CHAPTER 14

Probability Distribution

Random Variable

Let S be the sample space associated with a given random experiment. Then, a real valued function X which assigns each outcome $w \in S$ to a unique real number $X(w)$ is called a random variable.

In other words, the domain of a random variable is the sample space of a random experiment, while its codomain is the set of real numbers.

Thus, $X : S \to R$ is a random variable.

e.g. A coin is tossed ten times. The random variable X is the number of tails that are noted. Then, X can only take the values $0, 1, 2, \ldots, 10$.

Types of Random Variables

There are two types of random variables, namely discrete and continuous.

i. Discrete Random Variable

If the range of the real function $X : U \to R$ is a finite set or an infinite set of real numbers, then X is called a discrete random variable.

e.g. In tossing of two coins $S = \{HH, HT, TH, TT\}$, let X denotes number of heads in tossing of two coins, then $X(HH) = 2, X(TH) = 1, X(TT) = 0$

Note *The values of a discrete random variable are obtained by counting.*

ii. Continuous Random Variable

If the range of X is an interval (a, b), then X is called a continuous random variable.

e.g. Suppose temperature of a city varies between 20°C and 30°C. Thus, it can be take any value in the interval (20, 30).

Probability Distribution of Discrete Random Variables

The system which link the values of random variable with the probability of their occurrence is called probability distribution.

If a random variable X takes values $x_1, x_2, \ldots, x_n$ with respective probabilities $p_1, p_2, \ldots, p_n$, then

X	x_1	x_2	$x_3 \ldots$	x_n
$P(X)$	p_1	p_2	$p_3 \ldots$	p_n

is known as the probability distribution of X.

Probability distribution gives the values of the random variable along with the corresponding probabilities.

It satisfies the following conditions:

(i) $0 \le P(x_i) \le 1$ (ii) $\Sigma P(x_i) = 1$

e.g. Probability distribution of number of heads when two coins are tossed.

Let X denotes the number of heads occurred, then

$P(X = 0) =$ Probability of occurrence of zero head

$$= P(TT) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$$

$P(X = 1) =$ Probability of occurrence of one head

$$= P(HT) + P(TH) = \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{2}$$

$P(X = 2) =$ Probability of occurrence of two heads

$$= P(HH) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$$

Thus, the probability distribution of number of heads when two coins are tossed is as given below:

X	0	1	2
$P(X)$	1/4	1/2	1/4

Probability Mass Function (p.m.f)

A function that specifies the probability of each value of discrete random variable, is called a probability mass function (abbreviated as pmf). If $f(x)$ is the probability mass function of the random variable X, then $f(x) = P(X = x)$ has the following properties:

(i) $f(x) \geq 0$ for all values of X (ii) $\Sigma f(x) = 1$

Cumulative Distribution Function (c.d.f)

If X is a discrete random variable with pmf $f(x)$, its cumulative mass function (abbreviated as cmf) specifies the probability that an observed value of X will not be greater than x. i.e. if $F(x)$ is a cmf, then

$$F(x) = P(X \leq x)$$

If X is a continuous random variable with the pdf, $f(x)$, then its cumulative distribution function or simply distribution function $F(x)$ is defined as

$$\begin{aligned} F(x) &= P(X \leq x) \\ &= \sum_{x_i < x} P[X = x_i] \\ &= \sum_{x_i < x} P_i = \sum_{x_i < x} f(x_i) \end{aligned}$$

where $-\infty < x < \infty$.

Note (i) $0 \leq F(x) \leq 1 \; \forall \; x \in R$ (ii) $F'(x) = f(x)$

Expected Value and Variance of a Random Variable

The average value of a random variable is called the expected value of the random variable. Let X be a discrete random variable with probability distribution $P(X)$, then the expected value $E(X)$ is given by $E(X) = \Sigma X \cdot P(X)$

where, the elements are summed over all values of the random variable X.

In other words, if a discrete random variable X has possible values $x_1, x_2, \ldots, x_n$, with corresponding probabilities $p_1, p_2, \ldots, p_n$, then the expected value $E(X)$ is defined as

$$E(X) = x_1 p_1 + x_2 p_2 + \ldots + x_n p_n$$

Thus, the expected value of random variable X is merely the mean which may be denoted by μ. Similarly, the variance of the probability distribution of the random variable X is defined as the expected value of the squared deviations of the values of X from their mean. Thus, the variance of the discrete random variable X is given by

$$\begin{aligned} \text{Var}(X) = \sigma^2 &= E[X - E(X)]^2 \\ &= \Sigma [X - E(X)]^2 \, P(X) \\ &= E(X^2) - [E(X)]^2 \end{aligned}$$

The standard deviation, σ is the square root of the variance.

i.e. $\quad \text{SD}(X) = \sqrt{E(X^2) - [E(X)]^2}$