UNIVERSIDADE DE SÃO PAULO

A Bayesian Approach Using Nonhomogeneous Poisson Process for Software Reliability Models

JORGE A. ACHCAR
DIPAK K. DEY
MURALI NIVERTHI

Nº 27

NOTAS

Instituto de Ciências Matemáticas de São Carlos

Instituto de Ciências Matemáticas de São Carlos

ISSN - 0103-2577

A Bayesian Approach Using Nonhomogeneous Poisson Process for Software Reliability Models

JORGE A. ACHCAR
DIPAK K. DEY
MURALI NIVERTHI

Nº 27

NOTAS DO ICMSC
Série Estatística

São Carlos
Fev. /1996

RESUMO

Métodos Bayesianos usando processos não-homogeneos são considerados para modelagem de problemas de confiabilidade de software. Modelos de estatísticas de ordem gama generalizado e log-normal são considerados para modelar as épocas das falhas do software. Algoritmos de Metrópolis com etapas de Gibbs são propostos para desenvolver inferência Bayesiana para esses modelos. Algumas técnicas Bayesianas de diagnóstico são desevolvidas e incorporadas para verificar as suposições do modelo proposto. A seleção de modelos é baseada nos valores de predição ordenados. A metodologia desenvolvida neste artigo é exemplificada com um conjunto de dados introduzido por Jelinski e Morada (1972).

A Bayesian Approach Using Nonhomogeneous Poisson Process for Software Reliability Models

Jorge A. Achcar
Department of Computer Sciences and Statistics
ICMSC, University of São Paulo
C.P. 668, 13560-970
São Carlos, S.P., Brazil

Dipak K. Dey and Murali Niverthi
Department of Statistics
University of Connecticut
Storrs, CT 06269

February 22, 1996

Abstract

Bayesian approach using nonhomogeneous Poisson process is considered for modeling software reliability problems. A generalized gamma and lognormal order statistics models are considered to model epochs of the failures of software. Metropolis algorithms along with Gibbs steps are proposed to perform the Bayesian inference of such models. Some Bayesian model diagnostics are developed and incorporated to verify further modeling assumptions. Model selection based on a prequential likelihood of conditional predictive ordinates is considered. The methodology developed in this paper is exemplified with a software reliability data set introduced by Jelinski and Moranda (1972).

Key Words: Generalized Gamma intensity function, General order statistics, Gibbs sampling, Lognormal intensity function, Metropolis algorithm, Model diagnostics, Model selection, Nonhomogeneous Poisson process.

1 Introduction

The modelling of the number of failures of a software could be based on a point process to count failures (see for example, Singpurwalla and Wilson, 1994; or Musa, Iannino and Okumoto, 1987). Let $M(t)$ be the cumulative number of failures of the software that are observed during time $(0, t]$ and $M(t)$ is modeled by a nonhomogeneous Poisson process with mean value function $m(t)$. The different models of this type specify a different function $m(t)$. The cumulative number of failures $M(t)$ can also be specified by its intensity function $\lambda(t)$ which is the derivative of $m(t)$ with respect to t; either of these functions completely specify a particular nonhomogeneous Poisson process, with distribution,

$$P\{M(t) = n\} = \frac{\{m(t)\}^n}{n!} e^{-m(t)} \tag{1}$$

where $n = 0, 1, 2, ...$

If λ is a constant (so that $m(t)$ is linear), then $M(t)$ is called a homogeneous Poisson process; otherwise it is called a nonhomogeneous Poisson process (NHPP).

Many different choices for the mean value function $m(t)$ are considered in the literature. Goel and Okumoto (1979) assume that the expected number of software failures to time t, given by the mean value function $m(t)$, is nondecreasing and bounded above. Specifically, they consider the mean function,

$$m_1(t) = \theta(1 - e^{-\beta t}), \tag{2}$$

where θ represents the expected number of errors in the software and β is considered to be the fault detection rate. From (2), we have $\lambda_1(t) = m_1'(t) = \theta\beta e^{-\beta t}$.

Often, the rate of faults in software increases initially before eventually decreasing. In view of that, Goel (1983) proposes a generalization of model (2) given by the intensity function,

$$\lambda_2(t) = \theta\beta\alpha t^{\alpha-1} e^{-\beta t^{\alpha}} \tag{3}$$

where θ is still the total number of bugs and the parameters β and α describe the quality of testing.

From (3) it follows that the corresponding mean function

$$m_2(t) = \theta(1 - e^{-\beta t^\alpha}). \tag{4}$$

Observe that (2) and (4) can be written as a special case of a general form where the mean value function is given as

$$m(t) = \theta F(t). \tag{5}$$

Here we assume that there is an unknown number N of faults at the beginning of software debugging and we model the epochs of failures to be the first n order statistics taken from N i.i.d. observations with density f supported in R^+ and c.d.f. F (of a general order statistics model). Also observe that in this case, $\lim_{t\to\infty} m(t)$ is finite and this process can be denoted by NHPP-I.

From (5), we observe that if $F(t) = 1 - e^{-\beta t}$, we have the Goel and Okumoto (1979) process (an exponential order statistics model); if $F(t) = 1 - e^{-\beta t^\alpha}$, we have the Goel (1983) process (a Weibull order statistics model); whereas if $F(t) = 1 - (1 + \beta t)e^{-\beta t}$, we have the Ohba-Yamada (1982, 1983) process.

Other NHPP processes are also proposed to model software reliability data. In situations where new faults are introduced during debugging, we need to replace the general order statistics (GOS) model with a record value statistics model (RVS). In these cases, we have $m(t) \to \infty$ as $t \to \infty$, and we denote such as NHPP-II processes (see Yang, 1994). A special case is given with $m(t) = -\ln(1 - F(t))$,which reduces to the cases of the Musa-Okumoto (1984) process with $\lambda(t) = \alpha/(t+\beta)$, the Duane (1964) process with $\lambda(t) = \alpha\beta t^{\alpha-1}$ and the Cox and Lewis (1966) process with $\lambda(t) = \exp(\alpha+\beta t)$.

Recently, Kuo and Yang (1995) developed a unified approach for software reliability growth models considering the two different classes of models: (i) using general order statistics models (GOS) and, (ii) using record value statistics models (RVS). They also introduce a Bayesian analysis for NHPP-I software reliability models considering some special choices of $F(t)$ in (5) (exponential, Weibull, Pareto and extreme value order statistics models) and using Gibbs sampling with Metropolis-Hastings algorithms.

Kuo, Lee, Choi and Yang (1996) present Bayesian inference for the Ohba-Yamada (1982, 1983) process by considering a further extension given by the

NHPP-gamma-k model with,

$$F(t) = 1 - e^{-\beta t} \sum_{j=0}^{k-1} \frac{(\beta t)^j}{j!}. \tag{6}$$

They used Metropolis-within-Gibbs algorithm for the Bayesian analysis assuming a known value of k. Observe that if $k = 1$ in (6), we have the Goel and Okumoto (1979) process and if $k = 2$, we have the Ohba-Yamada process (1982, 1983).

In this paper we present Bayesian inference for software reliability models using Metropolis-within-Gibbs algorithms for two special classes of GOS models: a supermodel given by the generalized gamma order statistics model which incorporates some standard order statistics software reliability models and the log-normal order statistics model. The use of these two classes of models give great flexibility for the shape of the intensity function $\lambda(t)$, which implies in better fit for software reliability data. We also explore some Bayesian model selection criteria considering a software reliability data set introduced by Jelinski and Moranda (1972).

An outline of the paper is as follows. In Section 2, a Bayesian inference for generalized order statistics software reliability models are developed. Section 3 specializes to generalized gamma order statistics models. Gibbs algorithms are presented for special cases of such models. Section 4 discusses Bayesina inference for the lognormal order statistics models. Bayesian inference and model selection procedures are described in Section 5. Section 6 considers a numerical illustration of Bayesian model fitting along with model comparisons.

2 Bayesian Inference for GOS Software Reliability Models

Let us assume that the mean value function $m(t)$ is indexed by the unknown parameters θ and $\underset{\sim}{\beta}$.

Given the time truncated model testing until time t, the ordered epochs of the observed n failure times are denoted by $x_1, x_2, ..., x_n$.

The likelihood function is given by

$$L_{NHPP}(\theta, \underset{\sim}{\beta} | \, D_t) = \{\prod_{i=1}^{n} \lambda(x_i)\} \exp\{-m(t)\} \tag{7}$$

where $D_t = \{n; x_1, x_2, ..., x_n; t\}$ is the data set (see for example, Cox and Lewis, 1966; or Lawless, 1982, p. 495).

For the failure truncated model a similar expression to (7) can be applied with t replaced by x_n.

For Bayesian inference of GOS software reliability models (5), we consider the use of Metropolis-within-Gibbs algorithms (see for example, Gelfand and Smith, 1990). Since the presence of the expression $m(t) = \theta F(t)$ for the NHPP in the likelihood function (7) usually prevent us in specifying a convenient form for the conditional density of $\underset{\sim}{\beta}$ and θ given D_t needed in the Gibbs sampling, we consider the introduction of a latent variable $N' = N - n$ which has a Poisson distribution with parameter $\theta[1 - F(t \mid \underset{\sim}{\beta})]$ (see Yang, 1994; or Kuo and Yang, 1995).

In this way, the posterior distribution $p(\theta, \underset{\sim}{\beta} \mid D_t)$ can be obtained from the joint density $p(\theta, \underset{\sim}{\beta}, N' \mid D_t)$ by marginalization. The joint posterior density $p(\theta, \underset{\sim}{\beta}, N' \mid D_t)$ is approximated from the Gibbs samplers drawn from the following conditional densities: $p(N' \mid \theta, \underset{\sim}{\beta}, D_t)$, $p(\theta \mid N', \underset{\sim}{\beta}, D_t)$ and $p(\underset{\sim}{\beta} \mid N', \theta, D_t)$.

3 Bayesian Inference for the Generalized Gamma Order Statistics Model

Let us assume that $F(t)$ given in (5) is the c.d.f. of a generalized gamma distribution, that is, the corresponding mean function

$$m_{GG}(t) = \theta I_k(\beta t^{\alpha}), \tag{8}$$

where $I_k(s)$ is the incomplete gamma integral given by

$$I_k(s) = \frac{1}{\Gamma(k)} \int_0^s x^{k-1} e^{-x} dx.$$

From (8), we get the intensity function

$$\lambda_{GG}(t) = m'_{GG}(t) = \frac{1}{\Gamma(k)} \theta \beta^k \alpha t^{\alpha k - 1} e^{-\beta t^{\alpha}}. \tag{9}$$

Observe that several usual order statistics models which are already considered in the literature are special cases of the generalized gamma order statistics model:

(i) If $k = 1$, $\lambda_{GG}(t)$ reduces to $\lambda_2(t)$ given in (3) (a Goel (1983) process).

(ii) If $k = 1$, $\alpha = 1$, $\lambda_{GG}(t)$ reduces to $\lambda_1(t) = \theta\beta e^{-\beta t}$ (a Goel and Okumoto (1979) process).

(iii) If $\alpha = 1$, $\lambda_{GG}(t)$ reduces to,

$$\lambda_3(t) = \theta \frac{\beta^k}{\Gamma(k)} t^{k-1} e^{-\beta t} \tag{10}$$

which is a $NHPP - gamma - k$ process with $m_3(t) = \theta F(t)$ where $F(t)$ is given by (6). We denote this NHPP process by a gamma order statistics model.

(iv) If $k \to \infty$, we have a normal order statistics model (see for example, Lawless (1982)).

Considering the generalized order statistics model (8), the likelihood function for θ, α, β and k is given (from (7)) by,

$$\begin{aligned} &L_{NHPP}(\theta, \alpha, \beta, k \mid D_t) \\ &= \left\{ \prod_{i=1}^{n} \theta \tfrac{\beta^k}{\Gamma(k)} \alpha x_i^{\alpha k - 1} e^{-\beta x_i^\alpha} \right\} \exp\left\{ - \int_0^t \theta \tfrac{\beta^k}{\Gamma(k)} \alpha u^{\alpha k-1} e^{-\beta u^\alpha} du. \right\} \end{aligned} \tag{11}$$

That is,

$$\begin{aligned} L_{NHPP}(\theta, \alpha, \beta, k \mid D_t) = {} & \tfrac{\theta^n \alpha^n \beta^{kn}}{\{\Gamma(k)\}^n} \left\{ \prod_{i=1}^{n} x_i^{\alpha k-1} \right\} \\ & \times \exp\left\{ -\beta \sum_{i=1}^{n} x_i^\alpha - \theta I_k(\beta t^\alpha) \right\} \end{aligned} \tag{12}$$

For Bayesian inference, considering the introduction of a latent variable $N' = N - n$ (see section 2), we assume the following prior densities for $N', \theta, \alpha, \beta$ and k :

$$
\begin{aligned}
&(i)\ N' \sim P\{\theta[1 - I_k(\beta t^\alpha)]\},\\
&(ii)\ \theta \sim \Gamma(a, b);\ a, b \text{ known},\\
&(iii)\ \beta \sim \Gamma(c, d);\ c, d \text{ known},\\
&(iv)\ \alpha \sim \pi_1(\alpha), \text{ where } \pi_1(\alpha) \text{ is a prior density for } \alpha (\alpha > 0),\\
&(v)\ k \sim \pi_2(k), \text{ where } \pi_2(k) \text{ is a prior density for } k (k > 0).
\end{aligned} \tag{13}
$$

Here, $P(\lambda)$ denotes a Poisson distribution with parameter λ, $\Gamma(a, b)$ denotes a gamma distribution with mean a/b and variance a/b^2. We further assume independence among the parameters θ, α, β and k.

The joint posterior density is,

$$
\begin{aligned}
&P(N', \alpha, \beta, k, \theta \mid D_t) \propto \frac{\theta^{N'+n+a-1} \alpha^n \beta^{kn+c-1}}{N'! \{\Gamma(k)\}^n}\\
&\times \left\{ \prod_{i=1}^{n} x_i^{\alpha k - 1} \right\} \{1 - I_k(\beta t^\alpha)\}^{N'} e^{-(b+1)\theta - (d + \sum_{i=1}^{n} x_i^\alpha)\beta} \pi_1(\alpha) \pi_2(k)
\end{aligned} \tag{14}
$$

For the failure truncated model, similar expressions to (12) and (14) can be applied with t replaced by x_n.

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

$$
\begin{aligned}
&(i)\ N' \mid \alpha, \beta, k, \theta, D_t \sim P[\theta(1 - I_k(\beta t^\alpha))],\\
&(ii)\ \theta \mid N', \alpha, \beta, k, D_t \sim \Gamma[a + n + N', b + 1],\\
&(iii)\ p(\alpha \mid \theta, N', \beta, k, D_t) \propto \alpha^n \left\{ \prod_{i=1}^{n} x_i^{\alpha k - 1} \right\} e^{-\beta \sum_{i=1}^{n} x_i^\alpha} \{1 - I_k(\beta t^\alpha)\}^{N'} \pi_1(\alpha),\\
&(iv)\ p(\beta \mid N', \alpha, k, \theta, D_t) \propto \beta^{kn+c-1} e^{-\beta[\sum_{i=1}^{n} x_i^\alpha + d]} \{1 - I_k(\beta t^\alpha)\}^{N'}\\
&(v)\ p(k \mid N', \alpha, \beta, \theta, D_t) \propto \frac{\beta^{kn}}{\{\Gamma(k)\}^n} \left\{ \prod_{i=1}^{n} x_i^{\alpha k - 1} \right\} \{1 - I_k(\beta t^\alpha)\}^{N'} \pi_2(k).
\end{aligned} \tag{15}
$$

The variables α, β and k should be generated using Metropolis-Hastings algorithm (see for example, Chib and Greenberg, 1994).

3.1 Gibbs Algorithms for Special Cases of the Generalized Gamma Order Statistics Model

Assuming $k = 1$ in (8), we have a Weibull order statistics model (Goel (1983) process). In this case, $1 - I_k(\beta t^\alpha) = e^{-\beta t^\alpha}$ and the joint posterior density for

N', α, β and θ is given (from (14)) by,

$$p(N', \alpha, \beta, \theta \mid D_t) \propto \frac{1}{N'!}\theta^{N'+n+a-1}\alpha^n\beta^{n+c-1} \times \left\{\prod_{i=1}^{n} x_i^{\alpha-1}\right\} e^{-(b+1)\theta} e^{-(\sum_{i=1}^{n} x_i^{\alpha}+N't^{\alpha}+d)\beta}\pi_1(\alpha). \tag{16}$$

The marginal posterior densities for the Gibbs algorithm are given by

$$\begin{aligned}
&(i)\ N' \mid \theta, \alpha, \beta, D_t \sim P(\theta e^{-\beta t^{\alpha}}) \\
&(ii)\ \theta \mid N', \alpha, \beta, D_t \sim \Gamma(a+n+N', b+1) \\
&(iii)\ \beta \mid N', \alpha, \theta, D_t \sim \Gamma(n+c, N't^c + d + \textstyle\sum_{i=1}^{n} x_i^{\alpha}) \text{ and} \\
&(iv)\ p(\alpha \mid N', \theta, \beta, D_t) \propto \alpha^n (\textstyle\prod_{i=1}^{n} x_i^{\alpha}) e^{-(N't^{\alpha}+\sum_{i=1}^{n} x_i^{\alpha})\beta}\pi_1(\alpha).
\end{aligned} \tag{17}$$

Observe that when $k = 1$, we only need to use the Metropolis-Hastings algorithm to generate the variable α.

Assuming $\alpha = 1$ in (8), we have a gamma order statistics model. The joint posterior density for N', β, θ and k is given by

$$p(N', \beta, \theta, k \mid D_t) \propto \frac{1}{N'!\{\Gamma(k)\}^n}\theta^{N'+n+a-1}\beta^{kn+c-1} \times \left\{\prod_{i=1}^{n} x_i^{k-1}\right\} e^{-(b+1)\theta} e^{-(d+\sum_{i=1}^{n} x_i)\beta} \{1 - I_k(\beta t)\}^{N'} \pi_2(k). \tag{18}$$

The marginal posterior densities for the Gibbs algorithm are given by

$$\begin{aligned}
&(i)\ N' \mid \beta, \theta, k, D_t \sim P[\theta(1 - I_k(\beta t))], \\
&(ii)\ \theta \mid N', \beta, k, D_t \sim \Gamma[a+n+N', b+1], \\
&(iii)\ p(\beta \mid N', \theta, k, D_t) \propto \beta^{kn+c-1} e^{-(\sum_{i=1}^{n} x_i + d)\beta} \{1 - I_k(\beta t)\}^{N'} \\
\text{and}\ &(iv)\ p(k \mid N', \beta, \theta, D_t) \propto \frac{\beta^{kn}}{\{\Gamma(k)\}^n}\left\{\prod_{i=1}^{n} x_i^{k-1}\right\}\{1 - I_k(\beta t)\}^{N'} \pi_2(k).
\end{aligned} \tag{19}$$

When $\alpha = 1$, we need to use the Metropolis-Hastings algorithm to generate the variables β and k.

When $\alpha = 1$ and $k = 1$ (a Goel and Okumoto (1979) process), we have the exponential order statistics model. In this case, the joint posterior density for N', θ and β is given (from (14)) by

$$p(N', \theta, \beta \mid D_t) \propto \frac{1}{N'!}\theta^{N'+n+a-1}\beta^{n+c-1} e^{-(b+1)\theta} e^{-(\sum_{i=1}^{n} x_i + N't + d)\beta}. \tag{20}$$

With $\alpha = 1$, $k = 1$, the marginal posterior densities for the Gibbs algorithm are given by

$$\begin{array}{l}(i)\ N' \mid \theta, \beta, D_t \sim P(\theta e^{-\beta t}) \\ (ii)\ \theta \mid N', \beta, D_t \sim \Gamma[a + n + N', b + 1] \\ (iii)\ \beta \mid N', \theta, D_t \sim \Gamma[n + c, d + N't + \sum_{i=1}^{n} x_i].\end{array} \tag{21}$$

4 Bayesian Inference for the Log-Normal Order Statistics Model

Let us assume that $F(t)$ given in (5) is the c.d.f. of a log-normal distribution, that is, the mean function is

$$m_{LN}(t) = \theta \Phi_z \left(\frac{\ln t - \mu}{\sigma} \right), \tag{22}$$

where $\Phi_z(\cdot)$ is the c.d.f. of a standard normal distribution. The intensity function is given by

$$\lambda_{LN}(t) = \frac{\theta}{t\sqrt{2\pi}\sigma} \exp\left\{ -\frac{(\ln t - \mu)^2}{2\sigma^2} \right\}. \tag{23}$$

The likelihood function for θ, μ and σ is given (from (7)) by

$$\begin{array}{l} L_{NHPP}(\theta, \mu, \sigma) = \frac{\theta^n}{(2\pi)^{n/2}\sigma^n} \left\{ \prod_{i=1}^{n} 1/x_i \right\} \\ \times \exp\left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^{n} (\ln x_i - \mu)^2 \right\} \exp\left\{ -\theta \Phi_z\left(\frac{\ln t - \mu}{\sigma}\right) \right\}. \end{array} \tag{24}$$

Considering the introduction of a latent variable $N' = N - n$ (see section 2), we assume that following prior densities for N', θ, μ and σ :

(i) $N' \sim P[\theta(1 - \Phi_z(\frac{\ln t - \mu}{\sigma}))]$,
(ii) $\theta \sim \Gamma[a, b]$; a, b known,
(iii) $\mu \sim \pi_3(\mu)$, where $\pi_3(\mu)$ is a prior density for $\mu(-\infty < \mu < \infty)$, and
(iv) $\sigma \sim \pi_4(\sigma)$, where $\pi_4(\sigma)$ is a prior density for σ $(\sigma > 0)$.

(25)

The joint posterior density is

$$p(N',\theta,\mu,\sigma \mid D_t) \propto \frac{1}{\sigma^n N'!}\theta^{n+N'+a-1}e^{-\theta(b+1)}$$
$$\times\left\{1-\Phi_z\left(\frac{\ln t-\mu}{\sigma}\right)\right\}^{N'}\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(\ln x_i-\mu)^2\right\}\pi_3(\mu)\pi_4(\sigma). \tag{26}$$

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

$$\begin{aligned}
&(i)\ N' \mid \theta,\mu,\sigma,D_t \sim P[\theta(1-\Phi_z(\tfrac{\ln t-\mu}{\sigma}))],\\
&(ii)\ \theta \mid N',\mu,\sigma,D_t \sim \Gamma[a+n+N',b+1],\\
&(iii)\ \mu \mid N',\theta,\sigma,D_t \propto \exp\{-\tfrac{1}{2\sigma^2}\sum_{i=1}^{n}(\ln x_i-\mu)^2\}\times\{1-\Phi_z(\tfrac{\ln t-\mu}{\sigma})\}^{N'}\pi_3(\mu),\\
\text{and}\quad &(iv)\ \sigma \mid N',\mu,\theta,D_t \propto \exp\{-\tfrac{1}{2\sigma^2}\sum_{i=1}^{n}(\ln x_i-\mu)^2\}\times\sigma^{-n}\{1-\Phi_z(\tfrac{\ln t-\mu}{\sigma})\}^{N'}\pi_4(\sigma).
\end{aligned} \tag{27}$$

Observe that, we again need to use the Metropolis-Hastings algorithm to generate the variables μ and σ.

5 Bayesian Inference and Model Determination

We can use the Gibbs samplers to get inferences on the parameters of the software reliability model or functions of these parameters. In this case, we could approximate posterior moments of interest. As a special case, consider the number of remaining errors in the software given by $\varepsilon(t) = m(\infty) - m(t)$. Considering the generalized gamma order statistics model, we have $\varepsilon(t) = \theta(1 - I_k(\beta t^\alpha))$ and a Bayes estimator of $\varepsilon(t)$ with respect to the squared error loss function is given by

$$E[\varepsilon(t) \mid D_t] = E[\theta(1 - I_k(\beta t^\alpha)) \mid D_t]. \tag{28}$$

Let $\theta^{(r,s)}$, $k^{(r,s)}$, $\alpha^{(r,s)}$ and $\beta^{(r,s)}$ denote the variates for θ, k, α and β drawn in the r^{th} iteration and the s^{th} replication where R and S are respectively the total number of iterations and simulations of the Gibbs sampler. Then (28) can be estimated by

$$\widehat{\varepsilon}(t) = \frac{2}{RS}\sum_{s=1}^{S}\sum_{r=\frac{R}{2}+1}^{R}\theta^{(r,s)}(1 - I_{k^{(r,s)}}(\beta^{(r,s)}t^{\alpha^{(r,s)}})). \tag{29}$$

Similar inferences could be obtained for $m(t)$ and also by considering other order statistics models.

For model checking, observe that, if the model (5) is correct $m(t)/\theta = F(t)$ has a standard uniform distribution. Therefore, at each failure time we could consider Empirical Q-Q plots of the Monte Carlo estimates of $m(t)/\theta$ versus Uniform(0,1) distribution for each failure time and for each model. Departure from uniform distribution indicates model inadequacy.

For model selection, we could consider the prequential conditional predictive ordinate (PCPO) as suggested by Dawid (1984), for the future epoch x_{i+1} defined by $c_i = p(x_{i+1} \mid D_{x_i})$, the conditional density of X_{i+1} evaluated at the future observed time x_{i+1} given $(x_1, x_2, ..., x_i)$. The sequence $\{X_i\}_{i\geq 1}$ is Markov chain (see Yang, 1994, pp 52). Therefore, we can write $c_i = p(x_{i+1} \mid x_i)$. The PCPO can be computed by

$$\begin{aligned} p(x_{i+1} \mid x_i) &= \int p(x_{i+1} \mid \underset{\sim}{\beta}, D_{x_i}) p(\underset{\sim}{\beta}, D_{x_i}) d\, \underset{\sim}{\beta} \\ &= \int \lambda_{II}(x_{i+1}) \exp\{-m_{II}(x_{i+1}) + m_{II}(x_i)\} p(\underset{\sim}{\beta}, D_{x_i}) d\, \underset{\sim}{\beta} \end{aligned} \tag{30}$$

where $m_{II}(t) = -\ln(1 - F(t))$ is the mean value function of a NHPP-II process considering the record value statistics in a given interval and $\lambda_{II}(t)$ is the intensity function (also the hazard function of F, that is, $\lambda_{II}(t) = m'_{II}(t) = f(t)/(1 - F(t))$.

Considering the generalized gamma order statistics model (8) with $F(t) = I_k(\beta t^\alpha)$ we have $m_{II}(t) = -\ln(1 - I_k(\beta t^\alpha))$ and $\lambda_{II}(t) = \alpha\beta^k t^{\alpha k-1} e^{-\beta t^\alpha}/\{\Gamma(k)[1 - I_k(\beta t^\alpha)]\}$.

Thus, the PCPO for the future epoch x_{i+1}, considering the generalized gamma order statistics model is given

$$p(x_{i+1} \mid x_i) = \int\int\int \frac{\alpha\beta^k x_{i+1}^{\alpha k-1} e^{-\beta x_{i+1}^\alpha}}{\Gamma(k)\{1 - I_k(\beta x_i^\alpha)\}} p(\alpha, \beta, k \mid D_{x_i}) d\alpha d\beta dk. \tag{31}$$

Special cases of the generalized gamma order statistics model are easily obtained from (31); for example, if $k = 1$ (a Weibull order statistics model), we have $1 - I_k(\beta x_i^\alpha) = e^{-\beta x_i^\alpha}$ and,

$$p(x_{i+1} \mid x_i) = \int\int \alpha\beta x_{i+1}^{\alpha-1} \exp\left\{-\beta x_{i+1}^\alpha + \beta x_i^\alpha\right\} p(\alpha, \beta \mid D_{x_i}) d\alpha d\beta. \tag{32}$$

Using the Gibbs samplers (31) can be approximated by its Monte Carlo estimate

$$\widehat{p}(x_{i+1} \mid x_i) = \frac{2}{RS} \sum_{s=1}^{S} \sum_{r=\frac{R}{2}+1}^{R} \frac{\alpha^{(r,s)} \beta^{(r,s)^{k^{(r,s)}}} x_{i+1}^{\alpha^{(r,s)} k^{(r,s)}-1} e^{-\beta^{(r,s)} x_{i+1}^{\alpha^{(r,s)}}}}{\Gamma(k^{(r,s)})\{1-I_k(\beta^{(r,s)} x_i^{\alpha^{(r,s)}})\}}. \tag{33}$$

Considering the log-normal order statistics model with mean value function (22) we have from (30)

$$m_{II}(t) = -\ln\left\{1 - \Phi z(\frac{\ln t - \mu}{\sigma})\right\}$$

and

$$\lambda_{II}(t) = m'_{II}(t) = \frac{\exp\{-(\ln t - \mu)^2/2\sigma^2\}}{\sqrt{2\pi}\sigma t\{1 - \Phi z(\frac{\ln t-\mu}{\sigma})\}}.$$

In this case, the PCPO for the future epoch x_{i+1} considering the log-normal order statistics model is given by

$$p(x_{i+1} \mid x_i) = \int\int \frac{\exp\{-(\ln x_{i+1} - \mu)^2/2\sigma^2\} p(\mu, \sigma \mid D_{x_i}) d\mu d\sigma}{\sqrt{2\pi}\sigma x_{i+1}\{1 - \Phi z(\frac{\ln x_i - \mu}{\sigma})\}}. \tag{34}$$

We can use the obtained PCPO $c_i = p(x_{i+1} \mid x_i)$ in model selection. In this way, we could consider plots of c_i versus i ($i = 1, 2, \ldots$) for different models; large value of c_i (in average) indicates the better model.

We also could choose the model such that $c(\ell) = \prod_{i=1}^{n} c_i(\ell)$ is maximum (ℓ indexes models).

6 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 observe that $n = 26$, $\sum_{i=1}^{n} \ln x_i = 112.4776$ and $x_n = x_{26} = 250$.

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

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			

Assuming the generalized gamma order statistics model (8) and the failure truncated model with t replaced by $x_{26} = 250$, we consider (from (13)) the priors $N' \sim P\{\theta[1 - I_k(\beta(250)^\alpha)]\}$, $\theta \sim \Gamma(60, 2)$, $\beta \sim \Gamma(10, 2)$ and noninformative prior densities for α and k given by $\pi_1(\alpha) \propto 1/\alpha$, $\alpha > 0$ and $\pi_2(k) \propto 1/k$, $k > 0$. From the marginal posterior densities for $N', \theta, \alpha, \beta$ and k given in (15), we generated 10 separate Gibbs chains each of which ran for 1000 iterations, and we monitored the convergence of the Gibbs samplers 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 consider the 110^{th}, 120^{th}, ...,1000^{th} iteration, which for 10 chains yields a sample of size 900. In Table 2, we have the obtained posterior summaries for the parameters $N', \theta, \alpha, \beta$ and k and in figure 1 we have the approximate marginal posterior densities considering $S = 900$ Gibbs samples. It is interesting to observe that the maximum likelihood estimators for θ, α, β and k obtained by using the software SAS are given by $\hat{\theta} = 29.4301$, $k = 18.2010$, $\hat{\alpha} = 0.2748$ and $\hat{\beta} = 5.1279$.

Usually, software engineers consider special cases of the supermodel (8) to analyze software reliability data.

Considering $k = 1$ (a Goel (1983) process) and assuming the failure truncated model, assume the priors $N' \sim P(\theta e^{-(250)^\alpha \beta})$, $\theta \sim \Gamma(90, 3)$, $\beta \sim \Gamma(26, 160)$ and a noninformative prior density for α given by $\pi_2(\alpha) = 1/\alpha$, $\alpha > 0$. In table 3, we have posterior summaries for the parameters N', θ, α

and β using $S = 900$ samples obtained from 10 Gibbs chains each of 1000 iterations (from (17)).

When $\alpha = 1$ (a Gamma order statistics model) and assuming the prior densities $N' \sim P[\theta(1 - I_k(250\beta))]$, $\theta \sim \Gamma(90, 3)$, $\beta \sim \Gamma(12, 680)$ and a noninformative prior density for k proportional to $1/k$, $k > 0$, we have in table 4, the posterior summaries for the parameters of the model considering $S = 900$ samples obtained from 10 Gibbs chains each of 1000 iterations of the marginal posterior densities given by (19).

Considering $k = 1$ and $\alpha = 1$ (a Goel and Okumoto (1979) process) and assuming the prior densities $N' \sim P(\theta e^{-250\beta})$, $\theta \sim \Gamma(60, 2)$, $\beta \sim \Gamma(5, 1000)$, with θ independent of β, we have in table 5, posterior summaries for the parameters N', θ and β considering $S = 900$ samples obtained from 10 Gibbs chains each of 1000 iterations of the marginal posterior densities (21).

In table 6, we have posterior summaries for the parameters N', θ, μ and σ of the log-normal order statistics model. Considering $S = 1000$ samples obtained from the Metropolis-within-Gibbs algorithm generated using the marginal posterior densities (27) and assuming the priors $N' \sim P[\theta(1 - \Phi_z(\frac{\ln 250 - \mu}{\sigma}))]$, $\theta \sim \Gamma(36; 1.2)$ and the noninformative priors for μ and σ given by $\pi_3(\mu) \propto$ constant, $-\infty < \mu < \infty$ and $\pi_4(\sigma) \propto 1/\sigma$, $\sigma > 0$.

In table 7, we have approximated Bayes estimates for the mean value function $m(t)$ with respect to the squared error loss considering the different models and using the Gibbs samplers.

In figure 2, we have the plots of the PCPO (see (30)) c_i against i for all considered models and observe that the Generalized Gamma Order Statistics Model gives the largest PCPO values.

In table 8, we have the values of $c(\ell) = \prod_{i=2}^{n} c_i(\ell)$, $\ell = 1, 2, \ldots, 5$ considering the 5 different models.

We also have in table 8, the sums of relative errors defined by $RE(\ell) = \sum_{i=1}^{26} (n_i - \widehat{m}(x_i))^2/\widehat{m}(x_i)$, where n_i is the observed number of bugs in the interval $(0, x_i]$, $\widehat{m}(x_i)$ is the Bayesian estimator for the mean value function and ℓ indexes models. We observe better fit of the generalized gamma order statistics model for the NTDS data of table 1 (bigger value for $c(\ell)$ and smaller value for $RE(\ell)$).

From figure 3, it follows that the Generalized Gamma Order Statistics model gives values of $m(t)/\theta$ closest to the Uniform(0,1) model.

Table 2. Posterior Summaries for the Generalized Gamma Order Statistics Model

	Mean	Median	S.D.	95% Credible Interval
N'	1.997	1.000	2.472	(0, 8)
θ	28.78	28.21	5.190	(19.91, 39.58)
α	0.303	0.308	0.056	(0.195, 0.399)
β	5.188	4.732	2.107	(2.096, 10.170)
k	18.93	18.88	5.131	(10.42, 31.25)

Table 3. Posterior Summaries for the Weibull Order Statistics Model (Goel (1983) Process; k=1)

	Mean	Median	S.D.	95% Credible Interval
N'	6.994	7.000	3.96	(1, 16)
θ	31.02	30.89	2.93	(26.00, 37.37)
α	0.479	0.478	0.062	(0.356, 0.605)
β	0.111	0.111	0.023	(0.071, 0.159)

Table 4. Posterior Summaries for the Gamma Order Statistics Model ($\alpha = 1$)

	Mean	Median	S.D.	95% Credible Interval
N'	2.20	2.00	1.998	(0, 7)
θ	29.61	29.41	2.780	(24.65, 35.56)
β	0.017	0.0169	0.0042	(0.0099, 0.0262)
k	1.890	1.854	0.404	(1.188, 2.840)

Table 5. Posterior Summaries for the Exponential Order Statistics Model (Goel and Okumoto (1979) Process; $k = 1$, $\alpha = 1$)

	Mean	Median	S.D.	95% Credible Interval
N'	7.971	7.000	4.400	(1, 17)
β	0.0059	0.0057	0.0016	(0.0032, 0.0094)
θ	31.242	31.074	3.484	(25.09, 38.79)

Table 6. Posterior Summaries for the Lognormal Order Statistics Model

	Mean	Median	S.D.	95% Credible Interval
N'	4.21	2.00	5.440	(0, 21)
θ	30.12	29.73	4.405	(22.68, 40.21)
μ	4.57	4.48	0.413	(4.13, 5.96)
σ	0.852	0.779	0.299	(0.576, 1.694)

Table 7. Bayes Estimators for $m(x_i)$, $i = 1, 2, ..., 26$

i	x_i	n_i	Generalized Gamma	Gamma ($\alpha = 1$)	Weibull ($k = 1$)	Exponential ($k = 1, \alpha = 1$)	Log-Normal
1	9	1	0.3846	0.5818	8.4327	1.5897	0.1411
2	21	2	2.0812	2.1213	11.7309	3.5758	1.1853
3	32	3	4.3854	3.9483	13.6585	5.2713	3.0620
4	36	4	5.31023	4.6694	14.2275	5.8599	3.8828
5	43	5	6.9691	5.9725	15.1072	6.8558	5.4090
6	45	6	7.4451	6.3508	15.3360	7.1326	5.8561
7	50	7	8.6263	7.3020	15.8716	7.8099	6.9785
8	58	8	10.4612	8.8234	16.6368	8.8515	8.7491
9	63	9	11.5574	9.7633	17.0676	9.4772	9.8186
10	70	10	13.0122	11.0534	17.6201	10.3218	11.2487
11	71	11	13.2119	11.2346	17.6947	10.4396	11.4459
12	77	12	14.3654	12.3034	18.1224	11.1314	12.5889
13	78	13	14.5502	12.4782	18.1905	11.2443	12.7726
14	87	14	16.1163	14.0042	18.7678	12.2302	14.3377
15	91	15	16.7573	14.6534	19.0055	12.6515	14.9828
16	92	16	16.9124	14.8128	19.0633	12.7552	15.1393
17	95	17	17.3656	15.2839	19.2329	13.0628	15.5975
18	98	18	17.8009	15.7441	19.3972	13.3649	16.0392
19	104	19	18.6197	16.6318	19.7107	13.9530	16.8745
20	105	20	18.7497	16.7755	19.7612	14.0490	17.0077
21	116	21	20.0653	18.2748	22.2848	15.0679	18.3662
22	149	22	22.9720	21.9213	21.5821	17.7547	21.4633
23	156	23	23.4314	22.5449	21.8159	18.2606	21.9703
24	247	24	26.7075	27.3343	24.0499	23.2699	25.8712
25	249	25	26.7454	27.3909	24.0870	23.3534	25.9213
26	250	26	26.7640	27.4187	24.1054	23.3948	25.9460

Table 8. Values for c(l) and RE(l) for all the considered models

Models	c(l)	R.E(l)
Gen. Gamma Order Statistics Model	$9.576*10^{-50}$	6.775
Weibull Order Statistics Model (k = 1)	$9.874*10^{-58}$	68.589
Gamma Order Statistics Model ($\alpha = 1$)	$1.325*10^{-51}$	23.376
Exponential Order Statistics Model ($\alpha = 1$, k = 1)	$3.587*10^{-57}$	221.450
Lognormal Order Statistics Model	$1.130*10^{-51}$	7.972

ACKNOWLEDGMENTS

The first author acknowledges financial support of the Brazilian institution, FAPESP, to visit the Department of Statistics of the University of Connecticut

References

[1] Chib, S. and Greenberg, E. (1994). Understanding the Metropolis-Hastings algorithm. Manuscript.

[2] Cox, D.R. and Lewis, P.A. (1966). *Statistical Analysis of Series of Events.* Methuen.

[3] Dawid, A.P. (1984). Statistical theory: the prequential approach. Journal of the Royal Statistical Society, Ser. A, 147, 278-292.

[4] Duane, J.T. (1964). Learning curve approach to reliability monitoring. IEEE Transaction on Aerospace, AS-2(2), 563-566.

[5] Gelfand, A.E. and Smith, A.F.M. (1990). Sampling-based approaches to calculating marginal densities. Journal of the American Statistical Association, 85, 398-409.

[6] Gelman, A.E. and Rubin, D. (1992). Inference from iterative simulation using multiple sequences. Statistical Science, 7, 457-472.

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

[8] Goel, A.L. (1983). A guidebook for software reliability assessment. Technical Report RADC-TR-83-176.

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

[10] Jelinski, Z. and Moranda, P.B. (1972). Software Reliability Research. In *Statistical Computer Performance Evaluation*, ed. W. Freiberger, New York: Academic Press, pp. 465-497.

[11] Kuo, L. and Yang, T.Y. (1995). Bayesian computation for nonhomogeneous Poisson process in software reliability. To appear in Journal of the American Statistical Association.

[12] Kuo, L., Lee, J., Choi, K. and Yang, T.Y. (1996). Bayesian inference for S-shaped software reliability growth models. Technical Report #96-05, Department of Statistics, University of Connecticut.

[13] Lawless, J.F. (1982). *Statistical Models and Methods for Lifetime Data.* New York: John Wiley & Sons.

[14] Musa, J.D. and Okumoto, K. (1984). A logarithmic Poisson execution time model for software reliability measurement. In *Proceedings of Seventh International Conference on Software Engineering*, Orlando, pp. 230-238.

[15] Musa, J.D., Iannino, A. and Okumoto, K. (1987). *Software Reliability: Measurement, Prediction, Application.* New York: Wiley.

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

[17] Singpurwalla, N. and Wilson, S.P. (1994). Software reliability modeling. International Statistical Review, 62, 289-317.

[18] Yamada, S., Ohba, M. and Osaki, S. (1983). S-shaped reliability growth modeling for software error detection. IEEE Transactions on Reliability, 32, 475-478.

[19] Yang, T.Y. (1994). Computational approaches to Bayesian inference for software reliability. Ph.D. Thesis, Department of Statistics, University of Connecticut, Storrs.

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 $p(\theta_s \mid \underset{\sim}{\theta}_{(s)})$, $s = 1, 2, ..., k$ where $\underset{\sim}{\theta}_{(s)}$ denotes the random vector of $k-1$ random variables with the 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 $p(\theta_1 \mid \theta_2^{(0)}, \theta_3^{(0)}, ..., \theta_k^{(0)})$. Similarly, generate a value $\theta_2^{(1)}$ from the conditional density $p(\theta_2 \mid \theta_1^{(1)}, \theta_2^{(2)}, ..., \theta_k^{(0)})$ and continue up to the value $\theta_k^{(1)}$ from the conditional density $p(\theta_k \mid \theta_1^{(1)}, \theta_2^{(1)}, ..., \theta_{k-1}^{(1)})$.

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 $p(\theta_s \mid \underset{\sim (s)}{\theta})$ is available in closed form, then

$$\widehat{p}(\theta_s) = \frac{1}{B} \sum_{g=1}^{B} p(\theta_s \mid \underset{\sim (s)g}{\theta}),$$

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

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 $p(\theta_i \mid \underset{\sim (i)}{\theta})$. Observe that $p(\theta_i \mid \underset{\sim (i)}{\theta})$ 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 which 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 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 untilt he process reaches a stationary distribution.

FIGURE 1. Density plots for parameters of the Generalized Gamma Order Statistics Model

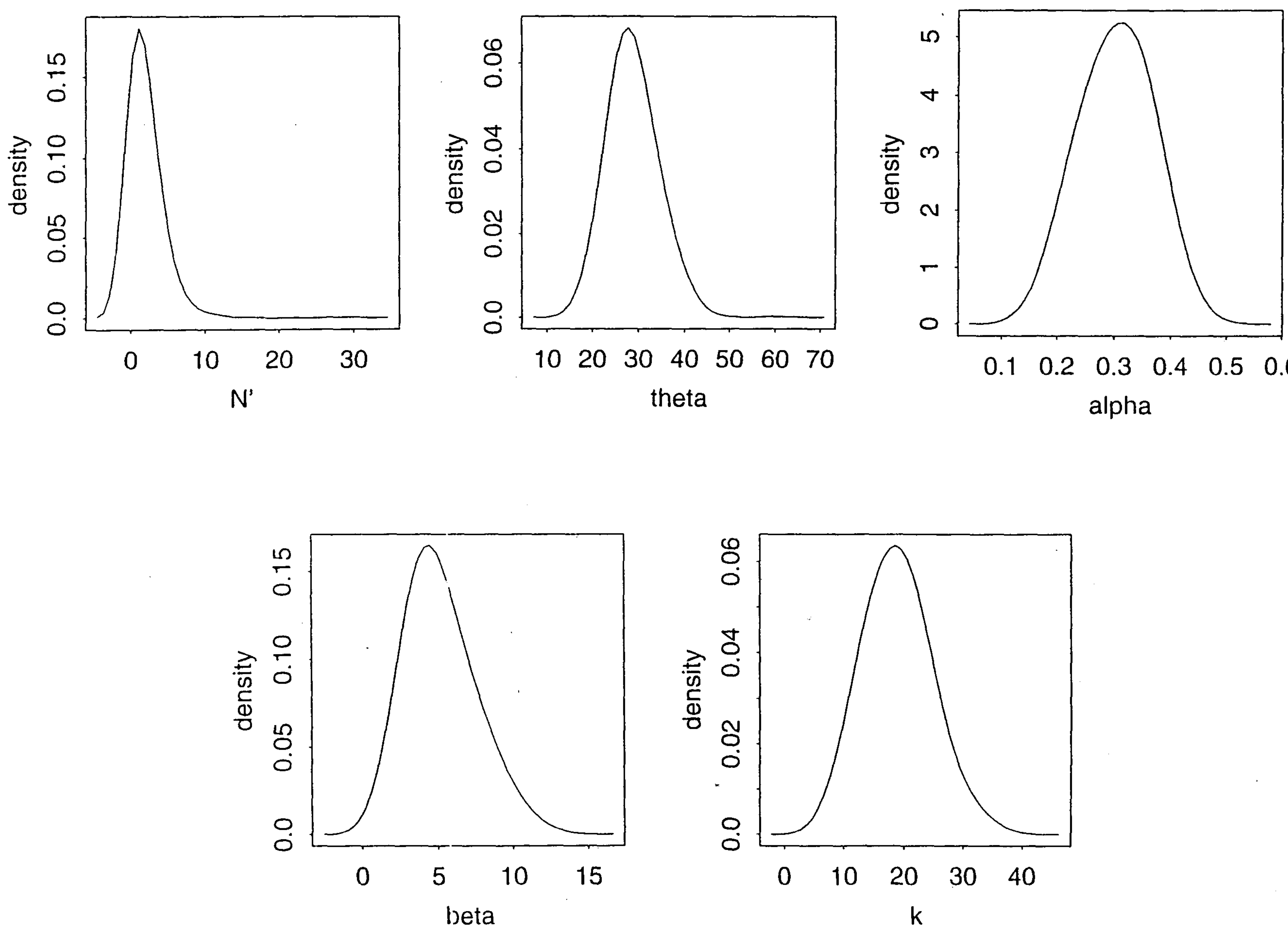


FIGURE 2. PCPO plots for the five models

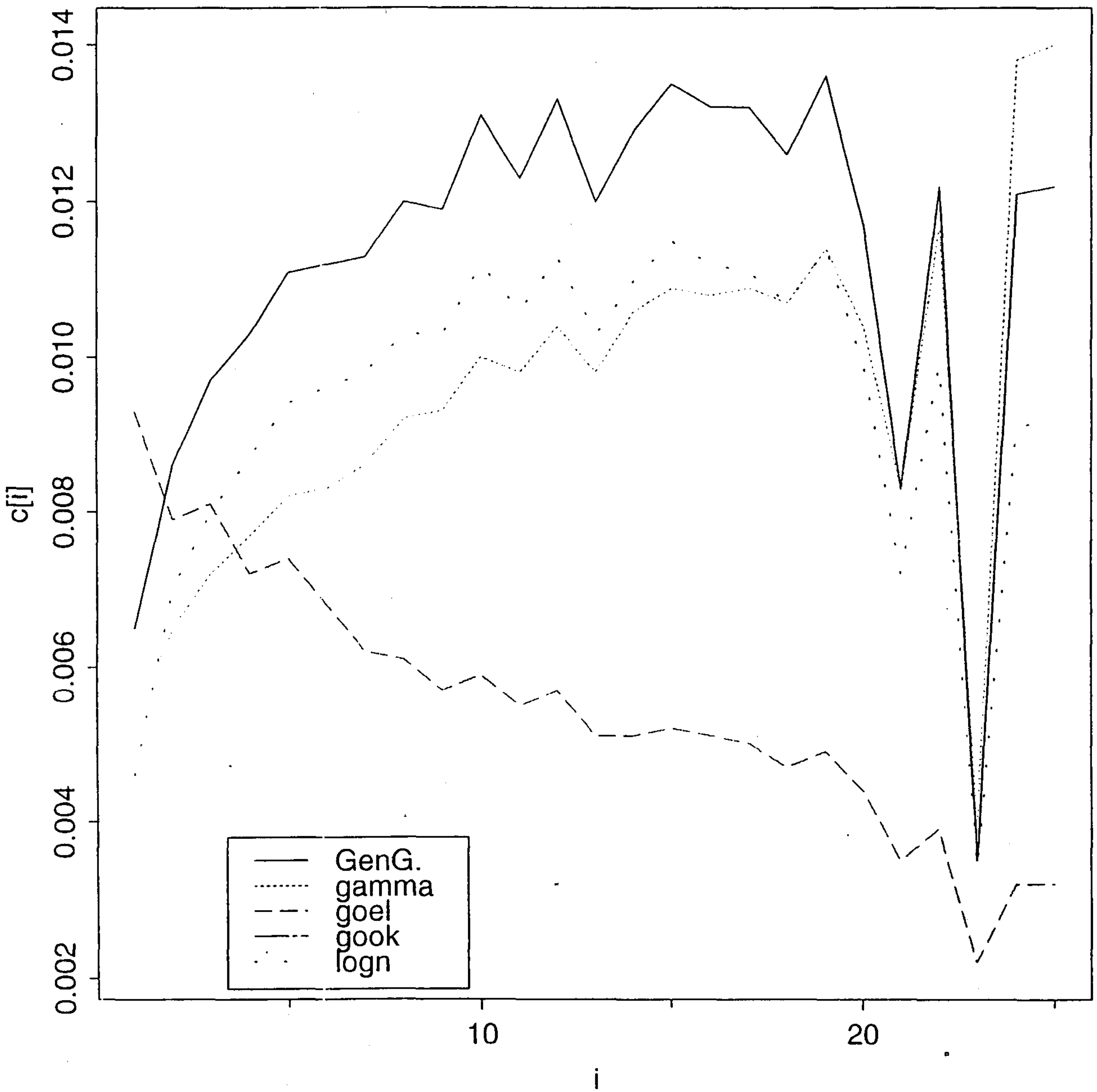
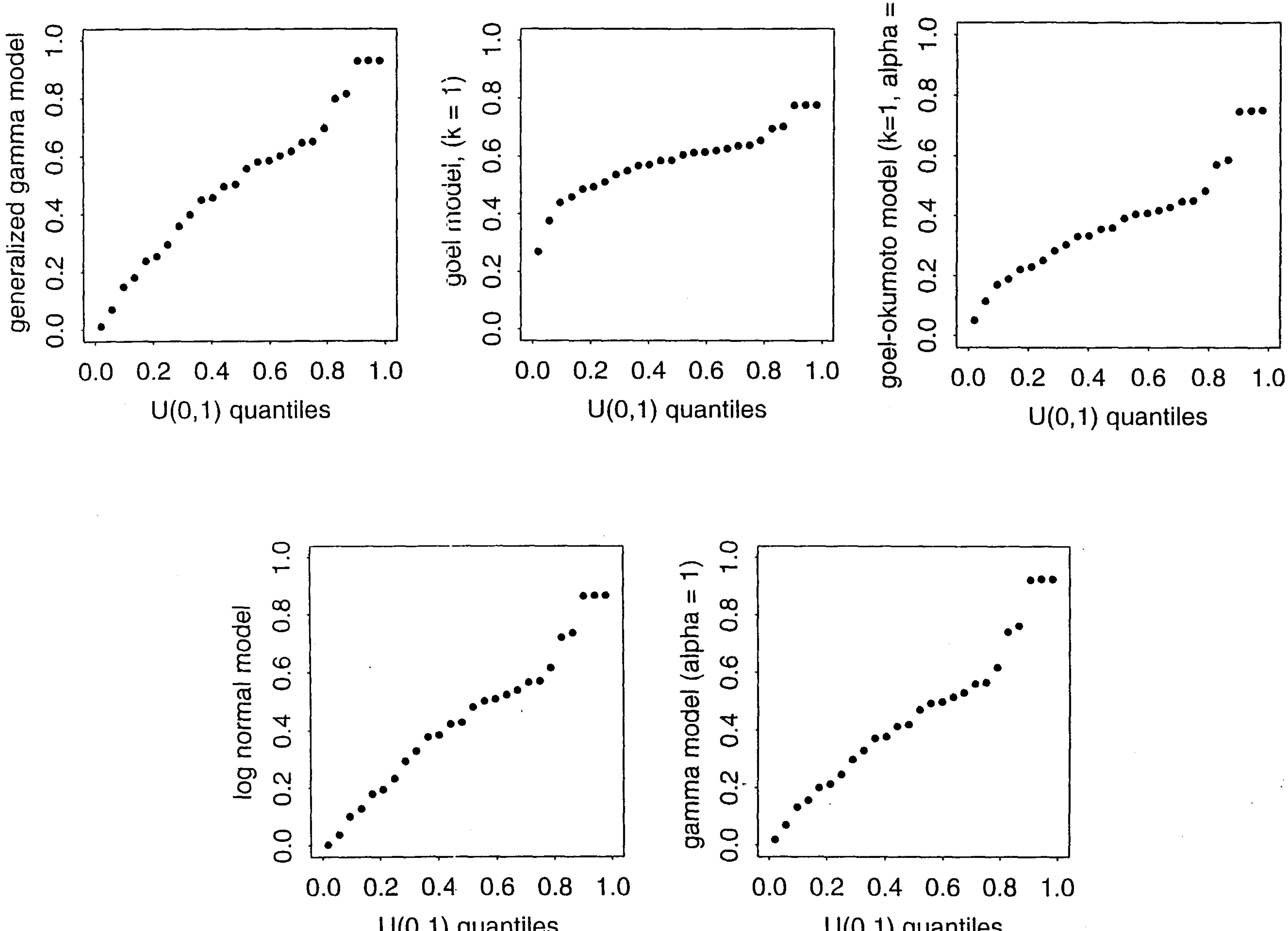


FIGURE 3. Empirical Q-Q plots for m(t)/theta vs uniform for 5 models.



NOTAS DO ICMSC

SÉRIE ESTATÍSTICA

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.

019/95 LEANDO, R.A.; ACHCAR, J.A.; - Generation of bivariate lifetime data assuming the Block & Basu exponential

018/95 ACHCAR, J.A.; - Use of approximate bayesian inference for software reliability

017/95 ACHCAR, J.A.; FAVORETTI, A.C. - Accurate inferences for the Michaelis-Menten model