Comparison among Methods of Modeling Epistemic Uncertainty in Reliability Estimation

Nam Ho Kim[1,*], Min Young Yoo[2] and Joo-Ho Choi[2]

[1]*Department of Mechanical and Aerospace Engineering, University of Florida, Gainesville, Florida 32611, USA*
[2]*Department of Aerospace and Mechanical Engineering, Korea Aerospace University, Goyang-city, 412-791, Korea*

Abstract

Epistemic uncertainty, the lack of knowledge, is often more important than aleatory uncertainty, variability, in estimating reliability of a system. While the probability theory is widely used for modeling aleatory uncertainty, there is no dominant approach to model epistemic uncertainty. Different approaches have been developed to handle epistemic uncertainties using various theories, such as probability theory, fuzzy sets, evidence theory and possibility theory. However, since these methods are developed from different statistics theories, it is difficult to interpret the result from one method to the other. The goal of this paper is to compare different methods in handling epistemic uncertainty in the view point of calculating the probability of failure. In particular, four different methods are compared; the probability method, the combined distribution method, interval analysis method, and the evidence theory. Characteristics of individual methods are compared in the view point of reliability analysis.

Keywords: Epistemic uncertainty, Reliability, Interval method, evidence theory, probability method

---

1. Introduction

Epistemic uncertainty, the lack of knowledge, is often more important than aleatory uncertainty, variability, in estimating reliability of a system. While the probability theory is widely used for modeling aleatory uncertainty, there is no dominant approach to model epistemic uncertainty. Different approaches have been developed to handle epistemic uncertainties using various theories, such as probability theory [1], fuzzy sets or possibility theory [2], and evidence theory [3,4]. However, since these methods are developed from different statistics theories, it is difficult to interpret the result from one method to the other. The goal of this paper is to compare different methods in handling epistemic uncertainty in the view point of calculating the probability of failure.

It should be stressed out that we focus on describing characteristics of different approaches instead of suggesting a good approach. Here, we examine four approaches in modeling epistemic uncertainty: the probability method, the combined distribution method, the interval analysis method and the evidence theory. The first two methods, also known as second-order probability approach, model epistemic uncertainty with a distribution. The first method separates epistemic uncertainty from aleatory uncertainty and uses double-loop Monte Carlo simulation (MCS) to calculate the distribution of reliability. The second method, on the other hand, integrates epistemic uncertainty with aleatory uncertainty to predict combined distribution, whose results are equivalent to the mean of reliability from the first method. The interval analysis method is the simplest way to propagate epistemic uncertainty into interval of reliability. This method is a special case of the probability based method with separated epistemic uncertainty modeling. The epistemic uncertainty is defined with a bound-ed uniform distribution and the corresponding reliability bounds. Evidence theory models epistemic uncertainty with sets of intervals. A basic probability is assigned for each interval and intervals may overlap each other or have a gap.

We demonstrate these four approaches with a simple example of failure strength estimation. There is aleatory uncertainty in the failure strength due to inherent material variability. The source of epistemic uncertainty is from estimating the distribution of failure strength using a finite number of samples; i.e., statistical uncertainty. We will present how to interpret the epistemic uncertainty in the probability of failure for different approaches.

2. Aleatory and Epistemic Uncertainty

In the case of time-independent structural systems, it is often convenient to use a limit-state function in order to describe a reliability of a system. In this paper, the following form of limit-state function is used to determine the safety and failure of the system:

$$G = S - R \tag{1}$$

where, in the case of strength of material, S is the failure strength of a material, and R is the stress applied to the structure. The system is considered safe when $G > 0$ and failed with $G \leq 0$. When the material strength or applied stress are uncertain, the safety of the system is evaluated in terms of reliability or, equivalently, the probability of failure defined as

$$P_F = P[G \leq 0] \tag{2}$$

where $P[\ ]$ stands for the probability of the event in the bracket.

In order to simplify the discussion, it is assumed that there is no uncertainty in the applied stress; that is, the failure strength is the only source of uncertainty. Due to inherent variability of

---

*Corresponding author. Tel.: 1-352-575-0665, Fax.: 1-352-392-7303
 E-mail address: nkim@ufl.edu

the material, it is assume that the failure strength is normally distributed; that is,

$$S_{\text{true}} \sim \mathrm{N}(\mu_{\text{true}}, \sigma_{\text{true}}) \tag{3}$$

where μ_{true} and σ_{true} are, respectively, the mean and standard deviation of the true strength of the material, which are also called the distribution parameters. The uncertainty S_{true} is inherent variability and called *aleatory uncertainty*.

Unfortunately, the true distribution parameters of failure strength are unknown, and it can only be estimated through tests. When n number of specimens are used to estimate the failure strength, the estimated distribution of the failure strength becomes

$$S_{\text{test}} \sim \mathrm{N}(\mu_{\text{test}}, \sigma_{\text{test}}) \tag{4}$$

where μ_{test} and σ_{test} are, respectively, the sample mean and standard deviation of test. Due to sampling error, the distribution parameters from test are different from the true ones. Of course, when infinite test samples are used, S_{test} will converge to S_{true}. However, since a finite number of samples are used, the estimated distribution parameters have *epistemic uncertainty* (i.e., sampling uncertainty or statistical uncertainty). From the classical probability theory, the uncertainty in the estimated mean and standard deviation can be expressed by

$$\mathrm{M}_{\text{est}} \sim N\left(\mu_{\text{test}}, \frac{\sigma_{\text{test}}}{\sqrt{n}}\right) \tag{5}$$

$$\Sigma_{\text{est}} \sim \frac{\sigma_{\text{test}}}{\sqrt{n-1}} \mathcal{X}(n-1) \tag{6}$$

where $\mathcal{X}(n-1)$ is the chi-distribution of order $n-1$. Figure 1 illustrates the concept of aleatory and epistemic uncertainty of the failure strength distribution.
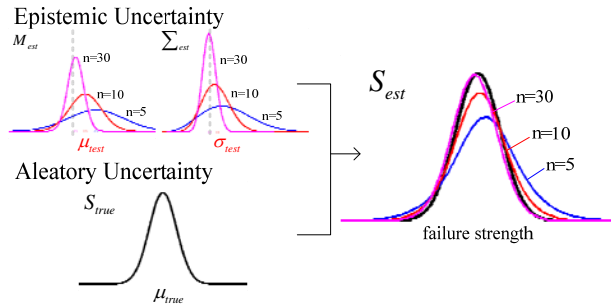

Fig. 1. Illustration of the concept of aleatory and epistemic uncertainty of the failure strength distribution.

The source of epistemic uncertainty in Eqs. (5) and (6) are from sampling error. In general, however, different sources of epistemic uncertainty exist, such as modeling error, numerical error, etc. Although the statistical uncertainty in Eqs. (5) and (6) is given in the form of a distribution, epistemic uncertainty is not random by nature; that is, the true mean and standard deviation will be a single value, but their values are unknown. In this regard, the PDF of the distributions in Eqs. (5) and (6) should be interpreted as the shape of knowledge about the parameter. For example, Eq. (5) can be interpreted that the likelihood of μ_{test} being μ_{true} is higher than any other values.

When both aleatory and epistemic uncertainties exist, designer must determine a conservative failure strength in order to compensate for both uncertainties. For example, when there exists aleatory uncertainty only, designer can determine the 90 percentile of the distribution as a conservative failure strength. When epistemic uncertainty also exists, however, the 90 percentile cannot be determined as a single value, rather it becomes a distribution by itself. A designer must consider the effect of epistemic uncertainty when he or she chooses a conservative failure strength.

The conservative estimate of failure strength for both aleatory and epistemic uncertainty has already been implemented aircraft design. When coupon tests are used to estimate a conservative failure strength of an aluminum material, it is required to be estimated using A- or B-basis approach [5]. In the case of B-basis, for example, the conservative failure strength is estimated by 90 percentile with 95% confidence. The 90 percentile is for aleatory uncertainty, while 95% confidence is for epistemic uncertainty.

In the case when the epistemic uncertainty is represented by a probability distribution as in Eqs. (5) and (6), the probability theory can be used to calculate the effect of both aleatory and epistemic uncertainty, but it may not be straightforward if different methods of representing epistemic uncertainty are employed. In the following sections, four different methods of representing epistemic uncertainty are explained in the view point of conservative estimation of the failure strength: (1) the probability method, (2) the combined distribution method, (3) the interval analysis method and (4) the Dempster-Shafer evidence theory.

In order to make the discussions in the following section simple and consistent, we further simplify the epistemic uncertainty. It is assumed that epistemic uncertainty only exists in the estimated mean, M_{est}, not in the estimated standard deviation. In addition, it is further assumed that the epistemic uncertainty in the mean is uniformly distributed. This simplification is necessary because some methods do not use probability distribution to represent epistemic uncertainty. Table 1 shows the distribution parameters for both aleatory and epistemic uncertainty

Table 1. Distribution parameters for aleatory and epistemic uncertainty.

Failure strength	$S_{\text{est}} \sim N(\mathrm{M}_{\text{est}}, \Sigma_{\text{est}})$
Estimated mean	$\mathrm{M}_{\text{est}} \sim U(200, 250)$
True mean	$\mu_{\text{true}} = 225$
Estimated standard deviation	$\sigma_{\text{est}} = \sigma_{\text{true}} = 20$
Applied stress	$R = 190$

3. Methods of Representing Epistemic Uncertainty

3.1 Probability method

The probability method is the most common approach in representing both epistemic and aleatory uncertainty. In this approach, the epistemic uncertainty is given in the form of a probability distribution. In the case of sampling error (i.e., statistical uncertainty), the epistemic uncertainties in the mean and standard deviation can be represented by normal and chi distributions, respectively, as shown in Eqs. (5) and (6). In the case of modeling error, since the epistemic uncertainty is related to the lack of knowledge, a uniform distribution is frequently used. However, it is important to note that even if the epistemic uncertainty is represented in the form of probability distribution, its interpretation should be different from that of aleatory uncertainty. That is, there is no randomness in epistemic uncertainty, but the probability distribution is used to shape the form of knowledge regarding the uncertain variable. Therefore this method is preferable when the information on the epistemic uncertainty is detail enough, such as the case of sampling error, so that the probability distribution of the epistemic uncertainty can be formed.

The problem formulation in Section 2 is based on the situation where the form of probability distribution for an uncertain variable is known, but the distribution parameters governing the distribution are uncertain. In such a case, the estimated failure strength essentially becomes a distribution of distributions. The estimated distribution of the failure strength can be obtained using a double-loop Monte Carlo simulation (MCS), as shown in Fig. 2. In the figure, the outer loop generates n samples from the estimated mean distribution, $M_{est} \sim U(200, 250)$, from which n sets of normal distributions, $N(\mu_{est}^i, \sigma_{est})$, can be defined. In the inner-loop, m samples of failure strengths are generated from each $N(\mu_{est}^i, \sigma_{est})$, which represents aleatory uncertainty. Since the failure strength is normally distributed, it is also possible that the inner-loop can be analytically calculated without generating samples.
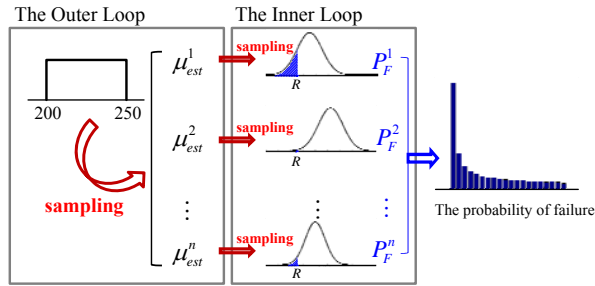

Fig. 2. Double-loop Monte Carlo simulation for estimated distribution of failure strength.

For each given sample from epistemic uncertainty, the aleatory uncertainty is used to build a probability distribution, $N(\mu_{est}^i, \sigma_{est})$, from which the probability of failure, P_F^i, can

be calculated. By collecting all samples, a distribution of probability of failure can be obtained, which represents the epistemic uncertainty. A conservative estimate of the probability of failure, P_F^{90}, can be obtained by taking the 90 percentile of the distribution. Therefore, the effect of aleatory uncertainty is considered by calculating P_F^i, while that of epistemic uncertainty is considered by calculating P_F^{90}.

For the given example, the PDF of the probability of failure and its 90 percentile conservative estimate is shown in Figure 3. It is noted that since the PDF of the probability of failure is highly skewed, the conservative estimate, P_F^{90}, is far from its mean value, P_F^m.
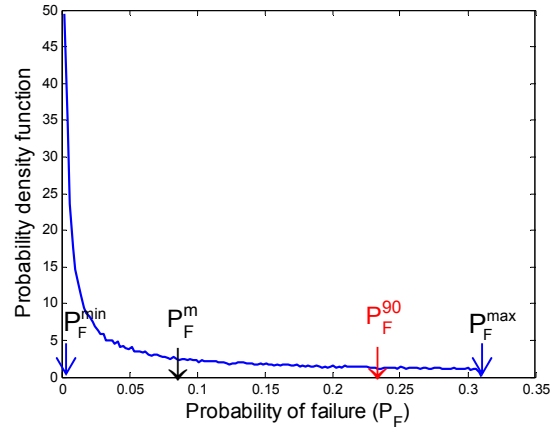

Fig. 3. PDF of the probability of failure and its 90 percentile conservative estimate.

In the probability-based method, the epistemic and aleatory uncertainties are treated separately, which can have both advantages and disadvantages. Disadvantages are the computational cost related to the double-loop MCS and the increase in dimensionality. That is, the number of uncertain input variables increases. Advantages are the separate treatment of epistemic and aleatory uncertainty such that it is clear to identify the sources of uncertainty.

3.2 Combined distribution method

In the combined distribution method, the epistemic and aleatory uncertainties are combined together and represented as a single distribution. Because of that, the advantages and disadvantages of the probability method are interchanged in this method. That is, the estimated true method is computationally inexpensive with a less number of uncertain input variables, while it cannot separate epistemic uncertainty from aleatory uncertainty.

If MCS-based sampling method is used to calculated the estimated true distribution, all $n \times m$ samples in Figure 2 are used to obtain the estimated distribution of failure strength, which includes both aleatory and epistemic uncertainty. However, the real advantage of the estimated true distribution is when an analytical method is used to calculate the combined distribution, which eliminates sampling error. In order to

3

model the above MCS process analytically, the estimated failure strength is firstly defined as a conditional distribution as

$$S_{\text{est}} \big| (\mu_{\text{est}}, \sigma_{\text{est}}) \sim N(\mu_{\text{est}}, \sigma_{\text{est}}) \quad (7)$$

where the left-hand side is a conditional random variable given μ_{est} and σ_{est}. The PDF of the conditional distribution can be calculated by integrating the distribution parameters. Since only the uncertainty in the mean is considered, the PDF of the failure strength can be expressed as

$$f_S(s_{\text{est}}) = \int_{-\infty}^{\infty} \phi(s_{\text{est}} \mid M_{\text{est}} = \mu_{\text{est}}) \, f_{M_{\text{est}}}(\mu_{\text{est}}) \, d\mu_{\text{est}} \quad (8)$$

where $\phi(s_{\text{est}} \mid M_{\text{est}} = \mu_{\text{est}})$ is the normal PDF of failure strength with given μ_{est}, and $f_{M_{\text{est}}}(\mu_{\text{est}})$ is the PDF of the mean parameter, which is uniformly distributed as given in Table 1. For mathematical details when both the mean and standard deviation have epistemic uncertainty, readers are referred to Park et al. [6]. Once the PDF of the estimated failure strength is obtained, the probability of failure can be calculated using Eq. (2) as

$$P_F^{\text{est}} = \int_{-\infty}^{R} f_S(s_{\text{est}}) \, ds_{\text{est}} \quad (9)$$

As the estimated distribution includes both aleatory and epistemic uncertainty, the probability of failure is a single value. Park et al. [6] showed that Gauss quadrature with 50 segments are accurate enough when the level of probability of failure is in the order of 10^{-7}, while MCS has more than 200% COV with a million samples.

It is interesting to note that the probability of failure in Eq. (9) is indeed the mean of the probability of failure distribution, P_F^m, from the probability method. It is relatively easy to show this fact by using MCS process, as

$$\frac{1}{n} \sum_{j=1}^{n} \left[\frac{1}{m} \sum_{i=1}^{m} I\left(G_{ij} < 0\right) \right] = \frac{1}{nm} \sum_{i=1}^{n \times m} I\left(G_i < 0\right) \quad (10)$$

In the above equation, the term on the left-hand side corresponds to $P_F^m = 0.0789$ from the probability method, while the term on the right-hand side is $P_F^{\text{est}} = 0.0786$ in Eq. (9).

In terms of computational cost, the estimated distribution method is more efficient than the probability method because the former can obtain the probability of failure through numerical integration. However, since the former can only estimate the expected value of epistemic uncertainty, it is difficult to take a conservative estimate. Especially when the distribution is severely skewed, such as the distribution of the probability of failure in Fig. 3, it can be dangerous to use a mean value. Therefore, it is not easy to find the confidence interval due to epistemic uncertainty.

3.3 Interval analysis method

The interval analysis method is considered as the simplest way to represent epistemic uncertainty. Compared to the probability method, the interval analysis method assumes that nothing is known about the input uncertainty except for the lower- and upper-bound [7]. Since the least amount of information is used in representing the input uncertainty, it is natural that the output uncertainty will also have the least amount of information; that is, the lower- and upper-bound of the limit-state function.

Although the input uncertainty is represented in the simplest method, it is not straightforward to find the minimum and maximum values of the limit-state function, unless it is a monotonic function of input variables. Optimization algorithms can be employed to find the minimum and maximum values of limit-state function. However, many optimization algorithms are limited to local optima, and finding global optima is often computationally challenging.

Although the interval analysis method does not assume any particular distribution type between the intervals, it is convenient to consider the input interval is uniformly distributed, and uniform samples are generated to find the minimum and maximum of limit-state samples. Of course, in this case, the accuracy depends on the number of samples. In this regard, the interval analysis method becomes similar to the probability method when the epistemic uncertainty is uniformly distributed between intervals. Therefore, the minimum and maximum values of the probability of failure from the double-loop MCS will be identical to that of the interval analysis method, which is in this case $[P_F^{\min}, P_F^{\max}] = [0.00135, 0.309]$, as shown in Fig. 3. Therefore, this method can be used only when very limited information on input epistemic uncertainty is available. Even so, it never used information that is generated during uniform sampling searching for the minimum and maximum values of probability of failure.

Since the probabilistic concept is not used in the interpretation of the interval analysis results, it is not possible to find confidence intervals. Instead, the maximum and minimum values of probability of failure can be used for the purpose of conservative design. As shown in Table 2, however, the maximum P_F from the interval analysis method is too conservative compared to P_F^{90} from the probability method. This is partly because the tail portion of the distribution is significantly skewed.

Table 2. Comparison of uncertainty in the probability of failure between different methods.

Method	Estimated probability of failure
Probability method	$P_F^m = 0.0789, \quad P_F^{90} = 0.23$
Combined distribution method	$P_F^{\text{est}} = 0.0786$
Interval analysis method	$[P_F^{\min}, P_F^{\max}] = [0.00135, 0.309]$
Evidence theory	$[Bel^{90}, Pl^{90}] = [0.154, 0.309]$

3.4 Evidence theory

Evidence theory, also called Dempster-Shafer theory, measures uncertainty with belief and plausibility, which are the lower- and upper-bounds of probability with evidence instead of probability distribution. It is close to the interval

4

analysis method in a sense that uncertainty is represented in the form of lower- and upper-bounds, while it is close to the probability method in a sense that each interval has an assigned probability. It is useful for epistemic uncertainty in situations where there is little information on which to evaluate a probability or when the information is nonspecific, ambiguous, or conflicting.

In Dempster-Shafer evidence theory, the epistemic uncertain input variables are modeled using a belief structure, which is a set of intervals with basic probability assignment (BPA) to each interval, which indicates the level of likelihood that the uncertain input falls within the interval. If the entire range of epistemic uncertainty is represented by a single interval with BPA = 1, it becomes identical to the interval analysis method.

In order to show epistemic uncertainty in evidence theory, a belief structure on the input failure strength is constructed as shown in Fig. 4(a), where each interval has a constant BPA of 0.2.
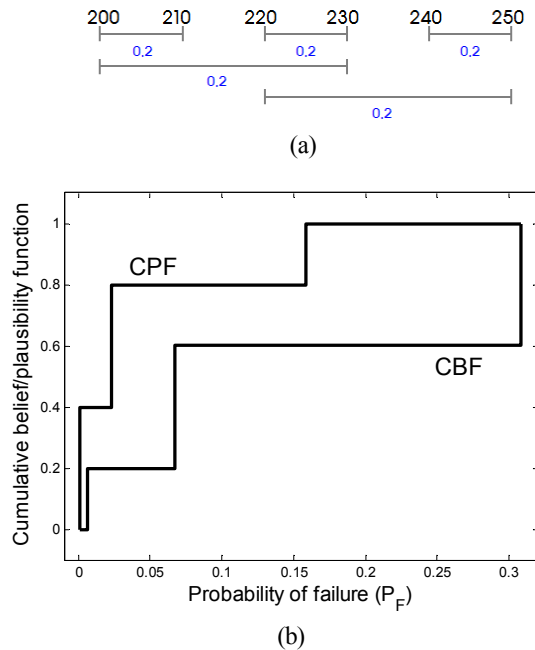

(a)

(b)

Fig. 4. CBF and CPF of the probability of failure. (a) Basic probability assignment for five overlapping intervals, and (b) cumulative plausibility and belief functions of the probability of failure.

The main process of evidence theory is to find the cumulative belief function (CBF) and the cumulative plausibility function (CPF) of the probability of failure. CBF is the cumulative belief that the uncertain probability of failure P_F is less than a given value y, $Bel(P_F \leq y)$. Similarly, CPF is the cumulative plausibility that the uncertain probability of failure P_F is less than a given value y, $Pl(P_F \leq y)$.

CBF and CPF can be found by performing a similar process with the interval analysis method: finding minimum and maximum values and accumulating BPA of each interval. Since the interval analysis method is used for each interval, the evi-

dence theory can be expensive, especially when multiple variables are involved.

Figure 4(b) shows the CBF and CPF for the probability of failure. Similar to the interval analysis method, it is not easy to define a confidence interval for the evidence theory. However, different from the interval analysis method, it is possible to estimate a narrower interval corresponding to 90 percentile from CBF and CPF graphs. For example, 90 percentile of CBF is $Bel^{90} = 0.15$, while that of CPF is $Pl^{90} = 0.309$. Therefore, the range of the probability is much narrower than that of the interval analysis method. In addition, this range also covers the 90 percentile from the probability method, $P_F^{90} = 0.23$.

The interpretation of CBF and CPF is a combination of the probability method and the interval analysis method. For example, the minimum likelihood that the probability of failure is less than 0.1 is 0.6, while the maximum likelihood is 0.8. Therefore, connecting with the probability method, it is possible to consider that CBF is the lower-bound of CDF, and CPF is the upper-bound. This observation becomes clear if the number of intervals increases. Figure 5 compares the CBF and CPF for different number of intervals. It is clear that as the number of intervals increases, the gap between CBF and CPF decreases and both the CBF and CPF converge to the distribution of probability of failure obtained from the probability method (black curve).
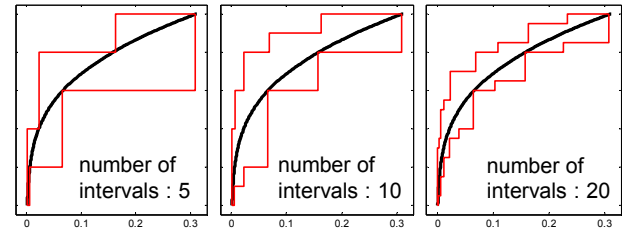

Fig. 5. Comparison between probability method and evidence theory with different numbers of intervals.

4. Conclusions

In this paper, four different methods of representing epistemic uncertainty are presented in estimating the uncertainty in the probability of failure. It is found that the probability method represents the uncertainty most accurately as it uses the full distribution of input uncertainty. On the other hand, the interval analysis method only provides a lower- and upper-bounds of the probability of distribution because of the lack of information in input epistemic uncertainty. Since computational costs of these two methods are almost identical, the choice depends on the availability of information of the input uncertainty.

The combined distribution method has a computational advantage but only provides the mean value of the distribution, which can be dangerous especially when the distribution is highly skewed. However, this method is the most computationally inexpensive. The evidence theory is located between the probability method and the interval analysis method, but the computa-

tional cost can be the highest among four methods. As more information is available for input uncertainty, this method converges to the probability method. Figure 6 illustrates the distributions of the probability of failure from the four methods.

Based on this study, it is concluded that it is important to choose a method based on the level of information available in input epistemic uncertainty.
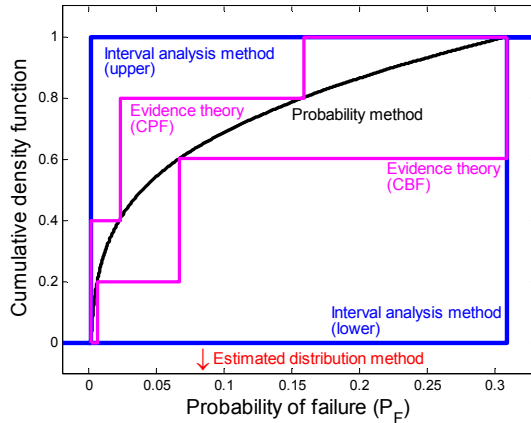

Fig. 6. CDF of the probability of failure.

Acknowledgment

Nomenclature

G : Limit-state function
S : Capacity (failure strength)

R : Response (applied stress)
μ_{true} : Mean of true failure strength
μ_{test} : Mean failure strength from test
μ_{est} : Estimate of true mean failure strength
σ_{true} : Standard deviation of true failure strength
σ_{test} : Standard deviation of failure strength from test
σ_{est} : Estimated standard deviation of true failure strength

References

[1] Helton, J.C., and R. J. Breeding, Calculation of reactor accident safety goals, Reliability Engineering and System Safety Vol. 39, pp. 129-158, 1993.

[2] L. A. Zadeh, Fuzzy sets as a basis for a theory of possibility, Fuzzy Sets and Systems 100: 9–23, 1999.

[3] G. Shafer, A mathematical theory of evidence, Princeton University Press, 1976.

[4] A. P. Dempster, Upper and lower probabilities induced by a multivalued mapping, The Annals of Mathematical Statistics 38 (2): 325–339, 1967.

[5] Guidelines for property testing of composites. Composite materials handbook MIL-HDBK-17. U.S. Department of Defense: Washington, DC; 2002.

[6] C. Y. Park, N. H. Kim, and R. T. Haftka, How coupon and element tests reduce conservativeness in element failure prediction, Reliability Engineering and System Safety 123 (2014) 123-136

[7] J. C. Helton, J. D. Johnson, W. L. Oberkampf, and C. J. Sallaberry. Representation of analysis results involving aleatory and epistemic Uncertainty, Sandia National Laboratories Technical Report SAND 2008-4379.

[8] L. P. Swiler, T. L. Paez, and R. L. Mayes, Epistemic uncertainty quantification tutorial, Proceedings of the 27th International Modal Analysis Conference. 2009.