MONTE CARLO METHODS

Melchers R.E., Structural reliability analysis and prediction. John Wiley & Sons 2001

The Monte Carlo method has for long been recognized as the most exact method for all the calculations that require the knowledge of the probability distribution of response of uncertain systems to uncertain inputs.
A general idea of the Monte Carlo method can be summarized as follows. Suppose that the following integral is evaluated

$$I = \int_D g(\mathbf{x})\, dx \tag{5.1}$$

where D is a region in high-dimensional space and $g(\mathbf{x})$ is the target function of interest. If independent and identically distributed random samples $\mathbf{x}_1, \ldots, \mathbf{x}_m$, uniformly simulated from D, an approximation of I can be obtained as

$$\hat{I}_m = \frac{1}{m}\left[g(\mathbf{x}_1) + \ldots + g(\mathbf{x}_m) \right] \tag{5.2}$$

According to the law of large numbers the average of many independent random variables with common mean and finite variances tends to stabilize at their common mean

$$\lim_{m\to\infty} \hat{I}_m = I, \quad \text{with probability } 1 \tag{5.3}$$

Its convergence rate can be assessed by the central limit theorem

$$\sqrt{m}\left(\hat{I}_m - I\right) \to N\left(0, \sigma^2\right) \tag{5.4}$$

where $\sigma^2 = \mathrm{var}[g(\mathbf{x})]$.

Hence, the error term of the Monte Carlo approximation is $O(m^{-1/2})$, regardless of the dimensionality of $\mathbf{x}$.

In the case of structural reliability analysis, this means, that each random variable vector $\mathbf{x}_i$ is randomly generated to obtain sample value $\hat{\mathbf{x}}_i$, and then the limit state function $G(\hat{\mathbf{x}}_i) = 0$ is checked.

If the limit state is violated, i.e. $G(\hat{\mathbf{x}}_i) \le 0$, the structure or structural element has “failed”.

The experiment is repeated many times.

If N trials are conducted, the probability of failure is given approximately by

$$p_f = \frac{n\left(G\left(\hat{\mathbf{x}}_i\right) \le 0\right)}{N} \tag{5.5}$$

where $n(G(\hat{\mathbf{x}}_i) \le 0)$ denotes the number of trials n for which $G(\hat{\mathbf{x}}_i) \le 0$. The number N of trials is related to the accuracy for p_f estimation.

To apply the Monte Carlo techniques to structural reliability it is necessary (Melchers 1999):

1) to develop simulation technique for numerical sampling of the basic variables $\hat{\mathbf{x}}_i$,

2) to consider the effect of the complexity of calculating the limit state function $G(\hat{\mathbf{x}}_i)$ and the number of basic variables on the simulation techniques used,

3) to determine the amount of sampling required to obtain a reasonable estimate of the structure probability of failure p_f.

Direct Monte Carlo method

The *direct sampling* or *Simple Random Sampling* is the simplest Monte Carlo approach in solving reliability problems. It can be graphically presented as so-called *ant-hill* (see Fig 1). It does not apply any reduction method to the generated set of variates, which allows for the statistical description of the structural behaviour without scarifying the description quality. Thus, this method can be fast enough for the reliability analysis of structures with a reduced number of degree of freedom but it is too costly for any large structure analysis.

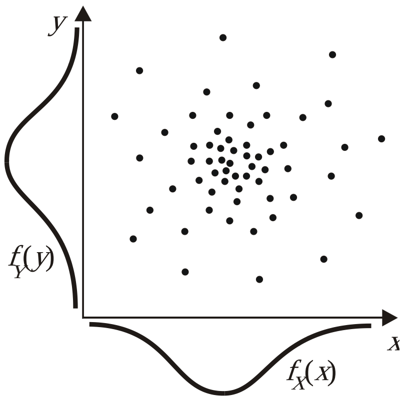


Fig. 1. Simple random sampling

In this case the probability of structure failure

$$p_f = P\left[G\left(\mathbf{X}\right) \le 0\right] = \int \ldots \int_{G(\mathbf{X}) \le 0} f_{\mathbf{X}}\left(\mathbf{x}\right) dx \tag{5.6}$$

may be expressed as (Melchers 1999)

$$p_f = J = \int \ldots \int I\left[G\left(\mathbf{X}\right) \le 0\right] f_{\mathbf{X}}\left(\mathbf{x}\right) d\mathbf{x} \tag{5.7}$$

where $I[\cdot]$ is an indicator function which equals 1 if $[\cdot]$ is "true" and 0 if $[\cdot]$ is "false". Thus, the indicator function identifies the integration domain.
The unbiased estimator of the expected value J and the estimator of standard deviation can be calculated as follows:

$$p_f \approx J_1 = \frac{1}{N} \sum_{i=1}^{N} I\left[G\left(\hat{\mathbf{x}}_i\right) \le 0\right] \tag{5.8}$$

$$\sigma_{J_1}^2 = \sum_{i=1}^{N} \frac{1}{N^2} \operatorname{var}\left[I\left(G \le 0\right)\right] = \frac{\sigma_{I(G \le 0)}^2}{N} \tag{5.9}$$

where $\hat{\mathbf{x}}_i$ represents the i-th vector of random observations from $f_{\mathbf{X}}(\cdot)$.

The standard deviation of J_1 and hence of the Monte Carlo estimate p_f (5.8) varies inversely with $N^{1/2}$ (see also Eq. (5.4)) These observations are important in determining the number of simulations required for a particular level of confidence.

On the basis of the central limit theorem, the following confidence statement can be made concerning the number of J_1 trails in which failure are possible (see Melchers 1999)

$$P\left(-k\sigma < J_1 - \mu < +k\sigma\right) = C \tag{5.10}$$

where μ is the expected value of J_1 given by Eq. (5.8) and σ is standard deviation expressed by (5.9).

The number N of simulations for a given confidence level C in the failure probability p_f can also be obtained from (see Melchers 1999)

$$N > \frac{-\ln\left(1-C\right)}{p_f} \tag{5.11}$$

Using Eq. (5.11) for a 95% confidence level and $p_f = 10^{-3}$ the required number of simulations is more than 3000.

It is not convenient to apply the above theoretical rules to the accuracy analysis in any particular Monte Carlo calculations. According to Melchers (1999) a useful tool for this purpose is to plot progressive results of the estimate of p_f and variance $\sigma^2_{J_1}$ (Eqs. (5.8) and (5.9) respectively). Such plots (see Fig. 2) will show that these measures decline when the number of samples rises and that a degree of stability is reached at a sufficiently high number of samples.

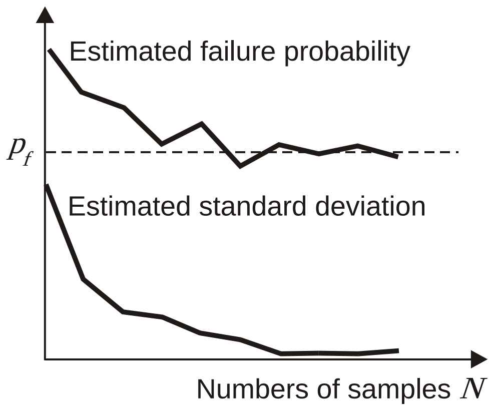


Fig. 2. Convergence of probability estimate with increasing sample size

The results may also be represented as a cumulative distribution function $F_G(g)$ (see Fig 3). The estimate of p_f in Eq. (5.8) may be improved by fitting an appropriate distribution function through the points for which $G(\cdot) \leq 0$, i.e. the left-hand tail in Fig. 3 (Melchers, 1999).

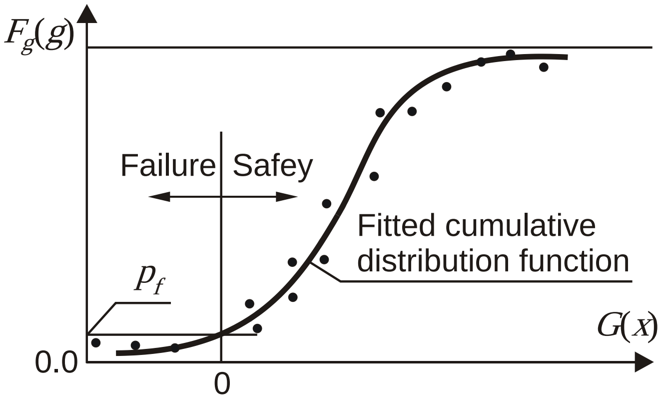


Fig. 3. Use of fitted cumulative distribution function to estimate D_1

Stratified sampling and Latin Hypercube Sampling

The *Stratified Sampling* and *Latin Hypercube Sampling* techniques have been proposed to reduce the Monte Carlo calculation.

In the case of the *Stratified Sampling* method the whole space of the variable is divided into subsets of equal probability. The acquired data are generated from each subset and an analysis is performed with corresponding sets of points (Fig. 4). A sample from inside the subset is taken either from the middle or randomly.

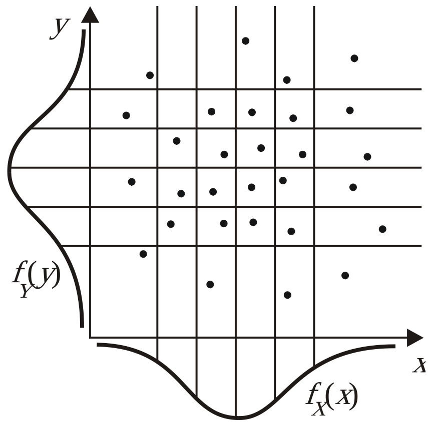


Fig. 4. Stratified Sampling

It should be stressed that the Direct Monte Carlo and the Stratified Sampling Method can also be applied to those cases in which the limit state function $G(\mathbf{X})$ is not known.

The *Latin Hypercube Sampling* method combines at random each subset number from each random variable with other subset numbers of the remaining variables only once (see Fig. 5).

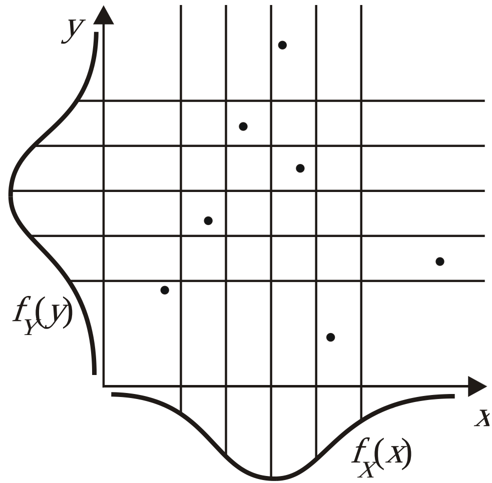


Fig. 5. Latin Hypercube Sampling

Importance sampling and search techniques

The integral (5.6) can be written using the indicator function $I[\cdot]$ as (Melchers 1999)

$$J = \int \ldots \int I\left[G\left(\mathbf{X}\right) \leq 0\right] \frac{f_{\mathbf{X}}\left(\mathbf{x}\right)}{h_{\mathbf{V}}\left(\mathbf{x}\right)} h_{\mathbf{V}}\left(\mathbf{x}\right) d\mathbf{x} \tag{5.12}$$

where $h_{\mathbf{V}}(\mathbf{x})$ is termed the “importance-sampling” probability density function.

An unbiased estimate of J is given by (cf. (5.8))

$$p_f \approx J_2 = \frac{1}{N} \sum_{i=1}^{N} \left\{ I\left[G\left(\hat{\mathbf{v}}_i\right) \leq 0\right] \frac{f_{\mathbf{X}}\left(\hat{\mathbf{v}}_i\right)}{h_{\mathbf{V}}\left(\hat{\mathbf{v}}_i\right)} \right\} \tag{5.13}$$

where $\hat{\mathbf{v}}_i$ is a vector of sample values taken from the importance sampling function $h_{\mathbf{V}}(\mathbf{v})$.

For a given level of confidence, far fewer sample points of $h_{\mathbf{V}}(\mathbf{v})$ are required than in the direct Monte Carlo method with $f_{\mathbf{X}}(\mathbf{x})$ as sampling distribution.

The derivation of optimal $h_{\mathbf{V}}(\mathbf{v})$ functions is difficult and they are often selected on a priori grounds.

Sometimes it is possible to estimate the point x^*, known as the point of "maximum likelihood" or the "design point", with $f_{\mathbf{X}}(\mathbf{x})$ having the largest influence on the limit state function (see Fig. 6).

The point x^* may be found by a direct application of the numerical maximization techniques or the search algorithms. Once x^* is identified, the most common approach of choosing $h_{\mathbf{V}}(\mathbf{v})$ is to use the distribution $f_{\mathbf{X}}(\mathbf{x})$ shifted so that its mean is at x^*.

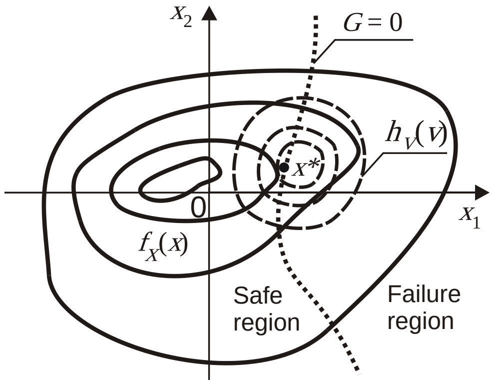


Fig. 6. Importance sampling function $I_{\mathbf{V}}[\ ]$ in x space.

Adaptive sampling techniques apply modification of $h_{\mathbf{V}}(\mathbf{v})$, depending on the information being obtained from the search process (see Melchers 1999).

First the initial location of $h_{\mathbf{V}}(\mathbf{v})$, described by a mean vector and a covariance matrix is assumed.

A limited amount of sampling is then carried out. The samples which fall into the failure domain are used to relocate and change the form of $h_{\mathbf{V}}(\mathbf{v})$.

In general, it requires good physical understanding of the problem being solved.

The importance sampling method makes allowance for the estimation of the sensitivity of failure probability to changes in random variables.

Generally, if the effect of changing one or more variables on the failure probability is required to be evaluated, two Monte Carlo calculations, with or without a change should be performed.

Such an analysis is unlikely to be very helpful. If the limit state function is analytical, then the differentials $\partial G / \partial X_i$ will give the sensitivity of $G(\mathbf{X})$ to a change in X_i.

In the case of the importance sampling the probability estimate for the modified problem with a changed random variable x_i is given by (Melchers, 1999)

$$p_f + \Delta p_i = \int_D f_{\mathbf{X}+\Delta X_i}(\mathbf{x})\,d\mathbf{x} \approx \frac{1}{N}\sum_{j=1}^{N} I\left[\hat{\mathbf{x}}_j\right]\frac{f_{\mathbf{X}+\Delta X_i}\left(\hat{\mathbf{x}}_j\right)}{h_{\mathbf{X}}\left(\hat{\mathbf{x}}_j\right)} \tag{5.14}$$

and the sensitivity can be estimated as follows

$$S_i = \left[\left(p_f + \Delta p_i\right) - p_i\right] \big/ \Delta x_i \tag{5.15}$$

Simulation methods – the Monte Carlo method

Szczepan Woliński, Krystyna Wróbel
Reliability of engineering structures (in Polish)
Publishers of Rzeszów University of Technology 2001

Structural reliability assessment by means of the Monte Carlo simulation method (MCS):

- generate a sequence of random numbers / random fields due to every random involved in the reliability analysis,
- state a reliability measure, being an outcome of physical experiments,
- classify the results to the zones of reliable or failed states,
- performing a sufficiently great population of realizations (N) compute the ratio of failed cases N_I to the total population N,
- the ratio $q = N_I / N$ is a structural unreliability measure, reliability $Q = 1 - q$).

Sufficient accuracy in reliability estimation requires a high population $N \geq (25 \div 100) q^{-1}$, e.g. 10000 ($q$ – anticipated failure probability)
Satisfactory results for practical problems may be achieved by means of a relatively low number of realizations, e.g. 10-30.

The Monte Carlo method is convenient for any structures, including those of nonlinear vector of structural performance, producing gross errors while linearized statistically.

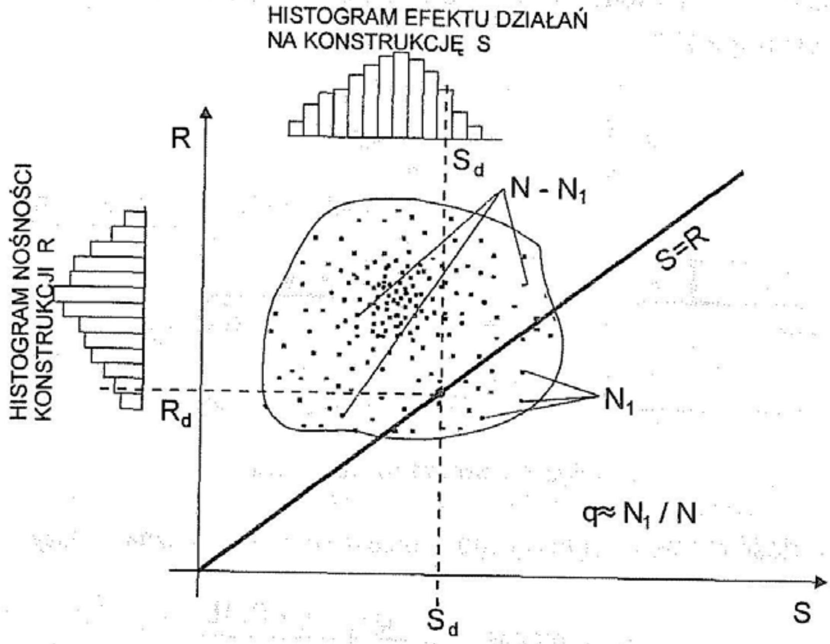
The advanced Monte Carlo techniques require a more comprehensive data on structural performance, including the failure regions. These methods improve significantly the result convergence. Examples are: importance sampling, directional, stratified, adaptive sampling, Latin Hypercube sampling.

The MCS was applied in a straightforward way to assess the structural reliability within the TERECO project (coordinator: prof. Pavel Marek, software: Milan Guštar - Prague, Polish participants: prof. Szczepan Woliński, prof. Ryszard Kowalczyk)

The assumptions prior to software making:

- all random variables: basic (material and geometric parameters, imperfections, structural actions), compound (multidimensional, correlated, action effects, limit parameters, etc.) and the results (reliability / safety, durability, serviceability measures, economic measures) are represented by bar histograms (Fig. 4.7),
- reliability check is the comparison of computed probability of failure / exceeding the limit values (ratio of the failed cases to the whole population) with the allowable probabilities.

The concept is illustrated in Fig. 4.7. (fundamental reliability case – two random variables: load effect *S* and resistance *R*).



Rys. 4.7. Symulacja Monte Carlo w dwuwymiarowej przestrzeni zmiennych losowych S i R

Fig. 4.7. Monte Carlo simulation in a two-dimensional space of S and R variables

The MCS package created by P. Marek and M. Guštar includes five procedures:

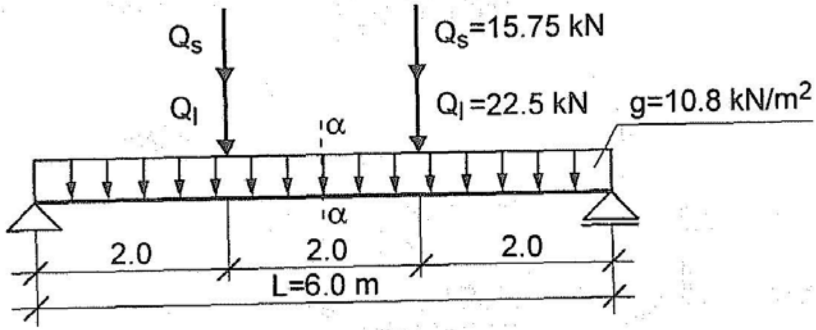
- LoadCom – load combination analysis, design loads due to various standards,
- M-Star – solver for a large family of algebraic, logarithmic, exponential and trigonometric equations composed of maximum 30 random variables expressed by bar histograms. It is a tool to analyze problems of load-carrying capacity assessment of elements and structures, load effect combinations (e.g. cross-sectional forces), failure probability, damage accumulation and serviceability criteria,
- AntHill – two- and multi-dimensional random variable analysis, e.g. reliability assessment, cross-sectional forces due to complex actions,

- DamAc – the impact of load duration to fatigue resistance of structures and reliability assessment incorporating rheological material phenomena,

- ResCom – structural load analysis.

Example

Apply the Monte Carlo method to compute the bending moment at a critical section $\alpha - \alpha$ of a beam, shown in Fig. 4.8 whose probability of exceedance is $q = 10^{-4}$.



Rys. 4.8. Schemat statyczny belki

Fig. 4.8. Static model of a beam

The maximum, midspan bending moment is given:

$$M_{sd} = \max M_{\alpha-\alpha} = \frac{ql^2}{8} + \frac{(Q_I + Q_S)L}{3} \tag{16}$$

Loads and the beam length are random variables –products of their nominal (design) values and the random factor, derived experimentally and represented by bar histograms, shown in Fig. 4.9.

– dead load: g = 10.8*Gvar , Gvar = Dead1

– long-lasting (sustained) live load: $Q_I = 22.5 * Q_{I,var}$

$Q_{I,var}$ = Long1

– short-lasting (transient) live load: $Q_S = 15.75 * Q_{S,var}$

$Q_{S,var}$ = Short1

The span: $L = 6.0 * L_{var}$ L_{var} = U1-05

The M-Star software was used for solution.

$M_{sd} = g*L\text{^}2/8+(Q_I + Q_S) *L/3$

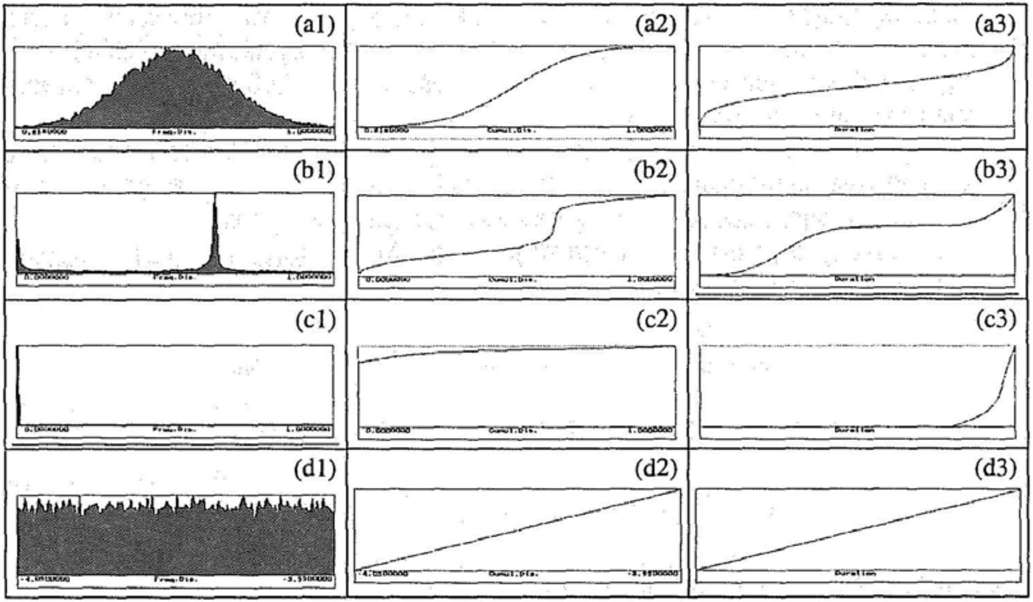
$g = 10.8*G_{var}$

$Q_I = 22.5 * Q_{I,var}$

$Q_{Sh} = 15.75*O_{s,var}$

$L = 6.0*L_{var}$

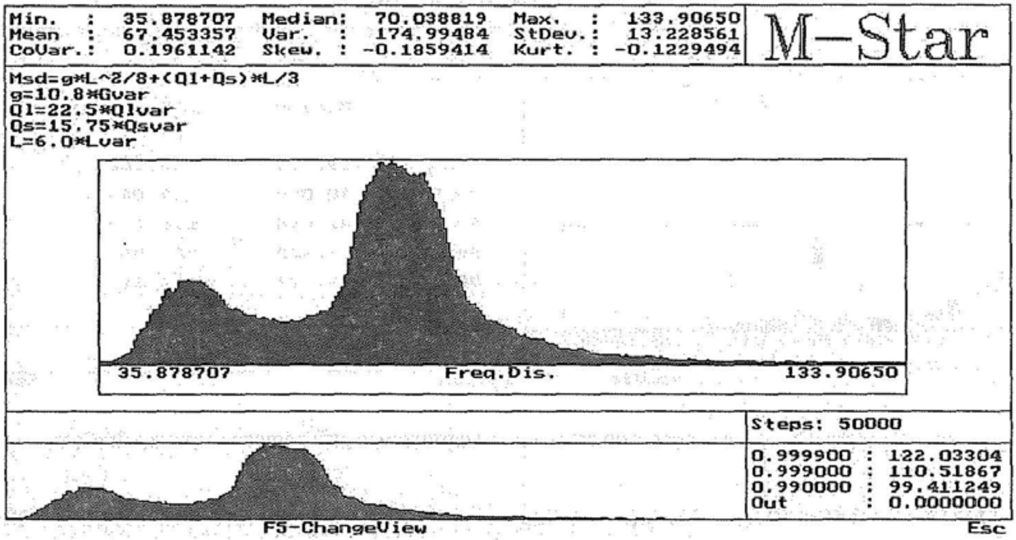
The basic variables were assumed according to Fig. 4.9.



Rys. 4.9. Histogramy słupkowe (a1, b1, c1, d1), dystrybuanty empiryczne (a2, b2, c2, d2) i krzywe rozkładu w czasie (a3, b3, c3, d3) bazowych zmiennych losowych Dead 1 (a), Long 1 (b), Short 1 (c), U1-05 (d)

Fig. 4.9. Bar histograms (a1, b1, c1, d1), empirical CDFs (a1, b1, c1, d1), time duration curves (a3, b3, c3, d3) of basic random variables: Dead 1(a). Long 1 (b), Short 1 (c), U1-05 (d).

The results of 50 000 realizations are shown in Fig. 4.10.



Rys. 4.10. Wydruk wyników obliczeń z przykładu 4.4 wykonanych za pomocą programu M-Star

Fig. 4.10. The M-Star output of an example

Histogram in Fig. 4.10 represents a random variable M_{sd} – bending moment at the critical section of a beam, due to loads g, Q_I i Q_s

being uncorrelated random variables whose distributions are known, considering randomly variable beam length L. The result may be also a CDF or a time duration curve of M_{sd}.
The value M_{sd} = 122.03 kNm may be exceeded with the probability $q = 10^{-4}$.

The second solution variant uses the ResCom software. Nominal (design) load values were taken from the previous case.
The main difference is a deterministic beam length.

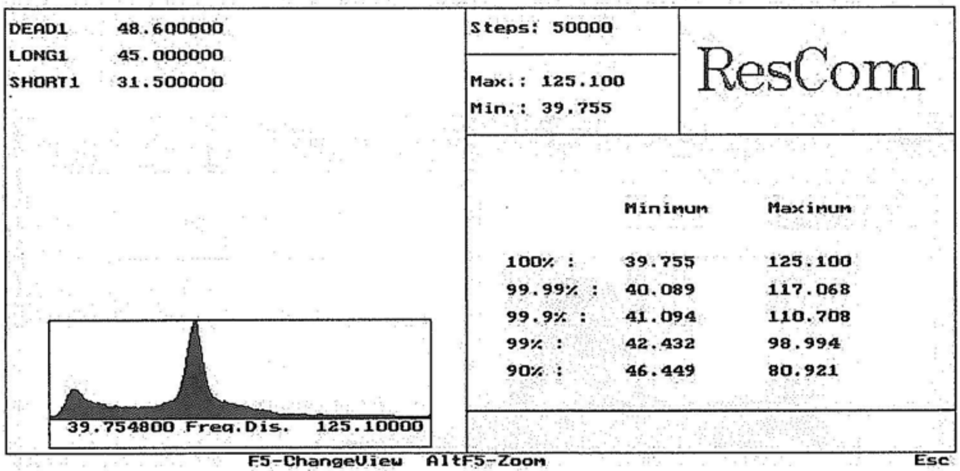
Critical bending moment is the sum
$M_{sd} = M_{sd}(g) + \mathrm{M}_{\mathrm{sd}}(O_I) + M_{sd}(O_s)$, the components are:

$$M_{sd}(g) = g \times L^2 / 8 = 10.8 * G_{var} * 6.0^2 / 8 = 48.60 * G_{var}$$

$$M_{sd}(Q_l) = Q_l \times L / 3 = 22.5 * Q_{l,var} * 6.0 / 3 = 45.0 * Q_{l,var}$$

$$M_{sd}(Q_s) = Q_s \times L / 3 = 15.75 * Q_{s,var} * 6.0 / 3 = 31.5 * Q_{s,var}$$

The result is shown in Fig. 4.11.

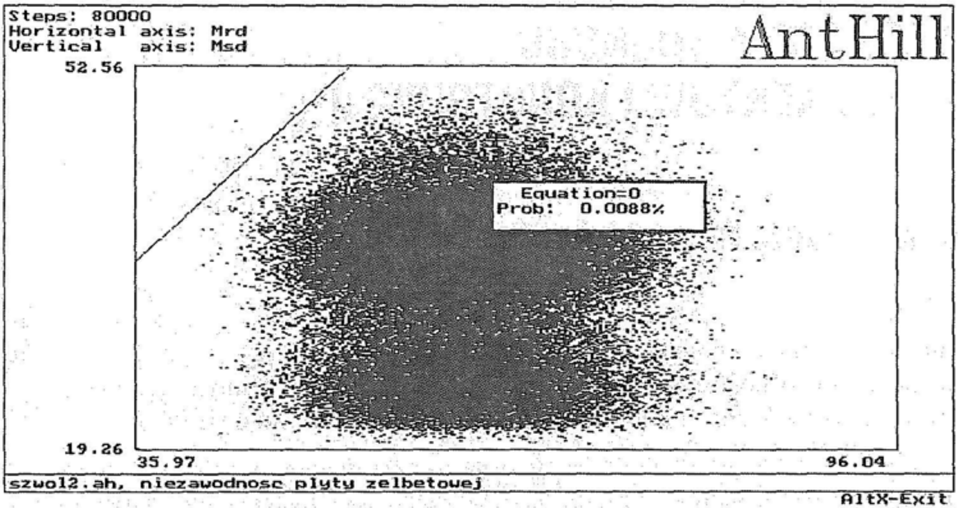


Rys. 4.11. Wydruk wyników obliczeń do przykładu 4.4 wykonanych przy pomocy programu ResCom

Fig. 4.11. The ResCom output of an example

The result is the value M_{sd} = 117.07 kNm
to be exceeded with a probability $q = 10^{-4}$.

A lower M_{sd} value occurs, referring to the same $q = 10^{-4}$. It yields from a deterministic beam length assumption of the second variant. The AntHill output



Rys. 4.13. Wydruk wyników obliczeń do przykładu 4.5, wykonanych za pomocą programu AntHill

Fig. 4.13. The AntHill output of an example