

zadanie:	1a	1b	Σ :
punkty:	30	30	30
wynik:			
efekty:			

1 Metoda Monte Carlo

Przeglądając literaturę dotyczącą metody Monte Carlo trudno jest znaleźć konkretną definicję. Trudność ta bierze się z faktu mnogości rodzajów metod Monte Carlo i ogromnego zakresu jej stosowalności. Należałoby zatem mówić o metodach Monte Carlo jako o szerokiej klasie algorytmów obliczeniowych, a nie jako o pojedynczej metodzie. W najbardziej ogólny sposób można powiedzieć, że głównym założeniem metody Monte Carlo jest zastąpienie problemu numerycznego, zadaniem z dziedziny prawdopodobieństwa, o takim samym rozwiązaniu. Metoda Monte Carlo jest bardzo przydatnym narzędziem ze względu na fakt, że obliczenia statystyczne, szczególnie przy wykorzystaniu możliwości nowoczesnych komputerów, pochłaniają często znacznie mniej czasu obliczeniowego, niż ich numeryczne odpowiedniki. Metoda ta znajduje zastosowanie w modelowaniu matematycznym procesów złożonych na tyle, że bardzo trudno uzyskać jest ich wyniki za pomocą podejścia analitycznego. Istotną rolę w tej metodzie odgrywa losowanie wielkości charakteryzujących proces, przy czym losowanie dokonywane jest zgodnie z rozkładem, który musi być znany.

2 Metoda Riemanna

W ogólności całkowanie metodą Riemanna, to zagadnienie numeryczne (deterministyczne) polegające na podziale zadanego obszaru na pewną liczbę prostokątów (lub innych figur), wyznaczenie obszaru dla każdego prostokąta i zsumowanie uzyskanych wartości. Otrzymana w ten sposób suma jest aproksymacją badanej całki.

3 Wyznaczyć $P(0 < Z_1 \leq 1, 0 < Z_2 \leq 1)$

3.1 Metoda analityczna

W celu wyznaczenia prawdopodobieństwa w pierwszej kolejności skorzystamy z własności niezależności zmiennych losowych Z_1 i Z_2 . Wiemy, że dla niezależnych

zmiennych losowych X i Y zachodzi

$$P(X \leq a, Y \leq b) = P(X \leq a)P(Y \leq b),$$

dla $a, b \in \mathbb{R}$. Zatem możemy zapisać

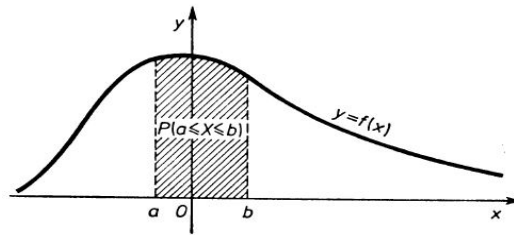
$$P(0 < Z_1 \leq 1, 0 < Z_2 \leq 1) = P(0 < Z_1 \leq 1)P(0 < Z_2 \leq 1). \quad (1)$$

Skorzystamy teraz faktu, że dla ciągłych zmiennych losowych zachodzi

$$P(a \leq X < b) = P(a < X \leq b) = P(a < X < b) = P(a \leq X \leq b) = F(b) - F(a)$$

$$P(a < X < b) = \int_a^b f(x)dx,$$

gdzie f jest funkcją gęstości prawdopodobieństwa, a F dystrybuantą. Geometryczną interpretację przedstawiono na rysunku 1.



Rysunek 1: Ilustracja $P(a \leq X \leq b)$ jako pola nad przedziałem (a, b) pod wykresem gęstości f .

Wtedy rozpisujemy (1) do postaci

$$(F_{Z_1}(1) - F_{Z_1}(0))(F_{Z_2}(1) - F_{Z_2}(0)), \quad (2)$$

gdzie F_{Z_1} i F_{Z_2} są dystrybuantami rozkładów zmiennych losowych Z_1 i Z_2 . Z treści zadania wiemy, że (Z_1, Z_2) ma dwuwymiarowy standardowy rozkład normalny, dodając do tego niezależność zmiennych możemy stwierdzić, że $Z_1, Z_2 \sim \mathcal{N}(0, 1)$. Oznacza to, że dystrybuanty F_{Z_1} i F_{Z_2} są sobie równe oraz odpowiadają dystrybuancie standardowego rozkładu normalnego Φ , której wartości są stabilizowane. Mając powyższe fakty na uwadze możemy zapisać

$$(\Phi(1) - \Phi(0))^2 \quad (3)$$

odczytujemy wartości z tablicy rozkładu normalnego: $\Phi(1) \approx 0,8413$, $\Phi(0) \approx 0,5$. Czyli

$$P(0 < Z_1 \leq 1, 0 < Z_2 \leq 1) = (0,3413)^2 = 0,1165. \quad (4)$$

3.2 Metoda Monte Carlo

Przyjmując dane w zadaniu parametry dokonujemy symulacji. Wykonujemy 100 wyciągów całki i wyciągamy średnią.

```

1 m <- 10000
2 pdf <- function(z1, z2) (1/(2*pi)*exp(-(z1^2 + z2^2)/2))
3
4 a1 <- 0
5 b1 <- 1
6
7 a2 <- 0
8 b2 <- 1
9
10
11 MC_a <- sapply(1:100, function(x) sum(pdf(runif(m, a1, b1),
12 runif(m, a2, b2)))) * (b1 - a1)*(b2 - a2)/m)
13
14 MC_mean_a <- mean(MC_a)
15 MC_mean_a

```

Uzyskany wynik to ≈ 0.1165187 , co w porównaniu do wartości uzyskanej w sposób analityczny stanowi błąd względny na poziomie 0.016%.

3.3 Metoda Riemanna

```
1 x1 <- seq(a1, b1, by=1/m)
2 x2 <- seq(a2, b2, by=1/m)
3
4 R_a <- sum(pdf(x1, x2))*(b1 - a1)^2/m
5 R_a
```

Uzyskany wynik to ≈ 0.1188716 , co w porównaniu do wartości uzyskanej w sposób analityczny stanowi błąd względny na poziomie 2.04%.

4 Wyznaczyć $P(Z_1^2 + Z_2^2 < 1)$

4.1 Metoda analityczna

Zacznijmy od przypomnienia własności cechującej funkcję gęstości prawdopodobieństwa rozkładów wektorów losowych o zmiennych niezależnych. Jeżeli (Z_1, Z_2) jest wektorem losowym oraz Z_1 i Z_2 są niezależne, to funkcję gęstości można wyrazić wzorem

$$f_{Z_1, Z_2}(z_1, z_2) = f_{Z_1}(z_1)f_{Z_2}(z_2). \quad (5)$$

Z treści zadania wiemy, że $Z_1, Z_2 \sim \mathcal{N}(0, 1)$, zatem przyjmując za funkcję gęstości wzór funkcję gęstości standardowego rozkładu normalnego dostajemy

$$f_{Z_1}(z_1)f_{Z_2}(z_2) = \frac{1}{2\pi} \exp\left(-\frac{(z_1^2 + z_2^2)}{2}\right). \quad (6)$$

Dziedzina, po której całkujemy jest kołem, zatem należy obliczyć całkę podwójną po kole jednostkowym z funkcji z (6). W celu uproszczenia obliczeń przejdźmy na współrzędne biegunowe

$$\begin{cases} z_1 = r \cos \varphi \\ z_2 = r \sin \varphi \\ |J| = r \end{cases} \quad (7)$$

oraz

$$\begin{cases} r \in (0, 1) \\ \varphi \in (0, 2\pi) \end{cases} \quad (8)$$

Przystępujemy do obliczenia całki

$$\begin{aligned} P(Z_1^2 + Z_2^2 < 1) &= \int_0^{2\pi} \int_0^1 \frac{1}{2\pi} \exp\left(-\frac{(r^2 \cos^2 \varphi + r^2 \sin^2 \varphi)}{2}\right) r dr d\varphi = \\ &= \frac{1}{2\pi} \int_0^{2\pi} \int_0^1 \exp\left(-\frac{r^2}{2}\right) r dr d\varphi = \\ &= \frac{1}{2\pi} 2\pi \int_0^1 \exp\left(-\frac{r^2}{2}\right) r dr \end{aligned}$$

podstawiając $t = -r^2/2$ dostajemy

$$\int_0^1 \exp\left(-\frac{r^2}{2}\right) r dr = \int_{\frac{-1}{2}}^0 \exp(t) dt = \exp(0) - \exp\left(\frac{-1}{2}\right) =$$

$$= 1 - \exp\left(\frac{-1}{2}\right) \approx 0,3935.$$

4.2 Metoda Monte Carlo

W tym przypadku generujemy liczby z okręgu jednostkowego o środku w punkcie $(0,0)$. Również wykonujemy 100 powtórzeń wyliczenia i uśredniamy wynik

```

1 # biegunowe
2 r <- 1 * sqrt(runif(m))
3 phi <- runif(m) * 2 * pi
4
5 # kartezjanskie (kolo jednostkowe o srodku w (0,0))
6 s1 <- r * cos(phi)
7 s2 <- r * sin(phi)
8
9 MC_b <- sapply(1:100, function(i) sum(pdf(s1, s2)) * pi*1/m)
10 MC_mean_b <- mean(MC_b)
11 MC_mean_b

```

Uzyskany wynik to ≈ 0.3936758 , co w porównaniu do wartości uzyskanej w sposób analityczny stanowi błąd względny na poziomie 0.045%.

4.3 Metoda Riemanna

W tym przypadku skorzystamy z części rozwiązania analitycznego w celu znalezienia całki.

```

1 # na współrzędnych biegunowych
2 pdf_polar <- function(r) (1/(2*pi))*exp(-(r^2)/2)*r
3 # na podstawie rownania z rozwiazania analitycznego w zad. 2
4
5 # promien
6 r1 <- 0
7 r2 <- 1
8
9 # kat
10 phi1 <- 0
11 phi2 <- 2*pi
12
13 v1 <- seq(r1, r2, by=r2/m)
14
15 R_b <- sum(pdf_polar(v1))*(r2 - r1)*(phi2 - phi1)/m
16 R_b

```

Uzyskany wynik to ≈ 0.3934997 , co w porównaniu do wartości uzyskanej w sposób analityczny stanowi błąd względny na poziomie 0.000085%.

5 Wyznaczyć $P(Z_1 > 0, Z_2 > 0, Z_1 + Z_2 < 1)$

5.1 Metoda analityczna

Niestety brak wyniku.

5.2 Metoda Monte Carlo

W tym przypadku musimy ograniczyć wylosowane punkty do trójkąta. Analogicznie jak poprzednio wykonujemy 100 wyliczeń całki i uśredniamy wynik.

```
1 a <- 0
2 b <- 1
3 h=b-a
4
5 MC_c <- c()
6
7 for (j in 1:100) {
8
9   MC_c_tmp <- c()
10  # generujemy liczby
11  c1 <- a + (b - a) * runif(m)
12  c2 <- a + (b - a) * runif(m)
13
14  for (i in 1:m) {
15
16    if (c1[i] + c2[i] < 1) { # sprawdzamy czy sa w trojkacie
17
18      MC_c_tmp = c(MC_c_tmp, pdf(c1[i], c2[i]))
19
20    }
21  }
22
23  MC_c = c(MC_c, sum(MC_c_tmp)*h^2/m)
24 }
25
26 MC_mean_c <- mean(MC_c)
27 MC_mean_c
```

Uzyskany wynik to ≈ 0.06757265 .

5.3 Metoda Riemanna

```

1 a <- 0
2 b <- 1
3
4 k <- seq(a, b, 1/m)
5
6 R_c_tmp <- 0
7
8 for (j in k) {
9   for(i in k) {
10     if(j + i < 1) { # czy sa w trojkacie
11       R_c_tmp <- R_c_tmp + pdf(j, i)
12     }
13   }
14 }
15
16 R_c <- R_c_tmp * (1/m)^2
17
18 R_c

```

Uzyskany wynik to ≈ 0.06773793 .

6 Eksperyment

Zbadamy dokładność wyników przy zwiększeniu liczby m . W tym celu przyjmujemy $m = 20000$. Uzyskane wyniki prezentują się następująco

	(a) M = 10000	(b) M = 10000	(c) M = 10000	(a) M = 20000	(b) M = 20000	(c) M = 20000
Metoda.analityczna	0.1165000	0.3935000	N/A	0.1165000	0.3935000	N/A
Metoda.Monte.Carlo	0.1165187	0.3936758	0.06757265	0.1164938	0.3935227	0.06776662
Metoda.Riemanna	0.1188716	0.3934997	0.06773793	0.1188662	0.3934845	0.06773398

Rysunek 2: Wyniki dla $m = 10000$ oraz $m = 20000$.

Tabela błędów

	Bład.względny
(a) MC: $M = 10000$	1.605242e-02
(b) MC: $M = 10000$	4.466976e-02
(c) MC: $M = 10000$	NA
(a) MC: $M = 20000$	5.319368e-03
(b) MC: $M = 20000$	5.763722e-03
(c) MC: $M = 20000$	NA
(a) R: $M = 10000$	2.035740e+00
(b) R: $M = 10000$	8.488259e-05
(c) R: $M = 10000$	NA
(a) R: $M = 20000$	2.031069e+00
(b) R: $M = 20000$	3.938159e-03
(c) R: $M = 20000$	NA

Rysunek 3: Błędy względne dla $m = 10000$ oraz $m = 20000$.

7 Wnioski

Przyglądając się wartościom uzyskanym poprzez metodę Monte Carlo i metodę Riemanna w stosunku do wartości teoretycznych zauważamy, że wartości całki z metody Monte Carlo są obarczone mniejszym błędem dla przypadku a), natomiast metoda Riemanna daje lepszy wynik dla punktu b).

Ponadto w naszym przypadku dokładniejsze wyniki uzyskiwaliśmy dla większej liczby generowanych punktów (poza zaskakująco dobrym wynikiem dla metody Riemanna dla $m = 1000$ w b)), tym niemniej wraz ze wzrostem ilości generowanych liczb zwiększał się czas wykonywania algorytmu. Zatem przy stosowaniu obu metod numerycznych należy mieć na uwadze poprawność rozwiązania w stosunku do możliwości komputera.