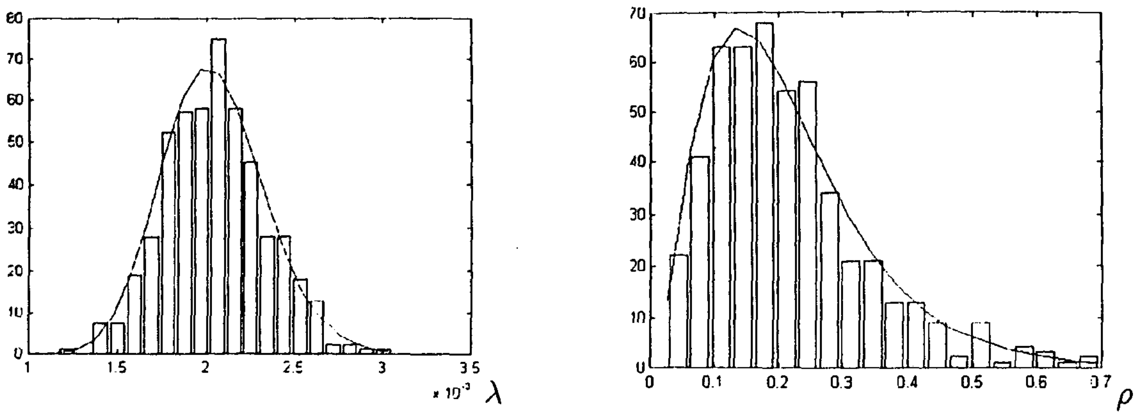

Figure 1 - Marginal posterior densities for λ and ρ (non-informative prior)

Assuming $\tau^* = 697$ known, consider now the prior density (9) with $a_1 = 4$, $b_1 = 2048$, $a_2 = 18$ and $b_2 = 86$. The conditional posterior densities (11) for the Gibbs algorithm are given by,

$$i)\, \lambda \backslash \rho, \tau^*, \mathcal{D} \sim \Gamma[\,55, 25559 + 6982\rho\,]$$

$$ii)\, \rho \backslash \lambda, \tau^*, \mathcal{D} \sim \Gamma[\,21, 86 + 6982\,\lambda] \qquad (22)$$

In table III, we have the obtained posterior summaries for the parameters λ and ρ, and in figure 2, we have the approximate marginal posterior densities considering 500 Gibbs samples.

	Mean	Mode	S.D.	95% Credible Interval
λ	0.0020	0.0018	8.4×10^{-5}	(0.001579 , 0.0026318)
ρ	0.2099	0.2404	1.99×10^{-3}	(0.12215 , 0.29894)

Table III - Posterior summaries (prior density (9))

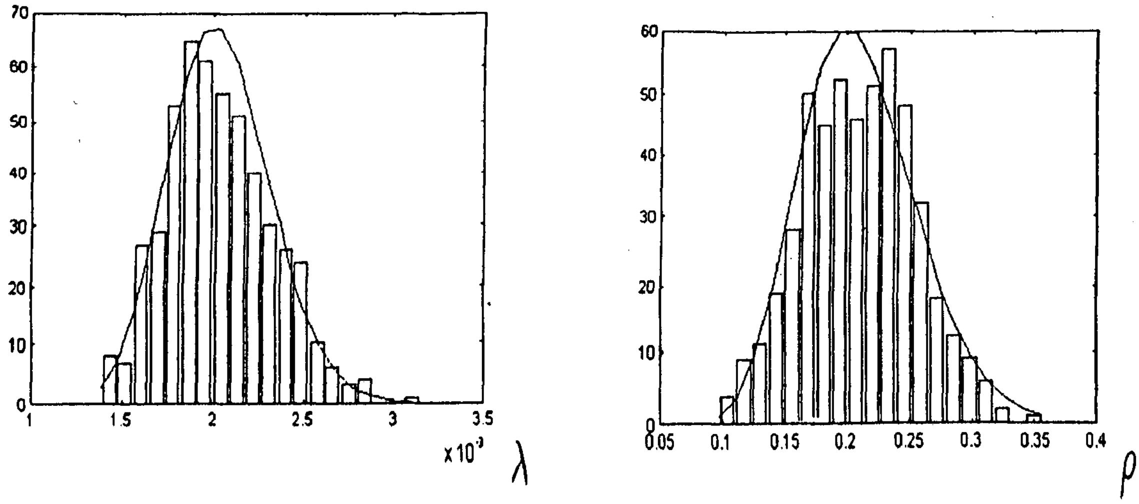

Figure 2 - Marginal posterior densities for λ and ρ (prior density (9))

5.2) An example with two samples

In table IV, we have the survival times of patients submitted to two different treatments. For treatment 1, we have $n_1 = 40$ patients and for treatment 2, we have $n_2 = 50$ patients.

Treatment 1: ($n_1 = 40$)

0.2	1.4	2.3	5.0	7.8	11.2	15.0	48.9	121.1	230.6
0.5	1.5	3.2	5.4	8.8	11.5	39.5	63.0	121.2	253.5
0.7	1.6	4.0	7.0	9.4	13.8	44.4	71.5	198.5	303.0
0.8	2.1	4.6	7.3	10.0	14.0	46.8	115.4	218.6	365.9

Treatment 2: ($n_2 = 50$)

0.15	1.26	1.81	6.00	7.20	10.5	287	1288	1909	3551
0.16	1.48	2.03	6.20	7.70	14.6	450	1361	1951	3946
0.82	1.55	2.90	6.40	8.50	16.9	860	1433	2531	5721
0.92	1.71	4.10	6.80	9.10	21.2	1041	1560	3004	6050
1.24	1.72	5.80	7.00	9.80	25.0	1246	1670	3045	6083

Table IV - Survival times in days.

In figure 3, we have the plot of $log[-\widehat{H}(t)]$ versus $log(t)$, where $\widehat{H}(t) = -log[\widehat{S}(t)]$, $\widehat{S}(t)$ is the Kaplan and Meier (1958) estimator of survival function. We observe that the model (12) is appropriated for the data. From figure 3, we also obtain approximate estimates for τ_1(treatment 1) and τ_2(treatment 2) given by $\widehat{\tau}_1 \approx 15$ and $\widehat{\tau}_2 \approx 25$.

Treatment 1 | Treatment 2

$log[-\widehat{H}(t)]$ | $log[-\widehat{H}(t)]$

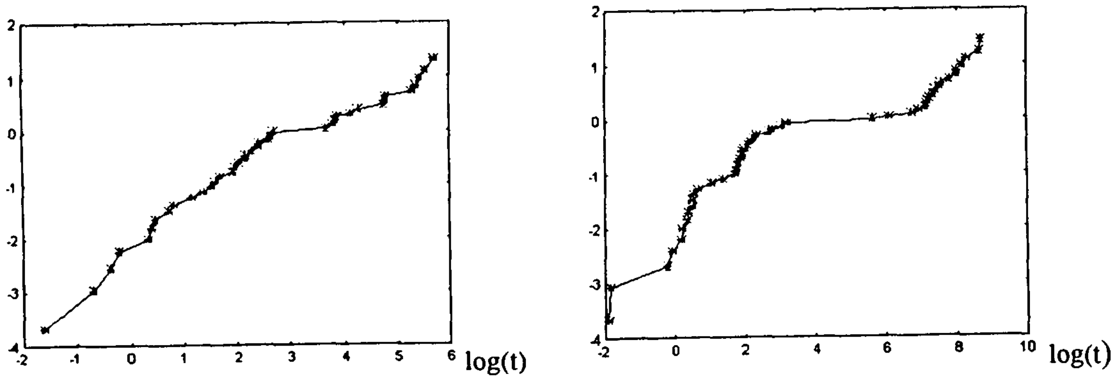

Figure 3 - $\log[-\hat{H}(t)]$ versus $\log(t)$.

Assuming $\tau_1=15$ and $\tau_2=25$ known, and the non-informative prior density (15) the conditional posterior densities (17) for the Gibbs algorithm are given by

$$
\begin{aligned}
&i)\ \lambda \backslash \rho, \phi, \mathcal{D} \sim \Gamma[90,\ 1064.65 + (2016.9 + 48487\phi)\rho] \\
&ii)\ \rho \backslash \lambda, \phi, \mathcal{D} \sim \Gamma[37,\ (2016.9 + 48487\phi)\lambda] \qquad (23) \\
&iii)\ \phi \backslash \lambda, \rho, \mathcal{D} \sim \Gamma[\ 27,\ 48487\lambda\rho\ ]
\end{aligned}
$$

where $\phi = e^{\beta}$.

We generated 5 separate Gibbs chains each of which ran for 2000 iterations. For each parameter, we considered the $515^{th}, 530^{th}, ..., 2000^{th}$ iterations, which for 5 chains yields a sample of size 500.

In table V, we have the obtained posterior summaries for each parameter λ,ρ and β and in figure 4, we have the approximate marginal posterior densities considering the 500 Gibbs samples. The maximum likelihood estimators for λ, ρ and β are given by $\hat{\lambda}$=0.04978, $\hat{\rho}$ = 0.1594 and $\hat{\beta}$ = -2.9079.

	Mean	Mode	S.D.	95% Credible Interval
λ	0.04962	0.04722	0.00664	(0.03616 , 0.06291)
ρ	0.1639	0.1352	0.0484	(0.04729 , 0.18227)
β	-2.9149	-3.0349	0.3470	(-2.9108 , -1.4325)

Table V - Posterior summaries (non-informative prior)

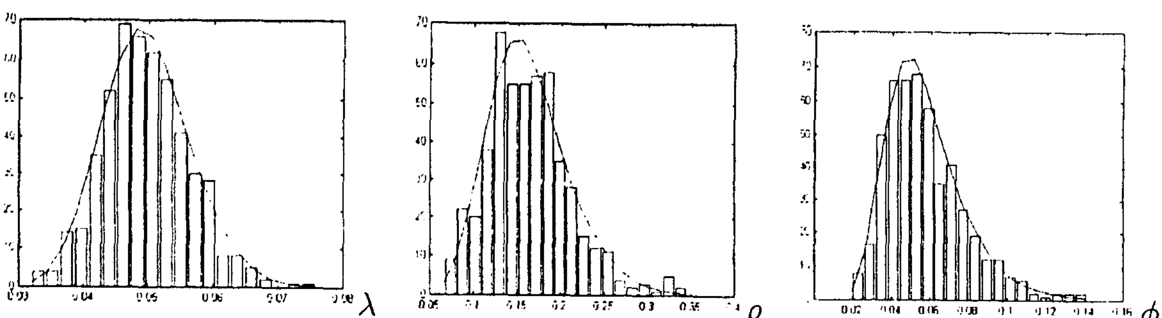

Figure 4 - Marginal posterior densities for λ, ρ and $\phi = e^{\beta}$ (prior density (15))

Also assuming $\tau_1 = 15$ and $\tau_2 = 25$ known and the prior densities (see (18)),

$$\begin{aligned} &i)\, \lambda \sim \Gamma[\, 53,\ 1065.9] \\ &ii)\, \rho \sim \Gamma[\, 12,\ 77\,] \\ &iii)\, \beta \sim N[\, -2.9079,\ 0.11\,], \end{aligned} \tag{24}$$

The conditional posterior densities (20) required for the Gibbs algorithm are given by

$$\begin{aligned} &i)\, \lambda \backslash \rho, \beta, \mathcal{D} \sim \Gamma[\, 143,\ 2130.65 + (2016.9 + 48487e^{\beta})\rho\,] \\ &ii)\, \rho \backslash \lambda, \beta, \mathcal{D} \sim \Gamma\,[49,\ 77 + (2016.9 + 48487e^{\beta})\lambda\,] \\ &iii)\, \beta \backslash \lambda, \rho, \mathcal{D} \sim N[\, -2.9079,\ 0.11]\, \psi(\lambda, \rho, \beta) \end{aligned} \tag{25}$$

here $\psi(\lambda, \rho, \beta) = exp\{21\beta - (2016.9 + 48487e^{\beta})\lambda\rho\}$

In table VI, we have the obtained posterior summaries for each parameter λ,ρ and β, considering 500 Gibbs samples generated from (25).

	Mean	Mode	S.D.	95% Credible Interval
λ	0.0501	0.0482	0.0045	(0.0422 , 0.0592)
ρ	0.1579	0.1630	0.0258	(0.1107 , 0.2116)
β	-2.8974	-2.9293	0.1035	(-3.1048 , -2.6923)

Table VI - Posterior summaries (prior density (24))

In figure 5, we have the approximate marginal posterior densities considering 500 Gibbs samples.

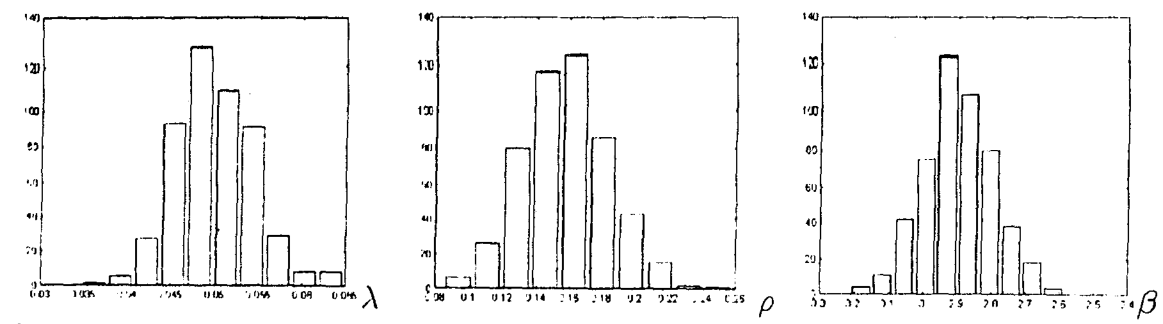

Figure 5 - Histograms for λ, ρ and β (prior density (24)).

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 $\underset{\sim}{\theta} = (\theta_1, \theta_2, ..., \theta_k)$, we want to generate a random sample from their joint distribution $p(\theta_1, \theta_2, ..., \theta_k)$ or $p(\underset{\sim}{\theta})$. By suppressing the dependence on the data, we need the complete conditional distributions $p(\theta_s, \underset{\sim}{\theta}_{(s)})$, $s = 1, 2, ..., k$ where $\underset{\sim}{\theta}_{(s)}$ denotes the random vector of $k-1$ random variables with s^{th} random variable being deleted. Given the initial values of $\theta_1^{(0)}, \theta_2^{(0)}, ..., \theta_k^{(0)}$, the Gibbs sampling algorithm proceeds as follows: generate a value $\theta_1^{(1)}$ from the conditional density $p(\theta_1 \backslash \theta_2^{(0)}, \theta_3^{(0)}, ..., \theta_k^{(0)})$. Similarly, generate a value $\theta_2^{(1)}$ from the conditional density $p(\theta_2 \backslash \theta_1^{(1)}, \theta_2^{(0)}, ..., \theta_k^{(0)})$ and continue up to the have $\theta_k^{(1)}$ from the conditional density $p(\theta_k \backslash \theta_1^{(1)}, \theta_2^{(1)}, ..., \theta_{k-1}^{(1)})$.

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

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

$$\widehat{p}(\theta_s) = \sum_{g=1}^{B} p(\theta_s, \underset{\sim}{\theta}_{(s)g}), \quad s = 1, 2, ..., k.$$

The Gibbs sampler involves drawing random samples form all full conditional densities of $p(\underset{\sim}{\theta})$. 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 $p(\theta_i, \underset{\sim}{\theta}_{(i)})$. Observe that $p(\theta_i, \underset{\sim}{\theta}_{(i)})$ only needs to be known up to a constant and we denote $p(\theta)$ as the target density, suppressing the subscript, and the conditional variables for brevity. Let us define a transition Kernel $q(\theta, X)$ which maps θ to X. If θ is a real variable with ranges in $(-\infty,\infty)$, we can construct q such that $X \leftarrow \theta + \sigma Z$, with Z being the standard normal random variable and σ^2 reflecting the conditional variance of θ in $p(\theta)$. If θ 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 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 he process reaches a stationary distribution.

REFERENCES

Achcar, J.A. ; Bolfarine, H. (1989), Constant hazard against a change-point alternative: a Bayesian approach with censored data. *Communications in Statistics - Theory and methods*, 18(10), 3801-3819.

Box, G.E. ; Tiao, G.C. (1973). Bayesian inference in Statistical Analysis, New York: Addison-Wesley.

Chib,S. ; Greenberg, E. (1995). Understanding the Metropolis-Hastings algorithm. *American Statistician*, 49,4,327-335.

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.

Kaplan, E. L. ; Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, 53, 457-481.

Mattews, D. E. ; Farewell, V. T. (1982). On testing for a constant hazard against a change-point alternative. *Biometrics*, 38, 463-468.

Mattews, D. E. ; Farewell, V. T. ; Pyke, R. (1985). Asymptotic score statistic processes and tests for constant hazard against a change-point alternative. *Annals of Statistics*, 13, 2, 583-591.

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

NOTAS DO ICMSC

SÉRIE ESTATÍSTICA

036 /97 MOALA, F..A.; RODRIGUES, J. - Bayesian inference of the Weibull reliability function via Laplace approximation

.035/96 ACHCAR, J.A.; STORANI, K. - Nonhomogeneous poisson processes assuming a inverse Gaussian order statistics model for software reliability data: a bayesian approach.

034/96 ACHCAR, J.A.; LEANDRO, R.A. - Regression models for bivariate survival data: a bayesian approach.

033/96 ACHCAR, J.A. - Use of gibbs-with-metropolis-hastings algorithms for a bayesian analysis of complex network reliability systems.

032/96 ACHCAR, J.A.; LEANDRO, R.A. - Use of markov chain Monte Carlo methods in a bayesian analysis of the Block and Basu bivariate exponential distribution.

031/96 CEREGATO, S.A.; RODRIGUES, J. - Utilização da inferência bayesiana em experimentos de captura-recaptura.

030/96 ACHCAR, J.A. - Bayesian inference for software reliability models considering interfailure time data.

029/96 ACHCAR, J.A. - Bayesian inference for software reliability models using homogeneous poisson process.

028/96 RODRIGUES, J.; LEITE, J.G. - Inference for the software reliability using imperfect recapture debbugging model.

027/96 ACHCAR, J.A.; DEY, D.K.; NIVERTHI, M. - A bayesian approach using nonhomogeneous poisson process for software reliability models.