Lognormal Uncertainty Estimation for Failure Rates

RAM X Training Summit
Huntsville, AL
November 8-9, 2017

Paul Britton [1]
Mohammad AL Hassan [1]
Rob Ring[2]

1. **NASA, Marshal Space Flight Center (MSFC)**
2. **Bastion Technologies, Inc.**

Presentation Outline

- Introduction
- Background Material
  - Probability Distributions
  - Central Tendency and Dispersion
  - Failure Rate Uncertainty
- The Bayesian Approach
- Uncertainty Estimation Examples
- Lognormal Uncertainty Estimation Cheat Sheet
- Conclusion

Introduction

When uncertainty estimates are expected to inform decision-makers, it is especially important to carefully consider, understand, and communicate the significance of the statistical parameters used in the characterization of failure probability distributions.

Uncertainty analysis aims to make a technical contribution to decision-making through the quantification of uncertainties in the relevant variables as well as through the propagation of these uncertainties up to the result.

We will illustrate key principles as we step through the quantification of uncertainty.

Finally, the risk implications of uncertainty estimation are summarized in a convenient reference card: Uncertainty Estimation Cheat Sheet.

Probability Distributions

- Informally, a **probability distribution** is a mathematical function that assigns probabilities to each element of the **sample space** (the set of all possible outcomes in an experiment).

- A **random variable** is a function that maps outcomes of an experiment to numerical quantities.

- For a continuous distribution, the **probability density function (pdf)** is the function that is used to generate the probability that a random variable ***X*** lies within an interval [*a*, *b*]:

$$\Pr[a \leq \boldsymbol{X} \leq b] = \int_a^b f(x)\, dx$$

Probability Distributions

- The probability density of the **exponential distribution** is:

$$f(t) = \begin{cases} \lambda e^{-\lambda t}, & t \geq 0 \\ 0, & t < 0 \end{cases}$$

- The pdf of the **normal (or Gaussian) distribution** is:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

- The pdf of the **lognormal distribution** parameterized with the mean (μ) and standard deviation (σ) of the underlying normal distribution is given as:

$$f(\lambda) = \frac{1}{\lambda\sigma\sqrt{2\pi}} \exp\left(-\frac{[\ln(\lambda) - \mu]^2}{2\sigma^2}\right), (0 < \lambda < \infty)$$

Central Tendency

- For a continuous distributions, the arithmetic **mean** is:

$$E[\boldsymbol{X}] = \int x f(x)\, dx,$$

where the weighting function *f*(*x*) is the pdf of ***X***.

- The **median** or 50th percentile is the midpoint where half of the probability (area under the pdf) lies to either side.

$$\int_{-\infty}^{median} f(x)\, dx = \int_{median}^{\infty} f(x)\, dx = \frac{1}{2}$$

- The **mode** is a local maximum or peak of the pdf.

Dispersion

- The **variance** is the expected value of the squared deviations about the mean:

$$Var[\boldsymbol{X}] = E[(\boldsymbol{X} - E[\boldsymbol{X}])^2]$$

- The square root of the variance is the **standard deviation**:

$$\sigma = \sqrt{variance}$$

- The **error factor (EF)** is defined as the square root of the 95th percentile divided by the 5th percentile. Equivalently, the EF is equal to the 50th divided by the 5th and the 95th divided by the 50th as summarized in the following equivalence:

$$EF = \sqrt{\frac{95^{th}\ percentile}{5^{th}\ percentile}} = \frac{95^{th}\ Percentile}{50^{th}\ Percentile} = \frac{50^{th}\ Percentile}{5^{th}\ Percentile}$$

Failure Rate Uncertainty

- Component failure rates (λ) are not physical quantities; that is, they cannot be measured directly but must be inferred.
- Previous research evaluated different distributions to represent the uncertainty of the parameter λ [1]. They found the lognormal distribution was appropriate for simple components with a single failure mode.
- Uncertainty has many sources in addition to variation among individuals within a population and lack knowledge due to sparse data. However, this paper examines the implications of applying uncertainty around central tendency estimates in order to quantify degree of belief – in particular when expressing degree of belief via the shape of the lognormal pdf.

The Bayesian Approach

- Application of classical life data analysis requires component data in the form of failures and exposure time or number of demands.
- Highly reliable components produced in small quantities, such as in space applications, do not have enough operating time and failure history to yield useful confidence bounds using classical statistical data analysis
- Bayesian approach is able to address the challenges described above because it admits prior experience into the estimation procedure in the form of a prior degree of belief about the likely values of the component in the form of a prior distribution.
- In our experience, engineers with specific discipline expertise are generally familiar with the normal probability distribution, but have little direct experience with skewed distributions, such as the lognormal.
- Subject matter experts who often assist PRA analysts in the quantification of the prior failure rate distribution must be educated to develop an intuitive understanding of how the lognormal distribution morphs as its centrality and dispersion measures are varied.

The Bayesian Approach

One of the main purposes of this paper is to illustrate with specific examples the effects of varying one of the parameters, such as the dispersion while holding another fixed to show the effect on the remaining parameters.

Specifying any two parameter values completely specifies the lognormal distribution. Thus we can solve for μ and σ and then fill in the remaining parameter values in the table using the formulas.

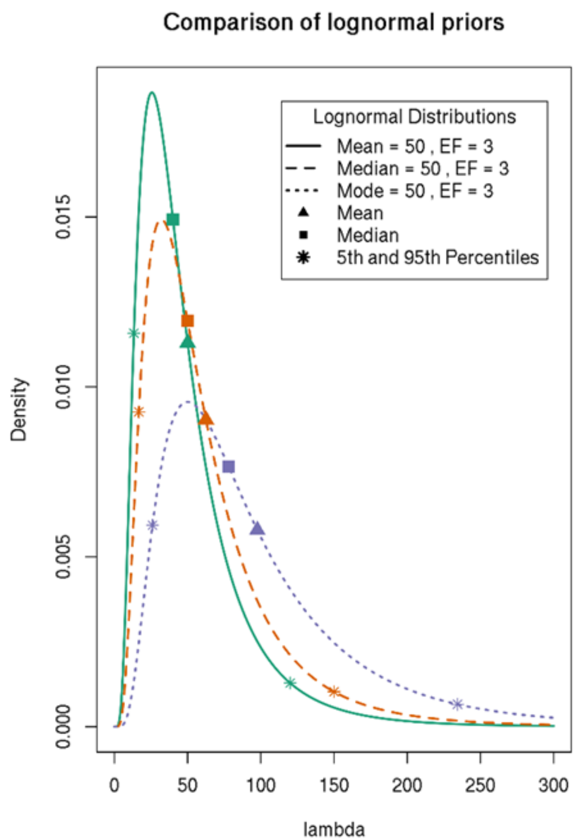
Parameter	**As a function of μ and σ**
Mean	$\exp\left(\mu + \frac{\sigma^2}{2}\right)$
Median	$\exp(\mu)$
Mode	$\exp(\mu - \sigma^2)$
Standard Deviation	$\sqrt{[\exp(\sigma^2) - 1]\exp(2\mu + \sigma^2)}$
Error Factor	$\exp(1.64485\sigma)$

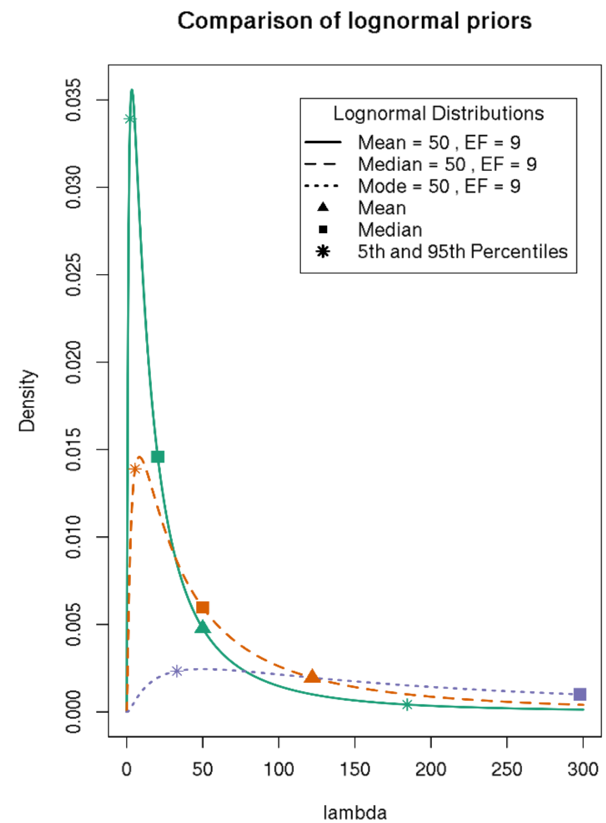
Uncertainty Estimation Examples

- Reliability predictions are reported as point estimates. The PRA has to estimate uncertainty to create the probability distribution that represent degree of belief. Heuristic approaches have been used. These approaches consider the data source applicability with respect to similarity. Multipliers can be applied to convert the data from the reported operating environment to a more applicable one.
- By fixing each of the mode, mean and median, while varying the error factor, we demonstrate the effect on the other measures of centrality as well as the risk implications to results via uncertainty propagation.

Uncertainty Estimation Example 1

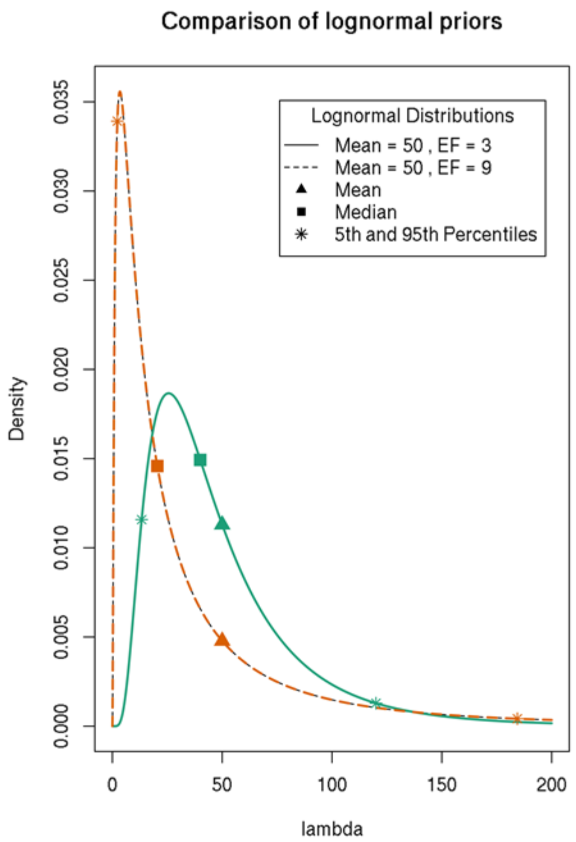
- These figures illustrate the effect on the resultant prior pdf for the three variations of the heuristic method using assumed measures of dispersion error factors of 3 and 9, respectively. What is given in the contractor's reliability analysis report is the point estimate for failure rate (λ) of 50 failures per million hours (FPMH) of exposure.

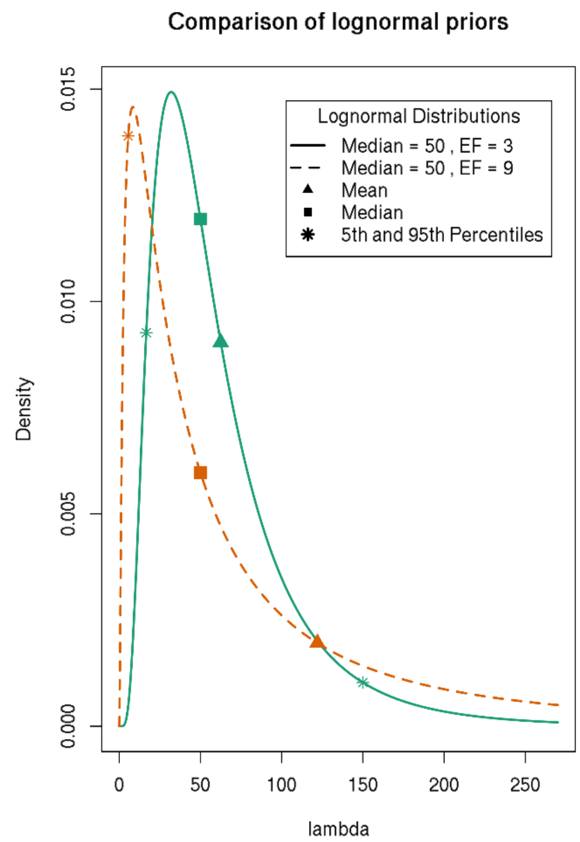


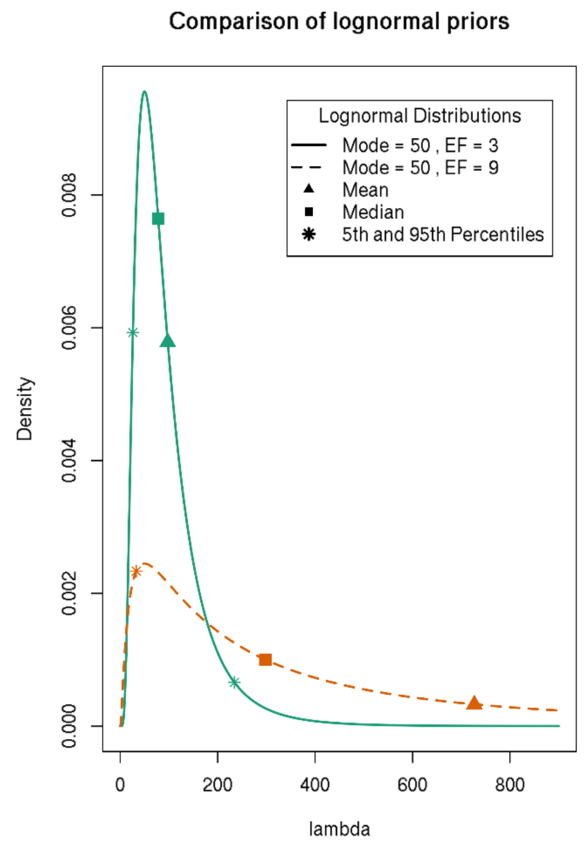


Uncertainty Estimation Example 2

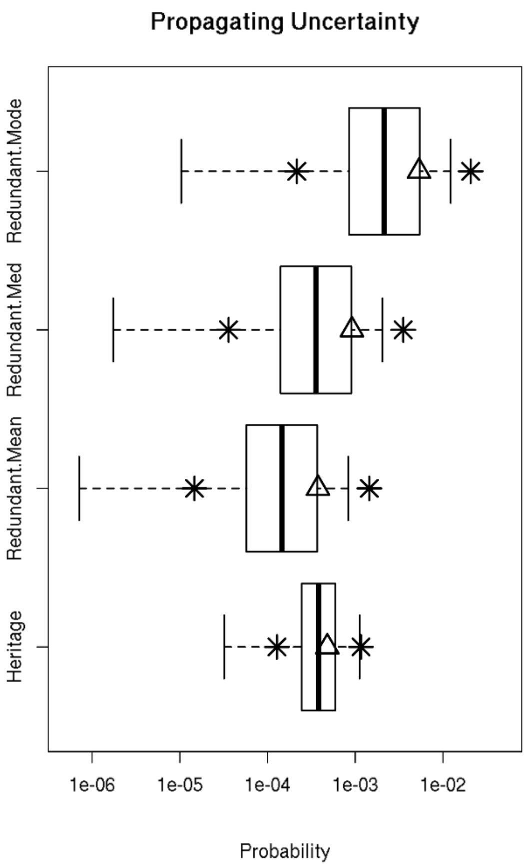
- The next set of results are similar comparisons, but allows us to view what is happening from a different perspective. The first case begins with holding the mean fixed to a value of 50 while varying the error factor between 3 and 9. In the other cases we hold the median and mode fixed while varying the error factor.







Uncertainty Estimation Example 3



This final example illustrates a hypothetical trade study. It compares a highly reliable, heritage, zero failure tolerant design with a retrofitted redundant option that is not only susceptible to common cause failure modes but is such that each leg of redundancy is considered less reliable than the heritage design.

The heritage design has a well-established lognormal failure rate distribution with a mean failure rate of 50 FPMH and an error factor of 3. The legs of the redundant design are estimated (through engineering judgement) to have failure rates of 150 and 200; and associated error factors of 6 and 9, respectively. A common cause failure basic event is also assumed to follow a beta distribution with a 5th percentile of 0.1 and a 95th percentile of 0.4. The time both options are exposed to failure is 8 hours.

Uncertainty Estimation Cheat Sheet

Lognormal Uncertainty Estimation Cheat Sheet

Uncertainty Increased	
Fixed	**Risk**
Mean	Lower
Median	Neutral
Mode	Higher

Uncertainty Decreased	
Fixed	**Risk**
Mean	Higher
Median	Neutral
Mode	Lower

The purpose of the cheat sheet is to reinforce an understanding of the cause and effect relationships between the adjusting of parameters (that measure central tendency and dispersion) and their risk implications. The cheat sheet is qualitative in nature and must be taken with a grain of salt. Ultimately, the choice of which measure of central tendency to hold fixed is subjective. However, it is important to understand and consider the risk implications of these choices within the context of the assumptions and beliefs of those involved in the estimation process.

Conclusion

- Although many cases are presented, the typical case for aerospace PRA is to assume the measure of central tendency is the mean and keeps it fixed while increasing the uncertainty (error factor).
- Unfortunately, this is often the case without strong rationale.
- Theoretical distributions do not always behave intuitively. Care must be taken when adjusting the parameters of a distribution as part of a heuristic or other method.
- Our recommended default for heuristic estimation of lognormal uncertainty is to quantify the median from the given data and then adjust the error factor accordingly.
- One ought to understand the relationships and effects of all relevant parameters as well as the risk implications.
- It is our hope that the Uncertainty Estimation Cheat Sheet (for Lognormal Uncertainty) will help those involved in the PRA process (such as managers, subject matter experts and PRA analysts) make effective technical contributions to decision-making.

Questions?

POC: Paul T. Britton

Paul.t.britton@nasa.gov

Office: +1-256-544-8301