PDF and CDF of uniform and gaussian distributions

Kushagra Gupta

1 Uniform Random Numbers

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

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.

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

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

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

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 $p_U(x)$:

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

We know that,

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

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

$$F_U(x) = \begin{cases} 0, & x \in (-\infty, 0) \\ x, & x \in (0,1) \\ 1, & x \in (1,\infty) \end{cases} \tag{1.4}$$

1.4 **Question:** 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:

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

$$\text{We know that } E(U) = \int_{-\infty}^{\infty} xd(F_U(x)) \tag{1.7}$$

$$\implies E(U) = \int_{0}^{1} xdx \tag{1.8}$$

$$\implies E(U) = 0.5 \tag{1.9}$$

$$E(U^1) = \int_{-\infty}^{\infty} x^2 d(F_u(x)) \tag{1.10}$$

$$\implies E(U^2) = \int_{0}^{1} x^2 dx \tag{1.11}$$

$$\therefore E(U^2) = \frac{1}{3} \tag{1.12}$$

We know that

$$var(U) = E(U^2) - (E(U))^2 \tag{1.13}$$

$$\implies var(U) = \frac{1}{3} - \frac{1}{4} \tag{1.14}$$

$$\therefore var(U) = \frac{1}{12} = 0.0825 \tag{1.15}$$

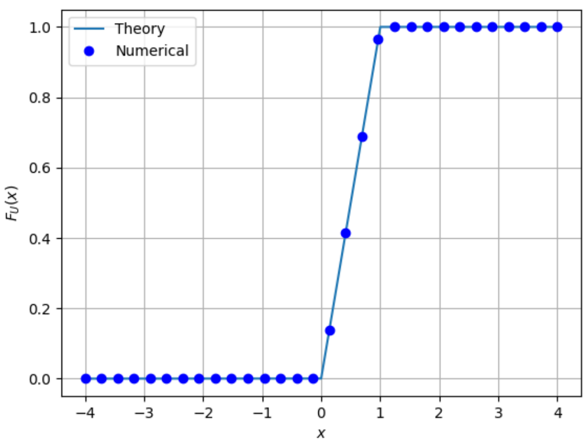

Fig. 1: The CDF of U

Code command are as follows:

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

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

1.5 **Question:**Verify your result theoretically given that

$$E\left[U^k\right] = \int_{-\infty}^{\infty} x^k dF_U(x) \tag{1.16}$$

We are given that

$$E[U^k] = \int_{-\infty}^{\infty} x^k dF_U(x) \tag{1.17}$$

$$\implies E[U^k] = \int_{-\infty}^{\infty} x^k p_U(x)\, dx \tag{1.18}$$

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

$$\mu = \int_{-\infty}^{\infty} x p_U(x)\, dx \tag{1.19}$$

$$\mu = \int_0^1 x\, dx \tag{1.20}$$

$$= \frac{x^2}{2}\Big|_0^1 \tag{1.21}$$

$$= \frac{1}{2} \tag{1.22}$$

$$var(U) = E((U - E(U))^2) \tag{1.23}$$

This can also be represented as

$$var(U) = E(U^2 - 2E(U)U + (E(U))^2) \tag{1.24}$$

$$= E(U^2) - 2(E(U))^2 + (E(U))^2 \tag{1.25}$$

$$= E(U^2) - (E(U))^2 \tag{1.26}$$

We can evaluate $E(U^2)$ using (1.18) as:

$$E(U^2) = \int_{-\infty}^{\infty} x^2 p_U(x)\, dx \tag{1.27}$$

$$= \int_0^1 x^2\, dx \tag{1.28}$$

$$= \frac{x^3}{3}\Big|_0^1 \tag{1.29}$$

$$= \frac{1}{3} \tag{1.30}$$

Using (1.22) and (1.26) we have

$$var(U) = \frac{1}{3} - \frac{1}{4} = \frac{1}{12} \tag{1.31}$$

2 Central Limit Theorem

2.1 **Question:** 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: 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.

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

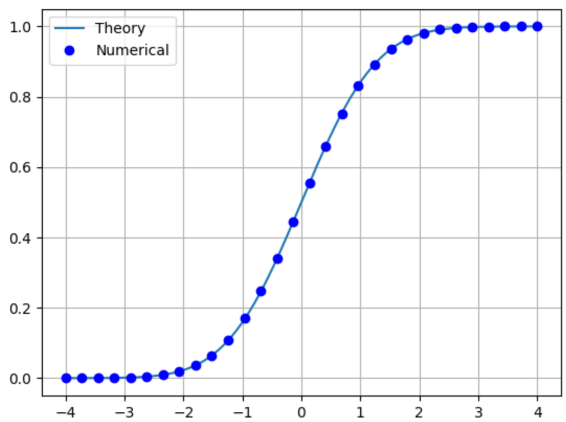

Fig. 2: The CDF of X

The properties of CDF are:

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

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

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

Solution:

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

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

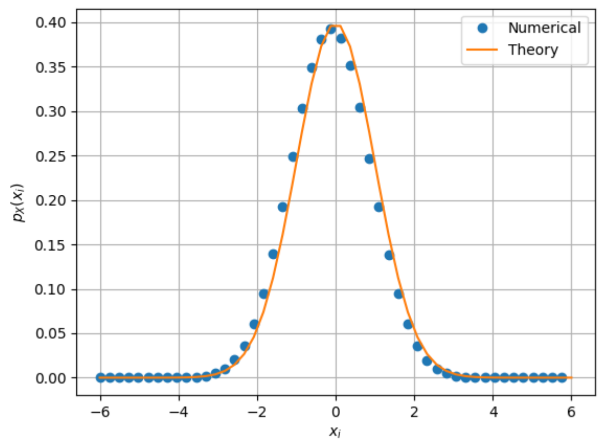

Fig. 2: The PDF of X

Solution:

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

2.5 **Question:** 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: :
For random variable X, we have been given the density function as follows

$$p_X(x) = \frac{1}{\sqrt{2\pi}}\exp\frac{-x^2}{2} \tag{2.4}$$

$$\tag{2.5}$$

For finding μ,

$$\mu = \int_{-\infty}^{\infty} x.p_X(x)dx \tag{2.6}$$

$$= \int_{-\infty}^{\infty} x.\frac{1}{\sqrt{2\pi}}e^{-x^2/2}dx \tag{2.7}$$

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

For finding the variance var,

$$var = E(X^2) \tag{2.8}$$
$$= \int_{-\infty}^{\infty} x^2.\frac{1}{\sqrt{2\pi}}e^{-x^2/2}dx \tag{2.9}$$
$$\text{let } x^2 = 2t \tag{2.10}$$
$$\implies var = \int_{-\infty}^{\infty} \sqrt{2t}.\frac{1}{\sqrt{2\pi}}e^{-t}dt \tag{2.11}$$
$$= \frac{2}{\sqrt{\pi}}\Gamma\left(\frac{3}{2}\right) \tag{2.12}$$
$$\implies var = 1 \tag{2.13}$$

2.6 Theoritical expression of CDF in terms of Q function is

$$Q_X(x) = \int_{x}^{\infty} \frac{1}{\sqrt{2\pi}}e^{\frac{-x^2}{2}}dx \tag{2.14}$$
$$= \frac{\text{erfc}(\frac{x}{\sqrt{2}})}{2} \tag{2.15}$$

The CDF is then:

$$F_X(x) = 1 - Q_X(x) \tag{2.16}$$
$$= 1 - \frac{\text{erfc}(\frac{x}{\sqrt{2}})}{2} \tag{2.17}$$

3 From Uniform to Other

3.1 **Question:**Generate samples of

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

Solution:
Code command are as follows:
`$ python3 cdf_plot.py`
This plots figure 4

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:

$$V = -2\ln(1-U) \tag{3.2}$$

Note that the obtained distribution function

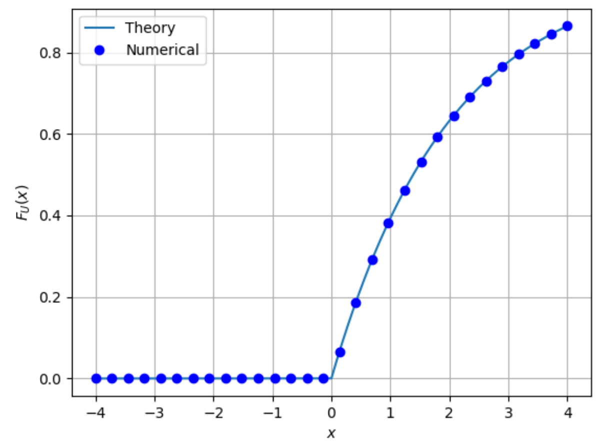

Fig. 3: The CDF of V

(CDF) for random variable U is:

$$F_U(x) = \begin{cases} 0, & x \in (-\infty, 0) \\ x, & x \in (0,1) \\ 1, & x \in (1,\infty) \end{cases} \tag{3.3}$$

We know for any random variable X

$$F_X(x) = \Pr(X \le x) \tag{3.4}$$

Hence, we can write:

$$F_V(x) = \Pr(V \le x) \tag{3.5}$$
$$= \Pr(-2\ln(1-U) \le x) \tag{3.6}$$
$$= \Pr(U \le 1 - \exp\frac{-x}{2}) \tag{3.7}$$
$$= F_U(1 - \exp\frac{-x}{2}) \tag{3.8}$$

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

$$f(x) \in \begin{cases} 0, & x \in (-\infty, 0) \\ (0,1) & x \in (0,\infty) \end{cases} \tag{3.9}$$

Hence we can write

$$F_V(x) = \begin{cases} 0, & x \in (-\infty, 0) \\ 1 - \exp\frac{-x}{2}, & x \in (0,\infty) \end{cases} \tag{3.10}$$

4 Triangular Distribution

4.1 **Question:** Generate

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

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 "tri.dat" which contains random data points for a triangular distribution.

4.2 **Question:** Find the CDF of T
Solution:
Code command are as follows:

```
$ python3 cdf_plot.py
```

This plots figure 5
e

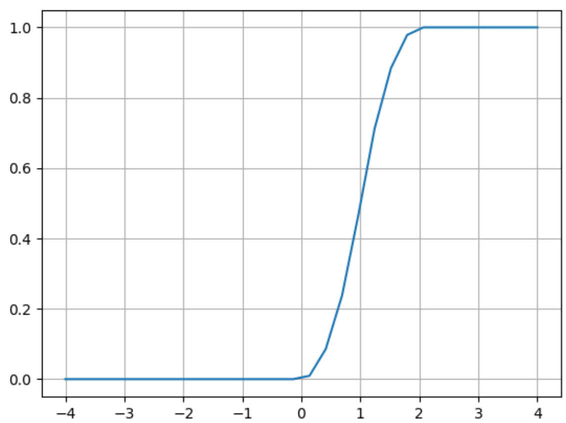

Fig. 4: CDF of T

4.3 **Question:** Find the PDF of T
Solution:
Code command are as follows:

```
$ python3 pdf_plot.py
```

This plots figure 6

4.4 **Question:** Find the theoritical expressions for the PDF and CDF of T.
Solution:

$$T = U_1 + U_2 \tag{4.2}$$

where T, U_1 and U_2 are random variables, we have:

$$p_T(t) = (p_{U_1} * p_{U_2})(t) \tag{4.3}$$

$$\tag{4.4}$$

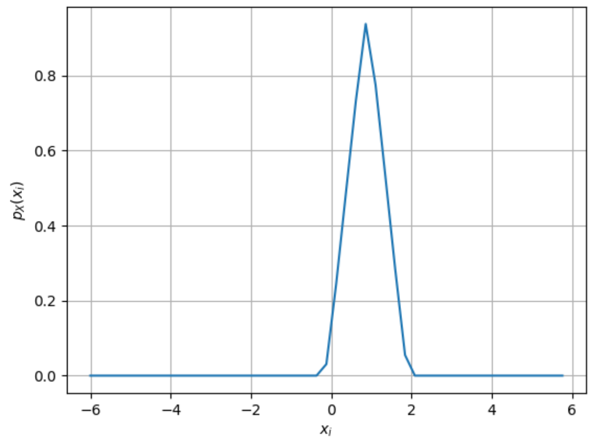

Fig. 4: PDF of T

Here, $p_{U_1}(t) = p_{U_2}(t) = p_U(t)$

$$p_T(t) = \int_{-\infty}^{\infty} p_U(\tau)p_U(t-\tau)d\tau \tag{4.5}$$

$$\tag{4.6}$$

When $t < 0$ and $t > 2$, the integral evaluates to 0. When $0 < t < 1$:

$$p_T(t) = \int_0^1 p_U(t-\tau)d\tau \tag{4.7}$$

$$= \int_0^t p_U(t-\tau)d\tau \tag{4.8}$$

$$= t \tag{4.9}$$

when $1 < t < 2$:

$$p_T(t) = \int_{t-1}^1 p_U(t-\tau)d\tau \tag{4.10}$$

$$= 2 - t \tag{4.11}$$

Therefore, we have:

$$p_T(x) = \begin{cases} x, & x \in (0,1) \\ 2-x, & x \in (1,2) \\ 0, & otherwise \end{cases} \tag{4.12}$$

To find the CDF, we use:

$$F_T(x) = \int_{-\infty}^{x} p_T(t)dt \tag{4.13}$$

We get:

$$F_T(x) = \begin{cases} 0, & x \in (-\infty, 0) \\ \frac{x^2}{2}, & x \in (0, 1) \\ -\frac{x^2}{2} + 2x - 1, & x \in (1, 2) \\ 1, & x \in (2, \infty) \end{cases} \tag{4.14}$$

4.5 **Question:** Verify your results through a plot.
Solution:
Code commands are as follows:

```
$ python3 cdf_plot.py
$ python3 pdf_plot.py
```

This plots figure 7 and 8

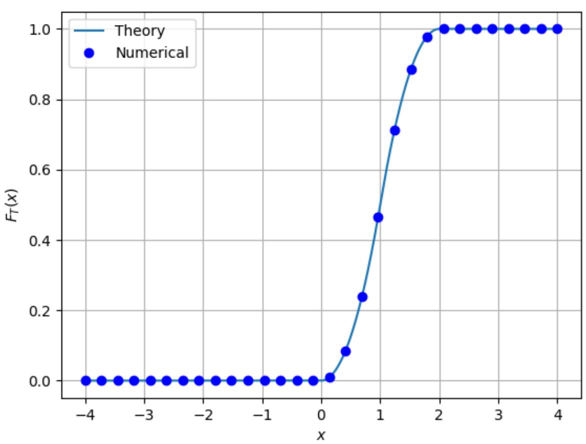

Fig. 4: CDF sim of T

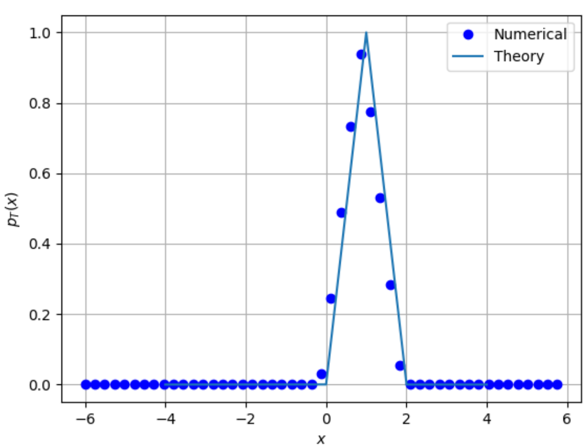

Fig. 4: PDF sim of T