Parameter Estimation for a Specific Software Reliability Model

Peter Spreij
Centre for Mathematics and Computer Science, Amsterdam

Key Words—**Software reliability, Parameter estimation, *s*-Confidence interval, Jelinski-Moranda model.**

Reader Aids—
Purpose: Widen state of the art
Special math needed for explanations: Advanced probability, Statistics
Special math needed to use results: Statistics
Useful to: Software reliability theoreticians

Summary & Conclusions—**The problem of maximum likelihood estimation in the Jelinski-Moranda software reliability model is studied. The distribution of the stochastic variable that completely determines the maximum likelihood estimate is obtained. *s*-Confidence intervals for the parameter of interest can then be constructed by using the same stochastic variable. An example is given using real data.**

1. INTRODUCTION

Various models are used to evaluate the reliability of complex computer programs. Along with the programs, the models also vary in complexity. One of the oldest models, originally proposed by Jelinski & Moranda [1], is frequently used. It is extremely simple and conclusions inferred from it have only a limited applicability. Some drawbacks of this model have been pointed out by Forman & Singpurwalla [2] and by Littlewood [3].

However, for reasons of simplicity, it remains attractive and as long as one keeps the limitations in mind, this model can be a useful tool in obtaining some basic ideas about the reliability of a program. An alternative approach that motivates the attraction of the Jelinski-Moranda model via the theory of "shock models and wear processes" is given by Langberg & Singpurwalla [10].

Although numerous models have been proposed in the literature to study the reliability of a program, very little attention has been paid to the important question of estimating parameters in the models.

This paper discusses the estimation of the initial error content of a program and gives a method to construct one sided *s*-confidence intervals for this parameter.
A numerical example involving real data is given.

Notation

n_t	number of observed failures in $[0, t]$
ϕ	failure rate (per error)
N	initial error content of the program
λ_t	failure rate of the program at time t, or intensity of the process n_t
$(N)_n$	$\prod_{i=1}^{n} (N - i + 1)$
t_i	random time elapsed between observation of failures i and $i - 1$
$[x]$	$\max\{n \in \mathbf{Z}: n \leq x\}$
χ_A	indicator function of the set A: $\chi_A(x) = \begin{cases} 1, \text{if } x \in A \\ 0, \text{otherwise} \end{cases}$
f^{k*}	k-fold convolution of the function f
ζ	$\sum_{i=1}^{n} (i - 1)t_i / \sum_{i=1}^{n} t_i$
gauf	Cdf of a standard normal random variable
X_i	random time until error i causes a failure
p_x	pdf of the random variable X

Other, standard notation is given in "Information for Readers & Authors" at rear of each issue.

Assumptions

A1. Errors behave *s*-independently, that is, the X_i's are *s*-independent random variables.

A2. X_i has an exponential distribution with rate ϕ.

A3. The failure rate λ_t is proportional with the residual number of errors in the program: $\lambda_t = \phi(N - n_t)$.

A4. The parameters ϕ and N are unknown constants

2. DESCRIPTION OF THE MODEL

Following the assumptions of Jelinski & Moranda, consider the failure process n_t (ie, the total number of observed failures in the time interval $[0, t]$) of a program as a self-exciting Poisson process which means that the intensity λ_t of the process depends in general on n_s, $s \in [0, t]$. For this model, $\lambda_t = \phi(N - n_t)$.

Suppose we observe n_t for $t \in [0, T]$. If one wants to estimate the unknown parameters N and ϕ using the method of maximum likelihood (*ML*), one has to maximize the likelihood function which, in this context, has the form [4]:

$$L = \exp(-\int_0^T \lambda_s \, ds + \int_0^T \log \lambda_s \, dn_s), \tag{2.1}$$

$$\lambda_{s-} = \lim_{s \uparrow u} \{\lambda_u\}.$$

0018-9529/85/1000-0323$01.00©1985 IEEE

However, one rarely observes the process n_t itself, but only the successive interfailure times $t_1, t_2, \ldots$. Suppose one wants an estimate of N after observation of the first n interfailure times $t_1, \ldots, t_n$, then one maximizes, instead of (2.1), a "stopped version" of the likelihood function, that is (2.1) evaluated at $T = \sum_1^n t_j$, which becomes, after evaluating the integrals:

$$L(N, \phi; t_1, \ldots, t_n) = \phi^n (N)_n \exp(-\phi \sum_{i=1}^{n} (N - i + 1)t_i), \quad (2.2)$$

For better understanding of the model it can be useful to look at the situation in another way. Consider the N random variables X_i and let $\tilde{X}_i$ be order statistic i from the sample $X = (X_1, \ldots, X_N)$, $\tilde{X} = (\tilde{X}_1, \ldots, \tilde{X}_N)$. Then one can compute the pdf of $(t_1, \ldots, t_N)$ as follows

$$P_{t_1,\ldots,t_N}(t_1, \ldots, t_N) = p_{\tilde{X}}(t_1, t_1 + t_2, \ldots, t_1 + \ldots + t_N)$$

$$= N!\, p_X(t_1, t_1 + t_2, \ldots, t_1 + \ldots + t_N)$$

$$= N! \prod_{j=1}^{N} [\phi \exp(-\phi \sum_{i=1}^{j} t_i)]$$

$$= N! \phi^N \exp\left(-\phi \sum_{j=1}^{N} (N - j + 1)t_j \right). \quad (2.3)$$

By integrating over $t_{n+1}, \ldots, t_N$ one obtains the pdf of $(t_1, \ldots, t_n)$

$$p_{t_1,\ldots,t_n}(t_1, \ldots, t_n) = \prod_{j=1}^{n} p^j(t_j),\ p^j(t_j) = \phi(N - j + 1)$$

$$\exp(-(N - j + 1)t_j) \quad (2.4)$$

Hence the t_j's are s-independent [5]. The conclusion is that the t_j's are indeed the interfailure times as they appear in the Jelinski-Moranda model. Thus (2.4) yields again (2.2).

From a Bayes point of view the parameters N and ϕ should be treated as random variables with a certain prior distribution. In this case the above derivation still holds given N and ϕ, if one replaces the s-independence of the X_i's and the t_i's by conditional s-independence given N and ϕ. An analysis of the implications of a Bayes viewpoint concerning the Jelinski-Moranda model can be found in [9-11].

3. ESTIMATION

In order to estimate the unknown parameters N and ϕ, note that the parameter space for (N, ϕ) is $\mathbf{N} \times \mathbf{R}_+$. However in obtaining ML estimates, N is treated as a continuous parameter. Let $(\hat{N}, \hat{\phi})$ denote the value of (N, ϕ) that maximizes $L(\cdot, \cdot; t_1, \ldots, t_n)$ as a function from $\mathbf{R}_+^2$ to $\mathbf{R}$; then the true ML estimate of N will be $[\hat{N}]$ or $[\hat{N}] + 1$, depending on which of the two yields the largest value of (2.2). To derive ML estimates for N and ϕ we solve (if possible) the equations $\frac{\partial L}{\partial \phi} = 0$, $\frac{\partial L}{\partial \mathrm{N}} = 0$. N is our main parameter of interest and its ML estimate is the solution of:

$$\sum_1^n \frac{1}{N - i + 1} - \frac{n}{N - \zeta} = 0,\ \zeta = \sum_1^n (i - 1)t_i / \sum_1^n t_i. \quad (3.1)$$

In [6] it is proved that (3.1) has a finite solution for N if and only if $\zeta > \frac{1}{2}(n - 1)$.

Eq. (3.1) shows that the solution $\hat{N}$ is completely determined by the value of ζ, and the statistical properties of $\hat{N}$ follow from those of ζ. An immediate problem is therefore to find the distribution of ζ. This distribution has been empirically investigated by Forman & Singpurwalla [2].

Theorem 3.1

The ζ is distributed with pdf $f_\zeta(z; N)$ w.r.t. Lebesque measure with support $[0, n - 1]$:

$$f_\zeta(z; N) = \frac{(N)_n}{(N - z)^n} \cdot \chi_{[0,1]}^{(n-1)^*}(z), \quad (3.2)$$

where $\chi_{[0,1]}^{(n-1)^*}$ is the $(n - 1)$-folded convolution of the indicator function of $[0, 1]$ with itself. Calculation of the convolution yields:

$$f_\zeta(z; N) = \frac{(N)_n}{(N - z)^n} \cdot \frac{1}{(n - 1)!} \sum_{j=0}^{[z]} \binom{n-1}{j}$$

$$(-1)^j (z - j)^{n-2}. \quad (3.3)$$

Theorem 3.1 is proved in the appendix.

Corollary 3.2

The Cdf $\{\zeta\}$ is:

$$F_\zeta(z; N) = \frac{(N)_n}{(N - z)^{n-1}} \cdot \frac{1}{(n - 1)!} \sum_{j=0}^{[z]} \binom{n-1}{j}$$

$$\frac{(-1)^j}{N - j} (z - j)^{n-1} \quad (3.4)$$

Application

Now we are able to compute the probability that when n interfailure times are observed the ML estimate of N becomes infinite. This probability is $F_\zeta(\frac{1}{2}(n - 1); N)$.

When $n \ll N$ then $F_\zeta(½(n-1); N) \approx F_\zeta(½(n-1; \infty) = ½$. This means, that for large systems, which presumably contain many errors, we anticipate, in the early phase of testing, infinite estimates of N in 50% of the cases.

Observe that the family of pdf's $\{f_\zeta(z; N)\}_N$ has (decreasing) monotone likelihood ratio, since:

$$f_\zeta \frac{(z; N+m)}{f_\zeta(z; N)} = \frac{(N+m)_n}{(N)_n} \left(1 - \frac{1}{N+m-z}\right)^n \quad (3.5)$$

which is a decreasing function of z. The following proposition can then be proved [8].

Proposition 3.3

If we express the dependence of ζ on N by writing ζ_N instead of ζ, then the sequence of random variables $\{\zeta_N\}_N$ is stochastically decreasing, ie, the sequence $\{F_\zeta(z; N)\}_N$ is increasing.

4. s-CONFIDENCE INTERVALS FOR N (1-SIDED)

Consider hypotheses of the form H_0: $N \geq N_0$. Suppose one tests such a hypothesis at s-significance level α and uses ζ as a test statistic. Then H_0 is rejected for large values of ζ, say $\zeta \geq c$, where $c = c(N_0, \alpha)$ is such that

$$\sup_{N \geq N_0} P_N(\zeta \geq c) = P_{N_0}(\zeta \geq c) = \alpha.$$

Suppose that $\zeta = z$ is observed and that H_0: $N \geq N_0$ is accepted, so $z \leq c$. Then one would also accept H_0': $N \geq N_0'$ with $N_0' < N_0$, because:

$$\sup_{N \geq N_0'} \mathrm{P}_N(\zeta \geq z) > P_{N_0}(\zeta \geq z) > P_{N_0}(\zeta \geq c) = \alpha$$

Here the first inequality follows from proposition 3.3. As a consequence of this, a $1-\alpha$ s-confidence interval for N based on $\zeta = z$ is $[n, \overline{N}]$ where —

$$\overline{N} = \sup\{N: F_\zeta(z; N) < 1 - \alpha\}. \quad (4.1)$$

It may well be possible that $\overline{N}$ is infinite. A necessary and sufficient condition for $\overline{N}$ to be finite is

$$F_\zeta(z; \infty) = \frac{1}{(n-1)!} \sum_{j=0}^{[z]} \binom{n-1}{j} (-1)^j (z-j)^{n-1} > 1 - \alpha \quad (4.2)$$

If one is interested only in finite s-confidence intervals, and $\overline{N}$ is not finite for a specific value of $1-\alpha$, given $\zeta = z$, then one can of course obtain a finite $\overline{N}$ by lowering the s-confidence level $1-\alpha$. The graph of the function $N \to F_\zeta(z; N)$ has a horizontal asymptote at level $F_\zeta(z; \infty)$. If $F_\zeta(z; \infty) > 1 - \alpha$ but very close to $1-\alpha$, then $\overline{N}$ is very sensitive to changes in the value of α. A minor increase (decrease) of $1-\alpha$ results in a disproportionately large increase (decrease) in the value of $\overline{N}$. Eq. (4.2) can be well approximated, provided n is not very small, by

$$\mathrm{gauf}\left(\frac{z - ½(n-1)}{\sqrt{1/12(n-1)}}\right) > 1 - \alpha, \quad (4.3)$$

because ζ_∞ is approximately s-normal with mean ½ $(n-1)$ and variance 1/12 $(n-1)$

5. CONTINUOUS-TIME PARAMETER ESTIMATION

The character of the estimation problem is completely changed if one has full knowledge of the process n_t over a given time interval [0, T]. Now one has to deal directly with (2.1). Again an equation can be written from which a *ML* estimate of N can be calculated.

Paralleling the procedure leading to (3.1) leads to:

$$\sum_{i=1}^{n_T} \frac{1}{N-i+1} - \frac{n_T}{N-\zeta_T} = 0, \; \zeta_T = \frac{1}{T}\int_0^T n_s ds. \quad (5.1)$$

In this case $\hat{N}$ is not determined by a single statistic but instead by (n_T, ζ_T), the distribution of which is hard to determine. Solving this problem requires an entirely different approach; I hope to treat it in another paper.

6. NUMERICAL EXAMPLE

The *J-M* model has been applied on data collected from an automization project at the Dutch Aerospace Laboratory (NLR), where the numerical computation has also been performed. The data are represented as consecutive interfailure times in table.

TABLE 1
Execution times between successive failures in seconds

880	3820	24300	14910
3430	14800	17500	14670
2860	1770	4450	16310
11760	24270	4860	38410
4750	4800	640	1120
240	470	3990	30560
2300	40	26840	6210
8570	10170	2270	120
4620	1120	200	20210
1060	980	39180	26400

Notation

$\hat{N}$ *ML* estimate of N
$\overline{N}$ upperbound of a *s*-confidence interval for N
n number of observed failures.

Work out the cases when n = 10, 14, 20, 30, 32, 34, 36, 38, 40 failures are detected. (Many of the other cases yield infinite *s*-confidence bounds.) This results in table 2:

TABLE 2
ML estimates and upperbounds for $1-\alpha$ *s*-confidence intervals for N after detection of n failures. Here "-" means that an infinite or a very large value was found.

n	$\hat{N}$	$\overline{N}$, $\alpha = 0.30$	$\overline{N}$, $\alpha = 0.05$
10	123	—	—
14	16	18	60
20	111	—	—
30	39	46	113
32	40	45	84
34	38	41	54
36	41	43	55
38	47	51	80
40	47	50	69

Figures 1 & 2 represent the above table among other results with the number of failures(n) and cumulative execution time, respectively, represented on the horizontal axis.

APPENDIX

Proof of theorem 3.1

First we will show that —

$$\frac{\partial}{\partial N} \log f_\zeta(z; N) = \sum_1^n \frac{1}{N - i + 1} - \frac{n}{N - z}. \quad \text{(A.1)}$$

Let $t = (t_1, \ldots, t_n) \in \mathbf{R}_+^n$ and let $f(t; N, \phi)$ denote the joint pdf of t: $f(t; N, \phi) = \phi^n (N)_n \exp(-\phi T(t) \cdot (N - z(t)))$, with $T(t) = \sum_1^n t_i$, $z(t) = \sum_1^n (i-1)t_i / \sum_1^n t_i$. The ζ does not depend on ϕ, for we can write $\zeta = \sum_1^n (i-1)\phi t_i / \sum_1^n \phi t_i$ where the ϕt_i's are *s*-independent random variables with exponential distribution determined by their respective means $1/(N - i + 1)$. The pdf of ζ is —

$$f_\zeta(z; N) = \frac{\partial}{\partial z} \int_{\zeta(t) \le z} \phi^n (N)_n e^{-\phi T(t) \cdot (N - \zeta(t))}\, dt. \quad \text{(A.2)}$$

Since $\frac{\partial}{\partial \phi} f_\zeta(z; N) = 0$ it follows from differentiating w.r.t. ϕ under the integral sign —

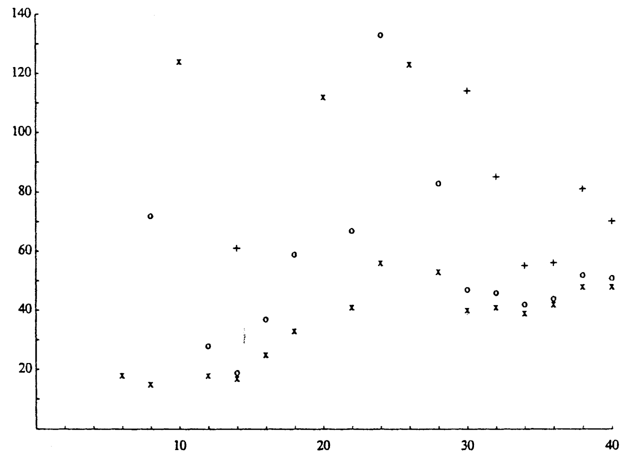

Fig. 1. ML estimates and *s*-confidence bounds for the parameter N in the JM model plotted against the number of observed failures. $\hat{N}$ is denoted by x, $\overline{N}$ for $\alpha = 0.30$ denoted by o and $\overline{N}$ for $\alpha = 0.05$ is denoted by $+$.

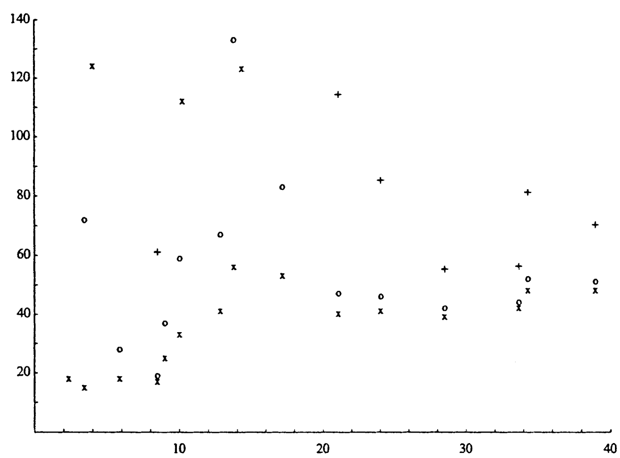

Fig. 2. ML estimates and *s*-confidence bounds for the parameter N in the JM model plotted against cumulative execution time $\times 10^4$ sec. at failure instants. $\hat{N}$ is denoted by x, $\overline{N}$ for $\alpha = 0.30$ denoted by o and $\overline{N}$ for $\alpha = 0.05$ is denoted by $+$.

$$\frac{\partial}{\partial z} \int_{\zeta(t) \le z} n e^{-\phi T(t) \cdot (N - \zeta(t))} dt = \frac{\partial}{\partial z} \int_{\zeta(t) \le z} \phi T(t)(N - \zeta(t)) e^{-\phi T(t)(N - \zeta(t))} dt \quad \text{(A.3)}$$

Consider now a coordinate transformation $t \to (\tilde{t}, \zeta)$, $\tilde{t} \in \mathbf{R}_+^{n-1}$ with Jacobi matrix J. Then (A.3) becomes:

$$\frac{\partial}{\partial z} \int_{\zeta \le z} \int_{\tilde{t}} n e^{-\phi T(\tilde{t}, \zeta)(N - \zeta)} |J(\tilde{t}, \zeta)| \, d\tilde{t} \, d\zeta = \frac{\partial}{\partial z} \int_{\zeta \le z} \int_{\tilde{t}} \phi T(\tilde{t}, \zeta)(N - \zeta) e^{-\phi T(\tilde{t}, \zeta)(N - \zeta)} |J(\tilde{t}, \zeta)| \, d\tilde{t} \, d\zeta. \quad \text{(A.4)}$$

Hence —

$$\int_{\tilde{t}} n e^{-\phi T(\tilde{t}, z)(N - z)} |J(\tilde{t}, z)| \, d\tilde{t} = \int_{\tilde{t}} \phi T(\tilde{t}, z)(N - z) e^{-\phi T(\tilde{t}, z)(N - z)} |J(\tilde{t}, z)| \, d\tilde{t}. \quad \text{(A.5)}$$

Now compute

$$\frac{\partial}{\partial N}\log f_\zeta(z;N) =$$

$$\frac{\frac{\partial}{\partial z}\int_{\zeta(t)\le z}\phi^n(N)_n\left[\sum_1^n \frac{1}{N-i+1} - \phi T(t)\right]e^{-\phi T(t)(N-\zeta(t))}\,dt}{\frac{\partial}{\partial z}\int_{\zeta(t)\le z}\phi^n(N)_n e^{-\phi T(t)(N-\zeta(t))}\,dt}$$

$$= \sum_1^n \frac{1}{N-i+1} - \frac{\frac{\partial}{\partial z}\int_{\zeta(t)\le z}\phi T(t)e^{-\phi T(t)(N-\zeta(t))}\,dt}{\frac{\partial}{\partial z}\int_{\zeta(t)\le z}e^{-\phi T(t)(N-\zeta(t))}\,dt}, \tag{A.6}$$

which becomes, by using the same transformation:

$$\sum_1^n \frac{1}{N-i+1} - \frac{\frac{\partial}{\partial z}\int_{\zeta\le z}\int_{\tilde t}\phi T(\tilde t,\zeta)e^{-\phi T(\tilde t,\zeta)(N-\zeta)}\,|J(\tilde t,\zeta)|\,d\tilde t d\zeta}{\frac{\partial}{\partial t}\int_{\zeta\le z}\int_{\tilde t}e^{-\phi T(\tilde t,\zeta)(N-\zeta)}\,|J(\tilde t,\zeta)|\,d\tilde t d\zeta}$$

$$= \sum_1^n \frac{1}{N-i+1} - \frac{\int_{\tilde t}\phi T(\tilde t,z)e^{-\phi T(\tilde t,z)(N-z)}\,|J(\tilde t,z)|\,d\tilde t}{\int_{\tilde t}e^{-\phi T(\tilde t,z)(N-z)}\,|J(\tilde t,z)|\,d\tilde t}, \tag{A.7}$$

Rewrite (A.7) by using (A.5) into:

$$\sum_1^n \frac{1}{N-i+1} - \frac{n}{N-z}. \tag{A.8}$$

From (A.1) it follows that $f_\zeta(z;N) = \frac{(N)_n}{(N-z)^n}\,h(z)$, where the function h does not depend on N.

From the invariance of the distribution of ζ with respect to ϕ it follows that we would have (A.1) again if the t_i's had s-expectation $1-(i-1)/N$ (take $\phi = 1/N$).

Let $N\to\infty$ and derive by means of a continuity argument (eg, dominated convergence applied to the calculation of f_ζ from $f(t;N,1/N)$) that —

$$h(z) = f_\zeta(z;\infty) = \frac{\partial}{\partial z}\int_{\zeta(t)\le z}e^{-T(t)}dt. \tag{A.9}$$

Stated otherwise, h is the pdf of $\zeta = \sum_1^n (i-1)\tau_i/\sum_1^n \tau_i$, where the τ_i's are i.i.d. with common exponential distribution (with mean 1). But then it is known [5, 7] that ζ is stochastically equal to $\sum_1^{n-1} U_i$, where the U_i are i.i.d. random variables with uniform distribution on [0, 1]. Hence $h = \chi_{[0,1]}^{(n-1)*}$. *Q.E.D.*

Remarks

The $\frac{\partial}{\partial N}\log f_\zeta = \frac{\partial}{\partial N}\log f(t;N,\phi(N,t))$, where $\phi(N,t)$ is the solution for ϕ of $\frac{\partial}{\partial\phi}f(t;N,\phi) = 0$. This is an appealing result, although not obvious, because ζ is not sufficient for N (in presence of the nuisance parameter ϕ).

REFERENCES

[1] Z. Jelinski, P. B. Moranda, "Software reliability research," in *Statistical Computer Performance Evaluation,* ed. W. Freiberger, Academic Press, 1972, pp 465-484.

[2] E. H. Forman, N.D. Singpurwalla, "An empirical stopping rule for debugging and testing computer software," *J. Amer. Statist. Assoc.,* vol 72, 1977 Dec, pp 750-757.

[3] B. Littlewood, "Theories of software reliability: How good are they and how can they be improved?," *IEEE Trans. Software Engineering,* vol SE-6, 1980 Sep, pp 489-500.

[4] D. L. Snyder, *Random Point Processes,* John Wiley & Sons, 1975.

[5] W. Feller, *An Introduction to Probability Theory and its Applications,* Vol II, 2nd ed. John Wiley & Sons, 1971.

[6] B. Littlewood, J. L. Verral, "Likelihood function of a debugging model for computer software reliability," *IEEE Trans. Reliability,* vol R-30, 1981 Jun, pp 145-148.

[7] R. E. Barlow, D. J. Bartholomew, J. M. Bremner, H. D. Brunk, *Statistical Inference under Order Restrictions,* John Wiley & Sons, 1978.

[8] E. L. Lehmann, *Testing Statical Hypotheses,* John Wiley & Sons, 1959.

[9] G. Koch, P. J. C. Spreij, "Software reliability as an application of martingale and filtering theory," *IEEE Trans. Reliability,* vol R-32, 1983 Oct, pp 342-345.

[10] N. Langberg, N. D. Singpurwalla, "Unification of some software reliability models via the Bayesian approach," to appear in *SIAM J.* Scientific & Statistical Computing.

[11] R. J. Meinhold, N. D. Singpurwalla, "Bayesian analysis of a commonly used model for describing software failures," *The Statistician,* vol 32, 1983, pp 168-173.

AUTHOR

Peter Spreij; Center for Mathematics and Computer Science; PO Box 4079; 1009 AB Amsterdam, NETHERLANDS.

Peter Spreij: For a biography see *IEEE Transactions on Reliability,* 1983 October, page 345.

Manuscript TR83-121 received 1983 August 6; revised 1985 August 2.

★★★

Book Review . . . . Ralph A. Evans, *Product Assurance Consultant*

IEEE Recommended Practice for Design of Reliable Industrial and Commercial Power Systems

(The Gold Book)
Power System Technologies Committee of the
IEEE Industry Applications Society
1980, 224 pages, *ca.* $20
Wiley-Interscience
LCCC 80-83819; ISBN 0-417-09261-4

Table of Contents

This is a very useful book for those who need: 1) this kind of data, and/or 2) the introduction to reliability concepts this book provides. The book itself is a pioneering effort and reports the results of some pioneering efforts.

Two of the excellent aspects of the book are:

1. The emphasis on planning, design, and corrective action.

2. The many examples and their illustrative numerics.

Both aspects will help engineers to learn that the reliability discipline is not a numbers game but is an important aid to traditional engineering tasks.

The treatment of probability and statistics is above average for an engineering book (but is deficient as explained below). Considering that the book is a pioneering work, the treatment is very good. Some areas for future improvement are:

1. The assumption of constant failure rate should be explicit.

2. The differences between statistical jargon and ordinary language should be explained better. For example, *expected* and *confidence* are used as statistical jargon, but *significance* is not. Three different words (expected value, mean, average) are all used to mean the same thing. In a technical treatise, synonyms can and should be given in a glossary, but in the text there should be a one-to-one correspondence between concepts and the names for those concepts.

3. FMEA (failure modes and effects analysis) should either be explained more conventionally or its differences from the conventional version should be explained.

4. The category of maintenance-induced failures should be introduced under preventive maintenance. Preventive maintenance should be used only where it is clear that more good than harm will come from actions done in its name.

5. The explanations and examples for *series* and *parallel* systems need more work. *Series* and *parallel* in this context refer to logic diagrams, not to schematics. Where there is more than one failure mode for a component then the concepts of *series* and *parallel* become more complicated.

For example, assume: a) a capacitor can fail only *short* or *open,* and b) a 2-capacitor bank can likewise fail only *short* or *open*. Two capacitors that are physically in parallel are in *series* for *shorts* (the 2-capacitor bank fails if either one *shorts*) and in *parallel* for *opens* (the 2-capacitor bank does not fail *open* unless both fail *open*).

6. The failure data for each class of component are not a random sample from a population that has a single, constant failure rate. There are several sub-populations, each with its own failure rate. Thus, the statistical treatment of confidence intervals is not appropriate. The explanations treat *accuracy* and *confidence* as synonomous; that is especially not true in the current treatment.

7. The MTBF (mean time between failures) is the reciprocal of the failure rate only when the anticipated lives are much greater than the reciprocal failure rate.

8. The problem of common-causes of failures is not treated adequately. For example, one violent electrical storm can damage several components; those failures are definitely not a random sample from a sub-population with a constant failure rate.

★★★