INFORMATICA, 1995, Vol. 6, No. 3, 313–322

ESTIMATION OF SMALL PROBABILITIES IN STRUCTURAL SYSTEMS

Vitalija RUDZKIENĖ

Vilnius Technical University
Saulėtekio al. 11, 2054 Vilnius, Lithuania

Abstract. The exact solution of the reliability of structures under stochastic loading is generally difficult, and various approximate methods have been developed. The most popular are the linearization method, the Monte-Carlo method and its numerous variants. In this paper new modification of the Monte-Carlo method based on asymptotical expansion is examined. Results of mathematical simulation are given.

Key words: failure probability, reliability, Monte-Carlo, linearization techniques, simulation techniques.

1. Introduction. In structural systems problems of calculation of probabilities of rare occurrence appeare. Let $X = (X_1, \ldots, X_m)$ is the vector of random variables with a known distribution and $G(x)$ is the known function of m variables. This paper deals with the calculation of probabilities $p = \mathbf{P}\{G(X) \geqslant 0\}$ using random mathematical simulation when the probability p is small and random vector X is Gaussian. This problem often occurres in the design of buildings where most of the factors of structures are described by Gaussian models.

In solution of this problem main attention ussually is turned to the linearization method (see, for example, the review paper Schueller and Stix, 1987). This method is based on Taylor series expansion of the function $G(x)$ around the closest to the origin point of the domain $\{G(x) \geqslant 0\}$. In this way an analitical approximation of the probability p is given. The solution of the linearization method ussually has systemic error, depending on curvature of surface. The Monte-Carlo method has any systemic error, but practically in the estimation of small probabilities certain difficulties appeare. Therefore numerous modified variants of the Monte-Carlo method have been proposed (for example, Melchers, 1990). In present paper the method combinating Taylor series expansion

with the Monte-Carlo method is suggested. Apparently, this idea wasn't adapted widely before. This method yields generally lower error than the linearization method and, on the other hand, it demands lower random samples than usual the Monte-Carlo method. The details are given below.

Let $X_1, \ldots, X_m$ are Gaussian, standard and independent components. Then the probability of structural failure may be defined by

$$p = \int_W \varphi_m(x)dx \stackrel{\text{def}}{=} p_W, \tag{1}$$

where $\varphi_m(x) = \frac{1}{(2\pi)^{m/2}} \exp\left\{-||x||^2/2\right\}$ is the Gaussian probability density function, $||x||$ – Euclidean norm, $W\{x : G(x) \geqslant 0\}$ – the failure domain in R^m.

In this direct way finding p it is necessary to calculate m-dimensional integral. If the number of m is large it requires considerable computer time. For many years the random mathematical simulation, called the Monte-Carlo method, has been used to calculate multidimensional integrals. The samples of independent, random vector X were generated and the statistical estimate $\widehat{p}$ of probability p was calculated as the empirical probability of the event $\{X \in W\}$:

$$\widehat{p} = \frac{1}{N}\sum_{j=1}^{N} \mathbf{1}_{\{X(j)\in W\}} \stackrel{\text{def}}{=} \widehat{p}_W. \tag{2}$$

It is known that the variance of a relative error $1-\widehat{p}/p$ is equal to the quantity $\frac{1-p}{Np} \cong 1/Np$. Therefore, a large amount of generated random samples, $N >> 1/p$, is required to give a small relative error. When probability p is small, the fulfillment of this condition requires a lot of computer time, and sometimes it cannot be realized through the limited possibilities of the random number generator. For example, if $p = 0.0001$, $m = 10$, it is necessary to generate 1 million independent random values to fulfill the condition $N > 0.1/p$. It is known that random number generators can to ensure only limited quantity of independent random values.

One possible solution is based on the generation the random vectors $Y(i)$ with probability density function $f_Y(x)$ including values from the domain A, what is a part of the space R^m. If conditions

$$A \supset W, \tag{3}$$

and

$$f_Y(x) = \frac{\varphi_m(x)}{p_A}, \quad x \subset A,$$
$$p_A = \int_A \varphi_m(x)dx, \tag{4}$$

are fulfilled, then the statistical estimate $p^* = \frac{p_A}{N} \sum_{j=1}^{N} \mathbf{1}_{\{Y(j)\in W\}}$ has no bias ($\mathbf{E}p^* = p$) and the variance of this relative error is equal to $p_A(1-p/p_A)/(Np)$, that is about p_A^{-1} times less than in case (2). Therefore, p_A^{-1} times less of generated values are needed to give the same precision. This method is valid practically in the case when the random vector Y, whose distribution is completely described in the domain A, can be constructively expressed by standard random values given using the random number generator. Sometimes the requirement (3) should be refused to ensure a comfortable calculation of the random vectors $Y(1), \ldots, Y(N)$. In this case the systematic error

$$p - \mathbf{E}p^* = \int_{W\setminus A} \varphi_m(x)dx \tag{5}$$

is found. Therefore, choosing A it is necessary to try to get the systematic error lesser than the desirable precision. On the other hand, if the domain A increases, p_A grows too and more random vectors $Y(i)$ should be generated. Next, the specific algorithm, when domain A is a part of the space R^m bounded by hyperplane will be described. This idea is related to the linearization method available for calculation of failure probabilities of structures.

2. Algorithm of evaluation. Firstly we will turn the attention to one evident inequality. Denote $a = \arg\min_{x\in W} ||x||$. Therefore, a is the point of the domain W, which is the closest to the origin. Inequality

$$\varphi_m(x) \leqslant \varphi_m(a) \exp\{-||a|| \cdot (||x|| - ||a||)\}, \quad x \in W \tag{6}$$

is valid, i.e., the density function decreases exponentially when distance from the origin increases. Therefore for a such functions G, which have distinctly expressed the closest point a, it is enough for estimation of p to integrate around the point a. The linearization method is based on this idea. If a function G is

uniformly differentiable around the point a, then Taylor series expansion of the function $G(x)$ around this point a can be described by

$$G(x) = G(a) + G'(a)(x - a) + o(\|x - a\|).$$

Here $G'(a) = \left(\frac{\partial G(a)}{\partial a_1}, \ldots, \frac{\partial G(a)}{\partial a_m}\right)$, $x \cdot y$ is scalar product.

Because of $G(a) = 0$,

$$\begin{aligned} p =& \mathbf{P}\{G(X) \geqslant 0\} \cong \mathbf{P}\{G'(a) \cdot (X - a) \geqslant 0\} \\ =& \mathbf{P}\{a \cdot X \geqslant \|a\|^2\} = 1 - \Phi(\|a\|), \end{aligned} \tag{7}$$

$\Phi(\cdot)$ is standardised Gaussian distribution function. Here the property that the direction of a gradient in the closest to the origin point a of domain W coincides with direction of the point a, is used. The situation is shown in Fig. 1.

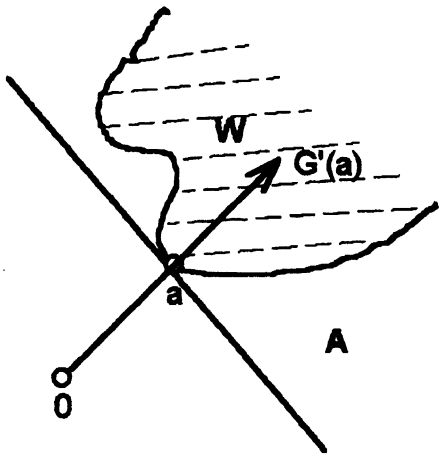

Fig. 1. Schematic sketch of two-variable system.

The tangential hyperplane L is denoted by equality

$$L = \{x : \; x \cdot a = \|a\|^2\}.$$

The estimation $\tilde{p} = 1 - \Phi(\|a\|)$ is obtained by the linearization method.

When N value is enough large, the error given by the Monte-Carlo method will be lower than the error given by the linearization method.

Let $A = \{x : x \cdot a \geqslant ||a||^2\}$. Fig. 1 illustrates the case of such function G, when $A \supset W$. In this case, when random vectors $Y(i)$ with the distribution density function (4) is generated, probability p can be estimated as precise as it is necessary, because here is no systemic error.

Now the algorithm of generation Y will be discussed. Let $m \times m$ matrix Q is orthogonal, i.e., $Q \cdot Q^T$ is unit matrix and its elements $Q_{mi} = a_i/||a||$. Let $X^* = QX$. Then X_i^* is independent standard normal variables, in addition $X - X_m^* a$ is the projection of vector X to plane L, and the vector $X_m^* a$ is the projection of X to the direction of the vector a. Let V is the random variable with distribution density function

$$f_v(v) = \begin{cases} \varphi_1(v)/p_A, & \text{if } v \geqslant ||a||, \\ 0, & \text{otherwise,} \end{cases} \tag{8}$$

where $\varphi_1(v) = \frac{1}{\sqrt{2\pi}} e^{-v^2/2}$.

The random vector Y has the distribution (4) in the case when equality $Y = Q^{-1}Y^*$ is satisfied, where $Y^* = (X_1^*, \ldots, X_{m-1}^*, V)$. The sample of $Y(t)$, $t = 1, 2, \ldots, N$ is generated in such way: standard normal components $X_1^*(t), \ldots, X_{m-1}^*(t)$ are obtained using the random number generator, and $V(t)$ is obtained from equality

$$\frac{1 - \Phi(V(t))}{p_A} = 1 - \xi(t), \quad t = 1, \ldots, N, \tag{9}$$

where $\xi(t)$ – independent random values, distributed uniformly in the interval (0,1).

In the case of such function G when expression (3) is not valid, this method of estimation, how it was mentioned, will give the systemic error, indicated in (4), that will be less than the bias $1 - \Phi(||a||) - p$ given by the linearization method. It can be decreased, defining the domain A by equality

$$A = \{x : x \cdot a \geqslant \alpha ||a||^2\}, \tag{10}$$

where $\alpha < 1$.

When the coefficient α is choosed less, the bias is lower, but variance of error is increasing, because in this case

$$p_A = 1 - \Phi(\alpha ||a||). \tag{11}$$

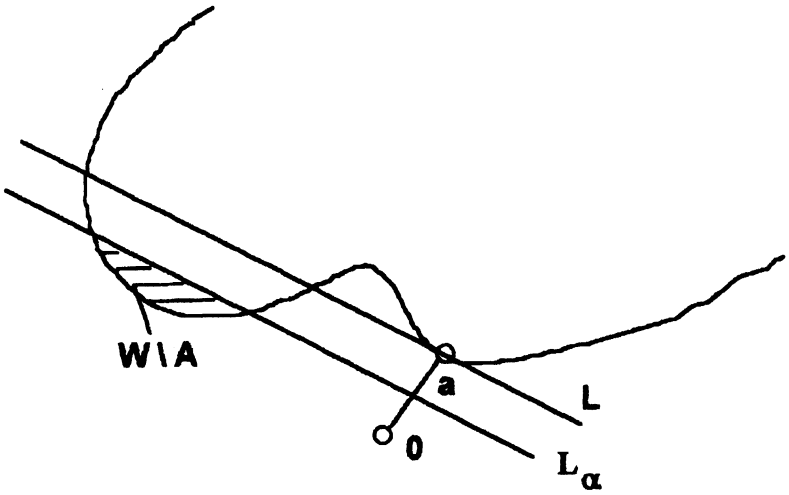

Fig. 2. Situation with bias. Here $L_\alpha = \{x : xa = \alpha||a||^2\}$.

The figure 2 illustrates the situation with a bias.

3. Generalization of method. Briefly we will outline the main ideas of Section 2 and will review possible generalizations. The method described earlier can be divided into these four parts.

3.1. The estimation with a possible bias to reduce variance. Probability p can be expanded

$$p = p_1 + p_2, \quad p_1 = \mathbf{P}\{x \in WA\},$$

where A is chosen as the domain of simple structure. Probability p_1 is estimated without bias and is accepted $\hat{p}_2 = 0$. In this way the bias is equal p_2, and estimation of p_1 is constructed with lower variance than the empirical probability $\hat{p}$ in Eq. 2.

3.2. The switch to new orthonormal coordinate system using the nearest point a of the domain W.

$$x^* = Qx,$$

where Q is orthogonal matrix, $Q_{mi} = a_i/||a||$, $i = 1, \ldots, m$. Let $d = m - 1$. Then x_m^* describes the projection of the vector x to the direction of the vector a, and $x_1, \ldots, x_d$ describe the projection of x to the tangential hyperplane of the curve $G(x) = 0$, what is drawn through the point a.

For every $A \subset \mathbb{R}_m$ we will denote $A^* = QA$.

3.3. The approximating domain is constructed based on Taylor series expansion of the function $G(x)$ around the point a.

It is very important to approximate the domain W around the point a as precise as possible, because $\max\limits_{x \in W} \varphi_m(x) = \varphi_m(a)$. If instead $G(x)$ the first member of its Taylor series would be taken

$$G_1(x) = G(a) + G'(a)(x-a) = G'(a)(x-a), \quad \text{because } G(a) = 0, \quad (12)$$

then corresponding to the W approximation $W_1 = \{x : G_1(x) \geqslant 0\}$ would be described by equality

$$W_1^* = \{x^* : x_m^* \geqslant \|a\|\}.$$

Therefore the simplest way is to describe the domain A by equality

$$A^* = \{x^* : x_m^* \geqslant \alpha\|a\|\}, \quad (13)$$

where the parameter $\alpha \leqslant 1$ is introduced for the decreasion of p_2 bias size.

3.4. The definition of estimate using conditional probabilities. If $X \in N(0, J_m)$, $X^* = QX$, and Q is orthogonal matrix, then $X^* \in N(0, J_m)$, too, where J_m is the unit matrix $m \times m$. We have $p_1 = \mathbf{P}\{X^* \in W^* \big| X^* \in A^*\} \cdot \mathbf{P}\{X^* \in A^*\}$ and $\mathbf{P}\{X^* \in A^*\} = 1 - \Phi(\alpha\|a\|)$. Denote $Z = (X_1^*, \ldots, X_d^*)$. Then the random vector Y can be defined by

$$Y^* = (Z, V), \quad (14)$$

where V is the random value, independent from Z, with distribution density

$$f_V(v) = \frac{\varphi_1(v)}{\Phi(1 - \alpha\|a\|)} \cdot \mathbf{1}_{\{v \in A^*\}}. \quad (15)$$

Then $\mathbf{P}\{X^* \in W^* \big| X^* \in A\} = \mathbf{P}\{Y^* \in W^*\} = \mathbf{P}(Y \in W)$ and

$$p_1 = \mathbf{P}\{Y \in W\} \cdot (1 - \Phi(\alpha\|a\|)). \quad (16)$$

In this way random vectors $Y^*(t)$, $t = 1, \ldots, N$ are generated and the estimate $\widehat{p}$ is defined by equality

$$\widehat{p} = \frac{1}{N} \sum_{t=1}^{N} \mathbf{1}_{\{G(Y(t)) \geqslant 0\}} \cdot (1 - \Phi(\alpha\|a\|)), \quad (17)$$

where $Y(t) = Q^{-1}Y^*(t)$.

We will outline natural generalization of the items 3 and 4. To define more precesily the approximation of the domain W, it is possible to use more members of Taylor series expansion:

$$G_1(x) = G(a) + \sum_{k=1}^{r} G^{(k)}(a)(x-a)^k/k! \tag{11a}$$

Here terms $G^{(k)}(a)(x-a)^k$ are understood as polynomial forms corresponding to the k-th derivative. From (11a) and from the definition of the domain $W_1^* = QW_1$, where $W_1 = \{x:\ G_1(x) \geqslant 0\}$ follows that

$$W_1^* = \{x^*:\ \mathbf{P}_z(x_m^*) \geqslant 0,\ z = (x_1^*, \ldots, x_d^*)\},$$

$$\mathbf{P}_z(u) = \sum_{k=0}^{u} c_k(z)u^k \text{ and coefficients } c_k(z)$$

are defined by derivatives $G^{(k)}(a)$, $k = 1, \ldots, r$, $z = (x_1^*, \ldots, x_d^*)$ and by the vector a. To increase the bias, it is possible to describe the domain A by equality

$$A^* = \{x^*:\ x_m^* \in B(z)\}, \tag{13a}$$

where $B(z) = \{u:\ \mathbf{P}_Z(u) + \Theta_Z(u) \geqslant 0\}$, $\Theta_Z(u)$ is some selected non-negative function. Then the algorithm of random samples generation and for counting of $\widehat{p}$, described in item 4, can be generalized in the following way.

Y is defined by Eq. 13, but V is optionally independent from Z. The conditional distribution density $f_V(v|Z = z)$ of random value V under the condition $Z = z$ satisfies the equality

$$f_v(v|Z = z) = \frac{\varphi_1(v)}{\Phi(B(z))} \cdot \mathbf{1}_{\{v \in B(z)\}}, \tag{15a}$$

where $\Phi(B) = \int_B \varphi_1(u)du$.

The equalities (16) and (17) change in the next equalities (16a) and (17a):

$$\begin{aligned} p_1 &= \int_{\mathbb{R}^d} \mathbf{P}\{X \in W|Z = z\} \cdot \varphi_d(z)dz \\ &= \int \mathbf{P}\{y \in W|Z = z\} \cdot \Phi(B(z)) \cdot \varphi_d(z)dz \\ &= \mathbf{E}\left[\mathbf{1}_{\{G(Y) \geqslant 0\}} \cdot \Phi(B(Z))\right], \end{aligned} \tag{16a}$$

$$\widehat{p}_1 = \frac{1}{N} \sum_{t=1}^{N} \mathbf{1}_{\{G(Y(t) \geqslant 0)\}} \cdot \Phi(B(Z(t))). \tag{17a}$$

Table 1. Estimation of the failure probabilities $\hat{p}$ and time of calculation T using different techniques

Dimension	Safety index	Theoretical	Linearization method Time T in s		By formula (4) Number of samples $N=1000$ Time T in s			Monte-Carlo method N – number of samples Time T is in hours			
n	β	$\hat{p}$	$\hat{p}$	T (s)	$\hat{p}^*$	95% confidence interval	T (s)	$N=10^4$		$N=10^5$	
								$\hat{p}$	T	$\hat{p}$	T
2	3	0.000914	0.00135	3	0.000909	[0.864; 0.954] E-3	17	0.0004	0.05	0.00037	0.33
2	2	0.0166	0.02275	3	0.0166	[0.161; 0.171] E-1	17	0.0179	0.05	0.0178	0.33
4	3	0.000403	0.00135	4	0.0004	[0.368; 0.432] E-3	32	–	0.08	–	0.67
4	2	0.00847	0.02275	4	0.00838	[0.779; 0.897] E-2	32	0.0065	0.08	0.0065	0.67
6	3	0.000169	0.00135	6	0.000177	[0.146; 0.208] E-3	51	–	0.12	0.00001	0.92
6	2	0.00406	0.02275	6	0.00425	[0.361; 0.489] E-2	51	0.003	0.12	0.00324	0.92
8	3	0.000669	0.0135	8	0.0000671	[0.501; 0.841] E-4	70	–	0.15	0.00001	1.12
8	2	0.00183	0.02275	8	0.00182	[0.152; 0.213] E-2	70	0.002	0.15	0.00026	1.12

– denotes that there were no values; the estimation of $\hat{p}$ was given from 10 samples.

Particular algorithms and results of mathematical simulation in the case when $r = 2$ in the first equality will be discussed in the next article.

4. Numerical example. In this example three methods for calculation of failure probabilities – such as usual Monte-Carlo method, the linearization method and the method, described by Eq. 4 – are compared in light of their accuracy and efficiency. To have the possibility to compare these methods, the hypersphere

$$Z = \sum_{i=1}^{m}(x_i - a)^2 - b^2 = 0$$

was taken as function $G(X)$ with $X_1, X_2, \ldots, X_m$ being standard Gaussian and $m = 2, 4, 6, 8$.

The data in Table 1 confirm the advantages of technique, defined by formula (1). Of course, these results are depended on the shape of failure surface. The precision of the solution given by the linearization method depend on number of variables of $G(x)$, and the Monte-Carlo method requires a lot of computer time.

REFERENCES

Melchers, R.E. (1990). Radial importance sampling for structural reliability. *Journal of Engineering Mechanics,* **116**(1), 189–203.

Schueller, G.I., and R. Stix (1987). A critical appraisal of methods to determine failure probabilities. *Structural Safety,* **4**, 293–309.

Received April 1995

V. Rudzkienė received Ph.D. from Kaunas Technological University in 1990. Currently she is an Associated Professor in the Applied Informatics Department of Vilnius Technical University. Her research interests lie in the fields of analysis and forecast of random processes and structures.

Mažų tikimybių vertinimas struktūrose

Vitalija RUDZKIENĖ

Skaičiuojant statybinių konstrukcijų patikimumą, naudojami įvairūs aproksimaciniai metodai. Populiariausi yra linearizacijos ir Monte-Karlo metodai bei jų modifikuoti variantai. Straipsnyje nagrinėjama nauja Monte-Karlo metodo modifikacija, kuri remiasi asimptotiniais skleidiniais. Pateikiami matematinio modeliavimo rezultatai.