Mathematica Balkanica

Mathematical Society of South-Eastern Europe
A quarterly published by
the Bulgarian Academy of Sciences – National Committee for Mathematics

---

Mathematica Balkanica

New Series Vol. 6, 1992, Fasc. 3

Minimization of the Probable Error of the Monte Carlo Method for Solving of Nonlinear Integral Equation

Todor Gjurov

Presented by P. Kenderov

The paper deals with the analysis of the probable error of Monte Carlo method for solving of nonlinear integral equation.

1. Introduction

Monte Carlo algorithms for calculation of functionals

(1) $$(g, \varphi)=\int g(x)\varphi(x)\,dx$$

are considred.

Everywhere below integration means integration in the domain $G \subset R^n$. The function $g(x)$ belongs to the space $L_2(G)$ and $x=(x^{(1)}, x^{(2)},\dots,x^{(n)})\in G$ is a point in the Euclidean space R^n. The function $\varphi(x)$ is a solution of the integral equation with polynomial nonlinearity.

(2) $$\varphi(x)=\int\dots\int K(x, y_1, y_2,\dots,y_m)\prod_{i=1}^{m}\varphi(y_i)\prod_{i=1}^{m}dy_i+f(x)$$

(2a) $$(\varphi=K\varphi^{(m)}+f),\ (m\geqq 2, \text{ is natural number})$$

where K, f are given functions and $f\in L_2(G)$, $K\in L_2(G\times\dots\times G)$.

It is assumed that this equation has an iteration solution corresponding to the iteration process

$$\varphi_{n+1}(x)=\int\dots\int K(x, y_1, y_2,\dots,y_m)\prod_{i=1}^{m}\varphi_n(y_i)\prod_{i=1}^{m}dy_i+f(x),$$

(3) $$\varphi_0(x)=f(x),$$

or

(3a) $$\varphi_{n+1}=K\varphi_n^{(m)}+f,\ \varphi_n^{(m)}=\prod_{i=1}^{m}\varphi_n(y_i),\ \varphi_0=f.$$

If the method of successive approximations

$$\varphi_{n+1}(x)=\int\ldots\int|K(x, y_1, y_2,\ldots,y_m)|\prod_{i=1}^{m}\varphi_n(y_i)\prod_{i=1}^{m}dy_i+|f(x)|,$$

$$\varphi_0(x)=|f(x)|,$$

converges, then the use of a branching process [1] (this process is described in the next section) enable us to construct random variables which mathematical expectations are equal to functionals (1).

2. Monte Carlo algorithms for estimation of functions (1)

In this section we present the Monte Carlo algorithms for estimation of (1). We refer to S. M. Ermakov, G. A. Mikhailov [1] for more details.

We consider a branching stochastic process included in the following scheme: Any particle with initial frequency function $p_0(x)\geq 0$ $(\int p_0(x)\,dx=1)$ is born in the domain $G\subset R^n$ in a random point $x=x_0$. In the next moment this particle either dies with probability $h(x)$ $(0\leq h(x)<1)$ or generates a posterity of m analogical particles $(y_1, y_2,\ldots,y_m)$ with probability $p_m(x)=1-h(x)$ and transition frequency function

$$\frac{p(x_0, y_1, y_2,\ldots,y_m)}{p_m(x)}, \text{ where}$$

$$\int\ldots\int p(x, y_1, y_2,\ldots,y_m)\,dy_1\ldots dy_m=1-h(x).$$

The generated particles behave in the next moment as the initial one and etc. The traces of such a process is a tree from the type sketched in Fig. 1. The description of this process is closely connected with the equation (2).

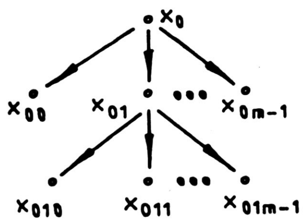

Fig. 1

Fig. 2

The used index numbers in Fig. 1 are called multi-indexes. The particle from the zero generation x is numerated with zero index (x_0). Its direct inheritors — with the indexes 00, 01, 02,.., $0m-1$ (first generation). If the particle from the k-th generation has the multi-index $\nu[k]$, then the multi-index of l-th inheritor of this particle has the following form $\nu[k+1]=(\nu[k], l)$, $l=0, 1,\ldots,m-1$. In this way the multi-index is a number written in m-th numerical system.

Now we consider a simple case when $m=2$ and the first two iterations of (3), (see [1], p. 254).

$$\varphi_0(x_0)=f(x_0)$$

$$\varphi_1(x_0)=f(x_0)+\iint K(x_0,\ x_{00},\ x_{01})f(x_{00})f(x_{01})\,dx_{00}\,dx_{01}$$

$$\varphi_2(x_0)=f(x_0)+\iint K(x_0,\ x_{00},\ x_{01})f(x_{00})f(x_{01})\,dx_{00}\,dx_{01}$$

$$+\iint K(x_0,\ x_{00},\ x_{01})f(x_{00})\Big(\iint K(x_{01},\ x_{010},\ x_{011})f(x_{010})f(x_{011})\,dx_{010}\,dx_{011}\Big)dx_{00}$$

$$\times dx_{01}+\iint K(x_0,x_{00},x_{01})f(x_{01})\Big(\iint K(x_{00},x_{000},x_{001})f(x_{000})f(x_{001})\,dx_{000}\,dx_{001}\Big)$$

$$\times dx_{00}\,dx_{01}+\iint K(x_0,\ x_{00},\ x_{01})\Big(\iint K(x_{01},\ x_{010},\ x_{011})f(x_{010})f(x_{011})\,dx_{010}\,dx_{011}\Big)$$

$$\times\Big(\iint K(x_{00},\ x_{000},\ x_{001})f(x_{000})f(x_{001})\,dx_{000}\,dx_{001}\Big)dx_{00}\,dx_{01}. \tag{4}$$

It is obvious, that the structure of φ_1 is linked with all trees, which appear till first generation (Fig. 2), and the structure of φ_2 is linked with all trees till the second generation (Fig. 3).

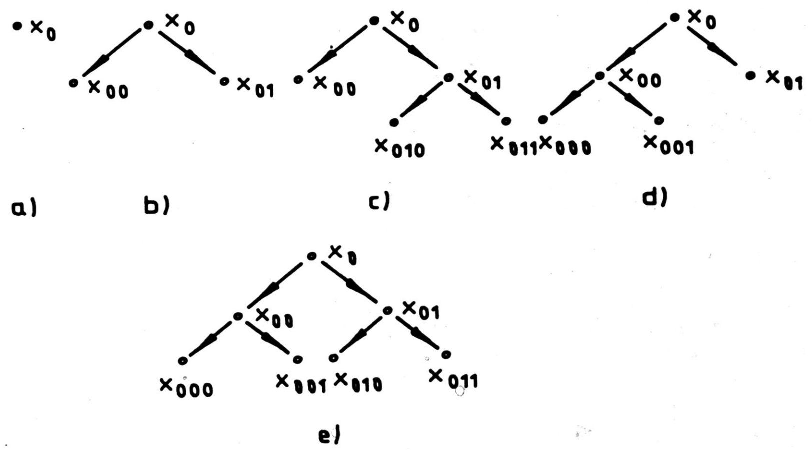

Fig. 3

This similarity allows to construct a procedure of a random choice of subtrees of some full tree and to calculate the values of some random variable (r. v.). This r. v. corresponds to the random choice subtree. Thus the arithmetic mean over N independent samples (N is very large) of such random variables is an estimate of the functional (1).

Definition. A full tree with n generations is called the tree Γ_n, where the dying of particles is not visible from zero to n-2nd generation, but all the generated particles of n-1st generation die.

Example. In fact Γ_3 is the tree γ_0 in Fig. 3e, and Γ_2 is the tree in Fig. 2b. Evidently, the following density corresponds to the tree γ_0 from Fig. 3e:

$$p_{\gamma_0}=p_0(x_0)p(x_0,\ x_{00},\ x_{01})p(x_{00},\ x_{000},\ x_{001})p(x_{01},\ x_{010},\ x_{011})$$

(5) $$\times h(x_{000})h(x_{001})h(x_{010})h(x_{011}).$$

Then the random variable which corresponds to γ_0 is:

(6) $$\theta_{[g]}(\gamma_0)=\frac{g(x_0)K(x_0,\ x_{00},\ x_{01})K(x_{00},\ x_{000},\ x_{001})K(x_{01},\ x_{010},\ x_{011})}{p_0(x_0)p(x_0,\ x_{00},\ x_{01})p(x_{00},\ x_{000},\ x_{001})p(x_{01},\ x_{010},\ x_{011})}\times\frac{f(x_{000})f(x_{001})f(x_{010})f(x_{011})}{h(x_{000})h(x_{001})h(x_{010})h(x_{011})}$$

This r. v. estimates the last of the additionals in (4) if the condition is that the realization of the random process appears to be a tree from the kind γ_0. Thus random variables are constructed, which correspond to trees of another type.

We consider the branching stochastic process in the general case ($m=$const). First some notation is introduced.

Let A denote the set of points, which generate new points and B denote the set of points, which die.

We suppose, the tree $\gamma\in\Gamma_n$, i. e. γ is a subtree of the full tree Γ_n ($n>1$), and that γ has more of one generation. (The case with one generation is not interesting.)

We receive any trees γ when modulating the branching process and we correspond to everyone of them the r. v. $\theta_{[g]}(\gamma)$ in the following way:

$$\theta_{[g]}(\gamma)=\frac{g(\xi_0)}{p_0(\xi_0)},\ \text{if the tree consists of the initial point.}$$

If the tree consists of other points, then $\theta_{[g]}(\gamma)$ is constructed simultaneously with the construction of γ. When we have a transition from point ξ to $\xi_1,\ \xi_2,\dots,\xi_m$, we multiply by $\dfrac{K(\xi,\ \xi_1,\dots,\xi_m)}{p(\xi,\ \xi_1,\dots,\xi_m)}$, and when the point ξ dies — by $\dfrac{f(\xi)}{h(\xi)}$.

Then the r. v. $\theta_{[g]}(\gamma)$ which corresponds to γ is:

(7) $$\theta_{[g]}(\gamma)=\frac{g(\xi_0)}{p_0(\xi_0)}\prod_{x_{\gamma[q]}\in A}\frac{K(\xi_{\gamma[q]})}{p(\xi_{\gamma[q]})}\prod_{x_{\gamma[p]}\in B}\frac{f(\xi_{\gamma[p]})}{h(\xi_{\gamma[p]})},$$

with the frequency function

(8) $$p_\gamma=p_0(\xi_0)\prod_{x_{\gamma[q]}\in A}p(\xi_{\gamma[q]})\prod_{x_{\gamma[p]}\in B}h(\xi_{\gamma[p]}),$$

$$(1\leqq q\leqq n-1,\ 1<p\leqq n-1),$$

where

$$K(\xi_{\gamma[q]})\stackrel{\text{def}}{=}K(\xi_{\gamma[q]},\ \xi_{\gamma[q]0},\dots,\xi_{\gamma[q]m-1}),$$

$$p(\xi_{\gamma[q]})\stackrel{\text{def}}{=}p(\xi_{\gamma[q]},\ \xi_{\gamma[q]0},\dots,\xi_{\gamma[q]m-1}).$$

Thus the mathematical expectation is

$$E\theta_{[g]}(\gamma)=(g,\ \varphi)=\int g(x)\varphi(x)\,dx,$$

where $\varphi(x)$ is the solution of (2) (see [1], [2]).

We propose with probability 1 that all the trees have finite number of generations and the arithmetic mean over particles which are born in the any generation is also finite. In this way the procedure of estimation exists really. The points ξ, ξ_j $(j=0, 1, 2, \dots, m)$ (see above) are random points in the domain G with a frequency function $p_0(x_0)$ (if $j=0$) and a transition frequency function $p(x, y_1, \dots, y_m)$ (if $j=1, 2, \dots, m$).

Besides, the initial frequency function $p_0(x_0)$ and the transition frequency function $p(x, y_1, \dots, y_m)$ are assumed to be admissible with respect to the functions $g(x)$ and $K(x, y_1, \dots, y_m)$, correspondingly.

We call the initial frequency function $p_0(x_0)$ admissible to the function $g(x)$ if

$$p_0(x)=\begin{cases} p_0(x)>0, & \text{for } x: g(x)\neq 0 \\ p_0(x)\geqq 0, & \text{for } x: g(x)=0 \end{cases}$$

and the transition frequency function $p(x, y_1, \dots, y_m)$ is called admissible to the kernel $K(x, y_1, \dots, y_m)$ if

$$p(x, y_1, \dots, y_m)=\begin{cases} p(x, y_1, \dots, y_m)>0, & \text{for } (x, y_1, \dots, y_m): K(x, y_1, \dots, y_m)\neq 0 \\ p(x, y_1, \dots, y_m)\geqq 0, & \text{for } (x, y_1, \dots, y_m): K(x, y_1, \dots, y_m)=0. \end{cases}$$

Then, the series $\frac{1}{N}\left(\sum_{s=1}^{N} \theta_{[g]}(\gamma)\right)_s \approx (g, \varphi)$ is an estimate of the functional (1); $(\theta_{[g]}(\gamma))_s$ is the value of $\theta_{[g]}(\gamma)$ for the "s" trial of the tree γ; N is the number of the trials.

In this case the probable error is [1]

$$r_n = k\sigma(\theta_{[g]}(\gamma))N^{-1/2},$$

where $k \approx 0.6745$ and $\sigma(\theta_{[g]}(\gamma))$ is the standard deviation.

3. Minimization of the probable error of the Monte Carlo method

The problem of optimization of the Monte Carlo algorithms consists in the minimization of $\sigma^2(\theta_{[g]}(\gamma))$.

We process to the minimization of the dispersion $\sigma^2(\theta_{[g]}(\gamma))$ by means of a suitably chosen frequency function p_γ of (8).

Let us introduce the functions

$$\Phi(x)=\left(\int\dots\int \frac{K^2(x, y_1, \dots, y_m)}{p(x, y_1, \dots, y_m)} \prod_{i=1}^{m} \Phi^2(y_i) \prod_{i=1}^{m} dy_i\right)^{1/2} \tag{9}$$

$$\hat{\Phi}(x)=\int\dots\int \left|K(x, y_1, \dots, y_m) \prod_{i=1}^{m} \Phi(y_i)\right| \prod_{i=1}^{m} dy_i, \tag{10}$$

where $p(x, y_1, \dots, y_m)$ is an admissible frequency function to the kernel $K(x, y_1, \dots, y_m)$.

Lemma 1. *The transition frequency function*

$$p(x, y_1,\dots,y_m)=\frac{|K(x, y_1,\dots,y_m)\prod_{i=1}^{m}\Phi(y_i)|}{\int\dots\int|K(x, y_1,\dots,y_m)\prod_{i=1}^{m}\Phi(y_i)|\prod_{i=1}^{m}dy_i}$$

minimizes $\Phi(x)$ *for any* x *from* G *and*

$$\min_p \Phi(x)=\hat{\Phi}(x).$$

Proof. Let us substitute $p(x, y_1,\dots,y_m)$ in (9)

$$\Phi(x)=\left(\int\dots\int\frac{K^2(x, y_1,\dots,y_m)\prod_{i=1}^{m}\Phi^2(y_i)}{|K(x, y_1,\dots,y_m)\prod_{i=1}^{m}\Phi(y_i)|}\left(\int\dots\int|K(x, y_1,\dots,y_m)\prod_{i=1}^{m}\Phi(y_i)|\right.\right.$$

$$\left.\left.\times\prod_{i=1}^{m}dy_i\right)\prod_{i=1}^{m}dy_i\right)^{1/2}=\int\dots\int|K(x, y_1,\dots,y_m)\prod_{i=1}^{m}\Phi(y_i)|\prod_{i=1}^{m}dy_i=\hat{\Phi}(x).$$

Now we must show that for any other admissible to the kernel $K(x, y_1,\dots,y_m)$ transition frequency function, there holds

$$\hat{\Phi}(x)\leqq\Phi(x).$$

In fact, if we multiply and divide the integrand in (10) by $p^{1/2}(x, y_1,\dots,y_m)>0$ and apply the Cauchy–Schwarz inequality, then we have

$$\hat{\Phi}^2(x)=\left(\int\dots\int|K(x, y_1,\dots,y_m)\prod_{i=1}^{m}\Phi(y_i)|p^{-1/2}(x, y_1,\dots,y_m)\right.$$

$$\left.\times p^{1/2}(x, y_1,\dots,y_m)\prod_{i=1}^{m}dy_i\right)^2$$

$$\leqq\int\dots\int K^2(x, y_1,\dots,y_m)\prod_{i=1}^{m}\Phi^2(y_i)p^{-1}(x, y_1,\dots,y_m)\prod_{i=1}^{m}dy_i$$

$$\times\int\dots\int p(x, y_1,\dots,y_m)\prod_{i=1}^{m}dy_i$$

$$\leqq\int\dots\int K^2(x, y_1,\dots,y_m)\prod_{i=1}^{m}\Phi^2(y_i)p^{-1}(x, y_1,\dots,y_m)\prod_{i=1}^{m}dy_i=\Phi^2(x),$$

because $\int\dots\int p(x, y_1,\dots,y_m)\prod_{i=1}^{m}dy_i=1-h(x)=p_m(x)<1$ for any $x\in G$.

Lemma 2. *The initial frequency function*

$$p_0(x_0)=|g(x_0)\Phi(x_0)|/\int|g(x_0)\Phi(x_0)|\,dx_0$$

minimizes the functional

$$\int g^2(x_0)\Phi^2(x_0)p_0^{-1}(x_0)\,dx_0.$$

The minimum of this functional is equal to

$$\left(\int|g(x_0)\Phi(x_0)|\,dx_0\right)^2.$$

Proof. The proof is similar to the proof of Lemma 1

$$\int g^2(x_0)\Phi^2(x_0)p_0^{-1}(x_0)\,dx_0$$
$$=\int g^2(x_0)\Phi^2(x_0)|g(x_0)\Phi(x_0)|^{-1}\left(\int|g(x_0)\Phi(x_0)|\,dx_0\right)dx_0$$
$$=\left(\int|g(x_0)\Phi(x_0)|\,dx_0\right)^2.$$

It remains to establish that for any another frequency function the inequality holds

$$\left(\int|g(x_0)\Phi(x_0)|\,dx_0\right)^2\leqq\int g^2(x_0)\Phi^2(x_0)p_0^{-1}(x_0)\,dx_0.$$

In fact

$$\left(\int|g(x_0)\Phi(x_0)|\,dx_0\right)^2=\left(\int|g(x_0)\Phi(x_0)|p_0^{-1/2}(x_0)p_0^{1/2}(x_0)\,dx_0\right)^2$$
$$\leqq\int g^2(x_0)\Phi^2(x_0)p_0^{-1}(x_0)\,dx_0\int p_0(x_0)\,dx_0$$
$$=\int g^2(x_0)\Phi^2(x_0)p^{-1}(x_0)\,dx_0. \qquad \blacksquare$$

Before we prove the main result we consider the r.v. $\theta_{[\sigma]}(\gamma_0)$ (6), which estimates the last of the additionals in (4) in private case when $m=2$. We prove the following theorem 1.

Theorem 1. *Introduce the next constant*

$$c_0=\left(\int|g(x_0)\hat{\Phi}(x_0)|\,dx_0\right)^{-1},$$

where $\hat{\Phi}(x)$ *is a function of kind* (10) *for* $m=2$ *and the next function*

$$\Phi(x)=\frac{|f(x)|}{h(x)^{1/2}}.$$

Then the frequency function

$$\hat{p}_{\gamma_0}=c_0|g(x_0)||K(x_0,x_{00},x_{01})||K(x_{00},x_{000},x_{001})||K(x_{01},x_{010},x_{011})|$$
$$\times h(x_{000})h(x_{001})h(x_{010})h(x_{011})\stackrel{\text{def}}{=}\hat{p}_0(x_0)\hat{p}(x_0,x_{00},x_{01})$$
$$\times\hat{p}(x_{00},x_{000},x_{001})\hat{p}(x_{01},x_{010},x_{011})\frac{h(x_{000})h(x_{001})h(x_{010})h(x_{011})}{\Phi(x_{000})\Phi(x_{001})\Phi(x_{010})\Phi(x_{011})}$$

minimizes the second moment $E\theta^2_{[\sigma]}(\gamma_0)$ *of the r.v.* $\theta_{[\sigma]}(\gamma_0)$, *i.e.*

$$\min_{p_{\gamma_0}} (E\theta^2_{[\sigma]}(\gamma_0)) = E\hat{\theta}^2_{[\sigma]}(\gamma_0), \textit{ when } p_{\gamma_0} = \hat{p}_{\gamma_0}.$$

Proof. Let $\hat{\theta}_{[\sigma]}(\gamma_0)$ is the r.v. $\theta_{[\sigma]}(\gamma_0)$ for which $p_{\gamma_0} = \hat{p}_{\gamma_0}$. Then

$$E\hat{\theta}^2_{[\sigma]}(\gamma_0) = \int \dots \int \theta^2_{[\sigma]}(\gamma_0) p_{\gamma_0} \, dx_0 \dots dx_{011}$$

$$= \int \dots \int \frac{g^2(x_0) K^2(x_0, x_{00}, x_{01}) K^2(x_{00}, x_{000}, x_{001}) K^2(x_{01}, x_{010}, x_{011})}{p_0^2(x_0) p^2(x_0, x_{00}, x_{01}) p^2(x_{00}, x_{000}, x_{001}) p^2(x_{01}, x_{010}, x_{011})}$$

$$\times \frac{f^2(x_{000}) f^2(x_{001}) f^2(x_{010}) f^2(x_{011})}{h^2(x_{000}) h^2(x_{001}) h^2(x_{010}) h^2(x_{011})} p_0(x_0) p(x_0, x_{00}, x_{01})$$

$$\times p(x_{00}, x_{000}, x_{001}) p(x_{01}, x_{010}, x_{011}) h(x_{000}) h(x_{001}) h(x_{010})$$

$$\times h(x_{011}) \, dx_0 \dots dx_{011}$$

$$= \int \dots \int \frac{g^2(x_0) K^2(x_0, x_{00}, x_{01}) K^2(x_{00}, x_{000}, x_{001}) K^2(x_{01}, x_{010}, x_{011})}{p_0(x_0) p(x_0, x_{00}, x_{01}) p(x_{00}, x_{000}, x_{001}) p(x_{01}, x_{010}, x_{011})}$$

$$\times \Phi^2(x_{000}) \Phi^2(x_{001}) \Phi^2(x_{010}) \Phi^2(x_{011}) \, dx_0 \dots dx_{011}$$

$$= \int \frac{g^2(x_0)}{p_0(x_0)} \left\{ \iint \frac{K^2(x_0, x_{00}, x_{01})}{p(x_0, x_{00}, x_{01})} \left[\iint \frac{K^2(x_{00}, x_{000}, x_{001})}{p(x_{00}, x_{000}, x_{001})} \Phi^2(x_{000}) \Phi^2(x_{001}) \right. \right.$$

$$\left. \times dx_{000} \, dx_{001} \right] \left[\iint \frac{K^2(x_{01}, x_{010}, x_{011})}{p(x_{01}, x_{010}, x_{011})} \Phi^2(x_{010}) \Phi^2(x_{011}) \, dx_{010} \, dx_{011} \right]$$

$$\left. \times dx_{00} \, dx_{01} \right\} dx_0 \tag{11}$$

$$E\hat{\theta}^2_{[\sigma]}(\gamma_0) = \int \dots \int \hat{\theta}^2_{[\sigma]}(\gamma_0) \hat{p}_{\gamma_0} \, dx_0 \dots dx_{011}$$

$$= \int \dots \int \left[\frac{g(x_0) K(x_0, x_{00}, x_{01}) K(x_{00}, x_{000}, x_{001}) K(x_{01}, x_{010}, x_{011})}{c_0 |g(x_0)| |K(x_0, x_{00}, x_{01})| |K(x_{00}, x_{000}, x_{001})| |K(x_{01}, x_{010}, x_{011})|} \right.$$

$$\left. \times \frac{f(x_{000}) f(x_{001}) f(x_{010}) f(x_{011})}{h(x_{000}) h(x_{001}) h(x_{010}) h(x_{011})} \right]^2 c_0 |g(x_0)| |K(x_0, x_{00}, x_{01})|$$

$$\times |K(x_{00}, x_{000}, x_{001})| |K(x_{01}, x_{010}, x_{011})| h(x_{000}) h(x_{001}) h(x_{010})$$

$$\times h(x_{011}) \, dx_0 \dots dx_{011}$$

$$= c_0^{-1} \int |g(x_0)| \{ \iint | K(x_0, x_{00}, x_{01})| [\iint | K(x_{00}, x_{000}, x_{001}) \Phi(x_{000}) \Phi(x_{001})|$$

$$\times dx_{000} \, dx_{001}] [\iint |K(x_{01}, x_{010}, x_{011}) \Phi(x_{010}) \Phi(x_{011})| \, dx_{010} \, dx_{011}]$$

$$\times dx_{00} \, dx_{01} \} dx_0$$

$$= \{ \int |g(x_0)| \iint | \{K(x_0, x_{00}, x_{01})| [\iint |K(x_{00}, x_{000}, x_{001}) \Phi(x_{000}) \Phi(x_{001})| \times dx_{000} dx_{001}] [\iint |K(x_{01}, x_{010}, x_{011}) \Phi(x_{010}) \Phi(x_{011})| dx_{010} dx_{011}] \times dx_{00} dx_{01} \} dx_0 \}^2 \tag{12}$$

According to (9), (10) the functions $\Phi(x)$ and $\hat{\Phi}(x)$ are non-negative for any $x \in G$, and so the frequency function p_{γ_0} may be written in the following form:

$$p_{\gamma_0} = c_0 |g(x_0)| |K(x_0, x_{00}, x_{01})| |K(x_{00}, x_{000}, x_{001})| |(K(x_{01}, x_{010}, x_{011})| \times h(x_{000}) h(x_{001}) h(x_{010}) h(x_{011})$$

$$= \frac{|g(x_0)\hat{\Phi}(x_0)|}{c_0^{-1}} \frac{|K(x_0, x_{00}, x_{01})\hat{\Phi}(x_{00})\hat{\Phi}(x_{01})|}{\hat{\Phi}(x_0)}$$

$$\times \frac{|K(x_{00}, x_{000}, x_{001})\hat{\Phi}(x_{000})\hat{\Phi}(x_{001})|}{\hat{\Phi}(x_{00})} \frac{|K(x_{01}, x_{010}, x_{011})\hat{\Phi}(x_{010})\hat{\Phi}(x_{011})|}{\hat{\Phi}(x_{01})}$$

$$\times \frac{h(x_{000}) h(x_{001}) h(x_{010}) h(x_{011})}{\Phi(x_{000}) \Phi(x_{001}) \Phi(x_{010}) \Phi(x_{011})}$$

$$= \frac{|g(x_0)\hat{\Phi}(x_0)|}{c_0^{-1}} \frac{|K(x_0, x_{00}, x_{01})\hat{\Phi}(x_{00})\hat{\Phi}(x_{01})|}{\iint |K(x_0, x_{00}, x_{01})\Phi(x_{00})\Phi(x_{01})| dx_{00} dx_{01}}$$

$$\times \frac{|K(x_{00}, x_{000}, x_{001})\hat{\Phi}(x_{000})\hat{\Phi}(x_{001})|}{\iint |K(x_{00}, x_{000}, x_{001})\Phi(x_{000})\Phi(x_{001})| dx_{000} dx_{001}}$$

$$\times \frac{|K(x_{01}, x_{010}, x_{011})\hat{\Phi}(x_{010})\hat{\Phi}(x_{011})|}{\iint |K(x_{01}, x_{010}, x_{011})\Phi(x_{010})\Phi(x_{011})| dx_{010} dx_{011}}$$

$$\times \frac{h(x_{000}) h(x_{001}) h(x_{010}) h(x_{011})}{\Phi(x_{000}) \Phi(x_{001}) \Phi(x_{010}) \Phi(x_{011})}$$

$$= p(x_0) p(x_0, x_{00}, x_{01}) p(x_{00}, x_{000}, x_{001}) p(x_{01}, x_{010}, x_{011}) \times \frac{h(x_{000}) h(x_{001}) h(x_{010}) h(x_{011})}{\Phi(x_{000}) \Phi(x_{001}) \Phi(x_{010}) \Phi(x_{011})},$$

Because $\hat{\Phi}(x) \geq 0$.

Lemma 1 (for $m=2$) should be applied to the frequency functions $p(x_{01}, x_{010}, x_{011})$, $p(x_{00}, x_{000}, x_{001})$ and $p(x_0, x_{00}, x_{01})$ in (11) and (12), because:

$$\iint \frac{K^2(x, y_1, y_2)\Phi^2(y_1)\Phi^2(y_2)}{p(x, y_1, y_2)} dy_1 dy_2 = \iint \frac{K^2(x, y_1, y_2)\hat{\Phi}^2(y_1)\hat{\Phi}^2(y_2)}{\hat{p}(x, y_1, y_2)} dy_1 dy_2$$

$$= \iint \frac{K^2(x, y_1, y_2)\hat{\Phi}^2(y_1)\hat{\Phi}^2(y_2)}{|K(x, y_1, y_2)\hat{\Phi}(y_1)\hat{\Phi}(y_2)|} \left(\iint |K(x, y_1, y_2)\Phi(y_1)\Phi(y_2)| dy_1 dy_2\right) dy_1 dy_2$$

$$= \iint |K(x, y_1, y_2)\hat{\Phi}(y_1)\hat{\Phi}(y_2)| dy_1 dy_2 \iint |K(x, y_1, y_2)\Phi(y_1)\Phi(y_2)| dy_1 dy_2 .$$

It should be taken into account that for

$$p(x, y_1, y_2) = \hat{p}(x, y_1, y_2) = \frac{|K(x, y_1, y_2)\hat{\Phi}(y_1)\hat{\Phi}(y_2)|}{\iint |K(x, y_1, y_2)\Phi(y_1)\Phi(y_2)| dy_1 dy_2}$$

the corresponding functions $\Phi(x)$ and $\hat{\Phi}(x)$ are equal to each generation point x in the branching process, i.e.

$$\Phi(x) = \hat{\Phi}(x),$$

because the corresponding frequency functions minimize $\Phi(x)$ and their minima are equal to the functions $\hat{\Phi}(x)$ for any $x \in G$.

So, for any frequency function $p_0(x_0)$ which is admissible to the function $g(x_0)$

$$\min_{\Phi} \left(\int g^2(x_0)\Phi^2(x_0)p_0^{-1} x_0) dx_0\right) = \int g^2(x_0)\hat{\Phi}^2(x_0)p_0^{-1}(x_0) dx_0$$

holds and in this case

$$p_{\gamma_0} = p_0(x_0)|K(x_0, x_{00}, x_{01})||K(x_{00}, x_{000}, x_{001})||K(x_{01}, x_{010}, x_{011})|$$
$$\times h(x_{000})h(x_{001})h(x_{010})h(x_{011}) \tag{13}$$

According to Lemma 2, the last functional is minimized by the frequency function (13) if

$$p_0(x_0) = \hat{p}_0(x_0) = |g(x_0)\hat{\Phi}(x_0)| / \int |g(x_0)\hat{\Phi}(x_0)| dx_0 ,$$

since

$$\Phi(x) = \hat{\Phi}(x)$$

under the condition (13). This completes the proof. ∎

We consider the branching process in the general case (m=const). Now we prove

Theorem 2. *Introduce the next constant*

$$c = \left(\int |g(x_0)\hat{\Phi}(x_0)| dx_0\right)^{-1},$$

where $\hat{\Phi}(x)$ *is a function of kind* (10) *and the next function*

$$\Phi(x)=\frac{|f(x)|}{h(x)^{1/2}}.$$

Then the frequency function

$$\hat{p}_\gamma=c|g(x_0)|\prod_{x_{\nu[q]}\in A}|K(x_{\nu[q]})|\prod_{x_{\nu[p]}\in B}h(x_{\nu[p]})$$

minimizes the second moment $E\theta^2_{[g]}(\gamma)$ *of the r.v.* $\theta_{[g]}(\gamma)$ *for any* $\gamma\in\Gamma_n$, *i.e.*

$$\min_{p_\gamma} E\theta^2_{[g]}(\gamma)=E\hat{\theta}_{[g]}(\gamma),\ \textit{when}\ p_\gamma=\hat{p}_\gamma.$$

Proof. Introduce the following function

$$F(x_{\nu[q]})=\begin{cases}\hat{\Phi}(x_{\nu[q]}), & \text{for } x_{\nu[q]}\in A\\ \Phi(x_{\nu[q]}), & \text{for } x_{\nu[q]}\in B\end{cases}$$

Let $\hat{\theta}_{[g]}(\gamma)$ is the r.v. $\theta_{[g]}(\gamma)$ for which $p_\gamma=\hat{p}_\gamma$. Then

$$E\theta^2_{[g]}(\gamma)=\int\dots\int\theta^2_{[g]}(\gamma)p_\gamma\prod_{x_{\nu[q]}\in A\cup B}dx_{\nu[q]} \tag{14}$$

$$=\int\dots\int\frac{g^2(g_0)}{p_0(x_0)}\prod_{x_{\nu[q]}\in A}\frac{K^2(x_{\nu[q]})}{p(x_{\nu[q]})}\prod_{x_{\nu[q]}\in B}F^2(x_{\nu[q]})\prod_{x_{\nu[q]}\in A\cup B}dx_{\nu[q]}$$

$$E\hat{\theta}^2_{[g]}(\gamma)=\int\dots\int\hat{\theta}^2_{[g]}(\gamma)\hat{p}_\gamma\prod_{x_{\nu[q]}\in A\cup B}dx_{\nu[q]} \tag{15}$$

$$=\left(\int\dots\int|g(x_0)|\prod_{x_{\nu[q]}\in A}|K(x_{\nu[q]})|\prod_{x_{\nu[q]}\in B}F(x_{\nu[q]})\prod_{x_{\nu[q]}\in A\cup B}dx_{\nu[q]}\right)^2.$$

According to (9), (10) the functions $\Phi(x)$ and $\hat{\Phi}(x)$ are non-negative for any $x\in G$, and so the frequency function p_γ may be written in the following form:

$$\hat{p}_\gamma=c|g(x_0)|\prod_{x_{\nu[q]}\in A}|K(x_{\nu[q]})|\prod_{x_{\nu[p]}\in B}h(x_{\nu[p]})$$

$$=\frac{|g(x_0)F(x_0)|}{c^{-1}}\prod_{x_{\nu[q]}\in A}\frac{|K(x_{\nu[q]})F(x_{\nu[q]0})\dots F(x_{\nu[q]m-1})|}{F(x_{\nu[q]})}\prod_{x_{\nu[p]}\in B}\frac{h(x_{\nu[p]})}{F(x_{\nu[p]})}$$

$$=\frac{|g(x_0)F(x_0)|}{c^{-1}}$$

$$\times\prod_{x_{\nu[q]}\in A}\frac{|K(x_{\nu[q]})F(x_{\nu[q]0})\dots F(x_{\nu[q]m-1})|}{\int\dots\int|K(x_{\nu[q]})F(x_{\nu[q]0})\dots F(x_{\nu[q]m-1})|\,dx_{\nu[q]0}\dots dx_{\nu[q]m-1}}$$

$$\times\prod_{x_{\nu[p]}\in B}\frac{h(x_{\nu[p]})}{F(x_{\nu[p]})}=\hat{p}_0(x_0)\prod_{x_{\nu[q]}\in A}\hat{p}(x_{\nu[q]})\prod_{x_{\nu[p]}\in B}\frac{h(x_{\nu[p]})}{F(x_{\nu[p]})}, \tag{16}$$

since $F(x)\geq 0$.

Let the total number in the set A is M ($m<\infty$, because the branching process is finite). Then Lemma 1 should be applied "M" times to the frequency function $\hat{p}(x_{\nu[q]})$ in (14) and (15), as we begin of the frequency functions, which describe the last generation and we go to the first generation. It can be separated in (14) terms of the following form

$$\int\ldots\int\frac{K^2(x, y_1,\ldots,y_m)}{p(x, y_1,\ldots,y_m)}\sum_{i=1}^{m}F^2(y_i)\prod_{i=1}^{m}dy_i,$$

depending on the generation for that we consider the corresponding frequency function $\hat{p}(x_{\nu[q]})$.

Thus Lemma 1 should be applied "M" times to the frequency function $p(x, y_1,\ldots,y_m)$ in (14), because:

$$\int\ldots\int\frac{K^2(x,y_1,\ldots,y_m)}{p(x, y_1,\ldots,y_m)}\prod_{i=1}^{m}F^2(y_i)\prod_{i=1}^{m}dy_i=\int\ldots\int\frac{K^2(x,y_1,\ldots,y_m)}{\hat{p}(x,y_1,\ldots,y_m)}\prod_{i=1}^{m}F^2(y_i)\prod_{i=1}^{m}dy_i$$

$$=\int\ldots\int\frac{K^2(x,y_1,\ldots,y_m)\prod_{i=1}^{m}F^2(y_i)}{|K(x,y_1,\ldots,y_m)\prod_{i=1}^{m}F(y)|}\left(\int\ldots\int|K(x,y_1,\ldots,y_m)\prod_{i=1}^{m}F(y_i)|\prod_{i=1}^{m}dy_i\right)\prod_{i=1}^{m}dy_i$$

$$=\left(\int\ldots\int|K(x, y_1,\ldots,y_m)\prod_{i=1}^{m}F(y_i)|\prod_{i=1}^{m}dy_i\right)^2.$$

It should be taken into account that for

$$p(x, y_1,\ldots,y_m)=\hat{p}(x, y_1,\ldots,y_m)=\frac{|K(x, y_1,\ldots,y_m)\prod_{i=1}^{m}F(y_i)|}{\int\ldots\int|K(x, y_1,\ldots,y_m)\prod_{i=1}^{m}F(y_i)|\prod_{i=1}^{m}dy_i}$$

the corresponding functions $\Phi(x)$ and $\hat{\Phi}(x)$ are equal to each generation point x in the branching process, i.e.

$$\Phi(x)=\hat{\Phi}(x)=F(x),$$

because the corresponding frequency functions minimize $\Phi(x)$ and their minima are equal to the functions $\hat{\Phi}(x)$ for any $x\in G$.

In the end, for any initial frequency function $p_0(x_0)$, which is admissible to the function $g(x_0)$ it holds:

$$\min_{\Phi}\left(\int g^2(x_0)\Phi^2(x_0)p_0^{-1}(x_0)\,dx_0\right)=\int d^2(x_0)\hat{\Phi}^2(x_0)p_0^{-1}(x_0)\,dx_0$$

and in this case

$$p_\gamma = p_0(x_0) \prod_{x_{\nu[q]} \in A} |K(x_{\nu[q]})| \prod_{x_{\nu[q]} \in B} h(x_{\nu[p]}). \tag{17}$$

According to Lemma 2, the last functional is minimizes by the frequency function (17) if

$$p_0(x_0) = \hat{p}_0(x_0) = |g(x_0)\hat{\Phi}(x_0)| / \int g(x_0)\hat{\Phi}(x_0)|\,dx_0,$$

because

$$\Phi(x) = \hat{\Phi}(x)$$

under the condition (17). This completes the proof. ■

Since

$$\sigma^2(\theta_{[g]}(\gamma)) = E\theta^2_{[g]}(\gamma) - (E\theta_{[g]}(\gamma))^2$$

it is easy to show, that the frequency function $\hat{p}_\gamma$ minimizes the standard deviation $\sigma(\theta_{[g]}(\gamma))$. It is so, because $E\theta_{[g]}(\gamma) = \text{const}$ for any admissible frequency function.

4. Algorithms with null probable error

These algorithms could be very important for solving calculation problems, which need a big calculation resource.

Let us write the expression for $\sigma^2(\theta_{[g]}(\gamma))$:

$$\sigma^2(\theta_{[g]}(\gamma)) = E\theta^2_{[g]}(\gamma) - (E\theta_{[g]}(\gamma))^2$$

$$= \int \dots \int \theta^2_{[g]}(\gamma) p_\gamma \prod_{x_{\nu[q]} \in A \cup B} dx_{\nu[q]} - \Big(\int \dots \int \theta_{[g]}(\gamma) p_\gamma \prod_{x_{\nu[q]} \in A \cup B} dx_{\nu[q]}\Big)^2.$$

It is easy to see that the necessary conditions for the zero probable error for the optimal frequency function (when $p_\gamma = \hat{p}_\gamma$) are: The functions $K(x, y_1, \dots, y_m)$, $f(x)$, $g(x)$ to be non-negative for any $x \in G$.

References

[1] S. M. Ermakov, G. A. Mikhailov. Statistical simulation. Moskow, Nauka, 1982 (In Russian).
[2] I. M. Sobol. Monte Carlo numerical methods. Moskow, Nauka, 1973 (In Russian).
[3] I. T. Dimov. Minimization of the probable error for Monte Carlo methods. Application of Mathematics in Technology. Different Equations and Applications. Varna 1986, Sofia 1987, 161-164. (In Bulgarian).
[4] S. M. Ermakov. On summation of series connected with integral equation. *Vestnik Leningrad Univ. Math.*, 16, 1984, 57-63. (In Russian).

Center of Informatics and Computer Technology,
Bulgarian Academy of Sciences,
Acad. G. Bonchev str., block 25-A,
1113 Sofia,
BULGARIA

Received 20.12.1991