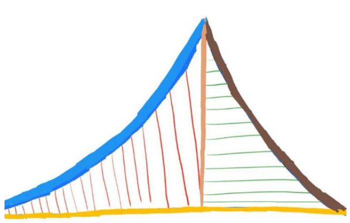
GATE PROBABILITY

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);

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

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

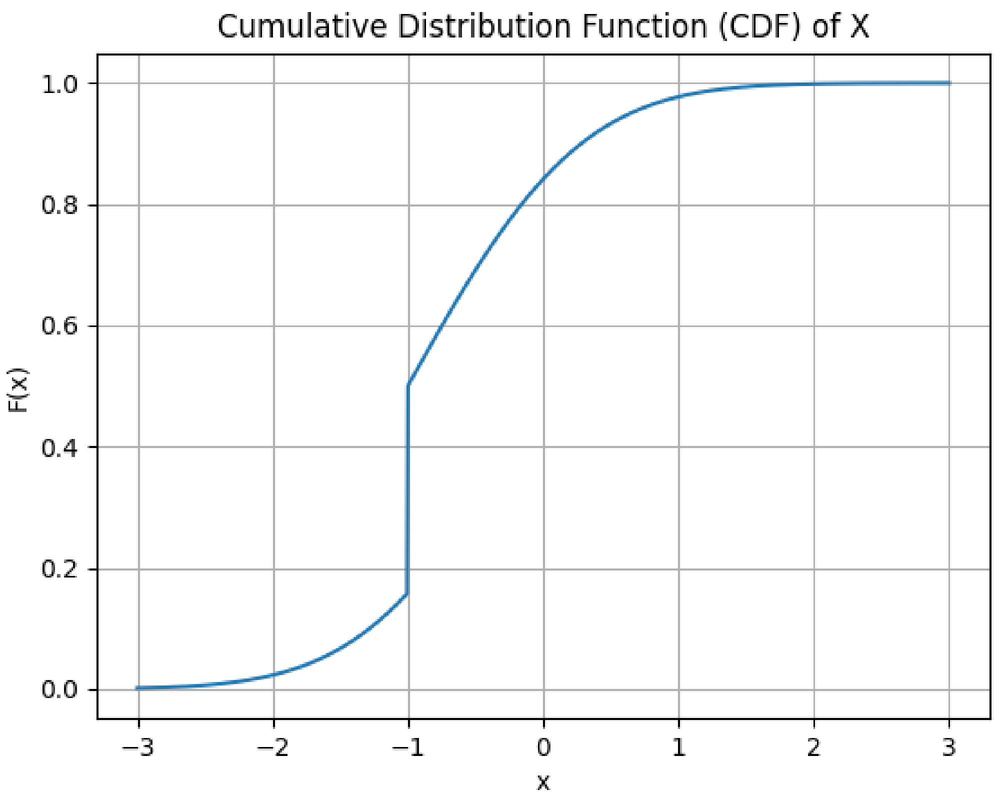


Figure 2.1:

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

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

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)

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

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

$$= 1 \tag{2.21}$$

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

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

Chapter 3

Conditional Probability

Chapter 4

Moments

4.1 Suppose that X has the probability density function

$$f(x) = \begin{cases} \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} & \lambda > 0 \\ 0 & otherwise \end{cases} \tag{4.1}$$

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)

$$E(X) = \int_{-\infty}^{\infty} x p_X(x) dx \tag{4.2}$$
$$= \int_{0}^{\infty} x \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} \tag{4.3}$$
$$= \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_{0}^{\infty} x^\alpha e^{-\lambda x} \tag{4.4}$$
$$\tag{4.5}$$

since we know that

$$\int_{0}^{\infty} x^{\alpha-1} e^{-\lambda x} dx = \frac{\Gamma(\alpha)}{\lambda^\alpha} \qquad \text{for } \lambda > 0, \alpha > 0 \tag{4.6}$$

$$E(X) = \frac{\lambda^\alpha}{\Gamma(\alpha)} \frac{\Gamma(\alpha+1)}{\lambda^{\alpha+1}} \tag{4.7}$$

Using the relation

$$\Gamma(x+1) = \Gamma(x)x \tag{4.8}$$

$$E(X) = \frac{\alpha}{\lambda} \tag{4.9}$$

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

(b)

$$Var(X) = E(X^2) - E(X)^2 \tag{4.10}$$

$$E(X^2) = \int_0^\infty x^2 \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}\, dx \tag{4.11}$$

$$= \int_0^\infty \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{(\alpha+2)-1} e^{-\lambda x}\, dx \tag{4.12}$$

$$= \int_0^\infty \frac{1}{\lambda^2} \frac{\lambda^{\alpha+2}}{\Gamma(\alpha)} x^{(\alpha+2)-1} e^{-\lambda x}\, dx \tag{4.13}$$

$$E(X^2) = \int_0^\infty \frac{\alpha(\alpha+1)}{\lambda^2} \frac{\lambda^{\alpha+2}}{\Gamma(\alpha+2)} x^{(\alpha+2)-1} e^{-\lambda x} dx \tag{4.14}$$

using the density of the gamma distribution, we get

$$E(X^2) = \frac{\alpha(\alpha+1)}{\lambda^2} \tag{4.15}$$

$$Var(X) = \frac{\alpha^2 + \alpha}{\lambda^2} - \frac{\alpha}{\lambda}^2 \tag{4.16}$$

$$= \frac{\alpha}{\lambda^2} \tag{4.17}$$

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

(c)

$$E\left(\frac{1}{X}\right) = \int_0^\infty \frac{1}{x} \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} \tag{4.18}$$

$$= \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^\infty x^{\alpha-2} e^{-\lambda x} \tag{4.19}$$

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

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

$$E(ln(1+X)) = E(X) - \frac{E(X^2)}{2} + \frac{E(X^4)}{4} - .. \tag{4.20}$$

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

$$E(X^n) = \frac{(\alpha)(\alpha+1)\ldots(\alpha+n-1)}{\lambda^n} \tag{4.21}$$

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

$$\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| = L \tag{4.22}$$

$$\left|\frac{E(X^{n+2})}{E(X^n)}\right| = \frac{\frac{(\alpha)(\alpha+1)\ldots(\alpha+n+1)}{\lambda^{n+2}}}{\frac{(\alpha)(\alpha+1)\ldots(\alpha+n-1)}{\lambda^n}} \tag{4.23}$$

$$= \frac{(\alpha+n)(\alpha+n+1)}{\lambda^2} \tag{4.24}$$

$$\lim_{n\to\infty}\left|\frac{E(X^{n+2})}{E(X^n)}\right| = \infty \tag{4.25}$$

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

4.2 Let X be a random variable with probability density function

$$f(x;\lambda) = \begin{cases} \frac{1}{\lambda}e^{-\frac{x}{\lambda}} & \text{if} x > 0 \\ 0 & \text{otherwise} \end{cases} \tag{4.26}$$

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

$$\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i \tag{4.27}$$

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 (4.26) 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,

$$E(Y_i) = E(X_i^2) \tag{4.28}$$

$$= \int_0^\infty x^2 \frac{1}{\lambda} e^{-\frac{x}{\lambda}} \tag{4.29}$$

$$= 2\lambda^2 \tag{4.30}$$

Method of moment is defined by (4.27) which gives,

$$\bar{Y} = E(Y_i) \tag{4.31}$$

$$= 2\lambda^2 \tag{4.32}$$

where

$$\lambda = \sqrt{\frac{\bar{Y}}{2}} \tag{4.33}$$

$\therefore$ Option (4.2a) 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(1,10)**

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

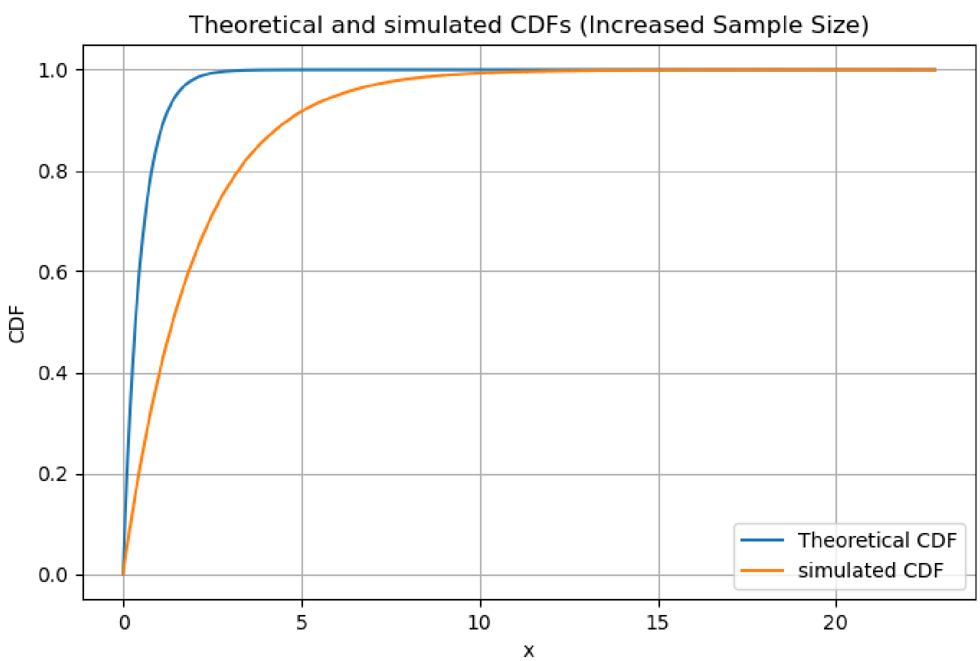


Figure 4.1: Figure1

Chapter 5

Random Algebra

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

$$p_{XY}(x, y) = \begin{cases} 8xy & if 0 < x < y < 1 \\ 0 & otherwise \end{cases} \tag{5.1}$$

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:

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

where

$$p_{X|Y} = \frac{p_{XY}(x,y)}{p_Y(y)} \tag{5.3}$$

$$p_Y(y) = \int_0^y p_{X|Y}(x,y)\,dx \tag{5.4}$$

for $0 < y < 1$

$$= \int_0^y 8xydx \tag{5.5}$$

$$= 8y\left[\frac{x^2}{2}\right]_0^y \tag{5.6}$$

$$= 4y^3 \tag{5.7}$$

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

$$p_{X|Y} = \frac{8xy}{4y^3} \tag{5.8}$$

$$= \frac{2x}{y^2} \tag{5.9}$$

and

$$E(X|Y=y_0) = \int_0^{y_0} x.\frac{2x}{y_0^2}dx \tag{5.10}$$

$$= \frac{2}{y_0^2}\left[\frac{x^3}{3}\right]_0^{y_0} \tag{5.11}$$

$$= \frac{2y_0}{3} \tag{5.12}$$

$$\implies \frac{2y_0}{3} = \frac{1}{2} \tag{5.13}$$

$$y_0 = \frac{3}{4} \tag{5.14}$$

Chapter 6

Hypothesis Testing

6.1 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:**

$$L = \prod_{i=1}^{1} f(x) = f(x) \tag{6.1}$$

To determine the most powerful test, we need to consider the likelihood ratio test

$$\frac{L(H_1)}{L(H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} k \tag{6.2}$$

$$\implies \frac{\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}}{\frac{1}{2} e^{-2|x|}} \underset{H_0}{\overset{H_1}{\gtrless}} k \tag{6.3}$$

$$\implies e^{\frac{x^2-2|x|}{2}} \underset{H_0}{\overset{H_1}{\gtrless}} k\frac{\sqrt{\pi}}{\sqrt{2}} \tag{6.4}$$

$$(|x|-1)^2 \underset{H_0}{\overset{H_1}{\gtrless}} 2\log\left(\frac{k\sqrt{\pi}}{\sqrt{2}}\right) + 1 \tag{6.5}$$

Taking square root on both sides,

$$||x|-1| \underset{H_0}{\overset{H_1}{\gtrless}} \sqrt{2\log\left(\frac{k\sqrt{\pi}}{\sqrt{2}}\right) + 1} \tag{6.6}$$

$$\implies |x| \underset{H_0}{\overset{H_1}{\gtrless}} 1 + \sqrt{2\log\left(\frac{k\sqrt{\pi}}{\sqrt{2}}\right) + 1} \tag{6.7}$$

Hence, the correct answer is (6.1c)

6.2 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 $\{i : X_i > \theta_0, 1 \le i \le n\}$. 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 6.1: Median θ is defined as

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

Definition 6.2: S is defined as

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

where $I(X_i > \theta_0$ represents an indicator function.

$$I(X_i > \theta_0) = \begin{cases} 1, & \text{if } X_i > \theta_0 \\ 0, & \text{if } X_i \leq \theta_0 \end{cases} \tag{6.8}$$

$$E(S) = E\left(\sum_{i=1}^{n} I(X_i > \theta_0)\right) \tag{6.9}$$

$$= \sum_{i=1}^{n} E(I(X_i > \theta_0)) \tag{6.10}$$

Since,

$$E(I(X_i > \theta_0)) = P(X_i > \theta_0) = \int_{\theta_0}^{\infty} f(x)\,dx \tag{6.11}$$

Therefore,

$$E(S) = \sum_{i=1}^{n} \int_{\theta_0}^{\infty} f(x)\,dx \tag{6.12}$$

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

(b) The power function can be expressed as:

$$\pi(\theta) = \Pr(\text{Reject } H_0 \,|\, H_1 \text{ is true}) \tag{6.13}$$

$$= \Pr(S > c|\theta) \tag{6.14}$$

Therefore, power function depends on value of θ.

6.3 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

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

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

(A) $\frac{1}{n}\sum_{i=1}^{n}\left(\ln X_i\right)^2$ 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:

$$L\left(\theta\right) = f\left(x_1, x_2, .., x_n;\theta\right) \tag{6.15}$$

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,

$$L(\theta) = \prod_{i=1}^{n} f(x_i; \theta) \tag{6.16}$$

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

$$\begin{aligned}
l(\theta) &= \ln L(\theta) &(6.17)\\
&= \ln\left(\prod_{i=1}^{n} f(x_i; \theta)\right) &(6.18)\\
&= \sum_{i=1}^{n} \ln f(x_i; \theta) &(6.19)\\
&= -n\ln 2 - n\ln\theta + \sum_{i=1}^{n} \ln(-\ln x_i) - \sum_{i=1}^{n} (\ln x_i) - \sum_{i=1}^{n} \frac{(\ln x_i)^2}{\theta} &(6.20)
\end{aligned}$$

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

$$\frac{\partial l(\theta)}{\partial \theta} = 0 \tag{6.21}$$

$$\frac{-n}{\theta} + \frac{1}{(\theta)^2} \sum_{i=1}^{n} (\ln X_i)^2 = 0 \tag{6.22}$$

$$\theta = \frac{1}{n} \sum_{i=1}^{n} (\ln X_i)^2 \tag{6.23}$$

Hence (A) is the true statement.

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

$$P(X = 0, Y = 0) = P(X = 1, Y = 1) = \theta, \tag{6.24}$$

$$P(X = 1, Y = 0) = P(X = 0, Y = 1) = \frac{1}{2} - \theta. \tag{6.25}$$

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,

$$H_0 : \quad \theta = \theta_0 = \frac{1}{4}, \tag{6.26}$$

$$H_1 : \quad \theta = \theta_1 = \frac{1}{3}. \tag{6.27}$$

and the pmf is given by

$$p_{XY}(0, 0) = p_{XY}(1, 1) = \theta \tag{6.28}$$

$$p_{XY}(0, 1) = p_{XY}(1, 0) = \frac{1}{2} - \theta \tag{6.29}$$

Then for the given random sample of data,

$$p_{X_i,Y_i}(x,y) = \begin{cases} 2\theta & x = y \\ 1 - 2\theta & x \neq y \end{cases} \tag{6.30}$$

$$\tag{6.31}$$

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

$$L(\theta_0 \mid data) = \prod_{i=1}^{n} p_{X_i,Y_i}(x,y) \tag{6.32}$$

$$= (2\theta_0)^m (1 - 2\theta_0)^{n-m} \tag{6.33}$$

$$= \left(\frac{1}{2}\right)^m \left(\frac{1}{2}\right)^{n-m} \tag{6.34}$$

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

$$L(\theta_1 \mid data) = \prod_{i=1}^{n} p_{X_i,Y_i}(x,y) \tag{6.35}$$

$$= (2\theta_1)^m (1 - 2\theta_1)^{n-m} \tag{6.36}$$

$$= \left(\frac{2}{3}\right)^m \left(\frac{1}{3}\right)^{n-m} \tag{6.37}$$

The likelyhood ratio will be

$$\lambda(data) = \frac{L(\theta_1 \mid x)}{L(\theta_0 \mid x)} \tag{6.38}$$

$$= \frac{\left(\frac{2}{3}\right)^m \left(\frac{1}{3}\right)^{n-m}}{\left(\frac{1}{2}\right)^m \left(\frac{1}{2}\right)^{n-m}} = (2)^m \left(\frac{2}{3}\right)^n \tag{6.39}$$

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

$$\implies \lambda(data) \underset{H_0}{\overset{H_1}{\gtrless}} c_1 \tag{6.40}$$

$$\tag{6.41}$$

From (6.39),

$$\implies (2)^m \left(\frac{2}{3}\right)^n \underset{H_0}{\overset{H_1}{\gtrless}} c_1 \tag{6.42}$$

$$\implies (2)^m \underset{H_0}{\overset{H_1}{\gtrless}} c_1 \left(\frac{2}{3}\right)^n \tag{6.43}$$

$$\implies m \underset{H_0}{\overset{H_1}{\gtrless}} \log_2 \left(c_1 \left(\frac{2}{3}\right)\right)^n \tag{6.44}$$

$$\implies m \underset{H_0}{\overset{H_1}{\gtrless}} c \quad \exists\, c \in \mathbb{R} \tag{6.45}$$

where,

$$c = \log_2 \left(c_1 \left(\frac{2}{3}\right)\right)^n \tag{6.46}$$

$\therefore$ From (6.45), Option A is correct and Options B,C,D are incorrect

Chapter 7

Bivariate Random Variables

Chapter 8

Random Processes

Chapter 9

Convergence

9.1 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,

$$lim_{n\to\infty} F_{X_n}(x) = F_X(x) \tag{9.1}$$

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

$$lim_{n\to\infty} \Pr\left(|X_n - X| > \epsilon\right) = 0 \tag{9.2}$$

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

$$lim_{n\to\infty} E(|X_n - X|^p) = 0 \tag{9.3}$$

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

$$B = \{x : |x - c| \geq \epsilon\} \tag{9.4}$$

Now,

$$\Pr\left(|X_n - c| \geq \epsilon\right) = \Pr\left(X_n \in B\right) \tag{9.5}$$

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

$$\begin{aligned}
\limsup_{n\to\infty} \Pr\left(X_n \in B\right) &\leq \Pr\left(c \in B\right) && (9.6)\\
&\leq \Pr\left(|0 - 0| \geq \epsilon\right) && (9.7)\\
&\leq \Pr\left(0 \geq \epsilon\right) && (9.8)\\
&\leq 0 && (9.9)\\
&= 0 && (9.10)\\
lim_{n\to\infty} \Pr\left(|X_n - c| > \epsilon\right) &= 0 && (9.11)
\end{aligned}$$

From (9.2), $X_n \xrightarrow{p} c$. So, we have

$$X_n \xrightarrow{d} c \implies X_n \xrightarrow{p} c \tag{9.12}$$

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 $X = 0$

$$lim_{n\to\infty} \Pr(|X_n - X| > \epsilon) = lim_{n\to\infty} \Pr(X_n > \epsilon) \tag{9.13}$$

$X_n > \epsilon$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,

$$\Pr(X_n > \epsilon) \leq \Pr(X_n = n) \tag{9.14}$$

$$lim_{n\to\infty} \Pr(|X_n - X| > \epsilon) \leq lim_{n\to\infty} \Pr(X_n = n) \tag{9.15}$$

$$\leq lim_{n\to\infty} \frac{1}{n} \tag{9.16}$$

$$\leq 0 \tag{9.17}$$

$$= 0 \tag{9.18}$$

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

$$\begin{aligned} lim_{n\to\infty} E(|X_n - X|^3) &= lim_{n\to\infty} E(X_n^3) && (9.19)\\ &= lim_{n\to\infty} 0^3 \left(1 - \frac{1}{n}\right) + n^3 \left(\frac{1}{n}\right) && (9.20)\\ &= lim_{n\to\infty} n^2 \neq 0 && (9.21)\end{aligned}$$

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

$$Z \sim \mathcal{N}(0,1) \tag{9.22}$$

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

$$lim_{n\to\infty} F_{Y_n}(x) = \Pr(Z \leq x) \tag{9.23}$$

For n odd

$$\begin{aligned} lim_{n\to\infty} F_{Y_n}(x) &= \Pr(-Z \leq x) && (9.24)\\ &= \Pr(Z \leq x) && (9.25)\end{aligned}$$

Proved. So,we have

$$F_{X_n+Y_n}(x) = \Pr\left(X_n + Y_n \leq x\right) \tag{9.26}$$
$$= \Pr\left(Z + (-1)^n Z \leq x\right) \tag{9.27}$$

For n is even

$$F_{X_n+Y_n}(x) = \Pr\left(2Z \leq x\right) \tag{9.28}$$
$$= \Pr\left(Z \leq \frac{x}{2}\right) \tag{9.29}$$
$$= 1 - \Pr\left(Z > \frac{x}{2}\right) \tag{9.30}$$
$$\approx 1 - Q\left(\frac{x}{2}\right) \tag{9.31}$$

For n is odd

$$F_{X_n+Y_n}(x) = \Pr\left(0 \leq x\right) \tag{9.32}$$
$$= \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases} = H(x) \tag{9.33}$$

So, on generalizing

$$F_{X_n+Y_n}(x) = \begin{cases} 1 - Q\left(\frac{x}{2}\right) & \text{if } n \text{ is even} \\ H(x) & \text{if } n \text{ is odd} \end{cases} \tag{9.34}$$

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

$$lim_{n\to\infty}E(X_n) = 0\left(1-\frac{1}{n}\right)+n\left(\frac{1}{n}\right) \tag{9.35}$$

$$= 1 \tag{9.36}$$

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

$$lim_{n\to\infty}X_n = 0 = X \tag{9.37}$$

To find 1^{st} mean convergennce of X_n. From (9.36)

$$lim_{n\to\infty}E(|X_n - X|) = lim_{n\to\infty}E(X_n) \tag{9.38}$$

$$= 1 \neq 0 \tag{9.39}$$

So, X_n does not converges in 1^{st} mean to X.

Chapter 10

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	$\frac{1}{128}$

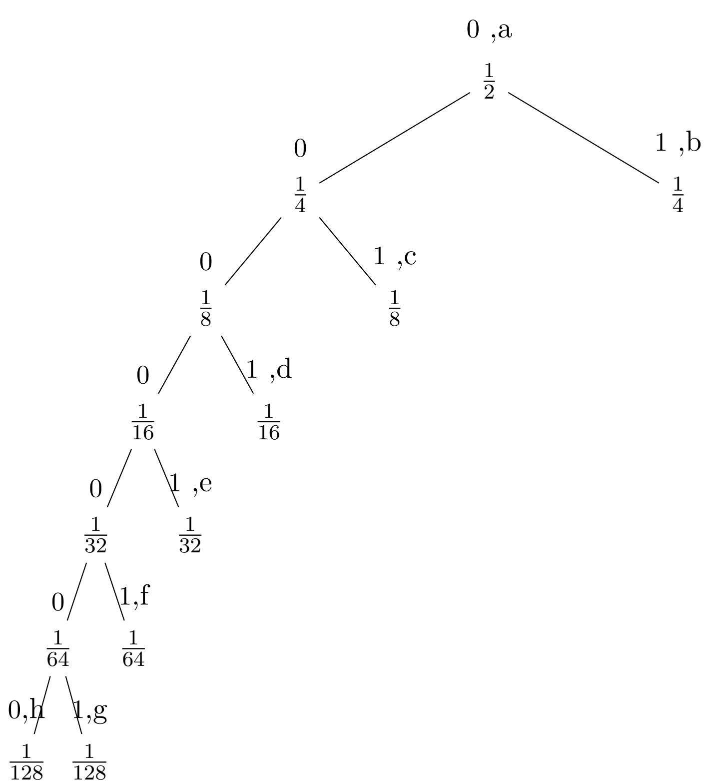
Solution:

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

$$
\begin{aligned}
H(X) &= \sum_i p_X(i) \log_b \left(\frac{1}{p_X(i)} \right) && (10.1)\\
&= \frac{1}{2}\log_2(2) + \frac{1}{4}\log_2(4) + ... + \frac{1}{128}\log_2(128) && (10.2)\\
&= 0.5 + 0.5 + 0.375 + ... + 0.0078125 && (10.3)\\
&= 1.984375 && (10.4)
\end{aligned}
$$

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



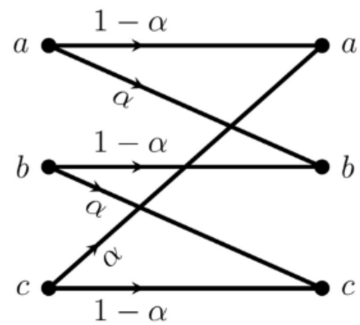
Using the above binary table following code is generated;

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 10.1: 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

$$C = \sup_{p_X(x)} I\left(X, Y\right) \tag{10.5}$$

$$\begin{aligned}
I\left(X, Y\right) &= \sum_{x,y} p\left(x, y\right) \log_2 \frac{p\left(x, y\right)}{p\left(x\right) p\left(y\right)} &(10.6)\\
&= \sum_{x,y} p\left(x, y\right) \log_2 \frac{p\left(\left. y \right| x\right)}{p\left(y\right)} &(10.7)\\
&= -\sum_{x,y} p\left(x, y\right) \log_2 p\left(y\right) + \sum_{x,y} p\left(x, y\right) \log_2 p\left(\left. y \right| x\right) &(10.8)\\
&= -\sum_{y} p\left(y\right) \log_2 p\left(y\right) - \left(-\sum_{x,y} p\left(x, y\right) \log_2 p\left(\left. y \right| x\right)\right) &(10.9)\\
&= H\left(Y\right) - H\left(\left. Y \right| X\right) &(10.10)
\end{aligned}$$

Now,

$$\sum_{i=0}^{2} p_i = 1 \tag{10.11}$$

$$\sum_{j=0}^{2} q_i = 1 \tag{10.12}$$

$$H(\mathbf{q}) = -\sum_{i=0}^{2} q_i \log_2 q_i \tag{10.13}$$

$$= -\left(q_0 \log_2 q_0 + q_1 \log_2 q_1 + q_2 \log_2 q_2\right) \tag{10.14}$$

$$H(Y|\, X) = -\sum_{i=0}^{2}\sum_{j=0}^{2} p_i p_{Y|\,X}\left(y_j|\, x_i\right)\log_2\left(p_{Y|\,X}\left(y_j|\, x_i\right)\right) \tag{10.15}$$

$$\begin{aligned} &= -p_0\left((1-\alpha)\log_2(1-\alpha) + \alpha\log_2\alpha\right) \\ &\quad - p_1\left((1-\alpha)\log_2(1-\alpha) + \alpha\log_2\alpha\right) \\ &\quad - p_2\left((1-\alpha)\log_2(1-\alpha) + \alpha\log_2\alpha\right) \end{aligned} \tag{10.16}$$

Using (10.14) and (10.16) in (10.10)

$$\begin{aligned} I(X,Y) &= -\left(q_0 \log_2 q_0 + q_1 \log_2 q_1 + q_2 \log_2 q_2\right) \\ &\quad + p_0\left((1-\alpha)\log_2(1-\alpha) + \alpha\log_2\alpha\right) \\ &\quad + p_1\left((1-\alpha)\log_2(1-\alpha) + \alpha\log_2\alpha\right) \\ &\quad + p_2\left((1-\alpha)\log_2(1-\alpha) + \alpha\log_2\alpha\right) \end{aligned} \tag{10.17}$$

$$\implies \frac{d}{d\alpha} I\left(X,Y\right) = p_0 \log_2\left(\frac{\alpha}{1-\alpha}\right) + p_1 \log_2\left(\frac{\alpha}{1-\alpha}\right) + p_2 \log_2\left(\frac{\alpha}{1-\alpha}\right) \tag{10.18}$$

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

$$\log_2\left(\frac{\alpha}{1-\alpha}\right)\left(p_0 + p_1 + p_2\right) = 0 \tag{10.19}$$

$$\implies \alpha = \frac{1}{2} \tag{10.20}$$

Chapter 11

Markov chain

11.1 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:

$$\begin{pmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{pmatrix} \tag{11.1}$$

$$\Pr\left(X_2 = 1 | X_1 = 1, X_3 = 2\right) = \Pr\left(X_2 = 1 | X_1 = 1\right) \tag{11.2}$$

$$= p_{11} \tag{11.3}$$

$$= 0.5 \tag{11.4}$$

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

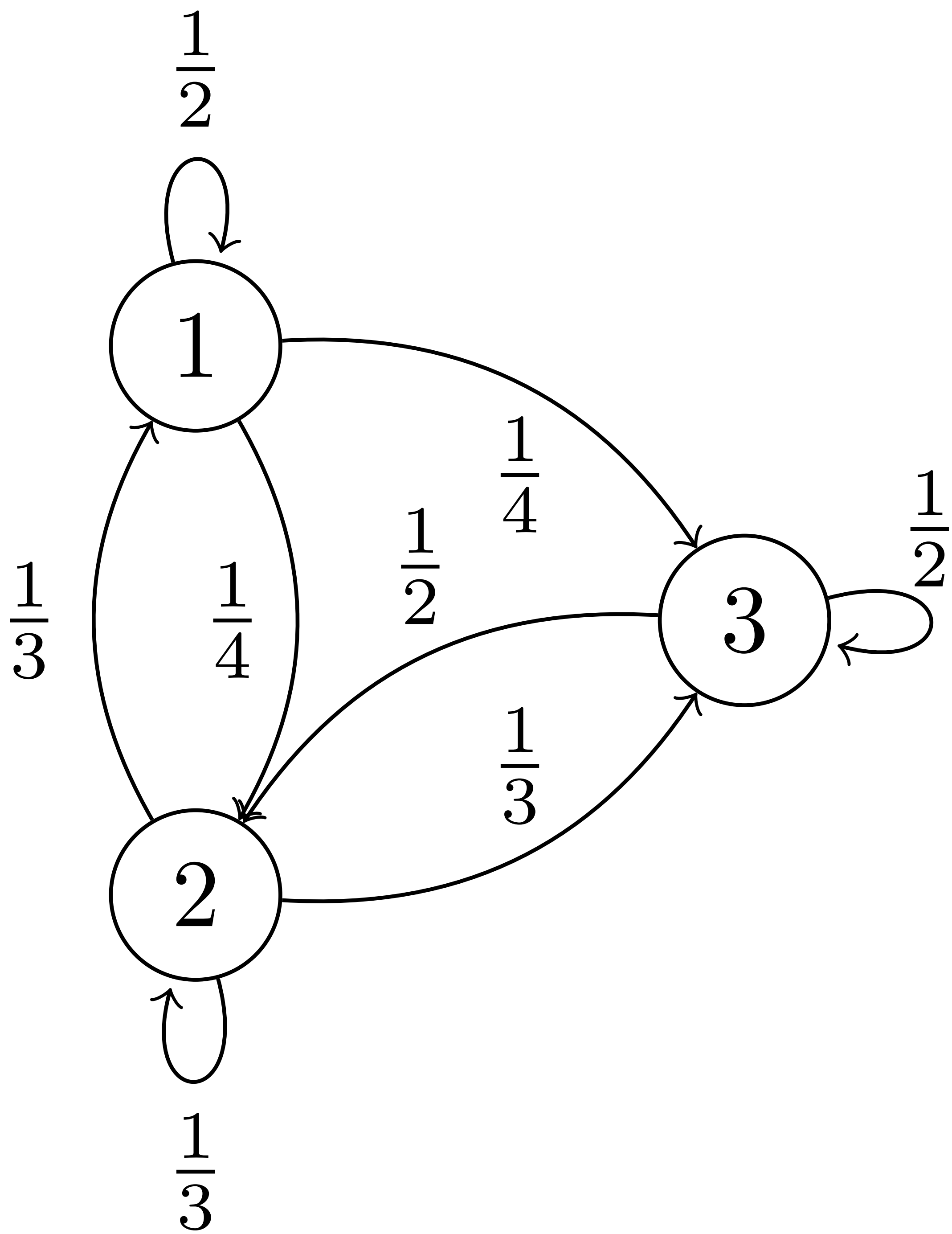


Figure 11.1: Markov Chain diagram