Inference and Prediction for a General Order Statistic Model With Unknown Population Size

ADRIAN E. RAFTERY*

Suppose that the first n order statistics from a random sample of N positive random variables are observed, where N is unknown. This, the general order statistic model, has been applied to the study of market penetration, capture–recapture, burn-in in repairable systems, software reliability growth, the estimation of the number of individuals exposed to radiation, and the estimation of the number of unseen species. Inference is to be made about the unknown parameters, especially N, and future observations are to be predicted. A Bayes empirical Bayes approach to inference is presented. This permits the comparison of competing, perhaps nonnested, models for the distribution of the random variables in a natural way. It also provides easily implemented inference and prediction procedures that avoid the difficulties of non-Bayesian methods. One such difficulty is that the maximum likelihood estimator of N may be infinite. Results are given for the case in which vague prior information about the model parameters is approximated by limiting, improper, prior forms. Applications to three software reliability data sets indicate that the much-used exponential order statistic may give rather optimistic estimates of system reliability and that the not previously considered Weibull order statistic model seems promising for such applications.

KEY WORDS: Bayes empirical Bayes; Bayes factor; Nonnested 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 its probability density function (pdf) at x equal to $\beta f_\theta(\beta x)$. Here β is a scalar precision parameter, θ is a, possibly vector, shape parameter, and N is unknown. In applications, X_i is often a length of time, such as a life length, 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 $0 \le t_1 \le \cdots \le t_n \le T$. T is the period of observation; there is no X_i such that $t_n < X_i \le T$. 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 in which $f_\theta(x) = \exp(-x)$, 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, in press). 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. The examples in Section 6, however, show that it may give rather optimistic estimates of system reliability.

The EOS model can be generalized 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 ξ equal to $\beta^{-1} w_\theta(\xi\beta^{-1})$. This is a special case of the GOS model, where

$$f_\theta(x) = \int y w_\theta(y) \exp(-xy)\, dy. \tag{1.1}$$

Miller (1986) pointed out that many proposed software reliability models are, in fact, of this form. When the ξ_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 nonnested, models for $f_\theta(x)$ in a natural way (Sec. 2) and providing easily implemented inference and prediction procedures that avoid the difficulties of non-Bayesian methods (Sec. 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 in which 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, in which the X_i have a Weibull distribution (Sec. 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.

* Adrian E. Raftery is Associate Professor, Department of Statistics, GN-22, University of Washington, Seattle, WA 98195. This work was supported by Office of Naval Research Contract N00014-84-C-0169. The author is grateful to W. S. Jewell for discussions, to Peter Guttorp, Harold Asher, the editor, and the associate editor for comments, and to an anonymous referee for a very careful reading of an earlier version, all of which considerably improved the manuscript.

Journal of the American Statistical Association
December 1987, Vol. 82, No. 400, Theory and Methods

2. MODEL COMPARISON

Consider the GOS model described in Section 1. In this section and the next one I assume that θ 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 nonhomogeneous Poisson process with $\lambda(s)$, the intensity function at time s, given by $\lambda(s) = \rho f(\beta s)$ $(\rho > 0)$. The likelihood is

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

where $F(x) = \int_0^x f(y)\, dy$.

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

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

the ratio of the marginal likelihoods. In (2.2),

$$p(t \mid M_i) = \int_0^\infty \int_0^\infty p(t \mid \rho, \beta, M_i) p(\rho, \beta \mid M_i)\, d\rho\, d\beta, \quad i = 1, 2. \quad (2.3)$$

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

I now develop an expression for B_{12} in the situation in which 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

$$B_{0i} = p(t \mid M_0)/p(t \mid M_i), \quad i = 1, 2, \quad (2.4)$$

and

$$p(t \mid M_0) = \int_0^\infty p(t \mid \mu, M_0) p(\mu \mid M_0)\, d\mu. \quad (2.5)$$

Then $B_{12} = B_{02}/B_{01}$. 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 μ,

$$p(\mu \mid M_0) = c_0 \mu^{-1} \quad (2.6)$$

(Jaynes 1968), and consider the evaluation of B_{01}. To provide a satisfactory approximation for vague knowledge over all scales, the prior distribution of (ρ, β) should yield a Bayes factor B_{01} that is *time-invariant*—that is, invariant to scale changes in the time variable.

Theorem 1. B_{01} is time-invariant iff there is a function $\phi(\cdot)$ such that

$$p(\rho, \beta \mid M_1) = c_1 \rho^{-2} \phi(\rho^{-1}\beta). \quad (2.7)$$

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

$$p(t \mid M_0) = c_0(n-1)!T^{-1}. \quad (2.8)$$

Using (2.1) and substituting ρT for ρ and βT for β in (2.3), and then dividing the result into (2.8), yields, by (2.4),

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

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

$$T^{-2} p(\rho T^{-1}, \beta T^{-1} \mid M_1) = p(\rho, \beta \mid M_1), \quad \rho, \beta, T > 0. \quad (2.10)$$

Setting $T = \rho$ in (2.10) yields (2.7), where $\phi(x) = p(\rho = 1, \beta = x \mid M_1)$. Moreover, 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 nonincreasing in ρ and β, then, by Theorem 1, $\phi(x)$ must be bounded above by γ_1 and below by $\gamma_2 x^{-2}$ for x sufficiently large, where γ_1 and γ_2 are positive constants. Consider now the case in which the likelihood (2.1) is of exponential family form when T is fixed. Then f is of the form

$$f(y) \propto y^a \exp\left(\sum_{j=1}^{J} b_j y^{d_j} \right). \quad (2.11)$$

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

$$p(\rho, \beta \mid M_1) \propto \beta^{ak_0} \exp\left(\sum_{j=1}^{J} k_j \beta^{d_j} \right) \times \rho^{k_{J+1}} \exp\{-k_{J+2}\rho\beta^{-1}F(\beta T)\}. \quad (2.12)$$

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

$$p(\rho, \beta \mid M_1) = c_1 \rho^{-2}. \quad (2.13)$$

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

$$B_{01} = c_{01}(n-1)h(u)^{-1}, \quad (2.14)$$

where

$$h(u) = \int_0^\infty y^{n-1} \left\{ \prod_{i=1}^{n} f(yu_i) \right\} F(y)^{-(n-1)}\, dy. \quad (2.15)$$

Equation (2.14), however, involves the arbitrary, undefined, multiplicative constant c_{01}, which appears because the priors used are improper. Akman and Raftery (1986a) showed how this may be assigned using the minimal imaginary training sample idea of Spiegelhalter and Smith (1982). This consists of imagining that there is available a data set that involves the smallest possible sample size permitting a comparison of M_0 and M_1, and that provides maximum possible support for M_0. It is then argued that the resulting Bayes factor, B_{01}, should be only slightly greater than 1. Raftery and Akman (1986) 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 was also applied to log-linear models for contingency tables by Raftery (1986).

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 maximize the value of B_{01} in (2.14). In practice, in all of the examples considered, B_{01} is maximized at either $v = 1$ or $v = 0$. When B_{01} is maximized 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 that suggests most strongly that the system is not being debugged. This yields

$$c_{01} \approx \int_0^\infty y f(y)^2 F(y)^{-1}\, dy. \tag{2.16}$$

Strictly speaking, any value of B_{12} less than 1 suggests that the data provide evidence against M_1 for M_2. As a rough order of magnitude interpretation, however, Jeffreys (1961, app. B) 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

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

$$\propto \{N(N-1)\}^{-1}\beta^{-1}. \tag{3.1}$$

In addition,

$$p(t \mid N, \beta) = \{N!/(N-n)!\}\, \beta^n \left\{\prod_{i=1}^n f(\beta t_i)\right\} \overline{F}(\beta T)^{N-n}, \tag{3.2}$$

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

$$p(M \mid t) \propto \{(M + n - 2)!/M!\}\, g(u, M), \qquad M = 0, 1, \ldots, \tag{3.3}$$

where

$$g(u, M) = \int_0^\infty y^{n-1} \left\{\prod_{i=1}^n f(yu_i)\right\} \overline{F}(y)^M\, dy. \tag{3.4}$$

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. Experience with the simple EOS model indicates, however, that point estimators of N are liable to perform badly (Raftery, in press). Interval estimators of N, such as highest posterior density regions, can readily be found from (3.3), and they 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 are 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

$$\begin{aligned}\Pr[Z > z \mid t] &= \sum_{M=0}^\infty \int_0^\infty \Pr[Z > z \mid t, M, \beta]\, p(M, \beta \mid t)\, d\beta \\ &= \Pr[M = 0 \mid t] + \{T/(T+z)\}^n \\ &\quad \times \sum_{M=1}^\infty \frac{g(uT/(T+z), M)}{g(u, M)}\, p(M \mid t).\end{aligned} \tag{3.5}$$

4. SHAPE PARAMETER UNKNOWN

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

The results of Sections 2 and 3 can be generalized to this situation by conditioning on θ 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). Then (2.16) becomes

$$c_{01} = \int_\Theta \int_0^\infty y f_\theta(y)^2 F_\theta(y)^{-1}\, dy\, p(\theta)\, d\theta. \tag{4.4}$$

For estimation of N, (3.3) becomes

$$p(M \mid t) \propto \{(M + n - 2)!/M!\}\, G(u, M), \qquad M = 0, 1, \ldots, \tag{4.5}$$

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

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

5. SPECIAL CASES

5.1 The Weibull Order Statistic Model

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

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

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

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

and $c_{01} = (\pi^2/6 - 1)E[\theta]$.

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

$$g_\theta(u, M) \propto \theta^{n-1} \left(\prod_{i=1}^n u_i\right)^{\theta-1} \left(\sum_{i=1}^n u_i^\theta + M\right)^{-n}.$$

5.2 The Pareto Order Statistic Model

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

$$w_\theta(y) = \Gamma(\theta)^{-1} y^{\theta-1} e^{-\theta y}; \quad (5.2)$$

so by (1.1),

$$f_\theta(y) = \theta(1 + y)^{-(\theta+1)}. \quad (5.3)$$

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

$$h_\theta(u) = \theta^n \int_0^\infty y^{n-1} \left\{\prod_{i=1}^n (1 + yu_i)^{-(\theta+1)}\right\} \times \{1 - (1 + y)^{-\theta}\}^{-(n-1)}\, dy$$

and (4.4) becomes

$$c_{01} = \int_0^\infty \int_0^\infty y(1 + y)^{-2(\theta+1)} \times \{1 - (1 + y)^{-\theta}\}^{-1}\, dy\, \theta^2 p(\theta)\, d\theta.$$

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

$$p(\theta) \propto \theta^{\kappa_1 - 1} e^{-\kappa_2 \theta}. \quad (5.4)$$

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

$$G(u, M) \propto \int_0^\infty y^{n-1} \left\{\prod_{i=1}^n (1 + yu_i)^{-1}\right\} \times \left\{\kappa_2 + \sum_{i=1}^n \log(1 + yu_i) + M \log(1 + y)\right\}^{-(n+\kappa_1)} dy.$$

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 article 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.

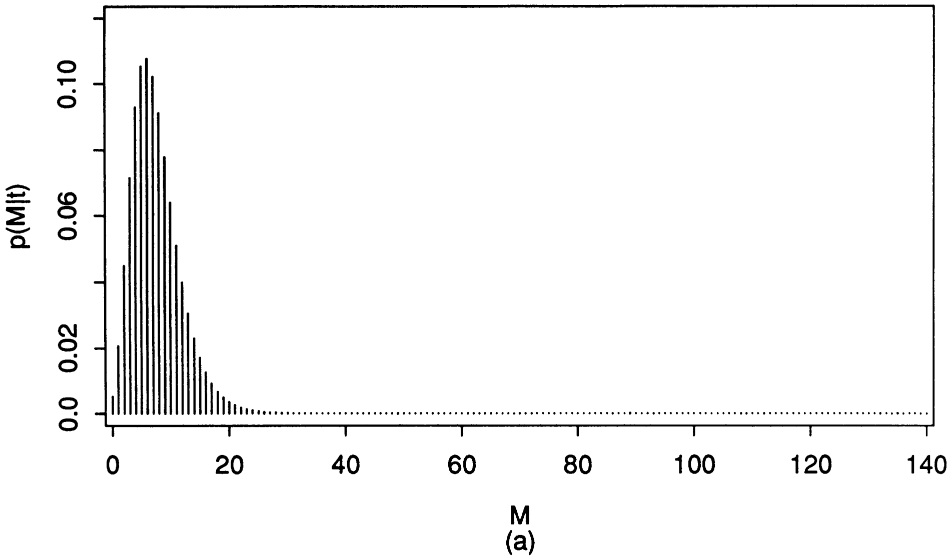
For the WOS model (5.1), only distributions with tails at least as heavy as exponential were considered, and $p(\theta)$ was taken to be uniform between $\frac{1}{2}$ and 1. $\theta = \frac{1}{2}$ corresponds to a quite heavy-tailed distribution, whereas $\theta = 1$ is the exponential distribution. With this prior, the WOS model can be thought of as representing a situation in which the bugs become harder to detect as the debugging process proceeds. Although such a discontinuous prior may seem, at first sight, somewhat artificial, the likelihood for θ is much smaller when $\theta < \frac{1}{2}$ than when $\frac{1}{2} \le \theta \le 1$, so reducing the lower bound would not change the final inference much. The same is true in Examples 2 and 3.

For the POS model (5.3), the prior distribution of θ was given by (5.4) with $\kappa_1 = 2$ and $\kappa_2 = \frac{1}{2}$, so about 95% of the prior distribution of θ was concentrated between $\frac{1}{2}$ and 10. $\theta = \frac{1}{2}$ in (5.2) corresponds to a heavy-tailed distribution for ξ_i, whereas $\theta = 10$ corresponds to a distribution for ξ_i that is close to normality.

Table 1 shows that no model performs markedly better than any other. Indeed, the EOS model, originally pro-

Table 1. Log_{10} (Bayes factor) for the Model Comparisons in Examples 1, 2, and 3

	Example		
Comparison	1	2	3
EOS vs. WOS	.4	−3.7	−8.3
EOS vs. POS	−.1	−1.4	−5.8
WOS vs. POS	−.5	2.3	2.5



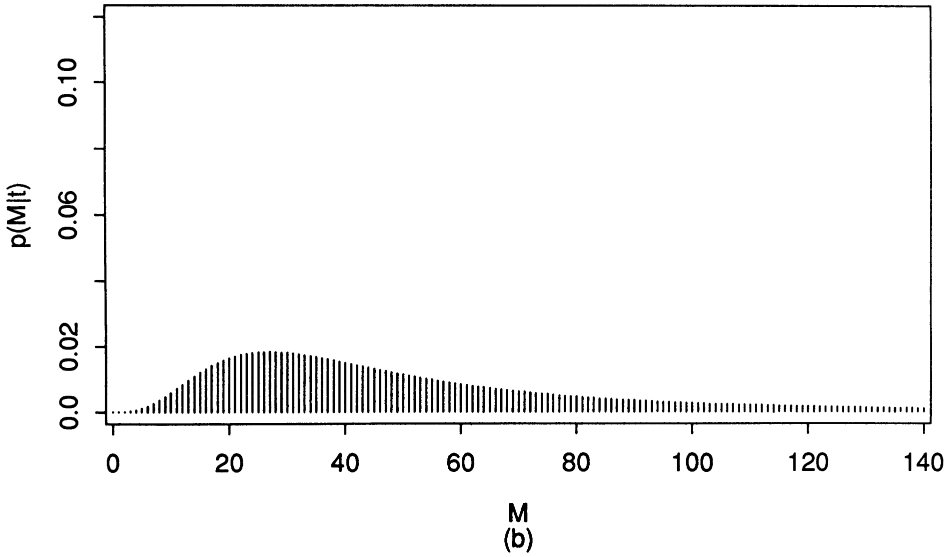


Figure 1. Posterior Distributions of M and the Number of Remaining Bugs in Example 2 Under (a) the EOS Model and (b) the WOS Model. The WOS model, which is favored by the data, estimates a much larger number of remaining bugs than the EOS model.

posed for these 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 they 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 its salient features are summarized in Table 2. It appears that the EOS model substantially underestimates the number of faults still present.

Table 2. Features of the Posterior Distribution of M and the Number of Remaining Bugs, Under the EOS and WOS Models, in Examples 2 and 3

		Feature			
Example	Model	Mode	Median	$Pr[M = 0 \mid t]$	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.

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 of 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 8 or more failures occurring was less than 10^{-4} under the EOS model but .18 under the WOS model.

Table 3. The Posterior Distribution of M and the Number of Remaining Bugs, Under the EOS and WOS Models, in Example 3

M	p(M | t)	
	EOS model	WOS model
0	.95	.27
1	.05	.27
2	—	.19
3	—	.12
4	—	.07
5	—	.04
6	—	.02
7	—	.01
8	—	.01
9	—	—
10	—	—

NOTE: A dash indicates that the posterior probability was less than .005.

The posterior distributions of the number of remaining faults under the EOS and WOS models are shown in Table 3. 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 .95, whereas under the WOS model it was only .27.

In addition to its capacity for representing slowly decreasing failure rates, the WOS model can also represent failure rates that 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) described software reliability data sets of which this is a feature.

[*Received August 1986. Revised April 1987.*]

REFERENCES

Akman, V. E., and Raftery, A. E. (1986a), "Bayes Factors for Non-Homogeneous Poisson Processes With Vague Prior Information," *Journal of the Royal Statistical Society*, Ser. B, 48, 322–329.

——— (1986b) "Asymptotic Inference for a Change-Point Poisson Process," *The Annals of Statistics*, 14, 1583–1590.

Anscombe, F. J. (1961), "Estimating a Mixed Exponential Response Law," *Journal of the American Statistical Association*, 56, 493–502.

Ascher, H., and Feingold, H. (1984), *Repairable Systems Reliability*, New York: Marcel Dekker.

Bazovsky, I. (1961), *Reliability Theory and Practice*, Englewood Cliffs, NJ: Prentice-Hall.

Blumenthal, S., and Marcus, R. (1975), "Estimating Population Size With Exponential Failure," *Journal of the American Statistical Association*, 70, 913–922.

Cozzolino, J. M. (1968), "Probabilistic Models of Decreasing Failure Rate Processes," *Naval Research Logistics Quarterly*, 15, 361–374.

Deely, J. J., and Lindley, D. V. (1981), "Bayes Empirical Bayes," *Journal of the American Statistical Association*, 76, 833–841.

Efron, B., and Thisted, R. (1976), "Estimating the Number of Unseen Species: How Many Words Did Shakespeare Know?" *Biometrika*, 63, 435–447.

Forman, E. H., and Singpurwalla, N. D. (1977), "An Empirical Stopping Rule for Debugging and Testing Computer Software," *Journal of the American Statistical Association*, 72, 750–757.

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

Goudie, I. B. J., and Goldie, C. M. (1981), "Initial Size Estimation for the Linear Pure Death Process," *Biometrika*, 68, 543–550.

Hoel, D. G. (1968), "Sequential Testing of Sample Size," *Technometrics*, 10, 331–341.

Jaynes, E. T. (1968), "Prior Probabilities," *IEEE Transactions on Systems Science and Cybernetics*, 4, 227–241.

Jeffreys, H. (1961), *Theory of Probability* (3rd ed.), Oxford, U.K.: Oxford University Press.

Jelinski, Z., and Moranda, P. B. (1972), "Software Reliability Research," in *Statistical Computer Performance Evaluation*, ed. W. Freiberger, London: Academic Press, pp. 465–484.

Jewell, W. S. (1985), "Bayesian Extensions to a Basic Model of Software Reliability," *IEEE Transactions on Software Engineering*, 12, 1465–1471.

Joe, H., and Reid, N. (1985), "Estimating the Number of Faults in a System," *Journal of the American Statistical Association*, 80, 222–226.

Littlewood, B. (1981), "Stochastic Reliability Growth: A Model for Fault Removal in Computer Programs and Hardware Designs," *IEEE Transactions on Reliability*, 30, 313–320.

Littlewood, B., and Verrall, J. L. (1981), "Likelihood Function of a Debugging Model for Computer Software Reliability," *IEEE Transactions on Reliability*, 30, 145–148.

Meinhold, R. J., and Singpurwalla, N. D. (1983), "Bayesian Analysis of a Commonly Used Model for Describing Software Failures," *The Statistician*, 32, 168–173.

Miller, D. R. (1986), "Exponential Order Statistic Models of Software Reliability Growth," *IEEE Transactions on Software Engineering*, 12, 12–24.

Morris, C. N. (1983), "Parametric Empirical Bayes Inference: Theory and Applications" (with discussion), *Journal of the American Statistical Association*, 78, 47–65.

Raftery, A. E. (1986), "A Note on Bayes Factors for Log-Linear Contingency Table Models With Vague Prior Information," *Journal of the Royal Statistical Society*, Ser. B, 48, 249–250.

——— (in press), "Analysis of a Simple Debugging Model," *Applied Statistics*, 37.

Raftery, A. E., and Akman, V. E. (1986), "Bayesian Analysis of a Poisson Process With a Change-Point," *Biometrika*, 73, 85–89.

Spiegelhalter, D. J., and Smith, A. F. M. (1982), "Bayes Factors for Linear and Log-Linear Models With Vague Prior Information," *Journal of the Royal Statistical Society*, Ser. B, 44, 377–387.