International Journal of Advances in Applied Mathematics and Mechanics
Volume 1, Issue 1 : (2013) pp. 35-46
Available online at www.ijaamm.com
ISSN: 2347-2529

Bayesian analysis of software reliability models with reference prior

Ashkan Ertefaie [a, *] and M. Reza Meshkani [b]

[a] *Department of Mathematics, Stuttgart University, Stuttgart, Germany*
[b] *Department of Mathematical science, Shahid Beheshti University, Tehran, Iran*

Received 09 August 2013; Accepted (in revised version) 10 September 2013

ABSTRACT

In this paper, we introduce a Bayesian analysis for non-homogeneous Poisson process in software reliability models. Posterior summaries of interest are obtained using Markov chain Monte Carlo methods. We compare the results obtained from using conjugate and reference priors. Model selection based on the prequential conditional predictive ordinates is developed.

Keywords: Gibs sampling, Metropolis algorithm, Model selection, non-homogeneous poisson process, Reference prior.

MSC 2010 codes: 62A01, 62J12.

1 Introduction

Numerous reliability growth models have been proposed for modeling the failure times in software testing, including Duane (1964), Jelinski and Moranda (1972), and Musa and Okumoto (1984). This article presents a Bayesian inference and comparison between the results obtained from using conjugate and reference priors.

Let $M(t)$ be the cumulative number of failures of the software that are observed during time $(0,t]$. Usually, $M(t)$ is modeled by a non-homogeneous poisson process (NHPP) with mean value function $m(t)$.

When we assume that there is an unknown number of faults N at the beginning of software testing and we model the observed epochs of failures to be the first n order statistics taken from N i.i.d observations with p.d.f f and c.d.f F (a general order statistic model), we can consider the mean value function given by

* Corresponding author
E-mail address: ertefaie@mathematik.uni-stuttgart.de (Ashkan Ertefaie), rmeshkani@hotmail.com (M. Reza Meshkani)

$$m(t) = \theta F(t) \quad (1)$$

where in addition , N is assumed to be distributed as a Poisson variable with mean θ . If f is an exponential density, then this is the Jelinski-Moranda model. We can consider other models using different *f*'s, such as Pareto and Weibull. Observe that in these cases, $\lim_{t\to\infty} m(t)$ is finite and this process can be denoted by NHPP-I. In situations where the new faults can occur during testing period, we need to replace the general order statistic (GOS) model with the record value statistic (RVS) model. In the latter case we have $m(t) \to \infty$ as $t \to \infty$, and we denote it by NHPP-II (see Kuo, 1996). A special case (Musa and Okumoto (1984), Duane (1964)) is given by

$$m(t) = -\ln(1 - F(t)). \quad (2)$$

In both models, we assume that $m(t)$ is differentiable, and $\lambda(t) = \partial m(t) / \partial t$ denotes the intensity function or the rate of occurrence of failures (ROCOF) for the NHPP. Conversely, one can show that failure epochs of a NHPP-II with ROCOF $\lambda(t)$ obey the RVS model derived from the i.i.d unobserved outcomes with the density $f(t) = \lambda(t)\exp(\int_0^t -\lambda(u)du)$. In particular, the set of observed failure epochs of the Duane process is the counting process of the RVS from the Weibull distribution which is called Weibull NHPP-II.

For a Bayesian inference, we can choose conjugate or reference priors for the parameters. In this paper, we focus on using reference priors because in most practical cases we may not have enough evidence to elicit a specific prior or we are not willing to use priors subjectively. In these cases, reference priors can be useful because they are obtained from the observed data. To compute Bayes estimates, Gibbs sampling is proposed for evaluating the features of the posterior distribution. Gibbs sampling is a Markov Chain Monte Carlo (MCMC) that helps us generate random samples according to a Markov chain with the stationary distribution as the desired posterior distribution. When the conditional densities are not easily identified the Metropolis (1953) algorithm may be used.

Since many models have been proposed in software reliability, it is natural to ask which model should be used. The criteria of the prequential conditional predictive ordinates are adapted here for model selection. Section 2 develops the reference prior briefly and clarifies some theorems in this matter. Section 3 discusses the likelihood of the NHPP process. Section 4 exhibits the Gibbs algorithm and a criterion for model selection. Section 5 presents some concluding remarks.

2 Reference prior

Under Bayesian paradigm, the outcome of any inference problem (the posterior distribution of he quality of interest) combines the information provided by the data with relevant available prior information. In many situations, either the available prior information on the quantity of interest is too vague or it is too subjective to be useful in scientific communication. It is therefore important to be able to identify the mathematical form of a "non-informative" prior, a prior that would have a minimal effect, relative to data, on the posterior inference. Notice that the concept of a "non-informative" prior is related to the information provided by the data.

Bernardo (1979) initiated the reference prior approach to development of noninformative priors, following in the tradition of Laplace and Jeffreys.
It has been found that the behavior of many important limiting processes, in both probability theory and statistical inference, is better described in terms of information-theory, related divergence measure, the intrinsic discrepancy (Bernardo and Rueda, 2002), which is now defined.

2.1 Definition: Intrinsic Discrepancy

The intrinsic discrepancy $\delta\{p_1, p_2\}$ between two probability distributions of a random vector $x \in \mathrm{X}$, specified by their density functions $p_1(x), x \in \mathrm{X}_1 \subset \mathrm{X},$ and $p_2(x), x \in \mathrm{X}_2 \subset \mathrm{X},$ with either identical or nested supports, is

$$\delta\{\mathrm{p}_1, \mathrm{p}_2\} = \min\left\{\int_{\mathrm{X}_1} \mathrm{p}_1(x)\log\frac{\mathrm{p}_1(x)}{\mathrm{p}_2(x)}dx, \quad \int_{\mathrm{X}_2} \mathrm{p}_2(x)\log\frac{\mathrm{p}_2(x)}{\mathrm{p}_1(x)}dx\right\}.$$

provided one of the integrals (or sums) is finite.

The posterior distribution of the parameter is often asymptotically normal [see e.g., Bernardo and Smith (1994, Sec. 5.3)]. In this case, the reference prior is easily obtained in terms of Fisher information matrix.

2.2 Definition: Permissible prior function

A positive function $\pi(w)$ is a permissible prior function for model $M \equiv \{p(x\,|\,\omega), x \in \mathrm{X}, \omega \in \Omega\}$ if for all $x \in X$ one has $\int_{\Omega} p(x\,|\,\omega)\pi(\omega)d\omega < \infty$, and for some increasing sequence $\{\Omega_i\}_{i=1}^{\infty}$ of subsets of Ω such that $\lim_{i\to\infty}\Omega_i = \Omega$, and $\int_{\Omega_i} \pi(\omega)d\omega < \infty$

$$\lim_{i\to\infty}\int_{\mathrm{X}} p_i(x)\delta\{\pi_i(\omega\,|\,x), \pi(\omega\,|\,x)\}dx = 0$$

where $\pi_i(\omega)$ is the renormalized restriction of $\pi(\omega)$ to Ω_i, $\pi_i(\omega\,|\,x)$ is the corresponding posterior, $p_i(x) = \int_{\Omega_i} p(x\,|\,\omega)\pi_i(\omega)d\omega$ is the corresponding predictive, and $\pi(\omega\,|\,x) \propto p(x\,|\,\omega)\pi(\omega).$

2.3 Theorem: Reference prior under asymptotic normality

Let data consist of a random sample from model $M \equiv \{p(y\,|\,\theta,\lambda), y \in Y, \theta \in \Theta, \lambda \in \Lambda\}$, and let P be the class of all continuous(joint) priors with support $\Theta * \Lambda$. If the posterior distribution of (θ,λ) is asymptotically normal with dispersion matrix $V(\hat{\theta}_n,\hat{\lambda}_n)/n$, wehere $(\hat{\theta}_n,\hat{\lambda}_n)$ is a consistent estimator of (θ,λ), define

$$V(\theta,\lambda)=\begin{pmatrix} v_{\theta\theta}(\theta,\lambda) & v_{\theta\lambda}(\theta,\lambda) \\ v_{\theta\lambda}(\theta,\lambda) & v_{\lambda\lambda}(\theta,\lambda)\end{pmatrix},\ \ H(\theta,\lambda)=V^{-1}(\theta,\lambda), \text{and}$$

$$\pi(\lambda\,|\,\theta)\propto h_{\lambda\lambda}^{1/2}(\theta,\lambda),\ \ \lambda\in\Lambda,$$

and, if $\pi(\lambda\,|\,\theta)$ is proper,

$$\pi(\theta)\propto \exp\left\{\int \pi(\lambda\,|\,\theta)\log[v_{\theta\theta}^{-1/2}(\theta,\lambda)]d\lambda\right\},\ \theta\in\Theta \tag{3}$$

then, if $\pi(\lambda\,|\,\theta)\pi(\theta)$ is a permissible prior function, the θ-reference prior is

$$\pi(\theta\,|\,M^n,P)\propto\pi(\lambda\,|\,\theta)\pi(\theta). \tag{4}$$

if $\pi(\lambda\,|\,\theta)$ is not proper, integration is performed on elements of an increasing sequences $\{\Lambda_i\}_{i=1}^{\infty}$ such that $\int_{\Lambda_i}\pi(\lambda\,|\,\theta)d\lambda<\infty$, to obtain the sequence $\{\pi_i(\lambda\,|\,\theta)\pi_i(\theta)\}_{i=1}^{\infty}$, where $\pi_i(\lambda\,|\,\theta)$ is the renormalization of $\pi(\lambda\,|\,\theta)$ to Λ_i , and the θ -reference prior $\pi^{\theta}(\theta\lambda)$ is defined as its corresponding intrinsic limit.

Proof. Bernardo (2005) □

Notice that under appropriate regularity conditions [see e.g. Bernardo and Smith (1994, Sec. 5.3)] the joint posterior distribution of $\{\theta,\lambda\}$ is asymptotically normal with precision matrix $nI(\hat{\theta}_n,\hat{\lambda}_n)$ where $I(\theta)$ is Fisher information matrix; in that case, the asymptotic dispersion matrix in theorem 2.3 is simply $V(\theta,\lambda)=I^{-1}(\theta,\lambda)/n$.

2.4 Theorem: Reference prior under factorization

In the conditions of Theorem 2.3, if (i) θ and λ are variation independent so that Λ does not depend on θ and (ii) both $h_{\lambda\lambda}(\theta,\lambda)$ and $v_{\theta\theta}(\theta,\lambda)$ factorize so that

$$v_{\theta\theta}^{-1/2}(\theta,\lambda)\propto f_{\theta}(\theta)g_{\theta}(\lambda),\quad h_{\lambda\lambda}^{1/2}(\theta,\lambda)\propto f_{\lambda}(\theta)g_{\lambda}(\lambda), \tag{5}$$

then the $\theta-$ reference prior is simply $\pi^{\theta}(\theta,\lambda)=f_{\theta}(\theta)g_{\lambda}(\lambda)$, even if the conditional reference prior $\pi(\lambda\,|\,\theta)=\pi(\lambda)\propto g_{\lambda}(\lambda)$ is improper.

Proof. Bernardo (2005) □

We use this two theorems to develop reference prior for Weibull NHPP-II. For further details about reference priors, see Bernordo (2005, 1997, 1979), Berger and Bernordo (1992).

3 Models

Here, we obtain the likelihood function of a NHPP under two assumptions, i.e., time truncation and failure truncation. We assume that $m(t)$ is indexed by the unknown positive real parameters θ and β , where sometimes β can be a vector. For convenience, we write $m(t)$ for $m(t\,|\,\theta,\beta)$.

Given the time truncated model, i.e. testing until some pre-specified time *t*, the ordered epochs of the observed *n* failure times are denoted by $x_1,...,x_n$. Therefore, the data set D_t consists of $\{n, x_1,...,x_n; t\}$. Given the failure truncated model, i.e. observing the process until some pre-specified nth failure, the data set D_{x_n} consists of $\{x_1,...,x_n\}$.

The likelihood function for the failure truncated model, where the ordered epochs of the observed n failures are denoted by $x_1,...,x_n$, [see for example Cox and Lewis (1966) or Kuo and Yang(1996)] is given by

$$L_{NHPP}(\theta,\beta\,|\,D)=\left\{\prod_{i=1}^{n}\lambda(x_i)\right\}\exp\{-m(x_n)\}. \tag{6}$$

For time truncated model this expression can be applied with x_n replaced by *t*.
In the following, we show general form of the likelihood functions for the NHPP-I and NHPP-II.

3.1 Likelihood functions

We consider two types of NHPP process according to the limiting behavior of $m(t)$: NHPP-I if $\lim_{t\to\infty} m(t)<\infty$, and NHPP-II, if $m(t)\to\infty$ as $t\to\infty$.

Suppose we use GOS model for our data and that *N* has a Poisson distribution with mean θ . It can be shown that $M(t)$ is a NHPP-I with mean function $m(t)=\theta F(t\,|\,\theta)$ (Kuo and Yang (1994); Musa et al. (1987). In this case, for the time truncated model in NHPP-I with $\lim_{t\to\infty} m(t)=\theta$ the likelihood function is

$$L_{NHPP-I}(\theta,\beta\,|\,N,D)=\frac{N!}{(N-n)!}\left\{\prod_{i=1}^{n}\frac{\lambda(x_i)}{\theta}\right\}\left\{1-\frac{m(t)}{\theta}\right\}^{N-n}. \tag{7}$$

The Jelinski-Moranda model is a special case of the GOS model with $f(t)=\beta\exp(-\beta t)$, and the corresponding point process that counts the number of failures in $(0,t]$ is a NHPP-I with mean function $\theta[1-\exp(-\beta t)]$. This NHPP was proposed by Goel and Okumoto (1979) to model software reliability. We can also use Weibull instead of exponential distribution. Note that the Yamada et al. (1983) S-shape software reliability models are contained in the NHPP-I with *f* as a gamma density with shape parameter equal to 2.

As we mentioned before, GOS models cannot be applied to situations where new faults might occur in the process of debugging. To extend the NHPP-I to the NHPP-II, where $m(t) \to \infty$ as $t \to \infty$, we need to replace GOS model with a record value statistics model (RVS).

In the RVS model, we assume that the epochs of failures are the record values of iid random variables distributed according to a density $f(S \mid \beta)$ with distribution function $F(S \mid \beta)$. Let $S_1, S_2, \ldots$ denote the iid random variables distributed according to $f(S \mid \beta)$. We denote the sequence of record values, $\{X_n\}_{n\geq 1}$ by

$$R_1 = 1,$$
$$R_{k+1} = \min\{i : S_i > S_{R_k}\}, \text{ where } k = 1, 2, 3, \ldots,$$

and

$$Set\ X_n = S_{R_n},\ \ for\ \ n \geq 1.$$

We model the observed epochs of failure $x_1 < x_2 \ldots < x_n$ as the record values $X_1, X_2, \ldots$.

Resnic (1987) showed that if F is continuous with support R^+, then the record values in $(0, t]$ constructed above from F are the points of a NHPP-II in $(0, t]$ with mean measure $m(t) = -\ln(1 - F(t))$, and the ROCOF in NHPP-II is $\lambda(t) = m'(t) = \dfrac{f(t)}{1 - F(t)}$.

Now by substituting for $\lambda(t)$ and $m(t)$ in (6), the likelihood of NHPP-II is obtained as

$$L_{NHPP-II}(\beta \mid N, D) = \left\{ \prod_{i=1}^{n} \frac{f(x_i \mid \beta)}{1 - F(x_i \mid \beta)} \right\} \{1 - F(x_n \mid \beta)\}. \tag{8}$$

For the time truncated model, the likelihood is similar to (8) with x_n replaced by *t*.

4 Bayesian inference

In this section we describe the Bayes estimators. In addition to making inference on unknown parameters, we also present two criteria for model selection. We divide our discussion into three subsections, one for Gibbs algorithm, one for using conjugate and reference prior for Weibull NHPP-II, and the other for Model selection.

4.1 Gibbs sampling

For inference on the parameters, we can use the Gibbs sampler, i.e. the mixture of the conditional densities of the parameters of interest given the rest of parameters. In order to sample from a posterior distribution, we need to use conditional distribution of each parameter. In this section we develop the conditional distribution used by Gibbs algorithm. By using this algorithm we drive samples from the posterior distribution. There are some methods for testing the convergence of the algorithm, such as replicating this algorithm with independent starting points. For further information on convergence see Casella and George

(1992), Gelman et al. (2004) and Gelfand and Smith (1990). When the conditional densities are not easily identified, such as cases without conjugate priors, the Metropolis (1953) algorithm can be applied [see Givens and Hoeting (2005)]. In this section we use time truncated case it suffices to change x_n to *t*.

4.2 Using conjugate and reference priors for Weibull NHPP-II

We discuss Weibull distribution for several reasons. First, it is probably the most widely used distribution for lifetimes. Second, the Weibull distribution is related to the power law process, a commonly used model for repairable systems. Third, if repairs bring a system back to a good-as-new sate, then the assumption that the times between failures $X_1,...,X_n$ are iid Weibull random variables may be reasonable. The pdf of Weibull is given as:

$$f(x) = \alpha\beta x^{\alpha-1} e^{(-\beta x^{\alpha})}. \tag{9}$$

Now by inserting (9) in (8) the likelihood is

$$L_{NHPP-II}(\alpha,\beta \mid D) = \left(\prod_{i=1}^{n} \alpha\beta x_i^{\alpha-1}\right) e^{-\beta x_n^{\alpha}}, \tag{10}$$

Conjugate priors for the parameters are

$$\beta \sim \Gamma(c,d)\ ;\ \alpha \sim \Gamma(a,b), \tag{11}$$

and conditional posterior densities are

$$\left.\begin{aligned} &\beta \mid \alpha, D \sim \Gamma(n+c, d+x_n^{\alpha}),\\ &\alpha \mid \beta, D \propto \alpha^{a+n-1}\Big(\prod_{i=1}^{n} x_i^{\alpha-1}\Big) e^{-(\beta x_n^{\alpha}+b\alpha)}, \end{aligned}\right\} \tag{12}$$

where α can be generated by Metropolis algorithm.

To obtain the reference prior for Weibull NHPP-II, we use theorem (2.4). The Fisher information for Weibull distribution is

$$I(\alpha,\beta) = \begin{pmatrix} \dfrac{1}{\alpha^2} + \dfrac{\varphi_1(\beta)}{\alpha^2} & \varphi_2(\alpha,\beta) \\ \varphi_2(\alpha,\beta) & \dfrac{1}{\beta^2} \end{pmatrix} \tag{13}$$

and

$$V(\alpha,\beta) = I^{-1}(\alpha,\beta) = \begin{pmatrix} \dfrac{6\alpha^2}{\pi^2} & \varphi_3(\alpha,\beta) \\ \varphi_3(\alpha,\beta) & \varphi_4(\beta) \end{pmatrix}, \tag{14}$$

where φ_1 and φ_4 are functions of β, and φ_2 and φ_4 are functions of α and β. With regard to (5) we have:

$$\left.\begin{aligned} v_{\alpha\alpha}^{-1/2}(\alpha,\beta) &\propto \frac{1}{\alpha} = f_\alpha(\alpha)g_\alpha(\beta) \\ h_{\beta\beta}^{1/2}(\alpha,\beta) &\propto \frac{1}{\beta} = f_\beta(\alpha)g_\beta(\beta) \end{aligned}\right\} \tag{15}$$

and using theorem(4) the reference prior for Weibull distribution is

$$\pi^\alpha(\alpha,\beta) = f_\alpha(\alpha)g_\beta(\beta) \propto \frac{1}{\alpha\beta} \tag{16}$$

Thus, the posterior and conditional densities are

$$\left.\begin{aligned} &\pi(\alpha,\beta \mid D) \propto \frac{1}{\alpha\beta}\alpha^n\beta^n \prod_{i=1}^n x_i^{\alpha-1} e^{(-\beta x_n^\alpha)}, \\ &\beta \mid \alpha, D \sim \Gamma(n, x_n^\alpha), \\ &\alpha \mid \beta, D \propto \alpha^{n-1} \prod_{i=1}^n x_i^{\alpha-1} e^{(-\beta x_n^\alpha)}. \end{aligned}\right\} \tag{17}$$

These results are used to compute the various moments of $(\alpha \mid D)$ and $(\beta \mid D)$ by MCMC methods.

4.3 Model selection

There are several models in software reliability. Which model or models are appropriate for a given situation? We address this issue using the prequential likelihood of the conditional predictive ordinates. The prequential conditional predictive ordinates (PCPO) for the future inter failure time t_{n+1} is defined to be $c_{n+1} = p(t_{n+1} \mid D_{t_n})$ for $i \geq 2$. For $i = 1$, we define $c_1 = p(t_1 \mid prior)$. The conditional density of T_{n+1} evaluated at the future observed time given $(t_1,\ldots,t_n)$. The PCPO is particularly relevant for model selection, because it assesses the predictability of T_{n+1} given the past data. the PCPO c_{n+1} can be computed by

$$p(t_{n+1} \mid D_{t_n}) = \int\int p(t_{n+1} \mid \alpha,\beta,D_{t_n})\, p(\alpha,\beta \mid D_{t_n})\, d\alpha d\beta \tag{18}$$

Equation (18) can be computed from Gibbs sampler. We can plot c_i versus i, $i = 1,\ldots,n$, for each model. The bigger the c are on the average, the better the model.

5 Numerical illustration

Numerical illustration based on a real-life data set is given here. The data set of 26 inter failure times originating with Jelinski and Moranda [cf. Goel and Okumoto (1979)] is listed in Table 1. Goel and Okumoto (1979) and Mazzuchi and Suyer (1988) and Kuo and Yang (1995) also have used data for illustration. The data were based on the trouble report for one of the larger modules of the Naval Tactical Data System. The times (days) between software failures found during the production phases are listed in Table 1. We apply Weibull models using conjugate and reference priors. We monitored the convergence of the Gibbs sampler using the Gelman and Rubin (1992) method that uses the analysis of variance technique to determine whether further iterations are needed. We found 5000 iteration to be sufficient. We use $\alpha \sim \Gamma(5,3)$ and $\beta \sim \Gamma(7,2)$ for conjugate priors. We list the estimates of α and β with respect to the squared error and LINEX loss functions in Table 2. The differences between conjugate and reference prior PCPO's are shown in Figure 1, which shows that when we increase the sample size to compute the PCPO's, conjugate prior works better.

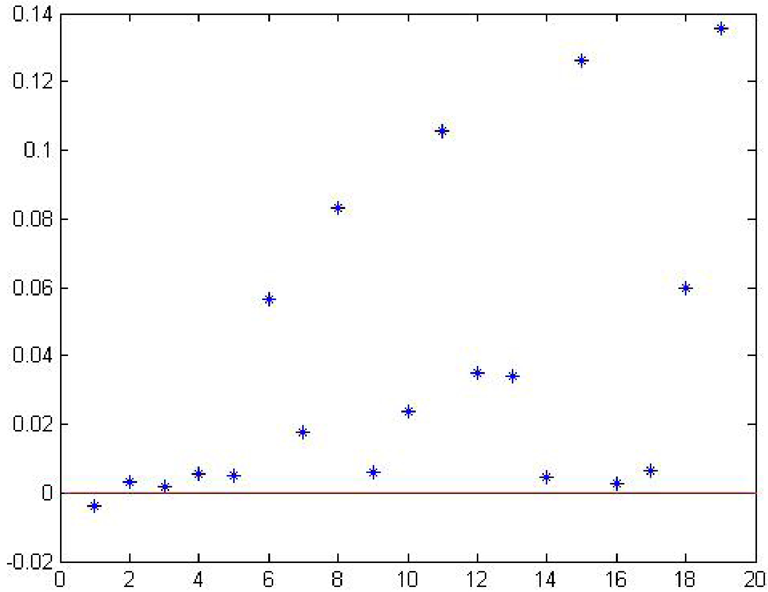

Figure 1: The differences of the PCPOs. [PCPO(c)-PCPO(r)]

5.1 Posterior risk

Posterior risk is one of the criteria for estimators in Bayesian analysis. Posterior risks of Weibuell process with conjugate and reference prior with respect to square and LINEX loss functions are shown in Table 3.

6 Conclusion

Software reliability growth models are used in developing and testing stages of software development to monitor the error content of the software. We use Weibull NHPP-II to describe the failure times which is a kind of RVS model under some assumptions. We introduce the reference priors and present Bayesian inference with LINEX and squared error loss functions using MCMC methods. Besides the Bayesian inference, we also consider model selection which plays a crucial role in choosing the "best" model or models. As pointed out by Singpurwalla and Wilson (1994), statistical inference for software reliability can be considered as a means to a goal; that is, to assist software engineers and managers to make informed decisions on when to release the software.

7 Tables

Table 1: Software failure data

Error No	Inter failure time	Error No	Inter failure time	Error No	Inter failure time
1	9	10	7	19	9
2	12	11	1	20	1
3	11	12	5	21	11
4	4	13	1	22	33
5	7	14	9	23	7
6	2	15	4	24	91
7	5	16	1	25	2
8	8	17	3	26	1
9	5	18	3		

Table 2: Bayesian Inference for α and β

	Square Loss	LINEX loss (a=1)	LINEX loss (a=-1)
α_c	0.4226	0.4203	0.4248
α_r	0.2528	0.2514	0.2541
β_c	1.7286	1.6488	1.8203
β_r	6.8417	5.2284	10.5704

α_c and β_c are the estimates using conjugate prior and α_r and β_r using reference prior.

Table 3: Posterior risks of α and β in Weibull process

	Square Loss	LINEX loss (a=1)	LINEX loss (a=-1)
α_c	0.0027	0.0014	0.0013
α_r	0.0021	0.0010	0.0011
β_c	0.1696	0.0795	0.1061
β_r	4.6313	1.6222	3.7091

α_c and β_c are the estimates using conjugate prior and α_r and β_r using reference prior.

References

Berger, J. O. and Bernardo, J. M. (1992): On the Development of Reference Priors, Bayesian Statistics 4, Oxford University Press, pp. 35-60.

Bernardo, J. M. (1979): Expected Information as Expected Utility, Ann. Statist. 7, pp.686-690.

Bernardo, J. M. (1997): Non-Informative Priors do not Exist, J. Statist. Planning and Inference 65, pp. 159-189.

Bernardo, J. M. (2005): Reference Analysis, Handbook of Statistics 25, pp. 17-90.

Bernardo, J. M. and Rueda, R. (2002): Bayesian hypothesis testing: A reference approach. Internat. Statist. Rev. 70, pp. 351-372.

Casella, G. and George, E. I. (1992): Explaining the Gibbs Sampler,The American Statistician 46, pp.167-174.

Cox, D. R., and Lewis, P. A. (1966): Statistical Analysis of Series of Events,London, Methuen.

Duane, J.T. (1964),Learning Curve Approach to Reliability Monitoring, IEEE Transactions on Aerospace AS-2, pp. 563-566.

Gelfand, A. E., Smith, A. F. M. (1990): Sampling-Based Approaches to calculating Marginal Densities, Journal of the American Statistical Association 85, pp. 398-409.

Gelman, A. E. and Rubin, D. (1992): Inference from Iterative Simulation Using Multiple Sequences, Statistical Science 7, pp. 457-472.

Givens, G. H. and Hoeting, I. A. (2005): Computational Statistics, John Wiley, New Jersey.

Goel, A. L. and Okumoto, K. (1979): Time Dependent Error Detection Rate Model for Software Reliability and Other Performance Measures, IEEE Transaction on Reliability 28, pp. 206-211.

Jelinski, Z. and Moranda, P. B. (1972): Software Reliability Research, in Statistical Computer Performance Evaluation,ed.W. Freiberger, New York: Academic Press, pp. 465-497.

Mazzuchi, T. A. and Soyer, R. (1988): A Bayes Empirical Bayes Model for Software Reliability, IEEE Transaction on Reliability37, pp. 248-254.

Musa, J. D., Iannino, A., and Okumoto, K. (1987): Software Reliability Measurement, Prediction, Application, McGrow-Hill, New York.

Musa, J. D. and Okumoto, K. (1984): A Logarithmic Poisson Execution Time Model for Software Reliability Measurement, in Proceeding of the Seventh International Conference on Software Engineering, pp. 230-238.

Raftery, A. E. (1987): Inference and Prediction for a General Order Statistic Model with unknown population Size, Journal of the American Statistical Association 82, pp. 1163-1168.

Resnic, S. I. (1987): Extreme Values, Regular Variation, and Point Process, Springer-Verlag, Berlin.

Singpurwalla, N. D. and Wilson, S. P. (1994): Software Reliability Modeling, International Statistical Review 62, pp. 289-317.

Kuo, T., and Yang, T. (1994): Bayesian Computation for Nonhomogeneous Poisson Processes in Software Reliability, Technical Report, University of Connecticut, Statistics Department.

Kuo, T., and Yang, T. (1995): Bayesian Computation of Software Reliability, Journal of Computational and Graphical Statistics 4, pp. 65-82.

Kuo, T., and Yang, T. (1996): Bayesian Computation for Nonhomogeneous Poisson Processes in software Reliability, Journal of the American Statistical Association .91, pp. 763-773.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E. (1953): Equation of State Calculations by Fast Computing Machines, Journal of Chemical Physics, 21, pp. 1087-1092.

Yamada, S., Ohba, M. and Osaki, S. (1983): S-Shaped Reliability Growth Modeling for Software Error Detection, IEEE Transaction and Reliability 32, pp. 475-478.