

# Całkowanie metodą Monte Carlo

Plan wykładu:

1. Podstawowa metoda Monte Carlo
2. Metody MC o zwiększonej efektywności
  - a) losowania ważonego
  - b) zmiennej kontrolnej
  - c) losowania warstwowego
  - d) obniżania krotności całki

Funkcja gęstości prawdopodobieństwa zmiennej losowej  $x$ .

Funkcja ta ma następujące własności:

$$\bigwedge_{x \in [a, b]} f(x) \geq 0$$

$$\int_a^b f(x) dx = 1$$

Przy jej pomocy można określić prawdopodobieństwo zdarzenia że zmienna  $x$  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

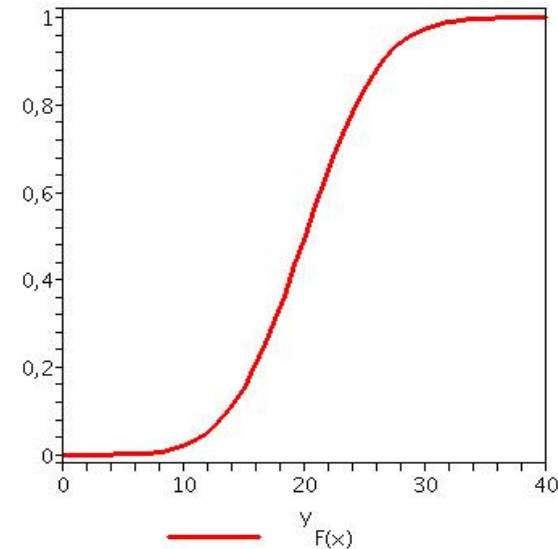
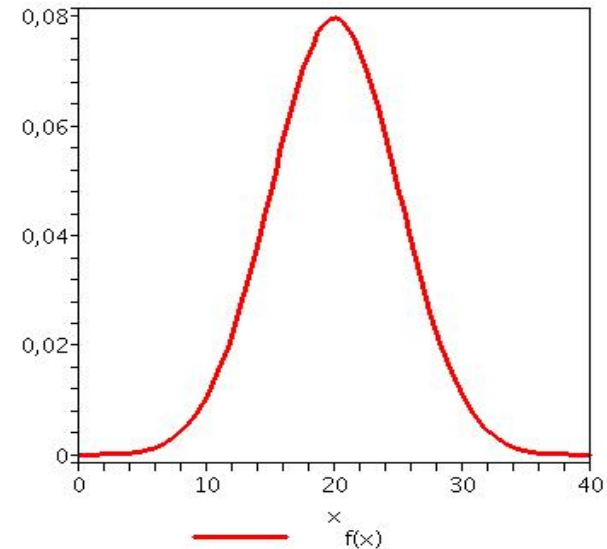
$$F(x) = \int_{-\infty}^x f(x') dx'$$

która jest funkcją prawostronnie ciągłą i niemalejącą.

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

Przykład. Rozkład normalny (Gaussa)

$$f(x) = \frac{1}{2\sqrt{\pi}} \exp \left( -\frac{(x - \mu)^2}{2\sigma^2} \right)$$



- wartość oczekiwana zmiennej losowej  $x$

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

- wariancja zmiennej losowej  $x$

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

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

- odchylenie standardowe

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

Wartość oczekiwana  $\mu$  oraz odchylenie standardowe  $\sigma$  są parametrami funkcji gęstości prawdopodobieństwa  $f(x)$ .

W podobny sposób możemy określić wartość oczekiwaną funkcji, której argumentem jest zmienna losowa  $x$  o funkcji gęstości prawdopodobieństwa  $f(x)$

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

i analogicznie jej wariancję

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

oraz 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  $f(x)$  to estymatorem wartości oczekiwanej  $\mu(z)$  zmiennej losowej  $z(x_i)$  jest średnia z próbki

$$\bar{z} = \frac{1}{N} \sum_{i=1}^N z(x_i)$$

z wariancją

$$s^2(z) = \frac{1}{N-1} \sum_{i=1}^N [z(x_i) - \bar{z}]^2$$

Uwaga:

(N-1) w mianowniku wynika z faktu że średnią wyliczamy z N wartości  $z(x_i)$  - znając jej wartość możemy wyliczyć dowolną  $z(x_i)$  dysponując N-1 pozostałymi wartościami. Liczba stopni swobody zmniejsza się o 1. W praktyce, dla dużych N jedynkę można pominąć.

$$\begin{aligned} s^2(z) &= \frac{1}{N-1} \sum_{i=1}^N (z_i - \bar{z})^2 \\ &= \frac{1}{N-1} \left( \sum_{i=1}^N z_i^2 - \frac{1}{N} \left( \sum_{i=1}^N z_i \right)^2 \right) \end{aligned}$$

Miarą „rozrzutu” zmiennych losowych  $z_i$  wokół wartości średniej jest odchylenie standardowe

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

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?

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

Do jego estymacji możemy użyć  $s(z)$

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

## Podstawowa metoda 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 funkcja gęstości  $g(z)$

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

a rozkład prawdopodobieństwa wektora  $\mathbf{x}$  opisuje funkcja gęstości  $f(\mathbf{x})$

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

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

Przy takich założeniach, zgodnie z CTG

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

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

a) wartość całki

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

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

b) błąd oszacowania

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

Zazwyczaj obszarem całkowania jest określony podzbiór przestrzeni  $R^M$ . W takim przypadku obliczaną całkę trzeba zapisać w nieco zmienionej postaci:

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

gdzie:  $\mathbf{1}_V(\mathbf{x})$

jest funkcją przynależności do zbioru  $V \subset \Omega$

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

### Kwadratura Monte Carlo (metoda orzeł-reszka)

$$I = \int_V z(\mathbf{x}) d\mathbf{x} = \int_{\Omega} \mathbf{1}_V(\mathbf{x}) z(\mathbf{x}) d\mathbf{x} \approx \frac{\Omega}{N} \sum_{i=1}^N \mathbf{1}_V(\mathbf{x}) z(\mathbf{x})$$

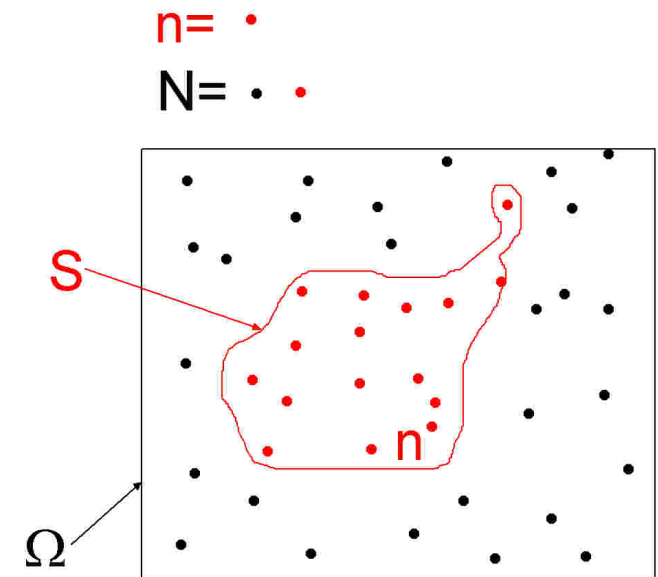
Uwagi:

a) w powyższym przypadku zakładamy, że funkcja gęstości prawdopodobieństwa jest stała w obszarze  $\Omega$

b) wydajność metody zależy od stosunku wielkości obszaru  $V$  i obszaru  $\Omega$ .

### Przykład

Wyznaczyć pole powierzchni obiektu o nieregularnym kształcie.



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

$$S = \frac{\Omega}{N} \sum_{i=1}^N \mathbf{1}_V = \Omega \frac{n}{N}$$

### Przykład

Należy obliczyć numerycznie wartość całki

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

a) metoda trapezów

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

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

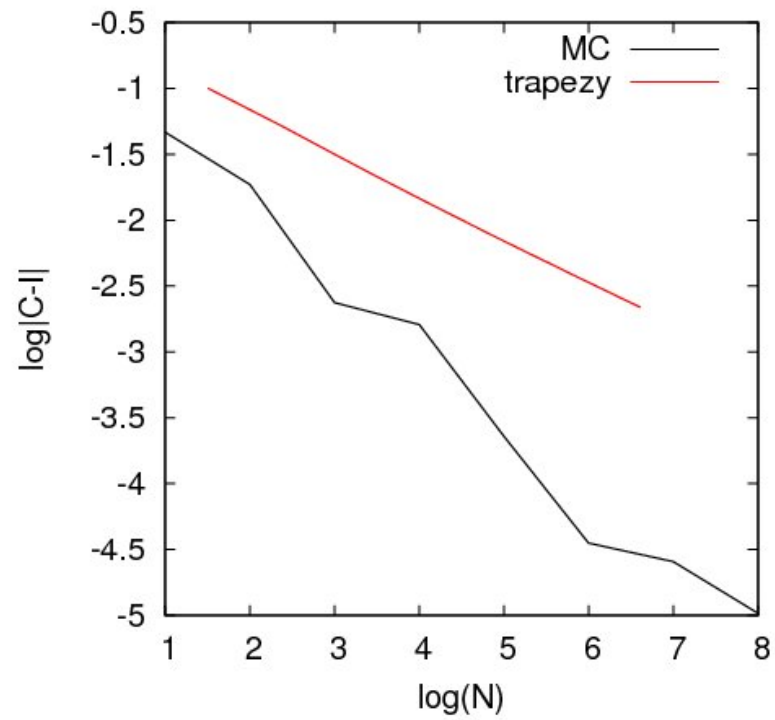
$$h = \frac{1}{n} \quad w_{i,j,k,l,m} = \begin{cases} \frac{1}{2} & i, j, k, l, m = 0, n \\ 1 & i, j, k, l, m = 1, 2, \dots, n-1 \end{cases}$$

b) Kwadratura Monte Carlo

$$I_{MC} = \frac{h^5}{N} \sum_{i=1}^N g(\mathbf{X})$$

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

jest wektorem, którego składowe są zmiennymi losowymi



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



### Przykład

Dzielnik napięcia powinien zapewniać tłumienie o wartości 0.5 z dokładnością 2%. Opory  $r_1$  i  $r_2$  mają rozrzuty produkcyjne które można reprezentować za pomocą niezależnych zmiennych

$$r_1 \quad r_2$$

o funkcjach gęstości prawdopodobieństwa

$$f_{r_1}(r_1) \quad f_{r_2}(r_2)$$

Wyznaczyć estymatę uzysku produkcyjnego  $\eta$ , czyli średniego odsetka układów sprawnych.

Tłumienie napięciowe dzielnika:

$$k = \frac{r_1}{r_1 + r_2}$$

Tłumienie jest realizacją zmiennej losowej:

$$k = \frac{r_1}{r_1 + r_2}$$

Rozkład tej zmiennej opisuje fgp:  $f_k(k)$

zależna od  $f_{r_1}(r_1) \quad f_{r_2}(r_2)$

Warunkiem sprawności układu (jednej z wielu realizowanych możliwości) jest :

$$k \in V$$

$$V = [0.49, 0.51]$$

Wykorzystujemy metodę MC do estymacji wartości oczekiwanej:

$$\eta = \int_{-\infty}^{\infty} \mathbf{1}_V(k) f_k(k) dk$$

$\underline{k}$  jest funkcją wektora losowego:

$$\mathbf{r} = [r_1, r_2]^T$$

dlatego uzysk produkcyjny można wyrazić wzorem na średnią wartość funkcji przynależności:

$$\eta = \int_{R^2} \mathbf{1}_V(k(\mathbf{r})) f_{\mathbf{r}}(\mathbf{r}) d\mathbf{r}$$

gdzie:

$$f_{\mathbf{r}} = f_{r_1}(r_1) f_{r_2}(r_2)$$

jest iloczynem ze względu na niezależność zmiennych losowych  $r_1$  i  $r_2$ .

Estymatę uzysku można obliczać jako średnią arytmetyczną

$$\hat{\eta} = \frac{1}{N} \sum_{n=1}^N \mathbf{1}_V(k(\mathbf{r}_n))$$

gdzie:

$$\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \dots$$

są niezależnymi realizacjami wektora losowego  $\mathbf{r}$

Algorytm wyznaczenia uzysku:

- 1) Wylosuj parę liczb:  $r_1$  i  $r_2$ , zwiększ  $N$  o 1
- 2) Jeśli obliczone  $k$  mieści się w obszarze  $V$  wówczas zwiększ  $N_s$  o 1
- 3) Uzysk oblicz jako wartość ułamka

$$\eta = \frac{N_s}{N}$$

### Przykład.

Wyznaczyć minimalną liczbę  $N$  próbek wystarczającą do wyznaczenia estymaty uzysku z trzysigmowym błędem względnym:

$$\delta = \frac{3\sigma_{\hat{\eta}}}{\hat{\eta}}$$

Dla  $\delta = 0.1\%, 1\%, 10\%$ .

Obliczamy wariancję estymatora:

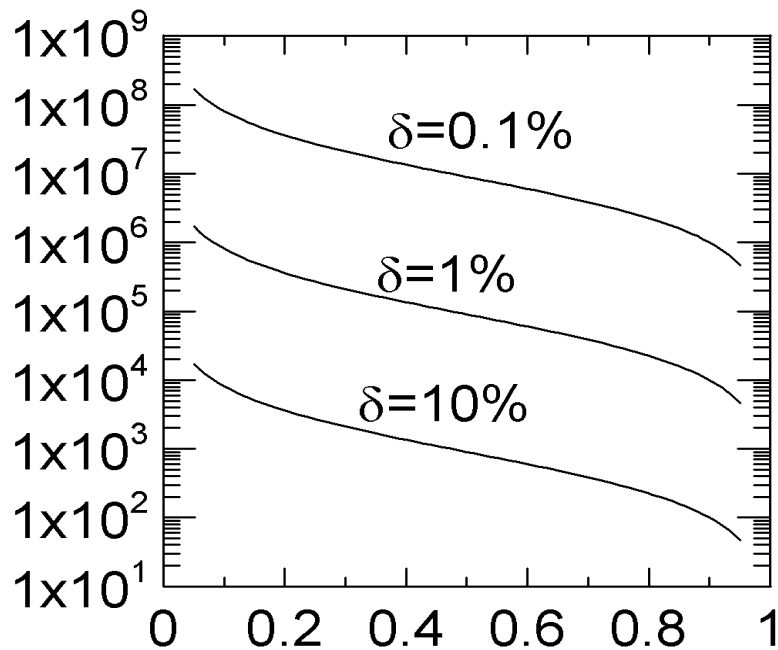
$$\begin{aligned} \sigma_{\hat{\eta}}^2 &= \frac{1}{N(N-1)} \left( \sum_{n=1}^N (\mathbf{1}_V(k(\mathbf{r}_n)))^2 \right. \\ &\quad \left. - \frac{1}{N} \left( \sum_{n=1}^N \mathbf{1}_V(k(\mathbf{r}_n)) \right)^2 \right) \\ &= \frac{\hat{\eta}(1 - \hat{\eta})}{N - 1} \end{aligned}$$

Błąd względny:

$$\delta = 3\sqrt{\frac{1 - \hat{\eta}}{(N - 1)\hat{\eta}}}$$

Przekształcając go można otrzymać wyrażenie na minimalną liczbę próbek potrzebną do uzyskania wymaganej dokładności:

$$N = \frac{1 - \hat{\eta}}{\hat{\eta}} \left( \frac{3}{\delta} \right)^2$$



Rys. Zależność minimalnej liczby próbek od założonego uzysku

## Metody zwiększania efektywności metody Monte Carlo

$$I = \int_{\Omega} G(\mathbf{x})f(\mathbf{x})d\mathbf{x} = \int_{R^M} \mathbf{1}_V(\mathbf{x})G(\mathbf{x})f(\mathbf{x})d\mathbf{x}$$

Dokładność wyznaczenia całki metodą MC zależy od liczby próbek  $N$  oraz wariancji zmiennej losowej:

$$z = \mathbf{1}_V(\mathbf{x})G(\mathbf{x})$$

Wydajność metody można zwiększyć ustalając  $N$  i dokonując takiej transformacji aby nowa zmienna losowa miała mniejszą wariancję.

## a) Metoda losowania ważonego

Zakładamy że  $g_{\mathbf{x}}(\mathbf{x})$  jest fgp dodatnio określoną dla

$$\mathbf{x} \in V$$

$$I = E(z) = \int_V \left\{ G(\mathbf{x}) \frac{f_{\mathbf{x}}(\mathbf{x})}{g_{\mathbf{x}}(\mathbf{x})} g_{\mathbf{x}}(\mathbf{x}) d\mathbf{x} \right\}$$

$$y = \mathbf{1}_V G(\mathbf{x}) f_{\mathbf{x}} / g_{\mathbf{x}}$$

Całkę estymujemy:

$$I = \frac{1}{N} \sum_{n=1}^N y(\mathbf{x}_n)$$

Zmienna losowa  $z$  ma taką samą wartość oczekiwaną jak zmienna losowa  $y$  oraz wariancję zależną od fgp:

$$g_{\mathbf{x}}(\mathbf{x})$$

Wariancję estymatora całki można zmniejszyć odpowiednio dobierając fgp.

Najmniejszą wartość wariancja osiąga dla:

$$g_{\mathbf{x}}(\mathbf{x}) = \frac{\mathbf{1}_V |G(\mathbf{x})| f_{\mathbf{x}}(\mathbf{x})}{\int_V |G(\mathbf{x})| f_{\mathbf{x}}(\mathbf{x}) d\mathbf{x}}$$

Jeżeli  $G(\mathbf{x})$  jest funkcją nieujemną, wówczas minimalna wariancja estymatora ważonego jest równa 0. Należałoby jednak w takim przypadku znać wartość całki w mianowniku. Zazwyczaj nie jest to możliwe, dlatego funkcję  $G(\mathbf{x})$  zastępuje się inną  $G_1(\mathbf{x})$ , której całka może być łatwo obliczona. Minimalizacja wariancji w takim przypadku zależy od jakości zastosowanego przybliżenia.

## b) Metoda zmiennej kontrolnej.

Metoda polega na dekompozycji całki:

$$I = \int_V \hat{G}(\mathbf{x}) f_{\mathbf{x}} d\mathbf{x} + \int_V [G(\mathbf{x}) - \hat{G}(\mathbf{x})] f_{\mathbf{x}} d\mathbf{x}$$

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 znacznie mniejszą wariancję niż  $G(\mathbf{x})$ .

## c) Losowanie warstwowe

W metodzie tej obszar całkowania  $V$  dzieli się na  $K$  rozłącznych podobszarów:

$$V_1, V_2, \dots, V_k$$

Całkę  $I$  oblicza się jako sumę całek w podobszarach.

$$\begin{aligned} I_k &= \int_{V_k} G(\mathbf{x}) f_{\mathbf{x}}(\mathbf{x}) d\mathbf{x} \\ &= \mu(V_k) \int_{V_k} G(\mathbf{x}) f_{\mathbf{x}} / \mu(V_k) d\mathbf{x} \end{aligned}$$

$$\mu(V_k) = \int_{V_k} f_{\mathbf{x}}(\mathbf{x}) d\mathbf{x}$$

gdzie:  $k=1,2,3,\dots,K$

Całki  $I_k$  można obliczać za pomocą podstawowej wersji metody MC

$$\hat{I}_k = \frac{\mu(V_k)}{N_k} \sum_{n=1}^{N_k} \mathbf{1}_{V_k}(\mathbf{x}_n^{(k)}) G(\mathbf{x}_n^{(k)})$$

Próbki  $\{\mathbf{x}_n^{(k)} | n = 1, 2, \dots, N_k\}$

są realizacjami wektora losowego  $\mathbf{x}$  o fgp

$$f_{\mathbf{x},k}(\mathbf{x}) = \frac{\mathbf{1}_{V_k} f_{\mathbf{x}}(\mathbf{x})}{\mu(V_k)}$$

#### d) Metoda obniżania krotności całki

Obniżenia krotności całki można dokonać gdy jest możliwa dekompozycja wektora oryginalnych zmiennych losowych:

$$\mathbf{x}^T = [\mathbf{u}^T \mathbf{v}^T]$$

oraz obszaru

$$V = V_u \times V_v$$

że zachodzi

$$f_{\mathbf{x}}(\mathbf{x}) = f_{\mathbf{u}}(\mathbf{u})f_{\mathbf{v}}(\mathbf{v})$$

$$\mathbf{u} \in V_u$$

$$\mathbf{v} \in V_v$$

$$I = \int_{V_u} \left\{ \int_{V_v} G(\mathbf{x}(\mathbf{u}, \mathbf{v})) f_{\mathbf{v}}(\mathbf{v}) d\mathbf{v} \right\} f_{\mathbf{u}}(\mathbf{u}) d\mathbf{u}$$

Zmienna losowa

$$z = \int_{V_v} G(\mathbf{x}(\mathbf{u}, \mathbf{v})) f_{\mathbf{v}}(\mathbf{v}) d\mathbf{v}$$

ma zazwyczaj mniejszą wariancję niż  $G(\underline{x})$  co pozwala dość łatwo obliczyć całkę zewnętrzną. Metoda jest skuteczna jeśli potrafimy dość dokładnie i szybko obliczyć całkę wewnętrzną (analitycznie lub numerycznie).

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

- a) symulacja komputerowa probabilistycznego modelu matematycznego/fizycznego (kwantowa dyfuzyjna metoda MC).
- b) Obliczanie wartości całek wielokrotnych (obliczanie objętości, momentów bezwładności itp. obiektów o nieregularnym kształcie)
- c) Optymalizacja (minimalizacja czasu oczekiwania pacjenta w kolejce do lekarza)
- d) Rozwiązywanie równań różniczkowych (rów. Poissona metodą błędzenia przypadkowego ze stałym lub zmiennym krokiem)