The Korean Communications
in Statistics Vol. 6, No. 3, 1999
pp. 801-811

Bayesian Analysis for Multiple Change-point Hazard Rate Models

Younshik Chung[1], Kwangmo Jeong[2] and Mihae Han[3]

Abstract

Change-point hazard rate models arise, for example, in applying "burn-in" techniques to screen defective items and in studing times until undesirable side effects occur in clinical trials. Sometimes, in screening defectives it might be sensible to model two stages of burn-in. In a clinical trial, there might be an initial hazard rate for a side effect which, after a period of time, changes to an intermediate hazard rate before settling into a long term hazard rate. In this paper, we consider the multiple change points hazard rate model. The classical approach's asymptotics can be poor for the small to all moderate sample sizes often encountered in practice. We propose a Bayesian approach, avoiding asymptotics, to provide more reliable inference conditional only upon the data actually observed. The Bayesian models can be fitted using simulation methods. Model comparison is made using recently developed Bayesian model selection criteria. The above methodology is applied to a generated data and the Lawless(1982) failure times of electrical insulation.

1. Introduction

In the reliability theory a widely accepted procedure is to apply "burn-in" techniques to screen out defective items and improve the lifetimes of surviving items. One helpful tool for capturing "burn-in" is to model the age process by the hazard function. See Cinlar(1975) for the age process. Let T denote the lifetime with density function $f(t)$ and survival function $\overline{F}(t)=P_r(T>t)$. We consider the case of more than one threshold if it were appropriate. For instance, in screening defectives it might be sensible to model two stages of burn-in. In a clinical trial, there might be an initial hazard rate for a side effect which, after a period of time, changes to an intermediate hazard rate before settling into a long term hazard rate. An

1) Associate Prof., Research Institutes of Information and Communication, Department of Statistics, Pusan National University, Pusan, 609-735 Korea
2) Professor, Research Institutes of Information and Communication, Department of Statistics, Pusan National University, Pusan, 609-735 Korea
3) Lecturer, Department of Statistics, Pusan National University, Pusan, 609-735 Korea

illustrative hazard function $h(t)=\frac{f(t)}{F(t)}$ would take the form

$$h(t)=\theta_1 I(0<t\leq\tau_1)+\theta_2 I(\tau_1<t\leq\tau_2)+\theta_3 I(t>\tau_2) \tag{1.1}$$

where

$$I(A)=\begin{cases}1 & \text{if } x\in A\\ 0 & otherwise.\end{cases}$$

More specifically, suppose the lifetimes of items are independent, but that first failures appear to occur at one rate and second failure (after some threshold time) appear to occur at another rate and so on. That is, it is assumed to be k change points in the given period. Then a suitable form for the hazard function $h(t)$ is

$$h(t) = h_1(t)I(0\leq t\leq\tau_1)+h_2(t)I(\tau_1<t\leq\tau_2)+\cdots+h_k(t)I(\tau_{k-1}<t\leq\tau_k)+h_{k+1}(t)I(t>\tau_k), \tag{1.2}$$

where $\tau=(\tau_1,\cdots,\tau_k)$ is the threshold parameter vector and $\tau\in(R^+)^k$. From(1.2), the cumulative hazard function is, for $\tau_l\leq t<\tau_{l+1}$,

$$\begin{aligned} H(t) &= \int_0^t h(x)dx \\ &= H_1(\min(t,\tau_1))+\sum_{i=1}^{l}[H_i(\tau_i)-H_i(\tau_{i-1})]^+ \\ &\quad +[H_{l+1}(t)-H_{l+1}(\tau_l)]^+ \end{aligned} \tag{1.3}$$

where $H_i(t)$ is the cumulative hazard function on $\tau_{i-1}<t\leq\tau_i$, set $H_1(\tau_1)-H_1(\tau_0)=0$ and

$$[a]^+=\begin{cases}a & \text{if } a>0\\ 0 & otherwise.\end{cases}$$

In practice neither τ nor h_i's will be known. The goal of this paper is to consider inference in this general case.

The literature to date focuses primarily on examination of this problem from a classical perspective. For instance, Nguyen et al. (1984) consider the case where $h_1(t)=\lambda_1$ and $h_2(t)=\lambda_2$ for $k=2$ in our case. Basu et al.(1988) extend this to allow a general h_1 keeping $h_2(t)=\lambda_2\leq h_1(t)$ .See also Ebrahimi(1991) and Loader(1991) in this regard. In all of this work estimates for τ and λ_1 are proposed and their asymptotic properties are examined. Such asymptotics can be poor for the small to moderate sample size often encountered in practical reliability situations. We adopt a fully Bayesian approach for this problem, avoiding asymptotics to provide more reliable inference conditional only upon the data actually observed. Indeed, for each feature of the model we obtain an entire posterior distribution enabling any desired inference about the feature. In order to deflect customary criticism of the Bayesian approach to subjectivity in the prior specification, we adopt rather vague priors so that our inference resembles that of a likelihood analysis.

The remainder of the paper is organized as follows. Section 2 introduce the Bayesian model with a sample of lifetime collected under (1.2). We discuss prior specifications for $\tau_1, \cdots, \tau_k$ and fitting using simulation methods. In Section 3, we discuss the issues of model selection using Bayes factor. In Section 4, we specialize to the case of $h_1(t)=\theta_1$, $h_2(t)=\theta_2$ and $h_3(t)=\theta_3$ which yields particularly simple fitting. Finally, we analyze a generated data and a real data set from Lawless(1982).

2. Bayesian formulation in change-point hazard rate model

Returning to (1,1), we adopt a semi-parametric modelling assuming $h_i(t)=h_i(t\ \Theta_i)$ for $i=1,\cdots,k+1$. That is, h_i's are, possibly distinct, parametric families of hazard functions indexed by Θ_i's.

For an uncensored sample $T_1, \cdots, T_n$ of lifetime, the likelihood takes the form

$$\begin{aligned} L_k(\theta_1,\cdots,\theta_{k+1},\tau_1,\cdots,\tau_k;t) &= \prod_{j=1}^{n} h(t_j)\exp(-H(t_j)) \\ &= \prod_{j=1}^{n}\prod_{l=1}^{k+1} h_l(t_j;\Theta_l)^{I(\tau_l < t_j < \tau_{l+1})} \\ &\times \exp\{-\sum_{j=1}^{n}(H_1(\min(t_j,\tau_1);\Theta_1)+\sum_{i=1}^{l}[H_i(\tau_i;\Theta_i)-H_i(\tau_{i-1};\Theta_i)]^+ \\ &\quad +[H_{l+1}(t;\Theta_{l+1})-H_{l+1}(\tau_l;\Theta_{l+1})]^+)\} \end{aligned} \tag{2.1}$$

where $t=(t_1,\cdots,t_n)$ denotes the observed values of the lifetimes.
For notational convenience, let

$$h(t: r_1,\cdots,r_k)=\prod_{i=1}^{r_1-1} h_1(t_{(i)}:\theta_1)\cdots\prod_{i=r_{k-1}}^{r_k-1} h_k(t_{(i)};\theta_k)\prod_{i=r_k}^{n} h_{k+1}(t_{(i)}:\theta_{k+1})$$

and

$$H(t: r_1,\cdots,r_k)=\sum_{j=1}^{r_1-1} H_1(t_{(j)}:\theta_1)+\cdots+\sum_{j=r_{k-1}}^{r_k-1} H_k(t_{(j)}:\theta_k)+\sum_{j=r_k}^{n} H_{k+1}(t_{(j)}:\theta_{k+1})$$

Then (2.1) becomes

$$\begin{aligned} & L_k(\theta_1,\cdots,\theta_{k+1},\tau_1,\cdots,\tau_k;t) \\ =\ & h(t: r_1,\cdots,r_k)\times\exp\{-H(t: r_1,\cdots,r_k) \\ & -\sum_{l=1}^{k}(k-l+1)[(r_l-r_{l-1})H_l(\tau_l:\theta_l)-(r_{l+1}-r_l)H_{l+1}(\tau_l:\theta_{l+1})]\} \end{aligned} \tag{2.2}$$

with $r_0=1$ and $r_{k+1}=n-1$.

For our model we need to uniquely define the notion of "no change". That is, $\tau_k < t_j,\ \forall j$, is not distinguishable from $\tau_1 \ge t_j,\ \forall j$; in either case there is no change-point during the period of observation. As a simple remedy, if we order the observation times, $t_{(1)} < t_{(2)} < \cdots < t_{(n)}$, we restrict the likelihood so that $\tau_1 \ge t_{(1)}$. Certainly of k change points during the period of observation would then add the further restriction, $\tau_1 < t_{(n-k+1)}$. Since there are k change-points during the period, the conditions such as $t_{(l)} < \tau_l < t_{(n-k+l)}$ for $l = 2, \cdots, k$ are added to the likelihood function.

To create a Bayesian model, we require a prior specification for $\Theta_1, \Theta_2, \cdots, \Theta_{k+1}$ and $\tau_1, \cdots, \tau_k$ We assume that it takes the general form

$$f(\Theta_1, \cdots, \Theta_{k+1}) \times f(\tau_1, \cdots, \tau_k). \tag{2.3}$$

and that it is proper to assure that the posterior $f(\Theta_1, \cdots, \Theta_{k+1}, \tau_1, \cdots, \tau_k | t)$ is proper. Our prior information on $\tau_1, \cdots, \tau_k$ places it on the interval $(0, b)$ with b possibly ∞. The actual support for τ_i is truncated according to the restrictions imposed by the likelihood.

Combining (2.2) and (2.3) provides the complete Bayesian specification and thus the posterior $f(\Theta_1, \cdots, \Theta_{k+1}, \tau_1, \cdots, \tau_k | t)$ which is proportional to

$$L_k(\Theta_1, \cdots, \Theta_{k+1}, \tau_1, \cdots, \tau_k; t) \times f(\Theta_1, \cdots, \Theta_{k+1}) \times f(\tau_1, \cdots, \tau_k). \tag{2.4}$$

In fact for each t, since $h_i(t; \Theta_i)$'s and $H_i(t; \Theta_i)$'s are all random variables, they all have posterior distributions which would be of interest as well. However, primary interest is in the posterior for τ, $f(\tau_i | t)$ and when a change is not certain, $\Pr(\tau_k > t_{(n)} | t)$.

The expression in (2.3) is not analytically tractable so we turn to simulation approaches for fitting such a model. That is, we seek to draw samples from the posterior in (2.3) in order to learn about its features. Customary iterative approaches using, e.g. a rejection method or weighted bootstrap, see Smith and Gelfand (1992), are difficult here. However, the form in (2.3) is nicely suited for Markov chain Monte carlo simulation using Gibbs sampling updating (Gelfand and Smith, 1990). To clarify we consider the full conditional distributions for Θ_i's and for τ_i.

In particular, we immediately have $f(\Theta_1 | \Theta_{(-1)}, \tau, t)$ proportional to

$$\prod_{j=1}^{r_1-1} h_1(t_{(j)}; \Theta_1) \times \exp\{-\sum_{j=1}^{r_1-1} H_1(t_{(j)}; \Theta_1)\} \times f(\Theta_1, \cdots, \Theta_{k+1}). \tag{2.5}$$

For $l = 2, \cdots, k+1$, the full conditional distribution of Θ_l is given by

$$\begin{aligned} \prod_{j=r_{l-1}}^{r_l-1} h_l(t_{(j)}; \Theta_l) \times f(\Theta_1, \cdots, \Theta_{k+1}) \times \exp\{-\sum_{j=r_{l-1}}^{r_l-1} H_l(t_{(j)}; \Theta_l) \\ - (r_l - r_{l-1})[(k-l+1)H_l(\tau_l; \Theta_l) - (k-l+2)H_l(\tau_{l-1}; \Theta_l)]\} \end{aligned} \tag{2.6}$$

As for τ_i's. suppose $t_{(r_1-1)} < \tau_1 < t_{(r_1)}$, $r_1 = 2, \cdots, n-k+1$. Then , on this interval the full conditional distribution is given as

$$\begin{aligned} &[\tau_1 | \tau_{(-1)}, \theta, t] \\ &\propto K_{r_1} \exp\{-k[(r_1-1)H_1(\tau_1 : \theta_1) - (r_2 - r_1)H_2(\tau_1 : \theta_2)]\} \times f(\tau) \end{aligned} \tag{2.7}$$

where

$$K_{r_i} = h(t: r_1, \cdots, r_k) \times \exp\{-H(t: r_1, \cdots, r_k) - \sum_{l=2}^{k}(k-l+1)[(r_l - r_{l-1})H_l(\tau_l: \theta_l) - (r_{l+1} - r_l)H_{l+1}(\tau_l: \theta_{l+1})]\} \tag{2.8}$$

for $i=1,\cdots,k$ In general, on the interval $t_{(r_i-1)} < \tau_i < t_{(r_i)}$, $r_i = i+1,\cdots,n-k+i$, the full conditional distribution of τ_i given $\tau_{(-i)}$, θ and t is given by

$$[\tau_i | \tau_{(-i)}, \theta, t] \propto K_{r_i} \exp\{-(k-i+1)[(r_i - r_{i-1})H_i(\tau_i: \theta_i) - (r_{i+1} - r_i)H_{i+1}(\tau_i: \theta_{i+1})]\} \times f(\tau) \tag{2.9}$$

where

$$K_{r_i} = h(t: r_1, \cdots, r_k)\exp\{-H(t: r_1, \cdots, r_k) - \sum_{l \neq i}^{k}(k-l+1)[(r_l - r_{l-1})H_l(\tau_l: \theta_l) - (r_{l+1} - r_l)H_{l+1}(\tau_l: \theta_{l+1})]\} \tag{2.10}$$

Hence, let

$$C_{r_i} = K_{r_i}\int_{t_{(r-1)}}^{t_{(r)}} f(\tau)\exp\{-(k-i+1)[(r_i - r_{i-1})H_i(\tau_i: \theta_i) - (r_{i+1} - r_i)H_{i+1}(\tau_i: \theta_{i+1})]\}d\tau \tag{2.11}$$

and $C_r = \sum_{r_i=i+1}^{n-k+i} C_{r_i}$ and $p_{r_i} = \frac{C_{r_i}}{C_r}$ .To sample τ_i we first randomly select an interval, i.e., we select $t_{(r_i-1)} < \tau_i < t_{(r_i)}$ with probability p_{r_i}. We then draw τ_i in this interval from the normalized version of (2.9). That is, from the density which is (2.9) devided by C_r. Note that K_{r_i} cancels out of this density. In special cases the integration for C_r can be done explicitly in which case the density over $t_{(r_i-1)} < \tau_i < t_{(r_i)}$ is routine to sample. Finally, for the prior on $\tau = (\tau_1, \cdots, \tau_k)$, we illustrate with the case of

$$f(\tau_1, \cdots, \tau_k) = \prod_{i=1}^{k} \frac{1}{t_{(n-k+i)} - t_{(i)}} I(t_{(i)} < \tau_i < t_{(n-k+i)}) . \tag{2.12}$$

We have described how to handle a general change-point hazard rate problem using a Bayesian approach in the absence of censoring. The case where some of the lifetimes are censored can also be conveniently handled. Suppose, for instance, that for item i, the actual lifetime is not observed. Rather, the item is removed from test after time w_i at which point it has not failed so, $T_i > w_i$. All we need to do is introduce T_i as latent variable into our setup in (2.1) and (2.2). The form in (2.2) is unchanged except that T_i is now unknown and restricted to be greater than w_i. (2.8) now yields a posterior for $\Theta_1, \cdots, \Theta_{k+1}, \tau_1, \cdots, \tau_k$ and T_i given $t_1, \cdots, t_{i-1}, w_i, t_{i+1}, \cdots, t_n$. Again adopting Gibbs sampling to fit the model, the full conditional distributions for Θ_j now given $T_i = t_i$ are exactly as in (2.5) and (2.6). Also given $T_i = t_i$ the full conditional distributions for τ_j is as in (2.9) and (2.10). Lastly, we need to sample T_i given $\Theta_1, \cdots, \Theta_{k+1}, \tau_1, \cdots, \tau_k$ with $T_i > w_i$, i.e. we need to sample T_i from the distribution with hazard (1.2) subject to the restriction. If we draw U_i which is an exponential(1) random variable and invert

$$U_i = H_1(\min(t_i, \tau_1)) + \sum_{i=1}^{l}[H_i(\tau_i) - H_i(\tau_{i-1})]^+ + [H_{l+1}(t_i) - H_{l+1}(\tau_l)]^+ \qquad (2.13)$$

to solve for t_i, then t_i is a realization from the unconstrained distribution. Hence, we retain t_i if $t_i > w_i$. If not we draw a new U_i.

3. Bayesian Model Selection

In this section, we try to find the number of change points in given data. For example, we consider testing M_0: one change v.s. M_1: two changes. That is, the model with one change and the model with two changes are compared using Bayes factor. In general, suppose that we are interested in comparing two models M_0 and M_1. The formal Bayesian model choice procedure goes as follows. Let w_i be the prior probability of M_i, $i=0,1$and let $f(y|M_i)$ be the predictive distribution for model M_i, i.e.

$$[m]_i = f(y|M_i) = \int f(y|\theta_i, M_i)\,\pi(\theta_i|M_i)\,d\theta_i.$$

If y is the observed data, then we choose the model yielding the larger $w_i f(y|M_i)$. Often we set $w_i = \frac{1}{2}$ and compute the Bayes factor (or M_0 with respect to M_1)

$$BF = \frac{f(y|M_0)}{f(y|M_1)} = \frac{[m]_0}{[m]_1}. \qquad (3.1)$$

Jeffreys(1961) and Kass and Raftery(1995) suggest interpretive ranges for the Bayes factor and in general, M_0 is supported if $BF > 1$.

More generally, we want to estimate $[m] = \int f(y|\beta)\pi(\beta)d\beta$ using the importance sampling method. Let us consider $\pi(\beta|y)$ as the importance sampling function. Then the Markov Chain Monte Carlo methods, particularly Metropolis algorithm and Gibbs sampler, are used to get the sample from the posterior density $\pi(\beta|y)$. Let $\{\beta^{(g)}\}_{g=1}^{G}$ be Gibbs outputs as above. Then by Monte Carlo method, the approximating marginal density of Y is $[\widehat{m}] = \frac{\sum_{g=1}^{G} w_g f(y|\beta^{(g)})}{\sum_{g=1}^{G} w_g}$ where $w_g = \frac{\pi(\beta^{(g)})}{\pi(\beta^{(g)}|y)}$. Since $\pi(\beta|y) = \frac{f(y|\beta)\pi(\beta)}{[m]}$, the approximation can be expressed as

$$[\widehat{m}] = \left[\frac{1}{G}\sum_{g=1}^{G}\frac{1}{f(y|\beta^{(g)})}\right]^{-1}. \qquad (3.2)$$

Also, this final form is mentioned in Kass and Raftery(1995).

For example, suppose that we want to test one change (M_0) v.s. two change (M_1) , then the approximating Bayes factor for favoring M_0 is given by

$$BF = \frac{[\frac{1}{G}\sum_{g=1}^{G}\{L_1(y|\beta_1^{(g)}\}^{-1}]^{-1}}{[\frac{1}{G}\sum_{g=1}^{G}\{L_2(y|\beta_2^{(g)}\}^{-1}]^{-1}} \tag{3.3}$$

where $L_1(\cdot)$ and $L_2(\cdot)$ denote the likelihood function under M_0 and M_1 in (2.1) or (2.2), respectively. $\beta_1^{(g)} = (\theta_1^{(g)}, \theta_2^{(g)}, \tau^{(g)})$ and $\beta_2^{(g)} = (\theta_1^{(g)}, \theta_2^{(g)}, \theta_3^{(g)}, \tau_1^{(g)}, \tau_2^{(g)})$ where $\beta_1^{(g)}, \beta_2^{(g)}$ are the Gibbs output under the models M_0 and M_1 , respectively.

4. Application

4.1. Constant hazard rate model

We consider the case where $h_1(t) = \theta_1$, $h_2(t) = \theta_2$, $h_3(t) = \theta_3$. That is,

$$h(t) = \theta_1 I(0 < t \le \tau_1) + \theta_2 I(\tau_1 < t \le \tau_2) + \theta_3 I(t > \tau_2). \tag{4.1}$$

Classical analysis for this case using asymptotic results is present in Ngyuyen et al.(1984) and Loader(1991). For an observed sample of values $t_1, \cdots, t_n$ (2.1) becomes

$$\begin{aligned} &L_2(\theta_1, \theta_2, \theta_3, \tau_1, \tau_2, ; t) \\ &= \theta_1^{\sum I(t_j \le \tau_1)} \theta_2^{\sum I(\tau_1 < t_j \le \tau_2)} \theta_3^{\sum I(t > \tau_2)} \times \exp\{-\theta_1 \sum_{j=1}^{r_1-1} t_{(j)} - \theta_2 \sum_{j=r_1}^{r_2-1} t_{(j)} - \theta_3 \sum_{j=r_2}^{n} t_{(j)} \\ &\quad -2((r_1-1)\theta_1\tau_1 - (r_2-r_1)\theta_2\tau_1) - ((r_2-r_1)\theta_2\tau_2 - (r_3-r_2)\theta_3\tau_2)\} \end{aligned} \tag{4.2}$$

For the prior on θ_1, θ_2 and θ_3, we take independent gamma specifications restricted to $\{\theta_3 < \theta_2 < \theta_1\}$, i.e., $f(\theta_1, \theta_2, \theta_3)$ is propotional to

$$\theta_1^{a_1-1}\theta_2^{a_2-1}\theta_3^{a_3-1}\exp\{-b_1\theta_1 - b_2\theta_2 - b_3\theta_3\} I(\theta_3 < \theta_2 < \theta_1) \tag{4.3}$$

Also for the prior on $\tau = (\tau_1, \tau_2)$, we illustrate with the case of

$$f(\tau_1, \tau_2) = \prod_{i=1}^{2} \frac{1}{t_{(n-2+i)} - t_{(i)}} I(t_{(i)} < \tau_i < t_{(n-2+i)}). \tag{4.4}$$

In applications we choose the a_i and b_i so that the prior provide little information. For convenient notation, $Gam(a, b)$ denotes the gamma distribution with two parameters a and b. Then the full conditional distributions of θ_i's given $\theta_{(-i)}, \tau_1, \tau_2, t$ are as follows;

$$[\theta_1 | \theta_2, \theta_3, \tau_1, \tau_2, t] = Gam(a_1 + r_1, (\sum_{j=1}^{r_1-1} t_{(j)} + 2(r_1 - 1)\tau_1 + b_1)^{-1}) \cdot I(\theta_2 < \theta_1 < \infty) \tag{4.5}$$

$$[\theta_2 | \theta_1, \theta_3, \tau_1, \tau_2, t] = Gam(a_2 + r_2 - r_1, (\sum_{j=r_1}^{r_2-1} t_{(j)} + (r_2 - r_1)(\tau_2 - 2\tau_1) + b_2)^{-1}) \cdot I(\theta_3 < \theta_2 < \theta_1) \tag{4.6}$$

and

$$[\theta_3 | \theta_1, \theta_2, \tau_1, \tau_2, t] = Gam(a_3 + n - r_2, (\sum_{j=r_2}^{n} t_{(j)} - (n - 1 - r_2)\tau_2 + b_3)^{-1}) \cdot I(0 < \theta_3 < \theta_2) \tag{4.7}$$

Then (2.7) becomes

$$[\tau_1 | \tau_2, \theta_1, \theta_2, \theta_3, t] \propto K_{r_1} \exp\{-2((r_1 - 1)\theta_1 - (r_2 - r_1)\theta_2)\tau_1\} \times \frac{I(t_{(1)} < \tau_1 < t_{(n-1)})}{t_{(n-1)} - t_{(1)}} \tag{4.8}$$

where

$$K_{r_1} = \theta_1^{r_1-1}\theta_2^{r_2-r_1}\theta_3^{n-r_2+1} \exp\{-\theta_1 \sum_{j=1}^{r_1-1} t_{(j)} - \theta_2 \sum_{j=r_1}^{r_2-1} t_{(j)} - \theta_3 \sum_{j=r_2}^{n} t_{(j)} - ((r_2 - r_1)\theta_2 - (r_3 - r_2)\theta_3)\tau_2\} \tag{4.9}$$

and

$$[\tau_2 | \tau_1, \theta_1, \theta_2, \theta_3, t] \propto K_{r_2} \times \exp\{-((r_2 - r_1)\theta_2 - (r_3 - r_2)\theta_3)\tau_2\} \frac{I(t_{(2)} < \tau < t_{(n-2)})}{t_{(n-2)} - t_{(2)}} \tag{4.10}$$

where

$$K_{r_2} = \theta_1^{r_1-1}\theta_2^{r_2-r_1}\theta_3^{n-r_2+1} \exp\{-\theta_1 \sum_{j=1}^{r_1-1} t_{(j)} - \theta_2 \sum_{j=r_1}^{r_2-1} t_{(j)} - \theta_3 \sum_{j=r_2}^{n} t_{(j)} - 2((r_2 - 1)\theta_1 - (r_2 - r_1)\theta_2)\tau_1\}. \tag{4.11}$$

Obviously, C_r hence p_r, can be computed explicitly so choosing an interval at random for τ_i is easily done. Then we can draw τ_i within this interval by simple cumulative distribution function.

4.2. Simulated data

In this subsection, we try to find the number of change point in a simulated data set. We assume the model (4.1) with two change points and sample size $n = 15$ and generate a data set given in Table 4.1 according to the following simulation design.

i) Set θ_1=3.0 , θ_2=2.0 and θ_3=1.0 as the values of scale parameters.

ii) Fix the location of change point, τ_1 and τ_2 such that $P(T \le \tau_1)$=0.3 and $P(T \le \tau_2)$=0.7. Therefore, we can find $\tau_1 = 11.8892$ and $\tau_2 = 54.2541$.

Table 4.1. Simulated data

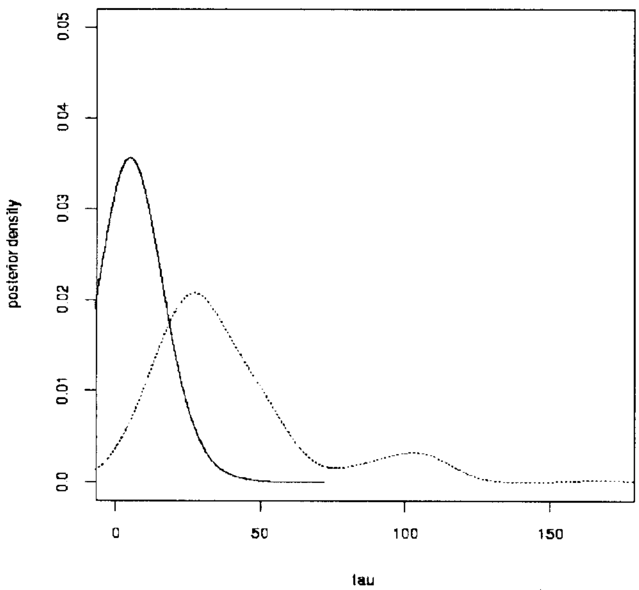
110.825	12.708	0.103	55.313	575.825	42.315	530.743	172.104
0.008	20.910	12.799	0.080	0.094	35.134	0.108	

First of all, we consider testing M_0: one change point against M_1: two change points in the model for the model selection. That is, the likelihood functions under M_0 is given by

$$L_1(\theta_1, \theta_2, \tau_1; t) = \theta_1^{\sum I(t_j \le \tau_1)} \theta_2^{\sum I(t > \tau_1)} \times \exp\{-\theta_1 \sum_{j=1}^{r_1 - 1} t_{(j)} - \theta_2 \sum_{j=r_1}^{n} t_{(j)} - ((r_1 - 1)\theta_1 \tau_1 - (r_2 - r_1)\theta_2 \tau_1)\} \tag{4.12}$$

and the likelihood function for M_1, $L_2(\theta_1, \theta_2, \theta_3, \tau_1, \tau_2; t)$ is given in (4.2). For diffuse prior of θ's, a's ans \$b\$'s are set to zero. Gibbs sampler was run for 5,000 iterations and the first 3,000 being discard as a burn-in period. Convergence of the Gibbs sampler was assessed via Geweke(1992) method, using the CODA(Best, Cowles and Vines, 1995) suitable of diagnostics in S-plus. The number of parameters is 5, such as $\theta_1, \theta_2, \theta_3, \tau_1$ and τ_2. Most of the parameters had Geweke statistics between -1.96 and 1.96, indicating convergence is plausible.

Figure 4.1. Posteriors of τ_1, τ_2



Thus, we compute the approximate Bayes factor(or M_0 with respect to M_1) in (3.3) as follows;

$$BF = \frac{\widehat{m}[M_0]}{\widehat{m}[M_1]} = 0.23779 \tag{4.13}$$

Since $BF < 1$ in (4.14), it is reasonable the model M_1 with two change points is well fitted. Therefore, we need only the posterior characteristics under M_1. Table 4.2 indicates the posterior means and 95 percent interval estimate for M_1 which has the two change points model. In Table 4.2, the first and second change point, τ_1 and τ_2, are estimated to 6.82 and 49.66, respectively. In Figure 4.1, the solid line means the marginal posterior density of τ_1 and the dotted line for τ_2 for simulated data.

Table 4.2 Inference summary with two change points

	posterior mean	posterior interval estimate
θ_1	2.7303	(1.2254, 4.8966)
θ_2	1.3459	(0.5628, 2.5250)
θ_3	0.5314	(0.2191, 0.9266)
τ_1	5.8241	(0.1028, 20.4497)
τ_2	39.6645	(12.7183, 109.3623)

4.3. A Real data set

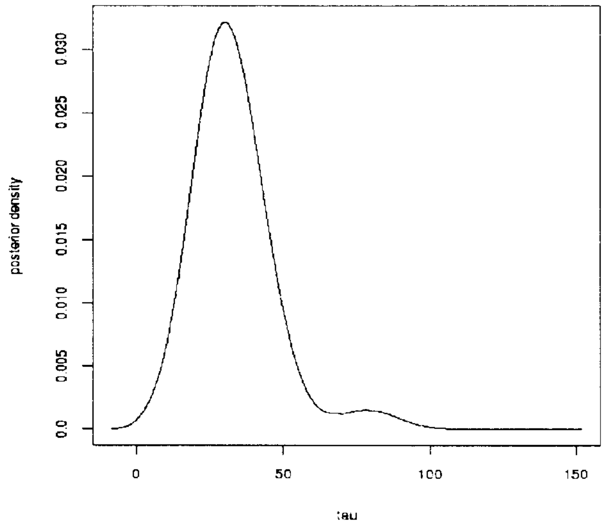
We have data which represent failure times in hours for electrical insulation in which the insulation was subjected to a continuously increasing voltage stresses.

Table 4.3. Electrical insulation data

219.3	79.4	86	150.2	21.7	18.5	121.9	40.5	147.1	35.1	42.3	48.7

Lawless(1982) fits a simple exponential model to these data. For diffuse prior of θ's, a's ans b's are set to zero. Here, Gibbs sampler is applied. Also, Geweke statistics is used for checking the convergence of Gibbs outputs. We consider the constant hazard rate model and compare M_0 : one change-point with M_1 : two change-points. Since the approximated Bayes factor in (3.3) is obtained as $2.289E+3$, the constant hazard model with one change point is adequate. For the one change-point model, the estimated value $\widehat{\tau_1}$ is 33.320. Figure 4.2 indicates the marginal posterior density of τ_1.

Figure 4.2. Posteriors of τ_1



References

[1] Basu, A.P., Ghosh, J.K. and Joshi, S.N.(1988). On estimating change point in a failure rate. *Statistics and Decision Theory*, 4, 239-252.

[2] Best, N.G., Cowles, M.K. and Vines, S.K.(1995) *Convergence Diagnosis and Output Analysis Software for Gibbs Sampling Output, Version 0.3*, Cambridge, MRC Biostatistics Unit.

[3] Cinlar, E.(1975). *Introduction to Stochastic Process*. Prentice-Hall, Englewood Cliffs, New Jergy.

[4] Ebrahimi, N.(1991). On estimating change-point in a mean residual life function. *Sankhya A*, 53, 209-219.

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

[6] Geweke, J.(1992) Evaluating the Accuracy of Sampling-Based Approaches to Calculating Posterior Moments, *In Bayesian Statistics 4*, ed. J.M. Bernado, J.O. Berger, A.P.Dawid and A.F.M. Smith, Oxford, UK; Oxford university Press, 169-193.

[7] Kass, R.E. and Raftery, A.E.(1995). Bayes factors. *Journal of American Statistical Association*, 90, 773-395.

[8] Lawless, J.F.(1982). *Satistical models and methods for lifetime data*. John Wiley, New York.

[9] Loader, C.R.(1991). Inference for a hazard rate change point. *Biometrika*, 78, 749-757.

[10] Nguyen, H.T., Rogers, G.S. and Walker, E.A.(1984). Estimation in change point hazard rate models. *Biometrika*, 71, 229-304.