2

The 95/95 Acceptance Criterion in Risk-Informed Regulation

The 95/95 Delusion

Yuri Orechwa
NRR/DSS/SNPB

June xx, 2019

I. Introduction

A previous White Paper (Ref. 1) discussed a systematic process for identifying and understanding the deterministic computation of the uncertainties associated with the formulation of an analysis framework, that enhances the traditional PRA analyses at L1, by exploiting the results obtained routinely today in licensing basis analyses of reactor transients. The results of such an analysis set the initial conditions for the subsequent L2 and L3 analyses, with which the safety to the public is assessed. This sequence of analyses estimates the amount of radioactive material released from the fuel elements throughout the core to the public under accident conditions. In the continuing spirit "that *understanding* the formulation of Risk-Based Analysis is the 'sine qua non' for *making* Risk-Informed Decisions", we further elaborate the analysis of Ref. 1. In particular, the previous analysis used a statistical hypothesis testing approach with respect to a fixed failure margin at the 95/95 acceptance level of a single fuel element. We intend to expand this analysis to the mean core damage, which is more appropriate for decision making through the application of statistical decision theoretic principles. (Ref. 2) Moreover, we shall address the appropriateness of 95/95 tolerance limit statistic in safety analyses in the context of nuclear reactor safety.

Our methodology is analogous to studies of the behavior of stock market returns. There the motivation and theory are based on that "Evidence on the form of the distributions of returns on securities and portfolios is important for several reasons. For the investor, the **form** of the distribution is a major factor in determining the risk of investment." (Ref.3, p. 17) The analogy to " the distributions of returns on securities" will be in our study of the probability density function of the figures of merit, in particular the peak cladding temperature, that is the independent variable of the loss (core damage) function. The expected core damage will be the measure of risk. The key observation is that **probability alone is not risk!**

Thus, at the heart of the discussion is the expression for the mean core damage (CD) estimated from a best-estimate plus uncertainty (BEPU) result of an analysis of a reactor transient:

$$\langle \mathrm{CD} \rangle = \int_{-\infty}^{+\infty} \mathrm{CD}(T_p)\, f(T_p \mid \mu_p,\ \sigma_p)\, d\,T_p. \qquad \text{Eq. I.1}$$

The objective is the deconstruction of this expression, and, thereby, gain insight to it's place in reactor safety analyses. It is recognized that the core damage is a function of a number of correlated figures of merit. For ease of illustration and discussion, we focus on one, the peak clad temperature T_p. The more realistic multivariate case is numerically messy, but a straightforward extension of the argument presented here. This allows us, without loss of generality, to focus on and illuminate the central issues in our argument of risk.

II. Formulation of the Mean Damage for L2 and L3 PRA Analysis based on L1 LBA Results

The analysis in Ref. 1 is based on a statistical hypothesis testing framework. In that context, the estimate of the level of damage of a single fuel element was quantified as $\alpha * \mathrm{M}$, where M is the damage at a threshold (the limiting peak clad temperature T_p, for example) and α the probability that the peak clad temperature T_p has exceeded the clad temperature at the damage threshold L_γ. We then have the following relationship

$$\mathrm{M} * \mathrm{Prob}\{T_p > L_\gamma \mid \mathrm{ES}, \mathrm{IE}\} = \alpha * \mathrm{M} * (1 - \beta),$$

where the damage is modified by $(1 - \beta)$, the sensitivity of the measurement instrument (i.e. computer code in this application) that the damage margin has been exceeded. This formulation is instructive, but we feel unsatisfactory for risk analysis.

The **fuel element** damage formulation from Ref. 1, however, is suggestive of a phenomenologically descriptive formulation of an estimate of **core** damage for use as an initial condition in L2 through L3 PRA.

To this end, we derive, to first order, the following decision theoretic formulation of the expected core damage <CD> at L1, based on the phenomenological analysis of a **given peak clad temperature probability density function** associated with a **single** fuel element, and a loss function in the form of a peak clad temperature core-damage function. This, in principle, allows an estimate, to first order, of a risk-informed assessment of the relative **mean source** of the health effects at L3 of new reactor designs and due to design changes of operating reactors.

Let us define the risk of core damage as the mean estimate of core damage based on a probability density function f(T_p| μ_p , σ_p), constructed on the basis of a best-estimate plus uncertainty (BEPU) transient analysis of a reactor accident sequence. The additional required information for the **risk** estimate is a "loss" function CD(T_p), that in our case we shall refer to as a "damage" function. This then results in the following expression (Eq. I.1) for the mean core damage:

$$\langle \mathrm{CD} \rangle = \int_{-\infty}^{+\infty} \mathrm{CD}(T_p)\, f(T_p \mid \mu_p,\ \sigma_p)\, d T_p. \qquad \text{Eq. I.1}$$

Let us examine the integrand of this expression by first considering separately the functions in this risk formulation; The two component functions , CD(T_p) the core damage as a function of the peak clad temperature, and f(T_p| μ_p, σ_p), the probability density function of the peak cladding temperature. The interaction between these two functions is critical in determining the implied mean core damage base on BEPU analyses results. At issue is the phenomenological behavior of each function, and the interpretation of it's contribution to the mean core damage for risk-informed decisions.

A. A Phenomenological Loss Function CD(T_p) that reflects Core Damage.

We develop our phenomenological core damage function CD(T_p) by considering, without loss of generality, a minimalist core model consisting of cylindrical geometry and a homogeneous composition. Let us assume a hemispherical flux distribution in the core. This results in a hemispherical power distribution, and for uniform flow as a hemispherical clad temperature distribution of T_p as shown in Fig. 1. We associate each planar ring with a ring of fuel pins with a the same T_p.

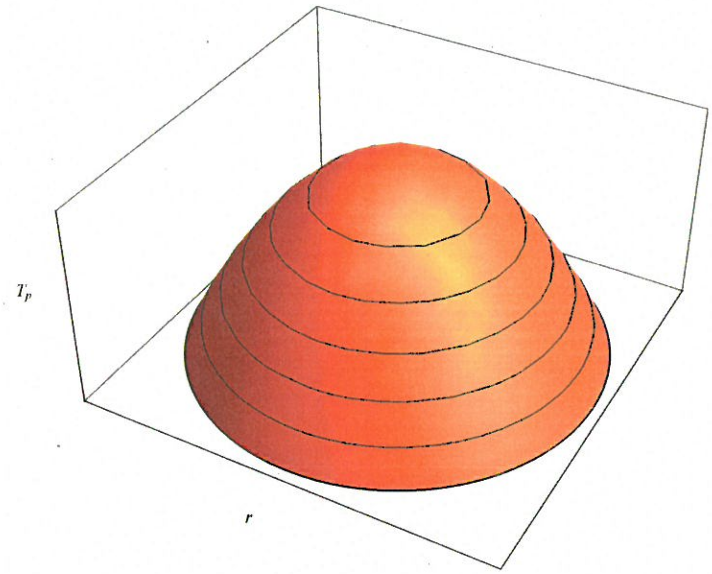

Fig. 1 Peak Cladding Temperature T_p as a Function of Core Radius

The relation between T_p and the core radius r is shown in Fig. 2, in an axial planar perspective, for no pin failures, partial core failure and full core failure for a fuel pin at radius r. The phenomenological effect this illustrates is that as the fuel performance code computes $T_p > L_\gamma$ for the lead pin at the center, as the **calculated** T_p of the center pin increases more and more adjacent pins fail.

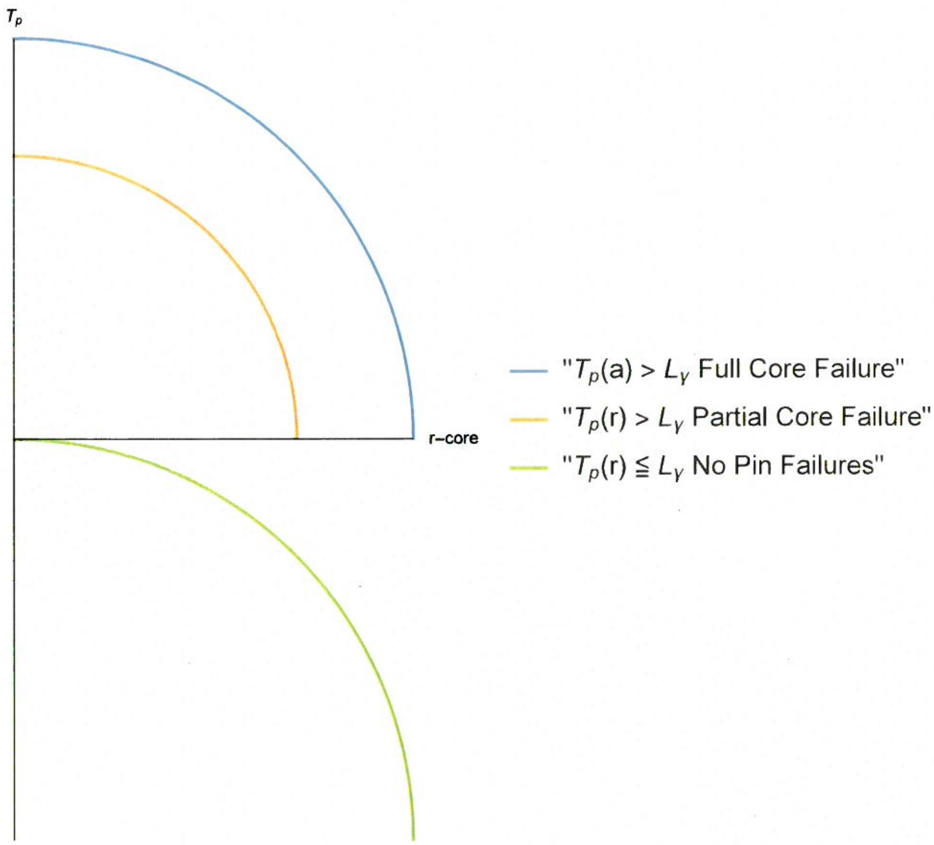

Fig. 2 Peak Cladding Temperature T_p as a Function of Core Radius

Let the core consist of N identical fuel elements, that are each associated with a unit cell area (UC). Thus, the planar active core region has an area $A_c = \text{N*UC}$. Furthermore, we consider that each fuel pin that fails (by $T_p > L_\gamma$) releases a fixed amount of 'damage' (PD) per fuel pin, say in the form radio active isotopes, upon exceeding the acceptance limit L_γ.

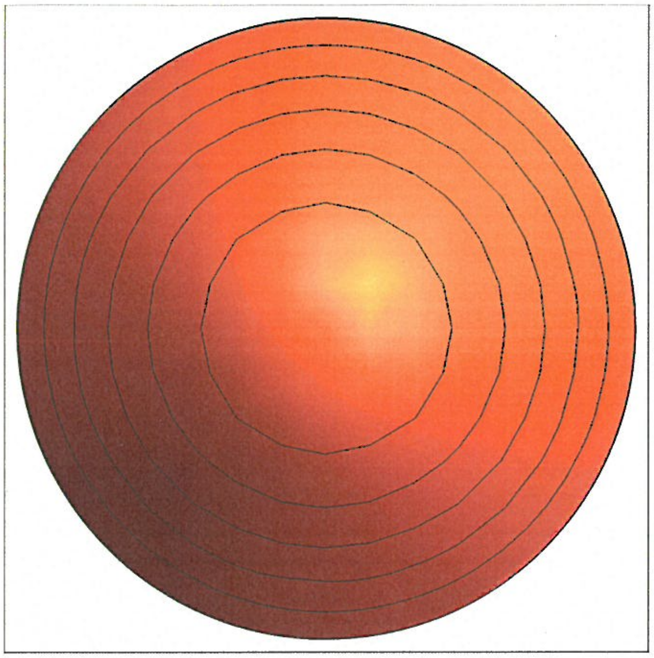

Fig. 3 Variation of the Area of Failed Fuel Pins for Peak Cladding Temperature T_p as a Function of Core Radius

Therefore, as the peak clad temperature of the peak pin (core center pin) increases (Fig. 3), the concentric area containing all the failed pins increases. We, therefore, can express the number of failed fuel pins for an area associated encompassed by T_p(r) (Fig. 2), as

$$\text{no. failed pins} = \pi r^2/\text{UC}, \quad \text{for } T_p(\text{r}).$$

We, therefore, have

$$\text{CD}(T_p(\text{r})) = \text{PD}^*(\pi r^2/\text{UC}). \qquad \text{Eq. II.2}$$

Given our assumption of a hemispheric core power distribution, we can scale the radius to T_p.

$$\text{r} = [(T_p(\text{a}) - L_\gamma) / \text{a}]^*[T_p(\text{r}) - L_\gamma],$$

and let $C_a \equiv [(T_p(\text{a}) - L_\gamma) / \text{a}]$. We have the necessary relationship

$$\text{r} = C_a{}^*[T_p(\text{r}) - L_\gamma].$$

Substituting this relation into Eq. (II.2) we obtain the necessary the core damage function

$$\mathrm{CD}(T_p(\mathrm{r})) \sim [T_p(r) - L_\gamma]^2. \qquad \text{Eq. II.3}$$

That is, CD is **quadratic** in $T_p(r)$.

It is, of course, recognized that in a realistic case the analytic relationship will not be as 'clean' as in Eq. II.3. However, it will be a **power law** dependent on the peak to average power distribution of the core under consideration.

III. Analysis of Core Damage Risk Based on a Perfect Information Probability Density Function of the Peak Cladding Temperature T_p and the Acceptance Limit L_γ.

As a generic representative for our argument, we shall consider the normal probability density function for ease and familiarity. Any continuous pdf will have similar phenomenological implications. For example, such as the lognormal pdf that is more appropriate in many cases.

As perfect information, we define two normal pdfs for which we know the true mean and standard deviation, say $N_B(10, 1)$ and $N_G(1, 6.45)$. (We shall not associate dimensions with these, since only their relations are of interest.) In practice, we have estimates of these at best. So any shortcomings in the methodology under perfect information will likely be more pronounced in practice. The parameters were chosen specifically so that both have the same 0.95 quantile. Thus, for both pdf's $\gamma = 0.95$, and, therefore, $L_\gamma = 11.65$. This is graphically presented in Fig. 4.

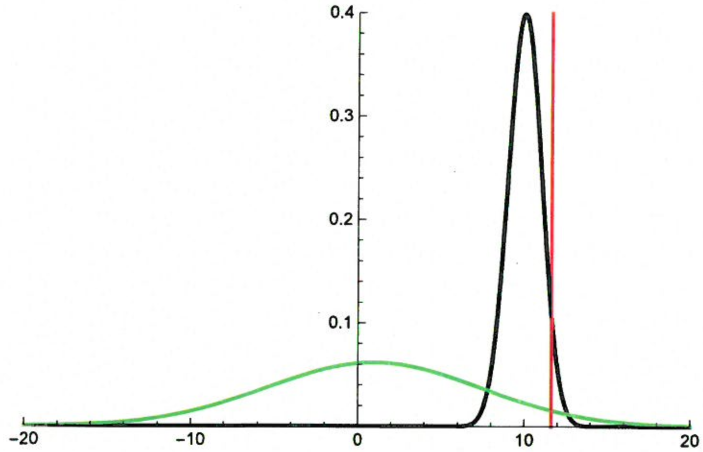

Fig. 4. N_B(10, 1) - Black, N_G(1, 6.45) - Green and $L_{0.95}$ - Red.

These two normal distributions are not unique with respect to having the same quantile value 11.65. We can, in principle, construct an infinite number of normal distributions that have the same quantile by adjusting the mean and variance appropriately.

In the previous section above, we developed a phenomenological relationship between the peak clad temperature of the peak pin and core damage. In BEPU analysis of reactor systems, this relationship is not considered. The acceptance criteria are stated in terms of the probabilities of two nonintersecting events: $\{T_p \leq L_\gamma\}$ and $\{T_p > L_\gamma\}$ as Prob$\{T_p \leq L_\gamma\} = 0.95$ and Prob$\{T_p > L_\gamma\} = 0.05$. Both distributions in Fig. 4, for example, meet this BEPU criterion. If we let the two distributions be associated with two different reactors, they are both acceptable. However, we might ask is one "riskier" than the other? Looking at the distributions in Fig. 4, an "interocular" test might conclude that the reactor associated with the black distribution is riskier. Since the frequency of high T_p close to L_γ is much greater than that for the reactor associated with the distribution in green. This conclusion could be based, for example, on the intuitive analogy with standing at the edge of a cliff as opposed to further back.

This BEPU decision framework is not only not unique, but also only assigns a same probability to infinite possible designs and not a unique expected damage to each. Let us consider Fig. 4 further, in the context of Eq. I.1 of the expected core damage associated with any BEPU result. Our interest is the interaction between CD(T_p(r)) and $f(T_p \mid \mu_p, \sigma_p)$. For this we can set CD(T_p(r)) = $[T_p(r) - L_\gamma]^2$ and capture the phenomenological behavior of interest as in Fig. 5..

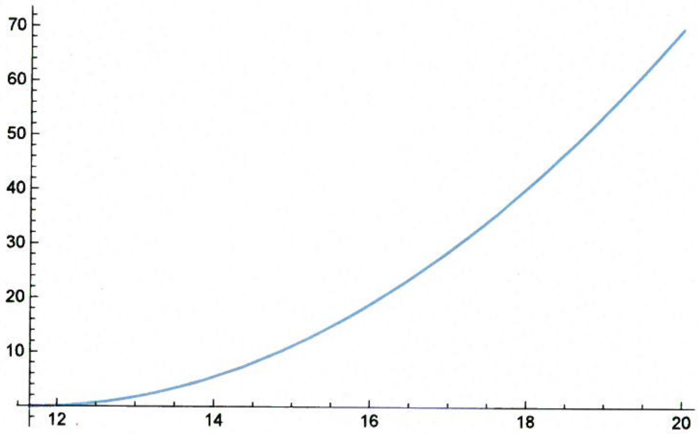

Fig. 5. Core Damage Function CD(T_p(r)) ≡ $[T_p(r) - L_\gamma]^2$ for event $\{T_p > L_\gamma\}$.

The expected damage for both systems is dependent only on the event $\{T_p > L_\gamma\}$ for a single fuel pin, while the two pdfs are dependent on the specific designs of the two reactor systems and are shown in Fig. 6.

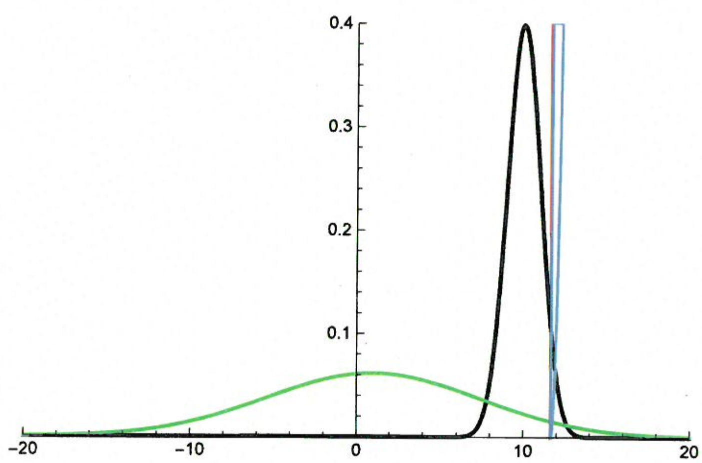

Fig. 6. Superposition of expected core damage components CD(T_p) and $f(T_p \mid \mu_p, \sigma_p)$.

Since the expected core damage for both systems is dependent only on the event $\{T_p > L_\gamma\}$, we need only examine the asymptotic relationship with respect to the abscissa, as shown in Fig. 7. Mathematically the asymptotic behaviors of the functions are: CD $\sim$ x^2 and the pdfs $\sim$ e^{-x^2}. Therefore, as x $\to \infty$, the pdfs $\to 0$ (black and green curves) faster than $x^2 \to \infty$ (blue curve).

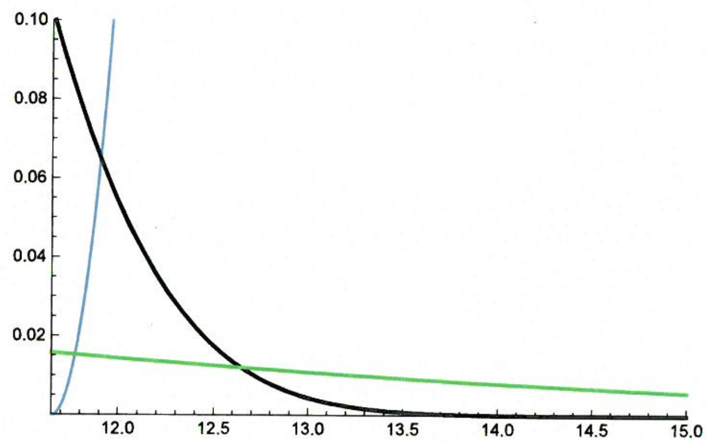

Fig. 7. Expected core damage components CD(T_p) and $f(T_p \mid \mu_p, \sigma_p)$ for event $\{T_p > L_\gamma\}$.

The integrand of Eq. I.1 for the two pdfs is shown in Fig. 8, and illustrates the importance of the contribution to $< \text{CD} >$ of the "tails" of the pdfs for event $\{T_p > L_\gamma\}$. The thicker tail of the green pdf contributes more to CD than the more rapidly decaying black pdf.

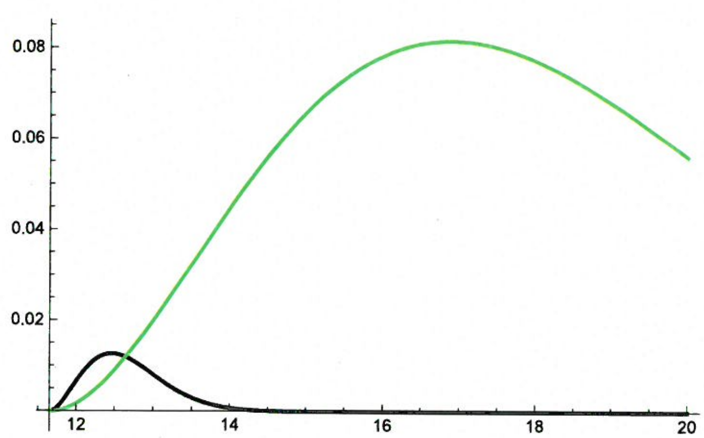

Fig. 8. Integrand of <CD> for event $\{T_p > L_\gamma\}$.

The mean core damage for the two systems given by Eq.I.1 is then:

$$< \text{CD} >_{\text{Black}} = 0.015$$

$$< \text{CD} >_{\text{Green}} = 0.467.$$

Thus, based on the traditional BEPU acceptance criterion Prob$\{T_p \leq L_\gamma\} = 0.95$, both systems are acceptable under perfect information. Our interocular test suggested that the system represented by the black pdf is 'riskier'. Phenomenological analysis, that takes core damage into account, shows that both are incorrect, and that the system represented by the green curve will have a much greater risk of core damage. Thus, in order to make a risk-informed decision based on BEPU results, it is necessary to introduce a physically relevant loss function, that gives the relationship between the peak clad temperature and the damage incurred as a result of exceeding an experimentally determined limiting temperature due to uncertainty.

IV. Analysis Based on an Imperfect (i.e. Sample) Information Probability Density Function of the Peak Cladding Temperature T_p and the Acceptance Limit L_γ.

In practice, we only have a random sample of size N from the true pdf. Thereby, BEPU analyses give only imperfect information for making a risk assessment with regard to core damage. (It is important to note that we assume that the uncertainty is due only to the random effects of the model's parameters. i.e. it is phenomenologically correct) In order to see how "imperfect" the results are we study two basic estimation methods: parametric and non-parametric. In the case of the former, we shall use a representative of the class of response surface methods. For the latter, in light of the current in-vogue approach at NRC, a tolerance interval method; in particular using the so called 95/95 methodology of Wilks' Theorem (Ref. 4) to quantify the level of safety. We assess these approaches by applying them to the previous analysis above of perfect information via the two specified pdfs shown in Fig. 4. The core damage function is the same and not an issue; we focus on the implication of the effect of the limited information content in the BEPU sample for making risk informed decisions with regard to reactor safety.

A. Base Case Sample Data

Our objective is to study the loss of information when we have only a random sample drawn from the “true” pdf of the figure of merit. to the end, let us first consider a base case sample from each of the two distribution functions $N_B(10, 1)$ and $N_G(1, 6.45)$ shown in Fig.4. The key objective is to demonstrate the weakness of the so called 95/95 methodology in light of the estimation of the mean core damage (Eq. I.1) for assessing nuclear reactor safety.

We draw random samples of size N = 59, so as to be consistent with the 95/95 methodology.

For $N_B(10,$
$1)$:

{10.2107, 9.38212, 10.1792, 10.1587, 12.1323, 10.6931, 10.3642, 8.52441, 10.4009, 10.1661,
10.5556, 10.4479, 9.92655, 9.21098, 10.4392, 9.7536, 10.9363, 9.60991, 10.9228, 9.15423,
10.1566, 10.7537, 9.60641, 11.1431, 9.28669, 11.4952, 8.81195, 10.4316, 10.9711, 9.844,
10.113, 10.0296, 9.15593, 9.92208, 10.7389, 8.27512, 9.93952, 10.8391, 8.56407, 9.66609,
11.2341, 7.58957, 10.4807, 7.23384, 11.1549, 8.25832, 11.2367, 9.15294, 10.001, 10.9748,
9.46498, 10.7563, 10.5194, 11.8907, 10.7448, 9.39109, 11.1513, 12.2136, 8.97911}

and for $N_G(1, 6.45)$
:

{12.5798, 9.66408, 0.740544, -0.9746, -3.46813, -4.83713, -3.14985, 2.52295, 4.69621, -0.944088,
-8.02446, 1.19877, 5.76318, 2.14461, -1.08628, -2.93579, -12.5708, -1.97193, 1.42524, 2.69249,
3.8633, -3.74272, -15.3513, 1.3456, 3.828, 1.05397, 5.98373, 7.16034, 3.80897, -1.21239,
12.0271, 0.0894731, 6.94144, 4.11539, -2.09345, 3.52815, -7.7629, 0.16986, 5.61613, 7.95565,
4.47253, 1.66266, 3.63721, 13.1487, -7.31508, 0.973678, -4.08423, 17.2706, 13.3061, 10.9996,
7.18058, -0.454851, 6.32628, -1.06634, -9.64521, -5.5967, -3.36068, 2.94812, -0.664466}

Table I. Basic Distribution Moments for Perfect vs. Sample Information in the Two Base Cases

	Perf. Inf.BLK	Perf. Inf.GRE	Sampel Inf.BLK	Sampel Inf.GRE
mean	10	1	10.09	1.53
st.dev.	1	6.45	1.03	6.37
skew	0	0	-0.45	-0.049
kurt.	3	3	3.26	3.31
T(59)	NA	NA	12.21	17.27

B. Non-Parametric Estimation (á la Wilks)

Recall from Ref.1 that our basic task is to distinguish between the probabilities of the two events: { $T_p \leq L_\gamma$ } No damage, and { $T_p > L_\gamma$ } damage. Thus, the following identity holds:

$$\text{Prob}\{\ T_p \leq L_\gamma\ \} + \text{Prob}\{\ T_p > L_\gamma\ \} = 1. \qquad \text{Eq. III.4}$$

In addition, the cumulative probability density function in the case of perfect information is given by the expression:

$$\int_{-\infty}^{L_\gamma} f(T \mid \mu, \sigma)\, dT = \gamma\ . \qquad \text{Eq. III.5}$$

Where γ is the fixed acceptable portion of the pdf f (T | μ, σ). Thus, with perfect information in the form of the analytic form of the pdf f(T| μ, σ), we can compute the quantile q_γ exactly with the relationship given in Eq.III.5. However, when we do not know the analytic form of f(T| μ, σ), we must estimate L_γ with a random sample of Ts via a BEPU analysis with a computer code.

It is instructive to derive the Wilks' relationship for estimating the quantile q_γ associated with γ from imperfect (i.e. sample) information. To this end, we consider a set of N runs of the BEPU code for a transient analysis to obtain N **different** values of T_p (Since T_p is a random variable in a BEPU methodology.)

Let us order the N calculated values as an ordered set $\{T(i)\}_{i=1}^{N}$, where T(1) < T(2) < T(3) < ...... < T(N). From identity Eq.III.4, it follows that

$$\text{Prob}\{\ T_p > L_\gamma\ \} = 1 - \text{Prob}\{\ T_p \leq L_\gamma\ \}. \qquad \text{Eq.III.6}$$

Since T(i)'s are assumed to be independent continuous random variables with common, but unknown, pdf,s we can express,

$$\text{Prob}\{\text{T(N)} \leqslant L_\gamma\} = \quad \text{Prob}\{\text{T(1)} \leqslant L_\gamma\}\times\text{Prob}\{\text{T(2)} \leqslant L_\gamma\}\times \ ...\ \times\text{Prob}\{\text{T(N)} \leqslant L_\gamma\}.$$

Therefore we get with the application of Eq.III.5 that

$$\text{Prob}\{\text{T(N)} \leqslant L_\gamma\} = \gamma^N.$$

Substituting this expression into Eq.III.6, and defining $\beta \equiv \text{Prob}\{\text{T(N)} > L_\gamma\}$, we have Wilk's expression for one random variate as

$$\beta = 1 - \gamma^N. \qquad \text{Eq.III.7}$$

There are two basic interpretations of this expression (Ref.):

A. β is the probability that the event $\{\text{T(N)} > L_\gamma\}$ is greater than that of the γ-quantile of the true but unknown pdf f(T| μ, σ) that generated the sample $\{T(i)\}_{i=1}^N$.

B. β is the probability that the interval ($-\infty$, T(N)] covers a larger portion than γ of the of the true but unknown pdf f(T| μ, σ) .

Recall in the case of perfect information in Sec.II.B, Fig. 4, we saw that an infinite number of unique pdfs are consistent with the event $\{T_p > L_\gamma\}$ for the same quantile L_γ. Similarly, for an outcome of a maximum value T(N) in our imperfect (sample) information analysis of N code runs, there is an infinite set of γ/β combinations that satisfy Wilks' formula for fixed N as shown by the Blue curve in Fig.9. Are the other points γ/β on the Blue curve equally "risky" to 95/95? Wilks' formula is silent with regard to the implications on the risk states associated with them. This is clear, since the basic premise is that the pdf is unknown, and, by extension, the shape of the tail for event $\{T_p > L_\gamma\}$, which is critical for the estimation of the risk.

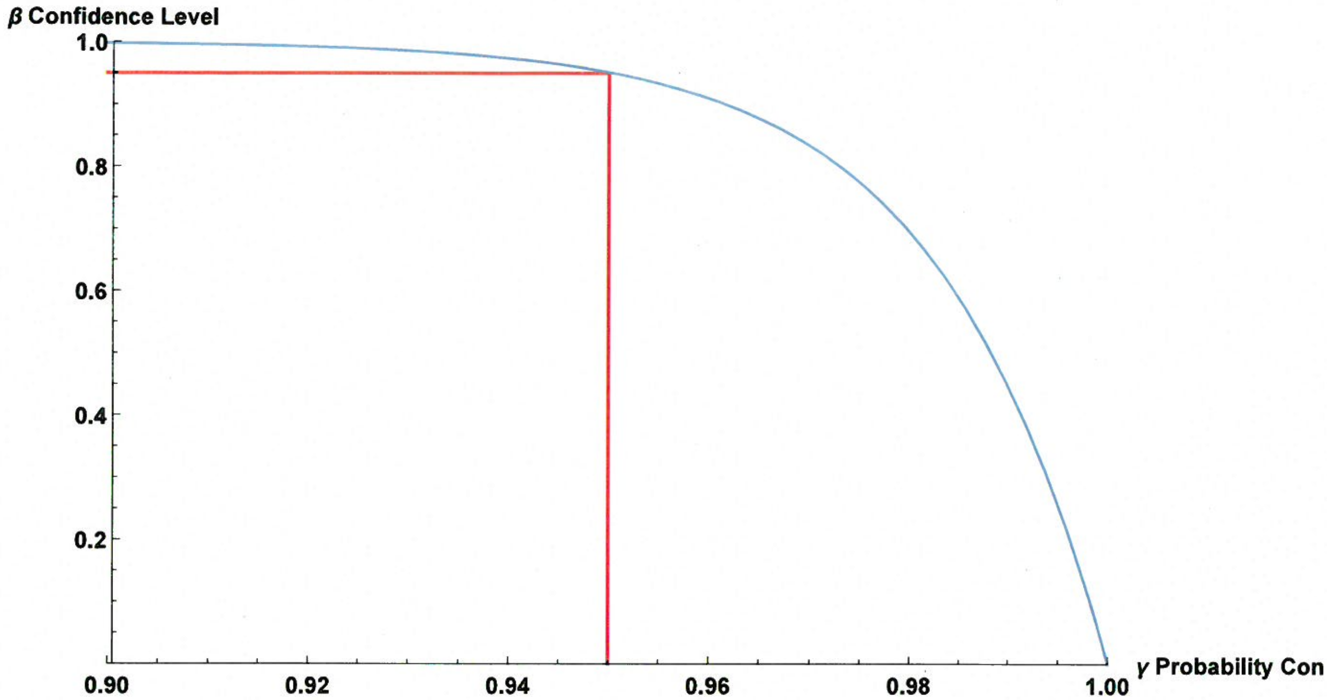

Fig. 9. Wilks' formula Eq.III.7 with N=59 (Blue) and 0.95/0.95 (Red).

Similarity Relations between Perfect Information and Wilks' γ/β Methodology Acceptance Criteria Formulations for Peak Clad Temperature.

Perfect Information		**Sample Information**
$\text{Prob}\{\, T_p > L_\gamma \,\} = \beta$	$\sim$	$\text{Prob}\{\, T(N) > L_\gamma \,\} = \beta$
β	$=$	β
L_γ − quantile (0.95)	$\sim$	

T (N) – quantile nonparametric estimate

$f(T \mid \mu, \sigma)$ – known analytic form ~ T (N) – random variable

$(-\infty, L_{\gamma})$ – specified interval ~
$(-\infty, T(N))$ – random interval

μ / σ – defines quantile ~ γ / β – defines N

References

1) Y. Orechwa, "Some Formal Considerations in the Application of Best-Estimate Plus Uncertainty Results in Licensing Basis Analysis to Risk "Informed" Decision Making for Reactor Safety," Internal NRC/NRR/DSS White Paper, June 14, 2018.

2) M. H. De Groot, *Optimal Statistical Decisions*, McGraw-Hill, New York, 1970.

3) Eugene F. Fama, *Foundations of Finance*, Basic Books, Inc. New York, 1976. (p. 17).

4) Y. Orechwa, "Best-Estimate Analysis and Decision Making Under Uncertainty," Proc. Int. Mtg. Updates inBest-Estimate Methods in Nuclear Installations Safety analysis (BE-2004), November 14-18, Washington D.C., American Nuclear Society (2004).

5) I. Guttman, *Statistical Tolerance Regions: Classical and Bayesian*, Griffin's Statistical Monographs & Courses, Hafner, Darien Conn., 1970.

6) U.S. Nuclear Regulatory Commission, "PRA Procedures Guide," NUREG/CR-2300, Vol. 1 (January 1983).

7) L. Pal and M. Makai, "Remarks on the statistical aspects of the safety analysis." arXiv:physics/0511140v1, November 2005.

8) A. Guba, Makai and L. Pal, "Statistical Aspects of Best-Estimate Methods - I," Reliab. Eng. Syst. Saf., 80, 217 (2003).

9) M. Makai, L. Pal, "Best estimate method and safety analysis II." Reliab. Eng. Syst. Saf., 91, 222 (2006).