INFERENCE AND PREDICTION
FOR A GENERAL ORDER STATISTIC MODEL
WITH UNKNOWN POPULATION SIZE

By

ABRAMO G. RAYKOV

TECHNICAL REPORT 84-29
August 1984

Department of Statistics, #11-32
University of Washington
Seattle, Washington 98195, USA

DISTRIBUTION STATEMENT A
Approved for public release
Distributed Unlimited
Inference and prediction for a general order statistic model with unknown population size

A Bayes empirical Bayes approach to inference is presented. This permits the comparison of competing, perhaps non-nested, models in a natural way, and also provides easily implemented inference and prediction procedures which avoid the difficulties of non-Bayesian methods. Applications to three software reliability data sets indicate that the much-used...
exponential order statistic model may give rather optimistic estimates of system reliability, while the, not previously considered, Weibull order statistic model seems promising for such applications.
Inference and Prediction for a General Order Statistic Model With Unknown Population Size

Adrian E. Raftery
Department of Statistics, GN-22,
University of Washington,
Seattle, WA 98195.

ABSTRACT

Suppose that the first $n$ order statistics from a random sample of $N$ positive random variables are observed, where $N$ is unknown. A Bayes empirical Bayes approach to inference is presented. This permits the comparison of competing, perhaps non-nested, models in a natural way, and also provides easily implemented inference and prediction procedures which avoid the difficulties of non-Bayesian methods. Applications to three software reliability data sets indicate that the much-used exponential order statistic model may give rather optimistic estimates of system reliability, while the, not previously considered, Weibull order statistic model seems promising for such applications.

KEY WORDS: Bayes empirical Bayes; Bayes factor; Non-nested models; Pareto order statistic model; Software reliability; Weibull order statistic model.
1. INTRODUCTION

Suppose $X_1, \ldots, X_N$ is a random sample of positive random variables from a distribution with probability density function (pdf) at $x$ equal to $\beta f_\theta(\beta x)$. Here $\beta$ is a scalar precision parameter, $\theta$ is a, possibly vector, shape parameter, and $N$ is unknown. In applications, $X_i$ is often a length of time, such as a lifelength, and $X_i = x$ corresponds to the occurrence of an event at time $x$. I shall use this temporal imagery without further explanation.

The first $n$ order statistics, $t = (t_1, \ldots, t_n)$, are observed, where $0 \leq t_1 \leq \cdots \leq t_n \leq T$. $T$ is the period of observation: there is no $X_i$ such that $t_n < X_i \leq T$. Inference is to be made about the unknown parameters, and future observations are to be predicted.

I shall call this the general order statistic (GOS) model. Special cases have been proposed as models for market penetration and capture-recapture studies (Anscombe 1961), burn-in in repairable systems (Bazovsky 1961, chap. 8; Cozzolino 1968), software reliability growth (Jelinski and Moranda 1972; Littlewood 1981), estimating the number of individuals exposed to radiation (Hoel 1968), and estimating the number of unseen species (Efron and Thisted 1976, and references therein).

Perhaps the simplest special case is the exponential order statistic (EOS) model where $f_\theta(x) = \exp(-x)$, statistical analysis of which has been extensively studied (Blumenthal and Marcus 1975; Forman and Singpurwalla 1977; Goudie and Goldie 1981; Jewell 1985; Joe and Reid 1985; Raftery 1986a). It has been used extensively as a simple, physical, debugging model for software reliability. In this context it is often called the Jelinski-Moranda model, and is based on the assumption that a system has $N$ faults, each of which causes a failure of the system, and is
then located and removed; the times at which the $N$ failures occur are independent and identically distributed exponential random variables. However, the examples in Section 6 show that it may give rather optimistic estimates of system reliability.

The EOS model can be generalised by assuming that $X_1, \ldots, X_n$ are independent exponential random variables with different means $\xi_1^{-1}, \ldots, \xi_N^{-1}$, where $\xi_1, \ldots, \xi_N$ is itself a random sample from a distribution with pdf at $\xi$ equal to $\beta^{-1}w(\xi\beta^{-1})$. This is a special case of the GOS model, where

$$f_\theta(x) = \int y \, w(y) \exp(-xy) \, dy$$

(1.1)

Miller (1986) has pointed out that many proposed software reliability models are, in fact, of this form. When the $\xi_i$ have a gamma distribution, the $X_i$ have a Pareto distribution. This, the Pareto order statistic (POS) model, is discussed in more detail in Section 5.2.

I adopt a Bayes empirical Bayes approach (Deely and Lindley 1981) to the problem of inference for the GOS model. This has the advantage of permitting comparisons between competing, perhaps non-nested, models for $f_\theta(x)$ in a natural way (Section 2), as well as providing easily implemented inference and prediction procedures which avoid the difficulties of non-Bayesian methods (Section 3). One such difficulty is that the maximum likelihood estimator of $N$ may be infinite. Indeed, Goudie and Goldie (1981) concluded that for the special case they considered, all standard non-Bayesian point estimation techniques are liable to fail. Attention is paid to the situation where vague prior information about the model parameters is approximated by limiting, improper, prior forms.
Some analytic simplification is possible for the *Weibull order statistic* (WOS) model, where the $X_i$ have a Weibull distribution (Section 5.1). The examples in Section 6 suggest that this model may be promising for software reliability applications, for which it has not previously been considered.

### 2. MODEL COMPARISON

Consider the GOS model described in Section 1. In this section and the next one I assume that $\theta$ is known and omit it from the notation; this assumption is relaxed in Section 4. I assume that $N$ has a Poisson distribution in the GOS model; this defines an empirical Bayes model in the sense of Morris (1983).

It is equivalent to a non-homogeneous Poisson process with $\lambda(s)$, the intensity function at time $s$, given by $\lambda(s) = \rho f(\beta s) \ (\rho > 0)$. The likelihood is

$$ p(t \mid \rho, \beta) = \rho^n \left\{ \prod_{i=1}^{n} f(\beta r_i) \right\} \exp\{-\rho \beta^{-1} F(\beta T)\} \quad (2.1) $$

where $F(x) = \int_{0}^{x} f(y) \, dy$.

Consider the problem of comparing competing, perhaps non-nested, models for $f(x)$, $M_1$ and $M_2$, say. Such comparisons will be based on the *Bayes factor*, or ratio of posterior to prior odds for $M_1$ against $M_2$,

$$ B_{12} = \frac{p(t \mid M_1)}{p(t \mid M_2)} \quad (2.2) $$

the ratio of the marginal likelihoods. In (2.2),
\[ p(t | M_i) = \int_0^\infty \int_0^\infty \int_0^\infty \int_0^\infty \int_0^\infty p(t | \rho, \beta, M_i) p(\rho, \beta | M_i) d\rho d\beta \quad (i=1,2) \]  

(2.3)

If the priors \( p(\rho, \beta | M_i) \) \( (i=1,2) \) are proper, (2.2) can be evaluated directly.

I now develop an expression for \( B_{12} \) in the situation where vague prior knowledge is approximated by limiting, improper, prior forms. This is done by comparing \( M_1 \) and \( M_2 \) in turn with the constant rate Poisson process, \( M_0: \lambda(s) = \mu \), which is nested within each of \( M_1 \) and \( M_2 \).

This yields Bayes factors \( B_{01} \) and \( B_{02} \), where

\[ B_{0i} = \frac{p(t | M_0)}{p(t | M_i)} \quad (i=1,2) \]  

(2.4)

and

\[ p(t | M_0) = \int_0^\infty \int_0^\infty p(t | \mu, M_0) p(\mu | M_0) d\mu \]  

(2.5)

Then \( B_{12} = B_{01} / B_{01} \). Comparison of \( M_0 \) with \( M_i \) using (2.4) may itself be of interest. For example, in the software reliability context, it provides a test of whether the system is, indeed, being debugged.

I use the standard vague prior for \( \mu \),

\[ p(\mu | M_0) = c_0 \mu^{-1} \]  

(2.6)

(Jaynes 1968), and consider the evaluation of \( B_{01} \). In order to provide a satisfactory approximation for vague prior knowledge over all scales, the prior distribution of \( (\rho, \beta) \) should yield a Bayes factor \( B_{01} \) which is time-invariant, i.e. invariant to scale changes in the time variable.
Theorem 1: \( B_{01} \) is time-invariant if and only if there is a function \( \phi(.) \) such that

\[
p(\rho, \beta | M_1) = c_1 \rho^{-2} \phi(\rho^{-1} \beta)
\]  
\[(2.7)\]

Proof: Suppose \( B_{01} \) is time-invariant. By (2.6)

\[
p(t | M_0) = c_0 (n-1)! T^{-n}
\]  
\[(2.8)\]

Using (2.1), and substituting \( \rho T \) for \( \rho \) and \( \beta T \) for \( \beta \) in (2.3), and then dividing the result into (2.8), yields, by (2.4),

\[
B_{01} = c_{01} \left[ \int_0^\infty \int_0^\infty \right] \rho^n \left\{ \prod_{i=1}^n f(\beta u_i) \right\} \exp\left\{-\rho \beta^{-1} F(\beta)\right\} T^{-2} p(\rho T^{-1}, \beta T^{-1} | M_1) d\rho d\beta
\]  
\[(2.9)\]

Thus

\[
T^{-2} p(\rho T^{-1}, \beta T^{-1} | M_1) = p(\rho, \beta | M_1) \quad (\rho, \beta, T > 0)
\]  
\[(2.10)\]

Setting \( T=\rho \) in (2.10) yields (2.7), where \( \phi(x) = p(\rho=1, \beta=x | M_1) \). Also, when (2.7) holds, the time-invariance of \( B_{01} \) follows by direct substitution in (2.9). This completes the proof.

If the prior is to be asymptotically non-increasing in \( \rho \) and \( \beta \), then, by Theorem 1, \( \phi(x) \) must be bounded above by \( \gamma_1 \) and below by \( \gamma_2 x^{-2} \) for \( x \) sufficiently large, where \( \gamma_1 \) and \( \gamma_2 \) are positive constants. Consider now the case where the likelihood (2.1) is of exponential family form, so that

\[
f(\beta x) = \exp\left\{ a(\beta) + a(x) + \sum_{j=1}^J (\beta x)^j + \text{const.} \right\}
\]  
\[(2.11)\]

This is quite a general family, and includes, for example, the gamma and Weibull distributions.

By (2.1), a natural family of conjugate prior distributions is...
\[ p(\rho, \beta | M_1) = c_1 \exp \{ k_0 \alpha (\beta) + \sum_{j=1}^{J} k_j \beta_j \} \rho \exp \{ -k_{f+2} F(\beta T) \} \] (2.12)

By Theorem 1, the unique prior of the form (2.12) which is independent of \( T \) and yields time-invariant Bayes factors for all models of the form (2.11) is

\[ p(\rho, \beta | M_1) = c_1 \rho^{-2} \] (2.13)

This prior is also independent of the shape parameter.

It follows from (2.9) that, with the priors (2.6) and (2.13), the Bayes factor has the form

\[ B_{01} = c_{01} (n-1) h(u)^{-1} \] (2.14)

where \( c_{01} = c_0 / c_1 \), \( u = (u_1, \ldots, u_n) \), \( u_i = t_i / T \) for \( i = 1, \ldots, n \), and

\[ h(u) = \int_{0}^{\infty} y^{n-1} \left( \prod_{i=1}^{n} f(y u_i) \right) F(y)^{-(n-1)} dy \] (2.15)

However, (2.14) involves the arbitrary, undefined, multiplicative constant \( c_{01} \), which appears because the priors used are improper. Akman and Raftery (1986a) have shown how this may be assigned using the minimal imaginary training sample idea of Spiegelhalter and Smith (1982). This consists of imagining that a data set is available which involves the smallest possible sample size permitting a comparison of \( M_0 \) and \( M_1 \), and provides maximum possible support for \( M_0 \). It is then argued that the resulting Bayes factor, \( B_{01} \), should be only slightly greater than one. Raftery and Akman (1986) have applied this approach to the change-point Poisson process; their results may be compared with the non-Bayesian solution of Akman and Raftery (1986b). This approach has also been applied to log-linear models for contingency tables by Raftery (1986b).
In the present situation, the appropriate imaginary data set consists of two observations at
the same value, $u = (u_1, u_2) = (v, v)$, where $v$ is chosen so as to maximise the value of $B_{01}$ in
(2.14). In practice, in all the examples considered, $B_{01}$ is maximised at either $v=1$ or $v=0$. When
$B_{01}$ is maximised at $v=0$, however, the maximum value is infinite. In such cases, I use the local
maximum at $v=1$, because this corresponds, in the software reliability situation, for example, to
the data set which suggests most strongly that the system is not being debugged. This yields

$$c_{01} = \int_0^\infty y f(y)^2 F(y)^{-1}dy$$  \hspace{1cm} (2.16)

Strictly speaking, any value of $B_{12}$ less than one suggests that the data provide evidence
against $M_1$ for $M_2$. However, as a rough order of magnitude interpretation, Jeffreys (1961,
Appendix B) has suggested that the evidence should be regarded as strong only if $B_{12} < 10^{-1}$,
and as decisive only if $B_{12} < 10^{-2}$.

3. ESTIMATION AND PREDICTION

I now consider estimation of $N$, and prediction of future observations for the GOS model.
The framework developed in Section 2 is used. It follows from (2.13) that

$$p(N, \beta) = \int_0^\infty p(N | \rho, \beta) p(\rho, \beta) d\rho$$

$$\propto \{N(N-1)\}^{-1} \beta^{-1}$$  \hspace{1cm} (3.1)

Also,
\[ p(t | N, \beta) = \{N!(N-n)\} \beta^n \left( \prod_{i=1}^{n} f(\beta t_i) \right) \bar{F}(\beta T)^{N-n} \]  

(3.2)

where \( \bar{F}(x) = 1 - F(x) \). Combining (3.1) with (3.2) and integrating over \( \beta \) yields the posterior distribution of the number of unobserved variables \( M = N - n \),

\[ p(M | t) \propto \{(M+n-2)!/M!\} g(u,M) \quad (M=0,1, \ldots) \]  

(3.3)

where

\[ g(u,M) = \int_0^\infty y^{n-1} \left( \prod_{i=1}^{n} f(yu_i) \right) \bar{F}(y)^M dy \]  

(3.4)

Point estimators of \( N \) may be obtained by combining (3.3) with an appropriate loss function; examples are the posterior mode and the posterior median. However, experience with the simple EOS model indicates that point estimators of \( N \) are liable to perform badly (Raftery 1986a). Interval estimators of \( N \), such as highest posterior density regions, can readily be found from (3.3), and may well be more useful.

Various prediction problems may be of interest, and can be solved, often quite easily, using the present approach. One example is finding the probability, given the data, that there is no \( X_i \) such that \( T < X_i < T + z \). In the software reliability context, this is the current reliability of the system for a task of length \( z \). If \( Z = t_{n+1} - T \), where \( t_{n+1} = \infty \), then

\[ P[Z > z | t] = \sum_{M=0}^{\infty} \int P[Z > z | t] p(M, \beta | t) d\beta \]

\[ = P[M=0 | t] + \left( \frac{T}{(T+x)} \right)^n \sum_{M=1}^{\infty} \{ g(uT/(T+x),M)/g(u,M) \} p(M | t) \]  

(3.5)
4. SHAPE PARAMETER UNKNOWN

Suppose now that the shape parameter \( \theta \) in the GOS model is unknown. I continue to use the framework of Sections 2 and 3, but quantities which depend on \( \theta \) are now written with a subscript \( \theta \). I know of no single prior which can provide a satisfactory approximation for vague prior knowledge about \( \theta \) in all situations. I therefore assume that

\[
p(\rho, \beta, \theta) = c_1 \rho^{-2} p(\theta) \tag{4.1}\]

where \( p(\theta) \) is proper. I denote the set of possible values of \( \theta \) by \( \Theta \).

The results of Sections 2 and 3 can be generalised to this situation by conditioning on \( \theta \) and using the total probability law in an appropriate way. Thus (2.14) becomes

\[
B_{01} = c_{01} (n-1) H(u)^{-1} \tag{4.2}
\]

where

\[
H(u) = \int_{\Theta} h_\theta(u) p(\theta) \, d\theta \tag{4.3}
\]

and \( h_\theta(u) \) is defined by (2.15). (2.16) becomes

\[
c_{01} = \int_{\Theta} \int \frac{y f_\theta(y)^2 F_\theta(y)^{-1}}{p(\theta)} \, dy \, d\theta \tag{4.4}
\]

For estimation of \( N \), (3.3) becomes

\[
p(M | t) \propto \{(M+n-2)!/M!\} G(u, M) \quad (M=0,1, \cdots) \tag{4.5}
\]

where

\[
G(u, M) = \int_{\Theta} g_\theta(u, M) p(\theta) \, d\theta \tag{4.6}
\]
and $g_\theta(u,M)$ is defined by (3.4).

For the prediction problem considered in Section 3, (3.5) becomes

$$P[Z > z \mid t] = P[M = 0 \mid t] + \{T/(T + x)\}^n \sum_{M=1}^{\infty} \{G(uT/(T + x), M) \land G(u, M)\} p(M \mid t) \quad (4.7)$$

5. SPECIAL CASES

5.1 The Weibull Order Statistic (WOS) Model

Among commonly used models for positive random variables, the Weibull distribution yields some analytic simplification of the results in Section 4. The WOS model is defined by setting

$$f_\theta(x) = \theta x^{\theta - 1} \exp(-x^\theta) \quad (\theta > 0) \quad (5.1)$$

in the GOS model. Then $B_{01}$ is given by (4.2), (4.3), and (4.4), where

$$h_\theta(u) = \theta^{n-1} (\prod_{i=1}^n u_i)^{\theta-1} \int_0^{\infty} \exp(-y \sum_{i=1}^n u_i^\theta) \left\{y/(1-e^{-y})\right\}^{n-1} dy$$

and $c_{01} = \left(\frac{n}{\theta}\right) - 1 E[\theta]$.

The solutions to the estimation and prediction problems are given by (4.5), (4.6), and (4.7), where

$$g_\theta(u,M) = \theta^{n-1} (\prod_{i=1}^n u_i)^{\theta-1} \left(\sum_{i=1}^n u_i^\theta + M\right)^{-n}$$
5.2 The Pareto Order Statistic (POS) Model

Consider the POS model described in Section 1, where in (1.1),

\[ w_\theta(y) = \Gamma(\theta)^{-1} y^{\theta-1} e^{-\theta y} \]  

(5.2)

so that, by (1.1),

\[ f_\theta(y) = \theta (1+y)^{-(\theta+1)} \]  

(5.3)

\( B_{01} \) is again given by (4.2) and (4.3), where

\[ h_\theta(u) = \theta^n \prod_{i=1}^n (1+\beta u_i)^{-(\theta+1)} \int_0^\infty y^{n-1} \{1-(1+y)^{-\theta}\}^{-(n-1)} dy \]

and (4.4) becomes

\[ c_{01} = \int \int_0^\infty y (1+y)^{-2(\theta+1)} \{1-(1+y)^{-\theta}\}^{-1} dy \theta^2 p(\theta) d\theta \]

The solutions to the estimation and prediction problems are somewhat simplified if a gamma prior for \( \theta \) is used, namely, in (4.1),

\[ p(\theta) \propto \theta^{\kappa_1-1} e^{-\kappa_2 \theta} \]  

(5.4)

The solutions are given by (4.5) and (4.7), where

\[ G(u, M) = \prod_{i=1}^n (1+\beta u_i)^{-1} \int_0^\infty y^{n-1} \{\kappa_2 + \sum_{i=1}^n \log(1+y u_i) + M \log(1+y)\}^{-(\kappa+\kappa_1)} dy \]

Most of the integrals in this section, which require numerical evaluation, could be replaced by convergent infinite series. However, this was not found to be computationally advantageous.
6. EXAMPLES

I now apply the techniques proposed here to three, previously analyzed, software reliability data sets.

Example 1: Goel and Okumoto (1979) gave the 31 failure times of a piece of software developed as part of the Naval Tactical Data System. The Bayes factors for comparing the models considered in this paper are shown in Table 1. As explained in Section 2, these were obtained as quotients of the Bayes factors for the constant rate Poisson process against each of the models individually, given by (4.2). The necessary single and double numerical integrations were carried out using the IMSL routines DCADRE and DBLIN, respectively.

| Table 1 about here |

For the WOS model (5.1), only distributions with tails at least as heavy as exponential were considered, and \( p(\theta) \) was taken to be uniform between \( \frac{1}{4} \) and 1. \( \theta = \frac{1}{4} \) corresponds to a quite heavy-tailed distribution, while \( \theta = 1 \) is the exponential distribution. With this prior, the WOS model can be thought of as representing a situation where the bugs become harder to detect as the debugging process proceeds.

For the POS model (5.3), the prior distribution of \( \theta \) was given by (5.4) with \( \kappa_1 = 2 \) and \( \kappa_2 = \frac{1}{2} \), so that about 95% of the prior distribution of \( \theta \) was concentrated between \( \frac{1}{2} \) and 10. \( \theta = \frac{1}{2} \) in (5.2) corresponds to a heavy-tailed distribution for \( \xi_i \), while \( \theta = 10 \) corresponds to a distribution for \( \xi_i \) which is close to normality.
Table 1 shows that no model performs markedly better than any other. Indeed, the EOS model, originally proposed for this data by Jelinski and Moranda (1972), seems quite acceptable.

**Example 2:** Meinhold and Singpurwalla (1983) gave the 136 failure times of a real-time command and control system, and analyzed them using the EOS model. The same priors are used as in Example 1. The Bayes factors in Table 1 suggest that the WOS model is better than both the EOS and POS models. The posterior distribution of $M$ for the EOS and WOS models is shown in Figure 1, and salient features are summarised in Table 2. It appears that the EOS model substantially underestimates the number of faults still present.

![](Figure 1 about here)

![](Table 2 about here)

**Example 3:** Forman and Singpurwalla (1977) analyzed a data set consisting of 107 failures using the EOS model. The priors used are the same as in the first two examples. The data were grouped, and I distributed the failures randomly according to a uniform distribution over the time intervals in which they occurred. The conclusions of all the model comparisons were the same for each of four different sequences of random numbers used to distribute the failure times; the results reported here are for one of these.
The WOS model was again the preferred one. There were other signs of the inadequacy of the EOS model. For example, after 99 of the 107 recorded failures, the probability of eight or more failures occurring was less than $10^{-4}$ under the EOS model, but 0.18 under the WOS model.

The posterior distributions of the number of remaining faults under the EOS and WOS models are shown in Figure 2. The EOS model gave rather optimistic estimates of the state of the system. For example, under the EOS model, the probability of the system having been fully debugged was 0.95, while under the WOS model it was only 0.27.

In addition to its capacity for representing slowly decreasing failure rates, the WOS model can also represent failure rates which increase and then decrease, when $\theta>1$ in (5.1). This possibility has not been exploited here, but Littlewood and Verrall (1981) and Ascher and Feingold (1984, pp.110-111) have described software reliability data sets of which this is a feature.

REFERENCES


Table 1. $\log_{10}(\text{Bayesfactor})$
for the model comparisons in Examples 1,2,3.

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Example 1</th>
<th>Example 2</th>
<th>Example 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>EOS vs. WOS</td>
<td>0.4</td>
<td>-3.7</td>
<td>-8.3</td>
</tr>
<tr>
<td>EOS vs. POS</td>
<td>-0.1</td>
<td>-1.4</td>
<td>-5.8</td>
</tr>
<tr>
<td>WOS vs. POS</td>
<td>-0.5</td>
<td>2.3</td>
<td>2.5</td>
</tr>
</tbody>
</table>
Table 2. Features of the posterior distribution of $M$, the number of remaining bugs, under the EOS and WOS models, in Examples 2 and 3.

| Example | Model | Mode | Median | $P[M=0|t]$ | 95% HPDR |
|---------|-------|------|--------|------------|----------|
| 2       | EOS   | 6    | 6.5    | .01        | 1-16     |
|         | WOS   | 27   | 40.7   | .00        | 6-122    |
| 3       | EOS   | 0    | .0     | .95        | 0        |
|         | WOS   | 1    | .9     | .27        | 0-6      |

NOTE: 95% HPDR is the 95% highest posterior density region. $i-j$ denotes the set of integers from $i$ to $j$ inclusive.
Table 2. Features of the posterior distribution of $M$, the number of remaining bugs, under the EOS and WOS models, in Examples 2 and 3.

| Example | Model | Mode | Median | $P[M=0|t]$ | 95% HPDR |
|---------|-------|------|--------|------------|----------|
| 2       | EOS   | 6    | 6.5    | .01        | 1-16     |
|         | WOS   | 27   | 40.7   | .00        | 6-122    |
| 3       | EOS   | 0    | .0     | .95        | 0        |
|         | WOS   | 1    | .9     | .27        | 0-6      |

NOTE: 95% HPDR is the 95% highest posterior density region. $i-j$ denotes the set of integers from $i$ to $j$ inclusive.
Captions for Figures 1 and 2:

Figure 1. Posterior distributions of $M$, the number of remaining bugs, in Example 2 under (a) the EOS model, and (b) the WOS model. The WOS model, which is favored by the data, estimates a much larger number of remaining bugs than the EOS model.

Figure 2. Posterior distributions of $M$ in Example 3 under (a) the EOS model, and (b) the WOS model. Under the EOS model, almost the entire posterior distribution of $M$ is concentrated at 0, while from the WOS model, which is favored by the data, it appears that there may be up to six remaining bugs with non-negligible probability.
Figure 2

(a) $p(M|t)$

(b) $p(M|t)$
END
7-89
Dtic