AIAA 2006-7038

CONSERVATIVE ESTIMATION OF PROBABILITY OF FAILURE

Victor Picheny, Nam-Ho Kim, Raphael T. Haftka
Dept. of Mechanical & Aerospace Engineering

Jorg Peters
Dept. of Computer Science and Engineering
University of Florida
Gainesville, Florida, 32611, USA

11th AIAA/ISSMO Multidisciplinary Analysis & Optimization Conference
6 – 8 Sep. 2006/Portsmouth, VA

CONSERVATIVE ESTIMATION OF PROBABILITY OF FAILURE

Victor Picheny[*], *Nam-Ho Kim*[†], *Raphael T. Haftka*[‡], Jorg Peters[§]
University of Florida, Gainesville FL-32611

Abstract:

The paper provides a review of how to estimate a probability of failure from a small sample of data, and shows that the usual estimators of the parameters of the cumulative distribution function are biased, and can lead to unconservative estimations. Then, it explores different ways to make this estimation conservative: one is based on adding constraints when distributions are fitted; the second is based on the use of bootstrap methods. We explore the relationship between the chance that the estimate is conservative and the accuracy of the estimate. In particular, we study the case when we want to achieve a 95% chance to have conservative estimators. Finally, these methods are applied to the problem of a composite panel under thermal loading.

I. Introduction

In mechanical analysis, uncertainties in input parameters—such as material properties or geometry—prevent the engineer from taking analysis results at face value. The quantification of the influence of these uncertainties on reliability is crucial in mechanical design. Engineering systems need to be designed so that the risk of failure will not exceed an acceptable value.

In the literature (e.g., Ref. [1]), many methods are proposed to estimate the reliability of structural systems and to apply them to design under uncertainties. However, there exist uncertainties in the uncertainty analysis itself. They are caused by various sources, such as variation in the samples and errors in fitting distributions. It has been shown that error in probability distributions due to insufficient information can have large effect on probability calculation (e.g., Ref. [2] [3]). In many engineering problems there is an incentive to obtain probability of failure estimates that will not be less than the real probability; that is *conservative* estimation. This approach can provide a method of uncertainty estimation with a confidence level.

Given the probability distribution of inputs, we can quantify by Monte-Carlo Simulations (MCS) the distribution of the output. The standard application of MCS to estimate the probability of failure is to generate a sample of outputs and calculate the number of values that exceed the limit. However, a small sample cannot evaluate directly the tails of the distributions; we can evaluate low probabilities only by fitting a distribution to the sample. Given a sample of system response $g_1, g_2, \ldots, g_n$, and a limit state $g = g_{\text{limit}}$ defining failure,

[*] PhD Student, department of Mechanical and Aerospace Engineering, Student Member AIAA
[†] Assistant Professor, department of Mechanical and Aerospace Engineering, Member AIAA
[‡] Distinguished Professor, department of Mechanical and Aerospace Engineering , Fellow AIAA
[§] Professor, department of Computer Science and Engineering

estimating the probability of failure is equivalent to estimating the Cumulative Distribution Function (CDF) F_g of g at $g = g_{\text{limit}}$

$$P_f = P(g \geq g_{\text{limit}}) = 1 - F_g(g_{\text{limit}}) \tag{1}$$

In this study, we consider several alternatives of estimating the probability of failure P_f such that the estimation $\hat{P}_f$ is likely to be no lower than the true P_f. To do so, we modify the methods of estimating the CDF from a sample in order to bias the estimator of P_f. We also explore the possibility of using the bootstrap method for probability of failure estimations, and define conservative estimators based on bootstrapping.

In the next section, we discuss how we use sampling to estimate the probability of failure. Section III shows how to use constraints to obtain conservative estimators. Section IV describes the bootstrap method and how to use it to define conservative estimator. The accuracy of such estimators are analyzed using a simple numerical example in Section V, and the conservative estimators are applied to an engineering problem in Section VI, followed by concluding remarks in Section VII.

II. Various estimators of CDF

Analytical estimators

Assuming a certain distribution, an analytical model of CDF can be fitted to the sample by adjusting its parameters. In the case of normal distribution, for example, we want to estimate the mean and the standard deviation, μ and σ. The classical estimators of the mean and the standard deviation from a sample of size n are:

$$m = \frac{1}{n}\sum_i x_i \ , \qquad s = \sqrt{\frac{1}{n-1}\sum_i (x_i - m)^2} \tag{2}$$

where estimate s of the standard deviation is normalized by $(n - 1)$ to make s^2 the best unbiased estimate of the variance. However, it is shown that s is a biased estimate of σ [6]. The expectation of s can be approximated by:

$$E(s) \approx \sigma\sqrt{\left(1 - \frac{\sigma^2}{2(n-1)}\right)} \tag{3}$$

As a consequence, the estimate of the standard deviation is likely to be underestimated; so the tail of the estimated CDF will be biased in the unconservative side.

Estimators based on fitting the empirical CDF

If we simulate a sample of n points and arrange the points in increasing order $(x_1 \leq x_2 \leq \ldots \leq x_n)$, the empirical CDF is defined as:

$$F_X(x) = \begin{cases} 0 & \text{for } x \leq x_1 \\ \dfrac{k}{n} & \text{for } x_k \leq x \leq x_{k+1} \\ 1 & \text{for } x \geq x_n \end{cases} \tag{4}$$

It is then possible to estimate the mean and standard deviation of the CDF that best approximates the empirical CDF. Two different ways of approximation are studied:

- Minimizing the RMS (root mean square) error between estimated CDF and empirical CDF
- Minimizing the Kolmogorov-Smirnov distance.

To minimize the *RMS error* between empirical and estimated CDF, the errors are calculated at sample points. In order to have an unbiased estimator, the values of the empirical CDF are chosen as (see Figure 1):

$$F_X(x_k) = \frac{k - \frac{1}{2}}{n}, \qquad 1 \leq k \leq n \tag{5}$$

The parameters (μ, σ) will then be calculated by solving the optimization problem:

$$\underset{\mu,\sigma}{\text{Minimize}} \left[\frac{1}{n} \sum_{k=1}^{n} \left(F_{\mu,\sigma}(x_k) - \frac{k - \frac{1}{2}}{n} \right)^2 \right] \tag{6}$$

where $F_{\mu,\sigma}$ is the value of the CDF of a normal distribution with parameters (μ, σ):

$$F_{\mu,\sigma}(x_k) = \frac{1}{\sigma.\sqrt{2\pi}} . \int_{-\infty}^{x_k} \exp\left[\frac{-(t-\mu)^2}{2.\sigma^2} \right] .dt \tag{7}$$

The *Kolmogorov-Smirnov* (K-S) *distance* is the classical way to test if a sample is representative of a distribution. The K-S distance is equal to the maximum distance between two CDFs (see Figure 2). The maximum distance occurs at one data point. The optimization problem for the K-S distance becomes:

$$\underset{\mu,\sigma}{\text{Minimize}} \left[\max_{1 \leq k \leq n} \left(\left| F_{\mu,\sigma}(x_k) - \frac{k}{n} \right|, \left| F_{\mu,\sigma}(x_k) - \frac{k-1}{n} \right| \right) \right] \tag{8}$$

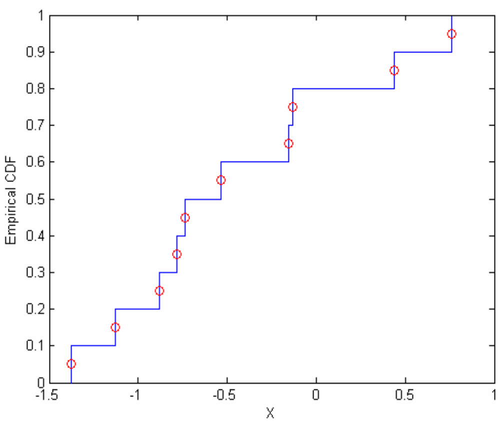

Figure 1: Example of points (red) chosen to fit an empirical CDF (blue) obtained by sampling 10 points from N(0,1).

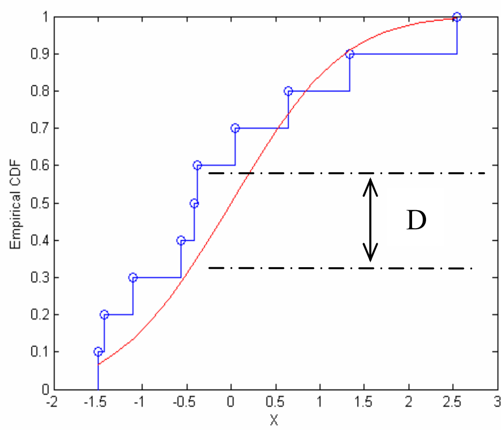

Figure 2: Example of a K-S distance between an empirical CDF (blue) and a standard normal CDF (red).

III. Conservative estimators from biased fitting

We have seen that fitting a distribution to a sample can be seen as an optimization problem. The key idea of this section is adding various constraints to the fitting problem so that the resulting estimator becomes conservative. We will also present the relationship between conservativeness and accuracy.

In most engineering problems, failure of a random variable occurs far from its mean in one direction. Therefore, we limit the requirement that the distribution is conservative to one side of the mean, i.e. the right half data points.

The milder conservative CDF can be obtained by constraining the estimator to pass *below the data points*. The more conservative CDF can be obtained by constraining the estimator to pass *below the entire empirical CDF*. They will be called, respectively, **RSPC (Right Sample Point Conservative)** and **RECC (Right Experimental CDF Conservative)**. The choice between the two constraints is a matter of balance between accuracy and conservativeness.

RSPC constraints: $$F_{\mu,\sigma}(x_i) - \frac{i}{n} \leq 0 \quad \text{for} \quad \frac{n}{2} \leq i \leq n \tag{9}$$

RECC constraints: $$F_{\mu,\sigma}(x_i) - \frac{i+1}{n} \leq 0 \quad \text{for} \quad \frac{n}{2} \leq i \leq n \tag{10}$$

When n is odd, the lower bound for i in the above two equations is modified to $(n-1)/2$.

Example

To illustrate conservative estimators we use an example with 10 sample points generated from N(0,1). P_f is defined as the probability that x is larger than 2. The exact probability of failure is 2.28%. Figure 3 shows the empirical CDF and the three estimators based on minimum RMS error.

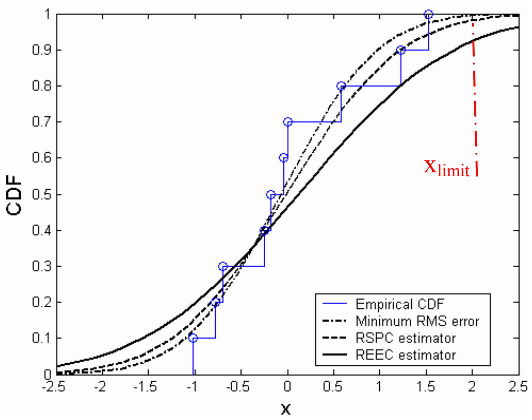

Figure 3: Example of estimators based on RMS error for a sample of 10 points generated from N(0,1)

Table 1: Values of mu, sigma and Pf of the different estimators of Figure 3

	Minimum RMS error	RSPC	RECC
μ	-0.07	0.12	-0.01
σ	0.81	1.31	0.97
P_f	0.56%	2.68%	7.65%

We can see on the graph the effect of the constraints: the RSPC estimator is shifted down to be under the 9th point, so that the values of the tail are decreased. The RECC estimator is shifted even more. The conservative curves are unconstrained on the left part of the graph, and cross the empirical curve. The mean, standard deviation, and P_f for three estimators are summarized in Table 1.

In this example, the minimum RMS error is strongly unconservative even if unbiased estimator is used. The RSPC and RECC estimators are conservative, and the RSPC is more accurate than the RECC. In order to generalize these results and come up to reliable conclusions, we will perform statistical experiments based on large number of simulations in Section V.

IV. Conservative estimators using bootstrap methods

The bootstrap principle

The bootstrap method is based on resampling; it allows us to estimate the distribution of any estimator $\hat{\theta}$ of a statistic θ (for example, the mean of a population) based on a single set of data. The bootstrap idea is to create many set of bootstrap samples by resampling with replacement from the original data and compute $\hat{\theta}$ for each bootstrap sample (Ref. [4] [5]). Then, the set of $\hat{\theta}$ provides an approximation of its distribution. Figure 4 shows a schematic representation of the bootstrap approach.

This approach allows us to estimate the distribution of any statistic without additional data. We can compute from that distribution standard error or confidence intervals. However, the bootstrap method provides an *approximation* of a distribution. Any value taken from bootstrap distribution (such as percentile or standard error) is *random* because it depends on the values of the initial sample. In order to obtain accurate results, sample sizes must not be below 100. A typical number of bootstrap can be from 500 to 5000.

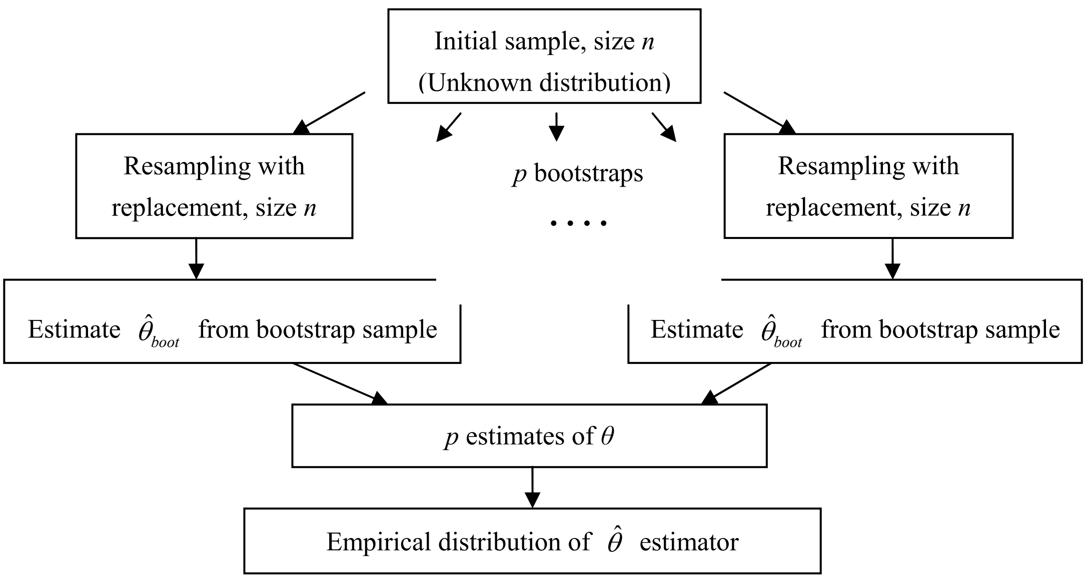

Figure 4: Schematic representation of bootstrapping. Bootstrap distribution of θ is obtained by multiple resampling (here p times) from a single set of data.

Bootstrap and probability of failure

We assume here that the distribution type of a random variable is known to be normal with unknown parameters μ and σ. We estimate these parameters by the average and the standard deviation of a sample from the population. We consider the same test case as in Section III: the data for a variable x is generated from standard normal distribution $N(0,1)$ and failure is defined as $x \geq 2$ (so the actual probability of failure is 2.28%). The sample size is taken as 100.

For each sample, we generate 5000 bootstrap resamples. We compute the mean and standard deviation of these bootstrap resamples and estimate the corresponding probability of failure. Such obtained set of 5000 $P_{f\mathrm{boot}}$ defines the empirical bootstrap distribution of the estimator of P_f.

We would like to use this distribution to minimize the risk of providing unconservative estimates of P_f. In other words, we want to find a procedure that maximizes the quantity:

$$\alpha = Prob(\hat{P}_f \geq P_f) \tag{11}$$

A procedure that satisfies Equation (11) provides an α-conservative estimator of P_f. We use here the bootstrap distribution to approximately satisfy Eq. (11). For example, if we desire α = 0.95, we can select P_f to be the 95 percentile of the bootstrap distribution of the probability of failure.

We define then two conservative estimators: the 95th bootstrap percentile and the mean of the 10% highest bootstrap values. We call these estimators *Bootstrap p95* and *Bootstrap CVaR 90* (see Figure 5). Note that any bootstrap quantile higher than 50% defines a conservative estimator; however, if we choose a very high α we may raise the mean value of $\hat{P}_f$ and be over-conservative.

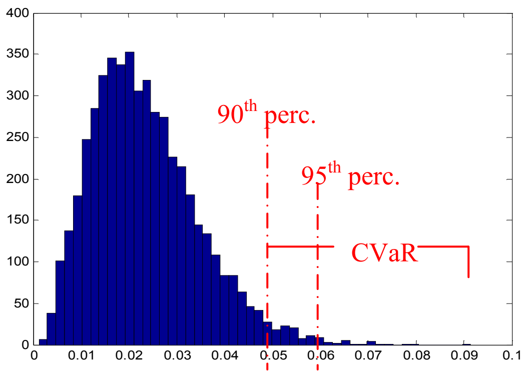

Figure 5: Conservative estimators of Pf from bootstrap distribution: 95th percentile and mean of the 10% highest values.

V. Numerical test case: samples generated from standard normal distribution

The goal of this section is to estimate the accuracy and the conservativeness of the estimators presented in sections III and IV, using a simple numerical example, where the actual distribution and P_f are known. To do so, we repeat the P_f computation a large number of times in order to extract statistical measures of our estimators.

We consider samples of size 100 generated from $N(0,1)$; the failure is defined for $x = 2$, which correspond to an actual probability of failure of 2.28%. For each sample, the different P_f estimators are computed. This procedure is repeated 1000 times. The *neutral estimator* is computed using the analytical expressions for the mean and the standard deviation. The *RSPC*, *RECC*, *Bootstrap p95* and *Bootstrap CVaR90 estimators* are computed as described in the previous sections.

We expect that most of the values will exceed the actual probability of failure, but we do not want to overestimate dramatically P_f. We present the results in the form of mean and 90% symmetric confidence interval in Table 2. The lower bound of the confidence interval shows the conservativeness of the estimator; the mean and the upper bound show the accuracy and the variability of the estimator.

Table 2: Means and confidence intervals of several estimators of P_f

	Statistics obtained over 5000 simulations	
Estimators	**90% confidence interval**	**Mean**
Neutral estimator	[0.0101 ; 0.0404]	0.0210
RSPC	[0.0153 ; 0.0786]	0.0364
RECC	[0.0217 ; 0.1230]	0.0507
Boot. p95	[0.0206 ; 0.0657]	0.0400
Boot. CVaR90	[0.0214 ; 0.0676]	0.0413
Actual	**0.0228**	

The neutral estimator of the standard deviation is, as we have seen analytically, biased; so the mean of the probability of failure estimator is less than the actual. Moreover, there is a five per cent chance to underestimate P_f by a factor of two (the lower bound of the confidence interval is 1.01%). Thus, classical estimators of CDF provide unconservatively biased estimators of P_f; this gives us a particular incentive on finding a way to improve the conservativeness of the estimations of distributions.

The RSPC and RECC estimators have a positive bias. As expected, the RECC is more conservative than the RSPC; but, as a consequence, the former is more biased and the risk of large overestimations is increased. The RECC confidence interval shows that there is a 5% chance to overestimate P_f by a factor of 6. The RECC estimators lead to about 95% conservative results for 100-point sample. The choice between the RSPC and RECC estimators will be a choice between accuracy and conservativeness.

From the confidence interval column, we see that 5% of the Bootstrap p95 are above 0.0206, which is not far from the actual probability of failure. Thus, the 95% conservativeness is approximately reached. The Bootstrap CVaR is a little bit more conservative. From the upper bound of the confidence interval, we see that for the bootstrap p95 and the CVaR90, the risk of overestimating P_f by a factor of three is of the order of 5%.

Bootstrap methods appear to be much more efficient than the optimization based methods (RSPC and RECC). For an equivalent level of conservativeness (95%), the bias is reduced and the risk of large overestimations is much lower. However, bootstrap cannot be used with very small sample sizes. In that case, optimization based methods can be used instead.

VI. Application to a composite panel under thermal loading

In this section, we apply the conservative estimate of probability to an example of a composite panel under mechanical and thermal loading, which is used for the wall of a hydrogen tank. The design of composite laminates for liquid hydrogen tanks involves several challenges: the cryogenic operating temperatures develop large residual strains due to the different coefficients of thermal expansion of the fiber and the matrix.

Qu et al. (2003) performed the deterministic and probabilistic design optimizations of composite laminates under cryogenic temperatures, using response surface approximations for probability of failure calculations. Acar and Haftka (2005) found that using CDF estimations for strains improves the accuracy of probability of failure calculation.

We analyze here the problem that is addressed by Qu et al. (2003); hence the geometry, material parameters and the loading conditions are taken from that paper. Our aim is to explore further the possibilities to improve the estimation of the probability of failure calculations in a conservative way.

Problem definition:
We consider the design of a composite panel that is the wall of a hydrogen tank. It is subject to resultant stress caused by mechanical loading (N_x is 33 MPa and N_y is 16 MPa) and thermal loading due to the operating temperature 20K (Figure 6). The objective is to minimize the weight of the composite panel that is a symmetric balanced laminate with two ply angles (that means an eight-layer composite). The design variables are the ply angles [$\pm\theta_1$, $\pm\theta_2$] and the ply thickness t_1 and t_2. The geometry and loading condition are shown in Figure 6.

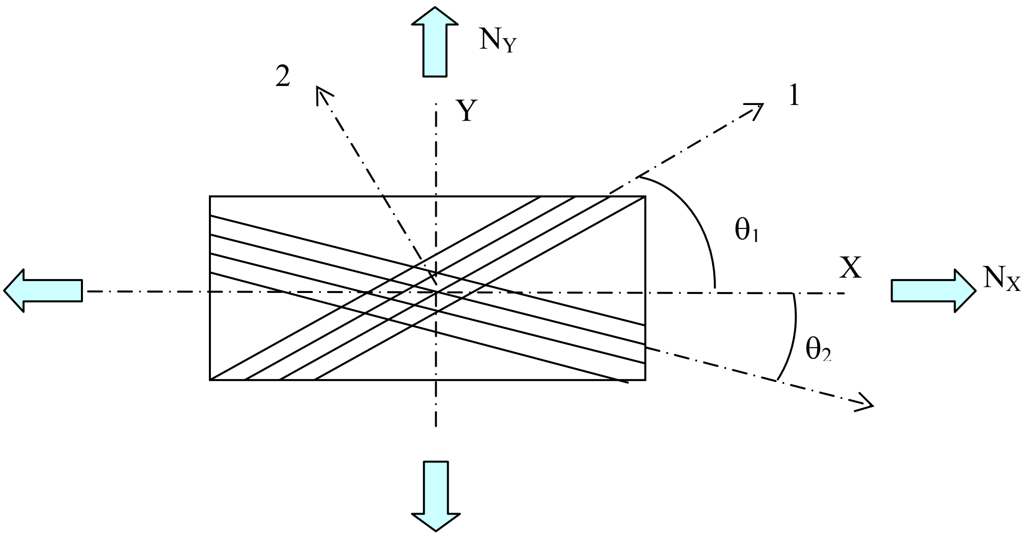

Figure 6: Geometry and loading of the cryogenic laminate

The material used in the laminates is IM600/133 graphite-epoxy, defined by the following mechanical properties:

- Elastic properties (E_1, E_2, G_{12} and ν_{12})
- Coefficients of thermal expansion (α_1 and α_2)
- Stress-free temperature T_{zero}
- Failure strains (ε_1^L, ε_1^U, ε_2^L, ε_2^U, γ_{12}^U).

The minimum thickness of each layer is taken as 0.05mm; that corresponds to manufacturing constraints, but this is also necessary to prevent hydrogen leakage. The failure is defined as the *first ply failure* that is when the strain values of the first ply exceed failure strains.

The deterministic optimization problem is formulated as:

$$\textbf{\textit{Minimize}} \quad \mathbf{h = 4(t_1 + t_2)}$$

$$\textbf{\textit{s.t.}} \quad \begin{aligned} & t_1, t_2 \geq 0.5\times10^{-3} \\ & \varepsilon_1^L \leq S_F\varepsilon_1 \leq \varepsilon_1^U \\ & \varepsilon_2^L \leq S_F\varepsilon_2 \leq \varepsilon_2^U \\ & S_F|\gamma_{12}| \leq \gamma_{12}^U \end{aligned} \qquad (12)$$

where S_F is a safety factor chosen at 1.4.

The deterministic and probabilistic optimizations were solved by Qu et al. (2003), as summarized in Table 3.

Table 3: Deterministic optima found by Qu et al (2003)

θ_1 (deg)	θ_2 (deg)	t_1 (mm)	t_2 (mm)	h (mm)
27.04	27.04	0.254	0.381	2.540
0	28.16	0.127	0.508	2.540
25.16	27.31	0.127	0.508	2.540

Calculation of the probability of failure

Given the material properties and the design variable, we calculate the ply strains using Classical Lamination Theory (CLT). Due to manufacturing variability, the material properties and failure strains are considered as random variables. All random variables are assumed to follow uncorrelated normal distributions. The coefficients of variation are given in Table 4. E_2, G_{12}, α_1 and α_2 are function of the temperature; since the design must be feasible for the entire range of temperature, strain constraints were applied at 21 different temperatures, which were uniformly distributed from 20 to 250K. As a consequence, we first calculate the mean value of the random variables for a given T and then generate the random number. The mean of the other parameters are given in Table 5.

Table 4: Coefficients of variation of the random variables

E_1, E_2, G_{12} and ν_{12}	α_1 and α_2	T_{zero}	ε_1^L and ε_1^U	ε_2^L, ε_2^U and γ_{12}^U
0.035	0.035	0.03	0.06	0.09

Table 5: Mean of random parameters

E_1	ν_{12}	T_{zero}	ε_1^L	ε_1^U	ε_2^L	ε_2^U	γ_{12}^U
$21.5\text{x}10^6$	0.359	300	-0.0109	0.0103	-0.013	0.0154	0.0138

The critical strain is the transverse strain on the first ply (direction *2* in Figure 6*)*. The variable we consider is the difference between the critical strain and the failure strain. The probability of failure is then $P_f = 1 - F(0)$, where F is the CDF of the difference between the strain and the strain failure.
The software ARENA is used to determine which distribution type fits the best the critical strain data; it shows that Normal distribution is the best choice.

Results

100 MCS are run in the configuration of the first optimum (27.04, 27.04) and the estimations of P_f are computed. In order to estimate the distribution of the estimators, 1000 simulations are done. Finally, 100,000 samples are generated to estimate the actual distribution of the strains.

Table 6: Percentiles of five estimators of Pf (x 10^{-4}) for deterministic optimum of cryogenic laminate

	Statistics obtained over 5000 simulations	
Estimators	**90% confidence interval**	**Mean**
Neutral estimator	[2.1 ; 26.5]	6.9
RSPC	[5.2 ; 82.7]	23.1
RECC	[8.1 ; 140]	40.9
Boot. p95	[7.7 ; 67.5]	29.8
Boot. CVaR90	[8.5 ; 72.7]	32.2
Actual	**8.2**	

Table 6 summarizes the results of five different estimators for the cryogenic laminate. The use of classical estimators of μ and σ lead here to a five per cent chance of underestimating P_f by a factor of four, whereas the *RECC* estimator is 95% conservative (the 5% percentile is equal to the actual P_f). However, the right tail of its distribution and the bias are very large. The *RSPC* estimator is less conservative but both variability and bias are substantially reduced.

Bootstrap p95 is almost 95% conservative; *CVaR90* is a little bit more than 95% conservative. The upper bounds of their confidence interval are two times lower than the *RECC* estimator, for an equivalent level of conservativeness.

The overall performance of the conservative estimators is not as good as for the numerical example of section IV. We explain these differences by two main reasons:

- First, the actual distribution of the strains may not be exactly normal, that increases the error in the CDF fitting
- Second, the actual probability of failure is of the order of 10^{-4} instead of 10^{-2} previously. Since we estimate here the value of the CDF at a farther point in the tail, the variability is logically increased.

VII. Concluding remarks

The estimation of the probability of failure of a structure is crucial in reliability based design. In a context of expensive numerical experiments, or when the data samples provided are small, the direct use of Monte-Carlo Simulation is not possible, and an estimation of continuous distributions is necessary. However, we have seen that the classical ways to estimate a CDF from a sample of values may lead to dangerous underestimations of the probability of failure.

In this paper, several methods of estimating CDF based on finite samples are tested. We first implemented a method constraining the estimated CDF to be under the empirical CDF. Then, we have shown how to use the bootstrap method to obtain distributions of probability of failure estimators, and how to use this bootstrap distribution to define conservative estimators.

In the case of samples generated from standard normal distribution, numerical test case shows that both methods improve the chance of the estimation to be conservative. Bootstrap based estimators appear to provide much better results than optimization based methods. However, optimization based methods can be used when sample size is very small, where bootstrap cannot be used.

We have also applied these procedures to estimate the probability of failure of composite laminates at cryogenic temperatures. We found that estimating the probability of failure from the mean and standard deviation of a sample lead to a five per cent chance of underestimating *Pf* by a factor of four. The conservative estimation allows us to reduce that risk and avoid the use of additional safety factors. However, controlling the uncertainty of the estimation is crucial to limit the risks of oversizing.

For both analytical example and the composite laminate, we found that conservative estimates based on a bootstrap approach outperform one-sided fits to the experimental CDF. That is, for the same confidence in the conservativeness of the probability estimate, the penalty in the accuracy of the estimate is substantially smaller.

Acknowledgements

Funding for this research was provided by National Science Foundation (Grant # 0423280).

References

[1] "Integration of Probabilistic Methods into the Design Process," Society of Automotive Engineers, Aerospace Information Rept. 5080, 1997

[2] Ben-Haim, Y., and Elishakoff, I., *Convex models of Uncertainty in Applied Mechanics,* Elsevier, Amsterdam, 1990

[3] Neal, D.M., Matthews, W. T. and Vangel, M.G., 1992, "Uncertainty in Obtaining High Reliability fro Stress-Strength Models," *Proceedings of the 9th DOD-/NASA/FAA Conference on Fibrious Composites in Structural Design*, Vol. 1, Dept. of Defense, Lake Tahoe, NV, 1991, pp. 503-521

[4] Efron, B., 1982, *The Jackknife, the Bootstrap, and other Resampling Plans*, SIAM, Philadelphia

[5] Chernick, M.R., 1999, *Bootstrap Methods, a Practitioner's guide*, (Wiley series in probability and statistics) A Wiley-Interscience publication

[6] Eric W. Weisstein. "*Sample Variance Distribution.*" From *MathWorld*--A Wolfram Web Resource. http://mathworld.wolfram.com/SampleVarianceDistribution.html