

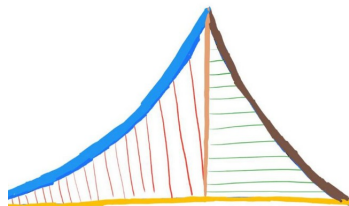
---

# GATE PROBABILITY

## Through Simulations

---

G. V. V. Sharma



Copyright ©2022 by G. V. V. Sharma.

<https://creativecommons.org/licenses/by-sa/3.0/>

and

<https://www.gnu.org/licenses/fdl-1.3.en.html>

# Contents

Introduction	iii
<b>1 Axioms</b>	<b>1</b>
<b>2 Distributions</b>	<b>3</b>
<b>3 Conditional Probability</b>	<b>87</b>
<b>4 Random Variable</b>	<b>93</b>
<b>5 Moments</b>	<b>97</b>
<b>6 Random Algebra</b>	<b>111</b>
<b>7 Hypothesis Testing</b>	<b>113</b>
<b>8 Bivariate Random Variables</b>	<b>129</b>
<b>9 Random Processes</b>	<b>131</b>
<b>10 Convergence</b>	<b>137</b>
<b>11 Information Theory</b>	<b>171</b>
<b>12 Markov chain</b>	<b>181</b>
<b>13 Estimation</b>	<b>185</b>

14 Multivariate Gaussian 227

14.1 Simulation Steps . . . . . 229



# Introduction

This book solves probability problems in GATE question papers.



# Chapter 1

## Axioms

1.1 Fabry disease in humans is a X-linked disease. The probability (in percentage) for a phenotypically normal father and a carrier mother to have a son with fabry disease is (GATE BT 2023)

**Solution:** before going into question let me clear out few things on chromosome distribution in general

parent	chromosome
father	XY
mother	XX

Table 1.1: without disease

now chromosome distribution when fabry disease is effected and there is a normal father and carrier mother

parent	chromosome
father	XY
mother	$X\bar{X}$

Table 1.2: with disease

as given in the question a son is to be born with fabry disease

let us denote a random variable  $Z$  as the event that the son is born with fabry disease.



offspring	chromosome	
son	with fabry disease	$\bar{X}Y$
	without fabry disease	$XY$

Table 1.3: chromosome of son

$$\Pr(Z) = \frac{1}{2}$$

hence the percentage of the probability is 50%

- 1.2 The probabilities of occurrences of two independent events A and B are 0.5 and 0.8 respectively. What is the probability of occurrence of at least A or B (rounded off to 1 decimal place)? (GATE CE 2023)

Given,

$$\Pr(A) = 0.5 \text{ and } \Pr(B) = 0.8$$

Probability of occurrence of at least A or B is given by  $\Pr(A + B)$

Since A and B are independent events, we can say that:

$$\Pr(A + B) = \Pr(A) + \Pr(B) - \Pr(AB) \quad (1.1)$$

$$= \Pr(A) + \Pr(B) - \Pr(A) \times \Pr(B) \quad (1.2)$$

$$= 0.5 + 0.8 - 0.5 \times 0.8 \quad (1.3)$$

$$= 0.9 \quad (1.4)$$

## Chapter 2

# Distributions

2.1 Let  $\phi(\cdot)$  denote the cumulative distribution function of a standard normal random variable. If the random variable  $X$  has the cumulative distribution function

$$F(x) = \begin{cases} \phi(x), & x < -1 \\ \phi(x+1), & x \geq -1 \end{cases} \quad (2.1)$$

then which one of the following statements is true?

(a)  $P(X \leq -1) = \frac{1}{2}$

(b)  $P(X = -1) = \frac{1}{2}$

(c)  $P(X < -1) = \frac{1}{2}$

(d)  $P(X \leq 0) = \frac{1}{2}$

(GATE ST 2023)

**Solution: Gaussian**

Q function is defined

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-\frac{u^2}{2}} du \quad (2.2)$$

From the question and (2.2);

$$F_X(x) = \begin{cases} Q(-x), & x < -1 \\ 1 - Q(x+1), & x \geq -1 \end{cases} \quad (2.3)$$

From (2.3);

(a)

$$\Pr(X \leq -1) = F_X(-1) = 1 - Q(0) \quad (2.4)$$

$$= 0.5 \quad (2.5)$$

So Option A i.e.,  $P(X < -1) = \frac{1}{2}$  is correct

(b) The pdf of X can be defined in terms of cdf as

$$\Pr(X = b) = F_X(b) - \lim_{x \rightarrow b^-} F_X(x) \quad (2.6)$$

From (2.6);

$$\Pr(X = -1) = F_X(-1) - \lim_{x \rightarrow -1^-} F_X(x) \quad (2.7)$$

$$= 1 - Q(0) - Q(-(-1)) \quad (2.8)$$

$$= 0.341 \quad (2.9)$$

So Option B i.e.,  $P(X = -1) = \frac{1}{2}$  is incorrect

(c)

$$\Pr(X < -1) = \lim_{x \rightarrow -1^-} F_X(x) = F_X(-1) \quad (2.10)$$

$$= Q(-(-1)) \quad (2.11)$$

$$= 0.159 \quad (2.12)$$

So Option C i.e.,  $P(X < -1) = \frac{1}{2}$  is incorrect

(d)

$$\Pr(X \leq 0) = F_X(0) = 1 - Q(1) \quad (2.13)$$

$$= 0.8413 \quad (2.14)$$

So Option D i.e.,  $P(X \leq 0) = \frac{1}{2}$  is incorrect

Gaussian CDF plot of  $X$  is given in

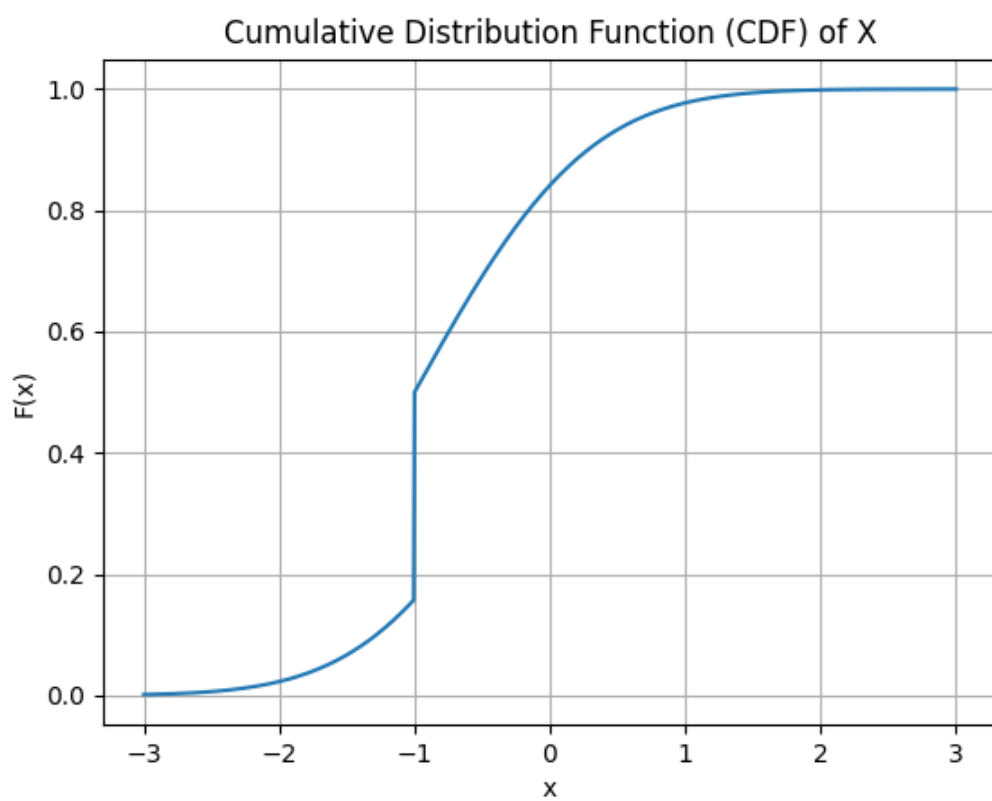


Figure 2.1:

### Box-Muller

**STATEMENT:** Suppose  $U_1$  and  $U_2$  are independent samples chosen from the uniform distribution on the unit interval  $(0, 1)$ . Let

$$Z_0 = R \cos(\Theta) = \sqrt{-2 \ln U_1} \cos(2\pi U_2) \quad (2.15)$$

and

$$Z_1 = R \sin(\Theta) = \sqrt{-2 \ln U_1} \sin(2\pi U_2) \quad (2.16)$$

Then  $Z_0$  and  $Z_1$  are independent random variables with a standard normal distribution. So, now we will generate  $Z_0$  using C Code as given below

**PROOF:** Let  $X$  and  $Y$  be independent standard normal variables

$$X, Y \sim \mathcal{N}(0, 1) \quad \text{and} \quad X \perp Y \quad (2.17)$$

The joint pdf of  $X$  and  $Y$  is given by

$$f_{XY}(X, Y) = f(x)f(y) \quad (2.18)$$

$$= \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} \quad (2.19)$$

$$= \frac{1}{2\pi} e^{-\frac{(x^2+y^2)}{2}} \quad (2.20)$$

The relationship between Cartesian coordinates  $(x, y)$  and polar coordinates  $(r, \theta)$  is

as follows

$$x = r \cos \theta \quad (2.21)$$

$$y = r \sin \theta \quad (2.22)$$

Change  $f_{XY}(x, y)$  to polar coordinates :

$$f_{XY}(x, y)dx dy = f_{R\theta}(r, \theta)dr d\theta \quad (2.23)$$

i.e.,

$$f_{R\theta}(r, \theta) = f_{XY}(x, y) \frac{dx dy}{dr d\theta} = f_{XY}(x, y) \left| \frac{\partial(x, y)}{\partial(r, \theta)} \right| \quad (2.24)$$

where J is the Jacobian

$$J = \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{vmatrix} \quad (2.25)$$

$$= \begin{vmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{vmatrix} = r \quad (2.26)$$

For  $r \geq 0$  and  $\theta \in [0, 2\pi)$ , we have

$$f_{R\theta}(r, \theta)dr d\theta = \frac{1}{2\pi} e^{\frac{-(r^2)}{2}} r dr d\theta \quad (2.27)$$

Now we change the variable from  $(r, \theta)$  to  $(r^2, \theta)$ ;

$$r dr = \frac{1}{2} dr^2 \quad (2.28)$$

Now it can be written as,

$$f_{R\theta}(r, \theta) dr d\theta = f_{R^2\theta}(r^2, \theta) dr^2 d\theta \quad (2.29)$$

$$= \frac{1}{2\pi} e^{-\frac{(r^2)}{2}} \frac{1}{2} dr^2 d\theta \quad (2.30)$$

$$= \left( \frac{1}{2} e^{-\frac{(r^2)}{2}} dr^2 \right) \left( \frac{1}{2} d\theta \right) \quad (2.31)$$

$$= f_{R^2}(r^2) dr^2 f_{\theta}(\theta) d\theta \quad (2.32)$$

from the above equation, we can say that

$$R^2 \perp \theta \text{ (i.e., } R^2 \text{ and } \theta \text{ are independent)} \quad (2.33)$$

Generate,  $\theta \sim Unif(0, 2\pi)$  Generate,  $V \sim Exp(\lambda = \frac{1}{2})$  (i.e.,  $V = R^2$ ) and compute

$$R = \sqrt{V} \quad (2.34)$$

Compute

$$X = R \cos \theta \quad (2.35)$$

$$Y = R \sin \theta \quad (2.36)$$

$$(2.37)$$

where,  $X, Y$  are i.i.d in  $\sim \mathcal{N}(0, 1)$  Then

$$\theta = 2\pi U_1 \quad (2.38)$$

$$V = -2\log(U_2) \quad (2.39)$$



We can prove  $V = -2\log(U_2)$  using the c.d.f definition of the exponential distribution. Let  $X$  be a random variable following an exponential distribution with rate parameter  $\lambda = \frac{1}{2}$ , denoted as  $X \sim \text{Exp}(\frac{1}{2})$ . The cumulative distribution function (CDF) of the exponential distribution is given by:

$$F(x) = 1 - e^{-\frac{x}{2}} \quad (2.40)$$

Now, suppose we have a random variable  $U$  following a uniform distribution in the interval  $[0, 1]$ , denoted as  $U \sim \text{U}(0, 1)$ . The CDF of the uniform distribution is simply:

$$F_U(u) = u, \text{ for } 0 \leq u \leq 1 \quad (2.41)$$

We can use the probability integral transform to express the exponential random variable  $X$  in terms of the uniform random variable  $U$ :

$$F(x) = F_U(u) \quad (2.42)$$

$$1 - e^{-\frac{x}{2}} = u \quad (2.43)$$

Now, solve for  $x$ :

$$e^{-\frac{x}{2}} = 1 - u \quad (2.44)$$

$$-\frac{x}{2} = \ln(1 - u) \quad (2.45)$$

$$x = -2\ln(1 - u) \quad (2.46)$$

So,  $X \sim -2\ln(1 - U)$  for  $U \sim \text{U}(0, 1)$ . This expression represents the exponential random variable  $X$  in terms of a uniform random variable  $U$ .

C and Python codes are given below :

```
#include <stdio.h>
#include <stdlib.h>
#include <math.h>
#include <time.h>

double sn() {
    double u1, u2;
    while (1) {
        u1 = ((double)rand() / RAND_MAX);
        if (u1 > 0 && u1 < 1)
            break;}
    while (1) {
        u2 = ((double)rand() / RAND_MAX);
        if (u2 > 0 && u2 < 1)
            break;}
    double z1 = sqrt(-2 * log(u1)) * cos(2 * M_PI * u2);
    return z1;
}

int main() {
    // Seed the random number generator
    srand(time(0));

    int numSamples = 100000; // You can change this to the desired number of
```

```

        samples

FILE *file = fopen("uni.dat", "w");

if (file == NULL) {
    printf("Unable to open file for writing.\n");
    return 1;
}

for (int i = 0; i < numSamples; i++) {
    double sample = sn();
    fprintf(file, "%lf\n", sample);
}

fclose(file);

return 0;
}

```

```

import numpy as np
import matplotlib.pyplot as plt

# Read data from the "uni.dat" file
x = np.genfromtxt("uni.dat")

# Create a histogram plot of the PDF

```

```
plt.hist(x, bins=50, density=True, alpha=0.6, color='b', label='PDF_(Histogram)')
plt.xlabel('x_(Standard_Normal_Random_Variable)')
plt.ylabel('Frequency')
plt.title('PDF_of_Standard_Normal_Distribution_(Histogram)')
plt.savefig('/home/sayyam/KHP/figs/figure1.png')
plt.show()
```

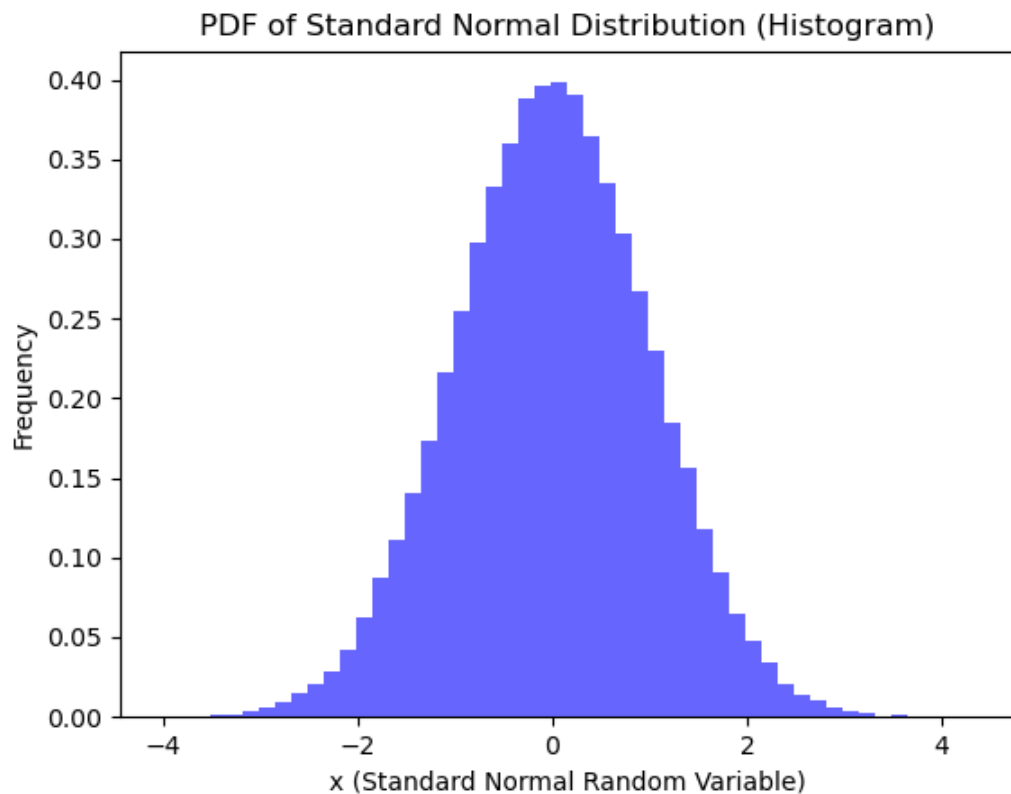


Figure 2.2: Histogram plot of density of  $Z_0$

2.2 Let  $X$  be a random variable with the probability density function  $f(x)$  such that

$$f(x) = \begin{cases} \frac{1}{2\sqrt{3}}, & -\sqrt{3} \leq x \leq \sqrt{3} \\ 0, & \text{otherwise} \end{cases} \quad (2.47)$$

Then the variance of  $X$  is?

(GATE XH-C1 2023)

**Solution:**

The mean of  $X$

$$\mu_X = \int_{-\infty}^{\infty} x f(x) dx \quad (2.48)$$

As the integrand is odd

$$\implies \mu_X = 0 \quad (2.49)$$

The variance of  $X$  is:

$$\sigma_X^2 = \mathbb{E}(X - \mu_X)^2 \quad (2.50)$$

From (2.49)

$$\implies \sigma_X^2 = \mathbb{E}(X^2) \quad (2.51)$$

$$= \frac{1}{2\sqrt{3}} \int_{-\sqrt{3}}^{\sqrt{3}} x^2 dx \quad (2.52)$$

$$= 1 \quad (2.53)$$

2.3 Two defective bulbs are present in a set of five bulbs. To remove the two defective

bulbs, the bulbs are chosen randomly one by one and tested. If  $X$  denotes the minimum number of bulbs that must be tested to find out the two defective bulbs, then  $\Pr(X = 3)$  (rounded off to two decimal places) equals

(GATE ST 2023)

**Solution:**

RV	Values	Description
A	0	1 <sup>st</sup> Bulb defective
	1	1 <sup>st</sup> Bulb non-defective
B	0	2 <sup>nd</sup> Bulb defective
	1	2 <sup>nd</sup> Bulb non-defective
C	0	3 <sup>rd</sup> Bulb defective
	1	3 <sup>rd</sup> Bulb non-defective

Table 2.1: Random variable declaration.

Here, the word "minimum" does not signify anything. Therefore we get

$$p_X(2) = p_{AB}(0, 0) \quad (2.54)$$

$$= \frac{2}{5} \times \frac{1}{4} \quad (2.55)$$

$$= \frac{1}{10} \quad (2.56)$$

$$p_X(3) = p_{ABC}(1, 0, 0) + p_{ABC}(0, 1, 0) + p_{ABC}(1, 1, 1) \quad (2.57)$$

$$= \frac{3}{5} \times \frac{2}{4} \times \frac{1}{3} + \frac{2}{5} \times \frac{3}{4} \times \frac{1}{3} + \frac{3}{5} \times \frac{2}{4} \times \frac{1}{3} \quad (2.58)$$

$$= \frac{3}{10} \quad (2.59)$$

$$p_X(4) = p_{ABC}(0, 1, 1) + p_{ABC}(1, 0, 1) + p_{ABC}(1, 1, 0) \quad (2.60)$$

$$= \frac{2}{5} \times \frac{3}{4} \times \frac{2}{3} + \frac{3}{5} \times \frac{2}{4} \times \frac{2}{3} + \frac{3}{5} \times \frac{2}{4} \times \frac{2}{3} \quad (2.61)$$

$$= \frac{6}{10} \quad (2.62)$$

Hence, The pmf of X is

$$p_X(k) = \begin{cases} 0 & k = 1 \\ \frac{1}{10} & k = 2 \\ \frac{3}{10} & k = 3 \\ \frac{6}{10} & k = 4 \\ 1 & k = 5 \end{cases} \quad (2.63)$$

2.4 Let  $X$  be a random variable with cumulative distribution function

$$F_X(x) = \begin{cases} 0 & \text{if } x < -1 \\ \frac{1}{4}(x+1) & \text{if } -1 \leq x < 0 \\ \frac{1}{4}(x+3) & \text{if } 0 \leq x < 1 \\ 1 & \text{if } x \geq 1 \end{cases} \quad (2.64)$$

Which one of the following statements is true?

(A)

$$\lim_{n \rightarrow \infty} \Pr \left( -\frac{1}{2} + \frac{1}{n} < X < -\frac{1}{n} \right) = \frac{5}{8} \quad (2.65)$$

(B)

$$\lim_{n \rightarrow \infty} \Pr \left( -\frac{1}{2} - \frac{1}{n} < X < \frac{1}{n} \right) = \frac{5}{8} \quad (2.66)$$

(C)

$$\lim_{n \rightarrow \infty} \Pr \left( X = \frac{1}{n} \right) = \frac{1}{2} \quad (2.67)$$

(D)

$$\Pr (X = 0) = \frac{1}{3} \quad (2.68)$$

(GATE ST 2023)

**Solution:**

$$f_X(x) = \begin{cases} 0 & \text{if } x < -1 \\ \frac{1}{4} & \text{if } -1 \leq x < 0 \\ \frac{1}{4} + \frac{1}{2}\delta(x) & \text{if } 0 \leq x < 1 \\ 0 & \text{if } x \geq 1 \end{cases} \quad (2.69)$$

(A)

$$\begin{aligned} \lim_{n \rightarrow \infty} \Pr \left( -\frac{1}{2} + \frac{1}{n} < X < -\frac{1}{n} \right) \\ = \lim_{n \rightarrow \infty} F_X \left( -\frac{1}{n} \right) - \lim_{n \rightarrow \infty} F_X \left( -\frac{1}{2} + \frac{1}{n} \right) \end{aligned} \quad (2.70)$$

$$= \lim_{n \rightarrow \infty} F_X \left( -\frac{1}{n} \right) - \lim_{n \rightarrow \infty} F_X \left( -\frac{1}{2} + \frac{1}{n} \right) \quad (2.71)$$

$$= \lim_{n \rightarrow \infty} \frac{1}{4} \left( -\frac{1}{n} + 1 \right) - \lim_{n \rightarrow \infty} \frac{1}{4} \left( -\frac{1}{2} + \frac{1}{n} + 1 \right) \quad (2.72)$$

$$= \frac{1}{8} \quad (2.73)$$



$\therefore (A)$  is not true.

(B)

$$\begin{aligned} \lim_{n \rightarrow \infty} \Pr \left( -\frac{1}{2} - \frac{1}{n} < X < \frac{1}{n} \right) \\ = \lim_{n \rightarrow \infty} F_X \left( \frac{1}{n} \right) - \lim_{n \rightarrow \infty} F_X \left( -\frac{1}{2} - \frac{1}{n} \right) \end{aligned} \quad (2.74)$$

$$= \lim_{n \rightarrow \infty} F_X \left( \frac{1}{n} \right) - \lim_{n \rightarrow \infty} F_X \left( -\frac{1}{2} - \frac{1}{n} \right) \quad (2.75)$$

$$= \lim_{n \rightarrow \infty} \frac{1}{4} \left( \frac{1}{n} + 3 \right) - \lim_{n \rightarrow \infty} \frac{1}{4} \left( -\frac{1}{2} - \frac{1}{n} + 1 \right) \quad (2.76)$$

$$= \frac{5}{8} \quad (2.77)$$

$\therefore (B)$  is true.

(C) From (2.69)

$$\lim_{n \rightarrow \infty} \Pr \left( X = \frac{1}{n} \right) = 0 \quad (2.78)$$

$\therefore (C)$  is not true.

(D) From (2.69)

$$\Pr(X = 0) = \frac{1}{2} \quad (2.79)$$

$\therefore (D)$  is not true.

Steps for the simulation of r.v  $X$ :

(a) Identify the point of discontinuity (0 here).

- (b) Define the simulation size for the simulation data set (`num_sim`).
- (c) Define the functions of CDF and PDF of  $X$ .
- (d) Find  $\Pr(X = 0)$  from the PDF of  $X$ .
- (e) For this simulation, the remaining numbers in  $[-1, 1)$  have probability of  $1 - \Pr(X = 0)$ .
- (f) Generate random sample in  $[-1, 1) - \{0\}$  of the size  $= \text{num\_sim} \times (1 - \Pr(X = 0))$ .
- (g) Generate sample containing only zeros of the size  $= \text{num\_sim} \times \Pr(X = 0)$ .
- (h) Combine all the generated samples to make a single sample and we generate the required r.v  $X$ .

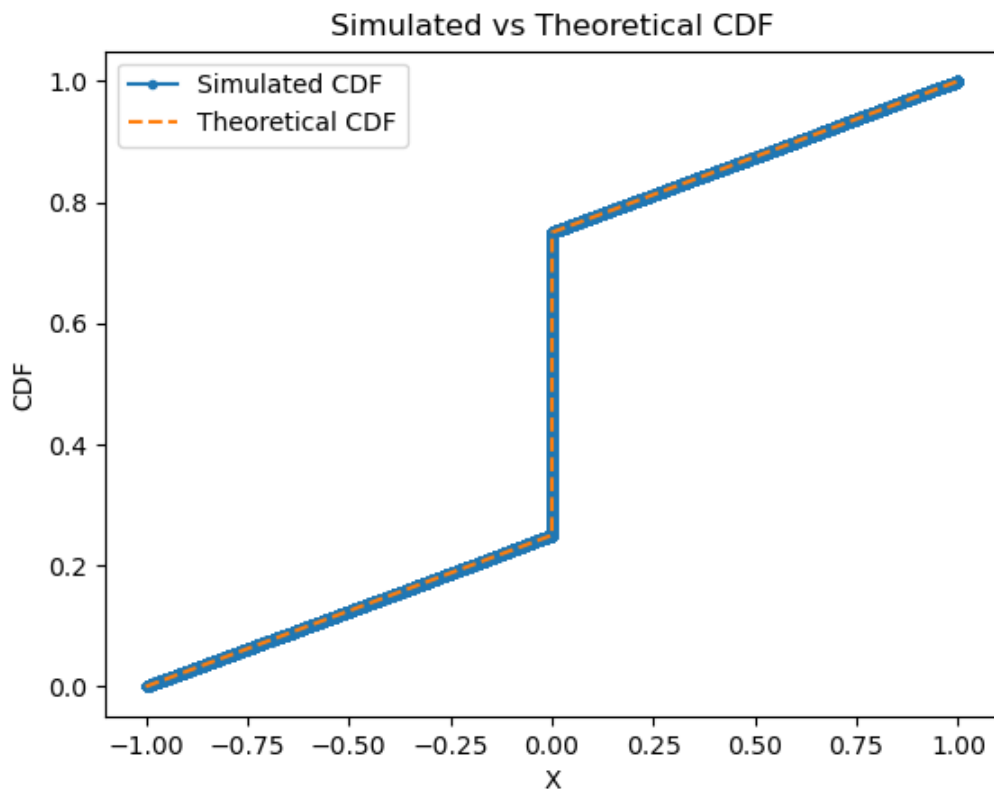


Figure 2.3: CDF of X-(simulation vs actual)

2.5 Three unbiased coins were tossed. Provided that at least two outcomes are tails, the probability of having all three outcomes as tails is  
(GATE PI 2023)

**Solution:**

Parameter	value	description
$X_i$	1	first coin
	2	second coin
	3	third coin
$n$	3	number of coins
$p, q$	$\frac{1}{2}$	toss result in heads/tails
$Y$	$\sum_{i=0}^3 X_i$	three coins

Table 2.2: Definition of  $Y$  and parameters.

$$\Pr(Y = 3|Y \geq 2) = \frac{\Pr(Y \geq 2, Y = 3)}{\Pr(Y \geq 2)} \quad (2.80)$$

$$= \frac{\Pr(Y = 3)}{\Pr(Y \geq 2)} \quad (2.81)$$

$$= \frac{p_Y(3)}{1 - F_Y(1)} \quad (2.82)$$

$$p_Y(k) = {}^nC_k p^k q^{n-k} \quad (2.83)$$

$$= {}^3C_k \left(\frac{1}{2}\right)^k \left(\frac{1}{2}\right)^{3-k} \quad (2.84)$$

$$\Rightarrow p_Y(k) = \begin{cases} \frac{{}^3C_k}{8}; k = \{0, 1, 2, 3\} \\ 0; otherwise \end{cases} \quad (2.85)$$

$$F_Y(k) = \Pr(Y \leq k) \quad (2.86)$$

$$= \sum_{k=0}^k p_Y(k) \quad (2.87)$$

$$\Rightarrow F_Y(k) = \begin{cases} 0; k < 0 \\ \sum_{k=0}^k \frac{{}^3C_k}{8}; k = \{0, 1, 2, 3\} \\ 1; k > 3 \end{cases} \quad (2.88)$$

$$\Rightarrow \Pr(Y = 3|Y \geq 2) = \frac{\left(\frac{1}{8}\right)}{\left(\frac{1}{2}\right)} \quad (2.89)$$

$$= \frac{1}{4} \quad (2.90)$$

∴ The probability of having all three outcomes as tails is 0.25.

2.6 Let  $X$  be a random variable with probability density function

$$p_X(x) = \begin{cases} e^{-x} & \text{if } x \geq 0 \\ 0 & \text{otherwise} \end{cases} \quad (2.91)$$

For  $a < b$ , if  $U(a, b)$  denotes the uniform distribution over the interval  $(a, b)$ , then which of the following statements is/are true?

- (A)  $e^{-X}$  follows  $U(-1, 0)$  distribution
- (B)  $1 - e^{-X}$  follows  $U(0, 2)$  distribution
- (C)  $2e^{-X} - 1$  follows  $U(-1, 1)$  distribution
- (D) The probability mass function of  $Y = [X]$  is  $\Pr(Y = k) = e^{-k}(1 - e^{-1})$  for  $k = 0, 1, 2, \dots$ , where  $[X]$  denotes the largest integer not exceeding  $x$

(GATE ST 2023)

**Solution:** Let  $Y \sim U(a, b)$ , then

$$p_Y(y) = \begin{cases} \frac{1}{b-a} & a < y < b \\ 0 & \text{otherwise} \end{cases} \quad (2.92)$$

and for  $a < y < b$

$$F_Y(y) = \Pr(Y \leq y) \quad (2.93)$$

$$= \int_a^y \frac{1}{b-a} dy \quad (2.94)$$

$$= \frac{y-a}{b-a} \quad (2.95)$$

Similarly, for  $x \geq 0$

$$F_X(x) = \Pr(X \leq x) \quad (2.96)$$

$$= \int_0^x e^{-x} dx \quad (2.97)$$

$$= 1 - e^{-x} \quad (2.98)$$

$$(A) \ Y = e^{-X} = U(a, b)$$

for  $a < y < b$

$$F_Y(y) = \Pr(e^{-X} \leq y) \quad (2.99)$$

$$= \Pr(X \geq -\ln y) \quad (2.100)$$

$$= 1 - F_X(-\ln y) \quad (2.101)$$

$$= 1 - (1 - y) \quad (2.102)$$

$$= y \quad (2.103)$$

Comparing this with CDF of Uniform distribution, we obtain

$$a = 0, b = 1 \quad (2.104)$$

$$\therefore Y \sim U(0, 1) \quad (2.105)$$

$$(B) \ Y = 1 - e^{-X} = U(a, b)$$

for  $a < y < b$

$$F_Y(y) = \Pr(1 - e^{-X} \leq y) \quad (2.106)$$

$$= \Pr(e^{-X} \geq 1 - y) \quad (2.107)$$

$$= \Pr(X \leq -\ln(1 - y)) \quad (2.108)$$

$$= F_X(-\ln(1 - y)) \quad (2.109)$$

$$= 1 - (1 - y) \quad (2.110)$$

$$= y \quad (2.111)$$

$$\implies Y \sim U(0, 1) \quad (2.112)$$

(C)  $Y = 2e^{-X} - 1 = U(a, b)$

for  $a < y < b$

$$F_Y(y) = \Pr(2e^{-X} - 1 \leq y) \quad (2.113)$$

$$= \Pr\left(X \geq -\ln\left(\frac{y+1}{2}\right)\right) \quad (2.114)$$

$$= 1 - F_X\left(-\ln\left(\frac{y+1}{2}\right)\right) = 1 - \left(1 - \frac{y+1}{2}\right) \quad (2.115)$$

$$= \frac{y+1}{2} \quad (2.116)$$

Comparing this with CDF of Uniform distribution, we obtain

$$a = -1, b = 1 \quad (2.117)$$

$$\therefore Y \sim U(-1, 1) \quad (2.118)$$

(D)  $Y = [X]$

$$\Pr(Y = k) = \Pr([X] = k) \quad (2.119)$$

$$= \Pr(k \leq X < k + 1) \quad (2.120)$$

$$= \int_k^{k+1} e^{-x} dx \quad (2.121)$$

$$= e^{-k} (1 - e^{-1}) \text{ for } k=0,1,2.. \quad (2.122)$$

(E) Generation of Random Variable  $X$  in C language

(i) `rand()` / `(double)RAND_MAX`:

This generates a random variable between 0 and `RAND_MAX` and divides it by `RAND_MAX` to obtain a uniform distribution between 0 and 1.

(ii) `-log(rand()) / (double)RAND_MAX` :

This transforms the uniform distribution between 0 and 1 into an exponential distribution by making the values vary from 0 to  $\infty$ .

(iii) Alternatively the Uniform distribution can be converted into Gaussian distribution using the Central Limit Theorem.

(iv) Gaussian is then converted into chi-square distribution with degree of freedom 2 which is similar to an exponential distribution.



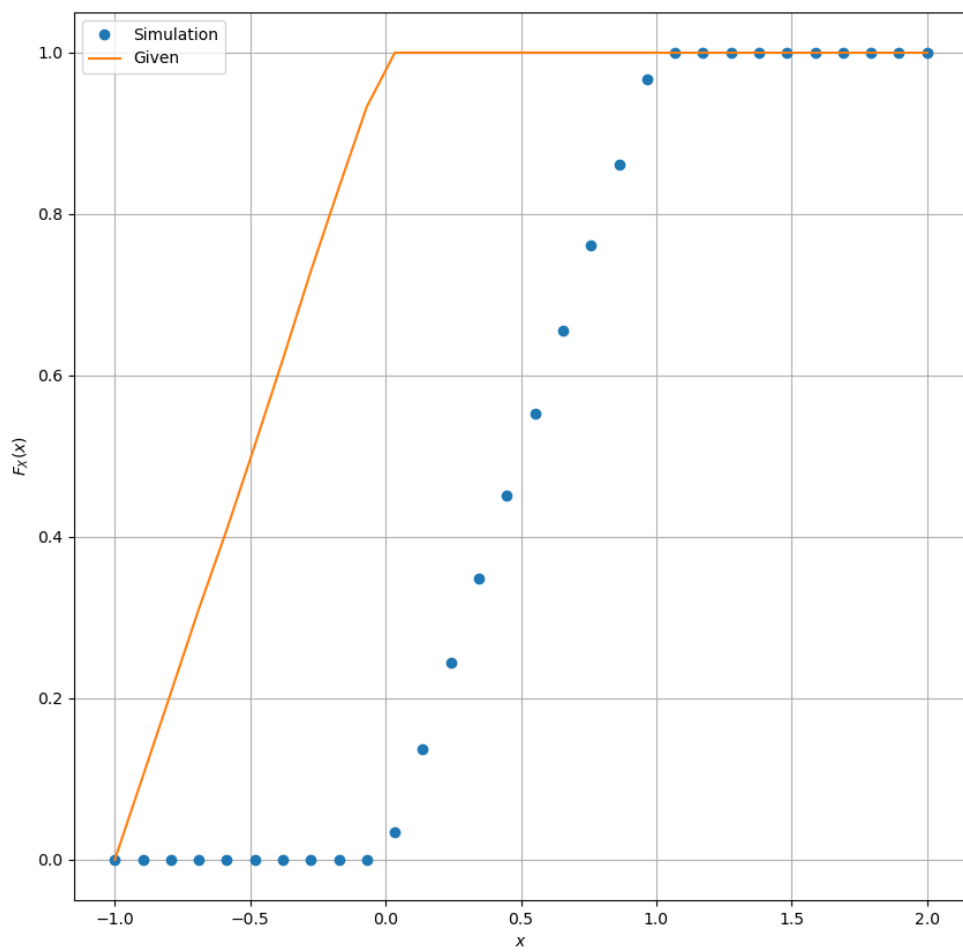


Figure 2.4:  $e^{-X}$  vs.  $U(-1, 0)$   
 Graphs don't match,  $\therefore$  wrong option

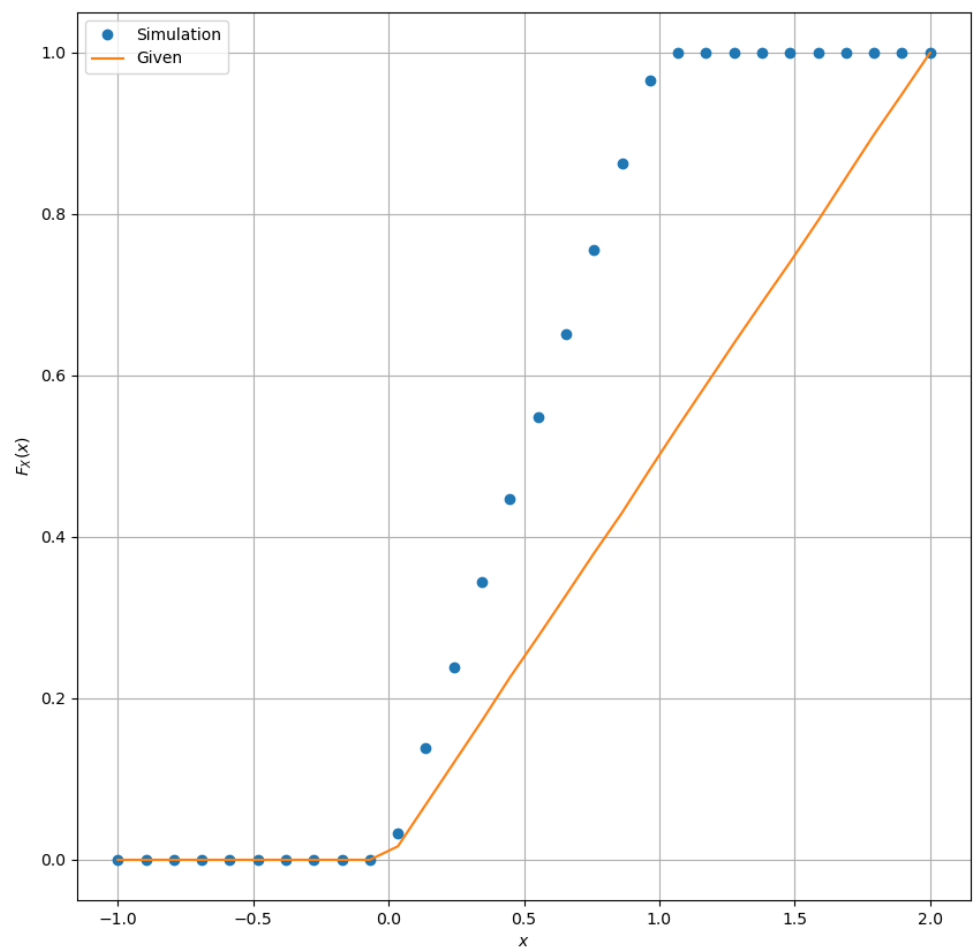


Figure 2.5:  $1 - e^{-X}$  vs.  $U(0, 2)$   
 Graphs don't match,  $\therefore$  wrong option

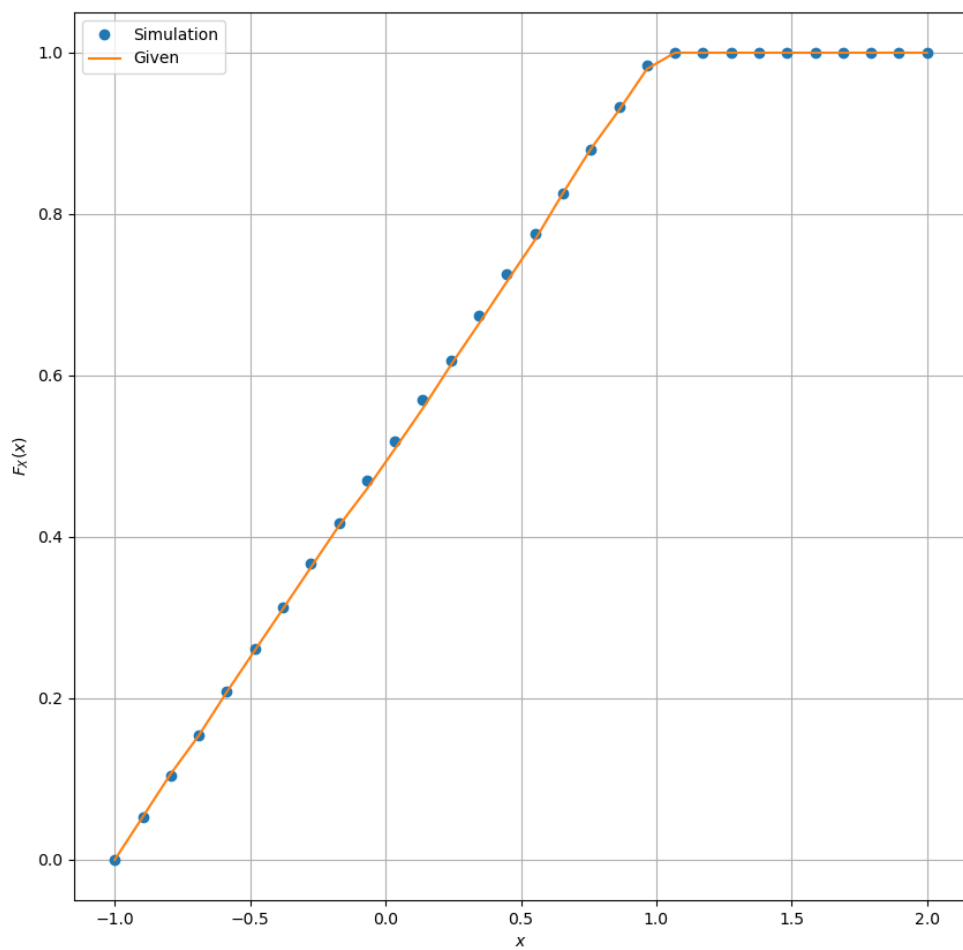


Figure 2.6:  $2e^{-X} - 1$  vs.  $U(-1, 1)$   
 Graphs match,  $\therefore$  correct option

2.7 Question: Let  $X$  be a positive valued continuous random variable with finite mean  $\mu$ .

If  $Y = [X]$ , the largest integer less than or equal to  $X$ , then which of the following

statements is/are true?

(A)  $\Pr(Y \leq \mu) \leq \Pr(X \leq \mu)$  for all  $\mu \geq 0$

(B)  $\Pr(Y \geq \mu) \leq \Pr(X \geq \mu)$  for all  $\mu \geq 0$

(C)  $E(X) < E(Y)$

(D)  $E(X) > E(Y)$

(GATE ST 2023)

**Solution:** Given that  $X$  is a positive valued random variable and  $Y = [X]$ . So,

$$X = Y + Z \quad (2.123)$$

Here,  $Z$  is an uniform distribution.

$$Z \sim U[0, 1) \quad (2.124)$$

$$F_Z(x) = x \quad (2.125)$$

$$E(Z) = \frac{1}{2} \quad (2.126)$$

Consider

(a)

$$\Pr(Y \leq \mu) = \Pr(X - Z \leq \mu) \quad (2.127)$$

$$= \Pr(Z \geq X - \mu) \quad (2.128)$$

$$= E(1 - F_Z(X - \mu)) \quad (2.129)$$

$$= E(1 - X + \mu) \quad (2.130)$$

$$= 1 - E(X) + \mu \quad (2.131)$$

$$= 1 \quad (2.132)$$

From option (A), we have  $1 \leq \Pr(X \leq \mu)$ . Option (A) is wrong since probability can't be greater than 1.

(b)

$$\Pr(Y \geq \mu) = \Pr(X - Z \geq \mu) \quad (2.133)$$

$$= \Pr(Z \leq X - \mu) \quad (2.134)$$

$$= E(F_Z(X - \mu)) \quad (2.135)$$

$$= E(X - \mu) \quad (2.136)$$

$$= E(X) - \mu \quad (2.137)$$

$$= 0 \quad (2.138)$$

From option B, we have  $\Pr(X \leq \mu) \geq 0$ . Option (B) is correct.

(c)

$$E(Y) = E(X - Z) \quad (2.139)$$

$$= E(X) - E(Z) \quad (2.140)$$

$$= \mu - \frac{1}{2} \quad (2.141)$$

$$= E(X) - \frac{1}{2} \quad (2.142)$$

$E(X) > E(Y)$ . Option (D) is correct and (C) is wrong.

#### Steps for Simulation:

- (a) Taking  $n$  samples, Generate  $n$  exponential random variable( $X$ ) samples.
- (b) Generate  $n$  samples of  $Y = [X]$  by floor to every sample of  $X$ .
- (c) Find number of samples of  $X$  where  $X \leq \mu$  and  $X \geq \mu$  and divide with  $n$  to get  $\Pr(X \leq \mu)$  and  $\Pr(X \geq \mu)$  respectively.
- (d) Find number of samples of  $Y$  where  $Y \leq \mu$  and  $Y \geq \mu$  and divide with  $n$  to get  $\Pr(Y \leq \mu)$  and  $\Pr(Y \geq \mu)$  respectively.
- (e) Sum the  $n$  samples of  $X$  and  $Y$  and divide with  $n$  to get  $E(X)$  and  $E(Y)$ .

**Note:** At  $x \in \text{integers}$ ,  $Y = X$ , so, CDF curves of  $Y$  and  $X$  are same. At non-integers we can see some difference in CDF curves in  $X$  and  $Y$ .

2.8 In a diploid angiosperm species, flower colour is regulated by the R gene. RR and Rr genotypes produce red flowers, whereas the rr genotype produces white flowers. If two individual plants are randomly selected from a large segregating population of a genetic cross between RR and rr parents, the probability of both the plants producing red flowers is

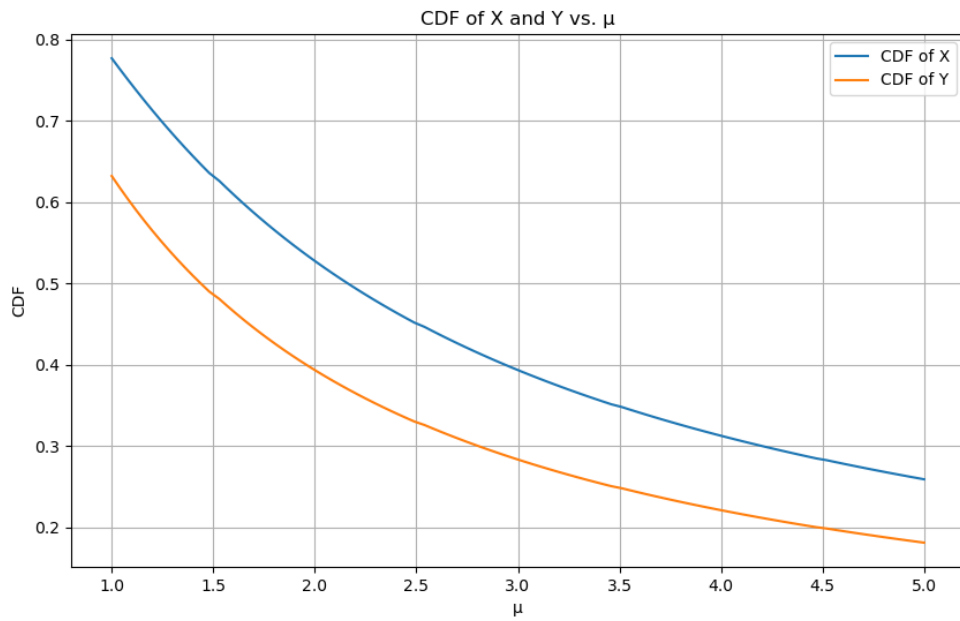


Figure 2.7: CDF'S of X and Y for varying  $\mu$  at  $x=1.5$

(GATE XL 2023)

**Solution:**

Gene	Representation
R	1
r	0

Table 2.4: Gene Representation.

For the parent genes:

Hence, we can see that it gives only Rr gene i.e.,10

For the children genes:

	1	1
0	10	10
0	10	10

Table 2.5: Gene of Parents.

	1	0
1	11	10
0	10	00

Table 2.6: Gene of Children.

$$p_X(k) = {}^nC_k p^k q^{n-k} \quad \forall k = 0, 1, 2 \quad (2.143)$$

$$= {}^2C_k \left(\frac{1}{2}\right)^k \left(\frac{1}{2}\right)^{2-k} \quad (2.144)$$

$$= {}^2C_k \left(\frac{1}{2}\right)^2 \quad (2.145)$$

we know that Red flower comes for RR and Rr i.e., 11 and 10

Therefore,

$$\Pr(X \leq 1) = 1 - \Pr(X = 2) \quad (2.146)$$

$$= 1 - \frac{1}{4} \quad (2.147)$$

$$= \frac{3}{4} \quad (2.148)$$

2.9 The frequencies for autosomal alleles  $A$  and  $a$  are  $p = 0.5$  and  $q = 0.5$ , respectively, where  $A$  is dominant over  $a$ . Under the assumption of random mating, the mating frequency among dominant parents is.

(GATE XL 2023)



RV	Values	Description
$X$	0	11
	1	10
	2	00

Table 2.7: Random varibale declaration

parameter	value
$n$	2
$p$	$\frac{1}{2}$
$q$	$\frac{1}{2}$

Table 2.8: Binomial parameters declaration

**Solution:** Given:  $A$  and  $a$  are two alleles where  $A$  is dominant one.

Let  $Y$  be a random variable depicting the number of dominant alleles in zygote( $AA, Aa, aA, aa$ ).

Parameter	Value	Description
$A$	1	dominant allele
$a$	0	Recessive allele
$n$	2	number of Alleles
$p$	0.5	frequency of dominant one
$q$	0.5	frquency of recessive one
$Y$	0,1,2	Number of dominant allele in zygote

(a) Theory:

Using binomial which states that

$$\Pr(Y = i) = {}^nC_i(p)^i(q)^{(n-i)} \quad (2.149)$$

For the mating frequency among the dominant parents, both parents must have

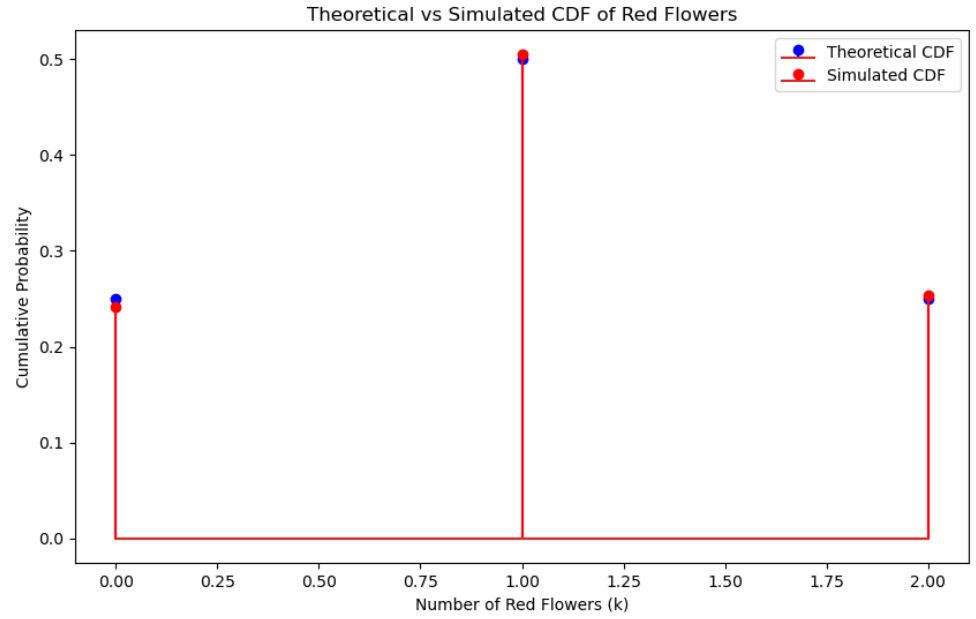


Figure 2.8: Simulation vs Theoretical

atleast one dominant allele. The probability of getting atleast 1 dominant allele in parent zygote :

$$\Pr(Y \geq 1) = \Pr(Y = 1) + \Pr(Y = 2) \quad (2.150)$$

$$= {}^2C_1(p)(q) + {}^2C_2(p)^2(1) \quad (2.151)$$

$$= 2pq + p^2 \quad (2.152)$$

$$= 0.5 + 0.25 \quad (2.153)$$

$$= 0.75 \quad (2.154)$$

(b) Step for Simulation of Random variable  $Y$ :

- i. Define the simulation size for simulation data set.
- ii. Generate two different random distribution to get two different bernoulli showing  $A$  as 1 and  $a$  as 0 each having frequency of 0.5.
- iii. The two bernoulli data for parents will mate with each other to form a zygote containig 1's and 0's.
- iv. Add up two bernouli to generate binomial distrubition for random variable  $Y$  showing the number of 1's or dominant allele in zygote.
- v. Count up the cases where there is atleast one 1 to generate the simulated probability.

2.10 Let  $X$  be a random variable having poisson distribution with mean  $\lambda > 0$ . Then

$E\left(\frac{1}{1+X} \mid X > 0\right)$  equals

(a)  $\frac{1-e^{-\lambda}-\lambda e^{-\lambda}}{\lambda(1-e^{-\lambda})}$

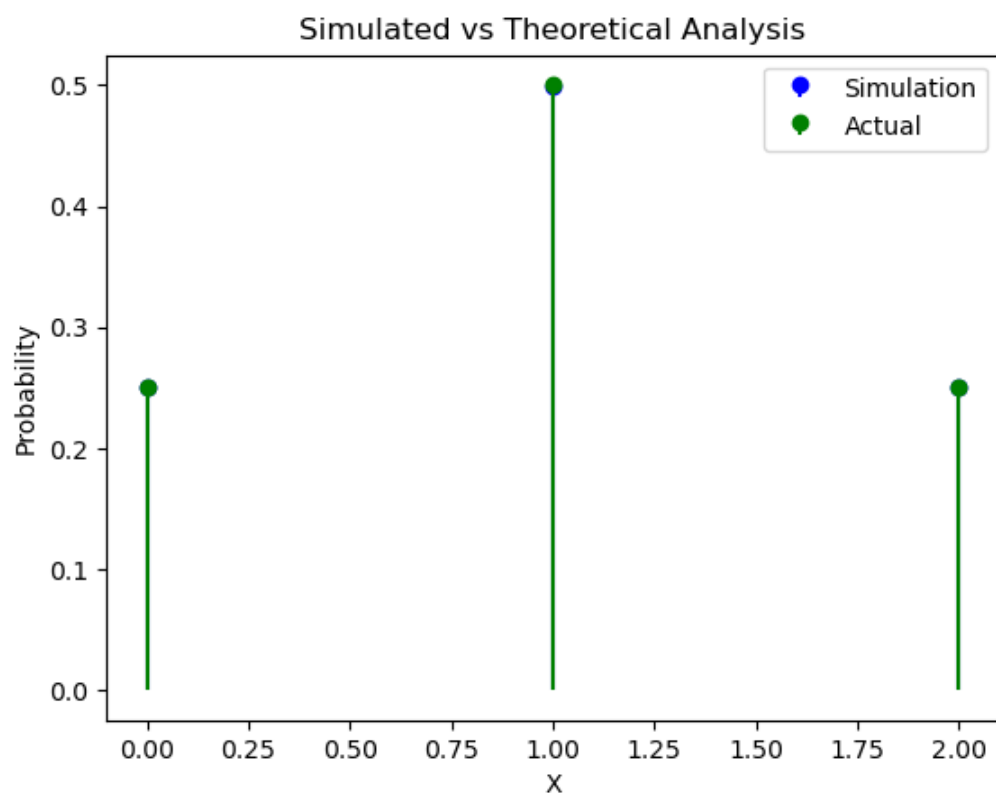
(b)  $\frac{1-e^{-\lambda}}{\lambda}$

(c)  $\frac{1-e^{-\lambda}-\lambda e^{-\lambda}}{\lambda}$

(d)  $\frac{1-e^{-\lambda}}{\lambda+1}$

(GATE ST 2023)

**Solution:**



(A) Theory

$$X \sim Pois(\lambda) \quad (2.155)$$

$$\Pr(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}; k \geq 0 \quad (2.156)$$

we know that

$$E(A|B) = \frac{E(A, B)}{\Pr(B)} \quad (2.157)$$

$$\Rightarrow E\left(\frac{1}{1+X} \middle| X > 0\right) = \frac{\sum_{k=1}^{\infty} \frac{1}{k+1} \Pr(X = k)}{\Pr(X > 0)} \quad (2.158)$$

$$= \frac{\sum_{k=1}^{\infty} \frac{1}{k+1} e^{-\lambda} \frac{\lambda^k}{k!}}{1 - \Pr(X \leq 0)} \quad (2.159)$$

$$= \frac{e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^k}{(k+1)!}}{1 - \Pr(X = 0)} \quad (2.160)$$

from equation (2.156)

$$= \frac{e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^k}{(k+1)!}}{1 - e^{-\lambda}} \quad (2.161)$$

Now simplifying Just the Summation

$$\Rightarrow \sum_{k=1}^{\infty} \frac{\lambda^k}{(k+1)!} \quad (2.162)$$

$$= \frac{1}{\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k+1}}{k+1} \quad (2.163)$$

Letting  $k + 1 = m$ ,

$$\Rightarrow \frac{1}{\lambda} \sum_{m=2}^{\infty} \frac{\lambda^m}{m!} \quad (2.164)$$

We Know from Taylor series

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!} \quad (2.165)$$

$$\Rightarrow \frac{1}{\lambda} \sum_{m=2}^{\infty} \frac{\lambda^m}{m!} = \frac{1}{\lambda} (e^{\lambda} - 1 - \lambda) \quad (2.166)$$

Substituting back we get,

$$= \frac{e^{-\lambda}}{(1 - e^{\lambda})} \left( \frac{1}{\lambda} (e^{\lambda} - 1 - \lambda) \right) \quad (2.167)$$

$$= \frac{e^{-\lambda}}{\lambda(1 - e^{\lambda})} (e^{\lambda} - 1 - \lambda) \quad (2.168)$$

$$= \frac{1 - e^{-\lambda} - \lambda e^{-\lambda}}{\lambda(1 - e^{\lambda})} \quad (2.169)$$

(B) Simulation

- (i) In the code, it simulates the generation of RV  $X$  using the CDF of Poisson distribution.

$$F(x) = \sum_{k=0}^x e^{-\lambda} \frac{\lambda^k}{k!} \quad (2.170)$$

- (ii) initialize  $X = 0$
- (iii) initialize  $F = e^{-\lambda}$  which is  $F(0)$  of a poisson distribution
- (iv) generate a uniform random variable between 0 and 1
- (v) enter the loop that continues as long as  $u > F$

Inside the loop,

- (vi) Increment  $X$  by 1

- (vii) Add next term of Poisson pmf ( $\Pr(X = k + 1)$ ) to  $F(x)$
- (viii) the loop continues until  $u$  is no longer greater than  $F$ . At this point,  $X$  represents the generated value of the poisson random variable that follows the desired poisson distribution with mean parameter  $\lambda$ .
- (ix) Save all the values of poisson random variable  $X$  in pois.dat so to open it in python an plot the cdf graph
- (x) Then for the second part Check if the generated value of  $X$  is greater than 0. If  $X$  is greater than 0, calculate the value  $Y$  as  $\frac{1}{X+1}$  and add it to the sumY.Increment the validCount to keep track of valid  $X$  values.
- (xi) If validCount is greater than 0, calculate the estimate of the conditional expectation by dividing sumY by validCount.

2.11 Consider the probability space  $(\Omega, \mathcal{G}, P)$ , where  $\Omega = \{1, 2, 3, 4\}$ ,  $\mathcal{G} = \{\emptyset, \Omega, \{1\}, \{4\}, \{2, 3\}, \{1, 4\}, \{1, 2, 3\}, \{2, 3, 4\}\}$ ,  $P(\{1\}) = \frac{1}{4}$ .

2.12 Let  $X$  be the random variable defined on the above probability space as  $X(1) = 1$ ,  $X(2) = X(3) = 2$ ,  $X(4) = 3$ . If  $P(X \leq 2) = \frac{3}{4}$ , then find  $P(\{1, 4\})$  (rounded off to two decimal places).

(GATE ST 2023)

**Solution:**

Table 2.9: Probability space

Probability space	Value
$\Omega$	$\{1, 2, 3, 4\}$
$\mathcal{G}$	$\{\emptyset, \Omega, \{1\}, \{4\}, \{2, 3\}, \{1, 4\}, \{1, 2, 3\}, \{2, 3, 4\}\}$
$P(\{1\})$	$\frac{1}{4}$
$P(X \leq 2)$	$\frac{3}{4}$

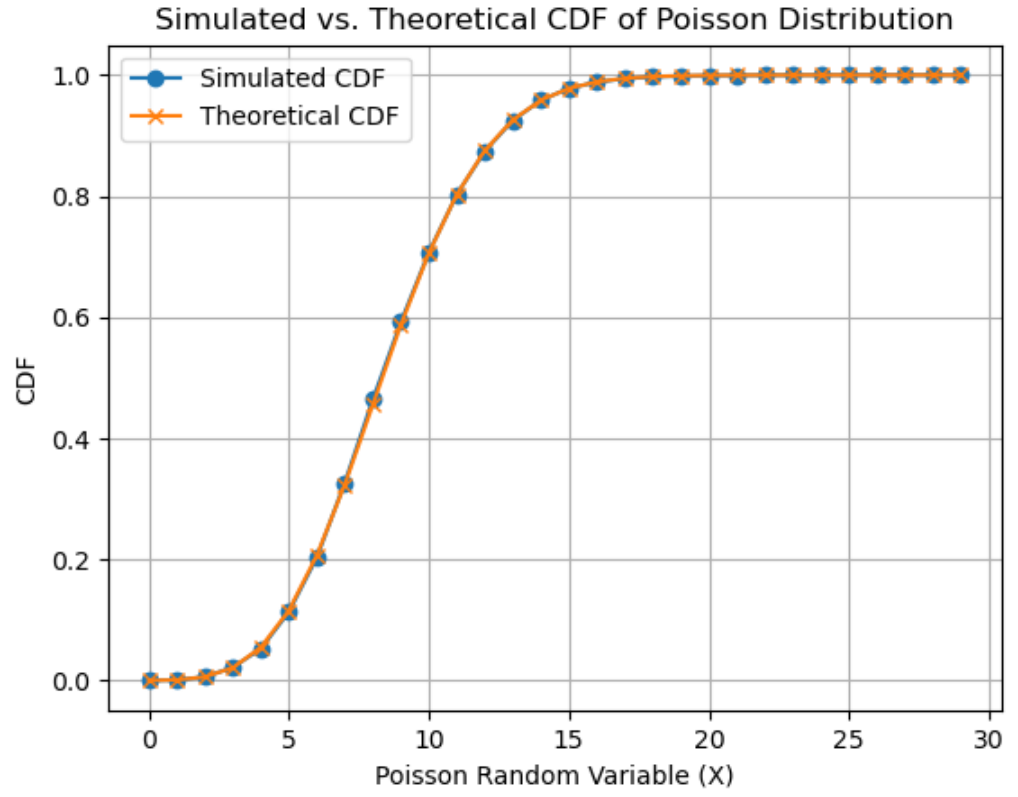


Figure 2.9: Simulation Vs theoretical cdf of poisson distribution with  $\lambda = 9$

Pmf is defined as

$$p_x(k) = \begin{cases} P(\{1\}) & , k = 1 \\ P(\{2, 3\}) & , k = 2 \\ P(\{4\}) & , k = 3 \end{cases} \quad (2.171)$$



Table 2.10: Random variable

$X(\Omega)$	$\Omega$
$\{1\}$	1
$\{2, 3\}$	2
$\{4\}$	3

Values of  $P(\{2,3\})$ ,  $P(\{4\})$  are unknown, so let  $p$ ,  $q$  be their respective values

$$p_X(k) = \begin{cases} \frac{1}{4} & , k = 1 \\ p & , k = 2 \\ q & , k = 3 \end{cases} \quad (2.172)$$

$$\Pr(\{1, 4\}) = p_X(1) + p_X(3) \quad (2.173)$$

We know

$$p_X(1) + p + q = 1 \quad (2.174)$$

We can express  $\Pr(X \leq 2)$  as:

$$\Pr(X \leq 2) = p_X(1) + p \quad (2.175)$$

$$(2.176)$$

We can expres above equations as:

$$\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} p \\ q \end{pmatrix} = \begin{pmatrix} \frac{3}{4} \\ \frac{1}{2} \end{pmatrix} \quad (2.177)$$

$$p = \frac{1}{2}, q = \frac{1}{4} \quad (2.178)$$

Finally

$$\Pr(\{1, 4\}) = P(\{1\}) + q \quad (2.179)$$

$$\Pr(\{1, 4\}) = \frac{1}{4} + \frac{1}{4} \quad (2.180)$$

$$\Pr(\{1, 4\}) = 0.5 \quad (2.181)$$

Steps for simulating random variable.

- (a) Define the simulation size for datast (samples).
- (b) Assign calculated probablity for each probablity space p1, p2, p3, p4.
- (c) Define Random to generate a random number between 0 and 1.
- (d) Define the loop such that it generated number 1, 2, 3 for defined probablity space.
- (e) Store the simulated data in a .dat file.
- (f) Using matplotlib lib of python generate a V-line graph from the data in .dat file by counting the number of 1, 2, 3 .

2.13 Let  $N(t)_{t \geq 0}$  be a Poisson process with rate 1. Consider the following statements.

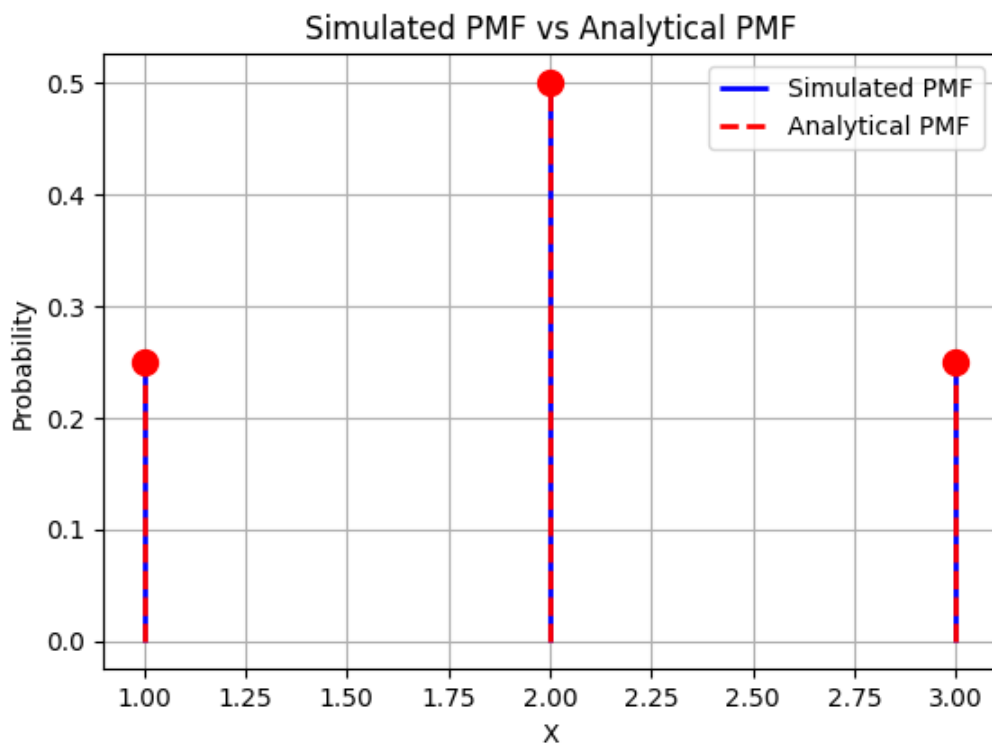


Figure 2.10: Analytical vs simulated

(a)  $\Pr(N(3) = 3 | N(5) = 5) = {}^5C_3 \left(\frac{3}{5}\right)^3 \left(\frac{2}{5}\right)^2$

(b) If  $S_5$  denotes the time of occurrence of the  $5^{th}$  event for the above Poisson process, then  $E(S_5 | N(5) = 3) = 7$

Which of the above statements is/are true?

- (i) only (a)
- (ii) only (b)
- (iii) Both (a) and (b)
- (iv) Neither (a) and (b)

( GATE ST 2023)

**Solution:**

Parameter	Values	Description
$X$	$N(t_1)$	poisson
$Y$	$N(t_2)$	random
$X + Y$	$N(t_1 + t_2)$	variables

Table 2.11: Table 1

(a) Using the Poisson probability formula,

$$\Pr(N(t) = k) = Po(t; k) = \frac{(\lambda t)^k e^{-\lambda t}}{k!} \quad (2.182)$$

here  $\lambda$  is 1

$$\Pr(N(t) = k) = \frac{(t)^k e^{-t}}{k!} \quad (2.183)$$

$$(2.184)$$

$X$  and  $Y$  are independent Poisson random variables, then  $X + Y$  is also Poisson

$$\Pr(X = k, X + Y = n) = pr(X = k, Y = n - k) \quad (2.185)$$

$$= \frac{(t_1)^k}{k!} e^{-t_1} \frac{(t_2)^{n-k}}{(n-k)!} e^{-t_2} \quad (2.186)$$

$$= e^{-(t_1+t_2)} \left( \frac{(t_1+t_2)^n}{n!} \right) {}^nC_k \left( \frac{t_1}{t_1+t_2} \right)^k \left( \frac{t_2}{t_1+t_2} \right)^{n-k} \quad (2.187)$$

$$\Pr(X + Y = n) = e^{-(t_1+t_2)} \left( \frac{(t_1 + t_2)^n}{n!} \right) \quad (2.188)$$

From conditional probability, from the equations (2.187) and (2.188)

$$\Pr(X = k | X + Y = n) = \frac{\Pr(X = k, Y = n - k)}{\Pr(X + Y = n)} \quad (2.189)$$

$$= {}^nC_k \left( \frac{t_1}{t_1 + t_2} \right)^k \left( \frac{t_2}{t_1 + t_2} \right)^{n-k} \quad (2.190)$$

For the given question,

Parameter	Values
$t_1$	3
$t_2$	5

Table 2.12: Table 1

$$\Pr(N(3) = 3 | N(5) = 5) = {}^5C_3 \left( \frac{3}{2+3} \right)^3 \left( \frac{2}{2+3} \right)^2 \quad (2.191)$$

$$= {}^5C_3 \left( \frac{3}{5} \right)^3 \left( \frac{2}{5} \right)^2 \quad (2.192)$$

Hence statement (a) is true.

Generation of poisson Random Variable from uniform in C language

(i) Define the Poisson Random Variable Generator Function:

In your program, define the poissonRandomVariable function to generate Poisson random variables with a given lambda parameter

(ii) "lambda" is the mean parameter for the Poisson distribution, representing the average rate of events in the given interval.

L is exp(-lambda), where exp is the exponential function. This value represents the probability of having zero events in the interval.

The function enters a loop that continues until p is less than or equal to L.

(iii) rand() / (double)RAND\_MAX:

This generates a random variable

(iv) the Main Function:

”numSamples” controls how many random samples you want to generate.

(v) srand(time(NULL)):

The ”srand(time(NULL))” line seeds the random number generator using the current time to ensure different random sequences each time you run the program.

2.14 Suppose that  $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n, \mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_n$  are independent and identically distributed random vectors each having  $N_p(\boldsymbol{\mu}, \Sigma)$  distributions, where  $\Sigma$  is non-singular,  $p > 1$  and  $n > 1$ . If  $\mathbf{X} = \frac{1}{n} \sum_{i=1}^n \mathbf{X}_i$  and  $\mathbf{Y} = \frac{1}{n} \sum_{i=1}^n \mathbf{Y}_i$ , then which one of the following statements is true?

- (a) There exists  $c > 0$  such that  $c(\mathbf{X} - \boldsymbol{\mu})^T \Sigma^{-1} (\mathbf{X} - \boldsymbol{\mu})$  has  $\chi^2$ -distribution with  $p$  degrees of freedom.
- (b) There exists  $c > 0$  such that  $c(\mathbf{X} - \mathbf{Y})^T \Sigma^{-1} (\mathbf{X} - \mathbf{Y})$  has  $\chi^2$ -distribution with  $(p - 1)$  degrees of freedom.
- (c) There exists  $c > 0$  such that  $c \sum_{i=1}^n (\mathbf{X}_i - \mathbf{X})^T \Sigma^{-1} (\mathbf{X}_i - \mathbf{X})$  has  $\chi^2$ -distribution with  $p$  degrees of freedom.
- (d) There exists  $c > 0$  such that  $c \sum_{i=1}^n (\mathbf{X}_i - \mathbf{Y}_i - \mathbf{X} + \mathbf{Y})^T \Sigma^{-1} (\mathbf{X}_i - \mathbf{Y}_i - \mathbf{X} + \mathbf{Y})$  has  $\chi^2$ -distribution with  $p$  degrees of freedom.

GATE ST Paper 2023

**Solution:**

We are given that,

$$\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n, \mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_n \sim N_p(\boldsymbol{\mu}, \Sigma) \quad (2.193)$$

Also,

$$\mathbf{X} = \frac{1}{n} \sum_{i=1}^n \mathbf{X}_i \quad (2.194)$$

$$\mathbf{Y} = \frac{1}{n} \sum_{i=1}^n \mathbf{Y}_i \quad (2.195)$$

The mean of  $\mathbf{X}$  is given by:

$$\boldsymbol{\mu}_{\mathbf{X}} = E(\mathbf{X}) \quad (2.196)$$

$$= \frac{1}{n} \sum_{i=1}^n E(\mathbf{X}_i) \quad (2.197)$$

$$= \boldsymbol{\mu} \quad (2.198)$$

Similarly,

$$\boldsymbol{\mu}_{\mathbf{Y}} = \boldsymbol{\mu} \quad (2.199)$$

The covariance of  $\mathbf{X}$  is given by:

$$\Sigma_{\mathbf{X}} = E \left[ (\mathbf{X} - \boldsymbol{\mu}) (\mathbf{X} - \boldsymbol{\mu})^T \right] \quad (2.200)$$

$$= E \left[ \left( \frac{1}{n} \sum_{i=1}^n \mathbf{X}_i - \boldsymbol{\mu} \right) \left( \frac{1}{n} \sum_{i=1}^n \mathbf{X}_i - \boldsymbol{\mu} \right)^T \right] \quad (2.201)$$

$$= \frac{1}{n^2} E \left[ \sum_{i=1}^n (\mathbf{X}_i - \boldsymbol{\mu}) (\mathbf{X}_i - \boldsymbol{\mu})^T \right] \quad (2.202)$$

$$= \frac{1}{n^2} \left[ \sum_{i=1}^n E (\mathbf{X}_i^2 + \boldsymbol{\mu}^2 - 2\boldsymbol{\mu}\mathbf{X}_i) \right] \quad (2.203)$$

$$= \frac{1}{n^2} \left[ \sum_{i=1}^n E (\mathbf{X}_i^2) + \sum_{i=1}^n E (\boldsymbol{\mu}^2) - 2\boldsymbol{\mu} \sum_{i=1}^n E (\mathbf{X}_i) \right] \quad (2.204)$$

$$= \frac{1}{n^2} [n\Sigma + n\boldsymbol{\mu}^2 + n\boldsymbol{\mu}^2 - 2\boldsymbol{\mu}^2] \quad \left[ \because E (\mathbf{X}_i^2) = \Sigma_{\mathbf{X}_i} + E (\mathbf{X}_i)^2 \right] \quad (2.205)$$

$$= \frac{\Sigma}{n} \quad (2.206)$$

Similarly,

$$\Sigma_{\mathbf{Y}} = \frac{\Sigma}{n} \quad (2.207)$$

(a) To check option (A):

let us say,

$$\mathbf{A} = c(\mathbf{X} - \boldsymbol{\mu})^T \Sigma^{-1} (\mathbf{X} - \boldsymbol{\mu}) \quad (2.208)$$

$$(2.209)$$



And,

$$\Sigma^{-1} = \mathbf{F}^T \mathbf{F} \quad (2.210)$$

$$\mathbf{y} = \mathbf{F} (\mathbf{X} - \boldsymbol{\mu}) \quad (2.211)$$

$$\implies \mathbf{A} = c \mathbf{y}^T \bar{\mathbf{y}} \quad (2.212)$$

$$= c \|\mathbf{y}\|^2 \quad (2.213)$$

Equation (2.213) shows that  $\mathbf{A}$  can have  $\chi^2$ -distribution.

To confirm that we will find the mean and covariance-matrix of  $\bar{\mathbf{y}}$ .

$$\boldsymbol{\mu}_{\mathbf{y}} = E(\mathbf{y}) \quad (2.214)$$

$$= E(\mathbf{F}(\mathbf{X} - \boldsymbol{\mu})) \quad (2.215)$$

$$= \mathbf{F} [E(\mathbf{X}) - E(\boldsymbol{\mu})] \quad (2.216)$$

$$= \mathbf{F} [\boldsymbol{\mu} - \boldsymbol{\mu}] \quad \text{from (2.198)} \quad (2.217)$$

$$= 0 \quad (2.218)$$

And,

$$\Sigma_{\mathbf{y}} = E[(\mathbf{y} - \boldsymbol{\mu}_{\mathbf{y}})(\mathbf{y} - \boldsymbol{\mu}_{\mathbf{y}})^T] \quad (2.219)$$

$$= E[(\mathbf{F}(\mathbf{X} - \boldsymbol{\mu}))(\mathbf{F}(\mathbf{X} - \boldsymbol{\mu}))^T] \quad (2.220)$$

$$= E[\mathbf{F}(\mathbf{X} - \boldsymbol{\mu})(\mathbf{X} - \boldsymbol{\mu})^T \mathbf{F}^T] \quad (2.221)$$

$$= \mathbf{F} [E[(\mathbf{X} - \boldsymbol{\mu})(\mathbf{X} - \boldsymbol{\mu})^T]] \mathbf{F}^T \quad (2.222)$$

$$= \mathbf{F} \Sigma \mathbf{F}^T \quad (2.223)$$

since,

$$\Sigma^{-1} = \mathbf{F}^T \mathbf{F} \quad (2.224)$$

$$\Sigma \Sigma^{-1} = \Sigma \mathbf{F}^T \mathbf{F} \quad (2.225)$$

$$\mathbf{I} = \Sigma \mathbf{F}^T \mathbf{F} \quad (2.226)$$

$$\mathbf{I} \mathbf{F}^{-1} = \Sigma \mathbf{F}^T \quad (2.227)$$

$$\mathbf{F} \mathbf{F}^{-1} = \mathbf{F} \Sigma \mathbf{F}^T \quad (2.228)$$

$$\mathbf{I} = \mathbf{F} \Sigma \mathbf{F}^T \quad (2.229)$$

So using (2.229),

$$\Sigma_{\mathbf{y}} = \mathbf{I} \quad (2.230)$$

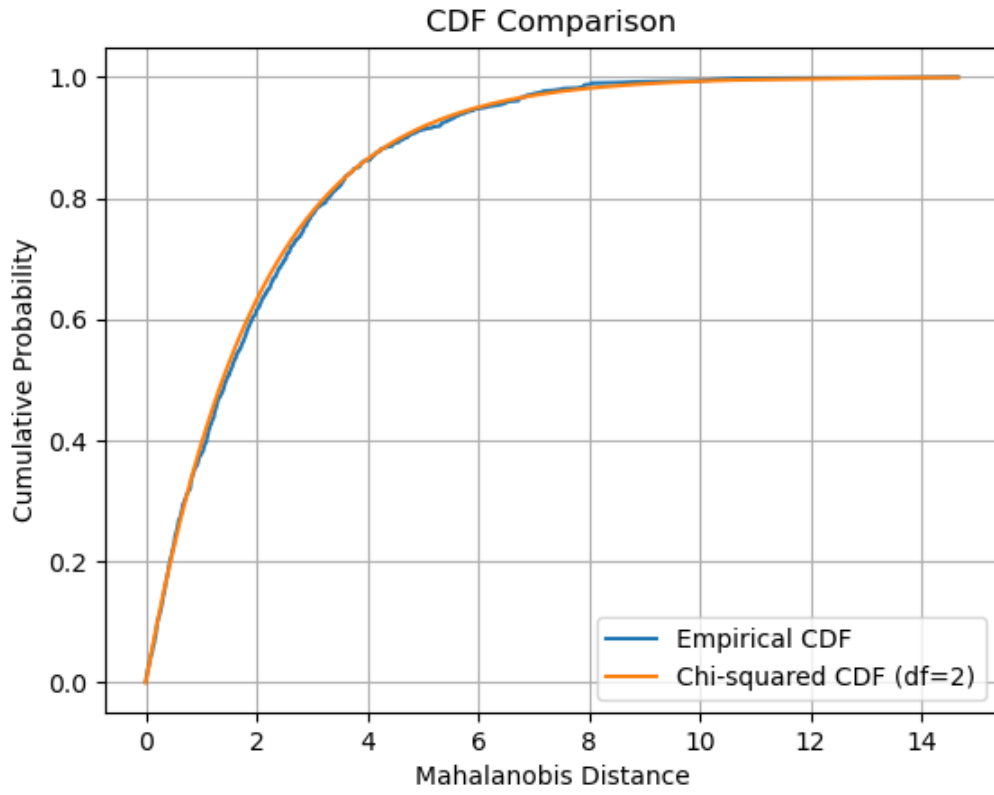
Hence, For  $c = 1$   $\mathbf{A}$  has  $\chi^2$ -distribution with  $p$  degrees of freedom.

So option (A) is correct.

(b) To check option (B):

Let us say,

$$\mathbf{B} = c(\mathbf{X} - \mathbf{Y})^T \Sigma^{-1} (\mathbf{X} - \mathbf{Y}) \quad (2.231)$$



And,

$$\Sigma^{-1} = \mathbf{F}^T \mathbf{F} \quad (2.232)$$

$$\mathbf{y} = \mathbf{F} (\mathbf{X} - \mathbf{Y}) \quad (2.233)$$

$$\Rightarrow \mathbf{B} = c \mathbf{y}^T \bar{\mathbf{y}} \quad (2.234)$$

$$= c \|\mathbf{y}\|^2 \quad (2.235)$$

Equation (2.235) shows that  $\mathbf{B}$  can have  $\chi^2$ -distribution.

To confirm that we will find the mean and covariance-matrix of  $\bar{\mathbf{y}}$ .

$$\boldsymbol{\mu}_{\mathbf{y}} = E(\mathbf{y}) \quad (2.236)$$

$$= E[F(\mathbf{X} - \mathbf{Y})] \quad (2.237)$$

$$= F[E(\mathbf{X}) - E(\mathbf{Y})] \quad (2.238)$$

$$= F[\boldsymbol{\mu} - \boldsymbol{\mu}] \quad (2.239)$$

$$= 0 \quad (2.240)$$

And,

$$\Sigma_{\mathbf{y}} = E[(\mathbf{y} - \boldsymbol{\mu}_{\mathbf{y}})(\mathbf{y} - \boldsymbol{\mu}_{\mathbf{y}})^T] \quad (2.241)$$

$$= E[(\mathbf{F}(\mathbf{X} - \mathbf{Y}))(\mathbf{F}(\mathbf{X} - \mathbf{Y}))^T] \quad (2.242)$$

$$= E[\mathbf{F}(\mathbf{X} - \mathbf{Y})(\mathbf{X} - \mathbf{Y})^T \mathbf{F}^T] \quad (2.243)$$

$$= \mathbf{F} \left[ E[(\mathbf{X} - \mathbf{Y})(\mathbf{X} - \mathbf{Y})^T] \right] \mathbf{F}^T \quad (2.244)$$

$$= \mathbf{F} \left[ E[\|\mathbf{X} - \mathbf{Y}\|^2] \right] \mathbf{F}^T \quad (2.245)$$

$$= \mathbf{F} [E(\mathbf{X}^2) + E(\mathbf{Y}^2) - E(2\mathbf{X}\mathbf{Y})] \mathbf{F}^T \quad (2.246)$$

$$= \mathbf{F} \left[ \frac{\Sigma}{n} + \boldsymbol{\mu}^2 + \frac{\Sigma}{n} + \boldsymbol{\mu}^2 - 2\boldsymbol{\mu}^2 \right] \mathbf{F}^T \quad \left[ \because E(\mathbf{X}^2) = \Sigma_{\mathbf{X}} + E(\mathbf{X})^2 \right] \quad (2.247)$$

$$= \frac{2}{n} \mathbf{F} \Sigma \mathbf{F}^T \quad (2.248)$$

$$= \frac{2}{n} \mathbf{I} \quad (2.249)$$

Hence, for  $c = \frac{n}{2}$ ,  $\mathbf{B}$  has  $\chi^2$ -distribution with p degrees of freedom.

So option (B) is incorrect.

(c) To check option (C):

let us say,

$$\mathbf{C} = c \sum_{i=1}^n (\mathbf{X}_i - \mathbf{X})^T \Sigma^{-1} (\mathbf{X}_i - \mathbf{X}) \quad (2.250)$$

And,

$$\Sigma^{-1} = \mathbf{F}^T \mathbf{F} \quad (2.251)$$

$$\mathbf{y} = \mathbf{F} \left( \sum_{i=1}^n (\mathbf{X}_i - \mathbf{X}) \right) \quad (2.252)$$

$$\implies \mathbf{C} = c \mathbf{y}^T \bar{\mathbf{y}} \quad (2.253)$$

$$= c \|\mathbf{y}\|^2 \quad (2.254)$$

Equation (2.254) shows that  $\mathbf{C}$  can have  $\chi^2$ -distribution.

To confirm that we will find the mean and covariance-matrix of  $\bar{\mathbf{y}}$ .

$$\mu_{\mathbf{y}} = E(\mathbf{y}) \quad (2.255)$$

$$= E \left[ \mathbf{F} \left( \sum_{i=1}^n (\mathbf{X}_i - \mathbf{X}) \right) \right] \quad (2.256)$$

$$= \mathbf{F} \left[ \sum_{i=1}^n (E(\mathbf{X}_i) - E(\mathbf{X})) \right] \quad (2.257)$$

$$= \mathbf{F} [E(X_1) - E(X) + E(X_2) - E(X) + \dots + E(X_n) - E(X)] \quad (2.258)$$

$$= 0 \quad (2.259)$$

And,

$$\Sigma_{\mathbf{y}} = E \left[ (\mathbf{y} - \boldsymbol{\mu}_{\mathbf{y}}) (\mathbf{y} - \boldsymbol{\mu}_{\mathbf{y}})^T \right] \quad (2.260)$$

$$= \mathbf{F} E \left[ \left( \sum_{i=1}^n (\mathbf{X}_i - \mathbf{X}) \right) \left( \sum_{i=1}^n (\mathbf{X}_i - \mathbf{X}) \right)^T \right] \mathbf{F}^T \quad (2.261)$$

$$= \mathbf{F} E \left[ \left( \sum_{i=1}^n \mathbf{X}_i - n\mathbf{X} \right) \left( \sum_{i=1}^n \mathbf{X}_i - n\mathbf{X} \right)^T \right] \mathbf{F}^T \quad (2.262)$$

$$= \mathbf{F} E \left[ (n\mathbf{X} - n\mathbf{X}) (n\mathbf{X} - n\mathbf{X})^T \right] \mathbf{F}^T \quad (2.263)$$

$$= \mathbf{F} \mathbf{0} \mathbf{F}^T \quad (2.264)$$

$$= \mathbf{0} \quad (2.265)$$

Hence, There is no value of  $c > 0$  for which  $\mathbf{C}$  have  $\chi^2$ -distribution.

So option (C) is incorrect.

(d) To check option (D):

let us say,

$$\mathbf{D} = c \sum_{i=1}^n (\mathbf{X}_i - \mathbf{Y}_i - \mathbf{X} + \mathbf{Y})^T \Sigma^{-1} (\mathbf{X}_i - \mathbf{Y}_i - \mathbf{X} + \mathbf{Y}) \quad (2.266)$$

And,

$$\Sigma^{-1} = \mathbf{F}^T \mathbf{F} \quad (2.267)$$

$$\mathbf{y} = \mathbf{F} \left( \sum_{i=1}^n (\mathbf{X}_i - \mathbf{Y}_i - \mathbf{X} + \mathbf{Y}) \right) \quad (2.268)$$

$$\implies \mathbf{C} = c \mathbf{y}^T \bar{\mathbf{y}} \quad (2.269)$$

$$= c \|\mathbf{y}\|^2 \quad (2.270)$$

Equation (2.270) shows that  $\mathbf{D}$  can have  $\chi^2$ -distribution.

To confirm that we will find the mean and covariance-matrix of  $\bar{\mathbf{y}}$ .

$$\boldsymbol{\mu}_{\mathbf{y}} = E(\mathbf{y}) \quad (2.271)$$

$$= E\left(\mathbf{F}\left(\sum_{i=1}^n (\mathbf{X}_i - \mathbf{Y}_i - \mathbf{X} + \mathbf{Y})\right)\right) \quad (2.272)$$

$$= \mathbf{F}E\left[\sum_{i=1}^n \mathbf{X}_i - \sum_{i=1}^n \mathbf{Y}_i - n\mathbf{X} + n\mathbf{Y}\right] \quad (2.273)$$

$$= \mathbf{F}\left[\sum_{i=1}^n E(\mathbf{X}_i) - \sum_{i=1}^n E(\mathbf{Y}_i) - nE(\mathbf{X}) + nE(\mathbf{Y})\right] \quad (2.274)$$

$$= \mathbf{F}[n\boldsymbol{\mu} - n\boldsymbol{\mu} - n\boldsymbol{\mu} + n\boldsymbol{\mu}] \quad (2.275)$$

$$= \mathbf{0} \quad (2.276)$$

And,

$$\Sigma_{\mathbf{y}} = E[(\mathbf{y} - \boldsymbol{\mu}_{\mathbf{y}})(\mathbf{y} - \boldsymbol{\mu}_{\mathbf{y}})^T] \quad (2.277)$$

$$= \mathbf{F}E\left[\left(\sum_{i=1}^n (\mathbf{X}_i - \mathbf{Y}_i - \mathbf{X} + \mathbf{Y})\right)\left(\sum_{i=1}^n (\mathbf{X}_i - \mathbf{Y}_i - \mathbf{X} + \mathbf{Y})\right)^T\right]\mathbf{F}^T \quad (2.278)$$

$$= \mathbf{F}E\left[\left(\sum_{i=1}^n \mathbf{X}_i - \sum_{i=1}^n \mathbf{Y}_i - n\mathbf{X} + n\mathbf{Y}\right)\left(\sum_{i=1}^n \mathbf{X}_i - \sum_{i=1}^n \mathbf{Y}_i - n\mathbf{X} + n\mathbf{Y}\right)^T\right]\mathbf{F}^T \quad (2.279)$$

$$= \mathbf{F}E[(n\mathbf{X} - n\mathbf{Y} - n\mathbf{X} + n\mathbf{Y})(n\mathbf{X} - n\mathbf{Y} - n\mathbf{X} + n\mathbf{Y})^T]\mathbf{F}^T \quad (2.280)$$

$$= \mathbf{F}\mathbf{0}\mathbf{F}^T \quad (2.281)$$

$$= \mathbf{0} \quad (2.282)$$

Hence, There is no value of  $c > 0$  for which  $\mathbf{D}$  have  $\chi^2$ -distribution.

So option (D) is incorrect.

### **Mahalanobis Distance:**

It is the measure of the distance between a point  $\mathbf{X}$  and a distribution  $Q$ . It is a multi-dimensional generalization of the idea of measuring how many standard deviations away  $\mathbf{X}$  is from the mean of  $Q$ . So, Given a probability distribution  $Q$  on  $R^N$  with,

$$\boldsymbol{\mu} = (\mu_1, \mu_2, \dots, \mu_N)^T \quad (2.283)$$

and positive covariance matrix  $\Sigma$ , the mahalanobis distance of a point,

$$\mathbf{X} = (X_1, X_2, \dots, X_N)^T \quad (2.284)$$

is given by,

$$d_M(\mathbf{X}, Q) = \sqrt{(\mathbf{X} - \boldsymbol{\mu})^T \Sigma^{-1} (\mathbf{X} - \boldsymbol{\mu})} \quad (2.285)$$

### **Steps for simulation:**

- (a) Firstly in the the file "gauss.c", I have generated 1000 random vectors with dimension 2 using Box-Muller method and listed the data in the file "randomvectors.dat".
- (b) Then in the file "distance.c", using the random vectors generated in the first step, I found the value of  $c(\mathbf{X} - \boldsymbol{\mu})^T \Sigma^{-1} (\mathbf{X} - \boldsymbol{\mu})$  distribution which will give us  $1 \times 1$  matrix.
- (c) So as we have generated 1000 random vectors in first step, we will have 1000 values of the distribution.



- (d) Then I have listed the values of the distribution  $c(\mathbf{X} - \boldsymbol{\mu})^T \Sigma^{-1} (\mathbf{X} - \boldsymbol{\mu})$  that I got in the file "mahalanobisdistances.dat".
- (e) Now in the file "cdf.py", I have plotted the cdf of the distribution  $c(\mathbf{X} - \boldsymbol{\mu})^T \Sigma^{-1} (\mathbf{X} - \boldsymbol{\mu})$  and also plotted the theoretical cdf plot of a  $\chi^2$  distribution.

The variables that are used in the simulation are:

Variable	Definition
$\mathbf{X}$	random vector
p	dimension of vector
n	number of vectors
$\boldsymbol{\mu}$	mean
$\boldsymbol{\Sigma}$	Covariance

2.15 In a locality 'A', the probability of a convective storm event is 0.7 with a density function,

$$f_{X_1}(x_1) = e^{-x_1}, \quad x_1 > 0 \quad (2.286)$$

The probability of tropical cyclone-induced storm in the same location is given by the density function,

$$f_{X_2}(x_2) = 2e^{-2x_2}, \quad x_2 > 0 \quad (2.287)$$

The probability of occurring more than 1 unit of storm event is  
(GATE AG 2023)

**Solution:**

**Laplace Transform**

Let  $X$  be a random variable such that

$$X = X_1 + X_2 \quad (2.288)$$

Given,

$$f_{X_1}(x) = e^{-x}u(x) \quad (2.289)$$

$$f_{X_2}(x) = 2e^{-2x}u(x) \quad (2.290)$$

Where,  $u(x)$  is unit step function.

CDF of  $X_1$ :

$$F_{X_1}(x) = \int_{-\infty}^x f_{X_1}(x) dx \quad (2.291)$$

$$= \int_0^x e^{-x} dx \quad (2.292)$$

$$= 1 - e^{-x} \quad (2.293)$$

CDF of  $X_2$ :

$$F_{X_2}(x) = \int_{-\infty}^x f_{X_2}(x) dx \quad (2.294)$$

$$= \int_0^x 2e^{-2x} dx \quad (2.295)$$

$$= 1 - e^{-2x} \quad (2.296)$$

Now,

$$M_X(s) = M_{X_1}(s) \cdot M_{X_2}(s) \quad (2.297)$$

$$M_{X_1}(s) = \int_{-\infty}^{\infty} f_{X_1}(x) \cdot e^{-sx} dx \quad (2.298)$$

$$= \int_0^{\infty} e^{-x} \cdot e^{-sx} u(x) dx \quad (2.299)$$

$$= \frac{1}{s+1} \quad (2.300)$$

Region of Convergence(ROC) of  $M_{X_1}(s)$ :

$$Re(s) > -1 \quad (2.301)$$

Now,

$$M_{X_2}(s) = \int_{-\infty}^{\infty} f_{X_2}(x) \cdot e^{-sx} dx \quad (2.302)$$

$$= \int_0^{\infty} 2e^{-2x} \cdot e^{-sx} u(x) dx \quad (2.303)$$

$$= \frac{2}{s+2} \quad (2.304)$$

ROC of  $M_{X_2}(s)$ :

$$Re(s) > -2 \quad (2.305)$$

Using (2.300) and (2.304) in (2.297)

$$M_X(s) = \frac{1}{s+1} \times \frac{2}{s+2} \quad (2.306)$$

$$= \frac{2}{(s+1)(s+2)} \quad (2.307)$$

$$p_X(x) = L^{-1}[M_X(s)] \quad (2.308)$$

$$= L^{-1}\left[\frac{2}{(s+1)(s+2)}\right] \quad (2.309)$$

$$= 2L^{-1}\left[\frac{1}{s+1} - \frac{1}{s+2}\right] \quad (2.310)$$

ROC of laplace transform:

$$Re(s) > -1 \cap Re(s) > -2 \quad (2.311)$$

$$\implies Re(s) > -1 \quad (2.312)$$

So, now

$$p_X(x) = 2(e^{-x} - e^{-2x})u(x) \quad (2.313)$$

CDF of X:

$$F_X(x) = \int_{-\infty}^x p_X(x) dx \quad (2.314)$$

$$= 2 \int_{-\infty}^x (e^{-x} - e^{-2x}) u(x) dx \quad (2.315)$$

But  $p_X(x)$  is integrable for  $x > 0$

$$F_X(x) = 2 \int_0^x (e^{-x} - e^{-2x}) dx \quad (2.316)$$

$$= 2 \left( -e^{-x} + \frac{1}{2}e^{-2x} \right) \Big|_0^x \quad (2.317)$$

$$= 2 \left( -e^{-x} + \frac{1}{2}e^{-2x} + \frac{1}{2} \right) \quad (2.318)$$

$$= -2e^{-x} + e^{-2x} + 1 \quad (2.319)$$

Now,

$$\Pr(X > 1) = F_X(1) \quad (2.320)$$

$$= -2e^{-1} + e^{-2} + 1 \quad (2.321)$$

$$= 0.39 \quad (2.322)$$

Steps for simulation the given distribution

(a) CDF of  $X_1$  is given as:

$$F_{X_1}(x) = \begin{cases} 0 & x \leq 0 \\ 1 - e^{-x} & x > 0 \end{cases} \quad (2.323)$$

(b) Declare a function inverse cdf ( $I(u)$ ) such that its input is any random number and output is random variable whose cdf equals that of the given distribution

For  $x \leq 0$

$$u = 0 \quad (2.324)$$

$$\because x \leq 0 \quad (2.325)$$

$$u \leq 0 \quad (2.326)$$

For  $x > 0$

$$u = 1 - e^{-x} \quad (2.327)$$

$$e^{-x} = 1 - u \quad (2.328)$$

$$x = -\ln(1 - u) \quad (2.329)$$

$$\because x > 0 \quad (2.330)$$

$$u > 0 \quad (2.331)$$

$$I(u) = \begin{cases} 0 & u \leq 0 \\ -\ln(1 - u) & u > 0 \end{cases} \quad (2.332)$$

- (c) Define three arrays `random_vars` , `cdf.values` , `theoretical_cdf_values` to store random variables, simulated cdf values and theoretical cdf values
- (d) Generate random numbers using `rand()` and calling inverse cdf function to generate our random variable
- (e) Calling cdf function to calculate the cdf of the generated random variable
- (f) Storing the random variable,theoretical cdf and generated cdf into their respective arrays
- (g) Storing the data of these three array into a `.dat` file
- (h) Plotting these `.dat` file in python
- (i) Repeat all these steps for  $X_2$  as well as  $X$

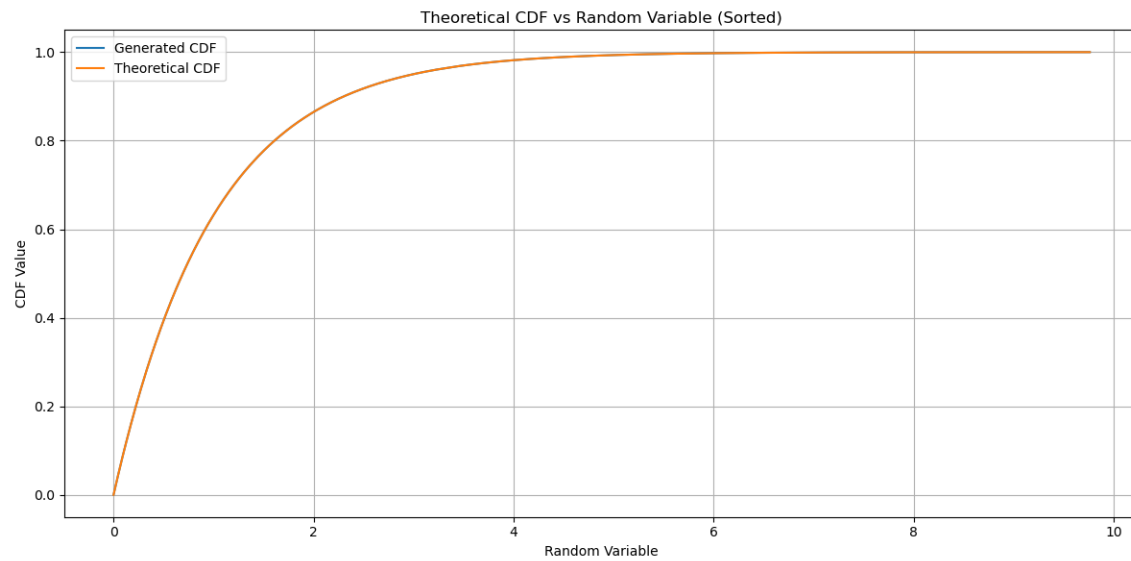


Figure 2.11: Theoretical vs Simulation Analysis wrt  $X_1$

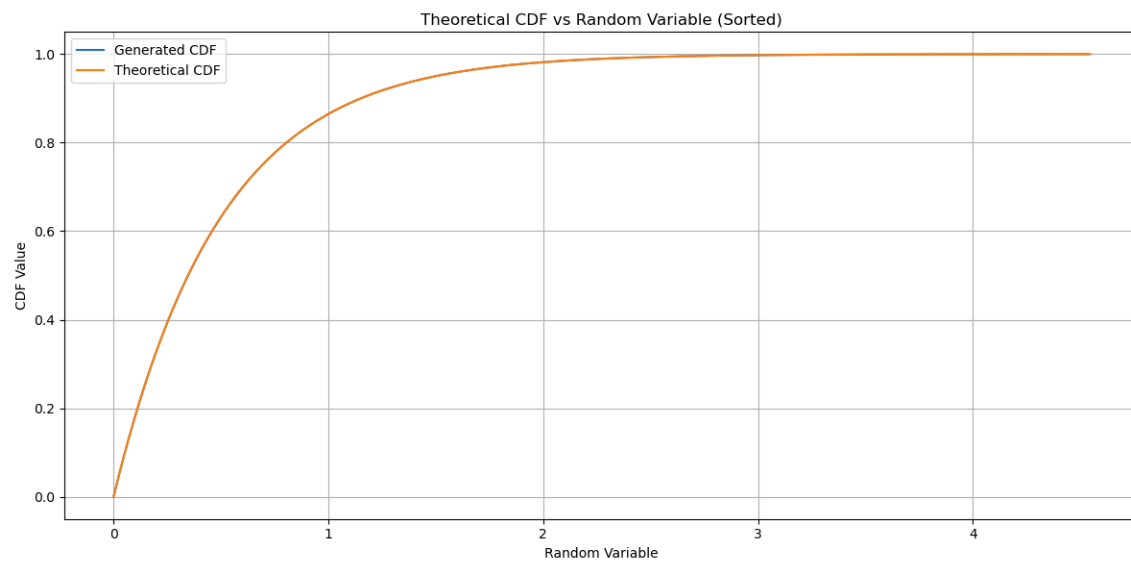


Figure 2.12: Theoretical vs Simulation Analysis wrt  $X_2$

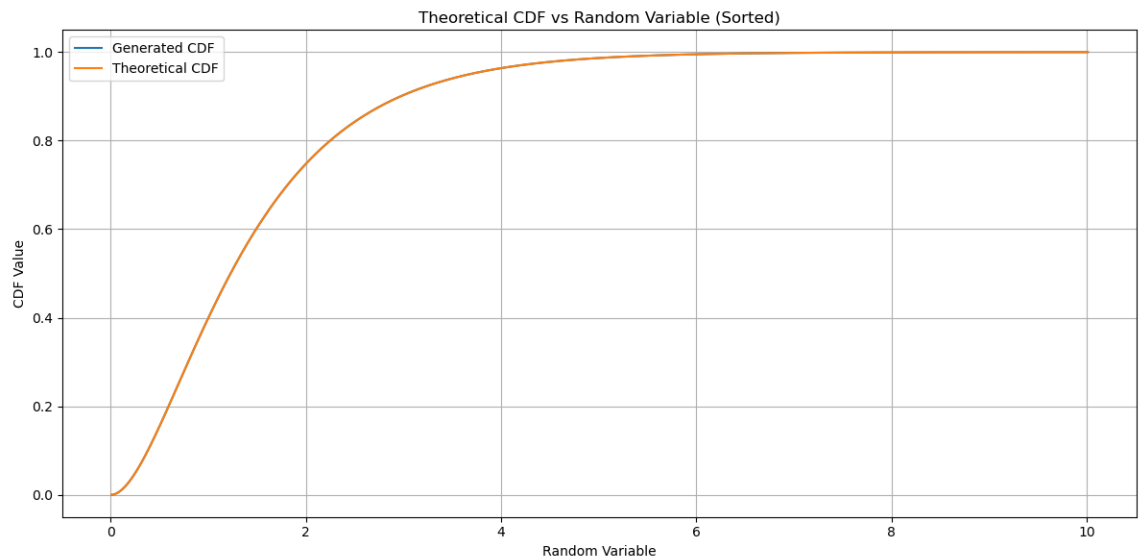


Figure 2.13: Theoretical vs Simulation Analysis wrt  $X$

2.16 Suppose that  $X$  is a discrete random variable with the following probability mass

$$P(X = 0) = \frac{1}{2} (1 + e^{-1}) \quad (2.333)$$

$$P(X = k) = \frac{e^{-1}}{2k!} \text{ for } k = 1, 2, 3, \dots \quad (2.334)$$

Which of the following is/are true?

- (a)  $E(X) = 1$
- (b)  $E(X) < 1$
- (c)  $E(X|X > 0) < \frac{1}{2}$
- (d)  $E(X|X > 0) > \frac{1}{2}$

(GATE ST 2023)

**Solution:**



(a) **Theory:**

i. As we know,

$$E(X) = \sum k p_X(k) \quad (2.335)$$

Therefore,

$$E(X) = 0 \cdot \frac{1}{2} (1 + e^{-1}) + \sum_{k=1}^{\infty} \frac{k e^{-1}}{2k!} \quad (2.336)$$

$$= \sum_{k=1}^{\infty} \frac{e^{-1}}{2(k-1)!} \quad (2.337)$$

$$= \frac{1}{2e} \sum_{k=1}^{\infty} \frac{1}{(k-1)!} \quad (2.338)$$

$$= \frac{1}{2e} \cdot e \quad (\text{Using standard result of exponential series})$$

$$= \frac{1}{2} \quad (2.339)$$

ii. To find  $E(X|X > 0)$ , first we need to find  $\Pr(X|X > 0)$  which can be given as:

$$\Pr(X|X > 0) = \frac{\Pr(X = k)}{1 - \Pr(X = 0)} \quad (2.340)$$

$$= \frac{e^{-1}}{2k!} \cdot \frac{1}{(1 - \frac{1}{2}(1 + e^{-1}))} \quad (2.341)$$

$$= \frac{e^{-1}}{2k!} \cdot \frac{2}{(1 - e^{-1})} \quad (2.342)$$

$$= \frac{e^{-1}}{k!(1 - e^{-1})} \quad (2.343)$$

$$= \frac{1}{k!(e - 1)} \quad (2.344)$$

Therefore,

$$E(X|X > 0) = \sum_{k=1}^{\infty} k \frac{1}{k!(e-1)} \quad (2.345)$$

$$= \frac{1}{e-1} \sum_{k=1}^{\infty} \frac{1}{(k-1)!} \quad (2.346)$$

$$= \frac{1}{e-1} \cdot e \quad (\text{Using standard result of exponential series})$$

$$= 1.582 \quad (2.347)$$

Referring to equations (2.339) and (2.347), we get that option (2) and (4) are correct.

(b) **Simulation:**

To make the simulation of the given question, generate a large set of random variables, say  $X$ , with the probability:

$$P(X = 0) = \frac{1}{2} (1 + e^{-1}) \quad (2.348)$$

$$P(X = k) = \frac{e^{-1}}{2k!} \text{ for } k = 1, 2, 3, \dots \quad (2.349)$$

In the simulation process, we use the concept of Inverse Transform Sampling. This involves generating a uniform random variable  $U$  from the range  $[0, 1]$  and then inverting the generalized form of CDF, i.e.  $F_X$ , to obtain  $X$ . The inversion is done using the formula:

$$X = F_X^{-1}(U) \quad (2.350)$$

Since, this distribution is discrete, we use following method to find the discrete

random variable:

$$\sum_{j=0}^{k-1} p_X(j) \leq U < \sum_{j=0}^k p_X(j) \quad (2.351)$$

In simpler terms, you're using  $U$  to navigate through the distribution's cumulative probabilities to find the corresponding value of  $X$ . Using this method, random number is compared to cumulative probabilities (CDF) for various values of the random variable  $k$  until the CDF exceeds it. This process continues until a match is found and  $k$  is returned as the generated random variable.

- i. To get  $E(X)$ , take the weighted sum of all possible values of  $X$  including 0, each multiplied by its respective probability. Expression for this can be given as:

$$E(X) = \sum k p_X(k) \quad (2.352)$$

- ii. As given in theory, to get  $E(X|X > 0)$ , take the weighted sum of all possible values of  $X$  excluding 0, each multiplied by its respective probability and dividing it by the probability of getting random variables greater than 0. Expression for this can be given as:

$$\Pr(X|X > 0) = \frac{\Pr(X = k)}{1 - \Pr(X = 0)} \quad (2.353)$$

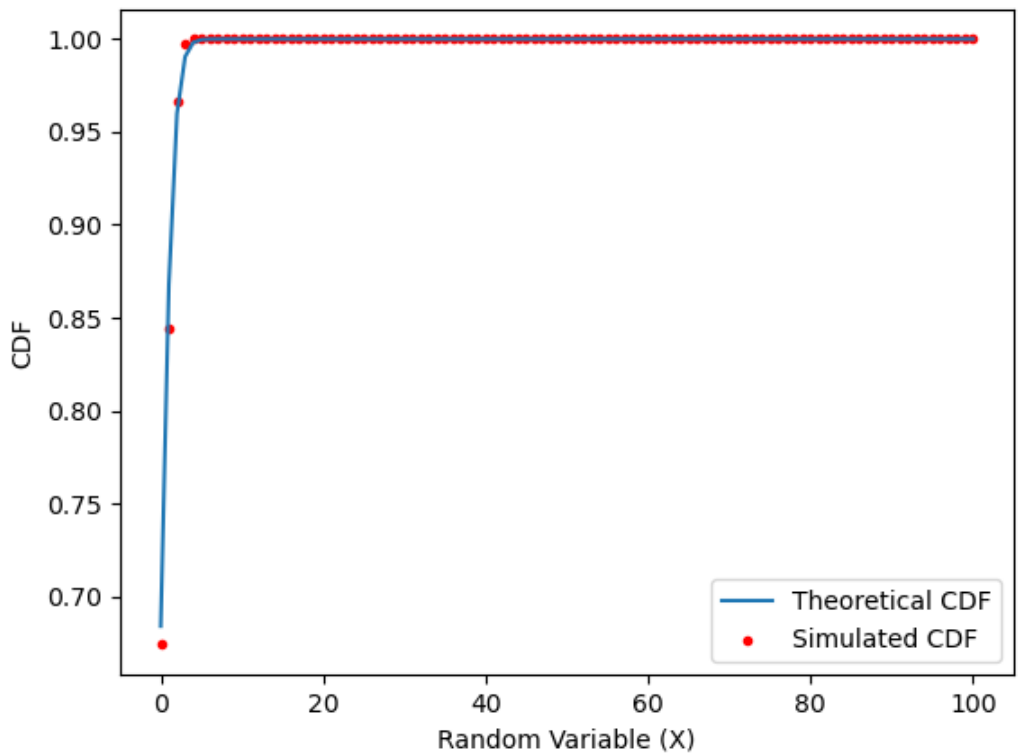


Figure 2.14: plot of corcumcirclr O and points A, B and C.

2.17 Let  $(X, Y)$  have joint probability mass function

$$p_{XY}(x, y) = \begin{cases} \frac{e^{-2}}{x!(y-x)!} & \text{if } x = 0, 1, \dots, y; y = 0, 1, 2, \dots \\ 0 & \text{otherwise.} \end{cases} \quad (2.355)$$

Then which of the following statements is/are true?

- (a)  $E(X|Y = 4) = 2$
- (b) The moment generating function of  $Y$  is  $e^{2(e^v-1)}$  for all  $v \in R$

(c)  $E(X) = 2$

(d) The joint moment generating function of  $(X, Y)$  is  $e^{-2+(1+e^u)e^v}$  for all  $(u, v) \in R^2$

(GATE ST 2023)

**Solution:** Given

$$p_{XY}(x, y) = \frac{e^{-2}}{x!(y-x)!} \quad (2.356)$$

Then,

$$p_Y(y) = \sum_{x=0}^{\infty} p_{XY}(x, y) dx \quad (2.357)$$

$$= \sum_{x=0}^{\infty} \frac{e^{-2}}{x!(y-x)!} \quad (2.358)$$

$$= \frac{e^{-2} 2^y}{y!} \quad (2.359)$$

$$(2.360)$$

(a)  $E(X|Y = 4) = 2$ :

$$E(X|Y = 4) = \sum_x x \frac{p_{XY}(x, y)}{p_Y(y)} \quad (2.361)$$

$$= \sum_{x=0}^4 x \frac{p_{XY}(x, 4)}{p_Y(4)} \quad (2.362)$$

$$= \sum_{x=0}^4 x \frac{4!}{2^4 x! (4-x)!} \quad (2.363)$$

$$= \sum_{x=0}^4 x \frac{24}{16 x! (4-x)!} \quad (2.364)$$

$$= \sum_{x=0}^4 x \frac{3}{2 x! (4-x)!} \quad (2.365)$$

$$= \frac{3}{2} \left( 0 + \frac{1}{6} + \frac{1}{2} + \frac{1}{2} + \frac{1}{6} \right) \quad (2.366)$$

$$= \frac{3}{2} \times \frac{4}{3} \quad (2.367)$$

$$= 2 \quad (2.368)$$

Hence, the given statement is true.

(b) The moment generating function of  $Y$  is  $e^{2(e^v-1)}$  for all  $v \in R$ :

The  $z$ -transform of  $Y$  is defined as:

$$M_Y(z) = E[z^{-Y}] \quad (2.369)$$

$$= \sum_{k=0}^{\infty} p_Y(k) z^{-k} \quad (2.370)$$

$$= \sum_{k=0}^{\infty} \frac{e^{-2} 2^k}{k!} z^{-k} \quad (2.371)$$

$$= e^{-2} \sum_{k=0}^{\infty} \frac{z^{-k} 2^k}{k!} \quad (2.372)$$

$$= e^{-2} \sum_{k=0}^{\infty} \frac{\left(\frac{2}{z}\right)^k}{k!} \quad (2.373)$$

$$= e^{-2+\frac{2}{z}} \quad (2.374)$$

For moment generating function,  $z$  is replaced by  $e^{-z}$ , we get

$$M_Y(z) = e^{-2+\frac{2}{e^{-z}}} \quad (2.375)$$

$$= e^{-2+2e^z} \quad (2.376)$$

$$= e^{2(e^z-1)} \quad (2.377)$$

$$\implies M_Y(v) = e^{2(e^v-1)} \quad (2.378)$$

Hence, the given statement is true.

(c)  $E(X) = 2$ :

$$p_X(x) = \sum_{y=x}^{\infty} p_{XY}(x, y) dy \quad (2.379)$$

$$= \sum_{y=x}^{\infty} \frac{e^{-2}}{x!(y-x)!} \quad (2.380)$$

$$= \frac{e^{-2}}{x!} \sum_{y=x}^{\infty} \frac{1}{(y-x)!} \quad (2.381)$$

$$= \frac{e^{-2}e}{x!} \quad (2.382)$$

$$= \frac{e^{-1}}{x!} \quad (2.383)$$

Moment generating function is defined as:

$$M_X(t) = E(e^{tx}) \quad (2.384)$$

$$= \sum_{k=0}^{\infty} p_X(k) e^{tk} \quad (2.385)$$

In  $z$ -transform we replace  $e^t$  with  $z^{-1}$ .



$$M_X(\nu) = E(\nu^{-x}) \quad (2.386)$$

$$= \sum_{k=0}^{\infty} p_X(k) \nu^{-k} \quad (2.387)$$

$$= \sum_{k=0}^{\infty} \frac{e^{-1}}{k!} \nu^{-k} \quad (2.388)$$

$$= e^{-1} \sum_{k=0}^{\infty} \frac{\nu^{-k}}{k!} \quad (2.389)$$

$$= e^{-1+\nu^{-1}} \quad (2.390)$$

$$\implies M_X(\nu^{-1}) = e^{-1+\nu} \quad (2.391)$$

Now,

$$E[X] = \frac{dM_X(\nu^{-1})}{dz} \Big|_{\nu=1} \quad (2.392)$$

$$= \frac{de^{-1+\nu}}{d\nu} \Big|_{\nu=1} \quad (2.393)$$

$$= e^{-1+\nu} \Big|_{\nu=1} \quad (2.394)$$

$$= e^0 \quad (2.395)$$

$$= 1 \quad (2.396)$$

Therefore,  $E(X) = 1$ .

Hence, the given statement is wrong.

(d) The joint moment generating function of  $(X, Y)$  is  $e^{-2+(1+e^u)e^v}$  for all  $(u, v) \in R^2$ :

$$M_{XY}(z_1, z_2) = E(e^{z_1x+z_2y}) \quad (2.397)$$

$$= \sum_{x=0}^{\infty} \sum_{y=x}^{\infty} e^{z_1x+z_2y} \frac{e^{-2}}{x!(y-x)!} \quad (2.398)$$

$$= e^{-2} \sum_{x=0}^{\infty} \frac{e^{z_1x}}{x!} \sum_{y=x}^{\infty} \frac{e^{z_2y}}{(y-x)!} \quad (2.399)$$

$$= e^{-2} e^{e^{z_2}} \sum_{x=0}^{\infty} \frac{e^{(z_1+z_2)x}}{x!} \quad (2.400)$$

$$= e^{-2+e^{z_2}} e^{e^{z_1}+z_2} \quad (2.401)$$

$$= e^{-2+(1+e^{z_1})e^{z_2}} \quad (2.402)$$

$$\implies M_{XY}(u, v) = E(e^{ux+vy}) \quad (2.403)$$

$$= e^{-2+(1+e^u)e^v} \quad (2.404)$$

Hence, the given statement is true.

Therefore, 1,2,4 statements are true.

### Simulation:

To make the simulation of the given question, generate a large set of random variables, say  $X$  and  $Y$ , with the probability:

$$P(X = k) = \frac{e^{-1}}{k!} \text{ for } k = 0, 1, 2, 3, \dots \quad (2.405)$$

$$P(Y = k) = \frac{e^{-2}2^k}{k!} \text{ for } k = 0, 1, 2, 3, \dots \quad (2.406)$$

In the simulation process, we use the concept of Inverse Transform Sampling. This involves generating a uniform random variable  $U$  from the range  $[0, 1]$  and then inverting the generalized form of CDF, i.e.  $F_X$ , to obtain  $X$  and  $F_Y$ , to obtain  $Y$ . The

inversion is done using the formula:

$$X = F_X^{-1}(U) \quad (2.407)$$

$$Y = F_Y^{-1}(U) \quad (2.408)$$

CDF of  $X$  is given as

$$F_X(x) = \sum_{i=0}^x p_X(i) \quad (2.409)$$

CDF of  $Y$  is given as

$$F_Y(y) = \sum_{i=0}^y p_Y(i) \quad (2.410)$$

Since, this distribution is discrete, we use following method to find the discrete random variable:

$$\sum_{i=0}^{k-1} p_X(i) \leq U < \sum_{i=0}^k p_X(i) \quad (2.411)$$

$$\sum_{i=0}^{k-1} p_Y(i) \leq U < \sum_{i=0}^k p_Y(i) \quad (2.412)$$

In simpler terms, you're using  $U$  to navigate through the distribution's cumulative probabilities to find the corresponding value of  $X$  and  $Y$ . Using this method, random number is compared to cumulative probabilities (CDF) for various values of the random variable  $k$  until the CDF exceeds it. This process continues until a match is found and  $k$  is returned as the generated random variable.

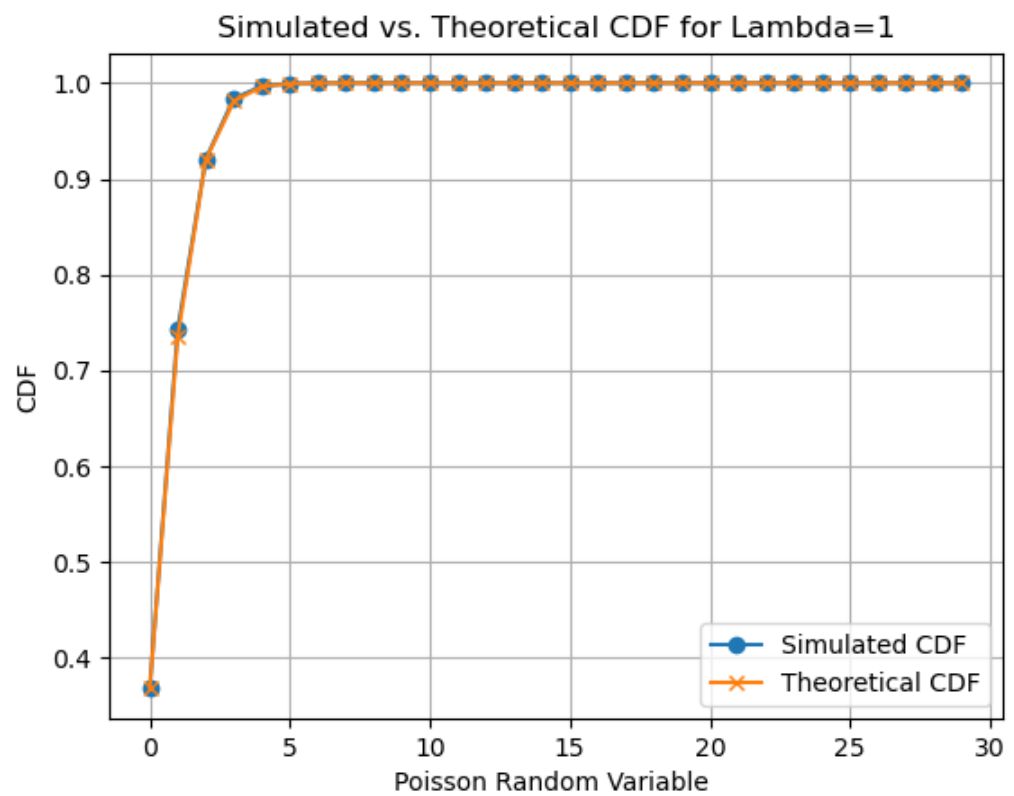


Figure 2.15: CDF of X

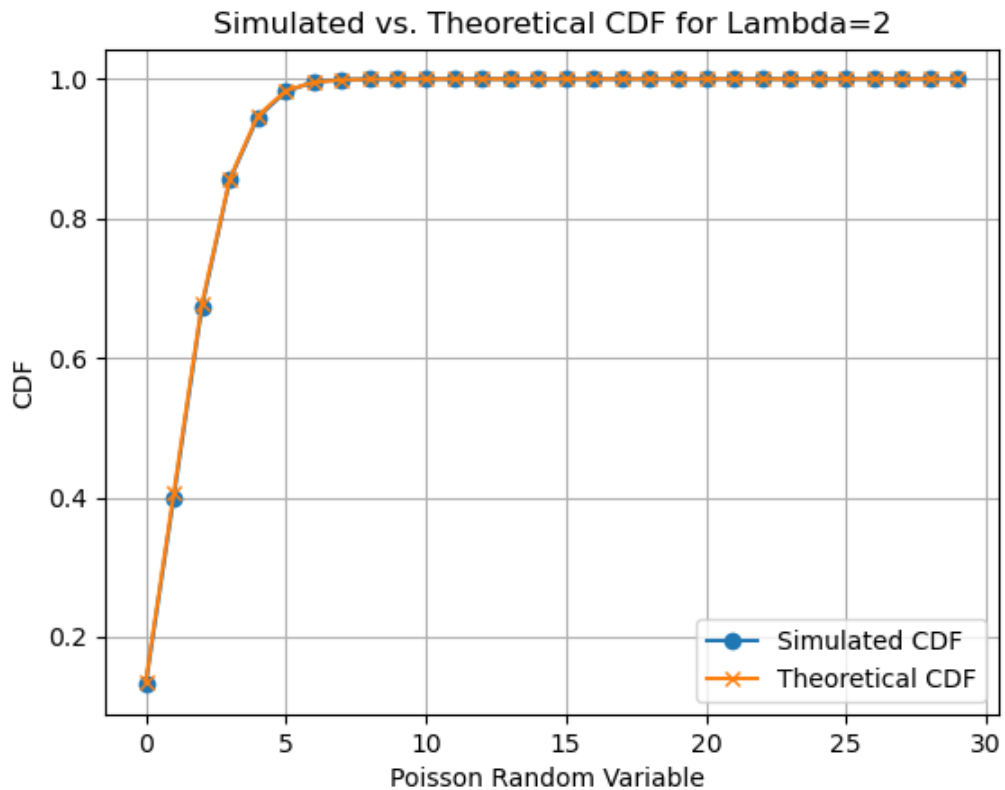


Figure 2.16: CDF of Y

2.18 Out of 1000 individuals in a town, 100 unidentified individuals are covid positive. Due to lack of adequate covid-testing kits, the health authorities of the town devised a strategy to identify these covid-positive individuals. The strategy is to:

- (a) Collect saliva samples from all 1000 individuals and randomly group them into sets of 5.
- (b) Mix the samples within each set and test the mixed sample for covid.
- (c) If the test done in (ii) gives a negative result, then declare all the 5 individuals to be covid negative.

- (d) If the test done in (ii) gives a positive result, then all the 5 individuals are separately tested for covid.

Given this strategy, no more than testing kits will be required to identify all the 100 covid positive individuals irrespective of how they are grouped

- (a) 700  
(b) 600  
(c) 800  
(d) 1000

(GATE BM 2023)

**Solution :**

Given :

Parameter	Value
Number of individuals	1000
Strength of each group	5
Number of groups	200
Number of Covid positive individuals	100

At first we ll have to examine all the groups and then test all 5 individuals of each group who tested positive.

Using binomial expansion to find the probability of a group testing positive

Parameter	Value	Expression
N	1000	Total number of individuals
n	5	Number of individuals in group
p	0.1	Probability of individual testing positive
q	0.9	Probability of individual testing negative

Probability of a group of 5 with i covid negative individuals is given as the PMF:

$$Pr(X = i) = {}^n C_i (q)^i (p)^{n-i} \quad (2.413)$$

$$= {}^5 C_i (0.9)^i (0.1)^{5-i} \quad (2.414)$$

Let us assume X is a random variable which denotes the number of individuals tested negative in a group. Probability of group having atmost k negative tested individuals the CDF :

$$F_X(k) = Pr(X \leq k) \quad (2.415)$$

$$= \sum_{i=0}^k ({}^n C_i) (q)^i (p)^{n-i} \quad (2.416)$$

$$= \sum_{i=0}^k ({}^5 C_i) (0.9)^i (0.1)^{5-i} \quad (2.417)$$

$$(2.418)$$

If a group is tested positive, atleast one of the 5 individuals should be covid positive which means atmost 4 individuals are covid negative.

Finding probability of a group testing positive

$$Pr = F_X(4) \quad (2.419)$$

$$= \sum_{i=0}^4 \binom{5}{i} (0.9)^i (0.1)^{5-i} \quad (2.420)$$

$$= 0.4096 \quad (2.421)$$

Parameter	Value
Number of tests in first round of testing	A
Number of groups tested positive in first round	b
Number of tests in second round of testing	B
Total number of tests in both rounds	T

In the first round we will test all the 200 groups.

Number of tests in first round(A) :

$$A = 200$$

Out of these 200 tests, assuming the estimated number of groups tested positive is b  
 If a group is tested positive, atleast one of the 5 is covid positive, which implies the group can only have atmost 4 covid negative individual.

Calculating b :

$$\Pr(\text{of group testing positive}) = \frac{b}{A} \quad (2.422)$$

$$F_X(4) = \frac{b}{A} \quad (2.423)$$

$$\implies b = 81.92 \quad (2.424)$$



In all the groups tested positive, atleast one individual is covid positive, to find all the covid positive individuals, we will test all the 5 individuals of the positive groups seperately in round 2.

So estimated number tests in round 2 (B) :

$$B = 5b \quad (2.425)$$

$$= (5) (81.92) \quad (2.426)$$

$$= 409.6 \quad (2.427)$$

Total number of tests estimated (T) :

$$T = A + B \quad (2.428)$$

$$= 200 + 409.6 \quad (2.429)$$

$$= 609.6 \quad (2.430)$$

∴ So 609.6 is the expected number of tests required to detect all the covid positive individuals.

According to the options, no more than 700 tests will be required.

**So, the correct answer is option A 700.**

**Simulation :**

- (a) Including standard C libraries - `stdio.h` (for input/output), `stdlib.h` (for dynamic memory allocation), `time.h` (for time-related functions), and `math.h` (for mathematical functions).

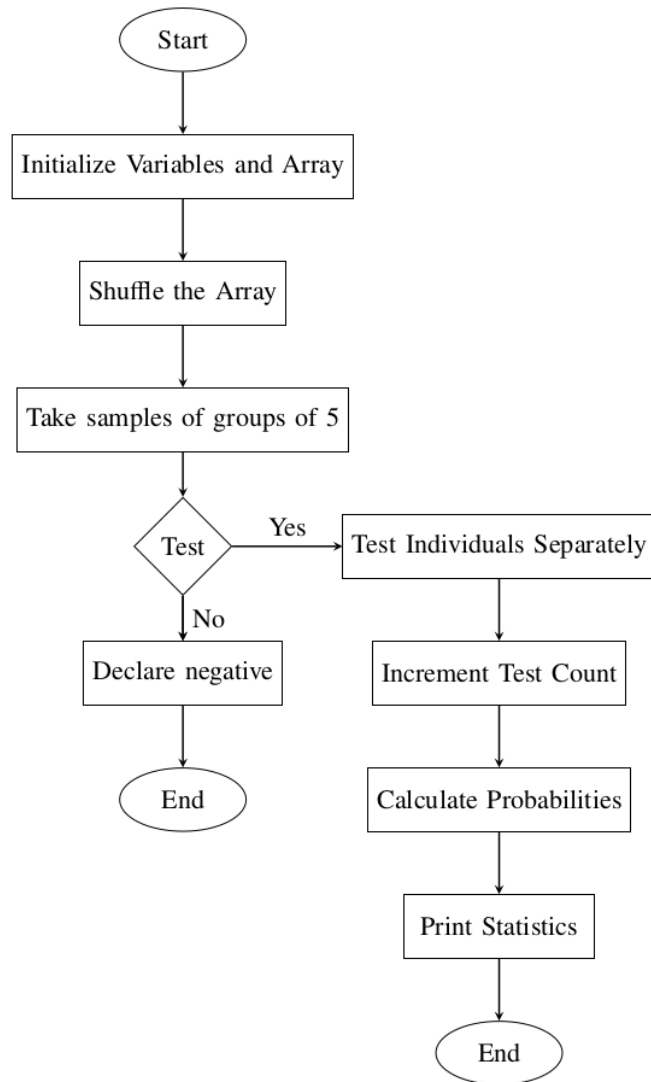
- (b) Defining constants for the total number of individuals (`TOTAL_INDIVIDUALS`), the number of initially positive individuals (`POSITIVE_INDIVIDUALS`), and the group size used in testing (`GROUP_SIZE`).
- (c) Declared a function named `covidTestingStrategy` that simulates a COVID-19 testing strategy using group testing. The function takes an array of individuals, the total number of individuals, and the number of initially positive individuals as parameters.
- (d) Initialized variables `testCount` and `groupTestTrueCount` to keep track of the total number of tests and the number of group tests that turned out to be true and shuffled the array of individuals randomly using the Fisher-Yates shuffle algorithm to have randomness into the order of testing.
- (e) We mixed the samples within each group using Bitwise "OR" and tested them.
  - i. If the mixed sample is negative (`groupResult == 0`), declares all 5 individuals in the group as COVID negative with one test.
  - ii. If the mixed sample is positive, tests each individual separately, and increments the test count accordingly.
- (f) Calculated the probability of the group test being true based on the initial probability of an individual being positive and the group size and Calculated the expected number of tests for one set, total expected number of sets, and total

expected number of tests based on the calculated probabilities.

(g) Printed various statistics related to the testing strategy.

(h) Seed the random number generator with the current time. Initialized an array representing individuals, marking the first `POSITIVE_INDIVIDUALS` as positive. Calls the `covidTestingStrategy` function with the array and other parameters.

Flowchart to explain the function `covidTestingStrategy` is given below :



```

snehil@snehil-ASUS-TUF-Gaming-A15-FA506IHRB-FA506IHRB:~/gate/CODE$ ./main
Total tests required: 222
Group test true count: 78
Probability of group test being true: 0.4095
Expected number of tests for one set: 3.05
Total expected number of sets: 200.00
Total expected number of tests: 609.51

```

Figure 2.17: output



## Chapter 3

# Conditional Probability

1. Let  $X$  be a random variable with probability density function

$$f(x) = \begin{cases} \frac{1}{x^2} & x \geq 1 \\ 0 & \text{otherwise.} \end{cases} \quad (3.1)$$

If  $Y = \log_e X$ , then  $\Pr(Y < 1 | Y < 2)$  equals (GATE ST 2023)

**Solution:** Given, the probability density function of  $X$  is

$$f(x) = \begin{cases} \frac{1}{x^2} & x \geq 1 \\ 0 & \text{otherwise.} \end{cases} \quad (3.2)$$

Also,  $Y = \log_e X$ .

Consider the cumulative distribution function(CDF) of  $X$ ,

$$F_X(x) = \Pr(X \leq x) \quad (3.3)$$

$$= \int_1^x \frac{1}{x^2} dx \quad (3.4)$$

$$= 1 - \frac{1}{x}, x \geq 1 \quad (3.5)$$

Now, we need to find the CDF of  $Y$ .

$$F_Y(y) = \Pr(Y \leq y) \quad (3.6)$$

$$= \Pr(\log_e X \leq y) \quad (3.7)$$

$$= \Pr(X \leq e^y) \quad (3.8)$$

$$= F_X(e^y) \quad (3.9)$$

$$= 1 - \frac{1}{e^y}, y \geq 0 \quad (3.10)$$

Now, we need to find  $\Pr(Y < 1 | Y < 2)$ . For that, we need to find  $F_Y(1)$  and  $F_Y(2)$ .

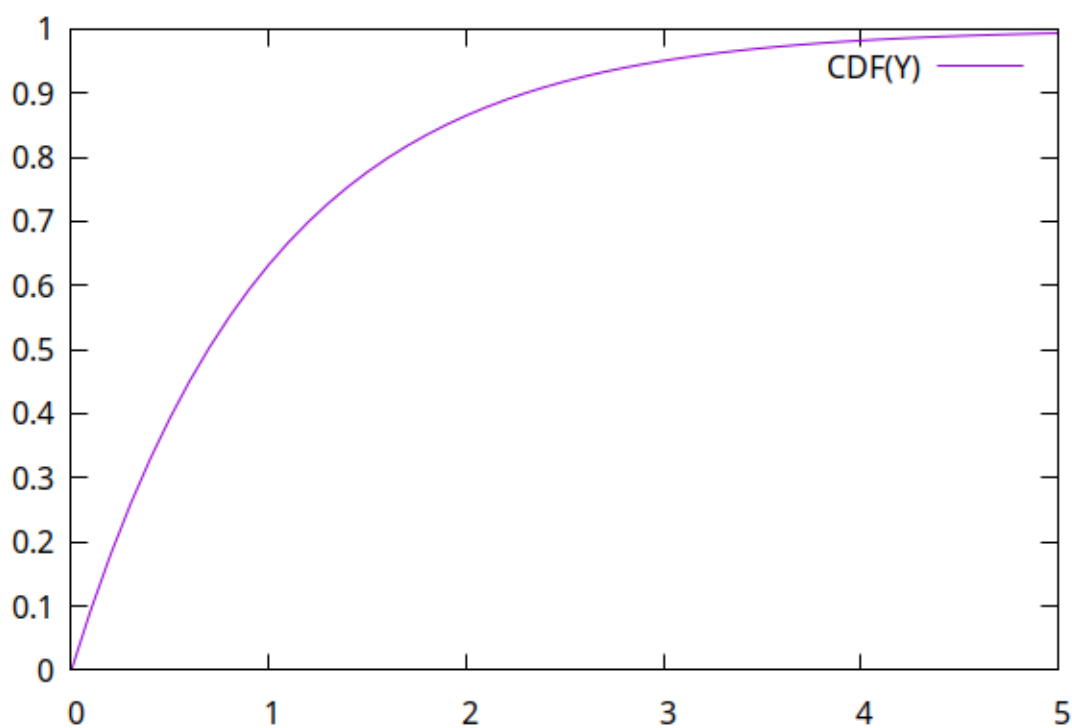


Figure 3.1: CDF of  $Y$

Using the equation for CDF,

$$F_Y(1) = 1 - \frac{1}{e} \quad (3.11)$$

and

$$F_Y(2) = 1 - \frac{1}{e^2} \quad (3.12)$$

Now, we can find  $\Pr(Y < 1|Y < 2)$  as follows,

$$\Pr(Y < 1|Y < 2) = \frac{\Pr(Y < 1, Y < 2)}{\Pr(Y < 2)} \quad (3.13)$$

$$= \frac{\Pr(Y < 1)}{\Pr(Y < 2)} \quad (3.14)$$

$$= \frac{F_Y(1)}{F_Y(2)} \quad (3.15)$$

$$= \frac{1 - \frac{1}{e}}{1 - \frac{1}{e^2}} \quad (3.16)$$

$$= \frac{e(e-1)}{e^2-1} \quad (3.17)$$

$$= \frac{e}{e+1} \quad (3.18)$$

**Steps for simulation.**

(a) Generate a uniform random variable between 0 and 1.

(b) Given the PDF of  $X$  as  $\frac{1}{x^2}$ , use the inverse transform method to generate  $X$ .

The inverse transform method is given by,

$$X = F^{-1}(U) \quad (3.19)$$



where  $U$  is a uniform random variable between 0 and 1 and  $F^{-1}$  is the inverse of CDF of  $X$ .

(c) The CDF of  $X$  is given by,

$$F(x) = \int_1^x \frac{1}{x^2} dx = 1 - \frac{1}{x} \quad (3.20)$$

The inverse of CDF of  $X$  is given by,

$$F^{-1}(x) = \frac{1}{1-x} \quad (3.21)$$

(d) Use  $X$  to generate  $Y = \log_e X$ .

(e) Calculate the CDF of  $Y$  using the equation  $1 - \frac{1}{e^y}$ .

(f) Store the values of CDF in data file.

(g) Plot the CDF using GNUPlot.

2. The probability of a person telling the truth is  $4/6$ . An unbiased die is thrown by the same person twice and the person reports that the numbers appeared in both the throws are same. Then the probability that actually the numbers appeared in both the throws are same is ? (GATE XE 2023)

**Solution:** Random variables on  $i \in \{1, 2\}$  defined as

Random Variable	Values	Description
$X_i$	$1 \leq X_i \leq 6$	number appeared on $i$ th die
$Y$	$\{0, 1\}$	person telling the truth or lie

$p_Y(0)$  = Probability that person telling the lie

$p_Y(1)$  = Probability that person telling the truth

Consider

$$Z = X_1 - X_2 \quad (3.22)$$

$Z$	0	$\neq 0$
$\Pr(Z)$	$\frac{1}{6}$	$\frac{5}{6}$

$$p_Y(i) = \begin{cases} \frac{2}{3} & \text{if } i = 1 \\ \frac{1}{3} & \text{if } i = 0 \\ 0 & \text{otherwise} \end{cases} \quad (3.23)$$

Given in the question that the person reports that the numbers appeared in both the throws are same. the probability that actually the numbers appeared in both the throws are same, that is simply, the probability of person's truth given that numbers on both dice are same that is,  $\Pr(Y = 1/Z = 0)$ .

$$\Pr(Y = 1/Z = 0) = \frac{\Pr((Z = 0).(Y = 1))}{\Pr(Z = 0)} \quad (3.24)$$

Since  $X_i$  and  $Y$  are independent events

$$\Pr(Y = 1/Z = 0) = p_Y(1) \quad (3.25)$$

$$= \frac{2}{3} \quad (3.26)$$

$$\approx 0.667 \quad (3.27)$$



# Chapter 4

## Random Variable

4.1 A cytoplasmic male-sterile female plant with the restorer (nuclear) genotype  $rr$  is crossed to a male-fertile male plant with the genotype  $RR$ . Both  $RR$  and  $Rr$  can restore the fertility, whereas  $rr$  cannot. When an  $F_1$  female plant with  $Rr$  genotype was test-crossed to a male-fertile male plant with the  $rr$  genotype, the percentage of the population that is male fertile would be? (GATE XL 2023)

**Solution:**  
Representing  $R$  and  $r$  as follows:

Gene	represent
R	1
r	0

Table 4.2: Table3:  $R=1,r=0$

On crossing between 00 and 11 we get:

	1	1
0	10	10
0	10	10

Table 4.4: Table1: Crossing btw  $RR$  and  $rr$

Which gives  $F_1$  as:

$$F_1 = \{10, 10, 10, 10\} \quad (4.1)$$

When  $F_1$  (10) is test-crossed with (00) we get:

	0	0
1	10	10
0	00	00

Table 4.6: Table2: Crossing btw 10 and 00

$$F_2 = \{10, 10, 00, 00\} \quad (4.2)$$

Probability that the population is male fertile(10) from (4.2) is given by:

$$\Pr(10) = \frac{1}{2} \quad (4.3)$$

∴ The percentage of the population that is male fertile would be 50%

4.2 Given a fair six-faced dice where the faces are labelled '1','2','3','4','5', and '6'. what is the probability of getting a '1' on the first roll of the dice and a '4' on the second roll ?  
(GATE XH 2023)

**Solution:** Let  $X_1$  and  $X_2$  be an bernoulli rv's defined as,

The probabbility follows:

Table 4.7: Declaration of rv's

Parameter	value	Description
$X_1$	1	getting 1 in 1st throw
	0	not getting 1 in 1st throw
$X_2$	1	getting 4 in 2nd throw
	0	not getting 4 in 2nd throw

$$p_{X_1}(k) = \begin{cases} \frac{1}{6}, & k = 1 \\ \frac{5}{6}, & k = 0 \end{cases} \quad (4.4)$$

$$p_{X_2}(k) = \begin{cases} \frac{1}{6}, & k = 1 \\ \frac{5}{6}, & k = 0 \end{cases} \quad (4.5)$$

Now,

$$\Pr(X_1 = 1, X_2 = 1) = \Pr(X_1 = 1) \Pr(X_2 = 1) \quad (4.6)$$

$$= \Pr(1) \Pr(1) \quad (4.7)$$

$$= \frac{1}{6} \cdot \frac{1}{6} \quad (4.8)$$

$$= \frac{1}{36} \quad (4.9)$$

$$= 0.028 \quad (4.10)$$

Hence, probability of getting a '1' on the first roll of the dice and a '4' on the second roll is 0.028

## Chapter 5

# Moments

5.1 Suppose that  $X$  has the probability density function

$$f(x) = \begin{cases} \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} & \lambda > 0 \\ 0 & \text{otherwise} \end{cases} \quad (5.1)$$

where  $\alpha > 0$  and  $\lambda > 0$ . Which one of the following statements is NOT true?

- (a)  $E(X)$  exists for all  $\alpha > 0$  and  $\lambda > 0$
- (b) Variance of  $X$  exists for all  $\alpha > 0$  and  $\lambda > 0$
- (c)  $E(\frac{1}{X})$  exists for all  $\alpha > 0$  and  $\lambda > 0$
- (d)  $E(\ln(1 + X))$  exists for all  $\alpha > 0$  and  $\lambda > 0$

(GATE ST 2023)

**Solution:**



(a)

$$E(X) = \int_{-\infty}^{\infty} xp_X(x)dx \quad (5.2)$$

$$= \int_0^{\infty} x \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} \quad (5.3)$$

$$= \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^{\infty} x^\alpha e^{-\lambda x} \quad (5.4)$$

$$(5.5)$$

since we know that

$$\int_0^{\infty} x^{\alpha-1} e^{-\lambda x} dx = \frac{\Gamma(\alpha)}{\lambda^\alpha} \quad \text{for } \lambda > 0, \alpha > 0 \quad (5.6)$$

$$E(X) = \frac{\lambda^\alpha}{\Gamma(\alpha)} \frac{\Gamma(\alpha+1)}{\lambda^{\alpha+1}} \quad (5.7)$$

Using the relation

$$\Gamma(x+1) = \Gamma(x)x \quad (5.8)$$

$$E(X) = \frac{\alpha}{\lambda} \quad (5.9)$$

Thus  $E(X)$  exists for all  $\alpha > 0$  and  $\lambda > 0$ .

(b)

$$Var(X) = E(X^2) - E(X)^2 \quad (5.10)$$

$$E(X^2) = \int_0^{\infty} x^2 \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} dx \quad (5.11)$$

$$= \int_0^{\infty} \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{(\alpha+2)-1} e^{-\lambda x} dx \quad (5.12)$$

$$= \int_0^{\infty} \frac{1}{\lambda^2} \frac{\lambda^{\alpha+2}}{\Gamma(\alpha)} x^{(\alpha+2)-1} e^{-\lambda x} dx \quad (5.13)$$

$$E(X^2) = \int_0^{\infty} \frac{\alpha(\alpha+1)}{\lambda^2} \frac{\lambda^{\alpha+2}}{\Gamma(\alpha+2)} x^{(\alpha+2)-1} e^{-\lambda x} dx \quad (5.14)$$

using the density of the gamma distribution, we get

$$E(X^2) = \frac{\alpha(\alpha+1)}{\lambda^2} \quad (5.15)$$

$$Var(X) = \frac{\alpha^2 + \alpha}{\lambda^2} - \frac{\alpha^2}{\lambda} \quad (5.16)$$

$$= \frac{\alpha}{\lambda^2} \quad (5.17)$$

Thus, Variance of  $X$  exists for all  $\alpha > 0$  and  $\lambda > 0$

(c)

$$E\left(\frac{1}{X}\right) = \int_0^{\infty} \frac{1}{x} \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} dx \quad (5.18)$$

$$= \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^{\infty} x^{\alpha-2} e^{-\lambda x} dx \quad (5.19)$$

For this,  $\alpha > 1$  is a must condition. Hence C is not a correct option. Hence C is the answer.

(d) For  $E(\ln(1 + X))$ ,

$$E(\ln(1 + X)) = E(X) - \frac{E(X^2)}{2} + \frac{E(X^4)}{4} - .. \quad (5.20)$$

We write the general expression for  $E(X^n)$

$$E(X^n) = \frac{(\alpha)(\alpha + 1) \dots (\alpha + n - 1)}{\lambda^n} \quad (5.21)$$

So by applying the ratio test to check the convergence of the sequence

$$\lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| = L \quad (5.22)$$

$$\left| \frac{E(X^{n+2})}{E(X^n)} \right| = \frac{\frac{(\alpha)(\alpha+1)\dots(\alpha+n+1)}{\lambda^{n+2}}}{\frac{(\alpha)(\alpha+1)\dots(\alpha+n-1)}{\lambda^n}} \quad (5.23)$$

$$= \frac{(\alpha + n)(\alpha + n + 1)}{\lambda^2} \quad (5.24)$$

$$\lim_{n \rightarrow \infty} \left| \frac{E(X^{n+2})}{E(X^n)} \right| = \infty \quad (5.25)$$

Thus  $E(\ln(1 + X))$  generates a divergent function and hence  $E(\ln(1 + X))$  does not exist for all  $\alpha > 0$  and  $\lambda > 0$ .

5.2 For a Binomial random variable  $X$ ,  $E(X)$  and  $\text{Var}(X)$  are the expectation and variance, respectively. Which one of the following statements CANNOT be true? (GATE BM 2023)

**Solution:**

$$X \sim \text{Bin}(n, p)$$

S.No	E(X)	Var(X)
1	20	16
2	6	5.4
3	10	15
4	64	12.8

Table 5.1:

We know ,

$$E(X) = np \quad (5.26)$$

$$Var(X) = np(1 - p) \quad (5.27)$$

$$0 \leq p \leq 1 \quad (5.28)$$

$$\implies -1 \leq -p \leq 0 \quad (5.29)$$

$$\implies 0 \leq 1 - p \leq 1 \quad (5.30)$$

$$\implies np(1 - p) \leq np \quad (5.31)$$

Therefore,

$$Var(X) \leq E(X) \quad (5.32)$$

From the four options,the statement that cannot be true is option (3)

### Simulation steps

#### Step 1: Generate a Range of Probabilities

The program generates a range of probabilities ( $p$ ) in increments of 0.0098, ranging from 0.01 to 0.99. This range is used to create different binomial distributions for subsequent calculations.

**Step 2: Generating binomial r.v from Uniform distribution**

$U \sim \text{Uniform}(0, 1)$ .

Defining a random variable  $X$  as:

$$X = \begin{cases} 1 & \text{if } U > p \\ 0 & \text{otherwise} \end{cases}$$

This becomes a Bernoulli rv. The count variable evaluates the Binomial r.v by the summation of Bernoulli r.v

$$count = \sum_{i=1}^n X$$

**Step 3: Calculate Variances for Each Probability**

For each probability ( $p$ ) in the generated range, the program generates a binomial random variable using the given mean and the inverse of the probability ( $mean/p$ ) from uniform distribution.

**Step 4: Find Maximum and Minimum Variances**

After calculating variances for each probability, the program identifies the maximum and minimum variances in the generated set.

**Step 5: User Input**

The program prompts the user to input a variance value for validation.

**Step 6: Check Validity**

The user-input variance is compared against the computed maximum and minimum variances. If the input variance falls within this range (inclusive of the minimum and exclusive of the maximum), the program outputs "Valid." Otherwise, it outputs "Invalid."

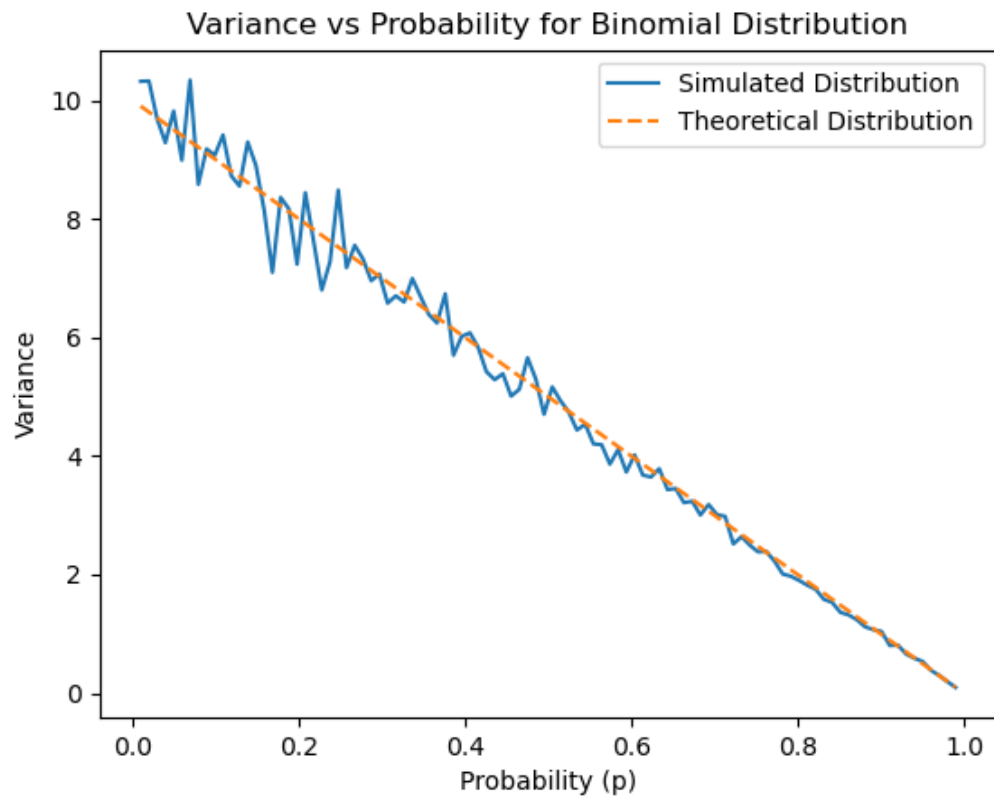


Figure 5.1: Variance for mean=10

5.3 The signal-to-noise ratio (SNR) of an ADC with a full-scale sinusoidal input is given to be 61.96 dB. The resolution of the ADC is  
(GATE EC 2023)

**Solution:**

Variable	Defination
$T$	Time Period
$f$	Frequency
$A$	Amplitude
$P_S$	Signal Power
$P_N$	Noise Power
$q$	Quantization step size
$n$	No of bits
$Y$	Quantization Error

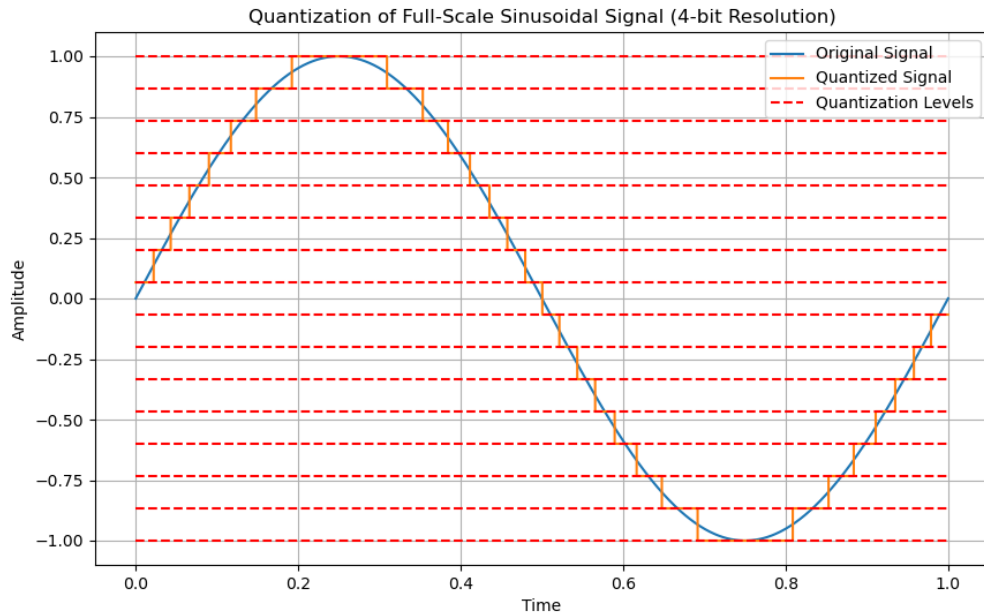


Figure 5.2: Quantization of Sinusoidal Signal

(a) Signal Power:

The power of a continuous-time signal is defined as the average value of the square of the signal over a certain time interval. For a sinusoidal signal  $x(t) =$

$A \sin(2\pi ft + \phi)$ , the power ( $P_S$ ) is calculated as:

$$P_S = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} |x(t)|^2 dt \quad (5.33)$$

where  $\phi$  is the phase of the signal.

$$P_S = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} |A \sin(2\pi ft + \phi)|^2 dt \quad (5.34)$$

$$P_S = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} A^2 \sin^2(2\pi ft + \phi) dt \quad (5.35)$$

$$P_S = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} A^2 \cdot \frac{1 - \cos(4\pi ft + 2\phi)}{2} dt \quad (5.36)$$

$$P_S = \frac{1}{2} \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} A^2 dt - \frac{1}{2} \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} A^2 \cos(4\pi ft + 2\phi) dt \quad (5.37)$$

$$P_S = \frac{1}{2} \cdot A^2 - 0 \quad (5.38)$$

$$P_S = \frac{A^2}{2} \quad (5.39)$$

$$(5.40)$$

Here,  $A=1$ , so:

$$P_S = \frac{1}{2} \quad (5.41)$$

(b) Noise Power:

No of Quantization levels is given by  $2^n$ , Where n is resolution or no of bits.

Distance between any two Quantization levels = Quantization step



(No of Quantization levels)\*(Quantization step) = Peak Distance (Refer to Fig 1)

$$q = \frac{\text{Peak distance}}{2^n} = \frac{1 - (-1)}{2^n} = \frac{2}{2^n} = 2^{-(n-1)} \quad (5.42)$$

We know, quantization error has a maximum value of plus or minus half the step size, so

$$|Y| \leq \frac{q}{2} \text{ and therefore, } |Y| \leq 2^{-n} \quad (5.43)$$

For a large enough number of quantization steps, the probability density function of the quantization error tends toward being flat 1. Pdf of error (Y) of quantization is defined as

$$p_Y(y) = \begin{cases} \frac{1}{q}, & \text{if } -\frac{q}{2} \leq y \leq \frac{q}{2} \\ 0, & \text{otherwise} \end{cases} \quad (5.44)$$

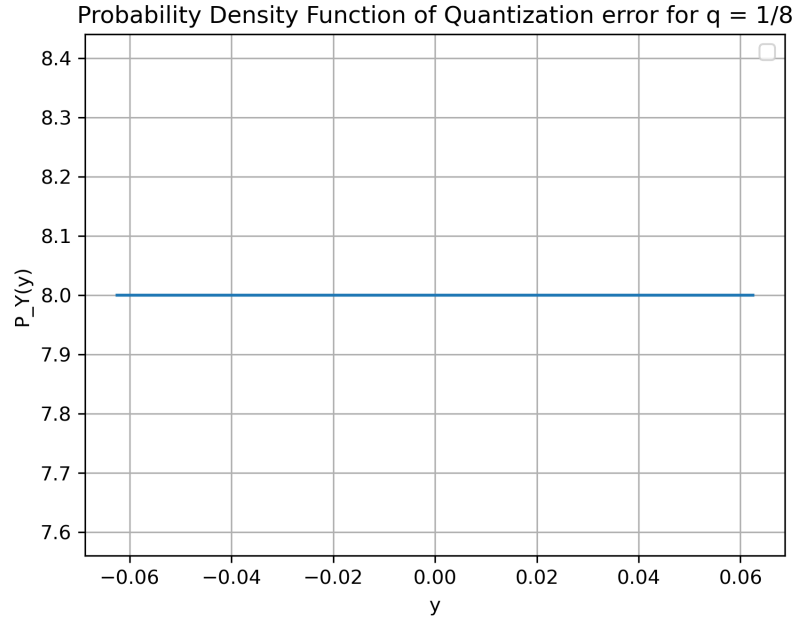


Figure 5.3: plot of pdf of Quantization Error

So, we can calculate its mean power or variance as the 2nd moment of its distribution.

Since, the distribution of error is uniform hence  $E[Y]=0$ .

$$E(Y) = \int_{-\frac{q}{2}}^{\frac{q}{2}} p_Y(y) y \, dy \quad (5.45)$$

$$= \frac{1}{q} \cdot \frac{1}{2} \left( \left( \frac{q}{2} \right)^2 - \left( -\frac{q}{2} \right)^2 \right) \quad (5.46)$$

$$= \frac{1}{q} \cdot \frac{1}{2} (0 - 0) \quad (5.47)$$

$$= 0 \quad (5.48)$$

$$E[Y^2] = \int_{-\frac{q}{2}}^{\frac{q}{2}} p_Y(y) y^2 dy \quad (5.49)$$

$$= \frac{1}{q} \cdot \frac{1}{3} \left( \left( \frac{q}{2} \right)^3 - \left( -\frac{q}{2} \right)^3 \right) \quad (5.50)$$

$$= \frac{1}{3q} \cdot \frac{q^3}{4} \quad (5.51)$$

$$= \frac{q^2}{12} \quad (5.52)$$

$$(5.53)$$

On putting  $q = 2^{-(n-1)}$ , we have:

$$E[Y^2] \approx \frac{2^{-2n}}{3} \quad (5.54)$$

$$P_Y \approx \frac{2^{-2n}}{3} \quad (5.55)$$

$$P_N \approx \frac{2^{-2n}}{3} \quad (5.56)$$

$$(5.57)$$

Thus, an ideal ADC would have a signal-to-noise ratio

$$SNR = \frac{P_S}{P_N} = 1.5 \cdot 2^{2n} \quad (5.58)$$

or, expressed in decibels,

$$SNR = 10 (\log_{10}(1.5 \cdot 2^n)) \quad (5.59)$$

$$= 10 (\log_{10}(1.5) + \log_{10}(2^{2n})) \quad (5.60)$$

$$= 10 (0.176 + 2n \cdot 0.3010) \quad (5.61)$$

$$= 1.76 + 6.02n \quad (5.62)$$

Substituting value of  $SNR$

$$61.96 = 1.76 + 6.02n \quad (5.63)$$

$$n = \frac{61.96 - 1.76}{6.02} \quad (5.64)$$

$$= \frac{60.2}{6.02} \quad (5.65)$$

$$= 10 \quad (5.66)$$

So, the resolution of the ADC is approximately 10 bits.

Simulation Steps :-

- i. Initialize the simulation parameters, including the given SNR in dB (which is 61.96 dB) and signal power ( $P_s$ ) (which is 0.5), and create an array of quantization error values ( $q$ -values) to test.
- ii. Generate random variables uniformly distributed over  $-\frac{q}{2}$  to  $\frac{q}{2}$ . For each quantization error value, calculate mean error, variance error, and average power (which is equivalent to noise power), and then use Noise Power obtained and Signal Power to calculate the SNR.
- iii. The number of bits required for quantization is calculated as  $1 - \log_2(q)$ .
- iv. Compare the calculated SNR to the target SNR\_dB, and when it's within a specified tolerance, print the corresponding number of bits required for quantization.

- v. The program breaks the loop and exits after finding the number of bits that achieves the target SNR<sub>dB</sub>.

## Chapter 6

# Random Algebra

1. Let  $(X, Y)$  have joint probability density function

$$p_{XY}(x, y) = \begin{cases} 8xy & \text{if } 0 < x < y < 1 \\ 0 & \text{otherwise} \end{cases} \quad (6.1)$$

if  $E(X|Y = y_0) = \frac{1}{2}$ , then  $y_0$  equals

- (a)  $\frac{3}{4}$
- (b)  $\frac{1}{2}$
- (c)  $\frac{1}{3}$
- (d)  $\frac{2}{3}$

(GATE ST 2023)

**Solution:**

$$E(X|Y) = \int_{-\infty}^{\infty} xp_{X|Y}dx \quad (6.2)$$

where

$$p_{X|Y} = \frac{p_{XY}(x, y)}{p_Y(y)} \quad (6.3)$$

$$p_Y(y) = \int_0^y p_{X|Y}(x, y) dx \quad (6.4)$$

for  $0 < y < 1$

$$= \int_0^y 8xy dx \quad (6.5)$$

$$= 8y \left[ \frac{x^2}{2} \right]_0^y \quad (6.6)$$

$$= 4y^3 \quad (6.7)$$

For  $0 < x < y < 1$ , on substituting  $p_Y(y)$  we get

$$p_{X|Y} = \frac{8xy}{4y^3} \quad (6.8)$$

$$= \frac{2x}{y^2} \quad (6.9)$$

and

$$E(X|Y = y_0) = \int_0^{y_0} x \cdot \frac{2x}{y_0^2} dx \quad (6.10)$$

$$= \frac{2}{y_0^2} \left[ \frac{x^3}{3} \right]_0^{y_0} \quad (6.11)$$

$$= \frac{2y_0}{3} \quad (6.12)$$

$$\Rightarrow \frac{2y_0}{3} = \frac{1}{2} \quad (6.13)$$

$$y_0 = \frac{3}{4} \quad (6.14)$$

## Chapter 7

# Hypothesis Testing

7.1 Suppose that  $x$  is an observed sample of size 1 from a population with probability density function  $f(\cdot)$ . Based on  $x$ , consider testing

$$H_0 : f(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}; \quad y \in \mathbb{R}$$

against

$$H_1 : f(y) = \frac{1}{2} e^{-|y|}; \quad y \in \mathbb{R}.$$

Then which one of the following statements is true?

- (a) The most powerful test rejects  $H_0$  if  $|x| > c$  for some  $c > 0$
- (b) The most powerful test rejects  $H_0$  if  $|x| < c$  for some  $c > 0$
- (c) The most powerful test rejects  $H_0$  if  $||x| - 1| > c$  for some  $c > 0$
- (d) The most powerful test rejects  $H_0$  if  $||x| - 1| < c$  for some  $c > 0$

(GATE ST 2023) **Solution:**

$$L = \prod_{i=1}^1 f(x) = f(x) \tag{7.1}$$



To determine the most powerful test, we need to consider the likelihood ratio test

$$\frac{L(H_1)}{L(H_0)} \underset{H_0}{\geq} k \quad (7.2)$$

$$\implies \frac{\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}}{\frac{1}{2} e^{-2|x|}} \underset{H_0}{\geq} k \quad (7.3)$$

$$\implies e^{\frac{x^2 - 2|x|}{2}} \underset{H_0}{\geq} k \frac{\sqrt{\pi}}{\sqrt{2}} \quad (7.4)$$

$$(|x| - 1)^2 \underset{H_0}{\geq} 2 \log \left( \frac{k\sqrt{\pi}}{\sqrt{2}} \right) + 1 \quad (7.5)$$

Taking square root on both sides,

$$||x| - 1| \underset{H_0}{\geq} \sqrt{2 \log \left( \frac{k\sqrt{\pi}}{\sqrt{2}} \right) + 1} \quad (7.6)$$

$$\implies |x| \underset{H_0}{\geq} 1 + \sqrt{2 \log \left( \frac{k\sqrt{\pi}}{\sqrt{2}} \right) + 1} \quad (7.7)$$

Hence, the correct answer is (7.1c)

7.2 Suppose that  $X_1, X_2, \dots, X_n$  are independent and identically distributed random variables, each having probability density function  $f(\cdot)$  and median  $\theta$ . We want to test  $H_0 : \theta = \theta_0$  against  $H_1 : \theta > \theta_0$

Consider a test that rejects  $H_0$  if  $S > c$  for some  $c$  depending on the size of the test, where  $S$  is the cardinality of the set  $\{i : X_i > \theta_0, 1 \leq i \leq n\}$ . Then which one of the following statements is true?

- (a) Under  $H_0$ , the distribution of  $S$  depends on  $f(\cdot)$ .
- (b) Under  $H_1$ , the distribution of  $S$  does not depend on  $f(\cdot)$ .
- (c) The power function depends on  $\theta$ .
- (d) The power function does not depend on  $\theta$ .

(GATE ST 2023)

**Solution:**

**Definition 7.1:** Median  $\theta$  is defined as

$$\Pr(X_i \leq \theta) = 0.5 \text{ for all } i \text{ from } 1 \text{ to } n.$$

**Definition 7.2:**  $S$  is defined as

$$S = \sum_{i=1}^n I(X_i > \theta_0)$$

where  $I(X_i > \theta_0)$  represents an indicator function.

$$I(X_i > \theta_0) = \begin{cases} 1, & \text{if } X_i > \theta_0 \\ 0, & \text{if } X_i \leq \theta_0 \end{cases} \quad (7.8)$$

$$E(S) = E\left(\sum_{i=1}^n I(X_i > \theta_0)\right) \quad (7.9)$$

$$= \sum_{i=1}^n E(I(X_i > \theta_0)) \quad (7.10)$$

Since,

$$E(I(X_i > \theta_0)) = P(X_i > \theta_0) = \int_{\theta_0}^{\infty} f(x) dx \quad (7.11)$$

Therefore,

$$E(S) = \sum_{i=1}^n \int_{\theta_0}^{\infty} f(x) dx \quad (7.12)$$

(a) From (6.12), under  $H_0$ , the distribution of  $S$  depends on  $f(\cdot)$ .

(b) The power function can be expressed as:

$$\pi(\theta) = \Pr(\text{Reject } H_0 \mid H_1 \text{ is true}) \quad (7.13)$$

$$= \Pr(S > c \mid \theta) \quad (7.14)$$

Therefore, power function depends on value of  $\theta$ .

7.3 Let  $X_1, X_2, X_3, \dots, X_n$  be a random sample of size  $n (\geq 2)$  from a population having probability density function

$$f(x; \theta) = \begin{cases} \frac{2}{\theta x} (\log_e x) e^{-\frac{(\log_e x)^2}{\theta}} & , 0 < x < 1 \\ 0 & , \text{otherwise} \end{cases}$$

where  $\theta > 0$  is an unknown parameter. Then which of the following statements is true,

- (A)  $\frac{1}{n} \sum_{i=1}^n (\ln X_i)^2$  is the maximum likelihood estimator of  $\theta$
- (B)  $\frac{1}{n-1} \sum_{i=1}^n (\ln X_i)^2$  is the maximum likelihood estimator of  $\theta$
- (C)  $\frac{1}{n} \sum_{i=1}^n \ln X_i$  is the maximum likelihood estimator of  $\theta$
- (D)  $\frac{1}{n-1} \sum_{i=1}^n \ln X_i$  is the maximum likelihood estimator of  $\theta$

(GATE ST 2023)

**Solution:**

$$L(\theta) = f(x_1, x_2, \dots, x_n; \theta) \quad (7.15)$$

The product of pdfs can be used to approximate the likelihood function even if the

variables are dependent. This is a general approach that is often used in practice to estimate MLE of  $\theta$ . Therefore,

$$L(\theta) = \prod_{i=1}^n f(x_i; \theta) \quad (7.16)$$

Maximizing  $L(\theta)$  is equivalent to maximizing the  $\ln L(\theta)$  as  $\ln$  is a monotonically increasing function.

$$l(\theta) = \ln L(\theta) \quad (7.17)$$

$$= \ln \left( \prod_{i=1}^n f(x_i; \theta) \right) \quad (7.18)$$

$$= \sum_{i=1}^n \ln f(x_i; \theta) \quad (7.19)$$

$$= -n \ln 2 - n \ln \theta + \sum_{i=1}^n \ln(-\ln x_i) - \sum_{i=1}^n (\ln x_i) - \sum_{i=1}^n \frac{(\ln x_i)^2}{\theta} \quad (7.20)$$

Maximizing  $l(\theta)$  with respect to  $\theta$  gives the MLE estimation, therefore

$$\frac{\partial l(\theta)}{\partial \theta} = 0 \quad (7.21)$$

$$\frac{-n}{\theta} + \frac{1}{(\theta)^2} \sum_{i=1}^n (\ln x_i)^2 = 0 \quad (7.22)$$

$$\theta = \frac{1}{n} \sum_{i=1}^n (\ln x_i)^2 \quad (7.23)$$

Hence (A) is the true statement.

7.4 Suppose that  $(X, Y)$  has joint probability mass function

$$P(X = 0, Y = 0) = P(X = 1, Y = 1) = \theta, \quad (7.24)$$

$$P(X = 1, Y = 0) = P(X = 0, Y = 1) = \frac{1}{2} - \theta. \quad (7.25)$$

where  $0 \leq \theta \leq \frac{1}{2}$  is an unknown parameter. Consider testing  $H_0 : \theta = \frac{1}{4}$  against  $H_1 : \theta = \frac{1}{3}$ ; based on a random sample  $(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$  from the above probability mass function. Let  $M$  be the cardinality of the set  $\{i : X_i = Y_i, 1 \leq i \leq n\}$ . If  $m$  is the observed value of  $M$ , then which one of the following statements is true?

- (a) The likelihood ratio test rejects  $H_0$  if  $m > c$  for some  $c$ .
- (b) The likelihood ratio test rejects  $H_0$  if  $m < c$  for some  $c$ .
- (c) The likelihood ratio test rejects  $H_0$  if  $c_1 < m < c_2$  for some  $c_1$  and  $c_2$ .
- (d) The likelihood ratio test rejects  $H_0$  if  $m < c_1$  or  $m > c_2$  for some  $c_1$  and  $c_2$ .

(GATE ST 2023)

**Solution:** Given that,

$$H_0 : \quad \theta = \theta_0 = \frac{1}{4}, \quad (7.26)$$

$$H_1 : \quad \theta = \theta_1 = \frac{1}{3}. \quad (7.27)$$

and the pmf is given by

$$p_{XY}(0, 0) = p_{XY}(1, 1) = \theta \quad (7.28)$$

$$p_{XY}(0, 1) = p_{XY}(1, 0) = \frac{1}{2} - \theta \quad (7.29)$$

Then for the given random sample of data,

$$p_{X_i, Y_i}(x, y) = \begin{cases} 2\theta & x = y \\ 1 - 2\theta & x \neq y \end{cases} \quad (7.30)$$

$$(7.31)$$

Then the likelihood of the data under  $H_0$  is given by:

$$L(\theta_0 \mid data) = \prod_{i=1}^n p_{X_i, Y_i}(x, y) \quad (7.32)$$

$$= (2\theta_0)^m (1 - 2\theta_0)^{n-m} \quad (7.33)$$

$$= \left(\frac{1}{2}\right)^m \left(\frac{1}{2}\right)^{n-m} \quad (7.34)$$

Then the likelihood of the data under  $H_1$  is given by:

$$L(\theta_1 \mid data) = \prod_{i=1}^n p_{X_i, Y_i}(x, y) \quad (7.35)$$

$$= (2\theta_1)^m (1 - 2\theta_1)^{n-m} \quad (7.36)$$

$$= \left(\frac{2}{3}\right)^m \left(\frac{1}{3}\right)^{n-m} \quad (7.37)$$

The likelihood ratio will be

$$\lambda(data) = \frac{L(\theta_1 \mid x)}{L(\theta_0 \mid x)} \quad (7.38)$$

$$= \frac{\left(\frac{2}{3}\right)^m \left(\frac{1}{3}\right)^{n-m}}{\left(\frac{1}{2}\right)^m \left(\frac{1}{2}\right)^{n-m}} = (2)^m \left(\frac{2}{3}\right)^n \quad (7.39)$$

Let the critical value be denoted by  $c_1$ , then the likelihood ratio test rejects  $H_0$  if

$$\implies \lambda(data) \underset{H_0}{\overset{H_1}{\geq}} c_1 \quad (7.40)$$

$$(7.41)$$

From (7.39),

$$\implies (2)^m \left(\frac{2}{3}\right)^n \underset{H_0}{\overset{H_1}{\geq}} c_1 \quad (7.42)$$

$$\implies (2)^m \underset{H_0}{\overset{H_1}{\geq}} c_1 \left(\frac{2}{3}\right)^n \quad (7.43)$$

$$\implies m \underset{H_0}{\overset{H_1}{\geq}} \log_2 \left( c_1 \left(\frac{2}{3}\right)^n \right) \quad (7.44)$$

$$\implies m \underset{H_0}{\overset{H_1}{\geq}} c \quad \exists c \in \mathbb{R} \quad (7.45)$$

where,

$$c = \log_2 \left( c_1 \left(\frac{2}{3}\right)^n \right) \quad (7.46)$$

$\therefore$  From (7.45), Option A is correct and Options B,C,D are incorrect

7.5 Let  $X$  be a random sample of size 1 from a population with cumulative distribution function

$$F_X(x) = \begin{cases} 0 & \text{if } x \leq 0 \\ 1 - (1 - x)^\theta & \text{if } 0 \leq x < 1 \\ 1 & \text{if } x \geq 1, \end{cases} \quad (7.47)$$

where  $\theta > 0$  is an unknown parameter. To test  $H_0 : \theta = 1$  against  $H_1 : \theta = 2$ , consider

using the critical region  $(x \in \mathbb{R} : x < 0.5)$ . If  $\alpha$  and  $\beta$  denote the level and power of the test, respectively, then  $\alpha + \beta$  (rounded off to two decimal places) equals (GATE ST 2023)

**Solution:** Given that,

$$H_0 : \theta = \theta_0 = 1 \quad (7.48)$$

$$H_1 : \theta = \theta_1 = 2 \quad (7.49)$$

PDF can be defined as:

$$p_X(x) = \frac{d}{dx} F_X(x) \quad (7.50)$$

$$= \begin{cases} \theta (1-x)^{\theta-1} & \text{if } 0 \leq x < 1 \\ 0 & \text{otherwise} \end{cases} \quad (7.51)$$

Level of test:

$$\alpha = \Pr(\text{reject } H_0 | H_0 \text{ is true}) \quad (7.52)$$

$$= \Pr(x < 0.5 | \theta_0) \quad (7.53)$$

$$= F_X(0.5) \quad (7.54)$$

$$= 1 - (1 - 0.5) \quad (7.55)$$

$$= \frac{1}{2} \quad (7.56)$$



Power of test:

$$\beta = \Pr(\text{reject } H_0 | H_1 \text{ is true}) \quad (7.57)$$

$$= \Pr(x < 0.5 | \theta_1) \quad (7.58)$$

$$= F_X(0.5) \quad (7.59)$$

$$= 1 - (1 - 0.5)^2 \quad (7.60)$$

$$= \frac{3}{4} \quad (7.61)$$

Now,

$$\alpha + \beta = \frac{1}{2} + \frac{3}{4} \quad (7.62)$$

$$= 1.25 \quad (7.63)$$

7.6 Using the Ordinary Least Squares (OLS) method, a researcher estimated the relationship between initial salary (S) of MBA graduates and their cumulative grade point average (CGPA) as

$$\hat{S}_i = \hat{\beta}_0 + \hat{\beta}_1 \text{CGPA}_i, i = 1, 2, \dots, 100$$

where  $\hat{\beta}_0 = 4543$  and  $\hat{\beta}_1 = 645.08$ . The standard errors of  $\hat{\beta}_0$  and  $\hat{\beta}_1$  are 921.79 and 70.01, respectively.

The t-statistic for testing the null hypothesis  $\beta_1 = 0$  is (GATE XH 2023)

**Solution:**

**Definition 7.3(*t-statistic*):** The t-statistic is the ratio of the difference between

the estimated value of a parameter from its hypothesized value to its standard error.

$$t_{\hat{\beta}_1} = \frac{\hat{\beta}_1 - \beta_1}{SE(\hat{\beta}_1)} \quad (7.64)$$

where,

- $\hat{\beta}_1$  is the point estimate.
- $\beta_1$  is the hypothesized value.
- $SE(\hat{\beta}_1)$  standard error of the estimator.

**Definition 7.4(Standard error):** It is a measure of how much the statistic is likely to vary from the true value of the parameter it is estimating.

$$SE(\hat{\beta}_1) = \sqrt{\frac{s^2}{n-2}} \quad (7.65)$$

where,

- $s^2$  is the variance
- $n$  is the sample size

Given that  $\hat{\beta}_1 = 645.08$  and  $SE(\hat{\beta}_1) = 70.01$ , we get

$$t_{\hat{\beta}_1} = \frac{645.08 - 0}{70.01} \quad (7.66)$$

$$t_{\hat{\beta}_1} = 9.21 \quad (7.67)$$

7.7 Let  $\{0.13, 0.12, 0.78, 0.51\}$  be a realization of a random sample of size 4 from a

population with cumulative distribution function  $F(\cdot)$ . Consider testing

$$H_0 : F = F_0 \quad \text{against} \quad H_1 : F \neq F_0 \quad (7.68)$$

where,

$$F_0(x) = \begin{cases} 0 & x < 0 \\ x & 0 \leq x < 1 \\ 1 & x \geq 1 \end{cases} \quad (7.69)$$

Let  $D$  denote the Kolmogorov-Smirnov test statistic. If  $P(D > 0.669) = 0.01$  under  $H_0$  and

$$\psi = \begin{cases} 1 & \text{if } H_0 \text{ is accepted at level } 0.01 \\ 0 & \text{otherwise} \end{cases} \quad (7.70)$$

then based on the given data, the observed value of  $D + \psi$  (rounded off to two decimal places) equals (GATE ST 2023)

**Solution:** Its given that random sample is of size 4, So

$$n = 4 \quad (7.71)$$

The cdf of the random sample is given as

$$F_X(x) = \begin{cases} 0 & x < 0 \\ x & 0 \leq x < 1 \\ 1 & x \geq 1 \end{cases} \quad (7.72)$$

The empirical distribution function(edf)  $G_n$  for  $n$  independent and identically distributed (i.i.d.) ordered observations  $X_i$  is defined as

$$G_n(x) = \frac{\text{no of (elements in the sample } \leq x)}{n} = \frac{1}{n} \sum_{i=1}^n 1(X_i \leq x) \quad (7.73)$$

where  $1(A)$  is the indicator of event  $A$  and in (7.73) it is defined as,

$$1(X_i \leq x) = \begin{cases} 1 & X_i \leq x \\ 0 & \text{otherwise} \end{cases} \quad (7.74)$$

From (7.71), (7.72) and (7.73), the edf for the given data will be

$$G_n(0.13) = \frac{1}{4} \sum_{i=1}^n 1(X_i \leq 0.13) = \frac{1}{2} \quad (7.75)$$

$$G_n(0.12) = \frac{1}{4} \sum_{i=1}^n 1(X_i \leq 0.12) = \frac{1}{4} \quad (7.76)$$

$$G_n(0.78) = \frac{1}{4} \sum_{i=1}^n 1(X_i \leq 0.78) = 1 \quad (7.77)$$

$$G_n(0.51) = \frac{1}{4} \sum_{i=1}^n 1(X_i \leq 0.51) = \frac{3}{4} \quad (7.78)$$

The Kolmogorov–Smirnov statistic for a given cdf  $F_X(x)$  is

$$D_n = \sup |G_n(x) - F_X(x)| \quad (7.79)$$

The difference between cdf and edf for the given data will be (i.e.,  $\forall x \in \{0.13, 0.12, 0.78, 0.51\}$ )

$$G_n(0.13) - F_X(0.13) = 0.37 \quad (7.80)$$

$$G_n(0.12) - F_X(0.12) = 0.25 \quad (7.81)$$

$$G_n(0.78) - F_X(0.78) = 0.22 \quad (7.82)$$

$$G_n(0.51) - F_X(0.51) = 0.24 \quad (7.83)$$

Then

$$D_n = \sup(0.37, 0.25, 0.22, 0.24) = 0.37 \quad (7.84)$$

Given that,

$$P(D > 0.669) = 0.01 \quad (7.85)$$

Then

$$H_0 = \begin{cases} \text{accepted at level 0.01} & \text{if } D_n \leq 0.669 \\ \text{rejected at level 0.01} & \text{if } D_n > 0.669 \end{cases} \quad (7.86)$$

From (7.84) and (7.86); We can say that  $H_0$  is accepted at level 0.01 and

$$\psi = 1 \quad (7.87)$$

$\therefore$  the value will be

$$\psi + D_n = 1 + 0.37 = 1.37 \quad (7.88)$$



## Chapter 8

# Bivariate Random Variables





## Chapter 9

# Random Processes

9.1 Let  $X(t)$  be a Gaussian noise with power spectral density  $\frac{1}{2}W/Hz$ . If  $X(t)$  is input to an LTI system with impulse response  $e^{-tu(t)}$ . The average power of the system is (rounded off to two decimal places). (GATE EC 2023)

**Solution:** The output power spectral density of a LTI system with impulse response  $h(t)$  and input  $X(t)$  and input power spectral density  $S_X(f)$  is given by:

$$S_Y(f) = |H(f)|^2 S_X(f) \quad (9.1)$$

where  $H(f)$  is frequency response of the system.

$H(f)$  can be found by taking fourier transform of  $h(t)$

$$H(f) \xleftrightarrow{\mathcal{F}} \frac{1}{j2\pi f + 1} \quad (9.2)$$

The average power of a signal with power spectral density  $S(f)$  is given by:

$$P_Y(f) = \int_{-\infty}^{\infty} S_Y(f) df \quad (9.3)$$

Substituting  $S_Y(f)$  in the equation we get:

$$P_Y(f) = \int_{-\infty}^{\infty} |H(f)|^2 \cdot S_X(f) df \quad (9.4)$$

$$= \int_{-\infty}^{\infty} \left| \frac{1}{j2\pi f + 1} \right|^2 \cdot \frac{1}{2} df \quad (9.5)$$

$$= \frac{1}{2} \int_{-\infty}^{\infty} \frac{1}{(2\pi f)^2 + 1} df \quad (9.6)$$

$$= \frac{1}{2} \times 2 \int_0^{\infty} \frac{1}{(2\pi f)^2 + (1)^2} df \quad (9.7)$$

$$= \int_0^{\infty} \frac{1}{(2\pi f)^2 + (1)^2} df \quad (9.8)$$

$$= \frac{1}{2\pi} \tan^{-1}(x) \Big|_0^{\infty} \quad (9.9)$$

$$= \frac{1}{2\pi} (\tan^{-1} \infty - \tan^{-1} 0) \quad (9.10)$$

$$= \frac{1}{2\pi} \left( \frac{\pi}{2} - 0 \right) \quad (9.11)$$

$$= \frac{1}{2\pi} \left( \frac{\pi}{2} \right) \quad (9.12)$$

$$= \frac{1}{4} \quad (9.13)$$

Rounded off to two decimal places, the average power of the system output is  $0.25W$ .

9.2 For a real signal, which of the following is/are valid power spectral density/densities?

(a)  $S_X(\omega) = \frac{2}{9+\omega^2}$

(b)  $S_X(\omega) = e^{-\omega^2} \cos^2 \omega$

(c) See Fig. 9.1

(d) See Fig. 9.2

(GATE EC 2023)

**Solution:**

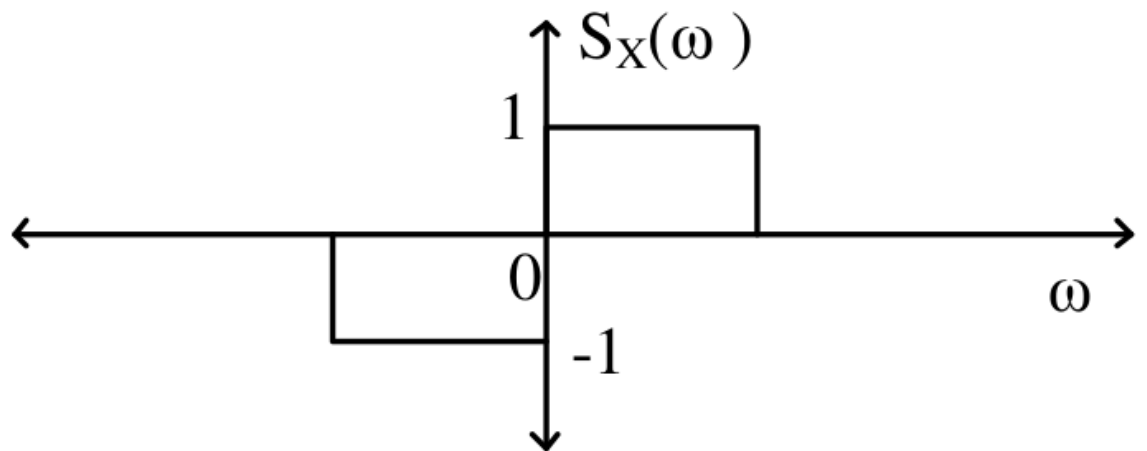


Figure 9.1: Figure1

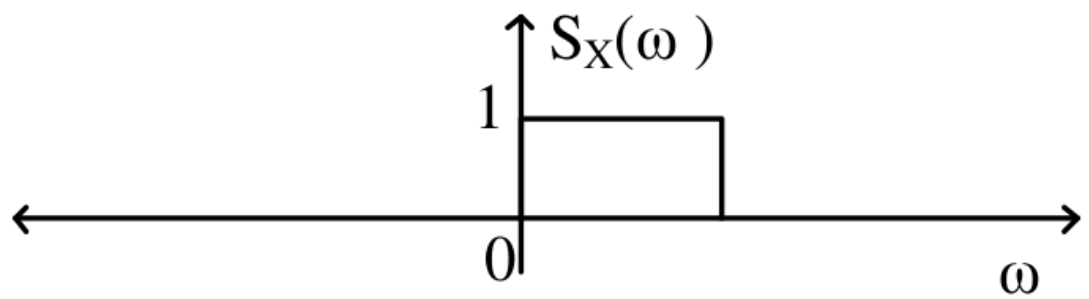


Figure 9.2: Figure2

The  $S_X$  and  $R(\tau)$  of a power signal form a Fourier transform pair, i.e.,

$$R(\tau) \stackrel{\mathcal{F}}{\rightleftharpoons} S_X \quad (9.14)$$

which gives :

$$S_X = \int R_X(\tau) e^{(-j\omega\tau)} d\tau \quad (9.15)$$

$$= \int [R_X(\tau) \cos(\omega\tau)] - j [R_X(\tau) \sin(\omega\tau)] d\tau \quad (9.16)$$

Now, For a real signal, the  $R_X(\tau)$  is real and even the fourier transform of  $R_X(\tau)$  will also exhibits the same properties:

$$\text{Im}(S_X(\omega)) = - \int j [R_X(\tau) \sin(\omega\tau) d\tau] \quad (9.17)$$

$$\implies 0 \quad (9.18)$$

$$\text{and, } S_X(-\omega) = S_X(\omega) \quad (9.19)$$

$$\int R_X(\tau) e^{j\omega\tau} d\tau = \int R_X(\tau) e^{-j\omega\tau} d\tau \quad (9.20)$$

So, the properties of  $S_X$  are:

$$(a) \text{ Im} S_X(\omega) = 0$$

$$(b) S_X(-\omega) = S_X(\omega)$$

Now,

$$(a) \text{ Plot for } S_X(\omega) = \frac{2}{9+\omega^2}$$

$$\text{Im} \left( \frac{2}{9+\omega^2} \right) = 0 \quad (9.21)$$

Also,

$$\frac{2}{9+\omega^2} = \frac{2}{9+(-\omega)^2} \quad (9.22)$$

So,  $S_X$  is valid.

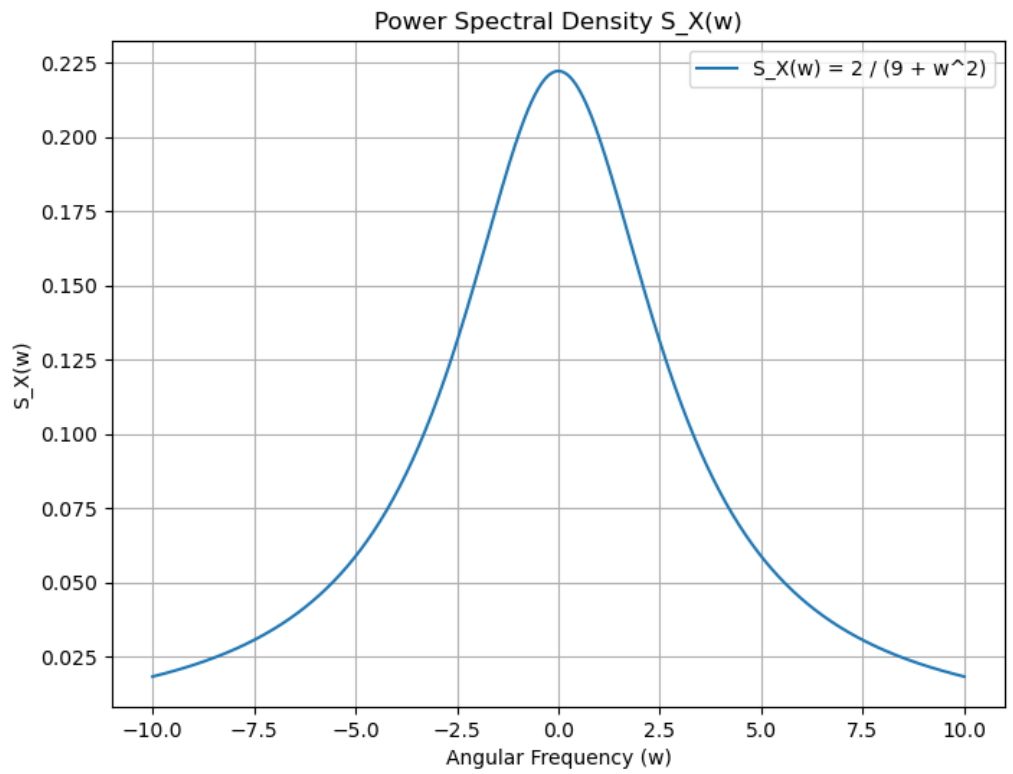


Figure 9.3: plot1

(b) Using ((a)) & ((b))

$$\text{Im}(e^{-\omega^2} \cos^2 \omega) = 0 \quad (9.23)$$

$$e^{-\omega^2} \cos^2 \omega = e^{-(-\omega)^2} \cos^2(-\omega) \quad (9.24)$$

It is also a valid  $S_X$ .

(c) Using ((a)) & ((b))

$$S_X(-\omega) = -S_X(\omega) \quad (9.25)$$

As, It is real but odd function. So, it is not a valid  $S_X$ .

(d)

$$S_X = \begin{cases} 1, 0 < \omega < \omega_o \\ 0, \text{otherwise} \end{cases} \quad (9.26)$$

Here,  $S_X$  is neither odd nor even. So, it is not valid.

$\therefore$  Option (9.2a) and (9.2b) are correct.

## Chapter 10

# Convergence

10.1 Let  $\{X_n\}_{n \geq 1}$  and  $\{Y_n\}_{n \geq 1}$  be two sequences of random variables and  $X$  and  $Y$  be two random variables, all of them defined on the same probability space. Which one of the following statements is true?

- (A) If  $\{X_n\}_{n \geq 1}$  converges in distribution to a real constant  $c$ , then  $\{X_n\}_{n \geq 1}$  converges in probability to  $c$ .
- (B) If  $\{X_n\}_{n \geq 1}$  converges in probability to  $X$ , then  $\{X_n\}_{n \geq 1}$  converges in  $3^{rd}$  mean to  $X$ .
- (C) If  $\{X_n\}_{n \geq 1}$  converges in distribution to  $X$  and  $\{Y_n\}_{n \geq 1}$  converges in distribution to  $Y$ , then  $\{X_n + Y_n\}_{n \geq 1}$  converges in distribution to  $X + Y$ .
- (D) If  $\{E(X_n)\}_{n \geq 1}$  converges to  $E(X)$ , then  $\{X_n\}_{n \geq 1}$  converges in  $1^{st}$  mean to  $X$ .

(GATE ST 2023) **Solution:**

- (a)  $X_n$  converges in distribution to  $X$ ,  $X_n \xrightarrow{d} X$ , then for all  $x$ ,

$$\lim_{n \rightarrow \infty} F_{X_n}(x) = F_X(x) \quad (10.1)$$

- (b)  $X_n$  converges in probability to  $X$ ,  $X_n \xrightarrow{p} X$ , then for all  $\epsilon > 0$ ,

$$\lim_{n \rightarrow \infty} \Pr(|X_n - X| > \epsilon) = 0 \quad (10.2)$$



(c)  $X_n$  converges in  $p^{th}$  mean to  $X$ , then we have

$$\lim_{n \rightarrow \infty} E(|X_n - X|^p) = 0 \quad (10.3)$$

(A) For  $\epsilon > 0$ ,  $B$  be defined as

$$B = \{x : |x - c| \geq \epsilon\} \quad (10.4)$$

Now,

$$\Pr(|X_n - c| \geq \epsilon) = \Pr(X_n \in B) \quad (10.5)$$

Using Portmanteau Lemma, if  $X_n \xrightarrow{d} c$ , we have

$$\limsup_{n \rightarrow \infty} \Pr(X_n \in B) \leq \Pr(c \in B) \quad (10.6)$$

$$\leq \Pr(|0 - 0| \geq \epsilon) \quad (10.7)$$

$$\leq \Pr(0 \geq \epsilon) \quad (10.8)$$

$$\leq 0 \quad (10.9)$$

$$= 0 \quad (10.10)$$

$$\lim_{n \rightarrow \infty} \Pr(|X_n - c| > \epsilon) = 0 \quad (10.11)$$

From (10.2),  $X_n \xrightarrow{p} c$ . So, we have

$$X_n \xrightarrow{d} c \implies X_n \xrightarrow{p} c \quad (10.12)$$

Option (A) is correct.

(B) Statement (B) is may or may not correct. Counter Example: Consider distribution

$X_n$	0	$n$
$\Pr(X_n)$	$1 - \frac{1}{n}$	$\frac{1}{n}$

For  $\epsilon > 0$ ,  $X_n$  converges in probability to  $X = 0$

$$\lim_{n \rightarrow \infty} \Pr(|X_n - X| > \epsilon) = \lim_{n \rightarrow \infty} \Pr(X_n > \epsilon) \quad (10.13)$$

$X_n > \epsilon$  is subset of  $X_n = n$  since every time  $X_n$  equals  $n$ , it's also true that  $X_n$  is greater than  $\epsilon$ . But there may be times when  $X_n$  is greater than  $\epsilon$  without  $X_n$  being equal to  $n$ . So,

$$\Pr(X_n > \epsilon) \leq \Pr(X_n = n) \quad (10.14)$$

$$\lim_{n \rightarrow \infty} \Pr(|X_n - X| > \epsilon) \leq \lim_{n \rightarrow \infty} \Pr(X_n = n) \quad (10.15)$$

$$\leq \lim_{n \rightarrow \infty} \frac{1}{n} \quad (10.16)$$

$$\leq 0 \quad (10.17)$$

$$= 0 \quad (10.18)$$

But  $X_n$  does not converges in  $3^{rd}$  mean to  $X = 0$ .

$$\lim_{n \rightarrow \infty} E(|X_n - X|^3) = \lim_{n \rightarrow \infty} E(X_n^3) \quad (10.19)$$

$$= \lim_{n \rightarrow \infty} 0^3 \left(1 - \frac{1}{n}\right) + n^3 \left(\frac{1}{n}\right) \quad (10.20)$$

$$= \lim_{n \rightarrow \infty} n^2 \neq 0 \quad (10.21)$$

(C) Statement (C) is may or may not correct. Counter Example: Consider distribution

$$Z \sim \mathcal{N}(0, 1) \quad (10.22)$$

Let  $\{X_n\}_{n \geq 1}$  and  $\{Y_n\}_{n \geq 1}$  be sequences of random variables such that they both converge in distribution as  $Z$  and  $(-1)^n Z$ . Proof that  $Y_n$  converges in distribution.

For  $n$  even

$$\lim_{n \rightarrow \infty} F_{Y_n}(x) = \Pr(Z \leq x) \quad (10.23)$$

For  $n$  odd

$$\lim_{n \rightarrow \infty} F_{Y_n}(x) = \Pr(-Z \leq x) \quad (10.24)$$

$$= \Pr(Z \leq x) \quad (10.25)$$

Proved. So, we have

$$F_{X_n+Y_n}(x) = \Pr(X_n + Y_n \leq x) \quad (10.26)$$

$$= \Pr(Z + (-1)^n Z \leq x) \quad (10.27)$$

For  $n$  is even

$$F_{X_n+Y_n}(x) = \Pr(2Z \leq x) \quad (10.28)$$

$$= \Pr\left(Z \leq \frac{x}{2}\right) \quad (10.29)$$

$$= 1 - \Pr\left(Z > \frac{x}{2}\right) \quad (10.30)$$

$$\approx 1 - Q\left(\frac{x}{2}\right) \quad (10.31)$$

For  $n$  is odd

$$F_{X_n+Y_n}(x) = \Pr(0 \leq x) \quad (10.32)$$

$$= \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases} = H(x) \quad (10.33)$$

So, on generalizing

$$F_{X_n+Y_n}(x) = \begin{cases} 1 - Q\left(\frac{x}{2}\right) & \text{if } n \text{ is even} \\ H(x) & \text{if } n \text{ is odd} \end{cases} \quad (10.34)$$

$\lim_{n \rightarrow \infty} F_{X_n+Y_n}(x)$  oscillate between  $1 - Q\left(\frac{x}{2}\right)$  and  $H(x)$ . This doesnot imply convergence.

(D) Statement (D) is may or may not correct. Counter Example: Consider

$X_n$	0	$n$
$\Pr(X_n)$	$1 - \frac{1}{n}$	$\frac{1}{n}$

$$\lim_{n \rightarrow \infty} E(X_n) = 0 \left(1 - \frac{1}{n}\right) + n \left(\frac{1}{n}\right) \quad (10.35)$$

$$= 1 \quad (10.36)$$

As  $n \rightarrow \infty$ ,  $E(X_n)$  converges to  $E(X) = 1$ .

$$\lim_{n \rightarrow \infty} X_n = 0 = X \quad (10.37)$$

To find 1<sup>st</sup> mean convergence of  $X_n$ . From (10.36)

$$\lim_{n \rightarrow \infty} E(|X_n - X|) = \lim_{n \rightarrow \infty} E(X_n) \quad (10.38)$$

$$= 1 \neq 0 \quad (10.39)$$

So,  $X_n$  does not converge in 1<sup>st</sup> mean to  $X$ .

10.2 Let  $\{X_n\}_{n \geq 1}$  be a sequence of independent and identically distributed random variables each having a mean 4 and variance 9. If  $Y_n = \frac{1}{n} \sum_{i=1}^n X_i$  for  $n \geq 1$ , then  $\lim_{n \rightarrow \infty} E \left[ \left( \frac{Y_n - 4}{\sqrt{n}} \right)^2 \right]$  (in integer) equals \_\_\_\_\_. (GATE ST 2023)

**Solution:**

(a) **Theory:** For all  $X_i$  which are i.i.d's, mean  $\mu = 4$  and variance  $\sigma^2 = 9$ ,

$$Y_n = \frac{1}{n} \sum_{i=1}^n X_i \quad (10.40)$$

The mean of a sum of i.i.d random variables is calculated as

$$E[Y_n] = E \left[ \frac{1}{n} \sum_{i=1}^n X_i \right] \quad (10.41)$$

$$= \frac{1}{n} \sum_{i=1}^n E[X_i] \quad (10.42)$$

$$= \frac{1}{n} (n\mu) \quad (10.43)$$

$$= \mu \quad (10.44)$$

The variance of a sum of i.i.d random variables is calculated as

$$\text{var}(Y_n) = \text{E} \left[ \left( \frac{1}{n} \sum_{i=1}^n X_i \right)^2 \right] - \left( \text{E} \left[ \frac{1}{n} \sum_{i=1}^n X_i \right] \right)^2 \quad (10.45)$$

$$= \frac{1}{n^2} \left\{ \text{E} \left[ \left( \sum_{i=1}^n X_i \right)^2 \right] - \left( \text{E} \left[ \sum_{i=1}^n X_i \right] \right)^2 \right\} \quad (10.46)$$

But

$$\text{E} \left[ \left( \sum_{i=1}^n X_i \right)^2 \right] = \text{E} \left[ \sum_{i=1}^n \sum_{j=1}^n X_i X_j \right] \quad (10.47)$$

$$= \sum_{i=1}^n \sum_{j=1}^n \text{E} [X_i X_j] \quad (10.48)$$

and

$$\left( \text{E} \left[ \sum_{i=1}^n X_i \right] \right)^2 = \left( \sum_{i=1}^n \text{E} [X_i] \right)^2 \quad (10.49)$$

$$= \sum_{i=1}^n \sum_{j=1}^n \text{E} [X_i] \text{E} [X_j] \quad (10.50)$$

Putting (10.48) and (10.50) in (10.46), and using the definition of covariance,

$$\text{var}(Y_n) = \frac{1}{n^2} \left\{ \sum_{i=1}^n \sum_{j=1}^n (\text{E} [X_i X_j] - \text{E} [X_i] \text{E} [X_j]) \right\} \quad (10.51)$$

$$= \frac{1}{n^2} \left\{ \sum_{i=1}^n \sum_{j=1}^n \text{cov}(X_i, X_j) \right\} \quad (10.52)$$

As all the variables are i.i.d's and are thus uncorrelated,

$$\text{cov}(X_i, X_j) = \begin{cases} 0 & \text{if } i \neq j \\ \text{var}(X_i) & \text{if } i = j \end{cases} \quad (10.53)$$

Putting (10.53) in (10.52),

$$\text{var}(Y_n) = \frac{1}{n^2} \left( \sum_{i=1}^n \text{cov}(X_i, X_i) \right) \quad (10.54)$$

$$= \frac{1}{n^2} \left( \sum_{i=1}^n \text{var}(X_i) \right) \quad (10.55)$$

$$= \frac{1}{n^2} \left( \sum_{i=1}^n \sigma^2 \right) \quad (10.56)$$

$$= \frac{\sigma^2}{n} \quad (10.57)$$

Consider the term  $\left(\frac{Y_n - \mu}{\sqrt{n}}\right)^2$ . Calculating its expectation,

$$\text{E} \left[ \left( \frac{Y_n - \mu}{\sqrt{n}} \right)^2 \right] = \frac{1}{n} \text{E} [(Y_n - \mu)^2] \quad (10.58)$$

$$= \frac{1}{n} \text{var}(Y_n) \quad (10.59)$$

$$= \frac{\sigma^2}{n^2} \quad (10.60)$$

Substituting  $\sigma^2 = 9$  and  $\mu = 4$ , we get

$$\lim_{n \rightarrow \infty} \text{E} \left[ \left( \frac{Y_n - 4}{\sqrt{n}} \right)^2 \right] = \lim_{n \rightarrow \infty} \frac{9}{n^2} = 0 \quad (10.61)$$

(b) **Simulation:** Any distribution with mean  $\mu = 4$  and variance  $\sigma^2 = 9$  can be used for the variable  $X_{ij}$  for all  $i, j \in \mathbb{N}$ ; as shown in the Theory part, the limit is always

zero regardless of the distribution. The most straightforward distribution that can be used for  $X_{ij}$  is:

$$p_{X_{ij}}(x) = \begin{cases} 0.5 & \text{if } x \in \{1, 7\} \\ 0 & \text{otherwise} \end{cases} \quad (10.62)$$

This distribution has the following characteristics:

$$\mu = E[X_{ij}] = 0.5 \times 1 + 0.5 \times 7 = 4 \quad (10.63)$$

$$\sigma^2 = E[X_{ij}^2] - (E[X_{ij}])^2 \quad (10.64)$$

$$= (0.5 \times 1^2 + 0.5 \times 7^2) - 4^2 \quad (10.65)$$

$$= 9 \quad (10.66)$$

A matrix  $X_{n \times m}$  is generated for all  $i \leq n$  and  $j \leq m$ . Using this matrix, a set of  $m$  values for  $Y_j$  is generated as

$$Y_j = \frac{1}{n} \sum_{i=1}^n X_{ij} \quad (10.67)$$

Now, the expression  $\frac{(Y_j-4)^2}{n}$  is calculated for all  $j \leq m$  and their expectancy is calculated as follows:

$$E \left[ \left( \frac{Y_n - 4}{\sqrt{n}} \right)^2 \right] = \frac{1}{m} \sum_{j=1}^m \frac{(Y_j - 4)^2}{n} \quad (10.68)$$

To calculate the limit  $n \rightarrow \infty$ , different values of  $n$  are taken, and the expected value is calculated (taking a fixed small value of  $m$  to reduce computational time) for each case. This output is plotted and is seen to be close to the curve  $\frac{9}{n^2}$ , as



derived in (10.61). In both cases, we can observe the limit tends towards zero.

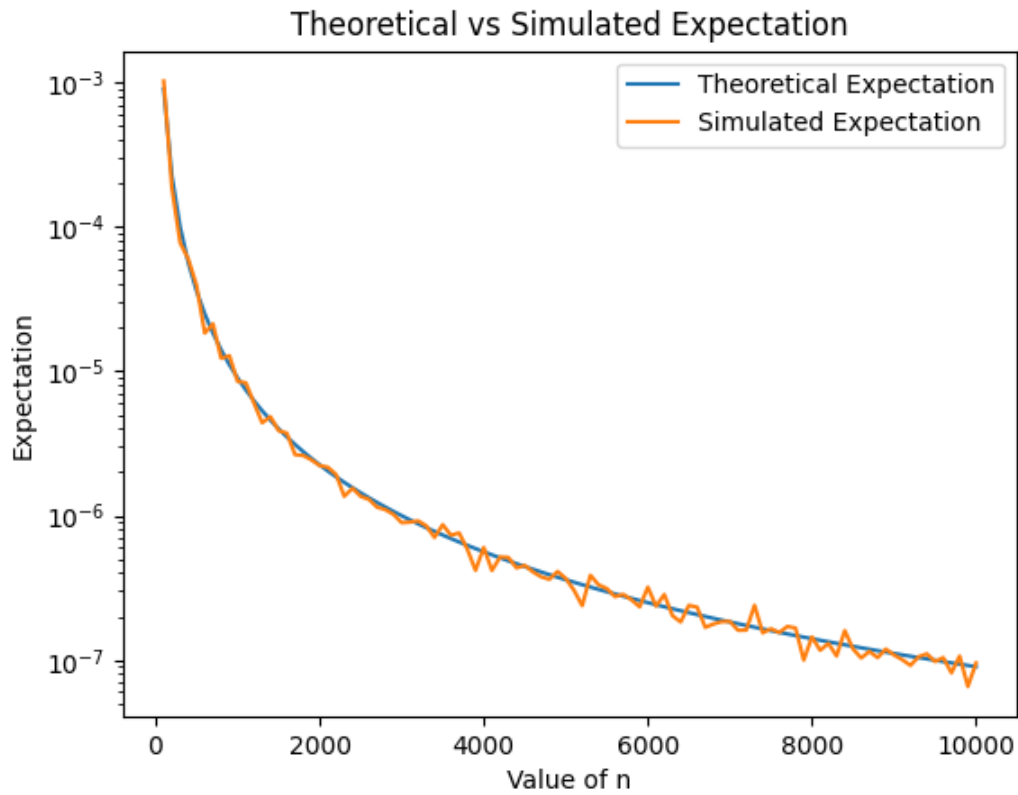


Figure 10.1: Expectation vs n

10.3 Let  $\{X_n\}_{n \geq 1}$  be a sequence of independent and identically distributed random variables with mean 0 and variance 1, all of them defined on the same probability space. For  $n=1,2,3,\dots$ , let

$$Y_n = \frac{1}{n}(X_1X_2 + X_3X_4 + \dots + X_{2n-1}X_{2n}) \quad (10.69)$$

Then which of the following statements is/are true?

(A)  $\{\sqrt{n}Y_n\}_{n \geq 1}$  converges in distribution to a standard normal random variable.

- (B)  $\{Y_n\}_{n \geq 1}$  converges in 2nd mean to 0.
- (C)  $\{Y_n + \frac{1}{n}\}_{n \geq 1}$  converges in probability to 0.
- (D)  $\{X_n\}_{n \geq 1}$  converges almost surely to 0.

(GATE ST 2023)

**Solution:**

- (a)  $X_n$  converges in distribution to a standard normal random variable  $X$ ,  $X_n \xrightarrow{d} X$ , then for all  $x$ ,

$$\lim_{n \rightarrow \infty} F_{X_n}(x) = F_X(x) \quad (10.70)$$

where,  $X \sim \mathcal{N}(0, 1)$

- (b)  $X_n$  converges in  $p^{th}$  mean to  $X$ , then we have

$$\lim_{n \rightarrow \infty} E(|X_n - X|^p) = 0 \quad (10.71)$$

- (c)  $X_n$  converges in probability to  $X$ ,  $X_n \xrightarrow{p} X$ , then for all  $\epsilon > 0$ ,

$$\lim_{n \rightarrow \infty} \Pr(|X_n - X| \geq \epsilon) = 0 \quad (10.72)$$

- (d)  $X_n$  converges almost surely to  $X$ ,  $X_n \xrightarrow{a.s.} X$ , then for all  $\epsilon > 0$ ,

$$\lim_{n \rightarrow \infty} \Pr(|X_n - X| < \epsilon) = 1 \quad (10.73)$$

$$E(X) = \int_{-\infty}^{\infty} x F_X x dx \quad (10.74)$$

$$\because X \text{ is positive-valued} \quad (10.75)$$

$$= \int_0^{\infty} x F_X x dx \quad (10.76)$$

$$\geq \int_{\epsilon^2}^{\infty} x F_X x dx \quad (10.77)$$

$$\text{for some } \epsilon^2 > 0 \quad (10.78)$$

$$\geq \int_{\epsilon^2}^{\infty} k F_X x dx \quad (10.79)$$

$$= \epsilon^2 \int_{\epsilon^2}^{\infty} F_X x dx \quad (10.80)$$

$$E(X) = \epsilon^2 \Pr(X \geq \epsilon^2) \quad (10.81)$$

This inequality is called Markov inequality

Now, we take

$$\sigma^2 = E(X^2) - [E(X)]^2 \quad (10.82)$$

$$= E(X^2) - \mu^2 \quad (10.83)$$

$$= E(X^2) - E(\mu^2) \quad (10.84)$$

$$\sigma^2 = E(X^2 - \mu^2) \quad (10.85)$$

using (10.81),

$$\sigma^2 = \epsilon^2 \Pr(X^2 - \mu^2 \geq \epsilon^2) \quad (10.86)$$

$$\frac{\sigma^2}{\epsilon^2} = \Pr(X^2 - \mu^2 \geq \epsilon^2) \quad (10.87)$$

$$\Pr(|X - \mu| \geq \epsilon) = \frac{\sigma^2}{\epsilon^2} \quad (10.88)$$

This above inequality is called Chebyshev's inequality

Now, as  $X_i$  is a sequence of independent and identically distributed random variables,

Let  $Z_i = X_{2i-1}X_{2i}$ .

$$p_{Z_i Z_j}(x) = p_{X_{2i-1} X_{2i} X_{2j-1} X_{2j}}(x) \quad (10.89)$$

$$= p_{X_{2i-1}}(x) p_{X_{2i}}(x) p_{X_{2j-1}}(x) p_{X_{2j}}(x) \quad (10.90)$$

$$= p_{X_{2i-1} X_{2i}}(x) p_{X_{2j-1} X_{2j}}(x) \quad (10.91)$$

$$= p_{Z_i}(x) p_{Z_j}(x) \quad (10.92)$$

$Z_i$  is an independent R.V. Similarly,

$$F_{X_{2i-1}}(x) = F_{X_{2i}}(x) = F_{X_{2j-1}}(x) = F_{X_{2j}}(x) \quad (10.93)$$

$$F_{X_{2i-1}}(x) F_{X_{2i}}(x) = F_{X_{2j-1}}(x) F_{X_{2j}}(x) \quad (10.94)$$

$$F_{Z_i}(x) = F_{Z_j}(x) \quad (10.95)$$

Thus,  $Z_i$  is an identical distributed R.V.

Hence,

$$Z_n = \sum_{i=1}^n Z_i \quad (10.96)$$

is an i.i.d

$$\mathbb{E}(Z_i) = \mathbb{E}(X_{2i-1}X_{2i}) \quad (10.97)$$

$$= \mathbb{E}(X_{2i-1})\mathbb{E}(X_{2i}) = 0 \quad (10.98)$$

$$\text{Var}(Z_i) = \mathbb{E}(Z_i^2) - [\mathbb{E}(Z_i)]^2 \quad (10.99)$$

$$= \mathbb{E}(Z_i^2) - 0 \quad (10.100)$$

$$= \mathbb{E}(X_{2i-1}^2 X_{2i}^2) \quad (10.101)$$

$$= \mathbb{E}(X_{2i-1}^2)\mathbb{E}(X_{2i}^2) = 1 \quad (10.102)$$

Now,

$$Y_n = \frac{1}{n} \left( \sum_{i=1}^n Z_i \right) \quad (10.103)$$

$$\mathbb{E}(Y_n) = \mathbb{E} \left( \frac{1}{n} \sum_{i=1}^n Z_i \right) \quad (10.104)$$

$$= \frac{1}{n} \sum_{i=1}^n \mathbb{E}(Z_i) = 0 \quad (10.105)$$

As all the variables are i.i.d's and are thus uncorrelated,

$$\text{cov}(X_i, X_j) = \begin{cases} 0 & \text{if } i \neq j \\ \text{Var}(X_i) & \text{if } i = j \end{cases} \quad (10.106)$$

$$\text{Var}(Y_n) = \frac{1}{n^2} \left( \sum_{i=1}^n \text{cov}(Z_i, Z_i) \right) \quad (10.107)$$

$$= \frac{1}{n^2} \left( \sum_{i=1}^n \text{var}(Z_i) \right) \quad (10.108)$$

$$= \frac{1}{n^2} \left( \sum_{i=1}^n 1 \right) \quad (10.109)$$

$$= \frac{1}{n} \quad (10.110)$$

(A) by central limit theorem,

$$\frac{Y_n - \mu}{\sigma} \sim \mathcal{N}(0, 1) \quad (10.111)$$

$$= \frac{Y_n - 0}{\sqrt{\frac{1}{n}}} \sim \mathcal{N}(0, 1) \quad (10.112)$$

$$= \sqrt{n}Y_n \sim \mathcal{N}(0, 1) \quad (10.113)$$

Thus,

$$\lim_{n \rightarrow \infty} F_{\sqrt{n}Y_n} = F_P(x) \quad (10.114)$$

where,  $P \sim \mathcal{N}(0, 1)$

So, option (A) is correct

(B) For 2nd mean to be converging to 0,

$$\lim_{n \rightarrow \infty} E(|Y_n - 0|^2) = \lim_{n \rightarrow \infty} E(Y_n^2) \quad (10.115)$$

$$\because \frac{1}{n} = E(Y_n^2) - [E(Y_n)]^2 \quad (10.116)$$

$$= \lim_{n \rightarrow \infty} \frac{1}{n} + [E(Y_n)]^2 \quad (10.117)$$

$$\because [E(Y_n)]^2 = 0 \quad (10.118)$$

$$= \lim_{n \rightarrow \infty} \frac{1}{n} = 0 \quad (10.119)$$

Thus,  $\{Y_n\}_{n \geq 1}$  converges in 2nd mean to 0 Hence, option (B) is correct.

(C) For  $\{Y_n + \frac{1}{n}\}_{n \geq 1}$  to be converging  $Y_n + \frac{1}{n} \xrightarrow{p} 0$

$$\lim_{n \rightarrow \infty} \Pr \left( \left| Y_n + \frac{1}{n} - 0 \right| \geq \epsilon \right) = \lim_{n \rightarrow \infty} \Pr \left( \left| Y_n + \frac{1}{n} \right| \geq \epsilon \right) \quad (10.120)$$

$$\leq \lim_{n \rightarrow \infty} \frac{\text{Var}(Y_n + \frac{1}{n})}{\epsilon^2} \quad (10.121)$$

$$\forall \epsilon > 0 \quad (10.122)$$

$$= \lim_{n \rightarrow \infty} \frac{\text{Var}(Y_n) + \text{Var}(\frac{1}{n})}{\epsilon^2} \quad (10.123)$$

$$= \lim_{n \rightarrow \infty} \frac{\frac{1}{n} + 0}{\epsilon^2} \quad (10.124)$$

$$= 0 \quad (10.125)$$

hence, option (C) is correct

(D) As all the variables are i.i.d's and are thus uncorrelated,

$$\text{Var}(X_n) = \sum_{i=1}^n \text{cov}(X_i, X_i) \quad (10.126)$$

$$= \sum_{i=1}^n \text{var}(X_i) \quad (10.127)$$

$$= \sum_{i=1}^n 1 \quad (10.128)$$

$$= n \quad (10.129)$$

$$\lim_{n \rightarrow \infty} \Pr(|X_n - X| < \epsilon) = 1 - \lim_{n \rightarrow \infty} \Pr(|X_n - 0| \geq \epsilon) \quad (10.130)$$

$$\leq 1 - \lim_{n \rightarrow \infty} \frac{\text{Var}(X_n)}{\epsilon^2} \quad (10.131)$$

$$\forall \epsilon > 0 \quad (10.132)$$

$$= \lim_{n \rightarrow \infty} 1 - \frac{\text{Var}(X_n)}{\epsilon^2} \quad (10.133)$$

$$= \lim_{n \rightarrow \infty} 1 - \frac{n}{\epsilon^2} = -\infty \quad (10.134)$$

Hence, option (D) is incorrect

### Simulation:

- (a) we take  $X_i$  as uniform R.V and generate it by using **generateUniformRandom()**
- (b) for A, we initialize  $Y_n$  in the **main()** and standardize  $Y_n$  and save it to file and plot it in python. Also, compare it to standard normal distribution.
- (c) for B, we define a function **calculate2ndmean** to compute 2nd mean to see its convergence
- (d) for C, we define a **calculateYnplus1overn** to define the sequence  $Y_n + \frac{1}{n}$  and



compute the proportion of simulation i.e  $\Pr(|Y_n + \frac{1}{n} - Y| \geq \epsilon)$

(e) for D, we define a **simulateXn** to define the sequence  $X_n$ .

(f) To check convergence almost surely by a function **checkconvergencealmost-surely** that computes  $\Pr(|X_n| < \epsilon)$  and if  $\geq 1$  then it does not converges almost surely.

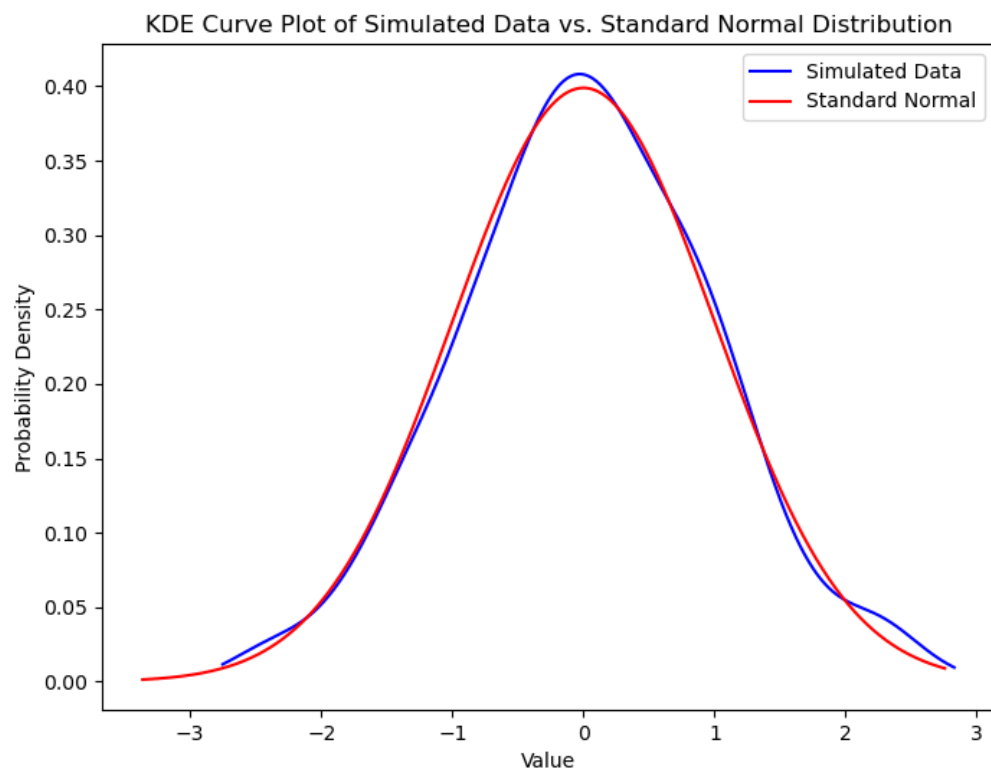


Figure 10.2: simulated vs standard normal

10.4 Let  $\{X_n\}_{n \geq 1}$  be a sequence of independent and identically distributed random variables

each having probability density function

$$f(x) = \begin{cases} e^{-x} & \text{if } x > 0 \\ 0 & \text{otherwise.} \end{cases}$$

For  $n \geq 1$ , let  $Y_n = |X_{2n} - X_{2n-1}|$ . If  $\bar{Y}_n = \frac{1}{n} \sum_{i=1}^n Y_i$  for  $n \geq 1$  and  $\{\sqrt{n}(e^{-\bar{Y}_n} - e^{-1})\}_{n \geq 1}$  converges in distribution to a normal random variable with mean 0 and variance  $\sigma^2$ , then  $\sigma^2$  (rounded off to two decimal places) equals (GATE ST 2023)

**Solution:**

(a) Let  $X, Y \sim \exp(1)$  and  $Z = X - Y$

$$p_X(x) = e^{-x} u(x) \quad (10.135)$$

$$M_X(s) = E(e^{-sX}) \quad (10.136)$$

$$= \int_0^\infty e^{-sx} e^{-x} dx \quad (10.137)$$

$$= \frac{1}{s+1} \quad (10.138)$$

ROC for  $M_X(s) : \operatorname{Re}(s) > -1$

Similarly,

$$M_Y(s) = \frac{1}{s+1} \quad (10.139)$$

$$M_Y(-s) = \frac{1}{-s+1} \quad (10.140)$$

ROC for  $M_Y(-s) : \operatorname{Re}(s) < 1$

$$M_Z(s) = E(e^{-sZ}) \quad (10.141)$$

Using,

$$Z = X - Y \quad (10.142)$$

$$\implies M_Z(s) = E \left( e^{-s(X-Y)} \right) \quad (10.143)$$

$$= E \left( e^{-sX} \right) E \left( e^{sY} \right) \quad (10.144)$$

$$= M_X(s) M_Y(-s) \quad (10.145)$$

$$= \frac{1}{s+1} \times \frac{1}{-s+1} \quad (10.146)$$

$$M_Z(s) = \frac{1}{1-s^2} \quad (10.147)$$

The ROC for the laplace transform :  $|\text{Re}(s)| < 1$

$$M_Z(s) = \frac{1}{2} \left( \frac{1}{1-s} + \frac{1}{1+s} \right) \quad (10.148)$$

Using Inverse Laplace transform,

$$P_Z(x) = \frac{1}{2} \left( e^x u(-x) + e^{-x} u(x) \right) \quad (10.149)$$

$$p_Z(x) = \frac{1}{2} e^{-|x|} \quad (10.150)$$

$$\implies Z \sim \text{Lap}(0, 1) \quad (10.151)$$

(b) Let  $T = |Z|$

$$p_Z(x) = \frac{1}{2}e^{-|x|} \quad (10.152)$$

$$F_Z(x) = \int_{-\infty}^x \frac{1}{2}e^{-|t|} dt \quad (10.153)$$

$$= \frac{1}{2} + \frac{1}{2}e^{-x} \quad (10.154)$$

$$F_T(x) = \Pr(T \leq x) \quad (10.155)$$

$$= \Pr(|Z| \leq x) \quad (10.156)$$

$$= \Pr(-x \leq Z \leq x) \quad (10.157)$$

$$F_T(x) = \frac{1}{2} - \frac{1}{2}e^{-x} - \left(-\frac{1}{2} + \frac{1}{2}e^{-x}\right) \quad (10.158)$$

$$F_T(x) = 1 - e^{-x} \text{ for } x > 0 \quad (10.159)$$

$$T \sim \exp(1) \quad (10.160)$$

$$\implies |Z| \sim \exp(1) \quad (10.161)$$

Using equations (10.142) and (10.161), we get:

$$|X_{2n} - X_{2n-1}| \sim \exp(1) \quad (10.162)$$

$$\implies Y_n \sim \exp(1) \quad (10.163)$$

$$M_{Y_n}(s) = \frac{1}{1+s} \quad (10.164)$$

$$\mathbb{E}(Y_n) = \mu_1 \quad (10.165)$$

$$\mu_1 = -\frac{dM_{Y_n}(s)}{ds} \quad (10.166)$$

$$= -\frac{d}{ds} \left( \frac{1}{s+1} \right) \Big|_{s=0} \quad (10.167)$$

$$= \frac{1}{(s+1)^2} \Big|_{s=0} \quad (10.168)$$

$$\mathbb{E}(Y_n) = 1 \quad (10.169)$$

$$\mathbb{E}(Y_n^2) = \mu_2 \quad (10.170)$$

$$\mu_2 = \frac{d^2 Y_n(s)}{ds^2} \quad (10.171)$$

$$= \frac{d^2}{ds^2} \left( \frac{-1}{(s+1)^2} \right) \Big|_{s=0} \quad (10.172)$$

$$= \frac{2}{(s+1)^3} \Big|_{s=0} \quad (10.173)$$

$$\mathbb{E}(Y_n^2) = 2 \quad (10.174)$$

$$\text{Var}(Y_n) = \mathbb{E}((Y_n - \mathbb{E}(Y_n))^2) \quad (10.175)$$

$$= \mathbb{E}((Y_n - 1)^2) \quad (10.176)$$

$$= \mathbb{E}(Y_n^2) - 2\mathbb{E}(Y_n) + 1 = 1 \quad (10.177)$$

(c) We know,

$$\bar{Y}_n = \frac{1}{n} \sum_{i=1}^n Y_i \quad (10.178)$$

$$\mathbb{E}(\bar{Y}_n) = \frac{1}{n} \sum_{i=1}^n \mathbb{E}(Y_i) \quad (10.179)$$

$$= \frac{1}{n} \cdot (n) = 1 \quad (10.180)$$

$$\mathbb{E}(\bar{Y}_n) = 1 \quad (10.181)$$

$$\text{var}(\bar{Y}_n) = \text{E} \left[ \left( \frac{1}{n} \sum_{i=1}^n Y_i \right)^2 \right] - \left( \text{E} \left[ \frac{1}{n} \sum_{i=1}^n Y_i \right] \right)^2 \quad (10.182)$$

$$= \frac{1}{n^2} \left\{ \text{E} \left[ \left( \sum_{i=1}^n Y_i \right)^2 \right] - \left( \text{E} \left[ \sum_{i=1}^n Y_i \right] \right)^2 \right\} \quad (10.183)$$

But

$$\text{E} \left[ \left( \sum_{i=1}^n Y_i \right)^2 \right] = \text{E} \left[ \sum_{i=1}^n \sum_{j=1}^n Y_i Y_j \right] \quad (10.184)$$

$$= \sum_{i=1}^n \sum_{j=1}^n \text{E} [Y_i Y_j] \quad (10.185)$$

and

$$\left( \text{E} \left[ \sum_{i=1}^n Y_i \right] \right)^2 = \left( \sum_{i=1}^n \text{E} [Y_i] \right)^2 \quad (10.186)$$

$$= \sum_{i=1}^n \sum_{j=1}^n \text{E} [Y_i] \text{E} [Y_j] \quad (10.187)$$

Putting (10.185) and (10.187) in (10.183), and using the definition of covariance,

$$\text{var}(\bar{Y}_n) = \frac{1}{n^2} \left\{ \sum_{i=1}^n \sum_{j=1}^n (\text{E} [Y_i Y_j] - \text{E} [Y_i] \text{E} [Y_j]) \right\} \quad (10.188)$$

$$= \frac{1}{n^2} \left\{ \sum_{i=1}^n \sum_{j=1}^n \text{cov}(Y_i, Y_j) \right\} \quad (10.189)$$

As all the variables are i.i.d's and are thus uncorrelated,

$$\text{cov}(Y_i, Y_j) = \begin{cases} 0 & \text{if } i \neq j \\ \text{var}(Y_i) & \text{if } i = j \end{cases} \quad (10.190)$$

Putting (10.190) in (10.189),

$$\text{var}(\bar{Y}_n) = \frac{1}{n^2} \left( \sum_{i=1}^n \text{cov}(Y_i, Y_i) \right) \quad (10.191)$$

$$= \frac{1}{n^2} \left( \sum_{i=1}^n \text{var}(Y_i) \right) \quad (10.192)$$

$$= \frac{1}{n^2} \cdot n = \frac{1}{n} \quad (10.193)$$

$$\text{var}(\bar{Y}_n) = \frac{1}{n} \quad (10.194)$$

$$\implies \text{E}(\bar{Y}_n) = 1 \text{ and } \text{Var}(\bar{Y}_n) = \frac{1}{n} \quad (10.195)$$

By the Central Limit Theorem,  $n \rightarrow \infty \implies \sqrt{n}(Y_n - \mu) \rightarrow \mathcal{N}(0, 1)$

$$\frac{\bar{Y}_n - 1}{\frac{1}{\sqrt{n}}} \sim \mathcal{N}(0, 1) \quad (10.196)$$

$$\sqrt{n}(\bar{Y}_n - 1) \sim \mathcal{N}(0, 1) \quad (10.197)$$

We know,

$$\sqrt{n}(Y_n - k) \sim \mathcal{N}(0, \sigma^2) \quad (10.198)$$

Let us write the Taylor expansion of  $g(Y_n)$  around  $k$

$$g(Y_n) = g(k) + g'(k)(Y_n - k) + \frac{1}{2}g''(k)(Y_n - k)^2 + \dots \quad (10.199)$$

Apply the Central Limit Theorem (CLT) to the standardized variable  $Z_n$

$$Z_n = \frac{\sqrt{n}g'(k)(Y_n - k)}{\sigma\sqrt{n}} \quad (10.200)$$

$$n \rightarrow \infty \implies Z_n \rightarrow \mathcal{N}(0, 1) \quad (10.201)$$

Compare with standardised variable we get,

$$\implies \sqrt{n}(g(Y_n) - g(k)) \sim \mathcal{N}(0, \sigma^2[g'(k)]^2) \quad (10.202)$$

$$g(x) = e^{-x} \implies g'(x) = -e^{-x} \quad (10.203)$$

Using equation (10.202), we get:

$$\sqrt{n}(e^{-\bar{Y}_n} - e^{-1}) \sim \mathcal{N}(0, e^{-2}) \quad (10.204)$$

$$\implies \sigma^2 = e^{-2} = 0.14 \quad (10.205)$$

### Steps for Simulation:

(a) `rand()` / `(double)RAND_MAX`:

This generates a random variable between 0 and `RAND_MAX` and divides it by `RAND_MAX` to obtain a uniform distribution between 0 and 1

(b) `-log(rand() / (double)RAND_MAX)` :

This transforms the uniform distribution between 0 and 1 into an exponential distribution by making the values vary from 0 to  $\infty$ .

(c) Generate '2n' samples of Random Variable  $X$  from the given probability density function.



- (d) Now generate 'n' samples of  $Y = |X_{2n} - X_{2n-1}|$ .
- (e) Now find  $\bar{Y}$  which is mean of 'n' samples of  $Y$ .
- (f) Now calculate  $\sqrt{n}(e^{-\bar{Y}_n} - e^{-1})$  as result.
- (g) Now repeat the process for 'm' simulations so that we get m results.
- (h) Calculate the mean and the variance of the 'm' results obtained earlier using basic mean and variance formula.

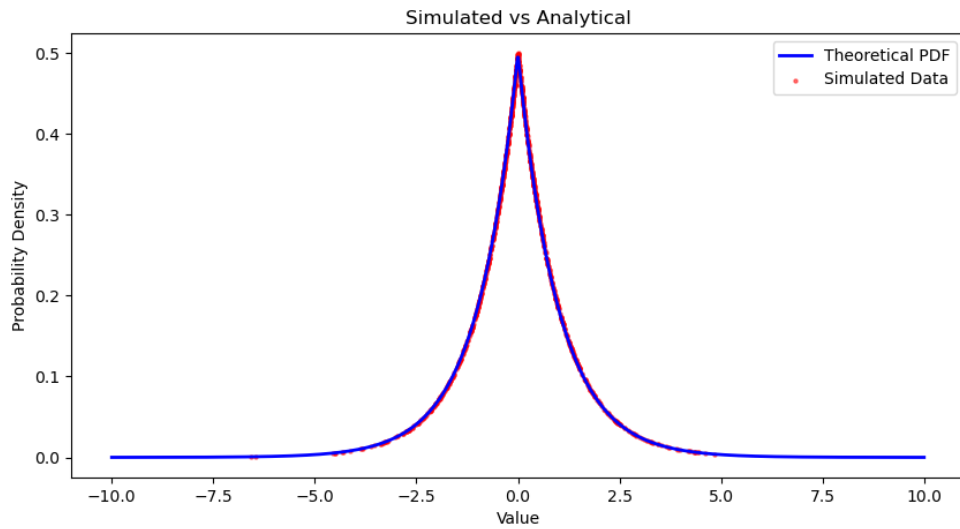


Figure 10.3: pdf of the laplacian

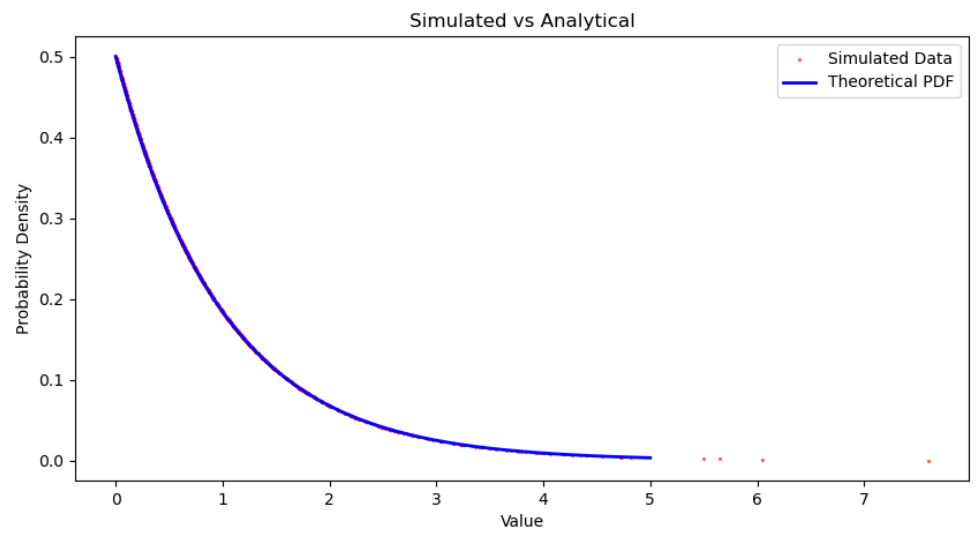


Figure 10.4: pdf of absolute of the laplacian

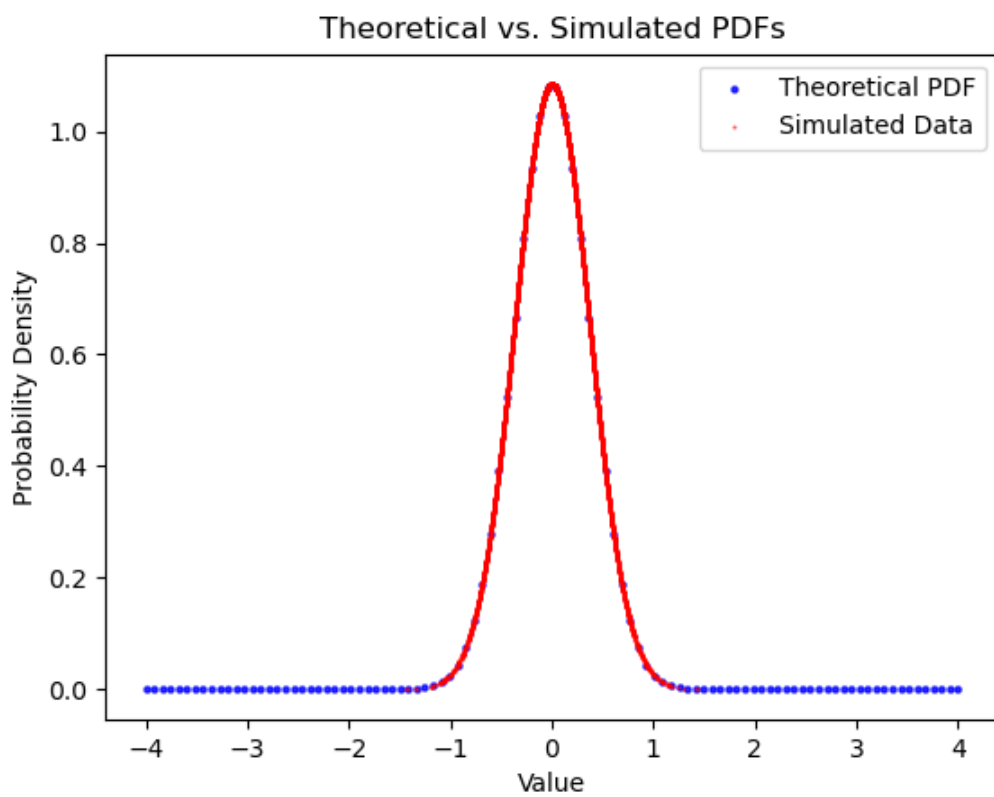


Figure 10.5: Gaussian pdf

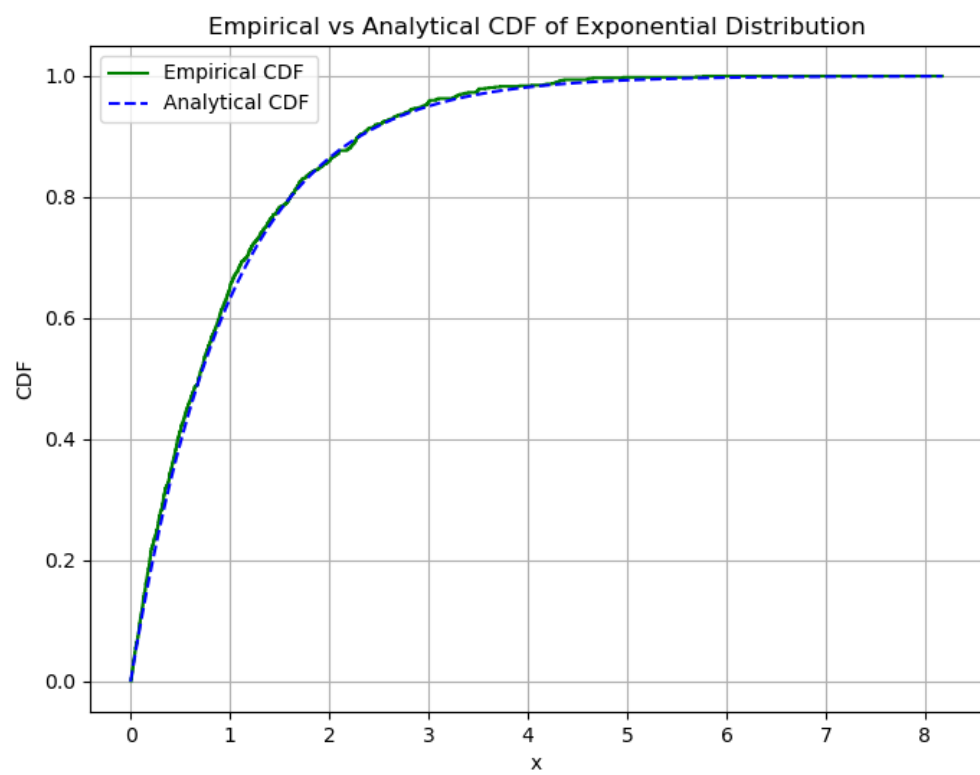


Figure 10.6: Cdf of X

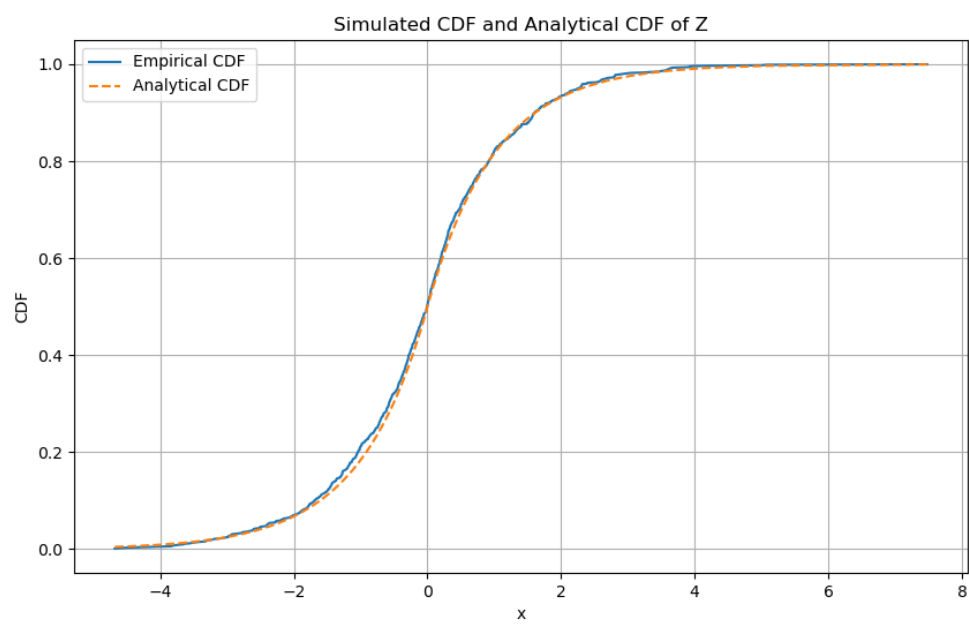


Figure 10.7: Cdf of Z

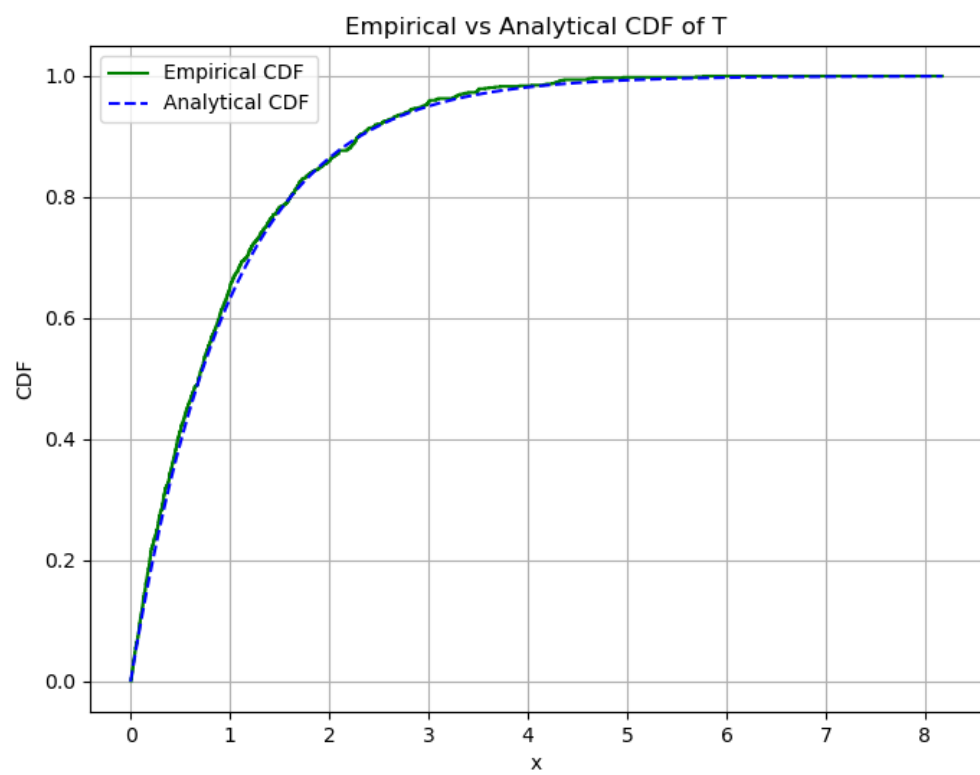


Figure 10.8: Cdf of T

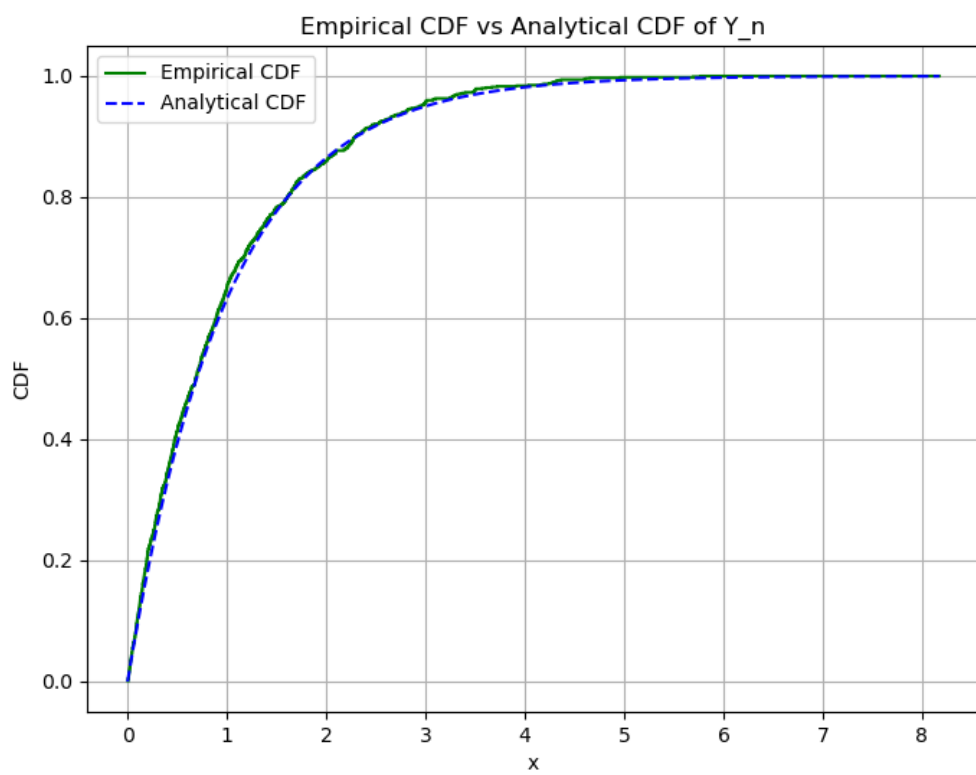


Figure 10.9: Cdf of  $Y_n$

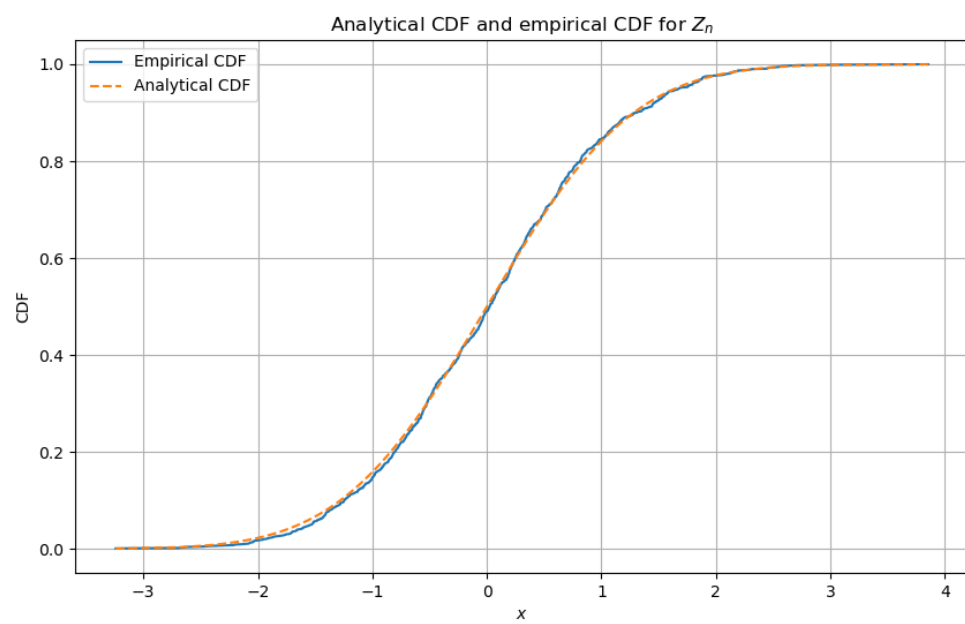


Figure 10.10: Cdf of  $Z_n$





## Chapter 11

# Information Theory

1. The frequency of occurrence of 8 symbols (a-h) is shown in the table below. A symbol is chosen and it is determined by asking a series of “yes/no” questions which are assumed to be truthfully answered. The average number of questions when asked in the most efficient sequence, to determine the chosen symbol, is

Symbols	Frequency of occurrence
a	$\frac{1}{2}$
b	$\frac{1}{4}$
c	$\frac{1}{8}$
d	$\frac{1}{16}$
e	$\frac{1}{32}$
f	$\frac{1}{64}$
g	$\frac{1}{128}$
h	$\frac{1}{128}$

**Solution:**

(GATE EC 2022)

Parameter	Value	Description
$X$	$1 \leq X \leq 8$	number of symbols
$l$	2	base of algorithm
$H(X)$	$\sum_i p_X(i) \log_l \left( \frac{1}{p_X(i)} \right)$	average number of question

$$H(X) = \sum_i p_X(i) \log_b \left( \frac{1}{p_X(i)} \right) \quad (11.1)$$

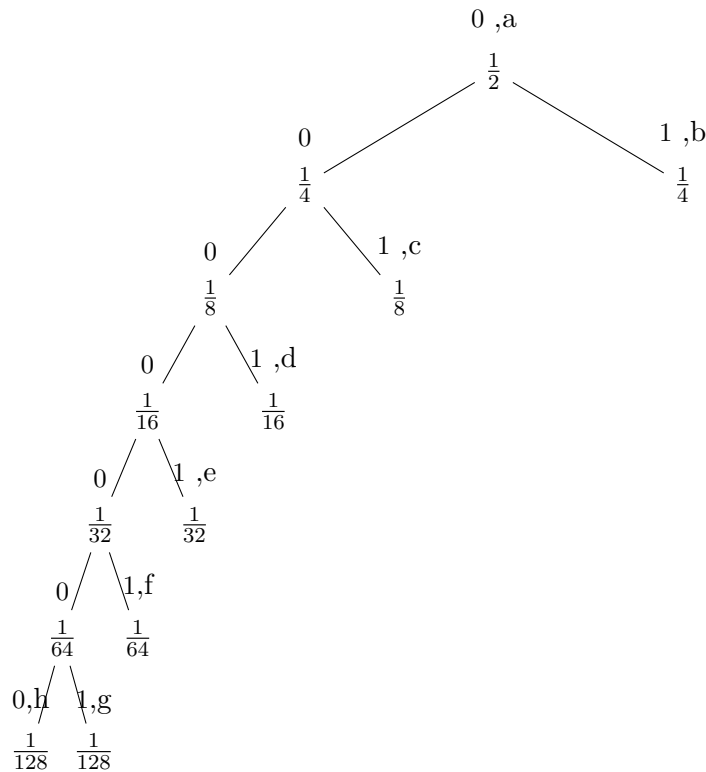
$$= \frac{1}{2} \log_2 (2) + \frac{1}{4} \log_2 (4) + \dots + \frac{1}{128} \log_2 (128) \quad (11.2)$$

$$= 0.5 + 0.5 + 0.375 + \dots + 0.0078125 \quad (11.3)$$

$$= 1.984375 \quad (11.4)$$

Now, finding the average using Huffman code, We start from a frequency of 1 and distribute it uniformly. The following conventions is used,

symbol	alloted bit
occured	1
not occured	0



Using the above binary table following code is generated;

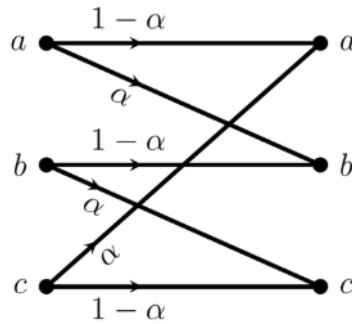
Symbols	Frequency	Code	Size
$a$	$\frac{1}{2}$	1	0.5
$b$	$\frac{1}{4}$	01	0.25
$c$	$\frac{1}{8}$	001	0.125
$d$	$\frac{1}{16}$	0001	0.0625
$e$	$\frac{1}{32}$	00001	0.03125
$f$	$\frac{1}{64}$	000001	0.015625
$g$	$\frac{1}{128}$	0000001	0.0078125
$h$	$\frac{1}{128}$	0000000	0.0078125

Table 11.1: Huffman table

The average number of question = Weighted path length = 1.9844

2. The transition diagram of a discrete memoryless channel with three input symbols and three output symbols is shown in the figure. The transition probabilities are as marked.

The parameter  $\alpha$  lies in the interval  $[0.25, 1]$ . The value of  $\alpha$  for which the capacity of this channel is maximized, is (GATE EC 2022) **Solution:**



Variable	Description	Value
$x_i$	Input	$x_0, x_1, x_2$
$y_i$	Output	$y_0, y_1, y_2$
$p_i$	Input probability	$p_0, p_1, p_2$
$q_i$	Output probability	$q_0, q_1, q_2$
$C$	Channel Capacity	$C$
$I$	Mutual Information	$I$
$H$	Entropy	$H$

$$C = \sup_{p_X(x)} I(X, Y) \quad (11.5)$$

$$I(X, Y) = \sum_{x,y} p(x, y) \log_2 \frac{p(x, y)}{p(x) p(y)} \quad (11.6)$$

$$= \sum_{x,y} p(x, y) \log_2 \frac{p(y|x)}{p(y)} \quad (11.7)$$

$$= - \sum_{x,y} p(x, y) \log_2 p(y) + \sum_{x,y} p(x, y) \log_2 p(y|x) \quad (11.8)$$

$$= - \sum_y p(y) \log_2 p(y) - \left( - \sum_{x,y} p(x, y) \log_2 p(y|x) \right) \quad (11.9)$$

$$= H(Y) - H(Y|X) \quad (11.10)$$

Now,

$$\sum_{i=0}^2 p_i = 1 \quad (11.11)$$

$$\sum_{j=0}^2 q_j = 1 \quad (11.12)$$

$$H(\mathbf{q}) = - \sum_{i=0}^2 q_i \log_2 q_i \quad (11.13)$$

$$= - (q_0 \log_2 q_0 + q_1 \log_2 q_1 + q_2 \log_2 q_2) \quad (11.14)$$

$$H(Y|X) = - \sum_{i=0}^2 \sum_{j=0}^2 p_i p_{Y|X}(y_j|x_i) \log_2 (p_{Y|X}(y_j|x_i)) \quad (11.15)$$

$$\begin{aligned} &= -p_0 ((1-\alpha) \log_2 (1-\alpha) + \alpha \log_2 \alpha) \\ &\quad - p_1 ((1-\alpha) \log_2 (1-\alpha) + \alpha \log_2 \alpha) \\ &\quad - p_2 ((1-\alpha) \log_2 (1-\alpha) + \alpha \log_2 \alpha) \end{aligned} \quad (11.16)$$

Using (11.14) and (11.16) in (11.10)

$$\begin{aligned} I(X, Y) &= - (q_0 \log_2 q_0 + q_1 \log_2 q_1 + q_2 \log_2 q_2) \\ &\quad + p_0 ((1-\alpha) \log_2 (1-\alpha) + \alpha \log_2 \alpha) \\ &\quad + p_1 ((1-\alpha) \log_2 (1-\alpha) + \alpha \log_2 \alpha) \\ &\quad + p_2 ((1-\alpha) \log_2 (1-\alpha) + \alpha \log_2 \alpha) \end{aligned} \quad (11.17)$$

$$\Rightarrow \frac{d}{d\alpha} I(X, Y) = p_0 \log_2 \left( \frac{\alpha}{1-\alpha} \right) + p_1 \log_2 \left( \frac{\alpha}{1-\alpha} \right) + p_2 \log_2 \left( \frac{\alpha}{1-\alpha} \right) \quad (11.18)$$

For Maxima or minima  $\frac{d}{d\alpha} I(X, Y) = 0$

$$\log_2 \left( \frac{\alpha}{1-\alpha} \right) (p_0 + p_1 + p_2) = 0 \quad (11.19)$$

$$\Rightarrow \alpha = \frac{1}{2} \quad (11.20)$$

3. let  $H(X)$  denote the entropy of a discrete random variable  $X$  taking  $K$  possible distinct real values. Which of the following statements is/are necessarily true?

(A)  $H(X) \leq \log_2 K$  bits

(B)  $H(X) \leq H(2X)$

(C)  $H(X) \leq H(X^2)$

(D)  $H(X) \leq H(2^X)$

**Solution:**

Random independent variable	value of R.V	Description
$X$	$X \in (x_1, x_2, \dots, x_K)$	Value of the discrete variable $X$

(a) For Option(A) we will find

We know that :

$$\begin{aligned} \max_{p_X(k)} \quad & H(X) \\ \text{s.t.} \quad & \sum_{k=0}^K p_X(k) = 1 \end{aligned}$$

$\Rightarrow$

$$\begin{aligned} \max_{p_X(k)} \quad & - \sum_{k=0}^K p_X(k) \log_2 p_X(k) \\ \text{s.t.} \quad & \sum_{k=0}^K p_X(k) = 1 \end{aligned}$$

Now, we use lagranges multiplier to find the maximum entropy subject to the lagranges multiplier constant  $\lambda$  and  $p_X(k)$

$$L(p_X(k), \lambda) = - \sum_{k=0}^K p_X(k) \log_2 p_X(k) + \lambda \left( \sum_{k=0}^K p_X(k) - 1 \right) \quad (11.21)$$

$$\frac{\partial L}{\partial p_X(k)} = - \log_2 p_X(k) - 1 + \lambda \quad (11.22)$$

Now, we take the derivative of L with respect to each  $p_X(k)$  equal to zero for  $H(X)_{max}$

$$\lambda = \log_2 \frac{2}{k} \quad (11.23)$$

$$p_X(k) = 1/K \quad (11.24)$$



On solving, we get the value of

$$H(X)_{max} = \log_2 K \quad (11.25)$$

$$H(X) \leq \log_2 K \quad (11.26)$$

Hence, Option(A) is correct

(b) Let's consider the discrete variable as follows

$X \in x_i$	$p_X(k)$
-1	$\frac{1}{4}$
0	$\frac{1}{2}$
1	$\frac{1}{4}$

$$H(X) = \frac{1}{4} \log_2 4 + \frac{1}{2} \log_2 2 + \frac{1}{4} \log_2 4 \quad (11.27)$$

$$H(X) = 1.5 \text{ units} \quad (11.28)$$

Now  $Y = 2X$

$Y \in y_i$	$p_Y(k)$
-2	$\frac{1}{4}$
0	$\frac{1}{2}$
2	$\frac{1}{4}$

$$H(Y) = \sum_{i=0}^2 p_Y(k) \log_2 \frac{1}{p_Y(k)} \quad (11.29)$$

$$H(Y) = 1.5 \text{ units} \quad (11.30)$$

$$H(Y) = H(2X) = H(X) \quad (11.31)$$

Hence, Option(B) is correct

(c) Similarly on substituting  $Y = X^2$

$Y \in y_i$	$p_Y(k)$
0	$\frac{1}{2}$
1	$\frac{1}{2}$

$$H(Y) = \sum_{i=0}^1 p_Y(k) \log_2 \frac{1}{p_Y(k)} \quad (11.32)$$

$$H(Y) = 1 \text{ units} \quad (11.33)$$

$$H(Y) = H(X^2) \leq H(X) \quad (11.34)$$

Hence, Option(C) is incorrect

(d) Now for  $Y = 2^X$

$Y \in y_i$	$p_Y(k)$
$2^{-1} = \frac{1}{2}$	$\frac{1}{4}$
$2^0 = 1$	$\frac{1}{2}$
$2^1 = 2$	$\frac{1}{4}$

$$H(Y) = \sum_{i=0}^2 p_Y(k) \log_2 \frac{1}{p_Y(k)} \quad (11.35)$$

$$H(Y) = 1.5 \text{ units} \quad (11.36)$$

$$H(Y) = H(2^X) = H(X) \quad (11.37)$$

Hence, Option(D) is correct

The ans is (A), (B), (D)

These options are correct for the particular example.

## Chapter 12

# Markov chain

12.1 Let  $X_{n \geq 1}$  be a Markov chain with state space  $\{1, 2, 3\}$  and transition probability matrix

$$\begin{pmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

Then  $\Pr(X_2 = 1 | X_1 = 1, X_3 = 2)$  equals

(GATE ST 2023)

**Solution:** Consider transition matrix as:

$$\begin{pmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{pmatrix} \quad (12.1)$$

$$\Pr(X_2 = 1 | X_1 = 1, X_3 = 2) = \Pr(X_2 = 1 | X_1 = 1) \quad (12.2)$$

$$= p_{11} \quad (12.3)$$

$$= 0.5 \quad (12.4)$$

(by markov's property and using transition probability matrix)

12.2 A state transition diagram with states  $A$ ,  $B$ , and  $C$ , and transition probabilities  $p_1, p_2, \dots, p_7$  is shown in the figure (e.g.,  $p_1$  denotes the probability of transition from state  $A$  to  $B$ ). For this state diagram, select the statement(s) which is/are universally true

(a)  $p_2 + p_3 = p_5 + p_6$

(b)  $p_1 + p_3 = p_4 + p_6$

(c)  $p_1 + p_4 + p_7 = 1$

(d)  $p_2 + p_5 + p_7 = 1$

**Solution:**

From, the given state diagram we can make a Transition matrix as:

$$M = \begin{pmatrix} p_7 & p_1 & p_4 \\ p_3 & p_2 & 0 \\ p_6 & 0 & p_5 \end{pmatrix} \quad (12.5)$$

And a valid transition matrix for a Markov Chain, the sum of rows should be 1 which gives:

$$p_7 + p_1 + p_4 = 1 \quad (12.6)$$

$$p_3 + p_2 = 1 \quad (12.7)$$

$$p_6 + p_5 = 1 \quad (12.8)$$

From (12.7) and (12.8) we have:

$$p_3 + p_2 = p_6 + p_5 \quad (12.9)$$

$\therefore$  Option (12.2a) and (12.2c) are correct.

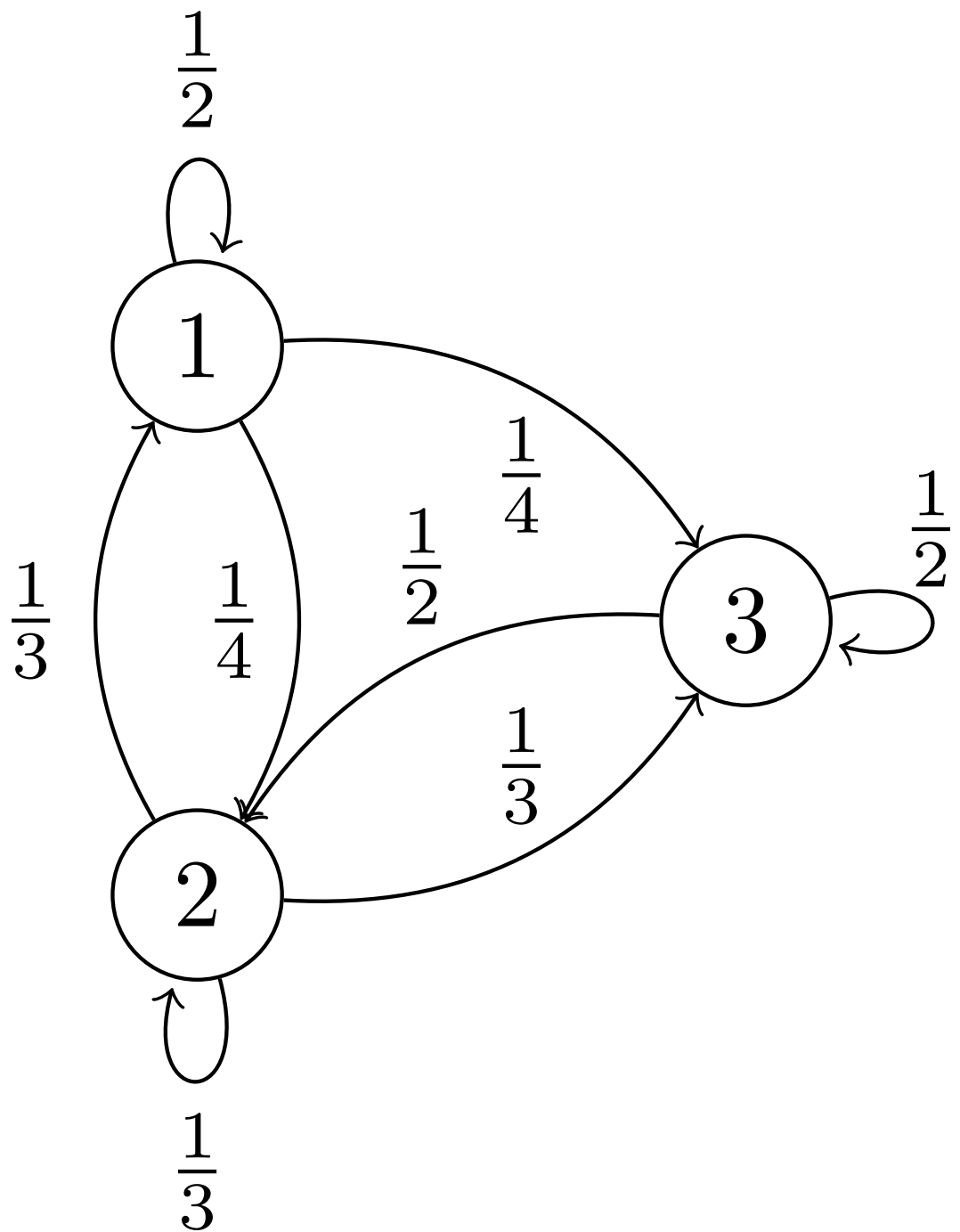


Figure 12.1: Markov Chain diagram

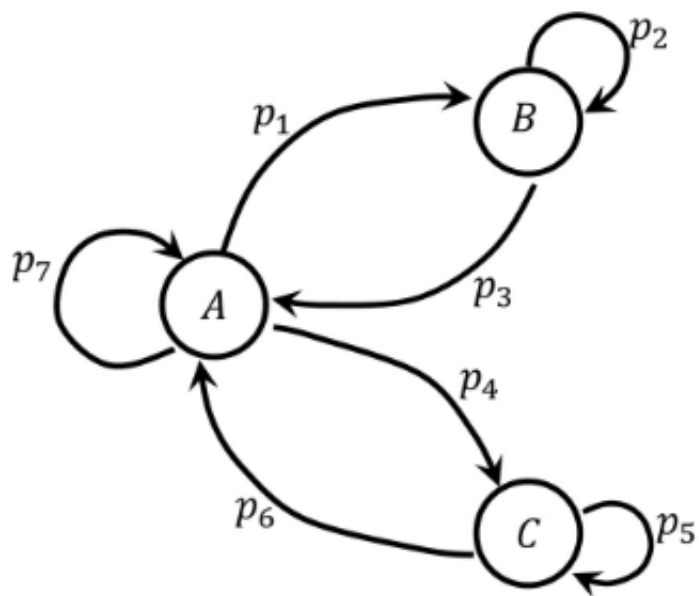


Figure 12.2: Figure1

# Chapter 13

## Estimation

13.1 Let  $\{-1, -\frac{1}{2}, 1, \frac{5}{2}, 3\}$  be a realization of a random sample of size 5 from a population having  $N(\frac{1}{2}, \sigma^2)$  distribution, where  $\sigma > 0$  is an unknown parameter. Let  $T$  be an unbiased estimator of  $\sigma^2$  whose variance attains the Cramer-Rao lower bound. Then, based on the above data, the realized value of  $T$  (rounded off to two decimal places) equals (GATE ST 2023)

**Definition 13.1:** Unbiased Estimator is defined as

$$E(\hat{\sigma}^2) = \sigma^2 \quad (13.1)$$

where,  $E(\hat{\sigma}^2)$  represents the expected value of the estimator  $\hat{\sigma}^2$  and  $\sigma^2$  represents the true parameter

**Definition 13.2:** The Cramér-Rao bound can be defined as follows:

$$\text{Var}(\sigma^2) \geq \frac{1}{I(\sigma^2)} \quad (13.2)$$

where  $I(\sigma^2)$  represents fisher information for the parameter  $\sigma^2$ . Mathematically,

$$I(\sigma^2) = -E \left[ \frac{\partial^2}{\partial(\sigma)^2} \log P_X(X|\sigma^2) \right]$$



where,  $E[\cdot]$  represents the expected value and  $P_X(X|\sigma^2)$  is the p.d.f of random variable  $X$  given the parameter  $\sigma^2$ .

$P_X(X|\sigma^2)$  is given by:

$$P_X(X|\sigma^2) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{(X - \frac{1}{2})^2}{2\sigma^2}\right) \quad (13.3)$$

$$\log p_X(X|\sigma^2) = \log\left(\frac{1}{2\pi\sigma^2} \exp\left(-\frac{(X - \frac{1}{2})^2}{2\sigma^2}\right)\right) \quad (13.4)$$

$$= -\frac{1}{2} \log(2\pi\sigma^2) - \frac{(X - \frac{1}{2})^2}{2\sigma^2} \quad (13.5)$$

$$\frac{\partial^2}{\partial(\sigma^2)^2} \log P_X(X|\sigma^2) = \frac{1}{2\pi\sigma^2} - \frac{3(X - \frac{1}{2})^2}{\sigma^4} \quad (13.6)$$

$$I(\sigma^2) = \frac{3}{\sigma^4} E[X^2] - \frac{3}{\sigma^4} E[X] + \frac{3}{4\sigma^4} - \frac{1}{2\pi\sigma^2} \quad (13.7)$$

$$E[X^2] = \sigma^2 + \left(\frac{1}{2}\right)^2 \quad (13.8)$$

$$E[X] = \frac{1}{2} \quad (13.9)$$

$$\Rightarrow I(\sigma^2) = \left(3 - \frac{1}{2\pi}\right) \frac{1}{\sigma^2} \quad (13.10)$$

Hence, Cramér-Rao bound is given as  $\frac{\sigma^2}{(3 - \frac{1}{2\pi})}$

**Definition 13.3:** Variance of  $T$  attains Cramer-Rao lower bound

$\Rightarrow T$  has attained minimum possible variance and  $T$  is an efficient estimator

$X_i$	-1	$-\frac{1}{2}$	1	$\frac{5}{2}$	3
$(X_i - \mu)^2$	$\frac{9}{4}$	1	$\frac{1}{4}$	4	$\frac{25}{4}$

Table 13.1: Table 1

Therefore,

$$T = \frac{\sum (X_i - \mu)^2}{n} \quad (13.11)$$

$$n = 5 \quad (13.12)$$

$$\mu = \frac{1}{2} \quad (13.13)$$

$$\sum (X_i - \mu)^2 = 13.75 \quad (13.14)$$

Hence,

$$T = 2.75 \quad (13.15)$$

Since,  $T$  is an unbiased estimator of  $\sigma^2$ ,

$$\text{Cramér-Rao bound} = \frac{T}{\left(3 - \frac{1}{2\pi}\right)} \quad (13.16)$$

$$= 0.968 \quad (13.17)$$

13.2 Let  $X$  be a random variable with probability density function

$$f(x; \lambda) = \begin{cases} \frac{1}{\lambda} e^{-\frac{x}{\lambda}} & \text{if } x > 0 \\ 0 & \text{otherwise} \end{cases} \quad (13.18)$$

where  $\lambda > 0$  is an unknown parameter. Let  $Y_1, Y_2, \dots, Y_n$  be a random sample of size  $n$

from a population having the same distribution as  $X^2$ . If

$$\bar{Y} = \frac{1}{n} \sum_{i=1}^n Y_i \quad (13.19)$$

then which of the following statements is true?

- (a)  $\sqrt{\frac{\bar{Y}}{2}}$  is a method of moments estimator of  $\lambda$
- (b)  $\sqrt{\bar{Y}}$  is a method of moments estimator of  $\lambda$
- (c)  $\frac{1}{2}\sqrt{\bar{Y}}$  is a method of moments estimator of  $\lambda$
- (d)  $2\sqrt{\bar{Y}}$  is a method of moments estimator of  $\lambda$  (GATE ST 2023)

**Solution:**

- (a) Using PDF in (13.18) we need to find an estimator for the unknown parameter  $\lambda$  in terms of sample mean  $\bar{Y}$

we know  $Y_i = X_i^2$  then,

$$E(Y_i) = E(X_i^2) \quad (13.20)$$

$$= \int_0^\infty x^2 \frac{1}{\lambda} e^{-\frac{x}{\lambda}} \quad (13.21)$$

$$= 2\lambda^2 \quad (13.22)$$

Method of moment is defined by (13.19) which gives,

$$\bar{Y} = E(Y_i) \quad (13.23)$$

$$= 2\lambda^2 \quad (13.24)$$

where

$$\lambda = \sqrt{\frac{\bar{Y}}{2}} \quad (13.25)$$

$\therefore$  Option (13.2a) is correct.

(b) The simulation steps to estimate  $\lambda$  using method of moment estimator in python.

- i. Generate a random value of  $\lambda$  within the specified range using **np.random.uniform(1**
- ii. Use the generated  $\lambda$  to create a random sample of  $X$  values following the given PDF using **np.random.exponential()**
- iii. Then, generate  $Y$  as  $Y = X^2$
- iv. calculate the mean ( $\bar{Y}$ ) as **np.mean(Y)**
- v. Hence, the estimated value of  $\lambda$  is **np.sqrt( $\frac{\bar{Y}}{2}$ )**

Graph of simulated CDF vs Theoretical CDF

13.3 Suppose from the estimation of a linear regression model

$$Y_i = \beta_0 + \beta_1 X_i + e_i$$

the residual sum of squares and the total sum of squares are obtained as 44 and 80, respectively. The value of coefficient of determination is (round off to two decimal places). (GATE XH 2023)

$$Y_i = \beta_0 + \beta_1 X_i + e_i \quad (13.26)$$

Here

**Definition 13.4:** Residual sum of squares(RSS):

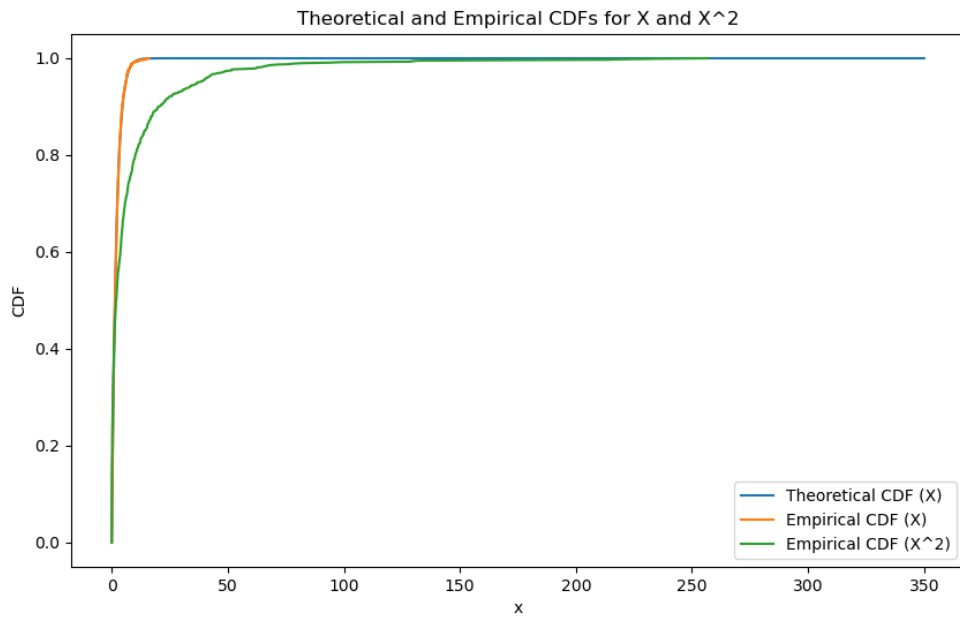


Figure 13.1: Figure1

It measures the extent of variability of observed data not predicted by the regression model. That is it estimates the variance in residual or error term's.

$$RSS = \sum (Y_i - \hat{Y}) \quad (13.27)$$

$$= \sum e_i^2 \quad (13.28)$$

$$= \sum (Y_i - \beta_0 - \beta_1 X_i)^2 \quad (13.29)$$

Here  $\hat{Y}$  = the value of Y on the line of regression.

**Definition 13.5:** Total sum of squares(TSS):

Table 13.2: Parameters

Parameters	Description
$Y_i$	Dependent variable
$X_i$	Independent variables
$\beta_0, \beta_1$	Constant variables
$e_i$	Error term

It measures the amount of variation measures in observed data. It is a measure of deviation from the mean. A low total sum of squares indicates little variation between data sets while a higher one indicates more variation.

$$TSS = \sum (Y_i - \bar{Y})^2 \quad (13.30)$$

where  $\bar{Y}$  = Mean of data

**Definition 13.6:** Coefficient of determination( $R^2$ ):

It is the proportion of the variance in the dependent variable that is predicted from the independent variable. It indicates the level of variation in the given data set.

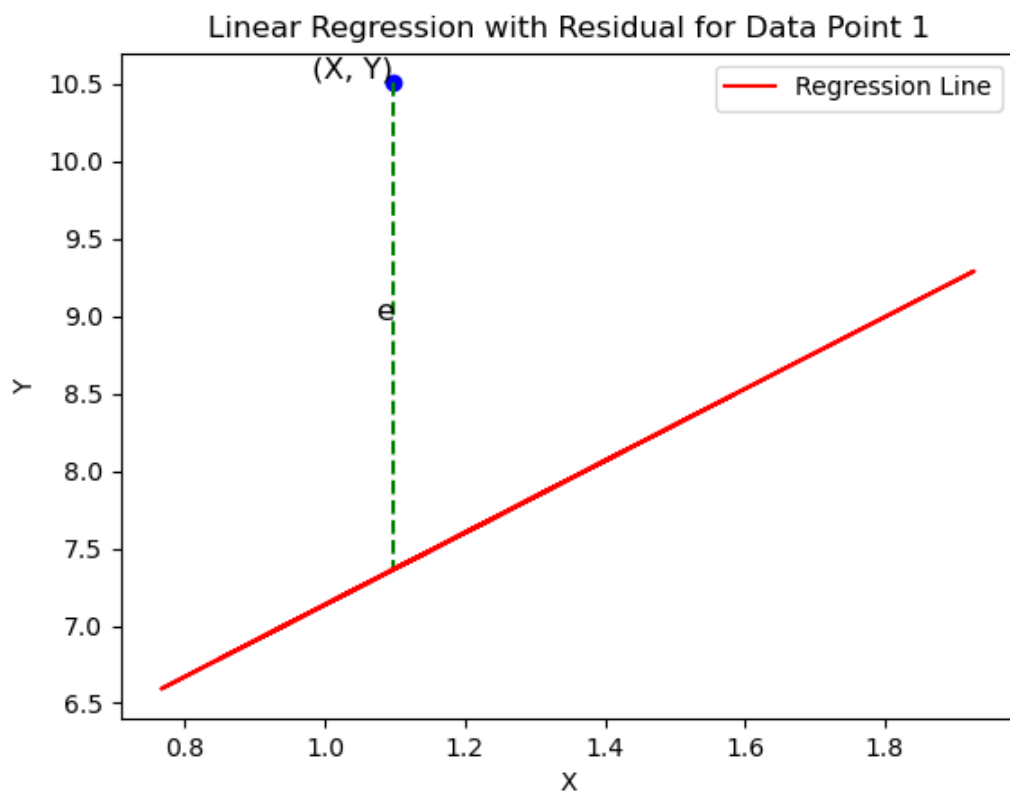
$$R^2 = 1 - \frac{RSS}{TSS} \quad (13.31)$$

$$= 1 - \frac{44}{80} \quad (13.32)$$

$$= 0.45 \quad (13.33)$$

45 percent of the variance in the Y variable is predicted from the X variable.

13.4 Consider the following regression model



$$y_k = \alpha_0 + \alpha_1 \log_e k + \epsilon_k, \quad k = 1, 2, \dots, n,$$

where  $\epsilon_k$ 's are independent and identically distributed random variables each having probability density function  $f(x) = \frac{1}{2}e^{-|x|}$ ,  $x \in \mathbb{R}$ . Then which one of the following statements is true?

- (A) The maximum likelihood estimator of  $\alpha_0$  does not exist
- (B) The maximum likelihood estimator of  $\alpha_1$  does not exist
- (C) The least squares estimator of  $\alpha_0$  exists and is unique
- (D) The least squares estimator of  $\alpha_1$  exists, but it is not unique

(GATE ST 2023)

**Solution:**

$$f(\epsilon_k) = \frac{1}{2}e^{-|\epsilon_k|} \quad (13.34)$$

$$\text{Likelihood function : } f(\epsilon_1 \epsilon_2 \dots \epsilon_n) = \prod_{k=1}^n f(\epsilon_k) \quad (13.35)$$

$$L = \prod_{k=1}^n \frac{1}{2}e^{-|\epsilon_k|} \quad (13.36)$$

$$L_1 = \ln L = \ln \left( \prod_{k=1}^n \frac{1}{2}e^{-|\epsilon_k|} \right) \quad (13.37)$$

$$= \sum_{k=1}^n \ln \left( \frac{1}{2}e^{-|\epsilon_k|} \right) \quad (13.38)$$

$$= \sum_{k=1}^n (-\ln 2 - |y_k - \alpha_0 - \alpha_1 \log_e k|) \quad (13.39)$$

$$= -n \ln 2 - \sum_{k=1}^n (|y_k - \alpha_0 - \alpha_1 \log_e k|) \quad (13.40)$$

$$L_1 = \text{function of } \alpha_0, \alpha_1 \quad (13.41)$$

(a) Maximum likelihood estimator

We need to find the value of  $\alpha_0$  and  $\alpha_1$  which will maximise the value of  $L_1$  i.e. the value of  $\alpha_0$  and  $\alpha_1$  which will minimise the value of  $\sum_{k=1}^n |y_k - \alpha_0 - \alpha_1 \log_e k|$

i. With respect to  $\alpha_0$

A. For  $y_k - \alpha_0 - \alpha_1 \log_e k > 0$

$$\begin{aligned} \min_{\alpha_0} \quad & y_k - \alpha_0 - \alpha_1 \log_e k \\ \text{s.t.} \quad & \alpha_0 \leq y_k - \alpha_1 \log_e k \end{aligned}$$



Using Lagrange multiplier method

$$L(\lambda) = y_k - \alpha_0 - \alpha_1 \log_e k - \lambda(\alpha_0 - y_k + \alpha_1 \log_e k) \quad (13.42)$$

$$\frac{\partial L}{\partial \alpha_0} = -1 - \lambda = 0 \quad (13.43)$$

$$\frac{\partial L}{\partial \lambda} = y_k - \alpha_0 - \alpha_1 \log_e k = 0 \quad (13.44)$$

$$\lambda = -1 \quad (13.45)$$

$$\alpha_0 = y_k - \alpha_1 \log_e k \quad (13.46)$$

B. For  $y_k - \alpha_0 - \alpha_1 \log_e k < 0$

$$\min_{\alpha_0} \quad -(y_k - \alpha_0 - \alpha_1 \log_e k)$$

$$\text{s.t.} \quad \alpha_0 \geq y_k - \alpha_1 \log_e k$$

Using Lagrange multiplier method

$$L(\lambda) = -(y_k - \alpha_0 - \alpha_1 \log_e k) - \lambda(\alpha_0 - y_k + \alpha_1 \log_e k) \quad (13.47)$$

$$\frac{\partial L}{\partial \alpha_0} = 1 - \lambda = 0 \quad (13.48)$$

$$\frac{\partial L}{\partial \lambda} = y_k - \alpha_0 - \alpha_1 \log_e k = 0 \quad (13.49)$$

$$\lambda = 1 \quad (13.50)$$

$$\alpha_0 = y_k - \alpha_1 \log_e k \quad (13.51)$$

As value of  $\alpha_0$  matches for both cases of modulus

$\therefore$  The maximum likelihood estimator of  $\alpha_0$  exist

ii. With respect to  $\alpha_1$

A. For  $y_k - \alpha_0 - \alpha_1 \log_e k > 0$

$$\begin{aligned} \min_{\alpha_1} \quad & y_k - \alpha_0 - \alpha_1 \log_e k \\ \text{s.t.} \quad & \alpha_1 \leq \frac{y_k - \alpha_0}{\log_e k} \end{aligned}$$

Using Lagrange multiplier method

$$L(\lambda) = y_k - \alpha_0 - \alpha_1 \log_e k - \lambda \left( \alpha_1 - \frac{y_k - \alpha_0}{\log_e k} \right) \quad (13.52)$$

$$\frac{\partial L}{\partial \alpha_1} = -\log_e k - \lambda = 0 \quad (13.53)$$

$$\frac{\partial L}{\partial \lambda} = - \left( \alpha_1 - \frac{y_k - \alpha_0}{\log_e k} \right) = 0 \quad (13.54)$$

$$\lambda = -\log_e k \quad (13.55)$$

$$\alpha_1 = \frac{y_k - \alpha_0}{\log_e k} \quad (13.56)$$

B. For  $y_k - \alpha_0 - \alpha_1 \log_e k < 0$

$$\begin{aligned} \min_{\alpha_1} \quad & -(y_k - \alpha_0 - \alpha_1 \log_e k) \\ \text{s.t.} \quad & \alpha_1 \geq \frac{y_k - \alpha_0}{\log_e k} \end{aligned}$$

Using Lagrange multiplier method

$$L(\lambda) = -(y_k - \alpha_0 - \alpha_1 \log_e k) - \lambda \left( \alpha_1 - \frac{y_k - \alpha_0}{\log_e k} \right) \quad (13.57)$$

$$\frac{\partial L}{\partial \alpha_1} = \log_e k - \lambda = 0 \quad (13.58)$$

$$\frac{\partial L}{\partial \lambda} = - \left( \alpha_1 - \frac{y_k - \alpha_0}{\log_e k} \right) = 0 \quad (13.59)$$

$$\lambda = \log_e k \quad (13.60)$$

$$\alpha_1 = \frac{y_k - \alpha_0}{\log_e k} \quad (13.61)$$

As value of  $\alpha_1$  matches for both cases of modulus

$\therefore$  The maximum likelihood estimator of  $\alpha_1$  exist

$\therefore$  Option (A) and (B) are incorrect

iii. Least square estimator

The least square estimator of  $\alpha_0$  and  $\alpha_1$  is  $\tilde{\alpha}_0$  and  $\tilde{\alpha}_1$  which will minimise

parameter	value	description
$\bar{y}$	$\frac{1}{n} \sum_{k=1}^n y_k$	Average value of $y_k$
$\bar{x}$	$\frac{1}{n} \sum_{k=1}^n \log_e k$	Average value of $\log_e k$

Table 13.3: Variables used

$$Q(\alpha_0, \alpha_1) = \sum_{k=1}^n (y_k - \alpha_0 - \alpha_1 \log_e k)^2 \quad (13.62)$$

$$\frac{\partial Q}{\partial \alpha_0} = -2 \sum_{k=1}^n (y_k - \alpha_0 - \alpha_1 \log_e k) = 0 \quad (13.63)$$

$$\sum_{k=1}^n (y_k - \alpha_0 - \alpha_1 \log_e k) = 0 \quad (13.64)$$

$$n\bar{y} - n\alpha_0 - \alpha_1 n\bar{x} = 0 \quad (13.65)$$

$$\implies \tilde{\alpha}_0 = \bar{y} - \tilde{\alpha}_1 \bar{x} \quad (13.66)$$

$$\frac{\partial Q}{\partial \alpha_1} = -2 \sum_{k=1}^n (y_k - \alpha_0 - \alpha_1 \log_e k) \log_e k = 0 \quad (13.67)$$

$$\implies \tilde{\alpha}_1 = \frac{\sum_{k=1}^n (\log_e k - \bar{x})(y_k - \bar{y})}{\sum_{k=1}^n (\log_e k - \bar{x})^2} \quad (13.68)$$

$\therefore$  Least square estimator of  $\alpha_0$  and  $\alpha_1$  exists and are unique

$\therefore$  Option (C) is correct and (D) is incorrect

iv. Steps for simulation the given distribution whose probability density function

is  $f(x) = \frac{1}{2}e^{-|x|}$

A. Write a function cdf for calculating the cdf of any random variable

$$p_X(x) = \begin{cases} \frac{1}{2}e^x & x \leq 0 \\ \frac{1}{2}e^{-x} & x > 0 \end{cases} \quad (13.69)$$

$$F_X(x) = \begin{cases} \int_{-\infty}^x \left(\frac{1}{2}e^x\right) dx & x \leq 0 \\ \int_{-\infty}^0 \left(\frac{1}{2}e^x\right) dx + \int_0^x \left(\frac{1}{2}e^{-x}\right) dx & x > 0 \end{cases} \quad (13.70)$$

$$F_X(x) = \begin{cases} \frac{1}{2}e^x & x \leq 0 \\ \frac{1}{2}(2 - e^{-x}) & x > 0 \end{cases} \quad (13.71)$$

B. Declare a function inverse cdf ( $I(u)$ ) such that its input is any random number and output is random variable whose cdf equals that of the given distribution

For  $x \leq 0$

$$u = \frac{1}{2}e^x \quad (13.72)$$

$$e^x = 2u \quad (13.73)$$

$$x = \ln 2u \quad (13.74)$$

$$\because x \leq 0 \quad (13.75)$$

$$u \leq 0.5 \quad (13.76)$$

For  $x > 0$

$$u = \frac{1}{2}(2 - e^{-x}) \quad (13.77)$$

$$2 - e^{-x} = 2u \quad (13.78)$$

$$e^{-x} = 2 - 2u \quad (13.79)$$

$$x = -\ln(2 - 2u) \quad (13.80)$$

$$\because x > 0 \quad (13.81)$$

$$u > 0.5 \quad (13.82)$$

$$I(u) = \begin{cases} \ln(2u) & u \leq 0.5 \\ -\ln(2-2u) & u > 0.5 \end{cases} \quad (13.83)$$

- C. Define three arrays `random_vars` , `cdf_values` , `theoretical_cdf_values` to store random variables, simulated cdf values and theoretical cdf values
- D. Generate random numbers using `rand()` and calling inverse cdf function to generate our random variable
- E. Calling cdf function to calculate the cdf of the generated random variable
- F. Storing the random variable, theoretical cdf and generated cdf into their respective arrays
- G. Storing the data of these three array into a .dat file
- H. Plotting these .dat file in python

13.5 Consider the following regression model

$$y_t = \alpha_0 + \alpha_1 t + \alpha_2 t^2 + \epsilon_t, \quad t = 1, 2, \dots, n \quad (13.84)$$

where  $\alpha_0$  ,  $\alpha_1$  and  $\alpha_2$  are unknown parameters and  $\epsilon_t$ 's are independent and identically distributed random variables each having  $\mathcal{N}(\mu, 1)$  distribution with  $\mu$  unknown. Then which of the following statements is/are true?

- (a) There exists an unbiased estimator of  $\alpha_1$
- (b) There exists an unbiased estimator of  $\alpha_2$
- (c) There exists an unbiased estimator of  $\alpha_0$
- (d) There exists an unbiased estimator of  $\mu$

(GATE ST 2023)

**Solution:** Assuming that the model is

$$y_t = \alpha_0 + \alpha_1 t + \alpha_2 t^2 + \epsilon_t \quad (13.85)$$

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & A_{11} & A_{12} \\ 1 & A_{21} & A_{22} \\ \vdots & & \\ 1 & A_{n1} & A_{n2} \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix} \quad (13.86)$$

Finding mean of the  $\mathbf{y}$ ,

$$\mathbf{y} = \mathbf{A}\mathbf{a} + \boldsymbol{\epsilon} \quad (13.87)$$

$$E(\mathbf{y}) = E(\mathbf{A}\mathbf{a} + \boldsymbol{\epsilon}) \quad (13.88)$$

$$= \mathbf{A}\mathbf{a} + E(\boldsymbol{\epsilon}) \quad (13.89)$$

$$E(\boldsymbol{\epsilon}) = \begin{bmatrix} E(\epsilon_1) \\ E(\epsilon_2) \\ \cdot \\ \cdot \\ E(\epsilon_n) \end{bmatrix} = \begin{bmatrix} \mu \\ \mu \\ \cdot \\ \cdot \\ \mu \end{bmatrix} \quad (13.90)$$

$$\boldsymbol{\mu}_\epsilon = \mu \begin{bmatrix} 1 \\ 1 \\ \cdot \\ \cdot \\ 1 \end{bmatrix} \quad (13.91)$$

Let  $\mathbf{A}_i$  represent the  $i^{th}$  row of  $\mathbf{A}$

$$\text{Cov}(y_i, y_j) = E((y_i - E(y_i))(y_j - E(y_j))) \quad (13.92)$$

$$= E((\mathbf{A}_i \mathbf{a} + \epsilon_i - E(\mathbf{A}_i \mathbf{a} + \epsilon_i))(\mathbf{A}_j \mathbf{a} + \epsilon_j - E(\mathbf{A}_j \mathbf{a} + \epsilon_j))) \quad (13.93)$$

$$= E((\epsilon_i - E(\epsilon_i))(\epsilon_j - E(\epsilon_j))) \quad (13.94)$$

$$= \text{Cov}(\epsilon_i, \epsilon_j) \quad (13.95)$$

$$\mathbf{C}_y = \mathbf{C}_\epsilon \quad (13.96)$$

Since the  $\epsilon_i$ 's are independent and identical vectors,

$$\text{Cov}(\epsilon_i, \epsilon_j) = 0, \quad \forall i \neq j \quad (13.97)$$



$$\text{Var}(\epsilon_i) = 1, \quad \forall 1 \leq i \leq n \quad (13.98)$$

$$\mathbf{C}_{\mathbf{y}} = I_{n \times n} \quad (13.99)$$

$$\mathbf{y} \sim \mathcal{N}(\mathbf{A}\mathbf{a} + \boldsymbol{\mu}_{\epsilon}, I) \quad (13.100)$$

$$p_{\mathbf{y}} = \frac{1}{\sqrt{(2\pi)^n \det(\mathbf{C}_{\mathbf{y}})}} \exp \frac{-(\mathbf{y} - \boldsymbol{\mu}_{\epsilon} - \mathbf{A}\mathbf{a})^T \mathbf{C}_{\mathbf{y}}^{-1} (\mathbf{y} - \boldsymbol{\mu}_{\epsilon} - \mathbf{A}\mathbf{a})}{2} \quad (13.101)$$

where  $\mathbf{C}_{\mathbf{y}}$  is the covariance matrix for  $\mathbf{y}$

The maximum likelihood function can be written as:

$$L(\mathbf{y}) = \frac{1}{(2\pi)^{\frac{n}{2}}} e^{\frac{-(\mathbf{y} - \boldsymbol{\mu}_{\epsilon} - \mathbf{A}\mathbf{a})^T (\mathbf{y} - \boldsymbol{\mu}_{\epsilon} - \mathbf{A}\mathbf{a})}{2}} \quad (13.102)$$

$$\ln L(\mathbf{y}) = -\frac{n}{2} \ln(2\pi) - \frac{1}{2} (\mathbf{y} - \boldsymbol{\mu}_{\epsilon} - \mathbf{A}\mathbf{a})^T (\mathbf{y} - \boldsymbol{\mu}_{\epsilon} - \mathbf{A}\mathbf{a}) \quad (13.103)$$

$$\frac{\partial \ln L(\mathbf{y})}{\partial \mathbf{a}} = \frac{\partial (-\mathbf{y}^T \mathbf{A}\mathbf{a} - \mathbf{a}^T \mathbf{A}^T \mathbf{y} + \mathbf{a}^T \mathbf{A}^T \mathbf{A}\mathbf{a} + \mathbf{a}^T \mathbf{A}^T \boldsymbol{\mu}_{\epsilon} + \boldsymbol{\mu}_{\epsilon}^T \mathbf{A}\mathbf{a})}{\partial \mathbf{a}} \quad (13.104)$$

$$= -2\mathbf{A}^T \mathbf{y} + 2\mathbf{A}^T \mathbf{A}\mathbf{a} + 2\mathbf{A}^T \boldsymbol{\mu}_{\epsilon} \quad (13.105)$$

The normal equation is

$$\frac{\partial \ln L(\mathbf{y})}{\partial \mathbf{a}} = 0 \quad (13.106)$$

$$\mathbf{a} = (\mathbf{A}^T \mathbf{A})^{-1} (\mathbf{A}^T \mathbf{y} - \mathbf{A}^T \boldsymbol{\mu}_{\epsilon}) \quad (13.107)$$

For unbiased estimator,

$$E(\mathbf{a}) = \mathbf{a} \quad (13.108)$$

$$E(\mathbf{a}) = E\left((\mathbf{A}^T \mathbf{A})^{-1} (\mathbf{A}^T \mathbf{y} - \mathbf{A}^T \boldsymbol{\mu}_\epsilon)\right) \quad (13.109)$$

$$= (\mathbf{A}^T \mathbf{A})^{-1} (\mathbf{A}^T E(\mathbf{y}) - \mathbf{A}^T \boldsymbol{\mu}_\epsilon) \quad (13.110)$$

$$= (\mathbf{A}^T \mathbf{A})^{-1} (\mathbf{A}^T (\mathbf{A} \mathbf{a} + \boldsymbol{\mu}_\epsilon) - \mathbf{A}^T \boldsymbol{\mu}_\epsilon) \quad (13.111)$$

$$= (\mathbf{A}^T \mathbf{A})^{-1} (\mathbf{A}^T \mathbf{A}) \mathbf{a} \quad (13.112)$$

$$= \mathbf{a} \quad (13.113)$$

Hence there exist unbiased estimator for  $\alpha_0, \alpha_1, \alpha_2$

For Maximum Likelihood Estimator of  $\mu$

$$\frac{\partial \ln L(\mathbf{y})}{\partial \boldsymbol{\mu}_\epsilon} = \frac{\partial (-\mathbf{y}^T \boldsymbol{\mu}_\epsilon - \mathbf{a}^T \mathbf{A}^T \boldsymbol{\mu}_\epsilon + \boldsymbol{\mu}_\epsilon \mathbf{y} + \boldsymbol{\mu}_\epsilon \mathbf{A} \mathbf{x} + \boldsymbol{\mu}_\epsilon^T \boldsymbol{\mu}_\epsilon)}{\partial \boldsymbol{\mu}_\epsilon} \quad (13.114)$$

$$= -2\mathbf{y} + 2\mathbf{A} \mathbf{a} + 2\boldsymbol{\mu}_\epsilon \quad (13.115)$$

The normal equation is

$$\frac{\partial \ln L(\mathbf{y})}{\partial \boldsymbol{\mu}_\epsilon} = 0 \quad (13.116)$$

$$\boldsymbol{\mu}_\epsilon = \mathbf{y} - \mathbf{A} \mathbf{a} \quad (13.117)$$

$$E(\boldsymbol{\mu}_\epsilon) = E(\mathbf{y} - \mathbf{A}\mathbf{a}) \quad (13.118)$$

$$= E(\mathbf{y}) - \mathbf{A}\mathbf{a} \quad (13.119)$$

$$= E(\boldsymbol{\epsilon}) \quad (13.120)$$

$$= \boldsymbol{\mu}_\epsilon \quad (13.121)$$

Since,

$$\boldsymbol{\mu}_\epsilon = \mu \begin{bmatrix} 1 \\ 1 \\ \cdot \\ \cdot \\ \cdot \\ 1 \end{bmatrix} \quad (13.122)$$

Hence there exists an unbiased estimator for  $\mu$  as well.

Simulation in C We get the following values:

Parameter	True Value	Simulated Value	Average Bias
$\alpha_0$	3	2.983275	0.016725
$\alpha_1$	2	2.002000	-0.002000
$\alpha_2$	0.5	0.500500	-0.000500
$\mu$	0.1	0.1	0

Table 13.4: Simulation results

- Defining some arbitrary values for  $\alpha_0, \alpha_1, \alpha_2$  and  $\mu$
- Generating the epsilon using Box-Muller method and using it to find  $y_i$
- Generating dependent variable  $y_i$  using  $y_i = \alpha_0 + \alpha_1 i + \alpha_2 i^2 + \epsilon_i, \quad i = 1, 2, \dots, n$

- (d) Storing the generated data into a *generated\_data.dat* file.
- (e)  $\epsilon_i$  is generated as a gaussian variable
- (f) Defining 4 variables *beta<sub>0</sub>hat*, *beta<sub>1</sub>hat*, *beta<sub>2</sub>hat* and *mu-hat* for storing the estimated values
- (g) Defining 4 variables *beta<sub>0</sub>bias*, *beta<sub>1</sub>bias*, *beta<sub>2</sub>bias* and *mu-bias* for keeping track of the deviation of the estimated values from the true values
- (h) Using nested for loops to ensure that average bias for each coefficient is calculated over multiple simulations
- (i) Dividing the total bias by the number of simulations to get the average bias per simulation

13.6 Let  $X_1, X_2, \dots, X_n$  be a random sample of size  $n$  from a population having uniform distribution over the interval  $(\frac{1}{3}, \theta)$ , where  $\theta > \frac{1}{3}$  is an unknown parameter. If  $Y = \max \{X_1, X_2, \dots, X_n\}$ , then which one of the following statements is true?

- (a)  $(\frac{n+1}{n})(Y - \frac{1}{3}) + \frac{1}{3}$  is an unbiased estimator of  $\theta$
- (b)  $(\frac{n}{n+1})(Y - \frac{1}{3}) + \frac{1}{3}$  is an unbiased estimator of  $\theta$
- (c)  $(\frac{n+1}{n})(Y + \frac{1}{3}) - \frac{1}{3}$  is an unbiased estimator of  $\theta$
- (d)  $Y$  is an unbiased estimator of  $\theta$

(GATE ST 2023)

**Solution:**

Any one of the above expressions is an unbiased estimator of  $\theta$  if

$$E(\text{estimator}) = \theta$$

We will first calculate expected value of  $Y$

The PDF of  $X_i$  can be expressed as

$$p_{X_i}(x) = \frac{1}{\theta - \frac{1}{3}}, \frac{1}{3} < \theta < 3 \quad (13.123)$$

$$F_{X_i}(y) = \int_{\frac{1}{3}}^y p_{X_i}(x) dx \quad (13.124)$$

$$= \int_{\frac{1}{3}}^y \frac{1}{\theta - \frac{1}{3}} dx \quad (13.125)$$

$$= \frac{y - \frac{1}{3}}{\theta - \frac{1}{3}} \quad (13.126)$$

$$F_X(y) = \Pr(X \leq y) \quad (13.127)$$

$$= \Pr(X_1 \leq y, X_2 \leq y, \dots, X_n \leq y) \quad (13.128)$$

$$= \Pr(X_1 \leq y) \Pr(X_2 \leq y) \dots \Pr(X_n \leq y) \quad (13.129)$$

$$= \left( \frac{y - \frac{1}{3}}{\theta - \frac{1}{3}} \right)^n \quad (13.130)$$

$$p_X(y) = \frac{d}{dy} F_X(y) \quad (13.131)$$

$$= \frac{n}{\left(\theta - \frac{1}{3}\right)^n} \left(y - \frac{1}{3}\right)^{n-1} \quad (13.132)$$

$$E(Y) = \int_{\frac{1}{3}}^{\theta} yp_X(y)dy \quad (13.133)$$

$$= \int_{\frac{1}{3}}^{\theta} y \frac{n}{(\theta - \frac{1}{3})^n} \left(y - \frac{1}{3}\right)^n dy \quad (13.134)$$

$$= \frac{n}{(\theta - \frac{1}{3})^n} \int_{\frac{1}{3}}^{\theta} y \left(y - \frac{1}{3}\right)^{n-1} dy \quad (13.135)$$

$$\text{Let } y - \frac{1}{3} = t \quad (13.136)$$

$$\implies y = t + \frac{1}{3} \implies dy = dt \quad (13.137)$$

$$\text{Therefore,} \quad (13.138)$$

$$E(Y) = \frac{n}{(\theta - \frac{1}{3})^n} \int_0^{\theta - \frac{1}{3}} \left(t^n + \frac{t^{n-1}}{3}\right) dt \quad (13.139)$$

$$= \frac{n}{(\theta - \frac{1}{3})^n} \left( \frac{(\theta - \frac{1}{3})^{n+1}}{n+1} + \frac{(\theta - \frac{1}{3})^n}{3n} \right) \quad (13.140)$$

$$= \frac{n}{n+1} \left(\theta - \frac{1}{3}\right) + \frac{1}{3} \quad (13.141)$$

$$= \frac{3n\theta + 1}{3(n+1)} \neq \theta \quad (13.142)$$

Therefore, fourth option is incorrect.

(a)

$$E\left(\left(\frac{n+1}{n}\right) \left(Y - \frac{1}{3}\right) + \frac{1}{3}\right) = \frac{n+1}{n} E(Y) - \frac{n+1}{3n} + \frac{1}{3} \quad (13.143)$$

$$= \frac{3n\theta + 1 - (n+1) + n}{3n} \quad (13.144)$$

$$= \theta \quad (13.145)$$

It is the unbiased estimator of  $\theta$ .

(b) Similarly,

$$E\left(\left(\frac{n}{n+1}\right)\left(Y - \frac{1}{3}\right) + \frac{1}{3}\right) = \frac{n}{n+1}E(Y) - \frac{n}{3(n+1)} + \frac{1}{3} \quad (13.146)$$

$$= \frac{n(3n\theta + 1) - n(n+1) + (n+1)^2}{3(n+1)^2} \quad (13.147)$$

$$= \frac{3n^2 + n}{3(n+1)^2} \neq \theta \quad (13.148)$$

It is not an unbiased estimator of  $\theta$ .

(c) In the same way,

$$E\left(\left(\frac{n+1}{n}\right)\left(Y + \frac{1}{3}\right) - \frac{1}{3}\right) = \frac{n+1}{n}E(Y) + \frac{n+1}{3n} - \frac{1}{3} \quad (13.149)$$

$$= \frac{3n\theta + 1 + n + 1 - n}{3n} \quad (13.150)$$

$$= \frac{3n\theta + 2}{3n} \neq \theta \quad (13.151)$$

It is not an unbiased estimator of  $\theta$ .

Hence, the first option is the correct option.

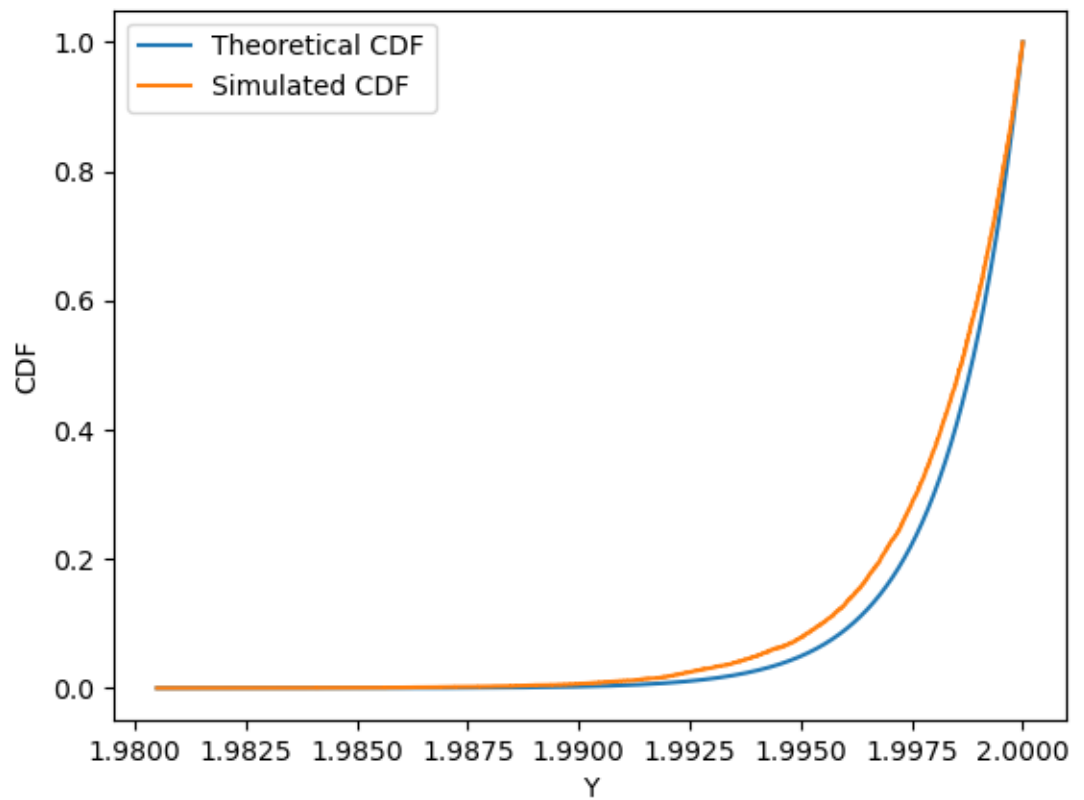


Figure 13.2: Theoretical and Simulated CDFs

Steps for the Simulation using C:

- (a) Include Header Files: Start the C program by including necessary header files, particularly '`<stdio.h>`' to perform standard input and output functions.
- (b) Declare Variables: Declare variables for sample size, number of simulations and  $\theta$ .
- (c) Open the file: Use 'fopen' to open a file for writing the simulated data. The file is named 'estimations.txt'.
- (d) Simulate data: Use loop to perform multiple simulations. In each iteration, gen-



erate random samples, calculate the estimator, and store the estimations in an array.

- (e) Write data to file: Inside the loop, use ‘fprintf’ to write each estimation to the opened file.
- (f) Close the file: After all simulations are complete and data is written, close the file using ‘fclose’ to save the data.

Python code for Plotting and Analysis:

Steps:

- (a) Include Required Libraries: In Python code, include libraries like NumPy for data manipulation and Matplotlib for plotting.
- (b) Load data: Use NumPy to load the data from the file ‘estimations.txt’ into a python array or NumPy array.
- (c) Define theoretical CDF: Define the function that calculates the theoretical CDF.
- (d) Sort the data: Sort the loaded data in ascending order.
- (e) Calculate the empirical CDF: Calculate the empirical CDF for the sorted data by dividing the rank of each value by the total number of data points.
- (f) Plot the CDFs: Use Matplotlib to create a plot that compares the theoretical and simulated CDFs.

13.7 Let  $X_1, X_2, \dots, X_n$  be a random sample of size  $n$  from a population having probability density function

$$p_X(x; \mu) = \begin{cases} e^{-(x-\mu)}, & \text{if } \mu \leq x < \infty \\ 0, & \text{otherwise,} \end{cases} \quad (13.152)$$

where  $\mu \in \mathbb{R}$  is an unknown parameter. If  $\hat{M}$  is the maximum likelihood estimator of

the median of  $X_1$ , then which one of the following statements is true?

A)  $\Pr(\hat{M} \leq 2) = 1 - e^{-n(1-\log_e 2)}$  if  $\mu = 1$

B)  $\Pr(\hat{M} \leq 1) = 1 - e^{-n \log_e 2}$  if  $\mu = 1$

C)  $\Pr(\hat{M} \leq 3) = 1 - e^{-n(1-\log_e 2)}$  if  $\mu = 1$

D)  $\Pr(\hat{M} \leq 4) = 1 - e^{-n(2 \log_e 2 - 1)}$  if  $\mu = 1$

(GATE ST 2023)

**Solution:** For continuous random variable X, median M is such that,

$$\Pr(X \leq M) = 0.5 \quad (13.153)$$

The pdf of X is given by,

$$p_X(x) = \begin{cases} e^{-(x-\mu)}, & \text{if } \mu \leq x < \infty \\ 0, & \text{otherwise,} \end{cases} \quad (13.154)$$

Hence, cdf is given by

$$F_X(x; \mu) = \int_{\mu}^x e^{-(t-\mu)} dt \quad (13.155)$$

$$= e^{\mu}[-e^{-x} + e^{-\mu}] \quad (13.156)$$

$$= 1 - e^{-(x-\mu)} \quad (13.157)$$

Now,

$$F_X(x; \mu) = 0.5 \quad (13.158)$$

$$\implies 1 - e^{-(M-\mu)} = 0.5 \quad (13.159)$$

$$\implies \hat{M} = \mu + \ln(2) \quad (13.160)$$

**Definition 13.7:**  $\hat{L}$ , the Maximum Likelihood Estimator of the distribution is given by,

$$L = \prod e^{-(x-\mu)} \quad (13.161)$$

$$= e^{-(\sum x_i - n\mu)} \quad (13.162)$$

$$(13.163)$$

For the Likelihood function to be maximum,  $\sum x_i - n\mu$  should be minimum Hence,

$$X_i > \mu \quad (13.164)$$

$$\implies \sum x_i > n\mu \quad (13.165)$$

$$\sum x_i - n\mu > 0 \quad (13.166)$$

$$\implies \mu = \frac{\sum x_i}{n} \quad (13.167)$$

Given,

$$p_X(x) = e^{-(x-\mu)} \quad (13.168)$$

$$(13.169)$$

$X_i$  follows an exponential distribution.

**Definition 13.8:** We know that if,

$$p_X(x) = \lambda_i e^{-\lambda_i x} \quad (13.170)$$

$$S = X_1 + X_2 + \dots + X_n \quad (13.171)$$

$$p_S(n) = \frac{\lambda^n x^{n-1} e^{-\lambda x}}{(n-1)!} \quad (13.172)$$

which is gamma distribution with parameters  $n$  and  $\lambda$

Hence, pdf of

$$Y = \sum_{i=1}^n x_i \quad (13.173)$$

$$\text{where,} \quad (13.174)$$

$$p_X(x) = e^{-(x-\mu)} \quad (13.175)$$

will follow gamma distribution with parameter  $n$  and  $\lambda = 1$ , given by,

$$p_Y(x; n, 1) = \frac{x^{n-1} e^{-x}}{(n-1)!} \quad (13.176)$$

Hence, cdf is given by,

$$F_Y(x; n) = \int_1^x \frac{t^{n-1} e^{-t}}{(n-1)!} dt \quad (13.177)$$

$$= 1 - \Gamma(n, x) \quad (13.178)$$

where  $\Gamma(n, x)$  is incomplete gamma function. Thus,

$$\Pr(\hat{M} \leq k) = \Pr\left(\frac{\sum x_i}{n} + \ln(2) \leq k\right) \quad (13.179)$$

$$= \Pr(Y/n + \ln(2) \leq k) \quad (13.180)$$

$$= \Pr(Y \leq n(k - \ln(2))) \quad (13.181)$$

$$= F_Y(n(k - \ln(2))) \quad (13.182)$$

$$= 1 - \Gamma(n, n(k - \ln(2))) \quad (13.183)$$

It needs a value of n to be computed.

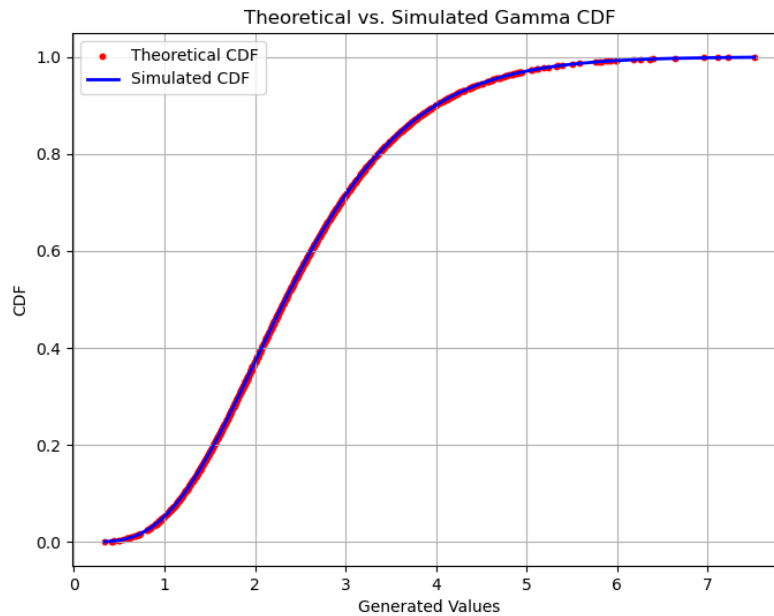


Figure 13.3: Verifying gamma cdf through simulation

Steps for simulation in C:

- (a) Import the necessary libraries, including 'stdio.h', 'stdlib.h' and 'math.h'.

- (b) Write functions for generating exponential distribution, gamma pdf and gamma cdf.
- (c) In the main function, the exponentials generated using the function are added and stored in variable 'gamma samples' for each sample, to get sum of exponentials.
- (d) Then the simulated gamma cdf values are calculated using 'gamma cdf' function and the calculated sum of exponentials.
- (e) The code is then compiled using GCC compiler in the terminal (`gcc simulation.c -o simulation -lm`), and the results are stored in an output.txt file. (`./simulation>output.txt`)
- (f) The output file is loaded into the python code with theoretical cdf values, and the final graph between theoretical and simulated cdf is plotted.
- (g) The simulated and theoretical cdf values match, which verifies the gamma cdf through simulation.

13.8 Let  $X_1, X_2, \dots, X_{10}$  be a random sample of size 10 from a population having  $\mathcal{N}(0, \theta^2)$  distribution, where  $\theta > 0$  is an unknown parameter. Let  $T = \frac{1}{10} \sum_{i=1}^{10} X_i^2$ . If the mean square error of  $cT$  ( $c > 0$ ) as an estimator of  $\theta^2$ , is minimized at  $c = c_0$ , then the value of  $c_0$  equals

- (i)  $\frac{5}{6}$
- (ii)  $\frac{2}{3}$
- (iii)  $\frac{3}{5}$
- (iv)  $\frac{1}{2}$

**Solution:** The mean of  $T$  is given by,

$$E(T) = E\left(\frac{1}{10} \sum_{i=1}^{10} X_i^2\right) \quad (13.184)$$

$$= \frac{1}{10} \sum_{i=1}^{10} E(X_i^2) \quad (13.185)$$

Since

$$E(X^2) = V(X) + (E(X))^2 \quad (13.186)$$

Using (13.186)

$$E(T) = \frac{1}{10} \sum_{i=1}^{10} V(X_i) + E(X_i)^2 \quad (13.187)$$

$$= \frac{1}{10} (10\theta^2) \quad (13.188)$$

$$= \theta^2 \quad (13.189)$$

Similarly, using (13.186), the variance of  $T$  is given by,

$$V(T) = E(T^2) - E(T)^2 \quad (13.190)$$

$$= \frac{1}{100} \left( E\left(\left(\sum_{i=1}^{10} X_i^2\right)^2\right) - \left(E\left(\sum_{i=1}^{10} X_i^2\right)\right)^2 \right) \quad (13.191)$$

But

$$E \left( \left( \sum_{i=1}^{10} X_i^2 \right)^2 \right) = E \left( \sum_{i=1}^{10} \sum_{j=1}^{10} X_i^2 X_j^2 \right) \quad (13.192)$$

$$= \sum_{i=1}^n \sum_{j=1}^n E (X_i^2 X_j^2) \quad (13.193)$$

and

$$\left( E \left( \sum_{i=1}^{10} X_i \right) \right)^2 = \left( \sum_{i=1}^{10} E (X_i) \right)^2 \quad (13.194)$$

$$= \sum_{i=1}^{10} \sum_{j=1}^{10} E (X_i) E (X_j) \quad (13.195)$$

Using (13.193) , (13.195), and the definition of covariance,

$$V (T) = \frac{1}{100} \left( \sum_{i=1}^{10} \sum_{j=1}^{10} (E [X_i^2 X_j^2] - E (X_i^2) E (X_j^2)) \right) \quad (13.196)$$

$$= \frac{1}{100} \left( \sum_{i=1}^{10} \sum_{j=1}^{10} \text{cov} (X_i^2, X_j^2) \right) \quad (13.197)$$

As all the variables are i.i.d's and are thus uncorrelated,

$$\text{cov} (X_i^2, X_j^2) = \begin{cases} 0 & \text{if } i \neq j \\ V (X_i^2) & \text{if } i = j \end{cases} \quad (13.198)$$



Now,

$$V(T) = \frac{1}{100} \sum_{i=1}^{10} \text{cov}(X_i^2, X_i^2) \quad (13.199)$$

$$= \frac{1}{100} \sum_{i=1}^{10} V(X_i^2) \quad (13.200)$$

Using (13.186) to find  $V(X_i^2)$ ,

$$V(X_i^2) = E(X_i^4) - E(X_i^2)^2 \quad (13.201)$$

Using moment generating function to find  $E(X_i^4)$  and  $E(X_i^2)$ ,

$$M_X(t) = E(e^{tX}) \quad (13.202)$$

$$= e^{\frac{1}{2}\theta^2 t^2} \quad (13.203)$$

Differentiating it with respect to  $t$ ,

$$\frac{dM_X(t)}{dt} = E(Xe^{tX}) \quad (13.204)$$

$$\frac{d^n M_X(t)}{dt^n} = E(X^n e^{tX}) \quad (13.205)$$

$$\frac{dM_X(t)}{dt} = \theta^2 t e^{\frac{1}{2}\theta^2 t^2} \quad (13.206)$$

$$\frac{d^2 M_X(t)}{dt^2} = \theta^2 e^{\frac{1}{2}\theta^2 t^2} + \theta^4 t^2 e^{\frac{1}{2}\theta^2 t^2} \quad (13.207)$$

$$\frac{d^3 M_X(t)}{dt^3} = \theta^4 t e^{\frac{1}{2}\theta^2 t^2} + 2\theta^4 t e^{\frac{1}{2}\theta^2 t^2} + \theta^6 t^3 e^{\frac{1}{2}\theta^2 t^2} \quad (13.208)$$

$$= 3\theta^4 t e^{\frac{1}{2}\theta^2 t^2} + \theta^6 t^3 e^{\frac{1}{2}\theta^2 t^2} \quad (13.209)$$

$$\frac{d^4 M_X(t)}{dt^4} = 3\theta^4 e^{\frac{1}{2}\theta^2 t^2} + 3\theta^6 t^2 e^{\frac{1}{2}\theta^2 t^2} + 3\theta^6 t^2 e^{\frac{1}{2}\theta^2 t^2} + \theta^8 t^4 e^{\frac{1}{2}\theta^2 t^2} \quad (13.210)$$

$$= 3\theta^4 e^{\frac{1}{2}\theta^2 t^2} + 6\theta^6 t^2 e^{\frac{1}{2}\theta^2 t^2} + \theta^8 t^4 e^{\frac{1}{2}\theta^2 t^2} \quad (13.211)$$

Now,

$$E(X^4) = \frac{d^4 M_X(t)}{dt^4} \Big|_{t=0} \quad (13.212)$$

$$= 3\theta^4 \quad (13.213)$$

$$E(X^2) = \frac{d^2 M_X(t)}{dt^2} \Big|_{t=0} \quad (13.214)$$

$$= \theta^2 \quad (13.215)$$

$$V(X_i^2) = 3\theta^4 - (\theta^2)^2 \quad (13.216)$$

$$= 2\theta^4 \quad (13.217)$$

$$V(T) = \frac{1}{100} \sum_{i=1}^{10} 2\theta^4 \quad (13.218)$$

$$= \frac{1}{5} \theta^4 \quad (13.219)$$

The mean square error of  $cT$  as an estimator of  $\theta^2$  is given by

$$f = E\left((cT - \theta^2)^2\right) \quad (13.220)$$

Since

$$E(X^2) = V(X) + (E(X))^2 \quad (13.221)$$

Now,

$$f = V(cT - \theta^2) + \left(E(cT - \theta^2)^2\right) \quad (13.222)$$

$$= V(cT) + \left(E(cT) - \theta^2\right)^2 \quad (13.223)$$

$$= c^2 V(T) + \left(cE(T) - \theta^2\right)^2 \quad (13.224)$$

$$= \frac{c^2}{5}\theta^4 + \theta^4(c-1)^2 \quad (13.225)$$

Minimizing the mean square error,

$$\frac{df}{dc} = 0 \quad (13.226)$$

$$\frac{df}{dc} = \frac{2c}{5}\theta^4 + 2\theta^4(c-1) \quad (13.227)$$

$$\frac{c}{5} + c - 1 = 0 \quad (13.228)$$

$$6c = 5 \quad (13.229)$$

$$c = \frac{5}{6} \quad (13.230)$$

To verify that  $cT$  is a good estimate for  $\theta$ , let

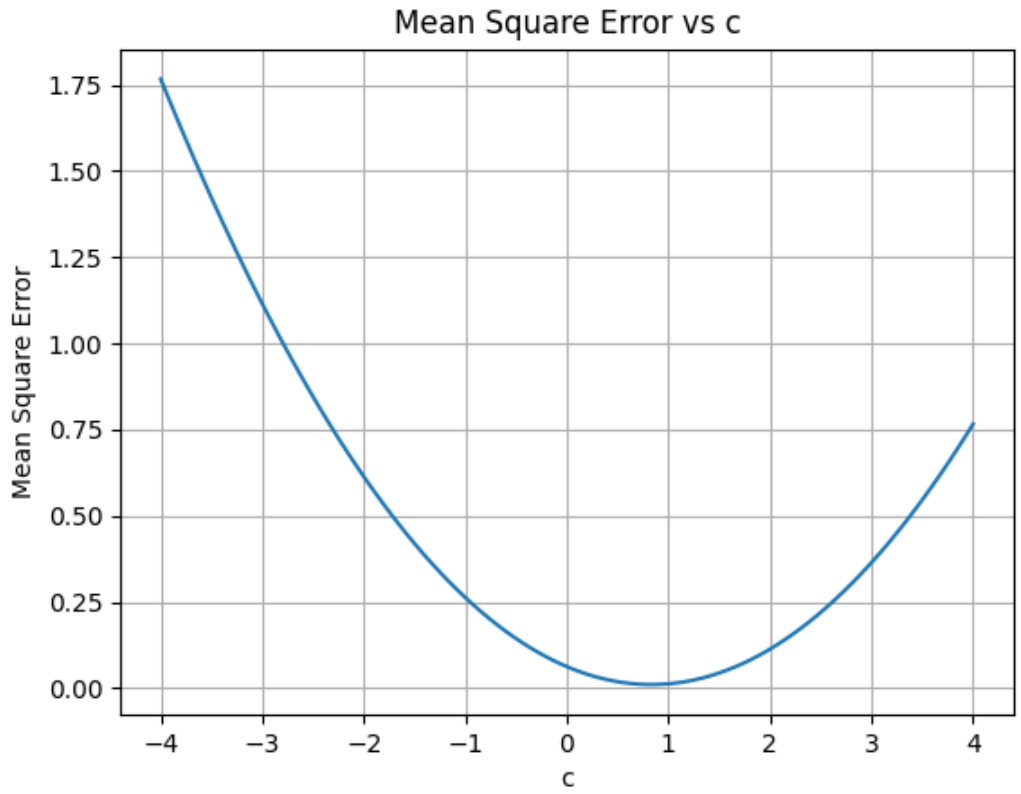
$$\theta = 0.5 \quad (13.231)$$

$$E\left((cT - \theta^2)^2\right) = \frac{1}{5} \left(\frac{25}{36}\right) (0.5)^4 + \left(\frac{-1}{6}\right)^2 (0.5)^4 \quad (13.232)$$

$$= \frac{1}{96} \quad (13.233)$$

$$= 0.0104 \quad (13.234)$$

The mean square error is small relative to the value of  $\theta$ , hence  $cT$  is a good estimate.



Using laplace transform to find CDF,

$$M_T(s) = M_{\frac{1}{10} \sum_{i=1}^{10} X_i^2}(s) \quad (13.235)$$

$$= \prod_{i=1}^{10} M_{\frac{X_i^2}{10}}(s) \quad (13.236)$$

Now,

$$M_{\frac{X^2}{10}}(s) = E\left(e^{\frac{-sX^2}{10}}\right) \quad (13.237)$$

$$= \int_{-\infty}^{\infty} e^{\frac{-sx^2}{10}} \frac{e^{\frac{-x^2}{2\theta^2}}}{\theta\sqrt{2\pi}} dx \quad (13.238)$$

$$= \frac{1}{\theta\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\frac{-x^2}{2}\left(\frac{1}{\theta^2} + \frac{s}{5}\right)} dx \quad (13.239)$$

Using the substitution,

$$u = x\sqrt{\frac{1}{\theta^2} + \frac{s}{5}} \quad (13.240)$$

$$du = \left(\sqrt{\frac{1}{\theta^2} + \frac{s}{5}}\right) dx \quad (13.241)$$

$$M_{\frac{X^2}{10}}(s) = \frac{1}{\theta\sqrt{\frac{1}{\theta^2} + \frac{s}{5}}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{\frac{-u^2}{2}} du \quad (13.242)$$

The integral evaluates to 1 since it is the pdf of a normal distribution,

$$M_{\frac{X^2}{10}}(s) = \frac{1}{\sqrt{1 + \frac{s\theta^2}{5}}} \quad (13.243)$$

$$M_T(s) = \frac{1}{\left(1 + \frac{s\theta^2}{5}\right)^5} \quad (13.244)$$

Taking inverse laplace transform gives us the pdf,

$$p_T(t) = L^{-1} [M_T(s)] \quad (13.245)$$

$$= L^{-1} \left[ \frac{1}{\left(1 + \frac{s\theta^2}{5}\right)^5} \right] \quad (13.246)$$

$$= \frac{3125t^4 e^{\frac{-5t}{\theta^2}}}{24\theta^{10}} u(t) \quad (13.247)$$

ROC of laplace transform:

$$Re(s) > \frac{-5}{\theta^2} \quad (13.248)$$

CDF of T:

$$F_T(t) = \int_{-\infty}^t p_T(t) dt \quad (13.249)$$

$$= \int_{-\infty}^t \frac{3125t^4 e^{\frac{-5t}{\theta^2}}}{24\theta^{10}} u(t) dt \quad (13.250)$$

$$= \int_0^t \frac{3125t^4 e^{\frac{-5t}{\theta^2}}}{24\theta^{10}} dt \quad (13.251)$$

Applying integration by parts to solve the integral,

$$F_T(t) = \frac{3125}{24\theta^{10}} \left( \frac{-\theta^2}{5} t^4 e^{\frac{-5t}{\theta^2}} + \int_0^t \frac{4\theta^2}{5} t^3 e^{\frac{-5t}{\theta^2}} dt \right) \quad (13.252)$$

$$= \frac{3125}{24\theta^{10}} \left( -e^{\frac{-5t}{\theta^2}} \left( \frac{\theta^2}{5} t^4 + \frac{4\theta^4}{25} t^3 \right) + \int_0^t \frac{12\theta^4}{25} t^2 e^{\frac{-5t}{\theta^2}} dt \right) \quad (13.253)$$

$$= \frac{3125}{24\theta^{10}} \left( -e^{\frac{-5t}{\theta^2}} \left( \frac{\theta^2}{5} t^4 + \frac{4\theta^4}{25} t^3 + \frac{12\theta^6}{125} t^2 \right) + \int_0^t \frac{24\theta^6}{125} t e^{\frac{-5t}{\theta^2}} dt \right) \quad (13.254)$$

$$= \frac{3125}{24\theta^{10}} \left( -e^{\frac{-5t}{\theta^2}} \left( \frac{\theta^2}{5} t^4 + \frac{4\theta^4}{25} t^3 + \frac{12\theta^6}{125} t^2 + \frac{24\theta^8}{625} t \right) + \int_0^t \frac{24\theta^8}{625} e^{\frac{-5t}{\theta^2}} dt \right) \quad (13.255)$$

$$= \frac{3125}{24\theta^{10}} \left( -e^{\frac{-5t}{\theta^2}} \left( \frac{\theta^2}{5} t^4 + \frac{4\theta^4}{25} t^3 + \frac{12\theta^6}{125} t^2 + \frac{24\theta^8}{625} t + \frac{24\theta^{10}}{3125} \right) + \frac{24\theta^{10}}{3125} \right) \quad (13.256)$$

$$= 1 - \frac{e^{\frac{-5t}{\theta^2}}}{24\theta^8} (625t^4 + 500\theta^2 t^3 + 300\theta^4 t^2 + 120\theta^6 t + 24\theta^8) \quad (13.257)$$

Let  $\theta = 0.5$  for simulation,

$$F_T(t) = 1 - \frac{e^{-20t}}{3} (20000t^4 + 4000t^3 + 600t^2 + 60t + 3) \quad (13.258)$$

Simulation procedure:

(i)

$$u_1 = (\text{double}) \frac{\text{rand}()}{RAND\_MAX} \quad (13.259)$$

$$u_2 = (\text{double}) \frac{\text{rand}()}{RAND\_MAX} \quad (13.260)$$

Generates a uniform distribution between 0 and 1.

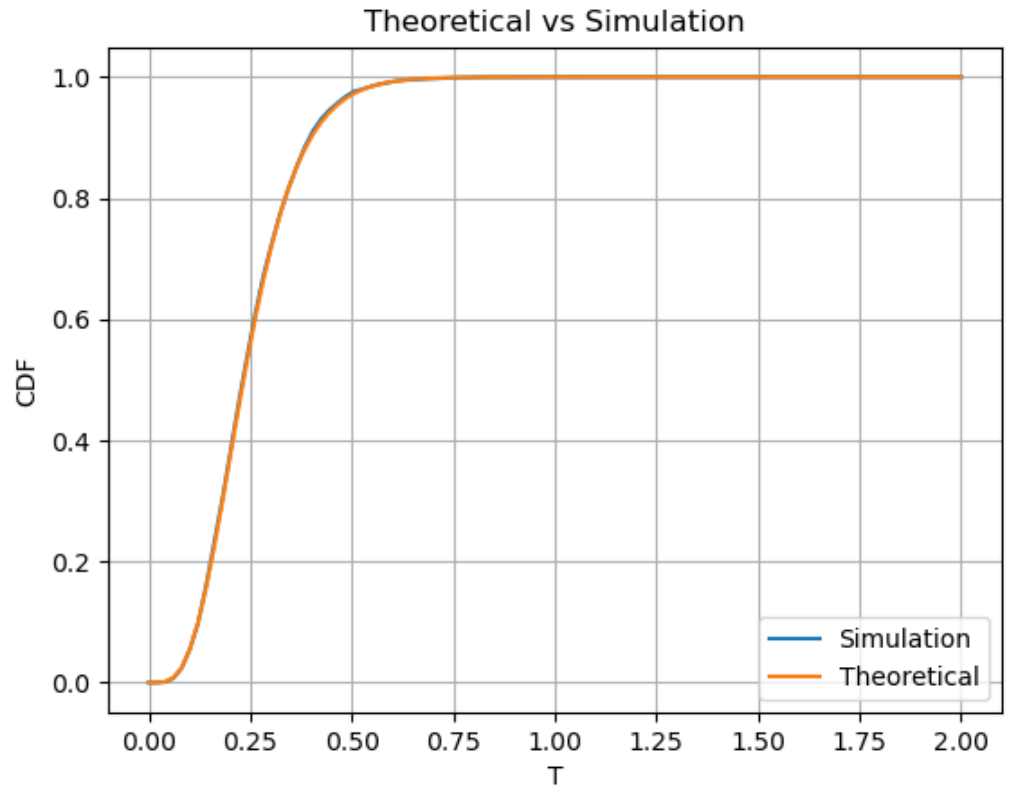


Figure 13.4: Theoretical CDF vs Simulation CDF

(ii)

$$X_i = \sqrt{\theta^2} \left( \sqrt{-2 \log u_1} \cos 2\pi u_2 \right) + \mu \quad (13.261)$$

Transforms the uniform distribution into gaussian distribution of desired mean and variance. Ten such random variables are generated.



(iii)

$$T = \frac{1}{10} \sum_{i=1}^{10} X_i^2 \quad (13.262)$$

The values of the random variables are squared and then averaged together to generate  $T$ .

- (iv) The value of  $c$  which minimizes the mean square error is found by calculating  $E\left((cT - \theta^2)^2\right)$  for a range of values of  $c$  and choosing that value of  $c$  which gives the minimum value for the expression.
- (v) The cdf is simulated by counting the number of samples less than a certain  $t$  and dividing it by the total number of samples. This gives the CDF of  $T$  at  $t$ .

## Chapter 14

# Multivariate Gaussian

14.1 Suppose that  $(\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3)$  has  $N_3(\mu, \Sigma)$  distribution with  $\mu = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$   $\Sigma = \begin{bmatrix} 2 & 2 & 1 \\ 2 & 5 & 1 \\ 1 & 1 & 1 \end{bmatrix}$

Given that  $\phi(-0.5)=0.3085$ , where  $\phi(\cdot)$  denotes the cumulative distribution function of a standard normal random variable,  $P\left((X_1 - 2X_2 + 2X_3)^2 < \frac{7}{2}\right)$  equals to  
(GATE ST 2023)

**Solution:**

$$\text{Let } \mathbf{Y} = \mathbf{X}_1 - 2\mathbf{X}_2 + 2\mathbf{X}_3 \text{ and } \mathbf{a} = \begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix}$$

$$\mu_Y = \mathbf{a}^T \boldsymbol{\mu} \quad (14.1)$$

$$= \begin{bmatrix} 1 & -2 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad (14.2)$$

$$= 0 \quad (14.3)$$

$$\sigma_Y^2 = \mathbf{a}^T \boldsymbol{\Sigma} \mathbf{a} \quad (14.4)$$

$$= \begin{bmatrix} 1 & -2 & 2 \end{bmatrix} \begin{bmatrix} 2 & 2 & 1 \\ 2 & 5 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix} \quad (14.5)$$

$$= 1 \quad (14.6)$$

$$\Pr \left( Y^2 < \frac{7}{2} \right) = \Pr \left( -\sqrt{\frac{7}{2}} < Y < \sqrt{\frac{7}{2}} \right) \quad (14.7)$$

$$\Phi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz \quad (14.8)$$

$$\Phi \left( \sqrt{\frac{7}{2}} \right) = \int_{-\infty}^{\sqrt{\frac{7}{2}}} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz \quad (14.9)$$

$$= \int_{-\infty}^{-\frac{1}{2}} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} + \int_{-\frac{1}{2}}^{\sqrt{\frac{7}{2}}} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz \quad (14.10)$$

$$= \Phi(-0.5) + \int_{-\frac{1}{2}}^{\sqrt{\frac{7}{2}}} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz \quad (14.11)$$

$$= 0.3085 + 0.66082 \quad (14.12)$$

$$= 0.96932 \quad (14.13)$$

$$\Phi\left(-\sqrt{\frac{7}{2}}\right) = \int_{-\infty}^{-\sqrt{\frac{7}{2}}} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz \quad (14.14)$$

$$= 0.03068 \quad (14.15)$$

$$\Pr\left(\left(X_1 - 2X_2 + 2X_3\right)^2 < \frac{7}{2}\right) = \Phi\left(\sqrt{\frac{7}{2}}\right) - \Phi\left(-\sqrt{\frac{7}{2}}\right) \quad (14.16)$$

$$= 0.96932 - 0.03068 \quad (14.17)$$

$$= 0.93864 \quad (14.18)$$

## 14.1. Simulation Steps

- (a) Generate Random Samples: Generate random samples for  $X_1$ ,  $X_2$ , and  $X_3$  from the given multivariate normal distribution with mean vector  $\mu = [0, 0, 0]$  and covariance matrix  $\Sigma$ .
- (b) Calculate  $Y$ : Calculate  $Y = X_1 - 2X_2 + 2X_3$ .
- (c) Count Samples: Count the number of samples for which  $Y^2 < \frac{7}{2}$ .
- (d) Calculate Estimated Probability: Calculate the estimated probability by dividing the count by the total number of samples (1000000).

14.2 Let  $\{W_t\}_{t \geq 0}$  be a standard Brownian motion. Then  $E(W_4^2 | W_2 = 2)$  in integer equals (GATE ST 2023)

**Solution:**

Parameter	Description
$\mu_x$	Mean of x
$Var(x)$	Variance of x
$Cov(x, y)$	Covariance between x and y
$\sigma_x$	Standard deviation of x
$\rho$	Co-Relation coefficient
$E(x)$	Expectation of x

In standard brownian motion,

$$W_i \sim N(0, i) \quad (14.19)$$

$$Cov(W_i, W_j) = \min(i, j) \quad (14.20)$$

Now, we know that,

$$E(Y^2|X) = Var(Y|X) + (E(Y|X))^2 \quad (14.21)$$

$X$  and  $Y$  can be represented as:

$$X = \sigma_X Z_1 + \mu_X \quad (14.22)$$

$$Y = \sigma_Y \left( \rho Z_1 + \sqrt{1 - \rho^2} Z_2 \right) + \mu_Y \quad (14.23)$$

where  $Z_1$  and  $Z_2$  are normal distributions.

$$Z_1, Z_2 \sim N(0, 1) \quad (14.24)$$

Writing the above equations in matrix form,

$$\begin{pmatrix} X \\ Y \end{pmatrix} = \begin{bmatrix} \sigma_X & 0 \\ \sigma_Y \rho & \sigma_Y \sqrt{1 - \rho^2} \end{bmatrix} \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} + \begin{pmatrix} \mu_X \\ \mu_Y \end{pmatrix} \quad (14.25)$$

This can be represented as,

$$\mathbf{x} = A\mathbf{z} + \boldsymbol{\mu} \quad (14.26)$$

Taking expectation both sides,

$$E(\mathbf{x}) = E(A\mathbf{z} + \boldsymbol{\mu}) \quad (14.27)$$

$$= AE(\mathbf{z}) + E(\boldsymbol{\mu}) \quad (14.28)$$

$$= \boldsymbol{\mu} \quad (14.29)$$

We know that covariance matrix for  $X$  and  $Y$  is given by:

$$\sigma_{\mathbf{z}} = E\left((\mathbf{x} - \boldsymbol{\mu})(\mathbf{x} - \boldsymbol{\mu})^T\right) \quad (14.30)$$

$$= E\left((A\mathbf{z})(A\mathbf{z})^T\right) \quad (14.31)$$

$$= E(A\mathbf{z}\mathbf{z}^T A^T) \quad (14.32)$$

$$= AE(\mathbf{z}\mathbf{z}^T)A^T \quad (14.33)$$

Multiplying  $\mathbf{z}$  and  $\mathbf{z}^T$  we get,

$$\mathbf{z}\mathbf{z}^T = \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} \begin{pmatrix} Z_1 & Z_2 \end{pmatrix} \quad (14.34)$$

$$= \begin{pmatrix} Z_1^2 & Z_1 Z_2 \\ Z_1 Z_2 & Z_2^2 \end{pmatrix} \quad (14.35)$$

We know that,

$$Var(Z_1) = E\left((Z_1 - \mu)^2\right) \quad (14.36)$$

$$E(Z_1^2) = 1 \quad (14.37)$$

Same goes for  $Z_2$  as  $Z_1$  and  $Z_2$  are both normal distributions.

Taking expectation both sides in equation (14.35),

$$E(\mathbf{z}\mathbf{z}^T) = \begin{pmatrix} E(Z_1^2) & E(Z_1 Z_2) \\ E(Z_1 Z_2) & E(Z_2^2) \end{pmatrix} \quad (14.38)$$

$$= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad (14.39)$$

Hence,

$$\sigma_{\mathbf{z}} = AA^T \quad (14.40)$$

$$= \begin{bmatrix} \sigma_X & 0 \\ \sigma_Y \rho & \sigma_Y \sqrt{1 - \rho^2} \end{bmatrix} \begin{bmatrix} \sigma_X & \sigma_Y \rho \\ 0 & \sigma_Y \sqrt{1 - \rho^2} \end{bmatrix} \quad (14.41)$$

$$= \begin{bmatrix} (\sigma_X)^2 & \sigma_X \sigma_Y \rho \\ \sigma_X \sigma_Y \rho & (\sigma_Y)^2 \end{bmatrix} \quad (14.42)$$

The Co-Relation Coefficient is given by:

$$\rho = \frac{Cov(X, Y)}{\sqrt{Var(X) Var(Y)}} \quad (14.43)$$

Substituting value of  $Z_1$  in  $Y$ ,

$$Y = \sigma_Y \rho \left( \frac{X - \mu_X}{\sigma_X} \right) + \sigma_Y \sqrt{1 - \rho^2} Z_2 + \mu_Y \quad (14.44)$$

This an equation of  $Y$  in terms of  $X$ . All the terms except  $Z_2$  are constants. Taking expectation on both sides,

$$E(Y|X = x) = E \left( \sigma_Y \rho \left( \frac{x - \mu_X}{\sigma_X} \right) + \sigma_Y \sqrt{1 - \rho^2} Z_2 + \mu_Y \right) \quad (14.45)$$

$$= E \left( \sigma_Y \rho \left( \frac{x - \mu_X}{\sigma_X} \right) + \mu_Y \right) + E \left( \sigma_Y \sqrt{1 - \rho^2} Z_2 \right) \quad (14.46)$$

$$= \mu_Y + \rho \left( \frac{\sigma_Y}{\sigma_X} \right) (x - \mu_X) + \sigma_Y \sqrt{1 - \rho^2} E(Z_2) \quad (14.47)$$



Now for variance,

$$Var(Y|X=x) = Var\left(\sigma_Y \rho \left(\frac{x - \mu_X}{\sigma_X}\right) + \sigma_Y \sqrt{1 - \rho^2} Z_2 + \mu_Y\right) \quad (14.48)$$

$$= Var\left(\sigma_Y \rho \left(\frac{x - \mu_X}{\sigma_X}\right) + \mu_Y\right) + Var\left(\sigma_Y \sqrt{1 - \rho^2} Z_2\right) \quad (14.49)$$

Variance of constants terms is 0.

$$Var(Y|X=x) = Var\left(\sigma_Y \sqrt{1 - \rho^2} Z_2\right) \quad (14.50)$$

$$= (1 - \rho^2) \sigma_Y^2 Var(Z_2) \quad (14.51)$$

$Z_2$  is a normal distribution so,

$$E(Z_2) = 0 \quad (14.52)$$

$$Var(Z_2) = 1 \quad (14.53)$$

By substituting these values in above equations,

$$E(Y|X=x) = \mu_Y + \rho \left(\frac{\sigma_Y}{\sigma_X}\right) (x - \mu_X) \quad (14.54)$$

$$Var(Y|X=x) = (1 - \rho^2) \sigma_Y^2 \quad (14.55)$$

In our case,

$$Y = W_4 \quad (14.56)$$

$$X = W_2 \quad (14.57)$$

$$x = 2 \quad (14.58)$$

Hence, we get that,

$$\mu_X = \mu_Y = 0 \quad (14.59)$$

$$\sigma_X = \sqrt{2} \quad (14.60)$$

$$\sigma_Y = 2 \quad (14.61)$$

$$\rho = \frac{2}{\sqrt{8}} \quad (14.62)$$

$$= \frac{1}{\sqrt{2}} \quad (14.63)$$

Substituting the values in above equations,

$$E(Y|X=2) = \frac{1}{\sqrt{2}} \cdot \frac{2}{\sqrt{2}} \cdot 2 \quad (14.64)$$

$$= 2 \quad (14.65)$$

$$Var(Y|X=2) = \left(1 - \frac{1}{2}\right) (2)^2 \quad (14.66)$$

$$= \frac{1}{2} \cdot 4 \quad (14.67)$$

$$= 2 \quad (14.68)$$

Substituting these values in (14.21),

$$E(Y^2|X=2) = 2 + (2)^2 \quad (14.69)$$

$$= 6 \quad (14.70)$$

$$E(W_4^2|W_2=2) = 6 \quad (14.71)$$

Steps for Simulation:

- (a) Set the number of samples to be generated as 10000
- (b) Write a function to generate a uniform distribution
- (c) Write a function to generate a normal distribution by using the uniform distribution function defined above through box muller method
- (d) Write a function to generate value of  $\mathbf{z}$  from one the column of the dist matrix
- (e) Write a function to calculate  $\mathbf{x}$  by matrix multiplication:  $\mathbf{x} = A\mathbf{z} + \mathbf{u}$ .
- (f) Write a function to store values of each vector  $\mathbf{x}$  obtained in a 2D array of size  $2*\text{numsamples}$
- (g) Make a temporary 2D array(dist) of size  $2 \times 10000$  to store the values of  $Z_1$  and  $Z_2$  obtained after calling the normal distribution function
- (h) Use a for loop to store the values obtained by calling the normal function multiple times
- (i)
- (j) Assume values for constants,

$$\sigma_X = 0.5 \quad (14.72)$$

$$\sigma_Y = 0.8 \quad (14.73)$$

$$\rho = 0.5 \quad (14.74)$$

$$\mu_X = 1 \quad (14.75)$$

$$\mu_Y = 1.5 \quad (14.76)$$

- (k) Create  $A$  matrix and vector  $\boldsymbol{\mu}$  and fill the values assumed above in these, the dimension for matrix is  $2 \times 2$  and the vector  $2 \times 1$

- (l) Make a 2D array(`ans`) to store the final values of  $\mathbf{x}$  obtained after matrix multiplication
- (m) For each iteration, create a  $\mathbf{z}$  assign values from one of the columns from `dist` to  $\mathbf{z}$  using the function defined above
- (n) For each iteration, Generate  $\mathbf{x}$  vector and call the function to solve to matrix multiplication and give value for  $\mathbf{x}$
- (o) Store the values obtained from matrix multiplication in the `ans` array using the store function defined above
- (p) The `ans` array has the values of  $\mathbf{x}$  for 10000 simulations.

