Volume: 1 2015 Issue: 3 Pages: 19-22 ISSN 1339-9640

USAGE OF SIMULATION FOR INTEGRALS CALCULATION

Gabriela Ižaríková
TU of Košice, Faculty SjF, Institute of Special Technical Sciences, Department of Applied Mathematics and Informatics, Letná 9, 042 00 Košice, gabriela.izarikova@tuke.sk

Keywords: simulation, Monte Carlo method, geometric method, estimation of certain characteristics
Abstract: The article deals with the simulation, more specifically with the method Monte Carlo. The term as a simulation, simulate, a simulator are well known in many scientific disciplines. The simulation model allows performing a number of experiments, analysing them, evaluating, optimizing and afterwards applying the results to the real system. By using the Monte Carlo method it is possible to compute one-dimensional and multi-dimensional integral. Monte Carlo method can be divided according to procedure solutions to geometrical methods and calculation based on the estimation of certain characteristics of random variable.

1 Introduction

Simulation is a method designed to mimic the real system. It is used when an analysis is mathematically difficult or expensive (there is no analytical solution), or in case where is not possible to made a real experiment. In order to simulate the values of the random variables are replaced by generating a large number of the implementation of the random variable and these are statistically processed. One method is Monte Carlo simulation through which it is possible to determine the value of the number π and compute one-dimensional and multi-dimensional integral, it is also applicable in estimating the development of finance and so on. Monte Carlo method can be divided according to procedure solutions to geometrical methods and calculation based on the estimation of certain characteristics of random variable.

2 Monte Carlo method

Geometric algorithm method is based on the fact that we want to calculate the content formation Ω, which lies inside the unit squares (Figure 1). In the square we choose n random points evenly distributed and see how many of them fall into the unit, the number denoted m. According to the theory of geometric probability it is then content formation Ω is approximately equal ratio m/n. When calculating the area of the process is that it is generated independently from the implementation $(x_i, y_i)\cdots(x_n, y_n)$ of the uniform distribution $(0,1)\, x\, (0,1)$ if true $(x_i, y_i)\in\Omega$ then the random variable $\varphi_i = 1$ otherwise $\varphi_i = 0$. Referred to below $m=\sum_{i=1}^{n}\varphi_i$ in which m has the binomial distribution with $n, p \approx Bi(n,p)$, with means $m=\sum_{i=1}^{n}\varphi_i$, $E(m)=np$ and dispersion $D(m)=np(1-p)$. Estimate of the area Ω is then approximately equal ratio m/n.

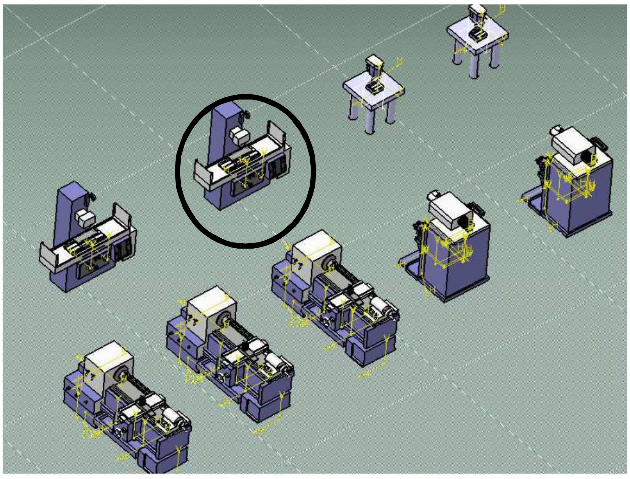

Figure 1 Example of area inside the unit squares

The calculation is based on the estimation of certain characteristics of the random variable is that the need to calculate the value of some unknown z. We model such a random variable φ to which is applied $E(\varphi)=z$ (mean random variable is equal to the unknown value). Furthermore, we expect the implementation of n independent $x_1, x_2, \ldots x_n$ random variables $\varphi_1, \varphi_2, \ldots \varphi_n$ with the same probability distribution as φ. Unknown value z is estimated using the arithmetic average $z=\frac{1}{n}\sum_{i=1}^{n}x_i$. We obtain profit in a suitable transformation of random numbers generated by the search value $x_1, x_2, \ldots x_n$.

Volume: 1 2015 Issue: 3 Pages: 19-22 ISSN 1339-9640

General scheme of Monte Carlo:

- Generating random numbers x_i with uniform distribution on the interval $(0,1)$
- Transformation of random numbers x_i to random numbers z_i to the necessary distribution.
- Calculation of estimates of characteristics of random variables X by random numbers z_i
- Statistical treatment of results.

Monte Carlo calculation method is based on a random process modelling and data-processing statistical methods. To achieve the necessary precision, it is necessary to repeat the simulation many times. At the same time the estimated value of the unknown is important to determine the accuracy of the estimate.

Accuracy of estimates may be determined as follows:

- Searched the unknown value X estimated by the implementation of the random variable $\overline{X}$ while $X \approx \overline{X}$. Unknown value is estimated using the arithmetic average, $\overline{X} = \frac{1}{N}\sum_{i=1}^{N} X_i$ which $X_1, X_2, \ldots X_n$ are independent random variables with means X and with the dispersion σ^2.
- The accuracy of approximate $X \approx \overline{X}$ has equality ε with the reliability α if the inequality $|X - \overline{X}| < \varepsilon$ is valid: $P(|X - \overline{X}| < \varepsilon) = \alpha$.
- According to the central limit theorems a random variable $\overline{X}$ asymptotically (for $N \to \infty$) a normal distribution with mean $E(\overline{X}) = X$ and variance $D(\overline{X}) = \frac{\sigma^2}{N}$, $P\left(\frac{|X - \overline{X}|}{\sigma^2}\sqrt{N} < u_\alpha\right) = \alpha$ where u_α the critical value of the standard normal distribution ($u_{0,95} = 1{,}644854$).
- For the estimate of the accuracy is valid: $\varepsilon = u_\alpha\sqrt{\frac{\sigma^2}{N}}$ ($N = u^2{}_\alpha\left(\frac{\sigma}{\varepsilon}\right)^2$ - number of attempts needed to achieve the necessary precision ε the level of significance α).
- If a random variable m has a binomial distribution with parameters n and p, with a mean value $E(m) = np$ and dispersion $D(m) = np(1-p)$ so the true size of the error $P\left(\frac{|X - \overline{X}|}{\sigma^2}\sqrt{N} < u_\alpha\right) = \alpha$. The accuracy of the estimate is valid $\varepsilon = u_\alpha\sqrt{\frac{p(1-p)}{n}}$.

2.1 Integral estimate using geometric methods

Estimate of the integral $\int_a^b f(x)dx$ of geometric methods we mean the detection of the equals (1) the area under the curve $f(x)$ in which (a,b), where $f(x)$ the density, to determine the probability that the random value X falls into the range (a,b).

$$P(a < X < b) = p = \int_a^b f(x)dx \qquad (1)$$

Using the Monte Carlo method first generate random numbers x_i from the distribution which has a density $f(x)$ then we determine the number of attempts (m) which are valid $a \le x_i \le b$ According to Bernoulli's sentence for sufficiently large n can take the integral estimate for the relative number m/n.

In calculating the difference whether integrated functions is implemented probability density variable or not.

a) If $f(x)$ is the probability density of a random variable: $\int_{0,3}^{0,7} e^{-x}dx$.

First, they generate random numbers x_i with exponential distribution (e^{-x} - density of a random variable with exponential distribution), and then determine the number of attempts *(m)* which are valid $a \le x_i \le b$. After conversion of attempts n is the approximate value of the integral m/n. The exact value of the integral is 0,244233 the approximate value of the integral depends on the number of attempts and transferred accuracy of the estimate are shown in Table 1. The accuracy of the estimate can be determined from the relation $\varepsilon = u_\alpha\sqrt{\frac{p(1-p)}{n}}$, where *p* is the exact value of the integral.

Table 1 The approximate value of the integral $\int_{0,3}^{0,7} e^{-x}dx$

n	10	50	100	500	1000
I	0,2	0,21	0,25	0,228	0,235
ε	0,0158	0,00303	0,00152	0,0003	0,00015

b) If $f(x)$ is not the probability density of a random variable $\int_0^1 f(x)dx$ whichever is valid $0 \le f(x) \le 1$

that the value of the integral equals the area under the curve $f(x)$. In this case, a pair generate random numbers (x_i, y_i) from a uniform interval $(0,1)\,x\,(0,1)$ and determine the number of attempts (m) which are valid $y_i < f(x)$. After conversion of attempts n is the approximate value of the integral m/n. The exact value is 0,310268 for the integral $\int_0^1 \sin(x^2)dx$ the approximate value of the depends on the number of attempts and transferred accuracy of the estimate are shown in Table 2.

Table 2 The approximate value of the integral $\int_0^1 \sin(x^2)dx$

n	10	50	100	500	1000
I	0,2	0,3	0,34	0,292	0,306
ε	0,0176	0,003520	0,00176	0,000352	0,000176

If $f(x)$ is not the probability density of a random variable and the limits of the integrals are a,b: $\int_a^b f(x)dx$ is necessary to make the substitution (2):

$x = a + (b-a)z$,

$$\int_a^b f(x)dx = (S-i)(b-a)\int_0^1 g(z)\,dz + (b-a)i \qquad (2)$$

Where S is a supremum and i infimum function $f(x)$ of the interval (a,b), $g(z) = \frac{f[a+(b-a)z]-i}{S-i}$ for the function $g(z)$ valid $0 \le g(z) \le 1$. The exact value is 0,467970 for the integral $\int_1^{1,5} \sin(x^2)dx$ the approximate value of the integral depends on the number of attempts and transferred accuracy of estimates are shown in Table 3.

Table 3 The approximate value of the integral $\int_1^{1,5} \sin(x^2)dx$

n	10	50	100	500	1000
I	0,4556	0,46449	0,464496	0,466271	0,465716
ε	0,0205	0,00409	0,00204	0,00041	0,00020

2.2 Integral calculation based on estimates of the mean value of random variable

The calculation is based on the model of such a random variable φ which is valid $E(\varphi) = z$, (mean of random variable is equal to the unknown value). Unknown value is integral (3):

$$z = I = \int_a^b g(x)f(x)dx \qquad (3)$$

First, we generate random numbers x_i from the distribution which has a density $f(x)$ then calculated $y_i = g(x_i)$. According to the law of large numbers in the form of sentences Khinchin a sufficiently large number of trials n can approximate value for the integral of the arithmetic average $g(x_i)$ (4):

$$I = \int_a^b g(x)f(x)dx \approx \frac{1}{n}\sum_{i=1}^{n} g(x_i) \qquad (4)$$

Accuracy of estimates can be determined by the relationship $\varepsilon = u_\alpha \frac{s}{\sqrt{n}}$.

When calculating the integral $J = \int_a^b h(x)dx$ we choose the distribution with density $f(x)$ $(\int_a^b f(x)dx = 1)$ and integral adjusted as follows (5):

$$J = \int_a^b \frac{h(x)}{f(x)} f(x)dx = \int_a^b g(x)f(x)\,dx \qquad (5).$$

If the final boundary of the interval, so we can $f(x)$ take for uniform distribution and density $f(x) = \frac{1}{b-a}$ of the integral holds: $J = (b-a)\int_a^b h(x)\frac{1}{b-a}dx$. In case of improper integrals for example $\int_0^\infty g(x)f(x)dx$ we take $f(x)$ the density exponential distribution with parameter 1, $f(x) = e^{-x}$.

For integral $\int_1^{1,5} \sin(x^2)dx$ is density for uniform distribution $f(x) = \frac{1}{1,5-1}$, generates a random number x_i from a uniform distribution, since the limits of the integral are 1 and 1.5 will transfer the substitution (6):

$$x_i = (1.5-1)RAND(x) + 1 \qquad (6),$$

calculate values $y_i = h(x_i) = \sin(x_i^2)$ and determine their arithmetic average $\overline{y}$. For the integral value true: $\int_1^{1,5} \sin(x^2)dx \approx (1,5-1)\overline{y}$. The exact value of the integral of 0,467970, the approximate value of the integral depends on the number of attempts and transferred accuracy of estimates is shown in Table 4.

Table 4 The approximate value of the integral $\int_1^{1,5} \sin(x^2)dx$

n	10	50	100	500	1000
$\overline{y}$	0,95271	0,92895	0,94041	0,93721	0,93864
I	0,47635	0,46447	0,47020	0,46860	0,46932
ε	0,05851	0,03016	0,02299	0,01051	0,00736

These methods may also be used in the calculation of multidimensional integral.

Conclusion

The article is an example of using the Monte Carlo, demonstrated the possibility of its usage compute one-dimensional and multi-dimensional integral, it is also applicable in estimating the development of finance and so on. Monte Carlo method can be compute one-dimensional integral, divided according to procedure solutions to geometrical methods and calculation based on the estimation of certain characteristics of random variable. The contributions are calculated integrally both methods and determining the accuracy of the estimate. Both methods have comparatively results. The result is more accurate with a count in simulation.

Acknowledgement

This article was created by implementation of the grant project VEGA 1/0708/16: "Development of a new research method for simulation, assessment, evaluation and quantification of advanced methods of production."

References

[1] ANDĚL, J.: Základy matematické štatistiky, MatfyzPress, p. 358, 2011. (Original in Slovak)
[2] BAČA, M., FEŇOVČÍKOVÁ, A.: Mathematics 2, SjF TU Košice, p. 278, 2010.
[3] MALINDŽÁK, D. a kol.: Modelovanie a simulácia v logistike /teória modelovania a simulácie/. Košice:TU-BERG, p. 181, 2009. (Original in Slovak)
[4] NELSON, B. L.: Stochastic Modeling: Analysis and Simulation, McGraw-Hill, 1995.
[5] TREBUŇA P. a kol.: Modelovanie v priemyselnom inžinierstve, TUKE, p. 195, 2015. (Original in Slovak)
[6] STRAKA, M.: Diskrétna a spojitá simulácia v simulačnom jazyku EXTEND [online], Košice: TU FBERG, Edičné stredisko/AMS. 2007. (Original in Slovak)

Review process

Single-blind peer reviewed process by two reviewers