Chapter

Stability of Algorithms in Statistical Modeling

Alexander A. Kronberg and Tatiana K. Kronberg

Abstract

In this paper, we investigate algorithms stability for calculation of multidimensional integrals using the statistical modeling methods. We considered issues of the algorithms optimization and we give sufficient conditions for the stability. We apply our approach to both calculation of integral from the regression function and the moments integral calculation. In all our numerical experiences, we used the mt19937 pseudorandom number generator.

Keywords: statistical modeling, pseudorandom numbers, optimal density, integral estimation, Monte Carlo methods

1. Introduction

One of the main problem of the statistical modeling method (the Monte Carlo method) is the problem of quality for pseudorandom numbers. In the paper, we consider a task of multidimensional integrals calculation by the statistical modeling method and give sufficient conditions for the stability of this task to quality of pseudorandom numbers. Included results of various numerical experiences with the mt19937 pseudorandom number generator. In our work, we discuss important issues of algorithms optimization in the statistical modeling. In particular, we apply the new approach to the following: a task of finding of integral functionals from solution of boundary-value problems for both the linear [1] or nonlinear [2] elliptic equations (the estimations are given near to a boundary).

The paper is organized as follows: In Section 2, we give the sufficient conditions of stability. Calculation of an integral of very large dimensions is discussed in Section 3. Rare events effect is the subject of Section 4. In Section 5, we describe calculation of integral moments. In Section 6, we apply our approach to calculate an integral of the regression function. In Section 7, we give the conclusion of our studies.

2. Sufficient conditions of stability

Let

$$I = \int_D f(x)dx \tag{1}$$

be the Riemann integral. Here D is a domain of the s-dimensional Euclidean space R^s. If the dimension s is large enough then we must use a statistical modeling method. In this case, our integral has form of the mathematical expectation for a random value $\eta = f(\xi)/p(\xi)$:

$$\int_D f(x)dx = \int_D p(x)\frac{f(x)}{p(x)}dx = \mathcal{E}\frac{f(x)}{p(x)} = \mathcal{E}\eta. \tag{2}$$

Here $p(x)$ is a density of random variable $\xi \in D$. We put $p(x) \neq 0$ for $f(x) \neq 0$, and we say that there exists integral

$$\int_D |f(x)|dx.$$

A variance of the random value η:

$$\sigma^2 = \mathbf{var}\eta = \mathcal{E}\eta^2 - (\mathcal{E}\eta)^2 = \int_D p(x)\left[\frac{f(x)}{p(x)}\right]^2 dx - I^2 = \int_D \frac{f^2(x)}{p(x)}dx - I^2. \tag{3}$$

We estimate the mathematical expectation $\mathcal{E}\eta$ by the sum $\sum_{i=1}^N \eta_i/N$, where η_i are independent realizations of the random value η. Suppose σ^2 is finite, and N is large enough; then from the classical central limit theorem, it follows that the random value $\sum_{i=1}^N \eta_i/N$ has distribution close to the normal distribution with a mathematical expectation I, and mean-square deviation $\sigma/\sqrt{N}$. This property above is useful to estimate error, e.g., using the 3σ rule. So we have

$$\left| I - \frac{1}{N}\sum_{i=1}^N \eta_i \right| \leq \frac{3\sigma}{\sqrt{N}} \tag{4}$$

with probability 0,997, approximately.

Let us η_i be realizable sampling values; then the value σ is estimated as the following:

$$\sigma^2 \approx \frac{1}{N}\sum_{i=1}^N \eta_i^2 - \left(\frac{1}{N}\sum_{i=1}^N \eta_i\right)^2. \tag{5}$$

Suppose $\mathcal{E}\eta^4$ is finite and in Eq. (3) we replace the σ by its approximate value. Then it changes the estimation of error in calculation of the integral in order of $O\left(\frac{1}{N}\right)$.

In practice, when we simulate random variables ξ, we receive simulation with some density $q(x)$ instead of simulation with the origin density $p(x)$. Now, we investigate the stability of the theoretical estimation Eq. (2). Let us consider the following expression:

$$\begin{aligned}\int_D p(x)\frac{f(x)}{p(x)}dx - \int_D p(x)\frac{q(x)}{p(x)}dx = \int_D \frac{f(x)}{p(x)}(p(x) - q(x))dx \leq \\ \leq \varepsilon \int_D \frac{|f(x)|}{p(x)}dx,\end{aligned} \tag{6}$$

where $\varepsilon = \sup_{x \in D}|p(x) - q(x)|$. By $I(|f|/p)$ denote the integral $\int \frac{|f(x)|}{p(x)}\,dx$. The both values ε and $I(|f|/p)$ provide the guaranteed proximity of the real estimation to the theoretical one of the integral.

Example 1. The inequality Eq. (6) is reduced to the equality if $D = D_1 \cup D_2, D_1 \cap D_2 = \varnothing$. For $x \in D_1$ we get $p(x) - q(x) \equiv \varepsilon > 0$, and $f(x) \equiv 0$ in D_2. If the condition $I(|f|/p) = +\infty$ holds, then the error of the real estimation of the integral will be infinity for any $\varepsilon > 0$.

Hence, a quality of pseudorandom variables (i.e., smallness of ε) does not yet guaranties the smallness of the error in general, as the integral $I(|f|/p)$ have to be both finite and not very great in magnitude.

Suppose we simultaneously make the estimations for both $I(|f|/p)$ and the origin integral Eq. (2) using the same density $p(x)$. Then we need to ask boundedness of the integral $\int_D \frac{|f(x)|}{p^2(x)}dx$ to get the guaranteed stability of the estimation for the integral $I(|f|/p)$, and so on. The qualitative comparison of simulation with both densities $p_1(x)$ and $p_2(x)$ can be provided not only by a magnitude of the variance estimation (here we do not pay attention to the complexity of random values simulation) but also magnitudes of both the integrals $I(|f|/p_1)$ and $I(|f|/p_2)$. On the other hand we have the Schwarz inequality:

$$\int_D \frac{f(x)}{p(x)}(p(x) - q(x))\,dx \leq \sqrt{\int_D [p(x) - q(x)]^2 dx}\sqrt{\int_D \frac{f^2(x)}{p^2(x)}\,dx}. \tag{7}$$

The sufficient condition for the estimation to be stability is that $\int_D \frac{f^2(x)}{p^2(x)}$ to be finite and not very great. The Schwarz inequality is reduced to the equality if and only if

$$\lambda \frac{f(x)}{p(x)} = p(x) - q(x), \tag{8}$$

where λ is a real number. From the above we get

$$\begin{gathered} q(x) = p(x) - \lambda \frac{f(x)}{p(x)}, \\ \int_D q(x)dx = \int_D p(x)dx - \lambda \int_D \frac{f(x)}{p(x)}\,dx. \end{gathered} \tag{9}$$

Therefore, the equality in Eq. (7) is reached under the necessary condition $\int_D \frac{f(x)}{p(x)}\,dx = 0$, when $\int_D p(x)dx = \int_D q(x)dx = 1$.

Example 2. The condition above is realized, e.g., if

$$\begin{gathered} D = [-1, 1], \quad D_1 = [-1, 0], \quad D_1 = [0, 1], \\ f(x) = -1 \,\text{in}\, D_1, \quad f(x) = 1 \,\text{in}\, D_2, \quad p(x) = p(-x). \end{gathered}$$

Let us η_k be $[f(\xi)/p(\xi)]^k$. Now we consider an estimation:

$$\mathcal{E}\eta_k = \int_D p(x)\left[\frac{f(x)}{p(x)}\right]^k dx = \int_D \frac{f^k(x)}{p^{k-1}(x)}\,dx, \quad k \geq 1. \tag{10}$$

This expectation is actually estimated by the integral: $\int_D q(x)[f(x)/p(x)]^k dx$:

$$\int_D \frac{f^k(x)}{p^{k-1}(x)}dx - \int_D q(x)\frac{f^k(x)}{p^k(x)}dx = \int_D [p(x)-q(x)]\frac{f^k(x)}{p^k(x)}dx \le \le \varepsilon \int_D \frac{|f^k(x)|}{p^k(x)}dx. \tag{11}$$

The last integral is assumed to be a finite, and not very large. These conditions are desirable. In Eq. (11) the equality is reached like to the Example 1.

For all cases above, the stability will be observed if $|f(x)/p(x)| \le M < +\infty$ for not very great M. From the Schwarz inequality we have:

$$\int_D [p(x)-q(x)]\frac{f^k(x)}{p^k(x)}dx = \int_D \frac{p(x)-q(x)}{p^\beta(x)} \cdot \frac{f^k(x)}{p^{k-\beta}(x)}dx \le \le \sqrt{\int_D \frac{[p(x)-q(x)]^2}{p^{2\beta}(x)}dx}\sqrt{\int_D \frac{f^{2k}(x)}{p^{2(k-\beta)}(x)}dx}, \tag{12}$$

where β is a real number. We get a family of proximity measures for the distribution densities:

$$\int_D \frac{[p(x)-q(x)]^2}{p^{2\beta}(x)}dx. \tag{13}$$

For $\beta = 0,5$ we obtain expression

$$\chi^2(p,q) = \int_D \frac{[p(x)-q(x)]^2}{p^{(x)}dx.} \tag{14}$$

that is well known in the mathematical statistics.

For $\beta = -0,5$ we get $\int_D [p(x)-q(x)]^2 p(x)dx$ and have the obvious inequalities

$$\int_D [p(x)-q(x)]^2 p(x)dx \le \sup_{x\in D} p(x) \cdot \int_D [p(x)-q(x)]^2 dx,$$
$$\int_D [p(x)-q(x)]^2 p(x)dx \le \sup_{x\in D}[p(x)-q(x)]^2 \int_D p(x)dx = \sup_{x\in D}[p(x)-q(x)]^2.$$

In Eq. (6) the equality is satisfied if and only if

$$\frac{p-q}{p^\beta} = \lambda \frac{f^k}{p^{k-\beta}}, \quad p-q = \lambda \frac{f^k p^\beta}{p^{k-\beta}}, \tag{15}$$

i.e., the necessary condition is $\int_D \frac{f^k(x)}{p^{k-2\beta}(x)}dx = 0$. This is realized in the Example 2.

Let us remark that for $k=1$ and $\beta = 1$ we have

$$\int_D \frac{p(x)-q(x)}{p(x)} f(x)dx = \int_D \frac{p-q}{\sqrt{p}} \cdot \frac{f}{\sqrt{p}} \le \le \sqrt{\int_D \frac{(p-q)^2}{p}dx}\sqrt{\int_D \frac{f^2}{p}dx} = \sqrt{\chi^2(p,q)} \cdot \mathcal{E}\eta_2. \tag{16}$$

We assume that integrals $\int_D [f(x) + \varepsilon_i(x)]\,dx$ are known and the subintegral functions $f(x) + \varepsilon_i(x) > 0$ close to a function $f(x) \geq 0$. Suppose also

$$1 - \delta_1(\varepsilon_i) < \frac{f(x)}{f(x) + \varepsilon_i(x)} < 1 + \delta_2(\varepsilon_i),$$
$$J(f) - \delta_3(\varepsilon_i) \leq J(f + \varepsilon_i) \leq J(f) + \delta_4(\varepsilon_i), \quad \delta_j \to 0, \text{ as } \varepsilon_i \to 0, \; j = 1, 2, 3, 4. \quad (17)$$
$$\int_D q(x)\frac{f(x)}{p(x)}\,dx = \int_D q(x)\frac{f(x)J(f + \varepsilon_i)}{f(x) + \varepsilon_i(x)}\,dx = \mathcal{E}\hat{\eta},$$

where the random value $\hat{\eta}$ has a form:

$$\hat{\eta} = \frac{f(\hat{\xi})J(f + \varepsilon_i)}{f(\hat{\xi}) + \varepsilon_i(\hat{\xi})}. \quad (18)$$

Here $\hat{\xi}$ is distributed with the density $q(x)$. Keeping the above factors in mind we get the following:

$$[J(f) - \delta_3][1 - \delta_1] \leq \eta \leq [1 + \delta_2][J(f) + \delta_4];$$

i. Regardless of $q(x)$, i.e., regardless of quality of a pseudorandom number generator we have $\mathcal{E}\hat{\eta} \to J(f), \boldsymbol{var}\hat{\eta} \to 0$, as$\varepsilon \to 0$;

ii. All moments $\mathcal{E}\hat{\eta}^k$ of the random variable are finite.

Let we calculate $\int_D \frac{|f(x)|}{p(x)}\,dx$ using the density

$$p_1(x) = \frac{|f(x)|}{I(|f|/p)p(x)},$$

then the estimation variance equals to zero.

Suppose we calculate the integral $\int_D f(x)\,dx$ with the density $p_1(x)$. In this case it would be interesting to know both the values $\int_D \frac{|f(x)|}{p_1}\,dx$ and $\int_D \frac{f^2}{p_1^2}\,dx$.

Proposition 1. $I(|f|/p_1) = I(|f|/p)$.

Proposition 2. $I(f^2/p_1^2) = I^2(|f|/p)I(p^2)$.

Now we consider the density

$$p_2(x) = \frac{f^2(x)}{p(x)I(f^2/p)}.$$

Using the density above for the estimation of the integral $\int_D \frac{f^2}{p}\,dx$ we obtain the estimation variance equals to zero.

Further we estimate the integral $\int_D f(x)\,dx$ with the density p_2: $\hat{\eta} = f(\xi_2)/p(\xi_2)$, ξ_2 is distributed with $p_2(x)$. Suppose $\eta = f(\xi)/p(\xi)$, ξ is distributed with $p(x)$; then $I(f^2/p_2) = I(f^2/p)$.

Proposition 3. $\boldsymbol{var}\,\eta = \boldsymbol{var}\,\hat{\eta}$.

Proposition 4. $\left[\int\limits_D \frac{f(x)}{p(x)}\,dx\right]^2 \le (vol\,D)\cdot \int\limits_D \frac{f^2(x)}{p^2(x)}\,dx,$where $vol\,D$ is the volume of the domain D.

Proposition 5. If $vol\,D = 1, p_3(x) = \frac{f^2(x)}{I(f^2)}, \ \eta_3 = \frac{f(\xi_3)}{p(\xi_3)},$ where ξ_3 is a random variable distributed with the density of $p_3(x)$; $\eta_4 = \frac{f(\xi_4)}{p(\xi_4)},$ where ξ_4 is a random variable distributed with the density of $p(x) \equiv 1,$ then $\boldsymbol{var}\,\eta_3 = \boldsymbol{var}\,\eta_4$.

In actual practice normalization constants are usually unknown for both $p_1(x)$ and $p_2(x)$. But using densities close to them we can get the approximate equalities in the Prepositions 1, 2, 3.

Now we consider

$$I(f) = \int\limits_{[0,1]^{10}} x_1 x_2 \dots x_{10}\,dx_1 \dots dx_{10}, \tag{19}$$

where the integration domain $D = [0,1]^{10}$ is the 10-dimensional unit cube, the subintegral function $f(x)$ is equals to $x_1x_2 \dots x_{10}$. To realize algorithms of the statistical modeling at a computer it is necessary to set a number N of realizations for random variable $\eta = f(\xi)/p(\xi)$, where ξ is distributed with the density $p(\xi)$. In fact we realize the discrete set of numbers $\xi_i, \ \ i = 1, \dots, N,$ which we can consider to be realizations of some distribution $q_N(x)$.

In all our numerical computations we use the pseudorandom number generator: *generator type* mt19937 [3]. For $s = 10, \ \ p(x) \equiv 1$ we have

$$I(f) = (1/2)^{10} \approx 9,7656 \cdot 10^{-4}, \quad I(f/p) = I(f),$$
$$I(f^2/p) = I(f^2/p^2) = (1/3)^{10} \approx 1,6935 \cdot 10^{-5},$$
$$\sigma = \sqrt{(1/3)^{10} - (1/2)^{20}} \approx 3,998 \cdot 10^{-3}.$$

Table 1 shows the empirical estimations $\hat{I}$ and $\hat{\sigma}$ for I and σ, respectively. Taking $p(x_i) = 3x_i^2$ over each coordinate we get

$$I(f^2/p) = (1/3)^{10}, \quad I(f/p) = \infty, \quad I(f^2/p^2) = \infty,$$
$$\sigma = \sqrt{(1/3)^{10} - (1/2)^{20}}.$$

The value σ is the same as one for $p(x) \equiv 1$.

N	$\hat{I}$	$\hat{\sigma}$
1,000,000	$9,785 \cdot 10^{-4}$	$3,991 \cdot 10^{-3}$
9,000,000	$9,768 \cdot 10^{-4}$	$4,001 \cdot 10^{-3}$
81,000,000	$9,766 \cdot 10^{-4}$	$3,997 \cdot 10^{-3}$
100,000,000	$9,765 \cdot 10^{-4}$	$3,995 \cdot 10^{-3}$

Table 1.
The results of numerical calculations for the integral Eq. (19) with the uniform density.

N	$\hat{I}$	$\hat{\sigma}$
1,000,000	$9,766 \cdot 10^{-4}$	$3,276 \cdot 10^{-3}$
9,000,000	$9,754 \cdot 10^{-4}$	$3,619 \cdot 10^{-3}$
81,000,000	$9,765 \cdot 10^{-4}$	$3,759 \cdot 10^{-3}$
100,000,000	Inf	nan

Table 2.
The results of numerical calculations for the integral Eq. (19) with the density $p(x_i) = 3x_i^2$.

For $N = 100000000$ the computer code outputs an error because of machine zero divide. The reason of this event is $\eta = \prod_1^{10} \left(1/\left(3\sqrt[3]{\alpha_i}\right)\right)$, where α_i are the pseudorandom numbers with the uniform density in $(0, 1)$ [3]. If the formulas of random numbers simulation generate division by very small numbers then such formulas are one more source of the algorithms instability in the statistical modeling. As seen in **Table 2** the value of integral $I(f)$ is successfully estimated, but the empirical estimations of $\hat{\sigma}$ are sufficiently different from the theoretical value σ. This result is explained by the following: $\mathcal{E}\eta^4 = \infty$, $I(f/p) = \infty$, $I\left(f^2/p^2\right) = \infty$. Taking $p(x_i) = 2, 6x_i^{1,6}$ we obtain $\sigma \approx 1, 223 \cdot 10^{-3}$, $I(f/p) \approx 0, 67556$, the finite value of $\mathcal{E}\eta^4$, and $I\left(f^2/p^2\right) = \infty$. Although the last estimation is infinite, but **Table 3** shows that both values $I(f)$ and $\hat{\sigma}$ are successfully calculated. The value of $\hat{\sigma}$ is very close to σ.

3. Integrals of very large dimensions

We are coming now to the question of calculation of an integral

$$I(f) = \int_0^\infty \dots \int_0^\infty e^{-(x_1 + x_2 + \dots + x_s)} dx_1 dx_2 \dots dx_{10} = 1 \tag{20}$$

with the distribution density $p(x) = \lambda^s e^{-\lambda(x_1 + x_2 + \dots + x_s)}$. For $\lambda \geq 2$, the estimation variance η is infinity. For $0 < \lambda < 2$, the variance will be finite. For $\lambda > 1$, we obtain $I(f/p) = \infty$ and $I\left(f^2/p^2\right) = \infty$. However, as seen in **Table 4**, the results of calculations for $N = 10000$ allow us to make the conclusion below. If we have the pseudorandom generator of the high quality and a good $p(x)$ then we can calculate the very high dimensional integrals.

N	$\hat{I}$	$\hat{\sigma}$
1,000,000	$9,769 \cdot 10^{-4}$	$1,212 \cdot 10^{-3}$
9,000,000	$9,760 \cdot 10^{-4}$	$1,220 \cdot 10^{-3}$
81,000,000	$9,766 \cdot 10^{-4}$	$1,223 \cdot 10^{-3}$

Table 3.
The results of numerical calculations for the integral Eq. (19) with the density $p(x_i) = 2, 6x_i^{1,6}$.

λ	s	$\hat{I}$	$\hat{\sigma}$	σ
1,01	1000	0,999	0,320	0,324
1,01	10,000	0,995	1,39	1,31
1005	20,000	1004	0,807	0,805
1003	40,000	0,994	0,649	0,658
1001	80,000	0,991	0,605	0,661

Table 4.
The results of numerical calculations for the integral Eq. (20).

4. Special integrals

Let us consider the following class of the integrals:

$$\int\limits_{[0,+\infty)^s} f(x_1, x_2, \ldots, x_s)\, dx_1 dx_2 \ldots dx_s, \tag{21}$$

We define the behavior of the subintegral function as follows: $f(x_1, x_2, \ldots, x_s)$ to be [label = ()]

1. close to 1 in the cube $[0, a]^s$ as $0 < a < 1$;

2. much less than unit as $a < x_i < b, \ \ b \geq a$;

3. equil to zero as $b \leq x_i < \infty$.

Note that very often the integration of functions can be reduced to the linear combination of the integrals similar to Eq. (21) using various replacements of variables.

Below, let us perform a theoretical and numerical analysis how to integrate a model function from our class. The model function is assumed to be $f \equiv 1$ as $0 \leq x_i \leq a$, otherwise $f = 0$. We take both distribution densities set $p_1(x_i) = \lambda e^{-\lambda x_i}, \ \ 0 \leq x_i < \infty$ and $p_2(x_i) = (\omega + 1)(1 - x_i)^{\omega}, \ \ 0 \leq x_i < 1$ to be examined. Our goal is to determine what of two densities provides the best accuracy of the integral computation with given model function.

If we simulate a random point $\xi = (\xi_1, \ldots, \xi_s)$ with densities $p_1(x_i) = \lambda e^{-\lambda x_i}$ then the integral estimation is given by

$$\begin{aligned} \eta_s &= \prod_{i=1}^{s} \lambda e^{\lambda \xi_i}, \\ \mathbf{var}\eta_s &= \mathcal{E}\eta_s^2 - (\mathcal{E}\eta_s)^2 = \left[\frac{1}{\lambda}\int\limits_0^a e^{\lambda x} dx\right]^s - a^s = \left[\frac{e^{\lambda a} - 1}{\lambda^2}\right]^s - a^s. \end{aligned} \tag{22}$$

Testing the variance $\mathbf{var}\eta_s$ for the extremum over λ we get the minimum condition

$$\lambda a e^{\lambda a} - 2e^{\lambda a} + 2 = 0. \tag{23}$$

Let A be λa; then the equation Eq. (23) is reduced to

$$Ae^A - 2e^A + 2 = 0. \tag{24}$$

The equation above has the unique root at $A \approx 1,593620$. It follows that $\lambda_{min} = A/a$. For such λ_{min} the relative error with the 3σ rule is given by

$$\frac{3\sigma}{\sqrt{N}a^s} = 3\left(\frac{\left(e^A - 1\right)^s}{A^{2s}} - 1\right)^{0,5} / \sqrt{N}. \tag{25}$$

Suppose $N = 9 \cdot 10^6, \ s = 10, \lambda = A/a$; then the theoretical value of the relative error is approximately $8,72 \cdot 10^{-3}$. The numerical estimation of the relative error is approximately $8,69 \cdot 10^{-3}$ as $a \in [0,1; 0,001]$. Thus, the numerical estimation of one gives a good fit to the predicted value over a wide range of a.

Now we discuss the use of the density $p_2(x_i) = (\omega + 1)(1 - x_i)^{\omega}$. First, we estimate the second moment of a random value η_s:

$$\mathcal{E}\eta_s^2 = \left[\frac{1 - (1 - a)^{1-\omega}}{1 - \omega^2}\right]^s. \tag{26}$$

The parameter ω is chosen to be A/a; then the expression above is rewritten as follows

$$\mathcal{E}\eta_s^2 = \left[\frac{1}{1 - A^2/a^2}\left(1 - (1 - a)^{1-A/a}\right)\right]^s. \tag{27}$$

Let us consider $\mathcal{E}\eta_s^2$ as $a \to 0$:

$$\begin{aligned} \lim_{a\to 0} \mathcal{E}\eta_s^2 &= \lim_{a\to 0}\left[\frac{a^2}{a^2 - A^2}\left(1 - (1 - a)^{1-A/a}\right)\right]^s = \\ &= -\lim_{a\to 0}\left[\frac{a^2}{A^2}\left(1 - (1 - a)(1 - a)^{-A/a}\right)\right]^s = -\lim_{a\to 0}\left[\frac{a^2}{A^2}\left(1 - (1 - a)\left[(1 - a)^{-1/a}\right]^A\right)\right]^s = \\ &= -\lim_{a\to 0}\left[\frac{a^2}{A^2}\left(1 - (1 - a)e^A\right)\right]^s = -\lim_{a\to 0}\left[\frac{a^2}{A^2}\left(1 - e^A\right)\right]^s \sim \left[\frac{\left(e^A - 1\right)}{A^2}a^2\right]^s. \end{aligned} \tag{28}$$

Comparison between Eqs. (22) and (28) allows to make the following conclusion. If w is chosen to be A/a then the asymptotics of variances, as $a \to 0$, are the same in the densities set of $p_1(x_i), p_2(x_i)$. In the numerical simulation the relative accuracy of $\approx 8,68 \cdot 10^{-3}$ is reached as $N = 9 \cdot 10^6, \ s = 10, \omega = A/a, a \in [0,01; 0,001]$.

Let us turn now to the integral

$$I(f) = \int\limits_{[0,1]^{10}} f(x_1, x_2, \ldots, x_{10})\, dx_1 dx_2 \ldots dx_{10}, \tag{29}$$

where

$$f(x_1, x_2, \ldots, x_{10}) = \begin{cases} 1, & 0 \le x_i \le 1/4, \\ 0, & \text{otherwise}. \end{cases} \tag{30}$$

We put $p(x) \equiv 1$; then $I(f) = I(f/p) = I(f^2/p) = (1/4)^{10} \approx 9,5367 \cdot 10^{-7}$. For $N = 640000$ all realizations are turned out to be equal to zero, i.e., $q_{640000}(x) = 0$ as $0 \le x_i \le 1/4$. In this case, we have

$$\int\limits_{[0,1]^{10}} p(x) \frac{f(x)}{p(x)} dx - \int\limits_{[0,1]^{10}} q_{640000} \frac{f(x)}{p(x)} dx = (1/4)^{10} - 0 = (1/4)^{10}. \tag{31}$$

In accordance with N, the realizations numbers of η_i are turned out to be equal to 1 as $N = 810000$; equal to 3 as $N = 4000000$; equal to 15 as $N = 16000000$.

We now take $p(x_i) = (\omega + 1)(1 - x_i)^{\omega}$ in the unit cube $[0, 1]^{10}$ $(\omega > 1)$. Such choice provides the gross realizations of points in $[0, 1/4]^{10}$ and as consequence, we get benefit in quality of random values (simultaneously, we have decrease of the estimation variance, and as consequence decrease of the statistical error with the 3σ rule.) **Table 5** shows the calculations results for $N = 9000000$. Note that $\hat{\sigma}$ reaches the minimum as $\omega = 5$. In this case, we have $\hat{p}(x) = 6(1 - x)^5$, $I(f/\hat{p}) = I(f^2/\hat{p}) \approx 3,49 \cdot 10^{-11}$. Making the more detailed research for both $\omega = 5$ and the theoretical value $\sigma \approx 5,83 \cdot 10^{-6}$ we get the results represented in **Table 6**. If the function

w	$\hat{I}$	rule "3σ"
3	$9,53 \cdot 10^{-7}$	$8,61 \cdot 10^{-9}$
4	$9,58 \cdot 10^{-7}$	$6,36 \cdot 10^{-9}$
5	$9,55 \cdot 10^{-7}$	$5,82 \cdot 10^{-9}$
6	$9,55 \cdot 10^{-7}$	$6,49 \cdot 10^{-9}$
7	$9,51 \cdot 10^{-7}$	$8,06 \cdot 10^{-9}$

Table 5.
The results of numerical calculations for the integral Eq. (29) at various ω values.

N	$\hat{I}$	$\hat{\sigma}$
640,000	$9,61 \cdot 10^{-7}$	$6,15 \cdot 10^{-6}$
810,000	$9,56 \cdot 10^{-7}$	$6,00 \cdot 10^{-6}$
1,000,000	$9,51 \cdot 10^{-7}$	$5,91 \cdot 10^{-6}$
4,000,000	$9,52 \cdot 10^{-7}$	$5,74 \cdot 10^{-6}$
16,000,000	$9,55 \cdot 10^{-7}$	$5,84 \cdot 10^{-6}$

Table 6.
The results of numerical calculations for the integral Eq. (29) at $\omega = 5$.

$f(x_1, x_2, \dots, x_{10})$ close to some constant in $[0, 1/4]^{10}$ and small out of this interval then we can advise to use $\hat{p} = 6(1-x)^5$ to calculate the integral in $[0, 1]^{10}$.

5. Moments calculation

We are now concerned with the following issue: to find the kth moments of a random value τ with the distribution density $p(x)$:

$$\mathcal{E}\tau^k = \int_a^b x^k p(x)\,dx. \tag{32}$$

In fact we have realizations of the random value ξ with a distribution density $q(x)$. With $p(x)$ replaced by $q(x)$ in Eq. (32) we get an error

$$\int_a^b x^k p(x)\,dx - \int_a^b x^k q(x)\,dx = \int_a^b x^k [p(x) - q(x)]\,dx. \tag{33}$$

Suppose $b = \infty$ and ξ_{max} are the maximum value of the random variable over the all realizations for fixed N; then value of ξ_{max} gives shift $\int_{\xi_{max}}^{\infty} x^k p(x)\,dx$ that increases both monotonically and without limit. The condition $q(x) = 0$ as $x > \xi_{max}$ determines the lower limit of the last integral.

Many solutions of the boundary-value problems for the elliptic and parabolic Equations [4, 5] have a form of the expectations for the random value moments. Meaning of these expectations is the first exit time of the Wiener process trajectories to the domain boundary.

Let a domain be the three-dimensional ball with the radius $r = 1$ and the Wiener trajectories start from the ball center; then a function of distribution of the first exit time for the Wiener trajectory is, in particular, given by [5].

$$F(t) = 1 + 2\sum_{k=1}^{\infty} (-1)^k \exp\left(-k^2\pi^2 t/2\right), \quad t \in [0, +\infty). \tag{34}$$

From the above, we obtain the distribution density:

$$p(t) = 2\sum_{k=1}^{\infty} (-1)^{k+1} \mu k^2 \exp\left(-\mu k^2 t\right), \quad \mu = \pi^2/2. \tag{35}$$

Assuming τ is distributed with this density and calculating the expectation of the kth moment we get

$$\mathcal{E}\tau^k = \int_0^{\infty} t^k p(t)\,dt. \tag{36}$$

In **Table 7**, we put the calculations results for $N = 1000000$. The kth moment expectation can be represented in a form.

$$\mathcal{E}\tau^k = \int_0^\infty q(x) t^k \frac{p(t)}{q(t)} dt, \tag{37}$$

where $q(t)$ is some density in $[0, \infty)$. Taking $q(t) = \lambda \exp(-\lambda t)$, for $\lambda = \pi^2/2$ we get that the number of realizations $\xi_i > 1$ will be almost twice as small as in the case of the modeling with the original $p(t)$. In this situation we should obtain degradation of the estimation for the high moments. The calculations results with $q(t)$ for $N = 1000000$ are represented in **Table 8**. However, in realizations at a computer we get the obvious

Moment	Simulation	Theory
1	$3,304 \cdot 10^{-1}$	$3,333 \cdot 10^{-1}$
2	$1,553 \cdot 10^{-1}$	$1,556 \cdot 10^{-1}$
3	$9,848 \cdot 10^{-2}$	$9,841 \cdot 10^{-2}$
4	$8,076 \cdot 10^{-2}$	$8,063 \cdot 10^{-2}$
5	$8,291 \cdot 10^{-2}$	$8,193 \cdot 10^{-2}$
6	$9,843 \cdot 10^{-2}$	$9,969 \cdot 10^{-2}$
7	$1,319 \cdot 10^{-1}$	$1,414 \cdot 10^{-1}$
8	$2,070 \cdot 10^{-1}$	$2,293 \cdot 10^{-1}$
9	$4,518 \cdot 10^{-1}$	$4,182 \cdot 10^{-1}$
10	$7,286 \cdot 10^{-1}$	$8,474 \cdot 10^{-1}$
11	$9,021 \cdot 10^{2}$	$4,251 \cdot 10^{3}$
12	$1,183 \cdot 10^{3}$	$1,637 \cdot 10^{4}$
13	$8,389 \cdot 10^{3}$	$6,634 \cdot 10^{4}$

Table 7.
The results of numerical calculations for the moments by the first way.

Moment	Simulation
5	$8,188 \cdot 10^{-2}$
6	$1,013 \cdot 10^{-1}$
7	$1,468 \cdot 10^{-1}$
8	$2,247 \cdot 10^{-1}$
9	$3,950 \cdot 10^{-1}$
10	$8,283 \cdot 10^{-1}$
18	$2,833 \cdot 10^{3}$
19	$1,056 \cdot 10^{4}$
20	$5,118 \cdot 10^{4}$

Table 8.
The results of numerical calculations for the moments by the second way.

improvement in quality of the moments estimation for all k from 5 to 20. Consider the choice of modeling strategy with regards to the variance. Suppose ξ and η be estimations of the statistical modeling for a value J, i.e., $\mathcal{E}\xi = \mathcal{E}\eta = J$ with the variances of $\sigma_1^2(\xi), \sigma_2^2(\eta)$ and the realizations of $\xi = (\xi_1 + \ldots + \xi_N)/N$, $\eta = (\eta_1 + \ldots + \eta_N)/N$. It would seem that for $\sigma_1(\xi) < \sigma_2(\eta)$ the real estimation of ξ will be occurred close to the origin value of J. But this statement does not need to be always true. Without loss of generality it can believed that $J = 0$. Additionally, if N is large enough then ξ, η are chosen be normal random variables with $N(0, \sigma_1)$ and $N(0, \sigma_2)$, respectively. The following theorem holds.

Proposition 6. Let ξ, η be normal random variables, and $\xi \sim N(0, \sigma_1), \eta \sim N(0, \sigma_2)$ then $P(\xi| > |\eta|) = \frac{2}{\pi} \arctan \frac{\sigma_1}{\sigma_2}$.

Proof:

$$P|\xi > |\eta| = \frac{1}{\sigma_1\sqrt{2\pi}} \int_{-\infty}^{0} e^{-\frac{y^2}{2\sigma_1^2}} dy \left\{ \frac{1}{\sigma_2\sqrt{2\pi}} \int_0^{y} e^{-\frac{x^2}{2\sigma_2^2}} dx + \frac{1}{\sigma_2\sqrt{2\pi}} \int_0^{|y|} e^{-\frac{x^2}{2\sigma_2^2}} dx \right\} +$$
$$+ \frac{1}{\sigma_1\sqrt{2\pi}} \int_0^{\infty} e^{-\frac{y^2}{2\sigma_1^2}} dy \left\{ \frac{1}{\sigma_2\sqrt{2\pi}} \int_0^{y} e^{-\frac{x^2}{2\sigma_2^2}} dx + \frac{1}{\sigma_2\sqrt{2\pi}} \int_0^{|y|} e^{-\frac{x^2}{2\sigma_2^2}} dx \right\} =$$
$$= \frac{2}{\sigma_1\sqrt{2\pi}} \int_0^{\infty} e^{-\frac{y^2}{2\sigma_1^2}} dy \left\{ \frac{2}{\sigma_2\sqrt{2\pi}} \int_0^{y} e^{-\frac{x^2}{2\sigma_2^2}} dx \right\} = \frac{4}{2\pi\sigma_1\sigma_2} \int_0^{\infty} e^{-\frac{y^2}{2\sigma_1^2}} dy \cdot \int_0^{y} e^{-\frac{x^2}{2\sigma_2^2}} dx.$$

Using Taylor expansion

$$e^{-\frac{y^2}{2\sigma_2^2}} = \sum_{n=0}^{\infty} (-1)^n \frac{y^{2n}}{2^n \sigma_2^{2n} n!},$$

we get

$$\int_0^{y} \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{2^n \sigma_2^{2n} n!} dx = \sum_{n=0}^{\infty} (-1)^n \frac{y^{2n+1}}{(2n+1)2^n \sigma_2^{2n} n!},$$
$$\int_0^{\infty} \sum_{n=0}^{\infty} (-1)^n \frac{y^{2n+1}}{(2n+1)2^n \sigma_2^{2n} n!} e^{-\frac{y^2}{2\sigma_1^2}} dy = \sum_{n=0}^{\infty} (-1)^n \frac{1}{(2n+1)2^n \sigma_2^{2n} n!} \int_0^{\infty} y^{2n+1} e^{-\frac{y^2}{2\sigma_1^2}} dy =$$
$$= \sum_{n=0}^{\infty} (-1)^n \frac{1}{(2n+1)2^n \sigma_2^{2n} n!} \cdot \frac{2^{n+1}\sigma_1^{2n+2} n!}{2}. \quad (*)$$

Note that the last equality is obtained with the help of the formula:

$$\int_0^{\infty} x^{2n+1} e^{-px^2} dx = \frac{n!}{2p^{n+1}}, \;\; p > 0.$$

In our case, p is $\frac{1}{2\sigma_1^2}$. We continue the equalities chain which is broken at (*):

Moment	Simulation
1	$3,297 \cdot 10^{-1}$
2	$1,556 \cdot 10^{-1}$
3	$9,842 \cdot 10^{-2}$
4	$8,066 \cdot 10^{-2}$
5	$8,194 \cdot 10^{-2}$
6	$9,968 \cdot 10^{-2}$
7	$1,414 \cdot 10^{-1}$
8	$2,293 \cdot 10^{-1}$
9	$4,181 \cdot 10^{-1}$
10	$8,474 \cdot 10^{-1}$
18	$4,251 \cdot 10^{3}$
19	$1,637 \cdot 10^{4}$
20	$6,634 \cdot 10^{4}$

Table 9.
The results of numerical calculations for the moments by the third way.

$$(*) = \sum_{n=0}^{\infty} (-1)^n \frac{1}{(2n+1)} \frac{\sigma_1^{2n+2}}{\sigma_2^{2n}} = \sum_{n=0}^{\infty} (-1)^n \frac{\sigma_1^2}{(2n+1)} \left(\frac{\sigma_1}{\sigma_2}\right)^{2n} =$$
$$= \frac{2}{\pi} \sum_{n=0}^{\infty} (-1)^n \frac{(\sigma_1/\sigma_2)^{2n+1}}{2n+1} = \frac{2}{\pi} \arctan \frac{\sigma_1}{\sigma_2}.$$

For the k-moment calculation we take

$$q_k(t) = \frac{\lambda^{k+1} t^k e^{-\lambda t}}{k!}, \lambda = \frac{\pi^2}{2}$$

and get the results shown in **Table 9**.

6. Integral from the regression function

Now, we consider the issue of calculation of an integral

$$\int_D f(x)dx, \qquad (38)$$

where the function $f(x)$ has no an analytical expression. Suppose there exists a random variable $\xi(x,w)$ such that its expectation is equals to $\mathcal{E}\xi(x,w) = f(x)$ for some fixed x. The random variable $\xi(x,w)$ may be realized neither as result of the physical measurements or some calculations (e.g., using the modeling statistical method). In this case the optimal density is given by [6].

$$p(x) = \frac{f(x)}{\sqrt{d(x) + \lambda}}, \tag{39}$$

where $d(x)$ is the variance of the random variable $\xi(x, w)$. Note that one should use the optimal density from [1] if complexity in calculations (experimental measurements) is much different from each other for any x. We determine the parameter λ from the condition $\int\limits_D f(x)/\sqrt{d(x)+\lambda}dx = 1$. Really in practice, we find a priori or a posteriori approaches to both $f(x)$ and $d(x)$. By $\bar{f}(x)$ and $\bar{d}(x)$ denote these approaches. Then the approach to the optimal $p(x)$ will look like

$$\overline{p}(x) = \frac{\bar{f}(x)}{\sqrt{\bar{d}(x) + \bar{\lambda}}} \text{ and } \int\limits_D \frac{\bar{f}(x)}{\sqrt{\bar{d}(x) + \bar{\lambda}}} dx = 1. \tag{40}$$

The parameter λ is often turned out to be find enough complicity [6]. If the domain D is the interval $[0, H]$ for small H then it is suppose to use the quasioptimal density $\overline{p}(x)$.

Example 3. We now consider the following issue: Suppose $f(x) = x$, $d(x) = 1/x$, $D = [0, H]$. The optimal density is given by

$$p(x) = \frac{x}{\sqrt{1/x + \lambda}} = \frac{x\sqrt{x}}{\sqrt{\lambda\sqrt{x} + 1}} \sim c \cdot x^{3/2}. \tag{41}$$

We take the quasioptimal density in the form $\overline{p}(x) = 5H^{5/2}x^{3/2}/2$. In this case, for $p(x) \equiv 1$ the estimation variance of the random value $\eta = f(x, w)/p(x)$:

$$\mathbf{var}\eta = \int\limits_0^H d(x)p(x)dx + \int\limits_0^1 \frac{f^2}{p(x)}dx - I^2 \tag{42}$$

is equals to ∞. Taking $p(x) = 2x/H^2$ we get $\mathbf{var}\eta = 2/H$. But if the function $f(x)$ was precisely known for the same density $p(x)$ then $\mathbf{var}\eta = 0$. If we choose the quasioptimal density $\overline{p}(x) = 5H^{5/2}x^{3/2}/2$ then the estimation variance of η is equals to $17H^4/12 + 4/(15H)$. For $H \to 0$ the variance behaves approximately as $4/(15H)$. It is much the better than $2/H$. For $H = 1$ the estimation variance with the density $p(x) = 2x/H^2$ is equals to 2, and the estimation variance with the quasioptimal density $\overline{p}(x)$ is equals to 101/60.

Suppose we practically realize calculation of the integral Eq. (38) with $d(x) = 1/x$; then one should discard the interval $[0, \delta]$ and to calculate $\int_\delta^H f(x)dx$ because of the values $|\xi(x, w)|$ can be the intolerably large. Also one should replace $f(x)$ by $\hat{f}(x)$:

$$\hat{f}(x) = \begin{cases} 0, & 0 \le x \le \delta, \\ x, & \delta < x \le H. \end{cases} \tag{43}$$

The shift is $\int\limits_0^\delta x dx = \delta^2/2$ and choosing $\delta \sim 1/\sqrt[4]{N}$ we get the total error $\delta^2/2 + 3\sigma/\sqrt{N}$ of oder $O(1/\sqrt{N})$.

In applications the estimation variance for the integral functionals (e.g., field flow calculation neither across the arc or the surface) from the solutions of the boundary-value problems for both the linear [1] or nonlinear [2] elliptic equations is of interest. For the above variance is $d(x) \sim B/x^2,\ f(x) \approx a_0 + a_1x + a_2x^2 + \ \dots$, where x is the distance to the domain boundary. Suppose $f(x) \approx a_1x + a_2x^2 + \ \dots$; then the optimal density is given by

$$p(x) = \frac{a_1x + a_2x^2 + \ \dots}{\sqrt{B/x^2 + \lambda}}. \tag{44}$$

The quasioptimal density has the form $\overline{p}(x) = 3x^2/H^3$ for small H in $[0, H]$. In applications, this case is of our main interest. Taking $\delta = 1/\sqrt[4]{N}$ like in the Example 3 we get the asymptotics of decrease for the total error as $O\left(1/\sqrt{N}\right)$.

Suppose $d(x) \sim B/x^2,\ f(x) \approx a_0 + a_1x + a_2x^2 + \ \dots$, and $a_0 \neq 0$ then there is no density kind of $\overline{p}(x) = (w+1)x^w,\ x \in [0, H]$ with the finite variance. The density $p(x) = |f(x)|/\sqrt{d(x) + \lambda}$ will be give the estimation with the infinity variance. Instead of calculation of the integral $\int_0^H f(x)dx$ we will be calculate the integral $\int_\delta^H f(x)dx$. For this integral we already can choice the quasioptimal density with the finite variance of the estimation: $\overline{p}(x) = 2x/\left(H^2 - \delta^2\right)$. For $\delta \sim O\left(\ln N/\sqrt{N}\right)$ the total error will have the asymptotics $O\left(\ln N/\sqrt{N}\right)$.

Example 4. Suppose $d(x) = 1/x^2,\ f(x) = 1,\ \ H = 1;$ then the asymptotics of the variance with the quasioptimal density has kind of $(-2, 5 \cdot \ln \delta)$.

In conditions of Example 4, choice of the optimal density in the form

$$p(x) = \frac{x}{\sqrt{1 - \delta^2}\sqrt{1 - x^2}},\ \ x \in [\delta, 1] \tag{45}$$

yields the following result: the estimation variance will have asymptotics $(-2 \cdot \ln \delta)$ for $\delta \to 0$.

Remark. If we know that a value of $f(x)$ in the interval $[0, \delta]$ close to the number f_0, then in Eq. (43) to use

$$\hat{f}(x) = \begin{cases} f_0, & 0 \le x \le \delta, \\ x, & \delta < x \le H, \end{cases} \tag{46}$$

more efficiently and also to take $\int_0^H f(x)dx \approx f_0\delta + \int_\delta^H f(x)dx$.

7. Conclusion

In the paper we describe the sufficient conditions of the stable calculations for the multidimensional integrals by the Monte Carlo method. We get the results of numerous numerical computations using the mt19937 pseudorandom number generator. The article results can be also useful in the practical solution of the boundary value problem, for both the elliptic and parabolic equations. The earlier suggested approach to the optimal choice of the density [1, 6] often needs to solve a complicated secondary task. In the paper we suggest the approach to choice of the quasioptimal densities that is of considerable interest in applied problems solution.

Author details

Alexander A. Kronberg*† and Tatiana K. Kronberg†
Dubna, Russia

*Address all correspondence to: kronberg.alexander@gmail.com

† These authors contributed equally.

References

[1] Kronberg AA. On the numerical finding of some functionals on solving of Laplace's equation. Izv. Vuz. Matematika. 1987;**31**:30-36

[2] Kronberg AA. Algorithms for solving elliptic boundary value problems. UUSR Computational Mathematics and Mathematical Physics. 1982;**22**:122-131. DOI: 10.1016/0041-5553(82)90103-3

[3] Matsumoto M, Nishimura T. Mersenne twister: A 623-dimensionally equidistributed uniform preudorandom number. ACM Transactions on Modeling and Computer Simulation (TOMACS). 1998;**8**:G1-G25. DOI: 10.1145/272991. 272995

[4] Dynkin EB. Markov Processes. Vol. 2. Berlin: Springer-Verlag; 1965. p. 274

[5] Haji-Sheiki A, Sparrow EM. The solution of heat conduction problems by probability methods. Transmission ASME Series C: Journal of Heat Transfer. 1967;**89**:121-131. DOI: 10.1115/1.3614330

[6] Kronberg AA. On the numerical solution of Laplace's equation in unbounded domain. Izv. Vuz. Matematika. 1984;**28**:45-48