Projekt nr B.4.2

Generacja liczb przypadkowych przy użyciu metody Metropolisa

Wprowadzenie

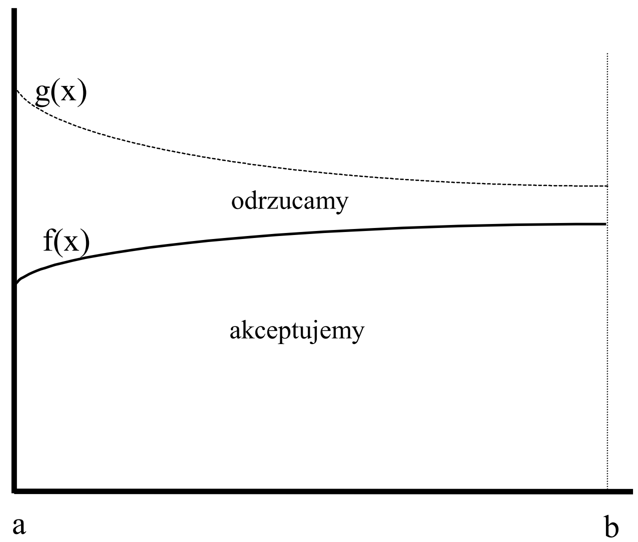
Potrzeba generowania zmiennych losowych podlegających pewnemu, z góry ustalonemu rozkładowi, pojawia się niezwykle często w zagadnieniach rozwiązywanych metodami Monte Carlo. Ze względu na znaczenie praktyczne tego problemu pokrótce omówione zostaną podstawowe metody stosowane w takich przypadkach.

Metoda losowania z odwróconej dystrybuanty

W przypadkach najprostszych, kiedy zmienna losowa opisywana jest rozkładem jednowymiarowym, skuteczna bywa metoda losowania z odwróconej dystrybuanty. Dla znanej funkcji gęstości prawdopodobieństwa $f(x)$ wyznacza się wówczas analitycznie lub numerycznie dystrybuantę $y = F(x)$, a następnie znajduje się funkcję odwrotną $x = F^{-1}(y)$. Jak można zauważyć, dziedziną ostatniej funkcji jest przedział (0,1). Stosując dowolny generator o rozkładzie równomiernym w tym przedziale, możemy wylosować kolejną wartość y_n: y_n = rand(), a następnie wyznaczyć $x_n = F^{-1}(y_n)$.

Metoda von Neumanna

Użyteczną metodą generowania jedno- i wielowymiarowych zmiennych losowych jest metoda eliminacji von Neumanna. Jej zaleta polega na tym, że nie ma potrzeby wyznaczania (i odwracania) dystrybuant rozkładów, a w miejsce tego stosowane jest dodatkowe losowanie. Załóżmy dla prostoty, że wyznaczać chcemy jednowymiarową zmienną losową $x \in (a,b)$ podlegającą rozkładowi $f(x)$.



Algorytm omawianej metody jest wówczas następujący:

1. W przedziale (a,b) określamy taki rozkład $g(x)$, że dla każdego $x \in (a,b)$: $g(x) > f(x)$. Zauważmy, że rozkład $g(x)$ może (choć nie musi) być równomierny.
2. Stosując dowolny generator o rozkładzie równomiernym w przedziale (a,b) losujemy nowy punkt x_n, a następnie wyznaczamy $w_n = f(x_n)/g(x_n)$. Oczywiście, $0 < w_n < 1$.
3. Stosując dowolny generator o rozkładzie równomiernym w przedziale (0,1) losujemy liczbę losową r_n.
4. Jeżeli $r_n < w_n$, punkt x_n przyjmujemy, jeżeli nie - odrzucamy.

Powyższa metoda działa analogicznie w przypadku rozkładów wielowymiarowych. Jak można zauważyć, jest ona efektywna, o ile liczba odrzuceń nie jest zbyt duża, a więc jeśli kształt rozkładu $g(\boldsymbol{x})$ jest zbliżony do kształtu rozkładu $f(\boldsymbol{x})$. Niestety, w praktyce sytuacja taka nie zdarza się często - w szczególności w przypadku rozkładów wielowymiarowych o ostrych maksimach lokalnych.

Metoda Metropolisa

Metoda opublikowana w 1953 roku przez Metropolisa jest powszechnie używana w przypadku rozkładów, dla których metoda von Neumanna okazuje się nieefektywna. W odróżnieniu od tej ostatniej, nie wymaga ona określania pomocniczego rozkładu $g(\boldsymbol{x})$, a zatem opiera się wyłącznie na zadanym rozkładzie $f(\boldsymbol{x})$. Podany niżej algorytm, dla przejrzystości, odnosi się do problemu jednowymiarowego. Jednak rozszerzenie go na zagadnienie wielowymiarowe jest oczywiste.

1. Wybieramy punkt startowy x_n. Możemy to zrobić losowo, albo preferując jakiś konkretny punkt. W punkcie tym wyznaczamy $f(x_n)$.
2. Wyznaczamy nowy punkt x_t losując równomiernie w otoczeniu x_n: $x_n - \Delta x \leq x_t \leq x_n + \Delta x$. W punkcie tym wyznaczamy $f(x_t)$.
3. Jeżeli $f(x_t) \geq f(x_n)$ to akceptujemy nowy punkt i przyjmujemy $x_{n+1} = x_t$
4. Jeżeli $f(x_t) < f(x_n)$, generujemy liczbę losową r z przedziału (0,1):
   - jeżeli $r < f(x_t)/f(x_n)$, akceptujemy nowy punkt i przyjmujemy $x_{n+1} = x_t$
   - jeżeli $r \geq f(x_t)/f(x_n)$, odrzucamy nowy punkt i przyjmujemy $x_{n+1} = x_n$
5. x_{n+1} jest nowym wylosowanym punktem (może być $x_{n+1} = x_n$).

Po wykonaniu odpowiednio dużej liczby kroków procedura doprowadza do zbioru punktów x zgodnego z założonym rozkładem $f(x)$.

Zadania do wykonania

1. Napisać procedurę generującą liczby pseudolosowe z jednowymiarowego rozkładu normalnego stosując algorytmy von Neumanna i Metropolisa. Przeprowadzając odpowiednie testy statystyczne (jakie?) sprawdzić, czy dla zadanego poziomu istotności uzyskane wyniki są zgodne z rozkładem normalnym. Porównać efektywność obu metod, rozumianą jako stosunek liczby zaakceptowanych zmiennych losowych do liczby wszystkich wygenerowanych zmiennych losowych (tj. zaakceptowanych + odrzuconych).

2. Napisać procedurę generującą liczby pseudolosowe z dowolnego rozkładu jednowymiarowego opisanego przy pomocy histogramu, wykorzystując algorytmy von Neumanna i Metropolisa. Porównać efektywność obu metod.

3. Napisać procedurę generującą liczby pseudolosowe z dwuwymiarowego rozkładu normalnego stosując algorytmy von Neumanna i Metropolisa. Porównać efektywność obu metod,

Literatura

1. Wykład
2. Koonin, Computational Physics