http://ansinet.com/itj

ITJ

ISSN 1812-5638

INFORMATION TECHNOLOGY JOURNAL

ANSInet
Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Information Technology Journal 12 (21): 6203-6207, 2013
ISSN 1812-5638 / DOI: 10.3923/itj.2013.6203.6207

Bayesian Method for Reliability Assessment Based on Zero-failure Data under Small Sample Size: Application for High Speed Railway Vehicle

Jianwei Yang, Jinhai Wang and Yidong Xie
School of electric-mechanical and automotive engineering,
Beijing University of Civil Engineering and Architecture, 100044, Beijing, China

Abstract: In order to analyze the reliability of parts and components based on zero-failure data under small sample size, Bayesian method is applied with Weibull distribution to solve this problem. Gamma distribution and Uniform distribution are chosen as prior distribution and hierarchical prior distribution. And then, the hybrid MCMC algorithm is proposed to compute the estimators of posterior distribution. After that, taken air compressor in the braking system of railway vehicle as an example, the zero-failure data is assessed using the method given above. Computing result shows that the method can assess the reliability of air compressor based on zero-failure data under small sample size.

Key words: Zero-failure, hierarchical bayesian, small sample size, hybrid MCMC algorithm, high speed railway vehicle, weibull distribution

INTRODUCTION

The reliability of high speed railway vehicle which is made of lots of complex system is significant for safety. Recently, more and more scholars have researched this problem. Three-parameter Weibull model with the least square method and correlation coefficient method is used to analyze key components and parts of railway equipment by Wang *et al*. (2008). Wu *et al*. (2009) has researched the failure rate of relay valve of braking system for reliability with two-parameter Weibull model. And then Wang and Wu (2010) assess the reliability of unit brake for urban rail vehicle according to the durability test. But, both of them have a same shortcoming that the accuracy of assessment cannot perform well if the sample size is small. To crack this hard nut, the sample size is increased from n = 2~3 to n = 12 using virtually expanded sample method for assessing reliability of structure of body bolsters railway vehicle in Bootstrap method by Tian (2008). Despite the approach can be satisfied to engineering application under extreme small sample circumstance, its limitation is that the method can be only used in assuming the failure time model based on Gaussian distribution.

Reliability assessment of high speed railway vehicle is a time-consuming and costly task due to a large number of samples required, n = 30, generally. It is impossible to get a large number of samples because of the feature of railway vehicle that the lifetime is long, the cost is expensive and the constructer of system is complex. Above all, an approach which can be applied to assess the reliability of lifetime cycle under small sample size is necessary.

BAYESIAN INFERENCE

For solving the problem of reliability assessment under small sample size, Bayesian method is used to combine sample information with prior information. Bayesian approach is applied in analysis of industrial experiment and the method is introduced in detail include Bayesian $\div^2$ goodness-of-fit test, model selection, response surface and so on by Weaver and Hamada (2008). Automotive reliability is analyzed based on past data and technical knowledge using Bayesian method by Guida and Pulcini (2002). The effect of small sample size and low mean values for modeling motor vehicle crash is studied in Bayesian method through simulation of a large number of data by Lord *et al*. (2008). The result of fatigue test is obtained for Shinkansen vehicle axle using Bayesian method to combine full-scale model's data and small specimens' data by Akama (2002). The reliability of space bearing is estimated under small sample size based on the failure that caused from lubricant volatilization by Jin (2010). The Bayesian formula is given as follows:

$$\pi(\theta|x)=\frac{L(x|\theta)\times\pi(\theta)}{\int L(x|\theta)\times\pi(\theta)} \quad (1)$$

Corresponding Author: Jianwei Yang, School of Electric-Mechanical and Automotive Engineering, Beijing University of Civil Engineering and Architecture, 100044, Beijing, China

where, $\pi(\theta|x)$ is posterior joint distribution, $\pi(\theta)$ is prior joint distribution, $L(x|\theta)$ is likelihood function where parameters information in sample are all included. It is obvious that Bayesian inference takes full advantage of historical data in estimating parameters.

Zero-failure model: Weibull distribution is widely applied on several fields. And its model is given as follows:

$$f(x|\lambda,\beta) = \lambda\beta x^{\beta-1}\exp\left(-\lambda x^{\beta}\right) \tag{2}$$

$$F(x) = 1-\exp\left(-\lambda x^{\beta}\right) \tag{3}$$

$$R(x) = \exp\left(-\lambda x^{\beta}\right) \tag{4}$$

$$h(x) = \lambda\beta x^{\beta-1} \tag{5}$$

With increase of lifetime and reliability of product, the lifetime data of failure is obtained less and less in type 1 censoring test. And there is a problem for zero-failure data because the data cannot be obtained.

Let $x_{1,\ldots,}$ x_n be the observed failure time of type 1 censoring test, the likelihood of Weibull distribution based on zero-failure data is given by:

$$L(\lambda,\beta|0) = \prod_{i=1}^{n}\left[1-F(x_i)\right]^{R_i} = \exp\left(-\sum_{i=1}^{n}R_i\lambda x_i^{\beta}\right) \tag{6}$$

For frequentist, the likelihood function contains all information about sample. It is, however, inadequate for statistical inference under small sample size and zero-failure data. For solving the problem, Bayesian method is proposed to combined sample data with prior data that include historical data of similar product, expert information and simulation information.

Bayesian estimation: The most important thing in Bayesian inference is selecting prior distribution and determining hyper parameter in reasonable way. Due to the same range, Gamma distribution is selected for most of parameters which their interval is from zero to infinity. Besides, some parameters that their range is a closed interval between two real number may apply Uniform distribution as prior distribution based on the principle of indifference which considers each possible value has same opportunity to be choose. The prior distribution of parameters λ and β are given as follows:

$$\pi(\lambda|a_\lambda,b_\lambda) = \frac{\lambda^{a_\lambda-1}}{\Gamma(a_\lambda)}\exp(-b_\lambda\lambda) \tag{7}$$

$$\pi(\beta|a_\beta,b_\beta) = \frac{\beta^{a_\beta-1}}{\Gamma(a_\beta)}\exp(-b_\beta\beta) \tag{8}$$

where, a_λ, b_λ, a_β and b_β are hyper parameters of prior distributions $\pi(\lambda)$ and $\pi(\beta)$ with $a_\lambda>0$, $b_\lambda>0$, $a_\beta>0$ and $b_\beta>0$, respectively. In general, it is difficult to computer definite hyper parameter based on prior information. To crack this hard nut, hierarchical Bayesian model is proposed for controlling the range for a_λ and b_λ using Uniform distribution as follows:

$$\pi(a_\lambda) = \frac{1}{c_a} \tag{9}$$

$$\pi(b_\lambda) = \frac{1}{c_b} \tag{10}$$

where, c_a and c_b are hyper parameters of hierarchical prior distribution.

From Eq. 7-10, the joint prior density distribution can be obtained as:

$$\begin{aligned}\pi(\lambda,\beta,a_\lambda,b_\lambda) &= \pi(\lambda|a_\lambda,b_\lambda)\pi(a_\lambda)\\ &\quad \pi(\beta|a_\beta,b_\beta)\pi(b_\lambda)\\ &= \iiiint\frac{1}{c_\lambda}\frac{b_\lambda^{a_\lambda}b_\beta^{a_\beta}}{\Gamma(a_\lambda)\Gamma(a_\beta)}\\ &\quad \exp\left(-b_\lambda\lambda-b_\beta\beta\right)\beta^{a_\beta-1}\\ &\quad \lambda^{a_\lambda-1}da_\lambda db_\lambda da_\beta db_\beta\end{aligned} \tag{11}$$

Substituting Eq. 6, 11-1, the posterior joint probability density can be obtained as follows:

$$\pi(\lambda,\beta,a_\lambda,b_\lambda|x) = \frac{\iiiint\frac{1}{c_\lambda}\frac{b_\lambda^{a_\lambda}b_\beta^{a_\beta}}{\Gamma(a_\lambda)\Gamma(a_\beta)}\beta^{a_\beta-1}\lambda^{a_\lambda-1}e^{\left(-b_\lambda\lambda-b_\beta\beta-\sum_{i=1}^{n}R_i\lambda x_i^{\beta}\right)}da_\lambda db_\lambda da_\beta db_\beta}{\iiiiint\!\!\int\frac{1}{c_\lambda}\frac{b_\lambda^{a_\lambda}b_\beta^{a_\beta}}{\Gamma(a_\lambda)\Gamma(a_\beta)}\beta^{a_\beta-1}\lambda^{a_\lambda-1}e^{\left(-b_\lambda\lambda-b_\beta\beta-\sum_{i=1}^{n}R_i\lambda x_i^{\beta}\right)}da_\lambda db_\lambda da_\beta db_\beta d\lambda d\beta} \tag{12}$$

In Bayesian statistics the posterior distribution $\pi(\lambda, \beta|x)$ contains all relevant information on the unknown parameters given observed data. From Eq. 12, it is almost impossible to get normalizing constant due to difficulty in high dimensional integration. So, it is also tough to obtain directly analytical solution because of complicated high dimensional integration and having no closed form. In order to solve this problem, MCMC algorithm which is a dynamic sampling algorithm has been proposed.

MARKOV CHAIN MONTE CARLO ALGORITHM

Markov chain Monte Carlo algorithm also called MCMC which provides an efficient and simple way to sample from target distribution that is always high dimensional and complex is often used in Bayesian inference. Metropolis *et al*. (1953) established the general framework of MCMC. And then Geman and Geman (1984) developed the algorithm named Gibbs sampling algorithm for special situation that conditional distribution can be sampled directly. Usually, for achieving a quick computation speed, Metropolis Hasting algorithm and Gibbs algorithm are combined for constructing a Markov chain. In this article, the full posterior conditional distribution for λ, β, a_λ and $b_\lambda$are proportional to, respectively:

$$\pi(\lambda \mid x, \beta, a_\lambda, b_\lambda) \propto \lambda^{a_\lambda - 1} e^{-\lambda\left(b_\lambda + \sum_{i=1}^{n} R_i x_i^\beta\right)} \tag{13}$$

$$\pi(\beta \mid x, a_\lambda, b_\lambda, \lambda) \propto \beta^{a_\beta - 1} e^{-b_\beta \beta - \sum_{i=1}^{n} R_i \lambda x_i^\beta} \tag{14}$$

$$\pi(a_\lambda \mid x, b_\lambda, \beta, \lambda) \propto \frac{(b_\lambda \lambda)^{a_\lambda}}{\Gamma(a_\lambda)} \tag{15}$$

$$\pi(b_\lambda \mid x, \lambda, \beta, a_\lambda) \propto b_\lambda^{a_\lambda} e^{-b_\lambda \lambda} \tag{16}$$

From Eq. 13-16, we know that Eq. 13 and 16 is Gamma distribution with parameters a_λ:

$$b_\lambda + \sum_{i=1}^{n} R_i x_i^\beta$$

$a_\lambda+1$ and λ, respectively and it can be sampled using Gibbs algorithm. However, rest of conditional posterior distributions of β and a_λ can only be sampled using Metropolis Hasting algorithm due to its form is not any distribution sampling directly. With the notations defined, the basic scheme of the Hybrid MCMC algorithm for these four posterior distributions is given as follows:

- Choose a starting point $\lambda(0)$, $\beta(0)$, $a_\lambda(0)$ and $b_\lambda(0)$ and set i = 1
- i = i+1
- Generate $\lambda(i)$ from $\Gamma(a_\lambda, b_\lambda+\Sigma^n_{i=1} R^i x_i^\beta)$
- Generate $\beta(i)$ from $U(0, M_\beta)$ using Metropolis Hating algorithm
- Generate $a_\lambda(i)$ from $U(0, M_\beta)$ using Metropolis Hating algorithm
- Generate $b_\lambda(i)$ from $\Gamma(a_\lambda+1, \lambda)$
- Repeat Eq. 2-6 T times

After that, we run a "burn-in" period $T_{burn-in}$ to delete these samples which do not converge to the target distribution. Finally, Monte Carlo approach is used to estimate the posterior mean of each parameter.

APPLICATION FOR AIR COMPRESSOR

A set of lifetime test data of air compressor of braking system of high speed railway vehicle is shown in Table 1. Because the lifetime of the air compressor is long and the cost of experiment is expensive, there is no any air compressor failure and the quantity of sample is only 20. So reliability assessment is under small sample and zero-failure data.

From historical data of similar air compressor and expert opinion, hyper parameters can be calculated by moment estimation method. Because expert opinion is not a definite value for parameter λ, the Uniform distribution is selected as hierarchical prior distribution. For limiting of length of article, there is no more detail of defineing prior information. Finally, the hyper parameters are given as $(c_a, c_b, a_\beta, b_\beta) = (10, 10, 1, 1)$.

After that, we use MATLAB to sample from the posterior distribution whose length of Markov chain is 200000 using the method given above. Figure 1 shows the iteration process of each parameter.

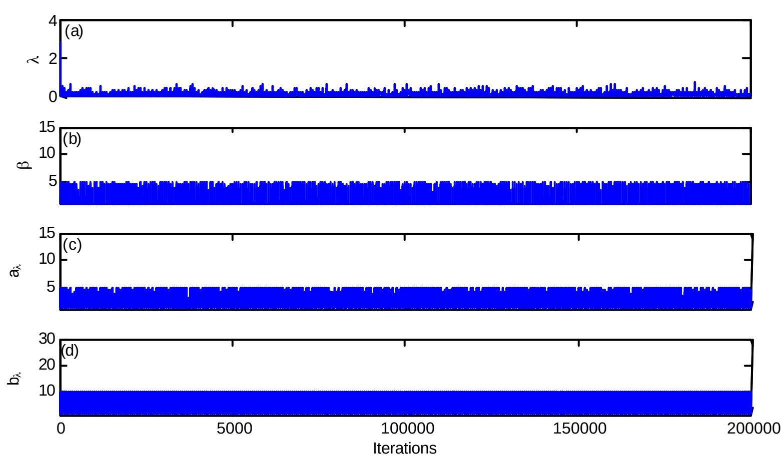

Fig. 1: Markov chain of paramteres λ, β, a_λ and b_λ for hybrid MCMC algorithm

Table 1: Data set of air compressor based on zero-failure data

i	1	2	3	4	5	6	7
x_i(kh)	0.11007	0.51657	0.60648	0.87669	1.1609	1.2911	1.4292
R_i	2	8	3	1	2	1	3

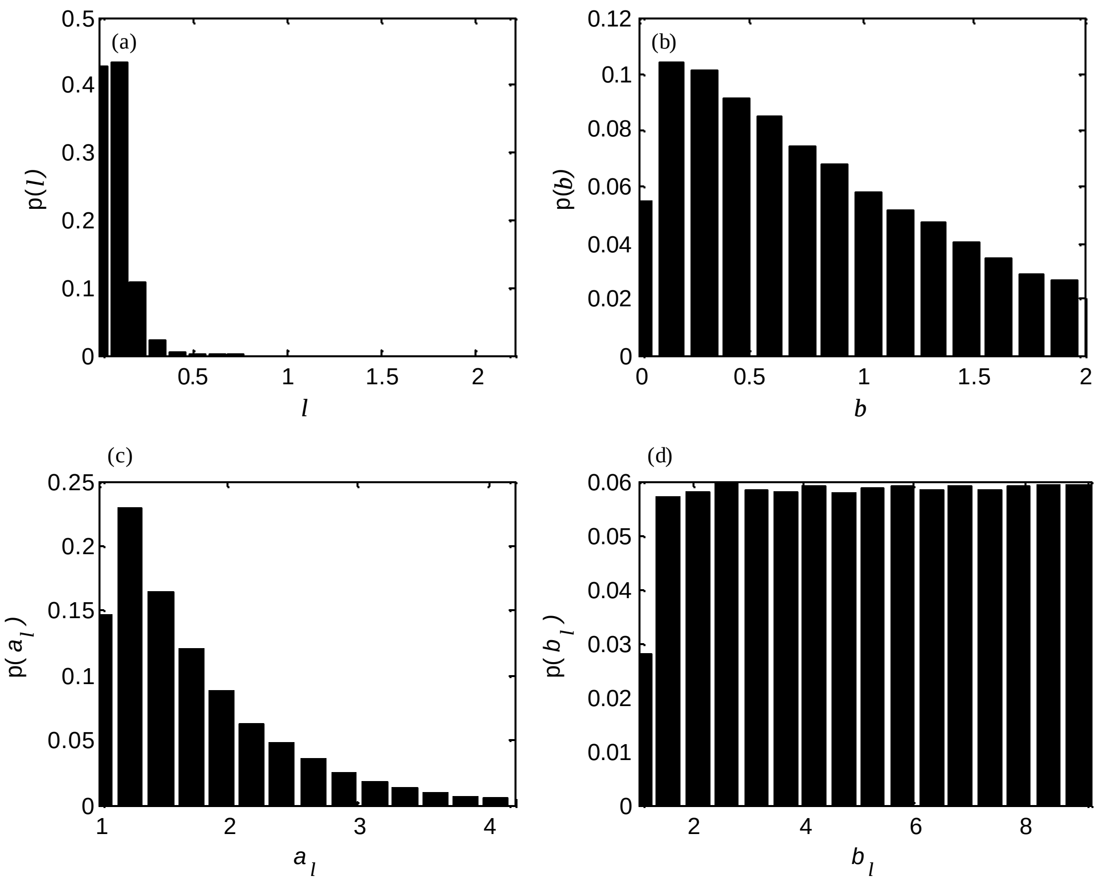

Fig. 2(a-d): Histogram of samples after "burn-in" period

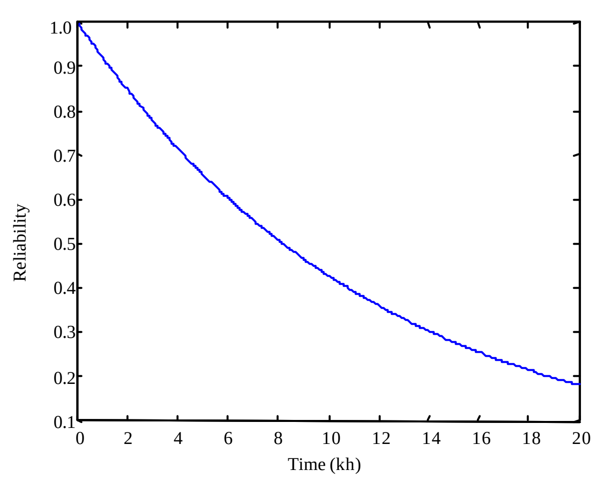

Fig. 3: Reliability curve of the air compressor based on zero-failure data

Based on the samples of Markov chain, the burn-in operation ($T_{burn-in}$=50000) is made to delete these samples which do not converge to the target distribution. Figure 2 shows all samples after "burn-in" period.

Using Monte Carlo approach, the estimator of each parameter is obtained as (λ^*, β^*, a_λ^*, b_λ^*) = (0.0815, 1.0170, 1.7449, 5.5260). Based on parameter estimators, the reliability curve of air compressor of braking system of high speed railway vehicle can be achieved as Fig. 3.

CONCLUSION

From the example above, we can conclude that Bayesian m ethod is able to assess the reliability of parts under small sample size and zero-failure data where it is difficult to analyze in classical method such as maximum likelihood method. Meanwhile, due to the high acceptance rate from Fig. 1, the hybrid MCMC algorithm can be used to achieve posterior estimators easily.

ACKNOWLEDGMENT

This study is supported by National Natural Science Foundation of China, under the Grant 51175028.

REFERENCES

Akama, M., 2002. Bayesian analysis for the results of fatigue test using full-scale models to obtain the accurate failure probabilities of the Shinkansen vehicle axle. Reliab. Eng. Syst. Saf., 75: 321-332.

Geman, S. and D. Geman, 1984. Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. IEEE Trans. Pattern Anal. Mach. Intell., 6: 721-741.

Guida, M. and G. Pulcini, 2002. Automotive reliability inference based on past data and technical knowledge. Reliab. Eng. Syst. Saf., 76: 129-137.

Jin, G., 2010. Performance reliability modeling and estimation for space bearing under small sample circumstance. J. Natl. Univ. Defense Technol., 32: 133-137.

Lord, D. and L.F. Miranda-Moreno, 2008. Effects of low sample mean values and small sample size on the estimation of the fixed dispersion parameter of Poisson-gamma models for modeling motor vehicle crashes: A Bayesian perspective. Saf. Sci., 46: 751-770.

Metropolis, N., A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller and A.E. Teller, 1953. Equation of state calculations by fast computing machines. J. Chem. Phys., 21: 1087-1091.

Tian, Y., 2008. Reliability analysis application for structure of beam of C70 railway vehicle under extreme small sample size. Master Thesis, Beijing Jiaotong University, Beijing, China.

Wang, L.Z., Y.G. Xu and J.D. Zhang, 2008. Research on reliability analysis model for key components and parts of railway equipment and its application. J. China Railway Soc., 30: 93-97.

Wang, X.Y. and M.L. Wu, 2010. On the reliability of unit brake for urban rail vehicle. Urban Mass Transit, 13: 52-53, 73.

Weaver, B.P. and M.S. Hamada, 2008. A Bayesian approach to the analysis of industrial experiments: An illustration with binomial count data. Qual. Eng., 20: 269-280.

Wu, M.L., X.Y. Wang and C. Tian, 2009. Reliability of relay valve of brake system for rail vehicles. J. Southwest Jiaotong Univ., 44: 365-369.