Random Numbers

AI1110: Probability and Random Variables

Indian Institute of Technology Hyderabad

Ankit Saha
AI21BTECH11004

30 Jun 2022

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 C source code by executing the following commands

```
wget https://github.com/Ankit−Saha−2003/
    AI1110/raw/main/Random−Numbers/
    codes/1.1.c
wget https://github.com/Ankit−Saha−2003/
    AI1110/raw/main/Random−Numbers/
    codes/header.h
```

Compile and run the C program by executing the following

```
cc −lm 1.1.c
./a.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: Download the following Python code that plots Fig. 1.2

```
wget https://github.com/Ankit−Saha−2003/
    AI1110/raw/main/Random−Numbers/
    codes/1.2.py
```

Run the code by executing

```
python 1.2.py
```

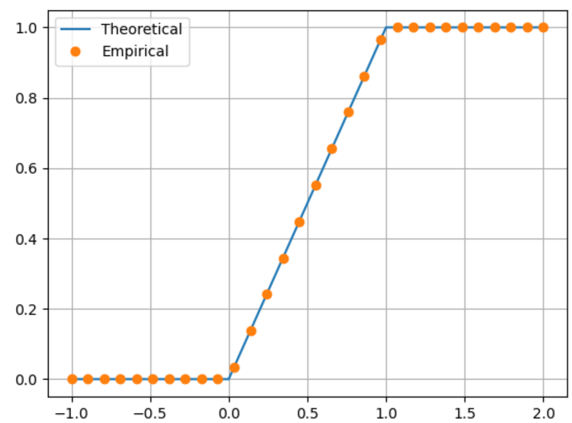

Fig. 1.2. The CDF of U

1.3 Find a theoretical expression for $F_U(x)$

Solution: The PDF of U is given by

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

The CDF of U is given by

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

If $x < 0$,

$$\int_{-\infty}^{x} p_U(x)\,\mathrm{d}x = \int_{-\infty}^{x} 0\,\mathrm{d}x = 0 \tag{1.4}$$

If c,

$$\int_{-\infty}^{x} p_U(x)\,\mathrm{d}x = \int_{-\infty}^{0} 0\,\mathrm{d}x + \int_{0}^{x} 1\,\mathrm{d}x \tag{1.5}$$

$$= 0 + x \tag{1.6}$$

$$= x \tag{1.7}$$

If $x > 1$,

$$\int_{-\infty}^{x} p_U(x)\ \mathrm{d}x = \int_{-\infty}^{0} 0\ \mathrm{d}x + \int_{0}^{1} 1\ \mathrm{d}x + \int_{1}^{x} 0\ \mathrm{d}x \tag{1.8}$$

$$\int_{-\infty}^{x} p_U(x)\ \mathrm{d}x = 0 + 1 + 0 \tag{1.9}$$

$$= 1 \tag{1.10}$$

Therefore, we obtain the CDF of U as

$$F_U(x) = \begin{cases} 0 & x < 0 \\ x & 0 \le x \le 1 \\ 1 & x > 1 \end{cases} \tag{1.11}$$

1.4 The mean of U is defined as

$$E[U] = \frac{1}{N}\sum_{i=1}^{N} U_i \tag{1.12}$$

and its variance as

$$\mathrm{Var}[U] = E[U - E[U]]^2 \tag{1.13}$$

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/Ankit−Saha−2003/
    AI1110/raw/main/Random−Numbers/
    codes/1.4.c
wget https://github.com/Ankit−Saha−2003/
    AI1110/raw/main/Random−Numbers/
    codes/header.h
```

Compile and run the C program by executing the following

```
cc −lm 1.4.c
./a.out
```

The output of the code is

$$\mu_{\text{emp}} = 0.500007 \tag{1.14}$$

$$\mu_{\text{the}} = 0.500000 \tag{1.15}$$

$$\sigma^2_{\text{emp}} = 0.083301 \tag{1.16}$$

$$\sigma^2_{\text{the}} = 0.083333 \tag{1.17}$$

1.5 Verify your result theoretically given that

$$E\left[U^k\right] = \int_{-\infty}^{\infty} x^k \mathrm{d}F_U(x) \tag{1.18}$$

Solution: The mean of U is given by

$$E[U] = \int_{-\infty}^{\infty} x\ \mathrm{d}F_U(x) \tag{1.19}$$

On differentiating the CDF of U, we get

$$\mathrm{d}F_U(x) = \begin{cases} 0 & x < 0 \\ \mathrm{d}x & 0 \le x \le 1 \\ 0 & x > 1 \end{cases} \tag{1.20}$$

$$\therefore E[U] = \int_{0}^{1} x\ \mathrm{d}x = \frac{1}{2} = 0.5 \tag{1.21}$$

Similarly,

$$\therefore E\left[U^2\right] = \int_{0}^{1} x^2\ \mathrm{d}x = \frac{1}{3} \tag{1.22}$$

Now, the variance of U is given by

$$\mathrm{Var}[U] \tag{1.23}$$

$$= E[U - E[U]]^2 \tag{1.24}$$

$$= E\left[U^2 - 2UE[U] + (E[U])^2\right] \tag{1.25}$$

By linearity of expectation, we have

$$E\left[U^2\right] + E[-2UE[U]] + E\left[(E[U])^2\right] \tag{1.26}$$

$$= E\left[U^2\right] - 2E[U]E[U] + (E[U])^2 \tag{1.27}$$

$$= E\left[U^2\right] - (E[U])^2 \tag{1.28}$$

$$= \frac{1}{3} - \left(\frac{1}{2}\right)^2 \tag{1.29}$$

$$= \frac{1}{12} \approx 0.083333 \tag{1.30}$$

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 $U_i, i = 1, 2, \ldots, 12$ 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/Ankit−Saha−2003/
    AI1110/raw/main/Random−Numbers/
    codes/2.1.c
wget https://github.com/Ankit−Saha−2003/
    AI1110/raw/main/Random−Numbers/
    codes/header.h
```

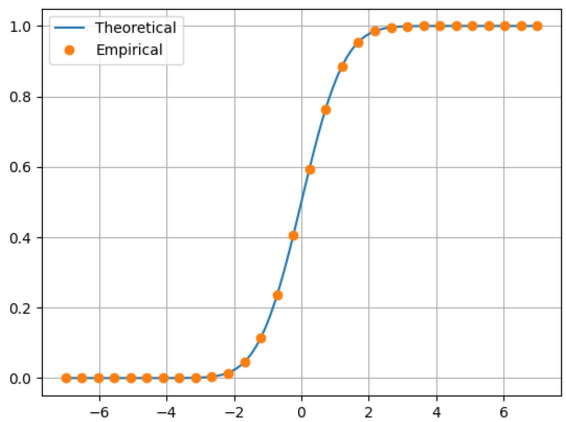

Fig. 2.2. The CDF of X

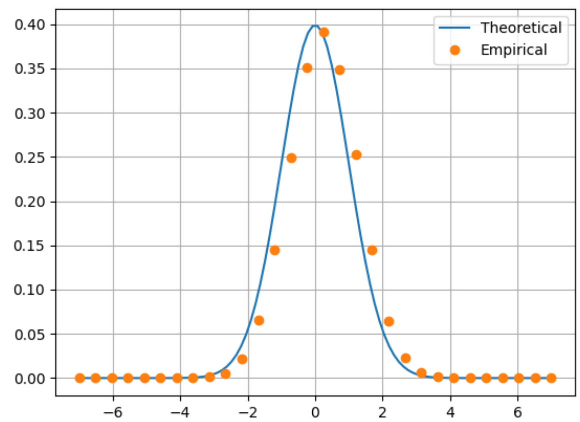

Fig. 2.3. The PDF of X

Compile and run the C program by executing the following

```
cc -lm 2.1.c
./a.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: Download the following Python code that plots Fig. 2.2

```
wget https://github.com/Ankit-Saha-2003/
    AI1110/raw/main/Random-Numbers/
    codes/2.2.py
```

Run the code by executing

```
python 2.2.py
```

Every CDF is monotone increasing and right-continuous. Furthermore,

$$\lim_{x\to-\infty} F_X(x) = 0 \qquad \lim_{x\to\infty} F_X(x) = 1 \tag{2.2}$$

Thus, every CDF is bounded between 0 and 1 and hence, convergent.
In this case, the CDF is also left-continuous. Therefore, X is a continuous random variable.

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

$$p_X(x) = \frac{\mathrm{d}}{\mathrm{d}x} F_X(x) \tag{2.3}$$

What properties does the PDF have?
Solution: Download the following Python code that plots Fig. 2.2

```
wget https://github.com/Ankit-Saha-2003/
    AI1110/raw/main/Random-Numbers/
    codes/2.3.py
```

Run the code by executing

```
python 2.3.py
```

Every PDF is bounded between 0 and 1 and

$$\int_{-\infty}^{\infty} p_X(x)\ \mathrm{d}x = 1 \tag{2.4}$$

In this case, the PDF is symmetric about $x = 0$

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/Ankit-Saha-2003/
    AI1110/raw/main/Random-Numbers/
    codes/2.4.c
wget https://github.com/Ankit-Saha-2003/
    AI1110/raw/main/Random-Numbers/
    codes/header.h
```

Compile and run the C program by executing the following

```
cc -lm 2.4.c
./a.out
```

The output of the code is

$$\mu_{\text{emp}} = 0.000294 \tag{2.5}$$
$$\mu_{\text{the}} = 0.000000 \tag{2.6}$$
$$\sigma^2_{\text{emp}} = 0.999560 \tag{2.7}$$
$$\sigma^2_{\text{the}} = 1.000000 \tag{2.8}$$

2.5 Given that

$$p_X(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right), -\infty < x < \infty, \tag{2.9}$$

repeat the above exercise theoretically
Solution: The mean of X is given by

$$E[X] = \int_{-\infty}^{\infty} x p_X(x) \mathrm{d}x \tag{2.10}$$
$$= \int_{-\infty}^{\infty} \frac{x}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) \mathrm{d}x \tag{2.11}$$

Now, let

$$g(x) = \frac{x}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) \tag{2.12}$$
$$\implies g(-x) = \frac{-x}{\sqrt{2\pi}} \exp\left(-\frac{(-x)^2}{2}\right) \tag{2.13}$$
$$= -\frac{x}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) \tag{2.14}$$
$$= -g(x) \tag{2.15}$$

Thus, $g(x)$ is an odd function

$$\therefore E[X] = \int_{-\infty}^{\infty} g(x) \mathrm{d}x = 0 \tag{2.16}$$

Now,

$$E\left[X^2\right] = \int_{-\infty}^{\infty} x^2 p_X(x) \mathrm{d}x \tag{2.17}$$
$$= \int_{-\infty}^{\infty} \frac{x^2}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) \mathrm{d}x \tag{2.18}$$
$$= 2\int_{0}^{\infty} \frac{x^2}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) \mathrm{d}x \tag{2.19}$$

since $\frac{x^2}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right)$ is an even function
Using integration by parts,

$$E\left[X^2\right] = \sqrt{\frac{2}{\pi}} \int_{0}^{\infty} x \cdot x \exp\left(-\frac{x^2}{2}\right) \mathrm{d}x \tag{2.20}$$

$$= \sqrt{\frac{2}{\pi}} \left(x \int x \exp\left(-\frac{x^2}{2}\right) \mathrm{d}x \right)\Bigg|_0^{\infty} - \sqrt{\frac{2}{\pi}} \int_0^{\infty} 1 \cdot \int x \exp\left(-\frac{x^2}{2}\right) \mathrm{d}x \tag{2.21}$$

Substitute $t = \frac{x^2}{2} \implies \mathrm{d}t = x\mathrm{d}x$

$$\int x \exp\left(-\frac{x^2}{2}\right) \mathrm{d}x = \int \exp(-t) \mathrm{d}t \tag{2.22}$$
$$= -\exp(-t) \tag{2.23}$$
$$= -\exp\left(-\frac{x^2}{2}\right) \tag{2.24}$$

Now,

$$-x \exp\left(-\frac{x^2}{2}\right)\Bigg|_0^{\infty} = 0 - 0 = 0 \tag{2.25}$$
$$\because \lim_{x\to\infty} x \exp\left(-\frac{x^2}{2}\right) = \lim_{x\to\infty} \frac{x}{\exp\left(\frac{x^2}{2}\right)} = 0 \tag{2.26}$$

as exponential function grows much faster than a polynomial function
Also,

$$\int_0^{\infty} -\exp\left(-\frac{x^2}{2}\right) \mathrm{d}x \tag{2.27}$$
$$\xleftrightarrow{x=t\sqrt{2}} \int_0^{\infty} -\exp(-t^2) \mathrm{d}t \sqrt{2} \tag{2.28}$$
$$= -\sqrt{2} \int_0^{\infty} \exp(-t^2) \mathrm{d}t \tag{2.29}$$
$$= -\sqrt{2} \frac{\sqrt{\pi}}{2} \tag{2.30}$$
$$= -\sqrt{\frac{\pi}{2}} \tag{2.31}$$

Therefore,

$$E\left[X^2\right] = 0 - \sqrt{\frac{2}{\pi}} \left(-\sqrt{\frac{\pi}{2}}\right) \tag{2.32}$$
$$= 1 \tag{2.33}$$
$$\therefore \mathrm{Var}[X] = E\left[X^2\right] - (E[X])^2 \tag{2.34}$$
$$= 1 - 0 \tag{2.35}$$
$$= 1 \tag{2.36}$$

3. From Uniform to Other

3.1 Generate samples of

$$V = -2 \ln(1 - U) \tag{3.1}$$

and plot its CDF
Solution: Download the C source code by executing the following commands

```
wget https://github.com/Ankit-Saha-2003/
    AI1110/raw/main/Random-Numbers/
    codes/3.1.c
```

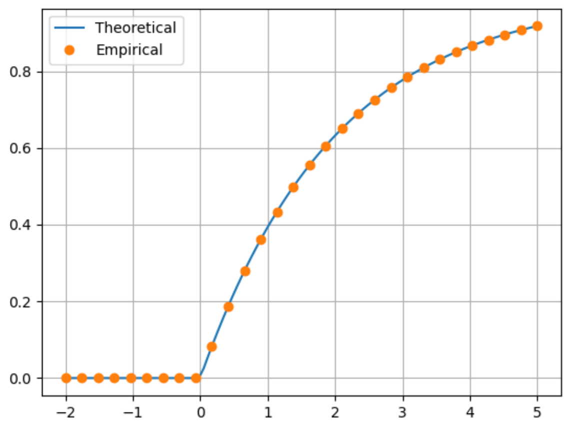

Fig. 3.1. The CDF of V

Compile and run the C program by executing the following

```
cc −lm 3.1.c
./a.out
```

Download the following Python code that plots Fig. 3.1

```
wget https://github.com/Ankit−Saha−2003/
    AI1110/raw/main/Random−Numbers/
    codes/3.1.py
```

Run the code by executing

```
python 3.1.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

$$T = U_1 + U_2 \tag{4.1}$$

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

```
wget https://github.com/Ankit−Saha−2003/
    AI1110/raw/main/Random−Numbers/
    codes/4.1.c
```

Compile and run the C program by executing the following

```
cc −lm 4.1.c
./a.out
```

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

$$F_T(t) = \Pr(T \le t) = \Pr(U_1 + U_2 \le t) \tag{4.2}$$

Since $U_1, U_2 \in [0,1] \implies U_1 + U_2 \in [0,2]$
Therefore, if $t \ge 2$, then $U_1 + U_2 \le t$ is always true and if $t < 0$, then $U_1 + U_2 \le t$ is always false.
Now, fix the value of U_1 to be some x

$$x + U_2 \le t \implies U_2 \le t - x \tag{4.3}$$

If $0 \le t \le 1$, then x can take all values in $[0,t]$

$$F_T(t) = \int_0^t \Pr(U_2 \le t-x)\, p_{U_1}(x) \mathrm{d}x \tag{4.4}$$
$$= \int_0^t F_{U_2}(t-x) p_{U_1}(x) \mathrm{d}x \tag{4.5}$$

$$0 \le x \le t \implies 0 \le t-x \le t \le 1 \tag{4.6}$$
$$\implies F_{U_2}(t-x) = t-x \tag{4.7}$$

$$F_T(t) = \int_0^t (t-x)\cdot 1 \cdot \mathrm{d}x \tag{4.8}$$
$$= tx - \left.\frac{x^2}{2}\right|_0^t \tag{4.9}$$
$$= \frac{t^2}{2} \tag{4.10}$$

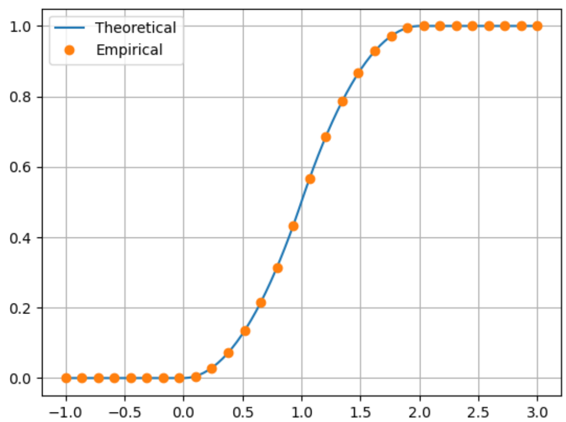

Fig. 4.2. The CDF of T

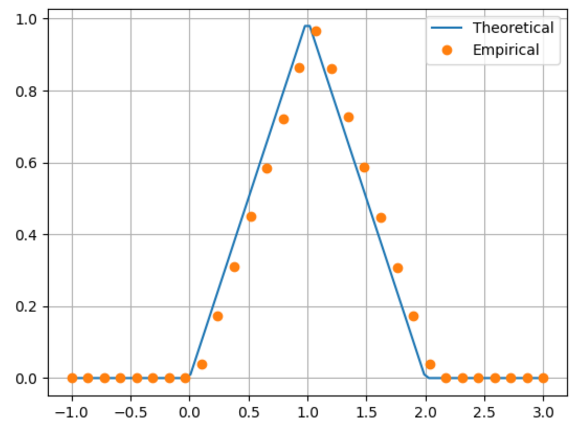

Fig. 4.3. The PDF of T

If $1 < t < 2$, x can only take values in $[0, 1]$ as $U_1 \leq 1$

$$F_T(t) = \int_0^1 F_{U_2}(t-x) \cdot 1 \cdot \mathrm{d}x \tag{4.11}$$

$$0 \leq x \leq t-1 \implies 1 \leq t-x \leq t \tag{4.12}$$

$$t-1 \leq x \leq 1 \implies 0 < t-1 \leq t-x \leq 1 \tag{4.13}$$

$$F_T(t) = \int_0^{t-1} 1\mathrm{d}x + \int_{t-1}^1 (t-x)\mathrm{d}x \tag{4.14}$$

$$= t-1+t(1-(t-1)) - \frac{1}{2} + \frac{(t-1)^2}{2} \tag{4.15}$$

$$= t-1+2t-t^2-\frac{1}{2}+\frac{t^2}{2}+\frac{1}{2}-t \tag{4.16}$$

$$= -\frac{t^2}{2}+2t-1 \tag{4.17}$$

Therefore,

$$F_T(t) = \begin{cases} 0 & t < 0 \\ \frac{t^2}{2} & 0 \leq t \leq 1 \\ 2t - \frac{t^2}{2} - 1 & 1 < t < 2 \\ 1 & t \geq 2 \end{cases} \tag{4.18}$$

4.3 Find the PDF of T

Solution: The PDF of T is given by

$$p_T(t) = \frac{\mathrm{d}}{\mathrm{d}t} F_T(t) \tag{4.19}$$

$$\therefore p_T(t) = \begin{cases} 0 & t < 0 \\ t & 0 \leq t \leq 1 \\ 2-t & 1 < t < 2 \\ 0 & t \geq 2 \end{cases} \tag{4.20}$$

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

Solution: The theoretical expressions for the CDF and PDF have been found in problems 4.2 and 4.3 respectively

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/Ankit-Saha-2003/
    AI1110/raw/main/Random-Numbers/
    codes/4.2.py
wget https://github.com/Ankit-Saha-2003/
    AI1110/raw/main/Random-Numbers/
    codes/4.3.py
```

Run the codes by executing

```
python 4.2.py
python 4.3.py
```