

Całkowanie metodą Monte Carlo w 2D z rozkładem eksponencjalnym.

Tomasz Chwiej

16 czerwca 2021

1 Opis problemu

Należy oszacować wartość całki

$$I = \int_0^1 \int_0^1 g(x, y) \cdot f(x) \cdot f(y) dx dy \quad (1)$$

gdzie:

$$g(x, y) = \frac{\sin(x+y)}{\ln(2+x+y)}, \quad f(x) = e^{-x}, \quad f(y) = e^{-y} \quad (2)$$

metodą Monte Carlo. Jako dokładną wartość przyjmujemy $I_{dok} = 0.2557840245$.

1.1 Sposób 1

Używamy bezpośrednio kwadratury Monte Carlo dla funkcji podcałkowej

$$I \approx \frac{1}{n} \sum_{i=1}^n z(x_i, y_i) = \frac{1}{n} \sum_{i=1}^n g(x_i, y_i) \cdot f(x_i) \cdot f(y_i) \quad (3)$$

gdzie $x_i, y_i \in U(0, 1)$, a n jest liczbą losowań.

W celu kontroli jakości rozwiązania obliczamy wariancję pojedynczego wyniku:

$$\sigma^2 = \frac{1}{n-1} \left[\sum_{i=1}^n z^2(x_i, y_i) - \frac{1}{n} \left(\sum_{i=1}^n z(x_i, y_i) \right)^2 \right] \quad (4)$$

oraz odchylenie standardowe wartości średniej:

$$\sigma_{sr} = \frac{\sigma}{\sqrt{n}} \quad (5)$$

1.2 Sposób 2

Zauważmy, że zmienne x oraz y mogą być opisane funkcjami gęstości prawdopodobieństwa $f(x)$ i $f(y)$ (**jeszcze nieunormowanymi**) które występują w całkowaniu

$$I = \int_0^1 \int_0^1 \frac{g(x, y)}{C_x C_y} \cdot C_x f(x) \cdot C_y f(y) dx dy \quad (6)$$

Zatem możemy do wyznaczenia całki możemy użyć także wzoru:

$$I \approx \frac{1}{n} \sum_{i=1}^n g(x_i, y_i) \frac{1}{C_x C_y} \quad (7)$$

gdzie $x_i, y_i \in Q$, gdzie Q jest rozkładem eksponencjalnym, n jest liczbą losowań, a stałe C_x i C_y wyznaczmy dla rozkładu Q . Wariancję pojedynczej zmiennej losowej $g(x, y)$ wyznaczamy ze wzoru

$$\sigma^2 = \frac{1}{n-1} \left[\sum_{i=1}^n \left(\frac{g(x_i, y_i)}{C_x C_y} \right)^2 - \frac{1}{n} \left(\sum_{i=1}^n \frac{g(x_i, y_i)}{C_x C_y} \right)^2 \right] \quad (8)$$

Generator dla zmiennej o rozkładzie eksponencjalnym skonstruujemy wykorzystując metodę odwracania dystrybuanaty. Najpierw zajmujemy się stałą normalizacji rozkładu:

$$f_{gp}(x) = C_x e^{-x} \rightarrow \int_0^1 f_{gp}(x) dx = 1 \quad (9)$$

skąd otrzymujemy

$$C_x = C_y = \frac{1}{1 - e^{-1}} \quad (10)$$

Zgodnie z metodą odwracania dystrybuanaty mamy warunek:

$$F(x) = \int_0^x f_{gp}(x') dx' = C_x (1 - e^{-x}) = X_U \in U(0, 1) \quad (11)$$

który pozwala wyznaczyć zmienną X_Q o rozkładzie eksponencjalnym Q :

$$X_Q = -\ln \left(1 - \frac{X_U}{C_x} \right) \quad (12)$$

2 Zadania do wykonania

1. Przyjmujemy $N = 10^5$ jako maksymalną liczbę losowań.
2. Oszacować wartość całki (1) wykorzystując wzór (3) oraz wyznaczyć błąd oszacowania wartości średniej (wykorzystując wzór 5). W trakcie obliczeń do pliku zapisać: n (aktualna ilość wykonanych losowań), $I(n)$ oraz $\sigma_{sr}(n)$ dla $n = 10, 10^2, 10^3, 10^4, 10^5$.
3. Oszacować wartość całki (1) wykorzystując wzór (7) oraz wyznaczyć błąd oszacowania wartości średniej (wykorzystując wzór 8). W trakcie obliczeń do pliku zapisać: n (aktualna ilość wykonanych losowań), $I(n)$ oraz $\sigma_{sr}(n)$ dla $n = 10, 10^2, 10^3, 10^4, 10^5$.
4. Testowanie generatora Q .
 - a) Przedział $[0, 1]$ podzielić na $k = 10$ podprzedziałów, następnie wylosować $N = 10^5$ liczb pseudolosowych tym generatorem.
 - b) Dla każdego podprzedziału wyznaczyć ilość liczb n_i , $i = 1, 2, \dots, k$, które wpadają do podprzedziału i .
 - c) Wyznaczyć teoretyczne prawdopodobieństwo wylosowania liczby z danego podprzedziału jako

$$P_i(x_{i-1} < x \leq x_i) = F(x_i) - F(x_{i-1}) \quad (13)$$

gdzie dystrybuanata $F(x)$ jest określona wzorem (11), a x_{i-1} oraz x_i wyznaczają krańce i -tego podprzedziału.

- d) Narysować histogramy p_i oraz P_i na jednym rysunku, np. w Gnuplocie

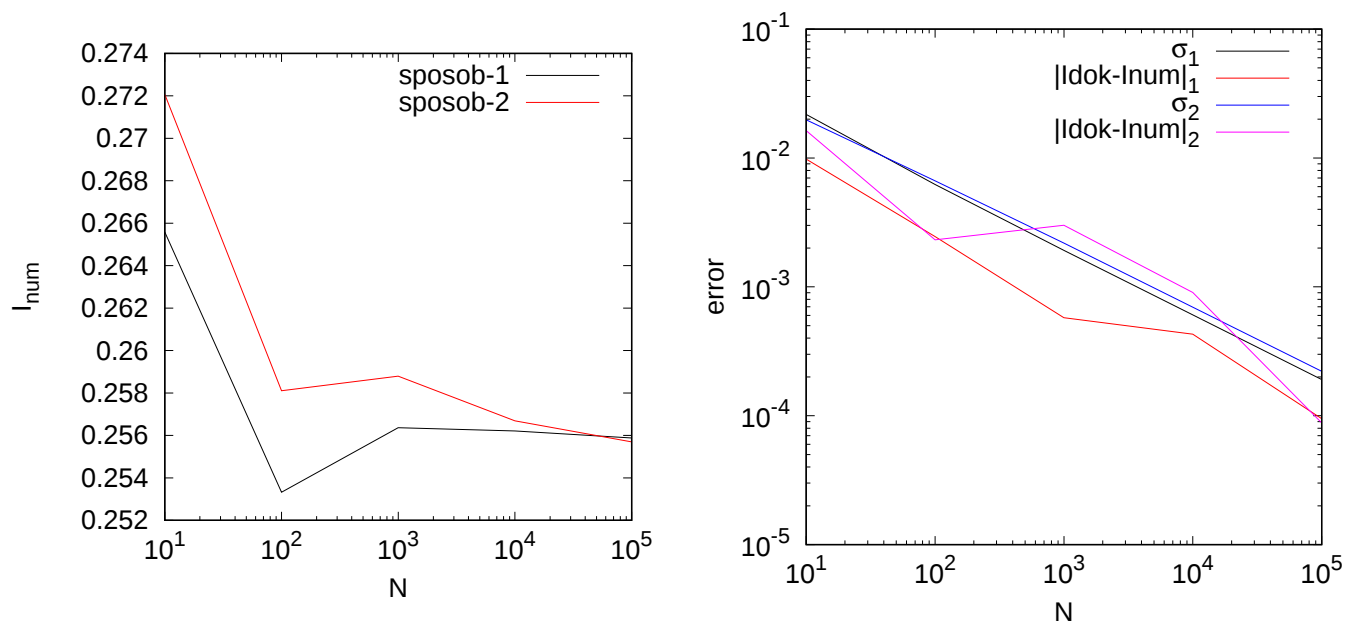
```
plot 'hist_pk.dat' u 1:2 w boxes lc rgb 'black' t 'num', \
    'hist_Pk.dat' u 1:3 w boxes lc rgb 'red' t 'dok'
```

- e) Obliczyć wartość statystyki χ_{k-1}^2 :

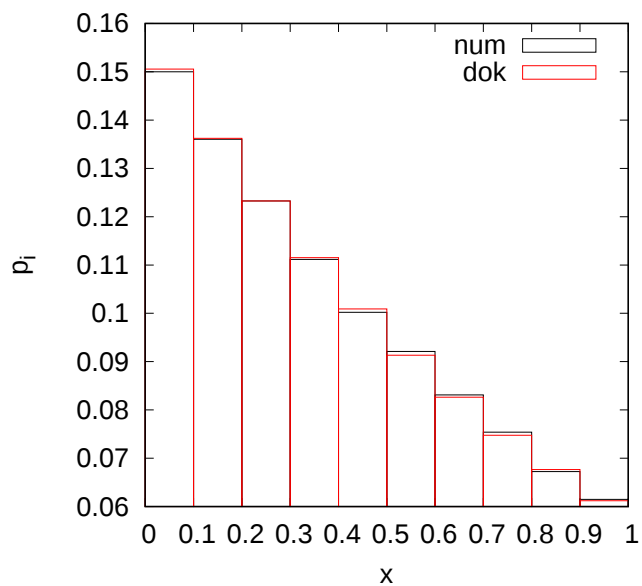
$$\chi_{k-1}^2 = \sum_{i=1}^k \frac{(n_i - N \cdot P_i)^2}{N \cdot P_i} \quad (14)$$

i zapisać do pliku.

3 Przykładowe wyniki



Rysunek 1: Wartość całki (lewy) i błąd (prawy) liczona 1 i 2 sposobem.



Rysunek 2: Histogram dla generatora eksponencjalnego Q .