M.Sc. (STATISTICS) (COURSE CODE: 138)

Paper - I: PROBABILITY AND DISTRIBUTION THEORY

SYLLABUS

UNIT – I

Probability as a set function, continuity axiom of probability, Borel - Cantelli lemma, random variable, distribution function and its properties, discrete and continuous distribution functions, convolutions of random variables, vector of random variables and statistical independence. Notion of mathematical expectation, conditional expectation, moment inequalities – Markov, chebyshev, Kolmogrov, Holder, minkowski. Characteristic function – Inversion theorem.

UNIT – II

Convergence of sequence of random variables – Types of convergence with interrelations. Weak Laws of Large numbers – Chebyshev and Khintchin's Laws. Kolomogrov's strong law of large numbers for independent random variables. Central limit property, central limit theorems – De Moivre's, Levy and Lindeberg forms of central limit theorem. Lindberg and Feller condition (Statement only). Liapounov's form of central limit theorem.

UNIT – III

Discrete distributions: Definitions, moment generating functions, probability generating functions, characteristic functions, means, variances, reproductive properties (if exist) and interrelations of multinamial, compound binomial, Compound Poisson.

UNIT – IV

Continuous distributions : Definitions, moment generating functions, probabillty generating functions, characteristic functions, means, variances, reproductive properties (if exist) and interrelations of Weibull, Laplace, lognormal, logistic.

UNIT – V

Sampling distributions central and non central chisquare, t and F. Order statistic - distribution function, Probability density function of a single order statistic, joint probability density function of two of more order statistics.

Note: Two Questions are to be set from each unit

CONTENTS

UNIT–I

NOTES

CHAPTER 1

PROBABILITY

OBJECTIVES

After going through this chapter, you should be able to:

- describe random experiment
- explain sample space
- know about algebra of events
- know probability of an event
- know about 'Odds in favour' and odds against an event
- know about addition theorem
- know about conditional probability and independent events
- know about probability mass function and probability density function
- know mean and variance, Joint and marginal probability mass function
- know about two dimensional distribution function and stochastic independence.

STRUCTURE

1.1 INTRODUCTION

The words **'Probability'** and **'Chance'** are quite familiar to everyone. Many a times, we come across statements like "Probably it may rain today", Chances of his visit to the university are very few", "It is possible that he may pass the examination with good marks". In the above statements, the words *probably, chance, possible* etc., convey the sense of uncertainty about the occurrence of some event. Ordinarily, it may appear that there cannot be any exact measurement for these uncertainties, but in Mathematics, we do have methods for calculating the degree of certainty of events in numerical values, provided certain conditions are satisfied.

1.2 PROBABILITY AS A SET FUNCTION

The theory of probability which closely relates the theory of sets, was proposed by A.N. Kolwogorov, a Russran mathematician in 1933. To understand the concept of "Probability as a set function, we need some basic terminology of various terms, like random experiment, descrete and continuous sample space, simple event, possible and impossible event. Now, we discuss all these terms stated above in the following sections.

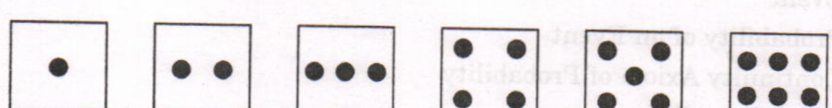
Random Experiment

When we perform experiments in science and engineering, repeatedly under very nearly identical conditions, we get almost the same result. Such experiments are called **deterministic experiments.**

There also exist experiments in which the results may not be essentially the same even if the experiment is performed under very nearly identical conditions. Such experiments are called **random experiments.** If we toss a coin, we may get 'head' or 'tail'. This is a random experiment. Throwing of a die is also a random experiment as any of the six faces of the die may come up. In this experiment, there are six possibilities (1 or 2 or 3 or 4 or 5 or 6).

A random experiment is also known as a **probabilistic experiment** or as a **non-deterministic experiment.**

Remark 1. A *die* is a small cube used in gambling. On its six faces, dots are marked as shown below :



Numbers on a die

Plural of the word die is *dice*. The outcome of throwing a die is the number of dots on its upper most face.

Remark 2. A *pack of cards* consists of four suits called *Spades, Hearts, Diamonds* and *Clubs.* Each suit consists of 13 cards, of which nine cards are numbered from 2 to 10, an ace, a king, a queen and a jack (or knave). Spades and clubs are *black faced cards*, while hearts and diamonds are *red faced cards*. The kings, queens and jacks are called **face cards.**

Sample Space

The **sample space** of a random experiment is defined as the set of all possible outcomes of the experiment. The possible outcomes are called **sample points.** The

NOTES

sample space is generally denoted by the letter S. The number of sample points in the sample space S is denoted by $n(S)$. A sample space is called *discrete*. If it contains only finitely many points and which can be arranged into a simple sequence $\omega_1, \omega_2, \ldots\ldots$. We list the sample space of some random experiments.

Random Experiment	Sample Space
1. Tossing of an unbiased coin	S = {H, T}
2. Tossing of unbiased coin twice	S = {HH, HT, TH, TT}

In S, the sample point HT represent 'head' on first toss and tail on second toss.

3. Tossing of two unbiased coins	S = {HH, HT, TH, TT}

In S, the sample point HT represent 'head' on first coin and 'tail' on second coin.

4. Tossing of three unbiased coins	S = {HHH, HHT, HTH, THH, HTT, THT, TTH, TTT}*
5. A family of two children	S = {BB, BG, GB, GG}.

In S, the sample point BG represent elder child, 'Boy' and younger child 'Girl'.

6. Throwing of a fair die	S = {1, 2, 3, 4, 5, 6}.
7. Throwing of two fair die	S = {(1,1) (1, 2), (1, 3), (1,4), (1, 5), (1, 6), (2, 1), (2, 2), (2, 3), (2, 4), (2, 5), (2, 6), (3, 1), (3, 2), (3, 3), (3, 4), (3, 5), (3, 6), (4, 1), (4, 2), (4, 3), (4, 4), (4, 5), (4, 6), (5, 1), (5, 2), (5, 3), (5, 4), (5, 5), (5, 6), (6, 1), (6, 2), (6, 3), (6, 4) (6, 5), (6, 6)}.

In S, the sample point (1, 2) represent the situation that 1 appeared on the first die and 2 on the second die. This sample space has 6 × 6 = 36 sample points.

Note. The sample space of a random experiment is either finite or infinite. In our present course, we shall restrict ourselves only to finite sample spaces.

ILLUSTRATIVE EXAMPLES

Example 1. *A bag contains 4 red balls. What is the sample space if the random experiment consists of choosing :*

(i) 1 ball *(ii) 2 balls*
(iii) 3 balls *(iv) 4 balls ?*

Sol. Let the red balls be denoted by R_1, R_2, R_3 and R_4.

(*i*) In this experiment, one ball is drawn.

No. of elements in $S = {}^4C_1 = 4$

$\therefore \quad S = \{R_1, R_2, R_3, R_4\}.$

(*ii*) In this experiment, two balls are drawn.

No. of elements in $S = {}^4C_2 = \dfrac{4 \times 3}{1 \times 2} = 6$

$\therefore \quad S = \{R_1R_2, R_2R_3, R_3R_4, R_4R_1, R_1R_3, R_2R_4\}.$

(*iii*) In this experiment ; three balls are drawn.

No. of elements in $S = {}^4C_3 = \dfrac{4 \times 3 \times 2}{1 \times 2 \times 3} = 4$

$\therefore \quad S = \{R_1R_2R_3, R_2R_3R_4, R_3R_4R_1, R_4R_1R_2\}.$

*If n coins are tossed, then the number of elements in its sample space is equal to 2^n.

NOTES

(*iv*) In this experiment, 4 balls are drawn.

No. of elements in $S = {}^4C_4 = 1$

$\therefore \quad S = \{R_1R_2R_3R_4\}.$

Example 2. *A bag contains 3 red and 2 white balls. What is the sample space if the random experiment consists of drawing (i) one ball (ii) two balls from the bag ?*

Sol. Let R_1, R_2, R_3, W_1, W_2 represent the red and white balls.

Total no. of balls = 3 + 2 = 5

(*i*) No. of elements in the sample space, $S = {}^5C_1 = 5$

$$S = \{R_1, R_2, R_3, W_1, W_2\}.$$

(*ii*) No. of elements in the sample space $= {}^5C_2 = \frac{5 \times 4}{1 \times 2} = 10$

$$S = \{R_1R_2, R_1R_3, R_1W_1, R_1W_2, R_2R_3, R_2W_1, R_2W_2, R_3W_1, R_3W_2, W_1W_2\}.$$

EVENTS AS SUBSETS OF SAMPLE SPACE

1.3 EVENT

An **event** of a random experiment is defind as a subset of the sample space of the random experiment. If the outcome of an experiment is an element of an event A, we say that the event A has occurred. An event is called an **elementary** (*or* **simple**) **event,** if it contains only one sample point. In the experiment of rolling a die, the event A of getting '3' is a simple event. We write A′ = {3}. An event is called an **impossible event,** if it can never occur. In the above example, the event B = {7} of getting '7' is an impossible event. On the other hand, an event which is sure to occur is called a **sure event.** In the above example of rolling a die, the event C of getting a number less than 7 is a sure event. A sure event is also called a **certain event.**

ILLUSTRATIVE EXAMPLES

Example 1. *There are 2 children in a family. Find the events that :*

(*i*) *both children are boys* (*ii*) *only one of the children is a girl*

(*iii*) *there is at least one girl* (*iv*) *the older child is a boy.*

Sol. Here S = {BB, BG, GB, GG}.

(*i*) Let A be the event that both children are boys.

$\therefore \quad$ A = {BB}.

(*ii*) Let A be the event that only one of the children is a girl.

$\therefore \quad$ A = {BG, GB}.

(*iii*) Let A be the event that there is at least one girl.

$\therefore \quad$ A = {BG, GB, GG}.

(*iv*) Let A be the event that the older child is a boy.

$\therefore \quad$ A = {BB, BG}.

Example 2. *An urn contains 4 red and 6 yellow balls. Two balls are drawn at random from the urn. Find the number of elements in the sample space. Also find the number of elements in the event of getting :*

(*i*) *both balls red* (*ii*) *one ball red and one ball yellow*

(*iii*) *both balls yellow.*

NOTES

Sol. Total number of balls = 4 + 6 = 10.

No. of elements is S = no. of ways of selecting 2 balls out of 10 balls

= no. of combinations of 10 things taking 2 at a time

$$= {}^{10}C_2 = \frac{10 \times 9}{1 \times 2} = 45.$$

4 Red
6 Yellow

(*i*) Let A be the events of getting 2 red balls.

$\therefore$ No. of elements in A $= {}^4C_2 \times {}^6C_0 = \frac{4 \times 3}{1 \times 2} \times 1 = 6.$

(*ii*) Let B be the event of getting one red and one yellow balls.

$\therefore$ No. of elements in B $= {}^4C_1 \times {}^6C_1 = 4 \times 6 = 24.$

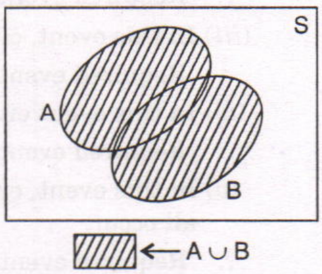
(*iii*) Let C be the event of getting 2 yellow balls.

$\therefore$ No. of elements in C $= {}^4C_0 \times {}^6C_2 = 1 \times \frac{6 \times 5}{1 \times 2} = 15.$

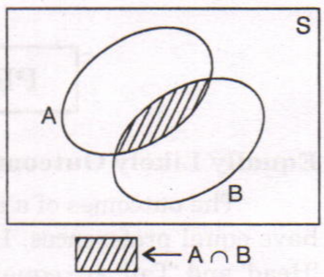
Algebra of Events

We know that the events of a random experiments are sets, being subsets of the sample space. Thus, we can use set operations to form new events.

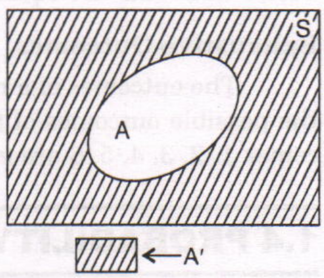
Let A and B be any two events associated with a random experiment.



The event of occurrence of either A or B or both is written as 'A or B' and is denoted by the subset $A \cup B$ of the sample space. In other words, $A \cup B$ represents the event of occurrence of at least one of A and B.



The event of occurrence of both A and B is written as 'A and B' and is denoted by the subset $A \cap B$ of the sample space. For simplicity the event $A \cap B$ is also denoted by 'AB'.



The event of non-occurrence of event A is written as 'not A' and is denoted by the set A′, which is the complement of set A. The event A′ is called the **complementary event** of the event A.

ILLUSTRATIVE EXAMPLES

Example 1. *A coin is tossed twice. If A denotes the event "number of heads is odd" and B denotes the event "number of tails is at least one". Find the cases favourable to the event $A \cap B$.*

Sol. Here S = {HH, HT, TH, TT}

A = {HT, TH}, B = {HT, TH, TT}

$\therefore$ $A \cap B$ = event of occurring both A and B

= {HT, TH}.

NOTES

Example 2. *A, B and C are three events associated with the sample space S of a random experiment. If A, B and C also denote the subsets of S representing these events, what are the sets representing the events :*

(*i*) *Out of the three events, only A occurs*

(*ii*) *Out of the three events, not more than two occur*

(*iii*) *Out of the three events, only one occurs*

(*iv*) *Out of the three events, exactly two events occur*

(*v*) *Out of the three events, at least two events occur.*

Sol. (*i*) In this event, A occurs and B, C do not occur?

$\therefore$ Required event = $A \cap B' \cap C'$.

(*ii*) In this event, all events do not occur simultaneously.

$\therefore$ Required event = $(A \cap B \cap C)'$.

(*iii*) In this event, either only A occur or only B occur or only C occur.

$\therefore$ Required event = $(A \cap B' \cap C') \cup (A' \cap B \cap C') \cup (A' \cap B' \cap C)$.

(*iv*) In this event either only A, B occur or only B, C occur or only A, C occur.

$\therefore$ Required event = $(A \cap B \cap C') \cup (A' \cap B \cap C) \cup (A \cap B' \cap C)$.

(*v*) In this event, either only A, B occur or only B, C occur or only A, C occur or all occur.

$\therefore$ Required event = $(A \cap B \cap C') \cup (A' \cap B \cap C) \cup (A \cap B' \cap C) \cup (A \cap B \cap C)$.

PROBABILITY OF AN EVENT

Equally Likely Outcomes

The outcomes of a random experiment are called **equally likely,** if all of these have equal preferences. In the experiment of tossing a unbiased coin, the outcomes, 'Head' and 'Tail' are equally likely.

Exhaustive Outcomes

The outcomes of a random experiment are called **exhaustive,** if these cover all the possible outcomes of the experiment. In the experiment of rolling a die, the outcomes 1, 2, 3, 4, 5, 6 are exhaustive.

1.4 PROBABILITY OF AN EVENT

Suppose in a random experiment, there are n exhaustive, equally likely outcome. Let A be an event and there are m outcomes (cases) favourable to the happening of it. The **probability** P(A) of the happening of the event A is defined as:

$$\mathbf{P(A) = \frac{Total\ number\ of\ cases\ favourable\ to\ the\ happening\ of\ A}{Total\ number\ of\ exhaustive\ equally\ likely\ cases} = \frac{m}{n}}.$$

It may be observed from this definition, that $0 \leq m \leq n$.

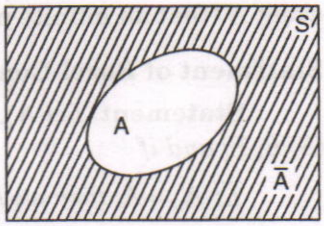
$$\therefore \quad 0 \leq \frac{m}{n} \leq 1 \quad \text{or} \quad \mathbf{0 \leq P(A) \leq 1.}$$

NOTES

The number of cases favourable of the non-happening of the event A is $n - m$.

$\therefore \quad P(\text{not } A) = \frac{n-m}{n} = \frac{n}{n} - \frac{m}{n} = 1 - \frac{m}{n} = 1 - P(A).$

$\therefore \quad P(A) + P(\text{not } A) = 1$ *i.e.*, $P(A) + P(\bar{A}) = 1$



If A is a sure event, then $P(A) = \frac{n}{n} = 1$ and if A happens to be an impossible event, then $P(A) = \frac{0}{n} = 0$.

From now onward, we shall always assume that the outcomes of any given random experiment are equally likely unless the contrary is stated explicitely.

Probability as a Set Function

A Purely mathematical definition of probability cannot give us the actual value of P(A), the probability of occurrence of the event A and hence, P(A) must be considered as a function defined on all events. To define P(A) as a set function, we need a *domain space* which is the σ-field B of the events, generated by S, a *range space* which is the closed interval [0, 1] on the real line and a rule which assigns a value to every element of the domain space B.

Definition. P(A) is the probability as a set function defined on a σ-field B of events if the following axioms are satisfied.

(*i*) For each $A \in B$, P(A) is defined and $P(A) \geq 0$ (axiom of non-negativity)

(*ii*) $P(S) = 1$ (axiom of certainty)

(*iii*) For $<A_n>$, the sequence of disjoint events, we have

$$P\left(\bigcup_{i=j}^{n} A_i\right) = \sum_{i=j}^{n} P(A_i) \qquad \text{(axiom of additivity)}$$

The set function 'P' defined on σ-field B, satisfying all the above axioms is called probability measure. The set (S, B, P) is known as the probability space. In our practical problems, S is taken as finite and the σ-field B is taken as the collection of all subsets of S.

1.5 CONTINUITY AXIOM OF PROBABILITY

Let $B_1, B_2, ..., B_n, ...$ be a countable sequence of events satisfying the following:

(*i*) $B_{i+1} \subset (B_i \ (i = 1, 2, 3, ...)$

(*ii*) $\bigcap_{i=1}^{\infty} B_i = \phi$

In words, (*i*) means that each succeeding event implies the proceeding event and (*ii*) means that the simultaneous occurrence of all B_i is an impossible event, then

$$\underset{n \to \infty}{\text{Lt}} P(B_n) = 0.$$

1.6 BOREL-CANTELLI LEMMA

Let $<A_n>$ be a sequence of events. Let A be the event "that an infinite number of An occur". This means that $\omega \in A$ if $\omega \in A_n$ for an infinite number of values of n (but not

NOTES

necessarily every n). The set of such ω is denoted by Lt sup An or $\overline{\text{Lt}}\ An$. Thus, the event A "that an infinite number of An occur" is denoted by $\overline{\text{Lt}}\ A_n$.

Statement of Borel-Contelli Lemma

Statement. *Let $A_1, A_2, \ldots, A_n$ be a sequence of events on the probability space (S, B, P) and if*

$\overline{Lt}\ A_n = A$, then we have

$$P(A) = 0, \text{ provided } \sum_{n=1}^{\infty} P(A_n) < \infty$$

In words, it states that if $\sum_{n=1}^{\infty} P(A_n)$ converges with probability one, then only a finite number of $A_1, A_2, \ldots, A_n$ occur.

Proof. Given $\overline{\text{Lt}}\ A_n = A \Rightarrow$ we can write

$$A = \overline{\text{Lt}}\ A_n = \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_n \Rightarrow A \subset \bigcup_{m=n}^{\infty} A_m \text{ for every } n.$$

This implies for each n, $P(A) \le \sum_{m=n}^{\infty} P(A_m)$

Now we are given that $\sum_{n=1}^{\infty} P(A_n)$ is convergent, therefore, $\sum_{m=n}^{\infty} P(A_m)$, being the remainder term of a convergent series tends to zero as $n \to \infty$. Hence $P(A) \le \sum_{m=n}^{\infty} P(A_m) \to 0$ as $n \to \infty \Rightarrow P(A) = 0$. Hence the theorem.

Converse of Borel-Contelli Lemma

Statement. *Let $\langle A_n \rangle$ be a sequence of independent events on the probability space (S, B, P) and if Lt $A_n = A$. Then*

$$P(A) = 1, \text{ provided } \sum_{n=1}^{\infty} P(A_n) = \infty$$

Proof. Let $\overline{A}_n$ denotes the complement of A_n, *i.e.*, $\overline{A}_n = S - A_n$. For any m, $n (m > n)$, we have

$$\bigcap_{k=n}^{\infty} \overline{A}_k \subset \bigcap_{k=n}^{m} \overline{A}_k$$

$$\Rightarrow P\left(\bigcap_{k=n}^{\infty} \overline{A}_k\right) \le P\left(\bigcap_{k=n}^{m} \overline{A}_k\right) \qquad | A \subset B \Rightarrow P(A) \le P(B)$$

$$= \prod_{k=n}^{m} P(\overline{A}_k) \qquad \begin{array}{|l} A_n, A_{n+1}, \ldots A_m \text{ are independent events} \\ \Rightarrow \overline{A}_n, \overline{A}_{n+1}, \ldots, \overline{A}_m \text{ are also independent events} \end{array}$$

$$\Rightarrow \quad P\left(\bigcap_{k=n}^{\infty} \overline{A}_k\right) \le \prod_{k=n}^{m} (1-P(A_k)) \le \prod_{k=n}^{m} e^{-P(A_k)} = e^{-\sum_{k=n}^{m} P(A_k)} \qquad ...(1)$$

$$\Big|\ 1-x \le e^{-x} \forall x \ge 0$$

$$\text{Now } \sum_{k=1}^{\infty} P(A_k) = \infty \qquad |\text{ given}$$

$$\Rightarrow \quad \sum_{k=n}^{m} P(A_k) \to \infty \text{ as } m \to \infty$$

$$\therefore \quad e^{-\sum_{k=n}^{m} P(A_k)} \to 0 \text{ as } m \to \infty$$

$$\therefore \quad \text{From (1), } P\left(\bigcap_{k=1}^{\infty} \overline{A}_k\right) = 0 \qquad ...(2)$$

$$\text{Also,} \qquad A = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k$$

$$\Rightarrow \qquad \overline{A} = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} \overline{A}_k \qquad |\text{ De-Morgan's Law}$$

$$\Rightarrow \qquad P(\overline{A}) \le \sum_{n=1}^{\infty} P\left(\bigcap_{k=n}^{\infty} \overline{A}_k\right) = 0 \Rightarrow P(\overline{A}) = 0 \qquad |\text{ Using (2)}$$

$$\Rightarrow \qquad 1 - P(A) = 0 \Rightarrow P(A) = 1$$

Hence the theorem.

1.7 ZERO-ONE LAW

Statement. *If $A_1, A_2, ..., A_n, ...$ are independent and if E belongs to the field generated by the class $(A_n, A_{n+1}, ...)$ for every n, then $P(E) = 0$ or 1.*

Proof. The proof directly follows from Borel-Cantelli lemma.

Example. *Find the probability that in a sequence of Bernouli trials with probability of success p for each trial, the pattern SFS, where S devotes success and F denotes failure, appears infinitely of ten ?*

Sol. Let A_n be the event that the trial number $n = k, k+1, k+2$, generates the sequence SFS ($k = 0, 1, 2, ...$). then the events A_n are not mutually independent but the sequence A_1, A_4, A_7, A_{10}, ... contains only mutually independent events. (As n_0 two of them do not depend upon the outcome of the same trials).

Also $\quad P(A_k) = P(SFS) = pqp = p^2q = pk$, is independent of k and hence the series $p_1 + p_4 + p_7 + ...$, diverges.

By converse of the Borel-cantelli theorems, the pattern SFS appears infinitely often with probability **one**.

NOTES

ILLUSTRATIVE EXAMPLES

Example 1. *Three coins are tossed simultaneously. Write the sample space and the probabilities of getting (i) no head and (ii) two heads.*

Sol. Here $S = \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}$

(*i*) P(no head) $= P(\{TTT\}) = \dfrac{n(|TTT|)}{n(S)} = \dfrac{1}{8}$.

(*ii*) P(two heads) $= P(\{HHT, HTH, THH\}) = \dfrac{3}{8}$.

Remark. For the sake of simplicity, P({TTT}) is written as P(TTT).

'Odds In Favour' And 'Odds Against' An Event

Let A be an event of a random experiment. The ratio $P(A) : P(\overline{A})$ is called the **odds in favour** of happening of the event A. The ratio $P(\overline{A}) : P(A)$ is called the **odds against** the happening of the event A.

Let odds in favour of an event A be $m : n$.

Let $P(A) = p$. $\therefore$ $p : 1 - p = m : n$

$$\Rightarrow \quad \frac{p}{1-p} = \frac{m}{n} \quad \Rightarrow \quad np = m - mp \quad \Rightarrow \quad p = \frac{m}{m+n} \text{ i.e., } P(A) = \frac{m}{m+n}.$$

$\therefore$ **If odds in favour of A are m : n, then** $\mathbf{P(A) = \dfrac{m}{m+n}}$.

Similarly, if odds against A are $m : n$ then odds in favour of A are $n : m$ and P(A) $= \dfrac{n}{n+m}$.

Remark. Odds in favour of event A are same as odds against the complement A′ of A and *vice versa*.

WORKING RULES FOR SOLVING PROBLEMS

I. Find the number of elements in the sample space S of the given random experiment. Write the sample space, if it is feasible to do so.

II. Out of the elements of the sample space, identify the elements which are favourable to the event, A (say), whose probability is required. Write the event A, if it is feasible to do so.

III. Divide $n(A)$ by $n(S)$. This is equal to the required probability, P(A).

Example 2. *A bag contains 5 white, 7 black and 8 red balls. A ball is drawn at random. Find the probability of getting :*

(*i*) *red ball* (*ii*) *non-white ball* (*iii*) *white ball or black ball.*

Sol. No. of white balls = 5

No. of black balls = 7

No. of red balls = 8

$\therefore$ Total number of balls = 5 + 7 + 8 = 20.

5 White
7 Black
8 Red

(*i*) Let R = event of getting red ball

$\therefore$ P(red ball) $= P(R) = \dfrac{n(R)}{n(S)} = \dfrac{8}{20} = \dfrac{2}{5}$.

(*ii*) Let W = event of getting white ball

$\therefore$ P(non-white ball) = $P(\overline{W}) = \dfrac{\text{no. of non-white balls}}{\text{total no. of balls}} = \dfrac{7+8}{20} = \dfrac{15}{20} = \dfrac{3}{4}$.

(*iii*) Let A = event of getting a white ball or a black ball

$\therefore$ P(white ball or black ball) = $P(A) = \dfrac{n(A)}{n(S)} = \dfrac{5+7}{20} = \dfrac{12}{20} = \dfrac{3}{5}$.

Example 3. *Find the probability that a leap year, selected at random, will contain 53 sundays.*

Sol. There are 366 days in a leap year. Now, $366 = 7 \times 52 + 2$.

$\therefore$ This leap year will contain at least 52 sundays. The possible combinations for the remaining two days are :

(*i*) sunday and monday
(*ii*) monday and tuesday
(*iii*) tuesday and wednesday
(*iv*) wednesday and thursday
(*v*) thursday and friday
(*vi*) friday and saturday
(*vii*) saturday and sunday.

Let A be the event of getting 53 sundays in the leap year. Therefore, only those combinations will be favourable to the event A which contain 'sunday'.

$\therefore$ The combinations (*i*) and (*vii*) are favourable to the happening of A.

$$\therefore \qquad P(A) = \frac{2}{7}.$$

ADDITION THEOREMS

Mutually Exclusive Events

Two events associated with a random experiment are said to be **mutually exclusive,** if both cannot occur together in the same trial. In the experiment of throwing a die, the events A = {1, 4} and B = {2, 5, 6} are mutually exclusive events. In the same experiment, the events A = {1, 4} and C = {2, 4, 5, 6} are not mutually exclusive because, if 4 appear on the die, then it is favourable to both events A and C. The definition of mutually exclusive events can also be extended to more than two events. We say that more than two events are mutually exclusive, if the happening of one of these, rules out the happening of all other events. The events A = {1, 2}, B = {3} and C = {6}, are mutually exclusive in connection with the experiment of throwing a single die.

n events $A_1, A_2, \ldots\ldots, A_n$ associated with a random experiment are said to **mutually exclusive events** if $A_i \cap A_j = \phi$ for all *i*, *j* and $i \neq j$.

For example, let a pair of dice be thrown and let A, B, C be the events that the sum is 7, sum is 8, sum is greater than 10 respectively,

$\therefore$ A = {(1, 6), (2, 5), (3, 4), (4, 3), (5, 2), (6, 1)}

B = {(2, 6), (3, 5), (4, 4), (5, 3), (6, 2)}

and C = {(5, 6), (6, 5), (6, 6)}

The events A, B and C are mutually exclusive.

NOTES

Example. *Two dice are rolled. A is the event that the sum of the numbers shown on the two dice is 5. B is the event that at least one of the dice show up a 3. Are the two events A and B (i) mutually exclusive, (ii) exhaustive ? Give arguments in support of your answer.*

Sol. Here S = {(1, 1), (1, 2), (1, 3) ,......., (6, 5), (6, 6)}

We have A = {(1, 4), (2, 3), (3, 2), (4, 1)}

and B = {(3, 1), (3, 2), (3, 3), (3, 4), (3, 5), (3, 6), (1, 3), (2, 3), (4, 3), (5, 3), (6, 3)}

(*i*) The events A and B are not mutually exclusive because (2, 3), (3, 2) lies in both A and B.

(*ii*) The events A and B are not exhaustive because there are same elements in S (like (1, 1), (1, 2), (1, 5)) which are neither in A nor in B.

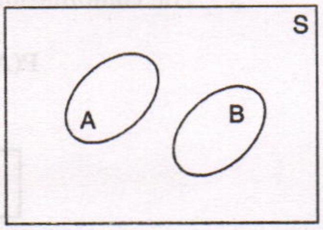
Addition Theorem (For Mutually Exclusive Events)

Statement. *If A and B are two mutually exclusive events associated with a random experiment, then*

$$P(A \cup B) = P(A) + P(B).$$

Proof. Let n be the total number of exhaustive, equally likely cases of the experiment.

Let m_1 and m_2 be the number of cases favourable to the happening of the events A and B respectively.



$$\therefore \quad P(A) = \frac{m_1}{n} \quad \text{and} \quad P(B) = \frac{m_2}{n}.$$

Since the events are given to be mutually exclusive, therefore, there cannot be any sample point common to both events A and B.

$\therefore$ The event A $\cup$ B can happen in exactly $m_1 + m_2$ ways.

$$\therefore \quad P(A \cup B) = \frac{m_1 + m_2}{n} = \frac{m_1}{n} + \frac{m_2}{n} = \mathbf{P(A) + P(B).}$$

Hence, $\mathbf{P(A \cup B) = P(A) + P(B).}$

Remark. This theorem can also be extended to more than two events.

Theorem I. *If A, B, C are three mutually exclusive events associated with a random experiment, then $P(A \cup B \cup C) = P(A) + P(B) + P(C)$.*

Proof. A, B, C are mutually exclusive elements.

$$\therefore \quad A \cap B = \phi, B \cap C = \phi, A \cap C = \phi.$$

We have $A \cap (B \cup C) = (A \cap B) \cup (A \cap C) = \phi \cup \phi = \phi$

$\therefore$ The events A and B $\cup$ C are mutually exclusive.

$\therefore$ By *addition theorem*, we have

$$P(A \cup (B \cup C)) = P(A) + P(B \cup C)$$
$$= P(A) + (P(B) + P(C))$$

(By applying *addition theorem*, for the *m.e.* events B and C)

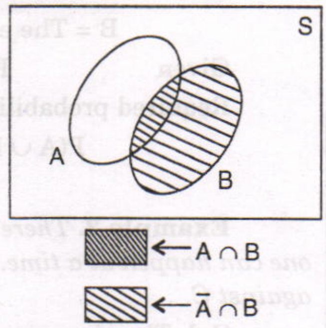
$$\therefore \quad \mathbf{P(A \cup B \cup C) = P(A) + P(B) + P(C).}$$

NOTES

Theorem II. *If A and B be two events associated with a random experiment, then show that :*

(i) $P(\overline{A} \cap B) = P(B) - P(A \cap B)$

(ii) $P(A \cap \overline{B}) = P(A) - P(A \cap B)$.



Proof. *(i)* $(\overline{A} \cap B) \cup (A \cap B) = (\overline{A} \cup A) \cap B$

$$= S \cap B = B \quad \text{...(1)}$$

Also $(\overline{A} \cap B) \cap (A \cap B) = (\overline{A} \cap A) \cap (B \cap B)$

$$= \phi \cap B = \phi$$

$\therefore$ The events $\overline{A} \cap B$ and $A \cap B$ are *m.e.*

By *addition theorem,* we have

$$P((\overline{A} \cap B) \cup (A \cap B)) = P(\overline{A} \cap B) + P(A \cap B)$$

$$\Rightarrow \quad P(B) = P(\overline{A} \cap B) + (P \cap B) \qquad \text{(Using (1))}$$

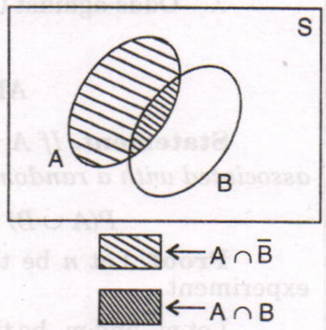
$$\Rightarrow \quad \mathbf{P(\overline{A} \cap B) = P(B) - P(A \cap B).}$$

(ii) $(A \cap \overline{B}) \cup (A \cap B) = A \cap (\overline{B} \cup B)$

$$= A \cap S = A \quad \text{...(1)}$$

Also $(A \cap \overline{B}) \cap (A \cap B) = (A \cap A) \cap (\overline{B} \cap B)$

$$= A \cap \phi = \phi .$$



$\therefore$ The events $A \cap \overline{B}$ and $A \cap B$ are *m.e.*

$\therefore$ By *addition theorem,* we have

$$P((A \cap \overline{B}) \cup (A \cap B)) = P(A \cap \overline{B}) + P(A \cap B)$$

$$\Rightarrow \quad P(A) = P(A \cap \overline{B}) + P(A \cap B)$$

(Using (1))

$$\Rightarrow \quad \mathbf{P(A \cap \overline{B}) = P(A) - P(A \cap B).}$$

WORKING RULES FOR SOLVING PROBLEMS

I. Find the number of elements in the sample space S of the given random experiment. Write the sample space, if it is feasible to do so.

II. Designate the events as A and B the probability of whose union, $A \cup B$ is to be found out. Out of the elements of the sample space, identify the elements which are favourable to the events A and B both.

III. Make sure that the events A and B are mutually exclusive, *i.e.,* the set $A \cap B$ is the empty set.

IV. Use $P(A \cup B) = P(A) + P(B)$. This gives the required probability.

ILLUSTRATIVE EXAMPLES

Example 1. *The probability that a bread prepared in a hotel is well baked is 0.81 and that it will have sufficient proteins is 0.54. Again the probability that it has both is 0.78. Find the probability that a well-baked bread will contain sufficient proteins.*

NOTES

Sol. Let A = The event that a bread is baked

B = The event that a bread has sufficient proteins.

Given $P(A) = 0.81, P(B) = 0.54, P(A \cap B) = 0.18.$

Required probability that a well baked bread will contain sufficient proteins

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$
$$= 0.81 + 0.54 - 0.18 = 1.17.$$

Example 2. *There are three events A, B, C one of which must happen and only one can happen at a time. The odds are 8 to 3 against A, 5 to 2 against B, find the odds against C.*

Sol. The given events A, B and C are mutually exclusive and exhaustive.

$\therefore \quad P(A \cup B \cup C) = P(A) + P(B) + P(C)$ and $P(A \cup B \cup C) = 1$

$$\therefore \quad P(A) + P(B) + P(C) = 1 \quad \text{...(1)}$$

Odds against A are 8 : 3. $\therefore \quad P(A) = \dfrac{3}{8+3} = \dfrac{3}{11}$

Odds against B are 5 : 2 $\therefore \quad P(B) = \dfrac{2}{5+2} = \dfrac{2}{7}$

$$\therefore \quad (1) \Rightarrow \quad \frac{3}{11} + \frac{2}{7} + P(C) = 1 \quad \Rightarrow \quad P(C) = 1 - \frac{3}{11} - \frac{2}{7} = \frac{77-21-22}{77} = \frac{34}{77}$$

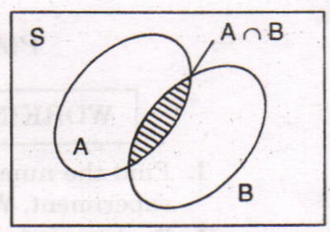
$$\therefore \quad \text{Odds against C } = P(\overline{C}) : P(C) = 1 - \frac{34}{77} : \frac{34}{77} = \frac{43}{77} : \frac{34}{77} = 43 : 34.$$

ADDITION THEOREM (GENERAL)

Statement. *If A and B are two events not necessarily mutually exclusive, associated with a random experiment, then*

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

Proof. Let n be the total number of exhaustive, equally likely cases of the experiment.



Let m_1 and m_2 be the number of cases favourable to the happening of the events A and B respectively.

$$\therefore \quad P(A) = \frac{m_1}{n} \quad \text{and} \quad P(B) = \frac{m_2}{n}.$$

Since the events are given to be not necessarily mutually exclusive, there may be some sample points common to both events A and B.

Let m_3 be the number of these common sample points. m_3 will be zero in case A and B are mutually exclusive.

$$\therefore \quad P(A \cap B) = \frac{m_3}{n}.$$

The m_3 sample points, which are common to both events A and B, are included in the events A and B separately.

$\therefore$ Number of sample points in the event $A \cup B = m_1 + m_2 - m_3$.

m_3 is subtracted from $m_1 + m_2$ to avoid counting of common sample points twice.

NOTES

$\therefore \quad P(A \cup B) = \frac{m_1 + m_2 - m_3}{n} = \frac{m_1}{n} + \frac{m_2}{n} - \frac{m_3}{n} = P(A) + P(B) - P(A \cap B).$

Hence, $\mathbf{P(A \cup B) = P(A) + P(B) - P(A \cap B).}$

Corollary. If events A and B happen to be mutually exclusive events, then $P(A \cap B) = 0$ and in this case *addition theorem* implies

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = P(A) + P(B) - 0 = P(A) + P(B)$$

This is the same as the *addition theorem* for mutually exclusive events.

Remark. This theorem can also be extended to more than two events.

Theorem I. *If A, B, C are three events associated with a random experiment, then*

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(B \cap C) - P(A \cap C) + P(A \cap B \cap C).$$

Proof. Considering '$B \cup C$' as one event and applying addition theorem to the events 'A' and '$B \cup C$', we have

$$P(A \cup B \cup C) = P(A \cup (B \cup C))$$
$$= P(A) + P(B \cup C) - P(A \cap (B \cup C))$$
$$= P(A) + (P(B) + P(C) - P(B \cap C)) - P((A \cap B) \cup (A \cap C))$$

(By using *addition theorem* for the events B and C)

$$= P(A) + P(B) + P(C) - P(B \cap C) - (P(A \cap B) + P(A \cap C) - P((A \cap B) \cap (A \cap C)))$$

(By applying *addition theorem* for the events $A \cap B$ and $A \cap C$)

$$= P(A) + P(B) + P(C) - P(B \cap C) - P(A \cap B) - P(A \cap C) + P(A \cap B \cap C)$$

$(\because \quad (A \cap B) \cap (A \cap C) = (A \cap A) \cap B \cap C = A \cap B \cap C)$

$$\therefore \quad \mathbf{P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(B \cap C) - P(A \cap C) + P(A \cap B \cap C).}$$

WORKING RULES FOR SOLVING PROBLEMS

I. Find the number of elements in the sample space S of the given random experiment. Write the sample space, if it is feasible to do so.

II. Designate the events as A and B the probability of whose union, $A \cup B$ is to be found out. Out of the elements of the sample space, identify the elements which are favourable to the events A and B both.

III. Write the event $A \cap B$ and find its probability.

IV. Use $P(A \cup B) = P(A) + P(B) - (A \cap B)$. This gives the required probability.

ILLUSTRATIVE EXAMPLES

Example 1. *Find the probability of 4 turning up for at least once in two tosses of a fair die.*

Sol. Here $S = \{(1, 1), (1, 2), \ldots\ldots, (6, 5), (6, 6)\}$.

Let $\quad A$ = event of getting 4 on the first die and $\quad$ B = event of getting 4 on the second die.

$\therefore$ $\quad A = \{(4, 1), (4, 2), (4, 3), (4, 4), (4, 5), (4, 6)\}$

and $\quad B = \{(1, 4), (2, 4), (3, 4), (4, 4), (5, 4), (6, 4)\}$.

NOTES

$\therefore \quad P(A) = \frac{n(A)}{n(S)} = \frac{6}{36} = \frac{1}{6}$ and $P(B) = \frac{n(B)}{n(S)} = \frac{6}{36} = \frac{1}{6}$.

The events A and B are not *m.e.* because the sample (4, 4) is common to both.

$\therefore \quad A \cap B = \{(4, 4)\} \quad \therefore \quad P(A \cap B) = \frac{1}{36}$

By *addition theorem*, the required probability of getting four at least once is

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = \frac{1}{6} + \frac{1}{6} - \frac{1}{36} = \frac{11}{36}.$$

Example 2. *Two dice are thrown once. Find the probability of getting an even number on the first die or a total of 8.*

Sol. Here $S = \{(1, 1), (1, 2), ..., (6, 5), (6, 6)\}$.

Let $\quad$ A = event of getting an even number on the first die

and $\quad$ B = event of getting a total 8.

$\therefore \quad A = \{(2, 1), (2, 2), (2, 3), (2, 4), (2, 5), (2, 6), (4, 1), (4, 2), (4, 3), (4, 4), (4, 5), (4, 6), (6, 1), (6, 2), (6, 3), (6, 4), (6, 5), (6, 6)\}$

and $\quad B = \{(2, 6), (3, 5), (4, 4), (5, 3), (6, 2)\}$

$\therefore \quad P(A) = \frac{18}{36} = \frac{1}{2}$ and $P(B) = \frac{5}{36}$

The events A and B are not *m.e.* because the sample points (2, 6), (4, 4), (6, 2) are common to both.

$\therefore \quad A \cap B = \{(2, 6), (4, 4), (6, 2)\} \quad \therefore \quad P(A \cap B) = \frac{3}{36} = \frac{1}{12}$

By *addition theorem*, the required probability of getting an even number on the first die or a total 8 is

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = \frac{1}{2} + \frac{5}{36} - \frac{1}{12} = \frac{5}{9}.$$

Example 3. *A and B are two non-mutually exclusive events. If $P(A) = \frac{1}{4}$, $P(B) = \frac{2}{5}$ and $P(A \cup B) = \frac{1}{2}$, find the values of $P(A \cap B)$ and $P(A \cap B^c)$.*

Sol. We have $\quad P(A) = \frac{1}{4}, P(B) = \frac{2}{5}, P(A \cup B) = \frac{1}{2}$

By *addition theorem*, $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

$$\therefore \quad \frac{1}{2} = \frac{1}{4} + \frac{2}{5} - P(A \cap B)$$

$$\therefore \quad P(A \cap B) = \frac{1}{4} + \frac{2}{5} - \frac{1}{2} = \frac{5 + 8 - 10}{20} = \frac{3}{20}.$$

We have, $\quad (A \cap B^c) \cap (A \cap B) = A \cap (B^c \cap B) = A \cap \phi = \phi$

$\therefore$ The events $A \cap B^c$ and $A \cap B$ are mutually exclusive and

$(A \cap B^c) \cup (A \cap B) = A \cap (B^c \cup B) = A \cap S = A \quad$ (S is the sample space)

$\therefore$ By *addition theorem*, $\quad P(A) = P(A \cap B^c) + P(A \cap B)$.

NOTES

$$\therefore \quad \frac{1}{4} = P(A \cap B^c) + \frac{3}{20}$$

or $$P(A \cap B^c) = \frac{1}{4} - \frac{3}{20} = \frac{1}{10}.$$

Example 4. *A, B, C are events such that $P(A) = 0.3$, $P(B) = 0.4$, $P(C) = 0.8$, $P(A \cap B) = 0.08$, $P(A \cap C) = 0.28$, $P(A \cap B \cap C) = 0.02$. If $P(A \cup B \cup C) \geq 0.75$, then find the minimum value of the probability of the event $B \cap C$.*

Sol. We have $0.75 \leq P(A \cup B \cup C) \leq 1$.

Using *addition theorem,* we have

$$0.75 \leq P(A) + P(B) + P(C) - P(A \cap B) - P(B \cap C) - P(A \cap C) + P(A \cap B \cap C) \leq 1$$

$$\Rightarrow \quad 0.75 \leq 0.3 + 0.4 + 0.8 - 0.08 - P(B \cap C) - 0.28 + 0.02 \leq 1$$

$$\Rightarrow \quad 0.75 \leq 1.16 - P(B \cap C) \leq 1$$

$$\Rightarrow \quad 0.75 - 1.16 \leq - P(B \cap C) \leq 1 - 1.16$$

$$\Rightarrow \quad - 0.41 \leq - P(B \cap C) \leq - 0.16$$

$$\Rightarrow \quad 0.41 \geq P(B \cap C) \geq 0.16$$

i.e., $$0.16 \leq P(B \cap C) \leq 0.41.$$

$\therefore$ Minimum value of $P(B \cap C) = 0.16$.

INDEPENDENT EVENTS

Introduction

In the present chapter, we shall study the method of evaluating probabilities of events relating to independent events and independent experiments. We shall also study random variables and their probability distributions.

Conditional Probability

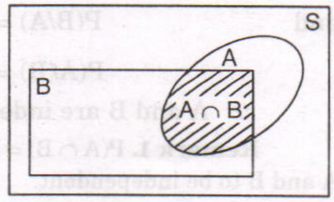
Let us consider the random experiment of throwing a die. Let A be the event of getting an odd number on the die.

$$\therefore \quad S = \{1, 2, 3, 4, 5, 6\} \text{ and } A = \{1, 3, 5\}.$$

$$\therefore \quad P(A) = \frac{3}{6} = \frac{1}{2}.$$

Let B = {2, 3, 4, 5, 6}. If, after the die is thrown, we are given the information that the event B has occurred, then the probability of event A will no more be $\frac{1}{2}$, because in this case, the favourable cases are three and the total number of possible outcomes will be five and not six. The probability of event A, with the condition that event B has happened will be 3/5. This conditional probability is denoted as P(A/B). Let us define the concept of conditional probability in a formal manner.

Let A and B be any two events associated with a random experiment. The probability of occurrence of event A when the event B has already occurred is called the **conditional probability** of A when B is given and is denoted as P(A/B). The conditional probability P(A/B) is meaningful only when $P(B) \neq 0$, *i.e.,* when B is not an impossible event.



NOTES

By definition,

P(A/B) = Probability of occurrence of event A when the event B as already occurred.

$$= \frac{\text{No. of cases favourable to B which are also favourable to A}}{\text{No. of cases favourable to B}}$$

$$\therefore \quad \mathbf{P(A/B)} = \frac{\textbf{No. of cases favourable to } \mathbf{A \cap B}}{\textbf{No. of cases favourable to B}}$$

Also,
$$P(A/B) = \frac{\dfrac{\text{No. of cases favourable to } A \cap B}{\text{No. of cases in the sample space}}}{\dfrac{\text{No. of cases favourable to B}}{\text{No. of cases in the sample space}}}$$

$$\therefore \quad \mathbf{P(A/B)} = \frac{\mathbf{P(A \cap B)}}{\mathbf{P(B)}}, \textbf{ provided } \mathbf{P(B) \neq 0.}$$

Similarly, we have

$$\mathbf{P(B/A)} = \frac{\mathbf{P(A \cap B)}}{\mathbf{P(A)}}, \textbf{ provided } \mathbf{P(A) \neq 0.}$$

Independent Events

Let A and B be two events associated with a random experiment. We have

$$P(B/A) = \frac{P(A \cap B)}{P(A)}, \text{ provided } P(A) \neq 0.$$

$$\therefore \quad P(A \cap B) = P(A)\, P(B/A).$$

In general P(B/A) may or may not be equal to P(B). When P(B/A) and P(B) are equal, then the events A and B are of special importance.

Two events associated with a random experiment are said to be **independent events** if the occurrence or non-occurrence of one event does not affect the probability of the occurrence of the other event. For example, the events A and B are independent events when P(A/B) = P(A) and P(B/A) = P(B).

Theorem II. *Let A and B be events associated with a random experiment. The events A and B are independent if and only if* $P(A \cap B) = P(A)\, P(B)$.

Proof. Let A and B be independent events.

$$\therefore \quad P(A \cap B) = \left(\frac{P(A \cap B)}{P(B)}\right) P(B) = P(A/B)\, P(B) \qquad \left(\because \quad P(A/B) = \frac{P(A \cap B)}{P(B)}\right)$$

$$= P(A)\, P(B) \qquad (\because \quad P(A/B) = P(A))$$

$$\therefore \quad P(A \cap B) = P(A)\, P(B).$$

Conversely, let $P(A \cap B) = P(A)\, P(B)$.

$$\therefore \quad P(A/B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)\, P(B)}{P(B)} = P(A)$$

and
$$P(B/A) = \frac{P(B \cap A)}{P(A)} = \frac{P(A \cap B)}{P(A)} = \frac{P(A)\, P(B)}{P(A)} = P(B).$$

$$\therefore \quad P(A/B) = P(A) \text{ and } P(B/A) = P(B).$$

$\therefore$ A and B are independent events.

Remark 1. $P(A \cap B) = P(A)\, P(B)$ is the necessary and sufficient condition for the events A and B to be independent.

NOTES

Remark 2. Let A and B be events associated with a random experiment.

(*i*) Let A and B be *m.e.* $\therefore$ $P(A \cap B) = 0$

$\therefore$ $P(A \cap B) \neq P(A)\,P(B)$ *i.e.*, A and B are not independent events.

$\therefore$ **Mutually exclusive events cannot be independent.**

(*ii*) Let A and B be independent.

$\therefore$ $P(A \cap B) = P(A)\,P(B)$ *i.e.*, $P(A \cap B) \neq 0$.

$\therefore$ A and B are not mutually exclusive events.

$\therefore$ **Independent events cannot be mutually exclusive.**

Important observation. If A and B be any two events associated with a random experiment, then their physical description is not sufficient to decide if A and B are independent events or not. A and B are declared to be independent events only when we have $P(A \cap B) = P(A)\,P(B)$.

Dependent Events

Let A and B be two events associated with a random experiment. If A and B are not independent events, then these are called **dependent events.**

$\therefore$ In case of dependent events, we have $\mathbf{P(A \cap B) = P(A)\,P(B/A)}$.

Multiplication Rule of Probability

If A and B be any two events associated with a random experiment, then we have

$$\mathbf{P(A \cap B) = P(A)\,P(B/A).}$$

This is called the **multiplication rule of probability.**

In particular, if the events A and B are independent, then the *multiplication rule of probability* becomes

$$\mathbf{P(A \cap B) = P(A)\,P(B).}$$

Theorem I. *If A, B, C are three events associated with a random experiment, then*

$$P(A \cap B \cap C) = P(A)\,P(B/A)\,P(C/A \cap B).$$

Proof. We have

$$P(C/A \cap B) = \frac{P(C \cap (A \cap B))}{P(A \cap B)} = \frac{P(A \cap B \cap C)}{P(A \cap B)}$$

$$\therefore \quad P(A \cap B \cap C) = P(A \cap B)\,P(C/A \cap B) \qquad ...(1)$$

Also, $$P(B/A) = \frac{P(B \cap A)}{P(A)} = \frac{P(A \cap B)}{P(A)}$$

$$\therefore \quad P(A \cap B) = P(A)\,P(B/A) \qquad ...(2)$$

(1) and (2) implies $\mathbf{P(A \cap B \cap C) = P(A)\,P(B/A)\,P(C/A \cap B)}$.

Definition. Three events A, B, C associated with a random experiment are called **independent** if A, B, C are pairwise independent and $P(A \cap B \cap C) = P(A)\,P(B)\,P(C)$.

Theorem II. *Let A and B be events associated with a random experiment. If A and B are independent, then show that the events (i)* $\overline{A}, B$ *(ii)* $A, \overline{B}$ *(iii)* $\overline{A}, \overline{B}$ *are also independent.*

Proof. The events A and B are independent.

$$\therefore \quad P(A \cap B) = P(A)\,P(B) \qquad ...(1)$$

NOTES

(*i*) $(A \cap B) \cap (\overline{A} \cap B) = (A \cap \overline{A}) \cap (B \cap B) = \phi \cap B = \phi$

and $(A \cap B) \cup (\overline{A} \cap B) = (A \cup \overline{A}) \cap B = S \cap B = B.$

$\therefore$ The events $A \cap B$ and $\overline{A} \cap B$ are *m.e.* and their union is B.

$\therefore$ By *addition theorem,* we have $P(B) = P(A \cap B) + P(\overline{A} \cap B)$.

$$\Rightarrow \quad P(\overline{A} \cap B) = P(B) - P(A \cap B) = P(B) - P(A)\,P(B) \qquad \text{(Using (1))}$$

$$= (1 - P(A))\,P(B) = P(\overline{A})\,P(B)$$

$\therefore$ $\mathbf{P(\overline{A} \cap B) = P(\overline{A})\ P(B)}$ *i.e.,* $\overline{A}$ and B are independent.

(*ii*) $(A \cap B) \cap (A \cap \overline{B}) = (A \cap A) \cap (B \cap \overline{B}) = A \cap \phi = \phi$

and $(A \cap B) \cup (A \cap \overline{B}) = A \cap (B \cup \overline{B}) = A \cap S = A.$

$\therefore$ The events $A \cap B$ and $A \cap \overline{B}$ are *m.e.* and their union is A.

$\therefore$ By *addition theorem,* we have

$$P(A) = P(A \cap B) + P(A \cap \overline{B}).$$

$$\Rightarrow \quad P(A \cap \overline{B}) = P(A) - P(A \cap B) = P(A) - P(A)\,P(B) \qquad \text{(Using (1))}$$

$$= P(A)(1 - P(B)) = P(A)P(\overline{B})$$

$\therefore$ $\mathbf{P(A \cap \overline{B}) = P(A)\ P(\overline{B})}$ *i.e.,* A and $\overline{B}$ are independent.

(*iii*) $(\overline{A} \cap B) \cap (\overline{A} \cap \overline{B}) = (\overline{A} \cap \overline{A}) \cap (B \cap \overline{B}) = \overline{A} \cap \phi = \phi$

and $(\overline{A} \cap B) \cup (\overline{A} \cap \overline{B}) = \overline{A} \cap (B \cup \overline{B}) = \overline{A} \cap S = \overline{A}.$

$\therefore$ The events $\overline{A} \cap B$ and $\overline{A} \cap \overline{B}$ are *m.e.* and their union is $\overline{A}$.

$\therefore$ By *addition theorem,* we have

$$P(\overline{A}) = P(\overline{A} \cap B) + P(\overline{A} \cap \overline{B}) \qquad ...(1)$$

$$\Rightarrow \quad P(\overline{A} \cap \overline{B}) = P(\overline{A}) - P(\overline{A} \cap B) = P(\overline{A}) - P(\overline{A})\,P(B) \qquad \text{(Using part (}i\text{))}$$

$$= P(\overline{A})\,(1 - P(B)) = P(\overline{A})\,P(\overline{B}).$$

$\therefore$ $\mathbf{P(\overline{A} \cap \overline{B}) = P(\overline{A})\ P(\overline{B})}$ *i.e.,* $\overline{A}$ and $\overline{B}$ are independent.

WORKING RULES FOR SOLVING PROBLEMS

I. Find the number of elements in the sample paper S of the given random experiment. Write the sample space, if it is feasible to do so.

II. Designate the events, as A and B whose 'independence' is to be checked. Out of the elements of the sample space, identify the elements which are favourable to the events A and B. Also find, the event $A \cap B$.

III. Find P(A), P(B) and $P(A \cap B)$.

IV. Find P(A) P(B). If P(A) P(B) is equal to $P(A \cap B)$, then declare that the given events A and B are independent events.

Example 1. *If A and B are independent events such that $P(A \cup B) = 0.6$ and $P(A) = 0.2$, find P(B).*

Sol. We have $P(A \cup B) = 0.6$ and $P(A) = 0.2$.

By *addition theorem,* we have

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

$$\Rightarrow \quad P(A \cup B) = P(A) + P(B) - P(A)\,P(B) \quad (\because \text{ A and B are independent})$$

$$\Rightarrow \quad 0.6 = 0.2 + P(B) - (0.2)\,P(B)$$

$$\Rightarrow \quad 0.4 = P(B)\,(1 - 0.2)$$

$$\Rightarrow \quad (0.8)\,P(B) = 0.4 \Rightarrow P(B) = \frac{0.4}{0.8} = \frac{1}{2} = 0.5.$$

NOTES

Example 2. *A pair of dice is thrown. A is the event : "the sum is 8" and B is the event : "at least one odd number is obtained". Show that the events A and B are dependent.*

Sol. Let S be the sample space.

$\therefore \quad S = \{(1, 1), (1, 2), ..., (6, 5), (6, 6)\}.$

$A = \{(2, 6), (3, 5), (4, 4), (5, 3), (6, 2)\}$

$B = \{(1, 1), (1, 2), (1, 3), (1, 4), (1, 5), (1, 6), (3, 1), (3, 2), (3, 3), (3, 4), (3, 5),$
$(3, 6), (5, 1), (5, 2), (5, 3), (5, 4), (5, 5), (5, 6), (2, 1), (2, 3), (2, 5), (4, 1),$
$(4, 3), (4, 5), (6, 1), (6, 3), (6, 5)\}.$

$$\therefore \quad P(A) = \frac{5}{36} \text{ and } P(B) = \frac{27}{36} = \frac{3}{4}$$

Also, $A \cap B = \{(3, 5), (5, 3)\}. \quad \therefore P(A \cap B) = \frac{2}{36} = \frac{1}{18}$

$$P(A)\,P(B) = \frac{5}{36} \times \frac{3}{4} = \frac{5}{48} \neq P(A \cap B).$$ The events A and B are dependent.

INDEPENDENT EXPERIMENTS

1.8 INDEPENDENT EXPERIMENTS

Two random experiments are said to be **independent** if, for every pair of events A and B where A is associated with the first and B with the second experiment, the probability of simultaneous occurrence of the events A and B, when the two experiments are performed, is equal to the product of the probabilities P(A) and P(B) calculated separately on the basis of two experiments.

The event 'A and B' of simultaneous occurrence of events A and B is denoted by $A \cap B$ or more briefly as AB.

Illustration. Let E_1 and E_2 be random experiments of throwing a die and tossing a coin respectively. Let S_1 and S_2 be their respective sample spaces.

$$\therefore \quad S_1 = \{1, 2, 3, 4, 5, 6\} \text{ and } S_2 = \{H, T\}.$$

If S represents the sample space of combined experiment of E_1 and E_2, then

$$S = \{1H, 2H, 3H, 4H, 5H, 6H, 1T, 2T, 3T, 4T, 5T, 6T\}.$$

The elementary events in each of E_1, E_2 and their combined experiment are equally likely.

NOTES

Let A = event of getting number less than 3, and
B = event of getting tail.

$$\therefore \quad P(A) = \frac{n(A)}{n(S_1)} = \frac{2}{6} = \frac{1}{3}, \quad P(B) = \frac{n(B)}{n(S_2)} = \frac{1}{2} \text{ and thus } P(A)\,P(B) = \frac{1}{3} \times \frac{1}{2} = \frac{1}{6}.$$

$$\text{Also, } AB = \{1T, 2T\} \text{ and } P(AB) = \frac{n(AB)}{n(S)} = \frac{2}{12} = \frac{1}{6}.$$

$\therefore$ P(AB) = P(A) P(B).

The above equality is not sufficient to infer that the experiments are independent, because by definition, the experiments would be independent if the equality P(AB) = P(A) P(B) holds for all possible events A in E_1 and B in E_2.

Next, we prove a theorem which will lay down a criterion for the independence of experiments.

Theorem III. *If the occurrence or non-occurrence of an event in one random experiment does not in any way affect the probability of the occurrence of any event in the other random experiment, then the experiments are independent.*

Proof. Since the occurrence or non-occurrence of an event in the first experiment is not affecting the probability of occurrence of any event in the second experiment, the sample spaces of the experiments are not affected by the events.

Let n_1 and n_2 be the numbers of elementary events in the first and second experiment respectively.

Let A and B be any events associated with first experiment and second experiment respectively.

Let m_1 be the number of cases favourable to the happening to the event A out of n_1 exhaustive, equally likely cases of the first experiment.

$$\therefore \quad P(A) = \frac{m_1}{n_1}$$

Let m_2 be the number of cases favourable to the happening of the event B out of n_2 exhaustive, equally likely cases of the second experiment.

$$\therefore \quad P(B) = \frac{m_2}{n_2}$$

By the **Fundamental principle of events,** the number of cases favourable to the happening of the event AB in this specified order is m_1m_2. Also the number of elementary events in the combined experiment is n_1n_2.

$$\therefore \quad P(AB) = \frac{m_1m_2}{n_1n_2} = \frac{m_1}{n_1} \cdot \frac{m_2}{n_2} = P(A)\,P(B).$$

$\therefore$ The experiments are independent.

Remark. If A and B are events associated with experiments which are not independent, then the probability of the event 'AB' is found by using the result :

P(AB) = P(A) P(B/A).

This result can also be extended to more than two experiments.

Important observation. If A and B be any two events associated with two different random experiments, then we may use the formula :

P(A in first experiment and B in second experiment)
= P(A in first experiment) . P(B in second experiment),

if on the basis of physical description of the random experiments, the occurrence or non-occurrence of an event in one random experiment does not affect the probability of occurrence of an event in the other random experiment.

NOTES

WORKING RULES FOR SOLVING PROBLEMS

I. If A and B are mutually exclusive events in a random experiment, then
$$P(A \cup B) = P(A) + P(B).$$

II. If A and B are events in a random experiment, then
$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

III. If A and B are independent events in a random experiment then
$$P(A \cap B) = P(A)\, P(B).$$

IV. If A and B are dependent events in a random experiment, then
$$P(A \cap B) = P(A)\, P(B/A).$$

V. If, on the basis of the physical description of two random experiments, the occurrence of an event in one random experiment does not affect the probability of the occurrence of an event in the other random experiment, we conclude that the random experiments are independent and make use of the result :

P(A in first experiment and B in second experiment)

= P(A in first experiment). P(B in second experiment).

ILLUSTRATIVE EXAMPLES

Example 1. *A and B appeared for an interview for two posts. Probability of A's rejection is 2/5 and that of B's selection is 4/7. Find the probability that only one of them is selected.*

Sol. The random experiments 'interview of A' and 'interview of B' are independent.

Let E = event that A is selected

and F = event that B is selected.

$$\therefore \qquad P(\overline{E}) = \frac{2}{5} \text{ and } P(F) = \frac{4}{7}.$$

Also, $P(E) = 1 - P(\overline{E}) = 1 - \frac{2}{5} = \frac{3}{5}$ and $P(\overline{F}) = 1 - P(F) = 1 - \frac{4}{7} = \frac{3}{7}.$

Required probability = P(only one is selected)

$= P(E\overline{F} \cup \overline{E}F) = P(E\overline{F}) + P(\overline{E}F)$

(Using **addition theorem**)

$= P(E)\, P(\overline{F}) + P(\overline{E})\, P(F)$

(Using **multiplication theorem**)

$$= \frac{3}{5} \times \frac{3}{7} + \frac{2}{5} \times \frac{4}{7} = \frac{9+8}{35} = \frac{17}{35}.$$

Example 2. *A husband and a wife appear in an interview for two vacancies for the same post. The probability of husband's selection is 3/5 and that of wife's selection is 1/5. Find the probability that :*

(*i*) *both are selected* (*ii*) *exactly one is selected*

(*iii*) *none is selected.*

Sol. The random experiments 'interview of husband' and interview of wife' are independent

Let $\quad H$ = event that 'Husband' is selected

and $\quad W$ = event that 'Wife' is selected.

$$\therefore \qquad P(H) = \frac{3}{5} \quad \text{and} \quad P(W) = \frac{1}{5}.$$

Also, $\quad P(\overline{H}) = 1 - P(H) = 1 - \frac{3}{5} = \frac{2}{5}$ and $P(\overline{W}) = 1 - P(W) = 1 - \frac{1}{5} = \frac{4}{5}$.

(*i*) P(both are selected) $= P(HW) = P(H)\,P(W)$

(Using **multiplication theorem**)

$$= \frac{3}{5} \times \frac{1}{5} = \frac{3}{25}$$

(*ii*) P(exactly one is selected) $= P(H\overline{W} \cup \overline{H}W)$

$= P(H\overline{W}) + P(\overline{H}W)$ (Using **addition theorem**)

$= P(H)P(\overline{W}) + P(\overline{H})P(W)$

(Using **multiplication theorem**)

$$= \frac{3}{5} \times \frac{4}{5} + \frac{2}{5} \times \frac{1}{5} = \frac{14}{25}$$

(*iii*) P(none is selected) $= P(\overline{H}\,\overline{W}) = P(\overline{H})\,P(\overline{W})$

(Using **multiplication theorem**)

$$= \frac{2}{5} \times \frac{4}{5} = \frac{8}{25}.$$

Note. In this example, it is worth while to note that $\frac{3}{25} + \frac{14}{25} + \frac{8}{25} = \frac{25}{25} = 1$.

This has happened because the events : both are selected, exactly one is selected and none is selected are *mutually exclusive* and *exhaustive*.

Example 3. *The odds in favour of one student passing a test are 3 : 7. The odds against another student passing it are 3 : 5. What is the probability that both pass the test ?*

Sol. Let $\quad A$ = event that first pass the test.

$$\therefore \qquad P(A) = \frac{3}{3+7} = \frac{3}{10}$$

Let $\quad B$ = event that second pass the test.

$$\therefore \qquad P(B) = \frac{5}{3+5} = \frac{5}{8}$$

The random experiments of results of students are independent.

$$\therefore \quad \text{P(both pass the test)} = P(AB) = P(A)\,P(B) = \frac{3}{10} \times \frac{5}{8} = \frac{3}{16}.$$

Example 4. *Three group of children contain respectively 3 girls and 1 boy, 2 girls and 2 boys, 1 girl and 3 boys. One child is selected at random from each group. Show that the chance that the group selected consists of 1 girl and 2 boys is* $\frac{13}{32}$.

Sol. There are three possibilities :

(*i*) Boy is selected from group-I and girls from group-II and group-III, or

(*ii*) Boy is selected from group-II and girls from group-I and group-III or

(*iii*) Boy is selected from group-III and girls from group-I and group-II.

I	II	III
1 B, 3 G	2 B, 2 G	3 B, 1 G

$\therefore$ Required probability of selecting a group of 1 boy and 2 girls

$$= \left(\frac{{}^1C_1}{{}^4C_1} \times \frac{{}^2C_1}{{}^4C_1} \times \frac{{}^1C_1}{{}^4C_1}\right) + \left(\frac{{}^2C_1}{{}^4C_1} \times \frac{{}^3C_1}{{}^4C_1} \times \frac{C_1}{{}^4C_1}\right) + \frac{{}^3C_1}{{}^4C_1} \times \frac{{}^3C_1}{{}^4C_1} \times \frac{{}^2C_1}{{}^4C_1}$$

$$= \frac{2}{3} + \frac{6}{64} + \frac{18}{64} = \frac{26}{64} = \frac{13}{32}.$$

Example 5. *A bag contains 5 white and 3 black balls and another bag contains 3 white and 4 black balls. If one ball is drawn from each bag, what is the probability that one is white and the other is black ?*

Sol. The random experiments 'drawing one ball from first bag' and 'drawing one ball from second bag' are independent.

Bag I	Bag II
5 White 3 Black	3 White 4 Black

Let W_i and B_i be the events of drawing white ball and black ball respectively from the ith bag. $i = 1, 2$.

Required probability $= P$ (one is white and one is black) $= P(W_1B_2 \text{ or } B_1W_2)$

$$= P(W_1B_2) + P(B_1W_2) = P(W_1)P(B_2) + P(B_1)P(W_2)$$

$$= \frac{5}{5+3} \times \frac{4}{3+4} + \frac{3}{5+3} \times \frac{3}{3+4} = \frac{20}{56} + \frac{9}{56} = \frac{29}{56}.$$

Example 6. *A can hit a target 3 times in 5 shots, B 2 times in 5 shots, C 3 times in 4 shots. Each fire a volley, what is the probability that 2 shots hit the target ?*

Sol. The three random experiments of hitting target by A, B, C are independent.

Let A, B, C also represent the events that A, B, C respectively hit the target.

$$\therefore \quad P(A) = \frac{3}{5}, \quad P(B) = \frac{2}{5} \quad \text{and} \quad P(C) = \frac{3}{4}$$

$$\therefore \quad P(\overline{A}) = 1 - \frac{3}{5} = \frac{2}{5}, \quad P(\overline{B}) = 1 - \frac{2}{5} = \frac{3}{5} \quad \text{and} \quad P(\overline{C}) = 1 - \frac{3}{4} = \frac{1}{4}.$$

Required probability $= P(2 \text{ shots hit the target}) = P(AB\overline{C} \text{ or } A\overline{B}C \text{ or } \overline{A}BC)$

$$= P(AB\overline{C}) + P(A\overline{B}C) + P(\overline{A}BC) = P(A)P(B)P(\overline{C}) + P(A)P(\overline{B})P(C) + P(\overline{A})P(B)P(C)$$

$$= \frac{3}{5} \times \frac{2}{5} \times \frac{1}{4} + \frac{3}{5} \times \frac{3}{5} \times \frac{3}{4} + \frac{2}{5} \times \frac{2}{5} \times \frac{3}{4} = \frac{6}{100} + \frac{27}{100} + \frac{12}{100} = \frac{45}{100} = \frac{9}{20}.$$

PROBABILITY DISTRIBUTION

Introduction

We have already studied a lot about frequency distributions. These distributions are based upon observations, *i.e.*, the frequencies for different values of the variable, under consideration, are based on actual observation. For example, if an unbiased coin is tossed 100 times, we may get head 57 times. Here, 57 is the observed frequency but theoretically we shall expect 'head', 50 times. In this section, we shall study *probability distributions* and *frequency distributions* which are based upon theoretical considerations.

1.9 RANDOM VARIABLE

NOTES

Let S be the sample space of a given random experiment. A real valued function 'x' defined on the sample space S is called a **random variable.**

Thus, if $s \in S$, then $x(s)$ is a unique real number.

Remark. The values of a random variable are real numbers, connected with the outcomes of the random experiment, under consideration.

In the random experiment of toss of two coins, if we define the random variable (x) as the *number of heads,* then the values of the random variable x are 0, 1, 1, 2 corresponding to the outcomes TT, TH, HT, HH respectively.

We write, $x(TT) = 0^*$, $x(TH) = 1$, $x(HT) = 1$, $x(HH) = 2$.

In case, there are three coins, then the values of this random variable are 0, 1, 1, 1, 2, 2, 2, 3 corresponding to the outcomes TTT, TTH, THT, HTT, HHT, HTH, THH, HHH respectively.

We can define any number of random variables on the same sample space. If x denotes the random variable, defined as the cube of the number of tails, in the experiment of toss of two coins, then we have

Sample points	HH	HT	TH	TT
x	$(0)^3 = 0$	$(1)^3 = 1$	$(1)^3 = 1$	$(2)^3 = 8$

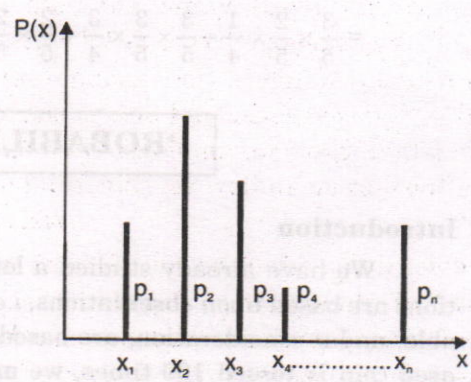
Random variables are of two types : (*i*) discrete random variable and (*ii*) continuous random variable.

(*i*) A random variable is called a **discrete random variable** if it can take only finitely many values. For example, in the experiment of drawing three cards from a pack of playing cards, the random variable "number of kings drawn" is a discrete random variable taking value either 0 or 1 or 2 or 3.

(*ii*) A random variable is called a **continuous random variable** if it can take any value between certain limits. For example, height, weight of students in a class are continuous random variables.

Probability Distribution of a Discrete Random Variable

Let x be a discrete random variable assuming values $x_1, x_2, x_3, ..., x_n$ corresponding to the various outcomes of a random experiment. If the probability of occurrence of $x = x_i$ is $P(x_i) = p_i$, $1 \le i \le n$ such that $p_1 + p_2 + p_3 + ... + p_n = 1$, then the function, $P(x_i) = p_i$, $1 \le i \le n$ is called the **probability function** of the random variable x and the set $\{P(x_1), P(x_2), P(x_3), ..., P(x_n)\}$ is called the **probability distribution** of x.



*__Why this step.__ $x(TT) = 0$, because the number of heads in the sample point 'TT' is zero.

. The graph of a probability distribution is also drawn as shown in the diagram. This is also known as a **bar-chart.**

Working Rules for Finding Probability Distribution

I. Identify the random variable and put it as x.

II. Find the possible values of x.

III. Find $P(x)$ for all possible values of x and write P.D. of x.

IV. Check that the sum of all probabilities in the P.D. is one. If this sum is not one, then some mistake is bound to have occurred in the calculation work. Remove the mistake and again verify that the sum of all probabilities is one.

1.10 DISTRIBUTION FUNCTION

Let X be a random variable. Define a function $F(x)$ by

$F(x) = P(X \le x) = P\{w : X(w) \le x\}, -\infty < x < \infty$ is called the distribution function of the random variable X. A distribution function, is also known as cumulative distribution function. We, some times, denote the distribution function of the random variable X, by $F_X(x)$.

The domain of $F(x)$ is $(-\infty, \infty)$ and range is $[0, 1]$.

Properties of Distribution Function

The properties of distribution function are discussed in the following theorems.

Theorem I. *Let F is the distribution function of the random variable X and if a < b, then*

$$P(a < X \le b) = F(b) - F(a)$$

Proof. Consider the events $a < X \le b$ and $X \le a$.

These two events are disjoint and their union is $X \le b$. By addition theorem of probability,

$$P(a < X \le b) + P(X \le a) = P(X \le b)$$

$$\Rightarrow \quad P(a < X \le b) = P(X \le b) - P(X \le a) = F(b) - F(a)$$

Theorem II. *If F is the distribution function of the random variable X and if a < b, then*

(i) $P(a \le X \le b) = F(b) - F(a) + P(X = a)$

(ii) $P(a < X < b) = F(b) - F(a) - P(X = b)$

(iii) $P(a \le X < b) = F(b) - F(a) + P(X = a) - P(X = b)$

(iv) If $P(X = a) = 0$ *and* $P(X = b) = 0$, *then (i), (ii) and (iii) have same probability* $F(b) - F(a)$

Proof. By above theorem,

$$P(a < X \le b) = F(b) - F(a)$$

$$(i)\ P(a \le X \le b) = P[(X = a) \cup (a < X \le b)]$$
$$= P(X = a) + P(a < X \le b) \qquad | \text{ Disjoint Events}$$
$$= P(X = a) + F(b) - F(a)$$

$$(ii)\ P(a < X < b) = P(a < X \le b) - P(X = b)$$
$$= F(b) - F(a) - P(X = b)$$

(*iii*) $P(a \le X < b) = P(a < X < b) + P(X = a)$
$= F(b) - F(a) + P(X = a) - P(X = b)$

(*iv*) If $P(X = a) = 0$ and $P(X = b) = 0$, then (*i*), (*ii*) and (*iii*) reduces to $F(b) - F(a)$.

NOTES

Theorem III. *If F is distribution function of one-dimensional random variable, then*

(*i*) $0 \le F(x) \le 1$ (*ii*) $F(x) \le F(y) \Rightarrow x \le y$

(*iii*) $F(-\infty) = 0$ (*iv*) $F(\infty) = 1$

Proof. (*i*) By definition of distribution function, the range of $F(x)$ is $[0, 1]$

$\therefore \quad 0 \le F(x) \le 1$

(*ii*) $F(x) = P(X \le x)$ and $F(y) = P(Y \le y)$

$\therefore \quad F(x) \le F(y) \Rightarrow P(X \le x) \le P(Y \le y)$

which holds if $x \le y$

(*iii*) Let S be the sample space of the random variable X, then

$$S = \left\{\bigcup_{n=1}^{\infty} (-n < X \le -n+1)\right\} \cup \left\{\bigcup_{n=0}^{\infty} (n < X \le n+1)\right\}$$

$$\Rightarrow \quad P(S) = \sum_{n=1}^{\infty} P(-n < X \le -n+1) + \sum_{n=0}^{\infty} P(n < X \le n+1)$$

$$\Rightarrow \quad 1 = \lim_{a\to\infty} \sum_{n=1}^{a} \{F(-n+1) - F(-n)\} + \lim_{b\to\infty} \sum_{n=0}^{b} \{F(n+1) - F(n)\}$$

$$= \lim_{a\to\infty} \{F(0) - F(-a)\} + \lim_{b\to\infty} \{F(b+1) - F(0)\}$$

$$= \{F(0) - F(-\infty)\} + \{F(\infty) - F(0)\}$$

$$\therefore \quad 1 = F(\infty) - F(-\infty) \qquad ...(1)$$

Since $-\infty < \infty$, $F(-\infty) \le F(\infty)$. Also $F(-\infty) \ge 0$ and $F(\infty) \le 1$

$$\therefore \quad 0 \le F(-\infty) \le F(\infty) \le 1 \qquad ...(2)$$

From (1) and (2), $F(-\infty) = 0$ and $F(\infty) = 1$.

Discrete Distribution Function or Probability Mass Function

If X is a discrete random variable with distinct values $x_1, x_2, ..., x_n$; then the function $p_X(x)$ defined by

$$p_X(x) = \begin{cases} P(X = x_i) = pi, & \text{if } x = x_i \\ 0 & , \quad x \ne x_i\ ; i = 1, 2, 3, ... \end{cases}$$

is known as Discrete distribution function or probability mass function (abbreviated as p.m.f.) of the random variable X.

The set of ordered pairs $\{(x_i, p(x_i))\ i = 1, 2, ... n, ...\}$ or $\{(x_1, p_1), (x_2, p_2), ... (x_n, p_n), ...\}$ specifies the probability distribution of the random variable X

Also $p(x_i) \ge 0\ \forall\ i$ and $\sum_{i=1}^{\infty} p(x_i) = 1$

Theorem IV. *If F is the distribution function of the discrete random variable X, then*

$$p(x_i) = P(X = x_i) = F(x_i) - F(x_{i-1}). \textit{ Interpret the result.}$$

NOTES

Proof. Let $x_1 < x_2 < x_3 < \ldots\ldots$, then

$$F(x_i) = P(X \le x_i) = \sum_{j=1}^{i} P(X = x_i) = \sum_{J=1}^{i} p(x_j)$$

$$F(x_{i-1}) = P(X \le x_{i-1}) = \sum_{j=1}^{i-1} P(X = x_{i-1}) = \sum_{j=1}^{i-1} p(x_j)$$

$$\therefore\ F(x_i) - F(x_{i-1}) = \sum_{j=1}^{i} p(x_j) - \sum_{j=1}^{i-1} p(x_j) = p(x_i).$$

Interpretation. Given the distribution function of the discrete random variable, its probability mass function can be obtained.

Continuous Distribution Function or Probability Density Function

Let X be a random variable and $f(x)$ be any continuous function of x so that $f(x)\,dx$ represents the probability that X falls in the infinitesimal interval $(x, x + dx)$ *i.e.,*

$$f_x(x)\,dx = P\,(x \le X \le x + dx),$$

then the function $f_x(x)$ is known as continuous distribution function or probability density function (abbreviated as p.d.f) or simply density function. Mathematically,

$$f_x(x) = \underset{\delta x \to 0}{\text{Lt}} \frac{P(x \le X \le x + \delta x)}{\delta x}$$

The probability density function (p.d.f.) $f_x(x)$ or $f(x)$ of a random variable X satisfies the following properties

(*i*) $f(x) \ge 0$ (*ii*) $\int_{-\infty}^{\infty} f(x)\,dx = 1$

Remark. In case of continuous random variables, the probability at a point is always zero. But, in case of discrete random variable, the probability at a point, *i.e.,* $P(x = c)$ is not zero for some fixed c. Hence

$$P\,(\alpha \le X \le \beta) = P\,(\alpha \le X < \beta) = P(\alpha < X \le \beta) = P\,(\alpha < X < \beta)$$

i.e., In case of continuous random variable, it does not matter whether we include the end points of the interval from α to β. However, this result is not true in case of discrete random variables. We are giving below some important formulae which will be frequently used in the present chapter. Let $f(x)$ denotes the probability density function of the random variable X, then

(*i*) **Arithmetic Mean** $= \int_a^b x\, f(x)\,dx$

(*ii*) **Harmonic Mean.** Harmonic mean H is given by : $\frac{1}{H} = \int_a^b \frac{1}{x} \cdot f(x)\,dx$

(*iii*) **Geometric Mean.** Geometric mean G is given by :

$$\log G = \int_a^b \log x.\, f(x)\,dx$$

(*iv*) The rth moment,

(*a*) μ'_r (about origin) $= \int_a^b x^r \cdot f(x)\,dx$

(*b*) μ'_r (about the point $x = A$) $= \int_a^b (x - A)^r \cdot f(x)\,dx$

(c) μ_r (about mean) $= \int_a^b (x - \text{mean})^r \cdot f(x)\, dx$

In particular,

NOTES

$$\mu'_1 \text{ (about origin)} = Mean = \int_a^b x\, f(x)\, dx \text{ and } \mu'_2 = \int_a^b x^2\, f(x)\, dx$$

Hence, $$\mu_2 = \mu'_2 - \mu_1'^2 = \int_a^b x^2\, f(x)\, dx - \left(\int_a^b x\, f(x)\, dx\right)^2$$

(v) **Median.** Median is the point which divides the entire distribution in two parts. In case of continuous distribution, median is the point which divides the area into two equal parts. Thus if M is the median, then

$$\int_a^M f(x)\, dx = \int_M^b f(x)\, dx = \frac{1}{2}$$

Thus solving $\int_a^M f(x)\, dx = \frac{1}{2}$ or $\int_M^b f(x)\, dx = \frac{1}{2}$

for M, we get the value of median.

(vi) **Mean Deviation.** Mean deviation about the mean μ_1' is given by :

$$\text{M.D} = \int_a^b |x - \text{mean}|\, f(x)\, dx$$

In general, mean deviation about an average 'A' is given by :

$$\text{M.D. about 'A'} = \int_a^b |x - \text{A}|\, f(x)\, dx$$

(vii) **Quartiles and Deciles.** Q_1 and Q_3 are given by the equations :

$$\int_a^{Q_1} f(x)\, dx = \frac{1}{4} \text{ and } \int_a^{Q_3} f(x)\, dx = \frac{3}{4}$$

D_i, *i*th decile is given by :

$$\int_a^{D_i} f(x)\, dx = \frac{i}{10};\ i = 1, 2, ..., 9$$

(viii) **Mode.** Mode is the value of x for which $f(x)$ is maximum. Mode is thus the solution of $f'(x) = 0$ and $f''(x) < 0$, provided it lies in $[a, b]$.

ILLUSTRATIVE EXAMPLES

Example 1. *Find the probability distribution of the random variable "number of heads" when :*

(i) two coins are tossed *(ii) one coin is tossed twice.*

Sol. (i) Let S be the sample space.

$\therefore$ S = {HH, HT, TH, TT}

Let x denotes the discrete random variable *"number of heads"*.

$\therefore$ The possible values of x are 0, 1, 2.

We have $P(x = 0) = P(\{TT\}) = \frac{1}{4}$

$$P(x = 1) = P(\{HT, TH\}) = \frac{2}{4} = \frac{1}{2}$$

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

NOTES

$\therefore$ The required probability distribution (P.D.) is

x	0	1	2
$P(x)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

(*ii*) Let H be the event of getting a head.

Let x denote the discrete random variable "*number of heads*" in two tosses.

$\therefore$ The possible values of x are 0, 1, 2.

We have $P(x = 0) = P(\overline{H}_1\overline{H}_2) = P(\overline{H}_1)\,P(\overline{H}_2) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$

$$P(x = 1) = P(H_1\overline{H}_2 \text{ or } \overline{H}_1H_2) = P(H_1\overline{H}_2) + P(\overline{H}_1H_2)$$

$$= P(H_1)\,P(\overline{H}_2) + P(\overline{H}_1)\,P(H_2) = \frac{1}{2} \times \frac{1}{2} + \frac{1}{2} \times \frac{1}{2} = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$$

$$P(x = 2) = P(H_1H_2) = P(H_1)\,P(H_2) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}.$$

$\therefore$ The required probability distribution (P.D.) is

x	0	1	2
$P(x)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

Example 2. *Two cards are drawn successively with replacement from a well shuffled pack of 52 cards. Find the probability distribution of number of queens.*

Sol. Let Q be the event of drawing a queen from the pack of cards.

Let x denotes the discrete random variable "*number of queens*" in two draws.

$\therefore$ The possible values of x are 0, 1, 2.

We have $P(x = 0) = P(\overline{Q}_1\overline{Q}_2) = P(\overline{Q}_1)\,P(\overline{Q}_2) = \frac{48}{52} \times \frac{48}{52} = \frac{144}{169}$

$$P(x = 1) = P(Q_1\overline{Q}_2 \text{ or } \overline{Q}_1Q_2) = P(Q_1\overline{Q}_2) + P(\overline{Q}_1Q_2)$$

$$= P(Q_1)\,P(\overline{Q}_2) + P(\overline{Q}_1)\,P(Q_2) = \frac{4}{52} \times \frac{48}{52} + \frac{48}{52} \times \frac{4}{52} = \frac{24}{169}$$

$$P(x = 2) = P(Q_1Q_2) = P(Q_1)\,P(Q_2) = \frac{4}{52} \times \frac{4}{52} = \frac{1}{169}.$$

$\therefore$ The required probability distribution (P.D.) is

x	0	1	2
$P(x)$	$\frac{144}{169}$	$\frac{24}{169}$	$\frac{1}{169}$

Mean and Variance of a Random Variable

We know the method of finding the mean and variance of frequency distributions. In a frequency distribution, we have frequencies corresponding to different values of the variable. Similarly, in a probability distribution, we have probabilities corresponding to different admissible values of the discrete random variable.

NOTES

Now, we shall extend the idea of mean and variance for probability distributions.

Let x be a discrete random variable assuming values $x_1, x_2, \ldots\ldots, x_n$ with respective probabilities $p_1, p_2, \ldots\ldots, p_n$ with $p_1 + p_2 + \ldots\ldots + p_n = 1$.

$$\text{We define, mean } (\mu) = \frac{\sum_{i=1}^{n} p_i x_i}{\sum_{i=1}^{n} p_i} = \sum_{i=1}^{n} p_i x_i \qquad \left(\because \sum_{i=1}^{n} p_i = 1\right)$$

and
$$\text{variance} = \frac{\sum_{i=1}^{n} p_i (x_i - \mu)^2}{\sum_{i=1}^{n} p_i} = \sum_{i=1}^{n} p_i (x_i - \mu)^2 .$$

$$\text{We have } \sum_{i=1}^{n} p_i (x_i - \mu)^2 = \sum_{i=1}^{n} p_i (x_i^2 + \mu^2 + 2\mu x_i) = \sum_{i=1}^{n} p_i x_i^2 + \mu^2 \sum_{i=1}^{n} p_i - 2\mu \sum_{i=1}^{n} p_i x_i$$

$$= \sum_{i=1}^{n} p_i x_i^2 + \mu^2 .1 - 2\mu.\mu = \sum_{i=1}^{n} p_i x_i^2 - \mu^2 .$$

$$\therefore \qquad \textbf{Mean } (\mu) = \sum_{i=1}^{n} p_i x_i \quad \textbf{and} \quad \textbf{variance } (\sigma^2) = \sum_{i=1}^{n} p_i x_i^2 - \mu^2 .$$

In short, we write, $\mu = \Sigma px$ and variance $= \Sigma px^2 - \mu^2$.

The mean of random variable x is called the **expected value** of x and is denoted by $E(x)$. The mean and variance of a random variable are also referred to as the mean and variance of the corresponding P.D.

Remark. S.D. of probability distribution $= \sqrt{\text{variance}} = \sqrt{\Sigma px^2 - \mu^2}$.

WORKING RULES FOR SOLVING PROBLEMS

I. Identify the random variable (x) and its possible values $x_1, x_2, \ldots\ldots$

II. Find probabilities for all values of the variable x.

III. Draw table and find Σpx and Σpx^2.

IV. Find mean and variance by using the formulae $\mu = \Sigma px$ and variance $= \Sigma px^2 - \mu^2$.

Example 3. *A random variable x has the following probability distribution :*

x	-2	-1	0	1	2	3
$P(x)$	0.1	k	0.2	$2k$	0.3	k

(i) Find the value of k.

(ii) Calculate mean and variance of x.

NOTES

Sol. **Calculation of mean and variance**

x	p	px	px^2
-2	0.1	-0.2	0.4
-1	k	$-k$	k
0	0.2	0	0
1	$2k$	$2k$	$2k$
2	0.3	0.6	1.2
3	k	$3k$	$9k$
	$\Sigma p = 4k + 0.6$	$\Sigma px = 4k + 0.4$	$\Sigma px^2 = 12k + 1.6$

(*i*) In a P.D., we have $\Sigma p = 1$.

$\therefore \quad 4k + 0.6 = 1 \quad \text{or} \quad 4k = 0.4 \text{ or } k = 0.1.$

(*ii*) Mean $(\mu) = \Sigma px = 4k + 0.4 = 4(0.1) + 0.4 = 0.8.$

Variance $= \Sigma px^2 - \mu^2 = (12k + 1.6) - (0.8)^2 = (12(0.1) + 1.6) - 0.64 = 2.16.$

Example 4. *Find the mean and variance of the number of heads in the two tosses of a coin.*

Sol. Let x denotes the random variable, "*number of heads*" in the two tosses.

$\therefore$ The possible values of x are 0, 1, 2.

We have $P(x = 0) = P(\text{no head}) = P(\overline{H}_1\overline{H}_2) = P(\overline{H}_1)P(\overline{H}_2) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$

where H_i is the event of getting head in the ith toss, $i = 1, 2$.

$$P(x = 1) = P(\text{one head}) = P(H_1\overline{H}_2 \text{ or } \overline{H}_1 H_2) = P(H_1)P(\overline{H}_2) + P(\overline{H}_1)P(H_2)$$

$$= \left(\frac{1}{2} \times \frac{1}{2}\right) + \left(\frac{1}{2} \times \frac{1}{2}\right) = \frac{1}{2}$$

$$P(x = 2) = P(\text{both heads}) = P(H_1H_2) = P(H_1)P(H_2) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}.$$

Calculation of mean and variance

x	p	px	px^2
0	$\frac{1}{4}$	0	0
1	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$
2	$\frac{1}{4}$	$\frac{1}{2}$	1
	$\Sigma p = 1$	$\Sigma px = 1$	$\Sigma px^2 = \frac{3}{2}$

$\therefore \quad$ Mean $(\mu) = \Sigma px = 1$

Variance $= \Sigma px^2 - \mu^2 = \frac{3}{2} - (1)^2 = 0.5.$

NOTES

Example 5. *Two cards are drawn successively with replacement from a well shuffled pack of 52 cards. Find the mean and standard deviation for the number of aces.*

Sol. Let A be the event of getting an ace in a draw of one card.

$\therefore \quad P(A) = \frac{4}{52} = \frac{1}{13}$ and $P(\overline{A}) = 1 - \frac{1}{13} = \frac{12}{13}$.

Let x denotes the random variable *"no. of aces"*.

$\therefore \quad x$ can take values 0, 1, 2.

$$P(x = 0) = P(\text{no ace}) = P(\overline{A}_1\,\overline{A}_2) = P(\overline{A}_1)\,P(\overline{A}_2) = \frac{12}{13} \times \frac{12}{13} = \frac{144}{169}$$

$$P(x = 1) = P(\text{one ace}) = P(A_1\,\overline{A}_2 \text{ or } \overline{A}_1\,A_2) = P(A_1)P(\overline{A}_2) + P(\overline{A}_1)\,P(A_2)$$

$$= \left(\frac{1}{13} \times \frac{12}{13}\right) + \left(\frac{12}{13} \times \frac{1}{13}\right) = \frac{24}{169}$$

$$P(x = 2) = P(\text{both aces}) = P(A_1A_2) = P(A_1)P(A_2) = \frac{1}{13} \times \frac{1}{13} = \frac{1}{169}.$$

Calculation of mean and S.D.

x	p	px	px^2
0	$\frac{144}{169}$	0	0
1	$\frac{24}{169}$	$\frac{24}{169}$	$\frac{24}{169}$
2	$\frac{1}{169}$	$\frac{2}{169}$	$\frac{4}{169}$
	$\Sigma p = 1$	$\Sigma px = \frac{2}{13}$	$\Sigma px^2 = \frac{28}{169}$

$\therefore \quad \text{Mean} = \Sigma px = \frac{2}{13}$

$$\text{S.D.} = \sqrt{\Sigma px^2 - \mu^2} = \sqrt{\frac{28}{169} - \left(\frac{2}{13}\right)^2} = \sqrt{\frac{28}{169} - \frac{4}{169}} = \frac{\sqrt{24}}{13} = \frac{2\sqrt{6}}{13}.$$

Example 6. *Two cards are drawn simultaneously (or successively without replacement) from a well shuffled pack of 52 cards. Compute the variance of the number of aces.*

Sol. Let two cards be drawn simultaneously from the well shuffled pack of 52 cards. Let x denotes the random variable *"no. of aces"*.

$\therefore \quad x$ can take values 0, 1, 2.

$$P(x = 0) = P(\text{no ace}) = \frac{{}^{48}C_2}{{}^{52}C_2} = \frac{188}{221}$$

$$P(x = 1) = P(\text{one ace}) = \frac{{}^{4}C_1 \times {}^{48}C_1}{{}^{52}C_2} = \frac{32}{221}$$

$$P(x = 2) = P(\text{both aces}) = \frac{{}^{4}C_2}{{}^{52}C_2} = \frac{1}{221}.$$

NOTES

Calculation of variance

x	p	px	px^2
0	$\frac{188}{221}$	0	0
1	$\frac{32}{221}$	$\frac{32}{221}$	$\frac{32}{221}$
2	$\frac{1}{221}$	$\frac{2}{221}$	$\frac{4}{221}$
	$\Sigma p = 1$	$\Sigma px = \frac{34}{221}$	$\Sigma px^2 = \frac{36}{221}$

Variance $= \Sigma px^2 - \mu^2 = \Sigma px^2 - (\Sigma px)^2$

$$= \frac{36}{221} - \left(\frac{34}{221}\right)^2 = \frac{7956-1156}{48841} = \frac{6800}{48841} = \frac{400}{2873}.$$

Example 7. *The diameter of an electric cable, say X, is assumed to be a continuous random variable with p.d.f. $f(x) = 6x(1-x)$, $0 \le x \le 1$*

(a) Examine whether f(x) is a probability density function ?

(b) Determine b such that $P(X < b) = P(X > b)$

Sol. (*a*) For probability density function, we should have

(*i*) $f(x) \ge 0$, which is true for $0 \le x \le 1$

(*ii*) $\int_0^1 f(x)\,dx = 1$, we check it.

Here $\int_0^1 f(x)\,dx = \int_0^1 6x(1-x)\,dx = 6\int_0^1 (x - x^2)\,dx$

$$= 6\left(\frac{x^2}{2} - \frac{x^3}{3}\right)_0^1 = 6\left(\frac{1}{2} - \frac{1}{3}\right) = 6 \cdot \frac{1}{6} = 1$$

(*b*) $P(X < b) = P(X > b)$

$\Rightarrow \int_0^b f(x)\,dx = \int_b^1 f(x)\,dx$

$\Rightarrow 6\int_0^b x(1-x)\,dx = 6\int_b^1 x(1-x)\,dx$

$\Rightarrow \left|\frac{x^2}{2} - \frac{x^3}{3}\right|_0^b = \left|\frac{x^2}{2} - \frac{x^3}{3}\right|_b^1$

$\Rightarrow \left(\frac{b^2}{2} - \frac{b^3}{3}\right) = \left[\left(\frac{1}{2} - \frac{1}{3}\right) - \left(\frac{b^2}{2} - \frac{b^3}{3}\right)\right]$

$\Rightarrow 3b^2 - 2b^3 = (1 - 3b^2 + 2b^3) \Rightarrow 4b^3 - 6b^2 + 1 = 0$

$\Rightarrow (2b-1)(2b^2 - 2b - 1) = 0$

$\therefore 2b - 1 = 0 \Rightarrow b = \frac{1}{2}$ or $2b^2 - 2b - 1 = 0$

$$\Rightarrow \qquad b = \frac{2 \pm \sqrt{4+8}}{4} = \frac{1 \pm \sqrt{3}}{2}.$$

But $b \in (0, 1)$ $\therefore$ $b = \frac{1}{2}$.

NOTES

Example 8. *Let X be a continuous random variable with probability density function given by*

$f(x) = kx\,(2 - x),\ 0 \le x \le 2$

(i) find mean, variance, β_1 and β_2 and hence show that the distribution is symmetrical

(ii) Find mean deviation about mean

(iii) For this distribution, $\mu_{2n+1} = 0$

(iv) Find mode, median and harmonic mean of the distribution.

Sol. As $f(x)$ is a probability density function

$$\therefore \qquad \int_{-\infty}^{\infty} f(x)\,dx = 1 \ \Rightarrow\ \int_0^2 k\,x\,(2-x)\,dx = 1$$

$$\Rightarrow \quad k\left| 2\frac{x^2}{2} - \frac{x^3}{3} \right|_0^2 = 1$$

$$\Rightarrow \qquad k\left(4 - \frac{8}{3}\right) = 1$$

$$\Rightarrow \qquad \frac{4}{3}k = 1 \ \Rightarrow\ k = \frac{3}{4}$$

The rth moment about origin is given by

$$\mu'_r = \int_0^2 x^r\, f(x)\,dx = \frac{3}{4}\int_0^2 x^{r+1}(2-x)\,dx = \frac{3.2^{r+1}}{(r+2)(r+3)} \qquad \ldots(1)$$

$$(i)\ \text{Mean} = \mu'_1 = \frac{3.2^2}{3.4} = 1,\quad \mu'_2 = \frac{3.2^3}{4.5} = \frac{6}{5},\quad \mu'_3 = \frac{3.2^4}{5.6} = \frac{8}{5},$$

and $$\mu'_4 = \frac{3.2^5}{6.7} = \frac{16}{7}$$

Hence, variance $= \mu_2 = \mu'_2 - \mu_1'^2 = \frac{6}{5} - 1 = \frac{1}{5}$

$$\mu_3 = \mu'_3 - 3\mu'_2\,\mu'_1 + 2\mu_1'^3 = \frac{8}{5} - 3\,.\frac{6}{5}\,.1 + 2 = 0$$

$$\mu_4 = \mu'_4 - 4\mu'_3\,\mu'_1 + 6\mu_2'\,\mu_1'^2 - 3\mu_1'^4$$

$$= \frac{16}{7} - 4\,.\frac{8}{5}\,.1 + 6\,.\frac{6}{5}\,.1 - 3\,.1 = \frac{3}{35}$$

$$\therefore \qquad \beta_1 = \frac{\mu_3^2}{\mu_2^3} = 0 \quad \text{and} \quad \beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{3/35}{(1/5)^2} = \frac{15}{7}$$

Since $\beta_1 = 0$, the distribution is *symmetrical*.

NOTES

(*ii*) Mean deviation about mean is given as

$$= \int_0^2 |x-1|\, f(x)\, dx = \int_0^1 |x-1|\, f(x)\, dx + \int_1^2 |x-1|\, f(x)\, dx$$

$$= \frac{3}{4}\left[\int_0^1 (1-x)\, x\,(2-x)\, dx + \int_1^2 (x-1)\, x\,(2-x)\, dx\right]$$

$$= \frac{3}{4}\left[\int_0^1 (2x-3x^2+x^3)\, dx + \int_1^2 (3x^2-x^3-2x)\, dx\right]$$

$$= \frac{3}{4}\left[\left| x^2 - \frac{3\,.x^3}{3} + \frac{x^4}{4}\right|_0^1 + \left| 3.\frac{x^3}{3} - \frac{x^4}{4} - \frac{2x^2}{2}\right|_1^2\right] = \frac{3}{8}$$

(*iii*) $$\mu_{2n+1} = \int_0^2 (x-\text{mean})^{2n+1} f(x)\, dx = \frac{3}{4}\int_0^2 (x-1)^{2n+1}\, x\,(2-x)\, dx$$

$$= \frac{3}{4}\int_{-1}^1 t^{2n+1}(t+1)(1-t)\, dt = \frac{3}{4}\int_{-1}^1 t^{2n+1}(1-t^2)\, dt \text{ where } t = x-1$$

Hence, $u_{2n+1} = 0.$

(*iii*) Mode is the solution of $f(x) = 0$ and $f''(x) < 0$

Now, $f'(x) = \frac{3}{4}(2-2x) = 0 \;\Rightarrow\; x = 1$ and $f''(x) = \frac{3}{4}(-2) = -\frac{3}{2} < 0.$

Hence mode = 1

Also, Harmonic mean H is given by :

$$\frac{1}{\text{H}} = \int_0^2 \frac{1}{x} f(x)\, dx = \frac{3}{4}\int_0^2 (2-x)\, dx = \frac{3}{2} \;\Rightarrow\; \text{H} = \frac{2}{3}$$

Finally, If M is the median,

$$\int_0^{\text{M}} f(x)\, dx = \frac{1}{2} \;\Rightarrow\; \frac{3}{4}\int_0^{\text{M}} x\,(2-x)\, dx = \frac{1}{2}$$

$$\Rightarrow \left| x^2 - \frac{x^3}{3}\right|_0^{\text{M}} = \frac{2}{3}$$

$$\Rightarrow 3\text{M}^2 - \text{M}^3 = 2 \Rightarrow \text{M}^3 - 3\text{M}^2 + 2 = 0 \Rightarrow (\text{M}-1)(\text{M}^2 - 2\text{M} - 2) = 0$$

$$\therefore \quad \text{M} = 1 \text{ or } \text{M} = \frac{2 \pm \sqrt{4+8}}{2} = 1 \pm \sqrt{3}$$

The only value of M lying in [0, 2] is M = 1. Hence median is 1.

Two-dimensional Random Variable

Let X and Y be two random variables defined on the same sample space S, then the function (X, Y) that assigns a point in R^2 (= R × R), is called a two-dimensional random variable.

Joint Probability Mass Function

If(X, Y) is a two dimensional (discrete) random variable, then the joint probability mass function of X, Y, denoted by $p_{XY}(x, y)$, is defined as

NOTES

$$p_{XY}(x_i, y_i) = \begin{cases} P(X = x_i, Y = y_i), & (x_i, y_i) \in (X, Y) \\ 0 & \text{otherwise} \end{cases}$$

Note (*i*) $\sum_i \sum_i p_{XY}(x_i, y_i) = 1.$

(*ii*) A two-dimensional random variable is called discrete if it takes at most countable number of points of R^2.

Marginal Probability Mass Function

Let (X, Y) be a discrete two-dimensional random variable, then the probability distribution of X is obtained as follows.

$$\begin{aligned} p_X(x_i) &= P(X = x_i) \\ &= P(X = x_i \cap Y = y_1) + P(X = x_i \cap Y = y_2) + \ldots\ldots + P(X = x_i \cap Y = y_m) \\ &= p_{i1} + p_{i2} + p_{i3} + \ldots\ldots + p_{im} = \sum_{j=1}^{m} p_{ij} = p_i, \end{aligned}$$

and is known as marginal probability mass function or discrete marginal density function of X

Also $\quad \sum_{i=1}^{n} p_i = \sum_{i=1}^{n} \sum_{j=1}^{m} p_{ij} = 1$

Similarly, we can prove that

$$p_Y(y_i) = P(Y = y_i) = \sum_{j=1}^{4} p_{ij} = \sum_{j=1}^{4} p(x_i, y_j) = p_j$$

which is known as marginal probability mass function of Y.

Remark. A necessary and sufficient condition for the discrete random variables X and Y to be independent is that

$$P(X = x_i, Y = y_j) = P(X = x_i)\, P(Y = y_j) \text{ for all } (x_i, y_j) \in (X, Y)$$

Two-dimensional Distribution Function

Let $F_{XY}(x, y)$ denotes the distribution function of the two-dimensional random variable (X, Y), then we define

$$F_{XY}(x, y) = P(X \le x, Y \le y)$$

Properties of Two-dimensional Distribution Function

$F_{XY}(x, y)$ satisfies the following properties

Property I. *For the real numbers a, b, c, d*

$$P(a < X \le b, c < Y \le d) = F(b, d) - F(b, c) - F(a, d) + F(a, c)$$

where $a < b$; $c < d$ and $F_{XY} = F$

Proof. Define the events :

Now, A : {X ≤ a} ; B : {X ≤ b} ; C = {Y ≤ c}; D = {Y ≤ d} ; for $a < b$; $c < d$.

$$\begin{aligned} P(a < X \le b \cap c < Y \le d) &= P[(B - A) \cap (D - C)] \\ &= P[B \cap (D - C) - A \cap (D - C)] \quad \ldots(1) \end{aligned}$$

(By distributive property of sets)

Also $E \subset F \Rightarrow E \cap F = E$, then

$$P(F - E) = P(\overline{E} \cap F) = P(F) - P(E \cap F) = P(F) - P(E)$$

NOTES

Obviously $\quad A \subset B \;\Rightarrow\; [A \cap (D - C)] \subset [B \cap (D - C)]$ ...(2)

Using (2) in (1), we have

$$P\,(a < X \le b \cap c < Y \le d) = P\,[B \cap (D - C)] - P[A \cap (D - C)]$$
$$= P\,[(B \cap D) - (B \cap C)] - P\,[(A \cap D) - (A \cap C)]$$
$$= P\,(B \cap D) - P\,(B \cap C) - P\,(A \cap D) + P\,(A \cap C) \qquad ...(3)$$

[From (2), since $C \subset D \;\Rightarrow\; (B \cap C) \subset (B \cap D)$ and $(A \cap C) \subset (A \cap D)$]

We have : $P\,(B \cap D) = P\,[X \le b \cap Y \le d] = F\,(b, d)$.

Similarly $P\,(B \cap C) = F\,(b, c)$; $P\,(A \cap D) = F\,(a, d)$ and $P\,(A \cap C) = F\,(a, c)$

Substituting in (3), we get

$$P(a < X \le b \cap c < Y \le d) = F(b, d) - F(b, c) - F(a, d) + F(a, c)$$

Property II. ***$F_{XY}(x, y)$ is a monotonic non-decreasing function.***

Proof. Consider the real numbers a, b, c, d such that $a < b$, $c < d$, then

$$(X \le a, Y \le b) + P\,(a < X \le c, Y \le b) = (X \le c, Y \le b)$$

As the events on the L.H.S are mutually exclusive

$\therefore \quad F(a, b) + P\,(a < X \le c, Y \le b) = F(c, b)$

$\Rightarrow \quad F(c, b) - F(a, b) = P\,(a < X \le c, Y \le b) \ge 0$

$\Rightarrow \quad F(c, b) > F(a, b)$

Similarly, $F(a, d) \ge F(a, b)$.

Property III. $F\,(-\infty, y) = 0,$

$F\,(x, -\infty) = 0,$

$F\,(-\infty, \infty) = 1$

Also $\dfrac{\partial^2 F}{\partial x \partial y} = f(x, y)$, provided

$f(x, y)$ is continuous at (x, y)

1.11 MARGINAL DISTRIBUTION FUNCTION

From the joint distribution function $F_{XY}(x, y)$, we can obtain the marginal distribution functions $F_X(x)$ and $F_Y(y)$, as below.

$$F_X(x) = P(X \le x) = P\,(X \le x, y < \infty) = \underset{y \to \infty}{\text{Lt}}\; F_{XY}(x, y)$$
$$= F_{XY}(x, \infty)$$

Also $\quad F_Y(y) = P\,(Y \le y) = P(X < \infty, Y \le y) = \underset{x \to \infty}{\text{Lt}}\; F_{XY}(x, y) = F_{XY}(\infty, y)$

In case of jointly discrete random variables,

$$F_X(x) = \sum_y P(X \le x, Y = y)$$

$$F_Y(y) = \sum_x P(X = x, Y \le y)$$

In case of jointly continuous random variable,

$$F_X(x) = \int_{-\infty}^{x}\int_{-\infty}^{\infty} f_{XY}(x, y)\, dy\, dx$$

$$F_Y(y) = \int_{-\infty}^{y}\int_{-\infty}^{\infty} f_{XY}(x, y)\, dx\, dy$$

NOTES

Joint Density Function

From the joint distribution function $F_{XY}(x, y)$ of two-dimensional continuous random variables, we can obtain the joint probability density function as below

$$f_{XY}(x, y) = \frac{\partial^2 F(x, y)}{\partial x \partial y} = \underset{\substack{\delta x \to 0 \\ \delta y \to 0}}{\text{Lt}} \frac{P(x \le X \le x + \delta x, y \le Y \le y + \delta y)}{\delta x \delta y}$$

and is known as joint probability density function of X and Y. The marginal probability function of X and Y can be obtained as below.

$$f_X(x) = \begin{cases} \sum_y p_{XY}(x, y), & \text{(In case of descrete variables)} \\ \int_{-\infty}^{\infty} f_{XY}(x, y)\, dy, & \text{(In case of continuous variables)} \end{cases}$$

$$f_Y(y) = \begin{cases} \sum_x p_{XY}(x, y), & \text{(for discrete variable)} \\ \int_{-\infty}^{\infty} f_{XY}(x, y)\, dx, & \text{(for continuous variables)} \end{cases}$$

Also $$f_X(x) = \frac{d\,F_X(x)}{dx} = \int_{-\infty}^{\infty} f_{XY}(x, y)\, dy$$

$$f_X(y) = \frac{d\,F_Y(y)}{dy} = \int_{-\infty}^{\infty} f_{XY}(x, y)\, dx$$

(a) Convolution of Random Variables

Def. *Probability generating function (p.g.f.)*

Let $<a_n>$ be a sequence of real numbers such that

$$A(s) = a_0 + a_1 s + a_2 s^2 + a_3 s^3 + \ldots$$

$= \sum_{n=1}^{\infty} a_n s^n$ converges in some interval $(-s_0, s_0)$, then the sum function $A(s)$ is known as the generating function of the sequence $<a_n>$.

(b) Convolution of Two Random Variables

Let x and y are non-negative independent discrete random variables and if $P(s)$ and $R(s)$ are the corresponding probability generating function (p.g.f.) such that

$$P(s) = \sum_{k=0}^{\infty} p_k s^k \text{ where } p_k = P(x = k)$$

$$R(s) = \sum_{k=0}^{\infty} r_k s^k \text{ where } r_k = P(y = k)$$

Take $z = x + y$ and $P(z = k) = w_k$

where $w_k = p_0 r_k + p_1 r_{k-1} + p_2 r_{k-2} + \ldots\ldots + p_k r_0$, $k \ge 0$ then the sequence $<w_k>$ is known as convolution of the sequences $<p_k>$ and $<r_k>$ where

$$<w_k> = <p_k> * <r_k>$$

NOTES

Theorem. *If <p_k> and <r_k> are the sequences with p.g.f. P(s) and R(s) respectively and <w_k> is their convolution,*

then $w(s) = P(s)\,R(s)$ *where*

$$w(s) = \sum_{k=0}^{\infty} w_k s^k \text{ is the p.g.f. of the sum } x + y = z.$$

Proof. By definition of convolution of two random variables, we know that

$$w_k = p_0 r_k + p_1 r_{k-1} + \ldots\ldots + p_{k-1} r_1 + p_k r_0$$

∴ The probability generating function

for $z = x + y$ is given by

$$w(s) = \sum_{k=0}^{\infty} w_k s^k = \mathrm{P}(s)\,\mathrm{R}(s)$$

Hence the theorem.

Stochastic or Statistical Independence

Two random variables X and Y with joint p.d.f (p.m.f) $f_{XY}(x, y)$ and marginal p.d.f. (p.m.f) $f_X(x)$ and $f_Y(y)$ respectively are said to be stochastically independent or statistical Independent if

$$f_{XY}(x, y) = f_X(x)\, f_Y(y)$$

or
$$F_{XY}(x, y) = F_X(x), F_Y(y)$$

where $F_{XY}(x, y)$ is the joint distribution function and $F_X(x)$ and $F_Y(y)$ are the marginal distribution functions respectively.

ILLUSTRATIVE EXAMPLES

Example 9. *For the joint probability distribution of two random variables X and Y, given below*

(i) Find the marginal distributions of X and Y

(ii) Find the conditional distribution of X given that Y = 1 and that of Y given that X = 2

X/Y	*1*	*2*	*3*	*4*	*Total*
1	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	$\frac{10}{36}$
2	$\frac{1}{36}$	$\frac{3}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{9}{36}$
3	$\frac{5}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{8}{36}$
4	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	$\frac{5}{36}$	$\frac{9}{36}$
Total	$\frac{11}{36}$	$\frac{9}{36}$	$\frac{7}{36}$	$\frac{9}{36}$	*1*

Sol. (*i*) Let $p_X(x)$ denotes the marginal distribution of X,

Then, $$p_X(x) = \mathrm{P}(X = x) = \sum_y \mathrm{P}(X = x, Y = y)$$

NOTES

$$\therefore \quad P(X=1) = \sum_y P(X=1, Y=y)$$

$$= P(X=1, Y=1) + P(X=1, Y=2) + P(X=1, Y=3) + P(X=1, Y=4)$$

$$= \frac{4}{36} + \frac{3}{36} + \frac{2}{36} + \frac{1}{36} = \frac{10}{36}$$

$$P(X=2) = \sum_y P(X=2, Y=y) = \frac{1}{36} + \frac{3}{36} + \frac{3}{36} + \frac{2}{36} = \frac{9}{36}$$

$$P(X=3) = \sum_y P(X=3, Y=y) = \frac{5}{36} + \frac{1}{36} + \frac{1}{36} + \frac{1}{36} = \frac{8}{36}$$

$$P(X=4) = \sum_y P(X=4, Y=y) = \frac{1}{36} + \frac{2}{36} + \frac{1}{36} + \frac{5}{36} = \frac{9}{36}$$

Hence, the marginal distribution of X is given below:

X	1	2	3	4
$p_X(x)$	$\frac{10}{36}$	$\frac{9}{36}$	$\frac{8}{36}$	$\frac{9}{36}$

Similarly, the marginal distribution of Y is given below:

Y.	1	2	3	4
$p_Y(y)$	$\frac{11}{36}$	$\frac{9}{36}$	$\frac{7}{36}$	$\frac{9}{36}$

(*ii*) The conditional distribution of X given Y is given by

$$P(X=x \mid Y=y) = \frac{P(X=x, Y=y)}{P(Y=y)}$$

$$\therefore \quad P(X=1 \mid Y=1) = \frac{P(X=1, Y=1)}{P(Y=1)} = \frac{4/36}{11/36} = \frac{4}{11}$$

$$P(X=2 \mid Y=1) = \frac{P(X=2, Y=1)}{P(Y=1)} = \frac{1/36}{11/36} = \frac{1}{11}$$

$$P(X=3 \mid Y=1) = \frac{5/36}{11/36} = \frac{5}{11},$$

$$P(X=4 \mid Y=1) = \frac{1/36}{11/36} = \frac{1}{11}$$

$\therefore$ **The conditional distribution of X given Y = 1 is given below:**

X	1	2	3	4
$P(X=x \mid Y=1)$	$\frac{4}{11}$	$\frac{1}{11}$	$\frac{5}{11}$	$\frac{1}{11}$

NOTES

Similarly, the conditional distribution of Y for X = 2 is given below:

Y	1	2	3	4
$P(Y = y \mid X = 2)$	$\frac{1}{9}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{2}{9}$

Since, $P(Y = 1 \mid X = 2) = \dfrac{P(Y = 1, X = 2)}{P(X = 2)} = \dfrac{1/36}{9/36} = \dfrac{1}{9}$ etc.

Example 10. *If the joint probability density function of two random variables X and Y is given by*

$$f(x, y) = \begin{cases} \dfrac{1}{8}(6 - x - y), & 0 \le x < 2, \quad 2 \le y < 4 \\ 0 & , \quad \text{otherwise} \end{cases}$$

(i) Find $P(X < 1 \cap Y < 3)$

(ii) Find $P(X + Y < 3)$

(iii) $P(X < 1 \mid Y < 3)$

Sol. *(i)* $P(X < 1 \cap Y < 3) = \int_{-\infty}^{1}\int_{-\infty}^{3} f(x, y)\, dx\, dy$

$$= \int_0^1\int_2^3 \frac{1}{8}(6 - x - y)\, dx\, dy = \frac{1}{8}\int_0^1 \left| 6x - \frac{x^2}{2} - yx \right|_2^3 dy$$

$$= \frac{1}{8}\int_0^1 \left[\left(18 - \frac{9}{2} - 3y\right) - (12 - 2 - 2y)\right]_{dy}$$

$$= \frac{1}{8}\int_0^1 \left(6 - \frac{5}{2} - y\right) dy = \frac{1}{8}\left| \frac{7}{2}y - \frac{y^2}{2} \right|_0^1$$

$$= \frac{1}{8}\left(\frac{7}{2} - \frac{1}{2}\right) = \frac{1}{8} . \frac{6}{2} = \frac{3}{8}$$

(ii) $P(X + Y < 3) = \int_0^1 \int_2^{3-x} \frac{1}{8}(6 - x - y)\, dy\, dx$

$$= \frac{1}{8}\int_0^1 \left| 6y - xy - \frac{y^2}{2} \right|_2^{3-x} dx$$

$$= \frac{1}{8}\int_0^1 \left\{ \begin{array}{l} \left[6(3 - x) - x(3 - x) - \dfrac{1}{2}(3 - x)^2\right] \\ \qquad - [(12 - 2x - 2)] \end{array} \right\} dx$$

$$= \frac{1}{8}\int_0^1 \left(18 - 6x - 3x + x^2 - \frac{1}{2}(9 + x^2 - 6x) - 12 + 2x + 2\right) dx$$

$$= \frac{1}{8}\int_0^1 \left(\frac{7}{2} - 4x + \frac{x^2}{2}\right) dx$$

NOTES

$$= \frac{1}{8}\left| \frac{7}{2}x - 2x^2 + \frac{x^3}{6} \right|_0^1 = \frac{1}{8}\left(\frac{7}{2} - 2 + \frac{1}{6}\right) = \frac{1}{8}\left(\frac{21-12+1}{6}\right)$$

$$= \frac{1}{8} \cdot \frac{10}{6} = \frac{5}{24}$$

(*iii*) $P(X < 1 \mid Y < 3) = \dfrac{P(X < 1 \cap Y < 3)}{P(Y < 3)} = \dfrac{3/8}{5/8} = \dfrac{3}{5}$ | Using (*i*) and (*ii*)

Example 11. *The joint probability density function of two-dimensional random variable (X, Y) is given below:*

$$f(x, y) = \begin{cases} 2, & 0 < x < 1, \quad 0 < y < x \\ 0, & \text{otherwise} \end{cases}$$

(*i*) *Find the marginal density functions of X and Y*

(*ii*) *Find the conditional density function of Y given X = x and conditional density function of X given Y = y*

(*iii*) *Check for independence of X and Y.*

Sol. The marginal density function of X, denoted by $f_X(x)$, is given by

$$f_X(x) = \begin{cases} \int_{-\infty}^{\infty} f_{XY}(x, y)\, dy = \int_0^x 2dy = 2x, & 0 < x < 1 \\ 0, & \text{otherwise.} \end{cases}$$

Similarly,

$$f_Y(y) = \begin{cases} \int_{-\infty}^{\infty} f_{XY}(x, y)\, dx = \int_y^1 2dx = 2(1-y), & 0 < y < 1 \\ 0, & \text{otherwise.} \end{cases}$$

(*ii*) The conditional density function of Y given X $(0 < x < 1)$ is

$$f_{Y/X}(y \mid x) = \frac{f_{XY}(x, y)}{f_X(x)} = \frac{2}{2x} = \frac{1}{x},\ 1 < y < x$$

Similarly, the conditional density function of X given y $(0 < y < 1)$ is

$$f_{X \mid Y}(x \mid y) = \frac{f_{XY}(x, y)}{f_Y(y)} = \frac{2}{2(1-y)} = \frac{1}{1-y},\ y < x < 1$$

(*iii*) Here $f_{XY}(x, y) = 2$ ...(1)

$f_X(x)\, f_Y(y) = 2x \,.\, 2(1-y)$ ...(2)

from (1) and (2) $f_{XY}(x, y) \neq f_X(x)\, f_Y(y)$

$\therefore$ X and Y are not independent

Example 12. *The joint probability density function of two random variables X and Y is given below:*

$$f(x, y) = \frac{9(1+x+y)}{2(1+y)^4(1+x)^4},\ 0 \le x < \infty,\ 0 \le y < \infty$$

Find (*i*) The marginal distribution of X and Y

(*ii*) The conditional distribution of Y for X = x

Sol. Let $f_X(x)$ denotes the marginal distribution of X and $f_Y(x)$ be the marginal distribution of Y, Then,

NOTES

$$f_X(x) = \int_0^\infty f(x, y)\, dy = \frac{9}{2(1+x)^4} \int_0^\infty \frac{(1+y)+x}{(1+y)^4}\, dy$$

$$= \frac{9}{2(1+x)^4} \int_0^\infty \{(1+y)^{-3} + x(1+y)^{-4}\}\, dy \qquad ...(1)$$

$$= \frac{9}{2(1+x)^4} \left(\left| \frac{-1}{2(1+y)^2} \right|_0^\infty + x \left| \frac{-1}{3(1+y)^3} \right|_0^\infty \right)$$

$$= \frac{9}{2(1+x)^4} \cdot \left(\frac{1}{2} + \frac{x}{3} \right) = \frac{3}{4} \cdot \frac{3+2x}{(1+x)^4};\ 0 < x < \infty.$$

Similarly,

$$f_Y(y) = \int_0^\infty f(x, y)\, dx = \int_0^\infty \frac{9(1+x+y)}{2(1+y)^4 (1+x)^4}\, dx$$

$$= \frac{9}{2(1+y)^4} \int_0^\infty \left[(1+x)^{-3} + y(1+y)^{-4}\right] dy$$

$$= \frac{3}{4} \cdot \frac{3+2y}{(1+y)^4},\ 0 < y < \infty \qquad | \text{ Proceed from (1)}$$

Also, the conditional distribution of Y for X = x is given by

$$f_{XY}(Y = y \mid X = x) = \frac{f_{XY}(x, y)}{f_X(x)} = \frac{9(1+x+y)}{2(1+x)^4 (1+y)^4} \times \frac{4(1+x)^4}{3(3+2x)}$$

$$= \frac{6(1+x+y)}{(1+y)^4 (3+2x)},\ 0 < y < \infty.$$

Example 13. ***The joint probability density function of two random variables X and Y is given below:***

$$f_{XY}(x, y) = \begin{cases} \frac{1}{4}(1+xy), & |x| < 1, |y| < 1 \\ 0, & \text{otherwise} \end{cases}$$

Show that X and Y are not independent, but X^2 and Y^2 are independent

Sol. **Let $f_X(x)$ and $f_Y(y)$ be the marginal density function of the random variables X and Y, then**

$$f_X(x) = \int_{-1}^{1} f(x, y)\, dy = \frac{1}{4} \left| y + \frac{xy^2}{2} \right|_{-1}^{1} = \frac{1}{2},\ -1 < x < 1; \qquad ...(1)$$

Similarly, $$f_Y(y) = \int_{-1}^{1} f(x, y)\, dx = \frac{1}{2},\ -1 < y < 1 \qquad ...(2)$$

From (1), (2) $f_{X,Y}(x, y) \neq f_X(x)\, f_Y(y)$, X and Y are not independent.

Also $$P(X^2 \le x) = P(|X| \le \sqrt{x}) = \int_{-\sqrt{x}}^{\sqrt{x}} f_X(x)\, dx$$

$$= \int_{-\sqrt{x}}^{\sqrt{x}} \frac{1}{2}\, dx = 2 \cdot \frac{1}{2} \int_0^{\sqrt{x}} dx = \sqrt{x} \qquad ...(3)$$

NOTES

$$P(X^2 \le x \cap Y^2 \le y) = P(|X| \le \sqrt{x} \cap |Y| \le \sqrt{y})$$

$$= \int_{-\sqrt{x}}^{\sqrt{x}} \left[\int_{-\sqrt{y}}^{\sqrt{y}} f(u, v)\, dv \right] du = \sqrt{x}\sqrt{y}$$

$$= P(X^2 \le x) \,.\, P(Y^2 \le y) \qquad \text{[From (3)]}$$

Hence, X^2 and Y^2 are independent.

SUMMARY

- Deterministic experiments are those experiments which give almost the same result when performed under very nearly indentical conditions.
- Non-deterministic experiments are those experiments which do not give the same results when performed under very nearly indentical conditions. Non-deterministic experiments are also known as random experiment.
- Two events associated with a random experiment are said to be *independent* if the occurrence or non-occurrence of one event does not affect the probability of the occurrence of the other event.
- Two random variables X and Y are said to be stochastically independent if

  $$f_{xy}(x, y) = f_x(x)\, f_y(y)$$

  where $f_{xy}(x, y)$ is the joint *p.d.f.* (*p.m.f.*), $f_x(x)$ and $f_y(y)$ are the marginal *p.d.f.* (*p.m.f.*) respectively.

GLOSSARY

- **Sample Space.** The set of all possible outcomes of a random experiment is known as sample space. The possible outcomes are known as sample points.
- **Exhaustive Outcomes.** The outcomes of a random experiment are called exhaustive if these cover all the possible outcomes of the experiment.
- **Mutually Exclusive Events.** Two events associated with a random experiment are said to be mutually exclusive if both cannot occur together in the same trial.
- **Addition Theorem.** If A and B are two events, not necessarily mutually exclusive, associated with a random experiment, then

  $$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

- **Independent Events:** Two events A and B associated with a random experiment are said to be independent if $P(A \cap B) = P(A) \,.\, P(B)$, otherwise, they are said to be dependent.
- **Multiplication Theorem.** If A and B are two events associated with a random experiment, then

  $$P(A \cap B) = P(A)\, P(B/A)$$

- **Discrete Random Variable.** A random variable if it can take only finite values is said to be a discrete random variable, *e.g.*, no. of heads in the toss of two coin.
- **Continuous Random Variable.** A random variable if it can take any values between certain limits, is said to be a continuous random variable, *e.g.*, height of students.

REVIEW QUESTIONS

1. Three coins are tossed simultaneously. List the sample space for this random experiment.

2. A bag contains 4 red balls and 3 black balls. What is the sample space if the random experiment consists of drawing (*i*) one ball (*ii*) two balls from the bag ?

3. From a group of 3 boys and 2 girls, we select two children. What would be the sample space of this experiment ?

4. From a group of 3 boys and 2 girls, we select two children. What would be the sample space of this random experiment ? Also, write the events of getting (*i*) both girls (*ii*) both boys.

5. Two dice are thrown simultaneously. Find the number of elements in the event of getting :

(*i*) sum 4 (*ii*) sum 7

(*iii*) sum 11 (*iv*) sum not greater than 5.

6. A coin is tossed. Find the events A′, B′, A ∪ B, A ∩ B, where :
A = event of getting no head and B = event of getting one head.

7. A die is thrown. If :
A = event of getting a prime number and B = event of getting number greater than 3, find the events A^c, B^c, A ∪ B and A ∩ B.

8. What is the probability of getting an even number in the throw of an unbiased die ?

9. A and B are mutually exclusive events of an experiment. If P('not A') = 0.65, P(A ∪ B) = 0.65, and P(B) = *p*, find the value of *p*.

10. A box contains 4 red balls, 4 green balls and 7 white balls. What is the probability that a ball drawn is either red or white ?

11. (*i*) If A and B are two events defined on a sample space such that

$$P(A \cup B) = \frac{5}{6},\ P(A \cap B) = \frac{1}{3},$$

$$P(B^c) = \frac{1}{3}, \text{ find } P(A).$$

(*ii*) If A and B are two events such that $P(A) = \frac{1}{4}$, $P(B) = \frac{1}{2}$ and $P(A \text{ and } B) = \frac{1}{8}$, find (*a*) P(A or B), (*b*) P(not A and not B).

12. For any two events A and B, prove that $P(A \cap B) \le P(A) \le P(A \cup B) \le P(A) + P(B)$.

13. For two events A and B, let P(A) = 0.4 and P(B) = *p* and P(A ∪ B) = 0.6

(*i*) Find *p* so that A and B are mutually exclusive.

(*ii*) Find *p* so that A and B are independent.

14. (*i*) A coin is tossed twice and all possible outcomes are assumed to be equally likely. E is the event : "at most one head has occurred" and F is the event : "at most one tail has occurred". Show that the events E and F are not independent.

(*ii*) A coin is tossed twice and all possible outcomes are assumed to be equally likely. A is the event : both head and tail have occurred and B is the event : "at least one tail has occurred". Show that A and B are not independent.

15. (*i*) An unbiased die is thrown twice. Find the probability of getting a 4, 5, or 6 in the first throw and a 1, 2, 3 or 4 in the second throw.

(*ii*) A die is thrown twice. Find the probability of getting an odd number in the first throw and a multiple of 3 in the second throw.

NOTES

16. One ball is drawn from a bag containing 3 red and 2 black balls. Its colour is noted and then it is put back in the bag. A second draw is made and the same procedure is repeated. Find the probability of drawing : (*i*) two red balls (*ii*) one red and one black ball, and (*iii*) two black balls.

17. From a well shuffled pack of 52 cards, 3 cards are drawn one-by-one with replacement. Find the probability distribution of number of queens.

18. Find the mean and the variance of the following probability distributions :

(*i*)

x	2	3	4	5
$P(x)$	0	0.4	0.1	0.5

(*ii*)

x	1	2	3	4	5	6
$P(x)$	0.1	0.1	0.2	0.3	0.1	0.2

19. Two urns contain 5 black, 4 white balls and 4 black, 5 white balls. One ball is drawn from each urn. Find the mean and variance of the probability distribution of the random variable "no. of white balls drawn".

20. Two-dimensional random variable (X, Y) have the joint density

$$f(x, y) = \begin{cases} 8\,xy, & 0 < x < y < 1 \\ 0, & \text{otherwise} \end{cases}$$

(*i*) Find $P\left(X < \frac{1}{2} \cap Y < \frac{1}{4}\right)$

(*ii*) Find the marginal and conditional distributions.

(*iii*) Are X and Y independent ? Give reasons for your answer.

21. The two random variables X and Y have the joint probability density function :

$$f(x, y) = \frac{1}{2x^2 y}, \text{ for } 1 \le x < \infty \text{ and } \frac{1}{x} < y < x.$$

Derive the marginal distributions of X and Y. Further obtain the conditional distribution of Y for X = x and also that of X given Y = y.

22. A random variable X has F(x) as its distribution function [$f(x)$ is the density function]. Find the distribution function and the density function of the random variable :

(*i*) $Y = a + bX$, a and b are real numbers,
(*ii*) $Y = X^{-1}$, $[P(X = 0) = 0]$,
(*iii*) Y = tan X, and (*iv*) Y = cos X.

23. A, B, C are three mutually exclusive and exhaustive events associated with a random experiment. Find P(A), given that $P(B) = \frac{3}{2}\,P(A)$ and $P(C) = \frac{1}{2}\,P(B)$.

24. (*i*) Two dice are tossed together. Find the probability of getting a doublet or a total of 6.
(*ii*) A pair of dice is rolled. Find the probability of getting a doublet or sum of numbers to be at least 10.

25. A coin is tossed three times and all possible outcomes are assumed to be equally likely. E is the event : "both heads and tails have occurred", and F is the event : "at most one tail has occurred". Show that the events E and F are independent.

26. (*i*) 'A' speaks truth in 65% cases and 'B' in 80% cases. In what percentage of cases are they likely to contradict each other in stating the same fact ?
(*ii*) A speaks truth in 75% cases and B in 80% cases. In what percentage of cases are they likely to contradict each other in stating the same fact.

NOTES

27. Let X be a continuous random variable with *p.d.f.* $f(x)$. Let $Y = X^2$. Show that the random variable Y has *p.d.f.* given by :

$$g(y) = \begin{cases} \dfrac{1}{2\sqrt{y}}\{f(\sqrt{y}) + f(-\sqrt{y})\}, & y > 0 \\ 0, & y \le 0 \end{cases}$$

28. Find the distribution and density functions for

(*i*) $Y = aX + b$, $a \neq 0$, b real

(*ii*) $Y = e^x$, assuming that $F(x)$ and $f(x)$, the distribution and the density functions respectively of X are known.

FURTHER READINGS

1. Introduction to Modern Probability Theory: B.R. Bhat : Wiley Eastern.
2. Introduction to probability and Mathematical Statics: V.K. Rohatagi: Wiley Eastern.
3. Discrete Distributions : N.L. Johnson and S. Kotz, John Wiley and Sons.
4. Continuous Univerate Distributions–1: N.L. Johnson and S.Kotz.
5. Continuous Univerate Distributions–2 : N.L. Johnson and S.Kotz, John Wiley and Sons.
6. Introduction to Probability Theory with Applications: W. Feller, Vol-1 : Wiley Astern.

CHAPTER

2

EXPECTATION

OBJECTIVES

After going through this chapter, you should be able to:

- find the expected value of a random variable.
- know the properties of expectation.
- know how to use Chebychev's Inequality.
- know moment generating function.
- know uniqueness theorem of moment generating function.
- know about characteristic function, cumulants and Inversion theorem.

STRUCTURE

2.1 BASIC NOTIONS OF MATHEMATICAL EXPECTATION

If $f(x)$ is the probability mass function of a discrete random variable X, then

$E(X) = \sum_x xf(x)$, is known as expected value of X, provided the series is absolutely convergent *i.e.*, provided $\sum_x |x\,f(x)| = \sum_x |x|\,f(x) < \infty$

If $f(x)$ is the probability density function of a continuous random variable, then

$E(X) = \int_{-\infty}^{\infty} x\, f(x)\, dx$, provided the integral is absolutely convergent *i.e.*, provided

$$\int_{-\infty}^{\infty} |x\, f(x)|\, dx = \int_{-\infty}^{\infty} |x|\, f(x)\, dx < \infty$$

Remark. E(X) exists if E |X| exists

2.2 EXPECTED VALUE OF FUNCTION OF A RANDOM VARIABLE

Let $f(x)$ is a probability density function (or probability mass function) of a random variable X, and if $g(X)$ is a random variable such that $E(g(X))$ exists. Then,

$$E(g(X)) = \int_{-\infty}^{\infty} g(x)\, f(x)\, dx, \text{ (for continuous random variable)}$$

$$= \sum_{x} g(x)\, f(x), \text{ (for discrete random variable)}$$

Particular Cases

I. If we take $g(X) = X^r$, r is a positive integer,

then $E(X^r) = \int_{-\infty}^{\infty} x^r\, f(x)\, dx = \mu'_r$, (about origin)

i.e., $\mu'_1 = E(X)$, $\mu'_2 = E(X^2)$

$\therefore$ Mean $= \bar{x} = \mu'_1 = E(X)$

Variance $\mu_2 = \mu'_2 - \mu'^2_1 = E(X^2) - (E(X))^2$

II. If we take $g(X) = (X - E(X))^r = (X - \bar{x})^r$,

then, $E[X - E(X)]^r = \int_{-\infty}^{\infty} (x - \bar{x})^r\, f(x)\, dx = \mu_r$

where μ_r = rth moment about mean

$\therefore$ $\mu_2 = E(X - E(X))^2 = \int_{-\infty}^{\infty} (x - \bar{x})^2\, f(x)\, dx$

III. If we take $g(X) = c$, then

$$E(g(X)) = E(c) = \int_{-\infty}^{\infty} c f(x)\, dx = C \int_{-\infty}^{\infty} f(x)\, dx = c \qquad \Big|\ \text{Since } \int_{-\infty}^{\infty} f(x)\, dx = 1$$

Remark. The corresponding formulae for the discrete random variable X can be obtained on replacing integration by summation over the given range of the variable X in above particular cases.

2.3 PROPERTIES OF EXPECTATION

Addition Theorem of Expectation

Property I. *If X and Y are two random variables, then E (X + Y) = E (X) + E(Y), provided E(X) and E(Y) exist.*

NOTES

Proof. Let X and Y are continuous random variable and $f_{XY}(x, y)$ is the joint probability density function along with marginal probability density function $f_X(x)$ and $f_Y(y)$ respectively, then

$$E(X + Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x + y)\, f_{XY}(x, y)\, dx\, dy$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x\, f_{XY}(x, y)\, dx\, dy + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y f_{XY}(x, y)\, dx\, dy$$

$$= \int_{-\infty}^{\infty} x \left[\int_{-\infty}^{\infty} f_{XY}(x, y)\, dy\right] dx + \int_{-\infty}^{\infty} y \left[\int_{-\infty}^{\infty} f_{XY}(x, y)\, dx\right] dy$$

$$= \int_{-\infty}^{\infty} x\, f_X(x)\, dx + \int_{-\infty}^{\infty} y\, f_Y(y)\, dy$$

$$= E(X) + E(Y)$$

Generalisation. If $X_1, X_2, ..., X_n$ are n random variables, where $E(X_i)$ exist for all i, then

$$E(X_1 + X_2 + ... + X_n) = E(X_1) + E(X_2) + ... + E(X_n)$$

or $$E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} E(X_i), \text{ provided all the expectation exist}$$

Remark. The proof of property I for discrete random variable is similar to above if we replace the integration sign by summation (Σ).

Property II. Multiplication Theorem of Expectation.

If X and Y are independent random variables, then

$$E(XY) = E(X) \,.\, E(Y)$$

Proof. Let $f_{XY}(x, y)$ is the joint probability density function of the random variables and $f_X(x)$ and $f_Y(y)$ are the marginal probabilities density function respectively, then

$$E(XY) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\, f_{XY}(x, y)\, dx\, dy$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\, f_X(x)\, f_Y(y)\, dx\, dy \qquad \left|\; \begin{array}{l} f_{XY}(x, y) = f_X(x)\, f_Y(y) \text{ as} \\ X, Y \text{ are independent} \end{array}\right.$$

$$= \int_{-\infty}^{\infty} x f_X(x)\, dx \int_{-\infty}^{\infty} y f_Y(y)\, dy$$

$$= E(X)\, E(Y)$$

Generalisation. If $X_1, X_2, ... X_n$ are n independent random variables, then $E(X_1X_2 ...... X_n) = E(X_1)\, E(X_2) ...... E(X_n)$

or $$E\left(\prod_{i=1}^{n} X_i\right) = \prod_{i=1}^{n} E(X_i), \text{ provided all } X_j\text{'s independent and } E(X_i) \text{ exist for all } i.$$

Property III. *If X is a random variable and c is any constant, then*

(i) $E(cg(X)) = c\, E(g(X))$

(ii) $E(c + g(X)) = c + E(g(X))$

where g(X) is a function of a random variable X and all expectations exist.

Proof. (*i*) Let $f(x)$ is the probability density function of the random variable X, then

$$E(cg(X)) = \int_{-\infty}^{\infty} cg(x)\, f(x)\, dx = c\int_{-\infty}^{\infty} g(x)\, f(x)\, dx = cE(g(X))$$

(*ii*)
$$E(c + g(X)) = \int_{-\infty}^{\infty} (c + g(x))\, f(x)\, dx$$

$$= \int_{-\infty}^{\infty} c\, f(x)\, dx + \int_{-\infty}^{\infty} g(x)\, f(x)\, dx$$

$$= c\int_{-\infty}^{\infty} f(x)\, dx + \int_{-\infty}^{\infty} g(x)\, f(x)\, dx = c + E(g(X)).$$

Particular Cases. If we take $g(X) = X$, then

(*i*) $E(aX) = aE(X)$

(*ii*) $E(a + X) = a + E(X)$

If we take $g(X) = 1$, then $E(a) = a$

Property IV. *If X is a random variable and a and b are constants, then* $E(aX + b) = aE(X) + b$, *provided E(X) exists*

Proof. Let $f(x)$ is the probability density functions of the random variable X, then

$$E(ax + b) = \int_{-\infty}^{\infty} (ax + b)\, f(x)\, dx$$

$$= a\int_{-\infty}^{\infty} xf(x)\, dx + b\int_{-\infty}^{\infty} f(x)\, dx = aE(X) + b$$

Particular Case. If we take $a = 1$, $b = -\overline{X} = -E(X)$, then

$$E(X - \overline{X}) = E(X) - \overline{X} = E(X) - E(X) = 0$$

Remark. If we take $g(X) = aX + b$, then

$g(E(X)) = a\, E(X) + b$

Also $E(g(X)) = E(aX + b) = aE(X) + b$

It implies $E(g(X)) = g\,[E(X)]$,

provided $g(X)$ is a linear function. However, this result is not true if $g(X)$ is non-linear.

For example,

$$E(1/X) \neq 1/E(X)$$
$$E(X^{1/2}) \neq (E(X))^{1/2}$$
$$E(\log X) \neq \log E(X)$$
$$E(X^2) \neq (E(X))^2$$

Property V. *If* $X_1, X_2, \ldots\ldots X_n$ *are n random variables and* $a_1, a_2, \ldots\ldots a_n$ *are n constants, then,*

$$E(a_1X_1 + a_2X_2 + \ldots\ldots + a_nX_n) = a_1E(X_1) + a_2E(X_2) + \ldots\ldots + a_nE(X_n),$$

provided all $E(X_i)$ *exist*

NOTES

Proof. $E(a_1X_1 + a_2X_2 + ... + a_nX_n)$

$$= E(a_1X_1) + E(a_2X_2) + ... + E(a_nX_n) \quad | \text{Property I}$$

$$= a_1E(X_1) + a_2E(X_2) + ... + a_nE(X_n) \quad | \text{Property III}$$

Property VI. *If $X \geq 0$, then $E(X) \geq 0$*

Proof. Let $f(x)$ is the probability density function of the random variable, then

$$E(X) = \int_{-\infty}^{\infty} x\,f(x)\,dx = \int_{-\infty}^{0} 0\,dx + \int_{-\infty}^{\infty} x\,f(x)\,dx$$

$$= \int_{-\infty}^{\infty} x\,f(x)\,dx \geq 0 \quad \left|\begin{array}{l}\text{If } X \geq 0, \text{ then} \\ f(x) = 0 \text{ for all } X < 0\end{array}\right.$$

$\Rightarrow \quad E(X) \geq 0$

Property VII. *Let X and Y are two random variables such that $X \leq Y$, then $E(X) \leq E(Y)$, provided all expectation exist*

Proof. Given $X \leq Y \Rightarrow X - Y \leq 0 \Rightarrow Y - X \geq 0$

$\Rightarrow \quad E(Y - X) \geq 0$ | Property VI

$\Rightarrow \quad E(Y) - E(X) \geq 0$

$\Rightarrow \quad E(Y) \geq E(X)$

$\Rightarrow \quad E(X) \leq E(Y)$

Property VIII. $|E(X)| \leq E|X|$, provided E(X) exists

Proof. For each X, $X \leq |X|$

$\Rightarrow \quad E(X) \leq E|X| \quad ...(1) \quad$ | Property VII

Also $\quad -X \leq |X|$

$\Rightarrow \quad E(-X) \leq E|X|$

$\Rightarrow \quad -E(X) \leq E|X| \quad ...(2)$

From (1) and (2), $|E(X)| \leq E|X|$

Theorem I. *If X is a random variable, then $V(aX + b) = a^2V(X)$, where a and b are constants and V(X) is the variance of X*

Proof. Let $\quad Y = aX + b$. Then

$$E(Y) = E(aX + b) = aE(X) + b$$

$$Y - E(Y) = aX + b - (aE(X) + b) = a[X - E(X)]$$

Squaring and taking expectation of both sides,

$$E[Y - E(Y)]^2 = a^2[X - E(X)]^2$$

$\Rightarrow \quad V(Y) = a^2\,V(X) \quad$ or $\quad V(aX + b) = a^2\,V(X)$

Particular Cases. (*i*) If $b = 0$, then $\text{Var}\,(aX) = a^2\,V(X)$ *i.e.*, variance is not independent of change of scale.

(*ii*) If $a = 0$, then $V(b) = 0$ *i.e.*, variance of a constant is zero

(*iii*) If $0 = 1$, then $V(X + b) = V(X)$ *i.e.*, variance is independent of change of origin

Def. Covariance. Let X and Y are two random variables, then covariance between X and Y, denoted by Cov(X, Y), is defined by

$$\text{Cov}(X, Y) = E\{X - E(X)\}\{Y - E(Y)\}$$

$$= E[XY - XE(Y) - YE(X) + E(X)\,E(Y)]$$

$$= E(XY) - E(X)\,E(Y) - E(Y)\,E(X) + E(X)\,E(Y)$$

$$= E(XY) - E(X)\,E(Y) \quad ...(1)$$

If X and Y are independent, then

$E(XY) = E(X)\,E(Y)$ and hence from (1),

$$\text{Cov}(X, Y) = E(X)\,E(Y) - E(X)\,E(Y) = 0$$

NOTES

Remark 1. $\text{Cov}(aX, bY) = E[\{aX - E(aX)\}\{bY - E(bY)\}]$
$$= E[a(X - E(X))\, b(Y - E(Y))]$$
$$= abE[(X - E(X))(Y - E(Y))]$$
$$= ab\,\text{Cov}(X, Y)$$

2. $\text{Cov}(X + a, Y + b) = \text{Cov}(X, Y)$

3. $\text{Cov}(aX + b, cY + d) = ac\,\text{Cov}(X, Y)$

4. $\text{Cov}(X + Y, Z) = \text{Cov}(X, Z) + \text{Cov}(Y, Z)$

5. $\text{Cov}(aX + bY, cX + dY) = ac\,Var(X) + bd\,Var(Y) + (ad + bc)\,\text{Cov}(X, Y)$

Theorem II. *If $X_1, X_2, ... X_n$ are n random variables, then*

$$Var(a_1X_1 + a_2X_2 + ... + a_nX_n)$$
$$= a_1^2 Var(X_1) + a_2^2 Var(X_2) + ... + a_n^2 Var(X_n)$$
$$+ 2\sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j \,\text{Cov}(X_i, X_j),\ i < j$$

Proof. Take $U = a_1X_1 + a_2X_2 + ... + a_nX_n$

$$\Rightarrow \quad E(U) = E(a_1X_1 + a_2X_2 + ... + a_nX_n)$$
$$= a_1E(X_1 + a_2E(X_2)) + ... + a_nE(X_n)$$
$$U - E(U) = a_1(X_1 - E(X_1)) + a_2(X_2 - E(X_2)) + ... + a_n(X_n - E(X_n))$$

Squaring and taking expectation both sides,

$$E[U - E(U)]^2 = a_1^2E[X_1 - E(X_1)]^2 + a_2^2\,E(X_2 - E(X_2))^2$$
$$+ ...... + a_n^2\,E[X_n - E(X_n)]^2 + 2\sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j\,E[\{X_i - E(X_i)\}\{X_j - E(X_j)\}]$$

$$\Rightarrow \quad Var(U) = a_1^2\,Var(X_1) + a_2^2\,Var(X_2) + ... + a_n^2\,Var(X_n)$$
$$+ 2\sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j\,\text{Cov}(X_i, X_j),\ i < j \quad ...(1)$$

Particular Cases. I. If $a_1 = a_2 = ... a_n = 1$,

Then, $Var(X_1 + X_2 + ... + X_n)$

$$= Var(X_1) + Var(X_2) + ... + Var(X_n) + 2\sum_{i=1}^{n}\sum_{j=1}^{n} \text{Cov}(X_i, X_j),\ i < j \quad ...(2)$$

II. If $X_1, X_2, ... X_n$ are pairwise independent, then $\text{Cov}(X_i, X_j) = 0$ for all $i \neq j$, then from (1)

$$Var(a_1X_1 + a_2X_2 + ... + a_nX_n) = a_1^2\,Var(X_1) + a_2^2\,Var(X_2) + ... + a_n^2\,Var(X_n)$$

Also from (2)

$$Var(X_1 + X_2 + ... + X_n) = Var(X_1) + Var(X_2) + ... + Var(X_n)$$

III. If $a_1 = 1, a_2 = 1$ and $a_3 = a_4 = ... a_n = 0$,

then from (1)

$$Var(X + Y) = Var(X) + Var(Y) + 2\,\text{Cov}(X, Y)$$

NOTES

If $a_1 = 1, a_2 = -1$, and $a_3 = a_4 = \ldots\ldots a_n = 0$,
then from (1)

$$\text{Var}(X - Y) = \text{Var}(X) + \text{Var}(Y) - 2\,\text{Cov}(X, Y)$$

$$\therefore \quad \text{Var}(X \pm Y) = \text{Var}(X) + \text{Var}(Y) \pm 2\,\text{Cov}(X, Y)$$

If X and Y are independent, then Cov(X, Y) = 0 and hence

$$\text{Var}(X \pm Y) = \text{Var}(X) + \text{Var}(Y)$$

ILLUSTRATIVE EXAMPLES

Example 1. *Let X be a random variable with the following probability distribution*

x	– 3	6	9
p(x)	*1/6*	*1/2*	*1/3*

(i) Find E(X) and E(X²)

(ii) Find E(2X + 1)²

Sol.

$$E(X) = \sum x\,p(x) = (-3)\,.\frac{1}{6} + 6\,.\frac{1}{2} + 9\,.\frac{1}{3} = \frac{11}{2}$$

$$E(X^2) = \sum x^2\,p(x) = 9\,.\frac{1}{6} + 36\,.\frac{1}{2} + 81\,.\frac{1}{3} = \frac{93}{2}$$

$$E(2X + 1)^2 = E(4X^2 + 1 + 4X)$$

$$= 4E(X^2) + 1 + 4\,E(X)$$

$$= 4\,.\frac{93}{2} + 1 + 4\,.\frac{11}{2} = \mathbf{209}$$

Example 2. *A random variable has the following probability distribution.*
$P(X = 0) = P(X = 2) = p$ and
$P(X = 1) = 1 - 2p$. For what value of p, Var(X) is maximum. It is given that

$$0 \le p \le \frac{1}{2}$$

Sol. The random variable takes the values 0, 1, 2 with respective probabilities $p, 1 - 2p, p$

$$\therefore \quad E(X) = 0 \times P + 1 \times (1 - 2p) + 2 \times p = 1$$

$$E(X^2) = 0 \times p + 1 \times (1 - 2p) + 4 \times p = 1 + 2p$$

$$\therefore \quad \text{Var}(X) = E(X^2) - (E(X))^2 = 1 + 2p - 1 = 2p$$

For $0 \le p \le \frac{1}{2}$, Var(X) is maximum when $p = \frac{1}{2}$. Also,

$$\text{Max. Var}(X) = 2\,.\frac{1}{2} = 1$$

Example 3. *Let X be a random variable with mean μ and variance σ². Show that $E(X - b)^2$, as a function of b, is minimised when $b = \mu'$ or sum of squares of deviations is minimum when taken about mean.*

NOTES

Sol. $E(X-b)^2 = E[(X-\mu)+(\mu-b)]^2$

$= E[(X-\mu)^2 + (\mu-b)^2 + 2(X-\mu)(\mu-b)]$

$= E(X-\mu)^2 + E(\mu-b)^2 + 2(\mu-b)\,E(X-\mu)$

$= E[X-E(X)]^2 + (\mu-b)^2$ $\qquad | E(X-\mu)=0$

$= \text{Var}(X) + (\mu-b)^2 \ge \text{Var}(X)$

$\Rightarrow \quad E(X-b)^2 \ge \text{Var}(X)$ $\qquad |$ Since $(\mu-b)^2 \ge 0$

$\therefore$ $E(X-b)^2$ is minimum when $\mu = b$ and its minimum value is

$E(X-\mu)^2 = \text{Var}(X)$

Example 4. *If t is any positive real number, show that the function defined by $p(x) = e^{-t}(1-e^{-t})^{x-1}$ can represent a probability function of a random variable X which takes the values 1, 2, 3, ... Find E(X) and Var (X) of the distribution.*

Sol. For $t > 0$, $e^t > 1 \Rightarrow e^{-t} < 1 \Rightarrow 1-e^{-t} > 0$

Also $\quad e^{-t} = \dfrac{1}{e^t} > 0$ for all $t > 0$

$\therefore \quad p(x) = e^{-t}(1-e^{-t})^{x-1} \ge 0$ for all $t > 0$, $\quad x = 1, 2, 3, ...$

Also $\quad \displaystyle\sum_{x=1}^{\infty} p(x) = \sum_{x=1}^{\infty} e^{-t}(1-e^{-t})^{x-1}$

$$= e^{-t}\sum_{x=1}^{\infty}(1-e^{-t})^{x-1}$$

$$= e^{-t}\sum_{x=1}^{\infty} a^{x-1} \text{ where } a = 1-e^{-t}$$

$$= e^{-t}(1 + a + a^2 + ...) = \frac{e^{-t}}{1-a}$$

$$= \frac{e^{-t}}{1-1+e^{-t}} = \frac{e^{-t}}{e^{-t}} = 1$$

Hence $p(x) = e^{-t}(1-e^{-t})^{x-1}$, $t \ge 0$, $x = 1, 2, 3, ...$ represents the probability function of the random variable X.

Also $\quad \displaystyle E(X) = \sum_{x=1}^{\infty} x\,p(x) = \sum_{x=1}^{\infty} x\,e^{-t}(1-e^{-t})^{x-1}$

$$= e^{-t}\sum_{x=1}^{\infty} x\,a^{x-1} \text{ where } a = 1-e^{-t}$$

$$= e^{-t}(1 + 2a + 3a^2 + 4a^3 + ...)$$

$$= e^{-t}(1-a)^{-2}$$

$$= e^{-t}(1-1+e^{-t})^{-2} = e^{-t}e^{2t} = e^t$$

Also $\quad \displaystyle E(X^2) = \sum_{x=1}^{\infty} x^2 p(x) = \sum_{x=1}^{\infty} x^2 e^{-t}(1-e^{-t})^{x-1}$

$$= e^{-t}\sum_{x=1}^{\infty} x^2 a^{x-1} \text{ where } a = 1-e^{-t}$$

NOTES

$$= e^{-t}(1 + 4a + 9a^2 + 16a^3 + \ldots\ldots \infty)$$
$$= e^{-t}(1 + a)(1 - a)^{-3}$$
$$= e^{-t}(1 + 1 - e^{-t})(e^{-t})^{-3}$$
$$= e^{-t}(2 - e^{-t})e^{3t} = e^{2t}(2 - e^{-t})$$

$$\therefore \quad \text{Var}(X) = E(X^2) - (E(X))^2 = e^{2t}(2 - e^{-t}) - e^{2t}$$
$$= e^{2t}(2 - e^{-t} - 1) = e^{2t}(1 + e^{-t})$$

2.4 CONDITIONAL EXPECTATION

The conditional expectation of a discrete random variable X given $Y = y_j$ is defined by

$$E(X \mid Y = y_j) = \sum_{i=1}^{\infty} x\, P(X = x_i \mid Y = y_j)$$

The conditional expectation of a continuous random variable X given $Y = y$ is defined as

$$E(X \mid Y = y) = \frac{\int_{-\infty}^{\infty} x\, f(x, y)\, dx}{f_Y(y)}$$

Similarly $E(Y \mid X = x) = \dfrac{\int_{-\infty}^{\infty} y\, f(x, y)\, dy}{f_X(x)}$

Theorem III. Chebychev's Inequality

If X is a random variable with mean μ and variance σ^2, then for a +ve number k,

$$P\{(X - \mu) \geq k\sigma\} \leq \frac{1}{k^2} \text{ or}$$

$$P\{(X - \mu) < k\sigma)\} \geq 1 - \frac{1}{k^2}$$

Proof. Case I. If X is a continuous random variable, then

$$\sigma^2 = \sigma_X{}^2 = E\{X - E(X)\}^2 = E(X - \mu)^2$$

$$= \int_{-\infty}^{\infty} (x - \mu)^2 f(x)\, dx, \text{ where } f(x) \text{ is } p.d.f. \text{ of X.}$$

$$= \int_{-\infty}^{\mu - k\sigma} (x - \mu)^2 f(x)\, dx + \int_{\mu - k\sigma}^{\mu + k\sigma} (x - \mu)^2 f(x)\, dx + \int_{\mu + k\sigma}^{\infty} (x - \mu)^2 f(x)\, dx$$

$$\geq \int_{-\infty}^{\mu - k\sigma} (x - \mu)^2 f(x) + dx + \int_{\mu + k\sigma}^{\infty} (x - \mu)^2 f(x)\, dx \qquad \ldots(1)$$

We know that :

$$x \leq \mu - k\sigma \quad \text{and} \quad x \geq \mu + k\sigma \iff |x - \mu| \geq k\sigma \qquad \ldots(2)$$

Substituting in (1), we get

$$\therefore \quad \sigma^2 \geq k^2\sigma^2 \left[\int_{-\infty}^{\mu - k\sigma} f(x)\, dx + \int_{\mu + k\sigma}^{\infty} f(x)\, dx\right] \qquad | \text{Using (2)}$$

$$= k^2\sigma^2 [P(X \leq \mu - k\sigma) + P(X \geq \mu + k\sigma)]$$

$$= k^2\sigma^2 . P(|X - \mu| \geq k\sigma) \qquad | \text{Using (2)}$$

$$\Rightarrow \quad P(|X-\mu| \geq k\sigma) \leq 1/k^2, \qquad \ldots(3)$$

Also since

$$P\{|X-\mu| \geq k\sigma\} + P\{|X-\mu| < k\sigma\} = 1, \text{ we get}$$

$$P\{|X-\mu| < k\sigma\} = 1 - P\{|X-\mu| \geq k\sigma\} \geq 1 - \{1/k^2\} \qquad |\text{Using (3)}$$

Hence the first case.

Case II. When X is a discrete random variable, then the proof follows directly as in case I on replacing integration by summation.

Particular Cases. Take $k\sigma = C$, $C > 0$, then from (1) and (2),

$$P\{(X-\mu) \geq C\} \leq \frac{\sigma^2}{C^2} \quad \text{and} \quad P\{|X-\mu| < C\} \geq 1 - \frac{\sigma^2}{C^2}$$

$$\text{or} \quad P\{(X - E(X) \geq C\} \leq \frac{\text{Var}(X)}{C^2} \text{ and } P\{|X - E(x) < C\} \geq 1 - \frac{\text{Var}(X)}{C^2}$$

Theorem IV. Markov's Inequality

Let $g(X)$ be a non-negative function of a random variable X, then for every $k > 0$, we have

$$P\{|X| \geq k\} \leq \frac{E(X)}{k}$$

Proof. Let S denotes the set of all X and $g(X) \geq k$. *i.e.*,

$$S = \{x : g(x) \geq k\}, \text{ then}$$

$$\int_S dF(x) = P(X \in S) = P\{g(X) \geq k\},$$

where $F(x)$ is the distribution function of X

$$\therefore \quad E(g(X)) = \int_{-\infty}^{\infty} g(x)\, dF(x) \geq \int_S g(x)\, dF(x)$$

$$\geq kP\{g(x) \geq k\} \qquad |g(x) \geq k \text{ on S}$$

$$\Rightarrow \quad P\{g(x) \geq k\} \leq \frac{E(g(X))}{k}$$

Take $g(X) = |X|$, then

$$P[|X| \geq k] \leq \frac{E(X)}{k}$$

2.5 MOMENT GENERATING FUNCTION

The moment generating function (*m.g.f*) of a random variable X (about origin) with probability distribution function $f(x)$, is defined by

$$M_X(t) = E(e^{tX}) = \begin{cases} \int e^{tx} f(x)\, dx, & \text{for continuous random variable} \\ \sum_x e^{tx} f(x), & \text{for discrete random variable} \end{cases}$$

Theorem I. *Show that* $M_X(t) = \sum_{r=0}^{\infty} \frac{t^r}{r!} \mu_r'$

NOTES

where $$\mu'_r = E(X^r) = \begin{cases} \int x^r f(x)\, dx \\ \sum_x x^r f(x) \end{cases}$$

Proof. $$M_X(t) = E(e^{tx}) = E\left(1 + tX + \frac{(tX)^2}{2!} + \frac{(tX)^3}{3!} + \ldots\ldots\right)$$

$$= 1 + t\,E(X) + \frac{t^2}{2!}E(X^2) + \frac{t^3}{3!}E(X^3) + \ldots\ldots$$

$$= 1 + t\mu'_1 + \frac{t^2}{2!}\mu'_2 + \ldots\ldots + \frac{t^r}{r!}\mu'_r + \ldots\ldots$$

$$= \sum_{r=0}^{\infty} \frac{t^r}{r!}\mu'_r$$

Thus, coefficient of $\frac{t^r}{r!}$ in $M_X(t)$ gives moments, hence $M_X(t)$ is know as moment generating function.

The moment generating function of the random variable about the point $X = a$ is defined as

$$M_X(t) = (\text{about } X = a) = E\,[e^{t(X-a)}]$$

$$= E\left[1 + t(X-a) + \frac{t^2}{2!}(X-a)^2 + \ldots\ldots + \frac{t^r}{r!}(X-a)^r + \ldots..\right]$$

$$= 1 + t\,\mu_1 + \frac{t^2}{2!}\mu_2 + \ldots\ldots + \frac{t^2}{r!}\mu_r + \ldots\ldots$$

where $\mu_r = E[(X-a)^r]$, rth moment about the point $X = a$.

Theorem II. *$M_{cX}(t) = M_X(ct)$, c is any constant*

Proof. $M_{cX}(t) = E(e^{tcX}) = \text{L.H.S.}$

Also $M_X(t) = E(e^{ctX}) = \text{R.H.S.}$

Hence the theorem

Theorem III. *The moment generating function of the sum of a number of independent random variables is equal to the product of their respective moment generating functions. i.e., if $X_1, X_2, \ldots\ldots X_n$ are independent random variables, then*

$$M_{X_1 + X_2 + \ldots\ldots + X_n}(t) = M_{X_1}(t)\, M_{X_2}(t) \ldots\ldots M_{X_n}(t)$$

Proof. By definition.

$$M_{X_1 + X_2 + \ldots\ldots + X_n}(t) = E\left[e^{t(X_1 + X_2 + \ldots\ldots + X_n)}\right]$$

$$= E\left(e^{tX_1} \cdot e^{tX_2} \ldots e^{tX_n}\right)$$

$$= E(e^{tX_1}) \cdot E(e^{tX_2}) \ldots\ldots E(e^{tX_n}) \quad | X_1, X_2, \ldots\ldots X_n \text{ are independent}$$

$$= M_{X_1}(t), M_{X_2}(t), \ldots\ldots M_{X_n}(t).$$

Theorem IV. *Effect of change of origin and scale on moment generating function.*

Proof. Let X be a random variable. Transform X to the new variable U by changing both the origin and scale in X by defining

$$U = \frac{X-a}{h}, \text{ where } a, h \text{ are constants.}$$

Then $$M_U(t) = E(e^{tU}) = E\left(e^{t\frac{X-a}{h}}\right)$$

$$= E\left(e^{\frac{tX}{h}} . e^{-\frac{ta}{h}}\right) = e^{-\frac{ta}{h}} E\left(e^{\frac{tX}{h}}\right)$$

$$= e^{-\frac{ta}{h}} M_X\left(\frac{t}{h}\right), \text{ where}$$

$M_X(t)$ is the moment generating function of X about the origin

Particular Cases : Take $a = E(X) = \mu$, $h = Var(X) = \sigma$,

then $$U = \frac{X - E(X)}{Var(X)} = \frac{X - \mu}{\sigma} = Z, \text{ say};$$

Then $$M_Z(t) = e^{-\frac{\mu t}{\sigma}} M_X\left(\frac{t}{\sigma}\right)$$

Theorem. Uniqueness Theorem of Moment Generating Function

The moment generating function of a distribution, if it exists, uniquely determines the distribution. This means, corresponding to a given probability distribution, there is only one moment generating function, (provided it exists) and corresponding to a given moment generating function, there is only one probability distribution. Hence

$$M_X(t) = M_Y(t) \Rightarrow \text{ X and Y are identically distributed.}$$

Cumulants

If $k(t) = \log M_X(t)$, then $k(t)$ is known as cumulant generating function, provided the right hand side can be expanded as a convergent series in powers of t.

Thus, $$k_X(t) = k_1 t + k_2 \frac{t^2}{2!} + \ldots\ldots + k_r \frac{t^r}{r!} + \ldots\ldots = \log M_X(t)$$

$$= \log\left(1 + \mu_1' t + \mu_2' . \frac{t^2}{2!} + \mu_3' . \frac{t^3}{3!} + \ldots\ldots + \mu_r' \frac{t^r}{r!} + \ldots\ldots\right)$$

where k_r = coefficient of $\frac{t^r}{r!}$ in $k_X(t)$, is called the rth cumulant. Hence

$$k_1 t + k_2 \frac{t^2}{2!} + \ldots\ldots + k_r \frac{t^r}{r!} + \ldots\ldots$$

$$= \left(\mu_1' t + \mu_2' . \frac{t^2}{2!} + \mu_3' . \frac{t^3}{3!} + \ldots\ldots\right) - \frac{1}{2}\left(\mu_1' t + \mu_2' . \frac{t^2}{2!} + \mu_3' \frac{t^3}{3!} + \ldots\ldots\right)^2$$

$$+ \frac{1}{3}\left(\mu_1' t + \mu_2' . \frac{t^2}{2!} + \mu_3' . \frac{t^3}{3!} + \ldots\ldots\right)^3 - \frac{1}{4}\left(\mu_1' t + \mu_2' \frac{t^2}{2!} + \mu_3' \frac{t^3}{3!} + \ldots\ldots\right)^4$$

$$\log(1 + x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \ldots\ldots$$

Equating the coefficients of like powers of t, we get

$$k_1 = \mu_1' = \text{mean}$$

$$\frac{k_2}{2!} = \frac{\mu_2'}{2!} - \frac{\mu_1'^2}{2!} \Rightarrow k_2 = \mu'_2 - \mu'_1{}^2 = \mu_2$$

NOTES

$$\frac{k_3}{3!} = \frac{\mu_3'}{3!} - \frac{1}{2} \cdot \frac{2\mu_1'\mu_2'}{2!} + \frac{\mu_1'^3}{3}$$

$\Rightarrow$ $$k_3 = \mu'_3 - 3\mu'_1\mu'_2 + 2\mu_1'^3 = \mu_3$$

Similarly, $\mu_4 = k_4 + 3k_2^{\,2}$

2.6 CHARACTERISTIC FUNCTION

The characteristic function of a random variable, X, denoted by $\phi_X(t)$, is defined by

$$\phi_X(t) = E(e^{itX}) = \begin{cases} \int e^{itx} f(x)\, dx, & \text{for continuous } r.v.X \\ \sum_x e^{itx} f(x), & \text{for discrete } r.v.X \end{cases}$$

If $F_X(x)$ is the distribution function of a continuous random variable X, then

$$\phi_X(t) = \int_{-\infty}^{\infty} e^{itx}\, dF(x).$$

Properties of Characteristic Function

Statement. *For all real t,*

(i) $\phi(0) = 1$ *(ii)* $|\phi(t)| \le \phi(0)$

Proof. $$\phi(t) = \int_{-\infty}^{\infty} e^{itx} f(x)\, dx$$

$\therefore$ $$\phi(0) = \int_{-\infty}^{\infty} f(x)\, dx = 1$$

Also $$|\phi(t)| = \left| \int_{-\infty}^{\infty} e^{itx} f(x)\, dx \right| \le \int_{-\infty}^{\infty} \left| e^{itx} \right| f(x)\, dx = \int_{-\infty}^{\infty} f(x)\, dx = 1$$

$$\left| e^{itx} \right| = |\cos tx + i \sin tx| = 1$$

$\Rightarrow$ $$|\phi(t)| \le 1 = \phi_0(t)$$

Since $|\phi(t)| \le 1$, characteristic function always exists.

2.7 KOLMOGOROV INEQUALITY

If $g(x)$ is continuous and convex function on an interval I and X is random variable with values in with probability 1, then

$$E(g(x)) \ge g(E(x)), \text{ provided } E(x) \text{ exists.}$$

2.8 HOLDER'S INEQUALITY

Statement. *If X and Y are two discrete random variables, then*

$$\left(\sum_{j=1}^{n} x_i y_i \right) \le \left(\sum_{j=1}^{n} x_i^{\,2} \right) \left(\sum_{j=1}^{n} y_i^{\,2} \right)$$

NOTES

Proof. Let G(t) be a real valued function defined for t, such that

$$G(t) = E(X + ty)^2 \quad ...(1)$$

Since $(x + ty)^2 \geq 0 \Rightarrow E(x + ty)^2 \geq 0 \Rightarrow G(t) \geq 0 \ \forall\ t.$

Now $G(t) = E(x^2 + t^2y^2 + 2txy) \geq 0$

$$= E(x)^2 + t^2\, E(y^2) + 2t\, E(xy)$$

$$= E(y^2).t^2 + 2t\, E(xy) + E(x^2),$$

which is quadratic int, say,

$$G(t) = At^2 + Bt + C, \text{ where}$$

$$A = E(y^2), \quad B = 2E(xy), \quad C = E(x^2)$$

Since $G(t) \geq 0$, It means that graph of G(t) will either touch t-axis at one point or does not touch t-axis.

$\therefore$ We should have $D \leq 0$, where D is the discrimant of G(t)

| Since G(t) has two distinct roots if $D > 0$

$\Rightarrow \quad B^2 - 4AC \leq 0$

$\Rightarrow \quad (2E\,(xy))^2 - 4.\, E(y^2).E(x^2) \leq 0$

$\Rightarrow \quad [E(xy)]^2 \leq E(x^2)\, E(y^2) \quad ...(2)$

Take $E(x^2) = \dfrac{1}{n}\sum\limits_{\Gamma=1}^{n} x_i^2,$

$$E(y^2) = \frac{1}{n}\sum_{\Gamma=1}^{n} y_i^2$$

$$E(xy) = \frac{1}{n}\sum_{\Gamma=1}^{n} x_i y_i$$

$\therefore$ (2) gives

$$\left(\frac{1}{n}\sum_{\Gamma=1}^{n} x_i y_i\right)^2 \leq \left(\frac{1}{n}\sum_{\Gamma=1}^{n} x_i{}^2\right)\left(\frac{1}{n}\sum_{\Gamma=1}^{n} y_i{}^2\right)$$

which proves the result.

2.9 MINKOWSKI'S INEQUALITY

If x and y are two random variables (discrete or continuous). Then

$$E(x + y) = E(x) + E(y)$$

If Case I when x and y are continuous random variable. Let $f_x(x)$, $f_y(y)$ denotes the probability density function of the random variables x and y and let $f_{xy}(x, y)$ be the joint p.d.f. of these random variables, then,

$$E(x) = \int_{-\infty}^{\infty} x f_x(x)\, dx,\ E(y) = \int_{-\infty}^{\infty} y f_y(y)\, dy$$

Also

$$E(x + y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x + y)\, f_{xy}(x, y)\, dx dy$$

NOTES

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x\, f_{xy}(x, y)dxdy + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y\, f_{xy}(x, y)dxdy$$

$$= \int_{-\infty}^{\infty} x\left[\int_{-\infty}^{\infty} f_{xy}(x, y)dy\right]dx + \int_{-\infty}^{\infty} y\left[\int_{-\infty}^{\infty} f_{xy}(x, y)dx\right]dy$$

$$= \int_{-\infty}^{\infty} x\, f_x(x)\,dx + \int_{-\infty}^{\infty} y\, f_y(y)\,dy$$

$$= E(x) + E(y)$$

Theorem. State and Prove Inversion Theorem.

Let $F(x)$ and $\phi(t)$ be the distribution function and the characteristic function respectively, of random variable X, then for $x \in (a - h, a + h)$,

$$F(a + h) - F(a - h) = \underset{T\to\infty}{\text{Lt}}\ \frac{1}{\pi}\int_{-T}^{T} \frac{\sin hT}{t}\, e^{-ita}\, \phi(t)\, dt$$

$$\text{Define } J = \frac{1}{\pi}\int_{-T}^{T} \frac{\sin ht}{t}\, e^{-ita}\, \phi(t)\, dt \qquad \ldots(1)$$

$$= \frac{1}{\pi}\int_{-T}^{T}\left\{\frac{\sin ht}{t}\, e^{ita}\int_{-\infty}^{\infty} e^{itx} f(x)\, dx\right\} dt$$

where $\qquad dF(x) = f(x)\, dx$, $f(x)$ being the probability density function of X.

$$\therefore \qquad J = \frac{1}{\pi}\int_{-T}^{T}\left\{\int_{-\infty}^{\infty}\frac{\sin ht}{t}\, e^{it(x-a)} f(x)\, dx\right\} dt \qquad \ldots(2)$$

Since the limits of integral over t are finite and the integral in the R.H.S. is absolutely convergent, *i.e.*,

$$\int_{-\infty}^{\infty}\left|\frac{\sin ht}{t}\,.\,e^{it(x-a)}\, f(x)\right| dx \le \int_{-\infty}^{\infty}\left|\frac{\sin ht}{t}\right| f(x)\, dx \qquad (\because \quad |\, e^{itx}\,| = 1)$$

$$= \int_{-\infty}^{\infty}\left|\frac{\sin ht}{ht}\,.\, h\right| f(x)\, dx \le |h| \int_{-\infty}^{\infty} f(x)\, dx = |\, h\,|,$$

Changing the order of integration in (2)

$$\therefore \qquad J = \frac{1}{\pi}\int_{-\infty}^{\infty}\left\{\int_{-T}^{T}\frac{\sin ht}{t}\,.\,e^{it(x-a)}\, f(x)\, dt\right\} dx \qquad |\text{ Fubini's theorem}$$

$$= \frac{1}{\pi}\left[\int_{-\infty}^{\infty}\left\{\int_{-T}^{T}\frac{\sin ht}{t}\,.\cos t(x-a)\, dt\right\} f(x)\, dx\right]$$

$$+ \frac{i}{\pi}\left[\int_{-\infty}^{\infty}\left\{\int_{-T}^{T}\frac{\sin ht}{t}\,.\sin t(x-a)\, dt\right\} f(x)\, dx\right]$$

Since $\dfrac{\sin ht}{t}\cos t\,(x - a)$ is an even function of t and $\dfrac{\sin ht}{t}\,.\sin t\,(x - a)$ is an odd function of t, the second integral vanishes and we get

$$J = \frac{2}{\pi}\int_{-\infty}^{\infty}\left\{\int_{0}^{T}\frac{\sin ht}{t}\,.\cos t\,(x-a)\, dt\right\} f(x)\, dx\Bigg]$$

$$= \int_{-\infty}^{\infty} g(x, T)\, f(x)\, dx, \qquad \ldots(3)$$

NOTES

where $$g(x, T) = \frac{2}{\pi}\int_0^T \frac{\sin ht}{t} . \cos t\,(x-a)\,dt$$

$$= \frac{1}{\pi}\int_0^T \frac{1}{t}\{2\cos t\,(x-a) . \sin ht\}\,dt$$

$$= \frac{1}{\pi}\int_0^T \frac{1}{t}\{\sin (x-a+h)\,t - \sin (x-a-h)\,t\}\,dt$$

$$= \frac{1}{2}\int_0^T \frac{2}{\pi}\,\frac{\sin (x-a+h)\,t}{t} . dt - \frac{1}{2}\int_0^T \frac{2}{\pi}\,\frac{\sin (x-a-h)\,t}{t} . dt$$

$$= \frac{1}{2} . S\,(x-a+h, T) - \frac{1}{2} S(x-a-h, T)$$

We know that if

$$S(h, t) = \frac{2}{\pi}\int_0^T \frac{\sin ht}{t}\,dt, \text{ where } h \text{ is real and } T > 0,$$

then

$$\underset{T\to\infty}{\text{Lt}}\ S(h, T) = \begin{cases} -1, & h < 0 \\ 0, & h = 0 \\ 1, & h > 0 \end{cases}$$

$$= \frac{1}{2} S\,\{x-(a-h), T\} - \frac{1}{2} S\,\{x-(a+h), T\},$$

$$= I_1 - I_2, \text{ (say)} \qquad \ldots(4)$$

Since $S(h, T)$ is bounded, the right hand side in (4) is bounded and consequently $g(x, T)$ is bounded, *i.e.*,

$$|g\,(x, T)\,| < |k|, \text{ say, } \forall\, T. \qquad \ldots(5)$$

From (3),

$$\lim_{T\to\infty} J = \lim_{T\to\infty}\int_{-\infty}^{\infty} g(x, T)\,f(x)\,dx$$

$$= \int_{-\infty}^{\infty} \lim_{T\to\infty} g(x, T)\,f(x)\,dx \qquad \ldots(6)$$

Now, consider the values of I_1 and I_2 in (4), for different ranges of x in the following table.

$x \to$	$x < a-h$	$x = a-h$	$a-h < x < a+h$	$x = a+h$	$x > a+h$	
I_1	– 1/2	0	1/2	1/2	1/2	as $T \to \infty$
I_2	– 1/2	– 1/2	– 1/2	0	1/2	as $T \to \infty$
$I_1 - I_2$	0	1/2	1	1/2	0	

NOTES

$$\therefore \qquad \lim_{T\to\infty} g(x, T) = \begin{cases} 0, & x < a-h \\ \dfrac{1}{2}, & x = a-h \\ 1, & a-h < x < a+h \\ \dfrac{1}{2}, & x = a+h \\ 0, & x > a+h \end{cases} \qquad \ldots(7)$$

Substituting these values in (6), we get

$$\lim_{T\to\infty} J = \int_{-\infty}^{a-h} \lim_{T\to\infty} g(x, T)\, f(x)\, dx + \int_{a-h}^{a+h} \lim_{T\to\infty} g(x, T)\, f(x)\, dx$$

$$+ \int_{a+h}^{\infty} \lim_{T\to\infty} g(x, T)\, f(x)\, dx$$

$$= 0 + \int_{a-h}^{a+h} f(x)\, dx + 0$$

$$= \int_{a-h}^{a+h} f(x)\, dx = P(a-h \le X \le a+h)$$

$$= P(X \le a+h) - P(X \le a-h)$$

$$= F(a+h) - F(a-h) \qquad \ldots(8)$$

Substituting from (8) in (1), we get

$$F(a+h) - F(a-h) = \lim_{T\to\infty} \frac{1}{\pi} \int_{-T}^{T} \frac{\sin ht}{t}\, e^{-ita}\, \phi(t)\, dt,$$

Hence the theorem.

SUMMARY

- If $f(x)$ is the probability mass function (or probability density function) of a random variable, then the expected value of x is given by

$$E(x) = \begin{cases} \sum_x x\, f(x), & \text{for discrete random variable} \\ \int x\, f(x)\, dx, & \text{for continuous random variable} \end{cases}$$

- If $g(x)$ is any linear function, then $E(g(x)) = g(E(x))$. But this result is not true if $g(x)$ is non-linear.
- If x and y are independent random variables, then cov $(x, y) = 0$. But the converse is not true.

GLOSSARY

- **m.g.f.** The moment generating function of a random variable x(about origin) with p.d.f. $f(x)$ is given by

$$M_x(t) = E(e^{tx}) = \begin{cases} \int e^{tx}\, f(x)\, dx, & \text{for continuous random variable} \\ \sum_x e^{tx}\, f(x), & \text{for discrete random variable} \end{cases}$$

- **Characteristic Function.** The characteristic function of a random variable x, is defined by

$$\phi_x(t) = E(e^{itx}) = \begin{cases} \int e^{itx} f(x)\, dx, & \text{for continuous random variable} \\ \sum_x e^{itx} f(x), & \text{for discrete random variable} \end{cases}$$

REVIEW QUESTIONS

1. Prove the following formulae
   (i) $k_1 = \mu_1^1$ mean (ii) $k_2 = \mu_2$
   (iii) $k_3 = \mu_3$ (iv) $k_4 = \mu_4 - 3k_2^2$
2. If $Z = \dfrac{X - \mu}{\sigma}$, then show that $M_2(t) = e^{-\frac{\mu t}{\sigma}} M_x\left(\dfrac{t}{\sigma}\right)$.
3. If X is a random variable with mean μ and variance σ^2, then
   $p\{(X - E(X)) \geq c\} \leq \dfrac{\text{Var (X)}}{c^2}$, where c is same positive constant.

FURTHER READINGS

1. Introduction to Modern Probability Theory: B.R. Bhat: Wiley Eastern.
2. Introduction to probability and Mathematical Statistics: V.K. Rohatgi: Wiley Eastern.
3. Discrete Distributions: N.L. Johnson and S.Kotz, John Wiley and Sons.
4. Continuous Univarate distributions—1: N.L. Johnson and S.Kotz.

UNIT–II

NOTES

CHAPTER

3

CONVERGENCE OF SEQUENCE OF RANDOM VARIABLES

OBJECTIVES

After going through this chapter, you should be able to:

- know about convergence of sequence of random variables in probability.
- know about WLLN (weak law of large numbers)
- know Markov's theorem for WLLN
- know the relation between central limit theorem and WLLN.

STRUCTURE

3.1 INTRODUCTION

A sequence of random variables $X_1, X_2, \ldots, X_n$ is said to converge in probability to a constant a, if for any $\varepsilon > 0$,

$$\underset{x \to \infty}{\text{Lt}} \; P(\,|\, X_n - a \,|\, < \varepsilon) = 1 \quad \text{or} \quad \underset{x \to \infty}{\text{Lt}} \; P(\,|\, X_n - a \,|\, \geq \varepsilon) = 0$$

and we write $X_n \xrightarrow{P} a$, as $n \to \infty$.

3.2 CONVERGENCE IN PROBABILITY

If there exists a random variable X such that $X_n - X \xrightarrow{P} a$, as $n \to \infty$, we say that the given sequence $<X_n>$ of random variables converges in probability to the random variables X.

3.3 TYPES OF CONVERGENCE AND INTERRELATIONS

Types of Convergence. There are two types of convergence, namely

(*i*) Convergence in probability

(*ii*) Ordinary convergence

Interrelation: The concept of convergence in probability is basically different from that of ordinary convergence of sequence of real numbers. But, the following results also hold for the convergence in probability.

If $X_n \xrightarrow{P} a, \dfrac{1}{n} \xrightarrow{P} b$, as $n \to \infty$, then

(*i*) $X_n \pm Y_n \xrightarrow{P} a \pm b$, as $n \to \infty$

(*ii*) $\dfrac{X_n}{Y_n} \xrightarrow{P} \dfrac{a}{b}, b \neq 0$, as $n \to \infty$.

Theorem I. Chebyshev's Theorem on Convergence in Probability.

Statement. *If $X_1, X_2, \ldots\ldots, X_n$ is a sequence of random variables and if mean μ_n and standard deviation σ_n of X_n exists for all n satisfying $\sigma_n \to 0$ as $n \to \infty$, then*

$$X_n - \mu_n \xrightarrow{P} 0, \text{ as } n \to \infty.$$

Proof. By Chebyshev's inequality, if X is a random variable with mean μ and variable σ^2,

Then for any positive number k,

$$P\{|\, X - \mu \,| \geq c\} \leq \frac{\text{Var}(X)}{c^2}$$

In our case, we have for $\varepsilon > 0$,

$$P\{|\, x_n - \mu_n \,| \geq \varepsilon\} \leq \frac{\sigma_n^{\,2}}{\varepsilon^2} \to 0, \text{ as } n \to \infty$$

$\therefore$ By definition of convergence in probability,

$$X_n - \mu_n \xrightarrow{P} 0, \text{ as } n \to \infty, \text{ provided } \sigma_n \to 0, n \to \infty$$

Hence the theorem.

NOTES

Theorem II. Weak law of large numbers (Chebyshev's Form)

Statement. *If $X_1, X_2, ..., X_n$ be a sequence of random variables and $\mu_1, \mu_2, ..., \mu_n$ be their respective means, satisfying* $\underset{n \to \infty}{Lt} \frac{B_n}{n^2} \to 0$ *where*

$$B_n = Var\,(X_1 + X_2 + ... + X_n) < \infty, \text{ then}$$

$$P\left\{\left|\frac{X_1 + X_2 + ... + X_n}{n} - \frac{\mu_1 + \mu_2 + ... + \mu_n}{n}\right| < \varepsilon\right\} \geq 1 - \eta$$

for all $n > n_0$ where ε and η are arbitrary small positive numbers.

Proof. By Chebyshev's inequality, if X is a random variable with mean μ and variance σ^2, then

$$P\{|X - E(X)| < c\} \geq 1 - \frac{Var(X)}{c^2}$$

In our case, replace X by $\frac{X_1 + X_2 + ... + X_n}{n}$,

We have, for $\varepsilon > 0$,

$$P\left\{\left|\left(\frac{X_1 + X_2 + ... + X_n}{n}\right) - E\left(\frac{X_1 + X_2 + ... + X_n}{n}\right)\right| < \varepsilon\right\} \geq 1 - \frac{B_n}{n^2\varepsilon^2}$$

where $\quad Var\left(\frac{X_1 + X_2 + ... + X_n}{n}\right)$

$$= \frac{1}{n^2}\, Var\,(X_1 + X_2 + ... + X_n) = \frac{B_n}{n^2}$$

or $\quad P\left\{\left|\left(\frac{X_1 + X_2 + ... + X_n}{n}\right) - \left(\frac{\mu_1 + \mu_2 + ... + \mu_n}{n}\right)\right| < \varepsilon\right\} \geq 1 - \frac{B_n}{n^2\varepsilon^2}$

Since ε is arbitrary, we assume $\frac{B_n}{n^2\varepsilon^2} \to 0$ as $n \to \infty$. Also choose small positive numbers ε and η and n_0 such that $\frac{B_n}{n^2\varepsilon^2} < n$ for $n > n_0$

$\therefore$ From above

$$P\left\{\left|\frac{X_1 + X_2 + ... + X_n}{n} - \frac{\mu_1 + \mu_2 + ... + \mu_n}{n}\right| < \varepsilon\right\} \geq 1 - \eta \quad \text{for all } n > n_0.$$

Another form of Weak Law of Large Numbers

Theorem III. *If $X_1, X_2, ..., X_n$ be a sequence of random variables with $\mu_1, \mu_2, ... \mu_n$ as means, then*

$$\overline{X}_n \xrightarrow{P} \overline{\mu}_n, \textit{ provided } \frac{B_n}{n^2} \to 0, \textit{ as } n \to \infty.$$

Proof. From above theorem, if

$$\overline{X}_n = \frac{X_1 + X_2 + \ldots + X_n}{n},$$

$$\overline{\mu}_n = \frac{\mu_1 + \mu_2 + \ldots + \mu_n}{n}, \text{ then}$$

$$P\{|\overline{X}_n - \overline{\mu}_n| < \varepsilon\} \geq 1 - \frac{B_n}{n^2\varepsilon^2} = 1 \qquad \Big| \because \frac{B_n}{n^2} \to 0, \text{ as } n \to \infty$$

$$\Rightarrow \qquad P\{|\overline{X}_n - \overline{\mu}_n| < \varepsilon\} = 1$$

$\therefore$ By definition, $\overline{X}_n \xrightarrow{P} \overline{\mu}_n$, provided $\frac{B_n}{n^2} \to 0$, as $n \to \infty$, where

$$B_n = \text{Var}(X_1 + X_2 + \ldots + X_n).$$

3.4 WEAK LAW OF LARGE NUMBERS FOR INDEPENDENT AND IDENTICALLY DISTRIBUTED RANDOM VARIABLES

Theorem IV. *If $X_1, X_2, \ldots, X_n$ are independent and identically distributed random variables i.e., if*

$$E(X_i) = \mu \text{ and } Var(X_i) = \sigma^2 \quad \text{for all } i = 1, 2, \ldots, n,$$

$$\text{then } \overline{X}_n \xrightarrow{P} \mu \text{ as } n \to \infty.$$

Proof. Here $B_n = \text{Var}(X_1 + X_2 + \ldots + X_n)$

$$= \text{Var}(X_1) + \text{Var}(X_2) + \ldots + \text{Var}(X_n)$$

$$= \sigma^2 + \sigma^2 + \ldots + \sigma^2 = n\sigma^2$$

$$| \because \quad X_1, X_2, \ldots, X_n \text{ are independent}$$

$$\therefore \qquad \underset{n\to\infty}{\text{Lt}} \frac{B_n}{n^2} = \underset{n\to\infty}{\text{Lt}} \frac{n\sigma^2}{n^2} = \underset{n\to\infty}{\text{Lt}} \frac{\sigma^2}{n} = 0$$

$$\therefore \qquad P\left\{\left|\frac{X_1 + X_2 + \ldots + X_n}{n} - \mu\right| < \varepsilon\right\} > 1 - \frac{B_n}{n^2\varepsilon^2} \ \forall\, n > n_0$$

or

$$P\left\{\left|\frac{X_1 + X_2 + \ldots + X_n}{n} - \mu\right| < \varepsilon\right\} \to 1, \text{ as } n \to \infty$$

$$\Rightarrow \qquad P\{|\overline{X}_n - \mu| \geq \varepsilon\} \to 0 \text{ as } n \to \infty$$

$$\Rightarrow \qquad \overline{X}_n \xrightarrow{P} \mu, \text{ as } n \to \infty$$

3.5 KOLMOGOROV'S STRONG LAW OF LARGE NUMBERS

If $x_1, x_2, \ldots, x_n$ are independent random variables with means $\mu_i = E(x_i)$ and variances σ^2.

Let $\qquad \delta n = E(x_1 + x_2 + \ldots + x_n)$ and

NOTES

$$E(\delta n) = E(x_1 + x_2 + ... + x_n) = E(x_1) + E(x_2) + ... + E(x_n)$$
$$= \mu_1 + \mu_2 + ... + \mu_n = m_n, \text{ say,}$$

then we say that the sequence $<\delta n>$ obeys the strong law of large numbers if for $\varepsilon > 0$, there exists $n \in N$ such that $\frac{|\delta_n - m_n|}{n} < \varepsilon \; \forall \; n > N$.

Markov's Theorem V. The weak law of large numbers holds if for some $\delta > 0$, $E(|X_i|^{1+\delta})$ exist and are bounded $\forall \; i = 1, 2, ...$

Markov's theorem gives only a necessary condition for the weak law of large numbers to hold good. It means that, if for $\delta > 0$, $E|X_i|^{1+\delta}$ is unbounded, then weak law of large numbers cannot hold for the sequence of random variables $X_1, X_2, ..., X_n$.

Khintchine's Theorem VI. If $X_1, X_2, ..., X_n$ are independent and identically distributed random variables, then the only condition necessary for the weak law of large numbers to hold is that $E(X_i)$, $i = 1, 2, ..., n$, should exist.

Necessary and sufficient condition for the sequence $<X_n>$ to satisfy weak law of large numbers.

Theorem VII. *A necessary and sufficient for the sequence $<X_n>$ to satisfy the weak law of large numbers is given by*

$$E\left(\frac{Y_n^2}{1 + Y_n^2}\right) \to 0, \text{ as } n \to \infty$$

where $$Y_n = \frac{S_n - E(S_n)}{n}, \; S_n = X_1 + X_2 + ... + X_n.$$

Proof. Let $E\left(\frac{Y_n^2}{1+Y_n^2}\right) \to 0$ as $n \to \infty$ and we show $<X_n>$, the sequence of random variables satisfies weak law of large numbers. For this, we define an event A, given by $A = \{|Y_n| \geq \varepsilon\}$

Let $\omega \in A \;\Rightarrow\; |Y_n| \geq \varepsilon \;\Rightarrow\; |Y_n|^2 \geq \varepsilon^2 > 0$

For $a \geq 0$, $b > 0$, we know $a \geq b \;\Rightarrow\; a + ab \geq b + ab$ ...(1)

$\therefore$ Taking $a = Y_n^2$ and $b = \varepsilon^2$ in (1), we define another event B as follows :

$$B = \left\{\frac{Y_n^2(1+\varepsilon^2)}{\varepsilon^2(1+Y_n^2)} \geq 1\right\} = \left\{\frac{Y_n^2}{(1+Y_n^2)} \geq \frac{\varepsilon^2}{1+\varepsilon^2}\right\}$$

Since $\omega \in A \;\Rightarrow\; \omega \in B, A \subseteq B \;\Rightarrow\; P(A) \leq P(B)$.

$$\therefore \quad P(|Y_n| > \varepsilon) \leq P\left(\frac{Y_n^2}{1+Y_n^2} \cdot \frac{1+\varepsilon^2}{\varepsilon^2} \geq 1\right)$$

$$\leq E\left\{\frac{Y_n^2/(1+Y_n^2)}{\varepsilon^2/(1+\varepsilon^2)}\right\} \quad \text{[By Markov's Inequality]}$$

$$\to 0 \text{ as } n \to \infty$$

$$P(|Y_n| \geq \varepsilon) \to 0 \text{ as } n \to \infty \;\Rightarrow\; \lim_{n \to \infty} P\left\{\left|\frac{S_n - E(S_n)}{n}\right| \geq \varepsilon\right\} \to 0$$

Thus weak law of large numbers holds for the sequence $\{X_n\}$ of random variables.

Converse. If $<X_n>$, the sequence of random variables, satisfies weak law of large numbers, we show

$E\left(\dfrac{Y_n^2}{1+Y_n^2}\right) \to 0$, as $n \to \infty$. For this, assume X_i's are continuous and if $f_n(y)$ be the probability density function of Y_n, then

$$E\left(\frac{Y_n^2}{1+Y_n^2}\right) = \int_{-\infty}^{\infty} \frac{y^2}{1+y^2} f_n(y)\,dy$$

$$= \int_{|y|\geq\varepsilon} \frac{y^2}{1+y^2} f_n(y)\,dy + \int_{|y|\leq\varepsilon} \frac{y^2}{1+y^2} f_n(y)\,dy$$

$$\Rightarrow \quad E\left(\frac{Y_n^2}{1+Y_n^2}\right) \leq \int_{|y|\geq\varepsilon} 1 . f_n(y)\,dy + \int_{|y|\leq\varepsilon} y^2 . f_n(y)\,dy$$

$$\left[\because \quad \frac{y^2}{1+y^2} < 1 \text{ and } \frac{y^2}{1+y^2} < y^2\right]$$

$$\leq P(A) + \varepsilon^2 \int_{A^c} f_n(y)\,dy \qquad (\because \quad \text{On } A^c : |y| < \varepsilon]$$

$$= P(A) + \varepsilon^2 . P(A^c) \leq P(A) + \varepsilon^2 \qquad [\because \quad P(A^c) < 1]$$

$$\Rightarrow \quad E\left\{\frac{Y_n^2}{1+Y_n^2}\right\} \leq P\{|Y_n| \geq \varepsilon\} + \varepsilon^2 \to 0, \text{ as } n \to \infty$$

Since $\{X_n\}$ satisfies weak law of large numbers, we have $\lim\limits_{n\to\infty} P|Y_n| \geq \varepsilon] \to 0$,

Hence the theorem.

ILLUSTRATIVE EXAMPLES

Example 1. *Consider a discrete random variable with probability density function f(x), given by*

$$f(x) = \frac{1}{8} . I_{-1}(x) + \frac{6}{8} . I_0(x) + \frac{1}{8} . I_1(x)$$

Evaluate $P[|X - \mu_x| \geq 2\sigma_x]$. *Also compare your result with that obtained on by using Chebyshev's inequality.*

Sol. Given probability distribution for the random variable X is shown in the following table.

x	-1	0	1
$p(x)$	$\frac{1}{8}$	$\frac{6}{8}$	$\frac{1}{8}$

$$\therefore \qquad E(X) = -1 \times \frac{1}{8} + 0 \times \frac{6}{8} + 1 \times \frac{1}{8} = 0 = \mu_x$$

$$E(X^2) = 1 \times \frac{1}{8} + 0 \times \frac{6}{8} + 1 \times \frac{1}{8} = \frac{2}{8} = \frac{1}{4}$$

$$\therefore \quad \text{Var(X)} = \text{E(X}^2) - (\text{E(X)})^2 = \frac{1}{4} - 0 = \frac{1}{4}$$

$$\Rightarrow \quad \sigma_x = \frac{1}{2}$$

$$\therefore \quad \text{P}\{|\text{ X} - \mu_x | \geq 2\sigma_x\} = \text{P}\{|\text{X}| \geq 1\} = 1 - \text{P}[|\text{ X }| < 1]$$
$$= 1 - \text{P}(-1 < \text{X} < 1) = 1 - \text{P(X} = 0)$$
$$= 1 - \frac{6}{8} = \frac{2}{8} = \frac{1}{4}$$

Also by Chebyshev's inequality,

$$\text{P}[|\text{ X} - \mu_x | \geq 2\,\sigma_x] \leq \frac{\text{Var(X)}}{(2\sigma_x)^2} = \frac{1}{4}$$

$\therefore$ In both cases, the results are same.

Example 2. *Consider the distribution $p(x) = 2^{-x}$, $x = 1, 2, 3, \ldots$ Using Chebyshev's inequality, show that*

$$P\{|X - 2| \leq 2\} > \frac{1}{2}, \text{ but the actual probability is } \frac{15}{16}.$$

Sol. We first find mean and variance of the given distribution.

Here $$\text{E(X)} = \sum_{x=1}^{\infty} x \,.\, p(x) = \sum_{x=1}^{\infty} x \,.\, \frac{1}{2^x}$$

$$= 1 \,.\, \frac{1}{2} + 2 \,.\, \frac{1}{2^2} + 3 \,.\, \frac{1}{2^3} + \ldots$$

$$= \frac{1}{2}\left(1 + 2 \,.\, \frac{1}{2} + 3 \,.\, \frac{1}{2^2} + \ldots\right) = \frac{1}{2}\left(1 - \frac{1}{2}\right)^{-2}$$

$$= \frac{1}{2} \,.\, 4 = 2 \qquad |\,(1 - x)^{-2} = 1 + 2x + 3x^2 + \ldots$$

$$\text{E(X}^2) = \sum_{x=1}^{\infty} x^2 \,.\, p(x) = \sum_{x=1}^{\infty} x^2 \,.\, \frac{1}{2^x}$$

$$= \frac{1}{2} + 2^2 \,.\, \frac{1}{2^2} + 3^2 \,.\, \frac{1}{2^3} + \ldots = \frac{1}{2}\left(1 + 4 \,.\, \frac{1}{2} + 9 \,.\, \frac{1}{2^2} + \ldots\right)$$

$$= \frac{1}{2}\,(1 + x)\,(1 - x)^{-3} \text{ where } x = \frac{1}{2} \qquad |\text{ See remark below}$$

$$= \frac{1}{2}\,\frac{3}{2} \,.\, \left(\frac{1}{2}\right)^{-3} = \frac{3}{4} \,.\, 8 = 6$$

$$\therefore \quad \text{Var(X)} = \text{E(X}^2) - (\text{E(X)})^2 = 6 - 4 = 2 \Rightarrow \sigma = \sqrt{2}$$

Using Chebyshev's inequality, we have

$$\text{P}\{|\text{ X} - \text{E(X)} | < k\sigma\} \geq 1 - \frac{1}{k^2}$$

Take $k\sigma = 2 \Rightarrow k = \frac{2}{\sigma} = \frac{2}{\sqrt{2}} = \sqrt{2}$, we have

$$P[\,|X - 2| < 2] \geq 1 - \frac{1}{2} = \frac{1}{2}$$

Also the actual probability is given by

$$P\{|X - 2| \leq 2\} = P(2 - 2 \leq X \leq 2 + 2)$$

$$= P(0 \leq X \leq 4) = P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4)$$

$$= 0 + \frac{1}{2} + \frac{1}{2^2} + \frac{1}{2^3} + \frac{1}{2^4} = \frac{15}{16}.$$

Remark: Consider $S = 1 + 2^2 \,.\, a + 3^2 \,.\, a^2 + 4^2 \,.\, a^3 + 5^2 \,.\, a^4 + ...$ ...(1)

Multiplying (1) by $-3a$, $3a^2$, $-a^3$ successively, we have

$$S = 1 + 4a + 9a^2 + 16a^3 + 25a^4 + ... \quad ...(2)$$

$$-3aS = -3a - 12a^2 - 27a^3 - 48a^4 - ...... \quad ...(3)$$

$$3a^2S = 3a^2 + 12a^3 + 27a^4 + ... \quad ...(4)$$

$$-a^3S = -a^3 - 4a^4 - ...... \quad ...(5)$$

Adding (2), (3), (4) and (5) vertically, we get

$$(1 - 3a + 3a^2 - a^3)\, S = 1 + a$$

or $$(1 - a)^3 S = 1 + a \Rightarrow S = (1 + a)(1 - a)^{-3}$$

Take $a = \frac{1}{2}$ in above, we get

$$1 + 4 \,.\, \frac{1}{2} + 9 \,.\, \frac{1}{2^2} + 16 \,.\, \frac{1}{2^3} + ... = \left(1 + \frac{1}{2}\right)\left(1 - \frac{1}{2}\right)^{-3}$$

$$= \frac{3}{2} \,.\, 8 = 12.$$

Example 3. *A symmetric die is thrown 600 times. Find the lower bound for the probability of getting 80 to 120 sixes.*

Sol. Let X denotes the total number of successes, p be the probability of getting a six, then $p = \frac{1}{6}$. Also $n = 600$.

$$\therefore \quad E(X) = np = 600 \times \frac{1}{6} = 100$$

$$Var(X) = npq = 100 \times \frac{5}{6} = \frac{500}{6} \qquad | \text{ Using Binomial distribution}$$

Using Chebyshev's inequality,

$$P\{|X - E(X)| < k\sigma\} \geq 1 - \frac{1}{k^2}$$

or $$P\left\{|X - 100| < k\sqrt{\frac{500}{6}}\right\} \geq 1 - \frac{1}{k^2}$$

or $$P\left(100 - k\sqrt{\frac{500}{6}} < X < 100 + k\sqrt{\frac{500}{6}}\right) \geq 1 - \frac{1}{k^2} \quad ...(1)$$

NOTES

Take $100 + k\sqrt{\frac{500}{6}} = 120 \Rightarrow k\sqrt{\frac{500}{6}} = 20$

$$\Rightarrow \quad k = 20\sqrt{\frac{6}{500}}$$

$\therefore$ From (1),

$$P(80 < X < 120) \geq 1 \frac{500}{6} \cdot \frac{1}{400} = \frac{19}{24}.$$

Example 4. *Consider a sequence of random variables defined by*

$$P(X_k = \pm 2k) = 2^{-2k-1},\ P(X_k = 0) = 1 - 2^{-2k}.$$

Examine whether weak law of large numbers can be applied to the sequence $<X_k>$ of random variables.

Sol. Since X_k assumes the values 2^k, $-2k$ and 0.

$$\therefore \quad E(X_k) = 2^k 2^{-(2k+1)} + (-2^k) \cdot 2^{-(2k+1)} + 0 \times (1 - 2^{-2k})$$
$$= 2^{-(2k+1)}(2^k - 2^k) = 0$$
$$E(X_k^2) = (2^k)^2 \cdot 2^{-(2k+1)} + (-2^k)^2 \cdot 2^{-(2k+1)} + 0^2 \times (1 - 2^{-2k})$$
$$= 2^{2k} \cdot 2^{-(2k+1)} + 2^2 \cdot 2^{-(2k+1)} = \frac{1}{2} + \frac{1}{2} = 1$$

Also $\quad \mathrm{Var}(X_k) = E(X_k^2) - \{E(X_k)\}^2 = 1 - 0 = 1$

$$\therefore \quad B_n = \mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathrm{Var}(X_i) = \sum_{i=1}^{n} (1) = n$$

[$X_1, X_2, ..., X_n$ are independent]

$$\Rightarrow \quad \lim_{n\to\infty} \frac{B_n}{n^2} = \lim_{n\to\infty} \frac{1}{n} \to 0$$

Hence Weak Law of large numbers, holds for the sequence $<X_k>$ of independent random variables.

Example 5. *Consider a sequence $<X_k>$ of independent and indentically distributed random variables defined as below.*

$$P\{X_i = (-1)^{k-1} \cdot k\} = \frac{6}{\pi^2 k^2},\ k = 1, 2, 3, ..., i = 1, 2, 3, ...$$

Examine whether weak law of large numbers can be applied to the sequence $<X_k>$ of random variables.

Sol. We shall apply Khintchine's theorem

Here $$E(X_i) = \sum_{k=1}^{\infty} (-1)^{k-1} \cdot k \cdot \frac{6}{\pi^2 k^2} = \frac{6}{\pi^2} \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} \quad ...(1)$$

R.H.S. of (1) is an alternating series with $a_k = \frac{1}{k} \to 0$, as $k \to \infty$. Also $a_{k+1} = \frac{1}{k+1}$

Since $k + 1 > k$

$$\Rightarrow \quad \frac{1}{k+1} < \frac{1}{k}$$

$$\Rightarrow \quad a_{k+1} < a_k \ \forall\ k$$

$\therefore$ (1) is convergent by using Leibnitz's test.

$\therefore$ From (1),

$$E(X_i) = \frac{6}{\pi^2}\left(1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} \ldots\right)$$

$$= \frac{6}{\pi^2} \log 2 = \text{finite}$$

By Khintchine's theorem, weak law of large numbers can be applied to the sequence $<X_i>$ of independent and identically distributed random variables.

Example 6. *If $X_1, X_2, \ldots, X_n$ be a sequence of random variables with equal expectations and finite variation. Examine whether weak law of large numbers can be applied to the sequence $<X_i>$. It is given that all the covariances σ_{ij} are negative.*

Sol. Since $\text{Var}(X_i) < \infty$, consider

$$\frac{B_n}{n^2} = \frac{\text{Var}\,(X_1 + X_2 + \ldots + X_n)}{n^2} = \frac{1}{n^2}\left\{\sum_{i=1}^{n} \sigma_i^{\,2} + 2\sum\sum_{i<j=1}^{n} \sigma_{ij}\right\}$$

$$< \frac{1}{n^2}\left(\sum_{i=1}^{n} \sigma_i^{\,2}\right), \text{ if all the covariances } \sigma_{ij} \text{ are negative}$$

$$\therefore \quad \lim_{n\to\infty} \frac{B_n}{n^2} < \lim_{n\to\infty} \frac{1}{n^2}\left(\sum_{i=1}^{n} \sigma_i^{\,2}\right) = 0 \qquad (\because \ \sigma_i^2;\ i = 1, 2, \ldots, n \text{ are finite})$$

Hence weak law of large numbers holds.

Statement of Central Limit Theorem

Theorem VIII. If $X_1, X_2, \ldots, X_n$ are n independent random variables with means $\mu_1, \mu_2, \ldots, \mu_n$ and variances $\sigma_1^{\,2}, \sigma_2^{\,2}, \ldots, \sigma_n^{\,2}$ respectively, then under certain good conditions, the sum $S_n = X_1 + X_2 + \ldots + X_n$ is asymptotically normal with means and standard deviation σ where

$$\mu = \mu_1 + \mu_2 + \ldots + u_n\ ;\ \sigma^2 = \sigma_1^{\,2} + \sigma_2^{\,2} + \ldots + \sigma_n^{\,2}.$$

Note. The above theorem was stated by Laplace and its proof was given by Liapounov's under certain general conditions.

Theorem IX. De Moivre's Central Limit Theorem

If $X_1, X_2, \ldots, X_n$ be a sequence of independent and identically distributed random variables such that

$$X_i = \begin{cases} 1, \text{with probability } p \\ 0, \text{with probability } q, \end{cases}$$

Then the sum $S_n = X_1 + X_2 + \ldots + X_n$ is asymptotically normal with mean μ and variance σ^2.

Proof. Let $M_{X_i}(t)$ denotes the moment generating function of the random variable X_i, then

$$M_{X_i}(t) = E(e^{tX_i}) = e^{t.1}\, p + e^{t.0}\,.\, q = pe^t + q.$$

NOTES

If $M_{S_n}(t)$ denotes the moment generating function of the sum $S_n = X_1 + X_2 + \ldots\ldots + X_n$, then

$$M_{S_n}(t) = M_{X_1 + X_2 + \ldots\ldots + X_n}(t)$$

$$= M_{X_1}(t), M_{X_2}(t) \ldots\ldots M_{X_n}(t)$$

$$= (M_{X_i}(t))^n \qquad | X_1, X_2, \ldots\ldots X_n \text{ are independent and identically distributed}$$

$$= (q + pe^t)^n$$

which is the moment generating function of the Binomial distribution with parameters n and p. Therefore, by uniqueness theorem of moment generating functions, $S_n \sim B(n, p)$

$\therefore$ $\qquad E(S_n) = np = \mu$, say, $V(S_n) = npq = \sigma^2$, say,

Take $\qquad Z = \dfrac{S_n - E(S_n)}{\sqrt{Var(S_n)}} = \dfrac{S_n - \mu}{\sigma}$

If $M_Z(t)$ denotes the moment generating function of Z, then

$$M_Z(t) = M_{\frac{S_n - \mu}{\sigma}}(t)$$

$$= e^{-\frac{\mu t}{\sigma}} M_{S_n}\left(\frac{t}{\sigma}\right)$$

$$= e^{-\frac{npt}{\sqrt{npq}}} M_{S_n}\left(\frac{t}{\sigma}\right), \text{ where } S_n \sim N(\mu, \sigma^2)$$

$$= e^{\frac{-npt}{\sqrt{npq}}}\left[q + pe^{\frac{t}{\sigma}}\right]^n$$

$$= e^{-\frac{npt}{\sqrt{npq}}}\left(q + pe^{\frac{t}{\sqrt{npq}}}\right)^n$$

$$= \left(q\, e^{-\frac{pt}{\sqrt{npq}}} + pe^{\frac{t}{\sqrt{npq}} - \frac{pt}{\sqrt{npq}}}\right)^n$$

$$= \left(q\, e^{-\frac{pt}{\sqrt{npq}}} + pe^{\frac{tq}{\sqrt{npq}}}\right)^n$$

$$= \left[q\left(1 - \frac{pt}{\sqrt{npq}} + \frac{p^2t^2}{npq}\cdot\frac{1}{2!}\ldots\ldots\right) + p\left(1 + \frac{tq}{\sqrt{npq}} + \frac{t^2q^2}{npq}\cdot\frac{1}{2!} + \ldots\ldots\right)\right]^n$$

$$= \left[(p+q) + \left(\frac{ptq}{\sqrt{npq}} - \frac{ptq}{\sqrt{npq}}\right) + \frac{t^2}{2n}(p+q) + O(n^{-3/2})\right]^n$$

$$= \left[1+\frac{t^2}{2n}+O(n^{-3/2})\right]^n$$

NOTES

where O $(n^{-3/2})$ represents terms involving $n^{3/2}$ and its higher power of n in the denominator. Taking $n \to \infty$, we get

$$\underset{n\to\infty}{\text{Lt}}\ M_Z(t) = \underset{n\to\infty}{\text{Lt}} \left[1+\frac{t^2}{2n}+O(n^{-3/2})\right]^n$$

$$= \underset{n\to\infty}{\text{Lt}} \left[1+\frac{t^2}{2n}\right]^n = e^{t^2/2}$$

which is the moment generating function of the standard normal variate. Therefore, by uniqueness theorem of moment generating functions, $Z = \dfrac{S_n - \mu}{\sigma}$ is asymptotically normal with mean 0 and variance 1.

$\therefore$ $S_n = X_1 + X_2 + \ldots\ldots + X_n$ is asymptotically normal with mean μ and variance σ^2.

3.6 STATE AND PROVE LINDERBERG-LEVY FORM OF CENTRAL LIMIT THEOREM

Theorem X. *If $X_1, X_2, \ldots\ldots, X_n$ are n independent and identically distributed random variables with mean μ and variance σ^2 i.e., $E(X_i) = \mu$ and $Var(X_i) = \sigma^2\ \forall i = 1, 2, \ldots\ldots n$, then the sum $S_n = X_1 + X_2 + \ldots\ldots + X_n$ is asymptotically normal with mean $n\mu$ and variance $n\sigma^2$ respectively. It is given that $E(X_i^2)$ exists.*

Proof. Given $\mu'_1 = E(X_i - \mu) = 0$

$\mu'_2 = E(X_i - \mu)^2 = \sigma^2$

Let $M(t)$ denotes the moment generating function of each of the deviation $X_i - \mu$, then

$$M(t) = 1 + \mu'_1 t + \mu'_2 \frac{t^2}{2!} + \mu'_3 \frac{t^3}{3!} + \ldots\ldots$$

$$= 1 + \sigma^2 \frac{t^2}{2} + O(t^3) \qquad \ldots(1)$$

where $O(t^3)$ contains terms in t^3 and its higher order

Take $$Z = \frac{S_n - n\mu}{\sqrt{n\sigma^2}} = \frac{X_1 + X_2 + \ldots\ldots + X_n - n\mu}{\sqrt{n}\sigma}$$

$$= \sum_{i=1}^{n} \left(\frac{X_i - \mu}{\sqrt{n}\,\sigma}\right)$$

$$\therefore \quad M_Z(t) = M_{\sum_{i=1}^{n} \frac{X_i-\mu}{\sqrt{n}\sigma}}(t) = M_{\sum_{i=1}^{n}(X_i-\mu)}(t/\sigma)$$

NOTES

$$= \prod_{i=1}^{n} M_{(X_i - \mu)}\left(\frac{t}{\sqrt{n}\sigma}\right) \qquad | X_1, X_2, ..., X_n \text{ are independent}$$

$$= \left[M\left(\frac{t}{\sqrt{n}\,\sigma}\right)\right]^n = \left[1 + \frac{t^2}{2n} + O(n^{-3/2})\right]^n \qquad | \text{Using (1)}$$

$$\underset{n\to\infty}{\text{Lt}}\ M_Z(t) = \underset{n\to\infty}{\text{Lt}} \left[1 + \frac{t^2}{2n} + O(n^{-3/2})\right]^n = \underset{n\to\infty}{\text{Lt}}\left(1 + \frac{t^2}{2n}\right)^n = e^{\frac{t^2}{2}}$$

where the term $O(n^{-3/2}) \to 0$, as $n \to \infty$.

$\therefore$ By uniqueness theorem of moment generating function, $Z = \dfrac{S_n - n\mu}{\sqrt{n}\sigma}$ is asymptotically normal with mean $n\mu$ and variance $n\sigma^2$.

3.7 LINDEBERG AND FELLER CONDITIONS (STATEMENT ONLY)

If $<X_n>$ be a sequence of independently and identically distributed random variables with mean $\mu = E(X_i)$ and variances σ^2, $(i = 1, 2, ..., n)$ then the sum $x_1 + x_2 + ... + x_n$ is asymptotically normal with mean $n\mu$ and variance $n\sigma^2$ under the following conditions, which are due to "Lindeberg and Feller"

(*i*) The random variables $x_1, x_2, ..., x_n$ are independent and identically distributed.

(*ii*) $E(x_i^2)$ exists for all $i = 1, 2, ..., n$.

2.8 ANOTHER FORMS OF CENTRAL UNIT THEOREM OR LIAPOUNOV'S FORM OF CENTRAL LIMIT THEOREM

Theorem. XI. If $X_1, X_2, ..., X_n$ be a sequence of n independent and identically distributed random variables with mean μ and variance σ^2, then

$$\underset{n\to\infty}{\text{Lt}}\ P\left[a \le \frac{S_n - n\mu}{\sqrt{n}\sigma} \le b\right] = \phi(b) - \phi(a) = \frac{1}{\sqrt{2\pi}} \int_{\infty}^{b} e^{-x^2/2}\, dx \qquad ...(1)$$

where $\phi(.)$ is the distribution function of the standard normal variate and

$$S_n = X_1 + X_2 + ... + X_n$$

(1) can also be written as

$$\underset{n\to\infty}{\text{Lt}}\ P\left\{a \le \frac{S_n - E(S_n)}{\sqrt{\text{Var}(S_n)}} \le b\right\} = \phi(b) - \phi(a).$$

Applications of Central Limit Theorem

Theorem XII. *If $X_1, X_2, ..., X_n$ are n independent and identically distributed random variables following Binomial distribution with parameter r and p. Let*

$$S_n = X_1 + X_2 + ... + X_n, \text{ then}$$

NOTES

$$\underset{n \to \infty}{\text{Lt}} \text{ P}\left\{a \le \frac{S_n - nrp}{\sqrt{nrp\,(1-p)}} \le b\right\} = \phi(b) - \phi(a),\ 0 < p < 1.$$

Proof. $E(S_n) = E(X_1 + X_2 + \ldots\ldots + X_n)$

$= E(X_1) + E(X_2) + \ldots\ldots + E(X_n)$

$= rp + rp + \ldots\ldots + rp$

$= nrp$ $\qquad X_i \sim B(r, p)$

$\therefore$ mean $= rp$ variance $= rpq$

Also $\quad \text{Var}(S_n) = \text{Var}(X_1 + X_2 + \ldots\ldots + X_n)$

$= \text{Var}(X_1) + \text{Var}(X_2) + \ldots\ldots \text{Var}(X_n)$

$= rpq + rpq + \ldots\ldots n$ times $\qquad | \ X_i$'s are independent

$= nrpq = nrp(1 - p)$

By Central Limit Theorem, we have

$$\underset{n \to \infty}{\text{Lt}} \text{ P}\left\{a \le \frac{S_n - E(S_n)}{\sqrt{\text{Var}(S_n)}} \le b\right\} = \phi(b) - \phi(a)$$

or

$$\underset{n \to \infty}{\text{Lt}} \text{ P}\left[a \le \frac{S_n - nrp}{\sqrt{nrp\,(1-p)}} \le b\right] = \phi(b) - \phi(a).$$

Theorem XIII. *If Y_n is a Binomial variate with parameters n and p, then*

$$\underset{n \to \infty}{Lt} \ P\left[a \le \frac{Y_n - np}{\sqrt{np\,(1-p)}} \le b\right] = \phi(b) - \phi(a),\ 0 < p < 1.$$

Proof. We know that if $X_1, X_2, \ldots\ldots X_n$ is a sequence of independent and identically distributed random variables following Bernouli's distribution *i.e.,* $B(1, p)$, then

$$X_1 + X_2 + \ldots\ldots + X_n \sim B(n, p)$$

But it is given $Y_n \sim B(n, p)$, therefore, take $Y_n = S_n$ and hence

$$E(Y_n) = np,\ \text{Var}(Y_n) = npq$$

By Central Limit Theorem,

$$\underset{n \to \infty}{\text{Lt}} \text{ P}\left[a \le \frac{S_n - E(S_n)}{\sqrt{\text{Var}(S_n)}} \le b\right] = \phi(b) - \phi(a)$$

$$\Rightarrow \quad \underset{n \to \infty}{\text{Lt}} \text{ P}\left[a \le \frac{Y_n - np}{\sqrt{npq}} \le b\right] = \phi(b) - \phi(a)$$

$$\Rightarrow \quad \underset{n \to \infty}{\text{Lt}} \text{ P}\left[a \le \frac{Y_n - np}{\sqrt{np\,(1-p)}} \le b\right] = \phi(b) - \phi(a),\ 0 < p < 1$$

Theorem XIV. *If Y_n is distributed as Poisson distribution with parameter n, then*

$$\underset{n \to \infty}{Lt} \ P\left[a \le \frac{Y_n - n}{\sqrt{n}} \le b\right] = \phi(b) - \phi(a)$$

Also $\quad P(Y_n \le n) \to \dfrac{1}{2}$ *as* $n \to \infty$.

NOTES

Proof. If $X_1, X_2, ..., X_n$ are n independent and identically distributed random variables following Poisson distributions with parameter n, then

$$X_1 + X_2 + ... + X_n \sim P(n)$$

$$\Rightarrow \quad S_n \sim P(n). \text{ Take } Y_n = S_n$$

$$\therefore \quad Y_n \sim P(n)$$

Now mean of Poisson distribution is n and also variance of Poisson distribution is also n

i.e., $\quad E(Y_n) = n, \text{Var}(Y_n) = n$

$\therefore$ By Central Limit Theorem,

$$\underset{n \to \infty}{\text{Lt}} P\left[a \le \frac{S_n - E(S_n)}{\sqrt{\text{Var}(S_n)}} \le b\right] = \phi(b) - \phi(a)$$

$$\Rightarrow \quad \underset{n \to \infty}{\text{Lt}} P\left[a \le \frac{Y_n - n}{\sqrt{n}} \le b\right] = \phi(b) - \phi(a)$$

Take $a = -\infty$, $b = 0$ in above, we get

$$\underset{n \to \infty}{\text{Lt}} P\left(a \le \frac{Y_n - n}{\sqrt{n}} \le b\right) = \underset{n \to \infty}{\text{Lt}} P\left(\frac{Y_n - n}{\sqrt{n}} \le 0\right)$$

$$= \underset{n \to \infty}{\text{Lt}} P(Y_n \le n) \quad ...(1)$$

$$\text{Also} \quad \phi(b) - \phi(a) = \phi(0) - \phi(-\infty) = \frac{1}{2} \quad ...(2)$$

From (1) and (2),

$$P(Y_n \le n) \to \frac{1}{2} \text{ as } n \to \infty.$$

ILLUSTRATIVE EXAMPLES

Example 1. *Let $X_1, X_2, ..., X_n$ are independent and identically distributed random variables following Poisson distribution with parameter λ. Use Central Limit Theorem, find $P(120 \le Sn \le 160)$ where $S_n = X_1 + X_2 + ... + X_n$, $\lambda = 2$, $n = 75$.*

Sol. Given $X_1, X_2, ..., X_n$ are independent and identically distributed Poisson variates with parameter λ, it implies

$$E(X_i) = \lambda, \text{Var}(X_i) = \lambda \ \forall \ i = 1, 2, ..., n$$

$$\therefore \quad E(S_n) = E(X_1 + X_2 + ... + X_n)$$

$$= E(X_1) + E(X_2) + ... + E(X_n) \quad | X_i\text{'s are independent}$$

$$= \lambda + \lambda + ... + \lambda$$

$$= n\lambda$$

$$\text{Var}(S_n) = \text{Var}(X_1 + X_2 + ... + X_n)$$

$$= \text{Var}(X_1) + \text{Var}(X_2) + ... + \text{Var}(X_n)$$

$$= \lambda + \lambda + ... + \lambda = n\lambda$$

Using Linderberg-Levy theorem, for equation, $S_n \sim N(n\lambda, n\lambda)$. But $n = 75$, $\lambda = 2$

$\therefore \quad S_n \sim N(150, 150) = N(\mu, \sigma^2)$ where $\mu = 150$, $\sigma^2 = 150$

NOTES

Take $\quad Z = \dfrac{X - 150}{\sqrt{150}}$

When $\quad X = 120,\ Z = \dfrac{120 - 150}{\sqrt{150}} = \dfrac{-30}{\sqrt{150}} = -2.45$

When $\quad X = 160,\ Z = \dfrac{160 - 150}{\sqrt{150}} = \dfrac{10}{\sqrt{150}} = 0.82$

$$\therefore \quad P(120 \le S_n \le 160) = P\left(\frac{120 - 150}{\sqrt{150}} \le Z \le \frac{160 - 150}{\sqrt{150}}\right)$$

$$= P(-2.45 \le Z \le 0.82) \qquad | \ Z \sim N(0, 1)$$

$$= P(-2.45 \le Z \le 0) + P(0 \le Z < 0.82)$$

$$= P(0 \le Z \le 2.45) + P(0 \le Z \le 0.82)$$

$$= 0.4929 + 0.2939 = 0.7868 \text{ (Using Normal Probability Tables)}$$

SUMMARY

- If a sequence of random variables $x_1, x_2, ..., x_n$ converges to x in probability then the following are true.

  If $X_n \xrightarrow{P} a,\ Y_n \xrightarrow{P} b,$ as $n \to \infty$, then

  (*i*) $X_n \pm Y_n \xrightarrow{P} a \pm b,$ as $n \to \infty$,

  (*ii*) $\dfrac{X_n}{Y_n} \xrightarrow{P} \dfrac{a}{b}, b \ne 0,$ as $n \to \infty$.

GLOSSARY

- **WLLN.** If the sequence $<X_i>$ of random variables has means $\mu_1, \mu_2, ..., \mu_n$. then

  $\overline{X}_n \xrightarrow{P} \overline{\mu}_n$, provided $\dfrac{B_n}{n^2} \to 0$ as $n \to \infty$

  where $B_n = N$ or $(X_1 + X_2 + ... + X_n)$

- **C.L.T.** If $<X_i>$ is a sequence of random variables with means $\mu_1, \mu_2, ..., \mu_n$ and variances $\sigma_1^2, \sigma_2^2, ..., \sigma_n^2$. Then the sum $S_n = X_1 + X_2 + ... + X_n$ is a normal variable with mean μ and variance σ^2, where

$$\mu = \mu_1 + \mu_2 + ... + \mu_n; \quad \sigma^2 = \sigma_1^2 + \sigma_2^2 + ... + \sigma_n^2$$

REVIEW QUESTIONS

1. A random variable X has a mean value of 5 and variance of 3.
   (*i*) What is the least value of Prob $\{| X - 5 | < 3\}$?
   (*ii*) What value of h guarantees that Prob $\{| X - 5 | < h\} \ge 0.99$?
   (*iii*) What is the least value of Prob $(| X - 5 | < 7.5)$?

NOTES

2. If X denote the sum of the numbers obtained when two dice are thrown, use Chebyshev's inequality to obtain an upper bounds for $P(|X - 7| > 4)$. Compare this with the actual probability.
3. An unbiased coin is tossed 100 times. Show that the probability that the number of heads will be between 30 and 70 is greater than 0.93.
4. Within what limits will the number of heads lie, with 95% probability, in 1,000 tosses of a coin which is practically unbiased?
5. A symmetric die is thrown 720 times. Use Chebyshev's inequality to find the lower bound for the probability of getting 100 to 140 sixes.
6. Use Chebyshev's inequality to determine how many times a fair coin must be tossed in order that the probability will be at least 0.95 that the ratio of the number of heads to the number of tosses will be between 0.45 and 0.55.
7. If X is a *r.v.* such that $E(X) = 3$ and $E(X^2) = 13$, use Chebyshev's inequality to determine a lower bound for $P(-2 < X < 8)$.
8. State and prove Chebyshev's inequality. Use it to prove that in 2,000 throws with a coin the probability that the number of heads lies between 900 and 1,100 is at least 19/20.
9. A random variable X has the density function e^{-x} for $x \geq 0$. Show that Chebychev's inequality gives, $P(|X - 1| > 2) < \frac{1}{4}$ and show that the actual probability is e^{-3}.
10. Let X have the p.d.f.
$$f(x) = \begin{cases} \frac{1}{2\sqrt{3}}, & -\sqrt{3} < x < \sqrt{3} \\ 0 \quad, & \text{elsewhere} \end{cases}$$
Find the actual probability $P\left(|X - \mu| \geq \frac{3}{2}\sigma\right)$ and compare it with the upper bound obtained by Chebyshev's inequality.
11. If X has the distribution with *p.d.f.* $f(x) = e^{-x}$, $0 \leq x < \infty$, use Chebyshev's inequality to obtain a lower bound to probability of the inequality $-1 \leq X \leq 3$, and compare it with the actual probability.
12. If $X_1, X_2, ..., X_n$ be *r.v.*'s. with means $\mu_1, \mu_2, ..., \mu_n$ and standard deviations $\sigma_1, \sigma_2, ..., \sigma_n$ respectively, and $[\text{Var}(X_1 + X_2 + .. + X_n)]/n^2 \to 0$ as $n \to \infty$, show that $\overline{X}_n - \overline{\mu}_n$ converges to zero in probability.
Hence show that if m is the number of successes in n independent trials, the probability of success at ith trial being p_i then m/n converges in probability to $(p_1 + p_2 + ... + p_n)/n$.
13. $\{X_k\}$, $k = 1, 2, ...$ is a sequence of independent random variables each taking the values : $-1, 0, 1$. Given that $P(X_k = 1) = \frac{1}{k} = P(X_k = -1)$, $P(X_k = 0) = 1 - \frac{2}{k}$, examine if the law of large numbers holds for this sequence.
14. Examine whether the weak law of large numbers holds good for the sequence X_n of independent random variables, where $P\left(X_n = \frac{1}{\sqrt{n}}\right) = \frac{2}{3}$, $P\left(X_n = -\frac{1}{\sqrt{n}}\right) = \frac{1}{3}$.
15. $\{X_n\}$ is a sequence of independent random variables such that
$$P\left(X_n = \frac{1}{\sqrt{n}}\right) = p_n,\ P\left(X_n = 1 + \frac{1}{\sqrt{n}}\right) = 1 - p_n$$
Examine whether the weak law of large numbers is applicable to the sequence $\{X_n\}$.

16. If X is a *r.v.* and $E(X^2) < \infty$, then prove that $P\{|X| \geq a\} \leq \frac{1}{a^2} E(X^2)$, for all $a > 0$.

Use Chebychev's inequality to show that for $n > 36$, the probability that in n throws of a fair die, the number of sixes lies between $\frac{1}{6}n - \sqrt{n}$ and $\frac{1}{6}n + \sqrt{n}$ is at least $\frac{31}{36}$.

17. Let $\{X_n\}$ be a sequence of mutually independent random variables such that :

$X_n = \pm 1$ with probability $\frac{1-2^{-n}}{2}$ and $X_n = \pm 2^{-n}$ with probability 2^{-n-1}.

Examine whether the weak law of large numbers can be applied to the sequence $\{X_n\}$.

18. Examine whether the laws of large numbers holds for the sequence $\{X_k\}$ of independent random variables defined by $P(X_k = \pm k^{-1/2}) = \frac{1}{2}$.

19. State and prove Weak Law of Large Numbers. Determine whether it holds for the following sequence of independent random variables :

$$P(X_n = \pm 1) = \frac{1}{2}(1 - 2^{-n}) = P(X_n = -1).$$

20. A distribution with unknown mean μ has variance equal to 1.5. Use central limit theorem to find how large a sample should be taken from the distribution in order that the probability will be at least 0.95 that the sample mean will be within 0.5 of the population mean.

21. The life time of a certain brand of an electric bulb may be considered a random variable with mean 1,200 hours and standard deviation 250 hours. Find the probability using central limit theorem, that the average life-time of 60 bulbs exceeds 1,400 hours.

22. Decide whether the central limit theorem holds for the sequence of independent random variables X_r with distribution defined as $P(X_r = 1) = p_r$ and $P(X_r = 0) = 1 - p_r$.

23. If $X_1, X_2, X_3, \ldots$ is a sequence of independent random variables having the uniform densities :

$$f_i(x_i) = \begin{cases} 1/(2 - i^{-1}), & 0 < x_i < 2 - i^{-1} \\ 0 & , \text{elsewhere} \end{cases}$$ show that the central limit theorem holds.

24. Let $\overline{X}_n$ be the sample mean of a random sample of size n from Rectangular distribution on [0, 1]. Let $U_n = \sqrt{n}\left(\overline{X}_n - \frac{1}{2}\right)$.

Show that $F(u) = \lim\limits_{n \to \infty} P(U_n < u)$ exists and determine it.

25. Let $X_1, X_2, \ldots\ldots$ be a sequence of independent, identically distributed non-negative random variables such that $E(\log X_1)^2$ is finite. $Z_n = (X_1, X_2 \ldots\ldots, X_n)^{1/n}$. Show that the positive constant c can be so chosen that the random variable $(c Z_n)^{\sqrt{n}}$ has a non-degenerate limit distribution function F(.) and determine F(.).

26. $\{X_n\}$ is a sequence of *i.i.d.* random variables. If n is a perfect square, then X_n is Cauchy variate with density : $\frac{1}{\pi} \cdot \frac{1}{1+x^2}, -\infty < x < \infty$. Otherwise X_n has a distribution function $F(x)$ with mean zero and finite variance σ^2. Discuss the asymptotic distribution of $(X_1 + X_2 + \ldots + X_n)/\sqrt{n}$.

27. Let $\{X_k\}$, $k \geq 1$ be a sequence of *i.i.d.* variates with $f(x) = \frac{1}{2} e^{-|x|}, -\infty < x < \infty$.

Find the constants a_n and b_n such that $\{| X_1 | + | X_2 | + ... + | X_n | - a_n\}/b_n \xrightarrow{L}$ N(0,1).

NOTES

28. Using C.L.T., show that : $\lim_{n \to \infty} \left(e^{-n} \sum_{k=0}^{n} \frac{n^k}{k!} \right) = \frac{1}{2} = \lim_{n \to \infty} \int_0^n \frac{e^{-t} . t^{n-1}}{(n-1)!} dt.$

29. Let $\{X_n, n = 1, 2, ...\}$ be a sequence of independent Bernoulli variates such that :

$$P(X_n = 1) = p_n = 1 - P(X_n = 0),\ n = 1, 2, ..., (q_n = 1 - p_n).$$

Show that if $\sum p_n q_n = \infty$, $(n = 1, 2, ..., \infty)$, then the CLT holds for the sequence $\{X_n\}$. What happens if $\sum p_n q_n < \infty$.

30. Derive weak low of large numbers from Chebychev's inequality.

31. State and prove WLLN for *i.i.d.* random variables.

32. The necessary condition for the WLLN to hold is that $E(X_i)$, $i = 1, 2, ..., n$, should exist. Write the name of this theorem.

FURTHER READINGS

1. Discrete Distributions: N.L. Johnson and S. Kotz, John Wiley and Sons
2. Continuous Univarate distribution–1: N.L. Johnson and S. Kotz
3. Continuous Univarate distributions–2: N.L. Johnson and S. Kotz, John Wiley and Sons
4. Introduction to Probability theory with applications: W. Feller, Vol-1: Wiley astern.

UNIT–III

CHAPTER 4

DISCRETE DISTRIBUTIONS

NOTES

OBJECTIVES

After going through this chapter, you should be able to:

- know the conditions for the applicability of binomial distribution
- know the techniques of solving the problems by using binomial distribution
- know the mean, varriance, S.D. and central moments for binomial distribution
- m.g.f. of the binomial distribution.

STRUCTURE

NOTES

I. BINOMIAL DISTRIBUTION

4.1 INTRODUCTION

The binomial distribution is a particular type of probability distribution. This was discovered by **James Bernoulli** (1654–1705) in the year 1700. This distribution mainly deals with attributes. An attribute is either present or absent with respect to elements of a population. For example, if a coin is tossed, we get either *head* or *tail*. The workers of a factory may be classified as *skilled* and *unskilled*.

4.2 CONDITIONS FOR APPLICABILITY OF BINOMIAL DISTRIBUTION

The following conditions are essential for the applicability of Binomial Distribution :

(*i*) **The random experiment is performed for a finite and fixed number of trials.** If in an experiment, a coin is tossed repeatedly or a ball is drawn from an urn repeatedly, then each toss or draw is called a **trial.** For example, if a coin is tossed 6 times, then this experiment has 6 trials. The number of trials is an experiment is generally denoted by 'n'.

(*ii*) **The trials are independent.** By this we mean that the result of a particular trial should not affect the result of any other trial. For example, if a coin is tossed or a die is thrown, then the trials would be independent. If from a pack of playing cards, some draws of one card are made without replacing the cards, then the trials would not be independent. But, if the card drawn is replaced before the next draw, the trials would be independent.

(*iii*) **Each trial must result in either "success" or "failure".** In other words, in every trial, there should be only two possible outcomes *i.e., success* or *failure*. For example, if a coin is tossed, then either *head or tail* is observed. Similarly, if an item is examined, it is either *defective* or *non-defective*.

(*iv*) **The probability of success in each trial is same.** In other words, this condition requires that the probability of *success* should not change in different trials. For example, if a sample of two items is drawn, then the probability of exactly one being defective will be constant in different trials provided the items are replaced before the next draw.

4.3 BINOMIAL VARIABLE

A random variable which counts the number of successes in a random experiment with trials satisfying above four conditions is called a **binomial variable.**

For example, if a coin is tossed 5 times and the event of getting head is *success*, then the possible values of the binomial variable are 0, 1, 2, 3, 4, 5. This is so, because, the minimum number of successes is 0 and the maximum number of successes is 5.

4.4 BINOMIAL PROBABILITY FUNCTION

When a fair coin is tossed, we have only two possibilities : head and tail. Let us call the occurrence of head as 'success'. Therefore, the occurrence of tail would be a 'failure'. Let this coin be tossed 5 times. Suppose we are interested in finding the probability of getting 4 heads and 1 tail *i.e.*, of getting 4 successes. If S and F denote 'success' and 'failure' in a trial respectively, then there are ${}^5C_4 = 5$ ways of having 4 successes.

These are : SSSSF, SSSFS, SSFSS, SFSSS, FSSSS.

The probability of getting 4 successes in each case is $\left(\frac{1}{2}\right)^4\left(\frac{1}{2}\right)$, because the trials are independent.

$\therefore$ By using *addition theorem*, the required probability of having 4 successes is ${}^5C_4\left(\frac{1}{2}\right)^4\left(\frac{1}{2}\right)$, which is equal to $\frac{5}{32}$. Now we shall generalise this method of finding the probabilities for different values of a *binomial variable*.

Let a random experiment satisfying the conditions of **binomial distribution** be performed. Let the number of trials in the experiment be n. Let p denotes the probability of *success* in any trial.

$\therefore$ Probability of failure, $q = 1 - p$.

Let x denotes the binomial variable corresponding to this experiment.

$\therefore$ The possible values of x are 0, 1, 2, ..., n.

If there are r successes in n trials, then there would be $n - r$ failures. One of the ways in which r successes may occur is

$$\underbrace{SS.............S}_{r\text{ times}}\ \underbrace{FF.............F}_{n-r\text{ times}}$$

where S and F denote success and failure in the trials.

Now, P(SS .... SFF ...... F) = P(S)P(S) ...... P(S)P(F)P(F) ..... P(F)

($\because$ the trials are independent)

$$= p.p \,.......\, p.q.q \,.......\, q = p^r q^{n-r}.$$

NOTES

We know that ${}^nC_r$ is the number of combinations of n things taking r at a time. Therefore, the number of ways in which r successes can occur in n trials is equal to the number of ways of choosing r trials (for successes) out of total n trials *i.e.*, it is ${}^nC_r$. Therefore, there are ${}^nC_r$ ways in which we get r successes and $n-r$ failures and the probability of occurrence of each of these ways is $p^r q^{n-r}$. Hence the probability of r successes in n trials in any order is

$$P(x = r) = p^r q^{n-r} + p^r q^{n-r} + \ldots\ldots {}^nC_r \text{ terms} \quad \text{(By addition theorem)}$$

or $\quad \mathbf{P(x = r) = {}^nC_r b^r q^{n-r}, 0 \le r \le n.}$

This is called the **binomial probability function.** The corresponding binomial distribution is

x	0	1	2 .... n
$P(x)$	${}^nC_0 p^0 q^n$	${}^nC_1 p^1 q^{n-1}$	${}^nC_2 p^2 q^{n-2}$ .... ${}^nC_n p^n q^0$

The probabilities of 0 success, 1 success, 2 successes, ..., n successes are respectively the 1st, 2nd, 3rd, ..., $(n + 1)$th terms in binomial expansion of $(q + p)^n$. This is why, it is called **binomial distribution.**

4.5 BINOMIAL FREQUENCY DISTRIBUTION

If a random experiment, satisfying the requirements of binomial distribution, is repeated N times, then the expected frequency of getting $r (0 \le r \le n)$ successes is given by

$$\mathbf{N.\ P(x = r) = N.\ {}^nC_r\ p^r q^{n-r},\ 0 \le r \le n.}$$

The *frequencies* of getting 0 success, 1 success, 2 successes, ..., n successes are respectively the 1st, 2nd, 3rd, ..., $(n + 1)$th terms in the expansion of $N(q + p)^n$.

4.6 HISTOGRAM OF BINOMIAL DISTRIBUTION

We know the method of drawing histogram of a frequency distribution. The method of drawing histogram of a binomial distribution is analogous to the procedure of drawing histogram of a frequency distribution. In case of a binomial distribution, we mark all the values of the random variable on the horizontal axis and their respective probabilities on the vertical axis. Rectangles of uniform width are constructed with values of the variable at centre and heights equal to their corresponding probabilities.

WORKING RULES FOR SOLVING PROBLEMS

I. Make sure that the trials in the random experiment are independent and each trial result in either 'success' or 'failure'.

II. Define the binomial variable and find the values of n and p from the given data. Also find q by using : $q = 1 - p$.

III. Put the values of n, p and q in the formula :

$$P(r \text{ successes}) = {}^nC_r\, p^r q^{n-r},\ r = 0, 1, 2, \ldots\ldots, n \qquad \ldots(1)$$

IV. Express the event, whose probability is desired, in terms of values of the binomial variable x. Use (1) to find the required probability.

NOTES

ILLUSTRATIVE EXAMPLES

Example 1. *An unbiased coin is tossed 10 times. Find, by using binomial distribution, the probability of getting at least 3 heads.*

Sol. Let p be the probability of success, *i.e.*, of getting head in the toss of the coin.

$$\therefore \quad n = 10, \; p = \frac{1}{2} \text{ and } q = 1 - p = 1 - \frac{1}{2} = \frac{1}{2}$$

Let x be the binomial variable, "*no. of successes*".

By **Binomial distribution,** $P(x = r) = {}^nC_r \, p^r q^{n-r}, 0 \le r \le n$.

$$\therefore \quad P(x = r) = {}^{10}C_r \left(\frac{1}{2}\right)^r \left(\frac{1}{2}\right)^{10-r} = {}^{10}C_r \left(\frac{1}{2}\right)^{10} = {}^{10}C_r \frac{1}{1024}, 0 \le r \le 10.$$

Now, P(at least 3 heads) $= P(x \ge 3) = 1 - P(x < 3)$

$$= 1 - [P(x = 0 \text{ or } x = 1 \text{ or } x = 2)]$$

$$= 1 - [P(x = 0) + P(x = 1) + P(x = 2)]$$

$$= 1 - \left[{}^{10}C_0 \frac{1}{1024} + {}^{10}C_1 \frac{1}{1024} + {}^{10}C_2 \frac{1}{1024}\right]$$

$$= 1 - \frac{1}{1024}[{}^{10}C_0 + {}^{10}C_1 + {}^{10}C_2]$$

$$= 1 - \frac{1}{1024}[1 + 10 + 45] = \frac{1024 - 56}{1024} = \frac{968}{1024} = \frac{121}{128}.$$

Example 2. *A coin is tossed 7 times. What is the probability that head appears an odd number of times.*

Sol. Let p be the probability of success, *i.e.*, of getting a head.

$$\therefore \quad n = 7, \; p = \frac{1}{2} \text{ and } q = 1 - p = 1 - \frac{1}{2} = \frac{1}{2}$$

Let x be the Binomial variable "*no. of successes*".

By **Binomial distribution,** $\quad P(x = r) = {}^nC_r \, p^r q^{n-r}, 0 \le r \le n$.

$$\therefore \quad P(x = r) = {}^7C_r \left(\frac{1}{2}\right)^r \left(\frac{1}{2}\right)^{7-r}$$

$$= {}^7C_r \left(\frac{1}{2}\right)^7 = {}^7C_r \left(\frac{1}{128}\right), 0 \le r \le 7.$$

Required probability = P(head appearing an odd number of times)

$$= P(x = 1 \text{ or } 3 \text{ or } 5 \text{ or } 7) = P(x = 1) + P(x = 3) + P(x = 5) + P(x = 7)$$

$$= {}^7C_1 \left(\frac{1}{128}\right) + {}^7C_3 \left(\frac{1}{128}\right) + {}^7C_5 \left(\frac{1}{128}\right) + {}^7C_7 \left(\frac{1}{128}\right)$$

$$= (7 + 35 + 21 + 1)\left(\frac{1}{128}\right) = \frac{64}{128} = \frac{1}{2}.$$

Example 3. *Draw a histogram for the binomial probability distribution of the number of heads in 5 tosses of coin.*

Sol. Let p be the probability of success, *i.e.*, of getting a head.

$$\therefore \quad n = 5, \; p = \frac{1}{2} \text{ and } q = 1 - p = 1 - \frac{1}{2} = \frac{1}{2}.$$

Let x be the Bionomial variable "no. of successes".

$\therefore \qquad x = 0, 1, 2, ..., 5.$

NOTES

By **Binomial distribution,**

$$P(x = r) = {}^nC_r \, p^r \, q^{n-r}, \; 0 \le r \le n.$$

$$\therefore \qquad P(x = r) = {}^5C_r \left(\frac{1}{2}\right)^r \left(\frac{1}{2}\right)^{5-r} = {}^5C_r \left(\frac{1}{2}\right)^5 = {}^5C_r \left(\frac{1}{32}\right), \; 0 \le r \le 5.$$

$$\therefore \qquad P(x = 0) = {}^5C_0 \left(\frac{1}{32}\right) = \frac{1}{32}, \qquad P(x = 1) = {}^5C_1 \left(\frac{1}{32}\right) = \frac{5}{32}$$

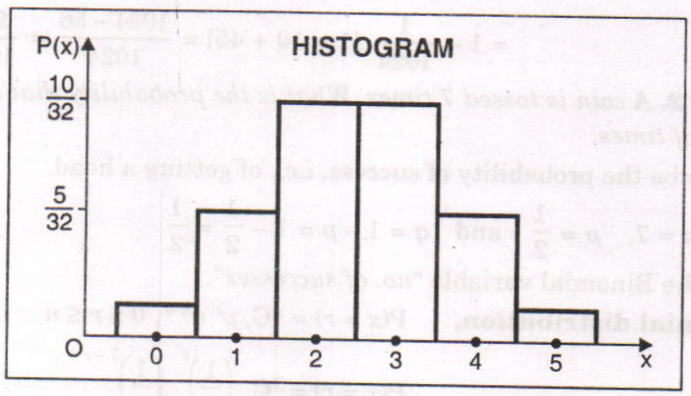
$$P(x = 2) = {}^5C_2 \left(\frac{1}{32}\right) = \frac{10}{32}, \qquad P(x = 3) = {}^5C_3 \left(\frac{1}{32}\right) = \frac{10}{32}.$$

$$P(x = 4) = {}^5C_4 \left(\frac{1}{32}\right) = \frac{5}{32}, \qquad P(x = 5) = {}^5C_5 \left(\frac{1}{32}\right) = \frac{1}{32}.$$

$\therefore$ The required probability distribution is

x	0	1	2	3	4	5
$P(x)$	$\frac{1}{32}$	$\frac{5}{32}$	$\frac{10}{32}$	$\frac{10}{32}$	$\frac{5}{32}$	$\frac{1}{32}$

The histogram of the Binomial Probability Distribution is shown in the figure :



Example 4. *In a hurdle race, a player has to cross 10 hurdles. The probability that he will clear each hurdle is 5/6. What is the probability that he will knock down fewer than 2 hurdles ?*

Sol. Let p be the probability of success, *i.e.*, of knocking down a hurdle.

$$\therefore \qquad n = 10, \;\; p = 1 - \frac{5}{6} = \frac{1}{6} \quad \text{and} \quad q = 1 - p = 1 - \frac{1}{6} = \frac{5}{6}.$$

Let x be the binomial variable *"no. of successes"*.

By **Binomial distribution,** $\qquad P(x = r) = {}^nC_r \, p^r \, q^{n-r}, \; 0 \le r \le n.$

$$\therefore \qquad P(x = r) = {}^{10}C_r \left(\frac{1}{6}\right)^r \left(\frac{5}{6}\right)^{10-r}, \; 0 \le r \le 10.$$

Now, P(knocking down fewer than 2 hurdles)

$$= P(x < 2) = P(x = 0) + P(x = 1) = {}^{10}C_0 \left(\frac{1}{6}\right)^0 \left(\frac{5}{6}\right)^{10} + {}^{10}C_1 \left(\frac{1}{6}\right)^1 \left(\frac{5}{6}\right)^9$$

$$= \left[1\times1\times\left(\frac{5}{6}\right)^{10}\right] + \left[10\times\frac{1}{6}\times\left(\frac{5}{6}\right)^{9}\right] = \left(\frac{5}{6}\right)^{9}\left[\frac{5}{6}+\frac{10}{6}\right] = \left(\frac{5}{6}\right)^{9}\left(\frac{5}{2}\right).$$

NOTES

Example 5. *The probability of a man hitting a target is 1/4. He fires 7 times. What is the probability of his hitting the target at least twice ?*

Sol. Let p be the probability of success, *i.e.*, of hitting the target.

$\therefore \quad n = 7, \quad p = \frac{1}{4}$ and $q = 1 - p = 1 - \frac{1}{4} = \frac{3}{4}.$

Let x be the binomial variable *"no. of successes"*.

By **Binomial distribution,** $\quad P(x = r) = {}^nC_r\, p^r\, q^{n-r},\ 0 \le r \le n.$

$$\therefore \quad P(x = r) = {}^7C_r \left(\frac{1}{4}\right)^r \left(\frac{3}{4}\right)^{7-r}, 0 \le r \le 7.$$

Now, $\quad$ P(hitting at least twice) = $P(x \ge 2)$

$$= 1 - P(x < 2) = 1 - [P(x = 0) + P(x = 1)]$$

$$= 1 - \left[{}^7C_0\left(\frac{1}{4}\right)^0\left(\frac{3}{4}\right)^7 + {}^7C_1\left(\frac{1}{4}\right)^1\left(\frac{3}{4}\right)^6\right]$$

$$= 1 - \left[1\times1\times\left(\frac{3}{4}\right)^7 + 7\times\frac{1}{4}\times\left(\frac{3}{4}\right)^6\right] = 1 - \left(\frac{3}{4}\right)^6\left(\frac{3}{4}+\frac{7}{4}\right) = \frac{4547}{8192}.$$

Example 6. *The probability that a bulb produced by a factory will fuse after 100 days of use is 0 05. Find the probability that out of 5 such bulbs :*

(i) none $\qquad$ *(ii) not more than one*

(iii) more than one $\qquad$ *(iv) at least one,*

will fuse after 100 days of use.

Sol. Let p be the probability of success, *i.e.*, the bulb being fused after 100 days.

$\therefore \quad n = 5, \quad p = 0.05 = \frac{5}{100} = \frac{1}{20}$ and $q = 1 - p = 1 - \frac{1}{20} = \frac{19}{20}.$

Let x be the binomial variable *"no. of successes"*.

By **Binomial distribution,** $\quad P(x = r) = {}^nC_r\, p^r\, q^{n-r},\ 0 \le r \le n.$

$$\therefore \quad P(x = r) = {}^5C_r\left(\frac{1}{20}\right)^r\left(\frac{19}{20}\right)^{5-r}, 0 \le r \le 5.$$

(i) P(none will fuse) = $P(x = 0) = {}^5C_0\left(\frac{1}{20}\right)^0\left(\frac{19}{20}\right)^5 = \left(\frac{19}{20}\right)^5.$

(ii) P(not more than one will fuse) = $P(x \le 1) = P(x = 0) + P(x = 1)$

$$= {}^5C_0\left(\frac{1}{20}\right)^0\left(\frac{19}{20}\right)^5 + {}^5C_1\left(\frac{1}{20}\right)^1\left(\frac{19}{20}\right)^4$$

$$= \left(\frac{19}{20}\right)^4\left[1\times1\times\frac{19}{20} + 5\times\frac{1}{20}\times1\right] = \left(\frac{19}{20}\right)^4\left(\frac{6}{5}\right).$$

NOTES

(*iii*) P(more than one will fuse) $= P(x > 1) = 1 - P(x \le 1) = 1 - \left(\frac{19}{20}\right)^4 \left(\frac{6}{5}\right)$

(Using part (*ii*))

(*iv*) P(at least one will fuse) $= P(x \ge 1) = 1 - P(x < 1) = 1 - P(x = 0)$

$$= 1 - {}^5C_0 \left(\frac{1}{20}\right)^0 \left(\frac{19}{20}\right)^5 = 1 - \left(\frac{19}{20}\right)^5.$$

Example 7. *A bag contains 25 items of which 5 are defective. A random sample of two is drawn (without replacement). What is the probability that (i) of both being good (ii) of both being bad (iii) at least one being good.*

Sol. Let p = The probability of getting a success *i.e.,* the probability of having a defective item.

Given, $$p = \frac{5}{25} = \frac{1}{5} = 0.2$$

$$q = 1 - p = 1 - 0.2 = 0.8$$

Let x be a random variable following the Binomial distribution, then

$$P(X = r) = {}^nC_r\, q^{n-r}\, p^r,\ 0 \le r \le n$$

(*i*) Required probability = P(both items are good)

$= 1 -$ P(none is good)

$= 1 -$ P(all items are defective)

$= 1 - P(x = 2)$

$= 1 - {}^2C_2\,(0.8)^0\,(0.2)^2$

$= 1 - 0.04 = 0.96$

(*ii*) Required probability = P(both items are bad)

$= P(X = 2)$

$= {}^2C_2\,(0.8)^0\,(0.2)^2$

$= 0.04$

(*iii*) Required probability = P(at least one item is good)

$= 1 -$ P(at most one item is bad)

$= 1 - P(X \le 1)$

$= 1 - (P(X = 0) + P(X = 1))$

$= 1 - P(X = 0) - P(X = 1)$

$= 1 - {}^2C_0\,(0.8)^2\,(0.2)^0 - {}^2C_1\,(0.8)^1\,(0.2)^1$

$= 1 - 0.64 - 2 \times 0.8 \times 0.2$

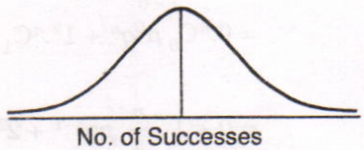
$= 1 - 0.64 - 0.32$

$= 1 - 0.96 = 0.04.$

NOTES

PROPERTIES OF BINOMIAL DISTRIBUTION

4.7 SHAPE OF BINOMIAL DISTRIBUTION



The shape of the binomial distribution depends upon the probability of success (p) and the number of trials in the experiment. If $p = q = \frac{1}{2}$, then the distribution will be symmetrical for every value of n. If $p \neq q$, then the distribution would be asymmetrical, *i.e.*, skewed. The magnitude of skewness varies as thedifference between p and q.

The probabilities in binomial distribution depends upon n and p. These are called the **parameters** of the distribution.

4.8 LIMITING CASE OF BINOMIAL DISTRIBUTION

As number of trials (n) in the binomial distribution increases, the number of successes also increases. If neither p nor q is very small, then as n approaches infinity, the skewness in the distribution disappears and it becomes continuous. Such a continuous, bell shaped distribution is called *normal distribution*. Thus, the normal distribution is limiting case of binomial distribution as n approaches infinity.

4.9 MEAN OF BINOMIAL DISTRIBUTION

Let x be a binomial variable and $P(x = r) = {}^nC_r p^r q^{n-r}$, $0 \leq r \leq n$.

Here n is the number of trials and p, the probability of success in a trial.

The *mean* of x is the average number of successes.

$\therefore$ Mean,

$$\mu = \sum_{r=0}^{n} r\,.P(x = r) = \sum_{r=0}^{n} r\,.{}^nC_r p^r q^{n-r}$$

$$= 0\,.\,{}^nC_0 p^0 q^n + 1.{}^nC_1 p^1 q^{n-1} + 2.{}^nC_2 p^2 q^{n-2} + \ldots.. + n.{}^nC_n p^n\,.\,q^0$$

$$= 0 + n\,.\,pq^{n-1} + \frac{n\,(n-1)}{1.2}\,p^2 q^{n-2} + \ldots\ldots. + n.1.p^n$$

$$= np\left\{q^{n-1} + \frac{n-1}{1}\,pq^{n-2} + \ldots\ldots\ldots + p^{n-1}\right\}$$

$$= np\,\{{}^{n-1}C_0\,p^0 q^{n-1} + {}^{n-1}C_1\,p^1 q^{n-2} + \ldots.. + {}^{n-1}C_{n-1}\,p^{n-1}\,q^0\}$$

$$= np(q + p)^{n-1} = np\,(1)^{n-1} = np.$$

$\therefore$ **Mean (μ) of x = np.**

4.10 VARIANCE AND S.D. OF BINOMIAL DISTRIBUTION

Let x be a binomial variable and $P(x = r) = {}^nC_r p^r q^{n-r}$, $0 \leq r \leq n$.

Here n is the number of trials and p, the probability of success in a trial.

The variance and standard deviation of x measures the dispersion of the binomial distribution and are given by

$$\text{Variance} = \sum_{r=0}^{n} r^2 .\text{P}(x=r) - \mu^2 \quad \text{and} \quad \text{S.D.} = \sqrt{\sum_{r=0}^{n} x^2 .\text{P}(x=r) - \mu^2}.$$

NOTES

$$\text{Now,} \quad \sum_{r=0}^{n} r^2 .\text{P}(x=r) = \sum_{r=0}^{n} r^2 .{}^n\text{C}_r\, p^r q^{n-r}$$

$$= 0.{}^n\text{C}_0\, p^0 q^n + 1^2.{}^n\text{C}_1\, p^1 q^{n-1} + 2^2\, {}^n\text{C}_2\, p^2 q^{n-2} + 3^2.{}^n\text{C}_3\, p^n q^{n-3} + \ldots\ldots + n^2 .\ {}^n\text{C}_n\, p^n q^0$$

$$= 0 + 1 . \frac{n}{1}\, pq^{n-1} + 2^2 . \frac{n\,(n-1)}{1.2}\, p^2 q^{n-2} + \frac{3^2 . n\,(n-1)(n-2)}{1.2.3}\, p^3 q^{n-3} + \ldots\ldots + n^2 . 1 . p^n\ 1$$

$$= np\left\{q^{n-1} + \frac{2(n-1)}{1}\, pq^{n-2} + \frac{3(n-1)(n-2)}{1\times 2}\, p^2 q^{n-3} + \ldots . + np^{n-1}\right\}$$

$$= np\left\{\left(q^{n-1} + \frac{(n-1)}{1}\, pq^{n-2} + \frac{(n-1)(n-2)}{1\times 2}\, p^2 q^{n-3} + \ldots . + p^{n-1}\right) + \left(\frac{(n-1)}{1}\, pq^{n-2} + \frac{2(n-1)(n-2)}{1\times 2}\, p^2 q^{n-3} + \ldots . + (n-1)\, p^{n-1}\right)\right\}$$

$$= np\ \{(q+p)^{n-1} + (n-1)\, p(q^{n-2} + (n-2)\, pq^{n-3} + \ldots\ldots + p^{n-2})\}$$

$$= np\ \{1 + (n-1)\, p(q+p)^{n-2}\} = np\ \{1 + (n-1)\, p.1\}$$

$$= np\ \{1 + np - p\} = np + n^2p^2 - np^2$$

$$\therefore \quad \text{Variance} = \sum_{r=0}^{n} x^2 .\text{P}(x=r) - \mu^2 = (np + n^2p^2 - np^2) - n^2p^2 = np - np^2$$

$$= np(1-p) = \mathbf{npq}.$$

$$\text{Also,} \qquad \textbf{S.D.} = \sqrt{\textbf{Variance}} = \sqrt{\textbf{npq}}\ .$$

Theorem I. *Show that the first four moments about origin for the Binomial distribution are given as*

$$\mu_1' = np,\ \mu_2' = n(n-1)\, p^2 + np$$

$$\mu_3' = n(n-1)\,(n-2)p^3 + 3n\,(n-1)\, p^2 + np$$

$$\mu_4' = n(n-1)\,(n-2)\,(n-3)\, p^4 + 6n\,(n-1)\,(n-2)\, p^3 + 7n\,(n-1)\, p^2 + np$$

From above results, derive the first four central moments viz, μ_1, μ_2, μ_3 *and* μ_4*. Also find the values of* β_1, β_2 *and* γ_1, γ_2 *respectively.*

Proof. Let X denotes the random variable which follows binomial distribution with parameters n and p, then

$$p(x) = \text{P}(\text{X} = x) = {}^n\text{C}_x\, q^{n-x} p^x$$

$\therefore$ The moments about origin are as unter

$$\mu_1' = \text{E}(\text{X}) = \sum_{x=0}^{n} x\ {}^n\text{C}_x\ p^x q^{n-x} = np \sum_{x=1}^{n} {}^{n-1}\text{C}_{x-1}\, p^{x-1} q^{n-x}$$

$$= np(q+p)^{n-1} = np$$

NOTES

$$^nC_x = \frac{n}{x} \cdot {}^{n-1}C_{x-1} = \frac{n(n-1)}{x(x-1)} \cdot {}^{n-2}C_{x-2} \text{ etc.}$$

Hence the mean of the binomial distribution is np.

$$\mu'_2 = E(X^2) \sum_{x=0}^{n} x^2 \; {}^nC_x \; p^x q^{n-x}$$

$$= \sum_{x=0}^{n} \{x(x-1) + x\} \frac{n(n-1)}{x(x-1)} \cdot {}^{n-2}C_{x-2} \; p^x q^{n-x}$$

$$= n(n-1) p^2 \left\{ \sum_{x=2}^{n} {}^{n-2}C_{x-2} \; p^{x-2} q^{n-x} \right\} + np$$

$$= n(n-1) p^2 (q+p)^{n-2} + np = n(n-1)p^2 + np$$

| Using Binomial theorem,

$$\mu'_3 = E(X^3) = \sum_{x=0}^{n} x^3 \cdot p(x) = \sum_{x=0}^{n} \{x(x-1)(x-2) + 3x(x-1) + x\} \; {}^nC_x \; p^x q^{n-x}$$

$$= n(n-1)(n-2) p^3 \sum_{x=3}^{n} {}^{n-3}C_{x-3} \; p^{x-3} q^{n-x}$$

$$+ 3n(n-1) p^2 \sum_{x=2}^{n} {}^{n-2}C_{x-2} \; p^{x-2} q^{n-x} + np$$

$$= n(n-1)(n-2) p^3 (q+p)^{n-3} + 3n(n-1) p^2 (q+p)^{n-2} + np$$

$$= n(n-1)(n-2)p^3 + 3n(n-1)p^2 + np$$

Consider $x^4 = Ax(x-1)(x-2)(x-3) + Bx(x-1)(x-2) + Cx(x-1) + Dx$

Put $x = 1$, in above, we get $D = 1$

Put $x = 2$, in above, we get $2C + 2D = 16 \;\Rightarrow\; C = 7$

Put $x = 3$, in above, we get $6B + 6C + 3D = 81$

$\Rightarrow \quad 6B = 81 - 42 - 3 = 36 \;\Rightarrow\; B = 6$

Equate the coefficient of x^4, $1 = A$

$\therefore \quad x^4 = x(x-1)(x-2)(x-3) + 6x(x-1)(x-2) + 7x(x-1) + x$

$$\therefore \quad \mu'_4 = E(X^4) = \sum_{x=0}^{n} x^4 \; {}^nC_x \; p^x q^{n-x}$$

$$= n(n-1)(n-2)(n-3) p^4 + 6n(n-1)(n-2) p^3 + 7n(n-1) p^2 + np$$

| Using Binomial theorem

To derive central Moments of Binomial Distribution:

$$\mu_2 = \mu'_2 - \mu_1'^2 = n^2p^2 - np^2 + np - n^2p^2 = np(1-p) = npq$$

$$\mu_3 = \mu'_3 - 3\mu'_2 \mu'_1 + 2\mu_1'^3$$

$$= \{n(n-1)(n-2)p^3 + 3n(n-1)p^2 + np\} - 3\{n(n-1)p^2 + np\} np + 2(np)^3$$

$$= np(-3np^2 + 3np + 2p^2 - 3p + 1 - 3npq)$$

$$= np\{3np(1-p) + 2p^2 - 3p + 1 - 3npq\}$$

$$= np(2p^2 - 3p + 1) = np(2p^2 - 2p + q) = npq(1 - 2p)$$

$$= npq\{q + p - 2p\} = npq(q - p)$$

NOTES

$$\mu_4 = \mu'_4 - 4\mu'_3\mu'_1 + 6\mu'_2\mu'^2_1 - 3\mu'^4_1 = npq\,\{1 + 3\,(n-2)\,pq\}$$

| Using $\mu'_1, \mu'_2, \mu'_3, \mu'_4$.

Also $$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{n^2p^2q^2\,(q-p)^2}{n^3p^3q^3} = \frac{(q-p)^2}{npq} = \frac{(1-2p)^2}{npq}$$

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{npq\,\{1+3\,(n-2)\,pq\}}{n^2p^2q^2} = \frac{1+3\,(n-2)\,pq}{npq} = 3 + \frac{1-6pq}{npq}$$

$$\gamma_1 = \sqrt{\beta_1} = \frac{q-p}{\sqrt{npq}} = \frac{1-2p}{\sqrt{npq}},\ \gamma_2 = \beta_2 - 3 = \frac{1-6pq}{npq}$$

Cor. For the Binomial distribution, variance is less than mean.

Proof. Let X ~ B(n, p) *i.e.,* X follows the Binomial distribution with parameters n and p, then, from above theorem,

mean = np, variance = $npq < np$ = mean. | $0 < q < 1$

Theorem II. *Find the mean deviation about mean of the Binomial distribution.*

Proof. Let X ~ B (n, p) *i.e.,* X follows the Binomial distribution with parameters n and p, then

$$p(x) = p(X = x) = {}^nC_x\, q^{n-x}\, p^x$$

Let n denotes the mean deviation about the mean np. Then

$$\eta = \sum_{x=0}^{n} |x-np|\,p(x) = \sum_{x=0}^{n} |x-np|\ {}^nC_x\ p^x q^{n-x} \qquad (x \text{ being an integer})$$

$$= \sum_{x=0}^{np} -(x-np)\ {}^nC_x\ p^x q^{n-x} + \sum_{x=np}^{n} (x-np)\ {}^nC_x\ p^x q^{n-x}$$

$$= 2\sum_{x=np}^{n} (x-np)\ {}^nC_x\ p^x q^{n-x}$$

$$\left[\because\ \sum_{x=0}^{n} x\ {}^nC_x\ p^x q^{n-x} = np \ \Rightarrow\ \sum_{x=0}^{n} (x-np)\ {}^nC_x\ p^x\ q^{n-x} = 0\right]$$

$$= 2\sum_{\mu}^{n} (x-np)\ {}^nC_x\ p^x q^{n-x},$$

$$= 2\sum_{\mu}^{n} \left[\{xq - (n-x)\,p\}\ {}^nC_x\ p^x q^{n-x}\right]$$

$$= 2\sum_{\mu}^{n} \left[\frac{n\,!}{(x-1)\,!\,(n-x)\,!}\,p^x\,q^{n-x+1} - \frac{n\,!}{x\,!\,(n-x-1)\,!}\,p^{x+1}q^{n-x}\right]$$

$$= 2\sum_{x=\mu}^{n} (t_{x-1} - t_x),\ \text{Take } t_x = \frac{n\,!}{x\,!\,(n-x-1)\,!}\,p^{x+1}q^{n-x}$$

$$= 2\,(t_{\mu-1} - t_n) = 2t_{\mu-1} \qquad |\ t_n = 0$$

$$= 2\,\frac{n\,!}{(\mu-1)\,!\,(n-\mu)\,!}\,.\,p^{\mu}\,q^{n-\mu+1} = 2npq\ {}^{n-1}C_{\mu-1}\,p^{\mu-1}\,q^{n-\mu}.$$

| where μ is greatest integer contained in $np + 1$.

NOTES

Moment Generating Function

Theorem III. *To find the moment generating function of the Binomial distribution.*

Proof. Let X ~ B (n, p) *i.e.,* X follows the Binomial distribution with parameters n and p, then

$$p(x) = P(X = x) = {}^nC_x\, q^{n-x}\, p^x$$

Let $M_X(t)$ denotes the moment generating function of the random variables X, then

$$M_X(t) = E\,(e^{tx}) = \sum_{x=0}^{n} e^{tx}\ {}^nC_x\, q^{n-x}\ p^x.$$

$$= \sum_{x=0}^{n} {}^nC_x (pe^t)^x\, q^{n-x} = (q + pe^t)^n.$$

Probability Generating Function

Let P(S) be the probability generating function of the Binomial distribution, then,

$$P(S) = \sum_{k=0}^{n} P(X = k)\, s^k = \sum_{k=0}^{n} ({}^nC_k p^k q^{n-k})\, s^k$$

$$= \sum_{k=0}^{n} {}^nC_k\, (ps)^k\, q^{n-k} = (ps + q)^4$$

$\therefore$ The p.g.f. is the nth power of $(ps + q)$.

Example. *Prove* $E\left(\frac{1}{x+a}\right) = \int_0^1 t^{a-1}\, G(t)\, dt, a > 0$ *where* $G(t)$ *is the p.g.f. of X. Find also* $E\left(\frac{1}{x+1}\right)$ *when* $x \sim B(n_1 p)$.

Sol. Consider $\int_0^1 t^{a-1}\, G(t)\, dt = \int_0^1 t^{a-1}\, E(t^x)\, dt$

$$= \int_0^1 t^{a-1}\left(\sum_{x=0}^{n} p(x)\, t^x\right) dt$$

$$= \sum_{x=0}^{n}\left[p(x)\int_0^1 t^{a-1+x}\, dt\right]$$

$$= \sum_{x=0}^{n}\left[p(x)\,.\left[\frac{t^{a-1+x+1}}{a-1+x+1}\right]_0^1\right]$$

$$= \sum_{x=0}^{n} p(x)\,.\,\frac{1}{a+x} = E\left(\frac{1}{x+a}\right) \qquad ...(1)$$

Also, If $x \sim B(n, p)$, then $G(t) = \sum_{x=0}^{n} t^x\ p(x)$

$$= \sum_{x=0}^{n} t^x \, {}^nC_x p^x q^{n-x} = \sum_{x=0}^{n} {}^nC_x (pt)^x q^{n-x} = (q+pt)^n \qquad \text{...(2)}$$

NOTES

For $a = 1$, from (1), we have

$$E\left(\frac{1}{x+1}\right) = \int_0^1 G(t)\,dt = \int_0^1 (q+bt)^n \qquad | \text{ Using (1)}$$

$$= \left[\frac{(q+pt)^{n+1}}{(n+1)p}\right]_0^1 = \frac{(q+p)^{n+1} - q^{n+1}}{(n+1)b}$$

$$= \frac{1-q^{n+1}}{(n+1)p}.$$

4.11 REPRODUCTIVITY PROPERTY (IF EXISTS)

Statement. If X ~ B(n_1, p_1) and Y ~ B(n_2, p_2) are two independent random variables then X + Y does not satisfy the additive or reproductive property. In other words, the sum of two independent binomial variates is not a binomial variate.

Proof. Given X is a Binomial variate with parameters n_1 and p_1, therefore its moment generating function is

$$M_X(t) = (q_1 + p_1 e^t)^{n_1}$$

Also, Y is a Binomial variate with parameters n_2 and p_2, its moment generating function is $M_Y(t) = (q_2 + p_2 e^t)^{n_2}$

Consider $M_{X+Y}(t) = M_X(t) \,.\, M_Y(t)$ | X,Y are independent

$$= (q_1 + p_1 e^t)^{n_1} (q_2 + p_2 e^t)^{n_2} \qquad \text{...(1)}$$

Now (1) cannot be expressed in the form $(q + pe^t)^n$. Therefore, by uniquencess Theorem of moment generating function, X + Y is not a Binomial variate.

Cor. If we take $p_1 = p_2 = p$, then from (1), $M_{X+Y}(t) = (q + pe^t)^{n_1 + n_2}$

Therefore, by uniquencess theorem of moment generating function, X + Y is a binomial variate with parameter $n_1 + n_2$ and p. Thus, the reproductive property holds when $p_1 = p_2$.

Generalisation. If $x_1, x_2, ..., x_n$ are independent binomial variates with parameters (n_i, p) $(i = 1, 2, ..., n)$, then their sum $x_1 + x_2 + ... + x_n$ is also a binomial variate with parameters $n_1 + n_2 + ... + n_n$ and p.

4.12 CHARACTERISTIC FUNCTION OF BINOMIAL DISTRIBUTION

Let $\phi_x(t)$ denotes the characteristic function of the Binomial distribution, then

$$\phi_x(t) = E(e^{tix}) = \sum_{x=0}^{n} e^{itx} p(x)$$

$$= \sum_{x=0}^{n} e^{itx} \, {}^nc_x p^x q^{n-x} = \sum_{x=0}^{n} {}^nC_x (pe^{it})^x q^{n-x} = (q + pe^{it})^x.$$

4.13 RECURRENCE FORMULA FOR BINOMIAL DISTRIBUTION

NOTES

Let x be a binomial variable and $P(x = r) = {}^nC_r p^r q^{n-r}$, $0 \le r \le n$.

For $0 \le k < n$, $\quad P(k) = {}^nC_k\, p^k\, q^{n-k}$

and $\quad P(k+1) = {}^nC_{k+1}\, p^{k+1}\, q^{n-(k+1)}$.

Dividing, we get
$$\frac{P(k+1)}{P(k)} = \frac{{}^nC_{k+1}\, p^{k+1}\, q^{n-k-1}}{{}^nC_k p^k q^{n-k}}$$
$$= \frac{n!}{(k+1)!(n-(k+1))!} \cdot \frac{k!(n-k)!}{n!} \cdot \frac{p}{q} = \frac{n-k}{k+1} \cdot \frac{p}{q}.$$

$$\therefore \quad \mathbf{P(k+1) = \frac{n-k}{k+1} \cdot \frac{p}{q}\ P(k) \quad for\ 0 \le k < n.}$$

This is the required **recurrence formula**.

ILLUSTRATIVE EXAMPLES

Example 1. *Find the expectation of the number of heads in 15 tosses of a coin.*

Sol. Here $n = 15$. Let p be the probability of getting a head in a trial, *i.e.*, in a loss.

$\therefore \quad p = \frac{1}{2}$.

Let x be the Binomial variable "*no. of heads*".

$\therefore \quad$ Expectation of $x = E(x) = \text{mean} = np = 15 \times \frac{1}{2} = 7.5$.

Example 2. *Obtain the binomial distribution whose mean is 10 and standard deviation is* $2\sqrt{2}$.

Sol. Let number of trials $= n$ and probability of success $= p$.

$\therefore \quad P(r \text{ successes}) = {}^nC_r\, p^r\, q^{n-r}$, $0 \le r \le n$.

We have mean $= np = 10$ and S.D. $= \sqrt{npq} = 2\sqrt{2}$.

$$\therefore \quad \sqrt{10q} = \sqrt{8} \;\Rightarrow\; q = \frac{8}{10} = \frac{4}{5}$$

$$\therefore \quad p = 1 - q = 1 - \frac{4}{5} = \frac{1}{5}$$

$$\therefore \quad np = 10 \;\Rightarrow\; n\left(\frac{1}{5}\right) = 10 \;\Rightarrow\; n = 50$$

$$\therefore \quad \mathbf{P(r\ successes) = {}^{50}C_r \left(\frac{1}{5}\right)^r \left(\frac{4}{5}\right)^{50-r}}, \ 0 \le r \le 50.$$

Example 3. *A discrete random variable x has mean score equal to '6' and variance equal to '2'. Assuming that the underlying distribution of x is binomial, what is the probability when* $5 \le x \le 6$.

NOTES

Sol. We have $\quad \text{mean} = np = 6 \quad$ ...(1)

and $\quad \text{variance} = npq = 2 \quad$ ...(2)

(1) and (2) $\Rightarrow \quad 6 \times q = 2 \Rightarrow q = \frac{2}{6} = \frac{1}{3}.$

$\therefore \quad p = 1 - q = 1 - \frac{1}{3} = \frac{2}{3}$

(1) $\Rightarrow \quad n\left(\frac{2}{3}\right) = 6 \Rightarrow n = 9.$

$\therefore \quad \text{P}(r \text{ successes}) = {}^n\text{C}_r\, p^r\, q^{n-r}, 0 \le r \le n$

$$= {}^9\text{C}_r \left(\frac{2}{3}\right)^r \left(\frac{1}{3}\right)^{9-r}, 0 \le r \le 9.$$

$\therefore \quad \text{P}(5 \le x \le 6) = \text{P}(x = 5 \text{ or } x = 6) = \text{P}(x = 5) + \text{P}(x = 6)$

$$= {}^9\text{C}_5 \left(\frac{2}{3}\right)^5 \left(\frac{1}{3}\right)^4 + {}^9\text{C}_6 \left(\frac{2}{3}\right)^6 \left(\frac{1}{3}\right)^3$$

$$= \frac{1}{3^9}\,[126 \times 32 + 84 \times 64] = \frac{9408}{3^9}.$$

Example 4. *(a) The sum of mean and variance of a binomial variance is 15 and the sum of their squares is 117. Find the distribution.*

(*b*) *The sum and the product of the mean and variance of a binomial distribution are 24 and 128 respectively. Find the distribution.*

(*c*) *If the probability of a defective bulb is 0.1, find the mean and the standard deviation of defective bulbs in a total of 900.*

Sol. (*a*) Let the binomial distribution be $\text{P}(x = r) = {}^n\text{C}_r\, p^r\, q^{n-r}, 0 \le r \le n$.

We have mean $= np$ and variance $= npq$.

$\therefore \quad np + npq = 15 \quad$ ...(1) $\quad$ and $\quad (np)^2 + (npq)^2 = 117 \quad$ ...(2)

(1) $\Rightarrow \quad np(1 + q) = 15 \quad$ ...(3) $\quad$ (2) $\Rightarrow \quad n^2p^2(1 + q^2) = 117 \quad$ ...(4)

$$\Rightarrow \quad \frac{n^2p^2(1+q^2)}{(np(1+q))^2} = \frac{117}{(15)^2} \Rightarrow \frac{1+q^2}{(1+q)^2} = \frac{117}{225} \quad \Rightarrow \quad \frac{1+q^2}{1+q^2+2q} = \frac{13}{25}$$

$$\Rightarrow \quad 25 + 25q^2 = 13 + 13q^2 + 26q \Rightarrow 6q^2 - 13q + 6 = 0 \Rightarrow q = \frac{2}{3}, \frac{3}{2}.$$

$\therefore \quad q = \frac{2}{3}$ because $q = \frac{3}{2} > 1$ is impossible.

$\therefore \quad p = 1 - q = 1 - \frac{2}{3} = \frac{1}{3}.$

Putting $p = \frac{1}{3}$, $q = \frac{2}{3}$ in (3), we get, $n \,.\, \frac{1}{3}\left(1 + \frac{1}{3}\right) = 15.$

$\Rightarrow \quad n \,.\, \frac{5}{9} = 15 \Rightarrow n = 27.$

$\therefore \quad P(x = r) = {}^{27}C_r \left(\frac{1}{3}\right)^r \left(\frac{2}{3}\right)^{27-r}, \; 0 \le r \le 27.$

(*b*) Let the Binomial distribution be

$$P(X = r) = {}^nC_r\, p^r q^{n-r},\; 0 \le r \le n$$

NOTES

Here mean = np and variance = npq

$\therefore$ Given, $\quad np + npq = 24 \qquad ...(1)$

and $\quad np \,.\, npq = 128 \qquad ...(2)$

Squaring (1) and dividing it by (2), we have

$$\frac{(np(1+q))^2}{n^2p^2.q} = \frac{576}{128} \quad \text{or} \quad \frac{n^2p^2(1+q)^2}{n^2p^2.q} = \frac{9}{2}$$

$$\Rightarrow \quad 2(1+q)^2 = 9q$$

$$\Rightarrow \quad 2q^2 + 2 + 4q = 9q$$

$$\Rightarrow \quad 2q^2 - 5q + 2 = 0$$

$$\Rightarrow \quad 2q^2 - 4q - q + 2 = 0$$

$$\Rightarrow \quad 2q(q-2) - (q-2) = 0$$

$$\Rightarrow \quad (q-2)(2q-1) = 0$$

$$\Rightarrow \quad q = 2, \frac{1}{2}$$

But $q \ne 2 \quad \therefore$ We take $q = \frac{1}{2} \quad \therefore \quad p = 1 - \frac{1}{2} = \frac{1}{2}$

Also from (1), $np(1+q) = 24$

$$\Rightarrow \quad n \,.\, \frac{1}{2}.\frac{3}{2} = 24$$

$$\Rightarrow \quad n = 32$$

Hence the required distribution is

$$P(X = r) = {}^{32}C_r \left(\frac{1}{2}\right)^r \left(\frac{1}{2}\right)^{32-r}, \; 0 \le r \le 32.$$

(*c*) Given p = the probability of a bulb being defective

$$= 0.1$$

$$q = 1 - p = 1 - 0.1 = 0.9,\; n = 900$$

Using Binomial distribution,

$$\text{mean} = np = 900 \times 0.1 = 90$$

$$\text{variance} = npq$$

$$= 90 \times 0.9 = 81$$

$\therefore \quad$ Standard deviation $= \sqrt{npq} = \sqrt{81}$

$$= 9.$$

Example 5. *If X is a binomial variate with parameters n and p, then show that*

(i) $E(X/n - p)^2 = \frac{pq}{n}$ $\qquad$ *(ii)* $Cov\left(\frac{X}{n}, \frac{n-X}{n}\right) = -\frac{pq}{n}.$

NOTES

Sol. Given X ~ B (n, p), therefore E(X) = np, Var(X) = npq

Also $E(X|n) = \frac{1}{n} E(X) = \frac{1}{n} \cdot np = p$

$$\text{Var}(X|n) = \frac{1}{n^2} \text{Var}(X) = \frac{1}{n^2} \cdot npq = \frac{pq}{n}$$

(*i*) $E(X|n - p)^2 = E(X/n - E(X/n))^2 = \text{Var}(X/n) = \frac{pq}{n}$

Usin Cov (X, Y) = E [(X – E(X)) (Y – E(Y))], we have

$$(ii)\ \text{Cov}\left(\frac{X}{n}, \frac{n-X}{n}\right) = E\left[\left(\frac{X}{n} - E\left(\frac{X}{n}\right)\right)\left(\frac{n-X}{n} - E\left(\frac{n-X}{n}\right)\right)\right]$$

$$= E\left[\left(\frac{X}{n} - p\right)\left(1 - \frac{X}{n} - E\left(1 - \frac{X}{n}\right)\right)\right]$$

$$= E\left[\left(\frac{X}{n} - p\right)\left(1 - \frac{X}{n} - 1 + p\right)\right]$$

$$= E\left[\left(\frac{X}{n} - p\right)\left(p - \frac{X}{n}\right)\right]$$

$$= -E\left(\frac{X}{n} - p\right)^2 = -E\left(\frac{X}{n} - E\left(\frac{X}{n}\right)\right)^2$$

$$= -\text{Var}\left(\frac{X}{n}\right) = \frac{-pq}{n}.$$

Example 6. (*i*) *Determine the Binomial distribution for which the mean is 4 and variance is 3. Also find its mode.*

(*ii*) *Show that for p = 0.50, the Binomial distribution has a maximum probability at* $X = \frac{1}{2}n$ *if n is even and at* $X = \frac{1}{2}(n-1), \frac{1}{2}(n+1)$ *if n is odd.*

Sol. (*i*) Given mean = $np = 4$, variance $npq = 3$

$$\Rightarrow \qquad q = \frac{3}{np} = \frac{3}{4}.$$

Also $$p = 1 - q = 1 - \frac{3}{4} = \frac{1}{4}$$

$$\therefore \qquad np = 4 \text{ gives } n = 16$$

The required distribution is given by

$$(q + p)^n = \left(\frac{3}{4} + \frac{1}{4}\right)^{16}$$

To find mode, consider $(n + 1)\,p = 17 \cdot \frac{1}{4} = 4.25$, which is not an integer. Hence the mode is unique and given by integeral part of $(n + 1)\,p = 4$.

NOTES

(*ii*) Here $p = \frac{1}{2}$ and consider $(n + 1)\,p$.

Case I. If n is even, say, $n = 2m$, then $(n + 1)\,p = (2m + 1)\,\frac{1}{2} = m + \frac{1}{2}$, not an integer. Therefore, the mode is unique and value of mode = integeral part of $m + \frac{1}{2}$

$$= m = \frac{n}{2}.$$

Case II. If n is odd, say, $n = 2m + 1$

Consider $(n + 1)\,p = (2m + 1 + 1)\,\frac{1}{2} = m + 1$, which is an integer. Therefore, the distribution is bimodal and the values of mode are given by

$$m + 1, m \quad \text{or} \quad \frac{n-1}{2} + 1, \frac{n-1}{2} \quad \text{or} \quad \frac{n+1}{2}, \frac{n-1}{2}.$$

Example 7. *If X is a binomial distribution with parameters n and p, what is the distribution of Y = n – X.*

Sol. We know that the moment generating function of the binomial variate is

$$M_X(t) = E(e^{tX}) = (q + pe^t)^n$$

$$\therefore \quad M_Y(t) = E(e^{tY}) = E\,(e^{t(n-X)})$$

$$= E\,(e^{nt} \cdot e^{-tX}) = e^{nt}\,E(e^{-tX})$$

$$= e^{nt} M_X(-t)$$

$$= e^{nt}\,(q + pe^{-t})^n$$

$$= (e^t\,(q + pe^{-t}))^n$$

$$= (qe^t + q)^n$$

By uniqueness theorem of moment generating function, $Y \sim B\,(n, q)$ *i.e.*, $n - X$ is also a Binomial distribution with parameters n and q.

Example 8. *If the independent random variables X and Y are binomially distributed with $n = 3$, $p = \frac{1}{3}$ and $n = 5$, $p = \frac{1}{3}$. Find $P\,(X + Y \geq 1)$.*

Sol. Given $\quad X \sim B\left(3, \frac{1}{3}\right), Y \sim B\left(5, \frac{1}{3}\right)$

$$\Rightarrow \quad X + Y \sim B\left(3+5, \frac{1}{3}\right) \text{ or } X + Y \sim B\left(8, \frac{1}{3}\right) \qquad | \text{ Additive property}$$

$$\therefore \quad P(X+Y=r) = {}^8C_r \left(\frac{2}{3}\right)^{8-r}\left(\frac{1}{3}\right)^r$$

NOTES

Hence $\quad P(X+Y \geq 1) = 1 - P(X+Y<1)$

$$= 1 - P(X+Y=0)$$

$$= 1 - {}^8C_0 \left(\frac{2}{3}\right)^8 \left(\frac{1}{3}\right)^0 = 1 - \left(\frac{2}{3}\right)^8.$$

Multinomial Distribution: This distribution is the generalisation of the binomial distribution. When there are more than two mutually exclusive outcomes of a trial, the observations lead to multinomial distribution. Let $E_1, E_2, \ldots\ldots, E_k$ are k mutually exclusive and exhaustive outcomes of a trial with respective probabilities $p_1, p_2, \ldots\ldots, p_k$.

Let the probability that

E_1 occurs x_1 times ;

E_2 occurs x_2 times ;

⋮

E_k occurs x_k times ;

in n independent observations, is given by

$$p(x_1, x_2, \ldots\ldots, x_k) = c\ p_1^{x_1} p_2^{x_2} \ldots\ldots p_k^{x_k}$$

where $\sum_{i=1}^{n} x_i = n$ and c is the number of permutations of the events $E_1, E_2, \ldots\ldots, E_k$

To determine c, we have to find the number of permutations of n objects of which x_1 are of one kind, x_2 of another kind, ..., x_k of the kth kind, which is given by:

$$c = \frac{n!}{x_1!\,x_2!\ldots\ldots x_k!}$$

$$\text{Hence } p(x_1, x_2, \ldots\ldots, x_k) = \frac{n!}{x_1!\,x_2!\ldots\ldots x_k!} \cdot p_1^{x_1} p_2^{x_2} \ldots\ldots p_k^{x_k}, \; 0 \leq x_i \leq n$$

$$= \frac{n!}{\prod_{i=1}^{k} x_i!} \prod_{i=1}^{k} p_i^{x_i} \quad \sum_{i=1}^{k} x_i = n$$

which is the required probability function of the multinomial distribution. It is so called since (8.30) is the general term in the multinomial expansion:

$$(p_1 + p_2 + \ldots\ldots + p_k)^n, \sum_{i=1}^{k} p_i = 1$$

Since, total probability is 1, we have

$$\sum_x p(x) = \sum_x \left[\frac{n!}{x_1!\,x_2!\ldots\ldots x_k!}\,p_1^{x_1}\,p_2^{x_2}\ldots\ldots p_k^{x_k}\right] = (p_1 + p_2 + \ldots\ldots + p_k)^n = 1$$

NOTES

Theorem IV. *Find the moment generating function of the multinomial distribution.*

Proof. Let X ~ Multinomial distribution and $M_X(t)$ denotes the *m.g.f.*, then,

The moment generating function is given by:

$$M_X(t) = M_{X_1, X_2, \ldots\ldots, X_k}(t_1, t_2, \ldots\ldots, t_k) = E\left[\exp\left\{\sum_{i=1}^{k} t_i X_i\right\}\right]$$

$$= \sum_x \left[\frac{n!}{x_1!\,x_2!\ldots\ldots x_k!}\,p_1^{x_1} p_2^{x_2}\ldots\ldots p_k^{x_k} \exp\left(\sum_{i=1}^{k} t_i x_i\right)\right]$$

$$= \sum_x \left[\frac{n!}{x_1!\,x_2!\ldots\ldots x_k!}\,(p_1 e^{t_1})^{x_1}\,(p_2 e^{t_2})^{x_2}\ldots\ldots (p_k e^{t_k})^{x_k}\right]$$

$$= (p_1 e^{t_1} + p_2 e^{t_2} + \ldots\ldots + p_k e^{t_k})^n$$

where $\quad X = (X_1, X_2, \ldots\ldots X_k)$.

4.14 COMPOUND BINOMIAL DISTRIBUTION

Compound distribution. Let X is a random variable such that its distribution depends on a single parameter θ, where θ, instead of being regarded as a fixed constant is also a random variable following a particular distribution. Then, we say, X follows a **compound distribution.**

Theorem V. *Find the mean and variance of the compound Binomial distribution.*

Proof. Let $X_1, X_2, \ldots\ldots, X_n$ are independent and indentically distributed Bernouli variates with $P(X_i = 1) = p$ and $P(X_i = \theta) = q = 1 - p$ then,

$X = X_1 + X_2 + \ldots\ldots + X_n \sim B(n, p)$ *i.e.,* X is a Binomial variate with parameters n and p and hence its probability density function is

$$P(X = r) = {}^nC_r\, q^{n-r}\, p^r,\ r = 0, 1, 2, \ldots\ldots, n \qquad \ldots(1)$$

where $P(X = r)$ is the probability of r successes in n independent trials with constant probability 'p' of success for each trial.

Now suppose that n, instead of being regarded as a fixed constant, is also a random variable following Poisson law with parameter λ. Then

$$P(n = k) = \frac{e^{-\lambda}\lambda^k}{k!};\ k = 0, 1, 2, \ldots\ldots \qquad \ldots(2)$$

NOTES

In such a case X is said to have compound binomial distribution. The joint probability function of X and n is given by:

$$P(X = r \cap n = k) = P(n = k)\, P(X = r \mid n = k) = \frac{e^{-\lambda}\lambda^k}{k!}\, {}^kC_r\, p^r q^{k-r},$$

| using (1) and (2)

Since $P(X = r \mid n = k)$ is probability of r successes in k trials. Obviously, $r \le k$

$\Rightarrow \quad k \ge r.$

The marginal distribution of X is given by :

$$P(X = r) = \sum_{k=r}^{\infty} P(X = r \cap n = k)$$

$$= e^{-\lambda} p^r \sum_{k=r}^{\infty} {}^kC_r \frac{\lambda^k q^{k-r}}{k!} = \frac{e^{-\lambda}(\lambda p)^r}{r!} \sum_{k=r}^{\infty} \frac{(\lambda q)^{k-r}}{(k-r)!}$$

$$= \frac{e^{-\lambda}(\lambda p)^r}{r!} \sum_{j=0}^{\infty} \frac{(\lambda q)^j}{j!} \qquad (j = k - r)$$

$$= \frac{e^{-\lambda}(\lambda p)^r}{r!} . e^{\lambda q} = \frac{e^{-\lambda p}(\lambda p)^r}{r!}$$

which is the probability function of a Poisson variate with parameter λp.

Hence $E(X) = \lambda p$ and $Var(X) = \lambda p$.

4.15 COMPOUND POISSON DISTRIBUTION

Theorem VI. *Discuss compound Poisson distribution or if X is a Poisson variate with parameter λ, then show that the compound Poisson destribution of X is a negative Binomial distribution with parameters (q, p).*

Proof. Given X is a Poisson variable with parameter λ and hence its probability density function is

$$P(X = r) = \frac{e^{-\lambda}\lambda^r}{r!}, \; r = 0, 1, 2, ...$$

Let us suppose that, λ, instead of being a fixed constant, is itself a continuous random variable with generalised gamma density, given by

$$g(\lambda) = \begin{cases} \dfrac{a\, e^{-a\lambda}\lambda^{v-1}}{\Gamma(V)}, & \lambda > 0, a > 0, V > 0 \\ 0, & \lambda \le 0 \end{cases}$$

Now, consider the two dimensional **random vector** (X, λ) in which one variable is discrete and the other is continuous. For a constant $h > 0$ and $\lambda_1 > 0$, the joint density of X and λ is given by :

$$P(X = r \cap \lambda_1 \le \lambda \le \lambda_1 + h) = P(\lambda_1 \le \lambda \le \lambda_1 + h)\, P(X = r \mid \lambda_1 \le \lambda \le \lambda_1 + h)$$

Dividing both sides by h and proceeding the limits as $h \to 0$, we get

$$\lim_{h \to 0} \frac{P(X = r \cap \lambda_1 \le \lambda \le \lambda_1 + h)}{h}$$

NOTES

$$= \lim_{h\to 0} P(X = r \mid \lambda_1 \le \lambda \le \lambda_1 + h) \times \lim_{h\to 0} \frac{P(\lambda_1 \le \lambda \le \lambda_1 + h)}{h}$$

Since $\lim_{h\to 0} \frac{P(\lambda_1 \le \lambda \le \lambda_1 + h)}{h} = \lim_{h\to 0} \frac{G(\lambda_1 + h) - G(\lambda_1)}{h} = G'(\lambda_1) = g(\lambda_1)$,

where $G(\cdot)$ is the distribution function and $g(\cdot)$ is *p.d.f.* of λ.

$$\therefore \quad \lim_{h\to 0} \frac{P(X = r \cap \lambda_1 \le \lambda \le \lambda_1 + h)}{h} = \frac{e^{-\lambda_1}\lambda_1^r}{r!} \cdot \frac{a^v}{\Gamma(v)} \lambda_1^{v-1} e^{-a\lambda_1}$$

Integrating *w.r.t.* to λ_1 over 0 to ∞ and using Gamma integral, the marginal probability function of X is given by :

$$P(X = r) = \frac{a^v}{\Gamma(v)\, r!} \int_0^\infty e^{-(1+a)\lambda} \lambda^{r+v-1}\, d\lambda = \frac{a^v}{\Gamma(v)\, r!} \cdot \frac{\Gamma(r+v)}{(1+a)^{r+v}}$$

$$= \left(\frac{a}{1+a}\right)^v \frac{v(v+1)(v+2)\ldots\ldots(v+r-1)}{(1+a)^r\, r!}$$

$$= \left(\frac{a}{1+a}\right)^v (-1)^r \binom{-v}{r} \left(\frac{1}{1+a}\right)^r = \binom{-v}{r} p^v (-q)^r;\ r = 0, 1, 2, \ldots\ldots$$

where $\quad p = a/(1+a),\ q = 1 - p = 1/(1+a)$.

Hence the marginal distribution of X is negative binomial with parameters (v, p).

Hence the theorem.

II. POISSON DISTRIBUTION

4.16 INTRODUCTION

The Poisson distribution is also a discrete probability distribution. This was discovered by French mathematician **Simon Denis Poisson** (1781–1840) in the year 1837. This distribution deals with the evaluation of probabilities of *rare* events such as "number of car accidents on road", "number of earthquakes in a year", "number of misprints in a book" etc.

4.17 CONDITIONS FOR APPLICABILITY OF POISSON DIS TRIBUTION

The Poisson distribution is derived as a limiting case of the binomial distribution. So, the conditions for the applicability of the Poisson distribution are same as those for the applicability of Binomial distribution. Here the additional requirement is that the probability of 'success' is quite near to zero.

NOTES

4.18 POISSON VARIABLE

A random variable which counts the number of successes in a random experiment with trials satisfying above conditions is called a **Poisson variable.** If the probability of an article being defective is 1/500 and the event of getting a defective article is *success* and samples of 10 articles are checked for defective articles, then the possible values of Poisson variable are 0, 1, 2, ..., 10.

4.19 POISSON PROBABILITY FUNCTION

Let a random experiment satisfying the conditions for Poisson Distribution be performed. Let the number of trials in the experiment be n, which is indefinitely large. Let p denotes the probability of *success* in any trial. We assume that p is indefinitely small, *i.e.*, we are dealing with a rare event. Let x denotes the Poisson variable corresponding to this random experiment.

$\therefore$ The possible values of x are 0, 1, 2, ..., n.

The Poisson distribution is obtained as a limiting case of the corresponding binomial distribution of the experiment under the conditions :

(*i*) n, the number of trials is indefinitely large, *i.e.*, $n \to \infty$.

(*ii*) p, the probability of success in a trial is indefinitely small, *i.e.*, $p \to 0$.

(*iii*) The product np of n and p is constant.

By **Binomial distribution**, $P(x = r) = {}^nC_r\, p^r q^{n-r},\ 0 \le r \le n$, where $q = 1 - p$.

Let $np = m$. $\therefore\ p = \dfrac{m}{n}$ and $q = 1 - p = 1 - \dfrac{m}{n}$.

$$\therefore \quad P(x = r) = \frac{n!}{r!\,(n-r)!}\left(\frac{m}{n}\right)^r\left(1-\frac{m}{n}\right)^{n-r}$$

$$= \frac{n(n-1)(n-2)\ \ldots\ldots\ (n-(r-1))(n-r)!}{r!\,(n-r)!}\cdot\frac{m^r}{n^r}\left(1-\frac{m}{n}\right)^{n-r}$$

$$= \frac{m^r}{r!}\cdot\frac{n}{n}\cdot\frac{n-1}{n}\cdot\frac{n-2}{n}\ \ldots\ldots\ \frac{n-(r-1)}{n}\cdot\left(1-\frac{m}{n}\right)^n\left(1-\frac{m}{n}\right)^{-r}$$

$$= \frac{m^r}{r!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\ \ldots\ldots\ \left(1-\frac{r-1}{n}\right)\left(1-\frac{m}{n}\right)^n\left(1-\frac{m}{n}\right)^{-r}$$

$$\therefore \quad \lim_{n\to\infty} P(x = r) = \frac{m^r}{r!}\lim_{n\to\infty}\left(1-\frac{1}{n}\right)\lim_{n\to\infty}\left(1-\frac{2}{n}\right)\ldots\ldots\lim_{n\to\infty}\left(1-\frac{r-1}{n}\right)$$

$$\times \lim_{n\to\infty}\left(1-\frac{m}{n}\right)^n\lim_{n\to\infty}\left(1-\frac{m}{n}\right)^{-r}$$

$$= \frac{m^r}{r!}\,(1-0)(1-0)\ \ldots..\ (1-0)e^{-m}(1-0)^{-r}$$

$$\left[\because \lim_{n\to\infty}\left(1-\frac{m}{n}\right)^n = e^{-m}\right]$$

$$= \frac{m^r}{r!}\cdot e^{-m}\cdot 1 = \frac{e^{-m}\,m^r}{r!}.$$

$\therefore$ When n is indefinitely large, we have $\mathbf{P(x = r) = \dfrac{e^{-m}\, m^r}{r\,!}}$, $\mathbf{r = 0, 1, 2, ...}$

This is called the **Poisson probability function.** The corresponding **Poisson distribution** is

NOTES

x	0	1	2	3 ......
$P(x)$	$\dfrac{e^{-m}m^0}{0!}$	$\dfrac{e^{-m}m^1}{1!}$	$\dfrac{e^{-m}m^2}{2!}$	$\dfrac{e^{-m}m^3}{3!}$ ....

The constant m is the product of n and p and is called the **parameter** of the Poisson distribution.

4.20 POISSON FREQUENCY DISTRIBUTION

If a random experiment, satisfying the requirements of Poisson distribution, is repeated N times, then the expected frequency of getting $r(0 \le r \le n)$ successes is given by

$$\mathbf{N \cdot P(x = r) = N\, \frac{e^{-m} r m^r}{r\,!},\ r = 0, 1, 2, ...}$$

WORKING RULES FOR SOLVING PROBLEMS

I. Make sure that the trials in the random experiment are independent and the success is a rare event and each trial result in either success or failure.

II. Define the Poisson variable and find the value of n and p from the given data. Find $m = np$. Sometimes, the value of m is directly given.

III. Put the value of m in the formula :

$$P(r \text{ successes}) = \frac{e^{-m}\, m^r}{r\,!},\ r = 0, 1, 2, ..., n. \qquad ...(1)$$

IV. Express the event, whose probability is desired in terms of values of the Poisson variable x. Use (1) to find the required probability.

Remark 1. The distribution to be used in solving a problem is generally given is the problem. If it is not given, then the student should make use of Poisson distribution only when the event in the problem is of rare nature, *i.e.*, the probability of happening of event is quite near to zero.

Remark 2. The value of e^{-m} required in any particular problem is generally given with the problem itself. Otherwise, the value of e^{-m} can be found out by using the table given in this chapter. In the examination hall, generally the table of e^{-m} is available for students. If at all the value of e^{-m} is neither given with the problem nor the table of e^{-m} is supplied in the examination hall, then the students are advised to retain their final result in terms of e^{-m}.

ILLUSTRATIVE EXAMPLES

Example 1. *Out of 100 bulbs sample, the probability of a bulb to be defective is 0.03. Using Poisson distribution, obtain the Probability that in a sample of 100 bulbs, none is defective.*

[Given $e^{-3} = 0.04979$]

Sol. Let x be the Poisson variable, *"no. of defective bulbs in a sample of 100 bulbs"*.

By Poisson distribution, $P(x = r) = \dfrac{e^{-m}\, m^r}{r\,!}$, $r = 0, 1, 2, \ldots\ldots$

Here $n = 100, p = 0.03$. $\therefore$ $m = np = 100 \times 0.03 = 3$.

NOTES

$$\therefore \quad P(x = r) = \frac{e^{-3}\,(3)^r}{r\,!}, \; r = 0, 1, 2, \ldots\ldots, 100.$$

$$\therefore \; P(\text{none is defective}) = P(x = 0) = \frac{e^{-3}\,(3)^0}{0\,!} = \frac{0.04979 \times 1}{1} = 0.04979.$$

Example 2. *There are 50 telephone lines in an exchange. The probability that any one of them will be busy is 0.1. What is the probability that all the lines are busy ?*

Sol. Let x be the Poisson variable, *"no of busy lines in the exchange"*.

By Poisson distribution, $P(x = r) = \dfrac{e^{-m}\, m^r}{r\,!}$, $r = 0, 1, 2, \ldots\ldots$

Here $n = 50, p = 0.1$. $\therefore$ $m = np = 50 \times 0.1 = 5$.

$$\therefore \quad P(x = r) = \frac{e^{-5}\,(5)^r}{r\,!}, \; r = 0, 1, 2, \ldots\ldots, 50.$$

$$\therefore \quad P(\text{all lines are busy}) = P(x = 50) = \frac{e^{-5}\,(5)^{50}}{50\,!}.$$

Example 3. *Eight per cent of the bolts produced in a certain factory turns out to be defective. Find the probability, using Poisson distribution, that in a sample of 25 bolts chosen at random, (i) exactly 3 (ii) more than 3, will be defective.*
[Take $e^{-2} = 0.135$]

Sol. Let x be the Poisson variable, *"no. of defective bolts in a sample of 25 bolts"*.

By Poisson distribution, $P(x = r) = \dfrac{e^{-m}\, m^r}{r\,!}$, $r = 0, 1, 2, \ldots\ldots$

Here $n = 25$, $p = \dfrac{8}{100} = \dfrac{2}{25}$. $\therefore$ $m = np = 25 \times \dfrac{2}{25} = 2$.

$$\therefore \quad P(x = r) = \frac{e^{-2}\,(2)^r}{r\,!}, \; r = 0, 1, 2, \ldots\ldots, 25.$$

(*i*) P(exactly 3 defectives)

$$= \frac{e^{-2}\,(2)^3}{3!} = \frac{0.135 \times 8}{6} = 0.18. \qquad (\text{Using } e^{-2} = 0.135)$$

(*ii*) P(more than 3 defectives) $= P(x > 3) = 1 - P(x \le 3)$

$$= 1 - P(x = 0 \;\text{ or }\; x = 1 \;\text{ or }\; x = 2 \;\text{ or }\; x = 3)$$

$$= 1 - [P(x = 0) + P(x = 1) + P(x = 2) + P(x = 3)]$$

$$= 1 - \left[\frac{e^{-2}(2)^0}{0!} + \frac{e^{-2}(2)^1}{1!} + \frac{e^{-2}(2)^2}{2!} + \frac{e^{-2}(2)^3}{3!}\right]$$

$$= 1 - e^{-2}\left[1 + 2 + \frac{4}{2} + \frac{8}{6}\right] = 1 - 0.135\left(1 + 2 + 2 + \frac{4}{3}\right)$$

$$= 1 - 0.135 \times \frac{19}{3} = 0.145.$$

NOTES

Example 4. *A box contains 200 tickets each bearing one of the numbers from 1 to 200. 20 tickets are drawn successively with replacement from the box. Find the probability that at most 4 tickets bear numbers divisible by 20.*

Sol. Let x be the Poisson variable, "*no. of tickets bearing number divisible by 20*".

By **Poisson distribution,** $P(x = r) = \dfrac{e^{-m}\, m^r}{r!}$, $r = 0, 1, 2, ...$

The numbers from 1 to 200 and divisible by 20 are 20, 40, 60, 80, 100, 120, 140, 160, 180, 200.

Let p be the probability of getting a ticket with number divisible by 20.

$$\therefore \quad p = \frac{10}{200} = \frac{1}{20}.$$

Also n = number of trials = 20.

$$\therefore \quad m = np = 20 \times \frac{1}{20} = 1$$

$$\therefore \quad P(x = r) = \frac{e^{-1}\,(1)^r}{r!} = \frac{e^{-1}}{r!},\ r = 0, 1, 2, ..., 20$$

$\therefore$ P(at most 4 tickets bear number divisible by 20) $= P(x \le 4)$

$$= P(x = 0) + P(x = 1) + P(x = 2) + P(x = 3) + P(x = 4)$$

$$= \frac{e^{-1}}{0!} + \frac{e^{-1}}{1!} + \frac{e^{-1}}{2!} + \frac{e^{-1}}{3!} + \frac{e^{-1}}{4!} = \frac{1}{e}\left[1 + 1 + \frac{1}{2} + \frac{1}{6} + \frac{1}{24}\right] = \frac{65}{24e}.$$

PROPERTIES OF POISSON DISTRIBUTION

4.21 SHAPE OF POISSON DISTRIBUTION

The shape of the Poisson distribution depends upon the parameter m, the average number of successes per unit. As value of m increases, the graph of Poisson distribution would get closer to a symmetrical continuous curve.

4.22 SPECIAL USEFULNESS OF POISSON DISTRIBUTION

The Poisson distribution is specially used when there are events which do not occur as outcomes of a definite number of trials in an experiment, rather occur randomly in nature. This distribution is used when the event under consideration is rare and casual. In finding probabilities by *Poisson distribution*, we require only the measure of average chance of occurrence (m) based on past experience or a small sample drawn for the purpose.

4.23 MEAN OF POISSON DISTRIBUTION

NOTES

Let x be a Poisson variable and $P(x = r) = \dfrac{e^{-m}m^r}{r!}$, $r = 0, 1, 2, ...$

The *mean of x* is the average numbers of successes.

$$\therefore \quad \text{Mean,} \quad \mu = \sum_{r=0} r.P(x=r) = \sum_{r=0} r.\frac{e^{-m}m^r}{r!}$$

$$= 0.\frac{e^{-m}m^0}{0!} + 1.\frac{e^{-m}m^1}{1!} + 2.\frac{e^{-m}m^2}{2!} + 3.\frac{e^{-m}m^3}{3!} + ...$$

$$= 0 + me^{-m}\left(\frac{1}{1!} + \frac{2m}{2!} + \frac{3m^2}{3!} + ...\right) = me^{-m}\left(1 + \frac{m}{1!} + \frac{m^2}{2!} + ...\right)$$

$$= me^{-m} \,.\, e^m = me^0 = m \,.\, 1 = m.$$

$\therefore$ **Mean (μ) of x = m.**

4.24 VARIANCE AND S.D. OF POISSON DISTRIBUTION

Let x be a Poisson variable and $P(x = r) = \dfrac{e^{-m}m^r}{r!}$, $r = 0, 1, 2, ...$

The variance and standard deviation of x measures the dispersion of the Poisson distribution and are given by

$$\text{Variance} = \sum_{r=0} r^2.P(x=r) - \mu^2 \quad \text{and} \quad \text{S.D.} = \sqrt{\sum_{r=0} r^2.P(x=r) - \mu^2}.$$

$$\therefore \sum_{r=0} r^2.P(x=r) = \sum_{r=0} r^2.\frac{e^{-m}m^r}{r!}$$

$$= 0^2.\frac{e^{-m}m^0}{0!} + 1^2\frac{e^{-m}m^1}{1!} + 2^2\frac{e^{-m}m^2}{2!} + 3^2\frac{e^{-m}m^3}{3!} + 4^2\frac{e^{-m}m^4}{4!} + ...$$

$$= 0 + me^{-m}\left(\frac{1}{1!} + \frac{2m}{1!} + \frac{3m^2}{2!} + \frac{4m^3}{3!} + ...\right)$$

$$= me^{-m}\left\{\left(1 + \frac{m}{1!} + \frac{m^2}{2!} + \frac{m^3}{3!} + ...\right) + \left(\frac{m}{1!} + \frac{2m^2}{2!} + \frac{3m^3}{3!} + ...\right)\right\}$$

$$= me^{-m}\left\{e^m + m\left(1 + \frac{m}{1!} + \frac{m^2}{2!} + ...\right)\right\}$$

$$= me^{-m}\{e^m + me^m\} = me^{-m}e^m(1+m) = me^0(1+m)$$

$$= m(1+m) = m + m^2.$$

$\therefore$ Variance $= \sum_{r=0} r^2.P(x = r) - \mu^2 = (m + m^2) - m^2 = \mathbf{m}.$

Also, **S.D.** $= \sqrt{\textbf{Variance}} = \sqrt{\mathbf{m}}$ **.**

NOTES

Theorem VII. *Show that the first four moments about the origin for the Poisson distribution with parameter λ, are given as*

$$\mu'_1 = \lambda, \mu'_2 = \lambda^2 + \lambda, \mu'_3 = \lambda^3 + 3\lambda^2 + \lambda$$
$$\mu'_4 = \lambda^4 + 6\lambda^3 + 7\lambda^2 + \lambda.$$

From the above result, derive the first four central moments and hence find β_1, β_2 and γ_1, γ_2 respectively.

Proof. Let X follows Poisson distribution with parameter λ, then

$$p(x) = \frac{e^{-\lambda}\lambda^x}{x!}, x = 0, 1, 2, \ldots\ldots$$

The moments about the origin are as under

$$\mu'_1 = E(X) = \sum_{x=0}^{\infty} xp(x, \lambda) = \sum_{x=0}^{\infty} x \cdot \frac{e^{-\lambda}\lambda^x}{x!} = \lambda e^{-\lambda}\left\{\sum_{x=0}^{\infty} \frac{\lambda^{x-1}}{(x-1)!}\right\}$$

$$= \lambda e^{-\lambda}\left(1 + \lambda + \frac{\lambda^2}{2!} + \frac{\lambda^3}{3!} + \ldots\ldots\right) = \lambda e^{-\lambda} . e^{\lambda}$$

$= \lambda =$ Mean of the Poisson distribution

$$\mu'_2 = E(X^2) = \sum_{x=0}^{\infty} x^2 p(x, \lambda) = \sum_{x=0}^{\infty} \{x(x-1) + x\} \frac{e^{-\lambda}\lambda^x}{x!}$$

$$= e^{-\lambda}\sum_{x=0}^{\infty} x(x-1)\frac{\lambda^x}{x!} + \sum_{x=0}^{\infty} x\frac{e^{-\lambda}\lambda^x}{x!} = \lambda^2 e^{-\lambda}\left[\sum_{x=0}^{\infty} \frac{\lambda^{x-2}}{(x-2)!}\right] + \lambda$$

$$= \lambda^2 e^{-\lambda} e^{\lambda} + \lambda = \lambda^2 + \lambda$$

$$\mu'_3 = E(X^3) = \sum_{x=0}^{\infty} x^3 p(x, \lambda)$$

$$= \sum_{x=0}^{\infty} \{x(x-1)(x-2) + 3x(x-1) + x\} \frac{e^{-\lambda}\lambda^x}{x!}$$

$$= \sum_{x=0}^{\infty} x(x-1)(x-2)\frac{e^{-\lambda}\lambda^x}{x!} + 3\sum_{x=0}^{\infty} x(x-1)\frac{e^{-\lambda}\lambda^x}{x!} + \sum_{x=0}^{\infty} x\frac{e^{-\lambda}\lambda^x}{x!}$$

$$= e^{-\lambda}\lambda^3\left\{\sum_{x=3}^{\infty} \frac{\lambda^{x-3}}{(x-3)!}\right\} + 3e^{-\lambda}\lambda^2\left\{\sum_{x=2}^{\infty} \frac{\lambda^{x-2}}{(x-2)!}\right\} + \lambda$$

$$= e^{-\lambda}\lambda^3 e^{\lambda} + 3e^{-\lambda}\lambda^2 e^{\lambda} + \lambda = \lambda^3 + 3\lambda^2 + \lambda$$

$$\mu_4' = E(X^4) = \sum_{x=0}^{\infty} x^4 \cdot p(x, \lambda)$$

NOTES

$$= \sum_{x=0}^{\infty} \{x(x-1)(x-2)(x-3) + 6x(x-1)(x-2) + 7x(x-1) + x\} \frac{e^{-\lambda}\lambda^x}{x!}$$

$$= e^{-\lambda}\lambda^4 \left\{\sum_{x=4}^{\infty} \frac{\lambda^{x-4}}{(x-4)!}\right\} + 6e^{-\lambda}\lambda^3 \left\{\sum_{x=3}^{\infty} \frac{\lambda^{x-3}}{(x-3)!}\right\} + 7e^{-\lambda}\lambda^2 \left\{\sum_{x=2}^{\infty} \frac{\lambda^{x-2}}{(x-2)!}\right\} + \lambda$$

$$= \lambda^4 (e^{-\lambda} e^{\lambda}) + 6\lambda^3 (e^{-\lambda} e^{\lambda}) + 7\lambda^2 (e^{-\lambda} e^{\lambda}) + \lambda$$

$$= \lambda^4 + 6\lambda^3 + 7\lambda^2 + \lambda$$

We now find the first four central moments as follows:

$$\mu_2 = \mu_2' - \mu_1'^2 = (\lambda^2 + \lambda) - \lambda^2 = \lambda$$

$$\mu_3 = \mu_3' - 3\mu_2'^3\mu_1 + 2\mu_1'^3 = (\lambda^3 + 3\lambda^2 + \lambda) - 3\lambda(\lambda^2 + \lambda) + 2\lambda^3 = \lambda.$$

$$\mu_4 = \mu_4' - 4\mu_3'\mu_1' + 6\mu_2'\mu_1'^2 - 3\mu_1'^4$$

$$= (\lambda^4 + 6\lambda^3 + 7\lambda^2 + \lambda) - 4\lambda(\lambda^3 + 3\lambda^2 + \lambda) + 6\lambda^2(\lambda^2 + \lambda) - 3\lambda^4 = 3\lambda^2 + \lambda$$

Also coefficients of skewness and kurtosis are given by :

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{\lambda^2}{\lambda^3} = \frac{1}{\lambda} \text{ and } \gamma_1 = \sqrt{\beta_1} = \frac{1}{\sqrt{\lambda}}$$

Also $\beta_2 = \dfrac{\mu_4}{\mu_2^2} = 3 + \dfrac{1}{\lambda}$ **and** $\gamma_2 = \beta_2 - 3 = \dfrac{1}{\lambda}.$

Moment Generating Function

Let $M_x(t)$ denotes the m.g.f. of the poisson distribution, then

$$M_x(t) = \sum_{x=0}^{\infty} e^{tx} p(x)$$

$$= \sum_{x=0}^{\infty} e^{tx} \frac{e^{-\lambda}\lambda^x}{x!} = \sum_{x=0}^{\infty} \frac{e^{-\lambda}(\lambda e^t)^x}{x!} = e^{-\lambda}\left\{1 + \lambda t + \frac{(\lambda e^t)^2}{2!} + \ldots\ldots\right\}$$

$$= e^{-\lambda} e^{\lambda e^t} = e^{\lambda(e^t - 1)}$$

Probability Generating Function

The p.g.f. of Poisson distribution can be obtained as follows.

$$P(S) = \sum_{k=0}^{\infty} P(x = k)s^k = \sum_{k=0}^{\infty} \frac{e^{-\lambda}\lambda^k}{k!} s^k = \sum_{k=0}^{\infty} \frac{e^{-\lambda}(\lambda s)^k}{k!}$$

$$= e^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda s)}{k!} s^k = e^{-\lambda} e^{\lambda s} = e^{\lambda(s-1)}.$$

4.25 CHARACTERISTIC FUNCTION OF THE POISSON DISTRIBUTION

NOTES

Let $\phi_x(t)$ denotes the characteristic function of the poisson distribution, then

$$\phi_x(t) = E(e^{itx}) = \sum_{x=0}^{\infty} e^{itx} p(x) = \sum_{x=0}^{\infty} e^{itx} \frac{e^{\lambda}\lambda^x}{x!}$$

$$= \sum_{x=0}^{\infty} e^{-\lambda} \frac{(\lambda e^{it})^x}{x!} = e^{-\lambda}\left\{1 + \lambda e^{it} + \frac{(\lambda e^{it})^2}{2!} + \ldots\right\} = e^{-\lambda} e^{\lambda e^{it}} = e^{\lambda(e^{it}-1)}$$

Reproductive or Additive property of independent Poisson variates

Theorem VIII. *If $X_1, X_2, \ldots\ldots, X_n$ are n independent poisson variates with parameters $\lambda_1, \lambda_2, \ldots\ldots, \lambda_n$, then $X_1 + X_2 + \ldots\ldots + X_n$ is also a Poisson variate with parameter $\lambda_1 + \lambda_2 + \ldots\ldots + \lambda_n$.*

Proof. Let $M_{X_i}(t)$ denotes the moment generating function of the random variable X_i, then

Consider $M_{X_1+X_2+\ldots\ldots+X_n}(t) = M_{X_1}(t)\, M_{X_2}(t) \ldots\ldots M_{X_n}(t)$,

[Since $X_1, X_2, \ldots\ldots, X_n$ are independent.]

$$= e^{\lambda_1(e^t-1)}\, e^{\lambda_2(e^t-1)} \ldots e^{\lambda_n(e^t-1)} = e^{(\lambda_1+\lambda_2+\ldots\ldots+\lambda_n)(e^t-1)}$$

which is the moment generating function of a Poisson variate with parameter $\lambda_1 + \lambda_2 + \ldots\ldots + \lambda_n$. Hence, by uniqueness theorem of moment generating functions $\sum_{i=1}^{n} X_i$ is also a Poisson variate with parameter $\lambda_1 + \lambda_2 + \ldots\ldots + \lambda_n$.

Cor. 1. Converse of reproductive property of Poisson distribution is also true.

Statement. *If $x_1, x_2, \ldots, x_n$ are independent random variables and $\sum_{\Gamma=1}^{n} x_i$ has a Poisson distribution, then each of the random variable $x_1, x_2, \ldots, x_n$ has a Poisson distribution.*

Proof. We Prove the result for I = 2 *i.e.*, If x_1 and x_2 are independent random variables such that $x_1 \sim P(\lambda_1)$ $x_1 + x_2 \sim P(\lambda_1 + \lambda_2)$, then we show that $x_2 \sim P(\lambda_2)$

AS x_1 and x_2 are independent,

$$M_{x_1+x_2} = M_{x_1}(t) + M_{x_2}(t)$$

$$\Rightarrow \quad e^{(\lambda_1+\lambda_2)(e^t-1)} = e^{\lambda_1(e^t-1)} . M_{x_2}(t)$$

$$\Rightarrow \quad e^{\lambda_1(e^t-1)} . e^{\lambda_2(e^t-1)} = e^{\lambda_1(e^t-1)} . M_{x_2(t)}$$

$$\Rightarrow \quad e^{\lambda_2(e^t-1)} = M_{x_2}(t)$$

which is the moment generating function of a Poisson variate. By uniquencess theorem of moment generating function,

$$x_2 \sim P(\lambda_2).$$

Cor. 2. Difference of two independent Poisson variates is not a Poisson variate.

NOTES

4.26 RECURRENCE FORMULA FOR POISSON DISTRIBUTION

Let x be a Poisson variable and $P(x = r) = \frac{e^{-m} m^r}{r!}$, $r = 0, 1, 2, \ldots$

For $k \geq 0$, $P(k) = \frac{e^{-m} m^k}{k!}$ and $P(k+1) = \frac{e^{-m} m^{k+1}}{(k+1)!}$.

Dividing, we get $\frac{P(k+1)}{P(k)} = \frac{e^{-m} m^{k+1}}{(k+1)!} \cdot \frac{k!}{e^{-m} m^k} = \frac{m}{k+1}$.

$$\therefore \quad \mathbf{P(k+1) = \frac{m}{k+1} P(k), \quad k = 0, 1, 2, \ldots\ldots}$$

This is the required **recurrence formula.**

4.27 APPLICATIONS OF POISSON DISTRIBUTION

This distribution is applied to problems concerning.

1. The number of persons born blind per year in a country.
2. The number of deaths by horse kick in an army corps.
3. The number of fragments from a shell hitting a target.
4. Demand pattern for certain spare parts.

ILLUSTRATIVE EXAMPLES

Example 1. *A pair of dice is thrown 200 times. If getting a sum of 9 is considered as success, using Poisson distribution, find the mean and variance of the number of successes.*

Sol. Let p be the probability of getting sum 9 in a throw of pair of dice. Out of total 36 outcomes, the favourable outcomes are (3, 6), (4, 5), (5, 4) and (6, 3).

$$\therefore \qquad p = \frac{4}{36} = \frac{1}{9}.$$

Also, $\qquad n = 200.$

$$\therefore \qquad m = np = 200 \times \frac{1}{9} = \frac{200}{9} = 22.22 \qquad | \because \text{Mean} = \text{Variance}$$

$\therefore$ Mean $= m = 22.22$ and variance $= m = 22.22$.

NOTES

Example 2. *(i) For a Poisson distribution, it is given that, P(X = 1) = P(X = 2). Find the value of mean of the distribution. Hence find P (X = 0) and P(X = 4).*

(ii) A random variable X follows a Poisson distribution with Parameter 4. Find the Probability that X assumes the values less than 2.

Sol. (*i*) Let $$P(X = r) = \frac{e^{-m}\, m^r}{r\,!},\ r = 0, 1, 2, \ldots \qquad \ldots(1)$$

where m is the average number of successes.

We have P(X = 1) = P(X = 2)

$$\Rightarrow \qquad \frac{e^{-m}\, m^1}{1!} = \frac{e^{-m}\, m^2}{2\,!}$$

$$\Rightarrow \qquad m = \frac{m^2}{2} \quad \Rightarrow \quad m = 2 \qquad (\because \quad m \neq 0)$$

$\therefore$ Mean of the distribution, $m = 2$.

Using (1), we have $$P(X = 0) = \frac{e^{-m}\, m^0}{0\,!} = e^{-2}$$

and $$P(X = 4) = \frac{e^{-m}\, m^4}{4\,!} = \frac{e^{-2}(2)^4}{24} = \frac{2}{3}e^{-2}.$$

(*ii*) Here $m = 4$. By using Poisson distribution

we know

$$P(X = r) = \frac{e^{-m}\, m^r}{r\,!},\ r = 0, 1, 2, \ldots \qquad \ldots(1)$$

Required Probability $= P(X < 2) = P(X \leq 1)$

$= P(X = 0) + P(X = 1) = e^{-m} + e^{-m} \,.\, m$ | Using (1)

$= e^{-4}\,(1 + 4) = 5e^{-4} = 5 \times 0.0183 = 0.09157.$

Example 3. *A telephone exchange receives on an average 4 calls per minute. Find the probabilities on the basis of Poisson distribution (m = 4), of :*

(i) 2 or less calls per minute *(ii) upto 4 calls per minute*

(iii) more than 4 calls per minute.

Sol. Let x be the Poisson variable *"no. of calls per minute"*.

By **Poisson distribution,** $$P(x = r) = \frac{e^{-m}\, m^r}{r\,!},\ r = 0, 1, 2, \ldots\ldots$$

Here m = Average number of successes *i.e.*, calls per minute = 4

$$\therefore \qquad P(x = r) = \frac{e^{-4}(4)^r}{r\,!},\ r = 0, 1, 2, \ldots\ldots$$

(*i*) P(2 or less calls per minute) $= P(x \leq 2) = P(x = 0) + P(x = 1) + P(x = 2)$

$$= \frac{e^{-4}\,.\,4^0}{0\,!} + \frac{e^{-4}\,.\,4^1}{1!} + \frac{e^{-4}\,.\,4^2}{2\,!} = e^{-4}\,\{1 + 4 + 8\} = 0.01832 \times 13 = 0.2382.$$

(*ii*) P(upto 4 calls per minute) $= P(x \leq 4) = P(x = 0) + P(x = 1) + P(x = 2) + P(x = 3) + P(x = 4)$

$$= \frac{e^{-4}\,.\,4^0}{0\,!} + \frac{e^{-4}\,.\,4^1}{1!} + \frac{e^{-4}\,.\,4^2}{2\,!} + \frac{e^{-4}\,.\,4^3}{3\,!} + \frac{e^{-4}\,.\,4^4}{4\,!}$$

$$= e^{-4}\left\{1+4+8+\frac{64}{6}+\frac{256}{24}\right\} = 0.01832 \times 34.3333 = 0.6289.$$

NOTES

(*iii*) P(more than 4 calls per minute) = $P(x > 4) = 1 - P(x \leq 4) = 1 - 0.6289 = 0.3711$.

Example 4. *A manufacturer of pins knows that on average 5% of his product is defective. He sells pins in boxes of 100, and guarantees that not more than 4 pins will be defective. What is the probability that a box will meet the guaranteed quality ? (e^{-5} = 0.0067)*

Sol. Let p = The probability that a pin is defective = 5% = 0.05

Also $\qquad n = 100$

$\therefore \qquad \lambda = np = 100 \times (0.05) = 5$

Using Poisson distribution, we have

$$P(X = r) = \frac{e^{-\lambda}\lambda^r}{r!}, r = 0, 1, 2, \ldots$$

The box will meet the guarantee if it contains at the most 4 pins defective.

$\therefore$ Required probability = $P(X \leq 4)$

$$= P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4)$$

$$= e^{-\lambda} + e^{-\lambda}.\ \lambda + e^{-\lambda}\ .\ \frac{\lambda^2}{2!} + e^{-\lambda}\ .\ \frac{\lambda^3}{6} + e^{-\lambda}\ .\ \frac{\lambda^4}{24}$$

$$= e^{-5}\left(1+5+\frac{25}{2}+\frac{125}{6}+\frac{625}{24}\right)$$

$$= e^{-5}\,(6 + 12.5 + 20.83 + 26.04)$$

$$= e^{-5}\,(65.37) = 0.0067 \times 65.37$$

$$= 0.44.$$

Example 5. *Red blood cells deficiency may be determined by examining a specimen of blood under a microscope. Suppose a certain small fixed volume contains on an average 20 red cells for normal persons. Using Poisson distribution, obtain the probability that a specimen from a normal person will contain less than 15 red cells.*

Sol. Let x be the Poisson variable, "*no. of red blood cells in the specimen*".

By **Poisson distribution,** $\quad P(x = r) = \dfrac{e^{-m}\, m^r}{r!}$, $r = 0, 1, 2, \ldots\ldots$

Here m = average number of R.B.C. in the specimen = 20

$$\therefore \qquad P(x = r) = \frac{e^{-20}(20)^r}{r!}, r = 0, 1, 2, \ldots\ldots$$

$\therefore$ P(less than 15 R.B.C. in the specimen)

$$= P(x < 15) = P(x = 0) + P(x = 1) + \ldots\ldots + P(x = 14)$$

$$= \frac{e^{-20}(20)^0}{0!} + \frac{e^{-20}(20)^1}{1!} + \ldots\ldots + \frac{e^{-20}(20)^{14}}{14!} = e^{-20}\sum_{k=0}^{14}\frac{(20)^k}{k!}.$$

Example 6. *An electric bulb manufacturer finds that 4% of the bulbs are defective. What is the probability that a random sample of 50 bulbs does not have a defective bulb. ?*

Sol. Let p = The probability that a bulb is defective

$$= 4\% = \frac{4}{100} = 0.04 \text{ (very small)}$$

Also $n = 50$

$\therefore$ $\lambda = np = 50 \times 0.04 = 2$

By **Poisson distribution,**

$$P(X = r) = \frac{e^{-\lambda}\lambda^r}{r!}, r = 0, 1, 2, ...$$

$$\text{Required probability} = P(X = 0) = \frac{e^{-\lambda}\lambda^0}{0!} = e^{-\lambda} = e^{-2} = 0.1353.$$

Example 7. *If X_1 and X_2 are independent random variable such that $X_1 \sim P(\lambda_1)$ and $X_1 + X_2 \sim P(\lambda_1 + \lambda_2)$, then show that $X_2 \sim P(\lambda_2)$.*

Sol. Let $M_X(t)$ denotes the moment generating function of the random variable X which follows Poisson distribution, then $M_X(t) = e^{\lambda(e^t - 1)}$

$$\therefore \quad M_{X_1 + X_2}(t) = M_{X_1}(t)\, M_{X_2}(t)$$

$$\Rightarrow \quad e^{(\lambda_1+\lambda_2)(e^t-1)} = e^{\lambda_1(e^t-1)}\, M_{X_2}(t) \qquad | X_1, X_2 \text{ are independent}$$

$$\Rightarrow \quad e^{\lambda_1(e^t-1)}\, e^{\lambda_2(e^t-1)} = e^{\lambda_1(e^t-1)}\, M_{X_2}(t)$$

$$\Rightarrow \quad M_{X_2}(t) = e^{\lambda_2(e^t-1)}$$

$$\Rightarrow \quad X_2 \sim P(\lambda_2) \qquad | \text{By uniqueness theorem of m.g.f.}$$

Example 8. *The Difference of two independent Poisson variates is not a Poisson variate.*

Sol. Let X_1 and X_2 are two Poisson variates with parameters λ_1 and λ_2, then

$$M_{X_1 - X_2}(t) = M_{X_1 + (-X_2)}(t)$$

$$= M_{X_1}(t) \,.\, M_{-X_2}(-t)$$

$$= M_{X_1}(t)\, M_{X_2}(t) \qquad | M_{CX_1}(t) = M_{X_1}(Ct)$$

$$= e^{\lambda_1(e^t-1)}\, e^{\lambda_2(e^{-t}-1)}$$

$$= e^{\lambda_1(e^t-1) + \lambda_2(e^{-t}-1)},$$

which cannot be put in the form $e^{\lambda(e^t-1)}$. By uniqueness theorem of moment generating function, $X_1 - X_2$ cannot follow Poisson distribution.

NOTES

The Poisson distribution, then the two modes of the Poisson distribution are given by $x = \lambda, \lambda - 1$. Given, the two modes are the points $x = 1, 2$. It implies $\lambda = 2$. Using Poisson distribution, $P(X = x)$

$$= \frac{e^{-\lambda}\lambda^x}{x!} = \frac{e^{-2}2^x}{x!}, x = 0, 1, 2, \ldots\ldots$$

$$\therefore \quad P(X = 1) = 2e^{-2}, P(X = 2) = \frac{e^{-2}2^2}{2} = 2e^{-2}$$

Hence required probability

$$= P(X = 1) + P(X = 2)$$

$$= 2e^{-2} + 2e^{-2} = 4e^{-2} = 0.542.$$

Example 9. *Show that in a Poisson distribution with unit mean, mean deviation about mean is $\frac{2}{e}$ times the standard deviation.*

Sol. Let $X \sim P(\lambda)$, then using Poisson distribution,

$$P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}, x = 0, 1, 2, \ldots\ldots$$

But given mean $= \lambda = 1$

$\therefore$ M.D. about mean $1 = E \mid X - 1 \mid$

$$= \sum_{x=0}^{\infty} |x - 1| P(X = x)$$

$$= \sum_{x=0}^{\infty} |x-1| \frac{e^{-1}1^x}{x!} = e^{-1} \sum_{x=0}^{\infty} \frac{|x-1|}{x!}$$

$$= e^{-1}\left(1 + 0 + \frac{1}{2!} + \frac{2}{3!} + \ldots\ldots\right)$$

$$= e^{-1}\left[1 + \left(1 - \frac{1}{2!}\right) + \left(\frac{1}{2!} - \frac{1}{3!}\right) + \left(\frac{1}{3!} - \frac{1}{4!}\right) + \ldots\ldots\right]$$

$$= e^{-1}(1 + 1) \qquad \text{Using } \frac{n}{(n+1)!} = \frac{n+1-1}{(n+1)!}$$

$$= \frac{2}{e} \qquad = \frac{n+1}{(n+1)!} - \frac{1}{(n+1)!}$$

$$= \frac{2}{e} \times 1 \qquad = \frac{1}{n!} - \frac{1}{(n+1)!}$$

$$= \frac{2}{e} \text{ standard deviation.}$$

NOTES

SUMMARY

- The Binomial distribution was discovered by **James Bernoulli** in 1700. The following are the essential conditions for the applicability of the Binomial distribution.

  * The number of trials should be finite
  * The tirals should be independent
  * Each trial must result in either "success or failure".
  * The probability of a success is constant in each trial.
- **Histogram for B.D.** To draw the histogram of a B-D, first mark all the values of the random variable on the x-axis and their respective probability on y-axis. Construct the rectangles of uniform width with values of the variables at the centre and heights equal to their corresponding probabilities.
- The Poisson distribution was discovered by a French mathematician, Simon Denis Poisson, in 1837. It is a discrete probability distribution. This distribution is used when there are events which do not occur as outcomes occur randomly in nature.
- Mean and variance of the Poisson distribution are always equal.
- If $X_1, X_2, ..., X_n$ are Poisson variates with Parameter $\lambda_1, \lambda_2, ..., \lambda_n$, respectively, then $\sum_{\Gamma=1}^{n} X_i$ is also a Poisson variate with parameter $\lambda_1 + \lambda_2 + ... + \lambda_n$.
- The Poisson distribution is the limiting case of the Binomial distribution.

GLOSSARY

- **Binomial Variable.** A random variable which counts the number of successes in a random experiment following Binomial distribution is called a Binomial variable.
- **Compound distribution.** If a random variable X is such that its distribution depends upon a single parameter Q, then the distribution of 'θ' will be known as compound distribution of Q.
- **Poisson Variate.** A random variable which follows the Poisson distribution is known as Poisson variate.
- **Poisson frequency distribution.** If a random experiment, is repeated N times, then the expected frequency of getting $r(0 \le r \le N)$ success is given by

$$N.P.(X = r) = \frac{N.e^{-\lambda}\lambda^r}{r!}, r = 0, 1, 2, ...$$

where λ is the parameter.

REVIEW QUESTIONS

NOTES

1. An unbiased coin is tossed 8 times. Find by using binomial distribution, the probability of getting at least 3 heads.
2. 5 dice are thrown simultaneously. If the occurrence of an even number in a single die is considered a success, find the probability of getting at most 3 successes.
3. A box contains 100 tickets each bearing one of the numbers from 1 to 100. If 5 tickets are drawn successively with replacement from the box, find the probability that all the tickets bear number divisible by 10.
4. A policeman fires 6 bullets on a dacoit. The probability that the dacoit will be killed by a bullet is 0.6. What is the probability that the dacoit is still alive ?
5. Assume that the probability that a bomb dropped from an aeroplane will strike a certain target is 0.2. If 6 bombs are dropped, find the probability that :
   (*i*) exactly 2 will strike the target (*ii*) at least 2 will strike the target.
6. For a binomial distribution with $p = \frac{1}{4}$ and $n = 10$, find mean and variance.
7. The mean and S.D. of a binomial distribution are 20 and 4 respectively, calculate n, p and q.
8. (*i*) Find the binomial distribution when the sum of its mean and variance for five trials is 4.8.
   (*ii*) If the sum of the mean and variance of a binomial distribution for 6 trials be $\frac{10}{3}$, find the distribution.
9. It is known that 60% of mice inoculated with a serum are protected from a certain disease. If 5 mice are inoculated, find the probability that :
   (*i*) none contact the disease (*ii*) more than 3 contact the disease.
10. Find the moment generating function of the standard binormal variate $(X - np) / \sqrt{npq}$ and obtain its limiting form as $n \to \infty$. Also interpret the result.
11. *Neyman's Contagious (Compound) Distribution.* Let $X \sim P(\lambda, y)$, where y itself is an observation of a variate $Y \sim P(\lambda_1)$. Find the unconditional distribution of X and show that its mean is less than its variance.
12. If X has Poisson distribution: $P(X = r) = \dfrac{e^{-\lambda}\lambda^r}{r!}$; $r = 0, 1, 2, ...$
    where the parameter λ is a random variable of the continuous type with the density function:
    $f(\lambda) = \dfrac{a^v}{\Gamma(v)} . e^{-a\lambda} \lambda^{(v-1)}$; $\lambda \geq 0, v > 0, a > 0$, derive the distribution of X.
    Show that the characteristic function of X is given by:
    $\Phi_X(t) = E(e^{itX}) = q^v (1 - pe^{it})^{-v}$, where $p = 1/(1 + a)$, $q = 1 - p$.
13. A company knows on the basis of its past experience that 3% of its bulbs are defective. Using Poisson's distribution, find the probability that in a sample of 100 bulbs, no bulb is defective.
    [Given $e^{-3} = 0.04979$]
14. Six per cent of the bolts produced in a certain factory turn out to be defective. Find the probability, using Poisson distribution, that in a sample of 10 bolts chosen at random (*i*) exactly 2 (*ii*) more than 2, will be defective. [Take $e^{-0.6} = 0.549$]

NOTES

15. Assume that the probability that a bomb dropped from an aeroplane will strike a certain target is 1/5. If 6 bombs are dropped, find the probability that :

    (*i*) exactly 2 will strike the target

    (*ii*) at least 2 will strike the target. [Use $e^{-1.2} = 0.3012$]

16. In a certain factory turning out razor blades, there is a small chance 1/500 for any blade to be defective. The blades are in packets of 10. Use Poisson distribution to calculate the approximate number of packets containing :

    (*i*) no defective (*ii*) one defective

    (*iii*) two defective blades in a consignment of 10000 packets. (*P.T.U., M.B.A. Dec. 2000*)

17. Comment on the following statement : "The mean and variance of a Poisson distribution are equal only if the average occurrence of the Poisson variance is ≤ 4".

18. The standard deviation of a Poisson distribution is 3. Find the probability of getting 3 successes.

19. A car-hire firm has two cars, which it hires out day by day. The number of demands for a car on each day is distributed as a Poisson distribution with mean 1.5. Calculate the proportion of days on which neither car is used and the proportion of days on which some demand is refused.

    ($e^{-1.5} = 0.2231$)

20. In a hospital, there are 20 kidney dialysis machines and that the chance of any one of them to be out of service during a day is 0.02. Determine the probability that exactly 3 machines will be out of service on the same day.

21. Write a short note on poisson distribution.

FURTHER READINGS

1. Continuous Univarate distributions–2: N.L. Johnson and S.Kotz, John Wiley and Sons.
2. Introduction to Probability theory with applications: W. Feller, Vol-1: Wiley astern.
3. Introduction to Modern Probability Theory: B.R. Bhat: Wiley Eastern.
4. Introduction to probability and Mathematical Statistics: V.K.Rohatgi: Wiley Eastern.

UNIT–IV

NOTES

CHAPTER 5

CONTINUOUS DISTRIBUTIONS

OBJECTIVES

After going through this chapter, you should be able to:

- know the basic properties of the normal distribution.
- know the area property of the normal distribution.
- know that moments of odd order for the normal distribution vanish.
- know that mean mode and median of the normal distribution are equal.
- know that the probable error (P.E.) is approximately equal to $\frac{2}{3}$ times the standard deviation.
- know about various distribution like log normal, standard Laplace, two parameter Laplace, weibul and logistic distribution.

STRUCTURE

NOTES

NORMAL DISTRIBUTIONS

5.1 INTRODUCTION

The normal distribution is a limiting case of the Binomial distribution under the following conditions :

(1) When n, the number of trials is very large and

(2) p, the probability of a success, is close to $\frac{1}{2}$.

Remark : (*i*) The normal distribution was first discovered by De Moivre, in 1733, a French mathematician.

(*ii*) The normal distribution is a continuous distribution.

5.2 NORMAL DISTRIBUTION

The general equation of the normal distribution is given by

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, -\infty < x < \infty, -\infty < \mu < \infty, \sigma > 0 ;$$

where $\sqrt{2\pi} = 2.5066, e = 2.7183$

The parameters μ and σ are respectively mean and standard deviation of the distribution.

5.3 NORMAL VARIATE

A random variable X is called a normal variate if it follows a normal distribution. If the random variable X follows a normal distribution with mean μ and S.D. σ, then we write $X \sim N(\mu, \sigma^2)$.

5.4 NORMAL CURVE AND ITS PROPERTIES

The graph of the normal distribution is called the **normal curve.**

Properties

(*a*) The graph of the normal distribution is bell-shaped and symmetrical about the line $x = u$.

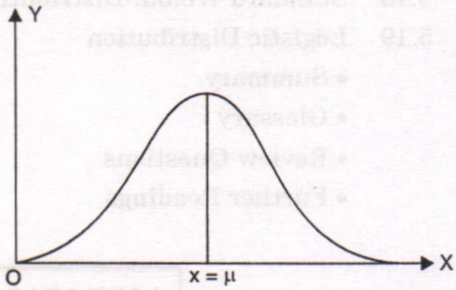
NOTES

(*i.e.*, If we fold the normal curve about the line $x = \mu$, the two halves coincide).

(*b*) The normal curve is unimodal.

(*c*) The line $x = \mu$ divides the area under the normal curves above x-axis into two equal parts (Fig.).

(*d*) The area under the normal curve between any two given ordinates $x = x_1$ and $x = x_2$, represents the probability of values falling into the given interval.



(*e*) The total area under the normal curve above the x-axis is 1.

5.5 BASIC PROPERTIES OF NORMAL DISTRIBUTION

The probability density function (p.d.f) of the normal variate X is given by

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \quad -\infty < x < \infty, -\infty < \mu < \infty, \sigma > 0$$

Then the curve $y = f(x)$, known as normal probability curve and satisfies the following properties.

(*a*) The normal distribution is symmetrical about the line $x = \mu$

(*b*) It is unimodal

(*c*) For a normal distribution, mean = median = mode

(*d*) The area bounded by the curve $y = f(x)$, and x-axis is 1 unit, *i.e.*,

$$\int_{-\infty}^{\infty} f(x)\, dx = 1. \text{ Also } f(x) \geq 0$$

(*e*) The points of inflexion of the normal curve (can be obtained by putting $\frac{d^2y}{dx^2} = 0$ and verifying that at these points, $\frac{d^3y}{dx^3} \neq 0$) are given by $x = \mu \pm \sigma$.

i.e., these points are equidistant from the mean on either side.

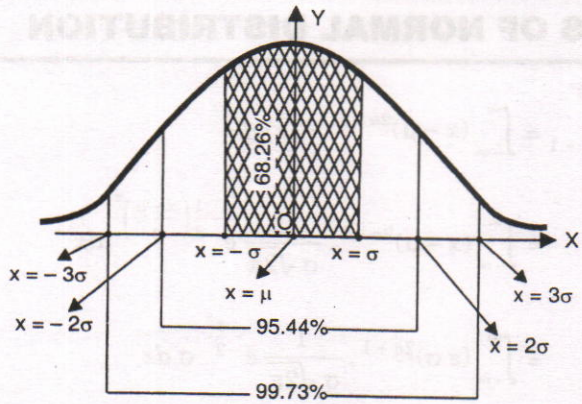
5.6 AREA PROPERTY OF NORMAL DISTRIBUTION

(1) The area under the normal curve between the ordinates $x = \mu - \sigma$ and $x = \mu + \sigma$, is 68.26%.

(2) The area under the normal curve between the ordinates $x = \mu - 2\sigma$ and $x = \mu + 2\sigma$ is 95.44%.

(3) The area under the normal curves between the ordinates $x = \mu - 3\sigma$ and $x = \mu + 3\sigma$ is 99.73%.

NOTES



Area property of normal curve

Theorem I. *For a normal distribution, the mean deviation about mean is* $\frac{4}{5}\sigma$.

Proof. We know that mean deviation about the mean μ is given by

$$\text{M.D.} = \int_{-\infty}^{\infty} |x-\mu|\, f(x)\, dx$$

$$= \int_{-\infty}^{\infty} |x-\mu| . \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx \qquad \Big|\ \text{Put } \frac{x-\mu}{\sigma} = z \ \Rightarrow\ dx = \sigma\, dz$$

$$= \int_{-\infty}^{\infty} |z\,\sigma| . \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{z^2}{2}}\, \sigma\, dz$$

$$= \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{\infty} |z|\, e^{-\frac{z^2}{2}}\, dz \qquad \Big|\ \sigma > 0$$

$$= \frac{2\sigma}{\sqrt{2\pi}} \int_{0}^{\infty} |z|\, e^{-\frac{z^2}{2}}\, dz \qquad \Big|\ |z|\, e^{-\frac{z^2}{2}} \text{ is an even function}$$

$$= \frac{\sqrt{2}}{\pi}\sigma \int_{0}^{\infty} z e^{-\frac{z}{2}}\, dz$$

$$= \frac{\sqrt{2}}{\pi}\sigma \int_{0}^{\infty} e^{-t} dt = \sqrt{\frac{2}{\pi}}\, \sigma \left| \frac{e^{-t}}{(-1)} \right|_{0}^{\infty}$$

$$= \sqrt{\frac{2}{\pi}}\, \sigma\, (0+1) \qquad \Big|\ \text{Put } \frac{z^2}{2} = t \ \Rightarrow\ z\, dz = dt$$

$$= \sqrt{\frac{2}{\pi}}\, \sigma = 0.7979\, \sigma \cong \frac{4}{5}\, \sigma.$$

5.7 MOMENTS OF NORMAL DISTRIBUTION

NOTES

Consider $\mu_{2n+1} = \int_{-\infty}^{\infty} (x-\mu)^{2n+1} . f(x)\, dx$

$$= \int_{-\infty}^{\infty} (x-\mu)^{2n+1} . \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

$$= \int_{-\infty}^{\infty} (z\sigma)^{2n+1} . \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{z^2}{2}} \sigma\, dz$$

$$\left| \text{Put } z = \frac{x-\mu}{\sigma} \Rightarrow dz = \frac{1}{\sigma} dx \right.$$

$$= \sigma^{2n+1} \int_{-\infty}^{\infty} z^{2n+1} e^{-\frac{z^2}{2}} dz$$

$$= 0, \text{ since } z^{2n+1} e^{-\frac{z^2}{2}} \text{ is an odd function.}$$

i.e., **all odd order moments about the mean vanish.**

Further, $\mu_{2n} = \int_{-\infty}^{\infty} (x-\mu)^{2n} . f(x)\, dx$

$$= \int_{-\infty}^{\infty} (x-\mu)^{2n} . \frac{1}{\sigma\sqrt{2\pi}} . e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

$$= \frac{\sigma^{2n}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z^{2n} . e^{-\frac{z^2}{2}} dz \qquad \left| \text{Put } \frac{x-\mu}{\sigma} = z \Rightarrow dx = \sigma\, dz \right.$$

$$= \frac{\sigma^{2n}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z^{2n-1} \left(e^{-\frac{z^2}{2}} z \right) dz \text{ Integrating by parts,}$$

$$= \frac{\sigma^{2n}}{\sqrt{2\pi}} \left[\left| z^{2n-1} \left(-e^{-\frac{z^2}{2}} \right) \right|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} (2n-1)\, z^{2n-2} e^{-\frac{z^2}{2}} dz \right]$$

$$= \frac{\sigma^{2n}}{\sqrt{2\pi}} \left[(0-0) + (2n-1) \int_{-\infty}^{\infty} z^{2n-2} e^{-\frac{z^2}{2}} dz \right]$$

$$= \frac{(2n-1) . \sigma^{2n}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z^{2n-2} e^{-\frac{z^2}{2}} dz$$

$$= (2n-1)\, \sigma^2 . \frac{\sigma^{2n-2}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z^{2n-2} . e^{-\frac{z^2}{2}} dz$$

Hence $\mu_{2n} = (2n-1)\, \sigma^2 . \mu_{2n-4}$ ...(1)

Changing n to $n - 1$ in (1), we get

$$\mu_{2n-2} = (2n-3)\,\sigma^2\,\mu_{2n-4} \qquad ...(2)$$

Using (2) in (1), we get

$$\mu_{2n} = (2n-1)(2n-3)\,\sigma^4 . \mu_{2n-4}$$

....

$$= (2n-1)(2n-3) \ldots\ldots 3.1 \,.\, \sigma^{2n}\,\mu_{2n-2n}$$

$$= (2n-1)(2n-3) \ldots\ldots 3.1 \,.\, \sigma^{2n}. \qquad | \;\mu_0 = 1$$

5.8 VARIANCE OF NORMAL DISTRIBUTION

If we put $n = 1, 2,$ in above, we get

Variance $\mu_2 = \sigma^2$, $\mu_4 = 3.1 \,.\, \sigma^4 = 3\,\sigma^4$

$$\therefore \qquad \beta_1 = \frac{\mu_3^{\,2}}{\mu_2^{\,3}} = 0 \qquad | \text{ As odd order moments about the mean vanish}$$

$$\beta_2 = \frac{\mu_4}{\mu_2^{\,2}} = \frac{3\,\sigma^4}{\sigma^4} = 3.$$

Theorem II. *For the normal distribution, show that the value of mode is given by $x = \mu$, where μ is the mean.*

Proof. Let $f(x)$ be the probability density function of the normal distribution, then

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \; -\infty < x < \infty, \; -\infty < \mu < \infty, \; \sigma > 0, \qquad ...(1)$$

Now, the mode is the solution of $f'(x) = 0$, and $f''(x) < 0$

Taking logarithm of (1), we get

$$\log f(x) = \log \frac{1}{C\sqrt{2\pi}} + \log e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

$$= -\log C\sqrt{2\pi} - \frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2$$

$$= A - \frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2 \quad \text{where } A = \log C\sqrt{2\pi}$$

Differentiating w.r.t. x, we get

$$\frac{f'(x)}{f(x)} = -\left(\frac{x-\mu}{\sigma}\right) . \frac{1}{\sigma} = -\frac{1}{\sigma^2}(x-\mu)$$

$$\Rightarrow \qquad f'(x) = -\frac{1}{\sigma^2}(x-\mu)\, f(x) \qquad ...(2)$$

Again differentiating,

$$f'(x) = -\frac{1}{\sigma^2}\left[(x-\mu)\, f'(x) + f(x)\right]$$

$$= \frac{-1}{\sigma^2}\left[\frac{-(x-\mu)^2}{\sigma^2}\cdot f(x) + f(x)\right]$$

NOTES

$$= -\frac{f(x)}{\sigma^2}\left[1 - \frac{(x-\mu)^2}{\sigma^2}\right] \quad ...(3)$$

For mode, from (2) $f'(x) = 0$ gives

$$\frac{-1}{\sigma^2}(x-\mu)\,f(x) = 0 \;\Rightarrow\; x-\mu = 0$$

$$\Rightarrow \qquad x = \mu$$

Also $\quad [f''(x)]_{x=\mu} = -\frac{1}{\sigma^2}\cdot[f(x)]_{x=\mu}\,(1-0)$

$$= \frac{1}{\sigma^2}\cdot\frac{1}{\sigma\sqrt{2\pi}} = \frac{-1}{\sigma^3\,.\sqrt{2\pi}} < 0$$

Hence the mode is given by $x = \mu$.

Theorem III. *Show that for the normal distribution, median is equal to the mean of the distribution.*

Proof. Let M denotes the median and $f(x)$ is the probability density function, then

$$\int_{-\infty}^{M} f(x)\,dx = \frac{1}{2}$$

$$\Rightarrow \qquad \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{M} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx = \frac{1}{2}$$

$$\Rightarrow \quad \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\mu} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx + \frac{1}{\sigma\sqrt{2\pi}}\int_{\mu}^{M} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx = \frac{1}{2} \quad ...(2)$$

Consider $\dfrac{1}{\sigma\sqrt{2\pi}}\displaystyle\int_{-\infty}^{\mu} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$

$$= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{0} e^{-\frac{1}{2}z^2}\,dz$$

$z = \dfrac{x-\mu}{\sigma} \;\Rightarrow\; dz = \dfrac{dx}{\sigma}$
When $x = \mu$, $z = 0$,
When $x \to -\infty$, $z \to -\infty$

$$= \frac{1}{\sqrt{2\pi}}\cdot\sqrt{\frac{\pi}{2}} = \frac{1}{2}$$

$\therefore$ from (2)

$$\frac{1}{2} + \frac{1}{\sigma\sqrt{2\pi}}\int_{\mu}^{M} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx = \frac{1}{2}$$

$$\Rightarrow \qquad \frac{1}{\sigma\sqrt{2\pi}}\int_{-\mu}^{M} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx = 0,$$

which is possible only when $\mu = M$ *i.e.*, for the normal distribution mean = median

Cor. Show that for the normal distribution, mean, median and mode consider. Hence, the normal distribution is symmetrical.

Proof. Use above two theorems, we get the required result.

Moment Generating Function

Theorem IV. *Find the moment generating function of the normal distribution.*

Proof. Let $f(x)$ denotes the probability density function of the normal distribution and if $M_X(t)$ denotes the moment generating function of the random variable X following normal distribution, then

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, -\infty < x < \infty, -\infty < \mu < \infty, \sigma > 0$$

$$\therefore \quad M_X(t) = \int_{-\infty}^{\infty} e^{tx} f(x)\, dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{tx} \exp\{-(x-\mu)^2/2\sigma^2\}\, dx$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\{t(\mu + \sigma z)\} \exp(-z^2/2)\, dz,$$

$$= e^{\mu t} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left\{-\frac{1}{2}(z^2 - 2t\sigma z)\right\} dz \qquad \Big| \text{Put } z = \frac{x-\mu}{\sigma}$$

$$= e^{\mu t} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2}\{(z-\sigma t)^2 - \sigma^2 t^2\}\right] dz \qquad \Big| \Rightarrow dz = \frac{dx}{\sigma}$$

$$= e^{\mu t + t^2\sigma^2/2} \times \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left\{-\frac{1}{2}(z-\sigma t)^2\right\} dz$$

$$= e^{\mu t + t^2\sigma^2/2} \times \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp(-u^2/2)\, du$$

$$= e^{\mu t + \frac{\sigma^2 t^2}{2}} \times \frac{1}{\sqrt{2\pi}} \cdot 2 \int_0^{\infty} \exp\left(-\frac{\mu^2}{2}\right) du$$

$$= e^{\mu t + \frac{\sigma^2 t^2}{2}} \cdot \frac{1}{\sqrt{2\pi}} \cdot 2 \cdot \sqrt{\frac{\pi}{2}} = e^{\mu t + \frac{\sigma^2 t^2}{2}}$$

$$\therefore \quad M_X(t) = e^{\mu t + \frac{\sigma^2 t^2}{2}}.$$

Cor. Find the moment generating function of the standard normal variate.

Proof. Let X ~ N (μ, σ^2) be a normal variate with parameters μ and σ.

Take $Z = \dfrac{X-\mu}{\sigma}$, then

$$M_Z(t) = M_{\frac{X-\mu}{\sigma}}(t) = e^{-\frac{\mu t}{\sigma}} M_{\frac{X}{\sigma}}(t) = e^{-\frac{\mu t}{\sigma}} M_X\left(\frac{t}{\sigma}\right)$$

$$= e^{-\frac{\mu t}{\sigma}} \exp\left[\frac{\mu t}{\sigma} + \frac{t^2}{\sigma^2} \cdot \frac{\sigma^2}{2}\right] = e^{-\frac{\mu t}{\sigma}} \cdot e^{\frac{\mu t}{\sigma} + \frac{t^2}{2}} = e^{\frac{t^2}{2}}.$$

NOTES

Theorem V. *Let $X_1, X_2, \ldots\ldots, X_n$ are n independent normal variates with mean $\mu_1, \mu_2, \ldots, \mu_n$ and variances $\sigma_1^2, \sigma_2^2, \ldots\ldots, \sigma_n^2$ respectively, then show that $a_1X_1 + a_2X_2 + \ldots\ldots + a_nX_n$ is also a normal variate with a_n $a_1\mu_1 + a_2\mu_2 + \ldots\ldots + \mu_n$ and variances $a_1^2\sigma_1^2 + a_2^2\sigma_2^2 + \ldots\ldots + a_n^2\sigma_n^2$ respectively.*

Proof. Let $X \sim N(\mu, \sigma^2)$ be a normal variate, then we know

$$M_X(t) = \exp\left(\mu t + \frac{t^2\sigma^2}{2}\right) \text{ and}$$

$$M_{X_i} = \exp\{\mu_i t + (t^2\sigma_i^2/2)\} \qquad \ldots(1)$$

The moment generating function of their linear combination $\sum_{i=1}^{n} a_i X_i$, where $a_1, a_2, \ldots, a_n$ are constants, is given by:

$$M_{\sum_i a_i X_i(t)} = \prod_{i=1}^{n} M_{a_i X_i(t)} \qquad (\because\ X_i's \text{ are independent})$$

$$= M_{X_1}(a_1t) \cdot M_{X_2}(a_2t) \ldots\ldots M_{X_n}(a_nt) \qquad [\because\ M_{cX}(t) = M_X(ct)]$$

From (1), we have $M_{X_i}(a_it) = e^{\mu_i a_i t + t^2 a_i^2 \sigma_i^2/2}$ ...(2)

$$\therefore \quad M_{\sum_i a_i X_i}(t) = \left[e^{\mu_1 a_1 + t^2 a_1^2 \sigma_1^2/2} \times e^{\mu_2 a_2 t + t^2 a_2^2 \sigma_2^2/2} \times \ldots\ldots \times e^{\mu_n a_n t + t^2 a_n^2 \sigma_n^2/2}\right]$$

[Using (2)

$$= \exp\left[\left(\sum_{i=1}^{n} a_i\mu_i\right)t + t^2\left(\sum_{i=1}^{n} a_i^2\ \sigma_i^2\right)/2\right],$$

which is the moment generating function of a normal variate with mean $\sum_{i=1}^{n} a_i\mu_i$ and variance $\sum_{i=1}^{n} a_i^2\sigma_i^2$.

Hence by uniqueness theorem of moment generating function.,

$$\sum_{i=1}^{n} a_iX_i \sim N\left[\sum_{i=1}^{n} a_i\mu_i, \sum_{i=1}^{n} a_i^2\ \sigma_i^2\right] \qquad \ldots(3)$$

Cor I. If $X_1 \sim N(\mu_1, \sigma_1^2)$ and $X_2 \sim N(\mu_2^2, \sigma_2^2)$ be two normal variates with means μ_1 and μ_2 and variances σ_1^2 and σ_2^2 respectively, then

(*i*) $X_1 + X_2 \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$

(*ii*) $X_1 - X_2 \sim N(\mu_1 - \mu_2, \sigma_1^2 + \sigma_2^2)$.

Proof. Take $a_1 = 1, a_2 = 1, a_3 = 0, a_4 = 0 \ldots\ldots a_n = 0$ in above theorem, we get (Using equation (3))

$$X_1 + X_2 \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$$

Again, take $a_1 = 1, a_2 = -1, a_3 = 0, a_4 = 0, ..., a_n = 0$ in (3), we get

$$X_1 - X_2 \sim N(\mu_1 - \mu_2, \sigma_1^2 + \sigma_2^2).$$

NOTES

5.9 REPRODUCTIVE PROPERTY

Cor II. If $X_1, X_2, ..., X_n$ are n independent normal variates with means $\mu_1, \mu_2, ..., \mu_n$ and variances $\sigma_1^2, \sigma_2^2, ..., \sigma_n^2$, then $X_1 + X_2 + ... + X_n$ is also a normal variate with wean $\mu_1 + \mu_2 + ... + \mu_n$ and variances $\sigma_1^2 + \sigma_2^2 + ... + \sigma_n^2$ respectively. This is called **additive property** or **reproductive property** of the normal distribution.

Proof. Take $a_1 = a_2 = ... = a_n = 1$ in (3) of above theorem, we get the required result.

Theorem VI. *If $X_1, X_2, ..., X_n$ are n identically and independently distributed normal variates with mean μ and variance σ^2, then their mean $\overline{X}$ is also a normal variate with mean μ and variance $\frac{\sigma^2}{n}$.*

Proof. Take $a_1 = a_2 = ... = a_n = \frac{1}{n}$ in (3) of above theorem, we get

$$\frac{X_1 + X_2 + ... + X_n}{n} \sim N\left(\frac{\mu + \mu + ... + \mu}{n}, \frac{\sigma^2 + \sigma^2 + ... + \sigma^2}{n^2}\right)$$

or $$\overline{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

5.10 PROBABILITY INTEGRAL OR ERROR FUNCTION

The integral $P(z) = \frac{1}{\sqrt{2\pi}} \int_0^z e^{-\frac{z^2}{2}} dz$, is known as probability integral or error function. We prove this.

We know that the probability of x lying between x_1 and x_2 is given by the area under the normal curve from x_1 to x_2 *i.e.,*

$$P(x_1 \le x \le x_2) = \frac{1}{\sigma\sqrt{2\pi}} \int_{x_1}^{x_2} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \qquad ...(1)$$

Put $\frac{x - \mu}{\sigma} = z \quad \Rightarrow \quad dx = \sigma d_z$

when $x = x_1, z = \frac{x_1 - \mu}{\sigma} = z_1$, say,

when $x = x_2, z = \frac{x_2 - \mu}{\sigma} = z_2$, say,

$\therefore$ (1) gives

$$P(x_1 \le x \le x_2) = \frac{1}{\sqrt{2\pi}} \int_{z_1}^{z_2} e^{-\frac{z^2}{2}} dz$$

NOTES

$$= \frac{1}{\sqrt{2\pi}} \int_0^{z_2} e^{-\frac{z^2}{2}} dz - \frac{1}{\sqrt{2\pi}} \int_0^{z_1} e^{-\frac{z^2}{2}} dz$$

$$= P_2(z) - P_1(z)$$

$$\therefore \quad P(z) = \frac{1}{\sqrt{2\pi}} \int_0^{z} e^{-\frac{z^2}{2}} dz,$$

which is called probability generating function.

Probable Error (λ)

It is defined as the deviation on either side of the arithmetic mean, the probability of occurrence of which is equal to 0.5

i.e., It is the value of λ, satisfying,

$$\frac{1}{\sigma\sqrt{2\pi}} \int_{\mu-\lambda}^{\mu+\lambda} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = 0.5$$

$$\Rightarrow \quad \frac{1}{\sigma\sqrt{2\pi}} \int_{-\frac{\lambda}{\sigma}}^{\frac{\lambda}{\sigma}} e^{-\frac{z^2}{2}} . \sigma\, dz$$

Put $\frac{x-\mu}{\sigma} = z \Rightarrow dx = \sigma\, dz$

When $x = \mu + \lambda$, $z = \frac{\lambda}{\sigma}$. When $x = \mu - \lambda$, $z = \frac{-\lambda}{\sigma}$

$$\Rightarrow \quad \frac{2}{\sqrt{2\pi}} \int_0^{\frac{\lambda}{\sigma}} e^{-\frac{z^2}{2}} dz = 0.5$$

$e^{-\frac{z^2}{2}}$ is an even function

$$\Rightarrow \quad \frac{1}{\sqrt{2\pi}} \int_0^{\frac{\lambda}{\sigma}} e^{-\frac{z^2}{2}} dz = 0.25$$

Using the table, $\frac{\lambda}{\sigma} = 0.67 \Rightarrow \lambda = 0.67\,\sigma \cong \frac{2}{3}\sigma$

Hence the Probable error $\lambda \cong \frac{2}{3}\sigma$.

5.11 APPLICATIONS OF NORMAL DISTRIBUTION

This distribution is applied to Problems concerning :

(1) Calculation of hit probability of a shot.

(2) Statistical inference in most branches of science.

(3) Calculation of errors made by chance in experimental measurements.

5.12 STANDARD FORM OF NORMAL DISTRIBUTION

If x is a normal random variable with mean μ and standard deviation σ, then the random variable defined by

$$z = \frac{x - \mu}{\sigma},$$

is said to be a standard normal variate with mean 0 and standard deviation 1, *i.e.*, $z \sim N(0, 1)$.

NOTES

The probability density function for the standard normal variate is given by

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}, -\infty < z < \infty.$$

The integral $\int_0^z f(z)\, dz$, cannot be evaluated analytically. The values of this integral for various positive values of z have been given in the table.

ILLUSTRATIVE EXAMPLES

Example 1. *Let z be a standard normal variate, then find*

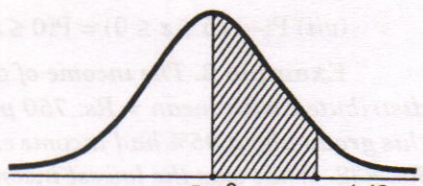
(*i*) $P(0 \le z \le 1.42)$ (*ii*) $P(z \ge -1.28)$
(*iii*) $P(|z| \le 0.5)$ (*iv*) $P(-0.73 \le z \le 0)$
(*v*) $P(0.81 \le z \le 1.94)$ (*vi*) $P(|z| \ge 10.5)$
(*vii*) $P(-0.75 \le z \le 0)$.

Sol. (*i*) We know that

$P(0 \le z \le 1.42)$ = Area under the standard normal curve between the ordinates $z = 0$ and $z = 1.42$

$= 0.4222$

(In the table given in the end, move down the column marked z until we get the entry 1.42 and then move right to column marked 2. The required entry is 0.4222.)



(*ii*) $P(z \ge -1.28)$ = Area under the standard normal curve to the right of $z = -1.28$

= (Area between $z = -1.28$ and $z = 0$)
+ (Area to the right of $z = 0$)

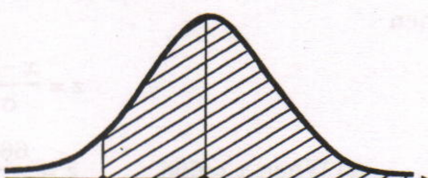
$= P(-1.128 \le z \le 0) + P(z \ge 0)$

$= P(0 \le z \le 1.28) + P(z \ge 0)$

(Due to symmetry.)

$= 0.3997 + 0.5$ | From normal table
| $P(z \ge 0) = 0.5$

$= 0.89997.$



(*iii*) $P(|z| \le 0.5) = P(0.5 \le z \le 0.5)$

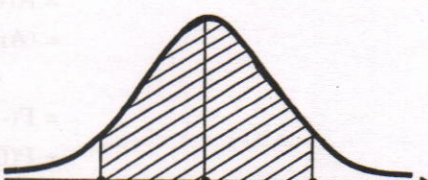
= Area between $z = -0.5$ and $z = 0.5$

= 2 (Area between $z = 0$ and $z = 0.5$)

$= 2P(0 \le z \le 0.5)$

$= 2(0.1915) = 0.3830.$

| From normal table



(*iv*) $P(-0.73 \le z \le 0)$

$= P(0 \le z \le 0.73)$ | By symmetry

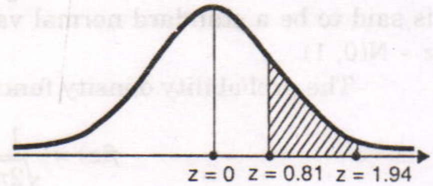
$= 0.2673.$ | From normal table

NOTES

(*v*) $P(0.81 \le z \le 1.94)$

= Area under the normal variate between $z = 0.81$ and $z = 1.94$

= (Area between $z = 0$ and $z = 1.94$) – (Area between $z = 0$ and $z = 0.81$)



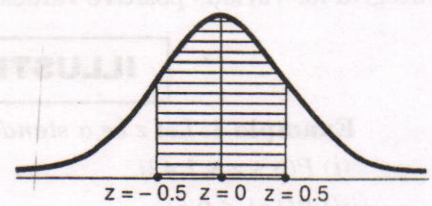
$= P(0 \le z \le 1.94) - P(0 \le z \le 0.81)$

$= 0.4738 - 0.2910$ | From normal table

$= 0.1828.$

(*vi*) $P(|z| \ge 0.5) = P(z \ge 0.5 \text{ or } z \le 0.5)$

| $|a| \ge b \Rightarrow a \ge b$ or $a \le -b$



$= P(z \ge 0.5) + P(z \le -0.5)$

= (Area to the right of $z = 0.5$) + (Area to the left of $z = -0.5$)

= 2(Area to the right of $z = 0.5$)

= 2 [(Area to the right of $z = 0$) – (Area between $z = 0$ and $z = 0.5$)]

$= 2[0.5 - P(0 \le z \le 0.5)$

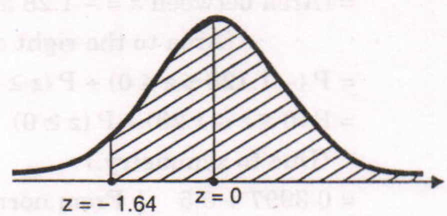
$= 2(0.5 - 0.1915) = 2\,(0.3085)$ | From normal table

$= 0.6170.$

(*vii*) $P(-.75 \le z \le 0) = P(0 \le z \le 0.75) = 0.2734.$ | Due to symmetry

Example 2. *The income of a group of 10,000 persons was found to be normally distributed with mean = Rs. 750 p.m. and standard deviation = Rs. 50. Show that of this group about 95% had income exceeding Rs. 668 and only 5% had income exceeding Rs. 832. What was the lowest income among the richest ?*

Sol. Let x denote the income then, given, x is a normal variate with mean $\mu = 750$ and S.D. $\sigma = 50$. Let z be the standard normal variate, then



$$z = \frac{x - \mu}{\sigma} = \frac{x - 750}{50}$$

(*i*) When $x = 668$, $z = \dfrac{668 - 750}{50}$

$$= \frac{-82}{50} = -1.64$$

$P(x > 668) = P(z > -1.64)$

= Area to the right of $z = -1.64$

= (Area between $z = -1.64$ and $z = 0$) + (Area to the right of $z = 0$)

$= P(-1.64 \le z \le 0) + P(z \ge 0)$

$= P(0 \le z \le 1.64) + P(z \ge 0)$

$= 0.4495 + 0.5 = 0.9495$ | See normal table

Hence required % of persons having income greater than Rs. 668

$= 94.95\% \cong 95\%$

(*ii*) When $x = 832$, $\quad z = \dfrac{832 - 750}{50} = 1.64$

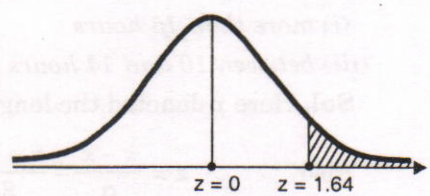
$\therefore \quad P(x > 832) = P(z > 1.64)$

$= \text{Area to the right of } z = 1.64$

$= (\text{Area to the right of } z = 0) - (\text{Area between } z = 0 \text{ and } z = 1.64)$

$= P(z \geq 0) - P(0 \leq z \leq 1.64)$

$= 0.5 - 0.4495 = 0.0505$



Hence required % of persons having income greater than Rs. 832 = 5.

Lastly, to find the lowest income among the richest 100, we need to find the value of r such that $P(x \geq r) = 0.01$

when $x = r$, $\quad z = \dfrac{x - \mu}{\sigma} = \dfrac{r - 750}{50} = z_1$, say

Now $\quad P(x \geq r) = 0.01$

$\Rightarrow \quad P(z \geq z_1) = 0.01$

$\Rightarrow \quad 0.5 - P(0 \leq z \leq z_1) = 0.01$

$\Rightarrow \quad P(0 \leq z \leq z_1) = 0.5 - 0.01 = 0.49$

$\Rightarrow \quad z_1 = 2.33$ | See normal table

$\Rightarrow \quad \dfrac{r - 750}{50} = 2.33 = r = 750 + 50\,(2.33) = 866.5$

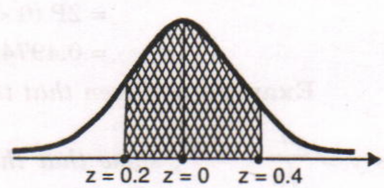
Hence the lowest income among the richest 100 = Rs. 866.50.

Example 3. *The weekly wages of 1,000 workmen are normally distributed with a mean of Rs. 70 and a standard deviation of Rs. 5. Estimate the number of workers whose weekly wages will be between Rs. 69 and 72.*

Sol. Let X be a random variable following normal distribution.

Given $\mu = 70$, $\sigma = 5$

Let $Z = \dfrac{X - \mu}{\sigma}$ be the standard normal variate



When X = 69, $Z = \dfrac{69 - 70}{5} = -0.2$

When X = 72, $Z = \dfrac{72 - 70}{5} = 0.4$

$\therefore \quad P(69 < X < 72) = P(-0.2 < Z < 0.4)$

$= P(-0.2 < Z < 0) + P(0 < Z < 0.4)$

$= P(0 < Z < 0.2) + P(0 < Z < 0.4)$ | Due to symmetry

$= 0.0793 + .1554 = 0.2347$ | Using normal table

NOTES

Hence, number of workers getting wages between Rs. 69 and 72

$$= (0.2347) \times 1000 = 234.7$$

$$\cong 234.$$

NOTES

Example 4. *A sample of 100 dry battery cells tested to find the length of life produced the following results*

$$\bar{x} = 12 \text{ hours}, \sigma = 3 \text{ hours}.$$

Assuming the data to be normally distributed, what percentage of battery cells are expected to have

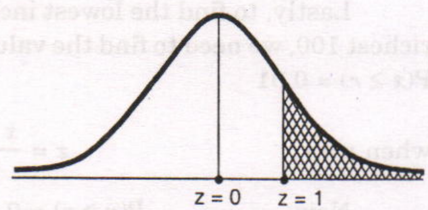
(i) more than 15 hours *(ii) less than 6 hours*

(iii) between 10 and 14 hours ?

Sol. Here x denoted the length of life of dry battery cells.

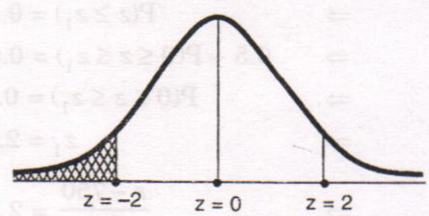
Also $z = \dfrac{x - \bar{x}}{\sigma} = \dfrac{x - 12}{3}$

(*i*) When $x = 15, z = 1$



$$\therefore \quad P(x > 15) = P(z > 1)$$
$$= P(0 < z < \infty) - P(0 < z < 1)$$
$$= 0.5 - 0.3413 = 0.1587$$
$$= 15.87\%.$$

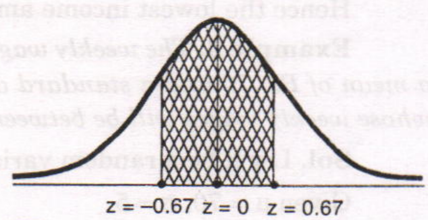
(*ii*) When $x = 6, z = -2$



$$\therefore \quad P(x < b) = P(z < -2)$$
$$= P(z > 2)$$
$$= P(0 < z < \infty) - P(0 < z < 2)$$
$$= 0.5 - 0.4772 = 0.0228$$
$$= 2.28\%.$$

(*iii*) When $x = 10, z = -\dfrac{2}{3} = -0.67$

When $x = 14, \quad z = \dfrac{2}{3} = 0.67$



$$P(10 < x < 14) = P(-0.67 < z < 0.67)$$
$$= 2P\,(0 < z < 0.67) = 2 \times 0.2487$$
$$= 0.4974 = 49.74\%.$$

Example 5. *Given that the probability of committing an error of magnitude x is* $y = \dfrac{h}{\sqrt{\pi}} e^{-h^2 x^2}$, *show that the probable error is 0.4769/h.*

Sol. Using normal distribution, we know

$$y = f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad \ldots(1)$$

Also given probability of committing an error is

$$y = \frac{h}{\sqrt{\pi}} e^{-h^2 x^2} \quad \ldots(2)$$

NOTES

From (1) and (2),

$$\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} = \frac{h}{\sqrt{\pi}} e^{-h^2 x^2} \Rightarrow \frac{1}{\sigma\sqrt{2\pi}} = \frac{h}{\sqrt{\pi}}$$

$$\Rightarrow \qquad h = \frac{1}{\sigma\sqrt{2}} \Rightarrow \sigma = \frac{1}{\sqrt{2}h}$$

$\therefore$ Required probable error

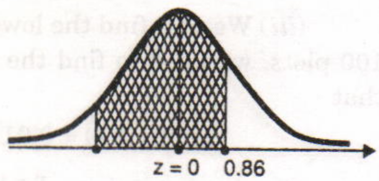
$$= \frac{2}{3}\sigma = \frac{2}{3} \cdot \frac{1}{\sqrt{2}h} = \frac{\sqrt{2}}{3h} = \frac{0.4769}{h}.$$

Example 6. *Assume that 4 percent of the population over 65 years old has Alzheimer's disease. Suppose a random sample of 3500 people over 65 is taken. Find the probability that fewer than 150 of them have the disease.*

Sol. Here $n = 3500, p = 4\% = 0.04$

$$q = 1 - p = 1 - 0.04 = 0.96$$

Here $\mu = \text{mean} = np = 3500 \cdot \frac{4}{100} = 140$



Standard deviation

$$= \sigma = \sqrt{npq} = \sqrt{3500 \cdot \frac{4}{100} \cdot \frac{96}{100}}$$

$$= \sqrt{134.4} = 11.6$$

Let x denote the number of people with Alzheimer's disease, then we required to find $P(x < 150)$

When $x = 150$, $z = \frac{x-\mu}{\sigma} = \frac{150-140}{11.6} = \frac{10}{11.6} = \frac{100}{116} = 0.86$

$\therefore \quad P(x < 150) = P(z < 0.86)$

$= \text{Area under standard normal variate to the left of } z = 0.86$

$= (\text{Area to the left of } z = 0) + (\text{Area between } z = 0 \text{ and } z = 0.86)$

$= 0.5 + P(0 \le z \le 0.86)$

$= 0.5 + 0.3051 = 0.8051.$

Example 7. *The mean yield for 1 acre plot is 662 kilos with S.D. of 32 kilos. Assuming normal distribution, how many 1 acre plots in a batch of 1000 plots would you expect to have yield (i) over 700 kilos (ii) below 650 kilos (iii) what is the lowest yield of best 100 plots?*

Sol. Given : mean $\mu = 662$. Standard deviation $\sigma = 32$. Let x be the normal variate. Also $N = 100$.

(*i*) When $x = 700$, $z = \frac{x-\mu}{\sigma} = \frac{700-662}{32} = \frac{38}{32} = \frac{19}{16} = 1.1875$

$\therefore$ Required probability $= N.P\,(x > 700)$

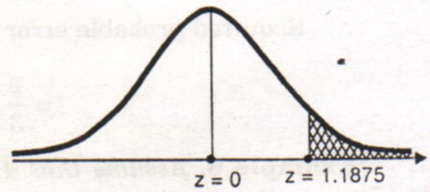
$= 1000 \,.\, P(z > 1.1875)$

$= 1000\,(\text{Area to the right of } z = 1.1875)$

$= 1000$ [(Area to the right of $z = 0$) – (Area between $z = 0$ and $z = 1.1875$)]
$= 1000\ [(P(z \geq 0) - P(0 \leq z \leq 1.1875)]$
$= 1000\ (0.5 - 0.3810] = 1000\ .\ (.1190) = 119$

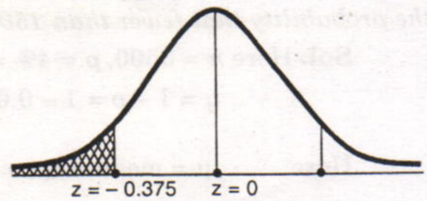
NOTES

(*ii*) When $x = 650$, $z = \dfrac{650 - 662}{32} = \dfrac{-12}{32} = \dfrac{-3}{8} = -0.375$



Required probability

$= N.P(x < 650)$
$= 1000.P\ (z < -0.375)$
$= 1000$ (Area to the right of $z = 0.375$)
$= 1000\ [(P\ (z \geq 0) - P(0 \leq z \leq 0.375)]$
$= 1000\ (0.5 - 0.1443)$
$= 1000 \times (0.3557) = 355.7$



(*iii*) We now find the lowest yield of best 100 plots, we need to find the value of r such that

$$P\ (x \geq r) = 0.01$$

When $x = r$, $\quad z = \dfrac{x - \mu}{\sigma} = \dfrac{r - 662}{32} = z_1$, say

Now $\quad P(x \geq r) = 0.01 \quad \Rightarrow \quad P(z \geq z_1) = 0.01$

$\Rightarrow \quad 0.5 - P(0 \leq z \leq z_1) = 0.01$

$\Rightarrow \quad P(0 \leq z \leq z_1) = 0.5 - 0.01 = 0.49$

$\Rightarrow \quad z_1 = 2.33$ | See table

$\Rightarrow \quad \dfrac{r - 662}{32} = 2.33$

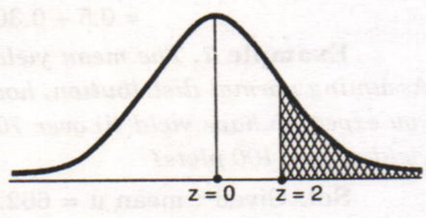
$\Rightarrow \quad r = 662 + 32\ (2.33) = 662 + 74.56 = 736.$

Example 8. *The income distribution of workers in a certain factory was found to be normal with mean of Rs. 500 and standard deviation of Rs. 50. There were 228 persons getting above Rs. 600. How many persons were there in all ?*

Sol. Let X denote the income of the workers. Given $\mu = 500$, $\sigma = 50$

Let $Z = \dfrac{X - \mu}{\sigma}$ be a standard normal variate.

When $X = 600$, $\quad Z = \dfrac{600 - 500}{50} = 2$



The probability of the persons getting above Rs. 600

$= P(X > 600) = P(Z > 2)$
$= 0.5 - P(0 < Z < 2)$
$= 0.5 - 0.4772$ | Using normal table
$= 0.0228$

Now, there are 228 persons getting salary above Rs. 600. Therefore total number of persons $= \dfrac{228}{0.0228} = 10{,}000.$

NOTES

Example 9. *The mean inside diameter of a sample of 500 washers produced by a machine is 5.02 mm. and the standard deviation is 0.05 mm. The purpose for which these washers are intended allows a maximum tolerance in the diameter of 4.96 to 5.08 mm, otherwise the washers are considered defective. Determine the percentage of defective washers produced by the machine, assuming the diameters are normally distributed.*

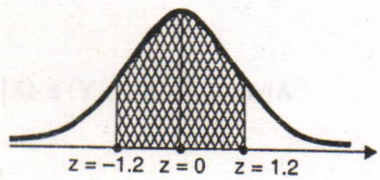
Sol. Given sample mean $\mu = 5.02$ mm

Standard deviation $\sigma = 0.05$ mm

Let $Z = \dfrac{X - \mu}{\sigma}$ be a standard normal variate.

When $X = 4.96$, $Z = \dfrac{4.96 - 5.02}{0.05} = -1.2$

When $X = 5.08$, $Z = \dfrac{5.08 - 5.02}{0.05} = 1.2$



Probability of non-defective washers

$$= P(4.96 < X < 5.08)$$
$$= P(-1.2 < Z < 1.2)$$
$$= P(-1.2 < Z < 0) + P(0 < Z < 1.2)$$
$$= P(0 < Z < 1.2) + P(0 < Z < 1.2) \qquad | \text{ Due to symmetry}$$
$$= 2P(0 < Z < 1.2)$$
$$= 2(0.3849) = 0.7699 \cong 0.77 \qquad | \text{Using normal table)}$$

$\therefore$ Percentage of non-defective washers = 77%

$\therefore$ Required percentage of defective washers

$$= 100 - 77 = 23.$$

Example 10. *For a certain normal distribution, the first moment about 10 is 40 and the fourth moment about 50 is 48. Find the arithmetic mean and standard deviation of the distribution.*

Sol. If μ'_1 denotes the first moment about the point X = A, then

$$\text{A.M.} = \mu'_1 + A = \text{mean}$$

But given μ'_1 (about the point X = 10) = 40

$\therefore$ $\text{mean} = 40 + 10 = 50$

Also μ'_4 (about X = 50) = 48

$\Rightarrow$ $\mu_4 = 48$

Also for a normal distribution, we know

$$\mu_4 = 3\sigma^4 \;\Rightarrow\; 48 = 3\sigma^4 \;\Rightarrow\; \sigma^4 = 16 \;\Rightarrow\; \sigma = 2.$$

Example 11. *If X and Y are independent normal variates with means 6, 7 and variance 9.16 respectively. Find the value of λ such that*

$$P(2X + Y \le \lambda) = P(4X - 3Y \ge 4\lambda).$$

Sol. Given $X \sim N(6, 9)$

$Y \sim N(7, 16)$

$\therefore \quad 2X + Y \sim N(2 \times 6 + 7, 4 \times 9 + 16) = N(19, 52)$

$4X - 3Y \sim N(4 \times 6 + (-3) \times 7, (16 \times 9 + 9 \times 16)$

$= N(3, 288)$ | See theorem

Now $P(2X + Y \le \lambda) = P(U \le \lambda)$ where

| $U = 2X + Y \sim N(19, 52)$

$$= P\left(Z \le \frac{\lambda - 19}{\sqrt{52}}\right) \text{ where } z \sim N(0, 1)$$

| Take $Z = \dfrac{U - 19}{\sqrt{52}}$

| When $U = \lambda$, $Z = \dfrac{\lambda - 19}{\sqrt{52}}$

Also $P[(4X - 3Y) \ge 4\lambda] = P(V \ge 4\lambda)$

$$= P\left(Z \ge \frac{4\lambda - 3}{12\sqrt{2}}\right) \text{ where } Z \sim N(0, 1)$$

| Where $V = 4X - 3Y \sim N(3, 288)$

| Take $Z = \dfrac{V - 3}{\sqrt{288}}$

| When $V = 4\lambda$, $Z = \dfrac{4\lambda - 3}{\sqrt{288}} = \dfrac{4\lambda - 3}{12\sqrt{2}}$

Now it is given that

$$P[(2X + Y) \le \lambda)] = P[(4X - 3Y) \ge 4\lambda]$$

$$\Rightarrow \quad P\left(Z \le \frac{\lambda - 19}{\sqrt{52}}\right) = P\left(Z \ge \frac{4\lambda - 3}{12\sqrt{2}}\right)$$

$$\Rightarrow \quad \frac{\lambda - 19}{\sqrt{52}} = -\left(\frac{4\lambda - 3}{12\sqrt{2}}\right)$$

$$\Rightarrow \quad \frac{\lambda - 19}{2\sqrt{13}} = \frac{-(4\lambda - 3)}{12\sqrt{2}}$$

| $P(Z \le a) = P(Z \ge b)$
| $\Rightarrow \quad a = -b$

$$\Rightarrow \quad (6\sqrt{2} + 4\sqrt{13})\lambda = 114\sqrt{2} + 3\sqrt{13}$$

$$\Rightarrow \quad \lambda = \frac{114\sqrt{2} + 3\sqrt{13}}{6\sqrt{2} + 4\sqrt{13}}$$

5.13 LOG-NORMAL RANDOM VARIABLE

If X is a normal variable with mean μ and variance σ^2, then $Y = e^X$ is called a log-normal random variable where $X = \log Y$ is a normal random variable.

Theorem VII. *Find the rth moment about origin i.e., μ'_r for the log-normal distribution.*

Proof. By definition, if log X is a log-normal variable, then

$$\mu'_r = E(X^r) = E(e^{rY})$$
$$= M_Y(r) \qquad | \text{ Log } X = Y \Rightarrow X = e^Y$$
$$= \exp\left(\mu r + \frac{1}{2} r^2\sigma^2\right) \qquad | \text{ m.g.f of } Y \sim N(\mu, \sigma^2)$$

NOTES

5.14 LOG-NORMAL DISTRIBUTION

A random variable X (X > 0) is said to have a log-normal distribution if $\log_e X$ is normally distributed.

Theorem VIII. To find the probability density function of the log-normal distribution.

Proof. Take $\log_e X \sim N(\mu, \sigma^2)$. For $x > 0$, we have

$$F_X(x) = P(X \le x) = P(\log_e X \le \log_e x) = P(Y \le \log_e x)$$
$$| \text{ log X is monotonic increasing function.)}$$
$$= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\log x} \exp\{-(y-\mu)^2 / 2\sigma^2\}\, dy \qquad | \; Y \sim N(\mu, \sigma^2)]$$
$$= \frac{1}{\sigma\sqrt{2\pi}} \int_0^x \exp\{-\log u - \mu)^2 / 2\sigma^2\} \frac{du}{u},$$

$$\left| \begin{array}{l} y = \log u \\ \Rightarrow \quad dy = \dfrac{1}{\mu}\, du \\ \text{When } y = \log x,\ u = x. \\ \text{When } y \to -\infty,\ u = e^{-\infty} = 0 \end{array} \right.$$

For $x \le 0$, $F_X(x) = P(X \le x) = 0$, because X is a positive random variable.

Define
$$f_X(u) = \begin{cases} \dfrac{1}{\sigma\sqrt{2\pi}} \cdot \dfrac{1}{u} \exp\left[\dfrac{-(\log u - \mu)^2}{2\sigma^2}\right], & u > 0 \\ 0, & u \le 0 \end{cases}$$

Then
$$F_X(x) = \int_{-\infty}^{x} f_X(u)\, du, \text{ where } f_X(u)$$

is the probability density function of X.

5.15 LAPLACE DOUBLE EXPONENTIAL DISTRIBUTION OR STANDARD LAPLACE DISTRIBUTION

A continuous random variable X is said to follow standard Laplace distribution if its probability density function is given by

$$f(x) = \frac{1}{2} e^{-|x|}, -\infty < x < \infty$$

NOTES

Theorem IX. *To find the characteristic function of the standard Laplace distribution.*

Proof. Let $\varphi_X(t)$ denotes the characteristic function of the standard Laplace distribution then, by definition

$$\varphi_X(t) = \int_{-\infty}^{\infty} e^{itx} f(x)\, dx = \frac{1}{2}\int_{-\infty}^{\infty} e^{itx}\, e^{-|x|}\, dx$$

$$= \frac{1}{2}\left[\int_{-\infty}^{\infty} \cos t\, tx \,.\, e^{-|x|}\, dx + i\int_{-\infty}^{\infty} \sin tx\, e^{-|x|} dx\right]$$

$$= \frac{1}{2}\,.\,2\int_{0}^{\infty} \cos tx \,.\, e^{-|x|}\, dx, \qquad \left|\begin{array}{l} \text{Sin } tx\,.\,e^{-|x|} \text{ is an odd function} \\ \therefore \int_{-\infty}^{\infty} \sin tx\,.\,e^{-|x|}\, dx = 0 \end{array}\right.$$

$$\therefore \quad \varphi_X(t) = \int_{0}^{\infty} e^{-x} \cos tx\, dx = 1 - t^2 \int_{0}^{\infty} e^{-x} \cos tx\, dx$$

(On integration by parts)

$$= 1 - t^2\, \varphi_X(t)$$

$$\Rightarrow \quad \varphi_X(t) = \frac{1}{1+t^2}.$$

Theorem X. *Find the mean and variance of the standard Laplace distribution and also obtain the following.*

$$\mu_3 = 0, \mu_4 = 36, \beta_1 = 0, \beta_2 = 9$$

Proof. By using the above theorem, the characteristic function of the standard Laplace distribution is given by

$$\varphi_X(t) = \frac{1}{1+t^2} = (1 + t^2)^{-1}$$

$$= 1 - t^2 + t^4 - t^6 + t^8 - \ldots\ldots \infty$$

$$= 1 + 2\,.\,\frac{(it)^2}{2} + 4\,!\,.\,\frac{(it)^4}{4\,!} + \ldots\ldots \infty$$

Here $\quad k_1 = k_3 = 0 \,;\, k_2 = 2,\, k_4 = 4\,! = 24$

$\therefore \quad$ mean $= k_1 = 0$, variance $= \mu_2 = k_2 = 2$

$$\mu_3 = k_3 = 0,\ \mu_4 = k_4 + 3k_2^{\,2} = 24 + 12 = 36$$

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = 0,\ \beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{36}{4} = 9.$$

Two Parameter Laplace Distribution

A continuous random variable X is said to have Laplace distribution with two parameters λ and μ if its probability density function is given by

$$f(x, \mu, \sigma) = \frac{1}{2\lambda} e^{-\frac{|x-\mu|}{\lambda}}, \quad -\infty < x < \infty, \lambda > 0$$

If X is a continuous random variable following Laplace distribution with two parameters λ and μ, then we write X ~ Lap (λ, μ).

Theorem XI. *To find the probability density function of the standard Laplace distribution by using two parameters Laplace distribution.*

Proof. Let X ~ Lap (λ, μ) and

$$Z = \frac{X - \mu}{\lambda} \quad \Rightarrow \quad X = \mu + \lambda Z$$

The probability density function of Z is given by

$$g(Z) = f(x) \cdot \left| \frac{dx}{dZ} \right| = \frac{1}{2\lambda} \cdot e^{-|Z|} \cdot \lambda$$

$$= \frac{1}{2} e^{-|Z|}, \; -\infty < Z < \infty$$

which is the *p.d.f* of standard Laplace distribution *i.e.,* If X ~ Lap (λ, μ), then

$$Z = \frac{X - \mu}{\lambda\sigma} \sim \text{Lap}(1, 0)$$

Theorem XII. *To find the characteristic function of the Laplace distribution with two parameters.*

Proof. Let X ~ Lap (λ, μ) and $\varphi_X(t)$ is the characteristic function of X, then

$$\varphi_X(t) = E(e^{itX}) = E\,[e^{it(\mu+\lambda Z)}], \quad \text{where } Z = \frac{X-\mu}{\lambda} \sim \text{Lap}(1, 0)$$

$$= e^{it\mu} \cdot E\,(e^{i.\lambda t.Z}) = e^{it\mu} \cdot \varphi_Z(\lambda t)$$

$$= \frac{e^{it\mu}}{1 + \lambda^2 t^2}$$

$$\left[\because \; Z \text{ is standard Laplace variate with } \varphi(t) = \frac{1}{1+t^2}.\right]$$

Theorem XIII. *Find* μ'_1, μ'_2 *and* μ_2 *for the two parameters Laplace distribution.*

Proof. Let X ~ Lap (λ, μ), then *r*th moment about origin is given by

$$\mu'_r = E(X^r) = \frac{1}{2\lambda} \int_{-\infty}^{\infty} x^r \exp\left(\frac{-|x-\mu|}{\lambda}\right) dx$$

$$= \frac{1}{2} \int_{-\infty}^{\infty} (Z\lambda - \mu)^r \exp(-|Z|)\, dz, \qquad \left| Z = \frac{x-\mu}{\lambda} \right.$$

$$= \frac{1}{2} \int_{-\infty}^{\infty} \left[\sum_{k=0}^{r} {}^rC_k\, (Z\lambda)^k\, \mu^{r-k} \right] \exp(-|Z|)\, dz$$

$$= \frac{1}{2} \sum_{k=0}^{r} \left[{}^rC_k\, \lambda^k \mu^{r-k} \int_{-\infty}^{\infty} z^k \exp(-|z|)\, dz \right]$$

$$= \frac{1}{2} \sum_{k=0}^{r} \left[{}^rC_k\, \lambda^k \mu^{r-k} \left\{ \int_{-\infty}^{0} z^k e^{-|z|}\, dz + \int_{0}^{\infty} z^k\, e^{-|z|}\, dz \right\} \right]$$

$$= \frac{1}{2} \sum_{k=0}^{r} \left[{}^rC_k\, \lambda^k \mu^{r-k} \left\{ (-1)^k \int_{0}^{\infty} e^{-z}\, z^k\, dz + \int_{0}^{\infty} e^{-z}\, z^k\, dz \right\} \right]$$

$$= \frac{1}{2} \sum_{k=0}^{r} \left[{}^rC_k\, \lambda^k \mu^{r-k}\, \Gamma(k+1) \{(-1)^k + 1\} \right]$$

$$\Rightarrow \qquad \mu'_r = \frac{1}{2}\sum_{k=0}^{r}\left[{}^rC_k\,\lambda^k\mu^{r-k}\,k!\,\{1+(-1)^k\}\right]$$

$\therefore$ Mean $= \mu'_1 = \mu$, $\mu'_2 = \mu^2 + 2\lambda^2$ and variance $= \mu_2 = \mu'_2 - \mu_1'^2 = 2\lambda^2$.

NOTES

5.16 WEIBUL VARIABLE

A random variable X is said to be a Weibul variable with three parameters $C > 0$, $\alpha > 0$ and μ, if $Y = \left(\frac{X-\mu}{\alpha}\right)^C$ has the exponential distribution with *p.d.f.*

$$P_Y(y) = e^{-y},\ y > 0$$

5.17 WEIBUL DISTRIBUTION

A continuous random variable X is said to follow Weibul distribution with three parameters C, α and μ if its probability density function is given by

$$f(x\,;\,C,\alpha,\mu) = \frac{C}{\alpha}\left(\frac{x-\mu}{\alpha}\right)^{C-1} e^{-\left(\frac{x-\mu}{\alpha}\right)^C}, \qquad \ldots(1)$$

$$C > 0,\ \alpha > 0,\ x > \mu$$

6.18 STANDARD WEIBUL DISTRIBUTION

If $\alpha = 1$, $\mu = 0$ in eq. (1) we get, the *p.d.f.* of standard Weibul distribution,

$$f(x, C) = Cx^{C-1}\,e^{-x^C},\ x > 0,\ C > 0.$$

Theorem XIV. *To find the mean and variance of standard Weibul distribution.*

Proof. The standard Weibul variate is given by $Y = \left(\frac{X-\mu}{\alpha}\right)^C$ where $\alpha = 1$, $\mu = 0$

or $\qquad Y = X^C$ where Y has the exponential distribution with *p.d.f* given by

$$P_Y(y) = e^{-y},\ y > 0$$

The rth moment μ'_r is given by

$$\mu'_r = E(X^r) = E(Y^{1/C})^r$$

$$= E\,(Y^{r/C}) = \int_0^\infty e^{-y}\,y^{r/C}\,dy$$

$$= \Gamma\left(\frac{r}{C}+1\right) \qquad \left|\ \Gamma(n) = \int_0^\infty e^{-x}x^{n-1}\,dx\right.$$

Put $r = 1$, $\quad \mu'_1 = \text{mean} = E(X) = \Gamma\left(\frac{1}{C}+1\right)$

Variance (X) $= E\,(X^2) - (E\,(X))^2$

$$= \Gamma\left(\frac{2}{C}+1\right) - \left(\Gamma\left(\frac{1}{C}+1\right)\right)^2$$

5.19 LOGISTIC DISTRIBUTION

NOTES

A continuous random variable X is said to follow logistic distribution with parameters α and β, if its distribution function is of the form

$$F_X(x) = \left[1 + e^{-\frac{x-\alpha}{\beta}}\right]^{-1}, \ \beta > 0$$

$$= \frac{1}{2}\left[1 + \tanh \frac{1}{2}\left(\frac{x-\alpha}{\beta}\right)\right], \ \beta > 0.$$

Theorem XV. *Find the probability density function of logistic distributions with parameters α and β (β > 0).*

Proof. The distribution function F(x) of logistic distribution is given by

$$F(x) = \left[1 + \exp P\left\{-\left(\frac{x-\alpha}{\beta}\right)\right\}\right]^{-1}$$

∴ The probability density function $f(x)$ of the logistic distribution is given by

$$f(x) = \frac{d}{dx}[F(x)] = \frac{d}{dx}\left[1 + \exp\left\{-\left(\frac{x-\alpha}{\beta}\right)\right\}\right]^{-1}$$

$$= \frac{1}{\beta}\left[1 + \exp\left\{-\left(\frac{x-\alpha}{\beta}\right)\right\}\right]^{-2} \exp\left[-\left(\frac{x-\alpha}{\beta}\right)\right] \qquad ...(1)$$

Consider $\tanh x = \dfrac{e^x - e^{-x}}{e^x + e^{-x}} = \dfrac{1 - e^{-2x}}{1 + e^{-2x}}$

∴ $\quad 1 + \tanh x = 1 + \dfrac{1 - e^{-2x}}{1 + e^{-2x}} = \dfrac{2}{1 + e^{-2x}}$

⇒ $\dfrac{1}{2}(1 + \tanh x) = (1 + e^{-2x})^{-1}$

Also $\quad 1 + e^{-2x} = \dfrac{2}{1 + \tanh x}$

⇒ $\quad e^{-2x} = \dfrac{2}{1 + \tanh x} - 1 = \dfrac{1 - \tanh x}{1 + \tanh x}$

∴ From (1)

$$f(x) = \frac{1}{\beta} \cdot \left(\frac{2}{1 + \tanh \frac{x-\alpha}{2\beta}}\right)^{-2} \left(\frac{1 - \tanh \frac{x-\alpha}{2\beta}}{1 + \tanh \frac{x-\alpha}{2\beta}}\right)$$

NOTES

$$= \frac{1}{4\beta}\left(1+\tanh\frac{x-\alpha}{2\beta}\right)^2\left(\frac{1-\tanh\frac{x-\alpha}{2\beta}}{1+\tanh\frac{x-\alpha}{2\beta}}\right)$$

$$= \frac{1}{4\beta}\left(1-\tanh^2\frac{x-\alpha}{2\beta}\right) = \frac{1}{4\beta}\operatorname{sech}^2\frac{x-\alpha}{2\beta}.$$

Theorem XVI. *To find the probability density function of the standard logistic variate.*

Proof. Let $\quad Y = \dfrac{X-\alpha}{\beta}$ where X is a logistic variable.

Let $g_Y(y)$ be the probability density function of the standard logistic variate Y, then

$$g_Y(y) = f(x)\,.\left|\frac{dx}{dy}\right| = e^{-y}(1+e^{-y})^{-2}, -\infty < y < \infty \qquad |\text{Using (1)}$$

$$= \frac{1}{4}\operatorname{sech}^2\frac{y}{2}, -\infty < y < \infty \qquad |\text{ Using (2)}$$

Theorem XVII. *Find the mean and variance of the standard logistic variable X with parameters α and β. Also find β_1 and β_2.*

Proof. We first find the moment generating function of the standard logistic variate Y where

$$Y = \frac{X-\alpha}{\beta} \text{ and X is a logistic variate}$$

Let $M_Y(t)$ denotes the moment generating function of Y, then

$$M_Y(t) = E(e^{tY}) = \int_{-\infty}^{\infty} e^{ty} g_Y(y)\,dy$$

$$= \int_{-\infty}^{\infty} e^{ty}e^{-y}(1+e^{-y})^{-2}\,dy = \int_{-\infty}^{\infty} e^{ty}e^{-y}\left(\frac{1+e^{y}}{e^{y}}\right)^{-2} dy$$

$$= \int_{-\infty}^{\infty} e^{ty}e^{y}(1+e^{y})^{-2}\,dy$$

Put $z = (1+e^y)^{-1} \quad \Rightarrow \quad 1+e^y = \dfrac{1}{z}$

$$\Rightarrow \qquad e^y = \frac{1}{z} - 1 = \frac{1-z}{z}$$

$$\Rightarrow \qquad e^y\,dy = \frac{z(-1)-(1-z)}{z^2}\,dz = \frac{-z-1+z}{z^2}\,dz = \frac{-1}{z^2}\,dz$$

When $\quad y = \infty, \quad z = 0$

$\qquad\quad y = -\infty, z = 1$

NOTES

$$= \int_1^0 \left(\frac{1-z}{z}\right)^t \cdot \left(1+\frac{1-z}{z}\right)^{-2} \cdot \left(\frac{-1}{z^2}\,dz\right)$$

$$= \int_0^1 \left(\frac{1-z}{z}\right)^t \cdot \frac{1}{z^{-2}} \cdot \frac{dz}{z^2}$$

$$= \int_0^1 z^{-t}\,(1-z)^t\,dz = \beta\,(1-t,\,1+t),\ 1-t>0$$

$$= \frac{\Gamma(1-t)\,\alpha(1+t)}{\Gamma(1-t+1+t)} = \frac{\Gamma(1-t)\,\alpha(1+t)}{\Gamma(2)}$$

$$= \Gamma\,(1-t)\,\alpha\,(1+t)$$

$$= \pi t \text{ cosec } \pi t,\ t<1 \qquad \text{...(3)}$$

But $\quad x \text{ cosec } x = \dfrac{x}{\sin x} = \dfrac{x}{x - \dfrac{x^3}{3\,!} + \dfrac{x^5}{5\,!} - \dfrac{x^7}{7\,!} + \ldots\ldots}$

$$= \frac{1}{1 - \dfrac{x^2}{3\,!} + \dfrac{x^4}{5\,!} - \dfrac{x^6}{7\,!} + \ldots\ldots}$$

$$= \left[1 - \left(\frac{x^2}{3\,!} - \frac{x^4}{5\,!} + \frac{x^6}{7\,!} \ldots\ldots\right)\right]^{-1}$$

$$= 1 + \left(\frac{x^2}{3\,!} - \frac{x^4}{5\,!} + \frac{x^6}{7\,!} \ldots\ldots\right) + \left(\frac{x^2}{3\,!} - \frac{x^4}{5\,!} + \frac{x^6}{7\,!} \ldots\ldots\right)^2 + \ldots\ldots$$

$$= 1 + \frac{x^2}{6} + x^4\left(\frac{1}{36} - \frac{1}{120}\right) + \ldots\ldots$$

$$= 1 + \frac{x^2}{6} + \frac{7}{360}x^4 + \ldots\ldots$$

$\therefore$ From (3)

$$M_Y(t) = 1 + \frac{(\pi t)^2}{6} + \frac{7}{360}(\pi t)^4 + \ldots\ldots \qquad \text{...(4)}$$

Hence mean of Y = coefficient of t in (4) = 0

$$\mu_2 = E(Y^2) = \text{coefficient of } \frac{t^2}{2\,!} = \frac{\pi^2}{3}$$

$$\mu_3 = E(Y^3) = \text{coefficient of } \frac{t^3}{3\,!} = 0$$

$$\mu_4 = E(Y^4) = \text{coefficient of } \frac{t^4}{4\,!} = \frac{7}{15}\pi^4$$

$\therefore$ for standard logistic distribution,

$$\text{mean} = 0, \text{ variance } \mu_2 = \frac{\pi^2}{3}$$

NOTES

Also $$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = 0, \beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{7\pi^4}{15} \times \frac{9}{\pi^4} = \frac{63}{15} = 4.2.$$

SUMMARY

- The graph of the normal distribution is known as normal curve. It is bell-shaped and symmetrical about the line $x = \mu$.
- The normal distribution is a continuous destribution.
- The normal curve is unimodal. This means that the normal curve has a unique mode.
- The total area under the normal curve above x-axis 4 units.
- For a normal distribution, mean = median = mode.
- For standard Laplace distribution,

  $\mu_3 = 0, \mu_4 = 36, \beta_1 = 0, \beta_2 = 9.$
- For standard logistic distribution, mean is zero.

GLOSSARY

- **Standard Normal Variate.** Let X be a normal variable with mean μ and variance σ^2, then $Z = \dfrac{X - \mu}{\sigma}$ is known as standard normal variable with mean 0 and variance only.
- **Error Function.** The error function is defined by the following integral

  $$P(Z) = \frac{1}{\sqrt{2\pi}} \int_0^Z e^{-\frac{z^2}{2}} dz.$$

  It is also known as probability integral.
- **Lognormal Variate.** Let X be normal variable X with mean μ and variable σ^2. Then e^X is called a log normal variate.
- **Weibul Variate.** A random variable X is said to be a Weibul variate with three parameters $C > 0$, $\alpha > 0$ and μ. If $Y = \left(\dfrac{X - \mu}{\alpha}\right)^c$ has the exponential distribution with p.d.f. $P_y(y) = e^{-y}, y > 0$.

REVIEW QUESTIONS

1. The mean height of 500 male students in a certain college is 151 cm and the standard deviation is 15 cm. Assuming the heights are normally distributed, find how many students have heights between 120 and 155 cm ?
2. Students of a class were given a mechanical aptitude test. This marks were found to be normally distributed with mean 60 and standard deviation 5. What per cent of students scored

NOTES

(i) more than 60 marks ? (ii) less than 56 marks ?

(iii) between 45 and 65 marks ?

3. In an examination taken by 500 candidates, the average and the standard deviation of marks obtained (normally distributed) are 40% and 10%. Find approximately

(i) how many will pass, if 50% is fixed as a minimum ?

(ii) what should be the minimum if 350 candidates are to pass ?

(iii) how many have scored marks above 60% ?

4. In a certain examination, the percentage of passes and distinction were 46 and 9 respectively. Estimate the average marks obtained by the candidate, the minimum pass and distinction marks, being 40 and 75 respectively. (Assume the distribution of marks to be normal).

5. Suppose the waist measurements X of 800 girls are normally distributed with mean 66 cm and standard deviation 5 cm. Find the number of girls with waist

(i) between 65 cm and 70 cm.

(ii) greater than or equal to 72 cm.

6. Write short note on normal distribution.

7. Distinguish between Binomial and normal distribution.

8. Distinguish between Poisson and normal distribution.

9. Show that, for a normal distribution, the mean deviation about mean is $\frac{4}{5}$ times the standard deviation.

10. For a normal distribution, $\beta_1 = 0, \beta_2 = 3$.

11. Show that, the moment generators function for the standard normal variate Z is given by $\mu_Z(t) = e^{\frac{t^2}{2}}$.

12. Write five applications of normal distribution.

13. For the standerd Laplace distribution show that $\beta_1 = 0, \beta_2 = 0$.

14. For the standard logistic distribution, $\beta_1 = 0, \beta_2 = 4.2$.

FURTHER READING

1. Introduction to Probability theory with applications: W.Feller, Vol-1: Wiley astern
2. continuous Univarate distribution-2: N.L. Johnson and S.Kotz, John Wiley and Sons
3. Introduction to Modern Probability Theory: B.R. Bhat: Wiley Eastern.

NOTES

UNIT–V

CHAPTER 6

SAMPLING DISTRIBUTIONS

OBJECTIVES

After going through this chapter, you should be able to:

- know about sampling distributions like X^2, t and F.
- know about central and non-central distributions like X^2, t and F.
- conditions for applying X^2, t and F test.
- order statistics of single and two or more order.

STRUCTURE

6.1 INTRODUCTION

In this chapter, we will discuss central distributions of the central statistics χ^2, t and F as well as non-central distributions of the non-central statistics χ^2, t and F respectively.

Central chi-square distribution. The square of a standard normal variate is known as chi-square variate with 1 degree of freedom (d.f.), *i.e.*, If $X \sim N(\mu, \sigma^2)$, then we know that the standard normal variate is given by

$$Z = \frac{X - \mu}{\sigma} \sim N(0, 1) \text{ and its square}$$

$$Z^2 = \left(\frac{X-\mu}{\sigma}\right)^2 \text{ is known as chi-square variate with 1 d.f.}$$

NOTES

Generally, if $X_1, X_2, X_3, \ldots\ldots, X_n$ are n independent normal variates with mean $\mu_1, \mu_2, \ldots\ldots, \mu_n$ and variances $\sigma_1^2, \sigma_2^2, \ldots\ldots, \sigma_n^2$, then

$$\chi^2 = \sum_{i=1}^{n} \frac{(X_i - \mu_i)^2}{\sigma_i^2}, \text{ is a chi-square variate with } n \text{ degree of freedom (d.f.)}$$

Theorem I. *Using the method of moment generating function, derive the chi-square distribution.*

Proof. Let $X_1, X_2, \ldots\ldots, X_n$ are n independent normal variates, *i.e.*, $X_i \sim N(\mu_i, \sigma_i^2)$, $i = 1, 2, \ldots\ldots, n$.

Then, by definition,

$$\chi^2 = \sum_{i=1}^{n} \left(\frac{X_i - \mu_i}{\sigma_i}\right)^2 = \sum_{i=1}^{n} u_i^2 \text{ where } u_i = \frac{X_i - \mu_i}{\sigma_i} \sim N(0, 1)$$

Given X_i's are independent, it follows u_i's are also independent. Therefore,

$$M_{\chi^2}(t) = M_{\Sigma u_i^2}(t) = \prod_{i=1}^{n} M_{u_i^2}(t) = [M_{u_i^2}(t)]^n,$$

Now $\quad M_{u_i^2}(t) = E[\exp(tu_i^2)] = \int_{-\infty}^{\infty} \exp(tu_i^2) f(x_i)\, dx_i$

$$= \int_{-\infty}^{\infty} \exp(tu_i^2) \frac{1}{\sigma\sqrt{2\pi}} \exp\{-(x_i - \mu)^2/2\sigma^2\}\, dx_i$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp(tu_i^2) \exp(-u_i^2/2)\, du_i, \qquad \left| \; u_i = \frac{x_i - \mu}{\sigma} \right.$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left\{-\left(\frac{1-2t}{2}\right)u_i^2\right\} du_i$$

$$= \frac{1}{\sqrt{2\pi}} \cdot \frac{\sqrt{\pi}}{\left(\frac{1-2t}{2}\right)^{1/2}} = (1-2t)^{-1/2} \qquad \left| \; \int_{-\infty}^{\infty} e^{-a^2x^2}\, dx = \frac{\sqrt{\pi}}{a} \right.$$

$$\therefore \quad M_{\chi^2}(t) = (1-2t)^{-n/2}, \qquad | \text{Using (1)}$$

which is the moment generating function of a Gamma variate with parameters $\frac{1}{2}$ and $\frac{1}{2}n$.

Hence, by uniqueness theorem of moment generating functions,

$$\chi^2 = \sum_{i}^{n} \left(\frac{x_i - \mu_i}{\sigma_i}\right)^2, \text{ is a Gamma variate with parameters } \frac{1}{2} \text{ and } \frac{1}{2}n.$$

NOTES

Therefore, we can write

$$dP(\chi^2) = \left(\frac{1}{2}\right)^{\frac{n}{2}} e^{-\frac{1}{2}\chi^2} (\chi^2)^{\frac{n}{2}-1} d\chi^2$$

$$= \frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)} e^{-\frac{1}{2}\chi^2} (\chi^2)^{\frac{n}{2}-1} d\chi^2, 0 \le \chi^2 < \infty$$

Remark. If a random variable X has a χ^2 distribution with n degree of freedom, we write $X \sim \chi^2(n)$ and its probability density function (p.d.f) is given as

$$f(x) = \frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)} e^{-\frac{x}{2}} x^{\frac{n}{2}-1}, \quad 0 \le x < \infty$$

Theorem II. *If X is a chi-square variate with n degree of freedom, then*

$$\frac{1}{2}X \sim r\left(\frac{1}{2}n\right)$$

Proof. Take $Y = \frac{1}{2}X$, then the probability density function of Y is given by

$$g(y) = f(x) \cdot \left|\frac{dx}{dy}\right| = \frac{1}{2^{n/2}\,\Gamma(n/2)} e^{-y} \cdot (2y)^{(n/2)-1} \cdot 2$$

$$= \frac{1}{\Gamma(n/2)} e^{-y} y^{(n/2)-1}; 0 \le y < \infty$$

$$\therefore \qquad Y = \frac{1}{2}X \sim \gamma\left(\frac{1}{2}n\right)$$

Hence the theorem.

Moment Generating Function

Theorem III. *Find the moment generating function of the chi-square distribution.*

Proof. Let X is a chi-square variate with n degree of freedom. If $M_X(t)$ denotes the moment generating function of X, then

Let $X \sim \chi^2_{(n)}$, then

$$M_X(t) = E(e^{tX}) = \int_0^\infty e^{tx} f(x)\, dx = \frac{1}{2^{n/2}\,\Gamma(n/2)} \int_0^\infty e^{tx} \cdot e^{-x/2} x^{(n/2)-1} dx$$

$$= \frac{1}{2^{n/2}\,\Gamma(n/2)} \int_0^\infty \exp\left[-\left(\frac{1-2t}{2}\right)x\right] \cdot x^{n/2-1} dx$$

$$= \frac{1}{2^{n/2}\,\Gamma(n/2)} \frac{\Gamma(n/2)}{[(1-2t)/2]^{n/2}} \qquad | \text{ Gamma Integral}$$

$$= (1-2t)^{-n/2}, \ |\ 2t\ | < 1$$

is the required moment generating function of the χ^2-variate with n degree of freedom.

Cor. $\mu_r' = n(n+2)(n+4) \ldots\ldots (n+2r-2)$

Proof. From above theorem,

$$M_X(t) = (1-2t)^{-\frac{n}{2}}$$

$$= 1 + \frac{n}{2}(2t) + \frac{\frac{n}{2}\left(\frac{n}{2}+1\right)}{2!}(2t)^2 + \ldots\ldots$$

$$+ \frac{\frac{n}{2}\left(\frac{n}{2}+1\right)\left(\frac{n}{2}+2\right)\ldots\ldots\left(\frac{n}{2}+r-1\right)}{r!}(2t)^r + \ldots\ldots$$

$\therefore$ $\mu_r' = $ Coefficient of $\frac{t^r}{r!}$ in the expansion of $M_X(t)$

$$= 2^r \frac{n}{2}\left(\frac{n}{2}+1\right)\left(\frac{n}{2}+2\right) \ldots\ldots \left(\frac{n}{2}+r-1\right)$$

$$= n(n+2)(n+4) \ldots\ldots (n+2r-2).$$

Hence proved

Theorem IV. *Find the cumulant generating function of χ^2-distribution.*

Proof. Let X is a chi-square variate with n degree of freedom, and $K_X(t)$ denotes the cumulant generating function of χ^2, then, we know

$$M_X(t) = (1-2t)^{-\frac{n}{2}}$$

$$\therefore \quad K_X(t) = \log M_X(t) = -\frac{n}{2}\log(1-2t) = \frac{n}{2}\left[2t + \frac{(2t)^2}{2} + \frac{(2t)^3}{3} + \frac{(2t)^4}{4} + \ldots\right]$$

$\therefore$ $\kappa_1 = $ Coefficient of t in $K(t) = n$, $\kappa_2 = $ Coefficient of $\frac{t^2}{2!}$ in $K(t) = 2n$,

$\kappa_3 = $ Coefficient of $\frac{t^3}{3!}$ in $K(t) = 8n$, and $\kappa_4 = $ Coefficient of $\frac{t^4}{4!}$ in $K(t) = 48n$

In general, $\kappa_r = $ Coefficient of $\frac{t^r}{r!}$ in $K(t) = n\, 2^{r-1}(r-1)\,!$ ...(1)

Putting $r = 1, 2, 3, 4$ in (1), we get

$$\text{Mean} = \kappa_1 = n, \quad \text{Variance} = \mu_2 = \kappa_2 = 2n$$

$$\mu_3 = \kappa_3 = 8n, \quad \mu_4 = \kappa_4 + 3\kappa_2^{\,2} = 48n + 12n^2$$

$$\beta_1 = \frac{\mu_3^{\,2}}{\mu_2^{\,3}} = \frac{8}{n} \quad \text{and} \quad \beta_2 = \frac{\mu_4}{\mu_2^{\,2}} = \frac{12}{n} + 3$$

Theorem V. *For large value of n, where n is the degree of freedom, χ^2 distribution tends to normal distribution.*

NOTES

Proof. Let X is a chi-square variate with n degree of freedom, then we know

$$M_X(t) = (1-2t)^{-\frac{n}{2}}, \ | \ 2t \ | < 1$$

Let $$Z = \frac{X-\mu}{\sigma}, \text{ then}$$

The moment generating function of Z, the standard χ^2-variate, is given as

$$M_Z(t) = M_{\frac{X-\mu}{\sigma}}(t) = e^{-\frac{\mu t}{\sigma}} M_X\left(\frac{t}{\sigma}\right)$$

$$= e^{-\mu t/\sigma}(1-2t/\sigma)^{-n/2} = e^{-nt/\sqrt{2n}}\left(1-\frac{2t}{\sqrt{2n}}\right)^{-n/2}$$

$$| \ \mu = n, \sigma^2 = 2n$$

$$\therefore \quad K_Z(t) = \log M_Z(t) = -t\sqrt{\frac{n}{2}} - \frac{n}{2}\log\left(1-t\sqrt{\frac{2}{n}}\right)$$

$$= -t\sqrt{\frac{n}{2}} + \frac{n}{2}\left[t\cdot\sqrt{\frac{2}{n}} + \frac{t^2}{2}\cdot\frac{2}{n} + \frac{t^3}{3}\left(\frac{2}{n}\right)^{3/2} + \ldots\ldots\right]$$

$$= -t\sqrt{\frac{n}{2}} + t\cdot\sqrt{\frac{n}{2}} + \frac{t^2}{2} + O(n^{-1/2}) = \frac{t^2}{2} + O(n^{-1/2}),$$

where $O(n^{-1/2})$ denotes the terms containing $n^{1/2}$ and higher powers of n in the denominator.

$$\therefore \quad \lim_{n\to\infty} K_Z(t) = \frac{t^2}{2} \quad \Rightarrow \quad M_Z(t) = e^{t^2/2} \text{ as } n \to \infty,$$

which is the moment generating function of a standard normal variate. Hence, by uniqueness theorem of moment generating function Z is asymptotically normal. In other words, standard χ^2 variate tends to standard normal variate as $n \to \infty$. Thus, χ^2, distribution tends to normal distribution for large degree of freedom.

Characteristic Function

Theorem VI. *Find the characteristic function of the χ^2-distribution.*

Proof. Let X is a chi-square variate and $\phi_X(t)$ denotes the characteristic function of χ^2, then

$$\phi_X(t) = E\{\exp(it X)\} = \int_0^\infty \exp(itx) f(x)\, dx$$

$$= \frac{1}{2^{n/2}\,\Gamma(n/2)} \int_0^\infty \exp\left\{-\left(\frac{1-2it}{2}\right)x\right\} (x)^{\frac{n}{2}-1} dx = (1-2it)^{-n/2}$$

is the required characteristic function of χ^2.

Theorem VII. *Prove that if X is a chi-square variate with n degree of freedom, then*

(i) mode of χ^2 is $n-2$

(ii) For $n \geq 1$, χ^2-distribution is positively skewed.

NOTES

Proof. (*i*) We know that if X is a chi-square variate with n degree of freedom, then,

$$f(x) = \frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)} e^{-\frac{x}{2}} x^{\frac{n}{2}-1}, 0 \le x < \infty \qquad ...(1)$$

Now mode of χ^2 is the solution of $f'(x) = 0$ and $f''(x) < 0$.

Taking logarithm of (1), we have

$$\log f(x) = -\log 2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right) + \log e^{-\frac{x}{2}} + \log x^{\frac{n}{2}-1}$$

Differentiating w.r.t x, we get

$$\frac{f'(x)}{f(x)} = 0 - \frac{1}{2} + \left(\frac{n}{2} - 1\right) \cdot \frac{1}{x} = \frac{n-2-x}{x}$$

$$\Rightarrow \qquad f'(x) = \frac{n-2-x}{2x} f(x).$$

For mode, $f'(x) = 0$ gives

$$\frac{n-2-x}{2x} f(x) = 0 \ \Rightarrow \ n-2-x = 0 \qquad | f(x) \ne 0$$

$$\Rightarrow \qquad x = n-2$$

Also $\qquad f'(x) = \frac{d}{dx}(f'(x)) = \frac{d}{dx}\left(\frac{n-2-x}{x} f(x)\right)$

$$= \frac{n-2-x}{x} f'(x) + \frac{x(-1)-(n-2-x)}{x^2} f(x)$$

$$\Rightarrow \qquad (f''(x))_{x=n-2} = 0 + \frac{-(n-2)}{(n-2)^2}(f(x))_{n-2} = -\frac{(f(x))_{x=n-2}}{n-2} < 0$$

$\therefore$ mode of χ^2-distribution is $n - 2$.

(*ii*) Karl Pearson's coefficient of skewness is given by :

$$\text{Skewness} = \frac{\text{Mean} - \text{Mode}}{\text{S.D.}} = \frac{n-(n-2)}{\sqrt{2n}} = \sqrt{\frac{2}{n}}$$

Since Pearson's coefficient of skewness is greater than zero for $n \ge 1$, the χ^2-distribution is positively skewed.

Theorem VIII. *Reproductive property or Additive property of independent χ^2-variates.*

If $X_1, X_2, \ldots\ldots X_{2k}$ are n independent χ^2-variates with $n_1, n_2, \ldots\ldots n_k$ degree of freedom, then $X_1 + X_2 + \ldots\ldots X_k$ is also a chi-square variate with $(n_1 + n_2 + \ldots\ldots + n_k)$ degree of freedom.

Proof. Here X_i is a chi-square variate and if $M_{X_i}(t)$ is the moment generating function of X_i, then we have $\quad M_{X_i}(t) = (1-2t)^{-n_i/2}$; $i = 1, 2, \ldots\ldots, k$.

NOTES

The moment generating function of the sum $\sum_{i=1}^{k} X_i$ is given by:

$$M_{\Sigma X_i}(t) = M_{X_1}(t)\, M_{X_2}(t) \,......\, M_{X_k}(t) \qquad [\because \; X_i\text{'s are independent}]$$

$$(1-2t)^{-n_1/2}(1-2t)^{-n_2/2} \ldots (1-2t)^{-n_k/2} = (1-2t)^{-(n_1+n_2+\ldots+n_k)/2}$$

which is the moment of generating function of a χ^2-variate with $(n_1 + n_2 + \ldots.. + n_k)$ d.f. Hence by uniqueness theorem of moment generating functions,

$(X_1 + X_2 + \ldots... + X_k)$ is also a chi-square variate with $(n_1 + n_2 + \ldots... + n_k)$ degree of freedom.

Cor. *Converse of the above theorem is also true. We prove it for $n = 2$. i.e.,*

If X and Y are independent non-negative variates such that X + Y follows chi-square distribution with $n_1 + n_2$ degree of freedom and if one of them say X is a χ^2-variate with n_1 degree of freedom then the other, viz., Y is a χ^2-variate with n_2 degree of freedom.

Proof. Since X and Y are independent variates, $M_{X+Y}(t) = M_X(t)\, M_Y(t)$

$$\Rightarrow \qquad (1-2t)^{-(n_1+n_2)/2} = (1-2t)^{-n_1/2} \,.\, M_Y(t)$$

$$[\because \; X + Y \sim \chi^2_{(n_1+n_2)}] \text{ and } X - X \sim \chi^2_{(n_1)}]$$

$$\therefore \qquad M_Y(t) = (1-2t)^{-n_2/2},$$

which is the moment generating function of χ^2-variate with n_2 degree of freedom. Hence by uniqueness theorem of moment generating function $Y \sim \chi^2_{(n_2)}$ *i.e.*, Y is a chi-square variate with n_2 degree of freedom. Hence the theorem.

6.2 CHI-SQUARE PROBABILITY CURVE

Theorem IX. *Define chi-square probability curve and mention its properties.*

Proof. The graph of the chi-square distribution is known as chi-square probability curve. If X is a chi-square variate, with degree of freedom n, then we know

$$f(x) = \frac{1}{2^{\frac{n}{2}}\,\Gamma\left(\frac{n}{2}\right)} e^{-\frac{x}{2}} x^{\frac{n}{2}-1}, \; 0 \le x < \infty$$

Taking logarithm both sides,

$$\log f(x) = -\log 2^{\frac{n}{2}\propto\left(\frac{n}{2}\right)} + \log e^{-\frac{x}{2}} + \log x^{\frac{n}{2}-1}$$

$$= -\frac{n}{2}\log 2 - \log \Gamma\left(\frac{n}{2}\right) - \frac{x}{2} + \left(\frac{n}{2}-1\right)\log x$$

Differentiating w.r.t. x, we get

$$\frac{f'(x)}{f(x)} = 0 - \frac{1}{2} + \frac{\left(\frac{n}{2}-1\right)}{x}$$

$$\Rightarrow \qquad f'(x) = \frac{n-2-x}{2x} f(x) \qquad \text{...(1)}$$

Case I. When $n \le 2$. Since $x > 0$ and $f(x)$ being probability density function is always non-negative, we get from (1),

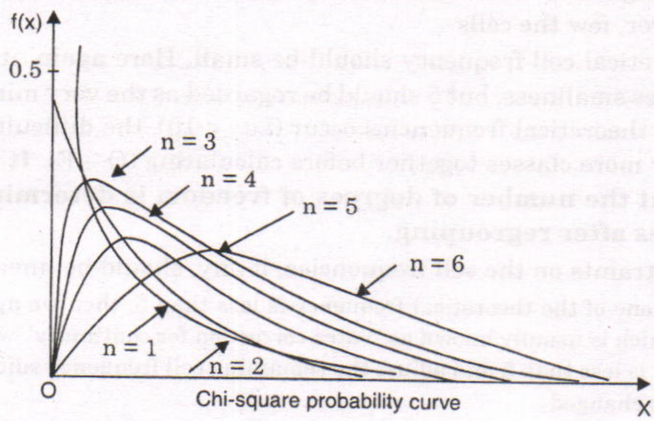
$$f'(x) < 0 \text{ if } (n-2) \le 0,$$

for all values of x. Thus the χ^2-probability curve for 1 and 2 degrees of freedom is monotonically decreasing.

Case II. When $n > 2$, then

$$f'(x) = \begin{cases} > 0, \text{ if } x < (n-2) \\ = 0, \text{ if } x = n-2 \\ < 0, \text{ if } x > (n-2) \end{cases}$$

This implies that for $n > 2$, $f(x)$ is monotonically increasing for $0 < x < (n-2)$ and monotonically decreasing for $(n-2) < x < \infty$, while at $x = n-2$, it attains the maximum value.



Chi-square probability curve

For $n \ge 1$, as x increases, $f(x)$ decreases rapidly and finally tends to zero as $x \to \infty$. Thus for $n > 1$, the χ^2-probability curve is positively skewed towards higher values of x. Moreover, x-axis is an asymptote to the curve. The shape of the curve for $n = 1, 2, 3, \ldots\ldots, 6$ is given in the figure. For $n = 2$, the curve will meet $y = f(x)$ axis at $x = 0$, *i.e.*, at $f(x) = 0.5$. For $n = 1$, it will be an inverted J-shaped curve.

Another form of Chi-square Distribution

If O_i and $E_i (i = 1, 2, \ldots, k)$, be a set of observed and expected frequencies, then

$$\chi^2 = \sum_{i=1}^{k} \left[\frac{(O_i - E_i)^2}{E_i} \right] \text{ where } \sum_{i=1}^{k} O_i = \sum_{i=1}^{k} E_i$$

follows chi-square distribution with $(k-1)$ *d.f.*

Another convenient form of this formula is as follows :

$$\chi^2 = \sum_{i=1}^{k} \left(\frac{O_i^2 + E_i^2 - 2O_iE_i}{E_i} \right) = \sum_{i=1}^{k} \left(\frac{O_i^2}{E_i} + E_i - 2O_i \right)$$

$$= \sum_{i=1}^{k} \frac{O_i^2}{E_i} + \sum_{i=1}^{k} E_i - 2\sum_{i=1}^{k} O_i = \sum_{i=1}^{k} \frac{O_i^2}{E_i} - N,$$

where $\sum_{i=1}^{k} O_i = \sum_{i=1}^{k} E_i = N$ (say), is the total frequency.

NOTES

Applications of χ^2 Test

χ^2 test is one of the simplest and the most general test known. It is applicable to a very large number of problems in practice which can be summed up under the following heads :

(*i*) as a test of goodness of fit.

(*ii*) as a test of independence of attributes.

(*iii*) as a test of homogeneity of independent estimates of the population variance.

(*iv*) as a test of the hypothetical value of the population variance σ^2.

(*v*) as a list of the homogeneity of independent estimates of the population correlation coefficient.

Conditions for Applying χ^2 Test

Following are the conditions which should be satisfied before χ^2 test can be applied.

(*a*) N, the total number of frequencies should be large. It is difficult to say what constitutes largeness, but as an arbitrary figure, we may say that **N should be at least 50,** however, few the cells.

(*b*) No theoretical cell-frequency should be small. Here again, it is difficult to say what constitutes smallness, but 5 should be regarded as the very minimum and **10 is better.** If small theoretical frequencies occur (*i.e.*, < 10), the difficulty is overcome by grouping two or more classes together before calculating (O – E). **It is important to remember that the number of degrees of freedom is determined with the number of classes after regrouping.**

(*c*) The constraints on the cell frequencies, if any, should be linear.

Note. If any one of the theoretical frequency is less than 5, then we apply a corrected given by F Yates, which is usually known as 'Yates correction for continuity', we add 0.5 to the cell frequency which is less than 5 and adjust the remaining cell frequency suitable so that the marginal total is not changed.

The χ^2 Distribution

For large sample sizes, the sampling distribution of χ^2 can be closely approximated by a continuous curve known as the chi-square distribution. The probability function of χ^2 distribution is given by

$$f(\chi)^2 = c(\chi^2)^{(\nu/2-1)}\, e^{-x^2/2}$$

where $e = 2.71828$, ν = number of the degrees of freedom ; c = a constant depending only on ν.

Symbolically, the degrees of freedom are denoted by the symbol ν or by d.f. and are obtained by the rule $\nu = n - k$, where k refers to the number of independent constraints.

In general, when we fit a Binomial distribution, the number of degrees of freedom is one less than the number of classes. When we fit a Poisson distribution, the degrees of freedom are 2 less than the number of classes, because we use the total frequency and the arithmetic mean to get the parameter of the Poisson distribution. When we fit a Normal curve, the number of degrees of freedom are 3 less than the number of classes, but in this fitting, we use the total frequency, mean the standard deviation.

NOTES

We may summarise the above explanation as follows :

If the data is given in series of "n" numbers. Then

In the case of Binomial distribution d.f. $= n - 1$

In the case of Poisson distribution d.f. $= n - 2$

In the case of Normal distribution d.f. $= n - 3$.

χ^2 Test as a Test of Goodness of Fit

χ^2 test enables us to ascertain how well the theoretical distributions such as Binomial, Poisson or Normal etc., fit empirical distributions, *i.e.*, distributions obtained from sample data. If the **calculated value of χ^2 is less than the table value** at a specified level (generally 5%) of significance, the **fit is considered to be good** *i.e.*, the divergence between actual and expected frequencies is attributed to fluctuations of simple sampling. If the calculated value of χ^2 is greater than the table value, the fit is considered to be poor.

ILLUSTRATIVE EXAMPLES

Example 1. *The following table gives the number of accidents that take place in an industry during various days of the week. Test if accidents are uniformly distributed over the week.*

Day	*Mon*	*Tue*	*Wed*	*Thu*	*Fri*	*Sat*
No. of accidents	*14*	*18*	*12*	*11*	*15*	*14*

Sol. Null hypothesis H_0. The accidents are uniformly distributed over the week.

Under this H_0, the expected frequencies of the accidents on each of these days

$$= \frac{84}{6} = 14$$

Observed frequency O_i	14	18	12	11	15	14
Expected frequency E_i	14	14	14	14	14	14
$(O_i - E_i)^2$	0	16	4	9	1	0

$$\therefore \qquad \chi^2 = \frac{\Sigma(O_i - E_i)^2}{E_i} = \frac{30}{14} = 2.1428.$$

Conclusion. Table value of χ^2 at 5% level for $(6 - 1 = 5$ d.f.) is 11.09. (see χ^2-table)

Since the calculated value of χ^2 is less than the tabulated value, H_0 is accepted *i.e.*, the accidents are uniformly distributed over the week.

Example 2. *Records taken of the number of male and female births in 800 families having four children are as follows :*

No. of male births	*0*	*1*	*2*	*3*	*4*
No. of female births	*4*	*3*	*2*	*1*	*0*
No. of families	*32*	*178*	*290*	*236*	*94*

NOTES

Test whether the data are consistent with the hypothesis that the Binomial law holds and the chance of male birth is equal to that of female birth, namely $p = q = 1/2$.

Sol. Null Hypothesis H_0 : The data are consistent with the hypothesis of equal probability for male and female births. *i.e.,* $p = q = 1/2$.

We use Binomial distribution to calculate theoretical frequency given by :

$$N(r) = N \times P(X = r)$$

where N is the total frequency, $N(r)$ is the number of families with r male children and

$$P(X = r) = {}^nC_r p^r q^{n-r},$$

where p and q are probability of male and female birth, n is the number of children.

$$\therefore \quad N(0) = \text{No. of families with 0 male children} = 800 \times {}^4C_0 \left(\frac{1}{2}\right)^4$$

$$= 800 \times 1 \times \frac{1}{2^4} = 50$$

$$N(1) = 800 \times {}^4C_1 \left(\frac{1}{2}\right)^1 \left(\frac{1}{2}\right)^3 = 200 \text{ ; } N(2) = 800 \times {}^4C_2 \left(\frac{1}{2}\right)^2 \left(\frac{1}{2}\right)^2 = 300$$

$$N(3) = 800 \times {}^4C_3 \left(\frac{1}{2}\right)^1 \left(\frac{1}{2}\right)^3 = 200 \text{ ; } N(4) = 800 \times {}^4C_4 \left(\frac{1}{2}\right)^0 \left(\frac{1}{2}\right)^4 = 50$$

Observed frequency O_i	32	178	290	236	94
Expected frequency E_i	50	200	300	200	50
$(O_i - E_i)^2$	324	484	100	1296	1936
$\dfrac{(O_i - E_i)^2}{E_i}$	6.48	2.42	0.333	6.48	38.72

$$\chi^2 = \frac{\Sigma(O_i - E_i)^2}{E_i} = 54.433.$$

Conclusion. Table value of χ^2 at 5% level of significance for $5 - 1 = 4$ d.f. is 9.49 (see χ_2-table). Since the calculated value of χ^2 is greater than the tabulated value, H_0 is rejected. *i.e.,* the data are not consistent with the hypothesis that the Binomial law holds and that the chance of a male birth is not equal to that of a female birth.

Note. Since the fitting is Binomial, the degrees of freedom $\nu = n - 1$ *i.e.,* $\nu = 5 - 1 = 4$.

Example 3. *Fit a Poisson distribution to the following :*

x :	*0*	*1*	*2*	*3*	*4*
f :	*46*	*38*	*22*	*9*	*1*

Sol. Null Hypothesis H_0 : Poisson fit is a good fit to the data.

Here mean of the given distribution $= \dfrac{\Sigma f_i x_i}{\Sigma f_i} = \dfrac{113}{116} = 0.97$

To fit a Poisson distribution, we require m

Here $\quad m = \bar{x} = 0.97$

By Poisson distribution, the frequency of r success is

$$N(r) = N \times e^{-m} \cdot \frac{m^r}{r\,!}, \text{ N is the total frequency}$$

i.e.,

$$N(0) = 116 \times e^{-0.97} = 116 \times 0.37 = 43.97\ ;$$

$$N(2) = 116 \times e^{-0.97} \times \frac{(0.97)^2}{2} = 116 \times 0.17 = 19.72$$

$$N(3) = 116 \times e^{-0.97} \times \frac{(0.97)^3}{3\,!} = 6.688$$

$$N(4) = 116 \times e^{-0.97} \times \frac{(0.97)^4}{4\,!} = 1.62$$

X	0	1	2	3	4
O_i	46	38	22	9	1
E_i	43.97	41.76	19.72	6.68	1.62
$\frac{(O_i - E_i)^2}{E_i}$	0.093	0.338	0.263	0.805	0.23

$$\therefore \qquad \chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} = 1.729$$

Conclusion. The calculated value of χ^2 is 1.729. Also the tabulated value of χ^2 at 5% level of significance for $\nu = 5 - 2 = 3$ *d.f.* is 7.815 (see χ^2-table). Since the calculated value of χ^2 at 5% level of significance is less than the tabulated value. Hence H_0 is accepted *i.e.*, Poisson distribution is a best fit to the given data.

Example 4. *Fit a Binomial distribution to the following frequency distribution.*

x	*0*	*1*	*2*	*3*	*4*	*5*	*6*
f	*13*	*25*	*52*	*58*	*32*	*16*	*4*

Sol. Null Hypothesis H_0 : Binomial distribution is a good fit to the given data.

We first find mean of the given distribution.

x_i	f_i	$f_i x_i$
0	13	0
1	25	25
2	52	104
3	58	174
4	32	128
5	16	80
6	4	24
	$N = \sum f_i = 200$	$\sum f_i x_i = 535$

$$\text{Here mean} = \frac{\sum f_i x_i}{\sum f_i}$$

$$\Rightarrow \qquad np = \frac{535}{200} = 2.675$$

$$\Rightarrow \qquad p = \frac{2.675}{7} = 0.38 \qquad |\ n = 7$$

$$\therefore \qquad q = 1 - p = 0.62$$

NOTES

When $r = 0$, $P(X = 0) = 200 \times {}^7C_0\,(0.62)^7 = 7.043$

When $r = 1$, $P(X = 1) = 200 \times {}^7C_1\,(0.38)\,(0.62)^6$
$= 200 \times 7 \times 0.38 \times 0.0568 = 30.2176$

When $r = 2$, $P(X = 2) = 200 \times {}^7C_2\,(0.38)^2\,(0.62)^5$
$= 200 \times 21 \times 0.1444 \times 0.0916 = 55.55$

When $r = 3$, $P(X = 3) = 200 \times {}^7C_3\,(0.38)^3(0.62)^4$
$= 200 \times 35 \times 0.054 \times 0.1477 = 55.83$

When $r = 4$, $P(X = 4) = 200 \times {}^7C_4\,(0.38)^4(0.62)^3$
$= 200 \times 35 \times 0.020 \times 0.23 = 33.32$

When $r = 5$, $P(X = 5) = 200 \times {}^7C_5\,(0.38)^5\,(0.62)^2$
$= 200 \times 21 \times 0.0079 \times 0.3844 = 12.75$

When $r = 6$, $P(X = 6) = 200 \times {}^7C_6 \times (0.38)^6(0.62)^1$
$= 200 \times 7 \times 0.0030 \times 0.62 = 2.64$

E_i	O_i	$(O_i - E_i)^2$	$\frac{(O_i - E_i)^2}{E_i}$
7.043	13	35.48	178.73
30.21	25	27.144	24.38
55.55	52	12.60	2.85
55.83	58	4.70	0.39
33.32	32	1.74	0.09
12.75	16	10.56	8.74
2.64	4	1.84	1.28
	Total = 200		

(adjust the observed frequencies such that their total sum = 200)

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} = 207.72$$

Conclusion : The tabulated value of χ^2 at 5 % level of significance for $7 - 1 = 6$ d.f 12.592 (See-chi-square table). Since the calculated value of χ^2 is the tabulated value of χ^2 at 5% level of significance.

Therefore H_0 is rejected *i.e.*, The Binomial law does not hold good according to the given data.

χ^2 Test as a Test of Independence

With the help of χ^2 test, we can find whether or not two attributes are associated. We take the null hypothesis that there is no association between the attributes under study, *i.e.*, **we assume that the two attributes are independent. If the calculated value of χ^2 is less than the table value** at a specified level (generally 5%) of significance, the hypothesis holds good, *i.e.*, **the attributes are independent** and do not bear any association. On the other hand, if the calculated value of χ^2 is greater than the table value at a specified level of significance, we say that the results of the experiment do not support the hypothesis. Thus a very useful application of χ^2 test is to investigate the relationship between trials or attributes which can be classified into two or more categories.

The sample data set out into two-way table, called **contingency table.**

Let us consider two attributes A and B divided into r classes $A_1, A_2, A_3, ..., A_r$ and B divided into s classes $B_1, B_2, B_3, ..., B_s$. If (A_i), (B_j) represents the number of persons possessing the attributes A_i, B_j respectively, $(i = 1, 2, ..., r, j = 1, 2, ..., s)$ and $(A_i B_j)$ represent the number of persons possessing attributes A_i and B_j. Also we have $\sum_{i=1}^{r} A_i = \sum_{j=1}^{s} B_j = N$, where N is the total frequency. The contingency table for $r \times s$ is given below :

A / *B*	A_1	A_2	A_3	...A_r	*Total*
B_1	(A_1B_1)	(A_2B_1)	(A_3B_1)	......(A_rB_1)	B_1
B_2	(A_1B_2)	(A_2B_2)	(A_3B_2)	......(A_rB_2)	B_2
B_3	(A_1B_3)	(A_2B_3)	(A_3B_3)	......(A_rB_3)	B_3
......	......	......	......	......	......
......	......	......	......	......	......
B_s	(A_1B_s)	(A_2B_s)	(A_3B_s)	......(A_rB_s)	(B_s)
Total	(A_1)	(A_2)	(A_3)	......(A_r)	N

H_0 : Both the attributes are independent. *i.e.*, A and B are independent under the null hypothesis, we calculate the expected frequency as follows :

$P(A_i)$ = Probability that a person possesses the attribute

$$A_i = \frac{(A_i)}{N} \quad i = 1, 2, ..., r$$

$P(B_j)$ = Probability that a person possesses the attribute

$$B_j = \frac{(B_j)}{N}$$

$P(A_iB_j)$ = Probability that a person possesses both attributes A_i and $B_j = \frac{(A_iB_j)}{N}$

If $(A_iB_j)_0$ is the expected number of persons possessing both the attributes A_i and B_j

$$(A_iB_j)_0 = NP(A_iB_j) = NP(A_i)(B_j)$$

$$= N\frac{(A_i)}{N}\frac{(B_j)}{N} = \frac{(A_i)(B_j)}{N} \qquad (\because \text{ A and B are independent})$$

Hence $$\chi^2 = \sum_{i=1}^{r}\sum_{j=1}^{s}\left[\frac{[(A_iB_j)_0 - (A_iB_j)]^2}{(A_iB_j)_0}\right]$$

which is distributed as a χ^2 variate with $(r - 1)(s - 1)$ degrees of freedom.

NOTES

Note 1. For a 2 × 2 contingency table where the frequencies are $\dfrac{a/b}{c/d}$ χ^2 can be calculated from independent frequencies as $\chi^2 = \dfrac{(a+b+c+d)(ad-bc)^2}{(a+b)(c+d)(b+d)(a+c)}$.

Note 2. If the contingency table is not 2 × 2, then the formula for calculating χ^2 as given in note 1, cannot be used. Hence, we have another formula for calculating the expected frequency $(A_iB_j)_0 = \dfrac{(A_i)(B_j)}{N}$

i.e., expected frequency in each cell is $= \dfrac{\text{Product of column total and row total}}{\text{whole total}}$.

Note 3. If $\dfrac{a \mid b}{c \;\; d}$ is the 2 × 2 contingency table with two attributes, $Q = \dfrac{ad-bc}{ad+bc}$ is called the coefficient of association.

If the attributes are independent then $\dfrac{a}{b} = \dfrac{c}{d}$.

Note 4. Yate's Correction. In a 2 × 2 table, if the frequencies of a cell is small, we make Yate's correction to make χ^2 continuous.

Decrease by $\dfrac{1}{2}$ those cell frequencies which are greater than expected frequencies, and increase by $\dfrac{1}{2}$ those which are less than expectation. This will not affect the marginal columns. This correction is known as Yate's correction to continuity.

After Yate's correction $\chi^2 = \dfrac{N\left(bc - ad - \frac{1}{2}N\right)^2}{(a+c)(b+d)(c+d)(a+b)}$ when $ad - bc < 0$

$$\chi^2 = \frac{N\left(ad - bc - \frac{1}{2}N\right)^2}{(a+c)(b+d)(c+d)(a+b)} \quad \text{when} \quad ad - bc > 0.$$

Example 5. *What are the expected frequencies of 2 × 2 contingency tables given below :*

(*i*)

a	b
c	d

(*ii*)

2	10
6	6

Sol. Observed frequencies → Expected frequencies

(*i*)

a	b	$a+b$
c	d	$c+d$
$a+c$	$b+d$	$a+b+c+d = N$

$\dfrac{(a+c)(a+b)}{a+b+c+d}$	$\dfrac{(b+d)(a+b)}{a+b+c+d}$
$\dfrac{(a+c)(c+d)}{a+b+c+d}$	$\dfrac{(b+d)(c+d)}{a+b+c+d}$

NOTES

Observed frequencies → Expected frequencies

(*ii*)

2	10	12
6	6	12
8	16	24

$\frac{8 \times 12}{24} = 4$	$\frac{16 \times 12}{24} = 8$
$\frac{8 \times 12}{24} = 4$	$\frac{16 \times 12}{24} = 8$

Example 6. *From the following table regarding the colour of eyes of fathers and sons test if the colour of son's eye is associated with that of the father.*

		Eye colour of son	
		Light	*Not light*
Eye colour of father	*Light*	*471*	*51*
	Not light	*148*	*230*

Sol. Null hypothesis H_0. The colour of son's eye is not associated with that of father. *i.e.,* they are independent.

Degree of freedom $v = (r - 1)(s - 1) = (2 - 1)(2 - 1) = 1$

Under H_0, we calculate the expected frequency in each cell as

$$= \frac{\text{Product of column total and row total}}{\text{Whole total}}$$

Expected frequencies are :

Eye colour of father \ *Eye colour of son*	*Light*	*Not light*	*Total*
Light	$\frac{619 \times 522}{900} = 359.02$	$\frac{289 \times 522}{900} = 167.62$	522
Not light	$\frac{619 \times 378}{900} = 259.98$	$\frac{289 \times 378}{900} = 121.38$	378
Total	619	289	900

$$\chi^2 = \sum_{i=1}^{4} \frac{(O_i - E_i)^2}{E_i}$$

$$= \frac{(471 - 359.02)^2}{359.02} + \frac{(51 - 167.62)^2}{167.62} + \frac{(148 - 259.98)^2}{259.98} + \frac{(230 - 121.38)^2}{121.38}$$

$= 261.498.$

Also $\chi^2_{0.05}$ = Tabulated value of χ^2 at 5% level for 1 d.f. is 3.841.

Conclusion. Since the calculated value of $\chi^2 >$ tabulated value of χ^2, H_0 is rejected. They are dependent *i.e.,* the colour of son's eye is associated with that of the father.

NOTES

Example 7. *A cigarette company interested in the effect of sex on the type of cigarettes smoked and has collected the following data from a random sample of 150 persons.*

Cigarette size	*Number of people smoked*		*Total*
	male	*female*	
Small	*25*	*30*	*55*
Medium	*40*	*15*	*55*
King size	*30*	*10*	*40*
Total	*95*	*55*	*150*

Test whether the type of cigarette smoked and sex are independent at level of significance $\alpha = 5\%$ *:*

Sol. Let H_0 : The type of cigarette smoked and sex are independent

$\therefore$ Degree of freedom $= (r - 1)(s - 1) = (2 - 1)(3 - 1) = 2$

The given data can be shown in the form of 3×2 contingency table.

Observed Frequencies

	Male	*Female*	*Total*
Small	25	30	55
Medium	40	15	55
King size	30	10	40
Total	95	55	150

Expected Frequencies

$\dfrac{95 \times 55}{150} = 34.83$	$\dfrac{55 \times 55}{150} = 20.16$
$\dfrac{95 \times 55}{150} = 34.83$	$\dfrac{55 \times 55}{150} = 20.16$
$\dfrac{95 \times 40}{150} = 25.3$	$\dfrac{55 \times 40}{150} = 14.66$

Calculate χ^2, we have the following table :

O_i	25	30	40	15	30	10
E_i	34.83	20.16	34.83	20.16	25.3	14.66
$\dfrac{(O_i - E_i)^2}{E_i}$	2.774	4.80	0.76	1.32	0.873	1.481

$$\therefore \chi^2 = \sum_{i=1}^{6} \frac{(O_i - E_i)^2}{E_i} = 2.774 + 4.80 + 0.76 + 1.32 + 0.873 + 1.48$$

$$= 12.008$$

NOTES

Also $\chi^2_{0.05}$ = The tabulated value of χ^2 at 5% level of significance for 2 d.f. = 5.991

Conclusion. The calculated value of χ^2 is greater than the tabulated value of χ^2 at 5% level of significance. Hence H_0 is rejected. *i.e.*, the type of cigarette smoked and sex depends on each other.

Example 8. *From the following data, use χ^2-test and conclude whether inoculation is effective in preventing tuberculosis.*

	Attached	*Not-attached*	*Total*
Inoculated	*31*	*469*	*500*
Not-inoculated	*185*	*1315*	*1500*
Total	*216*	*1784*	*2000*

Sol. Null Hypothesis H_0. The inoculation is not effective in presenting tuberculosis

Degree of freedom = $(r-1)(s-1) = (2-1)(2-1) = 1$

The given data can be shown in the form of 2 × 2 contingency table.

Observed Frequencies

	Attached	*Not-attached*	*Total*
Inoculated	31	469	500
Not-inoculated	185	1315	1500
Total	216	1784	2000

Expected Frequencies

$\frac{216 \times 500}{2000} = 54$	$\frac{1784 \times 500}{2000} = 446$
$\frac{216 \times 1500}{2000} = 162$	$\frac{1784 \times 1500}{2000} = 1338$

To calculate χ^2, we have the following table :

O_i	31	185	469	1315
E_i	54	162	446	1338
$(O_i - E_i)^2$	529	529	529	529
$\frac{(O_i - E_i)^2}{E_i}$	9.79	3.26	1.186	0.395

$$\therefore \quad \chi^2 = \sum_{i=1}^{4} \frac{(O_i - E_i)^2}{E_i} = 9.79 + 3.26 + 1.186 + 0.395 = 14.631$$

NOTES

Also $\chi^2_{0.05}$ = The tabulated value of χ^2 at 5% level of significance for 1 d.f. = 3.841.

Conclusion. The calculated value of χ^2 is greater than the tabulated value of χ^2 at 5% level of significance for 1 d.f. $\therefore$ H_0 is rejected *i.e.*, the inoculation is effective in preventing tuberculosis.

6.3 CENTRAL t-DISTRIBUTION OR STUDENT'S 't' DISTRIBUTION

If $X_1, X_2, \ldots\ldots, X_n$ be a random sample of size n from a normal population with mean μ and variance σ^2. Then, we define the student's 't' by

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}} \sim (n - 1 \text{ d.f.})$$

where $$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i, \text{ the sample mean}$$

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2,$$

an unbaised estimate of the population variance σ^2

Also, the probability density function is given by

$$f(t) = \frac{1}{\sqrt{v}\,\beta\left(\frac{1}{2}, \frac{v}{2}\right)} \cdot \frac{1}{\left(1 + \frac{t^2}{v}\right)^{\frac{v+1}{2}}}, -\infty < t < \infty$$

For $v = 1$,

$$f(t) = \frac{1}{\beta\left(\frac{1}{2}, \frac{1}{2}\right)} \cdot \frac{1}{1+t^2} = \frac{1}{\pi} \cdot \frac{1}{1+t^2}, -\infty < t < \infty$$

which is the probability density function of the standard Cauchy distribution.

$$\because \quad \beta\left(\frac{1}{2}, \frac{1}{2}\right) = \Gamma\left(\frac{1}{2}\right) \frac{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{1}{2}\right)}{\Gamma\left(\frac{1}{2} + \frac{1}{2}\right)}$$

$$= \frac{\left(\Gamma\left(\frac{1}{2}\right)\right)^2}{\Gamma(1)} = \frac{(\sqrt{\pi})^2}{1} = \pi$$

Hence for $v = 1$, student's 't' distribution tends to standard cauchy distribution.

Theorem X. *Derive student's t-distribution.*

Proof. By definition, the student's t is defined as

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}} \qquad \ldots(1)$$

Squaring (1), and using $s^2 = \dfrac{n}{n-1} S^2$, where S^2 is the sample variance, we have

$$t^2 = \frac{(\bar{x}-\mu)^2\, n}{s^2} = \frac{(\bar{x}-\mu)^2\,(n-1)}{S^2}$$

$$\Rightarrow \quad \frac{t^2}{n-1} = \frac{(\bar{x}-\mu)^2}{S^2} = \frac{\dfrac{(\bar{x}-\mu)^2}{\sigma^2/n}}{\dfrac{nS^2}{\sigma^2}}$$

Since $x_i (i = 1, 2, ..., n)$ is a random sample from the normal population with mean μ and variance σ^2, $\bar{x} \sim N(\mu, \sigma^2/n) \Rightarrow \dfrac{(\bar{x}-\mu)}{\sigma/\sqrt{n}} \sim N(0, 1)$

Hence $\dfrac{(\bar{x}-\mu)^2}{\sigma^2/n}$, being the square of a standard normal variate is a chi-square variate with 1 d.f.

Also $\dfrac{nS^2}{\sigma^2}$ is a χ^2-variate with $(n-1)$ d.f.

Further since $\bar{x}$ and S^2 are independently distributed, $\dfrac{t^2}{n-1}$ being the ratio of two independent χ^2-variates with 1 and $(n-1)$ *d.f.* respectively, is a $\beta_2\left(\dfrac{1}{2}, \dfrac{n-1}{2}\right)$ variate and its distribution is given by :

$$dF(t) = \frac{1}{B\left(\frac{1}{2}, \frac{v}{2}\right)} \cdot \frac{(t^2/v)^{\frac{1}{2}-1}}{\left(1+\dfrac{t^2}{v}\right)^{(v+1)/2}}\, d(t^2/v),\ 0 \le t^2 < \infty \qquad | \text{ where } v = (n-1)$$

$$= \frac{1}{\sqrt{v}\, B\left(\frac{1}{2}, \frac{v}{2}\right)} \cdot \frac{1}{\left(1+\dfrac{t^2}{v}\right)^{(v+1)/2}}\, dt;\ -\infty < t < \infty$$

the factor 2 disappearing since the integral from $-\infty$ to ∞ must be unity. This is the required probability density function of Student's t-distribution with $v = (n-1)$ d.f.

Remark. If $X_1, X_2, ..., X_n$ be a random sample from a normal population with mean μ and variance σ^2, then

(*i*) $\bar{X} \sim N(\mu, \sigma^2/n)$

(*ii*) $\dfrac{nS^2}{\sigma^2} = \displaystyle\sum_{i=1}^{n} \left(\frac{X_i - \bar{X}}{\sigma}\right)^2$ is a χ^2-variate with $(n-1)$ d.f.

Confidence Limits or Fiducial Limits for μ.

Let $t_{0.05}$ denotes the tabulated value of t for $v = (n-1)$ degree of freedom, at 5% level of significance, then

$$P(|\,t\,| > t_{0.05}) = 0.05 \Rightarrow P(|\,t\,| \le t_{0.05}) = 0.95,$$

NOTES

The 95% confidence limits for μ and given by :

$$| t | \leq t_{0.05}, \quad i.e., \quad \left| \frac{\bar{x} - \mu}{S/\sqrt{n}} \right| \leq t_{0.05} \Rightarrow \bar{x} - t_{0.05} \cdot \frac{S}{\sqrt{n}} \leq \mu \leq \bar{x} + t_{0.05} \frac{S}{\sqrt{n}}$$

Hence, 95% confidence limits for μ are : $\bar{x} \pm t_{0.05} \cdot (S/\sqrt{n})$

Similarly, 99% confidence limits for μ are : $\bar{x} \pm t_{0.01} \cdot (S/\sqrt{n})$

where $t_{0.01}$ is the tabulated value of t for $\upsilon = (n - 1)$ d.f. at 1% level of significance.

Theorem XI. *For student's 't' distribution, prove the following.*

(i) μ'_{2r+1} *(about origin) = 0, r = 0, 1, 2, ..., i.e., all the moments of odd order about the origin vanish.*

(ii) $\mu_{2r} = \dfrac{(2r-1)(2r-3)\ldots\ldots 3.1}{(n-2)(n-4)\ldots\ldots(n-2r)} n^r, \; \dfrac{n}{2} > r$

Hence find the values of $\mu_2, \mu_4, \beta_1, \beta_2$

Proof. *(i)* By definition, the probability density function for student's 't' is given as

$$f(t) = \frac{1}{\sqrt{n}\, \beta\left(\frac{1}{2}, \frac{n}{2}\right)} \cdot \frac{1}{\left(1 + \frac{t^2}{n}\right)^{\frac{n+1}{2}}}, -\infty < t < \infty$$

Since $f(t)$ is symmetrical about the line $t = 0$, all the moments of odd order about origin must vanish. Hence μ_{2r+1} (about origin) = 0, r = 0, 1, 2, ...

In particular, μ_1' (about origin) = 0 = mean

$\therefore$ Central moments coincide with moments about origin *i.e.*,

$$\mu_{2r+1} = 0 \; (r = 0, 1, 2, \ldots\ldots)$$

The moments of even order are given by :

$$\mu_{2r} = \mu'_{2r} \text{ (about origin)} = \int_{-\infty}^{\infty} t^{2r} f(t)\, dt = 2\int_0^{\infty} t^{2r} f(t)\, dt$$

$$= 2. \frac{1}{B\left(\frac{1}{2}, \frac{n}{2}\right)\sqrt{n}} \int_0^{\infty} \frac{t^{2r}}{\left(1 + \frac{t^2}{n}\right)^{(n+1)/2}} dt \quad \ldots(1)$$

The integral (1) is absolutely convergent if $2r < n$.

Put $1 + \dfrac{t^2}{n} = 1 \Rightarrow t^2 = \dfrac{n(1-y)}{y} \Rightarrow 2tdt = -\dfrac{n}{y^2} dy$

When $t = 0, y = 1$ and when $t = \infty, y = 0$. Therefore, (1) gives

$$\mu_{2r} = \frac{2}{\sqrt{n}\, B\left(\frac{1}{2}, \frac{n}{2}\right)} \int_1^0 \frac{t^{2r}}{(1/y)^{(n+1)/2}} \cdot \frac{-n}{2ty^2} dy$$

$$= \frac{n}{\sqrt{n}\, B\left(\frac{1}{2}, \frac{n}{2}\right)} \int_0^1 (t^2)^{(2r-1)/2}\, y^{[(n+1)/2]-2}\, dy$$

NOTES

$$= \frac{\sqrt{n}}{B\left(\frac{1}{2}, \frac{n}{2}\right)} \int_0^1 \left[n\left(\frac{1-y}{y}\right)\right]^{r-\frac{1}{2}} y^{[(n+1)/2]-2}\, dy$$

$$= \frac{n^r}{B\left(\frac{1}{2}, \frac{n}{2}\right)} \int_0^1 y^{\frac{n}{2}-r-1} (1-y)^{r-\frac{1}{2}}\, dy = \frac{n^r}{B\left(\frac{1}{2}, \frac{n}{2}\right)} . B\left(\frac{n}{2}-r, r+\frac{1}{2}\right), n > 2r.$$

$$= \frac{\Gamma[(n/2)-r]\, \Gamma\left(r+\frac{1}{2}\right)}{\Gamma(1/2)\, \Gamma(n/2)} n^r \qquad ...(2)$$

$$= \frac{\left(r-\frac{1}{2}\right)\left(r-\frac{3}{2}\right) ... \frac{3}{2}\frac{1}{2}\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{n}{2}-r\right)}{\Gamma(1/2)[(n/2)-1]\,[(n/2)-2] ... \left[\left(\frac{n}{2}\right)-r\right]\Gamma[(n/2)-r]} n^r$$

$$= \frac{(2r-1)\,(2r-3) ... 3.1}{(n-2)\,(n-4) ... (n-2r)} n^r, \ \frac{n}{2} > r \qquad ...(3)$$

Put $r = 1, 2, 3, 4$ in (3), we get

$$\mu_2 = n . \frac{1}{(n-2)} = \frac{n}{n-2}, (n > 2) \qquad ...(4)$$

and $$\mu_4 = n^2 \frac{3.1}{(n-2)\,(n-4)} = \frac{3n^2}{(n-2)\,(n-4)}, (n > 4) \qquad ...(5)$$

Hence $\beta_1 = \frac{\mu_3^{\,2}}{\mu_2^{\,3}} = 0$ and $\beta_2 = \frac{\mu_4}{\mu_2^{\,2}} = 3\left(\frac{n-2}{n-4}\right)$; $(n > 4)$.

Remark. As $n \to \infty$, $\beta_1 = 0$ and $\beta_2 = \lim_{n\to\infty} 3\left(\frac{n-2}{n-4}\right) = 3 \lim_{n\to\infty}\left[\frac{1-(2/n)}{1-(4/n)}\right] = 3.$

Moment Generating Function

Theorem III. *Moment generating function of t-distribution.*

Proof. From above theorem, we have

$$\mu_{2r} = \frac{(2r-1)\,(2r-3) ...... 3.1}{(n-2)\,(n-4) ...... (n-2r)} . n^r, \frac{n}{2} > r$$

If $\frac{n}{2} > r \Rightarrow 2r < n$, it implies all the moments of order $2r < n$ exist, but the moments of order $2r \geq n$ do not exist. Hence the moment generating function of t-distribution does not exist.

ILLUSTRATIVE EXAMPLES

Example 1. *If the random variables X_1 and X_2 are independent and follow chi-square distribution with n degree of freedom show that $\sqrt{n}\ (X_1 - X_2)/2\sqrt{X_1 X_2}$ is distributed as Student's 't' with n degree of freedom independently of $X_1 + X_2$.*

Sol. Since X_1 and X_2 are independent chi-square variates each with n degree of freedom their joint probability density function is given by :

NOTES

$$p(x_1, x_2) = p_1(x_1) \times p_2(x_2)$$

$$= \frac{1}{2^n [\Gamma(n/2)]^2} \cdot e^{-\frac{x_1+x_2}{2}} x_1^{\frac{n}{2}-1} x_2^{\frac{n}{2}-1} \; ; 0 \le x_1 < \infty, 0 \le x_2 < \infty$$

Put $\quad u = \dfrac{\sqrt{n}\,(x_1 - x_2)}{2\sqrt{x_1 x_2}}$ and $v = x_1 + x_2$

$$\Rightarrow \quad x_1 = \frac{v}{2}\left[1 + \frac{1}{\sqrt{\left(1+\frac{n}{u^2}\right)}}\right], x_2 = \frac{v}{2}\left[1 - \frac{1}{\sqrt{\left(1+\frac{n}{u^2}\right)}}\right]$$

The Jacobian of transformation is $J = \dfrac{\partial(x_1, x_2)}{\partial(u, v)} = \dfrac{v}{2\sqrt{n}\left(1+\dfrac{u^2}{n}\right)^{3/2}}$

The joint probability density function of u and v becomes

$$g(u, v) = p(x_1, x_2) \mid J \mid$$

$$= \frac{1}{2^{2n-1}\,\Gamma(n/2)\,\Gamma(n/2)\,\sqrt{n}} \frac{e^{-v/2}\, v^{n-1}}{\left(1+\frac{u^2}{n}\right)^{(n+1)/2}} \; ; -\infty < u < \infty, 0 \le v < \infty$$

By Legender's duplication formula

$$\Gamma(n) = 2^{n-1}\,\Gamma(n/2)\,\Gamma\left(\frac{n+1}{2}\right)\Big/\sqrt{\pi} \quad \Rightarrow \quad \Gamma(n/2) = \frac{\Gamma(n)\sqrt{\pi}}{2^{n-1}\,\Gamma\left(\frac{n+1}{2}\right)}, \text{ we get}$$

$$2^{2n-1}\,\Gamma(n/2)\,\Gamma(n/2)\,\sqrt{n} = \frac{2^{2n-1}\,\Gamma(n)\sqrt{\pi}}{2^{n-1}\,\Gamma\left(\frac{n+1}{2}\right)}\,\Gamma\left(\frac{n}{2}\right)\sqrt{n}$$

$$= 2^n\,\Gamma(n)\,\sqrt{n}\,B\left(\frac{1}{2}, \frac{n}{2}\right) \qquad \left|\; \Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}\right.$$

$$\therefore \quad g(u, v) = \left(\frac{1}{2^n\,\Gamma(n)}\, e^{-v/2}\, v^{n-1}\right)\left[\frac{1}{\sqrt{n}\,B\left(\frac{1}{2}, \frac{n}{2}\right)} \cdot \frac{1}{\left(1+\frac{u^2}{n}\right)^{(n+1)/2}}\right];$$

$$0 < v < \infty, -\infty < u < \infty$$

$$\Rightarrow \quad g(u, v) = g_1(u)\, g_2(v), \qquad \text{...(1)}$$

where
$$g_1(u) = \frac{1}{\sqrt{n}\,\beta\left(\frac{1}{2}, \frac{n}{2}\right)} \cdot \frac{1}{\left(1+\frac{u^2}{n}\right)^{(n+1)/2}}, -\infty < u < \infty \qquad \text{...(2)}$$

and
$$g_2(v) = \frac{1}{2^n\,\Gamma(n)}\, e^{-v/2}\, v^{n-1}, 0 < v < \infty \qquad \text{...(3)}$$

(1) $\Rightarrow$ $U = \sqrt{n}\ (X_1 - X_2)/2\sqrt{X_1X_2}$ and $V = X_1 + X_2$ are independently distributed.

(2) $\Rightarrow$ $U = \sqrt{n}\ (X_1 - X_2)/2\sqrt{X_1X_2} \sim t_n$, and

(3) $\Rightarrow$ $V = X_1 + X_2 \sim \gamma\left(a = \frac{1}{2}, n\right)$.

Example 2. *Show that for t-distribution with n degree of freedom mean deviation about mean is given by :*

$\sqrt{n}\ \Gamma\,[(n-1)/2]\,/\,\sqrt{\pi}\ \Gamma(n/2)$

Sol. M.D. (about mean) $= \int_{-\infty}^{\infty} |t|\, f(t)\, dt = \dfrac{1}{\sqrt{n}\ B\left(\frac{1}{2}, \frac{n}{2}\right)} \int_{-\infty}^{\infty} \dfrac{|t|\, dt}{\left(1 + \frac{t^2}{n}\right)^{(n+1)/2}}$

$$= \frac{2}{\sqrt{n}\ B\left(\frac{1}{2}, \frac{n}{2}\right)} \int_0^{\infty} \frac{t\, dt}{\left(1 + \frac{t^2}{n}\right)^{(n+1)/2}} = \frac{\sqrt{n}}{B\left(\frac{1}{2}, \frac{n}{2}\right)} \int_0^{\infty} \frac{dy}{(1+y)^{(n+1)/2}}, \quad \left|\ \frac{t^2}{n} = y\right.$$

$$= \frac{\sqrt{n}}{B\left(\frac{1}{2}, \frac{n}{2}\right)} \int_0^{\infty} \frac{y^{1-1}}{(1+y)^{\frac{n-1}{2}+1}}\, dy = \frac{\sqrt{n}}{B\left(\frac{1}{2}, \frac{n}{2}\right)} \cdot B\left(\frac{n-1}{2}, 1\right) = \frac{\sqrt{n}\ \Gamma[(n-1)/2]}{\sqrt{\pi}\ \Gamma(n/2)}$$

Theorem XII. *For large degree of freedom (n), show that t-distribution tends to standard normal distribution.*

Proof. The probability density function for t-distribution is given as

$$f(t) = \frac{1}{\sqrt{n}\ \beta\left(\frac{1}{2}, \frac{n}{2}\right)} \cdot \frac{1}{\left(1 + \frac{t^2}{n}\right)^{\frac{n+1}{2}}}, \ -\infty < t < \infty$$

Considere $\lim\limits_{n \to \infty} \dfrac{1}{\sqrt{n}\ B\left(\frac{1}{2}, \frac{n}{2}\right)} = \lim\limits_{n \to \infty} \dfrac{1}{\sqrt{n}} \dfrac{\Gamma[(n+1)/2]}{\Gamma(1/2)\,\Gamma(n/2)}$

$$= \frac{1}{\sqrt{n}} \cdot \frac{1}{\sqrt{\pi}} \left(\frac{n}{2}\right)^{\frac{1}{2}} = \frac{1}{\sqrt{2\pi}}$$

$$\left|\ \Gamma(1/2) = \sqrt{\pi} \text{ and } \lim_{n \to \infty} \frac{\Gamma(n+k)}{\Gamma(n)} = n^k,\ \beta(m, n) = \frac{\Gamma(m)\,\Gamma(n)}{\Gamma(m+n)}\right.$$

$$\therefore \quad \lim_{n \to \infty} f(t) = \lim_{n \to \infty} \frac{1}{\sqrt{n}\ B\left(\frac{1}{2}, \frac{n}{2}\right)} \cdot \lim_{n \to \infty} \left[\left(1 + \frac{t^2}{n}\right)^n\right]^{-\frac{1}{2}} \times \lim_{n \to \infty} \left(1 + \frac{t^2}{n}\right)^{-\frac{1}{2}}$$

$$= \frac{1}{\sqrt{2\pi}} \exp\,(-t^2/2),\ -\infty < t < \infty$$

which is the probability density function of the standard normal variate. Hence, we can say that, for large value of n, the degree of freedom, t-distribution tends to standard normal distribution.

NOTES

Probability Curve for *t*-distribution

The probability density function of *t*-distribution with *n* degree of freedom is given as

$$f(t) = \frac{1}{\sqrt{n}\,\beta\left(\frac{1}{2}, \frac{n}{2}\right)} \cdot \frac{1}{\left(1+\frac{t^2}{n}\right)^{\frac{n+1}{2}}}, -\infty < t < \infty$$

$$= c\left(1+\frac{t^2}{n}\right)^{-\frac{n+1}{2}}, \quad \text{where } c = \frac{1}{\sqrt{n}}\beta\left(\frac{1}{2}, \frac{n}{2}\right)$$

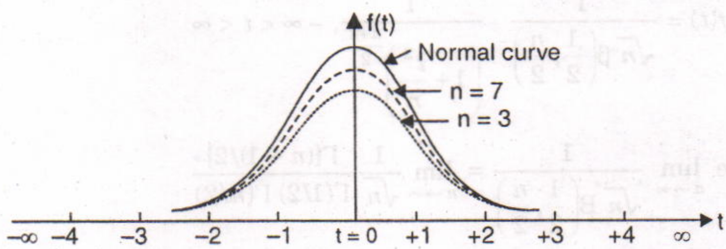
$$\Rightarrow \qquad f(-t) = c\left(1+\frac{t^2}{n}\right)^{-\frac{n+1}{2}} = f(t)$$

Therefore, the probability curve is symmetrical about the line $t = 0$, when t is large, $f(t)$ becomes small and tends to zero when $t \to \infty$. Hence t-axis is an asymptote to the curve. We know

$$\mu_2 = \frac{n}{n-2}, n > 2,$$

$$\beta_2 = \frac{3(n-2)}{n-4}, n > 4$$

For $n > 2$, $\mu_2 > 1$ *i.e.*, the variance of *t*-distribution is greater than that of standard normal distribution and for $n > 4$, $\beta_2 > 3$ and thus *t*-distribution is more flat on the top than the normal curve.



Probability Curve of *t*-distribution

i.e., the tails of the *t*-distribution have a greater probability (area) than the tails of standard normal distribution. Moreover for large *n* (d.f.), *t*-distribution tends to standard normal distribution.

Applications of *t*-distribution

The student's *t*-distribution can be used.

(*i*) To test if the sample mean ($\bar{x}$) differs significantly from the population mean μ.

(*ii*) To test the significance of the difference between two sample means.

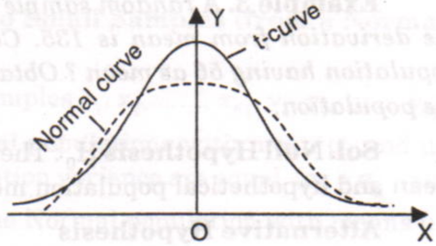
Properties of *t*-distribution

The following are the properties of *t*-distribution :

1. It is unimodal distribution.
2. The probability distribution curve is symmetrical about the line $t = 0$.

NOTES

3. It is bell-shaped curve just like a Normal curve with its tail a little higher above the abscissa than the Normal curve (see fig.)



4. The limiting form of t-distribution, when its degree of freedom $v \to \infty$, is given by

$$y = y_0 e^{\frac{1}{2}t^2}, \text{ which is a Normal curve.}$$

This means that t is normally distributed for large samples.

Critical value of t

The critical value or significant value of t at level of significance α, degrees of freedom γ for two tailed test is given by

$$P[\,|\,t\,| > t_\gamma(\alpha)] = \alpha$$
$$P[\,|\,t\,| \le t_\gamma(\alpha)] = 1 - \alpha$$

The significant value of t at level of significance α for a single tailed test can be got from those of two tailed test by referring to the values at 2α.

Test I : t-test of Significance of the Mean of a Random Sample

To test whether the mean of a sample drawn from a Normal population deviates significantly from a stated value when variance of the population is unknown.

H_0 : There is no significant difference between the sample mean $\bar{x}$ and the population mean μ *i.e.*, we use the statistic

$$t = \frac{\overline{X} - \mu}{s/\sqrt{n}}, \quad \text{where } \overline{X} \text{ is mean of the sample.}$$

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2 \text{ with degrees of freedom } (n-1).$$

At given level of significance α and degrees of freedom $(n-1)$, we refer to t-table t_α (two tailed or one tailed).

If calculated value $|\,t\,|$ is such that $|\,t\,| < t_\alpha$, the null hypothesis is accepted. *i.e.*, If $|\,t\,| > t_\alpha$, H_0 is rejected.

Fiducial Limits of Population Mean

If t_α is the tabulated value of t at level of significance α at $(n-1)$ degrees of freedom. Then $\left|\frac{\overline{X} - \mu}{s/\sqrt{n}}\right| < t_\alpha$ for acceptance of H_0.

$$\Rightarrow \quad \bar{x} - t_\alpha\, s/\sqrt{n} < \mu < \bar{x} + t_\alpha\, s/\sqrt{n} \text{ i.e.,}$$

95% confidence limits (level of significance 5%) are $\overline{X} \pm t_{0.05}\, s/\sqrt{n}$.

99% confidence limits (level of significance 1%) are $\overline{X} \pm t_{0.01}\, s/\sqrt{n}$.

Note. Instead of calculating s, we calculate S for the sample.

Since $\quad s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2 \;\therefore\; S^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2.$

$$\left[(n-1)s^2 = nS^2,\ s^2 = \frac{n}{n-1}S^2\right]$$

NOTES

Theorem XIII. *To derive Snedecor's F-distribution.*

Proof. Let X and Y are two independent chi-square variates with v_1 and v_2 degree of freedom respectively. Then their joint probability density function is given by

$$f(x, y) = \left\{\frac{1}{2^{v_1/2}\,\Gamma(v_1/2)} \exp(-x/2)\, x^{(v_1/2)-1}\right\} \times \left\{\frac{1}{2^{v_2/2}\,\Gamma(v_2/2)} \exp(-y/2)\, y^{(v_2/2)-1}\right\}$$

$$= \frac{1}{2^{(v_1+v_2)/2}\,\Gamma(v_1/2)\,\Gamma(v_2/2)} \exp\{-(x+y)/2\} \times x^{(v_1/2)-1}\, y^{(v_2/2)-1},\ 0 \le (x, y) < \infty$$

Consider the following transformation of variables :

$$F = \frac{x/v_1}{y/v_2} \text{ and } u = y, \text{ so that } 0 \le F < \infty,\ 0 < u < \infty \quad \therefore \quad x = \frac{v_1}{v_2} Fu \text{ and } y = u$$

The Jacobian J of transformation is given by :

$$J = \frac{\partial(x, y)}{\partial(F, u)} = \begin{vmatrix} \dfrac{v_1}{v_2} u & 0 \\ \dfrac{v_1}{v_2} F & 1 \end{vmatrix} = \frac{v_1 u}{v_2}$$

Hence, the joint probability density function of the transformed variables is :

$$g(F, u) = \frac{1}{2^{(v_1+v_2)/2}\,\Gamma(v_1/2)\,\Gamma(v_2/2)} \exp\left\{-\frac{u}{2}\left(1+\frac{v_1}{v_2}F\right)\right\} \times \left(\frac{v_1}{v_2}Fu\right)^{(v_1/2)-1} u^{(v_2/2)-1}\ |\ J\ |$$

$$= \frac{(v_1/v_2)^{v_1/2}}{2^{(v_1+v_2)/2}\,\Gamma(v_1/2)\,\Gamma(v_2/2)} \exp\left\{-\frac{u}{2}\left(1+\frac{v_1}{v_2}F\right)\right\} \times u^{[(v_1+v_2)/2]-1}\, F^{(v_1/2)-1};$$

$$0 < u < \infty,\ 0 \le F < \infty \quad ...(1)$$

Integrating (1) w.r.t. u over the range 0 to ∞, we get

$$g_1(F) = \frac{(v_1/v_2)^{(v_1/2)}\, F^{(v_1/2)-1}}{2^{(v_1+v_2/2)}\,\Gamma(v_1/2)\,\Gamma(v_2/2)} \times \left[\int_0^\infty \exp\left\{-\frac{u}{2}\left(1+\frac{v_1}{v_2}F\right)\right\} u^{[(v_1+v_2)/2]-1}\, du\right]$$

$$= \frac{(v_1/v_2)^{(v_1/2)}\, F^{(v_1/2)-1}}{2^{(v_1+v_2)/2}\,\Gamma(v_1/2)\,\Gamma(v_2/2)} \times \frac{\Gamma[(v_1+v_2)/2]}{\left[\dfrac{1}{2}\left(1+\dfrac{v_1}{v_2}F\right)\right]^{(v_1+v_2)/2}}$$

$$\therefore \quad g_1(F) = \frac{(v_1/v_2)^{v_1/2}}{B\left(\dfrac{v_1}{2}, \dfrac{v_2}{2}\right)} \cdot \frac{F^{(v_1/2)-1}}{\left(1+\dfrac{v_1}{v_2}F\right)^{(v_1+v_2)/2}},\ 0 \le F < \infty \qquad \left|\ \beta(m, n) = \frac{\Gamma(m)\,\Gamma(n)}{\Gamma(m+n)}\right.$$

which is the required probability density function of F-distribution with (v_1, v_2) d.f.

Theorem XIV. *Show that, for F-distribution,*

$$\mu_r' = \left(\frac{v_2}{v_1}\right)^r \cdot \frac{1}{\beta\left(\frac{v_1}{2}, \frac{v_2}{2}\right)} \beta\left(r + \frac{v_1}{2}, \frac{v_2}{2} - r\right) \frac{v_2}{r} > z,$$

Hence find (i) $\mu_1', \mu_2', \mu_3', \mu_4'$

(ii) $\mu_1, \mu_2, \mu_3, \mu_4$

Proof. The probability density function of F-distribution with (v_1, v_2) degree of freedom is given by

$$g(\mathrm{F}) = \frac{\left(\frac{v_1}{v_2}\right)^{\frac{v_1}{2}}}{\beta\left(\frac{v_1}{2}, \frac{v_2}{2}\right)} \cdot \frac{\mathrm{F}^{\frac{v_1}{2}-1}}{\left(1 + \frac{v_1}{v_2}\mathrm{F}\right)^{\frac{v_1+v_2}{2}}},\ 0 \le \mathrm{F} < \infty$$

$$\therefore \qquad \mu_r' \text{ (about origin)} = \mathrm{E}(\mathrm{F}^r) = \int_0^\infty \mathrm{F}^r g(\mathrm{F})\, d\mathrm{F}$$

$$= \frac{(v_1/v_2)^{v_1/2}}{\mathrm{B}\left(\frac{v_1}{2}, \frac{v_2}{2}\right)} \int_0^\infty \mathrm{F}^r \frac{\mathrm{F}^{(v_1/2)-1}}{\left(1 + \frac{v_1}{v_2}\mathrm{F}\right)^{\frac{(v_1+v_2)}{2}}}\, d\mathrm{F}$$

To evaluate the integral, put : $\frac{v_1}{v_2}\mathrm{F} = y$, so that $d\mathrm{F} = \frac{v_2}{v_1}\, dy$

$$\therefore \qquad \mu_r' = \frac{[v_1/v_2]^{v_1/2}}{\mathrm{B}\left(\frac{v_1}{2}, \frac{v_2}{2}\right)} \int_0^\infty \frac{\left(\frac{v_2}{v_1} y\right)^{r+(v_1/2)-1}}{(1+y)^{(v_1+v_2)/2}} \left(\frac{v_2}{v_1}\right) dy$$

$$= \frac{\left(\frac{v_2}{v_1}\right)^r}{\mathrm{B}\left(\frac{v_1}{2}, \frac{v_2}{2}\right)} \int_0^\infty \frac{y^{r+(v_1/2)-1}}{(1+y)^{(v_1/2)+r+[(v_2/2)-r]}}\, dy$$

$$= \left(\frac{v_2}{v_1}\right)^r \cdot \frac{1}{\mathrm{B}\left(\frac{v_1}{2}, \frac{v_2}{2}\right)} \cdot \mathrm{B}\left(r + \frac{v_1}{2}, \frac{v_2}{2} - r\right), v_2 > 2r$$

$$\therefore \qquad \mu_r' = \left(\frac{v_2}{v_1}\right)^r \cdot \frac{\Gamma[r + (v_1/2)\,\Gamma(v_2/2) - r]}{\Gamma(v_1/2)\,\Gamma(v_2/2)}; r < \frac{v_2}{2} \Rightarrow v_2 > 2r \qquad \text{...(1)}$$

Put $r = 1, 2, 3, 4$ in (1), we get

$$\mu_1 = \frac{v_2}{v_1} \cdot \frac{\Gamma[1 + (v_1/2)]\,\Gamma[(v_2/2) - 1]}{\Gamma(v_1/2)\,\Gamma(v_2/2)} = \frac{v_2}{v_2 - 2}, v_2 > 2$$

$$|\ \alpha(n) = (n-1)\,\alpha(n-1)$$

Thus the mean of F-distribution is independent of v_1.

$$\text{Also} \qquad \mu_2' = \left(\frac{v_2}{v_1}\right)^2 \cdot \frac{\Gamma[(v_1/2) + 2]\,\Gamma(v_2/2) - 2]}{\Gamma(v_1/2)\,\Gamma(v_2/2)}$$

$$= \left(\frac{v_2}{v_1}\right)^2 \cdot \frac{[(v_1/2)+1]\,(v_1/2)}{[(v_2/2)-1]\,[(v_2/2)-2]} = \frac{v_2^{\,2}(v_1+2)}{v_1(v_2-2)\,(v_2-4)},\ v_2 > 4.$$

NOTES

$$\therefore \qquad \mu_2 = \mu_2' - \mu_1'^2 = \frac{v_2^{\,2}(v_1+2)}{v_1(v_2-2)(v_2-4)} - \frac{v_2^{\,2}}{(v_2-2)^2}$$

$$= \frac{2v_2^{\,2}\,(v_2+v_1-2)}{v_1(v_2-2)^2\,(v_2-4)},\ v_2 > 4$$

Similarly, on putting $r = 3$ and 4 in μ_r', we get μ_3' and μ_4' respectively, from which the central moments μ_3 and μ_4 can be obtained.

Theorem XV. *Show that for F-distribution with (v_1, v_2) degree of freedom, mode exists if $fv_1 > 2$. Also mode of F-distribution is always less than unity.*

Proof. The Probability density function of F-distribution is given as

$$g(F) = \frac{\left(\frac{v_1}{v_2}\right)^{\frac{v_1}{2}}}{\beta\left(\frac{v_1}{2}, \frac{v_2}{2}\right)} \cdot \frac{F^{\frac{v_1}{2}-1}}{\left(1+\frac{v_1}{v_2}F\right)^{\frac{v_1+v_2}{2}}},\ 0 \le F < \infty$$

Taking logarithm, we get

$$\log\,(g(F)) = \log\left[\frac{\left(\frac{v_1}{v_2}\right)^{\frac{v_1}{2}}}{\beta\left(\frac{v_1}{2}, \frac{v_2}{2}\right)}\right] + \log\frac{F^{\frac{v_1}{2}-1}}{\left(1+\frac{v_1}{v_2}F\right)^{\frac{v_1+v_2}{2}}}$$

$$= c + \left(\frac{v_1}{2} - 1\right)\log F - \frac{v_1+v_2}{2}\log\left(1+\frac{v_1}{v_2}, F\right) \qquad \text{...(1)}$$

where $\qquad c = \log\left(\frac{v_1}{v_2}\right)^{\frac{v_1}{2}} - \log\beta\left(\frac{v_1}{2}, \frac{v_2}{2}\right)$

Differentiating (1) partially w.r.t. F, we get

$$\frac{g'(F)}{g(F)} = 0 + \left(\frac{v_1}{2} - 1\right) \cdot \frac{1}{F} - \frac{v_1+v_2}{2} \cdot \frac{\frac{v_1}{v_2}}{1+\frac{v_1}{v_2}F}$$

For mode, put $g'(F) = 0$

$$\Rightarrow \qquad \left(\left(\frac{v_1}{2} - 1\right) \cdot \frac{1}{F} - \frac{v_1+v_2}{2} \cdot \frac{v_1}{v_2+v_1F}\right) g(F) = 0$$

$$\Rightarrow \qquad \frac{v_1-2}{2F} - \frac{(v_1+v_2)\,v_1}{2(v_2+v_1F)} = 0$$

$$\Rightarrow \qquad \frac{(v_1-2)(v_2+v_1F) - (v_1+v_2)\,v_1F}{2F(v_2+v_1F)} = 0$$

$$\Rightarrow \qquad v_1v_2 + v_1^{\,2}F - 2v_2 - 2v_1F - v_1^{\,2}F - v_1v_2F = 0$$

$$\Rightarrow \qquad v_1v_2 - 2v_2 = (2v_1 + v_1v_2)\,F$$

NOTES

or $$F = \frac{v_2(v_1 - 2)}{v_1(v_2 + 2)}$$

Also at $$F = \frac{v_2(v_1 - 2)}{v_1(v_2 + 2)}, g''(F) < 0$$

Hence mode of F-distribution is given by

$$F = \frac{v_2(v_1 - 2)}{v_1(v_2 + 2)}$$

Further, since $F > 0$, we must have $v_1 > 2$

Also $$\text{mode} = \frac{v_2}{v_2 + 2} \cdot \frac{v_1 - 2}{v_1} < 1$$

Points of Inflexion

Theorem XVI. *Show that the points of inflexion of F-distribution exist for $v_1 > 4$ and are equidistant from mode.*

Proof. We have $$\frac{v_1}{v_2} F = \frac{X}{Y} \sim \beta_2(l, m), \quad ...(1)$$

where $l = v_1/2$ and $m = v_2/2$. We now find the points of inflexion of Beta distribution of second kind with parameters l and m. If $X \sim \beta_2(l, m)$, its Probability density function is :

$$f(x) = \frac{1}{\beta(l, m)} \cdot \frac{x^{l-1}}{(1+x)^{l+m}} ; 0 \le x < \infty \quad ...(2)$$

Points of inflexion are the solution of $f''(x) = 0$ and $f'''(x) \neq 0$

From (2), $\log f(x) = -\log \beta(l, m) + (l - 1) \log x - (l + m) \log (1 + x)$

Differentiating twice *w.r.* to x, we get

$$\frac{f'(x)}{f(x)} = \frac{l-1}{x} - \frac{l+m}{1+x} \quad ...(3)$$

Also $$\frac{f(x) f''(x) - [f'(x)]^2}{[f(x)]^2} = -\left(\frac{l-1}{x^2}\right) + \frac{l+m}{(1+x)^2}$$

$$f''(x) = 0, \text{ gives } \quad -\left[\frac{f'(x)}{f(x)}\right]^2 = -\left(\frac{l-1}{x^2}\right) + \frac{l+m}{(1+x)^2}$$

$$\Rightarrow \quad -\left[\frac{l-1}{x} - \frac{l+m}{1+x}\right]^2 = -\left(\frac{l-1}{x^2}\right) + \frac{l+m}{(1+x)^2} \quad | \text{ using (3)}$$

$$\Rightarrow \quad \frac{l-1}{x^2}(l-1-1) - 2\frac{(l-1)(l+m)}{x(1+x)} + \frac{l+m}{(1+x)^2} \times (l+m+1) = 0$$

$$\Rightarrow \quad (l-1)(l-2)(1+x)^2 - 2x(1+x)(l-1)(l+m) + x^2(l+m)(l+m+1) = 0 \quad ...(4)$$

which is a quadratic in x. It can be easily verified that at these values of x, $f'''(x) \neq 0$, if $l > 2$.

The roots of (4) give the points of inflexion of $\beta_2(l, m)$ distribution. The sum of the points of inflexion is equal to the sum of roots of (4) and is given by :

$$-\left[\frac{\text{Coeff. of } x \text{ in (4)}}{\text{Coeff. of } x^2 \text{ in (4)}}\right] = -\left[\frac{2(l-1)(l-2) - 2(l-1)(l+m)}{(l-1)(l-2) - 2(l-1)(l+m) + (l+m)(l+m+1)}\right]$$

Conclusion. If the calculated value of F exceeds $F_{0.05}$ for $((n_1 - 1), (n_2 - 1))$ degrees of freedom, we conclude that the ratio is significant at 5% level. *i.e.,* we conclude that the sample could have come from two Normal population with same variance.

NOTES

The assumptions on which F-test is based are :

1. The populations for each sample must be normally distributed.
2. The samples must be random and independent.
3. The ratio of σ_1^2 to σ_2^2 should be equal to 1 or greater than 1. That is why we take the larger variance in the Numerator of the ratio.

Method to Apply *f*-test

We defined $$F = \frac{s_1^2}{s_2^2}, \text{ where } s_1^2 = \frac{\Sigma(x_i - \bar{x})^2}{n_1 - 1};\ i = 1, 2, \ldots, n$$

$$s_2^2 = \frac{\Sigma(y_j - \bar{y})^2}{n_2 - 1},\ j = 1, 2, \ldots, m$$

Here s_1^2, s_2^2 are called the unbiased estimates of the population variances.

If the calculated value of F exceeds the tabulated value (which depends on the degree of freedom $\nu_1 = (n_1 - 1)$ and $\nu_2 = (n_2 - 1)$, then the null hypothesis H_0 is rejected, If the calculated value of F is less than the tabulated value, then the null hypothesis H_0 is accepted.

ILLUSTRATIVE EXAMPLES

Example 1. *In two independent samples of sizes 8 and 10, the sum of squares of deviations of the sample values from the respective sample means were 84.4 and 102.6. Test whether the difference of variances of the populations is significant or not.*

Sol. Null hypothesis $H_0 : \sigma_1^2 = \sigma_2^2 = \sigma^2$, *i.e.,* there is no significant difference between population variance.

Under H_0 : $$F = \frac{s_1^2}{s_2^2} \sim F(\nu_1, \nu_2 \text{ d.f}),$$

where $\nu_1 = n_1 - 1$, n_1 = Sample I size = 8 ; $\nu_2 = n_2 - 1$, n_2 = Sample II size = 10

$$\Sigma(X_1 - \bar{X}_1)^2 = 84.4 ;\ \Sigma(X_2 - \bar{X}_2)^2 = 102.6$$

Here $$s_1^2 = \frac{\Sigma(X_1 - \bar{X}_1)^2}{n_1 - 1} = \frac{84.4}{7} = 12.057 ;$$

$$s_2^2 = \frac{\Sigma(X_2 - \bar{X}_2)^2}{n_2 - 1} = \frac{102.6}{9} = 11.4$$

$$F = \frac{s_1^2}{s_2^2} \quad \because \quad s_1^2 > s_2^2 \quad \therefore \quad F = \frac{12.057}{11.4} = 1.0576.$$

Conclusion. The tabulated value of F at 5% level of significance for (7, 9) d.f is 3.29.

$\therefore$ $F_{0.05} = 3.29$ and $| F | = 1.0576 > 3.29 = F_{0.05}$ $\Rightarrow$ H_0 is accepted. *i.e.*, There is no significant difference between the variance of the populations.

NOTES

Example 2. *Two random samples are drawn from 2 Normal populations are as follows :*

A	*17*	*27*	*18*	*25*	*27*	*29*	*13*	*17*
B	*16*	*16*	*20*	*27*	*26*	*25*	*21*	

Test whether the samples are drawn from the same Normal population.

Sol. To test if two independent samples have been drawn from the same population, we have to test (*i*) equality of the means by applying *t*-test and (*ii*) equality of population variance by applying F-test.

Since *t*-test assumes that the sample variances are equal, we shall first apply F-test.

Null hypothesis H_0 : $\sigma_1^2 = \sigma_2^2$, *i.e.*, **the population variance do not differ significantly.**

Alternative hypothesis. $H_1 : \sigma_1^2 \neq \sigma_2^2$.

Test statistic : $F = \frac{s_1^2}{s_2^2}$, (if $s_1^2 > s_2^2$)

Computations for s_1^2 and s_2^2

X_1	$X_1 - \overline{X}_1$	$(X_1 - \overline{X}_1)^2$	X_2	$X_2 - \overline{X}_2$	$(X_2 - \overline{X}_2)^2$
17	– 4.625	21.39	16	– 2.714	7.365
27	5.735	28.89	16	– 2.714	7.365
18	– 3.625	13.14	20	1.286	1.653
25	3.375	11.39	27	8.286	68.657
27	5.735	28.89	26	7.286	53.085
29	7.735	54.39	25	6.286	39.513
13	– 8.625	74.39	21	2.286	5.226
17	– 4.625	21.39			

$$\overline{X}_1 = 21.625 \; ; n_1 = 8 \; ; \Sigma(X_1 - \overline{X}_1)^2 = 253.87$$

$$\overline{X}_2 = 18.714 \; ; n_2 = 7 \; ; \Sigma(X_2 - \overline{X}_2)^2 = 182.859$$

$$s_1^2 = \frac{\Sigma(X_1 - \overline{X}_1)^2}{n_1 - 1} = \frac{253.87}{7} = 36.267 \; ;$$

$$s_2^2 = \frac{\Sigma(X_2 - \overline{X}_2)^2}{n_2 - 1} = \frac{182.859}{6} = 30.47$$

$$F = \frac{s_1^2}{s_2^2} = \frac{36.267}{30.47} = 1.190.$$

Conclusion. The table value of F for $\nu_1 = 7$ and $\nu_2 = 6$ degrees of freedom at 5% level is 4.21. (See F-table) The calculated value of F is less than the tabulated value of F. $\therefore$ H_0 is accepted. Hence we conclude that the variability in two populations is same.

NOTES

***t*-test**

Null hypothesis. $H_0 : \mu_1 = \mu_2$, *i.e.*, the population means are equal.

Alternative hypothesis. $H_1 : \mu_1 \neq \mu_2$

$$s^2 = \frac{\Sigma(X_1 - \overline{X}_1)^2 + \Sigma(X_2 - \overline{X}_2)^2}{n_1 + n_2 - 2} = \frac{253.87 + 182.859}{8 + 7 - 2} = 33.594 \quad \therefore \quad s = 5.796$$

$$t = \frac{\overline{X}_1 - \overline{X}_2}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{21.625 - 18.714}{5.796\sqrt{\frac{1}{8} + \frac{1}{7}}} = 0.9704 \sim t(n_1 + n_2 - 2) \text{ d.f.}$$

Conclusion. The tabulated value of t at 5% level of significance for 13 d.f. is 2.16. (See t-table) the calculated value of t is less than the tabulated value. H_0 is accepted, *i.e.*, there is no significant difference between the population mean ; *i.e.*, $\mu_1 = \mu_2$. $\therefore$ We conclude that the two samples have been drawn from the same normal population.

Example 3. *Two independent sample of sizes 7 and 6 have the following values :*

Sample A	*28*	*30*	*32*	*33*	*31*	*29*	*34*
Sample B	*29*	*30*	*30*	*24*	*27*	*28*	

Examine whether the samples have been drawn from Normal populations having the same variance.

Sol. Null Hypothesis H_0 **:** The variance are equal, *i.e.*, $\sigma_1^2 = \sigma_2^2$ *i.e.*, the samples have been drawn from Normal populations with same variance.

Alternative Hypothesis $H_1 : \sigma_1^2 \neq \sigma_2^2$

Computations for s_1^2 and s_2^2

X_1	$X_1 - \overline{X}_1$	$(X_1 - \overline{X}_1)^2$	X_2	$X_2 - \overline{X}_2$	$(X_2 - \overline{X}_2)^2$
28	– 3	9	29	1	1
30	– 1	1	30	2	4
32	1	1	30	2	4
33	2	4	24	– 4	16
31	0	0	27	– 1	1
29	– 2	4	28	0	0
34	3	9			
		28			26

$$\overline{X}_1 = 31, \quad n_1 = 7\ ; \quad \Sigma(X_1 - \overline{X}_1)^2 = 28$$

$$\overline{X}_2 = 28, \quad n_2 = 6\ ; \quad \Sigma(X_2 - \overline{X}_2)^2 = 26$$

$$s_1^2 = \frac{\Sigma(X_1 - \overline{X}_1)^2}{n_1 - 1} = \frac{28}{6} = 4.666\ ; \quad s_2^2 = \frac{\Sigma(X_2 - \overline{X}_2)^2}{n_2 - 1} = \frac{26}{5} = 5.2$$

Under H_0, the test statistic $\quad F = \frac{s_2^2}{s_1^2} = \frac{5.2}{4.666} = 1.1158. \qquad (\because \quad s_2^2 > s_1^2)$

Conclusion. (See F-table) The tabulated value of F at $\nu_1 = 6 - 1$ and $\nu_2 = 7 - 1$ d.f. for 5% level of significance is 4.39. Since the tabulated value of F is less than the calculated value, H_0 is accepted, *i.e.*, there is no significant difference between the variance, *i.e.*, the samples have been drawn from the Normal population with same variance.

Example 4. *The two random samples reveal the following data :*

Sample no.	*Size*	*Mean*	*Variance*
I	*16*	*440*	*40*
II	*25*	*460*	*42*

Test whether the samples come from the same normal population.

Sol. A Normal population has two parameters namely mean μ and variance σ^2. To test whether the two independent samples have been drawn from the same Normal population, we have to test

(*i*) the equality of means $\qquad$ (*ii*) the equality of variance.

Since *t*-test assumes that the sample variance are equal, we first apply F-test.

Null hypothesis. $H_0 : \sigma_1^2 = \sigma_2^2$ *i.e.*, the population variance do not differ significantly.

Alternative hypothesis. $H_1 : \sigma_1^2 \neq \sigma_2^2$

Under H_0, the test statistic is given by $F = \frac{s_1^2}{s_2^2}, (s_1^2 > s_2^2)$

Given : $n_1 = 16, n_2 = 25\ ; S_1^2 = 40, S_2^2 = 42$

$$\therefore \qquad F = \frac{s_1^2}{s_2^2} = \frac{\frac{n_1 S_1^2}{n_1 - 1}}{\frac{n_2 S_2^2}{n_2 - 1}} = \frac{16 \times 40}{15} \times \frac{24}{25 \times 42} = 0.9752.$$

$$\Rightarrow \qquad F = 0.9752$$

Conclusion. The calculated value of F is 0.9752. The tabulated value of F at (16 – 1, 25 – 1 d.f.) at 5% level of significance is 2.11 (see F-table). Since the calculated value is less than that of the tabulated value, H_0 is accepted ; *i.e.*, the population variance are equal.

***t*-test**

Null hypothesis. $H_0 : \mu_1 = \mu_2$, *i.e.*, the population means are equal.

Alternative hypothesis. $H_1 : \mu_1 \neq \mu_2$

Given : $n_1 = 16, n_2 = 25, \overline{X}_1 = 440, \overline{X}_2 = 460$

$$s^2 = \frac{n_1 s_1^{\,2} + n_2 s_2^{\,2}}{n_1 + n_2 - 2} = \frac{16 \times 40 + 25 \times 42}{16 + 25 - 2} = 43.333 \quad \therefore \quad s = 6.582$$

NOTES

Under H_0, $$t = \frac{\overline{X}_1 - \overline{X}_2}{s\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}} \sim (n_1 + n_2 - 2)\text{ d.f.} = \frac{440 - 460}{6.582\sqrt{\dfrac{1}{16} + \dfrac{1}{25}}} = -9.490$$

$\Rightarrow \quad | t | = 9.490.$

Conclusion. The calculated value of $| t |$ is 9.490. The tabulated value of t at 39 d.f. for 5% level of significance is 1.96 (see t-table). Since the calculated value is greater than the tabulated value, H_0 is rejected. *i.e.*, there is significant difference between means, *i.e.*, $\mu_1 \neq \mu_2$.

Since there is significant difference between means and no significant difference between variance, we conclude that the samples do not come from the same Normal population.

6.5 NON-CENTRAL DISTRIBUTIONS

In previous section, we derived central distributions of the central statistics χ^2, t and F. In this section, we shall derive the non-central distributions of the non-central statistics χ^2, t and F.

Non-central chi-square. The non-central chi-square (χ^2) statistics is defined as

$$\chi^2 = \sum_{i=1}^{n} \frac{x_i^{\,2}}{\sigma_i^{\,2}}$$

where $X_1, X_2, \ldots\ldots, X_n$ are n independent normal variates with means $\mu_1, \mu_2, \ldots\ldots, \mu_n$ and variances $\sigma_1^{\,2}, \sigma_2^{\,2}, \ldots\ldots, \sigma_n^{\,2}$.

Theorem XVI. *Derive non-central χ^2 distribution of chi-square variate.*

Proof. Take $\psi^2 = \displaystyle\sum_{i=1}^{n}\left(\frac{\mu_i^{\,2}}{\sigma_i^{\,2}}\right)$ and transform the variables $x_i'^3$ to the variables $y_j'^2$ by assuming

$$y_i = \left(\frac{\mu_1}{\sigma_1}\frac{x_1}{\sigma_1} + \frac{\mu_2}{\sigma_2}\frac{x_2}{\sigma_2} + \ldots\ldots + \frac{\mu_n}{\sigma_n}\frac{x_n}{\sigma_n}\right) \Bigg/ \sqrt{\sum_{1}^{n} \frac{\mu_i^{\,2}}{\sigma_i^{\,2}}}$$

$$= \sum_{i=1}^{n}\left(\frac{\mu_i}{\sigma_i}.\frac{x_i}{\sigma_i}\right) \Bigg/ \sqrt{\sum_{i=1}^{n}\frac{\mu_i^{\,2}}{\sigma_i^{\,2}}},$$

$$y_j = C_{j1}.\frac{x_1}{\sigma_1} + C_{j2}.\frac{x_2}{\sigma_2} + \ldots\ldots + C_{jn}.\frac{x_n}{\sigma_n}$$

$$= \sum_{i=1}^{n} C_{ji}.\frac{x_i}{\sigma_i}, j = 2, 3, \ldots\ldots, n$$

where the coefficients C_{ji} satisfy the following.

$$\sum_{i=1}^{n} C_{ji}^2 = 1, \quad i.e., \ (C'_j = (C_{j1}, C_{j2}, \ldots\ldots, C_{jn}) \text{ are unitary}$$

$$\sum_{j} C_{ji}\, C_{ji}{}' = 0 = \Sigma\, C_{ji}\, C_{ji}{}', \ i' \neq i \ [i.e. \ C_j \text{ and } C_j', \text{ are orthogonal}]$$

NOTES

It is easy to verify that the variates y_i's are independent normal variates with

$$E(y_1) = \sum_{i=1}^{n} \frac{\mu_i}{\sigma_i} E\left(\frac{x_i}{\sigma_i}\right) \Big/ \sqrt{\sum_{i=1}^{n} \frac{\mu_i^{\,2}}{\sigma_i^{\,2}}} = \sqrt{\sum \frac{\mu_i^{\,2}}{\sigma_i^{\,2}}} = \psi$$

and $$E(y_j) = \sum_{i=1}^{n} C_{ji}\, E\left(\frac{x_i}{\sigma_i}\right) = \sum_{i=1}^{n} C_{ij} \frac{\mu_i}{\sigma_i} = 0, j = 2, 3, \ldots\ldots, n$$

$\because$ The vectors C_j' are mutually orthogonal to the unit vector

$$\left[\frac{\mu_1}{\sigma_1}, \frac{\mu_2}{\sigma_2}, \ldots\ldots, \frac{\mu_n}{\sigma_n}\right] \Big/ \sqrt{\sum_{1}^{n} \frac{\mu_i^{\,2}}{\sigma_i^{\,2}}}$$

Also $$\text{Var}\, y_1 = \sum_{i=1}^{n} \frac{\mu_i^{\,2}}{\sigma_i^{\,2}} \cdot \frac{1}{\sigma_i^{\,2}} \text{Var}\, x_i \Big/ \sum_{i=1}^{n} \frac{\mu_i^{\,2}}{\sigma_i^{\,2}} = 1.$$

$$\text{Var}\, y_j = \sum_{i=1}^{n} C_{ji}^2 \cdot \frac{1}{\sigma_i^{\,2}} \text{Var}\, x_i = \sum_{i=1}^{n} C_{ji}^2 = 1.$$

Thus the variable y_1 is $N(\psi, 1)$ and y_j is $N(0, 1), j = 2, 3, \ldots\ldots, n$.

Now $$\chi^2 = \sum_{i=1}^{n} \frac{x_i^{\,2}}{\sigma_i^{\,2}} = \sum_{i=1}^{n} y_i^{\,2} = y_1^{\,2} + \sum_{j=2}^{n} y_j^{\,2} = \chi_1^{\,2} + \chi_2^{\,2}, \text{ say}$$

where $\chi_i^{\,2}$ has the distribution

$$\frac{1}{\sqrt{2\pi}} e^{-(\chi_1^{\,2} + \psi^2)/2} \cos h\,(\psi\chi_1) \,.\, (\chi_1^{\,2})^{-\frac{1}{2}}\, d\chi_1^{\,2}$$

$$\Rightarrow \quad \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}(\chi_1^{\,2} + \psi^2)\right] \sum_{k=0}^{\infty} \frac{(\psi\, \chi_1)^{2k}}{2^k} \,.\, (\chi_1^{\,2})^{-1/2}\, d\chi_1^{\,2}$$

$$\Rightarrow \quad \frac{1}{2} \sum_{k=0}^{\infty} \exp\left[-\frac{1}{2}\psi^2\right] \frac{\left(\frac{1}{2}\psi^2\right)^k}{k!} \cdot \frac{\exp\left[-\frac{1}{2}\chi_1^{\,2}\right]\left(\frac{1}{2}\chi_1^{\,2}\right)^{k-\frac{1}{2}}}{\Gamma\left(k + \frac{1}{2}\right)} d\chi_1^{\,2}$$

and the distribution of $\chi_2^{\,2}$ is

$$\frac{1}{2^{\frac{n-1}{2}}\, \Gamma\left(\frac{n-1}{2}\right)} \exp\left[-\frac{1}{2}\chi_2^{\,2}\right] (\chi_2^{\,2})^{\frac{n-3}{2}}\, d\chi_2^{\,2}$$

Remark 1. For joint distribution of $\chi_1^{\,2}$ and $\chi_2^{\,2}$, let us make the transformations
$\chi_1 = \chi \cos\theta, \chi_2 = \chi \sin\theta, 0 < \chi < \infty, 0 < \theta < \pi/2$

NOTES

We find that the probability density function of χ^2 $(= \chi_1^2 + \chi_2^2)$ is

$$\exp\left[-\frac{1}{2}\psi^2\right]\exp\left[-\frac{1}{2}\chi^2\right]\sum_{k=0}^{\infty}\frac{\left(\frac{1}{2}\psi^2\right)^k}{k!}\cdot\frac{(\chi^2)^{k+\frac{(n-2)}{2}}}{2^{\frac{n}{2}+k}\,\Gamma\left(\frac{n}{2}+k\right)}$$

Remark 2. The *p.d.f.* of non-central χ^2 is the weighted average of central χ^2 *p.d.f.'s* with $(2k+n)$ *d.f.*, the weights being the Poisson probabilities $(\psi^2/2)^k \exp\left(-\frac{1}{2}\psi^2\right)\Big/ k\,!\; k = 0, 1, \ldots\ldots, n$.

3. If $\psi = 0$ and $k = 0$, then the distribution of non-central χ^2 is the central χ^2-distribution with $n.d.f.$

Moment Generating Function

Theorem XVII. *Find the moment generating function of the non-central χ^2 distribution.*

Proof. Assume $X_1, X_2, \ldots\ldots, X_n$ are n independent normal variates with mean $\mu_1, \mu_2, \ldots\ldots, \mu_n$ and variances unity, then the *p.d.f.* of X_i's is given by

$$f(x) = \frac{1}{\sqrt{2\pi}}\, e^{-\frac{1}{2}(x-\mu)^2}$$

$$\therefore\ M_x{}^2(t) = E[e^{tx^2}] = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{tx^2}\, e^{-(1/2)(x-\mu)^2}\,dx$$

$$= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\exp\left[-\left\{\left(\frac{1}{2}-t\right)x^2 - \mu x + \frac{1}{2}\mu^2\right\}\right]dx$$

$$= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\exp\left[-\left(\frac{1}{2}-t\right)\left\{\left(x-\frac{\mu}{1-2t}\right)^2 + \frac{\mu^2}{1-2t} - \frac{\mu^2}{(1-2t)^2}\right\}\right]dx$$

$$= \frac{1}{\sqrt{2\pi}}\exp\left(\frac{t\mu^2}{1-2t}\right)\int_{-\infty}^{+\infty}\exp\left[-\left(\frac{1-2t}{2}\right)\left(x-\frac{\mu}{1-2t}\right)^2\right]dx$$

$$= \frac{1}{\sqrt{2\pi}}\exp\left(\frac{t\mu^2}{1-2t}\right)\int_{-\infty}^{+\infty}\exp\left(\frac{1}{2}Z^2\right)\frac{dZ}{\sqrt{1-2t}}, \text{ where}$$

$$Z = \left(x - \frac{\mu}{1-2t}\right)\times\sqrt{1-2t} = (1-2t)^{-\frac{1}{2}}\exp\left(\frac{t\mu^2}{1-2t}\right),\ t < \frac{1}{2}.$$

Theorem XVIII. *State Reproductive property or additive property of non-central chi-squares variates.*

Statement. If $Y_1, Y_2, \ldots\ldots, Y_k$ are k independent non-central χ^2-variates with $n_1, n_2, \ldots\ldots, n_k$ degree of freedom, and non-central parameter $\psi_1, \psi_2, \ldots\ldots, \psi_k$

where $\qquad \psi_i^2 = \sum_{i=1}^{n}\frac{\mu_i^2}{\sigma_i^2}$, then

$y_1 + y_2 + \ldots\ldots + y_k$ is also a non-central χ^2-variate with $(n_1 + n_2 + \ldots\ldots + n_k)$ degree of freedom and non-central parameter $\chi = \sum_{i=1}^{n}\psi_i$.

Theorem XIX. *Find the cumulants of non-central χ^2-distribution.*

Proof. k_r = rth cumulant = coeff. of $\frac{t^r}{r!}$ in log $M_X{}^2(t)$ *i.e.*

$$K_{\chi^2}(t) = (n+2\psi)\,t + (n+4\psi)\,t^2 + \ldots\ldots + \left[\frac{2^{r-1}}{r}.\,n + 2\psi\, 2^{r-1}\right] t^r + \ldots\ldots$$

$$\therefore \qquad k_r = \text{coeff. of } \frac{t^r}{r!} \text{ in } K_{\chi^2}(t) = r!\left(\frac{n}{r}+2\psi\right)2^{r-1}$$

$$k_r = 2^{r-1}(r-1)!\,(n+2\psi r)$$

and

$$k_{r-1} = 2^{r-2}(r-2)!\,[n+2\psi(r-1)]$$

$$\therefore \qquad \frac{d}{d\psi}k_{r-1} = 2^{r-2}(r-2)!\,2(r-1) = 2^{r-1}(r-1)! = \frac{k_r}{n+2\psi r}$$

$$\Rightarrow \qquad k_r = (n+2\psi r)\frac{d}{d\psi}(k_{r-1}).$$

Non-central Student's *t*-distribution

Theorem XX. *If X is distributed as N(μ, σ²) and Y is an independent χ^2-variate with n.d.f., then*

$$t = \frac{X/\sigma}{\sqrt{Y}/\sqrt{n}}$$

has a non-central t-distribution with n.d.f. and with non-centrality parameter μ/σ.

Proof. Since X and Y are independent, their joint p.d.f. is

$$\frac{1}{\sigma\sqrt{2\pi}\,2^{\frac{n}{2}}\,|\frac{n}{2}}\,e^{-\frac{1}{2}y}.\,y^{\frac{1}{2}n-1}.\,e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}}$$

$$X = \frac{1}{\sigma\sqrt{2\pi}}.\frac{1}{2^{\frac{n}{2}}|\left(\frac{n}{2}\right)}\,e^{-\frac{1}{2}\frac{(\mu^2+x^2+y)}{\sigma^2}}\,y^{\frac{n}{2}-1}\sum_{j=0}^{\infty}\left(\frac{\mu}{\sigma}\right)^j\left(\frac{X^j}{j!}\right)$$

$$= \frac{1}{\sqrt{2\pi}}.\frac{1}{2^{\frac{n}{2}-1}\,\Gamma\left(\frac{n}{2}\right)}.\,e^{\frac{\mu^2}{2\sigma^2}}.\,e^{-\frac{1}{2}y}.\,y^{\frac{n}{2}-1}.\,e^{-\frac{1}{2}z^2}\times\sum_{j=0}^{\infty}\left(\frac{\mu}{\sigma}\right)^j\frac{z^j}{j!}, \text{ where } z = x/\sigma$$

Now make the transformation $(y, z) \to (t, u)$ such that

$$t = \frac{\sqrt{n}Z}{\sqrt{Y}} \text{ and } u = \sqrt{Y} \Rightarrow Y = u^2 \text{ and } Z = tu/\sqrt{n}$$

Jacobian J of transformation is $J = \frac{\partial(y,z)}{\partial(t,u)} = \begin{vmatrix} 0 & 2u \\ \frac{u}{\sqrt{n}} & \frac{t}{\sqrt{n}} \end{vmatrix} = \frac{2u^2}{\sqrt{n}}$

Hence the joint probability density function of t and u is

$$\frac{1}{2^{\frac{n-1}{2}}\,\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{1}{2}\right)}\,r = e^{-\frac{\mu^2}{2\sigma^2}}.\sum_{j=0}^{\infty}\frac{1}{j!}\left(\frac{\mu}{\sigma}\right)^j\times\exp\left[-\frac{u^2}{2}\left(1+\frac{t^2}{n}\right)\right]\left(\frac{tu}{\sqrt{n}}\right)^j u^{n-1}\,|J|$$

$$= \frac{1}{2^{\frac{(n-1)}{2}}\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{1}{2}\right)} e^{-\frac{\mu^2}{2\sigma^2}} \cdot \sum_{j=0}^{\infty} \frac{1}{j!}\frac{(\mu/\sigma)^j}{(j+1)/2} \times \exp\left[-\frac{u^2}{2}\left(1+\frac{t^2}{n}\right)\right] t^j\, u^{n+j}$$

NOTES

Integrating this w.r.t. u between the limits 0 and ∞, we get the *p.d.f.* of non-central t with *n.d.f.* as

$$\frac{1}{\sqrt{\pi}\,\Gamma\frac{n}{2}} \exp\left[-\frac{\mu^2}{2\sigma^2}\right] \times \sum_{j=0}^{\infty} \frac{1}{j!}\left(\frac{\mu}{\sigma}\right)^j \cdot n^{\frac{j}{2}} \frac{\Gamma\left(\frac{n+j+1}{2}\right)}{n^{(j+1)/2}}$$

$$\cdot \frac{1}{\left(1+\frac{t^2}{n}\right)^{(n+j+1)/2}} dt, -\infty < t < \infty$$

Non-central F-distribution

If χ_1^2 be a non-central χ^2 with degree of freedom n_1 and non-centre parameter χ^2 and let χ_2^2, independent of χ_1^2, is a central χ^2 with n_2 degree of freedom, then the ratio $\frac{\chi_1^2/n_1}{\chi_2^2/n_2}$ is known as non-central F statistics with degree of freedom (n_1, n_2) and non-central parameter χ^2.

Theorem XXI. *Derive the non-central distribution of* χ^2.

Proof. Now the joint *p.d.f.* of χ_1^2 and χ_2^2 is

$$\exp\left[-\frac{1}{2}(\chi_1^2+\chi_2^2)\right] \times \sum_{j=0}^{\infty} \frac{(\psi^2/2)^j}{j!} \times \frac{(\chi_1^2)^{\frac{n_1}{2}-1+j}\,(\chi_2^2)^{\frac{n_2}{2}-1}}{2^{\frac{1}{2}(n_1+n_2)+j}\,\Gamma\left(\frac{n}{2}+\mathrm{J}\right)\Gamma\left(\frac{n_2}{2}\right)}$$

Making the transformation $(\chi_1^2, \chi_2^2) \to (\mathrm{F}, v)$, such that

$$\mathrm{F} = \frac{\chi_1^2/n_1}{\chi_2^2/n_2} \quad \Rightarrow \quad \chi_1^2 = \frac{n_1}{n_2}\,\mathrm{F}v$$

and $\qquad v = \chi_2^2 \quad \Rightarrow \quad$ and $\quad \chi_2^2 = v$

the joint *p.d.f.* of F and v can be obtained.

Integrating with respect to v between 0 and ∞ we find the *p.d.f.* of F as

$$\frac{n_1}{n_2}\exp\left(-\frac{1}{2}\psi^2\right)\sum_{j=0}^{\infty}\left(\frac{\frac{1}{2}\psi^2}{j!}\right)^j \times \frac{\Gamma\left(\frac{n_1+n_2}{2}+\mathrm{J}\right)}{\Gamma\left(\frac{n_1}{2}+\mathrm{J}\right)\Gamma\left(\frac{n_2}{2}\right)}$$

$$\times \frac{\left(\frac{n_1}{n_2}\mathrm{F}\right)^{\frac{n_1}{2}-1+\mathrm{J}}}{\left(1+\frac{n_1}{n_2}\mathrm{F}\right)^{\frac{n_1+n_2}{2}+\mathrm{J}}}, 0 < \mathrm{F} < \infty$$

6.6 ORDER STATISTICS

NOTES

Def. Let $X_1, X_2, ..., X_n$ are n independent and identically distributed random variables and $F(x)$ be the common cumulative distribution function. We arrange X_i's in ascending order of magnitude as

$X_{(1)} \le X_{(2)} \le ...... \le X_{(n)}$, then $X_{(r)}$ is known as rth order statistic where r can take the values 1, 2, 3, ..., n. As there is equality sign, then $X_{(r)}$'s are necessarily dependent.

Notations. We write $X_{(1)} = \min(X_1, X_2, ..., X_n)$

$$= \text{The smallest of } X_1, X_2, ... X_n$$

$$X_{(n)} = \text{Max}\,(X_1, X_2, ..., X_n)$$

$$= \text{The largest of } X_1, X_2, ..., X_n$$

$$X_{(r)} = r\text{th smallest of } X_1, X_2, ..., X_n$$

Cumulative Distribution Function

Theorem XXII. *Find the cumulative distribution function of a single order statistic.*

Proof. Let $F_r(x)$ $(r = 1, 2, ..., n)$ denotes the cumulative distribution function of the rth order statistic $X_{(r)}$; therefore, the cumulative distribution function of the largest order statistic $X_{(n)}$ is given by

$$F_n(x) = P\left[X_{(n)} \le x\right] = P[\text{Max}\,(X_1, X_2, ..., X_n) \le x]$$

$$= P[X_i \le x, i = 1, 2, ..., n]$$

$$= P[X_1 \le x \cap X_2 \le x \cap ......, X_n \le x] \qquad |\, X_i\text{'s are independent}$$

$$= P(X_1 \le x)\, P(X_2 \le x) ......, P(X_n \le x)$$

$$= F(x) . F(x) ......, F(x) = (F(x))^n$$

Hence, the c.d.f. (comulative distribution function) of $X_{(n)} = (F(x))^n$ where $F(x)$ is the c.d.f. of $X_1, X_2, ..., X_n$.

We now find the c.d.f. of the smallest order statistic $X_{(1)}$. Here

$$F_1(x) = P(X_{(1)} \le x) = 1 - P(X_{(1)} \ge x)$$

$$= 1 - P[\min\,(X_1, X_2, ..., X_n) \ge x]$$

$$= 1 - P[X_i \ge x, i = 1, 2, ..., n]$$

$$= 1 - P[X_1 \ge x \cap X_2 \ge x \cap ......, X_n \ge x]$$

$$= 1 - P(X_1 \ge x) . P(X_2 \ge x) ......, P(X_n \ge x)$$

$$= 1 - \prod_{i=1}^{n} P(X_i \ge x)$$

$$= 1 - \prod_{i=1}^{n} [1 - P(X_i \le x)]$$

$$= 1 - (1 - F(x))^n$$

$|\because$ $X_1, X_2, ..., X_n$ independent and identically distributed random variables.

Theorem XXIII. *Find the cumulative distribution function of the rth order statistic* $X_{(r)}$.

Proof. Let $F_r(x)$ denotes the c.d.f. of the rth order statistic $X_{(r)}$, then

$$F_r(x) = P[X_{(r)} \le x]$$

$$= P(\text{atleast } r \text{ of the } X_i\text{'s} \le x]$$

$$= \sum_{J=r}^{n} \text{P[Exactly J out of } X_1, X_2, \ldots\ldots, X_n \text{ are} \le x]$$

NOTES

$$= \sum_{J=r}^{n} {}^{n}C_J \, F^J(x)\,(1 - F(x))^{n-J}$$

Remark. We can also write

$$F_r(x) = \frac{1}{\beta(r, n-r+1)} \int_0^{F(x)} t^{r-1}\,(1-t)^{n-r}\,dt \quad \Big| \text{ using Binomial Probability law}$$

Probability density function

Theorem XXIV. *Find the probability density function of a single order statistic.*

Proof. Let $X_1, X_2, \ldots\ldots, X_n$ are independent and identically distributed random variables and if $f(x)$ be the probability density function, then

$$f(x) = F'(x) \text{ where } F(x) \text{ is the cumulative distribution function of } X_i\text{'s.}$$

Let $f_r(x)$ denotes the *p.d.f.* of *r*th order statistic, then

$$f_r(x) = \frac{d}{dx}[F_r(x)] = \frac{d}{dx} \cdot \frac{1}{\beta(r, n-r+1)} \int_0^{F(x)} t^{r-1}\,(1-t)^{n-r}\,dt \qquad \ldots(1)$$

Take $\quad g(t) = \int t^{r-1}\,(1-t)^{n-r}\,dt \qquad \ldots(2)$

Differentiating (2) w.r.t. t, we get

$$g'(t) = t^{r-1}\,(1-t)^{n-r} \qquad \ldots(3)$$

Integrating w.r.t. t, under the limits 0 to $F(x)$,

$$\int_0^{F(x)} t^{r-1}\,(1-t)^{n-r}\,dt = \int_0^{F(x)} g'(t)\,dt = \Big| g(t) \Big|_0^{F(x)} = g(F(x)) - g(0)$$

$$\Rightarrow \quad \frac{d}{dx} \int_0^{F(x)} t^{r-1}\,(1-t)^{n-r}\,dt = \frac{d}{dx}\,[g(F(x)) - g(0)] = \frac{d}{dx}\,g(F(x))$$

$| \; g(0)$ is constant

$$= g'(F(x)) \,.\, \frac{d}{dx}\,F(x) = g'(F(x)) \,.\, f(x) \qquad \ldots(4)$$

Using (1) and (4), we have

$$f_r(x) = \frac{1}{\beta(r, n-r+1)} \,.\, (F(x))^{r-1}\,(1 - F(x))^{n-r} \,.\, f(x) \qquad | \text{ Using (3)}$$

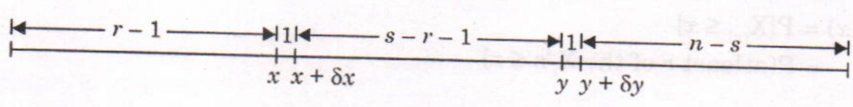
Joint Probability density function of two order statistics

Theorem XXV. *To find the joint probability density function of two order statistics.*

Proof. Let $f_{rs}\,(x, y)$ denotes the joint probability density function of $X_{(r)}$ and $X_{(s)}$ where $1 \le r < s \le n$, then

$$f_{rs}(x, y) = \lim_{\substack{\delta x \to 0 \\ \delta y \to 0}} \frac{P(x \le X_{(r)} \le x + \delta x \cap y \le X_{(s)} \le y + \delta y]}{\delta x\,\delta y} \qquad \ldots(1)$$

The event $E = \{x \le X_{(r)} \le x + \delta x \cap y \le X_{(s)} \le y + \delta y\}$ can be represented as follows :



NOTES

$X_i \le x$ for $r-1$ of the X_i's, $x < X_i \le x + \delta x$ for one X_i

$x + \delta x < X_i \le y$ for $(s-r-1)$ of X_i's, $y < X_i \le y + \delta y$ for one X_i,

and $X_i > y + \delta y$ for $(n-s)$ of the X_i's

By using multinomial probability law, we get

$$P(E) = P[x \le X_{(r)} \le x + \delta x \cap y \le X_{(s)} < y + \delta y]$$

$$= \frac{n!}{(r-1)!\,1!\,(s-r-1)!\,1!\,(n-s)}\, p_1^{r-1} \,.\, p_2\, p_3^{s-r-1} \,.\, p_4\, p_5^{n-s} \qquad ...(2)$$

where $p_1 = P(X_i \le x) = F(x)$, $p_2 = P(x < X_i \le x + \delta x) = F(x + \delta x) - F(x)$

$p_3 = P(x + \delta x < X_i \le y) = F(y) - F(x + \delta x)$

$p_4 = P(y < X_i \le y + \delta y) = F(y + \delta y) - F(y)$

$p_5 = p(X_i > y + \delta y) = 1 - P(X_i \le y + \delta y) = 1 - F(y + \delta y)$

Using (2) in (1), we have

$$f_{rs}(x, y) = \lim_{\substack{\delta x \to 0 \\ \delta y \to 0}} \frac{P(E)}{\delta x \, \delta y}.$$

$$= \frac{n!}{(r-1)!\,(s-r-1)!\,(n-s)!} \times F^{r-1}(x) \times \lim_{\delta x \to 0} \frac{[F(x+\delta x) - F(x)]}{\delta x}$$

$$\times \lim_{\delta x \to 0} [F(y) - F(x + \delta x)]^{s-r-1} \times \lim_{\delta y \to 0} \left[\frac{F(y+\delta y) - F(y)}{\delta y}\right] \times \lim_{\delta y \to 0} [1 - F(y + \delta y)]^{n-s}$$

$$= \frac{n!}{(r-1)!\,(s-r-1)!\,(n-s)!} F^{r-1}(x)\,.\,f(x)\,.\,[F(y) - F(x)]^{s-r-1} f(y)\,.\,[1 - F(y)]^{n-s}$$

Theorem XXVI. *To find the joint probability density function of k-order statistics.*

Proof. Let $f_{r_1, r_2, ..., r_k}(x_1, x_2, ..., x_k)$ denotes the joint probability density function of the k-order statistics $X_{(r_1)}, X_{(r_2)}, ..., X_{(r_k)}$, where $1 \le r_1 < r_2 < ......, < r_k \le n$ and $1 \le k \le n$, then

$$f_{r_1, r_2, ..., r_k}(x_1, x_2, ..., x_k) = \frac{n!}{(r_1 - 1)!\,(r_2 - r_1 - 1)!\,......\,(r_k - r_{k-1} - 1)!\,(n - r_k)!}$$

$$\times F^{r_1 - 1}(x_1) \times f(x_1) \times [F(x_2) - F(x_1)]^{r_2 - r_1 - 1} \times f(x_2)$$

$$\times [F(x_3) - F(x_2)]^{r_3 - r_2 - 1} \times f(x_3) \times ...... \times f(x_k)\,[1 - F(x_k)]^{n - r_k}$$

ILLUSTRATIVE EXAMPLES

Example 1. *Let $X_1, X_2, ..., X_n$ be a random sample from a population with continuous density. Show that $Y_1 = \min(X_1, X_2, ..., X_n)$ is exponential with parameter $n\lambda$ if and only if each X_i is exponential with parameter λ.*

Sol. Let $f(x)$ denotes the probability density function of the random sample $X_1, X_2, ..., X_n$ and if each X_i is exponentially distributed with parameter λ.

We show $Y_1 = \text{Min}(X_1, X_2, ..., X_n)$ is exponentially distributed with parameter $n\lambda$. Now

NOTES

$$f(x) = \lambda e^{-\lambda x} \; ; x \geq 0, \lambda > 0$$

$$F(x) = P(X \leq x) = \int_0^x f(u)\,du = \lambda \int_0^x e^{-\lambda u}\,du = 1 - e^{-\lambda x}$$

Distribution function G(.) of $Y_1 = \min (X_1, X_2, ..., X_n)$ is given by :

$$G_{Y_1}(y) = P(Y_1 \leq y) = 1 - [1 - F(y)]^n \qquad | \text{ Theorem}$$

$$= 1 - [1 - (1 - e^{-\lambda y})]^n = 1 - e^{-n\lambda y},$$

which is the distribution function of exponential distribution with parameter $n\lambda$.

Hence $Y_1 = \text{Min}(X_1, X_2, ..., X_n)$ has exponential distribution with parameter $n\lambda$.

Converse. Let $Y_1 = \text{Min}(X_1, X_2, ..., X_n)$ is exponentially distributed with parameter $n\lambda$. We show each X_i is exponentially distributed with parameter $n\lambda$.

Since $Y_1 = \text{Min}(X_1, X_2, ..., X_n) \sim \exp(n\lambda)$, we have

$$P(Y_1 \leq y) = 1 - e^{-n\lambda y} \;\; \Rightarrow \;\; P(Y_1 \geq y) = e^{-n\lambda y},$$

$$\Rightarrow \;\; P[\min (X_1, X_2, ..., X_n) \leq y] = e^{-n\lambda y},$$

$$\Rightarrow \;\; P[(X_1 \geq y) \cap (X_2 \geq y) \cap ...... (X_n \geq y)]^n = e^{-n\lambda y}$$

$$\Rightarrow \;\; \prod_{i=1}^{n} P(X_i \geq y) = e^{-n\lambda y} \;\; \Rightarrow \;\; [P(X_i \geq y)]^n = e^{-n\lambda y} \qquad [\because \; X_i\text{'s are } i.i.d.]$$

$$\therefore \qquad P(X_i \geq y) = e^{-\lambda y} \;\; \Rightarrow \;\; P(X_i \leq y) = 1 - e^{-\lambda y}$$

which is distribution function of exp (λ) distribution. Hence X_i's are *i.i.d.* exp (λ).

Example 2. *Show that for a random sample of size 2 from $N(0, \sigma^2)$ population,*

$$E[X_{(1)}] = -\sigma/\sqrt{\pi}.$$

Sol. For $n = 2$, the *p.d.f.* $f_1(x)$ of $X_{(1)}$ is given by

$$f_1(x) = \frac{1}{\beta(1, 2)} \{1 - F(x)\} f(x) = 2\{1 - F(x)\} . f(x) \; ; -\infty < x < \infty,$$

where $$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-x^2/2\sigma^2} \qquad | \;\; X \sim N(0, \sigma^2)$$

$$\therefore \qquad E[X_{(1)}] = \int_{-\infty}^{\infty} x . f_1(x)\,dx = 2\int_{-\infty}^{\infty} \{1 - F(x)\} . x f(x)\,dx \qquad ...(1)$$

Also $$\log f(x) = -\log(\sqrt{2\pi}\,\sigma) - \frac{x^2}{2\sigma^2}. \qquad ...(2)$$

Differentiating *w.r.* to x, we get

$$\frac{f'(x)}{f(x)} = -\frac{x}{\sigma^2} \;\; \Rightarrow \;\; \int x f(x)\,dx = -\sigma^2 \int f'(x)\,dx = -\sigma^2 f(x) \qquad ...(3)$$

Integrating (1) by parts and using (3), we get

$$E[X_{(1)}] = 2 . \left[\{1 - F(x)\}\{-\sigma^2 f(x)\}\right]_{-\infty}^{\infty} - 2\int_{-\infty}^{\infty} \{-\sigma^2 f(x)\}\{-f(x)\}\,dx$$

$$= -2\sigma^2 \int_{-\infty}^{\infty} [f(x)]^2\,dx = -\frac{1}{\pi}\int_{-\infty}^{\infty} e^{-x^2/\sigma^2}\,dx$$

$$= -\frac{1}{\pi} . \frac{\sqrt{\pi}}{(1/\sigma)} = -\frac{\sigma}{\sqrt{\pi}} \qquad \left| \int_{-\infty}^{\infty} e^{-a^2x^2}\,dx = \frac{\sqrt{\pi}}{a}\right.$$

NOTES

Example 3. *Show that in odd samples of size n from U[0, 1] population, the mean and variance of the distribution of median are 1/2 and 1/[4(n + 2)] respectively.*

Sol. We have $f(x) = 1 \; ; 0 \le x \le 1$

$$F(x) = P(X \le x) = \int_0^x f(u)\, du = \int_0^x 1.du = -x$$

Let $n = 2m + 1$ (odd), where m is a positive integer ≥ 1. Then median observation is $X_{(m+1)}$. Taking $r = (m + 1)$ the p.d.f. of median $X_{(m+1)}$ is given by :

$$f_{m+1}(x) = \frac{1}{\beta(m+1, m+1)} \cdot x^m (1-x)^m$$

$$\therefore \quad E[X_{(m+1)}] = \frac{1}{\beta(m+1, m+1)} \int_0^1 x\, x^m (1-x)^m\, dx = \frac{\beta(m+2, m+1)}{\beta(m+1, m+1)}$$

$$= \frac{\Gamma(m+2)\,\Gamma(m+1)}{\Gamma(2m+3)} \times \frac{\Gamma(2m+2)}{\Gamma(m+1)\,\Gamma(m+1)} = \frac{m+1}{2m+2} = \frac{1}{2}$$

$$E[X^2_{(m+1)}] = \int_0^1 x^2 f_{m+1}(x)\, dx = \frac{1}{\beta(m+1, m+1)} \cdot \int_0^1 x^{m+2} (1-x)^m\, dx$$

$$= \frac{\beta(m+3, m+1)}{\beta(m+1, m+1)} = \frac{m+2}{2(2m+3)}$$

$$\therefore \quad Var[X_{(m+1)}] = E[X^2_{(m+1)}] - [E\{X_{(m+1)}\}]^2 = \frac{m+2}{2(2m+3)} - \frac{1}{4} = \frac{1}{4(2m+3)} = \frac{1}{4(n+2)}.$$

SUMMARY

- The m.g.f. of the chi-square distribution is $M_X(t) = (1-2t)^{-n/2}$, $|\,2t\,| < 1$.
- X^2 test of independence is used to examine whether the attributes are independent or not
- The 95% confidence limits for the mean μ for t-test are given by $\bar{x} \pm t_{0.05}(S/jn)$
- The points of inflexion of $F(v_1, v_2)$ distribution exist for $v_1 > 4$ and are equidistant from the mode.
- For large value of n, (n is the degree of freedom). X^2 distribution tends to normal distribution.
- For $n \geq 1$, X^2-distribution is +ve skewed.
- The probability curve for the t-distribution is symmetrical about the line $t = 0$.

GLOSSARY

- **chi-square Variate.** The square of a standard normal variate is known as chi-square variate.
- **F-Variate.** If X and Y are two independent chi-square variates with v_1 and v_2 degree of freedom, then, the F-variate is defined by

$$F = \frac{X/v_1}{Y/v_2}$$

NOTES

• **Non-Central *t*-distribution.** If $X \sim N(\mu, \sigma^2)$, Y is independent chi-square variate with n.d.f., then $t = \dfrac{X/\sigma}{\sqrt{Y}/\sqrt{n}}$

has a non-central distribution with n.d.f.

REVIEW QUESTIONS

1. The sales in a supermarket during a week are given below. Test the hypothesis that the sales do not depend on the day of the week, using a significant level of 0.05.

Days	:	Mon	Tues	Wed	Thurs	Fri	Sat
Sales (in 1000 Rs.)	:	65	54	60	56	71	84

2. A survey of 800 families with 4 children each revealed the following distribution :

No. of boys	:	0	1	2	3	4
No. of girls	:	4	3	2	1	0
No. of families	:	32	178	290	236	64

Is this result consistent with the hypothesis that the male and female births are equally probable ?

3. Fit a Poisson distribution to the following data and test the goodness of fit :

x	:	0	1	2	3	4
f	:	109	65	22	3	1

4. The number of scooter accidents per month in a certain town was as follows :

12 8 20 2 14 10 15 6 9 4

Use chi-square test to determine if these frequencies are in agreement with the belief that accident conditions were the same during 10-month period.

5. 500 students at school were graded according to their intelligences and economic conditions of their homes. Examine whether there is any association between economic condition and intelligence, from the following data :

Economic conditions	*Intelligence*	
	Good	*Bad*
Rich	85	75
Poor	165	175

6. In an experiment on the immunisation of goats from anthrox, the following results were obtained. Derive your inferences on the efficiency of the vaccine.

	Died anthrox	*Survived*
Inoculated with vaccine	2	10
Not inoculated	6	6

NOTES

7. A survey among women was conducted to study the family life. The observation were as follows :

	Family Life	
	Happy	Not happy
Educated	70	30
Non Educated	60	40

Test whether there is any association between family life and education.

8. A sample of 300 students of under-graduate and 300 students of post-graduate classes of a university were asked to give their opinion towards the autonomous colleges. 190 of the under-graduate and 210 of the post graduate students favoured the autonomous status. Present the above fact in the form of a frequency table and test that opinions of under-graduate and post-graduate students on autonomous status of colleges are independent.

9. The following values give the lengths of 12 samples of egyptian cotton taken from a consignment : 48, 46, 49, 46, 52, 45, 43, 47, 47, 46, 45, 50. Test if the mean length of the consignment can be taken as 46.

10. A sample of 18 items has a mean 24 units and standard deviation 3 units. Test the hypothesis that it is a random sample from a Normal population with mean 27 units.

11. A random sample of 10 boys had the I.Q's 70, 120, 110, 101, 88, 83, 95, 98, 107 and 100. Do these data support the assumption of a population mean I.Q of 160 ?

12. The mean life of 10 electric motors was found to be 1450 hrs with S.D. of 423 hrs. A second sample of 17 motors chosen from a different batch showed a mean life of 1280 hrs with a S.D. of 398 hrs. Is there a significant difference between means of the two samples ?

13. A group of 10 boys fed on diet A and another group of 8 boys fed on a different diet B recorded the following increase in weight (kgs).

Diet A :	5	6	8	1	12	4	3	9	6	10
Diet B :	2	3	6	8	10	1	2	8		

Does it show the superiority of diet A over the diet B ?

14. To compare the prices of a certain product in two cities, 10 shops when related at random in each town. The price was noted below :

City 1 :	61	63	56	63	56	63	59	56	44	61
City 2 :	55	54	47	59	51	61	57	54	64	58

Test whether the average prices can be said to be the same in two cities.

15. The average number of articles produced by two machines per day are 200 and 250 with standard deviation 20 and 25 respectively on the basis of records of 25 days production. Can you regard both the machines equally efficient at 5% level of significance ?

16. The change in sleeping hours of 7 patients after taking a medicine are as follows :
0.7, 0.1, – 0.3, 1.2, 1.0, 0.3 and – 0.4 hrs.
Do these data give evidence that the medicine produces additional hours of sleep ?
[**Hint.** $|t| = 1.439$]

17. The daily wages in Rupees of skilled workers in two cities are as follows :

	Size of sample of workers	*S.D. of wages in the sample*
City A	16	25
City B	13	32

18. The standard deviation calculated from two random samples of sizes 9 and 13 are 2.1 and 1.8 respectively. Can the samples be regrated as drawn from normal populations with the same standard deviation ?

NOTES

19. Let $X_1, X_2, \ldots\ldots, X_n$ be n independent variates, X_i having a geometric distribution with parameter p_i, *i.e.*, $P(X_i = x_i) = q_i^{x_i - 1} \cdot p_i$; $q_i = 1 - p_i, x_i = 1, 2, 3, \ldots\ldots$

    Show that $X_{(1)}$ is distributed geometrically with parameter $(1 - q_1 q_2 \ldots\ldots q_n)$.

20. For a random sample of size n from a continuous population whose *p.d.f.* $p(x)$ is symmetrical at $x = \mu$, show that $f_r(\mu + x) = f_{n-r+1}(\mu - x)$, where $f_r(.)$ is the *p.d.f.* or $X_{(r)}$.

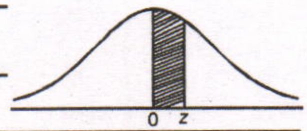
21. Show that the *c.d.f.* of the mid-point (or mid-range) $M = \frac{1}{2}[X_{(1)} + X_{(n)}]$, in random sample of size n from a continuous population with *c.d.f.* $F(x)$ is :

$$F(m) = P(M \le m) = n \int_{-\infty}^{m} [F(2m - x) - F(x)]^{n-1} \cdot f(x)\, dx$$

FURTHER READING

1. Introduction to probability and Mathematical Statistics: V.K. Rohatgi: Wiley Eastern.
2. Discrete Distributions: N.L. Johnson and S.Kotz, John Wiley and Sons
3. Continuous Univarate distributions-1: N.L.Johnson and S.Kotz
4. Continuous Univarate distributins-2: N.L.Johnson and S.Kotz, John Wiley

APPENDIX



NOTES

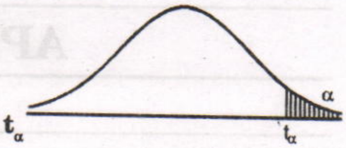
STATISTICAL TABLES

Table I: Area under the Normal curve from 0 to z = Φ (z)

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0754
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2258	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2518	0.2549
0.7	0.2580	0.2612	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2996	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4979	0.4980	0.4981
2.9	0.4981	0.4982	0.4982	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
3.0	0.4987	0.4987	0.4987	0.4988	0.4988	0.4989	0.4989	0.4989	0.4990	0.4990
3.1	0.4990	0.4991	0.4991	0.4991	0.4992	0.4992	0.4992	0.4992	0.4993	0.4993

NOTES

Table II : Values of t_α



υ	α = 0.10	α = 0.05	α = 0.025	α = 0.01	α = 0.005
1	3.078	6.314	12.706	31.821	63.657
2	1.886	2.920	4.303	6.965	9.925
3	1.638	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
6	1.440	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499
8	1.397	1.860	2.306	2.896	3.355
9	1.383	1.833	2.262	2.821	3.250
10	1.372	1.812	2.228	2.764	3.169
11	1.363	1.796	2.201	2.718	3.106
12	1.356	1.782	2.179	2.681	3.055
13	1.350	1.771	2.160	2.650	3.012
14	1.345	1.761	2.145	2.624	2.977
15	1.341	1.753	2.131	2.602	2.947
16	1.337	1.746	2.120	2.583	2.921
17	1.333	1.740	2.110	2.567	2.898
18	1.330	1.734	2.101	2.552	2.878
19	1.328	1.729	2.093	2.539	2.861
20	1.325	1.725	2.086	2.528	2.845
21	1.323	1.721	2.080	2.518	2.831
22	1.321	1.717	2.074	2.508	2.819
23	1.319	1.714	2.069	2.500	2.807
24	1.318	1.711	2.064	2.492	2.797
25	1.316	1.708	2.060	2.485	2.787
26	1.315	1.706	2.056	2.479	2.779
27	1.314	1.703	2.052	2.473	2.771
28	1.313	1.701	2.048	2.467	2.763
29	1.311	1.699	2.045	2.462	2.756

Table III : Values of χ^2 with level of significance α and degrees of freedom ν

ν \ α	0.99	0.95	0.50	0.30	0.20	0.10	0.05	0.01
1	0.0002	0.004	0.46	1.07	1.64	2.71	3.84	6.64
2	0.020	0.103	1.39	2.41	3.22	4.60	5.99	9.21
3	0.115	0.35	2.37	3.66	4.64	6.25	7.82	11.34
4	0.30	0.71	3.36	4.88	5.99	7.78	9.49	13.28
5	0.55	1.14	4.35	6.06	7.29	9.24	11.07	15.09
6	0.87	1.64	5.35	7.23	8.56	10.64	12.59	16.81
7	1.24	2.17	6.35	8.38	9.80	12.02	14.07	18.48
8	1.65	2.73	7.34	9.52	11.03	13.36	15.51	20.09
9	2.09	3.32	8.34	10.66	12.24	14.68	16.92	21.67
10	2.56	3.94	9.34	11.78	13.44	15.99	18.31	23.21
11	3.05	4.58	10.34	12.90	14.63	17.28	19.68	24.72
12	3.57	5.23	11.34	14.01	15.81	18.55	21.03	26.22
13	4.11	5.89	12.34	15.12	16.98	19.81	22.36	27.69
14	4.66	6.57	13.34	16.22	18.15	21.06	23.68	29.14
15	5.23	7.26	14.34	17.32	19.31	22.31	25.00	30.58
16	5.81	7.96	15.34	18.42	20.46	23.54	26.30	32.00
17	6.41	8.67	16.34	19.51	21.62	24.77	27.59	33.41
18	7.02	9.39	17.34	20.60	22.76	25.99	28.87	34.80
19	7.63	10.12	18.34	21.69	23.90	27.20	30.14	36.19
20	8.26	10.85	19.34	22.78	25.04	28.41	31.41	37.57
21	8.90	11.59	20.34	23.86	26.17	29.62	32.67	38.93
22	9.54	12.34	21.34	24.94	27.30	30.81	33.92	40.29
23	10.20	13.09	22.34	26.02	28.43	32.01	35.01	41.64
24	10.86	13.85	23.34	27.10	29.55	33.20	36.42	42.98
25	11.52	14.61	24.34	28.17	30.68	34.68	37.65	44.31
26	12.20	15.38	25.34	29.25	31.80	35.56	38.88	45.64
27	12.88	16.15	26.34	30.32	32.91	36.74	40.11	46.96
28	13.56	16.93	27.34	31.39	34.03	37.92	41.34	48.28
29	14.26	17.71	28.34	32.46	35.14	39.09	42.56	49.59
30	14.95	18.49	29.34	33.53	36.25	40.26	43.77	50.89

NOTES

NOTES

Table IV : Values of $F_{0.05}$

ν_2 = Degrees of freedom for denominator	ν_1 = Degrees of freedom for numerator																
	1	2	3	4	5	6	7	8	9	10	12	15	20	25	30	40	60
1	161	200	216	225	230	234	237	239	241	242	244	246	248	249	250	251	252
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.41	19.43	19.45	19.46	19.46	19.47	19.48
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.70	8.66	8.63	8.62	8.59	8.57
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.86	5.80	5.77	5.75	5.72	5.69
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.52	4.50	4.46	4.43
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.94	3.87	3.83	3.81	3.77	3.74
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.51	3.44	3.40	3.38	3.34	3.30
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.22	3.15	3.11	3.08	3.04	3.01
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01	2.94	2.89	2.86	2.83	2.79
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.85	2.77	2.73	2.70	2.66	2.62
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.79	2.72	2.65	2.60	2.57	2.53	2.49
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.62	2.54	2.50	2.47	2.38	2.38
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.60	2.53	2.46	2.41	2.38	2.34	2.30
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2.46	2.39	2.34	2.31	2.27	2.22
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40	2.33	2.28	2.25	2.20	2.16
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.35	2.28	2.23	2.19	2.15	2.11
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.38	2.31	2.23	2.18	2.15	2.10	2.06
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.27	2.19	2.14	2.11	2.06	2.02
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.23	2.16	2.11	2.07	2.03	1.98
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.20	2.12	2.07	2.04	1.99	1.95
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.25	2.18	2.10	2.05	2.01	1.96	1.92
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.23	2.15	2.07	2.02	1.98	1.94	1.89
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.20	2.13	2.05	2.00	1.96	1.91	1.86
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.18	2.11	2.03	1.97	1.94	1.89	1.84
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.16	2.09	2.01	1.96	1.92	1.87	1.82
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	2.01	1.93	1.88	1.84	1.79	1.74
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.00	1.92	1.84	1.78	1.74	1.69	1.64
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.92	1.84	1.75	1.69	1.65	1.59	1.53

NOTES

Table V : Values of $F_{0.01}$

ν_2 = Degrees of freedom for denominator	ν_1 = Degrees of freedom for numerator																
	1	2	3	4	5	6	7	8	9	10	12	15	20	25	30	40	60
1	4,052	5,000	5,403	5,625	5,764	5,859	5,928	5,982	6,023	6,056	6,106	6,157	6,209	6,240	6,261	6,287	6,313
2	98.50	99.00	99.17	99.25	99.30	99.33	99.36	99.37	99.39	99.40	99.42	99.43	99.45	99.46	99.57	99.47	99.48
3	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.35	27.23	27.05	26.87	26.69	26.58	26.50	26.41	26.32
4	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66	14.55	14.37	14.20	14.02	13.91	13.84	13.75	13.65
5	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16	10.05	9.89	9.72	9.55	9.45	9.38	9.29	9.20
6	13.75	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87	7.72	7.56	7.40	7.30	7.23	7.14	7.06
7	12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72	6.62	6.47	6.31	6.16	6.06	5.99	5.91	5.82
8	11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91	5.81	5.67	5.52	5.36	5.26	5.20	5.12	5.03
9	10.56	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35	5.26	5.11	4.96	4.81	4.71	4.65	4.57	4.48
10	10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85	4.71	4.56	4.41	4.31	4.25	4.17	4.08
11	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63	4.54	4.40	4.25	4.10	4.01	3.94	3.86	3.78
12	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39	4.30	4.16	4.01	3.86	3.76	3.70	3.62	3.54
13	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19	4.10	3.96	3.82	3.66	3.57	3.51	3.43	3.34
14	8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14	4.03	3.94	3.80	3.66	3.51	3.41	3.35	3.27	3.18
15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80	3.67	3.52	3.37	3.28	3.21	3.13	3.05
16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78	3.69	3.55	3.41	3.26	3.16	3.10	3.02	2.93
17	8.40	6.11	5.18	4.67	4.34	4.10	3.93	3.79	3.68	3.59	3.46	3.31	3.16	3.07	3.00	2.92	2.83
18	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60	3.51	3.37	3.23	3.08	2.98	2.92	2.84	2.75
19	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52	3.43	3.30	3.15	3.00	2.91	2.84	2.76	2.67
20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46	3.37	3.23	3.09	2.94	2.84	2.78	2.69	2.61
21	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40	3.31	3.17	3.03	2.88	2.79	2.72	2.64	2.55
22	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35	3.26	3.12	2.98	2.83	2.73	2.67	2.58	2.50
23	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30	3.21	3.07	2.93	2.78	2.69	2.62	2.54	2.45
24	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26	3.17	3.03	2.89	2.74	2.64	2.58	2.49	2.40
25	7.77	5.57	4.68	4.18	3.85	3.63	3.46	3.32	3.22	3.13	2.99	2.85	2.70	2.60	2.54	2.45	2.36
30	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98	2.84	2.70	2.55	2.45	2.39	2.30	2.21
40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89	2.80	2.66	2.52	2.37	2.27	2.20	2.11	2.02
60	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63	2.50	2.35	2.20	2.10	2.03	1.94	1.84