Generatory liczb losowych (2)

Mgr inż. Szymon Łukasik

szymonl@pk.edu.pl

Metody generowania liczb z rozkładów innych niż jednostajny

W zagadnieniach współczesnej nauki i inżynierii często pojawia się problem generowania liczb losowych z różnorakich rozkładów mających postać diametralnie różną od znanego już Państwu rozkładu jednostajnego.

Niech $U(0,1)$ oznacza zmienną losową o rozkładzie jednostajnym w przedziale (0,1), a $x \sim U(0,1)$ element próby otrzymanej z tego rozkładu. Celem niniejszego ćwiczenia jest zapoznanie Państwa z głównymi metodami generowania elementów próby *Y* charakteryzowanej przez odmienny rozkład zadany określoną funkcją gęstości $f(y)$ – oczywiście na podstawie próby pozyskanej z rozkładu $U(0,1)$. Podstawowe stosowane w tym zakresie techniki to:

– metoda odwracania dystrybuanty

oraz

– metoda eliminacji.

Metoda odwracania dystrybuanty

Niech $F(y)$ oznacza dystrybuantę charakteryzującą pewien rozkład zmiennej losowej *Y*. Ciąg elementów próby $y_1, y_2, \ldots, y_n$ zdefiniowany przez:

$$y_i = F^{-1}(x_i)$$

posiada rozkład $f(y)$ jeśli elementy $x_1, x_2, \ldots, x_n$ pozyskiwane są z rozkładu $U(0,1)$.

Przykład:

Dystrybuanta rozkładu wykładniczego (ekspotencjalnego) ma postać:

$$F(y) = 1 - e^{-\lambda y}$$

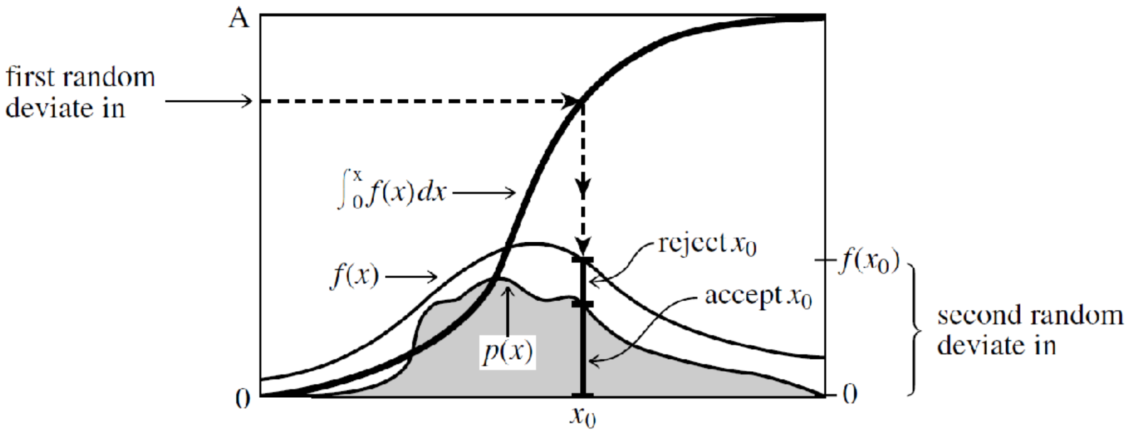
zatem zgodnie z powyższym aby otrzymać ciąg liczb losowych z tego rozkładu należy dokonać operacji:

$$y_i = -\frac{1}{\lambda} \ln(1 - x_i)$$

Problemem związanym z tą metodą jest kwestia uzyskiwania funkcji odwrotnej do dystrybuanty (nie zawsze jest to niestety możliwe – patrz: rozkład normalny).

Metoda eliminacji

Metoda ta stosowana jest dla generacji liczb z rozkładów o znanej gęstości *f(y)* w przedziale $y \in (a, b)$ w sytuacji w której dystrybuanta nie jest znana lub nie da się wyznaczyć funkcji do niej odwrotnej (wykazuje się niestety mniejszą efektywnością). Jeżeli znana jest metoda generacji liczb z rozkładu *g(y)* i $g(y) \geq f(y)\ \forall y \in (a, b)$, to po wygenerowaniu liczby *y* z tego rozkładu akceptuje się ją z prawdopodobieństwem *f(y)/g(y)*. Można udowodnić że tak pozyskana próba *Y* ma rozkład *g(y)*.



Ilustracja metody eliminacji (źródło: Numerical Recipes)

Jako rozkład *g(y)* można zastosować na przykład odpowiedni rozkład jednostajny *g(y)=c*. Procedura generacji liczb losowych ma wtedy postać:

1. Wygeneruj dwie liczby losowe y_1 i y_2 z rozkładów jednostajnych odpowiednio *U(a,b)* i *U(0,c).*
2. Sprawdź czy $y_2 \leq f(y_1)$. Jeśli tak przyjmij y_1 jako kolejny element generowanej próby losowej. Jeśli nie – powtórz krok 1.

Metodę tą można z powodzeniem zastosować na przykład do generacji liczb z rozkładu dwumianowego. Warto jednak nadmienić, że próbę z tego rozkładu można wygenerować również stosując inną bardziej bezpośrednią technikę. Skoro rozkład ten reprezentuje ilość sukcesów w *n* eksperymentach z prawdopodobieństwem pojedynczego sukcesu *p*, wystarczy jedynie wygenerować ciąg *n* liczb losowych $x \sim U(0,1)$ i wyznaczyć liczbę elementów *x* spełniających warunek $x \leq p$. Liczba ta jest pojedynczym elementem nowej zmiennej losowej *Y* o rozkładzie dwumianowym.

Inne metody transformacyjne

Aby uzyskać liczby o rozkładzie normalnym stosować można transformację Boxa-Mullera. Gdy U_1 i U_2 są dwiema zmiennymi losowymi o rozkładach jednostajnych $\sim U(0,1)$ to w wyniku transformacji:

$$Z_1 = \sqrt{-2 \ln U_1} \cos(2\pi U_2)$$

$$Z_2 = \sqrt{-2 \ln U_1} \sin(2\pi U_2)$$

otrzymuje się dwie niezależne zmienne losowe Z_1 i Z_2 o rozkładzie normalnym *N(0,1)* – czyli zerową wartością oczekiwaną i jednostkową wariancją. Aby rozkład wygenerowanych liczb miał postać $N(\mu,\sigma^2)$ należy pomnożyć otrzymaną liczbę Z_i przez σ i dodać μ.