Assignment

Karthikeya hanu prakash kanithi (EE22BTECH11026)

Question : Let $\phi(.)$ denote the cumulative distribution function of a standard normal random variable. If the random variable X has the cumulative distribution function

$$F(x) = \begin{cases} \phi(x), & x < -1 \\ \phi(x+1), & x \geq -1 \end{cases} \tag{1}$$

then which one of the following statements is true?

A) $P(X \leq -1) = \frac{1}{2}$
B) $P(X = -1) = \frac{1}{2}$
C) $P(X < -1) = \frac{1}{2}$
D) $P(X \leq 0) = \frac{1}{2}$

Solution: Gaussian

Q function is defined

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{\frac{-u^2}{2}} du \tag{2}$$

From (1) and (2);

$$F_X(x) = \begin{cases} Q(-x), & x < -1 \\ 1 - Q(x+1), & x \geq -1 \end{cases} \tag{3}$$

From (6);

A)

$$\Pr(X \leq -1) = F_X(-1) = 1 - Q(0) \tag{4}$$
$$= 0.5 \tag{5}$$

So Option A i.e., $P(X < -1) = \frac{1}{2}$ is correct

B) The pdf of X can be defined in terms of cdf as

$$\Pr(X = b) = F_X(b) - \lim_{x \to b^-} F_X(x) \tag{6}$$

From (6);

$$\Pr(X = -1) = F_X(-1) - \lim_{x \to -1^-} F_X(x) \tag{7}$$
$$= 1 - Q(0) - Q(-(-1)) \tag{8}$$
$$= 0.341 \tag{9}$$

So Option B i.e., $P(X = -1) = \frac{1}{2}$ is incorrect

C)

$$\Pr(X < -1) = \lim_{x \to -1^-} F_X(x) = F_X(-1) \tag{10}$$
$$= Q(-(-1)) \tag{11}$$
$$= 0.159 \tag{12}$$

So Option C i.e., $P(X < -1) = \frac{1}{2}$ is incorrect

D)

$$\Pr(X \leq 0) = F_X(0) = 1 - Q(1) \tag{13}$$
$$= 0.8413 \tag{14}$$

So Option D i.e., $P(X \leq 0) = \frac{1}{2}$ is incorrect

Guassian CDF plot of X is given in fig1

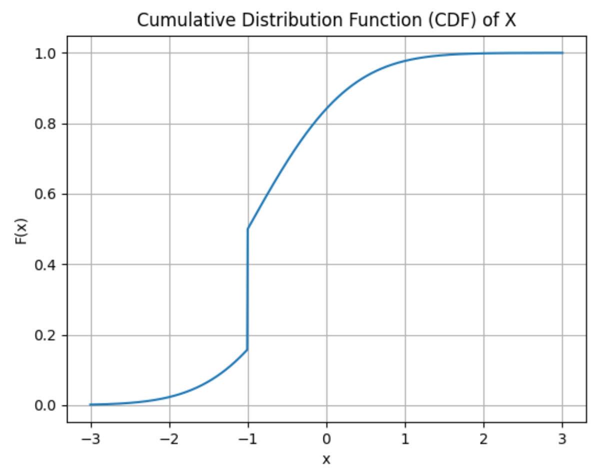

Fig. 1.

Box-Muller

STATEMENT: Suppose U1 and U2 are independent samples chosen from the uniform distribution on the unit interval (0, 1). Let

$$Z_0 = R\cos(\Theta) = \sqrt{-2\ln U_1}\cos(2\pi U_2) \tag{15}$$

and

$$Z_1 = R\sin(\Theta) = \sqrt{-2\ln U_1}\cos(2\pi U_2) \tag{16}$$

Then Z_0 and Z_1 are independent random variables with a standard normal distribution. So, now we will generate Z_0 using C Code as given below

PROOF: Let X and Y be independent standard normal variables

$$X, Y \sim \mathcal{N}(0,1) \quad and \quad X \perp Y \tag{17}$$

The joint pdf of X and Y is given by

$$\begin{aligned} f_{XY}(X,Y) &= f(x)f(y) && (18)\\ &= \frac{1}{\sqrt{2\pi}}e^{\frac{-x^2}{2}}.\frac{1}{\sqrt{2\pi}}e^{\frac{-y^2}{2}} && (19)\\ &= \frac{1}{2\pi}e^{\frac{-(x^2+y^2)}{2}} && (20) \end{aligned}$$

The relationship between Cartesian coordinates (x, y) and polar coordinates (r, θ) is as follows

$$\begin{aligned} x &= r\cos\theta && (21)\\ y &= r\sin\theta && (22) \end{aligned}$$

Change $f_{XY}(x, y)$ to polar coordinates :

$$f_{XY}(x,y)dxdy = f_{R\theta}(r,\theta)drd\theta \tag{23}$$

i.e.,

$$f_{R\theta}(r,\theta) = f_{XY}(x,y)\frac{dxdy}{drd\theta} = f_{XY}(x,y)\left|\frac{\partial(x,y)}{\partial(r,\theta)}\right| \tag{24}$$

where J is the Jacobian

$$\begin{aligned} J &= \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{vmatrix} && (25)\\ &= \begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix} = r && (26) \end{aligned}$$

For $r \geq 0$ and $\theta \in [0, 2\pi)$, we have

$$f_{R\theta}(r,\theta)drd\theta = \frac{1}{2\pi}e^{\frac{-(r^2)}{2}}rdrd\theta \tag{27}$$

Now we change the variable from (r, θ) to (r^2, θ);

$$rdr = \frac{1}{2}dr^2 \tag{28}$$

Now it can be written as,

$$\begin{aligned} f_{R\theta}(r,\theta)drd\theta &= f_{R^2\theta}(r^2,\theta)dr^2d\theta && (29)\\ &= \frac{1}{2\pi}e^{\frac{-(r^2)}{2}}\frac{1}{2}dr^2d\theta && (30)\\ &= \left(\frac{1}{2}e^{\frac{-(r^2)}{2}}dr^2\right)\left(\frac{1}{2}d\theta\right) && (31)\\ &= f_{R^2}(r^2)dr^2 f_\theta(\theta)d\theta && (32) \end{aligned}$$

from the above equation, we can say that

$$R^2 \perp \theta \text{ (}i.e., R^2 \text{ and } \theta \text{ are independent)} \tag{33}$$

Generate, $\theta \sim Unif(0, 2\pi)$ Generate, $V \sim Exp(\lambda = \frac{1}{2})$ (i.e., $V = R^2$) and compute

$$R = \sqrt{V} \tag{34}$$

Compute

$$\begin{aligned} X &= R\cos\theta && (35)\\ Y &= R\sin\theta && (36)\\ & && (37) \end{aligned}$$

where, X, Y are i.i.d in $\sim \mathcal{N}(0,1)$ Then

$$\begin{aligned} \theta &= 2\pi U_1 && (38)\\ V &= -2log(U_2) && (39) \end{aligned}$$

We can prove $V = -2log(U_2)$ using the c.d.f definition of the exponential distribution

Let X be a random variable following an exponential distribution with rate parameter $\lambda = \frac{1}{2}$, denoted as $X \sim \text{Exp}\left(\frac{1}{2}\right)$. The cumulative distribution function (CDF) of the exponential distribution is given by:

$$F(x) = 1 - e^{-\frac{x}{2}} \tag{40}$$

Now, suppose we have a random variable U following a uniform distribution in the interval $[0, 1]$, denoted as $U \sim \text{U}(0, 1)$. The CDF of the uniform distribution is simply:

$$F_U(u) = u, \text{ for } 0 \leq u \leq 1 \tag{41}$$

We can use the probability integral transform to express the exponential random variable X in terms of the uniform random variable U:

$$\begin{aligned} F(x) &= F_U(u) && (42)\\ 1 - e^{-\frac{x}{2}} &= u && (43) \end{aligned}$$

Now, solve for x:

$$\begin{aligned} e^{-\frac{x}{2}} &= 1 - u && (44)\\ -\frac{x}{2} &= \ln(1-u) && (45)\\ x &= -2\ln(1-u) && (46) \end{aligned}$$

So, $X \sim -2\ln(1 - U)$ for $U \sim \text{U}(0, 1)$. This expression represents the exponential random variable X in terms of a uniform random variable U.

C and Python codes are given below :

```
#include <stdio.h>
#include <stdlib.h>
#include <math.h>
#include <time.h>

double sn() {
    double u1, u2;
    while (1) {
        u1 = ((double)rand() / RAND_MAX);
        if (u1 > 0 && u1 < 1)
```

```c
            break;}
    while (1) {
        u2 = ((double)rand() / RAND_MAX);
        if (u2 > 0 && u2 < 1)
            break;}
    double z1 = sqrt(-2 * log(u1)) * cos(2 * M_PI * u2);
    return z1;
}

int main() {
    // Seed the random number generator
    srand(time(0));

    int numSamples = 100000; // You can change this to the desired number of samples
    FILE *file = fopen("uni.dat", "w");

    if (file == NULL) {
        printf("Unable to open file for writing.\n");
        return 1;
    }

    for (int i = 0; i < numSamples; i++) {
        double sample = sn();
        fprintf(file, "%lf\n", sample);
    }

    fclose(file);

    return 0;
}
```

```python
import numpy as np
import matplotlib.pyplot as plt

# Read data from the "uni.dat" file
x = np.genfromtxt("uni.dat")

# Create a histogram plot of the PDF
plt.hist(x, bins=50, density=True, alpha=0.6, color='b', label='PDF (Histogram)')
plt.xlabel('x (Standard Normal Random Variable)')
plt.ylabel('Frequency')
plt.title('PDF of Standard Normal Distribution (Histogram)')
plt.savefig('/home/sayyam/KHP/figs/figure1.png')
plt.show()
```

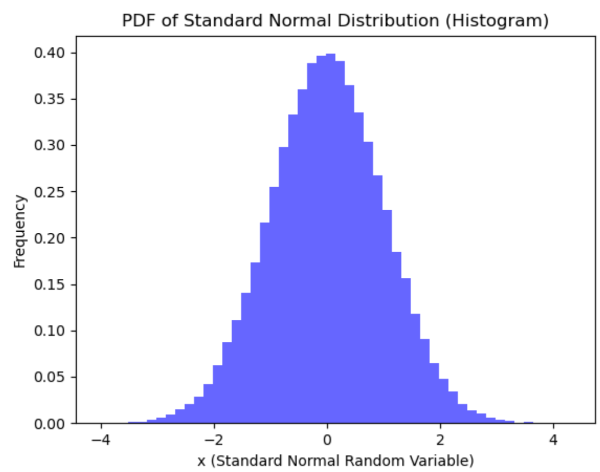

Fig. 2. Histogram plot of density of Z_0