Random Numbers
AI1110: Probability and Random Variables

J Sai Sri Hari Vamshi
AI21BTECH11014

CONTENTS

***Abstract*—This manual provides a simple introduction to the generation of random numbers**

1 Uniform Random Numbers

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

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

```
wget https://github.com/HARI-donk-EY/
    Rand_nums/blob/main/codes/1/unigen.c
wget https://github.com/HARI-donk-EY/
    Rand_nums/blob/main/codes/coeffs.h
```

Use the below command in the terminal to run code.

```
gcc unigen.c -lm -o unigen.out
./unigen.out
```

1.2 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

$$F_U(x) = \Pr(U \le x) \tag{1.1}$$

Solution: The following code plots Fig. 1.2

```
wget https://github.com/HARI-donk-EY/
    Rand_nums/blob/main/codes/1/
    cdf_plot_uni.py
```

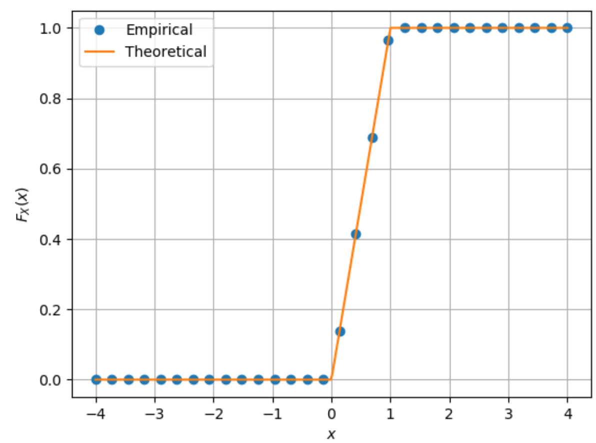

Fig. 1.2: The CDF of U

Use the below command in the terminal to run code.

```
python3 cdf_plot_uni.py
```

1.3 Find a theoretical expression for $F_U(x)$.
Solution: U is a Uniform Random Variable Distribution.
the PDF of the the distribution can be given as:

$$p_U(x) = \begin{cases} 1 & x \in [0,1] \\ 0 & \text{otherwise} \end{cases} \tag{1.2}$$

The CDF of U is given by

$$F_U(x) = \Pr(U \le x) = \int_{-\infty}^{x} p_U(x)\, dx \tag{1.3}$$

Therefore, we obtain the CDF of U as

$$F_U(x) = \begin{cases} 0 & x < 0 \\ x & 0 \le x \le 1 \\ 1 & x > 1 \end{cases} \tag{1.4}$$

1.4 The mean of U is defined as

$$E[U] = \frac{1}{N}\sum_{i=1}^{N} U_i \tag{1.5}$$

and its variance as

$$\text{var}[U] = E[U - E[U]]^2 \tag{1.6}$$

Write a C program to find the mean and variance of U.
Solution: Download the C source code by executing the following commands

```
wget https://github.com/HARI-donk-EY/
    Rand_nums/blob/main/codes/1/
    mean_var_cal.c
wget https://github.com/HARI-donk-EY/
    Rand_nums/blob/main/codes/coeffs.h
```

Use the below command in the terminal to run code.

```
gcc mean_var_cal.c -lm -o mean_var_cal.
    out
./mean_var_cal.out
```

From the code we get the output of the Mean as 0.500007, and Variance as 0.083301.

1.5 Verify your result theoretically given that

$$E\left[U^k\right] = \int_{-\infty}^{\infty} x^k dF_U(x) \tag{1.7}$$

Solution: We know that,

$$E[U] = \int_{-\infty}^{\infty} x dF_U(x) \tag{1.8}$$

On differentiation thr CDF obtained above, we get,

$$dF_U(x) = \begin{cases} 0 & x < 0 \\ dx & 0 \le x \le 1 \\ 0 & x > 1 \end{cases} \tag{1.9}$$

From this, we get,

$$E[U] = \int_0^1 x dx = \frac{1}{2} = 0.5 \tag{1.10}$$

Similarly,

$$E\left[U^2\right] = \int_0^1 x^2 dx = \frac{1}{3} \tag{1.11}$$

From varience,

$$\text{var}[U] = E\left[U^2\right] - (E[U])^2 \tag{1.12}$$

By substituting,

$$= \frac{1}{3} - \left(\frac{1}{2}\right)^2 \tag{1.13}$$

$$= \frac{1}{12} \approx 0.083333 \tag{1.14}$$

2 Central Limit Theorem

2.1 Generate 10^6 samples of the random variable

$$X = \sum_{i=1}^{12} U_i - 6 \tag{2.1}$$

using a C program, where $U_i, i = 1, 2, \dots, 12$ are a set of independent uniform random variables between 0 and 1 and save in a file called gau.dat
Solution: Download the C source code by executing the following commands

```
wget https://github.com/HARI-donk-EY/
    Rand_nums/blob/main/codes/2/gaugen.c
wget https://github.com/HARI-donk-EY/
    Rand_nums/blob/main/codes/coeffs.h
```

Compile and run the C program by executing the following

```
gcc gaugen.c -lm -o gaugen.out
./gaugen.out
```

2.2 Load gau.dat in python and plot the empirical CDF of X using the samples in gau.dat. What properties does a CDF have?
Solution: The following code plots Fig. 2.2

```
wget https://github.com/HARI-donk-EY/
    Rand_nums/blob/main/codes/2/
    cdf_plot_gau.py
```

Use the below command in the terminal to run code.

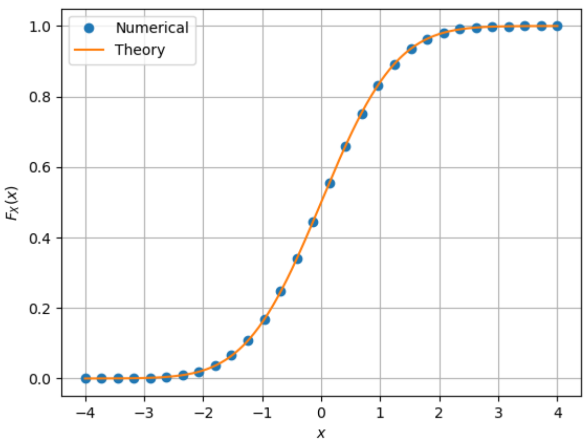

Fig. 2.2: The CDF of X

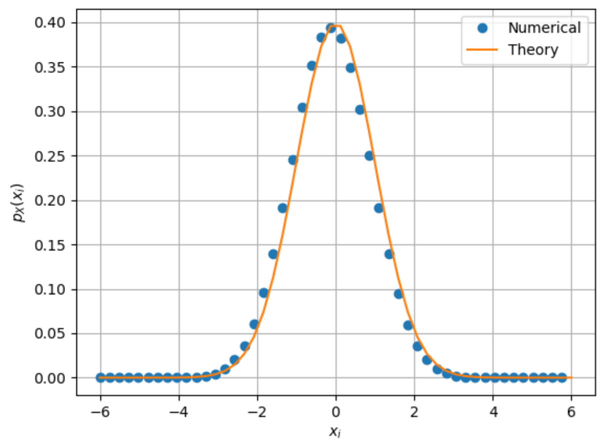

Fig. 2.3: The PDF of X

```
python3 cdf_plot_gau.py
```

Properties of this CDF are:

- $F_Z(x) = P(Z \leq x) = 1 - Q(x)$
- $\lim_{x\to\infty} F_Z(x) = 1, \quad \lim_{x\to-\infty} F_Z(x) = 0$
- $F_Z(0) = \frac{1}{2}$
- $F_Z(-x) = 1 - F_Z(x)$

2.3 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

$$p_X(x) = \frac{d}{dx} F_X(x) \tag{2.2}$$

What properties does the PDF have?
Solution: Download the following Python code that plots Fig. 2.3

```
wget https://github.com/HARI-donk-EY/
    Rand_nums/blob/main/codes/2/pdf_plot.
    py
```

Run the code by executing

```
python pdf_plot.py
```

The PDF graph is symmetric about $x = 0$ and bell shaped. The mean of the graph is situated at the apex of the bell.

2.4 Find the mean and variance of X by writing a C program.
Solution: Download the C source code by executing the following commands

```
wget https://github.com/HARI-donk-EY/
    Rand_nums/blob/main/codes/2/
    mean_var_cal.c
wget https://github.com/HARI-donk-EY/
    Rand_nums/blob/main/codes/coeffs.h
```

Use the below command in the terminal to run code.

```
gcc mean_var_cal.c -lm -o mean_var_cal.
    out
./mean_var_cal.out
```

From the code we get the output of the Mean as 0.000294, and Variance as 0.999560.

2.5 Given that

$$p_X(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right), -\infty < x < \infty, \tag{2.3}$$

repeat the above exercise theoretically.
Solution: From above, the Mean will be given as,

$$E[X] = \int_{-\infty}^{\infty} x p_X(x) dx \tag{2.4}$$

$$= \int_{-\infty}^{\infty} \frac{x}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) dx \tag{2.5}$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x e^{-\frac{x^2}{2}} dx \tag{2.6}$$

$$xe^{-\frac{x^2}{2}} \text{ is an odd function.}$$
$$\text{Hence, } E[X] = 0$$

Similarly,

$$E\left[X^2\right] = \int_{-\infty}^{\infty} x^2 p_X(x) dx \quad (2.7)$$
$$= \int_{-\infty}^{\infty} \frac{x^2}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) dx \quad (2.8)$$
$$= 2\int_{0}^{\infty} \frac{x^2}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) dx \quad (2.9)$$
$$= \sqrt{\frac{2}{\pi}} \int_{0}^{\infty} x\left(x.\exp\left(-\frac{x^2}{2}\right)\right) dx \quad (2.10)$$
$$(2.11)$$

Using integration by parts, we get,

$$= \sqrt{\frac{2}{\pi}}\left(x\int x\exp\left(-\frac{x^2}{2}\right)dx\right)\Big|_0^{\infty} - \sqrt{\frac{2}{\pi}}\int_0^{\infty} 1\cdot\int x\exp\left(-\frac{x^2}{2}\right)dx \quad (2.12)$$

$$= \sqrt{\frac{2}{\pi}}\left(\left[-x\exp-\frac{x^2}{2}\right]_0^{\infty} - \int_0^{\infty} -\exp\left(-\frac{x^2}{2}\right)\right)$$

Substituting x with $t\sqrt{2}$,

And dx with $dt\sqrt{2}$

$$= \sqrt{\frac{2}{\pi}}\left(0 - \left(-\sqrt{2}\int_0^{\infty} \exp(-t^2)\, dt\right)\right)$$
$$= \sqrt{\frac{2}{\pi}}\left(\sqrt{2}\int_0^{\infty} \exp(-t^2)\, dt\right)$$
$$= \sqrt{\frac{2}{\pi}}\left(\sqrt{2}\left(\frac{\sqrt{\pi}}{2}\right)\right)$$
$$= 1$$

Therefore, Variance of X will be,

$$var(X) = E[X]^2 - E\left[X^2\right]$$
$$= 0 - 1 = 1$$

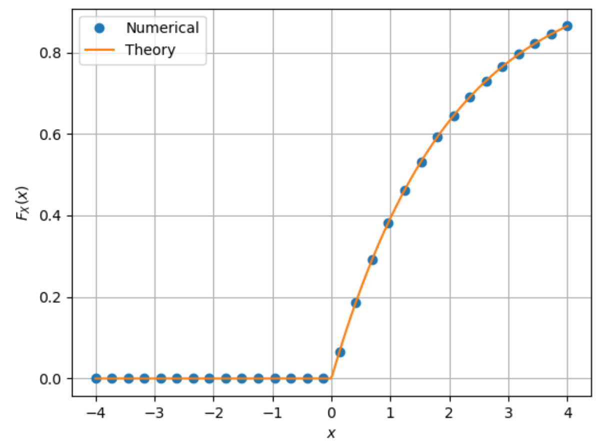

Fig. 3.1: The CDF of V

3 From Uniform to Other

3.1 Generate samples of

$$V = -2\ln(1 - U) \quad (3.1)$$

and plot its CDF.

Solution: Download the C source code by executing the following commands

```
wget https://github.com/HARI-donk-EY/
    Rand_nums/blob/main/codes/3/vgen.c
wget https://github.com/HARI-donk-EY/
    Rand_nums/blob/main/codes/coeffs.h
```

Compile and run the C program by executing the following

```
gcc vgen.c -lm -o vgen.out
./vgen.out
```

The following code plots Fig. 3.1

```
wget https://github.com/HARI-donk-EY/
    Rand_nums/blob/main/codes/3/
    cdf_plot_vdat.py
```

Use the below command in the terminal to run code.

```
python3 cdf_plot_vdat.py
```

3.2 Find a theoretical expression for $F_V(x)$.

Solution: We have

$$F_V(x) = \Pr(V \le x) \tag{3.2}$$
$$= \Pr(-2\ln(1-U) \le x) \tag{3.3}$$
$$= \Pr\left(\ln(1-U) \ge -\frac{x}{2}\right) \tag{3.4}$$
$$= \Pr\left(1-U \ge \exp\left(-\frac{x}{2}\right)\right) \tag{3.5}$$
$$= \Pr\left(U \le 1-\exp\left(-\frac{x}{2}\right)\right) \tag{3.6}$$
$$= F_U\left(1-\exp\left(-\frac{x}{2}\right)\right) \tag{3.7}$$

Now,

$$0 \le 1-\exp\left(-\frac{x}{2}\right) < 1 \qquad \text{if } x \ge 0 \tag{3.8}$$
$$1-\exp\left(-\frac{x}{2}\right) < 0 \qquad \text{if } x < 0 \tag{3.9}$$

Therefore,

$$F_V(x) = \begin{cases} 1-\exp\left(-\frac{x}{2}\right) & x \ge 0 \\ 0 & x < 0 \end{cases} \tag{3.10}$$

4 Triangular Distribution

4.1 Generate

$$T = U_1 + U_2$$

Solution: Download the C source code by executing the following commands

```
wget https://github.com/HARI-donk-EY/
    Rand_nums/blob/main/codes/4/tritrien.c
wget https://github.com/HARI-donk-EY/
    Rand_nums/blob/main/codes/coeffs.h
```

Compile and run the C program by executing the following

```
gcc trigen.c -lm -o trigen.out
./trigen.out
```

4.2 Find the *CDF* of T.
Solution: The *CDF* of T is given by:

$$F_T(t) = \Pr(T \le t) = \Pr(U_1 + U_2 \le t) \tag{4.1}$$

Since U_1, $U_2 \in [0,1]$, from which we get that, $U_1 + U_2 \in [0,2]$.
Therefore, if $t \ge 2$, $U_1 + U_2 \le t$ is always true and $U_1+U_2 \le t$ is always false if $t \le 0$, making their probabilities 1 and 0 respectivey.
For other cases, fix U_1 as x. Then,

$$x + U_2 \le t \implies U_2 \le t - x$$

If $0 \le t \le 1$, $x \in [0,t]$,

$$F_T(t) = \int_0^t \Pr(U_2 \le t-x)\, p_{U_1}(x)\ \mathrm{dx}$$
$$= \int_0^t F_{U_2}(t-x) p_{U_1}(x)\ \mathrm{dx}$$

Since,

$$0 \le x \le t \le 1 \implies t - x \le 1$$
$$\implies F_{U_2}(t-x) = t - x$$

We get,

$$F_T(t) = \int_0^t (t-x).1\ \mathrm{dx}$$
$$= \left[tx - \frac{x^2}{2}\right]_0^t$$
$$= \frac{t^2}{2}$$

similarly, if $1 < t < 2$, $x \in [0,1]$,

$$t < 2 \implies t - 1 < 1$$
$$\text{For,}\quad 0 < x < t-1 \implies 1 \le t-x \le t$$
$$\text{And for,}\quad t-1 < x < 1 \implies 0 < t-1 \le t-x \le 1$$

By substituting,

$$F_T(t) = \int_0^{t-1} 1\ \mathrm{dx} + \int_{t-1}^1 (t-x)\ \mathrm{dx}$$
$$= t - 1 + t - t^2 - t - \frac{1}{2} + \frac{(t-1)^2}{2}$$
$$= \frac{-t^2}{2} + 2t - 1$$

Therefore,

$$F_T(t) = \begin{cases} 0 & t < 0 \\ \frac{t^2}{2} & 0 \le t \le 1 \\ 2t - \frac{t^2}{2} - 1 & 1 < t < 2 \\ 1 & t \ge 2 \end{cases}$$

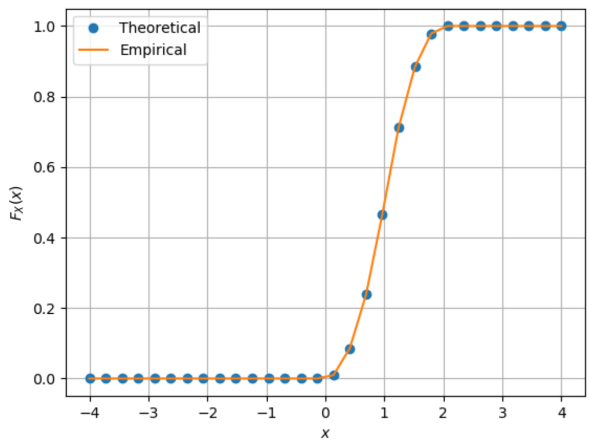

Fig. 4.2: The CDF of T

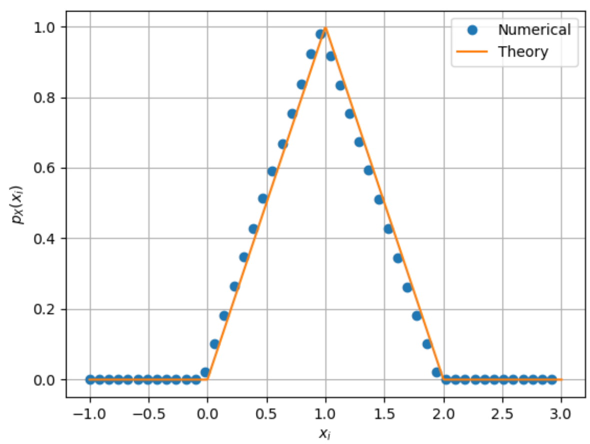

Fig. 4.3: The PDF of T

4.3 Find the PDF of T.

Solution: We know that *PDF* of T can be given as,

$$p_T(t) = \frac{\mathrm{d}}{\mathrm{dt}} F_T(t)$$

Therefore,

$$p_T(t) = \begin{cases} 0 & t < 0 \\ t & 0 \le t \le 1 \\ 2 - t & 1 < t < 2 \\ 0 & t \ge 2 \end{cases}$$

4.4 Find the theoretical expressions for the PDF and CDF of T.

Solution: The theoretical expressions for the CDF and PDF are as follws,
The CDF is,

$$F_T(t) = \begin{cases} 0 & t < 0 \\ \frac{t^2}{2} & 0 \le t \le 1 \\ 2t - \frac{t^2}{2} - 1 & 1 < t < 2 \\ 1 & t \ge 2 \end{cases}$$

The PDF is,

$$p_T(t) = \begin{cases} 0 & t < 0 \\ t & 0 \le t \le 1 \\ 2 - t & 1 < t < 2 \\ 0 & t \ge 2 \end{cases}$$

4.5 Verify your results through a plot.

Solution: Download the following Python codes that plot Fig. 4.2 and Fig. 4.3

```
wget https://github.com/HARI-donk-EY/
    Rand_nums/blob/main/codes/4/
    cdf_plot_tri.py
wget https://github.com/HARI-donk-EY/
    Rand_nums/blob/main/codes/4/pdf_plot.
    py
```

Run the code by executing

```
python3 cdf_plot_tri.py
python3 pdf_plot.py
```