Journal of the Korean Statistical Society (2005), **34**: 2, pp 85–98

A BAYESIAN ANALYSIS FOR PRODUCT OF POWERS OF POISSON RATES

HEA-JUNG KIM [1]

ABSTRACT

A Bayesian analysis for the product of different powers of k independent Poisson rates, written θ, is developed. This is done by considering a prior for θ that satisfies the differential equation due to Tibshirani and induces a proper posterior distribution. The Gibbs sampling procedure utilizing the rejection method is suggested for the posterior inference of θ. The procedure is straightforward to specify distributionally and to implement computationally, with output readily adapted for required inference summaries. A salient feature of the procedure is that it provides a unified method for inferencing θ with any type of powers, and hence it solves all the existing problems (in inferencing θ) simultaneously in a completely satisfactory way, at least within the Bayesian framework. In two examples, practical applications of the procedure is described.

AMS 2000 subject classifications. Primary 62F15; Secondary 62F25.
Keywords. Bayesian analysis, Product of powers of Poisson rates, Tibshirani's differential equation, the Gibbs sampling, the rejection method.

1. INTRODUCTION

Suppose that $X_i \sim P_o(\lambda_i)$, for $i = 1, \ldots, k$ are independent Poisson random variables with parameter λ_i. The parameter of interest is $\theta = \prod_{i=1}^k \lambda_i^{\alpha_i}$, the product of different powers of k Poisson rates. The estimate of θ has been applied in many statistical problems. An example is in the assessment of casualties due to traffic accidents. Assume that the casualties per accident, number of accidents per day, and number of days per week during which traffic accident is occurred are three independent Poisson random variables. The total casualties in a week is the product of three Poisson rates. When $\alpha_i = 1/k$ for all i, estimating θ can be

Received August 2004; accepted May 2005.
[1] Department of Statistics, Dongguk University, Pil-dong, 3-Ga, Chung-Gu, Seoul 100-715, Korea

recognized as the geometric mean of Poisson rates that arises in environmental applications (Environmental Protection Agency, 1986) and in economic applications (Kenneth et al., 1998). For $\alpha_i = 1$ and $\alpha_{i'} = -1$ for $i = 1, \ldots, k_1$ and $i' = k_1 + 1, \ldots, k$, θ can be viewed as a quotient of the Poisson rates that arises in the problem of exponential series system availability (Martz and Waller 1982). A couple of literatures are concerned with mathematical devices for the inference of θ, Harris (1971) and Harris and Soms (1973) derived conditional distributions that depends only on the quotient and the powers of poisson rates, respectively. Based on the distributions they suggested procedures for testing and obtaining confidence intervals for θ.

Even though results of the preceding literatures are concerned with mathematical devices for the inference of θ, each of them is confined to the case where $\{\alpha_i; i = 1, \ldots, k\}$ has a special set of integer values. Furthermore, the mathematical devices are not unified so that they vary according to integer values of α_i. A second problem of the results arises from the discreteness of the distributions involved in the mathematical devices. For any confidence level $1 - \gamma$, it is only possible to give intervals with probability of coverage exactly equal to $1 - \gamma$ by artificial randomization (cf., e.g., Lehmann 1959 and Harris and Soms 1973). A third problem is that the mathematical devices can not be extended to the case when values of α_i are not integer. This fact prevents us from developing applications involving various types of θ.

It is possible to deal simultaneously with all these problems in a completely satisfactory way, at least within the Bayesian framework. This article proposes a unified method for the inference of θ for any real values of α_i. As an alternative to Jeffreys' prior, Section 2 develops a noninformative prior for θ utilizing the differential equation by Tibshirani (1989) that asymptotically matches posterior and frequentist probabilities based on an Edgeworth expansion of a function of likelihood ratio statistic. Section 3 provides the Gibbs sampling procedure for inference of θ based on the noninformative prior leading to a proper posterior. In Section 4 posterior inferences of θ obtained from the prior and the Jeffreys' prior are compared in terms of frequentist coverage probabilities of the posterior credible sets. The calculations in both Sections 3 and 4 are made by a Markov chain Monte Carlo method, which is very successful in handling small to large k. Section 5 contains two examples which have been selected for the purpose of motivating the contents of this article.

2. THE PRIOR

Suppose we observe X_i, $i = 1, \dots, k$, as independent Poisson variables with parameter λ_i, the parameter of interest being $\theta = \prod_{i=1}^{k} \lambda_i^{\alpha_i}$, the product of different powers of k Poisson rates. Given a parameter vector $\lambda = (\lambda_1, \dots, \lambda_k)'$, we seek a noninformative prior $\pi_0(\lambda)$ so that the posterior interval for θ has asymptotically good coverage probabilty in the frequentist sense. Reasons for seeking this kind of prior is described in Stein (1985) and he derived nonrigorously a sufficient condition for an asymptotically optimal frequentist coverage prior. Through the use of orthogonal parameters, Tibshirani (1989) gave a differential equation that yields general form of the class of priors satisfying Stein's condition. In this section, as an alternative to the Jeffreys' prior, we obtain a noninformative prior of θ by deriving and solving the Tibshirani's differential equation.

Let $\theta = \prod_{i=1}^{k} \lambda_i^{\alpha_i}$ and $\zeta_i = \zeta_i(\lambda)$ for $i = 2, \dots, k$ and let $\zeta_i^j = \partial\zeta_i(\lambda)/\partial\lambda_j$ and $\eta_{(i)} = (\alpha_i/\lambda_i) \prod_{j=1}^{k} \lambda_j^{\alpha_j}$. The Jacobian matrix of this transformation is

$$\frac{\partial(\theta, \zeta)}{\partial(\lambda)} = \begin{bmatrix} \eta_{(1)} & \eta_{(2)} & \cdots & \eta_{(k)} \\ \zeta_2^1 & \zeta_2^2 & \cdots & \zeta_2^k \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \zeta_k^1 & \zeta_k^2 & \cdots & \zeta_k^k \end{bmatrix}.$$

Therefore, the inverse of the expected Fisher information matrix can be written as

$$I^{-1}(\theta, \zeta) = \left(\frac{\partial(\theta, \zeta)}{\partial(\lambda)}\right) I^{-1}(\lambda) \left(\frac{\partial(\theta, \zeta)}{\partial(\lambda)}\right)' = \begin{bmatrix} \sum_{i=1}^{k} \lambda_i \, \eta_{(i)}^2 & \phi' \\ \phi & A \end{bmatrix}, \tag{2.1}$$

where $\phi = (\sum_{i=1}^{k} \lambda_i \ \eta_{(i)}\zeta_2^i, \dots, \sum_{i=1}^{k} \lambda_i \ \eta_{(i)}\zeta_k^i)'$ and A is an $(k-1) \times (k-1)$ nonsingular matrix. Thus θ and ζ are orthogonal if and only if $\phi = 0$. This gives $k - 1$ homogeneous linear partial differential equation of first order. Any smooth function with form $\psi(\lambda_i/\alpha_i - \lambda_j/\alpha_j, \ i < j)$ could be a solution of the equations. For instance, one could take

$$\zeta_i = \frac{\lambda_1}{\alpha_1} - \frac{\lambda_i}{\alpha_i}, \quad i = 2, \dots, k \tag{2.2}$$

as the new *transformations*. Then θ and ζ are orthogonal and its Jacobian matrix can be written as

$$\frac{\partial(\theta,\zeta)}{\partial(\lambda)} = \begin{bmatrix} \eta_{(1)} & \eta_{(2)} & \cdots & \eta_{(k)} \\ \alpha_1^{-1} & -\alpha_2^{-1} & \cdots & 0 \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \alpha_1^{-1} & 0 & \cdots & -\alpha_k^{-1} \end{bmatrix}$$

and its absolute value of determinant is $\sum_{i=1}^k \eta_{(i)}\delta_{(i)}$, where $\delta_{(i)} = \prod_{j\neq i} \alpha_j^{-1}$. The Fisher information matrix, by (2.1), is

$$I(\theta,\zeta) = \begin{bmatrix} \sum_{i=1}^k \lambda_i\, \eta_{(i)}^2 & 0 & \cdots & 0 \\ 0 & \lambda_1/\alpha_1^2 + \lambda_2/\alpha_2^2 & \cdots & \lambda_1/\alpha_1^2 \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ 0 & \lambda_1/\alpha_1^2 & \cdots & \lambda_1/\alpha_1^2 + \lambda_k/\alpha_k^2 \end{bmatrix}^{-1} . \quad (2.3)$$

Using Tibshirani's method, a noninformative prior of form

$$\pi(\theta,\zeta) \propto g(\zeta) \left(\sum_{i=1}^k \lambda_i\, \eta_{(i)}^2 \right)^{-1/2} \quad (2.4)$$

will achieve the asymptotic optimal coverage property. Transforming back to the original parameter space λ, we obtain

$$\pi_T(\lambda) \propto g(\zeta(\lambda)) \left(\sum_{i=1}^k \eta_{(i)}\, \delta_{(i)} \right) \left(\sum_{i=1}^k \lambda_i\, \eta_{(i)}^2 \right)^{-1/2} \propto g(\zeta(\lambda)) \sqrt{\sum_{i=1}^k \lambda_i^{-1} \alpha_i^2}, \quad (2.5)$$

where $g(\zeta(\lambda)) > 0$ is arbitrary.

Up to now, we have given a procedure for deriving the prior $\pi_T(\lambda)$ for the balanced case where sample size from each of k Poisson populations are equal. Using the same procedure, we may obtain the prior for the unbalanced case where we observe X_{ij}, $i = 1, \ldots, k$, $j = 1, \ldots, n_i$, as independent Poisson variates with mean λ_i. Upon setting $\theta = \prod_{i=1}^k \lambda_i^{\alpha_i^*}$ with $\alpha_i^* \propto \sqrt{n_i}\alpha_i$ and applying the same procedure, we see that the corresponding prior for λ is exactly the same as (2.5). Therefore, the noninformative prior (2.5) for the product of powers of Poisson rates is valid for the unbalanced case as well as the balanced case. The choice $g(\zeta(\lambda)) = 1$ for (2.5) gives simple form of the prior, which also attains good frequentist coverage property compared to the Jeffreys' prior (see Section 4). The Jeffreys' prior for both balanced and unbalanced cases is given by

$$\pi_J(\lambda) \propto |I(\lambda)|^{1/2} = \left(\prod_{i=1}^{k} \lambda_i\right)^{-1/2}, \tag{2.6}$$

where $I(\lambda)$ is the information matrix associated with the likelihood function.

3. POSTERIOR ANALYSIS

3.1. *Propriety of the Posteriors*

Consider a prior distribution defined by

$$\pi_0(\lambda) \propto \left(\prod_{i=1}^{k} \lambda_i\right)^h \left(\sum_{i=1}^{k} \lambda_i^{-1}\alpha_i^2\right)^s, \quad \lambda_i > 0; i = 1, \ldots, k, \tag{3.1}$$

for $(h = -1/2, s = 0)$ and $(h = 0, s = 1/2)$. When $h = -1/2$ and $s = 0$, the prior is the Jeffreys' prior in (2.6), and $h = 0$ and $s = 1/2$ leads to the noninformative prior in (2.5) with the choice of $g(\zeta(\lambda)) = 1$.

Let $\{X_{ij}\}$ denote a random sample of size n_i from ith population having distribution $X_i \sim P_o(\lambda_i)$, $i = 1, \ldots, k$. Then the likelihood function of λ can be expressed by

$$L(\lambda|data) \propto \prod_{i=1}^{k} \lambda_i^{\sum_{j=1}^{n_i} x_{ij}} \exp\{-n_i\lambda_i\}, \quad \lambda_i > 0; \ i = 1, \ldots, k. \tag{3.2}$$

Therefore, using the prior from (3.1), the posterior distribution of λ is given by

$$\pi(\lambda|data) \propto \left(\prod_{i=1}^{k} \lambda_i\right)^h \left(\sum_{i=1}^{k} \lambda_i^{-1}\alpha_i^2\right)^s \prod_{i=1}^{k} \lambda_i^{\sum_{j=1}^{n_i} x_{ij}} \exp\{-n_i\lambda_i\}, \quad \lambda_i > 0; i = 1, \ldots, k \tag{3.3}$$

1. The density in (3.3) with $h = 0$ and $s = 1/2$ is always integrable over $C = \{\lambda : \ \lambda_i > 0, \ i = 1, \ldots, k\}$, because

$$(\sum_{i=1}^{k} \lambda_i^{-1}\alpha_i^2)^{1/2} \leq \sum_{i=1}^{k} \lambda_i^{-1/2}\alpha_i.$$

2. When $h = -1/2$ and $s = 0$, (3.3) is a kernel of product of k gamma densities, and hence is integrable over the range $C = \{\lambda : \lambda_i > 0; i = 1, \ldots, k\}$.

Therefore, the prior (3.3) yields proper posteriors, i.e. both noninformative prior and the Jeffreys' prior yield proper posteriors.

3.2. The Gibbs Sampling to Evaluate the Marginal Posterior Distribution of θ

In this section the informative prior and the uniform prior for θ are considered and the marginal posterior distribution is evaluated by Gibbs sampling simulation (Gelfand and Smith 1990). By the likelihood (3.3) and the prior of form (3.2), the conditional posterior distribution of λ_i given λ_{-i} where $\lambda_{-i} = (\lambda_1, \ldots, \lambda_{i-1}, \lambda_{i+1}, \ldots, \lambda_k)'$, is given by

$$\pi(\lambda_i|\lambda_{-i}, data) \propto \lambda_i^h \left(\sum_{i=1}^{k} \lambda_i^{-1}\alpha_i^2\right)^s \lambda_i^{\sum_{j=1}^{n_i} x_{ij}} \exp\{-n_i\lambda_i\}, \quad \lambda_i > 0; i = 1, \ldots, k. \tag{3.4}$$

For the noninformative prior with $h = 0$, and $s = 1/2$, (3.4) becomes

$$\pi(\lambda_i|\lambda_{-i}, data) \propto \left(1 + \lambda_i^{-1}\alpha_i^2 / \sum_{j \neq i} \lambda_j^{-1}\alpha_j^2\right)^{1/2} \lambda_i^{\sum_{j=1}^{n_i} x_{ij}} \exp\{-n_i\lambda_i\}, \quad \lambda_i > 0, \tag{3.5}$$

whereas for the uniform prior, in which $h = -1/2$, and $s = 0$, (3.4) is a gamma distribution of the density

$$\pi(\lambda_i|\lambda_{-i}, data) \propto \lambda_i^{\sum_{j=1}^{n_i} x_{ij} - 1/2} \exp\{-n_i\lambda_i\}, \quad \lambda_i > 0; i = 1, \ldots, k, \tag{3.6}$$

and hence it is straightforward to obtain a Gibbs sample from (3.6). To simulate random variates with densities (3.5), however, a rejection method is needed. The rejection method used in computations are described as follows.

[Simulation of X with Density $f(x) \propto (1 + \eta/x)^{1/2} x^\nu e^{-ax}$ for $x > 0$ and $h > 0$.]

Step 1. Generate x from a gamma distribution, $Gamma(\nu + 1/2, 1/(a - 1/2))$, where the mean of the distribution is $(\nu + 1/2)/(a - 1/2)$. If $u_1 \tau x^{\nu - 1/2} e^{-(a-1/2)x} \leq (1 + \sqrt{\eta/x}) x^\nu e^{-ax}$, then accept x, where u_1 is a $U(0, 1)$ variate, $\tau = (1 + \eta^{1/2}/y) y^{2\nu} e^{-ay^2} / \{y^{2\nu - 1} e^{-(a-1/2)y^2}\}$, and $y = \{-\eta^{1/2} + (\eta + 4)^{1/2}\}/2$.

Step 2. Simulate u_2 from $U(0,1)$. If $u_2 \leq (1+h/x)^{1/2}/(1+\sqrt{h/x})$ for x being derived from Step 1, then accept x.

Step 2 involves using a rejection method to generate a variate from $g(x) \propto (1+\sqrt{h/x})x^{\nu}e^{-ax}$ for $x > 0$. The maximum of the ratio $f(x)/g(x)$ when $x \in (0,\infty)$ is the ratio of the normalization constants of the two densities. Therefore, when a $U(0,1)$ random number $u_2 \leq f(x)/g(x)$, where x is from $g(x)$, then we accept x.

For the part of Step 1, let $g_1(x) \propto x^{\nu-1/2}e^{-(a-1/2)x}$. Then $g(x)/g_1(x) \propto (\sqrt{x}+\sqrt{h})e^{-x/2}$, and the maximum of $g(x)/g_1(x)$ is proportional to τ defined in Step 1. Here all the proportions mean that we are only ignoring the ratio of the normalization constants between the two density functions $g(x)$ and $g_1(x)$. The algorithm follows the regular rejection method after the maximum of $g(x)/g_1(x)$ has been determined.

Once we obtain a Gibbs sample $\{\lambda^{(m)} = (\lambda_1^{(m)},\ldots,\lambda_k^{(m)})', m = 1,\ldots,M\}$, let $\theta_m = \prod_{i=1}^k(\lambda_i^{(m)})^{\alpha_i}$ for $m = 1,\ldots,M$. Also let $\theta_{(m)}$ denote ordered value of $\theta^{(m)}$. Then the γth quantile of the marginal posterior distribution of θ can be estimated by

$$\hat{\theta}^{(\gamma)} = \begin{cases} \theta_{(1)} & if\ \ \gamma = 0 \\ \theta_{(m)} & if\ \ m-1 < \gamma \leq m \end{cases}. \tag{3.7}$$

Using (3.7), we compute

$$R_j(M) = (\hat{\theta}^{(j/M)}, \hat{\theta}^{(j+[1-\gamma]M)/M)}), \tag{3.8}$$

and a $100(1-\gamma)\%$ HPD interval of θ is $R_{j^*}(M)$ that has the smallest interval width among all $R_j(M)$'s. It is known that, as $M \to \infty$, $R_{j^*}(M)$ converges to true HPD interval of θ when the posterior density of θ is unimodal (see, Theorem 7.3.2 in Chen, Shao, and Ibrahim 2000).

4. FREQUENTIST COVERAGE PROBABILITY

An appropriate noninformative prior should have good frequentist properties. Many authors (Datta and Ghosh 1996; Datta, Ghosh and Mukerjee 2000; Mukerjee and Reid 1999; Stein 1985; Ye and Berger 1991 among others) suggested and argued those properties. One of them is that the frequentist coverage probability of a $(1-\gamma)$th posterior quantile should be close to $(1-\gamma)$. Using the Gibbs sampling method in Section 3, we investigate the property numerically

TABLE 4.1 *Frequentist Coverage Probabilities for .05(.95) Posterior Quantiles of θ for the Balanced Case*

$k = 3$

λ_0	(1, 1, 1)	(1, 2, 3)	(1, 5, 10)	(2, 2, 2)	(5, 5, 5)	(10, 10, 10)
π_T	.072(.984)	.065(.976)	.054(.961)	.047(.938)	.053(.947)	.052(.951)
π_J	.145(.996)	.122(.987)	.041(.983)	.038(.916)	.025(.928)	.040(.938)

$k = 8$

λ_0	(1, 2, 3, 4, 5, 6, 7, 8)	(1, 2, 3, 4, 5, 5, 5, 5)	(5, 5, 5, 5, 5, 5, 5, 5)
π_T	.041(.971)	.043(.967)	.056(.962)
π_J	.011(.977)	.014(.962)	.019(.901)
λ_0	(5, 5, 5, 5, 6, 7, 8, 9)	(5, 5, 5, 5, 10, 10, 10, 10)	(10, 10, 10, 10, 10, 10, 10, 10)
π_T	.053(.958)	.047(.954)	.049(.952)
π_J	.022(.907)	.023(.910)	.028(.919)

for the priors π_T and π_J. The computation of the frequentist coverage probability of a $(1-\gamma)$th posterior quantile of θ is based on the following algorithm for any fixed true $\lambda_0 = (\lambda_{10}, \ldots, \lambda_{k0})'$ and any predetermined probability value γ.

[Algorithm for Calculating the Frequentist Coverage Probability]

Step 1. Given a fixed true $\lambda_0 = (\lambda_{10}, \ldots, \lambda_{k0})'$, simulate random sample $\{x_{ij}\}$ of size n_i independently from $P_o(\lambda_{i0})$ distributions, $i = 1, \ldots, k$.

Step 2. Using the Gibbs sampler given in Section 3.2, simulate a posterior random vector $\lambda|$ $\mathbf{X}$, $\mathbf{X} = [\{x_{1j}\}, \ldots, \{x_{kj}\}]$, discarding the first 1,000 samples to 'burn-in' the sampler. For the balanced case (i.e. $n_1 = \cdots = n_k = n$), repeat the simulation m_1 times and calculate the proportion ρ for which $\theta = \prod_{i=1}^{k} \lambda_i^{\alpha_i} \leq \theta_0$, where $\theta_0 = \prod_{i=1}^{k} \lambda_{i0}^{\alpha_i}$. Note that, for the unbalanced case, α_i in θ and θ_0 needs to be replaced by α_i^*, where $\alpha_i^* \propto \sqrt{n_i}\alpha$.

Step 3. Repeat Step 1 and Step 2 m_2 times, and compute the proportion δ of $\rho \leq \gamma$ in these replications.

The quantity ρ is the estimate of the marginal posterior probability of θ for the interval $(0, \theta_0)$. On the other hand δ is the estimated frequentist coverage

TABLE 4.2 *Frequentist Coverage Probabilities for .05(.95)th Posterior Quantiles of θ for the Unbalanced Case*

$k = 3$						
λ_0	(1, 1, 1)	(1, 2, 3)	(1, 5, 10)	(2, 2, 2)	(5, 5, 5)	(10, 10, 10)
π_T	.052(.978)	.052(.979)	.054(.949)	.051(.950)	.051(.947)	.049(.945)
π_J	.022(.901)	.026(.914)	.028(.899)	.028(.900)	.031(.908)	.032(.913)

$k = 8$			
λ_0	(1, 2, 3, 4, 5, 6, 7, 8)	(1, 2, 3, 4, 5, 5, 5, 5)	(5, 5, 5, 5, 5, 5, 5, 5)
π_T	.049(.971)	.053(.970)	.048(.944)
π_J	.015(.857)	.014(.854)	.016(.850)
δ_0	(5, 5, 5, 5, 6, 7, 8, 9)	(5, 5, 5, 5, 10, 10, 10, 10)	(10, 10, 10, 10, 10, 10, 10, 10)
π_s	.053(.950)	.049(.949)	.052(.949)
π_u	.018(.866)	.020(.864)	.025(.891)

probability of the γth posterior quantile. The algorithm is applied to a balanced case with $n_i = 1$ and $\alpha_i = 1/k$ for all $i = 1, \ldots, k$. Table 4.1 shows the estimated frequentist coverage probabilities of $\gamma = .05(.95)$ posterior quantiles for different values of λ_0's and k obtained by using π_T and π_J. For the calculations of the entries in the table, m_1 is 10,000 and m_2 is 10,000. The maximum standard errors of estimations ρ and δ are .0035 and .005 respectively. From Table 4.1, clearly the noninformative prior π_T is better than the Jeffreys' prior π_J in most of the situations. Therefore, the noninformative prior is more appealing. When each coordinate of λ_0 is large, the frequentist coverage probabilities obtained from using π_T are almost close to the desired levels. The table also notes that, even though $n_i = 1$, the frequentist coverage probabilities of π_T are uniformly better than those for π_J in all the situations.

The algorithm is also applied to an unbalanced case with $n_i = 2^{2|i-4|}$ and $\alpha_i^* = 1$ so that $\alpha_i = 2^{2-|i-4|}$ according to the relation $\alpha_i^* \propto \sqrt{n_i}\alpha_i$ for $i = 1, \ldots, k$. Table 4.2 shows the estimated frequentist coverage probabilities of $\gamma = .05(.95)$ posterior quantiles for different values of λ_0's and k.

For the calculations of the entries in the table, we used $m_1 = m_2 = 10,000$. The maximum standard errors of estimations ρ and θ are .0042 and .006 respectively. From the table, we see that the frequentist coverage probabilities obtained from using the noninformative prior π_T are almost close to the desired levels, while those obtained from using π_J underestimates the levels. Furthermore, limited but informative comparison studies using various set of values of $\{n_i, \alpha_i; i = 1, \ldots, k\}$ and $\{n_i, \alpha_i^*; i = 1, \ldots .k\}$ revealed the same phenomena of Table 4.1 and Table 4.2, respectively. These results are not listed in tables for the sake of saving spaces. Therefore, for both balanced and unbalanced cases, the prior π_T satisfying the Tibshirani's differential equation is more appealing in the sense of the frequentist property.

5. Illustrative Examples

5.1. An Example in Environmental Statistics

As stated in Section 1, the estimate of a product of power of Poisson rates has been applied in many environmental statistical problems. An example is in the assessment of bacterial (Escherichia coli) counts in water due to contamination from sewage. The assessment is usually based upon the geometric mean of bacterial rates of samples taken over an observation period (see, information on bacteria standards in the U.S.;http://www.novaregion.org/4milerun/standards.htm).

In this example, estimating the geometric mean of bacterial rates is considered. Especially, the HPD interval for the geometric mean is estimated by the suggested Bayesian procedure. For this, the procedure is applied to Escherichia coli data set (year 2002) obtained from Cedar river down stream of Waterloo in Iowa State. The data is available from the IASTORET data base of University of Iowa (http://wqm.igsb.uiowa.edu/iastoret).

This example highlights the utility of the suggested Bayesian procedure in the following reasons: (i) Since the observations include zero count on Feb. 7, usual sample estimate of the geometric mean (the sample geometric mean of the counts) is not available. (ii) Unlike the methods for estimating integer powers of the Poisson rates in Harris (1971) and Harris and Som (1973), a frequentist approach for estimating a non-integer powers of the Poisson rates (such as the geometric mean of

the Poisson rates) has not been seen in literatures yet. To sample $\lambda = (\lambda_1, \ldots, \lambda_{12})'$ from the required posterior distributions, we use the suggested Gibbs sampler in Section 3 with $n_i = 1$, $i = 1, \ldots, k$. The diagnosics we used are described in Cowles and Carlin (1996).

For the Gibbs sampling, we used 5,000 iterations to 'burn-in' the sampler; the

TABLE 5.1 *Escherichia Coli Bacterial Counts Per 100 Milliliter of Water*

Obs. Date	Jan. 7	Feb. 7	Mar. 6	Apr. 3	May. 2	Jun. 6
Bacterial Count	18	0	40	18	70	2200
Obs. Date	Jul. 9	Aug. 5	Sep. 9	Oct. 3	Nov. 11	Dec. 5
Bacterial Count	40	150	20	400	18	27

decision is base on the trace plots. We took every 20th iterate to obtain a random sample of $M = 1,000$ iterates from the joint posterior posterior density. Trace and estimated marginal posterior density of $\theta = \prod_{i=1}^{12} \lambda_i^{1/12}$ are given in Figure 1. Figure 1 shows that the density is left-skewed and unimodal, and hence we need HPD interval for the interval estimation of θ. Based upon the Gibbs sample of size M, the 95% HPD interval for θ is obtained from (3.8), the method by Chen and Shao (1999). The 95% HPD interval using the suggested noninformative prior is $57.398 < \prod_{i=1}^{12} \lambda_i^{1/12} < 104.67462$. For the point estimate of θ, we may use either the posterior mean 82.86127 or the median 84.1397.

5.2. An Application to System Reliability

Despite the fact that the problem of inferencing the parameter θ may arise as a problem of interest in its own right, the inference may be of more interest and will presumably be applied more often as approximate solutions to the problem of establishing the reliability of systems of k independent parallel components. The example is selected for the purpose of illustrating this case.

Let p_i, $i = 1, \ldots, k$ be the probability that the ith component fails. Then the probability that the system fails is $\psi = \prod_{i=1}^{k} p_i$. Assume that n_i, $i = 1, \ldots, k$ independent Bernoulli trials are made on each component. Then under assumptions such that the binomial distribution can be satisfactorily approximated by the Poisson distribution, Bayesian estimation results in $\theta = \prod_{i=1}^{k} \lambda_i$ may be employed to treat the reliability systems of k independent parallel components through the relation

$$\psi = \theta / (\prod_{i=1}^{k} n_i). \tag{5.1}$$

Using the data listed in Table 2 of in Harris (1971), we give a number of numerical examples for the probability of failure of independent parallel compo-

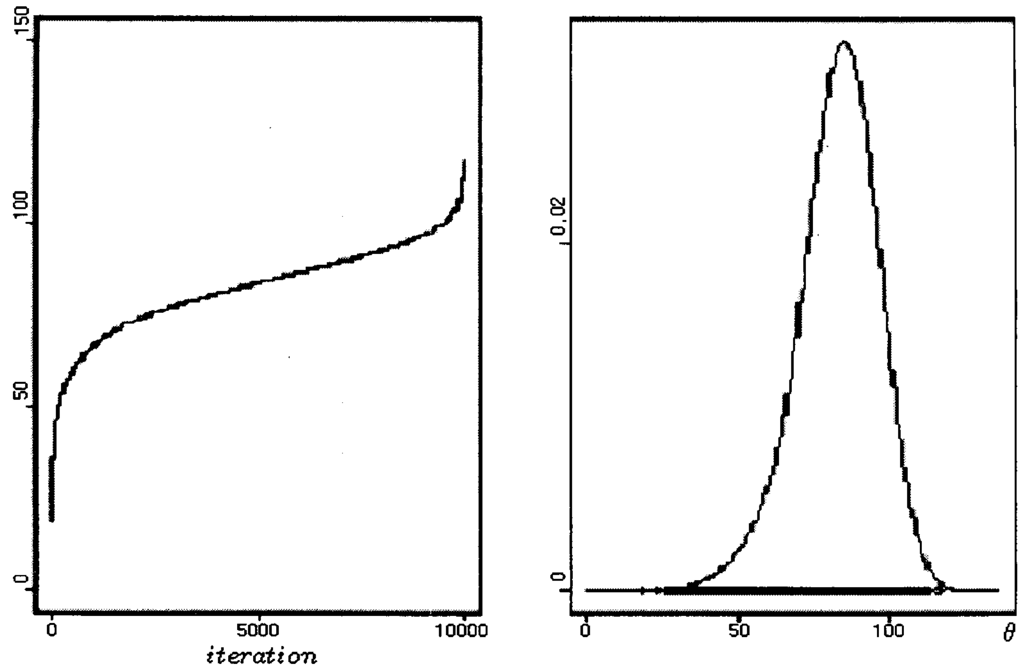

FIGURE 5.1 *Trace and Kernel Density Estimate of* θ

nents system. Assume that two systems are consists of two and three components in parallel, respectively. Then respective probabilities of system failure are $\psi_1 = p_1 p_2$ and $\psi_2 = p_1 p_2 p_3$, products of two and three binomial parameters. For confidence intervals of ψ_1 and ψ_2, numerical comparisons are provided for Harris's randomized method (1971), the likelihood ratio method of Madansky (1965) and the suggested Bayesian procedure. The utility of these comparisons is somewhat limited, since the techniques are all approximate and the exact confidence coefficient is not available.

For the Bayesian estimates, we first estimate the marginal posterior density of θ in order to check the skewness and unimodality of the distribution of θ, and then posterior quantities are obtained. Since estimated marginal posterior density of θ appears to be unimodal but severely skewed to the left we use the posterior median for the Bayesian estimate of θ. Three posterior quantities of ψ_1 and ψ_2, obtained from the relation (5.1), are presented in Table 5.2. The quantities are obtained from the same Gibbs sampling method described in Subsection 5.1. The first is γth quantile The second is the $\gamma \times 100\%$ HPD interval. The third

TABLE 5.2 *Upper Confidence Limits and the HPD interval for* $\prod_{i=1}^{k} p_i$ *with Confidence Coefficient* $1-\gamma=.9$; *the Posterior Median is in the Parenthesis*

$\psi_1 = p_1 p_2$

Sample sizes n_1, n_2	Observed x_1, x_2	Madansky's method	Harris's method	Bayesan (Median)	HPD Interval
100, 100	3, 5	.00433	.00416	.00406 (.00177)	[.000243, .004142]
100, 100	1, 4	.00188	.00184	.00172 (.00054)	[.000056, .001721]
100, 100	2, 2	.00168	.00170	.00157 (.00053)	[.000015, .001583]
150, 150	3, 3	.00133	.00128	.00124 (.00049)	[.000029, .001255]

$\psi_2 = p_1 p_2 p_3$

Sample sizes n_1, n_2, n_3	Observed x_1, x_2, x_3	Madansky's method	Harris's method	Bayesian (Median)	HPD
100, 100, 100	1, 2, 1	.000019	.000027	.000021 (4.9×10^{-6})	[$7.2 \times 10^{-9}, 2.2 \times 10^{-5}$]
100, 100, 100	2, 3, 5	.000133	.000145	.000132 (3.4×10^{-5})	[$1.7 \times 10^{-6}, 1.4 \times 10^{-4}$]

is the posterior median. The table notes that three methods provide quite close agreement. Even though the Bayesian method tends to give smaller upper limit for most of the cases, it is small enough to neglect. This fact and the results of Table 4.1 suggest that the Bayesian method can be an alternative method for estimating the system reliability.

REFERENCES

Chen, M. H., Shao, Q. M. (1999). "Monte Carlo estimation of Bayesian credible and HPD intervals." *Journal of Computational and Graphical Statistics*, **8**, 69–92.

Chen, M. H., Shao, Q. M., and Ibrahim, J. G. (2000). *Monte Carlo Methods in Bayesian Computation*, New York: Springer.

Cowles, M., and Carlin, B. (1996). "Markov chain Monte Carlo diagnostics: a comparative review," *Journal of the American Statistical Association*, **91**, 883–904.

Datta, G. S. and Ghosh, M. (1996). "On the invariance of noninformative priors," *The Annals of Staisitcs*, **24**, 141–159.

Datta, G. S. and Ghosh, M., and Mukerjee, R. (2000). "Some new results on probability matching priors", *Calcutta Statistical Association Bulletin*, **50**, 179–192.

Environmental Protection Agency (1986). "Bacteriological ambient water quality criteria for marine and fresh recreational waters," *Ambient Water Quality Criteria for Bacteria.*

Gelfand, A. E. and Smith, A. F. M. (1990). "Sampling based approaches to calculating marginal densities," *Journal of the American Statistical Association*, **85**, 398–409.

Harris, B. (1971). "Hypothesis testing and confidence intervals for products and quotients of Poisson parameters with applications to reliability," *Journal of the American Statistical Society*, **66**, 609–613.

Harris, B. and Soms, A. P. (1973). "The reliability systems of independent parallel components when some components are repeated," *Journal of the American Statistical Association*, **68**, 894–898.

Kenneth, V. D., John, S. D., and Kenneth, J. S. (1998). "Incorporating a geometric mean formula into CPI," *Monthly Labor Review*, October, 2–7.

Lehmann, E. L. (1959). *Testing Statistical Hypotheses*, John Wiley & Son, New York.

Madansky, A. (1965). "Approximate confidence limits for the reliability of series and parallel systems," *Technometrics*, **7**, 495–503.

Martz, H. F. and Waller, R. A. (1982). *Bayesian Reliability Analysis*, John Wiley & Son, New York.

Mukerjee, R. and Ghosh, M. (1997). "Second order probability matching priors," *Biometrika*, **84**, 970–975.

Mukerjee, R. and Reid, N. (1999). " On a property of probability matching priors : Matching the alternative coverage probabilities," *Biometrika*, **86**, 333–340.

Stein, C. (1985). "On the coverage probability of confidence sets based on a prior distribution," In *Sequential Methods in Statistics*, Banach center publications,**16**, Warsaw: PWN-Polish scientific publishers.

Tibshirani, R. (1989) "Noninformative priors for one parameter of many", *Biometrika*, **76**, 604–608.

Ye, K. and Berger, J. (1991). "Noninformative priors for inferences in exponential regression models," *Biometrika*, **78**, 645–656.