UNIVERSIDADE DE SÃO PAULO

Bayesian Inference for Software Reliability Models Considering Interfailure Time Data

JORGE ALBERTO ACHCAR

Nº 30

NOTAS

Instituto de Ciências Matemáticas de São Carlos

Instituto de Ciências Matemáticas de São Carlos

ISSN - 0103-2577

Bayesian Inference for Software Reliability Models Considering Interfailure Time Data

JORGE ALBERTO ACHCAR

Nº 30

NOTAS DO ICMSC
Série Estatística

São Carlos
Mai./96

RESUMO

Neste artigo,consideramos inferencia Bayesiana para alguns modelos de confiabilidade de software com dados de tempos entre falhas.Algoritmos Metropolis-Hastings com etapas de Gibbs sao propostos para desenvolver a inferencia Bayesiana para os modelos propostos.Tambem,exploramos alguns criterios Bayesianos de selecao de modelos considerando um conjunto de dados de confiabilidade de software introduzido por Jelinski e Moranda (1972).

Palavras-chaves: modelos de confiabilidade de software,dados de tempos entre falhas,Gibbs sampling,algoritmo de Metropolis.

Bayesian Inference for Software Reliability Models Considering Interfailure Time Data

Jorge Alberto Achcar
Universidade de São Paulo
ICMSC-USP, C. Postal 668
13560-970, São Carlos, S.P., Brazil

Abstract

In this paper, we consider Bayesian inference for some software reliability models considering interfailure time data. Metropolis - Hastings algorithms along with Gibbs steps are proposed to perform the Bayesian inference for the proposed models. We also explore some Bayesian model selection criteria considering a software reliability data set introduced by Jelinski and Moranda (1972).

Keywords: software reliability models, interfailure time data, Gibbs sampling, Metropolis algorithm.

1 Introduction

Software reliability is the probability of a computer program to be free of error in operation during a specified period of time. The software failures are related to errors in syntax or logic (see for example, Singpurwalla and Wilson, 1994). The literature in statistics and especially in software engineering presents many different models for software reliability. Among these stochastic models we have two strategies: (i) modelling times between successive failures or, (ii) modelling the number of failures of the software up to a given time.

For modelling interfailure time data, Jelinski and Moranda (1972) suppose that the total number of bugs in the program is N, and they assume that each time the software fails, one bug is corrected. The failure rate of the ith time between failures T_i is assumed to be a constant proportional to $N - i + 1$, which is the number of bugs remaining in the program.

Thus, the failure rate for T_i is given by

$$\lambda_i = \lambda_{JM}(N - i + 1) \tag{1}$$

where $i = 1, 2, 3, \ldots$.

That is, T_i has an exponential density,

$$f(t_i|\lambda_i) \;=\; \lambda_i \, e^{-\lambda_i t_i} \tag{2}$$

where $T_i \geq 0$ and λ_i is given by (1).

Some modifications of the JM model (1) are introduced in the literature. Moranda (1975), assume that the fixing of bugs that cause early failures in the system reduces the failure rate more than the fixing of bugs that occur later. Therefore, Moranda (1975) assume that the failure rate should remain constant for each T_i, but it decreases geometricaly in i after each failure, that is, for constants θ and k,

$$\lambda_i = \theta k^{i-1} \tag{3}$$

where $\theta > 0$ and $0 < k < 1$.

Goel and Okumoto (1978) propose a model similar to JM model (1), but assuming that there is a probability $p, 0 \leq p \leq 1$, of fixing a bug when it is encountered. Thus, the failure rate of T_i is given by

$$\lambda_i = \lambda_{GO}[N - p(i-1)] \tag{4}$$

When $p = 1$, we have the JM model (1).

Schick and Wolverton (1978) assume that the failure rate is proportional to the number of bugs remaining in the system and the time elapsed since the last failure. Thus,

$$\lambda_i = \lambda_{SW}(N - i + 1)t \tag{5}$$

Other model for software interfailure time data is introduced by Littlewood and Verrall (1973) to relax the assumption of perfect repair in the JM model (1). They assume that the ith time between failures has an exponential density (2) with failure rate λ_i, and that instead of λ_i be considered decreasing with certainty, as it is assumed in the JM model (1), they assume the sequence of $\lambda's$ to be stochastically decreasing, that is,

$P(\lambda_{i+1} < \lambda) \geq P(\lambda_i < \lambda)$, for $i = 1, 2, \ldots$ and $\lambda \geq 0$.

In this way, they consider a gamma distribution for λ_i with shape parameter α and scale $\psi(i)$, where $\psi(i)$ is a monotonically increasing function of i,

$$\pi(\lambda_i|\alpha, \psi(i)) = \frac{[\psi(i)]^\alpha}{\Gamma(\alpha)} \lambda_i^{\alpha-1} e^{-\psi(i)\lambda_i} \tag{6}$$

The function $\psi(i)$ is supposed to describe the quality of the programmer and the programming task. Mazzuchi and Soyer (1988) assume $\psi(i) = \beta_0 + \beta_1 i$.

For modelling the number of failures of the software up to a given time, the literature presents many different models assuming homogeneous or non-homogeneous Poisson process (see for example, Goel and Okumoto, 1979; Goel, 1983; Ohba and Yamada, 1982; Musa and Okumoto, 1984).

The use of Gibbs sampling with Metropolis-Hastings algorithms (see for example, Gelfand and Smith, 1990) has been used by many authors to get Bayesian inference for software reliability models (see for example, Kuo and Yang, 1995; Yang, 1994; Kuo, Lee, Choi and Yang, 1996; Achcar, Dey and Niverthi, 1996).

Yang (1994) considers the use of Metropolis-within - Gibbs algorithms to obtain Bayesian inferences on the Jelinski and Moranda (1972) model and on the Littlewood and Verral (1973) model, considering the introduction of latent variable (data augmentation technique) to simplify the conditional posterior densities required in the Gibbs algorithm.

In this paper, we present Bayesian inference for some special software reliability models using Metropolis-within- Gibbs algorithms (without the inclusion of latent variables): the Jelinski and Moranda (1972) model, the Moranda (1975) model, the Goel and Okumoto (1978) model, the Schick and Wolverton (1978) model and the Littlewood and Verral (1973) model. We also explore some Bayesian model selection criteria considering a software reliability data set introduced by Jelinski and Moranda (1972).

2 Bayesian Inference for the Jelinski-Moranda Model

Assuming a failure truncated model, that is, the test terminates when we observe the n^{th} failure, let $T_1, T_2, \ldots, T_n$ be a random sample of interfailure times of a software under the JM model (1). The likelihood function for λ_{JM} and N is given by

$$L(\lambda_{JM}, N) = \lambda_{JM}^n A(N) exp\{-\lambda_{JM} B(N)\} \tag{7}$$

where $A(N) = \Pi_{i=1}^n (N - i + 1)$ and $B(N) = \sum_{i=1}^n (N - i + 1) t_i$.

Observe that $t_i = x_i - x_{i-1}$, where x_i denotes the ordered epochs of failure time, $i = 1, 2, \ldots, n$. Therefore,

$$B(N) = \sum_{i=1}^{n} (N - i + 1) t_i = \sum_{i=1}^{n} x_i + (N - n) x_n.$$

Assuming prior independence, consider the following prior densities for λ_{JM} and N:

$$\begin{aligned} &\text{(i)} \qquad \lambda_{JM} \sim \Gamma[a_1, b_1] \\ &\text{(ii)} \qquad N \sim \mathcal{P}(\theta_1) \end{aligned} \tag{8}$$

where $\Gamma[a_1, b_1]$ denotes a gamma distribution with mean a_1/b_1, $\mathcal{P}(\theta_1)$ denotes a Poisson distribution with mean θ_1 and a_1, b_1 and θ_1 are known constants.

The joint posterior density for λ_{JM} and N is given by

$$\begin{aligned} \pi(\lambda_{JM}, N | \mathcal{D}_n) \propto\ & \frac{\lambda_{JM}^{n+a_1-1} A(N) \theta_1^N}{N!} \times \\ & \times\ exp\{-[b_1 + (N - n) x_n + \textstyle\sum_{i=1}^n x_i] \lambda_{JM}\} \end{aligned} \tag{9}$$

where $\lambda_{JM} > 0$; $N = n, n+1, \ldots$; and $\mathcal{D}_n$ denote the data set $\{x_1, x_2, \ldots, x_n\}$.

The conditional distributions for λ_{JM} and N required for the Gibbs algorithm are given by

$$\begin{aligned} &\text{(i)} \qquad \lambda_{JM} | N, \mathcal{D}_n \sim \Gamma\left[n + a_1, b_1 + (N-n)x_n + \textstyle\sum_{i=1}^{n} x_i\right] \\ &\text{(ii)} \qquad \pi\left(N | \lambda_{JM}, \mathcal{D}_n\right) \propto \frac{A(N)}{N!}\, \theta_1^N e^{-(N-n)x_n\lambda_{JM}} \end{aligned} \tag{10}$$

Observe that $A(N)/N! = 1/(N-n)!$. Therefore, considering $N' = N - n$, we get a Poisson distribution with mean $\theta_1 e^{-x_n\lambda_{JM}}$ for N' given λ_{JM} and $\mathcal{D}_n$.

That is,

$$\begin{aligned} &\text{(i)} \qquad \lambda_{JM} | N', \mathcal{D}_n \sim \Gamma\left[n + a_1, b_1 + N'x_n + \textstyle\sum_{i=1}^{n} x_i\right] \\ &\text{(ii)} \qquad N' | \lambda_{JM}, \mathcal{D}_n \sim \mathcal{P}\left(\theta_1 e^{-x_n\lambda_{JM}}\right) \end{aligned} \tag{11}$$

A sample of draws from the joint posterior density for λ_{JM} and N' can now be obtained by successively sampling λ_{JM} from $\pi\left(\lambda_{JM} | N', \mathcal{D}_n\right)$, and given this value of λ_{JM}, simulating N' from $\pi\left(N' | \lambda_{JM}, \mathcal{D}_n\right)$.

3 Bayesian Inference for the Moranda Model

Also assuming a failure truncated model and the MO model (3), the likelihood function for θ and k is given by

$$L(\theta, k) = \theta^n k^{\frac{n(n-1)}{2}} exp\left\{-\theta \sum_{i=1}^{n} k^{i-1} t_i\right\} \tag{12}$$

Assuming prior independence, consider the following prior densities for θ and k:

$$\begin{aligned} &\text{(i)} \qquad \theta \sim \Gamma\left[a_2, b_2\right] \\ &\text{(ii)} \qquad k \sim B\left[a_3, b_3\right] \end{aligned} \tag{13}$$

where $B[a_3, b_3]$ denotes a Beta distribution

$\pi(k) \propto k^{a_3-1}(1-k)^{b_3-1}, 0 \leq k \leq 1$, and a_2, b_2, a_3 and b_3 are known constants.

The joint posterior density for θ and k is given by

$$\begin{aligned} \pi(\theta, k|\mathcal{D}_n) \propto\ & \theta^{n+a_2-1} k^{a_3+\frac{n(n-1)}{2}-1} \times \\ & \times (1-k)^{b_3-1} exp\left\{-\left(b_2 + \sum_{i=1}^{n} k^{i-1}t_i\right)\theta\right\} \end{aligned} \quad (14)$$

The conditional distributions for θ and k required for the Gibbs algorithm are given by

$$\begin{aligned} &\text{(i)} \quad \theta|k, \mathcal{D}_n \sim \Gamma\left[n + a_2, b_2 + \sum_{i=1}^{n} k^{i-1}t_i\right] \\ &\text{(ii)} \quad \pi(k|\theta, \mathcal{D}_n) \propto k^{a_3+\frac{n(n-1)}{2}-1}(1-k)^{b_3-1} \times \\ &\qquad \times\ exp\left\{-\theta \sum_{i=1}^{n} k^{i-1}t_i\right\} \end{aligned} \quad (15)$$

Observe that the variable k should be generated using the Metropolis-Hastings algorithm (see for example, Chib and Greenberg, 1994).

In this way, observe that the conditional density for k given θ and $\mathcal{D}_n$ could be written by,

$$\pi(k|\theta, \mathcal{D}_n) \propto k^{a_3-1}(1-k)^{b_3-1}\psi_1(\theta, k) \quad (16)$$

where $\psi_1(\theta, k) = exp\left\{\frac{n(n-1)}{2} lnk - \theta \sum_{i=1}^{n} k^{i-1}t_i\right\}$

Thus, the value of k is simulated as: at the s$^{\text{th}}$ iteration (given the current value $\theta^{(s)}$, draw a candidate $k^{(s)}$ from a Beta distribution $B(a_3, b_3)$; if it satisfies stationarity, move to this point with probability

$$min\left\{\frac{\psi_1\left(\theta^{(s)}, k^{(s)}\right)}{\psi_1\left(\theta^{(s)}, k^{(s-1)}\right)}, 1\right\}, \quad (17)$$

and otherwise set $k^{(s)} = k^{(s-1)}$, where $\psi_1(\theta, k)$ is defined in (16).

4 Bayesian Inference for the Goel and Okumoto Model

Assuming the GO model (4), the likelihood function for λ_{GO}, p and N (also with a failure truncated model) is given by

$$L(\lambda_{GO}, N, p) = \lambda_{GO}^{n} A(N,p) exp\{-\lambda_{GO} B(N,p)\} \tag{18}$$

where $A(N,p) = \prod_{i=1}^{n}[N - p(i-1)]$ and $B(N,p) = \sum_{i=1}^{n}[N - p(i-1)]t_i$.

Since $t_i = x_i - x_{i-1}$, where x_i denotes the ordered epochs of failure time, we have $B(N,p) = p\sum_{i=1}^{n} x_i + (N - np)x_n$.

Assuming prior independence, consider the following prior densities for λ_{GO}, N and p:

$$\begin{array}{llll} \text{(i)} & \lambda_{GO} & \sim & \Gamma[a_4, b_4] \\ \text{(ii)} & N & \sim & \mathcal{P}(\theta_2), \\ \text{(iii)} & p & \sim & B[a_5, b_5], \end{array} \tag{19}$$

where a_4, b_4, a_5, b_5 and θ_2 are known constants.

The joint posterior density for λ_{GO}, N and p is given by

$$\begin{aligned} \pi(\lambda_{GO}, N, p|\mathcal{D}_n) \propto & \frac{\lambda_{GO}^{n+a_4-1} A(N,p)\theta_2^N}{N!} \times \\ & \times\ p^{a_5-1}(1-p)^{b_5-1} exp\{-[b_4 + p\sum_{i=1}^{n} x_i + (N-np)x_n]\lambda_{GO}\}\ , \end{aligned} \tag{20}$$

where $\lambda_{GO} > 0$; $N = n, n+1, \ldots$; and $0 \le p \le 1$.

In this case, the marginal posterior densities for the Gibbs algorithm are given by

$$\begin{array}{llll} \text{(i)} & \lambda_{GO}|N, p, \mathcal{D}_n & \sim & \Gamma[n + a_4, b_4 + p\sum_{i=1}^{n} x_i + (N - np)x_n] \\ \text{(ii)} & \pi(N|\lambda_{GO}, p, \mathcal{D}_n) & \propto & \frac{e^{-\theta_2}\theta_2^N}{N!}\psi_2(N, p, \lambda_{GO})\ , \end{array} \tag{21}$$

where

$$\psi_2(N,p,\lambda_{GO}) = exp\{lnA(N,p) - (N-np)x_n\lambda_{GO}\}.$$

and

(iii) $\pi(p|N,\lambda_{GO},\mathcal{D}_n) \propto p^{a_5-1}(1-p)^{b_5-1}\psi_3(N,p,\lambda_{GO})$,

where

$$\psi_3(N,p,\lambda_{GO}) = exp\{lnA(N,p) - p\lambda_{GO}\textstyle\sum_{i=1}^{n} x_i - (N-np)x_n\lambda_{GO}\}.$$

Observe that, the variables N and p should be generated using the Metropolis-Hastings algorithm.

5 Bayesian Inference for the Schick and Wolverton Model

Assuming the SW model (5), observe that the density function for the interfailure time T_i is given by,

$$f_i(t) = \lambda_{SW}t(N-i+1)exp\left\{-\lambda_{SW}(N-i+1)t^2/2\right\} \quad (22)$$

Also assuming a failure truncated model, the likelihood function for λ_{SW} and N is given by

$$L(\lambda_{SW},N) = \lambda_{SW}^n A_1(N)exp\left\{-\frac{\lambda_{SW}}{2}B_1(N)\right\} \quad (23)$$

where $A_1(N) = \prod_{i=1}^{n} t_i(N-i+1)$ and $B_1(N) = \sum_{i=1}^{n}(N-i+1)t_i^2$.

Assuming prior independence, consider the following prior densities for λ_{SW} and N:

$$\begin{aligned} &\text{(i)} \quad \lambda_{SW} \sim \Gamma\left[a_6, b_6\right], \\ &\text{(ii)} \quad N \sim \mathcal{P}(\theta_3) \end{aligned} \tag{24}$$

where a_6, b_6 and θ_3 are known constants.

The joint posterior density for λ_{SW} and N is given by,

$$\begin{aligned} \pi(\lambda_{SW}, N|\mathcal{D}_n) \propto \frac{\lambda_{SW}^{n+a_6-1} A_1(N)\theta_3^N}{N!} \times \\ \times\, exp\left\{-\lambda_{SW}\left[b_6 + \tfrac{N}{2}\sum_{i=1}^n t_i^2 - \tfrac{1}{2}\sum_{i=1}^n (i-1)t_i^2\right]\right\} \end{aligned} \tag{25}$$

where $\lambda_{SW} > 0$ and $N = n, n+1, \ldots$

The conditional posterior densities for the Gibbs algorithm are given by

$$\text{(i)} \quad \lambda_{SW}|N, \mathcal{D}_n \sim \Gamma\left[n + a_6, b_6 + \tfrac{N}{2}\sum_{i=1}^n t_i^2 - \tfrac{1}{2}\sum_{i=1}^n (i-1)t_i^2\right], \tag{26}$$

and

$$\text{(ii)} \quad \pi(N|\lambda_{SW}, \mathcal{D}_n) \propto \frac{e^{-\theta_3}\theta_3^N}{N!} A_1(N) exp\left\{-\frac{\lambda_{SW}}{2} N \sum_{i=1}^n t_i^2\right\}$$

Considering the transformation $N' = N - n$, we observe that $A_1(N)/N! = \prod_{i=1}^n t_i / N'!$. Therefore, we get,

$$\begin{aligned} &\text{(i)} \quad \lambda_{SW}|N', \mathcal{D}_n \sim \\ &\sim \Gamma\left[n + a_6, b_6 + \tfrac{N'}{2}\sum_{i=1}^n t_i^2 + \tfrac{n}{2}\sum_{i=1}^n t_i^2 - \tfrac{1}{2}\sum_{i=1}^n (i-1)t_i^2\right], \end{aligned} \tag{27}$$

and

$$\text{(ii)} \quad N'|\lambda_{SW}, \mathcal{D}_n \sim \mathcal{P}\left(\theta_3\, e^{-\lambda_{SW}\sum_{i=1}^n t_i^2/2}\right).$$

6 Bayesian Inference for the Littlewood and Verral Model

Assuming the LV model where the ith time between failures is assumed to have a exponential density with failure rate λ_i, and a gamma prior density for $\lambda_i \sim \Gamma[\alpha, \psi(i)]$, with $\psi(i) = \beta_0 + \beta_1 i$ (also considered by Mazzuchi and Soyer, 1988), let us consider the following prior densities for α, β_0 and β_1:

$$
\begin{aligned}
&\text{(i)} \qquad && \alpha \sim \Gamma\left[a_7, b_7\right], \\
&\text{(ii)} \qquad && \beta_1 \sim \Gamma\left[a_8, b_8\right], \text{ and} \\
&\text{(iii)} \qquad && \pi\left(\beta_0|\beta_1\right) \propto \left(\beta_0 + \beta_1\right)^{a_9 - 1} exp\left\{-b_9\left(\beta_0 + \beta_1\right)\right\}
\end{aligned}
\tag{28}
$$

where a_7, b_7, a_8, b_8, a_9 and b_9 are known constants.

Thus, a joint prior density for α, β_0 and β_1 is given by,

$$
\begin{aligned}
&\pi\left(\alpha, \beta_0, \beta_1\right) \propto \alpha^{a_7 - 1} e^{-b_7 \alpha}\left(\beta_0 + \beta_1\right)^{a_9 - 1} \times \\
&\times e^{-b_9(\beta_0 + \beta_1)} e^{-b_8} \, . \, \beta_1^{a_8 - 1}.
\end{aligned}
\tag{29}
$$

(this prior density for α, β_0 and β_1 also was assumed by Mazzuchi and Soyer, 1988).

The likelihood function for α, β_0 and β_1 is given by,

$$
L\left(\alpha, \beta_0, \beta_1\right) = \prod_{i=1}^{n} \int_0^\infty f\left(t_i|\lambda_i\right) \pi\left(\lambda_i|\alpha, \psi(i)\right) d\lambda_i \tag{30}
$$

where $f\left(t_i|\lambda_i\right) = \lambda_i e^{-\lambda_i t_i}$

and

$$
\pi\left(\lambda_i|\alpha, \psi(i)\right) = \frac{[\psi(i)]^\alpha}{\Gamma(\alpha)} \lambda_i^{\alpha - 1} e^{-\psi(i)\lambda_i}, \psi(i) = \beta_0 + \beta_1 i.
$$

That is,

$$
L\left(\alpha, \beta_0, \beta_1\right) \propto \alpha^n exp\left\{-\gamma_1 \alpha - \gamma_2\right\}, \tag{31}
$$

where $\gamma_1 = \sum_{i=1}^{n} ln\left(1+\frac{t_i}{\psi(i)}\right)$ and $\gamma_2 = \sum_{i=1}^{n} ln\left(t_i+\psi(i)\right), \psi(i) = \beta_0+\beta_1 i.$

The joint posterior density for α, β_0 and β_1 is given by

$$\begin{aligned} \pi\left(\alpha,\beta_0,\beta_1|\mathcal{D}_n\right) &\propto \alpha^{n+a_7-1}e^{-[b_7+\gamma_1]\ \alpha} \times \\ &\times e^{-\gamma_2}\left(\beta_0+\beta_1\right)^{a_9-1}e^{-(\beta_0+\beta_1)b_9}\beta_1^{a_8-1}e^{-b_8\beta_1} \end{aligned} \tag{32}$$

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

$$\begin{aligned} &\text{(i)} && \alpha|\beta_0,\beta_1,\mathcal{D}_n && \sim \Gamma\left[n+a_7,\gamma_1+b_7\right], \\ &\text{(ii)} && \pi\left(\beta_0|\alpha,\beta_1,\mathcal{D}_n\right) && \propto \left(\beta_0+\beta_1\right)^{a_9-1}e^{-(\beta_0+\beta_1)b_9} \times \\ &&& && \times\ \psi_4\left(\alpha,\beta_0,\beta_1\right) \end{aligned} \tag{33}$$

where $\psi_4\left(\alpha,\beta_0,\beta_1\right) = \exp\left\{-\gamma_1\alpha-\gamma_2\right\}$, and

$$\text{(iii)} \qquad \pi\left(\beta_1|\alpha,\beta_0,\mathcal{D}_n\right) \propto \beta_1^{a_8-1}e^{-b_8\beta_1}\psi_5\left(\alpha,\beta_0,\beta_1\right)$$

where

$$\begin{aligned} \psi_5\left(\alpha,\beta_0,\beta_1\right) &= exp\left\{(a_9-1)\,ln\left(\beta_0+\beta_1\right)-\right. \\ &\left.-b_9\beta_1-\gamma_1\alpha-\gamma_2\right\} \end{aligned}$$

From (33), observe that the variables β_0 and β_1 should be generated using the Metropolis-Hastings algorithm.

The marginal posterior density for $\lambda_i, i=1,2,\ldots,n$ is given by

$$\pi\left(\lambda_i|\mathcal{D}_n\right) = \int\int\int \pi\left(\lambda_i|\alpha,\psi(i)\right)\pi\left(\alpha,\beta_0,\beta_1|\mathcal{D}_n\right)d\beta_0 d\beta_1 d\alpha \tag{34}$$

where

$$\pi\left(\lambda_i|\alpha,\psi(i)\right) = \frac{\left(\beta_0+\beta_{1i}\right)^{\alpha}}{\Gamma(\alpha)}\lambda_i^{\alpha-1}e^{-(\beta_0+\beta_{1i})\lambda_i}\ .$$

Let $\alpha^{(r,s)}, \beta_0^{(r,s)}$ and $\beta_1^{(r,s)}$ denote the variates for α, β_0 and β_1 drawn in the r$^{\text{th}}$ replication and s$^{\text{th}}$ iteration, where R and S are respectively, the total number of simulations and iterations of the Gibbs Sampler. Thus, the marginal posterior density for λ_i can be approximated by

$$\hat{\pi}\left(\lambda_i|\mathcal{D}_n\right) =$$

$$= \frac{2}{RS}\sum_{r=1}^{R}\sum_{s=\frac{S}{2}+1}^{S} \frac{\left(\beta_0^{(r,s)}+\beta_1^{(r,s)}i\right)^{\alpha^{(r,s)}} \lambda_i^{\alpha^{(r,s)}-1} e^{-\left(\beta_0^{(r,s)}+\beta_1^{(r,s)}i\right)\lambda_i}}{\Gamma\left(\alpha^{(r,s)}\right)} \tag{35}$$

where $\lambda_i \geq 0$.

7 Some Considerations on Model Selection

For model selection, we could use the predictive density for t_i given $\mathcal{D}_{(i)}, i = 1,2,\ldots,n$ where $\mathcal{D}_{(i)}$ denotes the set of interfailure times not including t_i, that is, $\mathcal{D}_{(i)} = \{t_1,\ldots,t_{i-1},t_{i+1},\ldots,t_n\}$.

Considering the JM model (1), the predictive density for t_i given $\mathcal{D}_{(i)}$ is given by

$$c_i = f\left(t_i|\mathcal{D}_{(i)}\right) = \sum_{N'=0}^{\infty}\int_0^{\infty} f\left(t_i|N',\lambda_{JM}\right)\pi\left(N',\lambda_{JM}|\mathcal{D}_{(i)}\right)d\lambda_{JM} \tag{36}$$

Using the Gibbs samples, (36) can be approximated by its Monte-Carlo estimate,

$$\hat{f}\left(t_i|\mathcal{D}_{(i)}\right) = \frac{2}{RS}\sum_{r=1}^{R}\sum_{s=\frac{S}{2}+1}^{S} \lambda_{JM}^{(r,s)}\left(N'^{(r,s)}+n-i+1\right)e^{-\lambda_{JM}^{(r,s)}\left(N'^{(r,s)}+n-i+1\right)t_i} \tag{37}$$

where $i = 1,2,\ldots,n$; $\lambda_{JM}^{(r,s)}$ and $N'^{(r,s)}$ are generated by (11) for S iterations in each of R chains considering different initial values for λ_{JM} and N'.

Similary, we could use the Gibbs samples to obtain Monte Carlo estimates for the predictive densities for $t_i, i = 1,2,\ldots,n$ given $\mathcal{D}_{(i)}$ for the MO model (3), the GO model (4) and the SW model (5).

Considering the Littlewood and Verral (1973) model, the predictive density for t_i given $\mathcal{D}_{(i)}$ is

$$c_i = f\left(t_i|\mathcal{D}_{(i)}\right) = \int_0^\infty f\left(t_i|\lambda_i\right)\pi\left(\lambda_i|\mathcal{D}_{(i)}\right)d\lambda_i \tag{38}$$

From (34), we have

$$f\left(t_i|\mathcal{D}_{(i)}\right) = \int_0^\infty \lambda_i e^{-\lambda_i t_i}\left\{\int\int\int \pi\left(\lambda_i|\alpha,\psi(i)\right)\pi\left(\alpha,\beta_0,\beta_1|\mathcal{D}_{(i)}\right)d\alpha\beta_0 d\beta_1\right\}d\lambda_i, \tag{39}$$

where $\lambda_i|\alpha,\psi(i) \sim \Gamma[\alpha,\psi(i)]$ and $\psi(i) = \beta_0 + \beta_1 i$.

Changing the order of integration, we get,

$$\begin{aligned} f\left(t_i|\mathcal{D}_{(i)}\right) &= \\ &= \int\int\int \frac{\alpha\left[\beta_0+\beta_1 i\right]^\alpha}{\left(t_i+\beta_0+\beta_1 i\right)^{\alpha+1}}\pi\left(\alpha,\beta_0,\beta_1|\mathcal{D}_{(i)}\right)d\alpha d\beta_0 d\beta_1 \end{aligned} \tag{40}$$

Thus, a Monte-Carlo estimate for $c_i = f(t_i|\mathcal{D}_{(i)})$ based on the Gibbs samples generated from (33), is given by

$$\hat{f}\left(t_i|\mathcal{D}_{(i)}\right) = \frac{2}{RS}\sum_{r=1}^{R}\sum_{s=\frac{S}{2}+1}^{S}\frac{\alpha^{(r,s)}\left(\beta_0^{(r,s)}+\beta_1^{(r,s)}i\right)^{\alpha^{(r,s)}}}{\left(t_i+\beta_0^{(r,s)}+\beta_1^{(r,s)}i\right)^{\alpha^{(r,s)}+1}}, \tag{41}$$

where $i = 1,2,\ldots,n$ for S iterations in each of R simulations.

We can use $c_i = f(t_i|\mathcal{D}_{(i)})$ in model selection. In this way, we consider plots of c_i versus $i (i = 1,2,\ldots,n)$ for different models; large values of c_i (in average) indicates the better model. We also could choose the model such that $c(l) = \prod_{i=1}^{n} c_i(l)$ is maximum (l indexes models).

8 A Numerical Illustration

In table 1, we have a software reliability data set introduced by Jelinski and Moranda (1972). The data consists of the number of days between the 26 failures that occurred during the production phase of a software (NTDS data - Naval Tactical Data System).

From the data of table 1, we have $n = 26$ and $x_n = x_{26} = 250$.

i	t_i	x_i	i	t_i	x_i	i	t_i	x_i
1	9	9	11	1	71	21	11	116
2	12	21	12	6	77	22	33	149
3	11	32	13	1	78	23	7	156
4	4	36	14	9	87	24	91	247
5	7	43	15	4	91	25	2	249
6	2	45	16	1	92	26	1	250
7	5	50	17	3	95			
8	8	58	18	3	98			
9	5	63	19	6	104			
10	7	70	20	1	105			

Table 1 - NTDS data ($t_i = x_i - x_{i-1}$ is the interfailure time)

Assuming the JM model (1) and the failure truncated model with $x_{26} = 250$, we assume (see (8)) the prior densities $\lambda_{JM} \sim \Gamma[0.2, 20]$ and $N \sim \mathcal{P}(30)$. From the conditional distributions (see (11)), $\lambda_{JM}|N', \mathcal{D}_n \sim \Gamma[26.2, 20 + 250N' + \sum_{i=1}^{26} x_i]$ and $N'|\lambda_{JM}, \mathcal{D}_n \sim \mathcal{P}(30e^{-250\lambda_{JM}})$, where $N' = N - 26$, we generated 5 separate Gibbs chains each of which ran for 1000 iterations. Convergence of the Gibbs Samplers are monitored by the method of Gelman and Rubin(1992). For each parameter we selected the last 200 iterations, which for 5 chains yields a sample of size 1000. In table 2, we have the obtained posterior summaries for the parameters λ_{JM} and N', and in figure 1, we have the approximate marginal posterior densities considering the $S = 1000$ Gibbs samples. It is interesting to observe that the maximum likelihood estimators for λ_{JM} and N are given by $\hat{\lambda}_{JM} = 0.006904$ and $\hat{N} = 31.100(\hat{N}' = 5.1)$.

	Mean	Median	S.D.	95% Credible interval
λ_{JM}	0.00686	0.00667	0.00205	(0.00352,0.011329)
N'	5.933	5.000	3.720	(0,14)

Table 2 - Posterior summaries for the JM model.

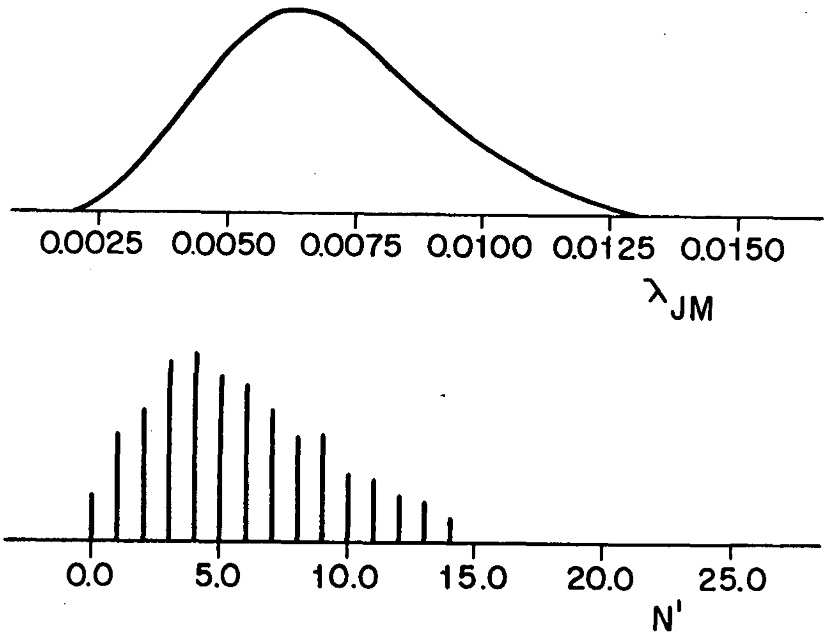

Figure 1 - Marginal posterior densities for λ_{JM} and N' (JM model)

Assuming the MO model (3), we consider (from (13)) the prior densities $\theta \sim \Gamma[24.2, 110]$ and $k \sim B[66.8, 5.8]$. From the conditional distributions (see (15)).

$\theta | k, \mathcal{D}_n \sim \Gamma\left[50.2, 110 + \sum_{i=1}^{26} k^{i-1} t_i\right]$ and $\pi(k|\theta, \mathcal{D}_n) \propto k^{66.8-1}(1-k)^{5.8-1}\psi_1(\theta, k)$, where $\psi_1(\theta, k) = exp\left\{325 lnk - \theta \sum_{i=1}^{26} k^{i-1} t_i\right\}$, we generated 5 separate Gibbs-within-Metropolis chains each of which ran for 1000 iterations. For each parameter, we selected the 205^{th}, 210^{th},..., 995^{th}, 1000^{th} iterations, which for 5 chains yields a sample of size 1000. In table 3, we have the obtained posterior summaries for the parameters θ and k, and in figure 2, we have the approximate marginal posterior densities considering the $S = 1000$ Gibbs samples.

	Mean	Median	S.D.	95% Credible interval
θ	0.22247	0.22230	0.03560	(0.15470,0.29684)
k	0.94526	0.94584	0.01266	(0.92061,0.96925)

Table 3 - Posterior summaries for the MO model.

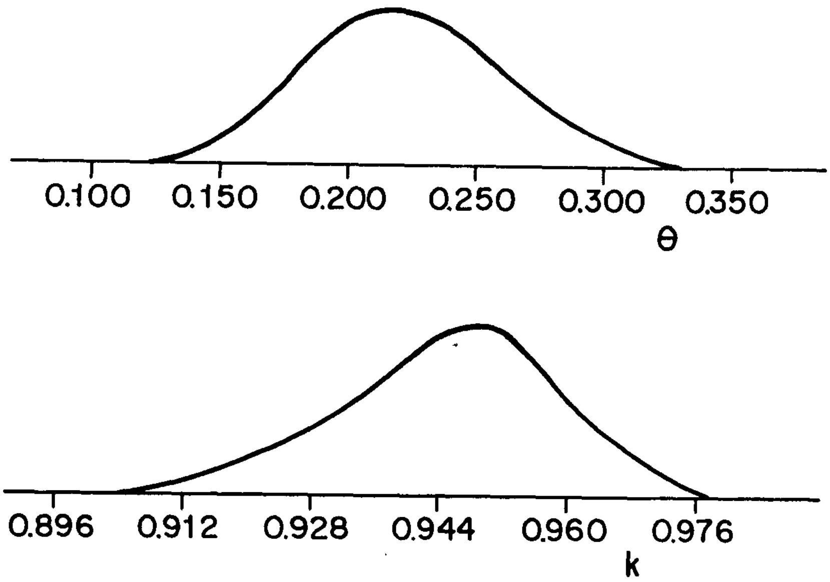

Figure 2 - Marginal posterior densities for θ and k (MO model)

Assuming the GO model (4), we consider (from (19)) the prior densities $\theta \sim \mathcal{P}(30)$, $\lambda_{GO} \sim \Gamma[0.2, 20]$ and $p \sim B[2.5, 2.5]$. From (21), the conditional distributions for the Gibbs-within-Metropolis algorithm are given by $\lambda_{GO} \mid N, p, \mathcal{D}_n \sim \Gamma[26.2, 20 + p\sum_{i=1}^{26} x_i + 250\,(N - 26p)]$, $\pi(N|\lambda_{GO}, p, \mathcal{D}_n) \propto (e^{-30}30^N/N!)\psi_2(N, p, \lambda_{GO})$ where $\psi_2(N, p, \lambda_{GO}) = exp\{ln\, A(n,p) - 250(N - 26p)\lambda_{GO}\}$, and $\pi(p|N, \lambda_{GO}, \mathcal{D}_n) \propto p^{2.5-1}(1-p)^{2.5-1}\psi_3(N, p, \lambda_{GO})$, where $\psi_3(N, p, \lambda_{GO}) = exp\{lnA(N,p) - p\lambda_{GO}\sum_{i=1}^{n} x_i - 250(N - 26p)\lambda_{GO}\}$. Also generating 5 separate Gibbs-within-Metropolis chains each of which with 1000 iterations, we selected for each parameter, the 205^{th}, 210^{th}, 215^{th},..., 995^{th}, 1000^{th} iterations, which for 5 chains yields a sample of size 1000. In table 4, we have the obtained posterior summaries for the parameters λ_{GO}, θ and N, and in figure 3 we have the approximate marginal posterior densities considering the $S = 1000$ Gibbs Samples.

	Mean	Median	S.D.	95% Credible interval
λ_{GO}	0.00606	0.00555	0.00249	(0.002728,0.01204)
N	28.743	29.000	5.417	(19,40)
p	0.61224	0.62766	0.17883	(0.25035,0.90689)

Table 4 - Posterior summaries for the GO model.

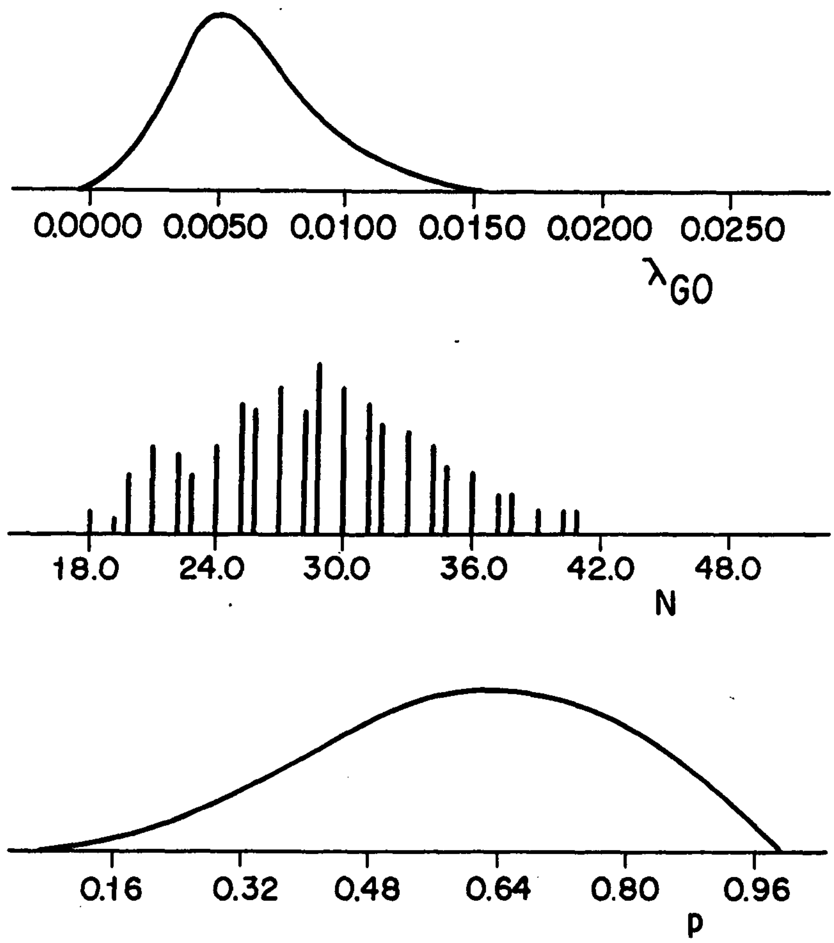

Figure 3 - Marginal posterior densities for λ_{GO}, N and p (GO model).

Assuming the SW model (5), consider (from (24)) the prior density $\lambda_{SW} \sim \Gamma[6.25, 625]$ and $N \sim \mathcal{P}(30)$. From the conditional distributions (27), $\lambda_{SW}|N', \mathcal{D}_n \sim \Gamma[32.25, 23950 + 5157N']$, and $N'|\lambda_{SW}, \mathcal{D}_n \sim \mathcal{P}(30e^{-5157\lambda_{SW}})$, where $N' = N - 26$, we generated 5 separate Gibbs chains each of which ran for 1000 iterations. For each parameter, we selected the last 200 iterations, which for 5 chains yields a sample of size 1000. In table 5, we have the obtained posterior summaries for the parameters λ_{SW} and N', and in figure 4, we have the marginal posterior densities considering the $S = 1000$ Gibbs Samples.

	Mean	Median	S.D.	95% Credible interval
λ_{SW}	0.00133	0.00131	0.00024	(0.00089,0.00186)
N'	0.05000	0.00000	0.23570	(0,1)

Table 5 - Posterior summaries for the SW model.

Assuming the LV model (see section 6), consider (from (28)) the prior densities

$\alpha \sim \Gamma[1, 0.3], \beta_1 \sim \Gamma[10, 100]$ and $\pi(\beta_0|\beta_1) \propto (\beta_0 + \beta_1)^{20-1} exp\{-(\beta_0 + \beta_1)\}$. From the conditional distributions (see (33)),

$$\alpha|\beta_0, \beta_1, \mathcal{D}_n \sim \Gamma[27, \gamma_1 + 0.3],$$

$$\pi(\beta_0|\alpha, \beta_1, \mathcal{D}_n) \propto (\beta_0 + \beta_1)^{20-1} exp\{-(\beta_0 + \beta_1)\}\psi_4(\alpha, \beta_0, \beta_1),$$

where $\psi_4(\alpha, \beta_0, \beta_1) = exp\{-\gamma_1\alpha - \gamma_2\}$ and $\pi(\beta_1|\alpha, \beta_0, \mathcal{D}_n) \propto \beta_1^{10-1} e^{-100\beta_1}\psi_5(\alpha, \beta_0, \beta_1)$, where $\psi_5(\alpha, \beta_0, \beta_1) = exp\{-19ln(\beta_0+\beta_1)-\beta_1-\gamma_1\alpha-\gamma_2\}$ and $\gamma_1 = \sum_{i=1}^{26} ln\left(1 + \frac{t_i}{\beta_0 + \beta_1 i}\right)$ and $\gamma_2 = \sum_{i=1}^{26} ln(t_i + \beta_0 + \beta_1 i)$, we generated 10 separate Gibbs-within-Metropolis chains each of which ran for 1000 iterations. For each parameter, we selected the 208^{th}, $216^{\text{th}}, \ldots$, iteractions which for 10 chains yields a sample of size 1000. In table 6, we have the obtained posterior summaries for the parameters α, β_0 and β_1 and in figure 4, we have the approximate marginal posterior densities considering the $S = 1000$ Gibbs Samples.

	Mean	Median	S.D.	95% Credible interval
α	3.3760	3.3179	0.8528	(1.87826,5.10630)
β_0	19.8450	19.5420	4.3670	(11.2302,29.3354)
β_1	0.10017	0.09627	0.03154	(0.04707,0.17062)

Table 6 - Posterior summaries for the LV model.

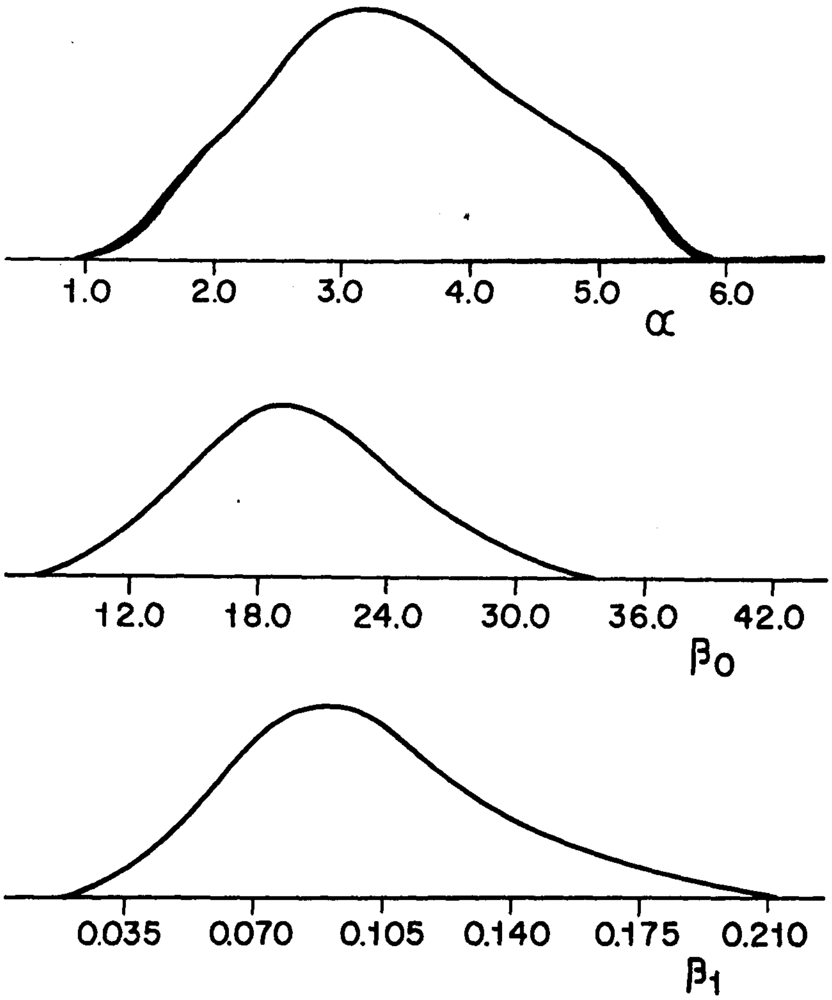

Figure 4 - Marginal posterior densities for α, β_0, and β_1 (LV model).

In table 7, we have the values of the predictive densities $c_i = f(t_i|\mathcal{D}_{(i)})$ (see section 7) evaluated at the observed values t_i (see table 1) and approximated by its Monte Carlo estimates based on $S = 1000$ Gibbs Samples for each considered model. In figure 5, we have the plots of c_i against i for all 5 models. We also have in table 7, the values for $c(l) = \prod_{i=1}^{26} c_i$, considering the different models.

From table 7 and figure 5, observe that $c(l)$ has the largest value considering the GO model (4), which indicates, in general, better fit for the software reliability data of table 1.

If we observe the last values of table 1 (also see figure 5), we observe that the LV model gives better prediction for the last observations, which is very important to prediction of future interfailure times.

i	LV model	JM model	SW model	GO model	MO model
1	0.032310	0.03111	0.075598	0.03643	0.03003
2	0.021150	0.01776	0.036750	0.02323	0.01713
3	0.024317	0.02238	0.050620	0.02749	0.02246
4	0.074348	0.08717	0.095400	0.08076	0.08769
5	0.044337	0.04951	0.098660	0.05092	0.05077
6	0.108993	0.12355	0.052920	0.10722	0.11908
7	0.061969	0.07120	0.094780	0.06798	0.07104
8	0.038010	0.04316	0.088760	0.04455	0.04477
9	0.061893	0.07029	0.088280	0.06701	0.06906
10	0.044479	0.05136	0.089980	0.05107	0.05187
11	0.131581	0.12472	0.021140	0.11088	0.11111
12	0.052265	0.05913	0.083050	0.05753	0.05759
13	0.130553	0.11456	0.018530	0.10505	0.10077
14	0.033028	0.03984	0.076270	0.04000	0.04026
15	0.073259	0.07245	0.056230	0.07028	0.06688
16	0.129043	0.09869	0.014600	0.09607	0.08689
17	0.087400	0.07529	0.037720	0.07511	0.06855
18	0.087214	0.07181	0.034180	0.07318	0.06584
19	0.052223	0.05169	0.052720	0.05303	0.04914
20	0.127089	0.07635	0.009330	0.08364	0.07118
21	0.025396	0.03229	0.053790	0.03263	0.03187
22	0.002960	0.00711	0.006230	0.00546	0.00714
23	0.044725	0.03962	0.032910	0.04466	0.04054
24	0.000176	0.00055	0.000000	0.00116	0.00038
25	0.103058	0.04397	0.005410	0.06223	0.05200
26	0.124274	0.04010	0.001390	0.06397	0.05268
$c(l)$	4.15×10^{-36}	3.29×10^{-36}	1.55×10^{-43}	1.73×10^{-35}	1.61×10^{-36}

Table 7 - Values for $c_i = f\left(t_i|\mathcal{D}_{(i)}\right), i = 1, 2, \ldots, 26.$

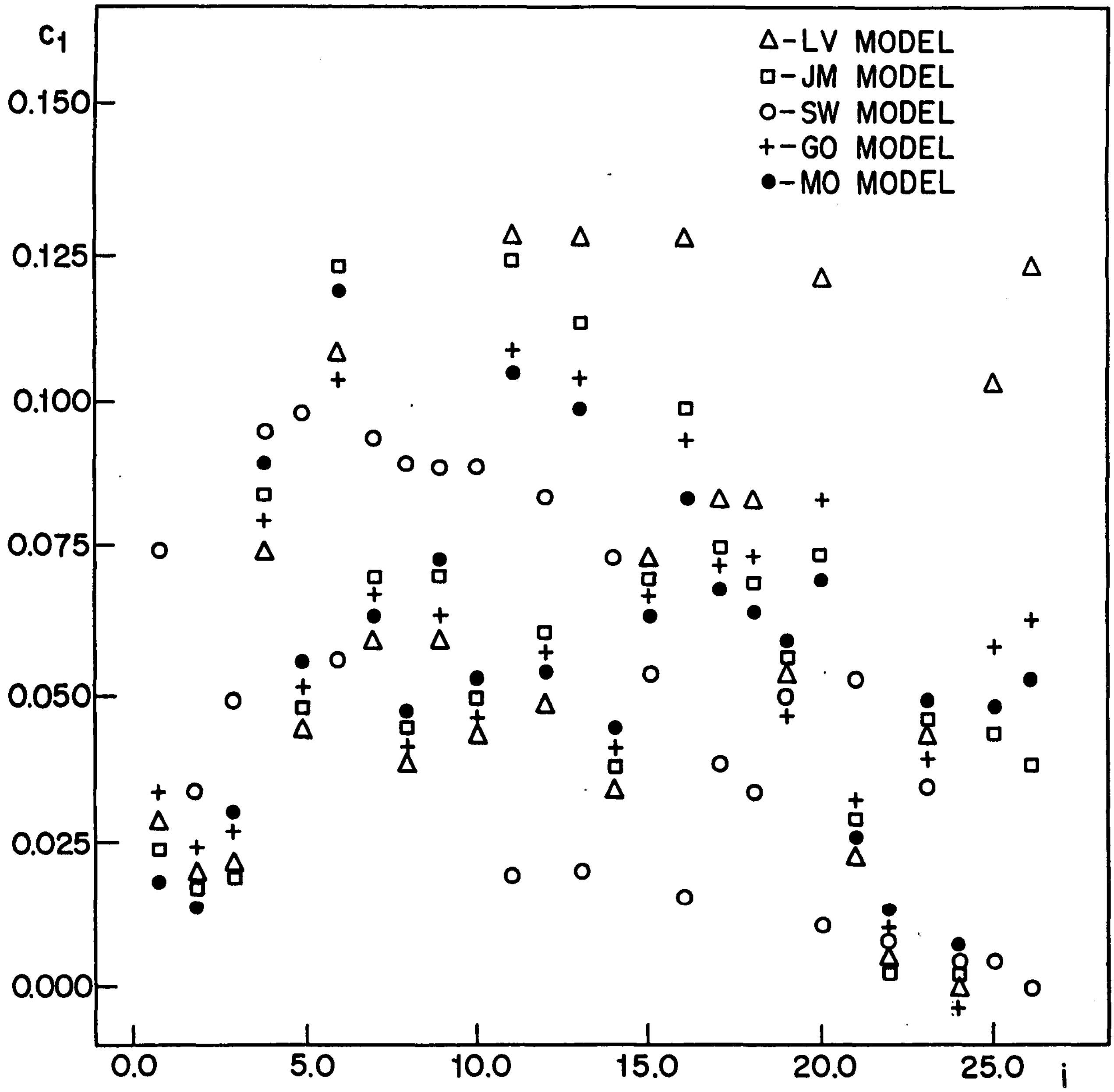

Figure 5 - Plots of c_i versus i for *LV* model, *JM* model, *SW* model, *GO* model and *MO* model.

References

Achcar, J.A.; Dey, D.K.; Niverthi, M. (1996) - A Bayesian approach using nonhomogeneous Poisson process for software reliability models, technical report # 96-07, Department of Statistics, University of Connecticut, Storrs, U.S.A.

Chib, S.; Greenberg, E. (1995) - Understanding the Metropolis - Hastings algorithm, The 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.

Goel, A.L. (1983) - A guidebook for software reliability assessment. Technical report RADC-TR-83-176, U.S.A.

Goel, A.L.; Okumoto, K. (1978) - An analysis of recurrent software failures on a real-time control system, 496-500. Proc. ACM Ann. Tech. Conf., Washington, D.C., U.S.A.

Goel, A.L.; Okumoto, K. (1979) - Time-dependent error detection rate model for software reliability and other performance measures, IEEE Transactions on Reliability, R-28, 206-211.

Jelinski, Z.; Moranda, P.B. (1972) - Software reliability research. In Statistical Computer performance evaluation, ed. W. Freiberger, New York: Academic Press, 465-497.

Kuo, L.; Yang, T.Y. (1995) - Bayesian Computation for nonhomogeneous Poisson process in software reliability. To appear in the Journal of the American Statistical Association.

Kuo, L.; Lee, J.; Choi, K.; Yang, T.Y. (1996) - Bayesian inference for S-shaped software reliability growth models. Technical report # 96-05, Department of Statistics, University of Connecticut, Storrs, U.S.A.

Littlewood, B.; Verral, J.L. (1973) - A Bayesian reliability growth model for computer software, Applied Statistics, 22, 332-346.

Mazzuchi, T.A.; Soyer, R. (1988) - A Bayes empirical-Bayes model for software reliability, IEEE Transactions on reliability, R-37:2,248-258.

Moranda, P.B. (1975) - Prediction of software reliability and its applications, Proceedings of the Annual Reliability and Maintainability Symposium, 327-332, Washington, D.C., U.S.A.

Musa, J.D.; Okumoto, K. (1984) - A logarithmic Poisson execution time model for software reliability measurement, Proceedings of the 7^{th} International Conference of Software engineering, Orlando, FL, U.S.A, 230-237.

Okba, M.; Yamada, S.; Takeda, K.; Osaki, S. (1982) - S-shaped software reliability growth curve: how good is it? COMPSAC'82, 38-44.

Schick, G.J.; Wolverton, R.W. (1978) - Assessment of software reliability, Proc. Oper. Res., 395-422. Wirzberg-Wien: Physica-Verlag.

Singpurwalla, N.D.; Wilson, S.P. (1994) - Software reliability modeling, International Statistical Review, 62,3,289-317.

Yang, T.Y. (1994) - Computational approaches to Bayesian inference for software reliability. PhD Thesis, Department of Statistics, University of Connecticut, Storrs, U.S.A.

NOTAS DO ICMSC

SÉRIE ESTATÍSTICA

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.

026/96 ANDRADE, M.G.; VAL, J.B.R. do - Um método numérico baseado na solução do valor médio para a equação de Helmholtz parte II: Rede triangular.

025/96 ANDRADE, M.G.; VAL, J.B.R. do - Um método numérico baseado na solução do valor médio para a equação de Helmholtz parte I: malha quadrada.

024/96 MAZUCHELLI, J.; ACHCAR, J.A. - Análise bayesiana para modelos de crescimento.

023/96 ANDRADE, M.G.; SOARES, S.; CRUZ Jr., G.; VINHAL, C.D.N - Uma abordagem estocástica para o planejamento a longo prazo da operação de sistema hidrotérmicos.

022/95 ACHCAR, J. - A generalized Moranda software reliability model: a bayesian approach.

021/95 RODRIGUES, J. - Inference for the software reliability using asymmetric loss functions: A hierarchical Bayes approach.

020/95 RODRIGUES, J.; BOLFARINE, H.; CORDEIRO, G.M. - Nonlinear quasi-bayesian theory and the inverse linear regression.