PDF and CDF of uniform and gaussian distributions

Kushagra Gupta

I. Uniform Random Numbers

Let U be a uniform random variable between 0 and 1.

A. Q 1.1

Question: Generate 10^6 samples of U using a C program and save into a file called uni.dat .
Download the following files and execute the C program.

```
wget https://github.com/gadepall/probability/raw/
    master/manual/codes/exrand.c
wget https://github.com/gadepall/probability/raw/
    master/manual/codes/coeffs.h
```

Solution:

```
$ gcc exrand.c -lm
$ ./a.out
```

Note: The flag `-lm` is to tell gcc to include the math library.
This code creates the file ”uni.dat” which contains random data points for a uniform distribution.

B. Q 1.2

Question: Load the uni.dat file into python and plot the empirical CDF of U using the samples in uni.dat. The CDF is defined as

$$F_U(x) = \Pr(U \leq x) \tag{1}$$

Download the follwing file for plotting cdf.

```
wget https://github.com/gadepall/probability/raw/
    master/manual/codes/cdf_plot.py
```

Solution:

```
$ python3 cdf_plot.py
```

This plots the numerical part of fig.1

C. Q 1.3

Question: Find a theoretical expression for $F_U(x)$.
Solution: Given U is a uniformly distributed random variable over the interval $(0, 1)$, we have the density function $p_U(x)$:

$$p_U(x) = \begin{cases} 1, & x \in (0,1) \\ 0, & otherwise \end{cases} \tag{2}$$

We know that,

$$F_U(x) = \int_{-\infty}^{x} p_U(x)\,dx \tag{3}$$

$\therefore$ We have the following expression for $F_U(x)$:

$$F_U(x) = \begin{cases} 0, & x \in (-\infty, 0) \\ x, & x \in (0,1) \\ 1, & x \in (1, \infty) \end{cases} \tag{4}$$

D. Q 1.4

Question: The mean of U is defined as

$$E[U] = \frac{1}{N}\sum_{i=1}^{N} U_i \tag{5}$$

and its variance as

$$\text{var}[U] = E[U - E[U]]^2 \tag{6}$$

Write a C program to find the mean and variance of U.

Solution:

$$\text{We know that } E(U) = \int_{-\infty}^{\infty} x d(F_U(x)) \tag{7}$$

$$\implies E(U) = \int_0^1 x dx \tag{8}$$

$$\implies E(U) = 0.5 \tag{9}$$

$$E(U^1) = \int_{-\infty}^{\infty} x^2 d(F_u(x)) \tag{10}$$

$$\implies E(U^2) = \int_0^1 x^2 dx \tag{11}$$

$$\therefore E(U^2) = \frac{1}{3} \tag{12}$$

We know that

$$var(U) = E(U^2) - (E(U))^2 \tag{13}$$
$$\implies var(U) = \frac{1}{3} - \frac{1}{4} \tag{14}$$
$$\therefore var(U) = \frac{1}{12} = 0.0825 \tag{15}$$

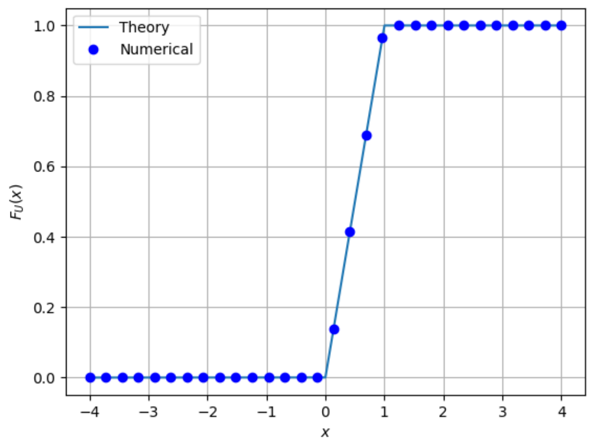

Fig. 1. The CDF of U

Code command are as follows:

```
$ gcc exrand.c -lm
$ ./a.out
```

Note: The flag `-lm` is to tell gcc to include the math library.

using ./code/exrand.c, we get the variance for the uniform distribution as 0.083301 and mean as 0.500007}

E. Q 1.5

Question: Verify your result theoretically given that

$$E\left[U^k\right] = \int_{-\infty}^{\infty} x^k dF_U(x) \tag{16}$$

We are given that

$$E[U^k] = \int_{-\infty}^{\infty} x^k dF_U(x) \tag{17}$$
$$\implies E[U^k] = \int_{-\infty}^{\infty} x^k p_U(x)\, dx \tag{18}$$

We know that mean μ is given by $E(U)$.
Hence,

$$\mu = \int_{-\infty}^{\infty} x p_U(x)\, dx \tag{19}$$
$$\mu = \int_0^1 x\, dx \tag{20}$$
$$= \frac{x^2}{2}\Big|_0^1 \tag{21}$$
$$= \frac{1}{2} \tag{22}$$

$$var(U) = E((U - E(U))^2) \tag{23}$$

This can also be represented as

$$var(U) = E(U^2 - 2E(U)U + (E(U))^2) \tag{24}$$
$$= E(U^2) - 2(E(U))^2 + (E(U))^2 \tag{25}$$
$$= E(U^2) - (E(U))^2 \tag{26}$$

We can evaluate $E(U^2)$ using (18) as:

$$E(U^2) = \int_{-\infty}^{\infty} x^2 p_U(x)\, dx \tag{27}$$
$$= \int_0^1 x^2\, dx \tag{28}$$
$$= \frac{x^3}{3}\Big|_0^1 \tag{29}$$
$$= \frac{1}{3} \tag{30}$$

Using (22) and (26) we have

$$var(U) = \frac{1}{3} - \frac{1}{4} = \frac{1}{12} \tag{31}$$

II. Central Limit Theorem

A. Q 2.1

Question: Generate 10^6 samples of the random variable

$$X = \sum_{i=1}^{12} U_i - 6 \tag{32}$$

using a C program, where $U_i, i = 1, 2, \ldots, 12$ are a set of independent uniform random variables between 0 and 1 and save in a file called gau.dat

Solution: Code command are as follows:

```
$ gcc exrand.c -lm
$ ./a.out
```

Note: The flag `-lm` is to tell gcc to include the math

library.
Note: The flag `-lm` is to tell gcc to include the math library.
This code creates the file "gau.dat" which contains random data points for a uniform distribution.

B. Q 2.2

Download the follwing file for plotting cdf.

```
wget https://github.com/gadepall/probability/raw/
    master/manual/codes/cdf_plot.py
```

Solution:
Code command are as follows:
`$ python3 cdf_plot.py`
This plots figure 2

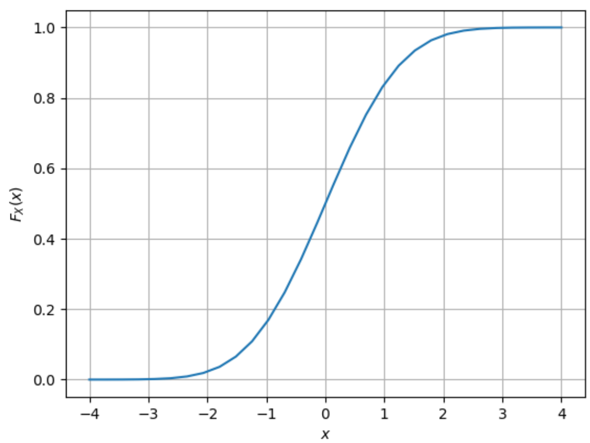

Fig. 2. The CDF of X

The properties of CDF are:

1) The CDF never decreases (cumulative)
2) $\lim_{x \to -\infty} F_X(x) = 0$
3) $\lim_{x \to \infty} F_X(x) = 1$

C. Q 2.3

Question: Load gau.dat in python and plot the empirical PDF of X using the samples in gau.dat. The PDF of X is defined as

$$p_X(x) = \frac{d}{dx} F_X(x) \tag{33}$$

Solution:

Code command are as follows:
`$ python3 pdf_plot.py`
This plots the figure 3.

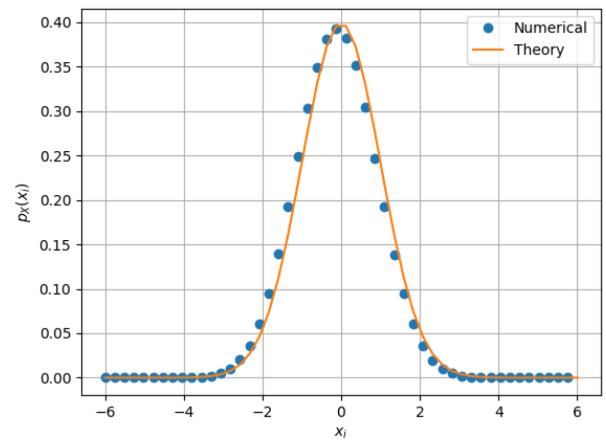

Fig. 3. The PDF of X

D. Q 2.4

Question: Find the mean and variance of X by writing a C program.
Solution:

> using code ./code/exrand.c, we get the mean as 0.000326 and variance as 0.000906.

E. Q 2.5

Question: Given that

$$p_X(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right), -\infty < x < \infty, \tag{34}$$

repeat the above exercise theoretically.
Solution: :
For random variable X, we have been given the density function as follows

$$p_X(x) = \frac{1}{\sqrt{2\pi}} \exp \frac{-x^2}{2} \tag{35}$$

$$\tag{36}$$

For finding μ,

$$\mu = \int_{-\infty}^{\infty} x.p_X(x)dx \tag{37}$$

$$= \int_{-\infty}^{\infty} x.\frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx \tag{38}$$

As $x.p_X(x)$ is an odd function the above integral is 0, therefore $\mu = 0$

For finding the variance var,

$$var = E(X^2) \tag{39}$$
$$= \int_{-\infty}^{\infty} x^2 . \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx \tag{40}$$
$$\text{let } x^2 = 2t \tag{41}$$
$$\implies var = \int_{-\infty}^{\infty} \sqrt{2t} . \frac{1}{\sqrt{2\pi}} e^{-t} dt \tag{42}$$
$$= \frac{2}{\sqrt{\pi}} \Gamma\left(\frac{3}{2}\right) \tag{43}$$
$$\implies var = 1 \tag{44}$$

III. From Uniform to Other

A. Q 3.1

Question: Generate samples of

$$V = -2 \ln (1 - U) \tag{45}$$

Solution:
Code command are as follows:

```
$ python3 cdf_plot.py
```

This plots figure 4

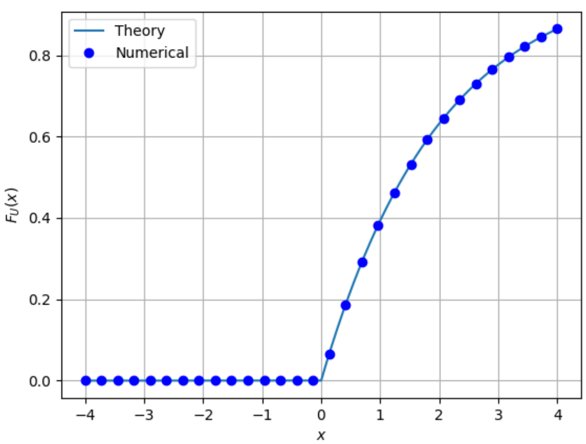

Fig. 4. The CDF of V

B. Q 3.2

Question: Find a theoretical expression for $F_V(x)$.
Solution:
We have been given that random variable V is a function of the random variable U as follows:

$$V = -2 \ln (1 - U) \tag{46}$$

Note that the obtained distribution function (CDF) for random variable U is:

$$F_U(x) = \begin{cases} 0, & x \in (-\infty, 0) \\ x, & x \in (0, 1) \\ 1, & x \in (1, \infty) \end{cases} \tag{47}$$

We know for any random variable X

$$F_X(x) = \Pr(X \leq x) \tag{48}$$

Hence, we can write:

$$F_V(x) = \Pr(V \leq x) \tag{49}$$
$$= \Pr(-2 \ln (1 - U) \leq x) \tag{50}$$
$$= \Pr(U \leq 1 - \exp \frac{-x}{2}) \tag{51}$$
$$= F_U(1 - \exp \frac{-x}{2}) \tag{52}$$

Note that the function $f(x) = 1 - \exp \frac{-x}{2}$ follows:

$$f(x) \in \begin{cases} 0, & x \in (-\infty, 0) \\ (0, 1) & x \in (0, \infty) \end{cases} \tag{53}$$

Hence we can write

$$F_V(x) = \begin{cases} 0, & x \in (-\infty, 0) \\ 1 - \exp \frac{-x}{2}, & x \in (0, \infty) \end{cases} \tag{54}$$