V. BĂDIN
The Academy of Economic Studies
Department of Mathematics
V. STOICESCU

Estimating Random Parameters in Some Economic Models through the Monte-Carlo Method

An effective control of production requires the replacement of the empirical methods by mathematical methods substantiated in a rigorous way.

I. The study of complex systems, which are to be understood as part or sum total of the means involved in the management and implementation of a certain project, is a matter of topical interest and outstanding economic importance. Such a study is generally achieved by drawing up mathematical models that cover the characteristic features of the processes taking place in the real system and thereby enable, through their careful examination, to reach in due time some results that can further be used in practical activities. We have in view in this line enterprises and factories pertaining to power industry, metallurgy, light industry, a.s.o.

A current mathematical model met with is the network graph drawn up in accordance with the plotting of the design and control process of the implementation of a new project. This graph facilitates an adequate mathematical description of the process and also helps us to find the algorithms required for the computation of the arising parameters, such as the time length of certain activities, the quantity of consumed labour, various costs, etc.

The more frequently used is the stochastic network model where, in contradistinction to the determinist model, the random character of certain parameters is taken into account. Such parameters are: the proximal time $t(k)$ of occurrence of an event k in which a number whatever of activities are involved, the strain coefficient $K_T(i, j)$ of a certain assignment, etc.*)

*) See article: V. Bădin & V. Stoicescu, *Modele economico-matematice in rețea cu parametri aleatori, folosite in planificarea unor obiective economice.* (Economic-Mathematical Models in Random Parameters Network for Planning Some Economic Projects). Studii și Cercetări Economice, No. 2, 1970.

To find the random parameters we shall use here the Monte-Carlo method, that consists in simulating the time of achievement of all the network assignments, in subsequently processing the corresponding data, in repeating the proceeding as many times as it may be necessary and, finally, in estimating the looked for parameters by means of adequate empirical distributions. The repeated simulation of the time required for the achievement of each activity (i, j) means to model the distribution law which is observed by the time length of the activity under consideration.

II. Applying the Monte-Carlo method to a network model means either to perform a n set number of simulations, or to carry out a number of simulations to be chosen in accordance with the probability convergency of the statistical characteristics of the parameters network under consideration.

We shall first examine the estimation of certain parameters met with in a network model by making use of a n set number of tests.

The network parameter we are to estimate first is the mean of the proximal term $t(k)$ for the occurrence of an event k in a network.

We write t_i for the value of variable $t(k) = t_i$ in the i-th occurrence. Let $\{t_i / 1 \leqslant i \leqslant n\}$ be the set of these values. The random variable

$$\bar{t}(k) = \sum_{i=1}^{n} \frac{1}{n} t_i = \bar{t} \tag{1}$$

is the estimation of mean value $M\,[t(k)]$ of the proximal term of occurrence of event k; we shall write m for this mean.

Dispersion $D[t(k)] = \sigma^2$ of random variable $t(k)$ has the estimation $S^2[t(k)] = S^2$ yielded by relationship

$$S^2 = \frac{1}{n} \sum_{i=1}^{n} (t_i - \bar{t})^2 \tag{2}$$

Because variables t_i are independent and $M\,(t_i) = m$, while $D(t_i) = \sigma^2$, we have:

$$M(\bar{t}) = M\left[\sum_{i=1}^{n} \frac{1}{n} t_i\right] = \frac{1}{n} nM(t_i) = m$$

and

$$D(\bar{t}) = \frac{1}{n} \cdot D[t(k)] = \frac{\sigma^2}{n}$$

In agreement with Lyapunov's theorem

$$\underset{n \to \infty}{P}\left(\frac{E - m}{\frac{\sigma}{\sqrt{n}}} < z\right) \to \frac{1}{2} + \frac{1}{\sqrt{2\pi}} \int_0^z e^{-\frac{x^2}{2}}\, dx$$

and

$$\underset{n\to\infty}{P}\left(\frac{E-m}{\frac{S}{\sqrt{n}}}<z\right)\to\frac{1}{2}+\frac{1}{\sqrt{2\pi}}\int_0^z e^{-\frac{x^2}{2}}\,dx \tag{3}$$

Consequently, for values of n sufficiently large we have:

$$P\left(|\bar{t}-m|<z\,\frac{S}{\sqrt{n}}\right)\approx 2\Phi(z) \tag{4}$$

where

$$\Phi(z)=\frac{1}{\sqrt{2\pi}}\int_0^z e^{-\frac{x^2}{2}}\,dx \quad (z>0)$$

is Laplace's function.

Given confidential probability γ in relationship $2\Phi(Z_\gamma)=\gamma$, we can find value z_γ so that:

$$P\left(\bar{t}-z_\gamma\frac{S}{\sqrt{n}}<m<\bar{t}+z_\gamma\frac{S}{\sqrt{n}}\right) \tag{5}$$

Thus, we have the unknown parameter $M\,[t(k)]=m$ of the network (the mean of proximal term of occurrence of event k) through a confidence interval; the confidential probability is γ.

Another network parameter to be estimated is the mean square deviation $\sigma[t\,(k)]=\sigma$ of the proximal term $t(k)=t_i$ of occurrence of event k. We have relationships:

$$M[t_i(k)-m]^2=M[t(k)-m]^2=D[t(k)]=\sigma^2$$

$$D(t_i-m)^2=D(t-m)^2=M(t-m)^4-M^2(t-m)=\mu_4-\mu_2^2=\sigma_2^2$$

We introduce now random variable

$$S_n^2=\frac{1}{n}\sum_{i=1}^{n}(t_i-m)^2. \tag{6}$$

The mean and dispersion of variable S_n^2 are

$$M(S_n^2)=M\left[\frac{1}{n}\sum_{i=1}^{n}(t_i-m)^2\right]=\frac{1}{n}\sum_{i=1}^{n}M(t_i-m)^2=$$

$$=\frac{1}{n}\,nM(t-m)^2=\sigma^2$$

$$D(S_n^2)=D\left[\frac{1}{n}\sum_{i=1}^{n}(t_i-m)^2\right]=\frac{1}{n^2}\sum_{i=1}^{n}D(t_i-m)^2=$$

$$=\frac{1}{n^2}\,nD(t_i-m)^2=\frac{\sigma_2^2}{n}$$

In agreement with the central limit theorem we have:

$$P\left(z_0 < \frac{S_n^2 - \sigma^2}{\frac{\sigma_2}{\sqrt{n}}} < z_1\right) \to \frac{1}{\sqrt{2\pi}} \int_{z_0}^{z_1} e^{-\frac{x^2}{2}} dx = \Phi(z_1) - \Phi(z_0) \tag{7}$$

that is the relationship from which it follows that variable $\frac{(S_n^2 - \sigma^2)\sqrt{n}}{\sigma_2}$ has an approximately normal distribution for large values of n.

Variables $\sqrt{n}\, S_n^2$ and $\sqrt{n}\, S^2$ have normal distributions for large values of n.

In order to prove this we compute the difference $S_n^2 - S^2$; S_n^2 and S^2 are yielded by relationships (2) and (6). We have:

$$S^2 = \frac{1}{n}\sum_{i=1}^{n} (t_i - \bar{t})^2 = \frac{1}{n}\sum_{i=1}^{n} (t_i - m + m - \bar{t})^2 =$$

$$= \frac{1}{n}\sum_{i=1}^{n} [(t_i - m)^2 + (\bar{t} - m)^2 - 2\,(t_i - m)\,(\bar{t} - m)] =$$

$$= \frac{1}{n}\sum_{i=1}^{n} (t_i - m)^2 + (\bar{t} - m)^2 - \frac{2}{n}(\bar{t} - m)\sum_{i=1}^{n}(t_i - m) = S_n^2 - (\bar{t} - m)^2,$$

because

$$\sum_{i=1}^{n} (t_i - m) = \sum_{i=1}^{n} t_i - nm = n\bar{t} - nm = n(\bar{t} - m).$$

Consequently,

$$S_n^2 - S^2 = (\bar{t} - m)^2.$$

On multiplying both members of the previous relationship by $\sqrt{n}$, we get

$$\sqrt{n}(S_n^2 - S^2) = \sqrt{n}(t - m)^2 \tag{8}$$

On the other hand, whatever $\varepsilon > 0$ may be, we have

$$\lim_{n\to\infty} P\left(|\,\bar{t} - m\,| > \frac{\varepsilon}{\sqrt[4]{n}}\right) = 0 \tag{9}$$

Indeed, when applying Chebyshev's inequality

$$P(|\,X - M(X)\,| > \varepsilon) \leqslant \frac{1}{\varepsilon^2} D(X)$$

to random variable $\bar{t}$ we get:

$$P\left(|\,\bar{t} - m\,| > \frac{\varepsilon}{\sqrt[4]{n}}\right) \leqslant \frac{\sqrt{n}}{\varepsilon^2} D(\bar{t}) = \frac{\sigma^2}{\varepsilon^2\sqrt{n}}$$

For $n \to \infty$ we get relationship (9).

So, it follows that variable $\sqrt[4]{n}\,|\,\bar{t} - m\,|$ converges in probability towards zero when $n \to \infty$.

Taking into account relationships (4) and (8) it follows that variables $\sqrt{n}\ S_n^2$ and $\sqrt{n}\ S^2$ have normal distributions for large values of n, that is to say

$$\lim_{n\to\infty} P\left(z_0 < \frac{S^2 - \sigma^2}{\frac{\sigma_2}{\sqrt{n}}} < z_1\right) = \frac{1}{\sqrt{2\pi}} \int_{z_0}^{z_1} e^{-\frac{x^2}{2}}\, dx. \tag{10}$$

Because $\sigma_2^2 = \mu_4 - \mu_2^2$, it follows that S^2 has a normal distribution with parameters $\left[\sigma^2, \frac{1}{n}(\mu_4 - \mu_2^2)\right]$.

Estimation S^2 is consequently an unshifted and consistent estimation of the unknown dispersion σ^2. Taking into account that variable $\frac{S}{\sigma}$ converges in probability towards 1 we get

$$\lim_{n\to\infty} P\left(Z_0 < \frac{S - \sigma}{\frac{\sigma_2}{2\sigma\sqrt{n}}} < z_1\right) = \frac{1}{\sqrt{2\pi}} \int_{z_0}^{z_1} e^{-\frac{x^2}{2}}\, dx \tag{11}$$

and, consequently, we infer that variable S for n sufficiently large is asymptotically normal with parameters $\left(\sigma, \frac{\sigma_2}{2\sigma\sqrt{n}}\right)$.

Let us estimate now value σ_2 using as estimations for the focussed moments μ_4 and μ_2 the corresponding selection moments m_4 and m_2.

We get the approximate equality

$$\sigma_2 \approx S_2 = \sqrt{\frac{\sum_{i=1}^{n} (t_i - \bar{t})^4}{n} - \left[\frac{1}{n}\sum_{i=1}^{n} (t_i - \bar{t})^2\right]}$$

or, taking into account relationship (2),

$$\sigma_2 \approx S_2 = \sqrt{\frac{1}{n}\left[\sum_{i=2}^{n} (t_i - \bar{t})\right]^4 - S^4} \tag{12}$$

We know that for a normal variable $\mu_4 = 3\sigma^4$; consequently $\sqrt{\mu_4 - \mu_2^2} = \sqrt{3\sigma^4 - \sigma^4} = \sigma^2\sqrt{2}$. It follows the approximate equality

$$\sigma_2 = \frac{\sigma_2}{2\sigma\sqrt{n}} = \frac{\sqrt{\mu_4 - \mu_2^2}}{2\sigma\sqrt{n}} \approx \frac{\sigma^2\sqrt{2}}{2\sigma\sqrt{n}} = \frac{\sigma}{\sqrt{2n}} \tag{13}$$

Relatioships (3) or (5) and (11) or (13) enable us to estimate number n of repeated simulations required for achieving the desired precision of estimations

$$m \approx \bar{t} \text{ and } \sigma^2 \approx S^2$$

For the mean, we have to assume a confidence coefficient γ, as well as estimation $\bar{t} = m$, so that relationship

$$P(|\bar{t} - m| < \varepsilon) > \gamma \tag{14}$$

shall take place.

Using relationship (4) and inequality $z_\gamma \frac{S}{\sqrt{n}} \leqslant \varepsilon$

we get

$$n \geqslant \frac{Z_\gamma^2 \sigma^2}{\varepsilon^2} = \frac{\Phi^{-1}\left(\frac{\gamma}{2}\right)\sigma^2}{\varepsilon^2} \tag{15}$$

If we assume ε of the form
$\varepsilon = q\,\sigma$, then relationship (15) becomes

$$n \geqslant \frac{\Phi^{-1}\left(\frac{\gamma}{2}\right)}{q^2} \tag{16}$$

or, further,

$$n \geqslant \frac{\Phi^{-1}\left(\frac{\gamma}{2}\right)S^2}{\varepsilon^2} = \frac{\Phi^{-1}\left(\frac{\gamma}{2}\right)}{q_1^2} \tag{17}$$

where

$$q_1 = \frac{\varepsilon}{S}$$

Similarly, we can estimate n starting from formulas (11) and (13) Writing relationship

$$P\left(|S - \sigma| < Z_\gamma \frac{\sigma}{\sqrt{n}}\right) = 2\Phi(z\gamma) \geqslant \gamma$$

and inequality $\frac{z_\gamma\,\sigma}{\sqrt{2n}} \leqslant \varepsilon_\gamma$ we obtain the estimation of n:

$$n \geqslant \frac{Z_\gamma^2 \sigma^2}{2\varepsilon_\sigma^2} = \frac{\Phi^{-1}\left(\frac{\gamma}{2}\right)\sigma^2}{2\varepsilon_\sigma^2} \tag{18}$$

or

$$n \geqslant \frac{\Phi^{-1}\left(\frac{\gamma}{2}\right)}{2q_\sigma^2} \tag{19}$$

where

$$q_\sigma = \frac{\varepsilon_\sigma}{\sigma}$$

III. Of particular practical importance is to find the number n of simulations necessary for learning to which of the α-quantile zones pertains a set activity (i, j). We recall that probabilities α_1 and α_2, $\alpha_2 \leqslant \alpha_1$ are assumed to be known, so that if $p_{i_j} > \alpha_1$ the activity pertains to the strain zone (the critical zone), if $p_{i_j} < \alpha_2$ the activity pertains to the reserve zone, and if $\alpha_2 \leqslant p_{i_j} < \alpha_1$ the activity (i, j) pertains to the intermediate zone. Number p_{i_j} stands for the probability that the strain coefficient K_T (i, j) of activity (i, j) assumes a value larger than $1 - \eta (\eta > 0)$ within the implementation of a project in the network.

Probability p_{i_j} is estimated by means of relative frequency $f_{i_j} = \frac{n_{i_j}}{n}$ where n_{i_j} is the number of cases out of n where value K_T (i, j) is larger than $1 - \eta$.

Let us assume that $p_{i_j} > \alpha_1$, that is activity (i, j) pertains to the strain zone. Based upon Moivre-Laplace's theorem we get

$$P\left\{\frac{n_{i_j}}{n} < \alpha_1 - K\sqrt{\frac{\alpha_1(1-\alpha_1)}{n}} < 1 - \frac{1}{\sqrt{2\pi}}\int_{-k}^{\infty} e^{-\frac{x^2}{2}}\,dx = \right. \tag{20}$$

$$= \frac{1}{\sqrt{2\pi}}\int_{\infty}^{-k} e^{-\frac{x^2}{2}}\,dx.$$

Consider the confidential probability γ that, if γ hypothesis $p_{i_j} > \alpha_1$ holds true, we shall not make — based on the ascertained empirical frequency $\frac{n_{i_j}}{n} = f_{i_j}$ — the decision of placing activity (i, j) in the second or third zone ($\gamma = 0.95$ or 0.99).

With a given γ we have relationship

$$P\left[\frac{n_{i_j}}{n} < \alpha_1 - k\sqrt{\frac{\alpha_1(1-\alpha_1)}{n}}\right] < 1 - \gamma$$

where k is inferred from relationship

$$F(k) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{k} e^{-\frac{x^2}{2}}\,dx = \gamma; \quad k = F^{-1}(\gamma) \tag{21}$$

For $\gamma = 0.95$ we get $k = 1.65$. We take notice that the difference between f_{i_j} and α_1 must not be, with probability γ, larger than

$$F^{-1}(\gamma)\sqrt{\frac{\alpha_1(1-\alpha_1)}{n}} = \varepsilon \tag{22}$$

Stating deviation ε, we estimate the value of n:

$$n \geqslant \frac{\alpha_1(1-\alpha_1)}{\varepsilon^2}[F^{-1}(\gamma)]^2 \tag{23}$$

Formula (23) enables us, based on the given confidence coefficient γ and the confidence interval $(\alpha_1 - \varepsilon, 1)$, to find n and subsequently compute

the relative frequency $f_{i_j} = \frac{n_{i_j}}{n}$. If $f_{i_j} < \alpha_1 - \varepsilon$, we consider the hypothesis to be false and activity (i, j) is not placed in the critical zone. Practice has proved that the use of this proceeding yields satisfactory results. It can also be used when one is interested to know if activity (i, j) pertains to the reserve zone or not. Confidence coefficient γ and deviation ε are stated; then, through n yielded by formula

$$n \geqslant \frac{\alpha_2(1 - \alpha_2)}{\varepsilon^2} [F^2(\gamma)]^2 \tag{24}$$

n simulations are performed. If value $\frac{n_{i_j}}{n} = f_{i_j}$ goes out of the interval $(0, \alpha_2 + \varepsilon)$, we assume activity (i, j) not to pertain to the reserve zone. For various values of γ and ε we get corresponding values for n. Thus, having $\gamma = 0.95$, $\varepsilon = 0.05$, $\alpha_1 = 0.5$ and $\alpha_2 = 0.2$, in order to decide if activity (i, j) pertains to the critical zone or not, we must carry out approximately $n = 270$ simulations, and in order to decide if activity (i, j) pertains to the reserve zone, $n = 220$ simulations are necessary.

Because values γ, ε and n are connected through relationship (23), the reverse problem — very often met with in practice — can also be approached: find γ(or ε) for n and ε (or γ) that have been given.

The placing of an activity (i, j) in the critical zone or the reserve zone depends obviously on the content we describe to the notions of strain and placing. It is usually assumed that an activity is placed (under strain) if during the implementation of a project in the network the respective activity is performed on the critical path or on a path whose length does not differ from the length of the critical path by more than q per cent.

Thus, it follows that, having found number n, we choose to place activity (i, j) in the critical zone if it is carried out on a path that does not differ by more that q per cent from the critical path, whereas in our previous reasoning we decided on the activities based on a strain coefficient larger that $1 - \eta$.

IV. We are often confronted with the problem of estimating number n. In order to solve it, we may use Kolmogorov's principle. This proceeding is particularly efficacious when we want to draw up the distribution function of the proximal term of achieving a certain activity k in the network model. Practice has proved that the main advantage of using Kolmogorov's principle lies in the possibility of previously finding the selection volume and the error with any degree of precision.

Some satisfactory estimations through statistical methods with a given confidence coefficient are reached by means of the empirical distribution of a finite number of observations $t_1(k)$, $t_2(k)$,...,$t_n(k)$ of the proximal term of achieving a certain activity in the network model using selections of a given volume n.

For the previous estimation of the selection volume n we must know coefficient γ with which we wish to draw the limits between which the distribution function of the general collectivity is to be comprised. In other

words, we must know the probability with which the theoretical distribution function (of the general collectivity) will be placed within the limits drawn based on the selection. The confidence coefficient γ is usually close to value 1; practically it is equal to 0.9 — 0.95.

It is known that $P\left\{max\ |F_n(x) - F(x)| < \frac{\lambda}{\sqrt{n}}\right\} = K(\lambda)$.

This relationship holds true for large values of n, $K(\lambda)$, having the expression

$$K(\lambda) = \sum_{k=-\infty}^{k=\infty} (-1)^k e^{-2k^2\lambda^2}$$

Consequently, having a confidence coefficient γ, we can find λ through formula $\gamma = K(\lambda)$.

Further, we must find the error of deviation $|F_n(x) - F(x)|$, where $F(x)$ is the distribution function and $F_n(x)$ is the curve of cumulated frequencies (the empirical distribution function) to be built based on the selection.

Consider

$$\varepsilon \geqslant max\,|F_n(x) - F(x)| = D_n \tag{25}$$

Hence we get

$$n \geqslant \frac{\lambda^2}{\varepsilon^2} = \frac{[K^{-1}(\gamma)]^2}{\varepsilon^2} \tag{26}$$

We see that n must not be smaller than the minimal admitted value n_{min}, from which the random variable $D_n\sqrt{n}$ is in accordance with a distribution close to Kolmogorov's distribution ($n_{min} \approx 100$). Therefore

$$n = max.\left(n_{min}, \frac{\lambda^2}{\varepsilon^2}\right) \tag{27}$$

We can also solve the reverse problem: Find ε (or γ) when γ and n (or ε and n) are known. It is often advisable to follow this way. If, for certain reasons, the selection volume is bounded, then $\varepsilon = \frac{\lambda}{\sqrt{n}}$, $\lambda = \varepsilon\sqrt{n}$ and, consequently, $\gamma = K(\varepsilon\sqrt{n})$. Further the volume selection n is effected and the curve of cumulated frequencies $F_n(x)$ is drawn up. Finally, the limits within which lies the diagram of the theoretical distribution function $F(x)$ are drawn:

$$F(x) = F_n(x) \pm \frac{\lambda}{\sqrt{n}}. \tag{28}$$

Because Kolmogorov's criterion is an integral one, $F(x)$ does not go out of the limits which have been found in any point of the field of possible values of x in agreement with probability γ.

The proceeding we have just described has, beside doubtless advantages, some shortcomings that must necessarily be taken into account when using it in practice.

One such shortcoming: the selection volume is lower bounded, that is $n \geqslant n_{min}$. Another shortcoming: we must often find limits for the distribution density $f(x)$ of the general collectivity, whereas with the aid of Kol-

mogorov's criterion, limits are currently drawn only for the distribution function $F(x)$.

Let us submit now a few remarks with regard to the estimation of the number tests n based on the information directly received in the process of modelling. Consider we must find the distribution density of the proximal term of carrying out a certain set activity k in a network project. In the beginning we state the field in which random values $t_i(k)$ are to be found, assuming i to be an index pointing out the number of activity simulations through the Monte-Carlo method. In particular, when the time length of all the network activities is assumed to be equal to the optimistic estimations of the given time lengths of the respective assignments, the proximal term of occurrence of event k is stated as lower limit of the field of the values, and an analogous parameter corresponding to the pessimistic estimations is stated as upper limit of the field.

The field thus stated may be considerably restricted because, with big networks the probability of attaining the extreme intervals at the boundaries of the values domain is nearly zero.

We divide into intervals the domain of values of the proximal term of achievement, whose number is to a certain extent arbitrary, stating it conveniently. After dividing the field into intervals, the first stage of testing is carried out, that is a selection is taken out of the set of values $t_i(k)$ of an arbitrary volume n_1 (that is to say, n_1 simulations are performed). We compute the empirical frequencies $\frac{m_i^{(1)}}{n_1}$ $(i = 1, 2, \ldots, l)$, where $m_i^{(1)}$ is the number of elements of the volume selection n_1 going into interval i. Further, a second stage of testing is carried out, that is a certain additional selection is taken out in view of bringing the total number of selection units to value n_2 (n_2 is also stated approximately). We compute again the empirical frequencies $\frac{m_i^{(2)}}{n_2}$, this time for a volume selection n_2. We carry out again an additional selection in view of bringing the selection volume to value n_3; we also compute frequencies $\frac{m_i^{(3)}}{n_3}$, a.s.o.

This iterative proceeding comes to an end when the maximal deviation (for all the intervals) of the frequency value during the last stage, as against the frequency value corresponding to the previous stage, becomes smaller than a set number δ, that is

$$\left| \frac{m_i^{(p)}}{n_p} - \frac{m_i^{(p-1)}}{n_{p-1}} \right| < \delta \tag{29}$$

for all i from 1 to l. In this case stage p will be the last.

This proceeding is based on the probability convergency of relative frequencies $\frac{m_i}{n}$ towards values p_i, where p_i are the probabilities yielded by relationship

$$p_i = \int_{x_i}^{x_{i+1}} p(x)\, dx,$$

where ($p(x)$ is the density of unknown theoretical distribution of variable $t(k)$ (the proximal term of achievement of activity k) and x_i and x_{i+1} are the extremities of interval i.

The empirical distribution drawn up based on relative frequencies $\frac{m_i}{n}$ in the last stage is assumed as approximate value of theoretical distribution density of variable $t(k)$.

Finally, if the selection volume n is bounded and the selection is carried out once, we advise for using the following elementary proceeding. As before, the domain is divided into l sub-intervals and empirical frequencies $\frac{m_i}{n}$ ($i = 1, 2, \ldots, l$) are computed. If $n \geqslant 20$, we can draw for each subinterval the following confidence limits for the distribution density:

$$\frac{m_i}{n} \pm 3\sqrt{\frac{m_i}{n}\left(1 - \frac{m_i}{n}\right)\frac{1}{n}} = \frac{m_i}{n} \pm \frac{3}{n}\sqrt{m_i\left(1 - \frac{m_i}{n}\right)} \tag{30}$$

Value $\frac{m_i}{n}\left(1 - \frac{m_i}{n}\right)$ approximates the empirical dispersion of frequencies distribution by which the theoretical dispersion is estimated. Joining the limits corresponding to all the intervals we get the general confidence limits for the whole density of empirical distribution. We can assume that the density of theoretical distribution is comprised between these limits.

We stress that, although this proceeding is frequently used in practice, it does not yield satisfactory results for certain distributions. Therefore we should be very cautious when using it.

When the estimation of the distribution density of the proximal term of achievement is carried out using at the same time various activities k_1, $k_2,\ldots,k_r$, formula (29) undergoes certain alterations. We write $m_i^{(p)}(K_j)$ for the number of elements in the volume selection n_p going into the subinterval i of the domain of values of the proximal term of occurrence of event k_j. Then formula (29) acquires the following form:

$$\max_{i,\ j}\left|\frac{m_i^{(p)}(k_j)}{n_p} - \frac{m_i^{(p-1)}(k_j)}{n_{p-1}}\right| < \varepsilon \tag{31}$$

$$1 \leqslant i \leqslant l, \quad 1 \leqslant j \leqslant r.$$

In other words, a selection is taken out as long as the relative estimations at the vertices of network $k_1, k_2,\ldots,k_r$ do not become smaller than a set error ε.

For computing the empirical distribution of the proximal term $i(t_p(i))$ of occurrence of the event, the following proceeding of estimation of the selection volume can be used:

We write $F_p^{(n)}(i)$ and $F_p(i)$ for the empirical distribution function and, respectively, the theoretical distribution function of $t_p(i)$. In agreement with Glivenko's theorem, for $n \to \infty$ we have:

$$P\{Sup\,|F_p^{(n)}(i) - F_p(i)| \to 0\} = 1$$

$$-\infty < t < +\infty$$

which expresses the convergency of the empirical distribution function towards the theoretical one.

Hence, the subsequent algorithm making use of Kolmogorov's criterion.

In the beginning a number n_1 of simulations for the value of variable t_p (i) is carried out, after which the distribution empirical function $F_p^{(n_1)}$ (i) is drawn; further, an additional number of simulations is carried into effect as long as their total number becomes n_2 and the empirical function $F_p^{(n_2)}$ (i) is drawn. By equalizing function $F_p(i)$ to the theoretical distribution function $F_p(i)$ we compute values:

$$D_{n_1} = max \mid F_p^{(n_2)}(i) - F_p^{(n_1)}(i) \mid \text{ and } \lambda(n_1) = \sqrt{n_1}\, D_{n_1}$$

We state a confidence probability $p(n_1) = K[\lambda(n_1)]$ and decide if the difference between $F_p^{(n_1)}$ (i) and $F_p^{(n_2)}$ (i) is significant or not.

If the difference is not significant, we assume $n = n_2$ and the process of modelling comes to an end. Otherwise, an additional number of testings is carried out on variable $t_p(i)$, increasing their number up to value n_3. Then we compare $F_p^{(n_3)}$ (i) to $F_p^{(n_2)}$ (i) a.s.o. up to a certain stage r when the difference between $F_p^{(n_r)}$ (i) and $F_p^{(n_{r+1})}$ (i)can be assumed to be insignificant. In this case, the necessary number of testings is $n = n_{r+1}$.

As a rule, the use of the Monte-Carlo method involves a large number of operations and calculi.

REFERENCES

1. BADIN V., STOICESCU V.: *Modele economico-matematice în rețele cu parametri aleatori utilizați în planificarea unor obiective economice. (Economic-Mathematical Models in Network with Random Parameters Used in Planning Some Economic Objectives)*. Studii și cercetări economice, No. 2, 1970.
2. GRAMER H. : *Mathematical Methods of Statistics. Stockholm*, 1946.
3. DEMIDOVICI B.P., MARON I.A.: *Osnovy vycislitelnoi matematiki*, Moscow, 1960.
4. NAIDOV JELEZOV: *Osnovy postroyeniya i parametri setevoi kompleksa operatsii*. Iz. Ekonomika, Moscow, 1967.
5. SIROINSKI E.V.: *Odin algoritm optimizatsii setevoi modeli metodom Monte Carlo*. Trudi M.G.P.I. im V.I.Lenina, 1968.

ESTIMATION DES PARAMETRES ALÉATOIRES DE CERTAINS MODÈLES ÉCONOMIQUES PAR LA MÉTHODE MONTE-CARLO

On y présente le mode d'application de la méthode Monte-Carlo pour estimer certains paramètres aléatoires d'un modèle économique en réseau, spécialement celle des paramètres liés à la durée de réalisation des évenements ou des activités du processus étudié. Y sont estimés le temps prochain $t(k)$ de réalisation d'un évènement auquel concourent un nombre quelconque d'activités, le coefficient de tension K_T (i, j) d'une certaine activité, ainsi que le nombre des simulations.

ОЦЕНКА СЛУЧАЙНЫХ ПАРАМЕТРОВ НЕКОТОРЫХ ЭКОНОМИЧЕСКИХ МОДЕЛЕЙ ПО МЕТОДУ МОНТЕ-КАРЛО

В настоящей статье представляется способ применения метода Монте-Карло для оценки случайных параметров экономической модели в сети, в особенности параметров в связи с продолжительностью осуществления событий или видов деятельности в изучаемом процессе. Производится оценка ближайшего времени $t(k)$ осуществления события k, которому способствует определенное число видов деятельности, коэффициента напряжения K_T (i, j) определенного вида деятельности и числа симулирований.