A Bayesian Inference for Power Law Process with a Single Change Point*

Kiwoong Kim[1], Inkwon Yeo[2], Sinsup Cho[1] and Jae Joo Kim[1]

[1] Department of Statistics, Seoul National University, Seoul, Korea
[2] Division of Mathematics and Statistical Informatics,
Chonbuk National University, Chonbuk, Korea

Abstract

The nonhomogeneous poisson process (NHPP) is often used to model repairable systems that are subject to a minimal repair strategy, with negligible repair times. In this situation, the system can be characterized by its intensity function. There have been many NHPP models according to intensity functions. However, the intensity function of system in use can be changed because of repair or its aging. We consider the single change point model as the modification of the power law process. The shape parameter of its intensity function is changed before and after the change point. We detect the presence of the change point using Bayesian methodology. Some numerical results are also presented.

Key Words: Repairable System, Nonhomogeneous Poisson Process, Change Point, Bayes Factor, Power Law Process, Markov Chain and Monte Carlo technique

1. Introduction

A system is defined as a collection of two or more parts designed to perform one or more functions, Ascher and Feingold (1984). There are two well-known systems in reliability, a repairable and a nonrepairable system. While a nonrepairable system can not be mended when it is out of order, a repairable system can be repaired and placed back in service. Since the system is returned to an operating condition to perform required functions, the model for the repairable system must be capable of reflecting the whole sequences of repeated failures and changes in reliability of the system as it ages. Poisson processes are often used in modeling the failure sequence of the repairable system, for instance, Thompson (1988). Involving the change in reliability, the pattern of interarrival times between failures is of interest. The system is called deteriorating and improving model when the interarriaval time tends to be shorter and longer, respectively, as it ages.

* Sinsup Cho's research was supported by a grant (1999-1-104-002-3) from the interdisciplinary research program of the Korean Science and Engineering Foundation.

Let $\{N(t)\}$ be the Poisson process with the intensity function, $\mu(t)$. According to Crow (1990), the realistic consideration of the system burn-in, wear-out, and maintenance policies often requires an approach that recognizes that the intensity function may not be constant over the operating time of interest. Hence, the proper intensity function of repairable system can include a decreasing, constant, and increasing function of time t. The intensity function decreasing in time provides a model for improving in reliability growth and the increasing function leads to a model for deteriorating system. A family of intensity functions indexed by a shape parameter was proposed by Duane (1964) and Cox and Lewis (1966). Some modifications and extensions have been investigated by Goel and Okumoto (1979), and Achcar *et al.* (1997). Among them, the power law by Duane (1964) and the log-linear by Cox and Lewis (1966) process are commonly used in modeling repairable systems because they have a simple form and can be improving or deteriorating over t by varying its shape parameter. The intensity function of power law process takes the form

$$\mu(t) = \lambda \beta t^{\beta-1},\ 0 < \lambda,\ 0 < \beta.$$

The shape parameter β plays an important role in determining the interarrival time between system failures. For the power law model, the intensity function with β=1 reduces to the constant λ for whole time t and the frequency of failures becomes time-independent. That is, the system follows a homogeneous poisson process (HPP). The intensity function with β>1 is increasing for time t so that the interarrival time between system failures tends to be decreased and the system is deteriorating over time. On the contrary, β<1 makes the frequency of failures decreases for time t so that the system becomes more reliable for larger t.

NHPP is often used to model repairable systems that are subject to a minimal repair strategy with negligible repair times. The minimal repair means that a failed system is restored just back to functioning state (as bad as old condition). In this paper, we consider a single change point problem in the power law process. Bayesian approaches are adopted in selection of the model and estimation of parameters. The Bayes factor is calculated to choose a model which has either no or one change point. The parameters in the selected model are estimated by Markov chain Monte Carlo (MCMC) techniques.

2. Single Change Point Model in Poisson Processes

Let M denote the single change point model in the time truncated situation. Then, the

intensity function for this model can be written as

$$\mu(t)=\begin{cases}\mu_b(t), & 0<t\le\tau,\\ \mu_a(t), & \tau<t\le T,\end{cases}$$

where τ is a change point and T is a predetermined terminal time. Let N be the number of failures over $(0,T]$ and let $X=(X_1,\cdots, X_N)^t$ be the ordered failure times, that is, $0<X_1<X_2<...<X_N\le T$. We assume that the number of failures follows NHPP with the intensity function, $\mu(t)$. Then, the joint density function of N and X under the model M is

$$\begin{aligned} f(n,x\mid M) &= f(x\mid n,M)f(n\mid M)\\ &= n\ !\left(\prod_{i=1}^{n}\frac{\mu(x_i)}{\int_o^T\mu(u)du}\right)\frac{\exp(-\int_o^T\mu(u)du)\left(\int_o^T\mu(u)du\right)^n}{n\ !}\\ &=\left(\prod_{i=1}^{n_\tau}\mu_b(x_i)\right)\exp\left(-\int_o^\tau\mu_b(u)du\right)\left(\prod_{i=n_\tau+1}^{n}\mu_a(x_i)\right)\exp(-\int_\tau^T\mu_a(u)du), \end{aligned} \tag{1}$$

where n_τ denotes the number of failures over $(0, \tau]$.

The problem of interest is the model selection in which the power law process has either no or exactly one change point at τ. We assume that the system under study may experience an abrupt change in its rate of failure occurrences due to the repair or replacement so that the intensity function may not be continuous in time. Carlin *et al.* (1992) proposed a restricted model in which the change point only occurs at failure time. We relax this restriction hence the proposed model can have a change point in any place on the time interval $(0,T]$.

Let M_0 be the model without the change point in the power law process and let M_1 be the model with a single change point. Then, the intensity functions of M_0 and M_1 are given as

$$M_0:\ \mu_0(t)=\lambda_0\beta_0 t^{\beta_0-1},\ 0<\lambda_0,\ 0<\beta_0,\ 0<t\le T,$$

and, for $\lambda>0$,

$$M_1:\ \mu_1(t)=\begin{cases}\lambda_1\beta_1 t^{\beta_1-1}, & 0<\beta_1,\ 0<t\le\tau,\\ \lambda_1\beta_2 t^{\beta_2-1}, & 0<\beta_2,\ \tau<t\le T,\end{cases}$$

respectively. Suppose that the number of failures is n and the failure times x_k, k=1,..., N are observed, $0<x_1<x_2<...<x_N \leq T$. Then, the likelihood function of $\theta=(\lambda_0,\ \beta_0)^t$ under M_0 is

$$L(\theta_0 \mid x) = \left(\prod_{i=1}^{n} \lambda_0 \beta_0 x_i^{\beta_0 - 1} \right) \exp\left(-\lambda_0 T^{\beta_0}\right)$$
$$= {\lambda_0}^n {\beta_0}^n \left(\prod_{i=1}^{n} x_i^{\beta_0 - 1} \right) \exp\left(-\lambda_0 T^{\beta_0}\right),$$

and the likelihood function of $\theta_1=(\lambda_1,\ \beta_1,\ \beta_2,\ \beta_3,\ \tau)^t$ under M_1 is, from (1),

$$L(\theta_1 \mid x) = \left(\prod_{i=1}^{n_\tau} \lambda_1 \beta_1 x_i^{\beta_1 - 1} \right) \exp\left(-\int_o^\tau \lambda_1 \beta_1 u^{\beta_1 - 1} du\right) \left(\prod_{i=n_\tau+1}^{n} \lambda_1 \beta_2 x_i^{\beta_2 - 1} \right) \exp\left(-\int_\tau^T \lambda_1 \beta_2 u^{\beta_2 - 1} du\right)$$
$$= {\lambda_1}^n {\beta_1}^{n_\tau} {\beta_2}^{n-n_\tau} \left(\prod_{i=1}^{n_\tau} x_i^{\beta_1 - 1} \right) \left(\prod_{i=n_\tau+1}^{n} x_i^{\beta_2 - 1} \right) \exp\left(-\lambda_1 \left(\tau^{\beta_1} + T^{\beta_2} - \tau^{\beta_2}\right)\right).$$

3. Model Selections and Estimations

In Bayesian model selection and hypothesis testing, the Bayes factor provides an important information to make a decision. However, the Bayes factor is not well defined when the prior distribution is improper since it depends on a ratio of two arbitrary constants. In this paper, we assume a conjugate prior for each parameter, to avoid this arbitrariness. Once, the model M_i is selected by comparing with Bayes factors, then θ_i is estimated by MCMC methods.

Suppose that we want to select the more plausible model out of M_0 and M_1 when observations x are given. Let $f_i(x \mid \theta_i)$ be the a parametric density function under model M_i where θ_i is the vector of unknown parameters on parameter space Θ_i, i=0,1. Let $\pi_i(\theta_i)$ be the prior density function of θ_i and let $P(M_i)$ be the prior probability that the model M_i is selected. Then, the posterior probability that the model M_0 is selected is

$$P(M_0 \mid x) = \left(1 + \frac{P(M_1)}{P(M_0)} B_{10}\right)^{-1},$$

where B_{10} is the Bayes factor of the model M_1 to the model M_0 and is defined as

$$B_{10} = \frac{m_1(x)}{m_0(x)} = \frac{\int_{\Theta_1} f_1(x \mid \theta_1)\pi_1(\theta_1)d\theta_1}{\int_{\Theta_0} f_0(x \mid \theta_0)\pi_0(\theta_0)d\theta_0}.$$

Here, $m_i(x)$ stands for the marginal (or prior predictive) density function of X under the model M_i, i=0,1 Assuming that $P(M_0)=P(M_1)$, we obtain the posterior probabilities

$$P(M_0 \mid x) = \frac{1}{1+B_{10}}$$

and

$$P(M_1 \mid x) = \frac{1}{1+B_{01}} = \frac{B_{10}}{1+B_{10}}$$

Consequently, we see that $P(M_0|x)<P(M_1|x)$ if and only if $B_{10}>1$. Therefore, we choose the model M_1, if $B_{10}>1$, and M_0, otherwise.

Let $g(\cdot;a,b)$ denote the gamma density function with mean a/b and let $\pi(\cdot)$ denote an arbitrary prior density function with positive support. Assuming priors $\lambda_0 \sim g(\cdot;a,b)$, $\beta_0 \sim \pi_0(\cdot)$, and λ_0 is independent of β_0, we see that the posterior density function of θ_0 given x is

$$\begin{aligned}\pi(\theta_0 \mid x) &\propto L(\theta_0 \mid x)g(\lambda_0;a_0,b_0)\pi_0(\beta_0) \\ &\propto \lambda_0^{n+a_0-1}\beta_0^n\left(\prod_{i=1}^n x_i^{\beta_0-1}\right)\exp\left(-\lambda_0\left(T^{\beta_0}+b_0\right)\right)\pi_0(\beta_0)\end{aligned}$$

and the marginal density function of M_0 is

$$m_0(x) = \frac{b_0^{a_0}\Gamma(n+a_0)}{\Gamma(a_0)}\int_0^\infty \frac{\beta_0^n\left(\prod_{i=1}^n x_i^{\beta_0-1}\right)}{(T^{\beta_0}+b_0)^{n+a_0}}\pi_0(\beta_0)d\beta_0,$$

where $\Gamma(\cdot)$ stands for the gamma function.

Under the model M_1, assuming the priors for $\theta_1=(\lambda_1,\ \beta_1,\ \beta_2,\ \tau)'$ as follows.

1. $\lambda_1 \sim g(\cdot;a,b)$, where a_1 and b_1 are known,
2. $\tau \sim U(0,T)$, where $U(0,T)$ denotes the uniform distribution over $(0,T)$,

3. $\beta_1 \sim \mu_1(\cdot)$ and $\beta_2 \sim \mu_2(\cdot)$, $\mu_1(\cdot)$ and $\mu_2(\cdot)$ are arbitrary density functions with positive support.

Note that, for λ_0 and λ_1, the gamma distribution is the conjugate prior distribution. The support of β_1 and β_2 should be positive and some candidate distributions for the prior of β including the lognormal, inverse gamma, and gamma distribution are available. When there is no information about whether the model is improving or deteriorating, it is reasonable to assign the same probability to two possibilities. Hence, we choose the lognormal distribution as the prior of β because it satisfies the condition that $P(\beta>1)=P(\beta<1)$. Suppose that λ_1, β_1, β_2 and τ are mutually independent. Then, the posterior density of θ_1 is

$$\pi(\theta_1 \mid x) \propto L(\theta_1 \mid x) g(\lambda_1; a_1, b_1) I_{(0,T)}(\tau) \pi_1(\beta_1) \pi_2(\beta_2)$$

$$\propto \lambda_1^{n+a_1-1} \beta_1^{n_\tau} \beta_2^{n-n_\tau} \left(\prod_{i=1}^{n_\tau} x_i^{\beta_1-1} \right) \left(\prod_{i=n_\tau+1}^{n} x_i^{\beta_2-1} \right) \exp\left(-\lambda_1 \left(\tau^{\beta_1} + T^{\beta_2} - \tau^{\beta_2} + b_1\right)\right) \pi_1(\beta_1) \pi_2(\beta_2)$$

and the marginal density function of M_1 is

$$m_1(x) = \frac{b_1^{a_1} \Gamma(n+a_1)}{\Gamma(a_1)} \int_0^\infty \int_0^\infty \int_0^T \frac{\beta_1^{n_\tau} \beta_2^{n-n_\tau} \left(\prod_{i=1}^{n_\tau} x_i^{\beta_1-1} \right) \left(\prod_{i=n_\tau+1}^{n} x_i^{\beta_2-1} \right)}{T(T^{\beta_2} + \tau^{\beta_1} - \tau^{\beta_2} + b_1)^{n+a_1}} \pi_1(\beta_1) \pi_2(\beta_2) d\tau d\beta_1 d\beta_2.$$

Since the integrations for $m_0(x)$ and $m_1(x)$ can not be analytically tractable, it can be computed by standard numerical integration techniques, approximated via Laplacian approximations with saddle point accuracy, or integrated by using the Monte Carlo simulation techniques. Once, a model is selected from the procedure above then the parameters of the selected model should be estimated. Since the posterior is analytically intractable, we apply MCMC technique in the estimation of parameters. Gibbs sampling is very useful for the implementation of the Bayesian estimation. It requires sampling from the full conditionals

$$\pi(\lambda_1 \mid \tau, \beta_1, \beta_2, x) = g(\lambda_1; n+a_1, T^{\beta_2} + \tau^{\beta_1} - \tau^{\beta_2} + b_1)$$

$$\pi(\tau \mid \lambda_1, \beta_1, \beta_2, x) \propto \beta_1^{n_\tau} \beta_2^{n-n_\tau} \left(\prod_{i=1}^{n_\tau} x_i^{\beta_1-1} \right) \left(\prod_{i=n_\tau+1}^{n} x_i^{\beta_2-1} \right) \exp\left(-\lambda_1 \left(\tau^{\beta_1} - \tau^{\beta_2}\right)\right)$$

$$\pi(\beta_1 \mid \lambda_1, \tau, \beta_2, x) \propto \beta_1^{n_\tau} \left(\prod_{i=1}^{n_\tau} x_i^{\beta_1-1} \right) \exp\left(-\lambda_1 \tau^{\beta_1}\right) \pi_1(\beta_1)$$

$$\pi(\beta_2 \mid \lambda_1, \tau, \beta_1, x) \propto \beta_2^{n-n_\tau} \left(\prod_{i=n_\tau+1}^{n} x_i^{\beta_2 - 1} \right) \exp\left(-\lambda_1 \left(T^{\beta_2} - \tau^{\beta_2} \right)\right) \pi_2(\beta_2)$$

Here, the marginal distributions of τ, β_1 and β_2 are nonstandard conditionals. We use the Metropolis-Hastings algorithm to generate samples, see Gilks *et al.*(1996).

4. Numerical Examples

We perform simulation studies in order to investigate the performance of our procedure in the simple case (β_0=1 in M_0 and β_1=1 in M_1). The numbers in parentheses(Table 1 - Table 4) are the standard deviations of estimators. Pseudo random variates are generated from the no change model M_0 with given parameter sets. Then the selection proportion (p in Table 1) are calculated in the following way: we count the number of Bayes factors $B_{10}=m_1(x)/m_0(x)$ which is less than 1 and divide it by the number of replications, 10000 in our case. This proportion is the measure of how the posterior probability supports the true model. Finally, the posterior mean, $\hat{\lambda}_0 = (a_0 + n)/(b_0 + T)$, is computed when the model M_0 is accepted. Throughout the section, $g(\cdot;0.5,0.5)$ is used as the prior of the intensity parameter λ. Numerical values are reported in Table 1. It is observed that the selection proportion of the true model is close to 0.75 and is maximized when the value of the intensity parameter λ is 1.0. Since we can guess the sample size using the mean value function the proposed procedure support the true model quite well as the sample size grows. We guess the sample size using the mean value function.

Table 1. Simulation results under M_0, T=10 and 10000 replications

λ_0	$\hat{\lambda}_0$	p
0.7	0.723(0.255)	0.747
0.8	0.818(0.281)	0.747
0.9	0.909(0.275)	0.740
1.0	1.003(0.303)	0.772
1.1	1.102(0.309)	0.745
1.2	1.184(0.335)	0.746
1.3	1.271(0.352)	0.743

We repeat the same procedures for model M_1 with given parameter sets. When model M_1 was accepted, we computed $\hat{\tau}$, posterior mean. The lognormal distribution, $\log(\beta_2) \sim \log(0, 0.2^2)$, is adopted as the prior of the shape parameter β_2 which means that we do not know whether the system is improving or deteriorating over the time. From Table 2, 3 and 4 we can see that the selection proportion p is incrased as we increase the value of the shape parameter β_2 from 1 (HPP).

Table 2. Simulation results under M_1, τ=3, T=10 and 10000 replications

λ_1	β_2	$\hat{\tau}$	p
1	1.4	4.351(0.814)	0.813
1	1.5	4.017(0.797)	0.951
1	1.6	3.696(0.835)	0.986
1	1.7	3.463(0.800)	0.999

Table 3. Simulation results under M_1, τ=5, T=10 and 10000 replications

λ_1	β_2	$\hat{\tau}$	p
1	1.4	4.833(0.673)	0.858
1	1.5	4.989(0.736)	0.982
1	1.6	4.829(0.716)	0.996
1	1.7	4.922(0.625)	1.000

Table 4. Simulation results under M_1, τ=7, T=10 and 10000 replications

λ_1	β_2	$\hat{\tau}$	p
1	1.4	5.604(0.777)	0.838
1	1.5	5.938(0.802)	0.957
1	1.6	6.347(0.732)	0.993
1	1.7	6.680(0.584)	0.999

We also observe that for τ=3, Table 2, the estimated posterior mean $\hat{\tau}$ is larger than the true value. This is because the sample size before the change point is too small to detect system's immediate change. On the other hand when τ=5 and τ=7 the estimated posterior mean $\hat{\tau}$ is smaller than the true value. But as we increase the value of the shape parameter β_2, $\hat{\tau}$ approaches the true value τ and we also attain the higher selection proportion.

4. References

1. Achcar, J. A., Dey, D. K. and Niverthi, M. (1997). "A Bayesian approach using nonhomogeneous Poisson process for software reliability model", Frontiers in Reliability, IAPQR.
2. Ascher, H and Feingold, H. (1984). "Repairable Systems Reliability : Modeling, Inference, Misconceptions and Their Causes", Marcel Dekker, New York.
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, 41, 389-405.
4. Cox, D. R. and Lewis, P. A. (1966). "The Statistical Analysis of Series of Events", Methuem, London.
5. Crow, L. H. (1990). "Evaluating the reliability of reparirable systems", Proceedings of the Annual Reliability and Maintainability Symposium: IEEE, 275-279.
6. Duane, J. T. (1964). "Learning curve approach to reliability monitoring", IEEE Transactions on Aerospace, 2, 263-566.
7. Gilks, W. R., S. Richardson, and D. J. Spiegelhalter (1996). "Markov chain Monte Carlo in practice", Chapman and Hall, New York.
8. 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.
9. Thompson Jr., W. A. (1988). "Point Process Models with Applications to Safety and Reliability", Chapman and Hall, New York.