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, El Paso, TX 79968, USA
ferregut@utep.edu, fjcampos@miners.utep.edu, vladik@utep.edu

1. Reliability

- *Failures* are ubiquitous, so *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.
- Reliability of a component is usually described by an *exponential model*:

$$P(t) \stackrel{\text{def}}{=} \text{Prob}(\text{system is intact by time } t) = \exp(-\lambda \cdot t).$$

- For this model, the average number of failures per unit time (*failure rate*) is equal to λ.
- Another important characteristic – *mean time between failure* (MTBF) θ – is, in this model, equal to $1/\lambda$.
- Usually, the MTBF is *estimated* as the average of observed times between failures.

2. Reliability in Aerospace Industry: A Challenge

- In aerospace industry, reliability is extremely important, especially for manned flights.
- Because of this importance, aerospace systems use unique, highly reliable components; but:
  - since each component is highly reliable,
  - we have few (≤ 5) failure records, not enough to make statistically reliable estimates of λ.
- So, we have to also use expert estimates.
- Problem: experts are over-conservative, their estimates for λ are higher than the actual failure rate.
- In this paper, we propose an algorithm that reduces the effect of this over-conservativeness.

3. Available Data and Main Assumptions

- For each of n 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.
- We usually assume the exponential distribution for the failure times, i.e., the probability density $\lambda_i \cdot \exp(-\lambda_i \cdot t)$.
- 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.
- Different components are assumed to be independent.
- Thus, the probability density ρ corresponding to all observed failures is equal to the product:

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

4. How Parameters Are Determined Now

- *Reminder:* prob. is $\rho = \prod_{i=1}^{n} \prod_{j=1}^{n_i} (\lambda_i \cdot \exp(-\lambda_i \cdot t_{ij}))$.
- *Maximum Likelihood Approach:* find λ_i with the highest probability ρ.
- *Idea:* $\rho \to \max$ if and only if $\psi \stackrel{\text{def}}{=} -\ln(\rho) \to \min$:

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

- So, $\psi(\lambda_i) = -\sum_{i=1}^{n} n_i \cdot \ln(\lambda_i) + \sum_{i=1}^{n} n_i \cdot \lambda_i \cdot t_i$, where

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

- Differentiating by λ_i and equating the derivative to 0, we get the traditional estimate $\lambda_i = \frac{1}{t_i}$.

5. What Is the Accuracy of This Estimate?

- *Central Limit Theorem:* when we have a large amount of data, the distribution of each parameter is $\approx$ normal:

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

- Thus, $\psi(\lambda_i) = \text{const} + \dfrac{(\lambda_i - \mu_i)^2}{2 \cdot \sigma_i^2}$, hence $\dfrac{\partial^2 \psi}{\partial \lambda_i^2} = \dfrac{1}{\sigma_i^2}$, and $\sigma_i^2 = \left(\dfrac{\partial^2 \psi}{\partial \lambda_i^2}\right)^{-1}$.

- For $\psi(\lambda_i) = -\sum\limits_{i=1}^{n} n_i \cdot \ln(\lambda_i) + \sum\limits_{i=1}^{n} n_i \cdot \lambda_i \cdot t_i$, we get $\dfrac{\partial^2 \psi}{\partial \lambda_i^2} = \dfrac{n_i}{\lambda_i^2}$, so the standard deviation is $\sigma_i = \dfrac{\lambda_i}{\sqrt{n_i}}$.

- So, the relative accuracy of the estimate λ_i is equal to

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

6. Confidence Interval

- Based on λ_i and σ_i, we can form an interval that contains the actual failure rate with a given confidence:

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

- 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.
- *Example:* for $n_i = 5$ and $k_0 = 2$, 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.

7. New Approach: Main Idea

- Experts provide *estimates* e_i for the failure rates λ_i.
- Expert *over-estimate*, i.e., $\lambda_i = k_i \cdot e_i$ for some $k_i < 1$.
- As usual, it is reasonable to assume that k_i are *normally distributed*, with unknown k and σ^2:

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

- Approximation errors $k_i - k$ corresponding to *different components* are independent.
- We then find λ_i, k, and σ from the *Maximum Likelihood Method* $\rho \to \max$:

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

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

8. Fuzzy Interpretation

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

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

- This formula is based on the *assumptions* of
  - Gaussian distribution, and
  - independence.
- A *similar formula* can be obtained if we simply use:
  - Gaussian membership functions, and
  - a product t-norm $f_\&(a, b) = a \cdot b$ to combine information about different components.
- In this case, we *do not need* independence *assumptions*.

9. Analysis of the Optimization Problem

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

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

- We reduce $\rho \to \max$ to $\psi \stackrel{\text{def}}{=} -\ln(\rho) \to \min$ and use t_i:

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

- Equating derivatives w.r.t. σ, k, and k_i to 0, we get:

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

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

10. Analysis of the Problem (cont-d)

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

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

- Multiplying both sides by k_i, we get a quadratic equation, with solution

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

- Thus, we start with some initial values $k_i^{(0)}$, and perform the following iterations until the process converges:
  - update k to $\frac{1}{n} \cdot \sum_{i=1}^{n} k_i$ and σ^2 to $\frac{1}{n} \cdot \sum_{i=1}^{n} (k_i - k)^2$;
  - update k_i to $\frac{k - n_i t_i e_i \sigma^2 + \sqrt{\ldots}}{2}$.

11. Resulting Accuracy of This Estimate

- The st. dev. σ_i of the estimate k_i is $\sigma_i^2 = \left(\dfrac{\partial^2 \psi}{\partial k_i^2}\right)^{-1}$.
- Here, $\psi = -\sum\limits_{i=1}^{n} n_i \cdot \ln(k_i) + \sum\limits_{i=1}^{n} k_i \cdot n_i \cdot e_i \cdot t_i + n \cdot \ln(\sigma) + \sum\limits_{i=1}^{n} \dfrac{(k_i - k)^2}{2\sigma^2}$, so $\dfrac{\partial^2 \psi}{\partial k_i^2} = \dfrac{n_i}{k_i^2} + \dfrac{1}{\sigma^2}$.
- Thus, $\dfrac{\sigma_i}{k_i} = \dfrac{1}{\sqrt{n_i + k_i^2 \cdot \sigma^{-2}}}$.
- The relative accuracy does not change if we multiply the k_i by e_i, i.e., go from k_i to $\lambda_i = k_i \cdot e_i$.
- So, the confidence interval for λ_i is:

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

12. Resulting Algorithm

- *Given:*
  - For each of component $i = 1, \ldots, n$, 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

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

  - Then, we compute the first approximation $k_i^{(0)}$ to the auxiliary parameter k_i: $k_i^{(0)} = \dfrac{1}{e_i \cdot t_i}$.

13. Algorithm (cont-d)

- *Iterations:* on each iteration p, $p = 0, 1, 2 \ldots$, based on the current approximations $k_i^{(p)}$, we compute:

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

$$z = k^{(p)} - n_i \cdot t_i \cdot e_i \cdot (\sigma^2)^{(p)}; \quad k_i^{(p+1)} = \frac{z + \sqrt{z^2 + 4 n_i \cdot (\sigma^2)^{(p)}}}{2}.$$

- We stop when $|k_i^{(p+1)} - k_i^{(p)}| \le \varepsilon \cdot k_i^{(p)}$ for all 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

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

14. Discussion.

- *Reminder:* we get the following confidence interval:

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

- The difference between this confidence interval and the confidence interval based only on observations is that:
  - we replace n_i in the denominator
  - with a larger value $n_i + k_i^2 \cdot \sigma^{-2}$.
- Thus, the new confidence interval is indeed narrower: expert estimates help.
- 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.

15. 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.