JOURNAL OF THE
CHUNGCHEONG MATHEMATICAL SOCIETY
Volume **21**, No. 4, December 2008

# A STUDY ON THE TECHNIQUES OF ESTIMATING THE PROBABILITY OF FAILURE

Yong-Kyun Lee* and Dae-Sik Hwang**

Abstract. In this paper, we introduce the techniques of estimating the probability of failure in reliability analysis. The basic idea of each technique is explained and drawbacks of these techniques are examined.

## 1. Introduction

Reliability analysis is simply an area of calculating the probability of failure. That is the multiple integration of joint probability density function(PDF) in failure surface of limit state function. In general, it is almost impossible to obtain the PDF of arbitrary random variables. Even if the PDF is available, calculating the multiple integration is also a difficult matter. Therefore, one approach for calculating the multiple integration is to use analytical approximation methods. The representative approximation integration techniques for this approach are fast probability integration technique and sampling technique.

The fast probability integration technique can be grouped into two types, namely, first-order reliability method(FORM) and second-order reliability method(SORM).

Another approximation integration technique is the sampling technique, such as the crude Monte Carlo sampling, the importance sampling and the descriptive sampling.

In this study, we will introduce the techniques of estimating the probability of failure. First, the probability of failure is explained. Then, the basic idea of the approximation of integration techniques is introduced for the multiple integration. Also, drawbacks of these techniques are found.

Received October 20, 2008; Accepted November 20, 2008.
2000 Mathematics Subject Classification: Primary 60K10; Secondary 62D05.
Key words and phrases: FORM, SORM, crude Monte Carlo sampling, importance sampling, descriptive sampling.

## 2. Techniques of estimating the probability of failure

The first step in evaluating the reliability or probability of failure of a structure is to decide on specific performance criteria and the relevant parameters, called the basic variables $X_i$, and the functional relationships among them corresponding to each performance criterion. Mathematically, this relationship of performance function can be described as

$$Z = g(X_1, X_2, \cdots, X_n). \tag{2.1}$$

The failure surface or the limit state of interest can then be defined as $Z = 0$. This is the boundary between the safe and unsafe regions in the design parameter space, and it also represents a state beyond which a structure can no longer fulfill the function for which it was designed. Assuming $X_1$ and $X_2$ are the two basic random variables, the failure surface and the safe and unsafe regions are shown in Figure 1.

Using Eq. (2.1), we find that failure occurs when $Z < 0$. Therefore, the probability of failure, $P_f$, is given by the integral

$$P_f = \int \cdots \int_{g(\mathbf{x})} f_X(x_1, x_2, \cdots, x_n) dx_1 dx_2 \cdots dx_n$$

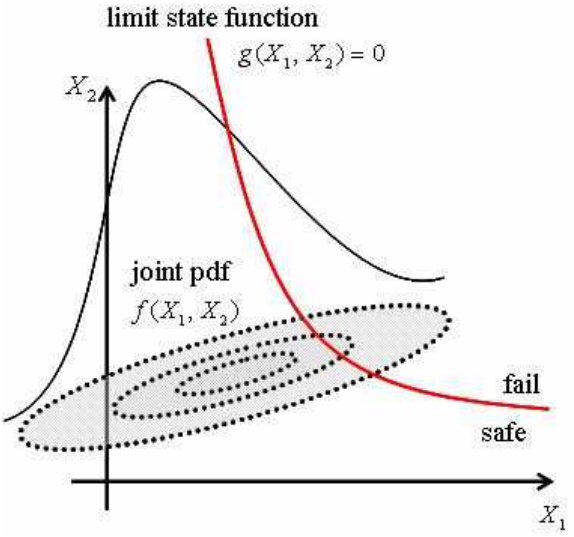


FIGURE 1. Probability of failure

in which $f_X(x_1, x_2, \cdots, x_n)$ is the joint probability density function for the basic random variables $X_1, X_2, \cdots, X_n$, and the integration is performed over the failure region, that is, $g(\mathbf{x}) < 0$. If the random variables are statistically independent, then the joint probability density function may be replaced by the product of the individual probability density functions in the integral.

The limit state function plays an important role in the development of structural reliability analysis methods. A limit state can be an explicit or implicit function of the basic random variables, and it can be in simple or complicated form. Reliability analysis techniques have been developed corresponding to limit states of different types and complexity.

### 2.1. Fast probability integration techniques

#### 2.1.1. Fast-order reliability method

Most of the earlier approximation used uncertain information represented only by their first two moments, i.e. its mean and standard deviation.

C. A. Cornell [4] proposed two ideas : (a) the use of only first- and second-moments to characterize the entire set of random variables, and (b) the linearization by means of the Taylor series expansion of the limit state function $g(X)$ at some appropriate checking point. The measure of reliability is given by the reliability or safety index

$$\beta = \frac{\bar{g}(X)}{\sigma(g)},$$

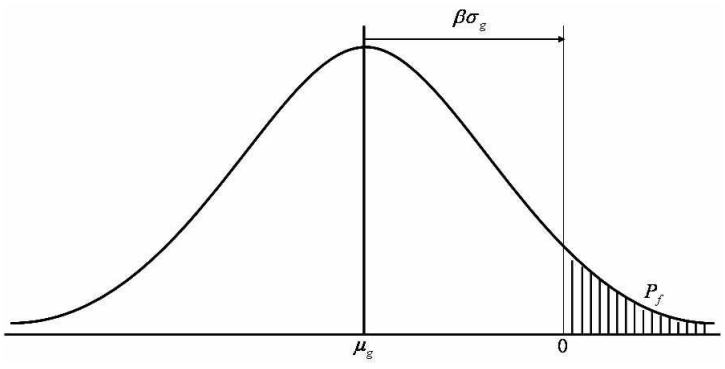


FIGURE 2. Reliability index and probability of failure

where $\bar{g}(X)$ is the mean value and $\sigma(g)$ is the standard deviation of $g(X)$. The method for evaluating $\beta$ by expanding around the mean value of the random variables is known as the mean value first order second-moment method(MVFOSM). This approximation technique changed the original complex probability problem into a simple problem. Figure 2 shows the relation between reliability index and probability of failure.

A. M. Hasofer and L. C. Lind [7] proposed to linearize about a point which lies on the failure surface and which corresponds to the maximum likelihood of failure occurrence(HL method). This point is known as the design point or most probable failure point(MPP). The reliability is measured through the Hasofer-Lind safety index and it is defined as the minimum distance from the origin to the failure surface as Figure 3.

R. Rackwitz and B. Fiessler [10] extended the HL method to include random variable distribution information, which is denoted as the HL-RF method. The HL-RF method requires the least amount of storage and computation in each step. For most situations this method does not only converges, but also converges faster.

But, three serious drawbacks of the FORM approach include : (1) evaluation of reliability by linearizing the limit state function about the mean values leads to erroneous estimates for limit state functions with highly nonlinear, or for large coefficients of variation, (2) the reliability

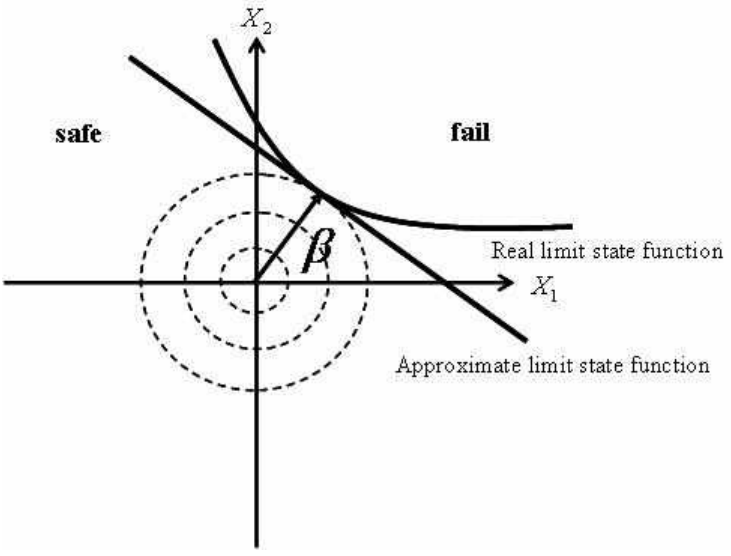


FIGURE 3. FORM technique

measure is distinct for different equivalent formulations of the same limit state function and (3) when the limit state function is complex and highly nonlinear, may converge slowly or even result in divergence.

### 2.1.2. Second-order reliability method

The SORM approach was first explored by B. Fiesslor et al. [5] using various quadratic approximations. A simple closed-form solution for the probability computation using a second-order approximation, $P_f$, was given by K. Breitung [2] using the theory of asymptotic approximations as

$$P_f = \Phi(-\beta) \prod_{i=1}^{n-1} (1 + \beta\kappa_i)^{-1/2}, \tag{2.2}$$

where $\kappa_i$ denotes the principal curvatures of the limit state at the minimum distance point, and $\beta$ is the reliability index using the FORM.

L. Tvedt [11] developed alternative SORM formulations to take care of these problems. Tvedt's method uses a parabolic and a general second-order approximation to the limit state, and it does not use asymptotic approximations.

A. D. Kiureghian et al. [8] approximated the limit state by two semiparabolas using curve-fitting at several discrete points around the design point and used both sets of curvature in Eq. (2.2). This strategy

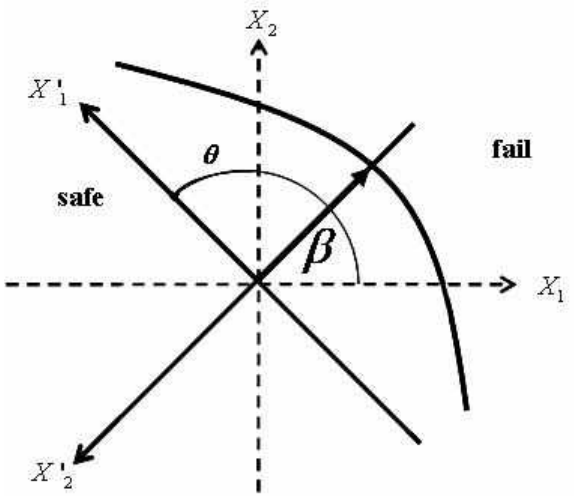


FIGURE 4. SORM technique

helps to avoid the computation of a full second-derivative matrix using the original limit state and is efficient for problems with a large number of random variables. Figure 4 shows the idea of the SORM. But, if the value of reliability index is low, the SORM estimate could be inaccurate. Also, the SORM approach could be obtained to the wrong curvature due to the numerical noise and the computation time is increased when the number of random variables increase.

### 2.2. Sampling techniques

#### 2.2.1. Crude Monte Carlo sampling

The crude Monte Carlo sampling is one of the techniques which estimate the probability of failure. A basic concept of this technique is shown in Figure 5.

This sampling selects the values of uncertain variables randomly according to their probability distribution functions. As an example, for uniform distribution any value within the valid range is selected by a uniform random number generator. For normal distributions, random values are selected, but values near the mean will be generated more frequently than those at the extremes. It is known that if the value of $g(\mathbf{x})$ is less than zero, it indicates failure. Let $n_f$ be the number of simulation cycles when $g(\mathbf{x})$ is less than zero and let $N$ be the total number of simulation cycles. Therefore, an estimate of the probability of failure

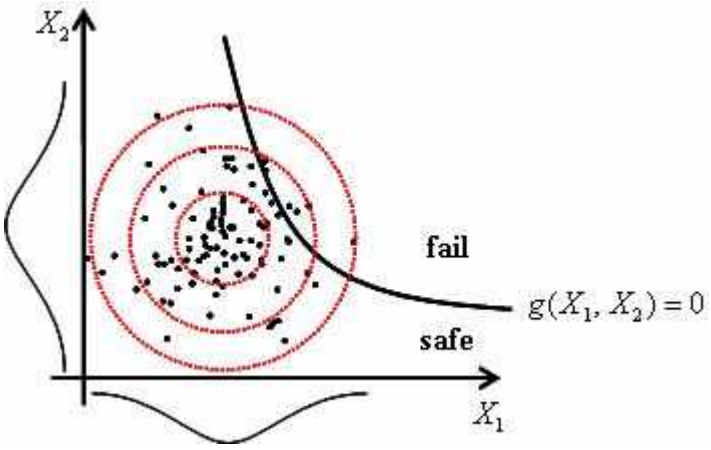


FIGURE 5. Crude Monte Carlo sampling

can be expressed as

$$P_f = \int \cdots \int I[g(\mathbf{x}) < 0] f_X(\mathbf{x}) d\mathbf{x}$$
$$\approx \frac{1}{N} \sum_{i=1}^{N} I[g(\mathbf{x}_i) < 0] \approx \frac{n_f}{N},$$

where $I[\cdot]$ is an indicator function which equals 1 if $[\cdot]$ is true and 0 if $[\cdot]$ is false(see [1], [3], [13]).

But the crude Monte Carlo sampling approach requires huge samples in order to obtain an accurate estimator and is therefore expensive.

### 2.2.2 Importance sampling

The basic idea of importance sampling is to concentrate the distribution of sampling points in the region of most importance, the area that mainly contributes to the failure probability, instead of spreading them out evenly among the whole range of possible values of the basic variables. One method to achieve this is illustrated in Figure 6.

For importance sampling, a new sampling PDF $h_X(\mathbf{x})$ is defined so as to obtain samples in the desired region. $h_X(\mathbf{x})$ is known as the sampling density function. The probability of failure is given by

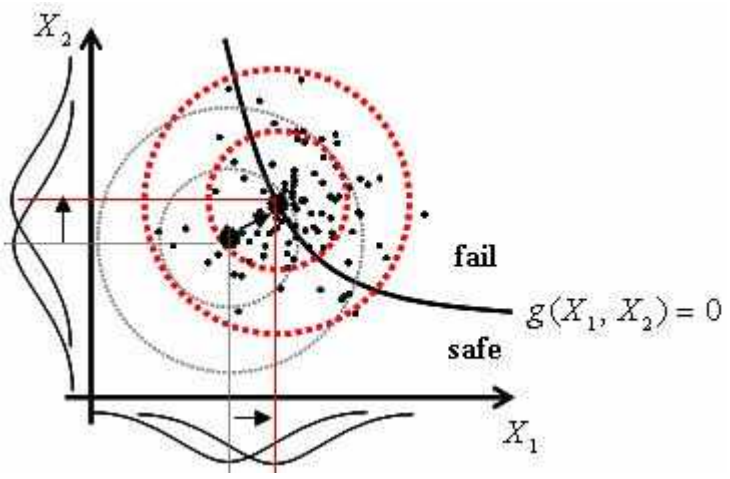


FIGURE 6. Importance sampling

$$P_f = \int \cdots \int I[g(\mathbf{x}) < 0] \frac{f_X(\mathbf{x})}{h_X(\mathbf{x})} h_X(\mathbf{x}) d\mathbf{x}$$
$$\approx \frac{1}{N} \sum_{i=1}^{N} I[g(\mathbf{x}_i) < 0] \frac{f_X(\mathbf{x}_i)}{h_X(\mathbf{x}_i)},$$

where $N$ is the number of simulations and $\mathbf{x}_i$, $(x_1, x_2, \cdots, x_n)_i$ , is the set of values of the basic random variables at the th simulation(see [6], [9]).

For importance sampling, the accuracy of the estimate depends on the choice of the sampling density function. But a clear choice method is the situation not to exist and have the defect to be altered according to the problem.

### 2.2.3 Descriptive sampling

The descriptive sampling was proposed in order to avoid the set variability in simulation studies by E. Saliby [12]. When using the standard simple random sampling or crude Monte Carlo approach, two kinds of variation are presented in a randomly generated sample - one is related to the set of values and the other to their sequences. The set variability refers to the deviations between true sample parameters and the corresponding assumed parameters for the input design variables. The sequence variability refers to the lack of order of randomness for the sample values.

The descriptive sampling is designed to remove, or at least to reduce set variability. It is based on the deterministic selection of input sample values and their random permutations. Symbolically, it follows that

*descriptive sampling = deterministic set × random sequence*

while

*simple random sampling = random set × random sequence.*

The only additional requirement to use the descriptive sampling instead of the crude Monte Carlo sampling is to know, in advance, the input sample size, which is related to a full size of simulation run. Once the sample size is known at least approximately, the set of values are defined for each input random variable $X_j$, $j = 1, 2, \cdots, n_v$, using the inverse transform method, so that

$$X_j^i = F_{X_j}^{-1}[(i - 0.5)/n_s], \quad i = 1, 2, \cdots, n_s,$$

where

$$F_X^{-1}(R), \quad R \in (0,1)$$

is the inverse transform for the particular input distribution.

Completing the descriptive sampling generation process, each of the $n_s$ sets of input values are used in a random sequence in each simulation run. Now, unlike with simple random sampling, set values are the same for all replicated runs in a simulation experiment. This random shuffling process is easily accomplished by sampling, without replacement, the descriptive set of values. Figure 7 shows the concept of the descriptive sampling.

But there are cases where the sample size determination is not trivial. For example, in a queue network simulation, with several service stations and complex routing rules, it is difficult to predict the sample size for the service times in each station. Another difficulty is that the sample size may vary among runs only due to the input sample variability. This error may lead to do the poor simulation results.

## 3. Conclusion

This paper introduced the techniques of estimating the probability of failure. In the techniques of estimating the probability, there are fast probability integration technique and sampling technique.

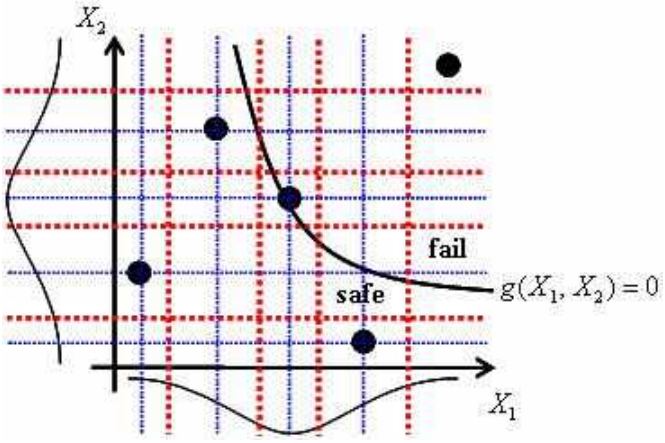


FIGURE 7. Descriptive sampling

The fast probability integration technique can be grouped into FORM and SORM. The FORM represents the information on mean, variance, and distribution of random variables in order to approximate reliability index, which is the relative measure about the probability of failure. The SORM that is an updated result of FORM includes additional information about the curvature of the limit state. However, the FORM has the limit in the accuracy though it is efficient. And the SORM is a complex method to calculate the curvature and has the numerical errors sometimes.

Another approximation integration technique is the sampling technique, such as the crude Monte Carlo, the importance sampling and the descriptive sampling. The crude Monte Carlo sampling is as well-known as a sampling technique. A basic concept of the crude Monte Carlo is simple and it can calculate the probability of failure without transformation about the limit state function. However, the crude Monte Carlo sampling requires huge samples and is expensive.

The importance sampling is used for variance reduction of the estimate of integration. It concentrates the distribution of sampling point in the region of most importance. But the drawback of the importance sampling is that the accuracy of the importance sampling estimate depends on the choice of the sampling density function.

The descriptive sampling is based on a deterministic and purposive selection of the input sample values. This sample value selection aims to achieve the closest fit with the represented distribution, instead of letting the sample histogram vary at random. However, because the descriptive sampling is difficult to apply to complex simulations, sample size is occasionally determined without thorough consideration by a user. Sample size may vary between runs due to input sample variability, leading to poor simulation results.

From this study, we should choose the suitable technique in reliability analysis. Also, we need to study the method to overcome the drawback of suggested techniques.

## References

[1] A. H. Ang and W. H. Tang, *Probability Concepts in Engineering Planning and Design*, John Wiley & Sons, New York, 1984.

[2] K. Breitung, *Asymptotic Approximations for Multinormal Integrals*, J. Engin. Mech., ASCE, **110** (1984), no. 3, 357-366.

[3] N. P. Buslenko, D.I. Golenko, Y.A. Shreider, I.M. Sobol and V.G. Sragowich, *The Monte Carlo Method*, Pergamon Press, New York, 1664.

[4] C. A. Cornell, *A Probability-based Structural Code*, ACI J. **66** (1969), no. 12, 974-985.
[5] B. Fiessler, H, J. Neumann and R.Rackwitz, *Quadratic Limit States in Structural Reliability*, J. Engin. Mech., ASCE, **1095** (1979), no. 4, 661-676.
[6] A. Harbitz, *An Efficient Sampling Method for Probability of Failure Calculation*, Structural Safety, **3** (1986), 109-115.
[7] A. M. Hasofer and L. C. Lind, *An Exact and Invariant First-order Reliability Format*, J. Engin. Mech., ASCE, **100** (1974), 111-121.
[8] A. D. Kiureghian, H. Z. Lin and S. J. Hwang, *Second Order Reliability Analysis Approximations*, J. Engin. Mech., ASCE, **113** (1987), no. 8, 1208-1225.
[9] R. E. Melcher, *Structural Reliability: Analysis and Prediction*, Ellis Horwood, Ltd , 1987.
[10] R. Rackwitz and B. Fiessler, *Structural Reliability under Combined Random Load Sequences*, Computers and Structures, **9** (1978), 489-494.
[11] L. Tvedt, *Distribution of Quadratic Forms in Normal Space-Application to Structural Reliability*, J. Engin. Mech., ASCE, **116** (1990), no. 6, 1183-1197.
[12] E. Saliby, *Descriptive Sampling: A Better Approach to Monte Carlo Simulation*, J. Operational Research Society. **41** (1990), no. 12, 1133-1142.
[13] M. L. Shooman, *Probability Reliability: An Engineering Approach*, McGraw-Hill, New York, 1968.

*
Department of Mathematics
Korea Air-Force Academy
Chungbuk 363-849, Republic of Korea
*E-mail*: mathyouth@afa.ac.kr

**
Department of Mathematics
Korea Air-Force Academy
Chungbuk 363-849, Republic of Korea
*E-mail*: dshwang@afa.ac.kr