

III.
CAŁKOWANIE METODAMI
MONTE CARLO

Janusz Adamowski

1 Nazwa metody

Podstawowym zastosowaniem w fizyce metody Monte Carlo (MC) jest opis złożonych układów fizycznych o dużej liczbie stopni swobody. Opis złożonych układów fizycznych musi mieć charakter statystyczny (stochastyczny). Bazuje on więc na pojęciach prawdopodobieństwa, rozkładu prawdopodobieństwa oraz generatorach liczb losowych.

Odpowiednie metody opisu nazywane są **metodami Monte Carlo** (od nazwy słynnego kasyna gier w Monaco). Zgodnie z obecną "modą" niektóre spośród metod stochastycznych nazywane są **metodami Las Vegas**. Nazwy te ma na celu podkreślenie przypadkowego (stochastycznego) charakteru metod.

Prawdopodobnie wynalazcą nazwy "metoda Monte Carlo" był polski matematyk Stanisław Ulam.

Obecna klasyfikacja algorytmów stochastycznych:

- (1) **algorytmy Monte Carlo**: szybkie, choć nie zawsze dokładne,
- (2) **algorytmy Las Vegas**: dokładne, choć nie zawsze szybkie.

Klasyfikacja ta jest zbyt ostra i raczej deprecjonująca wartość tych metod.

Większość zastosowań w fizyce tych metod polega na obliczaniu całek wielowymiarowych z użyciem różnych rozkładów statystycznych. W tej części wstępnej zajmijmy się obliczaniem całek metodami Monte Carlo/Las Vegas.

2 Całkowanie funkcji jednej zmiennej

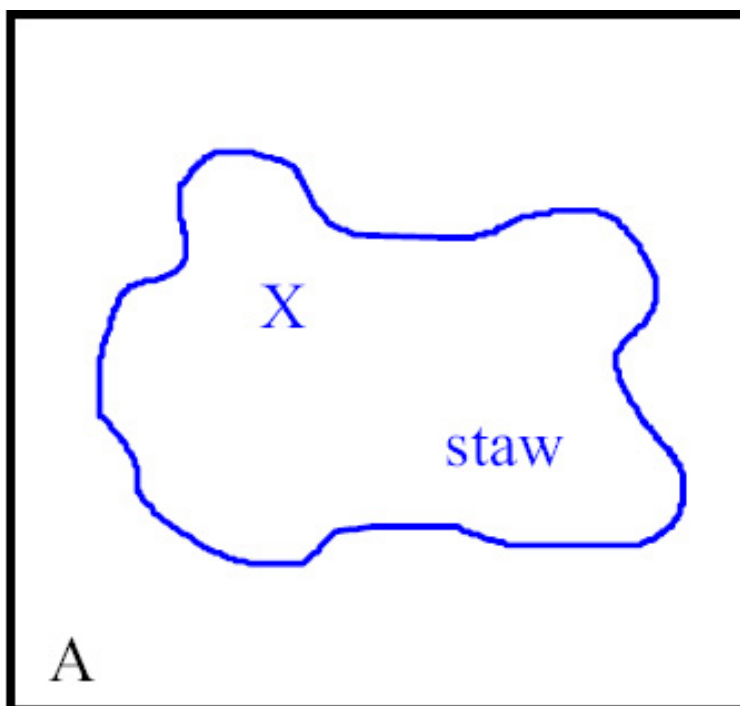
Idea metod MC może być zobrazowana na następującym przykładzie. Powiedzmy, że chcemy wyznaczyć powierzchnię stawu dysponując w tym celu jedynie stosem kamieni. Niech staw o nieznanym polu powierzchni X znajduje się wewnątrz ogrodu, np. w kształcie prostokąta, o znanym polu powierzchni A (Rys. 3.1). Czy można użyć stosu kamieni do oszacowania powierzchni stawu?

Tak, można. Należy w tym celu rzucać kamienie do ogrodu w sposób przypadkowy i liczyć liczbę plusków.

Powierzchnię stawu szacujemy następująco. Jeżeli N jest liczbą rzuconych kamieni, a N_p – liczbą plusków, to

$$\frac{X}{A} = \frac{N_p}{N} \quad (1)$$

Bezpośrednie zastosowanie tego pomysłu do obliczania całek prowadzi do tzw. **metody Monte Carlo chybił–trafił (hit-or-miss Monte Carlo)**.



3.1. Ilustracja idei metody Monte Carlo.

Rys.

3 Metoda Monte Carlo chybił–trafił

Chcemy obliczyć całkę z funkcji $f(x)$ w przedziale $[a, b]$, czyli

$$I = \int_a^b f(x) dx, \quad (2)$$

Niech $f(x)$ w tym przedziale spełnia nierówności

$$c \leq f(x) \leq d.$$

Używamy generatora liczb losowych o rozkładzie jednorodnym i losujemy N par niezależnych od siebie liczb pseudoprzypadkowych (x_i, y_i) takich, że spełnione są następujące nierówności:

$$a \leq x_i \leq b$$

oraz

$$c \leq y_i \leq d.$$

Następnie wyznaczamy liczbę par N_p spełniających warunek

$$y_i \leq f(x_i). \quad (3)$$

Zgodnie z **ideą metody Monte Carlo oszacowaniem Monte Carlo całki** (2), czyli pola pod krzywą $y = f(x)$, jest liczba

$$I_N = \frac{N_p}{N} A + c(b - a), \quad (4)$$

gdzie $A = (b - a)(d - c)$.

Oszacowanie Monte Carlo całki (2)

$$I_N \simeq I, \quad (5)$$

jest tym dokładniejsze, im większe jest N .

3.1 Podstawowa metoda Monte Carlo

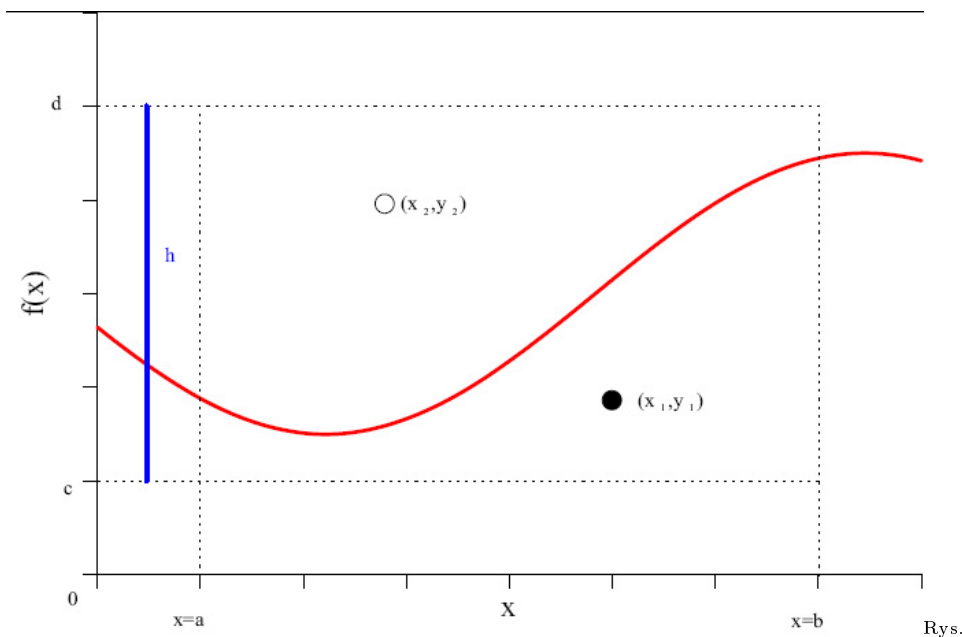
W celu oszacowania całki (2) opieramy się na twierdzeniu o wartości średniej, zgodnie z którym

$$I = (b - a) \langle f \rangle, \quad (6)$$

gdzie $\langle f \rangle$ jest wartością średnią funkcji $f(x)$. Wartość średnią liczymy metodą MC jako

$$\langle f \rangle = \frac{1}{N} \sum_{i=1}^N f(x_i), \quad (7)$$

gdzie x_i ($i = 1, \dots, N$) jest ciągiem N liczb przypadkowych o rozkładzie jednorodnym w przedziale $[a, b]$.



3.2. Ilustracja do metody Monte Carlo chybil–trafil. Punkt (x_1, y_1) spełnia warunek (3), punkt (x_2, y_2) tego warunku nie spełnia.

Otrzymujemy w ten sposób oszacowanie Monte Carlo całki (2)

$$I \simeq I_N = \frac{b-a}{N} \sum_{i=1}^N f(x_i). \quad (8)$$

W metodzie MC punkty x_i są przypadkowo rozrzucone w przedziale $[a, b]$ zgodnie z rozkładem jednorodnym, a N jest liczbą losowań.

Dla porównania wzór prostokątów dla całki (2) ma postać

$$I \simeq \frac{b-a}{N} \sum_{i=0}^{N-1} f(x_i), \quad (9)$$

przy czym $h = (b-a)/N$ oraz $x_i = x_0 + ih$, a $f(x_i)$ jest wartością funkcji $f(x)$ na początku każdego podprzedziału.

W odróżnieniu od metody MC, w metodzie prostokątów punkty x_i są rozłożone wewnątrz przedziału $[a, b]$ w sposób równomierny.

Błąd metody prostokątów w podprzedziałach $[x_{i-1}, x_i]$ jest rzędu $\mathcal{O}(h^2)$, a w przedziale $[a, b]$ jest rzędu $\mathcal{O}(h)$, czyli $\mathcal{O}(N^{-1})$.

Oszacowaniem błędów metod MC zajmiemy się później.

3.2 Metoda średniej ważonej

Metoda ta posiada nazwę angielską **importance sampling**.

Wprowadzimy ją na przykładzie obliczania całki jednowymiarowej (2).

Zgodnie z tą metodą wprowadzamy funkcję wagową $w(x)$ taką, że $w(x) \geq 0$ oraz

$$\int_a^b dx w(x) = 1 . \quad (10)$$

Funkcja $w(x)$ określa pewien rozkład prawdopodobieństwa w przedziale $[a, b]$.

Przekształcamy całkę (2) do postaci

$$I = \int_a^b dx w(x) \frac{f(x)}{w(x)} . \quad (11)$$

Oszacowaniem Monte Carlo całki (11) jest

$$I_N = \frac{b-a}{N} \sum_{i=1}^N \frac{f(x_i)}{w(x_i)} , \quad (12)$$

przy czym ciąg liczb przypadkowych x_i podlega rozkładowi prawdopodobieństwa $w(x)$ w przedziale $[a, b]$. Otrzymujemy w ten sposób możliwość redukcji wariancji, czyli błędu całkowania metodą Monte Carlo, ponieważ rozkład $w(x)$ można wybrać tak, aby przybliżał przebieg funkcji $f(x)$, a zatem redukował wariancję.

4 Całkowanie funkcji wielu zmiennych

Rozważmy funkcję $f(X)$ zależną od d zmiennych $X = (x_1, \dots, x_d)$, czyli

$$f(X) = f(x_1, \dots, x_d) .$$

Każda ze zmiennych x_j ($j = 1, \dots, d$) przyjmuje wartości z przedziału $[a_j, b_j]$.

Szukamy oszacowania całki d -wymiarowej

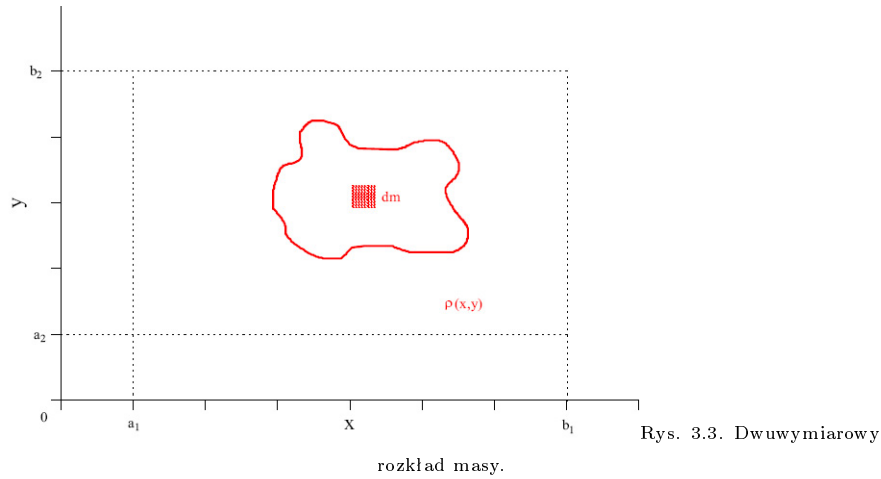
$$I = \int_{a_1}^{b_1} dx_1 \dots \int_{a_d}^{b_d} dx_d f(x_1, \dots, x_d) . \quad (13)$$

za pomocą metody MC.

Oszacowaniem Monte Carlo całki (13) jest

$$I_N = \frac{\Omega}{N} \sum_{i=1}^N f(x_{1i}, \dots, x_{di}) , \quad (14)$$

gdzie $\Omega = (b_1 - a_1) \times \dots \times (b_d - a_d)$ jest d -wymiarową objętością, a (x_{1i}, \dots, x_{di}) są wybrane losowo według rozkładu jednorodnego z przedziałów $[a_n, b_n]$, przy czym $n = 1, \dots, d$.



Przykład zastosowania metody Monte Carlo

Obliczmy współrzędne środka masy i moment bezwładności bryły sztywnej dla $d = 2$. Zakładamy pewien dwuwymiarowy rozkład masy $\varrho(x, y)$.

Masa elementu powierzchni $dA = dx dy$ dana jest wzorem

$$dm = \varrho(x, y) dx dy ,$$

gdzie $\varrho(x, y)$ jest dwuwymiarową gęstością masy. Całkowita masa bryły

$$M = \int \int_{\Omega} dx dy \varrho(x, y) ,$$

gdzie Ω jest objętością bryły. Współrzędne (X, Y) środka masy wyrażają się wzorami

$$X = \frac{1}{M} \int \int_{\Omega} dx dy x \varrho(x, y) ,$$

$$Y = \frac{1}{M} \int \int_{\Omega} dx dy y \varrho(x, y) .$$

Moment bezwładności bryły względem osi z przechodzącej przez środek masy dany jest wzorem

$$I_z^{CM} = \int dx \int dy (x^2 + y^2) \varrho(x, y) .$$

Dla dowolnego rozkładu masy $\varrho(x, y)$ każdą z powyższych całek można obliczyć metodą MC.

Np. oszacowaniem momentu bezwładności jest

$$I_z^{CM} \simeq I_N = (b_1 - a_1)(b_2 - a_2) \frac{1}{N} \sum_{i=1}^N (x_i^2 + y_i^2) \varrho(x_i, y_i),$$

gdzie $a_1 \leq x_i \leq b_1$ oraz $a_2 \leq y_i \leq b_2$, (x_i, y_i) jest parą niezależnych od siebie liczb przypadkowych generowanych zgodnie z rozkładem jednorodnym w przedziałach $[a_1, b_1]$ i $[a_2, b_2]$, a N jest liczbą losowań tych par.

5 Analiza błędów całkowania numerycznego

Metoda prostokątów w dwóch wymiarach

Obliczamy całkę z funkcji $f(x, y)$ po powierzchni A za pomocą dwuwymiarowej wersji metody prostokątów. Odpowiednim oszacowaniem tej całki jest suma objętości prostopadłościów o polu podstawy $\Delta x \Delta y$ i wysokościach $f(x_i, y_i)$, czyli

$$I = \int_A dx dy f(x, y) \simeq \Delta x \Delta y \sum_{i=0}^{N-1} f(x_i, y_i). \quad (15)$$

W celu oszacowania błędu całkowania rozwijamy funkcję $f(x, y)$ w szereg Taylora wokół punktu (x_i, y_i) . Otrzymujemy

$$\begin{aligned} f(x, y) &= f(x_i, y_i) + \frac{\partial f}{\partial x} \Big|_{(x_i, y_i)} (x - x_i) \\ &\quad + \frac{\partial f}{\partial y} \Big|_{(x_i, y_i)} (y - y_i) + \dots \end{aligned} \quad (16)$$

Błąd całkowania dla elementu powierzchni $\Delta\sigma_i = \Delta x \Delta y$ wyznaczony jest przez różnicę

$$\delta_i = \int_{\Delta\sigma_i} dx dy f(x, y) - f(x_i, y_i) \Delta x \Delta y, \quad (17)$$

przy czym powierzchnia $\Delta\sigma_i$ określona jest przez nierówności $x_i \leq x \leq x_i + \Delta x$ i $y_i \leq y \leq y_i + \Delta y$.

Podstawiamy (16) do (17) i całkujemy analitycznie

$$\begin{aligned} \delta_i &= f(x_i, y_i) \Delta x \Delta y + \frac{1}{2} \frac{\partial f}{\partial x} \Big|_{(x_i, y_i)} (\Delta x)^2 \Delta y \\ &\quad + \frac{1}{2} \frac{\partial f}{\partial y} \Big|_{(x_i, y_i)} (\Delta y)^2 \Delta x \\ &\quad - f(x_i, y_i) \Delta x \Delta y. \end{aligned} \quad (18)$$

Otrzymujemy stąd

$$\delta_i = \frac{1}{2} \left[\frac{\partial f}{\partial x} \Big|_{(x_i, y_i)} (\Delta x)^2 \Delta y + \frac{\partial f}{\partial y} \Big|_{(x_i, y_i)} (\Delta y)^2 \Delta x \right] \quad (19)$$

Korzystając z faktu, że $\Delta y = \mathcal{O}(\Delta x)$ są nieskończenie małymi tego samego rzędu dostajemy oszacowanie błędu całkowania metodą prostokątów związanego z pojedynczym prostopadłościanem

$$\delta_i \simeq \mathcal{O}[(\Delta x)^3]. \quad (20)$$

Jeżeli całkowita liczba prostopadłościanów wynosi N , przy czym

$$N \simeq \frac{A}{(\Delta x)^2},$$

to oszacowanie globalnego błędu całkowania po powierzchni A za pomocą metody prostokątów w dwóch wymiarach ma postać

$$\delta = N \delta_i \simeq \mathcal{O}[(\Delta x)] \simeq \mathcal{O}(N^{-1/2}). \quad (21)$$

Dla porównania oszacowany błąd metody prostokątów w jednym wymiarze wynosi

$$\delta \simeq \mathcal{O}(N^{-1}). \quad (22)$$

Wyniki (22) i (21) możemy uogólnić na przypadek całki po przestrzeni d -wymiarowej

$$\delta \simeq \mathcal{O}(N^{-1/d}) \quad (23)$$

Błędy całkowania w innych metodach klasycznych

Przypomnijmy, że metoda trapezów dla całki funkcji jednej zmiennej prowadzi do błędu

$$\delta \simeq \mathcal{O}(N^{-2}), \quad (24)$$

natomiast metoda Simpsona daje błąd

$$\delta \simeq \mathcal{O}(N^{-4}). \quad (25)$$

Wyrażenia na błędy (23), (24) i (25) możemy uogólnić na przypadek całkowania po przestrzeni d -wymiarowej. W przypadku całkowania funkcji d zmiennych szacowany błąd wynosi

$$\delta \simeq \mathcal{O}(N^{-m/d}), \quad (26)$$

gdzie m opisuje błąd całkowania w jednym wymiarze.

Jeżeli $d \rightarrow \infty$, to $\delta \rightarrow \mathcal{O}(N^0) = \mathcal{O}(1)$, a więc błąd całkowania metodami klasycznymi staje się rzędu wartości funkcji podcałkowej.

6 Błąd całkowania metodą Monte Carlo

Jeżeli $X = (x_1, \dots, x_d)$ oznacza wektor w przestrzeni d -wymiarowej, to oszacowaniem MC całki po przestrzeni d -wymiarowej

$$I = \int_{\Omega} d^d X f(X) \quad (27)$$

jest

$$I_N = \frac{\Omega}{N} \sum_{i=1}^N f(X_i), \quad (28)$$

gdzie punkty $X_i = (x_{1i}, \dots, x_{di})$ zostały wylosowane w sposób przypadkowy według d -wymiarowego rozkładu jednorodnego.

Wzór (28) można przepisać w postaci

$$I_N = \frac{1}{N} \sum_{i=1}^N \xi_i, \quad (29)$$

gdzie wielkości

$$\xi_i = \Omega f(X_i) \quad (30)$$

są zmiennymi losowymi podlegającymi pewnemu rozkładowi, który niekoniecznie jest rozkładem jednorodnym.

Wartości zmiennych losowych ξ_i będziemy traktować jako wyniki serii N losowań (N prób) bądź wyniki serii N pomiarów. Aby oszacować błąd serii N pomiarów, trzeba przeprowadzić kolejne serie pomiarów po N pomiarów każda. Oznaczmy przez M_S liczbę serii złożonych z N pomiarów (losowań) w każdej serii.

Wtedy $M_S N$ jest całkowitą liczbą pojedynczych pomiarów (losowań).

Niech ξ_{si} będzie wynikiem i -tego pomiaru w s -tej serii pomiarów, przy czym $s = 1, \dots, M_S$ oraz $i = 1, \dots, N$. Obliczamy wartość średnią dla s -tej serii pomiarów

$$\langle \xi_s \rangle = \frac{1}{N} \sum_{i=1}^N \xi_{si} \quad (31)$$

oraz wartość średnią dla całkowitej liczby (równej $M_S N$) pomiarów

$$\langle \xi \rangle = \frac{1}{M_S} \sum_{s=1}^{M_S} \langle \xi_s \rangle = \frac{1}{M_S N} \sum_{s=1}^{M_S} \sum_{i=1}^N \xi_{si}. \quad (32)$$

Uwaga

Ciąg wartości średnich arytmetycznych $\langle \xi_s \rangle$ podlega rozkładowi Gaussa, ponieważ jest to ciąg sum liczb przypadkowych.

Własność ta wynika z centralnego twierdzenia granicznego.

Różnica pomiędzy średnią (31) dla s -tej serii pomiarów a średnią (32) wszystkich pomiarów wynosi

$$\Delta_s = \langle \xi_s \rangle - \langle \xi \rangle, \quad (33)$$

natomiast różnica pomiędzy wynikiem pojedynczego pomiaru a średnią (32) dana jest wzorem

$$\delta_{si} = \xi_{si} - \langle \xi \rangle. \quad (34)$$

Obliczmy teraz wariancję dla ciągu średnich arytmetycznych $\langle \xi_s \rangle$, przy czym $s = 1, \dots, M_S$,

$$\sigma_S^2 = \frac{1}{M_S} \sum_{s=1}^{M_S} \Delta_s^2, \quad (35)$$

Natomiast wariancję wszystkich $M_S N$ pomiarów (M_S serii po N pomiarów każda) obliczamy jako

$$\sigma^2 = \frac{1}{M_S N} \sum_{s=1}^{M_S} \sum_{i=1}^N \delta_{si}^2. \quad (36)$$

Zgodnie z (33)

$$\Delta_s = \frac{1}{N} \sum_{i=1}^N \xi_{si} - \langle \xi \rangle = \frac{1}{N} \sum_{i=1}^N (\xi_{si} - \langle \xi \rangle), \quad (37)$$

ponieważ $\sum_{i=1}^N \langle \xi \rangle = N \langle \xi \rangle$. Ostatecznie

$$\Delta_s = \frac{1}{N} \sum_{i=1}^N \delta_{si}. \quad (38)$$

Podstawiamy (38) do (35) i otrzymujemy

$$\begin{aligned} \sigma_S^2 &= \frac{1}{M_S} \sum_{s=1}^{M_S} \frac{1}{N^2} \left(\sum_{i=1}^N \delta_{si} \right) \left(\sum_{j=1}^N \delta_{sj} \right) \\ &= \frac{1}{M_S N^2} \sum_{s=1}^{M_S} \sum_{i=1}^N \delta_{si}^2 + \frac{1}{M_S N^2} \sum_{s=1}^{M_S} \sum_{i,j=1}^N \delta_{si} \delta_{sj}. \end{aligned} \quad (39)$$

Drugi wyraz w (39) zeruje się, ponieważ jest on proporcjonalny do średniej

$$\langle \delta_{si} \delta_{sj} \rangle,$$

gdzie odstępstwa od średniej δ_{si} podlegają – podobnie jak ξ_{si} – rozkładowi przypadkowemu o średniej $\langle \delta \rangle = 0$. W związku z tym (39) prowadzi do następującego wzoru na wariancję

$$\sigma_S^2 = \frac{\sigma^2}{N}. \quad (40)$$

Miarą błędu dla pojedynczej serii N pomiarów jest **odchylenie standardowe**, czyli pierwiastek kwadratowy z wariancji (40)

$$\sigma_S = \frac{\sigma}{\sqrt{N}}. \quad (41)$$

Wynika stąd, że niezależnie od liczby wymiarów przestrzeni błąd całkowania za pomocą metody Monte Carlo jest rzędu

$$\mathcal{O}(N^{-1/2}). \quad (42)$$

Uwaga

Przyjmując poziom ufności $\alpha = 0.997$ dla rozkładu Gaussa średnich arytmetycznych $\langle \xi_s \rangle$ otrzymujemy następujące oszacowanie błędu metody Monte Carlo

$$\overline{\sigma_S} = \frac{3\sigma}{\sqrt{N}}. \quad (43)$$

Jest to tzw. **reguła trzech sigma**.

Otrzymane wyniki można ująć w tabeli podającej porównanie błędów całkowania otrzymanych za pomocą różnych metod.

Tabela 3.1. Porównanie błędów całkowania za pomocą różnych metod. Dla $d \geq d_0$ metoda MC prowadzi do mniejszego błędu.

metoda	błąd	d_0
Monte Carlo	$N^{-1/2}$	–
prostokątów	$N^{-1/d}$	3
trapezów	$N^{-2/d}$	5
Simpsona	$N^{-4/d}$	9

Z Tabeli 3.1 wynika, że metoda MC staje się dokładniejsza od każdej z metod klasycznych, gdy wymiar przestrzeni, po której całkujemy, przekracza d_0 .

Dodatkową zaletą metody MC jest możliwość zmniejszenia błędu całkowania przez redukcję odchylenia standardowego σ [por. wzór (41)].