

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

186,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Monte Carlo Simulation Method

Jaroslav Menčík

Additional information is available at the end of the chapter

<http://dx.doi.org/10.5772/62369>

Abstract

The Monte Carlo method studies random phenomena using numerous fictitious experiments with computer-generated random numbers. Its principle is explained and also the principle of generation of random numbers with various probability distributions. Also more complex cases, such as the response surface method and generation of correlated random quantities are explained. The use of the Monte Carlo method is illustrated on several examples.

Keywords: Probability, random numbers, Monte Carlo method, correlation, response surface method, probabilistic transformation

Random phenomena or processes can be successfully studied by the Monte Carlo method [1 - 4]. This is a probabilistic method based on performing numerous fictive experiments using random numbers. It is used in various branches of science and technology. For example, in reliability it serves for the analysis of load-carrying capacity or deformations of a construction, for the determination of time to failure, resonant frequency of a mechanical structure, or an electric circuit, or for the study of behavior of a complex transport or production system.

The Monte Carlo method is close to the engineering way of thinking. It is universal and does not need a special knowledge of probability theory. The only information it needs is the relationship between the output and input quantities,

$$y = f(x) , \text{ or } y = f(x_1, x_2, x_3, \dots) , \quad (1)$$

and the knowledge of probability distributions of the input variables. The method uses numerous repetitions of trials with computer-generated random numbers and the relevant

mathematical operations. In each "trial", the input variables x_1, x_2, \dots, x_n are assigned random values, but such that their distributions correspond to the probability distribution of each variable. With these input values, the output quantity y is calculated via Equation (1). From the results, a histogram can be constructed (Fig. 1), which corresponds to the distribution of y .

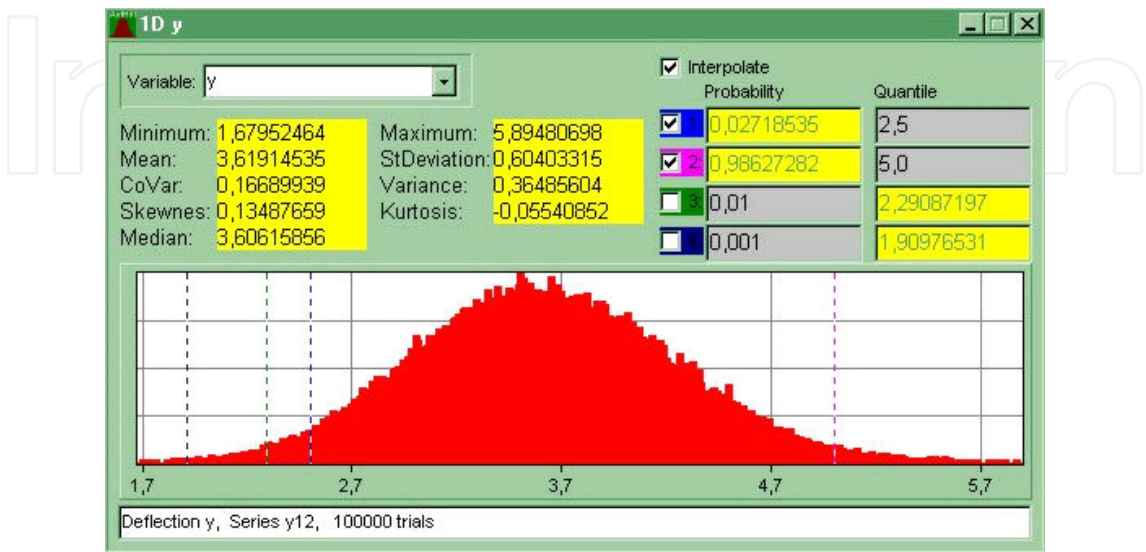


Figure 1. Histogram obtained by the Monte Carlo simulation program Ant-Hill [2].

The generated y values can be used for the determination of the average value and standard deviation, but also the probability that y will be lower or higher than a chosen value or lie within some interval. Also, the characteristic values that will be exceeded with higher probability than some allowable value (e.g. the guaranteed strength or time to failure). In a similar way, critical values can also be obtained, whose exceeding can be expected with only small probability (e.g. maximum load or deformation).

Today, various commercial computer programs exist for Monte Carlo simulations, but they can also be created. The base of such program is a generator of random numbers. Actually, they are not truly random but computer generated using a suitable deterministic algorithm. However, such algorithms are used, which generate numbers behaving nearly as if they were random.

The principle of these generators is simple. For example, the so-called congruential generator gives random numbers, distributed uniformly in the interval $(0; 1)$, in the following way. One number is chosen as the base for the series of random numbers u (e.g. $u_0 = 0.5284163$). Now, in the first step, this number is multiplied by some suitable number Q , for example 997. The product is $997 \times 0.5284163 = 526.8310511$. The first random number u_1 is then created as the part of the result, lying behind the decimal point; in our case, $u_1 = 0.8310511$. In the second step, u_1 is again multiplied by the same number Q , $997 \times 0.8310511 = 828.5579467$, and the second random number is created as the decimal part of the result (i.e. $u_2 = 0.5579467$). This procedure is repeated again and again. The formula for the random number in the j -th step is

$$u_j = \text{MOD}(Q \times u_{j-1}) ; \quad (2)$$

u_{j-1} is the random number from the previous step, and the symbol MOD (read "modulo") means the decimal part of the expression in brackets. The reader is encouraged to make several steps in this way; for a check, $u_3 = 0.2728599$. A long series of these numbers has approximately uniform distribution. Also other algorithms exist. For example, one generator of pseudorandom numbers with normal distribution uses the central limit theorem, etc. In any case, the use of "commercial" generators is strongly recommended, as they have undergone thorough statistical testing to prove that they behave as really "nearly random", and the period, after which the series of generated numbers is repeated, is very long, hundreds of millions of numbers. Generators of random numbers are usually a part of universal computer programs or languages. Even Excel can produce such numbers. Better tools are included in special packages for probabilistic analysis of reliability, such as FREET (www.freet.cz) or Ant-Hill (www.sbra-anthill.com), as well as in Matlab and other advanced software.

1. Creation of random numbers with nonstandard distributions

The commercial programs offer often-used distributions, such as uniform or normal. The random numbers, corresponding to other analytically defined distributions, can be generated via uniform distribution. The basic idea is that the distribution function F for any continuous random quantity is also random variable, distributed uniformly in the interval $(0; 1)$. Thus, if the distribution function of random quantity x is described by the expression $z = F(x)$, then the random numbers of x can be obtained from the random numbers z with uniform distribution in $(0; 1)$ using the inverse formula:

$$x = F^{-1}(z) . \quad (3)$$

Here, F^{-1} means inverse probabilistic transformation (Fig. 2). For example, the distribution function for exponential distribution is $z = F(x) = 1 - \exp(-x/x_0)$, with the parameter x_0 . The inverse transformation for this distribution is $x = -x_0 \ln(1 - z)$;

In some cases, the distribution of a random input quantity has a more complex shape and can be described by a histogram, obtained from experiments or monitoring. This histogram is then used for the construction of distribution function $F(x)$. This function can be approximated either by constant values of F in the individual subintervals of x (if the number of classes is high) or by interpolation within each class, such as

$$F(x) = F_i + \frac{F_{i+1} - F_i}{x_{i+1} - x_i} (x - x_i), \quad x = x_i + \frac{F(x) - F_i}{F_{i+1} - F_i} (x_{i+1} - x_i) \quad (4)$$

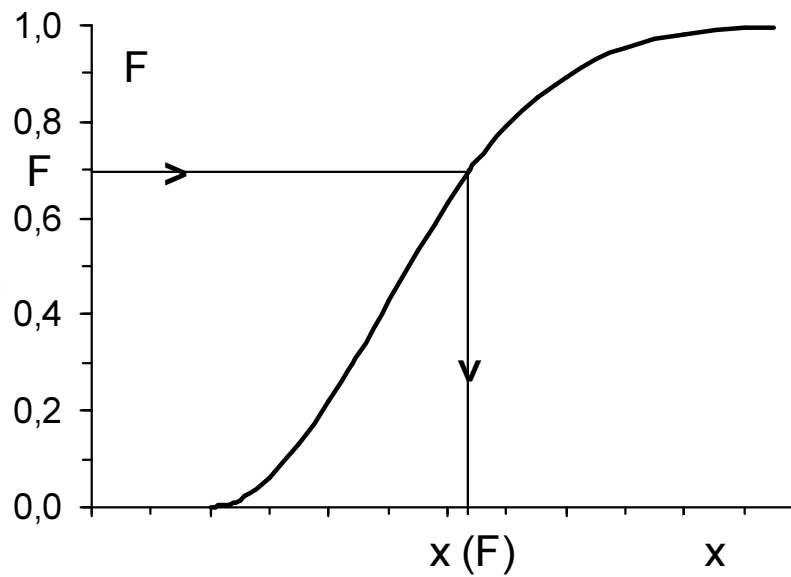


Figure 2. Generation of random numbers x by inverse probabilistic transformation.

$i = 1, 2, \dots, n$ denotes the interval. The formula on the right gives x corresponding to the probability F . The F values are then generated as random numbers with uniform distribution.

A typical feature of the Monte Carlo method is that the characteristic values (average, quantiles, probabilities corresponding to certain values of y , etc.), obtained as a result of n trials, are never the same as the results obtained in any other set of simulations. The results are thus only approximate and are closer to the actual values for a higher number of trials. The number of simulation steps n , needed for attaining some accuracy of results, is given approximately by the formula:

$$n = u_{\alpha/2}^2 (1 - P) / (P \delta^2) ; \quad (5)$$

P is the expected (estimated) probability of the investigated phenomenon, δ is the allowed relative error in the determination of P , $u_{\alpha/2}$ is the $\alpha/2$ -critical value of standard normal variable, and α is the probability that the actual value of P will lie outside the interval $P \pm \delta$. The necessary number of simulations thus grows significantly with decreasing probability. For example, if the assumed probability P is 0.01 and the allowed relative error δ is 10 % and confidence level $\alpha = 5\%$ (with $u_{\alpha/2} = 1.96$), then $\approx 40,000$ simulation trials are necessary. For $P = 0.0001$, it is as much as 4,000,000 trials, etc. [Note: Equation (5) is based on the fact that the number of outcomes of an event of probability P in n repetitions has binomial distribution, and this distribution can be approximated by normal distribution for high n .]

The characteristic features of the Monte Carlo method are illustrated on several examples at the end of this chapter. The reader is encouraged to work them out on a PC.

2. More complex cases — Response Surface Method

The direct use of the Monte Carlo method is suitable for simple relationships $y = f(x_1, x_2, \dots)$. Often, the response must be obtained by numerical solution (e.g. the finite-element method). If one trial lasts minutes or more, the thousands of simulations would consume too much time. In these cases, more effective is the combination of Monte Carlo with the response surface method (RSM). The principle is that the response is first calculated only for selected values of input variables, the results are fitted by a simple regression function (response surface; Fig. 3), and the Monte Carlo trials are done with this function.

The relationship between the output quantity y (deformation, load-carrying capacity, amplitude of vibrations, etc.) and the input variables can often be fitted by a polynomial function:

$$y = a_0 + \sum a_i x_i + \sum b_i x_i^2 + \dots + \sum c_{ij} x_i x_j + \dots \quad (6)$$

This approximation is suitable if the actual relationship between input and output has a similar character (e.g. $y \sim x^3$) or if the output quantity changes in the considered interval only little. If it differs significantly from a polynomial (e.g. $y \sim 1/x^3$ or $y \sim x^{1/2}$), expression (6) cannot give a good approximation in a wider interval. There are several ways for improvement, the starting point being a visual judgment of the character of the response. Linear or polynomial function may be used for the approximation of other relationships if suitable transformations are made. For example, the relationship $y = a/x^3$ can be expressed as $y = az$ by introducing a new variable $z = 1/x^3$; the relationship $y = ax_1/x_2^2$ can be converted to multiple linear regression $Y = A_0 + A_1 X_1 + A_2 X_2$ using logarithmic transformations, etc. Solvers in universal software (Excel, Mathcad, or Matlab) can find regression coefficients directly, without transformations.

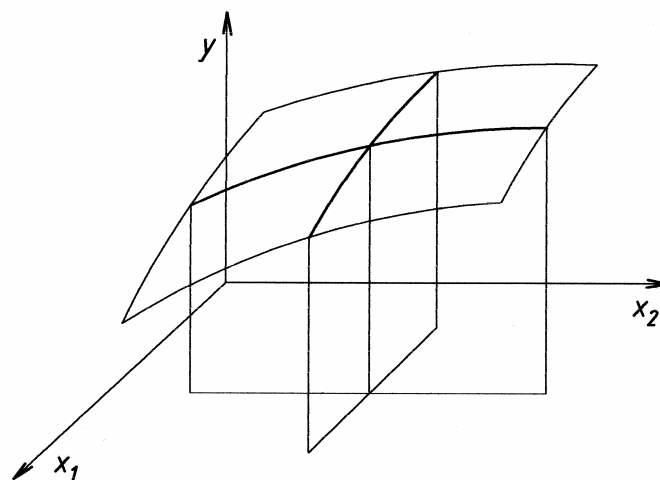


Figure 3. Response surface for two independent variables (a schematic, with cuts $x_1 = \text{const}$, $x_2 = \text{const}$).

The fit of response function can sometimes be improved by dividing the definition interval of some input quantities into subintervals and using different regression functions for each. This may be substantiated by the physical character of the problem. For example, the elastoplastic deflection of a beam obeys another law than purely elastic deformations.

The quality of the fit can be evaluated by means of residual standard deviation s_{res} , which characterizes the scatter of the measured values around the regression function. Also the maximum difference between the individual "accurate" values and those on the response surface can serve as a criterion. With good response surface, the individual differences are randomly positive and negative. Larger regions with differences of the same sign indicate that the chosen function does not correspond well to the character of the actual response.

3. Application of the Monte Carlo method for correlated quantities

The application of Monte Carlo simulations in problems with several input variables is simple if the individual input quantities are mutually independent (e.g. Young's modulus and the cross-section area of a beam). Sometimes, however, a relation between them exists; (e.g. between mass density and Young's modulus of concrete). In such case, one speaks about statistical dependence or correlation. A special case is the so-called autocorrelation, when the value of a random quantity at some point is related partly to the values at neighboring points or in preceding times. Examples are the properties of concrete or of soil at foundations or the temperature of a building structure: it varies during a day or from a day to day, but depends partly also on the season in the year.

The omission of correlations in the simulations can lead to errors. For example, a very low value of elastic modulus of concrete could be generated simultaneously with a very high value of strength, but this does not correspond to reality. If correlations are respected, the calculations reflect the reality better and the conclusions or predictions are more accurate, with smaller scatter. Sometimes, also, a quantity needed for the analysis is unavailable, but can be replaced by a correlated quantity. For example, if the direct measurement of the tensile strength of components in an existing massive steel structure is impossible, the information from hardness tests can sometimes be adapted.

The tightness of the relationship of two quantities is characterized by the correlation coefficient r , defined as

$$r_{ij} = cov(x_i, x_j) / (s_i s_j), \quad (7)$$

where $cov(x_i, x_j)$ is the covariance of x_i and x_j , and s_i and s_j are their standard deviations. The correlation coefficient r is nondimensional, ranging from -1 to $+1$. For $r = 0$, no mutual relationship exists, whereas $r = +1$ or -1 mean deterministic (functional) relationship. For $r > 0$, the x_j values grow with the growth of x_i and decrease for $r < 0$. (Note: The correlation coefficient is also equal to the square root of the coefficient of determination r^2 , given by programs for

curve fitting, available also in Excel.) Three examples with various values of r are shown in Figure 4.

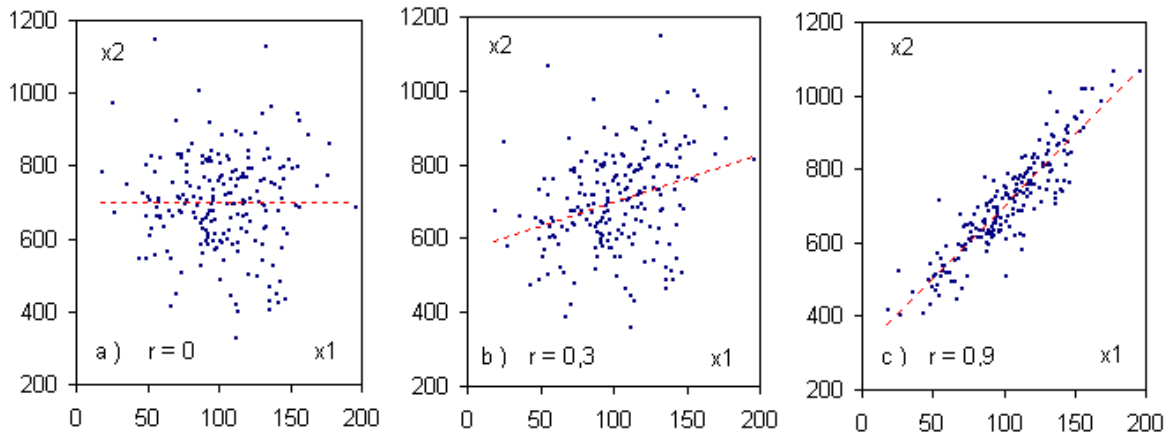


Figure 4. Two correlated quantities x_1 and x_2 with the same mean values ($\mu_1 = 100$ and $\mu_2 = 700$) and standard deviations ($\sigma_1 = 30$ and $\sigma_2 = 150$) and various correlation coefficients r [5].

If two correlated random quantities x_1 and x_2 should be generated, and if the regression function $x_{2,\text{reg}} = f(x_1)$ is known, as well as the coefficient of determination r^2 of this approximation, the following procedure may be used. First, the random value of x_1 is generated. Then, the corresponding value of x_2 is generated as [5, 6]

$$x_2 = f(x_1) + \Delta x_2 = f(x_1) + u s_{2,\text{res}} = f(x_1) + u s_2 (1 - r^2), \quad (8)$$

where $s_{2,\text{res}}$ is the residual standard deviation of quantity x_2 around the regression function f , and u is the quantile of standard normal distribution (provided that the distribution of individual values x_2 around f is normal). The right-hand part of Equation (8) uses the fact that the residual deviation $s_{2,\text{res}}$ of x_2 can be expressed by means of the standard deviation s_2 of x_2 and the coefficient of determination r^2 pertaining to the regression function $x_2 = f(x_1)$.

This approach may be used for linear, as well as nonlinear relationships between x_1 and x_2 . With some modification, it may also be used for multiple regression [6].

More information on the Monte Carlo method, especially on its use in the assessment of reliability of structures, including load-carrying capacity and lifetime, can be found in books [1 – 3] and proceedings of conferences [4], which contain many practical examples.

Example 1

Generate (e.g. using Excel) 500 random numbers with normal distribution with the mean $\mu = 5$ and standard deviation $\sigma = 1$. Calculate the sample average m and standard deviation s and compare them with μ and σ . Determine also the 5% quantile and the probability that x will be

larger than 8.0 and compare them with the exact values $x_{0,05} = 3.35515$, and $P(x > 8.0) = 0.99865$. Repeat the procedure and look at the new results. You can also make similar simulations with a lower number of generated values (e.g. $n = 50$) and also with a higher number ($n = 5000$).

Remark: Excel was mentioned here because it is ubiquitous and its use is easy. Everybody can thus try to solve such problems. The necessary routines Descriptive statistics, Histograms, Generation of random numbers, and Solver are installed in every Excel. However, they are not always directly available. If the command Data analysis does not appear on your screen after the command "Data", it must be activated. The procedure is as follows. Click on the button File, then on Possibilities (in this menu), then Add-Ins, then Analytical Tools (and Solver), and, finally, OK. After next pressing the command Data, the buttons Data Analysis and Solver appear in the upper part of the screen. It would be a pity not to use such powerful tools !

Example 2

Generate 10,000 random numbers (x) with uniform distribution in the interval (0; 1). Calculate the mean value and standard deviation (accurate values are $\mu = 0.5$ and $\sigma = 12^{-1/2}$), plot the histogram, and check if it corresponds to uniform distribution. Now, generate the second series of 10,000 numbers (Y) with the same parameters. Make the sums of two numbers with the same subscript j . Calculate now the parameters and create the histogram of the resultant distribution. (The accurate value of the mean is $\mu = \mu_1 + \mu_2 = 0.5 + 0.5 = 1$, and the distribution is triangular.) If you would make (in similar way) the sum of three and more random quantities with uniform distribution, you will see that the resultant distribution resembles more and more normal distribution, in agreement with the central limit theorem.

Example 3

Determine (by the Monte Carlo method) the mean time to failure from Example 5 in Chapter 5 (four elements in a series, each of exponential distribution with $\lambda_1 = 8 \times 10^{-6} \text{ h}^{-1}$, $\lambda_2 = 6 \times 10^{-6} \text{ h}^{-1}$, $\lambda_3 = 9 \times 10^{-6} \text{ h}^{-1}$, and $\lambda_4 = 2 \times 10^{-5} \text{ h}^{-1}$).

Solution.

The problem was solved using Excel. First, the mean times to failure of individual components were calculated (in standard way) as $m_{t,i} = 1/\lambda_i$ (i.e. $m_{t,1} = 125,000 \text{ h}$, $m_{t,2} = 166,667 \text{ h}$, $m_{t,3} = 111,111 \text{ h}$, and $m_{t,4} = 50,000 \text{ h}$). In the second step, four series of random values of time to failure of individual components were generated via inverse formula $t = -m_t \ln(1 - F)$ from random numbers F with uniform distribution in the interval (0; 1) and for the above parameters $m_{t,1}, \dots, m_{t,4}$. Each series had 1000 numbers. Then, in each simulation trial, the minimum of the four random numbers, corresponding to the times to failure of individual elements, was found, because the failure of the system of several components in a series occurs as the first element fails. In this way, the mean time to failure of the series system was $m_t = 23,376.3 \text{ h}$, which is very near the theoretical value 23,256 h. The standard deviation was $s_t = 24,137.7 \text{ h}$, again near the mean value, as typical for exponential distribution. (A histogram will confirm it.)

The mean values of the simulation series for the individual components are as follows (the numbers in brackets express the standard deviation): No. 1: 120,073 (126,158), No. 2: 161,217

(162,127), No. 3: 104,902 (103,727), and No. 4: 52,001 (53,723), everything in hours. One can see that the parameters of the generated variables are all near the parent parameters. With higher numbers of simulation trials, the differences would be even smaller.

Remark: Systems with parallel arrangement of components can be solved in a similar way, but instead of searching for the minimum of the times to failure in each trial, now the maximum will be sought, because the parallel system fails only after the failure of the component with the longest time to failure.

Author details

Jaroslav Menčík

Address all correspondence to: jaroslav.mencik@upce.cz

Department of Mechanics, Materials and Machine Parts, Jan Perner Transport Faculty,
University of Pardubice, Czech Republic

References

- [1] Hammersley J M, Handscomb D. Monte Carlo Methods. New York: John Wiley; 1964. 184 p.
- [2] Marek P, Brozetti J, Guštar M, Tikalsky P, editors. Probabilistic Assessment of Structures using Monte Carlo Simulation. Prague: ITAM CAS CR; 2003. 471 p.
- [3] Marek P, Guštar M, Anagnos T: Simulation-based reliability assessment for structural engineers. Boca Raton: CRC Press; 1996. 384 p.
- [4] Marek P (editor): Reliability of constructions (in Czech: Spolehlivost konstrukcí). Proceedings of conferences. Ostrava: Dům techniky Ostrava; 2000–2012. ISBN 80-0201708-0, 800201489-8, 80-0202007-3, 80-0286246-33-8, ISBN 80-0201551-7.
- [5] Menčík J. Simulation assessment of reliability with correlated quantities. (In Czech: Simulační posuzování spolehlivosti při korelovaných veličinách.) In: Spolehlivost konstrukcí (Reliability of structures), 23.-24. 4. 2003; Ostrava: Dům techniky; 2003. 151 – 156.
- [6] Čačko J, Bílý M, Bukovec J. Measurement, evaluation and simulation of random processes. (in Slovak: Meranie, vyhodnocovanie a simulácia náhodných procesov.) Bratislava: VEDA; 1984. 210 p.

