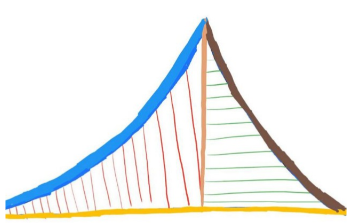
GATE PROBABILITY

Through Simulations

G. V. V. Sharma



Contents

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 occurances of two independent events A and B are 0.5 and 0.8 respectively. What is the probability of occurance of at least A or B(rounded off to 1 decimal place)? (GATE CE 2023)

Given,

$\Pr(A) = 0.5$ and $\Pr(B) = 0.8$

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

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

$$\Pr(A + B) = \Pr(A) + \Pr(B) - \Pr(AB) \tag{1.1}$$

$$= \Pr(A) + \Pr(B) - \Pr(A) \times \Pr(B) \tag{1.2}$$

$$= 0.5 + 0.8 - 0.5 \times 0.8 \tag{1.3}$$

$$= 0.9 \tag{1.4}$$

Chapter 2

Distributions

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

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

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

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

From question and (2.2);

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

From (2.3);

(a)

$$\Pr\left(X \leq -1\right) = F_X(-1) = 1 - Q\left(0\right) \tag{2.4}$$

$$= 0.5 \tag{2.5}$$

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

$$\Pr\left(X = b\right) = F_X(b) - \lim_{x \to b^-} F_X(x) \tag{2.6}$$

From (2.6);

$$\Pr\left(X = -1\right) = F_X(-1) - \lim_{x \to -1^-} F_X(x) \tag{2.7}$$

$$= 1 - Q\left(0\right) - Q\left(-(-1)\right) \tag{2.8}$$

$$= 0.341 \tag{2.9}$$

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

(c)

$$\Pr\left(X < -1\right) = \lim_{x \to -1^-} F_X(x) = F_X(-1) \tag{2.10}$$

$$= Q\left(-(-1)\right) \tag{2.11}$$

$$= 0.159 \tag{2.12}$$

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

(d)

$$\Pr\left(X \leq 0\right) = F_X(0) = 1 - Q\left(1\right) \tag{2.13}$$

$$= 0.8413 \tag{2.14}$$

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

Guassian CDF plot of X is given in fig2.1

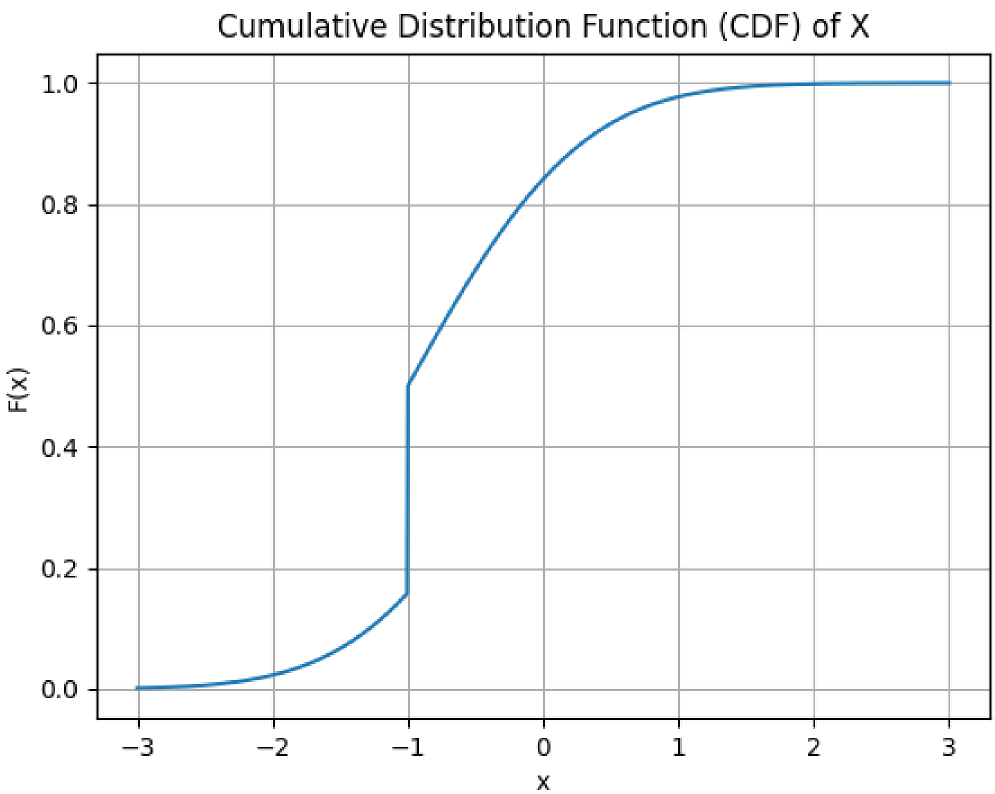


Figure 2.1:

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

$$f(x) = \begin{cases} \frac{1}{2\sqrt{3}}, & -\sqrt{3} \le x \le \sqrt{3} \\ 0, & \text{otherwise} \end{cases} \tag{2.15}$$

Then the variance of X is? (GATE XH-C1 2023)

Solution:

The mean of X

$$\mu_X = \int_{-\infty}^{\infty} x f(x) dx \tag{2.16}$$

As the integrand is odd

$$\implies \mu_X = 0 \tag{2.17}$$

The variance of X is:

$$\sigma_X^2 = \mathbb{E}\left(X - \mu_X\right)^2 \tag{2.18}$$

From (2.17)

$$\implies \sigma_X^2 = \mathbb{E}\left(X^2\right) \tag{2.19}$$

$$= \frac{1}{2\sqrt{3}} \int_{-\sqrt{3}}^{\sqrt{3}} x^2 dx \tag{2.20}$$

$$= 1 \tag{2.21}$$

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 $\Pr\left(X = 3\right)$ (rounded off to two decimal places) equals

(GATE ST 2023)

Solution:

RV	Values	Description
A	0	1^{st} Bulb defective
	1	1^{st} Bulb non-defective
B	0	2^{nd} Bulb defective
	1	2^{nd} Bulb non-defective
C	0	3^{rd} Bulb defective
	1	3^{rd} Bulb non-defective

Table 2.1: Random variable declaration.

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

$$p_X(2) = p_{AB}(0,0) \tag{2.22}$$
$$= \frac{2}{5} \times \frac{1}{4} \tag{2.23}$$
$$= \frac{1}{10} \tag{2.24}$$
$$p_X(3) = p_{ABC}(1,0,0) + p_{ABC}(0,1,0) + p_{ABC}(1,1,1) \tag{2.25}$$
$$= \frac{3}{5} \times \frac{2}{4} \times \frac{1}{3} + \frac{2}{5} \times \frac{3}{4} \times \frac{1}{3} + \frac{3}{5} \times \frac{2}{4} \times \frac{1}{3} \tag{2.26}$$
$$= \frac{3}{10} \tag{2.27}$$
$$p_X(4) = p_{ABC}(0,1,1) + p_{ABC}(1,0,1) + p_{ABC}(1,1,0) \tag{2.28}$$
$$= \frac{2}{5} \times \frac{3}{4} \times \frac{2}{3} + \frac{3}{5} \times \frac{2}{4} \times \frac{2}{3} + \frac{3}{5} \times \frac{2}{4} \times \frac{2}{3} \tag{2.29}$$
$$= \frac{6}{10} \tag{2.30}$$

Hence, The pmf of X is

$$p_X(k) = \begin{cases} 0 & k = 1 \\ \frac{1}{10} & k = 2 \\ \frac{3}{10} & k = 3 \\ \frac{6}{10} & k = 4 \\ 1 & k = 5 \end{cases} \tag{2.31}$$

2.4 Let X be a random variable with cumulative distribution function

$$F_X(x) = \begin{cases} 0 & \text{if } x < -1 \\ \frac{1}{4}(x+1) & \text{if } -1 \le x < 0 \\ \frac{1}{4}(x+3) & \text{if } 0 \le x < 1 \\ 1 & \text{if } x \ge 1 \end{cases} \tag{2.32}$$

Which one of the following statements is true?

(A)

$$\lim_{n\to\infty} \Pr\left(-\frac{1}{2} + \frac{1}{n} < X < -\frac{1}{n}\right) = \frac{5}{8} \tag{2.33}$$

(B)

$$\lim_{n\to\infty} \Pr\left(-\frac{1}{2} - \frac{1}{n} < X < \frac{1}{n}\right) = \frac{5}{8} \tag{2.34}$$

(C)

$$\lim_{n\to\infty} \Pr\left(X = \frac{1}{n}\right) = \frac{1}{2} \tag{2.35}$$

(D)

$$\Pr\left(X = 0\right) = \frac{1}{3} \tag{2.36}$$

(GATE ST 2023)

Solution:

$$f_X(x) = \begin{cases} 0 & \text{if } x < -1 \\ \frac{1}{4} & \text{if } -1 \leq x < 0 \\ \frac{1}{4} + \frac{1}{2}\delta(x) & \text{if } 0 \leq x < 1 \\ 0 & \text{if } x \geq 1 \end{cases} \tag{2.37}$$

(A)

$$\lim_{n\to\infty} \Pr\left(-\frac{1}{2} + \frac{1}{n} < X < -\frac{1}{n}\right) \\ = \lim_{n\to\infty} F_X\left(-\frac{1}{n}\right) - \lim_{n\to\infty} F_X\left(-\frac{1}{2} + \frac{1}{n}\right) \tag{2.38}$$

$$= \lim_{n\to\infty} F_X\left(-\frac{1}{n}\right) - \lim_{n\to\infty} F_X\left(-\frac{1}{2} + \frac{1}{n}\right) \tag{2.39}$$

$$= \lim_{n\to\infty} \frac{1}{4}\left(-\frac{1}{n} + 1\right) - \lim_{n\to\infty} \frac{1}{4}\left(-\frac{1}{2} + \frac{1}{n} + 1\right) \tag{2.40}$$

$$= \frac{1}{8} \tag{2.41}$$

$\therefore (A)$ is not true.

(B)

$$\lim_{n\to\infty} \Pr\left(-\frac{1}{2} - \frac{1}{n} < X < \frac{1}{n}\right)$$
$$= \lim_{n\to\infty} F_X\left(\frac{1}{n}\right) - \lim_{n\to\infty} F_X\left(-\frac{1}{2} - \frac{1}{n}\right) \quad (2.42)$$

$$= \lim_{n\to\infty} F_X\left(\frac{1}{n}\right) - \lim_{n\to\infty} F_X\left(-\frac{1}{2} - \frac{1}{n}\right) \quad (2.43)$$
$$= \lim_{n\to\infty} \frac{1}{4}\left(\frac{1}{n} + 3\right) - \lim_{n\to\infty} \frac{1}{4}\left(-\frac{1}{2} - \frac{1}{n} + 1\right) \quad (2.44)$$
$$= \frac{5}{8} \quad (2.45)$$

$\therefore (B)$ is true.

(C) From (2.37)

$$\lim_{n\to\infty} \Pr\left(X = \frac{1}{n}\right) = 0 \quad (2.46)$$

$\therefore (C)$ is not true.

(D) From (2.37)

$$\Pr\left(X = 0\right) = \frac{1}{2} \quad (2.47)$$

$\therefore (D)$ 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 = num_sim$\times(1 - \Pr(X = 0))$.

(g) Generate sample conatining only zeros of the size = 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.

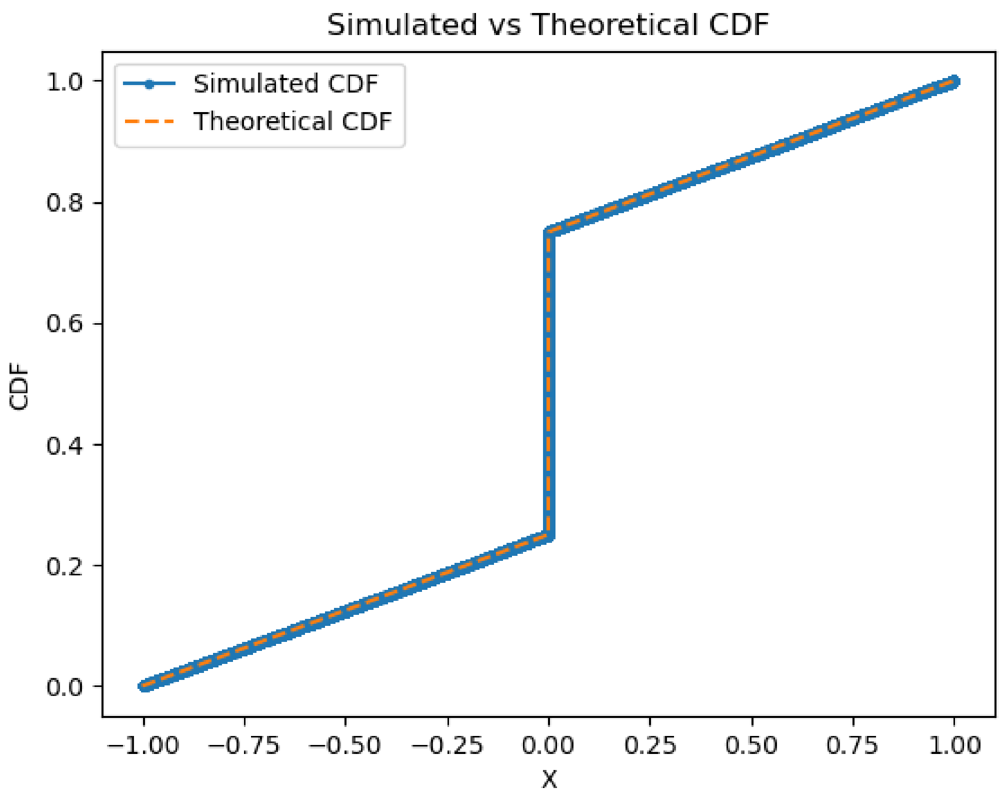


Figure 2.2: CDF of X-(simulation vs actual)

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.

$$\Pr\left(Y = 3 \middle| Y \geq 2\right) = \frac{\Pr\left(Y \geq 2, Y = 3\right)}{\Pr\left(Y \geq 2\right)} \tag{2.48}$$

$$= \frac{\Pr\left(Y = 3\right)}{\Pr\left(Y \geq 2\right)} \tag{2.49}$$

$$= \frac{p_Y(3)}{1 - F_Y(1)} \tag{2.50}$$

$$p_Y(k) = {}^nC_k p^k q^{n-k} \tag{2.51}$$

$$= {}^3C_k \left(\frac{1}{2}\right)^k \left(\frac{1}{2}\right)^{3-k} \tag{2.52}$$

$$\implies p_Y(k) = \begin{cases} \frac{{}^3C_k}{8}; k = \{0, 1, 2, 3\} \\ 0; otherwise \end{cases} \tag{2.53}$$

$$F_Y\left(k\right) = \Pr\left(Y \leq k\right) \tag{2.54}$$

$$= \sum_{k=0}^{k} p_Y(k) \tag{2.55}$$

$$\implies F_Y(k) = \begin{cases} 0; k < 0 \\ \sum_{k=0}^{k} \frac{{}^3C_k}{8}; k = \{0, 1, 2, 3\} \\ 1; k > 3 \end{cases} \tag{2.56}$$

$$\implies \Pr\left(Y = 3 \middle| Y \geq 2\right) = \frac{\left(\frac{1}{8}\right)}{\left(\frac{1}{2}\right)} \tag{2.57}$$

$$= \frac{1}{4} \tag{2.58}$$

$\therefore$ The probability of having all three outcomes as tails is 0.25.

2.6 Let X be a random variable with probability density function

$$p_X(x) = \begin{cases} e^{-x} & if x \geq 0 \\ 0 & otherwise \end{cases} \tag{2.59}$$

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 $\Pr(Y = k) = e^{-k}\left(1 - e^{-1}\right)$ for k= 0, 1, 2, $\ldots$, where $[X]$denotes the largest integer not exceeding x

(GATE ST 2023)

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

$$p_Y(y) = \begin{cases} \frac{1}{b-a} & a < y < b \\ 0 & \text{otherwise} \end{cases} \tag{2.60}$$

and for $a < y < b$

$$F_Y(y) = \Pr(Y \leq y) \tag{2.61}$$

$$= \int_a^y \frac{1}{b-a} dy \tag{2.62}$$

$$= \frac{y-a}{b-a} \tag{2.63}$$

Similarly,for $x \geq 0$

$$F_X(x) = \Pr(X \leq x) \tag{2.64}$$

$$= \int_0^x e^{-x} dx \tag{2.65}$$

$$= 1 - e^{-x} \tag{2.66}$$

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

for $a < y < b$

$$F_Y(y) = \Pr\left(e^{-X} \leq y\right) \tag{2.67}$$

$$= \Pr(X \geq -\ln y) \tag{2.68}$$

$$= 1 - F_X(-\ln y) \tag{2.69}$$

$$= 1 - (1 - y) \tag{2.70}$$

$$= y \tag{2.71}$$

Comparing this with CDF of Uniform distribution, we obtain

$$a = 0, b = 1 \tag{2.72}$$

$$\therefore Y \sim U(0, 1) \tag{2.73}$$

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

for $a < y < b$

$$F_Y(y) = \Pr\left(1 - e^{-X} \leq y\right) \tag{2.74}$$
$$= \Pr\left(e^{-X} \geq 1 - y\right) \tag{2.75}$$
$$= \Pr\left(X \leq -\ln(1-y)\right) \tag{2.76}$$
$$= F_X\left(-\ln(1-y)\right) \tag{2.77}$$
$$= 1 - (1-y) \tag{2.78}$$
$$= y \tag{2.79}$$
$$\implies Y \sim U(0,1) \tag{2.80}$$

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

for $a < y < b$

$$F_Y(y) = \Pr\left(2e^{-X} - 1 \leq y\right) \tag{2.81}$$
$$= \Pr\left(X \geq -\ln\left(\frac{y+1}{2}\right)\right) \tag{2.82}$$
$$= 1 - F_X\left(-\ln\left(\frac{y+1}{2}\right)\right) \qquad = 1 - \left(1 - \frac{y+1}{2}\right) \tag{2.83}$$
$$= \frac{y+1}{2} \tag{2.84}$$

Comparing this with CDF of Uniform distribution, we obtain

$$a = -1, b = 1 \tag{2.85}$$
$$\therefore Y \sim U(-1,1) \tag{2.86}$$

(D) $Y = [X]$

$$\Pr(Y = k) = \Pr([X] = k) \tag{2.87}$$
$$= \Pr(k \leq X < k+1) \tag{2.88}$$
$$= \int_k^{k+1} e^{-x} dx \tag{2.89}$$
$$= e^{-k}\left(1 - e^{-1}\right) \text{ for k=0,1,2..} \tag{2.90}$$

(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 ∞.

(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.

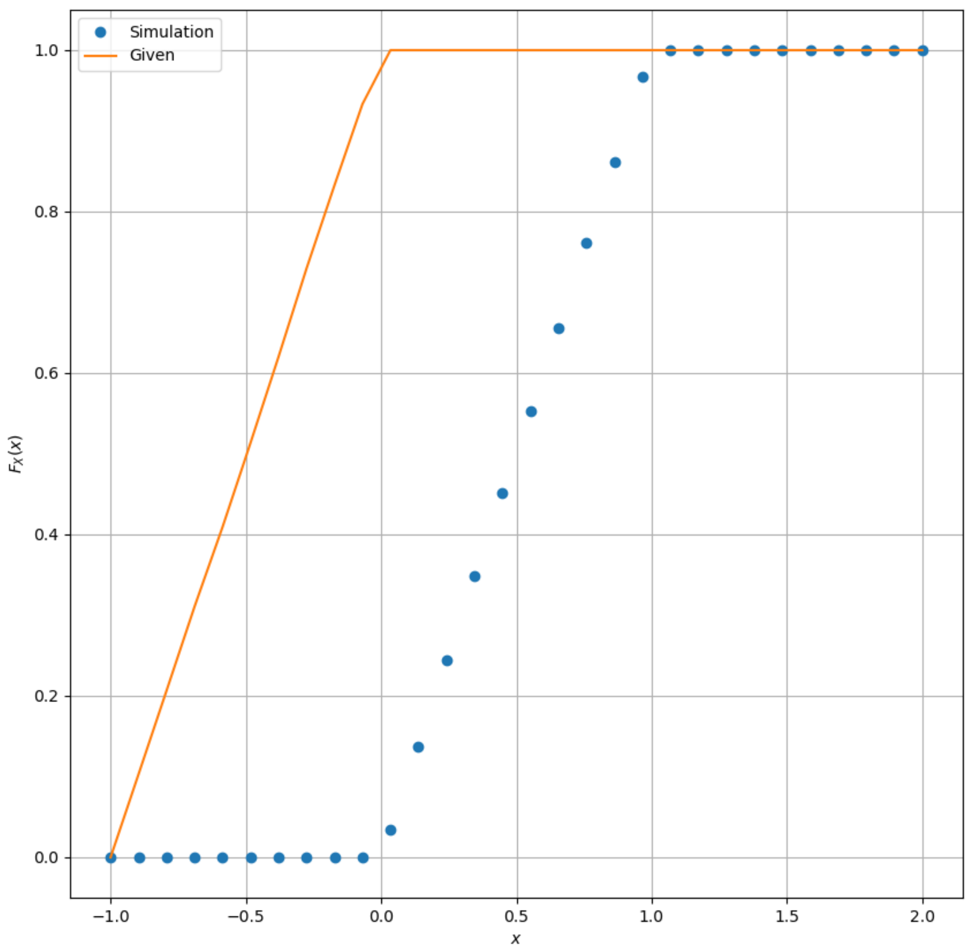


Figure 2.3: e^{-X} vs. $U(-1,0)$
Graphs don't match, $\therefore$ wrong option

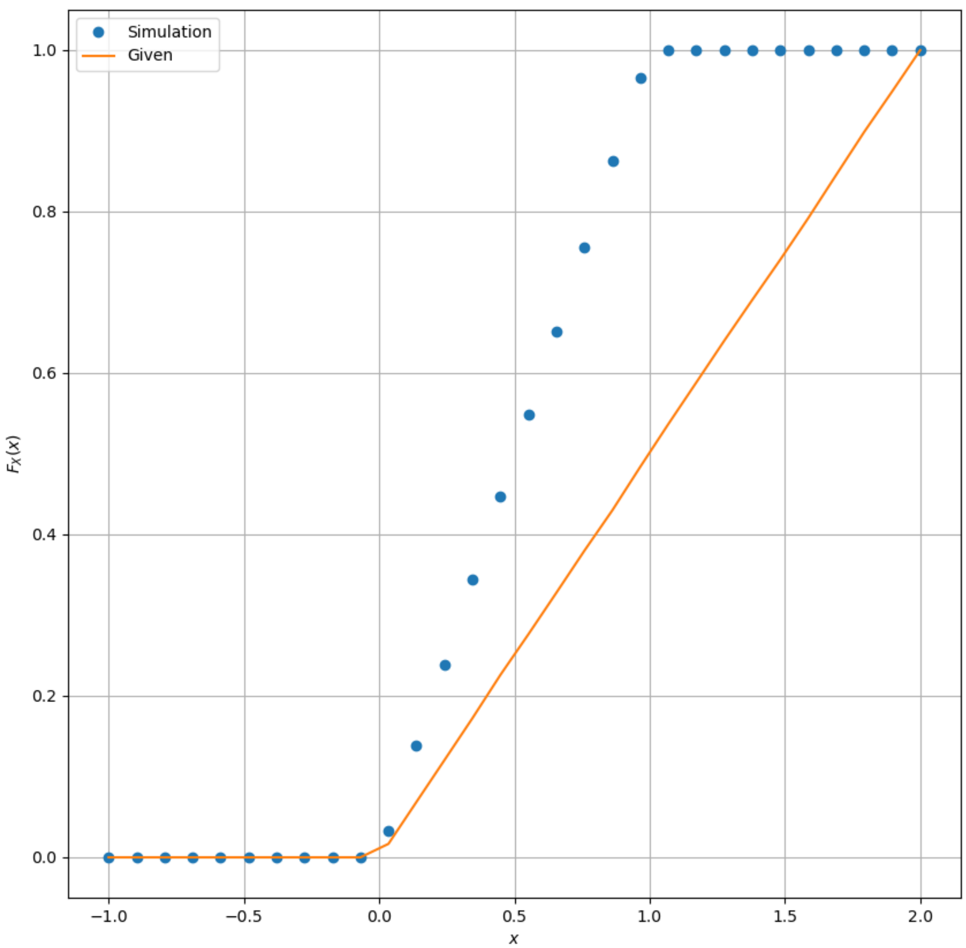


Figure 2.4: $1 - e^{-X}$ vs. $U(0,2)$
Graphs don't match, $\therefore$ wrong option

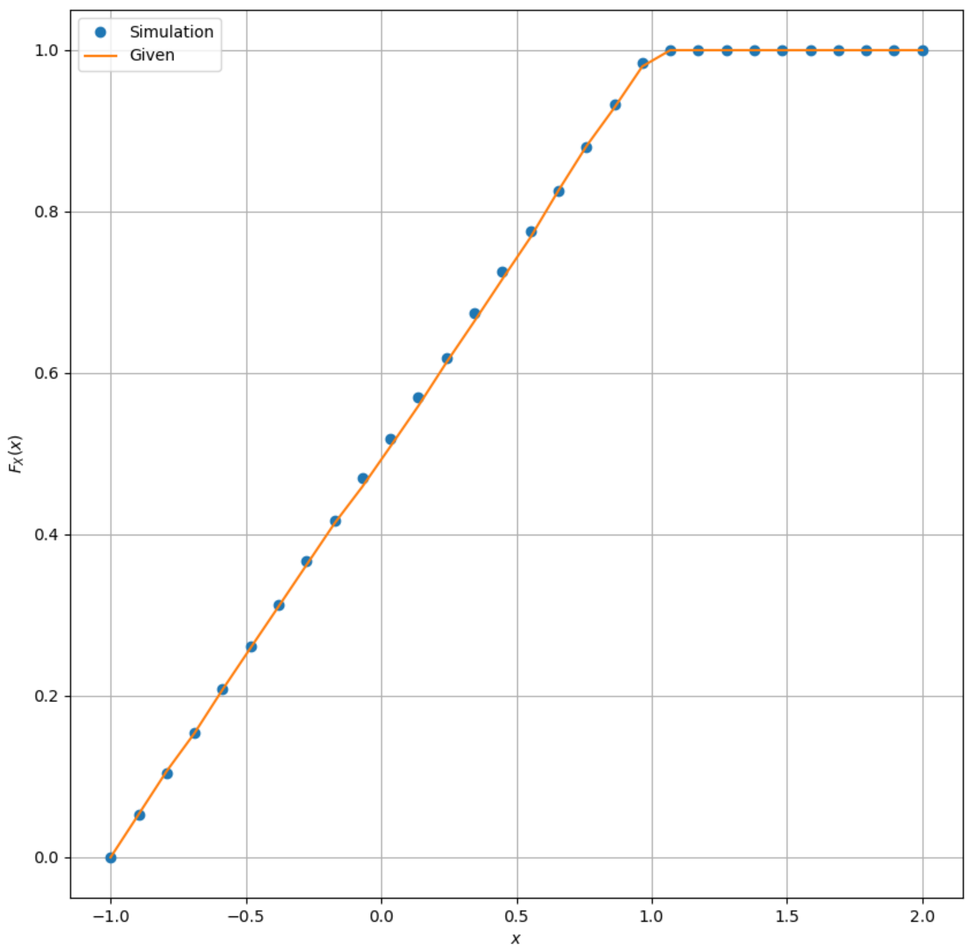


Figure 2.5: $2e^{-X}-1$ vs. $U(-1,1)$
Graphs match, $\therefore$ correct option

2.7 Question: Let X be a positive valued continuous random variable with finite mean μ. 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,

$$X = Y + Z \tag{2.91}$$

Here, Z is an uniform distrubtion.

$$Z \sim U[0, 1) \tag{2.92}$$

$$F_Z(x) = x \tag{2.93}$$

$$E(Z) = \frac{1}{2} \tag{2.94}$$

Consider

(a)

$$\Pr(Y \leq \mu) = \Pr(X - Z \leq \mu) \tag{2.95}$$
$$= \Pr(Z \geq X - \mu) \tag{2.96}$$
$$= E(1 - F_Z(X - \mu)) \tag{2.97}$$
$$= E(1 - X + \mu) \tag{2.98}$$
$$= 1 - E(X) + \mu \tag{2.99}$$
$$= 1 \tag{2.100}$$

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

(b)

$$\Pr(Y \geq \mu) = \Pr(X - Z \geq \mu) \tag{2.101}$$
$$= \Pr(Z \leq X - \mu) \tag{2.102}$$
$$= E(F_Z(X - \mu)) \tag{2.103}$$
$$= E(X - \mu) \tag{2.104}$$
$$= E(X) - \mu \tag{2.105}$$
$$= 0 \tag{2.106}$$

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

(c)

$$E(Y) = E(X - Z) \tag{2.107}$$
$$= E(X) - E(Z) \tag{2.108}$$
$$= \mu - \frac{1}{2} \tag{2.109}$$
$$= E(X) - \frac{1}{2} \tag{2.110}$$

$E(X) > E(Y)$. 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 $\in$ 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

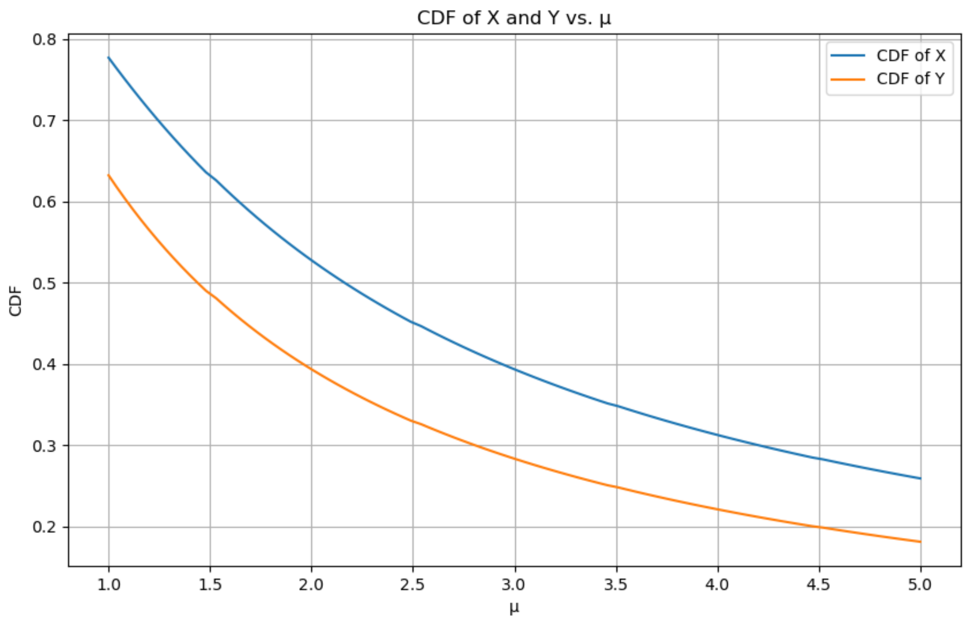


Figure 2.6: CDF'S of X and Y for varying μ at x=1.5

(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.

$$p_X(k) = {}^nC_k p^k q^{n-k} \qquad \forall k = 0, 1, 2 \tag{2.111}$$

$$= {}^2C_k \left(\frac{1}{2}\right)^k \left(\frac{1}{2}\right)^{2-k} \tag{2.112}$$

$$= {}^2C_k \left(\frac{1}{2}\right)^2 \tag{2.113}$$

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

Therefore,

$$\Pr(X \le 1) = 1 - \Pr(X = 2) \tag{2.114}$$

$$= 1 - \frac{1}{4} \tag{2.115}$$

$$= \frac{3}{4} \tag{2.116}$$

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
	0	11
X	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

$$\Pr\left(Y = i\right) = {}^{n}C_{i}(p)^{i}(q)^{(n-i)} \tag{2.117}$$

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

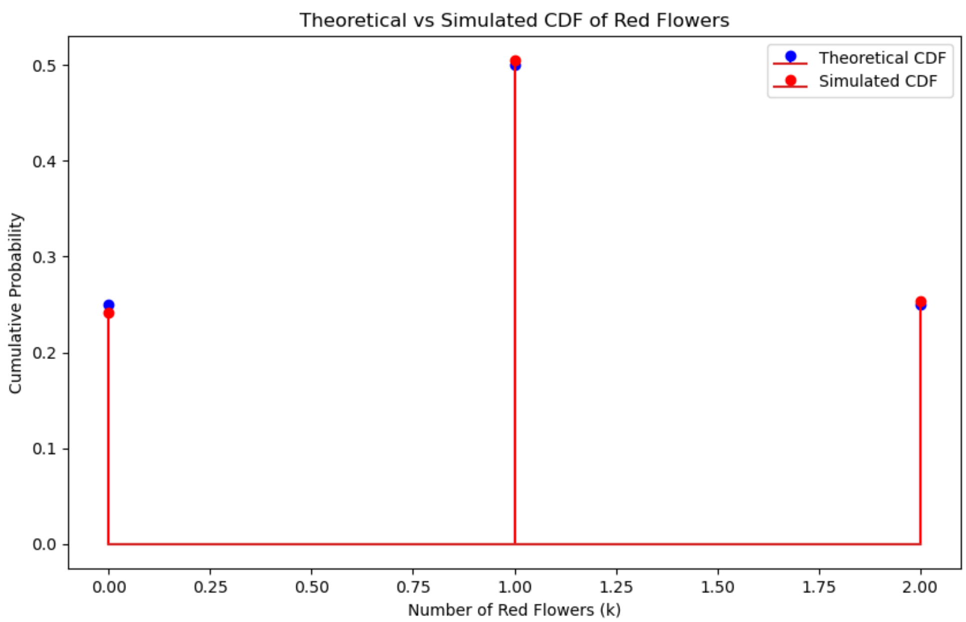


Figure 2.7: Simulation vs Theoretical

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

$$
\begin{aligned}
\Pr(Y \geq 1) &= \Pr(Y = 1) + \Pr(Y = 2) && (2.118)\\
&= {}^2C_1(p)(q) + {}^2C_2(p)^2(1) && (2.119)\\
&= 2pq + p^2 && (2.120)\\
&= 0.5 + 0.25 && (2.121)\\
&= 0.75 && (2.122)
\end{aligned}
$$

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

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

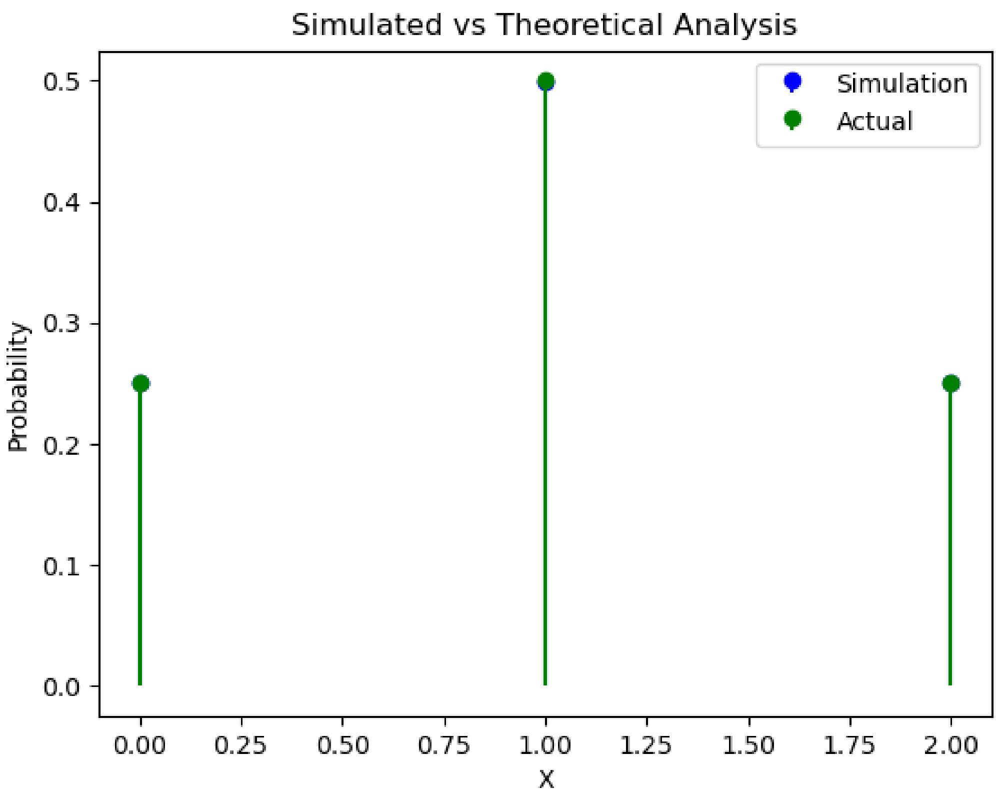
(b) $\frac{1-e^{-\lambda}}{\lambda}$

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

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

(GATE ST 2023)

Solution:



(A) Theory

$$X \sim Pois\left(\lambda\right) \tag{2.123}$$

$$\Pr\left(X = k\right) = e^{-\lambda}\frac{\lambda^k}{k!}; k \geq 0 \tag{2.124}$$

we know that

$$E\left(A|B\right) = \frac{E\left(A,B\right)}{\Pr\left(B\right)} \tag{2.125}$$

$$\implies E\left(\frac{1}{1+X}\middle|X>0\right) = \frac{\sum_{k=1}^{\infty}\frac{1}{k+1}\Pr\left(X=k\right)}{\Pr\left(X>0\right)} \tag{2.126}$$

$$= \frac{\sum_{k=1}^{\infty}\frac{1}{k+1}e^{-\lambda}\frac{\lambda^k}{k!}}{1-\Pr\left(X\leq 0\right)} \tag{2.127}$$

$$= \frac{e^{-\lambda}\sum_{k=1}^{\infty}\frac{\lambda^k}{(k+1)!}}{1-\Pr\left(X=0\right)} \tag{2.128}$$

from equation (2.124)

$$= \frac{e^{-\lambda}\sum_{k=1}^{\infty}\frac{\lambda^k}{(k+1)!}}{1-e^{\lambda}} \tag{2.129}$$

Now simplifying Just the Summation

$$\implies \sum_{k=1}^{\infty}\frac{\lambda^k}{(k+1)!} \tag{2.130}$$

$$= \frac{1}{\lambda}\sum_{k=1}^{\infty}\frac{\lambda^{k+1}}{k+1} \tag{2.131}$$

Letting $k+1=m$,

$$\implies \frac{1}{\lambda}\sum_{m=2}^{\infty}\frac{\lambda^m}{m!} \tag{2.132}$$

We Know from Taylor series

$$e^{x} = \sum_{k=0}^{\infty} \frac{x^{k}}{k!} \tag{2.133}$$

$$\implies \frac{1}{\lambda} \sum_{m=2}^{\infty} \frac{\lambda^{m}}{m!} = \frac{1}{\lambda} \left(e^{\lambda} - 1 - \lambda \right) \tag{2.134}$$

Substituting back we get,

$$= \frac{e^{-\lambda}}{(1 - e^{\lambda})} \left(\frac{1}{\lambda} \left(e^{\lambda} - 1 - \lambda \right) \right) \tag{2.135}$$

$$= \frac{e^{-\lambda}}{\lambda (1 - e^{\lambda})} \left(e^{\lambda} - 1 - \lambda \right) \tag{2.136}$$

$$= \frac{1 - e^{-\lambda} - \lambda e^{-\lambda}}{\lambda (1 - e^{\lambda})} \tag{2.137}$$

(B) Simulation

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

$$F(x) = \sum_{k=0}^{x} e^{-\lambda} \frac{\lambda^{k}}{k!} \tag{2.138}$$

(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 λ.

(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 $\Omega = \{1, 2, 3, 4\}$, $\mathcal{G} = \{\emptyset, \Omega, \{1\}, \{4\}, \{2, 3\}, \{1, 4$ $P(\{1\}) = \frac{1}{4}$.

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: Probablity space

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

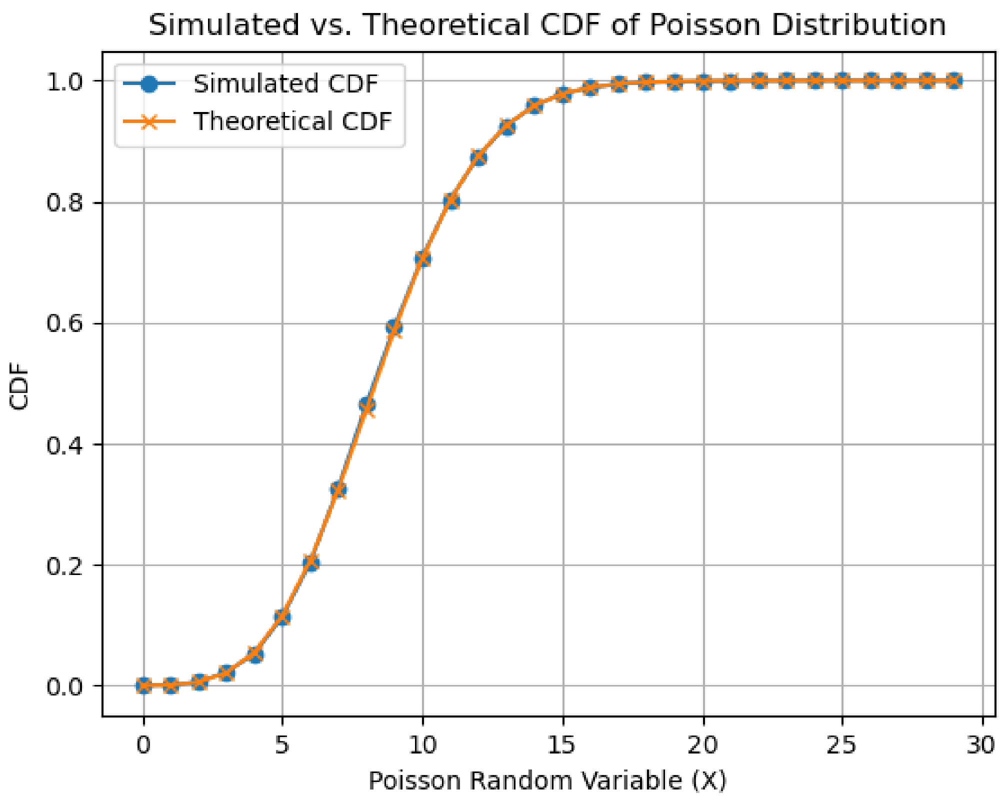


Figure 2.8: Simulation Vs theoretical cdf of poisson distribution with $\lambda = 9$

Pmf is defined as

$$p_x\left(k\right) = \begin{cases} P(\{1\}) & ,k=1 \\ P(\{2,3\}) & ,k=2 \\ P(\{4\}) & ,k=3 \end{cases} \tag{2.139}$$

Table 2.10: Random variable

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

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

$$p_X(k) = \begin{cases} \frac{1}{4} & ,k=1 \\ p & ,k=2 \\ q & ,k=3 \end{cases} \tag{2.140}$$

$$\Pr(\{1,4\}) = p_X(1) + p_X(3) \tag{2.141}$$

We know

$$p_X(1) + p + q = 1 \tag{2.142}$$

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

$$\Pr(X \le 2) = p_X(1) + p \tag{2.143}$$

$$\tag{2.144}$$

We can expres above equations as:

$$\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} p \\ q \end{pmatrix} = \begin{pmatrix} \frac{3}{4} \\ \frac{1}{2} \end{pmatrix} \tag{2.145}$$

$$p = \frac{1}{2}, q = \frac{1}{4} \tag{2.146}$$

Finally

$$\Pr\left(\{1,4\}\right) = P(\{1\}) + q \tag{2.147}$$

$$\Pr\left(\{1,4\}\right) = \frac{1}{4} + \frac{1}{4} \tag{2.148}$$

$$\Pr\left(\{1,4\}\right) = 0.5 \tag{2.149}$$

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.

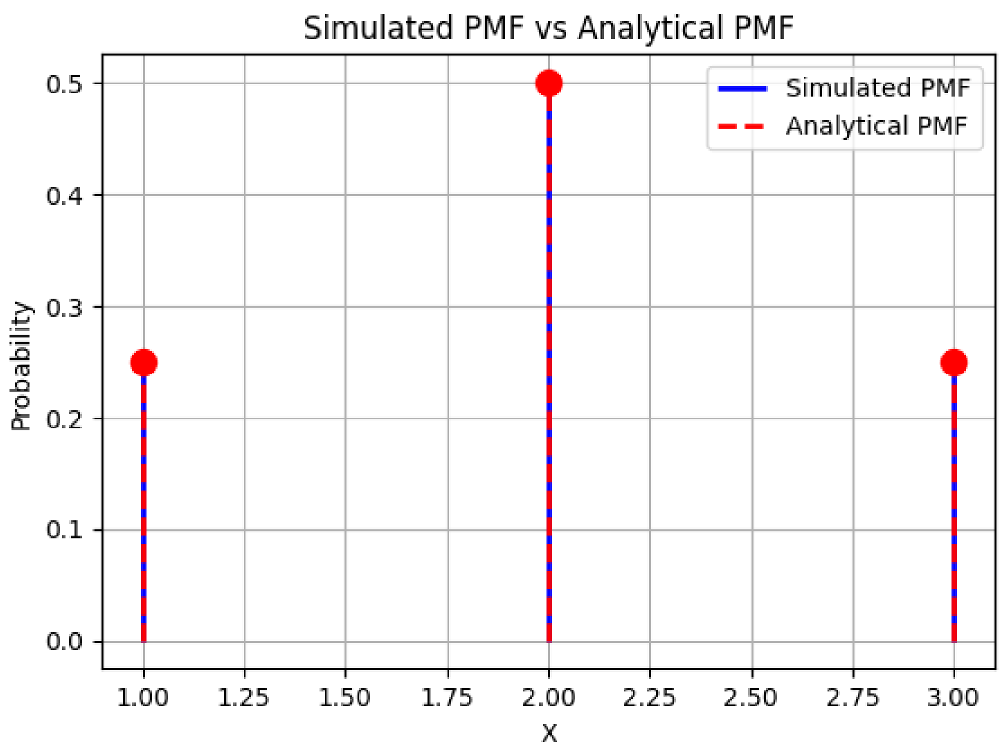


Figure 2.9: Analytical vs simulated

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

(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,

$$\Pr\left(N(t)=k\right)=Po(t;k)=\frac{(\lambda t)^k e^{-\lambda t}}{k!} \tag{2.150}$$

here λ is 1

$$\Pr\left(N(t)=k\right)=\frac{(t)^k e^{-t}}{k!} \tag{2.151}$$

$$\tag{2.152}$$

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

$$\Pr\left(X=k, X+Y=n\right)=pr\left(X=k, Y=n-k\right) \tag{2.153}$$

$$=\frac{(t_1)^k}{k!}e^{-t_1}\frac{(t_2)^{n-k}}{(n-k)!}e^{-t_2} \tag{2.154}$$

$$=e^{-(t_1+t_2)}\left(\frac{(t_1+t_2)^n}{n!}\right){}^nC_k\left(\frac{t_1}{t_1+t_2}\right)^k\left(\frac{t_2}{t_1+t_2}\right)^{n-k} \tag{2.155}$$

$$\Pr(X+Y=n) = e^{-(t_1+t_2)}\left(\frac{(t_1+t_2)^n}{n!}\right) \tag{2.156}$$

From conditional probability, from the equations (2.155) and (2.156)

$$\Pr(X=k|X+Y=n) = \frac{\Pr(X=k, Y=n-k)}{\Pr(X+Y=n)} \tag{2.157}$$

$$= {}^nC_k\left(\frac{t_1}{t_1+t_2}\right)^k\left(\frac{t_2}{t_1+t_2}\right)^{n-k} \tag{2.158}$$

For the given question,

Parameter	Values
t_1	3
t_1	5

Table 2.12: Table 1

$$\Pr(N(3)=3|N(5)=5) = {}^5C_3\left(\frac{3}{2+3}\right)^3\left(\frac{2}{2+3}\right)^2 \tag{2.159}$$

$$= {}^5C_3\left(\frac{3}{5}\right)^3\left(\frac{2}{5}\right)^2 \tag{2.160}$$

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 $\mathbf{X_1}, \mathbf{X_2}, \ldots, \mathbf{X_n}, \mathbf{Y_1}, \mathbf{Y_2}, \ldots, \mathbf{Y_n}$ are independent and identically distributed random vectors each having $N_p\left(\boldsymbol{\mu}, \Sigma\right)$ distributions, where Σ 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 $c(\mathbf{X}-\boldsymbol{\mu})^T\Sigma^{-1}\left(\mathbf{X}-\boldsymbol{\mu}\right)$ has χ^2-distribution with p degrees of freedom.

(b) There exists $c > 0$ such that $c(\mathbf{X}-\mathbf{Y})^T\Sigma^{-1}\left(\mathbf{X}-\mathbf{Y}\right)$ has χ^2-distribution with $(p-1)$ degrees of freedom.

(c) There exists $c > 0$ such that $c\sum_{i=1}^{n}\left(\mathbf{X_i}-\mathbf{X}\right)^T\Sigma^{-1}\left(\mathbf{X_i}-\mathbf{X}\right)$ has χ^2-distribution with p degrees of freedom.

(d) There exists $c > 0$ such that $c\sum_{i=1}^{n}\left(\mathbf{X_i}-\mathbf{Y_i}-\mathbf{X}+\mathbf{Y}\right)^T\Sigma^{-1}\left(\mathbf{X_i}-\mathbf{Y_i}-\mathbf{X}+\mathbf{Y}\right)$ has χ^2-distribution with p degrees of freedom. GATE ST Paper 2023

Solution:

We are given that,

$$\mathbf{X_1}, \mathbf{X_2}, \ldots, \mathbf{X_n}, \mathbf{Y_1}, \mathbf{Y_2}, \ldots, \mathbf{Y_n} \sim N_p\left(\boldsymbol{\mu}, \Sigma\right) \tag{2.161}$$

Also,

$$\mathbf{X} = \frac{1}{n}\sum_{i=1}^{n} \mathbf{X_i} \tag{2.162}$$

$$\mathbf{Y} = \frac{1}{n}\sum_{i=1}^{n} \mathbf{Y_i} \tag{2.163}$$

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

$$\boldsymbol{\mu}_{\mathbf{X}} = E\left(\mathbf{X}\right) \tag{2.164}$$

$$= \frac{1}{n}\sum_{i=1}^{n} E\left(\mathbf{X_i}\right) \tag{2.165}$$

$$= \boldsymbol{\mu} \tag{2.166}$$

Similarly,

$$\boldsymbol{\mu}_{\mathbf{Y}} = \boldsymbol{\mu} \tag{2.167}$$

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

$$\begin{aligned}
\Sigma_{\mathbf{X}} &= E\left[(\mathbf{X}-\boldsymbol{\mu})(\mathbf{X}-\boldsymbol{\mu})^T\right] && (2.168)\\
&= E\left[\left(\frac{1}{n}\sum_{i=1}^{n}\mathbf{X_i}-\boldsymbol{\mu}\right)\left(\frac{1}{n}\sum_{i=1}^{n}\mathbf{X_i}-\boldsymbol{\mu}\right)^T\right] && (2.169)\\
&= \frac{1}{n^2}E\left[\sum_{i=1}^{n}(\mathbf{X_i}-\boldsymbol{\mu})(\mathbf{X_i}-\boldsymbol{\mu})^T\right] && (2.170)\\
&= \frac{1}{n^2}\left[\sum_{i=1}^{n}E\left(\mathbf{X_i}^2+\boldsymbol{\mu}^2-2\boldsymbol{\mu}\mathbf{X_i}\right)\right] && (2.171)\\
&= \frac{1}{n^2}\left[\sum_{i=1}^{n}E\left(\mathbf{X_i}^2\right)+\sum_{i=1}^{n}E\left(\boldsymbol{\mu}^2\right)-2\boldsymbol{\mu}\sum_{i=1}^{n}E\left(\mathbf{X_i}\right)\right] && (2.172)\\
&= \frac{1}{n^2}\left[n\Sigma+n\boldsymbol{\mu}^2+n\boldsymbol{\mu}^2-2\boldsymbol{\mu}^2\right] \qquad \left[\because E\left(\mathbf{X_i}^2\right)=\Sigma_{\mathbf{X_i}}+E\left(\mathbf{X_i}\right)^2\right] && (2.173)\\
&= \frac{\Sigma}{n} && (2.174)
\end{aligned}$$

Similarly,

$$\Sigma_{\mathbf{Y}} = \frac{\Sigma}{n} \tag{2.175}$$

(a) To check option (A):

let us say,

$$\mathbf{A} = c(\mathbf{X}-\boldsymbol{\mu})^T\Sigma^{-1}(\mathbf{X}-\boldsymbol{\mu}) \tag{2.176}$$

$$\tag{2.177}$$

And,

$$\Sigma^{-1} = \mathbf{F}^T\mathbf{F} \tag{2.178}$$

$$\mathbf{y} = \mathbf{F}\left(\mathbf{X} - \boldsymbol{\mu}\right) \tag{2.179}$$

$$\implies \mathbf{A} = c\mathbf{y}^T\overline{\mathbf{y}} \tag{2.180}$$

$$= c\|\mathbf{y}\|^2 \tag{2.181}$$

Equation (2.181) shows that $\mathbf{A}$ can have χ^2-distribution.

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

$$\boldsymbol{\mu}_{\mathbf{y}} = E\left(\mathbf{y}\right) \tag{2.182}$$

$$= E\left(\mathbf{F}\left(\mathbf{X} - \boldsymbol{\mu}\right)\right) \tag{2.183}$$

$$= \mathbf{F}\left[E\left(\mathbf{X}\right) - E\left(\boldsymbol{\mu}\right)\right] \tag{2.184}$$

$$= \mathbf{F}\left[\boldsymbol{\mu} - \boldsymbol{\mu}\right] \quad from(2.166) \tag{2.185}$$

$$= 0 \tag{2.186}$$

And,

$$\Sigma_{\mathbf{y}} = E\left[\left(\mathbf{y} - \boldsymbol{\mu}_{\mathbf{y}}\right)\left(\mathbf{y} - \boldsymbol{\mu}_{\mathbf{y}}\right)^T\right] \tag{2.187}$$

$$= E\left[\left(\mathbf{F}\left(\mathbf{X} - \boldsymbol{\mu}\right)\right)\left(\mathbf{F}\left(\mathbf{X} - \boldsymbol{\mu}\right)\right)^T\right] \tag{2.188}$$

$$= E\left[\mathbf{F}\left(\mathbf{X} - \boldsymbol{\mu}\right)\left(\mathbf{X} - \boldsymbol{\mu}\right)^T\mathbf{F}^T\right] \tag{2.189}$$

$$= \mathbf{F}\left[E\left[\left(\mathbf{X} - \boldsymbol{\mu}\right)\left(\mathbf{X} - \boldsymbol{\mu}\right)^T\right]\right]\mathbf{F}^T \tag{2.190}$$

$$= \mathbf{F}\Sigma\mathbf{F}^T \tag{2.191}$$

since,

$$\Sigma^{-1} = \mathbf{F}^T\mathbf{F} \tag{2.192}$$

$$\Sigma\Sigma^{-1} = \Sigma\mathbf{F}^T\mathbf{F} \tag{2.193}$$

$$\mathbf{I} = \Sigma\mathbf{F}^T\mathbf{F} \tag{2.194}$$

$$\mathbf{I}\mathbf{F}^{-1} = \Sigma\mathbf{F}^T \tag{2.195}$$

$$\mathbf{F}\mathbf{F}^{-1} = \mathbf{F}\Sigma\mathbf{F}^T \tag{2.196}$$

$$\mathbf{I} = \mathbf{F}\Sigma\mathbf{F}^T \tag{2.197}$$

So using (2.197),

$$\Sigma_{\mathbf{y}} = \mathbf{I} \tag{2.198}$$

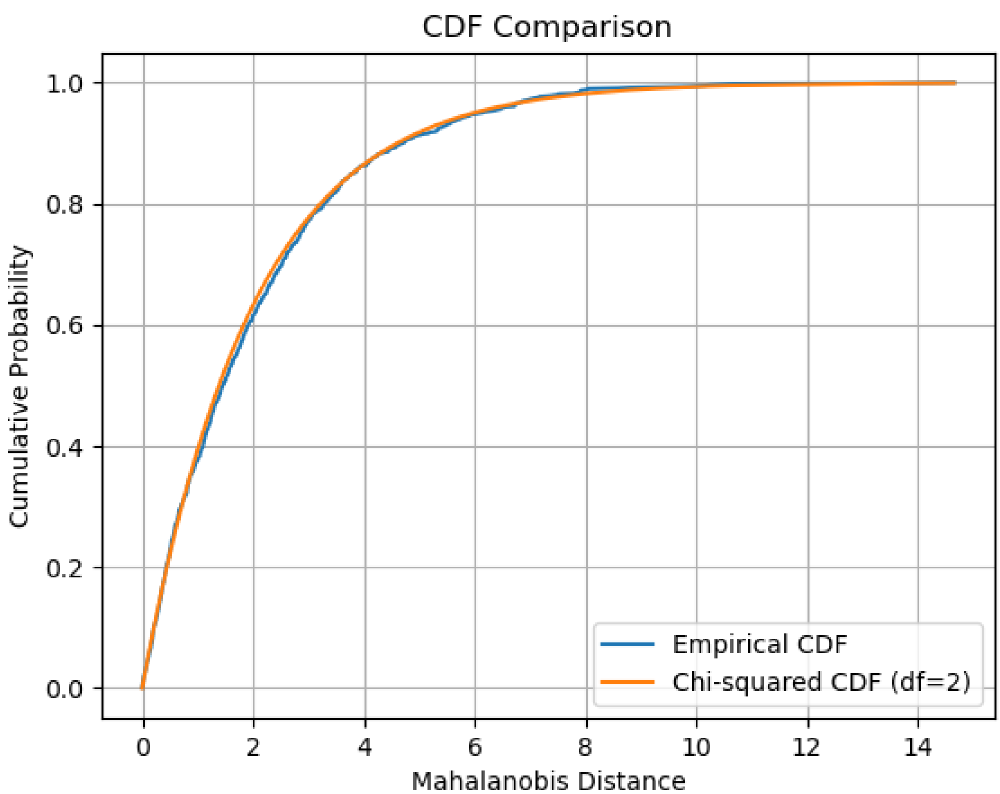
Hence, For $c = 1$ $\mathbf{A}$ has χ^2-distribution with p degrees of freedom.

So option (A) is correct.

(b) To check option (B):

Let us say,

$$\mathbf{B} = c(\mathbf{X} - \mathbf{Y})^T\Sigma^{-1}(\mathbf{X} - \mathbf{Y}) \tag{2.199}$$



And,

$$\Sigma^{-1} = \mathbf{F}^T\mathbf{F} \tag{2.200}$$

$$\mathbf{y} = \mathbf{F}\left(\mathbf{X} - \mathbf{Y}\right) \tag{2.201}$$

$$\implies \mathbf{B} = c\mathbf{y}^T\overline{\mathbf{y}} \tag{2.202}$$

$$= c\|\mathbf{y}\|^2 \tag{2.203}$$

Equation (2.203) shows that $\mathbf{B}$ can have χ^2-distribution.

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

$$\boldsymbol{\mu}_{\mathbf{y}} = E(\mathbf{y}) \tag{2.204}$$
$$= E[F(\mathbf{X} - \mathbf{Y})] \tag{2.205}$$
$$= F[E(\mathbf{X}) - E(\mathbf{Y})] \tag{2.206}$$
$$= F[\boldsymbol{\mu} - \boldsymbol{\mu}] \tag{2.207}$$
$$= 0 \tag{2.208}$$

And,

$$\Sigma_{\mathbf{y}} = E\left[(\mathbf{y} - \boldsymbol{\mu}_{\mathbf{y}})(\mathbf{y} - \boldsymbol{\mu}_{\mathbf{y}})^T\right] \tag{2.209}$$
$$= E\left[(\mathbf{F}(\mathbf{X} - \mathbf{Y}))(\mathbf{F}(\mathbf{X} - \mathbf{Y}))^T\right] \tag{2.210}$$
$$= E\left[\mathbf{F}(\mathbf{X} - \mathbf{Y})(\mathbf{X} - \mathbf{Y})^T\mathbf{F}^T\right] \tag{2.211}$$
$$= \mathbf{F}\left[E\left[(\mathbf{X} - \mathbf{Y})(\mathbf{X} - \mathbf{Y})^T\right]\right]\mathbf{F}^T \tag{2.212}$$
$$= \mathbf{F}\left[E\left[\|\mathbf{X} - \mathbf{Y}\|^2\right]\right]\mathbf{F}^T \tag{2.213}$$
$$= \mathbf{F}\left[E\left(\mathbf{X}^2\right) + E\left(\mathbf{Y}^2\right) - E(2\mathbf{X}\mathbf{Y})\right]\mathbf{F}^T \tag{2.214}$$
$$= \mathbf{F}\left[\frac{\Sigma}{n} + \boldsymbol{\mu}^2 + \frac{\Sigma}{n} + \boldsymbol{\mu}^2 - 2\boldsymbol{\mu}^2\right]\mathbf{F}^T \quad \left[\because E\left(\mathbf{X}^2\right) = \Sigma_{\mathbf{X}} + E(\mathbf{X})^2\right] \tag{2.215}$$
$$= \frac{2}{n}\mathbf{F}\Sigma\mathbf{F}^T \tag{2.216}$$
$$= \frac{2}{n}\mathbf{I} \tag{2.217}$$

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

So option (B) is incorrect.

(c) To check option (C):

let us say,

$$\mathbf{C} = c\sum_{i=1}^{n}\left(\mathbf{X_i} - \mathbf{X}\right)^T\Sigma^{-1}\left(\mathbf{X_i} - \mathbf{X}\right) \tag{2.218}$$

And,

$$\Sigma^{-1} = \mathbf{F}^T\mathbf{F} \tag{2.219}$$

$$\mathbf{y} = \mathbf{F}\left(\sum_{i=1}^{n}\left(\mathbf{X_i} - \mathbf{X}\right)\right) \tag{2.220}$$

$$\implies \mathbf{C} = c\mathbf{y}^T\overline{\mathbf{y}} \tag{2.221}$$

$$= c\|\mathbf{y}\|^2 \tag{2.222}$$

Equation (2.222) shows that $\mathbf{C}$ can have χ^2-distribution.

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

$$\boldsymbol{\mu}_{\mathbf{y}} = E\left(\mathbf{y}\right) \tag{2.223}$$

$$= E\left[\mathbf{F}\left(\sum_{i=1}^{n}\left(\mathbf{X_i} - \mathbf{X}\right)\right)\right] \tag{2.224}$$

$$= \mathbf{F}\left[\sum_{i=1}^{n}\left(E\left(\mathbf{X_i}\right) - E\left(\mathbf{X}\right)\right)\right] \tag{2.225}$$

$$= \mathbf{F}\left[E\left(X_1\right) - E\left(X\right) + E\left(X_2\right) - E\left(X\right) + \ldots + E\left(X_n\right) - E\left(X\right)\right] \tag{2.226}$$

$$= 0 \tag{2.227}$$

And,

$$\Sigma_{\mathbf{y}} = E\left[(\mathbf{y}-\boldsymbol{\mu}_{\mathbf{y}})(\mathbf{y}-\boldsymbol{\mu}_{\mathbf{y}})^T\right] \tag{2.228}$$

$$= \mathbf{F}E\left[\left(\sum_{i=1}^{n}(\mathbf{X_i}-\mathbf{X})\right)\left(\sum_{i=1}^{n}(\mathbf{X_i}-\mathbf{X})\right)^T\right]\mathbf{F}^T \tag{2.229}$$

$$= \mathbf{F}E\left[\left(\sum_{i=1}^{n}\mathbf{X_i}-n\mathbf{X}\right)\left(\sum_{i=1}^{n}\mathbf{X_i}-n\mathbf{X}\right)^T\right]\mathbf{F}^T \tag{2.230}$$

$$= \mathbf{F}E\left[(n\mathbf{X}-n\mathbf{X})(n\mathbf{X}-n\mathbf{X})^T\right]\mathbf{F}^T \tag{2.231}$$

$$= \mathbf{F}\mathbf{0}\mathbf{F}^T \tag{2.232}$$

$$= \mathbf{0} \tag{2.233}$$

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

So option (C) is incorrect.

(d) To check option (D):

let us say,

$$\mathbf{D} = c\sum_{i=1}^{n}(\mathbf{X_i}-\mathbf{Y_i}-\mathbf{X}+\mathbf{Y})^T\Sigma^{-1}(\mathbf{X_i}-\mathbf{Y_i}-\mathbf{X}+\mathbf{Y}) \tag{2.234}$$

And,

$$\Sigma^{-1} = \mathbf{F}^T\mathbf{F} \tag{2.235}$$

$$\mathbf{y} = \mathbf{F}\left(\sum_{i=1}^{n}(\mathbf{X_i}-\mathbf{Y_i}-\mathbf{X}+\mathbf{Y})\right) \tag{2.236}$$

$$\implies \mathbf{C} = c\mathbf{y}^T\overline{\mathbf{y}} \tag{2.237}$$

$$= c\|\mathbf{y}\|^2 \tag{2.238}$$

Equation (2.238) shows that $\mathbf{D}$ can have χ^2-distribution.

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

$$\boldsymbol{\mu}_{\mathbf{y}} = E\left(\mathbf{y}\right) \tag{2.239}$$

$$= E\left(\mathbf{F}\left(\sum_{i=1}^{n}\left(\mathbf{X_i} - \mathbf{Y_i} - \mathbf{X} + \mathbf{Y}\right)\right)\right) \tag{2.240}$$

$$= \mathbf{F}E\left[\sum_{i=1}^{n}\mathbf{X_i} - \sum_{i=1}^{n}\mathbf{Y_i} - n\mathbf{X} + n\mathbf{Y}\right] \tag{2.241}$$

$$= \mathbf{F}\left[\sum_{i=1}^{n}E\left(\mathbf{X_i}\right) - \sum_{i=1}^{n}E\left(\mathbf{Y_i}\right) - nE\left(\mathbf{X}\right) + nE\left(\mathbf{Y}\right)\right] \tag{2.242}$$

$$= \mathbf{F}\left[n\boldsymbol{\mu} - n\boldsymbol{\mu} - n\boldsymbol{\mu} + n\boldsymbol{\mu}\right] \tag{2.243}$$

$$= 0 \tag{2.244}$$

And,

$$\Sigma_{\mathbf{y}} = E\left[\left(\mathbf{y} - \boldsymbol{\mu}_{\mathbf{y}}\right)\left(\mathbf{y} - \boldsymbol{\mu}_{\mathbf{y}}\right)^T\right] \tag{2.245}$$

$$= \mathbf{F}E\left[\left(\sum_{i=1}^{n}\left(\mathbf{X_i} - \mathbf{Y_i} - \mathbf{X} + \mathbf{Y}\right)\right)\left(\sum_{i=1}^{n}\left(\mathbf{X_i} - \mathbf{Y_i} - \mathbf{X} + \mathbf{Y}\right)\right)^T\right]\mathbf{F}^T \tag{2.246}$$

$$= \mathbf{F}E\left[\left(\sum_{i=1}^{n}\mathbf{X_i} - \sum_{i=1}^{n}\mathbf{Y_i} - n\mathbf{X} + n\mathbf{Y}\right)\left(\sum_{i=1}^{n}\mathbf{X_i} - \sum_{i=1}^{n}\mathbf{Y_i} - n\mathbf{X} + n\mathbf{Y}\right)^T\right]\mathbf{F}^T \tag{2.247}$$

$$= \mathbf{F}E\left[\left(n\mathbf{X} - n\mathbf{Y} - n\mathbf{X} + n\mathbf{Y}\right)\left(n\mathbf{X} - n\mathbf{Y} - n\mathbf{X} + n\mathbf{Y}\right)^T\right]\mathbf{F}^T \tag{2.248}$$

$$= \mathbf{F}\mathbf{0}\mathbf{F}^T \tag{2.249}$$

$$= \mathbf{0} \tag{2.250}$$

Hence, There is no value of $c > 0$ for which $\mathbf{D}$ have χ^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,

$$\boldsymbol{\mu} = (\mu_1, \mu_2, \cdots, \mu_N)^T \tag{2.251}$$

and positive covariance matrix Σ, the mahalanobis distance of a point,

$$\mathbf{X} = (X_1, X_2, \cdots, X_N)^T \tag{2.252}$$

is given by,

$$d_M(\mathbf{X}, Q) = \sqrt{(\mathbf{X} - \boldsymbol{\mu})^T \Sigma^{-1} (\mathbf{X} - \boldsymbol{\mu})} \tag{2.253}$$

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 1×1 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 χ^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,

$$f_{X_1}(x_1) = e^{-x_1}, \quad x_1 > 0 \tag{2.254}$$

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

$$f_{X_2}(x_2) = 2e^{-2x_2}, \quad x_2 > 0 \tag{2.255}$$

The probability of occuring more than 1 unit of storm event is

(GATE AG 2023)

Solution:

Laplace Transform

Let X be a random variable such that

$$X = X_1 + X_2 \tag{2.256}$$

Given,

$$f_{X_1}(x) = e^{-x}u(x) \tag{2.257}$$

$$f_{X_2}(x) = 2e^{-2x}u(x) \tag{2.258}$$

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

CDF of X_1:

$$\begin{aligned} F_{X_1}(x) &= \int_{-\infty}^{x} f_{X_1}(x)\,dx && (2.259)\\ &= \int_{0}^{x} e^{-x}dx && (2.260)\\ &= 1 - e^{-x} && (2.261) \end{aligned}$$

CDF of X_2:

$$\begin{aligned} F_{X_2}(x) &= \int_{-\infty}^{x} f_{X_2}(x)\,dx && (2.262)\\ &= \int_{0}^{x} 2e^{-2x}dx && (2.263)\\ &= 1 - e^{-2x} && (2.264) \end{aligned}$$

Now,

$$M_X(s) = M_{X_1}(s) \cdot M_{X_2}(s) \tag{2.265}$$

$$M_{X_1}(s) = \int_{-\infty}^{\infty} f_{X_1}(x) \cdot e^{-sx} dx \tag{2.266}$$

$$= \int_{0}^{\infty} e^{-x} \cdot e^{-sx} u(x)\, dx \tag{2.267}$$

$$= \frac{1}{s+1} \tag{2.268}$$

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

$$Re(s) > -1 \tag{2.269}$$

Now,

$$M_{X_2}(s) = \int_{-\infty}^{\infty} f_{X_2}(x) \cdot e^{-sx} dx \tag{2.270}$$

$$= \int_{0}^{\infty} 2e^{-2x} \cdot e^{-sx} u(x)\, dx \tag{2.271}$$

$$= \frac{2}{s+2} \tag{2.272}$$

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

$$Re(s) > -2 \tag{2.273}$$

Using (2.268) and (2.272) in (2.265)

$$M_X(s) = \frac{1}{s+1} \times \frac{2}{s+2} \tag{2.274}$$

$$= \frac{2}{(s+1)(s+2)} \tag{2.275}$$

$$
\begin{aligned}
p_X(x) &= L^{-1}\left[M_X(s)\right] && (2.276)\\
&= L^{-1}\left[\frac{2}{(s+1)(s+2)}\right] && (2.277)\\
&= 2L^{-1}\left[\frac{1}{s+1} - \frac{1}{s+2}\right] && (2.278)
\end{aligned}
$$

ROC of laplace transform:

$$
\begin{aligned}
Re(s) > -1 \cap Re(s) > -2 && (2.279)\\
\implies Re(s) > -1 && (2.280)
\end{aligned}
$$

So, now

$$
p_X(x) = 2\left(e^{-x} - e^{-2x}\right)u(x) \tag{2.281}
$$

CDF of X:

$$
\begin{aligned}
F_X(x) &= \int_{-\infty}^{x} p_X(x)\,dx && (2.282)\\
&= 2\int_{-\infty}^{x} \left(e^{-x} - e^{-2x}\right)u(x)\,dx && (2.283)
\end{aligned}
$$

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

$$
\begin{aligned}
F_X(x) &= 2\int_{0}^{x} \left(e^{-x} - e^{-2x}\right)dx && (2.284)\\
&= 2\left(-e^{-x} + \frac{1}{2}e^{-2x}\right)\Bigg|_{0}^{x} && (2.285)\\
&= 2\left(-e^{-x} + \frac{1}{2}e^{-2x} + \frac{1}{2}\right) && (2.286)\\
&= -2e^{-x} + e^{-2x} + 1 && (2.287)
\end{aligned}
$$

Now,

$$\Pr\left(X>1\right)=F_X\left(1\right) \tag{2.288}$$

$$=-2e^{-1}+e^{-2}+1 \tag{2.289}$$

$$=0.39 \tag{2.290}$$

Steps for simulation the given distribution

(a) CDF of X_1 is given as:

$$F_{X_1}(x)=\begin{cases}0 & x\leq 0\\ 1-e^{-x} & x>0\end{cases} \tag{2.291}$$

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

For x ≤ 0

$$u=0 \tag{2.292}$$

$$\because x\leq 0 \tag{2.293}$$

$$u\leq 0 \tag{2.294}$$

For x > 0

$$u = 1 - e^{-x} \tag{2.295}$$

$$e^{-x} = 1 - u \tag{2.296}$$

$$x = -\ln(1-u) \tag{2.297}$$

$$\because x > 0 \tag{2.298}$$

$$u > 0 \tag{2.299}$$

$$I(u) = \begin{cases} 0 & u \leq 0 \\ -\ln(1-u) & u > 0 \end{cases} \tag{2.300}$$

(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 funtion 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

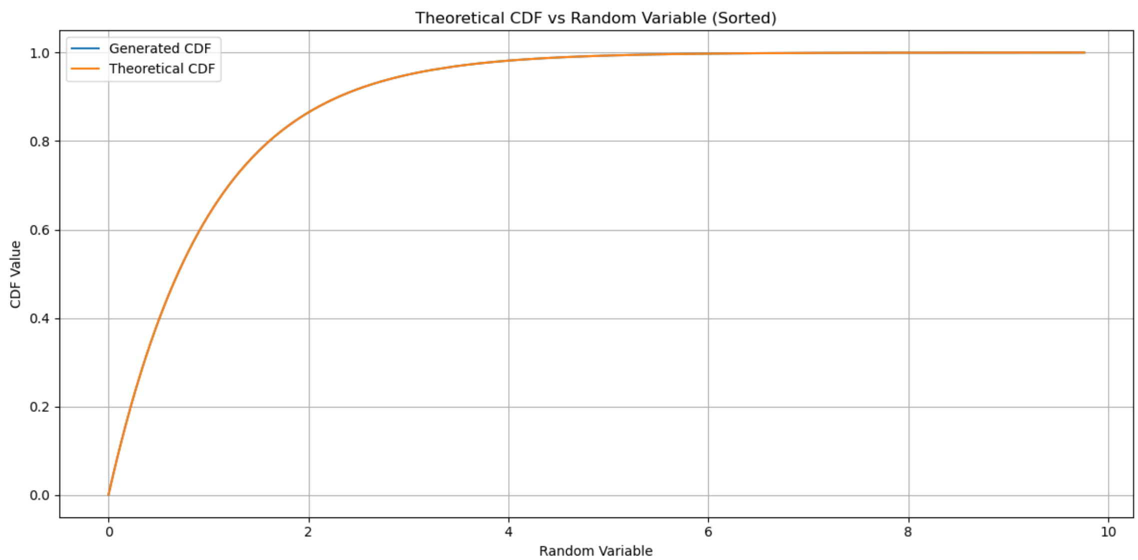


Figure 2.10: Theoretical vs Simulation Analysis wrt X_1

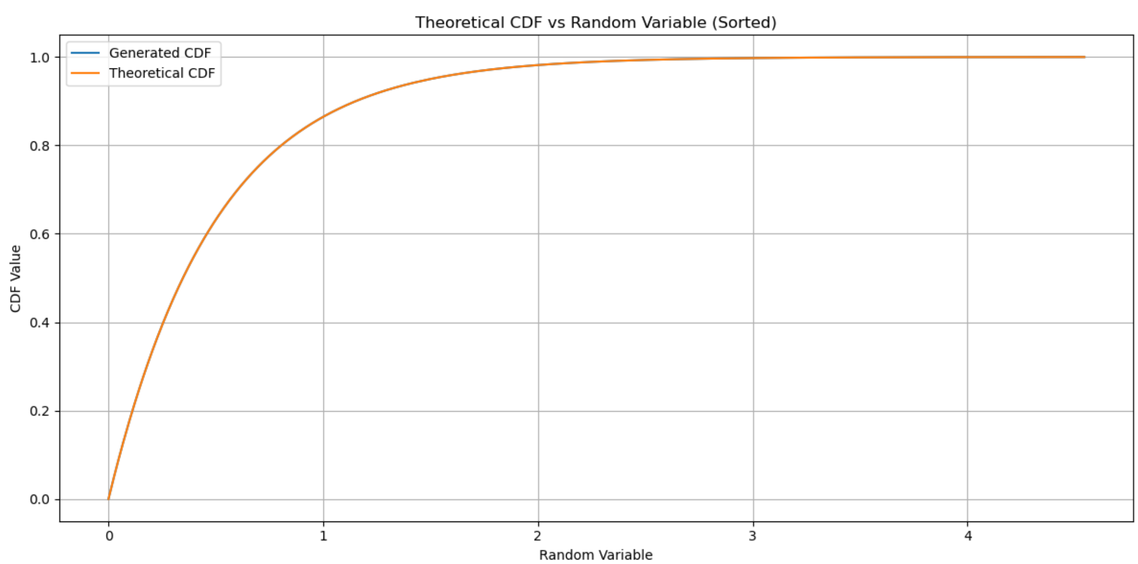


Figure 2.11: Theoretical vs Simulation Analysis wrt X_2

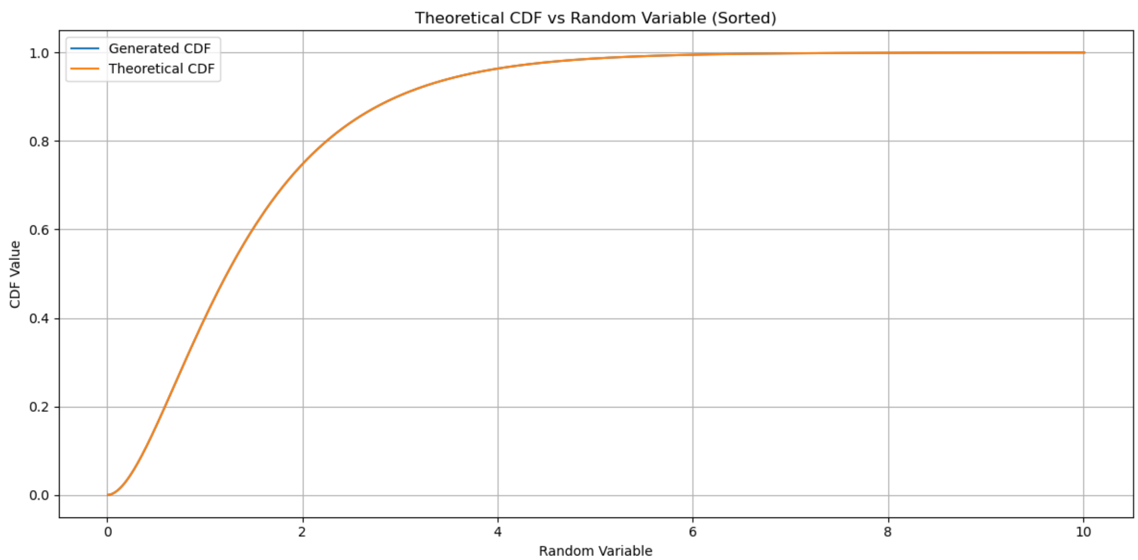


Figure 2.12: Theoretical vs Simulation Analysis wrt X

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

$$P(X = 0) = \frac{1}{2}\left(1 + e^{-1}\right) \tag{2.301}$$

$$P(X = k) = \frac{e^{-1}}{2k!} \ for\ k = 1, 2, 3, \dots . \tag{2.302}$$

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,

$$E(X) = \sum kp_X(k) \tag{2.303}$$

Therefore,

$$\begin{aligned}
E(X) &= 0 \cdot \frac{1}{2}\left(1 + e^{-1}\right) + \sum_{k=1}^{\infty} \frac{ke^{-1}}{2k!} && (2.304)\\
&= \sum_{k=1}^{\infty} \frac{e^{-1}}{2(k-1)!} && (2.305)\\
&= \frac{1}{2e} \sum_{k=1}^{\infty} \frac{1}{(k-1)!} && (2.306)\\
&= \frac{1}{2e} \cdot e && \text{(Using standard result of exponential series)}\\
&= \frac{1}{2} && (2.307)
\end{aligned}$$

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

$$\begin{aligned}
\Pr(X|X > 0) &= \frac{\Pr(X = k)}{1 - \Pr(X = 0)} && (2.308)\\
&= \frac{e^{-1}}{2k!} \cdot \frac{1}{(1 - \frac{1}{2}(1 + e^{-1}))} && (2.309)\\
&= \frac{e^{-1}}{2k!} \cdot \frac{2}{(1 - e^{-1})} && (2.310)\\
&= \frac{e^{-1}}{k!(1 - e^{-1})} && (2.311)\\
&= \frac{1}{k!(e - 1)} && (2.312)
\end{aligned}$$

Therefore,

$$
\begin{aligned}
E(X|X>0) &= \sum_{k=1}^{\infty} k \frac{1}{k!(e-1)} && (2.313)\\
&= \frac{1}{e-1} \sum_{k=1}^{\infty} \frac{1}{(k-1)!} && (2.314)\\
&= \frac{1}{e-1} \cdot e \qquad \text{(Using standard result of exponential series)}\\
&= 1.582 && (2.315)
\end{aligned}
$$

Refering to equations (**??**) and (**??**), 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:

$$
\begin{aligned}
P(X=0) &= \frac{1}{2}\left(1+e^{-1}\right) && (2.316)\\
P(X=k) &= \frac{e^{-1}}{2k!} \; for \; k = 1, 2, 3, \dots . && (2.317)
\end{aligned}
$$

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:

$$X = {F_X}^{-1}(U) \tag{2.318}$$

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

$$\sum_{j=0}^{k-1} p_X(j) \leq U < \sum_{j=0}^{k} p_X(j) \tag{2.319}$$

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:

$$E(X) = \sum k p_X(k) \tag{2.320}$$

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:

$$\Pr(X|X > 0) = \frac{\Pr(X = k)}{1 - \Pr(X = 0)} \tag{2.321}$$

$$\tag{2.322}$$

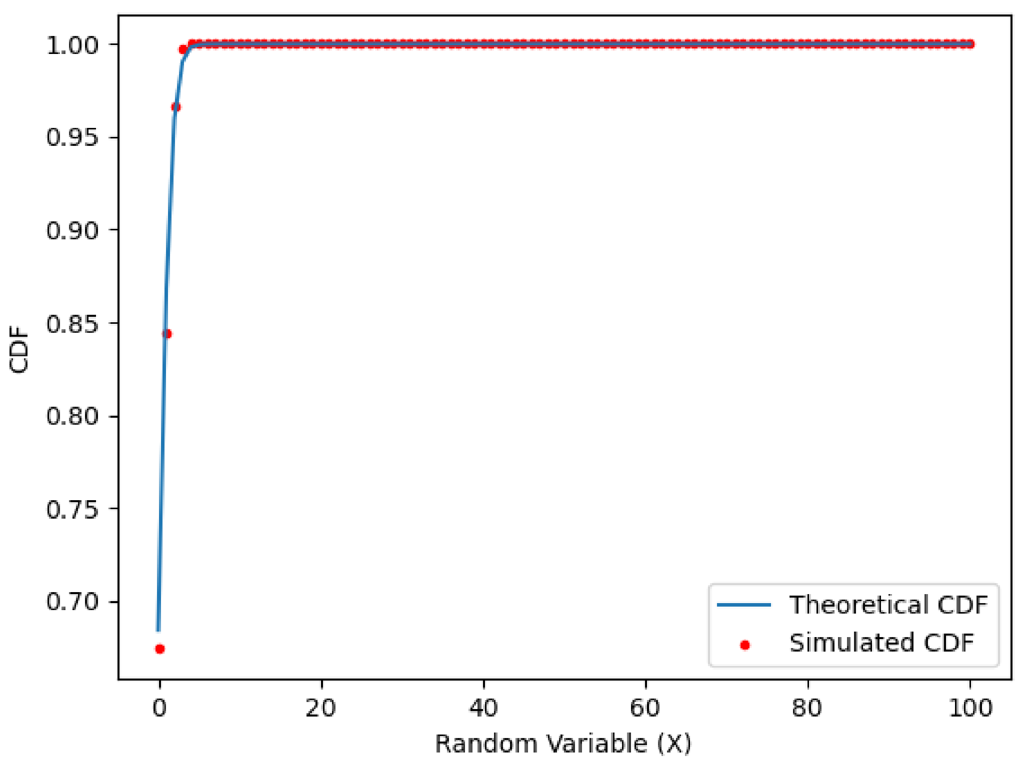


Figure 2.13: plot of corcumcirclr O and points A, B and C.

Chapter 3

Conditional Probability

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

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

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

Solution: Given, the probability density function of X is

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

Also, $Y = \log_e X$.

Consider the cumulative distribution function(CDF) of X,

$$F_X(x) = \Pr(X \leq x) \tag{3.3}$$

$$= \int_1^x \frac{1}{x^2} dx \tag{3.4}$$

$$= 1 - \frac{1}{x}, x \geq 1 \tag{3.5}$$

Now, we need to find the CDF of Y.

$$F_Y(y) = \Pr(Y \leq y) \tag{3.6}$$
$$= \Pr(\log_e X \leq y) \tag{3.7}$$
$$= \Pr(X \leq e^y) \tag{3.8}$$
$$= F_X(e^y) \tag{3.9}$$
$$= 1 - \frac{1}{e^y}, y \geq 0 \tag{3.10}$$

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

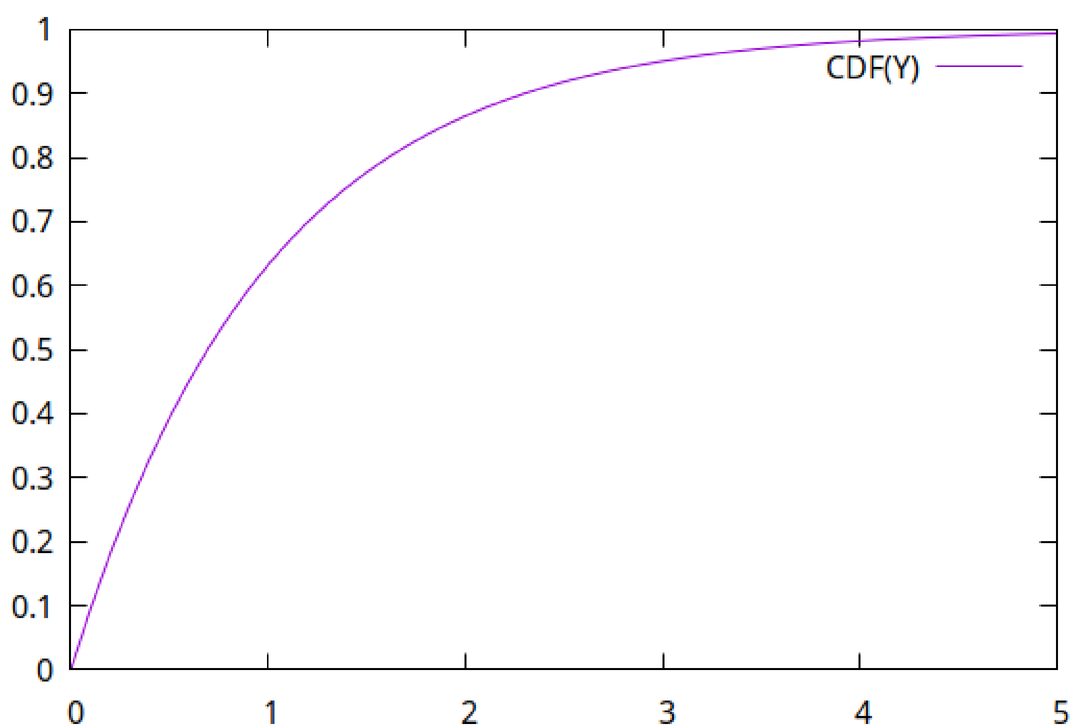


Figure 3.1: CDF of Y

Using the equation for CDF,

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

and

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

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

$$\begin{aligned}
\Pr(Y < 1 | Y < 2) &= \frac{\Pr(Y < 1, Y < 2)}{\Pr(Y < 2)} && (3.13)\\
&= \frac{\Pr(Y < 1)}{\Pr(Y < 2)} && (3.14)\\
&= \frac{F_Y(1)}{F_Y(2)} && (3.15)\\
&= \frac{1 - \frac{1}{e}}{1 - \frac{1}{e^2}} && (3.16)\\
&= \frac{e(e-1)}{e^2 - 1} && (3.17)\\
&= \frac{e}{e+1} && (3.18)
\end{aligned}$$

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,

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

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,

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

The inverse of CDF of X is given by,

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

(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 ith 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

$$Z = X_1 - X_2 \tag{3.22}$$

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

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

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, $\Pr(Y = 1/Z = 0)$.

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

Since X_i and Y are independent events

$$\Pr(Y = 1/Z = 0) = p_Y(1) \tag{3.25}$$

$$= \frac{2}{3} \tag{3.26}$$

$$\approx 0.667 \tag{3.27}$$

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 $F1$ 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:

$$F_1 = \{10, 10, 10, 10\} \tag{4.1}$$

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

$$F_2 = \{10, 10, 00, 00\} \tag{4.2}$$

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

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

$\therefore$ 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

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

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

Now,

$$\Pr(X_1 = 1, X_2 = 1) = \Pr(X_1 = 1)\Pr(X_2 = 1) \tag{4.6}$$
$$= \Pr(1)\Pr(1) \tag{4.7}$$
$$= \frac{1}{6} \cdot \frac{1}{6} \tag{4.8}$$
$$= \frac{1}{36} \tag{4.9}$$
$$= 0.028 \tag{4.10}$$

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

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

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)

$$
\begin{aligned}
E(X) &= \int_{-\infty}^{\infty} x p_X(x) dx && (5.2)\\
&= \int_0^{\infty} x \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} && (5.3)\\
&= \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^{\infty} x^\alpha e^{-\lambda x} && (5.4)\\
& && (5.5)
\end{aligned}
$$

since we know that

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

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

Using the relation

$$\Gamma(x+1) = \Gamma(x)x \tag{5.8}$$

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

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

(b)

$$Var(X) = E(X^2) - E(X)^2 \tag{5.10}$$

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

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

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

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

using the density of the gamma distribution, we get

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

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

$$= \frac{\alpha}{\lambda^2} \tag{5.17}$$

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

(c)

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

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

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

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

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

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

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

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

$$\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| = L \tag{5.22}$$

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

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

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

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 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

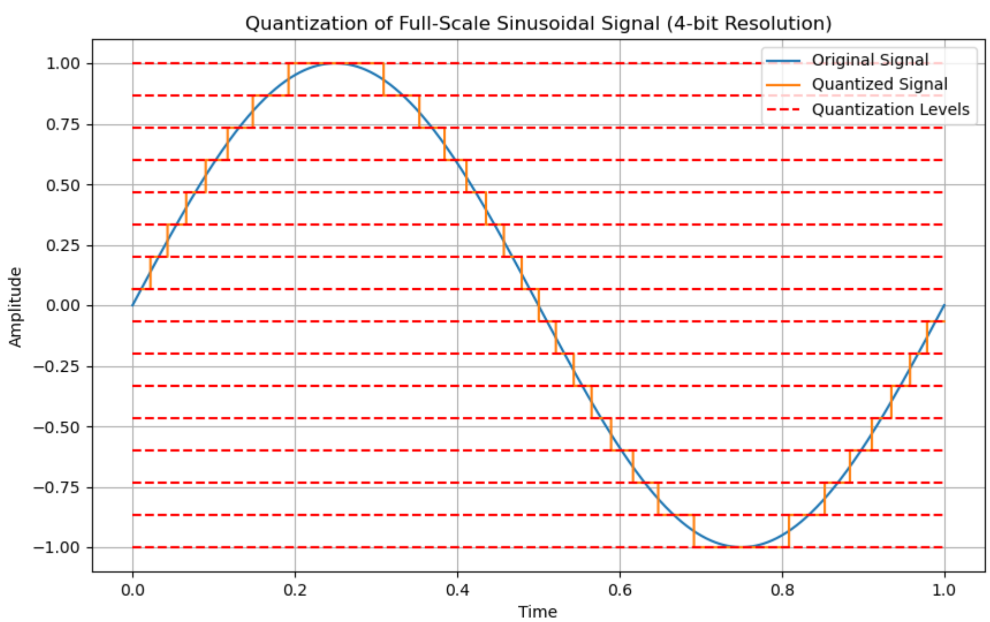


Figure 5.1: Quantization of Sinusoidal Signal

(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:

$$P_S = \lim_{T\to\infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} |x(t)|^2 dt \tag{5.26}$$

where ϕ is the phase of the signal.

$$P_S = \lim_{T\to\infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} |A\sin(2\pi ft + \phi)|^2 dt \tag{5.27}$$

$$P_S = \lim_{T\to\infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} A^2 \sin^2(2\pi ft + \phi) dt \tag{5.28}$$

$$P_S = \lim_{T\to\infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} A^2 \cdot \frac{1 - \cos(4\pi ft + 2\phi)}{2} dt \tag{5.29}$$

$$P_S = \frac{1}{2} \lim_{T\to\infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} A^2 dt - \frac{1}{2} \lim_{T\to\infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} A^2 \cos(4\pi ft + 2\phi) dt \tag{5.30}$$

$$P_S = \frac{1}{2} \cdot A^2 - 0 \tag{5.31}$$

$$P_S = \frac{A^2}{2} \tag{5.32}$$

$$\tag{5.33}$$

Here, A= 1, so:

$$P_S = \frac{1}{2} \tag{5.34}$$

(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

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

$$q = \frac{\text{Peak distance}}{2^n} = \frac{1-(-1)}{2^n} = \frac{2}{2^n} = 2^{-(n-1)} \tag{5.35}$$

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

$$|Y| \leq \frac{q}{2} \text{ and therefore, } |Y| \leq 2^{-n} \tag{5.36}$$

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

$$p_Y(y) = \begin{cases} \frac{1}{q}, & \text{if } -\frac{q}{2} \leq y \leq \frac{q}{2} \\ 0, & \text{otherwise} \end{cases} \tag{5.37}$$

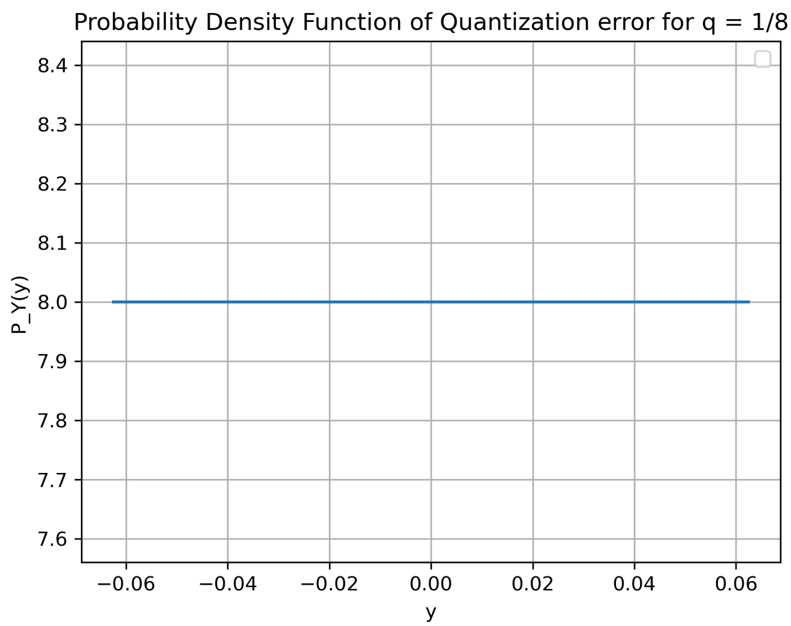


Figure 5.2: plot of pdf of Quantization Error

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.

$$E\left(Y\right) = \int_{-\frac{q}{2}}^{\frac{q}{2}} p_Y(y) y \, de \tag{5.38}$$

$$= \frac{1}{q} \cdot \frac{1}{2} \left(\left(\frac{q}{2}\right)^2 - \left(-\frac{q}{2}\right)^2 \right) \tag{5.39}$$

$$= \frac{1}{q} \cdot \frac{1}{2} \left(0 - 0\right) \tag{5.40}$$

$$= 0 \tag{5.41}$$

$$E[Y^2] = \int_{-\frac{q}{2}}^{\frac{q}{2}} p_Y(y) y^2 \, de \tag{5.42}$$
$$= \frac{1}{q} \cdot \frac{1}{3} \left(\left(\frac{q}{2}\right)^3 - \left(-\frac{q}{2}\right)^3 \right) \tag{5.43}$$
$$= \frac{1}{3q} \cdot \frac{q^3}{4} \tag{5.44}$$
$$= \frac{q^2}{12} \tag{5.45}$$
$$\tag{5.46}$$

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

$$E[Y^2] \approx \frac{2^{-2n}}{3} \tag{5.47}$$
$$P_Y \approx \frac{2^{-2n}}{3} \tag{5.48}$$
$$P_N \approx \frac{2^{-2n}}{3} \tag{5.49}$$
$$\tag{5.50}$$

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

$$SNR = \frac{P_S}{P_N} = 1.5 \cdot 2^{2n} \tag{5.51}$$

or, expressed in decibels,

$$SNR = 10 \left(\log_{10}(1.5 \cdot 2^n)\right) \tag{5.52}$$
$$= 10 \left(\log_{10}(1.5) + \log_{10}(2^{2n})\right) \tag{5.53}$$
$$= 10 \left(0.176 + 2n \cdot 0.3010\right) \tag{5.54}$$
$$= 1.76 + 6.02n \tag{5.55}$$

Substituting value of SNR

$$
\begin{aligned}
61.96 &= 1.76 + 6.02n &(5.56)\\
n &= \frac{61.96 - 1.76}{6.02} &(5.57)\\
&= \frac{60.2}{6.02} &(5.58)\\
&= 10 &(5.59)
\end{aligned}
$$

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_dB.

Chapter 6

Random Algebra

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

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

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:

$$E(X|Y) = \int_{-\infty}^{\infty} x p_{X|Y} dx \tag{6.2}$$

where

$$p_{X|Y} = \frac{p_{XY}(x, y)}{p_Y(y)} \tag{6.3}$$

$$p_Y(y) = \int_0^y p_{X|Y}(x, y)\, dx \tag{6.4}$$

for $0 < y < 1$

$$= \int_0^y 8xydx \tag{6.5}$$

$$= 8y \left[\frac{x^2}{2}\right]_0^y \tag{6.6}$$

$$= 4y^3 \tag{6.7}$$

For $0 < x < y < 1$,on substituting $p_Y(y)$ we get

$$p_{X|Y} = \frac{8xy}{4y^3} \tag{6.8}$$

$$= \frac{2x}{y^2} \tag{6.9}$$

and

$$E(X|Y = y_0) = \int_0^{y_0} x.\frac{2x}{y_0^2} dx \tag{6.10}$$

$$= \frac{2}{y_0^2}\left[\frac{x^3}{3}\right]_0^{y_0} \tag{6.11}$$

$$= \frac{2y_0}{3} \tag{6.12}$$

$$\implies \frac{2y_0}{3} = \frac{1}{2} \tag{6.13}$$

$$y_0 = \frac{3}{4} \tag{6.14}$$

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

$$\frac{L(H_1)}{L(H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} k \tag{7.2}$$

$$\implies \frac{\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}}{\frac{1}{2}e^{-2|x|}} \underset{H_0}{\overset{H_1}{\gtrless}} k \tag{7.3}$$

$$\implies e^{\frac{x^2-2|x|}{2}} \underset{H_0}{\overset{H_1}{\gtrless}} k\frac{\sqrt{\pi}}{\sqrt{2}} \tag{7.4}$$

$$(|x|-1)^2 \underset{H_0}{\overset{H_1}{\gtrless}} 2\log\left(\frac{k\sqrt{\pi}}{\sqrt{2}}\right)+1 \tag{7.5}$$

Taking square root on both sides,

$$||x|-1| \underset{H_0}{\overset{H_1}{\gtrless}} \sqrt{2\log\left(\frac{k\sqrt{\pi}}{\sqrt{2}}\right)+1} \tag{7.6}$$

$$\implies |x| \underset{H_0}{\overset{H_1}{\gtrless}} 1+\sqrt{2\log\left(\frac{k\sqrt{\pi}}{\sqrt{2}}\right)+1} \tag{7.7}$$

Hence, the correct answer is (7.1c)

7.2 Suppose that $X_1, X_2, \ldots, X_n$ are independent and identically distributed random variables, each having probability density function $f(\cdot)$ and median θ. 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 θ.

(d) The power function does not depend on θ.

(GATE ST 2023)

Solution:

Definition 7.1: Median θ is defined as

$$\Pr(X_i \leq \theta) = 0.5 \text{ for all i from 1 to n.}$$

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.

$$I(X_i > \theta_0) = \begin{cases} 1, & \text{if } X_i > \theta_0 \\ 0, & \text{if } X_i \leq \theta_0 \end{cases} \tag{7.8}$$

$$E(S) = E\left(\sum_{i=1}^{n} I(X_i > \theta_0)\right) \tag{7.9}$$

$$= \sum_{i=1}^{n} E(I(X_i > \theta_0)) \tag{7.10}$$

Since,

$$E(I(X_i > \theta_0)) = P(X_i > \theta_0) = \int_{\theta_0}^{\infty} f(x)\, dx \tag{7.11}$$

Therefore,

$$E(S) = \sum_{i=1}^{n} \int_{\theta_0}^{\infty} f(x)\, dx \tag{7.12}$$

(a) From (6.12), under H_0, the distribution of S depends on $f(\cdot)$.

(b) The power function can be expressed as:

$$\pi(\theta) = \Pr(\text{Reject } H_0 \mid H_1 \text{ is true}) \tag{7.13}$$

$$= \Pr(S > c|\theta) \tag{7.14}$$

Therefore, power function depends on value of θ.

7.3 Let X_1, X_2, X_3....,X_n be a random sample of size $n\,(\geq 2)$ from a population having probability density function

$$f\left(x;\theta\right) = \begin{cases} \frac{2}{\theta x}\left(\log_e x\right) e^{-\frac{\left(\log_e x\right)^2}{\theta}} & ,0 < x < 1 \\ 0 & ,otherwise \end{cases}$$

where $\theta > 0$ is an unknown parameter. Then which of the following statements is true,

(A) $\frac{1}{n}\sum_{i=1}^{n}\left(\ln X_i\right)^2$ is the maximum likelihood estimator of θ

(B) $\frac{1}{n-1}\sum_{i=1}^{n}\left(\ln X_i\right)^2$ is the maximum likelihood estimator of θ

(C) $\frac{1}{n}\sum_{i=1}^{n}\ln X_i$ is the maximum likelihood estimator of θ

(D) $\frac{1}{n-1}\sum_{i=1}^{n}\ln X_i$ is the maximum likelihood estimator of θ

(GATE ST 2023)

Solution:

$$L\left(\theta\right) = f\left(x_1, x_2, .., x_n;\theta\right) \tag{7.15}$$

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 θ. Therefore,

$$L(\theta) = \prod_{i=1}^{n} f(x_i; \theta) \tag{7.16}$$

Maximizing $L(\theta)$ is equivalent to maximizing the the $\ln L(\theta)$ as ln is a monotonically increasing function.

$$\begin{aligned}
l(\theta) &= \ln L(\theta) && (7.17)\\
&= \ln\left(\prod_{i=1}^{n} f(x_i; \theta)\right) && (7.18)\\
&= \sum_{i=1}^{n} \ln f(x_i; \theta) && (7.19)\\
&= -n\ln 2 - n\ln\theta + \sum_{i=1}^{n} \ln(-\ln x_i) - \sum_{i=1}^{n} (\ln x_i) - \sum_{i=1}^{n} \frac{(\ln x_i)^2}{\theta} && (7.20)
\end{aligned}$$

Maximizing $l(\theta)$ with respect to θ gives the MLE estimation, therefore

$$\frac{\partial l(\theta)}{\partial\theta} = 0 \tag{7.21}$$

$$\frac{-n}{\theta} + \frac{1}{(\theta)^2} \sum_{i=1}^{n} (\ln X_i)^2 = 0 \tag{7.22}$$

$$\theta = \frac{1}{n} \sum_{i=1}^{n} (\ln X_i)^2 \tag{7.23}$$

Hence (A) is the true statement.

7.4 Suppose that (X, Y) has joint probability mass function

$$P(X = 0, Y = 0) = P(X = 1, Y = 1) = \theta, \tag{7.24}$$
$$P(X = 1, Y = 0) = P(X = 0, Y = 1) = \frac{1}{2} - \theta. \tag{7.25}$$

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 $(X_1, Y_1), (X_2, Y_2), \ldots (X_n, Y_n)$ 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,

$$H_0 : \quad \theta = \theta_0 = \frac{1}{4}, \tag{7.26}$$
$$H_1 : \quad \theta = \theta_1 = \frac{1}{3}. \tag{7.27}$$

and the pmf is given by

$$p_{XY}(0, 0) = p_{XY}(1, 1) = \theta \tag{7.28}$$
$$p_{XY}(0, 1) = p_{XY}(1, 0) = \frac{1}{2} - \theta \tag{7.29}$$

Then for the given random sample of data,

$$p_{X_i,Y_i}(x,y) = \begin{cases} 2\theta & x = y \\ \\ 1 - 2\theta & x \neq y \end{cases} \tag{7.30}$$

$$\tag{7.31}$$

Then the likelihood of the data under H_0 is given by:

$$L(\theta_0 \mid data) = \prod_{i=1}^{n} p_{X_i,Y_i}(x,y) \tag{7.32}$$

$$= (2\theta_0)^m (1 - 2\theta_0)^{n-m} \tag{7.33}$$

$$= \left(\frac{1}{2}\right)^m \left(\frac{1}{2}\right)^{n-m} \tag{7.34}$$

Then the likelihood of the data under H_1 is given by:

$$L(\theta_1 \mid data) = \prod_{i=1}^{n} p_{X_i,Y_i}(x,y) \tag{7.35}$$

$$= (2\theta_1)^m (1 - 2\theta_1)^{n-m} \tag{7.36}$$

$$= \left(\frac{2}{3}\right)^m \left(\frac{1}{3}\right)^{n-m} \tag{7.37}$$

The likelyhood ratio will be

$$\lambda(data) = \frac{L(\theta_1 \mid x)}{L(\theta_0 \mid x)} \tag{7.38}$$

$$= \frac{\left(\frac{2}{3}\right)^m \left(\frac{1}{3}\right)^{n-m}}{\left(\frac{1}{2}\right)^m \left(\frac{1}{2}\right)^{n-m}} = (2)^m \left(\frac{2}{3}\right)^n \tag{7.39}$$

Let the critical value be denoted by c_1, then the likelihood ratio test rejects H_0 if

$$\implies \lambda(data) \underset{H_0}{\overset{H_1}{\gtrless}} c_1 \tag{7.40}$$

$$\tag{7.41}$$

From (7.39),

$$\implies (2)^m \left(\frac{2}{3}\right)^n \underset{H_0}{\overset{H_1}{\gtrless}} c_1 \tag{7.42}$$

$$\implies (2)^m \underset{H_0}{\overset{H_1}{\gtrless}} c_1 \left(\frac{2}{3}\right)^n \tag{7.43}$$

$$\implies m \underset{H_0}{\overset{H_1}{\gtrless}} \log_2 \left(c_1 \left(\frac{2}{3}\right)\right)^n \tag{7.44}$$

$$\implies m \underset{H_0}{\overset{H_1}{\gtrless}} c \quad \exists c \in \mathbb{R} \tag{7.45}$$

where,

$$c = \log_2 \left(c_1 \left(\frac{2}{3}\right)\right)^n \tag{7.46}$$

$\therefore$ 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

$$F_X(x) = \begin{cases} 0 & \text{if } x \leq 0 \\ 1 - (1-x)^\theta & \text{if } 0 \leq x < 1 \\ 1 & \text{if } x \geq 1, \end{cases} \tag{7.47}$$

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 α and β 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,

$$H_0 : \theta = \theta_0 = 1 \tag{7.48}$$

$$H_1 : \theta = \theta_1 = 2 \tag{7.49}$$

PDF can be defined as:

$$p_X(x) = \frac{d}{dx} F_X(x) \tag{7.50}$$

$$= \begin{cases} \theta (1-x)^{\theta-1} & \text{if } 0 \le x < 1 \\ 0 & \text{otherwise} \end{cases} \tag{7.51}$$

Level of test:

$$\alpha = \Pr(\text{reject } H_0 | \, H_0 \text{ is true}) \tag{7.52}$$

$$= \Pr(x < 0.5 | \, \theta_0) \tag{7.53}$$

$$= F_X(0.5) \tag{7.54}$$

$$= 1 - (1 - 0.5) \tag{7.55}$$

$$= \frac{1}{2} \tag{7.56}$$

Power of test:

$$\beta = \Pr\left(\text{reject } H_0 \middle| H_1 \text{ is true}\right) \tag{7.57}$$

$$= \Pr\left(x < 0.5 \middle| \theta_1\right) \tag{7.58}$$

$$= F_X\left(0.5\right) \tag{7.59}$$

$$= 1 - \left(1 - 0.5\right)^2 \tag{7.60}$$

$$= \frac{3}{4} \tag{7.61}$$

Now,

$$\alpha + \beta = \frac{1}{2} + \frac{3}{4} \tag{7.62}$$

$$= 1.25 \tag{7.63}$$

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

$$\hat{S}_i = \hat{\beta}_0 + \hat{\beta}_1 \text{CGPA}_i, i = 1, 2, \ldots, 100$$

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.

$$t_{\hat{\beta}_1} = \frac{\hat{\beta}_1 - \beta_1}{SE\left(\hat{\beta}_1\right)} \tag{7.64}$$

where,

- $\hat{\beta}_1$ is the point estimate.
- β_1 is the hypothesized value.
- $SE\left(\hat{\beta}_1\right)$ 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.

$$SE(\hat{\beta}_1) = \sqrt{\frac{s^2}{n-2}} \tag{7.65}$$

where,

- s^2 is the variance
- n is the sample size

Given that $\hat{\beta}_1 = 645.08$ and $SE\left(\hat{\beta}_1\right) = 70.01$, we get

$$t_{\hat{\beta}_1} = \frac{645.08 - 0}{70.01} \tag{7.66}$$

$$t_{\hat{\beta}_1} = 9.21 \tag{7.67}$$

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 $F(.)$. Consider testing

$$H_0 : F = F_0 \quad \text{against} \quad H_1 : F \neq F_0 \tag{7.68}$$

where,

$$F_0(x) = \begin{cases} 0 & x < 0 \\ x & 0 \leq x < 1 \\ 1 & x \geq 1 \end{cases} \tag{7.69}$$

Let D denote the Kolmogorov-Smirnov test statistic. If $P(D > 0.669) = 0.01$ under H_0 and

$$\psi = \begin{cases} 1 & \text{if } H_0 \text{ is accepted at level } 0.01 \\ 0 & \text{otherwise} \end{cases} \tag{7.70}$$

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

$$n = 4 \tag{7.71}$$

The cdf of the random sample is given as

$$F_X(x) = \begin{cases} 0 & x < 0 \\ x & 0 \le x < 1 \\ 1 & x \ge 1 \end{cases} \tag{7.72}$$

The empirical distribution function(edf) G_n for n independent and identically distributed (i.i.d.) ordered observations X_i is defined as

$$G_n(x) = \frac{\text{no of (elements in the sample} \le x)}{n} = \frac{1}{n}\sum_{i=1}^{n} 1(X_i \le x) \tag{7.73}$$

where $1(A)$ is the indicator of event A and in (7.73) it is defined as,

$$1(X_i \le x) = \begin{cases} 1 & X_i \le x \\ 0 & \text{otherwise} \end{cases} \tag{7.74}$$

From (7.71), (7.72) and (7.73), the edf for the given data will be

$$G_n(0.13) = \frac{1}{4}\sum_{i=1}^{n} 1(X_i \le 0.13) = \frac{1}{2} \tag{7.75}$$

$$G_n(0.12) = \frac{1}{4}\sum_{i=1}^{n} 1(X_i \le 0.12) = \frac{1}{4} \tag{7.76}$$

$$G_n(0.78) = \frac{1}{4}\sum_{i=1}^{n} 1(X_i \le 0.78) = 1 \tag{7.77}$$

$$G_n(0.51) = \frac{1}{4}\sum_{i=1}^{n} 1(X_i \le 0.51) = \frac{3}{4} \tag{7.78}$$

The Kolmogorov–Smirnov statistic for a given cdf $F_X(x)$ is

$$D_n = \sup |G_n(x) - F_X(x)| \tag{7.79}$$

The difference between cdf and edf for the given data will be (i.e., $\forall x \in \{0.13, 0.12, 0.78, 0.51\}$)

$$G_n(0.13) - F_X(0.13) = 0.37 \tag{7.80}$$

$$G_n(0.12) - F_X(0.12) = 0.25 \tag{7.81}$$

$$G_n(0.78) - F_X(0.78) = 0.22 \tag{7.82}$$

$$G_n(0.51) - F_X(0.51) = 0.24 \tag{7.83}$$

Then

$$D_n = \sup(0.37, 0.25, 0.22, 0.24) = 0.37 \tag{7.84}$$

Given that,

$$P(D > 0.669) = 0.01 \tag{7.85}$$

Then

$$H_0 = \begin{cases} \text{accepted at level } 0.01 & \text{if } D_n \leq 0.669 \\ \text{rejected at level } 0.01 & \text{if } D_n > 0.669 \end{cases} \tag{7.86}$$

From (7.84) and (7.86); We can say that H_0 is accepted at level 0.01 and

$$\psi = 1 \tag{7.87}$$

$\therefore$ the value will be

$$\psi + D_n = 1 + 0.37 = 1.37 \tag{7.88}$$

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:

$$S_Y(f) = |H(f)|^2 S_X(f) \tag{9.1}$$

where $H(f)$ is frequency response of the system.

$H(f)$ can be found by taking fourier transform of $h(t)$

$$H(f) \overset{\mathcal{F}}{\rightleftharpoons} \frac{1}{j2\pi f + 1} \tag{9.2}$$

The average power of a signal with power spectral density $S(f)$ is given by:

$$P_Y(f) = \int_{-\infty}^{\infty} S_Y(f)\, df \tag{9.3}$$

Substituting $S_Y(f)$ in the equation we get:

$$\begin{aligned}
P_Y(f) &= \int_{-\infty}^{\infty} |H(f)|^2 \cdot S_X(f)\, df && (9.4)\\
&= \int_{-\infty}^{\infty} \left|\frac{1}{j2\pi f + 1}\right|^2 \cdot \frac{1}{2}\, df && (9.5)\\
&= \frac{1}{2}\int_{-\infty}^{\infty} \frac{1}{(2\pi f)^2 + 1}\, df && (9.6)\\
&= \frac{1}{2} \times 2\int_{0}^{\infty} \frac{1}{(2\pi f)^2 + (1)^2}\, df && (9.7)\\
&= \int_{0}^{\infty} \frac{1}{(2\pi f)^2 + (1)^2}\, df && (9.8)\\
&= \frac{1}{2\pi} \tan^{-1}(x)\,|_0^{\infty} && (9.9)\\
&= \frac{1}{2\pi}\left(\tan^{-1}\infty - \tan^{-1} 0\right) && (9.10)\\
&= \frac{1}{2\pi}\left(\frac{\pi}{2} - 0\right) && (9.11)\\
&= \frac{1}{2\pi}\left(\frac{\pi}{2}\right) && (9.12)\\
&= \frac{1}{4} && (9.13)
\end{aligned}$$

Rounded off to two decimal places, the average power of the system output is $0.25W$.

Chapter 10

Convergence

10.1 Let $\{X_n\}_{n\geq 1}$ and Let $\{Y_n\}_{n\geq 1}$ 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,

$$lim_{n\to\infty} F_{X_n}(x) = F_X(x) \tag{10.1}$$

(b) X_n converges in probability to X, $X_n \xrightarrow{p} X$, then for all $\epsilon > 0$,

$$lim_{n\to\infty} \Pr\left(|X_n - X| > \epsilon\right) = 0 \tag{10.2}$$

(c) X_n converges in p^{th} mean to X, then we have

$$lim_{n \to \infty} E(|X_n - X|^p) = 0 \tag{10.3}$$

(A) For $\epsilon > 0$, B be defined as

$$B = \{x : |x - c| \geq \epsilon\} \tag{10.4}$$

Now,

$$\Pr(|X_n - c| \geq \epsilon) = \Pr(X_n \in B) \tag{10.5}$$

Using Portmanteau Lemma, if $X_n \xrightarrow{d} c$, we have

$$\begin{aligned} \limsup_{n \to \infty} \Pr(X_n \in B) &\leq \Pr(c \in B) && (10.6)\\ &\leq \Pr(|0 - 0| \geq \epsilon) && (10.7)\\ &\leq \Pr(0 \geq \epsilon) && (10.8)\\ &\leq 0 && (10.9)\\ &= 0 && (10.10)\\ lim_{n \to \infty} \Pr(|X_n - c| > \epsilon) &= 0 && (10.11) \end{aligned}$$

From (10.2), $X_n \xrightarrow{p} c$. So, we have

$$X_n \xrightarrow{d} c \implies X_n \xrightarrow{p} c \tag{10.12}$$

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 $X = 0$

$$lim_{n\to\infty} \Pr(|X_n - X| > \epsilon) = lim_{n\to\infty} \Pr(X_n > \epsilon) \tag{10.13}$$

$X_n > \epsilon$vis subset of $X_n = n$ since every time X_n equals n, it's also true that X_n is greater than ϵ. But there may be times when X_n is greater than ϵ without X_n being equal to n. So,

$$\Pr(X_n > \epsilon) \leq \Pr(X_n = n) \tag{10.14}$$
$$lim_{n\to\infty} \Pr(|X_n - X| > \epsilon) \leq lim_{n\to\infty} \Pr(X_n = n) \tag{10.15}$$
$$\leq lim_{n\to\infty} \frac{1}{n} \tag{10.16}$$
$$\leq 0 \tag{10.17}$$
$$= 0 \tag{10.18}$$

But X_n does not converges in 3^{rd} mean to $X = 0$.

$$lim_{n\to\infty} E(|X_n - X|^3) = lim_{n\to\infty} E(X_n^3) \tag{10.19}$$
$$= lim_{n\to\infty} 0^3 \left(1 - \frac{1}{n}\right) + n^3 \left(\frac{1}{n}\right) \tag{10.20}$$
$$= lim_{n\to\infty} n^2 \neq 0 \tag{10.21}$$

(C) Statement (C) is may or may not correct. Counter Example: Consider distribution

$$Z \sim \mathcal{N}(0,1) \tag{10.22}$$

Let $\{X_n\}_{n\geq1}$ and $\{Y_n\}_{n\geq1}$ 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

$$lim_{n\to\infty} F_{Y_n}(x) = \Pr\left(Z \leq x\right) \tag{10.23}$$

For n odd

$$\begin{aligned} lim_{n\to\infty} F_{Y_n}(x) &= \Pr\left(-Z \leq x\right) && (10.24) \\ &= \Pr\left(Z \leq x\right) && (10.25) \end{aligned}$$

Proved. So,we have

$$\begin{aligned} F_{X_n+Y_n}(x) &= \Pr\left(X_n + Y_n \leq x\right) && (10.26) \\ &= \Pr\left(Z + (-1)^n Z \leq x\right) && (10.27) \end{aligned}$$

For n is even

$$\begin{aligned} F_{X_n+Y_n}(x) &= \Pr\left(2Z \leq x\right) && (10.28) \\ &= \Pr\left(Z \leq \frac{x}{2}\right) && (10.29) \\ &= 1 - \Pr\left(Z > \frac{x}{2}\right) && (10.30) \\ &\approx 1 - Q\left(\frac{x}{2}\right) && (10.31) \end{aligned}$$

For n is odd

$$F_{X_n+Y_n}(x) = \Pr(0 \leq x) \tag{10.32}$$

$$= \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases} = H(x) \tag{10.33}$$

So, on generalizing

$$F_{X_n+Y_n}(x) = \begin{cases} 1 - Q\left(\frac{x}{2}\right) & \text{if } n \text{ is even} \\ H(x) & \text{if } n \text{ is odd} \end{cases} \tag{10.34}$$

$lim_{n\to\infty} F_{X_n+Y_n}(x)$ 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
$\Pr(X_n)$	$1 - \frac{1}{n}$	$\frac{1}{n}$

$$lim_{n\to\infty} E(X_n) = 0\left(1 - \frac{1}{n}\right) + n\left(\frac{1}{n}\right) \tag{10.35}$$

$$= 1 \tag{10.36}$$

As $n \to \infty$, $E(X_n)$ converges to $E(X) = 1$.

$$lim_{n\to\infty} X_n = 0 = X \tag{10.37}$$

To find 1^{st} mean convergennce of X_n. From (10.36)

$$lim_{n\to\infty}E(|X_n - X|) = lim_{n\to\infty}E(X_n) \tag{10.38}$$

$$= 1 \neq 0 \tag{10.39}$$

So, X_n does not converges in 1^{st} 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 $\lim_{n\to\infty} \mathrm{E}\left[\left(\frac{Y_n - 4}{\sqrt{n}}\right)^2\right]$ (in integer) equals __________. (GATE ST 2023)

Solution:

(a) **Theory:** For all X_i which as i.i.d's, mean $\mu = 4$ and variance $\sigma^2 = 9$,

$$Y_n = \frac{1}{n}\sum_{i=1}^{n} X_i \tag{10.40}$$

The mean of a sum of i.i.d random variables is calculated as

$$\mathrm{E}[Y_n] = \mathrm{E}\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] \tag{10.41}$$

$$= \frac{1}{n}\sum_{i=1}^{n} \mathrm{E}[X_i] \tag{10.42}$$

$$= \frac{1}{n}(n\mu) \tag{10.43}$$

$$= \mu \tag{10.44}$$

The variance of a sum of i.i.d random variables is calculated as

$$\operatorname{var}(Y_n) = \mathrm{E}\left[\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)^2\right] - \left(\mathrm{E}\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right]\right)^2 \tag{10.45}$$

$$= \frac{1}{n^2}\left\{\mathrm{E}\left[\left(\sum_{i=1}^{n} X_i\right)^2\right] - \left(\mathrm{E}\left[\sum_{i=1}^{n} X_i\right]\right)^2\right\} \tag{10.46}$$

But

$$\mathrm{E}\left[\left(\sum_{i=1}^{n} X_i\right)^2\right] = \mathrm{E}\left[\sum_{i=1}^{n}\sum_{j=1}^{n} X_i X_j\right] \tag{10.47}$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} \mathrm{E}\left[X_i X_j\right] \tag{10.48}$$

and

$$\left(\mathrm{E}\left[\sum_{i=1}^{n} X_i\right]\right)^2 = \left(\sum_{i=1}^{n} \mathrm{E}\left[X_i\right]\right)^2 \tag{10.49}$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} \mathrm{E}\left[X_i\right]\mathrm{E}\left[X_j\right] \tag{10.50}$$

Putting (10.48) and (10.50) in (10.46), and using the definition of covariance,

$$\operatorname{var}(Y_n) = \frac{1}{n^2}\left\{\sum_{i=1}^{n}\sum_{j=1}^{n}\left(\mathrm{E}\left[X_i X_j\right] - \mathrm{E}\left[X_i\right]\mathrm{E}\left[X_j\right]\right)\right\} \tag{10.51}$$

$$= \frac{1}{n^2}\left\{\sum_{i=1}^{n}\sum_{j=1}^{n} \operatorname{cov}\left(X_i, X_j\right)\right\} \tag{10.52}$$

As all the variables are i.i.d's and are thus uncorrelated,

$$\operatorname{cov}(X_i, X_j) = \begin{cases} 0 & \text{if } i \neq j \\ \operatorname{var}(X_i) & \text{if } i = j \end{cases} \tag{10.53}$$

Putting (10.53) in (10.52),

$$\operatorname{var}(Y_n) = \frac{1}{n^2}\left(\sum_{i=1}^{n} \operatorname{cov}(X_i, X_i)\right) \tag{10.54}$$

$$= \frac{1}{n^2}\left(\sum_{i=1}^{n} \operatorname{var}(X_i)\right) \tag{10.55}$$

$$= \frac{1}{n^2}\left(\sum_{i=1}^{n} \sigma^2\right) \tag{10.56}$$

$$= \frac{\sigma^2}{n} \tag{10.57}$$

Consider the term $\left(\frac{Y_n - \mu}{\sqrt{n}}\right)^2$. Calculating its expectation,

$$\mathrm{E}\left[\left(\frac{Y_n - \mu}{\sqrt{n}}\right)^2\right] = \frac{1}{n}\mathrm{E}\left[(Y_n - \mu)^2\right] \tag{10.58}$$

$$= \frac{1}{n}\operatorname{var}(Y_n) \tag{10.59}$$

$$= \frac{\sigma^2}{n^2} \tag{10.60}$$

Substituting $\sigma^2 = 9$ and $\mu = 4$, we get

$$\lim_{n\to\infty} \mathrm{E}\left[\left(\frac{Y_n - 4}{\sqrt{n}}\right)^2\right] = \lim_{n\to\infty} \frac{9}{n^2} = 0 \tag{10.61}$$

(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:

$$p_{X_{ij}}(x) = \begin{cases} 0.5 & \text{if } x \in \{1, 7\} \\ 0 & \text{otherwise} \end{cases} \tag{10.62}$$

This distribution has the following characteristics:

$$\mu = \mathrm{E}\left[X_{ij}\right] = 0.5 \times 1 + 0.5 \times 7 = 4 \tag{10.63}$$

$$\sigma^2 = \mathrm{E}\left[X_{ij}^2\right] - \left(\mathrm{E}\left[X_{ij}\right]\right)^2 \tag{10.64}$$

$$= \left(0.5 \times 1^2 + 0.5 \times 7^2\right) - 4^2 \tag{10.65}$$

$$= 9 \tag{10.66}$$

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

$$Y_j = \frac{1}{n} \sum_{i=1}^{n} X_{ij} \tag{10.67}$$

Now, the expression $\frac{(Y_j - 4)^2}{n}$ is calculated for all $j \leq m$ and their expectancy is calculated as follows:

$$\mathrm{E}\left[\left(\frac{Y_n - 4}{\sqrt{n}}\right)^2\right] = \frac{1}{m} \sum_{j=1}^{m} \frac{(Y_j - 4)^2}{n} \tag{10.68}$$

To calculate the limit $n \to \infty$, 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.

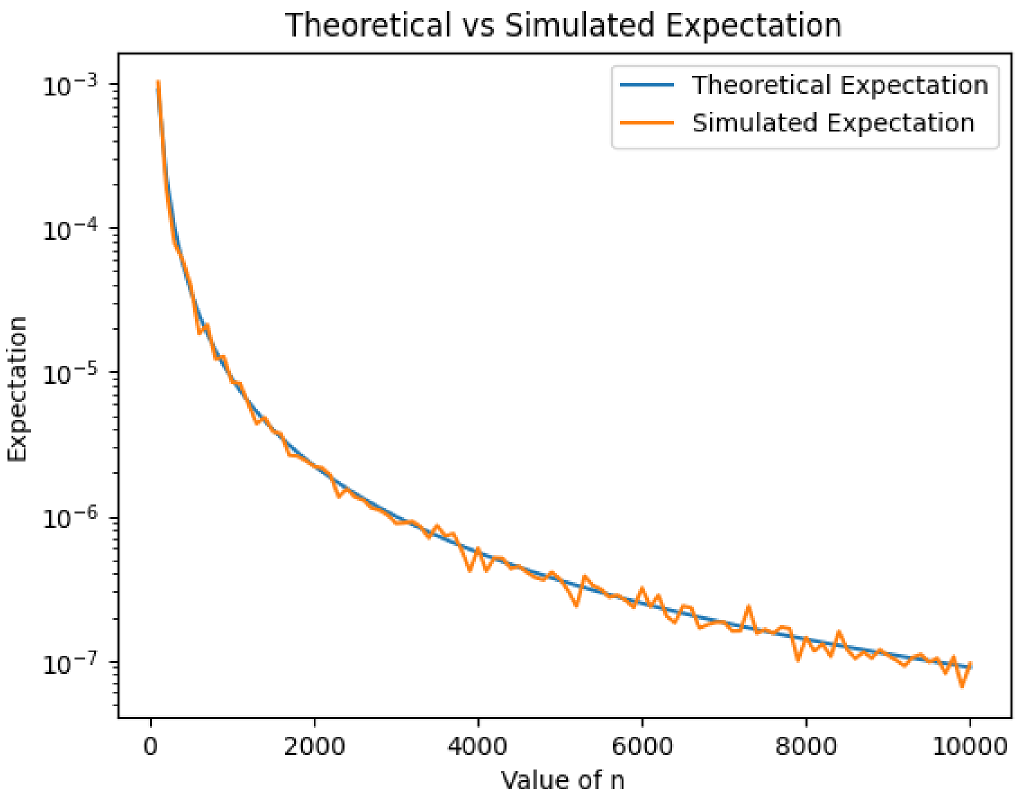


Figure 10.1: Expectation vs n

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,..., let

$$Y_n = \frac{1}{n}(X_1X_2 + X_3X_4 + ... + X_{2n-1}X_{2n}) \tag{10.69}$$

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,

$$\lim_{n\to\infty} F_{X_n}(x) = F_X(x) \tag{10.70}$$

where, $X \sim \mathcal{N}(0,1)$

(b) X_n converges in p^{th} mean to X, then we have

$$\lim_{n\to\infty} E(|X_n - X|^p) = 0 \tag{10.71}$$

(c) X_n converges in probability to X, $X_n \xrightarrow{p} X$, then for all $\epsilon > 0$,

$$\lim_{n\to\infty} \Pr\left(|X_n - X| \geq \epsilon\right) = 0 \tag{10.72}$$

(d) X_n converges almost surely to X, $X_n \xrightarrow{a.s} X$, then for all $\epsilon > 0$,

$$\lim_{n\to\infty} \Pr\left(|X_n - X| < \epsilon\right) = 1 \tag{10.73}$$

$$\mathrm{E}(X) = \int_{-\infty}^{\infty} x F_X x dx \tag{10.74}$$

$$\because X \text{ is positive-valued} \tag{10.75}$$

$$= \int_{0}^{\infty} x F_X x dx \tag{10.76}$$

$$\geq \int_{\epsilon^2}^{\infty} x F_X x dx \tag{10.77}$$

$$\text{for some } \epsilon^2 > 0 \tag{10.78}$$

$$\geq \int_{\epsilon^2}^{\infty} k F_X x dx \tag{10.79}$$

$$= \epsilon^2 \int_{\epsilon^2}^{\infty} F_X x dx \tag{10.80}$$

$$\mathrm{E}(X) = \epsilon^2 \Pr\left(X \geq \epsilon^2\right) \tag{10.81}$$

This inequality is called Markov inequality

Now, we take

$$\sigma^2 = \mathrm{E}(X^2) - [\mathrm{E}(X)]^2 \tag{10.82}$$

$$= \mathrm{E}(X^2) - \mu^2 \tag{10.83}$$

$$= \mathrm{E}(X^2) - \mathrm{E}(\mu^2) \tag{10.84}$$

$$\sigma^2 = \mathrm{E}(X^2 - \mu^2) \tag{10.85}$$

using (**??**),

$$\sigma^2 = \epsilon^2 \Pr\left(X^2 - \mu^2 \geq \epsilon^2\right) \tag{10.86}$$

$$\frac{\sigma^2}{\epsilon^2} = \Pr\left(X^2 - \mu^2 \geq \epsilon^2\right) \tag{10.87}$$

$$\Pr\left(|X - \mu| \geq \epsilon\right) = \frac{\sigma^2}{\epsilon^2} \tag{10.88}$$

This above inequality is called Chebyshev's inequaltiy

Now, as X_i is a sequence of independent and identically distributed random variables,

Let $Z_i = X_{2i-1}X_{2i}$.

$$p_{Z_iZ_j}(x) = p_{X_{2i-1}X_{2i}X_{2j-1}X_{2j}}(x) \tag{10.89}$$

$$= p_{X_{2i-1}}(x)p_{X_{2i}}(x)p_{X_{2j-1}}(x)p_{X_{2j}}(x) \tag{10.90}$$

$$= p_{X_{2i-1}X_{2i}}(x)p_{X_{2j-1}X_{2j}}(x) \tag{10.91}$$

$$= p_{Z_i}(x)p_{Z_j}(x) \tag{10.92}$$

Z_i is a independent R.V . Similarly,

$$F_{X_{2i-1}}(x) = F_{X_{2i}}(x) = F_{X_{2j-1}}(x) = F_{X_{2j}}(x) \tag{10.93}$$

$$F_{X_{2i-1}}(x)F_{X_{2i}}(x) = F_{X_{2j-1}}(x)F_{X_{2j}}(x) \tag{10.94}$$

$$F_{Z_i}(x) = F_{Z_j}(x) \tag{10.95}$$

Thus, Z_i is an identical distributed R.V.

Hence,

$$Z_n = \sum_{i=1}^{n} Z_i \tag{10.96}$$

is an i.i.d

$$\mathrm{E}\left(Z_i\right) = \mathrm{E}\left(X_{2i-1}X_{2i}\right) \tag{10.97}$$

$$= \mathrm{E}\left(X_{2i-1}\right)\mathrm{E}\left(X_{2i}\right) = 0 \tag{10.98}$$

$$\mathrm{Var}(Z_i) = \mathrm{E}(Z_i^2) - \left[\mathrm{E}\left(Z_i\right)\right]^2 \tag{10.99}$$

$$= \mathrm{E}(Z_i^2) - 0 \tag{10.100}$$

$$= \mathrm{E}(X_{2i-1}^2 X_{2i}^2) \tag{10.101}$$

$$= \mathrm{E}(X_{2i-1}^2)\mathrm{E}(X_{2i}^2) = 1 \tag{10.102}$$

Now,

$$Y_n = \frac{1}{n}\left(\sum_{i=1}^{n} Z_i\right) \tag{10.103}$$

$$\mathrm{E}(Y_n) = \mathrm{E}\left(\frac{1}{n}\sum_{i=1}^{n} Z_i\right) \tag{10.104}$$

$$= \frac{1}{n}\sum_{i=1}^{n}\mathrm{E}(Z_i) = 0 \tag{10.105}$$

As all the variables are i.i.d's and are thus uncorrelated,

$$\operatorname{cov}(X_i, X_j) = \begin{cases} 0 & \text{if } i \neq j \\ \\ \operatorname{Var}(X_i) & \text{if } i = j \end{cases} \tag{10.106}$$

$$\operatorname{Var}(Y_n) = \frac{1}{n^2}\left(\sum_{i=1}^{n} \operatorname{cov}(Z_i, Z_i)\right) \tag{10.107}$$

$$= \frac{1}{n^2}\left(\sum_{i=1}^{n} \operatorname{var}(Z_i)\right) \tag{10.108}$$

$$= \frac{1}{n^2}\left(\sum_{i=1}^{n} 1\right) \tag{10.109}$$

$$= \frac{1}{n} \tag{10.110}$$

(A) by central limit theorem,

$$\frac{Y_n - \mu}{\sigma} \sim \mathcal{N}(0,1) \tag{10.111}$$

$$= \frac{Y_n - 0}{\sqrt{\frac{1}{n}}} \sim \mathcal{N}(0,1) \tag{10.112}$$

$$= \sqrt{n}Y_n \sim \mathcal{N}(0,1) \tag{10.113}$$

Thus,

$$\lim_{n\to\infty} F_{\sqrt{n}Y_n} = F_P(x) \tag{10.114}$$

where, $P \sim \mathcal{N}(0,1)$

So, option (A) is correct

(B) For 2nd mean to be converging to 0,

$$\lim_{n\to\infty} E(|Y_n - 0|^2) = \lim_{n\to\infty} E(Y_n^2) \tag{10.115}$$

$$\because \frac{1}{n} = E(Y_n^2) - [E(Y_n)]^2 \tag{10.116}$$

$$= \lim_{n\to\infty} \frac{1}{n} + [E(Y_n)]^2 \tag{10.117}$$

$$\because [E(Y_n)]^2 = 0 \tag{10.118}$$

$$= \lim_{n\to\infty} \frac{1}{n} = 0 \tag{10.119}$$

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 $Y_n + \frac{1}{n} \xrightarrow{p} 0$

$$\lim_{n\to\infty} \Pr\left(\left|Y_n + \frac{1}{n} - Y\right| \geq \epsilon\right) = \lim_{n\to\infty} \Pr\left(\left|Y_n + \frac{1}{n}\right| \geq \epsilon\right) \tag{10.120}$$

$$\leq \lim_{n\to\infty} \frac{\text{Var}(Y_n + \frac{1}{n})}{\epsilon^2} \tag{10.121}$$

$$\forall \epsilon > 0 \tag{10.122}$$

$$= \lim_{n\to\infty} \frac{\text{Var}(Y_n) + \text{Var}(\frac{1}{n})}{\epsilon^2} \tag{10.123}$$

$$= \lim_{n\to\infty} \frac{\frac{1}{n} + 0}{\epsilon^2} \tag{10.124}$$

$$= 0 \tag{10.125}$$

hence, option (C) is correct

(D) As all the variables are i.i.d's and are thus uncorrelated,

$$\operatorname{Var}(X_n) = \sum_{i=1}^{n} \operatorname{cov}(X_i, X_i) \tag{10.126}$$

$$= \sum_{i=1}^{n} \operatorname{var}(X_i) \tag{10.127}$$

$$= \sum_{i=1}^{n} 1 \tag{10.128}$$

$$= n \tag{10.129}$$

$$\lim_{n\to\infty} \Pr(|X_n - X| < \epsilon) = 1 - \lim_{n\to\infty} \Pr(|X_n - 0| \geq \epsilon) \tag{10.130}$$

$$\leq 1 - \lim_{n\to\infty} \frac{Var(X_n)}{\epsilon^2} \tag{10.131}$$

$$\forall \epsilon > 0 \tag{10.132}$$

$$= \lim_{n\to\infty} 1 - \frac{Var(X_n)}{\epsilon^2} \tag{10.133}$$

$$= \lim_{n\to\infty} 1 - \frac{n}{\epsilon^2} = -\infty \tag{10.134}$$

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 intialize 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 $\Pr\left(\left|Y_n + \frac{1}{n} - Y\right| \geq \epsilon\right)$

(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 $\Pr\left(|X_n| < \epsilon\right)$ and if ≥ 1 then it does not converges almost surely.

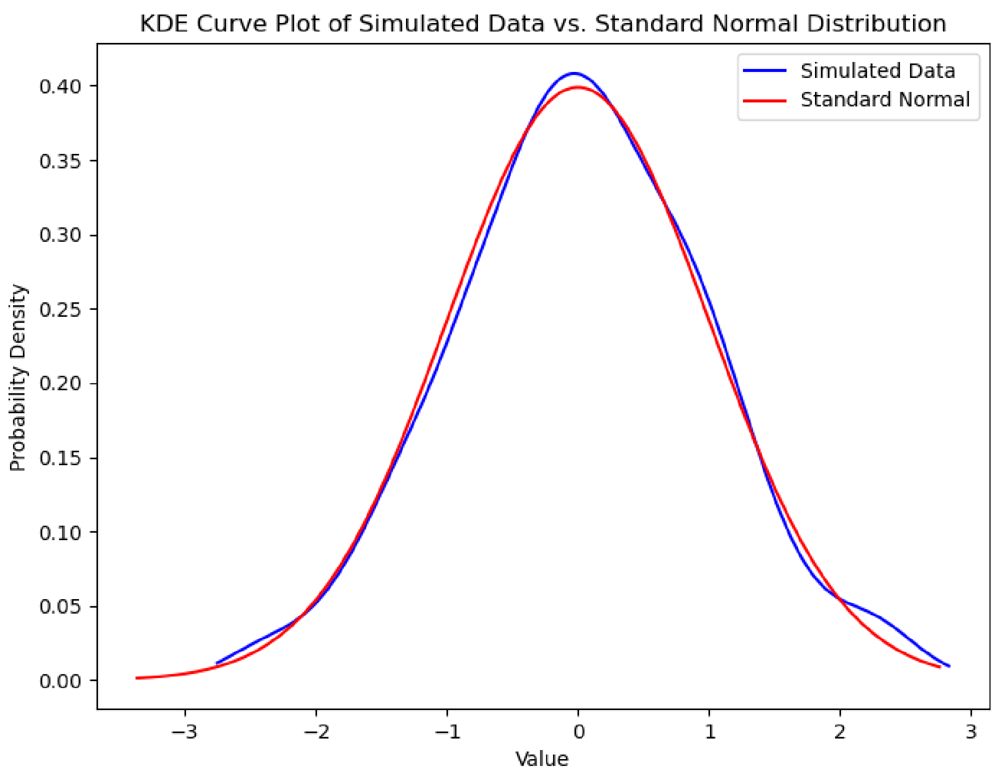


Figure 10.2: simulated vs standard normal

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 $\overline{Y}_n = \frac{1}{n}\sum_{i=1}^{n} Y_i$ for $n \geq 1$ and $\{\sqrt{n}(e^{-\overline{Y}_n} - e^{-1})\}_{n\geq 1}$ converges in distribution to a normal random variable with mean 0 and variance σ^2, then σ^2 (rounded off to two decimal places) equals (GATE ST 2023)

Solution:

(a) Let $X, Y \sim \exp(1)$ and $Z = X - Y$

$$p_X(x) = e^{-x}u(x) \tag{10.135}$$

$$M_X(s) = E\left(e^{-sX}\right) \tag{10.136}$$

$$= \int_0^\infty e^{-sx}e^{-x}\,dx \tag{10.137}$$

$$= \frac{1}{s+1} \tag{10.138}$$

ROC for $M_X(s) : \text{Re}(s) > -1$

Similarly,

$$M_Y(s) = \frac{1}{s+1} \tag{10.139}$$

$$M_Y(-s) = \frac{1}{-s+1} \tag{10.140}$$

ROC for $M_Y(-s) : \text{Re}(s) < 1$

$$M_Z(s) = E\left(e^{-sZ}\right) \tag{10.141}$$

Using,

$$Z = X - Y \tag{10.142}$$

$$\implies M_Z(s) = E\left(e^{-s(X-Y)}\right) \tag{10.143}$$

$$= E\left(e^{-sX}\right) E\left(e^{sY}\right) \tag{10.144}$$

$$= M_X(s) M_Y(-s) \tag{10.145}$$

$$= \frac{1}{s+1} \times \frac{1}{-s+1} \tag{10.146}$$

$$M_Z(s) = \frac{1}{1-s^2} \tag{10.147}$$

The ROC for the laplace transform : $|\text{Re}(s)| < 1$

$$M_Z(s) = \frac{1}{2}\left(\frac{1}{1-s} + \frac{1}{1+s}\right) \tag{10.148}$$

Using Inverse Laplace transform,

$$P_Z(x) = \frac{1}{2}\left(e^{x} u\left(-x\right) + e^{-x} u\left(x\right)\right) \tag{10.149}$$

$$p_Z(x) = \frac{1}{2} e^{-|x|} \tag{10.150}$$

$$\implies Z \sim \text{Lap}\left(0, 1\right) \tag{10.151}$$

(b) Let $T = |Z|$

$$p_Z(x) = \frac{1}{2}e^{-|x|} \tag{10.152}$$

$$F_Z(x) = \int_{-\infty}^{x} \frac{1}{2}e^{-|t|}\,dt \tag{10.153}$$

$$= \frac{1}{2} + \frac{1}{2}e^{-x} \tag{10.154}$$

$$F_T(x) = \Pr\left(T \le x\right) \tag{10.155}$$

$$= \Pr\left(|Z| \le x\right) \tag{10.156}$$

$$= \Pr\left(-x \le Z \le x\right) \tag{10.157}$$

$$F_T(x) = \frac{1}{2} - \frac{1}{2}e^{-x} - \left(-\frac{1}{2} + \frac{1}{2}e^{-x}\right) \tag{10.158}$$

$$F_T(x) = 1 - e^{-x} \text{ for } x > 0 \tag{10.159}$$

$$T \sim \exp\left(1\right) \tag{10.160}$$

$$\implies |Z| \sim \exp\left(1\right) \tag{10.161}$$

Using equations (10.77) and (10.96), we get:

$$|X_{2n} - X_{2n-1}| \sim \exp(1) \tag{10.162}$$

$$\implies Y_n \sim \exp(1) \tag{10.163}$$

$$M_{Y_n}(s) = \frac{1}{1+s} \tag{10.164}$$

$$\mathrm{E}\left(Y_n\right) = \mu_1 \tag{10.165}$$

$$\mu_1 = -\frac{dM_{Y_n}(s)}{ds} \tag{10.166}$$

$$= -\frac{d}{ds}\left(\frac{1}{s+1}\right)\Bigg|_{s=0} \tag{10.167}$$

$$= \frac{1}{(s+1)^2}\bigg|_{s=0} \quad (10.168)$$

$$\mathrm{E}\left(Y_n\right) = 1 \quad (10.169)$$

$$\mathrm{E}\left({Y_n}^2\right) = \mu_2 \quad (10.170)$$

$$\mu_2 = \frac{d^2 Y_n(s)}{ds^2} \quad (10.171)$$

$$= \frac{d^2}{ds^2}\left(\frac{-1}{(s+1)^2}\right)\bigg|_{s=0} \quad (10.172)$$

$$= \frac{2}{(s+1)^3}\bigg|_{s=0} \quad (10.173)$$

$$\mathrm{E}\left({Y_n}^2\right) = 2 \quad (10.174)$$

$$\mathrm{Var}(Y_n) = \mathrm{E}((Y_n - \mathrm{E}(Y_n))^2) \quad (10.175)$$

$$= \mathrm{E}((Y_n - 1)^2) \quad (10.176)$$

$$= \mathrm{E}({Y_n}^2) - 2\mathrm{E}(Y_n) + 1 = 1 \quad (10.177)$$

(c) We know,

$$\overline{Y}_n = \frac{1}{n}\sum_{i=1}^{n} Y_i \quad (10.178)$$

$$\mathrm{E}\left(\overline{Y}_n\right) = \frac{1}{n}\sum_{i=1}^{n} \mathrm{E}\left(Y_i\right) \quad (10.179)$$

$$= \frac{1}{n}\cdot(n) = 1 \quad (10.180)$$

$$\mathrm{E}\left(\overline{Y}_n\right) = 1 \quad (10.181)$$

$$\operatorname{var}\left(\overline{Y}_n\right) = \mathrm{E}\left[\left(\frac{1}{n}\sum_{i=1}^{n} Y_i\right)^2\right] - \left(\mathrm{E}\left[\frac{1}{n}\sum_{i=1}^{n} Y_i\right]\right)^2 \tag{10.182}$$

$$= \frac{1}{n^2}\left\{\mathrm{E}\left[\left(\sum_{i=1}^{n} Y_i\right)^2\right] - \left(\mathrm{E}\left[\sum_{i=1}^{n} Y_i\right]\right)^2\right\} \tag{10.183}$$

But

$$\mathrm{E}\left[\left(\sum_{i=1}^{n} Y_i\right)^2\right] = \mathrm{E}\left[\sum_{i=1}^{n}\sum_{j=1}^{n} Y_i Y_j\right] \tag{10.184}$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} \mathrm{E}\left[Y_i Y_j\right] \tag{10.185}$$

and

$$\left(\mathrm{E}\left[\sum_{i=1}^{n} Y_i\right]\right)^2 = \left(\sum_{i=1}^{n} \mathrm{E}\left[Y_i\right]\right)^2 \tag{10.186}$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} \mathrm{E}\left[Y_i\right]\mathrm{E}\left[Y_j\right] \tag{10.187}$$

Putting (10.120) and (10.122) in (10.118), and using the definition of covariance,

$$\operatorname{var}\left(\overline{Y}_n\right) = \frac{1}{n^2}\left\{\sum_{i=1}^{n}\sum_{j=1}^{n}\left(\mathrm{E}\left[Y_i Y_j\right] - \mathrm{E}\left[Y_i\right]\mathrm{E}\left[Y_j\right]\right)\right\} \tag{10.188}$$

$$= \frac{1}{n^2}\left\{\sum_{i=1}^{n}\sum_{j=1}^{n}\operatorname{cov}\left(Y_i, Y_j\right)\right\} \tag{10.189}$$

As all the variables are i.i.d's and are thus uncorrelated,

$$\operatorname{cov}\left(Y_i, Y_j\right) = \begin{cases} 0 & \text{if } i \neq j \\ \operatorname{var}\left(Y_i\right) & \text{if } i = j \end{cases} \tag{10.190}$$

Putting (10.125) in (10.124),

$$\operatorname{var}\left(\overline{Y}_n\right) = \frac{1}{n^2}\left(\sum_{i=1}^{n} \operatorname{cov}\left(Y_i, Y_i\right)\right) \tag{10.191}$$

$$= \frac{1}{n^2}\left(\sum_{i=1}^{n} \operatorname{var}\left(Y_i\right)\right) \tag{10.192}$$

$$= \frac{1}{n^2} \cdot n = \frac{1}{n} \tag{10.193}$$

$$\operatorname{var}\left(\overline{Y}_n\right) = \frac{1}{n} \tag{10.194}$$

$$\implies \mathrm{E}\left(\overline{Y}_n\right) = 1 \text{ and } \operatorname{Var}\left(\overline{Y}_n\right) = \frac{1}{n} \tag{10.195}$$

By the Central Limit Theorem, $n \to \infty \implies \sqrt{n}(Y_n - \mu) \to \mathcal{N}(0, 1)$

$$\frac{\overline{Y_n} - 1}{\frac{1}{\sqrt{n}}} \sim \mathcal{N}(0, 1) \tag{10.196}$$

$$\sqrt{n}(\overline{Y_n} - 1) \sim \mathcal{N}(0, 1) \tag{10.197}$$

We know,

$$\sqrt{n}(Y_n - k) \sim \mathcal{N}(0, \sigma^2) \tag{10.198}$$

Let us write the taylor expansion of $g\left(Y_n\right)$ around k

$$g(Y_n) = g(k) + g'(k)(Y_n - k) + \frac{1}{2}g''(k)(Y_n - k)^2 + \ldots \tag{10.199}$$

Apply the Central Limit Theorem (CLT) to the standardized variable Z_n

$$Z_n = \frac{\sqrt{n}g'(k)(Y_n - k)}{\sigma\sqrt{n}} \tag{10.200}$$

$$n \to \infty \implies Z_n \to \mathcal{N}(0,1) \tag{10.201}$$

Compare with standardised variable we get,

$$\implies \sqrt{n}(g(Y_n) - g(k)) \sim \mathcal{N}(0, \sigma^2[g'(k)]^2) \tag{10.202}$$

$$g(x) = e^{-x} \implies g'(x) = -e^{-x} \tag{10.203}$$

Using equation (10.137), we get:

$$\sqrt{n}(e^{-\overline{Y}_n} - e^{-1}) \sim \mathcal{N}(0, e^{-2}) \tag{10.204}$$

$$\implies \sigma^2 = e^{-2} = 0.14 \tag{10.205}$$

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 ∞.

(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 $\overline{Y}$ which is mean of 'n' samples of Y.

(f) Now calculate $\sqrt{n}(e^{-\overline{Y}_n} - e^{-1})$ 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.

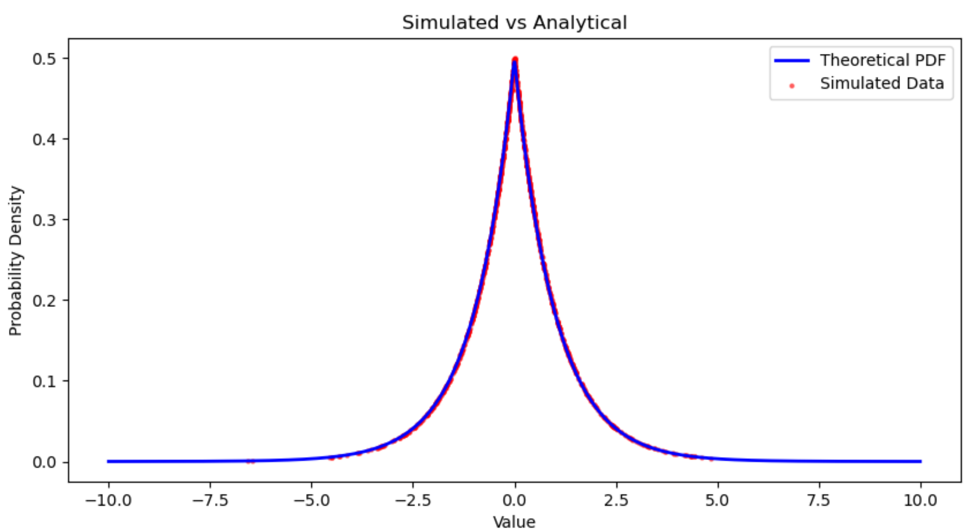


Figure 10.3: pdf of the laplacian

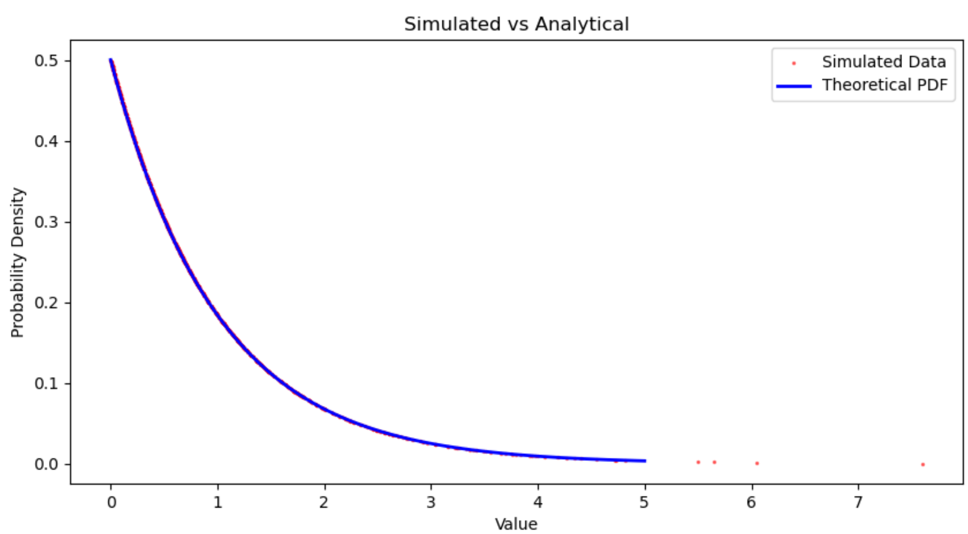


Figure 10.4: pdf of absolute of the laplacian

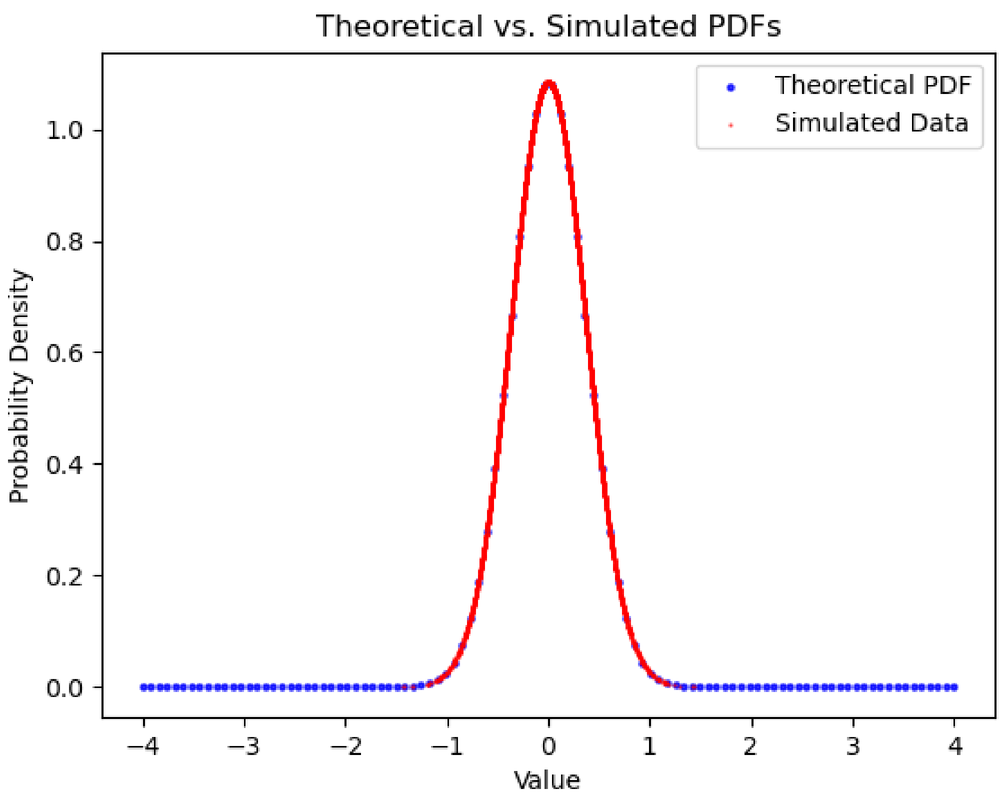


Figure 10.5: Gaussian pdf

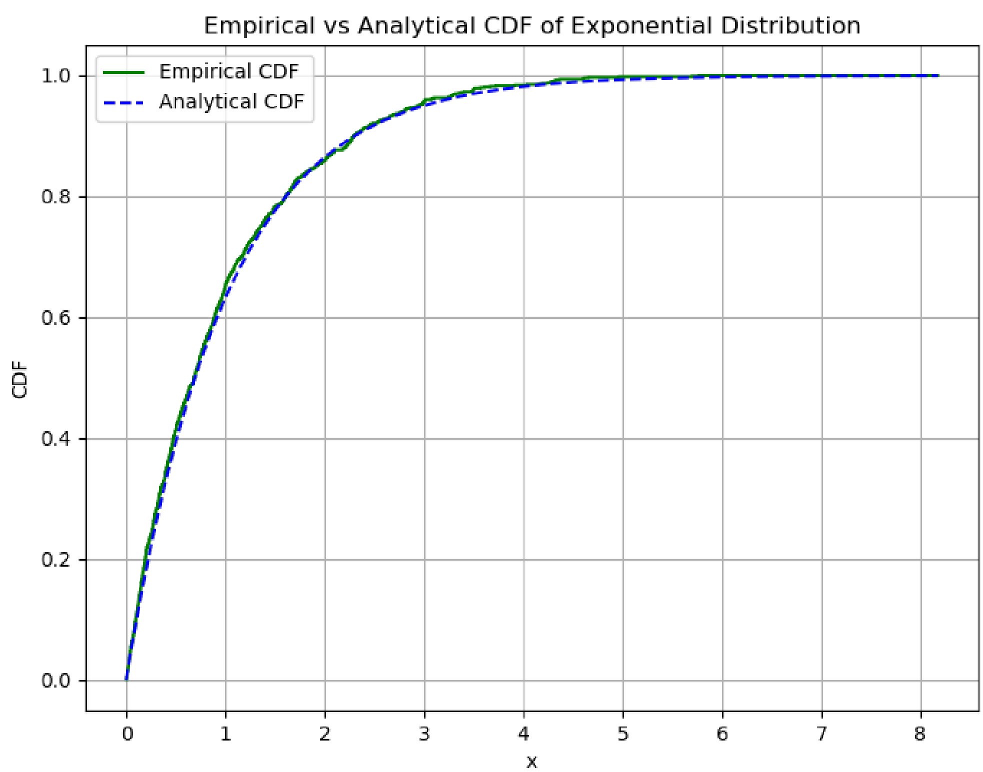


Figure 10.6: Cdf of X

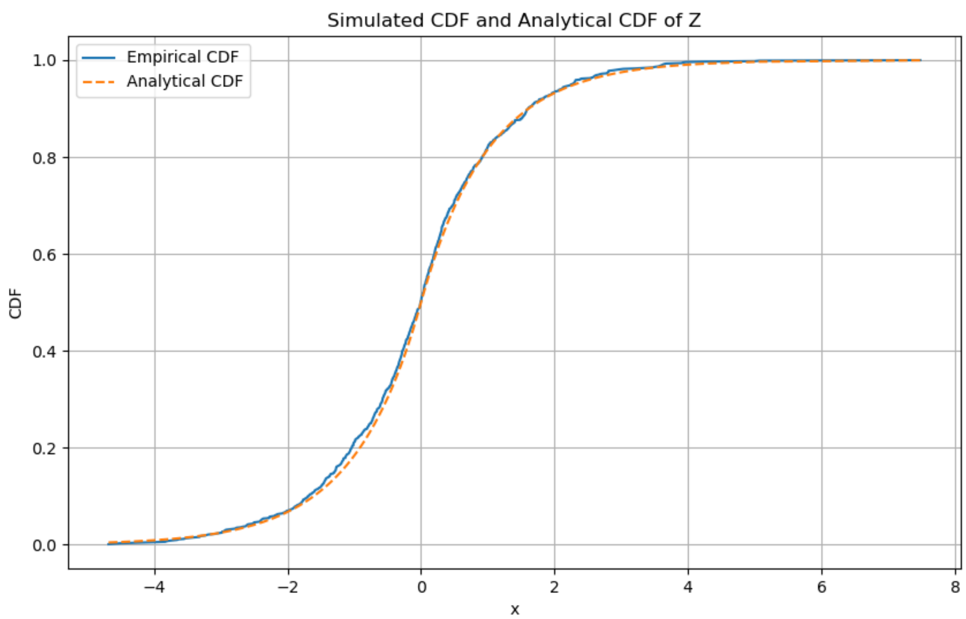


Figure 10.7: Cdf of Z

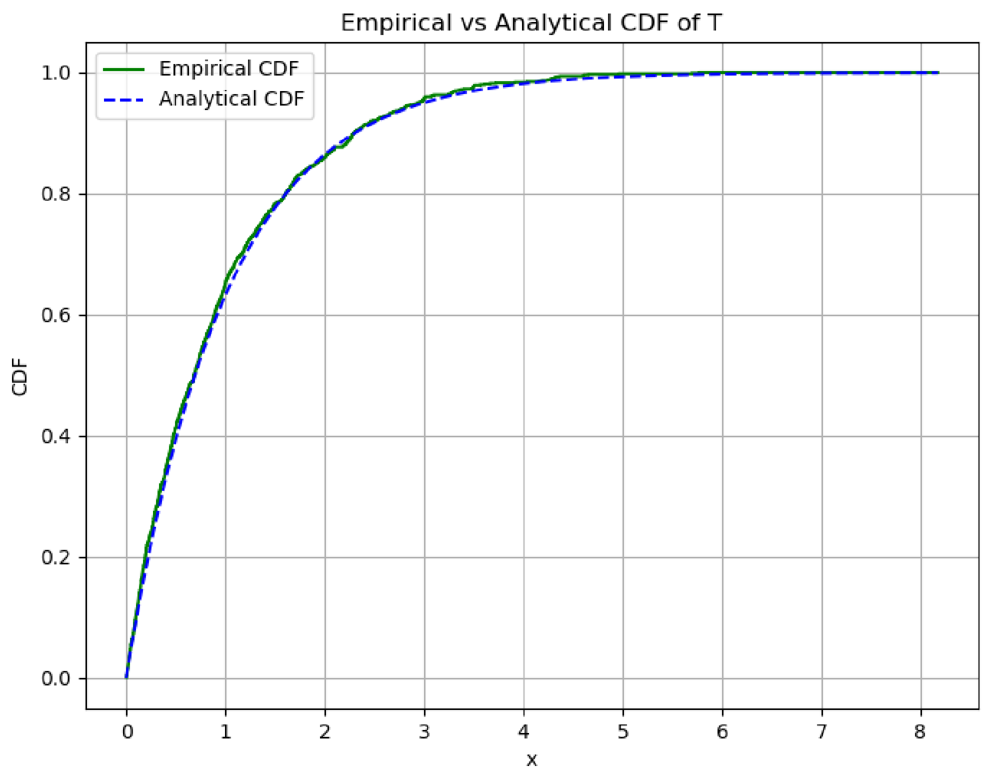


Figure 10.8: Cdf of T

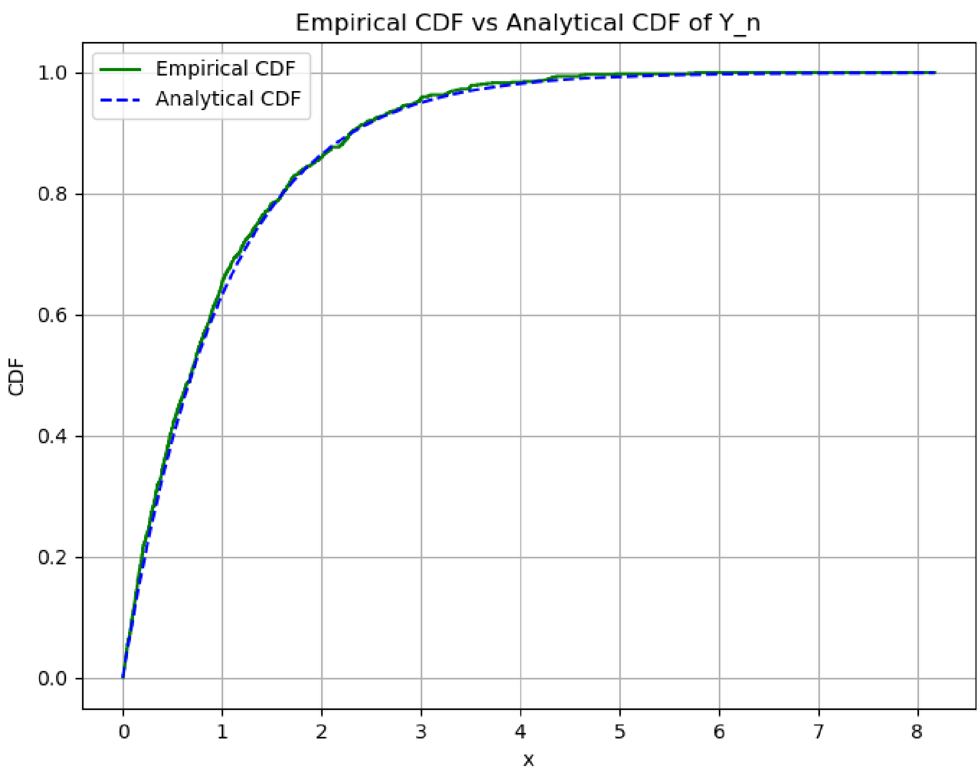


Figure 10.9: Cdf of Y_n

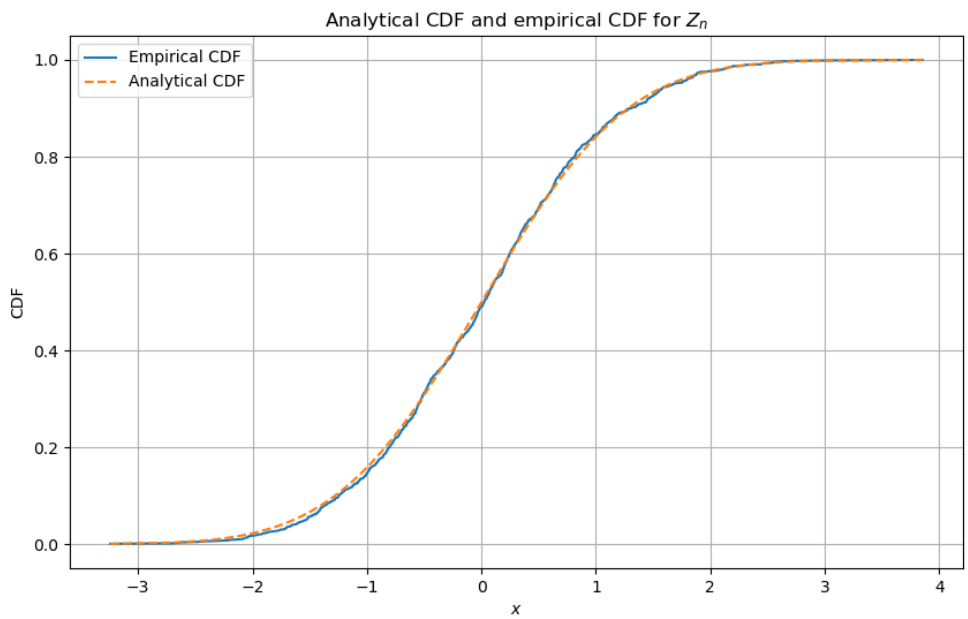


Figure 10.10: Cdf of Z_n

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 occurance**
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	$\frac{1}{128}$

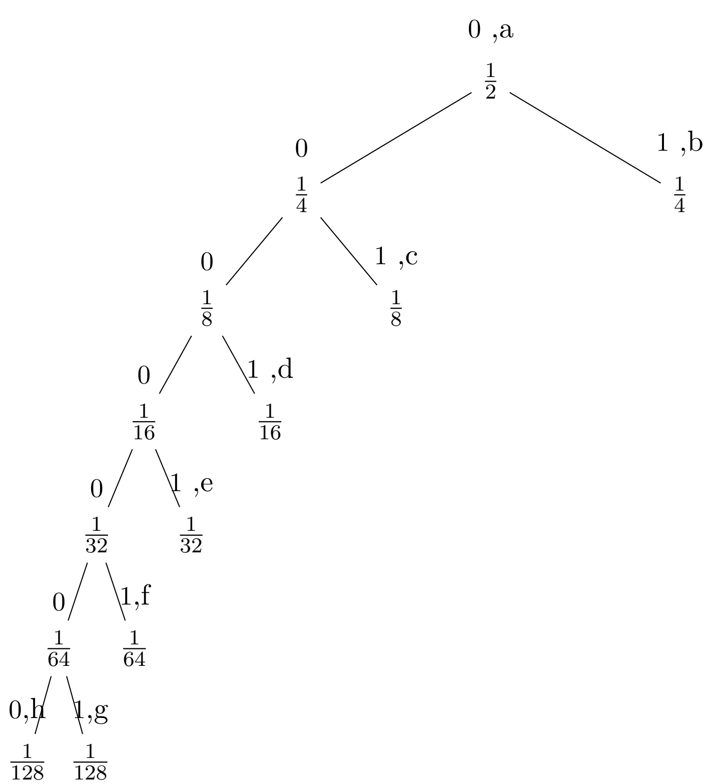
Solution: (GATE EC 2022)

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

$$
\begin{aligned}
H(X) &= \sum_i p_X(i) \log_b \left(\frac{1}{p_X(i)} \right) && (11.1) \\
&= \frac{1}{2} \log_2 (2) + \frac{1}{4} \log_2 (4) + ... + \frac{1}{128} \log_2 (128) && (11.2) \\
&= 0.5 + 0.5 + 0.375 + ... + 0.0078125 && (11.3) \\
&= 1.984375 && (11.4)
\end{aligned}
$$

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;

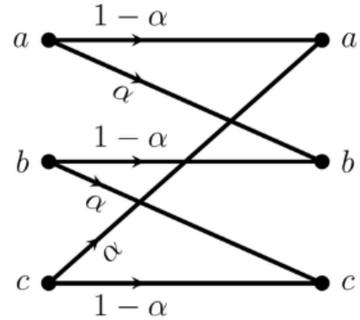
Symbols	**Frequency**	**Code**	**Size**
a	$\frac{1}{2}$	1	0.5
b	$\frac{1}{4}$	01	0.25
c	$\frac{1}{8}$	001	0.125
d	$\frac{1}{16}$	0001	0.0625
e	$\frac{1}{32}$	00001	0.03125
f	$\frac{1}{64}$	000001	0.015625
g	$\frac{1}{128}$	0000001	0.0078125
h	$\frac{1}{128}$	0000000	0.0078125

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 α lies in the interval $[0.25, 1]$. The value of α 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

$$C = \sup_{p_X(x)} I(X, Y) \tag{11.5}$$

$$\begin{aligned}
I(X, Y) &= \sum_{x,y} p(x, y) \log_2 \frac{p(x, y)}{p(x)\, p(y)} && (11.6)\\
&= \sum_{x,y} p(x, y) \log_2 \frac{p(y|\, x)}{p(y)} && (11.7)\\
&= -\sum_{x,y} p(x, y) \log_2 p(y) + \sum_{x,y} p(x, y) \log_2 p(y|\, x) && (11.8)\\
&= -\sum_{y} p(y) \log_2 p(y) - \left(-\sum_{x,y} p(x, y) \log_2 p(y|\, x)\right) && (11.9)\\
&= H(Y) - H(Y|\, X) && (11.10)
\end{aligned}$$

Now,

$$\sum_{i=0}^{2} p_i = 1 \tag{11.11}$$

$$\sum_{j=0}^{2} q_i = 1 \tag{11.12}$$

$$H(\mathbf{q}) = -\sum_{i=0}^{2} q_i \log_2 q_i \tag{11.13}$$

$$= -(q_0 \log_2 q_0 + q_1 \log_2 q_1 + q_2 \log_2 q_2) \tag{11.14}$$

$$H(Y|\, X) = -\sum_{i=0}^{2}\sum_{j=0}^{2} p_i p_{Y|X}(y_j|\, x_i) \log_2 \left(p_{Y|X}(y_j|\, x_i)\right) \tag{11.15}$$

$$\begin{aligned} = -p_0 ((1-\alpha)\log_2(1-\alpha) + \alpha \log_2 \alpha) \\ - p_1 ((1-\alpha)\log_2(1-\alpha) + \alpha \log_2 \alpha) \\ - p_2 ((1-\alpha)\log_2(1-\alpha) + \alpha \log_2 \alpha) \end{aligned} \tag{11.16}$$

Using (11.14) and (11.16) in (11.10)

$$\begin{aligned} I(X,Y) = -(q_0 \log_2 q_0 + q_1 \log_2 q_1 + q_2 \log_2 q_2) \\ + p_0 ((1-\alpha)\log_2(1-\alpha) + \alpha \log_2 \alpha) \\ + p_1 ((1-\alpha)\log_2(1-\alpha) + \alpha \log_2 \alpha) \\ + p_2 ((1-\alpha)\log_2(1-\alpha) + \alpha \log_2 \alpha) \end{aligned} \tag{11.17}$$

$$\implies \frac{d}{d\alpha} I(X,Y) = p_0 \log_2\left(\frac{\alpha}{1-\alpha}\right) + p_1 \log_2\left(\frac{\alpha}{1-\alpha}\right)$$
$$+ p_2 \log_2\left(\frac{\alpha}{1-\alpha}\right) \quad (11.18)$$

For Maxima or minima $\frac{d}{d\alpha} I(X,Y) = 0$

$$\log_2\left(\frac{\alpha}{1-\alpha}\right)(p_0 + p_1 + p_2) = 0 \quad (11.19)$$
$$\implies \alpha = \frac{1}{2} \quad (11.20)$$

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) $H(X) \leq H\left(X^2\right)$

(D) $H(X) \leq H\left(2^X\right)$

Solution:

Random independent variable	value of R.V	Description
X	$X \in (x_1, x_2 ..., x_K)$	Value of the discrete variable X

(a) For Option(A) we will find

We know that :

$$\begin{aligned} \max_{p_X(k)} \quad & H(X) \\ \text{s.t.} \quad & \sum_{k=0}^{K} p_X(k) = 1 \end{aligned}$$

$\Longrightarrow$

$$\begin{aligned} \max_{p_X(k)} \quad & -\sum_{k=0}^{K} p_X(k) \log_2 p_X(k) \\ \text{s.t.} \quad & \sum_{k=0}^{K} p_X(k) = 1 \end{aligned}$$

Now,we use lagranges multiplier to find the maximum entropy subject to the lagranges multiplier constant λ and $p_X(k)$

$$L(p_X(k), \lambda) = -\sum_{k=0}^{K} p_X(k) \log_2 p_X(k) + \lambda \left(\sum_{k=0}^{k} p_X(k) - 1 \right) \tag{11.21}$$

$$\frac{\partial L}{\partial p_X(k)} = -\log_2 p_X(k) - 1 + \lambda \tag{11.22}$$

Now, we take the derivative of L with respect to each $p_X(k)$ equal to zero for $H(X)\, max$

$$\lambda = \log_2 \frac{2}{k} \tag{11.23}$$

$$p_X(k) = 1/K \tag{11.24}$$

On solving, we get the value of

$$H(X)_{max} = \log_2 K \tag{11.25}$$

$$H(X) \le \log_2 K \tag{11.26}$$

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	$\frac{1}{4}$

$$H(X) = \frac{1}{4}\log_2 4 + \frac{1}{2}\log_2 2 + \frac{1}{4}\log_2 4 \tag{11.27}$$

$$H(X) = 1.5units \tag{11.28}$$

Now $Y = 2X$

$Y \in y_i$	$p_Y(k)$
-2	$\frac{1}{4}$
0	$\frac{1}{2}$
2	$\frac{1}{4}$

$$H(Y) = \sum_{i=0}^{2} p_Y(k) \log_2 \frac{1}{p_Y(k)} \tag{11.29}$$

$$H(Y) = 1.5 units \tag{11.30}$$

$$H(Y) = H(2X) = H(X) \tag{11.31}$$

Hence, Option(B) is correct

(c) Similarly on substituting $Y = X^2$

$Y \in y_i$	$p_Y(k)$
0	$\frac{1}{2}$
1	$\frac{1}{2}$

$$H(Y) = \sum_{i=0}^{1} p_Y(k) \log_2 \frac{1}{p_Y(k)} \tag{11.32}$$

$$H(Y) = 1 units \tag{11.33}$$

$$H(Y) = H\left(X^2\right) \leq H(X) \tag{11.34}$$

Hence, Option(C) is incorrect

(d) Now for $Y = 2^X$

$Y \in y_i$	$p_Y(k)$
$2^{-1} = \frac{1}{2}$	$\frac{1}{4}$
$2^0 = 1$	$\frac{1}{2}$
$2^1 = 2$	$\frac{1}{4}$

$$H(Y) = \sum_{i=0}^{2} p_Y(k) \log_2 \frac{1}{p_Y(k)} \tag{11.35}$$

$$H(Y) = 1.5 units \tag{11.36}$$

$$H(Y) = H\left(2^X\right) = H(X) \tag{11.37}$$

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:

$$\begin{pmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{pmatrix} \tag{12.1}$$

$$\Pr(X_2 = 1 | X_1 = 1, X_3 = 2) = \Pr(X_2 = 1 | X_1 = 1) \tag{12.2}$$

$$= p_{11} \tag{12.3}$$

$$= 0.5 \tag{12.4}$$

(by markov's property and using transition probability matrix)

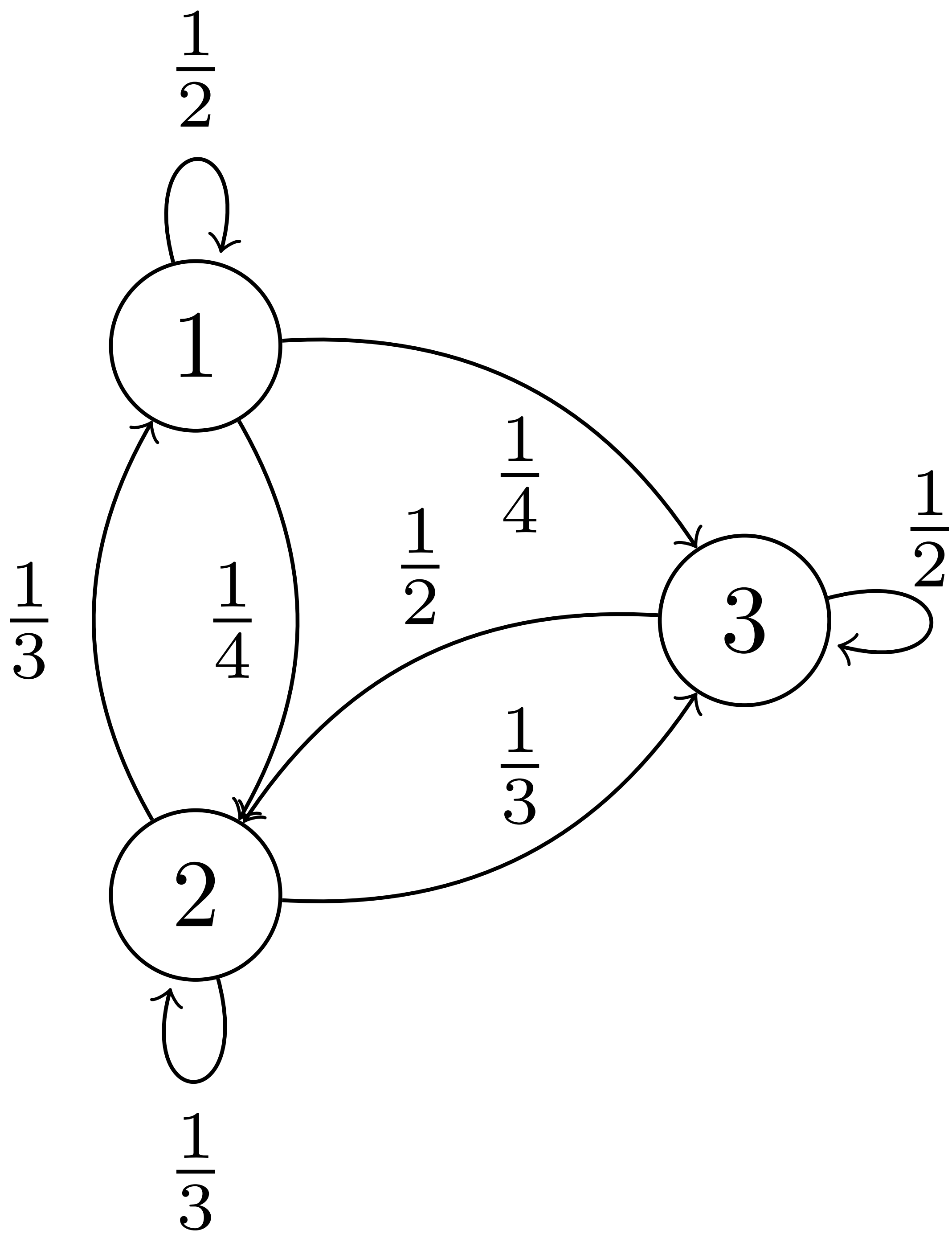


Figure 12.1: Markov Chain diagram

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 $N\left(\frac{1}{2}, \sigma^2\right)$ distribution, where $\sigma > 0$ is an unknown parameter. Let T be an unbiased estimator of σ^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

$$\mathrm{E}(\hat{\sigma^2}) = \sigma^2 \tag{13.1}$$

where, $E(\hat{\sigma^2})$ represents the expected value of the estimator $\hat{\sigma^2}$ and σ^2 represents the true parameter

Definition 13.2: The Cramér-Rao bound can be defined as follows:

$$\mathrm{Var}(\sigma^2) \geq \frac{1}{I(\sigma^2)} \tag{13.2}$$

where $I(\sigma^2)$ represents fisher information for the parameter σ^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 σ^2.

$P_X(X|\sigma^2)$ is given by:

$$P_X(X|\sigma^2) = \frac{1}{2\pi\sigma^2}\exp\left(-\frac{(X-\frac{1}{2})^2}{2\sigma^2}\right) \tag{13.3}$$

$$\log p_X(X|\sigma^2) = \log\left(\frac{1}{2\pi\sigma^2}\exp\left(-\frac{(X-\frac{1}{2})^2}{2\sigma^2}\right)\right) \tag{13.4}$$

$$= -\frac{1}{2}\log(2\pi\sigma^2) - \frac{(X-\frac{1}{2})^2}{2\sigma^2} \tag{13.5}$$

$$\frac{\partial^2}{\partial(\sigma^2)^2}\log P_X(X|\sigma^2) = \frac{1}{2\pi\sigma^2} - \frac{3(X-\frac{1}{2})^2}{\sigma^4} \tag{13.6}$$

$$I(\sigma^2) = \frac{3}{\sigma^4}E[X^2] - \frac{3}{\sigma^4}E[X] + \frac{3}{4\sigma^4} - \frac{1}{2\pi\sigma^2} \tag{13.7}$$

$$E[X^2] = \sigma^2 + \left(\frac{1}{2}\right)^2 \tag{13.8}$$

$$E[X] = \frac{1}{2} \tag{13.9}$$

$$\implies I(\sigma^2) = \left(3 - \frac{1}{2\pi}\right)\frac{1}{\sigma^2} \tag{13.10}$$

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

$\implies$ 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,

$$T = \frac{\sum(X_i - \mu)^2}{n} \tag{13.11}$$

$$n = 5 \tag{13.12}$$

$$\mu = \frac{1}{2} \tag{13.13}$$

$$\sum(X_i - \mu)^2 = 13.75 \tag{13.14}$$

Hence,

$$T = 2.75 \tag{13.15}$$

Since, T is an unbiased estimator of σ^2,

$$\text{Cramér-Rao bound} = \frac{T}{(3 - \frac{1}{2\pi})} \tag{13.16}$$

$$= 0.968 \tag{13.17}$$

13.2 Let X be a random variable with probability density function

$$f(x;\lambda) = \begin{cases} \frac{1}{\lambda}e^{-\frac{x}{\lambda}} & \text{if} x > 0 \\ 0 & \text{otherwise} \end{cases} \tag{13.18}$$

where $\lambda > 0$ is an unknown parameter. Let $Y_1, Y_2, ..., Y_n$ be a random sample of size n

from a population having the same distribution as X^2.If

$$\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i \tag{13.19}$$

then which of the following statements is true?

(a) $\sqrt{\frac{\bar{Y}}{2}}$is a method of moments estimator ofλ

(b) $\sqrt{\bar{Y}}$is a method of moments estimator ofλ

(c) $\frac{1}{2}\sqrt{\bar{Y}}$is a method of moments estimator of λ

(d) $2\sqrt{\bar{Y}}$is a method of moments estimator ofλ (GATE ST 2023)

Solution:

(a) Using PDF in (13.18) we need to find an estimator for the unknown parameter λ in terms of sample mean $\bar{Y}$

we know $Y_i = X_i^2$ then,

$$E(Y_i) = E(X_i^2) \tag{13.20}$$

$$= \int_0^\infty x^2 \frac{1}{\lambda} e^{-\frac{x}{\lambda}} \tag{13.21}$$

$$= 2\lambda^2 \tag{13.22}$$

Method of moment is defined by (13.19) which gives,

$$\bar{Y} = E(Y_i) \tag{13.23}$$

$$= 2\lambda^2 \tag{13.24}$$

where

$$\lambda = \sqrt{\frac{\bar{Y}}{2}} \tag{13.25}$$

$\therefore$ Option (13.2a) is correct.

(b) The simulation steps to estimate λ using method of moment estimator in python.

i. Generate a random value of λ within the specified range using **np.random.uniform(1**

ii. Use the generated λ 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 λ 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)

$$Y_i = \beta_0 + \beta_1 X_i + e_i \tag{13.26}$$

Here

Definition 13.4: Residual sum of squares(RSS):

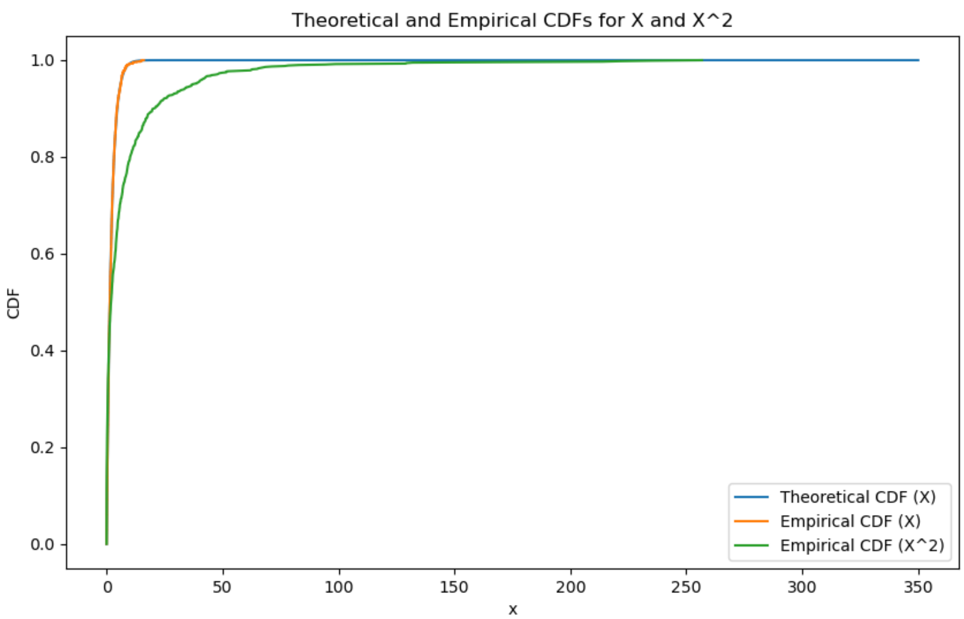


Figure 13.1: Figure1

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.

$$
\begin{aligned}
RSS &= \sum \left(Y_i - \hat{Y}\right) && (13.27)\\
&= \sum e_i^2 && (13.28)\\
&= \sum (Y_i - \beta_0 - \beta_1 X_i)^2 && (13.29)
\end{aligned}
$$

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	Independant variables
β_0, β_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.

$$TSS = \sum \left(Y_i - \bar{Y}\right)^2 \tag{13.30}$$

where $\bar{Y}$ = Mean of data

Definition 13.6: Coeffiecient 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.

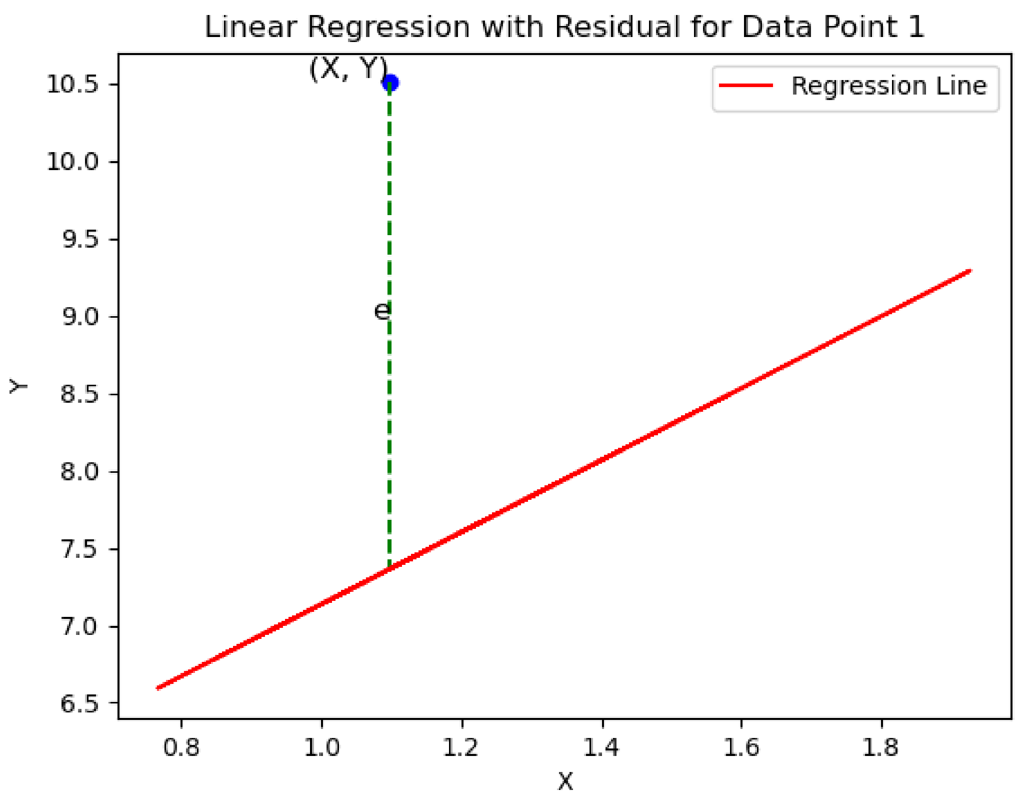
$$R^2 = 1 - \frac{RSS}{TSS} \tag{13.31}$$

$$= 1 - \frac{44}{80} \tag{13.32}$$

$$= 0.45 \tag{13.33}$$

45 percent of the variance in the Y variable is predicted from the X variable.

13.4 Consider the following regression model



$$y_k = \alpha_0 + \alpha_1 \log_e k + \epsilon_k, \qquad k = 1, 2, \ldots, n,$$

where ϵ_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 α_0 does not exist

(B) The maximum likelihood estimator of α_1 does not exist

(C) The least squares estimator of α_0 exists and is unique

(D) The least squares estimator of α_1 exists, but it is not unique

(GATE ST 2023)

Solution:

$$f(\epsilon_k) = \frac{1}{2}e^{-|\epsilon_k|} \tag{13.34}$$

$$\text{Likelihood function} : f(\epsilon_1\epsilon_2....\epsilon_n) = \prod_{k=1}^{n} f(\epsilon_k) \tag{13.35}$$

$$L = \prod_{k=1}^{n} \frac{1}{2}e^{-|\epsilon_k|} \tag{13.36}$$

$$L_1 = \ln L = \ln\left(\prod_{k=1}^{n} \frac{1}{2}e^{-|\epsilon_k|}\right) \tag{13.37}$$

$$= \sum_{k=1}^{n} \ln\left(\frac{1}{2}e^{-|\epsilon_k|}\right) \tag{13.38}$$

$$= \sum_{k=1}^{n} (-\ln 2 - |y_k - \alpha_0 - \alpha_1 \log_e k|) \tag{13.39}$$

$$= -n\ln 2 - \sum_{k=1}^{n} (|y_k - \alpha_0 - \alpha_1 \log_e k|) \tag{13.40}$$

$$L_1 = \text{function of } \alpha_0, \alpha_1 \tag{13.41}$$

(a) Maximum likelihood estimator

We need to find the value of α_0 and α_1 which will maximise the value of L_1 i.e. the value of α_0 and α_1 which will minimise the value of $\sum_{k=1}^{n} |y_k - \alpha_0 - \alpha_1 \log_e k|$

i. With respect to α_0

A. For $y_k - \alpha_0 - \alpha_1 \log_e k > 0$

$$\min_{\alpha_0} \quad y_k - \alpha_0 - \alpha_1 \log_e k$$

$$\text{s.t.} \quad \alpha_0 \leq y_k - \alpha_1 \log_e k$$

Using Lagrange multiplier method

$$L\left(\lambda\right) = y_k - \alpha_0 - \alpha_1 \log_e k - \lambda(\alpha_0 - y_k + \alpha_1 \log_e k) \tag{13.42}$$
$$\frac{\partial L}{\partial \alpha_0} = -1 - \lambda = 0 \tag{13.43}$$
$$\frac{\partial L}{\partial \lambda} = y_k - \alpha_0 - \alpha_1 \log_e k = 0 \tag{13.44}$$
$$\lambda = -1 \tag{13.45}$$
$$\alpha_0 = y_k - \alpha_1 \log_e k \tag{13.46}$$

B. For $y_k - \alpha_0 - \alpha_1 \log_e k < 0$

$$\begin{aligned} \min_{\alpha_0} \quad & -\left(y_k - \alpha_0 - \alpha_1 \log_e k\right) \\ \text{s.t.} \quad & \alpha_0 \geq y_k - \alpha_1 \log_e k \end{aligned}$$

Using Lagrange multiplier method

$$L\left(\lambda\right) = -\left(y_k - \alpha_0 - \alpha_1 \log_e k\right) - \lambda(\alpha_0 - y_k + \alpha_1 \log_e k) \tag{13.47}$$
$$\frac{\partial L}{\partial \alpha_0} = 1 - \lambda = 0 \tag{13.48}$$
$$\frac{\partial L}{\partial \lambda} = y_k - \alpha_0 - \alpha_1 \log_e k = 0 \tag{13.49}$$
$$\lambda = 1 \tag{13.50}$$
$$\alpha_0 = y_k - \alpha_1 \log_e k \tag{13.51}$$

As value of α_0 matches for both cases of modulus

$\therefore$ The maximum likelihood estimator of α_0 exist

ii. With respect to α_1

A. For $y_k - \alpha_0 - \alpha_1 \log_e k > 0$

$$\begin{aligned} \min_{\alpha_1} \quad & y_k - \alpha_0 - \alpha_1 \log_e k \\ \text{s.t.} \quad & \alpha_1 \le \frac{y_k - \alpha_0}{\log_e k} \end{aligned}$$

Using Lagrange multiplier method

$$L(\lambda) = y_k - \alpha_0 - \alpha_1 \log_e k - \lambda \left(\alpha_1 - \frac{y_k - \alpha_0}{\log_e k} \right) \tag{13.52}$$

$$\frac{\partial L}{\partial \alpha_1} = -\log_e k - \lambda = 0 \tag{13.53}$$

$$\frac{\partial L}{\partial \lambda} = -\left(\alpha_1 - \frac{y_k - \alpha_0}{\log_e k} \right) = 0 \tag{13.54}$$

$$\lambda = -\log_e k \tag{13.55}$$

$$\alpha_1 = \frac{y_k - \alpha_0}{\log_e k} \tag{13.56}$$

B. For $y_k - \alpha_0 - \alpha_1 \log_e k < 0$

$$\begin{aligned} \min_{\alpha_1} \quad & -(y_k - \alpha_0 - \alpha_1 \log_e k) \\ \text{s.t.} \quad & \alpha_1 \ge \frac{y_k - \alpha_0}{\log_e k} \end{aligned}$$

Using Lagrange multiplier method

$$L(\lambda) = -(y_k - \alpha_0 - \alpha_1 \log_e k) - \lambda\left(\alpha_1 - \frac{y_k - \alpha_0}{\log_e k}\right) \tag{13.57}$$

$$\frac{\partial L}{\partial \alpha_1} = \log_e k - \lambda = 0 \tag{13.58}$$

$$\frac{\partial L}{\partial \lambda} = -\left(\alpha_1 - \frac{y_k - \alpha_0}{\log_e k}\right) = 0 \tag{13.59}$$

$$\lambda = \log_e k \tag{13.60}$$

$$\alpha_1 = \frac{y_k - \alpha_0}{\log_e k} \tag{13.61}$$

As value of α_1 matches for both cases of modulus

$\therefore$ The maximum likelihood estimator of α_1 exist

$\therefore$ Option (A) and (B) are incorrect

iii. Least square estimator

The least square estimator of α_0 and α_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

$$Q(\alpha_0, \alpha_1) = \sum_{k=1}^{n} (y_k - \alpha_0 - \alpha_1 \log_e k)^2 \tag{13.62}$$

$$\frac{\partial Q}{\partial \alpha_0} = -2 \sum_{k=1}^{n} (y_k - \alpha_0 - \alpha_1 \log_e k) = 0 \tag{13.63}$$

$$\sum_{k=1}^{n} (y_k - \alpha_0 - \alpha_1 \log_e k) = 0 \tag{13.64}$$

$$n\bar{y} - n\alpha_0 - \alpha_1 n\bar{x} = 0 \tag{13.65}$$

$$\implies \tilde{\alpha_0} = \bar{y} - \tilde{\alpha_1}\bar{x} \tag{13.66}$$

$$\frac{\partial Q}{\partial \alpha_1} = -2 \sum_{k=1}^{n} (y_k - \alpha_0 - \alpha_1 \log_e k) \log_e k = 0 \tag{13.67}$$

$$\implies \tilde{\alpha_1} = \frac{\sum_{k=1}^{n} (\log_e k - \bar{x})(y_k - \bar{y})}{\sum_{k=1}^{n} (\log_e k - \bar{x})^2} \tag{13.68}$$

$\therefore$ Least square estimator of α_0 and α_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

$$p_X(x) = \begin{cases} \frac{1}{2}e^{x} & x \le 0 \\ \frac{1}{2}e^{-x} & x > 0 \end{cases} \tag{13.69}$$

$$F_X(x) = \begin{cases} \int_{-\infty}^{x} \left(\frac{1}{2}e^{x}\right) dx & x \le 0 \\ \int_{-\infty}^{0} \left(\frac{1}{2}e^{x}\right) dx + \int_{0}^{x} \left(\frac{1}{2}e^{-x}\right) dx & x > 0 \end{cases} \tag{13.70}$$

$$F_X(x) = \begin{cases} \frac{1}{2}e^x & x \le 0 \\ \frac{1}{2}\left(2 - e^{-x}\right) & x > 0 \end{cases} \tag{13.71}$$

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 ≤ 0

$$u = \frac{1}{2}e^x \tag{13.72}$$

$$e^x = 2u \tag{13.73}$$

$$x = \ln 2u \tag{13.74}$$

$$\because x \le 0 \tag{13.75}$$

$$u \le 0.5 \tag{13.76}$$

For x > 0

$$u = \frac{1}{2}\left(2 - e^{-x}\right) \tag{13.77}$$

$$2 - e^{-x} = 2u \tag{13.78}$$

$$e^{-x} = 2 - 2u \tag{13.79}$$

$$x = -\ln(2 - 2u) \tag{13.80}$$

$$\because x > 0 \tag{13.81}$$

$$u > 0.5 \tag{13.82}$$

$$I(u) = \begin{cases} \ln(2u) & u \le 0.5 \\ -\ln(2-2u) & u > 0.5 \end{cases} \tag{13.83}$$

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 funtion 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

$$y_t = \alpha_0 + \alpha_1 t + \alpha_2 t^2 + \epsilon_t, \qquad t = 1, 2, \ldots, n \tag{13.84}$$

where α_0 , α_1 and α_2 are unknown parameters and ϵ_t's are independent and identically distributed random variables each having $\mathcal{N}(\mu, 1)$ distribution with μ unknown. Then which of the following statements is/are true?

(a) There exists an unbiased estimator of α_1

(b) There exists an unbiased estimator of α_2

(c) There exists an unbiased estimator of α_0

(d) There exists an unbiased estimator of μ

(GATE ST 2023)

Solution: Assuming that the model is

$$y_t = \alpha_0 + \alpha_1 t + \alpha_2 t^2 + \epsilon_t \tag{13.85}$$

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & A_{11} & A_{12} \\ 1 & A_{21} & A_{22} \\ \vdots & & \\ 1 & A_{n1} & A_{n2} \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix} \tag{13.86}$$

Finding mean of the $\mathbf{y}$,

$$\mathbf{y} = \mathbf{A}\mathbf{a} + \boldsymbol{\epsilon} \tag{13.87}$$

$$E(\mathbf{y}) = E(\mathbf{A}\mathbf{a} + \boldsymbol{\epsilon}) \tag{13.88}$$

$$= \mathbf{A}\mathbf{a} + E(\boldsymbol{\epsilon}) \tag{13.89}$$

$$E(\boldsymbol{\epsilon}) = \begin{bmatrix} E(\epsilon_1) \\ E(\epsilon_2) \\ \cdot \\ \cdot \\ \cdot \\ E(\epsilon_n) \end{bmatrix} = \begin{bmatrix} \mu \\ \mu \\ \cdot \\ \cdot \\ \cdot \\ \mu \end{bmatrix} \tag{13.90}$$

$$\boldsymbol{\mu}_{\boldsymbol{\epsilon}} = \mu \begin{bmatrix} 1 \\ 1 \\ \cdot \\ \cdot \\ \cdot \\ 1 \end{bmatrix} \tag{13.91}$$

Let $\mathbf{A}_i$ represent the i^{th} row of $\mathbf{A}$

$$\begin{aligned} \mathrm{Cov}(y_i, y_j) &= E\left((y_i - E(y_i))\left(y_j - E(y_j)\right)\right) && (13.92) \\ &= E\left((\mathbf{A}_i\mathbf{a} + \epsilon_i - E(\mathbf{A}_i\mathbf{a} + \epsilon_i))(\mathbf{A}_j\mathbf{a} + \epsilon_j - E(\mathbf{A}_j\mathbf{a} + \epsilon_j))\right) && (13.93) \\ &= E\left((\epsilon_i - E(\epsilon_i))(\epsilon_j - E(\epsilon_j))\right) && (13.94) \\ &= \mathrm{Cov}(\epsilon_i, \epsilon_j) && (13.95) \end{aligned}$$

$$\mathbf{C}_{\mathbf{y}} = \mathbf{C}_{\boldsymbol{\epsilon}} \tag{13.96}$$

Since the ϵ_i's are independent and identical vectors,

$$\mathrm{Cov}(\epsilon_i, \epsilon_j) = 0, \qquad \forall i \neq j \tag{13.97}$$

$$\mathrm{Var}(\epsilon_i) = 1, \qquad \forall 1 \le i \le n \tag{13.98}$$

$$\mathbf{C_y} = I_{n\times n} \tag{13.99}$$

$$\mathbf{y} \sim \mathcal{N}\left(\mathbf{Aa} + \boldsymbol{\mu_\epsilon}, I\right) \tag{13.100}$$

$$p_{\mathbf{y}} = \frac{1}{\sqrt{(2\pi)^n \det(\mathbf{C_y})}} \exp \frac{-(\mathbf{y} - \boldsymbol{\mu_\epsilon} - \mathbf{Aa})^T \mathbf{C_y}^{-1} (\mathbf{y} - \boldsymbol{\mu_\epsilon} - \mathbf{Aa})}{2} \tag{13.101}$$

where $\mathbf{C_y}$ is the covariance matrix for $\mathbf{y}$

The maximum likelihood function can be written as:

$$L\left(\mathbf{y}\right) = \frac{1}{(2\pi)^{\frac{n}{2}}} e^{\frac{-(\mathbf{y} - \boldsymbol{\mu_\epsilon} - \mathbf{Aa})^T (\mathbf{y} - \mathbf{Aa} - \boldsymbol{\mu_\epsilon})}{2}} \tag{13.102}$$

$$\ln L\left(\mathbf{y}\right) = -\frac{n}{2}\ln(2\pi) - \frac{1}{2}(\mathbf{y} - \boldsymbol{\mu_\epsilon} - \mathbf{Aa})^T (\mathbf{y} - \boldsymbol{\mu_\epsilon} - \mathbf{Aa}) \tag{13.103}$$

$$\frac{\partial \ln L\left(\mathbf{y}\right)}{\partial \mathbf{a}} = \frac{\partial(-\mathbf{y}^T\mathbf{Aa} - \mathbf{a}^T\mathbf{A}^T\mathbf{y} + \mathbf{a}^T\mathbf{A}^T\mathbf{Aa} + \mathbf{a}^T\mathbf{A}^T\boldsymbol{\mu_\epsilon} + \boldsymbol{\mu_\epsilon}^T\mathbf{Aa})}{\partial \mathbf{a}} \tag{13.104}$$

$$= -2\mathbf{A}^T\mathbf{y} + 2\mathbf{A}^T\mathbf{Aa} + 2\mathbf{A}^T\boldsymbol{\mu_\epsilon} \tag{13.105}$$

The normal equation is

$$\frac{\partial \ln L\left(\mathbf{y}\right)}{\partial \mathbf{a}} = 0 \tag{13.106}$$

$$\mathbf{a} = \left(\mathbf{A}^T\mathbf{A}\right)^{-1}\left(\mathbf{A}^T\mathbf{y} - \mathbf{A}^T\boldsymbol{\mu_\epsilon}\right) \tag{13.107}$$

For unbiased estimator,

$$E\left(\mathbf{a}\right)=\mathbf{a} \tag{13.108}$$

$$\begin{aligned}
E\left(\mathbf{a}\right) &= E\left(\left(\mathbf{A}^T\mathbf{A}\right)^{-1}\left(\mathbf{A}^T\mathbf{y}-\mathbf{A}^T\boldsymbol{\mu_\epsilon}\right)\right) && (13.109)\\
&= \left(\mathbf{A}^T\mathbf{A}\right)^{-1}\left(\mathbf{A}^T E(\mathbf{y})-\mathbf{A}^T\boldsymbol{\mu_\epsilon}\right) && (13.110)\\
&= \left(\mathbf{A}^T\mathbf{A}\right)^{-1}\left(\mathbf{A}^T\left(\mathbf{A}\mathbf{a}+\boldsymbol{\mu_\epsilon}\right)-\mathbf{A}^T\boldsymbol{\mu_\epsilon}\right) && (13.111)\\
&= \left(\mathbf{A}^T\mathbf{A}\right)^{-1}\left(\mathbf{A}^T\mathbf{A}\right)\mathbf{a} && (13.112)\\
&= \mathbf{a} && (13.113)
\end{aligned}$$

Hence there exist unbiased estimator for $\alpha_0, \alpha_1, \alpha_2$

For Maximum Likelihood Estimator of μ

$$\begin{aligned}
\frac{\partial \ln L\left(\mathbf{y}\right)}{\partial \boldsymbol{\mu_\epsilon}} &= \frac{\partial(-\mathbf{y}^T\boldsymbol{\mu_\epsilon}-\mathbf{a}^T\mathbf{A}^T\boldsymbol{\mu_\epsilon}+\boldsymbol{\mu_\epsilon}\mathbf{y}+\boldsymbol{\mu_\epsilon}\mathbf{A}\mathbf{x}+\boldsymbol{\mu_\epsilon}^T\boldsymbol{\mu_\epsilon})}{\partial \boldsymbol{\mu_\epsilon}} && (13.114)\\
&= -2\mathbf{y}+2\mathbf{A}\mathbf{a}+2\boldsymbol{\mu_\epsilon} && (13.115)
\end{aligned}$$

The normal equation is

$$\frac{\partial \ln L\left(\mathbf{y}\right)}{\partial \boldsymbol{\mu_\epsilon}}=0 \tag{13.116}$$

$$\boldsymbol{\mu_\epsilon}=\mathbf{y}-\mathbf{A}\mathbf{a} \tag{13.117}$$

$$E(\boldsymbol{\mu}_{\boldsymbol{\epsilon}}) = E(\mathbf{y} - \mathbf{Aa}) \tag{13.118}$$

$$= E(\mathbf{y}) - \mathbf{Aa} \tag{13.119}$$

$$= E(\boldsymbol{\epsilon}) \tag{13.120}$$

$$= \boldsymbol{\mu}_{\boldsymbol{\epsilon}} \tag{13.121}$$

Since,

$$\boldsymbol{\mu}_{\boldsymbol{\epsilon}} = \mu \begin{bmatrix} 1 \\ 1 \\ \cdot \\ \cdot \\ \cdot \\ 1 \end{bmatrix} \tag{13.122}$$

Hence there exists an unbiased estimator for μ as well.

Simulation in C We get the following values:

Parameter	True Value	Simulated Value	Average Bias
α_0	3	2.983275	0.016725
α_1	2	2.002000	-0.002000
α_2	0.5	0.500500	-0.000500
μ	0.1	0.1	0

Table 13.4: Simulation results

(a) Defining some arbitrary values for $\alpha_0, \alpha_1, \alpha_2$ and μ

(b) Generating the epsilon using Box-Muller method and using it to find y_i

(c) Generating dependent variable y_i using $y_i = \alpha_0 + \alpha_1 i + \alpha_2 i^2 + \epsilon_i, \qquad i = 1, 2, \ldots, n$

(d) Storing the generated data into a *generated_data.dat* file.

(e) ϵ_i is generated as a gaussian variable

(f) Defining 4 variables $beta_0hat, beta_1hat, beta_2hat$ and $mu - hat$ for string the estimated values

(g) Defining 4 variables $beta_0bias, beta_1bias, beta_2bias$ 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 $X_1, X_2, ..., X_n$ be a random sample of size n from a population having uniform distribution over the interval $\left(\frac{1}{3}, \theta\right)$, where $\theta > \frac{1}{3}$ is an unknown parameter. If $Y = \max\{X_1, X_2, ..., X_n\}$, then which one of the following statements is true?

(a) $\left(\frac{n+1}{n}\right)\left(Y - \frac{1}{3}\right) + \frac{1}{3}$ is an unbiased estimator of θ

(b) $\left(\frac{n}{n+1}\right)\left(Y - \frac{1}{3}\right) + \frac{1}{3}$ is an unbiased estimator of θ

(c) $\left(\frac{n+1}{n}\right)\left(Y + \frac{1}{3}\right) - \frac{1}{3}$ is an unbiased estimator of θ

(d) Y is an unbiased estimator of θ

(GATE ST 2023)

Solution:

Any one of the above expressions is an unbiased estimator of θ if

$$E(\text{estimator}) = \theta$$

We will first calculate expected value of Y

The PDF of X_i can be expressed as

$$p_{X_i}(x) = \frac{1}{\theta - \frac{1}{3}} \ , \ \frac{1}{3} < \theta < 3 \tag{13.123}$$

$$\begin{aligned} F_{X_i}(y) &= \int_{\frac{1}{3}}^{y} p_{X_i}(x)dx & (13.124)\\ &= \int_{\frac{1}{3}}^{y} \frac{1}{\theta - \frac{1}{3}}dx & (13.125)\\ &= \frac{y - \frac{1}{3}}{\theta - \frac{1}{3}} & (13.126) \end{aligned}$$

$$\begin{aligned} F_X(y) &= \Pr(X \leq y) & (13.127)\\ &= \Pr(X_1 \leq y, X_2 \leq y, ..., X_n \leq y) & (13.128)\\ &= \Pr(X_1 \leq y) \Pr(X_2 \leq y) ... \Pr(X_n \leq y) & (13.129)\\ &= \left(\frac{y - \frac{1}{3}}{\theta - \frac{1}{3}}\right)^n & (13.130) \end{aligned}$$

$$\begin{aligned} p_X(y) &= \frac{d}{dy}F_X(y) & (13.131)\\ &= \frac{n}{(\theta - \frac{1}{3})^n}\left(y - \frac{1}{3}\right)^n & (13.132) \end{aligned}$$

$$E(Y) = \int_{\frac{1}{3}}^{\theta} y p_X(y) dy \tag{13.133}$$

$$= \int_{\frac{1}{3}}^{\theta} y \frac{n}{\left(\theta - \frac{1}{3}\right)^n} \left(y - \frac{1}{3}\right)^n dy \tag{13.134}$$

$$= \frac{n}{\left(\theta - \frac{1}{3}\right)^n} \int_{\frac{1}{3}}^{\theta} y \left(y - \frac{1}{3}\right)^{n-1} dy \tag{13.135}$$

$$\text{Let } y - \frac{1}{3} = t \tag{13.136}$$

$$\implies y = t + \frac{1}{3} \implies dy = dt \tag{13.137}$$

$$\text{Therefore,} \tag{13.138}$$

$$E(Y) = \frac{n}{\left(\theta - \frac{1}{3}\right)^n} \int_0^{\theta - \frac{1}{3}} \left(t^n + \frac{t^{n-1}}{3}\right) dt \tag{13.139}$$

$$= \frac{n}{\left(\theta - \frac{1}{3}\right)^n} \left(\frac{\left(\theta - \frac{1}{3}\right)^{n+1}}{n+1} + \frac{\left(\theta - \frac{1}{3}\right)^n}{3n}\right) \tag{13.140}$$

$$= \frac{n}{n+1} \left(\theta - \frac{1}{3}\right) + \frac{1}{3} \tag{13.141}$$

$$= \frac{3n\theta + 1}{3(n+1)} \neq \theta \tag{13.142}$$

Therefore, fourth option is incorrect.

(a)

$$E(\left(\frac{n+1}{n}\right)\left(Y - \frac{1}{3}\right) + \frac{1}{3}) = \frac{n+1}{n} E(Y) - \frac{n+1}{3n} + \frac{1}{3} \tag{13.143}$$

$$= \frac{3n\theta + 1 - (n+1) + n}{3n} \tag{13.144}$$

$$= \theta \tag{13.145}$$

It is the unbiased estimator of θ.

(b) Similarly,

$$E\left(\left(\frac{n}{n+1}\right)\left(Y-\frac{1}{3}\right)+\frac{1}{3}\right)=\frac{n}{n+1}E(Y)-\frac{n}{3(n+1)}+\frac{1}{3} \tag{13.146}$$

$$=\frac{n(3n\theta+1)-n(n+1)+(n+1)^2}{3(n+1)^2} \tag{13.147}$$

$$=\frac{3n^2+n}{3(n+1)^2}\neq\theta \tag{13.148}$$

It is not an unbiased estimator of θ.

(c) In the same way,

$$E\left(\left(\frac{n+1}{n}\right)\left(Y+\frac{1}{3}\right)-\frac{1}{3}\right)=\frac{n+1}{n}E(Y)+\frac{n+1}{3n}-\frac{1}{3} \tag{13.149}$$

$$=\frac{3n\theta+1+n+1-n}{3n} \tag{13.150}$$

$$=\frac{3n\theta+2}{3n}\neq\theta \tag{13.151}$$

It is not an unbiased estimator of θ.

Hence, the first option is the correct option.

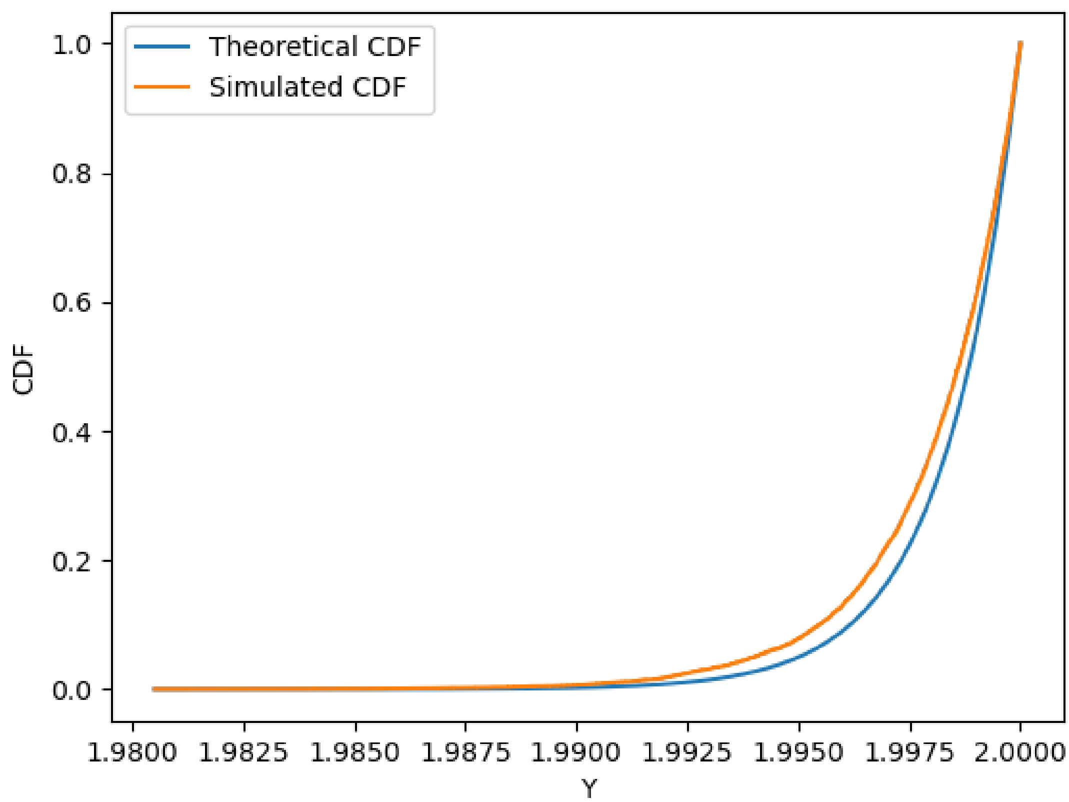


Figure 13.2: Theoretical and Simulated CDFs

Steps for the Simulation using C:

(a) Include Header Files: Start the C program by includung 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 θ.

(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, generate 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 pyhon 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 $X_1, X_2, \ldots, X_n$ be a random sample of size n from a population having probability density function

$$p_X(x;\mu) = \begin{cases} e^{-(x-\mu)}, & \text{if } \mu \leq x < \infty \\ 0, & \text{otherwise,} \end{cases} \tag{13.152}$$

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) $\Pr\left(\hat{M} \le 2\right) = 1 - e^{-n(1-\log_e 2)}$ if $\mu = 1$

B) $\Pr\left(\hat{M} \le 1\right) = 1 - e^{-n\log_e 2}$ if $\mu = 1$

C) $\Pr\left(\hat{M} \le 3\right) = 1 - e^{-n(1-\log_e 2)}$ if $\mu = 1$

D) $\Pr\left(\hat{M} \le 4\right) = 1 - e^{-n(2\log_e 2-1)}$ if $\mu = 1$

(GATE ST 2023)

Solution: For continuous random variable X, median M is such that,

$$\Pr(X \le M) = 0.5 \tag{13.153}$$

The pdf of X is given by,

$$p_X(x) = \begin{cases} e^{-(x-\mu)}, & \text{if } \mu \le x < \infty \\ 0, & \text{otherwise,} \end{cases} \tag{13.154}$$

Hence, cdf is given by

$$F_X(x;\mu) = \int_{\mu}^{x} e^{-(t-\mu)} dt \tag{13.155}$$

$$= e^{\mu}[-e^{-x} + e^{-\mu}] \tag{13.156}$$

$$= 1 - e^{-(x-\mu)} \tag{13.157}$$

Now,

$$F_X(x;\mu) = 0.5 \tag{13.158}$$

$$\implies 1 - e^{-(M-\mu)} = 0.5 \tag{13.159}$$

$$\implies \hat{M} = \mu + \ln(2) \tag{13.160}$$

Definition 13.7: L, the Maximum Likelihood Estimator of the distribution is given by,

$$L = \prod e^{-(x-\mu)} \tag{13.161}$$

$$= e^{-(\sum x_i - n\mu)} \tag{13.162}$$

$$\tag{13.163}$$

For the Likelihood function to be maximum, $\sum x_i - n\mu$ should be minimum Hence,

$$X_i > \mu \tag{13.164}$$

$$\implies \sum x_i > n\mu \tag{13.165}$$

$$\sum x_i - n\mu > 0 \tag{13.166}$$

$$\implies \mu = \frac{\sum x_i}{n} \tag{13.167}$$

Given,

$$p_X(x) = e^{-(x-\mu)} \tag{13.168}$$

$$\tag{13.169}$$

X_i follows an exponential distribution.

Definition 13.8: We know that if,

$$p_X(x) = \lambda_i e^{-\lambda_i x} \tag{13.170}$$

$$S = X_1 + X_2 + ... + X_n \tag{13.171}$$

$$p_S(n) = \frac{\lambda^n x^{n-1} e^{-\lambda x}}{(n-1)!} \tag{13.172}$$

which is gamma distribution with parameters n and λ

Hence, pdf of

$$Y = \sum_{i=1}^{n} x_i \tag{13.173}$$

$$\text{where,} \tag{13.174}$$

$$p_X(x) = e^{-(x-\mu)} \tag{13.175}$$

will follow gamma distribution with parameter n and $\lambda = 1$, given by,

$$p_Y(x; n, 1) = \frac{x^{n-1} e^{-x}}{(n-1)!} \tag{13.176}$$

Hence, cdf is given by,

$$F_Y(x; n) = \int_1^x \frac{t^{n-1} e^{-t}}{(n-1)!} \, dt \tag{13.177}$$

$$= 1 - \Gamma(n, x) \tag{13.178}$$

where $\Gamma(n,x)$ is incomplete gamma function. Thus,

$$\Pr\left(\hat{M} \leq k\right) = \Pr\left(\frac{\sum x_i}{n} + \ln(2) \leq k\right) \tag{13.179}$$

$$= \Pr\left(Y/n + \ln(2) \leq k\right) \tag{13.180}$$

$$= \Pr\left(Y \leq n(k - \ln(2))\right) \tag{13.181}$$

$$= F_Y(n(k - \ln(2))) \tag{13.182}$$

$$= 1 - \Gamma(n, n(k - ln2)) \tag{13.183}$$

It needs a value of n to be computed.

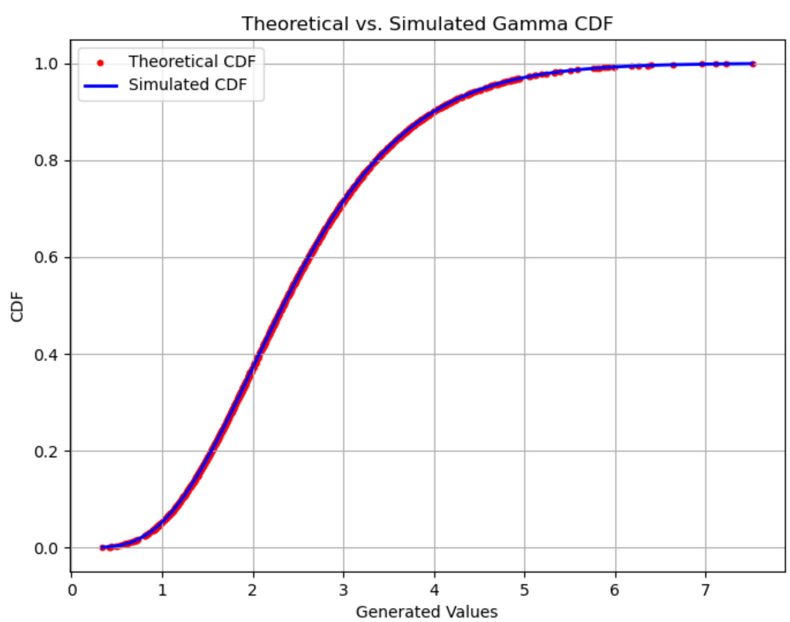


Figure 13.3: Verifying gamma cdf through simulation

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.

Chapter 14

Multivariate Gaussian

14.1 Suppose that $(\mathbf{X_1}, \mathbf{X_2}, \mathbf{X_3})$ has $N_3(\mu, \mathbf{\Sigma})$ distrubution with $\mu = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \Sigma = \begin{bmatrix} 2 & 2 & 1 \\ 2 & 5 & 1 \\ 1 & 1 & 1 \end{bmatrix}$

Given that $\phi(-0.5)$=0.3085, where $\phi(.)$ 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:

Let $\mathbf{Y} = \mathbf{X_1} - \mathbf{2X_2} + \mathbf{2X_3}$ and $\mathbf{a}= \begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix}$

$$\mu_Y = \mathbf{a}^T \mu \tag{14.1}$$

$$= \begin{bmatrix} 1 & -2 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \tag{14.2}$$

$$= 0 \tag{14.3}$$

$$\sigma_Y^2 = \mathbf{a}^T \mathbf{\Sigma a} \tag{14.4}$$

$$= \begin{bmatrix} 1 & -2 & 2 \end{bmatrix} \begin{bmatrix} 2 & 2 & 1 \\ 2 & 5 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix} \tag{14.5}$$

$$= 1 \tag{14.6}$$

$$\Pr\left(Y^2 < \frac{7}{2}\right) = \Pr\left(-\sqrt{\frac{7}{2}} < Y < \sqrt{\frac{7}{2}}\right) \tag{14.7}$$

$$\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz \tag{14.8}$$

$$\Phi\left(\sqrt{\tfrac{7}{2}}\right) = \int_{-\infty}^{\sqrt{\frac{7}{2}}} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz \tag{14.9}$$

$$= \int_{-\infty}^{-\frac{1}{2}} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} + \int_{-\frac{1}{2}}^{\sqrt{\frac{7}{2}}} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz \tag{14.10}$$

$$= \Phi(-0.5) + \int_{-\frac{1}{2}}^{\sqrt{\frac{7}{2}}} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz \tag{14.11}$$

$$= 0.3085 + 0.66082 \tag{14.12}$$

$$= 0.96932 \tag{14.13}$$

$$\Phi\left(-\sqrt{\frac{7}{2}}\right) = \int_{-\infty}^{-\sqrt{\frac{7}{2}}} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz \tag{14.14}$$

$$= 0.03068 \tag{14.15}$$

$$\Pr\left(\left(X_1 - 2X_2 + 2X_3\right)^2 < \frac{7}{2}\right) = \Phi\left(\sqrt{\frac{7}{2}}\right) - \Phi\left(-\sqrt{\frac{7}{2}}\right) \tag{14.16}$$

$$= 0.96932 - 0.03068 \tag{14.17}$$

$$= 0.93864 \tag{14.18}$$

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 Σ.

(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 $E\left({W_4}^2 \middle| W_2 = 2\right)$ in integer equals (GATE ST 2023)

Solution:

Parameter	Description
μ_x	Mean of x
$Var\left(x\right)$	Variance of x
$Cov\left(x,y\right)$	Covariance between x and y
σ_x	Standard deviation of x
ρ	Co-Relation coefficiant
$E\left(x\right)$	Expectation of x

In standard brownian motion,

$$W_i \sim N\left(0,i\right) \tag{14.19}$$

$$Cov\left(W_i,W_j\right) = \min\left(i,j\right) \tag{14.20}$$

Now, we know that,

$$E\left(Y^2\middle| X\right) = Var\left(Y\middle| X\right) + \left(E\left(Y\middle| X\right)\right)^2 \tag{14.21}$$

X and Y can be represented as:

$$X = \sigma_X Z_1 + \mu_X \tag{14.22}$$

$$Y = \sigma_Y\left(\rho Z_1 + \sqrt{1-\rho^2}Z_2\right) + \mu_Y \tag{14.23}$$

where Z_1 and Z_2 are normal distributions.

$$Z_1, Z_2 \sim N\left(0,1\right) \tag{14.24}$$

Writing the above equations in matrix form,

$$\begin{pmatrix} X \\ Y \end{pmatrix} = \begin{bmatrix} \sigma_X & 0 \\ \sigma_Y \rho & \sigma_Y \sqrt{1-\rho^2} \end{bmatrix} \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} + \begin{pmatrix} \mu_X \\ \mu_Y \end{pmatrix} \tag{14.25}$$

This can be represented as,

$$\mathbf{x} = A\mathbf{z} + \boldsymbol{\mu} \tag{14.26}$$

Taking expectation both sides,

$$E(\mathbf{x}) = E(A\mathbf{z} + \boldsymbol{\mu}) \tag{14.27}$$

$$= AE(\mathbf{z}) + E(\boldsymbol{\mu}) \tag{14.28}$$

$$= \boldsymbol{\mu} \tag{14.29}$$

We know that covariance matrix for X and Y is given by:

$$\sigma_{\mathbf{z}} = E\left((\mathbf{x} - \mu)(\mathbf{x} - \mu)^T\right) \tag{14.30}$$

$$= E\left((A\mathbf{z})(A\mathbf{z})^T\right) \tag{14.31}$$

$$= E\left(A\mathbf{z}\mathbf{z}^T A^T\right) \tag{14.32}$$

$$= AE\left(\mathbf{z}\mathbf{z}^T\right) A^T \tag{14.33}$$

Multiplying $\mathbf{z}$ and $\mathbf{z}^T$ we get,

$$\mathbf{z}\mathbf{z}^T = \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} \begin{pmatrix} Z_1 & Z_2 \end{pmatrix} \tag{14.34}$$

$$= \begin{pmatrix} {Z_1}^2 & Z_1 Z_2 \\ Z_1 Z_2 & {Z_2}^2 \end{pmatrix} \tag{14.35}$$

We know that,

$$Var\left(Z_1\right) = E\left(\left(Z_1 - \mu\right)^2\right) \tag{14.36}$$

$$E\left(Z_1^2\right) = 1 \tag{14.37}$$

Same goes for Z_2 as Z_1 and Z_2 are both normal distributions.

Taking expectation both sides in equation (14.35),

$$E\left(\mathbf{z}\mathbf{z}^T\right) = \begin{pmatrix} E\left(Z_1^2\right) & E\left(Z_1 Z_2\right) \\ E\left(Z_1 Z_2\right) & E\left(Z_2^2\right) \end{pmatrix} \tag{14.38}$$

$$= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{14.39}$$

Hence,

$$\sigma_{\mathbf{z}} = AA^T \tag{14.40}$$

$$= \begin{bmatrix} \sigma_X & 0 \\ \sigma_Y \rho & \sigma_Y \sqrt{1-\rho^2} \end{bmatrix} \begin{bmatrix} \sigma_X & \sigma_Y \rho \\ 0 & \sigma_Y \sqrt{1-\rho^2} \end{bmatrix} \tag{14.41}$$

$$= \begin{bmatrix} (\sigma_X)^2 & \sigma_X \sigma_Y \rho \\ \sigma_X \sigma_Y \rho & (\sigma_Y)^2 \end{bmatrix} \tag{14.42}$$

The Co-Relation Coefficiant is given by:

$$\rho = \frac{Cov\,(X, Y)}{\sqrt{Var\,(X)\,Var\,(Y)}} \tag{14.43}$$

Substituting value of Z_1 in Y,

$$Y = \sigma_Y \rho \left(\frac{X - \mu_X}{\sigma_X} \right) + \sigma_Y \sqrt{1-\rho^2} Z_2 + \mu_Y \tag{14.44}$$

This an equation of Y in terms of X. All the terms except Z_2 are constants. Taking expectation on both sides,

$$E\,(Y|\,X = x) = E \left(\sigma_Y \rho \left(\frac{x - \mu_X}{\sigma_X} \right) + \sigma_Y \sqrt{1-\rho^2} Z_2 + \mu_Y \right) \tag{14.45}$$

$$= E \left(\sigma_Y \rho \left(\frac{x - \mu_X}{\sigma_X} \right) + \mu_Y \right) + E \left(\sigma_Y \sqrt{1-\rho^2} Z_2 \right) \tag{14.46}$$

$$= \mu_Y + \rho \left(\frac{\sigma_Y}{\sigma_X} \right) (x - \mu_X) + \sigma_Y \sqrt{1-\rho^2} E\,(Z_2) \tag{14.47}$$

Now for variance,

$$Var\left(Y|\,X=x\right)=Var\left(\sigma_Y\rho\left(\frac{x-\mu_X}{\sigma_X}\right)+\sigma_Y\sqrt{1-\rho^2}Z_2+\mu_Y\right) \tag{14.48}$$

$$=Var\left(\sigma_Y\rho\left(\frac{x-\mu_X}{\sigma_X}\right)+\mu_Y\right)+Var\left(\sigma_Y\sqrt{1-\rho^2}Z_2\right) \tag{14.49}$$

Variance of constants terms is 0.

$$Var\left(Y|\,X=x\right)=Var\left(\sigma_Y\sqrt{1-\rho^2}Z_2\right) \tag{14.50}$$

$$=\left(1-\rho^2\right){\sigma_Y}^2Var\left(Z_2\right) \tag{14.51}$$

Z_2 is a normal distribution so,

$$E\left(Z_2\right)=0 \tag{14.52}$$

$$Var\left(Z_2\right)=1 \tag{14.53}$$

By substituting these values in above equations,

$$E\left(Y|\,X=x\right)=\mu_Y+\rho\left(\frac{\sigma_Y}{\sigma_X}\right)\left(x-\mu_X\right) \tag{14.54}$$

$$Var\left(Y|\,X=x\right)=\left(1-\rho^2\right){\sigma_Y}^2 \tag{14.55}$$

In our case,

$$Y=W_4 \tag{14.56}$$

$$X=W_2 \tag{14.57}$$

$$x=2 \tag{14.58}$$

Hence, we get that,

$$\mu_X = \mu_Y = 0 \tag{14.59}$$

$$\sigma_X = \sqrt{2} \tag{14.60}$$

$$\sigma_Y = 2 \tag{14.61}$$

$$\rho = \frac{2}{\sqrt{8}} \tag{14.62}$$

$$= \frac{1}{\sqrt{2}} \tag{14.63}$$

Substituting the values in above equations,

$$E\left(Y \middle| X = 2\right) = \frac{1}{\sqrt{2}} \cdot \frac{2}{\sqrt{2}} \cdot 2 \tag{14.64}$$

$$= 2 \tag{14.65}$$

$$Var\left(Y \middle| X = 2\right) = \left(1 - \frac{1}{2}\right)(2)^2 \tag{14.66}$$

$$= \frac{1}{2} \cdot 4 \tag{14.67}$$

$$= 2 \tag{14.68}$$

Substituting these values in (14.21),

$$E\left(Y^2 \middle| X = 2\right) = 2 + (2)^2 \tag{14.69}$$

$$= 6 \tag{14.70}$$

$$E\left({W_4}^2 \middle| W_2 = 2\right) = 6 \tag{14.71}$$

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*numsamples

(g) Make a temporary 2D array(dist) of size 2x10000 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,

$$\sigma_X = 0.5 \tag{14.72}$$

$$\sigma_Y = 0.8 \tag{14.73}$$

$$\rho = 0.5 \tag{14.74}$$

$$\mu_X = 1 \tag{14.75}$$

$$\mu_Y = 1.5 \tag{14.76}$$

(k) Create A matrix and vector $\boldsymbol{\mu}$ and fill the values assumed above in these, the dimension for matrix is 2x2 and the vector 2x1

(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.