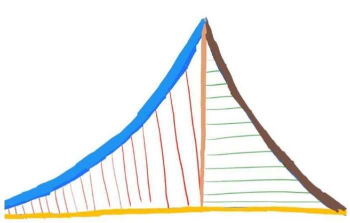
GATE PROBABILITY

Through Simulations

G. V. V. Sharma



Contents

Introduction

This book solves probability problems in GATE question papers.

Chapter 1

Axioms

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);

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

From (2.3);

(a)

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

$$= 0.5 \tag{2.5}$$

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

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

From (2.6);

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

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

$$= 0.341 \tag{2.9}$$

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

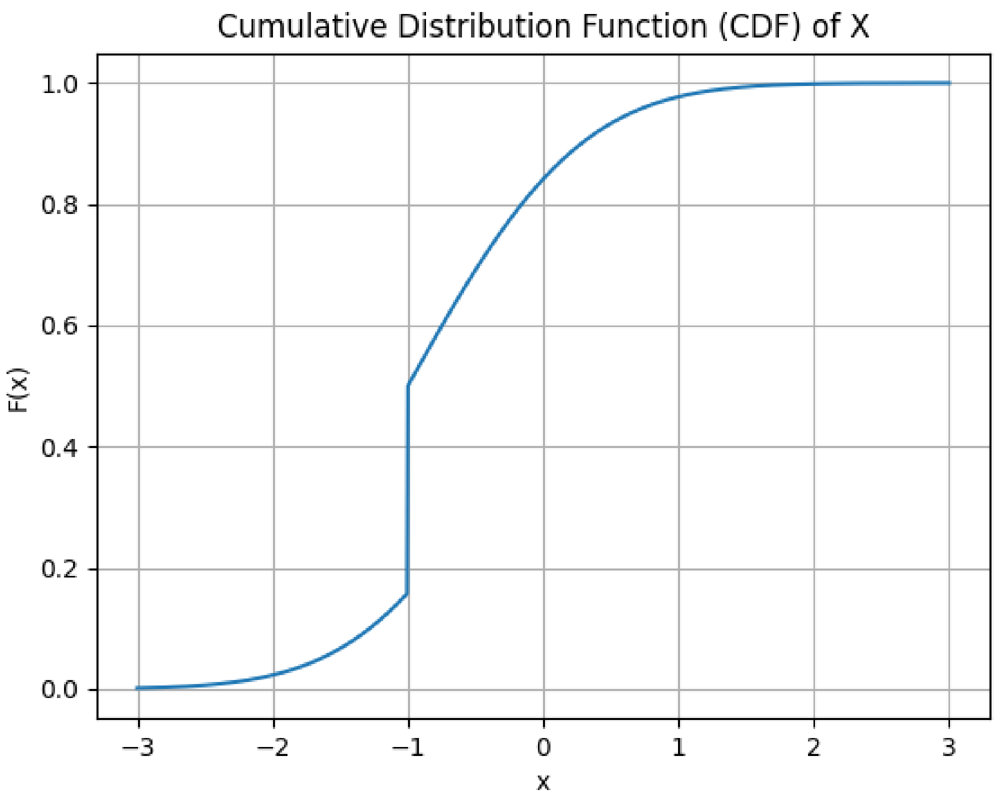


Figure 2.1:

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

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

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)

$$\begin{aligned} \implies \sigma_X^2 &= \mathbb{E}\left(X^2\right) && (2.19)\\ &= \frac{1}{2\sqrt{3}} \int_{-\sqrt{3}}^{\sqrt{3}} x^2 dx && (2.20)\\ &= 1 && (2.21) \end{aligned}$$

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

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

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:

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

(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)

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

$\therefore (D)$ 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.

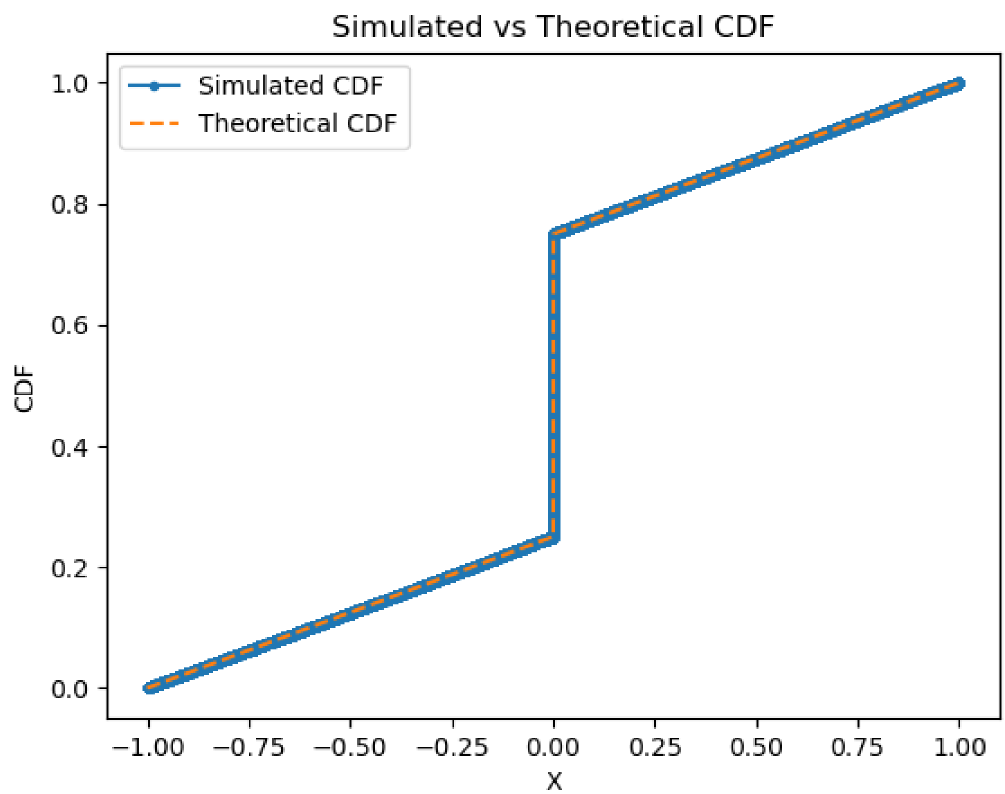


Figure 2.2: CDF of X-(simulation vs actual)

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.

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

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

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

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

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

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

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

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

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

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

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

$\therefore$ 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 $a < y < b$

$$
\begin{aligned}
F_Y(y) &= \Pr\left(1 - e^{-X} \le y\right) && (2.74)\\
&= \Pr\left(e^{-X} \ge 1 - y\right) && (2.75)\\
&= \Pr\left(X \le -\ln(1-y)\right) && (2.76)\\
&= F_X\left(-\ln(1-y)\right) && (2.77)\\
&= 1 - (1-y) && (2.78)\\
&= y && (2.79)\\
\implies Y &\sim U(0,1) && (2.80)
\end{aligned}
$$

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

for $a < y < b$

$$
\begin{aligned}
F_Y(y) &= \Pr\left(2e^{-X} - 1 \le y\right) && (2.81)\\
&= \Pr\left(X \ge -\ln\left(\frac{y+1}{2}\right)\right) && (2.82)\\
&= 1 - F_X\left(-\ln\left(\frac{y+1}{2}\right)\right) \qquad = 1 - \left(1 - \frac{y+1}{2}\right) && (2.83)\\
&= \frac{y+1}{2} && (2.84)
\end{aligned}
$$

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.

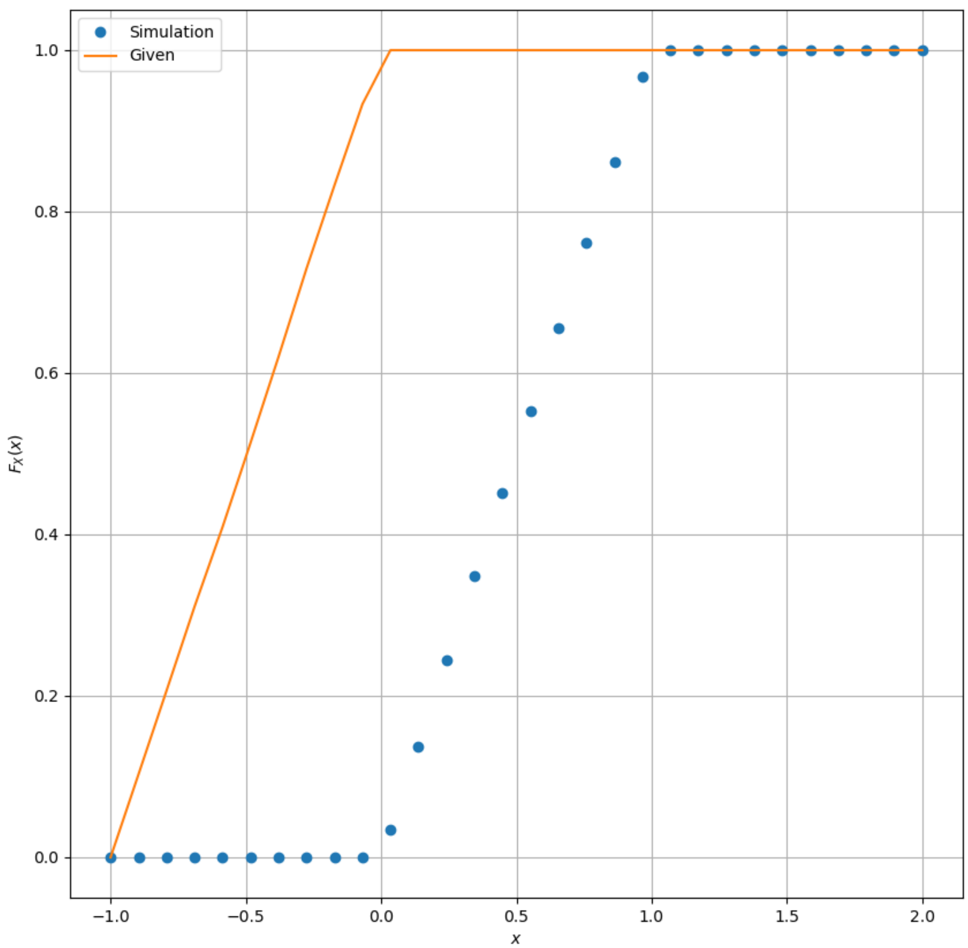


Figure 2.3: e^{-X} vs. $U(-1, 0)$
Graphs don't match, $\therefore$ wrong option

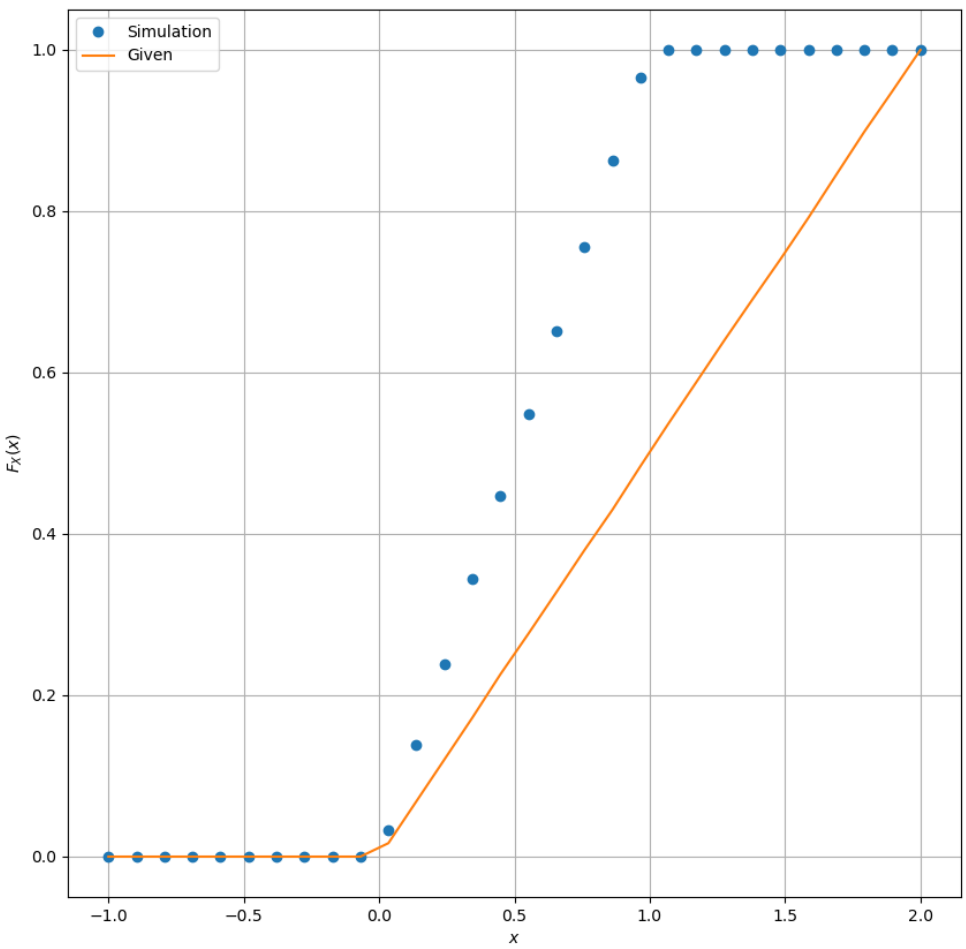


Figure 2.4: $1 - e^{-X}$ vs. $U\left(0, 2\right)$
Graphs don't match, $\therefore$ wrong option

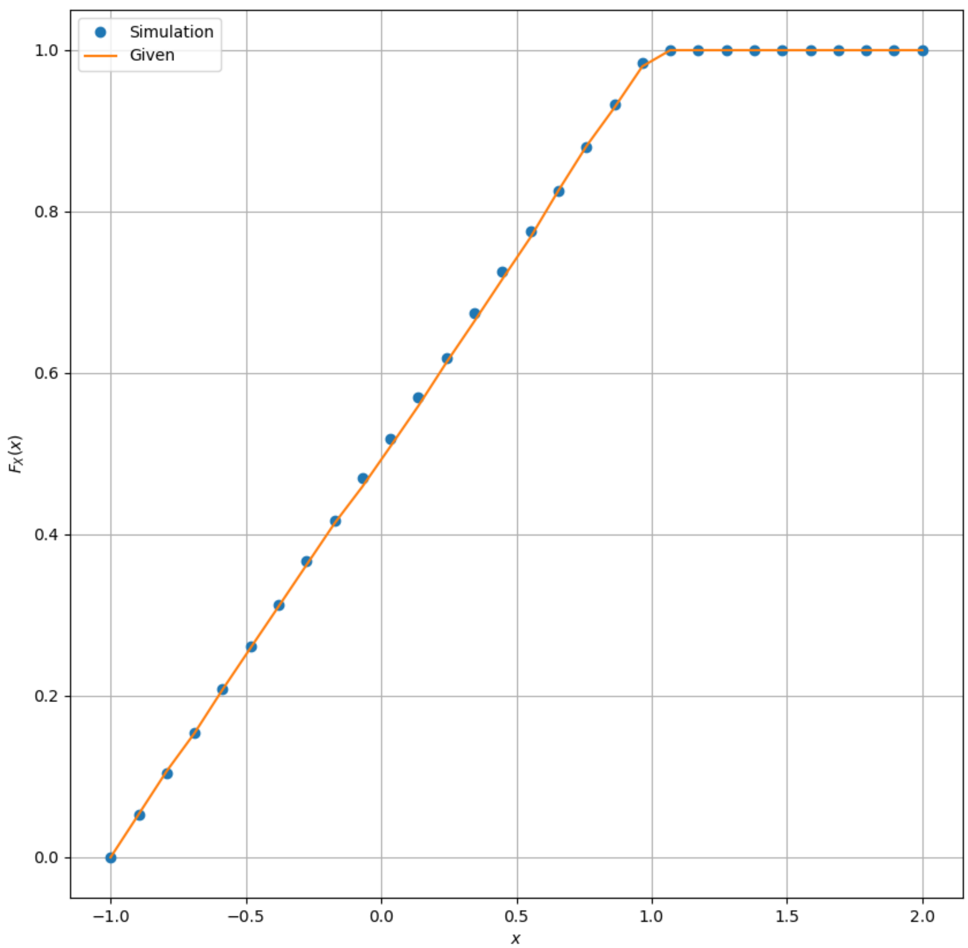


Figure 2.5: $2e^{-X} - 1$ vs. $U(-1, 1)$
Graphs match, $\therefore$ correct option

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

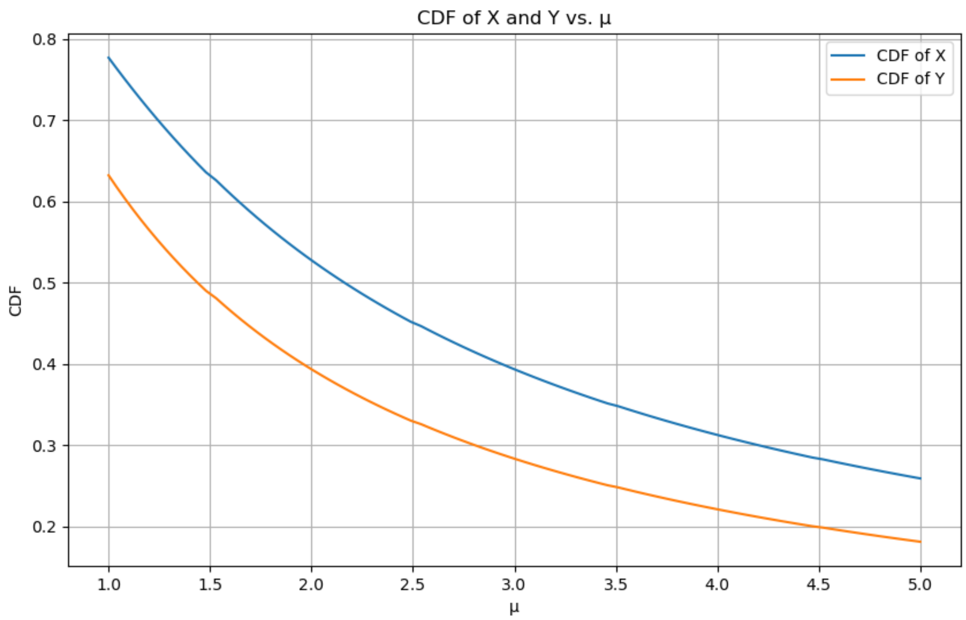


Figure 2.6: CDF'S of X and Y for varying μ at x=1.5

(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,

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

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

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

RV	Values	Description
X	0	11
	1	10
	2	00

Table 2.7: Random varibale declaration

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

Table 2.8: Binomial parameters declaration

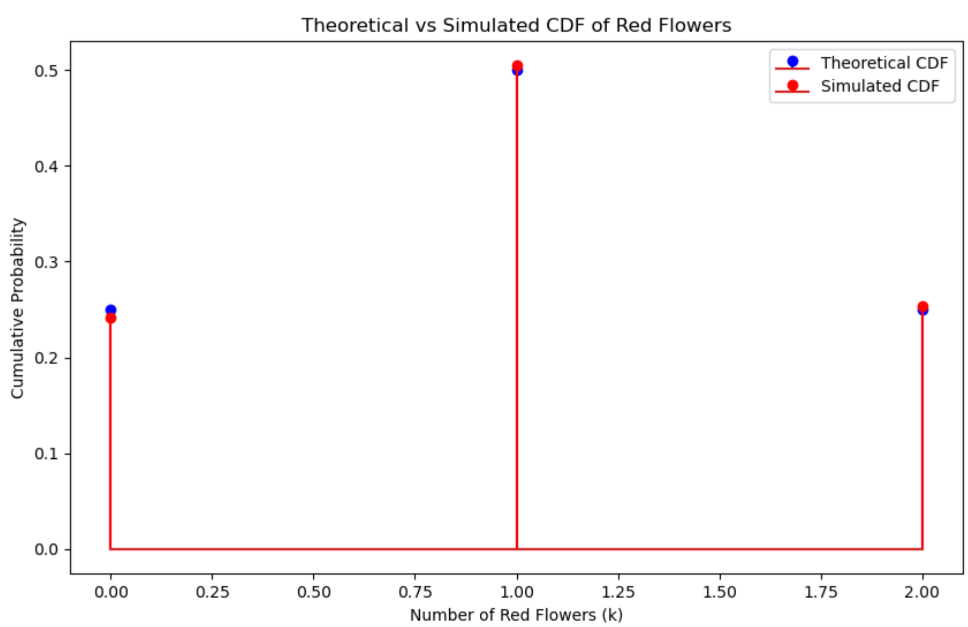


Figure 2.7: Simulation vs Theoretical

Chapter 3

Conditional Probability

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%

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)

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

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

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

$$\tag{5.5}$$

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 $\lambda > 0$.

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 $0 < y < 1$

$$\begin{aligned} &= \int_0^y 8xydx && (6.5)\\ &= 8y\left[\frac{x^2}{2}\right]_0^y && (6.6)\\ &= 4y^3 && (6.7) \end{aligned}$$

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

$$\begin{aligned} p_{X|Y} &= \frac{8xy}{4y^3} && (6.8)\\ &= \frac{2x}{y^2} && (6.9) \end{aligned}$$

and

$$\begin{aligned}
E\left(X|Y=y_0\right) &= \int_0^{y_0} x.\frac{2x}{y_0^2}dx && (6.10)\\
&= \frac{2}{y_0^2}\left[\frac{x^3}{3}\right]_0^{y_0} && (6.11)\\
&= \frac{2y_0}{3} && (6.12)\\
\implies \frac{2y_0}{3} &= \frac{1}{2} && (6.13)\\
y_0 &= \frac{3}{4} && (6.14)
\end{aligned}$$

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

$\Pr(X_i \leq \theta) = 0.5$ 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

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

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

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

(B) $\frac{1}{n-1}\sum_{i=1}^{n}(\ln X_i)^2$ 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:

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

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:

$$\begin{aligned} L(\theta_0 \mid data) &= \prod_{i=1}^{n} p_{X_i,Y_i}(x,y) && (7.32) \\ &= (2\theta_0)^m \, (1-2\theta_0)^{n-m} && (7.33) \\ &= \left(\frac{1}{2}\right)^m \left(\frac{1}{2}\right)^{n-m} && (7.34) \end{aligned}$$

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

$$\begin{aligned} L(\theta_1 \mid data) &= \prod_{i=1}^{n} p_{X_i,Y_i}(x,y) && (7.35) \\ &= (2\theta_1)^m \, (1-2\theta_1)^{n-m} && (7.36) \\ &= \left(\frac{2}{3}\right)^m \left(\frac{1}{3}\right)^{n-m} && (7.37) \end{aligned}$$

The likelyhood ratio will be

$$\begin{aligned} \lambda(data) &= \frac{L(\theta_1 \mid x)}{L(\theta_0 \mid x)} && (7.38) \\ &= \frac{\left(\frac{2}{3}\right)^m \left(\frac{1}{3}\right)^{n-m}}{\left(\frac{1}{2}\right)^m \left(\frac{1}{2}\right)^{n-m}} = (2)^m \left(\frac{2}{3}\right)^n && (7.39) \end{aligned}$$

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

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

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

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

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

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

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

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

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

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

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

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 $\{X_n\}_{n\geq1}$ and Let $\{Y_n\}_{n\geq1}$ 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 $\{X_n\}_{n\geq1}$ converges in distribution to a real constant c, then $\{X_n\}_{n\geq1}$ converges in probability to c.

(B) If $\{X_n\}_{n\geq1}$ converges in probability to X, then $\{X_n\}_{n\geq1}$ converges in 3^{rd} mean to X.

(C) If $\{X_n\}_{n\geq1}$ converges in distribution to X and $\{Y_n\}_{n\geq1}$ converges in distribution to Y, then $\{X_n + Y_n\}_{n\geq1}$ converges in distribution to $X + Y$.

(D) If $\{E(X_n)\}_{n\geq1}$ converges to $E(X)$, then $\{X_n\}_{n\geq1}$ 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,

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

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

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

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 $\{X_n\}_{n\geq 1}$ and $\{Y_n\}_{n\geq 1}$ be sequences of random variables such that they both converge in distribution as Z and $(-1)^n Z$. Proof that Y_n converges in distribution.

For n even

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

For n odd

$$lim_{n\to\infty} F_{Y_n}(x) = \Pr(-Z \leq x) \tag{10.24}$$

$$= \Pr(Z \leq x) \tag{10.25}$$

Proved. So,we have

$$F_{X_n+Y_n}(x) = \Pr\left(X_n + Y_n \leq x\right) \tag{10.26}$$

$$= \Pr\left(Z + (-1)^n Z \leq x\right) \tag{10.27}$$

For n is even

$$F_{X_n+Y_n}(x) = \Pr\left(2Z \leq x\right) \tag{10.28}$$

$$= \Pr\left(Z \leq \frac{x}{2}\right) \tag{10.29}$$

$$= 1 - \Pr\left(Z > \frac{x}{2}\right) \tag{10.30}$$

$$\approx 1 - Q\left(\frac{x}{2}\right) \tag{10.31}$$

For n is odd

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

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

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 $\lim\limits_{n\to\infty} \mathrm{E}\left[\left(\frac{Y_n-4}{\sqrt{n}}\right)^2\right]$ (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

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

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

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

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

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

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

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

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,

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

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

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.

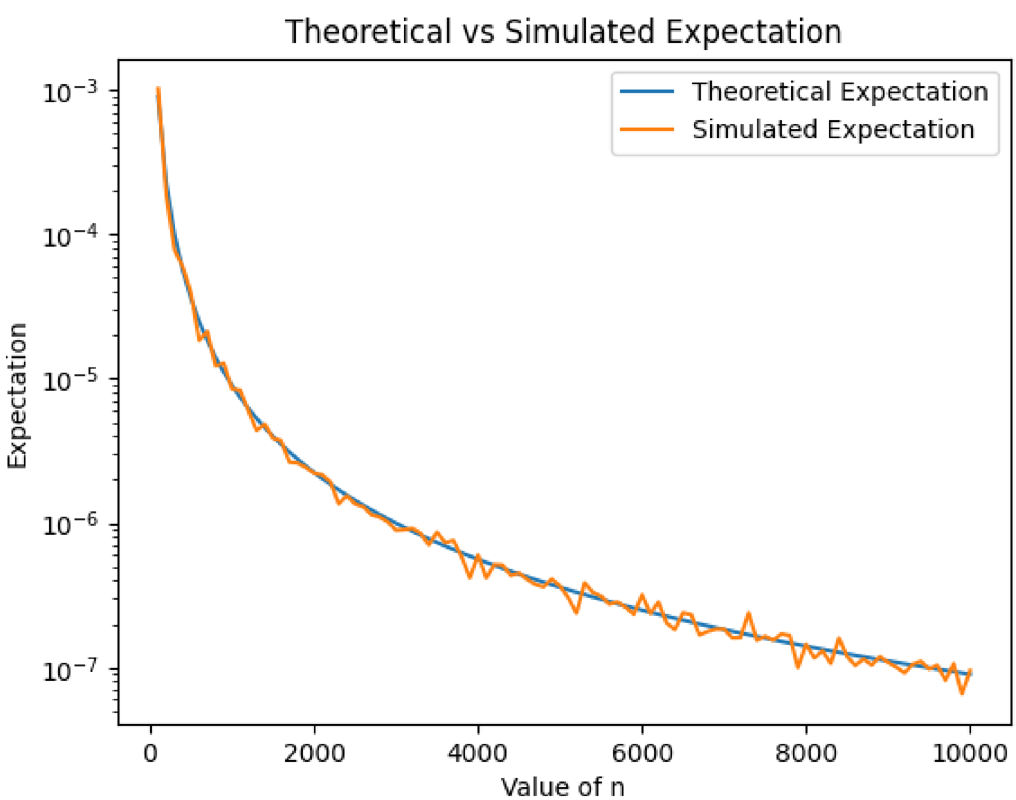


Figure 10.1: Expectation vs n

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:

Parameter	**Value**	**Description**
X	$1 \leq X \leq 8$	number of symbols
l	2	base of algorithm
$H(X)$	$\sum_i p_X(i) \log_l \left(\frac{1}{p_X(i)}\right)$	average number of question

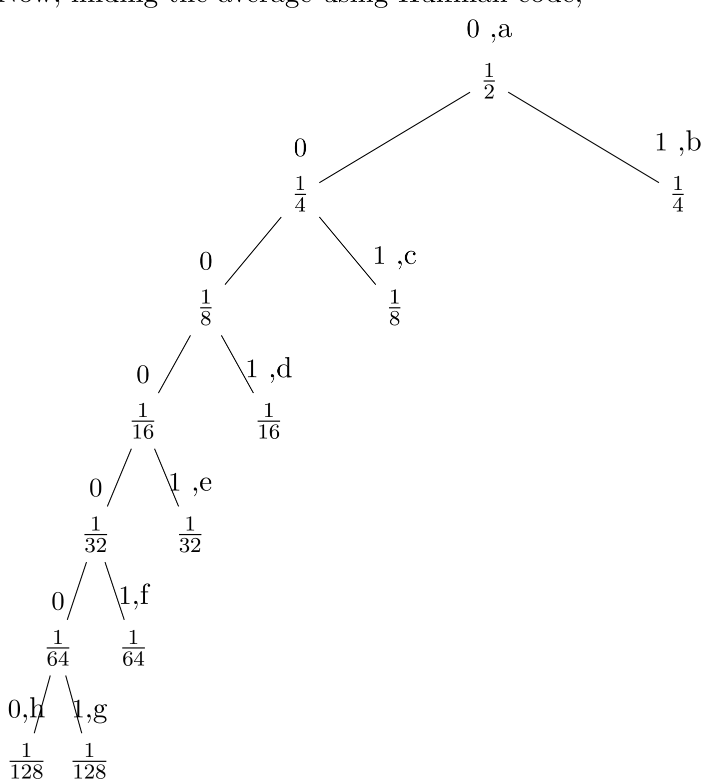
$$H(X) = \sum_i p_X(i) \log_b \left(\frac{1}{p_X(i)} \right) \tag{11.1}$$

$$= \frac{1}{2} \log_2 (2) + \frac{1}{4} \log_2 (4) + ... + \frac{1}{128} \log_2 (128) \tag{11.2}$$

$$= 0.5 + 0.5 + 0.375 + ... + 0.0078125 \tag{11.3}$$

$$= 1.984375 \tag{11.4}$$

Now, finding the average using Huffman code,



Using the above binary table following code is generated;

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.

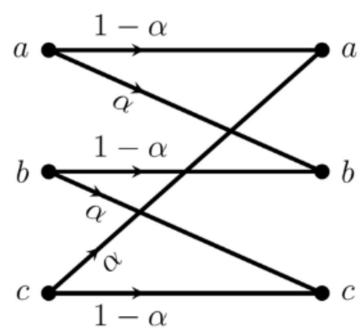
 The parameter α lies in the interval $[0.25, 1]$. The value of α for which the capacity

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

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

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

$$I\left(X, Y\right) = \sum_{x,y} p\left(x, y\right) \log_2 \frac{p\left(x, y\right)}{p\left(x\right) p\left(y\right)} \tag{11.6}$$

$$= \sum_{x,y} p\left(x, y\right) \log_2 \frac{p\left(y \mid x\right)}{p\left(y\right)} \tag{11.7}$$

$$= -\sum_{x,y} p\left(x, y\right) \log_2 p\left(y\right) + \sum_{x,y} p\left(x, y\right) \log_2 p\left(y \mid x\right) \tag{11.8}$$

$$= -\sum_{y} p\left(y\right) \log_2 p\left(y\right) - \left(-\sum_{x,y} p\left(x, y\right) \log_2 p\left(y \mid x\right)\right) \tag{11.9}$$

$$= H\left(Y\right) - H\left(Y \mid X\right) \tag{11.10}$$

Now,

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

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

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

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

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

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

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

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

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

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

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

Solution:

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

(a) For Option(A) we will find

We know that :

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

$\Longrightarrow$

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

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

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

$$H\left(X\right) \leq \log_2 K \tag{11.26}$$

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	$\frac{1}{4}$

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

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

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 $Y = 2^X$

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

$$H\left(Y\right) = \sum_{i=0}^{2} p_Y\left(k\right) \log_2 \frac{1}{p_Y\left(k\right)} \tag{11.35}$$

$$H\left(Y\right) = 1.5units \tag{11.36}$$

$$H\left(Y\right) = H\left(2^X\right) = H\left(X\right) \tag{11.37}$$

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 $\Pr\left(X_2 = 1 | X_1 = 1, X_3 = 2\right)$ equals

(GATE ST 2023)

Solution: Consider transition matrix as:

$$\begin{pmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{pmatrix} \tag{12.1}$$

$$\Pr\left(X_2 = 1 | X_1 = 1, X_3 = 2\right) = \Pr\left(X_2 = 1 | X_1 = 1\right) \tag{12.2}$$

$$= p_{11} \tag{12.3}$$

$$= 0.5 \tag{12.4}$$

(by markov's property and using transition probability matrix)

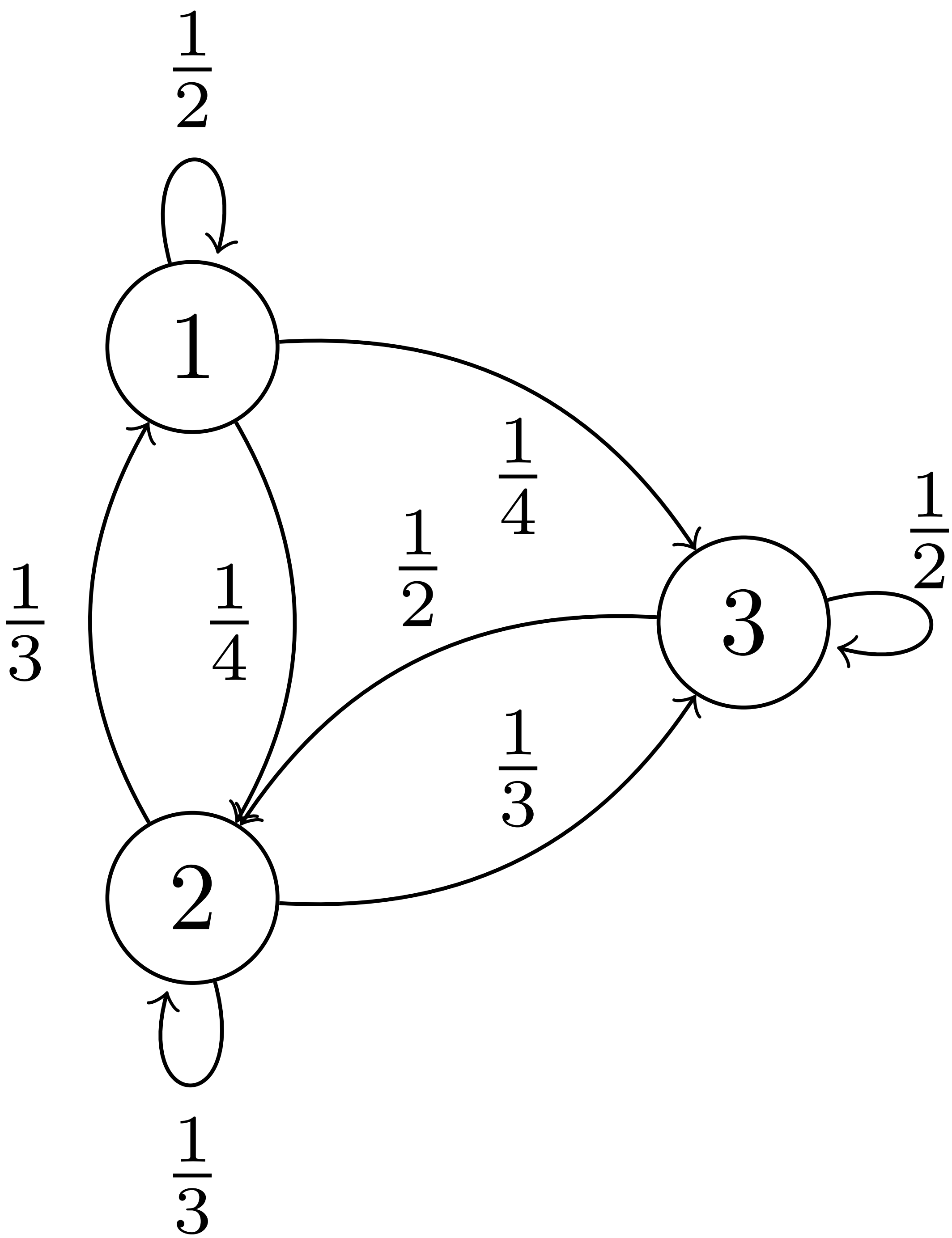


Figure 12.1: Markov Chain diagram

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,

$$\text{Cramér-Rao bound} = \frac{T}{\left(3 - \frac{1}{2\pi}\right)} \tag{13.16}$$

$$= 0.968 \tag{13.17}$$

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(**

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):

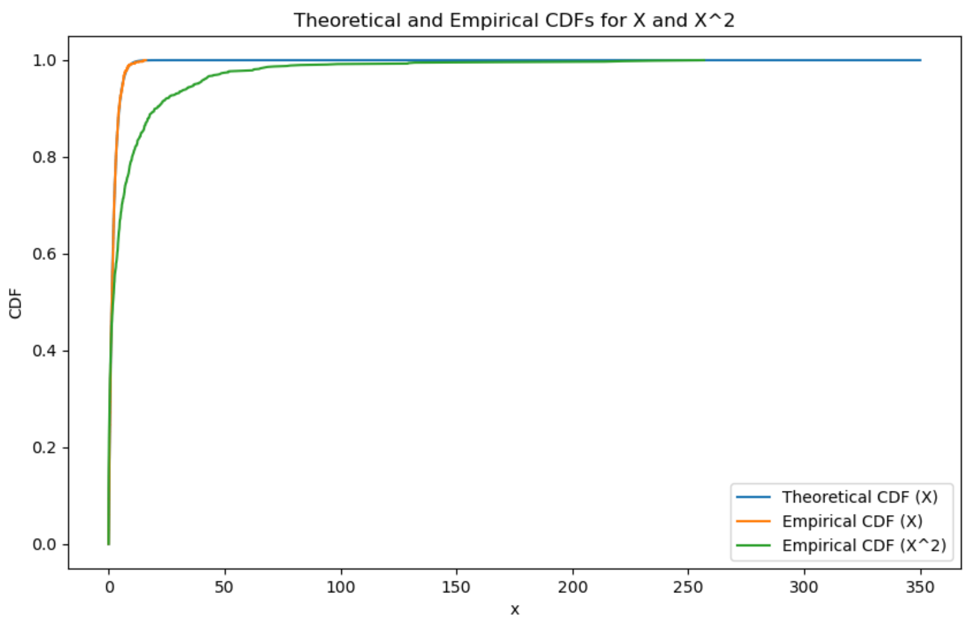


Figure 13.1: Figure1

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.

$$RSS = \sum \left(Y_i - \hat{Y}\right) \tag{13.27}$$

$$= \sum e_i^2 \tag{13.28}$$

$$= \sum (Y_i - \beta_0 - \beta_1 X_i)^2 \tag{13.29}$$

Here $\hat{Y}$ = the value of Y on the line of regression.

Definition 13.5: Total sum of squares(TSS):

Table 13.2: Parameters

Parameters	Description
Y_i	Dependent variable
X_i	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):

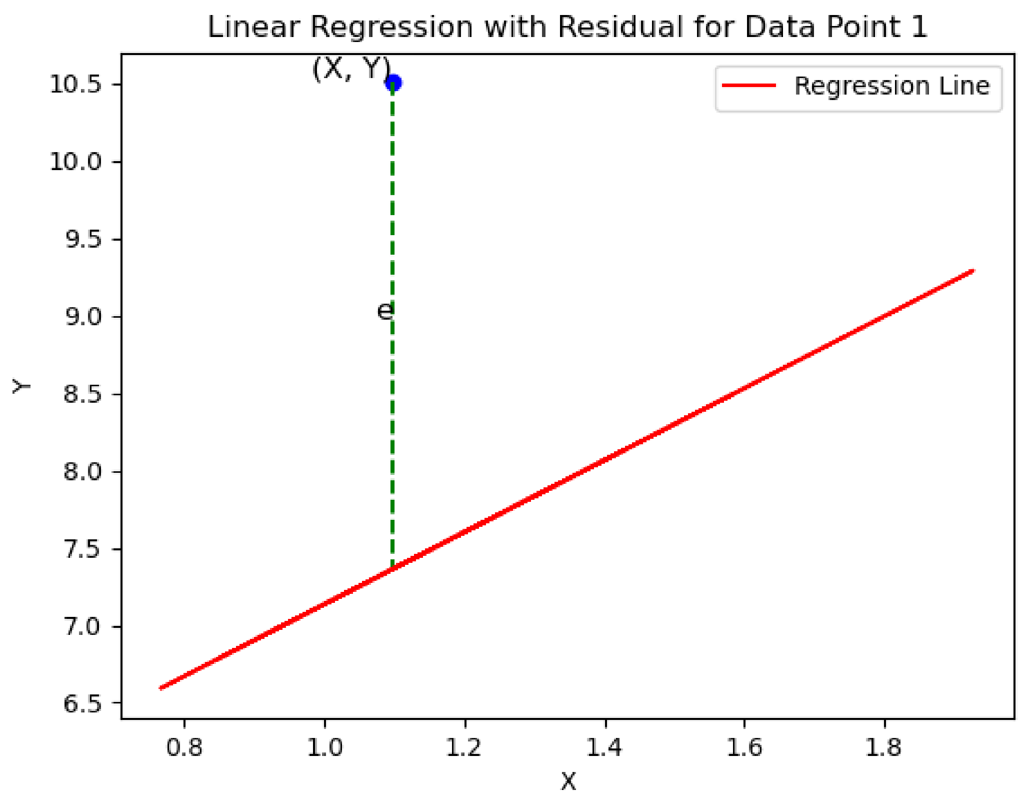
It is the proportion of the variance in the dependent variable that is predicted from the independent variable. It indicates the level of variation in the given data set.

$$R^2 = 1 - \frac{RSS}{TSS} \tag{13.31}$$

$$= 1 - \frac{44}{80} \tag{13.32}$$

$$= 0.45 \tag{13.33}$$

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 $y_k - \alpha_0 - \alpha_1 \log_e k > 0$

$$\begin{aligned} \min_{\alpha_0} \quad & y_k - \alpha_0 - \alpha_1 \log_e k \\ \text{s.t.} \quad & \alpha_0 \leq y_k - \alpha_1 \log_e k \end{aligned}$$

Using Lagrange multiplier method

$$L(\lambda) = y_k - \alpha_0 - \alpha_1 \log_e k - \lambda(\alpha_0 - y_k + \alpha_1 \log_e k) \tag{13.42}$$
$$\frac{\partial L}{\partial \alpha_0} = -1 - \lambda = 0 \tag{13.43}$$
$$\frac{\partial L}{\partial \lambda} = y_k - \alpha_0 - \alpha_1 \log_e k = 0 \tag{13.44}$$
$$\lambda = -1 \tag{13.45}$$
$$\alpha_0 = y_k - \alpha_1 \log_e k \tag{13.46}$$

B. For $y_k - \alpha_0 - \alpha_1 \log_e k < 0$

$$\begin{aligned} \min_{\alpha_0} \quad & -(y_k - \alpha_0 - \alpha_1 \log_e k) \\ \text{s.t.} \quad & \alpha_0 \geq y_k - \alpha_1 \log_e k \end{aligned}$$

Using Lagrange multiplier method

$$L(\lambda) = -(y_k - \alpha_0 - \alpha_1 \log_e k) - \lambda(\alpha_0 - y_k + \alpha_1 \log_e k) \tag{13.47}$$
$$\frac{\partial L}{\partial \alpha_0} = 1 - \lambda = 0 \tag{13.48}$$
$$\frac{\partial L}{\partial \lambda} = y_k - \alpha_0 - \alpha_1 \log_e k = 0 \tag{13.49}$$
$$\lambda = 1 \tag{13.50}$$
$$\alpha_0 = y_k - \alpha_1 \log_e k \tag{13.51}$$

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 $y_k - \alpha_0 - \alpha_1 \log_e k > 0$

$$\begin{aligned} \min_{\alpha_1} \quad & y_k - \alpha_0 - \alpha_1 \log_e k \\ \text{s.t.} \quad & \alpha_1 \leq \frac{y_k - \alpha_0}{\log_e k} \end{aligned}$$

Using Lagrange multiplier method

$$L(\lambda) = y_k - \alpha_0 - \alpha_1 \log_e k - \lambda \left(\alpha_1 - \frac{y_k - \alpha_0}{\log_e k} \right) \tag{13.52}$$

$$\frac{\partial L}{\partial \alpha_1} = -\log_e k - \lambda = 0 \tag{13.53}$$

$$\frac{\partial L}{\partial \lambda} = -\left(\alpha_1 - \frac{y_k - \alpha_0}{\log_e k} \right) = 0 \tag{13.54}$$

$$\lambda = -\log_e k \tag{13.55}$$

$$\alpha_1 = \frac{y_k - \alpha_0}{\log_e k} \tag{13.56}$$

B. For $y_k - \alpha_0 - \alpha_1 \log_e k < 0$

$$\begin{aligned} \min_{\alpha_1} \quad & -\left(y_k - \alpha_0 - \alpha_1 \log_e k\right) \\ \text{s.t.} \quad & \alpha_1 \geq \frac{y_k - \alpha_0}{\log_e k} \end{aligned}$$

Using Lagrange multiplier method

$$L(\lambda) = -(y_k - \alpha_0 - \alpha_1 \log_e k) - \lambda \left(\alpha_1 - \frac{y_k - \alpha_0}{\log_e k}\right) \tag{13.57}$$

$$\frac{\partial L}{\partial \alpha_1} = \log_e k - \lambda = 0 \tag{13.58}$$

$$\frac{\partial L}{\partial \lambda} = -\left(\alpha_1 - \frac{y_k - \alpha_0}{\log_e k}\right) = 0 \tag{13.59}$$

$$\lambda = \log_e k \tag{13.60}$$

$$\alpha_1 = \frac{y_k - \alpha_0}{\log_e k} \tag{13.61}$$

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

$$p_X(x) = \begin{cases} \frac{1}{2}e^{x} & x \le 0 \\ \frac{1}{2}e^{-x} & x > 0 \end{cases} \tag{13.69}$$

$$F_X(x) = \begin{cases} \int_{-\infty}^{x} \left(\frac{1}{2}e^{x}\right) dx & x \le 0 \\ \int_{-\infty}^{0} \left(\frac{1}{2}e^{x}\right) dx + \int_{0}^{x} \left(\frac{1}{2}e^{-x}\right) dx & x > 0 \end{cases} \tag{13.70}$$

$$F_X(x) = \begin{cases} \frac{1}{2}e^x & x \leq 0 \\ \frac{1}{2}\left(2 - e^{-x}\right) & x > 0 \end{cases} \tag{13.71}$$

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

$$u = \frac{1}{2}e^x \tag{13.72}$$

$$e^x = 2u \tag{13.73}$$

$$x = \ln 2u \tag{13.74}$$

$$\because x \leq 0 \tag{13.75}$$

$$u \leq 0.5 \tag{13.76}$$

For x > 0

$$u = \frac{1}{2}\left(2 - e^{-x}\right) \tag{13.77}$$

$$2 - e^{-x} = 2u \tag{13.78}$$

$$e^{-x} = 2 - 2u \tag{13.79}$$

$$x = -\ln\left(2 - 2u\right) \tag{13.80}$$

$$\because x > 0 \tag{13.81}$$

$$u > 0.5 \tag{13.82}$$

$$I(u) = \begin{cases} \ln(2u) & u \le 0.5 \\ -\ln(2-2u) & u > 0.5 \end{cases} \tag{13.83}$$

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