Using Accelerated Life Tests Results to Predict Product Field Reliability

William Q. Meeker
Dept. of Statistics
Center for Nondestructive Evaluation
Iowa State University
Ames, IA 50011

Luis A. Escobar
Dept. of Experimental Statistics
Louisiana State University
Baton Rouge, LA 70803

Yili Hong
Dept. of Statistics
Iowa State University
Ames, IA 50011

June 22, 2008

Abstract

Accelerated life tests (ALTs) provide timely assessments of the reliability of materials, components, and subsystems. ALTs can be run at any of these levels or at the full-system level. Sometimes ALTs generate multiple failure modes. A frequently asked question, coming near to the end of an ALT program, is "What do these test results say about field performance?" ALTs are carefully controlled whereas the field environment is highly variable. Products in the field see, for example, different average use rates across the product population. With good characterization of field use conditions, it may be possible to use ALT results to predict the failure time distribution in the field. When such information is not available but both life test data and field data (e.g., from warranty returns) are available, it may be possible to find a model to relate the two data sets. Under a reasonable set of practical assumptions, this model can then be used to predict the failure time distribution for a future component or product operating in the same use environment. This paper describes a model and methods for such situations. The methods will be illustrated by an example to predict the failure time distribution of a newly designed product with two failure modes.

Key words: Bivariate Lognormal Distribution, Censored Data, Competing Risks, Fatigue, Maximum Likelihood, Multiple Failure Modes, Reliability, Warranty.

1 INTRODUCTION

1.1 Background

Often the manufacturer of a product will experience higher than expected numbers of warranty returns. The root cause is often traced to an untested weakness in the product design or unexpected harmful environmental conditions. Generally, an engineering change to correct the problem is easy to implement. Although the particulars of the situation can vary greatly, invariably, accelerated life tests (ALTs) to assess the effect of the change is accompanied by a question from management that sounds something like "After these changes are made, what proportion of our future product will be returned for warranty service?"

1.2 Relationship Between ALT Results and Field Reliability

Finding a relationship between ALT results and field returns (usually from warranty data) involves some challenging technical tasks. In general, using ALT results to predict product field returns requires

- A model to describe the effect that acceleration and other environmental variables (e.g., increased cycling rate, stress, or temperature) will have on the lifetime distribution of the life-limiting component(s) of the product.
- An understanding of environmental and operating conditions in the product's use environment.

The most challenging task in the application of accelerated testing is to find an appropriate model to relate the accelerating variable(s) to product life. Commonly used regression relationships used in many other statistical applications may not be appropriate. Nelson (1990) discusses many useful ALT models as well as statistical methods for ALTs. Chapters 18-21 of Meeker and Escobar (1998) provide additional material on ALTs that complements Nelson (1990). Escobar and Meeker (2006) provide a detailed discussion of many ALT models.

The ability to quantify environmental and operating conditions varies greatly across different kinds of products. For example, stress and temperature patterns experienced by aircraft jet engines in commercial airliners are well understood and reasonably consistent across a product population. On the other hand, similar patterns in military aircraft are less controlled and therefore more highly variable. The operating environment and use rate of computer printers and copying machines (measured in number of pages printed per hour) varies tremendously from one unit in the product population to another. The amount of use that a home appliance (e.g., a dish washer, washing machine, or vacuum cleaner) experiences also varies from one household to another. The number of miles driven by different automobiles in a week across a product population may vary by more than two orders of magnitude. Some automobiles are regularly driven at high speeds while others may never experience such speeds.

The following steps should be used to establish a model relating ALT results and field performance.

1. Attempt to develop ALTs that generate the same failure modes, driven by the same failure mechanism, when compared with the actual use environment.

2. Carefully compare failures from ALTs with field failures to make sure that the failure modes are the same. For mechanical, electrical, and micro electronic failure modes, this is generally done by "autopsy" and physical failure mode analysis. For studies involving chemical degradation, analytical chemical measurements may be used to check that the same chemical reactions are seen in the ALTs as are causing failure in the use environment.

3. If discrepancies between ALT results and failures returned from the field are discovered, it is important to determine the cause of the differences and modify testing procedures so that test results (e.g., failure modes) agree.

4. Use physical/chemical knowledge about the failure mechanisms, information about the use environment, and available data to find a model to relate ALT results to field failures.

1.3 Related Literature

Nelson (2001) describes an application and model for using ALT data to predict field reliability of a population of units that are subjected to differing dynamic stresses. Yang and Zaghati (2006) describe models for use-rate ALTs that can be used when the cycles-to-failure distribution depends on the use rate. Mannaa, Pala, and Sinhab (2007) describe and illustrate applications for a use-rate model to predict returns from a two-dimensional warranty, like those offered by US automobile manufacturers. Wu and Hamada (2000) and Condra (2001) describe the use of designed experiments to improve product reliability. Some of the work in this paper, concerning multiple failure modes, uses results from the theory of competing risks. Useful books in this area include David and Moeschberger (1978) and Crowder (2001).

1.4 Overview

The rest of this paper is organized as follows. Section 2 describes a simple use-rate model and shows how it can be used with ALT data to predict the field performance of a component in an appliance. Section 3 presents a second example, from a different appliance, Appliance B, which had two causes of failure. Section 4 shows how to extend the use-rate model to products with multiple failure modes. Section 5 gives methods for maximum likelihood (ML) estimation for the multiple failure mode use-rate model. Section 6 gives a summary of the ML estimation results for Appliance B. Section 7 shows how the multiple failure modes use-rate model was used to make reliability predictions for proposed design changes to a turbine device in Appliance B. Section 8 investigates the importance of properly quantifying the dependence between failure modes. Section 9 contains concluding remarks and describes areas for further research.

The examples used in this paper are real. In order to protect proprietary or sensitive information, however, it was necessary for us to change names and modify the time scale of the data.

2 A SIMPLE USE-RATE MODEL FOR PREDICTING FIELD RELIABILITY

This section uses a simple use-rate model to predict the life of a component in the field. Section 4 presents a generalization of this model for more complicated prediction problems with two failure modes.

2.1 Time Scales in ALTs

ALTs, for most failure modes, record time in terms of a variable that is proportional to the amount of product "use." Use time should generally defined by the physics of the underlying failure mechanism. For example

- Hours of operation of a motor bearing.
- Number of takeoff-landing cycles for a jet engine or an airframe.
- Cycles of use for a washing machine.
- Number of power-on cycles for an electronic component failing from a thermal-fatigue process.
- Number of operations of a toaster.
- UV dosage, which is proportional to the number of photons hitting the surface, for a coating subject to photodegradation.
- Number of miles driven for an automobile engine.
- Number of *sets* of pages printed by a printer or a copier. A set might be defined as the average number of pages printed per day across the product's customer population and thus might differ from one product to another within a company.

Note that while the definition of *use* should depend on the failure mechanism, the time *scale* definition (e.g., hours versus days) is somewhat arbitrary and is generally set to make it easy to interpret the ALT results.

In this paper we do not consider, directly, the more complicated situation where a single failure mode could have more than one relevant use variable (e.g., the number of hours that a light bulb is on and the number of times the light bulb is turned on and off). Models for such situations are described, for example, by Kordonsky, and Gertsbakh (1993), Oakes (1995) and Duchesne and Lawless (2000).

2.2 A Problem with Component A

A home appliance was experiencing a higher-than-expected warranty return rate. The cause of the vast majority of the returns was discovered to be a design defect in a particular component that we will call Component A. A design change, expected to lengthen the life of Component A, was made and an ALT was conducted to estimate the lifetime of units manufactured under the new design. The engineers responsible for the reliability of this product provided the following information.

- The failure mode seen in the field could be accurately reproduced in the ALT bench test running at increased cycling rate and matching closely the operating parameters of the component when it is installed in the appliance.

- The increased cycling rate was chosen to be high enough to give test results in a reasonable amount of time, but not so high that it would change the cycles-to-failure distribution of Component A. In particular, Component A was given enough time to "cool down" between each test cycle.

- The cycles-to-failure distribution could be estimated with good precision from a relatively small number of test units because, under the carefully controlled laboratory conditions, there was relatively little variability in the number of cycles needed to make Component A fail (when compared with variability in the weeks-to-failure data from the field).

- The average number of uses per week of the appliance varied from household to household according to a distribution that was estimated from survey data that had been obtained from the company's marketing department. Figure 1 is the histogram of the estimated cycles-per-week distribution. The distribution is a discretized lognormal distribution, truncated at 20 uses; the number of households with more than 20 uses per week was negligible.

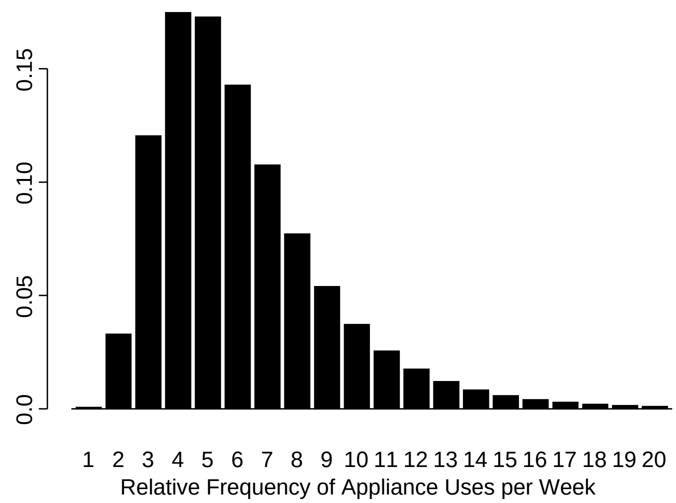


Figure 1: Component A Use-Rate Distribution (discretized lognormal distribution).

2.3 A Use-Rate Model for Field Reliability

Suppose that the cycles-to-failure distribution of the component in an ALT has a log-location-scale distribution

$$F_C(c) = \Pr(C \leq c) = \Phi\left[\frac{\log(c/\eta_C)}{\sigma_C}\right] = \Phi\left[\frac{\log(c) - \log(\eta_C)}{\sigma_C}\right]$$

where Φ is a completely specified cumulative distribution function. Important special cases of the log-location-scale family include the frequently used Weibull and lognormal distributions (as described, for example, in Chapter 4 of Meeker and Escobar 1998). Here η_C is a scale parameter having units of test cycles (or other units of testing time, as described in Section 2.1) used in the ALT. The parameter η_C is the median for the lognormal distribution and the approximate 0.63 quantile of the Weibull distribution and σ_C is a spread or shape parameter. For the Weibull distribution, the reciprocal of σ_C is the usual shape parameter.

Suppose also that the distribution of average use rates R in households can be described by a multinomial distribution with possible values $\mathbf{R} = (R_1, R_2, \ldots, R_k)$ cycles per unit time (e.g., uses per week), and corresponding probabilities $\boldsymbol{\pi} = (\pi_1, \pi_2, \ldots, \pi_k)$ where $\sum_{i=1}^{k} \pi_i = 1$ (similar to Figure 1). If a test cycle in an ALT has the same effect on lifetime as a service-use cycle, the fraction failing at real time t in the product population (e.g. weeks in service) can be described by the finite mixture

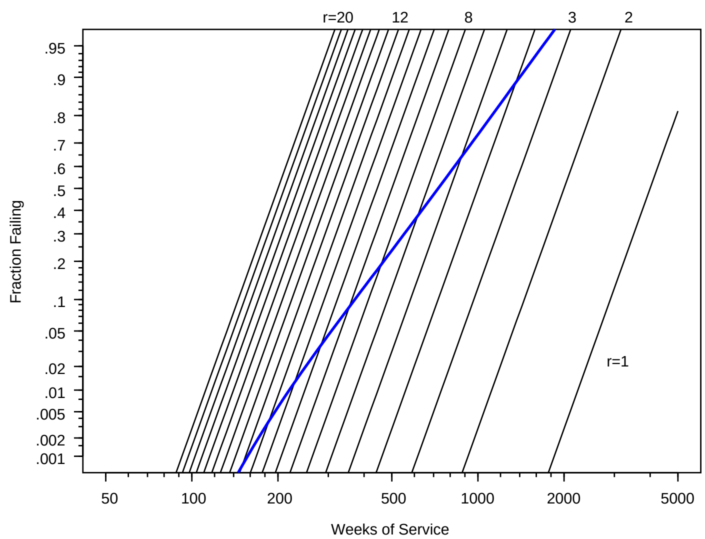


Figure 2: Fraction Failing as a Function of Weeks in Service $F_T(t; \eta_C, \sigma_C, \boldsymbol{\pi}, \mathbf{R})$ for Appliance A as a Weighted Average of Conditional Constant-Rate Field Distributions (a discrete lognormal mixture of lognormal distributions).

distribution

$$F(t; \eta_C, \sigma_C, \boldsymbol{\pi}, \mathbf{R}) = \Pr(T \leq t) = \sum_{i=1}^{k} \pi_i \, \Phi \left[\frac{\log (R_i \times t / \eta_C)}{\sigma_C} \right]. \tag{1}$$

Figure 2 provides an illustration of such a mixture distribution. Each of the parallel lines represents the ML estimate, based on the Component A ALT results, of the lognormal cdf, conditional on a given value of cycles per week ($r = 1, 2, \ldots, 20$). The curve going through the lines corresponds to the ML estimate of the mixture distribution in (1), giving the unconditional cdf as a function of weeks in service across the population, based on the use-rate distribution shown in Figure 1. Note that the curve has a smaller slope, showing that the weeks-to-failure distribution for the field product population will have larger relative variability.

It is also possible to use a continuous mixture distribution to describe such a use-rate model, but the discrete mixture is especially useful for explaining the concept to engineers and others who have not been exposed to continuous mixture distributions. One useful characteristic of the continuous mixture distributions is that a lognormal mixture of lognormal distributions is again a lognormal distribution. From Figure 2 we see that because the curve corresponding to (1) is nearly a straight line, a discretized lognormal mixture of lognormal distributions is approximately lognormal.

Although we did not need to do it for this application, one could use the use-rate model presented here, combined with methods such as those described in Escobar and

Table 1: Summary of the Appliance B Warranty Data and ALT.

Data Source	Failure Mode	Number	Five-Point Summary (Days/Cycles)				
			min	0.25	0.50	0.75	max
Warranty	Wear	93	53	205	278	389	641
	Crack	20	76	224	286	439	588
	Censored	4615	6	169	320	449	728
Constant-run ALT	Wear	8	98	157	325	542	686
	Censored	2	796	796	796	796	796
Over-load ALT	Crack	20	218	323	459	482	483

Meeker (1999), to generate prediction intervals for the cumulative number of future failures for an existing product population and a particular warranty policy.

3 APPLIANCE B RELIABILITY PROBLEMS

3.1 Appliance B Background

The model used in the rest of this paper was initially motivated by a product reliability problem in what we will call Appliance-B. Appliance-B had a turbine-device component that was failing at a higher-than-expected rate. Warranty-return projections for Appliance B had been based on cycles-to-failure experimental results from an industry-standard ALT and some subjective average use-rate information that had been provided by the company's marketing department. The industry-standard test was known to produce a *crack* failure mode.

During the investigation to learn why the number of units being returned for warranty service was higher than expected, it was discovered that fewer than 20% of the units returned for warranty repair had failed from the *crack* failure mode. The others had failed from a *wear* failure mode. Table 1 provides a summary of the data. The units of time is days in service for the warranty data and use cycles for the ALT data.

Actually twelve different failure modes had been observed from the field returns, but these were usefully collapsed into *crack* and *wear* categories because the failure modes within these two categories were similar (e.g. some of the different modes were characterized by the precise location of the crack). Such collapsing of failure mode definitions is a common practice, used to arrive at a definition of failure mode groups that allows failure modes to be treated as if they are independent (e.g., page 34 of David and Moeschberger 1978). Also the engineers studying the failures believed that there were two different mechanisms involved for the two different failure modes.

In particular, the industry-standard ALT that generated the crack mode ran the product continuously with a substantial amount of additional load on the turbine-device component.

Subsequently, the company conducted another ALT in which the turbine device was run continuously using the amount of load usually seen in actual application. This ALT always generated wear failures. From this knowledge and other physical modeling efforts the engineers then believed that, for the most part, the wear failure mode would be generated from normal product *use* but that crack failure mode was being caused, primarily, by product *abuse*. For practical reasons, however, because of the way that Appliance B is typically used, it had been designed to be robust to a reasonable amount of abuse.

As an analogy, consider a washing machine operating in spin mode. Under normal operation of a properly-designed bearing, the bearing will suffer wear during normal operation. Eventually, after a very long time, the amount of wear will accumulate and the bearing will fail. On the other hand if the load in the washer is unbalanced (on one side or the other of the drum) when operating in spin mode, the bearing will experience much higher than usual stresses. Owners who often operate their washer in this mode of operation cannot expect their product to last as long as someone who is careful not to allow the machine to stay in the unbalanced condition very long.

3.2 Appliance B Initial Field Data Analysis

Appliance B had a two-year warranty. Units in the product population had been inserted into the field over time (staggered entry). At the data-freeze date, 448 units had been in service more than 600 days and 57 units had been in service more than 700 days. The warranty data beyond two years (730 days) were deemed unreliable because customers had other options for having Appliance B repaired and the company had no information about units that were repaired through alternative channels (other than the presumption that they had survived two years without failure). Thus the available data consist of failure times for those units that had failed in their first two years of service and time in operation (with a maximum of two years) for those units that had not failed before the data-freeze date. A small number of units had been returned more than once, but we analyzed only the times to first failure.

Figure 3 is a lognormal probability plot showing the nonparametric estimates of the marginal distributions of the crack and wear failure-time distributions. Figure 3 also shows the corresponding ML estimates of the marginal parametric distributions computed under the assumption that the crack and wear failure modes in the field data are independent. The Appliance B engineers were confident that the crack and wear failure modes could be adequately modeled as being independent in *laboratory* ALTs. Another way to state this is that the failure-modes are independent, conditional on a fixed use rate. However, as explained in detail in Section 4, the variability in use

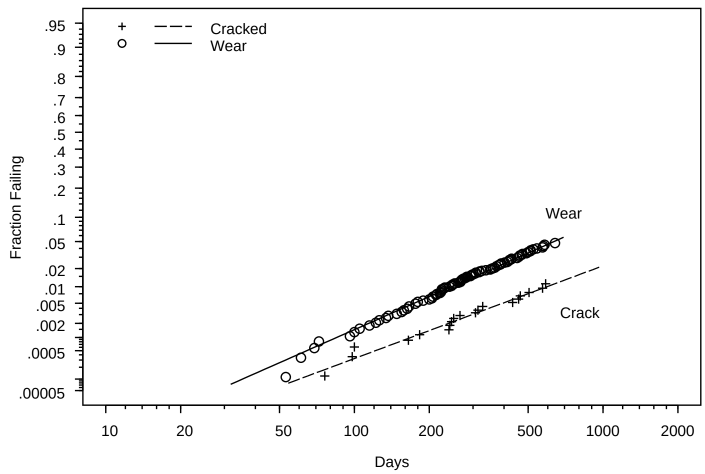


Figure 3: Lognormal Probability Plot of the Appliance B Independent Failure Mode Analysis of Wear and Crack Field Failure Modes.

rates in the field will induce some correlation between the crack and wear failure-time random variables. The amount of correlation depends on the variability in the use-rate distributions.

3.3 Appliance B Comparison of Field and ALT Data

Figure 4 is a lognormal probability plot, providing a comparison of the wear failure mode warranty-return data and the constant-run ALT that generates the wear failure mode. Figure 5 is a similar lognormal probability plot, providing a comparison of crack failure mode warranty-return data and the industry-standard ALT (constant-run with over loading) that generates the crack failure mode.

As indicated in Section 2.1, once a measure of use is chosen for a particular failure mode, the exact definition of the time scale for an ALT is somewhat arbitrary, as long as the variable is proportional to the defined measure of use. The definition of a “use cycle” for the Appliance B ALT data presented here and shown in Figures 4 and 5 was chosen in such a way that one cycle corresponds to a certain amount of use (or abuse) in one day and, making it easier to compare the slopes of the distributions from the two data sources on these probability plots.

As expected, the relative variability (recall steeper slopes in probability plots correspond to less variability) is smaller in the ALT than it is in the field returns. Note, however, that the difference in slopes is much larger for the crack failure mode in Figure 5, when compared to the wear failure mode in Figure 4. This is evidence

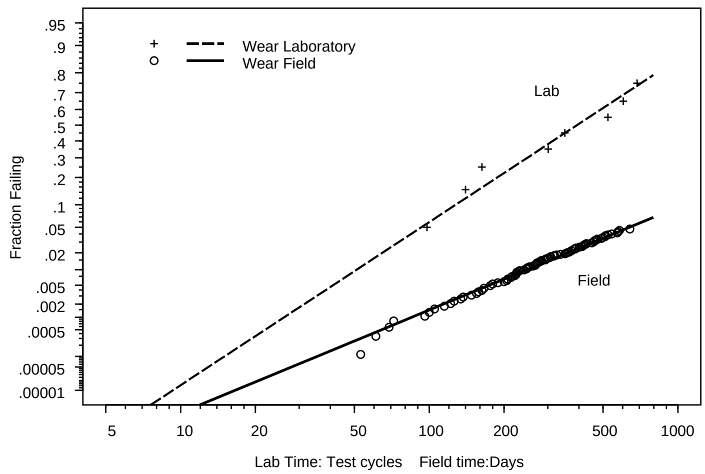


Figure 4: Lognormal Probability Plot Comparing the Wear ALT (unloaded continuous test cycling) and Wear Field Failure Kaplan-Meier Estimates for Appliance B. The fitted lines correspond to the Independent ($\rho = 0$) Bivariate Lognormal/Lognormal Use-Rate Model Described in Section 6.

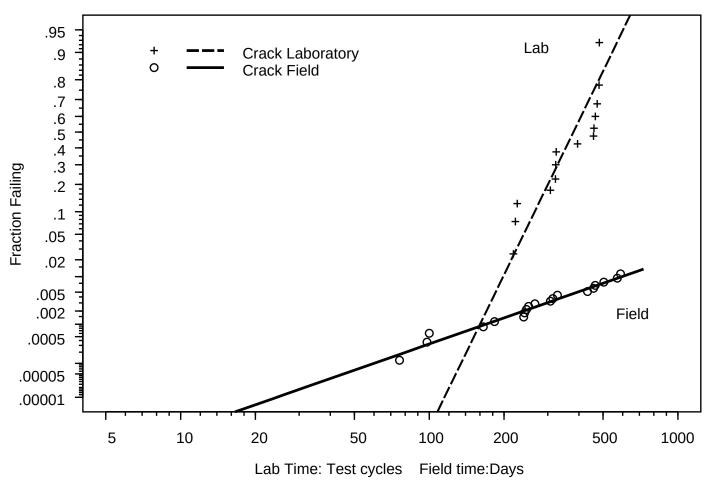


Figure 5: Lognormal Probability Plot Comparing the Wear ALT (loaded continuous test cycling) and Wear Field Failure Kaplan-Meier Estimates for Appliance B. The fitted lines correspond to the Independent ($\rho = 0$) Bivariate Lognormal/Lognormal Use-Rate Model Described in Section 6.

that there is much more variability, across the product population, in the amount that Appliance B is abused, when compared with the amount that it is used. The next section shows how to use the data in Figures 4 and 5 to estimate the parameters of the use-rate distribution.

4 A MULTIPLE FAILURE MODE USE-RATE MODEL

This section shows how to extend the use-rate model described in Section 2 to products, such as the turbine device in Appliance B, that have more than one failure mode. In order to keep the presentation reasonably simple, the development in this paper is for two failure modes. The extension to more than two failure modes would, however, be straightforward. We also use log-location-scale distributions here, although it would be possible to use any life distributions that has a scale parameter.

We represent the multiple failure modes of the turbine device in Appliance-B as a series-system model with two components. The system fails when the first component fails, and we are not able to see the failure time of the component that does not fail. For a more generic and concise presentation in the rest of this section, we index the components by number, where component 1 corresponds to a the wear failure mode and component 2 corresponds to the crack failure mode.

4.1 Cycles-to-Failure Distributions and Conditional Time-in-Service Distributions

Suppose that the lifetimes $C_j, j = 1, 2$, of the two components are independent in the *cycles* time scale with log-location-scale distributions

$$F_{C_j}(c) = \Phi\left[\frac{\log(c) - \log(\eta_{C_j})}{\sigma_{C_j}}\right]. \tag{2}$$

In the field, different components in the product may be subjected to different types of forces with different use rates. In the Appliance B example there is a mixture of regular use and over-loaded use (or abuse) that cause the two different failure modes. Thus there are different use-rate distributions for the two failure modes.

We denote the lifetime of component j in the real-time scale (days in service for Appliance B) by $T_j = C_j/R_j$, $j = 1, 2$, where, R_j is the use rate for component j. Using a simple scale change from (2), for *fixed* use rates $(R_1 = r_1, R_2 = r_2)$ for the two components, (T_1, T_2) are independent and have conditional distributions

$$F_j(t|R_j = r_j) = \Phi\left[\frac{\log(r_j \times t) - \log(\eta_{C_j})}{\sigma_{C_j}}\right], \quad j = 1, 2.$$

Here $r_j \times t$ is time in cycles corresponding to the fixed use-rate r_j.

4.2 Joint Use-Rate Distributions and Time-in-Service Distributions

Suppose that the use rates for the two different classes of use (regular use and abuse due to heavy load), can be quantified by a joint cumulative distribution $G(r_1, r_2)$. Let $g(r_1, r_2)$ denote the corresponding joint density. The joint field-failure cdf in the service time scale (e.g., days of service) is then

$$F\left(t_1, t_2; \boldsymbol{\theta}_C, \boldsymbol{\theta}_R\right) = \int_0^\infty \int_0^\infty F_1(t_1|r_1) F_2(t_2|r_2) g(r_1, r_2) dr_1 dr_2$$

where the parameters $\boldsymbol{\theta}_C = (\eta_{C_1}, \sigma_{C_1}, \eta_{C_2}, \sigma_{C_2})$ are the cycles-to-failure distribution parameters and $\boldsymbol{\theta}_R = (\eta_{R_1}, \sigma_{R_1}, \eta_{R_2}, \sigma_{R_2}, \rho)$ are the use-rate distribution parameters. Note that we expect R_1 and R_2 to be positively correlated because products that are used more also tend to be abused more. The bivariate survival function for the real-time scale is

$$S\left(t_1, t_2; \boldsymbol{\theta}_C, \boldsymbol{\theta}_R\right) = P(T_1 > t_1, T_2 > t_2) = 1 - F_1(t_1) - F_2(t_2) + F(t_1, t_2). \quad (3)$$

The cdf of $T = \min\{T_1, T_2\}$, the lifetime of the series system, is

$$F(t) = F(t; \boldsymbol{\theta}_R, \boldsymbol{\theta}_C) = 1 - S(t, t; \boldsymbol{\theta}_R, \boldsymbol{\theta}_C). \quad (4)$$

5 MAXIMUM LIKELIHOOD ESTIMATION FOR THE MULTIPLE FAILURE MODE USE-RATE MODEL

5.1 Notation

In order to define the likelihood for the multiple failure mode use-rate model we need to extend some of the notation that we introduced in Section 4. Let $S_j(t|r_j)$ denote the conditional survival function for T_j at fixed use rate r_j, $j = 1, 2$. Also, $f_j(t|r_j)$ will be used to denote the corresponding conditional pdf for T_j at fixed use rate r_j, and $g_j(r)$ will be used to denote the marginal pdf of the random use rates R_j with parameter $(\eta_{R_j}, \sigma_{R_j})$, and $g_j(r|r_k)$ is the conditional pdf for use rate R_j for fixed use rate $R_k = r_k, (j \neq k)$.

The warranty-return data are censored and the units fail due to one of the two different failure modes. Now using the subscript i on T to index the observational units, we observe $(T_i, \delta_{i1}, \delta_{i2})$, $i = 1, 2, \cdots n$. Here $T_i = \min\{T_{i1}, T_{i2}, t_{ic}\}$ where T_{ij} is the failure time corresponding to failure mode j for unit i. The censoring indicator is $\delta_{ij} = 1$ for a type j failure and $\delta_{ij} = 0$ otherwise. If unit i is censored at time t_{ic}, $\delta_{i1} = \delta_{i2} = 0$.

Recall that units tested in the two separate ALTs fail only due to the particular failure mode for which the ALTs were designed. Thus the ALT failure times in $(C_{ij}, \tau_{ij}), i = 1, 2, \cdots, m_j$ (in units of test cycles) for the two failure modes are independent. The ALT for failure mode j may also be censored at time t_{cj} with censoring indicator τ_{ij} (our notation corresponds to the usual situation in an ALT where all units in the test have the same censoring time, but this is, of course, not really a constraint). In particular, $\tau_{ij} = 1$ if unit i failed due to failure type j and $\tau_{ij} = 0$ otherwise.

5.2 The Likelihood

The bivariate survival function in (3) can be re-expressed as

$$S(t_1, t_2) = \int_0^\infty \int_0^\infty S_1(t_1|r_1)\, S_2(t_2|r_2) g(r_1, r_2) dr_1 dr_2.$$

Then the likelihood for field observation i at time t_i can be expressed as

$$L_i = \left[-\frac{\partial S(t, t_i)}{\partial t} \Big|_{t=t_i} \right]^{\delta_{i1}} \left[-\frac{\partial S(t_i, t)}{\partial t} \Big|_{t=t_i} \right]^{\delta_{i2}} [S(t_i, t_i)]^{(1-\delta_{i1})(1-\delta_{i2})} \quad (5)$$

where, for example, the contribution for a failure of type 1 is

$$\left[-\frac{\partial S(t, t_i)}{\partial t} \Big|_{t=t_i} \right] = \int_0^\infty f_1(t_i|r_1) g_1(r_1) \int_0^\infty S_2(t_i|r_2) g_2(r_2|r_1) dr_2 dr_1.$$

See Figure 6 for an illustration of the likelihood contributions for failures of type 1, for failures of type 2, and for censored observations. The total log-likelihood for the field data is

$$\begin{aligned} \mathcal{L}_{\text{FIELD}}(T; \boldsymbol{\theta}_R, \boldsymbol{\theta}_C) &= \sum_{i=1}^{n} \delta_{i1} \log \left[-\frac{\partial S(t, t_i)}{\partial t} \Big|_{t=t_i} \right] + \delta_{i2} \log \left[-\frac{\partial S(t_i, t)}{\partial t} \Big|_{t=t_i} \right] \\ &\quad + (1 - \delta_{i1} - \delta_{i2}) \log[S(t_i, t_i)]. \end{aligned}$$

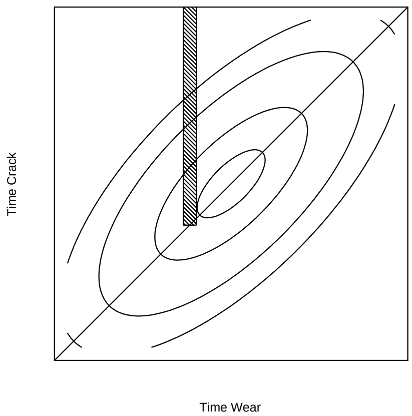
The log-likelihood for the ALT data is

$$\begin{aligned} \mathcal{L}_{\text{ALT}}(C; \boldsymbol{\theta}_C) &= \sum_{j=1}^{2} \sum_{i=1}^{m_j} \tau_{ij} \log \left(f_{C_j} \left[\frac{\log(c_{ij}) - \log(\eta_{C_j})}{\sigma_{C_j}} \right] \right) \\ &\quad + (1 - \tau_{ij}) \log \left(1 - \Phi \left[\frac{\log(c_{ij}) - \log(\eta_{C_j})}{\sigma_{C_j}} \right] \right). \end{aligned}$$

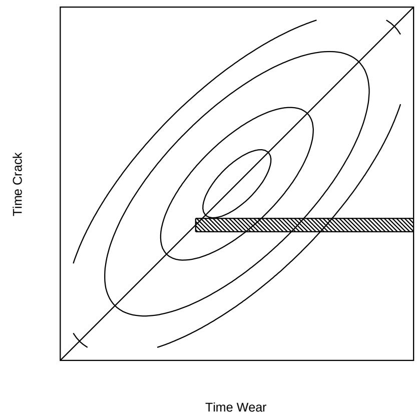
The total log likelihood for the field data and the ALT data is

$$\mathcal{L}(T, C; \boldsymbol{\theta}_R, \boldsymbol{\theta}_C) = \mathcal{L}_{\text{FIELD}}(T; \boldsymbol{\theta}_R, \boldsymbol{\theta}_C) + \mathcal{L}_{\text{ALT}}(C; \boldsymbol{\theta}_C). \quad (6)$$

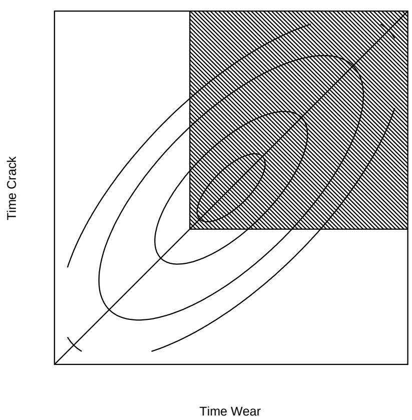
The ML estimates, denoted by $\widehat{\boldsymbol{\theta}}_R, \widehat{\boldsymbol{\theta}}_C$, can be obtained by maximizing (6).



(a)



(b)



(c)

(a): Probability of a Wear (Type 1) Failure Observation

(b): Probability of a Crack (Type 2) Failure Observation

(c): Probability of a Censored Observation

Figure 6: Multiple Failure Mode Continuous Mixture Use-Rate Model Likelihood Contributions for the Data from $F(t_1, t_2)$.

5.3 Estimate of the Appliance B cdf with Both Failure Modes Active

The ML estimator of $F(t; \boldsymbol{\theta}_R, \boldsymbol{\theta}_C)$, the cdf of T in (4) is

$$\widehat{F}(t) = F(t; \widehat{\boldsymbol{\theta}}_R, \widehat{\boldsymbol{\theta}}_C).$$

There is no closed-form expression for the quantile functions, but estimates of quantiles are, of course, easy to compute numerically. Approximate confidence intervals for either $F(t; \boldsymbol{\theta}_R, \boldsymbol{\theta}_C)$ or quantiles of the distribution can be computed by using the delta method, via bootstrap simulation-based methods, or by inverting a likelihood-ratio test. In our example, we have used the likelihood-based method.

5.4 Likelihood Expressions for the Lognormal/Lognormal Special Case

Figures 4 and 5 show that lognormal distributions provide an excellent description of the failure times for both failure modes from both the field and the ALT data. This suggests that a joint lognormal distribution can also be used to describe the joint use-rate distribution. Results of fitting Weibull distributions (not shown here) were not nearly as good.

In this section we present special-case formulas for the lognormal distribution. Use of these expressions simplifies programming and reduces computational burden. They also allowed us to check computer codes written for the more general expressions given in Section 5.2.

Letting the cycles (C_1, C_2) random variables and the use-rate random variables (R_1, R_2) both have bivariate lognormal distributions,

$$(T_1, T_2, R_1, R_2)' \quad \sim \quad \text{MVLOGNOR}\left(\boldsymbol{\eta}, \begin{bmatrix} \boldsymbol{\Sigma}_T & \boldsymbol{\Sigma}_R \\ \boldsymbol{\Sigma}_R & \boldsymbol{\Sigma}_R \end{bmatrix}\right)$$

where

$$\boldsymbol{\eta} = \left[\frac{\eta_{C_1}}{\eta_{R_1}}, \frac{\eta_{C_2}}{\eta_{R_2}}, \eta_{R_1}, \eta_{R_2}\right]'$$

is a vector of medians and

$$\boldsymbol{\Sigma}_T = \begin{bmatrix} \sigma^2_{C_1} + \sigma^2_{R_1} & \rho\sigma_{R_1}\sigma_{R_2} \\ \rho\sigma_{R_1}\sigma_{R_2} & \sigma^2_{C_2} + \sigma^2_{R_2} \end{bmatrix} \quad \boldsymbol{\Sigma}_R = \begin{bmatrix} \sigma^2_{R_1} & \rho\sigma_{R_1}\sigma_{R_2} \\ \rho\sigma_{R_1}\sigma_{R_2} & \sigma^2_{R_2} \end{bmatrix}.$$

From this the marginal cdf of T_j is

$$F_j(t) \quad = \quad \Phi_{\text{nor}}\left[\frac{\log(t) - \log\left(\frac{\eta_{C_j}}{\eta_{R_j}}\right)}{\sqrt{\sigma^2_{C_j} + \sigma^2_{R_j}}}\right], \quad j = 1, 2,$$

and the joint cdf of (T_1, T_2) is

$$F(t_1, t_2; \boldsymbol{\theta}_C, \boldsymbol{\theta}_R) \quad = \quad \Phi_2(z_1, z_2, \rho_{TT})$$

where $\Phi_2(\cdot, \cdot, \cdot)$ is the standardized bivariate normal distribution cdf. Here $z_j = \left[\log(t_j) - \log(\eta_{C_j}/\eta_{R_j})\right] / \sqrt{\sigma^2_{C_j} + \sigma^2_{R_j}}$, $j = 1, 2$, and

$$\rho_{TT} = \text{Corr}[\log(T_1), \log(T_2)] = \frac{\rho\sigma_{R_1}\sigma_{R_2}}{\sqrt{(\sigma^2_{C_1} + \sigma^2_{R_1})(\sigma^2_{C_2} + \sigma^2_{R_2})}}. \tag{7}$$

For this special-case formulation, there are explicit expressions for the likelihood terms in (5). In particular,

$$\left[-\frac{\partial S(t, t_i)}{\partial t}\bigg|_{t=t_i}\right] = \frac{\phi_{\text{nor}}(z_{i1})}{t_i\sqrt{\sigma^2_{C_1} + \sigma^2_{R_1}}}\,\Phi_{\text{nor}}\left(-\frac{z_{i2} - \rho_{TT}z_{i1}}{\sqrt{1 - \rho^2_{TT}}}\right)$$

$$\left[-\frac{\partial S(t_i, t)}{\partial t}\bigg|_{t=t_i}\right] = \frac{\phi_{\text{nor}}(z_{i2})}{t_i\sqrt{\sigma^2_{C_2} + \sigma^2_{R_2}}}\,\Phi_{\text{nor}}\left(-\frac{z_{i1} - \rho_{TT}z_{i2}}{\sqrt{1 - \rho^2_{TT}}}\right)$$

where

$$z_{ij} = \frac{\log(t_i) - \log\left(\frac{\eta_{C_j}}{\eta_{R_j}}\right)}{\sqrt{\sigma^2_{C_j} + \sigma^2_{R_j}}}, \quad j = 1, 2.$$

5.5 Correlation Between Use Rates for Different Failure Modes

Nonparametrically, the joint distribution in a multiple failure mode problem cannot be completely identified in the usual situation when only the minimum time to failure can be observed (e.g., Tsiatis 1975). With a parametric assumption, the joint distribution might be identifiable. Nádas (1971) and Basu and Ghosh (1978) describe technical results relating to identifiability of the bivariate normal distribution under the competing risks model.

Even with a parametric assumption, such as the joint lognormal distribution used in the Appliance B example, there little information in the data about ρ. As shown in Figure 7, the lognormal/lognormal use-rate model profile likelihood for ρ for the Appliance B data is almost flat. This due to the heavy censoring in the data (the fraction failing in the field data was only about 2%) and the limited amount of data. Thus it will be necessary to make some assumption about ρ and to use sensitivity analysis to assess the effects of departures from such assumptions.

An assumption of $\rho = 0$ might be realistic in some special case applications. For example, time to automobile coating failures caused by UV exposure may be

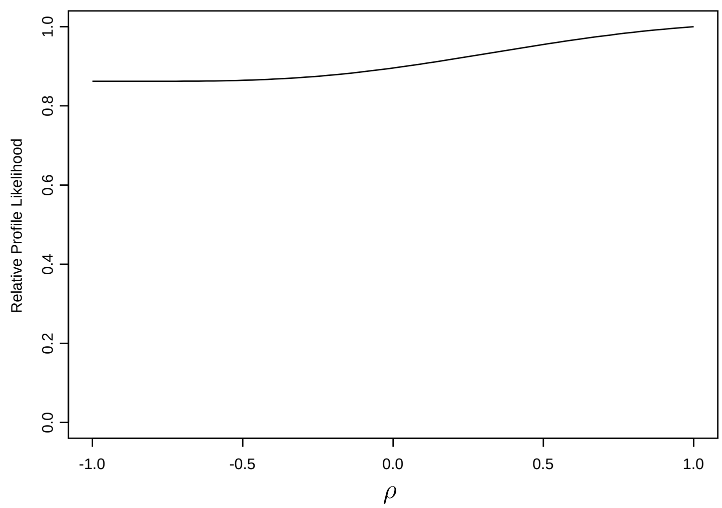


Figure 7: Relative Profile Likelihood for ρ for the Bivariate Lognormal/Lognormal Use-Rate Model.

approximately independent of bearing failures caused by miles driven). Recall the earlier washing machine example, however. The correlation between in-balance use versus out-of-balance abuse rates will almost certainly be positive (i.e., $\rho > 0$). The situation is similar for the turbine device in Appliance B.

An important special case, which we see arising frequently in practice, is when two or more failure modes have exactly the same underlying definition of use (as described in Section 2.1) and when use affects all failure modes in the same way (e.g., two different bearings in the same automobile). In this case, $\rho = 1$, and one can take advantage of some simplifications in the formulation of the problem.

For applications like the washing machine and the turbine device in Appliance-B the predominant failure modes have two different mechanisms with different definitions of use (use and abuse). We expect units with more use to generally have more abuse, implying a positive correlation between R_1 and R_2. In this case, a more reasonable assumption would be that the ratio of the amount of time in the abusive state to the amount of time in the use state is independent of the use rate for normal use. For our use-rate model, this implies that the ratio R_2/R_1 is independent of R_1. Then $\log(R_2) - \log(R_1)$ is independent of $\log(R_1)$. For the joint lognormal use-rate model, this implies that

$$\mathrm{Cov}[\log(R_2) - \log(R_1), \log(R_1)] = \rho\sigma_{R_1}\sigma_{R_2} - \sigma^2_{R_1} = 0. \tag{8}$$

Thus, for the Appliance B example, (8) implies that $\rho = \sigma_{R_{\mathrm{Wear}}}/\sigma_{R_{\mathrm{Crack}}} > 0$ and thus $\sigma_{R_{\mathrm{Crack}}} > \sigma_{R_{\mathrm{Wear}}}$. In our analyses in the following section we will use this model assumption, along with some sensitivity analyses to compare results. We call this the *wear-ratio independence* model.

6 ML ESTIMATION RESULTS FOR APPLIANCE B

This section presents ML estimates for Appliance B under different assumptions for the dependence structure of the joint use-rate distribution. The results are summarized in Table 2. In this table we give estimates of $t_{0.001}$, $t_{0.01}$, and $t_{0.20}$ for the marginal days-to-failure distributions for each failure mode and for the series system (corresponding to the situation when both failure modes are active). We chose these particular quantiles because span the range of interest for the application. We only go up to $t_{0.20}$ because this quantile is approximately 20 years, which is considerably beyond the expected *technological life* of Appliance B. The $\widehat{\sigma}$ column gives the estimate of the lognormal shape parameter (standard deviation of the log failure times) for the corresponding failure mode for the ALT or warranty part of the model. In particular $\widehat{\sigma} = \widehat{\sigma}_{C_j}$ for the ALT lines and $\widehat{\sigma} = \sqrt{\widehat{\sigma}^2_{C_j} + \widehat{\sigma}^2_{R_j}}$ for the Warranty lines in the table where, as before, $j = 1$ for the wear failure mode and $j = 2$ for the crack failure mode. The $\widehat{\rho}_{TT}$ column gives the ML estimate of the correlation between the failure times for the two different failure modes and was computed by evaluating (7) at $\widehat{\boldsymbol{\theta}}_R, \widehat{\boldsymbol{\theta}}_C$. In Table 2, we report $\widehat{\rho}_{TT}$ instead of $\widehat{\rho}$ because some observed effects in our sensitivity analyses are better described by this parameter. The units of time is days for the warranty rows and use cycles for the ALT rows in the table.

The estimates of the warranty returns for the two different failure modes (straight lines going through the nonparametric field return estimates) in Figures 4 and 5 were obtained by maximizing the lognormal/lognormal total log-likelihood function in (6) with fixed $\rho = 0$, giving $\widehat{\boldsymbol{\theta}}_C = (\widehat{\eta}_{C_1}, \widehat{\sigma}_{C_1}, \widehat{\eta}_{C_2}, \widehat{\sigma}_{C_2})$ and $\widehat{\boldsymbol{\theta}}_R = (\widehat{\eta}_{R_1}, \widehat{\sigma}_{R_1}, \widehat{\eta}_{R_2}, \widehat{\sigma}_{R_2}, 0)$. We also fit the wear-ratio independence model by maximizing (6) after replacing σ_{R_2} by σ_{R_1}/ρ, as implied by (8). Interestingly, as seen in Table 2, ML estimates for the *series system* (i.e., both failure modes active) quantiles for the $\rho = 0$ and the wear-ratio independence models are almost exactly the same. Visually, the corresponding probability plots were so similar to what we see in Figures 4 and 5 that we do not present them here. Of course this is not surprising given the results described in Section 5.5. There are some important differences in the estimates of the marginal distributions. Section 8 contains more comparisons between these models and further discussion.

Figure 8 is a lognormal probability plot showing the series-system model (both failure modes active) ML estimate of $F(t)$ extrapolated to 10 years, computed using the $(\widehat{\boldsymbol{\theta}}_R, \widehat{\boldsymbol{\theta}}_C)$ from the wear-ratio independent use-rate model. The ML estimate agrees well with the Kaplan-Meier estimate from the warranty data. The other two lines show the ML estimates of the marginal cdfs for the wear and crack failure modes. The

Table 2: Summary of ML Estimation Results for Appliance B.

Figure	Data Source	Failure Mode	Model	Total Likelihood	Days/Cycles $\widehat{t}_{0.001}$	$\widehat{t}_{0.01}$	$\widehat{t}_{0.20}$	$\widehat{\sigma}$	$\widehat{\rho}_{TT}$
			Current Product Generation						
4	ALT	Wear			26	51	188	0.88	
5	ALT	Crack			160	198	300	0.28	
4	Warranty	Wear	Fixed $\rho = 0$	-600.1	85	246	1945	1.39	0
5	Warranty	Crack			169	595	6868	1.65	
	Warranty	System			78	223	1612	NA	
	ALT	Wear			27	52	189	0.87	
	ALT	Crack			160	198	300	0.28	
8	Warranty	Wear	Ratio Ind.	-600.0	85	243	1877	1.38	0.54
8	Warranty	Crack			162	522	5095	1.53	
8	Warranty	System			79	223	1651	NA	
	ALT	Wear			26	52	188	0.87	
	ALT	Crack			160	198	300	0.28	
	Warranty	Wear	Fixed $\rho = 1$	-600.0	84	239	1821	1.37	0.76
	Warranty	Crack			148	450	3911	1.46	
	Warranty	System			79	223	1661	NA	
			Future Product Generation						
	ALT	Wear			131	256	940	0.88	
	ALT	Crack			320	396	601	0.28	
	Warranty	Wear	Fixed $\rho = 0$	-600.1	425	1231	9724	1.39	0
	Warranty	Crack			338	1190	13736	1.65	
10	Warranty	System			277	831	6105	NA	
	ALT	Wear			133	258	943	0.87	
	ALT	Crack			320	396	601	0.28	
9	Warranty	Wear	Ratio Ind.	-600.0	424	1215	9387	1.38	0.54
9	Warranty	Crack			324	1045	10191	1.53	
9,10	Warranty	System			275	805	6157	NA	
	ALT	Wear			132	258	942	0.87	
	ALT	Crack			320	396	601	0.28	
	Warranty	Wear	Fixed $\rho = 1$	-600.0	421	1195	9106	1.37	0.76
	Warranty	Crack			296	900	7821	1.46	
10	Warranty	System			266	771	5947	NA	

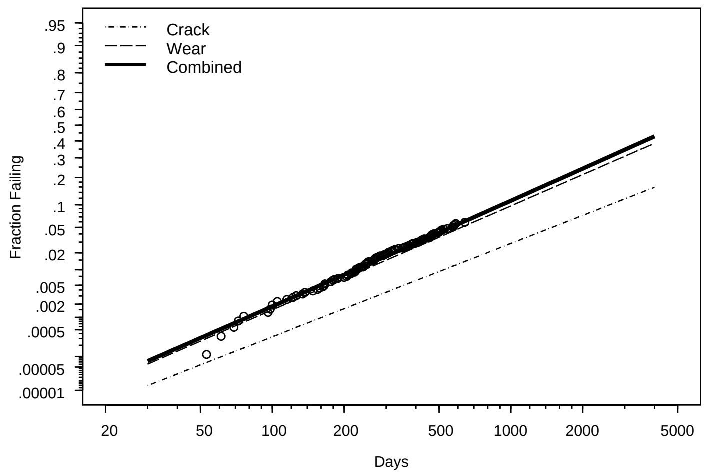


Figure 8: ML Estimate of the Original Design Appliance B Marginal cdfs and Combined Failure Modes Series-System Model cdf $F(t) = \Pr(\min(T_1, T_2) \leq t) = 1 - S(t, t)$ Under the Wear-Ratio Independent ($\widehat{\rho} = 0.71$, $\widehat{\rho}_{TT} = 0.54$) Use-Rate Model.

marginal cdf for the wear (crack) failure modes can be interpreted as the system cdf if the crack (wear) failure mode could be eliminated. Table 2 also gives ML estimates for $\rho = 1$. We also use these results and those for the other dependence structure assumptions in Section 7.

7 RELIABILITY PREDICTION FOR THE NEWLY DESIGNED TURBINE DEVICE

7.1 Design Changes

The manufacturers of Appliance B wanted to improve the reliability of this product by redesigning the life-limiting turbine device. Having good quantitative information about the effect of the individual failure modes (e.g., as given in Table 2 and Figure 8) allowed the engineers to decide how to focus their efforts and how to best apportion possible added cost to manufacture the improved component. For example, Figure 8 made it clear that it would not be cost effective to attempt to move the cdf for the crack failure mode to the right, unless the cdf for the wear mode could also be moved substantially to the right.

Given good base-line information about the failure time distribution of a failure mode (as was available for the wear and crack failure modes in the Appliance B turbine device), design engineers have tools (like finite element modeling) that will allow them to obtain reasonably accurate predictions about the effects that simple geometrical or size changes to a mechanical product will have on reliability. In particular, for proposed design changes to the turbine device, the engineers believed that

- The wear failure mode cycles-to-failure distribution would improve by a factor of approximately $\nu_1 = 5$ in the cycles scale.
- The crack failure mode cycles-to-failure distribution would improve by a factor of $\nu_2 = 2$ in the cycles scale.

In order to do a cost/benefit analysis and to plan for future warranty costs, management wanted an estimate of the life time distribution of Appliance B for the new design.

7.2 Reliability Prediction for New Design of Appliance B

The cdf of T, the failure-time of Appliance B with the newly-designed turbine device is

$$F\left(t; \boldsymbol{\theta}_R, \boldsymbol{\theta}_C^{\text{new}}\right) = 1 - S\left(t, t; \boldsymbol{\theta}_R, \boldsymbol{\theta}_C^{\text{new}}\right)$$

where $\boldsymbol{\theta}_C^{\text{new}} = (\nu_1 \eta_{C_1}, \sigma_{C_1}, \nu_2 \eta_{C_2}, \sigma_{C_2})$. Here ν_1 is the wear failure mode improvement factor and ν_2 is the crack failure mode improvement factor. So the estimator of the cdf of field life of the new product is $F\left(t; \widehat{\boldsymbol{\theta}}_R, \widehat{\boldsymbol{\theta}}_C^{\text{new}}\right)$ where $\widehat{\boldsymbol{\theta}}_C^{\text{new}} = (\nu_1 \widehat{\eta}_{C_1}, \widehat{\sigma}_{C_1}, \nu_2 \widehat{\eta}_{C_2}, \widehat{\sigma}_{C_2})$. Figure 9 shows the estimates of the marginal distributions for the wear and the crack failure modes as well as combined failure modes series-system model cdf using the wear-ratio independent use-rate model for the new design of the turbine device in Appliance B. The bottom of Table 2 contains a numerical summary of some of the important ML estimates for these reliability improvement projections.

Figure 10 gives the ML estimate along with likelihood-based pointwise 95% confidence intervals for the estimated combined failure modes series-system cdf $\widehat{F}(t)$ for the new design of the turbine device in Appliance B, out to more than 20 years. There are separate lines for the wear-ratio independent model and for the models with fixed values of ρ equal to 0 and 1. There is, however, little difference in the estimates for different assumed values of ρ indicating that, in this example, the predictions are insensitive to the assumption about the correlation between the different use rates. This will be discussed further in Section 8.

7.3 Assumptions for Prediction of Field Reliability

There are a number of important assumptions needed for the validity of the Appliance B predictions presented in this section, and other similar predictions in other

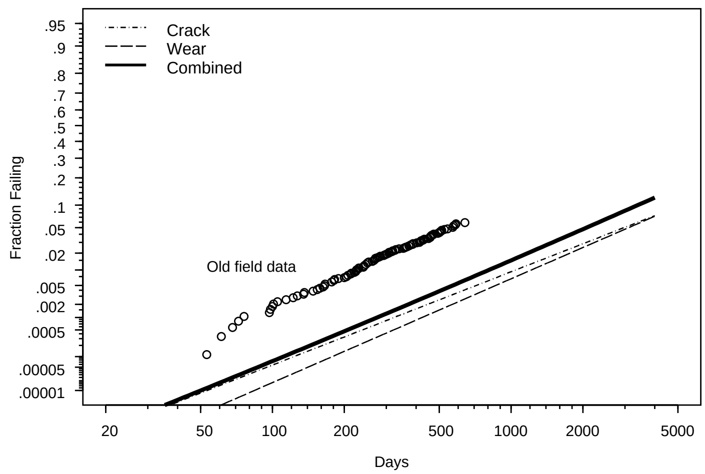


Figure 9: ML Estimate of the New Design Appliance B Marginal cdfs and Combined Failure Modes Series-System Model cdf $F(t) = \Pr(\min(T_1, T_2) \leq t) = 1 - S(t, t)$ Under the Wear-Ratio Independent ($\widehat{\rho} = 0.71$, $\widehat{\rho}_{TT} = 0.54$) Use-Rate Model, Compared with the Old Field Data.

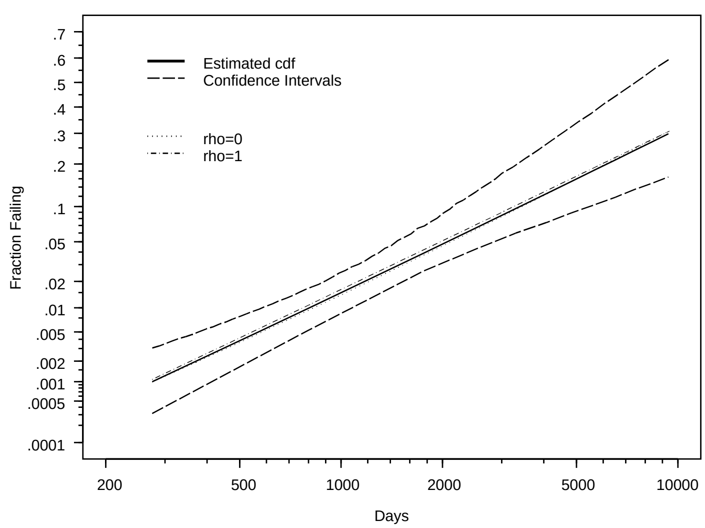


Figure 10: Likelihood-Based Pointwise Approximate 95% Confidence Intervals for the cdf of New Design of Appliance B on Lognormal Probability Paper
.

applications. The distributional, correlational, and other assumptions have already been mentioned and we have used the data, as much as is possible, for assessment. Here we elaborate on these assumptions and list some other assumptions. All of these assumptions were deemed to be reasonable by the engineers involved in the study.

- The engineering models used to predict improvement of proposed design changes are reasonably accurate. Of course these models are not perfect. Sensitivity analyses under different assumptions (details not given here) are important in such situations.

- The cycles-to-failure distribution does not depend on cycling rate. In the accelerated tests, units were allowed to "cool down" between uses, mimicking actual use and assuring that there would be no excessive heat build-up that could cause the cycles-to-failure distribution to change at the increased use rate.

- Wear and cracking are the primary failure types. It is always possible that other failure modes may arise in the future, but because about 10% (448 out of 4728) of the units in the field had survived more then 600 days (out of a 730 day warranty period) and because of the careful failure mode analysis that had been done on each failure, there was a good deal of confidence that there would be no other important failure mode arising during the warranty period. Of course, with the extension of the life for the wear and crack failure modes, it is possible that an underlying "masked" failure mode might begin to arise in the new design.

- ALTs adequately mimic the field-failure mechanisms. With the addition of the second accelerated life test to produce the wear failure modes, the engineers responsible for the reliability of Appliance B were confident that they were accurately reproducing the important failure modes in their accelerated tests. For example, the kind of accelerated tests done with Appliance B would tend to inhibit failures due to corrosion. But corrosion was not, for this product, a problem in the field. If corrosion had been a problem, a separate accelerated test for this failure mode might have been needed.

- Distributions of use rates and patterns of use in the field do not change over time (i.e., from the previous design to the new design). One concern in applications where use-rate-based predictions are used is that improvements in performance of a product may actually increase use rates and therefore cause bias in warranty predictions! Because the main improvement to Appliance B was in terms of reliability itself, this was not thought to be an unreasonable assumption.

8 THE IMPORTANCE OF USING THE CORRECT CORRELATIONAL MODEL IN APPLICATIONS WITH COMPETING FAILURE MODES

Sensitivity analyses showed that our conclusions about the future reliability of Appliance B were not strongly dependent on the assumption about the correlation between the two different use rates when the correlation was varied over the physically plausible range of 0 to 1. We do not want to leave the impression that this is always true. It is not. Because the conclusions are more general than our use rate model and example, we will discuss this issue in terms of the distributions of the field returns for two different, potentially correlated, failure modes.

Suppose that a system has two components and that the failure times of these components can be described by a joint log-location-scale distribution with similar shape parameters for the two marginal distributions. The distribution of the minimum of the two random variables can be approximated by the same log-location-scale distribution with a shape parameter that is between those for the marginal distributions and the adequacy of the approximation does not depend strongly on the correlation between the random variables. This result is exact for two independent Weibull distributions with the same shape parameter.

In light of this approximate equivalence result and the identifiability result mentioned in Section 5.5, it is not surprising that when two failure modes have marginal distributions with similar shape parameters (similar slopes when displayed on a probability plot), estimates of $F(t)$, the series-system cdf, do not depend strongly on the assumed amount of association (correlation for a joint lognormal distribution). This can be seen for the Appliance B example in Figure 8. There could, however, be important biases in estimating the individual marginal distributions if the $\rho = 0$ model is used incorrectly.

All other things being equal, the effect of using an incorrect $\rho = 0$ model will be stronger when ρ_{TT} is larger. For the lognormal distribution, the expression for ρ_{TT} in (7) suggests that

- ρ_{TT} is proportional to ρ and
- ρ_{TT} tends to be larger when the use-rate variability dominates the cycles-to-failure variability.

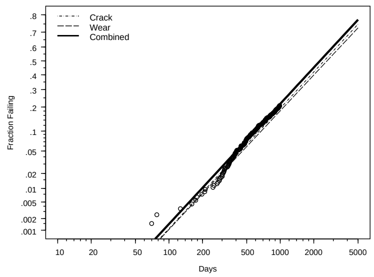
To provide insight into situations other than our particular example, we performed simulations to study the effect of making different amounts of improvement to the components of a system and two different values of ρ_{TT}. We used relatively large sample sizes (1000 units in the field and 100 units in each ALT) so that sampling variability would not be important. For continuity of discussion we continue to refer to the failure modes as crack and wear.

Figures 11 and 12 display the results of two of our simulations. In Figure 11, we simulated from a model in which the crack and wear marginal distributions are similar and thus, under the wear-ratio independence model, ρ_{TT} will be large. For our simulated data in Figure 11, $\widehat{\rho}_{TT} = 0.98$. For the data in Figure 12, we have $\widehat{\rho}_{TT} = 0.45$.

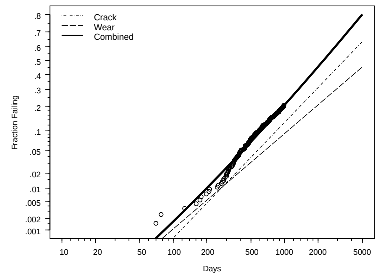
Figures 11 and 12 compare the effect that changes to the cycles-to-failure distributions have on the estimates of the field failure-time marginal cdf estimates ($\widehat{F}_{\text{Wear}}(t)$ and $\widehat{F}_{\text{Crack}}(t)$) and on the field failure-time series-system cdf estimate ($\widehat{F}(t)$). Both figures show the comparison for all four combinations of improvements (similar to the analyses done for Appliance B) in the wear and crack failure modes equal to $1\times$ and $3\times$. For each data set (one data set for each figure) we fit the correct (wear-ratio independence) model and the $\rho = 0$ model to the simulated data set. The left-hand column of Figures 11 and 12 gives the correct model estimates and the right-hand column gives the $\rho = 0$ model estimates for the corresponding data set (one for each figure). The rows correspond to the differing amounts of improvement to the system components. Figure 11 (a) and (b) compare parameter estimates for the correct and $\rho = 0$ models, respectively, in the situation where there is no improvement in the turbine device. The comparison shows that changing models has a strong effect on the marginal cdfs ($\widehat{F}_{\text{Wear}}(t)$ and $\widehat{F}_{\text{Crack}}(t)$), but almost no effect on the series-system cdf estimate ($\widehat{F}(t)$). This is as expected because with the high correlation between the failure modes, if you eliminate just one failure mode, the other one is unaffected. Physically, such a fix would correspond to fixing a failure mode symptom, rather than the root cause.

Figure 11 (g) and (h) provide a similar comparison for the situation where there is an equal $3\times$ improvement for both failure modes in the turbine device. Again there is a strong effect on $\widehat{F}_{\text{Wear}}(t)$ and $\widehat{F}_{\text{Crack}}(t)$, but almost no effect on $\widehat{F}(t)$. When, however, the improvement comes to one failure mode marginal distribution or the other, using the incorrect ($\rho = 0$) model causes serious bias not only in the marginals, but also in $\widehat{F}(t)$.

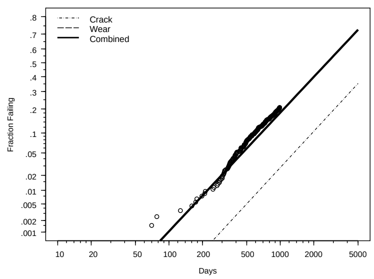
In Figure 12, the results are similar, but much less dramatic. In all cases there are, however, still important differences between the ML estimates of the marginal distributions when comparing the correct wear-ratio independent model with the incorrect $\rho = 0$ model. With equal changes to the marginal distributions, we again see little difference in the series-system estimates. For unequal changes to the marginal distributions, the effect tends to be larger on one side of the distribution than the other because of the differences in the shape parameters (reflected in the slopes of the marginal cdf estimates). But, as expected, the differences are not as large as they were in Figure 11 with the larger value of $\widehat{\rho}_{TT}$.



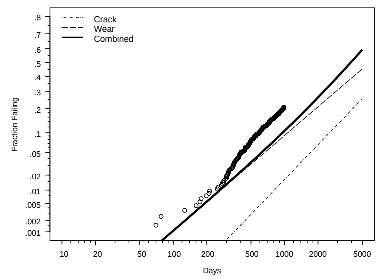
(a) Wear × 1 Crack × 1 $\widehat{\rho}_{TT} = 0.98$



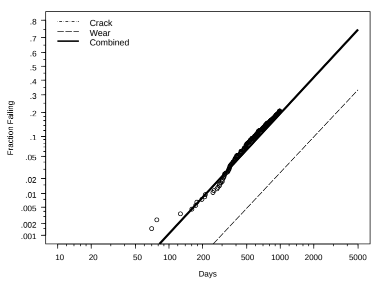
(b) Wear × 1 Crack × 1 $\rho_{TT} = 0$



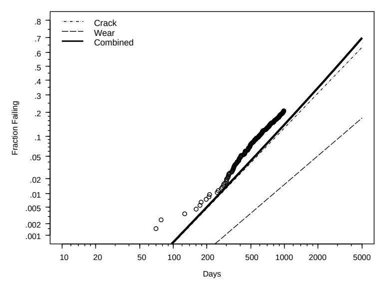
(c) Wear × 1 Crack × 3 $\widehat{\rho}_{TT} = 0.98$



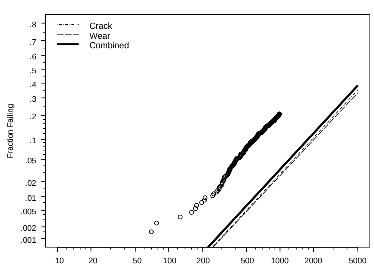
(d) Wear × 1 Crack × 3 $\rho_{TT} = 0$



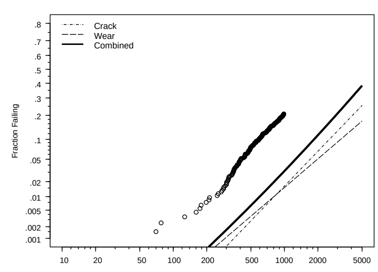
(e) Wear × 3 Crack × 1 $\widehat{\rho}_{TT} = 0.98$



(f) Wear × 3 Crack × 1 $\rho_{TT} = 0$

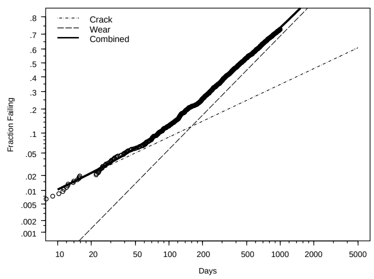


(g) Wear × 3 Crack × 3 $\widehat{\rho}_{TT} = 0.98$

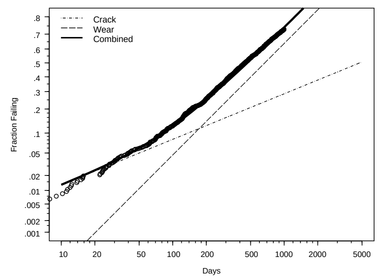


(h) Wear × 3 Crack × 3 $\rho_{TT} = 0$

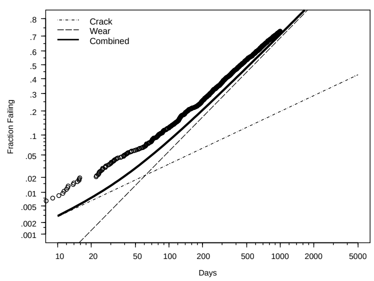
Figure 11: Comparison of the Effect of Using the $\rho = 0$ Model to Estimate Marginal Distributions and System Reliability for Different Levels of Improvement of System Components for Simulated Data with $\widehat{\rho}_{TT} = 0.98$.



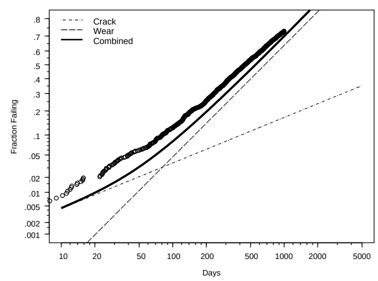
(a) Wear × 1 Crack × 1 $\widehat{\rho}_{TT} = 0.45$



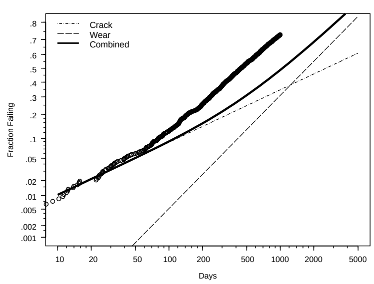
(b) Wear × 1 Crack × 1 $\rho_{TT} = 0$



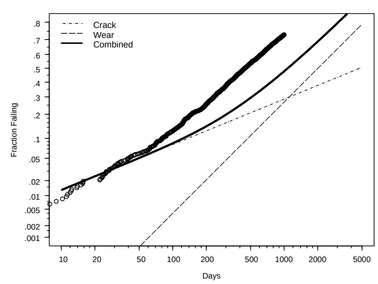
(c) Wear × 1 Crack × 3 $\widehat{\rho}_{TT} = 0.45$



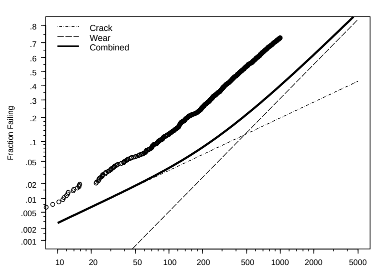
(d) Wear × 1 Crack × 3 $\rho_{TT} = 0$



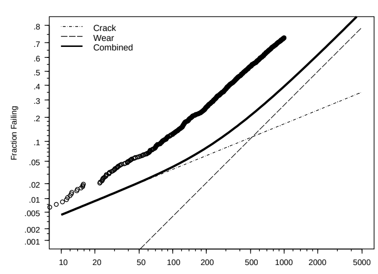
(e) Wear × 3 Crack × 1 $\widehat{\rho}_{TT} = 0.45$



(f) Wear × 3 Crack × 1 $\rho_{TT} = 0$



(g) Wear × 3 Crack × 3 $\widehat{\rho}_{TT} = 0.45$



(h) Wear × 3 Crack × 3 $\rho_{TT} = 0$

Figure 12: Comparison of the Effect of Using the $\rho = 0$ Model to Estimate Marginal Distributions and System Reliability for Different Levels of Improvement of System Components for Simulated Data with $\widehat{\rho}_{TT} = 0.45$.

9 CONCLUDING REMARKS AND AREAS FOR FURTHER RESEARCH

It is possible to predict field performance by using ALT data, if done carefully and with physically-motivated models. This has been illustrated in this paper with a use-rate model and two examples. There are a number of extensions to more complicated situations that we have either seen or anticipate seeing in other products. These include the following.

- With high reliability components, few or no failures would be expected in ALTs of reasonable length. In some applications it may be possible to observe degradation on test units. Models and inferences, similar to those presented in the paper could be developed for degradation data.

- Our model to relate accelerated tests and field reliability was relatively simple. Aging or chemical degradation of materials and exposure to variables that affect aging or degradation and that vary over time make matters considerably more complicated. We have already mentioned Nelson (2001). Martin, Nguyen, and Wood (2005) describe an application in the area of accelerated tests for organic paints and coatings subjected to outdoor weathering where experimental and environmental variables would include UV spectrum and intensity, temperature, and humidity. In some applications it will be necessary to extend this kind of work to handle multiple failure modes.

- For the Appliance B application it was reasonable to assume that each use had the same effect on time-to-wear life and that each abusive use had the same effect on time-to-crack life. In some applications each use has an associated stress that can be described by a probability distribution or, more generally, as stochastic process, with the parameters varying over the population of units. It would be possible to extend the model used here to allow nonconstant stress spectrum to describe the actual uses. Miner's rule (as described in Chapter 10 of Nelson 1990) might provide a useful model for doing this in some applications.

- Instead of doing sensitivity analysis to assess the effects of different assumptions about the value of ν_1 and ν_2, one could use a Bayesian approach to inference in this kind of application. Using such an approach would also allow the incorporation of possible prior information on the value of ρ and perhaps the values of the distribution shape parameters, as similar failure modes often give similar shape parameter values across different applications, providing increased precision in estimates or predictions.

Acknowledgments

Meeker and Hong's work on this paper was partially supported by funds from NSF Award CNS0540293 to Iowa State University. We would like to thank Rob Easterling and Tim Davis for helpful comments and discussion on earlier presentations of this work. The computations needed for the analyses in this paper were done with SPLIDA (Meeker and Escobar 2004). The Appliance B data are available in the SPLIDA download.

References

Basu, A. P. and Ghosh, J. K. (1978), Identifiability of the Multinormal Distribution Under Competing Risks Model. *Journal of Multivariate Analysis* 8, 413-429.

Condra, L. W. (2001), *Reliability Improvement with Design of Experiments, 2nd ed., rev. and expanded*, Series: Quality and Reliability 59, Marcel Dekker Inc: New York.

Crowder, M. (2001), *Classical Competing Risks*, Boca Raton: Chapman and Hall/CRC.

David, H. A., and Moeschberger, M. L. (1978), *The Theory of Competing Risks*, London: Griffin.

Duchesne T. and Lawless J. F. (2000), Alternative Time Scales and Failure Time Models. *Lifetime Data Analysis*, 6, 157-79.

Escobar, L. A. and Meeker, W. Q. (1999), Statistical Prediction Based on Censored Life Data. *Technometrics* 41, 113-124.

Escobar, L. A. and Meeker, W. Q. (2006), A Review of Accelerated Test Models, *Statistical Science*, 21, 552-577.

Kordonsky K. B., and Gertsbakh, I. (1993), Choice of the Best Time Scale for System Reliability Analysis, *European Journal of Operational Research*, 65, 235-246.

Mannaa, D. K. Pala, S. and Sinhab, S. (2007) A Use-Rate Based Failure Model for Two-Dimensional Warranty *Computers & Industrial Engineering*, 52, 229-240.

Martin, J. W. Nguyen, T. and Wood, K. A. (2005), "Unresolved Issues Related to Predicting the Service Life of Polymeric Materials," in *Service Life Prediction:*

Challenge the Status Quo, Martin, J.W., Ryntz, R. A. and Dickie, R. A. (Eds.), Federation of Societies for Coatings Technology, FSCT, Blue Bell, PA.

Meeker, W. Q. and Escobar, L. A. (1998), *Statistical Methods for Reliability Data*. John Wiley & Sons: New York.

Meeker, W. Q. and Escobar, L. A. (2004), *SPLIDA User's Manual*. Available from www.public.iastate.edu/˜stat533/splida.html.

Nádas, A. (1971), The Distribution of the Identified Minimum of a Normal Pair Determines the Distribution of the Pair, *Technometrics* 13, 201-202.

Nelson, W. (1990), *Accelerated Testing: Statistical Models, Test Plans, and Data Analysis*. John Wiley & Sons: New York.

Nelson, W. (2001), Prediction of Field Reliability of Units, Each Under Differing Dynamic Stresses, from Accelerated Test Data. In *Handbook of Statistics,* Vol 20. N. Balakrishnan and C. R. Rao, eds. Elsevier.

Oakes, D. (1995), Multiple time scales in survival analysis. *Lifetime Data Analysis*, 1, 7-18.

Singpurwalla, N. D. (1995), Survival in Dynamic Environments, *Statistical Science* 10, 86-103.

Tsiatis, A. (1975), A Nonidentifiability Aspect of the Problem of Competing Risks. *Proceedings of the National Academy of Sciences,* 72, 20-22.

Wu, C. F. J. and Hamada, M. (2000), *Experiments: Planning, Analysis, and Parameter Design Optimization*. John Wiley & Sons, New York.

Yang, G. and Zaghati, Z. (2006), Accelerated Life Tests at Higher Usage Rates: a Case Study, In the proceedings of the *Reliability and Maintainability Symposium*, 313-317, IEEE, Philadelphia.