

BAYESIAN INFERENCE FOR THE POWER LAW PROCESS

SHAUL K. BAR-LEV, IDIT LAVI AND BENJAMIN REISER

Department of Statistics, University of Haifa, Haifa 31999, Israel

(Received December 17, 1990; revised November 8, 1991)

Abstract. The power law process has been used to model reliability growth, software reliability and the failure times of repairable systems. This article reviews and further develops Bayesian inference for such a process. The Bayesian approach provides a unified methodology for dealing with both time and failure truncated data. As well as looking at the posterior densities of the parameters of the power law process, inference for the expected number of failures and the probability of no failures in some given time interval is discussed. Aspects of the prediction problem are examined. The results are illustrated with two data examples.

Key words and phrases: Power law process, Bayesian inference, prediction, repairable system.

1. Introduction

The power law process can be described as a nonhomogeneous Poisson process $\{N(t), t \geq 0\}$ with intensity function $\nu(t) = \beta t^{\beta-1}/\alpha^\beta$, $\alpha > 0$, $\beta > 0$, and mean value function $m(t) = E(N(t)) = (t/\alpha)^\beta$. This has been widely used in the literature to model reliability growth (Crow (1982)), software reliability (Kyparisis and Singpurwalla (1985)), and more generally repairable systems (Ascher and Feingold (1984), Engelhardt and Bain (1986), Rigdon and Basu (1989)). We follow Ascher (1981) in using the term power law process being convinced by his arguments that this is superior to the more frequently used Weibull process terminology.

An alternate way of describing the power law process is to consider the sequence of successive failure times T_1, T_2, \dots , where T_i is the time of the i -th failure. Then the time of the first failure T_1 has a Weibull distribution with scale and shape parameters α and β , respectively, while the failure time T_i , conditional on $T_1 = t_1, \dots, T_{i-1} = t_{i-1}$, has a truncated Weibull distribution with left truncation point t_{i-1} .

Inference on the power law process has generally been considered in the literature from a frequency viewpoint. Two sampling schemes are usually considered; failure truncation and time truncation. For these sampling schemes the large literature on point estimation, confidence intervals, and tests of hypotheses for the parameters α , β , and the intensity function at the end of the testing period is

reviewed by Rigdon and Basu (1989). Prediction for times of future failures is discussed by Lee, L. and Lee, K. (1978) and Engelhardt and Bain (1978), while Miller (1984) examines the prediction of the future intensity parameter. Calabria *et al.* (1988) discuss modified maximum likelihood estimators of the expected number of failures in a given time interval and of the failure intensity and compare their mean squared errors with those of the maximum likelihood estimators.

From a Bayesian perspective, Guida *et al.* (1989) discuss point and interval estimation for α and β assuming failure truncation data and using several different choices of priors, both informative and noninformative. Kyparisis and Singpurwalla (1985) analyse both interval and failure truncation data by employing informative priors on α and β , and derive prediction distributions of future failure times and the number of failures in some future time interval. Their predictive and posterior distributions generally require complicated numerical computations. Calabria *et al.* (1990) also derive predictive distributions for future failure times using both informative and noninformative priors and note the numerical equivalence with classical methods when noninformative priors are used. The above three references usually assume that the prior distributions on α and β are independent. Alternatively Calabria *et al.* (1990) consider independent priors on α and the mean value function $m(t)$, with t fixed.

The above references are concerned with data from a single system for which inference is required either on the parameters of the model for that system or on prediction for the future of that particular system. Crow (1974) and Bain (1978) discuss multisample estimation procedures where data are available on several independent systems which are assumed to be equivalent in the sense that the same power law process applies to them all. The related question of prediction from one system on which data are available to another equivalent system for which no data are available (in this context see Remark in Section 4) has apparently not been considered in the literature. This problem can arise, for example, if data are available on, say, an aircraft generator and another generator of the same model is to begin working under similar conditions as the first.

In this paper we further develop Bayesian inference for the power law process. We use vague noninformative priors which are motivated in Section 2. In many situations only weak prior information is available. Frequently, strong prior knowledge is only readily available to the manufacturer of the system, while the data analysis is of interest to the buyer who does not wish to use the prior of the manufacturer since in a sense he is testing how good the manufacturer is. In addition, analysis based on vague priors can be used as a base line to examine the effect of informative priors. We further show that these priors generally provide mathematically tractable results.

In Section 2 we indicate how a Bayesian analysis handles both failure and time truncation data in the same manner, in contrast to the frequency approach. Previous Bayesian results have not considered time truncation. Section 3 presents posterior densities for the parameters α and β and compares these results with those obtainable from the frequency approach. These posterior densities are similar to those previously obtained by Guida *et al.* (1989) and Kyparisis and Singpurwalla (1985), but include both failure and time truncation data. Section 4 discusses

Table 1. Failure times in hours for aircraft generator.

$i =$ Failure number	$t_i =$ Failure time
1	55
2	166
3	205
4	341
5	488
6	567
7	731
8	1308
9	2050
10	2453
11	3115
12	4017
13	4596

Table 2. Software failure times in seconds.

$i =$ Failure number	$t_i =$ Failure time	i	t_i	i	t_i	i	t_i
1	115	11	1955	21	6162	31	36800
2	115	12	2026	22	6552	32	37363
3	198	13	2632	23	8415	33	40133
4	376	14	3821	24	9752	34	40785
5	570	15	3861	25	14260	35	46378
6	706	16	4649	26	15094	36	58074
7	1783	17	4871	27	18494	37	64798
8	1798	18	4943	28	18500	38	67344
9	1813	19	5558	29	23061		
10	1905	20	6147	30	26229		

prediction of the number of failures in some specified time interval and of future failure times. Both predictions for the particular system under observation and for a different equivalent system (see Remark in Section 4) are considered. Although some of our results require numerical integration, most are given in a mathematically explicit form in terms of series which are easier to compute. Posterior distributions for system reliability, for the expected number of failures in a specified time interval and for the intensity function are derived in Section 5.

Throughout this paper we illustrate some of the results with two numerical examples. The first example involves the failure times of a “complex type of aircraft generator”. These data, due to Duane (1964), and discussed by Rigdon and Basu (1989), are presented in Table 1. They will be used to illustrate results relating to a new system equivalent to the one from which the data are taken, as discussed above.

The second example involves software failures data, taken from Musa (1979) (also presented as Table 1 in Kyparisis and Singpurwalla (1985)), and is given in our Table 2. For software failures, the particular "system" being analyzed is of interest, and for this system interest is focused on system reliability and future prediction.

2. The likelihood function and the choice of priors for (α, β)

We discuss two different sampling protocols which provide data on the power law process: (i) failure truncation and (ii) time truncation. For (i), the number of failures, say n , is fixed before testing begins and ordered system failure times T_1, T_2, \dots, T_n with observed values $t_1 < t_2 < \dots < t_n$ are obtained. For (ii), let t be a predetermined time and suppose $n \geq 1$ failures are observed during $(0, t]$ with failure times $t_1 < t_2 < \dots < t_n$.

The resulting likelihoods for both protocols can be written in a unified form (see for example, Crow (1982)) as

$$(2.1) \quad L(\alpha, \beta | \underline{t}) = (\beta/\alpha)^n \left(\prod_{i=1}^n t_i/\alpha \right)^{\beta-1} e^{-(y/\alpha)^\beta}, \quad \underline{t} = (t_1, \dots, t_n),$$

where $y = t_n$ for failure truncation and $y = t$ for time truncation. (The case of $n = 0$ for time truncation gives a different likelihood but is of no inferential interest and thus will not be considered.)

From the Bayesian perspective both failure and time truncation data can be handled in the same manner through (2.1) and result in the same type of posterior inference on α and β in contrast to the classical frequency approach (Crow (1982)) in which each case must be treated separately and different types of results are obtained.

We consider two types of noninformative vague priors for α and β . We motivate our choice of such priors by just considering the time T_1 to the first failure. This has a Weibull distribution with scale parameter α and shape parameter β . The use of the distribution of T_1 allows us to use the location-scale model and is justified by the fact that our prior knowledge is not affected by whether we observe T_1 or T_1, \dots, T_n . Suggestions for informative priors for the Weibull distribution can be found in Singpurwalla (1988). Indeed, the distribution of $\log T_1$ can be written in the form of a location-scale distribution with location parameter $\mu = \log \alpha$ and scale parameter $\alpha = \beta^{-1}$. Following Jeffrey's rule for noninformative priors in the location-scale situation (see Box and Tiao (1973), pp. 56–57), we consider two cases: (a) μ and σ are known to be independent *a priori*, and (b) the prior independence assumption is ignored. The noninformative prior of (μ, σ) will be then proportional to σ^{-1} , for (a), and to σ^{-2} , for (b). Transforming $(\mu, \sigma) \rightarrow (\alpha, \beta)$ results then in $f(\alpha, \beta) \propto (\alpha\beta)^{-1}$ and $f(\alpha, \beta) \propto \alpha^{-1}$, as the noninformative priors of (α, β) for cases (a) and (b), respectively.

For notational convenience we consider the prior

$$(2.2) \quad f(\alpha, \beta) \propto (\alpha\beta^\gamma)^{-1}, \quad \alpha > 0, \quad \beta > 0,$$

with $\gamma = 0$ and $\gamma = 1$ corresponding to the cases of dependence and independence, respectively. The case $\gamma = 1$ is most commonly used (Box and Tiao (1973)). The prior $(\alpha\beta)^{-1}$ has been used by Guida *et al.* (1989).

A referee has pointed out that the priors in (2.2) can alternatively be easily obtained (following Box and Tiao (1973)) by searching for a transformation for each parameter such that in terms of this transformation the likelihood function is data translated and then taking the transformed parameter to be locally uniform. We feel that our above argument based on the location-scale properties of the transformed Weibull distribution has some interest and helps clarify the role of the prior independence assumption.

3. Inference on (α, β)

From (2.1) and (2.2) the posterior density of (α, β) can be seen to be

$$(3.1) \quad f(\alpha, \beta | \underline{t}) = c_\gamma(\underline{t})\beta^{n-\gamma} \left(\prod_{i=1}^n t_i \right)^\beta e^{-(y/\alpha)^\beta} / \alpha^{n\beta+1}, \quad \alpha > 0, \quad \beta > 0,$$

where

$$(3.2) \quad c_\gamma(\underline{t}) = \left(\ln \prod_{i=1}^n y/t_i \right)^{n-\gamma} / \left(\Gamma(n)\Gamma(n-\gamma) \right), \quad \gamma = 0, 1.$$

Note that for $\gamma = 0$ and $n = 1$ this posterior density does not exist (i.e., is improper). Heuristically, one may argue that with only vague prior information the case $n = 1$ should not allow inference on both parameters.

Marginal inference on β is often of particular interest especially for reliability growth models. When $\beta = 1$, the power law process becomes a homogeneous Poisson process. For $\beta > 1$, the frequency of failures increases with time and thus the system is degrading in a reliability sense, while for $\beta < 1$ the failure frequency decreases with time and thus the system is improving. If the posterior density of β is heavily concentrated over $(0, 1)$, this would imply that there is an improvement over time. The marginal posterior density of β can be obtained by integrating out α in (3.1). This gives that *a posteriori*

$$(3.3) \quad \beta | \underline{t} \sim (\hat{\beta}/2n)\chi_{2(n-\gamma)}^2,$$

where

$$(3.4) \quad \hat{\beta} = n / \sum_{i=1}^n \ln y/t_i.$$

The $\gamma = 1$ case has been given by Guida *et al.* (1989). From (3.3), one or two-sided Bayesian probability intervals can be readily obtained. Note that for $\gamma = 1$, these will be numerically the same as confidence intervals obtained for failure truncation

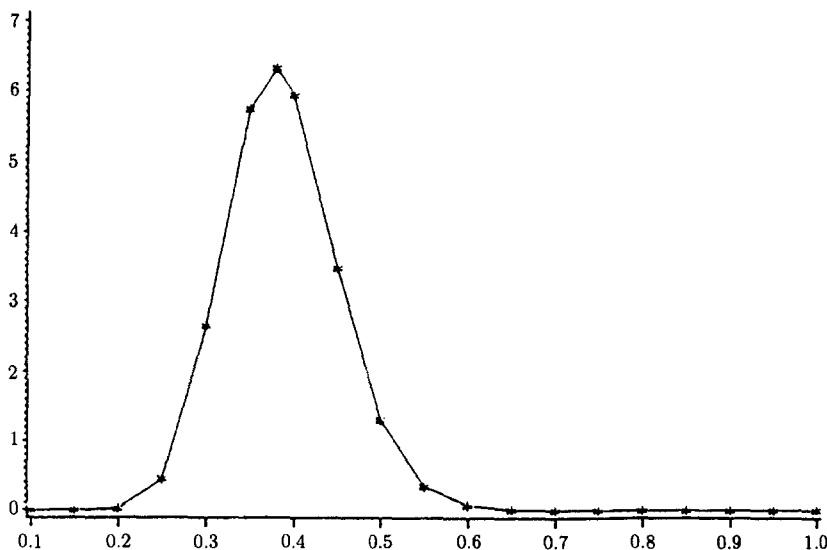


Fig. 1. The posterior density of β for the data of Table 2.

data, but will differ slightly from confidence intervals obtained for time truncation data (e.g., Rigdon and Basu (1989)). Figure 1 displays the posterior density of β for the data of Table 2. It can be seen that this posterior is concentrated around 0.4 and indicates a growth in software reliability. Kyparisis and Singpurwalla (1985) analysed these data using several distinctly different priors and obtained fairly similar results. For these data the details of the prior information does not tend to influence the posterior significantly.

The posterior density of α is obtained from (3.1) to be

$$(3.5) \quad f(\alpha | \underline{t}) = c_{\gamma}(\underline{t})\alpha^{-1} \int_0^{\infty} \beta^{n-\gamma} \left(\prod_{i=1}^n t_i/\alpha \right)^{\beta} e^{-(y/\alpha)^{\beta}} d\beta.$$

Bayesian probability intervals can be computed by either numerically integrating (3.5) or as follows. Following Lee, L. and Lee, K. (1978), we define $\hat{\alpha} = yn^{-1/\hat{\beta}}$ and $W = (\hat{\alpha}/\alpha)^{\beta}$, where $\hat{\beta}$ is given by (3.4). After some algebraic manipulation it can be shown that

$$(3.6) \quad P(W \leq w | \underline{t}) = \int_0^{\infty} G((nw)^{x/2n}) g_{\gamma}(x) dx,$$

where

$$(3.7) \quad g_{\gamma}(x) = (1/2)^{n-\gamma} x^{n-\gamma-1} e^{-x/2} / \Gamma(n-\gamma), \quad G(x) = \int_0^x v^{n-1} e^{-v} / \Gamma(n) dv.$$

For $\gamma = 1$ and $y = t_n$, (3.6) coincides with the expression obtained by Lee, L. and Lee, K. (1978) for the distribution of W , a pivotal quantity for α . Accordingly,

for the failure truncation case, with $\gamma = 1$, the Bayesian probability intervals are numerically equivalent to confidence intervals for α in the classical approach. Of course, their interpretation differs.

4. Prediction distributions

In this section we derive the prediction distributions of the number of future failures in some time interval and of future failure times. Two cases are of particular interest:

(I) A system has been observed until time y and we want to predict the number of future failures in some interval $(s_1, s_2]$, $y \leq s_1$, and the future failure times for the particular system we are observing.

(II) Data on one system have been obtained. Another equivalent system (in the sense that it has exactly the same intensity function) is to begin operating and we want to predict failure times and the number of failures of the new system over the interval $(0, s]$, for some s of common interest.

Remark. As noted above, various aspects of prediction for the system under observation have been discussed in the literature. We feel that inference for an "equivalent" system is often of interest. For example, we may have conducted an experiment on a number of printers of the same type working under the same operating conditions and be interested in another printer (of the same type) just beginning to work under similar conditions. Here, we only consider the problem with data available on one system. The more general problem is currently under research. One of the referees has commented that this problem should not really be considered a prediction problem but just a "procedure by which properties conferred from given sample data are attached to any item which is assumed to be equivalent". Furthermore the referee points out, quite correctly, that the posterior obtained for the system under observation is actually an informative prior for the equivalent system. Consequently, in the absence of data on the equivalent system, inference on the number of failures occurring on a time interval for the equivalent system (and other quantities of interest) can be obtained by standard probability arguments and should not be considered prediction in the same way as in case (I). We basically agree with this distinction but still find the word prediction to be useful at least in an informal sense for case (II). We note that Geisser (1986) states that "Statistical prediction is the process by which values for unknown observables (potential observations yet to be made or past ones which are no longer available) are inferred based on current observations and other information at hand." Our case (II) falls under this definition.

In the following we frequently make use of the fact that for $c > 0$,

$$(4.1) \quad \int_0^{\infty} \beta^l c^\beta d\beta = \begin{cases} \Gamma(l+1)/(-\ln c)^{l+1}, & l+1 > 0, \quad 0 < c < 1 \\ \infty, & \text{otherwise,} \end{cases}$$

which facilitates calculations.

4.1 Prediction of the number of failures in some time interval

Let $N(s_1; s_2)$ denote the number of failures occurring in the interval $(s_1, s_2]$, then

$$(4.2) \quad \begin{aligned} P[N(s_1; s_2) = r \mid \alpha, \beta] \\ = [(s_2/\alpha)^\beta - (s_1/\alpha)^\beta]^r \exp\{-[(s_2/\alpha)^\beta - (s_1/\alpha)^\beta]\}/r!, \\ r = 0, 1, \dots \end{aligned}$$

Consequently, the predictive distribution of $N(s_1; s_2)$ is

$$(4.3) \quad \begin{aligned} P[N(s_1; s_2) = r \mid \underline{t}] &= \int_0^\infty \int_0^\infty P[N(s_1; s_2) = r \mid \alpha, \beta] f(\alpha, \beta \mid \underline{t}) d\alpha d\beta \\ &= \frac{\Gamma(n+r)}{\Gamma(r+1)} c_\gamma(\underline{t}) \int_0^\infty \beta^{n-\gamma-1} \left(\prod_{i=1}^n t_i \right)^\beta \\ &\quad \cdot (s_2^\beta - s_1^\beta)^r (s_2^\beta - s_1^\beta + y^\beta)^{-(n+r)} d\beta, \\ &\quad n > \gamma, \quad r = 0, 1, \dots \end{aligned}$$

In general this would need to be computed numerically. However, for the important special case of (I) where $s_1 = y$, (4.3) with the help of (4.1) and the binomial formula, reduces to

$$(4.4) \quad \begin{aligned} P[N(y; s_2) = r \mid \underline{t}] \\ = \binom{n+r-1}{r} \sum_{k=0}^n \binom{r}{k} (-1)^k \left[\frac{\ln \prod_{i=1}^n y/t_i}{\ln \prod_{i=1}^n s_2/t_i + \ln(s_2/y)^k} \right]^{n-\gamma}, \\ n > \gamma, \quad r = 0, 1, \dots \end{aligned}$$

Using a different prior, Kyparisis and Singpurwalla (1985) obtain a complicated alternative to (4.4) which requires special numerical techniques to compute. Figure 2 displays the predictive probability function of $N(y; s_2)$ given in (4.4) for the data of Table 2 with $y = 67344$, $s_2 = 80000$ and $\gamma = 1$. These particular values were chosen to allow comparison with Fig. 3.3 of Kyparisis and Singpurwalla (1985). The results are quite similar, both having a modal value 2 and very small probability of more than 8 failures in this interval.

For case (II), $s_1 = 0$, $s_2 = s$, (4.3) becomes

$$(4.5) \quad \begin{aligned} P[N(0; s) = r \mid \underline{t}] \\ = c_\gamma(\underline{t}) \frac{\Gamma(n+r)}{\Gamma(r+1)} \int_0^\infty \beta^{n-\gamma-1} \left(\prod_{i=1}^n t_i/s \right)^\beta (1 + (y/s)^\beta)^{-(n+r)} d\beta, \\ n > \gamma, \quad r = 0, 1, \dots \end{aligned}$$

$$P(N(67,344; 80,000) = r)$$

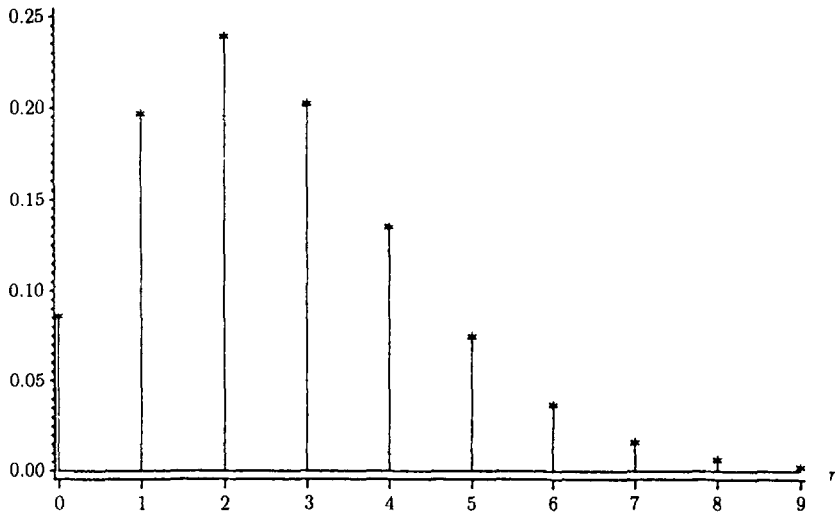


Fig. 2. The predictive probability function of $N(67, 344; 80, 000)$ for the data of Table 2.

The integral in (4.5) can be computed numerically. Consider, however, the two special cases: $s > y$ and $s = y$. Using (4.1) and the Taylor expansion for $(1 + a)^{-m}$, $|a| < 1$, (4.5) for $s > y$ can be written as an infinite sum,

$$(4.6) \quad P[N(0; s) = r | \underline{t}]$$

$$= \binom{n+r-1}{r} \sum_{j=0}^{\infty} \binom{-n-r}{j} \left[\frac{\ln \prod_{i=1}^n y/t_i}{\ln \prod_{i=1}^n s/t_i + \ln(s/y)^j} \right]^{n-\gamma},$$

$s > y, \quad n > \gamma, \quad r = 0, 1, \dots$

Whereas for $s = y$, (4.5) reduces to

$$(4.7) \quad P[N(0; y) = r | \underline{t}] = \binom{n+r-1}{n-1} (1/2)^{n+r}, \quad n > \gamma, \quad r = 0, 1, \dots$$

i.e., $N(0; y)$ is a negative binomial variate counting the number of failures occurring prior to observing the n -th success. (The result in (4.7) could have been obtained also by using combinatorial methods). Note that the distribution of $N(0; y)$ depends neither on the t_i 's, the actual observed failure times, nor on γ , but does depend on the number of failures n in the interval $(0, y]$. The distribution in (4.7) has two modes at $r = n - 1$ and $r = n$. Figure 3 displays (4.7) for the data of Table 1 ($n = 13$) for $r = 1(1)24$. Other values of r each have probability less than 0.0092. Consequently we would not expect more than 24 failures for a generator equivalent to ours operating over 4596 hours.

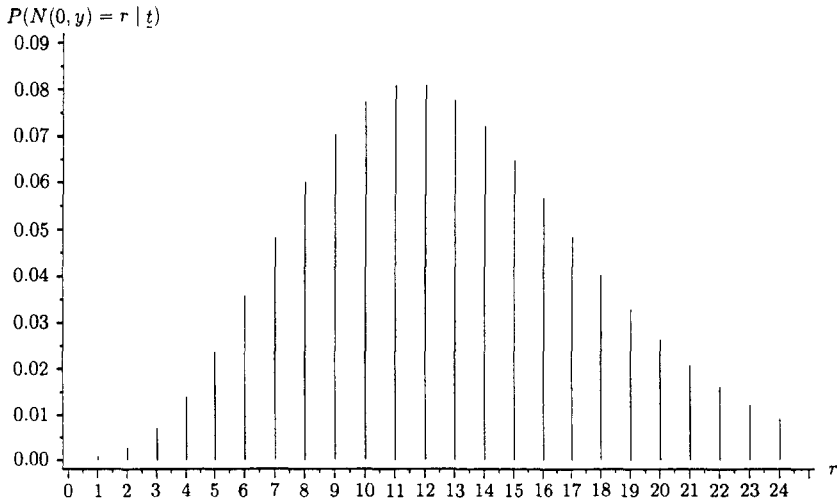


Fig. 3. The predictive probability function (4.7) of $N(0; 4596)$ for the data of Table 1.

4.2 Prediction of future failure times

We first consider case (I). For this case we restrict ourselves to the failure truncation case for which $y = t_n$ (the time truncation case leads to cumbersome calculations). We are interested, given current available data t_1, \dots, t_n , to predict the future $(n+r)$ -th failure time T_{n+r} . Often T_{n+1} , the next failure is of particular interest.

Define $Z_r = T_{n+r} - T_n$. Conditional on the observation $T_n = t_n$, predicting T_{n+r} is equivalent to predicting Z_r . Calabria *et al.* (1990) obtain the required predictive distribution using the noninformative prior (2.2) with $\gamma = 1$. Earlier, Kyparisis and Singpurwalla (1985) discussed this problem using a different prior. For the prior (2.2) it is readily shown that

$$(4.8) \quad P(Z_r \leq z | t) = r \binom{n+r-1}{r} \sum_{k=0}^{r-1} \frac{1}{n+k} \binom{r-1}{k} (-1)^k \cdot \left[1 - \left(\frac{\ln \prod_{i=1}^n t_n/t_i}{\ln \prod_{i=1}^n \frac{(t_n+z)}{t_i} + \ln \left(\frac{t_n+z}{t_n} \right)^k} \right)^{n-\gamma} \right],$$

$n > \gamma, \quad z \geq 0, \quad r = 1, 2, \dots$

For the special case $r = 1$, (4.8) reduces to

$$(4.9) \quad P(Z_1 \leq z | t) = 1 - \left[\frac{\ln \prod_{i=1}^n t_n/t_i}{\ln \prod_{i=1}^n \frac{(t_n+z)}{t_i}} \right]^{n-\gamma}, \quad z \geq 0, \quad n > \gamma,$$

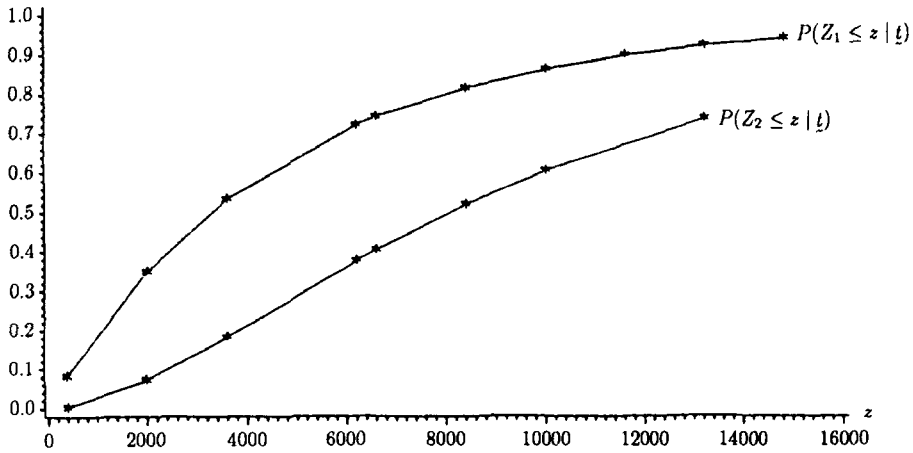


Fig. 4. The predictive distributions of Z_1 and Z_2 for the data of Table 1.

i.e., $Z_1 | \underline{t}$ equals in distribution to the minimum of $n - \gamma$ i.i.d. r.v.'s with common cumulative distribution

$$1 - \left[\frac{\ln \prod_{i=1}^n t_n/t_i}{\ln \prod_{i=1}^n \frac{(t_n + z)}{t_i}} \right].$$

This result is useful for simulation purposes.

The predictive distribution of T_{n+r} can be computed through the relation

$$(4.10) \quad P(T_{n+r} \leq \tau | \underline{t}) = P(Z_r \leq \tau - t_n | \underline{t}),$$

where the right-hand side of (4.10) can be computed from (4.8).

Kyparisis and Singpurwalla (1985) obtained results which require substantial numerical integration in contrast to the explicit solution given here. Figure 4 displays the predictive distribution of Z_r given by (4.8) for $r = 1, 2$ and the data of Table 2. It can be seen that for Z_1 the results are quite similar to Fig. 3.4 of Kyparisis and Singpurwalla (1985). The next failure will almost certainly occur within the next 14000 seconds and within the next 2000 seconds there is a probability of about 1/3 of a failure occurring.

We now consider case (II) for both failure and time truncation. Let W_r denote the r -th failure time of an equivalent new system which is to begin operating. Using the fact that $(W_r/\alpha)^\beta$ has a gamma distribution with scale parameter 1 and shape parameter r , the prediction distribution of W_r given \underline{t} is

$$(4.11) \quad P(W_r \leq w | \underline{t}) = \int_0^\infty \int_0^\infty \left(\int_0^{(w/\alpha)^\beta} \frac{x^{r-1} e^{-x}}{(r-1)!} dx \right) f(\alpha, \beta | \underline{t}) d\alpha d\beta$$

$$\begin{aligned}
 &= \frac{c_\gamma(\underline{t})}{(r-1)!} \Gamma(n+r) \\
 &\quad \cdot \int_0^\infty \beta^{n-\gamma-1} \left(\prod_{i=1}^n t_i/w \right)^\beta \left[\int_0^1 \frac{x^{r-1}}{[x + (y/w)\beta]^{n+r}} dx \right] d\beta, \\
 &\hspace{15em} n > \gamma, \quad w \geq 0, \quad r = 1, 2, \dots
 \end{aligned}$$

Two special cases might be of interest: $w = y$ and $r = 1$. For $w = y$, we get, using Gradshteyn and Ryzhik ((1965), 3.194(8), p. 285) and (4.1), that

$$(4.12) \quad P(W_r \leq y) = r \binom{n+r-1}{r} 2^{-r} \sum_{k=0}^{n-1} \binom{n-1}{k} \frac{(-2)^{-k}}{r+k},$$

$n > \gamma, \quad r = 1, 2, \dots,$

a result that could alternatively be obtained by a combinatorial method.

For $r = 1$ and $w > y$, we get by using (4.1) and the Taylor expansion of $(1+a)^{-n}$, $|a| < 1$, that

$$(4.13) \quad P(W_1 \leq w | \underline{t}) = 1 - \sum_{j=0}^{\infty} \binom{-n}{j} \left[\frac{\ln \prod_{i=1}^n y/t_i}{\ln \prod_{i=1}^n w/t_i + \ln(w/y)^j} \right]^{n-\gamma},$$

$w > y, \quad n > \gamma.$

5. Posterior distributions for some parameter functions

We derive here posterior distributions of functions of (α, β) which are of particular interest. The parameter functions considered are: system reliability, expected number of failures on a given interval and the intensity function.

We define system reliability to be the probability of no failures over a specified time interval. For a given repairable system for which data have been collected, a high reliability over some future time of interest will effect decisions on replacement and maintenance policy and on the purchase of spares. Clearly, similar considerations arise when dealing with a new equivalent system. In terms of reliability growth (hardware or software) a high reliability for some period of interest may imply that it is worthwhile ending the development process. Similar considerations make the expected number of failures in a given interval of interest.

As Crow (1975) points out once a development project ends and no further modifications are made in the system, then failures are often assumed to occur with a constant failure rate which takes on the current value of the intensity function. Thus inference on the current intensity function value is of interest.

5.1 *The posterior distribution for system reliability*

As in Section 4 we consider two cases of particular concern: (I) Data on system are available over $(0, y)$, where $y = t_n$ or t depending on whether failure or time truncation was used and we want to make inference for the future; i.e., for the interval $(y, s]$ and (II) Data are available on a system as in (I), on which we want to base inference on an equivalent system which is just beginning to operate; i.e., over the time interval $(0, s]$. For cases (I) and (II) the system reliability, as defined above, is obtained from (4.2) to be, respectively,

$$(5.1) \quad w = w(y, s) = P[N(y; s) = 0] = \exp\{-[(s/\alpha)^\beta - (y/\alpha)^\beta]\}, \quad s > y,$$

and

$$(5.2) \quad w = w(s) = P[N(s) = 0] = \exp\{-(s/\alpha)^\beta\}.$$

For case (I), the posterior density of w can then be shown to be

$$(5.3) \quad f(w | \underline{t}) = c_\gamma(\underline{t})(-\ln w)^{n-1}w^{-1} \int_0^\infty \frac{\beta^{n-\gamma-1} \left(\prod_{i=1}^n t_i/s\right)^\beta w^{[(s/y)^\beta - 1]^{-1}}}{[1 - (y/s)^\beta]^n} d\beta, \\ 0 < w < 1, \quad s > y, \quad n > \gamma.$$

For case (II), it is

$$(5.4) \quad f(w | \underline{t}) = c_\gamma(\underline{t})(-\ln w)^{n-1}w^{-1} \int_0^\infty \beta^{n-\gamma-1} \left(\prod_{i=1}^n t_i/s\right)^\beta w^{(y/s)^\beta} d\beta, \\ 0 < w < 1, \quad n > \gamma,$$

which, for $s = y$, reduces to

$$(5.5) \quad f(w | \underline{t}) = (-\ln w)^{n-1}/\Gamma(n), \quad 0 < w < 1.$$

For given data, a numerical integration of (5.4), even with double precision, is cumbersome and requires an extensive amount of computing time for convergence. This computation, however, is significantly facilitated by writing (5.4), for $y < s$, in terms of an infinite sum as follows. Write

$$w^{(y/s)^\beta} = \exp\{(y/s)^\beta \ln w\} = \sum_{i=0}^\infty (y/s)^{i\beta} (\ln w)^i / i!.$$

Substituting this in (5.4), interchanging the order of integration and infinite summation (the legitimacy of which can easily be verified), and using (4.1), yields

$$(5.6) \quad f(w | \underline{t}) = \frac{(-\ln w)^{n-1}w^{-1}}{\Gamma(n)} \sum_{i=0}^\infty \frac{(\ln w)^i}{i!} \left[\frac{\ln \left(\prod_{j=1}^n t_j/y\right)}{\ln \left(\prod_{j=1}^n t_j/s\right) + \ln(y/s)^i} \right]^{n-\gamma}, \\ 0 < w < 1, \quad s > y, \quad n > \gamma.$$

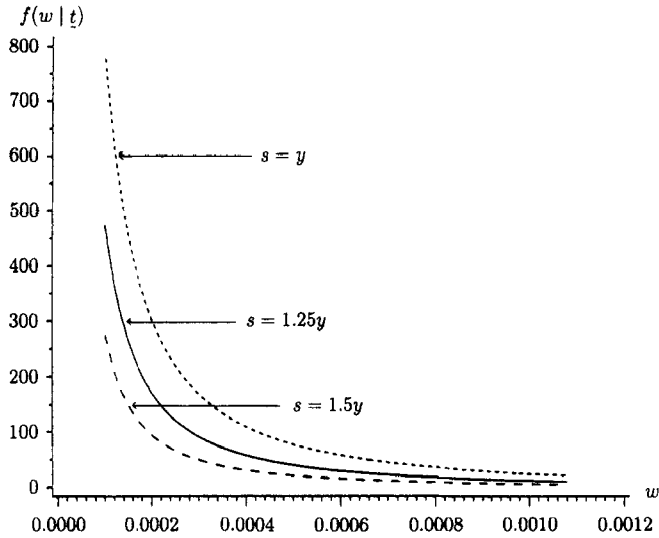


Fig. 5. The posterior density of $w = P(N(s) = 0)$ for $s = y = 4596$ (equation (5.5)), $s = 1.25y, 1.5y$ (equation (5.6)) and $\gamma = 1$ for the data of Table 1.

Figure 5 displays the posterior densities (5.5) (with $s = y$) and (5.6) with $s = 1.25y, s = 1.5y$ and $\gamma = 1$ for the data of Table 1 (in which $y = 4596$). It can be seen that they are all highly concentrated around a system reliability of zero indicating that it is extremely unlikely that an equivalent system would survive as long as 4596 or more hours without a failure.

5.2 Posterior distributions of the expected number of failures in future time interval

The expected number of failures over the intervals $(y, s]$ and $(0, s)$ are $\mu = \mu(y, s) = (s/\alpha)^\beta - (y/\alpha)^\beta$ and $\mu = \mu(s) = (s/\alpha)^\beta$, respectively. Comparing those with (5.1) and (5.2), respectively, we find that $w(y, s) = e^{-\mu(y,s)}$ for case (I) and $w(s) = e^{-\mu(s)}$ for case (II). We therefore immediately obtain from (5.3)–(5.6) that, for case (I), the posterior density of μ is

$$(5.7) \quad f(\mu | \underline{t}) = c_\gamma(\underline{t})\mu^{n-1} \int_0^\infty \frac{\beta^{n-\gamma-1} \left(\prod_{i=1}^n t_i/s \right)^\beta e^{-\mu((s/y)^\beta - 1)^{-1}}}{[1 - (y/s)^\beta]^\beta} d\beta, \quad \mu > 0, \quad n > \gamma.$$

For case (II), it is

$$(5.8) \quad f(\mu | \underline{t}) = c_\gamma(\underline{t})\mu^{n-1} \int_0^\infty \beta^{n-\gamma-1} \left(\prod_{i=1}^n t_i/s \right)^\beta e^{-\mu(y/s)^\beta} d\beta, \quad \mu > 0, \quad n > \gamma,$$

which for $s = y$, reduces to

$$(5.9) \quad \mu \mid \underline{t} \sim \text{gamma}(n, 1).$$

As in (5.6), (5.8) for $s > y$ can be written in terms of an infinite sum as

$$(5.10) \quad f(\mu \mid \underline{t}) = (\mu^{n-1} / \Gamma(n)) \sum_{i=0}^{\infty} \frac{(-\mu)^i}{i!} \left[\frac{\ln \left(\prod_{j=1}^n t_j / y \right)}{\ln \left(\prod_{j=1}^n t_j / s \right) + \ln(y/s)^i} \right]^{n-\gamma},$$

$\mu > 0, \quad s > y, \quad n > \gamma.$

5.3 The posterior distribution for the intensity function

In the context of reliability growth models, as discussed by Crow (1975), if after time y no further improvements are to be made on the system, then it can be assumed that the system will have a constant failure rate taking the value of the intensity function at y ; i.e., $\nu = \nu(y) = (\beta/\alpha)(y/\alpha)^{\beta-1}$. Consequently, if we let y be shifted to the origin, the system reliability for a future interval $(0, t_0]$ will be

$$(5.11) \quad R(t_0) = e^{-\nu(y)t_0}.$$

By using (3.1) and (4.1), we find the posterior density of ν given \underline{t} to be

$$(5.12) \quad f(\nu \mid \underline{t}) = c_\gamma(\underline{t}) y^n \nu^{n-1} \int_0^\infty \beta^{-\gamma-1} \exp \left\{ -\beta \ln \prod_{i=1}^n y/t_i - \nu y / \beta \right\} d\beta,$$

$\nu > 0, \quad n - \gamma > 0.$

Employing Gradshteyn and Ryzhik ((1965), 3.471(9), p. 340), (5.12) can be written as

$$(5.13) \quad f(\nu \mid \underline{t}) = 2c_\gamma(\underline{t}) y^n \nu^{n-1-\gamma/2} \cdot \left(y / \ln \prod_{i=1}^n y/t_i \right)^{-\gamma/2} K_{-\gamma} \left(2 \left(\nu y \ln \prod_{i=1}^n y/t_i \right)^{1/2} \right),$$

$\nu > 0, \quad n - \gamma > 0.$

where $K_{-\gamma}(z)$ is the modified Bessel function of the second kind of order $-\gamma$ with argument z , which is available in terms of infinite series.

Employing Gradshteyn and Ryzhik ((1965), 8.486(16), p. 970, 8.446, p. 961 and 8.447(3), p. 961), we find that $K_0(z)$ and $K_{-1}(z)$ in (5.13) can be expressed as

$$(5.14) \quad K_{-1}(z) = K_1(z) = z^{-1} \sum_{k=0}^{\infty} \frac{(z/2)^{2k+1}}{k!(k+1)!} [\ln z/2 - (1/2)\psi(k+1) - (1/2)\psi(k+2)],$$

and

$$(5.15) \quad K_0(z) = \sum_{k=0}^{\infty} \frac{(z/2)^{2k}}{(k!)^2} [\psi(k+1) - \ln z/2],$$

where $\psi(k+1) = -C + \sum_{j=1}^k j^{-1}$, and C is Euler's constant. (5.14) and (5.15) are quite tractable for plotting (5.13) for specific data and obtaining Bayesian probability intervals. A posterior distribution for $R(t_0)$ can be readily obtained from (5.11).

Now let $\delta = \nu(y)/\hat{\nu}(y)$, where $\hat{\nu}(y) = (\hat{\beta}/\hat{\alpha})(y/\hat{\alpha})^{\hat{\beta}-1}$, $\hat{\alpha} = yn^{1/\hat{\beta}}$ and $\hat{\beta}$ is given by (3.4), then by using (5.12) we obtain

$$(5.16) \quad P(\delta \leq x | t) = \int_0^{\infty} G(2n^2x/z)g_{\gamma}(z)dz, \quad x \geq 0,$$

where g_{γ} and G are defined in (3.7).

For failure truncation; i.e., $\gamma = 1$ and $y = t_n$, (5.16) coincides with the expression obtained by Lee, L. and Lee, K. (1978) for the distribution of δ , a pivotal quantity for ν . Thus, for the failure truncation case, the Bayesian probability intervals are the same numerically as the confidence intervals for ν obtained by Lee, L. and Lee, K. (1978).

6. Conclusions

We have shown how the Bayesian approach provides a unified methodology for inference problems on the power law process. In particular, failure and time truncation data are handled similarly in contrast to the standard frequency methodology which deals with each quite separately and violates the likelihood principle.

In addition to the usual inferential questions which have previously been discussed in the literature, we have developed posterior distributions for the expected number of failures and the probability of no failures over a specified time interval. Also we have shown how data on one system can be used to make predictions on a second equivalent system. The case where data are available independently on a number of equivalent systems and inference is desired on a new system equivalent to the others is open for further research. In principle, the Bayesian approach can be applied to this problem but the resulting multiple integrations cannot be carried out explicitly and require either numerical integration or some accurate approximation methods.

Acknowledgements

We would like to thank Professor Richard N. Schmidt for his advice on the numerical aspects of this work and for carrying out all the computations in this paper. The comments of the referees led to improvements in this work. We particularly want to thank one referee for his detailed remarks and suggestions and his most helpful insights on the prediction problem.

REFERENCES

- Ascher, H. (1981). Weibull distribution vs. Weibull process, *Annual Reliability and Maintainability Symposium, IEEE*, 426–431, Chicago, IL.
- Ascher, H. and Feingold, H. (1984). *Repairable Systems Reliability*, Dekker, New York.
- Bain, L. J. (1978). *Statistical Analysis of Reliability and Life-Testing Models*, Dekker, New York.
- Box, G. E. P. and Tiao, G. C. (1973). *Bayesian Inference in Statistical Analysis*, Addison-Wesley, Reading, Massachusetts.
- Calabria, R., Guida, M. and Pulcini, G. (1988). Some modified maximum likelihood estimators for the Weibull process, *Reliability Engineering and System Safety*, **23**, 51–58.
- Calabria, R., Guida, M. and Pulcini, G. (1990). Bayes estimation of prediction intervals for a power law process, *Comm. Statist. Theory Methods*, **19**, 3023–3035.
- Crow, L. H. (1974). Reliability analysis for complex repairable systems, *Reliability and Biometry* (eds. F. Proschan and D. J. Serfling), 379–410, SIAM, Philadelphia.
- Crow, L. H. (1975). Tracking reliability growth, Proceedings of the Twentieth Conference on Design of Experiments, *ARO Rep.* 75, 741–754.
- Crow, L. H. (1982). Confidence interval procedures for the Weibull process with applications to reliability growth, *Technometrics*, **24**, 67–72.
- Duane, J. T. (1964). Learning curve approach to reliability monitoring, *IEEE Trans. Aerospace Electron. Systems*, **2**, 563–566.
- Engelhardt, M. and Bain, L. J. (1978). Prediction intervals for the Weibull process, *Technometrics*, **20**, 167–169.
- Engelhardt, M. and Bain, L. J. (1986). On the mean time between failures for repairable systems, *IEEE Transactions on Reliability*, **R-35**, 419–422.
- Geisser, S. (1986). Predictive analysis, *Encyclopedia of Statistical Sciences*, Vol. 7, 158–170, Wiley, New York.
- Gradshteyn, I. S. and Ryzhik, I. M. (1965). *Tables of Integrals, Series, and Products*, Academic Press, New York.
- Guida, M., Calabria, R. and Pulcini, G. (1989). Bayes inference for a non-homogeneous Poisson process with power intensity law, *IEEE Transactions on Reliability*, **R-38**, 603–609.
- Kyparisis, J. and Singpurwalla, N. D. (1985). Bayesian inference for the Weibull process with applications to assessing software reliability growth and predicting software failures, *Computer Science and Statistics: Proceedings of the Sixteenth Symposium on the Interface* (ed. L. Billard), 57–64, Elsevier Science Publishers B. V. (North Holland), Amsterdam.
- Lee, L. and Lee, K. (1978). Some results on inference for the Weibull process, *Technometrics*, **20**, 41–45.
- Miller, G. (1984). Inference on a future reliability parameter with the Weibull process model, *Naval Res. Logist. Quart.*, **31**, 91–96.
- Musa, J. D. (1979). *Software Reliability Data*, deposited in IEEE Computer Society Repository, New York.
- Rigdon, S. E. and Basu, A. P. (1989). The power law process: a model for the reliability of repairable systems, *Journal of Quality Technology*, **21**, 251–260.
- Singpurwalla, N. D. (1988). An interactive PC-based procedure for reliability assessment incorporating expert opinion and survival data, *J. Amer. Statist. Assoc.*, **83**, 43–51.