Review of Draft Report on LOCA Frequency Estimates by Expert Elicitation

Corwin L. Atwood, Statwood Consulting
November 2, 2004

TABLE OF CONTENTS

SUMMARY

This is a review of the draft report on a task to estimate frequencies of various kinds of loss-of-coolant accidents (LOCAs) through elicitation of expert judgment and aggregation of the judgments of the experts. As requested during a two-day meeting when the report methodology was presented, the review focuses primarily on the analysis of the elicitation responses and the quantitative results. The review considers comments made at that meeting and at a meeting on Sept. 29, 2004, but it does not review the final report, which has not yet been completed. The primary conclusions of this review are the following, with item 1 being the most important.

1. There are two possible criteria for the aggregate distribution. Criterion 1 looks for a compromise consensus, a distribution that is "not too wide" but that "every panelist can live with." Criterion 2 looks, instead, for a complete description of both the individual panelists' uncertainty and also of all the variability between panelists. Even if the report authors choose to advocate one criterion, both should be stated explicitly, so that readers can see the choice that was made and possibly consider the effect of using the other criterion. I see three possible aggregation methods, though variations of each method can be considered as sensitivity studies.
   (a) Construct an aggregate distribution by taking the geometric means of the percentiles of the individual distributions. This is the report's "baseline" method. The aggregate distribution's median is the geometric mean of the individual medians, and its error factor is the geometric mean of the individual error factors (as shown in Sec. 1.2.1 below). Thus, this aggregate distribution can be thought of as the distribution of a single hypothetical representative panelist. This distribution is put forth in the report as satisfying Criterion 1. It certainly meets the condition of being "not too wide", but I am concerned that it discounts the outlying panelists too completely, by assigning no additional variance at all to reflect differences among panelists.
   (b) Aggregate all the panelists' distributions by using the arithmetic mean of the individual distributions. This can also be called an equally weighted mixture of the individual distributions. This distribution satisfies Criterion 2, accounting for within-panelist and between-panelist uncertainty in a mathematically quantified way.
   (c) Aggregate most or all of the individual distributions, in one of the above ways. Then, in addition, show the outlying distributions. For example, if method (a) is used for aggregation, the reader of the report will see the compromise consensus, but will also see the outliers and the extent to which they disagree with the consensus. This is analogous to a commission of various members that publishes a majority report but with individual dissenting opinions. I recommend that the results be presented in this way.

2. The procedure to elicit opinions seems to have been well planned and well executed. In particular, the meetings of the panelists allowed potential misunderstandings to be corrected and allowed all issues to be presented and discussed. However, the panelists gave their opinions and conclusions in private interviews, free from the influence of others.

3. The split lognormal distribution is a good way to construct a distribution from the three values given by a panelist. However, because the upper tail of the distribution is so very long when the error factor is large, the split lognormal is unrealistically conservative in some

cases. In the cases when the upper tail goes into the incredible region, the split lognormal should be regarded as giving an upper bound on the mean. A lower bound would be given by a split log-triangular distribution.

When some panelists' means have substantial uncertainty, the overall mean is also uncertain. This uncertainty must be presented as part of the final conclusions of the project.

Finally, we give some exact formulas to replace approximations used in the report. Exact formulas for the mean of the split lognormal and split log-triangular should be used. Also, when geometric means of the percentiles are used, the mean of the aggregate distribution is different from (and apparently always smaller than) the geometric mean of the individual means; in this case the exact formula should be used.

4. The methods of approximate error propagation — using only the upper tail of the distribution for the 95th percentile and median, and treating the result at every step as being split lognormal — are probably adequate for the panelists who report low frequencies, or for small contributors to a panelist's total frequency. However, for some cases, including cases of panelists who report frequencies so high that they influence the conclusions, more care should be taken. A few of these should be analyzed as test cases. For these test cases the approximation should be compared with nearly exact simulations, and the analysts should then be guided as to whether the approximations are good enough in every case.

5. Other minor details, such as constructing a hypothetical new panel member and constructing a confidence interval for the overall mean, are also discussed in this review.

The above topics are treated more fully in Sections 1 through 5 below. Summary recommendations and conclusions are scattered through the report in **bold face**.

1. AGGREGATION OF DISTRIBUTIONS OF VARIOUS PANELISTS

1.1 Criteria for Aggregate Distributions

Aggregation of panelists' distributions is controversial when those distributions differ greatly, so that each panelist denies the correctness of some other panelists. The panelists' distributions cannot all be correct. We reviewers were specifically asked to review the section on aggregating the distributions of the various panelists.

It became clear during the discussions of this topic that there is no universally accepted criterion of what makes a good aggregation. This disagreement was at the root of some passionate discussions about which aggregation method to use. The two criteria that have been advocated are the following.

Criterion 1. Construct a compromise consensus distribution. This should be a distribution that every panelist can live with, but it should not be much wider than the distribution of a typical panelist. Those who desire this kind of distribution include the authors of the report and one of the reviewers.

Criterion 2. Construct a distribution that accurately represents all the variation and differences among the panelists. If the panelists disagree greatly, the resulting distribution must reflect this, and must necessarily be considerably wider than the distribution of any single panelist. Those who have desired this kind of distribution include one NRC staff member and one of the reviewers (me).

The criteria can be compared in terms of the weight assigned to each panelist's opinion. Criterion 2 corresponds to assigning an equal vote to each of the n panel members. Each gets a weight of $1/n$ on the probability that λ is in any prespecified range. The average of the panelists' probabilities is the mixture probability. In contrast to this, Criterion 1 says that outliers should be discounted to some extent. Thus, outliers are given less than $1/n$ voting weight, and the panelists who agree in the middle are each given somewhat higher weights. (The calculations may not actually involve weighting of the probabilities, but the effect of discounting outliers is the same.) The crux of the difference between the criteria is the issue of whether it is wise or mistaken to downweight outliers.

I do not know which kind of aggregate distribution will be desired by the final decision makers. I think that **the above two criteria for a distribution should be presented clearly and explicitly in the report, so that the decision makers can be made aware of their own assumptions and preferences and therefore act appropriately.** Decision making is much more rational when the decision makers are aware of their own assumptions and desires, when these are explicit rather than implicit.

As a result, there are three possible classes of aggregate distributions that should be presented in this report.

Class *a*. Present the aggregations that are based on taking geometric means of the percentiles. This includes what is currently called the "baseline distribution" and nearly all of the sensitivity studies. These are the aggregations that are consensus compromises, intended to meet Criterion 1.

Class *b*. Present the aggregation that is based on the mixture (unweighted average) of the panelists' distributions. This is the mathematically justified distribution that accounts for all the variation among the panelists, satisfying Criterion 2.

Class *c*. As a third alternative, present both the aggregate distribution and any outliers. The aggregation may include all the panelists or only those panelists who are in substantial agreement. However, the outliers are presented in either case. Let the decision makers wrestle with outlying opinions and deal with them as they wish. This class of aggregation is for those who do not wish to choose either Criterion 1 or Criterion 2. It is also for those who worry that important information is lost when only the summary distribution is presented.

Because the above three classes reflect the criteria for what constitutes a good distribution (first criterion, second criterion, or less than full endorsement of either), **I recommend that the final report be structured to follow the outline given by the above three classes.** It would be misleading to present a distribution from one class as THE baseline, because the distributions in the different classes are attempting to meet different criteria. Furthermore, **I recommend that**

the final results consist of the aggregate distribution (such as the aggregate by geometric means) together with any outlying distributions, so that the readers can see the extent to which panelists dissented from the "consensus." Appendix A of this review works out some examples of the three kinds of estimate, and compares them to the two criteria.

The relations between the criteria and the first two classes of distributions has simply been stated above. Mathematical derivations are given in the next section.

1.2 Relations Between Criteria and Aggregation Methods

1.2.1 Criterion 1 and Geometric Means of Percentiles

To develop the properties of geometric-mean aggregation, suppose that each panelist provides some distribution, not necessarily lognormal or of any other parametric form. Denote the median and 95th percentiles of the ith distribution by m_i and $\lambda_{0.95,i}$, and denote the error factor by $\mathrm{EF}_i = \lambda_{0.95,i}/m_i$. We construct the aggregate distribution by taking geometric means of all the percentiles. In particular, the aggregate median m is

$$m = \exp[(1/n)\Sigma \ln(m_i)],$$

and the 95th percentile is

$$\begin{aligned}\lambda_{0.95} &= \exp[(1/n)\Sigma \ln(\lambda_{0.95,i})] \\ &= \exp[(1/n)\Sigma \ln(m_i \times \mathrm{EF}_i)] \\ &= \exp[(1/n)(\Sigma \ln(m_i) + \Sigma \ln(\mathrm{EF}_i))] \\ &= m \exp[(1/n)\Sigma \ln(\mathrm{EF}_i)]\end{aligned}$$

Therefore, the error factor of the aggregate distribution, $\lambda_{0.95}/m$, equals the geometric mean of the individual error factors.

This shows that the geometric means of the percentiles form a distribution whose median is the geometric mean of the individual medians, and whose error factor is the geometric mean of the individual error factors. It is as if we were constructing a distribution for a hypothetical "representative" panelist, one who has a typical median and a typical error factor. The percentiles of the aggregation are the percentiles of the uncertainty distribution for this hypothetical representative panelist.

Such a distribution certainly meets the Criterion 1 condition of being not too wide, because the error factor is within the range of error factors of the individual panelists. The only question is whether it is wide enough so that "every panelist can live with it." I am bothered by the fact that this method may completely discount outliers. Section A.2.2 gives an example in which the geometric-means distribution is identical with the distribution of the central panelists, and the outlying panelists seem to have no effect at all. It appears to me that the compromise consensus constructed in this way can be too concentrated.

For the split lognormal distribution, we can say, in addition, that the aggregate distribution is also split lognormal. The formulas are as follows. The ith underlying normal distribution has median $\mu_i = \ln(m_i)$ and $\sigma_i = \ln(\mathrm{EF}_i)/1.645$. (More precisely, if the distribution is split it has both an upper and a lower σ_i.) The 100pth percentile of the split lognormal distribution is $\exp(\mu_i + z_p\sigma_i)$, where z_p is the 100pth percentile of the standard normal distribution. It is easy to calculate that the

100pth percentile of the aggregate distribution equals $\exp(\bar{\mu} + z_p \bar{\sigma})$, where $\bar{\mu}$ is the arithmetic mean of the μ_i values and $\bar{\sigma}$ is the arithmetic mean of the σ_i values — more precisely, with split distributions use upper or lower σ_i values depending on whether $p > 0.5$ or $p < 0.5$. It follows that the aggregate distribution is split lognormal.

Similar results can be shown to hold for the split log-triangular distribution. For any panelist's distribution, let λ_p be the 100p percentile of a log-triangular distribution with median M, 95th percentile U, and 5th percentile L. That is, $\Pr(\lambda \le \lambda_p) = p$. For $p \ge 0.5$, it can be shown that

$$\ln(\lambda_p) = \ln(M) + 1.462\left(1 - \sqrt{2 - 2p}\right)\ln(U / M) \ .$$

The expression (U/M) is the error factor. The average of the $\ln(\lambda_p)$ values for several such distributions is of the same form, but with the average of the $\ln(M)$ values and the average of the $\ln(U/M)$ terms. A similar result holds for $p \le 0.5$. Therefore, using geometric means to aggregate split log-triangular distributions results in a split log-triangular distribution having the geometric mean of the medians and the geometric mean of the error factors.

Incidentally, the mean of the aggregate distribution is not equal to the geometric mean of the individual means, unless the individual error factors are all equal. This is discussed more fully in Section 3.2.3, where bounds on the aggregate mean are given for lognormal and log-triangular distributions.

1.2.2 Criterion 2 and Mixture of Individual Distributions

In the general setting, we present a scenario to some experts, and each expert uses his or her own judgments, assumptions, and perhaps even a small PRA analysis to arrive at a distribution for λ. Each distribution is conditional on the judgments and analyses of one expert. This may be regarded as using various "models," with each model consisting of the judgments, assumptions, and analyses of a particular expert. It is not obvious that the models are mutually exclusive, or exhaustive of the logical possibilities. However, when n experts have been interviewed, one can consider statements of the form "expert i has the best approach of the n under consideration", which we abbreviate as "approach i is best" in the equations below. These statements are mutually exclusive statements, and they are exhaustive within the universe of the n experts.

Let us assign probability $1/n$ to each of the n statements "approach i is best". This treats the experts as all having equal credibility, and uses a subjective probability to express this. The probability describes epistemic uncertainty, lack of knowledge, rather than aleatory randomness. That should not be a problem, because each of the distributions for λ given by the experts also is a representation of epistemic uncertainty.

Now suppose that we are willing to treat all these epistemic probabilities together, following the usual laws of probability. Some people are not willing to treat model uncertainty in this way, especially if the models give qualitatively different conclusions. However, here in this subsection we do so. Then we get

$\Pr(\lambda \text{ in } A) = \Sigma_i \Pr(\lambda \text{ in } A \mid \text{approach } i \text{ is best}) \Pr(\text{approach } i \text{ is best})$.

Because we give equal probability to all n models, we get

$$\Pr(\lambda \text{ in } A) = \sum_{i=1}^{n} \Pr(\lambda \text{ in } A \mid \text{approach } i \text{ is best}) / n \ .$$

This is the mixture aggregation of the probabilities. In particular, if A is the interval [0, a], each probability is a cumulative distribution function (c.d.f.) evaluated at a, and the equation is

$$F(a) = \sum_{i=1}^{n} F_i(a)/n$$

where F_i is the c.d.f. for expert i and F is the c.d.f. for the group as a whole.

If we do this with the entire population of all experts, the resulting F is the aggregate distribution for the population, and $F(a)$ is the mean over the entire population of values $F_i(a)$. If instead, we take a random sample of experts, the resulting sample mean is an unbiased estimator of the population mean $F(a)$. The unbiasedness can be shown from first principles.

Similarly, it can be shown that the density for the population is the mean of the densities of the experts, and the mean for the population is the arithmetic average of the experts' means. Also, the mean of a random sample of densities is an unbiased estimator of the population density, and the arithmetic average of the means of the sampled experts is an unbiased estimator of the population mean.

The 100pth percentile of the mixture distribution, λ_p, has the following interpretation. Let Pr_i denote the probability according to the ith panelist. Then
average[$\mathrm{Pr}_i(\lambda \le \lambda_p)$] $= p$,
where the average is over all the panelists. Another way of saying this is that if a panelist is picked at random and asked to give a random λ, there is exactly a probability p that this λ will be $\le \lambda_p$.

The mixture aggregation has been constructed to include equal parts of each distribution. Even if several distributions strongly disagree, the aggregate has been constructed to contain some of each rather than discarding both and possibly choosing only the middle ground. Therefore, its spread is wider than that of any individual distribution, satisfying Criterion 2. Examples are worked out in Appendix A of this review.

2. PROCEDURE FOR ELICITING OPINIONS

The panelists were chosen to represent a diversity of methodologies and professional affiliations.

Great effort was made ensure that
- the questions were clear to the panelists,
- all the panelists understood the technical issues, through group discussion
- the panelists' qualitative rationales were clearly explained to the analysts in private
- the panelists' judgments were given in private, not influenced directly by any other panelists
- inconsistencies were challenged by the analysts, and the answers reworked if necessary
- the methods and conclusions of each panelist were presented to the other panelists for group discussion, and each panelist had the opportunity to modify his/her work in private.

In addition, some base cases were worked out by four panelists. For the cases that were more difficult than the base cases, such as some type of LOCA frequency after 40 years, the panelists were told to only present the ratio of their answers to the base-case frequencies. This is a good idea, because people like to anchor on some value, and the process made this explicit. Unfortunately, however, the four base-case panelists were far from unanimous in their answers. Therefore, each panelist was allowed to decide which base case to anchor on when presenting a ratio, and some panelists did not use any of the base-case results. I agree that it was good to give the panelists freedom to answer the questions using any method they chose, and do not see how the analysts could have dealt better with the base-case results that they received. My only request is that the report clearly explain what was actually done — I had to ascertain some of the above information by telephone conversation.

In summary, the expert elicitation seems to have been planned and executed very well, but some details need to be explained more fully in the report.

3. CONSTRUCTING DISTRIBUTIONS FROM PANELISTS' NUMBERS

3.1 Use of Split Lognormal

For each frequency or ratio, for example the ratio (LOCA frequency)/(baseline frequency) for some component when the plant is of a certain age, the panelists were asked to give three numbers: a lower bound L, an upper bound U, and a middle value M. These were interpreted as the 5th, 95th, and 50th percentiles of an uncertainty distribution for the estimated quantity X.

Because people think easily in terms of order-of-magnitude, the lognormal distribution was considered. However, $\ln(U/M)$ and $\ln(M/L)$ were not necessarily equal. Therefore a split normal distribution was assumed for $\ln X$: the upper tail of a normal distribution was used to the right of $\ln M$, and the lower tail of a *possibly different* normal distribution was used to the left of $\ln M$. The 5th percentile and 95th percentile were $\ln L$ and $\ln U$, respectively.

This seems reasonable and appropriate. However, a lognormal distribution can have a very large mean if the error factor is large. In some cases the error factor U/M was as large as 1000. In this case, the mean of X is approximately 6700 times the median, beyond the 98th percentile of the distribution. The value of the mean is strongly influenced by the portion of the tail beyond the 99.99th percentile; that is, if the distribution were truncated at the 99.99th percentile, the mean would be reduced by a factor of about 3.

Surely no panelist would be willing to state with confidence what the distribution is beyond the 99.99th percentile. Therefore, the mean of the distribution is not exactly knowable. It can be calculated assuming the lognormal form of the tail, but this form is an assumption, nothing more. In such a case, the analysis method should try to put bounds on the mean, and not treat the calculated mean as a known value.

3.2 Bounds on the Mean

3.2.1 Upper Bound

As an example, consider panelist C for a Category 6 LOCA in PWRs that are 25 years old. The reported distribution is split lognormal with μ = -20.39, σ_L = 2.13, σ_U = 4.20. Note, the upper error factor is slightly over 1000.

The mean of this split lognormal is 9.51E−6, which is greater than the 95th percentile, 1.4E−6. This mean is strongly dependent on the extreme right tail of the distribution. In particular, the mean is found as an integral, and (using Lemma 1 in Appendix B) it can be shown that about half of the total integral consists of the portion of the integral beyond the 99.998th percentile.

Let us examine whether this extreme portion of the upper tail is credible. Suppose that the frequency were actually equal to the 99.998th percentile, λ = 4.36E−2/reactor-yr. The inverse, $1/\lambda$, would be 23 reactor-yrs per LOCA. Thus, in an industry with about 69 PWRs of average age 25 years, we would be seeing about three Category 6 LOCAs per year. That is not happening. Therefore, we conclude with high confidence that the distribution of λ should not extend so far on the right. Truncating the lognormal at the 99.998th percentile, for example, reduces the mean by a factor of 2. Thus, either the split lognormal distribution or the truncated split lognormal distribution provides an upper bound for the panelist's actual mean.

In this example, numerical calculations showed that the tail of the distribution extends into the incredible region. With other examples that will not necessarily be the case. However, it is only crucial to know the mean for those panelist distributions with the largest means. These largest means can be important in two ways.

- The largest means may show up as outliers.
- In addition, if aggregation is based on a mixture of the individual distributions, the largest means can strongly influence the mean of the mixture distribution.

Each case must be examined on its own, but there are grounds to think that the distributions *that matter* can be treated in a way similar to the example of this section.

One might point out that we usually worry about *over*confidence, and ask whether we should be doing so here, lengthening instead of shortening the tail of the distribution. The answer is no, panelist overconfidence is not a concern when the error factor is already so very large.

3.2.2 Lower Bound

For the lower bound on the mean, a short-tailed distribution can be used, and I suggest a split log-triangular distribution. This distribution is based on a split triangular density for ln$X$, defined as follows. The density of ln$X$ is zero outside of some range. Within the range, the density increases linearly until ln$X$ = ln$M$, then it has a possible discontinuity at ln$M$, and finally it decreases linearly to the right of ln$M$. (Full disclosure: The use of a split log-triangular is not my idea. It was suggested at the meeting we had, and I like it.)

3.2.3 Bounds on Aggregate Distribution

If the aggregate distribution is constructed as a mixture, the mean of the mixture distribution is the average of the individual means. If the aggregate distribution is constructed using the geometric means of the individual percentiles, the mean of the aggregate distribution is *not* the same as the geometric mean of the individual means, unless all the individual error factors are identical.

The second statement is shown for the lognormal distribution as follows, using the notation $\bar{\mu}$ and $\bar{\sigma}$ for the averages of the μ_i values and the σ_i values, as introduced in Section 1.2.1.. The mean of the ith lognormal distribution is $\exp(\mu_i + \sigma_i^2/2)$, and the mean of the aggregate is $\exp(\bar{\mu} + \bar{\sigma}^2/2)$. Because it is always true that

$$(1/n)\Sigma(\sigma_i^{\ 2}) \geq [(1/n)\Sigma\sigma_i]^2 \ ,$$

the geometric mean of the individual means is an *upper bound* on the mean of the aggregate distribution. More exactly,

$$(1/n)\Sigma(\sigma_i^{\ 2}) = [(1/n)\Sigma\sigma_i]^2 + (1/n)\Sigma(\sigma_i - \bar{\sigma})^2 \ .$$

That is, the mean of the squared σ terms equals the square of the mean σ plus the variance of the σ terms. If the panelists have widely different error factors, the geometric mean of the individual means will differ substantially from the mean of the aggregate distribution.

As for the mean of the log-triangular, I have not been able to develop a general formula. However in several examples the geometric mean of the individual log-triangular means is an upper bound on the mean of the aggregate distribution, just as for the lognormal distribution.

The above arguments are refer to the unsplit distributions. However, the upper half of a split distribution dominates the mean. Therefore, the above results for unsplit distributions probably hold, at least to good approximation, for the split distributions.

I have calculated one real example, for Category 6 LOCAs at age 25 years. The results are shown in Table 1.

Table 1. Approximate and exact means of aggregated distribution, assuming two distributional forms for the individual distributions.

	split log-triangular	split lognormal
geometric mean of means	9.53E-9	2.54E-8
exact mean of aggregate	5.24E-9	7.49E-9

In summary, the mean of the aggregate distribution is apparently bounded above by the geometric mean of the individual means. Therefore, I suggest that means of aggregated distributions be calculated exactly, when using geometric means to combine split lognormal or split log-triangular distributions. This is done as follows.

1. Assume either split lognormal or split log-triangular distributions.
2. Calculate the geometric means of the medians and of the upper and lower error factors.

3. Calculate the mean of the lognormal distribution or the log-triangular distribution having that median and those error factors, using Equation (1) or (2) in Section 3.3.
4 Use the mean of the aggregate split lognormal distribution as an upper bound on the aggregate mean, and the mean of the aggregate split log-triangular as a lower bound on the aggregate mean.

If the distributions to be aggregated have some other form, not a neat parametric form but something that resulted from earlier calculations or simulations, then it seems likely that the same relation holds, that the geometric mean of the means is an upper bound on the mean of the aggregate. If this is unacceptable, the exact mean of the aggregate must be calculated directly:

1 Calculate the 100*p*th percentile of each of the contributing distributions, for many values of p between 0 and 1.
2 For each of these values of p, calculate the percentile of the aggregate as the geometric mean of the individual percentiles.
3 Calculate the mean of the aggregate distribution numerically.

3.2.4 Application of the Bounds

The mean is only highly uncertain when the upper error factor is large. Also, as discussed in Section 3.2.1, the mean for one panelist only is very important if that one panelist's mean is one of the very largest of all the panelists' means. Therefore, this bounding procedure is very important only when both these conditions hold for a panelist: large error factor and large mean. Table 2 shows certain cases that satisfy these conditions, based on the Excel spreadsheet corresponding to the draft report. A lognormal mean is greater than the 95th percentile whenever EF $\geq$ 224. Therefore the Table shows all the cases with upper EF that large. If the mean is the largest of all the panelists' means for this category and plant age, the Table also shows the ratio of the mean to the next largest mean.

The Table shows in bold face all the cases for which the mean exceeds the other seven or eight means by a factor of 10 or more. These bold-face cases are the ones where it is most critical to show the uncertainty in the mean. For those cases, the entire range between the (possibly truncated) lognormal mean and the log-triangular mean could be displayed when presenting the uncertainty in the means. Figure A.4 is an example where this is done.

When the overall mean is found, the uncertainty in any panelist's mean will introduce a corresponding uncertainty into the overall mean. **When the final conclusions of the project are presented, the uncertainty in the value of the mean must be presented as part of the conclusions.**

Table 2. Cases with Large Error Factor and/or Large Mean

Reactor Type	LOCA Cat.	Panelist	Plant Age	Upper EF	Mean/Next
BWR	3	H	60	527	3.3
BWR	4	H	60	532	NA[a]
BWR	5	C	25	280	2.8
BWR	5	C	40	356	3.8
BWR	5	C	60	860	2.2
BWR	**6**	**C**	**25**	**1000**	**16.6**[b]
BWR	**6**	**C**	**40**	**1000**	**39.1**[b]
BWR	**6**	**C**	**60**	**1000**	**567.**[b]
BWR	6	H	60	1092	NA[a]
PWR	3	H	60	500	NA[a]
PWR	**4**	**J**	**60**	**263**	**5300.**[b]
PWR	4	H	60	500	NA[a]
PWR	**5**	**C**	**25**	**1004**	**14.5**[b]
PWR	**5**	**C**	**40**	**1005**	**33.4**[b]
PWR	5	C	60	1006	1.6
PWR	5	J	25	301	NA[a]
PWR	5	J	40	299	NA[a]
PWR	5	J	60	273	NA[a]
PWR	6	C	25	1004	5.0
PWR	**6**	**C**	**40**	**1004**	**11.7**[b]
PWR	**6**	**C**	**60**	**1005**	**24.4**[b]
PWR	6	H	60	500	NA[a]
PWR	6	J	25	1038	NA[a]
PWR	6	J	40	1109	NA[a]
PWR	6	J	60	1105	NA[a]

a. This mean is not the largest.
b. This mean exceeds the other seven or eight by a factor of 10 or more.

3.3 Exact Formulas for the Mean

Exact formulas are given below. **These exact formulas should be used, not approximations, even though the approximations may be close in most cases.** The derivations of these exact formulas are given in Appendix B.

3.3.1 Mean of Truncated and/or Split Lognormal

The truncated split lognormal distribution for X is defined as follows. The distribution of $\ln X$ is truncated split normal. That is, it is split at the median, μ. The upper tail is proportional to a normal(μ, σ_U^2) density, truncated at some upper bound b_U, that is, set to zero beyond b_U. We take b_U to be a high percentile, so that the normal area beyond b_U is a specified small number p. For example, if $p = 0.001$ the corresponding b_U is the 99.9th percentile of the normal distribution. The area under the normal density from μ to b_U is $(0.5 - p)$, so the density of the truncated split lognormal must be renormalized in the upper tail by dividing by $(1 - 2p)$, so that the integral from μ to b_U is 0.5.

Similarly, the lower tail of the truncated split normal is a normal(μ, σ_L^2) density, truncated at some lower bound b_L, and renormalized by dividing by $(1 - 2p)$. We require p to be the same in both tails, so if, for example, b_U is the 99.9th percentile of a normal(μ, σ_U^2) distribution, then b_L is the 0.1th percentile of a normal(μ, σ_L^2) distribution.

As a special case, we get a truncated lognormal (not split) when $\sigma_U = \sigma_L$. We get a split lognormal (not truncated) when $p = 0$, so $b_L = -\infty$ and $b_U = +\infty$.

The mean of a truncated split lognormal is

$$\frac{\exp(\mu)}{1-2p}\left\{\exp(\sigma_L^2/2)\left[\Phi(-\sigma_L)-\Phi\left(-z_{1-p}-\sigma_L\right)\right]+\exp(\sigma_U^2/2)\left[\Phi(\sigma_U)-\Phi\left(-z_{1-p}+\sigma_U\right)\right]\right\} ,$$

where z_{1-p} is the standard normal value corresponding to *upper* tail probability p. For example, if $p = 0.05$, then $z_{1-p} = +1.645$. In the special case when the distribution is not truncated, this becomes

$$\exp(\mu)\left\{\exp(\sigma_L^2/2)\Phi(-\sigma_L)+\exp(\sigma_U^2/2)\Phi(\sigma_U)\right\} . \tag{1}$$

In the special case when $\sigma_U = \sigma_L$, so the distribution is not split, the mean is

$$\frac{\exp(\mu+\sigma^2/2)}{1-2p}\left[\Phi\left(z_{1-p}+\sigma\right)-\Phi\left(-z_{1-p}+\sigma\right)\right] .$$

The derivation of this formula is given in Appendix B.

As a numerical example, consider Panelist C for the case of LOCA Category 6 in a PWR. His values were
$L = 4.18\text{E}{-11}$, $M = 1.39\text{E}{-9}$, $U = 1.40\text{E}{-6}$.
Therefore, the upper error factor is approximately 1007, and the lower error factor is approximately 33. Table 3 shows some means that can be found:

Table 3. Exact means for several distributions obtained from L, M, and U.

Distribution	Mean
Given by draft report	9.0E−6
Split lognormal	9.5E−6
Lognormal matching the upper tail	9.5E−6
Split lognormal, truncated at 0.01 and 99.99 percentiles	3.0E−6
Split lognormal, truncated at 0.1 and 99.9 percentiles	1.3E−6
Split log-triangular (defined below)	3.3E−7

To more significant figures, the split lognormal has mean 9.51E−6, whereas the lognormal matching the upper tail has mean 9.49E−6. That is, the lower tail has negligible influence on the mean.

3.3.2 Mean of Split Log-triangular

The split log-triangular distribution is defined for X as follows. Let X have 5th, 50th, and 95th percentiles denoted L, M, and U. The density of $\ln X$ is discontinuous at $m = \ln(M)$. To the right of m the density of $\ln X$ falls linearly until it equals 0 at some point $m + b_U$. The right side of the density of $\ln X$ is shown in Figure 1. The 95th percentile, $m + q$, is equal to $\ln U$, so $q = \ln(U/M)$. To the left of m, the density of $\ln X$ is zero at some point $m - b_L$ and it rises linearly. The left side is not shown in the figure.

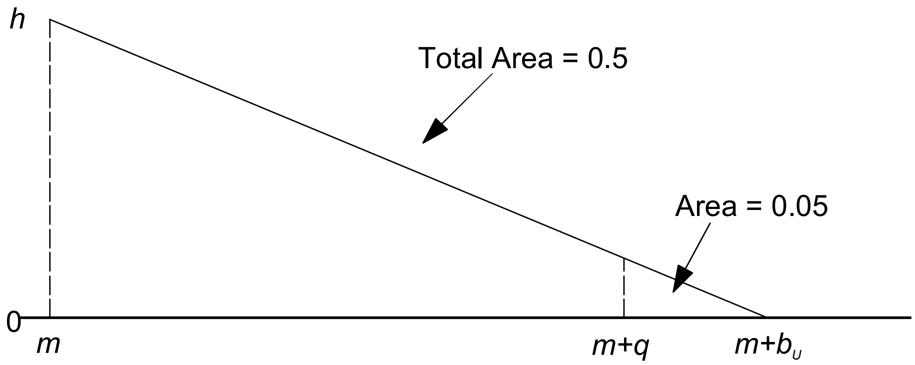

Figure 1. Right half of a split triangular distribution

The following facts can be shown:

$b_U = 1.462\ln(U/M)$
$b_L = 1.462\ln(M/L)$

The slope of the density of $\ln X$ is $-1/b_U^2$ on the right of m, and $+1/b_L^2$ on the left of m.

The mean of X, i.e. the expectation of $\exp(\ln X)$, is

$$M\times\{\ [\exp(-b_L) - 1 + b_L]/b_L^2 + [\exp(b_U) - 1 - b_U]/b_U^2\ \}\ . \qquad (2)$$

The derivations of these formulas are given in Appendix B of this review. An example application of the formula for the mean was given above in Table 3. In that example, the mean of the log-triangular distribution was smaller than the mean of any of the distributions based on a lognormal. This is to be expected, because the log-triangular distribution has a finite range.

4. APPROXIMATE AND (NEARLY) EXACT PROPAGATION OF UNCERTAINTY

The draft report makes many approximations, most of which are probably good enough. They treat a product of two independent split lognormals as if it were also split lognormal, and find the parameters of the product by using only one tail at a time. They treat a sum of split lognormals (independent or perfectly dependent) as if it were split lognormal, again finding the parameters of the sum using only the upper tails or only the lower tails.

A more exact method is to do a simulation. This can be made as exact as desired by increasing the size of the sample, and the accuracy can also be improved by using a highly stratified sample, similar to a Latin hypercube sample (LHS) but with only two variables. That is, divide the unit interval into n subintervals each of length $1/n$, and let p_i be the midpoint of the ith interval. This definition of p_i is a little simpler than randomly sampling for each p_i as LHS would do, and it gives almost the same answers. Let X have cumulative distribution function (c.d.f.) F, and let $x_i = F^{-1}(p_i)$. Construct Y in a similar way. If X and Y are to be independent, randomly shuffle the order of the x_i or y_i values. If X and Y are to be perfectly dependent, leave the order alone. Then construct $z_i = x_i y_i$. The z_i values represent a sample from the distribution of XY. The percentiles of the distribution can be estimated from the sample.

To see the effect of the draft report's approximation, I considered a hypothetical example in which X and Y are independent with the same lognormal distribution, with $M = 1$, $L = 0.1$, and $U = 300$. That is, the lower error factor is 10 and the upper is 300. The exact and approximate methods yielded the following percentiles.

Table 4. Exact and approximate percentiles of product of independent split lognormals.

	Exact (strat. sample)	Approximate
$\ln X + \ln Y$		
5th percentile	−3.26	−3.26
median	1.20	0
95th percentile	8.12	8.07
$X\,Y$		
5th percentile	3.8E−2	3.9E−2
median	3.31	1
95th percentile	3.4E+3	3.2E+3

In this example the percentiles agree quite well, except for the median. The reason for the median's behavior is seen by considering $\ln X + \ln Y$. If both terms are positive the sum is positive. If both terms are negative the sum is negative. But if one term is positive and one is negative, the sum is probably positive, because in this example the right tail is much longer than the left. Therefore, the sum is positive more often than not.

This sort of comparison should be made in at least a few of the cases where the distribution makes a difference to the final answer. This will guide the analysts in deciding whether the approximation is good enough in every case or not. If not, the exact method should be used.

An example when "the distribution makes a difference to the final answer" is when the distribution for a panelist is far to the right of the other distributions. The bold-face entries in Table 2 would be good candidates for the test cases.

The above comparisons do not consider the means. That is because the mean of a sum always equals the sum of the means, and when the terms are statistically independent the mean of a product equals the product of the means.

5. OTHER ISSUES AND DETAILS

5.1 Constructing a Hypothetical New Panelist

This is sometimes called the problem of estimating the 13th panelist, because 12 panelists were actually used. The draft report suggests using the median of each percentile and the median of the means. **This seems reasonable to me. However, aggregation by geometric means of the percentiles already constructs a 13th panelist in one way. Therefore, construction of a 13th**

panelist in a second way is a low-priority issue. The only difference between the two approaches to constructing a new panelist is that the proposed method would work with every step of the panelists' analyses, and the geometric means are applied only to the final conclusions of the individual panelists.

It must be recognized that these approaches construct an *average* panelist. In reality, any new panelist would not be average. The whole process is like trying to predict a new observation of a random process by taking the average of past observations — a realistic uncertainty band around the prediction may need to be very large. **Please recognize that any one panelist, even the hypothetical average panelist, cannot necessarily be regarded as a spokesman for the entire community.** The average panelist can speak for the entire panel only if the panelists do not disagree strongly. If the panelists have strong disagreements, the full picture is revealed only when the outlying distributions are also displayed along with the "average" distribution.

5.2 A Confidence Interval for the Overall Mean

5.2.1 Basic Considerations

Confidence intervals are used to express uncertainty about a parameter, based on an estimate using random data. So we must begin by defining what is random and what the parameter to be estimated is.

- The panelists are assumed to be chosen randomly from a population of all possible panelists. This is the randomness to be quantified by the confidence interval.
- The different aggregation methods estimate different distributions, wide or narrow, which have different means. Therefore, the parameter being estimated depends on the aggregation method.
  - If aggregation by geometric means of percentiles (Class *a*) is used, the parameter to be estimated is the mean of the distribution that results from taking the geometric means of the percentiles of all possible panelists.
  - If aggregation by a mixture (Class *b*) is used, the parameter to be estimated is the overall mean for the population, the average of the means of all the possible panelists.

The authors of the report suggest using the confidence interval for the mean as a measure of the diversity of panelists. Other, more direct measures of diversity should also be considered, such as the 90% interval of the mixture distribution, or the standard deviation of the panelists' log-means or log-medians. Possibly several such measures could be presented.

In my opinion, the primary value of the confidence interval for the mean is to qualify the presentation of the final results. A major result of the project is an estimate of the mean of the aggregated distribution. **This point estimate of the mean should be accompanied by a statement of the uncertainty in this estimate, such as a confidence interval.** The point estimate for the mean and the associated confidence bounds may form the basis for NRC regulations.

The confidence intervals given below assume that the panelists' individual means are known. For this project, however, the final confidence interval must account for the fact that some of the individual means can only be bounded. One way would be to use the union of all the confidence intervals obtained when different possible sets of individual values are used.

We now consider calculation of the confidence interval for the two classes of aggregation.

5.2.2 A Confidence Interval When Aggregating by Mixture

In this situation, the point estimate is simply the mean of the panelists' individual means. This is familiar and easy to think about. Therefore it is presented first, and geometric means of the percentiles are presented second.

Ordinary parametric methods do not work, because the distribution of the means is highly skewed, and a distribution for them is hard to justify. If a lognormal distribution for the individual means is accepted, then a confidence interval procedure is given by Land (1973). Alternatively, a generally applicable nonparametric confidence interval is a bootstrap confidence interval. There are several varieties of bootstrap confidence interval, but one, the BC_a interval (see Efron and Tibshirani 1993) has been implemented in an Excel spreadsheet, which I have sent to the primary authors of the draft report.

5.2.3 A Confidence Interval When Aggregating by Geometric Means of Percentiles

Here, the mean of the aggregated distribution depends on the geometric means of the individual medians and error factors. It also depends on the form of the distribution. Equations (1) and (2) give the mean when the distributions are split lognormal and split log-triangular, respectively. I have not worked out an example, but it appears that the BC_a interval also can be used in this case. To do this, construct a large number of bootstrap samples. For each bootstrap sample, calculate the geometric means of the medians and error factors, and the resulting mean of the aggregated distribution using Equation (1) or (2). Then construct the BC_a interval based on these means.

5.3 Adjustment for Overconfidence

The authors of the draft report tried several methods of adjusting those panelists' intervals that seem too short, because it is well established that people have a tendency to be overconfident. The method that seemed to give adjustments that were not unrealistic was to calculate the geometric mean of the error factors, and then to raise every smaller error factor to be that large. **I like this idea**, for the following reasons:

- It addresses a known problem.
- It gives results that are credible when applied to the actual data.
- It uses the experts themselves to decide how much uncertainty is reasonable for each parameter.
- It may be more tolerable to the panelists if their submittals are adjusted because of what the other panelists said, rather than because of some arbitrary criterion (e.g. EF = 10) that was established by the NRC.

5.4 Other Details

I sent two other files containing comments on the draft, on July 25 and August 4. Because most of those comments are minor, and the few more substantial comments are already in the process of being implemented, they are not repeated here.

6. REFERENCES

Efron, Bradley, and Robert J. Tibshirani, 1993. *An Introduction to the Bootstrap*. New York: Chapman & Hall. Section 14.3.

Land, Charles E., 1973. "Standard Confidence Intervals for Linear Functions of the Normal Mean and Variance," *J. Amer. Statistical Assoc.*, Vol. 68, pp 960-963.

APPENDIX A: Examples Using Different Aggregation Methods

This review considers the following two criteria for a good aggregated distribution.

Criterion 1. Construct a compromise consensus distribution. This should be a distribution that every panelist can live with, but it should not be much wider than the distribution of a typical panelist.

Criterion 2. Construct a distribution that accurately represents all the variation and differences among the panelists. If the panelists disagree greatly, the resulting distribution must reflect this, and must necessarily be considerably wider than the distribution of any single panelist.

Two primary techniques of aggregation method have been suggested, aggregation by taking geometric means of the percentiles and aggregation by taking a mixture of the distributions. In addition, this review suggests showing outliers alongside any aggregated distribution. This appendix considers some hypothetical and actual examples, shows how the different types of aggregation perform in those examples, and compares the results to the two criteria.

A.1 Example 1: A Hypothetical Example with No Consensus.

This example is very simple and idealized, but instructive nevertheless.

A.1.1 The Example

We present a LOCA scenario to a group of experts, with the following results.

- Half the experts say that λ is approximately 1E-5. More precisely, they agree that the distribution is lognormal with median 1E-5, error factor (EF) 10, and mean about 2.7E-5. The basis of their belief is that a LOCA occurrence is almost always produced by a certain mechanism “M”.
- The other half of the experts say that mechanism M is impossible. Therefore, these experts give λ a lognormal distribution with median 1E-11, EF = 10, and mean about 2.7E-11.

Note, the two groups of experts disagree by six orders of magnitude. The two 90% intervals, with median and mean, are shown on a logarithmic scale in the lower part of Figure A.1. The same values are shown on a linear scale in the lower part of Figure A.2. In Figure A.2, the values that appear to be zero there are not exactly zero, just extremely small compared to the scale of the graph.

Ideally, this situation should instigate a separate study of whether mechanism M can occur, but in the meantime, we consider how to aggregate the results that we have.

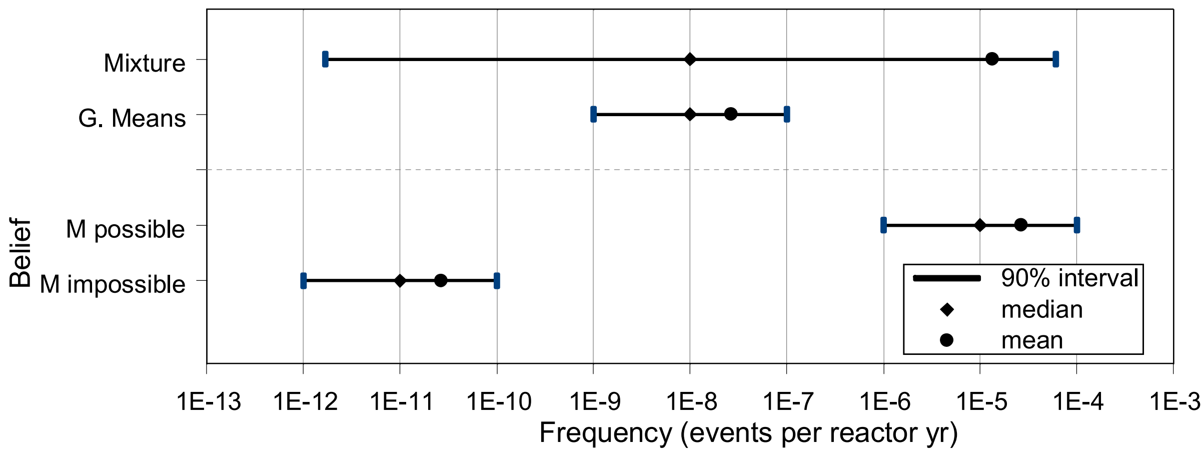

Figure A.1. Log-scale display of 90% intervals with medians and means, for hypothetical example.

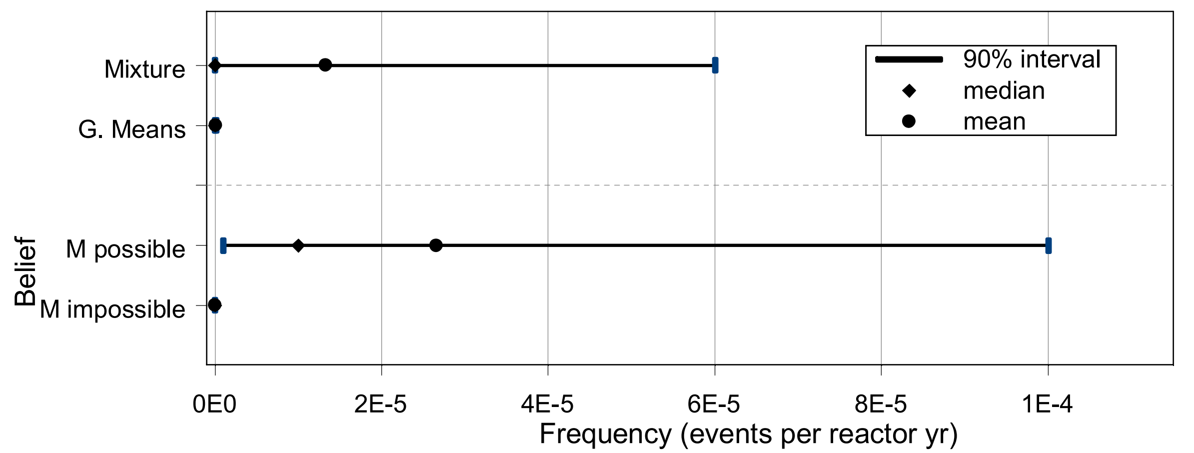

Figure A.2. Linear-scale display of 90% intervals with medians and means, for hypothetical example.

A.1.2 Mixture Aggregation in the Example

If it is conceptually possible to assign a probability to the proposition that mechanism M can occur, it seems unavoidable to set Pr(mechanism M can occur) = 0.5. This is, of course, a subjective probability reflecting epistemic uncertainty. It implies that
$\Pr(\lambda \sim 1\text{E-}5) = \Pr(\lambda \sim 1\text{E-}11) = 0.5$, where the symbol "~" means "is on the order of".

For any interval A, it follows that

$$\Pr(\lambda \text{ in } A) = \Pr(\lambda \text{ in A} \mid \text{M can occur}) \Pr(\text{M can occur}) + \Pr(\lambda \text{ in A} \mid \text{M cannot occur}) \Pr(\text{M cannot occur})$$

For any lognormal distribution with EF = 10, the 90th percentile is about 6 times the median. It follows that

$$\begin{aligned}\Pr(\lambda > 6\text{E-}5) &= \Pr(\lambda > 6\text{E-}5 \mid \text{M can occur}) \Pr(\text{M can occur}) \\ &\quad + \Pr(\lambda > 6\text{E-}5 \mid \text{M cannot occur}) \Pr(\text{M cannot occur}) \\ &= 0.10\times0.5 \ + \ 0\times0.5\end{aligned}$$

$= 0.05$.

Thus, the 95th percentile of λ is the 90th percentile given by the experts who believe in the possibility of mechanism M.

Similarly, the 5th percentile of λ is the 10th percentile given by the experts who disbelieve in the possibility of mechanism M. Also the aggregate mean is the average of the two conditional means, $(2.7E\text{-}5)\times 0.5 + (2.7E\text{-}11)\times 0.5 = 1.3E\text{-}5$.

In summary, the 90% probability interval for λ is (2E-12, 6E-5), the mean is about 1.3E-5, and the median is 1E-8. This interval, with the median and mean, is labeled *Mixture* at the top parts of Figure A.1 and Figure A.2.

A.1.3 Geometric Means of the Percentiles

If we take geometric means of the percentiles, the resulting distribution of λ is lognormal with median 1E-8 and EF = 10. This is a kind of "horizontal" averaging, contrasted with the "vertical" averaging presented above. The mean of this distribution is 2.7E-8, which is also the geometric mean of the two conditional means. The 90% probability interval is (1E-9, 1E-7). This interval, with median and mean, is labeled *G.Means* in Figure A.1 and Figure A.2.

The spread of this distribution is based on the geometric mean of the individual error factors, corresponding to a hypothetical "representative" expert. The distribution does not reflect the diversity of the community of experts. For example the aggregated 90% interval goes from 1E-9 to 1E-7, but half of the experts believe that λ is less than 1E-9 and half of the experts believe that λ is greater than 1E-7.

If half of the experts are correct (though we do not know which half) then the community as a whole gives a biased estimator of the true λ, because half the experts are about right and the other half are all wrong in the same direction (though we do not know which direction). The truth is that λ is either near 1E-11 or near 1E-5. The geometric mean, near 1E-8, is off by three orders of magnitude, but we do not know in which direction. An uncertainty distribution that is to include the truth must include both possibilities, 1E-11 and 1E-5.

A.1.4 Discussion

Because the numbers from the experts differ by orders of magnitudes, it is customary to view them on a log scale as in Figure A.1. However, the applications to PRA use λ, not $\ln\lambda$, giving a reason for using the linear scale of Figure A.2.

In Figure A.1, the mixture 90% interval looks unreasonably long, and the geometric-mean 90% interval looks comfortably short. In Figure A.2, on the other hand, the mixture 90% interval is slightly over half as long as the interval when mechanism M can occur, and the geometric-mean interval is invisibly short compared to the interval advocated by half of the experts.

Now consider the means. In Figure A.1 the mean of the mixture appears to be uncomfortably far to the right, but in Figure A.2, it is midway between the two conditional means, which appears right. Similarly the geometric-mean mean looks perfect in Figure A.1, but is invisibly small in Figure A.2.

The lesson for me is that we should not be swayed too much by pictures. If we are to aggregate, we must decide if we want to satisfy Criterion 1 or Criterion 2. An alternative is not to aggregate, but instead to simply present the disparate results to the decision maker.

This example has no central consensus, differing from the examples of the report. Therefore, we next consider an example that does have a central consensus.

A.2 Example 2: Same Hypothetical Example with Consensus in the Middle

We present a LOCA scenario to a group of experts, with the following results.

- Half the experts say that the distribution for λ is lognormal with median 1E-8, and error factor (EF) 10.
- One fourth of experts say that the median is 1E-11 and one fourth say that the median is 1E-5. They all have lognormal distributions with EF = 10.

The distributions, similar to those in the previous section, are summarized in Figure A.3.

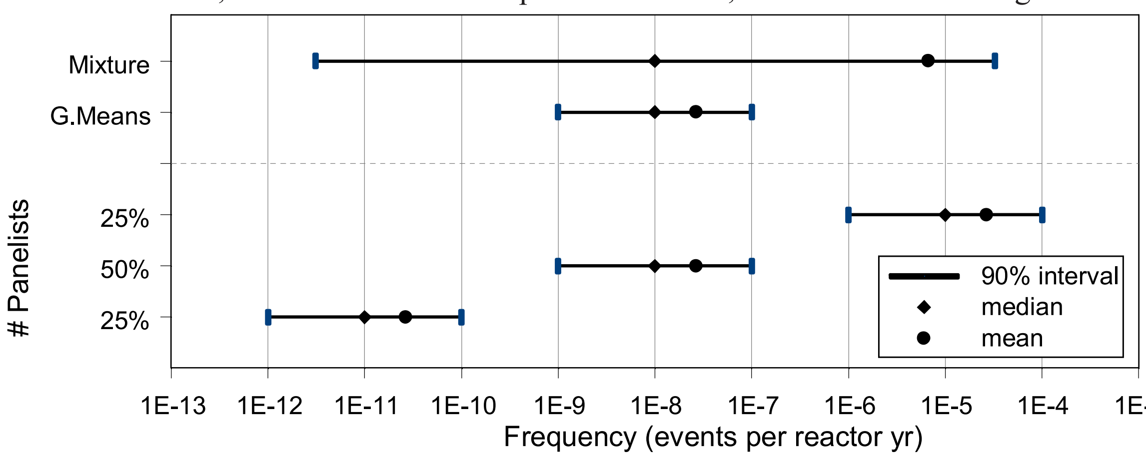

Figure A.3. Log-scale display of 90% intervals with medians and means, for second hypothetical example.

This shows that even though there is a consensus, with 50% of the panelists in the middle, the outliers are completely discounted in the geometric-mean distribution. The outlying medians have effects that cancel each other, and the error factors are the same for all the panelists, so the outlying distributions do not affect the error factor. The mixture distribution still includes much of the outlying distributions, although not quite as much as in Figure A.1.

This example illustrates the fact that averaging percentiles by geometric means can *completely* discount outliers. In this example, 50% of the panelists are outliers.

A.3 Example 3: Category 6 LOCAs in PWRs of Age 25 Years

Consider now a real example, Category 6 LOCAs in PWRs of age 25 years, using the numbers from the Excel spreadsheet containing the results for this example. The intervals and medians are plotted in Figure A.4. For each panelist the possible range of the means is also shown, using the mean of a split lognormal as an upper bound and the mean of a split log-triangular as the lower bound, as discussed in Section 3. The intervals are arranged by descending mean, where the mean is the reported value in the draft report.

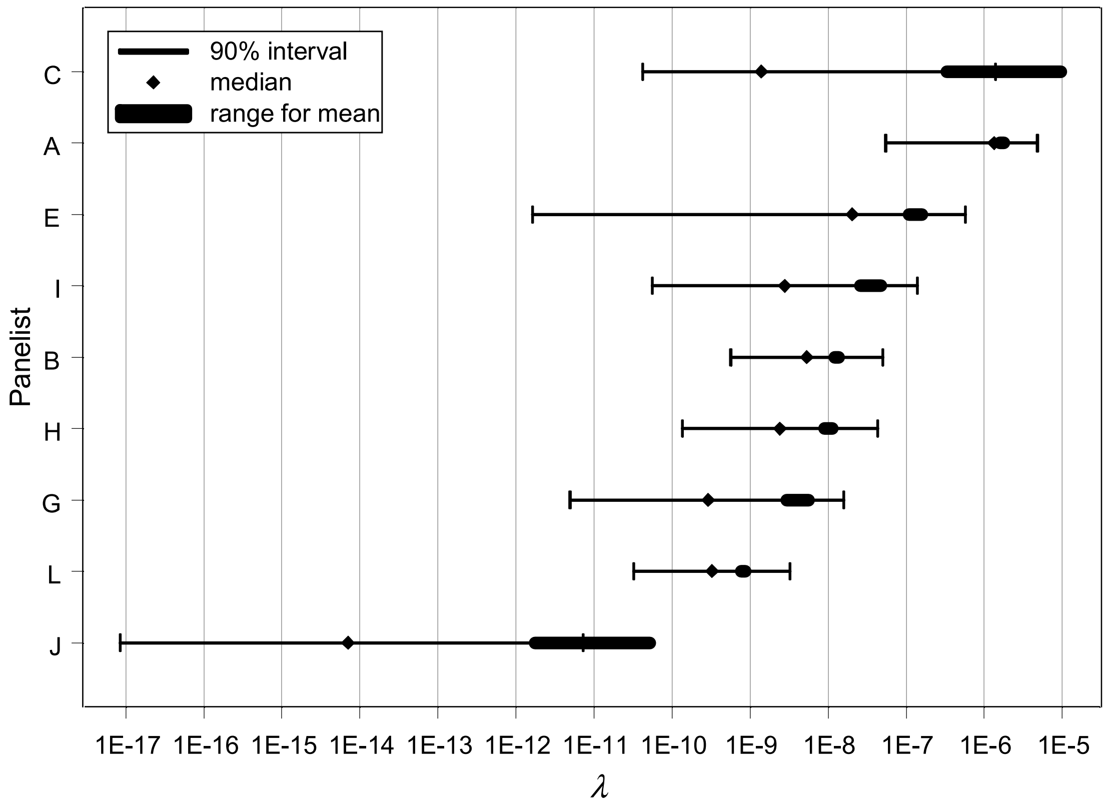

Figure A.4. Results for Category 6 LOCAs at age 25 years, with no aggregation.

It can be seen that the panelists with the longest intervals, that is, the largest error factors, have the most uncertainty in the means, as discussed in Section 3. Clearly, panelist J is an outlier at the low end. At the high end, panelist C may be considered an outlier because of the large mean and extreme skewness, or panelist A may be considered an outlier because of the large median. Identification of the outliers, especially on the right, is somewhat subjective.

Figure A.5 shows the results of aggregating all the panelists by using geometric means of the percentiles, and showing the outliers J, A, and C separately. The bounds on the aggregated mean are calculated exactly, following the procedure given in Section 3.2.3. This figure shows that two of the outlying 90% intervals (J and A) do not overlap the aggregate 90% interval at all. Also, two outlying means are over two orders of magnitude higher than the aggregate mean, one because of a large uncertainty by panelist C and one because of an overall high estimate by panelist A. Displaying both the aggregated distribution and the outliers gives the decision makers a better understanding of the panelists' results. Incidentally, the aggregate 90% interval of the six panelists without J, A, and C is similar to the aggregate 90% interval of all nine panelists.

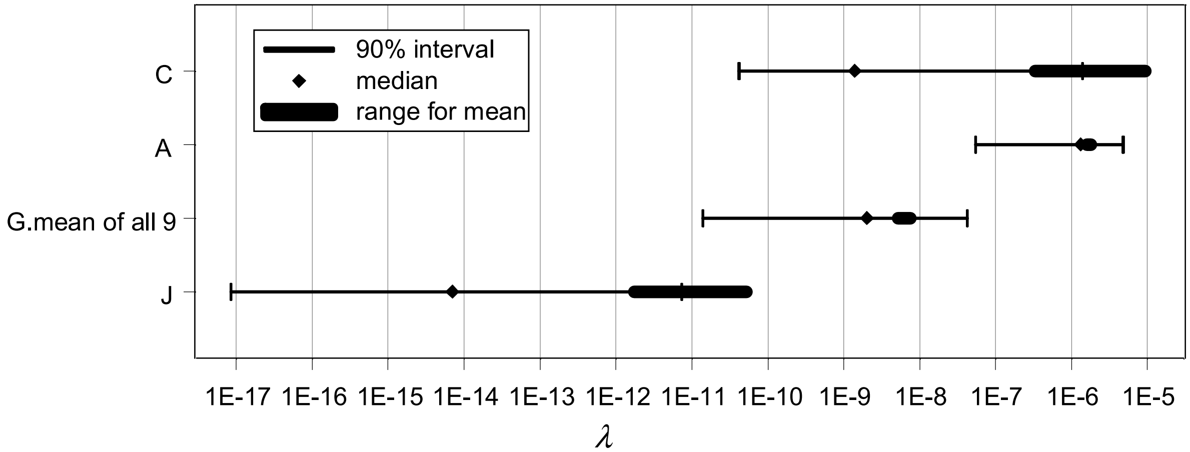

Figure A.5. Results from using geometric means of percentiles to aggregate distributions, and also displaying J, A, and C individually. Data are from Figure A.4.

Now consider aggregation by a mixture instead of by geometric means of percentiles. Figure A.6 shows the results from using mixture aggregation if J is regarded as an outlier, A and C are regarded as an outlying pair that can be aggregated, and the remaining distributions are aggregated. Other possible identifications of outliers could be shown, but are not. Figure A.6 was calculated using Excel by constructing the c.d.f. for each panelist, and then averaging the appropriate c.d.f.s.

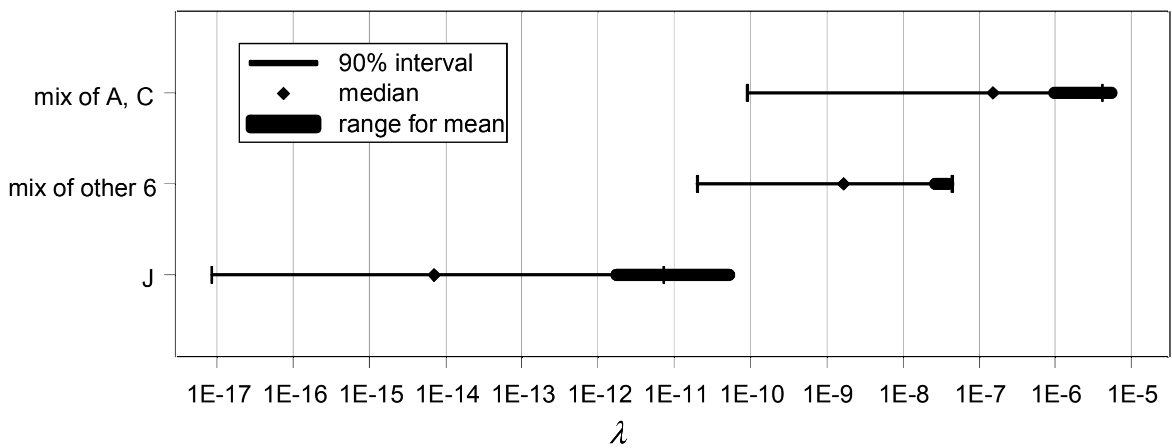

Figure A.6. Results from using mixtures to aggregate distributions into three groups: J, A and C, and all others. Data are from Figure A.4.

Observe that very similar 90% intervals are given by two different aggregations: the aggregation of all 9 distributions by geometric means of the percentiles in Figure A.5, and the aggregation of the 6 *central* distributions by a mixture in Figure A.6. However the two aggregations have means that differ from each other by half an order of magnitude.

In both figures, the outliers on the right reveal that some panelists believe that λ can be large. This fact should not be dodged. Use of only the geometric-means aggregation method or use of a mixture of only the central panelists would hide the large values that some of the panelists believe.

A.4 Summary on Aggregation

If the geometric-mean of percentiles is used for aggregation, outliers may be completely discounted. Therefore, I recommend that the individual outlying cases be presented in addition to the geometric-mean aggregated distribution.

If, instead, a mixture is used for aggregation, the resulting distribution (especially the mean) can be dominated by a large outlier. Therefore, I recommend that only the central distributions be aggregated by a mixture distribution, and that the outliers be shown individually.

A complication arises in the present project, because the panelists' individual means may have substantial uncertainty. This uncertainty is propagated into the overall mean, and a range must be given for the overall mean, not a single point.

APPENDIX B: Derivations of Formulas for the Mean

B.1 Mean of the Truncated Split Lognormal Distribution

This section derives the formula for the mean of the truncated split lognormal distribution. We will repeatedly use the following lemma.

Lemma 1. If X has normal(μ, σ^2) distribution, with density $f(z)$, then

$$\int_{-\infty}^{a} e^x f(x)dx = \exp(\mu + \sigma^2/2) \times \Phi\left(\frac{a-\mu}{\sigma} - \sigma\right),$$

where Φ is the standard normal cumulative distribution function.

Proof.

$$\begin{aligned}
\int_{-\infty}^{a} e^x f(x)dx &= \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{a} e^x e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx \\
&= \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{a} \exp\left[-\frac{1}{2}\left(\frac{-2x\sigma^2 + x^2 - 2x\mu + \mu^2}{\sigma^2}\right)\right]dx \\
&= \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{a} \exp\left[-\frac{1}{2}\left(\frac{[x-(\mu+\sigma^2)]^2 - 2\mu\sigma^2 - \sigma^4}{\sigma^2}\right)\right]dx \\
&= \exp(\mu + \sigma^2/2)\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{a} \exp\left[-\frac{1}{2}\left(\frac{x-(\mu+\sigma^2)]}{\sigma}\right)^2\right]dx \\
&= \exp(\mu + \sigma^2/2)\Phi\left(\frac{a-\mu-\sigma^2}{\sigma}\right),
\end{aligned}$$

which equals the asserted expression.

The mean of a truncated split lognormal is then derived as follows. The mean is

$$\begin{aligned}
&\frac{1}{1-2p}\frac{1}{\sqrt{2\pi}\sigma_L}\int_{b_L}^{\mu} e^x \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma_L}\right)^2\right]dx + \frac{1}{1-2p}\frac{1}{\sqrt{2\pi}\sigma_U}\int_{\mu}^{b_U} e^x \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma_U}\right)^2\right]dx \\
&= \frac{\exp(\mu+\sigma_L^2/2)}{1-2p}\left[\Phi(-\sigma_L) - \Phi\left(\frac{b_L-\mu}{\sigma_L} - \sigma_L\right)\right] + \frac{\exp(\mu+\sigma_U^2/2)}{1-2p}\left[\Phi\left(\frac{b_U-\mu}{\sigma_U} - \sigma_U\right) - \Phi(-\sigma_U)\right] \\
&= \frac{\exp(\mu+\sigma_L^2/2)}{1-2p}\left[\Phi(-\sigma_L) - \Phi\left(-z_{1-p} - \sigma_L\right)\right] + \frac{\exp(\mu+\sigma_U^2/2)}{1-2p}\left[\Phi\left(z_{1-p} - \sigma_U\right) - \Phi(-\sigma_U)\right] \\
&= \frac{\exp(\mu)}{1-2p}\left\{\exp(\sigma_L^2/2)\left[\Phi(-\sigma_L) - \Phi\left(-z_{1-p} - \sigma_L\right)\right] + \exp(\sigma_U^2/2)\left[\Phi(\sigma_U) - \Phi\left(-z_{1-p} + \sigma_U\right)\right]\right\}
\end{aligned}$$

as stated in Section 3.3.1.

B.2 Mean of the Split Log-Triangular Distribution

Consider now the split log-triangular distribution. Figure 1 of the body of this review is repeated here as Figure B.1, for convenience.

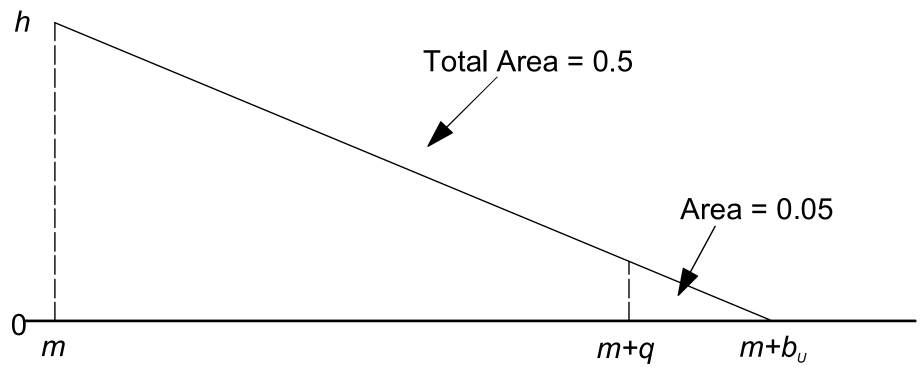

Figure B.1. Right half of a split triangular distribution

From this figure, showing a total area of 0.5, the height h of the density at m must satisfy hb_U = 1. Therefore,
$h = 1/b_U$.
The density has a constant slope which we will call c. The density is of the form $f(x) = 1/b_U + c(x-m)$, with $f(b_U + m) = 0$. Therefore, $c = -1/b_U^2$, as asserted in Section 3.3.2, and we have $f(x) = 1/b_U - (x-m)/b_U^2 = (b_U - x + m)/b_U^2$.

At the 95th percentile $q + m$, the density must satisfy $f(q + m)\times(b_U-q) = 0.1$. Therefore, we have

$$\frac{(b_U - q)^2}{b_U{}^2} = 0.1$$

$$1 - \frac{q}{b_U} = \sqrt{0.1}$$

$$b_U = 1.462q = 1.462 \times \ln(U/M)$$

as stated in Section 3.3.2. The value of b_L is derived similarly. This completely determines the split triangular density f.

The mean of exp$X$ is now found as follows, when X has the above split triangular distribution. We have

$$E[\exp(X)] = \int_{m-b_L}^{m} e^x f(x)dx + \int_{m}^{m+b_U} e^x f(x)dx \;,$$

where f is the split triangular density. The integral on the right equals

$$\frac{1}{b_U^2}\int_{m}^{m+b_U} e^x (b_U - x + m)dx$$

$$= \frac{1}{b_U^2}(b_U + m)e^x\Big|_{m}^{m+b_U} - \frac{1}{b_U^2}\int_{m}^{m+b_U} xe^x dx$$

The integral of xe^x can found by integrating by parts. Then the total expression is

$$\frac{1}{b_U^2}\left[(b_U + m)(e^{m+b_U} - e^m) - \left(xe^x \Big|_m^{m+b_U} - \int_m^{m+b_U} e^x\,dx \right)\right]$$

$$= \frac{1}{b_U^2}\left[(b_U + m)(e^{m+b_U} - e^m) - [(m + b_U)e^{m+b_U} - me^m)] + e^{m+b_U} - e^m\right]$$

$$= \frac{M}{b_U^2}\left[e^{b_U} - b_U - 1\right],$$

with $M = e^m$. The integral on the left is found similarly. The mean of exp$X$ is the sum of the integrals on the left and on the right,

$$M\left\{\frac{1}{b_L^2}\left[e^{-b_L} + b_L - 1\right] + \frac{1}{b_U^2}\left[e^{b_U} - b_U - 1\right]\right\},$$

as asserted in Section 3.3.2.