

Całkowanie metodą Monte Carlo

Plan wykładu:

- podstawowa metoda Monte Carlo
- metody MC o zwiększonej efektywności
 - losowania ważonego
 - zmiennej kontrolnej
- przykłady szacowania wartości całek

Przypomnienie wiadomości ze statystyki

- funkcja gęstości prawdopodobieństwa (**fgp**)
- wartość oczekiwana
- wariancja (odchylenie standardowe)
- funkcja gęstości prawdopodobieństwa (**fgp**) ta ma następujące własności

$$\bigwedge_{x \in [a, b]} f(x) \geq 0 \qquad \int_a^b f(x) dx = 1$$

przy jej pomocy można określić prawdopodobieństwo zdarzenia, że zmienna losowa x przyjmie wartość pomiędzy x a $x+dx$:

$$P \{x \leq x_i \leq x + dx\} = f(x)dx$$

- dla danej funkcji gęstości prawdopodobieństwa można określić dystrybuantę rozkładu

$$F(x) = \int_{-\infty}^x f(x') dx'$$

która jest funkcją prawostronnie ciągłą i niemalejącą

$$P \{x_1 \leq x \leq x_2\} = \int_a^b f(x) dx = F(x_2) - F(x_1)$$

- **wartość oczekiwana zmiennej losowej x** (pierwszy moment rozkładu)

$$\langle x \rangle = E(x) = \mu(x) = \int_a^b x f(x) dx$$

- **wariancja zmiennej losowej x** (drugi moment centralny)

$$\begin{aligned} \sigma^2(x) &= \langle [x - \langle x \rangle]^2 \rangle \\ &= \int_a^b [x - \langle x \rangle]^2 f(x) dx \\ &= \int_a^b [x^2 - 2x\langle x \rangle + \langle x \rangle^2] f(x) dx \\ &= \underbrace{\int_a^b x^2 f(x) dx}_{\langle x^2 \rangle} - 2\langle x \rangle \underbrace{\int_a^b x f(x) dx}_{\langle x \rangle} + \langle x \rangle^2 \underbrace{\int_a^b f(x) dx}_1 = \langle x^2 \rangle - \langle x \rangle^2 \end{aligned}$$

- **odchylenie standardowe** (miara rozrzutu wokół wartości oczekiwanej)

$$\sigma^2(x) = \langle x^2 \rangle - \langle x \rangle^2 \quad \longrightarrow \quad \sigma(x) = \sqrt{\langle x^2 \rangle - \langle x \rangle^2}$$

- zazwyczaj chcemy określić wartość oczekiwaną zmiennej losowej z będącej funkcją np. innej zmiennej losowej x , czyli $z=z(x)$:

$$\langle z \rangle = \mu(z) = \int_{-\infty}^{\infty} z g(z) dz$$

ale ... zazwyczaj $g(z)$ jest nieznana.

- można jednak skorzystać z prawa przenoszenia prawdopodobieństwa

$$g(z) dz = f(x) dx$$

i powyższą całkę przekształcić do postaci

$$\langle z \rangle = \mu(z) = \int_{-\infty}^{\infty} z g(z) dz = \int_a^b z(x) f(x) dx$$

analogicznie drugi moment

$$\langle z^2 \rangle = \int_a^b z^2(x) f(x) dx$$

- wariancja

$$\sigma^2(z) = \langle z^2 \rangle - \langle z \rangle^2$$

- odchylenie standardowe

$$\sigma(z) = \sqrt{\langle z^2 \rangle - \langle z \rangle^2}$$

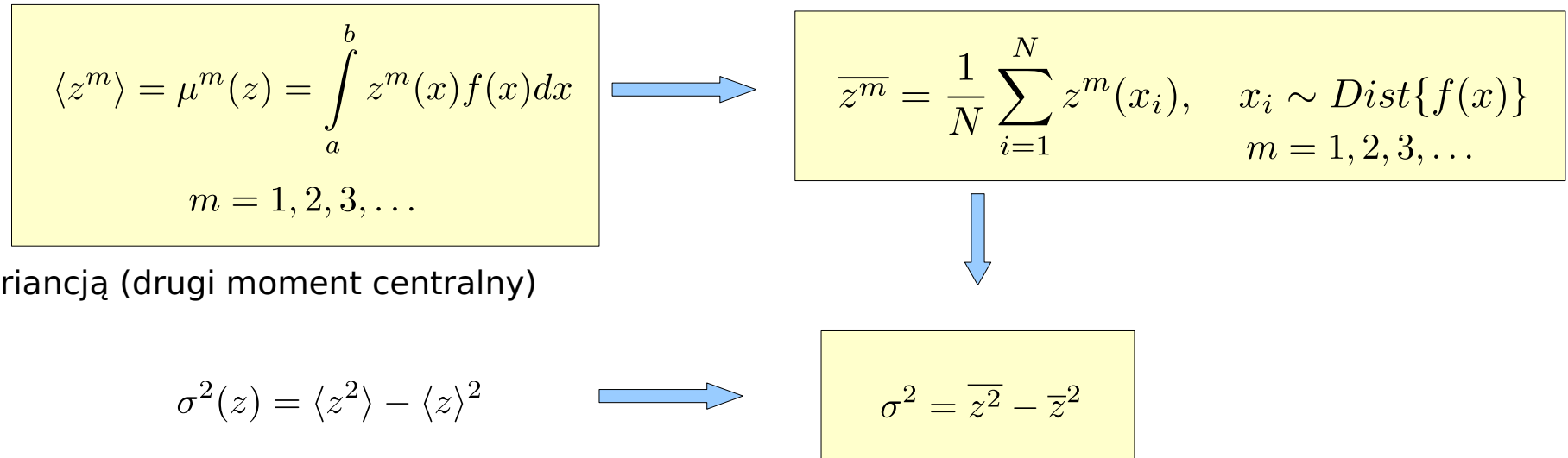
- jeśli ciąg liczb

$$\{x_i\} = \{x_i | n = 1, 2, \dots, N\}$$

stanowią zmienne losowe o funkcji gęstości prawdopodobieństwa $f(x)$

(**LOSUJEMY JE GENERATOREM O ROZKŁADZIE $f(x)$**)

to **estymatorem wartości oczekiwanej** m -tego momentu zmiennej losowej $z(x_i)$ jest średnia z próby



- powyższe wzory stanowią podstawowy sposób całkowania MC, tj. określania przybliżonej wartości całki oraz oszacowania jakości tego przybliżenia

- miarą „rozrzutu” zmiennych losowych z_i wokół wartości średniej jest **odchylenie standardowe**

$$\sigma = \sqrt{\sigma^2(z)}$$

ale \bar{z} też jest zmienną losową, ponieważ konstruujemy ją ze zmiennych z_i (każda z nich ma identyczną wariancję)

- jakie jest **odchylenie standardowe średniej**?

$$\sigma^2(\bar{z}) = \sigma^2 \left(\frac{1}{N} \sum_{i=1}^N z_i \right) = \frac{1}{N} \left(\frac{1}{N} \sum_{i=1}^N \sigma^2(z_i) \right) = \frac{1}{N} \sigma^2(z), \quad \sigma(z_i) = \sigma(z)$$

do jego estymacji możemy użyć $\sigma(z)$

$$\sigma(\bar{z}) = \frac{\sigma(z)}{\sqrt{N}}$$

- w Monte Carlo jest to **miara błędu/niepewności**, łatwo więc oszacować ile losowań musimy wykonać, aby zmniejszyć błąd

Podstawowa metoda całkowania Monte Carlo

Interesuje nas wyznaczenie (a raczej estymacja) wartości oczekiwanej zmiennej losowej

$$z = z(\mathbf{x})$$

która jest funkcją wektora zmiennych (losowych):

$$\mathbf{x} = [x_1, x_2, \dots, x_m]$$

Rozkład prawdopodobieństwa zmiennej losowej z opisuje (**nieznana**) fgp $g(z)$

$$\int_{-\infty}^{\infty} g(z(\mathbf{x})) dz = 1$$

a rozkład prawdopodobieństwa wektora \mathbf{x} opisuje **znana** fgp $f(\mathbf{x})$

$$\int_V f(\mathbf{x}) d\mathbf{x} = 1$$

$$\langle z \rangle = \int_{-\infty}^{\infty} z g(z) dz = \int_V z(\mathbf{x}) f(\mathbf{x}) d\mathbf{x}$$

Przy takich założeniach, zgodnie z CTG **dystrybuanta wartości oczekiwanej** ma rozkład normalny (niezależnie od rozkładu zmiennej losowej z)

$$\lim_{N \rightarrow \infty} P \left\{ \frac{|\bar{z} - \langle z \rangle|}{\frac{\sigma(z)}{\sqrt{N}}} \leq \lambda \right\} = \frac{1}{\sqrt{2\pi}} \int_{-\lambda}^{\lambda} e^{-\frac{u^2}{2}} du$$

metodę Monte Carlo szacowania wartości całek w wersji podstawowej definiują wzory:

a) wartość całki

$$I = \int_V z(\mathbf{x}) f(\mathbf{x}) d\mathbf{x} \approx \frac{1}{N} \sum_{i=1}^N z(\mathbf{x}_i), \quad \mathbf{x}_i \sim f(\mathbf{x})$$

Uwaga: \mathbf{x} - jest wektorem, którego składowe są niezależnymi zmiennymi losowymi o określonych funkcjach gęstości prawdopodobieństwa

b) błąd oszacowania

$$\sigma(I) = \sqrt{\int_V (z - \langle z \rangle)^2 f(\mathbf{x}) d\mathbf{x}} \approx \frac{\sigma(z)}{\sqrt{N}}$$

- gdy obszarem całkowania jest określony **nieregularny** podzbiór przestrzeni \mathbb{R}^M , wówczas obliczaną całkę trzeba zapisać w zmienionej postaci

$$V \subset \Omega \quad \longrightarrow \quad I = \int_V z(\mathbf{x}) f(\mathbf{x}) d\mathbf{x} = \int_{\Omega} \mathbf{1}_V(\mathbf{x}) z(\mathbf{x}) f(\mathbf{x}) d\mathbf{x}$$

gdzie: $\mathbf{1}_V(\mathbf{x})$ jest funkcją przynależności do zbioru

$$\mathbf{1}_V(\mathbf{x}) = \begin{cases} 1 & \iff \mathbf{x} \in V \\ 0 & \iff \mathbf{x} \notin V \end{cases}$$

- „kwadratura” Monte Carlo (metoda orzeł-reszka)

$$\begin{aligned} I &= \int_V z(\mathbf{x}) d\mathbf{x} = \int_{\Omega} \mathbf{1}_V(\mathbf{x}) z(\mathbf{x}) d\mathbf{x} \\ &= \int_{\Omega} \underbrace{\frac{1}{\Omega}}_{f(\mathbf{x})} \mathbf{1}_V(\mathbf{x}) z(\mathbf{x}) d\mathbf{x} \\ &\approx \frac{\Omega}{N} \sum_{i=1}^N \mathbf{1}_V(\mathbf{x}) z(\mathbf{x}), \quad \mathbf{x} \sim f(\mathbf{x}) \end{aligned}$$

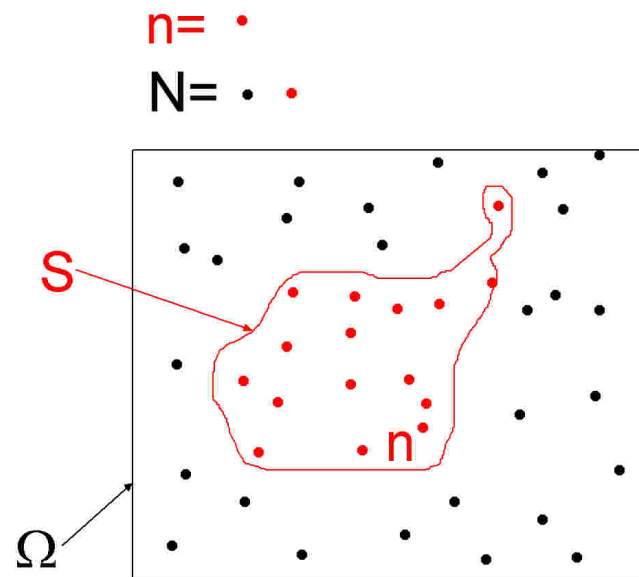
warunek unormowania $f(\mathbf{x})$

$$\int_{\Omega} f(\mathbf{x}) d\mathbf{x} = \int_{\Omega} \frac{1}{\Omega} d\mathbf{x} = 1$$

Uwagi:

- w powyższym przypadku zakładamy, że fgp jest stała w obszarze Ω
- „kwadratura” – wzór podobny jak w całkowaniu numerycznym, ale tu położenia węzłów są losowane

przykład: Wyznaczyć pole powierzchni obiektu o **nieregularnym** kształcie



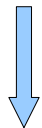
Uwagi:

- brzeg obszaru V (tu powierzchnia S) może być opisywany zestawem skomplikowanych równań
- Obszar Ω powinien mieć regularny kształt (np. n -wymiarowy prostokąt lub kula) aby w łatwy sposób wygenerować punkty równomiernie go pokrywające
- metoda mało efektywna w $d=3$ i więcej wymiarach (jest to odpowiednik metody eliminacji von Neumanna)

$$S = \int_V 1 d^2\mathbf{r} = \int_{\Omega} 1_V(\mathbf{r}) d^2\mathbf{r} = \int_{\Omega} \Omega \frac{1}{\Omega} 1_V(\mathbf{r}) d^2\mathbf{r}$$

$$= \Omega \int_{\Omega} f(\mathbf{r}) 1_V(\mathbf{r}) d^2\mathbf{r}$$

- funkcją podcałkową jest funkcja przynależności 1_V



$$S = \frac{\Omega}{N} \sum_{i=1}^N 1_V(\mathbf{r}_i) = \Omega \frac{n}{N},$$

$$\mathbf{r} \sim f(\mathbf{r}) = U(x_1, x_2) \times U(y_1, y_2)$$

Przykład

Należy obliczyć numerycznie wartość całki

$$I = \int_0^1 dx_1 \int_0^1 dx_2 \int_0^1 dx_3 \int_0^1 dx_4 \int_0^1 dx_5 g(x_1, x_2, x_3, x_4, x_5)$$

$$g(x_1, x_2, x_3, x_4, x_5) = \sqrt{x_1 + x_2 + x_3 + x_4 + x_5}$$

- metoda trapezów

$$\int_0^1 f(y) dy = h \left[\frac{f(y_0) + f(y_n)}{2} + \sum_{i=1}^{n-1} f(y_i) \right] \quad \begin{array}{l} y_0 = 0 \\ y_n = 1 \end{array}$$

$$I_{trapez} = h^5 \sum_{i=0}^n w_i \sum_{j=0}^n w_j \sum_{k=0}^n w_k \sum_{l=0}^n w_l \sum_{m=0}^n w_m \cdot g(x_i, x_j, x_k, x_l, x_m)$$

$$h = \frac{1}{n} \quad w_{i,j,k,l,m} = \begin{cases} \frac{1}{2} & i, j, k, l, m = 0, n \\ 1 & i, j, k, l, m = 1, 2, \dots, n-1 \end{cases}$$

- „kwadratura” Monte Carlo

$$I_{MC} = \frac{\Omega}{N} \sum_{i=1}^N g(\mathbf{X})$$

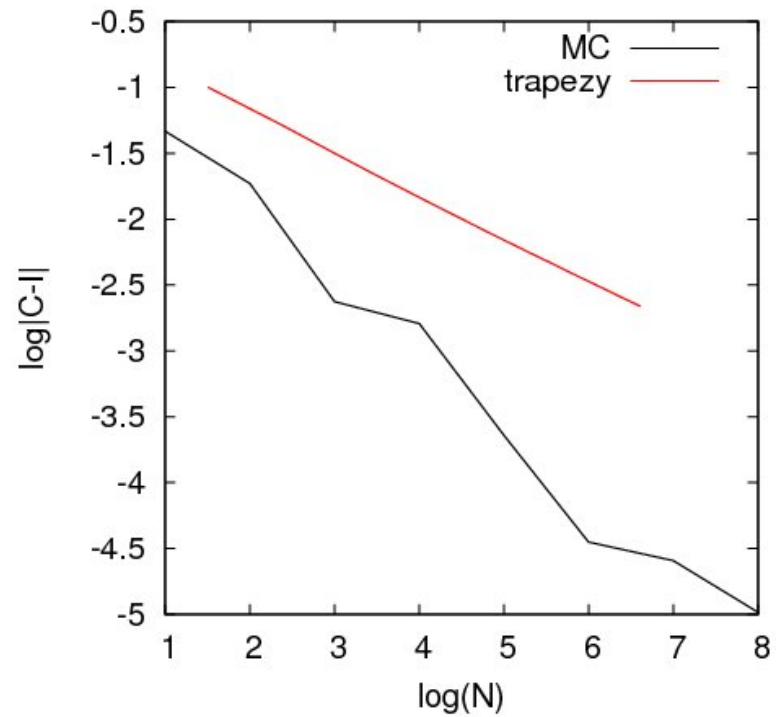
$$\Omega = \Delta^5$$

$$\Delta = 1$$

- objętość obszaru regularnego obliczeniowego (kostka 5-wymiarowa)

gdzie: $\mathbf{X} = [X_1, X_2, X_3, X_4, X_5]$

jest wektorem, którego składowe są zmiennymi losowymi



Rys. Wykres błędów oszacowania wartości całki w zależności od liczby węzłów (trapezy)/losowań (MC).

C - wartość dokładna całki

I - wartość całki wyznaczona numerycznie

Metody zwiększania efektywności/wydajności metody Monte Carlo

$$I = \int_{\Omega} z(\mathbf{x}) f(\mathbf{x}) d\mathbf{x}$$

- dokładność wyznaczenia całki metodą MC zależy od:
 - bezpośrednio od liczby próbek N (wzór na wariancję średniej)
 - sposobu liczenia wariancji zmiennej losowej
- **co oznacza wydajność?** - to dokładniejszy wynik (mniejszy błąd) uzyskany dla tej samej liczby losowań zmiennej niezależnej
- **wydajność metody** można zwiększyć ustalając N i dokonując takiej transformacji funkcji podcałkowej, aby nowa zmienna losowa (funkcja) miała mniejszą wariancję

Metoda losowania ważonego

Pierwotna postać całki

$$I = E(z) = \int_V G(\mathbf{x}) f(\mathbf{x}) d\mathbf{x}$$

Pod całkę wprowadzamy nową fgp $f^*(\mathbf{x})$

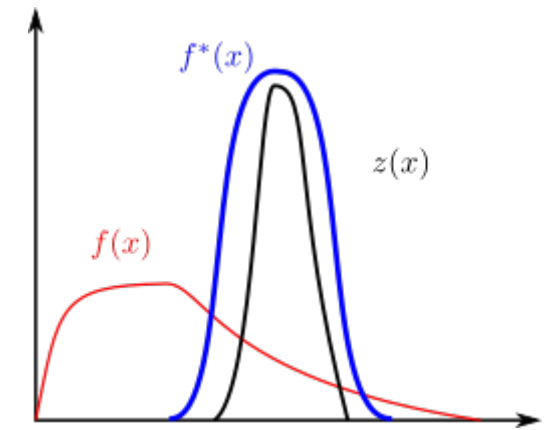
$$I = E(z) = \int_V z(\mathbf{x}) \frac{f(\mathbf{x})}{f^*(\mathbf{x})} f^*(\mathbf{x}) d\mathbf{x}$$

i wprowadzamy nową zmienną losową y

$$y(\mathbf{x}) = z(\mathbf{x}) \frac{f(\mathbf{x})}{f^*(\mathbf{x})}$$

wówczas całkę I możemy obliczyć losując wektor \mathbf{x} z rozkładem $g_{\mathbf{x}}(\mathbf{x})$ i sumując uzyskane wartości y_n

$$I = \frac{1}{N} \sum_{i=1}^N y(\mathbf{x}_i), \quad \mathbf{x}_i \sim f^*(\mathbf{x})$$



$$f^*(\vec{x}) \geq 0, \quad \int_V f^*(\vec{x}) d^n x = 1$$

Zmienna losowa y ma taką samą wartość oczekiwaną jak zmienna losowa G oraz wariancję zależną od nowej fgp $g_{\mathbf{x}}(\mathbf{x})$.

Wariancję estymatora całki można zmniejszyć odpowiednio dobierając nową fgp.

przykład losowania ważonego

$$\langle z \rangle = \int_0^1 \cos\left(\frac{\pi x}{2}\right) dx, \quad f(x) = \frac{1}{\Omega}, \quad \Omega = 1$$

$$\sigma_z^2 = \langle z^2 \rangle - \langle z \rangle^2 = \frac{1}{2} - \left(\frac{2}{\pi}\right)^2 = 0.0947152$$

$$\cos\left(\frac{\pi x}{2}\right) = 1 - \frac{\pi^2 x^2}{8} + \frac{\pi^4 x^4}{384} - \dots$$

Na podstawie rozwinięcia proponujemy funkcję nieujemną i znormalizowaną

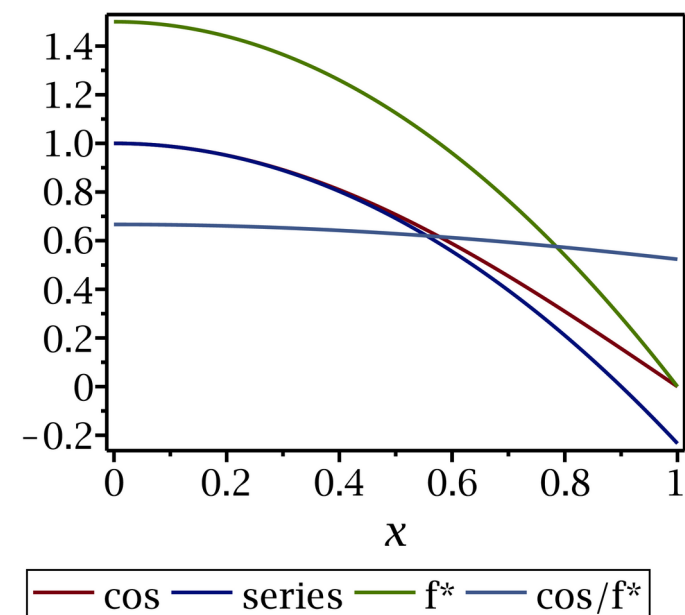
$$f^*(x) = \frac{3}{2}(1 - x^2), \quad \int_0^1 f^*(x) dx = 1$$

$$y(x) = \frac{z(x)}{f^*(x)} = \frac{2 \cos\left(\frac{\pi x}{2}\right)}{3(1 - x^2)}$$

$$\sigma_y^2 = \langle y^2 \rangle - \langle y \rangle^2 = 0.00099083$$

wskaźnik redukcji wariancji

$$\frac{\sigma_z^2}{\sigma_y^2} \approx 96$$



Metoda zmiennej kontrolnej

- metoda polega na dekompozycji całki

$$I = \int_V \left[G(\mathbf{x}) + \underbrace{\hat{G}(\mathbf{x}) - \hat{G}(\mathbf{x})}_{=0} \right] f(\mathbf{x}) d\mathbf{x} = \int_V \hat{G}(\mathbf{x}) f(\mathbf{x}) d\mathbf{x} + \int_V [G(\mathbf{x}) - \hat{G}(\mathbf{x})] f(\mathbf{x}) d\mathbf{x}$$

gdzie: $\hat{G}(\mathbf{x})$ jest aproksymacją funkcji $G(\mathbf{x})$ umożliwiającą łatwe obliczenie pierwszego wyrazu po prawej stronie (**analitycznie** lub **numerycznie**)

- wariancja zmiennej losowej

$$y = G(\mathbf{x}) - \hat{G}(\mathbf{x})$$

ma mniejszą wartość niż pierwotna zmienna (funkcja) $G(\mathbf{x})$

Przykład zastosowania metody zmiennej kontrolnej

- należy oszacować wartość całki

$$\langle g \rangle = \int_1^2 \ln(x) dx = 2 \ln(2) - 1 = 0.386294 \quad \langle g^2 \rangle = 0.1883 \quad \sigma_g = \langle g^2 \rangle - \langle g \rangle^2 = 0.03909$$

- wybieramy funkcję $h(x)=x$

$$\langle z \rangle = \int_1^2 [\ln(x) - x] dx = -1.1137 \quad \langle z^2 \rangle = 1.24906 \quad \sigma_z = \langle z^2 \rangle - \langle z \rangle^2 = 0.00872$$

- krotność redukcji wariancji $\frac{\sigma_g}{\sigma_z} = 4.48$

- wyberzmy inną funkcję $h(x)=x-1$

$$\langle z \rangle = \int_1^2 [\ln(x) - (x - 1)] dx = -0.1137$$

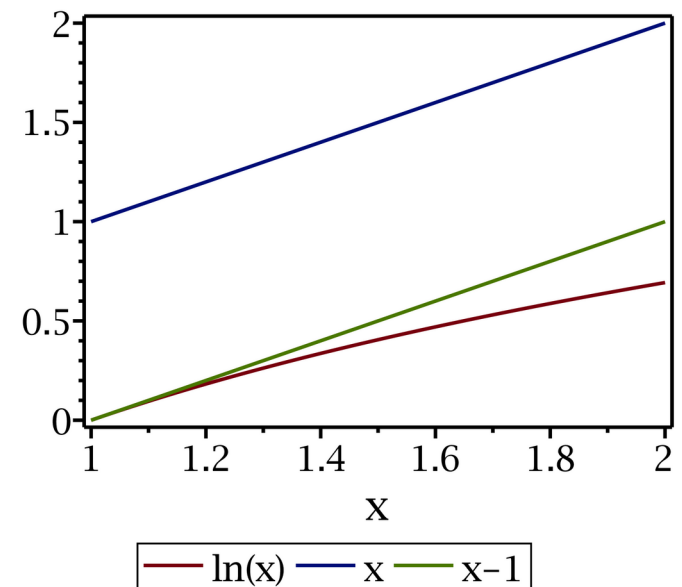
$$\langle z^2 \rangle = 0.02165$$

$$\sigma_z = \langle z^2 \rangle - \langle z \rangle^2 = 0.00872$$

- wynik nie uległ zmianie

$$\frac{\sigma_g}{\sigma_z} = 4.48$$

- stopień redukcji wariancji zależy od lokalnych zmian wartości $|g(x)-h(x)|$ a dokładniej: jak dobrze $h(x)$ aproksymuje $g(x)$
- wartość $|g(x)-h(x)|$ jest mniej istotna (patrz rysunek)



Metoda MC wymaga zastosowania generatora liczb pseudolosowych o zadanym rozkładzie gęstości prawdopodobieństwa. Generatory (a raczej ciągi generowanych liczb) muszą spełniać określone warunki (korelacja, okres, fgp itp.).

Zastosowania metody Monte Carlo

Ogólnie: w problemach nauki/techniki, w których zawodzą metody deterministyczne

- a) **sumulacja komputerowa** probabilistycznego modelu matematycznego/fizycznego (np. model Isinga, rozpraszanie neutronów w reaktorze, symulacja zmian wartości akcji w czasie).
- b) **obliczanie wartości całek wielokrotnych** (obliczanie objętości, momentów bezwładności itp. obiektów o nieregularnym kształcie, całki oddziaływania w chemii/fizyce kwantowej)
- c) **optymalizacja** (np. minimalizacja czasu oczekiwania pacjenta w kolejce do lekarza, problem komiwojażera, poszukiwanie minimum energii układu - stan stacjonarny)
- d) **rozwiązywanie równań różniczkowych zależnych** i niezależnych od czasu (np. równanie Poissona metodą błędzenia przypadkowego, dyfuzja w obszarze o skomplikowanej geometrii)

przykład - obliczanie momentu bezwładności kuli

obszar kuli: $(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2 \leq R^2$

def. momentu
bezwładności:
$$I = \int_M \rho^2 dm$$

$$I = \gamma \int_V (\vec{\rho} - \vec{\rho}_0)^2 dV$$

γ - ciężar właściwy materiału kuli

$$I = \frac{V_s \gamma}{N} \sum_{i=1}^N (\vec{\rho}_i - \vec{\rho}_0)^2 \theta_i$$

Wariancja:
$$\sigma_I^2 = \frac{1}{N} \left[\sum_{i=1}^N (V_s \gamma (\vec{\rho}_i - \vec{\rho}_0)^2 \theta_i)^2 - \frac{1}{N} \left(\sum_{i=1}^N V_s \gamma (\vec{\rho}_i - \vec{\rho}_0)^2 \theta_i \right)^2 \right]$$

