The Power Law Process: A Model for the Reliability of Repairable Systems

STEVEN E. RIGDON

Southern Illinois University at Edwardsville, Edwardsville, Illinois 62026

ASIT P. BASU

University of Missouri, Columbia, Missouri 65211

The power law process, often misleadingly called the Weibull process, is a useful and simple model for describing the failure times of repairable systems. We present elementary properties of the power law process, such as point estimation of unknown parameters, confidence intervals for the parameters, and tests of hypotheses. It is shown that an appropriate transformation of the failure times can lead to a goodness-of-fit test. We also discuss some of the terminology used for repairable systems and contrast this with terminology used for nonrepairable systems.

Introduction

A LARGE amount of the literature on reliability deals with units which can fail only once, for example, light bulbs, computer chips, or fan belts. After failure these items are discarded. Compared to the vast literature on the reliability of these nonrepairable items, there is relatively little written on the reliability of repairable systems, and what little there is, is often mired in misconceptions and poor terminology. Ascher and Feingold (1984, p. 133) state that ". . . the prevalent terminology could scarcely be more misleading if it had been designed to mislead—specifically, it has engendered such deep-seated misconceptions that it is extraordinarily difficult to supplant it with improved nomenclature." The objectives of this article are (1) to point out to practicing reliability and quality engineers the usefulness, as well as the simplicity, of one particular model for repairable systems, the power law process and (2) to clarify some of the misconceptions and poor terminology.

The terms used to describe the power law process present one point of confusion. The power law process has been called the Weibull-restoration process (Bassin [1973]), the nonhomogeneous Poisson process with Weibull intensity function (Crow [1974]), the Rasch-Weibull process (Moller [1976]), the Weibull Poisson process (Bain and Engelhardt [1986]; Bain, Engelhardt, and Wright [1985]; and Rigdon and Basu [1988]), and most frequently the Weibull process (Bain [1978] and many others). The term power law process, a phrase describing the functional form of the intensity function (defined in the next section), was introduced by Ascher (1981).

The power law process has proved to be a useful model for several reasons. First, it can be used to model systems which are deteriorating (the times between failures are getting shorter) as well as systems which are improving in time (the times between failures are getting larger). Second, Duane (1964) showed that many systems developed at General Electric fit a model that is closely related to the power law process. He showed that on log-log paper a plot of cumulative operating time t against $t/N(t)$, where $N(t)$ is the number of failures before time t, was nearly linear for many of these systems. Such a condition is necessary but not sufficient for the power law process. Duane assumed that the average number of failures before time t is a certain increasing function of t; thus his model is deterministic and is not equivalent to the power law process. A third reason for the popularity of the power law process is the simplicity of statistical inference procedures for it. Point estimators for the parameters have simple closed form expressions and confidence intervals can, in most cases, be obtained using existing tables.

Dr. Rigdon is an Assistant Professor in the Department of Mathematics and Statistics. He is an Associate Member of ASQC.

Dr. Basu is a Professor in the Department of Statistics. He is a Member of ASQC.

Models for Repairable Systems

The two most common models for the reliability of repairable systems are the *renewal process* and the *nonhomogeneous Poisson process* (NHPP). These models differ in the way that the system performs immediately after failure and repair.

If the repaired system's reliability is just like that of a brand new system, then the system is modeled by a renewal process. The phrase *good-as-new* has been used to describe a renewal process. This terminology is misleading because some systems whose initial reliability is poor may be *bad-as-new* after repair. Ascher and Feingold (1984) recommend the phrase *same-as-new* to describe a renewal process. Because a repaired system is in the same condition as a new system, a renewal process cannot be used to model a system which is deteriorating or a system which is experiencing reliability improvement.

If the system's reliability does not change after a repair, that is, a repaired system is in the same condition after the repair as just before the failure, then the appropriate model is an NHPP. The phrase *bad-as-old* has been used to describe an NHPP. Here again this terminology is misleading because a system in the development phase which is experiencing reliability growth should be *good-as-old.* The phrase *same-as-old,* recommended by Ascher and Feingold (1984), seems more appropriate. The term *minimal repair* has also been used to describe the effect of a failure and subsequent repair on a system modeled by an NHPP.

An important characteristic of the NHPP is the *failure intensity function,* or just the *intensity function,* which is denoted by $u(t)$. This is the (limit of the) probability of failure in a small time interval divided by the length of the interval; that is,

$$u(t)dt = \Pr[\text{a failure in } (t, t + dt)].$$

An axiomatic derivation of the Poisson process is given in Parzen (1962, Chapter 4). The failure intensity function depends only on the cumulative time t since the initial start-up of the system, and not on the previous pattern of failure times. Ascher and Feingold (1984) distinguish between *global time* t, which is the time since the system began operating, and the *local time* x, which is the time since the most recent failure. Thus the intensity function depends only on the global time t. If the intensity function is increasing, then the probability of failure in an interval of fixed length is increasing. This leads to failures becoming more frequent and corresponds to the situation in which the system is deteriorating. On the other hand if the intensity function is decreasing then the probability of failure in an interval of fixed length is decreasing. This leads to failures becoming less frequent and corresponds to the situation in which the system is experiencing reliability improvement.

Another approach to defining the intensity function, an approach which might provide more motivation into the nature of the intensity function, is as follows. Let $N(t)$ be the random number of failures up to time t for any stochastic point process. Ascher and Feingold (1984) define the rate of occurrence of failures (ROCOF) as

$$v(t) = \frac{d}{dt} E[N(t)].$$

Note that the intensity function could be defined for any stochastic point process, not just for the NHPP. For any orderly process, that is, a process that excludes the possibility of simultaneous independent failures, the intensity function $u(t)$ and the ROCOF $v(t)$ are equal, provided they both exist. (See Leadbetter [1970] or Thompson [1988, p. 15] for a proof.) Since the intensity function can be written as a derivative with respect to time, we can see why $u(t)$ is sometimes called the *failure rate* of the NHPP. Thompson (1981) states that this time derivative is "probably the meaning which an engineering user will attribute to the words failure rate regardless of how we define them." The problem with calling $u(t)$ the failure rate is that the term failure rate has been used to mean so many different things in the reliability literature that it has become almost meaningless. Specifically, the term failure rate has usually been used for $h(x)$, which is defined in the next section. Ascher and Feingold (1984, p. 157) have reached the conclusion that "the track record of the reliability field clearly demonstrates that 'failure rate' should not be used at all."

The power law process is just an NHPP with a particular form of intensity function. If t is global time, then the intensity function is

$$u(t) = (\beta/\theta)(t/\theta)^{\beta-1} \quad t > 0. \tag{1}$$

Thus the power law process depends on two unknown parameters, β and θ. When $\beta = 1$ the power law process becomes the *homogeneous Poisson process* (HPP) with mean time between failures equal to θ. The HPP is characterized by times between failure which are independent and identically distributed with an exponential distribution. Only in this special case is the NHPP also a renewal process. When $\beta > 1$ the intensity function is increasing; this corresponds to the situation in which the times between failures become shorter. When $\beta < 1$ the intensity function is decreasing; this corresponds to the situation in which the times between failures become longer.

Neither the renewal process (same-as-new) nor the NHPP (same-as-old) model seems perfectly realistic for most systems. In the case of deterioration, restoring an old system to a good-as-new state after repair seems overly optimistic. On the other hand a repaired system is usually in better condition after the repair than it was just before the failure. In the case of reliability growth, especially in a "Test, Analyze, and Fix" situation, the repaired system should be better after the repair than before the failure. In addition, a repaired system should be better than a new system if bugs are being removed from the system. Of the two models the NHPP is the more realistic model, a priori, in most situations. When available, however, data should be used to choose among a renewal process, an NHPP, or some other model.

A Comparison with Models for Nonrepairable Parts

Typically, to test nonrepairable parts, several parts are placed on test at the same time. Under these circumstances, it is reasonable to assume that the random lifetimes of the parts on test are independent and identically distributed. Usually, the goal is to learn about the distribution of lifetimes. Since all units are placed on test simultaneously, the shortest lifetime will be the first failure observed, the second shortest lifetime will be the second failure observed, etcetera. Thus the lifetimes are observed in order from shortest to longest. This similarity with repairable systems, for which successive failures on the single system under study are observed, is one of the causes of confusion between models for repairable and nonrepairable items.

The *hazard function,* or the *force of mortality,* for a nonrepairable part is the (limit of the) probability that a nonrepairable part fails in a small interval beginning at time x, conditioned on the event that it survives to time x, divided by the length of the interval. In symbols,

$$h(x)dx = \Pr[\text{failure in } (x, x + dx) \mid X > x].$$

Compare the definition of the hazard function with that of the intensity function given in the previous section. The definitions look nearly identical, the major difference being the conditional nature of the hazard function. In practice, they usually look even more nearly alike because both t and x are usually denoted t. Add to this the fact that both have been called the failure rate and it is easy to see why there has been so much confusion of terminology in the reliability literature.

Ascher and Feingold (1984) have suggested that a distinction be made between part and system degradation. They have proposed that parts with an increasing hazard function *wear out,* and systems with an increasing intensity function *deteriorate.* Note however that a system can improve even though its parts wear out. As a simple example, consider the system composed of a single replaceable part which fails at the (nonrandom) times 1, 3, 6, 10, ... In this case part wearout exists along with system improvement.

Examples

Throughout this article we illustrate properties of the power law process with two examples that differ in the way that the data were collected. The first example involves the failure times (read from Figure 2 in Duane [1964]) of a "complex type of aircraft generator;" these data are summarized in Table 1. For this set of data we will assume that the testing was terminated after the thirteenth failure. Data collected in this manner, in which testing stops after a predetermined number of failures, are called *failure truncated.* The second data set was presented by Martz (1975) and gives the failure times, or interruption times, of the 115 kV transmission circuit from Cunningham Generating Station, located near Hobbs, New Mexico, to Eddy County Interchange, located near Artesia, New Mexico. These data are summarized in Table 2. We assume that data collection was terminated on December 31, 1971. (We have no basis for this assumption other than the guess that this type of data is published yearly.) Data collected in this manner, with testing terminated at a predetermined time, are called time truncated. It is important to distinguish between these approaches to data collection because statistical inference procedures are different for the two situations.

Estimation of Parameters and Testing of Hypotheses

Failure Truncated Data

When the data are failure truncated the number of failures is fixed before testing begins and the time of the conclusion of the test is random. (The reverse is true for time truncated data.) If the power law process is assumed then the likelihood function for the failure times, measured in global time, is

TABLE 1. Failure Times in Hours for Aircraft Generator

i = Failure Number	t_i = Failure Time	$\ln(4596/t_i)$
1	55	4.426
2	166	3.321
3	205	3.110
4	341	2.601
5	488	2.243
6	567	2.093
7	731	1.839
8	1308	1.257
9	2050	0.807
10	2453	0.628
11	3115	0.389
12	4017	0.135
13	4596	0.000

$$f(t_1, t_2, \cdots, t_n \mid \beta, \theta) = (\beta/\theta^\beta)^n \left(\prod_{i=1}^{n} t_i\right)^{\beta-1} \exp[-(t_n/\theta)^\beta],$$

$$0 < t_1 < t_2 < \cdots < t_n < \infty$$

from which the maximum likelihood estimators (MLE's) can be found to be

$$\hat{\beta} = \frac{n}{\sum_{i=1}^{n-1} \ln(t_n/t_i)} \tag{2}$$

and

$$\hat{\theta} = t_n / n^{1/\hat{\beta}}. \tag{3}$$

A useful result for deriving confidence intervals and tests of hypotheses for β is that $2n\beta/\hat{\beta}$ has a chi-square distribution with $2(n-1)$ degrees of freedom. Thus a $100 \times (1-\alpha)\%$ confidence interval for β is

$$\left(\frac{\hat{\beta}\chi^2_{2(n-1),\alpha/2}}{2n}, \frac{\hat{\beta}\chi^2_{2(n-1),1-\alpha/2}}{2n} \right). \tag{4}$$

Here $\chi^2_{\nu,\gamma}$ is the $100 \times \gamma\%$ point of the chi-square distribution with ν degrees of freedom. It is often important to test the hypothesis $H_0 : \beta = 1$ versus $H_1 : \beta \neq 1$ because the power law process reduces to the HPP if $\beta = 1$. A size α test for testing any particular value of β can be constructed using the result that $2n\beta/\hat{\beta}$ has a chi-square distribution with $2(n-1)$ degrees of freedom. The rule is to reject $H_0 : \beta = \beta_0$ if either

$$\hat{\beta} < 2n\beta_0 / \chi^2_{2(n-1),1-\alpha/2}$$

or

$$\hat{\beta} > 2n\beta_0 / \chi^2_{2(n-1),\alpha/2}.$$

Confidence intervals and tests of hypotheses for θ can be based on $W = (\hat{\theta}/\theta)^{\hat{\beta}}$, because the distribution of W does not depend on the unknown parameters β and θ, although it does depend on n. The percentile points of the distribution of W were determined (using simulation) and tabulated by Finkelstein (1976). Lee and Lee (1978) showed that the distribution function of W can be determined analytically. They evaluated the distribution function for several values of n and α (the tail probability) and found close agreement with Finkelstein's table. Thus to find a $100 \times (1-\alpha)\%$ confidence interval for θ, the $\alpha/2$ and $1-\alpha/2$ percentile points, $w_{n,\alpha/2}$ and $w_{n,1-\alpha/2}$, must be read from Finkelstein's table. The confidence interval is then

$$(\hat{\theta}/w_{n,1-\alpha/2}^{1/\hat{\beta}}, \hat{\theta}/w_{n,\alpha/2}^{1/\hat{\beta}}). \tag{5}$$

A size α test of $H_0 : \theta = \theta_0$ versus $H_1 : \theta \neq \theta_0$ is to reject H_0 if either

$$\hat{\theta} < \theta_0 w_{n,\alpha/2}^{1/\hat{\beta}} \quad \text{or} \quad \hat{\theta} > \theta_0 w_{n,1-\alpha/2}^{1/\hat{\beta}}.$$

For the data in Table 1 on the failure times of an aircraft generator the estimates of β and θ are

$$\hat{\beta} = \frac{13}{\sum_{i=1}^{13-1} \ln(4596/t_i)} = 0.569$$

and

$$\hat{\theta} = \frac{4596}{13^{1/0.569}} = 50.7.$$

The estimate for β is smaller than one, which is to be expected since the times between failures are getting larger; this implies that the system is experiencing reliability improvement.

TABLE 2. Interruption Dates and Times Between Interruptions, In Years, for 115 kV Transmission Line

i = Failure Number	Failure Date	t_i = Failure Time	$\ln(8.463/t_i)$
0	15-Jul-63		
1	31-Aug-63	0.129	4.184
2	08-Sep-63	0.151	4.026
3	18-Apr-64	0.762	2.408
4	27-May-64	0.869	2.276
5	22-Jun-66	2.937	1.058
6	12-Aug-66	3.077	1.012
7	18-May-67	3.841	0.790
8	02-Jul-67	3.964	0.758
9	04-May-68	4.802	0.567
10	08-Jun-68	4.898	0.547
11	28-May-71	7.868	0.073
12	19-Dec-71	8.430	0.004
Trunc. Time	31-Dec-71	8.463	

A confidence interval for β can be constructed using (4). The 2.5 and 97.5 percent points of the chi-square distribution with 24 degrees of freedom are 12.4 and 39.4. This yields the 95% confidence interval (0.294, 0.934) for β. Since this interval excludes one, the test of $H_0 : \beta = 1$ would be rejected at the $\alpha = 0.05$ level. Thus there is statistically significant evidence that the system is improving in time. The p-value for this two-sided test is about 0.01.

The table for $w_{n,\gamma}$ in Finkelstein (1976) does not give the 2.5 or the 97.5 percent points; instead it gives the 2.0 and 98.0 percentile points. As a result this table cannot be used to calculate a 95% confidence interval for θ. It can, however, be used to calculate a 96% confidence interval. For failure times of the aircraft generator, $n = 13$. Interpolation in Finkelstein's table gives

$$w_{13,0.02} = 0.355 \quad \text{and} \quad w_{13,0.98} = 23.$$

Thus application of (5) yields (0.20, 312) as the confidence interval for θ. This is an extremely wide confidence interval. It is usually not necessary, however, to estimate θ with high precision since θ has no simple interpretation.

Time Truncated Data

Suppose now that system testing stops at some predetermined time t. In this case the number of failures N is random before testing begins, and the time of the conclusion of testing is fixed. If at least one failure occurs before time t then the MLE's of β and θ exist and are

$$\hat{\beta} = \frac{N}{\sum_{i=1}^{N} \ln(t/t_i)} \tag{6}$$

and

$$\hat{\theta} = t/N^{1/\hat{\beta}}. \tag{7}$$

Note the similarities and differences between (2) and (3), and (6) and (7). The formulas have similar form, however in (2) and (3) the n is fixed and the t_n is random, whereas in (6) and (7) the N is random and the t is fixed. Note also that for the time truncated case: (1) The numerator in the logarithm function of (6) is the truncation time t, not the time of the last failure T_N and (2) the sum in (6) goes up to $i = N$, not to $i = n - 1$ as in (2).

Confidence intervals and tests of hypotheses for β are based on the result that $2n\beta/\hat{\beta}$ has a chi-square distribution with $2n$ degrees of freedom. Thus a $100 \times (1 - \alpha)\%$ confidence interval for β is

$$\left(\frac{\hat{\beta}\chi^2_{2n,1-\alpha/2}}{2n}, \frac{\hat{\beta}\chi^2_{2n,\alpha/2}}{2n} \right) \tag{8}$$

and a size α test for $H_0 : \beta = \beta_0$ versus $H_1 : \beta \neq \beta_0$ is to reject if either

$$\hat{\beta} < 2n\beta_0/\chi^2_{2n,1-\alpha/2}$$

or

$$\hat{\beta} > 2n\beta_0/\chi^2_{2n,\alpha/2}.$$

Just as for failure truncated data, it is often important to test $H_0 : \beta = 1$ because the power law process reduces to the HPP when $\beta = 1$.

Obtaining confidence intervals for θ is more difficult when the data are time truncated. There appears to be no exact method of obtaining confidence intervals for θ alone, although Bain and Engelhardt (1980) suggest the following approximate method. After observing $N = n$ failures look up $w_{n+1,\alpha/2}$ and $w_{n+1,1-\alpha/2}$ in Finkelstein's (1976) table. (This is the same table used to obtain confidence intervals for θ for the failure truncated case. Note that the first subscript in w is $n + 1$, not n.) The approximate lower confidence limit for θ is then

$$\theta_L = \hat{\theta}\{n[(n+1)w_{n+1,1-\alpha/2}]^{-n/(n+1)}\}^{1/\hat{\beta}} \tag{9a}$$

and the upper confidence limit is

$$\theta_U = \hat{\theta}\{n[(n+1)w_{n+1,\alpha/2}]^{-n/(n+1)}\}^{1/\hat{\beta}}. \tag{9b}$$

In the second example, on the interruption times of the 115 kV transmission line, we assume that the data are time truncated on December 31, 1971. Since the time when data collection started was not given by Martz (1975) we use the time of the first failure, July 15, 1963, as the time zero, making a total of 8.463 years during which failure times were observed. The point estimates for β and θ are

$$\hat{\beta} = \frac{12}{\sum_{i=1}^{12} \ln(8.463/t_i)} = 0.678$$

and

$$\hat{\theta} = \frac{8.463}{12^{1/0.678}} = 0.217.$$

The 2.5 and 97.5 percent points of the chi-square distribution with 24 degrees of freedom are 12.4 and 39.4, respectively. Application of (8) yields (0.350, 1.147) as the 95% confidence interval for β. Since this interval includes one there is not strong evidence that the data do not come from a HPP. The p-value for the test of $H_0 : \beta = 1$ is between 0.1 and 0.2.

An approximate 96% confidence interval for θ, can be constructed using the approximation suggested by Bain and Engelhardt (1980). Interpolation in Finkelstein's table yields $w_{13,0.02} = 0.355$ and $w_{13,0.98} = 23$. The interval is then computed to be (0.004, 1.057).

Estimating the Intensity Function

Another important quantity is the value of the intensity function when the testing stops. For a system experiencing reliability growth during the development stage, this quantity represents the intensity, in failures per time unit, for the system if it were to be put into service with no further improvements. For this reason an estimate of this value is often used to decide when sufficient testing and improvements have been made. For a system which is deteriorating this value of the intensity function can be used to decide when to remove a system from use or when to overhaul it. Just as for the parameters β and θ, point and interval estimation procedures for this value of the intensity function are different for failure truncated and time truncated data.

Failure Truncated Data

Conditioned on the time of the n^{th} failure, that is $T_n = t_n$, the MLE of $u(t_n)$ is

$$\hat{u} = n\hat{\beta}/t_n. \tag{10}$$

Note that this does not involve the estimate of the parameter θ. Tables for constructing confidence intervals for $u(t_n)$ have been developed by Crow (1975) using simulation. Lee and Lee (1978) developed an analytic way to obtain the entries in Crow's table and showed that the simulations done by Crow were accurate. Later, Crow (1982) developed tables for obtaining confidence intervals for $M(t_n) = 1/u(t_n)$, a quantity which he called the "mean time between failures." The tables in Crow (1982) can also be used to find a confidence interval for $u(t_n)$. For confidence coefficients of 80%, 90%, 95%, and 98% he gives values ρ_1 and ρ_2 for $n = 2(1)30, 30(5)50, 60, 70, 80, 100$ such that the confidence interval for $M(t_n)$ is

$$(\rho_1\hat{M}(t_n), \rho_2\hat{M}(t_n)) \tag{11}$$

where

$$\hat{M}(t_n) = 1/\hat{u} = t_n/n\hat{\beta}.$$

The confidence interval for $u(t_n)$ is then

$$(\hat{u}/\rho_2, \hat{u}/\rho_1). \tag{12}$$

For values of n beyond 100 the normal approximation of Crow (1982) can be used to approximate ρ_1 and ρ_2. These confidence intervals are interpreted in the following manner: a confidence interval calculated in this way will include the true value of $u(T_n)$ (or $M(T_n)$) with probability $1 - \alpha$ when viewed before testing begins.

For the data on failure times of an aircraft generator the MLE of β was found to be 0.569 and the time of the last failure was 4596. Using (10) the estimate of the intensity function after 4596 hours is found to be

$$\hat{u} = \frac{13 \times 0.569}{4596} = 0.00161.$$

We can also construct a 95% confidence interval for $u(t_n)$. From Crow's (1982) table we find that for $n = 13$, $\rho_1 = 0.5488$ and $\rho_2 = 2.721$; thus the 95% confidence interval for $u(t_n)$ is

$$(0.00161/2.721, 0.00161/0.5488) = (0.00059, 0.00293).$$

Time Truncated Data

If the data are time truncated at time t then the MLE of $u(t)$ is

$$\hat{u} = N\hat{\beta}/t. \tag{13}$$

If no failures are observed before time t, then $\hat{u}$ is taken to be zero. Here again, this estimator does not depend on the estimator of θ. Crow (1982) gives tables for constructing 80%, 90%, 95%, and 98% confidence intervals for $M(t) = 1/u(t)$. For the observed value of $N = n$, the table gives Π_1 and Π_2 such that the confidence interval for $M(t)$ is

$$(\Pi_1\hat{M}(t), \Pi_2\hat{M}(t)) \tag{14}$$

where

$$\hat{M}(t) = 1/\hat{u} = t/N\hat{\beta}.$$

The confidence interval for $u(t)$ is then

$$(\hat{u}/\Pi_2, \hat{u}/\Pi_1).$$

Bain and Engelhardt (1980) and Crow (1982) give (slightly different) normal approximations which can be used to approximate Π_1 and Π_2 when $N > 100$.

For the interruption times of the 115 kV transmission line the estimate of β is 0.678 and the truncation time is 8.463 years. This yields

$$\hat{u} = \frac{12 \times 0.678}{8.463} = 0.961$$

as the estimate for the value of the intensity on December 31, 1971. From Crow's (1982) table we see that for $n = 12$, $\Pi_1 = 0.453$, and $\Pi_2 = 2.699$. Thus a 95% confidence interval for $u(t)$ is

$$(0.961/2.699, 0.961/0.453) = (0.356, 2.121).$$

The units are in failures per year.

Goodness-of-Fit Tests

By making the appropriate transformation of the failure times, exact goodness-of-fit tests can be constructed for the power law process. If the data are failure truncated the appropriate transformation is

$$U_j = \ln(T_n/T_{n-j}) \quad j = 1, 2, \cdots, n-1$$

and if the data are time truncated the appropriate transformation (conditioned on $N = n$) is

$$W_j = \ln(t/T_{n-j+1}) \quad j = 1, 2, \cdots, n.$$

The U's are distributed as the $n - 1$ order statistics of a sample of size $n - 1$ from an exponential distribution with (unknown) mean $1/\beta$, and the W's are distributed as the n order statistics from an exponential distribution with (unknown) mean $1/\beta$. Thus once the appropriate transformation is made one of the many goodness-of-fit tests for the exponential distribution can be used to assess the adequacy of the power law process. Here we illustrate only Lilliefors's (1969) test, although any of the following tests could be applied after the appropriate transformation is made (see Shapiro and Wilk [1972] and Stephens [1974, 1978]):

1. Kuiper's V Test
2. Cramér-von Mises W^2 Test
3. Watson's U^2 Test
4. Anderson-Darling A^2 Test
5. Shapiro-Wilk W Test
6. Stephens's W^* Test.

Notice that these transformations "get rid of" the nuisance parameter θ, and thus reduce the problem from one involving two unknown parameters to one involving just one unknown parameter!

Lilliefors's test of exponentiality, also described in Conover (1980), is a Kolmogorov-type test. That is, the test statistic is the largest absolute difference between the estimated cumulative distribution function and the empirical distribution function. To apply the goodness-of-fit test for the power law process first make the appropriate transformation given above. Once these order statistics from an exponential distribution are obtained, estimate the mean of the exponential distribution. (Note that this exponential distribution occurs because of the transformation made on the original failure times, and has no physical interpretation.) When the data are failure truncated this mean is estimated by

$$\hat{\mu}_F = \sum_{j=1}^{n-1} U_j/(n-1)$$

and when the data are time truncated it is estimated by

$$\hat{\mu}_T = \sum_{j=1}^{n} W_j/n.$$

Using these estimates, the estimated cumulative distribution function is

$$F_U^*(u) = 1 - \exp(-u/\hat{\mu}_F) \quad u > 0$$

or

$$F_W^*(w) = 1 - \exp(-w/\hat{\mu}_T) \quad w > 0.$$

The empirical distribution function is the step function

$$S_U(u) = \begin{cases} 0 & \text{if} \quad u < u_1 \\ j/(n-1) & \text{if} \quad u_j \le u < u_{j+1} \\ & \qquad j = 1, 2, \cdots, n-2 \\ 1 & \text{if} \quad u \ge u_{n-1} \end{cases}$$

or

$$S_W(w) = \begin{cases} 0 & \text{if} \quad w < w_1 \\ j/n & \text{if} \quad w_j \le w < w_{j+1} \\ & \qquad j = 1, 2, \cdots, n-1 \\ 1 & \text{if} \quad w \ge w_n. \end{cases}$$

The test statistic is then the largest absolute difference between $F_U^*(u)$ and $S_U(u)$ in the failure truncated case, or between $F_W^*(w)$ and $S_W(w)$ in the time truncated case. This largest absolute difference can be found by computing

$$T_1 = \max_{1 \le j \le n-1} |F_U^*(u_j) - S_U(u_j)|$$

and

$$T_2 = \max_{1 \le j \le n-1} |F_U^*(u_j) - S_U(u_{j-1})|$$

where $S_U(u_0)$ is defined to be zero. (See Daniel [1978, pp. 272–273].) For the time truncated case T_1 and T_2 are calculated as above with W in place of U and n in place of $n - 1$. Once these statistics are calculated, the text statistic is $T = \max(T_1, T_2)$. The value of T is then compared to the critical value found in the table given by Lilliefors (1969) or, preferably, in the table given by Conover (1980). Conover's table is more complete and more accurate than Lilliefors's. Originally, Lilliefors used Monte Carlo simulation to obtain the entries in the table. Later, Durbin (1975) developed an analytic method for obtaining the entries. Conover's table is adapted from the table in Durbin's paper.

Table 3 shows the U's for the failure times of the aircraft generator. Recall that U_j is defined as a function of T_{n-j} and T_n; as a result the U's are listed from U_{12} at the top of Table 3, to U_1 at the bottom. Also displayed in Table 3 are the estimated distribution function $F_U^*(u)$ and the empirical distribution function $S_U(u)$. The last two columns show the absolute differences

$$|F_U^*(u_j) - S_U(u_j)|$$

and

$$|F_U^*(u_j) - S_U(u_{j-1})|.$$

The statistics T_1 and T_2, the largest numbers in the last two columns, are $T_1 = 0.119$ and $T_2 = 0.202$. The test statistic is then the larger of these numbers; that is, $T = 0.202$. To find out whether this is statistically significant, refer to Conover's table using $n - 1$ in place of n. This is required since we have a sample of size $n - 1$ from an exponential distribution. Since $n = 13$ failures occurred, use $n = 12$ in the table to find that the critical value for a test with $\alpha = 0.05$ is 0.2981. Since 0.202 is less than 0.2981 there is no evidence of a departure from the power law process. In fact, the p-value for this test is greater than 0.30.

Next consider the time truncated data set on the interruption times of the 115 kV transmission line. The W's, along with $F_W^*(w)$, $S_W(w)$

$$|F_W^*(w_j) - S_W(w_j)|$$

and

$$|F_W^*(w_j) - S_W(w_{j-1})|$$

are shown in Table 4. In this example $T_1 = 0.155$ and $T_2 = 0.143$, so the test statistic is $T = 0.155$. To assess the statistical significance of this result, look up in Conover's table the critical value for a size $\alpha = 0.05$ test. This time, since the data are time truncated, use n equal to the number of failures. For this example $n = 12$ so the critical value is 0.2981. Since T is less than this value, there is no evidence of a departure from the power law process. Here again, the p-value is greater than 0.30.

Crow (1974) has developed a goodness-of-fit test from another approach. He suggests the transformations

$$Z_i = T_i/T_n \quad i = 1, 2, \cdots, n-1$$

for the failure truncated case and

$$Z_i = T_i/t \quad i = 1, 2, \cdots, N$$

for the time truncated case. Using the unbiased estimator of β

$$\bar{\beta} = \frac{n-2}{n}\hat{\beta}$$

when the data are failure truncated and

$$\bar{\beta} = \frac{n-1}{n}\hat{\beta}$$

as an estimator which is unbiased conditioned on $N = n$, when the data are time truncated, he then computes the modified Cramér-von Mises statistic

$$C^2 = \frac{1}{12M} + \sum_{j=1}^{M}\left(Z_j^{\bar{\beta}} - \frac{2j-1}{2M}\right)^2$$

TABLE 3. Calculations Involved in Computing Lilliefors's Test Statistic for Data on Aircraft Generator

Failure No. i	Failure Time T_i	$U_j = \ln\left(\frac{T_n}{T_i}\right)$	j	$F_U^*(u_j)$	$S_U(u_j)$	$\lvert F_U^*(u_j) - S_U(u_j)\rvert$	$\lvert F_U^*(u_j) - S_U(u_{j-1})\rvert$
1	55	4.426	12	0.902	1.000	0.098	0.015
2	166	3.321	11	0.825	0.917	0.092	0.008
3	205	3.110	10	0.805	0.833	0.028	0.055
4	341	2.601	9	0.745	0.750	0.005	0.078
5	488	2.243	8	0.692	0.667	0.025	0.109
6	567	2.093	7	0.667	0.583	0.084	0.167
7	731	1.839	6	0.619	0.500	0.119	0.202
8	1308	1.257	5	0.483	0.417	0.066	0.150
9	2050	0.807	4	0.345	0.333	0.012	0.095
10	2453	0.628	3	0.281	0.250	0.031	0.114
11	3115	0.389	2	0.185	0.167	0.018	0.102
12	4017	0.135	1	0.068	0.083	0.015	0.068
13	4596						

TABLE 4. Calculations Involved in Computing Lilliefors's Test Statistic for the Data on Interruption Times of 115 kV Transmission Line

Failure No. i	Failure Time T_i	$W_j = \ln\left(\frac{t}{T_i}\right)$	j	$F_W^*(w_j)$	$S_W(w_j)$	$\lvert F_W^*(w_j) - S_W(w_j)\rvert$	$\lvert F_W^*(w_j) - S_W(w_{j-1})\rvert$
1	0.129	4.184	12	0.941	1.000	0.059	0.024
2	0.151	4.026	11	0.935	0.917	0.018	0.102
3	0.762	2.408	10	0.805	0.833	0.028	0.055
4	0.869	2.276	9	0.786	0.750	0.036	0.119
5	2.937	1.058	8	0.512	0.667	0.155	0.071
6	3.077	1.012	7	0.496	0.583	0.087	0.004
7	3.841	0.790	6	0.415	0.500	0.085	0.002
8	3.964	0.758	5	0.402	0.417	0.015	0.069
9	4.802	0.567	4	0.319	0.333	0.014	0.069
10	4.898	0.547	3	0.310	0.250	0.060	0.143
11	7.868	0.073	2	0.048	0.167	0.119	0.035
12	8.430	0.004	1	0.003	0.083	0.080	0.003

where $M = n - 1$ if the data are failure truncated, and $M = n$ if the data are time truncated. Using the techniques of Darling (1955) he obtains the result that the distribution of the test statistic C^2 is independent of any unknown parameter. He gives a table of critical values for C^2, calculated from simulation, for $M = 2$ to 60.

Smith and Oren [1980] give yet another approach for testing goodness-of-fit for the power law process. They divide the time axis into several, say k, intervals and compare the observed number of failures O_i with the expected number of failures E_i for each interval $i = 1, 2, \cdots, k$. To assess the statistical significance they compute the chi-square statistic

$$\chi^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}.$$

This test requires that a system be observed for a large number of failures because the expected number of failures in each interval should be at least five.

Summary

In this article we have presented some of the elementary properties of the power law process. Other properties of the power law process have been presented by Engelhardt and Bain (1978) (prediction intervals for times of future failures), Lee (1980) (comparing several systems whose failure processes are power law processes), Higgins and Tsokos (1981) (a "quasi-Bayes" estimator for the failure intensity), Bain and Engelhardt (1982) (a sequential test for β), Engelhardt and Bain (1987) and Rigdon (1985) (allowing θ to vary among different copies of a system), Rigdon and Basu (1988) (comparing several estimators of the failure intensity function) and Rigdon and Basu (1990) (the effect of assuming a homogeneous Poisson process when the true process is a power law process).

The development of the power law process has been motivated by the need to model the failure times of repairable systems, however it is possible that the power law process could also be used to model the times of "lost work day cases" in a factory, the occurrences of earthquakes of a given magnitude, or other events that occur at random points in time.

One final note on using the power law process, or any other model for event times. Clearly indicate the time that data collection started and the time that it ceased. This is necessary so that the appropriate analysis, that is, an analysis based on failure truncated or time truncated data, can be applied and maximum information can be obtained from the data. For time truncated data, the time between the last failure and the termination of the test contains some information that shouldn't be wasted.

Acknowledgments

We would like to thank the two referees for their helpful comments. Their comments on some confusing notation in an earlier draft of this paper have certainly led to a better presentation of the material. This research has been supported in part by the Air Force Office of Scientific Research, Air Force Systems Command USAF, under grant number AFOSR-87-0139. The U.S. government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright notation thereon.

References

ASCHER, H. (1981). "Weibull Distribution Vs. Weibull Process." *Annual Reliability and Maintainability Symposium.* IEEE, Chicago, IL, pp. 426–431.

ASCHER, H. and FEINGOLD, H. (1984). *Repairable Systems Reliability.* Marcel Dekker, New York, NY.

BAIN, L. J. (1978). *Statistical Analysis of Reliability and Life Testing Models.* Marcel Dekker, New York, NY.

BAIN, L. J. and ENGELHARDT, M. (1980). "Inferences on the Parameters and Current System Reliability for a Time Truncated Weibull Process." *Technometrics* 22, pp. 421–426.

BAIN, L. J. and ENGELHARDT, M. (1982). "Sequential Probability Ratio Tests for the Shape Parameter of a Nonhomogeneous Poisson Process." *IEEE Transactions on Reliability* R-31, pp. 79–83.

BAIN, L. J. and ENGELHARDT, M. (1986). "On the Asymptotic Behavior of the Mean Time Between Failures for Repairable Systems." *Reliability and Quality Control* (Ed. A. P. Basu), Elsevier, Amsterdam, North-Holland, pp. 1–7.

BAIN, L. J.; ENGELHARDT, M.; and WRIGHT, F. T. (1985). "Tests for an Increasing Trend in the Intensity Function of a Poisson Process: A Power Study." *Journal of the American Statistical Association* 80, pp. 419–422.

BASSIN, W. M. (1973). "A Bayesian Optimal Overhaul Model for the Weibull Restoration Process." *Journal of the American Statistical Association* 68, pp. 575–578.

CONOVER, W. J. (1980). *Practical Nonparametric Statistics.* 2nd Edition. John Wiley & Sons, New York, NY.

CROW, L. H. (1974). "Reliability Analysis for Complex Systems." *Reliability and Biometry,* (Eds. F. Proschan and R. J. Serfling), SIAM, Philadelphia, pp. 379–410.

CROW, L. H. (1975). "Tracking Reliability Growth." *Proceedings of the Twentieth Conference on Design of Experiments* ARO Report 75-2, pp. 741–754.

CROW, L. H. (1982). "Confidence Interval Procedures for the Weibull Process with Applications to Reliability Growth." *Technometrics* 24, pp. 67–72.

DANIEL, W. W. (1978). *Applied Nonparametric Statistics.* Houghton-Mifflin, Boston, MA.

DARLING, D. A. (1955). "The Cramér-Smirnov Test in the Parametric Case." *Annals of Mathematical Statistics* 9, pp. 1–20.

Duane, J. T. (1964). "Learning Curve Approach to Reliability." *IEEE Transactions on Aerospace* 2, pp. 563–566.

DURBIN, J. (1975). "Kolmogorov-Smirnov Tests When Parameters Are Estimated with Applications to Tests of Exponentiality and Tests on Spacings." *Biometrika* 62, pp. 5–22.

ENGELHARDT, M. and BAIN, L. J. (1978). "Prediction Intervals for the Weibull Process." *Technometrics* 20, pp. 167–169.

ENGELHARDT, M. and BAIN, L. J. (1987). "Statistical Analysis of a Compound Power Law Model for Repairable Systems." *IEEE Transactions on Reliability* R-36, pp. 392–396.

FINKELSTEIN, J. M. (1976). "Confidence Bounds on the Parameters of the Weibull Process." *Technometrics* 18, pp. 115–117.

HIGGINS, J. J. and TSOKOS, C. P. (1981). "A Quasi-Bayes Estimate of the Failure Intensity of a Reliability Growth Model." *IEEE Transactions on Reliability* R-30, pp. 471–475.

LEADBETTER, M. R. (1970). "On Basic Results of Point Process Theory." *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, Vol. III,* (Eds. L. Le Cam, J. Neyman, and E. Scott), University of California Press, Berkeley, CA, pp. 449–462.

LEE, L. (1980). "Comparing Rates of Several Independent Weibull Processes." *Technometrics* 22, pp. 427–430.

LEE, L. and LEE, S. K. (1978). "Some Results on Inference for the Weibull Process." *Technometrics* 20, pp. 41–45.

LILLIEFORS, H. W. (1969). "On the Kolmogorov-Smirnov Test for the Exponential Distribution with Mean Unknown." *Journal of the American Statistical Association* 64, pp. 387–390.

MARTZ, H. F. (1975). "Pooling Life Test Data by Means of the Empirical Bayes Method." *IEEE Transactions on Reliability* R-24, pp. 27–30.

MOLLER, S. K. (1976). "The Rasch-Weibull Process." *Scandinavian Journal of Statistics* 3, pp. 107–115.

PARZEN, E. (1962). *Stochastic Processes.* Holden-Day, San Francisco, CA.

RIGDON, S. E. (1985). "On the Reliability of Repairable Systems." Unpublished Ph.D. Dissertation, University of Missouri, Columbia, MO.

RIGDON, S. E. and BASU, A. P. (1988). "Estimating the Intensity Function of the Weibull Poisson Process at the Current Time: Failure Truncated Case." *Journal of Statistical Computation and Simulation* 30, pp. 17–38.

RIGDON, S. E. and BASU, A. P. (1990). "The Effect of Assuming a Homogeneous Poisson Process When the True Process is a Power Law Process." *Journal of Quality Technology,* to appear.

SHAPIRO, S. S. and WILK, M. B. (1972). "An Analysis of Variance Test for the Exponential Distribution (Complete Samples)." *Technometrics* 14, pp. 355–370.

SMITH, S. A. and OREN, S. S. (1980). "Reliability Growth of Repairable Systems." *Naval Research Logistics Quarterly* 27, pp. 539–547.

STEPHENS, M. A. (1974). "EDF Statistics for Goodness of Fit and Some Comparisons." *Journal of the American Statistical Association* 69, pp. 730–737.

STEPHENS, M. A. (1978). "On the *W* Test for Exponentiality with Origin Unknown." *Technometrics* 20, pp. 33–35.

THOMPSON, W. A. (1981). "On the Foundations of Reliability." *Technometrics* 23, pp. 1–13.

THOMPSON, W. A. (1988). *Point Process Models with Applications to Safety and Reliability.* Chapman and Hall, New York, NY.

Key Words: *Intensity Function, Nonhomogeneous Poisson Process, Repairable Systems, Weibull Process.*