Probability Distributions

Abhay Shankar K : cs21btech11001

Contents

List of Figures

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/Abhay478/GVV_things/blob/master/code/coeffs.h
wget https://github.com/Abhay478/GVV_things/blob/master/code/exrand.c
```

```
gcc exrand.c -o exrand && ./exrand
```

Or download the sample file

```
wget https://github.com/Abhay478/GVV_things/blob/master/data/uni.dat
```

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

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

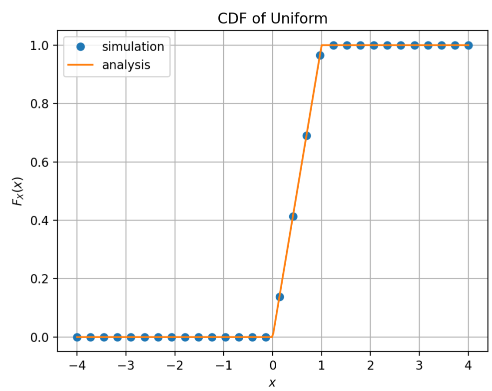
Solution:

The following code plots the figure 1

```
wget https://github.com/Abhay478/GVV_things/blob/master/code/cdf_plot_unif.py
```

```
python3 cdf_plot_unif.py
```

Fig. 1: The CDF of U



1.3 Find a theoretical expression for $P_U(x)$.

Solution: We know that

$$P_U(x) = \int_{-\infty}^{\infty} p_U(x) \tag{1.2}$$

Therefore,

$$\begin{aligned} P_U(x) &= \int_{-\infty}^{\infty} U(x)dx \\ &= \int_{-\infty}^{\infty} u(x) - u(x-1)dx \end{aligned}$$

$$\therefore P_U(x) = \begin{cases} 0 \ \forall x < 0 \\ x \ \forall x \in [0,1] \\ 1 \ \forall x > 1 \end{cases} \tag{1.3}$$

Here, $u(x)$ is the step function.

1.4 The mean of U is defined as

$$E[U] = \frac{1}{N}\sum_{i=1}^{N} U_i \tag{1.4}$$

and its variance as

$$\text{var}[U] = E[U - E[U]]^2 \tag{1.5}$$

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

Solution:

```
wget https://github.com/
    Abhay478/GVV_things/
    blob/master/code/mean-
    var.c
```

```
gcc mean-var.c -o mean-
    var && ./mean-var
```

1.5 Verify your result theoretically given that

$$E\left[U^k\right] = \int_{-\infty}^{\infty} x^k dF_U(x) \tag{1.6}$$

Solution:

The mean, $E[U]$,

$$\begin{aligned} E[U] &= \int_{-\infty}^{\infty} xU(x)dx \\ &= \int_0^1 xdx \\ &= 0.5 \end{aligned} \tag{1.7}$$

The experimental result was 0.500.

The variance, $Var[U] = E[U^2] - E^2[U]$,

$$\begin{aligned} E[U^2] &= \int_{-\infty}^{\infty} x^2U(x)dx \\ &= \int_0^1 x^2dx \\ &= \frac{1}{3} \end{aligned} \tag{1.8}$$

$$\begin{aligned} \therefore Var[U] &= \frac{1}{3} - \left(\frac{1}{2}\right)^2 \\ &= \frac{1}{12} \end{aligned} \tag{1.9}$$

The experimental result was 0.083.

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:

```
gcc exrand.c -o exrand &&
    ./exrand
```

Or downlload the sample file.

```
wget https://github.com/
    Abhay478/GVV_things/
    blob/master/data/gau.
    dat
```

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:

- All CDFs are monotonically increasing
- All CDFs are continuous over their domain
- $\lim_{n\to\infty} P_X(x) = 1$
- $\lim_{n\to-\infty} P_X(x) = 0$

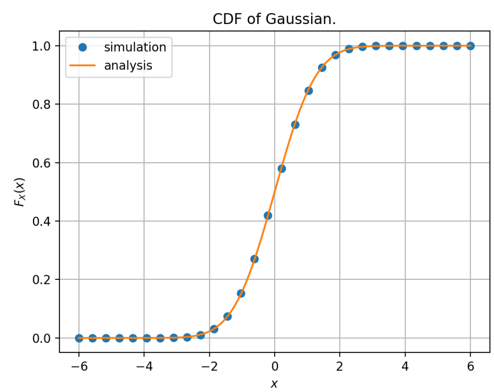
The CDF of X is plotted below.

.

```
wget https://github.com/
    Abhay478/GVV_things/
    blob/master/code/cdf_
    plot_gau.py
```

```
python3 cdf_plot_gau.py
```

Fig. 2: The CDF of 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

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

What properties does the PDF have?

Solution:

Properties of a PDF:

- The PDF adheres to unitarity, i.e. its integral over the real line is 1
- The PDF is always nonnegative
- The PDF is finite at all points

```
wget https://github.com/Abhay478/GVV_things/blob/master/code/pdf_plot_gau.py
```

```
python3 pdf_plot_gau.py
```

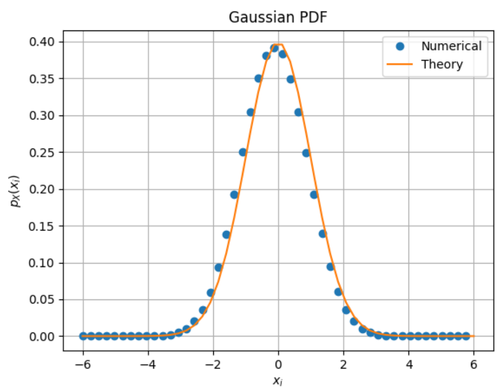


Fig. 3: The PDF of X

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

```
wget https://github.com/Abhay478/GVV_things/blob/master/code/mean-var.c
```

```
gcc mean-var.c -o mean-var && ./mean-var
```

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:

Given

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

We use the substitution $u = \frac{x^2}{2}$, whence $du = xdx$.

The mean,

$$\begin{aligned} E[X] &= \int_{-\infty}^{\infty} \frac{x}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} ue^{-u} du \\ &= 0 \end{aligned} \tag{2.5}$$

The variance, $Var[X] = E[X^2]$, since $E[X] = 0$.

$$\begin{aligned} Var[X] &= \int_{-\infty}^{\infty} \frac{x^2}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \\ &= \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} \sqrt{u} e^{-u} du \end{aligned} \tag{2.6}$$

Which is clearly $2\Gamma\left(\frac{3}{2}\right)$

$$= \frac{2}{\sqrt{\pi}} \times \frac{\sqrt{\pi}}{2} = 1 \tag{2.7}$$

3 From Uniform to Other

3.1 Generate samples of

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

and plot its CDF.

```
gcc exrand.c -o exrand &&
    ./exrand
```

Or download the sample file

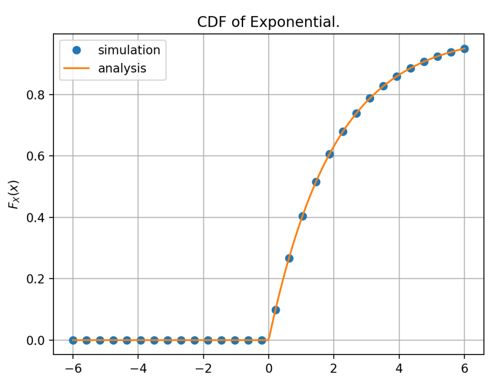
```
wget https://github.com/
    Abhay478/GVV_things/
    blob/master/data/exp.
    dat
```

The following python code plots the CDF.

```
wget https://github.com/
    Abhay478/GVV_things/
    blob/master/code/cdf_
    plot_exp.py
```

```
python3 cdf_plot_exp.py
```

Fig. 4: CDF of Exponential



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

Solution:

Let

$$v = -2ln(1-u)$$
$$\implies u = 1 - e^{-\frac{v}{2}} \tag{3.2}$$

Let

$$f(t) = -2ln(1-t) \tag{3.3}$$

$$\implies f'(t) = \frac{2}{1-t} \text{ which is increasing} \tag{3.4}$$

$$\therefore Pr(V < v) = Pr\left(U < 1 - e^{-\frac{v}{2}}\right) \tag{3.5}$$

We know that $Pr(U < f(x)) = f(x)$ if $f(x) \leq 1$

$$P_V(x) = \begin{cases} 0 \text{ if } x < 0 \\ 1 - e^{-\frac{x}{2}} \text{ if } x \geq 0 \end{cases} \tag{3.6}$$

Differentiating,

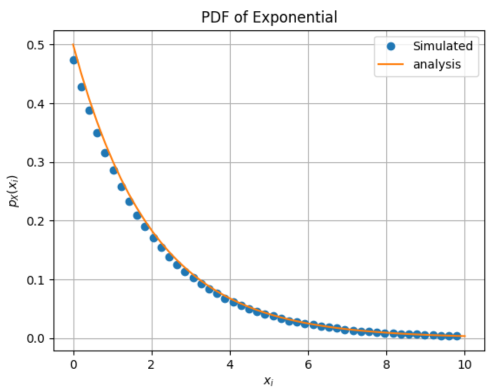
$$p_V(x) = u(x)\frac{e^{-\frac{x}{2}}}{2} \tag{3.7}$$

The PDF is clearly as exponential distribution with $\lambda = 0.5$.

```
wget https://github.com/
    Abhay478/GVV_things/
    blob/master/code/pdf_
    plot_exp.py
```

```
python3 pdf_plot_exp.py
```

Fig. 5: PDF of Exponential



4 Triangular Distribution

4.1 Generate

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

Solution:

```
gcc exrand.c -o exrand &&
    ./exrand
```

Or download the sample file

```
wget https://github.com/
   Abhay478/GVV_things/
   blob/master/data/tri.
   dat
```

4.2 Find the CDF of T.

Solution:

```
wget https://github.com/
   Abhay478/GVV_things/
   blob/master/code/cdf_
   plot_tri.py
```

```
python3 cdf_plot_tri.py
```

4.3 Find the PDF of T.

Solution:

```
wget https://github.com/
   Abhay478/GVV_things/
   blob/master/code/pdf_
   plot_tri.py
```

```
python3 pdf_plot_tri.py
```

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

Solution: The CDF:

$$P_T(t) = \int_{-\infty}^{\infty} P_{U_1}(t-u)\, p_{U_2}(u)\, du$$

$$= \int_{t-1}^{t} dx \begin{cases} 0 & \forall x < 0 \\ x & \forall x \in [0,1] \\ 1 & \forall x > 1 \end{cases} \tag{4.2}$$

$$= \begin{cases} 0 & \forall t < 0 \\ \frac{t^2}{2} & \forall t \in [0,1] \\ 2t - \frac{t^2}{2} - 1 & \forall t \in [1,2] \\ 1 & \forall t > 2 \end{cases} \tag{4.3}$$

Differentiating the equation (4.3),

$$p_T(t) = \begin{cases} 1 - |1-t| & \forall t \in [0,2] \\ 0 & otherwise \end{cases} \tag{4.4}$$

4.5 Verify your results through a plot.

Fig. 6: CDF of Triangular

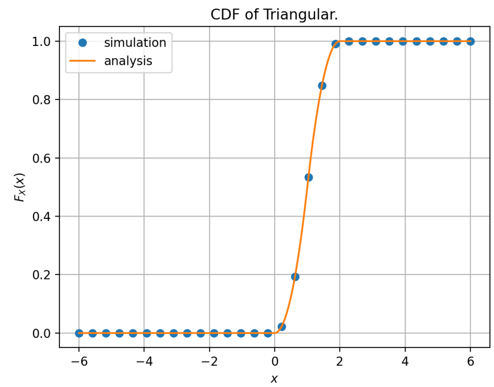
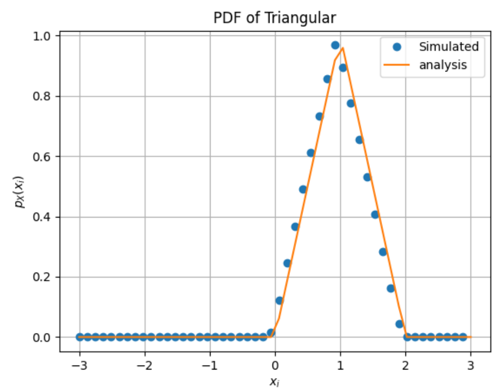


Fig. 7: PDF of Triangular



5 Maximum Likelihood

5.1 Generate equiprobable $X \in \{1, -1\}$.

```
gcc exrand.c -o exrand &&
    ./exrand
```

Or download the sample files.

```
wget https://github.com/
    Abhay478/GVV_things/
    blob/master/data/plm.
    dat
```

5.2 Generate

$$Y = AX + N, \tag{5.1}$$

where $A = 5$ dB, and $N \sim \mathcal{N}(0, 1)$.

Or download the sample files

```
wget https://github.com/
    Abhay478/GVV_things/
    blob/master/data/noi.
    dat
```

5.3 Plot Y using a scatter plot.

Solution:

```
wget https://github.com/
    Abhay478/GVV_things/
    blob/master/code/
    scatter.py
```

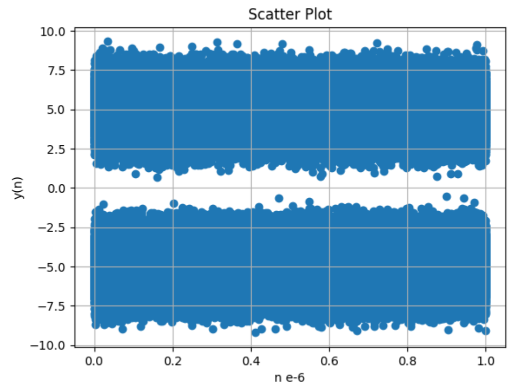
```
python3 scatter.py
```

5.4 Guess how to estimate X from Y.

Solution:

From the scatter plot, it is evident that $Y > 0$ usually correlates to $X = 1$, and $Y < 0$ to $X = -1$

Fig. 8: Scatterplot



5.5 Find

$$P_{e|0} = \Pr\left(\hat{X} = -1 \,|\, X = 1\right) \tag{5.2}$$

and

$$P_{e|1} = \Pr\left(\hat{X} = 1 \,|\, X = -1\right) \tag{5.3}$$

Solution:

The following code prints the number of samples that are incorrectly estimated. It is zero.

It also prints the theoretical value, which is an order of magnitude too small for our sample size.

```
wget https://github.com/
    Abhay478/GVV_things/
    blob/master/code/estim
    .py
```

```
python3 estim.py
```

Therefore, according to the simulation,

$$P_{e|0} = 0 \tag{5.4}$$

and

$$P_{e|1} = 0 \tag{5.5}$$

5.6 Find P_e assuming that X has equiprobable symbols.

$$\begin{aligned} P_{e|0} &= \Pr\left(\hat{X} = -1 \,|\, X = 1\right) && (5.6)\\ &= \Pr\left(Y < 0 | X = 1\right) \\ &= \Pr\left(N < -A\right) \\ &= \mathbf{Q}\left(A\right) && (5.7)\end{aligned}$$

Similarly, due to the symmetry of the gaussian, $P_{e|1} = \mathbf{Q}\left(A\right)$.

With $A = 5dB$,

$$P_{e|0} = \mathbf{Q}\left(5\right) \tag{5.8}$$

and

$$P_{e|1} = \mathbf{Q}\left(5\right) \tag{5.9}$$

where

$$\mathbf{Q}\left(5\right) = 2.866 \cdot 10^{-7} \tag{5.10}$$

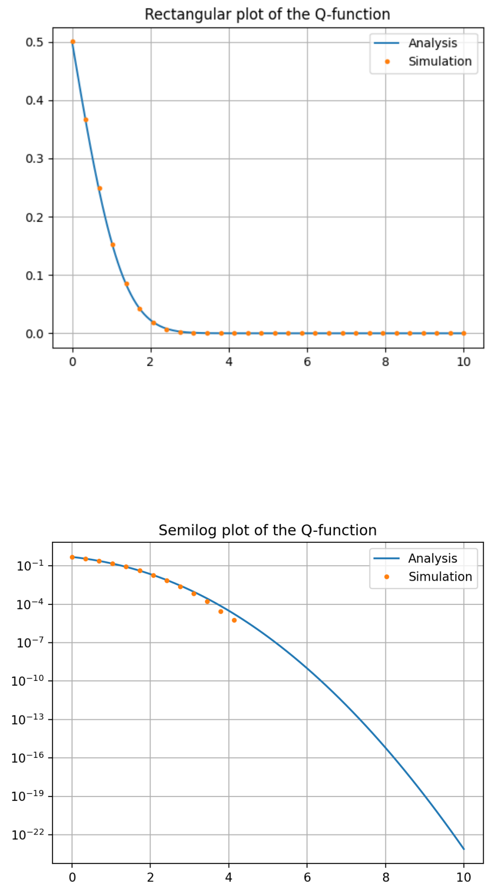
5.7 Verify by plotting the theoretical P_e with respect to A from 0 to 10 dB.

```
wget https://github.com/
    Abhay478/GVV_things/
    blob/master/code/Q_
    func.py
```

```
python3 Q_func.py
```

5.8 Now, consider a threshold δ while estimating X from Y. Find the value of δ that minimizes the theoretical P_e.

Fig. 9: Q function



Considering the estimation threshold to be δ, we have

$$\begin{aligned} P_e &= p_X\left(-1\right) P_N\left(A + \delta\right) \\ &\quad + p_X\left(1\right) P_N\left(A - \delta\right) && (5.11)\\ &= \frac{P_N\left(A + \delta\right) + P_N\left(A - \delta\right)}{2} && (5.12)\end{aligned}$$

Differentiating wrt δ and setting it to zero, we have:

$$\frac{dP_e}{d\delta} = \frac{p_N(A-\delta) - p_N(A+\delta)}{2}$$
$$\implies \delta = 0\,[\forall A \neq 0] \tag{5.13}$$

Differentiating again,

$$\frac{d^2P_e}{d\delta^2} = \frac{(A-\delta)e^{-\frac{(A-\delta)^2}{2}} + (A+\delta)e^{-\frac{(A+\delta)^2}{2}}}{2\sqrt{2\pi}} \tag{5.14}$$

which is clearly positive for the given parameters, ensuring that P_e is minimized.

5.9 Repeat the above exercise when

$$p_X(0) = p \tag{5.15}$$

From (5.11),

$$p p_N(A-\delta) = (1-p)p_N(A+\delta)$$
$$\implies e^{2A\delta} = \frac{1-p}{p}$$
$$\implies \delta = \frac{1}{2A}\ln\left(\frac{1-p}{p}\right) \tag{5.16}$$

5.10 Repeat the above exercise using the MAP criterion.

Using Bayes' theorem,

$$P(X=1\,|\,Y=y) = \frac{P(N=y-A)P_X(1)}{P_Y(y)}$$
$$= \frac{p p_N(y-A)}{p p_N(y-A) + (1-p)p_N(y+A)}$$
$$= \frac{p}{p+(1-p)e^{-2yA}} \tag{5.17}$$

Due to unitarity,

$$P(X=-1\,|\,Y=y) = \frac{1-p}{(1-p)+pe^{2yA}} \tag{5.18}$$

Hence,

If $X = 1$

$$\frac{p}{p+(1-p)\,e^{-2yA}} > \frac{1-p}{(1-p)+pe^{2yA}}$$
$$\implies \frac{(1-p)^2}{p^2} > e^{4yA}$$
$$\implies y > \frac{1}{2A}\ln\left(\frac{1-p}{p}\right) \tag{5.19}$$

If $X = -1$

$$\frac{p}{p+(1-p)\,e^{-2yA}} < \frac{1-p}{(1-p)+pe^{2yA}}$$
$$\implies \frac{(1-p)^2}{p^2} < e^{4yA}$$
$$\implies y < \frac{1}{2A}\ln\left(\frac{1-p}{p}\right) \tag{5.20}$$

6 Gaussian to Other

6.1 Let $X_1 \sim \mathcal{N}(0,1)$ and $X_2 \sim \mathcal{N}(0,1)$. Plot the CDF and PDF of

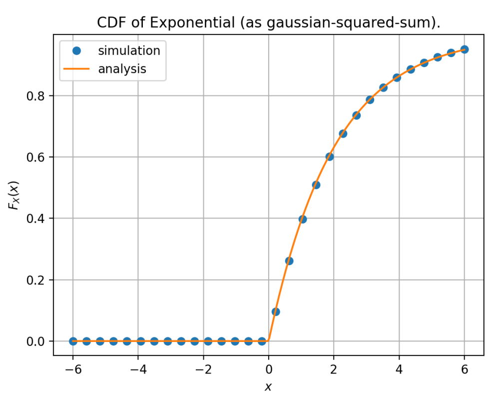
$$V = X_1^2 + X_2^2 \tag{6.1}$$

```
gcc exrand.c -o exrand &&
    ./exrand
```

Or download the sample file

```
wget https://github.com/
    Abhay478/GVV_things/
    blob/master/data/gss.
    dat
```

Fig. 10: CDF of Exponential *again*



6.2 If

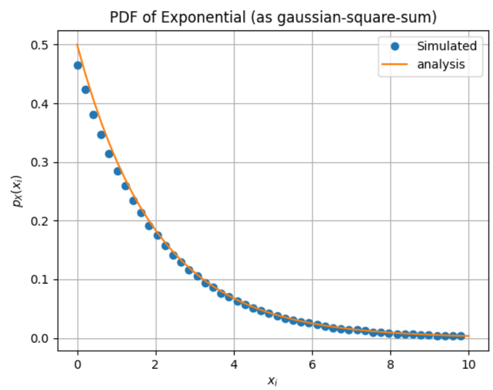
$$P_V(x) = \begin{cases} 1 - e^{-\alpha x} & x \geq 0 \\ 0 & x < 0, \end{cases} \tag{6.2}$$

find α.

Solution:

Making standard trigonometric substitutions :

Fig. 11: PDF of Exponential *again*



- $X_1 = R\cos(\Theta)$
- $X_2 = R\sin(\Theta)$

Within the domains

- $R \in \mathbb{R}^+$
- $\Theta \in [-\pi, \pi]$

Hence, the Jacobian.

$$\mathbb{J} = \begin{pmatrix} \frac{\partial X_1}{\partial R} & \frac{\partial X_2}{\partial R} \\ \\ \frac{\partial X_1}{\partial \Theta} & \frac{\partial X_2}{\partial \Theta} \end{pmatrix}$$

$$= \begin{pmatrix} \cos\Theta & \sin\Theta \\ -R\sin\Theta & R\cos\Theta \end{pmatrix} \tag{6.3}$$

$$\implies |\mathbb{J}| = R \tag{6.4}$$

Whence

$$\begin{aligned}
p_{R,\Theta}(r,\theta) &= Rp_{X_1}(x_1)\,p_{X_2}(x_2) \\
&= \frac{r}{2\pi}e^{-\frac{x_1^2+x_2^2}{2}} \\
&= \frac{r}{2\pi}e^{\frac{-r^2}{2}} && (6.5) \\
\implies p_R(r) &= \int_{-\pi}^{\pi}\frac{r}{2\pi}e^{\frac{-r^2}{2}}\,d\theta \\
&= re^{\frac{-r^2}{2}} && (6.6)
\end{aligned}$$

We now have $V = R^2$.

$$\begin{aligned}
P_V(x) &= P_R\left(\sqrt{x}\right) \\
&= \int_0^{\sqrt{x}} re^{\frac{-r^2}{2}}\,dr \\
&= 1 - e^{\frac{-x}{2}} && (6.7)
\end{aligned}$$

Thus, clearly, $\alpha = 0.5$.

6.3 Plot the CDF and PDf of

$$A = \sqrt{V} \qquad (6.8)$$

From (6.6), we obtain

$$p_A(a) = ae^{\frac{-a^2}{2}} \qquad (6.9)$$

Integrating,

$$\begin{aligned}
P_A(a) &= \int_0^a te^{\frac{-t^2}{2}}\,dt \\
&= 1 - e^{\frac{-a^2}{2}} && (6.10)
\end{aligned}$$

```
gcc exrand.c -o exrand &&
    ./exrand
```

Or get the sample file

```
wget https://github.com/
    Abhay478/GVV_things/
    blob/master/data/ral.
    dat
```

Generate the plots by downloading and executing the following

```
wget https://github.com/
    Abhay478/GVV_things/
    blob/master/code/cdf_
    plot_ral.py
wget https://github.com/
    Abhay478/GVV_things/
    blob/master/code/pdf_
    plot_ral.py
```

```
python3 cdf_plot_ral.py
python3 pdf_plot_ral.py
```

Fig. 12: CDF of Raleigh

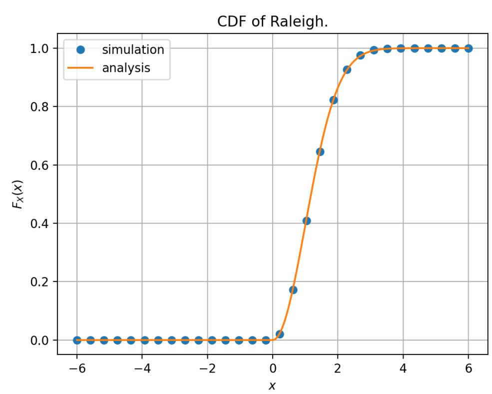
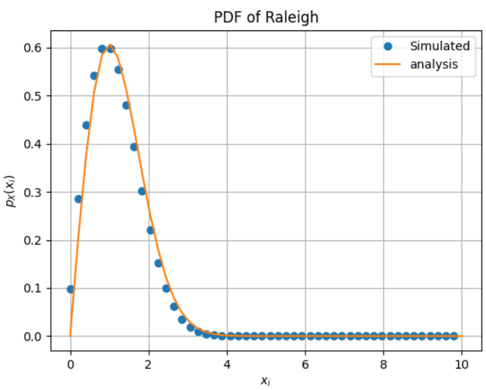


Fig. 13: PDF of Raleigh



7 Conditional Probability

7.1 Plot

$$P_e = \Pr\left(\hat{X} = -1 \,|\, X = 1\right) \tag{7.1}$$

for

$$Y = AX + N, \tag{7.2}$$

where A is Raleigh with $E\left[A^2\right] = \gamma, N \sim \mathcal{N}(0,1), X \in (-1,1)$ for $0 \leq \gamma \leq 10$ dB.

```
gcc exrand.c -o exrand &&
      ./exrand
```

Or download the sample files

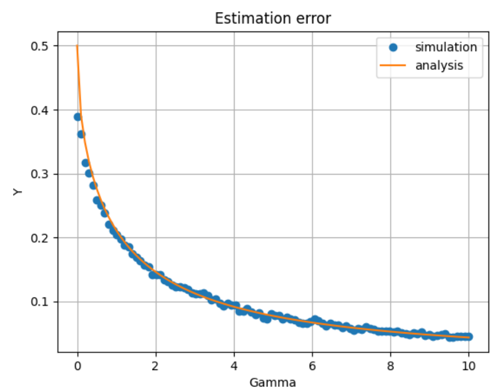
```
wget https://github.com/
    Abhay478/GVV_things/
    blob/master/data/con.
    dat
wget https://github.com/
    Abhay478/GVV_things/
    blob/master/data/con_
    log.dat
```

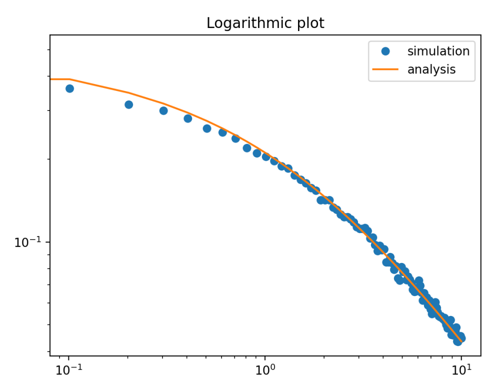
The program for plotting the graphs

```
wget https://github.com/
    Abhay478/GVV_things/
    blob/master/code/
    conditional.py
```

```
python3 conditional.py
```

Fig. 14: Conditional probability





7.2 Assuming that N is a constant, find an expression for P_e. Call this $P_e(N)$

Solution:

We rewrite the equation (5.6) as

$$\begin{aligned} P_e &= P(A < -N) \\ &= P_A(-N) \\ &= u(-N)\left(1 - e^{-\frac{N^2}{\gamma}}\right) \end{aligned} \tag{7.3}$$

7.3 For a function g,

$$E[g(X)] = \int_{-\infty}^{\infty} g(x) p_X(x)\, dx \tag{7.4}$$

Find $P_e = E[P_e(N)]$.

Solution:

From (7.3),

$$\begin{aligned} E[P_e] &= \int_0^{\infty} P_A(x)\, p_N(x)\, dx \\ &= \int_0^{\infty} \left(1 - e^{-\frac{x^2}{\gamma}}\right)\left(\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\right) \\ &= 0.5 - \frac{1}{\sqrt{2\pi}} \int_0^{\infty} e^{-\frac{x^2(\gamma+2)}{2\gamma}} \\ &= 0.5\left(1 - \sqrt{\frac{\gamma}{\gamma+2}}\right) \end{aligned} \tag{7.5}$$

7.4 Plot P_e in problems 7.1 and 7.3 on the same graph w.r.t γ. Comment.

Solution: The required graphs have been plotted (figure 14).

8 Two Dimensions

Let

$$\mathbf{y} = A\mathbf{x} + \mathbf{n}, \tag{8.1}$$

where

$$x \in (\mathbf{s}_0, \mathbf{s}_1), \mathbf{s}_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \mathbf{s}_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{8.2}$$

$$\mathbf{n} = \begin{pmatrix} n_1 \\ n_2 \end{pmatrix}, n_1, n_2 \sim \mathcal{N}(0, 1). \tag{8.3}$$

8.1 Plot

$$\mathbf{y}\,|\,\mathbf{s}_0 \text{ and } \mathbf{y}\,|\,\mathbf{s}_1 \tag{8.4}$$

on the same graph using a scatter plot.

Solution: Here, we have taken A = 5.

Download the sample file

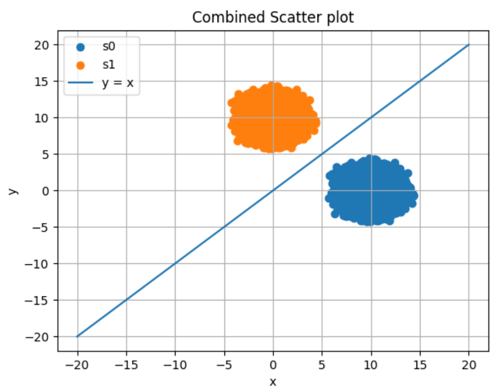
```
wget https://github.com/
    Abhay478/GVV_things/
    blob/master/data/gau.
    dat
```

Program to plot the graph

```
wget https://github.com/
    Abhay478/GVV_things/
    blob/master/code/2d.py
```

```
python3 2d.py
```

Fig. 15: Combined plot



8.2 For the above problem, find a decision rule for detecting the symbols $\mathbf{s}_0$ and $\mathbf{s}_1$.

Solution:

$$\hat{x} = \begin{cases} y > x \implies s_1 \\ y < x \implies s_0 \end{cases} \tag{8.5}$$

8.3 Plot

$$P_e = \Pr\left(\hat{\mathbf{x}} = \mathbf{s}_1 \,|\, \mathbf{x} = \mathbf{s}_0\right) \tag{8.6}$$

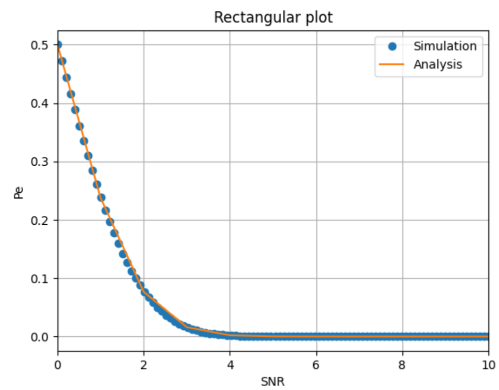
with respect to the SNR from 0 to 10 dB.

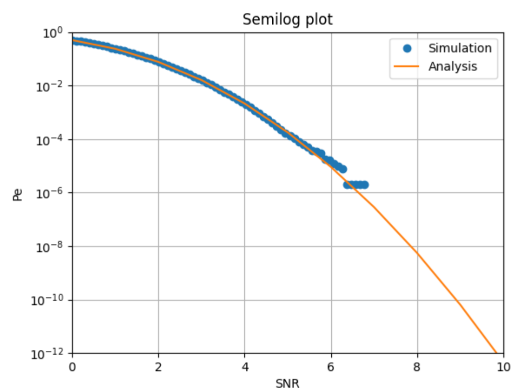
Program to plot the graph

```
wget https://github.com/
    Abhay478/GVV_things/
    blob/master/code/2d_
    estim.py
```

```
python3 2d_estim.py
```

Fig. 16: Estimation error





8.4 Obtain an expression for P_e. Verify this by comparing the theory and simulation plots on the same graph.

Solution:

$$\begin{aligned} P_e &= \Pr\left(\hat{\mathbf{x}} = \mathbf{s}_1 \,|\, \mathbf{x} = \mathbf{s}_0\right) \\ \implies P_e &= Pr(n_2 > n_1 + A) \\ &= \mathbf{Q}\left(\frac{A}{\sqrt{2}}\right) \end{aligned} \tag{8.7}$$

Plotted in figure 16