

Bootstrapping and Separable Monte Carlo Simulation Methods Tailored for Efficient Assessment of Probability of Failure of Structural Systems

Musarrat Jehan and Efstratios Nikolaidis
University of Toledo

ABSTRACT

There is randomness in both the applied loads and the strength of systems. Therefore, to account for the uncertainty, the safety of the system must be quantified using its reliability. Monte Carlo Simulation (MCS) is widely used for probabilistic analysis because of its robustness. However, the high computational cost limits the accuracy of MCS. Smarslok et al. [2010] developed an improved sampling technique for reliability assessment called Separable Monte Carlo (SMC) that can significantly increase the accuracy of estimation without increasing the cost of sampling. However, this method was applied to time-invariant problems involving two random variables. This paper extends SMC to problems with multiple random variables and develops a novel method for estimation of the standard deviation of the probability of failure of a structure. The method is demonstrated and validated on reliability assessment of an offshore wind turbine under turbulent wind loads. The results show the method can reduce the computational cost by two orders of magnitude compared to standard MCS for failure probabilities as low as 10^{-4} .

CITATION: Jehan, M. and Nikolaidis, E., "Bootstrapping and Separable Monte Carlo Simulation Methods Tailored for Efficient Assessment of Probability of Failure of Structural Systems," *SAE Int. J. Mater. Manf.* 8(3):2015, doi:10.4271/2015-01-0420.

INTRODUCTION

Many of the existing methods require extensive computational effort of standard Monte Carlo Simulation (MCS). However, the accuracy of MCS analysis is limited by the sample size, and to increase the accuracy in the estimation of the probability of failure using MCS, very large samples are required. MCS may not be capable of providing the desired accuracy when sampling is limited due to expensive structural analysis especially when the calculation of the limit state function involves Finite Element Analysis (FEA) of large-scale models.

Smarslok et al. [2006, 2008, 2010] and Ravishankar [2010] developed a sampling method named Separable Monte Carlo (SMC) simulation. When the performance function of the system can be written in terms of functions with no common variables, SMC can take advantage of this property to increase the sampling domain. An example is system whose performance function is the difference of the capacity (yield stress) and the applied stress (demand). In SMC method, the response and capacity are sampled separately as opposed to response and capacity being sampled as a pair in standard MCS. Hence, the user can compare many or all possible combinations of values of the response and capacity thereby increasing the accuracy in the estimator of the probability of failure without increasing the computational cost. In SMC method, sample values of the random variables are reused to increase efficiency of calculation. However, reusing the variables introduces correlation between the failure indication functions. The user needs to estimate the correlation coefficient between the failure indicator values in order to find the standard deviation of the probability of failure. This estimation is

challenging because it requires the correlation coefficient of the failure indicator functions is small (e.g., 10^{-4}) but still affects the standard deviation of the probability of failure significantly.

The paper presents a new, efficient method to assess the reliability of systems under random loading. The method is based on SMC simulation. A closed form equation is developed to estimate the standard deviation of the probability of failure. This method is applicable to static problems involving multiple variables.

METHOD

Problem Formulation

It is impractical to estimate the standard deviation of the probability of failure for small correlation coefficients (for example 10^{-3}) using standard MCS. The standard error of the correlation coefficient assuming a large sample is Freund and Williams, [1996]:

$$\sigma_\rho = \frac{1 - \rho^2}{\sqrt{n - 1}} \quad (1)$$

In the above equation ρ is the correlation coefficient, n is the sample size and σ_ρ the standard deviation of the correlation coefficient. A sample size of 10^8 is needed to estimate the correlation coefficient with a prescribed coefficient of variation of 0.1.

The key idea of the proposed method is to calculate the correlation coefficients between a pair of failure indicator function values separately for each number of common elements that the indicator functions may share. In SMC, some pairs of failure indicator function values may share common elements of capacity or response. Sharing one or more common elements introduces correlation between failure indicator functions. Not all the failure indicator values produced in SMC method are correlated. Some of the values have no common variables; consequently the correlation coefficient between the failure indicator functions that share no common element is zero. Some of the failure indicator values involve common variables. The correlation coefficient increases with the number of common elements that the failure indicator functions share. Thus, the cases where there are one or more common elements are considered separately, and the corresponding conditional correlation coefficients are calculated independently. The weighted average correlation coefficient is then calculated, and used in the calculation of the standard deviation of the probability of failure. The benefit of separation of the failure indicator functions into groups based on the number of common elements they share is that the method eliminates all the uncorrelated pairs from the calculation. This reduces the problem to the calculation of a much larger correlation coefficient.

Calculation of the Correlation Coefficient of Two Failure Indicator Functions

First, the paper demonstrates the derivation of the equation for the standard deviation of the probability of failure in SMC, on a problem of rod under tensile load where three replications are performed. Then the paper presents an equation for the standard deviation for n replications. The two random variables involved in the performance function are stress (S) in the rod and strength (S_y) of the material the rod is made of. Stress is considered as demand and strength as supply. Figure 1 depicts the probability distributions of stress and strength and the failure zone.

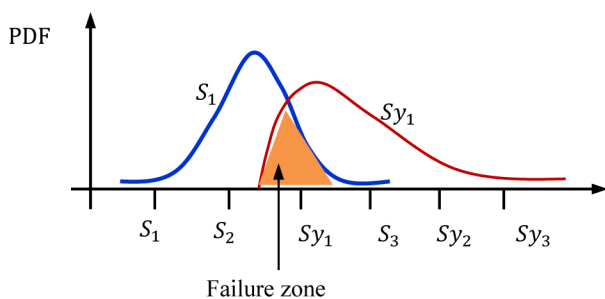


Figure 1. Stress and strength distribution with the failure zone

Three random sample values of stress (S_1, S_2, S_3) and strength (S_{y1}, S_{y2}, S_{y3}) are generated. The method will then be presented in the general case for N random values. The failure criterion is that failure occurs when the demand (stress) exceeds the supply (strength), that is, failure occurs when $S - S_y < 0$. The limit state function is $G[(S) - (S_y)]$.

The failure indicator function $I(G)$ is defined as follows;

$I = 1$, when $G < 0$ and

$I = 0$, otherwise.

In SMC simulation, all three sample values of stress are compared with all three sample values of strength. Therefore, a total of nine (3×3) distinct combinations are obtained. Nine failure indicator functions (I) are calculated for all nine combinations as shown in Table 1.

The first subscript of I indicates stress and the second indicates strength. Therefore, symbol I_{ij} denotes the failure indicator function for the i^{th} sample value of the stress and the j^{th} sample value of the strength.

Some of the pairs of failure indicator (I) are correlated because they share the same stress or same strength values or both. The remaining pairs are uncorrelated because they have no common variables. For example I_{11} and I_{12} are correlated because the same stress value S_1 is compared with two strength sample values when calculating these indicator functions. Similarly, I_{12} and I_{22} are correlated because they share a common strength value. Failure indicator functions I_{11} and I_{22} are uncorrelated because they have different stress and strength values.

Table 1. SMC sampling method using two variables (S_i and S_{yj}) and all possible combinations of stress and strength.

Combinations	S_i	S_{yj}	I_{ij}
1	S_1	S_{y1}	I_{11}
2		S_{y2}	I_{12}
3		S_{y3}	I_{13}
4	S_2	S_{y1}	I_{21}
5		S_{y2}	I_{22}
6		S_{y3}	I_{23}
7	S_3	S_{y1}	I_{31}
8		S_{y2}	I_{32}
9		S_{y3}	I_{33}

On the other hand, I_{11} and I_{11} has a correlation coefficient of one since these failure indicator functions share both the stress and strength.

To identify all correlated pairs of failure indicator values, a total of (9×9) combinations of all possible pairs of failure indicator functions are listed in Table 2. The correlation coefficient of values of the failure indicator function with common stress is denoted by ρ_s , and the correlation of indicator function values with common strength is denoted by ρ_{sy} . The uncorrelated pairs are marked by 0, and the indicator values that are perfectly correlated are equal to one in the table.

The pairs are arranged into four sets based on the common terms.

Table 2. Combination between failure indicator pairs

	I_{11}	I_{12}	I_{13}	I_{21}	I_{22}	I_{23}	I_{31}	I_{32}	I_{33}
I_{11}	1	ρ_s	ρ_s	ρ_{sy}	0	0	ρ_{sy}	0	0
I_{12}		1	ρ_s	0	ρ_{sy}	0	0	ρ_{sy}	0
I_{13}			1	0	0	ρ_{sy}	0	0	ρ_{sy}
I_{21}				1	ρ_s	ρ_s	ρ_{sy}	0	0
I_{22}					1	ρ_s	0	ρ_{sy}	0
I_{23}						1	0	0	ρ_{sy}
I_{31}							1	ρ_s	ρ_s
I_{32}								1	ρ_s
I_{33}									1

Symmetric

Set 1. Pairs of Identical Terms

$(I_{11}, I_{11}), (I_{22}, I_{22}), (I_{33}, I_{33}), (I_{12}, I_{12}), (I_{13}, I_{13}), (I_{21}, I_{21}), (I_{23}, I_{23}), (I_{31}, I_{31}), (I_{32}, I_{32})$

The terms in this set are those pairs that share both common stress and common strength. The total number of terms in this set is 3^2 . These terms can be represented by (I_{ij}^2) where $(i = 1 \text{ to } 3)$ and $(j = 1 \text{ to } 3)$ or by (I_{ij}, I_{ij}) where $(i = 1 \text{ to } 3), (j = 1 \text{ to } 3)$. The number of terms in the general case of N indicator functions is N^2 .

The expected (mean) value of the square of the failure indicator function is:

$$E(I_{ij}^2) = \widehat{pf} \cdot 1^2 + (1 - \widehat{pf}) \cdot 0^2 = \widehat{pf} \quad (2)$$

In this case, the correlation coefficient between the failure indicator values is one.

Set 2. Pairs with Common Stress

$(I_{11}, I_{12}), (I_{11}, I_{13}), (I_{12}, I_{13}), (I_{21}, I_{22}), (I_{21}, I_{23}), (I_{22}, I_{23}), (I_{31}, I_{32}), (I_{31}, I_{33}), (I_{32}, I_{33})$

The failure indicator functions of these pairs have the same I^s subscript.

The total number of terms in this set is 3^2 (or $\frac{3^3 - 3^2}{2} = \frac{3^2(3-1)}{2}$). The above equation for the total number of terms for a set of N elements is $(\frac{N^2(N-1)}{2})$.

The terms can be represented by (I_{ij}, I_{ik}) where $(i = 1 \text{ to } 3), (j = 1 \text{ to } 3), (k = 1 \text{ to } 3)$. The expectation of the product of failure indicator functions that share the same stress is:

$$E(I_{ij}, I_{ik}) = \rho_s \cdot \widehat{pf} \cdot (1 - \widehat{pf}) + \widehat{pf}^2 \quad (3)$$

Set 3. Pairs with Common Strength

$(I_{11}, I_{21}), (I_{11}, I_{31}), (I_{21}, I_{31}), (I_{12}, I_{22}), (I_{12}, I_{32}), (I_{22}, I_{32}), (I_{13}, I_{23}), (I_{13}, I_{33}), (I_{23}, I_{33})$

The terms in this set share the same strength. They have the same 2^{nd} subscript. The total number of terms in this set is 3^2 (or $\frac{3^3 - 3^2}{2} = \frac{3^2(3-1)}{2}$). For N pairs of stress and strength, the total number of terms is $(\frac{N^2(N-1)}{2})$.

The terms are represented by (I_{ij}, I_{kj}) where $(i = 1 \text{ to } 3), (j = 1 \text{ to } 3), (k = 1 \text{ to } 3)$. The expectation of the product of failure indicator functions that share the same strength is:

$$E(I_{ij}, I_{kj}) = \rho_{sy} \cdot \widehat{pf} \cdot (1 - \widehat{pf}) + \widehat{pf}^2 \quad (4)$$

Set 4. Pairs with No Common Stress or Strength

The terms in this set share no common stress or strength. The total number of terms in this set is the rest of the terms, $81 - 3 \times 9 = 54$. For N pairs of stress and strength, the total number of terms with no common variable is:

$$(N^2)^2 - (N^2 + N^2 + N^2) = (N^4 - 3N^2) = (N^4 - N \cdot N^2) = N^3 (N - 1).$$

These pairs are uncorrelated. Therefore, the correlation coefficient between the failure indicator values with no common elements is 0. The terms can be represented by (I_{ij}, I_{kl}) where $(i = 1 \text{ to } 3), (j = 1 \text{ to } 3), (k = 1 \text{ to } 3), (l = 1 \text{ to } 3)$. The expectation of the product of terms with no common elements is,

$$E(I_{ij} I_{kl}) = \widehat{pf}^2 \quad (5)$$

The estimator of the probability of failure in SMC, when using all possible combinations of the generated random variables, is:

$$\widehat{pf} = \frac{1}{3 \times 3} \sum_{i=1}^3 \sum_{j=1}^3 I_{ij} \quad (6)$$

The variance of $\widehat{pf}$ is estimated as follows:

The variance of the estimator of the probability of failure is,

$$\begin{aligned} \sigma_{\widehat{pf}}^2 &= Var(\widehat{pf}) = E[(\widehat{pf} - E[\widehat{pf}])^2] \text{ or} \\ \sigma_{\widehat{pf}}^2 &= E[\widehat{pf}^2] - \widehat{pf}^2 \end{aligned} \quad (7)$$

Substituting the value of $\widehat{pf}$ from equation (7) yields,