University of Mumbai

INSTITUTE OF DISTANCE AND OPEN LEARNING

PSMT/PAMT 205

PROBABILITY THEORY

M.Sc. (MATHEMATICS)

SEMESTER - II

PAPER - V

PROBABILITY THEORY

M.Sc. (Mathematics) SEMESTER - II
PAPER - V

Programme Co-ordinator, I/c Faculty of Science & Technology

Asst. Prof. Mandar Bhanushe
Department of Mathematics, University of Mumbai, IDOL

Course Writers

Prof. Kalpana Phal
B. N. Bandodkar College, Thane.

Published by
Dr. Prakash Mahanwar, Director Incharge
on behalf of Institute of Distance & Open Learning,University of Mumbai, Mumbai
and Printed at
Name & Address.

DTP Composed by 7skills

CONTENTS

PROBABILITY THEORY
M.SC. (MATHEMATICS) SEMESTER - II PAPER - V
SYLLABUS

PSMT205/PAMT205 PROBABILITY THEORY

Course Outcomes

1. Students will understand the concept of Modelling Random Experiments, Classical probability spaces, -fields generated by a family of sets, -field of Borel sets, Limit superior and limit inferior for a sequence of events.

2. Students will be able to know about probability measure, Continuity of probabilities, First Borel-Cantelli lemma, Discussion of Lebesgue measure on -field of Borel subsets of assuming its existence, Discussion of Lebesgue integral for non-negative Borel functions assuming its construction.

3. Students will be able to earn knowledge of discrete and absolutely continuous probability measures, conditional probability, total probability formula, Bayes formula.

4. Students will learn distribution of a random variable, distribution function of a random variable, Bernoulli, Binomial, Poisson and Normal distributions,

5. Students will be able to understand Chebyshev inequality, Weak law of large numbers, Convergence of random variables, Kolmogorov strong law of large numbers, Central limit theorem and Application of Probability Theory.

Unit I. Probability basics (15 Lectures)
Modelling Random Experiments: Introduction to probability, probability space, events. Classical probability spaces: uniform probability measure, fields, finite fields, finitely additive probability, Inclusion-exclusion principle, σ-fields, σ-fields generated by a family of sets, σ-field of Borel sets, Limit superior and limit inferior for a sequence of events.

Unit II. Probability measure (15 Lectures)
Probability measure, Continuity of probabilities, First Borel-Cantelli lemma, Discussion of Lebesgue measure on σ-field of Borel subsets of assuming its existence, Discussion of Lebesgue integral for non-negative Borel functions assuming its construction. Discrete and absolutely continuous probability measures, conditional probability, total probability formula, Bayes formula, Independent events.

Unit III. Random variables (15 Lectures)
Random variables, simple random variables, discrete and absolutely continuous random variables, distribution of a random variable, distribution function of a random variable, Bernoulli, Binomial, Poisson and Normal distributions, Independent random variables, Expectation and variance of random variables both discrete and absolutely continuous.

Unit IV. Limit Theorems (15 Lectures)
Conditional expectations and their properties, characteristic functions, examples, Higher moments examples, Chebyshev inequality, Weak law of large numbers, Convergence of random variables, Kolmogorov strong law of large numbers (statement only), Central limit theorem (statement only).

Recommended Text Books

1. M. Capinski, Tomasz Zastawniak: Probability Through Problems.
2. J. F. Rosenthal: A First Look at Rigorous Probability Theory, World Scientic.
3. Kai Lai Chung, Farid AitSahlia: Elementary Probability Theory, Springer Verlag.
4. Ross, Sheldon M. A first course in probability(8th Ed), Pearson.

1

BASICS OF PROBABILITY

Unit Structure

1.0 Objectives

After going through this chapter you will learn

- What is random experiment? How it forms the basis for the "probability"
- Notion of sample space and its types
- Various types of events
- Operations of the events and the laws these operations obey.
- Mathematical and Statistical definition of probability and their limitations.

1.1 Introduction

In basic sciences we usually come across deterministic experiments whose results are not uncertain. Theory of probability is based on Statistical or random experiments. These experiments have peculiar features

Definition 1.1. *Random Experiment*: *A non deterministic experiment is called as a random experiment if*

1. *It is not known in advance, what will be the result of a performance of trial of such experiment.*
2. *It is possible to list out all possible of this experiment outcomes prior to conduct it.*
3. *Under identical conditions, it w possible to repeat such experiment as many times as one wishes.*

Definition 1.2. ***Sample space*****:** *Collection of all possible outcomes of a random experiment is known as sample space*

Sample space is denoted by Ω. And an element of Ω by ω

1. Each ω represents a single outcome of the experiment.
2. Number of elements of Ω are called sample points, and total number of sample points are denoted by $\#(\Omega)$
3. Number of elements of Ω may be finite, or it may have one one correspondence with the $\mathbb{N}$, or with $\mathbb{R}$.
4. Depending on its nature Ω is called as finite, countable or uncountable.

Example 1.1.

1. *A roulette wheel with pointer fixed at the center is spinned. When it comes to rest, the angle made by pointer with positive direction is noted. This experiment is random. Since we do not know before spinning where the pointer would rest. But it may make angle any where between* $(O, 3600)$ *thus here sample space* $\Omega = \left(0, 360^0\right)$ *It is subset of* $\mathbb{R}$ *. It w uncountable*

2. *A coin is tossed until it turns up head. Number of tosses before we get head are noted. This is a random experiment. The corresponding sample space* $\Omega = \left(0,1,2,3\ldots\right)$ *has one one correspondence with the* $\mathbb{N}$ *, so it is countable*

3. *A gambler enters a casino with initial capital "* C *" .If his policy is to continuing to bet for a unit stake until, either his fortune reaches to "* C *" or his funds are exhausted . Gambler' s fortune after any game is though uncertain we can list it out. The sample space of this random experiment is* $\Omega = \left(0,1,2,3\ldots,C\right)$ *. Here sample space is finite.*

1.2 Some Terminologies and notations

Event: Any subset of Ω is termed as an event. Thus corresponding to random experiment a phenomenon may or may not be observed as a result of a random experiment is called as an event.

Note: Event is made up of one or many outcomes Outcomes which entails happening of the event is said to be favorable to the event. An event is generally denoted by alphabets. Number of sample points in an event "A" is denoted by $\#(A)$.

Algebra of events: Since events are sets algebraic operations on sets work for the events.

- Union of two events: A and B are two events of Ω ,then their union is an event representing occurrence of (at least one of them) A or B or both and denoted by $A \cup B$

  Thus, $A \cup B = \{\omega : \omega \in \text{A or } \omega \in \text{B or } \in \text{A and 13 } both\}$

- Intersection of two events: A and B are two events of Ω ,then their Intersection is an event representing simultaneous occurrence of A and B both and denoted by $A \cap B$

  Thus, $A \cap B = \{\omega : \omega \in \text{A and} \omega \in B\}$

- Complement of an event: Non occurrence of an event is its complementary event. Complement of an event is denoted by $\bar{A}$. It contains ω that are not in A. Thus , $\bar{A} = \{\omega : \omega$ does not belong to A $\}$

- Relative complementarity: Out of the two events occurrence of exactly one event is relative complement of the other. In particular if an event A occurs but B does not, it is relative complement of B relative to A. It is denoted by $A - B$ or $A \cap \bar{B}$ .This event contains all sample points of A that are not in B. Similarly $B - A$ or $B \cap \bar{A}$ represents an event that contains all sample points of 13 that are not in A. Thus , $A - B = \{\omega : \omega \in$ A and ω does not belong to 13 $\}$

- Finite Union and Countable Union: $A_1, A_2, \ldots, A_n$ be the events of the sample space $U_{i=1}^{n} A_i$ is called as finite union of the events.

  If $n \to \infty$ we have $U_{i=1}^{\infty} A_i$ which is called as countable union of the events

- Finite intersection and Countable intersection: $A_1, A_2, \ldots, A_n$ be the events of the sample space $\bigcap_{i=1}^{n} A_i$ is called as finite intersection of the events

  If $n \to \infty$ we have $\bigcap_{i=1}^{\infty} A_i$ which is called as countable intersection of the events

Laws of Operations: Union and intersection are the set operations, they obey following laws.

- Commutative law i) $A \cup B = B \cup A$ and ii) $A \cap B = B \cap A$
- Reflexive law i) $A \cup A = A$ and ii) $A \cap A = A$
- Associative law i) $(A \cup B) \cup C = A \cup (B \cup C)$

  ii) $(A \cap B) \cap C = A \cap (B \cap C)$
- Distributive law i) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

  ii) $A \cap (13 \cup C) = (A \cap B) \cup (A \cap C)$
- De Morgan's Law i) $(\overline{A \cup B}) = (\bar{A} \cap \bar{B})$. ii) $(\overline{A \cap B}) = (\bar{A} \cup \bar{B})$

Impossible event: An event corresponding to an empty set.

Certain event: An event corresponding to Ω.

Mutually Exclusive Event: When occurrence of one event excludes the occurrence of the other for all choices of it then the two events are called as Mutually exclusive events. Alternately, when the two events do not occur simultaneously then the two events are called as Mutually exclusive events. Here $A \cap B = \varphi$.

Exhaustive events: The two events are said to be Exhaustive events if they together form the sample space Alternately when all sample points are included in them they are called Exhaustive events. Here $AUB = \Omega$

Equally likely events: If we have no reason to expect any of the events in preference to the others, we call the events as Equally likely events.

Indicator function: Indicator function of an event denoted by $I_A(\omega)$ and defined as

$$I_A(\omega) = \begin{cases} 1 & \omega \in A \\ 0 & \omega \in \bar{A} \end{cases} \qquad (1.1)$$

Partition of sample space: $A_1, A_2, \ldots, A_n$ be the events of the sample space such that they are Mutually exclusive and Exhaustive then are said to form (finite) partition of the sample space.

So, if $A_1, A_2, \ldots A_n$ are forming partition of a sample space, for every $i \neq j = 1,2,\ldots n$; $\left(A_i \cap A_j\right) = \varphi$ And $\left(\bigcup_{i=1}^{n} A_i\right) = \Omega$

Note: Concepts of Mutually Exclusive Event and Exhaustive events and hence for partition can be generalized for countable events $A_1, A_2, \ \ldots$

Example 1.2. $\Omega = \{e_1, e_2, e_3, e_4, e_5\}$. *If* $A = \{e_1, e_3, e_5\}$, *and* $B = \{e_1, e_2, e_3, e_4\}$. *Answer the following* (*i*) *Are A*, *B mutually exclusive*? (*ii*) *Are A*, *B exhaustive9*. (*iii*) *If* $C = \{e_2, e_4\}$. *find* $A \cup (B \cap C)$ *and* $\left(\overline{A} \cap B\right)$

Solution**:** (*i*) *Since* $(A \cap B) = \{e_1, e_3\}$ *which is non null*, *so A*, *B are not mutually exclusive*. (*ii*) $(A \cup B) = \Omega$, *so* A, B *are exhaustive.*

(*iii*) $A \cup (B \cap C) = \Omega$ *and* $\left(\overline{A} \cap B\right) = C$

1.3 Different Approaches of Probability

Definition 1.3. *Classical or Mathematical definition* (*Leplace*): *If a random experiment is conducted results into N mutually exclusive*, *exhaustive and equally likely outcomes*, *M of which are favorable to the occurrence of the event A* , *then probability of an event A is defined as the ratio* $\frac{M}{N}$, *and denoted by* $P(A)$

$$P(A) = \frac{\#(A)}{\#(\Omega)} = \frac{M}{N}$$

This definition has limitations

1 It is not applicable when outcomes are not equally likely.

2 We may not always come across a random experiment that results into a finite number of outcomes.

3 Even if outcomes are finite, can not be enumerated or the number favorable to the event of interest may not be possible to count.

Definition 1.4. ***Empirical or Statistical definition*** **(*Von Mises*):***If a random experiment is conducted N times, out of which M times it results into outcomes favorable to an event A, then the limiting value of the ratio* $\frac{M}{N}$ *is called probability of A.*

$$P(A) = \lim_{n \to \infty} \frac{M}{N}$$

This definition also has limitations.

1 This definition gives a stabilized value of the relative frequency, and overcomes to some extent the drawbacks of classical approach

2 This definition also has some limitations first is, it may not be possible to repeat the experiment under identical conditions large number of times ,due to budgeted time and cost.

3 In repeatation of the experiment large number of times conditions no more remain identical.

4 Since it is based on concept of limit, drawbacks of limit are there with the definition also. However it it works satisfactorily and is widely used

Example 1.3.

What is the probability that a positive integer selected at random from the set of positive integers not exceeding 100 is divisible by (*i*) *5*, (*ii*)*5 or 3* (*iii*)*5 and 3* ‘.?

***Solution*:** $\Omega = \{1, 2, \ldots, 100\}$ *so*, $\#(\Omega) = 100$

(*i*) *Let A be an event that no. is divisible by 5, so A* , $= \{5, 10 \ldots, 100\}$

so, $\#(A) = 20$

$$P(A) = \frac{\#(A)}{\#(\Omega)} = \frac{20}{100} = 0.2$$

(*ii*) *Let B be an event that no. is divisible by 3, so* $B = \{3, 6, \ldots, 99\}$

so, $\#(B) = 33$

$$P(B) = \frac{\#(B)}{\#(\Omega)} = \frac{33}{100} = 0.33$$

(iii) *Let C be an event that no. is divisible by 5 or 3,* $C = A \cup B$
so, $\#(A \cup 13) = 47$

$$P(C)\frac{\#(C)}{\#(\Omega)} = \frac{47}{100} = 0.47$$

(iv) *Let* D *be an event that no. is divisible by 5 and 3,* $D = A \cap B$
so, $\#(A \cap B) = 6$

$$P(D)\frac{\#(D)}{\#(\Omega)} = \frac{6}{100} = 0.06$$

Example 1.4. *What is the probability that in a random arrangement of alphabets of word* ("*REGULATIONS*"

(i) *All vowels are together.* (ii)*No two vowels are together ?*

Solution**:** *Since there are 11 letters in the word, they can be arranged in 11!* *distinct ways so,* $\#(\Omega) = 11!$

Let A *be an event that the random arrangement has all vowels together. Since the 5 vowels is one group to be kept together and remaining 6 consonants, which is random arrangement of* 7 *entities in all can be done in* 7! *ways. In the group of 5 vowels the random arrangement can be done 5!. ways. so,* $\#(A) = 7!X5!$

$$P(A) = \frac{\#(A)}{\#(\Omega)} = \frac{7!5!}{11!} = .01515$$

(ii) *Let B be an event that the random arrangement no to vowels together. The consonants can be arranged as* $*C^*C^*C^*C^*C^*C^*$ *, where C stands for consonants. 5 vowels can be arranged a t* 7*, * positions in* ${}^7P_5$ *ways and 6 consonants in* 6! *ways, all such random arrangements* $7P_5X6!$ *ways so,*

$$\#(B) = \frac{7!X6!}{2!}$$

$$P(B) = \frac{\#(B)}{\#(\Omega)} = \frac{7!6!}{11!2^1} = .04545$$

Example 1.5. *From a pack of well shuffled 52 cards four cards are selected without replacing the selected card. Jack, queen, king or ace cards are treated as honor card.* $a)$ *What is the probability that are there are* $i)$ *all honor cards* $ii)$ *More* *honor cards 9. b) What will be these probabilities if cards are drawn with replacement?*

***Solution*:**

a) Since there are 52 cards in the pack of cards,4 can be selected without replacement in ${}^{52}C_4$ *distinct ways so,* $\#(\Omega) = {}^{52}C_4$

i) *Let A be an event that the random selection has 4 honor cards. Since there are in all* $4X4$ *honour cards,* $\#(A) = {}^{16}C_4$

$$P(A) = \frac{\#(A)}{\#(\Omega)} = \frac{16C_4}{52C_4} = 0.0067$$

ii) *Let B be an event that the random selection has more , that is 4 or 3 honor cards.*

$$\#(B) = {}^{16}C_4 + {}^{36}C_1^{16}C_3 = 21980$$

$$P(B) = \frac{\#(B)}{\#(\Omega)} = \frac{21980}{270725} = 0.08119$$

b) 4 *cards can be selected with replacement in* 52^4 *ways so,* $\#(\Omega) = 52^4$

i) *Let C be an event that the random selection has 4 honor cards.* $\#(C) = 16^4$

$$P(C) = \frac{\#(C)}{\#(\Omega)} = \frac{16^4}{52^4} = 0.00896$$

ii) *Let D be an event that the random selection has more) that is 4 or 3 honor cards.*

$$\#(D) = 16^4 + 36X16^3 = 212992$$

$$P(D) = \frac{\#(D)}{\#(\Omega)} = \frac{212992}{7311616} = 0.02913$$

Example 1.6. *In a party of 22 people, find the probability that (i) All have different birtday (ii) Two persons have sme bithday (iii) 11 persons have birtday in same month.*

***Solution*:** *We assume that none of them have birthday on 29th February.*

(*i*) *Since all 22 people can have any of 365 days as their bithday in* 365^{22} *ways. Thus* $\#(\Omega)=365^{22}$

A be the event that all have different birhday, $\#(A)={}^{365}P_{22}$

Hence $P(A)=0.5243$

(*ii*) *B be the event that two have same birthday and remaining 20 have different birhday, Any 2 out of 22 can be chosen to have same birthday in* $22C_2$ *ways, and remaining 21 different birthdays can be chosen from 365 days in* $365P_{22}$ *ways.*

$$\#(B)={}^{365}P_{21}\times{}^{22}C_2$$

Hence $P(B)=0.352$

(*iii*) *C be the event that 11 have birhday in same month and remaining 11 in different months.*

Now $\#(\Omega)=12^{22}$

And $\#(C)={}^{12}P_{11}\times{}^{22}C_{11}$

Hence $P(C)=0.000011$

The notion of probability is given modern approach which is based on measure theory. For this it is necessary to introduce class of sets of Ω

In next chapter we will discuss various classes of sets.

1.4 Chapter End Exercises

1. Cards are to be prepared bearing a four digit number formed by choosing digits among 1, 4, 5, 6 and 8. Γind the probability that a randomly chosen cards among them bear (i) An even number (ii) A number divisible by 4 (iii)A number has all four digits same.

2. A sample of 50 people surveyed for their blood group. If 22 people have 'A' blood group, 5 have 'B' blood group, 21 have 'O' blood group and 2 have 'AB' blood group. Find the probability that a randomly chosen person has (i) Either 'A' or '13' blood group (ii)Neither 'A' nor 'B' blood group.

3. A roulette wheel has 40 spaces numbered from 1 to 40. Find the probability of getting (i) number greater than 25(ii) An odd number (iii) A prime number.

4. A, B, C forms a partition. If the event A is twice as likely as B ,and event C is thrice as likely as A . Find their respective probabilities.

5. What is the probability that in a random arrangement of alphabets of word “CHILDREN”

   (i) All vowels are together.

   (ii) No two vowels are together?

6. A committee of 5 is to be formed from among a coordinator, chairperson, five research guides and three research students. What is the probability that committee (i) Do not have coordinator and chairperson. (ii) All research guides (iii) None of the students

7. 9 people are randomly seated at a round table. What is the probability that a particular couple sit next to each other?

8. In a box there are 10 bulbs out of which 4 are not working. An electrician selects 3 bulbs from that box at random what is the probability that at least one of the bulb is working?

9. $\Omega = \{1,2,\ldots,50\}$ A denote number divisible by 5, B denotes number up to 30 ,C is number greater than 25 and D is number less than or equal to 4. Answer the following

   (i) Which events are exhaustive?

   (ii) Which events are mutually exclusive?

   (iii) Give a pair of events which is mutually exclusive but not exhaustive.

   (iv) Give a pair of events which is not mutually exclusive but exhaustive.

   (v) Give a pair of events which is neither mutually exclusive nor exhaustive.

10. A pair of fair dice is thrown what is the probability that the sum of the numbers on faces of the dice is (i) 6, 7 or 8. (ii) Divisible by 5.(iii)a prime number?

11. What is the probability that in a group of 25 people (i) all have different birtdays (ii)11 have birhday in different month and 14 in the same month?

12. Five letters are to be kept in five self addressed envelopes. What is the probability that (i) All goes to correct envelope(ii)none of them goes to correct envelope?

13. The coefficients a,b,c of the quadratic equation $ax^2 + bx + c = 0$, are obtained by throwing a die thrice. Find the probability that equation has real roots.

14. What is the probability that there are 53 Thursdays and 53 Fridays in a leap year?

15. A sequence of 10 bits is randomly generated. What is the probability that (i) atleast one of these bits is 0? (ii) a sequence has equal number of 0 and 1.

16. The odds against an event A are 3: 5, the odds in favor of an event 13 are 7: 5, What are the probabilities of the events?

17. In a group of 12 persons what is the probability that (i) each of them have different birthday (ii) each of them have birthday in different calendar month?

18. A, B, C are mutually exclusive. $\mathrm{P}(\mathrm{A}) = \frac{1-3\mathrm{x}}{2}, \mathrm{P}(\mathrm{B}) = \frac{1+4\mathrm{x}}{3}$ and $\mathrm{P}(\mathrm{C}) = \frac{1+\mathrm{x}}{6}$. (i)Show that the range for x is, $\frac{-1}{4} < x < \frac{1}{3}$ (ii) are they exhaustive?

19. Express $A - (B \cap C)$ as union of three events.

2

FIELDS AND SIGMA FIELDS

Unit Structure

2.0 Objectives

After going through this chapter you will learn

- A class of sets and variouus closure properties that it may follow.
- Concept of field and its properties.
- Sigma field and its properties.
- Borel Sigma field, minimal Sigma field.
- Limit superior and limit inferior of sequence of events.

2.1 Class of Sets

Before introducing modern approach of probability, we need to define some terms from measure theory. Subsequent sections are also explaining their role in probability theory.

A collection of subsets of Ω is termed as Class of subsets of Ω .It plays an important role in measure theory. They have some closure properties with respect to different set operations.

A be the class of subsets of Ω.

Complement: **A** is said to be closed under the complement, if for any set $\mathrm{A} \in A$, $\bar{A}$ is $\in A$

Union: A is said to be closed under the union if for any sets $\mathrm{A}, B \in A, A \cup B$ is $\in A$ Intersection : A is said to be closed under the intersection if for any sets A , $B \in A$, $A \cup B$ is $\in A$

Finite Union and Countable Union : A is said to be closed under the finite union if for any sets A_1, A_2 , ..., $A_n . \in A$, $\bigcup_{i=1}^{n} A_i$ is $\in A$ .Further if $n \to \infty$ and if we have $\mathrm{U}_{i=1}^{\infty} A_i$ is $\in A, A$ is said to be closed under countable unions

Finite intersection and Countable intersection: A is said to be closed under the finite intersection if for any sets A_1, A_2 , ..., $A_n . \in A,$

$\bigcap_{i=1}^{n} A_i$ is $\in A$ .Further if $n \to \infty$ and if we have $\bigcap_{i=1}^{\infty} A_i$ is $\in A, A$ is said to be closed under countable intersection. Note: Closure property for countable operation implies closed for finite operation

2.2 Field

Definition 2.1. ***Field*** **:**

A class $\mathcal{F}$ of subsets of a non - empty set Ω is called a field on Ω if

1. $\Omega \in \mathcal{F}$.

2. It is closed under complement.

3. It is closed under finite Union

Notationally

A class $\mathcal{F}$ of subsets of a non - empty set Ω is called a field on Ω if

1. $\Omega \in \mathcal{F}$

2. for any set $\mathrm{A} \in \mathcal{F}$, $\bar{A} \in \mathcal{F}$.

3. for any sets A_1, A_2 , ... , $A_n . \in \mathcal{F}, \bigcup_{i=1}^{n} A_i \in \mathcal{F}$

Following points we should keep, in mind regarding field.

- Closure for complement and finite Union implies closure for intersection.So,field is closed under finite intersections.
- $\{\varphi, \Omega\}$ is a field.

- Power set $P(\Omega)$,which is set of all subsets of Ω is a field.
- For any $\mathrm{A} \in \Omega, \{\varphi, \Omega, A, \overline{A}\}$ is smallest field containing A.
- For any sets $\mathrm{A}, \mathrm{B} \in \mathcal{F}$ $\mathrm{A} - \mathrm{B} \in \mathcal{F}$,hence $\mathrm{A}\ \Delta B \in \mathcal{F}$.
- $\mathcal{F}_1$ and $\mathcal{F}_2$ are two fields on Ω, then $\mathcal{F}_1 \mathrm{n} \mathcal{F}_2$ is a field.
- Field is also called as an Algebra.

Example 2.1. $\Omega = \{1,2,3\}$, $\mathcal{F}_1 = \{\varphi, \Omega, \{1\}, \{2,3\}\}$ *and* $\mathcal{F}_2 = \{\varphi, \Omega, \{2\}, \{2,3\}\}$ *are two fields on* Ω. *Is union of these two fields is a field*[i)]

Solution: $\mathcal{F}_1 \cup \mathcal{F}_2 = \{\varphi, \Omega, \{1\}, \{2\}, \{2, 3\}\}$

let $A = \{1,2\} = \{1\} \bigcup \{2\} \notin \mathcal{F}_1 \mathrm{U} \mathcal{F}_2$

$\therefore f_1 \bigcup \mathcal{F}_2$ *is not a field.*

Example 2.2. $\mathcal{F} = \{ A \in \Omega$ *such that* A *is finite*$\}$. *Is* $\mathcal{F}$ *a field*?

Solution: *No, if* Ω *is infinite then,* Ω *does not belong to* $\mathcal{F}$ *and hence* $\mathcal{F}$ *cannot be a field.*

Example 2.3. *Complete the following class to obtain a field. Given* $\Omega = [0,1]$ *and*

$$\mathcal{F} = \left\{\varphi, 1, \left[0, \frac{1}{2}\right)\right\}$$

Solution: Add $\left\{[0,1], \left[\frac{1}{2}, 1\right), \left[0, \frac{1}{2}\right) \mathrm{U} \{1\}, [0,1), \left[\frac{1}{2}, 1\right]\right\}$ in $\mathcal{F}$ to make it a field

2.3 $\alpha -$ field and Borel $\alpha -$ field

Definition 2.2. $\sigma -$ *field*: *A class* C *of subsets of a non - empty set* Ω *is called a* σ *- field on* Ω *if*

1. $\Omega \in C$.
2. *It is closed under complement.*
3. *It is closed under countable Unions.*

Notationally

A class C of subsets of a non - empty set Ω is called a σ - field on Ω if

1. $\Omega \in C$
2. for any set $\mathrm{A} \in C$, $\overline{A} \in C$.
3. for any sets A_1, A_2 , ..., . $\in C$, then $\bigcup_{i=1}^{\infty} A_i \in C$

- Field which is closed under countable unions is a α - field.
- Like fields the intersection of arbitrary σ - fields is also σ - field but their union is not a σ - field
- Power set $P(\Omega)$,which is collection of all subsets of Ω is ao . - field.
- Given a class of sets consisting all countable and complements of countable sets is a σ - field

Example 2.4. *A class C of subsets A of Ω such that either A or its complement is finite. Is C is (I) a field d^p (II) a $\sigma -$ field 9.*

***Solution*:**

(*I*) $C = \{ A \subseteq \Omega \mid A$ *is finite or* - *is finite*$\}$

Note that $(i) C$ *is closed under complementation*, *since either of* A *or* - *is finite* (ii) If $A_, B \in C$ *both finite then* $A \bigcup B$ *is finite*

If A *is finite* B *is infinite* $A \bigcup B$ *is infinite*. But $\overline{A \bigcup B} = \overline{A} \bigcap \overline{B}$ *.Since* - *is finite* - *is infinite*

$\overline{A \bigcup \mathrm{B}}$ *finite*, *hence* A U B. *Similarly, we can check the case when both* A *and* B *are infinite*.

Thus, For any $A, B \in C$, $A \bigcup B$ *is also* $\in C$. *So* C *is a field*

(*II*) *But if* A_i *are finite* $\bigcup_{i=1}^{\infty} A_i$ *does not belong to* C.

C *is not closed for countable unions*, *hence can*not *be* $0^{\cdot}$ - *field*.

Definition 2.3. *Minimal* $\sigma -$ *field* : *A class* C *of subsets of* Ω *is called a minimal* σ - *field on* Ω , *if it is the smallest* $0^{-} -$ *field containing* C

- Minimal α - field can be generated by taking intersection of all the α - fields containing C
- If A is family of subsets of Ω , and $C_A = \text{n}\{C \mid A \subset C\}$,which is intersection of all 0 - fields containing A then C_A is a minimal σ - field.
- If A itself is a σ - field, then $C_A = A$

Hence onwards we term the pair (Ω, C) as a sample space

In theory of probability $\Omega =$ IR has specific features and sample space (IR, B) plays vital role.

Definition 2.4. ***Borel*** $\sigma -$ ***field***: *Let* C *is the class of all open intervals* $(-\infty, x)$ *where* $x \in \mathbb{R}$, *a minimal* $\sigma -$ *field generated by* C *is called as Borel* σ - *field, and denoted by* B.

Borel σ - field has following features.

1. It is clear that $[x, \infty)$ is complement of $(-\infty, x)$ but it does not belong to C .Thus C is not closed under complement C is also not closed under countable intersections, as

   $\bigcap_{n=1}^{\infty}\left(x - \frac{1}{n}, \infty\right) = [x, \infty)$. But B is closed under complements as well as countable unions or intersections.

   Hence $\mathcal{B}$ contains all intervals of the type $[x, \infty)$

2. $(-\infty, x] = \bigcap_{n=1}^{\infty}\left(-\infty, x + \frac{1}{n}\right)$ .So B contains all intervals of the type $(-\infty, x]$

3. (x, ∞) is complement of $(-\infty, x]$. Thus B contains all intervals of the type (x, ∞) . 4. $(a, b) = (-\infty, b) \bigcap (a, \infty)$, where $a < b$. So $\mathcal{B}$ contains all intervals of the type (a, b) . And contains even intervals of the type $[a, b)$, $(a, b]$ for all $\text{a}, \text{b} \in \mathbb{R}$.

Note that sets of B are called as borel sets.

2.4 Limit of sequence of events

In this chapter the concept of limit of a sequence of events is introduced.

Definition 2.5. *Limit Superior*: $\{A_n\}$ *be the sequence of events of space* (Ω, C). *Limit superior of* A_n *is an event which contains all points of* Ω *that belong to* A_n *for infinitely many* n *and it is denoted by* $\lim\ \sup A_n or - lim A_n$, *termed as limit superior of* A_n

- $\omega \in \lim\ \sup A_n$ iff for each $\mathrm{n} \geq 1$ there exists an integer $\mathrm{m} \geq \mathrm{n}$ such that $\omega \in A_m$ for all $\mathrm{m} \geq \mathrm{n}$
- Thus $\lim\ \sup = \mathrm{n}_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m$
- It can be clearly seen that $\lim\ \sup A_n \in C$
- $\overline{\lim} A_n = \{A_n i.0\}$, where i.o $=$ infinitely often.

Definition 2.6. *Limit Inferior*: $\{A_n\}$ *be the sequence of events of space* (Ω, C) . *Limit inferior of* A_n *is an event which contains all points of* Ω *that belongs to* A_n *but for finite values of* n *. and it is denoted by* $\lim\ \inf A_n or \overline{lim} A_n$, *termed as limit inferior of* A_n

- $\omega \in \lim\ \inf A_n$ iff there exists some $\mathrm{n} \geq 1$ such that $\omega \in A_m$ for all $\mathrm{m} \geq \mathrm{n}$
- Thus $\lim\ \inf A_n = \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} A_m$
- It can be clearly seen that $\lim \mathrm{in} \notin A_n \in C$
- $\lim\ \inf A_n subseteq \lim\ \sup A_n$
- If $\lim A_n$ exists, $\lim A_n = \lim\ \inf A_n = \lim\ \sup A_n$.

Definition 2.7. $\{A_n\}$ *be the sequence of events of space* (Ω, C) *such that* $A_1 \subset A_2 \ldots$, *then* $\{A_n\}$ *is called as expanding or increasing sequence and* $\lim A_n = \bigcup_{n=1}^{\infty} A_n$

Definition 2.2. $\{A_n\}$ *be the sequence of events of space* (Ω, C) *such that* $A_1 \supset A_2 \ldots$, *then* $\{A_n\}$ *is called as contracting or decreasing sequence and* $\lim A_n = \bigcap_{n=1}^{\infty} A_n$

Remark 2.1. $\{A_n\}$ *be the sequence of events of space* (Ω, C) *then* (i) $C_n = \bigcup_{m=n}^{\infty} A_m$ *is decreasing sequence.* $C_n \downarrow C$, *where* $C = \bigcap_{n=1}^{\infty} c_n = \lim\ \sup A_n$ (ii) $B_n = \bigcap_{m=}^{\infty} A_m$ *is increasing sequence.* $B_n \uparrow B$, *where* $B = \bigcup_{n=1}^{\infty} B_n = \lim\ \inf A_n$

Remark 2.2. $\{A_n\}$ *be the sequence of events of space* (Ω, C) *then* $\bigcup_{m=n}^{\infty} A_m$ *is also called as* $Sup_{m \geq n} A_m$, *and* $\bigcap_{m=n}^{\infty} A_m$ *is also called as* $Inf_{m \geq n} A_m$.

2.5 Chapter End Exercises

1. $\mathcal{F}$ is a field .If $A, B \in \mathcal{F}$ then show that $A - B$ and $A \triangle B$ are also events off

2. $\Omega = \{1, 2, 3, 4\}$.

   Which of the following classes is a field on Ω?

   (i) $\mathcal{F}_1 = \{\varphi, \{1,4\}, \{2,3\}\}$

   (ii) $\mathcal{F}_2 = \{\varphi, \Omega, \{2\}, \{3\}, \{2,3\}, \{1,4\}, \{1,2,4\}, \{1,3,4\}\}$.

3. Complete the following class to obtain a field. Given $\Omega = (0,1)$ and (i) $\mathcal{F} = \left\{\varphi, (0,1), \left(0, \frac{1}{2}\right), \left(\frac{1}{2}, 1\right)\right\}$ (ii) $\mathcal{F} = \{\varphi, (0,1), \left(0, \frac{1}{2}\right), \left[\frac{1}{2}\right) 1) \left(0, \frac{2}{3}\right], \left(\frac{2}{3}, 1\right)\}$

4. A class C of subsets A of Ω such that either A or its complement is countable. Is C is a σ - field?

5. $\mathrm{A}, B, \mathrm{C}$ forms partition of Ω , obtain a smallest field containing $A, 13, C$

6. Show that following are the Borel sets.

   (i) $[a,b]$(ii)$\{a\}$ (iii) Any finite set (iv)Any countable set (v) A set of rational numbers(v) A set of natural numbers

7. C is $\sigma -$ field on $\Omega = [0,1]$ such that $\left[\frac{1}{n,}, \frac{1}{n+1}\right] \in C$ for n = 1,2,....Show that following are the events of C. (i) $(\frac{1}{n}, 1]$ (ii) $(0, \frac{1}{n}]$

8. $A_n = \begin{cases} A & n = 1,3,5, \\ B & n = 2,4,6\ldots \end{cases}$

   Find $\lim\ \inf A_n$, $\lim\ \sup A_{71}$ and show that $\lim A_n$ does not exists.

9. Prove that (i) $\overline{\lim}\left(A_n \cup B_n\right) = \overline{lim}A_n \cup \overline{lim}B_n$, (ii) $\underline{lim}\bar{A}_n = \overline{\overline{\lim}\, A_n}$, (iii) $\overline{\lim}\left(A_n \cap B_n\right) \subset$

$\overline{lim}A_n \cap \overline{lim}B_n$

10. Are the above results true for $\lim$ inf?

11. If $A_n \to A$ then $- n \to$

❖❖❖

3

PROBABILITY MEASURE AND LEBESGUE MEASURE

Unit Structure

3.0 Objectives

After going through this chapter you will learn

- A function defined on sample space called as probability measure.
- Types of probability measure ,discrete and continuous.
- Lebesgue measure and Lebesgue integral.
- Properties of probability function
- Probability of limit of sequence of events.

3.1 Probability Measure

The modern approach of probability is based on measure theory. Following definition is due to Kolmogorov (1933)

Definition 3.1. *Axiomatic definition of probability*: *C be the Cf field associated with the sample space Ω. A function $P(.)$ defined on C to $[0,1]$ is called as probability measure or simply probability if it satisfies following axioms*.

1. $P(A) \geq 0$ *for all* $A \in C$.
2. $P(\Omega) = 1$
3. A_1, A_2 , ... *is sequence of mutually exclusive events of* C *then*

   $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \Sigma_{i=1}^{\infty} P(A_i)$ (3.1)

- Axioms are respectively called as non - negativity, normality and countable additivity.
- In this definition probabilities have been already assigned to the events, by some methods or by past information.
- The triplet (Ω, C, P) is called as a probability space
- Depending on Ω different types of probability space are decided.
- If Ω is finite or countable(at most countable)probability space is discrete.
- If Ω has one one correspondence with IR probability space is continuous.

Properties of the Probability function.

Complement

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

Proof:

$$\Omega = A \cup \overline{A}$$

A and $\overline{\text{A}}$ are mutually exclusive events So

$$P(\Omega) = P(A) + P(\overline{A})$$

…by countable additivity axiom

L. H. $\text{S} = 1$ ….. by normality axiom

R.H. $\text{S} = P(A) + P(\overline{A})$ Hence

$$P(A) + P(\overline{A}) = 1$$

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

Monotone A and B are $\in C$ such that$A$ $A \subset B$ then $P(A) \leq P(B)$

Proof:

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

Since A and $(B \cap \overline{A})$ are mutually exclusive, so by countable additivity axiom

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

So, $\text{P}(\text{A}) \leq \text{P}(\text{B})$ as $P(B \cap \overline{A}) \geq 0$

Substantivity A and B are $\in C$ such that $A \subset B$. Then $P(B-A)=P(B)-P(A)$

Proof: From the above proof

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

Thus

$$P(B-A)=P(B\cap \bar{A})=P(B)-F(A)$$

Similarly we can say that

A and B are $\in C$ such that $B\subset A$ Then $P(A-B)=P(A)-P(B)$

Continuity $\lim_{n\to\infty} A_n = A$,then $\lim_{n\to\infty} P(A_n)=P(A)$

Theorem 3.1. $(i)\{A_j\}$ *is expanding or increasing sequence of events of space* (Ω, C) *then*

$$\lim_{n\to\infty} P(A_n)=P\left(\bigcup_{n=1}^{\infty} A_n\right) \qquad (3.2)$$

(ii) $\{A_j\}$ *is the contracting or decreasing sequence events of space* (Ω, C) *then*

$$\lim_{n\to\infty} P(A_n)=P\left(\bigcap_{n=1}^{\infty} A_n\right) \qquad (3.3)$$

Proof: $(\mathrm{i})\{A_j\}$ be the sequence of increasing events, so $A_1\subset A_2\ldots$. Let $B_j = A_j \mathrm{n} \overline{A_{j+1}} B_j$ are mutually exclusive.

So $A_n = \bigcup_{j=1}^{n} B_j$

$$\bigcup_{j=1}^{\infty} A_j = \bigcup_{j=1}^{\infty} B_j \qquad (3.4)$$

By (9.4)

$$P\left(\bigcup_{j=1}^{\infty} A_j\right)=P\left(\bigcup_{j=1}^{\infty} B_j\right)$$

$= \sum_{j=1}^{\infty} P(B_j)$ By countable additivity

$= \lim_{n\to\infty}\sum_{j=1}^{n} P(B_j)$ By definition of sum of series

$= \lim_{n\to\infty} P\left(\bigcup_{j=1}^{n} B_j\right)$ By finite additivity

$= \lim_{n\to\infty} P(A_n)$ By definition of A_n

$$\lim_{n\to\infty} P(A_n)=P\left(\bigcup_{n=1}^{\infty} A_n\right)$$

$(\mathrm{ii})\{A_j\}$ be the sequence of decreasing events, so $A_1 \supset A_2\ldots$. hence $\{-j\}$ be the sequence of increasing events, so $-1\subset\ -2\cdots\cdot$ Applying result in (i) to $\{-j\}$

$$P\left(\bigcup_{j=1}^{\infty} -j\right)=\lim_{n\to\infty} P(-n) \qquad (3.5)$$

$= 1 - P\left(\bigcap_{j=1}^{\infty} A_j\right)$ L.H.S by De Morgan's law

$= 1 - \lim_{n\to\infty}\left[1 - P\left(A_n\right)\right]$ R.H.S By complementation

$= 1 - \lim_{n\to\infty} P\left(A_n\right)$

$$\lim_{n\to\infty} P\left(A_n\right) = P\left(\mathrm{n}_{n=1}^{\infty} A_n\right)$$

Theorem 3.2. *The continuity property of probability*.

$$\lim_{n\to\infty} A_n = A \text{, } then \ \lim_{n\to\infty} P\left(A_n\right) = P\left(\lim_{n\to\infty} A_n\right) = P\left(A\right) \tag{3.6}$$

Proof: $\lim\ \inf A_n = \bigcup_{m=1}^{\infty} \mathrm{n}_{m=n}^{\infty} A_m = \bigcup_{n=1}^{\infty} B_n$

were

$$B_n = \bigcap_{m=n}^{\infty} A_m$$

These B_ns are increasing events $\uparrow \bigcup_{n=1}^{\infty} B_n = B$ say

using (3.2)

$$\lim_{n\to\infty} P\left(B_n\right) = P\left(\bigcup_{n=1}^{\infty} B_n\right) = P\left(B\right)$$

$\lim\ \sup = \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m = P\left(\bigcap_{n=1}^{\infty} C_n\right)$

where

$$C_n = \bigcup_{m=n}^{\infty} A_m$$

These C_ns are decreasing events $\downarrow \bigcap_{n=1}^{\infty} c_n = Csay$

using (3.3)

$$\lim_{n\to\infty} P\left(C_n\right) = P\left(\bigcap_{n=1}^{\infty} C_n\right) = P\left(C\right)$$

Now consider,

$$\bigcap_{m=n}^{\infty} A_m \subset A_n \subset \mathrm{U}_{m=n}^{\infty} A_m \tag{3.7}$$

$$B_n \subset A_n \subset C_n$$

By monotone property of P

$$P\left(B_n\right) \le P\left(A_n\right) \le P\left(C_n\right)$$

taking limits

$$\lim_{n\to\infty} P\left(B_n\right) \le \lim_{n\to\infty} P\left(A_n\right) \le \lim_{n\to\infty} P\left(C_n\right)$$

So,

$$P\left(B\right) \le \lim_{n\to\infty} P\left(A_{rt}\right) \le P\left(C\right)$$

But, $\lim A_n = A$

$$A = \lim A_{\gamma t} = \lim\ \inf A_n = B = \lim\ \sup A_n = C$$

implies

$$P(B) = P(A) = P(C)$$

$$\lim_{n\to\infty} P(A_n) = P(A) = P\left(\lim_{n\to\infty} A_n\right) \tag{3.8}$$

Example 3.1. *Which of the following are Probability functions*?

(i) $\Omega = \{1,2,3\ldots\}$, , C *is* $\sigma -$ *field on* Ω. *A function* P

defined on space (Ω, C) *as*

$$P(i) = \frac{1}{2^i}$$

for $i \in \Omega$

Solution: *a*)

$$P(\Omega) = \sum_{i=1}^{\infty} \frac{1}{2^i} = 1$$

b)

$$P(A) \geq 0 \text{ for all } A \in C$$

c)*Let us define mutually exclusive events*, $A_i = i$ *we can verify countable additivity*.

$$P\left(\cup_{i=1}^{\infty} A_i\right) = \Sigma_{i=1}^{\infty} P(A_i) \tag{3.9}$$

By a) ,b)and c) Pis Probability function.

$(ii)\Omega = (0,\infty)$, *Borel* $\sigma -$ *field* B *defined on* Ω.*A function* F

defined on space (Ω, B) *as, for any* $I \in \mathcal{B}$

$$P(I) = \int_I e^{-x} dx \tag{3.10}$$

Solution: *a*) $P(\Omega) = \int_0^{\infty} e^{-x} d_x = 1$

b) $P(A) \geq 0$ *for all* $A \in B$

c)$A_i = (i, i+1)$ *we can verify countable additivity*

$$. \; P\left(\cup_{i=1}^{\infty} A_i\right) = \int_{\cup_{i=1}^{\infty}} e^{-x} d_x \tag{3.11}$$

$A_i s$ *are mutually exclusive*. From *the properties of integrals*,

$$= \sum_{i=1}^{\infty} f^{i+1} e^{-x} d_x = \sum_{i=1}^{\infty} P(A_i) \tag{3.12}$$

By a) ,b), c) P is Probability function.

(iii) $\Omega = (-\infty, \infty)$, $\sigma -$ *field* C *defined on* Ω *A function* P *defined on space* (Ω, C) *as, for any* $I \in C$

$$P(I) = \begin{cases} 0 & I \subset (-\infty, 1) \\ \frac{1}{2} & I \subset [1\infty) \end{cases}$$

Solution: a)

$$P(\Omega) = \int_1^\infty \frac{1}{2} \neq 1$$

P *is not a Probability function.*

Theorem 3.3. *Borel Catelli lemma*

If $\sum P(A_i) < \infty$ *then* $P\left(\overline{\lim} A_n\right) = P(A_n i.0) = 0$

Proof:

$$P\left(\overline{\lim} A_n\right) = P\left(\bigcap_{n=1}^\infty \bigcup_{m=n}^\infty A_m\right)$$

$$\leq P\left(\bigcup_{m=n}^\infty A_m\right)$$

$$\leq \sum_{m=n}^\infty P(A_m) \text{as events need not be mutually exclusive}$$

If $\sum P(A_i) < \infty$, then $\sum P(A_m)$ tends to zero as $n \to \infty$ Hence the proof.

Remark 3.1. *Other half of the above result is stated as follows. But it needs independent events.*

For A_i *are independent events of the sample space, If* $\sum P(A_i) = \infty$ *then* $P\left(-\lim A_n\right) = P(A_n i.0) = 1$

3.2 Lebesgue Measure and integral

Definition 3.2. ***Lebesgue Measure*** *A function* μ *defined on space* $(\mathbb{R}, B)$ *is called as Lebesgue Measure if it satisfies following*

1. $\mu(a, b] = b - a$
2. $\mu(\varphi) = 0$
3. $E_i s$ *are mutually exclusive intervals of* B *then,*

$$\mu\left(\bigcup_{i=1}^\infty E_? \cdot\right) = \mathrm{X}_{i=1}^\infty \mu(E_t)$$

Example 3.2. *Find Lebesgue Measure for following sets.*

(i) $\left[\frac{1}{2},\frac{1}{6}\right)$ (ii) $\left[\frac{1}{3},\frac{1}{2}\right)$ (iii) $\left[\frac{2}{3},\frac{7}{9}\right]$ (iv) $\frac{1}{n},: n=1,2,.\}$ (v) $\{x: |x-n| < \frac{1}{2^n}\ forn \in \mathbb{N}\}$

Solution:

(i) $\mu\left(\frac{1}{9},\frac{1}{6}\right] = \frac{1}{6}-\frac{1}{9} = \frac{1}{18}$

Similarly (ii) $\frac{1}{9}$ *(iii)* $\frac{1}{9}$ *(iv)* $\mu\{\frac{1}{n}: n=1,2,...\} = \sum_{i=1}^{\infty} \mu\left\{\frac{1}{n}\right\} = 0$

(v) $|x-n| < \frac{1}{2^n}$ *implies* $n-\frac{1}{2^n} < x < n+\frac{1}{2^n}$

x lies in mutually exclusive intervals of length $\frac{1}{2^{n-1}}$

$$\mu\left\{x: |x-n| < \frac{1}{2^n}\right\} \text{ for } n \in \mathbb{N} = \sum_{i=0}^{\infty} \mu\left(n-\frac{1}{2^n}, n+\frac{1}{2^n}\right) = 2$$

Remark 3.2.

- *If* $\Omega = [0,1]$ *then* $\mu = P$ *is Probability measure.*
- μ *is actually an extended measure. It is* σ *- finite measure,*
- *Since* $\mu(x) = \mu(\lim(x-\frac{1}{n}, x])$

$$\mu(x) = \lim \mu\left(x-\frac{1}{n}, x\right] = \lim \frac{1}{n}$$

by continuity

Thus $\mu(x) = 0$

- *From above* $\mu(a,b] = \mu(a,b) = b-a$
- *The sets whose* μ *measure is zero is called as* μ *- null set.*

Definition 3.3. $f(.)$ *is a function defined on* $\mathbb{R}$ is *a called as Borel function if inverse image is a Borel set.*

Definition 3.4. ***Lebesgue integral***: *Lebesgue integral is a mapping on non-negative burel function* f *which satisfy following,*

1 $\int f d\mu \in (0,\infty)$

2 $\int I_A d\mu = \mu(A)$ *for any* $A \in \mathbb{R}$

3 $\int (f+g)d\mu = \int fd\mu + \int gd\mu$ *and* $\int cfd\mu = c\int fd\mu$. *Where* $c \geq 0$

4 $\lim\int f_n d\mu = \int fd\mu$ *if* $\lim f_n(x) = f(x)$ *for any* $x \in \mathbb{R}$

- For any nonnegative piecewise continuous function f

$$J_{[a,b]} fd\mu = \int_{\alpha}^{b} f(x)dx = F(b) - F(a), \qquad (3.13)$$

  Where P is antiderivative of f.

- P is nondecreasing function with $F(\infty) - F(-\infty)$ is finite.
- If we divide F by $F(\infty) - F(-\infty)$ we get probability measure.
- We will revisit this function in next chapters.

3.3 Discrete and absolutely continuous probability measures

Definition 3.5. *Density function*: *A non-negative Borel function* $f : \mathbb{R} \to [0,\infty)$ *is called as a density if*

$$\int fd\mu = 1 \qquad (3.14)$$

Theorem 3.4. *If* f *is a density then* P *satisfying*

$$P(A) = \int_A fd\mu \qquad (3.15)$$

is a probability measure on Borel subsets A *of* $\mathbb{R}$

proof: Since f is a density on $\mathbb{R}$

$$P(\mathbb{R}) = \int_{\mathbb{R}} fd\mu = 1 \qquad (3.16)$$

Now consider $A_1, A_2\ldots$ be mutually exclusive Borel sets. Let $B_n = \bigcup_{i=1}^{n} A_{\mathrm{i}}$ and $B = \bigcup_{i=1}^{\infty} A_i$ Since $f\mathrm{I}_{B_n} \nearrow f\mathrm{I}_B$ By monotone convergence of Lebesgue measure

$$P(B_n) = \int_{B_n} fd\mu = \int f\mathrm{I}_{B_n} d\mu \nearrow \int f\mathrm{I}_B d\mu = \int_{B} fd\mu = P(B) \qquad (3.17)$$

Thus P is countably additive, hence it is Probability measure.

Example 3.3. *Find the constant* k *if following are the density functions*.

(i) $f(x) = k\mathrm{I}_{[-2,3]}(x)$

(*ii*) $f(x) = ke^{-2x}; x > 0$

Solution:(*i*) *The density is* 0 *outside* [- 2, 3] *and on* [- 2, 3] *it is 1*

$$. \int_{\mathbb{R}} f(x)dx = k\int_{-2}^{3} 1dx = 5k = 1 \tag{3.18}$$

$$k = \frac{1}{5}$$

(*ii*) *The density is* 0 *outside* $(0,\infty)$ *and on* $(0,\infty)$ *it ısf* $(x) = ke^{-2x}$

$$. \int_{\mathbb{R}} f(x)dx = k\int_{0}^{\infty} e^{-2x}dx = \frac{k}{2} = 1 \tag{3.19}$$

$$k = 2$$

Definition 3.6. ***Absolutely Continuous Probability Measure*:**

P *is a probability measure on Borel subsets* A *of* $\mathbb{R}$ *is said to be Absolutely Continuous probability Measure*, *if there exists a density* f *such that*

$$P(A) = \int_A f d\mu \tag{3.20}$$

Definition 3.7. ***Dirac Measure*:**

Let Ω *be at most countable arbitrary set*, $\mathcal{F}$ *be the family of subsets of* Ω. *A measure* δ_w *on* $\mathcal{F}$ *defined as*

$$\delta_\omega(A) = \begin{cases} 1 & \omega \in A \\ 0 & \omega \in \overline{A} \end{cases} \tag{3.21}$$

is called as Dirac Measure concentrated at w

Definition 3.8. Ω *be at most countable arbitrary set*) $\mathcal{F}$ *be the family of subsets* Ω *A Dirac measure on* $\mathcal{F}$ *say* δ_ω *A probability measure* P *defined as*

$$P(A) = \sum_{k=1}^{\infty} \alpha_k \delta_k \tag{3.22}$$

such that $\alpha_k \geq 0$ *and*

$$\sum_{k=1}^{\infty} \alpha_k = 1$$

is *said to be discrete probability measure*.

Remark 3.3. *Dirac Measure is a probability measure*.

- $\Omega - \{\omega\}$ *is largest set with measure O, and its every subset has also measure* 0. *Smallest set with measure 1 is* w

- *For* $\omega_1, \omega_2 \in \Omega$, $P(A) = \alpha\delta_{\omega_1}(A) + (1-\alpha)\delta_{\omega_2}(A)$ *is a probability measure. Where* $0 < \alpha < 1$.
- *Sometimes we come across measures which are neither discrete nor absolutely continuous. Following theorem is for such mixed probability measures.*

Theorem 3.5. P_1 , *and* P_2 *are the two probability measures,*

$P(A) = \alpha P_1(A) + (1-\alpha)P_2(A)$ *is a probability measure. Where* $0 < \alpha < 1$.

Proof:

1. $P(\Omega) = \alpha P_1(\Omega) + (1-\alpha)P_2(\Omega) = 1$,as both P_1 and P_2 are probability measures.
2. $P(A) = \alpha P_1(A) + (1-\alpha)P_2(A) \geq 0$ as $P_i(A) \geq 0, i = 1,2$.
3. A_n be the countable sequence of mutually exclusive events. I3y countable additivity of P_1 and P_2 ,

$$P\left(\cup_{n=1}^{\infty} A_n\right) = \alpha P_1\left(\cup_{n=1}^{\infty} A_n\right) + (1-\alpha)P_2\left(\cup_{n=1}^{\infty} A_n\right)$$

$$= \alpha\sum_{n=1}^{\infty} P_1(A_n) + (1-\alpha)\sum_{n=1}^{\infty} P_2(A_n)$$

$$= \sum_{n=1}^{\infty} P(A_n)$$

This shows that $\mathrm{P(A)}$ is a probability measure.

Remark 3.4. *Generalization of above result can be stated as* $: P_i, s$ *are the probability measures,* $P(A) = \sum \alpha_i P(A_i)$ *is a probability measure. Where* $0 < \alpha_i < 1$ *and* $\sum \alpha_i = 1$

Example 3.4. *Find the* $P(A)$ *if following are the density functions and* F *is absolutely Continuous probability measure* $w.r.t$ *it and* $A = (0,2\,]$.

(i) $f(x) = \frac{1}{6}\mathrm{I}_{[-3,3]}(x)$. (ii) $f(x) = e^{-x})x > 0$

Solution:(i)

$$P\left((0,2]\right) = \frac{1}{6}\int_0^2 1dx = \frac{1}{3}$$

(ii)

$$P\left((0,2]\right) = \int_0^2 e^{-x}dx = 0.8647$$

Example 3.5. *Define* P *by* $P(A)=\frac{1}{8}\delta_1(A)+\frac{3}{8}\delta_2(A)+\frac{1}{2}P_3(A)$,

P_3 *has density* $f(x)=\frac{1}{6}\mathrm{I}_{[-3,3]}(x)$ *Compute* $([2,3])$.

Solution: $\delta_1([2,3])=1, \delta_2([2,3])=1$, *and*

$P_3([2,3])=\frac{1}{6}\int_2^3 1dx=\frac{1}{6}P([2,3])=\frac{1}{8}+\frac{3}{8}+\frac{1}{12}=\frac{7}{12}.$

3.4 Chapter End Exercises

1. Find the constant k if following are the density functions.

   (i) $f(x)=k\mathrm{I}_{[-\sqrt{5},\sqrt{5}]}(x)$

   (ii) $f(x)=kx^3; 0<x<1$

2. Find the $\mathrm{P}(\mathrm{A})$, if following are the density functions and P is absolutely Continuous probability measure w.r. t it. $\mathrm{A}=(-1,0.5\,]$.

   (i) $f(x)=\frac{1}{4}\mathrm{I}_{[-1,4]}(x)$

   (ii) $f(x)=6x(1-x); 0<x<1$

3. $\Omega=(-\infty,\infty)\sigma$ - field C defined on Ω .A function P defined on space (Ω,C) as , for any $I\in C$

   $$P(I)=\begin{cases}1 & Iisfinite\\ 0 & Iisinfinite\end{cases}$$

4. Find Lebesgue Measure for following sets $(\mathbb{N}),[0,1]$, (, $\frac{1}{8},\frac{2}{5}$].

5. Show that Dirac Measure is a probability measure.

6. Define Paneity $(A)=\frac{1}{4}\delta_1(A)+\frac{1}{4}\delta_2(A)+\frac{1}{2}P_3(A)$

   P_3 has density $f(x)=2e^{-2x}$ Compute $P[1,3])$.’

7. Define P by $P(A)=\frac{1}{8}\delta_1(A)+\frac{3}{8}P_2(A)+\frac{1}{2}P_3(A)$, P_2 has density $f(x)=\frac{1}{2}\mathrm{I}_{[-1,1]}(x)$ and P_3 has density $f(x)=x^2;0<x<1$ .Compute P $([0,1])$.

8. If $P(A_n\Delta A)\to 0$ as $n\to\infty$ then show that $P(A_n)\to P(A)$.

9. Show that $P(AnBnC)\geq 1-P(-)-P(-)-P(-)$.

10. If A and B implies C then show that $P\ (-)\leq P(-)+P(-)$.

❖❖❖

4

CONDITIONAL PROBABILITY AND INDEPENDENCE

Unit Structure

4.0 Objectives

After going through this chapter, you will learn

- Conditional probability and its role in finding probability of simultaneous occurrence of events.
- Notion of independence of events and its consequences
- Total probability theorem.
- Bayes' theorem and its use to land posterior probabilities.

4.1 Conditional Probability and multiplication theorem

Let 13 be arbitrary set of Ω . Let A be the class of events of Ω.

$A_B = \{B \bigcap A \mid A \in A\}$

We can easily verify that A_B is a σ - field. And (B, A_B) is a measurable space. P measure on this space is not a probability as $P(B) \neq 1$,

let P_B is defined as

$$P_B(A) = \frac{P(A \bigcap B)}{P(B)} \qquad (4.1)$$

P_B is called as conditional probability measure or simply conditional probability of an event A

Theorem 4.1. *B be arbitrary set of Ω A be the class of events of Ω. $A_B = \{B \cap A \mid A \in A\}$*

$$P_B(A) = \frac{P(A \cap B)}{P(B)} \qquad (4.2)$$

P_B is a probability measure on (B, A_B)

proof:

1. $P_B(A) \geq 0$ for all $A \in A_B$

2. $P_B(B) = 1$

3. $A_1, A_2 \ldots$ be mutually exclusive sets of $\in A_B$

$$P_B(\cup_{i=1}^{\infty} A_i) = \frac{P(\cup_{i=1}^{\infty} A_i \cap B)}{P(B)} \qquad (10.3)$$

$$= \frac{P(\cup_{i=1}^{\infty}(A_i \cap B))}{P(B)} = \frac{\sum_{i=1}^{\infty} P(A_i \cap B)}{P(B)} = \sum_{i=1}^{\infty} P_B(A_i) \qquad (10.4)$$

From above it is clear that P_B is a probability measure on (B, A_B)

Remark 4.1. *Conditional probability is denoted by $P(A/B)$and called as conditional probability of an event A given event B has occurred Thus it is necessary to have $P(B) > 0$*

- $P(A/A) = 1$
- $P(A/\Omega) = 1$
- $P(A/B) \neq P(B/A)$
- *From the definition of conditional probability, it follows that*

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

which is also known as a Multiplication theorem on probability.

- *For three events Multiplication theorem on probability is stated as*

  $= P(A_1 \cap A_2 \cap A_3) = P(A_1 / A_2 \cap A_3)P(A_2 / A_3)P(A_3)$

The conditional probability is not defined when probability of given event is zero.

The conditional probability leads to another concept related with events, known as independence.

Example 4.1. *Show that*

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

Solution:

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

.. By definition of conditional prob.

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

.. By Distributive law

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

.. By Addition theorem on probability.

$$= \frac{P(A \cap C)}{P(C)} + \frac{P(\mathrm{B} \cap C)}{P(C)} - \frac{P(A \cap C \cap B)}{P(C)}$$

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

By definition of conditional prob.

Example 4.2. *Probability that it rains today is 0.4) probability that it will rain tomorrow is is 0.5, probability that it will rain tomorrow and rains today is 0.3. Given that it has rained today, what is the probability that it will rain tomorrow.?*

Solution: Let

$$P(A) = P \text{ (it } \textit{rains today}) = 0.4$$

$$P(B) = P \text{ (it } \textit{will rain tomorrow}) = 0.5$$

$$P(A \cap B) = P \; (\textit{it will rain tomorrow and rains today}) = 0.3$$

Required probability is

$$P(A / B) = \frac{P(A \cap B)}{P(B)} = 0.6$$

Example 4.3

A box contains cards numbered 1 to 25. A card bearing even number was drawn, but the number was not known. What /s the probability that it is card bearing number divisible by $5'$?

Solution: $\Omega = \{1, 2, 3 \ldots 25\}$,

$$P(B) = P\ (even\ no.\ card) = \frac{12}{25}$$

$$P(A) = P\ (card\ \ with\ no.\ divisible\ by\ 5) = \frac{5}{25}$$

$$P(A \cap B) = P\ (An\ \ even\ no.\ divisible\ by\ 5) = \frac{2}{25}$$

Required probability is

$$P(A/B) = \frac{P(A \cap B)}{P(B)} = \frac{2}{12}$$

4.2 Independence of the events

The occurrence and nonoccurrence of the event, when does not depend on occurrence and nonoccurrence of the other event the two events are said to be independent. Since occurrence and nonoccurrence of the event is measured in terms of probability. Instead of only independence we say stochastic independence or independence in probability sense. Let us first define independence of two events. we will later call it as pair wise independence

Definition 4.1. *Independence of the events*: *Let* (Ω, A, P) *be a probability space.* invents *A and B of this space are said to be stochastically independent or independent in probability sense if and only if* $P(A \cap B) = P(A)P(B)$

- Above definition works for any pair of events even when either $\mathrm{P}(\mathrm{A})$ or $\mathrm{P}(B)$ is equal to zero.
- Property of independence is reflexive.
- If A and B are independent then conditional probability and unconditional probabilities are same. That means if A is independent of B $P(A/B) = P(A)$ and $P(B/A) = P(B)$

Theorem 4.2. *If events A and B are independent so are*(i) A *and* $\overline{B}\left(ii\right)B$ *and* $\overline{A}$ (iii) $\overline{A}$ *and* $\overline{B}$.

proof: (i) Consider

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

Since A and B are independent

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

Consider

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

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

Thus

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

So, A and $\overline{B}$ are independent. Similarly we can prove (ii)

(iii) Consider

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

....By De Morgan's law

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

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

.......since A and I3 are independent

Thus

$$P\left(\overline{A}\cap\overline{B}\right)=[1-\left[P\left(A\right)\right]\left[1-P\left(B\right)\right]$$

$$=P\left(\overline{A}\right)P\left(\overline{B}\right)$$

So, $\overline{A}$ and $\overline{B}$ are independent.

Definition 4.2. *$A_t, i=1\ldots n$ events are mutually or completely independent if and only if for every sub collection*

$$P\left(A_1 n A_2 \cdots n A_k\right)=\coprod_{i=1}^{k}P\left(A_i\right)$$

for $k=2\ldots n$

Remark 4.2. *If the above condition holds for $k=2$ we say that events are pairwise independent. There are such nC_2 pairs, and those many conditions have to be checked. And for n events, to be completely independent, there are $2^n - n - 1$. conditions have to be checked.*

Remark 4.3. *If A, B, C are three events*

- *They are pairwise independent if*
  1. $P(A \cap B) = P(A)P(B)$
  2. $P(A \cap C) = P(A)P(C)$
  3. $P(B \cap C) = P(B)P(C)$
- *They are completely independent if*
  1. $P(A \cap B) = P(A)P(B)$
  2. $P(A \cap C) = P(A)P(C)$
  3. $P(B \cap C) = P(B)P(C)$ *And*
  4. $P(A \cap B \cap C) = P(A)P(B)P(C)$

4.3 Bayes' Theorem

It is possible to find probability of an event if conditional probabilities of such event given various situations. The situations need to be exhaustive and non-overlapping

Example 4.4. *An urn contains 5 white and 7 black balls.3 balls are drawn in succession. What is the probability that all are white'? If ball are drawn* (*i*) *with replacement* (*ii*) *Without replacement.*

Solution: *Let A_i be the event that i^{th} drawn ball is white, $i = 1,2,3$.* (*i*) *When balls are drawn Without replacement, events are not independent. Using multiplication theorem, Required prob.* $= P(A_1 \cap A_2 \cap A_3) =$

$$P(A_1)P(A_2 / A_1)P(A_3 / A_1 \cap A_2) = \frac{5}{12} \times \frac{4}{11} \times \frac{3}{10} = \frac{1}{22}$$

(*ii*) *When balls are drawn with replacement, events are independent*

$$= P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2)P(A_3) = \left(\frac{5}{12}\right)^3$$

Example 4.5. *A problem is given to three students whose chances of solving the problem are* 0.2,0.3 *and 0.5 respectively. If all of them solve the problem independently, find the probability that (i) None of them solves it. (ii) the problem is solved by exactly two students*

(iii) the problem is solved.

Solution: *Let* A_i *be the event that* $i^{th}, i=1,2,3$. *student solves the problem* $i=1,2,3$. (i) P $(\textit{None of them solves it}) = P\left(\overline{A_1} \cap \overline{A_2} \cap \overline{A_3}\right)$

Since they solve the problem independently, A_i *are independent, so are* $\overline{A_i}$

$$= P\left(\overline{A_1}\right)P\left(\overline{A_2}\right)P\left(\overline{A_3}\right) = 0.8 \times 0.7 \times 0.5 = 0.28$$

(ii) P (*the problem is solved by exactly two students*)

$$= P\left(A_1 \cap \overline{A_2} \cap \overline{A_3}\right) + P\left(\overline{A_1} \cap A_2 \cap \overline{A_3}\right) + P\left(\overline{A_1} \cap \overline{A_2} \cap A_3\right)$$
$$= 0.2 \times 0.7 \times 0.5 + 0.8 \times 0.3 \times 0.5 + 0.8 \times 0.7 \times 0.5 = 0.47$$

(iii) P (*the problem is solved*) $= 1 - P$ $(\textit{None of them solves it}) = 0.72$

Example 4.6. $\Omega = \{\{1,1,1\}\{1,2,1\}\{1,1,2\}\{2,1,1\}\}$, A : *first no. is 1*, B: *second no. is*1 , *C*:*third no. is 1*, *examine whether* A,B,C *are completely independent*?.

Solution:

$$P(A) = P(B) = P(C) = 0.5, P(A \cap B) = P(B \cap C) = P(A \cap C) = 0.25P(A \cap B)$$
$$= 0.25P(A \cap B) = P(A)P(B), P(A \cap C) = P(A)P(C), P(B \cap C) = P(B)P(C)$$

so A,B,C *are pairwise independent. But* $P(A \cap B) = 0.25 \neq P(A)P(C)P(B) = 0.125$, *hence they are not completely independent.*

Theorem 4.3. ***Theorem of total probability***: $A_1, A_2, .A_n$ *are forming partition of a* Ω, *Let* B *be another event,* $B \subset \Omega$ *then we can find probability of* B *by following relation*

$$P(B) = \sum_{i=1}^{n} P(B / A_i) P(A_i) \qquad (4.5)$$

proof: $A_1, A_2, .\ A_n$ are forming partition of a sample space, for every $i \neq j = 1,2$, ... n ; $\left(A_i \cap A_j\right) = \phi$ And $\left(\bigcup_{i=1}^{n} A_i\right) = \Omega$

$B \subset \Omega$

$$\therefore P(B) = P(B \cap \Omega)$$

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

Since A_is forms partition

$$= P\left(\mathrm{U}_{i=1}^{n} B \cap (A_i)\right)$$

...... By distributive law.

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

By finite additivity of Probability function

$$= \sum_{i=1}^{n} P(B / A_i) P((A_i))$$

.. ... By multiplication theorem. Hence

$$P(B) = \sum_{i=1}^{n} P(B / A_i) P(A_i)$$

- Though the theorem is proved for finite partition, it is also true for countable partition.
- At least two A_is are should have nonzero probability.
- If B is an effect and A_is are different causes $P(B)$ summarizes chance of the effect due to all possible causes.

Example 4.7. *Screws are manufactured by two machines A and B. Chances of producing defective screws are by machine A and B are* $4r_0$ *and* $1r_0$ *respectively. For a large consignment A produced* $70r_0$ *and B produced* $30r_0$ *screws. What is the probability that a randomly selected screw from this consignment is defective*?

Solution: *Let* A *be the event that screws are manufactured by two machines* A *and* B *be the event that screws are manufactured by machines B.* D *be the event that* *defective screws are manufactured. Given* $P(A)=0.7$, $P(B)=0.3$, $P(D/A)=0.04$, $P(D/B)=0.01$. *By Theorem of total probability* $P(D)=P(A)P(D/A)+P(B)P(D/B)=.031$

Example 4.8. *A ball is selected at random a box containing 3 white 7 black balls. If a ball selected is white it is removed and then second ball is drawn. If the first ball is black it is put back with 2 additional black balls and then second ball is drawn. What is the probability that second drawn ball is white' ?*

Solution: *Let* A_W *be the event that ball drawn at the first draw is white* A_B *be the event that ball drawn at the first draw is black.* D *be the event that ball drawn at the second draw is white Given* $P(A_W)=0.3, P(A_B)=0.7, P(D/A_W)=\frac{2}{9}$,

$P(D/A_B)=\frac{3}{12}$. *By total probability theorem,*

$$P(D)=P(A_W)P(D/A_W)+P(A_B)P(D/A_B)=0.2417$$

Theorem 4.4. ***Bayes' Theorem*** : $A_1, A_2, .A_n$ *are forming partition of a* Ω, *Let* B *be another event,* $B \subset \Omega$ *then we can find probability of* B *by following relation*

$$P(A_j/B)=\frac{(P(B/A_j)P(A_j)}{\sum P(B/A_i)P((A_i))} \tag{4.6}$$

proof:

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

By multiplication theorem.

And using $\mathrm{P(B)}$ from total probability theorem the proof of theorem follows.

- This theorem is useful for posterior analysis of cause and effect.
- Given $P(A_i)$, which are prior probabilities of the i^{th} cause.Where as $P(A_i/B)$ are posterior probability of the cause A_i given that I3 is effect observed.

Example 4.9. *Three people* X, Y, Z *have been nominated for the Manager' s post. The chances for getting elected for them are* 0 . 4, *0.35 and 0.25 respectively. If* X *will be selected the probability that he will introduce Bonus scheme is 0.6 the respective chances in respect of* Y *and* Z *are 0.3 and 0.4 respectively. If it w* *known that Bonus scheme has been introduced, what is the probability that X is selected as a Manager?*

Solution: *Let B be the event that bonus scheme is introduced. X, Y, Z denotes respectively that X, Y, Z are elected. Thus given $P(X) = 0.4$,, $P(Y) = 0.35, P(Z) = 0.25, P(B/X) = 0.6, P(B/Y) = 0.3, P(B/Z) = 0.4$ By Bayes Theorem,*

$$P(X/B) = \frac{P(B/X)P(X)}{P(B/X)P(X) + P(B/Y)P(Y) + P(B/Z)P(Z)}$$

$$= \frac{0.4 \times 0.6}{0.4 \times 0.6 + 0.25 \times 0.3 + 0.35 \times 0.4} = 0.5275$$

Example 4.10. 1% *of the population suffer from a dreadful disease. A suspected person undergoes a test. However the test making correct diagnosis* 90% *of times. Find the probability that person who has really caught by that disease given that the test resulted positive? Solution: Let A_1 be the event that person was really caught by that disease*

A_2 be the event that person was healthy

D be the event that person got the test positive

$P(D/A_1) = 0.9$, $P(D/A_2) = 0.1, A_1 = 0.01, A_2 = 0.99$

$$P(D) = P(A_1)P(D/A_1) + P(A_2)P(D/A_2) = 0.108$$

Required Probability is

$$P(A_1/D) = \frac{(P(A_1)P(D/A_1)}{(P(A_1)P(D/A_1) + (P(D/A_2)P(A_2)} = 0.08333$$

4.4 Chapter End Exercises

1. What is the probability that(i)husband, wife and daughter have same birthday (ii)two children have birthday in March?

2. 4 soldiers A,B,C and D fire at a target .Their chances of hitting the target are 0.4,0.3,0.75, and 0.6 respectively. They fire simultaneously. What is the chance that(i) the target is not hit? (ii) the target is hit by exactly one of them.

3. If A, B, C are independent show that, (i) A, $B \cap C$ are independent(ii)A, $B \cup C$ are independent(iii) A , $B - C$ are independent

4. $\Omega = \{1,2,3,4\}$, $A = \{1,2\}$ List all]3 such that A, B are independent.

5. If $P(A/B) < P(A)$ then $P(A/\bar{B}) > P(A)$, and vice versa.

6. Show that

$$P(A/\bar{B}) = \frac{P(A) - P(AnB)}{1 - P(B)}$$

, $P(B) \neq 0$,hence prove $P(AnB) \geq P(A) + P(B) - 1$

7. Examine for pairwise and mutual independence of events K, R , and S which are respectively getting of a king, red and spade card in a random draw from a well shuffled pack of 52 cards.

8. Urn A contains numbers 1 to 10 and B contains numbers 6 to 15. An urn is selected at random and from it a number is drawn at random. What is the probability of urn $l3$ was selected, if the number drawn is less than 7.

9. In a population of 55 % males and 45% females, 4% of the males and 1% of females are colorblind. Γind the probability a randomly selected person is colorblind person 7.

10. A man is equally likely to drive by one of the three routes A,B,and C from his home to office.The chances of being late to the office are 0.2, 0.4, 0.3 ,provided he has chosen the routes A, B, c respectively.If he was late on riday what is the prob. that he has chosen route C ?

5

RANDOM VARIABLE AND ITS DISTRIBUTION FUNCTION

Unit Structure

5.0 Objectives

After going through this chapter you will learn

- A real valued function defined on sample space,known as random variable
- Discrete and continuous r.v
- Distribution function of a r.v and its association with probability measure.
- Properties of Distribution function
- Probability mass function of a r.v.
- Probability density function of a r.v.

5.1 Random Variable

Definition 5.1. *Random Variable* (Ω, C) *be a measurable space. A real valued function* X *defined on this space is called as a random variable if every inverse image is a Borel set. That*$\uparrow$ *s for all* $B \in \mathrm{B}$ *we have*

$$X^{-1}(B) = \{\omega \mid X(w) \in B\} \in C$$

- Random variable is abbreviated by ‘ r.v ’
- X is a r.v iff for each $x \in \mathbb{R}, \{X \leq x\} \in C.$
- $X : \Omega \to \mathbb{R}$ .Further $\{\omega \in \Omega : X(\omega) \in B\}$ is an event or it is $\in C.$

- In chapter 2 we have seen that all intervals semi - open,semi - closed,single tones are $\in B$ .That is B may be $(-\infty,a)$ or $[a,\infty\ )$ etc.,so $\{x\},\{a<X<b\},\{a<X\leq b\},\{a\leq X<b\},\{a\leq X\leq b\}$ are all $\in\mathcal{B}$ hence are events.

Example 5.1. *Show that Indicator function is a $r.v.$*

Solution: *Indicator function is defined for a $A\subseteq\Omega$*

$$I_A(\omega)=\begin{cases}1 & \omega\in A\\ 0 & \omega\in\overline{A}\end{cases} \tag{5.1}$$

$I_A(\omega)$ is a $r.v$ on (Ω,C) iff $A\in C.$

Example 5.2. *Consider (Ω,C) be a sample space, where $\Omega=\{HH\ ,HT.TH,\ TT\}$ If $X(w)=$ Number of heads in ω,C is sigma field, Is X a $r.v$ ' ?*

Solution: *If $X(\omega)=$ Number of heads $in=\omega$ then $X(HH)=2,X(fi[T\ )$ $=1,X(TH)=1,X(I^{\gamma}T)=0$*

$$X^{-1}(-\infty,x]=\begin{cases}\varphi & x<0\\ TT & 0\leq x<1\\ HT,TH,HH & 1\leq x<2\\ \Omega & x\geq 2\end{cases} \tag{5.2}$$

All X^{-1} are events, so X is a $r.v.$

Example 5.3. *If X is a $r.v$ are following functions $r.v$?* (i) $aX+b$ (ii) $\dfrac{1}{X}$

Solution: *If X is a $r.v\{\omega:X\leq x\}\in C$*

Case *$I:a>0,b\in\mathbb{R}$ then $\left\{\omega:X\leq\dfrac{x-b}{a}\right\}\in C$*

$$\{\omega:aX+b\leq x\}\in C$$

Case *II: $a<0,b\in\mathbb{R}$ then as a complement of $\left\{\omega:X\leq\dfrac{x-b}{a}\right\}\in C$*

$$\left\{\omega:X>\frac{x-b}{a},a<0\right\}\in C$$

$$\{\omega\ :\ aX\leq x-b\}\in C$$

$$\{\omega : aX + b \le x\} \in C$$

Case *III*: $a = 0$

$$\{\omega : aX + b \le x\} = \{\omega : X \le x - b\} = \begin{cases} \Omega & x - b \ge 0 \\ \varphi & x - b < 0 \end{cases} \qquad (5.3)$$

Thus $aX + b$ is *a* *r.v.*

(ii) Let

$$\left[\frac{1}{X} \le x\right] = \left\{\frac{1}{X} \le x, X > 0\right\} \mathrm{U} \left\{\frac{1}{X} \le x, X < 0\right\} \mathrm{U} \left\{\frac{1}{X} \le x, X = 0\right\}$$

$$\left[w : \frac{1}{X} \le x\right] \begin{cases} \{X < 0\} & x = 0 \\ \{X \le x, X < 0\} \mathrm{U} \{X \ge \frac{1}{x,0}, X > 0\} & when\, x\, positive \\ \left\{X \ge \frac{1}{x}\right\} \mathrm{U} \left\{X \le \frac{1}{x}, X > 0\right\} & when\, x\, is\, negative \end{cases}$$

All events are $\in C$ *So*, $\left[\frac{1}{X} \le x\right]$ *is an event*, *hence* $\frac{1}{X}$ *is ar. v*

Example 5.4. $\Omega = \{1,2,3,4\}, C = \{\varphi, \Omega, \{1\}\{2,3,4\}\}$

Is $X(\omega) = 1 + \omega$ *is a random variable with respect to* (Ω, C) ?

Solution: *inverse image of* $\{\omega \in \Omega : X(\omega) = 3\} = \{2\} \notin C$

So $X(\omega) = 1 + \omega$ *is not a random variable*.

5.2 Distribution Function

Define a probability measure P_X by $P_X = P(X \in B)$,this is a mapping P_X : $\mathcal{B} \to [0,1].\mathcal{B}$ is a sigma field of borel sets. We define a point function P associated with probability space (Ω, C, P)

Definition 5.2. *Distribution Function of a random variable*: *A mapping* F_X :

IR $\mathbb{R} \to [0,1]$ *defined by* $F_X = P\left[\omega : X(\omega) \le x\right]$ *is called as a distribution function* $F_X(x)$ *of X.*

Example 5.5. *One of the numbers 2, ...* 1,2 *is chosen at random by throwing a pair of dice and adding the numbers shown on the two faces. You win $9 in case 2,3, 11 or 12 comes out or lose $ 10 if the outcome is* 7, *otherwise you do not lose or win anything. Find* $P[X > 0]$ *and* $P[X < 0]$

Solution $: \Omega = \{(a,b) : a,b = 1,2,3,4,5,6\}\, X \;:\; \Omega \to \mathbb{R}$ *define as*

$$X((v) = X(a,b) = \begin{cases} 9 & \text{if } a+b = 2,3,11,12 \\ -10 & \text{if } a+b = 7 \\ 0 & \text{if } a+b = 4,5,6,8,9,10 \end{cases} \qquad (5.4)$$

$$P[X > 0] = \{\omega : X(\omega) = 9\} = P\left[(1,1),(1,2),(2,1),(6,5)(5,6)(6,6)\right] = \frac{1}{6} P[X < 0] = \{\omega : X(\omega) = 7\} = \frac{1}{6}$$

Example 5.6. $X(\omega) = 1$ *for* $(\omega \in A, X(\omega) = -2$ *for* $\omega \in B, X(\omega) = 2$ *otherwise.*

B *are dis / oint . Find d.f of X.* $P(A) = \frac{1}{3}, P(B) = \frac{1}{2}$

Solution: x *is r.v , inverse image must be an event,*

$$X^{-1}(-\infty, x] = \begin{cases} 0 & x < -2 \\ B & -2 \le x < 1 \\ A & 1 \le x < 2 \\ \Omega & x \ge 2 \end{cases}$$

$d.fF_X(x)$ *of* X *, then* $F_X(x) = P_X(-\infty, x\,] =$

$$F_X(x) = P(-\infty, x] = \begin{cases} 0 & x < -2 \\ P(B) = \frac{1}{2} & -2 \le x < 1 \\ P(A) + P(B) = \frac{1}{2} + \frac{1}{3} = \frac{5}{6} & 1 \le x < 2 \\ P(\Omega) = 1 & x \ge 2 \end{cases}$$

- We can establish suitable correspondence between P and F as $F_X = P_X(-\infty, x\,]$

- $P[a < X \le b] = F_X^{\gamma}(b) - F_X(a)$

Distribution function (d.f) has following properties.

- $F_X(x)$ is non negative for all $x \in \mathbb{R}$

proof: We can easily verify this as $F_X = P[X \le x]$,and P is a probability measure.

- $F_X(x)$ is Monotonically non decreasing.

  proof: Let $x_1, x_2 \in \mathbb{R}$ such that $x_1 \le x_2$

  $(-\infty, x_1) \subset (-\infty, x_2)$

  Using monotone property of P, $P_X(-\infty, x_1) \le P_X(-\infty, x_2)$

  $F_X(x_1) \le F_X(x_2)$

- $F_X(x)$ is right continuous.

  proof: consider a sequence $x_n \downarrow$ x such that $x_1 > x_2 \ldots > x_n$, and events $B_n = (x, x_n\ ]$ now as B_n are decreasing events as seen in chapter 1, $B_n \downarrow \text{n}_n B_n = \varphi$, and using continuity property of P

  $$0 = P(\varphi) = P\left(\lim_n B_n\right) = \lim_n P(B_n) = \lim_n P(x, x_n] = \lim_n F_X(x_n) - F_X(x)$$

This implies right continuity,

$$\lim_n F_X(x_n) = F_X(x) \qquad (5.5)$$

(i) $\lim F_X(x) = F_X(-\infty) = 0.$

(ii) $\lim = F_X(\infty) = 1.$

proof: (i)Let $x_1 > x_2 \ldots > x_n$, and events $B_n = (-\infty, x_n\ ]$ now as B_n are decreasing events as seen in chapter 1, $B_n \downarrow \text{n}_n B_n = \varphi$, and using continuity property of P

$$0 = P(\varphi) = P\ (\ \lim_n B_n)\ = \lim_n P(B_n) = \lim_n P(-\infty, x_n\ ] \quad (5.6)$$

$$0 = \lim_n F_X(x_{n}) - F_X^{\gamma}(-\infty)$$

$$\lim_{n\to-\infty} F_X(x_n) = F_X(-\infty) = 0 \qquad (5.7)$$

(ii)Let $x_1 < x_2 \ldots < x_n$, and events $B_n = (-\infty, x_n\ ]$ now as B_n are increasing events as seen in chapter 1, $B_n \uparrow \text{U}_n B_n = \Omega$, and using continuity property of P

$$1 = P(\Omega) = P(\mathrm{U}_n B_n) = P(\lim_n B_n) = \lim_n P(B_n) \qquad (5.8)$$

$$1 = \lim_n P(-\infty, x_n] = \lim_n F_X(x_n) - F_X(-\infty)$$

But $F_X(-\infty) = 0$

$$\lim_{n\to\infty} F_X(x_n) = 1 \qquad (5.9)$$

Theorem 5.1. *Every d.f is a d.f of some r.v.*

Remark 5.1. *If X is a r.v on (Ω, CP) with $F_X = P[\omega : X(\omega) \le x]$ is associated d.f.*

By above theorem for every r.v we associate a d.f.on some prob. space.Thus given a r.v there exists a d.f and conversely.

Example 5.7. *1. Write d.f of the following r.v.s*

(i) $X(\omega) = C$, *for all* $\omega \in \Omega C$ *is constant*

(ii) X *is no. heads in tossing two coins.*

Solution: $(i) F_X(x) = P[X \le x] = P_X(-\infty, x] = 0 \ \ if\, x < C$

$F_X(x) = P[X \le x] = P_X(-\infty, x] = 1 \ \ if\, x \ge C$

(ii) $X = No.\ of\ heads,$ $\Omega = \{HH, HT, TH, TT\}$ $P[X = 0] = \frac{1}{4}$) $P[X = 1] = \frac{1}{2}, P[X = 2] = \frac{1}{4}$

$$F_X(x) = P[X \le x] \begin{cases} 0, & x < 0; \\ \frac{1}{4}) & 0 \le x < 1) \\ \frac{3}{4}, & 1 \le x < 2; \\ 1 & x \ge 2. \end{cases}$$

5.3 Discrete random variable and its p.m.f

Definition 5.3. *Discrete random variable: A random variable X is called as discrete if there exist an at most countable set D such that* $P[X \in D] = 1$

- Set D contains countable points $\{X = x_i\}$ .They have non negative mass .They are called as jump points or the point of increase of d.f.As seen before in chapter 1, $\{X = x_i\} \in \mathcal{B}$. And $\{\omega : X(\omega) = x_i\}$ is an event.We can assign$P_X\left[\omega : X(\omega) = x_i\right]$ denoted by $p(x_i)$ such that (i) $p(x_i) \geq 0$ and $\sum p(x_i) = 1$

- X is a discrete random variable if and only if $P_X$is a discrete probability measure.
- The distribution function of a discrete r.v is a step function.As $P[X = x] = F_X(x) - F_X(x_-)$ this jump at x.Where$F_X(x_-) = \lim F_X(x-h)$
- Random variable has its characteristic probability law. Γor discrete r.v it is also called as probability ass function (p.m.f)
- For X discrete, (i)$P(a < X \leq b) = F_X(b) - F_X(a)$

  (ii)$P(a < X < b) = F_X(b) - F_X(a) - P[X = b]$

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

  (iv)$P(a \leq X < b) = F_X(b) - F_X(a) + P[X = a] - P[X = b]$

Definition 5.4. *Probability mass function*: *A collection $p(x_i)$ which is representing $P(X = x_i)$ satisfying (i) $p(x_i) \geq 0$ and $\sum p(x_i) = 1$ is called as probability mass function $(p.m.f)$ of a discrete random variable X.*

Example 5.8. *Let X be no. of tosses of a coin up to and including the toss showing head first time. (i) Write p.m.f of X hence find P [X is even]. (ii) Also write $d.f$ of X.*

Solution: (i) *Let P ' be the chance of showing head. $1 - p = q$) is chance of showing tail.*

$$P[X = x] = P\ \{x-1 \text{ tosses are tail, } x^{th} \text{ toss is head}\} = pq^{x-1};\text{for } x = 1,2\ldots$$

$$P\ [X \text{ is even}] = P(\{w : x = even\}) = P\left(\cup_{i=even}\{x = i\}\right) = \sum_{i=even} P(\{x = i\})$$

$$P\ [X \text{ is even}] = p\left[q^1 + q^3 + q^5 \ldots\right] = pq\left[\sum_{i=1}^{\infty}(q^2)^i\right]$$

Using infinite geometric series with common ratio q^2

$$P\ [X\ \ is\ even] = \frac{pq}{1-q^2} = \frac{q}{q+1}$$

$(ii)\ d.f\ \ of\ X\ \ is$

$$F_X(x) = P[X \le x] = \sum_{x=1}^{n} p(x) = 1 - q^{7l}$$

$$F_X(x) = P[X \le x]\begin{cases} 0 & x < 1 \\ 1-q^n & n \le x < n+1; n = 1,2, \end{cases}$$

5.4 Continuous random variable and its p.d.f

Definition 5.5. ***Continuous random variable*** : *Reject random variable* X *is defined on* (Ω, CP) *with* $d.f\ F_X$ *is said to be continuous if* F *is absolutely continuous*.

Γ is absolutely continuous if there exists a density $f_X : X \to [0,1]$ defined as

$$P[X \in (a,b]] = P_X(a,b] = \int_a^b f_X(x)dx \quad (5.10)$$

For every a,b $\in \mathbb{R}$ This function f is called as probability density function (p.d.f)of a continuous r.v X

Definition 5.6. ***Probability density function***: *If f is* $p.d.f$ *of a continuous r.vX with* $d.f\ F$, *it satisfies*

1. $f \ge 0$
2. $\int_{-\infty}^{\infty} f_X(x)dx = 1$
3. $P[X \in (a,b]] = \int_a^b f_X(x)dx = F_X(b) - F_X(a)$

- For P absolutely continuous and f continuous for all x then

$$f(x) = \frac{dF_X}{dx} \qquad (5.11)$$

- For continuous r.v F_X is continuous, right as well as left. $F_X(x) = F_X(x_-) = F_X(x_+)$ Where $F_X^{\gamma}(x_+) = \lim F_X(x+h)$

- From above it is clear that $P\left[X=x\right]=0$ for continuous r.v
- For X continuous, $P(a<X<b)=P(a<X\le b)=P\left(a\le X\le b\right)=P(a\le X<b)=$ $F_X\left(b\right)-F_X\left(a\right)$

Example 5.9. *Coin is tossed. If it shows head you pay Rs.2 . If it show tail you spin a wheel which gives the amount to you, distributed with uniform prob. between Rs.0 to 10 you gain or loss is a random variable. Find the distribution function and use it to compute the probability that you will win at least 5.*

Solution: $P\left[X=-2\right]=\frac{1}{2}$ *and for* $\left[0,10\right], f\left(x\right)=\frac{1}{10}$, *so* $F_X\left(x\right)=\frac{x}{10}$

$$F_X\left(x\right)=\begin{cases}0 & x<-2\\ \overline{0}\,\frac{1}{2,1x} & 0\le x<10-2\le x<0\\ 1 & x\ge 10.\end{cases}$$

$P\ [X \text{ is at least } 5]=1-F\left(5\right)=\frac{1}{2}.$

Example 5.10. ***A r.X has p.d.f***

$$f\left(x\right)=\begin{cases}\dfrac{k}{x^2} & x\ge 100\\ 0 & otherwise\end{cases}$$

Find $\left(i\right)k\left(ii\right)P[50<X<200]$ (iii) M *such that* $P[X<M]=\frac{1}{2}$

Solution: (i) $\int_{-\infty}^{\infty} f_X\left(x\right)dx=1$ *So,*

$$\int_{\mathrm{i}00}^{\infty}\frac{k}{x^2}dx=k[-x^{-1}|_{100}^{\infty}=\frac{k}{100}$$

$\therefore \frac{k}{100}=1$, *gives* $k=100$

(ii)

$$P[50<X<200]=\int_{150}^{200}\frac{100}{x^2}dx=\frac{1}{3}$$

$\left(iii\right)M$ *such that* $P[X<M]=\frac{1}{2}$ *so*

$$\int_{100}^{M} \frac{100}{x^2} dx = \frac{1}{2}$$

$gives\ M = 200$

5.5 Chapter End Exercises

1. Find the smallest σ - field on Ω .Let $X(\alpha f) = |c \supset|$ is a random variable on $\Omega =$
$\{-2,-1,0,1,2\}$

2. Two dice are rolled. Let r.v X be the larger of the two numbers shown. Compute
$P_X([2,4])$.

3. $\Omega = [0,1]$ and C is a σ - field of borel sets in Ω

Is(i) $X((v) = (:)(\text{ii}) X(w) = \left|w - \frac{1}{2}\right|$ is a random variable on Ω w.r.t C

4. A r.X has p.d.f.Γind its d.f . Hence $P[|X| > 0.5]$

$$f(x) = \begin{cases} 1+x & -1 < x < 0 \\ 1-x & 0 \le x < 1 \\ 0 & \text{otherwise} \end{cases}$$

5. A r.X has d.f , find its p.d.f

$$F_X(x) = \begin{cases} 0 & 0 < x \\ \frac{x}{2} & 0 \le x < 1 \\ \frac{1}{2} & 1 \le x < 2 \\ \frac{3-x}{2} & 2 \le x < 3 \\ 1 & x \ge 3 \end{cases}$$

6. ff X is a r.v are following functions r.v?$(\text{i}) X^2 (\text{ii}) |X| (\text{iii}) \sqrt{X}$

❖❖❖

6

SOME SPECIAL RANDOM VARIABLES AND THEIR DISTRIBUTIONS

Unit Structure

6.0 Objectives

After going through this chapter you will learn

- Bernoulli and Binomial distribution and their properties.
- Poisson distribution its relation with Binomial distribution
- Normal distribution and its applications.

6.1 Bernoulli and Binomial distribution

In this chapter we will come across some typical r.v s and their distributions. In real life situation we come across many experiments which result into only two mutually exclusive outcomes. Generally the outcome of interest is called as ' Success' and other as 'Failure'. We assign a positive probability 'p' to success and $(1-p)$ to failure.

Definition 6.1. *Bernoulli r.v : A r.v X assuming values 1 and 0 with probabilities 'p' and $(1-p)$ is called as Bernoulli r.v*

- Thus Bernoulli r.v is same indicator function I_A with 'A' as success.
- The probability law of Bernoulli is also written as

$$P[X=x]=p^x(1-p)^{(1-x)}\ x=0,1$$

- Hence onward we denote $1-p$ by 'q'. Note that $p+q=1$

Example 6.1. *1. An indicator function* I_A *is a Bernoulli r.v , if we assign probability* $P(A) = p$ *and* $P(\overline{A}) = 1 - p$

When the trial of a Bernoulli experiment is repeated independently finite number of times say 'n' times, it gives rise to Binomial situation. If we count total number of successes in such n trials it is Binomial r.v .The probability law for Binomial r.v has n^{th} term of binomial expansion of $(p+q)^n$

Definition 6.2. ***Binomial distribution***: *A r.vX assuming values* $0,1,2\ldots n$ *is said to follow Binomial distribution if its p.m.f is given by* $P[X = x1 =$

$$\begin{cases} nC_x p^x q^{(n-x)} x = 0,1,\ldots n; 0 < p < 1, p+q = 1 \\ 0 otherwise \end{cases} \tag{6.1}$$

- The notation $X \to B(n,p)$ is used for showing X follows Binomial distribution with parameters n and p.
- Bernoulli r.v is particular case of Binomial with $n = 1$. And Binomial arises from sum of n independent Bernoulli r.v.s.
- The Binomial probabilities can be evaluated by the following recurrence relation, starting from $P[X = 0] = q^n$ And then using recursive formula,

$$P[X = x+1] = \frac{n-x}{x+1}\left(\frac{p}{q}\right)P[X = x] \tag{6.2}$$

This is forward formula. We can also start from $P[X = n]$ and use the equation as a backward formula.

- A r.v counting number of successes in n bernoulli trials follows 13(n,p) and counting number of failures $n - X = Y$ say, in n bernoulli trials follows B(n,q).
- We can easily verify that $P\left[X \in (0,1\ldots n)\right] = 1$, hence Binomial is a discrete r.v.

Example 6.2.

$X \to B(n,p)$ *and if* $y = n - x, Y \to B(n,q)$, *then show that*

$P_X[X = r] = P_Y[Y = n - r] = ^n C_r p^r q^{n-r}$

Solution:

$$P_X\left[X=r\right]={}^{n}C_r p^r q^{n-r}$$

$$P_Y\left[Y=n-r\right]={}^{n}C_{n-r} q^{n-r} p^r$$

But $nC_{n-r}={}^{n}C_r$

so,

$$P_X\left[X=r\right]=P_Y\left[Y=n-r\right]={}^{n}C_r p^r q^{n-r}$$

Head is thrice as likely as tail for a coin. It is flipped 4 times (i) *Write p.m.f of X, representing number of heads observed in this experiment.* (ii) *find probability of getting 3 or 4 heads.*

Solution : $P(H)=3P(T)$, *so* $P(H)=\frac{3}{4}=p$ *,tossing coin once is a Bernoulli trial.* X *counting number of heads observed in tossing such bais coin* 4 *times.* $(i)X$ *follows* $B\left(n=4, p=\frac{3}{4}\right)$

$P\left[X=x\right]=$

$$\begin{cases} {}^{4}C_x\left(\frac{3}{4}\right)^x\left(\frac{1}{4}\right)^{(n-x)} & x=0,1,2,3,4 \\ 0 & otherwise \end{cases} \qquad (6.3)$$

Required Prob $=P\left[X\geq 3\right]=1-P\left[X\leq 2\right]$

$$P\left[X\leq 2\right]=\left(\frac{1}{4}\right)^4+4\left(\frac{1}{4}\right)^3\left(\frac{3}{4}\right)+6\left(\frac{1}{4}\right)^2\left(\frac{3}{4}\right)^2=0.9198$$

So, $P\left[X\geq 3\right]=0.08016$

Example 6.3. *It is found that* $60r_0$ *of the health care victims are senior citizens . If a random sample of 10 victims is taken, what is the probability of getting exactly 3 senior citizens in this sample ,*

Solution: X *is number of victims who are senior citizens in this sample.* X *follows* $B(n=10, p=0.6)P\left[X=3\right]={}^{10}C_3 0.6^3 0.4^7=0.04247$

Example 6.4. X *follows* $B(n=6, p)$ *such that* $P\left[X=2\right]=P\left[X=4\right]$ *Find* p.

Solution: $P[X=2]=P[X=4]$

$$^6C_2 p^2 q^{(4)} = ^6 C_4 p^4 q^{(2)}$$

$p^2 = q^2$ *means* $p = q = \frac{1}{2}$

Example 6.5. *X follows $B(n,p)$, Y follows $B(m,p)$, If X,Y are independent. Find the probability distribution of $X+Y$. Solution: Consider $P[X+Y=k]$, for $k=0,1,\ldots m+n$*

$$P[X+Y=k] = P\left[\mathrm{U}_{x=0}^{n} X=x, Y=k-x\right]$$

$$=\sum\nolimits_{x=0}^{n} P[X=x, Y=k-x] \tag{6.4}$$

$$= \sum\nolimits_{x=0}^{n} P[X=x]P[Y=k-x] \text{ Since r.v.s are independent}$$

$$=\sum\nolimits_{x=0}^{n}\left[C_x p^x q^{(n-x)}\right]\left[C_x p^{k-x} q^{(m-k+x)}\right] \tag{6.5}$$

$$=\sum\nolimits_{x=0}^{n}\left[C_x^m C_{k-x} p^k q^{(m+n-k)}\right] \tag{6.6}$$

$$=\sum\nolimits_{x=0}^{n}\left[C_k^m p^k q^{(m+n-k)}\right] \tag{6.7}$$

Thus $X+Y$ follows $B(n+m,p)$.

6.2 Poisson distribution

Now let us introduce another commonly used discrete r.v . Many a times we come across a r.v counting number of occurrences in a fixed duration of time. Гor example, Number of deaths due to Maleria per month in Mumbai, Number of accidents per hour on a express highway. Thc number of defects in cloth per square metor is similar occasion where Poisson distribution is appropriate

Definition 6.3. *Poisson distribution: A discrete r.vX assuming values $0,1,2\ldots\infty$ is said to follow Poisson distribution if its p.m.f is given by*

$$P[X=x]=\begin{cases}\dfrac{e^{-\lambda}\lambda^x}{x!}; x=0,1,\ldots\infty; \lambda>0\\ 0 \quad otherwise\end{cases} \tag{6.8}$$

- The notation $X \to \mathcal{P}(\lambda)$ is used for showing X follows Poisson distribution with parameter λ
- The Poisson probabilities can be evaluated by the following recurrence relation,starting from $P[X=0]=e^{-\lambda}$ And then using recursive formula,

$$P[X=x+1]=\frac{\lambda}{x+1}P[X=x] \qquad (6.9)$$

all probabilities can be evaluated. Tables are also available for various values of λ

- We can easily verify that $P[X \in (0,1\ldots\infty)]=1$, hence Poisson is a discrete r.v.
- In Binomial situation if number of successes are very large and the chance of success is very small, but average number 'np' is fixed say λ then, Binomial probabilities tends to Poisson as n becomes very large

Theorem 6.1. *Poisson as a limiting distribution of Binomial* $X \to B(n,p)$, *then if* p *is very small and* n *becomes very large, but* 'np' *remains constant*$=\lambda$ *then Binomial probabilities tends to Poisson probabilities*

$$\lim_{n\to\infty} nC_x p^x q^{(n-x)} = \frac{e^{-\lambda}\lambda^x}{x!} x=0,1,\ \ldots\ \infty \qquad (6.10)$$

Proof: $X \to B(n,p)$, so $P[X=x]={}^n C_x p^x q^{(n-x)}$

By putting $p=\dfrac{\lambda}{n}$

$$\frac{n!}{x!n-x!}\left(\frac{\lambda}{n}\right)^x\left(1-\frac{\lambda}{n}\right)^n\left(1-\frac{\lambda}{n}\right)^{-x}$$

$$\lim_{n\to\infty}\frac{n!}{x!n-x!}=\frac{n(n-1)(n-2)\ldots(n-x+1)}{x!}\frac{(\lambda)^x}{{}^x n}=\frac{\lambda^x}{x^1}$$

$$\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^n\left(1-\frac{\lambda}{n}\right)^{-x}=e^{-\lambda}$$

Hence

$$\lim_{n\to\infty}{}^n C_x p^x q^{(n-x)}=\frac{e^{-\lambda}\lambda^x}{x!} x=0,1,\ \ldots\ \infty. \qquad (6.11)$$

Limiting distribution of Binomial is Poisson.

Example 6.6. *A sales firm receives on an average three toll - free calls per hour. For any given hour find the probability that firm receives* (i) *At most three calls.* (ii) *At least three calls*

Solution: $X = No$ *. of toll free calls a firm receives* $X \to \mathcal{P}(\lambda = 3)$

(i) *required prob*

$$P[X \le 3] = \sum_{x=0}^{3} \frac{e^{-\lambda}\lambda^x}{x!} = 0.6472$$

(ii) *required prob*

$$P[X \ge 3] = \sum_{x=3}^{\infty} \frac{e^{-\lambda}\lambda^x}{x!} = 1 - P[X \le 2] = 1 - 0.5768 = 0.4232$$

Example 6.7. *A* $r.vX\mathcal{P}(\lambda)$, *such that* $P[X = 4] = P[x = 5]$. *Find* λ *Solution*: *As* $P[X = 4] = P[x = 5]$.

$$\frac{e^{-\lambda}\lambda^4}{4!} = \frac{e^{-\lambda}\lambda^5}{5!}$$

$$\text{So}, \lambda = 5$$

Example 6.8. *A safety device in laboratory is set to activate an alarm, if it register 5 or more radio active particles within one second. If the back ground radiation is such that, the no of particles reaching the device has the poison distribution with parameter* $\lambda = 0.5$ flow *likely is it that alarm will be activated within a given one second period*?

Solution: *Let* X *be number of particles reaching safety device within a one sec. period .* $X \to \mathcal{P}(\lambda = 0.5)$ *The alarm will be activate if* $X \ge 5$

$P[x \ge 5] =$

$$\sum_{x=5}^{\infty} \frac{e^{-\lambda}\lambda^x}{x!} = 0.112$$

Example 6.9. $X_1 \to \mathcal{P}(\lambda_1)$ *and an independent* $r.vX_2 \to \mathcal{P}(\lambda_2)$

Show that $X_1 + X_2 \to \mathcal{P}(\lambda_1 + \lambda_2)$.

Solution: *Consider* $P[X + Y = k]$, *for* $k = 0, 1, \ldots \infty$

$$P[X + Y = k] = P\left[\cup_{x=0}^{\infty} X_1 = x, X_2 = k - x\right] = \sum_{x=0}^{\infty} P[X_1 = x, X_2 = k - x] =$$

$$\sum_{x=0}^{\infty} P[X_1 = x] P[X_2 = k - x] \text{ Since } r.vs \text{ are independent}$$

$$= \sum_{x=0}^{\infty} \left[\frac{e^{-\lambda_1}(\lambda_1)^x}{x!}\right]\left[\frac{e^{-\lambda_2}(\lambda_2)^{(k-x)}}{(k-x)!}\right] = \sum_{x=0}^{\infty} \left[\frac{e^{-(\lambda_1+\lambda_2)}(\lambda_1 + \lambda_2)^k}{k!}\right]$$

Thus $X_1 + X_2 \to \mathcal{P}(\lambda_1 + \lambda_2)$.

Example 6.10. 2% *students are left - handed. In a class of 200 students find the probability that exactly 5 are left - handed.*

Solution: *X is no. of left - handed in 200, $p=.02$ and $n=200$ thus X has $B(n=200, p=.02)$. Using Poisson as a limiting distribution of Binomial,*

$$X \to \mathcal{P}(\lambda = np = 4)$$

$$P[x \geq 5] = \frac{e^{-4}\lambda^5}{5!} = 0.1563$$

6.3 Normal Distribution

Definition 6.4. *Normal Distribution: A continuous r.v is said to follow Normal distribution if its p.d.f is given by*

$$f(x) = \begin{cases} \dfrac{e^{-1}z\left(\dfrac{x-\mu}{\sigma}\right)^2}{\sqrt{2\pi\sigma}} - \infty < x\infty; \mu \in \mathbb{R}, \sigma > 0 \end{cases} \qquad (6.12)$$

- The notation $X \to N(\mu, 0^2)$ is used to show that X follows normal distribution with parameters μ and σ^2
- Normal distribution is applicable to wide range of situations in real life.
- $\mu =$ Mean of X and $0^{-2} =$ Variance of X
- When $\mu = 0$ and $\sigma^2 = 1$ it is called as standard normal distribution. The tables for $P[X \leq x]$ are available for this distribution.
- Since any $X \to N(\mu, \sigma^2)$ its linear combination $Y = aX + b$ also has Normal distribution with parameters $a\mu + b$ and $a^2\sigma^2$
- If $X \to N(\mu, \sigma^2)$ then $Z = \dfrac{X-\mu}{\alpha}$ has standard normal distribution.
- We denote by $\varphi(z) = P[Z \leq z]$ as d.f of $N(0,1)$.
- $X_1 \to N(\mu_1, \sigma_1^2)$ and $X_2 \to N(\mu_2, \sigma_2^2)$,if X_1, X_2 are independent r.vs,then

  $X_1 + X_2 \to N(\mu_1 + \mu_1, \sigma_2^2 + \sigma_1^2)$

Example 6.11. *The score of the test are normally distributed with* $N(\mu = 100, \sigma = 15)$ *. Find the probability that score is below 112.*

Solution**:** *Score is denoted by* $X, X \to N(\mu = 100, \sigma = 15)$

$$P[X < 112] = P[Z < \frac{112 - 100}{15}] = \varphi(0.8) = 0.7881$$

, By normal tables. Since $Z = \dfrac{X - \mu}{\sigma}$ *has standard normal distribution.*

Example 6.12. $X_1 \to N(4, 1.5^2)$ *and* $X_2 \to N(-2, 2^2)$, *if* X_1, X_2 *are independent r.vs , then find* $P[X_1 + X_2 \geq 1]$

Solution**.** $X_1 \to N(\mu_1, \alpha_{1^2})$ *and* $X_2 \to N(\mu_2, 0_{2^2})$, *if* X_1, X_2 *are independent r.v.s., then*

$X_1 X_2 \to N(2, (2.25 + 4)) as X_1, X_2 are independenat\, r.v.s$

$P[X_1 + + X_2 \geq 1] = P\left[Z \geq \left(\dfrac{1-2}{25}\right)\right] = 1 - \varphi(-0.4) = \phi(0.4) = 0.6554$

6.4 Chapter End Exercises

1. In order to qualify police academy, candidates must have score in top 10%in the general ability test. If the test scores $X \to N(\mu = 200, \alpha = 20)$, find the lowest possible score to qualify.

2. $X_1 \to N(10, 2.5^2)$ and $X_2 \to N(12, 2^2)$,if X_1, X_2 are independent r.vs,then,find mean and variance of $X_1 + X_2$

3. On an average 0.2% of the screws are defective. Гind the probability that in a random sample of such 200 screws we get exactly 3 defective screws.

4. $X \to \mathcal{P}(\lambda)$. Гindλ if $\dfrac{P[X=4]}{P[X=3]} = \dfrac{3}{8}$.

5. X follows $B(n = 5, p)$. Find p if $\dfrac{P[X=4]}{P[X=3]} = \dfrac{3}{8}$.

6. A video tape on an average one defect every 1000 feet.What is the probability of atleast one defect in 3000 feet?

7. 3% of all the cars fail emission inspection. Γind the probability that in a sample of 90 cars three will fail.Use(i) Binomial distribution (ii)Poisson approximation to Binomial.

8. If a student randomly guesses at five multiple choice questions, find the probability of getting three or more correct answers .There are four possible options for each question.

9. $X \to N(\mu = 10, \sigma = 3)$. Find the probability that

(i) X is less than 13 but X is greater than 7.$(\text{ii})\, Y = 2X + 3$,then $Y < 26$.

$(\text{iii})\, X^2 100 (\text{iv}) |X| > 8$.

10. X_i follows $B\left(n = 5, \frac{i}{3}\right)$. where $i = 1, 2.(\text{i})$ Write p.m f of $X_1 + X_2$

(ii) Find $P[X_1 + X_2 \leq 3]$

❖❖❖

7

TWO-DIMENSIONAL R.V.S

Unit Structure

7.0 Objectives

After going through this chapter, you will learn

- Two dimensionaldiscrete r.v.s and its Joint Probability mass function.
- Two-dimensional continuous r.v.s and its joint Probability density function.
- From joint probability function of two-dimensional r.v.s finding marginal Probability laws.
- The conditional distributions of the r.v.s.
- Notion of independence of r.v.s and its consequences.

7.1 Probability Distributions of two-dimensional discrete r.v.s

The notion of r.v can be extended to multivariate case .In particular if X and Y are two r.v.s. defined on same probability spacc $(\Omega, C,, P)$ then $\{(X,Y) \in B\} \in C$, for any borel set B in $\mathbb{R}^2$ .Note that this Borel set is a $0^{\cdot}$ field generated by rectangles $(a,b)X(c,d)$ The mapping (X,Y) : $(\Omega, C) \to \mathbb{R}^2$ is a two dimensional r.v.

Definition 7.1. *Joint Probability Distribution of two dimensional r.v.s : The probability measure $P_{X,Y}$ defined on $\mathbb{R}^2$ is called as Joint Probability Distribution of two dimensional $r.v.s(X,Y)$ where*

$$P_{X,Y}(B) = P\left[(X,Y) \in B\right] \textit{ for every borel set } B \subset \mathbb{R}^2 \tag{7.1}$$

Definition 7.2. *Two dimensional discrete r.v . The two*-dimensional *random variable (X,Y) is called as discrete if there exist an at most countable set D such that* $P_{X,Y}(D) = 1$

Two r.v.s are jointly discrete if and only if they are discrete.

- The joint probability law $P_{X,Y}$ *O*f two dimensional discrete r.v s satisfy following
  1. $P_{X,Y}(x,y) \geq 0$ for all x,y
  2. $\sum_x \sum_y P_{X,Y}(x,y) = 1$
- The joint probability law $P_{X,Y}$ *O*f two dimensional discrete r.v s is also called as joint probability mass function of (X,Y)
- $P_X(x) = \sum_y P_{X,Y}(x,y)$ is called as Marginal p.m.f of X.
- $P_Y(y) = \sum_x P_{X,y}(x,y)$ is called as Marginal p.m.f of Y.
- Marginal p.mfs are proper p.m.fs of one dimensional discrete r.v.s.

Example 7.1. *Given following verify which are joint probability mass function of joint p.m.f , if so, find the constant K.*

(i) $P_{X,Y}(x,y) = K(x+y)$ *for* $x = 1,2,3$ $y = 1,2$.

(ii) $P_{X,Y}(x,y) = Kxy$ *for* $x = -1,0,1; y = -1,0,1$

Solution: $(i) P_{X,Y}(x,y) \geq 0$ *for all* x, y *if* $K > 0$

And

$$\sum_x \sum_y P_{X,Y}(x,y) =$$

$$P(1,1) + P(1,2) + P(2,1) + P(2,2) + P(3,1) + P(3,2) = 21K = 1$$

So, for $\mathrm{P_{X,Y}}(\mathrm{x},\mathrm{y})$ to be proper joint p.m.f $\mathrm{K} = \frac{1}{21}$,

(ii) $P(1,-1) < 0$, we can not have positive prob *for remaining pairs*, *if K is selected negative*. *Which means that for no K* , $P_{X,y}(x,y) > 0$ $P_{X,y}(x,y)$ *is not proper joint p.m.f*

Example 7.2. *Two cards are drawn from a pack of cards*. *Let X denotes no*. *of heart cards and Y no*. *of red cards*. *Find the joint p.m.f of r.v(X,Y)* . *Hence* $P[X = Y]$

Solution: $x, y = 0,1,2$. *following will be* / *ointp.m.f of* (X,Y)

$$P(1,2) = P[1heart, 2red] = \frac{13\times 25}{52C_2} = \frac{25}{102}$$

and so on.

$Y\downarrow\ X\rightarrow$	0	1	2	$P[Y=y]$
0	$\frac{25}{102}$	0	0	$\frac{25}{102}$
1	$\frac{26}{102}$	$\frac{26}{102}$	0	$\frac{52}{102}$
2	$\frac{6}{102}$	$\frac{13}{102}$	$\frac{6}{102}$	$\frac{25}{102}$
$P[X=x]$	$\frac{57}{102}$	$\frac{39}{102}$	$\frac{6}{102}$	1

And

$$P[X=Y] = P(0,0) + P(1,1) + P(2,2) = \frac{25+26+6}{102} = \frac{19}{34}$$

Example 7.3. *Using the above joint p.m.f of* X,Y

Solution: $P_X(x) = \sum P_{X,y}(x,y)$ *find marginal p.m.f.s of r.v.s* X,Y. *is as Marginal p.m.f of* X.

and $P_Y(y) = \sum P_{X,Y}(x,y)$ *is Marginal p.m.f of* Y.

X	0	1	2
Px(x)	$\frac{19}{34}$	$\frac{13}{34}$	$\frac{1}{17}$

Y	0	1	2
Py(x)	$\frac{25}{102}$	$\frac{52}{102}$	$\frac{25}{102}$

7.2 Probability Distributions of two dimensional continuous r.v.s

Definition 7.3. *Two dimensional continuous r.v . and their joint probability density function*: *The two dimensional random variable* (X,Y) *is called as continuous if there exists a function* $f_{X,Y} : \mathbb{R}^2 \rightarrow [0,1]$

satisfying

1. $f_{X,Y}(x,y) \geq 0$ *for all* x, y

2. $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dxdy = 1$

3. $\int_{a}^{b}\int_{c}^{d} f_{X,Y}\,dxdy = P[a \leq X \leq b, c \leq Y \leq d]$ *where* a,b,c, *and* d *are* $\in \mathbb{R}$

$g_X(x) = \int_{-\infty}^{\infty} f_{X,\text{y}}(x,y)\,dy$ is called as Marginal p.d.f of X.

$h_{\text{y}}(y) = \int_{-\infty}^{\infty} f_{X,\text{y}}(x,y)\,dx$ is called as Marginal p.d.f of Y.

Marginal p.d.fs are proper p.m.fs of one dimensional continuous r.v.s

Example 7.4. *Given following verify which are joint probability density function of* $(X,Y))$ *if so, find the constant* K.

(i) $f_{X,\text{y}}(x,y) = Ke^{-(x+y)}$ *for* $x \geq 0\, y \geq 0$.

(ii) $f_{X,\text{y}}(x,y) = Kxy$ *for* $0 < x < y < 1$

Solution: (i) $f_{X,Y}(x,y) \geq 0$ *for all* x, y *if* $K > 0$

And

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dxdy =$$

$$\int_{0}^{\infty}\int_{0}^{\infty} Ke^{-(x+y)}\,dxdy = K.\int_{0}^{\infty} e^{-y}(-e^{-x}\,|_{0}^{\infty}\,dy = K$$

So, *for* $f_{X,Y}(x,y)$ *to be proper joint* $p.d.f K = 1$

(ii) $f_{X,Y}(x,y) \geq 0$ *for all* x, y *if* $K > 0$

And

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dxdy = \int_{0}^{1}\int_{0}^{y} Kxydxdy =$$

$$K\int_{0}^{1} y.\frac{x^2}{2}\,|_{0}^{y}\,d_y = K\frac{y^4}{8}\,|_{0}^{1} = \frac{K}{8}$$

So, *for* $f_{X,Y}(x,y)$ *to be proper joint* $p.d.f K = 8$

Example 7.5. *For the above two joint* $p.d\,.\,fs$ *find* $(I)P[X < 0.5, \text{y} < 0.5]$ (II)*Marginal* $p\ d.f.s$ *of* X *and* Y.

Solution:(i) $f_{X,\text{y}}(x,y) = e^{-(x+y)}$ *for* $x \geq 0\, y \geq 0$.

(I)

$$P[X < 0.5, Y < 0.5] = \int_{0}^{0.5}\int_{0}^{0.5} e^{-(x+y)}\,dxdy = \left(1 - e^{-0.5}\right)^2$$

$(II)\, g_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy$ *is Marginal p.d.f of X.*

Marginal p.d.f of X. $g_X(x) = \int_0^{\infty} e^{-(x+y)}dy = e^{-x}$ *for* $x \geq 0$

$h_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx$ *is Marginal p.d.f of Y.*

Marginal p.d.f of Y. $h_Y(y) = \int_0^{\infty} f_{X,Y}(x,y)\,dx = e^{-y}$ *for* $y \geq 0$

(*ii*) $f_{X,y}(x,y) = 8xy$ *for* $0 < x < y < 1$

(*I*)

$$P[X < 0.5, y < 0.5] = \int_0^{0.5}\int_0^{y} 8\ xydxdy$$

$$= 8\int_0^{0.5} y\frac{x^2}{2}\Big|_0^y\, dy = 8\int_0^{0.5}\frac{y^3}{2}dy = (0.5)^4 = 0.0625$$

(*II*) *Marginal p.d.f of X.* $g_X(x) = \int_x^1 8xydy = 4x(1-x^2)$ *for* $0 < x < 1$

Marginal p.d.f of Y. $h_Y(y) = \int_0^y 8xydx = 4y^3$ *for* $0 < y < 1$

Example 7.6. *The joint probability density function of* (X,Y)

$$f(x,y) = \begin{cases} \dfrac{1+xy}{4} & |x| < 1, |y| < 1 \\ 0 & Othrewise \end{cases}$$

Find $P[X^2 < u, Y^2 < v]$

Solution:

$$P[X^2 < u, Y^2 < v] = P[-\sqrt{u} < X < \sqrt{u}, \sqrt{v} < Y < \sqrt{v}]$$

$$= \int_{-\sqrt{u}}^{\sqrt{u}}\int_{-\sqrt{v}}^{\sqrt{v}} \frac{1+xy}{4}dxdy = \sqrt{uv}$$

7.3 Conditional Probability distributions

Definition 7.4. *Conditional Probability mass function*: *Let the joint probability law of a two dimensional discrete r.v* (X,Y) *be* $P_{X,Y}$ *and the marginal p.m.f of X and Y be* $P_X(x), P_Y(y)$ *respectively, then the conditional p.m.f of X given* $Y = y$ *is given by*

$$P_{X/y=y}(x) = \frac{P_{X,y}(x,y)}{P_y(y)} \text{ for all } x, \text{provided} P_y(y) \neq 0 \qquad (7.2)$$

And the conditional p.m.f of Y given $X = x$ is given by

$$P_{Y/X=x}(y) = \frac{P_{X,Y}(x,y)}{P_X(x)} \text{ for all } y \text{, provided } P_X(x) \neq 0 \qquad (7.3)$$

Note that conditional p.m.f s are proper p.m.f.

Definition 7.5. *Conditional Probability density function* : *Let the joint probability law of a two dimensional continuous r.v.(X,Y) is $f_{X,Y}(x,y)$, and the marginal p.d.f of X and Y be $g_X(x), h_Y(y)$ respectively, then the conditional p.d.f of X given $Y = y$ is given by*

$$g_{X/y=y}(x) = \frac{f_{X,Y}(x,y)}{h_Y(y)} \text{ for all } x \text{, provided } h_Y(y) \neq 0 \qquad (7.4)$$

And the conditional p.d.f of Y given $X = x$ is given by

$$h_{Y/X=x}(y) = \frac{f_{X,Y}(x,y)}{g_X(x)} \text{ for all } y \text{, provided } g\ x(x) \neq 0 \qquad (7.5)$$

Note that conditional p.d.f s are proper p.d.f.s

Example 7.7. *There are 4 tickets in a bowl, two are numbered 1 and other numbered numbers 2. Two tickets are chosen at random from the bowl. X denotes the smaller of the numbers on the tickets drawn, and Y denotes the smaller of the numbers on the tickets drawn.*

(i)Find the joint p.m.f . of r.v.(X,Y)

(ii)Find the conditional p.m.f of Y given $X = 2$ (iii) Find the conditional p.m.f of X given $Y = 2$.

Solution: $\Omega = \{(1,1),(1,2),(2,1),(2,2)\}$

joint p . m.f. of r.v.(X,y)

$Y\downarrow\ X\rightarrow$	1	2	$P_Y(y)$
1	$\frac{1}{4}$	0	$\frac{1}{4}$
2	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{3}{4}$
$P_X(x)$	$\frac{3}{4}$	$\frac{1}{4}$	1

The conditional p.m.f of Y given $X = 2$ is given by

$$P_{Y/X=2}(y) = \frac{P_{X,Y}(x,y)}{\frac{1}{4}} \text{ for } y = 1,2\ . \qquad (7.6)$$

y	1	2
$P_{Y/X=2(y)}$	0	1

The conditional p.m.f of X given $Y=2$ is given by

$$P_{X/Y=2}(x)=\frac{P_{X,Y}(x,y)}{\frac{3}{4}} \text{ for } x=1,2. \quad (7.7)$$

x	1	2
$P_{X/Y=2(x)}$	$\frac{2}{3}$	$\frac{1}{3}$

Example 7.8. *The joint p.d.f of* $f_{X,Y}(x,y)=8xy$ *for* $0<x<y<1$ $r.v.(X,Y)$

(i)Find the conditional p.d.f of Y given $X=x$ (ii) Find the conditional p.d.f of X given $Y=y$.

Solution: $f_{X,y}(x,y)=8xy$ *for* $0<x<y<1$

Marginal p.d.f of X. $g_X(x)=\int_x^1 8xydy=4x(1-x^2)$ *for* $0<x<1$

Marginal p.d.f of y.$h_Y(y)=\int_0^y 8xydx=4y^3$ *for* $0<y<1$

(i) *The conditional p.d.f of X given $Y=y$ is given by*

$$g_{X/y=y}(x)=\frac{f_{X,y}(x,y)}{h_Y(y)}=\frac{8xy}{4y^3}=\frac{2x}{y^2} \text{ for } 0<x<y \quad (7.8)$$

And (ii) the conditional p.d.f of Y given $X=x$ is given by

$$h_{Y/X=x}(y)=\frac{f_{X,Y}(x,y)}{g_X(x)}=\frac{8xy}{4x(1-x^2)}=\frac{2y}{1-x^2} \text{ for } x<y<1 \quad (7.9)$$

7.4 Independence of r.v.s

Definition 7.6. *Independence of r.v.s*: (X,Y) *be two dimensional r.vs they are said to be independent if and only if, the events $X\in A$ and $Y\in B$ are independent for any Borel sets $A,B\in\mathbb{R}$*

- For the two dimensional discrete r.v s(X,Y) are independent if and only if the joint probability mass function is equal to the product of marginal p.m .fs that is

$$P_{X,Y}(x,y)=P_X(x)P_y(y) \text{ for all x,y} \quad (7.10)$$

- For the two dimensional continuous r.v s(X,Y) are independent if and only if the joint probability density is equal to the product of marginal p.d .fs that is

$$f_{X,Y}(x,y)=g_X(x)h_Y(y) \text{ for all x,y} \tag{7.11}$$

- When the r.v.s are independent their conditional p.m.f. s / p.d.fs are same as marginal p.m.f. s / p.d.fs

Example 7.9. *1. Verify whether* (X, Y) *are independent* $r\cdot v.s.$

(i) *The joint* $p.m.f$ *. of* (X,Y) *is* $P(0,0)=\frac{1}{9}$ $P(1,1)=\frac{1}{9}, P(0,1)=\frac{5}{9}, P(1,0)=\frac{2}{9}$

(ii) *The joint* $p.d.f$ *of* $f_{X,Y}(x,y)=8xy$ *for* $0<x<y<1$

Solution:(i)*The joint* $p.m.f$ *. of* (X,Y) *is*

$Y\downarrow$ $X\rightarrow$	O	1	$P_Y(y)$
0	$\frac{1}{9}$	$\frac{2}{9}$	$\frac{1}{3}$
1	$\frac{5}{9}$	$\frac{1}{9}$	$\frac{2}{3}$
$P_X(x)$	$\frac{2}{3}$	$\frac{1}{3}$	1

$P_X(0)P_Y(0)=\frac{2}{3}\times\frac{1}{3}$

$P_{X,Y}(0,0)=\frac{1}{9}$

$P_X(0)P_Y(0)\neq P_{X,Y}(0,0)$

$\therefore X,Y$ *are not independent.*

(ii) *The joint density is* $f_{X,Y}(x,y)=8xy$ for $0<x<y<1$ *Marginal* $p.d.f$ *of*

$X.g_X(x)=\int_x^1 8xydy=4x(1-x^2)$ *for* $0<x<1$

Marginal $p.d.f$ *of* $Y.$ $h_Y(y)=\int_0^y 8xydx=4y^3$ *for* $0<y<1$

$f_{X,Y}(x,y)\neq g_X(x)\times h_Y(y)$

X,Y *are not independent.*

7.5 Chapter End Exercises

1. Given following verify which are joint probability mass function of (X,Y), if so, find the constant K.

   (i) $P_{X,Y}(x,y) = Kx$ for $x = 1,2,3; y = 0,1,2.$

   (ii) $P_{X,Y}(x,y) = K\frac{x}{y}$ for $x = 1,2; y = 1,4$

2. Given following verify which are joint probability density function of (X,Y) . If so,

   (i) find the constant K.

   (ii) find marginal p.d.f of X, Y,

   (iii) Verify whether they are independent.

   (a) $f_{X,y}(x,y) = K$ for $x, y \in [0, 1],$

   (b) $f_{X,Y}(x,y) = \lambda\mu.e^{-(\lambda x+\mu y)}$ for $x \geq 0 y \geq 0; .\lambda, \mu > 0$

3. Given following joint probability mass function of (X, y)

   (i) Find the conditional p.m.f of Y given X = 2

   (ii) Find the conditional p.m.f of X given Y = 2.

   (iii) Also verify their independence.

   (I)

Y ↓ X - →	1	2	3	$P[Y=y]$
0	$\frac{1}{18}$	$\frac{2}{18}$	$\frac{3}{18}$	$\frac{1}{3}$
1	$\frac{1}{18}$	$\frac{2}{18}$	$\frac{3}{18}$	$\frac{1}{3}$
2	$\frac{1}{18}$	$\frac{2}{18}$	$\frac{3}{18}$	$\frac{1}{3}$
$P[X=x]$	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{2}$	1

   (II) Two fair dice are tossed . X is maximum of the number on two faces and Y is sum of the number on them.

4. The joint p.d.f of $f_{X,Y}(x,y)=2$ for $0<x<y<1$

r.v.(X,Y)

(ii)Find the conditional p.d.f of Y given X $=$ x (iii) Find the conditional p..f of X given Y $= y$

5. Find the constant K, if the joint p.m.f of (X,Y) is given as

$P(x,y)=K\left(3^{x-1}4^{y}\right)^{-1}$ for $x,y=1,2$, ...Also verify whether X, Y are independent.

❖❖❖

8

EXPECTATION, VARIANCE AND THEIR PROPERTIES

Unit Structure

8.0 Objectives

After going through this chapter you will learn

- The expected value of functions of r.v.s.
- Properties of expectation.
- Variance and its role in studying r.v
- Characteristic function and its properties.

8.1 Expectation of a r. v.

Definition 8.1. *Expectation*:

Case (I)Expected value of a discrete $r.v X$ assuming values $x_1, x_2 \ldots x_{rt}$ and with $p.m.f\ P[X = x_t]$ is defined as

$$B(X) = \sum_{i=1}^{n} x_i p(x_i) \qquad (8.1)$$

Provided the sum is convergent.

Case (II) Expected value of a continuous $r.v X$ with $p.d.f. f(x)$ is defined as

$$E(X) = \int_{-\infty}^{\infty} xf(x)dx \qquad (8.2)$$

Provided the integral is convergent.

Expected value of a r.v is its average, simply called as a mean of r.v

Example 8.1. *Find expectation of* X *if*

(i) A $r.v\,X$ *assuming values* $0,1,2,\ldots\, n$ *with probability proportional to* ${}^{n}C_{x}$

(ii) $X \to B(n,p)$

(iii) $X \to \mathcal{P}(\lambda)$

(iv) X *be no. of tosses of a coin up to and including the first toss showing heads.*

Solution*:(*i*)A* $r.vX$ *assuming values* $0,1,2,\ldots\, n$ *with probability proportional to* ${}^{n}C_{x}$

$$P[X=x]=K(C_x)$$

$$K\sum_{x=0}^{n}(C_x)=K2^{n}=1$$

so, $K=2^{-n}$

By definition of expectation,

$$E(X)=\sum_{x=0}^{n}xp(x)$$

$$=2^{-n}\sum_{x=0}^{n}x\left({}^{n}C_{x}\right)=$$

As

$$\left({}^{n}C_{x}\right)\frac{n}{x}=\left({}^{n-1}C_{x-1}\right)$$

$$2^{-n}n\sum_{x=1}^{n}\left({}^{n-1}C_{x-1}\right)=2^{-n}n2^{n-1}=\frac{n}{2}$$

(ii) $X \to \mathcal{P}(\lambda)$

$$E(X)=\sum_{x=0}^{\infty}xp(x)=\sum_{x=0}^{\infty}x\frac{e^{-\lambda}\lambda^{x}}{x!}$$

$$\lambda\sum_{x=1}^{\infty}\frac{e^{-\lambda}\lambda^{x-1}}{(x-1)!}$$

Using

$$\left[1+\frac{\lambda}{1^{1}}+\frac{\lambda^{2}}{2^{1}}+\ldots\right]=e^{\lambda}$$

$$E(X)=\lambda e^{-\lambda}e^{\lambda}=\lambda.$$

(iii) Let X *be no. of tosses of a coin up to and including the first toss showing heads. Let* 'p' *be the chance of showing head.* $1-p=q$, *is chance of showing tail.*

$$P[X=x]=pq^{x-1}$$

for $x = 1, 2 \ldots$

$$E(X) = \sum_{x=1}^{\infty} xp(x) = \sum_{x=1}^{\infty} xpq^{x-1}$$

$$= p\left[1 + 2q + 3q^2 + \ldots\right] = \frac{p}{(1-q)^2} = \frac{1}{p}$$

Example 8.2. *Find expectation of* $r.v\,X$ *if*

(i) A $r.vX$ *assuming values* $(0,\infty)$ *with* $p.d.f f(x) = \lambda e^{-\lambda x}$.[*This is Exponential distribution with parameter* λ]

(ii) A $r.vX$ *assuming values* $(a,b), a < b$ *real numbers, with* $p.d.f f(x) =$ *constant*. [*This is Rectagular or Uniform distribution]*

(iii) $X \to N(\mu, \sigma^2)$

***Solution*:**(*i*) *Since* X *is absolutely continuous* $r.v$ *with density* $f(x)$

$$E(X) = \int_{-\infty}^{\infty} xf(x)\,dx \qquad (8.3)$$

$$E(X) = \int_{0}^{\infty} \lambda xe^{-\lambda x}dx = \qquad (8.4)$$

$$x\lambda \frac{-e^{-\lambda x}}{\lambda}\Big|_0^{\infty} - \lambda\int_0^{\infty} \frac{-e^{-\lambda x}}{\lambda}dx = \frac{1}{\lambda}$$

(ii) *Since density is constant over* (a,b),

$$\int_0^b Kdx = 1$$

gives $f(x) = K = \dfrac{1}{b-a}$

$$B(X) = \int_{\alpha}^{b} \frac{x}{b-a}dx = \frac{x^2}{2(b-a)}\Big|_a^b = \frac{b^2-a^2}{2(b-a)} = \frac{b+a}{2} \qquad (8.5)$$

(*iii*)

$$f(x) = \begin{cases} \dfrac{e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}}{\sqrt{2\pi\alpha}} - \infty \le x \le \infty; \mu \in \mathbb{R}, \sigma > 0 \end{cases} \qquad (8.6)$$

$$E(X) = \int_{-\infty}^{\infty} x\frac{e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}}{\sqrt{2\pi\sigma}}dx = \qquad (8.7)$$

Put $z = \dfrac{x-\mu}{\alpha}$ *then* $x = \sigma z + \mu$ *and* $dx = \sigma dz$

$$= \int_{-\infty}^{\infty} (\sigma z + \mu)\frac{e^{-\frac{1}{2}z^2}}{\sqrt{2\pi}}dz$$

$$= \alpha\int_{-\infty}^{\infty} z \frac{e^{-\frac{1}{2}z^2}}{\sqrt{2\pi}} dz + \mu\int_{-\infty}^{\infty} \frac{e^{-\frac{1}{2}z^2}}{\sqrt{2\pi}} dz \qquad (8.8)$$

$$= \sigma \times 0 + \mu = \mu . \qquad (8.9)$$

Since the first integral is an even function, and

$$\int_{-\infty}^{\infty} \frac{e^{-\frac{1}{2}z^2}}{\sqrt{2\pi}} dz = 1$$

Property 8.1. ***Properties of Expectation***

- *Expectation of a function of a* $r.v$: *If* $g(X)$ *is a monotonic function of a* $r.v$ *then expected value of* $g(X)$ *deoted by* $E(g(X))$ *is defined as Case* (*I*) *Discrete* $r.v$

$$E[g(X)] = \sum_{i=1}^{n} g(x_i)\, p(x_i) \qquad (8.10)$$

Provided the sum is convergent.

Case (*II*) *Continuous* $r.v$

$$E[g(X)] = \int_{-\infty}^{\infty} g(x) f(x) dx \qquad (8.11)$$

Provided the integral is convergent.

- *Expectation of a constant* : $E[C] = C$ *Where* C *is constant.*
- *Effect of change of origin* : $E[X + A] = E(X) + A$ *Where* A *is constant.*
- *Effect of change of Scale* : $Pi[AX] = APi(X)$ *Where* A *is constant.*
- *Linearity*: *Combining above two we may write* $E[AX + B] = AE(X) + B$ *Where* A, B *are constants*
- *Monotone If* $X \geq Y, E[X] \geq E[Y]$

Example 8.3. *A* $r.vX$ *has mean* -1, *find mean of following* $r.v.s$, *(i)* $-X$, *(ii)* $2X$ *(iii)* $\frac{X+3}{2}$, *(iv)* $\frac{2-X}{2}$

Solution: (i) $E(-X) = -E(X) = 1$ (ii) $E(2X) = 2E(X) = -2$

(*iii*)

$$E\left(\frac{X+3}{2}\right) = \frac{E(X)+3}{2} = 1$$

(*iv*)

$$E\left(\frac{2-X}{2}\right) = \frac{2-B(X)}{2} = 1.5$$

8.2 Variance of a r.v.

The function of r.v $g(X) = X^r$, has special role in the study of a r.v

Definition 8.2. r^{th} ***raw moment of a*** *r.v* : *r^{th} raw moment of a r.v X is defined as$E(X^r)$ and is denoted by $\mu_r^{'}$*

- We can check that for $'r = 1$ we get first raw moment and it is equal to mean $E(X) = \mu_1^{'}$
- r^{th} raw moment of a r.v X is also called as moment about zero
- Moments can also be defined around arbitrary origin, that is $E[X - A]^r$
- In particular if arbitrary origin is mean the moments are called as central moments So $E\left[X - \mu_1^{'}\right]^{\tau}$ is called as r^{th} central moment. Second central moment is important tool in the study of r.v . and it gives the idea about spread or scatterness of the values of the variable.

Example 8.4. *Show that first central moment* $= E\left(X - \mu_1^{'}\right) = 0$

***Solution*:**

$$E\left(X - \mu_1^{'}\right) = E(X) - \mu_1^{'} = 0$$

, *since* $E(X) = \mu_1^{'}$

Example 8.5. *Show that* $E(X - a)^2$ *is minimum when* $a = E(X)$, *hence variance is least mean square*

***Solution*:** *Consider*

$$\frac{dfij(X-a)^2}{da} = E\left(\frac{d}{da}(X-a)^2\right) = -2E(X-a) = 0$$

when $E(X) = a$, *provided* $\dfrac{d^2E(X-a)^2}{da^2} > 0$. *Thus* $E(X - E(X))^2 = V(X) =$ *mean square deviation about mean is minimum*

Definition 8.3. ***Variance of a*** *r.v* **.** : *Variance of a r.v is its second central moment*

- [Variance of a r.v] If X is a r.v then variance of X denoted by V(X) is defined as: **Case (I)** Discrete r.v :

$$V(X) = \sum_{i=1}^{n}\left[x_i - \mu_1^{'}\right]^2 p(x_i) \qquad (8.12)$$

Provided the sum is convergent.

Case (II) Continuous r.v . :

$$V(X) = \int_{-\infty}^{\infty}\left(X - \mu_1^{'}\right)^2 f(x)\,dx \qquad (8.13)$$

Provided thc integral is convergent.

- $V(X)=E(X^2)-[E(X)]^2$, for computational purpose we use this formula.
- Variance of a constant : $V[C]=0$ Where C is constant.
- Effect of change of origin : $V[X+A]=V(X)$ Where A is constant.
- Effect of change of Scalc : $V[AX]=A^2V(X)$ Where A is constant.
- Combining above two we may write $V[AX+B]=A^2V(X)$ Where A,B are constants.
- Positive square root of variance is called as standard deviation (s.d)of the r.v.

Example 8.6. *A r.vX has variance 4, find variance and s.d of following r.v.s (i)* $-X$, *(ii)* $2X$, *(iii)* $\frac{X+3}{2}$, *(iv)* $\frac{2-X}{2}$

Solution:

(i) $V(-X)=V(X)=4, s.d=2.$

(ii) $V(2x)=2^2V(X)=16, s.d=4$

(iii) $V\left(\frac{X+3}{2}\right)=\frac{V(X)}{2^2}=1$

(iv) $V\left(\frac{2-X}{2}\right)=\frac{V(X)}{2^2}=1, s.d=1$

Example 8.7. *Find variance of following r.v.s.*

(i) $X \rightarrow \mathcal{P}(\lambda)$

(ii) X *has Exponential distribution with parameter* λ

(iii) X *has Uniform* (a,b)

Solution*:(i)* $X \rightarrow \mathcal{P}(\lambda)$ *so , as shown in above exercise* $E(X)=\lambda$

Now consider,

$$E(X^2)=E(X(X-1))+E(X)$$

$$E(X(X-1))=\sum_{x=0}^{\infty}(x(x-1))p(x)=\sum_{x=1}^{\infty}x(x-1)\frac{e^{-\lambda}\lambda^x}{x!}$$

$$\lambda^2\sum_{x=2}^{\infty}\frac{e^{-\lambda}\lambda^{x-2}}{(x-2)^1}$$

Using

$$\left[1+\frac{\lambda}{1^1}+\frac{\lambda^2}{2^1}+\ldots\right]=e^{\lambda}$$

$$E\left(X\left(X-1\right)\right)=\lambda^2 e^{-\lambda}e^{\lambda}=\lambda^2$$

So,

$$E\left(X^2\right)=E\left(X\left(X-1\right)\right)+E\left(X\right)=\lambda^2+\lambda.$$

$$V\left(X\right)=E\left(X^2\right)-\left[E\left(X\right)\right]^2=\lambda^2+\lambda-\lambda^2=\lambda$$

Thus $X\to\mathcal{P}\left(\lambda\right)$, *then* $V\left(X\right)=\lambda$

(ii) X *has Exponential distribution with parameter* λ

$$E\left(X^2\right)=\int_0^{\infty}\lambda x^2 e^{-\lambda x}dx= \tag{8.14}$$

$$x^2\lambda\frac{-e^{-\lambda x}}{\lambda}\Big|_0^{\infty}-2\lambda\int_0^{\infty}x\frac{-e^{-\lambda x}}{\lambda}dx=\frac{2}{\lambda^2}$$

Since

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

$$V\left(X\right)=Pi\left(X^2\right)-\left[E\left(X\right)\right]^2=\frac{2}{\lambda^2}-\left(\frac{1}{\lambda}\right)^2=\frac{1}{\lambda^2}$$

(iii) Since X *has Uniform over* $\left(a,b\right)$, *as seen above,* $E\left(X\right)=\dfrac{b+a}{2}$

$$E\left(X^2\right)=\int_a^b\frac{x^2}{b-a}dx=\frac{x^3}{3\left(b-a\right)}\Big|_a^b=\frac{b^3-a^3}{2\left(b-a\right)}=\frac{b^2+ab+a^2}{3} \tag{8.15}$$

$$V\left(X\right)=E\left(X^2\right)-\left[H\left(X\right)\right]^2=\frac{b^2+ab+a^2}{3}-\left(\frac{b+a}{2}\right)^2=\frac{\left(b-a\right)^2}{12}$$

8.3 Characteristic function of a r.v.

A complex valued function of a r.v that is useful to study various properties of a r.v is known as characteristic function(ch.f).

Definition 8.4. ***Characteristic function*****:** *X be a r.v a complex valued function denoted by* $\Phi_X\left(t\right)$ *is defined as* $\Phi_X\left(t\right)=E\left(e^{ztX}\right)$ *where* $t\in\mathbb{R}$ *and* $i=\sqrt{-1}$

I) For discrete r.v : A discrete r.v X having p.m.f P_X, then its ch.f is given by

$$\Phi_X\left(t\right)=\sum_{x=0}^{\infty}P_X\left(x\right)e^{itx} \tag{8.16}$$

$t\in\mathbb{R}$ and $i=\sqrt{-1}$

II) For continuous r.v A continuous r.v X having p.d.f $f_X(x)$, then its ch. f is given by

$$\Phi_X(t) = \int_{-\infty}^{\infty} f_X(x) e^{itx} dx \qquad (8.17)$$

$t \in \mathbb{R}$ and $i = \sqrt{-1}$

- We can also write $\Phi_X(t) = \int_{-\infty}^{\infty} e^{itx} dF_X(x)$ which includes all r.v.
- $\Phi_X(t) = E(e^{itX}) = E(cos(tX) + iE(sin(tX))$
- $R_e(\Phi_X(t)) = E(cos(tX))$, which is real part of $\Phi_X(t)$ .And $I_m\Phi_X(t) = E(sin(tX))$ is the imaginary part of $\Phi_X(t)$

Example 8.8. *Find ch. f of following r.v.s*

(i) $X \to B(n, p)$

(ii) X has $p.m.fp(x) = pq^x x = 0,1,2\ldots$ $(Geometric\ distribution\ with\ parameter\ p)$

(iii) $X \to N(\mu, \sigma^2)$

Solution:(i)

$$\Phi_X(t) = \sum_{x=0}^{n} (C_x) p^x q^{(n-x)} e^{itx} \qquad (8.18)$$

$t \in \mathbb{R}$ and $i = \sqrt{-1}$

$$= \Phi_X(t) = \sum_{x=0}^{n} (C_x) q^{(n-x)} (pe^{it})^x = (q + pe^{it})^n \qquad (8.19)$$

$Using\ Binomial\ expansion.\ Ch.f\ \ of\ X \to B(n, p)\ is\ \Phi_X(t) = (q + pe^{it})^n$ (ii)

$$\Phi_X(t) = \sum_{x=0}^{\infty} pq^x e^{itx} = \frac{p}{1 - qe^{it}} \qquad (8.20)$$

$Using\ geometric\ series\ with\ common\ ratio\ qe^{i}t < 1$

$So,\ Ch.f\ \ of\ Geometric\ distribution\ with\ parameter\ p\ is\ \Phi_X(t) = \dfrac{p}{1 - qe^{t}}$ (iii)

$$\Phi_X(t) = \int_{-\infty}^{\infty} e^{itx} \frac{e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}}{\sqrt{2\pi\sigma}} dx = \tag{8.21}$$

Put $z = \dfrac{x-\mu}{\sigma}$ *then* $x = \sigma z + \mu$ *and* $dx = \sigma dz$

$$= \int_{-\infty}^{\infty} e^{it(\sigma z+\mu)} \frac{e^{-\frac{1}{2}z^2}}{\sqrt{2\pi}} dz$$

$$= e^{it\mu - \frac{1}{2}\alpha^2 t^2} \int_{-\infty}^{\infty} \frac{e^{-\frac{1}{2}(z-it\alpha)^2}}{\sqrt{2\pi}} dz \tag{8.22}$$

$$= e^{it\mu - \frac{1}{2}\sigma^2 t^2} \tag{8.23}$$

Since

$$\int_{-\infty}^{\infty} \frac{e^{-\frac{1}{2}(z-it\sigma)^2}}{\sqrt{2\pi}} dz = 1$$

$X \to N(\mu, \sigma^2)$ *then its ch.* $f \Phi_X(t) = e^{it\mu - \frac{1}{2}\sigma^2 t^2}$

Property 8.2. *Properties of ch. f*

- $\Phi_X(0) = 1.$
- $|\Phi_X(t)| \le 1$ *for all* $t \in \mathbb{R}.$
- $\Phi_X(-t) = \Phi_{-X}(t) = \overline{\Phi_X(t)}$ *which is complex conjugate of* $\Phi_X(t)$.
- $\Phi_X(t)$ *is uniformly continuous function of t.*
- $\Phi_{aX+b}(t) = e^{ibl}\Phi_X(at).$
- $\Phi_X(t)$ *generates moments of r.v X. The coefficient of* $(it)^r$ *in the expansion of* Φ_X *is* r^{th} *moment of X. We can also get it from* r^{th} *derivative of* Φ_X .

$$\mu_r' = \frac{1}{i^r}\frac{d^r}{dt^r}\Phi_X(0) \tag{8.24}$$

x	–2	0	2
$p(x)$	- 41	- 21	- 41

- *If X and Y are independent r.vsch.f of* $X+Y$ *is equal to product of their ch. fs*
- *Product of any two ch. fs is also a ch, f. Thus any power of* $\Phi_X(t)$ *is also a ch, f.*

Example 8.9. *Find ch. f of X which is Uniform r.v over $(0,1)$. Hence that of - X. Solution* : $f(x)=1$ *for* $0<x<1$

$$\Phi_X(t)=\int_0^1 e^{itx}dx=\frac{e^{it}-1}{it}$$

$$\Phi_{-X}(t)=\frac{e^{-it}-1}{it}$$

Example 8.10. *Using above result find ch. f of X - Y if X,Y are i.i.d Uniform* $(0,1)$.

Solution: $\Phi_X(t)=\Phi_Y(t)=\dfrac{e^{it}-1}{it}$

$$\Phi_{-Y}(t)=\frac{e^{-it}-1}{it}$$

X and Y are independent, X , - Y are also independent, by property of ch. f

$$\Phi_{X-Y}(t)=\Phi_X(t)\Phi_{-Y}(t)\frac{e^{-it}-1}{it}\frac{e^{it}-1}{it}=\frac{e^{it}+e^{-it}-2}{t^2}$$

Example 8.11. *Is $cos^2(t)$ a ch. f?*

Solution:

$$cost=\frac{e^{il}+e^{-il}}{2}\ \ So,\ \ cos^2(t)=\frac{e^{2il}+e^{-2il}+2}{4}=E\left(e^{itX}\right)$$

Where X has p.m.f .

X	-2	0	2
p(x)	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

$cos^2(t)$ is a ch. f.

Theorem 8.1. *$X_1,X_2\ldots X_n$ be the r.vs with d . fs $F_{X_1},F_{X_2}\ldots F_{X_n}$ respectively and ch. fs $\Phi_{X_1},\Phi_{X_9}\ldots\Phi_{X_n}$ respectively, then for the constants $a_1,a_2\ldots a_n$ such that $a_i\geq 0$, and $\sum a_i=1,\sum \Phi_{X_{\dot{x}}}a_i$ is a ch. f of $\sum a_iF_{X_x}$*

Following theorem characterizes the ch. f and its density or d.f. Uniquely.

Theorem 8.2. ***Inversion theorem***: *X is absolutely continuous r.v such that*

$$\int_{-\infty}^{\infty} |\Phi_X(t)|\, dt < \infty$$

then its p.d.f is given by

$$f_X(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx}\Phi_X(t)\,dt \qquad (8.25)$$

Example 8.12. *Show that $e^{\Phi(t)-1}\dot{w}$ a ch.f . if $\Phi(t)$ is ch.f of some d.f $F(x)$*

Solution:

$$e^{\Phi(t)-1} = \frac{1+\Phi(t)+\frac{\Phi(t)^2}{2!}+\frac{\Phi(t)^3}{3!}+}{e}$$

$$= \sum_{j=1}^{\infty} a_j \Phi_j(t)$$

where $a\,a_j = \frac{1}{j!e}$ *with* $\sum_{j=1}^{\infty} a_1 = 1$ *and* $\Phi_j(t) = \Phi^j(t)$, *ch. f . of j -fold convolution of* $F'(x)$ *. hence* $e^{\Phi(t)-1}$ *is a ch.f*

Example 8.13. $f(x) = \frac{e^{-|x|}}{2}$ *for* $-\infty < x < \infty$ *then find its ch. f.*

Solution:

$$\Phi_X(t) = \int_{-\infty}^{\infty} e^{itx}\frac{e^{-|x|}}{2}dx =$$

$$= \int_{-\infty}^{0} e^{itx}\frac{e^{x}}{2}dx + \int_{0}^{\infty} e^{itx}\frac{e^{-x}}{2}dx = \frac{1}{2}\left(\frac{1}{1-it}+\frac{1}{1+it}\right) = \frac{1}{1+t^2}$$

$\Phi_X(t) = e^{-|t|}$ *find its p.d.f*

Solution: *Since from the above example for* $f(x) = e^{-}|x|$ *ch f is* $\frac{1}{1+t^2}$, *we write*

$$\frac{1}{1+t^2} = \int_{-\infty}^{\infty} e^{itx}\frac{e^{-|x|}}{2}dx$$

replace t by - y in above equation gives

$$\frac{1}{1+y^2} = \int_{-\infty}^{\infty} e^{iyx}\frac{e^{-|x|}}{2}dx$$

replace x by - t in above equation and multiply by $\frac{1}{\pi}$ *gives*

$$\frac{1}{\pi}\frac{1}{1+y^2}=\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{-iyt}e^{-|t|}dt$$

By inversion theorem, $e^{-|t|}$ *is a ch. f of r.v whose density is*

$f(y)=\frac{1}{\pi}\frac{1}{1+y^2}$ *for* $-\infty<x<\infty$ *This is Cauchy density.*

8.4 Chapter End Exercises

1. Find expectation and variance of X if

   (i) $X \to B(n,p)$

   (ii) X has Uniform over $(-1,1)$

   (iii) p.d.f of X is

   $$f(x)=\begin{cases}1+x & \text{for} -1<x\le 0\\ 1-x & \text{for} 0<x\le 1\\ 0 & \text{Otherwise}\end{cases} \qquad (8.26)$$

   (iv) X has p.m.f $p(x)=pq^x x=0,1,2\ldots$

2. Find ch. f of (i)X has Poisson with parameter λ. (ii) X_1-X_2, where $X_1 \to N(\mu_1,\sigma_{1^2})$ and $X_2 \to N(\mu_2,\sigma_{2^2})$ are independent r.v.s.

3. Are following ch. f? (i) cost (ii) $R_e\Phi_X(t)$ (iii) $P(\Phi_X(t))=\sum \Phi_X(t)^k p_k$.

   (iv) $\frac{1}{1+t}$

❖❖❖

9

THEOREMS ON EXPECTATION AND CONDITIONAL EXPECTATION

Unit Structure

9.0 Objectives

9.1 Expectation of a function of two dimensional r.v.s

9.2 Conditional Expectation

9.3 Chapter End Exercises

9.0 Objectives

After going through this chapter you will learn

- Expectation of a function of two dimensional r.v
- Theorems on expectation.
- Some inequalities based on expectations.
- Conditional expectation and its relation with simple expectation.

9.1 Expectation of a function of two dimensional r.v.s

Definition 9.1. ***Expectation of*** $g(X,\mathrm{Y})$: *Let the function of two dimensional r.vs* (X,Y) *be* $g(X,Y)$ *its expected value denoted by* $E\left[g(X,Y)\right]$ *is defined as*

Case (*I*)*Discrete r v . with joint p.m.f* $P_{X,\mathrm{y}}\left(x_i, y_j\right)$

$$Pi\left[g(X,Y)\right]=\sum_{j=1}^{m}\sum_{i=1}^{n}g\left(x_i,y_j\right)p\left(x_i,y_j\right) \qquad (9.1)$$

Provided the sum is convergent. Case (*II*)*Continuous r v . with joint* $p.d.f.f_{X,\mathrm{y}}(x,y)$

$$Pi\left[g(X,Y)\right]=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}g(x,y)\,f(x,y)\,dxdy \qquad (9.2)$$

Provided the integral is convergent.

Theorem 9.1. *Addition Theorem on expectation:* (X,Y) *be two dimensional r.vs then*

$$E(X+Y)=E(X)+E(Y) \qquad (9.3)$$

Proof: We assume that the r.v.s (*X*, Y) are continuous with joint p.d.f. $f_{X,Y}(x,y)$ and marginal p.d.f.s $g_X(x), h_Y(y)$.

$$E(X+\mathrm{y})=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x+y)f_{X,Y}(x,y)dxdy$$

$$=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xf_{X,\mathrm{y}}(x,y)dxdy+\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}yf_{X,Y}(x,y)dxdy$$

$$=\int_{-\infty}^{\infty}x\left[\int_{-\infty}^{\infty}f_{X,Y}(x,y)dy\right]dx+\int_{-\infty}^{\infty}y\left[\int_{-\infty}^{\infty}f_{X,Y}(x,y)dx\right]dy$$

$$=\int_{-\infty}^{\infty}xg_X(x)dx+\int_{-\infty}^{\infty}yh_Y(y)dy\Gamma\text{rom the definition } g_X(x),h_Y(y)$$

$$= Pi(X)+Pi(Y) \text{ From the definition of } \mathrm{E}[\mathrm{X}] \text{ and } \mathrm{E}[\mathrm{Y}]$$

Hence the proof.

Theorem 9.2. ***Multiplication Theorem on expectation*:** (X,Y) *be two dimensional independent r.vs then*

$$E(XY)=E(X)E(Y) \qquad (9.4)$$

Proof: We assume that the r.v.s(X,Y) are continuous with joint p.d.f. $f_{X,Y}(x,y)$ and marginal p.d.f.s $g_X(x), h_Y(y)$.

$$E(XY)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xyf_{X,\mathrm{y}}(x,y)dxdy$$

$$=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xyf_{X,Y}(x,y)dxdy$$

$$=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xyg_X(x)h_Y(y)dxdy \text{ By independence of X and Y}$$

$$=\int_{-\infty}^{\infty}xg_X(x)\int_{-\infty}^{\infty}yh_{\mathrm{y}}(y)dxdy$$

$$= E(X)Pi(Y)\Gamma\text{rom the definition of } E[\mathrm{X}] \text{ and } \mathrm{E}[\mathrm{Y}]$$

Hence the proof.

Above theorems can be generalized for n variables $X_1, X_2, \ldots X_n$

$$E(X_1 + X_2 + \ldots X_n) = E(X_1) + E(X_2) + \ldots E(X_n) \qquad (9.5)$$

Γorn independent random variables $X_1, X_2, \ldots X_n$

$$E(X_1 X_2 \ldots X_n) = E(X_1) E(X_2) \ldots E(X_n) \qquad (9.6)$$

Example 9.1. *Show that* $E(aX + bY + c) = aE(X) + bE(Y) + c$

Solution: *Consider continuous* $r.v.(X,Y)$ *with joint* $p.d.f.f_{X,Y}(x,y)$, *and marginal* $p.d.f.s$

$g_X(x), h_Y(y)$

$Pi[aX + bY + c)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(ax + by + c) f(x,y) dxdy$ (9.7)

$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} axf_{X,Y}(x,y)dxdy + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} by \ f_{X,y}(x,y)dxdy + c\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)dxdy$

$= a\int_{-\infty}^{\infty} x\left[\int_{-\infty}^{\infty} f_{X,y}(x,y)dy\right]dx + b\int_{-\infty}^{\infty} y\left[\int_{-\infty}^{\infty} f_{X,Y}(x,y)dx\right]dy + cBy$ *property of joint*

$p.d.f = a\int_{-\infty}^{\infty} xg_X(x)dx + b\int_{-\infty}^{\infty} yh_Y(y)dy + c$ *By the definition* $9x(x),\ h_Y(y)$

$= aE(X) + bE(Y) + c$ *By the definition of* $\sqrt{\ } l^X 1$ *and* $E[Y]$

If $X \geq Y$, *then prove that* $E[X] \geq E[Y]$

Solution: If $X \geq \mathrm{y}$, *then* $[X - Y] \geq 0$, *and hence* $E[X - Y] \geq 0$

From the property of expectation and addition theorem,

$E[X] - E[Y] \geq 0$ *or* $E[X] \geq E[Y]$.

Example 9.2. *Prove for any two* $r.v.sX, Y$,

$$E[X\mathrm{y}]^2 \leq E\left[X^2\right] E\left[Y^2\right]$$

[*This is Cauchy Schwarz' s inequality.1*

Solution: *consider a function* $h(a) = E[X - aY]^2$

$$h(a) = E\left[x^2\right] + a^2 E\left[Y^2\right] - 2aE[XY]$$

$$\frac{dh(a)}{da} = 2aE\left[Y^2\right] - 2E[XY] = 0$$

gives

$$a = \frac{pj[XY]}{E\left[Y^2\right]}$$

And

$$\frac{d^2h(a)}{da^2} = 2E\left[Y^2\right] \geq 0$$

Thus $h(a)$ *is minimum vjhen* $a = \dfrac{E[XY]}{E\left[Y^2\right]}$

$$h(a) \geq E\left[X - \frac{E[XY]}{E\left[Y^2\right]}Y\right]^2 = fij\left[X^2\right] + \frac{E[XY]}{E\left[Y^2\right]}E\left[Y^2\right]2 - 2\frac{E[XY]}{E\left[Y^2\right]}\, fii\ [XY]$$

$$E\left[X^2\right] - \frac{E[XY]^2}{pj\left[Y^2\right]} \geq 0$$

gives

$$E[XY]^2 \leq E\left[X^2\right]E\left[Y^2\right]$$

Example 9.3. *Show that*

$$E\left(\frac{1}{X^2}\right) \geq \frac{1}{E\left[X^2\right]}$$

Solution**:** *Using Cauchy Schwarz's inequality for* $\mathrm{y} = \dfrac{1}{X}$

$$1 \leq E\left[X^2\right]E\left[\frac{1}{X^2}\right]$$

divide inequality by $E\left[X^2\right] > 0$ *to get*,

$$B\left(\frac{1}{X^2}\right) \geq \frac{1}{E\left[X^2\right]}$$

Example 9.4. *Show that for any two* $r.v.sX, Y$,

$$\sqrt{E[X+Y]^2} \leq \sqrt{E\left[X^2\right]} + \sqrt{E\left[Y^2\right]}$$

Solution**:** *Consider*,

$$E[X+Y]^2 = E\left[X^2 + 2X\mathrm{y} + Y^2\right]$$

$$= E\left[X^2\right] + 2E[XY] + E\left[Y^2\right]$$

$$\leq E\left[X^2\right] + 2\sqrt{E\left[X^2\right]E\left[Y^2\right]} + E\left[Y^2\right] \textit{By Cauchy Schwarz's inequality}$$

$$= \left[\sqrt{E\left[X^2\right]} + \sqrt{E\left[Y^2\right]}\right]^2$$

Taking square root,

$$\sqrt{pj[X+Y]^2} \leq \sqrt{B\left[X^2\right]} + \sqrt{pj\left[Y^2\right]}$$

9.2 Conditional Expectation

Definition 9.2. ***Conditional Expectation of*** X :

Case (*I*) *Let* (X,Y) *be the two dimensional discrete* $r.v with / oint$ *probability mass function* $P_{X,Y}(x_i, y_j) i = 1\ ton,\ j = 1 tom$ *the conditional expected value of* X *given* $Y = y_j$ *denoted by* $E(X / \mathrm{y} = y_j)$ *and defined as*

$$E(X / Y = y_j) = \sum_{i=1}^{n} x_i P(X = x_i / Y = y_j) \qquad (9.8)$$

Case (*II*) *Let* (X,Y) *be the two dimensional continuous* $r.v$ *with joint probability density function* $f_{X,Y}(x,y) x, y \in \mathbb{R}^2$ *the conditional expected value of* X *given* $Y = y$ *denoted by* $E(X / \mathrm{y} = y)$ *and defined as*

$$E(X / Y = y) = \int_{-\infty}^{\infty} xg(X / Y = y) dx \qquad (9.9)$$

Similarly using conditional p.m.f s or p.d.f s of Y given X we can define conditional expected value of Y.

Definition 9.3. ***Conditional Expectation of*** Y **:** *Case* (*I*) *Let* (X,Y) *be the two dimensional discrete* $r.v$ *with joint probability mass function* $P_{X,Y}(x_i, y_j) i = 1 ton,$ $j = 1 tom$ *the conditional expected value of* Y *given* $X = x_i$ *denoted by* $Pi(Y / X = x_i)$ *and defined as*

$$E(Y / X = x_i) = \sum_{j=1}^{m} y_j P(Y = y_j / X = x_i) \qquad (9.10)$$

Case (*II*) *Let* (X,Y) *be the two dimensional continuous* $r.v$ *with joint probability density function* $f_{X,Y}(x,y) x, y \in \mathbb{R}^2$ *the conditional expected value of* X *given* $Y = y$ *denoted by* $E(Y / X = x)$ *and defined as*

$$E(Y / X = x) = \int_{-\infty}^{\infty} yh(Y / X = x) dy \qquad (9.11)$$

Theorem 9.3. *For any two* $r.v.sX$ *and* Y

$E_y E(X / Y = y) = E(X)$

and $E_x E(Y / X = x) = E(Y)$

Proof: Consider

$$E\left(X/\mathrm{y}=y\right)=\int_{-\infty}^{\infty}xg\left(X/y\right)dx$$

Multiply both sides of above equation by $h\left(y\right)$ and itegrate with respect to y,to get

$$\int_{-\infty}^{\infty}h\left(y\right)Pi\left(X/y\right)dy=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xh\left(y\right)g\left(X/y\right)dxdy$$

L.H.S becomes $E_yE\left(X/y\right)=E\left(X\right)$ and since $h\left(y\right)g\left(X/y\right)=f_{X,\mathrm{y}}\left(x,y\right)$ R.H.S becomes

$$=\int_{-\infty}^{\infty}x\left[\int_{-\infty}^{\infty}f_{X,Y}\left(x,y\right)dy\right]dx$$

$$=\int_{-\infty}^{\infty}xg_X\left(x\right)dx \text{ By definition of } g_X\left(x\right) \qquad (9.12)$$

$$=Pi\left(X\right) \qquad (9.13)$$

Hence the proof. We can similarly prove $E_xE\left(Y/X=x\right)=E\left(Y\right)$

Example 9.5. *Find conditional means of* X *and* Y *of the following r.v. The joint* $p.d.f$ *of* $f_{X,Y}\left(x,y\right)=8xy$ *for* $0<x<y<1$

Solution: *From* $\left(13.8\right)$ *the conditional* $p,d.f$ *of X givenY = yis*

$$g_{X/Y=y}\left(x\right)=\frac{2x}{y^2} \text{ for } 0<x<y$$

$$E\left(X/y\right)=\int_0^y x\frac{2x}{y^2}dx=\frac{2}{3}y$$

From (13.9) *the conditional* $p.d.f$ *of Ygiven* $X=x$*is*

$$h_{Y/X=x}\left(y\right)=\frac{2y}{1-x^2} \text{ for } x<y<1$$

$$E\left(Y/x\right)=\int_x^1 y\frac{2y}{1-x^2}dy=\frac{2\left(1+x+x^2\right)}{3\left(1+x\right)}$$

Example 9.6. *Find* $E\left(X/Y=2\right)$, *and* $E\left(Y/X=2\right)$ *if the joint* $p.m.f$. *of* $r.v.\left(X,Y\right)$ *is as given below.*

$Y\downarrow\ X\rightarrow$	1	2	$P_Y(y)$
1	$\frac{1}{4}$	0	$\frac{1}{4}$
2	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{3}{4}$
$P_X(x)$	$\frac{3}{4}$	$\frac{1}{4}$	1

Solution: *The conditional* $p.m.f$ *of* Y *given* $X=2$ *is given by*

Y	1	2
$P_{X/Y=2}(y)$	0	1 - 31

$E(Y/X=2)=2:yP(Y=y/X=2)=2$

The conditional p.m.f of X given $Y=2$ is given by

x	1	2
$P_{X/Y=2}(x)$	$\frac{2}{3}$	$\frac{1}{3}$

$$E(X/Y=2)=\sum xP(X=x/Y=2)=\frac{4}{3}$$

9.3 Chapter End Exercises

1. If X and Y are independent then show that conditional means are same same as simple means.

2. Show that $E|X-\mu| \le \sqrt{E(X-\mu)^2}$

3. Find conditional means of X and Y of the following r.v. The joint p.d.f of $f_{X,Y}(x,y)=2$ for $0<x<y<1$

4. Given the joint p.m.f of X,Y as

   $P[X=x,Y=y]=\frac{x+3y}{24}$ for $x,y=1,2$. Find conditional mean of Y given$X$ $=1$ and conditional mean of X given $Y=1$

5. For the above problem 3 and 4,find $\mathrm{Ei}(\mathrm{XY}),\mathrm{E}(\mathrm{X}+\mathrm{Y})$.

6. Two balls are drawn from an urn containing one yel of blue balls drawn.low,two red and three blue balls. Let X be the number of red balls, and Y be the number of blue balls drawn. Γind the joint p.m.f of X, Y ,hence $E(X/Y=2)$ and $E(Y/X=2)$.

❖❖❖

10

LAW OF LARGE NUMBERS AND CENTRAL LIMIT THEOREM

Unit Structure

10.0 Objectives

After going through this chapter you will learn

- Chebyshev' s inequality and its applications
- Various modes of convergence and their interrelations.
- Weak law of large numbersand necessary condition for asequence to obey this law.
- Strong law of large numbersand necessary condition for asequence to obey this law.
- CLT : An important theorem for finding probabilities using Normal approximation.

10.1 Chebyshev's inequality

In this chapter we will study asymptotic behavior of the sequence of r.v.s.

Theorem 10.1. *X is a non - negative r.v with finite mean then for any $C > 0$*

$$P[X \geq C] \leq \frac{E(X)}{C} \qquad (10.1)$$

proof :

$$E[X]=\int_{-\infty}^{\infty} x f_X(x)\,dx \tag{10.2}$$

Consider $\Omega=\{X \geq C\} \cup \{X < C\}$ So,

$$E[X]=\int_{X\geq C} x f_X(x)\,dx+\int_{X<C} x f_X(x)\,dx \tag{10.3}$$

$$\geq \int_{X\geq C} x f_X(x)\,dx$$

$$\geq \int_{X\geq C} C f_X(x)$$

$$= CP[X\geq C]$$

Hence we get

$$P[X\geq C]\leq \frac{B(X)}{C} \tag{10.4}$$

Following inequality is directly followed from above theorem

Theorem 10.2. ***Chebyshev*'s *inequality*** *X is a r.v with mean μ and variance $=\sigma^2$ then for any $C>0$*

$$P\left[|X-\mu|\geq C\right]\leq \frac{\sigma^2}{C^2}$$

proof: $|X-\mu|\geq C$ implies $(X-\mu)^2\geq C^2$

So

$$E\left[|X-\mu|\geq C\right]\leq P\left[(X-\mu)^2\geq C^2\right]\leq \frac{E((X-\mu)^2}{C^2}$$

Using above theorem for $(X-\mu)^2$

$$=\frac{V(X)}{C^2}$$

Hence

$$P\left[|X-\mu|\geq C\right]\leq \frac{\sigma^2}{C^2} \tag{10.5}$$

- Inequality can also be written as

$$P[|X-\mu|<C]\geq 1-\frac{\sigma^2}{C^2} \tag{10.6}$$

we get an lower bound on the probability that r.v deviates from its mean by C

- If C is replaced by $k_{(\gamma}$, where $k > 0$ then inequality reduces to give an upper bound

$$P\left[\left|X-\mu\right| \geq k_{Cf}\right] \leq \frac{1}{k^2} \qquad (10.7)$$

- By complementation can also write a lower bound.

$$P[\left|X-\mu\right| < k\sigma] \geq 1-\frac{1}{k^2} \qquad (10.8)$$

- If $k=2$, lower bound is $\frac{3}{4}$, which means that 75% of the times r.v assumes values in $(\mu-2\alpha, \mu+2\alpha)$
- The bounds given by Chebyshev's inequality are theoretical and not practical, in the sense that the bounds are rarely attained by the r.v.
- Inequality is useful when the information regarding probability distribution of r.v is not available but the mean and variance is known.

Example 10.1. *A r.vX has mean* 40 *and variance* 12. *Find the bound for the probability* $P\left[X \leq 32\right]+P\left[X \geq 48\right]$.

***Solution*:**

$$P\left[X \leq 32\right]+P\left[X \geq 48\right]=P\left[\left|X-40\right| \geq 8\right]$$

By Chebyshev's *inequality*,

$$P\left[\left|X-40\right| \geq 8\right] \leq \frac{variance}{8^2}=0.1875$$

Example 10.2. *A unbiased coin is tossed* 400 *times,find the probability that number of heads lie between* (160, 240).

***Solution*:** *X has* $B\left(400,\frac{1}{2}\right)$. *X has Mean 200 and Variance 100.*

So, $P\left[160 \leq X \leq 240\right]=P\left[\left|X-200\right| \leq 40\right]$ *,By Chebyshev' s inequality, with* $k=4$ *and* $\sigma=10$

$P\left[\left|X-200\right| \leq 40\right] \geq 1-\frac{1}{16}$. *This gives the lower bound* 0.9375

10.2 Modes of Convergence

Th modes of convergence are itroduced so as to define further laws of large numbers.

Definition 10.1. ***Convergence in Probability*:** (Ω, A, P) *be a probability space* $\{X_n\}$ *is a sequence of a r.v.s* X_n *is said to converge in probability to a r.vX*, *from the same space, if for any* $\varepsilon > 0$

$$\lim_{n\to\infty} P(|X_n - X|\varepsilon) = 0 \quad (10.9)$$

We say that $X_n \xrightarrow{P} X$

Definition 10.2. ***Almost Sure Convergence*** (Ω, A, P) *be a probability space* $\{X_n\}$ *is a sequence of a r.v.s* X_n *is said to converge almost surely to a r.vX*, *from the same space, if for any* $\varepsilon > 0$

$$P(\lim_{n\to\infty} |X_n - X| > \varepsilon) = 0 \qquad (10.10)$$

We say that $X_n \xrightarrow{a.s} X$

Definition 10.3. ***Convergence in Distribution*:** (Ω, A, P) *be a probability space* $\{X_n\}$ *is a sequence of a r.v.s with d.f.s* $\{F_n\}$. X_n *is said to converge in distribution to a r.vX*, *from the same space, if there exists a d.fF of* X *such that* F_n *converges to* F *at all continuity points of* F.

- Almost sure convergence implies convergence in probability.
- Convergence in probability implies convergence in distribution.

10.3 Laws of large numbers

Theorem 10.3. ***Weak law of Large Numbers*(*WLLN*):** $X_1, X_2 \ldots X_n$ *be the independent r.vs with means* $\mu_1, \mu_2 \ldots \mu_n$ *respectively and finite variances* $\sigma_1^2, \sigma_2^2 \ldots \sigma_n^2$ *respectively and* $S_n = \sum X_i$ *for any* $\varepsilon > 0$, *if*

$$\lim_{n\to\infty} P\left(\left|\frac{S_n}{n} - \frac{\sum_{i=1}^{n}\mu_i}{n}\right| > \varepsilon\right) = 0 \qquad (10.11)$$

We say the Weak Law of Large Numbers (*WLLN*) *holds for the sequence of r.v.s* $\{X_i\}$

proof: Consider

$$P(\left|\frac{S_n}{n} - \frac{\sum_i^n = 1\,\mu_i}{n}\right| > \varepsilon) = P(\left|\frac{S_n}{n} - E\left(\frac{S_n}{n}\right)\right| > \varepsilon)$$

$$\leq \frac{Var\left(\frac{S_n}{n}\right)}{\varepsilon^2}$$

Because of independence of r.vs and by Chebyshev's Inequality .Further

$$= \frac{\sum \sigma_i^2}{n^2\varepsilon^2} \tag{10.12}$$

Taking limit as n tends to ∞ of both sides of inequality we get R.H. S limit zero since variances are finite, finally

$$\lim_{n\to\infty} P(\left|\frac{S_n}{n} - \frac{\sum \mu_i}{n}\right| > \varepsilon) = 0 \tag{10.13}$$

Theorem 10.4. ***Khintchine's Weak law of Large Numbers(WLLN)*** $X_1, X_2 \ldots X_n$

be the i.i.dr.vs with common mean μ *, then* $S_n = \sum X_i$ *for any* $\varepsilon > 0$ *, if*

$$\lim_{n\to\infty} P(\left|\frac{S_n}{n} - \mu\right| > \varepsilon) = 0 \tag{10.14}$$

We say the Weak Law of Large Numbers (WLLN) holds for the sequence of r.v . s $\{X_i\}$

The law can be equivalently stated using complementation as

$$\lim_{n\to\infty} P(\left|\frac{S_n}{n} - \mu\right| < \varepsilon) = 1 \tag{10.15}$$

- In short when WLLN holds for thc sequence if $S_n \to P\sum \mu_i$.
- The limiting value of chance that average values of the r.v.s becomes close to the mean is nearly 1,as n approaches to ∞
- Assumption of finite variance is required for non identically distributed r.v.s.The condition for WLLN to hold for such sequence is that

  $$\frac{V(S_{rt})}{n^2}$$

  tends to zero as n approaches to infinity. Γor i.i.d.r.vs only existence of finite mean is required.
- Above law is a weak law in the sense that there is another law which implies this law

Example 10.3. *Examine whether WLLN holds for the following sequence of the independent r.v.s.*

1.

$$X_k = \begin{cases} -\sqrt{2k-1} & withprob = \frac{1}{2} \\ \sqrt{2k-1} & withprob = \frac{1}{2} \end{cases}$$

Solution: $E\left[X_k\right] = 0$ *and* $V\left[X_k\right] = Zk - 1, V\left(S_n\right) = n^2$

$\frac{V\left(S_n\right)}{n^2}$ *does not tend to zero as* n *approaches to* ∞ . *WLLN does not hold for the sequence*

2.

$$X_k = \begin{cases} \pm 2^k & withprob = \frac{1}{2^{2k+1}} \\ 0 & withprob = 1 - \frac{1}{2^{2k}} \end{cases}$$

Solution: $E\left[X_k\right] = 0$ *and* $V\left[X_k\right] = 1, V\left(S_n\right) = n$

$\frac{V\left(S_{\mathrm{n}}\right)}{n^2}$ *tends to zero as* n *approaches to* ∞, *therefore WLLN holds for the sequence*

Theorem 10.5. ***Strong Law of Large Numbers*(*SLLN*):** $X_1, X_2 \ldots X_n$ *be the independent r.vs with means* $\mu_1, \mu_2 \ldots \mu_{7l}$ *respectively and finite variances* $\sigma_1^2, \sigma_2^2 \ldots \sigma_n^2$ *respectively and* $S_n = \sum X_i$ *for any* $\varepsilon > 0$, *if*

$$P[\lim_{n\to\infty} | \frac{S_n}{n} - \frac{\sum_{i=1}^{n} \mu_i}{n} | > \varepsilon] = 0 \tag{10.16}$$

We say that Strong Law of Large Numbers (*SLLN*) *holds for the sequence of* $r.v.s\left\{X_i\right\}$

- In short when SLLN holds for the sequence if $S_n \to as \sum \mu_i$.
- The average values of the r.v.s becomes close to the mean as n approaches to ∞ with very high probability.That is almost surely.
- Assumption of finite variance is required for non identically distributed r.v.s. The condition for SI͵LN to hold for such sequence is that

$$\sum_{i=1}^{\infty} \frac{V(X_i)}{i^2} < \infty$$

as n approaches to infinity.This condition is known as Kolmogorov's Condition. For i.i.d.r.v.s only existence of finite mean is required.

- Above law is a Strong law in the sense that which implies Weak law

Example 10.4. *Examine whether SLLN holds for the following sequence of the independent r.v.s.*

1.

$$X_k = \begin{cases} \pm k & withprob = \dfrac{1}{2\sqrt{k}} \\ 0 & withprob = 1 - \dfrac{1}{\sqrt{k}} \end{cases}$$

Solution: $E[X_k] = 0$ *and* $V[X_k] = k^{\frac{3}{2}}, \sum \frac{V(X_k)}{k^2} 4\infty$ *SLLN does not hold for the sequence*

2.

$$X_k = \begin{cases} -2^k & withprob = \dfrac{1}{2} \\ 2^k & withprob = \dfrac{1}{2} \end{cases}$$

Solution: $E[X_k] = 0$ *and* $V[X_k] = 2^{2k}, \sum \frac{V(X_k)}{k^2} < \infty$ *SLLN holds for the sequence*

Weak law of large numbers gives an idea about whether the difference between average value of r.v.s and their mean becomes small But following theorem gives the limiting probability of this becomes less than small number ε

10.4 Central Limit Theorem

Central Limit Theorem is basically used to find the approximate probabilities, using Normal distribution.The theorem was initially proved for bernoulli r.v.s It has been proved by many mathematicians and statisticians, by imposing different conditions.

Theorem 10.6. ***Central Limit Theorem by Lindberg - Léévy(CLT):***

$X_1, X_2 \ldots X_n$ *be the i.i.dr.vs with common mean* μ *and common variance* σ^2, *let* $S_n = \sum X_i$ *for any* $a \in \mathbb{R} > 0$,

$$\lim_{n\to\infty} P\left(\left|\frac{\frac{S_n}{n}-\mu}{\frac{\alpha}{\sqrt{n}}}\right| < a\right) = \frac{1}{\sqrt{2_{71}}}\int_{-a}^{a} e^{-\frac{1}{2}x^2}\,dx \qquad (10.17)$$

We say that CLT *holds for the sequence of* $r.v\ .\ s\{X_i\}$

- This theorem is useful to find the probabilities using normal approximation. Normal distribution tables are available for d.f of $N(0,1)$ for all $a \in \mathbb{R} > 0$

$$\varphi(a) = \frac{1}{\sqrt{2_{71}}}\int_{-\infty}^{a} e^{-\frac{1}{2}x^2}\,dx \qquad (10.18)$$

- $$P\left(\left|\frac{\frac{S_n}{n}-\mu}{\frac{\sigma}{\sqrt{n}}}\right| < a\right) = P\left(\left|\frac{S_n}{n}-\mu\right| < \frac{\sigma}{\sqrt{n}}a\right) \qquad (10.19)$$

  Now if the choice of ε is arbitrary take $\varepsilon = \frac{\sigma}{\sqrt{n}}a$ that is $a = \frac{\sqrt{n}}{\sigma}\varepsilon$, as n approaches to ∞ the above probability becomes,

  $$\lim_{n\to\infty} P\left(\left|\frac{S_n}{n}-\mu\right| < \varepsilon\right) = \varphi(\infty) = 1 \quad (10.20)$$

- $S_n \to P\sum_{i=1}^{n}\mu_i$

  .Thus WLLN holds CLT gives probability bound for $\left|\frac{S_n}{n}-\mu\right|$,where as WLLN gives only the limiting value.

- If $X_1, X_2 \ldots X_n$ be the i.i. d bernoulli r.v s CLT becomes

$$\lim_{n\to\infty} P\left(\frac{\frac{S_n}{n}-p}{\sqrt{\frac{pq}{n}}} < a\right) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{o} e^{-\frac{1}{2}x^2}\,dx \qquad (10.21)$$

Example 10.5. *A fair coin is tossed 10, 000 independently. Find the probability that number of heads* (i)*differs by less than* 1% *from 5000* (ii)*is greater than 5100*

Solution: *S_n be the number of heads in 10, 000 independent tosses of a fair coin.*

$Pi(S_n) =$ 5000 *and* $V(S_n) = 2500$

By $CLT(i)$

$$\lim_{n\to\infty} P(|S_n - 5000| < 50) = \lim_{n\to\infty} P(\frac{|S_n - 5000|}{\sqrt{2500}} < 1) \qquad (10.22)$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-1}^{1} e^{-\frac{1}{2}x^2} dx = 2\varphi(1) - 1 = 0.6826 \qquad (10.23)$$

(ii)

$$\lim_{n\to\infty} P(S_n > 5100) = \lim_{n\to\infty} P(\frac{S_n - 5000}{\sqrt{2500}} > \frac{5100 - 5000}{\sqrt{2500}}) \quad (10.24)$$

$$= \frac{1}{\sqrt{2_7 \mathrm{r}}} \int_{2}^{\infty} e^{-\frac{1}{2}x^2} dx = 1 - \varphi(2) = 0.0228 \qquad (10.25)$$

Example 10.6. *How many independent tosses of a fair die are required for the probability that average number of sixes differ from* $\frac{1}{6}$ *by less than 6% to be at least 0.95?*

Solution. $\frac{S_n}{n} = \frac{5th}{36n} \frac{S_n}{)n}$ *bee number average of sixes in* n *independent tosses of a fair die.* $E\left(\frac{S_n}{n}\right) =$ *By* $CLT(i)$

$$\lim_{n\to\infty} P(\left|\frac{S_n}{n} - \frac{1}{6}\right| < .01\sqrt{\frac{5}{36n}}) \geq 0.95 = \qquad (10.26)$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-.01\sqrt{\frac{5}{36n}}}^{.01\sqrt{\frac{5}{36n}}} e^{-\frac{1}{2}x^2} dx = 2\varphi\left(.01\sqrt{\frac{5}{36n}}\right) - 1 \geq 0.95 \quad (10.27)$$

$$2\varphi(1.96) - 1 = \varphi(1.96) - \varphi(-1.96) = 0.95 \qquad (10.28)$$

So,$.01\sqrt{\frac{5}{36n}} \geq 1.96$ *or it gives* $n > \frac{196^2 x5}{36x.01^2} = 5336$ *Toss the die at least 5336 times to get the result.*

10.5 Chapter End Exercises

1. X is a r.v assuming values -1, 0,1 with probabilities 0.125, 0.75, 0.125 respectively.Find the bounds on $P[|X|>1]$

2. Find K such that probability of getting hcad between $450 to K$ is 0.9, in 1000 tosscs of a fair coin.

3. 3. $f(x)=e^{-}xx\geq 0$Γind the bound on the probability $P[|X-1|>2]$, and compare it with actual probability

4. Examine whether SLI͟N holds for the following sequence of the independent r.v.s.

$$X_k=\begin{cases}\pm 2^k & withprob=\dfrac{1}{2^{2k+1}}\\ 0 & withprob=1-\dfrac{1}{2^k}\end{cases}$$

5. Examine whether WI͟LN holds for the following sequence of the independent r.v.s.

$$X_k=\begin{cases}-\sqrt{k} & withprob=\dfrac{1}{2}\\ \sqrt{k} & withprob=\dfrac{1}{2}\end{cases}$$

6. Suppose a large lot contains l% defectives. By using CLT, find approximate probability of getting at least 20 defectives in a random sample of size 1000 units.

7. $\{X_i\}$ is sequence of independent r.vs such that $E(X_i)=0$ and $V(X_i)=\frac{1}{3}S_{100}=\sum_{i=1}^{100}X_i$ approximately the $P[S_n>0.2]$

❖❖❖