BAYESIAN ANALYSIS FOR THE POLY-WEIBULL DISTRIBUTION

by

James O. Berger and Dongchu Sun
Department of Statistics, Purdue University, West Lafayette, IN 47907-1399
Department of Statistics, University of Michigan, Ann Arbor, MI 48104

Technical Report # 92-05C

Department of Statistics
Purdue University

January, 1992

△

Bayesian Analysis for the Poly-Weibull Distribution *

James O. Berger and Dongchu Sun

Purdue University and The University of Michigan

Abstract

In this paper, Bayesian analysis for a Poly-Weibull distribution using informative and non-informative priors is discussed. This distribution typically arises when the data is the minimum of several Weibull failure times from competing risks. A general recursive formula is developed for exact computation of the posterior probability density function, posterior moments and the predictive reliability, when the shape parameters are known. Computation by simulation using the Gibbs sampler is also considered, for both known and unknown shape parameters, and is compared with the exact formula.

Keywords: Bi-Weibull distribution, competing risks, Bayesian inference, predictive distribution, Gibbs sampler, log-concave densities.

*This research was supported by NSF grants DMS-8702620, DMS-8717799 and DMS-8923071.

1 Basic Problem

1.1 Examples and the Bi-Weibull Distribution

Let us start with two examples.

Example 1.1 (in reliability theory). Suppose that the failure of a product can occur either because of an initial cause (e.g., manufacturing error) or long-term cause (e.g., wearout). It is frequently reasonable to assume that these two causes affect the product independently and have (Weibull) $\mathcal{W}(\theta_1, \beta_1)$ and $\mathcal{W}(\theta_2, \beta_2)$ distributions, respectively. The failure time of the product is then the minimum of these (independent) failure times, and the resulting reliability function (1− the c.d.f.) of the product is

$$R(t) = \exp\left\{ -\left(\frac{t}{\theta_1}\right)^{\beta_1} - \left(\frac{t}{\theta_2}\right)^{\beta_2} \right\}, \quad t > 0. \tag{1}$$

The corresponding distribution is called the Bi-Weibull distribution. As another example, Chuck, Goodrich, Hecht and McCulum (1990) investigated the mechanical behavior of selected Silicon Nitride (Si_3N_4) ceramics, which are candidate materials for advanced heat engines. They found that, at high temperatures, the time-to-failure seems to follow a Bi-Weibull distribution, with slow crack growth and damage accumulation being the two (independent) causes of failure.

Example 1.2 (competing risks in survival analysis.) Assume that a patient can die either because of a stroke or a heart attack. If the survival times under the two risks are assumed to be independently $\mathcal{W}(\theta_i, \beta_i)$ distributed, then the survival time of the patient has a Bi-Weibull distribution. Mendenhall and Hader (1958) and Cox (1959) are among the early papers concerning competing risks. Cox (1959) gave other examples where the above type of distribution arises. For recent papers concerning competing risks, see Greenhouse and Wolfe (1984), Goetghebeur and Ryan (1990). For simplicity, in this paper, we will utilize reliability language, rather than attempting to simultaneously use reliability and survival analysis language.

The hazard rate corresponding to (1) is

$$\frac{-\frac{d}{dt}R(t)}{R(t)} = \frac{\beta_1 t^{\beta_1 - 1}}{\theta_1^{\beta_1}} + \frac{\beta_2 t^{\beta_2 - 1}}{\theta_2^{\beta_2}}, \quad t > 0.$$

The shape of the hazard rate is determined by β_1 and β_2. If $\max(\beta_1, \beta_2) \in (0,1)$ ($\min(\beta_1, \beta_2) > 1$), the hazard rate is decreasing (increasing). If $\beta_1 < 1$ and $\beta_2 > 1$, the hazard rate is typically a bathtub curve. Figure 1 illustrates the Bi-Weibull reliability functions and hazard rates for several choices of β_1 in $(0,1)$ and β_2 in $(1,\infty)$. Note that if β_1 is close to 0, the hazard rate dramatically decreases in the initial failure region (see Martz and Waller (1982, p81) for the definition of initial failure region). Therefore, it is typically known that both β_1 and β_2 are bounded away from 0; this will be relevant in our later choice of the prior distribution.

Of course, more than two causes of failure or competing risks can be considered. Thus we will consider the general Poly-Weibull model.

1.2 The Poly-Weibull Model

Suppose that a product (or a system) consists of m (≥ 2) different elements. The age (time, miles, cyles, etc.) of the jth element is X_j, which has a Weibull, $\mathcal{W}(\theta_j, \beta_j)$, distribution with reliability function

$$R_j(t) \equiv P(X_j > t) = \exp\Big\{-\big(t/\theta_j\big)^{\beta_j}\Big\}, \text{ for } t > 0. \tag{2}$$

Here θ_j is the characteristic life and β_j is the shape parameter. Suppose that the m elements are connected in a series, and we do not know which element has failed when the system fails. (This is often the case in reliability contexts, especially with field – as opposed to experimental–data, but is perhaps less common in survival contexts.) Thus failure occurs at $X = \min(X_1, \cdots, X_m)$. Assume that the m failure ages $X_1, \cdots, X_m$ are independent so that the overall reliability of the product is

$$R(t) = \prod_{j=1}^{m} R_j(t) = \exp\Big\{-\sum_{j=1}^{m}\big(t/\theta_j\big)^{\beta_j}\Big\}, \quad t > 0, \tag{3}$$

and the probability of failure before t is $1 - R(t)$. Then X is said to have a Poly-Weibull distribution, and its density is given by

$$\begin{aligned} f(t|\beta_j, \theta_j, j = 1, \cdots, m) &= \frac{d}{dt}R(t) \\ &= \sum_{j=1}^{m} \frac{\beta_j t^{\beta_j - 1}}{\theta_j^{\beta_j}} \exp\Big\{-\sum_{k=1}^{m}\Big(\frac{t}{\theta_k}\Big)^{\beta_k}\Big\}, \quad t > 0. \end{aligned} \tag{4}$$

Here $\theta_1, \cdots, \theta_m$ are unknown parameters; we will treat both the cases where $\beta_1, \cdots, \beta_m$ are known and unknown. When the β_j are equal, the Poly-Weibull distribution is a regular Weibull distribution, and the parameters $\theta_1, \cdots, \theta_m$ are not identifiable. Thus we will assume that $\beta_1, \cdots, \beta_m$ are distinct. Further details about $(\beta_1, \cdots, \beta_m)$ will be discussed in Section 1.4.

1.3 Data and Statistics

Let $\vec{\theta} = (\theta_1, \cdots, \theta_m)$ and $\vec{\beta} = (\beta_1, \cdots, \beta_m)$. Assume that r units are tested independently, with their ages having the common p.d.f. (4). Let $t_1, \cdots, t_n$ be the observed failure times, and $t_1^*, \cdots, t_{r-n}^*$ the observed times-on-test of units that have not yet failed. The likelihood function is

$$L(\vec{\theta}, \vec{\beta}) = \Big\{\prod_{i=1}^{n}\sum_{j=1}^{m}\frac{\beta_j t_i^{\beta_j - 1}}{\theta_j^{\beta_j}}\Big\}\exp\Big\{-\sum_{k=1}^{m}\frac{S(\beta_k)}{\theta_k^{\beta_k}}\Big\}, \tag{5}$$

where

$$S(\beta_k) = \sum_{i=1}^{n} t_i^{\beta_k} + \sum_{i=1}^{r-n}(t_i^*)^{\beta_k}. \tag{6}$$

Note that simple sufficient statistics do not exist and classical approaches to the problem are difficult.

1.4 The Prior Distribution

Assume that $\theta_1, \cdots, \theta_m$, given $\vec{\beta}$, are independent and that the prior density of θ_j depends only on β_j and has the form

$$\pi_{1j}(\theta_j|\beta_j) \equiv \pi_{1j}(\theta_j|\beta_j, a_j, b_j) = \frac{\beta_j b_j^{a_j}}{\Gamma(a_j)}\,\theta_j^{-(1+\beta_j a_j)}\exp\Big\{-\frac{b_j}{\theta_j^{\beta_j}}\Big\}, \text{ for } \theta_j > 0. \tag{7}$$

where $a_j > 0$ and $b_j > 0$. Thus $\theta_j^{\beta_j}$ has the inverse Gamma, $\mathcal{IG}(a_j, b_j)$, distribution. Then the joint density of $\vec{\theta}$, given $\vec{\beta}$, is

$$\pi_1(\vec{\theta}|\vec{\beta}) = \prod_{j=1}^{m}\pi_{1j}(\theta_j|\beta_j). \tag{8}$$

Furthermore, assume that $\beta_1, \cdots, \beta_m$ are independent, and that β_j has density $\pi_{2j}(\beta_j)$ with respective to some σ-finite measure λ_j. Thus the prior density of $\vec{\beta}$ is

$$\pi_2(\vec{\beta}) = \prod_{j=1}^{m}\pi_{2j}(\beta_j). \tag{9}$$

Choice of the distributions in (8) and (9) will frequently be based on engineering knowledge or knowledge of previous similar products. No specific form for $\pi_2(\vec{\beta})$ is required, but, for the reasons given in Section 1.1, it is reasonable to assume that the support of β_j is (c_j, ∞) for some fixed

$c_j > 0$. This will be seen to be needed for posterior moments to exist. In Section 2.6.2, indications will be given that the analysis is often not sensitive to reasonable choices of the c_j. Note that, when only limited data is available, answers can depend significantly on the choice of π_2, so that careful modeling of subjective information about $\vec{\beta}$ is important. Noninformative priors will not be considered here; the standard noninformative priors for the ordinary Weibull distribution yield improper posteriors in the Poly-Weibull problem.

Here are some methods for eliciting the priors in (7), (i.e., determining the a_j and b_j), once the marginal prior $\pi_2(\vec{\beta})$ has been determined.

Option 1: Specify the first two marginal moments for θ_j. Assume that μ_{1j} and μ_{2j} are the first and second moments of θ_j. Since

$$E(\theta_j^k|\beta_j) = b_j^{k/\beta_j}\Gamma(a_j - \frac{k}{\beta_j})/\Gamma(a_j), \quad \text{if} \quad 0 \le k < a_j\beta_j, \tag{10}$$

a_j and b_j can be determined by solving the equations

$$\mu_{jk} = E(\theta_j^k) = \frac{1}{\Gamma(a_j)}\int_{c_j}^{\infty} b_j^{k/\beta_j}\Gamma(a_j - \frac{k}{\beta_j})\pi_j(\beta_j)d\lambda_j(\beta_j), \quad k = 1, 2. \tag{11}$$

Note that a solution exists only if $a_j > 2/c_j$.

Option 2: Specify two marginal quantiles of θ_j. Note that

$$\begin{aligned} P(\theta_j \le x) &= E\{P(\frac{2b_j}{\theta_j^{\beta_j}} \ge \frac{2b_j}{x^{\beta_j}}|\beta_j)\} \\ &= \int_0^{\infty}\Big[1 - \chi^2(2a_j; \frac{2b_j}{x^{\beta_j}})\Big]\pi_{2j}(\beta_j)d\lambda_j(\beta_j), \end{aligned} \tag{12}$$

where $\chi^2(2a_j;\cdot)$ is the c.d.f. of the χ^2 distribution with $2a_j$ degrees of freedom. If $q_j(\alpha_1)$ and $q_j(\alpha_2)$ are the α_1-quantile and the α_2-quantile of the marginal distribution of θ_j, respectively, then a_j and b_j can be obtained by solving the equations

$$\int_{c_j}^{\infty} \chi^2(2a_j; \frac{2b_j}{[q_j(\alpha_k)]^{\beta_j}})\pi_{2j}(\beta_j)d\lambda_j(\beta_j) = 1 - \alpha_k, \quad k = 1, 2.$$

Option 3: Specify two predictive quantiles. Note that the predictive reliability for the jth component failure time is

$$\begin{aligned} P(T_j \ge s) &= \int_{c_j}^{\infty}\int_0^{\infty} P(T_j \ge s|\theta_j, \beta_j)\pi_{1j}(\theta_j|\beta_j)\pi_{2j}(\beta_j)d\theta_j d\lambda_j(\beta_j) \\ &= \int_{c_j}^{\infty}\int_0^{\infty} \exp\Big\{-\frac{s^{\beta_j}}{\theta_j^{\beta_j}}\Big\}\frac{\beta_j b_j^{a_j}}{\Gamma(a_j)}\,\theta_j^{-(1+\beta_j a_j)}\exp\Big\{-\frac{b_j}{\theta_j^{\beta_j}}\Big\}\pi_{2j}(\beta_j)d\lambda_j(\beta_j) \\ &= \int_{c_j}^{\infty}\frac{1}{(1 + s^{\beta_j}/b_j)^{a_j}}\pi_{2j}(\beta_j)d\lambda_j(\beta_j), \quad s > 0. \end{aligned}$$

If $q_j^*(\alpha_1)$ and $q_j^*(\alpha_2)$ are the α_1-quantile and the α_2-quantile of the predictive distribution of the jth component, respectively, then a_j and b_j can be obtained by solving the equations

$$\int_{c_j}^{\infty} \frac{1}{(1+q_j^*(\alpha_k)^{\beta_j}/b_j)^{a_j}} \pi_{2j}(\beta_j) d\lambda_j(\beta_j) = 1-\alpha_k, \quad k=1,2. \tag{13}$$

The prior assessment for known $\vec{\beta}$, considered by Sun and Berger (1991a), is the special case of choosing a degenerate point mass for $\vec{\beta}$; then both a_j and b_j depend on the chosen $\vec{\beta}$.

Example 1.3 Consider a case where the age of a product has a Bi-Weibull distribution. Suppose that, from engineering knowledge and/or knowledge of previous similar products, the best guess for the mean of θ_1 is $\mu_{11} = 1,000$ (hours) and the standard deviation for the guess is $1,580$ (hours) so that $\mu_{12} = 1,000^2 + 1,580^2 = 3,496,400$. Prior information indicates that β_1 is between 0.1 and 0.9, with 0.5 considered the "most likely" value (mode); this is modeled by assuming that $(\beta_1 - 0.1)/(0.9-0.1)$ has the Beta(15, 15) distribution. It follows from equation (11) that $a_1 = 28.0$ and $b_1 = 684.0$ (using a Fortran subroutine from IMSL for the integration). Similarly, suppose that the best guess for the mean of θ_2 is $\mu_{21} = 2,000$ (hours) and the standard deviation for the guess is $2,830$ (hours) so that $\mu_{22} = 2,000^2 + 2,830^2 = 12,008,900$. It is known that $\beta_2 > 1$, and the "most likely" value (mode) is specified to be 2; this is modeled by assuming that $\beta_2 - 1$ is Gamma(15, 14). It then follows that $a_2 = 10.0$ and $b_2 = 18,600,000$.

1.5 Preview

Expanding the product and summation in (5) results in an expression with m^n terms. For example, if $m = 3$ and $n = 20$, there are $3^{20} = 3,486,784,399$ terms. Computation via brute force expansion is thus typically not feasible. The purpose of this paper is to show how Bayesian analysis can, nevertheless, be done. Two techniques, an exact iterative computational scheme and Gibbs sampling, are developed and compared.

In Section 2, an iterative computational scheme will be developed for closed form Bayesian analysis in the case of known $\vec{\beta}$. For an example of such a situation, see Townsend (1989). Analysis for the Bi-Weibull case will be presented first. A general formula is then introduced for the Poly-Weibull case, allowing closed form computation of the relative posterior probability density function, the posterior moments and the predictive reliability. This formula is recursive, with each step of the recursion corresponding to incorporation of an additional data point, and hence is completely compatible with sequential or multistage experimentation. The total number of computations needed is roughly $2n\binom{n+m-1}{m-1} + 3mn^2$, which is typically much smaller than the brute force m^n computations. (When $m = 3$ and $n = 20, 2n\binom{n+m-1}{m-1} + 3mn^2 = 11,760$.)

Unknown $\vec{\beta}$ is also addressed in Section 2 with two "standard" techniques. The first is consideration of the marginal likelihood function of $\vec{\beta}$ and the "Type II" maximum likelihood estimate. The second is extension of the recursive formula to deal with unknown $\vec{\beta}$, but with some involvement of numerical integration.

In Section 3, the Gibbs sampling approach to Bayesian computation for the Poly-Weibull case with unknown $\vec{\beta}$ is developed. With introduction of auxiliary variables and utilization of log–concave rejection sampling, the analysis is quite efficient computationally.

Finally, the exact computation and Gibbs sampling are compared in Section 4. Also, some generalizations of the problem are discussed.

2 Closed-Form Bayesian Analysis

2.1 Introduction

We initially assume that $\vec{\beta}$ is given. The case where $\vec{\beta}$ is unknown will be studied in Section 2.6. The following notation will be used throughout this section. For $j = 1, \cdots, m$, define

$$T_{\beta_j} = S(\beta_j) + b_j; \tag{14}$$

$$H_j(x;\beta_j) = \frac{\beta_j}{x^{1+\beta_j a_j}} \exp\left\{-T_{\beta_j}/x^{\beta_j}\right\}, \ x > 0, \tag{15}$$

where $S(\beta_j)$ is given by (6).

2.2 Posterior for the Bi-Weibull Case

Because of its importance and comparative simplicity, we first present results for the Bi-Weibull distribution. Define

$$\widetilde{W}(1;0) \equiv \widetilde{W}(1;0;\beta_1,\beta_2) = t_1^{\beta_2} \quad \text{and} \quad \widetilde{W}(1;1) \equiv \widetilde{W}(1;1;\beta_1,\beta_2) = t_1^{\beta_1}.$$

For $n \geq 2$ and $0 \leq i \leq n$, define

$$\widetilde{W}(n;i) \equiv \widetilde{W}(n;i;\beta_1,\beta_2) = \begin{cases} \widetilde{W}(n-1;0)t_n^{\beta_2}, & \text{if } j=0, \\ \widetilde{W}(n-1;i-1)t_n^{\beta_1} + \widetilde{W}(n-1;i)t_n^{\beta_2}, & \text{if } 1 \leq i < n, \\ \widetilde{W}(n-1;n-1)t_n^{\beta_1}, & \text{if } i=n. \end{cases} \quad (16)$$

Note that $\widetilde{W}(n;0) = [\prod_{i=1}^n t_i]^{\beta_2}$ and $\widetilde{W}(n;n) = [\prod_{i=1}^n t_i]^{\beta_1}$. Because of the recursive nature of (16), computation of posterior quantities can be done efficiently, requiring only $O(n^2)$ computations. For $\mu_1(< a_1\beta_1)$, $\mu_2(< a_2\beta_2)$ and $t(\geq 0)$, define

$$J(\mu_1,\mu_2;t) \equiv J(\mu_1,\mu_2;t;\beta_1,\beta_2) = \sum_{i=0}^{n} \frac{\widetilde{W}(n;i)\beta_1^i\beta_2^{n-i}\Gamma(a_1+i-\frac{\mu_1}{\beta_1})\Gamma(a_2+n-i-\frac{\mu_2}{\beta_2})}{(T_{\beta_1}+t^{\beta_1})^{i+a_1-\mu_1/\beta_1}(T_{\beta_2}+t^{\beta_2})^{n-i+a_2-\mu_2/\beta_2}}.$$

Expressions for several posterior quantities of interest are given in Table 1. These follow from the general expressions given in Section 2.3.

Table 1: Formulas for Posterior Quantities for a Bi-Weibull Distribution, Given β_1 and β_2

joint density	$\pi(\theta_1,\theta_2\|\beta_1,\beta_2;\text{data})$	$\frac{H_1(\theta_1;\beta_1)H_2(\theta_2;\beta_2)}{J(0,0;0)}\sum_{j=0}^{n}\widetilde{W}(n;j)\left[\frac{\beta_1}{\theta_1^{\beta_1}}\right]^j\left[\frac{\beta_2}{\theta_2^{\beta_2}}\right]^{n-j}$
marginal density	$\pi(\theta_1\|\beta_1,\beta_2;\text{data})$	$\frac{H_1(\theta_1;\beta_1)}{J(0,0;0)}\sum_{j=0}^{n}\frac{\widetilde{W}(n;j)\beta_2^{n-j}\Gamma(n-j+a_2)}{T_{\beta_2}^{n-j+a_2}}\left[\frac{\beta_1}{\theta_1^{\beta_1}}\right]^j$
	$\pi(\theta_2\|\beta_1,\beta_2;\text{data})$	$\frac{H_2(\theta_2;\beta_2)}{J(0,0;0)}\sum_{j=0}^{n}\frac{\widetilde{W}(n;j)\beta_1^j\Gamma(j+a_1)}{T_{\beta_1}^{j+a_1}}\left[\frac{\beta_2}{\theta_2^{\beta_2}}\right]^{n-j}$
mean	$E(\theta_1\|\beta_1,\beta_2;\text{data})$	$\frac{J(1,0;0)}{J(0,0;0)}$
	$E(\theta_2\|\beta_1,\beta_2;\text{data})$	$\frac{J(0,1;0)}{J(0,0;0)}$
variance	$\text{Var}(\theta_1\|\beta_1,\beta_2;\text{data})$	$\frac{J(2,0;0)}{J(0,0;0)}-\left\{\frac{J(1,0;0)}{J(0,0;0)}\right\}^2$
	$\text{Var}(\theta_2\|\beta_1,\beta_2;\text{data})$	$\frac{J(0,2;0)}{J(0,0;0)}-\left\{\frac{J(0,1;0)}{J(0,0;0)}\right\}^2$
covariance	$\text{Cov}(\theta_1,\theta_2\|\beta_1,\beta_2;\text{data})$	$\frac{J(1,1;0)}{J(0,0;0)}-\frac{J(1,0;0)J(0,1;0)}{(J(0,0;0))^2}$
predictive reliability	$\hat{R}(t)=E[R(t)\|\beta_1,\beta_2;\text{data}]$	$\frac{J(0,0;t)}{J(0,0;0)}$

Example 2.1 Consider a case where the age of a product has a Bi-Weibull distribution with $\beta_1 = 0.5$ and $\beta_2 = 2.0$. For example, see Townsend (1989). Assume that the hyperparameters for the prior distributions are $a_1 = 15.0, a_2 = 1.90, b_1 = 430$ and $b_2 = 10,575,000$. The following sample of size 20 ($n = r = 20$) is simulated from the Bi-Weibull distribution with $\theta_1 = 750$ and $\theta_2 = 3000$:

8.96, 2189.49, 384.42, 1792.82, 2891.43, 844.82, 243.04, 982.33, 1660.83, 88.32,
1037.78, 406.86, 130.21, 449.15, 129.80, 355.16, 111.81, 392.48, 304.68, 75.98.

The marginal prior densities (with solid curves) and marginal posterior densities (with dotted curves) of θ_1 and θ_2 can be seen from Figure 2. The posterior moments are given in Table 2. The predictive reliability is shown by the solid curve in Figure 3.

Table 2: Prior and Posterior Moments of (θ_1, θ_2) for a Bi-Weibull Distribution, given (β_1, β_2)

Prior Distribution		Posterior Distribution	
$\mathrm{E}(\theta_1 \vert \beta_1, \beta_2)$	1015.94	$E(\theta_1 \vert \beta_1, \beta_2; \mathrm{data})$	1002.60
$\mathrm{E}(\theta_2 \vert \beta_1, \beta_2)$	3000.02	$E(\theta_2 \vert \beta_1, \beta_2; \mathrm{data})$	2466.85
$\sqrt{\mathrm{Var}(\theta_1 \vert \beta_1, \beta_2)}$	625.27	$\sqrt{\mathrm{Var}(\theta_1 \vert \beta_1, \beta_2; \mathrm{data})}$	439.74
$\sqrt{\mathrm{Var}(\theta_2 \vert \beta_1, \beta_2)}$	1658.28	$\sqrt{\mathrm{Var}(\theta_2 \vert \beta_1, \beta_2; \mathrm{data})}$	887.46
$\mathrm{Cov}(\theta_1, \theta_2 \vert \beta_1, \beta_2)$	0	$\mathrm{Cov}(\theta_1, \theta_2 \vert \beta_1, \beta_2; \mathrm{data})$	−75823.10
$\mathrm{Corr}(\theta_1, \theta_2 \vert \beta_1, \beta_2)$	0	$\mathrm{Corr}(\theta_1, \theta_2 \vert \beta_1, \beta_2; \mathrm{data})$	−0.1943

2.3 Posterior for the Poly-weibull Case

2.3.1 A Recursive Formula and Notation

The key to avoiding a combinatorial explosion in the posterior analysis is the recursive formula presented in this section. Let $\mathcal{N}$ denote the set of all nonnegative integers. For $m, n \in \mathcal{N}, m \geq 2, n \geq 1$, denote all the partitions $(i_1, \cdots, i_m)$ of n by

$$\Delta_{m,n} = \Big\{(i_1, \cdots, i_m) : i_j \in \mathcal{N}, \ \sum_{j=1}^{m} i_j = n\Big\}. \tag{17}$$

Suppose that the failure times $t_1, \cdots, t_n$ are observed. The recursive formula is defined by the following two steps.

Step 1. For $(i_1, \cdots, i_m) \in \Delta_{m,1}$, define

$$W(1; i_1, \cdots, i_m; \vec{\beta}) = \begin{cases} t_1^{\beta_1}, & \text{if } (i_1, i_2, \cdots, i_m) = (1, 0, \cdots, 0), \\ \cdots & \cdots\cdots \\ t_1^{\beta_m}, & \text{if } (i_1, \cdots, i_{m-1}, i_m) = (0, \cdots, 0, 1). \end{cases} \tag{18}$$

Step 2. For $(i_1, \cdots, i_m) \in \Delta_{m,k}$ $(2 \le k \le n)$, define

$$W(k; i_1, \cdots, i_m; \vec{\beta}) = \sum_{i_j \ge 1} W(k-1; i_1, \cdots, i_{j-1}, i_j - 1, i_{j+1}, \cdots, i_m; \vec{\beta})\, t_k^{\beta_j}. \tag{19}$$

It is easy to see that (19) is well defined. The case $m = 2$ is equivalent to (16).

Note that, in sequential experimentation, each new incoming failure t_i calls for updating the previous $W's$ by Step 2. There are only $\binom{m+i-1}{m-1}$ terms in this updating of the $W's$ by an incoming t_i. Thus one does not have to cycle again through the induction, making evaluation of the posterior in a sequential context especially inexpensive.

For a real vector $\vec{\mu} = (\mu_1, \cdots, \mu_m)$ $(\mu_j < \beta_j(a_j + i_j))$ and $t \ge 0$, let

$$J(\vec{\mu}; t) \equiv J(\vec{\mu}; t; \vec{\beta}) = \sum_{(i_1, \cdots, i_m) \in \Delta_{m,n}} W(n; i_1, \cdots, i_m; \vec{\beta}) \prod_{j=1}^{m} \frac{\beta_j^{i_j} \Gamma\left(a_j + i_j - \frac{\mu_j}{\beta_j}\right)}{(T_{\beta_j} + t^{\beta_j})^{i_j + a_j - \mu_j/\beta_j}}. \tag{20}$$

2.3.2 The Poly-Weibull Posterior Distribution and Moments

Theorem 2.1 The posterior density of $\vec{\theta}$, given $\vec{\beta}$ and the data, is

$$\pi(\vec{\theta} | \vec{\beta}; \text{data}) = \frac{1}{J(\vec{0}; 0; \vec{\beta})} \sum_{(i_1, \cdots, i_m) \in \Delta_{m,n}} W(n; i_1, \cdots, i_m; \vec{\beta}) \prod_{k=1}^{m} H_k(\theta_k; \beta_k) \left\{ \frac{\beta_k}{\theta_k^{\beta_k}} \right\}^{i_k}, \tag{21}$$

and the marginal posterior density of θ_j $(1 \le j \le m)$, given $\vec{\beta}$ and the data, is

$$\pi_j(\theta_j | \vec{\beta}; \text{data}) = \frac{H_j(\theta_j; \beta_j)}{J(\vec{0}; 0; \vec{\beta})} \sum_{(i_1, \cdots, i_m) \in \Delta_{m,n}} W(n; i_1, \cdots, i_m; \vec{\beta}) \left(\frac{\beta_j}{\theta_j^{\beta_j}} \right)^{i_j} \prod_{k \ne j}^{m} \frac{\beta_k^{i_k} \Gamma(a_k + i_k)}{T_{\beta_k}^{a_k + i_k}}. \tag{22}$$

PROOF. The posterior density of $(\theta_1, \cdots, \theta_m)$ is proportional to

$$\left\{ \prod_{i=1}^{n} \sum_{k=1}^{m} \frac{\beta_k t_i^{\beta_k}}{\theta_k^{\beta_k}} \right\} \left\{ \prod_{k=1}^{m} \frac{\beta_k}{\theta_k^{1 + \beta_k a_k}} \right\} \exp\left\{ - \sum_{k=1}^{m} \frac{T_{\beta_k}}{\theta_k^{\beta_k}} \right\}. \tag{23}$$

By induction, it can be shown that (23) is equal to

$$\sum_{(i_1, \cdots, i_m) \in \Delta_{m,n}} W(n; i_1, \cdots, i_m; \vec{\beta}) \prod_{k=1}^{m} H_k(\theta_k; \beta_k) \left\{ \frac{\beta_k}{\theta_k^{\beta_k}} \right\}^{i_k}. \tag{24}$$

Since

$$\int_0^\infty \theta_k^y H_k(\theta_k;\beta_k)\theta_k^{-\beta_k i_k} d\theta_k = \Gamma(a_k + i_k - \frac{y}{\beta_k})/T_{\beta_k}^{i_k+a_k}, \tag{25}$$

the normalization constant for the joint posterior density of $\vec{\theta}$ is $J(\vec{0};0)$. This proves the first part. The second part follows immediately. □

The marginal posterior density of (θ_i,θ_j) can be written similarly. For example, the marginal posterior density of (θ_1,θ_2) is

$$\begin{aligned}\pi(\theta_1,\theta_2|\vec{\beta};\text{data}) &= \frac{H_1(\theta_1;\beta_1)H_2(\theta_2;\beta_2)}{J(\vec{0};0;\vec{\beta})} \\ &\times \sum_{(i_1,\cdots,i_m)\in\Delta_{m,n}} W(n;i_1,\cdots,i_m;\vec{\beta})\Big[\frac{\beta_1}{\theta_1^{\beta_1}}\Big]^{i_1}\Big[\frac{\beta_2}{\theta_2^{\beta_2}}\Big]^{i_2}\prod_{j=3}^{m}\frac{\beta_j^{i_j}\Gamma(a_j+i_j)}{T_{\beta_j}^{a_j+i_j}}.\end{aligned}$$

Theorem 2.2 The posterior moments, given $\vec{\beta}$ and the data, are

$$\text{E}(\theta_j|\vec{\beta};\text{data}) = \frac{J(\vec{0}_{(1)};0;\vec{\beta})}{J(\vec{0};0;\vec{\beta})}, \tag{26}$$

$$\text{Var}(\theta_1|\vec{\beta};\text{data}) = \frac{J(2\vec{0}_{(j)};0;\vec{\beta})}{J(\vec{0};0;\vec{\beta})} - \Big[\frac{J(\vec{0}_{(j)};0;\vec{\beta})}{J(\vec{0};0;\vec{\beta})}\Big]^2, \tag{27}$$

$$\text{Cov}(\theta_i,\theta_j|\vec{\beta};\text{data}) = \frac{J(\vec{0}_{(i,j)};0;\vec{\beta})}{J(\vec{0};0;\vec{\beta})} - \frac{J(\vec{0}_{(i)};0;\vec{\beta})J(\vec{0}_{(j)};0;\vec{\beta})}{(J(\vec{0};0;\vec{\beta}))^2}, \tag{28}$$

where $\vec{0}_{(j)} = \{(x_1,\cdots,x_m) : x_j = 1, x_k = 0, k \neq j\}$ and $\vec{0}_{(i,j)} = \{(x_1,\cdots,x_m) : x_i = x_j = 1, x_k = 0, k \neq i,j\}$.

The simple proof of Theorem 2.2 is omitted.

The number of terms in the expression for $J(\vec{0};0)$ is $\#(\Delta_{m,n})$, which equals $\binom{n+m-1}{m-1}$. In Section 4.1, we will see that the recursive formula effectively reduces the total number of computations for a Poly-Weibull distribution from an exponential rate in the sample size n to a polynomial rate.

2.4 Approximation for Posterior Quantiles of $\theta_1,\cdots,\theta_m$

There are no nice formulas for posterior quantiles of $\theta_1,\cdots,\theta_m$. One could, of course, use numerical integration to determine quantiles of $\pi(\theta_j|\vec{\beta},\text{data})$, if desired. However, the following approximation seems to work quite well. Let m_j and V_j be the posterior mean and the posterior variance of θ_j. Approximate the posterior distribution by the distribution of form (7) which has moments m_j and V_j. Thus the approximation is

$$\pi(\theta_j|\vec{\beta}, \text{data}) \approx \frac{\beta_j \tilde{b}_j^{\tilde{a}_j}}{\Gamma(\tilde{a}_j)}\, \theta_j^{-(1+\beta_j\tilde{a}_j)} \exp\left\{-\frac{\tilde{b}_j}{\theta_j^{\beta_j}}\right\}, \tag{29}$$

where $(\tilde{a}_j, \tilde{b}_j)$ satisfy

$$\begin{cases} \Gamma(\tilde{a}_j - 2/\beta_j)\Gamma(\tilde{a}_j)/\Gamma^2(\tilde{a}_j - 1/\beta_j) = 1 + V_j/m_j^2, \\ \tilde{b}_j = \left[m_j\Gamma(\tilde{a}_j)/\Gamma(\tilde{a}_j - 1/\beta_j)\right]^{\beta_j}. \end{cases}$$

Note that an approximate value of $\tilde{a}_j$ can be determined by iteratively solving

$$\tilde{a}_j = 0.5 + \left[\ln(1 + \frac{V_j}{m_j^2}) + \frac{2}{\beta_j}\ln\left(1 - \frac{1}{\tilde{a}_j\beta_j - 1}\right)\right]\Big/\ln\left[1 - \frac{1}{(\tilde{a}_j\beta_j - 1)^2}\right], \tag{30}$$

starting from the initial value

$$\tilde{a}_j^0 = \frac{1}{2\beta_j}\left(3 + \frac{1}{1 - (1 + \frac{V_j}{m_j^2})^{-0.5\beta_j}}\right). \tag{31}$$

Actually, $\tilde{a}_j^0$ is typically accurate enough.

Finally, the approximate α^{th} posterior quantile of θ_j is given by the α^{th} quantile of (29), which can be shown to be

$$\hat{q}_j(\alpha) = \left[\frac{2\tilde{b}_j}{\chi^2_{2\tilde{a}_j}(1-\alpha)}\right]^{1/\beta_j} = \frac{m_j\Gamma(\tilde{a}_j)}{\Gamma(\tilde{a}_j - 1/\beta_j)}\left[\frac{2}{\chi^2_{2\tilde{a}_j}(1-\alpha)}\right]^{1/\beta_j}, \tag{32}$$

where $\chi_j^2(1-\alpha)$ is the $(1-\alpha)$th quantile of the χ^2 distribution with j degrees of freedom.

Example 2.1 (continued). From Table 2, $(m_1, V_1) = (1002.60, 439.74^2)$ and $(m_2, V_2) = (2466.85, 887.46^2)$. Computation yields $(\tilde{a}_1^0, \tilde{b}_1^0) = (26.2383, 783.1516)$ and $(\tilde{a}_2^0, \tilde{b}_2^0) = (2.9316, 13{,}360{,}913)$. After 10 iterations of (30), we get $(\tilde{a}_1, \tilde{b}_1) = (25.2535, 751.9596)$ and $(\tilde{a}_2, \tilde{b}_2) = (2.9893, 13{,}709{,}647)$; these are accurate through the given digits. Figure 4 indicates the quality of the quantile approximations. The true quantiles and their approximations (32) are represented by the solid and dashed lines, respectively. There are no noticeable differences between the true quantiles and their approximations.

2.5 Predictive Reliability

Often, the predicted time to failure of the product is of considerable interest. Let T be a future observation of the product, which is assumed to be independent of current data. Then the predictive reliability, with $\vec{\beta}$ given, is defined by

$$\widehat{R}(t;\vec{\beta}) \equiv P(T > t|\vec{\beta};\text{data}) = \int_{\Re_n^+} R(t)\pi(\vec{\theta}|\vec{\beta},\text{data})d\vec{\theta}, \quad t > 0,$$

where $R(t)$ is given by (3) and $\Re_n^+ = \{(y_1,\cdots,y_n) : y_j > 0\}$. (Under squared error loss, $\widehat{R}(t;\vec{\beta})$ is the Bayes estimate of $R(t)$, given $\vec{\beta}$.)

Theorem 2.3 $\widehat{R}(t;\vec{\beta}) = J(\vec{0};t;\vec{\beta})/J(\vec{0};0;\vec{\beta})$.

PROOF. Note that

$$R(t)\pi(\vec{\theta}|\vec{\beta},\text{data}) = \sum_{(i_1,\cdots,i_m)\in\Delta_{m,n}} \frac{W(n;i_1,\cdots,i_m;\vec{\beta})}{J(\vec{0};0;\vec{\beta})} \prod_{j=1}^{m} \frac{\beta_j^{i_j+1}\exp\{-(T_{\beta_j}+t^{\beta_j})/\theta_j^{\beta_j}\}}{\theta_j^{1+\beta_j(a_j+i_j)}}.$$

The result immediately follows from (25). □

2.6 Bayesian Analysis When $\vec{\beta}$ is Unknown

We may face a case where the prior $\pi_2(\vec{\beta})$ is either known or unknown. These two situations are discussed separately in the following two subsections.

2.6.1 The "Type II" Maximum Likelihood Method

When $\pi_2(\vec{\beta})$ is unknown, the method of determining $\pi_1(\vec{\theta}|\vec{\beta})$ in Section 1.4 cannot be used as stated. Another possibility, however, is to treat $\vec{\beta}$ as a known value, which is equivalent to letting the prior for $\vec{\beta}$ be a degenerate point mass. Based on available information about $\vec{\theta}$, as discussed in section 1.4, one can then obtain $a_j = a_j(\beta_j)$ and $b_j = b_j(\beta_j)$. Finally, consider the marginal likelihood function for $\vec{\beta}$, which is given by

$$L^*(\vec{\beta}) \equiv \int_{\Re_n^+} L(\vec{\theta},\vec{\beta})\pi(\vec{\theta}|\vec{\beta})d\vec{\theta} = \prod_{j=1}^{m} \frac{b_j(\beta_j)^{a_j(\beta_j)}}{\Gamma(a_j(\beta_j))} J(\vec{0};0;\vec{\beta}). \tag{33}$$

The easiest way to estimate $\vec{\beta}$ would be to use maximum likelihood theory. If $\vec{\beta}^*$ satisfies

$$L^*(\vec{\beta}^*) = \max_{\vec{\beta}} L^*(\vec{\beta}),$$

then $\vec{\beta}^*$ is the "Type II" maximum likelihood estimate of $\vec{\beta}$. For $m = 2$, a contour graph of $L^*(\beta_1,\beta_2)$ is also revealing.

Recall that, a_j and b_j, determined from either of the three options in Section 1.4, depend on β_j in a complicated way. In order to find the "Type II" MLE under Option 1, it is far simpler (and quite accurate) to use the approximation (31), i.e.,

$$\tilde{a}_j(\beta_j) \equiv \frac{1}{2\beta_j}\Big[3 + \frac{1}{1 - 1/(1+\frac{V_j}{m_j^2})^{0.5\beta_j}}\Big],$$

$$\tilde{b}_j(\beta_j) \equiv \Big[m_j \Gamma(\tilde{a}_j(\beta_j))/\Gamma(\tilde{a}_j - 1/\beta_j)\Big]^{\beta_j}.$$

Then the "Type II" MLE can be obtained by maximizing

$$\tilde{L}(\vec{\beta}) \equiv \prod_{j=1}^{m} \frac{\tilde{b}_j(\beta_j)^{\tilde{a}_j(\beta_j)}}{\Gamma(\tilde{a}_j(\beta_j))} J(\vec{0};0;\vec{\beta}).$$

The resulting "Type II" MLE, $\vec{\beta}^*$, would then typically be treated as the known $\vec{\beta}$ for the Bayesian analysis.

Example 2.1 (continued). As in Table 2, the first two moments of the prior distribution are 1015 and 1,423,097 for θ_1, and 3000 and $11,749,893$ for θ_2, respectively. Plots (including a contour plot) of $\tilde{L}(\vec{\beta})$ versus β_1 and β_2 are given in Figure 5. The Type-II maximum likelihood estimate is $\vec{\beta}^* = (0.51, 2.02)$, which is close to the actual $\vec{\beta} = (0.50, 2.00)$.

2.6.2 A Fully Bayesian Analysis

When $\pi_2(\vec{\beta})$ is known, the marginal posterior density of $\vec{\beta}$ is

$$\pi(\vec{\beta}|\text{data}) \propto \Big\{\prod_{j=1}^{m} \frac{b_j^{a_j}}{\Gamma(a_j)}\Big\} J(\vec{0};0;\vec{\beta})\pi_2(\vec{\beta}). \tag{34}$$

Define

$$G(\vec{\mu};t) = \int \sum_{(i_1,\cdots,i_m)\in\Delta_{m,n}} W(n;i_1,\cdots,i_m;\vec{\beta}) \prod_{j=1}^{m} \frac{\beta_j^{i_j}\Gamma\Big(a_j + i_j - \frac{\mu_j}{\beta_j}\Big)\pi_2(\vec{\beta})}{(T_{\beta_j} + t^{\beta_j})^{i_j + a_j - \mu_j/\beta_j}} d\vec{\beta}; \tag{35}$$

here, the region of integration is $[c_1,\infty)\times\cdots\times[c_m,\infty)$ (see Section 1.4). Then the marginal posterior density of $\vec{\beta}$ is $\pi(\vec{\beta}|\text{data}) = J(\vec{0};0;\vec{\beta})\pi_2(\vec{\beta})/G(\vec{0};0)$. The marginal posterior moments of θ_j can be computed. For instance, $\text{E}(\theta_1|\text{data}) = G(\vec{0}_{(1)};0)/G(\vec{0};0)$, $\text{Var}(\theta_1|\text{data}) = G(2\vec{0}_{(1)};0)/G(\vec{0};0) - [G(\vec{0}_{(1)};0)/G(\vec{0};0)]^2$, $\text{Cov}(\theta_1,\theta_2|\text{data}) = G(\vec{0}_{(1,2)};0)/G(\vec{0};0) - G(\vec{0}_{(1)};0)G(\vec{0}_{(2)};0)/(G(\vec{0};0))^2$. The predictive reliability (the Bayes estimate of $R(t)$ under squared error loss) is given by $\hat{R}(t) =$

$G(\vec{0};t)/G(\vec{0};0)$. The marginal posterior density of θ_j has the following form:

$$\pi_j(\theta_j|\text{data}) = \frac{1}{G(\vec{0};0)}\int \sum_{(i_1,\cdots,i_m)\in\Delta_{m,n}} W(n;i_1,\cdots,i_m;\vec{\beta})\frac{H_j(\theta_j;\beta_j)\beta_j^{i_j}}{\theta_j^{\beta_j i_j}}\prod_{k\neq j}^{m}\frac{\beta_k^{i_k}\Gamma(a_k+i_k)}{T_{\beta_k}^{a_k+i_k}}\pi_2(\vec{\beta})d\lambda(\vec{\beta}).$$

Evaluation of $G(\vec{\mu};t)$ usually requires numerical integration. Details concerning the complexity of this computation will be given in Section 4.1, when the scheme is compared with Gibbs sampling.

In order to indicate the sensitivity of the posterior quantities to the choice of the $c_j's$, let us look at an example.

Example 2.2 Suppose that $\pi_1(\vec{\theta}|\vec{\beta})$ is as in Example 1.3, and that the data in Example 2.1 is available. Plots (including a contour plot) of $\pi(\vec{\beta}|\text{data})$, the marginal posterior density of $\vec{\beta}$, are given in Figure 6. It can be seen that the posterior mode of (β_1,β_2) is $(0.49, 1.91)$. Furthermore, the joint posterior density is unimodal and almost vanishes near $(0,0)$. Therefore, the posterior quantities are quite robust to the choice of the c_j in this example.

3 Bayesian Analysis via Gibbs Sampling

3.1 Introduction and Notation

Gibbs sampling, which is described generically in the appendix, can be considered as an alternative approach to computation of posterior features, and can even be more efficient than the closed form formulas for larger m and if great accuracy is not needed. In the case of unknown $\vec{\beta}$, Gibbs sampling is especially appealing. Note, however, that the conditional density of β_j does not have a nice form for any nondegenerate prior distribution of $\vec{\beta}$, so that Gibbs sampling here is non-trivial.

We will consider the setup in Sections 1.3 and 1.4, and assume that $\vec{\beta}$ is unknown. For known $\vec{\beta}$, one can just set $\vec{\beta}=\vec{\beta}^o$ in the following Gibbs sampling scheme. Assume that $\beta_1,\cdots,\beta_m$ are independent and that β_j has a log-concave prior density $\pi_{2j}(\beta_j)$, i.e., $\log\{\pi_{2j}(\beta_j)\}$ is concave with respect to β_j. Most commonly used densities are log-concave. For example, if β_j has a Gamma(α,μ) distribution $(\alpha\geq 1)$, then $\pi_{2j}(\beta_j)$ is log-concave. For more details about common log-concave densities, see Table 2 of Gilks and Wild (1990).

In this section, it will be shown that Gibbs sampling can be applied efficiently by introducing several auxiliary random variables. Assume that $t_i=\min(X_{i1},\cdots,X_{im})$, where $X_{i1},\cdots,X_{im}$ $(1\leq$

$i \leq n$) are independent random variables for given $\vec{\theta}$ and $\vec{\beta}$, and $X_{ij} \sim \mathcal{W}(\theta_j, \beta_j)$. For $1 \leq i \leq n$ and $1 \leq j \leq m$, define $I_{ij} = I(t_i = X_{ij})$, where $I(\cdot)$ is the indicator function, and denote $\vec{I}_i = (I_{i1}, \cdots, I_{im})$, $\vec{I} = (\vec{I}_1, \cdots, \vec{I}_n)$, $\vec{I}_{(-i)} = \{\vec{I}_k : 1 \leq k \leq n, k \neq i\}$, $\vec{X}_i = (\vec{X}_{i1}, \cdots, \vec{X}_{im})$, $\vec{X} = (\vec{X}_1, \cdots, \vec{X}_n)$, $\vec{X}_{(-i)} = \{\vec{X}_k : 1 \leq k \leq n, k \neq i\}$, $\vec{\theta}_{(-j)} = \{\theta_k : 1 \leq k \leq m, k \neq j\}$, and $\vec{\beta}_{(-j)} = \{\beta_k : 1 \leq k \leq m, k \neq j\}$.

There are two possible methods of utilizing Gibbs sampling.

1. *Use the indicators $\vec{I} = (\vec{I}_1, \cdots, \vec{I}_n)$ as auxiliary random variables.* Let $x = (\vec{t}, \vec{t}^*)$ and $\xi = (\vec{\theta}, \vec{\beta}, \vec{I}_1, \cdots, \vec{I}_n)$ (see the Appendix), and sample recursively from the conditional distributions $\pi(\theta_j | \vec{t}, \vec{t}^*, \vec{\theta}_{(-j)}, \vec{\beta}, \vec{I})$, $\pi(\beta_j | \vec{t}, \vec{t}^*, \vec{\theta}, \vec{\beta}_{(-j)}, \vec{I})$, and $\pi(\vec{I}_i | \vec{t}, \vec{t}^*, \vec{\theta}, \vec{\beta}, \vec{I}_{(-i)})$.

2. *Use $\vec{X} = (\vec{X}_1, \cdots, \vec{X}_n)$ as auxiliary random variables.* Let $x = (\vec{t}, \vec{t}^*)$ and $\xi = (\vec{\theta}, \vec{\beta}, \vec{X}_1, \cdots, \vec{X}_n)$, and sample recursively from the conditional distributions $\pi(\theta_j | \vec{t}, \vec{t}^*, \vec{\theta}_{(-j)}, \vec{\beta}, \vec{X})$, $\pi(\beta_j | \vec{t}, \vec{t}^*, \vec{\theta}, \vec{\beta}_{(-j)}, \vec{X})$, and $\pi(\vec{X}_i | \vec{t}, \vec{t}^*, \vec{\theta}, \vec{\beta}, \vec{X}_{(-i)})$.

For our problem, the first method is considerably more efficient than the second. In the following subsections, the conditional distributions used in the Gibbs sampler will be listed and studied in the same order that they will be simulated in the algorithm. In particular, the conditional densities of θ_j and $\vec{I}_i$ and their generation are given in Section 3.2, and in Section 3.3 it will be shown that the conditional distribution of β_j is log-cancave. The method of rejection sampling for a univariate log-concave probability density function, utilized to generate the β_j, is given in Appendix A2.

3.2 Available Conditional Distributions

3.2.1 The Conditional Density of θ_j

Theorem 3.1 The conditional density of θ_j, given $(\vec{t}, \vec{t}^*, \vec{\theta}_{(-j)}, \vec{\beta}, \vec{I})$, is

$$\pi(\theta_j | \vec{t}, \vec{t}^*, \vec{\theta}_{(-j)}, \vec{\beta}, \vec{I}) = \frac{\beta_j (S(\beta_j) + b_j)^{a_j + N_j}}{\Gamma(a_j + N_j)} \frac{1}{\theta_j^{1 + \beta_j (a_j + N_j)}} \exp\Big[-\frac{S(\beta_j) + b_j}{\theta_j^{\beta_j}}\Big], \tag{36}$$

where $S(\beta_j)$ is given by (6), and

$$N_j = \sum_{i=1}^{n} I_{ij}. \tag{37}$$

PROOF. Since $\pi(\theta_j|\vec{t},\vec{t^*},\vec{\theta}_{(-j)},\vec{\beta},\vec{I})$ is proportional to

$$\pi_{1j}(\theta_j|\beta_j)\Big[\prod_{i:I_{ij}=1} f_j(t_i|\theta_j,\beta_j)\Big]\Big\{\prod_{i:I_{ij}=0}[1-F_j(t_i|\theta_j,\beta_j)]\Big\}\Big\{\prod_{k=1}^{r-n}\prod_{l=1}^{m}[1-F_l(t_k^*|\theta_l,\beta_l)]\Big\},$$

where

$$f_j(t_i|\theta_j,\beta_j) = \frac{\beta_j t^{\beta_j-1}}{\theta_j^{\beta_j}}\exp\Big\{-\Big(\frac{t_i}{\theta_j}\Big)^{\beta_j}\Big\}, \tag{38}$$

$$1-F_j(t_i|\theta_j,\beta_j) \equiv \int_{t_i}^{\infty} f_j(s|\theta_j,\beta_j)ds = \exp\Big\{-\Big(\frac{t_i}{\theta_j}\Big)^{\beta_j}\Big\}, \tag{39}$$

(36) follows immediately. □

Note that the conditional distribution of $\theta_j^{\beta_j}$, given $(\vec{t},\vec{t^*},\vec{\theta}_{(-j)},\vec{\beta},\vec{I})$, is $\mathcal{IG}(a_j+N_j, S(\beta_j)+b_j)$. Hence, to simulate an observation from the density (36), first, generate a random variable Y from the Gamma$(a_j+N_j,1)$ distribution; then $[(S(\beta_j)+b_j)/Y]^{1/\beta_j}$ has the desired density.

3.2.2 The Conditional Density of $\vec{I}_i$

An argument similar to that in the proof of Theorem 3.1 yields

Theorem 3.2 The conditional density of $\vec{I}_i$ (with respect to counting measure), given $(\vec{t},\vec{t^*},\vec{\theta},\vec{\beta},\vec{I}_{(-i)})$, is

$$\begin{aligned}\pi(\vec{I}_i|\vec{t},\vec{t^*},\vec{\theta},\vec{\beta},\vec{I}_{(-i)}) &= \pi(\vec{I}_i|t_i,\vec{\theta},\vec{\beta}) \\ &= \begin{cases}\beta_1\left[\frac{t_i}{\theta_1}\right]^{\beta_1}\Big/q(t_i;\vec{\theta},\vec{\beta}), & \text{if } \vec{I}_i=(1,0,\cdots,0),\\ \cdots & \cdots \\ \beta_m\left[\frac{t_i}{\theta_m}\right]^{\beta_m}\Big/q(t_i;\vec{\theta},\vec{\beta}), & \text{if } \vec{I}_i=(0,\cdots,0,1),\end{cases}\end{aligned} \tag{40}$$

where $q(x;\vec{\theta},\vec{\beta})=\sum_{k=1}^{m}\beta_k(x/\theta_k)^{\beta_k}$. □

Simulating an observation from the discrete distribution (40) is easy.

3.3 Conditional Densities of the β's

Lemma 3.1 The conditional density of β_j, given $(\vec{t},\vec{t^*},\vec{\theta},\vec{\beta}_{(-j)},\vec{I})$, is log-concave.

PROOF. Note that the conditional density of β_j, given $(\vec{t},\vec{t^*},\vec{\theta},\vec{\beta}_{(-j)},\vec{I})$, is proportional to

$$\pi_{2j}(\beta_j)\pi_{1j}(\theta_j|\beta_j)\Big[\prod_{i:I_{ij}=1} f_j(t_i|\theta_j,\beta_j)\Big]\Big\{\prod_{i:I_{ij}=0}[1-F_j(t_i|\theta_j,\beta_j)]\Big\}\Big\{\prod_{k=1}^{r-n}\prod_{l=1}^{m}[1-F_l(t_k^*|\theta_l,\beta_l)]\Big\},$$

where $f_j(t_i|\theta_j, \beta_j)$ and $1 - F_j(t_i|\theta_j, \beta_j)$ are given by (38) and (39), respectively. Thus

$$\log(\pi(\beta_j|\vec{t}, \vec{t}^*, \vec{\theta}, \vec{\beta}_{(-j)}, \vec{I})) \propto \log(\pi_{2j}(\beta_j)) + \left\{(N_j+1)\log(\beta_j) + C_j(\theta_j)\beta_j - \frac{S(\beta_j)+b_j}{\theta_j^{\beta_j}}\right\}, \quad (41)$$

where N_j and $S(\beta_j)$ are given by (37) and (6), respectively, and $C_j(\theta_j) = \log[(\prod_{i:I_{ij}=1} t_i)/\theta_j^{a_j+N_j}]$. The second derivative of the second term in (41) is

$$-\frac{N_j}{\beta_j^2} - \frac{1}{\beta_j^2} - \frac{\sum_{i=1}^{n}[\log(t_i/\theta_j)]^2 t_i^{\beta_j} + \sum_{l=1}^{r-n}[\log(t_l^*/\theta_j)]^2 (t_l^*)^{\beta_j} + b_j(\log(\theta_j))^2}{\theta_j^{\beta_j}}. \quad (42)$$

This is negative, so that the second derivative of the second term in (41) is log-concave. The result follows from the assumption that $\pi_{2j}(\beta_j)$ is log-concave and the fact that the product of log-concave functions is log-concave. □

An algorithm for sampling from the density $\pi(\beta_j|\vec{t}, \vec{t}^*, \vec{I}, \vec{\theta}, \vec{\beta}_{(-j)})$, based on its log-concavity, is given in Appendix A2.

4 Discussion

4.1 Comparison

In order to judge the efficiency of Gibbs sampling, we compared the closed form computation and the Gibbs sampling scheme theoretically and in our numerical examples.

Comparison for known $\vec{\beta}$.

Consider the situation in Example 2.1. The posterior moments of (θ_1, θ_2), the marginal posterior densities, and the predictive reliability functions are compared in Table 3, Figure 7, and Figure 3, respectively. For the Gibbs sampler, we choose $(M_1, M) = (100, 1000)$, that is, we iterate 1000 times, discarding the first 100 samples. Note that 1000 iterations of the Gibbs sampler seems to be accurate enough, erring by no more than 1%, which is typically quite satisfactory in those situations.

For comparison of the two methods, it follows from (48) and (49) in the Appendix that, if $m << n$, the total number of computations involved in use of the closed form expression and Gibbs sampling are approximately $\frac{2}{(m-1)!}n^m + 3mn^2$ and $(3m+1)nM$, respectively. Thus use of the closed-form expression is recommended if

$$\frac{2}{(m-1)!}n^m + 3mn^2 \leq (3m+1)nM; \quad (43)$$

Table 3: Posterior Moments of (θ_1, θ_2) for a Bi-Weibull Distribution, Given β_1 and β_2

	Exact	Gibbs sampler ($M = 1000$)
$\mathrm{E}(\theta_1 \vert \beta_1, \beta_2, \text{data})$	1002.60	1000.58
$\mathrm{E}(\theta_2 \vert \beta_1, \beta_2, \text{data})$	2466.85	2465.16
$\sqrt{\mathrm{Var}(\theta_1 \vert \beta_1, \beta_2, \text{data})}$	439.74	436.47
$\sqrt{\mathrm{Var}(\theta_2 \vert \beta_1, \beta_2, \text{data})}$	887.46	879.54
$\mathrm{Cov}(\theta_1, \theta_2 \vert \beta_1, \beta_2, \text{data})$	−75823.10	−75618.49
$\mathrm{Corr}(\theta_1, \theta_2 \vert \beta_1, \beta_2, \text{data})$	−0.1943	−0.1970

otherwise Gibbs sampling should prove faster. Table 4 gives the suggested method for several values of m and M.

Table 4 : The Recommend Method of Computation

m	M	Recommended Method	
		Closed-form Expression	Gibbs Sampling
2	1000	if $n \leq 875$	if $n > 875$
3	1000	if $n \leq 95$	if $n > 96$
4	2000	if $n \leq 43$	if $n > 43$
5	2000	if $n \leq 25$	if $n > 25$

Comparison for Unknown $\vec{\beta}$ in Example 2.2

We follow the assumptions in Example 2.2. Exact computation by the fully Bayesian analysis described in Section 2.6.2 and Bayesian analysis via the Gibbs sampler are compared. As mentioned before, exact computation for unknown $\vec{\beta}$ involves numerical integration. A two-dimensional quadrature subroutine from IMSL is used here. Note that evaluation at any single point of the integrand of (35) requires running through the closed form iteration, which takes about 4 minutes on a SUN 3/60 workstation. For the Gibbs sampler, we chose $(M_1, M) = (100, 2000)$, i.e., we iterated 2000 times and discarded the first 100 samples. It takes about one minute to complete 2000 iterations.

The posterior moments of (θ_1, θ_2), as computed by the two methods, are given in Table 5. The corresponding marginal prior and posterior densities are compared in Figure 9. The marginal prior density of θ_j involves only one-dimensional integration, but its marginal posterior density involves two-dimensional integration. We computed the densities at only 60 points and connected the values. Note that 2000 iterations of the Gibbs sampler for unknown (β_1, β_2) seems to be accurate enough, erring by no more than 2.5%, which is quite satisfactory in these situations.

Table 5: Prior and Posterior Moments of (θ_1, θ_2) for a Bi-Weibull Distribution

Prior Moments		Posterior Moments		
			Numerical integration	Gibbs sampling $(M, M_1) = (1900, 100)$
$E(\theta_1)$	1002.91	$E(\theta_1\|\text{data})$	939.31	928.23
$E(\theta_2)$	2004.30	$E(\theta_2\|\text{data})$	2432.15	2396.78
$\sqrt{\text{Var}(\theta_1)}$	1580.61	$\sqrt{\text{Var}(\theta_1\|\text{data})}$	678.463	662.45
$\sqrt{\text{Var}(\theta_2)}$	2826.12	$\sqrt{\text{Var}(\theta_2\|\text{data})}$	1120.51	1092.73
$\text{Cov}(\theta_1, \theta_2)$	0	$\text{Cov}(\theta_1, \theta_2\|\text{data})$	−14938.46	−14338.04
$\text{Corr}(\theta_1, \theta_2)$	0	$\text{Corr}(\theta_1, \theta_2\|\text{data})$	−0.1965	−0.1980

It is necessary to evaluate 250 points or more to reform the two-dimensional integration by numerical integration with enough accuracy. Thus the number of computations needed in this approach is roughly $250(\frac{2}{(m-1)!}n^m + 3mn^2)$. For Gibbs sampling, it is required to obtain only mM additional $\beta_j's$ so that the number of computations remains approximately $(3m+1)nM$. It can indeed be shown that Gibbs sampling, for unknown $\vec{\beta}$, is much fast than the numerical integration approach unless the sample size n is small enough (≤ 7).

It should be mentioned that Gibbs sampling needs substantially more than 2000 iterations when $\vec{\beta}$ is unknown and has a rather vague prior, and the number of observations is small.

4.2 Generalization

The technique of this paper can be used to analyze the case where the component failure times are from the following exponential family:

$$\Big[H'(t)/Q(\theta)\Big]\exp\Big\{-H(t)/Q(\theta)\Big\},$$

where $H(\cdot)$ is an increasing function satisfying $H(0^+) = 0$ and $\lim_{t\to\infty} H(t) = \infty$, and $Q(\cdot)$ is a strictly increasing function. Note that the Weibull and Pareto pdf's are special cases. More details concerning this family can be found in Sun and Berger (1991a, b).

Appendix

A1. The Gibbs Sampling Scheme.

Let $p(\vec{\xi}|x)$ be a general posterior density with $\vec{\xi} = (\xi_1, \cdots, \xi_d) \in \Re_d$. We are interested in $\int f(\vec{\xi})p(\vec{\xi}|x)d\vec{\xi}$. Suppose that the full conditional distributions, $p_i(\xi_i|x, \xi_j,\ j \neq i)$, are available for sampling. Geman and Geman (1984) introduced an algorithm to compute $\int f(\vec{\xi})p(\vec{\xi}|x)d\vec{\xi}$, referred to as the Gibbs sampler. Their algorithm is a Markovian updating scheme and proceeds as follows.

Starting from a set of any initial values $(\xi_1^{(0)}, \xi_2^{(0)}, \cdots, \xi_d^{(0)})$, generate $\xi_1^{(1)}$ from $p_1(\xi_1|x, \xi_2^{(0)}, \cdots, \xi_d^{(0)})$, then $\xi_2^{(1)}$ from $p_2(\xi_2|x, \xi_1^{(1)}, \xi_3^{(0)}, \cdots, \xi_d^{(0)})$, $\cdots$ and so on up to $\xi_d^{(1)}$ from $p_d(\xi_d|x, \xi_1^{(1)}, \cdots, \xi_{d-1}^{(1)})$. Then repeat the process, using $(\xi_1^{(1)}, \xi_2^{(1)}, \cdots, \xi_d^{(1)})$ as the initial values. Continue iterating, ending with $(\xi_1^{(M)}, \xi_2^{(M)}, \cdots, \xi_d^{(M)})$ after M such iterations. Under mild conditions, $\frac{1}{M-M_1}\sum_{j=M_1+1}^{M} f(\xi_1^{(j)}, \xi_2^{(j)}, \cdots, \xi_d^{(j)}) \overset{(M\to\infty)}{\longrightarrow} \int f(\vec{\xi})p(\vec{\xi}|x)d\vec{\xi}$; here, M_1 is the discarded "burn-in" sample size. Furthermore, the marginal density of ξ_k can be estimated by the finite mixture density $\frac{1}{M-M_1}\sum_{j=M_1+1}^{M} p_k(\xi_k|x, \xi_s^{(j)}, s \neq k)$; this is called "Rao-Blackwellization" in Gelfand and Smith (1990). For further discussions of the Gibbs sampler, see Diebold and Robert (1990), Gelfand and Smith (1990), Geyer (1991), Müller (1991), and Tierney (1991a, b).

A2. Sampling from the Conditional Density of β_j

We will present an algorithm for sampling from the density $\pi(\beta_j|\cdot) = \pi(\beta_j|\vec{t}, \vec{t}^*, \vec{I}, \vec{\theta}, \vec{\beta}_{(-j)})$. The algorithm does not require obtaining the supremum of $\pi(\beta_j|\cdot)$. This is an explicit version of the accept-reject algorithm for sampling from a log-concave density (see Devroye (1986)). Since the prior density of β_j is log-concave, the support of β_j should be an interval and is denoted by (c_j, d_j), where $0 < c_j \leq d_j \leq \infty$. Denote the right hand side of (41) by $h(\beta_j)$. Then $\pi(\beta_j|\cdot)$ equals $\exp(h(\beta_j))$, up to a normalization constant.

- **Step 1**: Choose s_1 and s_3 so that $h'(s_1) > 0 > h'(s_3)$. Note that, if the originally selected s_1 has $h'(s_1) \leq 0$, one can simply replace s_1 by $(c_j + s_1)/2$. From Lemma 3.1, an s_1 with

$h'(s_1) > 0$ can be found in a finite number of such steps. Similarly, s_3 satisfying $h'(s_3) < 0$ can be found.

- **Step 2**: Compute the following quantities:

$$\begin{aligned} s_2 &= [h(s_3) - h(s_1) - s_3 h'(s_3) + s_1 h'(s_1)]/[h'(s_1) - h'(s_3)]; \\ u_k &= [h(s_{k+1}) - h(s_k) - s_{k+1} h'(s_{k+1}) + s_k h'(s_k)]/[h'(s_k) - h'(s_{k+1})], \ k = 1, 2; \\ V_k &= \frac{1}{h'(s_k)} e^{h(s_k) - s_k h'(s_k)} \{e^{u_k h'(s_k)} - e^{u_{k-1} h'(s_k)}\}, \quad k = 1, 2, 3, \end{aligned}$$

  where $u_0 \equiv c_j$ and $u_3 \equiv d_j$.

- **Step 3**: Generate a value U from the Uniform (0, 1) distribution and, independently, generate a value X^* from the following piecewise exponential probability density function:

$$p(s) = \tfrac{1}{V_1+V_2+V_3} \sum_{k=1}^{3} \exp\{h(s_k) - s_k h'(s_k) + s h'(s_k)\} I_{[u_{k-1}, u_k)}(s). \tag{44}$$

- **Step 4**: Define an *envelope* function and a *squeezing* function by

$$h_U(s) = \sum_{k=1}^{3} \{h(s_k) + h'(s_k)(s - s_k)\Big\} I_{[u_{k-1}, u_k)}(s); \tag{45}$$

  and

$$h_L(s) = \begin{cases} \frac{h(s_2)-h(s_1)}{s_2-s_1} s + \frac{s_2 h(s_1) - s_1 h(s_2)}{s_2 - s_1}, & \text{if } s_1 \le s \le s_2, \\ \frac{h(s_3)-h(s_2)}{s_3-s_2} s + \frac{s_3 h(s_2) - s_2 h(s_3)}{s_3 - s_2}, & \text{if } s_2 < s \le s_3, \\ -\infty, & \text{otherwise,} \end{cases} \tag{46}$$

  respectively. Compute $h_L(X^*)$ and $h_U(X^*)$. If $U \le \exp(h_L(X^*) - h_U(X^*))$, then let $\beta_j = X^*$. Otherwise, compute $h(X^*)$ and, if $U \le \exp(h(X^*) - h_U(X^*))$, then let $\beta_j = X^*$. If this fails, return to Step 3.

The β_j generated in Step 4 has the desired conditional density, $\pi(\beta_j | \vec{t}, \vec{t}^*, \vec{I}, \vec{\theta}, \vec{\beta}_{(-j)})$. An example of an *envelope* function and a *squeezing* function can be seen from Figure 6.

Remarks.

1. It follows from Lemma 3.1 and the fact that $h'(s_1) > 0 > h'(s_3)$ that $c_j \equiv u_0 < s_1 < u_1 < s_2 < u_2 < s_3 < u_3 \equiv d_j$ and $V_1, V_2, V_3 > 0$.

2. In Step 3, X^* can be generated by the following two stage scheme. Generate a discrete random variable Z based on the density

$$P(Z = k) = V_k/(V_1 + V_2 + V_3), \quad k = 1, 2, 3;$$

if $Z = k$, generate the value of X^* from the density $p_k(s)$, where

$$p_k(s) \quad = \quad \frac{h'(s_k)}{e^{u_k h'(s_k)} - e^{u_{k-1} h'(s_k)}} e^{s h'(s_k)} I_{(u_{k-1}, u_k)}(s).$$

It is easy to see that, if U^* is a $U(0,1)$ random variable, then $\{\log[U^* e^{u_k h'(s_k)} + (1 - U^*) e^{u_{k-1} h'(s_k)}]\}/h'(s_k)$ has the density $p_k(\cdot)$.

A3. The Number of Computations for Known $\vec{\beta}$

We first find the total number of computations involved in use of the closed form expression. The number of terms in the expression for $J(\vec{0}; 0)$ is $\#(\Delta_{m,n})$, which equals $\binom{n+m-1}{m-1}$. Recall that, for each failure t_k in the iteration, one must compute $t_k^{\beta_1}, \cdots, t_k^{\beta_m}$. From the definition of $W(k; i_1, \cdots, i_m; \vec{\beta})$, it follows that one must compute $\sum_{j=1}^m \binom{m}{j}\binom{k-1}{j-1} j$ products and $\sum_{j=1}^m \binom{m}{j}\binom{k-1}{j-1}$ $(j-1)$ sums. From basic combinatorial formulas, these are, respectively,

$$\begin{aligned}
\sum_{j=1}^{m} \binom{m}{j}\binom{k-1}{j-1} j &= m \sum_{j=0}^{m-1} \binom{m-1}{j}\binom{k-1}{j} = m\binom{m+k-2}{m-1}, \\
\sum_{j=1}^{m} \binom{m}{j}\binom{k-1}{j-1} &= \#(\Delta_{m,k}) = \binom{m+k-1}{m-1}.
\end{aligned}$$

Therefore, the number of computations in the updating of $W(k; \cdots)'s$ is $2m\binom{m+k-2}{m-1} - \binom{m+k-1}{m-1} + m$, and the total number computations for finding all $W(k; \cdots)'s$ $(1 \le k \le n)$ is

$$\begin{aligned}
& \sum_{k=1}^{n} \left[2m\binom{m+k-2}{m-1} - \binom{m+k-1}{m-1} + m\right] \\
= \; & 2m \sum_{k=0}^{n-2} \binom{m+k-2}{m-1} - \sum_{k=0}^{n-1} \binom{m+k-1}{m-1} + mn \\
= \; & 2m\binom{m+n-1}{m} - \binom{m+n}{m} + mn. \qquad (47)
\end{aligned}$$

Suppose, now, that we are interested in all the first two posterior marginal moments of θ_j $(1 \le j \le m)$, so that we also need to compute $\Gamma(a_j + k - \mu/\beta_j), \beta_j^k$, and $T_{\beta_j}^{a_j + k - \mu/\beta_j}, j = 1, \cdots, m, k =$

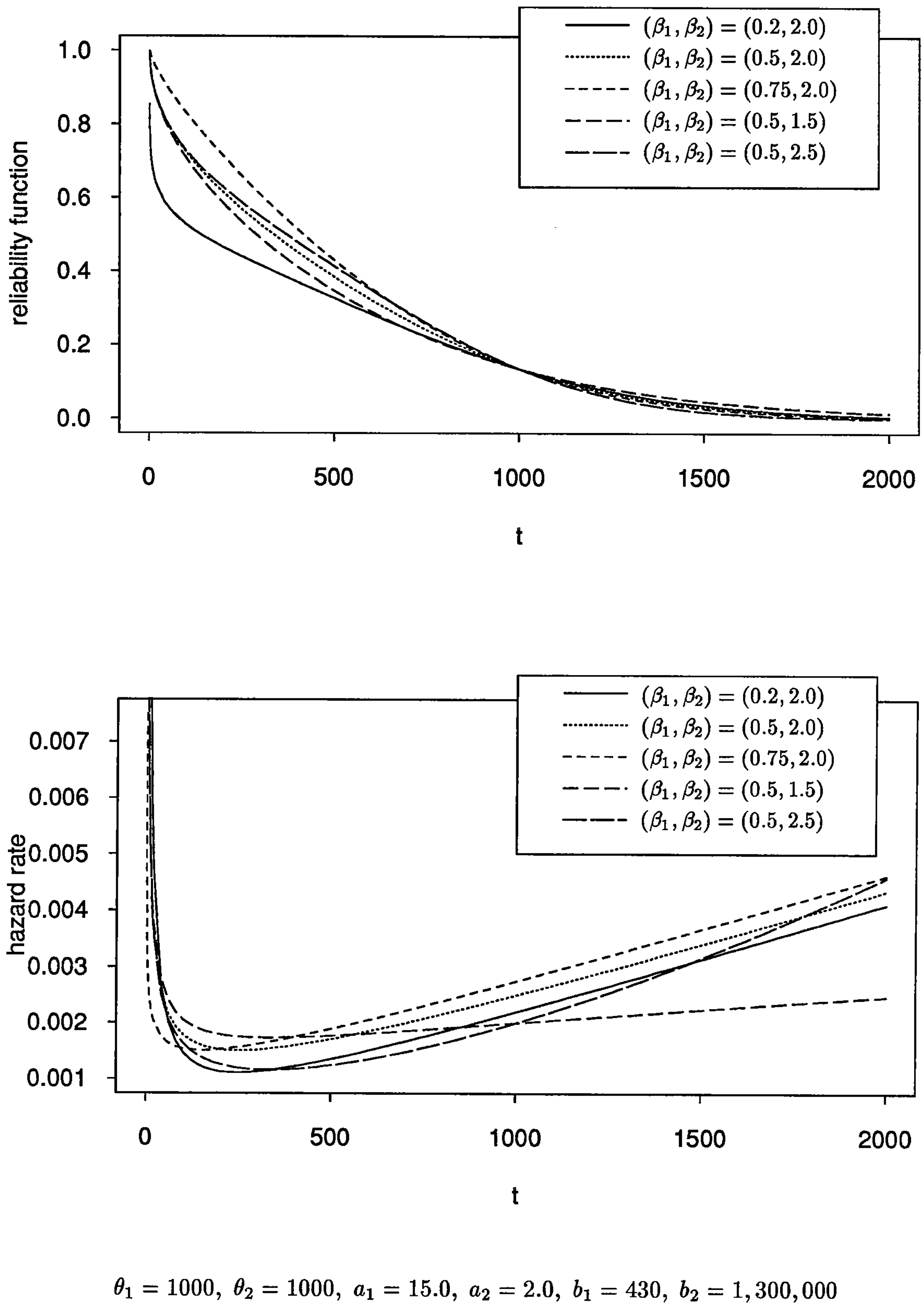
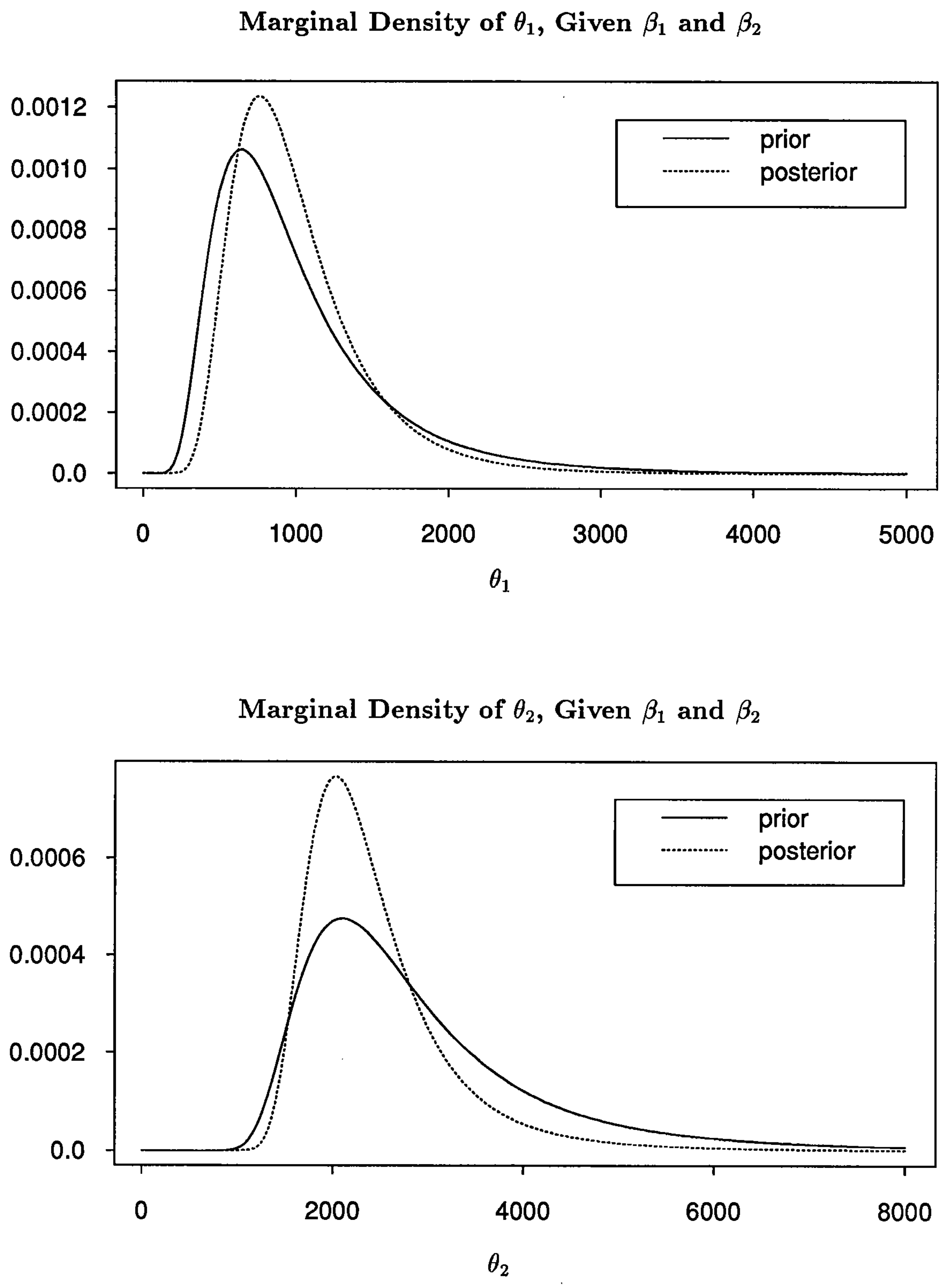
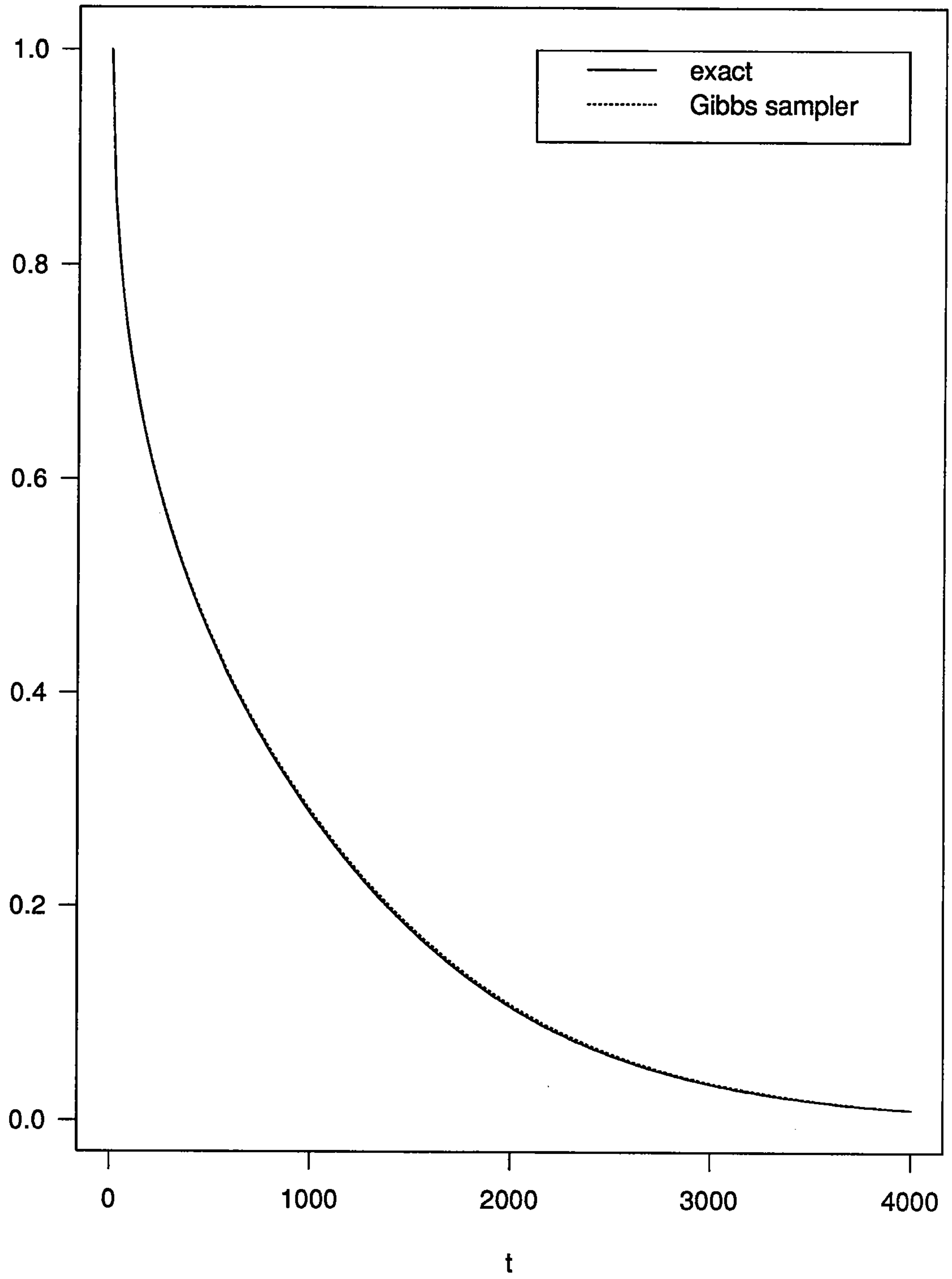


Figure 1: Reliability Funtion and Hazard Rate for Bi-Weibull Distributions



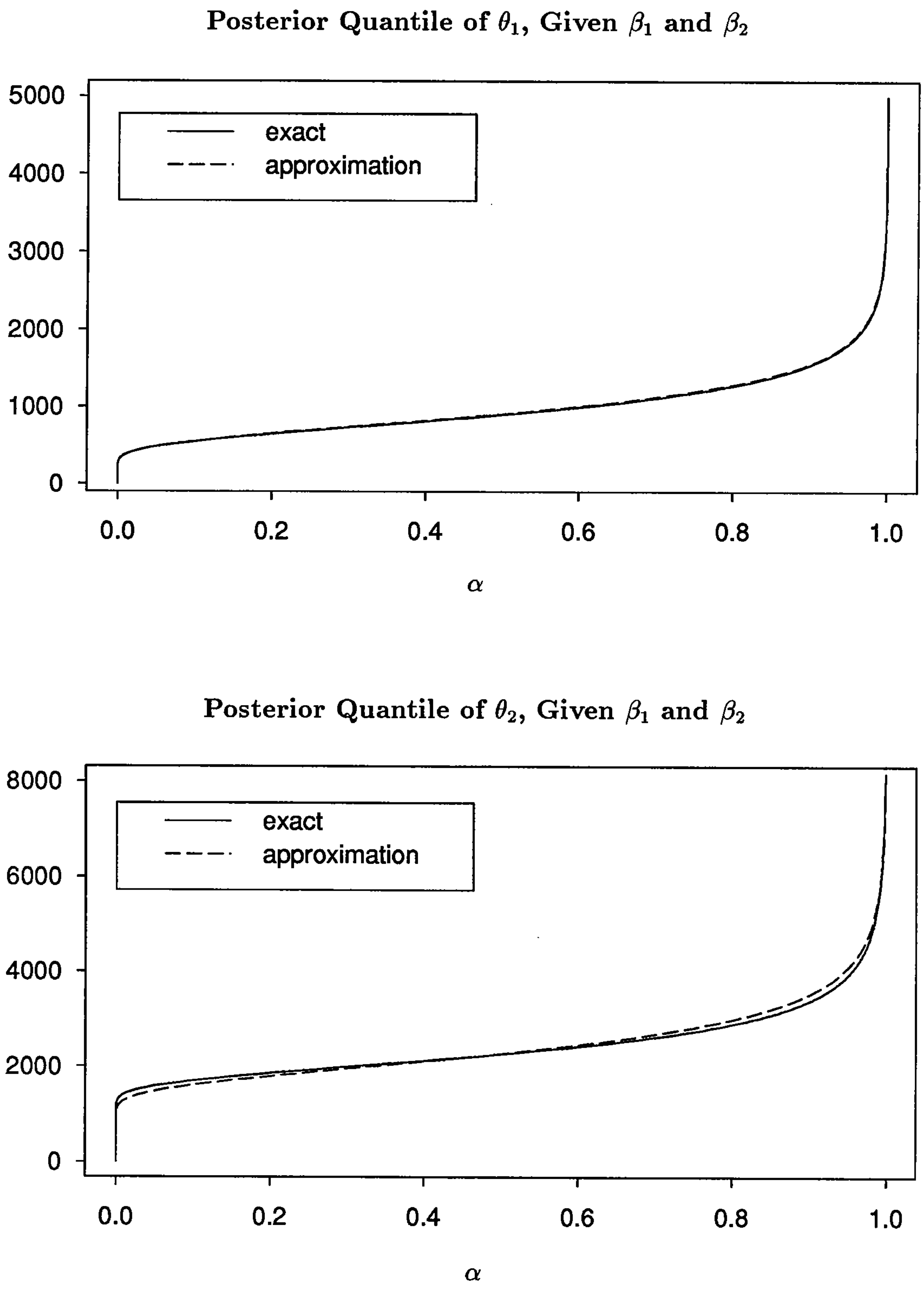
$\beta_1 = 0.5,\ \beta_2 = 2.0,\ a_1 = 15.0,\ a_2 = 1.90,\ b_1 = 430,\ b_2 = 10,575,000$

Figure 2: Marginal Prior and Posterior Densities of θ_1 and θ_2, Given β_1 and β_2



$\beta_1 = 0.5,\ \beta_2 = 2.0,\ a_1 = 15.0,\ a_2 = 1.9,\ b_1 = 430,\ b_2 = 10,575,000$

Figure 3: Predictive Reliability $\hat{R}(t|\beta_1, \beta_2)$



$\beta_1 = 0.5,\ \beta_2 = 2.0,\ a_1 = 15.0,\ a_2 = 1.90,\ b_1 = 430,\ b_2 = 10,575,000$

Figure 4: Posterior Quantiles of θ_1 and θ_2, Given β_1 and β_2

$1, \cdots, n, \mu = 0, 1, 2$. That needs $4mn + 3mn^2$ computations. An additional $3m\binom{n+m-1}{m-1}$ multiplications and $\binom{n+m-1}{m-1}$ sums are required for $J(\vec{0}; 0)$. Similarly for the other $J's$. Therefore the total number of computations needed to determine these moments is

$$
\begin{aligned}
& 2m\binom{m+n-1}{m} - \binom{m+n}{m} + 5mn + 3mn^2 + (2m+1)(3m+1)\binom{n+m-1}{m-1} \\
= \; & \left[2n - \frac{n}{m} - 1 + (2m+1)(3m+1)\right]\binom{n+m-1}{m-1} + 3mn^2 + 5mn. \qquad (48)
\end{aligned}
$$

For example, if $m=3$ and $n=20$, the right hand side of (48) equals 11,799 $<< 3^{20} = 3,486,784,401$, the latter being the number of terms in a brute force expansion of the expression in (23). The recursive formula effectively reduces the total number of computations for a Poly-Weibull distribution from an exponential rate to a polynomial rate. The time for computing a gamma function and a power function are usually 6 times and 3 times that for computing a sum or product or simulating a uniform (0,1) random variable, respectively. But, since the total number of computations for computing sums or products is much larger than that for the Gamma function, we ignore this difference.

For Gibbs sampling, we need to simulate $M(n+m)$ uniform (0,1) random variables. There are then 3mMn and 4mM computations to obtain the $N_j's$ and $\theta_j's$, respectively. In addition, 6mM computations are needed for averages of the $\theta_j's$ or of the $(\theta_j^2)'s$. The total number of computations for Gibbs sampling is thus

$$
(n + 11m + 3mn)M. \qquad (49)
$$

References

[1] Berger, J.O. (1985). *Statistical Decision Theory and Bayesian Analysis, 2nd ed.* Springer-Verlag, New York.

[2] Cox, D.R. (1959). The Analysis of Exponentially Distributed Lifetimes with Two Types of Failures. *J. R. Statist. Soc. B* **21**, 411-421.

[3] Chuck, L., Goodrich, S.M., Hecht, N.L., and Dale E.M. (1990). High Temperature Tensile Strength and Tensile Stress Rupture Behavior of Norton/TRW NT-154 Silcon Nitride. *Ceram. Eng. Sci. Proc.* **11**, 1007-1027.

[4] Devroye, L., (1986). *Non-uniform Random Variate Generation.* Springer-Verlag, New York.

[5] Diebolt, J. and Robert C. (1990). Bayesian Estimation of Finite Mixture Distributions, Part II: Sampling Implementation. Technical Report, Universite Paris VI, L.S.T.A..

[6] 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.

[7] German, S. and German, A. (1984). Stochastic Relaxation, Gibbs Distributions and the Bayesian Restoration of Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **6**, 721-740.

[8] Geyer, C.J. (1991). Markov Chain Monte Carlo Maximum Likelihood, (Submitted).

[9] Gilks, W.R. and Wild, P. (1990). Adaptive Rejection Sampling for Gibbs Sampling, TR UR 90-01, Medical Research Council Biostat Unit, UK Cambridge, CB2 2BW, UK.

[10] Goetghebeur, E. and Ryan, L. (1990). A Modified Log Rank Test for Competing Risks with Missing Failure Type. *Biometrika* **77**, 207-211.

[11] Greenhouse, J.B. and Wolfe, R.A. (1984). A Competing Risks Deviation of a Mixture Model for the Analysis of Survival Data. *Comm. Stat., Part A, Theory and Method* **13**, 3133-3154.

[12] Elandt-Johnson, R.C. and Johnson, N.L. (1980). *Survival Models and Data Analysis.* John Wiley Sons, New York.

[13] Martz, H.F. and Waller, R.A. (1982). *Bayesian Reliability Analysis.* John Wiley Sons, New York.

[14] Mendenhall, W. and Hadel, R.J. (1958). Estimation of Parameters of Mixed Exponentially Distributed Failure Time Distributions from Censored Life Test Data. *Biometrika* **45**, 504-520.

[15] Müller P. (1991) A Generic Approach to Posterior Integration and Gibbs Sampling. Technical Report 91-09, Department of Statistics, Purdue University.

[16] Sun, D. and Berger, J.O. (1991a). Bayesian Sequential Reliability for Weibull and Related Distributions, Technical Report **91-40c**, Department of Statistics, Purdue University.

[17] Sun, D. and Berger, J.O. (1991b). Recent Developments in Bayesian Sequential Reliability Demonstration Testing, Technical Report **201**, Department of Statistics, The University of Michigan.

[18] Tierney, L. (1991a). Exploring Posterior Distributions using Markov Chains, To appear in Proceedings of 1991 Interface.

[19] Tierney, L. (1991b). Markov Chains for Exploring Posterior Distributions. Technical Report 560, School of Statistics, University of Minnesota.

[20] Townsend, A.R. (1989). The Bi-Weibull Reliability Distribution. Technical Report. Indianapolis.

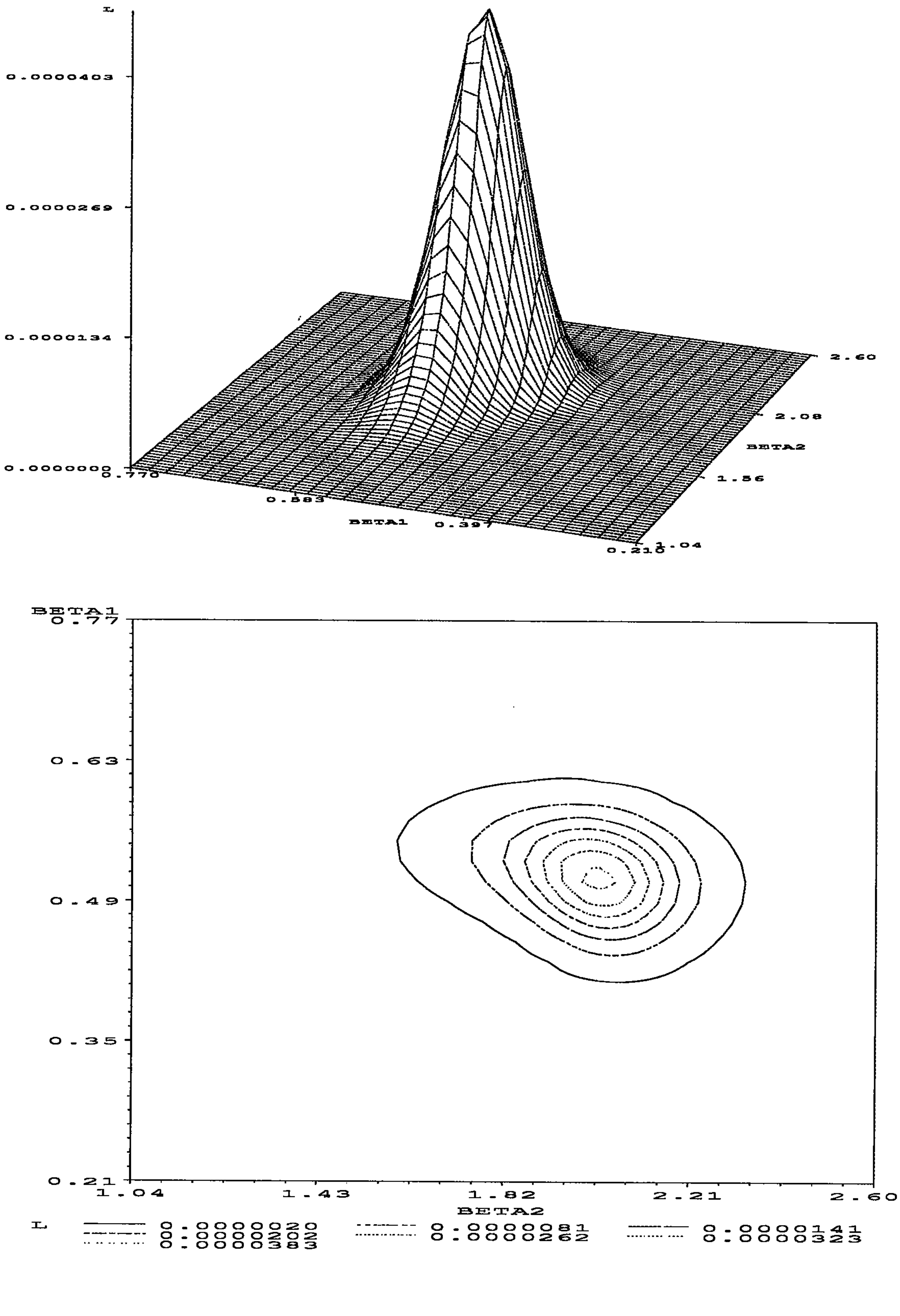
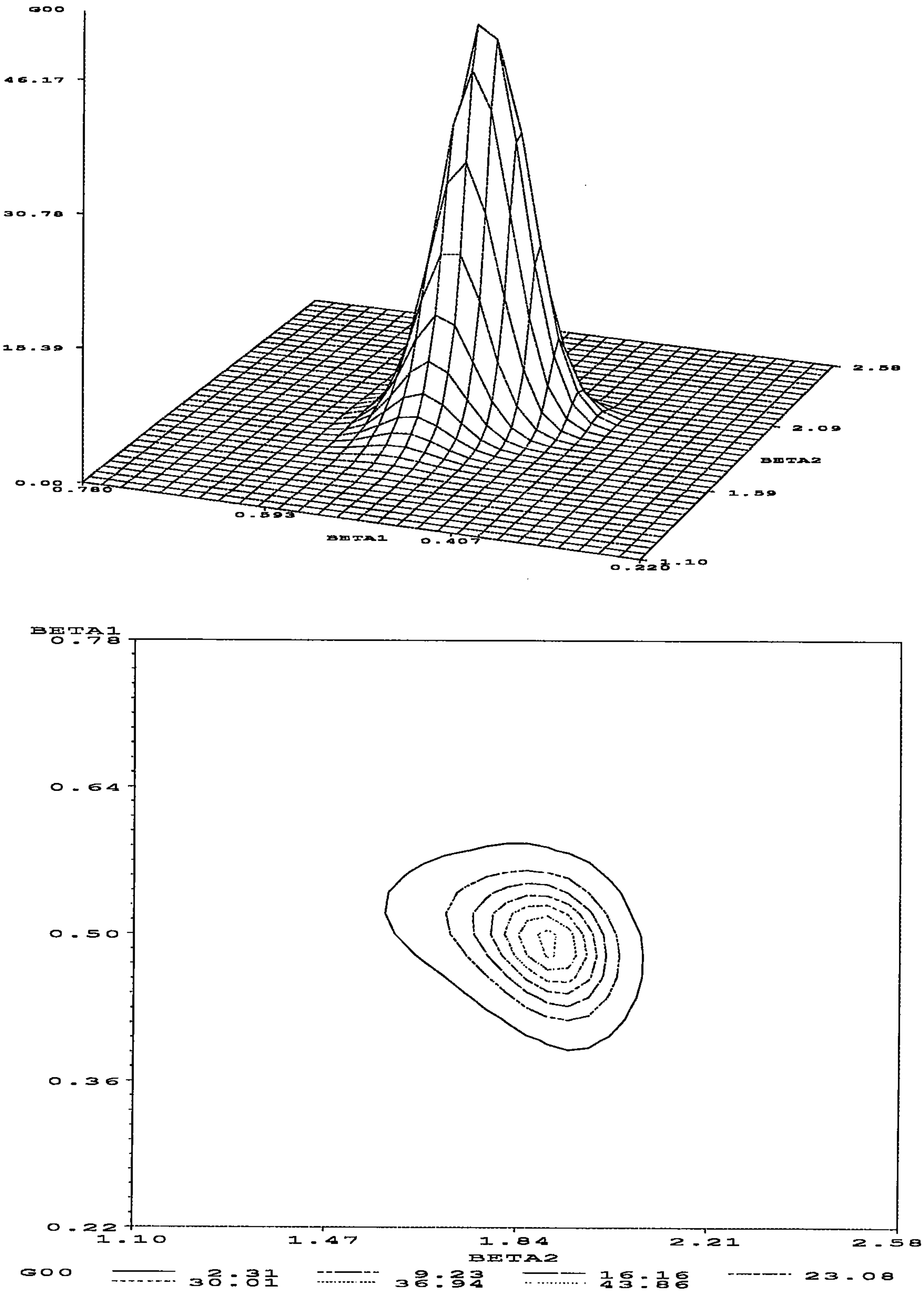
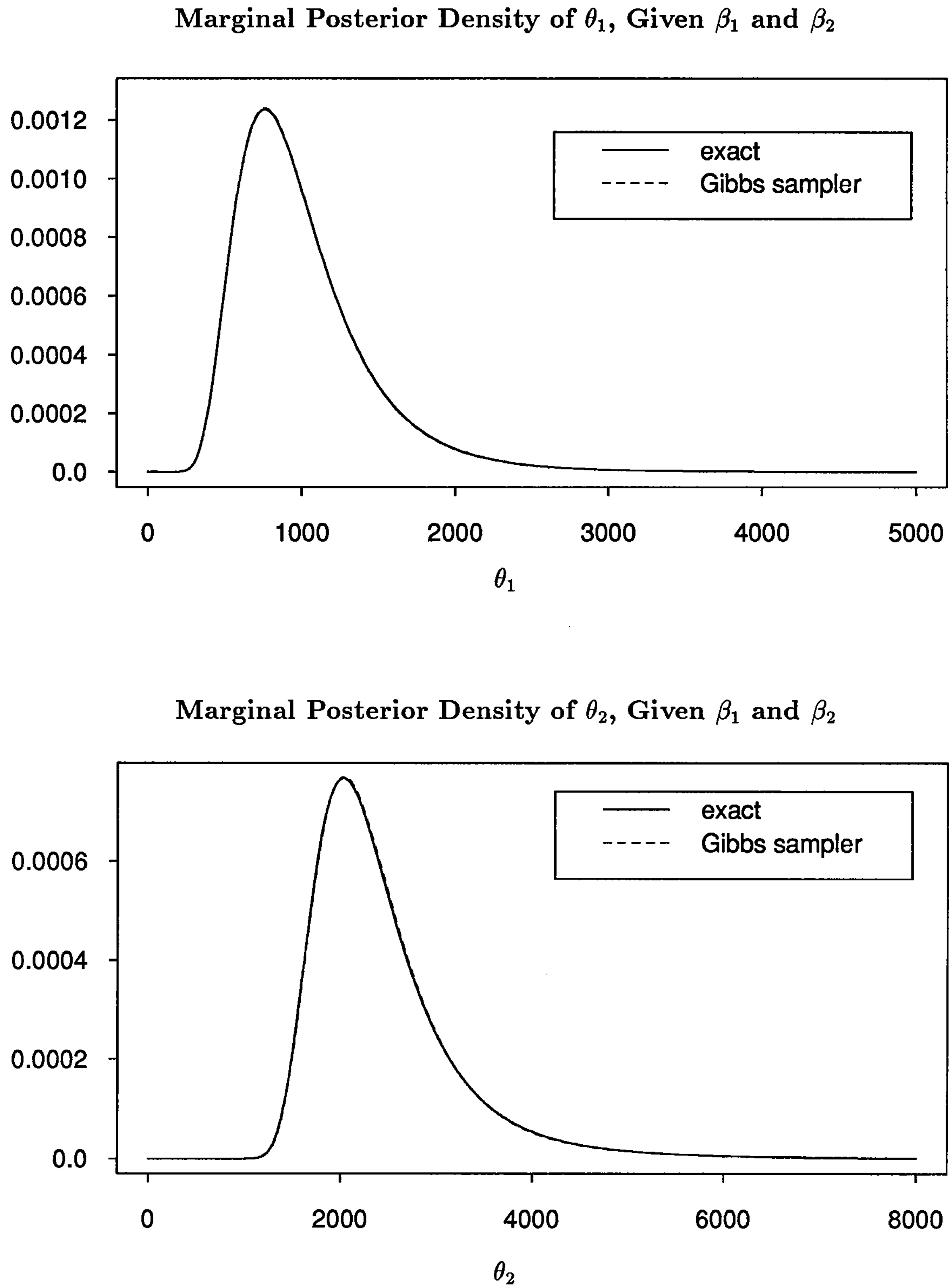


Figure 5: The "Type II" Likelihood Function of (β_1, β_2)



$a_1 = 28.0,\ a_2 = 10.0,\ b_1 = 684,\ b_2 = 18,600,000$

Figure 6: The Marginal Posterior Density Function of (β_1, β_2)



$\beta_1 = 0.5,\ \beta_2 = 2.0,\ a_1 = 15.0,\ a_2 = 1.90,\ b_1 = 430,\ b_2 = 10{,}575{,}000$

Figure 7: Marginal Posterior Densities of θ_1 and θ_2, Given β_1 and β_2

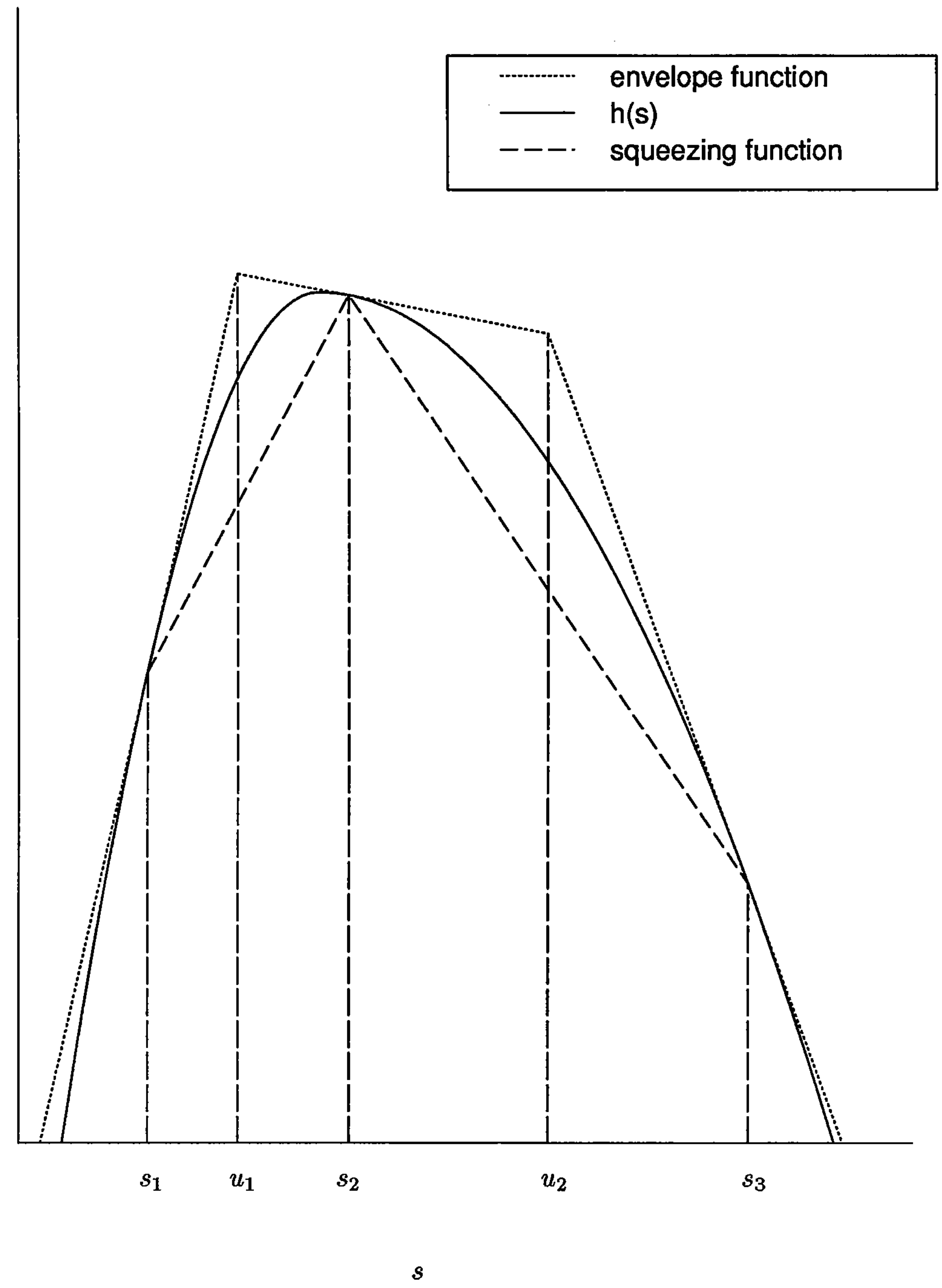
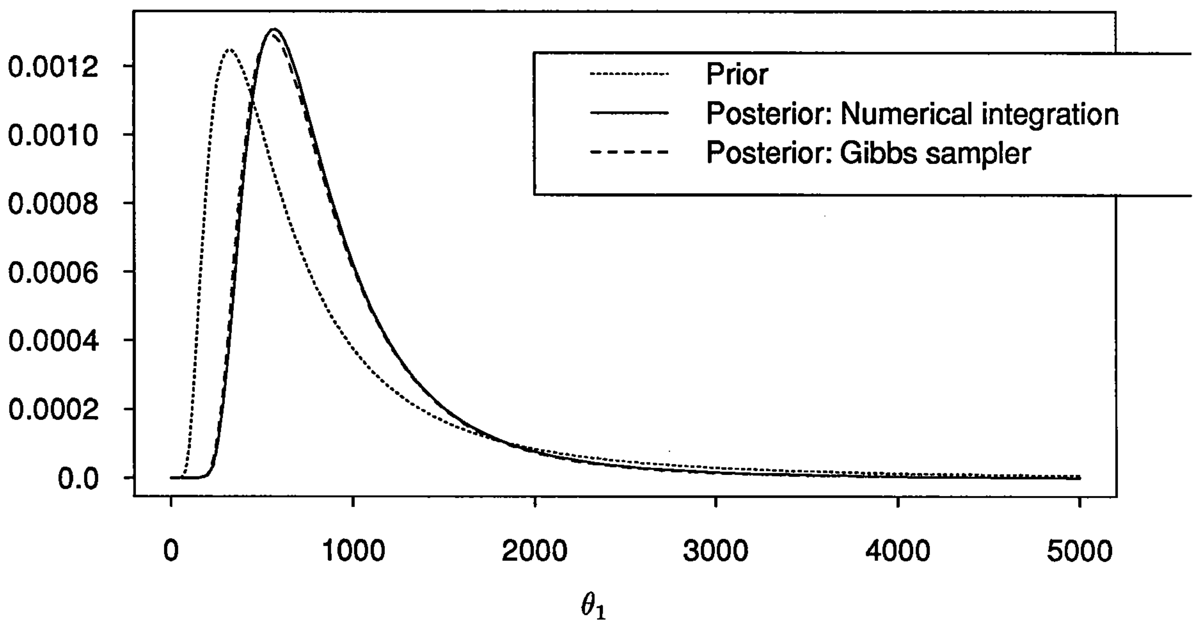
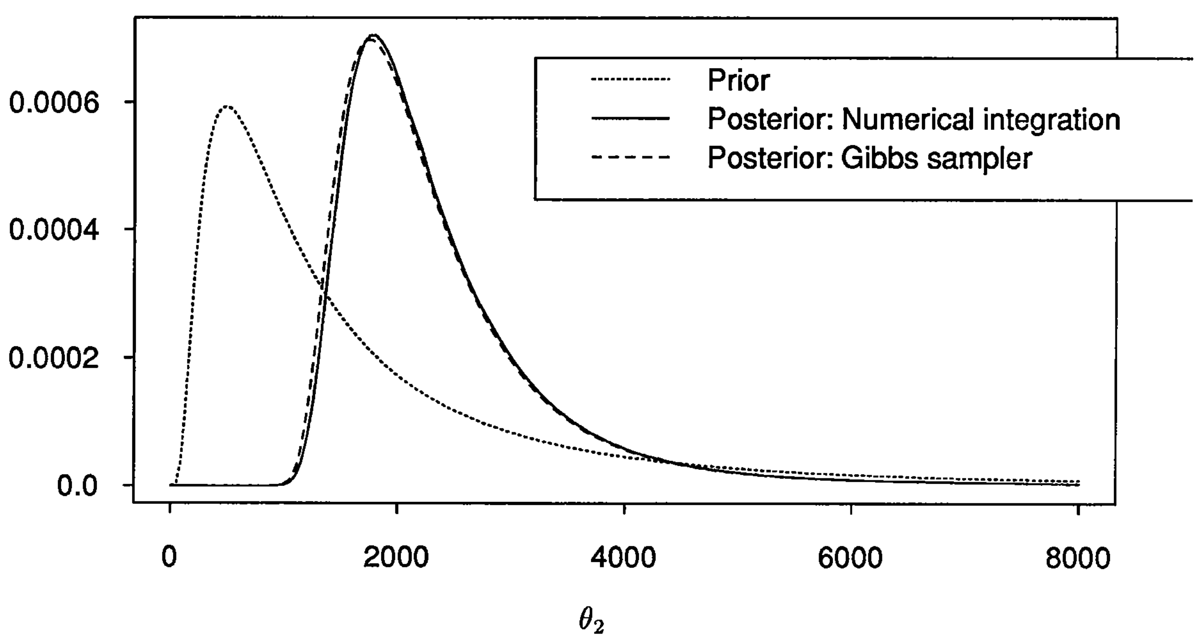


Figure 8: An Example of the Envelope and Squeezing Functions

Marginal Density of θ_1



Marginal Density of θ_2



$a_1 = 28.0,\ a_2 = 10.0,\ b_1 = 684,\ b_2 = 18,600,000$

Figure 9: Marginal Prior and Posterior Densities of θ_1 and θ_2