REVSTAT – Statistical Journal
Volume 5, Number 2, June 2007, 209–226

INTERFAILURE DATA WITH CONSTANT HAZARD FUNCTION IN THE PRESENCE OF CHANGE-POINTS

Authors: Jorge Alberto Achcar
– Departamento de Estatística, Universidade Federal de São Carlos, São Carlos-SP, Brazil
jachcar@power.ufscar.br

Selene Loibel
– Departamento de Matemática Aplicada e Estatística, Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos, S.P., Brazil
sloibel@terra.com.br

Marinho G. Andrade
– Departamento de Matemática Aplicada e Estatística, Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo-USP, São Carlos, S.P., Brazil
marinho@icmc.usp.br

Received: October 2006 Revised: February 2007 Accepted: April 2007

Abstract:

- Markov Chain Monte Carlo (MCMC) methods are used to perform a Bayesian analysis for interfailure data with constant hazard function in the presence of one or more change-points. We also present some Bayesian criteria to discriminate different models. The methodology is illustrated with a data set originally reported in Maguire, Pearson and Wynn [8].

Key-Words:

- *constant hazard; change-points; Gibbs sampling; MCMC algorithms.*

AMS Subject Classification:

- 62J02, 62F10, 62F03.

1. INTRODUCTION

Applications of change-point models are given in many areas of interest. For example, medical researchers usually have interest to know if a new therapy of leukemia produces a departure from the usual experience of a constant relapse rate after the induction of a remission (see for example, Matthews and Farewell [9], Matthews *et al.* [10] or Henderson and Matthews [6]). Bayesian analysis for change-point models has been introduced by many authors. A Bayesian analysis for a homogeneous Poisson process with a change-point has been introduced by Raftery and Akman [11]. A Bayesian interval estimator has been derived for a change-point in a Poisson process by West and Ogden [15] and a Bayesian approach for lifetime data with a constant hazard function and censored data in the presence of a change point by Achcar and Bolfarine [1]. Recently Loschi and Cruz [7] presented a Bayesian approach to the multiple change point identification problem in Poisson data.

In this paper, we consider the presence of two or more change-point for lifetime with constant hazards, generalizing previous work (see for example, Achcar and Bolfarine [1]).

Consider a homogeneous Poisson process with one or more change-points at unknown times. With a single change-point, the rate of occurrence at time s is given by

$$\lambda(s) = \begin{cases} \lambda_1, & 0 \le s \le \tau\,, \\ \lambda_2, & s > \tau\,. \end{cases} \tag{1.1}$$

The analysis of the Poisson process is based on the counting data in the period $[0,T]$, where $N(T) = n$ is the number of events that occur at the ordered times $t_1, t_2, ..., t_n$.

With two change-points at unknown times τ_1 and τ_2 the rate of occurrences are given by

$$\lambda(s) = \begin{cases} \lambda_1, & 0 < s \le \tau_1\,, \\ \lambda_2, & \tau_1 < s \le \tau_2\,, \\ \lambda_3, & \tau_2 < s \le T\,. \end{cases} \tag{1.2}$$

We also could have homogeneous Poisson processes with more than two change-points.

The use of Bayesian methods has been considered by many authors for homogeneous or nonhomogeneous Poisson processes in the presence of one change-point (see for example, Raftery and Akman [11] or Ruggeri and Sivaganesan [13]).

Observe that times between failures for a homogeneous Poisson process follow an exponential distribution.

In this paper, we present a Bayesian analysis for interfailure data with constant hazard function assuming more than one change-point and using MCMC methods (see for example [4]).

The paper is organized as follows: in Section 2, we introduce the likelihood function; in Section 3, we introduce a Bayesian analysis for the model, in Section 4, we present some consideration on model selection; in Section 5, we introduce an example with real data and finally, in Section 6, we present some conclusions.

2. THE LIKELIHOOD FUNCTION

Let $x_i = t_i - t_{i-1}$, $i = 1, 2, ..., n$ where $t_0 = 0$, be the interfailure times and assume a single-change-point model (1.1). In this way, we observe that x_i has an exponential distribution with parameter λ_1 for $\sum_{k=1}^{i} x_k \leq \tau$ and an exponential distribution with parameter λ_2 for $\sum_{k=1}^{i} x_k > \tau$, $i = 1, 2, ..., n$. Assuming that the change-point τ is taking the values t_i, the likelihood function for λ_1, λ_2 and τ is given by

$$L(\lambda_1, \lambda_2, \tau) = \prod_{i=1}^{N(T)} (\lambda_1 e^{-\lambda_1 x_i})^{\epsilon_i} (\lambda_2 e^{-\lambda_2 x_i})^{1-\epsilon_i} \tag{2.1}$$

where $\epsilon_i = 1$ if $\sum_{j=1}^{i} x_j \leq \tau$ and $\epsilon_i = 0$ if $\sum_{j=1}^{i} x_j \geq \tau$. That is,

$$L(\lambda_1, \lambda_2, \tau) = \lambda_1^{N(\tau)} e^{-\lambda_1 \tau} \lambda_2^{N(T)-N(\tau)} e^{-\lambda_2 (T-\tau)} \tag{2.2}$$

where $N(\tau) = \sum_{i=1}^{N(T)} \epsilon_i$, $N(T) = n$, $\tau = \sum_{i=1}^{N(T)} x_i \epsilon_i$ and $T - \tau = \sum_{i=1}^{N(T)} x_i (1 - \epsilon_i)$.

Let us assume a two-change-point model (1.2) with the change-points τ_1 and τ_2 taking discrete values $\tau_1 = t_i$, $\tau_2 = t_j$ $(t_i < t_j,\ i \neq j)$ with $k_1 = N(\tau_1)$ and $k_2 = N(\tau_2)$. The likelihood function for λ_1, λ_2, λ_3, τ_1 and τ_2 is given by

$$L(\lambda_1, \lambda_2, \lambda_3, \tau_1, \tau_2) = \prod_{i=1}^{n} (\lambda_1 e^{-\lambda_1 x_i})^{\epsilon_{1,i}} (\lambda_2 e^{-\lambda_2 x_i})^{\epsilon_{2,i}} (\lambda_3 e^{-\lambda_3 x_i})^{\epsilon_{3,i}} \tag{2.3}$$

where

$$\epsilon_{1,i} = \begin{cases} 1 & \text{if } \sum_{k=1}^{i} x_k \leq \tau_1 , \\ 0 & \text{if } \sum_{k=1}^{i} x_k > \tau_1 , \end{cases} \tag{2.4}$$

$$\epsilon_{2,i} = \begin{cases} 1 & \text{if } \tau_1 < \sum_{k=i+1}^{j} x_k \leq \tau_2 , \\ 0 & \text{if } \sum_{k=i+1}^{j} x_k \leq \tau_1 \text{ or } \sum_{k=i+1}^{j} x_k > \tau_2 , \end{cases} \tag{2.5}$$

$$\epsilon_{3,i} = \begin{cases} 1 & \text{if } \tau_2 < \sum_{k=j+1}^{n} x_k , \\ 0 & \text{if } \tau_2 \geq \sum_{k=j+1}^{n} x_k . \end{cases} \tag{2.6}$$

That is,

$$(2.7) \qquad L(\lambda_1,\lambda_2,\lambda_3,\tau_1,\tau_2) = \\ = \lambda_1^{N(\tau_1)} e^{-\lambda_1\tau_1}\, \lambda_2^{N(\tau_2)-N(\tau_1)} e^{-\lambda_2(\tau_2-\tau_1)}\, \lambda_3^{N(T)-N(\tau_2)} e^{-\lambda_3(T-\tau_2)}$$

where $\sum_{i=1}^{N(T)} \epsilon_{1,i} = N(\tau_1)$, $\sum_{i=1}^{N(T)} \epsilon_{2,i} = N(\tau_2) - N(\tau_1)$, $\sum_{i=1}^{N(T)} \epsilon_{3,i} = N(T) - N(\tau_2)$ and $N(T) = n$. Observe that $\tau_1 = \sum_{i=1}^{N(T)n} x_i\, \epsilon_{1,i}$, $\tau_2 - \tau_1 = \sum_{i=1}^{N(T)n} x_i\, \epsilon_{2,i}$ and $T - \tau_2 = \sum_{i=1}^{N(T)n} x_i\, \epsilon_{3,i}$.

In the same way, we could generalize for more than two change-points.

3. A BAYESIAN ANALYSIS

Assume the change-point model (1.1) with a single change-point τ.

Assume that τ is independent from λ_1 and λ_2, and also that λ_1 is conditionally independent from λ_2, given $\tau = t_i$. Considering a noninformative prior distribution for λ_1 and λ_2 given τ (see for example, Box and Tiao [2]), we have

$$(3.1) \qquad \pi\big(\lambda_1,\lambda_2,\tau=t_i\big) = \pi\big(\lambda_1,\lambda_2\,|\,\tau=t_i\big)\,\pi\big(\tau=t_i\big) \propto \frac{1}{\lambda_1\lambda_2}\,\pi\big(\tau=t_i\big)$$

where $\lambda_1, \lambda_2 > 0$.

Assuming an uniform prior distribution $\pi_0(\tau = t_i) = 1/n$, the joint posterior distribution for λ_1, λ_2 and τ is given by

$$(3.2) \qquad \pi\big(\lambda_1,\lambda_2,\tau\,|\,\mathcal{D}\big) \propto \lambda_1^{N(\tau)-1} e^{-\lambda_1\tau}\, \lambda_2^{n-N(\tau)-1} e^{-\lambda_2(T-\tau)}$$

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

Observe that we are using a data dependent prior distribution for the discrete change-point (see for example Achcar and Bolfarine [1]). Also observe that the event $\{\tau = t_i\}$ is equivalent to $\{N(t_i) = i\}$, where the t_i are the ordered occurrence epochs of failures. We also could consider an informative gamma prior distribution for the parameters λ_1 and λ_2.

The marginal posterior distribution for τ is, from (3.2), given by

$$(3.3) \qquad \pi\big(\tau\,|\,\mathcal{D}\big) \propto \frac{\Gamma\big[N(\tau)\big]\,\Gamma\big[n-N(\tau)\big]}{\tau^{N(\tau)}\,(T-\tau)^{n-N(\tau)}} .$$

Assuming $\tau = \tau^*$ known, the marginal posterior distribution for λ_1 and λ_2 are given by

$$(3.4) \qquad \begin{aligned} &\text{(i)} \quad \lambda_1|\tau^*, \mathcal{D} \sim \text{Gamma}\big[N(\tau^*), \tau^*\big] , \\ &\text{(ii)} \quad \lambda_2|\tau^*, \mathcal{D} \sim \text{Gamma}\big[n-N(\tau^*), T-\tau^*\big] , \end{aligned}$$

where Gamma$[a,b]$ denotes a gamma distribution with mean a/b and variance a/b^2.

Assuming τ unknown, since the marginal posterior distribution for τ is obtained analytically (see(3.3)), we use a mixed Gibbs sampling and Metropolis–Hastings algorithm to generate the posterior distributions of λ_1 and λ_2. The conditional posterior distributions for the Gibbs sampling algorithm are given by

$$\text{(3.5)} \qquad \begin{aligned} &\text{(i)} \quad \lambda_1 | \lambda_2, \tau, \mathcal{D} \sim \text{Gamma}\big[N(\tau), \tau\big] , \\ &\text{(ii)} \quad \lambda_2 | \lambda_1, \tau, \mathcal{D} \sim \text{Gamma}\big[n - N(\tau), T - \tau\big] . \end{aligned}$$

Starting with initial values $\lambda_1^{(0)}$ and $\lambda_2^{(0)}$, we follow the steps:

(i) Generate $\tau^{(i)}$ from (3.3).

(ii) Generate $\lambda_1^{(i+1)}$ from $\pi\big(\lambda_1 | \lambda_2^{(i)}, \tau^{(i)}, \mathcal{D}\big)$.

(iii) Generate $\lambda_2^{(i+1)}$ from $\pi\big(\lambda_2 | \lambda_1^{(i+1)}, \tau^{(i)}, \mathcal{D}\big)$.

We could monitor the convergence of the Gibbs samples using Gelman and Rubin's method that uses the analysis of variance technique to determine whether further iterations are needed (see [5] for details).

A great simplification to get the posterior summaries of interest for the constant hazard function model in the presence of a change-point is to use the *WinBugs* software (see, Spiegelhalter *et al.* [14]) which requires only the specification of the distribution for the data and prior distributions for the parameters.

Consider now, the change-point model (1.2) with two change-points τ_1 and τ_2 (with $\tau_1 < \tau_2$). The prior density for λ_1, λ_2, λ_3, τ_1 and τ_2 is given by

$$\text{(3.6)} \qquad \pi\big(\lambda_1, \lambda_2, \lambda_3, \tau_1, \tau_2\big) = \\ = \pi\big(\lambda_1, \lambda_2, \lambda_3 \,|\, \tau_1 = t_i, \tau_2 = t_j\big)\; \pi_0\big(\tau_1 = t_i, \tau_2 = t_j\big)\, I_{\{t_i < t_j\}} ,$$

given $\tau_1 = t_i$, $\tau_2 = t_j$, $(t_i < t_j,\ i \neq j)$.

Assuming τ_1 and τ_2 independent from λ_1, λ_2 and λ_3, and also that λ_1, λ_2 and λ_3 are conditionally independent given τ_1 and τ_2, a noninformative joint prior distribution for λ_1, λ_2, λ_3 and τ_1 and τ_2 is given by

$$\text{(3.7)} \qquad \pi\big(\lambda_1, \lambda_2, \lambda_3, \tau_1 = t_i, \tau_2 = t_j\big) \propto \frac{1}{\lambda_1 \lambda_2 \lambda_3}\, \pi_0\big(\tau_1 = t_i, \tau_2 = t_j\big)\, \mathrm{I}_{\{t_i < t_j\}}$$

where $\lambda_1, \lambda_2, \lambda_3 > 0$, $\mathrm{I}_{\{t_i < t_j\}} = 1$ if $t_i < t_j$ and $\mathrm{I}_{\{t_i < t_j\}} = 0$ otherwise, for all $i \neq j$.

Assuming an uniform prior distribution for the discrete variables $\tau_1 = t_i$ and $\tau_2 = t_j$, where $t_i < t_j$, $i, j = 1, ..., n$, that is $\pi_0(\tau_1 = t_i, \tau_2 = t_j) = 2/n(n-1)$, the joint posterior distribution for λ_1, λ_2, λ_3, τ_1 and τ_2 is given by

$$\text{(3.8)} \qquad \pi\big(\lambda_1, \lambda_2, \lambda_3, \tau_1, \tau_2 \,|\, \mathcal{D}\big) \propto \\ \propto \lambda_1^{N(\tau_1)-1} e^{-\lambda_1 \tau_1}\, \lambda_2^{N(\tau_2)-N(\tau_1)-1} e^{-\lambda_2(\tau_2-\tau_1)}\, \lambda_3^{N(T)-N(\tau_2)-1} e^{-\lambda_3(T-\tau_2)}$$

where $\lambda_1, \lambda_2, \lambda_3 > 0$ and $\tau_1 < \tau_2$.

The joint marginal posterior distribution for τ_1 and τ_2 is given by

$$\pi\big(\tau_1,\tau_2\,|\,\mathcal{D}\big) \;=\; \frac{\Gamma\big[N(\tau_1)\big]\;\Gamma\big[N(\tau_2)-N(\tau_1)\big]}{\tau_1^{N(\tau_1)}(\tau_2-\tau_1)^{N(\tau_2)-N(\tau_1)}}\;\frac{\Gamma\big[N(\tau_2)-N(\tau_1)\big]}{(T-\tau_2)^{N(T)-N(\tau_2)}}\;. \tag{3.9}$$

We use the Metropolis–Hastings algorithm to generate τ_1,τ_2 from the joint marginal posterior distribution (3.9) and the Gibbs sampling algorithm to generate λ_1, λ_2 and λ_3. The conditional posterior distribution for the Gibbs sampling algorithm are given by

$$\lambda_1|\;\lambda_2,\lambda_3,\tau_1,\tau_2,\mathcal{D} \;\sim\; \text{Gamma}\big[N(\tau_1),\tau_1\big]\;, \tag{3.10}$$

$$\lambda_2|\;\lambda_1,\lambda_3,\tau_1,\tau_2,\mathcal{D} \;\sim\; \text{Gamma}\big[N(\tau_2)-N(\tau_1),\,\tau_2-\tau_1\big]\;, \tag{3.11}$$

$$\lambda_3|\;\lambda_1,\lambda_2,\tau_1,\tau_2,\mathcal{D} \;\sim\; \text{Gamma}\big[N(T)-N(\tau_2),\,T-\tau_2\big]\;. \tag{3.12}$$

This marginalization process should be made with careful choice of the lower and upper limits of summation as well as of the number of minimum points between τ_1 and τ_2. We consider $\tau_1=t_i$ for $i=1,...,m_1-1$, $\tau_2=t_i$ for $i=m_2+1,...,n$, where $\tau_1<\tau_2$ and $m_j\;(j=1,2)$ is a positive integer such that $t_{m_j}=\tau_j$. Note that once τ_1, (τ_2) is known, possible candidates of τ_1, (τ_2) are limited within $\{t_1,...,t_{m_1-1}\}$, $(\{t_{m_2+1},...,t_n\})$.

Starting with the initial values $\lambda_1^{(0)}$, $\lambda_2^{(0)}$ and $\lambda_3^{(0)}$, we follow the steps:

(i) Generate $\tau_1^{(i)}$ and $\tau_2^{(i)}$ from the marginal posterior distributions (3.9).
(ii) Generate $\lambda_1^{(i+1)}$ from $\pi\big(\lambda_1|\lambda_2^{(i)},\lambda_3^{(i)},\tau_1^{(i)},\tau_2^{(i)},\mathcal{D}\big)$.
(iii) Generate $\lambda_2^{(i+1)}$ from $\pi\big(\lambda_2|\lambda_1^{(i+1)},\lambda_3^{(i)},\tau_1^{(i)},\tau_2^{(i)},\mathcal{D}\big)$.
(iv) Generate $\lambda_3^{(i+1)}$ from $\pi\big(\lambda_3|\lambda_1^{(i+1)},\lambda_2^{(i+1)},\tau_1^{(i)},\tau_2^{(i)},\mathcal{D}\big)$.

Observe that the choices for m_1 and m_2 could have been made empirically based on a preliminary analysis of the data set (empirical Bayesian methods). In this way, we could use plots of the accumulated number of failures against time of occurrence to get some information on the change-point.

4. SOME CONSIDERATIONS ON MODEL SELECTION

For model selection, we could use the predictive density for the interfailure time x_i given $\underset{\sim}{x}_{(i)}=(x_1,...,x_{i-1},x_{i+1},...,x_n)$. The predictive density for x_i given $\underset{\sim}{x}_{(i)}$ is

$$c_i \;=\; f\big(x_i|\underset{\sim}{x}_{(i)}\big) \;=\; \int f\big(x_i|\underset{\sim}{\theta}\big)\;\pi\big(\underset{\sim}{\theta}\,|\underset{\sim}{x}_{(i)}\big)\;d\underset{\sim}{\theta} \tag{4.1}$$

where $\pi\big(\underset{\sim}{\theta}\,|\underset{\sim}{x}_{(i)}\big)$ is the posterior density for a vector of parameters $\underset{\sim}{\theta}$ given the data $\underset{\sim}{x}_{(i)}$.

Using the Gibbs samples, (4.1) can be approximated by its Monte Carlo estimates,

$$\widehat{f}\big(x_i \,|\, \underset{\sim}{x}_{(i)}\big) \;=\; \frac{1}{M} \sum_{j=1}^{M} f\big(x_i \,|\, \underset{\sim}{\theta}^{(j)}\big) \;, \tag{4.2}$$

where $\underset{\sim}{\theta}^{(j)}$ are the generated Gibbs samples, $j = 1, 2, ..., M$.

We can use $c_i = \widehat{f}\big(x_i \,|\, \underset{\sim}{x}_{(i)}\big)$ in model selection. In this way, we consider plots of c_i versus i $(i = 1, 2, ..., n)$ for different models; large values of c_i (in average) indicates a better model. We could also have choosen the model such that $P_l = \prod_{i=1}^{n} c_i(l)$ is maximum (l indexes models). We could also have considered (see Raftery [12]) the marginal likelihood of the whole data set $\mathcal{D}$ for a model M_l given by

$$P(\mathcal{D} \,|\, M_l) \;=\; \int_{\theta_l} L(\mathcal{D} \,|\, \theta_l, M_l)\; \pi(\theta_l \,|\, M_l)\; d\theta_l \tag{4.3}$$

where $\mathcal{D}$ is the data, M_l is the model specification (the number of change points), θ_l is the vector of the parameters in M_l, $L(\mathcal{D} \,|\, \theta, M_l)$ is the likelihood function and $\pi(\theta_l \,|\, M_l)$ is the prior.

The Bayes factor criterion prefers model M_1 to model M_2 if $P(\mathcal{D} \,|\, M_2) < P(\mathcal{D} \,|\, M_1)$. A Monte Carlo estimate for the marginal likelihood $P(\mathcal{D} \,|\, M_l)$ is given by

$$\widehat{P}(\mathcal{D} \,|\, M_l) \;=\; \frac{1}{M} \sum_{j=1}^{M} L\big(\mathcal{D} \,|\, \theta_l^{(j)}, M_l\big) \tag{4.4}$$

where $\theta_l^{(j)}$, $j = 1, 2, ..., M$, could have been generated through the use of importance sampling. The simplest estimator of this type results from taking the prior as the importance sampling function (see Raftery [12]).

Other ways to estimate the marginal likelihood $P(\mathcal{D} \,|\, M_l)$ are proposed by Raftery [12].

Considering a sample from the posterior distribution, we have

$$\widehat{P}(\mathcal{D} \,|\, M_l) \;=\; \left(\frac{1}{M} \sum_{j=1}^{M} \frac{1}{L\big(\mathcal{D} \,|\, \theta_l^{(j)}, M_l\big)} \right)^{-1} . \tag{4.5}$$

In this case, the importance-sampling function is the posterior distribution.

A modification of the harmonic mean estimator (4.5) is proposed by Gelfand and Dey [3], given by

$$\widehat{P}(\mathcal{D} \,|\, M_l) \;=\; \left(\frac{1}{M} \sum_{j=1}^{M} \frac{f\big(\theta_l^{(j)}\big)}{L\big(\mathcal{D} \,|\, \theta_l^{(j)}, M_l\big)\; \pi_0\big(\theta_l^{(j)}\big)} \right)^{-1} \tag{4.6}$$

where $f(\theta_l)$ is any probability density and $\pi_0(\theta_l)$ is a prior probability density.

5. AN EXAMPLE

In this section, we analyze a data set related to the number of mine accidents in England from 1875 to 1951. To analyze this data set, we have assumed the validity of a homogeneous Poisson process in the presence of change-points. Considering the time intervals between explosions in mines, we introduced a Bayesian analysis to get inference for the parameter of the exponential distributions and for the finite change-points.

In Table 1, we have the time intervals (in days) between explosions in mines, involving more than 10 men killed, from December 6, 1875 to May 29, 1951 (data introduced by Maguire, Pearson and Wynn [8]).

Table 1: Time intervals in days between explosions in mines.

378	36	15	31	215	11	137	4	15	72	96
124	50	120	203	176	55	93	59	315	59	61
1	13	189	345	20	81	286	114	108	188	233
28	22	61	78	99	326	275	54	217	113	32
23	151	361	312	354	58	275	78	17	1205	644
467	871	48	123	457	498	49	131	182	255	195
224	566	390	72	228	271	208	517	1613	54	326
1312	348	745	217	120	275	20	66	291	4	369
338	336	19	329	330	312	171	145	75	364	37
19	156	47	129	1630	29	217	7	18	1357	

From a plot of $N(t_i)$ versus t_i, $i = 1, 2, ..., 109$ (see Figure 1), we observe the presence of two or more change-points. We could also have assumed the presence of a random number of change-points (see for example, Ruggeri and Sivaganesan [13]) but this case is beyond the scope of this paper. As an illustration of the proposed model introduced in Section 1, we assume the presence of two change-points. Assuming the two change-points model (1.2) to analyze the data set of Table 1 and from Figure 1, we see that these two change-points are approximately $\widehat{\tau}_1 = t_{45} = 5231$ and $\widehat{\tau}_2 = t_{81} = 19053$. We also assume the presence of only one change-point and use Bayesian discrimination methods to decide for the best model.

In Figure 1, we also have empirical estimates for the rates λ_j, $j = 1, 2, 3$, obtained from the usual definition of the homogeneous Poisson processes $N(t) \propto \lambda t + o(n)$, where $N(t)$ is the accumulated number of occurrences in the interval $(0, t)$.

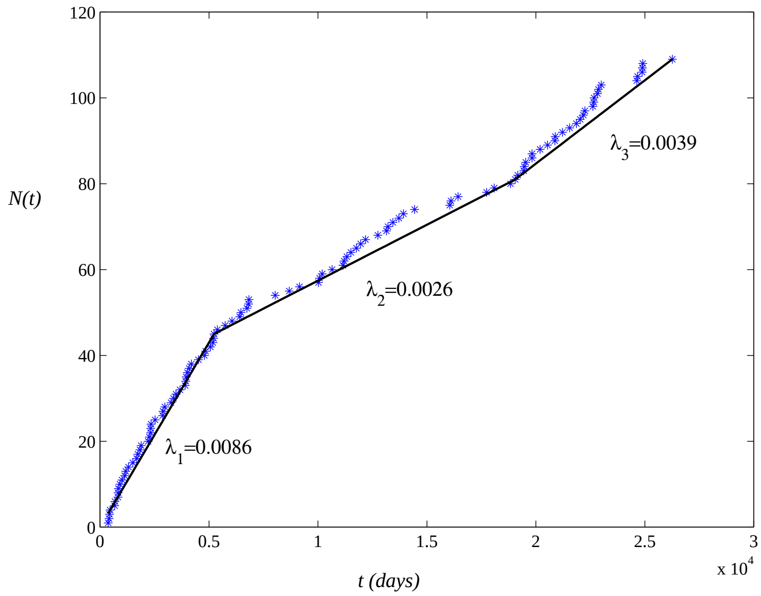

Figure 1: Plot of $N(t_i)$ versus $t_i(days)$.

If we assume the change-point model (1.1) with a single change-point τ with an uniform discrete prior, the mode of the marginal posterior distribution for τ (see (3.3)) is given by $\tau^* = 5382$ (see Figure 2). Assuming τ^* known, the mean of the marginal posterior distributions (3.4) are given by $\widetilde{\lambda}_1 = 0.008361$ and $\widetilde{\lambda}_2 = 0.003065$.

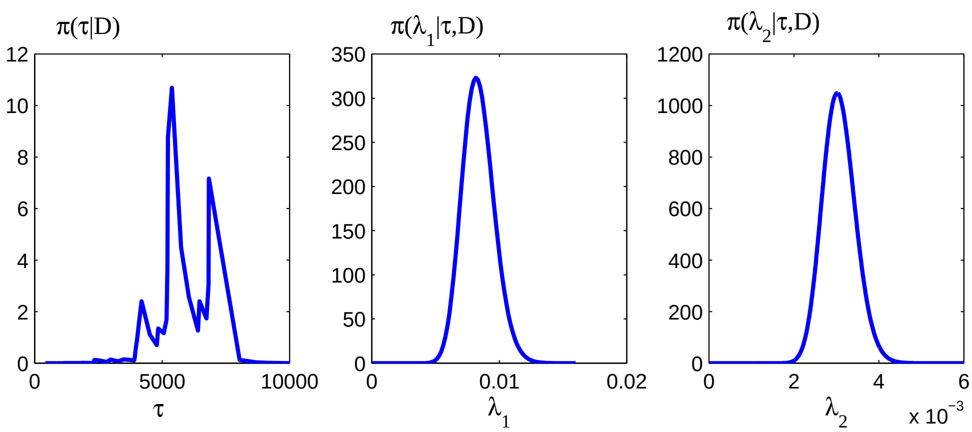

Figure 2: Marginal posterior distribution for τ and, λ_1 and λ_2 with $\tau = \tau^*$.

Assuming one or two unknown change-points, we have obtained posterior summaries (see Tables 2, 3, 4 and 5) through the use of MCMC algorithms.

In all cases, we have considered a "burn-in-sample" of size 5,000; after this, we have simulated 50,000 mixed Metropolis–Hastings and Gibbs samples taking every 10^{th} sample, to get approximated uncorrelated samples. The convergence of the mixed algorithms was monitored using graphical methods and standard existing indexes (see, for example, Gelman and Rubin [5]).

Considering the change-point model (1.1) with only one change-point τ, we have in Table 2, the posterior summaries for the parameters τ, λ_1 and λ_2 assuming the noninformative prior (3.1). In Figure 3, we have the approximate marginal posterior densities.

Table 2: Posterior summaries (change-point model 1.1).

	Mean	S.D.	95% Cred. Inter.
τ	5813	932	(4086 ; 7364)
λ_1	0.008059	0.001285	(0.005814 ; 0.010786)
λ_2	0.003047	4.011E-4	(0.002289 ; 0.003884)

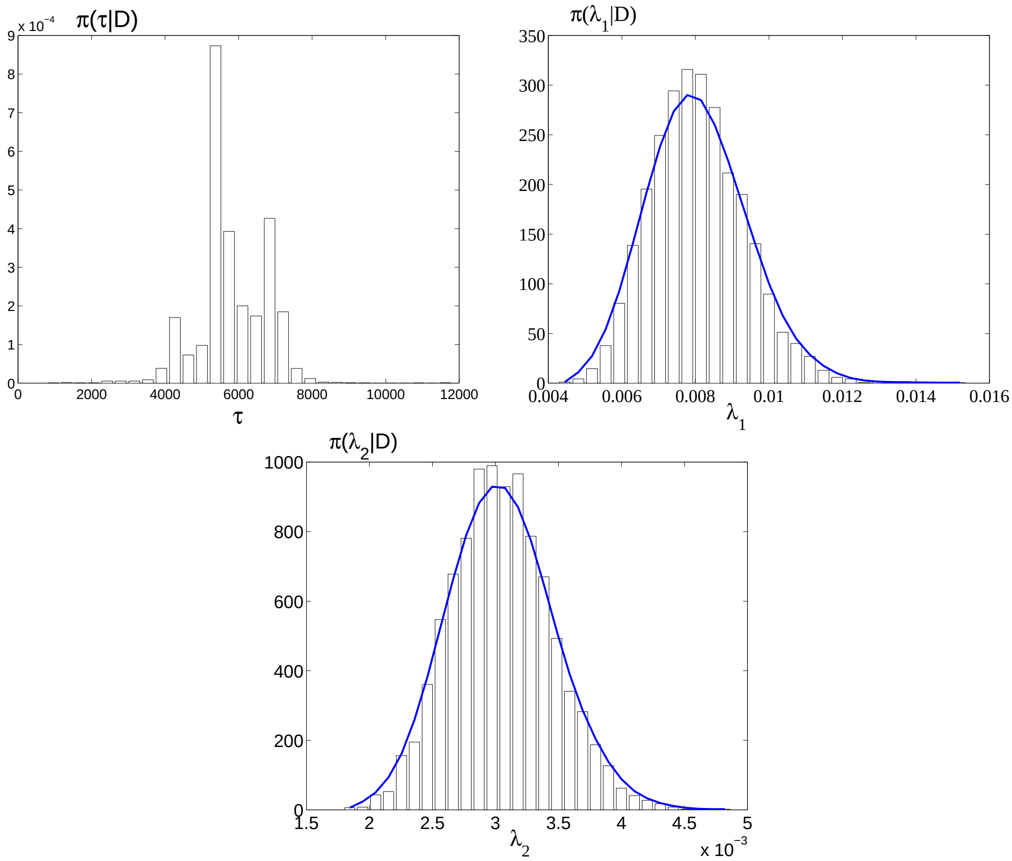

Figure 3: Marginal posterior distribution (change-point model 1.1).

Similar results could also have been obtained from the parametrization $k = N(t_k)$, λ_1 and λ_2. Assuming an uniform prior distribution for $N(t_i)$ taking the values $\{1, 2, ..., n\}$ and Gamma$(0.1, 0.1)$ prior distributions for λ_1 and λ_2, we obtain by Gibbs sampling algorithms the approximate marginal posterior densities for τ, λ_1 and λ_2. In Table 3 we have the posterior summaries of interest using the *WinBugs* software. The code of the *WinBugs* program is given in Appendix 1, assuming $k = N(t_k)$. Observe that $k \cong 46$ corresponds to $\tau = 5382$. That is, we have obtained results similar to the previous ones.

Table 3: Posterior summaries (gamma priors for λ_1 and λ_2).

	Mean	S.D.	95% Cred. Inter.
k	45.63	5.186	(35.0 ; 53.0)
λ_1	0.008322	0.001315	(0.006085 ; 0.01120)
λ_2	0.003056	3.975E-4	(0.002344 ; 0.003892)

In Figure 4, we have the approximated marginal posterior densities considering the 5,000 generated Gibbs samples.

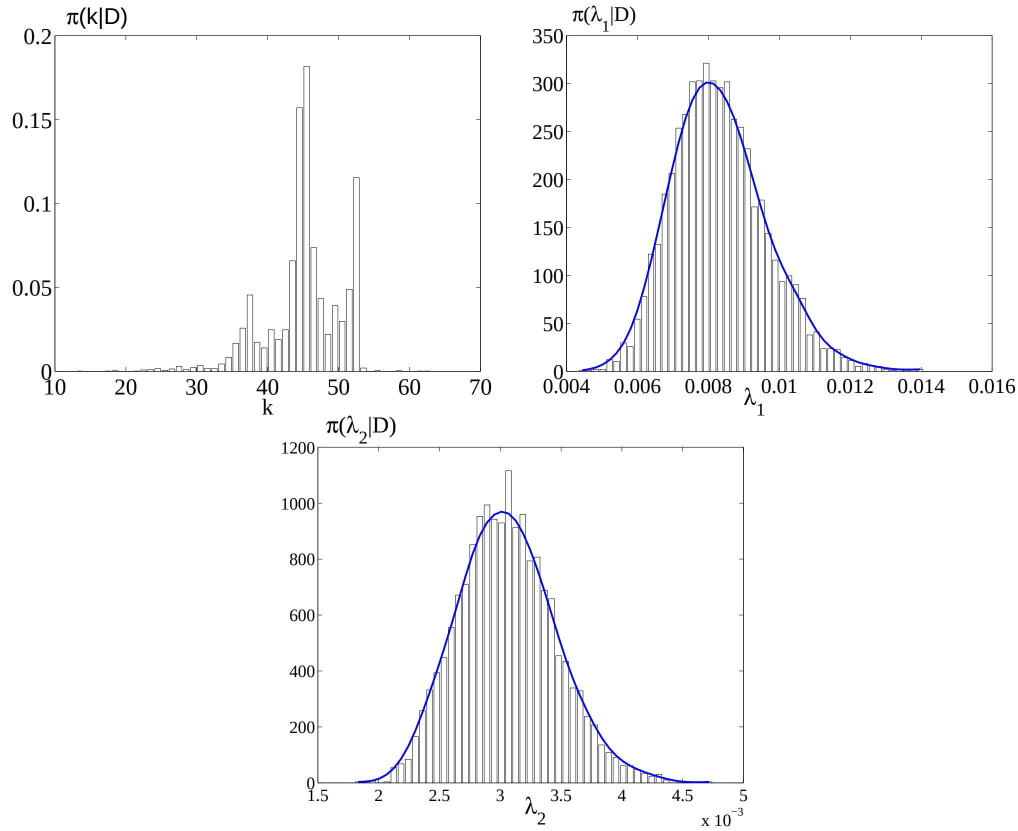

Figure 4: Marginal posterior distribution (gamma prior distribution for λ_1 and λ_2).

Assuming the two change-point model (1.2), we have in Table 4, the posterior summaries for the parameters λ_1, λ_2, λ_3, τ_1 and τ_2 obtained from the 5,000 generated Gibbs samples using the conditional posterior distributions (3.10)–(3.12). In Figure 5 we have the approximate marginal posterior densities.

Table 4: Posterior summaries (change-point model 1.2).

	Mean	S.D.	95% Cred. Inter.
τ_1	5990	876	(4176 ; 7354)
τ_2	17459	3162	(11287 ; 22741)
λ_1	0.008036	0.001262	(0.005765 ; 0.010703)
λ_2	0.002713	6.080E-4	(0.001655 ; 0.004053)
λ_3	0.003450	7.646E-4	(0.002103 ; 0.005082)

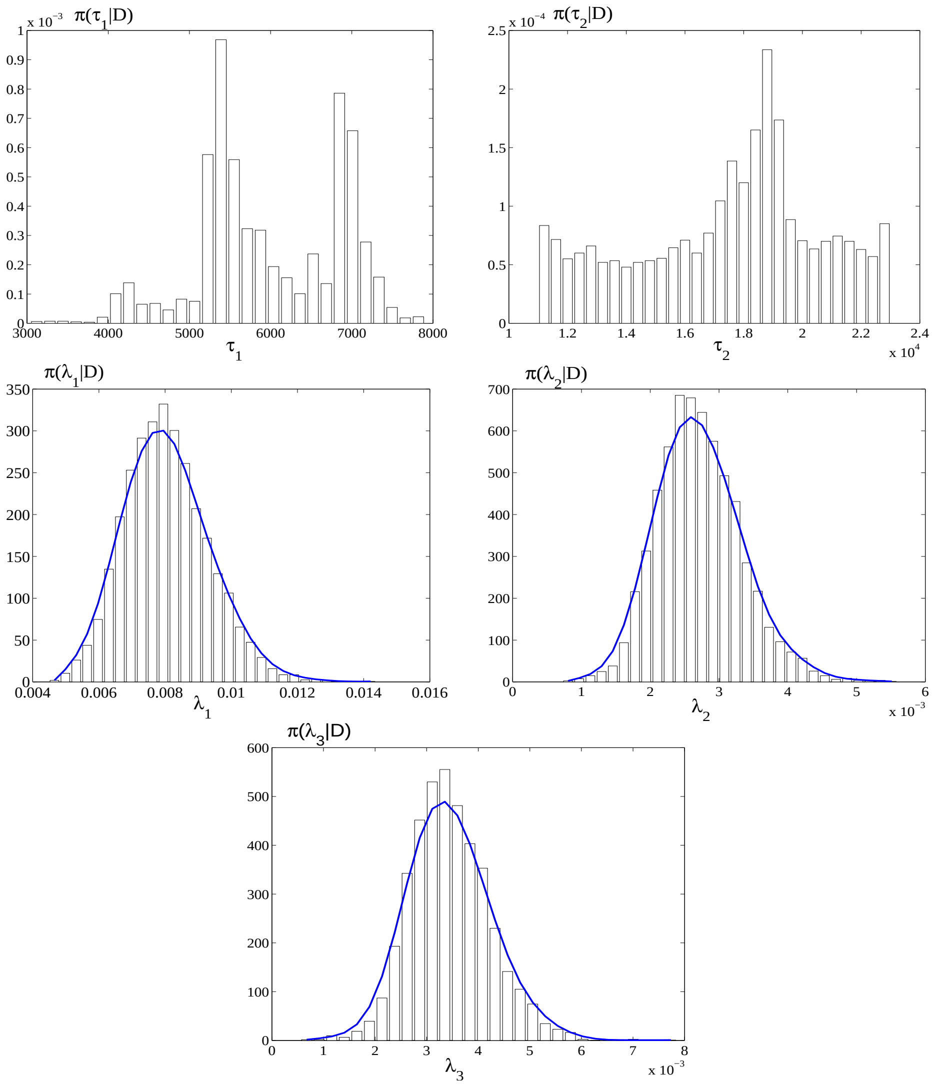

Figure 5: Marginal posterior distributions (change-point model 1.2).

Similar results have been obtained from the parametrization $k_1 = N(t_{k_1})$, $k_2 = N(t_{k_2})$, λ_1, λ_2 and λ_3. In Table 5, we have the posterior summaries of interest, obtained using the *WinBugs* software (code in Appendix 1), informative discrete prior distributions for the two change-points and independent Gamma(0.1, 0.1) prior distributions for λ_1, λ_2 and λ_3. Observe that $k_1 \cong 46$ corresponds to $\tau_1 = 5382$ and $k_2 \cong 78$ corresponds to $\tau_2 = 17743$. In Figure 6, we have the approximate marginal posterior distributions considering the 5,000 generated Gibbs samples.

Table 5: Posterior summaries (two change-point and gamma priors for λ_1, λ_2 and λ_3).

	Mean	S.D.	95% Cred. Inter.
k_1	46.22	4.237	(37.0 ; 53.0)
k_2	78.29	10.45	(58.0 ; 97.0)
λ_1	0.008349	0.001298	(0.006077 ; 0.01115)
λ_2	0.002780	6.378E-4	(0.001606 ; 0.004134)
λ_3	0.003445	7.392E-4	(0.002195 ; 0.005079)

In Figure 7, we have plots of the predictive densities $c_i = f(x_i | \underset{\sim}{x}_{(i)})$, $i = 1, 2, ..., n$, approximated by the Monte Carlo estimates (4.2) for both models M_1 (a single change-point model) and M_2 (two change-points model). For model M_1, we have $P_1 = \prod_{i=1}^n \widehat{c}_{1i} = 7.896 \times 10^{-303}$ and for model M_2 we have $P_2 = \prod_{i=1}^n \widehat{c}_{2i} = 9.5536 \times 10^{-302}$. The ratio of these values is given by $P_2/P_1 = 12.09$.

In Table 6, we have different estimates (see (4.5) and (4.6)) for the marginal likelihood functions considering models M_1 (single change-point model) and M_2 (two change-point model).

Table 6: Estimate values of the marginal likelihood.

Model	$P(\mathcal{D}\|M_l)$ using (4.5)	$P(\mathcal{D}\|M_l)$ using (4.6)
M_1	7.7716×10^{-305}	4.6420×10^{-304}
M_2	3.1256×10^{-304}	2.5020×10^{-302}

From Table 6, we calculate the Bayes factors $B_{ij} = P(\mathcal{D}|M_i)/P(\mathcal{D}|M_j)$, $i, j = 1, 2$. The Bayes factors are given by $B_{21} = 4.02$ (using (4.5)) and $B_{21} = 53.9$ (using (4.6)). If compared to one change-point model, we observe a better fit of the two change-point model M_2 for the data set of Table 1, considering the three model selection procedures.

It is important to point out that better models also could be considered to analyze the data set of the Table 1, considering more than two change-points.

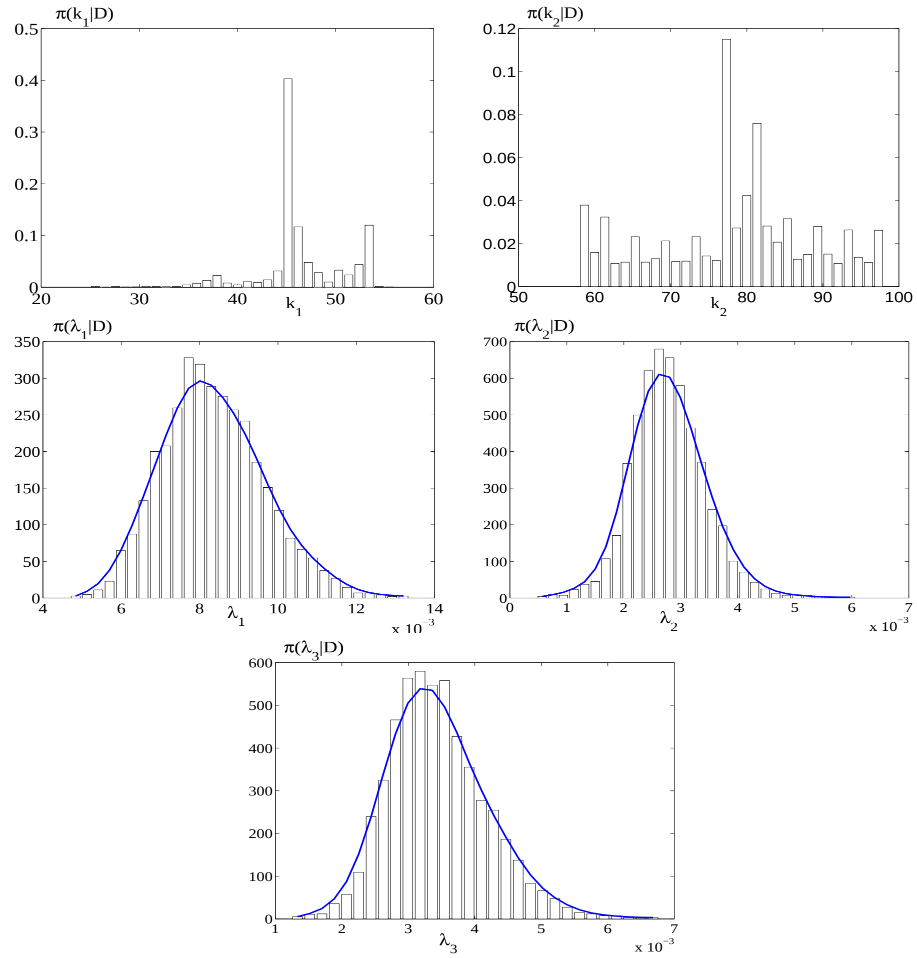

Figure 6: Marginal posterior distributions (gamma prior distributions for λ_1, λ_2 and λ_3 an informative discrete prior distribution for τ_1 and τ_2).

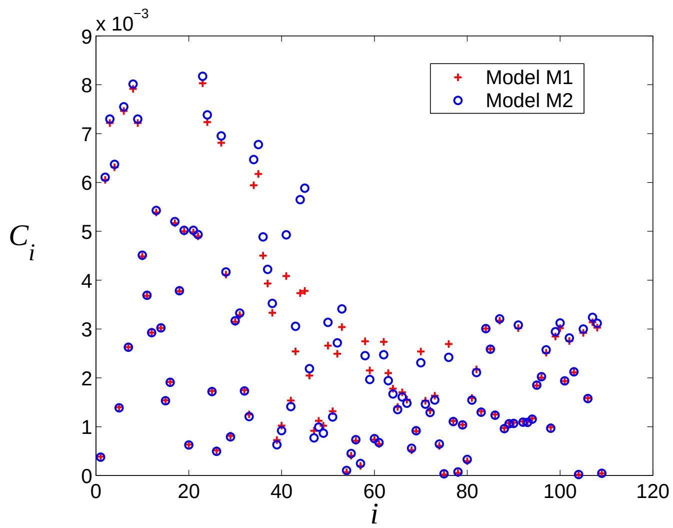

Figure 7: Plot of c_i versus i $(M_1\colon +, M_2\colon \circ)$.

6. CONCLUDING REMARKS

In this paper, we have observed that Bayesian inference for the parameters of change-point models is easily obtained through the use of Markov Chain Monte Carlo methods.

The use of recent software, such as *WinBugs*, to simulate samples for the joint posterior distribution of interest gives a great simplification in the computational work. It is important to point out that the usual classical inference procedures usually are not appropriate for change-point models (see for example, Mattews *et al.* [10]).

The proposed Bayesian methodology could also have been considered directly using the counting data modeled by homogeneous Poisson processes in the presence of one or more change-points in place of the inter-failure data (see for example, Raftery and Akman [11]).

Similar results could have been obtained for interfailure data with constant hazards and more than two change-points.

The use of Monte Carlo estimates for the predictive densities $f\big(x_i\,|\,\underset{\sim}{x}_{(i)}\big)$, $i = 1, 2, ..., n$, or for the marginal likelihood of the whole data set $\mathcal{D}$ for a model M_l, gives simple ways to discriminate the different change-point models, a problem of great practical interest.

APPENDIX

A. *WinBugs* code (one change-point)

```
Model
      {
      for(i in 1 : N) {
          t[i] ∽ dexp(lambda[J[i]])
          J[i]< −1+step(i−k−0.5)
          punif[i]<−1/N
                        }
      for(j in 1 : 2) {
          lambda[j]∽ dgamma(0.1, 0.1)
                        }
      k ∽ dcat(punif[ ])
      }
```

list(t=c(378, 36, 15, 31, 215, 11, 137, 4, 15, 72, 96, 124, 50, 120, 203, 176, 55, 93, 59, 315, 59, 61, 1, 13, 189, 345, 20, 81, 286, 114, 108, 188, 233, 28, 22, 61, 78, 99, 326, 275, 54, 217, 113, 32, 23, 151, 361, 312, 354, 58, 275, 78, 17, 1205, 644, 467, 871, 48, 123, 457, 498, 49, 131, 182, 255, 195, 224, 566, 390, 72, 228, 271, 208, 517, 1613, 54, 326, 1312, 348, 745, 217, 120, 275, 20, 66, 291, 4, 369, 338, 336, 19, 329, 330, 312, 171, 145, 75, 364, 37, 19, 156, 47, 129, 1630, 29, 217, 7, 18, 1357), N=109)

list(k=50, lambda=c(0.5, 0.5))

B. *WinBugs* code (two change-point)

```
Model
    {
    for(i in 1 : N) {
        t[i] ∽ dexp(lambda[J[i]])
        J[i]< −1+step(i−k1−0.5)+step(i−k2−0.5)
                }
    for(j in 1 : 3) {
        lambda[j]∽ dgamma(0.1, 0.1)
                }
    k1∽dcat(p1[ ])
    k2∽dcat(p2[ ])
    }
```

list(t=c(378, 36, 15, 31, 215, 11, 137, 4, 15, 72, 96, 124, 50, 120, 203, 176, 55, 93, 59, 315, 59, 61, 1, 13, 189, 345, 20, 81, 286, 114, 108, 188, 233, 28, 22, 61, 78, 99, 326, 275, 54, 217, 113, 32, 23, 151, 361, 312, 354, 58, 275, 78, 17, 1205, 644, 467, 871, 48, 123, 457, 498, 49, 131, 182, 255, 195, 224, 566, 390, 72, 228, 271, 208, 517, 1613, 54, 326, 1312, 348, 745, 217, 120, 275, 20, 66, 291, 4, 369, 338, 336, 19, 329, 330, 312, 171, 145, 75, 364, 37, 19, 156, 47, 129, 1630, 29, 217, 7, 18, 1357), p1=c(0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0.0225, 0.1, 0.0225, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0), p2=(0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0.0225, 0.10, 0.0225, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0), N=109)

list(k1=45,k2=78,lambda=c(0.5,0.5,0.5))

ACKNOWLEDGMENTS

The authors acknowledge helpful suggestions by the editor and the referees of this paper.

REFERENCES

[1] Achcar, J.A. and 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.

[2] Box, G.E. and Tiao, G.C. (1973). *Bayesian Inference in Statistical Analysis*, Addison-Wesley, New York.

[3] Gelfand, A.E. and Dey, D.K. (1994). Bayesian model choice: asymptotics and exact calculations, *Journal of the Royal Statistical Society, B*, **56**, 501–514.

[4] 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.

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

[6] Henderson, R. and Matthews, J.N.S. (1993). An investigation of change-points in the annual number of case of haemolytic uraemic sydrom, *Applied Statistics*, **42**, 461–471.

[7] Loschi, R.H. and Cruz, F.R.B. (2005). Byesian identification of multiple change points in Poisson data, *Advances in Complex Systems*, **8**, 4, 465–482.

[8] Maguire, B.A.; Pearson, E.S. and Wynn, A.H.A. (1952). The time intervals between industrial accidents, *Biometrika*, **39**, 168–180.

[9] Matthews, D.E. and Farewell, V.T. (1982). On testing for a constant hazard against a change-point alternative, *Biometrics*, **38**, 463–468.

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

[11] Raftery, A.E. and Akman, V.E. (1986). Bayesian analysis of a Poisson process with a change-point, *Biometrika*, **73**, 1, 85–89.

[12] Raftery, A.E. (1996). *Hypothesis testing and model selection*. In "Markov Chain Monte Carlo in Practice" (W.R. Gilks, S. Richardson and D.J. Spiegelhalter, Eds.), Chapman & Hall, London, 163–187.

[13] Ruggeri, F. and Sivaganesan, S. (2005). On modelling change-point in non-homogeneous Poisson process, *Statistical Inference for Stochastic Processes*, **8**, 311–329.

[14] Spiegelhalter, D.J.; Thomas, A. and Best, N.G. (1999). *WinBugs: Bayesian Inference Using Gibbs Sampling*, Cambridge: MCR Biostatistics Unit.

[15] West, R.W. and Ogden, R.T. (1997). Continous-time estimation of a change-point in a Poisson process, *Journal of Statistical Computational Simulation*, **56**, 293–302.