Losowanie z rozkładu normalnego (czyli Gaussa) N(μ,σ)

Najczęściej występujący w naturze. Zmienna w przedziale od Parametry rozkład to:

Wartość średnia <x>=μ

Odchylenie standardowe σ   jest związane z szerokością rozkładu (wykres górny z Wikipedii)

Reguła 1,2,3σ na wykresie dolnym ze strony www.pronost.pl. Prawdopodobieństwo znalezienia x z przedziału <μ-σ, μ+σ > wynosi 68,28%, itd

Funkcja gęstości prawdopodobieństwa opisana wzorem

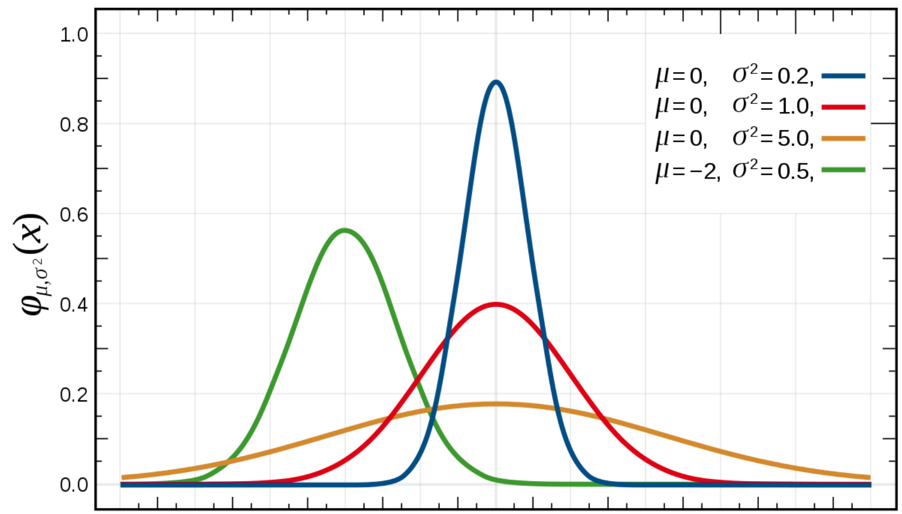
$$f_{\mu,\sigma}(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(\frac{-(x-\mu)^2}{2\sigma^2}\right).$$

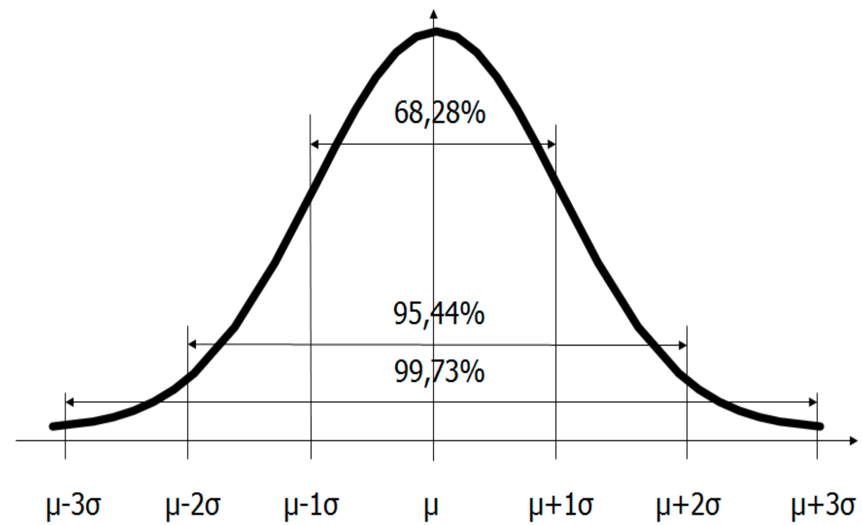
To jest już znormalizowana funkcja

Przez podstawienie z=(x-μ)/σ możemy uzyskać rozkład

N(0,1) – standardowy rozkład normalny (parametry jak dla czerwonego).

Ten rozkład jest stablicowany (zarówno wartości funkcji jak i dystrybuanta)





Losowanie z rozkładu normalnego N(μ,σ)

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

Rozkład normalny o dowolnych parametrach (średniej i wariancji) można sprowadzić do N(0,1) poprzez podstawienie

$$z = \frac{x-\mu}{\sigma}$$

$$f(z)\,\mathrm{d}z = \frac{1}{2\pi}\mathrm{e}^{-\frac{z^2}{2}}\,\mathrm{d}z \qquad z \in (-\infty, +\infty)$$

Jeśli będziemy poprawnie losować zmienną ***z*** to przez podstawienie odwrotne uzyskamy zmienna *x* , która podlega N(μ,σ). Wystarczy przeliczyć

$$x = z\sigma + \mu$$

Losowanie z rozkładu N(0,1)

1. Nie można analitycznie policzyć dystrybuanty (całka z funkcji typu $e^{(-x2)}$ nie została jeszcze policzona). Zwykła metoda odwróconej dystrybuanty nie działa
2. Zakres zmiennej $(-\infty, \infty)$ ogranicza również metodę eliminacji (można ograniczyć się do przedziału co najmniej $<\mu-3\sigma, \mu+3\sigma >$, ale to jest zbyt ważny rozkład na stosowanie ograniczeń)

I. Losowanie z rozkładu N(0,1) – metoda Boxa-Mullera

Zakładamy, że mamy dwa niezależne rozkłady normalne zmiennych losowych N(0,1) z których jedne opisuje współrzedną x a drugi współrzedną y punktu na płaszczyźnie. Skoro zmienne niezależne to ps-stwa się mnoży i tak prawdopodobieństwo że wspólrzędne punktu są z przedziałów (x,x+dx) i (y,y+dy) wynosi:

$$f(x)dxf(y)dy = \frac{1}{2\,\pi} \quad \exp(-(x^2+y^2)\,/2) \;\; dx\,dy$$

Przed liczeniem dystrybuanty przechodzimy do układu biegunowego współrzędnych i dzięki pojawieniu się jakobianu całkę można łatwo analitycznie wyznaczyć.

I. Losowanie z rozkładu N(0,1) – metoda Boxa-Millera

Losujemy dwie liczby pseudolosowe z przedziału (0,1> - rozkład równomierny równomiernego np: u1, u2

Obliczany zmienną

R= sqrt(-2 ln(u1)) ... i dlatego przedział (0,1>

Theta= 2 π u2

Zmienne

z_1=R sin(theta) oraz z_2= R cos(theta) podlegaja rozkładowi N(0,1)

Dowolny rozkład N(μ,σ) dostaniemy jak zawsze poprzez wstawienie obliczonej liczby z (nie ważne czy 1 czy 2) do wzoru

$$\boxed{x = z\sigma + \mu}$$

II. W oparciu o centralne twierdzenie graniczne

Centralne twierdzenie graniczne to twierdzenie matematyczne mówiące, że jeśli x_i są niezależnymi zmiennymi losowymi o jednakowym rozkładzie, takiej samej wartości oczekiwanej μ oraz dodatniej i skończonej wariancji σ^2, to zmienna losowa y (uzyskana z odpowiedniej **sumy m zmiennych x_i**)

zbiega według rozkładu do standardowego rozkładu normalnego N(0,1), gdy m rośnie do nieskończoności.

$$y = \frac{\frac{1}{m}\Sigma x_i - \mu}{\frac{\sigma}{\sqrt{m}}}$$

II. W oparciu o centralne twierdzenie graniczne

$$y = \frac{\frac{1}{m}\ \Sigma x_i\ - \mu}{\frac{\sigma}{\sqrt{m}}}$$

Dobrym przybliżeniem jest juz losowanie 6-ciu liczb z U(0,1).

Dla tego rozkładu:

μ=0,5

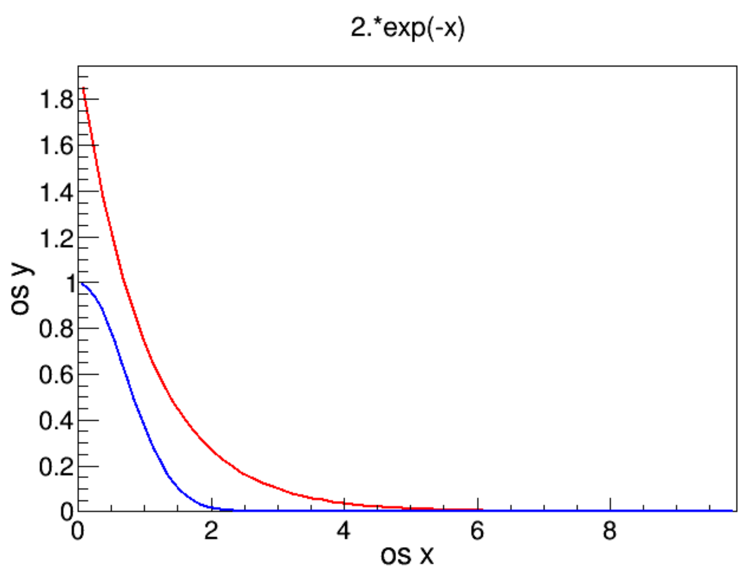
σ^2=1/3 -1/4= 1/(12)

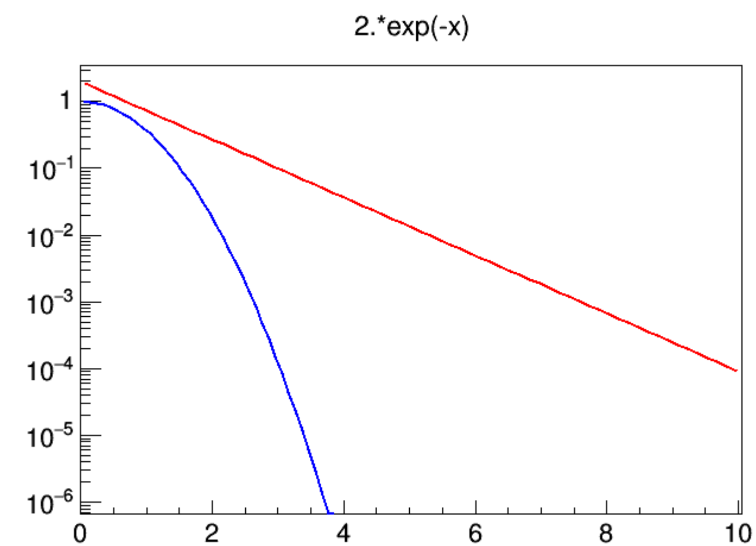
III. Poprzez połaczone metody odw. Dystrybuanty i eliminacji

Algorytm 3.1 w Wieczorkowski i inni: „ Komputerowe generatory liczb losowych”

- wytłumaczenie metody na wykresie, szukam funkcji ograniczającej z góry moją funkcje gęstości dla N(0,1)

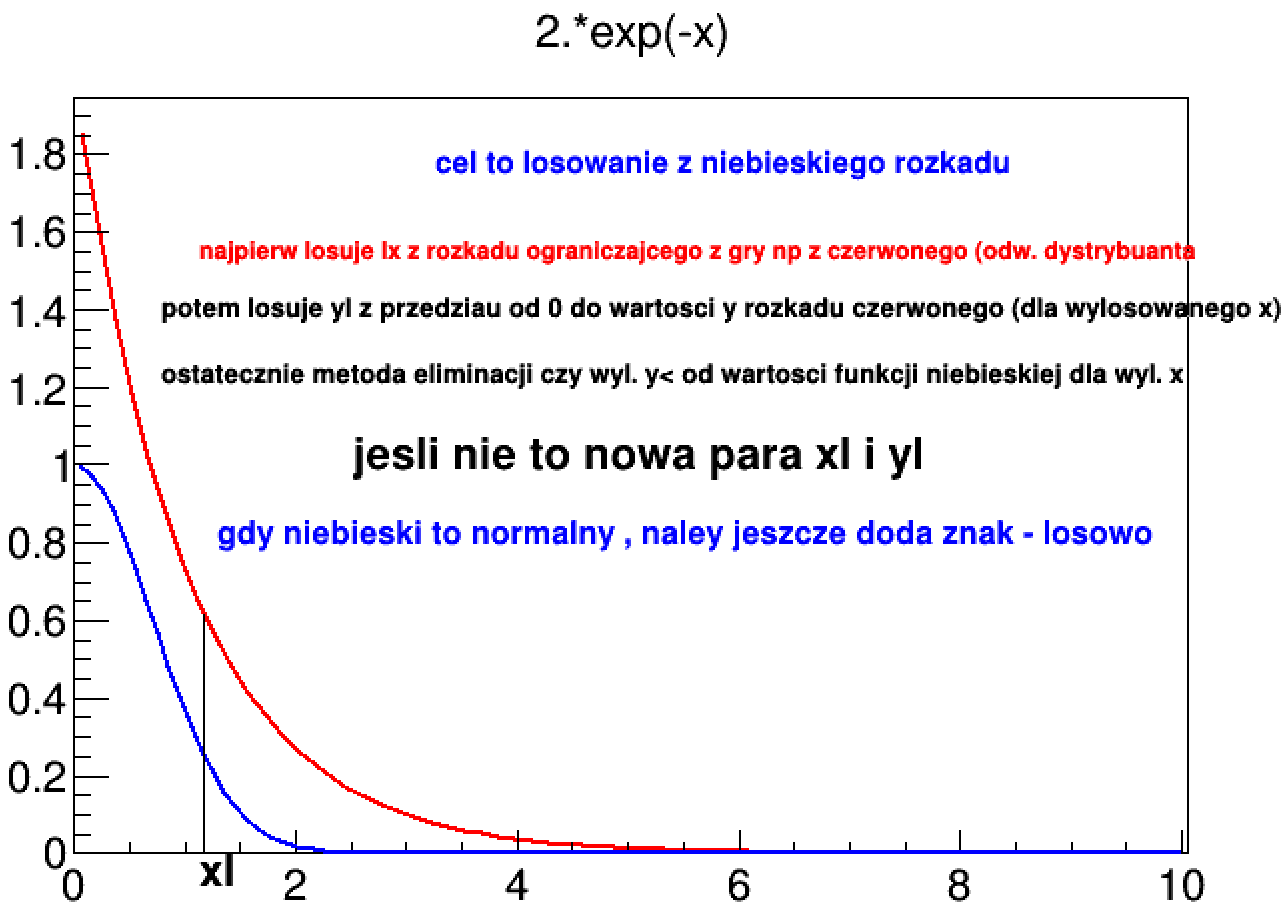
np.   2*exp(-x)





III. Poprzez połączone metody odw. Dystrybuanty i eliminacji

Algorytm 3.1 w Wieczorkowski i inni: „ Komputerowe generatory liczb losowych”



IV. Z biblioteki random

```
#include <random>
...
std::mt19937 GEN(time(nullptr));
...
    std::normal_distribution<> dodatkowa_drzemka(80., 20);
...
```