Probabilistic Aspects of LIFE PREDICTION

STP 1450

EDITORS:
W. Steven Johnson and Ben M. Hillberry

ASTM INTERNATIONAL
Standards Worldwide

STP 1450

Probabilistic Aspects of Life Prediction

W. Steven Johnson and Ben M. Hillberry, editors

ASTM Stock Number: STP1450

ASTM International
100 Barr Harbor Drive
PO Box C700
West Conshohocken, PA 19428-2959

Printed in the U.S.A.

Library of Congress Cataloging-in-Publication Data

(To come)

Photocopy Rights

Peer Review Policy

Each paper published in this volume was evaluated by two peer reviewers and at least one editor. The authors addressed all of the reviewers' comments to the satisfaction of both the technical editor(s) and the ASTM International Committee on Publications.

To make technical information available as quickly as possible, the peer-reviewed papers in this publication were prepared "camera-ready" as submitted by the authors.

The quality of the papers in this publication reflects not only the obvious efforts of the authors and the technical editor(s), but also the work of the peer reviewers. In keeping with long-standing publication practices, ASTM International maintains the anonymity of the peer reviewers. The ASTM International Committee on Publications acknowledges with appreciation their dedication and contribution of time and effort on behalf of ASTM International.

Printed in Lancaster, PA
October 2004

Foreword

The Symposium on *Probabilistic Aspects of Life Prediction* was held in Miami, FL on 6–7 November 2002. ASTM International Committee E8 on Fatigue and Fracture served as sponsor. Symposium chairmen and co-editors of this publication were W. Steven Johnson, Georgia Institute of Technology, Atlanta, GA and Ben Hillberry, Purdue University, West Lafayette, IN.

Overview

As fatigue and fracture mechanics approaches are used more often for determining the useful life and/or inspection intervals for complex structures, realization sets in that all factors are not well known or characterized. Indeed, inherent scatter exists in initial material quality and in material performance. Furthermore, projections of component usage in determination of applied stresses are inexact at best and are subject to much discrepancy between projected and actual usage. Even the models for predicting life contain inherent sources of error based on assumptions and/or empirically fitted parameters. All of these factors need to be accounted for to determine a distribution of potential lives based on a combination of the aforementioned variables, as well as other factors. The purpose of this symposium was to create a forum for assessment of the state-of-the-art in incorporating these uncertainties and inherent scatter into systematic probabilistic methods for conducting life assessment.

This is not the first ASTM symposium on this subject. On 19 October 1981 ASTM Committees E9 on Fatigue and E24 on Fracture Testing (today they are combined into Committee E8 an Fatigue and Fracture) jointly sponsored a symposium in St. Louis, MO. The symposium resulted in an ASTM STP 798, "*Probabilistic Fracture Mechanics and Fatigue Methods: Applications for Structural Design and Maintenance.*" The STP contained 11 papers. Both of the editors of this current STP were present. At that time, we were very involved with deterministic crack growth predictions under spectrum loading, trying to be as accurate as possible. We had little use for the statistics and probability. One thing that stood out in my listening to the speakers was the level of probability that they were predicting using the ASME boiler and pressure vessel code (author was G. M. Jouris). Some of their estimated probabilities of failure were on the order of 1×10^{-11}. A member of the audience noted that the inverse of this number was greater than the number of atoms in the universe. The audience laughed.

As time went by, a greater appreciation was developed for all the uncertainties in real world applications (as opposed to a more controlled laboratory testing environment). This confounded by needs to assure safety, avoid costly litigation suits, set meaningful inspection intervals, and establish economic risks, have brought more emphasis to the need to use probability in the lifing of components. Since the aforementioned symposium was almost 20 years ago, ASTM Committee E8 agreed to sponsor this symposium. The response was outstanding.

On 6–7 November 2002, in Miami, FL, 29 presentations were given. Lively discussions followed essentially all the talks. The presentations collectively did a great job on assessing the current state of the art in probabilisitc fatigue life prediction methodology. We would like to take this opportunity to recognize and thank our session chairs: Dr. Christos Chamis, Dr. Duncan Shepherd, Dr. James Larsen, Prof. Wole Soboyejo, Mr. Shelby Highsmith, Jr., Dr. Fred Holland, and Mr. Bill Abbott. A special thanks to Dr. Chamis for organizing a session.

Due to a number of factors, including paper attrition and a tough peer review process, only 17 papers have made it through the process to be included in this Special Technical Publication. The 17 papers have been divided into three topical groups for presentation in this publication: four papers are

in the section on *Probabilistic Modeling*; seven papers are in the section on *Material Variability*; and six papers are in the section on *Applications*.

We sincerely hope that you find this publication useful and that it helps make the world a safer place.

Prof. W. Steven Johnson
School of Materials Science and Engineering
George W. Woodruff School of Mechanical Engineering
Georgia Institute of Technology
Atlanta, GA

Prof. Ben M. Hillberry
School of Mechanical Engineering
Purdue University
West Lafayette, IN

Contents

Overview vii

SECTION I: PROBABILISTIC MODELING

Probabilistic Life Prediction *Isn't* as Easy as It Looks—C. ANNIS 3

Probabilistic Fatigue: Computational Simuation—C. C. CHAMIS AND S. S. PAI 15

The Prediction of Fatigue Life Distributions from the Analysis of Plain Specimen Data—D. P. SHEPHERD 30

Modeling Variability in Service Loading Spectra—D. F. SOCIE AND M. A. POMPETZKI 46

SECTION II: MATERIAL VARIABILITY

Probabilistic Fracture Toughness and Fatigue Crack Growth Estimation Resulting From Material Uncertainties—B. FARAHMAND AND F. ABDI 61

Predicting Fatigue Life Under Spectrum Loading in 2024-T3 Aluminum Using a Measured Initial Flaw Size Distribution—E. A. DEBARTOLO AND B. M. HILLBERRY 75

Extension of a Microstructure-Based Fatigue Crack Growth Model for Predicting Fatigue Life Variability—M. P. ENRIGHT AND K. S. CHAN 87

Scatter in Fatigue Crack Growth Rate in a Directionally Solidified Nickel-Base Superalloy—S. HIGHSMITH, JR. AND W. S. JOHNSON 104

Mechanism-Based Variability in Fatigue Life of Ti-6A1-2Sn-4Zr-6Mo—S. K. JHA, J. M. LARSEN, A. H. ROSENBERGER, AND G. A. HARTMAN 116

Predicting the Reliability of Ceramics Under Transient Loads and Temperatures with CARES/Life—N. N. NEMETH, O. M. JADAAN, T. PALFI, AND E. H. BAKER 128

Fatigue Life Variability Prediction Based on Crack Forming Inclusions in a High Strength Alloy Steel—P. S. SHARPE, B. M. HILLBERRY, AND B. A. CRAIG 150

SECTION III: APPLICATIONS

Preliminary Results of the United States Nuclear Regulatory Commissions Pressurized Thermal Shock Rule Reevaluation Project—T. L. DICKSON, P. T. WILLIAMS, B. R. BASS, AND M. T. KIRK 167

Corrosion Risk Assessment of Aircraft Structures—M. LIAO AND J. P. KOMOROWSKI 183

A Software Framework for Probabilistic Fatigue Life Assessment of Gas Turbine Engine Rotors—R. CRAIG MCCLUNG, M. P. ENRIGHT, H. R. MILLWATER, G. R. LEVERANT, AND S. J. HUDAK, JR. 199

Application of Probabilistic Fracture Mechanics in Structural Design of Magnet Components Parts Operating Under Cyclic Loads at Cryogenic Temperatures—M. YATOMI, A. NYILAS, A. PORTONE, C. SBORCHIA, N. MITCHELL, AND K. NIKBIN 216

A Methodology for Assessing Fatigue Crack Growth in Reliability of Railroad Tank Cars—W. ZHAO, M. A. SUTTON, AND J. PENÃ 240

Effect of Individual Component Life Distribution on Engine Life Prediction—E. V. ZARETSKY, R. C. HENDRICKS, AND S. M. SODITUS 255

Author Index 273

Subject Index 275

PROBABILISTIC MODELING

Journal of ASTM International, Feb. 2004, Vol. 1, No. 2
Paper ID JAI11557
Available online at: www.astm.org

Charles Annis [1]

Probabilistic Life Prediction *Isn't* as Easy as It Looks

ABSTRACT: Many engineers effect "probabilistic life prediction" by replacing constants with probability distributions and carefully modeling the physical relationships among the parameters. Surprisingly, the *statistical* relationships among the "constants" are often given short shrift, if not ignored altogether. Few recognize that while this simple substitution of distributions for constants will indeed produce a nondeterministic result, the corresponding "probabilities" are often woefully inaccurate. In fact, even the "trend" can be wrong, so these results can't even be used for sensitivity studies. This paper explores the familiar Paris equation relating crack growth rate and applied stress intensity to illustrate many statistical realities that are often ignored by otherwise careful engineers. Although the examples are Monte Carlo, the lessons also apply to other methods of probabilistic life prediction, including FORM/SORM (First/Second Order Reliability Method) and related "fast probability integration" methods.

KEYWORDS: life prediction, crack growth, Paris equation, probability, statistics, simulation, Monte Carlo, nondeterministic, probabilistic, joint, conditional, marginal, multivariate

There is more to probabilistic life prediction than replacing constants with probability densities. The purpose of this study is to demonstrate this by comparing the observed distribution of lives of 68 nominally identical crack growth specimens with Monte Carlo (MC) simulations of lives based on the distributions of their Paris law parameters. It will be shown that several common MC sampling techniques produce wildly inaccurate results, one with a standard deviation that is *7X* larger than was exhibited by the specimen lives themselves. The cause of such aberrant behavior is explained. It is further observed that the Paris law parameters are jointly distributed as bivariate normal, and a Monte Carlo simulation using this joint density reproduces the specimen mean and standard deviation to within a few percent. The lessons here apply to *any* regression model, not just to these data, nor only to crack growth rate models, nor are they limited only to MC.

The Data

In the mid-1970s Dennis Virkler, then a Ph.D. student of Professor Ben Hillberry at Purdue, conducted 68 crack growth tests of 2024-T3 aluminum [1,2]. These tests were unusual for several reasons. They were conducted expressly to observe random behavior in fatigue. While almost all crack growth tests measure crack length after some number of cycles, Virkler measured cycle count at 164 specific crack lengths. This provided a direct measure of variability in cycles, rather than the usually observed variability in crack length at arbitrary cyclic intervals. While two of the specimens appear to stand out from their brethren, the purpose of this investigation is not to play Monday Morning

Manuscript received Aug. 29 2002; accepted for publication Aug. 29 2003; published February 2004.
Presented at ASTM Symposium on Probabilistic Aspects of Life Prediction on Nov. 6, 2002 in Miami Beach, FL; W. S. Johnson and B. Hillberry, Guest Editors.

[1] Principal, Charles Annis, P.E., Statistical Engineering, Palm Beach Gardens, FL 33418-7161. Charles.Annis@StatisticalEngineering.com

Quarterback 25 years after the game, and there is no reason not to consider all 68 specimens here. In any event their exclusion changes only the numeric details. The fundamental results are not affected, nor are they affected by using a normal, rather than lognormal density to describe them.

It is common practice to fit a single *da/dN vs.* ΔK curve through multiple specimens of the same material tested under the same conditions of temperature, stress ratio, and frequency. In the study reported here, however, 68 individual Paris models were used. Fitting a single curve describes the mean trend behavior very well, but it obscures random specimen-to-specimen differences. Since real applications are subjected to similar randomness, it is necessary to capture that effect as well.

Fatigue Lives Are Lognormal

It has been long recognized that fatigue lives are satisfactorily modeled using the lognormal density. For these 68 specimens that model is less than optimal and there is some evidence that the probability density may be a mixture of two densities. It is not the purpose of this paper to repeat the earlier work by Virkler, Hillberry and Goel [2], and as it turns out, the actual form of the distribution of the specimen lives themselves only influences the numeric details of this study, since each specimen's crack growth rate curve was treated individually. (Treating the data as normal, however, results in a bias in the simulated mean of about 5%. The bias using the lognormal is negligible.)

Conventional Monte Carlo Simulation

Unlike many engineering analytical results, probability estimates are difficult to verify experimentally. This unfortunate reality has perpetuated the misuse of a valid statistical tool, and the consequences may not be apparent for years to come.

Most engineering Monte Carlo simulations are performed this way.

1. Set up a conventional deterministic analysis;
2. Replace constants with probability distributions;
3. Sample once from each distribution;
4. Compute the deterministic result and store the answer;
5. Repeat steps 3 and 4 many times;
6. Compute the mean and standard deviation of the collected results.

Sadly, many engineers are unfamiliar with the implicit *statistical* assumptions that are at the foundation of Monte Carlo simulation, but as been observed elsewhere [3] "Simply not understanding the nature of the assumptions being made does not mean that they do not exist."

What possibly could be wrong with this paradigm? Luckily we (the engineering community) have a dataset that is nearly perfect for answering this question, *viz.* the data collected by Virkler and Hillberry, as part of Virkler's Ph.D. dissertation. Professor Hillberry graciously made these available for further study.

Monte Carlo Modeling Specifics

After fitting individual Paris equations to each of the 68 specimens, the mean and standard deviation for the individual Paris parameters, intercept, C, and slope, n, were computed. The well-known Paris model for fatigue crack growth is given in equation 1

$$da/dN = 10^{C}(\Delta K)^{n} \tag{1}$$

where *da/dN* is the crack growth rate, in mm per cycle, and ΔK is the applied stress intensity factor, in MPa$\sqrt{m}$, given by equation 2.

$$\Delta K = \Delta\sigma\sqrt{\pi a}\, f(a \mid geometry) \tag{2}$$

Here, $\Delta\sigma$ is the testing stress range, $\sigma_{max} - \sigma_{min}$, a is the crack length, and $f(\,)$ is a function of the specimen (or component) geometry and crack length. Of course, when equation 1 is plotted on a *log-log* grid this is a straight line with intercept C and slope n. Assuming for the sake of simplicity that there was no variation in the starting crack size, the final crack size, or the test stress, the calculated cyclic lifetime can be computed from the individual Paris fits using equation 3.

$$\begin{aligned} da/dN &= 10^{C}\left[\Delta\sigma\sqrt{\pi a}\, f(a \mid geometry)\right]^{n} \\ dN &= da/\left\{10^{C}\left[\Delta\sigma\sqrt{\pi a}\, f(a \mid geometry)\right]^{n}\right\} \\ N &= \int_{a_0}^{a_{final}} 10^{-C}\left[\Delta\sigma\sqrt{\pi a}\, f(a \mid geometry)\right]^{-n} da \end{aligned} \tag{3}$$

In practice this integration is usually carried out numerically.

To conduct the usual MC simulation N_i is computed from $h(C_i\,,\ n_i)$ where $h(\,)$ is equation 3, and i ranges from 1 to say 1000 (or 10 000).

Many MC practitioners then calculate a mean and standard deviation for N, or $log_{10}(N)$, report the results and stop there, since there is nothing against which to compare the distribution of computed values for N_i. Virkler's data show the observed distribution of actual specimen lives and thus provide a direct comparison for these calculations.

The Paris Law is Adequate

Before going further it is prudent to check the goodness-of-fit of the Paris equation itself. If the underlying model for crack growth rate is inadequate there is little hope for accurate life prediction based on it. The sigmoidal shape of the *da/dN* ΔK data (Fig. 1) suggests a model such as the SINH [*4*] might do a better job than the straight line Paris model (and it does, increasing the ratio of standard deviations of calculated lives, 0.918 for Paris, to 0.957 for the SINH by reducing the disagreement between calculated and observed specimen lives from 8.2% to 4.3%). The added model complexity, however, obscures the real issue here, namely the abysmal performance of a rather common Monte Carlo simulation (700% error in predicted scatter). Since the Paris law is adequate it is used here for simplicity.

A Note on Modeling

Statisticians often assess the efficacy of a mathematical model by decomposing the sums-of-squares of differences between the model and the observations. We, however, are less interested in the differences between the measured crack growth rates, da/dN_i, and their Paris model, than we are in their integrated collective behavior, as given by equation 3. Such an integrated metric summarizes all sources of "error" - material variability, lack-of-fit, testing uncertainties – into the difference between the observed specimen life, and that provided by equation 3. We thus have traded the potential for better arithmetic diagnostics (scrutiny of the Paris model) for a more direct measure of what we are really interested in - life prediction performance.

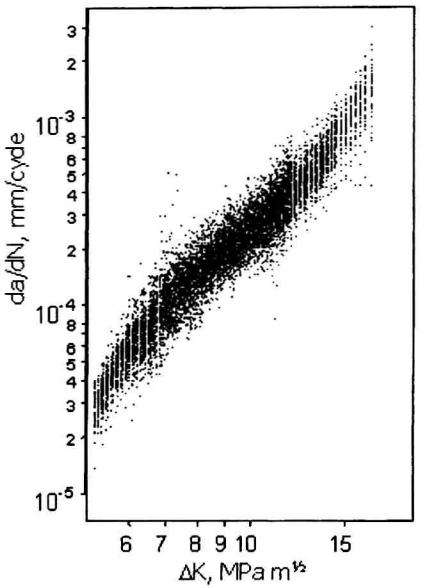

FIG.1—*da/dN vs. DK are S-shaped.*

How Well Does the Conventional Monte Carlo Algorithm Perform?

The conventional MC simulation of 1000 samples, with independent model parameters, *C* and *n*, did an acceptable job predicting the *mean* lifetime, after the *log* transform. Because the data are skewed to the right, as all fatigue data are, the untransformed simulated results overestimate means of the symmetrical normal models slightly.

The simulated *standard deviations* were another matter: The actual observed standard deviation for 68 specimens is 0.03015 log_{10} units (18 447 cycles)[2]. The conventional MC simulation of 1000 samples, with independent model parameters, *C* and *n*, produced a standard deviation of 0.19778 log_{10} units (140 261 cycles), *6.6X too large!*

A closer look shows the situation gets even worse. To be fair, the best possible Paris model would use the 68 individual Paris fits, since no simulation could be expected to be better than the actual specimens' behavior. Using the 68 Paris equations in equation 3 produces a standard deviation of 0.02769 log_{10} units (16 332 cycles), which is *smaller* than the observed standard deviation by about 8%. Why?

Of the 68 specimens, two seemed to exhibit longer lives than what might have been inferred by from the behavior of the other 66. All 68 specimens were used here. Since the actual specimen life doesn't directly influence its *da/dN vs. ΔK* behavior, predicted lives based on these two Paris fits would be more like their sister specimens, resulting in the smaller standard deviation for the integrated Paris equations. So to provide a fair

[2] The analyses were carried out using log_{10}*(cycles)*, and again using untransformed cycles. The reported log_{10} result can not, of course, be determined simply by taking the *log* of the mean and standard deviation of the untransformed results. All calculations are summarized in Tables 1 and 2 and Fig. 5.

comparison with simulated Paris models, the behavior of the 68 integrated Paris laws should be the baseline. Thus the baseline scatter is 0.02769 log_{10} units.

Comparing the simulation's standard deviation of 0.19778 log_{10} units with the integrated Paris law baseline shows the simulation to have overestimated the scatter by 0.19778 / 0.02769 or about *7.1X*. This is awful. Such a simulation would be worse than useless since it would likely compel a costly redesign. Put in perspective, the probability of failure before about 207 000 cycles is 0.1%, determined from the mean and standard deviation of the 68 specimens' (*log*-transformed) lives. The MC simulation puts this failure rate at about 33%, an overestimation of failure rate of over *300X*.

This absurd simulation result has been observed by every engineer who has performed similar MC simulations, since it doesn't require any statistics to detect an answer that is wrong by a factor approaching an order of magnitude in standard deviation. Sadly the most common palliatives proposed as remedies do not perform much better.

What Went Wrong?

The model parameters, C and n, are assumed to be normally distributed. Is this a good assumption in this case?

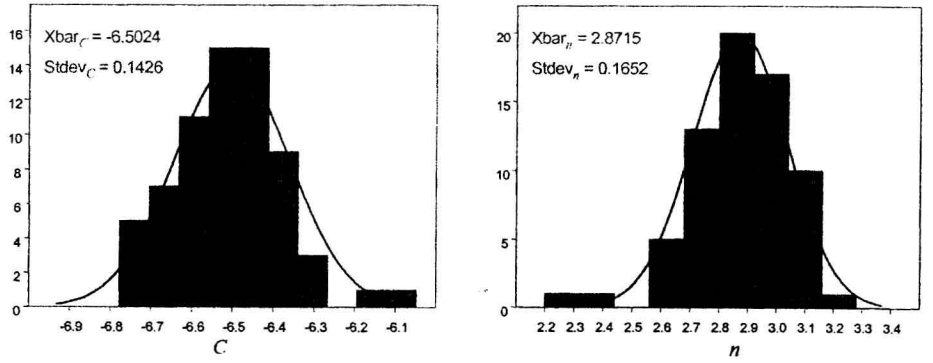

FIG. 2—*Histograms of Paris Model Parameters C and n.*

Figure 2 presents histograms of both model parameters. While somewhat approximate, the normal density is not an altogether improper model; surely these departures from the normal could not have caused the *7X* inflation of the standard deviation. A closer look at the figures provides a clue. There are two observations that are *high* for parameter C, and two that are *low* for parameter n. Perhaps these should be considered as pairs, rather than as independent observations. Figure 3, a schematic plot of crack growth rate *vs.* stress intensity on a *log-log* grid, shows why C and n behave in tandem: when the slope, n, is shallow the intercept, C, must be larger for the resulting line to go through the data. Similarly, a steeper slope requires a smaller intercept.

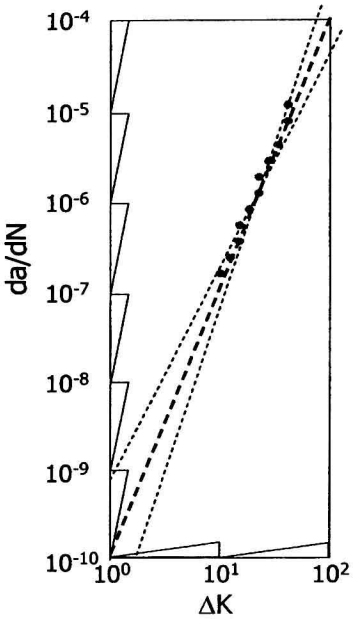

FIG. 3—*Schematic showing why Paris Parameters must be correlated. Note that in this schematic the intercept is C = log10(da/dN)= -10, at log10(DK)=0.*

Possible Remedies (All of Them Wrong)

Assuming C and n to be independent, when they obviously are not (the most common error in Monte Carlo modeling), results in unacceptable error in simulated lifetime scatter. Possible remedies that have been suggested are:

1. n assumed fixed, C is normal
2. C assumed fixed, n is normal
3. C assumed a linear function of n.

Fixing either n or C seems at first blush like a reasonable solution, and it does reduce the over-prediction of scatter from *7.1X* to *5.1X* (n fixed) or *5.4X* (C fixed). While this is an obvious improvement, the error remains wildly unacceptable. Sadly, it is at this stage when the standard deviation of C or n is arbitrarily "adjusted," *i.e.*, fudged until a believable result is achieved.

Figure 4 also shows why assuming either C or n as fixed is not reasonable. The horizontal line is at $n = 2.87$, the average of 68 Paris slopes. This is a reasonable value only when $-6.58 < C < -6.45$. When C is outside this range, as it will be often, the resulting simulated combination is very, *very* improbable. In fact observations in either the first or third quadrants (large n with large C, or small n with small C) are exceedingly unlikely in reality but occur about half the time in uncorrelated simulation.

Another option for remedy suggests itself since the two parameters are obviously so closely related: let one be a function of the other. A linear fit of $C=b_1 + b_2 n$, with n being sampled from a normal density, does indeed improve things. But this time the resulting error ratio is 0.51, *i.e.*: the scatter has been *over*-corrected, and now is *underestimated* by almost half. Clearly this nonconservative result is also unacceptable.

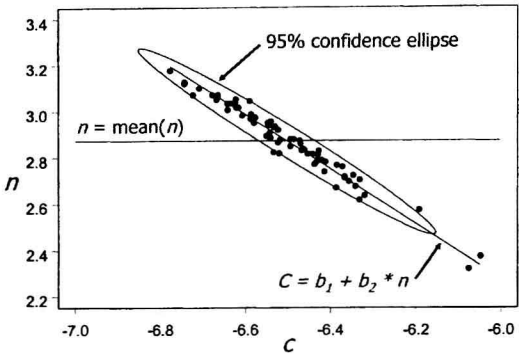

FIG. 4—*Paris Parameters C and n are obviously correlated (r=0.982).*

To understand why such an appealing suggestion should have such an undesirable result, look again at Fig. 4 which also shows the 95% confidence ellipse for the *C* and *n* pairs. Assuming that one is a linear function of the other, in effect collapses this ellipse into a line, thus *underestimating* the overall variability. (The confidence ellipse also suggests that two of the tests may be different from the others, as was noted earlier.)

The Right Way

We have considered four very common oversights in Monte Carlo modeling. So, how do you do it correctly?

Parameters estimates for *C* and *n* are jointly distributed. (Notice that this is not optional. It is how regression model parameters naturally behave. You can't choose the ratio of a circle's circumference to its diameter to be an integer because it might be more convenient. The fact is that π is inconveniently transcendental. Similarly, regression parameter estimates are asymptotically multivariate normal, and correlated, so any realistic simulation must sample from their correlated joint density.) Modeling them as bivariate normal in a MC simulation produces a standard deviation of 0.02802 in *log10* integrated lifetime for 1000 samples, which is very close to the standard deviation of the integrated individual Paris fits, 0.02769. The ratio of standard deviations is 1.012. *In other words, correctly modeling the joint behavior reduces the greater than 700% error in the estimate of the standard deviation to about 1%.*

Notice, too, that replacing a constant *n* (the horizontal line in Fig. 4) with a (conditional) probability density has the paradoxical effect of *decreasing* the resulting variability in calculated lifetime, since it corrects Mistake #2 (see Tables 1 and 2). This refutes the common misconception that replacing a constant with a probability density in a Monte Carlo simulation always results in increased scatter in the output. All these results are summarized in Tables 1 and 2 and in Fig. 5.

TABLE 1—*MC Results Assuming Cycles are LogNormally Distributed.*

			Correct	Mistake #1	Mistake #2	Mistake #3	Mistake #4
	Actual N	Eqn 3 N	C, n joint	C, n indept.	n fixed	C fixed	$C=b_0+b_1*n$
mean	5.40916	5.39773	5.39909	5.41404	5.39414	5.39911	5.42217
stdev	0.03015	0.02769	0.02802	0.19778	0.14084	0.14872	0.01426