Bayesian Changepoints Detection for the Power Law Process with Binary Segmentation Procedures[1]

Hyunsoo Kim[2], Seong W. Kim[3], and Hakjin Jang[4]

Abstract

We consider the power law process which is assumed to have multiple changepoints. We propose a binary segmentation procedure for locating all existing changepoints. We select one model between the no-changepoints model and the single changepoint model by the Bayes factor. We repeat this procedure until no more changepoints are found. Then we carry out a multiple test based on the Bayes factor through the intrinsic priors of Berger and Pericchi (1996) to investigate the system behaviour of failure times. We demonstrate our procedure with a real dataset and some simulated datasets.

Keywords : Binary segmentation, Changepoint, Model selection, Intrinsic prior, Power law
process

1. Introduction

The counting process is often used in modeling a repairable system. There are several different types of counting processes such as the homogeneous Poisson process (HPP), the renewal process, and the nonhomogeneous Poisson process (NHPP). The HPP and the renewal process are characterized by the times between failures. The failure times are independent and identically distributed with an exponential distribution for the HPP and an arbitrary distribution for the renewal process respectively. Upon failure it must be assumed that every repair restores the system to the same condition as new one.

In the NHPP, the interarrival times are neither independent nor identically distributed.

---

2) Research Professor, School of Information and Communications Engineering, Sungkyunkwan
University, 440-746, Suwon, Korea.
E-mail: bayes1@hanmail.net
3) Assistant Professor, Division of Applied Mathematics, Hanyang University, Ansan 426-791, Korea.
4) Ph. D Student, Division of Applied Mathematics, Hanyang University, Ansan 426-791, Korea.

According to Ascher and Feingold (1984), among the class of NHPP models, the power law process is the most commonly discussed in the literature. Duane (1964) analyzed data available for several repairable systems (including complex hydro-mechanical devices, complex aircraft generators, and a complete aircraft jet engine) in an effort to determine if any systematic changes occurred during development of these systems. Duane's analysis revealed that the plots of the cumulative failure rate as a function of cumulative operating hours are approximately linear on log-log paper. Crow (1974) expanded Duane's results by formulating a model for the reliability growth of repairable systems. The power law process (PLP) has the following form of the mean value function,

$$m(t) = \frac{1}{\eta}t^\beta, \eta > 0, \beta > 0.$$

Let $X_{(0,T]} = (x_1, \cdots, x_n)$ be the first n failure times of the PLP with $0 < x_1 \leq \cdots \leq x_n < T$. We assume that there exist K unknown changepoints. Further, we assume that for each K+1 subdivision the data follow the PLP with the same scale parameter η and different shape parameters $\beta_i, i = 1, \cdots, K+1$. It is reasonable to assume the same scale parameter in each subdivision since only the shape parameter determines the pattern of system failures. Our primary interest is to detect existing changepoints. In Bayesian perspectives, the Bayes factor is quite feasible as a model selection tool. However, computation could be a major factor due to huge number of parameters. To circumvent this difficulty, we employ a binary segmentation procedure, which are extensively used by several authors including Vostrikova (1981), Chen and Gupta (1997), and Yang and Kuo (2001).

The procedure can be described as follows: first, we compare the models between no changepoints and a single changepoint in $(0, T]$ using the Bayes factor. If one selects the no-changepoints model, stop the procedure, and move on to the next interesting feature. That is, we perform a multiple test about the shape parameter to see the behavior of failure times. If not, then we estimate the changepoint denoted by $\hat{\tau}$. Then we divide the data into two parts: one is denoted by $X_{(0,\hat{\tau}]}$, and the other contains all the event times between $\hat{\tau}$ and T inclusive. We now run two Bayes factor tests similar to what we have done above; one is based on $X_{(0,\hat{\tau}]}$ and the other is based on $X_{(0,\hat{\tau}]}$. Using data splitting in this procedure, we continue testing until no more changepoints are found in all of further subdivisions. Again, we perform multiple tests on the shape parameter in all subdivisions. In this procedure we only need to compare the no-changepoints model with the single-changepoint model. Furthermore, when we determine there is no single changepoint in a subsegment, we do not need to continue testing for the data in that subsegment. This dwindle down the sample size down quite significantly for locating changepoints in the remaining regions. The procedure is efficient and quite easy to implement.

A considerable amount of literature has been found in changepoint problems. In the case of Bayesian statistical analysis for a single changepoint in the Poisson process, there are some

work such as Akman and Raftery (1986), Carlin, Gelfand, and Smith (1992), and Raftery (1994). In particular, Akman and Raftery (1986) consider asymptotic inference for a changepoint Poisson process when it cannot be assumed that the changepoint occurs at an event time. Green (1995) proposes a reversible jump Markov chain Monte Carlo (MCMC) algorithm for computation in estimating multiple changepoints.

This article is presented as follows. In Section 2, we introduce the single changepoint model in the PLP and review the intrinsic Bayes factor and the intrinsic prior methodologies. In Section 3, we compute the Bayes factor, estimate the changepoint, and derive a general set of intrinsic priors. In Section 4, we present numerical results from real data and simulated datasets. We finish this article with a brief discussion in Section 5.

2. Preliminaries

2.1 The Changepoint Model of the NHPP

Consider a repairable system that is put into operation at time $t=0$. When the system has failed, it will be replaced or restored. The repair time is assumed to be so short that it may be neglected. A random variable of special interest is $N(t)$, the number of failures in the time interval (0,t]. The intensity function of a counting process $\{ N(t),\ t \geq 0 \}$ is defined as

$$\nu(t) = m'(t) = \frac{d}{dt} E[N(t)],$$

where $m(t)$ denotes the mean number of failures in the interval $(0,t]$, often called the mean value function. The power law process has the following form of the intensity function,

$$\nu(t) = \frac{\beta}{\eta} t^{\beta-1}, \quad \eta > 0,\ \beta > 0,\ t > 0. \tag{1}$$

Consider two models, the no-changepoints model (M_0) and the single changepoint model (M_1). For the data $X_{(0,T]}$ of failure times, the intensity functions are given respectively by

$$M_0 : \nu(t) = \frac{\beta_0}{\eta_0} t^{\beta_0-1}, \quad 0 < t \leq T,$$

and

$$M_1 : \nu(t) = \begin{cases} \dfrac{\beta_1}{\eta_1} t^{\beta_1-1} & ,0 < t \leq \tau, \\ \dfrac{\beta_2}{\eta_1} t^{\beta_2-1} & ,\tau < t \leq T, \end{cases}$$

where τ is a changepoint.

2.2 The Intrinsic Bayes Factor and the Intrinsic Prior

Suppose that we wish to compare q models for the given data X,

$$M_i: X \sim f_i(x|\Theta_i), i = 1, \cdots, q.$$

Then the Bayes factor is defined as

$$B_{ji}(x) = \frac{m_j(x)}{m_i(x)} = \frac{\int_{\Theta_j} f_j(x \mid \theta_j)\pi_j^N(\theta_j)d\theta_j}{\int_{\Theta_i} f_i(x \mid \theta_i)\pi_i^*(\theta_i)d\theta_i}, \tag{2}$$

where $\pi_i^N(\theta_i)$ is the noninformative prior, Θ_i is the parameter space for θ_i, and $m_i(x)$ is the marginal or predictive density of X under M_i. Now we can select the most plausible model using the posterior probabilities

$$P(M_i \mid X) = \left[\sum_{j=1}^{q}\frac{p(M_j)}{p(M_i)}B_{ji}\right]^{-1}, i = 1, \cdots, q,$$

where $p(M_i)$ is called the *prior* probability that model M_i is true.

However, noninformative priors are often improper. This makes the Bayes factor in (2) not to be well defined. Berger and Pericchi (1996) proposed the arithmetic intrinsic Bayes factor (AIBF), which is given by

$$B_{ji}^{AI}(x) = B_{ji}(x) \cdot CFA_{ij}(x), \tag{3}$$

where the correction factor is CFA_{ij} is

$$CFA_{ij} = \frac{1}{L}\sum_{l=1}^{L} B_{ij}(x(l)). \tag{4}$$

Here, $x(l)$ is called the minimal training sample, in the sense that the marginal density is finite for all possible models and L is the total number of training samples. In this setting, the size of the minimal training sample is same as the number of unknown parameters. Further, we notice that the ordinary Bayes factor requires proper priors such as conjugate priors. However, this Bayes factor could not be robust in perspective of the change of hyperparameters.

Note that the AIBF in (3) sometimes requires heavy computation. In particular, when the number of parameters is large, computation of the IBF (AIBF) becomes a bit heavier. However, if one can find a set of (hopefully) proper priors, one just needs to compute the usual Bayes factor defined in (2). Berger and Pericchi (1996) suggest a set of intrinsic priors denoted by (π_1^I, π_2^I) which is the solution of the following system of equations:

$$\begin{cases} \dfrac{\pi_2^I(\phi_2(\theta_1))\pi_1^N(\theta_1)}{\pi_2^N(\phi_2(\theta_1))\pi_1^I(\theta_1)} = B_1^*(\theta_1), \\[3mm] \dfrac{\pi_2^I(\theta_2)\pi_1^N(\phi_1(\theta_2))}{\pi_2^N(\theta_2)\pi_1^I(\phi_1(\theta_2))} = B_2^*(\theta_2), \end{cases} \tag{5}$$

where for $i=1,2$,

$$B_i^*(\theta_i) = \lim_{n \to \infty} CFA \text{ under } M_i,$$

and for $i \neq j$,

$$\phi_i(\theta_j) = \lim_{n \to \infty} E_{\theta_j}^{M_j}(\hat{\theta}_i) \quad \text{under } M_j,$$

with $\hat{\theta}_i$ being the MLE under M_i. We note that solutions are not necessarily unique nor proper.

3. Main Results

3.1 No-changepoint vs. A Single Changepoint

The likelihood function under M_0 is

$$f(x \mid \eta_0, \beta_0) = (\frac{\beta_0}{\eta_0})^n \left[\prod_{i=1}^n x_i \right]^{\beta_0 - 1} \exp\left\{ -T^{\beta_0}/\eta_0 \right\}. \tag{6}$$

And the likelihood function under M_1 is

$$f(x \mid \eta_1, \beta_1, \beta_2) = (\frac{1}{\eta_1})^n \beta_1^j \left[\prod_{i=1}^j x_i \right]^{\beta_1 - 1} \exp\left\{ -\tau^{\beta_1}/\eta_1 \right\}$$

$$\cdot \beta_2^{n-j} \left[\prod_{i=j+1}^n x_i \right]^{\beta_2 - 1} \exp\left\{ -(T^{\beta_2} - \tau^{\beta_2})/\eta_1 \right\}.$$

We note that the distribution of the likelihood ratio statistic cannot be obtained in analytic form for small sample sizes. Furthermore, the MLE of τ does not satisfy regularity conditions required to apply standard asymptotic likelihood ratio theory. Thus, we use the Bayes factor for comparing the no-change model (M_0) with the single change model (M_1).

Given the model M_0, we assume that $\eta_0 \sim Inverse\ Gamma(\xi, \nu)$ and the prior density for β_0 is $p_0(\beta_0) = 2/(1+\beta_0)^3$. (How do you decide $p_0(\beta_0) = 2/(1+\beta_0)^3$)

Then the marginal density of the full sample under M_0 is

$$m_0(x) = \frac{\nu^\xi \Gamma(n+\xi)}{\Gamma(\xi) \left[\prod_{i=1}^n x_i \right]} \int_0^\infty \frac{\beta_0^n \left[\prod_{i=1}^n x_i \right]^{\beta_0}}{(T^{\beta_0} + \nu)^{n+\xi}} p_0(\beta_0) d\beta_0$$

Given the model M_1, we assume the prior distribution on η_1, β_1, β_2, and τ are independent with $\eta_1 \sim Inverse\ \Gamma(\xi, \nu), p_i(\beta_i) = 2/(1+\beta_i)^3$ for $i = 1, 2$, and $\tau \sim U(0, T)$. Then the marginal density of the full sample under M_1 is

$$m_1(x) = \frac{\nu^{\xi}\Gamma(n+\xi)}{\Gamma(\xi)\left[\prod_{i=1}^{n}x_i\right]}\int_0^{\infty}\int_0^{\infty}\int_0^T \beta_1^j\beta_2^{n-j}\frac{\left[\prod_{i=1}^{j}x_i\right]^{\beta_2}\left[\prod_{i=j+1}^{n}x_i\right]^{\beta_2}}{K_1}p_1(\beta_1)p_1(\beta_2)\frac{1}{T}d\tau d\beta_1 d\beta_2,$$

where

$$K_1 = (\tau^{\beta_1} + T^{\beta_2} - \tau^{\beta_2} + \nu)^{n+\xi}.$$

The Bayes factor B_{10} for M_1 against M_0 is then $B_{10}(x) = A/B$, where

$$A = \int_{[0,x_1)}\int_0^{\infty}\int_0^{\infty}\frac{2}{(1+\beta_1)^3}\cdot\frac{2\beta_2^n}{(1+\beta_2)^3}\cdot\frac{\left[\prod_{i=1}^{n}x_i\right]^{\beta_2}}{K_1}d\beta_1 d\beta_2\frac{1}{T}d\tau$$

$$+ \sum_{j=1}^{n-1}\int_{[x_j,x_{j+1})}\int_0^{\infty}\int_0^{\infty}\frac{2\beta_1^j}{(1+\beta_1)^3}\cdot\frac{2\beta_2^{n-j}}{(1+\beta_2)^3}\cdot\frac{q_j}{K_1}d\beta_1 d\beta_2\frac{1}{T}d\tau$$

$$+ \int_{[x_n,T)}\int_0^{\infty}\int_0^{\infty}\frac{2\beta_1^n}{(1+\beta_1)^3}\cdot\frac{2}{(1+\beta_2)^3}\cdot\frac{\left[\prod_{i=1}^{n}x_i\right]^{\beta_1}}{K_1}d\beta_1 d\beta_2\frac{1}{T}d\tau,$$

and

$$B = \int_0^{\infty}\frac{2\beta_0^n}{(1+\beta_0)^3}\cdot\frac{\left[\prod_{i=1}^{n}x_i\right]^{\beta_0}}{(T^{\beta_0}+\nu)^{n+\xi}}d\beta_0,$$

with

$$q_j = \left[\prod_{i=1}^{j}x_i\right]^{\beta_1}\cdot\left[\prod_{i=j+1}^{n}x_i\right]^{\beta_2}, \text{for } j = 1,\cdots,n-1.$$

In this procedure, if the Bayes factor $B_{10} < 1$, we select the model M_0 and stop the procedure. After that, we move on testing multiple hypotheses regarding the shape parameter β_0 in (6). Otherwise, we select the model M_1, we estimate τ by using

$$\hat{\tau} = E(\tau \mid X) = \frac{C}{D}, \tag{7}$$

where

$$C = \int_{[0,x_1)}\int_0^\infty\int_0^\infty \frac{2}{(1+\beta_1)^3} \cdot \frac{2\beta_2^n}{(1+\beta_2)^3} \cdot \frac{\left[\prod_{i=1}^n x_i\right]^{\beta_2}}{K_1} d\beta_1 d\beta_2 \tau d\tau$$

$$+ \sum_{j=1}^{n-1}\int_{[x_j,x_{j+1})}\int_0^\infty\int_0^\infty \frac{2\beta_1^j}{(1+\beta_1)^3} \cdot \frac{2\beta_2^{n-j}}{(1+\beta_2)^3} \cdot \frac{q_j}{K_1} d\beta_1 d\beta_2 \tau d\tau$$

$$+ \int_{[x_n,T)}\int_0^\infty\int_0^\infty \frac{2\beta_1^n}{(1+\beta_1)^3} \cdot \frac{2}{(1+\beta_2)^3} \cdot \frac{\left[\prod_{i=1}^n x_i\right]^{\beta_1}}{K_1} d\beta_1 d\beta_2 \tau d\tau$$

and

$$D = \int_{[0,x_1)}\int_0^\infty\int_0^\infty \frac{2}{(1+\beta_1)^3} \cdot \frac{2\beta_2^n}{(1+\beta_2)^3} \cdot \frac{\left[\prod_{i=1}^n x_i\right]^{\beta_2}}{K_1} d\beta_1 d\beta_2 d\tau$$

$$+ \sum_{j=1}^{n-1}\int_{[x_j,x_{j+1})}\int_0^\infty\int_0^\infty \frac{2\beta_1^j}{(1+\beta_1)^3} \cdot \frac{2\beta_2^{n-j}}{(1+\beta_2)^3} \cdot \frac{q_j}{K_1} d\beta_1 d\beta_2 d\tau$$

$$+ \int_{[x_n,T)}\int_0^\infty\int_0^\infty \frac{2\beta_1^n}{(1+\beta_1)^3} \cdot \frac{2}{(1+\beta_2)^3} \cdot \frac{\left[\prod_{i=1}^n x_i\right]^{\beta_1}}{K_1} d\beta_1 d\beta_2 d\tau.$$

Now we divide the data into two parts. Repeat the procedure until no further changepoint is detected in all of the subdivisions. Note that the shape parameter β in (1) determines the behavior of failure times. In the next subsection, we extensively discuss the multiple test about the shape parameter β.

3.2 The Intrinsic Priors for the Power Law Model

For a repairable system, hypothesis testing about the shape parameter β in (1) plays an important role in determining the times between system failures. When $\beta = 1$, the intensity function is a constant; i.e. $\nu(t) = 1/\eta$, so that the power law process becomes a HPP. In this case, the frequency of failures is time independent and the system is experiencing no change over time. When $0<\beta<1$, the frequency of failures is a decreasing function of time so that the time between system failures increases and the system is improving over time. When $\beta>1$, the frequency of failures is a increasing function of time so that the time between system failures decreases and the system is deteriorating over time. Thus, we consider the following three hypotheses:

$$H_1 : \beta = 1; H_2 : 0 < \beta < 1; H_3 : \beta > 1. \tag{8}$$

The multiple test in (8) is extensively discussed by Lingham and Sivaganesan (1997) and

Kim and Sun (2000). Obviously, hypotheses H_2 and H_3 are more complex than H_1. An encompassing hypothesis $H_0 (0 < \beta < \infty)$ should be used in order for the AIBF to be valid (cf. Berger and Pericchi, 1996). From Berger and Pericchi (1996) the encompassing AIBF is defined as

$$B_{ji}^{0AI} = \frac{B_{0i}^{AI}(x)}{B_{0j}^{AI}(x)} = B_{ji}(x)\left(\frac{CFA_{i0}}{CFA_{j0}}\right), 1 \leq i,j \leq 3,$$

where $B_{ji}^{AI}(x)$ and CFA_{i0} are given by (3) and (4) respectively. However, this AIBF requires considerably heavy computation. Kim and Sun (2000) derived a general form of intrinsic priors for testing hypotheses in (8). It is presented in the next theorem.

Theorem 1 (Kim and Sun, 2000) Assume that Jeffreys's prior are used as starting priors. Then a set of intrinsic priors is

$$\pi_1^I(\eta) = g(\eta), 0 < \eta < \infty,$$
$$\pi_2^I(\eta,\beta) = g(\eta)\Delta_2(\eta,\beta)1_{\{0 < \beta < 1, 0 < \eta < \infty\}}, \qquad (9)$$
$$\pi_3^I(\eta,\beta) = g(\eta)\Delta_3(\eta,\beta)1_{\{\beta > 0, 0 < \eta < \infty\}},$$

where $g(\cdot)$ is proper on $(0,\infty)$, and

$$\Delta_2(\eta,\beta) = \frac{1}{\beta+1} \text{ and } \Delta_3(\eta,\beta) = \frac{1}{\beta(\beta+1)(\beta+2)}. \qquad (10)$$

Further, the normalizing constants of intrinsic priors π_2^I and π_3^I are $c_2 = 0.72135$ and $c_3 = 3.47605$, respectively.

Corollary 1 When $g(\cdot)$ is the probability density function of $Inverse\,Gamma\,(\xi,\nu)$, the set of intrinsic priors is

$$\pi_1^I(\eta) = \frac{\nu^\xi}{\Gamma(\xi)}\frac{1}{\eta^{\xi+1}}\exp\{-\nu/\eta\}, 0 < \eta < \infty,$$

$$\pi_2^I(\eta,\beta) = c_2\frac{\nu^\xi}{\Gamma(\xi)}\frac{1}{\eta^{\xi+1}}\exp\{-\nu/\eta\}\cdot\frac{2}{\beta+1}, 0 < \eta < \infty, 0 < \beta < 1, \qquad (11)$$

$$\pi_3^I(\eta,\beta) = c_3\frac{\nu^\xi}{\Gamma(\xi)}\frac{1}{\eta^{\xi+1}}\exp\{-\nu/\eta\}\cdot\frac{2}{\beta(\beta+1)(\beta+2)}, 0 < \eta < \infty, 1 < \beta,$$

where c_2 and c_3 are normalizing constants in Theorem 1.

As a matter of fact, the priors used in subsection 3.1 are all intrinsic priors except for τ. We only need to derive a set of intrinsic priors for testing H_1 against H_0. We do not present all the derivation details to come up with equation (5). These can be computed using the similar arguments in Appendix of Kim and Sun (2000). After taking limit, equation (5) becomes

$$\frac{\pi_0^I(\eta,\beta) \cdot 1/\eta}{1/(\eta\beta) \cdot \pi_1^I(\eta)} = \Delta_0(\eta,\beta). \tag{12}$$

Theorem 2 For any proper $g(\cdot)$ on $(0, \infty)$,

$$\pi_1^I(\eta) = g(\eta), 0 < \eta < \infty,$$

$$\pi_0^I(\eta,\beta) = g(\eta)\Delta_0(\eta,\beta)1_{\{0 < \beta < \infty, 0 < \eta < \infty\}}, \tag{13}$$

is the solution of (12), where $\Delta_0(\eta, \beta) = 1/(1+\beta)$. Further, the normalizing constant of the intrinsic prior π_0^I is $c_0 = 1/2$.

Proof: The proof is easily followed from the identical arguments as the proof of Theorem 1.

Remark 1 It is virtually impossible to find intrinsic priors for testing M_0 against M_1 in Subsection 3.1. However, we need sort proper priors to avoid heavy (training sample) computation for B_{10}. Because the shape parameter is not truncated in both models M_0 and M_1, a set of intrinsic priors in (13) could be used in a heuristic sense.

4. Numerical Results

4.1 Real Data Analysis

We analyzed the data of Jarrett (1979) on the time intervals between coal-mining disasters recorded from March 15, 1851 to March 22, 1962. There were 191 accidents in this period of 40,550 days. This dataset has been extensively used by several authors including Raftery and Akman (1986), Carlin et *al* (1992), Green (1995), and Yang and Kuo (2001). We fit the data for the power law process with the Bayes factor B_{10}. We consider a relatively diffuse piror for the scale parameter η. That is, $\eta_i \sim$ *Inverse Gamma* $(0.5, 0.0000001)$. For the complete data in $(0, 40550]$, the Bayes factor B_{10} is 1.072×10^3 and the estimated changepoint turned out to be 14450 by (7), where it is located between the 124th accident and the 125th accident. Next, we divide the complete data into two sub-divisions $(0, 14450]$ and $(14450, 40550]$. After calculating two Bayes factors, it appeared that no further changepoint is detected. That is, the Bayes factor B_{10} is 0.0632 in $(0, 14450]$, whereas the Bayes factor B_{10} is 0.5225 in $(14450, 40550]$. These results are quite similar to those of Yang and Kuo (2001), where they assume that the intensity function is a step function.

The next interesting feature is to see the behavior of failure times. In particular, we conduct a multiple test for the hypotheses: $H_1(\beta = 1), H_2(0 < \beta < 1), H_3(\beta > 1)$. We compute the posterior probabilities with the set of intrinsic priors in (11) assuming equal prior

model probabilities. Numerical results are reported in Table 1. It seems that the entire data follow the homogeneous Poisson process with a single changepoint. We also note that the MLEs of β for each sub-division are 0.9991 and 0.9935 respectively.

We use two more different hyperparameters in order to see if the values are stable. They are $(\xi, \nu) = (0.01, 0.01)$ and $(0.1, 0.1)$. The estimated changepoints are 14302 and 14315, where they are found in the same location. Furthermore, the Bayes factors and the posterior probabilities are quite close. Numerical values are reported in Table 2.

4.2 Simulation

We performed a simulation study. We generate datasets from the power law process with the intensity function in (1). These data consist of the following two datasets for the fixed scale parameter $\eta = 0.1$. We generate a dataset with size of 20 for $\beta = 0.3$ and generate a dataset with size of 20 for $\beta = 2$. So the complete data is size of 40. Since the MLEs of (η, β) are (0.0928, 0.2866) and (0.1033, 1.9988) respectively, it seems that the data are fairly well generated. The time interval of complete data is (0, 10.0936]. In (0, 10.0936], the Bayes factor B_{10} is 8.223×10 and the estimated changepoint is 8.8232, where it is located between the 20th observation and the 21th observation. We divide the complete data into two segments (0, 8.8233] and (8.8233, 10.0936]. In (0, 8.8233], the Bayes factor B_{10} is 0.9330. In (8.8233, 10.0936], the Bayes factor B_{10} is 0.3611. Now we move on to a multiple test to see the behavior of time intervals. We compute the posterior probabilities assuming equal prior model probabilities. Numerical values are reported in Table 3. The results are quite congruent with what we would expect from the data. Figure 1 and Figure 2 show the intensity functions with simulated datastes. As expected, the intensity functions are decreasing before the changepoint and increasing after the changepoint.

We generate datasets assuming there are two changepoints. We fix the scale parameter $\eta = 1.0$. The first dataset of size 20 is generated with $\beta = 1.0$. The second dataset of size 25 is generated with $\beta = 0.5$. The last dataset of size 20 is generated with $\beta = 2.0$. So the complete data is size of 65. The simulated data are plotted in Figure 3. The MLEs of (η, β) for each dataset are (0.8945, 1.0458), (1.0170, 0.4818), and (1.0428, 2.1601) respectively. The time interval of complete data is (0, 845.116]. In (0, 845.116], the Bayes factor B_{10} is 6.190×10^{31} and the estimated changepoint $\hat{\tau}_2$ is 841.355. This is located between the 45th observation and the 46th observation. We divide the complete data into two segments (0, 841.355] and (841.355, 845.116]. In (0, 841.355], the Bayes factor B_{10} is 4.499×10^4 and the estimated changepoint $\hat{\tau}_1$ is 17.3755, which is located between the 22th observation and 23th observation. In Table 4 we report three Bayes factors for each subdivision and corresponding posterior probabilities.

Table 1. The posterior probabilities and Bayes factors; coal-mining disasters data.

$(\xi, \nu)=(0.5, 0.0000001)$	$P(H_1|X)$	$P(H_2|X)$	$P(H_3|X)$	B_{10}
data before $\hat{\tau}$	0.8330	0.0672	0.0998	0.0623
data after $\hat{\tau}$	0.7745	0.1381	0.0844	0.5225

Table 2. The posterior probabilities and Bayes factors; coal-mining disasters data.

$(\xi, \nu)=(0.01, 0.01)$	$P(H_1|X)$	$P(H_2|X)$	$P(H_3|X)$	B_{10}
data before $\hat{\tau}$	0.8331	0.0679	0.0990	0.0535
data after $\hat{\tau}$	0.7763	0.1400	0.0837	0.7635
B_{10} for full data		5.971×10^3		
$(\xi, \nu)=(0.1, 0.1)$	$P(H_1|X)$	$P(H_2|X)$	$P(H_3|X)$	B_{10}
data before $\hat{\tau}$	0.8343	0.0723	0.0934	0.0545
data after $\hat{\tau}$	0.7679	0.1546	0.0775	0.7024
B_{10} for full data		4.431×10^3		

Table 3. The posterior probabilities and Bayes factors for the simulated data; a single change-point.

$(\xi, \nu)=(0.5, 0.0000001)$	$P(H_1|X)$	$P(H_2|X)$	$P(H_3|X)$	B_{10}
data before $\hat{\tau}$	7.888×10^{-10}	0.9999	1.800×10^{-11}	0.9330
data after $\hat{\tau}$	0.0265	0.0016	0.9719	0.3611
B_{10} for full data		8.223×10^{10}		

Table 4. The posterior probabilities and Bayes factors for the simulated data; two change-points.

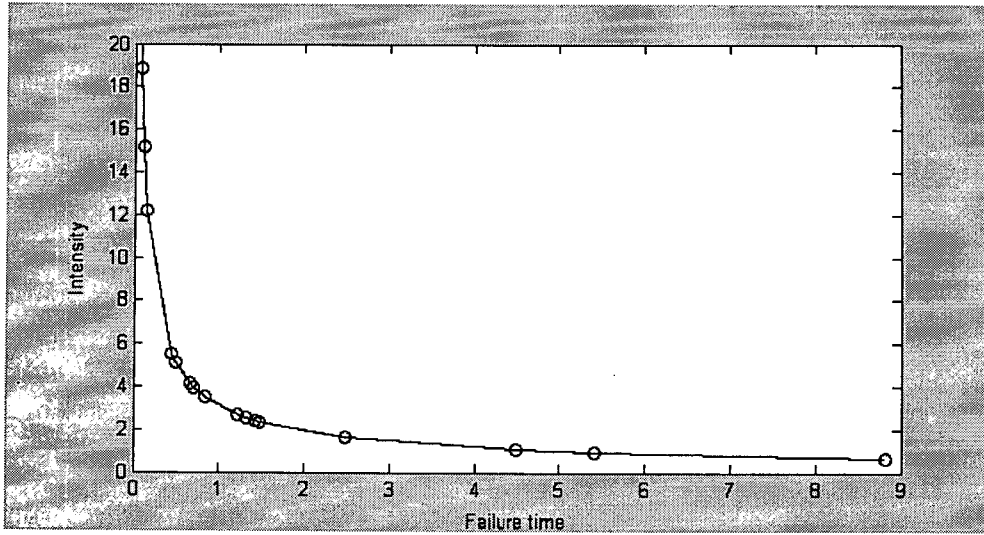| $(\xi, \nu)=(0.5, 0.0000001)$	$P(H_1|X)$	$P(H_2|X)$	$P(H_3|X)$	B_{10}
data before $\hat{\tau_1}$	0.6335	0.0678	0.2987	0.6143
data between $\hat{\tau_1}$ & $\hat{\tau_2}$	0.4149	0.5460	0.0391	0.2597
data after $\hat{\tau_2}$	0.0015	0.0001	0.9984	0.2952

Figure1. Plot of intensity as data before changepoint.
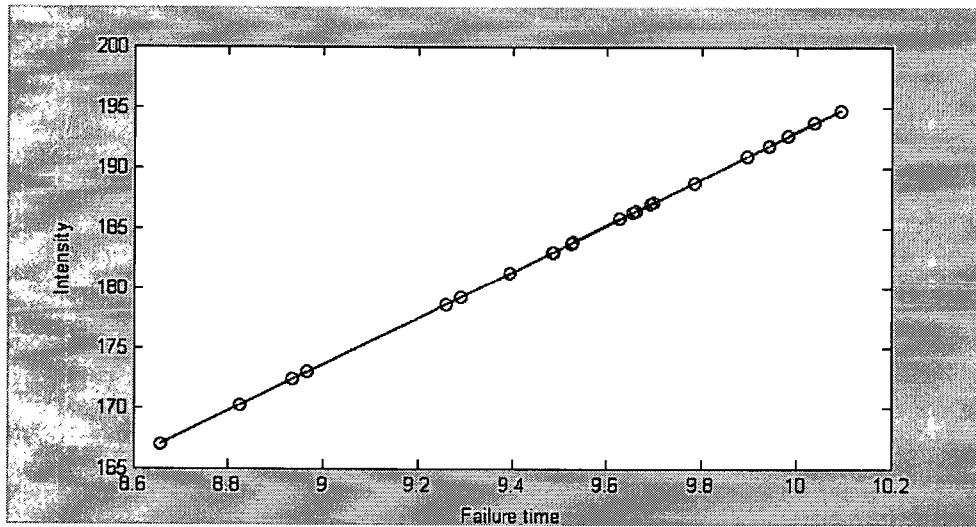

Figure2. Plot of intensity as data after changepoint.

5. Discussion

We proposed a Bayesian approach for detecting multiple changepoints in the power law process. In particular, we assume the changes in the shape parameter, which determines system behavior of the process. At each comparison in the binary segmentation procedure, we only need to compare a single changepoint model to a no-changepoint model. Thus, this methodology circumvents computational difficulties occurred in unknown multi-dimensional processes. We also conduct a multiple test based on the intrinsic priors. The computational

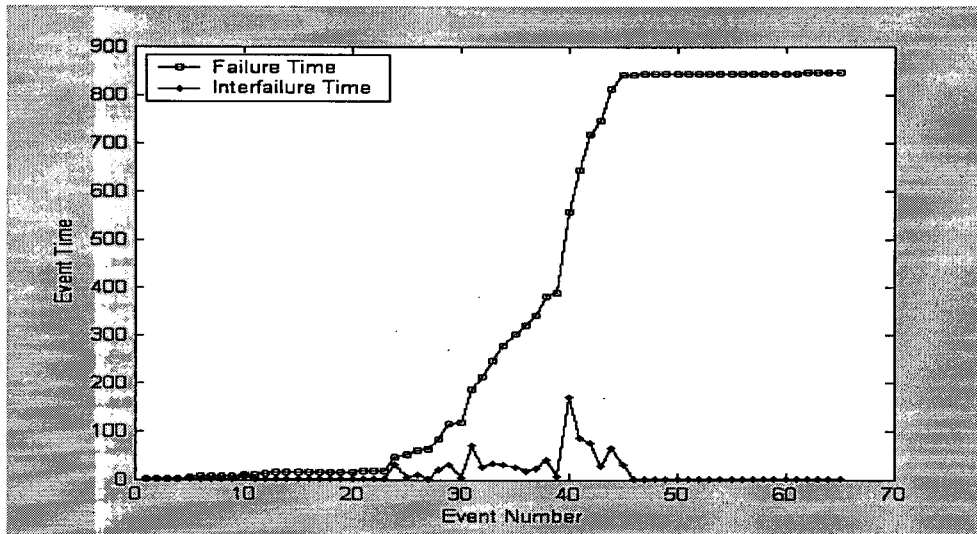results show that our procedure is quite feasible both for real data and simulated datasets.

Figure 3. Plot of simulated data with two changepoints.

References

[1] Ascher, H. and Feingold, H. (1984). Repairable Systems Reliability : Modeling, Inference, *Misconceptions and Their Causes*, Marcel Dekker, New York.

[2] Berger, J. O. and Pericchi, L. (1996). The intrinsic Bayes factor for model selection and prediction, *Journal of the American Statistical Association*, Vol. 91, 109-122.

[3] Carlin, B. P., Gelfand A. E., and Smith, A. F. M. (1992), Hierarchical Bayesian analysis of change point problems, *Journal of the Royal Statistical Society C*, Vol. 41, 389-405.

[4] Chen, J. and Gupta, A. K. (1997). Testing and locating variance changepoints with application to stock prices, *Journal of the American Statistical Association*, Vol. 92, 739-747.

[5] Chib, S. (1998). Estimation and comparison of multiple change-point models, *Journal of Econometrics*, Vol. 86, 221-241.

[6] Crow, L. H. (1974). Reliability Analysis of Complex Repairable Systems, *Reliability and Biometry*, F. Proschan and R. J. Serfling , (Eds.), 379-410, SIAM, Philadelphia.

[7] Duane, J. T. (1964). Learning curve approach to reliability monitoring, *IEEE Transactions on Aerospace*, Vol. 2, 563-566.

[8] Green, P, J. (1995). Revisible Jump Markov Chanin Monte Carlo Computation and Bayesian Model Determination, *Biometrika*, Vol. 82, 711-732.

[9] Jarrett, R.G. (1979). A note on the Intervals Between Coal-Mining Disasters, *Biometrika*, 66, 191-193.

[10] Kim, S. W. and Sun, D. (2000). Intrinsic priors for model selection using an encompassing model with applications to censored failure time data, *Lifetime Data Analysis*, Vol. 6, 251-269.

[11] Lingham, R. T. and Sivaganesan, S. (1997). Testing hypotheses about the power law process under failure truncation using intrinsic Bayes factors, *Annals of the Institute of Statistical Mathematics*, Vol. 49, 693-710.

[12] Raftery, A. E. (1994). Change point and change curve modeling in stocastic processes and spatial statistics, *Journal of Applied Statistical Science*, Vol. 1, 403-424.

[13] Raftery, A. E. and Akman, V. E. (1986). Bayesian analysis of a Poisson process with a changepoint, *Biometrika*, Vol. 73, 85-89.

[14] Yang, T. Y. and Lynn, Kuo (2001). Bayesian Binary Segmentation Procedure for a Poisson Process with Multiple Changepoints, *Journal of Computational and Graphical Statistics*, Vol. 10, 772-785.

[15] Vostrikova, L.J. (1981). Detecting disorder in multidimensional random processes, *Soviet Mathematics Doklady*, Vol. 24, 55-59.