B.A. Third Year
Mathematics

Course Code : MATH313TH
New Syllabus : CBCS

Probability and Statistics

Units 1 to 20

Centre for Distance and Online Education (CDOE)
Himachal Pradesh University,
Gyan Path, Summer Hill, Shimla - 171005

MATH313TH

Probability and Statistics

CONTENTS

SYLLABUS

Himachal Pradesh University

B.A. with Mathematics

Syllabus and Examination Scheme

Course Code	MATH313TH
Credits	4
Name of the Course	Probability and Statistics
Type of the Course	Skill Enhancement Course
Assignments	Max. Marks:30
Yearly Based Examination	Max Marks: 70
	Maximum Times: 3 hrs.

Instructions

Instructions for Candidates: Candidates are required to attempt five questions in all. Section A is Compulsory and from Section B they are required to attempt one question from each of the Units I, II, III and IV of the question paper.

SEC 3 : Probability and Statistics

Unit-I

Sample space, probability axioms, real random variables (discrete and continuous), cumulative distribution function, probability mass/density functions.

Unit-II

Mathematical expectation, moments, moment generating function, characteristic function, discrete distributions : uniform

Unit-III

Binomial, Poisson, continuous distributions : uniform, normal, exponential.

Unit-IV

Joint cumulative distribution function and its properties, joint probability density functions, marginal and conditional distributions, expectation of function of two random variables, conditional expectations, independent random variables.

Books Recommended

1. Robert V. Hogg, Joseph W. Mckean and Allen T. Craig, *Introduction to Mathematical Statistics*, Pearson Education, Asia, 2007.
2. Irwin Miller and Marylees Miller, John E. Freund, *Mathematical Statistics with Application*, 7th Ed., Pearson Education, Asia, 2006.
3. Sheldon Ross, *Introduction of Probability Model*, 9th Ed., Academic Press Indian Reprint, 2007.

Unit - 1

Basic Concepts of Probability

Structure

1.1 Introduction

If an experiment is repeated under essentially homogeneous and similar conditions, then we generally come across two types of situations.

(i) The result or what is usually known as the outcome is unique or certain.

(ii) The result is not unique but may be one of the several possible outcomes.

The phenomena covered by (i) are known as deterministic or predictable phenomena. By a deterministic phenomenon we mean one in which the result can be predicted with certainty e.g.

(a) For a perfect gas

$$V \alpha \frac{1}{P}$$

i.e. PV = constant .....(1)

Where V is one volume and P is the pressure of the gas,

provided the temperature remains the same.

(b) The velocity 'ν' of a particle after time t is given by

$$\nu = u + at, \quad ......(2)$$

where u is the initial velocity and a is the acceleration. Equation (2) uniquely determines ν if the right hand quantities are known.

(c) Ohm's law, viz.,

$$c = \frac{E}{R}, \qquad \text{......(3)}$$

where c is the flow of current, E the potential difference between the two ends of the conductor and R the resistance, uniquely determines the value c as soon as E and R are given.

A deterministic model is defined as a model, which stipulates that the conditions under which an experiment is performed determine the outcome of the experiment. For a number of situations the deterministic model suffices.

However, there are phenomena, as covered by (ii) above, which do not lend themselves to deterministic approach and are known as unpredictable or probabilistic phenomena, e.g.,

(i) In tossing of a coin one is not sure if a head or a tail will be obtained.

(ii) If a light-tube has lasted for m hours, nothing can be said about its further life. It may fail to function any moment.

In such cases we talk of chance or probability which is taken to be a quantitative measure of certainty.

The word probability may be used in two different contest. Firstly, it may be used in regard to some proposition. Take, for instance, the statement, "It is very probable that India will adhere to the democratic system of government till the end of this century" or, "It is very improbable that the county's brain drain will stop in the near future". Probability here means the degree of belief in the proposition of the person making the statement. This is called the subjective probability.

Alternatively, the word may be used in regard to the results of an experiment that can, conceivable, be repeated an infinite number of times under essentially similar conditions. The results will be called events. The probability of an event here refers to the proportion of cases in which the event occurs in such repetitions of the experiments. This type of probability is called the objective probability, being a part of the real world, and it is with this sense of the word that we shall be concerned in the present discussion.

1.2 Learning Objectives

After reading this unit, you should be able to:-

- Discuss the various terms that are frequently used in the theory of probability.
- Discuss the concept of probability and discuss to important approaches by means of cohich we can estimate the probability of an event i.e. discuss the classical approach or 'A Priori' approach and statistical or empirical probability.
- Discuss the limitations of classical definition of probability and empirical definition of probability.
- Do some basic questions of probability.

1.3 Basic Terminology

The following terms are frequently used in the theory of probability:-

If an experiment, when repeated under identical conditions, does not produce the same outcome every time but the outcome in a trial is one of the several possible outcome, then such an experiment is called a random experiment.

Examples of random experiments of are : fossing a coin, throwing a die, selecting a card from a pack of playing cards etc. In all these cases, there are number of possible results which can occur but there is an uncertainty as to which one of them will actually occur.

Outcome

The result of a random experiment coill be called an outcome.

Trial and Event

Any particular performance of a random experiment is called a trial and outcome or combinations of outcomes are termed as events e.g. is a coin is tossed repeatedly, the result is not unique-we may get any one of the two faces, heat or tail. Thus tossing of a coin is a random experiment or trial and getting of a head or tail is an event.

Elementary Event

If a random experiment is performed, them each of its outcomes is known as an elementary event.

Sample Space

A sample space is defined as the set of all possible outcomes of an experiment and is denoted by S. For example:-

(i) In tossing a coin, sample space is given by

$S = \{H, T\}$

(ii) In tossing two coins simultaneously, sample space is given by

$S = \{(H,T) \times (H, T)\} = \{HH, HT, TH, TT\}$

(iii) In tossing three coins simultaneously, sample space is given by

$S = \{(H,T) \times (H,T) \times (H,T)\}$

$= \{(HH, HT, TH, TT) \times (H,T)\}$

$= \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$

(iv) In throwing a die, sample space is

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

(v) In throwing a die twice or in a single throw of two die, same space is

Sample Point

Elements of sample space S are known as sample points.

If we roll a die, them

S = {1,2,3,4,5,6}

and sample points are 1,2,3,4,5.6

Event

A subset of the sample space associated with a random experiment is called an event.

Certain (or Sure) Event

An event associated with a random experiment is called a certain event if it always occurs whenever one experiment is performed.

Impossible Event

An event associated with a random experiment is called an impossible event if it never occurs whenever the experiment is performed.

Compound Event

An event associated with a random experiment is a compound event if it is the disjount union of two or more elementary events.

Exhaustive Events or Cases

The total number of possible outcomes of a random experiment is known as the exhaustive events or cases. For example,

(i) In tossing of a coin, there are two exhaustive cases, viz., head and tail come possibility of the coin standing on an edge being ignored.

(ii) In throwing of a die, there are 6 exhaustive events since any of the six faces may come uppermost.

(iii) In drawing two cards from a pack of cards, the exhaustive number of events is 52_{c_2}, since 2 cards can be drawn out of 52 cards in 52_{c_2} ways.

(iv) In throwing of two dice, the exhaustive number of cases is 6^2 since any of the 6 numbers 1 to 6 on the first die can be associated with any of the 6 numbers on the other die. In general, in throwing of dice, n the exhaustive number of cases is 6^n.

Mutually Exclusive Events

Events are said to be mutually exclusive or incompatible, if the happening of any one of them precludes the happening of all the others i.e. if no two or more of them can happen simultaneously in the same trial. For example:-

(i) In tossing a coin the events head and till are mutually exclusive

(ii) In throwing a die, all the 6 faces numbered 1 to 6 are mutually exclusive since if any one of these 6 faces comes, the possibility of others in the same trial, is ruled out.

Mutually Exclusive and Exhaustive System of Events

If there are events such that one of them must occur and the occurrence of one rules out the possibilities of the occurrence of the others, them the events are said to be mutually exclusive and exhaustive system of events.

Favorable Events or Cases

The number of cases favorable to an event in a trial is the number of outcomes which ensure the happening of the event. For example:

(i) In drawing a card from a pack of cards the number of cases favorable for drawing of a queen is 4, for drawing a spade is 13 and for drawing a black card is 26.

(ii) In throwing of two dice, the number of cases favorable to get a sum 5 is (1, 4), (4, 1), (2, 3), (3, 2) i.e. 4.

Equally Likely Events

Outcomes of trial are said to be equally likely is taking into consideration all the relevant evidence, there is no reason to expect any one in preference to other. For example:

(i) In tossing an unbiased coin, head or tail are equally likely events.

(ii) In throwing an unbiased die, all the six faces are equally likely to come.

Complimentary event (or negation) of E

Given an event E, the event which occurs when, and only when, E does not occur is called the event "not E". This event "not E" is also called the complimentary event of E or negation of E and is denoted by E^c. For example, consider die is thrown. The two events. "the number is even" and "the number is odd" are such that at least one of the events has to occur and only one occurs. If the first event does not occur, then the second must occur and the non-occurrence of the second means the first event must have occurred.

Simple Event, Compound Event

A single event is called a simple event and when two or more than two events occur in connection with each other then their simultaneous occurrence is called a compound event.

1.4 The Concept of Probability

In any random experiment there is always uncertainty as to whether a particular event will or will not occur. As a measure of the chance or probability, with which we can expect the event to occur, it is convenient to assign a number between 0 and 1.

If we are sure or certain that one event will occur, we say that its probability is 100% or 1, but if we are sure that the event will not occur, we say that its probability is zero. If, for example, the probability is 1/4. We could say that there is 25% chance it will occur and a 75% chance that it will not occur. Equivalently, we can say that the odds against its occurrences are 75% to 25% or 3 to 1. There are two important approaches by means of which we can estimate the probability of an event.

1. Classical Approach or 'A Priors' Approach

If a random experiment or a trial results in 'n' exhaustive, mutually exclusive and equally likely outcomes (or cases), out of which m are favorable to the occurrence of an event E, then the probability 'p' of occurrence (or happing) of E, usually denoted by P(E), is given by,

$$p = P(E) = \frac{Number\ of\ Favourable\ Cases}{Total\ number\ of\ exhaustive\ cases} = \frac{m}{n} \qquad ......(1)$$

Remarks

(i) Since m $\geq$ 0, n > 0 and m $\leq$ n, we get from (1) P(E) $\geq$ 0 and P(E) $\leq$ 1 $\Rightarrow$ 0 $\leq$ P(E) $\leq$ 1

(ii) Sometimes we express (1) by saying the odds in favour of E are m : (n-m) or the odds against E are (n-m) : m^1.

(iii) The non-happening of the event E is called the complementary event of E and is denoted by $\overline{E}$ or E^c.

(iv) Probability 'p' of the happening of an event is also known as the probability of success and the probability 'q' of the non-happening of the event as the probability of failure i.e. p + q = 1.

(v) If P(E) = 1, E is called a certain event and if P(E) = 0, E is called an impossible event.

2. Statistical or Empirical Probability

If an experiment is performed repeatedly under essentially homogeneous and identical conditions, them the limiting value of the ratio of the number of times the event occurs to the number of trials, as the number of trials becomes indefinitely large, is called the probability of happening of the event, it being assumed that the limit is finite and unique.

Symbolically, if in N trials an event E happens M times, then the probability of the happening of E, denoted by P(E), is given by:

$$P(E) = \lim_{N \to \infty} \frac{M}{N} \qquad .....(2)$$

Limitations of These Definitions

(I) The classical definition breaks down if

(i) The various outcomes of the random experiment are not equally likely. For example:

(a) The probability that a ceiling fan in a room will fall is not 1/2, since the events of the fan 'falling' and 'not falling' though mutually exclusive and exhaustive, are not equally likely. In fact, the probability of the fan falling will be almost zero.

(b) If a person jumps from a running train, then the probability of his survival will not be 50%, since in this case the events survival and death, though exhaustive and mutually exclusive, are not equally likely.

(ii) The exhaustive number of outcomes of the random experiment is infinite or unknown.

II. The Empirical definition has the following limitations:

(i) For large repetition of any experiment, the experimental conditions may not remain identical and homogeneous.

(ii) The result generally differs for different set of experiments of the same type:

Let us improve our understanding of these results by looking at some following examples:-

Example 1: Three coins are tossed once. find the probability that

(i) head and tails appear alternatively

(ii) at least one head and one tail occur.

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

$\therefore$ Total number of outcomes = 8

(i) Favourable cases are HTH, THT

$\therefore$ Number of favourable cases = 2

$\therefore$ Required probability = $\frac{2}{8} = \frac{1}{4}$

(ii) Favourable cases are HHT, HTH, THH, TTH, THT, HTT

$\therefore$ Number of favourable cases = 6

$\therefore$ Required probability = $\frac{6}{8} = \frac{3}{4}$

Example 2: In a single throw of two unbiased dice, what is the probability of obtaining:

(i) a total of 7 ? (ii) a total of 13 ?

(iii) a total as even number ?

Sol. We have

$$S = \begin{Bmatrix} (1,1) & (1,2) & (1,3) & (1,4) & (1,5) & (1,6) \\ (2,1) & (2,2) & (2,3) & (2,4) & (2,5) & (2,6) \\ (3,1) & (3,2) & (3,3) & (3,4) & (3,5) & (3,6) \\ (4,1) & (4,2) & (4,3) & (4,4) & (4,5) & (4,6) \\ (5,1) & (5,2) & (5,3) & (5,4) & (5,5) & (5,6) \\ (6,1) & (6,2) & (6,3) & (6,4) & (6,5) & (6,7) \end{Bmatrix}$$

$\therefore$ Total number of outcomes = 36

(i) Favourable outcomes are (1,6), (2,5), (3,4), (4,3), (5,2), (6,1)

$\therefore$ Number of favourable outcomes = 6

$\therefore$ Required probability = $\frac{6}{36} = \frac{1}{6}$

(ii) A total of 13 is an impossible event as the sum f two numbers on the two dice can not be 13.

$\therefore$ Number of favourable outcomes = 0

$\therefore$ Required probability = $\frac{0}{36} = 0$

(iii) Favourable outcomes are

(1,1), (1,3), (1,5), (2,2), (2,4), (2,6), (3,1), (3,3), (3,5), (4,2), (4,4), (4,6), (5,1), (5,3), (5,5), (6,2), (6,4), (6,6).

$\therefore$ Number of favourable outcomes = 18

$\therefore$ Required probability = $\frac{18}{36} = \frac{1}{2}$

Example 3: What is the choice that a leap year selected at random will contain 53 Sundays?

Sol. Leap Year contains 366 days.

$\therefore$ There are 52 complete weeks and two days other. The following are the possibilities of these two 'over' days:

(i)	Sunday and Monday	(ii)	Monday and Tuesday
(iii)	Thursday and Friday	(iv)	Wednesday and Thursday
(v)	Thursday and Friday	(vi)	Friday and Saturday
(vii)	Saturday and Sunday		

Now there will be 53 Sundays in a leap year when one of the two over days is a Sunday.

$\therefore$ out of 7 possibilities, two are favourable to this event.

$\therefore$ Required probability = $\frac{2}{7}$

Example 4: Tickets are numbered from 1 to 10. Two tickets are drawn one after the other with replacement. Find the probability that the number on one of the tickets is a multiple of 5 and the other a multiple of 4.

Sol. Two tickets are drawn from tickets numbered 1 to 10.

$\therefore$ S = {(1,1), (1,2), (1,3),......, (10,9), (10,10)}

$\therefore$ Total number of possible outcomes

$= 10 \times 10 = 100$

Let E denote the event that number on one of the tickets is a multiple of 5 and the number on the other is a multiple of 4.

$\therefore$ A = {(5,4), (5,8), (10,4), (10,8), (4,5), (4,10), (8,5), (8,10)}

$\therefore$ Number of favourable cases = 8

$\therefore$ Required probability = $\frac{8}{100} = \frac{2}{25}$

Example 5: In a random arrangement of the letters of word 'Mathematics', find the probability that all the vowels are together.

Sol. Given word is 'MATHEMATICS'

No. of given letters = 11

No. of M's = 2

No. of T's = 2

No. of A's = 2

$\therefore$ Required number of arrangements = $\frac{\lfloor 11}{\lfloor 2 \lfloor 2 \lfloor 2}$

$= \frac{11\times10\times9\times8\times7\times6\times5\times4\times3\times2\times1}{(2\times1)\times(2\times1)\times(2\times1)} = 4989600$

$\therefore$ Total number of cases = 4989600

Consider the four vowels as one letter.

$\therefore$ 8 letters can be arranged in $\frac{\lfloor 8}{\lfloor 2 \lfloor 2}$ ways.

Also four vowels can be arranged in $\frac{\lfloor 4}{\lfloor 2}$ ways.

$\therefore$ Number of words in which vowels are always together = $\frac{\lfloor 8}{\lfloor 2 \lfloor 2} \times \frac{\lfloor 4}{\lfloor 2}$

$= \frac{8\times7\times6\times5\times4\times3\times2\times1}{(2\times1)\times(2\times1)} \times \frac{4\times3\times2\times1}{2\times1} = 120960$

$\therefore$ Number of favourable cases = 120960

$\therefore$ Required probability = $\frac{120960}{4989600} = \frac{4}{165}$

Example 6: 20 books are placed at random in a shelf. Find the probability that a particular pair of books is

(i) always together (ii) never together

Sol: Total no. of books = 20

Number of ways in which these can be arranged = $\lfloor\underline{20}$

$\therefore$ Total number of cases = $\lfloor\underline{20}$

(i) Consider the two particular books as one. So 19 books can be arranged in $\lfloor\underline{19}$ ways. Also two particular books can be arranged in themselves in $\lfloor\underline{2}$ = 2 × 1 = 2 ways.

$\therefore$ Number of ways in which two particular books are always together = $2\lfloor\underline{19}$

$\therefore$ Number of favourable cases = $2\lfloor\underline{19}$

$\therefore$ Required probability = $\frac{2\lfloor\underline{19}}{20\times\lfloor\underline{19}} = \frac{1}{10}$

(ii) Number of ways in which 2 particular books are never together = $\lfloor\underline{20}$ - $2\lfloor\underline{19}$

= $20\times\lfloor\underline{19}$ - $2\lfloor\underline{19}$ = $18\lfloor\underline{19}$

$\therefore$ Number of favourable cases = $18\lfloor\underline{19}$

$\therefore$ Required probability = $\frac{18\lfloor\underline{19}}{\lfloor\underline{20}} = \frac{18\lfloor\underline{19}}{20\times\lfloor\underline{19}} = \frac{9}{10}$

Example 7: Find the probability that when a hand of 7 cards is dealt from a well-shuffled deck of 52 cards, it contains.

(i) all 4 kings (ii) exactly 3 kings

(iii) at least 3 kings

Sol: Total number of cards = 52

Number of cards to be taken = 7

$\therefore$ Total number of possible hands = 52_{c_7}

$\therefore$ Total number of possible hands = has 4 kings

(i) P (a hand has 4 kings) = P (a hand has 4 kings and three cards from non-kings)

$$= \frac{4_{c_4} \times 48_{c_3}}{52_{c_7}}$$

$$= \frac{1 \times \frac{48 \times 47 \times 46}{1 \times 2 \times 3}}{\frac{52 \times 51 \times 50 \times 49 \times 48 \times 47 \times 46}{1 \times 2 \times 3 \times 4 \times 5 \times 6 \times 7}}$$

$$= \frac{1}{7735}$$

(ii) Number of hands with 3 kings and 4 cards from non-kings = $\frac{4_{c_3} \times 48_{c_4}}{52_{c_7}}$

$$= \frac{\frac{4}{1} \times \frac{48 \times 47 \times 46 \times 45}{1 \times 2 \times 3 \times 4}}{\frac{52 \times 51 \times 50 \times 49 \times 48 \times 47 \times 46}{1 \times 2 \times 3 \times 4 \times 5 \times 6 \times 7}}$$

$$= \frac{9}{1547}$$

(iii) P (at least 3 kings) = P(3 kings) + P(4 kings)

$$= \frac{1}{7735} + \frac{9}{1547}$$

$$= \frac{1+45}{7735} = \frac{46}{7735}$$

Dear Students, now try the following exercises:-

1.4 Self Check Exercise

Q.1 In a lottery of 50 tickets numbered 1 to 50, two tickets are drawn simultaneously. Find the probability when

(i) both the tickets drawn have prime numbers

(ii) none of the tickets drawn has prime number.

Q.2 If n biscuits are distributed among N beggars, find the probability that a particular beggar receives r (< n) biscuits.

1.5 Summary

We conclude this unit by summarizing what we have covered in it:-

(1) Defined and discussed the various terms that are used in the theory of probability.

Various terms defined are: Random experiment, outcome, trial, event, sample space, sample point, certain event, impossible event, compound event, exhaustive, events, mutually exclusive events, favourable events. equally likely events, complimentary events, simple event, compound event.

(2) Discussed the concept of probability by classical approach or 'A Priori' approach and statistical or empirical probability.

(3) Did some basic questions of probability.

1.6 Glossary:

1. If an experiment, when repeated under identical conditions, does not produce the same outcome every time but the outcome in a trial is one of the several possible outcomes, these such as experiment is called a random experiment.
2. Any particular performance of a random experiment is called a trial and outcome or combination of outcomes are termed as events.
3. A sample space is defined as the set of all possible outcomes of an experiment
4. An event associated with a random experiment is called a certain event if it always occurs whenever the experiment is performed.
5. The total number of possible outcomes of a random experiment is known as the exhaustive events or cases:
6. Events are said to be mutually exclusive if the happening of any one of them precludes the happening of all the others.

1.7 Answer to Self Check Exercise

Ans.1 $\frac{21}{245}$

Ans.2 $\frac{119}{245}$

Ans.3 Required probability = $\frac{n_{c_r} \times (N-1)^{n-r}}{N^n}$

1.8 References/Suggested Readings

1. Robert V. Hogg, Joseph w. Mckean and Allen T. Craig, Introduction to Mathematical Statistics, Pearson Education, Asia, 2007.
2. Sheldon Ross, Introduction to Probability Model, 9th Ed., Academic Press, Indian Reprint, 2007
3. Irwin Miller, Marylees Miller and John E. Freund, Mathematical Statistics with Application, 7th Ed., Pearson Education, Asia, 2006.

1.9 Terminal Questions

1. Two cards are drawn together from a back of 52 cards at random. What is the probability that

   (i) both are spades?

   (ii) both are kings?

   (iii) exactly one is king?

2. If p, q are chosen randomly from the set :1,2,3,4,5,6,7,8,9,10}. Determine the probability that roots of equations $x^2 + p\ x + q = 0$ are real.

3. A fair coin is tossed four times, and a person wins Rs. 1 for each head and loses Rs. 1.50 for each tail that turns up. Calculate how many different amounts of money he can have after four tosses and the probability of having each of these amounts.

Unit - 2

Probability Axioms

Structure

2.1 Introduction

The probability axioms are set of fundamental principles that govern one mathematical theory of probability. These axioms were developed to provide a rigorous and consistent framework for defining and analyzing probabilities. The three main probability axioms are : Non-negativity axiom; Normalization axiom; Additively axiom. These three axioms form the foundation of probability theory and are used to derive various properties and theorem in probability. These probability axioms provide a solid mathematical foundation for probability theory and enable the development of more advanced concepts and applications in the field of probability and statistics.

2.2 Learning Objectives

After studying this unit, you should be able to:

- Define and discuss the three axioms of probability.
- Define and discuss the probability set function.
- Prove some important theorems of probability.
- Do questions related to probability by using theorems.

2.3 Axioms of Probability And The Probability Set Function

Suppose we have a sample space s. If S is discrete, all subsets correspond to events and conversely, but if S is non-discrete, only special subsets (called measurable) correspond to events. To each event A in the class C of events, we associate real number P(A). Then P is called a probability set function, and P(A) the probability of the event A, if the following axioms are satisfied:

Axiom 1 : For every event A in the class C, $P(A) \geq 0$. .....(1)

Axiom 2 : For the sure or certain event in the class C, P(S) = 1 ......(2)

Axiom 3 : For any number of mutually exclusive events $A_1, A_2, ...$, in the class C,

$$P(A_1 \cup A_2 \cup ....) = P(A_1) + P(A_2) + ... \qquad ....(3)$$

In particular, for two mutually exclusive events A_1, A_2,

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) \qquad .....(4)$$

The above three axioms give the Axiomatic definition of probability.

A more rigorous definition of a Probability set Function is:

P(A) is called the probability set function defined on a σ - field $\subset$ of events if the following properties or axioms hold:-

1. Axiom of non-negativity : For each $A \in C$, P(A) is defined, is real and P(A) > 0
2. Axiom of certainty : P (s) = 1
3. Axiom of additivity : If {An} is any finite or infinite sequence of disjoint events in c, then

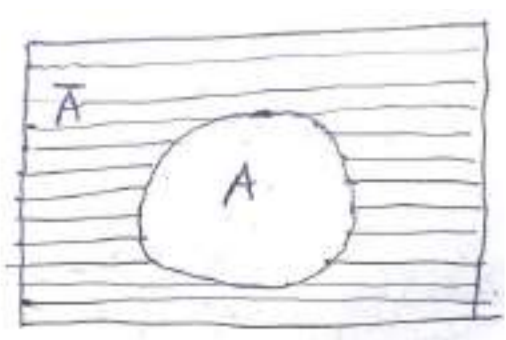
$$P = \left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i)$$

2.4 Some Important Theorems on Probability

Theorem 1 : For each a belonging to the class c of events,

$$P(\bar{A}) = 1 - P(A)$$

or $$P(A) = 1 - P(\bar{A})$$



Proof : Let S be the sample space .

Since $A \cap \bar{A} = \phi$, A and $\bar{A} \subset S$

$\therefore \quad S = A \cup \bar{A}$

By axioms (2) and (3), we have

$P(s) = 1$

and $\quad P(A \cup \bar{A}) = P(A) + P(\bar{A})$

Thus, we get

$1 = P(a) + P(\bar{A})$

$\Rightarrow \quad P(\bar{A}) = 1 - P(a)$

Or equivalently

$P(A) = 1 - P(\bar{A})$

Theorem 2 : The probability of the null set or empty set is zero, that is $P(\phi) = 0$

Proof : Let S be the sample space.

Then $S \cap \phi = \phi$ and $S \cup \phi = S$

$\Rightarrow \quad P(S \cup \phi) = P(S)$

Now Axiom (3) $\Rightarrow P(S \cup \phi) = P(S) + P(\phi)$

and Axiom (2) $\Rightarrow P(S) = 1$

$\therefore$ Equation (1) becomes

$P(S) + P(\phi) = P(S)$

$\Rightarrow \quad 1 + \Pi(\phi) = 1$

$\Rightarrow \quad P(\phi) = 0$

Hence the result

Theorem 3 : For any two events A and B in the class C of events, show that if $B \subseteq A$, then

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

(ii) $\quad P(B) \leq P(A)$

Proof : (i) Since $B \subseteq A$, therefore B and $A \cap \overline{B}$ are mutually exclusive events,

i.e. $\quad B \cap (C \cap \overline{B}) = \phi$

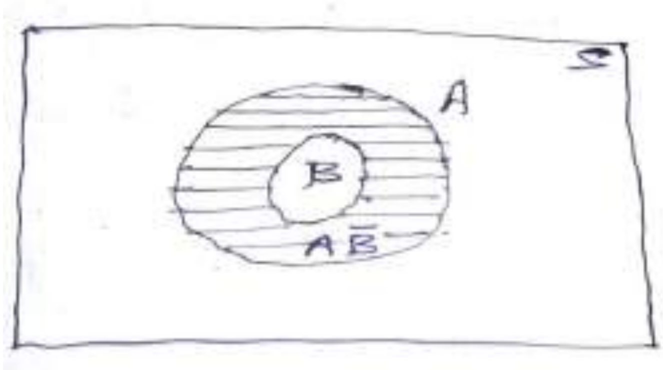
Moreover,

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

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

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

$\Rightarrow$ $P(A\cap \bar{B}) = P(A) - P(B)$ ....(1)



Since $A \cap \bar{B} = A - B$

Therefore, (i) can also be written as

$P(A - B) = P(A) - P(B)$

(ii) From Axiom (i), probability of any event in s is greater than or equal to zero.

In Particular, $P(A - B) \geq 0$

$\therefore$ Using (2), we have $P(A) - P(B) \geq 0$

or $P(B) < P(A)$ ....(3)

Hence the result

Theorem 4 : For every event A,

$0 \leq P(A) \leq 1$,

i.e. a probability is between 0 and 1

Proof : We know that for any event A in the class C of events, $A \subset S$ where S is the sample space.

Now $P(A) \leq P(S)$

But $P(S) = 1$ [By Axiom (2)]

$\therefore$ $P(A) \leq 1$

Also, by Axiom (1), we have

$P(A) \geq 0$

$\therefore$ On combining these two, we get

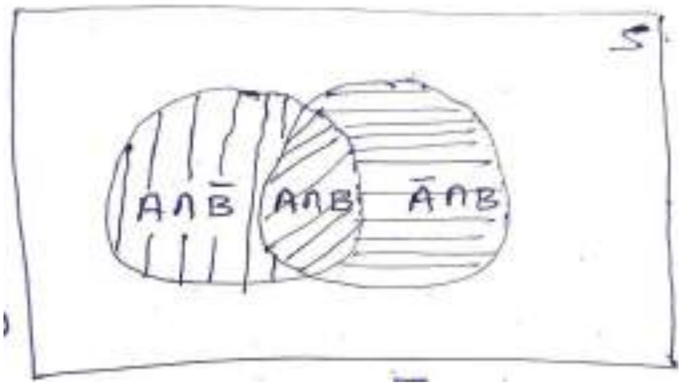
$0 \leq P(A) \leq 1$

Hence the result.

Theorem 5 : For any two events A and B, prove that

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

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



Proof: (i) We have $\overline{A}$ ∩ B and A ∩ B are disjoint events i.e.

($\overline{A}$ ∩ B) ∩ B) = ϕ Also

($\overline{A}$ ∩ B) ∪ (A ∩ B) = B

∴ P$\left[(\overline{A}\cap B)\cup(A\cap B)\right]$ = P(B)

⇒ P($\overline{A}$ ∩ B) + P(A∩B) = P(B) [∵ of Axiom (3)]

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

Hence the result.

(ii) Similarly Starting with

(A ∩ $\overline{B}$) ∩ (A ∩ B) = ϕ

and (A ∩ $\overline{B}$) ∪ (A ∩ B) = A

We can show that

Theorem 6: Addition Theorem of Probability

Statement: If A and B are any two events ⊂ subsets of sample space ≤ ⊃, then from the class ⊂ containing the events A and B cohich are not disjoint, we have

P(A ∪ B) = P(A) + P(B) - P(A ∩ B)

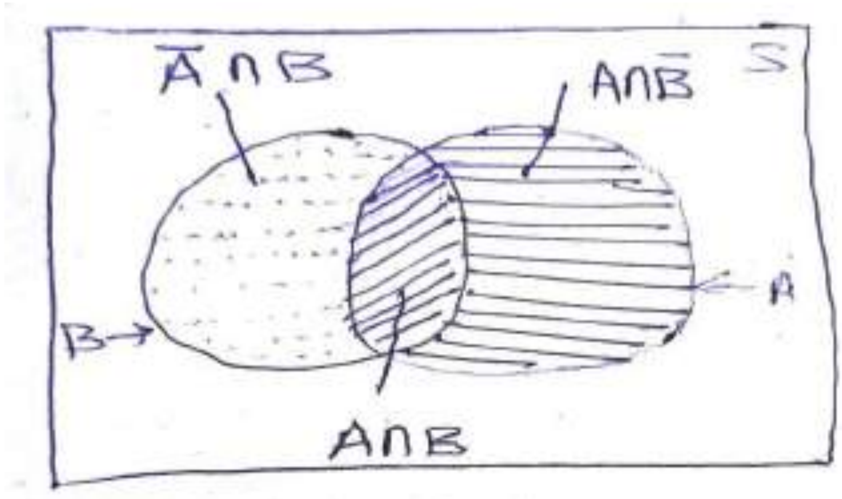
Proof: From the Venn diagram, we have

A∪B = A∪($\overline{A}$ ∩ B),cohere A and $\overline{A}$ ∩ B are mutually disjoint.

∴ P(A ∪ B) = P$\left[A\bigcup(\overline{A}\bigcap B)\right]$

= P(A) + P($\overline{A}$ ∩ B) [By Axiom 3] ......(*)

= P(A) + P(B) - P(A ∩ B) $\left[P(\overline{A}\bigcap B)=P(B)-p(A\bigcap B)\right]$



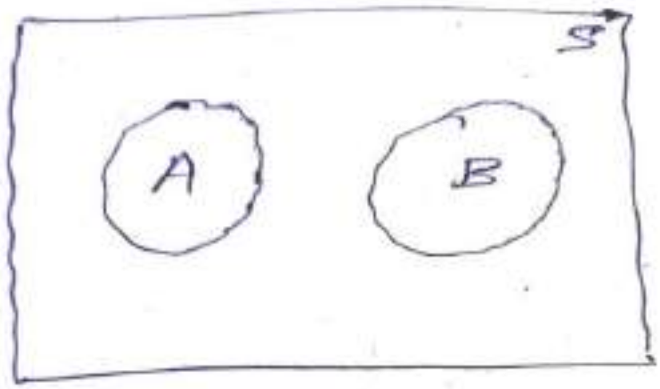
OR From (*) onwards

P(A ∪ B) = P(A) + $\left[P(\overline{A}\bigcap B)=P(A\bigcap B)+P(A\bigcap B)\right]$ - P(A ∩ B)

= P(A) +P$\left[(\overline{A}\bigcap B)\bigcup(A\bigcap B)-P(A\bigcap B)\right]$

[∵ ($\overline{A}$ ∩ B) and (A ∩ B) are disjoint]

⇒ P(A ∪ B) = P(A) + P(B) - P(A ∩ B)



Cor. 1: If the events A and B are mutually exclusive, then

$A \cap B = \phi \Rightarrow P(A \cap B) = P(\phi) = 0$

so that above result becomes

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

which is the third axiom of probability.

Cor 2: For three not mutually exclusive events A, B and C, we have

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

Proof : We have

$P(A \cup B \cup C) = P[A \cup (B \cup C)]$

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

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

$= P[(A \cap B) \cup (A \cap C)]$

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

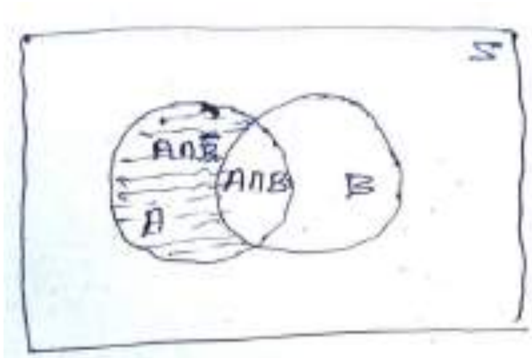
$- [P(A \cap B) + P(A \cap C) - P\{(A \cap B) \cap (B \cap C)\}$

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

$- P(C \cap A) + P[A \cap B \cap C]$

Hence the result.

Theorem 7: For any events A and B, prove that $P(A) = P(A \cap B) + P(A \cap \overline{B})$



Proof : We clearly have $A \cap B$ and $A \cap \overline{B}$ are disjoint sets.

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

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

$$\therefore \; P(A) = P[(A \cap B) \cup (A \cap \overline{B})]$$

By Axiom (2) of probability, we have

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

Hence the result.

Theorem 8 : Generalised Addition Theorem

For any n events $A_1, A_2, \ldots, A_n$ in a sample space, we have

$$P\left[\bigcup_{i=1}^{n} A_i\right] = \sum_{i=1}^{n} P(A_i) - \sum_{1 \le i < j \le n} \sum P(A_i \cap A_j)$$

$$+ \sum_{1 \le i < j < k \le n} \sum \sum P(A_i \cap A_j \cap A_k) + \ldots +$$

$$+ (-1)^{n-1} P(A_1 \cap A_2 \cap \ldots \cap A_n)$$

Proof : We shall prove the result by induction on n.

For n = 2, we have

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2) \qquad ...(1)$$

Which is always true

$\therefore$ the result holds for n = 2.

Let the result hold for n = r

i.e. $$P\left(\bigcup_{i=1}^{r} A_i\right) = \sum_{i=1}^{r} P(A_i) - \sum_{1\le i<j\le r}\sum P(A_i \cap A_j)$$

$$+ \sum_{1\le i<j<k\le r}\sum P(A_j \cap A_j \cap A_k) + ... +$$

$$+ .... (-1)^{r-1} P(A_1 \cap A_2 \cap ..... \cap A_r) \qquad ...(2)$$

is true

We shall now show that the result is also true for n = r + 1.

Consider

$$P\left(\bigcup_{i=1}^{r+1} A_i\right) = P\left(\left(\bigcup_{i=1}^{r} A_i\right) \cup A_{r+1}\right)$$

Using equation (1), we get

$$P\left(\bigcup_{i=1}^{r+1} A_i\right) = P\left(\bigcup_{i=1}^{r} A_i\right) + P(A_{r+1}) - P\left(\left(\bigcup_{i=1}^{r} A_i\right) \cap A_{r+1}\right)$$

$$= P\left(\bigcup_{i=1}^{r} A_i\right) + P(A_{r+1}) - P\left(\bigcup_{i=1}^{r} (A_i \cap A_{r+1})\right)$$

Using equation (2), we get

$$P\left(\bigcup_{i=1}^{r+1} A_i\right) = \sum_{i=1}^{r} P(A_i) - \sum_{1\le i<j\le r}\sum P(A_i \cap A_j)$$

$$+ \sum_{1\le i<j<k\le r}\sum\sum P(A_j \cap A_j \cap A_k) + ... +$$

$+ (-1)^{r-1} P(A_1 \cap A_2 \cap ..... \cap A_r) + P(A_{r+1}) - P\left(\bigcup_{i=1}^{r}(A_i \cap A_{r+1})\right)$

$$= \sum_{i=1}^{r+1} P(A_i) - \sum\sum_{1 \le i < j \le r} P(A_i \cap A_j) + \sum\sum\sum_{1 \le i < j < k \le r} P(A_j \cap A_j \cap A_k)$$

$$+ ...... + (-1)^{r-1} P(A_1 \cap A_2 \cap ..... \cap A_r)$$

$$- \left[\sum_{i=1}^{r} P(A_i \cap A_{r+1}) - \sum\sum_{1 \le i < j < k \le r} P(A_i \cap A_{r+1} \cap A_j)\right.$$

$$+ ...... + (-1)^{r-1} P(A_1 \cap A_2 \cap ..... \cap A_{r+1})]$$

[$\because$ of (2)]

$$= \sum_{i=1}^{r+1} P(A_i) - \left[\sum\sum_{1 \le i < j \le r} P(A_i \cap A_j) + \sum_{i=1}^{r} P(A_i \cap A_{r+1}) + \right.$$

$$\sum\sum\sum_{1 \le i < j < k \le r} P(A_i \cap A_j \cap A_k) + \sum\sum_{1 \le i < j \le r} P(A_i \cap A_j \cap A_{r+1})$$

$$+ (-1)^{r-1} P(A_1 \cap A_2 \cap ..... \cap A_{r+1}) + (-1)^r P(A_1 \cap A_2 \cap ..... \cap A_{r+1})]$$

$$= \sum_{i=1}^{r+1} P(A_i) - \sum\sum_{1 \le i < j \le r+1} P(A_i \cap A_j)$$

$$+ \sum\sum\sum_{1 \le i < j < k \le r+1} P(A_j \cap A_j \cap A_k) + ... + (-1)^r P(A_1 \cap A_2 \cap ..... \cap A_{r+1})$$

$\therefore$ the result is true for n = r + 1.

$\therefore$ by principle of Mathematical induction, the result is true $\forall$ n $\in$ N

Theorem 9 : Boole's Inequality : For any n events $A_1, A_2, ..., A_n$, we have

(i) $P\left(\bigcap_{i=1}^{n} A_i\right) \ge \sum_{i=1}^{n} P(A_i) - (n - 1)$ (ii) $\left(\bigcap_{i=1}^{n} A_i\right) \le \sum_{i=1}^{n} P(A_i)$

Proof : (i) For any two events A_1, A_2

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)$$

$\Rightarrow \quad P(A_1 \cap A_2) = P(A_1) + P(A_2) - P(A_1 \cup A_2)$ ....(1)

By axiom (1) of probability

$$P(A) \leq 1 \Rightarrow - P(A) \geq - 1$$

and in particular $- P(A_1 \cup A_2) \geq -1$

$\Rightarrow \quad - P(A_1) + P(A_2) - P(A_1 \cap A_2) \geq -1$

$\Rightarrow \quad P(A_1 \cap A_2) \geq P(A_1) + P(A_2) - 1$ ....(2)

$\Rightarrow$ The result is true for n = 2

Let us suppose that the result is true for n = r

i.e. $P\left(\bigcap_{i=1}^{r} A_i\right) \geq \sum_{i=1}^{r} P(A_i) - (r-1)$ ...(3)

Now consider n = r + 1

$$P\left(\bigcap_{i=1}^{r+1} A_i\right) = P\left(\left(\bigcap_{i=1}^{r} A_i\right) + P(A_{r+1}) - 1\right) + P(A_{r+1}) - 1 \qquad [\because \text{ of (2)}]$$

$$\geq \sum_{i=1}^{r} P(A_i) - (r-1) + P(A_{r+1}) - 1 \qquad [\because \text{ of (3)}]$$

$$= \sum_{i=1}^{r+1} P(A_i) - (r+1-1)$$

i.e. $P\left(\bigcap_{i=1}^{r+1} A_i\right) > \sum_{i=1}^{r+1} P(A_i) - r$

$\Rightarrow$ the result is true for n = r + 1

Thus the mathematical induction the result is true $\forall$ n $\in$ N.

(ii) We shall proceed by induction on n.

We already have

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)$$

$\Rightarrow \quad P(A_1 \cap A_2) = P(A_1) + P(A_2) - P(A_1 \cup A_2)$

$\because \quad P(A_1 \cap A_2) \geq 0$

$\therefore \quad P(A_1) + P(A_2) - P(A_1 \cup A_2) \geq 0$

$\Rightarrow \quad P(A_1 \cap A_2) \leq P(A_1) + P(A_2) \qquad \text{...(3)}$

$\therefore$ The result is true for n = 2.

Let us suppose that the result is true for n = r.

i.e. $P\left(\bigcap_{i=1}^{r} A_i\right) \geq \sum_{i=1}^{r} P(A_i) \qquad \text{...(4)}$

Now we shall show that the result is true for n = r+1.

For this we have

$$P\left(\bigcup_{i=1}^{r+1} A_i\right) = P\left[\left(\bigcup_{i=1}^{r} A_i\right) \cup A_{r+1}\right]$$

$$\leq P\left(\bigcup_{i=1}^{r} A_i\right) + P(A_{r+1}) \qquad [\because \text{ of (3)}]$$

$$\leq \sum_{i=1}^{r} P(A_i) + P(A_{r+1}) \qquad [\because \text{ of (4)}]$$

i.e. $P\left(\bigcup_{i=1}^{r+1} A_i\right) \leq \sum_{i=1}^{r+1} P(A_i)$

$\therefore$ by mathematical induction the result is true for all $n \in N$

i.e. $P\left(\bigcup_{i=1}^{n} A_i\right) \leq \sum_{i=1}^{n} P(A_i)$

Theorem 10 : For any n events $A_1, A_2, \ldots, A_n$, we have

$$P\left(\bigcup_{i=1}^{n} A_i\right) \geq \sum_{i=1}^{n} P(A_i) - \sum\sum_{1 \leq i < j \leq n} P(A_i \cap A_j)$$

Proof : We shall prove the result by induction.

We have that

$$P(A_1 \cup A_2 \cup A_3) = P(A_1) + P(A_2) + P(A_3)$$

$$- P(A_1 \cap A_2) + P(A_2 \cap A_3) + P(A_3 \cap A_1) + P(A_1 \cap A_2 \cap A_3)$$

$$\geq P(A_1) + P(A_2) + P(A_3) - [P(A_1 \cap A_2) + P(A_2 \cap A_3) + P(A_3 \cap A_1)$$

[As $P(A_1 \cap A_2 \cap A_3) \leq 1$]

i.e. $P\left(\bigcup_{i=1}^{3} A_i\right) \geq \sum_{i=1}^{3} P(A_i) - \sum\sum_{1 \leq i < j \leq 3} P(A_i \cap A_j)$

$\therefore$ The result holds for n = 3.

Let us suppose that the result is true for n = r.

i.e. P $P\left(\bigcup_{i=1}^{r} A_i\right) \geq \sum_{i=1}^{r} P(A_i) - \sum\sum_{1 \leq i < j \leq r} P(A_i \cap A_j)$ ...(1)

We shall prove that the result is also true for n = r + 1.

For this, consider

$$P\left(\bigcup_{i=1}^{r+1} A_i\right) = P\left(\left(\bigcup_{i=1}^{r} A_i\right) \cup A_{r+1}\right)$$

$$= P\left(\bigcup_{i=1}^{r} A_i\right) + P(A_{r+1}) - P\left(\left(\bigcup_{i=1}^{r} A_i\right) \cup A_{r+1}\right)$$

$$> \left[\sum_{i=1}^{r} P(A_i) - \sum\sum_{1 \leq i < j \leq r} P(A_i \cap A_j)\right] + P(A_{r+1}) - P\left(\left(\bigcup_{i=1}^{r} A_i\right) \cap A_{r+1}\right)$$

[$\because$ of (1)]

$$\sum_{i=1}^{r+1} P(A_i) - \sum\sum_{1\le i<j\le r} P(A_i \cap A_j) - P\left(\left(\bigcup_{i=1}^{r} A_i\right) \cap A_{r+1}\right)$$

By Boole's inequality, we have

$$P\left(\bigcup_{i=1}^{r} (A_i \cap A_{r+1})\right) \le \sum_{i=1}^{r} P(A_i \cap A_{r+1})$$

$$\Rightarrow \quad - P\left(\bigcup_{i=1}^{r} (A_i \cap A_{r+1})\right) \ge - \sum_{i=1}^{r} P(A_i \cap A_{r+1}) \qquad ...(3)$$

$\therefore$ from (2) and (3), we get

$$P\left(\bigcup_{i=1}^{r+1} A_i\right) - \sum_{i=1}^{r+1} P(A_i) + \sum\sum_{1\le i<j\le r} P(A_i \cap A_j) \ge - \sum_{i=1}^{r} P(A_i \cap A_{r+1})\Bigg]$$

$$P\left(\bigcup_{i=1}^{r+1} A_i\right) \ge \sum_{i=1}^{r+1} P(A_i) - \left[\sum\sum_{1\le i<j\le r} P(A_i \cap A_j) + \sum_{i=1}^{r} P(A_i \cap A_{r+1})\right]$$

$$= \sum_{i=1}^{r+1} P(A_i) - \sum\sum_{1\le i<j\le r+1} P(A_i \cap A_j)$$

Hence the result.

Let us now do some examples to have better idea of the concept : -

Example 1 : If A and B are two events defined on a sample space such that P(A∪B) = $\frac{1}{3}$, then find P(A).

Sol. Here P(A∪B) = $\frac{5}{6}$, P(A∩B) = $\frac{1}{3}$, P($\overline{B}$) = $\frac{1}{3}$

Now P($\overline{B}$) = $\frac{1}{3} \Rightarrow \frac{1}{3}$ 1 - P(B) = $\frac{1}{3} \Rightarrow$ P(B) = $\frac{2}{3}$

Also P(A∪B) = P(A) + P(B) 0 P (A∩B)

$\therefore \qquad \frac{5}{6} = P(A) + \frac{2}{3} + \frac{1}{3}$

$\Rightarrow \qquad P(A) = \frac{5}{6} - \frac{2}{3} + \frac{1}{3}$

$= \frac{3}{6} = \frac{1}{2}$

Example 2: A ball is drawn at random from a box containing 6 red balls, 4 white balls and 5 blue balls. Determine the probability that ball drawn is

(a) red (b) white

(c) blue (d) red or white

Sol: Let A be the event of drawing a red ball, B be the event of drawing a white ball and C be the event of drawing a blue ball.

Then $A \cup B$ is the event of drawing either red ball or a white ball.

(a) P (red ball) = $P(A) = \frac{6}{15} = \frac{2}{5}$

(b) P(white ball) = $P(B) = \frac{4}{15}$

(c) P (blue ball) = $P(C) = \frac{5}{15} = \frac{1}{3}$

(d) P (red ball or white ball)

$= P(A \cup B) = P(A) + P(B)$ [$\because$ A and B are mutually exclusive)

$= \frac{6}{15} + \frac{4}{15} = \frac{10}{15}$

$= \frac{2}{3}$

Example 3: The odds in favour of standing first of three students appearing at an examination are 1 : 2, 2 : 5 and 1 : 7 respectively. Find the probability that either of them stands first.

Sol. Let A, B, C denote the events of standing first of the three students respectively.

$\therefore \qquad P(A) = \frac{1}{1+2} = \frac{1}{3}, \; P(B) = \frac{2}{2+5} = \frac{2}{7},$

$P(C) = \frac{1}{1+7} = \frac{1}{8}$

$\therefore$ Required probability = P(A∪B∪C)

= P(A) + P(B) + P(C)

[$\because$ events A, B and C are mutually exclusive)

$= \frac{1}{3} + \frac{2}{7} + \frac{1}{8}$

$= \frac{125}{168}$

2.4 Self Check Exercise

Q.1 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 $P(A \cap \overline{B})$.

Q.2 One card in drawn from a pack of 52 cards, each of 52 cards being equally likely to be drawn. Find the probability of

(i) the card drawn is red

(ii) the card drawn is a king

(iii) the card drawn is a red and a king

(iv) the card drawn is either red on a king.

2.5 Summary

We conclude this unit by summarizing what we have covered in it:-

1. Defined and discussed three axioms of probability i.e. non-negativity axiom; normalization axiom and additivity axiom.
2. Defined and discussed the probability set function.
3. Proved some basic and important theorems on probability.
4. Did some equations on probability by using basic theorems and axioms of probability.

2.6 Glossary:

1. Suppose we have a sample space S. If S is discrete, all subsets correspond to events and conversely, but if S is non-discrete, only special subsets c(called measurable) correspond to events. To each event A in the class C of events, we

associate real number P(A). Then P is called a probability set function and P(A) the probability of the events if it satisfies three axioms.

2. If A and B are any two events, then from the class C containing the events A and B which are not disjoint, we have

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

2.7 Answer to Self Check Exercise

Ans.1 $P(A\cap B) = \frac{3}{20}$

and $P(A\cap \bar{B}) = \frac{1}{10}$

(i) $\frac{1}{2}$ (ii) $\frac{1}{3}$

(iii) $\frac{1}{26}$ (iv) $\frac{7}{13}$

2.8 References/Suggested Readings

1. Sheldon Ross, Introduction to Probability Model, 9th Ed., Academic Press, Indian Reprint, 2007

2. Irwin Miller, Marylees Miller and John E. Freund, Mathematical Statistics with Application, 7th Ed., Pearson Education, Asia, 2006.

3. Robert V. Hogg, Joseph w. Mckean and Allen T. Craig, Introduction to Mathematical Statistics, Pearson Education, Asia, 2007.

2.9 Terminal Questions

1. A and B are two events such that P(A) = 0.54, P(B) = 0.69 and $P(A\cap B)$ = 0.35, find (i) $P(A\cup B)$ (ii) $P(\bar{A}\cap \bar{B})$

   (iii) $P(A\cap \bar{B})$ (iv) $P(B\cap \bar{A})$

2. A card is drawn from 52 cards at random. Find the probability that card drawn is a heat or a face card or an ace.

3. M and N are two events. Show that the probability that one of them occur is P(M) + P(N) - 2 $P(M\cap N)$.

4. Two unbiased dice are thrown. Find the probability that neither a doublet nor a total of 10 will appear.

Unit - 3

Conditional Probability, Multiplication Theorem of Probability And Independence

Structure

3.1 Introduction

The probability P(A) of an event A represents the likelihood that a random experiment will result in an outcome in the set A relative so the sample space S of the random experiment. However, quite often, while evaluating some event probability, we already have same information stemming from the experiment. For example, if we have prior information that the outcome of the random experiment must be in a set B of S, then this information must be used to re-appraise the likelihood that the outcome will also be in B. This re-appraised probability is denoted by P(A/B) and is read as the conditional probability of the event A, given that the event B has already happened. For example, let us consider a random experiment of drawing a card from a pack of cards. Then the probability of happening of the event A : "The card drawn is a king", is given by : P(A) = $\frac{4}{52} = \frac{1}{13}$

Now suppose that a card is drawn and 49 we are informed that the drawn card is red. How does this information effect the likelihood of the event A? Obviously, if the event B, the card drawn is red, has happened, the event 'Black card' is not possible. Hence the probability of the event A must be computed relative to the new sample space 'B' which consists of 26 sample points (red cards only). Among these 26 red cards, there are two red kings. Hence the required probability P(A|B) = $\frac{2}{26} = \frac{1}{13}$.

From this illustrations we observe that some additional information may change the probability of the happening of some event.

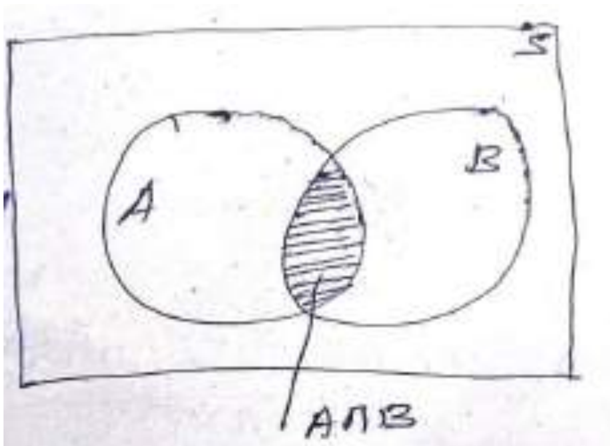
3.2 Learning Objectives

After studying this unit, you should be able to:

- Defined and discuss conditional probability and do questions related to it.
- State and prove multiplication theorem of probability.
- Define and discuss independent events.
- State and prove multiplication theorem of probability for independent events.
- Define and discuss pairwise and mutually independent events.
- Prove the theorem on law of total probability.
- Prove Baye's theorem.
- Do questions related to these theorem's.

3.3 Conditional Probability

Let A and B two events and S be the sample space. We denote by P(B|A) the probability of B given that A has occurred called the conditional probability of B given A. Since A is known to have occurred, it becomes the new sample space replacing the original S. we thus have conditional probability of B given A



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

We can also define P(B|A) as a probability set function, defined for subset of A as follows:

1. $P(B|A) \geq 0$

2. $P(B_1 \cup B_2 \cup ....|A) = P(B_1|A) + P(B_2|A) + ...$, provided that $B_1, B_2, ...$are mutually disjoint sets.

3. $P(A|A) = 1$

The statements 1 and 3 are obvious and for the 2 statement, we note that, by def.

$$P(B_1 \cup B_2 \cup ....|A) = \frac{P\left[A \cap (B_1 \cup B_2 ......)\right]}{P(A)}$$

$$= \frac{P\left[(A \cap B_1) \cup (A \cap B_2) \cup ......)\right]}{P(A)}$$

Here $(A \cap B_i) \cap (A \cap B_j) = A \cap (B_i \cap B_j) = A \cap \phi = \phi$

$\therefore$ the above relation becomes.

$$P(B_1 \cup B_2 \cup ....|A) = \frac{P(A \cap B_1) + P(A \cap B_2) + ...}{P(A)}$$

$$= \frac{P(A \cap B_1)}{P(A)} + \frac{P(A \cap B_2)}{P(A)} + ...$$

$\Rightarrow$ $P(B_1 \cup B_2 \cup ....|A) = P(B_1|A) + P(B_2|A) + ...$

Thus, statements is also proved and P(B|A) is a probability set function. This may thus be called the conditional probability set function, relative to the hypothesis A.

Similarly, P(A|B) = conditional probability of A and B = $\frac{P(A \cap B)}{P(B)}$

Let us do some examples to have better idea of the concept:-

Example 1: If A and B are two events such that P(A) = 0.5, P(B) = 0.6 and $P(A \cup B) = 0.8$, find P(A|B) and P(B|A)

Sol. Here P(A) = 0.5, P(B) = 0.6, $P(A \cup B) = 0.8$

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

$\Rightarrow$ $0.8 = 0.5 + 0.6 - P(A \cap B)$

$\Rightarrow$ $P(A \cap B) = 0.5 + 0.6 - 0.8$

$= 0 - 3$

$\therefore$ $P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{0.3}{0.6} = \frac{1}{2} = 0.5$

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{0.3}{0.5} = \frac{3}{5} = 0.6$$

Example 2: A die is thrown twice and the sum of the numbers appearing is observed to be 6. What is the conditional probability that the number 4 has appeared at least once?

Sol: Let the events be

E : number 4 appears at past once

F : the sum of the numbers appearing is 6

The elementary event favorable the occurrence of E are

(4,1), (4,2), (4,3), (4,4), (4,5), (4,6), (1,4), (2,4), (3,4), (5,4), (6,4)

The elementary events favourable to the occurrence of F are

(1,5), (2,4), (3,3), (4,2), (5,1)

The elementary events favourable to the occurrence of both E and F are

(2,4), (4,2)

$\therefore$ $P(E) = \frac{11}{36}$, $P(F) = \frac{5}{36}$, $P(E \cap F) = \frac{2}{36}$

$\therefore$ Required probability $= \frac{P(E \cap F)}{P(F)} = \frac{2}{5}$

Example 3: In a hostel, 60% of the students read Hindi newspaper, 40% read English newspaper and 20% read both Hindi and English newspaper. A students at random.

(a) find the probability that he reads neither Hindi nor English newspaper.

(b) If he reads Hindi newspaper, find the probability that he reads English newspaper.

(c) If he reads English newspaper, find the probability that he reads Hindi newspaper.

Sol. Let A and B be two events such that

A : student reads Hindi newspaper

B : a student reads English newspaper

$\therefore P(A) = \frac{60}{100} = \frac{3}{5}$; $P(B) = \frac{40}{100} = \frac{2}{5}$

$P(A \cap B) = \frac{20}{100} = \frac{1}{5}$

(a) P (a student reads neither Hindi nor English newspaper) = $P(A^c \text{ and } B^c)$

$= P\left[(A \cup B)^c\right] = 1 - P(A \cup B)$

$= 1 - \left[P(A)+P(B)-P(A\cap B)\right]$

$= 1 - \left[\frac{3}{5}+\frac{2}{5}-\frac{1}{5}\right]$

$= 1 - \frac{4}{5}$

$= \frac{1}{5}$

(b) P[he reads, English newspaper when it is given that he reads Hindi newspaper]

$= P(B|A) = \frac{P(B\cap A)}{P(A)} = \frac{P(A\cap B)}{P(A)} = \frac{1/5}{3/5}$

$= \frac{1}{3}$

(c) P(he reads Hindi newspaper when it is given that he reads English newspaper)

$= P(A|B) = \frac{P(A\cap B)}{P(B)} = \frac{1/5}{2/5} = \frac{1}{2}$

Example 4: If P(A) = p, P(B) = 2, then show that P(A|B) $\geq \frac{p+q-1}{q}$

Sol. Let S be the sample space we note that

P(A∪B) = P(A) +(B) - P(A∩B)

$\Rightarrow$ P(A∪B) = p + q - P(A∪B)

$\Rightarrow$ p + q - P(A∪B) = P(A∪B) ......(1)

Now P(A∪B) $\leq$ 1

$\Rightarrow$ - P(A∪B) $\geq$ -1

Now p + q - p(A∪B) $\geq$ p + q - 1

$\Rightarrow$ P(A∩B) $\geq$ p + q - 1 [$\because$ of (1)]

We know that

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

$\therefore$ P(A|B) $\geq \frac{P+q-1}{q}$ [$\because$ of (2)]

Hence the result

Self Check Exercise-1

Q.1 If A and B are two events such that P(A) = 0.3, P(B) = 0.5 and P(A|B) = 0.4, find P(A∩B) and P(B|A)

Q.2 A die is thrown three times. Events A and B are defined as below:-

A : 4 on the third throw

B : 6 on the first and 5 on the second throw.

Find the probability of A given that B has already occwared.

Q.3 Find the probability of drawing a king, when a card is drawn from a well shuffled pack of cards. It is also given that the card drawn is a face card.

3.4 Multiplication Theorem of Probability

For two events A and B in a sample space S,

$$\left.\begin{array}{l} P(A\cap B)=P(A).P(B\,|\,A), P(A)>0 \\ \text{and } P(A\cap B)=P(B).P(A\,|\,B), P(B)>0 \end{array}\right\} \quad .....(1)$$

Where P(A) = probability of occurrence of A

P(B|A) = conditional probability of occurrence of B given A

P(A|B) = conditional probability of occurrence of A given B.

The statement (1) is called the Multiplication theorem of probability. This can also be stated as "the probability of simultaneous occurrence of two events A and B is equal to the product of the probability of the other, given that the first one has occurred."

Proof: With usual notations, we have

P(A) = $\frac{n(A)}{n(S)}$, P(B) = $\frac{n(B)}{n(S)}$, P(A∩B) = $\frac{n(A\cap B)}{n(S)}$

Also P(B|A) = $\frac{\text{number of elements common to A and B}}{\text{number of elements in A}}$

$= \frac{n(A\cap B)}{n(A)}$

$\therefore$ R.H.S. of (1) is

P(A).P(B|A) = $\frac{n(A)}{n(S)} \cdot \frac{n(A\cap B)}{n(A)} = \frac{n(A\cap B)}{n(S)}$

$= P(A \cap B)$

Similarly $P(B).P(A|B) = P(A \cap B)$

Hence the result.

For n events $A_1, A_2, ..., A_n$, the multiplication theorem of probability can be stated as:

For n events $A_1, A_2, ..., A_n$, we have $P(A_1 \cap A_2 \cap ...... \cap A_n) = P(A_1)\ P(A_2|A_1)\ P(A_3|A_1 \cap A_2) .... P(A_n|A_1 \cap A_2 \cap .... \cap A_{n-1})$ ....(2)

Where $P(A_i|A_j \cap A_k \cap .... \cap A_l)$ represents the conditional probability of the event Ai given that the events A_j, A_k,......,A_l have already occurred.

3.5 Independent Events

Definition: Two or more event are said to be independent if the happening or non-happening of any one of them, does not in any way, affect the happening of others. In terms of probability, we can say that, an event A is said to be independent cor statistically independent) of another event B, if the conditional probability of A given B, i.e. P(A|B) is equal to the unconditional probability of A i.e. if

$$P(A|B) = P(A)$$

and similarly, B is said to be independent of A, if

$$P(B|A) = P(B)$$

Multiplication Theorem of Probability for Independent Events

If A and B are two events with positive probabilities, then A and B are independent iff

$$P(A \cap B) = P(A)\ P(B)$$

Proof: We have, by def.

$$P(A \cap B) = P(A)\ P(B|A) = P(B)\ P(A|B) \quad ......(1)$$

where $P(A) > 0$, $P(B) > 0$

If A and B are independent, then

$$P(A|B) = P(A) \text{ and } P(B|A) = P(B) \quad ....(2)$$

$\therefore$ From (1) and (2), we get

$$P\ (A \cap B) = P(A)\ P(B)$$

Hence the result

Conversely, if

$$P(A \cap B) = P(B)\ P(B)$$

then $\dfrac{P(A \cap B)}{P(B)} = P(A) \Rightarrow P(A|B) = P(B)$ ....(3)

and $\dfrac{P(A \cap B)}{P(A)} = P(B) \Rightarrow P(B|A) = P(B)$ ....(4)

(3) and (4) $\Rightarrow$ A and B are independent events.

Hence proved.

For n events, it is stated as: n events $A_1, A_2, ..., A_n$ are independent if and only if

$P(A_1 \cap A_2 \cap .... \cap A_n) = P(A_1)\, P(A_2) .... P(A_n)$

3.6 Pairwise And Mutually Independent Events

The n events $A_1, A_2, ..., A_n$ defined on a sample space S with $P(A_i) > 0$; i = 1,2,.....,n are said to be pairwise independent if every pair of two events is independent i.e.

$P(A_i \cap A_j) = P(A_i)\, P(A_j)$; $i \neq j = 1,2,......,n$.

The n events $A_1, A_2, ..., A_n$ defined on a sample space S with $P(A_i) = 0$, i = 1,2,....,n are said to be mutually independent if the probability of the simultaneous occurrence of (any) finite number of these is equal to the product of their separate probabilities i.e.

$P(A_{i_1} \cap A_{i_2} \cap .... \cap A_{i_k}) = P(A_{i_1})\, P(A_{i_2}) .... P(A_{i_k})$; k = 2,3,.......,n

Remarks:- 1. The total number of pairwise conditions for mutual indolence of A1,A2,.....,An is $2^n - 1 - n$.

2. Mutual independence of events implies that they are pairwise independent, but not conversely.

Let us do some examples:-

Example 5:- Assuming the probability of a male birth as $\frac{1}{2}$, find the chances that a family of 3 children will have

(i) atleast one girl, (ii) two boys and one girl and

(iii) atmost two girls.

Sol: Here it is given that probability of male birth = $\frac{1}{2}$

$\therefore$ probability of female birth = $\frac{1}{2}$

(i) P (at least one girl) = 1 - P (no girl birth)

= 1 - p (3 male births)

= 1 - P(BBB)

$= 1 - \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2}$

$= 1 - \frac{1}{8}$

$= \frac{7}{8}$

(ii) P (two boys and one girl) = P(BBG) + P(BGB) + P(GBB)

$= \frac{1}{8} + \frac{1}{8} + \frac{1}{8}$

$= \frac{3}{8}$

(iii) P (at most two girls) = 1 - P(all three girls)

= 1 - P(GGG)

$= 1 - \frac{1}{8}$

$= \frac{7}{8}$

Example 6: 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 E, F denote the two events of passing the test by two students.

$\therefore$ $P(E) = \frac{3}{3+7} = \frac{3}{10}$, $P(F) = \frac{5}{3+5} = \frac{5}{8}$

$\therefore$ Required probability = P(E and F)

= P(E) P(F)

$= \frac{3}{10} \times \frac{5}{8} = \frac{3}{16}$

Example 7: A problem in statistics is given to three students A, B and C whose chances of solving it are $\frac{1}{2}$, $\frac{3}{4}$ and $\frac{1}{4}$ respectively. What is the probability that the problem will be solved if all of them try independently.

Sol: Let A, B, C denote the events that the problem is solved by the students A, B, C respectively, then

$P(A) = \frac{1}{2}$, $P(B) = \frac{3}{4}$ and $P(C) = \frac{1}{4}$

The problem will be solved if at least one of them solves the problem. Thus the required probability is

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

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

Since A, B and C are independent events

$\therefore$ $\overline{A}, \overline{B}, \overline{C}$ are also independent

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

$$= 1 - \left(1-\frac{1}{2}\right)\left(1-\frac{3}{4}\right)\left(1-\frac{1}{4}\right)$$

$$= 1 - \frac{3}{32}$$

$$= \frac{29}{32},$$

Which is the required probability.

Example 8: A and B throw a coin alternatively till one of them gets a head and wins the game. If A starts the game, find their respective probabilities of winning.

Sol: Let P(A), $P(\overline{A})$ be probabilities of A's getting the head and not getting the head respectively, then

$$P(A) = \frac{1}{2} \Rightarrow P(\overline{A}) = 1-P(A) = 1 - \frac{1}{2} = 1$$

Similarly $P(B) = \frac{1}{2}$ and $P(\overline{B}) = \frac{1}{2}$

Let A start the game. He can win in the first throw, 3rd throw, 5th throw and so on.

Probability of A's winning in first throw = $P(A) = \frac{1}{2}$

Probability of A's winning in 3rd throw = $P(\overline{A})\, P(\overline{B})\, P(A)$

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

Probability of A's winning in 5th throw

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

$$= \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \left(\frac{1}{2}\right)^5$$

....

....

Since all these cases are mutually exclusive

$\therefore$ Probability of A's winning the game first is

$$= \frac{1}{2} + \left(\frac{1}{2}\right)^3 + \left(\frac{1}{2}\right)^5 + \ldots\ldots\ldots\ldots\infty$$

$$= \frac{1/2}{1-\left(\frac{1}{2}\right)^2} \qquad [\because S = \frac{a}{1-r}]$$

$$= \frac{1/2}{3/4} = 2/3$$

Since either A or B wins

$\therefore$ Probability of B's winning the game first

$$= 1 - 2/3 = 1/3$$

Self-Check Exercise - 2

Q.1 A fair die is tossed twice. Find the probability of getting 4,5 or 6 on the first toss and 1,2,3, or 4 on the second toss.

Q.2 A town has two doctors A and B operating independently. If the probability that doctor A is available is 0.9 and that for doctor B is 0.8, what is the probability that at least one doctor is available when needed?

3.7 Law of Total Probability

Statement: If an event A must result in one of the mutually exclusive events $E_1, E_2, \ldots, E_n$, then

$$P(A) = P(E_1)\, P(A|E_1) + P(E_2)\, P(A|E_2) + \ldots\ldots\ldots\ldots + P(E_n)\, P(A|E_n)$$

Proof: If an event A must result in one of the mutually exclusive events $E_1, E_2, \ldots\ldots, E_n$, then

$$A = A \cap (E_1 \cup E_2 \cup \ldots\ldots \cup E_n)$$

i.e. $A = (A \cap E_1) \cup (A \cap E_2) \cup \ldots\ldots \cup (A \cap E_n)$

Since E_i and E_j, $i \neq j$ are mutually exclusive

$\therefore$ $A \cap Ei$ and $A \cap Ej$, $i \neq j$ are also mutually exclusive

$\therefore$ By axiom 3 of probability, we have

$$P(A) = P[(A \cap E_1) \cup (A \cap E_2) \cup \ldots\ldots \cup (A \cap E_n)]$$

$$P(A \cap E_1) + P(A \cap E_2) + \ldots\ldots\ldots + P(A \cap E_n)$$

$\Rightarrow$ P(A) = $P(E_1)\,P(A|E_1) + P(E_2)\,P(A|E_2) + \ldots\ldots\ldots + P(E_n)\,P(A|E_n)$

or P(A) = $\sum_{i=1}^{n} P(E_i)P(A\,|\,E_i)$

Hence the result.

3.7 Baye's Theorem

Statement:- If $E_1, E_2, \ldots\ldots, E_n$ are mutually disjoint events with $P(E_i) \neq 0$, (i = 1, 2,....,n), then for any arbitrary event A which is a subset of $\sum_{i=1}^{n}(E_i)$ such that P(A) > 0, we have

$$P(E_i|A) = \frac{P(E_i)P(A\,|\,E_i)}{P(A)} = \frac{P(E_i)P(A\,|\,E_i)}{\sum_{i=1}^{n} P(E_i)P(A\,|\,E_i)}$$

Where i = 1, 2,....., n.

Proof: We are given that

$$A \subset \bigcup_{i=1}^{n} E_i$$

$\therefore$ We have using distributive law

$$A = A \cap \left(\bigcup_{i=1}^{n} E_i\right) = \bigcup_{i=1}^{n}(A \cap E_i)$$

Since $(A \cap E_i) \subset E_i$, (i = 1, 2,......, n)

and E_i's being mutually exclusive

$\therefore$ $A \cap E_i$'s are also mutually exclusive.

$\therefore$ Using axiom 3 of the definition of probability

$$P(A) = P\bigcup_{i=1}^{n}(A \cap E_i)$$

$$= P\left[(A \cap E_1) + P(A \cap E_2) \cup \ldots\ldots \cup (A \cap E_n)\right]$$

$\Rightarrow$ P(A) = $P(A \cap E_1) + P(A \cap E_2) + \ldots\ldots + P(A \cap E_n)$

and using multiplication theorem of probability, we get

$$P(A) = P(E_1)P(A|E_1) + P(E_2)P(A|E_2) + \ldots\ldots\ldots + P(E_n).P(A|E_n)$$

$$\Rightarrow \quad P(A) = \sum_{i=1}^{n} P(E_i)P(A\,|\,E_i) \qquad \ldots..(1)$$

Also, we have

$P(A \cap E_i) = P(A)\, P(E_i|A) = P(E_i)\, P(A|E_i)$ (By def.)

$$\Rightarrow \quad P(E_i|A) = \frac{P(A \cap E_i)}{P(A)} \qquad ....(2)$$

Now, on using equation (1) in (2), we get

$$P(E_i|A) = \frac{P(A \cap E_i)}{\sum_{i=1}^{n} P(E_i)P(A|E_i)} \qquad ......(3)$$

$$\Rightarrow \quad P(E_i|A) = \frac{P(E_i)P(A|E_i)}{\sum_{i=1}^{n} P(E_i)P(A|E_i)}$$

Hence the result.

Note: The probability of the occurrence of another event c,

$P(C|A \cap E_1)$, $P(C|A \cap E_2)$,..........$P(C|A \cap E_n)$ is given by

$$P(C|A) = \frac{\sum_{i=1}^{n} P(E_i)P(A|E_i)P(C|E_i \cap A)}{\sum_{i=1}^{n} P(E_i)P(A|E_i)}$$

Remarks:-

1. The probabilities $P(E_1)$, $P(E_2)$,........$P(E_n)$ are term as the 'a prior probabilities' because they are known before the happening of the experiment itself.
2. The probabilities $P(A|E_i)$, i = 1,2,......,n are called 'Likelihoods' because they indicate how likely the event A under consideration is to occur, given each and every a priori probability.
3. The probabilities $P(E_i|A)$, i = 1, 2,....., n are called 'Posterior probabilities' because they are determined offer the result of the experiment are known.

Let us now do some examples to have better idea of the concept:-

Example 9:- An win contains 10 white and 3 black balls, while another urn contains 3 white and 5 black balls. Two balls are drawn from the first urn and put into the second urn and them a ball is drawn from the latter. What is the probability that it is white ball?

Sol:

	White	Black
urn I	10	3

urn II	3	5

Let E_1, E_2, E_3 denote the events that two balls drawn from urn I are both white, one white and one black, both black.

$$\therefore \quad P(E_1) = \frac{10_{c_2}}{13_{c_2}} = \frac{\frac{10\times 9}{1\times 2}}{\frac{13\times 12}{1\times 2}} = \frac{15}{26}$$

$$P(E_2) = \frac{10_{c_1}\times 3_{c_1}}{13_{c_2}} = \frac{\frac{\frac{10}{1}\times\frac{3}{1}}{1\times 2}}{\frac{13\times 12}{1\times 2}} = \frac{5}{13}$$

$$P(E_3) = \frac{3_{c_2}}{13_{c_2}} = \frac{3\times 12}{13\times 12} = \frac{1}{26}$$

Alter 2 balls have been drawn from urn I, urn II. will have

(i) 5 white, 5 black balls or (ii) 4 white, 6 black balls or (iii) 3 white, 7 black balls.

Let E denote the event that a white ball is drawn from urn II.

$$\therefore \quad P(E|E_1) = \frac{5}{10} = \frac{1}{2},\ P(E|E_2) = \frac{4}{10} = \frac{2}{5},\ P(E|E_3) = \frac{3}{10}$$

Now $P(E) = P(E_1)\,P(E|E_1) + P(E_2)\,P(E|E_2) + P(E_3).\ P(E|E_3)$

$$= \frac{15}{26}\times\frac{1}{2} + \frac{5}{13}\times\frac{2}{5} + \frac{1}{26}\times\frac{3}{10}$$

$$= \frac{75+40+3}{260} = \frac{118}{260} = \frac{59}{130}$$

Example 10: Suppose that 5% of men and 0.25% of women have grey hair. A grey haired person is selected at random. What is the probability of this person being male? Assume that there are equal number of males and females.

Sol: Let E_1, E_2, E be the events as

E_1 : 'Selected person is a male'

E_2 : 'Selected person is a female'

E_3 : 'Selected person is grey haired'

$\therefore$ P(E_1) = P(E_2) = $\frac{1}{2}$

and P(E|E_1) = $\frac{5}{100} = \frac{1}{20}$; P(E|E_2) = $\frac{0.25}{100} = \frac{1}{400}$

Required probability = P(E_1|E) = $\frac{P(E_i)P(A|E_i)}{P(E_i)P(E|E_i)+P(E_2)P(E|E_2)}$

$$= \frac{\frac{1}{2}\times\frac{1}{20}}{\frac{1}{2}\times\frac{1}{20}+\frac{1}{2}\times\frac{1}{400}} = \frac{1/20}{1/20+\frac{1}{400}} = \frac{1/20}{\frac{20+1}{400}}$$

$$= \frac{1/20}{\frac{21}{400}} = 1/20 \times \frac{400}{21} = \frac{20}{21}$$

Example 11: The contents of urns I, II and III are as follow:

1 white, 2 black and 3 ked balls,

2 white, 1 black and 1 red ball, and

3 white, 5 black and 3 red balls.

One urn is chosen at random and two balls drawn from it. They happen to be white and red. What is the probability that they come from urns I, II or III.

Sol: Let E_1, E_2 and E_3 denote the events that one urns I, II and III are chosen, respectively Let A be the event that the two balls taken from the selected urn are white and red.

Then P(E_1) = P(E_2) = P(E_3) = $\frac{1}{3}$

P(A|E_1) = $\frac{1\times 3}{6_{c_2}} = \frac{1}{5}$, P(A|$E_2$) = $\frac{2\times 1}{4_{c_2}} = \frac{1}{3}$

and P(A|E_3) = $\frac{4\times 3}{12_{c_2}} = \frac{2}{11}$

Also P(A) = $\sum_{i=1}^{3} P(E_i)P(A|E_i)$

$$= \frac{1}{3}\times\frac{1}{5}+\frac{1}{3}\times\frac{1}{3}+\frac{1}{3}\times\frac{2}{11}$$

$$= \frac{1}{15} + \frac{1}{9} + \frac{2}{33} = \frac{33+55+30}{495} = \frac{188}{495}$$

$\therefore$ Required probabilities are

$$P(E_1|A) = \frac{P(E_1)P(A|E_1)}{P(A)} = \frac{\frac{1}{3}\times\frac{1}{5}}{118/495} = \frac{33}{118}$$

$$P(E_2|A) = \frac{P(E_2)P(A|E_2)}{P(A)} = \frac{\frac{1}{3}\times\frac{1}{3}}{118/495} = \frac{55}{118}$$

$$P(E_3|A) = \frac{P(E_3)P(A|E_3)}{P(A)} = \frac{\frac{1}{3}\times\frac{2}{11}}{118/495} = \frac{30}{118}$$

Example 12: A company has two plants to manufacture scooters. Plant I manufactures 70% of scooters and plant II manufactures 30%. At plant I, 80% of the scooters are rated as of standard quality and at plant II, 90% of the scooters are rated as of standard quality. A scooter is chosen at random and is found to be of standard quality. What is the probability that it has come from.

(i) Plant I (ii) Plant II ?

Sol: Let E_1, E_2 be the events that the scooter is produced by plants I, II respectively.

$\therefore$ $P(E_1) = \frac{70}{100}$, $P(E_2) = \frac{30}{100}$

Let E be the event of scoter being of standard quality.

$\therefore$ $P(E|E_1) = \frac{80}{100}$, $P(E|E_2) = \frac{90}{100}$

(i) $$P(E_1|E) = \frac{P(E_1)P(E|E_1)}{P(E_1)P(E|E_1)+P(E_2)P(E|E_2)}$$

$$= \frac{\frac{70}{100}\times\frac{80}{100}}{\frac{70}{100}\times\frac{80}{100}+\frac{30}{100}\times\frac{90}{100}} = \frac{5600}{5600+2700}$$

$$= \frac{5600}{8300} = \frac{56}{83}$$

(ii) $P(E_2|E) = \dfrac{P(E_2)P(E|E_2)}{P(E_1)P(E|E_1)+P(E_2)P(E|E_2)}$

$$= \frac{\frac{30}{100}\times\frac{90}{100}}{\frac{70}{100}\times\frac{80}{100}+\frac{30}{100}\times\frac{90}{100}} = \frac{2700}{5600+2700}$$

$$= \frac{2700}{8300} = \frac{27}{83}$$

Self-Check Exercise-3

Q.1 A bag A contains 3 white and 4 black balls. Another bag B contains 5 white and 7 black balls. A ball is transferred from the bag A to the bag B and one ball is drawn from the second bag B. Find the probability that it will be white.

Q.2 Suppose a girl throws a die. If she gets a 5 or 6, she tosses a coin three times and notes the number of heads. If she gets 1,2,3 or 4, she tosses a coin once and notes whether a head or tail is obtained. If she obtained exactly one head, what is the probability that she threw 1,2,3 or 4 with the die?

3.9 Summary

We conclude this unit by summarizing what we have covered in it:-

1. Defined and discussed in detail the conditional probability.
2. Did some examples related to conditional probability to clarify the concept.
3. Proved multiplication theorem of probability.
4. Defined independents events.
5. Proved multiplication theorem of probability for independent events.
6. Defined and discussed pairwise and mutually independent events.
7. Performed some examples to clarify the concept.
8. Proved the theorem on Law of total probability.
9. Proved Baye's theorem and did some examples by using this theorem.

3.10 Glossary:

1. Let A and B be two events and s be the sample space. Then P(B/A) g denoted the probability of B given that A has occurred, called the conditional probability of B given A and it is given by P(B/A) = $\dfrac{P(A\cap B)}{P(A)}$

2. For two events A and B in a sample space 3,

   $P(A\cap B) = P(A) \,.\, P(B/A),\ P(A) > 0$

and $P(A \cap B) = P(B) . (PA/B), P(B) > 0$

and it is called the multiplication theorem of probability.

3. Two or more events are said to be independent if the happening or non-happening of any one of them, does not in any way, asset the happening of others.

3.11 Answer to Self Check Exercise

Self Check Exercise-1

Ans.1 $P(A \cap B) = 0.2$

and $P(B/A) = \frac{2}{3}$

Ans. 2 $P(A/B) = \frac{1}{6}$

Ans. 3 $\frac{1}{4}$

Self Check Exercise-2

Ans.1 Required probability = $\frac{2}{3}$

Ans. 2 P (at least one doctor is available) = 0.98

Self Check Exercise-3

Ans.1 = $\frac{38}{91}$

Ans. 2 = $\frac{8}{11}$

3.12 References/Suggested Readings

1. Robert V. Hogg, Joseph w. Mckean and Allen T. Craig, Introduction to Mathematical Statistics, Pearson Education, Asia, 2007.

2. Irwin Miller, Marylees Miller and John E. Freund, Mathematical Statistics with Application, 7th Ed., Pearson Education, Asia, 2006.

3.13 Terminal Questions

1. If $P(A) = 0.4$, $P(B) = 0.8$, $P(B/A) = 0.6$,

   Find $P(A/B)$ and $P(A \cup B)$

2. A bag contain 10 gold and 8 silver coins. Two successive drawing of 4 coins are made such that

(i) coins are replaced before the second trail

(ii) the coins are not replaced before the second trial.

Find the probability that the first drawing will give 4 gold and the second 4 silver coins.

3. A purse contains 2 silver and 4 copper coins. A second purse contains 4 silver and 3 copper coins. If a coin is pulled one at random from one of the two purses, what is the probability that it is silver coin?

4. The probability of hitting a target by three marksmen are $\frac{1}{2}$, $\frac{1}{3}$ and $\frac{1}{4}$ respectively. Find the probability that one and only one of them will hit the target when they fire simultaneously.

5. An urn contains a white chips and b blue chips. A chip is chosen at random from the urn, discarded and replaced by one of opposite colour and then a second chip is drawn. Find the probability that second chip drawn is blue.

6. The probabilities of x, y and z becoming managers are $\frac{4}{9}$, $\frac{2}{9}$ and $\frac{1}{3}$ respectively. The probabilities that the bonus scheme will be introduced if x, y and z becomes managers are $\frac{3}{10}$, $\frac{1}{2}$ and $\frac{4}{5}$ respectively.

(i) What is the probability that bonus scheme will be introduced, and

(ii) If the bonus scheme has been introduced, what is the probability that the manager appointed was x?

Unit - 4

Probability Distribution of A Discrete Random Variable

Structure

4.1 Introduction

A Random variable is a numerical quantity that is one result of a random experiment or process. It is a variable that can take on different values with certain probabilities. A discrete random variable is a variable that can only take n a countable number of distinct values. The probability distribution of a random variable describes the likelihood of the variable taking on different values. For discrete random variables, the probability distribution is typically given as a probability mass function, which assigns a probability to each possible value. The expected value (or mean) of a random variable is the average or central tendency of the distribution and the variance of a random variable is a measure of the spread or dispersion of the distribution around the expected value.

4.2 Learning Objectives

After studying this unit, you should be able to:

- Define and discuss random variable.
- Define independent random variable.
- Define and discuss discrete random variable.
- Discuss probability distribution of discrete random variable.
- Define distributive function and able to discuss properties of distributive function.

- Do examples to have better idea if the concept.

4.3 Random Variable - Definition

By a random variable (r.v) we means a real number x associated with the outcomes of a random experiment. If can take any one of the various possible values each with definite probability. For example in tossing of a coin if x denotes the number of heads, then x is a random variable which can take any one of the two values ; 0 (No head i.e. tail) or 1 (i.e. head), each with equal probability $\frac{1}{2}$. Now consider a random experiment of three tosses of a coin (or three coins tossed simultaneously).

Then S = {HHH, HTH, THH, TTH, HHT, HTT, THT, TTT}

Let us consider the variable x, which is the number of heads obtained. The, x is a random variable which can take any one of the values 0, 1, 2, 3.

Sample point or outcome :	HHH	HTH	THH	TTH	HHT	HTT	THT	TTT
Values of X :	3	2	2	1	2	1	1	0

If the sample point in the above order be denoted by w_1, w_2, ... , w_8, then to each outcome w of the random experiment, we can assign a real number x = x (w). For example $x(w_1) = 3$, $x(w_2) = 2$, $x(w_3) = 2$, $x(w_4) = 1$, $x(w_5) = 2$, $x(w_6) = 1$, $x(w_7) = 1$, $x(w_8) = 0$. Therefore, Random variable may be defined as a real valued function on the sample space, taking values on the real time R $(-\infty, \infty)$. In other words, random variable is a function which takes real values which are determined by the outcomes of the random experiment.

Def. A function $X : S \rightarrow R$ is called a random variable, where S is the sample space of any random experiment.

Or

Random variable is a real values function defined on a sample space whose range is non-empty set of real numbers.

A variable whose value is a number determined by the outcome of an experiment is called random variable.

If $x \in R$, then the set of all $w \in s$ such that $X(w) = x$ is denoted by writing $X = x$.

$\therefore P(X = x) = P\{w : X(w) = x\}$

similarly, $P(X \leq a) = P\{w:X(w) \in (-\infty, a]\}$

and $P(A < X \leq b) = P\{w: X(w) \in (a, b]\}$.

Independent Random Variable

A random variable that does not have an effect on the other random variable in an experiment is called independent random variable.

4.4 Discrete Random Variable And Its Probability Distribution

Discrete Random Variable

If the random variable X assumes only a finite or at the most countable number of values, then it is known as discrete random variable. In other words, a real valued function described on a discrete sample space is called a discrete random variable. For example, marks obtained by students in a test, the number of students in a college, the number of accidents taking place on a busy road, etc. are all discrete random variables.

Probability Distribution

Consider a discrete random variable X which can take the possible values $x_1, x_2, \ldots\ldots, x_n$ and with each value of the variable X, associate a number.

$p_i = P(X = X_i)$; $i = 1,2,\ldots\ldots,n$ which is known as the probability of X_i and satisfies the following conditions:

(i) $p_i = P(X = X_i) \geq 0$, $(i = 1,2,\ldots.,n)$

i.e. p_i's are all non-negative and

(ii) $\{p_i = p_1 + p_2 + \ldots\ldots + p_n = 1$

i.e. the total probability is one.

Again, let X be a discrete random variable and $p(x) = P(X = x)$ such that $p(x) \geq 0$ and $\{p(x) = 1$, summation being taken over various values of the variable.

The function $pi = P(X = X_i)$ or $p(x)$ is called the probability function or probability mass function (p.m.s.) of the random variable X and the set of del possible ordered pairs $\{x, p(x)\}$ is called the probability distribution of the random variable X.

Def. If a random variable X takes values $x_1, x_2, \ldots.., x_n$ with respective probabilistic $p_1, p_2, \ldots.., p_n$, then the tabular description.

X:	x_1	x_2	....	x_n
P(X):	p_1	p_2	....	p_n

is called probability distribution of the random variable X.

4.5 Distribution Function

Let the random variable X takes the values $x_1, x_2, \ldots\ldots., x_n$ with respective probabilities $p_1, p_2, \ldots.., p_n$ and let $x_1 < x_2 < \ldots\ldots. < x_n$. The distributive function F(x) is defined as

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

where $P[X \leq x_i] = p_1 + p_2 + \ldots\ldots.. + p_i$

Another definition: If X is a discrete random variable then the function defined for all real x, is given by

$$F(x) = P[X \leq x] = \sum_{w \leq x} f(w)$$

is called distribution function.

Note1. Distributive function is abbreviated as d.f.

2. Distribution function is also sometimes called cumulative distribution function.

If X is a discrete r.v. with probability function p(x), them the distribution function, usually denoted by F(x) is defined as $F(x) = P(X \leq x)$

If X takes integral values i.e. 1,2,3,.........,x, then

$F(x) = P(X = 1) + P(X = 2) + ... + P(X = x)$

or $F(x) = p(1) + p(2) + p(3) + ... + p(4)$

4.6 Properties of Distributive Function

Property I : The distribution function F(x) is non-decreasing [i.e. $F(x_1) \leq F(x_2)$ if $x_1 < x_2$].

Proof: Let $x_1 < x_2$

Then $\{x : x \leq x_2\} = \{x : x \leq x_1\} \cup \{x : x_1 < x \leq x_2\}$ and therefore

$P(X < x_2) = P(X \leq x_1) + P(x_1 < X \leq x_2)$

i.e. $F(x_2) = F(x_1) + P(x_1 < X \leq x_2)$

$\Rightarrow F(x_2) - F(x_1) = P(x_1 < X \leq x_2) > 0$

$\Rightarrow F(x_2) > F(x_1)$

Thus F(x) is a non-decreasing function of x.

Property II : If F is the distribution function of a random variable X and if a < b, then

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

Proof: The events $a < X \leq b$ and $X \leq a$ are disjoint and their union is the event $X \leq b$.

Hence by addition theorem of probability.

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

$$\Rightarrow P(a < X \leq b) = P(X \leq b) - P(X \leq a)$$

$$= F(b) - F(a)$$

Hence the result

Property III: $0 \leq F(x) \leq 1$

Proof: We have, $F(x) = P(X \leq x\}$

Since probability of any event lies between 0 and 1.

$\therefore 0 \leq F(x) \leq 1$

Property IV: If F is the distribution function of the random variable X and if a < b, then

$P(a \leq X \leq b) = P(X = a) + [F(b) - F(a)]$

Proof: We have

$P(a \le X \le b) = P\{(X = a) \cup (a < X \le b)\}$

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

and using property II, we get

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

Property V: For the distribution function F of a random variable X,

(i) $\lim\limits_{x \to -\infty} F(x) = 0$

(ii) $\lim\limits_{x \to \infty} F(x) = 1$

Proof: The whole sample space S can be expressed as a countable union of disjoint events as follows:

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

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

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

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

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

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

Since $-\infty < \infty$, $F(-\infty) \le F(\infty)$

Also $F(-\infty) \ge 0$ and $F(\infty) \le 1$

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

From (1) and (2), we get

$F(-\infty) = 0$ and $F(\infty) = 1$

Hence the result.

Let us now do examples to have better idea of the concept:-

Example 1: Check whether the following can define probability mass function and explain your answer.

$f(x) = \dfrac{5 - x^2}{6}$; x = 0,1,2,3.

Sol: Here $f(x) = \frac{5-x^2}{6}$; x = 0,1,2,3.

For $f(x)$ to be a probability mass function, we have

$\sum_{x=0}^{3} f(x) = 1$ and $f(x) \geq 0$

Now $\sum_{x=0}^{3} f(x) = f(0) + f(1) + f(2) + f(3)$

$= \frac{5-0}{6} + \frac{5-1}{6} + \frac{5-4}{6} + \frac{5-9}{6}$

$= \frac{5}{6} + \frac{4}{6} + \frac{1}{6} - \frac{4}{6}$

$= \frac{5}{6} + \frac{1}{6} = 1$

$\therefore \sum_{x=0}^{3} f(x) = 1$

Again $f(3) = \frac{5-9}{6} = -\frac{4}{6} = -\frac{2}{3} < 0$

$\therefore$ $f(x) < 0$ for x = 3

$\therefore$ Given function is not a probability moss function.

Example 2: Given that $f(x) = \frac{k}{2^x}$ is a probability function for a random variable which can take on the values x = 0,1,2,3 and 4, Find k.

Sol. Here $f(x) = \frac{k}{2^x}$, x = 0,1,2,3,4

Since $f(x)$ is probability function

$\therefore$ $f(x) \geq 0 \; \forall \; x \Rightarrow k \geq 0$

and $\sum_{x=0}^{4} f(x) = 1$

$\Rightarrow$ $f(0) + f(1) + f(2) + f(3) + f(4) = 1$

$\Rightarrow$ $\frac{k}{2^0} + \frac{k}{2^1} + \frac{k}{2^2} + \frac{k}{2^3} + \frac{k}{2^4} = 1$

$\Rightarrow$ k + $\frac{k}{2}$ + $\frac{k}{4}$ + $\frac{k}{8}$ + $\frac{k}{16}$ = 1

$\therefore$ $\frac{16k+8k+4k+2k+k}{16}$ = 1

$\Rightarrow$ k $\left[\frac{16+8+4+2+1}{16}\right]$ = 1

or k $\left(\frac{31}{16}\right)$ = 1

$\Rightarrow$ k = $\frac{16}{31}$

Example 3: Find the constant c so that f(x) satisfies the condition of being a p.f. of the random variable X:

$$f(x) = \begin{cases} c\left(\frac{2}{3}\right)^x ; x = 1, 2, 3, ...; elsewhere \end{cases}$$

Sol. Here $f(x) = \begin{cases} c\left(\frac{2}{3}\right)^x ; x = 1, 2, 3, ...0; elsewhere \end{cases}$

Since f(x) is p.f.

$\therefore$ $\sum f(x) = 1$

$\therefore$ f(1) + f(2) + f(3) +........... = 1

$\Rightarrow$ c$\left(\frac{2}{3}\right)$ + c$\left(\frac{2}{3}\right)^2$ + c$\left(\frac{2}{3}\right)^3$ +............ = 1

$\Rightarrow$ c $\left[\frac{2}{3} + \left(\frac{2}{3}\right)^1 + \left(\frac{2}{3}\right)^2 + \left(\frac{2}{3}\right)^3 + ...\right]$ = 1

$\Rightarrow$ c$\left[\frac{2/3}{1-2/3}\right]$ = 1

$\Rightarrow$ $\left[\frac{2/3}{1/3}\right]$ = 1

$\Rightarrow$ c (2) = 1

$\Rightarrow$ c = $\frac{1}{2}$

Example 4: The probability mass function of a random variable X is given by

$$f(x) = \frac{CX^x}{\lfloor x}, \; x = 0,1,2,..........$$

Where λ is some positive number.

Find (i) P(X = 0) (ii) P(X > 2)

Sol. Now $\sum_{x=0}^{\infty} f(x) = 1$

$\Rightarrow$ $\sum_{x=0}^{\infty} \frac{CX^x}{\lfloor x} = 1$

$\Rightarrow$ c $\sum_{x=0}^{\infty} \frac{\lambda^x}{\lfloor x} = 1$

$\Rightarrow$ $ce^{\lambda} = 1$ $\left[\because e^x = \sum_{x=0}^{\infty} \frac{x^i}{\lfloor i} \right]$

$\Rightarrow$ c = $e^{-\lambda}$

$\Rightarrow$ $f(x) = \frac{e^{-\lambda}\lambda^x}{\lfloor x}$

(i) P(X = 0) = $f(0) = \frac{e^{-\lambda}\lambda^0}{\lfloor 0} = \frac{e^{-\lambda}(1)}{(1)} = e^{-\lambda}$

(ii) P(X > 2) = 1 - P(X $\leq$ 2)

= 1 - P(X = 0) -P(X = 1) -P(x = 2)

= 1- $\frac{e^{-\lambda}\lambda^0}{\lfloor 0}$ - $\frac{e^{-\lambda}\lambda}{\lfloor 1}$ - $\frac{e^{-\lambda}\lambda^2}{\lfloor 2}$

= $\frac{1-e^{-\lambda}(1)}{1}$ - $\frac{e^{-\lambda}\lambda}{1}$ - $\frac{e^{-\lambda}\lambda^2}{1\times 2}$

$$= 1 - e^{-\lambda} - \lambda\, e^{-\lambda} - \frac{\lambda^2 e^{-\lambda}}{1\times 2}$$

Example 5: A random variable X has the following probability functions:-

x:	0	1	2	3	4	5	6	7	8
p(x):	λ	3λ	5λ	7λ	9λ	11λ	13λ	15λ	17λ

(i) Find the value of λ

(ii) Evaluate $P(X \geq 6)$, $P(3 < X < 5)$

(iii) Calculate the minimum value of λ, such that $P(X \leq 3) > 0.5$.

Sol. (i) Since p(x) is a probability function

$$\therefore \quad \sum_{x=0}^{8} p(x) = 1$$

$$\Rightarrow \quad \lambda + 3\lambda + 5\lambda + 7\lambda + 9\lambda + 11\lambda + 13\lambda + 15\lambda + 17\lambda = 1$$

$$\therefore \quad 81\lambda = 1 \quad \Rightarrow \quad \lambda = \frac{1}{81}$$

(ii) $P(X > 6) = P(X = 6) + P(X = 7) + P(X = 8)$

$$= 13\lambda + 15\lambda + 17\lambda$$

$$= 45\lambda = \frac{45}{81} = \frac{5}{9}$$

$P(3 < X < 5) = P(X = 4)$

$$= 9\lambda = \frac{9}{81} = \frac{1}{9}$$

(iii) $P(X \leq 3) = P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3)$

$$= \lambda + 3\lambda + 5\lambda + 7\lambda = 16\lambda$$

Now $\quad P(X \leq 3) > 0.5 \Rightarrow 16\lambda > \frac{1}{2}$

or $\quad \lambda > \frac{1}{32}$

$\therefore \quad$ Minimum Value of λ is $\frac{1}{32}$

Example 6: In a single throw of two dice, find the probability function and the distribution function of the random variable X, where X denotes the total obtained with the pair of dice. Draw the graph of the distribution function and name it.

Sol. In a single throw of two dice, the possible values of the random variable X are 2,3,4,....., 12 and the corresponding probabilities are $\frac{1}{36}, \frac{2}{36}, \frac{3}{36}, \ldots\ldots, \frac{1}{36}$.

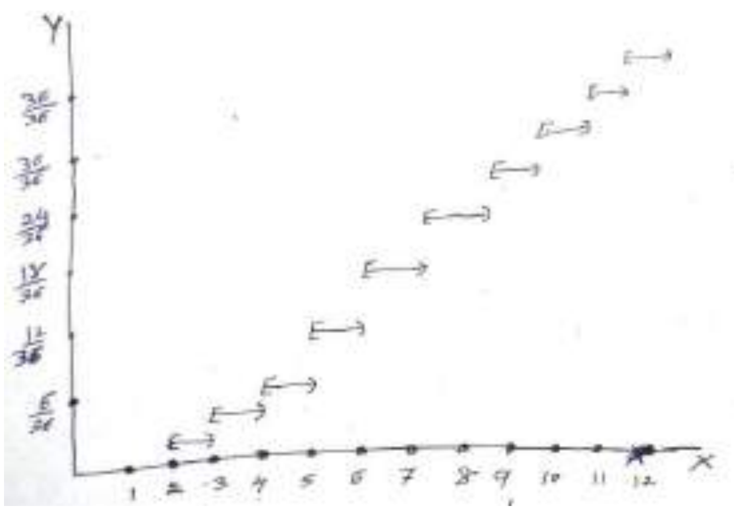
These results are shown in the table:

X=x	2	3	4	5	6	7	8	9	10	11	12
P[x=x]=f(x)	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$
F(x)=P(X≤x)	$\frac{1}{36}$	$\frac{3}{36}$	$\frac{6}{36}$	$\frac{10}{36}$	$\frac{15}{36}$	$\frac{21}{36}$	$\frac{26}{36}$	$\frac{30}{36}$	$\frac{33}{36}$	$\frac{35}{36}$	$\frac{36}{36}$

The above table can also be written in the following form:

$$
\begin{cases}
0 \quad for \ x < 2 \\
\frac{1}{36} \quad for \ 2 \le x < 3 \\
\frac{6}{36} \quad for \ 4 \le x < 5 \\
\frac{10}{36} \quad for \ 5 \le x < 6 \\
\frac{15}{36} \quad for \ 6 \le x < 7 \\
\frac{21}{36} \quad for \ 7 \le x < 8 \\
\frac{26}{36} \quad for \ 8 \le x < 9 \\
\frac{30}{36} \quad for \ 9 \le x < 10 \\
\frac{33}{36} \quad for \ 10 \le x < 11 \\
\frac{35}{36} \quad for \ 11 \le x < 12 \\
\frac{36}{36} \quad for \ 12 \le x
\end{cases}
$$

The graph of the distribution function is shown in the figure:



Distributive function is a step function

Example 7: A random Variable X has the following probability function:

x:	-2	-1	0	1	2	3
f(x):	0-1	k	0.2	2k	0.3	3k

(i) Find k

(ii) Evaluate P(X < 2), P(X $\geq$ 2), P(-2 < X < 2)

(iii) Determine the distribution function F(x) of X.

Sol. (i) Since f(x) is a probability function

$\therefore$ $$\sum_{x=-2}^{3} f(x) = 1$$

$\Rightarrow$ 0.1 + k + 0.2 + 2k + 0.3 + 3k = 1

$\Rightarrow$ 0.6 + 6k = 1

$\Rightarrow$ 6k = 1 - 0.6

$\Rightarrow$ 6k = 0.4

$\Rightarrow$ k = $\frac{0.4}{6} = \frac{4}{60}$

$\Rightarrow$ k = $\frac{1}{15}$

(ii) P(X < 2) = P(X = -2) + P(x = -1) + P(x = 0) + P(x = 1)

= 0.1 + k + 0.2 + 2k

= 0.3 + 3k

$= \frac{3}{10} + \frac{3}{15} = \frac{3+2}{10} = \frac{5}{10} = \frac{1}{2}$

(iii) P(X $\geq$ 2) = P(x = 2) + P(x = 3)

= 0.3 + 3k

$= \frac{3}{10} + \frac{3}{15} = \frac{3}{10} + \frac{1}{5}$

$= \frac{3+2}{10} = \frac{5}{10} = \frac{1}{2}$

P(-2 < X < 2) = P(x = -1) +P(x = 0) + P(x = 1)

= k + 0.2 + 2k

$= 3k + \frac{2}{10}$

$= \frac{3}{15} + \frac{2}{10} = \frac{1}{5} + \frac{1}{5} = \frac{2}{5}$

(iii) The distribution function F(x) is shown in the table given below:

x	-2	-1	0	1	2	3
f(x)	$\frac{3}{30}$	$\frac{2}{30}$	$\frac{6}{30}$	$\frac{4}{30}$	$\frac{9}{30}$	$\frac{5}{30}$
F(x)	$\frac{3}{30}$	$\frac{5}{30}$	$\frac{11}{30}$	$\frac{15}{30}$	$\frac{24}{30}$	$\frac{30}{30}$

Example 8: A bag contains two one rupee coins and 3 twenty paise coins. A person draws two coins. Find probability mass function of value of coins drawn.

Sol. Let X denote the value of the coin drawn X can take following value

$x_1 = 1 + 1 = 2$Rs. [When both one Rs. coins are drawn]

$x_2 = 1 + \frac{20}{100} = \frac{6}{5}$ Rs. [When one Rs. and one twenty paise coin is drawn]

$x_3 = \frac{20}{100} + \frac{20}{100} = \frac{2}{5}$ Rs. [When both twenty paise coins are drawn]

$\therefore$ P (X = x_1) = $\frac{2_{c_2} \times 3_{c_0}}{5_{c_2}} = \frac{1}{10}$

P(X = x_2) = $\frac{2_{c_1} \times 3_{c_1}}{5_{c_2}} = \frac{6}{10}$

P(X = x_3) = $\frac{2_{c_0} \times 3_{c_2}}{5_{c_2}} = \frac{3}{10}$

$\therefore$ Probability mass function of X is given by

X	$x_1 = 2$	$x_2 = \frac{6}{5}$	$x_3 = \frac{2}{5}$
P(X=x)=f(x)	P(X=2) = $\frac{1}{10}$	P(X=$\frac{6}{5}$) = $\frac{6}{10}$	P(X=$\frac{2}{5}$) = $\frac{3}{10}$

Example 9: From a lot of 12 items, having 4 defective items, 5 are drawn at random without replacement. If X denote the number of defectives in the sample:

(i) Write the probability distribution of X

(ii) Find P(1 < X < 3)

Sol. (i) 5 items can be drawn from a lot of 12 items in 12_{c_5} ways. If x denote the number of defective items, then x defectives can be chosen out of 4 defectives and from the remaining 8 items, 5 - x non-defective items can be chosen in $4_{c_x} \times 8_{c_{5-x}}$ ways.

Hence the probability function is

$$p(x) = \frac{4_{c_x} \times 8_{c_{5-x}}}{12_{c_5}}, \; x = 0,1,2,3,4$$

(ii) $P(1 < X < 3) = P(X = 2) = \dfrac{4_{c_2} \times 8_{c_3}}{12_{c_5}}$

$$= \frac{\dfrac{4\times3}{1\times2} \times \dfrac{8\times7\times6}{1\times2\times3}}{\dfrac{12\times11\times10\times9\times8}{1\times2\times3\times4\times5}} = \frac{14}{33}$$

Self-Check Exercise

Q.1 A bag contains two one rupee coins and 3 twenty five paise coins. A person draws two coins. Find probability mass function of value of coins drawn.

Q.2 Suppose that an urn contains a red and 4 blue balls. If 5 balls are selected at random without replacement, determine the probability function of the number of red balls that will be obtained.

Q.3 For the following, find constant $\subset$ so that f(x) satisfies the condition of probability mass function.

$$f(x) = \begin{cases} ex; x = 0,1,2,3,4,5,6 \\ 0; elsewhere \end{cases}$$

Q.4 Find the probability distribution of the number of heads when three coins are tossed simultaneously

4.7 Summary

We conclude this unit by summarizing what we have covered in it:-

1. Defined and discussed random variable. Also defined independent random variable.

2. Defined and discussed discrete random variable.

3. Discussed probability distribution of discrete random variable.

4. Defined distribution function. Discussed in detail the properties of distributive function.

5. Did some examples related to each topic so that concept be clarified.

4.8 Glossary:

1. By a random variable (r.v.) we mean a real number X associated with the outcomes of a random experiment It can take any one of the various possible values each with definite probability.

2. If a random variable X assumes only a finite or at the most countable number of values, then it is known as discrete random variable.

4.9 Answer to Self Check Exercise

Ans.1 Probability mass function of X is given by

X	$x_1 = 2$	$x_2 = \frac{5}{4}$	$x_3 = \frac{1}{2}$
P(X=x) = f(x)	P(X=2) = $\frac{1}{10}$	P(X=$\frac{6}{5}$) = $\frac{6}{10}$	P(X=$\frac{2}{5}$) = $\frac{3}{10}$

Ans. 2 Probability function is given by

X	1	2	3	4	5
P(X)	$\frac{9}{1287}$	$\frac{144}{1287}$	$\frac{504}{1287}$	$\frac{504}{1287}$	$\frac{126}{1287}$

Ans. 3 $\subset = \frac{1}{21}$

Ans. 4 Probability distribution table is

X	0	1	2	3
P(X)	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

4.10 References/Suggested Readings

1. Robert V. Hogg, Joseph w. Mckean and Allen T. Craig, Introduction to Mathematical Statistics, Pearson Education, Asia, 2007.

2. Sheldon Ross, Introduction to Probability Model, 9th Ed., Academic Press, Indian Reprint, 2007.

3. Irwin Miller, Marylees Miller and John E. Freund, Mathematical Statistics with Application, 7th Ed., Pearson Education, Asia, 2006.

4.11 Terminal Questions

1. The bad eggs are mixed with 10 good ones. Three eggs are drawn at random without replacement. Determine the probability function of number of bad eggs that will be obtained.

2. Four balls are drawn from a bag containing 5 black, 6 white and 7 red balls. Let X be number of white balls drawn. Find probability mass function or probability distribution of X.

3. Check whether the following can define probability mass function and explain your answer:

$$f(x) = \frac{5 - x^2}{6}, \quad x = 0,1,2,3$$

4. A fair die is rolled once. Find the probability function and the distribution function of the number of points appearing on the top face.

5. Suppose that a box contain 7 red and 3 blue balls. If 5 balls are selected at random with replacement, determine the probability function of the number of red balls that will be obtained.

Unit - 5

Probability Distribution of A Continuous Random Variable

Structure

5.1 Introduction

A continuous random variable is a variable that can take on any value within a specified range or interval. Unlike discrete random variables, which can only take on specific, countable values, continuous random variables can assume an infinite number of possible values. The probability distribution of a continuous random variable is described by a probability density function (pds) denoted as $f(x)$. The pdf represents the relative likelihood of the random variable taking on a particular value within the specified range. The pdf $f(x)$ is non-negative for all values of the random variable x. The total area under the probability density function curve is equal to 1. The expected value (or mean) of a continuous random variable is the weighted average of all possible values, where the weights are the probabilities associated with each value. Understanding continuous probability distributions has a wide range of real world applications across various fields like in Engineering and Physics; Quality control and Reliability engineering; Biology and Ecology: Healthcare and Medicine etc.

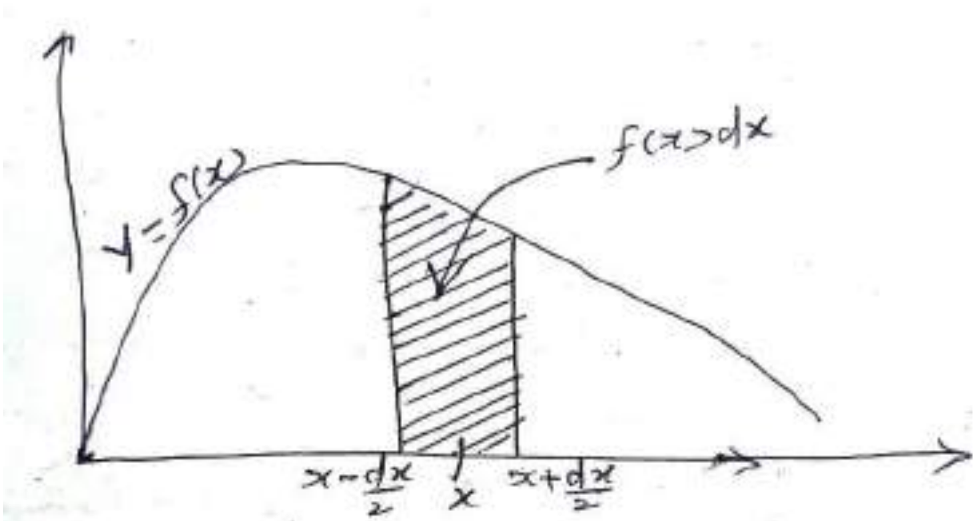
5.2 Learning Objectives

After studying this unit, you should be able to:-

- Define continuous random variable
- Discuss probability density function of a continuous random variable
- Define continuous distribution function and discuss properties of continuous distribution function.
- Do questions related to these concepts.

5.3 Continuous Distribution

A random variable X is said to be continuous if it can take all possible values between certain limits. In other words, a random variable is said to be continuous when its different values cannot be put in 1-1 correspondence with a set of positive numbers.



Probability Density Function

The probability density function f(x) of a continuous random variable X is defined as

$$f(x) = \lim_{\delta x \to 0} \frac{P(x \le X \le x + \delta x)}{\delta x}$$

Where $P(x \le X \le x + \delta x)$

Properties of Probability density function f(x)

(i) $f(x) > 0$

(ii) $\int_{-\infty}^{\infty} f(x)dx = 1$

(iii) $P(a < X < b) = \int_{a}^{b} f(x)dx$

(iv) For any event E,

$P(E) = \int_{E} f(x)dx$ is well defined.

5.4 Continuous Distribution Function

If X is a continuous random variable with the probability distribution function f(x), then the function.

F(x) = P(X < X) = $\int_{-\infty}^{x} f(t)dt, -\infty < t < \infty$

is called the distribution function or sometimes the Cumulative distribution function of the random variable X.

Properties of Distribution Function:

The distribution function defined by equation (1) as above obeys the following properties:

1. $0 < F(x) < 1, -\infty < x < \infty$

2. $F'(x) = \frac{d}{dx}F(x) = f(x) > 0$

   i.e. F(x) is a non-decreasing function of x.

3. $F(-\infty) = \underset{x \to -\infty}{Lim} F(x) = \underset{x \to -\infty}{Lim} \int_{-\infty}^{x} f(t)dt$

   $= \int_{-\infty}^{-\infty} f(t)dt = 0$

   and $F(+\infty) = \underset{x \to +\infty}{Lim} F(x) = \underset{x \to +\infty}{Lim} \int_{-\infty}^{x} f(t)dt$

   $= \int_{-\infty}^{\infty} f(t)dt = 1$

4. F(x) is a continuous function of x on the right

5. The discontinuities of F(x) are at the most countable

6. It may be noted that

   $P(a \leq X \leq b) = \int_{a}^{b} f(x)dx = \int_{-\infty}^{b} f(x)dx - \int_{-\infty}^{a} f(x)dx$

   $= P(X \leq b) - P(X \leq a)$

   $= F(b) - F(a)$

   $= F(b) - F(a)$

and $P(a < X < b) = P(a < X \leq b) = P(a \leq X < b)$

$= \int_{a}^{b} f(t)dt$

Let us now do some examples to have better idea of the concept:-

Example 1: Test if following is a probability density function:

$$f(x) = \begin{cases} x, 0 \le x < 1 \\ 2x, 1 \le x < 2 \end{cases}$$

Sol. If f(x) is a probability distribution function, it must satisfy

(i) $f(x) \geq 0$, which is true for all the given values of x.

(ii) $\int_{-\infty}^{\infty} f(x)dx = 1$

Now $\int_{-\infty}^{\infty} f(x)dx = \int_{0}^{2} f(x)dx = \int_{0}^{1} xdx + \int_{1}^{2} 2xdx$

$$= \left[\frac{x^2}{2}\right]_0^1 + \left[\frac{2x^2}{2}\right]_1^2$$

$$= \left[\frac{1}{2} - 0\right] + (4 - 1)\ \frac{1}{2} + 3 = \frac{7}{2} \neq 1$$

$\therefore$ f(x) is not a probability distribution function

Example 2: Find the constant c so that f(x) satisfies the conditions of being a p.d.f. of a random variable X:

$$f(x) = \begin{cases} exe^{-x}, 0 < x < \infty \\ 0, elsewhere \end{cases}$$

Sol. We have

$$f(x) = \begin{cases} cxe^{-x}, 0 < x < \infty \\ 0, elsewhere \end{cases}$$

Now $\int_{-\infty}^{\infty} f(x)dx = 1$

$\Rightarrow$ $\int_{-\infty}^{0} f(x)dx + \int_{0}^{\infty} f(x)dx = 1$

$\Rightarrow$ $\int_{0}^{\infty} c\,x\,e^{-x}dx = 1$

$\therefore$ c $\int_0^\infty x e^{-x} dx = 1$

$\Rightarrow$ c $\left\{ \left[\frac{xe^{-x}}{-1} \right]_0^\infty - \int_0^\infty 1.\frac{e^{-x}}{-1} dx \right\}$ = 1

or c $\left\{ \left[\frac{-x}{e^x} \right]_0^\infty - \int_0^\infty e^{-x} dx \right\}$ = 1

$\Rightarrow$ c $\left\{ (0-0) + \left[\frac{e^{-x}}{-1} \right]_0^\infty \right\}$ = 1

$\Rightarrow$ c $\left[-\frac{1}{e^x} \right]_0^\infty$ = 1

$\Rightarrow$ c [0 + 1] = 1

$\Rightarrow$ c = 1

Example 3: If the density function of a random variable X is given by

$$f(x) = \begin{cases} x, 0 < x < 1 \\ 2 - x, 1 \le x < a \\ 0, \text{ elsewhere} \end{cases}$$

Find (i) The value of a

(ii) The distribution function of x.

(iii) P (0.8 < X < 0.6 a).

Sol. (i) Here

$$f(x) = \begin{cases} x, 0 < x < 1 \\ 2 - x, 1 \le x < a \\ 0, \text{ elsewhere} \end{cases}$$

$\therefore$ $\int_{-\infty}^{\infty} f(x) dx = 1$

$\Rightarrow$ $\int_0^1 f(x) dx$ + $\int_1^a f(x) dx$ + $\int_a^\infty f(x) dx$ = 1

$\Rightarrow \quad \int_0^1 x dx + \int_1^a (2-x) dx + 0 = 1$

$\Rightarrow \quad \left[\frac{x^2}{2}\right]_0^1 + \left[2x - \frac{x^2}{2}\right]_1^a = 1$

$\Rightarrow \quad \left(\frac{1}{2} - 0\right) + \left[\left(2a - \frac{a^2}{2}\right) - \left(2 - \frac{1}{2}\right)\right] = 1$

$\Rightarrow \quad \frac{1}{2} + 2a - \frac{a^2}{2} - \frac{3}{2} = 1$

or $\quad 2a - \frac{a^2}{2} = 2$

or $\quad 4a - a^2 = 4$

$\Rightarrow \quad a^2 - 49 + 4 = 0$

$\Rightarrow \quad (a - 2)^2 = 0$

$\Rightarrow \quad a = 2$

$\therefore$ the p.d.f.

$$f(x) = \begin{cases} x, 0 < x < 1 \\ 2 - x, 1 \le x < 2 \\ 0, \, elsewhere \end{cases}$$

(ii) $\quad F(x) = P(X < x) = \int_{-\infty}^{x} f(x) dx$

For $\quad x < 0, F(x) = 0$

For $\quad 0 < x < 1, F(x) = \int_{-\infty}^{x} f(x) dx = \int_{-\infty}^{0} 0\, dx + \int_0^x x\, dx$

$= 0 + \left[\frac{x^2}{2}\right]_0^x = 0 + \frac{x^2}{2} - 0 = \frac{x^2}{2}$

For $\quad 1 \le x < 2, F(x) = \int_{-\infty}^{x} f(x) dx = \int_{-\infty}^{0} 0\, dx + \int_0^1 x\, dx + \int_1^x (2-x) dx$

$$= 0 + \left[\frac{x^2}{2}\right]_0^1 + \left[2x - \frac{x^2}{2}\right]_1^x$$

$$= 0 + \left(\frac{1}{2} - 0\right) + \left(2x - \frac{x^2}{2}\right) - \left(2 - \frac{1}{2}\right)$$

$$= 2x - \frac{x^2}{2} - 1$$

For $x \geq 2$, $F(x) = \int_{-\infty}^{x} f(x)dx$

$$= \int_{-\infty}^{0} 0\,dx + \int_{0}^{1} x\,dx + \int_{1}^{2} (2-x)dx + \int_{2}^{x} 0\,dx$$

$$= 0 + \left[\frac{x^2}{2}\right]_0^1 + \left[2x - \frac{x^2}{2}\right]_1^2 + \left[2x - \frac{x^2}{2}\right]_1^2 + 0$$

$$= \left(\frac{1}{2} - 0\right) + \left[(4-2) - \left(2 - \frac{1}{2}\right)\right]$$

$$= \frac{1}{2} + 2 - \frac{3}{2} = 1$$

$$\therefore \quad F(x) = \begin{cases} 0, x \leq 0 \\ \frac{x^2}{2}, x \leq 1 \\ 2x - \frac{x^2}{2} - 1, \ x \leq 2 \\ 1, \ x \geq 2 \end{cases}$$

(iii) $P(0.8 < X < 0.69) = P(0.8 < X < 1.2)$

$$= \int_{0.8}^{1.2} f(x)dx = \int_{0.8}^{1} x\,dx + \int_{1}^{1.2} (2-x)dx$$

$$= \left[\frac{x^2}{2}\right]_{0.8}^1 + \left[2x - \frac{x^2}{2}\right]_1^{1.2}$$

$= \dfrac{1-0.64}{2} + \left[\left\{2(1.2)-0.\dfrac{(1.2)^2}{2}\right\}-\left\{2-\dfrac{1}{2}\right\}\right]$

= 0.18 + 0.18

= 0.36

Example 4: A random variable X has the following density function

$$f(x) = \begin{cases} \dfrac{k}{1+x^2}, if -\infty < x < \infty \\ 0 \ , \ \ otherwise \end{cases}$$

Determine k and the distribution function. Also find the probability that X^2 lies between $\dfrac{1}{3}$ and 1.

Sol. We have

$$f(x) = \begin{cases} \dfrac{k}{1+x^2}, -\infty < x < \infty \\ 0 \ , \ \ otherwise \end{cases}$$

Now $\displaystyle\int_{-\infty}^{\infty} f(x)dx = 1$

$\Rightarrow$ $\displaystyle k\int_{-\infty}^{\infty} \frac{1}{1+x^2}dx = 1 \Rightarrow k\left[\tan^{-1} x\right]_{-\infty}^{\infty} = 1$

$\Rightarrow$ $k\left[\tan^{-1}\infty - \tan^{-1}(-\infty)\right] = 1$

$\Rightarrow$ $k\left[\dfrac{\pi}{2}-\left(-\dfrac{\pi}{2}\right)\right] = 1 \Rightarrow k\pi = 1$

$\Rightarrow$ $k = \dfrac{1}{\pi}$

Again $F(x) = \displaystyle\int f(x)dx$

$= k\displaystyle\int \frac{1}{1+x^2}dx$

$\therefore$ $F(x) = k\left[\tan^{-1} x + c\right]$ ......(1)

Now $F(-\infty) = 0 \Rightarrow k\left[\tan^{-1}\infty + c\right] = 0$

$\Rightarrow \quad \frac{1}{\pi}\left[-\frac{1}{\pi} + c\right] = 0$

$\Rightarrow \quad c = \frac{\pi}{2}$

$\therefore$ from (1), we have

$$F(x) = \frac{1}{\pi}\left[\tan^{-1}x + \frac{\pi}{2}\right]$$

The probability distribution function is

$$f(x) = \frac{1}{\pi(x^2+1)}, \ -\infty < x < \infty$$

Now, if $\frac{1}{3} \le X^2 \le 1$

then $X^2 \le 1 \Rightarrow |X| \le 1$

$\Rightarrow \quad -1 \le X \le 1$ .....(2)

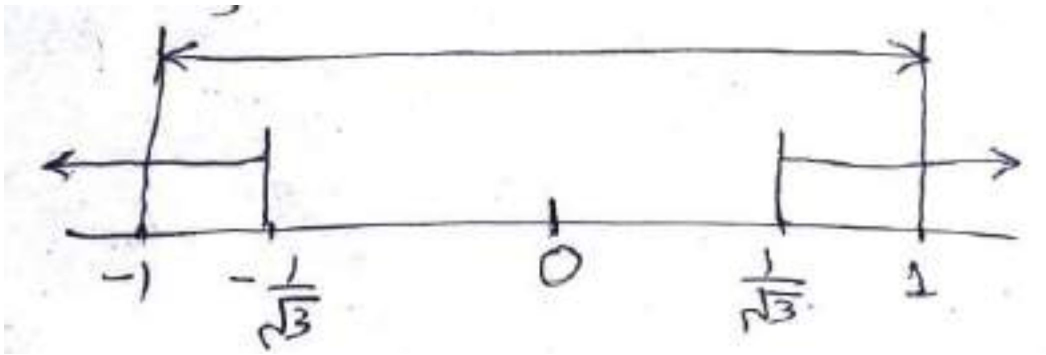
and $X^2 \ge \frac{1}{3}$

$\Rightarrow \quad |X| \ge \frac{1}{\sqrt{3}}$

$\Rightarrow \quad X \ge \frac{1}{\sqrt{3}}$ or $X \le -\frac{1}{\sqrt{3}}$ ......(3)

Combining (2) and (3), we have

$$\frac{\sqrt{3}}{3} \le X \le 1 \quad \text{or} \quad -1 \le X \le -\frac{\sqrt{3}}{3}$$



Thus, the required probability is

P$\left(\frac{1}{3} \le X^2 \le 1\right)$ = P $\left(\frac{\sqrt{3}}{3} \le X \le 1 \; or \; -1 \le X \le -\frac{\sqrt{3}}{3}\right)$

= P$\left(\frac{\sqrt{3}}{3} \le X \le 1\right)$ + P $\left(-1 \le X \le -\frac{\sqrt{3}}{3}\right)$

$$= \frac{1}{\pi} \int_{\sqrt{3}/3}^{1} \frac{dx}{x^2+1} + \frac{1}{\pi} \int_{-1}^{-\sqrt{3}/3} \frac{dx}{x^2+1}$$

$$= \frac{2}{\pi} \int_{\sqrt{3}/3}^{1} \frac{dx}{x^2+1} = \frac{2}{\pi} \left[\tan^{-1} x\right]_{\sqrt{3}/3}^{1}$$

$$= \frac{2}{\pi} \left[\tan^{-1}(1) - \tan^{-1}\left(\frac{\sqrt{3}}{3}\right)\right]$$

$$= \frac{2}{\pi} \left(\frac{\pi}{4} - \frac{\pi}{6}\right) = \frac{1}{6}$$

Example 5: The probability density function of a random variable X is given by

$$f(x) = \begin{cases} \frac{c}{\sqrt{x}}, \; 0 < x < 4 \\ 0 \;\;, \;\; elsewhere \end{cases}$$

Where c is a constant. Find c and then compute

(i) P$\left(x < \frac{1}{4}\right)$ (ii) P(X > 1)

Sol. Now f(x) is probability density function of X

$\therefore \quad f(x) \geq 0 \quad \forall x \quad \Rightarrow \subset \geq 0$

Also $\displaystyle\int_0^4 f(x)dx = 1$

$\therefore \quad \displaystyle\int_0^4 \frac{e}{\sqrt{x}} dx = 1 \quad \Rightarrow \quad c\int_0^4 x^{-1/2} dx = 1$

$\Rightarrow \quad c\left[\dfrac{x^{1/2}}{1/2}\right]_0^4 = 1$

$\Rightarrow \quad 2c\left[\sqrt{x}\right]_0^4 = 1$

$\Rightarrow \quad 2c\,[2 - 0] = 1 \quad \Rightarrow \quad 4c = 1$

$\Rightarrow \quad c = \dfrac{1}{4}$

$\therefore \quad f(x) = \begin{cases} \dfrac{1}{4\sqrt{x}}, & 0 < x < 4 \\ 0, & elsewhere \end{cases}$

(i) $\quad P\left(x < \dfrac{1}{4}\right) = \displaystyle\int_0^{1/4} \frac{1}{4\sqrt{x}} dx = \frac{1}{4}\int_0^{1/4} x^{-1/2} dx$

$= \dfrac{1}{4}\left[\dfrac{x^{1/2}}{1/2}\right]_0^{1/4}$

$= \dfrac{1}{2}\left[\sqrt{x}\right]_0^{1/4}$

$= \dfrac{1}{2}\left[\dfrac{1}{2} - 0\right] = \dfrac{1}{4}$

(ii) $\quad P(X > 1) = \displaystyle\int_1^4 \frac{1}{4\sqrt{x}} dx = \frac{1}{4}\int_1^4 x^{-1/2} dx$

$= \dfrac{1}{4}\left[\dfrac{x^{1/2}}{1/2}\right]_1^4 = \dfrac{1}{2}\left[\sqrt{x}\right]_1^4$

$$= \frac{1}{2}(2 - 1) = \frac{1}{2}$$

Example 6: Let $f(x) = \begin{cases} \frac{2x}{9}, & 0 < x < 3 \\ 0, & elsewhere \end{cases}$

Then find $P(A_1)$, $P(A_2)$ and $P(A_1 \cup A_2)$,

Where $A_1 = \{x: 0 < x < 1\}$ and $A_2 = \{x : 2 < x < 3\}$

Sol. Here $f(x) = \begin{cases} \frac{2x}{9}, & 0 < x < 3 \\ 0, & elsewhere \end{cases}$

Now $P(A_1) = P(0 < x < 1) = \int_0^1 f(x)dx = \int_0^1 \frac{2x}{9}dx$

$$= \frac{2}{9}\int_0^1 xdx$$

$$= \frac{2}{9}\left[\frac{x^2}{2}\right]_0^1 = \frac{1}{9}\left[x^2\right]_0^1 = \frac{1}{9}(1 - 0) = \frac{1}{9}$$

$P(A_2) = P(2 < x < 3) = \int_2^3 f(x)dx = \int_2^3 \frac{2x}{9}dx$

$$= \frac{2}{9}\int_2^3 xdx = \frac{2}{9}\left[\frac{x^2}{2}\right]_2^3 = \frac{1}{9}\left[x^2\right]_2^3$$

$$= \frac{1}{9}(9 - 4) = \frac{5}{9}$$

Also $A_1 \cap A_2 = \phi \Rightarrow P(A_1 \cap A_2) = 0$

Now $P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)$

$$= \frac{1}{9} + \frac{5}{9} - 0$$

$$= \frac{1+5}{9} = \frac{6}{9} = \frac{2}{3}$$

Example 7 : A continuous random variable x has p.d.f. $f(x) = 3x^2, 0 \leq x \leq 1$

Find a and b such that

(i) P(x $\leq$ a) = P (x > a), and

(ii) P(x > b) = 0.05

Sol. - (i) Here f(x) = $3x^2$, $0 \leq x \leq 1$

Now P(x $\leq$ a) = P (x > a)

$\Rightarrow \quad \int_0^a f(x)dx = \int_a^1 f(x)dx$

$\therefore \quad \int_0^a 3x^2 dx = \int_a^1 3x^2 dx \Rightarrow \int_0^a x^2 dx = \int_a^1 x^2 dx$

$\Rightarrow \quad \left[\frac{x^3}{3}\right]_0^a = \left[\frac{x^3}{3}\right]_a^1$

$\Rightarrow \quad a^3 = 1 - a^3 \Rightarrow 2a^3 = 1$

$\therefore \quad a^3 = \frac{1}{2} \Rightarrow a = \left(\frac{1}{2}\right)^{1/3}$

(ii) Again P(x > b) = 0.05

$\Rightarrow \quad \int_b^1 f(x)dx = 0.05$

$\Rightarrow \quad \int_b^1 3x^2 dx = 0.05$

or $\quad \left[\frac{x^3}{3}\right]_b^1 = \frac{5}{100}$

$\Rightarrow \quad 1 - b^3 = \frac{5}{100}$

$\Rightarrow \quad b^3 = 1 - \frac{5}{100}$

$\therefore \quad b^3 = \frac{95}{100}$

$\Rightarrow \quad b = \left(\frac{19}{20}\right)^{1/3}$

Example 8 : The amount of bread (in hundreds of pounds) x that a certain bakery is able to sell in a day in found to be a numerical valued random phenomenon, with a probability function specified by the probability density function f(x), given by

$$f(x) = \begin{cases} Ax, \text{ for } 0 \le x < 5 \\ A(10-x), \text{ for } 5 \le x < 10 \\ 0, \text{ otherwise} \end{cases}$$

(i) Find the value of A such that f(x) is p.d.f.

(ii) What is the probability that the number of pounds of bread that will be sold tomorrow is

(a) more than 500 pounds,

(b) less than 500 pounds,

(c) between 250 and 750 pounds?

Sol. : (i) Here

$$f(x) = \begin{cases} Ax, \text{ for } 0 \le x < 5 \\ A(10-x), \text{ for } 5 \le x < 10 \\ 0, \text{ otherwise} \end{cases}$$

Since f(x) is p.d.f

$$\therefore \quad \int_{-\infty}^{\infty} f(x)dx = 1 \Rightarrow \int_{10}^{5} f(x)dx + \int_{5}^{10} f(x)dx + \int_{10}^{\infty} f(x)dx = 1$$

$$\Rightarrow \quad \int_{10}^{5} f(x)dx + \int_{5}^{10} f(x)dx = 1$$

$$\Rightarrow \quad \int_{0}^{5} Axdx + \int_{5}^{10} A(10-x)dx = 1$$

$$\text{or} \quad A\int_{0}^{5} xdx + A\int_{5}^{10} (10-x)dx = 1$$

$$\Rightarrow \quad A\left[\frac{x^2}{2}\right]_0^5 + A\left[10x - \frac{x^2}{2}\right]_5^{10} = 1$$

$$\Rightarrow \quad A\left[\frac{25}{2} - 0\right] + A\left[\left(100 - \frac{100}{2}\right) - \left(50 - \frac{25}{2}\right)\right] = 1$$

$$\Rightarrow \quad A\left(\frac{25}{2}\right) + A\left[100 - \frac{100}{2} - 50 + \frac{25}{2}\right] = 1$$

$\therefore$ A $\left(\frac{25}{2}\right)$ + A $\left(\frac{25}{2}\right)$ = 1

$\Rightarrow$ A(25) = 1 $\Rightarrow$ A = $\frac{1}{25}$

(ii) (a) P(x > 500) = $\int_5^{10} A(10-x)dx$

$= A\left[10x - \frac{x^2}{2}\right]_5^{10}$

$= \frac{1}{25}\left[100 - \frac{100}{2} - 50 + \frac{25}{2}\right]$

$= \frac{1}{25}\left(\frac{25}{2}\right) = \frac{1}{2}$

$\therefore$ P(x > 500) = 0.5

(b) P(x < 500) = $\int_0^5 Axdx$ = A$\left[\frac{x^2}{2}\right]_0^5$

$= \frac{1}{25}\left[\frac{25}{2} - 0\right] = \frac{1}{25}\left(\frac{25}{2}\right) = \frac{1}{2}$ = 0.5

$\therefore$ P(X < 500) = 0.5

(c) P(250 < X < 750) = $\int_{2.5}^{3} f(x)dx + \int_{5}^{7.5} f(x)dx$

$= \int_{2.5}^{5} Axdx + \int_{5}^{7.5} A(10-x)dx$

= A $\int_{2.5}^{5} xdx$ + A $\int_{5}^{7.5} (10-x)dx$

$= \frac{1}{25}\left[\frac{x^2}{2}\right]_{2.5}^{5} + \frac{1}{25}\left[10x - \frac{x^2}{2}\right]_{5}^{7.5}$

$= \frac{1}{25}\left[\frac{25}{2} - \frac{6.25}{2}\right] + \frac{1}{25}\left[75 - \frac{56.25}{2} - (50 - \frac{25}{2}\right]$

$$= \frac{1}{25}\left(\frac{25}{2}-\frac{6.25}{2}\right) + \frac{1}{25}\left(75-\frac{56.25}{2}-(50-\frac{25}{2}\right)$$

$$= \frac{1}{25}\left(\frac{25-6.25}{2}\right) + \frac{1}{25}\left(25-\frac{31..25}{2}\right)$$

$$= \frac{1}{25}\left(\frac{18.75}{2}\right)+\frac{1}{25}\left(\frac{50-31.25}{2}\right)$$

$$= \frac{1}{25}\left(\frac{18.75}{2}\right)+\frac{1}{25}\left(\frac{18.75}{2}\right)$$

$$= \frac{1}{25}\left(\frac{18.75}{2}+\frac{18.75}{2}\right) = \frac{18.75}{25}$$

= 0.75

Self-Check Exercise

Q.1 Find the value of $\subset$ such that f(x) = $\subset e^{-x}$; $0 < x < \infty$ represents a probability density function.

Q.2 Let $f(x) = \begin{cases} c\,x\,e^{-2x}, x \geq 0 \\ 0\,, elsewhere \end{cases}$

be the probability density function of a random variable X. Then find

(i) Constant c (ii) the distribution F(x)

(iii) P(2 $\leq$ X < 3) (iv) P(X $\geq$ 1)

Q.3 Let f(x) be the p.d.f. of a random variable X, find the distribution function F(x) of X if $f(x) = \frac{2}{x^3}$, $1 < x < \infty$, zero elsewhere.

Q.4 Let X be the number of gallons of ice-cream that is required at a certain store on a hot summer day.

Let $f(x) = \begin{cases} \frac{12x(1000-x)^2}{12^{12}}, 0 < x < 1000 \\ 0\,, elsewhere \end{cases}$

be the probability density function if X. How many gallons of ice-cream should the store have in hand each of these days so that the probability of exhausting its supply on a particular day is 0.05

5.5 Summary

We conclude this unit by summarizing what we have covered in it:-

1. Defined continuous random variable.
2. Discussed probability density function of continuous random variable.
3. Defined continuous distribution function and discussed in detail she properties of continuous distribution function.
4. Performed some examples related to each topic so that each concept be clarified further.

5.6 Glossary:

1. A random variable X is said to be continuous if it can take all possible values between certain limits.
2. The probability density function f(x) of a continuous random variable X is defined as

$$f(x) = \lim_{\delta x \to 0} \frac{(x \leq X \leq x + \delta x}{\delta x}$$

Where P(x $\leq$ X $\leq$ x + δx) = f(x) dx

3. If X is a continuous random variable with the probability distribution function f(x), then the function

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)dt \text{ , } -\infty < t < \infty$$

is called the distribution function or sometimes the cumulative distribution function of the random variable X.

5.7 Answers To Self-Check Exercise

Ans.1 $\subset$ = 1

Ans.2 (i) $\subset$ = 4

(ii) The distribution of function is

$$F(x) = \begin{cases} 0, \quad x < 0 \\ 1 - 2xe^{-2x} - e^{-2x} \; ; 0 \leq x < \infty \\ 1 \; ; \quad x \to \infty \end{cases}$$

(iii) P(2 $\leq$ X $\leq$ 3) = $5e^{-4}$ - 76-6

(iv) P(X $\geq$ 1) = $3e^{-2}$

Ans.3 Distribution function

$$F(x) = \begin{cases} 0, & x < 1 \\ 1 - \dfrac{1}{x^2}, & 1 \le x \end{cases}$$

Ans.4 Required number of gallons of ice-cream = 751.395 nearly.

5.8 References/Suggested Readings

1. Irwin Miller, Marylees Miller and John E. Freund, Mathematical Statistics with Application, 7th Ed., Pearson Education, Asia, 2006.
2. Robert V. Hogg, Joseph w. Mckean and Allen T. Craig, Introduction to Mathematical Statistics, Pearson Education, Asia, 2007.
3. Sheldon Ross, Introduction to Probability Model, 9th Ed., Academic Press, Indian Reprint, 2007.

5.9 Terminal Questions

1. Is the function defined as follows a probability density function?

$$f(x) = \begin{cases} 0, & x < 2 \\ \dfrac{1}{8}(3+2x), & 2 \le x \le 4 \\ 0, & elsewhere \end{cases}$$

2. Show that the function

$$f(x) = \begin{cases} |x|, & -1 < x < 1 \\ 0, & elsewhere \end{cases}$$

is a possible probability distribution function. Also find its distribution function.

3. A continuous random variable X has the distribution function

$$F(x) = \begin{cases} 0, & if\ x \le 1 \\ k(x-1)^4, & if\ 1 < x \le 3 \\ 1, & if\ x > 3 \end{cases}$$

4. The distribution function for a random variable X is

$$F(x) = \begin{cases} 0, & x < 0 \\ 1 - e^{-2x}, & x \ge 0 \end{cases}$$

Find

(i) the probability density function of X

(ii) the probability that X > 2

(iii) the probability that $-3 < X \leq 4$

5. If $f(x) = c\,x^2$; $0 < x < 1$ is the probability density function of a continuous random variable X. Find

(i) constant c (ii) $P\left(\frac{1}{3} < x < \frac{1}{2}\right)$

(iii) Find 'a' such that $P(X \leq a) = P(X > a)$

(iv) Find 'b' such that $P(X > b) = 0.05$

6. A petrol pump is supplied with petrol once in a day. If its daily volume of sale (X) is thousand liters is distributed as

$$f(x) = 5\,(1 - x)^4 \; ; \; 0 \leq x \leq 1$$

What must be the capacity of its tank in order that the probability that its supply will be exhausted in a given day shall be 0.01.

Unit - 6

Expectation For Discrete And Continuous Random Variable

Structure

6.1 Introduction

The mathematical expectation, or expected value, of a discrete random variable X is denoted as E(X) or μ. It represents one average or long-term average value of the random variable, and is calculated as the sum of the products of each possible value of the random variable and its corresponding probability. For continuous random variable, it is defined as the integral of the product of the random variable and its probability density function (p.d.f) over the entire range of the random variable. The expected value of a continuous random variable represent the 'average' or 'central' value of the distribution, and it is a measure of the location or central tendency of the distribution. In gambling and casino games, expected value is used to axalyse the long-term expected payoff of different bets or strategies, which can inform decision-making and risk management. In lottery and other probability-based games, expected value is used to evaluate the fairness of the game and the potential return to players. In project management, expected value is used to evaluate the potential outcomes and is used to evaluate the potential outcomes and profitability of different projects or investments. In quality control, expected value is used to assess the potential impact of defects or failures and to make informed decisions about quality assurance measures.

6.2 Learning Objectives

After studying this unit, you should be able to:

- Define mathematical expectation for discrete random variable
- Define mathematical expectation for continuous random variable
- Define expectation of function of a random variable

6.3 Mathematical Expectation

Gambling was the origin of word expectation. Originally, expectation was defined as:

If p represents a person's chance of success in any venture and A the amount which he will receive in case of a success, then the money equivalent to pA is called expectation.

Mathematical Expectation for Discrete Random Variable

Let X be a random variable with p.d.f. f(x). Then its mathematical expectation, denoted by E(x), is given by

$$E(X) = \sum_{-\infty}^{\infty} x\, f(x)$$

Mathematical Expectation for Continuous Random Variable

Let X be a random variable with p.d.f. f(x). Then its mathematical expectation, denoted by E(X), is given by

$$E(X) = \int_{-\infty}^{\infty} x\, f(x)\, dx$$

Note: Here it is understood that $\int_{-\infty}^{\infty} x\, f(x)\, dx$ exists and $\sum_{-\infty}^{\infty} x\, f(x)$ is absolutely convergent, otherwise these definition are not valid.

In simple words if random variable X takes the value $x_1, x_2, \ldots\ldots, x_n$ with corresponding probabilities $p_1, p_2, \ldots\ldots\ldots, p_n$, then

$$E(X) = p_1x_1 + p_2x_2 + \ldots\ldots + p_nx_n$$

$$= \sum_{i=1}^{n} p_i x_i$$

Note: E(X) is also called the mean of X or the population mean and is denoted by μ.

Expectation of Function of a Random Variable

Let X be a random variable with p.d.f. f(x) and distribution function F(x). If g is a function such that g(X) is a random variable and E[g(X)] is defined, then

$$E[g(X)] = \int_{-\infty}^{\infty} g(x)\, f(x) dx \qquad \text{(for continuous random variable)}$$

$$= \sum_{-\infty}^{\infty} g(x) f(x) \qquad \text{(for discrete random variable)}$$

Let us improve our understanding of these results by looking at some of the following examples:-

Example 1: Find the expectation of the number on a die when thrown.

Sol. Let X be the random variable representing the number on a die when thrown. Then X can take any one of the values 1,2,3,4,5,6 each with equal probability $\frac{1}{6}$ as shown in the table:-

X = x	1	2	3	4	5	6
p	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

$$E(X) = (1)\left(\frac{1}{6}\right) + (2)\left(\frac{1}{6}\right) + (3)\left(\frac{1}{6}\right) + (4)\left(\frac{1}{6}\right) + (5)\left(\frac{1}{6}\right) + (6)\left(\frac{1}{6}\right)$$

$$= \frac{1}{6}[1 + 2 + 3 + 4 + 5 + 6]$$

$$= \frac{1}{6} \times 21 = \frac{7}{2}$$

Example 2: In four tosses of a coin, Let X be the number of heads. Tabulate the 16 possible outcomes with the corresponding values of X. By simple counting derive the distribution of X and hence calculate the expected value of X.

Sol. Let H represent a head, T a tail and X, then random variable denoting the number of heads.

S. No.	**Outcomes**	**No. of Heads (X)**	**S.No.**	**Outcomes**	**No. of Heads (X)**
1	HHHH	4	9	HTHT	2
2	HHHT	3	10	THTH	2
3	HHTH	3	11	THHT	2
4	HTHH	3	12	HTTT	1
5	THHH	3	13	THTT	1
6	HHTT	2	14	TTHT	1
7	HTTH	2	15	TTTH	1
8	TTHH	2	16	TTTT	0

The random variable X takes the values 0,1,2,3 and 4. Since, from the above table, we find that the number of cases favourable to the coming of 0,1,2,3 and 4 heads are 1,4,6,4 and 1, respectively. Then, we have

$$P(X = 0) = \frac{1}{16},\ P(X = 1) = \frac{4}{16} = \frac{1}{4},\ P(X = 2) = \frac{6}{16} = \frac{3}{8},$$

$$P(X = 3) = \frac{4}{16} = \frac{1}{4} \text{ and } P(X = 4) = \frac{1}{16}$$

Thus, the probability distribution of X is

x	0	1	2	3	4
p(x)	$\frac{1}{16}$	$\frac{1}{4}$	$\frac{3}{8}$	$\frac{1}{4}$	$\frac{1}{16}$

$$E(X) = \sum_{x=0}^{4} x\,p(x) = 1.\frac{1}{4} + 2.\frac{3}{8} + 3.\frac{1}{4} + 4.\frac{1}{16}$$

$$= \frac{1}{4} + \frac{3}{4} + \frac{3}{4} + \frac{1}{4} = 2$$

Example 3: A coin is tossed until a head appears. What is the expectation of the number of tosses required?

Sol. Let X denote the number of tosses required to get the first head. The probability distribution of X is

Event	x	Probability p(x)
H	1	$\frac{1}{2}$
TH	2	$\frac{1}{2}\times\frac{1}{2}=\frac{1}{4}$
TTH	3	$\frac{1}{2}\times\frac{1}{2}\times\frac{1}{2}=\frac{1}{8}$
......	......	......
......	......	......
......	......	......

$$\therefore \qquad E(X) = \sum_{x=1}^{\infty} x\,p(x)$$

$= 1 \times \frac{1}{2} + 2 \times \frac{1}{4} + 3 \times \frac{1}{8} + 4 \times \frac{1}{16} + ...$ .....(1)

This is an arithmetic-geometric series with ratio of GP being r = $\frac{1}{2}$

Let $S = 1.\ \frac{1}{2} + 2.\ \frac{1}{4} + 3.\ \frac{1}{8} + 4.\ \frac{1}{16} + ...$ ....(2)

$\therefore \quad \frac{1}{2} S = \frac{1}{4} + 2.\ \frac{1}{8} + 3.\ \frac{1}{16} + ...$ ....(3)

Subtracting (3) from (2), we get

$$\left(1-\frac{1}{2}\right) S = \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + ...$$

$$\Rightarrow \quad \frac{1}{2} S = \frac{1/2}{1-\left(\frac{1}{2}\right)} = 1$$

[Since the sum of an infinite G.P. with first term a and common ratio r (<) is $\frac{a}{1-r}$].

$\Rightarrow \quad S = 2$

$\therefore$ from (1), we get

E(X) = 2

Example 4: A and B throw with one die for a prize of Rs. 11 which is to be won by player who first throws 6. If A has the first throw, what are their respective expectation?

Sol. The chances of throwing a six are as follows:-

A	B
$\frac{1}{6}$	$\frac{5}{6} \times \frac{1}{6}$
$\left(\frac{5}{6}\right)^2 \times \left(\frac{1}{6}\right)$	$\left(\frac{5}{6}\right)^3 \times \frac{1}{6}$
......	......
......	......
......	......

A's chances of success = $\frac{1}{6} + \left(\frac{5}{6}\right)^2 . \frac{1}{6} + \left(\frac{5}{6}\right)^4 . \frac{1}{6} + \ldots\ldots\ldots\ldots\ldots\ldots$

$= \frac{1}{6}\left[1+\left(\frac{5}{6}\right)^2+\left(\frac{5}{6}\right)^4+\ldots\ldots\ldots\infty\right]$

$= \frac{1}{6} . \frac{1}{1-\left(\frac{5}{6}\right)^2} = \frac{6}{11}$

B's chances of success = 1 - $\frac{6}{11} = \frac{5}{11}$

A's expectation = $\frac{6}{11} \times 11$ = Rs. 6

B's expectation = $\frac{5}{11} \times 11$ = Rs. 5

Example 5: A person draws cards one by one from a pack until he draws all the aces. How many cards he may be expected to draw?

Sol. Suppose he has to make n draws for all the aces. It means that in n-1 draws, he draws three aces and in the nth one ace. The probability of such an occurrence.

$$= \frac{4_{c_3} \times 48_{c_{n-4}}}{52_{c_{n-1}}} \times \frac{1}{52-(n-1)}$$

$$= \frac{4 \times \lfloor 48 \times \lfloor n-1 \times \lfloor 52-n+1}{\lfloor n-4 \; \lfloor 48-n+4 \; \lfloor 52} \times \frac{1}{52-n+1}$$

$$= \frac{4(n-1)(n-2)(n-3)}{49 \times 50 \times 51 \times 52}$$

The least number of draws he has to make is 4 and the maximum number 52. Hence n ranges from 4 to 52.

The expected number of draws

$$= \sum_{n=4}^{52}\left[n.4\frac{(n-1)(n-2)(n-3)}{49 \times 50 \times 51 \times 52}\right]$$

$$= \frac{4}{49 \times 50 \times 51 \times 52}\left[\sum_{n=4}^{52} n^4 - 6\sum_{4}^{52} n^3 - 11\sum_{4}^{52} n^2 - 6\sum_{4}^{52} n\right]$$

Self-Check Exercise

Q.1 Two unbiased dice are thrown. Find the expected values of the sum of numbers of points on them.

Q.2 An urn contains 3 black marbles and 2 white marbles. Four persons A,B,C,D in order, draw one marble without replacement. The first to draw white marble gets Rs. 1. compute their expectations.

6.4 Summary

We conclude this unit by summarizing what we have covered in it:-

1. Defined mathematical expectation for discrete random variable
2. Defined and discussed expectation for continuous random variable
3. Defined expectation of function of a random variable

6.5 Glossary:

1. Let X be a random variable with p.d.f. f(x). Then its mathematical expectation denoted by E(X), is given by

$$E(X) = \begin{bmatrix} \sum_{-\infty}^{\infty} x\, f(x), \ \textit{for discrete random variable} \\ \int_{-\infty}^{\infty} x\, f(x)\, dx, \ \textit{for continuous random variable} \end{bmatrix}$$

2. Expectation of function of a random variable is defined as:

   Let X be a random variable with p.d.f. f(x) and distribution function F(x). If g is a function such that g(X) is a random variable and E[g(X)] is defined, then.

$$E[g(X)] = \begin{bmatrix} \sum_{-\infty}^{\infty} g(x)\, f(x), \ \textit{for discrete random variable} \\ \int_{-\infty}^{\infty} g(x)\, f(x)\, dx, \ \textit{for continuous random variable} \end{bmatrix}$$

6.6 Answers To Self-Check Exercise

Ans.1 7

Ans.2 A's expectation : Rs. 4

B's expectation : Rs. 3

C's expectation : Rs. 2

D's expectation : Rs. 1

6.7 References/Suggested Readings

1. Robert V. Hogg, Joseph w. Mckean and Allen T. Craig, Introduction to Mathematical Statistics, Pearson Education, Asia, 2007.

2. Irwin Miller, Marylees Miller and John E. Freund, Mathematical Statistics with Application, 7th Ed., Pearson Education, Asia, 2006.

3. Sheldon Ross, Introduction to Probability Model, 9th Ed., Academic Press, Indian Reprint, 2007.

6.8 Terminal Questions

1. If X has the uniform density

$$f(x) = \begin{cases} \frac{1}{2}, & for\ 2 < x < 4 \\ 0, & elsewhere \end{cases}$$

find its mathematical expectation.

2. Show that the expected number of failures preceding the first success in a series of Bernoullian trials with a constant probability 'p' of success is $\frac{1-p}{p}$.

3. A random variable X takes values

$$x_i = \frac{(-1)^i - 2^i}{i},$$

with probability $p_i = 2^{-i}$, i = 1, 2, 3, ...

Find the expected value of X.

Unit - 7

Addition And Multiplication Theorems Of Expectation

Structure

7.1 Introduction

The addition theorem of expectations states that for any two random variables X and Y, the expectation of their sum is equal to the sum of their individual expectations. This theorem holds true regardless of whether the random variables X and Y are independent or dependent. The multiplication theorem of expectations states that for any two random variables X and Y, the expectation of their product is equal to the product of their individual expectations, but only if the random variables are independent. In manufacturing and engineering, the addition theorem is used to calculate the expected failure rate of a system composed of multiple components. In finance, the addition theorem is used to calculate the expected return of a portfolio of assets. The multiplication theorem is used in various statistical tests, such as t-test determine the significance of the relationship between variables. In the insurance industry, the addition theorem is used to calculate the expected total claims for a portfolio of insurance policies.

7.2 Learning Objectives

After studying this unit, you should be able to:

- Prove addition theorem of expectation
- Prove multiplication theorem of expectation
- Define variance and standard deviation of a random variable
- Prove some standard results on the mathematical expectation, variance.

7.3 Addition Theorem of Expectation

If X and Y be random variables, then

E(X + Y) = E(X) + E(Y)

Proof: Let X take the values $x_1,x_2,....,x_n$ with respective probabilities $p_1,p_2,......,p_n$ and Y takes the values $y_1,y_2,.....,y_n$ with probabilities $p_1',p_2',.....,p_m'$. Then clearly X+Y is also a random variable which takes the nm values $x_i + y_j$, i = 1,2,....., n; j = 1,2,....., m. Also let p_{ij} be probability of X takes the value x_i and Y taking the value y_j simultaneously.

When X takes the value x_i, Y can take any one of the values $y_1,y_2,....,y_m$, therefore, the sum $\sum_{j=1}^{m} p_{ij}$ will represent the probability p_i of X taking the value x_i i.e. $\sum_{i=1}^{n} p_{ij} = p_i$. Giving the similar arguments, $\sum_{i=1}^{n} p_{ij}$ represents the probability p_j' of Y taking the value y_j i.e. $\sum_{i=1}^{n} p_{ij} = p_j$

Now, we have

$$E(X + Y) = \sum_{i=1}^{n} \sum_{j=1}^{m} p_{ij}(x_i + y_j)$$

$$= \sum_{j=1}^{m} p_{ij}x_i + \sum_{i=1}^{n} \sum_{j=1}^{m} p_{ij}y_j$$

$$= \sum_{i=1}^{n} \left(\sum_{j=1}^{m} p_{ij}\right)x_i + \sum_{j=1}^{m} \left(\sum_{i=1}^{n} p_{ij}\right)y_j$$

$$= \sum_{i=1}^{n} p_{ij}x_i + \sum_{j=1}^{m} p_i y_j$$

$$= E(X) + P(Y)$$

Note. 1. If X, Y, Z ...... be discrete random variables, then

E(X + Y + Z + ...) = E(X) + E(Y) + E(Z) +........

2. If X and Y be random variables, then E(aX + bY) = a E(X) + B E(Y), where a and b are constant.

7.4 Multiplication Theorem of Expectation

If X and Y be two independent random variables, then

E(XY) = E(X) E(Y)

Proof: Let X take the values x_1, x_2, ... x_n with respective probabilities p_1, p_2,...., p_n and Y take the values y_1, y_2,....., y_m with probabilities p_1, p_2,.......,p_m.

Since the variables X and Y are independent, therefore, the probability that X takes the values x_i and that the variable Y takes the value y_j simultaneously is p_i p_j'.

Thus, E(XY) = $\sum_{i=1}^{n} \sum_{j=1}^{n} p_i p_j x_i y_j$

= $\left(\sum_{i=1}^{n} p_i x_i \right) \left(\sum_{j=1}^{m} p_j ' y_j \right)$

= E(X) E(Y)

Note:- The above theorem may be extended to any number of independent variables i.e. if X,Y,Z,....... be independent random variables, then

E (X Y Z ....) = E(X) E(Y) E(Z)............

7.5 Variance, Standard Deviation

Variance of random variable X is the expected value of the non-negative random variable $(X - \mu)^2$.

$\therefore$ $V(X) = E(X - \mu)^2$

or Var (X) = $\sum p_i (x_i - \overline{X})^2$

The positive square root of the variance of X is called the standard deviation of X and is denoted by σ.

$\therefore$ Var (X) = $\sigma^2 = E(X = \mu)^2 = \sum p_i (xi - \overline{X})^2$

7.6 Some Standard Results

Result 1:- The mathematical expectation of the sum of n random variables is equal to the sum of their expectations, provided all the expectations exist.

Proof: We have to prove that

$E(X_1 + X_2 + ... + X_n) = E(X_1) + E(X_2) + ... + E(X_n)$ ......(1)

For the random variables X1, X2

E(X1 + X2) = E(X1) + E(X2)

$\therefore$ Result (1) is true for n = 2.

Assume that result (1) is true for n = m

$\therefore$ E $\left(\sum_{i=1}^{m} X_i \right) = \sum_{i=1}^{m} E(X_i)$ ....(2)

Now E$\left(\sum_{i=1}^{m+1} X_i \right)$ = E$\left(\sum_{i=1}^{m} X_i + X_{m+1} \right)$ = E$\left(\sum_{i=1}^{m} X_i \right)$ + $E(X_{m+1})$

$= \sum_{i=1}^{m} E(X_i)$ + $E(X_{m+1})$ [$\because$ of (2)]

$= \left(\sum_{i=1}^{m+1} E(X_i) \right)$

$\therefore$ Result (1) is true for n = m + 1.

$\therefore$ If the result (1) is true for n = m, then it is also true for n = m + 1.

i.e. if the result (1) is true for any integer, then it is also true for the next higher integer. But result (1) is true for n = 2

$\therefore$ By the method of induction, result (1) is true for all $n \in I$.

Result 2: If X is a random variable and a is constant, then

(i) $E[a\,\psi(X)] = a\,E[\psi(X)]$

(ii) $E[\psi(X) + a] = E[\psi(X)] + a$

where $\psi(X)$, a function of X, is a random variable and all the expectations exist.

Proof: (i) $E[a\,\psi(x)] = \int_{-\infty}^{\infty} a\psi(x) f(x)dx$

$= a\int_{-\infty}^{\infty} \psi(x) f(x)dx = aE[\psi(X)]$

(ii) $E[\psi(X) + a] = \int_{-\infty}^{\infty} \left[\psi(x)+a\right] f(x)dx$

$= \int_{-\infty}^{\infty} \psi(x) f(x)dx + a\int_{-\infty}^{\infty} f(x)dx$

$= E[\psi(X)] + a$ $\left[\because \int_{-\infty}^{\infty} f(x)dx = 1\right]$

Result 3: If X is a random variable and a and b are constants, then

$E[aX + b] = a\,E(X) + b$

Proof: By definition, we have

$E[aX + b] = \int_{-\infty}^{\infty} (ax+b) f(x)dx$

$$= a\int_{-\infty}^{\infty} xf(x)dx + b\int_{-\infty}^{\infty} f(x)dx$$

$$= a\,E(X) + b$$

Result 4: With usual notations

$$\sigma^2 = \sum_{i=1}^{n} p_i x_i^2 - \left(\sum_{i=1}^{n} p_i x_i\right)^2$$

Proof: $\sigma^2 = \sum_{i=1}^{n} p_i (x_i - \overline{X})^2$

$$= \sum_{i=1}^{n} p_i (x_i^2 - 2x_i\overline{X} + \overline{X}^2)$$

$$= \sum_{i=1}^{n} p_i x_i^2 - 2\overline{X}\sum_{i=1}^{n} p_i x_i + \overline{X}^2 \sum_{i=1}^{n} p_i$$

$$= \sum_{i=1}^{n} p_i x_i^2 - 2\overline{X}^2 + \overline{X}^2 \qquad \left[\because \overline{X} = \sum_{i=1}^{n} p_i x_i \text{ and } \sum_{i=1}^{n} p_i = 1\right]$$

$$= \sum_{i=1}^{n} p_i x_i^2 - \overline{X}^2$$

$$= \sum_{i=1}^{n} p_i x_i^2 - \left(\sum_{i=1}^{n} p_i x_i\right)^2$$

Result 5: For any random variable X,

$$\text{Var}(X) = E(X^2) - [E(X)]^2$$

Proof: $\text{Var}(X) = E(X - \mu)^2$

$$= E(X^2 - 2\mu X + \mu^2)$$

$$= E(X^2) - 2\mu E(X) + \mu^2 E(1)$$

$$= E(X^2) - 2\mu.\mu + \mu^2.1 \qquad [\because E(X) = \mu, E(1) = 1]$$

$$= E(X^2) - 2\mu^2 + \mu^2$$

$$= E(X^2) - \mu^2$$

$$= E(X^2) - [E(X)]^2 \qquad [\because \mu = E(X)]$$

Hence Var (X) = $E(X^2) - [E(X)]^2$

Result 6: For any constants a and b,

Var (a X + b) = a^2 Var (X)

Proof: We have

Var (a X + b) = E $[a X + b]^2 - [E (a X + b)]^2$

$= E(a^2X^2 + 2abX + b^2) - [a E (X) + b]^2$ [$\because$ E(1) = 1]

$= a^2E (X^2) + 2abE (X) + b^2 - a^2 [E(X)]^2 - b^2 - 1 - 2ab E (X)$

$= a^2 \{E(X^2) - [E(X)]^2\}$

$= a^2$ Var (X)

Hence the result

Result 7: If X_1, X_2 are two independent random variables having $E(X_1) = \mu_1$ and $E(X_2) = \mu_2$, then Var $(X_1 + X_2)$ = Var (X_1) + Var (X_2)

Proof: Var $(X_1 + X_2) = E(X_1 + X_2)^2 - [E(X_1 + X_2)]^2$

$= E(X_1^2) + E(X_2^2) + 2E(X_1 X_2) - \{[E(X_1(]^2 + [E(X_2)]^2 + 2E(X_1) E(X_2)\}$

$= E(X_1^2) + E(X_2^2) + 2E(X_1) E(X_2) - [E(X_1)]^2 - [E(X_2)]^2 - 2E(X_1) E (X_2)$

$= \{E(X_1^2) - [E(X_1)]^2\} + \{E(X_2^2) - [E(X_2)]^2$

= Var (X_1) + Var (X_2)

Hence Var $(X_1 + X_2)$ = Var (X_1) + Var (x_2)

Let us improve our understands of these results by looking at some of the following examples:-

Example 1: For the following probability distribution of the random variable X

X	8	12	16	20	24
p(X)	$\frac{1}{8}$	$\frac{1}{6}$	$\frac{3}{8}$	$\frac{1}{4}$	$\frac{1}{12}$

Find (i) E (X) (ii) E (X^2) (iii) E $(X - \overline{X})^2$

Proof: (i) $E(X) = \sum x_i p_i = 8 \times \frac{1}{8} + 12 \times \frac{1}{6} + 16 \times \frac{3}{8} + 20 \times \frac{1}{4} + 24 \times \frac{1}{12}$

= 16

(ii) $E(X^2) = \sum x_i^2 p_i = 8^2 \times \frac{1}{8} + 12^2 \times \frac{1}{6} + 16^2 \times \frac{3}{8} + 20^2 \times \frac{1}{4} + 24^2 \times \frac{1}{12}$

$= 276$

$V(X) = E(X^2) - [E(X)]^2$

$= 276 - (16)^2 = 20$

$\therefore$ Mean = E(X) = 16, Variance = V(X) = 20

(iii) $E(x - \overline{x})^2 = E(x^2 + \overline{x}^2 - 2x\overline{x})$

$= E(x^2) + \overline{x}^2 - 2E\,(x\overline{x})$

$= E(x^2) + \overline{x}^2 - 2x\overline{x}$

$= E(x^2) + \overline{x}^2$

$= E(x^2) + [E(x)]^2$

$= 276 - (16)^2 = 276 - 256$

$= 20$

Example 2: A s/x - sided die is tossed. Find the variance of the number of dots on the top face.

Sol. : Here s = {1, 2, 3, 4, 5, 6}

Let x be random variable denoting the number of points taking values 1, 2, 3, 4, 5, 6.

X	1	2	3	4	5	6
$f(x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

$E(x) = \sum x\, f(x)$

$= 1\left(\frac{1}{6}\right) + 2\left(\frac{1}{6}\right) + 3\left(\frac{1}{6}\right) + 4\left(\frac{1}{6}\right) + 5\left(\frac{1}{6}\right) + 6\left(\frac{1}{6}\right)$

$= \frac{1}{6} + \frac{2}{6} + \frac{3}{6} + \frac{4}{6} + \frac{5}{6} + \frac{6}{6}$

$= \frac{1+2+3+4+5+6}{6} = \frac{21}{6} = \frac{7}{2}$

$E(x^2) = \sum x^2\, f(x) = 1\left(\frac{1}{6}\right) + 4\left(\frac{1}{6}\right) + 9\left(\frac{1}{6}\right) + 16\left(\frac{1}{6}\right) + 25\left(\frac{1}{6}\right) + 36\left(\frac{1}{6}\right)$

$= \frac{1}{6} + \frac{4}{6} + \frac{9}{6} + \frac{16}{6} + \frac{25}{6} + \frac{36}{6}$

$$= \frac{1+4+9+16+25+36}{6} = \frac{91}{6}$$

Var (x) = $E(x^2) - [E(x)]^2$

$$= \frac{91}{6} - \frac{49}{4} = \frac{182-147}{12} = \frac{35}{12}$$

Example 3 : A coin is tossed until a tail appears. What is the expectation of the number of tosses required?

Sol. :- Let denote the number of tosses required to get the first tail. The probability distribution of x is

Event	x	Probability
H	1	$\frac{1}{2}$
TH	2	$\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$
TTH	3	$\frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{8}$
......	......	......
......	......	......
......	......	......

$\therefore \quad E(X) = \sum_{x=1}^{\infty} x\, p(x)$

$$= 1 \times \frac{1}{2} + 2 \times \frac{1}{4} + 3 \times \frac{1}{8} + 4 \times \frac{1}{16} + \text{................} \qquad \text{.....(1)}$$

This is an arithmetic-geometric series with ratio of G.P. being r = $\frac{1}{2}$

Let $\quad S = 1.\frac{1}{2} + 2.\frac{1}{4} + 3.\frac{1}{8} + 4.\frac{1}{16} + \text{..........} \qquad \text{.......(2)}$

$\therefore \quad \frac{1}{2}S = \frac{1}{4} + 2.\frac{1}{8} + 3.\frac{1}{16} + \text{...........} \qquad \text{.......(3)}$

Subtracting (3) from (2), we get

$$\left(1-\frac{1}{2}\right)S = \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \text{...........}$$

$$\Rightarrow \qquad \frac{1}{2} S = \frac{\frac{1}{2}}{1-\left(\frac{1}{2}\right)} = 1$$

[Since the sum of an infinite G.P. with first term a and common ratio r (<) is $\frac{a}{1-r}$].

$\Rightarrow$ S = 2

$\therefore$ from (1), we get E(x) = 2

Example 4 :- Three urns contain respectively 3 green and 2 white ball, 5 green and 6 white balls and 2 green and 4 white balls. One ball is drawn from each urn. Find the expected number of white balls drawn out?

Sol. :-

	Number of green balls	Number of white balls
Urn I	3	2
Urn II	5	6
Urn III	2	4

E (one white ball drawn from each urn)

$$= 1 \times \frac{2}{5} + 1 \times \frac{6}{11} + 1 \times \frac{4}{6}$$

$$= \frac{2}{5} + \frac{6}{11} + \frac{2}{3}$$

$$= \frac{66+90+110}{165} = \frac{266}{165}$$

Example 5 : Let x have the p.d.f.

$$f(x) = \begin{cases} \frac{1}{2}(x+1), & -1 < x < 1 \\ 0, & \text{elsewhere} \end{cases}$$

Find the mean and variance of x

Sol. : - Here f(x) = $\begin{cases} \frac{1}{2}(x+1), & -1 < x < 1 \\ 0, & \text{elsewhere} \end{cases}$

$\therefore$ E(x) = $\int_{-\infty}^{\infty} xf(x)dx$ = $\int_{-1}^{1} xf(x)dx$ = $\int_{-1}^{1} \frac{x(x+1)}{2} dx$

= $\frac{1}{2} \int_{-1}^{1} (x^2 + x)dx$ = $\frac{1}{2} \left[\frac{x^3}{3} + \frac{x^2}{2} \right]_{-1}^{1}$

= $\frac{1}{2} \left[\left(\frac{1}{3} + \frac{1}{2} \right) - \left(-\frac{1}{3} + \frac{1}{2} \right) \right]$ = $\frac{1}{2} \times \frac{2}{3}$ = $\frac{1}{3}$

$\therefore$ Mean = $\frac{1}{3}$

Now E(x^2) = $\int_{-1}^{1} x^2 f(x)dx$ = $\frac{1}{2} \int_{-1}^{1} x^2(x^2+1)dx$

= $\frac{1}{2} \int_{-1}^{1} (x^3 + x^2)dx$

= $\frac{1}{2} \left[\frac{x^4}{4} + \frac{x^3}{3} \right]_{-1}^{1}$

= $\frac{1}{2} \left[\left(\frac{1}{4} + \frac{1}{3} \right) - \left(\frac{1}{4} - \frac{1}{3} \right) \right]$

= $\frac{1}{2} \times \frac{2}{3}$ = $\frac{1}{3}$

Variance x = E(x^2) - [E(x)]^2

= $\frac{1}{3}$ - $\left(\frac{1}{3} \right)^2$ = $\frac{1}{3}$ - $\frac{1}{9}$ = $\frac{3-1}{9}$ = $\frac{2}{9}$

Example 6 :- Find the mean and variance of the distribution that has the distribution function

$$F(x) = \begin{cases} 0, & x < 0 \\ \frac{x}{8}, & 0 \le x < 2 \\ \frac{x^2}{16}, & 2 \le x < 4 \\ 1, & 4 \le x \end{cases}$$

Sol. :- Here

$$F(x) = \begin{cases} 0, & x < 0 \\ \frac{x}{8}, & 0 \le x < 2 \\ \frac{x^2}{16}, & 2 \le x < 4 \\ 1, & 4 \le x \end{cases}$$

which is continuous function of x

$\therefore$ The p.d.f. of x is given by $f(x) = F'(x)$

$$\therefore \quad f(x) = \begin{cases} \frac{1}{8}, & 0 \le x < 2 \\ \frac{x}{8}, & 2 \le x < 4 \\ 0, & \text{elsewhere} \end{cases}$$

Now $\therefore$ $E(x) = \int_{-\infty}^{\infty} xf(x)dx = \int_0^2 xf(x)dx + \int_2^4 xf(x)dx$

$$= \int_0^2 \frac{x}{8} dx + \int_2^4 \frac{x^2}{8} dx$$

$$= \frac{1}{8} \int_0^2 xdx + \frac{1}{8} \int_2^4 x^2 dx$$

$$= \frac{1}{16} \left[x^2\right]_0^2 + \frac{1}{24} \left[x^3\right]_2^4$$

$$= \frac{1}{16} [4 - 0] + \frac{1}{24} [64 - 8]$$

$$= \frac{4}{16} + \frac{56}{24} = \frac{1}{4} + \frac{7}{3} = \frac{31}{12}$$

and E(x^2) = $\int_{-\infty}^{\infty} x^2 f(x)dx$ = $\int_0^2 x^2 \frac{1}{8}dx$ + $\int_2^4 x^2 \frac{x}{8}dx$

$$= \frac{1}{8} \int_0^2 x^2 dx + \frac{1}{8}\int_2^4 x^3 dx$$

$$= \frac{1}{24} \left[x^3\right]_0^2 + \frac{1}{32}\left[x^4\right]_2^4$$

$$= \frac{1}{24} [8 - 0] + \frac{1}{32}[256 - 16] = \frac{8}{24} + \frac{240}{32}$$

$$= \frac{1}{3} + \frac{15}{2} = \frac{2+45}{6} = \frac{47}{6}$$

Now variance = E(x^2) - [E(x)]$_2$

$$= \frac{47}{6} - \left(\frac{31}{12}\right)^2 = \frac{47}{6} - \frac{961}{144}$$

$$= \frac{1128-961}{144} = \frac{167}{144}$$

Example 7 :- Find the mean and variance of a random variable that takes the values 1, 2, 3,....,n each with probability $\frac{1}{n}$.

Sol. :- Here

x	1	2	3	.....	n
P(x)	$\frac{1}{n}$	$\frac{1}{n}$	$\frac{1}{n}$	.....	$\frac{1}{n}$

E(x) = $\sum$ x f(x)

$$= 1\left(\frac{1}{n}\right) + 2\left(\frac{1}{n}\right) + 3\left(\frac{1}{n}\right) + \ldots\ldots x\left(\frac{1}{n}\right)$$

$$= \frac{1}{n} (1 + 2 + 3 + \ldots\ldots. + n)$$

$$= \frac{1}{n}\left[\frac{n(n+1)}{2}\right] = \frac{n+1}{2}$$

$\therefore$ Mean = E(x) = $\frac{n+1}{2}$

$E(x^2) = \sum x^2 f(x)$

$$= 1^2\left(\frac{1}{n}\right) + 2^2\left(\frac{1}{n}\right) + 3\left(\frac{1}{n}\right) + \ldots\ldots x^2\left(\frac{1}{n}\right)$$

$$= \frac{1}{n}\ (1^2 + 2^2 + \ldots\ldots. + n^2)$$

$$= \frac{1}{n}\left[\frac{n(n+1)(2n+1)}{6}\right]$$

$$= \frac{(n+1)(2n+1)}{6}$$

$\therefore$ Var (x) = $E(x^2) = [E(x)]^2$

$$= \frac{(n+1)(2n+1)}{6} - \left[\frac{n+1}{2}\right]^2$$

$$= \left(\frac{n+1}{2}\right)\left[\frac{2n+1}{3} - \frac{n+1}{2}\right]$$

$$= \left(\frac{n+1}{2}\right)\left[\frac{4n+2-(3n+3)}{6}\right]$$

$$= \left(\frac{n+1}{2}\right)\left[\frac{4n+2-3n-3}{6}\right]$$

$$= \left(\frac{n+1}{2}\right)\frac{n-1}{6} = \frac{(n+1)(n-1)}{12}$$

$$= \frac{n^2-1}{12}$$

Example 8 :- If x is random variable such that E(x) = 10, v(x) = 25, find the positive numbers a and b such that y = a x -b has mean zero and variance 1.

Sol. :- E(y) = E(x - b) = a E(x) - b = 10a - b = 0

$\Rightarrow$ 10 a = b ....(1)

Also $V(y) = a^2\, v(x) = a^2 \times 25 = 1$

$\Rightarrow$ $a^2 = \frac{1}{25}$ $\Rightarrow a = \frac{1}{5}$ ....(2)

Putting the value of a from (2) in (1), we have

b = 2

$\therefore \quad a = \frac{1}{5}$, b = 2

Example 9 : Show that if y and z are independent random values of a random variable x, then

$E(y - z)^2 = 2\ v(x)$

Sol. :- Let y and z are independent random values of variable x.

$\therefore \quad E(yz) = E(y)\ E(z)$

and $\quad E[y - z]^2 = E(y^2) + E(z^2) - 2E\ (yz)$

[$\therefore$ y and z are independent]

$= E(y^2) + E(z^2) - 2E\ (y)\ E(z)$

$= E(x^2) + E(x^2) - 2E\ (x)\ E(x)$

[$\therefore$ $E(x) = E(y) = E(z) = \mu$ (say)

and $E(x^2) = E(y^2) = E(z^2)$)

$=2[E(x^2) - [E(x)]^2]$

$= 2\ v(x)$

$\therefore \quad E(y - z)^2 = 2\ v(x)$

Self-check exercise

Q. 1 If x_1 and x_2 are two independent random variables having variance k and 2 respectively. If the variance of $y = 3x_2 - x1$ is 25. Find k.

Q.2 Let y = 3x - 5 and E(x) = 4, var (x) = 2. What is the mean and variance of y?

Q.3 Let the probability function of the random variable x be of the following form, where c is some constant: $f(x) = c \binom{5}{x}$, x = 0,1,2,3,4,5

(i) Determine the value of c.

(ii) Find E(x)

(iii) Find var (x)

Q.4 A and B thrown with one die for a stake of Rs. 44 which is to be won by the player who first throws a six. If A has the first throw, what are their respective expectations?

7.7 Summary

We conclude this unit by summarizing what we have covered in it:-

1. Proved addition theorem of expectation of two random variables.
2. Proved multiplication theorem of expectation of two independent random variables.

3. Defined variance and standard deviation of a random variable.

4. Proved some standard results of expectation and variance.

5. Some examples are given related to each topic so that the contents be clarified further.

7.8 Glossary:

1. Variance of a random variable x is the expected value of the non-negative random variable $(x - \mu)^2$ i.e.

   $\text{Var}(x) = E(X - \mu)^2$

   or $\text{Var}(x) = \sum p_i (xi - \overline{X})^2$

   The positive square root of the variance of x is called the standard deviation of x and is denoted by σ i.e.

   $\text{Var}(x) = \sigma^2 = E(x = 4)^2 = \sum p_i (xi - \overline{X})^2$

2. If x and y be random variables, then

   E (x + y) E (x) + e (y)

3. If x and y be two independent random variables, then

   E (xy) = E (x) E (y)

7.9 Answers To Self-Check Exercise

Ans.1 k = 7

Ans.2 Mean = 7 ; variance = 18

Ans.3 (i) $c = \frac{1}{3}$

(ii) $E(x) = \frac{5}{2}$

(iii) $\text{Var}(x) = \frac{5}{4}$

Ans.4 A's expectation = Rs. 24

B's expectation = Rs. 20

7.10 References/Suggested Readings

1. Robert V. Hogg, Joseph w. Mckean and Allen T. Craig, Introduction to Mathematical Statistics, Pearson Education, Asia, 2007.

2. Irwin Miller, Marylees Miller and John E. Freund, Mathematical Statistics with Application, 7th Ed., Pearson Education, Asia, 2006.

3. Sheldon Ross, Introduction to Probability Model, 9th Ed., Academic Press, Indian Reprint, 2007.

7.11 Terminal Questions

1. Let the random variable x have the distribution

   P(x = 0) = P (x = 2) = p ; P (x = 1) = 1 - 2p, for $0 \leq p \leq \frac{1}{2}$

   For what p is the V(x) a maximum.

2. Let the probability function of the random variable x be of the following form where c is some constant:

   f(x) = c x, x = 3,4,5,6

   (i) Determine the value of c

   (ii) Find E (x)

   (iii) Find Var (x)

3. If $f(x) = \begin{cases} x \ , \ 0 \leq x \leq 1 \\ 2 - x, 1 \leq x \leq 2 \end{cases}$

   Find the mean and variance of a random variable x.

4. Let x be a random variable with the following probability distribution

x	-3	6	9
p(x)	$\frac{1}{6}$	$\frac{1}{2}$	$\frac{1}{3}$

   Find E (x) and E (x^2) and using the laws of expectation, evaluate E $(2x + 1)^2$

5. An urn contains 3 black and 2 white marbles. Four persons A,B,C,D in order, draw one marble without replacement. The first to draw a white gets Rs. 10. compute their expectations.

Unit - 8

Moments

Structure

8.1 Introduction

Moment are numerical characteristics that describe the shape and properties of a probability distribution. They provide a quantitative way to analyse and understood the distribution of random variables. The most used moments are: mean (first moment); variance (second moment): skewness (third moment) and kurtosis (fourth moment). The mean or the first moment represents the contra) tendency of the distribution. It is the average or expected value of the random variable. The variance or the second moment, measures the spread or dispersion of the distribution around the mean. Skewness measures the asymmetry or lack of symmetry in the distribution. If indicates the direction and degree of the distribution's departure from symmetry kurtosis measures the 'peakedness' or 'flatness' of the distribution compared to a normal distribution. Positive kurtosis indicated a 'peaked' distribution with heavier tails, while negative kurtosis indicates a 'flat' distribution with lighter tails.

Moments are used to describe and characterize the shape and properties of probability distributions, such as symmetry, dispersion and tail behaviour. Moments are often used to estimate the parameters of probability distributions, such as the mean and standard deviation, from sample data.

8.2 Learning Objectives

After Studying this unit, you should be able to:

- Defined moments about origin and moments about mean for discrete and continuous probability distribution

- Find the relation between the moments about the mean in terms of moments about the origin.
- Find the relation between the moments about the origin in terms of moments about the mean

8.3 Moments

In case of discrete probability distribution, r[th] moment about origin, denoted by μ_r', is defined as

μ_r' = E [x - μ)^r] = $\sum (x-\mu)^r f(x)$

In case of continuous probability distribution,

$$\mu_r' = E(xr) = \int_{-\infty}^{\infty} x^r f(x)dx$$

and $\mu r = E[(x-\mu)^r] = \int_{-\infty}^{\infty} (x=\mu)^r f(x)dx$

Note I.

If r = 1, then $\mu_r' = \sum x^r f(x)$ given

$\mu_1' = E(x) = \sum x f(x)$,

which is the expected value of the random variable x and is denoted by μ.

$\therefore \quad \mu = \mu_1' = E(x) = \sum x f(x)$

Note II.

If r = 1, then $\mu_r = \sum (x-\mu) f(x)$ gives

$\mu_1 = E(x = \mu) = \sum (x-\mu) f(x)$ gives

$= \sum x f(x) - \mu \sum f(x)$

$= \mu - \mu.1 \qquad [\because \sum f(x) = 1]$

$= 0$

$\therefore$ First moment about mean is always zero

Note III.

If r = 2, then

$\mu_r = \sum (x-\mu)^r f(x)$ gives

μ_2 = E (x - μ)^2 = $\sum (x-\mu)^2 f(x) = \sigma^2$

8.4 Relation Between the Moment About the Mean in Terms of Moments About the Origin

We know that

$\mu_r = \sum (x-\mu)^r f(x)$

Put r = 1

$\therefore \quad \mu_1 = \sum (x-\mu) f(x)$

$= \sum x f(x) - \mu \sum f(x)$

= μ - μ- 1 $\quad \left[\because \sum f(x) = 1\right]$

= 0

$\therefore \quad \mu_1 = 0$

μ_2 = E (x - μ)^2 = E (x^2) - 2μ E (x) + μ^2 E (1)

= E (x^2) - 2μ E (x) + μ^2 $\quad$ [$\because$ E (x) =μ and E (1) = 1]

= E (x^2) - 2μ^2 + μ^2

= E (x^2) - μ^2 = μ_2 ' + μ_1^2

μ_3 = E (x - μ)^3 = E (x^3 - 3μ x^2 + 3μ^2 x - μ^3)

= E (x^3) - 3μ E (x^2) + 3μ^2 E (x) - μ^3 E (1)

= E (x^3) - 3μ.μ$_2$ + 3μ^2.μ - μ_1^3 $\quad$ [$\because$ μ= μ_1^1]

= μ_3 '- 3 μ_1 ' μ_2 '+ 2 μ_1^3

μ_4 = E (x -μ)^4

= E (x^4 - 4 x^{3}μ+ 6x^{2}μ^2 - 4xμ^3 +μ^4)

= E (x^4) - 4μE(x^3) + 6μ^2 E (x^2) - 4μ^3 E (x) + μ^4 E (1)

= μ_4^1 - 4μ μ_3^1 + 6μ^2 μ_2^1 - 4μ^3 μ_1^1 + μ^4

= μ_4 '- 4 μ_1 ' μ_3 '+ 6 μ_2 ' μ'^2_1 - 4 μ'^4_1 + μ'^4_1 $\quad$ [$\because$ μ = μ_1^1)

= μ_4^1 - 4 μ_3^1 μ_1^1 + 6 μ_2^1 μ'^2_1 - 3 μ'^4_1

8.5 Relation Between the Moments About the Origin In Terms of Moments About the Mean

We have

μ_1^1 = E (x) = μ

μ_2^1 = E (x^2) = E [(x -μ) +μ]^2

= E[(x =μ)^2 + 2μ (x -μ) + μ^2]

= E [(x -μ)^2] + 2μE (x = μ) + μ^2 E (1)

= μ_2 + 2μ(0) + μ^2 (1) [∵ E(x -μ) = 0, E(1) = 1]

∴ $\mu_2^1 = \mu_2 + \mu^2$

μ_3^1 = E (x^3) = E[(x = μ) + μ]^3

= E [(x -μ)^3 + 3μ(x =μ)^2 + 3μ^2 [x = μ) + μ^3]

= E [(x -μ)^3] + 3μE (x -μ)^2 + 3μ^2 E (x - μ) +μ^3 E (1)

= μ_3 + 3μ μ_2 + 3μ^2 (0) +μ^3 (1)

∴ $\mu_3^1 = \mu_3 + 3\mu\,\mu_2 + \mu^3$

μ_4^1 = E (x^4) = E [(x -μ) +μ]^4

= E [(x - μ)^4 + 4μ(x -μ)^3 + 6μ^2 (x - μ)^2 + 4μ^3 (x -μ) + μ^4]

= E[(x -μ)^4 + 4μE [(x -μ)^3] + 6μ^2 E [(x -μ)^2]

+ 4μ^3 E (x -μ) + μ^4 E (1)

= μ_4 + 4 μ_3 + 6μ^2 μ_2 + 4μ^2 (0) +μ4 (1)

∴ $\mu_4^1 = \mu_4 + 4\mu\,\mu_3 + 6\mu^2\,\mu_2 + \mu^4$

Let us improve our understanding of these results by looking at some of the following examples:-

Example 1: The first four moments of a distribution about the value 4 are -1.5, 17, -30, 108. Calculate the moments about the mean.

Sol. Here the rth moment about any point '4' is

$\mu_r{}' = E\,(x - 4)^r$

Putting r = 1,2,3,4, we obtain

$\mu_1^1 = E\,(x - 4)$

or $\quad -1.5 = E\,(x) - 4$

or $\quad E\,(x) = 4 - 1.5$

or $\quad \mu = 2.5 \qquad [\because \mu = E\,(x)]$

Put r = 2 $\qquad \mu_2^1 = E\,(x - 4)^2$

or $\quad 17 = E\,(x^2 - 8x + 16)$

$= E\,(x^2) - 8\,E\,(x) + 16\,E\,(1)$

$= E\,(x^2) - 8\mu + 16$

$\therefore \quad E\,(x^2) = 17 + 8\mu - 16$

$= 17 + 8\,(2.5) - 16$

$= 17 + 20 - 16 = 21$

Now $\quad \mu 2 = E\,[x^2] - [E(x)]^2$

$= 21 - (2.5)^2$

$= 21 - 6.25$

$= 14.75$

$\therefore \quad \mu_2 = 14.75$

Put r = 3; $\qquad \mu_3{}' = E\,(x - 4)3$

or $\quad -30 = E\,[x^3 - 12x^2 + 48x - 64]$

$= E\,(x^3) - 12\,E\,(x^2) + 48E\,(x) - 64E\,(1)$

$= E\,(x^3) - 12\,(21) + 48\,(2.5) - 64$

or $\quad E\,(x^3) = -30 + 252 - 120 + 64$

$\therefore \quad E\,(x^3) = 166$

Now $\quad \mu_3 = E\,(x^3) - 3E\,(x^2)\,E\,(x) + 2\,[E\,(x)]^2$

$= 166 - 3\,(21)\,(2.5) + 2\,(2.5)^2$

$= 166 - 157.5 + 31.25$

$= 197.25 - 157.50$

= 39.75

Pat r = 4; μ_4^1 = E [x - 4]3

or 108 = E [x^4 - 16x^3 + 96x^2 - 256x + 256]

= E(x^4) - 16E (x^3) + 96E (x^2) - 256 E (x) + 256 E (1)

= E (x^4) - 16 (166) + 96 (21) - 256 (2.5) + 256

or E (x4) = 108 + 2656 - 2016 + 640 - 256

= 1132

Now μ_4 = E (x^4) - 4 E (x^3) E (x) + 6E (x^2) $[E(x)]^2$ - 3 $[E(x)]^4$

= 1132 - 4 (166) (2.5) + 6(21) $(2.5)^2$ - 3 $(2.5)^4$

= 1132 - 1660 + 787.50 - 117.1875

or μ_4 = 142.3125

Example 2: Following are the four moments about the true mean, which is 5.2 :

μ_1 = 0, μ_2 = 5.16, μ_3 = -2.304, μ_4 = 59.8032.

Find the four moments about the origin.

Sol. Here μ= 5.2, μ_1 = 0, μ_2 = 5.16, μ_3 = -2.304,

μ_4 = 59.8032

Now μ_1^1 = μ= 5.2

$\mu_2^1 = \mu_2 + \mu_1^2$ = 5.16 + $(5.2)^2$ =5.16 + 27.04 = 32.2

$\mu_3^1 = \mu_3 + 3\mu\,\mu_2 + \mu^3$ = -2.304 + 3(5.2) (5.16) + $(5.2)^3$

= -2.304 + 80.496 + 140.608 = 218.8

$\mu_4^1 = \mu_4 + 4\mu\,\mu_3 + 6\mu^2 + \mu^4$

= 59.8032 + 4(5.2) (-2.304) + 6 $(5.2)^2$ (5.16) + $(5.2)^4$

= 59.8032 - 47.9232 + 837.1584 + 731.1616

= 1580.2

Example 3: The first four moments of a distribution about arbitrary origin 4 are 1,3.5,8.5,33.5 respectively.

Calculate (i) μ_2, μ_3, μ_4 (ii) the first four moments about zero.

Sol. (i) Here E (x - 4) = 1 $\Rightarrow$ E (x) - 4 = 1

$\Rightarrow$ $E(x) = 1 + 4$

$= 5$

$\therefore$ $\mu = 5$

$E(x - 4)^2 = 3.5$

$\Rightarrow$ $E[x^2 + 16 - 8x] = 3.5$

$\Rightarrow$ $E(x^2) + 16\,E(1) - 8E(x) = 3.5$

$\Rightarrow$ $E(x^2) + 16 - 8(5) = 3.5$

$\Rightarrow$ $E(x^2) = 27.5$

Now $E(x - 4)^3 = 8.5$

or $E[x^3 - 12x^2 + 48x - 64] = 8.5$

or $E(x^3) - 12E(x^2) + 48\,E(x) - 64 = 8.5$

or $E(x^3) - 12(27.5) + 48(5) - 64 = 8.5$

or $E(x^3) - 330 + 240 - 64 = 8.5$

or $E(x^3) = 162.5$

and

$E(x - 4)^4 = 33.5$

$\Rightarrow$ $E[x^4 - 16x^3 + 96x^2 - 256x + 256] = 33.5$

or $E(x^4) - 16\,E(x^3) + 96\,E(x^2) - 256\,E(x) + 256 = 33.5$

or $E(x^4) - 16(162.5) + 96(27.5) - 256(5) + 256 = 33.5$

or $E(x^4) - 2600 + 2640 - 1280 + 256 = 33.5$

$\Rightarrow$ $E(x^4) = 1017.5$

Now $\mu_2 = \mu_2^1 - (\mu_1^1)2 = 27.5 - (5)2 = 27.5 - 25 = 2.5$

$\mu_3 = \mu_3' - 3\mu_2'\mu_1' + 2\mu_1'^3$

$= 162.5 - 3(27.5)(5) + 2(5)^3$

$= 162.5 - 412 + 250$

$= 0$

$\mu_4 = \mu_4' - 4\mu_3'\mu_1' + 6\mu_2'(\mu_1')^2 - 3(\mu_1')^4$

$= 1017.5 - 4(162.5)(5) + 6(27.5)(5)^2 - 3(5)^4$

$= 1017.5 - 3250 + 4125 - 1875$

$= 17.5$

(ii) μ_4 ' - μ= 5, μ_2 ' = E (x^2) = 27.5, μ_3 '= E (x^3) = 162 - 5,

μ_4 ' = E (x^4) = 1017.5

Example 4: The first three moments of a distribution about the value 2 are 1, 16 and - 40. Show that the first three moments about zero are 3,24,76.

Sol. Here E (x - 2) = 1

$\Rightarrow$ E (x) - 2 = 1

$\therefore$ E (x) = 3

E $(x - 2)^2$ = 16

$\Rightarrow$ E (x^2 + 4 - 4 x) = 16

or E (x^2) + 4 - 4 E (x) = 16

$\Rightarrow$ E (x^2) + 4 - 4 (3) = 16

or E (x^2) = 24

Also E $(x - 2)^3$ = -40

$\Rightarrow$ E [x^3 - 8 - 6x^2 + 12x] = - 40

or E (x^3) - 8 - 6 E (x^2) + 12 E (x) = - 40

or E (x^3) - 8 - 6 (24) + 12 (3) = - 40

or E (x^3) - 8 - 144 + 36 = - 40

$\Rightarrow$ E (x^3) = 76

Example 5: Find the first four moments

(a) about the origin (b) about the mean,

for a random variable x having density function

$$f(x) = \begin{cases} 4x(9-x^2)/81, & 0 \le x \le 3 \\ 0 & , \ elewhere \end{cases}$$

Sol: (a) By definition

$$\mu_4' = E(x) = \int_0^3 x\, f(x)dx = \frac{4}{81}\int_0^3 x^2(9-x^2)dx$$

$$= \frac{4}{81}\left[9\frac{x^3}{3} - \frac{x^5}{5}\right]_0^3$$

$$= \frac{4}{81}\left[\frac{81}{1} - \frac{243}{5}\right] = \frac{8}{5} = \mu$$

μ_2 ' = E (x^2) = $\int_0^3 x^2 \, f(x)dx = \frac{4}{81}\int_0^3 x^3(9-x^2)dx$

$= \frac{4}{81}\left[9\frac{x^4}{4}-\frac{x^6}{6}\right]_0^3$

$= \frac{4}{81}\left[\frac{9\times 81}{4}-\frac{9\times 81}{6}\right] = \frac{4}{81} \times \frac{9\times 81}{1} \times \frac{1}{12} = 3$

μ_3 ' = E (x^3) = $\int_0^3 x^3 \, f(x)dx = \frac{4}{81}\int_0^3 x^4(9-x^2)dx$

$= \frac{4}{81}\left[9\frac{x^5}{5}-\frac{x^7}{7}\right]_0^3 = \frac{4}{81}\left[\frac{2187}{5}-\frac{2187}{7}\right]$

$= \frac{4}{81} \times 2187 \left[\frac{2}{35}\right] = \frac{216}{35}$

μ_4 ' = E (x^4) = $\int_0^3 x^4 \, f(x)dx = \frac{4}{81}\int_0^3 x^5(9-x^2)dx$

$= \frac{4}{81}\left[9\frac{x^6}{6}-\frac{x^8}{8}\right]_0^3$

$= \frac{4}{81}\left[\frac{9\times 729}{6}-\frac{651}{8}\right] = \frac{4}{81}\times 6561 \times \frac{1}{24} = \frac{27}{2}$

(b) We know that $\mu_1' = \mu$ and $\mu_0' = 1$

$\mu_2 = \mu_2' - \mu_1'^2$

$\mu_3 = \mu_3' - 3\mu_2'\mu_1' + 2\mu_1'^3$

$\mu_4 = \mu_4' - 4\mu_3'\mu_1' + 6\mu_2'\mu_1'^2 - 3\mu_1'^4$

Thus, we have

$\mu_1 = 0$

$\mu_2 = 3 - \left(\frac{8}{5}\right)^2 = \frac{11}{25} = \sigma^2$

$$\mu_3 = \frac{216}{35} - 3\,(3)\left(\frac{8}{5}\right) + 2\left(\frac{8}{5}\right)^3 = -\frac{32}{875}$$

$$\mu_4 = \frac{27}{2} - 4\left(\frac{216}{35}\right)\left(\frac{8}{5}\right) + 6(3)\left(\frac{8}{5}\right)^2 - 3\left(\frac{8}{5}\right)^4$$

$$= \frac{3693}{8750}$$

Self-check exercise

Q. 1 The first four moments of a distribution about x = 2 are 1, 2.5, 5.5 and 16. Calculate the first four moments about the mean.

Q.2 Let the first, second and third moments of the distribution about the point 7 be 3,11 and 15 respectively. Determine the mean μ of x and then find the first, second and third moments of the distribution about the point μ.

8.6 Summary

We conclude this unit by summarizing what we have covered in it:-

1. Defined moments about origin and moments about mean for discrete and continuous probability distribution.
2. Derived the relation between the moments about the mean in terms of moments about the origin.
3. Derived the relation between the moments about the origin in terms of moments about the mean.
4. Some examples are given related to each topic so that the contents be clarified further.

8.7 Glossary:

1. Moments are numerical characteristics that describe the shape and properties of a probability distribution.
2. In case of discrete probability distribution r^{th} moment about origin, denoted by μ_r', is defined as

$$\mu_r' = E(x^r) = \sum x^r f(x)$$

In case of continuous probability distribution,

$$\mu_r' = E(x^r) = \int_{-\infty}^{\infty} x^r f(x)dx$$

Also, r^{th} moment about mean, denoted by μ_r, is defined as

μ_r = E [(x - μ)^r] = $\int_{-\infty}^{\infty} (x-\mu)^r f(x)$ [For discrete probability distribution]

μ_r = E[(x - μ)^r] = $\int_{-\infty}^{\infty} (x-\mu)^r f(x)dx$ [For continuous probability distribution]

8.8 Answers To Self-Check Exercise

Ans.1 μ_1 = 0, μ_2 = 1.5, μ_3 = 0, μ_4 = 6

Ans.2 μ= 10; μ_1 = 0, μ_2 = 2, μ_3 = -30

8.9 References/Suggested Readings

1. Robert V. Hogg, Joseph w. Mckean and Allen T. Craig, Introduction to Mathematical Statistics, Pearson Education, Asia, 2007.
2. Irwin Miller, Marylees Miller and John E. Freund, Mathematical Statistics with Application, 7th Ed., Pearson Education, Asia, 2006.
3. Sheldon Ross, Introduction to Probability Model, 9th Ed., Academic Press, Indian Reprint, 2007.

8.10 Terminal Questions

1. The first four moments of a distribution are 1,4,10 and 46 respectively. Compute the first four central moments and the Beta constants.
2. The first four moments of a distribution about the value 4 are -1.5, 17, -30 and 108. Calculate the moments about origin.
3. In a continuous distribution whose relative frequency density is given by f(x) = $\frac{3}{4}$ x (2 - x), the variable ranges from 0 to z, show that μ_3 = 0.
4. The first four moments of a distribution about the value 5 of the variable are 2, 20, 40 and 50.

   Show that μ_3 = - 64, μ_4 = 162

Unit - 9

Moment Generating Function

Structure

9.1 Introduction

The moment generating function is a powerful mathematical tool used in probability theory and statistics to characterize the probability distribution of a random variable. It uniquely determines the probability distribution of the random variable i.e. if two random variables have the same moment generating function, then they have the same probability distribution. The n^{th} moment of the random variable x can be obtained by taking the n^{th} derivative of the moment generating function with respect to t and evaluating it at t = 0. The moment generating function is useful for studying the behaviour of transformed random variables. The moment generating function is widely used in deriving probability distribution; studying properties of random variables; analyzing sums and transformations of random variables; constructing confidence intervals and hypothesis tests.

9.2 Learning Objectives

After studying this unit, you should be able to:

- Define moment generating function (m.g.f.) for discrete and continuous random variable
- Discuss moment generating function
- Discuss different properties of moment generating function
- Discuss limitation of moment generating function

9.3 Moment Generating Function (m.g.f.)

Def: The moment generating function of the distribution of a random variable x (if it exists), is given by the expected value of e^{tx}, named by

$M_x(t) = E(e^{tx}) = \sum_x e^{tx}.f(x) =$ .....(1)

When x is discrete, and

$M_x(t) = E(e^{tx}) = \int_{-\infty}^{\infty} e^{tx} f(x)dx$ ......(2)

When x is continuous

Substituting in (1), the expansion of e^{tx}, we obtain for the discrete case

$$M_x(t) = \sum_x \left[1 + tx + \frac{t^2}{\lfloor 2} x^2 + ... + \frac{t^r}{\lfloor r} x^r + ...\right] f(x)$$

$$= \sum f(x) + t \sum x\, f(x) + \frac{t^2}{\lfloor 2} \sum x^2\, f(x) + ... + \frac{t^r}{\lfloor r} \sum x^r\, f(x) + ...$$

$$= 1 + \mu_1' t + \mu_2' \frac{t^2}{\lfloor 2} + ... + \mu_r' \frac{t^r}{\lfloor r} + ...$$

Where μ_1', μ_2'....., μ_r',..... are the first, second,....., r^{th} moment about the origin and this explains the term 'moments generating function'. It can be seen from (t), however, that for any random variable x, the m.g.f. $M_x(t)$ must exist at the point t = 0 and at that point its value must be $M_x(0) = E(1) = 1$.

Suppose now that the m.g.f. of a random variable x exists for all values of t in some interval around the point t = 0. It can be shown that the derivative $M_x'(t)$ then exists at the point t = 0, and that at t = 0, the derivative of the expectation in (1) must be equal to the expectation of the derivative. That is

$$[M_x'(t)]_{t=0} = \left[\frac{d}{dt} E(e^{tx})\right]_{t=0} = E\left[\frac{d}{dt}(e^{tx})\right]_{t=0}$$

$$= E(x), \qquad \left[\because \left(\frac{d}{dt} e^{tx}\right)_{t=0} = (x\,e^{tx})_{t=0} = x\right]$$

Therefore, it follows that

$$M'_x(0) = E(x)$$

In words, it means that the derivative of the on g.f. $M_x(t)$ at t = 0 is the mean of x. Similarly, it can be shown that for n = 1,2,....., the n^{th} derivative $Mx^{(n)}(t)$ at t = 0 will satisfy the following relation.

$$M_x^{(n)}(0) = \left[\frac{d^n}{dt^n} E(e^{tx})\right]_{t=0} = E\left[\frac{d^n}{dt^n} e^{tx}\right]_{t=0}$$

$$= E\left[(x^n e^{tx})\right]_{t=0} = E(x^n)$$

Thus, $M'_x(0) = E(x)$, $M''_x(0) = E(x^2)$, ... and soon,

9.4 Properties of Moment Generating Function

Property I : If a and b are constants, then

(a) $M_{x+a}(t) = e^{at} M_x(t)$

(b) $M_{bx}{}^{(t)} = M_x{}^{(bt)}$

(c) $M_{\frac{x+a}{b}}(t) = \frac{t^a}{e^b} M_x\left(\frac{t}{b}\right)$

Proof: (a) We have

$M_{x+a}(t) = E[e^{t(x+a)}] = E(e^{tx}.\ e^{at})$

$= e^{at}.\ E(e^{tx}) = e^{at}.\ M_x(t)$

(b) $M_{bx}(t) = E[e^{t.bx}] = E[e^{tbx}] = M_x(tb)$

(c) $M_{\frac{x+a}{b}}(t) = E\left[e^{t\frac{x+a}{b}}\right] = E\left[e^{\frac{t}{b}x}.e^{\frac{ta}{b}}\right]$

$= e^{\frac{ta}{b}} E\left[e^{\frac{t}{b}x}\right] = e^{\frac{ta}{b}} M_x\left(\frac{t}{b}\right).$

Property II: Effect of change of origin and scale on Moment Generating Function

Let us transform x to the new variable U by changing both the origin and scale in x as follows:

U : $\frac{x-a}{h}$, where a and h are constants. Moment generating function of U (about origin) is given by:

$$M_U(t) = E(e^{tU}) = E\left[\exp\left\{\frac{t(x-a)}{h}\right\}\right]$$

$$= E\left[e\frac{tx}{h}.e\frac{-at}{h}\right]$$

$$= e\frac{-at}{h}.\ E\left[e^{\frac{tx}{h}}\right]$$

$$= e\frac{-at}{h}.\ M_x\left(\frac{t}{h}\right)$$

$\Rightarrow \qquad M_U(t) = e^{\frac{-at}{h}} \cdot M_x\left(\frac{t}{h}\right)$

Where $M_x(t)$ is the m.g.f. of x about origin.

In particular, if we take a = E(x) = μ (say)

and $\quad h = \sigma_x = \sigma$ (say), then

$$U = \frac{x - E(x)}{\sigma_x} = \frac{x-\mu}{\sigma} = z(\text{say})$$

is known as a standard variable. Thus the m.g.f. of a standard variate z is given by

$$M_z(t) = e^{-\mu t/\sigma} \cdot M_x\left(\frac{t}{\sigma}\right)$$

Property III : The m.g.f. of the sum of a number of independent random variables is equal to the product of their respective moment generating functions.

Proof: Let $x_1, x_2, \ldots\ldots, x_n$ be independent random variables. Then by definition of a m.g.f. applied on $x_1 + x_2 + \ldots\ldots + x_n$, we have

$$M_{x_1+x_2+\ldots\ldots+x_n}(t) = E\left[e^{t(x_1 x_2+\ldots\ldots+x_n)}\right]$$

$$= E\left[e^{tx_1} \cdot e^{tx_2} \ldots\ldots e^{tx_n}\right]$$

$$= E\left(e^{tx_1}\right) E\left(e^{tx_2}\right) \ldots\ldots E\left(e^{tx_n}\right)$$

($\because$ $x_1, x_2, \ldots\ldots, x_n$ are independent)

$$= M_{x_1}(t)\, M_{x_2}(t) \ldots\ldots M_{x_n}(t)$$

i.e. $\qquad M_{x_1+x_2+\ldots\ldots+x_n}(t) = M_{x_1}(t)\, M_{x_2}(t) \ldots\ldots M_{x_n}(t)$

Hence the result

Property IV: Moment generating function of a random variable x about point x = a generate r^{th} moment about the point x = a.

Proof: Moment generating function about x = a is given by

$$M_x(t) = E[e^{t(x-a)}]$$

$$= E\left[1 + t(x-a) + \frac{t^2}{\lfloor 2}(x-a)^2 + \ldots\ldots + \frac{t^r}{\lfloor r}(x-a)^r + \ldots\ldots\right]$$

$$= 1 + t\mu_1' + \frac{t^2}{\lfloor 2}\mu_2' + \ldots\ldots + \frac{t^r}{\lfloor r}\mu_r' + \ldots\ldots$$

Which clearly shows that coefficient of $\frac{t^r}{\lfloor r} = \mu_r'$ (about x = a)

Property V: Property of Uniqueness

The m.g.f. of a distribution, if it exists, uniquely determines the distribution a given probability distribution, there is only one m.g.f. (provided it exists) and corresponding to a given m.g.f., there is only one probability distribution. Hence $M_x(t) = M_y(t) \Rightarrow$ x and y are identically distributed.

Property VI: Limitation of the Moment Generating Function

Moment generating function suffers from some drawbacks which has restricted its use in statistics. These are

1. A random variable x may have no moments although its m.g.f. exists.
2. A random variable x can have m.g.f. and some (or all) moments, yet the m.g.f. does not generate the moments.
3. A random variable x can have all or some moments, but m.g.f. does not exist except perhaps at one point.

Let us improve our understanding of these results by looking at some of the following examples:

Example 1: The random variable x can assume the value 1 and -1 with probability $\frac{1}{2}$ each. Find

(a) the moment generating function

(b) the first four moments about the origin

Sol. Given information in the tabular form can be put as under:

x	f(x)
1	1/2
-1	1/2

(a) $M^x(t) = E(e^{tx}) = \sum_x e^{tx} f(x)$

$= e^{t(1)} \left(\frac{1}{2}\right) + e^{t(-1)} \left(\frac{1}{2}\right)$

$= \frac{1}{2} \left(e^t + e^{-t}\right)$

(b) We have

$$e^t = 1 + t + \frac{t^2}{\lfloor 2} + \frac{t^3}{\lfloor 3} + \frac{t^4}{\lfloor 4} + \ldots\ldots\ldots\ldots$$

$$e^{-t} = 1 - t - \frac{t^2}{\lfloor 2} - \frac{t^3}{\lfloor 3} + \frac{t^4}{\lfloor 4} \ldots\ldots\ldots\ldots$$

Now $$\frac{1}{2}(e^t \times e^{-t}) = 1 + \frac{t^2}{\lfloor 2} + \frac{t^4}{\lfloor 4} + \ldots\ldots\ldots\ldots \qquad \ldots(1)$$

$$\therefore M_x(b) = 1 + \frac{t^2}{\lfloor 2} + \frac{t^4}{\lfloor 4} + \ldots\ldots\ldots\ldots$$

But $$M_x(b)\ 1 + \mu_1^1 t + \mu_2^1 \frac{t^2}{\lfloor 2} + \mu_3^1 \frac{t^3}{\lfloor 3} + \mu_4^1 \frac{t^4}{\lfloor 4} + \ldots\ldots\ldots\ldots \qquad \ldots.(2)$$

Comparing (1) and (2), we have

$\mu_1^1 = 0$, $\mu_2^1 = 1$, $\mu_3^1 = 0$, $\mu_4^1 = 1$, ...

The odd moments are all zero, and even moments are all one.

Example 2 : Find m.g.f. of Binomial variate and hence find mean and variance.

Sol. : - Let x be a binomial random variable

$\therefore P(x = r) = n_{c_r}\ p^r q^{n-r}$, r = 0, 1, 2, ..., n

p + q = 1, p, q > 0

By definition

$$M_x(t) = E(e^{tx}) = \sum_{r=0}^{n} e^{tr} P(x = r)$$

$$= \sum_{r=0}^{n} e^{tr} n_{c_r}\ p^r q^{n-r}$$

$$= \sum_{r=0}^{n} n_{c_r}\ (p\,.\,e^t)^r q^{n-r}$$

$$= n_{c_0}(p\,.\,e^t)^0\,.\,q^n + n_{c_1}(pe^t)\,q^{n-1} + n_{c_2}(pe^t)^2 q^{n-2} + \ldots\ldots.$$

$$+\ n_{c_n}(pe^t)^n q^0$$

$$= (q + pe^t)^n$$

$$\therefore \quad Mx(t) = (q + pe^t)^n$$

$\Rightarrow \quad \dfrac{dM_x(t)}{dt}$ = n(q + pe^t)$^{n-1}$ × pe^t

$\therefore \quad \mu_1^1 = \left|\dfrac{dM_x(t)}{dt}\right|_{t=0}$ = n(q + p)$^{n-1}$ p = np

[∴ q + p = 1]

$\Rightarrow \quad \mu_1^1$ = np

Also $\dfrac{d^2M_x(t)}{dt^2}$ = np[(q + pe^t)$^{n-1}$ e^t + e^t (n - 1).

(q + pe^t)$^{n-1}$ pe^t]

$\Rightarrow \quad \left|\dfrac{d^2M_x(t)}{dt^2}\right|_{t=0}$ = np[(q+p)$^{n-1}$ + (n-1) (q + p) $^{n-2}$ p]

$\Rightarrow \quad \mu_2^1 = \left|\dfrac{d^2M_x(t)}{dt^2}\right|_{t=0}$ = np[(q+p)$^{n-1}$ + (n-1) (q + p) $^{n-2}$ p]

$\Rightarrow \quad \mu_2^1$ = np[1+ (n-1)p] = np [1+ np - p]

$\therefore \quad \mu_2 = \mu_2^1 - \mu_1^{2}$ = np + n^{2}p^2 - np^2 - n^2 p^2

= np - np^2 = np(1- p) = npq

∴ Mean = np, variance npq, $M_x(t) = (q + pe^t)^n$

Example 3 : In a continuous distribution whose relative frequency density is given by f (x) = $\dfrac{3}{4}$ × (2 - x), the variable ranges from 0 to 2. Show that the distribution is symmetrical with mean = 1 and variance = $\dfrac{1}{5}$. Show that the third moment about x = 0 is $\dfrac{5}{8}$. Verify $\mu_3 = 0$.

Sol. : $\displaystyle\int_{-\infty}^{\infty} f(x)dx = \int_0^2 \frac{3}{4}(2-x)dx$

$$= \int_0^2 \frac{3}{4}(2x - x^2)dx$$

$$= \frac{3}{4}\left[x^2 - \frac{x^3}{3}\right]_0^2 = 1$$

and f (x) $\geq$ 0 for 0 $\leq$ x $\leq$ 2

$\therefore$ f (x) is the density function

Now using μ_r^1 (about the origin) = $\int_a^b f(x)dx$

we have

$$\mu_1^1 \text{ (about the origin) } = \frac{3}{4}\int_0^2 x^2(2-x)dx$$

$$= \frac{3}{4}\int_0^2 (2x^2 - x^3)dx$$

$$= \frac{3}{4}\left[\frac{2x^3}{3} - \frac{x^4}{4}\right]_0^2$$

$$= \frac{3}{4}\left[\frac{16}{3} - \frac{16}{4}\right] = \frac{3}{4} \times 16\left[\frac{1}{12}\right] = 1$$

$$\mu_2^1 \text{ (about the origin) } = \frac{3}{4}\int_0^2 x^3(2-x)dx$$

$$= \frac{3}{4}\int_0^2 (2x^3 - x^4)dx$$

$$= \frac{3}{4}\left[\frac{2x^4}{4} - \frac{x^5}{5}\right]_0^2$$

$$= \frac{3}{4}\left[\frac{32}{4} - \frac{32}{5}\right] = \frac{3}{4} \times 32 \times \frac{1}{20} = \frac{5}{6}$$

$$\mu_3^1 \text{ (about the origin) } = \frac{3}{4}\int_0^2 x^4(2-x)dx$$

$$= \frac{3}{4}\int_0^2 (2x^4 - x^5)dx$$

$$= \frac{3}{4}\left[\frac{2x^5}{5} - \frac{x^6}{6}\right]_0^2$$

$$= \frac{3}{4}\left[\frac{64}{5} - \frac{64}{6}\right] = \frac{3}{4} \times 64 \times \frac{1}{30} = \frac{5}{8}$$

$\therefore$ Mean $\mu_1^1 = 1$;

and variance $\mu_2 = \mu_2^1 - (\mu_1^1)^2 = \frac{6}{5} - 1 = \frac{1}{5}$,

$$\mu_3 = \mu_3^1 - 3\mu_2^1 \quad \mu_1^1 + 2\ \mu_1^{13} = \frac{8}{5} - \frac{18}{5} + 2 = 0$$

Since μ_3, which is measure of skewness is zero, the distribution is symmetrical about the mean.

Example 4 : For the Bernouth distribution

$$f(x;\theta) = \begin{cases} \theta^k (1-\theta)^{1-k}, & x = 0,1 \\ 0, & \text{elsewhere} \end{cases}$$

Find $M_x(t)$ and hence find μ_1^1 and μ_2. Also find β_1 and β_2.

Sol. :- We have

$$M_x(t) = E(e^{tx}) = \sum_x e^{tx} f(x;\theta)$$

$$= \sum_{x=0}^{1} e^{tx}\theta^x (1-\theta)^{1-x}$$

$$= \sum_{x=0}^{1} (\theta e^t)^x (1-\theta)^{1-x}$$

$= [(1 - \theta) + \theta e^t]$, which is the required m.g.f.

μ_1^1 (mean) $= \left[M_x^{'}(t)\right]_{t=0} = [\theta e^t]_{t=0} = \theta$

$\mu_2^1 = \left[M_x^{''}(t)\right]_{t=0} = [\theta\ e^t]_{t=0} = \theta$

Therefore, μ_2 (variance) $= \mu_2^1 - \mu_1^{12}$

$= \theta - \theta^2$

$= \theta\ (1 - \theta)$

Thus, the mean and variance of the Bernoulli distribution θ, $\theta\ (1 - \theta)$ respectively.

Also, $\beta_1 = \frac{\mu_3}{\mu_2^{3/2}}$, $\beta_2 = \frac{\mu_4}{\mu_2^2}$

Now, $\mu_3^1 = \left[M_x^{III}(t)\right]_{t=0} = [\theta\ e^t]_{t=0} = \theta$

$\mu_4^1 = \left[M_x^{iv}(t)\right]_{t=0}$ = $[\theta\, e^t]_{t=0} = \theta$

Thus, $\mu_3 = \mu_3^1 - 3\mu_2^1\,\mu_1^1 + 2\mu_1^{1\ 3}$

$= \theta - 3\,\theta\,.\,\theta + 2\theta^3$

$= \theta\,(1 - 3\theta + 2\theta^2)$

$= \theta\,(1 - \theta)\,(1 - 2\theta)$

$\mu_4 = \mu_4^1 - 4\mu_3^1\,\mu_1^1 + 6\mu_2^1\,\mu_1^{1\ 2}\,,\ 3\mu_1^{1\ 4}$

$= \theta - 4\theta.\theta + 6\theta.\theta^2 - 3\theta^4$

$= \theta\,(1 - 4\theta + 6\theta^2 - 3\theta^3)$

$= \theta(1 - \theta)\,(1 - 3\theta^2)$

$= \theta(1 - \theta)\,[1 - 3\theta\,(1 - \theta)]$

Hence, $\beta_1 = \dfrac{\mu_3}{\mu_2^{3/2}} = \dfrac{\theta(1-\theta)(1-2\theta)}{[\theta(1-\theta)]^{3/2}} = \dfrac{1-2\theta}{\sqrt{\theta(1-\theta)}}$

$\beta_2 = \dfrac{\mu_4}{\mu_2^2} = \dfrac{\theta(1-\theta)[1-3\theta(1-\theta)]}{\theta^2(1-\theta)^2}$

$= \dfrac{1-3\theta(1-\theta)}{\theta(1-\theta)}$

Example 5: Find the moment generating function for the distribution defined by

$$f(x) = \begin{cases} x\ ,\ 0 \le x \le 1 \\ 2-x, 1 \le x \le 2 \\ 0,\ elsewhere \end{cases}$$

Sol. Here $\quad f(x) = \begin{cases} x\ ,\ 0 \le x \le 1 \\ 2-x, 1 \le x \le 2 \\ 0,\ elsewhere \end{cases}$

Moment generating function = M(t) = E(e^{tx})

$$= \int_{-\infty}^{\infty} e^{tx} f(x)dx = \int_{-\infty}^{0} e^{tx} 0dx + \int_{0}^{1} e^{tx} x\,dx + \int_{1}^{2} e^{tx}(2-x)\,dx + \int_{2}^{\infty} e^{tx} 0\,dx$$

$$= \int_0^1 xe^{tx}\,dx + \int_1^2 (2-x)e^{tx}dx$$

$$= \left\{\left[x\frac{e^{tx}}{t}\right]_0^1 - \int_0^1 1.\frac{e^{tx}}{t}\right\} + \left\{\left[(2-x)\frac{e^{tx}}{t}\right]_1^2 - \int_1^2 (-1)\frac{e^{tx}}{t}\right\}$$

$$= \left[\frac{xe^{tx}}{t}\right]_0^1 - \frac{1}{t^2}\left[e^{tx}\right]_0^1 + \left[\frac{(2-x)e^{tx}}{t}\right]_1^2 + \frac{1}{t^2}\left[e^{tx}\right]_1^2$$

$$= \left[\frac{1}{t}e^t - 0\right] - \frac{1}{t^2}(e^t - e^0) + (0 - \frac{1}{t}e^t) + \frac{1}{t^2}(e^{2t} - e^t)$$

$$= \frac{1}{t}e^t - \frac{1}{t^2}e^t + \frac{1}{t^2} - \frac{1}{t}e^t + \frac{1}{t^2}e^{2t} - \frac{1}{t^2}e^t$$

$$= \frac{e^{2t}}{t^2} - 2\frac{e^t}{t^2} + \frac{1}{t^2}$$

$$= \left[\frac{e^t - 1}{t}\right]^2, \ t \neq 0$$

Example 6: Let $f(x) = \begin{cases} \frac{1}{3}, & -1 < x < 2 \\ 0, & elsewhere \end{cases}$

be the p.d.f. of random variable x. Show that m.g.f. of x is

$$M(t) = \begin{cases} \frac{e^{2t} - e^{-1}}{3t}, & t \neq 0 \\ 1, & t = 0 \end{cases}$$

Sol. Moment generating function

$$M(t) = E(e^{tx}) = \int_{-\infty}^{\infty} e^{tx} f(x)dx$$

$$= \int_{-1}^{2} e^{tx}\frac{1}{3}dx = \frac{1}{3}\left[\frac{e^{tx}}{t}\right]_{-1}^{2} = \frac{e^{2t} - e^{-t}}{3t}, \ t \neq 0$$

For t = 0, M(t) = $\int_{-1}^{2} e^{ox} \frac{1}{3} dx = \frac{1}{3} \int_{-1}^{2} 1 dx = \frac{1}{3}[x]_{-1}^{2} = \frac{1}{3}$ (2 + 1) = 1

$\therefore$ M(t) = $\begin{cases} \frac{e^{2t} - e^{-t}}{3t}, & t \neq 0 \\ 1, & t = 0 \end{cases}$

Example 7: The probability density function of the random variable x follows the probability law:

$$p(x) = \frac{1}{2\theta} \exp\left(-\frac{1x-\theta 1}{\theta}\right), -\infty < x < \infty$$

Find the m.g.f. of x. Hence or otherwise find E(x) and V(x).

Sol. The m.g.f. of x is

$$M_x(t) = E(e^{tx}) = \int_{-\infty}^{\infty} \frac{1}{2\theta} \exp\left(-\frac{1x-\theta 1}{\theta}\right) e^{tx}\, dx$$

$$= \int_{-\infty}^{\theta} \frac{1}{2\theta} \exp\left(-\frac{\theta - x}{\theta}\right) e^{tx}\, dx + \int_{\theta}^{\infty} \frac{1}{2\theta} \exp\left(-\frac{x-\theta}{\theta}\right) e^{tx}\, dx$$

For $x \in (-\infty, \theta)$, $x - \theta < 0 \Rightarrow \theta - x > 0$

$\therefore$ $|x - \theta| = \theta - x \ \forall\ x \in (-\infty, \theta)$

Similarly $|x - \theta| = x - \theta \ \forall\ x \in (\theta, \infty)$

$\therefore$ $M_x(t) = \frac{e^{-1}}{2\theta} \int_{-\infty}^{\theta} \exp\left\{x\left(t + \frac{1}{\theta}\right)\right\} dx + \frac{e}{2\theta} \int_{\theta}^{\infty} \exp\left\{-x\left(\frac{1}{\theta} - t\right)\right\} dx$

$$= \frac{e^{-1}}{2\theta} \frac{1}{t + \frac{1}{\theta}} \exp\left\{\theta\left(t + \frac{1}{\theta}\right)\right\} + \frac{e}{2\theta} \frac{1}{\left(\frac{1}{\theta} - t\right)} \exp\left\{-\theta\left(\frac{1}{\theta} - t\right)\right\}$$

$$= \frac{e^{\theta t}}{2(\theta t + 1)} + \frac{e^{\theta t}}{2(1 - \theta t)} = \frac{e^{\theta t}}{1 - \theta^2 t^2} = e^{\theta t}(1 - \theta^2 t^2)^{-1}$$

$$= \left(1 + \theta t + \frac{\theta^2 t^2}{\lfloor 2} + \ldots\ldots\right)(1 + \theta^2 t^2 + \theta^4 t^4 + \ldots\ldots\ldots)$$

$$= 1 + \theta t + \frac{3\theta^2 t^2}{\lfloor 2} + \ldots\ldots\ldots\ldots$$

$\therefore$ E(x) = μ_1^1 = Coeff. of t in $M_x(t) = \theta$

$\therefore$ Mean = θ

Now μ_2^1 = Coeff. of $\frac{t^2}{\underline{|2}}$ in $M_x(t) = 3\theta^2$

Hence var (x) = $\mu_2^1 - \mu_1^{1\ 2} = 3\theta^2 - \theta^2 = 2\theta^2$

$\therefore$ Variance (x) = $2\theta^2$

Self-check exercise

Q. 1 For the discrete uniform distribution

$$f(x) = \begin{cases} \frac{1}{k}, \ for x = 1, 2, ..., k \\ 0 \ , \ elsewhere \end{cases}$$

Find the m.g.f. and hence find μ_1^1 and M_2 (i.e. mean and variance).

Q.2 Let the random variable x assume the value 'r' with the probability law:

P (x = r) = 2^{r-1}. p ; r = 1,2,3,........

Find the moment generating function of x and hence its mean and variance.

Q.3 A random variable x has probability density function given by

$$\boldsymbol{f(x) =} \frac{1}{2} e^{-|x|}, -\infty < x < \infty$$

Find its m.g.f. Hence find the variance of the distribution.

9.5 Summary

We conclude this unit by summarizing what we have covered in it:-

1. Defined moments generating function (m.g.f.) for discrete and continuous random variable.
2. Discussed moment generating function further.
3. Discussed in detail different properties of moment generating function
4. Discussed limitations of moment generating function.
5. Some examples are given related to each topic so that the contents be clarified further.

9.6 Glossary:

1. The moment generating function (m.g.f.) of the distribution of a random variable x (if it exists), is given by the expected value of e^{tx}, named by

$$M_x(t) = E(e^{tx}) = \sum_x e^{tx} f(x),$$

When x is discrete, and

$$Mx(t) = E(etx) = \int_{-\infty}^{\infty} e^{tx} f(x)dx,$$

When x is continuous.

2. The m.g.f. of the sum of a number of independent random variables is equal to the product of their respective moment generating functions.

9.7 Answers To Self-Check Exercise

Ans.1 Moment generating function (m.g.f.)

$$= M_x(t) = \frac{e^t(1-e^{tk})}{k(1-e^t)}$$

$$\mu_1^1 = \frac{k+1}{2}$$

$$\mu_2 = \frac{k^2-1}{12}$$

Ans.2 Moment generating function (m.g.f.)

$$= M_x(t) = \frac{pe^t}{1-qe^t}$$

Mean = μ_1^1 (about origin) = $\frac{1}{p}$

Variance = $\mu_2 = \frac{q}{p^2}$

Ans.3 Moment generating function = $\frac{1}{1-t^2}$, $|t| < 1$ Variance = 2

9.8 References/Suggested Readings

1. Robert V. Hogg, Joseph w. Mckean and Allen T. Craig, Introduction to Mathematical Statistics, Pearson Education, Asia, 2007.

2. Irwin Miller, Marylees Miller and John E. Freund, Mathematical Statistics with Application, 7th Ed., Pearson Education, Asia, 2006.

3. Sheldon Ross, Introduction to Probability Model, 9th Ed., Academic Press, Indian Reprint, 2007.

9.9 Terminal Questions

1. A random variable x has density function given by

$$f(x) = \begin{cases} 2e^{-2x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

Find (a) the moment generating function

(b) the first four moments about the origin

2. Let x is a Bernoulli Variate with parameter p i.e. P(x = 1) = p, P(x = 0) = q; p,q > 0 and p+q = 1.

Find moment generating function of x and hence find mean and variance.

3. Let $f(x) = \begin{cases} \left(\frac{1}{2}\right)^x, & x = 1, 2, 3, \ldots.. \\ 0, & elsewhere \end{cases}$

be the p.d.f. of random variable x. Find m.g.f., mean and variance of x.

4. Find the r^{th} moment of the distribution that has m.g.f.

$M(t) = (1 - t)^{-3}, \; t < 1$

Unit - 10

Cumulant Generating Function And Characteristic Function

Structure

10.1 Introduction

Cumulants are a set of statistical measures that characterize the shape of a probability distribution. They are defined as the coefficients in the Taylor series expansion of the natural logarithm of the Characteristic function. They provide a way to quantity the deviation of a distribution from a normal distribution. The first cumulant is the mean, the second cumulant is the variance, the third cumulant is the skewness and the fourth cumulant is the kurtosis. The cumulant generating function is the natural logarithm of the characteristic function of a random variable. Cumulants are often used in statistical inference, signal processing and time series analysis. The eumulant generating function provides a concise way to characterize the entire distribution of a random variable.

The Characteristic function is a powerful mathematical tool that plays a crucial role in probability theory and its applications. It provides an alternative representation of the probability distribution of the random variable X. It exists for any random variable X with a finite first moment. The characteristic function uniquely determines the probability distribution of the random variable X. The characteristic function is a continuous function of the parameter t.

10.2 Learning Objectives

After studying this unit, you should be able to:

- Define cumulant generating function
- Find the series expansion of cumulant generating function

- Discuss properties of cumulants
- Define characteristic function
- Find the series expansion of the characteristic function
- Discuss properties of characteristic function

10.3 Cumulants And Cumulant Generating Function

Definition: Cumulant generating function k(t) is defined as:

$K_x(t) = \log_e M_x(t) = l_n[M_x(t)]$

provided the right hand side can be expanded as a convergent series in powers of t.

Series Expansion:

$$K_x(t) = K_1 t + K_2 \frac{t^2}{\lfloor 2} + \ldots\ldots + K_r \frac{t^r}{\lfloor r} + \ldots\ldots = \log M_x(t)$$

$$= \log \left(1 + \mu_1^1 t + \mu_2^1 \frac{t^2}{\lfloor 2} + \mu_3^1 \frac{t^3}{\lfloor 3} + \ldots\ldots + \mu_r^r \frac{t^r}{\lfloor r} + \ldots\ldots\right)$$

Where K_r = coefficient of $\frac{t^r}{\lfloor r}$ in $K_x(t)$ is called the r^{th} cumulant. Hence

$$K_1 t + K_2 \frac{t^2}{\lfloor 2} + K_3 \frac{t^3}{\lfloor 3} + \ldots\ldots\ldots + Kr \frac{t^r}{\lfloor r}$$

$$= \left(\mu_1^1 t + \mu_2^1 \frac{t^2}{\lfloor 2} + \mu_3^1 \frac{t^3}{\lfloor 3} + \ldots\ldots + \mu_r^r \frac{t^r}{\lfloor r} + \ldots\ldots\right) - \frac{1}{2}\left(\mu_1^1 t + \mu_2^1 \frac{t^2}{\lfloor 2} + \mu_3^1 \frac{t^3}{\lfloor 3} + \ldots\ldots\right)^2$$

$$+ \frac{1}{3}\left(\mu_1^1 t + \mu_2^1 \frac{t^2}{\lfloor 2} + \ldots\ldots\right)^3 - \frac{1}{4}\left(\mu_1^1 t + \mu_2^1 \frac{t^2}{\lfloor 2} + \ldots\ldots\right)^4 + \ldots\ldots$$

comparing the coefficients of like powers of t on both sides, we get the relationship between the moments and cumulants.

Hence, we have

$K_1 = \mu_1^1$ = mean

$$\frac{k_2}{\lfloor 2} = \frac{\mu_2^1}{\lfloor 2} - \frac{\mu^{1\ 2}}{\lfloor 2}$$

$\therefore$ $K_2 = \mu_2^1 - \mu_1^{1\ 2} = \mu_2$ = variance

$$\frac{k_3}{\lfloor 3} = \frac{\mu_3^1}{\lfloor 3} - \frac{1}{2} \cdot \frac{2\mu_1^1\mu_2^1}{\lfloor 2} + \frac{\mu_1^{1\ 3}}{3}$$

$\therefore$ $K_3 = \mu_3^1 - 3\mu_2^1\,\mu_1^1 + 2\mu_1^{1\ 3} = \mu_3$

Also $\frac{k_4}{\lfloor 4} = \frac{\mu_4^1}{\lfloor 4} = \frac{1}{2}\left(\frac{\mu_2^{1\ 2}}{4} + \frac{2\mu_1^1\mu_3^1}{\lfloor 3}\right) + \frac{1}{3}\,\frac{3\mu_1^{1\ 2}\mu_2^1}{2} - \frac{\mu_1^{1\ 4}}{4}$

$\Rightarrow$ $K_4 = \mu_4^1 - 3\mu_2^{1\ 2} - 4\mu_1^1\,\mu_3^1 + 12\mu_1^{1\ 2}\mu_2 - 6\mu_1^{1\ 4}$

$= (\mu_4^1 - 4\mu_3^1\,\mu_1^1 + 6\mu_2^1\,\mu_1^{1\ 2} - 3\mu_1^{1\ 4}) - 3(\mu_2^{1\ 2} - 2\mu_2^1\,\mu_1^{1\ 2} + \mu_1^{1\ 4})$

$= \mu_4 - 3(\mu_2^1 - \mu_1^{1\ 2})^2$

$[\because\ \mu_4 = \mu_4^1 - 4\mu_3^1\,\mu_1^1 + 6\mu_2^1\,\mu_1^{1\ 2} - 3\mu_1^{1\ 4}]$

$= \mu_4 - 3\mu_2^2$

$= \mu_4 - 3k_2^2$ $[\because\ \mu_2 = k_2]$

$\therefore$ $\mu_4 = k_4 + 3k_2^2$

Hence, we obtained

Mean = k_1

Variance = $\mu_2 = k_2$

$\mu_3 = k_3$

$\mu_4 = k_4 + 3k_2^2$

If we differentiate both sides of (1) w.r.t. t, 'r' times and then put t = 0, we get

$$k_r = \left|\frac{d^r}{dt^r}k_x(t)\right|_{t=0}$$

10.4 Properties of Cumulants

Property I: Additive Property of Cumulants

The r^{th} cumulant of the sum of the independent random variables is equal to the sum of the r^{th} cumulants of the individual variables. Symbolically,

$k_r(x_1 + x_2 + \ldots\ldots + x_n) = k_r(x_1) + k_r(x_2) + \ldots\ldots + k_r(x_n)$

Where x_i ; i = 1,2,....., n are independent random variables.

Proof: Since x_i is are independent, $M_{x_1+x_2+\ldots\ldots+x_n}(t) = M_{x_1}(t)\,M_{x_2}(t)\ \ldots\ldots\ M_{x_n}(t)$

Taking logarithms of each side, we get

$K_{x_1+x_2+.......+x_n}(t) = K_{x_1}$ (t) + K_{x_2} (t) +.......+ K_{x_n} (t)

Differentiating both sides w.r.t t 'r' times and putting t = 0, we get

$$\left[\frac{d^r}{dt^r}K_{x_1+x_2+.......+x_n}(t)\right]_{t=0} = \left|\frac{d^r}{dt^r}K_{x_1}(t)\right|_{t=0} + \left|\frac{d^r}{dt^r}K_{x_2}(t)\right|_{t=0} +.........+ \left|\frac{d^r}{dt^r}K_{x_n}(t)\right|_{t=0}$$

$\therefore$ K_r (x_1 + x_2 +........+ x_n) = K_r (x_1) + K_r (x_2) +.......+ K_r (x_n)

which establishes the result

Property II : Effect of change of origin and scale and Cumulants

If we take U = $\frac{x-a}{h}$, then

$$M_U(t) = \exp\left(-\frac{at}{h}\right) M_x\left(\frac{t}{h}\right)$$

$$\therefore \quad K_U(t) = \log M_U(t) = -\frac{at}{h} + K_x\left(\frac{t}{h}\right)$$

$$\Rightarrow \quad k_1^1 t + k_2^1 \frac{t^2}{\lfloor 2} +........+ k_r^1 \frac{t^r}{\lfloor r} + ... = -\frac{at}{h} + K_1\left(\frac{t}{h}\right)$$

$$+ K_2\left(\frac{t}{h}\right)^2 +..........+ K_r \frac{\left(\frac{t}{h}\right)^r}{\lfloor r} +.........$$

Where k_r^1 and k_r are r^{th} cumulants of U and X respectively.

Comparing Coefficients, we get

$$k_1^1 = \frac{k_1 - a}{h} \text{ and } k_r^1 = \frac{k_r}{h^r}\text{ ; r = 2,3,.......}$$

Thus we see that except the first cumulant, all cumulants are independent of change of origin. But the cumulants are not invariant of the change of scale as the r^{th} cumulant of U is $\left(\frac{1}{h^r}\right)$ times the r^{th} cumulant of the distribution of X.

10.5 Characteristic Function

Definition : Characteristic function is defined as

$\phi_x(t) = E(e^{itx})$

$$= \begin{cases} \int e^{itx} f(x)dx \text{ (for continuous probability distribution)} \\ \sum_x e^{itx} f(x) \text{ (for discrete probability distribution)} \end{cases}$$

If F(x) is the distribution function of a continuous random variable x, then

$$\phi_F(t) = \int_{-\infty}^{\infty} e^{itx} dF(x)$$

Obviously $\phi(t)$ is a complex valued function of real variable t. It may be noted that

$$|\phi(t)| = \left| \int_{-\infty}^{\infty} e^{itx} f(x)dx \right| \leq \int_{-\infty}^{\infty} \left|e^{itx}\right| f(x)dx$$

$$= \int_{-\infty}^{\infty} f(x)dx = 1 \qquad \text{......(1)}$$

[$\because$ $|e^{itx}| = |\cos tx + i \sin t x|^{1/2}$

$= [(\cos^2 tx + \sin^2 tx)^{1/2}]^2 = 1$]

Since $|\phi(t)| \leq 1$, characteristic function $\phi_x(t)$ always exists.

Series Expansion of the Characteristic Function :

$\phi_x(t) = E(e^{itx})$

$$= E\left[1 + |\, itx + \frac{(it)^2}{\lfloor 2} x^2 + \ldots\ldots + \frac{(it)^r}{\lfloor r} X^r + \ldots\ldots\ldots\right]$$

$$= 1 + it\, E(x) + \frac{(it)^2}{\lfloor 2} E(x^2) + \ldots\ldots\ldots + \frac{(it)^r}{\lfloor r} E(x^r) + \ldots\ldots$$

$$= 1 + it\, \mu_1^1 + \frac{(it)^2}{\lfloor 2} \mu_2^1 + \ldots\ldots\ldots + \frac{(it)^r}{\lfloor r} \mu_r^1 + \ldots\ldots\ldots$$

Where $\mu_r^1 = E(x^r)$, is the r^{th} moment of x about origin.

$\therefore \qquad \mu_r^1$ = coefficient of $\frac{(it)^r}{\lfloor r}$ in $\phi_x(t)$.

Hence, like m.g.f., the characteristic function $\phi(t)$ also generates moments

Rewriting (1), we get

$$\phi_x(t) = \sum_{r=0}^{\infty} \frac{(it)^r}{\lfloor r} \mu_r^1$$

10.6 Properties of Characteristic Function

Property I : For all real t, we have

(i) $\phi_{(0)} = \int_{-\infty}^{\infty} df(x) = 1$ (ii) $|\phi(t)| \leq 1 = \phi(0)$

Property II : ϕ (t) is continuous everywhere, i.e. $\phi(t)$ is continuous function of t in $(-\infty, \infty)$. Rather $\phi(t)$ is uniformly continuous in 't'.

Property III : $\phi_x(t)$ and $\phi_x(t)$ are conjugate functions,

i.e. $\phi_x(-t) = \phi_x(t)$, where $\bar{a}$ is the complex conjugate of a.

Property IV : If the distribution function of a random variable x is symmetrical about zero i.e. if

$1 - F(x) = F(-x) \Rightarrow f(-x) = f(x)$ then $\phi_x(t)$ is real valued and even function of t.

Proof: By definition, we have

$$\phi_x(t) = \int_{-\infty}^{\infty} e^{itx} f(x)dx = \int_{-\infty}^{\infty} e^{ity} f(-y)dy$$

$$= \int_{-\infty}^{\infty} e^{ity} f(y)dy$$

[Putting x = -y]

[$\because f(-y) = f(y)$]

$= \phi_x(-t)$

$\Rightarrow$ $\phi_x(t)$ is an even function of t.

From property III and equation (1), we get

$\phi x(t) = \phi x(-t) = \overline{\phi_x(t)}$

Hence $\phi_x(t)$ is a real valued and even function of t.

Property V : If x is some random variable with characteristic function $\phi_x(t)$ and if $\mu_r^1 = E(x^r)$ exists, then

$$\mu_r^1 = (-i)^r \left| \frac{\partial^r}{\partial t^r} \phi(t) \right|_{t=0}$$

Proof : $\phi(t) = \int_{-\infty}^{\infty} e^{itx} f(x)dx$

Differentiating (under the integral sign) 'r' times w.r.t. t, we have

$$\frac{\partial^r}{\partial t^r}\phi(t) = \int_{-\infty}^{\infty} (ix)^r . e^{itx} f(x) dx$$

$$= (i)^r \int_{-\infty}^{\infty} x^r . e^{itx} f(x) dx$$

$$\therefore \qquad \left| \frac{\partial^r}{\partial t^r}\phi(t) \right|_{t=0} = (i)^r \left| \int_{-\infty}^{\infty} x^r . e^{itx} f(x) dx \right|_{t=0}$$

$$= (i)^r \int_{-\infty}^{\infty} x^r f(x) dx$$

$$= i^r E(x^r) = i^r . \mu_r^1$$

Hence

$$\mu_r^1 = \left(\frac{1}{i}\right)^r \left| \frac{\partial^r}{\partial t^r}\phi(t) \right|_{t=0} = (i)^r \left| \frac{\partial^r}{\partial t^r}\phi(t) \right|_{t=0}$$

Property VI : ϕ_{cx} (t) = ϕ_x (ct), c being a constant.

Property VII : If x_1 and x_2 are independent random variables, then

$$\phi_{x_1+x_2}(t) = \phi_{x_1}(t) * \phi_{x_2}(t)$$

More generally, for n independent random variables

x_i, i = 1, 2, ... , n, we have

$$\phi_{x_1+x_2+\ldots\ldots+x_n}(t) = \phi_{x_1}(t)\ \phi_{x_2}(t) \ldots\ldots \phi_{x_n}(t)$$

Property VIII : Effect of change of origin and scale on characteristic function

If U = $\frac{x-a}{h}$, a and h being constants, then

$$\phi_U(t) = e^{-iat/h}\,\phi_x\left(\frac{t}{h}\right)$$

In particular if we take a = E(x) = μ (say) and h = σ_x = σ, then the characteristic function of the standard variate z = $\frac{x-E(x)}{\sigma_x} = \frac{x-\mu}{\sigma}$, is given by

$$\phi z(t) = e^{i\mu t/\sigma} . \phi_x\left(\frac{t}{\sigma}\right)$$

Property IX : Necessary and sufficient conditions for a function ϕ (t) to be a Characteristic Function.

ϕ(t) is a characteristic function if

(i) $\phi(0) = 1$

(ii) $\phi(t) = \phi(-t)$

(iii) $\phi(t)$ is continuous

(iv) $\phi(t)$ is convex for t > 0 i.e. for t1, t2 > 0

$$\phi\left[\frac{1}{2}(t_1+t_2)\right] \le \frac{\phi(t_1)+\phi(t_2)}{2}$$

(v) $\lim_{t\to\infty} \phi(t) = 0$

Let us improve our understanding of these results f by looking at some of the following examples :-

Example 1 : If μ_r^1 is the r^th^ moment about origin, prove that

$$\mu_r^1 = \sum_{j=1}^{r} \binom{r-1}{j-1} \mu_{r-j}^1 \; K_j, \text{ where } K_j \text{ is the } j^{th} \text{ cumulant.}$$

Sol. : By definition

$K_x(t) = \ln[M_x(t)]$

Differentiating both sides of

$$\frac{K_1 t + K_2 \frac{t^2}{\lfloor 2} + \ldots\ldots\ldots + K_r \frac{t^r}{\lfloor r} + \ldots\ldots\ldots\ldots}{r}$$

$$= \log\left(1 + \mu_1^1 t + \mu_2^1 \frac{t^2}{\lfloor 2} + \mu_3^1 \frac{t^3}{\lfloor 3} + \ldots\ldots\ldots + \mu_r^1 \frac{t^r}{\lfloor r} + \ldots\ldots\ldots\right)$$

w.r.t. t, we get

$$K_1 + K_2 t + K_3 \frac{t^2}{\lfloor 2} + \ldots\ldots\ldots\ldots + K_r \frac{t^{r-1}}{(r-1)!} + \ldots\ldots\ldots\ldots$$

$$= \frac{\mu_1^1 + \mu_2^1 t + \mu_3^1 \frac{t^2}{\lfloor 2} + \ldots\ldots\ldots + \mu_r^1 \frac{t^{r-1}}{(r-1)!} + \ldots\ldots\ldots\ldots}{1 + \mu_1^1 t + \mu_2^1 \frac{t^2}{\lfloor 2} + \mu_3^1 \frac{t^3}{\lfloor 3} + \ldots\ldots\ldots + \mu_r^1 \frac{t^r}{\lfloor r} + \ldots\ldots\ldots}$$

$$\Rightarrow \quad \left[K_1 + K_2 t + K_3 \frac{t^2}{\lfloor 2} + \ldots\ldots\ldots + K_r \frac{t^{r-1}}{\lfloor r-1} + \ldots\ldots\right]\left[1 + \mu_1^1 t + \mu_2^1 \frac{t^2}{\lfloor 2} + \ldots\ldots\ldots + \mu_r^1 \frac{t^r}{\lfloor r} + \ldots\ldots\right]$$

$$= 1 + \mu_1^1 t + \mu_2^1 t + \mu_3^1 \frac{t^2}{\lfloor 2} + \ldots\ldots\ldots + \mu_r^1 \frac{t^{r-1}}{\lfloor r-1} + \ldots\ldots$$

Comparing the coefficients of $\frac{t^{r-1}}{\lfloor r-1}$ on both sides, we get

μ_r^1 = K$_1$. μ_{r-1}^1 + (r-1) K$_2$ μ_{r-2}^1 + $\binom{r-1}{2}$ K$_3$ μ_{r-3}^1 +.......+K$_r$ = $\binom{r-1}{0}$ μ_{r-1}^1 K$_1$ + $\binom{r-1}{1}$

μ_{r-2}^1 K$_2$ + $\binom{r-1}{2}$ μ_{r-3}^1 K$_3$ +.........+ $\binom{r-1}{r-1}$ μ_0^1 K$_r$

$$= \sum_{j=1}^{r} \binom{r-1}{j-1} \mu_{r-j}^1 K_j$$

Hence the required result

Example 2: Find the characteristic function of the random variable x whose probability density function is given as

$$f(x) = \begin{cases} e^{-x} & , \ x > 0 \\ 0 & , \ otherwise \end{cases}$$

and hence find mean and variance of X.

Sol. We know that

$$\phi_x(t) = \int_{-\infty}^{\infty} e^{itx} f(x) dx = \int_{0}^{\infty} e^{itx} . e^{-x} dx$$

$$= \int_{0}^{\infty} e^{-(1-it)^x} dx = \left[\frac{\bar{e}(1-it)^x}{-(1-it)}\right]_{x=0}^{\infty}$$

$= \phi_x(t) = (1 - it)^{-1}$

Now the rth moment about origin can be obtained by

$$E(x^r) = \mu_r^1 = \left[(-i)^r \frac{d^x}{dt^r} \phi_x(t)\right]_{t=0}$$

$$\therefore E(x) = \mu_1^1 = \left[(-i)^r \frac{d}{dt} \phi_x(t)\right]_{t=0}$$

$$= \left[(-i)^1 \frac{d}{dt} (1 - it)^{-1} \right]_{t=0}$$

$$= \left[(-i) \left[-(1 - it)^{-2} (-i) \right]_{t=0} \right] = 1$$

$\therefore$ E(x) = μ_1^1 = 1 $\Rightarrow$ Mean = 1

Also E(x^2) = $\mu_2^1 = \left[(-i)^2 \frac{d^2}{dt^2} \phi_x(t) \right]_{t=0}$

$$= \left[(-1) \frac{d^2}{dt^2} (1 - it)^{-1} \right]_{t=0}$$

$$= \left[(-1) \frac{d}{dt} \left[i(1 - it)^{-2} \right] \right]_{t=0}$$

$$= \left[(-1) \left[-2i(1 - it)^{-3} (-i) \right]_{t=0} \right]$$

= (-1) (-2) = 2

$\therefore$ E(x^2) = μ_2^1 = 2

$\therefore$ Variance = μ_2 = $\mu_2^1 - \mu_1^{1\ 2}$ = 2 - $(1)^2$ = 2 - 1 = 1

Example 3: Find the characteristic function of the random variable x having density function given by

$$f(x) = \begin{cases} 2^{\frac{-x}{a}} & , |x| < a \\ 0 & , \text{ otherwise} \end{cases}$$

Sol. The characteristic function is given by

$$E(e^{iwx}) = \int_{-\infty}^{\infty} e^{iwx} f(x) dx$$

$$= \frac{1}{2a} \int_{-a}^{a} e^{iwx} dx$$

$$= \frac{1}{2a} \left[\frac{e^{iwx}}{iw} \right]_{-a}^{a}$$

$$= \frac{e^{iaw} - e^{-iaw}}{2iaw} = \frac{\sin aw}{aw}$$

using Euler's formula with θ = aw.

Example 4: Find the density function $f(x)$ corresponding to the characteristic function

$$\phi(t) = \begin{cases} 1-|t|\ , & |t| \le 1 \\ 0\ , & |t| > 1 \end{cases}$$

Sol. We know that

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx}\phi(t)dt$$

$$= \frac{1}{2\pi} \int_{-1}^{1} e^{-itx}(1-|t|)dt$$

$$= \frac{1}{2\pi} \int_{-1}^{0} e^{-itx}(1+t)dt + \int_{0}^{1} e^{-itx}(1-t)dt \qquad ....(1)$$

$$= \frac{1}{2\pi}\left[\left\{(1+t)\frac{e^{-itx}}{-ix} - \left(\frac{e^{-itx}}{i^2x^2}\right)\right\}_{-1}^{0} + \left\{(1-t)\frac{e^{-itx}}{-ix} + \left(\frac{e^{-itx}}{i^2x^2}\right)\right\}_{0}^{1}\right]$$

$$= \frac{1}{2\pi}\left[\left\{\frac{1}{-ix} - \frac{1}{-ix^2}\right\} - \left\{0 - \frac{e^{ix}}{-x^2}\right\} + \left\{0 + \frac{e^{-ix}}{-x^2}\right\} - \left\{\frac{1}{-ix} + \frac{1}{-ix^2}\right\}\right]$$

$$= \frac{1}{2\pi}\left[-\frac{1}{-ix} + \frac{1}{x^2} - \frac{e^{ix}}{x^2} - \frac{e^{-ix}}{x^2} + \frac{1}{ix} + \frac{1}{x^2}\right]$$

$$= \frac{1}{2\pi}\left[\frac{2}{x^2} - \frac{e^{ix} + e^{-ix}}{x^2}\right]$$

$$= \frac{1}{\pi}\left(\frac{1-\cos x}{x^2}\right), -\infty \le x \le \infty.$$

Self Check Exercise

Q. 1 What is the characteristic function of random variable x which takes values -1 and 1 with probability $\frac{1}{2}$.

Q. 2 Find the characteristic function of the random variable having p.d.f. as f (x) = ce=a|x|,-∞ < x < ∞, where a > 0, and c is a suitable constant.

10.7 Summary

We conclude this unit by summarizing what we have covered in it:-

1. Defined cumulant generating function.
2. Derived the series expansion of cumulant generating function.
3. Discussed and proved properties of cumulants.
4. Defined characteristic function.
5. Derived the series expansion of the characteristic function
6. Discussed in detail different properties of characteristic function.
7. Some examples are given related to each topic so that the contents be clarified further.

10.8 Glossary:

1. Cumulant generating function is defined as :

   $K_x (t) = \log_e M_x (t) = \ln [M_x (t)]$

   provided the right hand side can be expanded as a convergent series in powers of t.

2. characteristic function is defined as $\phi_x (t) = E(e^{itx})$

$$= \begin{cases} \int e^{itx} f(x)dx & \text{(for continuous probability distribution)} \\ \sum_x e^{itx} f(x) & \text{(for discrete probability distribution)} \end{cases}$$

10.9 Answer to Self Check Exercise

Ans.1 $\phi_x (t)$ = cost

Ans.2 $E(e^{itx}) = \dfrac{a^2}{a^2 + t^2}$

10.10 References/Suggested Readings

1. Robert V. Hogg, Joseph W. Mckean and Allen T. Craig, Introduction to Mathematical Statistics, Pearson Education, Asia, 2007.
2. Irwin Miller, Marylees Miller and John E. Freund, Mathematical Statistics with Application, 7th Ed., Pearson Education, Asia, 2006.
3. Sheldon Ross, Introduction to Probability Model, 9th Ed., Academic Press, Indian Reprint, 2007

10.11 Terminal Questions

1. Show that

$$e^{itx} = 1 + (e^{it} - 1)\, x^{(1)} + (eit - 1)^2 \,.\, \frac{x^{(2)}}{\lfloor 2} + \ldots\ldots$$

$$+ (e^{it} - 1)^n + \ldots\ldots \frac{x^{(r)}}{\lfloor r}$$

where $x^{(r)} = x (x - 1) (x - 2) \ldots\ldots (x - r + 1)$

Hence show that

$$\mu_r^1 = [D^r \phi (t)]_{t=0}, \text{ where } D = \frac{d}{d(e^{it})}$$

and μ_r^1 is the r^{th} factorial moment.

2. Show that the characteristic function of Laplace distribution

$$f(x) = \frac{1}{2} e\text{-}|x|, \; -\infty < x < \infty$$

is $\phi_x (t) = \dfrac{1}{1+t^2}$

Find also the mean, the variance and mean deviation about the mean.

Unit - 11

Binomial Distribution

Structure

11.1 Introduction

Binomial distribution is a discrete probability distribution that describes the number of successes in a fixed number of independent Bernoulli trials, where each trial can result in either success or failure with a constant probability of success. In Binomial distribution (B. D.) the number of trials, denoted as n, is fixed and known in advance. Here each trial is independent of the others and has only two possible outcomes : success or failure. The probability of success in each trial of B.D., denoted as p, is constant and does not change throughout the trials.

The binomial distribution used in a variety of real0world scenarios. In manufacturing, the binomial distribution can be used to model the number of defective items produced in a fixed number of units inspected e.g. monitoring the number of defective light bulbs in a batch of 100 light bulbs produced. In medical research, the binomial distribution is used to model the number of patients who experience a certain outcome (e.g. recovery, adverse effect) in a clinical trial with a fixed number of participants. When conducting surveys, the binomial distribution can be used to model the number of respondents who answer a particular question in a certain way, given a fixed sample size e.g. estimating the proportion of voters who support a particular candidate in a few examples of the many real-world applications of the binomial distribution. The key is that the situation involves a fixed number of independent trials, each with a constnat probability of success or failure.

11.2 Learning Objectives:

After studying this unit, students will be able to:

- Define discrete uniform distribution.
- Discuss properties of discrete uniform distribution
- Define Bernoulli Variate and Bernoulli distribution.
- Discuss probability density function of Bernoulli distribution
- Discuss moments and moment generating function of a Bernoulli Variate.
- Define and discuss Binomial distribution
- Find the mean and Variance of Binomial distribution.
- Discuss moment generating function of Binomial distribution.
- Discuss moments, central moments and mode of binomial distribution.
- Define standard binomial Variate and able to discuss theorem related to it.

11.3 Discrete Uniform Distribution and Its Properties

A random variable x is said to have a discrete uniform distribution over the range [1, n] if its probability mass function is expressed as:

$$P(X = x) = \begin{cases} \frac{1}{n} \; for x = 1, 2, ..., n \\ 0 \;\; otherwise \end{cases} \qquad .....(1)$$

Here n is known as the parameter of the distribution and lies in the set of all positive integers. The given distribution is also called a discrete rectangular distribution.

Such distributions can be conceived in practice of under the given experimental conditions, the different values of the random variable becomes equally likely. Thus for a dice experiment and for an experiment with a deck of cards such distribution is appropriate.

Properties:

(i) $E(X) = \frac{1}{n}\sum_{x=1}^{n} x = \frac{n+1}{2}$

$E(X^2) = \frac{1}{n}\sum_{x=1}^{n} x^2 = \frac{(n+1)(2n+1}{6}$

(ii) $V(X) = E(X^2) - [E(X)]^2 = \frac{(n+1)(n-1}{12}$

(iii) The moment generating function of X is :

$$M_X(t) = E(e^{tX}) = \frac{1}{n}\sum_{x=1}^{n} e^{tx} = \frac{e^t(1-e^{nt})}{n(1-e^t)}$$

11.4 Bernoulli Random Variable

A random variable X which takes two values 0 and 1, with probabilities q and p respectively, i.e., P (X = 1) = p, P(X = 0) = q, where q = 1 - p is called a Bernoulli variate and is said to have a Bernoulli distribution.

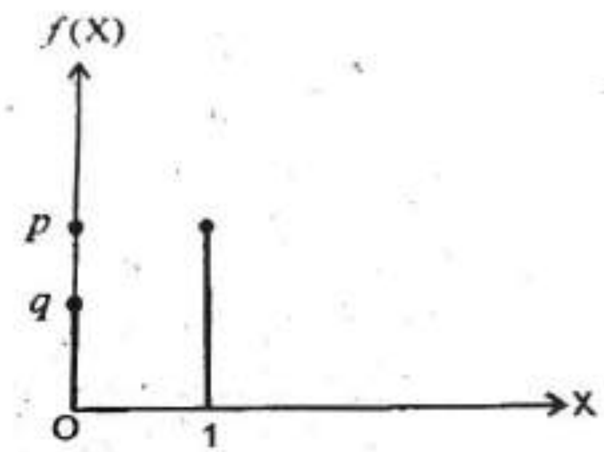
It is a discrete distribution and is usually written as X ⊔ B (1,p) so that

$$\left.\begin{aligned} &P(X=1)=p \\ &P(X=0)=q=1-p \\ &\text{where } 0<p<1 \end{aligned}\right\} \quad \Rightarrow$$

x	P(X = x)
1	p
0	1-p

P.D.F of a Bernoulli Distribution

The p.d.f. of X can be written as



$$f(\mathrm{x}) = \begin{cases} 1-p, & x=0 \\ p, & x=1 \\ 0\ , & \text{elsewhwere} \end{cases}$$

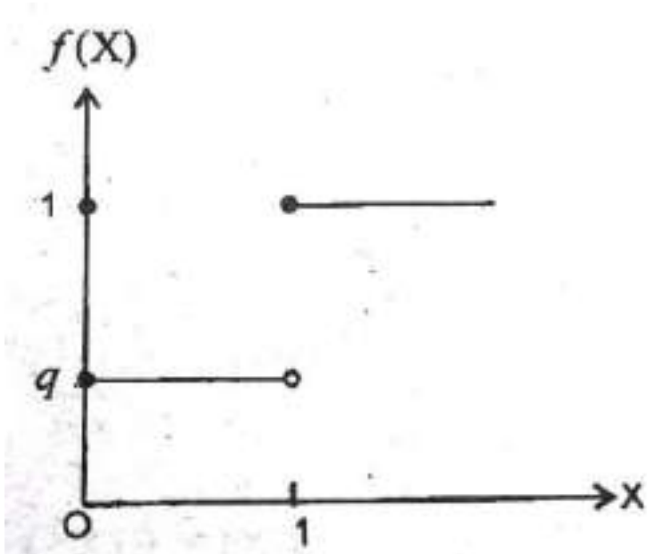
$$= \begin{cases} p^x (1-p)^{1-x}, & x=0,1 \\ 0 & , \quad \text{elsewhwere} \end{cases}$$

$$= \begin{cases} p^x\ q^{1-x}, & x=0,1 \\ 0 & , \quad \text{elsewhwere} \end{cases}$$

where q = 1 - p.

The graph of p.d.f. of Bernoulli's distribution is shown in the figure.

C.D.F. of a Bernoulli Distribution



The cumulative distribution function of Bernoulli's distribution is given by

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

When x < 0, F(X) = P(X < 0) - 0,

When 0$\leq$ x < 1, F(X) = P(X = 0) = q.

When x $\geq$ 1. F(X) = P(X = 0) + P(X = 1)

= q + p = 1

$$\therefore \quad F(X) = \begin{cases} 0, & for \quad x<0 \\ q, & for\ 0 \leq x < 1 \\ 1, & for\ x \geq 1 \end{cases}$$

The graph c.d. f. F(x) is shown in figure.

Moments and Moment Generating Function of a Bernoulli Variate

The rth moment about origin is

$\mu_r' = E(X^r) = \sum x^r (P(X = r) = 0r \times (1 - p) + 1^r \times P = p,$

where r = 1, 2, 3, 4

$\therefore \quad \boxed{\mu_r' = \text{Mean} = p}$ ......(1)

$\mu_2' = 0^2 \times (1 - p) + t^2 \times p = p$

$\mu_3' = 0^3 \times (1 - p) + t^3 \times p = p$

$\mu_4' = 0^4 \times (1 - p) + 1^4 \times p = p$

and so on.

Variance = $\mu_2' - \mu_1'^2 = p - p^2 = (1 - p) = pq$ [$\because$ q = 1 - p]

$\Rightarrow \quad \boxed{\text{Variance} = pq}$

$\mu_3 = \mu_3' - 3\mu_2' \mu_1' + 2 \mu_1'^3 = p - 3 p.p + 2p^3 = 2p^3 - 3p^2 + p$

= p (p2 - 3 p + 1)

$\mu_4 = \mu_4' - 4\mu_3' \mu_1' + 6\mu_2' \mu_1'^2 - 3\mu_1'^4 = p - 4p^2 + 6p^3 - 3p^4$

The moment generating function about origin is

$M_x(t) = E(e^{tX}) = \sum e^{tX} .P(X = x) = e^{t0} (1 - p) + e^{t1} p = 1 - p + e^t p$

$\boxed{M_X(t) = 1 + p(e^t - 1) = q + pe^t}$ ....(2)

11.5 Binomial Distribution

Binomial distribution was discovered by James Bernoulli (1654 - 1705) in the year 1700 and was first published posthumously in 1713, (eight years after his death). Let a random experiment be performed repeatedly and let the occurrence of an event in a trial be called a success and its non-occurrence a failure. Consider a set of n independent. Bernoullian trials (n being finite), in which the probability 'p' of success in any trial is constant for each trial. Then q = 1 - p, is probability of failure in any trial.

The probability of x successes and consequently (n - x) failures in n independent trials, in a specified order (say) SSFSFFFS......FSF (where S represents success and F failure) is given by the compound probability theorem by the expression:

P (SSFSFFS........FSF) = P(S)P(S)P(F)P(S)P(F)P(F)P(F)P(S) $\times$.......$\times$ P(F)P(S)P(F)

= p . p . q . p . q . q . q . p ....q . p . q

= (p . p ....to x factors) (q . q ...... to $\overline{n-x}$ factors) = $p^x q^{n-x}$

But x successes in n trials can occur in $\binom{n}{x}$ ways and the probability for each of these ways is $p^x q^{n-x}$. Hence the probability of x successes in n trials in any order whatsoever is given by the addition theorem of probability by the expression:

$$\binom{n}{x} p^x q^{n-x}$$

The probability distribution of the number of successes, so obtained is called the Binomial probability distribution, the probabilities of 0, 1, 2, ..., n successes, i.e., q^n, $\binom{n}{1}$ qn-1 p, $\binom{n}{2} q^{n-2} p^2$,......, p^n, are the successive terms of the binomial expansion $(q + p)^n$.

Definition. A random variable X is said to have binomial distribution if it assumes only non-negative values and its probability mass function is given by

$$P(X = x) = p(x) = \begin{cases} \binom{n}{x} p^x q^{n-x}; \ x = 0, 1, 2, ..., n; q = 1 - p \\ 0 \qquad\qquad otherwise \end{cases}$$

The two independent constants n and p in the distribution are known as the parameters of the distribution 'n' is also, sometimes, known as the degree of the binomial distribution.

Binomial distribution is a discrete distribution as X can take only the integral values, i.e. 0, 1, 2,....., n. Any variable which follows binomial distribution is known as binomial variate.

The probability p(x) is also sometimes denoted by b(x ; n, p) or f(x ; n, p) or f(x).

Let us consider the following examples to clear the idea:-

Example 1: If, on the average, 1 ship in every 10 is sunk, find the chance that out of 5 ships expected, 4 at least will arrive safely.

Sol. Let p be the probability that a vessel will arrive safely.

$\therefore \quad p = \frac{9}{10} = 0.9$

$\therefore \quad q = 1 - p = 1 - 0.9 = 0.1$

Now n = 5, p = 0.9, q = 0.1

Let X be the number of ships reaching safely.

The probability that out of 5 ships, x ships will arrive safely is

$f(x) = b(x ; 5, 0.9 = \binom{5}{x} (0.9)^x (0.1)^{5-x}$, x = 0, 1, 2, 3, 4, 5

Required probability = $P(X \geq 4)$

$= P(X = 4) + P(X = 5)$

$= f(4) + f(5)$

$= \binom{5}{4} (0.9)^4 (0.1)1 + \binom{5}{5} (0.9)^5$

$= \binom{5}{1} (0.9)^4 + 1(0.9)^5 = \frac{5}{1} \times (0.9)^4 + 1 \times (0.9)^5$

$= (0.9)^4 [5 + 0.9] = (0.9)^4 (5.9) = (5.9) (0.9)^4$

Example 2: From a lot containing 20 items, five of which are defective, four items are drawn with replacement. What is the probability of getting.

(i) exactly one defective item?

(ii) at least one defective item?

Sol. Let X be the random variable denoting the number of defective items drawn. Then the possible values of X are 0, 1, 2, 3, 4

Let p be the probability of defective item drawn.

$\therefore$ $p = \frac{5}{20} = \frac{1}{4}$

and $q = 1 - p = 1 - \frac{1}{4} = \frac{3}{4}$

Now $n = 4, p = \frac{1}{4}, q = \frac{3}{4}$

$\therefore$ $f(x) = b(x ; n, p) = (x; 4 \frac{1}{4})$

$= \binom{4}{x} \left(\frac{1}{4}\right)^x \left(\frac{3}{4}\right)^{4-x}$, x = 0, 1,2,3,4

(i) Probability of getting exactly one defective item

$P(X = 1) = f(1)$

$$= \binom{4}{1}\left(\frac{1}{4}\right)^{1}\left(\frac{3}{4}\right)^{3}$$

$$= \frac{4}{1} \times \frac{1}{4} \times \frac{27}{64}$$

$$= \frac{27}{64}$$

(ii) Probability of getting at least one defective item = P(X $\geq$ 1)

= P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4)

= f(1) + f(2) + f(3) + f(4)

= 1 - f(0)

$$= 1 - \binom{4}{0}\left(\frac{1}{4}\right)^{0}\left(\frac{3}{4}\right)^{4}$$

$$= 1 - x \times 1 \times \frac{81}{256}$$

$$= \frac{175}{256}$$

Example 3: If m things are distributed among 'a' men and 'b' women. Find the chance that the number of things received by men is odd.

Sol. Probability that the man gets a thing = $\frac{a}{a+b}$ = p. Probability that the woman gets a thing = $\frac{b}{a+b}$ = q Probability that r things are received by men = P(r) = m_{c_r} p^r q^{m-r}.

Since men are to receive odd number of things i.e. 1 or 3 or 5 or....., there required probability is = P(1) + P(3) + P(5) +........

$$= m_{c_1}\ p.q^{m-1} + m_{c_3}\ p^3\ q^{m-3} + m_{c_5}\ p^5\ q^{m-5} + \text{...........}$$

$$= \frac{1}{2}\left\{(p+q)^m - (q-p)^m\right\} \qquad \left[\because p+q=1 and q-p=\frac{b-a}{a+b}\right]$$

$$= \frac{1}{2}\left\{1-\left(\frac{b-a}{a+b}\right)^m\right\}$$

Example 4: A pair of dice is thrown 4 times. If a doublet is considered a success, find the probability of 2 successes.

Sol. Here a pair of dice is thrown 4 times.

$\therefore$ $n = 4$

Let p be the probability that doublet is obtained.

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

$= \frac{6}{36} = \frac{1}{6}$

and $q = 1 - p = 1 = \frac{1}{6} = \frac{5}{6}$

$\therefore$ Required probability = P(2) = $4_{c_2}\left(\frac{5}{6}\right)^2\left(\frac{1}{6}\right)^2$

$= \frac{4\times3}{1\times2} \times \frac{25}{36} \times \frac{1}{36}$

$= \frac{25}{216}$

Example 5: A coin is tossed 5 times. What is the probability that head appears an odd number of times?

Sol. Here a coin is tossed 5 times.

$\therefore$ $n = 5$

$\therefore$ p = P (head in one toss) = $\frac{1}{2}$

and $q = 1 - p = 1 = \frac{1}{2} = \frac{1}{2}$

Now P (head appearing odd number of times)

= P(r = 1, 3, or 5)

= P(1) + P(3) + P(5)

$= 5_{c_1}\left(\frac{1}{2}\right)^1\left(\frac{1}{2}\right)^{5-1} + 5_{c_3}\left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right)^{5-3} + 5_{c_5}\left(\frac{1}{2}\right)^0\left(\frac{1}{2}\right)^5$

$= \frac{5}{1} \times \frac{1}{16} \times \frac{1}{2} + \frac{5\times4\times3}{1\times2\times3} \times \frac{1}{4} \times \frac{1}{8} + 1 \times 1 \times \frac{1}{32}$

$= \frac{1}{32}(5 + 10 + 1)$

$= \frac{16}{32} = \frac{1}{2}$

Example 6: There are 5 percent defective items in a large bulk of items. What is the probability that a sample of 10 items will include not more than one defective item?

Sol. Here $p = \frac{5}{100} = \frac{1}{20}$,

$q = 1 - p = 1 - \frac{1}{20} = \frac{19}{20}$

and $n = 10$

Now Required probability = P(no defective) +P (1 defective)

$$= 10_{c_0}\left(\frac{19}{20}\right)^{10} + 10_{c_1}\left(\frac{19}{20}\right)^{9}\left(\frac{1}{20}\right)^{1}$$

$$= 1 \times \left(\frac{19}{20}\right)^{10} + \frac{10}{1} \times \left(\frac{19}{20}\right)^{9} \times \frac{1}{20}$$

$$= \left(\frac{19}{20}\right)^{9}\left[\frac{19}{20}+\frac{10}{20}\right]$$

$$= \left(\frac{19}{20}\right)^{9} \times \frac{29}{20}$$

Example 7: Fire cards are drawn successively with replacement from a well shuffled deck of 52 cards. What is the probability that

(a) all the five cards are spades?

(b) only three cards are spades?

(c) none is spade?

Sol. Here $p = P(\text{spade}) = \frac{13}{52} = \frac{1}{4}$,

$q = 1 - p = 1 - \frac{1}{4} = \frac{3}{4}$,

and $n = 5$

(a) (5 spades) = $5_{c_5}\left(\frac{1}{4}\right)^{5} = 1 \times \frac{1}{1024} = \frac{1}{1024}$

(b) P (3 spades) = $5_{c_3} \left(\frac{1}{4}\right)^3 \left(\frac{3}{4}\right)^2$ = $5_{c_2} \times \frac{1}{64} \times \frac{9}{16}$

$= \frac{5\times 4}{1\times 2} \times \frac{1}{64} \times \frac{9}{16} = \frac{45}{512}$

(c) P (none is spade) = P(0) = $5_{c_0} \left(\frac{3}{4}\right)^5 = 1 \times \frac{243}{1024}$

$= \frac{243}{1024}$

Example 8: A pair of dice is thrown 7 times. If getting a total of 7 is considered a success, what is the probability of

(a) no success? (b) 6 successes?

(c) at least 6 successes? (d) at the most 6 successes?

Sol. Here n = 7

Let p be the probability of getting a total of 7.

$\therefore$ p = $\frac{3}{36} = \frac{1}{6}$

and q = $1 - \frac{1}{6} = \frac{5}{6}$

(a) P (no success) = P(0) = $7_{c_0} \left(\frac{1}{6}\right)^0 \left(\frac{5}{6}\right)^{7-0}$

$= \left(\frac{5}{6}\right)^7$

(b) P (6 successes) = P(6) = $7_{c_6} \left(\frac{1}{6}\right)^6 \left(\frac{5}{6}\right)^{7-6}$

$= \frac{7}{1} \times \frac{5}{6} \times \left(\frac{1}{6}\right)^6$

$= 36 \left(\frac{1}{6}\right)^7$

(c) P (at least 6 successes) = P(6) + P(7)

$$= 7_{c_6}\left(\frac{1}{6}\right)^6\left(\frac{5}{6}\right)^{7-6} + 7_{c_7}\left(\frac{1}{6}\right)^7\left(\frac{5}{6}\right)^{7-7}$$

$$= 7 \times \frac{5}{6} \times \left(\frac{1}{6}\right)^6 + 1 \times 1 \times \left(\frac{1}{6}\right)^7$$

$$= 36 \times \left(\frac{1}{6}\right)^7$$

$$= \left(\frac{1}{6}\right)^5$$

(d) P (at the most 6 successes) = P (not 7 success)

$= 1 - P(7)$

$$= 1 - 7_{c_7}\left(\frac{1}{6}\right)^7\left(\frac{5}{6}\right)^{7-7}$$

$$= 1 - \left(\frac{1}{6}\right)^7$$

Example 9: The probability that a bulb produced by a factory will fuse after 150 days of use is 0.05. Find the probability that out of 5 such bulbs.

(a) none (b) not more than one

(c) more than one (d) at least one

will fuse after 150 days of use.

Sol. Here

$$p = 0.05 = \frac{5}{100} = \frac{1}{20},$$

$$q = 1 - p = 1 - \frac{1}{20} = \frac{19}{20},$$

and $n = 5$

(a) P (none will fuse after 150 days)

$$= P(0) = 5_{c_0}\left(\frac{1}{20}\right)^0\left(\frac{19}{20}\right)^{5-0}$$

$$= 1 \times \left(\frac{19}{20}\right)^5 \times 1 = \left(\frac{19}{20}\right)^5$$

(b) P (not more than one bulb will fuse after 150 days)

$$= P(0) + P(1)$$

$$= 5_{c_0}\left(\frac{1}{20}\right)^0\left(\frac{19}{20}\right)^{5-0} + 5_{c_1}\left(\frac{1}{20}\right)^1\left(\frac{19}{20}\right)^{5-1}$$

$$= 1 \times \left(\frac{19}{20}\right)^5 + \frac{5}{1} \times \frac{1}{20} \times \left(\frac{19}{20}\right)^4$$

$$= \left(\frac{19}{20}\right)^4\left[\frac{19}{20}+\frac{5}{20}\right] = \left(\frac{19}{20}\right)^4 \times \frac{24}{20}$$

$$= \frac{6}{5}\left(\frac{19}{20}\right)^4$$

(c) P(more than one bulb will fuse after 150 days)

$$= 1 - [P(0) + P(1)]$$

$$= 1 - \frac{6}{5}\left(\frac{19}{20}\right)^4 \qquad [\because \text{ of (b)}]$$

(d) P (at least one bulb will fuse after 150 days)

$$= 1 - P(0) = 1 - \left(\frac{19}{20}\right)^5 \qquad [\because \text{ of (a)}]$$

Self Check Exercise-1

Q.1 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 numbers divisible by 10.

Q.2 In a family of five children, what is the probability that there will be exactly two boys, assuming that the sexes are equally likely?

Q.3 The probability of a man hitting a target is $\frac{1}{4}$. He fires 7 times. What is the probability of his hitting at least twice the target?

Q.4 A bag contains 5 white, 7 red and 8 black balls. If four balls are drawn one by one with replacement, what is the probability that

(i) none is white? (ii) all are white?

(iii) only 2 are white? (iv) at least one is white?

11.6 Mean and Variance of Binomial Distribution

Let X be the random variable of the binomial

$$f(x) = \binom{n}{x} p^x q^{n-x}, x = 0, 1, 2, \ldots\ldots n$$

$$\text{Mean} = E(X) = \sum_{x=0}^{n} x\, f(x)$$

$$= \sum_{x=0}^{n} x \binom{n}{x} px\, q^{n-x}$$

$$= 0. \binom{n}{0} p^0 q^n + 1. \binom{n}{1} p^1 q^{n-1} + 2. \binom{n}{2} p^2 q^{n-2} + \ldots\ldots + n. \binom{n}{n} p^n q^0$$

$$= 0 + n\, q^{n-1} + 2\, \frac{n(n-1)}{1.2} p^2 q^{n-2} + 3\, \frac{n(n-1)(n-2)}{1.2.3} p^3 q^{n-3} + \ldots\ldots + np^n$$

$$= n\, p \left[q^{n-1} + (n-1)\, p\, q^{n-2} + \frac{(n-1)(n-2)}{1.2} p^2 q^{n-2} + \ldots\ldots.. + p^{n-1} \right]$$

$$= n\, p \left[q^{n-1} + {}^{(n-1)}C_1 p\, q^{n-2} + {}^{n-1}C_2 p^2 q^{n-2} + \ldots\ldots.. + {}^{n-1}C_{n-1} p^{n-1} \right]$$

$= n\, p\, (q + p)^{n-1}$ $[\because\ p + q = 1]$

$= n\, p\, (1)^{n-1}$

$= n\, p\, (1) = np$

$\therefore \quad \mu = n\, p$

Variance $= E(X - \mu)^2$

$= E(X^2) - [E(X)]^2$

$$= \sum_{x=0}^{n} x^2 f(x) - \mu^2 = \sum_{x=0}^{n} [x + x(x-1).\binom{n}{x} p^x.\, q^{n-x} - \mu^2$$

$$= \sum_{x=0}^{n} x\, .\, {}^nC_x\, p^x q^{n-x} + \sum_{x=0}^{n} x\, (x - 1)\, {}^nC^x\, p^x q^{n-x} - \mu^2$$

$= n\, p + 2.1\, {}^nC_2\, p^2 q^{n-2} + 3.2\, {}^nC_3\, p^3 q^{n-3} + \ldots\ldots.. + n\, (n - 1)\, p^n - \mu^2$

$$= n\, p + 2 \times 1 \times \frac{n(n-1)}{1 \times 2} p^2 q^{n-2}$$

$$+ 3 \times 2 \times \frac{n(n-1)(n-2)}{1\times 2\times 3} p^3 q^{n-3} + ... + n(n-1)p^n - \mu^2$$

$$= np + n(n-1)p^2 [q^{n-2} + (n-2)q^{n-3} p + ... + p^{n-2}] - \mu^2$$

$$= np + n(n-1)p^2 [q^{n-2} + {}^{n-2}C_1 q^{n-3} p + ... + {}^{n-2}C_{n-2} p^{n-2}] - \mu^2$$

$$= np + n(n-1)p^2 (q+p)^{n-2} - \mu^2$$

$$= np + (n^2 - n)p^2 (1)^{n-2} - \mu^2 \qquad [\because p + q = 1]$$

$$= np + (n^2 - n)p^2 - (np)^2 \qquad [\because \mu = np]$$

$$= np + n^2p^2 - np^2 - n^2p^2 = np - np^2 = np(1-p)$$

$\therefore$ Variance = n p q $\qquad [\because q = 1 + p]$

11.7 Moment Generating Function of Binomial Distribution

$M_x(t) = E(e^{tx})$, where t is real

$$= \sum_{x=0}^{n} e^{tx} \binom{n}{x} p^x q^{n-x}$$

$$= \sum_{x=0}^{n} e^{tx} \binom{n}{x} (pe^t)^x q^{n-x}$$

$$= \binom{n}{0} q^n + \binom{n}{1} (pe^t) q^{n-1} + \binom{n}{2} (pe^t)^2 q^{n-2} + ... + \binom{n}{n} (pe^t)^n$$

$$= (q + pe^t)^n$$

11.8 Moments

The first four moments about origin of binomial distribution are obtained as follows:

$$\mu'_1 = E(X) = \sum_{x=0}^{n} x \binom{n}{x} p^x q^{n-x} = np \sum_{x=0}^{n} \binom{n-1}{x-1} p^{x-1} q^{n-x}$$

$$= np(q+p)^{n-1} = np(1)^{n-1} = np$$

$$\mu'_2 = E(X^2) = \sum_{x=0}^{n} x^2 \binom{n}{x} p^x q^{n-x}$$

$$= \sum_{x=0}^{n} \{x(x-1) + x\} \frac{n(n-1)}{x(x-1)} \binom{n-2}{x-2} p^x q^{n-x}$$

$$= n(n-1)p^2 \left[\sum_{x=0}^{n} \binom{n-2}{x-2} p^{x-2} q^{n-x} \right] + np$$

$= n (n - 1) p^2 (q + p)^{n\ 2} + n p = n(n - 1) p^2 + n p$

$$\mu'_3 = E(X^3) = \sum_{x=0}^{n} x^3 \binom{n}{x} p^x q^{n-x}$$

$$= \sum_{x=0}^{n} \{x(x-1)(x-2) + 3x(x-1) + x\} p^x q^{n-x}$$

$$= n (n - 1) (n - 2) p^3 \sum_{x=0}^{n} \binom{n-3}{x-3} p^{x-3} q^{n-x}$$

$$+ 3n (n - 1) p^2 \sum_{x=0}^{n} \binom{n-2}{x-2} p^{x-2} q^{n-x} + n p$$

$= n (n - 1) (n - 2) p^3 (q + p)^{n-3} + 3n (n - 1)p^2 (q + p)^{n-2} + n p$

$= n (n - 1) (n - 2) p^3 + 3 n (n - 1) p^2 + n p$

Similarly

$x^4 = x (x - 1) (x - 2) (x - 3) + 6x (x -1) (x - 2) + 7x (x - 1) + x$

(Let $x^4 = Ax (x - 1) (x - 2) (x - 3) + Bx (x - 1) (x - 2) + Cx (x 1) + x$

By giving to x the values 1, 2 and 3 respectively, we find the values of arbitrary constant A, B and C]

$$\mu'_4 = E(X^4) = \sum_{x=0}^{n} x^4 \binom{n}{x} p^x q^{n-x}$$

$= n (n - 1) (n - 2) (n - 3) p^4 + 6n (n - 1) (n - 2) p^3 + 7 n (n - 1) p^2 + n p$

Central Moments of Binomial Distribution

$\mu_2 = \mu_2^1 - \mu_1^{1\ 2} = n^2 p^2 - np^2 + np - n^2 p^2 = np (1 - p) = n p q$

$\mu_3 = \mu_3^{1\ 1} - 3 \mu_2^1 \mu_1^1 + 2 \mu_1^{1\ 3}$

$= \{n (n - 1) (n - 2) p^3 + 3n (n - 1) p^2 + n p\} - 3 \{n (n - 1) p^2 + p\} np + 2 (np)^3$

$= np [- 3np^2 + 3np + 2p^2 - 3p + 1 - 3npq]$

$= np [3np (1 - p) + 2p^2 - 3p + 1 - 3npq]$

$= np [2p^2 - 3p + 1] = np (2p^2 - 2p + q)$

$= npq (1 - 2p)$

$= npq (q + p - 2p)$

$= npq (q - n)$

$\mu_4 = \mu_4^1 - 4 \mu_3^1 \mu_1^1 + 6 \mu_2^1 \mu_1^{1\ 2} \mu_1^{1\ 2} - 3 \mu_1^{1\ 4}$

= npq [1 + 3 (n - 2) pq]

Now

$$\beta_1 = \frac{\mu_3^{1\,2}}{\mu_2^{1\,3}} = \frac{n^2p^2q^2(q-p)^2}{n^3p^3q^3} = \frac{(q-p)^2}{npq} = \frac{(1-2p)^2}{npq}$$

$$\beta_2 = \frac{\mu_4}{\mu_2^{2\ 3}} = \frac{npq\{1+3(x-2)pq}{n^2p^2q^2}$$

$$= \frac{1+3(x-2)pq}{npq} = 3 + \frac{1-6pq}{npq}$$

$$V_1 = \sqrt{\beta_1} = \frac{q-p}{\sqrt{npq}} = \frac{1-2p}{\sqrt{npq}},\ V_2 = \beta_2 - 3 = \frac{1-6pq}{npq}$$

11.9 Mode of the Distribution

Let random variable x has a discrete distribution for which the probability function is f(x), then the value of x for which f(x) is maximum is called the mode of the distribution. If same maximum value of x is attained at more than one value of x, then all such values of x are called modes of the distribution.

Mode of Binomial Distribution

Suppose that the mode of the binomial distribution lies at X = r.

$\therefore$ P(X = r - 1) $\leq$ P (X = r) $\geq$ P (X = r + 1)

i.e. $n_{c_{r-1}}\, p^{r-1}\, q^{n-r+1} \leq n_{c_r}\, p^r\, q^{n-r} \geq n_{c_{r+1}}\, p^{r+1}\, q^{n-r-1}$

i.e. $$\frac{\lfloor n}{\lfloor r-1\ \lfloor n-r+1}\, p^{r-1}\, q^{n-r+1} \leq \frac{\lfloor n}{\lfloor r\ \lfloor n-r}\, p^r\, q^{n-r} \geq \frac{\lfloor n}{\lfloor r+1\ \lfloor n-r-1}\, p^{r+1}\, q^{n-r-1}$$

i.e. $$\frac{r}{n-r+1}\ \frac{p}{q} \leq 1 \geq \frac{n-r}{r+1}\ \frac{p}{q} \qquad ....(1)$$

Now $\frac{r}{n-r+1}\ \frac{q}{p} \leq 1$

$\Rightarrow$ r q < n p - p r + p

$\Rightarrow$ (p + q) r $\leq$ (n + 1) p

$\Rightarrow$ r $\leq$ np + p ......(2)

Also $1 \geq \frac{n-r}{r+1}\ \frac{p}{q}$

$\Rightarrow$ rq + q $\geq$ np - pr

$\Rightarrow$ (p + q) r $\geq$ np - q

$\Rightarrow$ r $\geq$ np - q ......(3)

From (1), (2), (3), we get

(n + 1) p - 1 $\leq$ r $\leq$ (n + 1) p

Now two cases arise:-

Case I:- (n + 1) p = k, where k is an integer

$\therefore$ there are two modes and the distribution is said to be Bimodal.

Two modes are at points np - q and np + p

Case II:- (n + 1) p = k + f, where k is an integer and f is a proper fraction.

$\therefore$ x = k = [(n + 1) p] = [np + p]

Where [np + p] mean greatest integer $\leq$ (np + p)

11.10 Standard Binomial Variate

Let a random variable X have binomial distribution with mean np and variance npq. Then the random variable $\geq$ = $\frac{X-np}{\sqrt{npq}}$ having binomial distribution with mean 0 and variance 1 is called standard binomial variate.

Theorem:- If a random variable X has binomial distribution with mean np and variance npq; prove that the random variable Z defined by > = $\frac{X-np}{\sqrt{npq}}$ has a binomial variate with mean 0 and variance 1.

Proof: Here Z = $\frac{X-np}{\sqrt{npq}}$

$$E(Z) = E\left(\frac{X-np}{\sqrt{npq}}\right)$$

$$= \frac{1}{\sqrt{npq}} E(X - np) \qquad [\because E(a x) = a E(X)]$$

$$= \frac{1}{\sqrt{npq}} [E(X) - np E(1)]$$

$$= \frac{1}{\sqrt{npq}} [np - np (1)] \qquad [\because E(X) = np, E(1) = 1]$$

$= \frac{1}{\sqrt{npq}}$ (np - np) = $\frac{1}{\sqrt{npq}}$ (0)

$\therefore$ E(Z) = 0

$\therefore$ Mean of Z is 0

Var (Z) = Var $\left[\frac{X-np}{\sqrt{npq}}\right]$

$= \frac{1}{\sqrt{npq}}$ Var (X - np) [$\because$ Var (aX) = a^2 Var(X)]

$= \frac{1}{\sqrt{npq}}$ Var (X) [$\because$ Var (X + a) = Var (X)]

$= \frac{1}{npq}$ npq = 1

$\therefore$ Variance of Z is 1

Let us improve our understanding of these results by looking at some of the following examples:-

Example 10: The mean and variance of a binomial distribution are 4 and $\frac{4}{3}$ respectively. Find P(X $\geq$ 1)

Sol. We have

np = 4 .... (1)

and npq = $\frac{4}{3}$ ....(2)

Dividing (2) by (1), we get

q = $\frac{1}{3}$

$\therefore$ p = 1 - q = 1 - $\frac{1}{3}$ = $\frac{2}{3}$

$\therefore$ from (1), n$\left(\frac{2}{3}\right)$ = 4

$\therefore$ n = 6

Now P(X $\geq$ 1) = 1 - P(X = 0) = 1 - q^n

$= 1 = \left(\frac{1}{3}\right)^6$

$= 1 - \frac{1}{729} = \frac{728}{729}$

Example 11: The m.g.f. of a random vairable X is $\left(\frac{2}{3}+\frac{1}{3}e^t\right)^9$. Show that

$$P(\mu - 2\sigma < X < \mu + 2\sigma) = \sum_{x=1}^{5} \binom{9}{x}\left(\frac{1}{3}\right)^x\left(\frac{2}{3}\right)^{9-x}$$

Sol. We have

$M_x(t) = \left(\frac{2}{3}+\frac{1}{3}e^t\right)^9 = (q + pe^t)^n$

$\therefore$ n = 9, q = $\frac{2}{3}$, p = $\frac{1}{3}$

$\mu = np = 9 \times \frac{1}{3} = 3$

$\sigma^2 = npq = 9 \times \frac{1}{3} \times \frac{2}{3} = 2$

$\Rightarrow$ $\sigma = \sqrt{2}$

$\mu - 2\sigma = 3 - 2\sqrt{2} = 3 - 2 \times 1.4 = 3 - 2.8 = 0.2$

and $\mu + 2\sigma = 3 + 2\sqrt{2} = 3 + 2 \times 1.4 = 3 + 2.8 = 5.8$

$\therefore$ $P(\mu - 2\sigma < x < \mu + 2\sigma) = p(0.2 < X < 5.8)$

$= P(1 \leq X \leq 5)$

$= \sum_{x=1}^{5} p(x) = \sum_{x=1}^{5} n_{c_x} p^x q^{n-x}$

$= \sum_{x=1}^{5} 9_{c_x}\left(\frac{1}{3}\right)^x\left(\frac{2}{3}\right)^{9-x}$

Example 12: If X □ B (n, p), show that

$$E\left(\frac{x}{n}-p\right)^2 = \frac{pq}{n} \text{ ; Cov } \left(\frac{x}{n}, \frac{n-x}{n}\right) = -\frac{pq}{n}$$

Sol. Since X □ B (n, p), E(X) = np

and Var (X) = npq

$\therefore$ $E\left(\frac{x}{n}\right) = \frac{1}{n} E(X) = \frac{1}{n} . np = p$

$$\text{Var}\left(\frac{x}{n}\right) = \frac{1}{n^2} \text{ Var (X)} = \frac{pq}{n} \qquad .....(1)$$

(i) $E\left(\frac{x}{n}-p\right)^2 = E\left[\frac{x}{n}-E\left(\frac{x}{n}\right)\right]^2 = \text{Var}\left(\frac{x}{n}\right) = \frac{pq}{n}$ [From (1)

(ii) Using Cov (X, Y) = E [X - E (X)] [Y = E(Y)]

$$\text{Cov.}\left[\frac{x}{n}, \frac{n-x}{n}\right] = E\left[\left\{\frac{x}{n}-E\left(\frac{x}{n}\right)\right\}.\left\{\frac{n-x}{n}-E\left(\frac{n-x}{n}\right)\right\}\right]$$

$$= E\left[\left\{\frac{x}{n}-E\left(\frac{x}{n}\right)\right\}\left\{\frac{n-x}{n}-E\left(\frac{n-x}{n}\right)\right\}\right]$$

$$= E\left[\left(\frac{x}{n}-p\right)\left(1-\frac{x}{n}-(1-p)\right)\right]$$

$$= E\left[\left(\frac{x}{n}-p\right)\left\{\left(-\frac{x}{n}-p\right)\right\}\right]$$

$$= -E\left(\frac{x}{n}-p\right)^2$$

$$= -\text{Var}\left(\frac{x}{n}\right)$$

$$= -\frac{pq}{n}$$

Example 13: Bring out the fallacy, if any, in the statement:

"The mean of a binomial distribution is 5 and its variance is 9".

Sol. Mean = 5

$\therefore$ np = 5

Variance = 5

$\therefore$ mpq = 9 ....(1)

Dividing (2) by (1), we get

$$q = \frac{9}{5}$$

This is not possible as $0 \leq q \leq 1$

$\therefore$ The given statement is wrong

Example 14: Determine the binomial distribution whose mean is 9 and whose standard deviation is $\frac{3}{2}$.

Sol. Now mean = 9, S.D. = $\frac{3}{2}$

$\therefore$ np = 9 ......(1)

and $\sqrt{npq} = \frac{3}{2}$

$\therefore$ npq = $\frac{9}{4}$ ......(2)

Dividing (2) by (1), we get

$$q = \frac{1}{4}$$

$\therefore$ $p = 1 - q = 1 - \frac{1}{4} = \frac{3}{4}$

Putting value of p in (1), we get

$$n \frac{3}{4} = 9$$

$\Rightarrow$ $n = 9 \times \frac{4}{3} = 12$

$\therefore$ n = 12

$\therefore$ Binomial distribution is $\left(\frac{4}{3} + \frac{3}{4}\right)^{12}$

Example 15: Compute the mode of a binomial distribution b$\left(7, \frac{1}{2}\right)$

Sol. Binomial distribution is b $\left(7, \frac{1}{2}\right)$

$\therefore$ n = 7, p = $\frac{1}{2}$

$\therefore$ q = 1 - p = 1 - $\frac{1}{2}$ = $\frac{1}{2}$

Let 'r' be the mode of the binomial distribution. We know that

np + p - 1 < r < np + p

Now np + p = 7 $\left(\frac{1}{2}\right)$ + $\frac{1}{2}$ = $\frac{7}{2}$ + $\frac{1}{2}$ = $\frac{8}{2}$ = 4

and np + p - 1 = 4 - 1 = 3

$\therefore$ from (1),

$3 \leq r \leq 4$

$\therefore$ r = 3 or 4

Example 16: The m.g.f. of binomial variate about origin was found to be $\left(\frac{3}{4}+\frac{1}{4}e^t\right)^8$, Find

(i) Mean, S.D. and coefficient of variation

(ii) Mode

(iii) P (X = 3)

Sol. We know that if X ~ B (n, p), then its moment generating function is

$M_x(t) = (q + pe^t)^n$

Comparing with $\left(\frac{3}{4}+\frac{1}{4}e^t\right)^8$, we have

q = $\frac{3}{4}$, p = $\frac{1}{4}$, n = 8

(i) Now, mean = np = 8 × $\frac{1}{4}$ = 2

Variance = npq = $8 \times \frac{1}{4} \times \frac{3}{4} = \frac{3}{2}$

S.D. = $\sqrt{\frac{3}{2}}$

and coefficient of variation = $\frac{\sqrt{\frac{3}{2}}}{2} = 0.612$

(ii) The mode is given by (n + 1) = (8 + 1) $\times \frac{3}{4}$

$= 9 \times \frac{3}{4} = \frac{27}{4} = 6.75$

which is not an integer

$\therefore$ The only mode = Integer part of (n + 1) q = 6

$\therefore$ Mode = 6

(iii) By definition P(x = 3) = $8_{c_3} \left(\frac{1}{4}\right)^3 \left(\frac{3}{4}\right)^5 = 0.2076$

Self-Check Exercise-2

Q.1 In eight throws of a die, 5 or 6 is considered a success. Find the mean number of success and its standard deviation.

Q.2 Find the binomial distribution whose mean is 10 and standard deviation $2\sqrt{2}$.

Q.3 If X is binomially distributed with parameters n and p. Find m.g.f. of Y = n - X.

11.11 Summary:

We conclude this unit by summarizing what

1. We have covered in it :
2. Defined Bernoulli variate and Bernoulli distribution. Discussed probability density function of Bernoulli distribution.
3. Discussed in detail the moments and m.g.f. of a Bernoulli variate.
4. Defined and discussed in detail Binomial distribution.
5. Find the mean and variance of Binomial distribution.
6. Discussed m.g.f. of Binomial distribution
7. Discussed moments, central moments and mode of Binomial distribution.
8. Defined standard binomial variate and prove a theorem related to it.

9. Did some examples related to each topic so that the contents be clarified further.

11.12 Glossary:

1. A random variable X is said to have a discrete uniform distribution over the range [1, n] if its probability mass function is expressed as:

$$P(X = x) = \begin{cases} \frac{1}{n} \ forx = 1, 2, \ldots.., n \\ 0 \ \ otheresiwe \end{cases}$$

2. A random variable X which takes two values 0 and 1, with probabilities q and p respectively, i.e. P(X = 1) = p, P(X = 0) = q where q = 1 - p is a Bernoulli variate and is said to have a Bernoulli distribution.

3. A random variable X is said to have Binomial distribution if it assumes only non-negative values and its probability mass function is given by

$$P(X = x) = p(x) = \begin{cases} \binom{n}{x} p^x g^{n-x}, x = 0, 1, \ldots\ldots n; q = 1 - p \\ 0 \qquad otheresiwe \end{cases}$$

The two independent constants n and p in the distribution are known as the parameters of the distribution

4. Let a random variable X have binomial distribution with mean np and variance npq. Then the random variable Z = $\frac{X - np}{\sqrt{npq}}$ having binomial distribution with mean 0 and variance 1 is called standard binomial variate.

11.13 Answer to Self Check Exercise

Self-Check Exercise-1

Ans.1 $\left(\frac{1}{10}\right)^5$

Ans.2 $\frac{5}{16}$

Ans.3 $\frac{4547}{8192}$

Ans.4 (i) $\frac{81}{256}$ (ii) $\frac{1}{256}$

(iii) $\frac{27}{128}$ (iv) $\frac{175}{256}$

Self-Check Exercise-2

Ans.1 Mean: $\frac{16}{9}$ and standard deviation: $\frac{4}{3}$

Ans.2 Binomial distribution is $\left(\frac{4}{5}+\frac{1}{5}\right)^{50}$

Ans.3 Y = (n - x) ∽ (n, q)

11.14 References/Suggested Readings

1. Robert V. Hogg, Joseph W. Mckean and Allen T. craig, Introduction to Mathematical statistics, Pearson Education, Asia, 2007.
2. Irwim Miller, Marylees Miller and John E. Freund, Mathematical Statistics with Application, 7th Ed., Pearson Education, Asia, 2006.
3. Sheldon Ross, Introduction to Probability Model, 9th Ed., Academic Press, Indian Reprint, 2007.

11.15 Terminal Questions

1. With the usual notation, find p for a binomial variate x, if

   n = 6 and 9P (x = 4) = P(x = 2)

2. A bag contains 10 balls each marked with one of the digits 0 to 9. If four balls are drawn successively with replacement from the bag, what is the probability that none is marked with the digit 0?
3. A coin is tossed 5 times. What is the probability that head appears:

   (i) an even number of times

   (ii) an odd number of times

   (You may regard 0 as an even number).

4. A and B play a game in which their chances of winning are in the ratio 3:2. Find A's chance of winning at least three games out of the five games played.
5. Comment on the following:

   "The mean of a binomial distribution is 3 and variance is 4".

6. If a random variable X follows binomial distribution with parameters n and p, prove that

   $P(X = \text{even}) = \frac{1}{2}[1 + (q - p)^n]$

7. Determine the binomial distribution for which mean is 4 and variance 3 and also find its mode.

8. Obtain the m.g.f. of Binomial distribution with $n = 7$, $p = 0.6$. find the first three moments of the distribution.

Unit - 12

Poisson Distribution

Structure

12.1 Introduction

Poisson distribution was discovered by the French Mathematician and physicist. Simeon Denis Poisson (1781 - 1840) who published it in 1837. The Poisson distribution is a discrete distribution that is used to model the number of events occurring in a fixed interval of time or space, when the events occur with a known constant mean rate and independently of the time since the last event. It assumes that the events occur independently of each other, meaning the occurrence of one event does not affect the probability of another event occurring. Poisson distribution is a limiting case of the binomial distribution under the following conditions:

(i) n, the number of trials is very large i.e. $n \to \infty$

(ii) p, the constant probability of success for each trial, is indefinitely small i.e. $p \to 0$

(iii) $np = \lambda$, say is finite.

Poisson distribution is used to model the arrival of customers or requests in queuing system, such as call centers, traffic flow, and computer networks. It can also be used to model the number of claims or events in insurance and finance, such as the number of insurance claims in a given period or the number of defaults in a portfolio of loans. Some instances where Poisson distribution may be successfully employed are : Number of deaths from a disease such as heart attack or cancer in a large city; Number of suicides reported in a particular city; Number of printing mistakes at each page of the book, Number of defective material in a packing manufactured by a good concern etc.

12.2 Learning Objectives

After reading this unit, you should be able to:

- Define Poisson distribution
- Derive Poisson distribution from Binomial distribution
- Find the mean and variance of Poisson distribution
- Find the m.g.f., moments and mode of Poisson distribution
- Discuss property of Poisson distribution

12.3 Derivation of Poisson Distribution From Binomial Distribution

Poisson Distribution - Def.

A random variable X is said to have a Poisson distribution with parameter λ, if its probability function is given by

$$f(x) = f(x ; \lambda) = P(X = x) = \frac{e^{-\lambda}\lambda^{x}}{\lfloor x}, \; x = 0, 1, 2, \ldots, x > 0 \qquad \ldots(1)$$

Here λ is known as the parameter of the distribution.

Note 1. m, instead of λ, is also sometimes taken as the parameter of the distribution.

Note 2. Sum of probabilities = $\sum_{x=0}^{\lambda} \frac{e^{-\lambda}\lambda^{x}}{\lfloor x}$

$$= e^{-\lambda} \sum_{x=0}^{\lambda} \frac{\lambda^{x}}{\lfloor x} = e^{-\lambda}\left(\frac{1}{\lfloor 0} + \frac{\lambda}{\lfloor 1} + \frac{\lambda^{2}}{\lfloor 2} + \ldots\right)$$

$$= e^{-\lambda}\left(1 + \frac{\lambda}{\lfloor 1} + \frac{\lambda^{2}}{\lfloor 2} + \ldots\right) = e^{-\lambda}(e^{\lambda}) = e^{0} = 1$$

Derivation of Poisson Distribution from Binomial Distribution

In the Binomial distribution, the probability of x successes is given by

$$f(x) = {}^{n}C^{x}\, p^{x}\, q^{n-x}$$

$$= \frac{\lfloor n}{\lfloor x \lfloor n-x}\, p^{x}\,(1-p)^{n-x} \qquad [\because q = 1 - p]$$

$$= \frac{n(n-1)(n-2).....(n-x+1)\lfloor n-x}{\lfloor x \lfloor n-x} \; p^x (1-p)^{n-x}$$

$$= \frac{n(n-1)(n-2).....(n-x+1)}{\lfloor n-x} \left(\frac{\lambda}{n}\right)^x \left(1-\frac{\lambda}{n}\right)^{n-x}$$

$$= \frac{n(n-1)(n-2).....(n-x+1)}{\lfloor x} \cdot \frac{\lambda^x}{n^x} \cdot \left(1-\frac{\lambda}{n}\right)^{n-x}$$

$$= \frac{n(n-1)(n-2).....(n-x+1)}{\lfloor x} \cdot \left(1-\frac{\lambda}{n}\right)^{n-x} \frac{\lambda^x}{\lfloor x}$$

$$= \left(\frac{n}{n}\right)\left(\frac{n-1}{n}\right)\left(\frac{n-2}{n}\right)........\left(\frac{n-x+1}{n}\right)\left(1-\frac{\lambda}{n}\right)^{n-x} \frac{\lambda^x}{\lfloor x}$$

$$= (1) \left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)......\left(1-\frac{x-1}{n}\right)\left(1-\frac{\lambda}{n}\right)^{n}\left(1-\frac{\lambda}{x}\right)^{-x} \frac{\lambda^x}{\lfloor x}$$

Let $n \to \infty$

$\therefore$ each of $1-\frac{1}{n}$, $1-\frac{2}{n}$,...... $1-\frac{x-1}{n}$, $\left(1-\frac{\lambda}{n}\right)^{-x} \to 1$

Also $\underset{n\to\infty}{Lt}\left(1-\frac{\lambda}{n}\right)^n = \underset{n\to\infty}{Lt}\left[\left(1-\frac{\lambda}{n}\right)^{-\frac{n}{\lambda}}\right]^{-\lambda} = e^{-\lambda}$ $\left[\because \underset{n\to\infty}{Lt}\left(1+\frac{1}{x}\right)^x = e\right]$

$\therefore$ $f(x) = (1)\,(1-0)\,(1-0)\,(1-0)\,......\,(1-0)\,e^{-\lambda}\,(1)\,\frac{\lambda^x}{\lfloor x}$

or $f(x) = \frac{e^{-\lambda}\lambda^x}{\lfloor x}$

The following examples will illustrate the idea more clearly:

Example 1: If a bank receives on the average 6 bad cheques per day, what is the probability that it will receive four bad cheques on any given day.

(Use $e^{-6} = 0.0025$)

Sol. Here $\lambda = 6$

Required probability = $P(X = 4) = f(4)$

$$= \frac{e^{-\lambda}(6)^4}{\underline{|4}} \qquad \left[\because f(x) = \frac{e^{-\lambda}\lambda^x}{\underline{|x}}\right]$$

$$= (0.0025) \times \frac{6\times6\times6\times6}{4\times3\times2\times1} = (0.0025) \times 54 = .135$$

Example 2: Find the probability that at the most 5 defective fuses will be found in a box of 200 fuses, if experience shows that 2% of such fuses are defective.

($e^{-4} = 0.0183$)

Sol. We have

$$n = 200,\ p = \frac{2}{100} = 0.02$$

$$\therefore \qquad \lambda - n\,p = 200 \times \frac{2}{100} = 4$$

Probability that at the most 5 defective fuses are found = $P(X \leq 5) = \sum_{x=0}^{5} f(x)$

$$= \sum_{x=0}^{5} \frac{e^{-\lambda}\lambda^x}{\underline{|x}} = e^{-4} \sum_{x=0}^{5} \frac{(4)^x}{\underline{|x}}$$

$$= e^{-4}\left[1+4+\frac{4^2}{\underline{|2}}+\frac{4^3}{\underline{|3}}+\frac{4^4}{\underline{|4}}+\frac{4^5}{\underline{|5}}\right]$$

$$= e^{-4}\left[1+4+8+\frac{3^2}{3}+\frac{3^2}{3}+\frac{128}{15}\right]$$

$$= (0.0183)\left(\frac{643}{15}\right) = 0.785$$

Example 3: A car fire firm has cars which it hires out by the day. The number of demands for a car on each day is distributed as a Poisson distribution with mean equal to 2. Calculate the proportion of days on which none of the cars is used, and proportion of days on which some demand is refused.

Sol. Let X denote the number of demand for a car on each day.

It is given that X is a Poisson variate with mean $\lambda = 2$

Therefore, its probability distribution is given by

$$P(X = x) = \frac{2^x e^{-2}}{\underline{|x}},\ x = 0, 1, 2, \ldots\ldots$$

Proportion of days on which none of the cars is used = Probability that there was no demand for a car

= P (X = 0)

$$= \frac{2^0 e^{-2}}{\lfloor 0} = e^{-2} = \frac{1}{e^2}$$

Also, Proportion of days when some demand is refused = Probability that there were more than 3 demands for cars

($\because$ the firm has only 3 cars)

= 1 - [Probability that the demand was for either no car or one car or two cars or for three cars]

= 1 - [P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3)]

$$= 1 - \left[\frac{2^0.e^{-2}}{\lfloor 0} + \frac{2^1.e^{-2}}{\lfloor 1} + \frac{2^2.e^{-2}}{\lfloor 2} + \frac{2^3.e^{-2}}{\lfloor 3}\right]$$

$$= 1 - e^{-2}\left(\frac{19}{3}\right)$$

$$= 1 - \frac{19}{3e^2}$$

Example 4: Assuming the probability that a bomb dropped from an airplane will strike a certain target is $\frac{1}{5}$. If 6 bombs are dropped, find the probability that at least 2 will strike the target.

[Use $e^{-1.2} = 0.3012$]

Sol. Here p = $\frac{1}{5}$, n = 6

$$\therefore \quad m = np = 6 \times \frac{1}{5} = 1.2$$

Probability that at least 2 will strike

= 1 - [P(0) + P(1)]

$$= 1 - \left[\frac{(1.2)^0}{\lfloor 0} e^{-1.2} + \frac{(1.2)^1}{\lfloor 1} e^{-1.2}\right]$$

$$= 1 - e^{-1.2}\,[1 + 1.2]$$

= 1 - (0.3012) (2.2)

$= 1 - 0.6626$

$= 0.3374$

Example 5: A typist on the average makes three errors per page. What is the probability of typing a page

(i) with no errors?

(ii) with at least two errors?

Sol. Here m = 3

(i) P (No error) = $\frac{(3)^0}{\lfloor 0} e^{-3}$ $\left[\because P(x) = \frac{m^x}{\lfloor x} e^{-m}\right]$

$= \frac{1}{1} \times (0.04979) = 0.04979$

(ii) P (at least two errors) = P(2) + P(3)

$= 1 - [P(0) + P(1)]$

$= 1 - [0.04979 + 3\ (0.04979)]$

$= 1 - 0.19916 = 0.80084$

Example 6: Assume that the probability of an individual coal-miner being killed in a mine accident during a year is $\frac{1}{2400}$. Using Poisson distribution, calculate the probability that in a mine employing 200 miners there will be at least one total accident in a year.

Sol. Here p= $\frac{1}{2400}$, n = 200

$\therefore$ m = np = $200 \times \frac{1}{2400} = \frac{1}{12} = 0.083$

Now P(x) = $\frac{m^x e^{-m}}{\lfloor x}$

$\therefore$ P(0) = $\frac{(0.083)^0}{\lfloor 0} e^{-0.083} = e^{-0.083} = 0.92$

P (at least one total accident) = 1 - P (no total accident)

$= 1 - P(0)$

$= 1 - 0.92$

$= 0.08$

Self-Check Exercise-1

Q.1 Six coins are tossed 6400 times. Using the Poisson distribution, find the approximate probability of getting six heads r times.

Q.2 If a random variable X has a Poisson distribution such that P(X = 1) = P(X = 2), then compute P(X = 4)

Q.3 If 5% of the electric bulbs manufactured be a company are defective, use Poisson distribution to find the probability that in a sample of 100 bulbs:-

(a) none is defective

(b) 5 bulbs will be defective

(use $e^{-5} = 0.007$)

12.4 Mean and Variance of Poisson Distribution

$$\text{Mean} = E(X) = \sum_{x=0}^{\infty} x \frac{e^{-\lambda}\lambda^{x}}{\lfloor x}$$

$$= \lambda^{e-\lambda} \sum_{x=0}^{\infty} \frac{\lambda^{x-1}}{\lfloor x-1} = \lambda\, e^{-\lambda} \left(1+\frac{\lambda}{\lfloor 1}+\frac{\lambda^2}{\lfloor 2}+\frac{\lambda^3}{\lfloor 3}+......\right)$$

$$= \lambda\, e^{-\lambda} (e^{\lambda}) = \lambda\, e^{0} = \lambda$$

$\therefore$ mean of Poisson distribution is λ.

$$\text{Variance} = E(X - \mu)^2$$

$$= E(X^2) - [E(X)]^2$$

$$= \sum_{x=0}^{\infty} x^2 f(x) - \lambda^2 = \sum_{x=0}^{\infty} [x(x-1)+x] \frac{e^{-\lambda}\lambda^{x}}{\lfloor x} - \lambda^2$$

$$= e^{-\lambda} \sum_{x=0}^{\infty} x(x-1) \frac{\lambda^{x}}{\lfloor x} + \sum_{x=0}^{\infty} x \frac{e^{-\lambda}\lambda^{x}}{\lfloor x} - \lambda^2$$

$$= \lambda^2 e^{-\lambda} \left[\sum_{x=0}^{\infty} \frac{\lambda^{x-2}}{\lfloor x-2}\right] + \lambda - \lambda^2$$

$$= \lambda^2 e^{-\lambda} \left(1+\frac{\lambda}{\lfloor 1}+\frac{\lambda^2}{\lfloor 2}+\frac{\lambda^3}{\lfloor 3}+......\right) + \lambda - \lambda^2$$

$$= \lambda^2 e^{-\lambda} (e^{\lambda}) + \lambda - \lambda^2$$

$$= \lambda^2 + \lambda - \lambda^2 = \lambda$$

$\therefore$ Variance = λ

Note : Mean and Variance are equal.

12.5 Moment Generating Function of Poisson Distribution

We have

$$M_x(t) = E(e^{tx})$$

$$= \sum_{x=0}^{\infty} e^{tx} f(x) = \sum_{x=0}^{\infty} e^{tx} \frac{e^{-\lambda}\lambda^x}{\lfloor x}$$

$$= e^{-\lambda} \sum_{x=0}^{\infty} \frac{(\lambda e^t)^x}{\lfloor x} = e^{-\lambda}\left[1 + \frac{(\lambda e^t)}{\lfloor 1} + \frac{(\lambda e^t)^2}{\lfloor 2} + \frac{(\lambda e^t)^3}{\lfloor 3} + \ldots\ldots\right]$$

$$= e^{-\lambda}\left(e^{\lambda e^t}\right) = e^{\lambda e^{t-\lambda}} = e^{\lambda}(e^t - 1)$$

12.6 Moments of the Poisson Distribution

$$\mu_1' = E(X) = \sum_{x=0}^{\infty} x\, f(x, \lambda)$$

$$= \sum_{x=0}^{\infty} x. \frac{e^{-\lambda}\lambda^x}{\lfloor x} = \lambda\, e^{-\lambda}\left[\sum_{x=0}^{\infty} \frac{\lambda^{x-1}}{\lfloor x-1}\right]$$

$$= \lambda\, e^{-\lambda}\left(1 + \lambda + \frac{\lambda^2}{\lfloor 2} + \frac{\lambda^2}{\lfloor 3} + \ldots\ldots\right) = \lambda\, e^{-\lambda} . e^{\lambda} = \lambda$$

Hence the mean of the Poisson distribution is λ.

$$\mu_2' = E(X^2) = \sum_{x=0}^{\infty} x^2 f(x, \lambda) = \sum_{x=0}^{\infty} \{x(x-1) + x\} \frac{e^{-\lambda}\lambda^x}{\lfloor x}$$

$$= e^{-\lambda} \sum_{x=0}^{\infty} x(x-1) \frac{\lambda^x}{\lfloor x} + \sum_{x=0}^{\infty} x \frac{e^{-\lambda}\lambda^x}{\lfloor x}$$

$$= \lambda^2 e\text{-}\lambda \left[\sum_{x=2}^{\infty} \frac{\lambda^{x-2}}{\lfloor x-2}\right] + \lambda = \lambda^2 e^{-\lambda} e^{\lambda} + \lambda = \lambda^2 + \lambda$$

$$\mu_3' = E(X^3) = \sum_{x=0}^{\infty} x^3 f(x, \lambda) = \sum_{x=0}^{\infty} \{x(x-1)(x-2) + 3\, x(x-1) + x\} \frac{e^{-\lambda}\lambda^x}{\lfloor x}$$

$$= \sum_{x=0}^{\infty} x(x-1)(x-2) \frac{e^{-\lambda}\lambda^x}{\lfloor x} + 3\sum_{x=0}^{\infty} x(x-1) \frac{e^{-\lambda}\lambda^x}{\lfloor x} + \sum_{x=0}^{\infty} x \frac{e^{-\lambda}\lambda^x}{\lfloor x}$$

$$= e^{-\lambda}\lambda^3\left[\sum_{x=3}^{\infty}\frac{\lambda^{x-3}}{\lfloor x-3}\right] + 3e^{-\lambda}\lambda^2\left[\sum_{x=2}^{\infty}\frac{\lambda^{x-2}}{\lfloor x-2}\right] + \lambda$$

$$= e^{-\lambda}\lambda^3 e^{\lambda} + 3e^{-\lambda}\lambda^2 e^{\lambda} + \lambda = \lambda^3 + 3\lambda^2 + \lambda$$

$$\mu_4' = E(X^4) = \sum_{x=0}^{\infty} x^4 f(x, \lambda)$$

$$= \sum_{x=0}^{\infty} \{x(x-1)(x-2)(x-3) + 6x(x-1)(x-2) + 7x(x-1) + x\}\frac{e^{-\lambda}\lambda^x}{\lfloor x}$$

$$= e^{-\lambda}\lambda^4\left[\sum_{x=4}^{\infty}\frac{\lambda^{x-4}}{\lfloor x-4}\right] + 6e^{-\lambda}\lambda^3\left[\sum_{x=3}^{\infty}\frac{\lambda^{x-3}}{\lfloor x-3}\right]$$

$$+ 7e^{-\lambda}\lambda^2\left[\sum_{x=2}^{\infty}\frac{\lambda^{x-2}}{\lfloor x-2}\right] + \lambda$$

$$= \lambda^4(e^{-\lambda}e^{\lambda}) + 6\lambda^3(e^{-\lambda}e^{\lambda}) + 7\lambda^2(e^{-\lambda}e^{\lambda}) + \lambda = \lambda^4 + 6\lambda^3 + 7\lambda^2 + \lambda$$

Now $\mu_2 = \mu_1'^2 = (\lambda^2 + \lambda) - \lambda^2 = \lambda$

Thus the mean and the variance of the Poisson distribution are each equal to λ

$\mu_3 = \mu_3' - 3\mu_1'\mu_2' + 2\mu_1'^3$

$= (\lambda^3 + 3\lambda^2 + \lambda) - 3\lambda(\lambda^2 + \lambda) + 2\lambda^3 = \lambda$

$\mu_4 = \mu_4' - 4\mu_3'\mu_1' + 6\mu_2'\mu_1'^2 - 3\mu_1'^4$

$= (\lambda^4 + 6\lambda^3 + 7\lambda^2 + \lambda) - 4\lambda(\lambda^3 + 3\lambda^2 + \lambda) + 6\lambda^2(\lambda^2 + \lambda) - 3\lambda^4$

$= 3\lambda^2 + \lambda$

Co-efficient 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 $\beta_2 = \frac{\mu_4}{\mu_2^2} = 3 + \frac{1}{\lambda}$ and $\gamma_2 = \beta_2 - 3 = \frac{1}{\lambda}$

Hence the Poisson distribution is alwyas a skewed distribution

Proceeding to the limit as $\lambda \to \infty$, we get

$\beta_1 = 0$ and $\beta_2 = 3$

12.7 Mode of Poisson Distribution

We know that the mode is that value of x for which P(X = x) is greater than the P(X = x + 1) and also P(X = x - 1). If x is the mode of Poisson distribution, then

$$P(x - 1) \leq P(x) \geq P(x + 1)$$

i.e. $$\frac{e^{-\lambda}\lambda^{x-1}}{\lfloor x-1} \leq \frac{e^{-\lambda}\lambda^{x}}{\lfloor x} \geq \frac{e^{-\lambda}\lambda^{x+1}}{\lfloor x+1}$$

$$\Rightarrow \quad \frac{e^{-\lambda}\lambda^{x-1}}{\lfloor x-1} \leq \frac{e^{-\lambda}\lambda^{x}}{\lfloor x} \quad \text{or} \quad \frac{e^{-\lambda}\lambda^{x}}{\lfloor x} \geq \frac{e^{-\lambda}\lambda^{x+1}}{\lfloor x+1}$$

$$\Rightarrow \quad 1 \leq \frac{\lambda}{x} \quad \text{or} \quad 1 \geq \frac{\lambda}{x+1}$$

i.e. $x \leq \lambda$ or $x \geq \lambda - 1$

Thus, if x is an integer, then both x - 1 and x represent modes.

And if x is not an integer, then mode is the integral value between $\lambda - 1$ and λ.

12.8 Property of the Poisson Distribution

If X_1 and X_2 be independent variates having Poisson distribution with means λ_1 and λ_2, respectively, then their sum $X_1 + X_2$ is again a Poisson variate with mean $\lambda_1 + \lambda_2$.

Proof: Let $M_1(t)$, $M_2(t)$ and M(t) be the moment generating functions of the variates X_1, X_2, and $X_1 + X_2$ respectively.

Then, $M_1(t) = \exp\{\lambda_1 (e^t - 1)\}$

$M_2(t) = \exp\{\lambda_2 (e^t - 1)\}$

Now, the moment generating function of $X_1 + X_2$ is given by

$$M(t) = M_1(t) \times M_2(t)$$

$$= \exp\{\lambda_1 (e^t - 1)\} \exp\{\lambda_2 (e^t - 1)\}$$

$$= \exp\{\lambda_1 (e^t - 1) + \lambda_2 (e^t - 1)\}$$

$$= \exp\{((\lambda_1 + \lambda_2) e^t - (\lambda_1 + \lambda_2)\}$$

$$= \exp[(\lambda_1 + \lambda_2) \{e^t - 1\}]$$

Which we know is the moment generating function of a Poisson distribution having mean $\lambda_1 + \lambda_2$.

Hence the result

Dear students let us improve our understanding of these results by looking at some following examples:-

Example 7:- If X is a Poisson variate such that P(X = 1) = 2 P(X = 2), find

(i) P(X = 0) (ii) E(X) (iii) V(X)

Sol. Here

P(X = 1) = 2 P(X = 2)

$\Rightarrow \quad \dfrac{e^{-\lambda}.\lambda}{\lfloor 1} = 2\left(\dfrac{e^{-\lambda}\lambda^2}{\lfloor 2}\right)$

$\Rightarrow \quad \dfrac{e^{-\lambda}\lambda}{1} = 2x\dfrac{e^{-\lambda}\lambda^2}{2}$

$\therefore \quad \lambda = \lambda^2 \quad \Rightarrow \quad \lambda^2 - \lambda = 0 \quad \Rightarrow \quad \lambda(\lambda - 1) = 0$

$\Rightarrow \quad \lambda = 0, 1$

$\therefore \quad \lambda = 1 \qquad [\because \lambda > 0]$

(i) $\quad P(X = 0) = \dfrac{e^{-1}\lambda^0}{\lfloor 0} = \dfrac{e^{-1}1}{1} = \dfrac{1}{e}$

(ii) $\quad$ Now $E(X) = \text{Mean} = \lambda = 1$

(iii) $\quad V(X) = \lambda = 1$

Example 8: Find the m.g.f. about mean of the Poisson distribution and from it derive the first four moments about the mean.

Sol. Let X be a Poisson variate with mean λ, then the m.g.f. of X about λ is

$M_{x-\lambda}(t) = e^{-\lambda t} M_x(t)$

Where $M_x(t)$ is the m.g.f. of X about the origin

But $\quad M_x(t) = e^{\lambda(e^t - 1)}$

$\therefore \quad M_{x-\lambda}(t) = e^{-\lambda t} e^{\lambda} (e^t - 1)$

$\qquad = e^{\lambda} (e^t - 1 - t)$

$\therefore \quad 1 + \mu_1 t + \mu_2 \dfrac{t^2}{\lfloor 2} + \mu_3 \dfrac{t^3}{\lfloor 3} + \ldots\ldots = e^{\lambda}\left(\dfrac{t^2}{\lfloor 2} + \dfrac{t^3}{\lfloor 3} + \dfrac{t^4}{\lfloor 4} + \ldots\ldots\right)$

$= 1 + \lambda\left(\dfrac{t^2}{\lfloor 2} + \dfrac{t^3}{\lfloor 3} + \dfrac{t^4}{\lfloor 4} + \ldots\ldots\right) + \dfrac{t^2}{\lfloor 2}\left(\dfrac{t^2}{\lfloor 2} + \dfrac{t^3}{\lfloor 3} + \ldots\ldots\right) + \ldots\ldots$

Equating coefficient of t, $\dfrac{t^2}{\lfloor 2}, \dfrac{t^3}{\lfloor 3}, \dfrac{t^4}{\lfloor 4}$, ..., we get

$\mu_1 = 0$

$\mu_2 = \lambda$

$\mu_3 = \lambda$

$\mu_4 = \lambda + 3\lambda^2$

Example 9: Show that in a Poisson distribution with unit mean, the mean deviation about its mean is $\frac{2}{e}$ times the standard deviation.

Sol. Let X be a Poisson variate having mean $\lambda = 1$

and p.f. = $\frac{e^{-\lambda}\lambda^x}{\lfloor x}$, x = 0,1,2,3,......

$\therefore$ Mean deviation = E $|(X - \lambda)|$

$$= \sum_{x=0}^{\infty} |x - \lambda|\, f(x) = \sum_{x=0}^{\infty} |x - \lambda| \frac{e^{-\lambda}\lambda^x}{\lfloor x}$$

$$= \sum_{x=0}^{\infty} |x - 1| \frac{e^{-1}1^x}{\lfloor x}$$

$$= \frac{1}{e}\sum_{x=0}^{\infty} \frac{|x-1|}{\lfloor x}$$

$$= \frac{1}{e}\left[1+0+\frac{1}{\lfloor 2}+\frac{2}{\lfloor 3}+\frac{3}{\lfloor 4}......\right]$$

$$= \frac{1}{e}\left[1+\left(1-\frac{1}{\lfloor 2}\right)+\left(\frac{1}{\lfloor 2}-\frac{1}{\lfloor 3}\right)+......\right]$$

$$\left[\because \frac{n}{\lfloor n+1} = \frac{(n+1)-1}{\lfloor n+1} = \frac{1}{\lfloor n} - \frac{1}{\lfloor n+1}\right]$$

$$= \frac{1}{e}[1 + 1] = \frac{2}{e}$$

$$= \frac{2}{e} \times 1$$

$$= \frac{2}{e} \times (\lambda) = \frac{2}{e} \text{ (standard deviation)}$$

Example 10: If x is a Poisson variate with mean m, what would be the expectation of e^{-kx}, where k is a constant. Find also the expectation of e^{-kx} kx.

Sol: The probabilities of x variate in Poisson distribution are given to be successive terms of $\frac{m^x e^{-m}}{\lfloor x}$,

Where x = 0,1,2,3,.......

The expectation of e^{-kx} is

$$E(e^{-kx}) = \sum_{x=0}^{\infty} e^{-kx} \frac{m^x e^{-m}}{\lfloor x}$$

$$= e^{-m} \sum_{x=0}^{\infty} \frac{(me^{-k})^x}{\lfloor x}$$

$$= e^{-m} \left[\frac{1}{1} + \frac{me^{-k}}{\lfloor 1} + \frac{(me^{-k})^2}{\lfloor 2} + ...\infty \right]$$

$$= e^{-m}.e^{m.e^{-k}} = e^{-m+me^{-k}} = e^{-m(1-e^{-k})}$$

$$\text{Further} \, E(kxe^{-kx}) = \sum_{x=0}^{\infty} \frac{kxe^{-kx} m^x e^{-m}}{\lfloor x}$$

$$= ke^{-m} \sum_{x=1}^{\infty} \frac{e^{-kx} m^x}{\lfloor x-1}$$

$$= ke^{-m} \left[me^{-k} + \frac{m^2 e^{-2k}}{\lfloor 1} + \frac{m^3 e^{-2k}}{\lfloor 2} + ... \right]$$

$$= ke^{-m}.me^{-k} e^{me^{-k}}$$

$$= mk(e)^{-m-k+me^{-k}}$$

$$= mk(e)^{-m(1-e^{-k})-k}$$

Example 11: if X and Y are independent variates such that P(X = 1) = P(X = 2) and P(Y = 2) = P(Y = 3) Find the variance of X - 2y.

Sol. Let X ~ P(λ) and Y ~ P (μ). Then we have

$$P(X = x) = \frac{e^{-\lambda} \lambda^x}{\lfloor x}, \; x = 0,1,2,......, \lambda > 0$$

and $$P(Y = y) = \frac{e^{-\mu} \mu^y}{\lfloor y}, \; y = 0,1,2,......, \mu > 0$$

Since P(X = 1) = P(X = 2) and P(Y = 2) = P(Y = 3)   (given)

$\therefore \quad \lambda e^{-\lambda} = \dfrac{\lambda^2 e^{-\lambda}}{\lfloor 2}$ and $\dfrac{\mu^2 e^{-\mu}}{\lfloor 2} = \dfrac{\mu^3 e^{-\mu}}{\lfloor 3}$

Solving these, we get

$\lambda e^{-\lambda} (\lambda - 2) = 0$ and $\mu^2 e^{-\mu} (\mu - 3) = 0$

$\Rightarrow \quad \lambda = 2$ and $\mu = 3$; since $\lambda > 0, \mu > 0$

Now Var (X) = λ = 2 and Var (Y) = μ = 3

$\therefore$ Var (X - 2Y) = 1^2. Var(X) + $(-2)^2$. Var (Y)

[Using $V(ax + by) = V(ax) + V(by) = a^2V(X) + b^2V(Y)$]

Covariance term vanishes since X and Y are independent.

Thus, we have Var (X - 2Y) = $2 + 4 \times 3 = 14$

Self-Check Exercise-2

Q.1 If X and Y are two independent Poisson variates having means 1 and 3 respectively, then find V(3x + y)

Q.2 If a Poisson distribution has a double mode at x = 1 and x = 2, find the probability that X. will have one or the other of these two values.

Q.3 If the variance of the Poisson distribution is 2, find the distribution for x = 1,2,3,4 and 5. (use $e^{-2} = 0.1356$)

12.9 Summary

We conclude this unit by summarizing what we have covered in it:

1. Defined Poisson distribution and derived Poisson distribution from Binomial distribution.
2. Find the mean and Variance of Poisson distribution.
3. Discussed the m.g.f., moments and mode of Poisson distribution and derived the formula to calculate each.
4. Discussed property of Poisson distribution
5. Did some examples related to each topic so that the contents be clarified further.

12.10 Glossary:

1. A random variable X is said to have a Poisson distribution with parameter λ, if its probability function is given by

$$f(x) = f(x ; \lambda) = P(X = x) = \frac{e^{-\lambda}\lambda^x}{\lfloor x}, \ x = 0,1,2,...., x, \ x > 0.$$

2. Mean of Poisson distribution is λ.

3. Variance of Poisson distribution is λ.

4. Moment generating function of Poisson distribution is e^{λ} $(e^t - 1)$

12.11 Answer to Self Check Exercise

Self-Check Exercise-1

Ans.1 P(X = r) = $\dfrac{e^{-100}.(100)^r}{\lfloor\underline{r}}$; r = 0,1,2,......

Ans.2 P(X = 4) = $\dfrac{2}{3e^2}$

Ans.3 (a) P(none is defective) = 0.007

(b) P(5 defective bulbs) = 0.1822

Self-Check Exercise-2

Ans.1 V(3X + Y) = 12

Ans.2 P(X = 1 or X = 2) = 0.542

Ans.3 Required distribution

$$= (0.1358) \left[\frac{2}{\lfloor\underline{1}} + \frac{2^2}{\lfloor\underline{2}} + \frac{2^3}{\lfloor\underline{3}} + \frac{24}{\lfloor\underline{4}} + \frac{2^5}{\lfloor\underline{5}}\right]$$

12.12 References/Suggested Readings

1. Robert V. Hogg, Joseph w. Mckean and Allen T. craig, Introduction to Mathematical statistics, Pearson Education, Asia, 2007.

2. Irwim Miller, Marylees Miller and John E. Freund, Mathematical Statistics with Application, 7th Ed., Pearson Education, Asia, 2006.

3. Sheldon Ross, Introduction to Probability Model, 9th Ed., Academic Press, Indian Reprint, 2007.

12.13 Terminal Questions

1. A block has 200 pages and 200 misprints distributed at random. What is the probability that a page contains

   (a) exactly two misprints ?

   (b) fewer than two misprints ?

   (Use e^{-1} = 0.09195)

2. Suppose the probability that an item produced by a particular machine is defective equals 0.2. If 10 items produced by this machine are selected at random, what is the probability that no more than one defective item is found.

(Use $e^{-2} = 0.01353$)

3. In a book of 520 pages, 390 type-graphical errors occur. Assuming Poisson distribution for the number of errors per page, find the probability that a random sample of 5 pages will contain no error.

4. If X is a Poisson variate such that P(X = 2) = 9 P(X = 4) + 90 P(X = 6)

   Find (i) λ, the mean of X and variance

   (ii) β, the coefficient of skewness,

5. Bring out the fallacy, if any, in the following statement:-

   "The mean of a Poisson distribution is 5 while its standard deviation is 4".

6. A Poisson distribution has a double mode at x = 4 and x = 5. Find the probability that x will have either of these values.

7. If X is a Poisson variate with parameter m and μ_r is the r^{th} central moment, prove that

$$m\left[r_{c_1}\mu_{r-1} + r_{c_2}\mu_{r-2} + \text{.........} + r_{c_r}\mu_0\right] = \mu_{r+1}.$$

Unit - 13

Uniform Distribution

Structure

13.1 Introduction

A continuous probability distribution is a type of probability distribution that describes the probability of a random variable taking on a value in a continuous range. Unlike discrete probability distributions, which deal with random variables that can only take on specific, countable values, continuous probability distributions deal with random variables that can take on any value within a specified interval.

The uniform or rectangular continuous distribution is a probability distribution where the random variable can take on any value within a specified interval, and all values within that interval are equally likely to occur. In other words, the probability density function (PDF) of a uniform distribution is constant over the interval. The uniform distribution is often used in various fields, such as computer science (for generating random numbers), decision-making (for modeling uncertainty), and physics (for describing the distribution of particles in a container). The uniform distribution is a continuous probability distribution, which means that the random variable can take on any value within the specified interval, rather than just discrete values.

13.2 Learning Objectives

After reading this unit, you should be able to:

- Define uniform (rectangular) distribution
- Discuss characteristics of uniform distribution
- Discuss properties of uniform distribution

13.3 Definition and Characteristics of Uniform or Rectangular Distribution

Def:- A random variable X is said to follow a continuous uniform (rectangular) distribution over interval (a, b) if its probability density function is constant k (say) over entire interval or range

$$f(x) = \begin{cases} k, & a \le x \le b \\ 0, & otherwise \end{cases}$$

For $f(x)$ to be a probability density function, we must have

$$\int_a^b f(x)dx = 1 \Rightarrow \int_a^b k\,dx = 1 \Rightarrow k = \frac{1}{b-a}$$

$$\therefore \quad f(x) = \begin{cases} \frac{1}{b-a}, & a < x < b \\ 0 & , otherwise \end{cases} \qquad ....(1)$$

Characteristics:-

1. a and b, (a < b) are the two parameters of the distribution. The distribution is called uniform distribution on (a, b) since it assumes a constant (uniform) value for all x in (a, b).

2. The distribution is also known as rectangular distribution, since the curve y = $f(x)$ describes a rectangle over the x-axis and between the ordinates x = a and x = b

3. A uniform or rectangular variate X on the interval (a, b) is written as

   $X \backsim \cup [a, b]$ or $X \backsim R [a, b]$

4. The cumulative distribution function F(x) can be calculated as follows:-

Case I: When $x \le a$

$$F(x) = P[X \le x] = \int_{-\infty}^{x} f(x)dx = \int_{-\infty}^{x} 0\,dx = 0$$

Case II: When a < x < b

$$F(x) = P[X \le x] = \int_{-\infty}^{a} f(x)dx + \int_{a}^{x} f(x)dx$$

$$= \int_{-\infty}^{a} 0\,dx + \int_{a}^{x} \frac{1}{b-a}dx = \frac{x-a}{b-a}$$

Case III: When $x \ge b$

$$F(x) = P[X \le x] = \int_{-\infty}^{a} f(x)dx + \int_{a}^{b} f(x)dx + \int_{b}^{x} f(x)dx$$

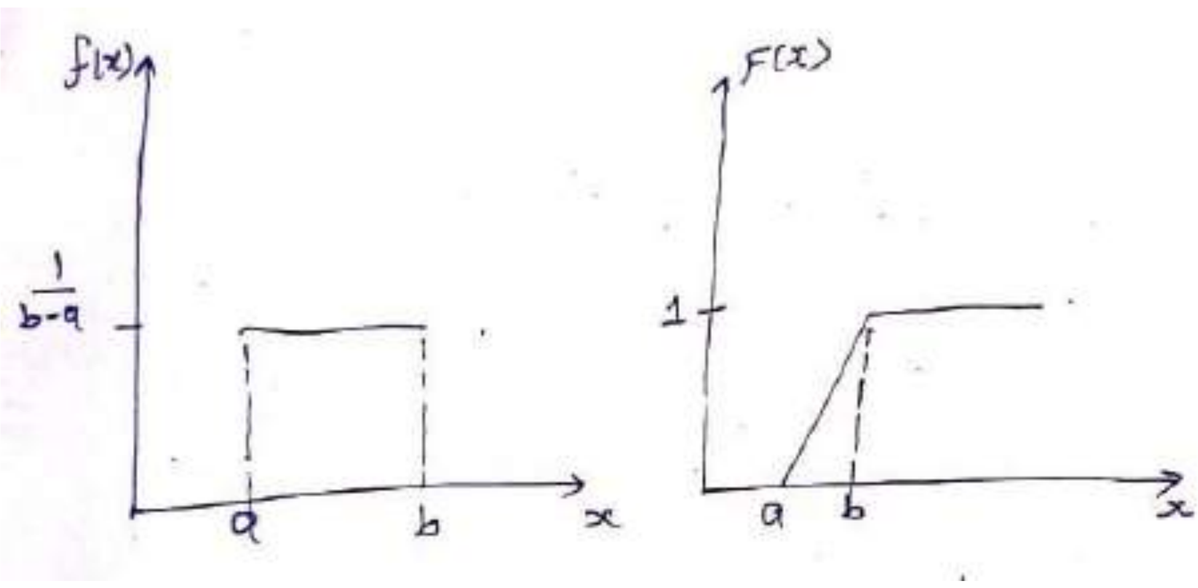
$$= \int_{-\infty}^{a} 0\,dx + \int_{a}^{b} \frac{1}{b-a}dx + \int_{b}^{x} 0\,dx = \frac{b-a}{b-a} = 1$$

∴ The cumulative distribution function F(x) is given by

$$f(x) = \begin{cases} 0\ , & x \le a \\ \dfrac{x-a}{b-a}, & a < x < b \\ 1\ , & x \ge b \end{cases}$$

Since F(x) is not continuous at x = a and x = b, it is not differentiable at these points. Thus $\frac{d}{dx}$F(x) = f(x) = $\frac{1}{b-a} \neq 0$, exists everywhere except at the points x = a and x = b and consequently probability density function f(x) is given by (1).

5. The graphs of uniform probability density function f(x) and the corresponding distribution function F(x) are given below:-



6. For a rectangular or uniform variate X in (-a, a), the probability density function (p.d.f.) is given by

$$f(x) = \begin{cases} \dfrac{1}{2a}, & -a < x < a \\ 0\ , & otherewise \end{cases}$$

13.4 Properties of a Uniform Distribution

Property I : Moments of Rectangular Distribution

Let X ~ U [a, b]

$\therefore \quad \mu_r^1 = E(X^r) = \int_a^b x^r f(x)dx = \frac{1}{b-a}\int_a^b x^r dx = \frac{1}{b-a}\left(\frac{b^{r+1}-a^{r+1}}{r+1}\right)$

In particular

On taking r = 1; $\quad \mu_1' = \frac{1}{b-a}\left(\frac{b^2-a^2}{2}\right) = \frac{b+a}{2}$

$\Rightarrow \quad$ Mean = $\mu_r^1 = \frac{a+b}{2}$

and on taking r = 2; $\mu_2' = \frac{1}{b-a}\left(\frac{b^3-a^3}{3}\right) = \frac{1}{3}(b^2 + ab + a^2)$

$\therefore \quad$ Variance = $\mu_2' - \mu_1'^2 = \frac{1}{3}(b^2 + ab + a^2) - \left\{\frac{1}{2}(b+a)\right\}^2$

$\Rightarrow \quad$ Variance = $\mu_2 = \frac{1}{12}(b - a)^2$

Property II: Moment generating function of Rectangular Distribution is given by

$M_x(t) = E[e^{tx}] = \int_a^b e^{tx} f(x)dx = \int_a^b \frac{e^{tx}}{b-a}dx = \frac{e^{bt}-e^{at}}{t(b-a)}$, $t \neq 0$

$$= \frac{\left(1+bt+\frac{b^2t^2}{2!}+.......+\frac{b^rt^r}{r!}+.......\right)-\left(1+at+\frac{a^2t^2}{2!}+.......+\frac{a^rt^r}{r!}+.......\right)}{t(b-a)}$$

$$= \frac{(b-a)t+(b^2-a^2)\frac{t^2}{2!}+.......+(b^r-a^r)-\frac{t^r}{r!}+(b^{r+1}-a^{r+1})\frac{t^{r+1}}{r+1!}+.......}{t(b-a)}$$

$\therefore \quad \mu_r^1$ = Coeff. of $\frac{t^r}{r!} = \frac{b^{r+1}-a^{r+1}}{(b-a)(r+1)}$

$\therefore \quad$ In particular on taking r = 1,2,3,4,.....; we get

$$\mu_1^1 = \frac{a+b}{2}$$

$$\mu_2^1 = \frac{1}{3}\ (b^2 + ab + a^2)$$

$$\mu_3^1 = \frac{1}{4}(b + a)\ (b^2 + a^2)$$

$$\mu_4^1 = \frac{1}{5}\ \frac{b^5 - a^5}{(b-a)} = \frac{b^4 + b^3a + b^2a^2 + ba^3 + a^4}{5}$$

∴ Various order Central moment are

$$\mu_1 = \mu_1^1 - \mu_1^1 = 0$$

$$\mu_2 = \mu_2^1 - (\mu_1^1)^2 = \frac{(b-a)^3}{12}$$

$$\mu_3 = \mu_3^1 - 3\mu_2^1\ \mu_1^1 + 2(\mu_1^1)^3 = 0$$

$$\mu_4 = \mu_4^1 - 4\mu_3^1\ \mu_1^1 + 6\mu_2^1(\mu_1^1)^2 - 3(\mu_1^1)^4 = \frac{(b-a)^4}{80}$$

∴ Coeff. of skewness = $\beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{0}{\mu_2^3} = 0$

Coeff. of Kurtosis = $\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{(b-a)^4/80}{[(b-a)^2/12]^2} = \frac{(12)^2}{80}$

$$= \frac{9}{5} = 1.8$$

Also $\gamma_1 = \sqrt{\beta_1} = \sqrt{0} = 0$

and $\gamma_2 = \beta_2 - 3 = 1.8 - 3 = -1.2 < 0$

Thus, rectangular distribution is symmetrical and platykurtic in nature.

Property III: Characteristic Function of Rectangular Distribution is given by:

$$\phi_X(t) = \int_a^b e^{itx} f(x)dx = \frac{e^{ibt} - e^{iat}}{it(b-a)}, \ t \neq 0$$

Property IV: Mean Deviation about Mean (M.D.) of Uniform Distribution is given by

M.D. = E|X - Mean| = $\int_a^b |x - Mean| f(x)dx \int_a^b \left|x - \frac{a+b}{2}\right| \frac{1}{b-a}$ dx

$$= \frac{1}{b-a}\int_a^b \left|x - \frac{a+b}{2}\right| dx$$

On taking x - $\frac{a+b}{2}$ = t

$\therefore$ dx = dt, we get

$$\therefore \quad \text{M.D.} = \frac{1}{b-a}\int_{-(b-a)/2}^{(b-a)/2} |t|\, dt$$

$$\text{M.D.} = \frac{1}{b-a} 2 \int_0^{\frac{(b-a)}{2}} t\, dt = \frac{b-a}{4}$$

$$\Rightarrow \quad \text{M.D.} = \frac{b-a}{4}$$

Property V: The quartiles of uniform distribution are given by:

First Quartile Q_1 is given by

$$\int_a^{Q_1} f(x)dx = \frac{1}{4}$$

$$\Rightarrow \quad \int_a^{Q_1} \frac{1}{b-a} dx = \frac{1}{4} \Rightarrow \frac{Q_1 - a}{b-a} = \frac{1}{4}$$

$\therefore$ Q_1 = a + $\frac{1}{4}$ (b - a)

Second Quartile Q_2 or Median is given by

$$\int_a^{Q_2} f(x)dx = \frac{1}{2}$$

$$\Rightarrow \quad \int_a^{Q_2} \frac{1}{b-a} dx = \frac{1}{2} \Rightarrow \frac{Q_2 - a}{b-a} = \frac{1}{2}$$

$\therefore$ Q_2 = Median = a + $\frac{1}{2}$ (b - a)

Third Quartile Q_3 is given by

$$\int_a^{Q_3} f(x)dx = \frac{3}{4}$$

$\Rightarrow$ $$\int_a^{Q2} \frac{1}{b-a}dx = \frac{3}{4} \Rightarrow \frac{Q_3 - a}{b-a} = \frac{3}{4}$$

$\therefore$ $Q_3 = a + \frac{3}{4}(b - a)$

Let us improve our understanding of these results by looking at some following examples:-

Example 1: If X is uniformly distributed with mean 1 and variance $\frac{4}{3}$, find P(X < 0).

Sol. Let X ∽ ∪ [a, b], so that $f(x) = \frac{1}{b-a}$, a < x < b

Now Mean = 1 [Given]

$\Rightarrow$ $\frac{1}{2}(b + a) = 1 \Rightarrow b + a = 2$ ....(1)

And Variance = $\frac{4}{3}$ [Given]

$\Rightarrow$ $\frac{1}{12}(b - a)^2 = \frac{4}{3}$

$\Rightarrow$ $(b - a)^2 = 16$

or $b - a = \pm 4$

Since a < b, so b - a is always positive

Therefore, b - a = 4,

On solving (1) and (2), we get a = -1 and b = 3

$\therefore$ $f(x) = \frac{1}{4}$; $-1 < x < 3$

Thus, $P(X < 0) = \int_1^0 f(x)dx = \int_{-1}^0 \frac{1}{4}dx =$

$$= \frac{1}{4}\,|x|_{-1}^{0}$$

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

Example 2: Suppose X is uniformly distributed over (-a, a), where a > 0. Determine a so that $P\left(X < \frac{1}{2}\right) = 0.3$

Sol. Since the random variable X follows uniform distribution over (-a, a), a > 0. Therefore, its p.d.f. is given by

$$f(x) = \begin{cases} \frac{1}{2a}; -a \le x \le a \\ 0 \ ; \ otherwise \end{cases}$$

Now $P\left(X < \frac{1}{2}\right) = 0.3$

$\Rightarrow \quad \int_{-a}^{1/2} f(x)dx = 0.3$

$\Rightarrow \quad \int_{-a}^{1/2} \frac{1}{2a} dx = 0.3$

$\Rightarrow \quad \frac{\frac{1}{2}+a}{2a} = 0.3$

$\Rightarrow \quad 0.5 + a = 0.6\, a$

$\Rightarrow \quad a - 0.6\, a = -0.5$

$\Rightarrow \quad 0.4\, a = -\, 0.5$

$\Rightarrow \quad a = -\frac{5}{4}$

Example 3: A random variable X has a uniform distribution over (-3, 3). Find k for which $P(X \ge k) = \frac{1}{3}$

Sol: Since the random variable X follows uniform distribution over (-3, 3). Therefore, its p.d.f. is given by

$f(x) = \begin{cases} \frac{1}{6}; -3 \le x \le 3 \\ 0 \;\; ; \; elsewhere \end{cases}$

Given P(X $\geq$ k) = $\frac{1}{3}$

$\Rightarrow \quad \int_k^3 f(x)dx = 1$

$\Rightarrow \quad \int_k^3 \frac{1}{6}dx = \frac{1}{3}$

$\Rightarrow \quad \frac{3-k}{6} = \frac{1}{3}$

$\Rightarrow$ 3 - k = 2

$\Rightarrow$ k = 1

Example 4: For rectangular distribution f(x) = 1, 1 $\leq$ x $\leq$ 2. Find A.M., G.M. H.M. and S.D. and show that A.M. > G.M. > H.M.

Sol. Given f(x) = 1; 1 $\leq$ x $\leq$ 2

i.e. a = 1, b = 2

$\therefore$ A.M. = E(X) = $\int_1^2 x.f(x)dx = \int_1^2 x\,dx = \left[\frac{x^2}{2}\right]_1^2$ = 1.5

Variance = E[X - E(x)]^2 = E[X - 1.5]^2 = $\int_1^2 (x-1.5)^2 f(x)\,dx$

= $\left[\frac{(x-1.5)^3}{3}\right]_1^2 = \frac{1}{3}\left[\left(\frac{1}{2}\right)^3 - \left(-\frac{1}{2}\right)^3\right] = \frac{1}{12}$

$\therefore$ S.D. = $\sqrt{\frac{1}{12}}$

If G is G.M., then

$\log_e$ G = E ($\log_e$ X) = $\int_1^2 (\log_e x)\,dx = \left[x\log_e x\right]_1^2 - \int_1^2 x.\frac{1}{x}dx$

$= 2\ \log_e{}^2 - \ x\Big|_1^2\ =\ \log_e{}^2 - 1 = 2 \times 0.6931 - 1 = 0.3862$

$\therefore$ G.M. = Anti log ($\log_e$ G) = Antilog (0.3862) = 1.4713

If H is H.M., then

$$\frac{1}{H} = E\left(\frac{1}{H}\right) = \int_1^2 \frac{1}{x} f(x)dx = \log\ x\Big|_1^2$$

$$= \log_e{}^2 - \log_e{}^1 = 0.6932$$

$\therefore$ $H = \dfrac{1}{0.6931} = 1.4427$

Hence A.M. = 1.5; G.M. = 1.4713; H.M. = 1.4427

or A.M. > G.M. > H.M. and S.D. = $\sqrt{\dfrac{1}{12}}$

Example 5: If X and Y are independent uniform (rectangular) variates on [0, 1], find the distribution of (i) $\dfrac{x}{y}$ and (ii) xy.

Sol. We are given that X ∽ ∪ [0, 1] and Y ∽ ∪ [0, 1], therefore the p.d.f. is of X and Y can be written as $f_1(x) = 1,\ 0 \le x \le 1$ and $f_2(y) = 1,\ 0 \le y \le l$.

Since X and Y are independent variables, the joint p.d.f. of X and Y is given by

$f(x, y) = f_1(x)\ f_2(y) = 1,\ 0 \le x \le 1,\ 0 \le y \le 1$ ......(1)

(i) To find the p.d.f. of $\dfrac{x}{y}$, we transform the system in terms of U and V where $u = \dfrac{x}{y}$ and V = y i.e. x = u v and y = v

$$\text{Here}\quad J = \frac{\partial(x,y)}{\partial(u,v)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial y}{\partial u} \\ \frac{\partial x}{\partial v} & \frac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} v & 0 \\ u & 1 \end{vmatrix} = V$$

The limits x = 0 maps to u = 0, v = 0;

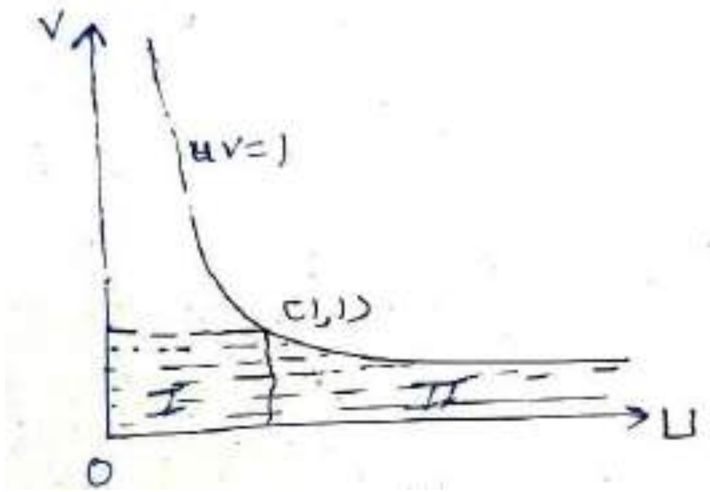
x = 1 maps to uv = 1 (Rectangular hyperbola); y = 0 maps to v = 0; and y = 1 maps to v = 1

Then the joint p.d.f. of U and V is given by

$y(u, v) = f(x, y)\ |J|$

$= v,\ 0 \leq u < \infty,\ 0 \leq v \leq 1$ with $uv = 1$

These limits can be shown in the figure below as:



Now to find marginal p.d.f. of U, we integrate g(u,v) with respect to v, within the specified limits. When we integrate w.r.t. v, the area under curve consists of two regions I and II as shown in figure.

In region I:

$$g_1(u) = \int_0^1 g(u,v)dv = \int_0^1 v\,dv = \frac{1}{2},\ 0 \leq u \leq 1$$

In region II:

$$g_1(u) = \int_0^{1/u} g(u,v)\,dv = \int_0^{1/u} v\,dv = \frac{1}{2u^2},\ 1 < u < \infty$$

Hence the distribution of $U = \frac{x}{y}$ is given by

$$g_1(u) = \begin{cases} \frac{1}{2}, 0 \leq u \leq 1 \\ \frac{1}{2u^2}, 1 < u < \infty \end{cases}$$

(ii) To find the p.d.f. of XY, we transform the system (1) in terms of U and V defined as u = xy and v = x

i.e. $x = v,\ y = \frac{u}{v}$

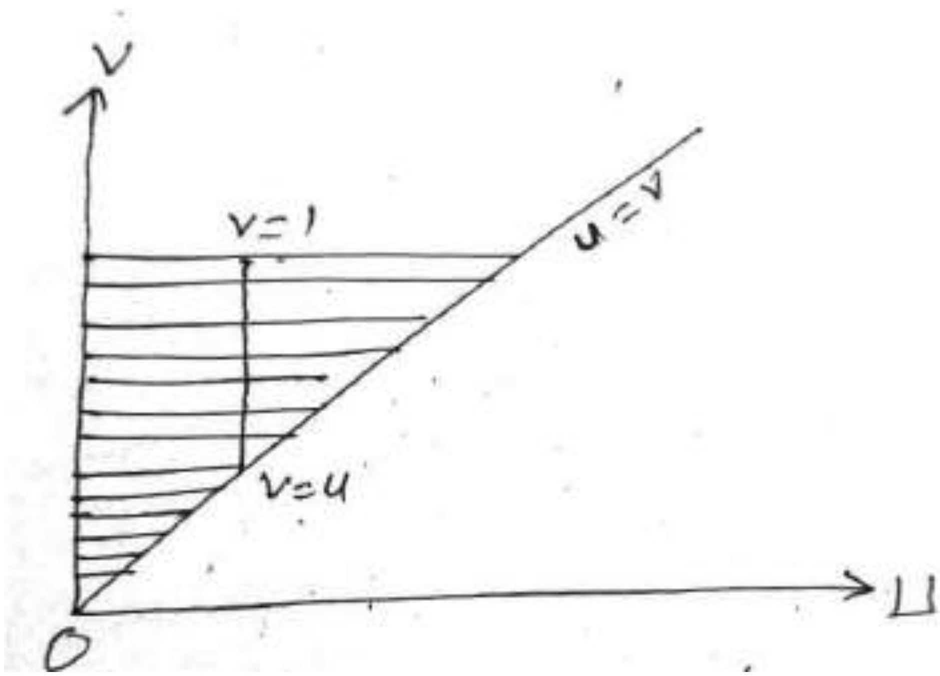
Here $\quad J = \dfrac{\partial(x,y)}{\partial(u,v)} = \begin{vmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial y}{\partial u} \\ \dfrac{\partial x}{\partial v} & \dfrac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} 0 & \dfrac{1}{v} \\ 1 & \dfrac{-u}{v^2} \end{vmatrix} = -\dfrac{1}{v}$

The limits x = 0 maps to v = 0 ; x = 1 maps to v = 1 ; y = 0 maps to u = 0 and y = 1 maps to u = v. The joint p.d.f. of U and V is given by

$g(u, v) = f(x, y)\, |J| = \dfrac{1}{v}$; $0 < u < 1$, $0 < v < 1$ with $u = v$

Now, to find marginal p.d.f. of U, we integrate g(u, v) with respect to v as below:-

$g_1(u) = \displaystyle\int_u^1 \frac{1}{v}\,dv = \left|\log v\right|_u^1 = -\log u,\ 0 < u < 1$



Example 6: Find the density function f(x) corresponding to the characteristic function

$$\phi(t) = \begin{cases} 1-|t|, & |t| \le 1 \\ 0, & |t| > 1 \end{cases}$$

Sol. We know that

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{=itx}\phi(t)\,dt = \frac{1}{2\pi}\int_{-1}^{1} e^{-itx}(1-|t|)\,dt$$

$$= \frac{1}{2\pi}\left[\int_{-1}^{0} e^{-itx}(1+t)dt + \int_{0}^{1} e^{-itx}(1-t)dt\right] \qquad ....(1)$$

$$= \frac{1}{2\pi}\left[\left[\left\{(1+t)\frac{e^{-itx}}{-ix} - \left(\frac{e^{-itx}}{i^2x^2}\right)\right\} + \left\{(1-t)\frac{e^{-itx}}{-ix} + \left(\frac{e^{-itx}}{i^2x^2}\right)\right\}\right]_0^1\right]$$

$$= \frac{1}{2\pi}\left[\left[\left\{\frac{1}{-ix} - \frac{1}{-x^2}\right\}\right] - \left\{0 - \frac{e^{-ix}}{-x^2}\right\} + \left\{0 + \frac{e^{-ix}}{-x^2}\right\} - \left[\left\{\frac{1}{-ix} + \frac{1}{-x^2}\right\}\right]\right]$$

$$= \frac{1}{2\pi}\left[-\frac{1}{ix} + \frac{1}{x^2} - \frac{e^{ix}}{x^2} - \frac{e^{ix}}{x^2} - \frac{e^{-ix}}{x^2} + \frac{1}{ix} + \frac{1}{x^2}\right]$$

$$= \frac{1}{2\pi}\left[\frac{2}{x^2} - \frac{e^{ix} + e^{-ix}}{x^2}\right]$$

$$= \frac{1}{2\pi}\left[\frac{2}{x^2} - \frac{2\cos x}{x^2}\right]$$

$$= \frac{1}{\pi}\left[\frac{1-\cos x}{x^2}\right], \ -\infty \le x \le \infty$$

Self-Check Exercise

Q.1 If X is uniformly distributed with mean $\frac{1}{2}$ and variance $\frac{25}{12}$, find P(x > 0) and P(x < 1).

Q.2 Calculate the mean and variance of the rectangular distribution given by the probability density function f(x) = $\frac{1}{2h}$ in 10-h < x < 10 + h and 0 elsewhere. What is the distribution function of the variable X of which f(x) is p.d.f.?

Q.3 If X has uniform distribution in [0, 1]. find probability distribution function of -2 $\log_e$ X.

13.5 Summary

We conclude this unit by summarizing what we have covered in it:

1. Defined continuous uniform (rectangular) distribution over interval (a, b)
2. Discussed different characteristics of uniform distribution
3. Discussed in detail different properties of uniform distribution
4. To improve understanding of different results we did some examples also.

13.6 Glossary:

1. A random variable X is said to follow a continuous uniform (rectangular) distribution over interval (a, b) if its probability density function is constant k (say) over entire interval or range $f(x) = \begin{cases} k, a \le x \le b \\ 0, otherwise \end{cases}$

2. For $f(x)$ to be a probability density function, k = $\dfrac{1}{b-a}$

$$\therefore \qquad f(x) = \begin{cases} \dfrac{1}{b-a}, a < x < b \\ 0, otherwise \end{cases}$$

13.7 Answer to Self Check Exercise

Ans.1 $P(X > 0) = \dfrac{3}{5}$

$P(X < 1) = \dfrac{3}{5}$

Ans.2 Mean = 10

Variance = $\dfrac{1}{3}h^2$

Distribution function

$$F(x) = \begin{cases} \dfrac{x-(10-h)}{2h}; 10-h \le x \le 10+h \\ 1 \quad ; \qquad x \ge 10+h \end{cases}$$

Ans.3 Probability distribution function of -2 loge X = $\dfrac{1}{2} e^{\frac{-y}{2}}$; $0 < y < \infty$

13.8 References/Suggested Readings

1. Robert V. Hogg, Joseph w. Mckean and Allen T. craig, Introduction to Mathematical statistics, Pearson Education, Asia, 2007.

2. Irwim Miller, Marylees Miller and John E. Freund, Mathematical Statistics with Application, 7th Ed., Pearson Education, Asia, 2006.

3. Sheldon Ross, Introduction to Probability Model, 9th Ed., Academic Press, Indian Reprint, 2007.

13.9 Terminal Questions

1. A random variable X has a uniform distribution over (-3, 3). Compute:

   (i) $P(X < 2)$ and (ii) $P(|X-2| < 2)$

2. Let the current (in MA) measured in a copper wire follows uniform distribution over the interval [0, 20]. Write down the probability density function of random variable X representing the current. Also calculate mean, variance and cumulative distribution function of X.

3. If $X \backsim U(-a, a)$. Show that its m.g.f. is given by $M_x(t) = \frac{1}{at} \sinh(at)$ and also show that $\mu_{2n+1} = 0$, $\mu_{2n} = \frac{a^{2n}}{2n+1}$; n is a positive integer.

4. Find the cumulative generating function of the rectangular distribution and the first four cumulants.

Unit - 14

Normal Distribution

Structure

14.1 Introduction

The normal distribution occupies the central position in probability and statistics. The normal distribution is the most frequently used of all probability distributions. The normal distribution was first discovered in 1733 by English Mathematician De-Moivre, who obtained this continuous distribution as a limiting case of the binomial distribution and applied it to problems arising in the game of chance. The normal distribution is a continuous probability distribution that is symmetric about the mean, and has a bell-shaped curve. It is defined by two parameters mean and standard deviation. Many natural and man-made phenomena follow a normal distribution, such as heights, IQ scores, measurement errors etc. In statisfical inference, the sampling distribution of many test statistics (e.g. sample means) follow a normal distribution, enabling the use of powerful statistical methods. It is widely used to model continuous random variables in fields like finance, engineering, biology and more. Its mathematical properties make it tractable for analysis and modeling.

14.2 Learning Objectives

After studying the unit, you should be able to:

- Define Normal distribution; Beta-function; Gamma-function
- Find relation between Beta and Gamma function
- Find mean and variance of normal distribution
- Discuss mode of normal distribution
- Discuss median, moments about the mean and m.g.f. of normal distribution
- Find moments of normal distribution
- Prove some theorems related to normal variate.
- Discuss 95% confidence interval for the mean of the population
- Discuss Area property (Normal probability integral)

14.3 Some Definitions

Normal Distribution

A continuous random variables X is said to be normally distributed if its p.d.f. is given by

$$f(x;\mu,\sigma) = f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \; -\infty < x < \infty, -\infty < \mu < \infty, \sigma > 0.$$

Here μ (called mean) and σ^2 (called variance) are parameters.

Now we have to prove that $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$, $-\infty < x < \infty, \sigma > 0$ is a p.d.d.

Now $f(x) \geq 0$ for all x

and $\int_{-\infty}^{\infty} f(x)dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$

$$= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{-y^2}{2}} \sigma\, dy$$

[Put $\frac{x-\mu}{\sigma}$ = y so that dx = σ dy]

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{-y^2}{2}} dy$$

$$= \frac{1}{\sqrt{2\pi}} \sqrt{2\pi} = 1$$

$\therefore \quad \int_{-\infty}^{\infty} f(x) = 1$

$\therefore \quad f$(x) satisfies the conditions of being a p.d.f. of continuous random variable.

Hence f(x) is a p.d.f.

Beta-Function

Beta - function is defined as

B(l, m) = $\int_{0}^{1} x^{l-1}(1-x)^{m-1} dx$

Take l = $\frac{1}{2}$, m = $\frac{1}{2}$

$\therefore \quad B\left(\frac{1}{2}, \frac{1}{2}\right) = \int_{0}^{1} x^{-1/2}(1-x)^{-1/2} dx$

Put x = $\sin^2$ Q

$\therefore \quad B\left(\frac{1}{2}, \frac{1}{2}\right) = 2 \int_{0}^{\pi/2} 1\, do \Rightarrow B\left(\frac{1}{2}, \frac{1}{2}\right) = \pi$

Gamma - Function

Gamma function is defined as

Γ(l) = $\int_{0}^{\infty} e^{-x} x^{l-1} dx, l > 0$

It can be proved that

Γ(l) = (l - 1) Γ(l - 1)

Relation between Beta and Gamma Function

B (l, m) = $\frac{\Gamma(l)\Gamma(m)}{\Gamma(l+m)}$

$\therefore \quad B\left(\frac{1}{2},\frac{1}{2}\right) = \dfrac{\Gamma\frac{1}{2}\Gamma\frac{1}{2}}{\Gamma(1)}$

$\Rightarrow \quad \pi = \left[\Gamma\left(\frac{1}{2}\right)\right]^2$ $\qquad$ [$\because$ $\Gamma(1) = 1$]

$\Rightarrow \quad \Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$

Note 1: The normal distribution with mean μ and variance σ^2 is abbreviated with the symbol N(μ, σ^2).

Note 2: Since f(x) is a p.d.f.

$\therefore \quad \sum_{-\infty}^{\infty} f(x)dx = 1$

$\Rightarrow$ total area under the curve f(x) and above the x-axis is equal to 1.

Note 3: The graph of y = f(x) is known as the normal probability curve or simply normal curve. It is a bell-shaped curve. The top of the bell is directly above the mean μ. For large values of σ, the curve tends to flatten out and for small values of σ, it has a sharp peak.

14.4 Mean and Variance of Normal Distribution

We have

Mean = E(X) = $\int_{-\infty}^{\infty} x\, f(x)dx$

$= \int_{-\infty}^{\infty} x.\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$ dx

Put $\frac{x-\mu}{\sigma}$ = y $\qquad$ or $\qquad$ x - μ = σ y

$\therefore$ dx = σ dy

$\therefore$ Mean = $\int_{-\infty}^{\infty} (\mu+\sigma y)\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{-y^2}{2}}$.σ dy

= $\mu \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}$ dy + $\sigma \int_{-\infty}^{\infty}$ y $\frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}$ dy

$\therefore$ Mean = $\dfrac{\mu}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}}$ dy + $\dfrac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}}$ y dy ....(1)

Let $I_1 = \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}}$ dy

Also $I_1 = \int_{-\infty}^{\infty} e^{\frac{-x^2}{2}}$ dx

$\therefore$ $(I_1)^2 = \int_{-\infty}^{\infty} e^{-\frac{1}{2}}$ $(x^2 + y^2)$ dx dy

$= \int_{0}^{2\pi} \int_{0}^{\infty} r^{e^{-\frac{r^2}{2}}} dr\, d\theta$ (changing to polar coordinates)

$= \int_{0}^{2\pi} d\theta \int_{0}^{\infty} r^{e^{\frac{-r^2}{2}}} dr$

$= (2\pi - 0) \left[-e^{-r^2/2}\right]_0^{\infty}$

$= - 2\pi (- 1 - 0)$

$= 2\pi$

$\therefore$ $I_1 = \sqrt{2\pi}$

Let $I_2 = \int_{-\infty}^{\infty} e^{-\frac{-y^2}{2}}$ y dy

= 0 [$\because$ integrand is an odd function]

$\therefore$ from (1), we get

Mean = $\dfrac{\mu}{\sqrt{2\pi}} \sqrt{2\pi} + \dfrac{\sigma}{\sqrt{2\pi}} . 0$

$= \mu$

Var (X) = E $(X - \mu)^2 = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)$ dx

$= \int_{-\infty}^{\infty} (x-\mu)^2 . \dfrac{1}{\sigma\sqrt{2\pi}} . \pi^{e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}} dx$

Put $\frac{x-\mu}{\sigma}$ = y or x - μ = σ y

∴ dx = σ dy

∴ Var (X) = $\int_{-\infty}^{\infty} (\sigma y)^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{-y^2}{2}}$ dy

= $\frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{0} y^2 e^{-\frac{-y^2}{2}}$ dy

= $\frac{2\sigma^2}{\sqrt{2\pi}} \int_{0}^{\infty} y^2 e^{\frac{-y^2}{2}}$ dy [∵ integrand is even function]

Put $\frac{y^2}{2}$ = z or y^2 = 2z

∴ 2ydy = 2 dz ⇒ dy = $\frac{1}{y}$ dz ⇒ dy = $\frac{1}{\sqrt{2z}}$ dz

∴ Var (X) = $\frac{2\sigma^2}{\sqrt{2\pi}} \int_{0}^{\infty}$ (2z)e^{-z} $\frac{dz}{\sqrt{2z}}$

= $\frac{2\sigma^2}{\sqrt{\pi}} \int_{0}^{\infty} e^{-z} z^{\frac{1}{2}}$ dz

= $\frac{2\sigma^2}{\sqrt{\pi}} \int_{0}^{\infty} e^{-z} z^{\frac{3}{2}-1}$ dz

= $\frac{2\sigma^2}{\sqrt{\pi}} \Gamma\left(\frac{3}{2}\right) = \frac{2\sigma^2}{\sqrt{\pi}} \frac{1}{2} \Gamma\left(\frac{1}{2}\right)$

= $\frac{\sigma^2}{\sqrt{\pi}} \sqrt{\pi} = \sigma^2$

∴ Var (X) = σ^2

14.5 Mode of Normal Distribution

The p.d.f. of a normal distribution is given by

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

Taking logarithm, we have

$$\log f(x) = -\log\left(\sigma\sqrt{2\pi}\right) - \frac{1}{2\sigma^2}(x - \mu)^2$$

Differentiating w.r.t. x, we get

$$\frac{f'(x)}{f(x)} = -\frac{1}{2\sigma^2}(x - \mu).\ 2$$

or $\quad f'(x) = -\frac{x-\mu}{\sigma^2} f(x) \qquad$ ......(1)

Again differentiating w.r.t. x, we get

$$f''(x) = \frac{1}{\sigma^2}\left[\left(x-\mu\right) f'(x) + f(x)\right]$$

$$= -\frac{1}{\sigma^2}\left[-(x-\mu)\frac{(x-\mu)}{\sigma^2} f(x) + f(x)\right] \qquad [\because \text{ of } (1)]$$

$$= \frac{f(x)}{\sigma^2}\left[1-\left(\frac{x-\mu}{\sigma}\right)^2\right]$$

From (1), $f'(x) = 0 \Rightarrow x = \mu$

Also $[f''(x)]_{\text{at } x=\mu} = -\frac{1}{\sigma^2}[f(x)]_{\text{at } x=\mu}$

$$= \frac{1}{\sigma^2}\left\{\frac{1}{\sigma\sqrt{2\pi}}\right\} = -\frac{1}{\sigma^3}\frac{1}{2\pi} < 0$$

Thus, mode of the normal distribution is μ, the mean of the distribution.

14.6 Median of the Normal Distribution

Let $\propto$ be the median of the normal distribution

Then $\quad \int_{-\infty}^{\infty} f(x)dx = \frac{1}{2} = \int_{-\infty}^{\infty} f(x)dx$

or $\quad \int_{-\infty}^{\mu} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx + \int_{\mu}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx = \frac{1}{2}$

or $\quad \frac{1}{2} + \int_{\mu}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx = \frac{1}{2}$

[I^st integral may be evaluated be putting $x = \mu$]

or $$\int_{\mu}^{\propto} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx = 0$$

$\therefore \quad \propto = \mu$

Thus, the median of the normal distribution is equal to μ, the mean.

Hence, for the normal distribution all the three median, mode and mean coincide.

14.7 Moments About the Mean

In a normal distribution

(i) all the odd moments about the mean μ vanish.

and (ii) all the even moments about the mean are given by $\mu_{2n} = (2n - 1)\sigma^2 \mu_{2n-2}$.

The even order moments about the mean μ of the normal distribution are given by

$\mu_{2n} = (2n - 1)\sigma^2 \mu_{2n-2}$.

The even order moments about the mean μ of the normal distribution are given by

$$\mu_{2n} = \int_{-\infty}^{\infty} (x-\mu)^{2n} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

$$= \frac{\sigma^{2n}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z^{2n-1}.ze^{-\frac{1}{2}z^2} dz \left(\text{Putting Z=}\frac{x-\mu}{\sigma}\right)$$

$$= \frac{\sigma}{\sqrt{2\pi}} \left[-z^{2n-1}e^{-\frac{1}{2}z^2}\right]_{-\infty}^{\infty} + \frac{\sigma^{2n}}{\sqrt{2n}} (2n - 1) \int_{-\infty}^{\infty} z^{2n-1}e^{-\frac{1}{2}z^2} dz$$

$$= \frac{\sigma}{\sqrt{2\pi}} (0) + [2n - 1]\sigma^2 \mu_{2n-2}.$$

$$= (2n - 1)\sigma^2 \mu_{2n-2}$$

Thus, $\mu_{2n} = (2n - 1)\sigma^2.\mu_{2n-2}$

Successively applying the recurrence formula, we get

$\mu_{2n-2} = (2n - 3)\sigma^2 \mu_{2n-4}$

....

....

$\mu_4 = 3 . \sigma^2 \mu_2$

$\mu_2 = \sigma_0^2 \mu_0 = \sigma^2.1$

Thus $\mu_{2n} = (2n - 1)(2n - 3)..........3.1.\sigma^{2n}$

Hence $\mu_2 = \sigma^2$, $\mu_4 = 3\sigma^4$

The odd order moments about the mean μ, are given as

$$\mu_{2n+1} = \int_{-\infty}^{\infty} (x-\mu)^{2n+1} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

$$= \frac{\sigma^{2n+1}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z^{2n+1} e^{-\frac{1}{2}z^2} dz \qquad \left[\text{Putting } Z = \frac{x-\mu}{\sigma}\right]$$

$$= \frac{\sigma^{2n+1}}{\sqrt{2\pi}} \times 0 \qquad [\because \text{ Integrant is an odd function of z}]$$

$$= 0$$

14.8 Moment Generating Function of Normal Distribution

We have

$$M_X(t) = E(e^{tx}) = \int_{-\infty}^{\infty} e^{tx} f(x)dx = \int_{-\infty}^{\infty} e^{tx} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2\sigma^2}[(x-\mu)^2 - 2\sigma^2 tx]} dx$$

Now $(x - \mu)^2 - 2\sigma^2 t x = x^2 - x(\mu + \sigma^2 t) x + \mu^2$

$= [x - (\mu + \sigma^2 t)]^2 + \mu^2 - (\mu + \sigma^2 t)^2$

$= [x - (\mu + \sigma^2 t)]^2 - 2\mu\sigma^2 t - \sigma^4 t^2]$

$$\therefore \qquad M_x(t) = e^{\mu t} + \frac{\sigma^2 t^2}{\underline{|2}} \left\{ \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}\left\{\frac{x-(\mu+\sigma^2 t)}{\sigma}\right\}^2_{dx}} \right\}$$

$$\Rightarrow \qquad M_x(t) = e^{ut + \frac{\sigma^2 t^2}{\underline{|2}}} \left[\because \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}\left\{\frac{x-(\mu+\sigma^2 t)}{\sigma}\right\}^2_{dx}} = 1 \right]$$

Cor. Put $\mu = 0$, $\sigma^2 = 1$

$$\therefore \qquad M_x(t) = e^{\frac{t^2}{\underline{|2}}}$$

$\therefore$ If X is a standard normal variate i.e. X is N (0, 1), then its p.d.f. and m.g.f. are respectively given by

$f(x) = \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}}$, $-\infty < x < \infty$ and $M_x(t) = e^{\frac{t^2}{\lfloor 2}}$

14.9 Moments of Normal Distribution

The m.g.f. (about mean) is given by $E[e^{t(x-\mu)}] = e^{-\mu t} E(e^{tx}) = e^{-\mu t} M_x(t)$ where $M_x(t)$ is the m.g.f. (about origin)

$\therefore$ Moment generating function (about mean)

$$= e^{-\mu t} e^{\mu t + t^2\sigma^2/2} = e^{t^2\sigma^2/2}$$

$$= \left[1 + (t^2\sigma^2/2) + \frac{(t^2\sigma^2/2)^2}{\lfloor 2} + \frac{(t^2\sigma^2/2)^3}{\lfloor 3} + ... + \frac{(t^2\sigma^2/2)^n}{\lfloor n} + ...\right] \quad ...(1)$$

The coefficient of $\frac{t^r}{r!}$ in (1) gives μ_r, the rth moment about mean. Since there is no term with odd powers of t in (1), all moments of odd order about mean vanish.

i.e. $\mu_{2n+1} = 0$; n = 0, 1, 2, ...(2)

and μ_{2n} = coefficient of $\frac{t^{2n}}{\lfloor (2n)}$ in (1)

$$= \frac{\sigma^{2n} \times \lfloor 2n}{2^n . \lfloor n}$$

$$= \frac{\sigma^{2n}}{2^n . \lfloor n} [2n (2n - 1) (2n - 2) .... 4.3.2.1]$$

$$= \frac{\sigma^{2n}}{2^n . \lfloor n} [1. 3. 5. ..... (2n -1)] [2.4.6. .... (2n -2) 2n]$$

$$= \frac{\sigma^{2n}}{2^n . \lfloor n} [1.3.5. ..... (2n -1)] 2^n [1.2.3. ..... 4]$$

$\Rightarrow$ $\mu_{2n} = 1.3.5. .... (2n -1) \sigma^{2n}$ ....(3)

14.10 Important Theorems

Theorem 1 : If x is a normal variate with mean μ and variance $\sigma^2 > 0$, then show that a new random variable z defined by $Z = \frac{x-\mu}{\sigma}$ is a variate with mean 0 and variance 1. Also find the m.g.f. of Z.

Proof : Here Z = $\frac{x-\mu}{\sigma}$

$\therefore$ Mean = E(Z) = E $\left(\frac{x-\mu}{\sigma}\right) = \frac{1}{\sigma}$ E(X - μ)

$= \frac{1}{\sigma}[E(x) - \mu\, E(1)] = \frac{1}{\sigma}[\mu - \mu\,(1)] = 0$

Var (z) = Var $\left(\frac{x-\mu}{\sigma}\right) = \frac{1}{\sigma^2}$ Var (x - μ)

$= \frac{1}{\sigma^2}$ Var (x) $= \frac{1}{\sigma^2}\,\sigma^2 = 1$

The m.g.f. of Z is

$M_z(t) = M_{\frac{x-\mu}{\sigma}}(t) = e^{\frac{-\mu t}{\sigma}} M_x\left(\frac{t}{\sigma}\right)$

$= e^{\frac{-\mu t}{\sigma}} . e^{\frac{\mu t}{\sigma}} + \frac{t^2}{\underline{|2}}$

$= e^{\frac{t^2}{\underline{|2}}}$

Theorem 2 : If $x_1, x_2, x_3, \ldots\ldots. x_n$ are independent variables having the same distribution with the mean μ and the variance σ^2, then

$E(\overline{x}) = \mu$ and $Var(\overline{x}) = \frac{\sigma^2}{n}$.

Proof : $\overline{x} = \frac{x_1 + x_2 + \ldots. + x_n}{n}$

$E(\overline{x}) = E\left(\frac{x_1 + x_2 + \ldots. + x_n}{n}\right) = \frac{1}{n} E(x_1, x_2, x_3, \ldots\ldots. x_n)$

$= \frac{1}{n}[E(x_1) + E(x_2) + \ldots\ldots. + E(x_n)]$

$= \frac{1}{n}[\mu + \mu + \ldots\ldots + \mu]$

$= \frac{1}{n}[\mu + \mu + \ldots\ldots +$ to x terms]

$$= \frac{1}{n} \cdot n\mu = \mu$$

$$\text{Var}(\bar{x}) = \text{Var}\left(\frac{x_1 + x_2 + \dots + x_n}{n}\right)$$

$$= \frac{1}{n^2} \text{Var}(x_1 + x_2 + \dots\dots x_n)$$

$$= \frac{1}{n^2} [\text{Var}(x_1) + \text{Var}(x_2) + \dots\dots + \text{Var}(x_n)]$$

$$= \frac{1}{n^2} [\sigma^2 + \sigma^2 + \dots\dots + \sigma^2]$$

$$= \frac{1}{n^2} \cdot n\sigma^2$$

$$= \frac{\sigma^2}{n}$$

14.11 95% Confidence Interval for the Mean of the Population

To introduce the idea of a confidence interval by mean of an example, let us refer to the sample distribution $\bar{x}$ for random samples of size n from a normal population with the mean μ and the known variance σ^2.

$\bar{x}$ is distributed normally with the mean μ and the variance $\frac{\sigma^2}{n}$ i.e. x is

$N\left(\mu, \frac{\sigma^2}{n}\right)$ or $Z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$ is N (0, 1)

Now P(-1.96 < Z < 1.96) 0.95

or $P\left(-1.96 < \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} < 1.96\right) = 0.95$

or $P\left(\bar{x} - 1.96 \frac{\sigma}{\sqrt{n}} < \mu \; \bar{x} + 1.96 \frac{\sigma}{\sqrt{n}}\right) = 0.95$

where $\bar{x}$ is value of $\bar{x}$ which we actually obtained in a sample.

Hence 95% confidence interval for the mean (μ) of the population is

$$\left(\bar{x} - 1.96 \frac{\sigma}{\sqrt{n}}, \; \bar{x} + 1.96 \frac{\sigma}{\sqrt{n}}\right)$$

14.12 Area Property (Normal Probability Integral)

If X ~ N (μ, σ^2), then the probability that random value of x will be between x = μ and x = x_1 is given by :

$$P(\mu < x < x_1) = \int_{\mu}^{x_1} f(x)dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{\mu}^{x_1} e^{-(x-\mu)^2/2\sigma^2} dx$$

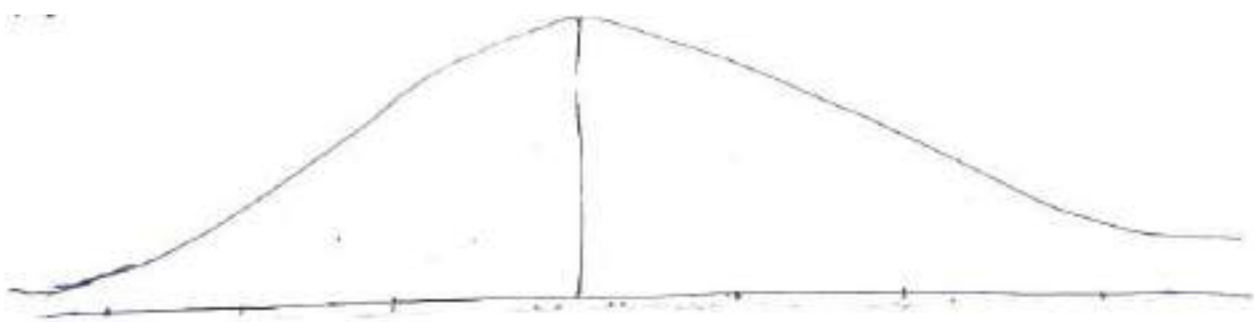
Put $\dfrac{x-\mu}{\sigma}$ = z $\Rightarrow$ x - μ = σz

Also when x = μ, z = 0 and X = x, Z = $\dfrac{x_1-\mu}{\sigma}$ = z_1 (say)

$\therefore$ $$P(\mu < X < x_1) = P(0 < Z < z1) = \frac{1}{\sqrt{2\pi}} \int_0^{z_1} e^{-z^2/2} dz$$

$$= \int_0^{z_1} \phi(z)dz$$

Where $\phi(z) = \dfrac{1}{\sqrt{2\pi}} e^{-z^2/2}$, is the probability density function of the standard normal variate. The definite integral $\int_0^{z_1} \phi(z)dz$ is known as normal probability integral and gives the area under standard normal crurve between the ordinates Z = 0 and Z = z1,



x=μ-3σ	x=μ-2σ	x=μ-σ	x=μ	x=μ+σ	x=μ+2σ	x=μ+3σ
Z= -3	Z= -2	Z= -1	Z= 0	Z= 1	Z= 2	Z= 3

In particular

(i) The probability that a random value of x lies in the interval (μ - σ, u + σ) is given by

$$P(\mu - \sigma < x < u + \sigma) = \int_{\mu-\sigma}^{\mu+\sigma} f(x)dx$$

$$\Rightarrow \quad P(-1 < Z < 1) = \int_{-1}^{1} \phi(z)dz$$

$$[\therefore z = \frac{x-\mu}{\sigma}]$$

$= 2 \int_0^1 \phi(z)dz$ [By symmetry]

$= 2 \times 0.3413 = 0.6826$ (From tables)

(ii) $P(\mu - 2\sigma < x < u + 2\sigma)$

$= P(-2 < Z < 2) = \int_{-2}^{2} \phi(z)dz$

$= 2 \int_0^2 \phi(z)dz$ [By symmetry]

$= 2 \times 0.4772 = 0.9544$

(iii) $P(\mu - 3\sigma < x < u + 3\sigma) = P(-3 < Z < 3)$

$= \int_{-3}^{3} \phi(z)dz$

$= 2 \int_0^3 \phi(z)dz = 2 \times 0.498665$

$= 0.9973$

Thus the probability that a normal variate x lies outside the range $\mu \pm 3\sigma$ is given by :

$P(|x - \mu| > 3\sigma) = P(|z| > 3) = 1 - P(-3 \leq Z \leq 3)$

$= 0.0027$

Let us improve our understanding of these results by looking at some following examples:

Example 1 : Suppose the height of 1000 soldiers in a regiment are distributed normally with mean 68.22 inches and variance 10.8 inches. How many soldiers have height-

(i) Over 6 feet (ii) below 5.5 feet?

Sol. : We have been given that

$\mu = 68.22$ and $\sigma = \sqrt{10.8} = 3.29$

(i) for x = 6 feet = 72 inches

$$\therefore z = \frac{x-\mu}{\sigma} = \frac{72-68.22}{3.29} = \frac{3.78}{3.29} = 1.15$$

$\therefore P(x > 72) = P(z > 1.15) = 0.5 - 0.3749$

$= 0.1251$

$\therefore$ Number of soldiers with height more than 72 inches

= 1000 x 0.1251 = 125.1 = 125

(ii) for x = 5.5 feet = 65 inches

$\therefore$ z = $\frac{x-\mu}{\sigma}$ = $\frac{65-68.22}{3.29}$ = - $\frac{3.22}{3.29}$ = -0.9787

$\therefore$ P(x < 65) = P (z < -0.9787) = 0.5 - P(0$\leq$ z $\leq$ 0.98)

= 0.5 -03365 = 0.1635

$\therefore$ Number of soldiers with height below 65 inches

= 1000 $\times$ 0.1635 = 163.5 = 163

Example 2 : Estimate the 95% confidence interval for the mean of a normal population having the variance σ2 = 100 and a random sample of size n = 25 yields $\overline{x}$ = 67.33

Sol. We have been gives that

σ^2 = 100 , n = 25, $\overline{x}$ = 67.53

95% confidence interval for the mean μ of the normal population is

$$\left(\overline{x}-1.96\frac{\sigma}{\sqrt{n}}, \overline{x}+1.96\frac{\sigma}{\sqrt{n}}\right)$$

or $$\left(67.53-(1.96)\left(\frac{10}{5}\right), 67.53+(1.96)\left(\frac{10}{5}\right)\right)$$

or 67.53 - 3.92, 67.53 + 3.92)

or (63.61, 71.45)

Example 3 : Let x be a normal random variable with mean 10 and standard deviation 4. Determine the probability.

(i) P(12 $\leq$ x $\leq$ 15) (ii) P(x $\geq$ 7)

Sol. : (i) Here μ = 10, σ = 4

For x = 12, we have

z = $\frac{x-\mu}{\sigma}$ = $\frac{12-10}{4}$ = $\frac{2}{4}$ = $\frac{1}{2}$ = 0.5

For z = 15, we have

z = $\frac{x-\mu}{\sigma}$ = $\frac{15-10}{4}$ = $\frac{5}{4}$ = 1.25

$\therefore$ P(12 $\leq$ x $\leq$ 15) = P(0.5 $\leq$ z $\leq$ 1.25)

= P(0 $\leq$ z $\leq$ 1.25) - P(0 $\leq$ z $\leq$ 0.5)

= 0.3944 - 0.1415 = 02029

(ii) Here μ = 10, σ = 4

For x = 7, we have

z = $\frac{x-\mu}{\sigma}$ = $\frac{7-10}{4}$ = - $\frac{3}{4}$ = -0.75

$\therefore$ P (x $\geq$ 7) = P(z $\geq$ -0.75)

= P(-0.75 $\leq$ z $\leq$ 0) + P (z > 0)

= 0.2734 + 0.5

= 0.7734

Example 4 : Let x denote the number of scores of a feet. It is normally distributed with mean 100 and standard deviation 15, find the probability that x does not exceed 130.

Sol. : Here μ = 100, σ = 15

Let x be changed into standard normal variable z.

$\therefore$ z = $\frac{x-\mu}{\sigma}$ = $\frac{x-100}{15}$

When x = 130, z = $\frac{130-100}{15}$ = $\frac{30}{15}$ = 2

Now P(x $\leq$ 130) = P (z $\leq$ 2) = F(2) = 0.9772

Example 5 : Find the mean and S.D. of an examination in which grade of 70 and 88 corresponding to standard score of -0.6 and 1.4 respectively.

Sol. : Let z be the normal variable corresponding to x when μ is mean and σ is standard deviation.

$\therefore$ x = 70, z = -0.6 and x = 88, z = 1.4

$\therefore$ z = $\frac{x-\mu}{\sigma}$ gives

- 0.6 = $\frac{70-\mu}{\sigma}$ $\Rightarrow$ 70 - μ = -0.6σ ....(1)

and 1.4 = $\frac{88-\mu}{\sigma}$ $\Rightarrow$ 88 - μ = 1.4 σ ...(2)

Subtracting (1) from (2), we get

$18 = 2\sigma \Rightarrow \sigma = 9$

$\therefore$ From (1), $70 - \mu = -5.4 \Rightarrow \mu$ 75.4

$\therefore$ $\mu = 75.4$ and $\sigma = 9$

Example 6 : The marks obtained by students in an examination are normally distributed. If 10% students have marks more than 75 and 60% have marks more than 50, find mean and variance of the distribution.

Sol. Let the mean be μ and S.D. by σ.

Probability of students getting marks more than 75 = $\frac{10}{100}$ = 0.1

Probability of students getting marks between mean and 75 = 0.5 - 0.1 = 0.4

$\therefore$ Value of Z corresponding to area 0.4 from the table is Z, = 1.29

Similarly probability of getting marks more than 50 = $\frac{60}{100}$ = 0.6

$\therefore$ Probability of getting marks between mean and 50 = 0.6 - 0.5 = 0.1

$\therefore$ Value of Z corresponding to area 0.1 is $Z_2 = 0.25$

Now $Z_1 = \frac{X-\mu}{\sigma}$

or $75 = \mu = 1.29$ .....(1)

and $Z_2 = \frac{50-\mu}{\sigma}$

$\therefore$ $-0.25 = \frac{50-\mu}{\sigma}$

or $50 - \mu = -0.25\sigma$ .....(2)

Subtracting (2) from (1), we get

$25 = 1.54\sigma \Rightarrow \sigma = 16.23$

From (1), $75 - \mu = (1.29)(16.23)$

or $75 - \mu = 20.94$

or $\mu = 75 - 20.94$

$= 54.06$

$\therefore$ Mean $\mu = 54.06$ and $\mu = 16.23$

Example 7: If X has the m.g.f. $\exp(2t + 32t^2)$, then find the mean and variance of X.

Sol. Given m.g.f. of X is $M_x(t) = e^{2t+32t^2}$ .....(1)

We know that m.g.f. of normal variable X with mean μ and variance σ^2 is

$$M_x(t) = e^{\mu t + \frac{\sigma^2 t^2}{\lfloor 2}} \quad ......(2)$$

From (2) and (1), we get

$$e^{\mu t + \frac{\sigma^2 t^2}{\lfloor 2}} = e^{2t+32t^2} \quad \Rightarrow \quad \mu = 2, \frac{\sigma^2}{\lfloor 2} = 32$$

$\therefore$ $\mu = 2, \sigma^2 = 32 \times 2 \times 1 \Rightarrow \mu = 2, \sigma^2 = 64$

$\therefore$ Mean = 2 and Variance = 64

Self-Check Exercise

Q.1 If the level of education among adults in a certain region is normally distributed with mean 8 and S.D. 5, what is the probability that in a sample of 100 adults, you will find an average level of education

(i) between 10 to 14 years

(ii) more than 14 years

Q.2 If X is a normal variable with mean 25 and standard deviation 5, find the probability that

(i) $X \leq 10$ (ii) $15 \leq X \leq 30$ (iii) $|X - 30| \geq 10$

Q.3 In a sample of 200 cases, the mean of a certain test is 14 and standard deviation is 2.5 Assuming normal distribution, find

(i) how many condidated score between 12 and 15?

(ii) how many score below 8?

[Given Z : 0.4 0.8 2.4 P(Z) : 0.1554 0.2881 0.4918]

Q.4 In a normal distribution 31% items are under 45 and 8% are over 64. Find the mean and standard deviation of the distribution.

14.13 Summary

We conclude this unit by summarizing what we have covered in it:

1. Defined normal distribution; Beta function; Gamma - function.
2. Derived the relation between Beta and Gamma function.
3. Discussed mean and variance of normal distribution and derived the formulae for these.

4. Discussed mode, median, moments about the mean, m.g.f. and moments of normal distribution and derived the formulae for these.
5. Proved some theorems related to normal variate.
6. Discussed in detail 95% confidence interval for the mean of the population
7. Discussed in detail the area property normal property integral.
8. Did some examples related to each topic so that the contents be clarified further.

14.14 Glossary:

1. A continuous random variable X is said to be normally distributed if its p.d.f. is given by

$$f(x;\mu,\sigma) = f(x) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2},\ \infty < x < \infty,\ -\infty < \mu < \infty,\ \sigma > 0$$

2. Beta - function is defined as

$$B(l, m) = \int_0^1 x^{l-1}(1-x)^{m-1}dx$$

3. Gamma - function is defined as

$$\Gamma(l) = \int_0^\infty e^x x^{l-1}dx,\ l > 0$$

4. If X is a normal variate with mean μ and variance $\sigma^2 > 0$, then a new random variable Z defined by $Z = \frac{x-\mu}{\sigma}$ is a variate with mean 0 and variance 1.

14.15 Answer to Self Check Exercise

Ans.1 (i) Number of students between 10 to 14 years = 23

(ii) Number of students with more than 14 years = 12

Ans. 2 (i) P $(X \leq 10) = 0.0013$

(ii) P $(1\ X - 301 \geq 10) = 0.16$

Ans. 3 (i) Number of students scoring between 12 and 15 = 89

(ii) Number of students scoring less than 8

Ans. 4 $\mu = 50$ and $\sigma = 10$

14.16 References/Suggested Readings

1. Robert V. Hogg, Joseph w. Mckean and Allen T. craig, Introduction to Mathematical statistics, Pearson Education, Asia, 2007.

2. Irwim Miller, Marylees Miller and John E. Freund, Mathematical Statistics with Application, 7th Ed., Pearson Education, Asia, 2006.

3. Sheldon Ross, Introduction to Probability Model, 9th Ed., Academic Press, Indian Reprint, 2007.

14.17 Terminal Questions

1. In a distribution, exactly normal, 7% of the items are under 35 and 89% are under 63. What are the mean and standard deviation of the distribution?

2. The mean weight of 500 male students at a certain college is 65.6 kg and the standard deviation is 10 kg. Assuming that the weights are normally distributed, find how many students weigh

   (i) more than 75.5 kg

   (ii) between 55.5 and 75.5 kg

3. In a statistics examination the mean score was 78 and S.D. was 10.

   (i) Determine standard score of school boys whose score was 93 and 62 respectively.

   (ii) Determine the score of students standard deviation of whose score was -0.6 and 1.4 respectively

4. Estimate the 95% confidence interval for the mean of a normal population having the variance $\sigma_2 = 225$ and a random size of n = 20 yields $\overline{x} = 64.3$

Unit - 15

Exponential Distribution

Structure

15.1 Introduction

The exponential distribution is a fundamental probability distribution in statistics and probability theory. It is widely used to model the time between independent, randomly occurring events, such as the time between arrivals in a Poisson process or the time between failures in reliability engineering. It is a continuous probability distribution, meaning it can take on any value greater than or equal to 0-. It has the memory less property, which means that the probability of an event occurring in a given time interval is independent of the time elapsed since the last event. This makes the exponential distribution suitable for modeling processes where the occurrence of events is random and independent of time. This distribution is characterized by a single parameter, known as the rate parameter (λ), which determines the average rate of occurrence of the events.

It is commonly used to model the time between events in a wide range of fields, such as reliability engineering, queuing theory, and survival analysis. It is closely related to the Poisson process which is a model for the occurrence of independent, randomly occurring events. It is mathematically tractable, meaning that it has simple and well-understood properties, which makes it useful for analytical and computational purposes in various fields of study.

15.2 Learning Objectives

After studying this unit, you should be able to:

- Define exponential distribution

- Discuss different properties of exponential distribution. Properties like distribution function, moment generating function, characteristic function, cumulant generating function, Quartiles, mean deviation about mean, you should able to discuss
- Discuss some important theorems of exponential random variables.

15.3 Exponential Distribution-Definition

A continuous random variable X assuming non-negative values is said to follow an exponential distribution with parameter $\lambda > 0$, if its probability density function (p.d.f.) is given by

$$f(x) = \begin{cases} \lambda e^{-\lambda x} \ ; \ x > 0 \\ 0 \ ; otherwise \end{cases}$$

A random variable is called exponential random variable if its probability function follows exponential distribution.

15.4 Properties of Exponential Distribution

Property I. Distribution Function:- The distribution function of exponential random variable is given by

$$F_x(x) = P(X \leq x) = \int_0^x f(x)dx = \lambda \int_0^x e^{-\lambda x}dx = 1 - e^{-\lambda x}$$

$$\therefore \quad F_x(x) = \begin{cases} 1 - e^{-\lambda x} \ ; \ x > 0 \\ 0 \ ; otherwise \end{cases}$$

Property II. Moment Generating Function: The m.g.f. of the exponential distribution about origin is given by

$$M_x(t) = E(e^{tx}) = \int_0^\infty e^{tx} f(x)dx$$

$$= \lambda \int_0^\infty e^{tx} e^{-\lambda x} dx = \lambda \int_0^\infty e^{-(\lambda - t)x} dx$$

$$= \frac{\lambda}{\lambda - t}$$

$$\therefore \quad M_x(t) = \frac{\lambda}{\lambda\left(1 - \frac{t}{\lambda}\right)} = \left(1 - \frac{t}{\lambda}\right)^{-1}$$

$$= 1 + \frac{t}{\lambda} + \left(\frac{t}{\lambda}\right)^2 + \left(\frac{t}{\lambda}\right)^3 + \ldots\ldots\ldots$$

$$= \sum_{r=0}^{\infty} \left(\frac{t}{\lambda}\right)^r$$

$\therefore$ $$M_x(t) = \left(1 - \frac{t}{\lambda}\right)^{-1} = \sum_{r=0}^{\infty} \left(\frac{t}{\lambda}\right)^r$$

Property III: The constants of exponential distribution are :

Mean = $\frac{1}{\lambda}$; Variance = $\frac{1}{\lambda^2}$; coefficient of skewness V_1 = 2 and coefficient of kurtosis V_2 = 6

Property IV: Characteristic Function: The characteristic function of the exponential distribution about origin is given by

$$\phi_x(t) = E[e^{itx}]$$

$$= \int_0^{\infty} e^{itx} f(x)dx = \lambda \int_0^{\infty} e^{itx}.e^{-\lambda x} dx$$

$$= \lambda \int_0^{\infty} e^{-(\lambda - it)x} dx = \frac{\lambda}{\lambda - it}$$

$\therefore$ $$\phi_x(t) = \frac{\lambda}{\lambda - it} = \frac{\lambda}{\lambda\left(1 - \frac{it}{\lambda}\right)} = \left(1 - \frac{it}{\lambda}\right)^{-1}$$

$$= 1 + \left(\frac{it}{\lambda}\right) + \left(\frac{it}{\lambda}\right)^2 + \left(\frac{it}{\lambda}\right)^3 + \ldots\ldots\ldots\ldots$$

$$= \sum_{r=0}^{\infty} \left(\frac{it}{\lambda}\right)^r$$

$\therefore$ $$\phi_x(t) = \sum_{r=0}^{\infty} \left(\frac{it}{\lambda}\right)^r$$

Property V: The cumulant generating function:

The cumulant generating function is

$$K_x(t) = \log M_x(t) = -\log \left(1 - \frac{t}{\lambda}\right)$$

$$= \left(\frac{t}{\lambda}\right) + \frac{1}{2}\frac{t^2}{\lambda^2} + \ldots\ldots + \frac{1}{r}\frac{t^2}{\lambda^2} + \ldots\ldots\ldots\ldots$$

or $\quad K_x(t) = \frac{1}{\lambda} . t + \frac{1}{\lambda^2} . \frac{t^2}{\lfloor 2} + \frac{\lfloor 2}{\lambda^3} . \frac{t^3}{\lfloor 3} + \frac{\lfloor 3}{\lambda^4} . \frac{t^4}{\lfloor 4} + ...+ \frac{\lfloor r-1}{\lambda^r} . \frac{t^r}{\lfloor r} + ...$

$\therefore \quad K_r = \frac{\lfloor r-1}{\lambda^r}$; r = 1, 2, 3,.......

Mean = coeff of t = $K_1 = \frac{1}{\lambda}$

Variance = coeff of $\frac{t^2}{\lfloor 2} = K_2 = \frac{1}{\lambda^2} = \mu_2$

Coeff. of $\frac{t^3}{\lfloor 3} = K_3 = \mu_3 = \frac{2}{\lambda^3}$

Coeff. of $\frac{t^4}{\lfloor 4} = K_4 = \frac{\lfloor 3}{\lambda^4} = \frac{6}{\lambda^4}$

or $\quad \mu_4 = K_4 + 3K_2^2 = \frac{9}{\lambda^4}$

Property VI: Quartiles:

Here $f(x) = \lambda e^{-\lambda x}$, x > 0

If Q_1 and Q_3 be the lower and upper quartiles respectively, then for the lower quartile

$$\int_0^{Q_1} f(x) = \frac{1}{4}$$

$$\Rightarrow \quad \int_0^{Q_1} \lambda e^{-\lambda x} dx = \frac{1}{4}$$

$$\Rightarrow \quad \left[\lambda \frac{e^{-\lambda x}}{-\lambda}\right]_0^{Q_1} = \frac{1}{4}$$

$$\Rightarrow \quad \left(-e^{-\lambda Q_1} + 1\right) = \frac{1}{4}$$

$$\Rightarrow \quad -e^{-\lambda Q_1} = \frac{3}{4}$$

$$\Rightarrow \quad Q_1 = -\frac{1}{\lambda} \log_e \frac{3}{4}$$

For the upper quartile

$$\int_0^{Q_3} f(x)dx = \frac{3}{4}$$

$$\Rightarrow \qquad \lambda \int_0^{Q3} e^{-\lambda x} dx = \frac{3}{4}$$

$$\Rightarrow \qquad \left[\lambda \frac{e^{-\lambda x}}{-\lambda} \right]_0^{Q_1} = \frac{3}{4}$$

$$\Rightarrow \qquad -e^{-\lambda Q_3} + 1 = \frac{3}{4}$$

$$\Rightarrow \qquad e^{-\lambda Q_3} = \frac{1}{4}$$

$$\Rightarrow \qquad -\lambda\, Q_3 = \log_e \left(\frac{1}{4} \right)$$

$$\Rightarrow \qquad Q_3 = \frac{1}{\lambda} \log_e 4$$

Property VII: Mean deviation about mean:

The mean deviation about mean is given by M.D. = $\frac{2}{\lambda}$ e^{-1}

Proof: Let X ∽ exp (λ)

$$\therefore \qquad E(X) = \text{Mean} = \frac{1}{\lambda}$$

M.D. about mean = E (|X - E(X)|) = E (|X - mean|)

$$= E \left| X - \frac{1}{\lambda} \right|$$

$$= \int_0^{\infty} \left| X - \frac{1}{\lambda} \right| f(x)dx$$

$$= \lambda \int_0^{\infty} \left| X - \frac{1}{\lambda} \right| e^{-\lambda x} dx = \int_0^{\infty} | \lambda x - 1 | e^{-\lambda x} dx$$

Put $\lambda x = y \Rightarrow$ $dx = \dfrac{dy}{\lambda}$

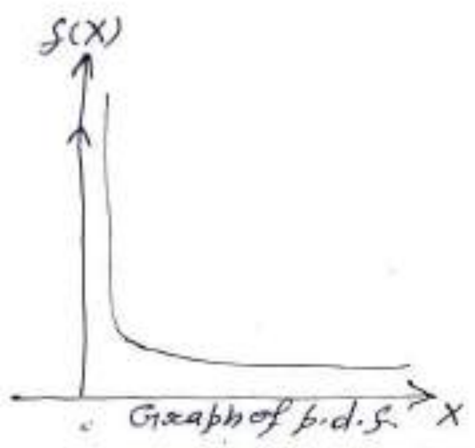
$$\therefore \quad \text{M.D.} = \frac{1}{\lambda}\int_0^\infty |y-1|e^{-y}dy = \frac{1}{\lambda}\left[\int_0^\infty (y-1)e^{-y}dy + \int_1^\infty (y-1)e^{-y}dy\right]$$

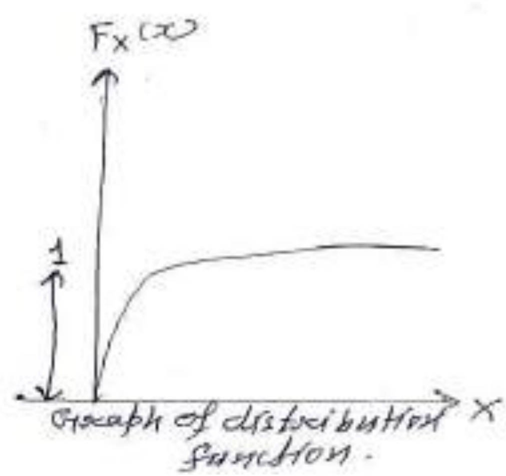
$$= \frac{1}{\lambda}[e^{-1} + e^{-1}] = \frac{2}{\lambda} e^{-1}$$

Property VIII: Graphs for Exponential Distribution:

The graph of exponential probability density function and distribution function are as under:

x	0	1	2	3	....	∞
$f(x) = \lambda e^{-\lambda x}$	λ	$\lambda^{e-\lambda}$	$\lambda e^{-2\lambda}$	$\lambda e^{-3\lambda}$	....	0
$F_x(x) = P(X \leq x) = 1 - e^{-\lambda x}$	0	$1-e^{-\lambda}$	$1-e^{-2\lambda}$	$1-e^{-3\lambda}$	....	1





15.5 Some Important Theorems

Theorem 1: If $X_1, X_2, \ldots, X_n$ are n independent exponential random variables with parameters $\lambda_1, \lambda_2, \ldots, \lambda_n$ respectively. Show that $Z = \text{Min}(X_1, X_2, \ldots, X_n)$ has exponential distribution with parameter $\lambda = \sum_{i=1}^{n} x_i$.

Proof: Let $X_i \sim \exp(\lambda_i)$; i = 1, 2, ..., n and Xi is are independent

Define $Z = \text{Min}(X_1, X_2, \ldots, X_n)$

The distribution function G of random variable Z is

$G_z(z) = P(Z \leq z) = 1 - P(Z > z)$

$= 1 - P[\text{Min}(X_1, X_2, \ldots, X_n) \geq z)$

$= 1 - [P(X_1 > z) \cap P(X_2 > z) \cap \ldots\ldots \cap P(X_n > z)]$

$= 1 - \prod_{i=1}^{n} P(X_i > z)$ [$\because$ X_i is are independent]

$= 1 - \prod_{i=1}^{n} [1 - P(X_i \leq z)]$

$= 1 - \prod_{i=1}^{n} [1 - F_{X_i}(z)]$

Where F is cumulative distribution function of random variable x_i

$= 1 - \prod_{i=1}^{n} [1 - (1 - e^{-\lambda_i z})]$

$= 1 - \prod_{i=1}^{n} e^{-\lambda_i z}$

$= 1 - e^{-\lambda_1 z} \cdot e^{-\lambda_2 z} \cdot e^{-\lambda_3 z} \ldots\ldots e^{-\lambda_n z}$

$= 1 - e^{-z(\lambda_1 + \lambda_2 + \ldots\ldots + \lambda_n)}$

$= 1 - e^{-\left(\sum_{i=1}^{n} \lambda_i\right) z}$

$\therefore$ The cumulative distribution function of random variable Z is

$G_z(z) = 1 - e^{-\left(\sum_{i=1}^{n} \lambda_i\right) z}$ for z > 0

The probability density function of random variable Z is

$g_z(z) = \frac{d}{dz} G_z(z) = \left(\sum_{i=1}^{n} \lambda_i\right) \exp\left\{\left(\sum_{i=1}^{n} \lambda_i\right) z\right\}$ for z > 0

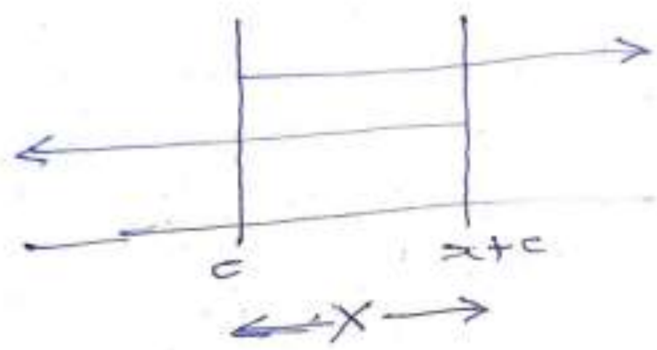
$\therefore$ $Z = \text{Min}(X_1, X_2, \ldots, X_n)$ is exponential with parameter $\lambda = \sum_{i=1}^{n} \lambda_i$

Theorem 2: The exponential distribution is a 'lacks memory' distribution i.e. if X is a exponential random variable, then for every constant $\subset > 0$,

$P(Y \leq x \ 1 \ X \geq \subset) = P(X \leq x)$ for all x, where $Y = X - \subset$.

Proof: Let X ∽ exp (X). Therefore probability distribution function of random variable X is

$f(x) = \lambda e^{-\lambda x}$; $x \geq 0$



∴ $P(Y \leq x \cap X \geq \subset) = P(X - \subset \leq x \cap X \geq \subset)$ [∵ $Y = X - \subset$]

$= P(X \leq x + \subset \cap X \geq \subset)$

$= P(\subset \leq X \leq x + \subset)$

$$= \int_{c}^{x+c} f(x)dx$$

$$= \int_{c}^{x+c} e^{-\lambda x}dx$$

$$= \lambda \int_{c}^{x+c} e^{-\lambda x}dx$$

$= e^{-c\lambda}(1-e^{-\lambda x})$

Also $P(X \geq c) = \int_{c}^{\infty} f(x)dx = \int_{c}^{\infty} \lambda e^{-\lambda x}dx = \left[-e^{-\lambda x}\right]_{c}^{\infty} = e^{-c\lambda}$

∴ $P(Y \leq x \, 1X \geq \subset) = \dfrac{P(Y \leq x \cap X \geq \subset)}{P(X \geq \subset)}$

$$= \frac{e^{-\subset\lambda}(1-e^{-\lambda x})}{e^{-\subset\lambda}} \qquad = 1 - e^{-\lambda x}$$

∴ $P(Y \leq x \mid X \geq \subset) = 1 - e^{-\lambda x}$

Also $P(X \leq x) = \int_{0}^{\infty} f(x)dx = \int_{0}^{\infty} \lambda e^{-\lambda x}dx$

$$= \left[-e^{-\lambda x}\right]_0^{\infty} = 1 - e^{-\lambda x} \qquad .....(2)$$

From (1) and (2), P(Y $\leq$ x | X > ⊂) = P (X $\leq$ x)

Hence, exponential distribution is a lacks memory distribution.

The following examples will illustrate the idea more clearly:-

Example 1: A random variable X follow the following probability law

$$f(x) = \begin{cases} 4e^{-4x} \ ; \ x > 0 \\ 0 \ ; otherwise \end{cases}$$

Find (i) P (X $\leq$ 3) (ii) Coefficient of Variation.

Sol. The random variable X has the following probability law

$$f(x) = \begin{cases} 4e^{-4x} \ ; \ x > 0 \\ 0 \ ; otherwise \end{cases} \qquad \text{[Given]}$$

(i) $$P(X \leq 3) = \int_0^3 f(x)dx = \int_0^3 4e^{-4x}dx = \left[-e^{-4x}\right]_0^3$$

= [e^{-12} - e^{0}]

= 1 - e^{-12}

(ii) Coefficient of variation (C.V.) = $\frac{S.D.}{Mean}$

$$\text{Mean} = E(X) = \int_0^{\infty} xf(x)dx = \int_0^{\infty} x4e^{-4x}dx$$

$$= 4\left[x\frac{e^{-4x}}{(-4)} - \frac{e^{-4x}}{16}\right]_0^{\infty} = 4\left[0 - \left\{0 - \frac{1}{16}\right\}\right] = \frac{1}{4}$$

Also $$E(X^2) = \int_0^{\infty} x^2 f(x)dx = 4\int_0^{\infty} x^2 e^{-4x}dx$$

$$= 4\left[x^2\left(\frac{e^{-4x}}{-4}\right) - 2x\left(\frac{e^{-4x}}{16}\right) + 2\left(\frac{e^{-4x}}{-64}\right)\right]_0^{\infty}$$

$$= 4\left[0 - \left\{0 - 0 + \frac{2}{(-64)}\right\}\right] = \frac{1}{8}$$

∴ Variance = E(x2) - [E(x)]3

$$= \frac{1}{8} - \left(\frac{1}{4}\right)^2 = \frac{1}{16}$$

and thus standard deviation S.D. = $\sqrt{\text{Variance}} = \frac{1}{4}$

$\therefore$ Coefficient of variation = $\frac{S.D.}{\text{Mean}} = \frac{1/4}{1/4}$ = 1

Example 2 : The daily consumption of petrol in Delhi in excess of 50,000 liters is distribution as exponential with parameter $\lambda = \frac{1}{10,000}$. The city has stock of 30,000 liters of petrol. Find the probability that there is shortage of petrol on a particular day.

Sol. : Let x be the random variable represents the consumption of petrol on a particular day in Delhi. Since, given the daily consumption of petrol in excess of 50,000 liters is follows exponential distribution with parameter $\lambda = \frac{1}{10,000}$. Therefore, the probability law for random variable x is

$$P(x) = \begin{cases} \frac{1}{10,000} e^{-\frac{x}{10,000}} & , \quad x \geq 0 \\ 0 & , \quad \text{otherwise} \end{cases}$$

$\therefore$ Required probability = P (there is a shortage of petrol)

= P (Consumption is more than 30,000 liters)

= P (x > 30,000)

$$= \int_{30,000}^{\infty} p(x)dx$$

$$= \int_{30,000}^{\infty} \frac{1}{10,000} e^{-\frac{x}{10,000}} dx.$$

Example 3 : What are the p.d.f., the mean, and the variance of x, if the m.g.f. of x is given by M(t) = $\frac{1}{1-3t}$, t < $\frac{1}{3}$

Sol. Given the random variable x has the m.g.f. M(t) = $\frac{1}{1-3t} = \frac{1/3}{\frac{1}{3}-t}$

It is of the form $\frac{\lambda}{\lambda - t}$, where λ = $\frac{1}{3}$, the m.g.f. of the exponential distribution with parameter λ = $\frac{1}{3}$

Therefore by uniqueness theorem on m.g.f. of the random variable x follows exponential distribution with parameter λ = $\frac{1}{3}$

∴ $f(x) = \lambda e^{-\lambda x}$; $0 \leq x < \infty$

i.e. $f(x) = \frac{1}{3} e^{-\frac{x}{3}}$; $0 \leq x < \infty$

∴ Mean (x) = $\frac{1}{\lambda} = \frac{1}{1/3} = 3$

and variance (x) = $\frac{1}{\lambda^2} = \frac{1}{(1/3)^2} = 9$

Example 4 : The mileage (in thousand of miles), which the car owners get with a certain kind of tyres is a random variable having probability density function.

$$f(x) = \begin{cases} 0-10e^{-0.10x} & ; \quad x > 0 \\ 0 & ; \quad \text{otherwise} \end{cases}$$

Find the probability that one of these tyres will last

(a) almost 5000 miles and (b) between 8,000 and 12,000.

Sol. : Let x be the random variable representing the mileage (in thousand miles) which a car owner can get with a certain kind of tyres.

$$\therefore \quad f(x) = \begin{cases} 0-10e^{-0.10x} & ; \quad x > 0 \\ 0 & ; \quad \text{otherwise} \end{cases}$$

(a) Required probability = P (almost 5000 miles)

= P (x < 5)

$$= \int_0^5 f(x)dx = \int_0^5 0.10e^{-0.10x}dx$$

$$= \left| \frac{e^{-0.10x}}{(-1)} \right|_0^3 = 1 - e^{-0.5} = 1 - 0.6065$$

= 0.3935

(b) P (between 8,000 and 12,000)

$= P(8 \le x \le 12)$

$= \int_8^{12} f(x)dx = \int_8^{12} 0.10\, e^{-0.10x} dx$

$= \left| \dfrac{e^{-0.10x}}{(-1)} \right|_8^{12}$

$= e^{-0.8} - e^{-1.2} = 0.4493 - 0.3012$

$= 0.1481$

Self-Check Exercise

Q.1 A random variable x has exponential distribution with parameter $\lambda = 3$. Find

(i) P (x > 4)

(ii) Find S. D. and coefficient variation.

Q.2 Suppose that the life is a certain type of electronic component has an exponential distribution with mean life of 500 hours. If x denotes the life of this component. Suppose that the component has been in operation for 300 hours. Find the conditional probability that it will last for another 600 hours.

15.6 Summary

We conclude this unit by summarizing what we have covered in it:

1. Defined exponential distribution.
2. Discussed in detail different properties of exponential distribution. Properties discussed are distribution function, moment generating function, characteristic function, cummulant generating function, Quartiles and mean deviation about mean.
3. Discussed and proved some important theorems of exponential random variables.
4. Some examples are given related to each topic so that the contents be clarified further.

15.7 Glossary:

1. A continuous random variable x assuming non-negative values is said to follow an exponential distribution with parameter $\lambda > 0$, if its probability density function is given by

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & ; \quad x > 0 \\ 0 & ; \quad \text{otherwise} \end{cases}$$

2. The distribution function of exponential random variable is given by

$$F_x(x) = \begin{cases} 1 - e^{-\lambda x} & ; \quad x > 0 \\ 0 & ; \quad \text{otherwise} \end{cases}$$

3. The characteristic Function of the exponential distribution about origin is given by

$$\phi_x(t) = \sum_{r=0}^{\infty} \left(\frac{it}{\lambda}\right)^r$$

15.8 Answer to Self Check Exercise

Ans.1 (i) $P(x \geq 4) = \frac{1}{e^{12}}$

(ii) S.D. = $\frac{1}{3}$

Coefficient of variation = 1

Ans. 2 (i) Required probability = conditional probability that it will last for another 600 hours = $e^{-6/5}$

15.9 References/Suggested Readings

1. Robert V. Hogg, Joseph w. Mckean and Allen T. craig, Introduction to Mathematical statistics, Pearson Education, Asia, 2007.
2. Irwim Miller, Marylees Miller and John E. Freund, Mathematical Statistics with Application, 7th Ed., Pearson Education, Asia, 2006.
3. Sheldon Ross, Introduction to Probability Model, 9th Ed., Academic Press, Indian Reprint, 2007.

15.10 Terminal Questions

1. If families are selected randomly in a certain thickly populated area and their monthly income in excess of Rs. 4,000 is treated as exponential random variable with parameter $\lambda = \frac{1}{2000}$. What is the probability that 3 out of 4 families selected in the area have income in excess of Rs. 5,000?
2. What are the p.d.f., the mean and the variance of x if the m.g.f. of x is given by

   $M(t) = \frac{3}{3-t}$, t < 3.
3. Customers arrive in a certain shop according to an approximate poisson process at a mean rate of 20 per hour. What is the probability that the shopkeeper will have to wait more than 5 minutes for the arrival of the first customer?

Unit - 16

Joint Distribution

Structure

16.1 Introduction

In probability theory and statistics, the joint distribution refers to the combined probability distribution of two or more random variables. It describes the likelihood of various combination of values for the random variables occurring together. The joint distribution of two random variables x an y is typically denoted as P(x, y) or f (x, y), depending on whether the variables are discrete or continuous. For discrete random variables, the joint probability mass function (PMF) gives the probability of each possible combination of values of x and y. For continuous random variables, the joint probability density function (PDF) describes the relative likelihood of different combinations of values for x and y. The individual probability distributions of the random variables x and y can be obtained by summing/integrating the joint distribution over the other variable. If the joint distribution factors into the product of the individual distributions, then the variables are independent. Understanding joint distributions is essential for analyzing relationships between multiple random variables, as well as for constructing and working with multivariate probability models.

16.2 Leaning Objectives

After studying the Unit, you should be able to :

- Define two dimensional random variable.
- Discuss discrete and continuous cases of two dimensional random variable.
- Discuss and find joint distribution function for discrete case.
- Discuss and find joint distribution function and marginal density functions for continuous case.

16.3 Joint Distributions

The ideas an distribution developed previously can be easily generalised to two or more random variables. To understand the idea of a Joint Distribution, we consider the typical case of two random variables that are either both discrete or both continuous. Before proceeding further, let us define a 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 R2 ($x \times R$), is called a two dimensional random variable.

We shall now consider discrete and continuous cases separately.

(a) Discrete Case : Let x and y be two discrete random variables. We define the joint probability function of x and y b

$$P(X = x, Y = y) = f(x, y) \qquad \text{....(1)}$$

where

(i) $f(x, y) \geq 0$

(ii) $\sum_x \sum_y f(x, y) = 1$

i.e. the sum over all values of x and y is one.

Suppose that X can assume any one of the m values $x_1, x_2, \ldots, x_m$ and y can assume any one of the n values $y_1, y_2, \ldots, y_n$. Then the probability of the event that $X = x_j$ and $Y = y_k$ is given by

$$P(X = x_j, Y = y_k) = f(x_j, y_k) \qquad \text{.......(2)}$$

A joint probability function for X and Y can be represented by a joint probability table as below:

X \ Y	y_1	y_2	....	y_n	Total ↓
x_1	$f(x_1, y_1)$	$f(x_1, y_1)$	......	$f(x_1, y_n)$	$f_1(x_1)$
x_2	$f(x_2, y_1)$	$f(x_2, y_2)$	......	$f(x_2, y_n)$	$f_1(x_2)$
$\vdots$	......	.....	.....	......	.....
x_m	$f(x_m, y_1)$	$f(x_m, y_2)$	......	$f(x_m, y_n)$	$f_1(x_m)$
Total →	$f_2(y_1)$	$f_2(y_2)$	......	$f_2(y_n)$	1

The probability that $X = x_j$ is obtained by adding all the entries in the row corresponding to x_j and is given by

$$P(X = x_j) = f_1(x_j) - \sum_{k=1}^{n} (x_j, y_k) \qquad .....(3)$$

for j = 1, 2, ..., m, these are indicated by the entry total in extreme right hand column or margin of the above table.

Similarly, probability that $Y = y_k$ is obtained by adding all entries in the column corresponding to y_k and is given by

$$P(Y = y_k) = f_2(y_k) = \sum_{j=1}^{m} f(x_j, y_k) \qquad .....(4)$$

for k = 1, 2, ..., n, these are indicated by the entry total in the bottom row or margin of the above table.

Since, the probabilities (3) and (4) are obtained from the margins of the table, we often refer to $f_1(x_j)$ and $f_2(y_k)$ [or simply $f_1(x)$ or $f_X(x)$ and $f_2(y)$ or $f_Y(y)$] as the marginal probability functions of X and Y respectively. It should also be noted that

$$\sum_{j=1}^{m} f_1(x_j) = 1, \ \sum_{k=1}^{n} f_2(y_k) = 1 \qquad ......(5)$$

which can be written

$$\sum_{j=1}^{m} \sum_{k=1}^{n} f(x_j, y_k) = 1 \qquad ......(6)$$

This is simply the statement that the total probability of all entries is 1. The grand total of 1 is indicated in the lower right hand corner of the table.

The joint distribution function of X and Y is defined by

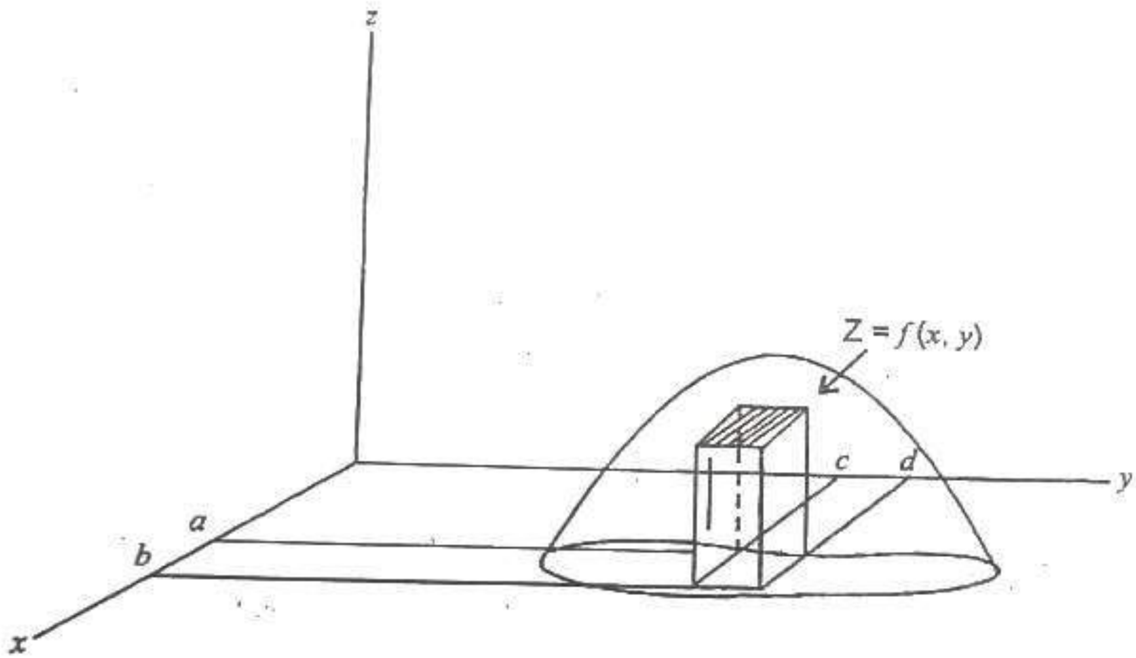
$$F(x, y) = P(X \leq x, Y \leq y) = \sum_{u \leq x} \sum_{v \leq y} f(y, v) \qquad .....(7)$$

In the above table, F(x, y) is the sum of all the entries for which $x_j \leq x$ and $y_k \leq y$.

(b) Continuous Case : The joint probability density function for the random variables X and Y (or, as it is more commonly called, the joint density function of X and Y) is defined by

(i) $f(x, y) \geq 0$ (ii) $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)\,dx\,dy = 1$

Graphically $z = f(x, y)$ represents a surface, called the probability surface, as shown in the following figure.



The total volume bounded by this surface and the XY plane is equal to 1 in accordance with property (ii) above. The probability that X lies between a and b while Y lies between c and d is given graphically by the shaded volume of the figure and mathematically by

$$P(a < X < b, c < Y < d) = \int_{x=a}^{b} \int_{y=c}^{d} f(x, y)\,dx\,dy \qquad ....(8)$$

More generally, if it represents any event, there will be a region R_A of the XY plane that corresponds to it. In such case we can find the probability of A by the following integral.

$$P(A) = \int_R \int_A f(x, y)\,dx\,dy \qquad ....(9)$$

The joint distribution function of X and Y in the continuous case is defined by

$$F(x, y) = P(X \leq x, Y \leq y) = \int_{u=-\infty}^{x} \int_{v=-\infty}^{y} f(u, v)\,du\,dv \qquad ....(10)$$

it is to be noted that

$$\frac{\partial^2 F}{\partial x \partial y} = f(x, y) \qquad ....(11)$$

i.e., the density function is obtained by differentiating the distribution function w.r.t. x and y.

From (11), we obtain

$$P(X \le x) = F_1(x) = \int_{u=-\infty}^{x} \int_{v=-\infty}^{\infty} f(u,v)\,du\,dv \qquad \text{....(12)}$$

$$\text{and } P(Y \le y) = F_2(y) = \int_{u=-\infty}^{\infty} \int_{v=-\infty}^{y} f(u,v)\,du\,dv \qquad \text{.....(13)}$$

We call $F_1(x)$ and $F_2(y)$ as the marginal distribution functions, or simply the distribution functions, of X and Y, respectively. The derivatives of (12) and (13) w.r.t. x and y are then called the marginal density functions, i.e.

$$f_1(x) = \frac{dF_1}{dx} = \int_{v=-\infty}^{\infty} f(x,v)\,dv \qquad \text{.....(14)}$$

$$\text{and} \quad f_2(y) = \frac{dF_2}{dy} = \int_{u--\infty}^{\infty} f(u,y)\,du \qquad \text{.....(15)}$$

it is to be noted that the individual distributions can be obtained from the joint distribution but the converse is not true.

Further, we also use the notations $f_X(x)$ and $g_Y(y)$ for $f_1(x)$ and $f_2(y)$ respectively.

The following examples will illustrated the idea more clearly:-

Example 1: For the following bivariate probability distribution of X and Y, find:

(i) $P(X \le 1, Y = 2)$ (ii) $P(X \le 1)$

(iii) $P(Y \le 3)$ and (iv) $P(X \le 3, Y \le 4)$

X \ Y	1	2	3	4	5	6
0	0	0	$\frac{1}{32}$	$\frac{2}{32}$	$\frac{2}{32}$	$\frac{3}{32}$
1	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$
2	$\frac{1}{32}$	$\frac{1}{32}$	$\frac{1}{64}$	$\frac{1}{64}$	0	$\frac{2}{64}$

Sol: The distribution table can be written as

Y / X	1	2	3	4	5	6	$\sum_x$
0	0	0	$\frac{1}{32}$	$\frac{2}{32}$	$\frac{2}{32}$	$\frac{3}{32}$	$\frac{8}{32}$
1	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{10}{16}$
2	$\frac{1}{32}$	$\frac{1}{32}$	$\frac{1}{64}$	$\frac{1}{64}$	0	$\frac{2}{64}$	$\frac{8}{64}$
$\sum_y$	$\frac{3}{32}$	$\frac{3}{32}$	$\frac{11}{64}$	$\frac{13}{64}$	$\frac{6}{32}$	$\frac{16}{64}$	1

(i) $P(X \le 1, Y = 2) = P(X = 0, Y = 2) + P(X = 1, Y = 2)$

(ii) $P(X \le 1) = P(X = 0) + P(X = 1) = \frac{8}{32} + \frac{10}{16} = \frac{7}{8}$

(iii) $P(Y \le 3) = P(Y = 1) + P(Y = 2) + P(Y = 3) = \frac{3}{32} + \frac{3}{32} + \frac{11}{64} = \frac{23}{64}$

(iv) $P(X \le 3, Y < 4) = P(X = 0, Y = 1) + P(X = 0, Y = 2) + P(X = 0, Y = 3) + P(X = 0, Y = 4) + P(X = 1, Y = 1) + P(X = 1, Y = 2) + P(X = 1, Y = 3) + P(X = 1, Y = 4) + P(X = 2, Y = 1) + P(X = 2, Y = 2) + P(X = 2, Y = 3) + P(X = 2, Y = 4)$

$$= \left(\frac{1}{32}+\frac{2}{32}\right) + \left(\frac{1}{16}+\frac{1}{16}+\frac{1}{8}+\frac{1}{8}\right) + \left(\frac{1}{32}+\frac{1}{32}+\frac{1}{64}+\frac{1}{64}\right) = \frac{9}{16}$$

Example 2: The joint probability distribution of random variables (X, Y) is given by

X / Y	1	2	3	4	5	6
0	0	0	$\frac{1}{32}$	$\frac{2}{32}$	$\frac{2}{32}$	$\frac{3}{32}$
1	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$
2	$\frac{1}{32}$	$\frac{1}{32}$	$\frac{1}{64}$	$\frac{1}{64}$	0	$\frac{2}{64}$

Find:

(i) $P(X = 2, Y \leq 1)$ (ii) $P(Y \leq 1)$

(iii) $P(X = 3)$ (iv) $P(X \leq 3)$

(v) $P(X \leq 4, Y \leq 3)$

Sol: The distribution table can be written as

Y \ X	1	2	3	4	5	6	$\sum_x$
0	0	0	$\frac{1}{32}$	$\frac{2}{32}$	$\frac{2}{32}$	$\frac{3}{32}$	$\frac{8}{32}$
1	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{10}{16}$
2	$\frac{1}{32}$	$\frac{1}{32}$	$\frac{1}{64}$	$\frac{1}{64}$	0	$\frac{2}{64}$	$\frac{8}{64}$
$\sum_y$	$\frac{3}{32}$	$\frac{3}{32}$	$\frac{11}{64}$	$\frac{13}{64}$	$\frac{6}{32}$	$\frac{16}{64}$	1

Now

(i) $P(X = 2, Y \leq 1) = P(X = 2, Y = 0) + P(X = 2, Y = 1) = 0 + \frac{1}{16} = \frac{1}{16}$

(ii) $P(Y \leq 1) = P(Y = 0) + P(Y = 1) = \frac{8}{32} + \frac{10}{16} = \frac{28}{32}$

(iii) $P(X = 3) = \frac{11}{64}$

(iv) $P(X \leq 3) = P(X = 1) + P(X = 2) + P(X = 3) = \frac{3}{32} + \frac{3}{32} + \frac{11}{64} = \frac{23}{64}$

(v) $P(X \leq 4, Y < 3) = P(X = 1, Y = 0) + P(X = 1, Y = 1) + P(X = 1, Y = 2) + P(X = 2, Y = 0) + P(X = 2, Y = 1) + P(X = 2, Y = 2) + P(X = 3, Y = 0) + P(X = 3, Y = 1) + P(X = 3, Y = 2) + \{(X = 4, Y = 0) + P(X = 4, Y = 1) + P(X = 4, Y = 2)$

$$= \left(0 + \frac{1}{16} + \frac{1}{32}\right) + \left(0 + \frac{1}{16} + \frac{1}{32}\right) + \left(\frac{1}{32} + \frac{1}{8} + \frac{1}{64}\right) + \left(\frac{2}{32} + \frac{1}{8} + \frac{1}{64}\right)$$

$$= \frac{3}{32} + \frac{3}{32} + \frac{11}{64} + \frac{13}{64} = \frac{36}{64}$$

Example 3: The joint probability distribution of two random variables X and Y is given by

P(X = 0, Y = 1) = $\frac{1}{3}$, P(X = 1, Y = -1) = $\frac{1}{3}$

and P(X = 1, Y = 1) = $\frac{1}{3}$

Find marginal distribution of X and Y

Sol: The joint probability distribution of two random variables X and Y is given as follows:

Y \ X	0	1	$\sum_x$
1	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{2}{3}$
-1	0	$\frac{1}{3}$	$\frac{1}{3}$
$\sum_y$	$\frac{1}{3}$	$\frac{2}{3}$	1

Now

Marginal distribution of X is

Y	P(Y = y)
1	$\frac{2}{3}$
-1	$\frac{1}{3}$
Total	1

Marginal distribution of Y is

Y	P(Y = y)
1	$\frac{2}{3}$
-1	$\frac{1}{3}$
Total	1

Example 4: Determine the value of k for which the function given by p(x, y) = k x y for x = 1, 2, 3 and y = 1, 2 can serve as joint probability mass function.

Sol: The probabilities for different values of X and Y can be written in the table as

Y \ X	1	2	3
1	k	2k	3k
2	2k	4k	6k

For p(x, y) = k x y to be the joint probability mass function, we must have

$$\sum_x \sum_y p(x, y) = 1 \Rightarrow k + 2k + 3k + 2k + 4k + 6k = 1$$

$\Rightarrow$ 18k = 1

$\Rightarrow$ $k = \frac{1}{18}$

Example 5: If the non-negative function g(x) has the property that

$$\int_0^\infty g(x)dx = 1$$

then show that

$$f(x_1, x_2) = \begin{cases} \dfrac{2g\left(\sqrt{x_1^2 + x_2^2}\right)}{\pi\sqrt{x_1^2 + x_2^2}}, 0 < x_1 < \infty, 0 < x_2 < \infty \\ 0 \qquad , \quad elsewhere \end{cases}$$

satisfies the conditions of being a probability density function of two continuous type random variables X_1 and X_2.

Sol: Clearly $f(x_1, x_2) > 0$

Now, to find

$$\int_0^\infty \int_0^\infty \frac{2g\left(\sqrt{x_1^2 + x_2^2}\right)}{\pi\sqrt{x_1^2 + x_2^2}} dx_1\, dx_2$$

Put $x_1 = r \cos \theta$, $x_2 = r \sin \theta$

therefore, $|J| = r$. Also r varies from 0 to ∞ and θ varies from 0 to $\pi/2$.

Thus, we get

$$\int_{r=0}^{\infty} \int_{\theta=0}^{\pi/2} \frac{2g(r)}{\pi r} r\, dr\, d\theta = \frac{2}{\pi} \int_{r=0}^{\infty} |\theta|_0^{\pi/2} g(r) dr$$

$$= \frac{2}{\pi} \cdot \frac{\pi}{2} \int_{r=0}^{\infty} g(r) dr = 1,$$

by using given value.

Therefore, $f(x_1, x_2)$ satisfies the conditions of probability density function.

Hence the required result

Example 6: For what value of k the function $f(x, y) = kx (x - y)$ for $0 < x < 1$, $-x < y < x$ is a joint p.d.f. Also, find both the marginal probability density functions.

Sol: By def, for $f(x, y)$ to be joint p.d.f., we must have

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dy\, dx = 1$$

i.e. $$\int_0^1 \int_{-x}^{x} f(x, y) dy\, dx = 1$$

$$\Rightarrow \quad \int_0^1 \int_{-x}^{x} k\, x(x-y)\,dy\,dx = 1$$

$$\Rightarrow \quad k \int_0^1 x \,|\, xy - y^2 \,|_{-x}^{x}\, dx = 1$$

$$\Rightarrow \quad k \int_0^1 x(2x^2)\,dx = 1$$

$$\Rightarrow \quad 2k \left| \frac{x^4}{4} \right|_0^1 = 1$$

$$\Rightarrow \quad \frac{2k}{4} = 1$$

$$\Rightarrow \quad k = 2$$

Marginal p.d.f. of X = $f_X(x) = \int_{-n}^{x} f(x, y)\,dx$

$$= \int_{-n}^{x} 2x(x-y)\,dy$$

$$= 2x \left| xy - y^2 \right|_{-x}^{x}$$

$= 2x\,(x^2) = 2x^3$; $0 < x < 1$

Similarly, Marginal p.d.f. of Y = $g_Y(y)$

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

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

$$= 2 \left(\frac{1}{3} - \frac{y}{2} \right)$$

$= \dfrac{2-3y}{3}$; $-x < y < x$, $0 < x < 1$.

Self Check Exercise

Q.1 The joint probability distribution of a pair of random variables is given by

Y \ X	0	1	2
-1	0.1	0.1	0.2
0	0.2	0.1	0.1
1	0.1	0.1	0.0

Find (i) The marginal distributions of X and Y

(ii) P(X + Y < 2)

Q.2 Find k so that f(x, y) = k x y, $1 \leq x \leq y \leq 2$ will be the probability density function.

16.4 Summary

We conclude this unit by summarizing what we have covered in it:

1. Defined two dimensional random variable.
2. Discussed in detail discrete two dimensional random variable case and find the marginal probability function and joint distribution function of two random variables.
3. Discussed in detail continuous two dimensional random variable case and find the marginal density functions and joint distribution function of two random variables.
4. Some examples are given related to each topic so that the contents be clarified further.

16.5 Glossary:

1. Let X and Y be two random variables defined on the same sample space S, them the function (X, Y) that assigns a point in R^2 ($R \times R$) is called a two dimensional random variable.
2. Let X and Y be two discrete random variables. Then the joint probability function of X and Y is given by P(X = x, Y = y) = f(x, y)

   where f(x, y) > 0 and $\sum_x \sum_y f(x, y) = 1$
3. The joint probability density function for the random variables X and Y is defined by f(x, y) > 0 and $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx\, dy = 1$

16.6 Answer to Self Check Exercise

Ans.1 (i) The marginal distribution of X is

X	0	1	2	Total
P(X=x)	0.4	0.3	0.3	1

The marginal distribution of Y is

Y	-1	0	1	Total
P(Y=y)	0.4	0.4	0.2	1

(ii) $P(X + Y < 2) = 0.8$

Ans. 2 $k = \frac{6}{5}$

16.7 References/Suggested Readings

1. Robert V. Hogg, Joseph w. Mckean and Allen T. craig, Introduction to Mathematical statistics, Pearson Education, Asia, 2007.
2. Irwim Miller, Marylees Miller and John E. Freund, Mathematical Statistics with Application, 7th Ed., Pearson Education, Asia, 2006.
3. Sheldon Ross, Introduction to Probability Model, 9th Ed., Academic Press, Indian Reprint, 2007.

16.8 Terminal Questions

1. The joint probability distribution of random variables (X, Y) is given by

Y \ X	-1	1
-1	$\frac{1}{8}$	$\frac{1}{2}$
0	0	$\frac{1}{4}$
1	$\frac{1}{8}$	0

Find the marginal distribution of X and Y.

2. Let the joint p.m.f. of X and Y be

$$p(x, y) = \begin{cases} \dfrac{x+y}{21}, & x = 1, 2, 3;\ y = 1, 2 \\ 0, & \text{otherewise} \end{cases}$$

3. A gun is aimed at a certain point (origin of the coordinate system) and because of the random factors, the actual hit point can be any point (X_1, X_2) in a circle of radius R about the origin. If the joint density of X_1 and X_2 is constant in this circle given by

$$f(x_1, x_2) = \begin{cases} k, & \text{for } x_1^2 + x_2^2 \le R^2 \\ 0, & \text{elsewhere} \end{cases}$$

Then (a) Compute k; and

(b) show that

$$f_1(x_1) = \begin{cases} \dfrac{2}{\pi R}\left\{1-\left(\dfrac{x_1}{R}\right)^2\right\}^{1/2}, & \text{for } -R \le x_1 \le R \\ 0, & \text{elsewhere} \end{cases}$$

Unit - 17

Conditional Distribution

Structure

17.1 Introduction

In probability theory and statistics, the conditional distribution refers to the distribution of a random variable given the value of one or more other random variables. Formally, if we have two random variables X and Y, the conditional distribution of Y given X = x is denoted as P(Y|X = x) or f(y|x). The conditional distribution describes the probability or probability density function of Y when the value of X is known. It provides information about the relationship between the two variables and how the distribution of one variable changes based on the value of the other. Conditional distributions are essential for drawing inferences about the relationship between variables. They allow as to predict the likely values of one variable (the dependent variable) based on the known values of the other variable (3) (the independent variable (1)). Conditional distributions play a crucial role in Bayesian inference, where they are used to update the prior beliefs about a parameter or a random variable based on observed data. Conditional distributions can also help us model complex relationships between variables, especially in multivariate settings.

17.2 Learning Objectives

After studying this unit, you should be able to:

- Discuss the conditional distribution for the case of two dimensional discrete random variable
- Define conditional distribution function
- Discuss the conditional distribution for the case of two dimensional continuous random variable

17.3 Discrete Case

Let (x, y) be a two dimensional discrete random variable. Then the conditional discrete density function or the conditional probability moss function of x, given Y = y, denoted by

$f_{x/y}$ (x/y) or simply f (x/y) is denied as

$f_{x/y}$ (x/y) = $P_{x/y}$ (x/y)

$$= \frac{P(X = x, Y = y)}{P(Y = y)} = \frac{f(x, y)}{f_2(y)} \qquad ...(1)$$

Simply for y and given X = x, we have

$$f_{y/x}(y/x) = P_{y/x}(y/x) = \frac{P(X = x, Y = y)}{P(X = x)}$$

$$= \frac{f(x, y)}{f_1(x)} \qquad ...(2)$$

It is to be noted that

$$\sum_x \frac{f(x, y)}{f_2(y)} = \frac{1}{f_2(y)} \sum_x f(x, y) = \frac{f_2(y)}{f_2(y)} = 1$$

and $\frac{f(x, y)}{f_2(y)} > 0 \ \forall$ x, y for which $f_2(y) \neq 0$ and this (1) respondents a probability function. Similarly (2) is also a probability function.

17.4 Continuous Case

We firstly define the conditional distribution function and then differentiate to obtain the conditional density function.

The conditional distribution function

F (y/x) (or $F_{y/x}$ (y/x)) denotes the distribution function of Y when X has already assumed the particular value of x i.e.

$$F_{y/x}(y/x) = P(Y < y/ X = x) = \frac{\int_{u=-\infty}^{\infty} \int_{v=-\infty}^{\infty} f(u, v)\, du\, dv}{\int_{u=x}^{x+dx} \int_{v=-\infty}^{\infty} f(u, v)\, du\, dv}$$

and a similar expression for $F_{X/Y}$(x/y).

The conditional probability density function of Y given X for two random variables X and Y which are jointly continuously distributed is defined as follows for two real numbers x and y as:

$f_{Y/X}(y/x) = \frac{\partial}{\partial y} F_{Y/X}(y/x) =$

It can also be written directly as

$f_{Y/X}(y/x) = \frac{(x, y)}{f_1(x)}$

Where $f(x, y)$ is the joint density function of X and Y and $f_1(x)$ is the marginal density function of X.

Let us improve our understanding of these results by looking at some following examples:-

Example 1: The joint probability distribution of a pair of random variables is given by

Y \ X	0	1	2
-1	0.1	0.1	0.2
0	0.2	0.1	0.1
1	0.1	0.1	0.0

Find the conditional distribution of X given Y = 0

Sol: The probability distribution table can be written as

Y \ X	0	1	2	$\sum_x$
-1	0.1	0.1	0.2	0.4
0	0.2	0.1	0.1	0.4
1	0.1	0.1	0.0	0.2
$\sum_y$	0.4	0.3	0.3	1

Now, we have

$P(X = x/y = 0) = \frac{P(X = x \cap Y = 0)}{P(Y = 0)}$

$\therefore$ P(X = 0/Y = 0) = $\frac{P(X=0\bigcap Y=0)}{P(Y=0)}$ = $\frac{0.2}{0.4}$ = $\frac{1}{2}$

P(X = 1/Y = 0) = $\frac{P(X=1\bigcap Y=0)}{P(Y=0)}$ = $\frac{0.1}{0.4}$ = $\frac{1}{4}$

and P(X = 2/Y = 0) = $\frac{P(X=2\bigcap Y=0)}{P(Y=0)}$ = $\frac{0.1}{0.4}=\frac{1}{4}$

$\therefore$ conditional distribution of X given Y = 0

X/Y = 0	P(X = x/y = 0)
x = 0/y = 0	$\frac{1}{2}$
x = 1/y = 0	$\frac{1}{4}$
x = 2/y = 0	$\frac{1}{4}$
Total	1

Example 2: Find marginal p.m.f's of X_1 and X_2 whose joint p.m.f. is

	0	1	2
0	$\frac{1}{12}$	$\frac{1}{6}$	$\frac{1}{24}$
1	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{40}$
2	$\frac{1}{8}$	$\frac{1}{20}$	0
3	$\frac{1}{120}$	0	0

Also find conditional distribution of X_2 given $X_1 = 0$

Sol: The distribution table can be written as

Y \ X	0	1	2	$\sum_x$
0	$\frac{1}{12}$	$\frac{1}{6}$	$\frac{1}{24}$	$\frac{7}{24}$
1	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{40}$	$\frac{21}{40}$
2	$\frac{1}{8}$	$\frac{1}{20}$	0	$\frac{7}{40}$
$\sum_y$	$\frac{56}{120}$	$\frac{28}{60}$	$\frac{8}{120}$	1

(i) The marginal p.m.f. of X_1 is

X_1	$P(X_1 = x_1)$
0	$\frac{56}{120}$
1	$\frac{28}{60}$
1	$\frac{8}{120}$
Total	1

(ii) The marginal p.m.f. of X_2 is

X_2	$P(X_2 = x_2)$
0	$\frac{56}{120}$
1	$\frac{28}{60}$
1	$\frac{8}{120}$
Total	1

(iii) The conditional distribution of X_2 given $X_1 = 0$

$X_2/X_1 = 0$	$P(X_2=x_2/x_1=0) = \frac{P(X_2 = x_2 \cap X_1 = 0)}{P(X_1 = 0)}$
$x_2 = 0/x_1 = 0$	$\frac{P(X_2 = 0 \cap X_1 = 0)}{P(X_1 = 0)} = \frac{\frac{1}{12}}{\frac{56}{120}} = \frac{10}{56}$
$x_2=1/x_1=0$	$\frac{P(X_2 = 1 \cap X_1 = 0)}{P(X_1 = 0)} = \frac{\frac{1}{4}}{\frac{56}{120}} = \frac{30}{56}$
$x_2=2/x_1=0$	$\frac{P(X_2 = 2 \cap X_1 = 0)}{P(X_1 = 0)} = \frac{\frac{1}{8}}{\frac{56}{120}} = \frac{1}{56}$
$x_2=3/x_1=0$	$\frac{P(X_2 = 3 \cap X_1 = 0)}{P(X_1 = 0)} = \frac{\frac{1}{120}}{\frac{56}{120}} = \frac{1}{56}$
Total	1

Example 3: If the joint p.d.f. of X and Y is given by

$$p(x, y) = \begin{cases} \frac{1}{4}(2x+y); 0 < x < 1, 0 < y < 2 \\ 0 \quad ; \quad otherwise \end{cases}$$

Find (i) marginal density functions

(ii) the conditional density of Y given X = $\frac{1}{4}$

Sol: The joint p.d.f. of X and Y is given by

$$p(x, y) = \begin{cases} \frac{1}{4}(2x+y); 0 < x < 1, 0 < y < 2 \\ 0 \quad ; \quad otherwise \end{cases}$$

(i) Marginal density function of X is

$f_1(x) = \int_{-\infty}^{\infty} f(x, y)dy$

$= \frac{1}{4} \left| 2xy + \frac{y^2}{2} \right|_0^2 = \frac{1}{4}$ (4x + 2)

$= \frac{2x+1}{2}$

$\therefore$ $f_1(x) = \frac{2x+1}{2}$; 0 < x < 1

(ii) Marginal density function of Y is

$f_2(y) = \int_{-\infty}^{\infty} f(x, y)dx = \int_0^1 \frac{1}{4}(2x + y)dx$

$= \frac{1}{4} \left| x^2 + xy \right|_0^1 = \frac{1+y}{4}$

$\therefore$ $f_2(y) = \frac{1+y}{2}$; 0 < y < 2

(iii) The conditional density of Y given X = $\frac{1}{4}$

$= f_{Y/X}\left(y / x = \frac{1}{4} \right) = \frac{f\left(Y = y \cap x = \frac{1}{4} \right)}{f_1\left(x = \frac{1}{4} \right)} = \frac{(f(x, y))x = \frac{1}{4}}{f_1\left(\frac{1}{4} \right)}$

$= \frac{\frac{1}{4}\left(2\left(\frac{1}{4} \right) + y \right)}{\frac{2\left(\frac{1}{4} \right) + 1}{2}} = \frac{(2y+1)}{\frac{3}{2}}$

$\therefore$ fY/X $\left(y / x = \frac{1}{4} \right) = \frac{2}{3}$ (2y + 1); 0 < y < 2

Example 4: Given the joint density of X_1 and X_2 as

$f(x_1, x_2) = \frac{k}{(1+x_1+x_2)^3}$, for $x_1 > 0$ and $x_2 > 0$,

= 0, elsewhere

Find the value of k and the marginal densities of X_1 and X_2. Also find the conditional density of X_1 given that X_2 assumes the value x_2.

Sol: We must have

$$\int_0^\infty \int_0^\infty \frac{k}{(1+x_1+x_2)^3} dx_1 dx_2 = 1$$

or $$-\frac{1}{2} k \int_0^\infty \left[\frac{k}{(1+x_1+x_2)^2} \right]_0^\infty dx_1 = 1$$

or $$-\frac{1}{2} k \int_0^\infty \frac{1}{(1+x_1)^2} dx_1 = 1$$

or $$-\frac{1}{2} k \int_0^\infty \left[\frac{1}{1+x_1} \right]_0^\infty = 1$$

or $$\frac{1}{2} k = 1$$

or $$k = 2$$

Now, the marginal density of X_1 is

$$f_1(x_1) = \int_0^\infty f(x_1, x_2) dx_2 = 2 \int_0^\infty \frac{1}{(1+x_1+x_2)^3} dx_2$$

$$= - \left[\frac{1}{(1+x_1+x_2)^2} \right]_0^\infty$$

$$= \frac{1}{(1+x_1)^2}, \text{ for } x_1 > 0$$

and 0, elsewhere

Similarly, it can be shown that the marginal density of X_2 is

$$f_2(x_2) = \frac{1}{(1+x_2)^2}, \text{ for } x_2 > 0$$

$= 0$, elsewhere

Further, the conditional density of X_1 given X_2 assumes the value x_2 is

$$f_{x_1/x_2}(x_1/x_2) = \frac{f(x_1, x_2)}{f_2(x_2)}$$

$= 2(1 + x_1 + x_2)^{-3}(1 + x_2)^2$, for x1 > 0 and 0, elsewhere

Example 5: The joint probability function of the two dimensional random variable (X, Y) is given by

$$f(x, y) = \begin{cases} \frac{8}{9}xy; 1 \le x \le y \le 2 \\ \quad 0 \ \ ; elsewhere \end{cases}$$

(i) Find the marginal density functions of X and Y

(ii) Find the conditional density functions of Y given X = x and of X given Y = y.

Sol: Given, the joint p.d.f. of random variable's (X, Y) is

$$f(x, y) = \begin{cases} \frac{8}{9}xy; 1 \le x \le y, 1 \le y \le 2 \\ \quad 0 \ \ ; elsewhere \end{cases}$$

(i) Marginal density function of X = $f_X(x)$

$$= \int_{-\infty}^{\infty} f(x, y)dy = \int_{1}^{2} \frac{8}{9} dy\, dy$$

$$= \frac{8}{9} \text{ x } \left|\frac{y^2}{2}\right|_1^2 = \frac{8}{9} \text{x} \left(\frac{3}{4}\right) = \frac{2}{3} \text{x}$$

(ii) Marginal density function of Y = $g_Y(y)$

$$= \int_{-\infty}^{\infty} f(x, y)dx$$

$$= \int_{1}^{y} \frac{8}{9} xy\, dx = \frac{8}{9} \text{y} \left|\frac{x^2}{2}\right|_1^y$$

$$= \frac{4}{9} \text{y } (y^2 - 1); 1 \le y \le 2$$

(iii) Conditional density function of Y, given X = x

$$f_{Y/X}(y/x) = \frac{f(x,y)}{f_x(x)} = \frac{\frac{8}{9}xy}{\frac{2}{3}x} = \frac{4}{3}y \; ; \; 1 \le y \le 2$$

(iv) Conditional density function of X given Y = y

$$g_{X/Y}(x/y) = \frac{f(x,y)}{g_y(y)} = \frac{\frac{8}{9}xy}{\frac{4}{9}y(y^2-1)} = \frac{2x}{y^2-1} \; ; \; 1 \le x \le y$$

Self-Check Exercise

Q.1 The joint probability mass function of discrete random variable's (X, Y) given by p(1,1) = 0.5, p(1,2) = 0.1, p(2,1) = 0.1, p(2,2) = 0.3

Find Conditional p.m.f. of X given Y = 1.

Q.2 Obtain the marginal and conditional probability functions, if the joint density function is

$$f(x, y) = \begin{cases} 2(2-x-y); 0 \le x \le y \le 1 \\ \quad 0 \quad ; \textit{elsewhere} \end{cases}$$

17.5 Summary

We conclude this unit by summarizing what we have covered in it:

1. Discussed the conditional distribution for the case of two dimensional discrete random variable and find the formulae for conditional discrete density function or the conditional probability mass functions of X, given Y, and for Y given X = x.
2. Defined conditional distribution function.
3. Discussed the conditional continuous random variable.
4. Some examples are given related to each topic so that the contents be clarified further.

17.6 Glossary:

1. If we have two random variables X and Y, the conditional distribution of Y given X = x is denoted as P(Y/X = x) or f(y/x).
2. The conditional distribution function F(y/x) (or $F_{Y/X}$) denotes the distribution function of Y when X has already assumed the particular value of x.
3. Conditional probability mass function of X, given Y = y, denoted by $f_{X/Y}$(x/y) or simply f(x/y) is defined as

$$f_{X/Y}(x/y) = P_{X/Y}(x/y) = \frac{P(X=x, Y=y)}{P(Y=y)} = \frac{f(x,y)}{f_2(y)}$$

17.7 Answer to Self Check Exercise

Ans.1 Conditional p.m.f. of X given Y = 1 is

X/Y = 1	P(X=x/y=1) = $\frac{P[(X=x \cap Y=1)]}{P(Y=1)}$
x=0/y = 1	$\frac{5}{6}$
x=1/y=1	$\frac{1}{6}$
Total	1

Ans. 2 Marginal probability function of X

= $3x^2 - 6x + 3$, for $0 \leq x \leq 1$

and 0, elsewhere

Marginal probability function of Y

= $4y - 3y^2$, for $0 \leq y \leq 1$

and 0, elsewhere

Conditional density function of Y given X

$(0 \leq x \leq 1)$ is $\frac{2(2-x-y)}{3x^2-6x+3}$, $0 \leq y \leq 1$

and conditional density function of X given Y,

$(0 \leq y \leq 1)$ is $\frac{2(2-x-y)}{4y-3y^2}$, $0 \leq x \leq 1$

17.8 References/Suggested Readings

1. Robert V. Hogg, Joseph w. Mckean and Allen T. craig, Introduction to Mathematical statistics, Pearson Education, Asia, 2007.

2. Irwim Miller, Marylees Miller and John E. Freund, Mathematical Statistics with Application, 7th Ed., Pearson Education, Asia, 2006.

17.9 Terminal Questions

1. If a two-dimensional random variable (X_1, X_2) have a bivariate distribution given by

$$f(x_1, x_2) = \frac{1}{27}(2x_1 + x_2)$$

Where x_1 and x_2 can assume only the integer values 0,1,2.

Then find the conditional distribution of X_2 for $X_1 = x_1$

2. Two discrete random variables X and Y have the joint probability function as

$$f(x, y) = \frac{\lambda^x e^{-\lambda} p^y (1-p)^{x-y}}{\lfloor y \, \lfloor x-y} \text{ ; } y = 0, 1, 2, \ldots\ldots, x = 0, 1, 2, \ldots\ldots.$$

where λ, p are constants with $\lambda > 0$ and $0 < p < 1$.

(a) Find the marginal probability functions for X and Y.

(b) Find the conditional probability functions for X and Y.

Unit - 18

Stochastic Independence

Structure

18.1 Introduction

Stochastic independence is a fundamental concept in probability theory and statistics. It describes a situation where the occurrence or non-occurrence of one random event does not influence the probability of another random event. In other words, if two events are stochastically independent, the knowledge of one event occurring or not occurring does not provide any information about the liklihood of the other event. Formally, two events A and B are said to be stochastically independent if the probability of their joint occurrence is equal to the product of their individual probabilities. Stochastic independence is a stronger condition than uncorrelation, which only requires the correlation coefficient between the two events to be zero. When events are stochastically independent, it simplifies the calculation of probabilities, as the joint probability can be expressed as the product of individual probabilities. Stochastic independence is an important assumption in many statistical techniques, such as hypothesis testing, regression analysis, and time series analysis.

18.2 Learning Objectives

After studying this unit, you should be able to:

- Define stochastic independence of two random variables
- Prove the theorem on stochastic independence of two random variables
- Do questions related to stochastic independence.

18.3 Stochastic Independence Definition

Two random variables X and Y with joint probability density function (or probability mass function) $f_{XY}(x,y)$ (or $f(x,y)$) and marginal probability density functions (or p.m.f. is) $f_X(x)$ and $g_Y(y)$ respectively are said to be stochastically independent if and only if

$f_{XY}(x, y) = f_X(x)\ g_Y(y)\ \forall$ x and y i.e. two random variables X and Y are said to be independent or stochastically independent if their joint probability function is equal to the product of their individual probability functions.

Further, in terms of the distribution function, two jointly distributed random variables X and Y are stochastically independent if and only if their joint distribution function F(x, y) is the product of their marginal distribution function $F_X(x)$ and $G_Y(y)$ i.e. $F_{XY}(x,y) = F_X(x)\ G_Y(y)\ \forall$ x and y.

The variables which are not stochastically independent are said to be stochastically dependent.

18.4 Theorem on Stochastic Independences

The random variables X and Y with joint probability density function $f_{XY}(x, y)$ are stochastically independent if and only if $f_{XY}(x, y)$ can be expressed as the product of a non-negative function of x alone and a non-negative function of y alone i.e. if

$$f_{XY}(x, y) = h_X(x), k_Y(y) \qquad ....(1)$$

where $H_X(x) > 0$ and $k_Y(y) > 0\ \forall$ x and y.

Proof: If X and Y are independent, then by definition

$$fXY(x, y) = fX(x).\ gY(y) \qquad ....(2)$$

where $f_x(x)$ and $g_Y(y)$ are the marginal probability density function of X and Y respectively. Thus the condition (1) holds, then we have to prove that X and Y are independent i.e. we have to show that

$$f_{XY}(x, y) = f_X(x)\ g_Y(y)$$

i.e. the joint probability function is equal to the product of individual probability functions.

Now, for continuous random variables X and Y, the marginal probability density functions are given by:

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)dy = \int_{-\infty}^{\infty} h(x)k(y)dy$$

$$= h(x) \int_{-\infty}^{\infty} k(y)dy$$

$$= c_1\ h(x), \text{ where } c_1 = \int_{-\infty}^{\infty} k(y)dy \qquad ......(3)$$

and $\quad g_Y(y) = \int_{-\infty}^{\infty} f(x,y)dx = \int_{-\infty}^{\infty} h(x)k(y)dx$

$= k(y) \int_{-\infty}^{\infty} h(x)dx$

$c_2\, k(y)$, say where $c_2 = \int_{-\infty}^{\infty} h(x)dx$ ....(4)

Here c_1 and c_2 are constants independent of x and y.

Moreover,

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)dx\,dy = 1 \Rightarrow \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x)k(y)dx\,dy = 1 \qquad [\because f(x, y) = h(x)\,k(y)]$$

$$\Rightarrow \left(\int_{-\infty}^{\infty} h(x)dx\right)\left(\int_{-\infty}^{\infty} k(y)dy\right) = 1$$

$\Rightarrow \quad c_1\, c_2 = 1$ ......(5)

Since [By (3) and (4)]

$f_{XY}(x, y) = h_X(x)\, k_Y(y) = c_1\, c_2\, h_X(x)\, k_y(y)$

$= [c_1 h_x(x)]\, [c_2 k_Y(y)]$

$= f_X(x)\, g_Y(y)$

where $f_X(x)$ is the marginal probability density

function of X and $g_Y(y)$ is the marginal probability density function of Y.

$\Rightarrow$ X and Y are stochastically independent.

The following examples will illustrate the idea more clearly:-

Example 1: The joint density function of (X, Y) is

$$f(x, y) = \begin{cases} Ae^{-x-y}; 0 \le x \le y, 0 \le y < \infty \\ 0 \;;\; \text{otherwise} \end{cases}$$

(i) Determine A

(ii) Find the marginal density function of X

(iii) Find the marginal density function of Y

(iv) Examine if X and Y are independent

(v) Find the conditional density function of Y given X = 2

Sol: The joint density function of (X, Y) is given by

$$f(x, y) = \begin{cases} Ae^{-x-y}; 0 \le x \le y, 0 \le y < \infty \\ \quad 0 \;\; ; \;\; \text{otherwise} \end{cases}$$

(i) Since $f(x, y)$ is the joint density function of (X, Y), we must have

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x, y)dx\,dy = 1$$

$$\Rightarrow \quad \int_0^{\infty}\int_0^{y} Ae^{-x-y}dx\,dy = 1$$

$$\Rightarrow \quad A \int_{-\infty}^{\infty} \left[\frac{e^{-(x+y)}}{-1}\right]_0^y dy = 1$$

$$\Rightarrow \quad A \int_{-\infty}^{\infty} (e^{-y} - e^{-2y})dy = 1$$

$$\Rightarrow \quad A \left[\frac{e^{-y}}{(-1)} + \frac{e^{-2y}}{2}\right]_0^{\infty} = 1$$

$$\Rightarrow \quad A \left[(0) - \left(-1 + \frac{1}{2}\right)\right] = 1$$

$$\Rightarrow \quad A \left(\frac{1}{2}\right) = 1$$

$$\Rightarrow \quad A = 2$$

(ii) Marginal density function of X = f_1 (x) =

$$= \int_{-\infty}^{\infty} f(x, y)dy = \int_x^{\infty} 2e^{-(x+y)}dy$$

$$= 2\left[\frac{e^{-(x+y)}}{-1}\right]_x^{\infty}$$

$= 2[0 + e^{-2x}]$

$= 2^{e-2x}$; $0 \le x \le y$, $0 \le y < \infty$

(iii) Marginal density function of Y = f_2(y)

$$= \int_{-\infty}^{\infty} f(x, y)dx$$

$= \int_{x}^{y} 2e^{-(x+y)} dx = 2\left[\frac{e^{-(x+y)}}{-1}\right]_0^y$

$= 2[-e^{-2y} + e^{-y}]$; $0 < y < \infty$

(iv) Since $f(x, y) \neq f_1(x)\, f2(y)$

Therefore X and Y are not independent

(v) Conditional density function of Y given X = 2 is

$$f(Y/X = 2) = \frac{[f(X,Y)]_{X=2}}{f_1(X=2)} = \frac{2e^{-2-y}}{2e^{-2(2)}}$$

$= e^{2-y}$; $0 < y < \infty$

Example 2: The joint density function of the random variables X and Y is given by

$$f(x, y) = \begin{cases} 8xy, 0 \le x \le 1, 0 \le y \le x \\ 0 \;\; ; \;\; \text{otherwise} \end{cases}$$

(a) Find the marginal density of X

(b) Find the marginal density of Y

(c) Find the conditional density of X given Y

(d) Find the conditional density of Y given X

(e) Check whether X and Y are independent.

Sol: (a) To obtain the marginal density of X, we f_ix x and integrate w.r.t. y from 0 to x

i.e. $f_X(x) = \int_{y=0}^{x} 8xy\, dy = 4x^3$ for $0 < x < 1$

and for all other values of x, $f_X(x) = 0$

(b) Similarly, the marginal density of Y.

$g_Y(y) = \int_{x=y}^{1} 8xy\, dy = 4y(1 - y^2)$, $0 < y < 1$

and for all other values of y, $g_Y(y) = 0$

(c) The conditional density function of X is, for $0 < y < 1$

$$f_{X/Y}(x/y) = \frac{f(x,y)}{g_Y(y)} = \begin{cases} \frac{2x}{1-y^2}, y \le x \le 1 \\ 0 \;\; ; \;\; \text{otherwise} \end{cases}$$

where $g_Y(y) \neq 0$

(d) Similarly,

$$f_{Y/X}(y/x) = \frac{f(x,y)}{g_X(x)} = \begin{cases} \frac{2y}{x^2}; 0 \le y \le x \\ 0 \;;\; \text{other y} \end{cases}$$

where $f_X(x) \neq 0$

(e) Here $f(x, y) = 8\,xy$, $f_X(x) = 4x^3$, $g_Y(y) = 4y\,(1 - y^2)$

Therefore, $f(x, y) \neq f_X(x)\, g_Y(y)$ and hence X and Y are dependent i.e. not independent.

Example 3: Let $f(x_1, x_2) = \begin{cases} c(x_1x_2 + e^{x_1}); 0 < (x_1, x_2)^2 < 1 \\ 0 \;;\; \text{otherwise} \end{cases}$

(i) Determine c

(ii) Examine whether X_1 and X_2 are stochastically independent

Sol: The joint probability function of variables X_1 and X_2 is given by

$$f(x_1, x_2) = \begin{cases} c(x_1x_2 + e^{x_1}); 0 < (x_1, x_2)^2 < 1 \\ 0 \;;\; \text{otherwise} \end{cases}$$

Since $f(x_1, x_2)$ is the probability function. Therefore, we must have

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x_1, x_2)\,dx_1\,dx_2 = 1$$

$$\Rightarrow \quad \int_0^1\int_x^1 c(x_1x_2 + e^{x_1})\,dx_1 dx_2 = 1$$

$$\Rightarrow \quad c\int_0^1 \left[\frac{x_1^2}{2}x_2 + e^{x_1}\right]_0^1 dx_2 = 1$$

$$\Rightarrow \quad c\int_0^1 \left(\frac{x_2}{2} + e^1\right) dx_2 = 1$$

$$\Rightarrow \quad c\left[\frac{x^4}{4} + ex_2\right]_0^1 = 1$$

$$\Rightarrow \quad c\left[\frac{1}{4} + e\right] = 0 \;\Rightarrow\; c = \frac{4}{1+4e}$$

$\therefore$ $f(x_1, x_2) = \begin{cases} \dfrac{4}{1+4e}(x_1x_2 + e^{x_1}); 0 < (x_1, x_2) < 1 \\ \quad 0 \;\; ; \quad \text{otherwise} \end{cases}$

(ii) The random variables X_1 and X_2 are statistically independent if

$f(x_1, x_2) = f_{X_1}(x_1) \; f_{X_2}(x_2)$

i.e. the joint probability function is equal to the product of marginal probability functions.

The marginal probability function of X_1 = $f_{X_1}(x_1) = \int_{-\infty}^{\infty} f(x_1, x_2)dx_2$

i.e. $f_{X_1}(x_1) = \int_0^1 \dfrac{4}{1+4e}(x_1x_2 + e^{x_1})dx_2$

$$= \frac{4}{1+4e}\left[x_1\frac{x_2^2}{2} + x_2e^{x_1}\right]_0^1$$

$$= \frac{4}{1+4e}\left[\frac{x_1}{2} + e^{x_1}\right]; 0 < x_2 < 1$$

The marginal probability function of X_2 = $f_{X_2}(x_2)$

$$= \int_{-\infty}^{\infty} f(x_1, x_2)dx_1$$

i.e. $f_{X_2}(x_2) = \int_0^1 \dfrac{4}{1+4e}(x_1x_2 + e_{x_1})dx_1$

$$= \frac{4}{1+4e}\left[\left(\frac{x_2}{2} + e\right) - 1\right]; 0 < x_1 < 1$$

Since $f(x_1, x_2) \neq f_{X_1}(x_1) \; f_{X_2}(x_2)$

$\therefore$ Variables X_1 and X_2 are not stochastically independent.

Example 4: Let random variables X_1 and X_2 have the joint probability density function

$$f(x_1, x_2) = \begin{cases} 12x_1x_2(1-x_2); 0 < x_1 < 1 \\ \qquad\qquad\qquad\quad 0 < x_2 < 1 \\ \quad 0 \quad , \quad elsewhere \end{cases}$$

Show that the random variables are independent.

Sol: Here $f_1(x_1) = \int_0^1 12x_1x_2(1-x_2)dx_2$

$= 12\, x_1 \int_0^1 (x_2 - x_2^2)dx$

$= 12\, x_1 \left[\frac{x_2^2}{2} - \frac{x_2^3}{3}\right]_0^1$

$= 12\, x_1 \left(\frac{1}{2} - \frac{1}{3}\right)$

$= 12\, x_1$

$f_2(x_2) = \int_0^1 12x_1x_2(1-x_2)dx_1$

$= 12x_2(1 - x_2)\left[\frac{x_1^2}{2}\right]_0^1$

$= 12x_2(1 - x_2)\,\frac{1}{2}$

$= 6x_2(1 - x_2)$

By using these values, we conclude that

$f(x_1, x_2) = f_1(x_1)\, f_2(x_2)$

Hence, the random variables X_1 and X_2 are independent

Example 5: The random vriables X and Y are jointly distributed as

$f(x, y) = e^{-(x+y)}$, $x > 0$, $y > 0$

(a) Are X and Y independent?

(b) Find $P(X > 1)$

(c) Find $P(X < Y/X < 2Y)$

(d) Find $P(1 < X + Y < 2)$

Sol: Here $f_X(x) \int_{y=0}^{\infty} e^{-(x+y)}dy = e^{-x}\int_0^{\infty} e^{-y}dy = e^{-x}$

and $f_Y(y) = \int_{x=0}^{\infty} e^{-(x+y)}dx = e^{-y}\int_0^{\infty} e^{-x}dx = e^{-y}$

(a) Here $f(x, y) = e^{-(x+y)} = (e^{-x})\ (e^{-y})$

$= f_X(x)\ f_y(y)$

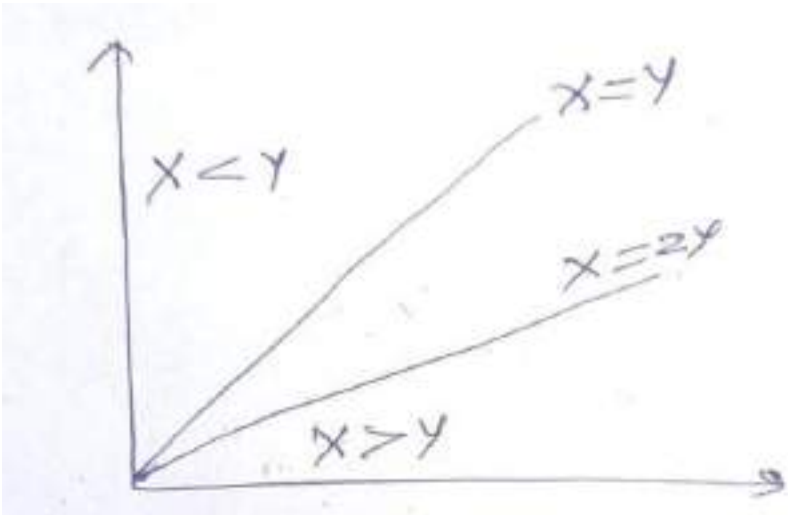
which implies that X and Y are independent

(b) $P(X > 1) = \int_1^\infty f_X(x)dx = \int_1^\infty e^{-x}dx = \left[\frac{x^{-x}}{-1}\right]_1^\infty = \frac{1}{e}$

(c) P(X < Y/X < 2 Y)

$= \frac{P(X < Y \cap X < 2Y)}{P(X < 2Y)}$

$= \frac{P(X < Y)}{P(X < 2Y)}$



Now, $P(X < Y) = \int_{y=0}^{\infty}\int_{x=0}^{y} e^{-(x+y)}dx\,dy$

$= \int_{y=0}^{\infty} e^{-y} \left[\int_0^y e^{-x}dx\right]dy$

$= -\int_{y=0}^{\infty} e^{-y}(e^{-y}-1)dy$

$= \left[\frac{e^{-2y}}{-2}+e^{-y}\right]_0^\infty$

$$= -\left[0-\left(-\frac{1}{2}+1\right)\right] = \frac{1}{2}$$

and P(X < 2Y) = $\int\limits_{y=0}^{\infty}\int\limits_{x=0}^{2y} e^{-(x+y)}dx\,dy$

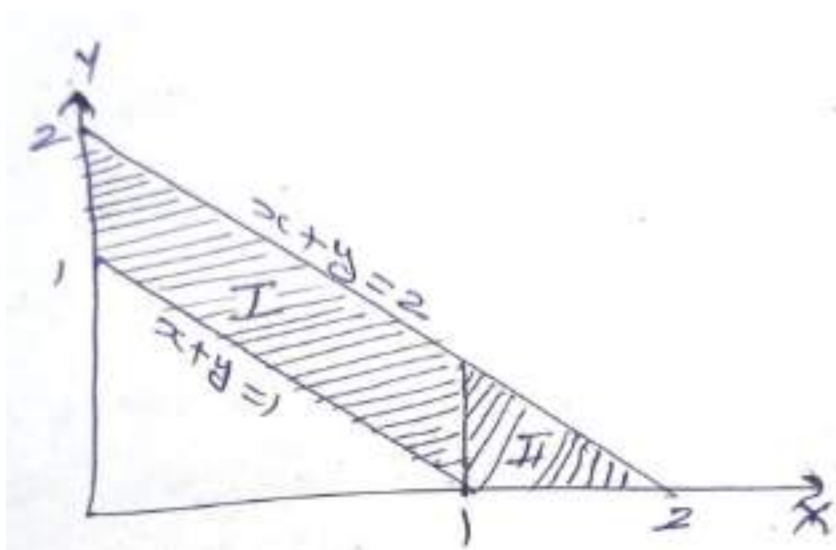
$$= \int\limits_{y=0}^{\infty} e^{-y}\left[\int\limits_{0}^{2y} e^{-x}dx\right] dy$$

$$= \int\limits_{y=0}^{\infty} e^{-y}\,(e^{-2y}-1)\,dy$$

$$= \left[\frac{e^{-3y}}{-3}+e^{-y}\right]_0^{\infty} = 1 - \frac{1}{3} = \frac{2}{3}$$

Therefore, P(X < Y/X < 2Y) = $\dfrac{1/2}{2/3} = \dfrac{3}{4}$

(d) P(1 < X + Y < 2) = $f(x,y)dxdy$ + $\iint\limits_{II} f(x,y)dxdy$



$$= \int\limits_{0}^{1}\left[\int\limits_{1=x}^{2-x} f(x,y)dy\right]dx + \int\limits_{1}^{2}\left[\int\limits_{0}^{2-x} f(x,y)dy\right] dx$$

$$= \int\limits_{1}^{2} e^{-x}\left[\int\limits_{1-x}^{2-x} e^{-y}dy\right] dx + \int\limits_{1}^{2} e^{-x}\left[\int\limits_{0}^{2-x} e^{-y}dy\right] dx$$

$$= \int_0^1 \frac{e^{-x}}{-1} (e^{x-2} - e^{x-1})\, dx + \int_1^2 \frac{e^{-x}}{-1} (e^{x-2}-1)\, dx$$

$$= - (e^{-2} - e^{-1}) \int_0^1 1dx - \int_1^2 (e^{-2} - e^{-x})dx$$

$$= - (e^{-2} - e^{-1}) \left[x\right]_0^1 - \left[e^{-2} + e^{-x}\right]_1^2$$

$$= \frac{2}{e} - \frac{3}{e^2}$$

Self-Check Exercise

Q.1 The joint probability density function of a two-dimensional random variable (X, Y) is given by

$$f(x, y) = \begin{cases} 2, 0 < x < 1, 0 < y < x \\ 0 \quad , \quad elsewhere \end{cases}$$

(a) Find the marginal density functions of X and Y.

(b) Find the conditional density function of Y given

X = x and conditional density function of X given Y = y.

(c) Check for independence of X and Y.

Q.2 Let X_1 and X_2 be jointly distributed with probability density function

$$f(x_1, x_2) = x_1 + x_2,\ 0 < x_1 < 1,\ 0 < x_2 < 1$$
$$= 0, \text{ elsewhere}$$

show that the random variables X1 and X2 are not independent.

18.5 Summary

We conclude this unit by summarizing what we have covered in it:

1. Defined stochastic intendance of two random variables.
2. Discussed stochastic independence of two random variables in detail.
3. Proved the theorem on stochastic independence of two random variables.
4. Examples are given related to each topic so that the contents be clarified further.

18.6 Glossary:

1. Two random variables X and Y with joint probability density function (or probability mass function) $f_{XY}(x, y)$ (or $f(x, y)$) and marginal probability density

functions $f_X(x)$ and $g_Y(y)$ are said to be stochastically independent if and only if $f_{XY}(x, y) = f_X(x)\, g_Y(y)\ \forall$ x and y.

2. Two jointly distributed random variables X and Y are stochastically independent if and only if their joint distribution function F(x, y) is the product of their marginal distribution function $F_X(x)$ and $G_Y(y)$ i.e. $F_{XY}(x, y) = F_X(x)\, G_Y(y)\ \forall$ x and y.

18.7 Answer to Self Check Exercise

Ans.1 (a) The marginal probability density function of X is $f_X(x) = 2x, 0 < x < 0$

$= 0$, other x.

The marginal probability density function of Y is

$g_Y(y) = 2(1 - y), 0 < y < 1$

$= 0$, other y

(b) The conditional density function of Y given X,

(0 < x < 1) is $f_{Y/X}(y/x) = \dfrac{1}{x}, 0 < y < x$

The conditional density function of X given Y,

(0 < y < 1) is $f_{X/Y}(x/y) = \dfrac{f(x,y)}{g_Y(y)} = \dfrac{1}{1-y}, y < x < 1$

(c) X and Y are not independent

Ans.2 $f(x_1, x_2) = x_1 + x_2, 0 < x_1 < 1, 0 < x_2 < 1$

$= 0$, elsewhere

$f_1(x_1) = x_1 + \dfrac{1}{2}, 0 < x_1 < 1$

$= 0$, elsewhere

$f_2(x_2) = \dfrac{1}{2} + x_2, 0 < x_2 < 1$

$= 0$, elsewhere

$f(x_1, x_2) \neq f_1(x_1)\, f_2(x_2)$

$\therefore$ X and Y are not independent

18.8 References/Suggested Readings

1. Robert V. Hogg, Joseph w. Mckean and Allen T. craig, Introduction to Mathematical statistics, Pearson Education, Asia, 2007.

2. Irwim Miller, Marylees Miller and John E. Freund, Mathematical Statistics with Application, 7th Ed., Pearson Education, Asia, 2006.

3. Sheldon Ross, Introduction to Probability Model, 9th Ed., Academic Press, Indian Reprint, 2007.

18.9 Terminal Questions

1. The joint probability function of two discrete random variables X and Y is given by

$$f(x, y) = \begin{cases} \subset (2x+y), 0 \le x \le 2, 0 \le y \le 3 \\ 0 \quad , \quad elsewhere \end{cases}$$

   (a) Find the value of the constant $\subset$

   (b) Find P(X = 2, Y = 1)

   (c) Find P(X $\geq$ 1, Y $\leq$ 2)

   (d) Find the marginal probability function of X

   (e) Find the marginal probability function of Y

   (f) Find f(y/2)

   (g) Find P(Y = 1/X = 2)

   (h) Show that X and Y are dependent

2. Let X_1 and X_2 be jointly distributed with probability density function

$$f(x_1, x_2) = \begin{cases} \frac{1}{4}(1+x_1x_2), |x_1|<1, |x_2|<1 \\ 0 \quad , \quad elsewhere \end{cases}$$

   Show that X_1 and X_2 are not independent whereas X_1^2 and X_2^2 are independent.

3. Given the independent random variables X_1, X_2 and X_3 with probability densities:-

$$f_1(x_1) = \begin{cases} e^{-x_1}, x_1 > 0 \\ 0 \ , \ elsewhere \end{cases}$$

$$f_2(x_2) = \begin{cases} 2e^{-2x_2}, x_2 > 0 \\ 0 \ , \ elsewhere \end{cases}$$

$$f_3(x_3) \begin{cases} 3e^{-3x_3}, x_3 > 0 \\ 0 \ , \ elsewhere \end{cases}$$

   Find P(X_1 + X_2 $\leq$ 1, X_3 > 1)

Unit - 19

Expectation of Function of Two Random Variables

Structure

19.1 Introduction

Bivariate expectation, also known as the expected value of a bivariate random variable, is a fundamental concept in probability and statistics that deals with the joint distribution of two random variables. It is a generalization of the univariate expectation, which is the expected value of a single random variable. In the case of a bivariate random variable (X, Y), the bivariate expectation is denoted as E[XY] and is defined as the average or expected value of the product of the two random variables, X and Y. In probability and statistics, bivariate expectation is an essential concept. It helps in understanding the relationship between two random variables and their joint distribution. It is used in the calculation of covariance and correlation, which are measures of the linear relationship between two random variables. It is a fundamental building block for more advanced statistical concepts, such as regression analysis and multivariate analysis.

19.2 Learning Objectives

After studying this unit, you should be able to:

- Define bivariate expectation
- Prove addition theorem of expectation for two random variables
- Prove multiplication theorem of expectation for two random variables
- Prove theorem of expectation of a linear combination of random variables
- Prove some other important theorems of two random variables.

19.3 Bivariate Expectation-Definition

Let f(x, y) be a function of two dimensional random variable. (X, Y) and p(x, y) denotes the joint p.m.f. or joint p.d.f. of random variables (X, Y). The joint expectation of function f(x, y_ of two random variables (X, Y) is denoted by E[f(X, Y)' and is defined as

$$E[f(X, Y)] = \begin{cases} \sum_i \sum_j f(x, y)\, p(x, y); if\ X\ and\ Y\ are\ discreter\ r.v.'s \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)\, p(x, y) dx\, dy; if\ X\ and\ Y\ are\ continuous\ r.v.'s \end{cases}$$

provided exist.

19.4 Theorems of Expectation

Theorem I: Addition Theorem of Expectation

If X and Y are random variables, then

E(X + Y) = E(X) + E(Y)

provided all the expectations exist.

Proof: Let X and Y be continuous random variables with joint p.d.f. h(x, y) and marginal probability density functions f(x) and g(y) respectively. Then by definition

$$f(x) = \int_{-\infty}^{\infty} h(x, y) dy\ and\ g(y)$$

$$\int_{-\infty}^{\infty} h(x, y) dx$$

and $$E(X) = \int_{-\infty}^{\infty} x\, f(x) dx \quad .....(1)$$

and $$E(Y) = \int_{-\infty}^{\infty} y\, g(y) dy \quad ......(2)$$

$$\therefore \quad E(X + Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x + y) h(x, y) dx\, dy$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x\, h(x, y) dx\, dy + \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y\, h(x, y) dx\, dy$$

$$= \int_{-\infty}^{\infty} x \left[\int_{-\infty}^{\infty} h(x, y) dy \right] dx + \int_{-\infty}^{\infty} y \left[\int_{-\infty}^{\infty} h(x, y) dx \right] dy$$

$$= \int_{-\infty}^{\infty} x\, f(x) dx + \int_{-\infty}^{\infty} y\, g(y) dy$$

$\Rightarrow$ E(X + Y) = E(X) + E(Y)

Hence the result

Generalization: The mathematical expectation f the sum of n random variables is equal to the sum of their expectations, provided all the expectations exist. Symbolically, if X_1, X_2, ..., X_n are random variables, then

$$E(X_1 + X_2 + \dots + X_n) = E(X_1) + E(X_2) + \dots + E(X_n)$$

or $E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E(X_i)$, if all the expectations exist.

Theorem II: Multiplication Theorem of Expectation

If X and Y are independent random variables, then E(XY) = E(X) E(Y)

Proof: Let h(x, y) be the joint probability density function of X and Y.

Let f(x) and g(y) be the marginal probability density functions of X and Y respectively. Since X and Y are independent, therefore,

$$h(x, y) = f(x)\, g(y) \qquad ....(1)$$

$$\therefore \quad E(XY) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x\, y\, h(x, y) dx\, dy$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x\, y\, f(x) g(y) dx\, dy \qquad [\because \text{ X, Y are independent}]$$

$$= \int_{-\infty}^{\infty} x\, f(x) dx \int_{-\infty}^{\infty} y\, g(y) dy$$

$\Rightarrow$ E(XY) = E(X) E(Y)

When X and Y are independent

Hence the result

Generalization: The mathematical expectation of the product of a number of independent random variables is equal to the product of their expectations i.e. if X1, X2,......, Xn are n independent random variables then

$$E(X_1, X_2, \dots, X_n) = E(X_1)\, E(X_2) \dots E(X_n)$$

i.e. $E\left(\prod_{i=1}^{n} X_i\right) = \prod_{i=1}^{n} E(X_i)$

provided all the expectations exists.

Theorem III: Expectation of a Linear Combination of Random Variables

If X_1, X_2, ..., X_n be any n random variables and if a_1, a_2, ..., a_n are an n constants, then

$$E\left[\sum_{i=1}^{n} a_i X_i\right] = \sum_{i=1}^{n} a_i \left(X_i\right)$$, provided all the expectations exist.

Proof: Since E(a X) = a E (X)

and $E\left(\sum x_i\right) = \sum E\left(x_i\right)$

$$\therefore \quad E\left[\sum_{i=1}^{n} a_i X_i\right] = \sum_{i=1}^{n}\left(a_i X_i\right) = \sum_{i=1}^{n} a_i E\left(X_i\right)$$

Hence the result

Theorem IV: If X and Y are two random variables s.t. X $\geq$ Y, then E(X) $\geq$ E(Y), provided that the expectations exist.

Proof: Since X $\geq$ Y, $\therefore$ X - Y $\geq$ 0

Hence, E(X - Y) $\geq$ 0 [$\because$ if X > 0 then E(X) $\geq$ 0]

Since E(X - Y) = E(X) - E(Y)

$\therefore$ E(X - Y) $\geq$ 0

$\Rightarrow$ E(X) - E(Y) $\geq$ 0

$\Rightarrow$ E(X) $\geq$ E(Y)

Theorem V: |E(X)| $\leq$ E|X|, provided all the expectations exist.

Proof: By Theorem IV, we note that since

E(X) $\leq$ E|X|, X $\leq$ |X|

and - X $\leq$ |X| ....(1)

$\Rightarrow$ E(-X) $\leq$ E|X|

i.e. - E(X) $\leq$ E|X| ......(2)

$\therefore$ from (1) and (2), we get

Max {E(x), - E(X)} $\leq$ E|X|

$\Rightarrow$ |E(x)| $\leq$ E|X|

Theorem VI: If X and Y are independent random variables, then

E[h(X). k(Y)] = E[h(X)] E[k(Y)]

where h is a function of X alone and k is a function of Y alone, provided expectations on both sides exist.

Proof: Let f_1(x) and f_2(y) be the marginal probability density functions of X and Y respectively. Since X and Y are independent, their joint probability density function f(x, y) is given by

$$f(x, y) = f_1(x)\, f_2(y)$$

∴ By definition

$$E[h(X)\, k(Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x)k(y)\, f(x, y)\, dx\, dy$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x)k(y)\, f_1(x)\, f_2(y)\, dx\, dy$$

Since E[h(X) k(Y)] exists, the R.H.S. is absolutely convergent and thus the order of integration can be changed to get

$$E\,[h(X)\, k(Y)] = \left[\int_{-\infty}^{\infty} h(x)\, f_1(x)dx\right]\left[\int_{-\infty}^{\infty} k(y)\, f_2(y)dy\right]$$

⇒ E [h(X) k(Y)] = E [h(X)] E [k(Y)]

Hence the result

Let us consider the following examples to clear the idea:

Example 1: Show that the expected value of X is equal to the expectation of the conditional expectation of X given Y. Symbolically,

$$E(X) = E\{E(X/Y)\}$$

Sol: We shall prove this result for discrete case.

By definition

$$E\{E(X/Y)\} = E\left\{\sum_i x_i P(X = x_i / Y = y_j)\right\}$$

$$= E\left\{\sum_i x_i \frac{P(X = x_i \cap Y = y_j)}{P(Y = y_j)}\right\}$$

$$= \sum_j \left[\sum_i \left\{x_i \frac{P(X = x_i \cap Y = y_j)}{P(Y = y_j)}\right\}\right] P(Y = y_j)$$

$$= \sum_j \sum_i x_i P\left(X = x_i \cap Y = y_j\right)$$

$$= \sum_i \left[x_i \left\{ \sum_j P(X = x_i \cap Y = y_j) \right\} \right]$$

$$= \sum_j x_i P(X = x_i)$$

= E(X)

Hence the result

Example 2: Let A and B be two mutually exclusive events, then

$$E(X/A\cup B) = \frac{P(A)E(X/A) + P(B)E(X/B)}{P(A \cup B)}$$

Sol: By definition

$$E(X/A\cup B) = \frac{1}{P(A \cup B)} \sum_{x_i \in A \cup B} x_i P\left(X = x_i\right) \qquad ......(1)$$

Fince A and B are mutually exclusive events,

$$\sum_{x_i \in A \cup B} x_i P\left(X = x_i\right) = \sum_{x_i \in A} x_i P\left(X = x_i\right) + \sum_{x_i \in B} x_i P\left(X = x_i\right)$$

$\therefore$ Using this in (1), we have

$$E(X/A\cup B) = \frac{1}{P(A \cup B)} \left[\sum_{x_i \in A} x_i P\left(X = x_i\right) + \sum_{x_i \in B} x_i P\left(X = x_i\right) \right]$$

$$= \frac{1}{P(A \cup B)} \left[P(A) \frac{\sum_{x_i \in A} x_i P\left(X = x_i\right)}{P(A)} + P(B) \frac{\sum_{x_i \in B} x_i P\left(X = x_i\right)}{P(B)} \right]$$

$$\Rightarrow \quad E(X/A\cup B) = \frac{P(A)E(X/A) + P(B)E(X/B)}{P(A \cup B)}$$

Hence the result

Example 3: The joint probability distribution function of two discrete random variables X and Y is given by

$$f(x, y) = \begin{cases} \dfrac{2x + y}{4^2}; 0 \le x \le 2, 0 \le y \le 3 \\ \quad 0 \quad ; \quad elsewhere \end{cases}$$

Find

(i) E(X) (ii) E(Y) (iii) E(XY)

(iv) E(X^2) (v) E(Y^2)

Sol: (i) E(X) = $\sum_x \sum_y x\, f(x,y) = \sum_x x\left[\sum_y f(x,y)\right]$

$$= \sum_{x=0}^{2}\sum_{y=0}^{3}\frac{x(2x+y)}{42}$$

$$= \sum_{x=0}^{2} x\left[\frac{2x}{42}+\frac{2x+1}{42}+\frac{2x+2}{42}+\frac{2x+3}{42}\right]$$

$$= \sum_{x=0}^{2} x\left[\frac{2x+6}{42}\right]$$

$$= 0.\frac{6}{42} + 1.\frac{14}{42} + 2.\frac{22}{42}$$

$$= \frac{58}{42} = \frac{29}{21}$$

(ii) E(Y) = $\sum_x \sum_y y\, f(x,y) = \sum_y y\left[\sum_y f(x,y)\right]$

$$= \sum_{x=0}^{3}\sum_{y=0}^{2}\frac{y(2x+y)}{42}$$

$$= \sum_{y=0}^{3} y\left[\frac{y}{42}+\frac{2+y}{42}+\frac{4+y}{42}\right]$$

$$= \sum_{y=0}^{3} y\left[\frac{6+3y}{42}\right]$$

$$= 0.\frac{6}{42} + 1.\frac{9}{42} + 2.\frac{12}{42} + 3.\frac{15}{42}$$

$$= \frac{78}{42} = \frac{13}{7}$$

(iii) E(Y) = $\sum_x \sum_y xy\, f(x,y) = \sum_{x=0}^{2}\sum_{y=0}^{3} xy\left[\frac{2x+y}{42}\right]$

$$= \sum_{x=0}^{2} \left[0 + \frac{(2x+1)}{42} + \frac{2x(2x+2)}{42} + \frac{3x(2x+3)}{42}\right]$$

$$= \sum_{x=0}^{2} \left[\frac{12x^2+14x}{42}\right] = 0 + \frac{12+14}{42} + \frac{48+28}{42}$$

$$= \frac{102}{42} + \frac{17}{7}$$

(iv) $E(X^2) = \sum_x \sum_y x^2 f(x, y) = \sum_x x^2 \left[\sum_y f(x,y)\right]$

$$= 0^2 . \frac{6}{42} + 1^2 . \frac{14}{42} + 2^2 . \frac{22}{42}$$

$$= \frac{102}{42} = \frac{17}{7}$$

(v) $E(Y^2) = \sum_x \sum_y y^2 f(x, y) = \sum_y y^2 \left[\sum_x f(x,y)\right]$

$$= 0^2 . \frac{6}{42} + 1^2 . \frac{9}{42} + 2^2 . \frac{12}{42} + 3^2 . \frac{15}{42}$$

$$= \frac{192}{42}$$

$$= \frac{32}{7}$$

Self-Check Exercise

Q.1 The joint density function of two continuous random variables X and Y is given by

$$f(x, y) = \begin{cases} \frac{2x+y}{210} ; 2 < x < 6, 0 < y < 5 \\ 0 \quad ; \quad elsewhere \end{cases}$$

Find

(i) E(X) (ii) E(Y) (iii) E(XY)

(iv) $E(X^2)$ (iv) $E(Y^2)$

Q.2 If X and Y are two random variables s.t. X $\geq$ Y, then prove that E(X) $\geq$ E(Y), provided that the expectations exists.

19.5 Summary

We conclude this unit by summarizing what we have covered in it:

1. Defined joint expectation of function $f(x, y)$ of two random variables (X, Y).
2. Proved addition theorem of expectation of two random variables.
3. Proved multiplication theorem of expectation of two random variables.
4. Proved theorem on expectation of a linear combination of random variables.
5. Proved some other important theorems of expectations of two random variables.
6. Some examples are given related so that the contents be clarified further.

19.6 Glossary:

1. Let $f(x, y)$ be a function of two dimensional random variable (X, Y) and p(x, y) denotes the joint p.m.f. or joint p.d.f. of random variable (X, Y). The joint expectation of function $f(x, y)$ of two random variables (X, Y) is denoted by $E[f(X, Y)]$ and is defined us

$$E[f(X, Y)] = \begin{cases} \sum_i \sum_j f(x,y)p(x,y); \text{if X and Y arediscrete r.v.'s} \\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)p(x,y)dx\,dy; \text{if X and Y are continuous r.v.'s} \end{cases}$$

2. If X and Y are random variables, then E(X + Y) = E(X) + E(Y), provided all the expectations exist.
3. If X and Y are independent random variables, then E(XY) = E(X) E(Y)

19.7 Answer to Self Check Exercise

Ans.1 (i) $\frac{268}{63}$ (ii) $\frac{170}{63}$ (iii) $\frac{80}{7}$

(iv) $\frac{1220}{63}$ (v) $\frac{1175}{126}$

19.8 References/Suggested Readings

1. Robert V. Hogg, Joseph w. Mckean and Allen T. craig, Introduction to Mathematical statistics, Pearson Education, Asia, 2007.
2. Irwim Miller, Marylees Miller and John E. Freund, Mathematical Statistics with Application, 7th Ed., Pearson Education, Asia, 2006.
3. Sheldon Ross, Introduction to Probability Model, 9th Ed., Academic Press, Indian Reprint, 2007.

19.9 Terminal Questions

1. Let X and Y be two random variables each taking three values -1, 0 and 1, and having the joint probability distribution.

X Y	-1	0	1	Total
-1	0	0.1	0.1	0.2
0	0.2	0.2	0.2	0.6
1	0	0.1	0.1	0.2
Total	0.2	0.4	0.4	1.0

Show that X and Y have different expectations.

2. State and prove addition theorem of expectation.

3. If X_1, X_2, ..., X_n be any n random variables and if a_1, a_2, ..., a_n are any n constants, then prove that

$$E\left[\sum_{i=1}^{n} a_i X_i\right] = \sum_{i=1}^{n} a_i E(X_i)$$, provided all the expectations exist.

Unit - 20

Covariance, Conditional Expectations and Conditional Variance

Structure

20.1 Introduction

Covariance is a measure of the strength and direction of the linear relationship between two random variables. It is defined as the expected value of the product of the deviations of two random variables from their respective means. The covariance can take positive, negative or zero values, indicating a positive, negative or no linear relationship between the variables, respectively. It is useful in portfolio optimization, risk management and multivariate statistical analysis.

Conditional expectation is the expected value of a random variable given the value of another random variable or set of random variables. It is a powerful tool in statistical inference as it allows us to make predictions about one variable based on the information provided by another variable. It is also used in the calculation of conditional variance and the derivation of regression models.

Conditional variance is the variance of a random variable given the value of another random variable or set of random variables. It measures the spread or dispersion of the distribution of X around its conditional mean, given the value of Y.

These concepts are fundamental in probability theory, statistics and many areas of applied mathematics and data science. Understanding and applying them effectively can provide valuable insights and enable more informed decision making.

20.2 Learning Objectives

After studying this unit, you should be able to:

- Define covariance
- Discuss properties of covariance
- Prove theorem of variance of a linear combination of random variables
- Define conditional expectation and conditional variance for discrete random variables
- Define conditional expectation and conditional variance for continuous random variables.

20.3 Covariance

Definition: If X and Y are two random variables, them the covariance between them is defined as

$$\text{Cov}(X, Y) = E[\{X - E(X)\}\{Y - E(Y)\}]$$

provided the expectations exist

Properties of Covariance

Property I: $\text{Cov}(X, Y) = E(XY) - E(X)E(Y)$

Proof: $\text{Cov}(X, Y) = E[\{X - E(X)\}\{Y - E(Y)\}]$

$$= E[XY - XE(Y) - YE(X) + E(X)E(Y)]$$

$$= E(XY) - E(X)E(Y) - E(Y)E(X) + E(X)E(Y)$$

$\Rightarrow$ $\text{Cov}(X, Y) = E(XY) - E(X)E(Y)$

Hence the result

Cor. : If X and Y are independent, then

$$E(XY) = E(X)E(Y)$$

$\therefore$ $\text{Cov}(X, Y) = E(XY) - E(X)E(Y)$

$$= E(X)E(Y) - E(X)E(Y)$$

$$= 0$$

$\Rightarrow$ $\text{Cov.}(X, Y) = 0$

Property II: $\text{Cov}(aX, bY) = ab\,\text{Cov}(X, Y)$

Proof: $\text{Cov}(aX, bY) = E[\{aX - E(aX)\}\{bY - E(bY)\}]$

$$= E[a\{X - E(X)\}\, b\{Y - E(Y)\}]$$

$$= abE[\{X - E(X)\}\{Y - E(Y)\}]$$

$$= ab\,\text{Cov}(X, Y)$$

i.e. covariance is not independent of scale.

Property III: Cov (X + a, Y + b) = Cov (X, Y)

Proof: Cov (X + a, Y + b) = E [{(X + a) = E (X + a)}] [{(X + b) - E (X + b)}]

= E[{X + a - E (X) - a} {Y + b - E(Y) - b}]

= E[{X - E(X)} {Y - E(Y)}]

= Cov (X, Y)

i.e. covariance is independent of change of origin.

Property IV: Cov $\left(\frac{X-\mu}{\sigma_X}, \frac{Y-\mu}{\sigma_Y}\right) = \frac{1}{\sigma_X \sigma_Y}$ Cov (X, Y)

Proof: Follows from property II and III.

Property V: Cov (aX + b, cY + d) = ac Cov (X, Y)

or Cov (X + Y, Z) = Cov (X, Z) + Cov (Y, Z)

or Cov (aX + bY, cX + dY) = ac σ_X^2 + bd σ_Y^2 + (ad + bc) Cov (X, Y)

Property VI: If X and Y are independent, then Cov (X, Y) = 0. However, the converse is not true.

Proof: Let if possible

X = U - V and Y = U + V

∴ $\bar{X} = \bar{U} - V$ and $\bar{Y} = \bar{U} + \bar{V}$

∴ $X - \bar{X} = (U - \bar{U}) - (V - \bar{V})$

and $Y - \bar{Y} = (U - \bar{U}) + (V - \bar{V})$

∴ Cov (X, Y) = E[(X - E(X)) (Y - E(Y))]

$= E[(X - \bar{X}) (Y - \bar{Y})]$

$= E[\{(U - \bar{U}) - (V - \bar{V})\} \{U - \bar{U} + (V - \bar{V})\}]$

$= E[\{(U - \bar{U})^2 - (V - \bar{V})^2\}]$

$= E[(U - \bar{U})^2] - E[(V - \bar{V})^2]$

= Var (U) - Var (V)

i.e. Cov (X, Y) = Var (U) - Var (V)

If we choose U and V such that Var (U) = Var (V), then Cov (X, Y) = 0, however X = U - V and Y = U + V or X + Y = 2 U i.e. X and Y are dependent.

i.e. if for two random variables X and Y, Cor (X, Y) = 0, then X and Y need not be independent.

Hence, converse is not true.

20.4 Variance of a Linear Combination of Random Variables

Statement: Let $X_1, X_2, \ldots, X_n$ be n random variables, then

$$\text{Var}\left[\sum_{i=1}^{n} a_i X_i\right] = \sum_{i=1}^{n} a_i^2 Var(X_i) + 2\sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j \,.\, \text{Cov}(X_i, X_j)$$

Proof: Let $Y = a_1X_1 + a_2X_2 + \ldots\ldots + a_nX_n$ ......(1)

$\Rightarrow$ $E(Y) = a, E(X_1) + a_2E(X_2) + \ldots\ldots + a_n E(X_n)$ ....(2)

$\therefore$ (1) - (2) implies

$$Y - E(Y) = a_1\left[X_1 - E(X_1)\right] + a_2\left[X_2 - E(X_2)\right] + \ldots\ldots + a_n\left[X_n - E(X_n)\right]$$

Squaring and taking expectation of both side, we get

$$E[Y - E(Y)]^2 = E[a_1\{X_1 - E(X_1)\} + a_2\{X_2 - E(X_2) + \ldots\ldots + a_n\{X_n - E(X_n)\}]^2$$

$$\Rightarrow \quad E[Y - E(Y)]^2 = a_1^2 E[X_1 - E(X_1)]^2 + a_2^2 E[X_2 - E(X_2)]^2 + \ldots\ldots +$$

$$a_n^2 E[X_n - E(Xn)]^2 + 2\sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j E[\{X_i - E(X_i)\}\{X_j - E(X_j)\}]$$

$$\Rightarrow \quad \text{Var } Y = a_1^2 \text{Var}(X_1) + a_2^2 \text{Var}(X_2) + \ldots\ldots + a_n^2 \text{Var}(X_n) + 2\sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j \text{Cov}(X_i, X_j)$$

$$\Rightarrow \quad \text{Var}\left[\sum_{i=1}^{n} a_i X_i\right] = \sum_{i=1}^{n} a_i^2 \text{Var}(X_i) + 2\sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j \text{Cov}(X_i, X_j)$$

Hence the result

20.5 Conditional Expectation and Conditional Variance

(i) For Discrete Random Variables:

The conditional expectation or mean value of a function g(X, Y) given that $Y = y_j$, is defined by

$$E[g(X, Y)/Y = y_j] = \sum_{i=1}^{\infty} g(x_i, y_j) P(X = x_i/Y = y_j)$$

$$= \frac{\sum_i g(x_i, y_j(P(X = x_i \cap Y = y_j}{P(Y = y_j)}$$

In particular,

$$E(X/Y = y_j) = \sum_{i=1}^{\infty} x_i P(X = x_i / Y = y_j)$$

The conditional variance of X given $Y = y_j$ is given by

$V(X/Y = y_j) = E[\{X/Y = y_j)\}^2 \; Y = y_j]$

(ii) For continuous Random Variables:

The conditional expectation of g(X, Y) given Y = y is defined by

$$E[g(X, Y)/ Y = y] = \sum_{-\infty}^{\infty} g(x, y) f_{X/Y}(x / y) dx$$

$$= \int_{-\infty}^{\infty} \frac{g(x, y) f(x, y)}{f_Y(y)} dx$$

In particular,

$$E(X/Y = y) = \int_{-\infty}^{\infty} \frac{x\, f(x, y) dx}{f_Y(y)}$$

Similarly, $E(Y.X = x) = \int_{-\infty}^{\infty} \frac{y\, f(x, y)}{f_X(x)} dy$

The conditional variance of X given Y = y be defined as

$$V(X/Y = y) = E[\{X - E(X/Y = y)\}^2 / Y = y]$$

and similarly,

$$V(Y/X = x) = E[\{Y - E[Y/X = x)\}^2] \; X = x]$$

Let us improve our understanding of these results by looking at some of the following examples:-

Example 1: Two random variables X and Y have the following joint probability density function

$$f(x, y) = \begin{cases} 2 - x - y; 0 \le x \le 1, 0 \le y \le 1 \\ \quad 0 \quad ; \; otherwise \end{cases}$$

Find (i) marginal probability density functions of X and Y

(ii) Conditional density functions

(iii) Var (X) and Var (Y); and

(iv) Covariance between X and Y

Sol: (i) By definition

$$f_X(x) = \sum_{-\infty}^{\infty} f(x, y) dy$$

$$= \int_0^1 (2 - x - y) dy$$

$= \frac{3}{2} - x$

$\therefore \quad f_X(x) = \begin{cases} \frac{3}{2} - x; 0 < x < 1 \\ 0 \quad ; \ otherwise \end{cases}$

Similarly, $\quad f_Y(y) = \begin{cases} \frac{3}{2} - y; 0 < y < 1 \\ 0 \quad ; \ otherwise \end{cases}$

(ii) By definition

$$f_{X/Y}(x/y) = \frac{F_{XY}(x,y)}{f_Y(y)} = \frac{(2-x-y)}{\left(\frac{3}{2} - y\right)}, \ 0 < (x, y) < 1$$

$$\therefore \quad f_{Y/X}(y/x) = \frac{F_{XY}(x,y)}{f_X(y)} = \frac{(2-x-y)}{\left(\frac{3}{2} - x\right)}, \ 0 < (x, y) < 1$$

(iii) $E(X) = \int_0^1 x \, f_X(x) dx = \int_0^1 x\left(\frac{3}{2} - x\right) dx = \frac{5}{12}$

$E(Y) = \int_0^1 y \, f_Y(y) dy = \int_0^1 y\left(\frac{3}{2} - y\right) dy = \frac{5}{12}$

$E(X^2) = \int_0^1 x^2\left(\frac{3}{2} - x\right) dx = \left|\frac{3}{6}x^3 - \frac{x^4}{4}\right|_0^1 = \frac{1}{4}$

$\therefore \quad V(X) = E(X^2) - \{E(X)\}^2$

$= \frac{1}{4} - \frac{25}{144} = \frac{11}{144}$

Similarly, $V(Y) = \frac{11}{144}$

(iv) $E(XY) = \int_0^1\int_0^1 xy(2-x-y)dx\,dy$

$= \int_0^1 \left[2\frac{x^2y}{2} - \frac{x^3y}{3} - \frac{x^2y^2}{2}\right]_{x=0}^{x=1} dy$

$$= \int_0^1 x \left(\frac{2}{3}y - \frac{1}{2}y^2\right) dy$$

$$= \left[\frac{y^2}{3} - \frac{y^3}{4}\right]_0^1 = \frac{1}{6}$$

$\therefore$ Cov (X, Y) = E(XY) - E(X) E(Y)

$$= \frac{1}{6} - \frac{5}{12} \cdot \frac{5}{12} = -\frac{1}{144}$$

Example 2: The joint p.d.f. of bivariate random variable (X, Y) is given by

$$f_{XY}(x, y) = \begin{cases} 6xy(2-x-y); 0 < x < 1, 0 < y < 1-y; 0 < y < 1 \\ 0 \qquad ;\ otherwise \end{cases}$$

Find conditional expectation of X given Y = y where 0 < y < 1.

Sol: Here $f_Y(Y) = \int_{-\infty}^{\infty} f(x, y)dx$

$$= \int_0^1 6xy(2-x-y)dx$$

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

$$= 6y \left\{\frac{(2-y)}{2} - \frac{1}{3} - 0\right\}$$

$\therefore$ The conditional expectation of X given Y = y is

$$E(X/Y = y) = \int_{-\infty}^{\infty} x.f_{X/Y}\left(\frac{x}{y}\right)dx$$

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

$$= \int_{-\infty}^{\infty} x\frac{6xy(2-x-y)}{(4-3y)y}dx$$

$$= \frac{6}{4-3y}\int_0^1 \left\{(2-y)x^2 - x^3\right\}dx$$

$$= \frac{6}{4-3y}\left|(2-y)\frac{x^3}{3}-\frac{x^4}{4}\right|_0^1$$

$$= \frac{6}{4-3y}\left\{\frac{(2-y)}{3}-\frac{1}{4}\right\}$$

$$= \frac{5-4y}{8-6y}$$

Example 3: The joint probability density function of X_1 and X_2 is

$f(x_1, x_2) = 2,\ 0 < x_1 < x_2 < 1$

$= 0$, elsewhere

Find,

(i) The marginal probability density functions of X_1 and X_2

(ii) The conditional probability density function of X_1, given $X_2 = x_2$

(iii) The conditional mean and conditional variance of X_1, given $X_2 = x_2$; and

and $P\left(0 < X_1 < \frac{1}{2}\right)$

Sol: (i) The marginal probability density function of X_1 is

$$f_1(x_1) = \int_{x_1}^{1} 2dx_2$$

$$= \begin{cases} 2(1-x_1); 0 < x < 1 \\ \quad 0 \quad ; \quad elsewhere \end{cases}$$

The marginal probability density function of X_2 is

$$f_2(x_2) = \int_{x_1=0}^{x_2} 2dx_1$$

$= 2x_2$; $0 < x_2 < 1$

- 0 ; elsewhere

(ii) The conditional probability density function of X_1, given $X_2 = x_2$, $0 < x_2 < 1$ is

$$f_{X_i/X_2}(x_1/x_2) = \frac{f(x_1, x_2)}{f_2(x_2)} = \frac{2}{2x_2}$$

$= \frac{1}{x_2}$; $0 < x_1 < x_2$

= 0 ; elsewhere

(iii) The conditional mean of X_1, given $X_2 = x_2$ is

$$E(X_1/X_2 = x_2) = \int_{-\infty}^{\infty} x_1 f_{X_1/X_2}(x_1/x_2)dx_1$$

$$= \int_{x_1=0}^{x_2} x_1\left(\frac{1}{x_2}\right)dx_1$$

$$= \frac{x_2^2}{2} \cdot \frac{1}{x_2}$$

The conditional variance of X_1 given $X_2 = x_2$ is

$$\text{Var}(X_1/X_2 = x_2) = E\left(\frac{X_1^2}{X_2 = x_2}\right) - \left[E\left(\frac{X_1}{X_2 = x_2}\right)\right]^2$$

$$= \int_{x_1=0}^{x_2} x_1^2 \ \frac{1}{x_2} dx_1 - \left(\frac{x_2}{2}\right)^2$$

$$= \left|\frac{x_1^3}{3}\right|_0^{x_2} \frac{1}{x_2} - \frac{x_2^2}{2^2}$$

$$= \frac{x_2^2}{3} - \frac{x_2^2}{4}$$

$$= \frac{x_2^2}{12}, \text{ where } 0 < x_2 < 1$$

(iv) $$P\left(0 < X_1 < \frac{1}{2} / X_2 = \frac{3}{4}\right) = \int_0^{1/2} f_{X_1/X_2}(x_1/3/4)dx_1$$

$$= \int_0^{1/2} \left(\frac{4}{3}\right)dx_1$$

$$P\left(0 < X_1 < \frac{1}{2}\right) = \int_0^{1/2} f_1(x_1)dx_1 = \int_0^{1/2} 2(1-x_1)dx_1$$

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

Example 4: Let X_1 and X_2 have the joint probability density function

$f(x_1, x_2) = 6x_2$, $0 < x_2 < x1 < 1$

= 0 , elsewhere

Find

(i) The marginal probability density function of X_1;

(ii) The conditional probability density function of X_2, given $X_1 = x_1$;

(iii) If conditional mean of X_2 given $X_1 = x_1$ is random variable, Y(say), then find the distribution function of Y and probability density function of Y;

(iv) The mean and variance of Y.

Sol: (i) The marginal probability density function of X_1 is

$$f_1(x_1) = \int_0^{x_1} 6x_2 dx_2 = 3x_1^2, \quad 0 < x_1 < 1$$

= 0 ; elsewhere

(ii) The conditional probability density function of X_2, given $X_1 = x_1$ is

$$f_{X2/X1}(x_2/x_1) = \frac{f(x_1, x_2)}{f_1(x_1)} = \frac{2x_2}{3x_1^2}$$

$$= \frac{2x_2}{x_1^2}, \quad 0 < x_2 < x_1$$

= 0 ; elsewhere, where $0 < x_1 < 1$

(iii) The conditional mean of X_2, given $X_1 = x_1$ is

$$E(X_2/X_1 = x_1) = \int_0^{x_1} x_2 f_{X_2/X_1}(x_2 / x_1) dx_2$$

$$= \int_0^{x_1} x_2 \frac{6x_2}{3x_1^2} dx_2$$

$$= \frac{6}{3x_1^2} \left| \frac{x_2^3}{3} \right|_0^{x_1}$$

$$= \frac{2x_1}{3}, \quad 0 < x_1 < 1$$

Now, $E(X_2/X_1) = \dfrac{2x_1}{3}$ is a random variable Y.

The distribution function of Y = $\frac{2x_1}{3}$ is

$$G(y) = \Pr(Y \leq y)$$

$$= \Pr\left(X_1 \leq \frac{3y}{2}\right),\ 0 \leq y \leq \frac{2}{3}$$

From the probability density function $f_1(x_1)$, we have

$f_1(x_1)$ = we have

$$G(y) = \int_0^{3y/2} 3x_1^2 dx_1 = \frac{27y^3}{8},\ 0 \leq y \leq \frac{2}{3}$$

Also, $G(y) = 0$, if $y < 0$

and $G(y) = 1$, if $\frac{2}{3} < y$.

Therefore, probability density function of

Y = $\frac{2x_1}{3}$ is

$$g(y) = G'(y) = \frac{81y^2}{8},\ 0 \leq y \leq \frac{2}{3}$$

(iv) Mean of Y $\left(= \frac{2x_1}{3}\right)$ is

$$E(Y) = \int_0^{2/3} \left(\frac{81y^2}{8}\right) dy$$

$$= \frac{81}{8} \left|\frac{y^4}{4}\right|_0^{2/3} = \frac{1}{2}$$

Variance of Y $\left(= \frac{2x_1}{3}\right)$ is

$$\operatorname{Var}(Y) = E(Y^2) - [E(Y)]^2$$

$$= \int_0^{2/3} y^2 \left(\frac{81y^2}{8}\right) dy - \left(\frac{1}{2}\right)^2$$

$$= \frac{81}{8}\left|\frac{y^5}{5}\right|_0^{2/3} - \left(\frac{1}{2}\right)^2$$

$$= \frac{81}{40}\left(\frac{32}{243}\right) - \frac{1}{4}$$

$$= \frac{4}{15} - \frac{1}{4}$$

$$= \frac{1}{60}$$

Example 5: Let the joint probability density function of random variables X and Y be

$f(x, y)$ e^{-y}, $0 < x < y < \infty$

= 0, elsewhere

Find moment generating function of this joint distribution and hence find the correlation coefficient of X and Y

Sol: Given joint probability density function is

$f(x, y) = e^{-y}$, $0 < x < y < \infty$

= 0, elsewhere

Now, the moment generating function of this joint distribution is

$M(t_1, t_2) = E\left(e^{t_1x+t_2y}\right)$

$$= \int_{x=0}^{\infty}\int_{y=x}^{\infty} e^{t_1x}e^{t_2y}e^{-y}\,dy\,dx$$

$$= \int_{x=0}^{\infty} e^{t_1x}\left|\frac{e^{-y(1-t_2)}}{-(1-t_2)}\right|_{y=x}^{\infty} dx \qquad \text{[Here } t_2 > 1\text{]}$$

$$= \int_{x=0}^{\infty} e^{t_1x}\,\frac{6^{-x(1-t_2)}}{-1-t_2}\,dx$$

$$= \frac{1}{1-t_2}\int_{x=0}^{\infty} e^{-x(1-t_1-t_2)}\,dx$$

$$= \frac{1}{1-t_2}\left|\frac{e^{-(1-t_1-t_2)}}{-(1-t_1-t_2)}\right|_0^{\infty}$$

provided that $t_1 + t_2 < 1$.

and $t_2 < 1$

$$= \frac{1}{(1-t_2)(1-t_1-t_2)}\text{, provided that } t_1 + t_2 < 0 \text{ and } t_2 < 1 \qquad ......(1)$$

Further

$$\mu_1 = E(X) = \frac{\partial M(0,0)}{\partial t_1} = \left[\frac{1}{(1-t_1-t_2)^2(1-t_2)}\right]_{(0,0)}$$

$$\mu_2 = E(Y) = \frac{\partial M(0,0)}{\partial t_2} = \left[\frac{1}{(1-t_2)^2} + \frac{1}{(1-t_1-t_2)(1-t_2)^2}\right]_{(0,0)} = 2\ ;$$

$$E(X^2) = \frac{\partial^2 M(0,0)}{\partial t_1^2} = \left[\frac{2}{(1-t_1-t_2)^3(1-t_2)}\right]_{(0,0)} = 2\ ;$$

$$E(Y^2) = \frac{\partial^2 M(0,0)}{\partial t_2^2} = 6\ ;$$

$$\sigma_1^2 = E(X^2) - \mu_1^2 = 2 - 1 = 1\ ;$$

$$\sigma_2^2 = E(Y^2) - \mu_2^2 = 6 - 4 = 2\ ;$$

$$\frac{\partial^2 M(0,0)}{\partial t_1 \partial t_2} = E(XY) = 3;$$

$$E[(X - \mu_1)(Y - \mu_2)] = \frac{\partial^2 M(0,0)}{\partial t_1 \partial t_2} - \mu_1 \mu_2$$

$$= 3 - 2 = 1$$

Now, the correlation coefficient of X and Y is

$$P = \frac{E\left[X-\mu_1)(Y-\mu_2)\right]}{\sigma_1\sigma_2} = \frac{1}{\sqrt{2}}$$

Example 6: Let the two variates X_1 and X_2 have the joint density function $f(x_1, x_2)$, then prove that the conditional mean of X_2 (given X_1) coincides with (unconditional) mean only if the random variables X_1 and X_2 are independent (stochastically)

Sol: The conditional mean of X_2 given X_1 is given by $E(X_2/X_1 = x_1) = \int_{x_2} x_2 f(x_2/x_1)dx_2$ .....(1)

where $f(x_2/x_1)$ is the conditional probability density function of X_2 given by $X_1 = x_1$.

The joint probability density function of X_1 and X_2 is given by

$f(x_1, x_2) = f_1(x_1)\, f(x_2/x_1)$

or $\quad f(x_2/x_1) = \dfrac{f(x_1, x_2)}{f_1(x_1)}$

where $f_1(x_1)$ is the marginal probability density function of X_1

Substituting this value in (1), we get

$$E(X_2/X_1 = x_1) = \int_{x_2} \left[\frac{x_2 f(x_1, x_2)}{f_1(x_1)}\right] dx_2 \qquad .....(2)$$

unconditional mean of X_2 is given by

$$E(X_2) = \int_{x_2} x_2 f_2(x_2) dx_2 \qquad ......(3)$$

From (2) and (3), we conclude that the conditional mean of X_2 (given X_1) will coincide with unconditional mean of X_2 only if

$$\frac{f(x_1, x_2)}{f_1(x_1)} = f_2(x_2)$$

i.e. if $f(x_1, x_2) = f_1(x_1) f_2(x_2)$

i.e. if X_1 and X_2 are independent (stochastic tally)

Hence the required result

Self-Check Exercise

Q.1 The joint p.d.f. of (X, Y) is given by $f_{XY}(x, y) = \begin{cases} 4y(x-y)e^{-(x+y)}; 0 < x < \infty, 0 \le \infty \\ 0 \quad ; \quad otherwise \end{cases}$

Find $E(X/Y = y)$

Q.2 If the random variable X have the marginal density

$$f_1(x) = 1, -\frac{1}{2} < x < \frac{1}{2}$$

and the conditional density of the vairable Y is

$$f(y/x) = \begin{cases} 1, x < y < x+1, -\dfrac{1}{2} < x < 0 \\ 1, -x < y < 1-x, 0 < x < \dfrac{1}{2} \end{cases}$$

then show that variables X and Y are uncorrelated.

Q.3 Let $f(x, y) = \begin{cases} 20x^2y^3; 0 < x < y < 1 \\ 0 \ ; \ otherwise \end{cases}$

be the joint p.d.f. of X and Y. Find the conditional mean and variance of X given Y = y, 0 < y < 1

20.6 Summary

We conclude this unit by summarizing what we have covered in it:

1. Defined covariance
2. Proved different properties of covariance
3. Proved theorem of variance of a linear combination of random variables
4. Defined conditional expectation and conditional variance for discrete random variables
5. Defined conditional expectation and conditional variance for continuous random variables
6. Some examples are given related to each topic so that the contents be clarified further.

20.7 Glossary:

1. If X and Y are two random variables, then the covariance between them is defined as

   Cov (X, Y) = E [{X - E(X)} [Y - E(Y)}],

   provided the expectations exists

2. The conditional expectation or mean value of a function g(X, Y) given that Y = y_j, is defined as

   E[g(X, Y)/Y = y_j] = $\sum_{i=1}^{\infty} g(x_1, y_j)$ P (X = xi/Y = yj)

   $$= \frac{\sum_i g(x_1, y_j) P(X = x_1 \cap = y_j)}{P(Y = y_j)} \qquad \text{[For discrete random variables]}$$

The conditional variance of X given Y = y_j is given by

V(X/Y = yj) = E[{X - E(X/Y = y_j)}^{2}Y = y_j] [For discrete random variables)

3. The conditional expectation of g(X, Y) given Y = y is defined as

   E[g(X, Y)/Y = y] = $\sum_{-\infty}^{\infty} g(x, y) f_{X/Y}(x/y)dx$

$$= \sum_{-\infty}^{\infty} \frac{g(x,y)\, f(x,y)}{f_y(y)} \text{dx} \qquad \text{[For continuous random variables]}$$

In particular

$$E(X/Y = y) = \int_{-\infty}^{\infty} \frac{x\, f(x,y)}{f_y(y)} dx$$

Similarly, $E(Y/X = x) = \int_{-\infty}^{\infty} \frac{y\, f(x,y)}{f_x(y)} dy$ [For continuous random variables]

The conditional variance of X given Y = y is defined as

$$V(X/Y = y) = E[\{X - E(X/Y= y)\}^2/Y=y]$$

and similarly,

$$V(Y/X = x) = E[\{Y - E\,(Y/X = x)\}^2/X = x] \qquad \text{[For continuous random variables]}$$

20.8 Answer to Self Check Exercise

Ans.1 $E[X/Y = y] = \frac{y-2}{y-1}$

Ans. 2 P(X, Y) = 0, hence variables X and Y are uncorrelated.

Ans. 3 Conditional mean of X given Y = y

$$= \frac{3}{4}\, y,\ 0 < y < 1$$

Conditional variance of X given Y = y

$$= \frac{3y^2}{80},\ 0 < y < 1$$

20.9 References/Suggested Readings

1. Robert V. Hogg, Joseph w. Mckean and Allen T. craig, Introduction to Mathematical statistics, Pearson Education, Asia, 2007.
2. Irwim Miller, Marylees Miller and John E. Freund, Mathematical Statistics with Application, 7th Ed., Pearson Education, Asia, 2006.
3. Sheldon Ross, Introduction to Probability Model, 9th Ed., Academic Press, Indian Reprint, 2007.

20.10 Terminal Questions

1. Let X and Y are two independent random variables, then show that the correlation coefficient of X and Y is zero.

2. Joint density function of Bivariate random variable (X, Y) is given by

$$f_{XY}(x, y) = \begin{cases} \frac{1}{2} ye^{-xy}; 0 < x < \infty, 0 < y < 2 \\ 0 \ ; \ \ otherwise \end{cases}$$

Find E $\left[e^{\frac{x}{2}} / Y = 1 \right]$

3. Let the random variables X and Y have the joint probability density function

$f(x, y) = 2, 0 < x < y, 0 < y < 1$

$= 0$, elsewhere

Then, show that

(i) The conditional means are, respectively, $(1 + x)/2, 0 < x < 1$ and $\frac{y}{2}$, $0 < y < 1$;

(ii) The variance of the conditional distribution of Y given X = x is

$\frac{(1-x)^2}{12}$, $0 < x < 1$

and the variance of the conditional distribution of X, given Y = y is $\frac{y^2}{12}$, $0 < y < 1$

4. If the joint probability density function of X and Y is given by

$$f(x, y) = \begin{cases} \frac{1}{4}(2x + y) & , \quad 0 < x < 1, 0 < y < 2 \\ 0 & , \quad otherwise \end{cases}$$

Find the conditional mean and conditional variance of Y given X = $\frac{1}{4}$.
