BAYESIAN INFERENCE FOR A CLASS OF POLY-WEIBULL DISTRIBUTIONS

by

James O. Berger and Dongchu Sun[1]
Department of Statistics, Purdue University
The University of Missouri-Columbis

Technical Report #93-17C

March 1993

James O. Berger is Richard M. Brumfield Distinguished Professor of Statistics, Purdue University, West Lafayette, IN 47907. Dongchu Sun is Assistant Professor, Department of Statistics, University of Missouri-Columbia, Columbia, MO 65211. This research was supported by NSF grants DMS-8702620, DMS-8717799 and DMS-8923071. Sun's work was partially supported by a Research Board Grant from the University of Missouri System.

*

Bayesian Inference for a Class of Poly-Weibull Distributions

James O. Berger and Dongchu Sun [1]

Purdue University and The University of Missouri–Columbia

Abstract

In this paper, Bayesian inference for a class of Poly-Weibull distributions is discussed. These distributions include the distributions with polynomial failure rates, and also arise when the data is the minimum of several Weibull failure times from competing risks (with known Weibull shape parameters). A general recursive formula is developed for exact computation of the posterior probability density function, posterior moments and the predictive reliability.

Keywords: Bi-Weibull distribution, competing risks, Bayesian inference, predictive distribution.

[1]James O. Berger is Richard M. Brumfield Distinguished Professor of Statistics, Purdue University, West Lafayette, IN 47907. Dongchu Sun is Assistant Professor, Department of Statistics, University of Missouri–Columbia, Columbia, MO 65211. This research was supported by NSF grants DMS-8702620, DMS-8717799 and DMS-8923071. Sun's work was partially supported by a Research Board Grant from the University of Missouri System.

1 Introduction

1.1 Background

It is well known that the cumulative distribution function for any positive and continuous random variable from survival data may be expressed as $F(t) = 1 - \exp\left\{-\int_0^t h(x)dx\right\}$, $0 < t < \infty$, where $h(t) = F'(t)/[1 - F(t)] = -R'(t)/R(t)$ is the failure rate function at time t. The function $R(t) = 1 - F(t)$ is the reliability or probability of survival of a product or a patient until time t and the failure rate may be interpreted as the instantaneous failure rate, or the conditional density of failure at time t, given the product has survived until time t.

A constant failure rate leads to the exponential distribution. The Weibull distribution is most commonly used distribution that allows increasing or decreasing failure rates. Unfortunately, even this model is often inadequate, as nonmonotone failure rates are not uncommon in practice.

Many of the commonly considered nonmonotone failure rates have bathtub shapes. For example, Prentice (1975) introduced a generalized F distribution; Hjorth (1980) proposed a three parameter family with increasing, decreasing and bathtub shaped failure rates; Stacy (1982) considered a generalized gamma family; Beetz (1982) and Own, Subramanian and Saunders (1986) considered mixtures of Weibull distributions. A reasonably comprehensive account of the models for bathtub shaped failure rates can be found in Rajarshi and Rajarshi (1988). For the related area of extreme value analysis, the use of a two-component extreme value distribution, e.g, Fiorentino, Versace and Rossi (1985) and Arnell and Gabriele (1988), has appeared in the hydrological literature.

Another possible approach towards developing more flexible models is to generate a class of distributions by considering low order polynomials as the failure rate function. The linear polynomial, $h(t) = a + bt$, generalizes the exponential distribution and has been applied by Kodlin (1967), Vaupel and Yashin (1985), and Bain (1974). This model allows for an initial positive failure rate, $h(t) = a$, while $h(0) = 0$ for most other increasing failure rate situations, such as Weibull distribution. The Rayleigh distribution is also a special case of the linear failure rate model ($a = 0$), and its application to life testing can be found in Polovko (1968). A quadratic failure rate, $h(t) = a + bt + ct^2$, is the lowerest order polynomial which allows for a nonmonotonic failure rate, and has been discussed by Bain (1974, 1978), and Gore *et. al.* (1986). More generally, Krane (1963) has considered a model with failure rate $h(t) = \sum_{j=1}^{m} \alpha_j t^j$, ($\alpha_j \geq 0$) for the analysis of survival data.

In this paper, we will consider a generalization of the above distributions, namely the class of Poly-Weibull distributions whose failure rate functions have the following form:

$$h(t) = \sum_{j=1}^{m} (t/\theta_j)^{\beta_j}, \tag{1}$$

where the $\beta_j's$ are known positive constants and the $\theta_j's$ are unknown constants. This model was first considered by Canfield and Borgman (1975) to model the situation where failure can be associated with an extreme random phenomenon such as the largest flaw or impurity. Many complex failure rate functions can be formed by proper choice of the parameters θ_j and β_j. For example, the Poly-Weibull distributions with $m = 2$ is called the Bi-Weibull distribution. The distributions with linear failure rates are special cases of the Bi-Weibull distribution. A Bi-Weibull distribution with a bathtub failure rate results if $\min(\beta_1, \beta_2) < 1 < \max(\beta_1, \beta_2)$. For $m = 3$, a bathtub failure rate results if $\beta_1 < 1, \beta_2 = 1$ and $\beta_3 > 1$.

The Poly-Weibull distribution can also arise from other scenarios, such as competing Weibull risks. Consider the situation where possible causes of death for individuals in a population are grouped into $m(\geq 2)$ classes of competing risks (cancer, heart disease, etc.). If the survival times under the m risks are assumed to be independently $\mathcal{W}(\theta_i, \beta_i)$ distributed, then the survival time of the patient has a Poly-Weibull distribution. Of course, the population in question need not be a human or animal population but may consist of products liable to m causes or modes of failure. For example, in reliability one might have a system of

m components arranged in series, with failure of one or more components leading to failure of the system. It is frequently reasonable to assume that these m causes affect the product independently, and have (Weibull) $\mathcal{W}(\theta_j, \beta_j)$ distributions. In manufacturing, it is common to encounter failure due either to manufacturing error ("early" failure) or wearout; this leads to the Bi-Weibull distribution. Note that the Poly-weibull distribution arises in these settings only if the cause of failure is unknown. This is often the case in reliability contexts, but is less common in survival analysis. For other references concerning competing risks, see David and Moeschberger (1978), and Basu and Klein (1982).

1.2 The Likelihood Function

The survival function of a product corresponding to the failure rate function (1) is $R(t) = \prod_{j=1}^{m} R_j(t) = \exp\left\{-\sum_{j=1}^{m}\left(t/\theta_j\right)^{\beta_j}\right\},\ t > 0,$ and its density is given by

$$f(t|\beta_j, \theta_j, j = 1, \cdots, m) = \sum_{j=1}^{m} \frac{\beta_j t^{\beta_j - 1}}{\theta_j^{\beta_j}} \exp\left\{-\sum_{k=1}^{m}\left(\frac{t}{\theta_k}\right)^{\beta_k}\right\}, \quad t > 0. \tag{2}$$

To avoid identifiability problems, we assume that the known $\beta_j's$ are all distinct.

Let $\boldsymbol{\theta} = (\theta_1, \cdots, \theta_m)$ and $\boldsymbol{\beta} = (\beta_1, \cdots, \beta_m)$. Assume that r units are tested independently, with their ages having the common p.d.f. (2). 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 then

$$L(\boldsymbol{\theta}) \ = \ \left\{\prod_{i=1}^{n}\sum_{j=1}^{m} \frac{\beta_j t_i^{\beta_j - 1}}{\theta_j^{\beta_j}}\right\}\exp\left\{-\sum_{k=1}^{m} \frac{S(\beta_k)}{\theta_k^{\beta_k}}\right\}, \tag{3}$$

where

$$S(\beta_k) = \sum_{i=1}^{n} t_i^{\beta_k} + \sum_{i=1}^{r-n} (t_i^*)^{\beta_k}. \tag{4}$$

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

1.3 Preview

Expanding the product and summation in (3) 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.

In Section 2, an iterative computational scheme will be developed for closed form Bayesian analysis. 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 $n = 20$, $2n\binom{n+m-1}{m-1} + 3mn^2 = 12,840$.) Some approximations for posterior quantiles of $\theta_1, \cdots, \theta_m$ are also proposed.

A numerical example is given in Section 3, to illustrate the efficiency of the exact computation. Finally, generalization to unknown $\boldsymbol{\beta}$ and more general distributional forms is considered in Section 4.

2 Closed-Form Bayesian Analysis

2.1 Prior Distributions

Assume that $\theta_1, \cdots, \theta_m$ are independent and that the prior density of θ_j is

$$\pi_j(\theta_j) = \frac{\beta_j b_j^{a_j}}{\Gamma(a_j)} \theta_j^{-(1+\beta_j a_j)} \exp\Big\{-\frac{b_j}{\theta_j^{\beta_j}}\Big\}, \tag{5}$$

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 $\boldsymbol{\theta}$, is $\pi(\boldsymbol{\theta}) = \prod_{j=1}^{m} \pi_j(\theta_j)$. Choice of the a_j, b_j will frequently be based on engineering knowledge or knowledge of previous similar products. Methods of eliciting the a_j and b_j can be found from Berger and Sun (1993) and Sun and Berger (1993).

The following notation will prove to be convenient. For $j = 1, \cdots, m$, define

$$T_{\beta_j} = S(\beta_j) + b_j; \tag{6}$$

$$H_j(x) = \frac{\beta_j}{x^{1+\beta_j a_j}} \exp\left\{-T_{\beta_j}/x^{\beta_j}\right\}, \; x > 0, \tag{7}$$

where $S(\beta_j)$ is given by (4).

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

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

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

$$\widetilde{W}(n;i) = \begin{cases} \widetilde{W}(n-1;0)t_n^{\beta_2}, & \text{if } j = 0, \\ \widetilde{W}(n-1;i-1)t_n^{\beta_1} + \widetilde{W}(n-1;i)t_n^{\beta_2}, & \text{if } 1 \leq i < n, \\ \widetilde{W}(n-1;n-1)t_n^{\beta_1}, & \text{if } i = n. \end{cases} \tag{8}$$

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 (8), 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

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

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

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

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

$$\Delta_{m,n} = \Big\{(i_1, \cdots, i_m) : i_j \in \mathcal{N}, \ \sum_{j=1}^{m} i_j = n\Big\}. \tag{9}$$

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

$$W(1; i_1, \cdots, i_m) = \begin{cases} t_1^{\beta_1}, & \text{if } (i_1, i_2, \cdots, i_m) = (1, 0, \cdots, 0), \\ \cdots & \cdots\cdots \\ t_1^{\beta_m}, & \text{if } (i_1, \cdots, i_{m-1}, i_m) = (0, \cdots, 0, 1). \end{cases} \tag{10}$$

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

$$W(k; i_1, \cdots, i_m) = \sum_{i_j \ge 1} W(k-1; i_1, \cdots, i_{j-1}, i_j - 1, i_{j+1}, \cdots, i_m)\, t_k^{\beta_j}. \tag{11}$$

It is easy to see that (11) is well defined. The case $m = 2$ is equivalent to (8).

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 $\boldsymbol{\mu} = (\mu_1, \cdots, \mu_m)$ $(\mu_j < \beta_j(a_j + i_j))$ and $t \ge 0$, let

$$J(\boldsymbol{\mu}; t) = \sum_{(i_1,\cdots,i_m)\in\Delta_{m,n}} W(n; i_1, \cdots, i_m) \prod_{j=1}^{m} \frac{\beta_j^{i_j}\Gamma\left(a_j + i_j - \frac{\mu_j}{\beta_j}\right)}{(T_{\beta_j} + t^{\beta_j})^{i_j + a_j - \mu_j/\beta_j}}. \tag{12}$$

2.3.2 The Poly-Weibull Posterior Distribution and Moments

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

$$\pi(\boldsymbol{\theta}) = \frac{1}{J(\mathbf{0}; 0)} \sum_{(i_1,\cdots,i_m)\in\Delta_{m,n}} W(n; i_1, \cdots, i_m) \prod_{k=1}^{m} H_k(\theta_k) \left\{ \frac{\beta_k}{\theta_k^{\beta_k}} \right\}^{i_k}, \tag{13}$$

and the marginal posterior density of θ_j $(1 \le j \le m)$ is

$$\pi_j(\theta_j|\text{data}) = \frac{H_j(\theta_j)}{J(\mathbf{0}; 0)} \sum_{(i_1,\cdots,i_m)\in\Delta_{m,n}} W(n; i_1, \cdots, i_m) \left(\frac{\beta_j}{\theta_j^{\beta_j}} \right)^{i_j} \prod_{k \ne j}^{m} \frac{\beta_k^{i_k}\Gamma(a_k + i_k)}{T_{\beta_k}^{a_k + i_k}}. \tag{14}$$

PROOF. The posterior density of $\boldsymbol{\theta}$ is proportional to

$$\left\{ \prod_{i=1}^{n} \sum_{k=1}^{m} \frac{\beta_k t_i^{\beta_k}}{\theta_k^{\beta_k}} \right\} \left\{ \prod_{k=1}^{m} \frac{\beta_k}{\theta_k^{1+\beta_k a_k}} \right\} \exp\left\{ -\sum_{k=1}^{m} \frac{T_{\beta_k}}{\theta_k^{\beta_k}} \right\}. \tag{15}$$

By induction, it can be shown that (15) is equal to

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

Since

$$\int_0^\infty \theta_k^y H_k(\theta_k)\theta_k^{-\beta_k i_k}d\theta_k = \Gamma(a_k+i_k-\frac{y}{\beta_k})/T_{\beta_k}^{i_k+a_k}, \tag{16}$$

the normalization constant for the joint posterior density of $\boldsymbol{\theta}$ is $J(\mathbf{0};0)$. 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

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

Theorem 2.2 The posterior moments are

$$\begin{aligned}
\mathrm{E}(\theta_j|\text{data}) &= \frac{J(\mathbf{0}_{(1)};0)}{J(\mathbf{0};0)},\\
\mathrm{Var}(\theta_1|\text{data}) &= \frac{J(2\mathbf{0}_{(j)};0)}{J(\mathbf{0};0)}-\Big[\frac{J(\mathbf{0}_{(j)};0)}{J(\mathbf{0};0)}\Big]^2,\\
\mathrm{Cov}(\theta_i,\theta_j|\text{data}) &= \frac{J(\mathbf{0}_{(i,j)};0)}{}-\frac{J(\mathbf{0}_{(i)};0)J(\mathbf{0}_{(j)};0)}{(J(\mathbf{0};0))^2},
\end{aligned}$$

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

The simple proof of Theorem 2.2 is omitted.

The number of terms in the expression for $J(\mathbf{0};0)$ is $\#(\Delta_{m,n})$, which equals $\binom{n+m-1}{m-1}$. In the Appendix 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 $\pi(\theta_j|\text{data})$, 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 (5) which has moments m_j and V_j. Thus the approximation is

$$\pi_j(\theta_j|\text{data}) \approx \frac{\beta_j \tilde{b}_j^{\tilde{a}_j}}{\Gamma(\tilde{a}_j)}\, \theta_j^{-(1+\beta_j\tilde{a}_j)} \exp\Big\{-\frac{\tilde{b}_j}{\theta_j^{\beta_j}}\Big\}, \tag{17}$$

where $(\tilde{a}_j, \tilde{b}_j)$ satisfy the following two equations that define the moments of (17):

$$\begin{cases} \Gamma(\tilde{a}_j - 2/\beta_j)\Gamma(\tilde{a}_j)/\Gamma^2(\tilde{a}_j - 1/\beta_j) = 1 + V_j/m_j^2, \\ \tilde{b}_j = \big[m_j\Gamma(\tilde{a}_j)/\Gamma(\tilde{a}_j - 1/\beta_j)\big]^{\beta_j}. \end{cases} \tag{18}$$

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

$$\tilde{a}_j = 0.5 + \Big[\ln(1 + \frac{V_j}{m_j^2}) + \frac{2}{\beta_j}\ln\Big(1 - \frac{1}{\tilde{a}_j\beta_j - 1}\Big)\Big]\Big/\ln\Big[1 - \frac{1}{(\tilde{a}_j\beta_j - 1)^2}\Big], \tag{19}$$

starting from the initial value

$$\tilde{a}_j^0 = \frac{1}{2\beta_j}\Big(3 + \frac{1}{1 - (1 + \frac{V_j}{m_j^2})^{-0.5\beta_j}}\Big). \tag{20}$$

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 (17), which can be shown to be

$$\hat{q}_j(\alpha) = \Big[\frac{2\tilde{b}_j}{\chi^2_{2\tilde{a}_j}(1-\alpha)}\Big]^{1/\beta_j} = \frac{m_j\Gamma(\tilde{a}_j)}{\Gamma(\tilde{a}_j - 1/\beta_j)}\Big[\frac{2}{\chi^2_{2\tilde{a}_j}(1-\alpha)}\Big]^{1/\beta_j}, \tag{21}$$

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

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 is

$$\widehat{R}(t) \equiv P(T > t|\text{data}) = \int_{\mathbb{R}_m^+} R(t)\pi(\boldsymbol{\theta}|\text{data})d\boldsymbol{\theta}, \quad t > 0,$$

where $R(t) = \exp\{-\sum_{j=1}^{m}(t/\theta_j)^{\beta_j}\}$ and $\mathbb{R}_m^+ = \{(y_1, \cdots, y_m) : y_j > 0\}$. (Under squared error loss, $\widehat{R}(t)$ is the Bayes estimate of $R(t)$, given $\boldsymbol{\beta}$.)

Theorem 2.3 $\widehat{R}(t) = J(\mathbf{0}; t)/J(\mathbf{0}; 0)$, for $t > 0$.

PROOF. Note that

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

The result immediately follows from (16). □

3 A Numerical Example

Suppose that failure time of a product has the Bi-Weibull distribution with $\beta_1 = 0.5$ and $\beta_2 = 2.0$. Assume that the hyperparameters for the prior distributions (5) are $a_1 = 15.0, a_2 = 1.90, b_1 = 430$ and $b_2 = 10,575,000$. (This would correspond to subjectively specified prior means and variances of 1016 and 390963 for θ_1, and 3000 and 2,749,892 for θ_2; these could be prior opinions for, say, failure due to manufacturing error and wearout, respectively.) The following sample of size 20 ($n = r = 20$) is drawn 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.

Marginal Posterior Density and Moments

The marginal prior densities (solid curves) and marginal posterior densities (dotted curves) of θ_1 and θ_2 are given in Figure 1. The posterior moments are given in Table 2. The predictive reliability is shown in Figure 2.

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

Prior Distribution		Posterior Distribution	
$\mathrm{E}(\theta_1)$	1015.94	$E(\theta_1 \mid \text{data})$	1002.60
$\mathrm{E}(\theta_2)$	3000.02	$E(\theta_2 \mid \text{data})$	2466.85
$\sqrt{\mathrm{Var}(\theta_1)}$	625.27	$\sqrt{\mathrm{Var}(\theta_1 \mid \text{data})}$	439.74
$\sqrt{\mathrm{Var}(\theta_2)}$	1658.28	$\sqrt{\mathrm{Var}(\theta_2 \mid \text{data})}$	887.46
$\mathrm{Cov}(\theta_1, \theta_2)$	0	$\mathrm{Cov}(\theta_1, \theta_2 \mid \text{data})$	-75823.10
$\mathrm{Corr}(\theta_1, \theta_2)$	0	$\mathrm{Corr}(\theta_1, \theta_2 \mid \text{data})$	-0.1943

Posterior Quantiles

From Table 2, $(m_1, V_1) = (1002.60, 439.74^2)$ and $(m_2, V_2) = (2466.85, 887.46^2)$. Solving the equations in (18), using (19) and (20), yields $(\tilde{a}_1, \tilde{b}_1) = (25.2535,\ 751.9596)$ and $(\tilde{a}_2, \tilde{b}_2) = (2.9893,\ 13,709,647)$. Figure 3 indicates the quality of the quantile approximations. The true quantiles and their approximations (21) are represented by the solid and dashed lines, respectively. There are no noticeable differences between the true quantiles and their approximations.

4 Generalizations

4.1 Unknown $\boldsymbol{\beta}$

When $\boldsymbol{\beta}$ is unknown, one could, of course, perform a full Bayesian analysis by specifying a prior distribution for $\boldsymbol{\beta}$. The resulting analysis is computationally very intensive, however, usually requiring Gibbs sampling; see Berger and Sun (1993). Another possibility is to estimate $\boldsymbol{\beta}$ by likelihood methods, and use the resulting estimate in our previous formulas.

Type II maximum likelihood is typically the preferred method in these setting. Thus we first integrate out $\boldsymbol{\theta}$, using the prior $\pi(\boldsymbol{\theta}|\boldsymbol{\beta})$ in (5), obtaining the marginal likelihood

$$L^*(\boldsymbol{\beta}) \equiv \int_{\mathbf{R}_m^+} L(\boldsymbol{\theta},\boldsymbol{\beta})\pi(\boldsymbol{\theta}|\boldsymbol{\beta})d\boldsymbol{\theta} = \prod_{j=1}^{m} \frac{b_j(\beta_j)^{a_j(\beta_j)}}{\Gamma(a_j(\beta_j))} J(\mathbf{0};0). \tag{22}$$

In (22), we have written a_j and b_j as functions of β_j, since typically one would desire to have $\pi(\boldsymbol{\theta})$ preserve moments or other subjectively specified features of the prior information for $\boldsymbol{\theta}$. In the case of moment specifications, (18) would define the functions a_j and b_j. Again, however, it is far simpler (and quite accurate) to use the approximation (20), i.e.,

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

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

Then the Type II MLE can be obtained by maximizing

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

For the example in Section 3, plots of $\tilde{L}(\boldsymbol{\beta})$ are given in Figure 4. The Type II maximum likelihood estimate is $\boldsymbol{\beta}^* = (0.51, 2.02)$; note that this is quite close to the true $\boldsymbol{\beta} = (0.50, 2.00)$.

4.2 The Exponential Family

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. Details concerning this family can be found in Sun and Berger (1993). More generally, the methods here apply to the following failure rate model:

$$h(t) = \sum_{i=1}^{m} H_i(t)/Q(\theta_i).$$

This includes the following useful models:

(i) Let $h(t) = \theta_1 + \theta_2 t + \theta_3/t$. This model is a generalization of gamma and truncated normal densities. An application of the model can be found in Glaser (1980) and Cobb *et. al.* (1983).

(ii) Let $h(t) = \theta_1/(1+t) + \theta_2 t^{\beta-1}$, for some known $\beta > 2$. See Ranjarshi and Ranjarshi (1988).

(iii) Let $h(t) = \theta_1/(\alpha + t) + \theta_2 t + \theta_3/t$, for some known $\alpha > 0$, considered by Gavar and Acar (1979).

(iv) Let $h(t) = \theta_1 e^{\mu t} + \theta_2/(1+t)$. This model yields a stochastic failure rate version of Makeham's curve. For details, see Ranjarshi and Ranjarshi (1988).

Appendix. The Number of Computations

We now find the total number of computations involved in use of the closed form expression. The number of terms in the expression for $J(\mathbf{0}; 0)$ 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 $W(k; i_1, \cdots, i_m)$, 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

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

$$= \quad 2m\binom{m+n-1}{m} - \binom{m+n}{m} + mn. \tag{23}$$

Suppose, now, that we are interested in all the first two posterior marginal moments of θ_j $(1 \leq j \leq m)$, so that we also need to compute $\Gamma(a_j + k - \mu/\beta_j), \beta_j^k$, and $T_{\beta_j}^{a_j+k-\mu/\beta_j}, j = 1, \cdots, m, k = 1, \cdots, n, \mu = 0, 1, 2$. 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 $J(\mathbf{0}; 0)$. Similarly for the other $J's$. Therefore the total number of computations needed to determine these moments is

$$\begin{aligned} & 2m\binom{m+n-1}{m} - \binom{m+n}{m} + 5mn + 3mn^2 + (2m+1)(3m+1)\binom{n+m-1}{m-1} \\ = \quad & \left[2n - \frac{n}{m} - 1 + (2m+1)(3m+1)\right]\binom{n+m-1}{m-1} + 3mn^2 + 5mn. \end{aligned} \tag{24}$$

For example, if $m = 3$ and $n = 20$, the right hand side of (24) 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 (15). 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.

References

[1] Arnell, N.W. and Gabriele, S. (1988). The performance of the two-component extreme value distribution in regional flood frequency analysis. *Water Resources Research*, 24, 879–887.

[2] Bain, L.J. (1974). The null distribution for a test of constant versus BT failure rate. *Technometrics*, 16, 551–560.

[3] Bain, L.J. (1978). *Statistical Analysis of Reliability and Life-Testing Models*, Marcel Dekker, Inc., New York.

[4] Basu, A. and Klein, J. (1982). *Some Recent Development in Competing Risks Theory.* In *Survival Analysis*, ed. by Crowley and Johnson, IMS, Hayward, CA, 216–229.

[5] Beetz, C.P. (1982). The Analysis of carbon fibre strength distributions exhibiting multiple modes of failure. *Fibre Science and Technology*, 16, 45–59.

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

[7] Berger, J.O. and Sun, D. (1993). Bayesian analysis for the Poly-Weibull distribution. To appear in *J. Amer. Statistist. Assoc.*

[8] Canfield, R.V. and Borgman, L.E. (1975). Some distributions of time to failure for reliability applications. *Technometrics*, 17, 263–268.

[9] Cobb, L., Koppstein, P. and Chen, N.H. (1983). Estimation amd moment recursion relations for multimodal distributions of the exponential families. *J. Amer. Statistist. Assoc.*, 78, 124–130.

[10] Cox, D.R. (1959). The analysis of exponentially distributed lifetimes with two types of failures. *J. R. Statist. Soc. B*, **21**, 411-421.

[11] 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 silicon nitride. *Ceram. Eng. Sci. Proc.*, **11**, 1007-1027.

[12] David, H.A. and Moeschberger, M.L. (1978). *The Theory of Competing Risks*, Griffin's Statistical Monographs & Courses No. 39, Charles Griffin and Company LTD.

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

[14] Fiorentino, M., Versace, P., and Rossi, F. (1985). Regional flood frequency estimation using the two–component extreme value distribution. *Hydrological Sciences Journal*, 30, 51–60.

[15] Gavar, D.P. and Acar, M. (1979). Analytical hazard representation for use in reliability, mortality and simulation studies. *Commun. Statist.– Simula. Computa.*, B8 (2), 91–111.

[16] Glaser, R.E. (1980). Bathtub and related failure rate characterization. *J. Amer. Statistist. Assoc.*, 75, 667–672.

[17] Gore, A.P., Paranjpe, S.A., Rajarshi, M.B. and Gadgil, M. (1986). Some methods for summarizing survivalship data in non-standard situations. *Biom. J.*, 28, 577-586.

[18] Hjorth, U. (1980). A reliability distribution with increasing, decreasing, constant and bathtub shaped failure rates. *Technometrics*, 22, 99–107.

[19] Kodlin, D. (1967). A new response time distribution. *Biometrics*, 2, 227–239.

[20] Krane, S.A. (1963). Analysis of survival data by regression technique. *Technometrics*, 5, 161–174.

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

[22] 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.

[23] Own, S.H., Subramanian, R.V. and Saunders, S.C. (1986). A bimodal lognormal model of the distribution of strength of carbon fibres: effects of electrodeposition of titanium DI (dioctyl pyrophosphate) oxyacetate. *Journal of Materials Science*, 21, 3912–3920.

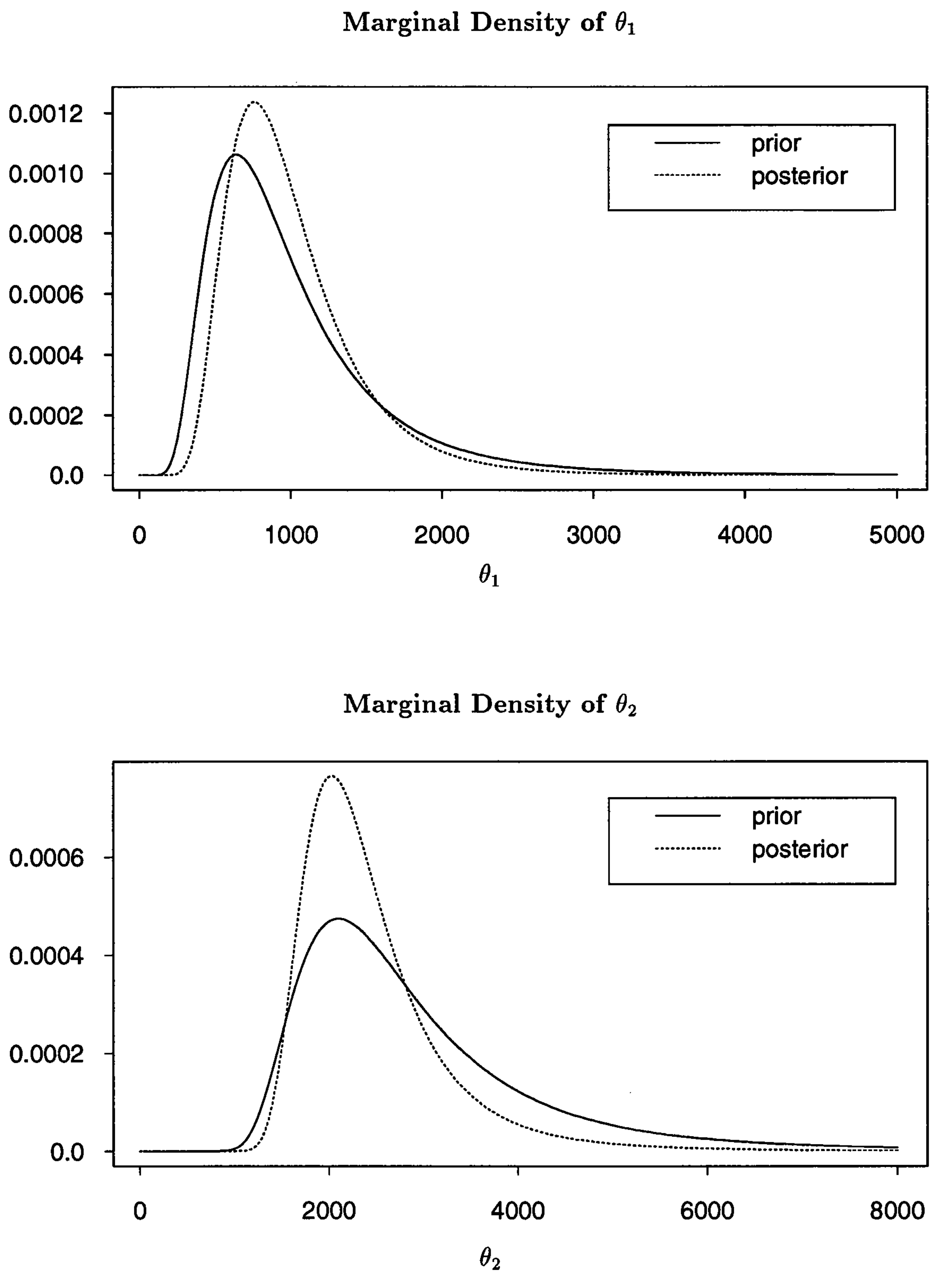
[24] Prentice, R.L. (1975). Discrimination among some parametric models. *Biometrika*, 62, 607–614.

[25] Polovko, A.M. (1968). *Fundamentals of Reliability Theory.* Academic Press, New York.

[26] Rajarshi, S., and Rajarshi, M.B. (1988). Bathtub distribution: a review. *Communications in Statistics, Part – A, Theory and Methods*, 17 (8), 2597–2621.

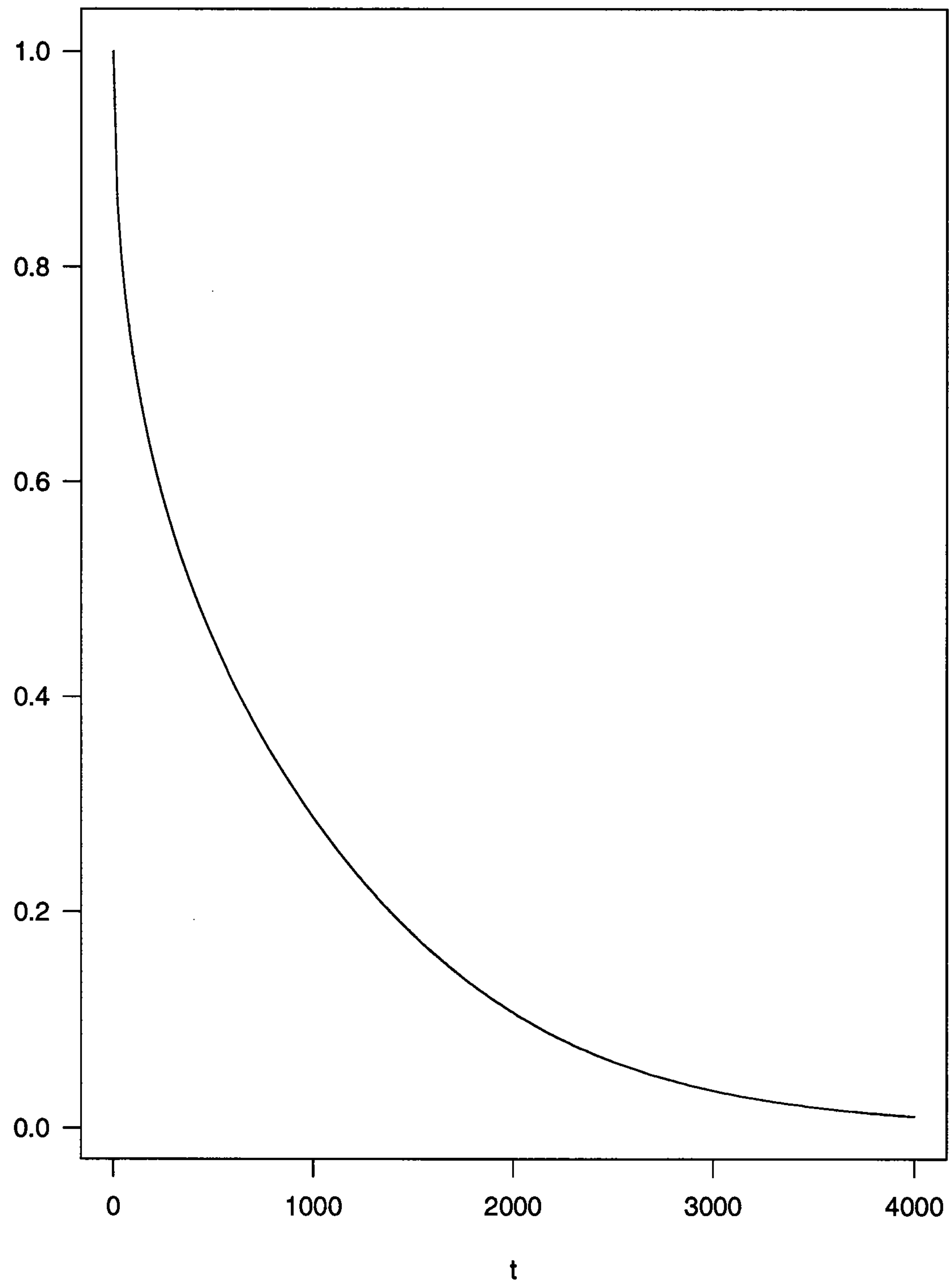
[27] Sun, D. and Berger, J.O. (1993). Recent developments in Bayesian sequential reliability demonstration tests. In *Advances in Reliability,* ed. by Basu, A., Noth-Holland, Amsterdam, 379-394.

[28] Vaupel, J.W. and Yashin, A.I. (1985). Some surprising effects of selection on population dynamics. *American Statistician*, 39, 176–184.



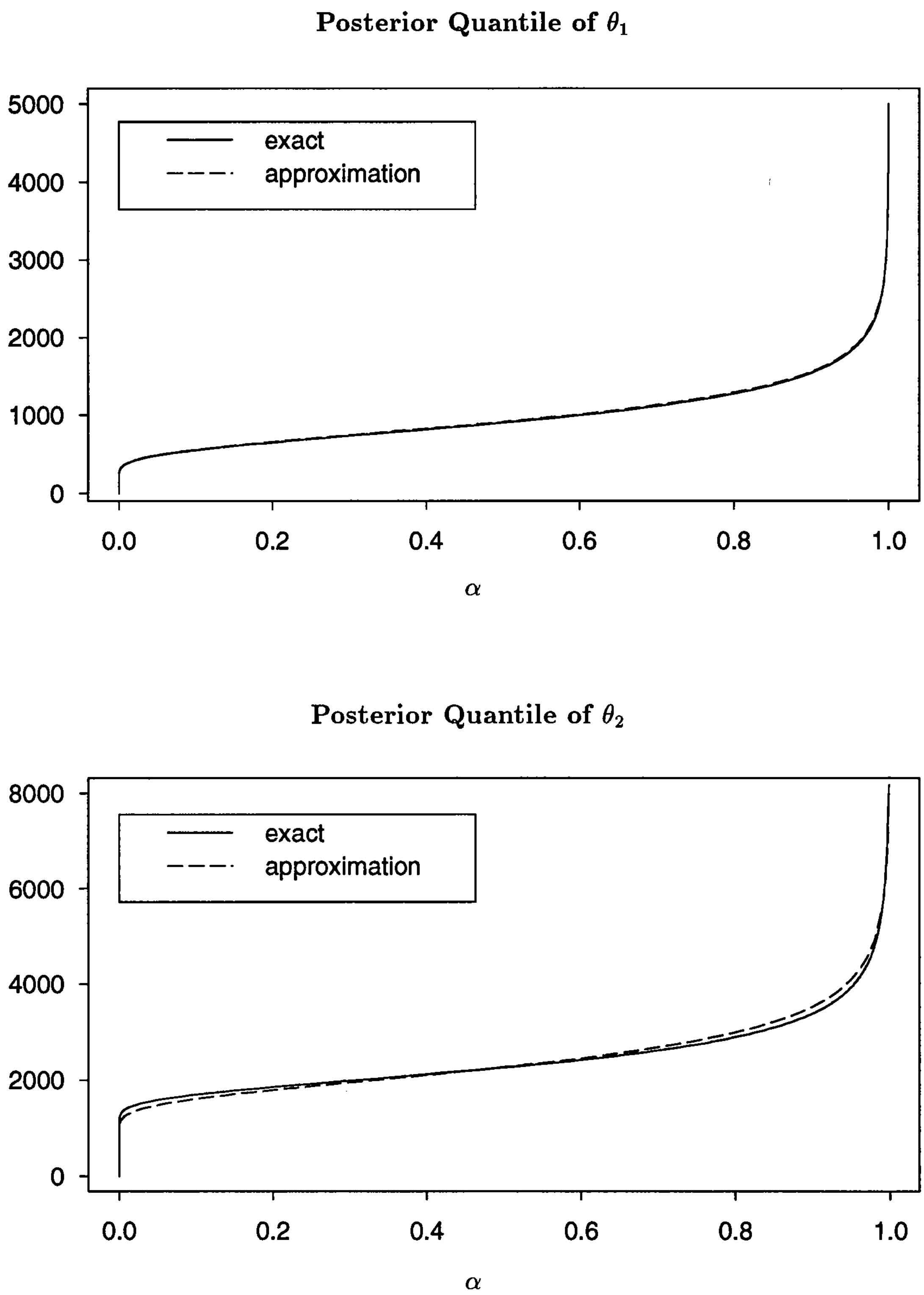
$\beta_1 = 0.5,\ \beta_2 = 2.0,\ a_1 = 15.0,\ a_2 = 1.90,\ b_1 = 430,\ b_2 = 10,575,000$

Figure 1: Marginal Prior and Posterior Densities of θ_1 and θ_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 2: Predictive Reliability $\hat{R}(t)$



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

Figure 3: Posterior Quantiles of θ_1 and θ_2

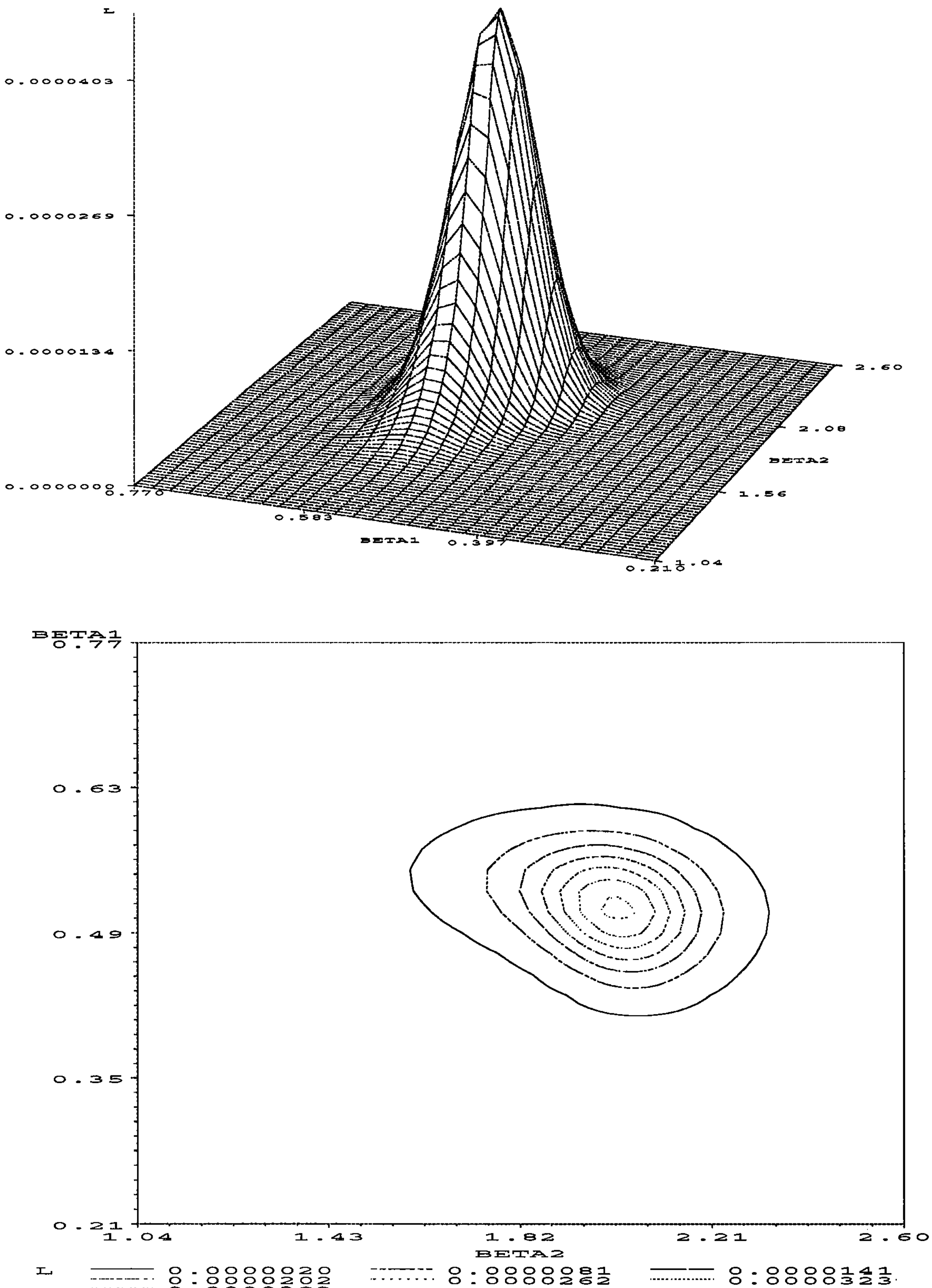


Figure 4: The "Type II" Likelihood Function of (β_1, β_2)