

Metody Monte Carlo – wzory

Maciej Romaniuk*

29 lipca 2014

1 Wybrane rozkłady prawdopodobieństwa

1.1 Rozkład dwupunktowy

Doświadczenie losowe ma tylko dwa możliwe wyniki, zazwyczaj zapisywane jako „1” i „0” (tak / nie, sukces / porażka, prawidłowy, nieprawidłowy, itd.). Prawdopodobieństwo „sukcesu” jest równe p , gdzie oczywiście $0 \leq p \leq 1$. Stąd rozkład prawdopodobieństwa dany jest wzorem

$$P(X = 1) = p, P(X = 0) = 1 - p. \quad (1)$$

Ważniejsze charakterystyki: $EX = p$, $\text{Var } X = p(1 - p)$.

Zastosowanie: rzut monetą, kontrola sprawności pojedynczego elementu na linii produkcyjnej, zaliczenie egzaminu.

1.2 Rozkład dwumianowy

Założmy, że mamy n niezależnych powtórzeń takiego doświadczenia losowego, które ma tylko dwa możliwe wyniki (zwane tradycyjnie **porażką** i **sukcesem**). Oznacza to, że n razy powtarzamy doświadczenie z rozkładu dwupunktowego. Przez p , jak poprzednio, oznaczmy prawdopodobieństwo sukcesu w pojedynczej próbie. Wtedy prawdopodobieństwo zajścia k sukcesów w n próbach (czyli zdarzenia $X = k$) określone jest wzorem

$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}. \quad (2)$$

Ważniejsze charakterystyki: $EX = np$, $\text{Var } X = np(1 - p)$. Tradycyjnie rozkład ten zapisujemy skrótowo $\text{Bin}(n; p)$.

Zastosowanie: wielokrotne rzuty monetą, kontrola sprawności n elementów na linii produkcyjnej, strzelanie do tarczy (trafienie / pudło).

1.3 Rozkład geometryczny

Założmy, że wykonujemy niezależne powtórzenia doświadczenia losowego, które ma tylko dwa możliwe wyniki, **aż do osiągnięcia sukcesu**. Przez p oznaczmy prawdopodobieństwo zajścia sukcesu w pojedynczej próbie. Wtedy

* e-mail: mroman@ibspan.waw.pl

liczba wykonanych doświadczeń ma rozkład geometryczny. Niech X będzie tą liczbą prób do momentu zajścia pierwszego sukcesu. Prawdopodobieństwo zdarzenia $X = k$ (czyli na początku nastąpiło $k - 1$ porażek, a potem pierwszy sukces) dane jest wzorem

$$P(X = k) = p(1 - p)^{k-1}, \quad (3)$$

gdzie $k = 0, 1, \dots$. Ważniejsze charakterystyki: $\mathbb{E} X = 1/p$, $\text{Var } X = (1 - p)/p^2$.

Zastosowanie: liczba rzutów monetą do momentu wypadnięcia pierwszego orła, liczba elementów na taśmie produkcyjnej zanim nie natrafimy na wadliwy, liczba wypełnionych losów TotoLotka, zanim po raz pierwszy nie trafimy „szóstki”.

1.4 Rozkład Poissona

Jeśli zmienna pochodzi z rozkładu Poissona, to jej rozkład prawdopodobieństwa opisany jest wzorem

$$P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda} \quad (4)$$

dla $k = 0, 1, \dots$, gdzie $\lambda > 0$ jest parametrem tego rozkładu. Tradycyjnie rozkład ten oznaczamy skrótem $\text{Poiss}(\lambda)$. Ważniejsze charakterystyki: $\mathbb{E} X = \lambda$, $\text{Var } X = \lambda$.

Zastosowanie: ilość wypadków, do których doszło w pewnym ustalonym przedziale czasowym, liczba cząstek wyemitowanych przez radioaktywny materiał w pewnym przedziale czasowym, liczba zgłoszeń klientów w sieci w pewnym okresie (np. w ciągu godziny). Rozkład ściśle związany z rozkładem wykładniczym.

Jeśli dla ustalonego λ stworzymy wykres funkcji $P(X = k)$ względem k (czyli wykres funkcji prawdopodobieństwa), będzie on malejący (tzn. wraz ze wzrostem k odpowiednie słupki pokazujące prawdopodobieństwo będą coraz niższe).

1.5 Rozkład jednostajny (równomierny)

Najprostszy z ciągłych rozkładów prawdopodobieństwa, oznaczany zazwyczaj skrótem $U[a; b]$. Jego gęstość na przedziale $[a; b]$ opisana jest wzorem

$$f(t) = \frac{1}{b - a}. \quad (5)$$

Oznacza to zatem, że prawdopodobieństwo zaobserwowania wartości zmiennej z dowolnego, małego przedziału o długości dx jest stałe i takie samo na całym przedziale $[a; b]$. Ważniejsze charakterystyki: $\mathbb{E} X = \frac{a+b}{2}$, $\text{Var } X = \frac{(b-a)^2}{12}$.

Zastosowanie: przy „równomierności” zdarzeń na przedziale losowym, np. przypadkowy wybór liczby z przedziału $[0; 100000]$.

Wykres funkcji gęstości dla tego rozkładu jest stały na przedziale $[a; b]$ i równy zero poza nim.

1.6 Rozkład wykładniczy

Zmienna losowa X pochodzi z rozkładu wykładniczego (co zapisujemy $X \sim \text{Exp}(\lambda)$), jeśli gęstość $f(\cdot)$ jest równa

$$f(t) = \lambda e^{-\lambda t} \quad (6)$$

dla $t \geq 0$ i $f(t) = 0$ dla $t < 0$. Ważniejsze charakterystyki: $\mathbb{E} X = \frac{1}{\lambda}$, $\text{Var} X = \frac{1}{\lambda^2}$.

Zastosowanie: jeśli w pewnym ustalonym przedziale czasowym liczba wystąpień jakiegoś zdarzenia jest zmienną z rozkładu Poissona, to okres czasu pomiędzy kolejnymi zdarzeniami jest właśnie zmienną z rozkładu wykładniczego, np. czas do wyemitowania radioaktywnej cząstki, czas do kolejnego zgłoszenia klienta w sieci.

Wykres funkcji gęstości dla tego rozkładu maleje (wykładniczo), począwszy od wartości przyjmowanej dla $t = 0$.

1.7 Rozkład normalny

Jeden z najważniejszych w statystyce rozkładów. Zmienna losowa X pochodzi z rozkładu normalnego (co zapisujemy $X \sim N(\mu, \sigma^2)$), jeśli gęstość $f(\cdot)$ jest równa

$$f(t) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(t-\mu)^2}{2\sigma^2}\right), \quad (7)$$

gdzie $\sigma > 0$. Parametr μ nazywamy wartością oczekiwaną (lub średnią), a σ^2 – wariancją.

Zastosowanie: bardzo różnorodne zastosowania, np. modelowanie wzrostu osób, ocen z egzaminu, cen akcji, ilości opadów, temperatury, itd. Ogólnie rzecz biorąc, rozkład ten stosuje się wtedy, gdy pewna zmienna losowa jest wynikiem sumarycznego działania wielu „małych” i „niezależnych” czynników.

Wykres funkcji gęstości dla tego rozkładu ma postać dzwonu o maksimum w $t = \mu$.

1.8 Rozkład t-Studenta (rozkład t)

Rozkład bardzo często wykorzystywany w wielu testach statystycznych. Zmienna losowa X pochodzi z rozkładu t-Studenta (w skrócie *rozkładu t*, co zapisujemy $X \sim t(n)$), jeśli gęstość $f(\cdot)$ ma postać

$$f(t) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\sqrt{n\pi}} \left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}}, \quad (8)$$

gdzie parametr $n \in \mathbb{N}_+$ zwany jest stopniami swobody (lub liczbą śladów).

Zastosowanie: testy statystyczne.

Ważniejsze charakterystyki (dla $n > 2$, dla mniejszej liczby stopni swobody niektóre momenty nie istnieją): $\mathbb{E} X = 0$, $\text{Var} X = \frac{n}{n-2}$.

Dla $n \rightarrow \infty$ wykres gęstości tego rozkładu coraz bardziej przypomina gęstość standardowego rozkładu normalnego.

1.9 Rozkład χ^2 (chi-kwadrat)

Rozkład bardzo często wykorzystywany w wielu testach statystycznych. Zmienna losowa X pochodzi z rozkładu chi-kwadrat, co zapisujemy $X \sim \chi^2(n)$, jeśli gęstość $f(\cdot)$ ma postać

$$f(t) = \frac{t^{\frac{n}{2}-1} e^{-\frac{t}{2}}}{2^{\frac{n}{2}} \Gamma(\frac{n}{2})} \quad (9)$$

dla $t > 0$, przy czym parametr $n \in \mathbb{N}_+$ zwany jest ilością śladów. Wykres gęstości ma postać „wolno przesuwaną się górki”.

Ważniejsze charakterystyki: $\mathbb{E} X = n$, $\text{Var } X = 2n$.

2 Generatory liniowe

Generatorem liniowym nazywamy generator opisany funkcją

$$X_{n+1} = (a_1 X_n + a_2 X_{n-1} + \dots + a_k X_{n-k+1} + c) \mod m \quad (10)$$

gdzie $a_1, a_2, \dots, a_k, c, m$ są liczbami całkowitymi, zwanymi *parametrami generatora*. Jeśli stała $c = 0$, to mówimy o **generatorze multiplikatywnym**, a dla $c \neq 0$ mówimy o **generatorze mieszanym**.

3 Okres

Lemat 1. *Generator multiplikatywny*

$$X_{n+1} = (a X_n) \mod m \quad (11)$$

ze stałą $m = 2^L$ dla $L \geq 4$ osiąga maksymalny okres równy 2^{L-2} wtedy i tylko wtedy, gdy X_0 jest liczbą nieparzystą oraz $a = 3 \mod 8$ (czyli a jest podzielne przez 8 z resztą 3) lub $a = 5 \mod 8$.

Lemat 2. *Generator mieszany osiąga pełny okres m wtedy i tylko wtedy, gdy są spełnione wszystkie trzy warunki:*

1. *liczby c i m nie mają wspólnych dzielników,*
2. *$a = 1 \mod p$ dla każdego czynnika pierwszego liczby m (czyli każdej liczby pierwszej dzielącej m),*
3. *$a = 1 \mod 4$, jeżeli 4 jest dzielnikiem liczby m .*

4 Generatory Fibonacciego

Tzw. ALFG (Additive Lagged Fibonacci Generator) ma postać

$$X_n = (X_{n-s} + X_{n-r}) \mod m \text{ dla } n \geq r, r > s \geq 1. \quad (12)$$

Dla $m = 2^L$ maksymalny okres ALFG wynosi $(2^r - 1)2^{L-1}$.

Generatory postaci (12) można uogólnić do wzoru

$$X_n = (X_{n-s} \diamond X_{n-r}) \mod m, \quad (13)$$

gdzie \diamond jest pewną operacją (np. dodawaniem, odejmowaniem, mnożeniem). W przypadku mnożenia mówimy o MLFG (Multiplicative Lagged Fibonacci Generator), a jego maksymalny okres wynosi $(2^r - 1)2^{L-3}$.

5 Generatory nieliniowe

Przykładem może być tutaj generator postaci

$$X_n = (a\check{X}_{n-1}^{-1} + b) \mod m, \quad (14)$$

gdzie m jest liczbą pierwszą, a operacja \check{X}^{-1} jest *odwracaniem modulo*. Operacja ta jest zdefiniowana następująco: dla $x = 0$ zachodzi $\check{x}^{-1} \mod m = 0$, a dla $x \neq 0$ liczba \check{x}^{-1} musi spełniać warunek $x \cdot \check{x}^{-1} \mod m = 1$.

Przykładem może też być generator oparty na obliczaniu kwadratów, opisany wzorem

$$X_n = X_{n-1}^2 \mod m. \quad (15)$$

6 Metoda odwracania dystrybuanty

W metodzie tej wykorzystujemy następujące twierdzenie.

Twierdzenie 3. *Niech U będzie zmienną losową z rozkładu jednostajnego na przedziale jednostkowym, a $F_X(\cdot)$ – ciągłą i ściśle rosnącą dystrybuantą pewnego rozkładu prawdopodobieństwa. Wtedy zmienna losowa X określona warunkiem*

$$X = F_X^{-1}(U) \quad (16)$$

ma rozkład dany dystrybuantą $F_X(\cdot)$.

Powyższe twierdzenie prowadzi bezpośrednio do następującego algorytmu

Algorytm 4.

```
U = GenerujU
X = F-1(U)
return X
```

gdzie $F^{-1}(\cdot)$ jest wartością funkcji $F_X^{-1}(\cdot)$.

Twierdzenie 3 można rozszerzyć na przypadek dowolnych, niekoniecznie ciągłych i ściśle rosnących dystrybuant $F_X(\cdot)$. Wystarczy jednak posłużyć się *uogólnioną definicją funkcji odwrotnej do dystrybuanty*. Niech

$$F_X^-(t) = \inf\{x : t \leq F(x)\}. \quad (17)$$

Zgodnie z tą definicją, $F_X^{-1}(t)$ jest to najmniejsza wartość funkcji dystrybuanty $F_X(x)$, która przekracza lub jest równa t .

Twierdzenie 5. *Niech U będzie zmienną losową z rozkładu jednostajnego na przedziale jednostkowym, a $F_X(\cdot)$ dystrybuantą pewnego rozkładu prawdopodobieństwa. Wtedy zmienna losowa X zdefiniowana warunkiem*

$$X = F_X^-(U) \quad (18)$$

ma rozkład określony dystrybuantą $F_X(\cdot)$.

7 Metoda eliminacji

Założmy, że interesuje nas wygenerowanie zmiennej losowej X o gęstości zadanej funkcją $f_X(t)$, która to funkcja jest większa od zera tylko na przedziale $[0; 1]$ i równa zero poza tym przedziałem. Przypuśćmy ponadto, że funkcja $f_X(t)$ przyjmuje na przedziale jednostkowym wartości ograniczone przez pewną stałą M . W takim przypadku następujący algorytm

Algorytm 6.

```
repeat
  {U1 = GenerujU;
  U2 = GenerujU }
until M * U2 <= f (U1)
X = U1
return X
```

generuje zmienną o rozkładzie zadany gęstością $f_X(\cdot)$.

Ogólny wariant tej metody zakłada, że umiemy generować zmienną Y o rozkładzie określonym funkcją gęstości $g_Y(t)$, zwaną gęstością dominującą. Założmy ponadto, że interesuje nas zmienna X opisana gęstością $f_X(t)$, przy czym na całym przedziale określoności tej gęstości zachodzi $f_X(t) \leq M g_Y(t)$ dla pewnej stałej M . Wtedy następujący algorytm

Algorytm 7.

```
repeat
  {U = GenerujU
  Y = GenerujG }
until M * U * g (Y) <= f (Y)
X = Y
return X
```

generuje zmienną X z rozkładu o gęstości $f_X(\cdot)$. Zauważmy, że w algorytmie tym wykorzystujemy funkcję **GenerujG**, która służy do generowania zmiennej z rozkładu o gęstości $g_Y(\cdot)$.

8 Metoda szybkiej eliminacji i szeregów

Skuteczniejsze okazać się może znalezienie prostszych funkcji ograniczających, postaci

$$\alpha_1(x) \leq \frac{f_X(x)}{M g_Y(x)} \leq \beta_1(x) \quad (19)$$

dla dowolnego x .

Algorytm ma wtedy postać

Algorytm 8.

```
flaga = 0
repeat
  {U = GenerujU
```

```

Y = GenerujG
if U <= alpha (Y) then
  flaga = 1
else
  if U <= beta (Y) then
    if M * U * g (Y) <= f (Y) then
      flaga = 1}
until flaga = 1
X = Y
return X

```

Metodę tą można uogólnić na ciąg funkcji przybliżających warunek akceptacji. Jest to tzw. metoda szeregów. Przypuśćmy, że dla dowolnych x i $n \in \mathbb{N}$ zachodzi

$$\underline{f}_n(x) \leq f_X(x) \leq \overline{f}_n(x) , \quad (20)$$

zatem gęstość docelowa $f_X(\cdot)$ jest przybliżana przez odpowiedni ciąg funkcji. Przykładowy algorytm ma wtedy postać

Algorytm 9.

```

repeat
  {U = GenerujU
  Y = GenerujG
  n = 0
  repeat
    {n = n + 1
    if M * U * g (Y) <= f_n (Y) then return Y}
  until M * U * g (Y) > f~n (Y)
until false

```

Bazując na (19) metodę tą możemy zapisać również jako ciąg warunków szybkiej akceptacji i szybkiego odrzucenia

$$\alpha_k(x) \leq \dots \leq \alpha_1(x) \leq \frac{f_X(x)}{Mg_Y(x)} \leq \beta_1(x) \leq \dots \beta_l(x) \quad (21)$$

dla dowolnego x .

Metoda szeregów w postaci (20) może zostać uogólniona na przykład szeregów zbieżnych. Przypuśćmy, że funkcja gęstości $f_X(\cdot)$ wyraża się jako granica zbieżnego szeregu nieskończonego

$$f_X(x) = \sum_{i=1}^{\infty} S_i(x) , \quad (22)$$

przy czym potrafimy zawsze oszacować bezwzględną wartość reszty szeregu

$$\left| \sum_{i=n}^{\infty} S_i(x) \right| \leq R_n(x) \quad (23)$$

dla dowolnego x . Odpowiedni algorytm ma wtedy postać

Algorytm 10.

```

repeat
  {U = GenerujU
  Y = GenerujG
  S = 0
  n = 0
  repeat
    {n = n + 1
    S = S + S_n (Y)}
  until | S - M * U * g (Y) | > R_{n+1} (Y)}
until M * U * g (Y) <= S (Y)
X = Y
return X

```

9 Metoda ilorazu równomiernego

Metoda *ilorazu równomiernego* (*RI*, *ROU*, czyli *ratio-of-uniforms*) bazuje na następującym twierdzeniu.

Twierdzenie 11. *Niech $f_X(\cdot)$ będzie nieujemną i skończenie całkowną funkcją i niech*

$$\mathcal{C}_f = \left\{ (u, v) : 0 \leq u \leq \sqrt{f_X\left(\frac{v}{u}\right)} \right\} . \quad (24)$$

Jeśli punkt (U, V) ma rozkład równomierny na zbiorze \mathcal{C}_f , to zmienna losowa $X = \frac{V}{U}$ ma rozkład o gęstości $\frac{f_X(\cdot)}{\int f_X(t) dt}$.

Algorytm jest zatem następujący

```

repeat
  {(U, V) = GenerujCf
  X = V / U}
until U^2 <= f(X)
return X

```

gdzie funkcja **GenerujCf** służy do generowania punktu z jednostajnego rozkładu na zbiorze \mathcal{C}_f .

Stosując podstawienie

$$z = \frac{v}{u} \quad (25)$$

możemy w parametryczny sposób przedstawić ograniczenie zbioru \mathcal{C}_f . Z (25) i ograniczenia zbioru w twierdzeniu 11 mamy bowiem

$$u = \sqrt{f(z)} , v = z\sqrt{f(z)} , \quad (26)$$

stąd

$$0 \leq u \leq \sup_z \sqrt{f(z)} , \inf_z z\sqrt{f(z)} \leq v \leq \sup_z z\sqrt{f(z)} , \quad (27)$$

co daje warunki na „zawarcie” zbioru \mathcal{C}_f w pewnym prostokącie.

10 Metoda superpozycji rozkładów

Metoda superpozycji (kompozycji) rozkładu polega na przedstawieniu rozważanej gęstości $f_X(\cdot)$ w postaci

$$f_X(t) = \sum_{i=1}^{\infty} p_i f_i(t) , \quad (28)$$

gdzie $p_i > 0$, $\sum_{i=1}^{\infty} p_i = 1$, a $f_i(\cdot)$ są gęstościami pewnych rozkładów prawdopodobieństwa.

Prowadzi to do następującego algorytmu

Algorytm 12.

```
K = GenerujK
X = GenerujF (K)
return X
```

przy czym funkcja **GenerujK** generuje zmienną z rozkładu dyskretnego zadanego prawdopodobieństwami p_1, p_2, \dots , a funkcja **GenerujF** (K) to rodzina funkcji generujących zmienną z rozkładu określonego gęstością $f_K(\cdot)$.

W praktyce zazwyczaj wzór (28) upraszcza się do postaci ze skończonym rozkładem dyskretnym

$$f_X(t) = \sum_{i=1}^n p_i f_i(t) . \quad (29)$$

W takim przypadku cały przedział określoności $f_X(\cdot)$ dzieli się na rozłączne podzbiory $\mathcal{A}_1, \mathcal{A}_2, \dots, \mathcal{A}_n$, tak, aby generowanie z poszczególnych gęstości $f_i(\cdot)$ było jak najłatwiejsze.

Zauważmy, że dla (29) mamy

$$p_i = \int_{\mathcal{A}_i} f_X(t) dt , \quad f_i(t) = \frac{f_X(t)}{p_i} \mathbb{1}_{\mathcal{A}_i}(t) . \quad (30)$$

Wzór (28) możemy w ogólniejszy sposób zapisać jako całkę

$$f_X(t) = \int_{\mathcal{Z}} f_z(t) h(z) dz , \quad (31)$$

gdzie $f_z(\cdot)$ dla każdej wartości parametru z jest pewną gęstością, a $h(z)$ jest gęstością pewnego rozkładu określonego na zbiorze \mathcal{Z} (np. prostej rzeczywistej).

Odpowiedni algorytm dla (31) ma postać

Algorytm 13.

```
Z = GenerujH
X = GenerujF (Z)
return X
```

najpierw następuje więc wygenerowanie parametru z z gęstości $h(\cdot)$, a w drugim kroku do generowania wykorzystywana jest odpowiednia funkcja gęstości $f_z(\cdot)$.

11 Metody generowania z rozkładów dyskretnych

Najprostszym algorytmem jest metoda bazująca bezpośrednio na metodzie odwracania dystrybucyj

Algorytm 14.

```
S = 0
U = GenerujU
I = 0
while S <= U do
  {I = I + 1
   S = S + p_I}
X = I
return X
```

Bardziej zaawansowanym algorytmem jest algorytm ALIAS, który może być stosowany do dyskretnych rozkładów o skończonej liczbie wartości. Załóżmy, że oprócz prawdopodobieństw

$$P(X = 1) = p_1, P(X = 2) = p_2, \dots, P(X = m) = p_m, \quad (32)$$

dysponujemy ciągiem q_1, q_2, \dots, q_m , takim, że $0 \leq q_i \leq 1$ dla $i = 1, 2, \dots, m$ oraz ciągiem $A(1), A(2), \dots, A(m)$ o wartościach w zbiorze $\{1, 2, \dots, m\}$. Oba te ciągi spełniają przy tym warunek

$$p_i = \left(q_i + \sum_{j:A(j)=i} (1 - q_j) \right) \frac{1}{m} \quad (33)$$

dla $i = 1, 2, \dots, m$. Algorytm ma wtedy następującą postać

Algorytm 15.

```
I = GenerujU[m]
U = GenerujU
if U < q_I then X = I
  else X = A (I)
return X
```

gdzie funkcja **GenerujU[m]** generuje wartość z rozkładu jednostajnego na zbiorze $\{1, 2, \dots, m\}$.

12 Metody szczegółowe

Jeśli zmienna X pochodzi z rozkładu normalnego standardowego, to zmienna $Y = \sigma X + \mu$ ma już dowolny rozkład $N(\mu; \sigma^2)$. Możemy się zatem skupić tylko na zagadnieniu uzyskiwaniu zmiennej z rozkładu $N(0; 1)$.

Algorytm prymitywny Jednym z najprostszych generatorów rozkładu normalnego ma postać

Algorytm 16.

```

X = 0
for I = 1 to 12 do
  X = X + GenerujU
X = X - 6
return X

```

Algorytm Boxa-Mullera Metoda ta pozwala na wygenerowanie *dwóch* zmiennych niezależnych X_1, X_2 z rozkładu $N(0; 1)$.

Algorytm 17.

```

U1 = GenerujU
U2 = GenerujU
Phi = 2 * Pi * U1
R = Sqrt ( - 2 * Ln ( U2 ))
X1 = R * Cos ( Phi )
X2 = R * Sin ( Phi )
return X1, X2

```

Algorytm Marsaglii Ten algorytm jest w wielu punktach podobny do poprzedniego.

Algorytm 18.

```

repeat
  {U1 = GenerujU;
  U2 = GenerujU;
  U1 = 2 * U1 - 1;
  U2 = 2 * U2 - 1;
  W = U1^2 + U2^2}
until W < 1
C = Sqrt ( - 2 * W^(-1) * Ln ( W ) )
X1 = C * U1
X2 = C * U2
return X1, X2

```

13 Wielowymiarowe zmienne losowe

Odpowiedni wektor losowy o p składowych $(X^{(1)}, X^{(2)}, \dots, X^{(p)})$ oznaczać będziemy przez \mathbb{X} . W ten sposób

$$\mathbb{X}_i = (X_i^{(1)}, X_i^{(2)}, \dots, X_i^{(p)}) \quad (34)$$

oznaczać będzie i -tą zmienną w ciągu o p składowych.

Dla przypadku $p = 2$ istnieje algorytm bazujący na współrzędnych biegunowych

Algorytm 19.

```

Phi = 2 * Pi * GenerujU
X1 = cos (Phi)
X2 = sin (Phi)
u = GenerujU
Y1 = Sqrt (U) * X1
Y2 = Sqrt (U) * X2
return ( X1,X2 )

```

W algorytmie tym najpierw generujemy punkt położony równomiernie na okręgu o współrzędnych (X_1, X_2) . Następnie skalujemy go, otrzymując punkt położony równomiernie w kole o współrzędnych (Y_1, Y_2) .

Częściej stosuje się algorytm bazujący na odpowiednim unormowaniu zmiennych. Jeśli rozpatrujemy rozkład jednostajny na kuli, to odpowiedni algorytm ma postać

Algorytm 20.

```

for i = 1 to p do
  Z[i] = GenerujNormalnyStd
R^2 = Z[1]^2 + ... + Z[p]^2
for i = 1 to p do
  Y[i] = Z[i] / R
U = GenerujU
R1 = U^(1/p)
for i = 1 to p do
  X[i] = Y[i] * R1
return ( X[1], ..., X[p] )

```

Algorytm ten generuje p zmiennych losowych Z_i , każda z niezależnego standardowego rozkładu normalnego. Są one następnie normowane względem odległości R w celu otrzymania zmiennej (Y_1, \dots, Y_p) , która pochodzi z rozkładu równomiernego na sferze jednostkowej. Następnie losowana jest dodatkowa zmienna z rozkładu jednostajnego na przedziale jednostkowym. Po jej przekształceniu, wcześniej uzyskany punkt ze sfery jest przesuwany do nowej pozycji (X_1, \dots, X_p) tak, aby uzyskać rozkład jednostajny na kuli.

13.1 Wielowymiarowy rozkład normalny

Jeśli chcemy wygenerować zmienne losowe

$$\mathbb{X}_1, \mathbb{X}_2, \dots \stackrel{iid}{\sim} N(\mu, \sigma^2 \cdot \mathbb{I}) , \quad (35)$$

czyli o niezależnych poszczególnych składowych (tzn. $\text{Cov}(X^{(i)}, X^{(j)}) = 0$ dla $i \neq j$), to jest to możliwe przy wykorzystaniu wcześniej poznanych metod dla jednowymiarowych zmiennych losowych. Jeśli jednak macierz kowariancji

$$\text{VAR } \mathbb{X} = \mathbb{W} = \begin{pmatrix} \text{Var } X^{(1)} & \text{Cov}(X^{(1)}, X^{(2)}) & \dots & \text{Cov}(X^{(1)}, X^{(p)}) \\ \text{Cov}(X^{(2)}, X^{(1)}) & \text{Var } X^{(2)} & \dots & \text{Cov}(X^{(2)}, X^{(p)}) \\ \vdots & \vdots & \ddots & \vdots \\ \text{Cov}(X^{(p)}, X^{(1)}) & \text{Cov}(X^{(p)}, X^{(2)}) & \dots & \text{Var } X^{(k)} \end{pmatrix} , \quad (36)$$

ma bardziej skomplikowaną postać, niezbędne jest zastosowanie innych metod.

Dla ułatwienia załóżmy, że $\mathbb{E} \mathbb{X} = 0$ i macierz kowariancji \mathbb{W} daje się przedstawić jako iloczyn

$$\mathbb{W} = \mathbb{C}\mathbb{C}^T \quad (37)$$

dla pewnej nieosobliwej macierzy \mathbb{C} . W takim przypadku, jeśli \mathbb{Y} ma wielowymiarowy rozkład normalny $N(0, 1 \cdot \mathbb{I})$, to po użyciu przekształcenia

$$\mathbb{X} = \mathbb{C}\mathbb{Y} \quad (38)$$

zachodzi

$$\mathbb{X} \sim N(0, \mathbb{W}) . \quad (39)$$

Wymagany krok jest skonstruowanie odpowiedniej nieosobliwej macierzy \mathbb{C} . Możemy w tym celu wykonać dekompozycję na macierz dolnotrójkątną (czyli z samymi zerami powyżej głównej przekątnej), wykorzystując metodę zwaną dekompozycją Choleskiego. Jest ona zdefiniowana wzorami

$$c_{i,i} = \sqrt{\left(w_{i,i} - \sum_{k=1}^{i-1} c_{i,k}^2 \right)} \quad (40)$$

$$c_{j,i} = \frac{w_{j,i} - \sum_{k=1}^{i-1} c_{j,k} c_{i,k}}{c_{i,i}} , \quad (41)$$

gdzie $w_{i,j}$ i $c_{i,j}$ są odpowiednimi wyrazami w komórkach macierzy \mathbb{W} i \mathbb{C} .