Uncertainty Estimation Cheat Sheet for Probabilistic Risk Assessment

Paul Britton, National Aeronautics and Space Administration (NASA); paul.t.britton@nasa.gov

Mohammad Al Hassan; National Aeronautics and Space Administration (NASA);

Robert Ring; Bastion Technologies Incorporated, Huntsville, Alabama USA

ABSTRACT

"Uncertainty analysis itself is uncertain, therefore, you cannot evaluate it exactly," Source Uncertain

Quantitative results for aerospace engineering problems are influenced by many sources of uncertainty. 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. Uncertainty can be thought of as a measure of the 'goodness' of a result and is typically represented as statistical dispersion.

This paper will explain common measures of centrality and dispersion; and—with examples—will provide guidelines for how they may be estimated to ensure effective technical contributions to decision-making.

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. This paper will focus on a few fundamental examples and the principles paramount in achieving successful uncertainty estimation. After establishing definitions and providing background material, 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.

1. Definitions and Background Material

In an attempt to make later concepts accessible to a broader audience, this section provides a synopsis of some of the relevant concepts of probability theory. Here and in what follows, **boldface** indicates a word or phrase that is being defined or explained.

1.1 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). Probability distributions are either discrete, or continuous, or a mixture of both types. However, the topics, herein, require only some basic knowledge of continuous distributions.

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 $\boldsymbol{X}$ lies within an interval $[a, b]$:

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

In this paper we will discuss the following continuous distributions: The exponential, normal, and lognormal.

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

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

Assuming an exponential failure model for a component's mission exposure time, the probability of failure of the component before time T is given by the cumulative distribution function calculated as:

$$F(T, \lambda) = \int_0^T \lambda e^{-\lambda t}\,dt = 1 - e^{-\lambda T}$$

The exponential is a simple model with one parameter, and its properties are widely understood. It is commonly used to model components operating within their service life because of the following characteristics:

1) The hazard function (instantaneous failure rate) is constant with respect to time and is equal to λ, which implies:

2) The memoryless property; that is, components do not wear out: they function as good as new regardless of how long they have been in service.

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

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

Where the parameters μ and σ are the mean and the standard deviation, respectively. The normal pdf is symmetric about x = μ. An example of data that follows a normal pdf are repeated measurements of a physical characteristic of a part (such as, weight, length, thickness).

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

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

An important property of the lognormal is its relationship to the normal distribution. If ***Y*** is lognormal, then ***X*** = ln(***Y***) is normal. The lognormal is a good choice for representing failure rate uncertainty because it is strictly defined on the positive x axis, and the distribution has a heavy (stretched out) right tail (i.e., right skewed).

1.2 Central Tendency

A measure of **central tendency** (or **centrality**) is a single value that describes a central or typical value for a probability distribution or set of data. It may refer to the center of probability (median) or center of probability density (mean) or a most probable value (mode). The mean, median and mode are the most common examples and are defined below. In subsequent sections, we will look at the mean, mode and median, and explain some of the conditions for their appropriate usage.

In the case of a probability distribution $p(x)$ defined on a finite set $\{x_1, x_2, \ldots, x_n\}$, the **arithmetic mean** or expected value of x is a weighted sum:

$$E[\boldsymbol{X}] = \sum xp(x)$$

For a continuous distributions, the arithmetic mean is:

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

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

Noteworthy is the fact that the mean is susceptible to the influence of outliers. These are values that lie far from the central body of the distribution. The effect of large outliers is to pull the mean away from the median towards the outlier. This can be understood by considering that the mean is analogous to the center of mass.

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.

1.3 Dispersion

Dispersion refers to the spread of the distribution. A measure of dispersion is a non-negative real number that quantifies the deviation from the central tendency. A large deviation (relative to the magnitude of the central tendency) is indicative of a distribution that is spread out or dispersed. Examples to be discussed are the variance, the standard deviation, and the error factor.

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**:

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

A nice property of the standard deviation is that it has the same units as the quantity being measured.

A frequently used measure of dispersion for the lognormal is the **error factor (EF)**. The EF defines dispersion about the median. The 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:

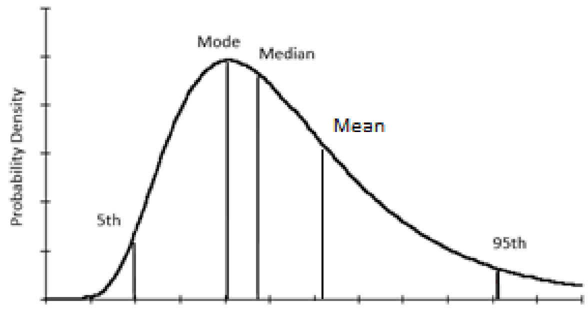
$$EF = \sqrt{\frac{95^{th}\ percentile}{5^{th}\ percentile}} = \frac{95^{th}\ Percentile}{50^{th}\ Percentile} = \frac{50^{th}\ Percentile}{5^{th}\ Percentile}$$

The EF has a minimum value of one, which represents certainty.

1.4 Failure Rate Uncertainty

Component failure rates (λ) are not physical quantities; that is, they cannot be measured directly but must be inferred. Operational or test reliability data expressed as total exposure time and number of failures from which to infer the failure rate for highly reliable components, are scarce. Failure rates have to be inferred by similarity to generic sources, such as those published in component reliability databases. In some cases, estimates are developed using engineering judgment or by eliciting the estimates from subject matter experts.

In the Bayesian interpretation, the parameter value is random (synonymous with uncertain) and is represented by a probability distribution. 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. Basic Event quantification for PRA then becomes one of estimating the central tendency and dispersion of the lognormal distribution, and then using Monte Carlo methods to simulate the probability distribution for failure of the component. The lognormal failure rate uncertainty pdf is illustrated in the figure below.



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.

2. The Bayesian Approach

Application of classical life data analysis requires component data in the form of failures and exposure time or number of demands. The data is fit to a distributional model of times to failure, goodness of fit tests are performed, and measures of centrality and dispersion are estimated. The model's parameters are assumed to be fixed (but unknown) quantities. The neighborhood in which the parameter value lies is estimated from the sampling distribution and is expressed as a confidence interval. Classical prediction methods rely solely on the data and do not permit prior knowledge to influence the estimate.

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 methods.

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 parameter in the form of a prior distribution. Specific data in the form of a likelihood function is then applied through Bayes Theorem to update the prior belief to yield the posterior uncertainty pdf.

Bayesian updating produces normative results in that if one believes the prior distribution, then one ought to believe the posterior distribution. It is important to note the implication of the previous statement. If after Bayesian updating, one does not believe the posterior distribution, then the prior is likely wrong. Hence it is important for the prior distribution to be developed and reviewed in a deliberative process with the help of subject matter experts to assure credibility.

In our experience, engineers with specific discipline expertise are familiar with the shape and properties of the normal probability distribution, but have little direct experience with skewed distributions, such as the lognormal. Recall, the normal probability density function is symmetric about its global maximum (mode). The median (50^{th} percentile), and mean coincide with the mode. Perhaps due to the coincidence of measures of central tendency of the normal distribution, subject matter experts do not understand the relationship of these measures when using the lognormal pdf. 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 central tendency and dispersion measures are varied. 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	$\exp(1.64485\sigma)$

3. Fundamental Examples of Uncertainty Estimation

There is strong evidence that we perceive things logarithmically [2]. In other words, when we think something is twice as big, it might be more like eight times as big.

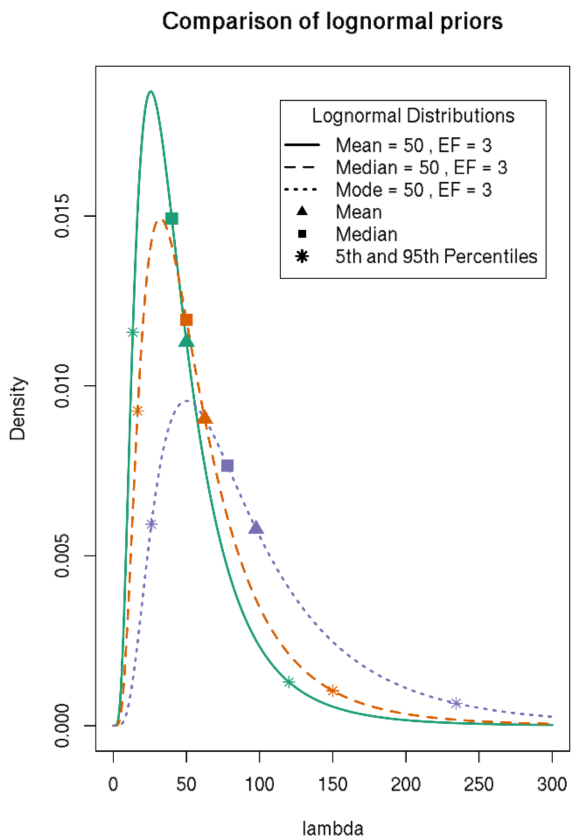
PRA analysts are often in the position of asking experts to estimate the percentiles of the uncertainty pdf. Experts are prone to overconfidence bias. For example, if we elicit the 5th and 95th percentiles, experts are likely to give answers that are closer to the 25th and 75th percentiles. This is especially true for rare events.

Component reliability data developed to support reliability allocation goals are an important source of data to help develop prior uncertainty estimates for use in the PRA. 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. The heuristic method proscribes using the provided point estimate, which is typically assumed to be an estimate of the mean failure rate and then, depending on data source applicability, apply an assumed error factor based on heuristic guidelines. The resultant prior distribution is then assumed to be a lognormal distribution with the provided mean and the heuristic error factor.

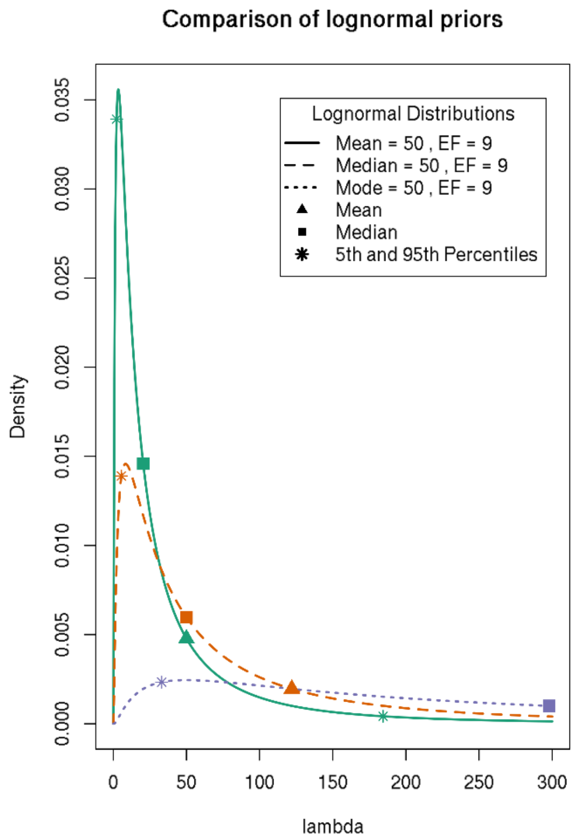
Clearly, any recipe for developing a prior needs to be augmented by visualization to confirm that the resulting distribution is credible. The discussion that follows illustrates a concern that when the point estimate is assumed to be the mean failure rate, the resultant pdf is counterintuitive – and is in fact non-conservative. PRA methodology tries to achieve the best estimate of the risk. But, in questionable cases should err on the side of conservatism.

The discussion that follows compares several modifications to the aforementioned heuristic approach. The comparison cases assume the point estimate represents one of the measures of central location (i.e. the mean, median or the mode) and is fixed, while the error factor is applied or varied. We then illustrate with a hypothetical trade study in which the recommended alternative, from a risk standpoint, depends on the measure of central location that is held fixed.

The figures below 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. The solid green curve is the base case assuming the point estimate is the mean; the dashed orange curve assumes the point estimate is the median; and the dotted purple curve assumes the point estimate is the mode. The square point on each of the resultant lognormal pdf curves is the median (50th percentiles) and the triangle is the mean.



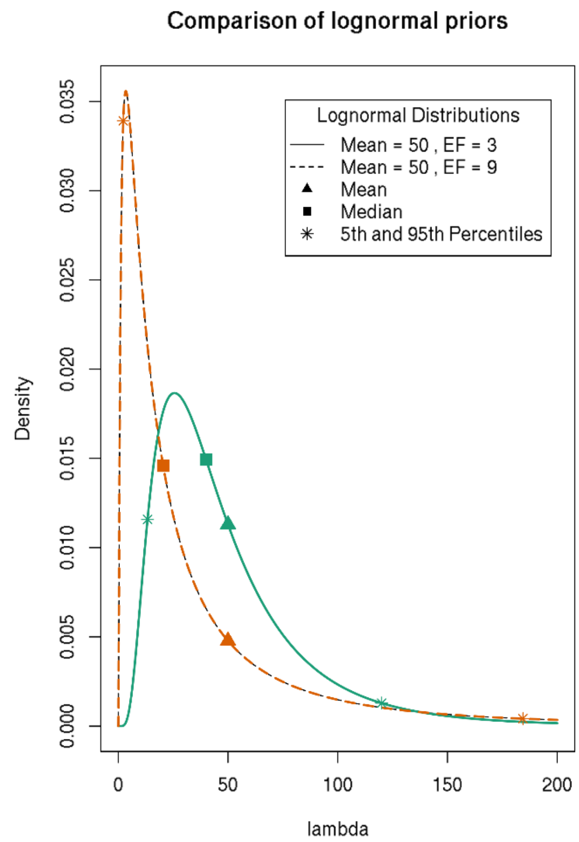
Central Value = 50, Error Factor = 3			
Fixed →	Mean	Median	Mode
Mean	50.00000	62.49422	97.62917
Median	40.00370	50.00000	78.11056
Mode	25.60710	32.00592	50.00000
5th	13.33457	16.66667	26.03685
95th	120.0111	150.0000	234.3317

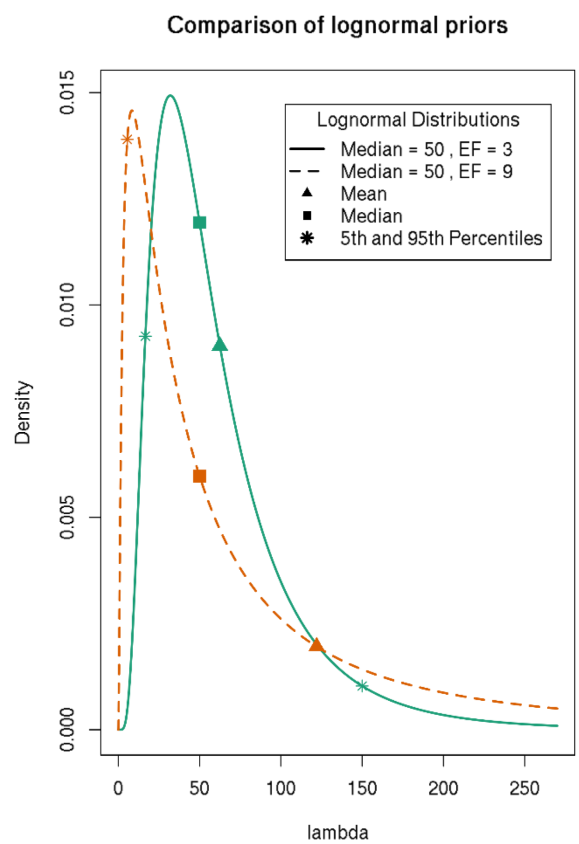


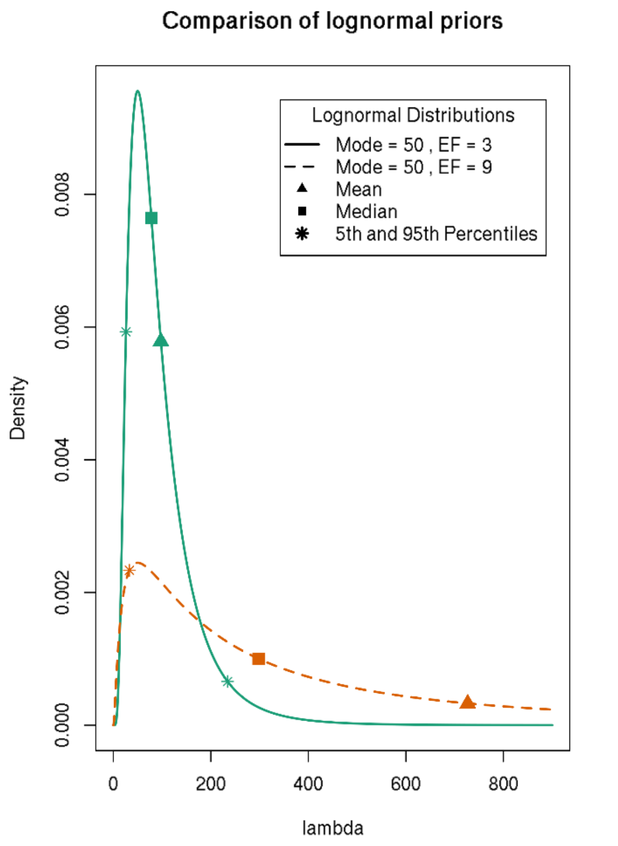
Central Value = 50, Error Factor = 9			
Fixed →	Mean	Median	Mode
Mean	50.00000	122.0252	726.7891
Median	20.48757	50.00000	297.80292
Mode	3.439787	8.394814	50.00000
5th	2.276397	5.555556	33.089213
95th	184.3882	450.0000	2680.2263

Notice that when the mean is assumed fixed, the resultant lognormal pdf is forced to the left. This may not be the desired result.

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.





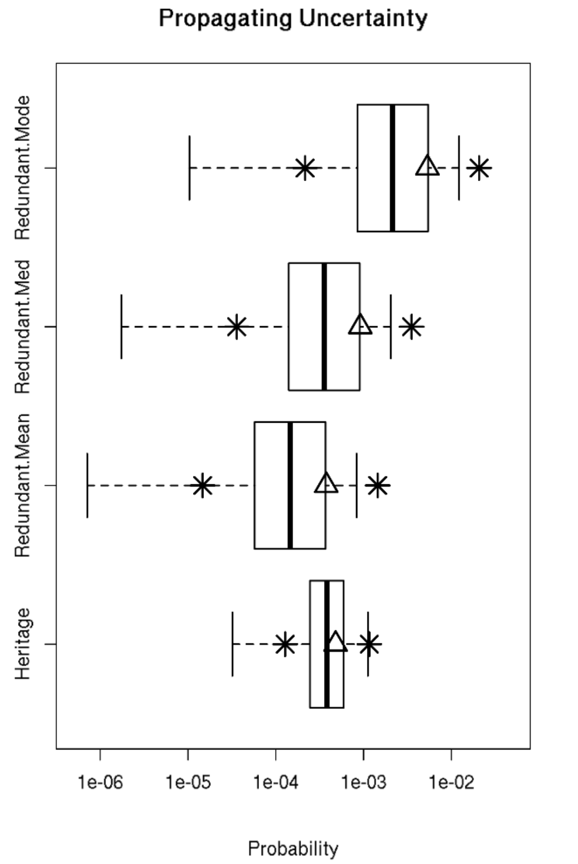


We observe that the general shape of the density curves are quite sensitive to holding the mean and mode fixed. Therefore, we strongly discourage heuristics that do not involve steps that require analysists and subject matter experts to assess visualizations of the resulting density curves.

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 5^{th} percentile of 0.1 and a 95^{th} percentile of 0.4 [3]. The time both options are exposed to failure is 8 hours.

The following figure compares the risk of the heritage option with three alternative methods of quantifying the uncertainty of the risk of the redundant design option: by assuming the provided failure rates are fixed and represent either the mean, median or mode.



Option -->	Redundant			Heritage
Fixed -->	Mean	Median	Mode	NA
Mean	3.77E-04	9.18E-04	5.34E-03	4.80E-04
Median	1.46E-04	3.55E-04	2.14E-03	3.83E-04
5th	1.47E-05	3.59E-05	2.16E-04	1.28E-04
95th	1.45E-03	3.53E-03	2.07E-02	1.16E-03

These results demonstrate that the effects on risk-informed decisions by the mere choice of the central parameter about which uncertainty is estimated can, in fact, be pivotal!

4. Uncertainty Estimation Cheat Sheet (for Lognormal Uncertainty)

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.

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

5. 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. 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. The median automatically remains fixed since it is independent of the error factor. Our recommendation holds even in the case where data uncertainty is absent and only a central value is given. In this case, the data tells us the mean, median and mode coincide. Therefore, we proceed by fixing the median to the given central value while estimating uncertainty about the median to obtain results.

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. 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.

6. References

[1] Parameter Uncertainty for ASP Models, Knudsen, James K. and Smith Curtis L., Idaho National Engineering Laboratory, October 1995

[2] Why do we perceive logarithmically? Authors Lav R Varshney, John Z Sun Publication date 2013/2/1 Journal Significance Volume 10 Issue 1 Pages 28-31

[3] Common Cause Failure Modeling. 8th IAASS Conference. Authors: Frank Hark, Steven Novack, Rob Ring, Paul Britton