Reducing Over-Conservative Expert Failure Rate Estimates in the Presence of Limited Data: A New Probabilistic/Fuzzy Approach

Carlos Ferregut, F. Joshua Campos, and Vladik Kreinovich
Future Aerospace Science & Technology (FAST) Center
University of Texas at El Paso
500 W. University
El Paso, TX 79968, USA
ferregut@utep.edu, fjcampos@miners.utep.edu,
vladik@utep.edu

***Abstract*—Unique highly reliable components are typical for aerospace industry. For such components, due to their high reliability and uniqueness, we do not have enough empirical data to make statistically reliable estimates about their failure rate. To overcome this limitation, the empirical data is usually supplemented with expert estimates for the failure rate. The problem is that experts tend to be – especially in aerospace industry – over-cautious, over-conservative; their estimates for the failure rate are usually much higher than the actual observed failure rate. In this paper, we provide a new fuzzy-related statistically justified approach for reducing this over-estimation.**

I. Formulation of the Problem

Reliability: how it is usually described and evaluated. Failures are ubiquitous. As a result, reliability analysis is an important part of engineering design.

In reliability analysis of a complex system, it is important to know the reliability of its components; see, e.g., [1]. Reliability of a component is usually described by an *exponential model*, in which the probability $P(t)$ for a system to be intact by the time t is equal to $\exp(-\lambda \cdot t)$ for some constant λ. For this model, the average number of failures per unit time is equal to λ; as a result, this value is called a *failure rate*. Another important characteristic – mean time between failure (MTBF) θ – is, in this model, equal to $1/\lambda$.

Usually, the failure rate λ (or, equivalently, the MTBF θ) are determined by analyzing the records of actual failures. When we observe a sufficient number of failures, we can then take an arithmetic average of the observed times between failures – and this average is a statistically justified estimate for θ.

Reliability estimates in aerospace industry: a challenge. In aerospace industry, especially in designing spaceships for manned flights, reliability is extremely important. Because of this importance, aerospace systems use unique, highly reliable components.

This reliability, however, leads to a challenge: since the components are unique and highly reliable, we do not have enough failure records to make statistically reliable estimates about their failure rate: in most cases, we have up to 5 failures. This scarcity of data is especially critical on the stage when we are still designing a spaceship.

Need to use expert estimates. To overcome this limitation, the empirical data is usually supplemented with expert estimates for the failure rate.

Expert estimations are over-conservative: a problem. A problem with expert estimates is that experts tend to be – especially in aerospace industry – over-cautious, over-conservative. The experts' estimates for the failure rate are usually much higher than the actual observed failure rate.

What we do in this paper. In this paper, we propose an algorithm that reduces this over-conservativeness.

II. Available Data

We have n components. For each of these components $i = 1, \ldots, n$, we have n_i observed times-between-failures $t_{i1}, \ldots, t_{in_i}$. We also have expert estimates $e_1, \ldots, e_n$ for the failure rate of each component.

III. Traditional Engineering Approach to Reliability Estimation: a Brief Reminder

Before describing our new approach, let us briefly recall the main formulas and ideas of the traditional engineering approach to reliability; see, e.g., [1].

Assumptions about the corresponding probabilities. Let λ_i denote the actual (unknown) failure rate of the i-th component. It is usually assumed that the failure rate has an exponential probability distribution, i.e., that the probability density is equal to $\lambda_i \cdot \exp(-\lambda_i \cdot t)$.

Thus, the probability density corresponding to each observation t_{ij} is equal to $\lambda_i \cdot \exp(-\lambda_i \cdot t_{ij})$.

Different observations are assumed to be independent – and different components are also assumed to be independent. Thus, the probability density ρ corresponding to all observed

failures is equal to the product of the corresponding probabilities:

$$\rho = \prod_{i=1}^{n} \prod_{j=1}^{n_i} (\lambda_i \cdot \exp(-\lambda_i \cdot t_{ij})). \tag{1}$$

How parameters are determined: idea. The desired values λ_i are usually determined from the Maximum Likelihood approach (see, e.g., see, e.g., [2], [3]), according to which we select the values λ_i for which the probability density (18) takes the largest possible value. In other words, we find the values λ_i that maximize the expression (18).

Analysis of the corresponding optimization problem. To solve the corresponding optimization problem, we can use the usual statistical trick according to which, since logarithm is a monotonic function, maximizing ρ is equivalent to minimizing $\psi \stackrel{\text{def}}{=} -\ln(\rho)$. Since the logarithm of a product is equal to the sum of the logarithms, the product (18) transforms into a sum – and terms in this sum which do not depend on the unknowns λ_i can be safely ignored.

As a result, we arrive at the problem of minimizing the following function:

$$\psi(\lambda_i) = -\sum_{i=1}^{n} n_i \cdot \ln(\lambda_i) + \sum_{i=1}^{n} \sum_{j=1}^{n_i} \lambda_i \cdot t_{ij}. \tag{2}$$

If we denote the arithmetic average of the values t_{ij} by

$$t_i \stackrel{\text{def}}{=} \frac{1}{n_i} \cdot \sum_{j=1}^{n_i} t_{ij}, \tag{3}$$

then we get $\sum\limits_{j=1}^{n_i} t_{ij} = n_i \cdot t_i$ and thus, the formula (4) takes a simplified form:

$$\psi(\lambda_i) = -\sum_{i=1}^{n} n_i \cdot \ln(\lambda_i) + \sum_{i=1}^{n} n_i \cdot \lambda_i \cdot t_i. \tag{4}$$

Differentiating this expression by λ_i and equating the derivative to 0, we get

$$\frac{\partial \psi}{\partial \lambda_i} = -\frac{n_i}{\lambda_i} + n_i \cdot t_i = 0, \tag{5}$$

thus

$$\lambda_i = \frac{1}{t_i}. \tag{6}$$

Resulting formula for the failure rate. For each component i, based on the observed times-to-failure $t_{i1}, \ldots, t_{in_i}$, we compute the average time-to-failure $t_i = \dfrac{1}{n} \cdot \sum\limits_{j=1}^{n_i} t_{ij}$ and then estimate the failure rate as $\lambda_i = \dfrac{1}{t_i}$.

How to determine the accuracy of this estimate. We have found the most probable values λ_i. This value is based on a finite sample, and is, thus, only an approximation to the actual (unknown) value λ_i.

To determine the accuracy with which this most probable value represents the actual value λ_i, it is reasonable to determine the standard deviation σ_i of this estimate. Due to the Central Limit Theorem (see, e.g., [2], [3]), when we have a large amount of data, the distribution of all the values is close to normal. In particular, the distribution of λ_i takes the form

$$\rho(\lambda_i) = \text{const} \cdot \exp\left(-\frac{(\lambda_i - \mu_i)^2}{2 \cdot \sigma_i^2}\right) \tag{7}$$

for some μ_i ad σ_i.

Thus, for $\psi = -\ln(\rho)$, we have

$$\psi(\lambda_i) = \text{const} - \frac{(\lambda_i - \mu_i)^2}{2 \cdot \sigma_i^2}. \tag{8}$$

Differentiating both sides by λ_i, we get

$$\frac{\partial \psi}{\partial \lambda_i} = \text{const} - \frac{\lambda_i - \mu_i}{\sigma_i^2}. \tag{9}$$

We have determined our maximum Likelihood estimate from the condition that this derivative is equal to 0. When this derivative is equal to 0, we get $\lambda_i = \mu_i$. Thus, μ_i is exactly our estimate for λ_i.

Differentiating the equality (9) with respect to λ_i once again, we conclude that

$$\frac{\partial^2 \psi}{\partial \lambda_i^2} = \frac{1}{\sigma_i^2}. \tag{10}$$

Thus, the desired standard deviation can be determined from the condition

$$\sigma_i^2 = \left(\frac{\partial^2 \psi}{\partial \lambda_i^2}\right)^{-1}. \tag{11}$$

In particular, for the expression (12), we have

$$\frac{\partial^2 \psi}{\partial \lambda_i^2} = \frac{n_i}{\lambda_i^2}, \tag{12}$$

thus,

$$\sigma_i^2 = \left(\frac{\partial^2 \psi}{\partial \lambda_i^2}\right)^{-1} = \frac{\lambda_i^2}{n_i} \tag{13}$$

and

$$\sigma_i = \frac{\lambda_i}{\sqrt{n_i}}. \tag{14}$$

Resulting accuracy of this estimate. The standard deviation of the Maximum Likelihood estimate is equal to $\sigma_i = \dfrac{\lambda_i}{\sqrt{n_i}}$. Thus, the relative accuracy of this estimate – i.e., the ratio of the standard deviation to the estimate – is equal to

$$\frac{\sigma_i}{\lambda_i} = \frac{1}{\sqrt{n_i}}.$$

Confidence interval. In general, once we have an estimate λ_i and an estimate σ_i for its standard deviation, we can form a confidence interval $[\lambda_i - k_0 \cdot \sigma_i, \lambda_i + k_0 \cdot \sigma_i]$ that contains the actual failure rate λ_i with a given confidence. Here, the

value of k_0 depends on the desired degree of confidence. For example:

- we take $k_0 = 2$ if we want 90% confidence;
- we take $k_0 = 3$ if we want 99.9% confidence;
- we take $k_0 = 6$ if we want 99.9999999% = $1 - 10^{-8}$ confidence.

In the traditional approach, we have an estimate $\lambda_i = \dfrac{1}{t_i}$ and an estimate (33) for σ_i. In this case, the confidence interval takes the form

$$\left[\lambda_i \cdot \left(1 - \frac{k_0}{\sqrt{n_i}}\right), \lambda_i \cdot \left(1 + \frac{k_0}{\sqrt{n_i}}\right)\right]. \tag{15}$$

Example. For $n_i = 5$ and $k_0 = 2$, we have $\sqrt{n_i} \approx 2$, so $\dfrac{k_0}{\sqrt{n_i}} \approx 1$, and the confidence interval is approximately equal to $[0, 2\lambda_i]$. In other words, the actual failure rate can be 0 or it can be twice higher than what we estimated. Thus, if we only have 5 measurements, we cannot extract much information about the actual failure rate.

IV. New Approach: Main Idea

Our main idea is to use the fact that here, we have *two* sources of knowledge:

- the empirical failure data t_{ij}, and
- the expert estimates $e_1, \ldots, e_n$ for the failure rates of different components.

For each individual component i, we do not have enough data to provide us with a meaningful statistically significant estimate for its failure rate λ_i.

However, as we will show,

- when we combine all these data together, we will get enough data points to gauge the accuracy of an expert – as an instrument for estimating the failure rates,
- as a result of this statistical analysis, an expert becomes a statistically justified estimation tool; so, we can add the expert estimates to the observed times t_{ij}; this additional data allows us to get better estimates for λ_i.

Discussion. One may ask a natural question: if the empirical data t_{ij} are not sufficient to make statistically reliable estimates about the failure rate, why these data are considered sufficient for gauging/reducing the over-conservativeness of experts' estimates? The answer to this question is as follows.

Failures of different components are considered statistically independent. Thus, *in the absence of expert estimates*, to find the failure rate λ_i, we can only use the values t_{ij} corresponding to this component. Since we have only $n_i = 5$ such values, this data is not sufficient to make accurate statistically reliable estimates about λ_i.

On the other hand, the over-conservativeness of *an expert* is reflected in the expert's estimates of the failure rates of all the components. Thus, to estimate this over-conservativeness, we can use the data from all the components. We may have about 5 measurement values for each component, but since we have dozens of components, we thus have hundreds of values t_{ij} that can be used to estimate this over-conservativeness, and a hundred data points is already enough to make statistically reliable estimates.

V. New Fuzzy/Statistical Model and the Resulting Optimization Problem

Experts over-estimate. As we have mentioned, expert estimates are usually over-conservative, they over-estimate the failure rates. In terms of our notations, this means that $\lambda_i \approx k \cdot e_i$ for some $k < 1$. In other words, experts as a collective "measuring instrument" are characterized by an (unknown) parameter k.

Over-estimation may be different for different components. Of course, the relation $\lambda_i \approx k \cdot r_i$ is only approximate. The actual ratio $k_i \stackrel{\text{def}}{=} \dfrac{\lambda_i}{e_i}$, in general, differs from a component to a component.

Normal distribution for each k_i. As usual in statistics (see, e.g., [2], [3]), it is reasonable to assume that the ratios k_i are normally distributed, with mean k and an (unknown) standard deviation σ^2.

Thus, we have

$$\lambda_i = k_i \cdot e_i, \tag{16}$$

and the corresponding Gaussian probability density has the form

$$\frac{1}{\sqrt{2 \cdot \pi} \cdot \sigma} \cdot \exp\left(-\frac{(k_i - k)^2}{2\sigma^2}\right). \tag{17}$$

Comment. A similar formula can be obtained if we do not use any assumptions about the probability distributions, but simply use a Gaussian membership function to describe the inaccuracy of expert estimates.

Approximation errors corresponding to different components are independent. It is natural to assume that, in contrast to the "bias" k, the approximation errors $k_i - k$ corresponding to different components are independent.

Resulting formula. Thus, the probability density corresponding to all the components is equal to the product of expressions (17), i.e., to the expression

$$\rho' = \prod_{i=1}^{n} \frac{1}{\sqrt{2 \cdot \pi} \cdot \sigma} \cdot \exp\left(-\frac{(k_i - k)^2}{2\sigma^2}\right). \tag{18}$$

Substituting the formula (16) into (1) and multiplying the result by the probability (18), we get the final expression for the probability density function ρ:

$$\rho = \left[\prod_{i=1}^{n} \prod_{j=1}^{n_i} (k_i \cdot e_i \cdot \exp(-(k_i \cdot e_i \cdot t_{ij}))\right] \cdot \rho', \tag{19}$$

where ρ' is determined by the formula (18).

Comment. In the fuzzy case, a similar formula can be obtained without any independence assumption, if we use a product t-norm $f_{\&}(a, b) = a \cdot b$ to combine information about different components.

Resulting optimization problem. Similarly to the traditional engineering approach to reliability, we will use the Maximum Likelihood Method to find the desired values λ_i.

Specifically,

- we find the values of k, k_i, and σ that maximize the expression (19), and then
- we estimate the values λ_i as $\lambda_i = k_i \cdot e_i$.

VI. How to Solve the New Optimization Problem: Analysis

First simplification. Let us first observe that in the expression (19), the observed times t_{ij} occur only in product with the expert estimates e_i. Thus, the expression (19) can be somewhat simplified if we introduce new parameters $T_{ij} \stackrel{\text{def}}{=} e_i \cdot t_{ij}$. In terms of these new parameters T_{ij}, the expression (19) takes a simplified form

$$\left[\prod_{i=1}^{n}\prod_{j=1}^{n_i}(k_i \cdot e_i \cdot \exp(-(k_i \cdot T_{ij}))\right] \cdot \left[\prod_{i=1}^{n}\frac{1}{\sqrt{2\cdot\pi}\cdot\sigma}\cdot\exp\left(-\frac{(k_i-k)^2}{2\sigma^2}\right)\right]. \quad (20)$$

First simplification. Let us use the fact that maximizing L is equivalent to minimizing $\psi \stackrel{\text{def}}{=} -\ln(\rho)$. Since the logarithm of a product is equal to the sum of the logarithms, the product (19) transforms into a sum – and terms in this sum which do not depend on the unknowns k, k_i, and σ can be safely ignored.

As a result, we arrive at the problem of minimizing the following function:

$$\psi = -\sum_{i=1}^{n} n_i \cdot \ln(k_i) + \sum_{i=1}^{n}\sum_{j=1}^{n_i} k_i \cdot e_i \cdot t_{ij} + n\cdot\ln(\sigma) + \sum_{i=1}^{n}\frac{(k_i-k)^2}{2\sigma^2}. \quad (21)$$

Second simplification. If we denote the arithmetic average of the times-between-failures by

$$t_i \stackrel{\text{def}}{=} \frac{1}{n_i}\cdot\sum_{j=1}^{n_i} t_{ij}, \quad (22)$$

then we get $\sum\limits_{j=1}^{n_i} t_{ij} = n_i \cdot i_i$. Then, the function ψ takes the following form:

$$\psi = -\sum_{i=1}^{n} n_i \cdot \ln(k_i) + \sum_{i=1}^{n} k_i \cdot n_i \cdot e_i \cdot t_i + n\cdot\ln(\sigma) + \sum_{i=1}^{n}\frac{(k_i-k)^2}{2\sigma^2}. \quad (23)$$

From the optimization problem to the system of equations. A function attains its minimum when its partial derivatives with respect to all its unknown are equal to 0.

Differentiating the above expression (23) relative to σ and equating the derivative to 0, we conclude that

$$\frac{n}{\sigma} - \frac{1}{\sigma^3}\cdot\sum_{i=1}^{n}(k_i-k)^2 = 0. \quad (24)$$

Differentiating the expression (21) by k, we conclude that

$$\sum_{i=1}^{n}\frac{k-k_i}{\sigma^2} = 0. \quad (25)$$

Finally, differentiating the expression (21) by k_i, we get

$$\frac{\partial\psi}{\partial k_i} = -\frac{n_i}{k_i} + n_i\cdot e_i\cdot t_i + \frac{k_i-k}{\sigma^2} = 0. \quad (26)$$

So, in order to find the values k, k_i, and σ, it is sufficient to solve a system consisting of the equations (24), (25), and (26).

Simplifying the resulting system of equations: equation coming from differentiating by σ. Multiplying both sides of the equation (24) by σ^3 and dividing by n, we get the following formula

$$\sigma^2 = \frac{1}{n}\cdot\sum_{i=1}^{n}(k_i-k)^2. \quad (27)$$

Simplifying the resulting system of equations: equation coming from differentiating by k. The equation (25) is equivalent to $\sum\limits_{i=1}^{n}(k-k_i) = 0$ and

$$k = \frac{1}{n}\cdot\sum_{i=1}^{n} k_i. \quad (28)$$

Simplifying the resulting system of equations: equation coming from differentiating by k_i. Multiplying both sides of the equation (26) by the common denominator $k_i\cdot\sigma^2$, we get the following equation:

$$-n_i\cdot V + n_i\cdot e_i\cdot t_i\cdot\sigma^2\cdot k_i + k_i^2 - k_i\cdot k = 0.$$

This equation is quadratic w.r.t. k_i:

$$k_i^2 + k_i\cdot(n_i\cdot t_i\cdot e_i\cdot\sigma^2 - k) - n_i\cdot\sigma^2 = 0. \quad (29)$$

Once we know k and σ, we can explicitly solve the quadratic equation and get k_i for all i, as

$$k_i = \frac{k - n_i t_i e_i \sigma^2 + \sqrt{(k-n_i t_i e_i\sigma^2)^2 + 4n_i\sigma^2}}{2}. \quad (30)$$

Comment. In principle, a quadratic equation has two roots, but the second root (with a minus in front of the square root) is negative, while the failure rate λ_i is always positive and so, $k_i = e_i/\lambda_i > 0$; thus, we only consider the positive square root.

Simplified system of equations. As a result of the above simplifications, we get the following system of equations: (27), (28), and (29) (or (30)).

How we can actually solve this system of equations. We need to solve the equations (27), (28), and (30) to find the unknown k, k_i, and σ^2. Each of these equations explicitly describes how to find one of the unknowns if we know others:

- once we know k and k_i, the formula (27) enables us to compute σ^2;
- once we know k_i, the equation (28) enables us to compute k; and
- once we know k and σ^2, the equation (30) enables us to compute k_i.

Thus, it is reasonable to find all the unknown as follows. We start with some initial values $k_i^{(0)}$ for k_i. For example, as the first approximation, we can take the values coming from processing the observed times-between-failures. In this approximation, $\lambda_i = \dfrac{1}{t_i}$, and, since $\lambda_i = k_i \cdot e_i$, we take $k_i^{(0)} = \dfrac{\lambda_i}{e_i} = \dfrac{1}{e_i \cdot t_i}$.

Then, on each iteration p, $p = 0, 1, 2 \ldots$:

- we use the formula (28) and the current approximations $k_i^{(p)}$ to k_i to compute a approximation to k and the formula (27) to compute the corresponding approximation to k and σ^2:

$$k^{(p)} = \frac{1}{n} \cdot \sum_{i=1}^{n} k_i^{(p)}; \quad (\sigma^2)^{(p)} = \frac{1}{n} \cdot \sum_{i=1}^{n} (k_i^{(p)} - k^{(p)})^2;$$

- after that, we use the formula (30) to compute the next approximation $k_i^{(p+1)}$ to k_i:

$$z = k^{(p)} - n_i \cdot t_i \cdot e_i \cdot (\sigma^2)^{(p)};$$

$$k_i^{(p+1)} = \frac{z + \sqrt{z^2 + 4 n_i \cdot (\sigma^2)^{(p)}}}{2}.$$

We stop when the values on the two consequent iterations are close to each other with a given accuracy, i.e., when $|k_i^{(p+1)} - k_i^{(p)}| \le \varepsilon \cdot k_i^{(p)}$ for all i, where $\varepsilon > 0$ is the relative accuracy with which we want to determine the values k_i.

After that, we can find the estimate for λ_i; by definition of the factor k_i, we take $\lambda_i = k_i \cdot e_i$.

Resulting accuracy of this estimate. Similarly to the standard approach, the standard deviation σ_i of the estimate k_i can be determined by the formula

$$\sigma_i^2 = \left(\frac{\partial^2 \psi}{\partial k_i^2} \right)^{-1}.$$

In particular, for the expression (26), we have

$$\frac{\partial^2 \psi}{\partial k_i^2} = \frac{n_i}{k_i^2} + \frac{1}{\sigma^2}, \tag{31}$$

thus,

$$\sigma_i^2 = \left(\frac{\partial^2 \psi}{\partial \lambda_i^2} \right)^{-1} = \frac{k_i^2}{n_i + k_i^2 \cdot \sigma^{-2}} \tag{32}$$

and

$$\sigma_i = \frac{k_i}{\sqrt{n_i + k_i^2 \cdot \sigma^{-2}}}. \tag{33}$$

Thus, the relative accuracy $\dfrac{\sigma_i}{k_i}$ of this estimate is equal to

$$\frac{\sigma_i}{k_i} = \frac{1}{\sqrt{n_i + k_i^2 \cdot \sigma^{-2}}}.$$

The relative accuracy does not change if we simply multiply the value by e_i. Thus, for $\lambda_i = k_i \cdot e_i$, we have the same relative accuracy. As a result, we get the following confidence interval.

Confidence interval. Based on the estimate λ_i and on the standard deviation σ_i, we can form a confidence interval $[\lambda_i - k_0 \cdot \sigma_i, \lambda_i + k_0 \cdot \sigma_i]$ that contains the actual failure rate λ_i with a given confidence. In our case, the confidence interval takes the form

$$\left[\lambda_i \cdot \left(1 - \frac{k_0}{\sqrt{n_i + k_i^2 \cdot \sigma^{-2}}} \right), \right.$$
$$\left. \lambda_i \cdot \left(1 + \frac{k_0}{\sqrt{n_i + k_i^2 \cdot \sigma^{-2}}} \right) \right]. \tag{34}$$

VII. Resulting Algorithm

Available data: reminder. For each of these components $i = 1, \ldots, n$, we have n_i observed times-between-failures $t_{i1}, \ldots, t_{in_i}$. We also have expert estimates $e_1, \ldots, e_n$ for the failure rate of each component.

Pre-processing. First, for each component, we compute the average of the observed times-between-failures

$$t_i = \frac{1}{n_i} \cdot \sum_{j=1}^{n_i} t_{ij}. \tag{35}$$

Then, we compute the first approximation $k_i^{(0)}$ to the auxiliary parameter k_i describing the expert's over-estimation of failure rate for the i-th component:

$$k_i^{(0)} = \frac{1}{e_i \cdot t_i}. \tag{36}$$

Iterations. On each iteration p, $p = 0, 1, 2 \ldots$, based on the current approximations $k_i^{(p)}$ to k_i, we do the following:

- we compute the corresponding approximation to the auxiliary variables k and σ^2:

$$k^{(p)} = \frac{1}{n} \cdot \sum_{i=1}^{n} k_i^{(p)}; \tag{37}$$

$$(\sigma^2)^{(p)} = \frac{1}{n} \cdot \sum_{i=1}^{n} (k_i^{(p)} - k^{(p)})^2; \tag{38}$$

- after that, we compute the next approximation $k_i^{(p+1)}$ to k_i:

$$z = k^{(p)} - n_i \cdot t_i \cdot e_i \cdot (\sigma^2)^{(p)};$$

$$k_i^{(p+1)} = \frac{z + \sqrt{z^2 + 4n_i \cdot (\sigma^2)^{(p)}}}{2}. \quad (39)$$

We stop when the values on the two consequent iterations are close to each other with a given accuracy, i.e., when

$$|k_i^{(p+1)} - k_i^{(p)}| \le \varepsilon \cdot k_i^{(p)}$$

for all i, where $\varepsilon > 0$ is the relative accuracy with which we want to determine the values k_i.

Once we have $k_i = k_i^{(p)}$, we then estimate λ_i as $k_i \cdot e_i$, and the corresponding confidence interval as

$$\left[\lambda_i \cdot \left(1 - \frac{k_0}{\sqrt{n_i + k_i^2 \cdot \sigma^{-2}}}\right), \lambda_i \cdot \left(1 + \frac{k_0}{\sqrt{n_i + k_i^2 \cdot \sigma^{-2}}}\right)\right]. \quad (40)$$

Discussion. The difference between this new confidence interval and the confidence interval based only on the observed times-between-failures is that we replace n_i in the denominator with a larger value $n_i + k_i^2 \cdot \sigma^{-2}$. Thus, the resulting confidence interval is indeed narrower.

When the values k_i are very close and $\sigma \approx 0$, this denominator tends to ∞, so we get very narrow confidence intervals for λ_i even when we have the same small number of observations – because we also use expert estimates e_i as the additional source of information about the components' reliability.

ACKNOWLEDGMENTS

This work was supported by a contract from the Boeing Company. It is also partly supported by the National Science Foundation grants HRD-0734825 and DUE-0926721 and by Grant 1 T36 GM078000-01 from the National Institutes of Health. The authors are thankful to the anonymous referees for valuable suggestions.

REFERENCES

[1] E. E. Barlow, *Engineering Reliability*, Society of Industrial and Applied Mathematics (SIAM), Philadelphia, Pennsylvania, and American Statistical Association (ASA), Alexandria, Virginia, 1998.

[2] S. Rabinovich, *Measurement Errors and Uncertainties: Theory and Practice*, American Institute of Physics, New York, 2005.

[3] D. J. Sheskin, *Handbook of Parametric and Nonparametric Statistical Procedures*, Chapman & Hall/CRC, Boca Raton, Florida, 2004.