Chapter 1

Probability

Definition 1. *A* **Sample Space***, S, is a set that contains all possible out comes of an experiment. Any element of the sample space S is called sample point. An* **event** *E is any subset of S.*

Definition 2. Axioms of Probability

1. *For any event E, $P(E) \geq 0$ i.e. the probability of any event cannot be non-negative.*
2. *For the sample space S, $P(S) = 1$.*
3. *If $E_1, E_2, E_3, \ldots$ are disjoint events then*

$$p(E_1 \cup E_2 \cup E_3 \cup \ldots) = p(E_1) + p(E_2) + p(E_3) + \ldots$$

Note* The axioms are fundamental rules of probability which are used to prove all other results. For example, the following theorems have been proved using the above axioms.

Theorem 1. *For any event E, $p(E) + p(\overline{E}) = 1$.*

Proof. Since E and $\overline{E}$ are disjoint and $E \cup \overline{E} = S$, we have

$$p(E \cup \overline{E}) = p(S) \implies P(E) + p(\overline{E}) = 1$$

This also gives $p(\overline{E}) = 1 - p(E)$.

Note* Sometimes, it is easier to calculate $p(\overline{E})$ rather than $p(E)$. So we calculate $p(\overline{E})$ and use $p(E) = 1 - p(\overline{E})$.

Theorem 2. *For any events E_1 and E_2 $p(E_1 \cup E_2) = p(E_1) + p(E_2) - p(E_1 \cap E_2)$.*

Proof. Write $E_1 \cup E_2 = E_1 \cup (\overline{E_1} \cap E_2)$. Since E_1 and $\overline{E_1} \cap E_2$ are disjoint, we have

$$p(E_1 \cup E_2) = p(E_1 \cup (\overline{E_1} \cap E_2)) = p(E_1) + p(\overline{E_1} \cap E_2) \tag{1.1}$$

Now, similarly, E_2 can be written as $E_2 = (E_1 \cap E_2) \cup (\overline{E}_1 \cap E_2)$ as union of disjoint sets. This gives us

$$p(E_2) = p(E_1 \cap E_2) + p(\overline{E}_1 \cap E_2) \implies p(\overline{E}_1 \cap E_2) = p(E_2) - p(E_1 \cap E_2)$$

Hence, we get

$$p(E_1 \cup E_2) = p(E_1) + p(E_2) - p(E_1 \cap E_2)$$

1.1 Classical Approach (Theoritical Probability)

In classical approach, we consider that all outcomes of sample space S are equally likely. In this case the probability of an event E is given by

$$p(E) = \frac{n(E)}{n(S)} \tag{1.2}$$

where $n(E)$ is the size of set E (number of elements of E) and $n(S)$ is the size of S. We employ various **counting techniques** as mentioned below to calculate the sizes of event E and set S.

Empirical Probability : Sometimes, the probability has to be calculated from repeating an experiment or from observation. Such probability is called **empirical probability**.

1.1.1 Inclusion Exclusion Principle

For two sets A and B, we have

$$n(A \cup B) = n(A) + n(B) - n(A \cap B) \tag{1.3}$$

Similarly for three sets A, B and C,

$$\begin{aligned} n(A \cup B \cup C) = n(A) + n(B) + n(C) - n(A \cap B) - n(B \cap C) \\ - n(A \cap C) + n(A \cap B \cap C) \end{aligned}$$

1.1.2 Basic Counting Principle

If there are n_1 ways of doing something and n_2 ways of doing another thing then there are total ways of $n_1 n_2$ both actions together.

1.1.3 Combinatorics

From a collection of n items, if r items are **arranged in order** then it is a permutation. A **permutation** of r elements from n element is given by

$$^nP_r = \frac{n!}{(n-r)!} \tag{1.4}$$

From a collection of n items, if r items are simply selected (without ordering) then it is a combination. A **combination** of r elements from n element is given by

$$^nC_r = \frac{n!}{r!(n-r)!} \tag{1.5}$$

Assignment 1. *In a construction company, there are* 20 *Civil engineers,* 12 *IT officers, and* 18 *finance officers. A committee of* 5 *has to be formed from a random selection. What is the probability that the committee consists of*

1. *All civil engineers.*
2. *One civil engineer.*
3. *At least one civil engineer.*

4. No civil engineer.

5. Three civil engineers and two from each sector.

6. Two civil engineers, two IT officers, and a finance officer.

[Hint] If there are different options then make different groups for each optoins and calculate choices for each group. Then multiply the choices for each group (basic counting principle) to calculate the choices. Also make use of $p(E) = 1 - p(\overline{E})$.

Assignment 2. *A random number is chosen from from* $1, 2, 3, \ldots, 100$. *What is the probability that the number is neither divisible by* 2 *nor by* 3.

[Hint] Calculate $p(\overline{A \cup B})$.

Assignment 3. *In a certain town, 850 people are surveyed which showed that* 230 *use gas stove,* 175 *use induction stove and* 40 *use both. If a random person is selected, find the probability that the person*

1. uses at least one of fuel.

2. use neither of fuel.

3. use gas stove only.

4. use electricity only.

5. use exactly one.

[Hint] Make use of set theory.

1.2 Conditional Probability

Let E_1 and E_2 be two events, the probability of E_1 given that E_2 has already happened is given by

$$p(E_1|E_2) = \frac{p(E_1 \cap E_2)}{p(E_2)} \tag{1.6}$$

Also, this gives the multiplication rule for $p(E_1 \cap E_2)$ as

$$p(E_1 \cap E_2) = p(E_2)p(E_1|E_2)$$

1.2.1 Independent Events

The two events E_1 and E_2 are said to be independent events if occurrence of one does not affect the other. In such case $p(E_1|E_2) = p(E_1)$ and $p(E_2|E_1) = p(E_2)$. In this case

$$p(E_1|E_2) = \frac{p(E_1 \cap E_2)}{p(E_2)} \implies p(E_1) = \frac{p(E_\cap E_2)}{p(E_2)}$$

So, we get

$$p(E_1 \cap E_2) = p(E_1)p(E_2)$$

Therefore, it can be said that events $E_1, E_2, \ldots E_n$ are independent events if

$$p(E_1 \cap E_2 \cap \cdots \cap E_n) = p(E_1)p(E_2)\ldots p(E_n) \tag{1.7}$$

Assignment 4. *A statistics problem is given to students A, B and C and their probabilities for solving the problem is $0.75, 0.6$ and 0.5. Suppose they solve it independently then find the probability that*

1. *all can solve the problem.*
2. *none can solve the problem.*
3. *problem can be solved.*
4. *exactly one can solve the problem.*

Assignment 5.

An oil exploration company has two active projects, project A and project B. Let A represent an event where project A is successful and similarly B represent the event where project B is successful. Assume that project are run independently and has the probability of succeeding $p(A) = 0.4$ and $p(B) = 0.7$.

1. *If project A is not successful then that is probability that project B is also not successful.*
2. *What is the probability that at least one of the project is successful.*
3. *Given that at least one of the project is successful, what is the probability that project A is successful.*

1.3 Bayes Theorem

1.3.1 Mutually Exclusive Events

The events E_1 and E_2 are said to be mutually exclusive events if $p(E_1 \cap E_2) = 0$. In such case

$$p(E_1 \cap E_2) = p(E_1) + p(E_2)$$

The events E_1 and E_2 are mutually exclusive if both events cannot occur simultaneously i.e. if E_1 occurs then E_2 cannot occur and vice versa. Thus, the mutually exclusive events are disjoint as well.

1.3.2 Exhaustive Events

Let S be a sample space. The events $E_1, E_2, \dots, E_n$ are exhaustive events if

$$E_1 \cup E_2 \cup \cdots \cup E_n = S$$

i.e. at least one of the event occurs from the given collection.

Theorem 3. *(Total Probability Rule) If $E_1, E_2, E_3, \dots, E_n$ are mutually exclusive and exhaustive events then for any event A*

$$p(A) = p(A|E_1) + p(A|E_2) + \cdots + p(A|E_n)$$

Proof. Since $E_1, E_2, E_3, \ldots, E_n$ are exhaustive events,

$$\begin{aligned} A &= A \cap (E_1 \cup E_2 \cup E_3 \cup \cdots \cup E_n) \\ &= (A \cap E_1) \cup (A \cap E_2) \cup \cdots \cup (A \cap E_n) \end{aligned}$$

Since $E_1, E_2, \ldots E_n$ are mutually exclusive (disjoint), using Axiom 3 of probability, we get

$$p(A) = p(A \cap E_1) + p(A \cap E_2) + \cdots + p(A \cap E_n)$$

Using multiplication rule for intersection of events, we get

$$p(A) = p(E_1)p(A|E_1) + p(E_2)p(A|E_2) + \cdots + p(E_n)p(A|E_n)$$

Theorem 4. *(Bayes Theorem) If $E_1, E_2, E_3, \ldots, E_n$ are mutually exclusive and exhaustive events then for any event A*

$$p(E_i|A) = \frac{p(E_i \cap A)}{\sum_{i=1}^{n} p(A|E_i)p(E_i)} = \frac{p(A|E_i)p(E_i)}{\sum_{j=1}^{n} p(A|E_j)p(E_j)}$$

Proof. The conditional probability is given by

$$p(E_i|A) = \frac{p(E_i \cap A)}{p(A)}$$

From Total Probability Rule, we have

$$p(A) = \sum_{j=1}^{n} p(A|E_j)p(E_j)$$

which gives us

$$p(E_i|A) = \frac{p(E_i \cap A)}{\sum_{j=1}^{n} p(A|E_j)p(E_j)}$$

and using the multiplication rule for intersection of events,

$$p(E_i \cap A) = p(A \cap E_i) = p(A|E_i)p(E_i)$$

Hence, we get

$$p(E_i|A) = \frac{p(A|E_i)p(E_i)}{\sum_{j=1}^{n} p(A|E_j)p(E_j)} \tag{1.8}$$

Note* If asked on exam, include the proof of Total Probability Rule as well.

Assignment 6. *Two plants, plant A and plant B, are contracted by Nepal Electricity Authority for make wire cables. Plant A is contracted to make 40% of cables and meets 95% strength specifications. Plant B makes 60% of cables and meets 98% strength specifications.*

1. *A random cable is selected and found to meet the required strength specifications. What is the probability that wire came from plant A.*

2. *Let A represent an event where a randomly selected cable comes from plant A and B represent for plant B. Let Y represent an event that the randomly selected cable meets the strength standards and N represent event where the cable does not meet the standard. Are events A and Y independent.*

Assignment 7. *In a certain recruitment test, there are multiple choice questions. There are four possible options to each question in which one of them is correct. A student knows* 80% *of the answer to the test.*

1. *What is the probability of getting a correct answer?*
2. *If student correctly answers a randomly chosen question then what is the probability that the student is guessing?*
3. *If student correctly answers a randomly chosen question then what is the probability that the student knows the answer?*

Assignment 8. *The content (marbles) of urns A, B and C of different colors are as follows. If two marbles are drawn from randomly chosen urns are found to be red and white, what is the probability that they came from urn A?*

Urn	White	Black	Red
A	10	20	30
B	20	10	10
C	40	50	30

Assignment 9. *The following table shows the percent of good review for different product.*

Success	Market Share (%)	Good Review (%)
High	40	95
Moderate	35	60
Low	25	10

1. *A new product is launched. Find the probability that the product attains a good review.*
2. *If a new product attains a good review, what is the probability that it will be highly successful product.*

Assignment 10. *A company produces machine components which pass through an automatic testing machine. Of all components entering the testing machine,* 5% *are defective. However, the testing machine is not reliable. If the component is defective then there is* 4% *chance that it will not be rejected. If a component is not defective then there is* 7% *probability that it will be rejected.*

1. *What fraction of the component is rejected.*
2. *What fraction of the components rejected are actually not defective.*

Chapter 2

Discrete Random Variables and Probability Distributions

Definition 3. *A* **random variable (rv)** *is any rule that associates each outcome of sample space S of any experiment to a number. Mathematically,* **rv** *is a function whose domain is sample space and range is a set of real numbers i.e. $X : S \to (-\infty, \infty)$*

Definition 4. *A rv X is* **discrete** *rv if its possible values is a finite or countable set. A rv X is continuous if it's possible values constitute an interval or union of interval in $(-\infty, \infty)$*

Note* We will use set S to represent a sample space and $X(S)$ will represent the range of X. Also, $s \in S$ represent element of sample space S and $x \in X(S)$ represent possible value of rv X.

Definition 5. *The* **probability distribution** *or* **probability mass function (pmf)** *of discrete rv is defined for every number $x \in X(S)$ by $p(x) = p(X = x) = p(alls \in S : X(s) = x)$.*

Note* The **pmf** gives (how) the distributions of probability over $X(S)$. If $p(x)$ is probability function than $\sum_{x \in X(S)} p(x) = 1$ i.e. the sum of all probabilities is 1.

Definition 6. *Let X be a discrete rv with pmf $p(x)$. The* **expected value** *of any function, denoted by $E[X]$ or μ_x or just μ, is given by*

$$E[f(X)] = \sum_{x \in X(S)} f(x) \cdot p(x)$$

If $f(x) = x$ then then $E[X] = \sum_{x \in X(S)} xp(x)$ is called **mean** *and if $f(x) = (x - \mu)^2$ then*

$$E[(x - \mu)^2] = \sum_{x \in X(S)} (x - \mu)^2 p(x) = E[X^2] - E[X]^2$$

is called **variance**. *The variance of rv $f(X)$ is given by*

$$V[f(X)] = \sum_{x \in X(S)} (f(x) - E[f(X)])^2 p(x)$$

Definition 7. *The cumulative distribution function* **(cdf)** *is of pmf $p(x)$ is given by*

$$c(x) = \sum_{y \leq x} p(y)$$

2.1 Binomial Distribution

The conditions for **Binomial distribution** are as follows

1. The number of observations n is fixed.
2. Each observation is independent.
3. Each observation represents one of two outcomes (”success” or ”failure”).
4. The probability of ”success” p is the same for each outcome.

The probability mass function (pmf) of binomial distribution of r successes with n observations and success probability s is given by

$$p(x; n; s) = {}^{n}C_{r} s^{r}(1-s)^{n-r}$$

which is also the probability distribution function.

2.2 Poisson Distribution

Definition 8. *A discrete random variable X is said to have a Poisson distribution with parameter $\mu(\mu > 0)$ if the pmf of X is*

$$p(x; \mu) = \frac{e^{-\mu} \cdot \mu^{x}}{x!} \quad x = 0, 1, 2, 3, \ldots$$

Poisson distribution can be obtained as limit of Binomial distribution with $n \to \infty, np \to \mu$.

2.3 Hypergeometric Distribution

Assumption of Hypergeometric Distribution

1. The population or set to be sampled consists of N individuals, objects, or elements (a finite population).
2. Each individual can be characterized as a success (S) or a failure (F), and there are M successes in the population.
3. A sample of n individuals is selected without replacement in such a way that each subset of size n is equally likely to be chosen.

Definition 9. *If X is the number of $S's$ in a completely random sample of size n drawn from a population consisting of $MS's$ and $(N-M)F's$, then the probability distribution of X, called the hypergeometric distribution, is given by*

$$P(X = x) = h(x; n, M, N) = \frac{{}^{M}C_{x} {}^{N-M}C_{n-x}}{{}^{N}C_{n}}$$

for x, an integer, satisfying $\max(0, n-N+M) \leq x \leq \min(n, M)$.

2.4 Negative Binomial Distribution

1. The experiment consists of a sequence of independent trials.
2. Each trial can result in either a success (S) or a failure (F).
3. The probability of success is constant from trial to trial, so $P(S \text{ on trial } i) = p$ for $i = 1, 2, 3, \ldots$
4. The experiment continues (trials are performed) until a total of r successes have been observed, where r is a specified positive integer.

Definition 10. *The pmf of the negative binomial rv X with parameters $r =$ number of $S's$ and $p = P(S)$ is*

$$nb(x; r, p) = {}^{x+r-1}C_{r-1}\, p^r (1-p)^x \quad x = 0, 1, 2, \ldots$$

Distribution	Given parameters (or to be calculated)	pmf ($p(x)$)	$E[X]$	$V[X]$
Binomial	n, p	${}^nC_x p^x (1-p)^x$	np	$np(1-p)$
Poisson	μ	$e^{-\mu} \frac{\mu^x}{x!}$	μ	μ
Hypergeometric	N, M, n	$\frac{{}^MC_x\, {}^{N-M}C_{n-x}}{{}^NC_n}$	$n \cdot \frac{M}{N}$	$\left(\frac{N-n}{N-1}\right) \cdot n \cdot \frac{M}{N} \cdot \left(1 - \frac{M}{N}\right)$
Negative Binomial	r, p	${}^{x+r-1}C_{r-1}\, p^r (1-p)^x$	$\frac{r(1-p)}{p}$	$\frac{r(1-p)}{p^2}$