NATIONAL OPEN UNIVERSITY OF NIGERIA

SCHOOL OF SCIENCE AND TECHNOLOGY

COURSE CODE: STT 211

COURSE TITLE: PROBABILITY DISTRIBUTION I

Course Code   STT 211

Course Title   Probability Distribution I

Writer   Arowolo. OLATUNJI
School of Science and Technology
Lagos State Polytechnic. Ikorodu..

Course Editing Team

PROBABILITY DISTRIBUTION I

MODULE 1

UNIT 1

PREREQUISITES

The main prerequisites for understanding the content of this book is a knowledge of elementary algebra, set theory, mathematical induction, differentiation and integration.

Set

A set is any well-defined list or collection of objectives. The objectives comprising the set are called its elements or members.

A set will be denoted by capital letters or symbols such as X,Y,A,B,….. and its elements will be denoted by lower case letters x, y, a,b,…..

Example 1.1 Toss a cubical die once. There are six possible numbers that can appear. We can write

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

Where Ω is a set consisting of six elements 1, 2, 3, 4, 5, 6 called the element of the set Ω.

There are essentially two ways of specify a particular set. One way if possible is by listing its elements as in example1.1 above. The other way is by stating properties which characterize the elements in the set. The above set Ω can be written as:

$$\Omega - \{x: \text{ is an integer}, 1 \le x \le 6\}$$

If x is an element Ω, the notation $x \in \Omega$ means that x belongs to Ω. The negation of this assertion i.e. the statement that x does not belong to Ω will be denoted by

$$x \in \Omega$$

thus, for the above example $2 \in \Omega$ but $8 \in \Omega$.

Definition 1.2

Subset: A set A is a subset of a set if each element in A also belongs to Ω.

In example 1.1 above, the set

$$A = \{1,3,5\} = (x: x \in \Omega \text{ and x is odd}\}$$

Is a subset of Ω, that is each element of A is in Ω.

Two sets A and B are called equal if and only if they contain exactly the same elements.

3

Throughout this book, whenever the word set is used, itwill be interpreted to mean a subset of a given set denoted by Ω. The set which contains no elements is called the Null set.

Definition

Let A be a set. The elements which are not included in A also constitute a subset. This is known as the complement of A and is denoted by A^c.

In example 1,1 if A {1,3,5}

$$A^c = \{2,4,6\} = \{x: x \; \Omega, x \text{ even}\}.$$

Definition:

Two sets A, B define two related sets. One of these is the set of all elements which belongs to both sets A and B. this is called *the intersection* of A and B and is denoted by A∩B. the other is the set of all the elements which occur in either A or B or both. This is called the union of A and B denoted by A B.

Example 1.2

Let $\Omega = \{1,2,3,4,5,6\}$, A = {1,3,5} and B = {2,3,5}

Then

$$A \cap B = \{3,5\}$$

$$A \cap B = \{1,2,3,5\}$$

A and B contain 3 element each while A ∩ B contains 2 elements and A ∩B contains 4 laments.

Note that the number of elements in A ∩ B is not the sum of the number of elements in A and B.

Definition: *Difference of Two sets*

The different of A and B is the set of elements which belong to A but not to B and is denoted by A/B

In example 1.2 above $A/B = \{x: x \; \in A, x \; \in B\}\Omega$

A/B = {1}

B/A = {2}

The union of two sets A B can be divided into three disjoint sets

A/B, A ∩ B, B/A.

That is

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

The number of element in a set will be denoted by nA. Thus,

$$n(A \cap B) = n(A/B) + n(A \cap B) \cap (B/A)$$

Since

(A/B), (A ∩ B), (B/A) are disjoint sets

$$nA = n(A/B) + n(A \cap B)$$

$$nB = n(B/A) + n(A \cap B)$$

$$n(A/B\_ = nA - n(A \cap B)$$

$$n(B/A) = nB - n(A \cap B)$$

hence,

$$n(A \cap B) = nA - n(A \cap B) + n(A \cap B) + nB - n(A \cap B) = nA + nB - n(A \cap B)$$

note: $A/B = A \cap B^c$, $B/A = B \cap A^c$

for any set A,

$$A \cap A^c = \Phi, A \cap A^c = \Omega.$$

For any two sets A and B, we have the following decomposition:

$$B = B \cap \Omega = B \cap (A \cap A^c) = (B \cap A) \cap (B \cap A^c),$$

Since A ∩ B and $A^c \cap B$ are disjoints, we have

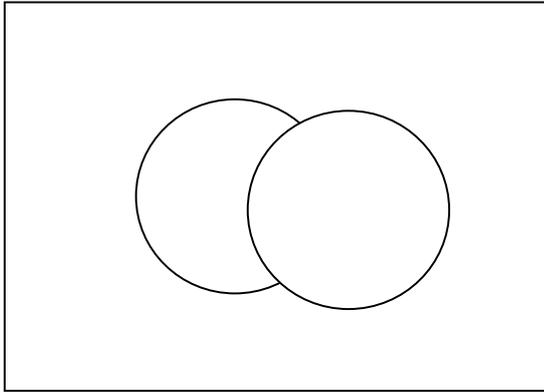$$nB = n(A \cap B) + n(A^c \cap B)$$

de Morgan's Law

  (i) For any two sets A and B
$$(A \cap B)^c = A^c \cap B^c$$

Proof:

$$A^c = (A \cap B)^c \cap (B/A),$$

$$B^c = (A \cap B)^c \cap (A/B).$$

Thus,

$$A^c \cap B^c = (A \cap B)^c$$

Since B/A and A/B are disjoint, Alternatively,

$$(A \ B)^c = \{ x: x \ A \text{ and } X \ B \} = A^c \cap B^c$$

In general, if $A_1, A_2, \ldots A_n$ are any n sets

$$(A_1 \cap A_2 \cap \ldots \cap A_n)^c = A_1^c \cap A_2^c \ldots A_n^c.$$

Thus

Series

A sequence is a set of numbers occurring in order, and there is a simple rule by which the terms are obtained. For example, 1, x, x^2… is a sequence. If the terms of a sequence are considered as a sum, for instance, $1 + x + x^2 + \ldots$ The expression is called a series. A series with a finite number of terms is called a finite series otherwise it is called a finite series otherwise it is called an infinite series. The summation is shown by the symbol.

When the sum is taken from the first term (r = 1) to thee nth term (r = n).

Example

The series $1 + x + x^2 + \ldots + x^{10}$ can be written as

6

The above series is called a geometric series with common ratio x. the sum

$$X^{r-1} = \frac{1 - xn}{1 - x} \quad x = 1$$

$$N, \qquad x = 1$$

The commonratio is x. if $-1 < x < 1$, $x^n \to 0$ as $n \to =$, then the series converges to

$$\frac{1}{1 - x}$$

The infinite series

Example

Find the sum of the series

Solution

$$(1 - p)^{r-1} = 1 + (1 - p) + 1 - p)^2 + \ldots$$

The common ratio is $1 - p$, from (1) let $x = 1 - p$, we have

$$(1 - p)^{r-1} = \frac{1}{1 - (1 - p)} = \frac{1}{p}$$

And

$$(1 - p)^{r-1} = \frac{1 - (1 - p)^n}{1 - (1 - p)} = \frac{1 - (1 - p)^n}{P}$$

Exponential series

The function $y = e^x$ is called an exponential function. This function is one of the special functions of analysis and can be defined as the sum of an infinite series. It is defined by

$$E^x = 1 + x + \frac{x^2}{2!} + \ldots + \frac{x^n}{n!} + \ldots = \frac{x^r}{r!}$$

The notation r! is called factorial r and is defined as

$$R! = r(r - 1)(r-2) \ldots \ldots \qquad\qquad 2.1$$

For example

$$5! = 5 \times 4 \times 3 \times 2 \times 1$$

Where $0! = 1$ and $1! = 1$. R! defined above has no meaning unless r itself is a positive integer.

The factorial of any negative integer is infinite.

Gamma Function

It is easily shown by direct integration that when m is an integer

$$M! =$$

Provided that $m > -1$. It can be proved that (2) will also have a meaning for fraction m. for instance, it is known that

Since $m! = m(m-1)!$. By setting $m = \dfrac{1}{2}$ we have

The Binomial Theorem

If n is a positive integer,

$$(a+b)^2 = a^n + {}^nC_1a^{n-1} B + {}^nC_2a^{n-2}b^2 + \ldots + b^n$$

Where ${}^nC_r = \dfrac{n!}{(n-r)!r!}$.   Using the summation notation, we have

$(a+b)^n$ using the summation notation, we have

Example

Find the sum of the series

$$S_n = P^n + {}^nC_1p^{n-1} (1-P) + {}^nC_2p^{n-2}(1-P)^2 + \ldots + (1-p)^n {}_nCr_P^{n-r} (1-P)^r$$

Substituting p and 1-p for a, b in the binomial theorem above, we have

Thus,

Generalization of the Bionomial theorem is the ***multinomial*** theorem. If n is a positive integer

$$(a_1+a_2 + \ldots a_k)^n = \sum \ldots \ldots \sum \dfrac{n!}{N_1! \, n_2! \ldots n_k!} a_1^{n1}a_2^{n2} \ldots a_k^{nk}$$

Sum is over all $n_1, n_2 \ldots n_k$ where $n_1 + n_2 + \ldots + n_k = n$.

Product Notation

The product of the terms of a sequence x_1, x_2…., x_n can be written as $x_1 x_2 x_3$….x_n. This product is shown by the symbol.

Where the product is taken from x_1, (the first term) to the nth term, x_n. for example,

$$N_1! \ N_2!...n_k! \ \text{and} \ a_{12}...a_k$$

Can be written as

Thus the multinomial theorem can be written as

Example

$$(a_1 + a_2 + a_3)^2 = \sum\sum \frac{2!}{3}$$

Where n=2, k = 3. Possible values of n_1 are

$$n_1 = 2, \ n_2 = 0, \ n_3 = 0$$
$$n_1 = 0, \ n_2 = 2, \ n_3 = 0$$
$$n_1 = 0, \ n_2 = 0, \ n_3 = 2$$
$$n_1 = 1, \ _n 2 = 1, \ n_3 = 0$$
$$n_1 = 1, \ n_2 = =, \ n_3 = 1$$
$$n_1 = 0, n_2 = 1, \ n_3 =$$

hence, we have

$$(a_a + a_2 + a_3)^2 = a_1{}^2 + a^2{}_2 + a^2{}_3 + 2a_1a_2 + 2a_{1a}3_{+2a}2_a3$$

Exponential Functions

These are function in which the variable occur in the index, for example $e^x 3^{2x}$ are called exponential functions, Let $y = e^{h(x)}$ then

For example, if $y = e^{3x2}$ then

$$\frac{dy}{dx} = 6xe^{3x2}$$

Derivative of $\log_e h(x)$

If $y = \log_e 2 \ x^3$ then

9

$$\frac{dy}{dx} = 6x \cdot \frac{1}{2x^3} = \frac{3}{x^2}$$

UNIT TWO

MATHEMATICS OF COUNTING

Probability had its origin in games of chance such as disc and card games. The number of definitions you can find for it is limited only by the number of books you may wish to consult. Probability can be defined as a measure put on occurrence of a random phenomenon. probability theory is developed as study of the outcomes of trail of an experiment.

Definition

An experiment is a phenomenon to be observed according to a clearly defined procedure. Probabilities are numbers between 0 and 1, inclusive that reflect the chances of a particular physical events occurring. If a die is tossed once, the possible outcomes are 1,2,3,4,5,6, let Ω = {1,2,3,4,5,6}. Then Ω is a set consisting of all possible outcome of tossing a die once. This set is given a name, it is called a sample space.

Example 1.1

Write down the staple space for each of the following experiments

(i) Toss a coin 3 times and observe the total number of heads

(ii) A box contains 6 items of which 2 are defectives. One item is chosen one after the other without replacement until the last defective items is chosen. We observe the total number of items removed from the box.

Solution

(i) Possible outcomes are: number of heads is 0 when we have TTT, number of heads is 1 when we have HTT or THT or TTH, number of heads is 2 when we have HHT, THH, HTH and number of heads are 0,1,2,3. Hence the sample space Ω = {0,1,2,3)

(ii) Total number of items removed is 2 if we have "the first is defective (D) and the second is defective D (Note that the total number of items removed can not be 0 or 1) denoted by DD. The total number of items removed is 3 if we have

GDD or DGD (where G denoted good item), and so on. Thus, the sample space

$$\Omega = \{2,3,4,5,6\}$$

Definition 1.2 Event

Any subset A of a sample space Ω is called an event, where $A < \Omega$.

Example 1.2

The following are examples of events

(i) An odd number occurs when a die is rolled once

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

(ii) Toss a coin 3 times, the total number of heads observed is even

$$\Omega = \{0,1,2,3\} \, A = \{0,2\}$$

Definition 1.3 Trial of an Experiment

A trial is a single performance of an experiment

There are basically two methods for assigning probabilities to events. These are

(i) Relative frequency approach

(ii) Classical approach

(i) **Relative Frequency Approach**

If we are interested in the occurrence of an event A, we could perform a large number of trials and define the relative frequency of A as

$$RF(a) = \frac{\text{Number of trials in which A occurs}}{\text{Total number of trials}}$$

The ratio RF(A) is called the empirical probability. If the number of trials is very large RF(A) will (in most cases0 tend to a particular value called the probability that the event A will occur, that is RF (A) is an estimate of P(A). the probability that A will occur.

Example 1.3

Suppose we are interested in the occurrence of a "6" when we roll a well balanced die (fair die) the die is then rolled 100 times and "6" appeared 15 times. The relative frequency of "6" is thus $\frac{15}{100}$ which is the empirical probability of getting 6

(ii) **Classical Approach:**

In the above example, there are six ways the die can fall when it is rolled once.
So we can define a theoretical probability by

$$P(6) = \frac{\text{number of ways of getting a "6" when rolled once}}{\text{Total number of possible outcomes when rolled once}} = \frac{1}{6}$$

The classical approach to probability gives

$$P(A) = \frac{nA}{n\,\Omega}$$

This definition is valid only when outcomes are equally likely.

Example 1.4

A coin is rolled thrice. The possible outcomes assumed to be equally likely are

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

Let A be the event that 2 heads occur. Then

$$A = \{HHT, HTH, THH\}$$
$$n\,\Omega = 8, \quad nA = 3$$

hence

$$P(A) = \frac{3}{8}$$

This classifiableapproach when applicable (the possible outcomes are equally likely) has the advantage of being exact. Thus to compute a probability by using the above definition one must be able to count.

  (i) $n\,\Omega$ the total number of possible outcomes in the sample space and

  (ii) nA, the number of ways in which event A can occur

equally like outcomes are also called equally probable outcomes. If an event cannot occur its probability is 0, if it must occur its probability is .

the computation of nA and $n\,\Omega$ is easy if Ω has only a few possible outcome, as the number of possible outcomes becomes large, this method of counting all possible outcomes are cumbersome and time consuming.

Alternative methods of counting must therefore be developed. For example if one asks for the probability of getting sum of numbers showing to be 30 when 7 dice are rolled

one must determine how many different ways are possible to get the sum to be 30. Such possible ways include

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

In this chapter weintroduce nontechnical discussion of techniques of mathematics of counting frequently needed in problems of finding nA and n Ω

Exercise 1.1

1. A die is rolled once. What are the probabilities of getting
   (i)   An even number
   (ii)   A prime number
   (iii)   An off prime number
   (iv)   An odd number
   (v)   An even number

Fundamental Principle of Counting

First Law of Counting

If an event A_1 can occur in n_1 ways and thereafter an event A_2 can occur in n_2 ways, "both A_1 and A_2 can occur in this order in n_1n_2 ways.

Example 1.5

Roll a die first and then a con

A_1 can occur in 6 ways (1,2,3,4,5,or 6) and A_2 can occur in 2 ways (H or T).

Thus, by the above law, there are

$$6 \times 2 = 12$$

Possible ways for the outcomes

The outcomes are

(1,H) (1,T)   (5, H) (5,T)

(2,H) (2,T)   (6,H) (6, T)

(3, H) (3, T)   (4, H) (4,T)

In general, if an event A_1 can occur in n_1 in different ways and if following this an event A_2 can occur in n_2 different ways, and if following this second event, an event A_3 can

occur in n_3 different ways and so forth, then the events A_1 and A_2 and A_3…. And A_k can occur in this order in $n_1 n_{2…} n_k$ ways.

Example 1.6

If a die is rolled 10 times. Let A_1 denote the outcome of the roll, i= 1,2 ,…. 10. A_1 can occur in 6 ways.

Thus thenumber of ways A_1, A_2,…. And A_{10} can occur is 6^{10} possible outcome.

That is

$$N \Omega = 6^{10}$$

Second Law of Counting

If an eventA_1 can occur in n_1 different ways and an event A_2 can occur in n_2 ways then either A_1 orA_2 can occur in $n_1 + n_2$ different ways.

Example 1.7

Let us toss a die or a coin once. let a_1 ne the event "the die shows an even number and a_2 be the event "the coin lands heads". a_1 can occur in 3 ways (2, 4or6) and a_2 can occur only ocne.

\the number of ways in which an even number or a head be obtain is

$$3 + 1 = 4$$

A_1 can occur in 3 ways A_2 can occur in 1 way

Therefore A_1 or A_2 can occur in 4 ways

In general, if events A_1, A_2….,A_k can occur in n_1….n_k different ways, then either or A_1 + n_2 +… + n_k different ways

Example 1.8

Four people enter a restaurant for lunch in which there are six chars. In how many ways can they be seated.

Let A_1, A_2,A_3,A_4 denote the events "choice of chair by the four people", the suffix denoting order of seating. The first person to sit down has six choices. He can decide tosit on any of the six vacant chairs. Therefore, there are 6 different ways A_1 can occur, after

the first person has seated, the second person can sit on any of the remaining 5 chairs. Thereafter the third person can sit on the remaining 4 chairs.

Thus, using the notation of law of counting

$$N_1 = 6, n_2 = 5, n_3 = 4, n_4 = 3$$

$$6 \times 5 \times 4 \times 3 = 360 \text{ ways}$$

They can be seated .

Example 1.9

How many 4 digits numbers can be formed from the digits 0,1,2,3,4,5 if the first digit must not be 0 and repetition of digits are not allowed.

Let A_1, A_2, A_3, A_4, denote mthe events: select the first second, third and fourth digits respectively.since the first digits can not be 0, the first digit can e either 1,2,3,4, or 5 therefore A_1 can occur in 5 ways, having chosen the first digit the second digit can be selected from the remaining 5 digits. there the first and second digits have been chosen, there remains 4 digits. The third digit can be chosen from the remaining 4 digits and the fourth can be chosen from the remaining 3 digits. Therefore,A_1, A_2, A_3 and A_4can occur in 5,5,4,3 ways respectively

Thus there

$5 \times 5 \times 4 \times 3 = 300$ numbers

Example 1.10

Ten candidates are eligible to fill 4 vacant position. How many ways are there of filling them?

Let A_1,A_2, A_3,A_4 be the events denoting candidates that fill positions 1,2,3,4 respectively. Candidate filling position 1 can be any of the tend candidates, therefore A_1 can occur in 10 ways. Following this, the next position 2 can be filled by any of the remaining 9 candidates since position 1 had been filled by one candidate. Therefore A_2 can occur in ways. Similarly, A_3 and A_4 can occur in 8 and 7 ways respectively. Thus there are

$10 \times 9 \times 8 \times 7 = 5040$ ways

Exercise 1.2

1. A student is to answer all the five questions in an examination. It is believed that the sequence in which the questions are answered may have a considerable effect

on the performance of the student. In how many different order can the question be answered

2. If a woman has 10 blouses and 6 skirts, in how many ways can she choose a dress assuming any combination of blouse and skirt matches

3. In a study of plants, five characteristics are to be examined. If there are six recognizable differences in each of four characteristics and eight, recognizable difference in the remaining characteristics. How many plants can be distinguished by these five characteristics?

4. A bus starts with 6 people and stops at 10 different stops.how many different ways can the 6 people depart if

   (i) Any passenger can depart at any bud stop

   (ii) No two passengers can leave at the same bus stop

5. Show that the number of ways of choosing r objects from n objects with replacement is given by n"

Permutation and Combination

A permutation is an arrangement of objects in a definite order (called ordered sample). A combination is a selection of objects without regard to order (unordered sample)

A group of objects, with regard to permutation and combination has three characteristics:

1. The way the objects in a group are arranged
2. The kind of objects in the group
3. The number of objects of each kind in the group

1.2 permutations:

Suppose that a set contain n objects. We are often interested in arranging the objects in a definite order.

Two groups containing n object are said to form different permutations if they differ in arrangements

Consider groups of letters a, b, c, d.

$$(a, b, c, d): (b, c, d, a): (c, a, d, b)$$

They are all different permutations because the arrangement of the letters is different in each group.

Example 1.11

How many permutations of four letters can be formed from the letter a, b, c, d.

To answer this question we reason as follows:

Since are permuting 4 letters, four events are involved. Let A_1 be the events denoting the letter to occupy the ith position. A_1 can occur in 4 ways, that is the first letter can be either a, b, c, or d. after this event has taken place, A_2 can occur in three ways (that is, after having chosen the first letter, the letter occupying the second position can be chosen from the remaining three letters. After the second event, there remains only two letters from which one is to be chosen tooccupy the third position. So A_3 is 2 and it remains only one letter.

A_4 can occur only in one way. Thus, there are 4 x 3 x 2 x 1 = 24 permutation of the four letters. Permutation of four objects from four objects is called permutation 4 and is denoted by $^4P_4$

Permutation of n Distinct Objects

Consider a set consisting of n district objects. permutation of this set consist of n events A_1, A_2, ….A_n, where the object occupying the ith position is the outcome of A_1, I = 1, 2,…..n A_1 can occur in n_1 ways, A_2 can occur in n-1 ways and so on.

Thus there are n(n-1) (n-2)… x 3 x 2 x 1 permutation of the n objects. The number of permutations of n distinct objects is n!.

Example 1.12

If n balls are distributed at random into r boxes, in how many ways can this be done if

    (i) Each of the balls can go into any of the r boxes

    (ii) No box has more than one ball.

Solution

    (i) Let A_1, A_2….A_n where the object occupying the ith position is the outcome of A_1 I = 1,2,…..n. A_1 can occur in n_1 way, A_2 can occur in n-1 ways and so on.

(ii) The first ball can go into any of the r boxes, the second any one of the remaining (r-1) boxes etc, so in all there are r (r-1) (r-2)… (r-n+1) different ways (n ≤ r).

Example 1.13

How many permutation of three letters can be formed from the letters a, b, c, d, e. abc, bae, cba, cdb,….. are few of the required permutation. Since the permutation consist of 3 letters, 3 events are involved, A_1, A_2 A_3. The first letter can be any of the 5 letters, therefore A_1 can occur in 5 ways. Similarly A_3 can occur in 3 ways. Thus, there are

$$5 \times 4 \times 3 = 60$$

Permutation of three letters from 5 letters. This is denoted by $^5P_3$ (5 permutation 3).

Definition 1.4

The number of permutation of r ($r \leq n$) objects from n district objects is called ^{n}n permutation r and is denoted by ^np_r

It can easily be shown that

$$^nP_r = n(n-1)(n-2)\ldots(n-r+1)$$

to see this we can argue as follows. Since we are permuting r objects, there are r events A_1, A_2…..A_3 involved.

The first position can be occupied by any of the n objects, following this the second position can be occupied by any of the remaining n-1 objects and so on. Therefore A_1, $A_{2,…}A_r$ can occur in n, n-1, n-2, …n-(r-1) ways respectively. thus the number ofpermutation is

$$n(n-1)(n-20\ldots(n-r+1)$$

But

$$n! = n(n-1)(n-2)\ldots(n-r+1)(n-r)(n-r-1)\ldots1 \qquad (3.2.1)$$

$$(n-r)! = (n-r)(n-r-1)\ldots2.1$$

$$\underline{\;n!\;} = n(n-1)(n-2)\ldots(n-r+1).$$
$$(n-r)!$$

Therefore

$$^nP_r = n(n-1)(n-2)\dots(n-r+1)$$

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

In example 1.13 above, n = 5, r = 3

$$^5P_3 = \frac{5!}{(5-3)} = \frac{5!}{2} = 5 \times 4 \times 3 = 60$$

Example 1.14

A group of students consist of 5 men and 3 women. The students are ranked according to their performance in a quiz competition. Assuming no two students obtain the same score

(i) How many different ranking are possible?

(ii) If the men are ranked just among themselves and the women among themselves, how many different rankings are possible?

Solution

(i) A possible ranking corresponds to a permutation of the students. The number of possible permutation of the 8 students gives the number of different ranking possible. Thus the answer is

$$^8P_8 = \frac{8!}{(8-8)!} = 8! = 40320.$$

(ii) There are $^5P_5 = 5! = 120$ possible rankings of the men and $^3P_3 = 3! = 6$ possible rankings of the women. It follows from the fundamental principle of counting that there are

$$^5P_5 \times {}^3P_3 = 5! \times 3! = 72\text{- possible rankings}$$

Example 1.15

Four digits numbers are to be formed using any of the digits 1, 2, 3, 4, 5, 6, (No repetition of digit is allowed).

(i) How many four digit number can be formed

(ii) How many 4 digit numbers greater than 3000 can be formed?

(iii) How many 4-digit even numbers can be formed?

(iv) How many of these even 4-digit numbers are greater than 3000?

Solution

(i) This is permutation of 4 digits from 6 digits. Therefore the different permutation is

$$^6P_4 = \frac{6!}{2!} = 360.$$

Alternatively, we can reason as follow:

The first digit to be selected can be any of the six given digits, so $n_1 = 6$. The second digits to be select can be any of the remaining 5 digits (since no reparation is allowed) so $n_2 = 5$. Similarly, $n_3 = 4$ and $n_4 = 3$. Thus, the answer is

$$N_1 \times n_2 \times n_3 \times n_4 = 6 \times 5 \times 4 \times 3 = 360$$

(ii) Since the number must be greater than 3000, the first digit must be chosen room 3, 4, 5 or 6 so $n_1 = 4$. The second digit can be any of the remaining 5 digits, so $n_2 = 5$ similarly, $n_3 = 4$, $n_4 = 3$. Thus, there are

$$4 \times 5 \times 4 \times 3 = 240$$

4 digit numbers greater than 3000 that can be formed using the digits 1, 2, 3, 4,5, 6,

(iii) A number can be defined to be an even number of the last digit of the number is even. Going by this definition, we see that the required number is even if the last digit is 2,4, or 6. Since there is a restriction on the last digit. We have to select the last digit first. The last digit can be selected from 2,4 or 6, so we have 3 choices. Having chosen the last digit, we can now select the first, second and third digit. After chosen the last digits there remains 5 digits, so $n_1 = 5$ similarly, $n_2 = 4, n_3 = 3$. Thus, the answer is $\quad$ 5x 4 x 3x 3 = 180

(iv) The last digit must be 2, 4 or 6 and the first digit must be 3, 4, 5or 6. If the last digits is 2, then the number of ways of selecting the first digit is 4. However, if the last digit is selected is 2 or 6 the number of ways of selecting the first digit is 3 (that is 3, 5 or 6, 3, 4 or 5)

Case 1: $\quad n_4 = 1$ $n_1 = 4$ $n_2 = 4$, $n_3 = 3$

The number of 4 digits numbers that can be formed in this case is

$$4 \times 4 \times 3 \times 1 = 48$$

Case II: $\quad n_4 = 2$, $n_1 = 3$ $n_2 = 4$, $n_3 = 3$

The number of 4 digit numbers that can be formed in this case is

3 x 4 x 3 x 2 = 72

Hence, the total numbers of even 4 digit numbers greater than 3000 that can be formed is 48 +72= 120.

Example 1.16

The letters A,B,C, D and E are placed at random to form a five letters word (without repetition). How many ways can a word be formed such that

(i) D directly follows A.

(ii) A and D follow each other

(iii)   A, D and E follow each other

Solution

(i) For D to directly follow A, we must always have AD appearing in the word so formed AD can be regarded as a letter so that we now have the letters B,C E and AD. The number of ways of rearranging tehse new four letters is 4! Thus there are 4! = 24 ways of forming a word such that D directly follows A

(ii) A and D follow each other; either we have AD or Da each having 4! Ways. Thus there are 2x 4! Ways of forming a word such that A and follow each other

(iii)   ADE can be regarded as a letter so that we now have the letters B.C ADE. There are 3! Permutation of the new 3 letters. ADE can also be permuted in 3! Ways, that is

ADE, AED, DAE, DEA, EAD, EDA.

Thus there are

3! X 3! = 36 ways

Permutation of Indistinguishable Objects

Consider n objects where n_1 are of type 1, of type 2,…. N_k of type k. ,in shown many ways can the n object be arranged.

For example in how many ways can the letters of the word book be arranged. here n= 4, k = 3, n_1= 1, n_2= 2, n_3 = 1. First give the two O's suffixes bo_1o_2k. then treating the O's as different, the 4 letters may be arranged in 4! Ways. In every distinct arrangement, the

2O's may be rearranged amongst themselves in 2! Ways without altering the permutation for instance o_1bo_2k are the same when the suffixes are removed.

Therefore, the number of permutations of the letters of the word book is

$$\frac{4!}{2!} = 6$$

Book, ookb, oobk, obok, okob, koob.

In general, the number of ways in which n objects where n_1 are of type 1, n_2 of type 2…, n_k of type k can be arranged is given by

$$\frac{n!: n_{+n}2_{+…+n}k}{n_1!n_2!...n_k!}$$

Example 1.17

How many distinct permutations are there of the letters of the word Television?

The ten letters to be permuted consist of 2e's, 2i's,IT, I, v, s, o, n. thus the number of distinct permutation is

$$\frac{10!}{2!\ 2!\ 1!\ 1!\ 1!\ 1!\ 1!\ 1!} = \frac{10!}{2!\ 2!}$$

1.3 Combinations

Definition 1.5

Two groups are said to form different "combination" if they differ in the number of any kind of object in the groups. Consider a group of 4 letters a, b, c, d. the combination abcd, bcda, cadb are identical combinations each of them contains the same number of a, b, c, d: one b, one c and one d.

The combinations of the 4 letters taken 3 at a time are:

Abc, acd, abd, cbd

Therefore, there are 4 distinct combination for three letters from the four letters. Each of tehse combinations has 3! = 6 permutation. For instance.

abc	=	abc,	acb	cab	cba	bac	bca
acd	=	acd	adc	cad	cda	dac	dca
abd	=	abd	adb	bad	bda	dab	dba

24

| | | cdb | = | cbd | cdb | bcd | bdc | dbc | dcb |

The number of distinct permutation is $^4P_3 = 24$

The number of distinct combinations is 4.

Therefore.

$$\text{Number of combinations} = \frac{\text{number of permutations}}{3!}$$

The number of distinct combination of 4 objects takn 3 at a time is denoted by $^4C_3$.

Thus,

$$^4C_3 = {}^4P_3/3!$$

In general,

$$^nC_r = {}^nP_r/r! = \frac{n!}{(n-r)!r}$$

Theorem 1.1

The number of distinct combinations of n objects taken r at a time (that is the number of ways of choosing r objects out of n, disregarding order and without replacement) is given by

Proof:

There are $^nP_r = \dfrac{n!}{(n-r)!}$ permutations of n objects taken r at a time. If we disregard order among the r objects, there are r! permutation that will give the same combination. Therefore the number of combinations is the number of permutation divided by r! thus.

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

Example 1.18

A club consist of 15 members. In how many ways can a committee of 3 to be chosen?

Solution

This can be done in $^{15}C_3$ ways

$$^{15}C_3 = \frac{15!}{} = 15 \times 7 \times 13 \qquad = 455 \text{ ways}$$

25

$$3! \, 12! \qquad\qquad 3$$

Example 1.19

A club consist of 10 men and 5 women, in how many ways can a committee of 6 consisting of 4 men and 2 women be chosen. The 4 men can be chosen from the 10 men in $^{10}C_4$ ways, the 2 women can be chosen from the 5 women in $^5C_2$ ways. Hence the committee can be chosen in (by the fundamental principle f counting)

$$^{10}C_4 \; x \; ^5C_2 = \underline{10!} \qquad X \qquad \underline{15!} \qquad = 2,100 \text{ ways}$$
$$6!4! \qquad\qquad 2!3!$$

Exmaple1.20

Suppose we have a box containing n balls of which r are black and the remaining white. A random sample of size k is selected without replacement. In how many ways can the sample be selected such that it contains x black balls

The x black balls can be selected in ^rC_x ways and k-x white balls n be selected from n-r white balls. Thus, there are

$$^rC_x \; X \; ^{n-r}C_{k-r}$$

ways of selecting a sample such that it contains x black balls and k-x white balls

Example 1.21

A committee of 4 men and 2 women is selected from 10 men and 5women. If two of the men are feuding and will not serve on the committee together, in how many ways can the committee be selected

Solution

The number of ways of selecting 4 men and 2 women is

$$^{10}C_4 \; x \; ^5C_2 = 2100$$

The number of ways of selecting the committee such that the two men are in the committee is

$$^8C_2 \; x \; ^5C_2 = 280$$

26

Hence, the number of different committees that can be formed such that the two men are not in the committee together is

$$2100 - 280 = 1820$$

Another method is to consider 3 cases

Case I: The two men say, A and B are not in the committee

$$^{8}C_4 \times {}^{5}C_2 = 700$$

Case II: A is the committee but not B:

$$^{8}C_4 \times {}^{5}C_2 = 560$$

Case III: B is the committee but not A

$$^{8}C_4 \times {}^{5}C_2 = 560$$

Thus, the total number of ways the committee can be selected is

$$700 + 560 + 560 = 1,820$$

Example 1.22

How man subsets can be formed, containing at least one member from a set of n element

Solution

There are $^{n}C_k$ subsets of size k that can be formed, the total number of subsets containing is least one member is

Example 1.23

A student is to answer 5 out of 8 question in an examination. How many if he must answer at least 2 of the first for 4 questions?

Solution

(i) By the combination law the answer of $^{8}C_5 = 56$

(ii) Possible choices are (2,3) (3,2) (4,1) where (2,3) means answer 2 questions from the first 4 questions and 3 from the remaining 4 questions. The number of ways of doing this is $^{4}C_2 \times {}^{4}C_3$.

Thus the answer is

$$(^4C_2 \text{ x } ^4C_3) + (^4C_3 \text{ x } ^4C_2) + (^4C_4 \text{ x } ^4C_1) = 52$$

1.4 Partitioning

Dividing a population or a sample of n objects into k ordered parts of which the first contain r object, the second r_2 objects and so on is called *ordered partition.* If the division is into k **unordered parts,** then the partition is said to be unordered. For example suppose a class contains 15 students and we want to divide the class into 3 tutorial groups of 5 each. Three lecturers are available and each is to take each group. In otherwise, we want to divide the 15 students into 3 ordered groups (A,B,C) this is an ordered partition since there are

$$3! = 6 \text{ ways}$$

The lecturers can be assigned to take any partition, for instance groups (A, B, C) can be taken by

$$(L_1, L_2, L_3) \text{ or } (L_2, L_1, L_3) \text{ or } (L_1, L_3, L_2)$$

$$\text{Or } (L_2, L_3, L_1) \text{ or } (L_3, L_2, L_1) \text{ or } (L_3, L_1, L_2)$$

Where L_1 means lecturer I ad (L_1, L_2, L_3) means L_1 takes groups A, L_2 takes group B, L_3 takes group C and so on

There are $^{15}C_5$ ways of selecting those students those students to be in group A, following this, there are 10 students left and so there are $^{20}C_5$ selecting those to be in the second group, therefore by the fundamental principle of counting, there are

$$^{15}C_5 \text{ x } ^{10}C_5 \text{x } ^5C_5 = \frac{15!}{10! \, 5!} \text{ X } \frac{10!}{5! \, 5!} \text{ X } \frac{5!}{5!} = \frac{15!}{5! \, 5! \, 5!}$$

Ordered partitions. The number of ways in which n object can be divided into **ordered parts** of which the fist contains r_1 objects the second r_2 objects and so on is

$$^nC_{r1.} \, ^{n-r}_1C_{r2.} \, ^{n-r1-r2}C_{r3...} \, ^{n-r1-r2-...-rk-1}C_{rk}$$

$$= \frac{n!}{(n-r_1)! r_1!} \cdot \frac{(n-r_2)!}{r_2!(n-r_1-r_2)!} \frac{(n-r_1-r_2)!}{(r_3!)n \, r_1-r_2-r_3)!} \frac{(n-r_1-r_2...r_{k-1}!}{R_k!(n-r_1-r_2...r_{k-1}-r_k)!} = \frac{n!}{r_1! r_2! r_3!...r_k!}$$

Where $n = r_1 + r_2 + ... + r_k$.

Example 1.24

In how many ways can three committees of five, three and two persons be formed from 10 persons.

We seek the number of ordered partitions of the 10persons

This is given by

$$\frac{10!}{5!\ 3!\ 2!} = 2520$$

Example 1.25

In how many ways can 9 toys be divided among three children if each gets 3 toys

If the toys are numbered 1 through 9, the partition {(1, 2, 3), (4, 5, 6), (7, 8, 9)) means child A gets toys 1, 2, 3 child B gets toys (4, 5,6) while child C gets toys 8, 8, 9. We distinguish between {(1, 2, 3), (4, 5,6) (7, 8, 9)} and {(4, 5, 6), (1, 2, 3), (7,8, 9) so these are ordered partitions. Thus there are

$$\frac{9!}{3!\ 3!\ 3!} = 1680 \text{ ways}$$

When $r_1 = r_j$ for $I = j$, we can distinguish between ordered and unordered partition for example partition a set consisting of 8 objects numbered 1 to 8 into 3 parts of which the first contains 2 objects, the second 4 objects and the third 2 objects. A partition is

$$\{(1, 2), (3, 4, 5,6), (7, 8)\}.$$

In an ordered partition we distinguish between the partition

$$\{(1,2), (3,4, 5, 6), (7, 8)\} \text{ and } \{(7, 8), (3, 4, 5, 6), (1, 2)\}$$

But they are the same for unordered partition. Therefore, the number of unordered partition is

$$\frac{8!}{2!\ 4!\ 2!} \cdot \frac{1}{2!} = 210$$

Example 1.26

In how many ways can a family of 9 divide itself into 3 groups so that each group contains 3 persons?

Solution

We are seeking for unordered partitions $r_1 = 3$, $r_2 = 3$ $r_3 = 3$

The number of unordered partitions is

$$\frac{9!}{3!\,3!3!} \cdot \frac{1}{3!} = 280$$

Since the three parts contain the same number of objects

The same relationship that exists between permutation and combination exist between ordered and unordered partition of a set.

Example 1.27

In how many ways can a family of 10 be divided into three groups, one containing 4 and the others 3?

Solution

$$R_1 = 4, r_2 =, r_3 = 3$$

The number of ordered partitions is

$$\frac{10!}{4!\,3!\,3!} = 4,200$$

And the number of unordered partition is

$$\frac{10!}{4!\,3!\,3!} \cdot \frac{1}{2!} = 2,100$$

Since only two of the three parts contain the same number of objects

1.5 **Selection of Non distinct Objects**

The number of ways r districts balls can be distributed into n cells in r. the number of ways if a specified cell contains exactly k balls (k = 0, 1,2…r) is $^rC_k\,(n-1)^{r-k}$

That is the balls can be chosen in ^rC_k ways and the remaining r-k balls can be paced into the remaining n-1 cells in $(n-1)^{r-k}$ ways

Now, suppose the balls are non-district (indistinguis-shable) we can only talk about number of balls in the ith cells

Let X_1, x_2…, x_n denote the number of balls in the ith cell, then

$$x_1 + x_2 + … + x_n = r$$

The number of distinct distribution in which no cell remains empty is

$$^{r-1}C_{n-1}$$

to see this let us assume that the r non-distinct objects are lined by and bars used to divide them into groups. The r balls r-1 space of which n-1 are to be occupied by bars. For example if r =9 and n = 5, we have

$$0 \cap 0 \cap 0 \cap 0 \cap 0 \cap 0 \cap 0 \cap 0 \cap 0$$

Thus 0/00/0/000/00 corresponds to

$$X_1 = 1, x_2 = 2, x_3 = 1, x_4 = 3, x_5 = 2.$$

And there are $^8C_4$ possible distribution of the bars. another possible distribution is

$$000/0/00/00$$

If $x_1 \geq 0$, that is cells can remain empty, the number of non-negative solutions of

$$X_1 + X_2 + … + x_n = r$$

Is the same as the number of positive solutions of

$$Y_1 + y_2 + … + y_n = r + n$$

Where $y_1 = x_1 + 1$. Thus, there are

$$^{r+n-1}C_{n-1} \qquad (1.2)$$

Distinct solution satisfying $x_1 + x_2 + … + x_n = r$

The number of distinct distribution is the number of ways n – 1 spaces can be selected out of the n + r-1 spaces

Application to Runs

Definition

A run is any ordered sequence of elements of two kinds

For example, by a run of wins we mean a consecutive sequence of wins. The sequence WWWLWWLLWLWW gives 4 runs of wins. The first run is length 3, the second run of length 2, the third of length 1 and the fourth of length 2.

31

Suppose now that we have two letters (W and L) n non-distinct letters (L) and m non-district letter (W). the total number of distinct orderings of W and L is $^{(n+m)}C_{n\text{-}}{}^{r}$ r runs of W is equivalent to arranging the letters W into r cells none of which is empty. If there are r runs on W,

The number of L run is necessarily r+ 1, r-1 or. Thus from (1.1) we have $^{(m\text{-}1)}C_{r\text{-}1}$ distinct was of having r runs of W. hence there are.

$^{(m\text{-}1)}C_{r\text{-}1}{}^{(n\text{-}1)}C_{r}$ ways of having r runs of W and (r + 1) runs of 1…

A sequence representing r runs of W is

LL…	WW…W	LL…	WW…W	…	WW…	L..L
Y_1	X_1	Y_2	X_2		X_r	Y_{r+1}

Where $y_1 + y_2 +… y_{r+1} = n$, $y_i \geq x_1 + x+2+ +… + x_r = m$, $x_i \geq 0$.

Let $y_1 + 1$, $y_i = y_i$, $I = 2, …, r$, $y_{r+1} = y_{r+1} + 1$

The number of non-negative solution to

$$Y_1\ y_2 + … + y_{1+1} = n, (y_i \geq 0)$$

Is the same as number of positive solution to

$$Y_1 + y_2 +…+ y_{r+1} = n + 2$$

Thus, the number of outcomes that result n r runs is $^{n+1}C_r$

Hence, $^{n\text{-}1}C_r{}^{m\text{-}1}C_{r\text{-}1}$ is the total number of ways of having r runs of W

Exercise 1.3

1. How many 4 digits numbers can be formed from the digits 0,1, 2, 3, 4,5, 6, 7,8, 9 if the first digit must not be 0 and repetitions is allowed. how many of these number are (i) even (ii) less than 5,000?

2. A code consist of five symbols. The first three symbols are letters and the last two are digits. How many codes can be made if no letter nor digit is repeated in any code word?

3. Suppose n objects are permuted at random among themselves. In how ways can this be done such that k specified objects occupy k specified positions

4. How many distinctarrangements are there of the letters of word (1) University (ii) biology (iii) Mississippi

5. The number 1,2,3… n are arranged in random order. in how many ways can this be done such that

    (i)   Q and 2 follow each other

    (ii)   2, 3 and 4 follow each other in that order

6. Show that the number of distinct ordered sample of size r that can be drawn from a population with n objects is

    (i) N^r if sampling is with replacement and

    (ii) $\dfrac{n1}{n-r!}$ if sampling is without replacement

7. A total of n balls are randomly placed into n cells. In how many ways can this be done. In how many ways if each cell is occupied

8. A warehouse has 6 different containers to be distributed among 10 retailers. In how many ways can this be done? How many way if no retailer receives more than one container?

9. Prove the following identities

    a. $^nC_x = {}^nC_{n-x}$ (b) $^{m+1}C_x = {}^nC_x + {}^nC_{x-1}$

10. (a) In how many ways can 4 boys and 2 girls be arranged to sit in a row?

    (b) In how many ways if only the boys must sit together?

    (c) In how many ways if the boy and no girl must sit together?

    (d) In how many ways if no boy and no girl must sit together?

11. In howmany ways can a football team be selected from 15 players. In how many ways if 6 particular players must be included in the team?

12. From a box containing 5 red, 4 white and 3 black marbles, three marbles are drawn one after the other without replacement, how many ways can this be done if

    (i)   all are white (ii) 2 re white and 1 is red (iii) at least one is black

13. A pair of dice is rolled once. In how many ways can

    (i)   The sum of the two numbers appearing exceeds 8

    (ii)   The maximum of the two numbers is greater than 4

    (iii)   The minimum of the two numbers is greater than 4

14. A disciplinary committee of four is to be chosen from six men and five women. One particular man and one particular woman refuses to serve if the other person is on the committee. How many committees may be formed

15. Eleven people are to travel in two cars- salon and station wagon. The saloon has 4 sets and the station going 7 seats. In how many ways can party be split up?

16. In how many ways can a committee of 6 composing of 3 full professors, 2 associate professors and 1 senior lecturer be selected from 5 full professor, 10 associate professors and 20 seniors lecturers

17. A box contain 12 balls labeled 1, 2, 3… 12, suppose a random sample of size 4 is selected. In how many ways can the sample be selected if balls labeled 2, 3, are among the four selected.

18. Suppose a random sample of size r is drawn from a population of n objects. In how many ways can this be done if k given object must be included in the sample and (a) sampling is without replacement (b) Sampling is with replacement

19. I bought 2 tickets to a lottery for which n tickets were sold and 4 prizes to be given in how many ways can the tickets be drawn such that 1 win at least a prize?

20. A committee of 8 is to be formed from 10 couples (10men and 10 women). In how many ways can the committee be formed if no husband serves on it with his wife.

21. Lines are drawn to pass through six points. In how many ways can this be done if each line passes a through only two points?

22. Interchanges may occur between any two of the n chromosomes of a cell

    (a)   In how many ways can exactly one interchange occur?

    (b)   In how many ways can exactly k interchanges occur?

    (c)   If n = 5, in how many ways can at most three inter changes occur?

23. Prove that $^{(n_1-n_2)}C_k = {}^{n_1}C_r {}^{n_2}C_{k-r}$ where $k \leq n_1, n_2$

    HINT: Select k objects from $n_1 + n_2$ obejcts of n_1 of type 1 and n_2 of typeII

24. A company is considering building additional warehouse at new locations. There are ten satisfactory location and the company must decide how many and which ones to select. How many choices are there?

25.   Are there more samples obtainable in five draws from 10 objects with replacement than 12 objects without replacement?

Each of fifty items is tested and found to be defective or non-defective. How many possible outcomes are there?

MODULE TWO

UNIT ONE

ELEMENTARY PRINCIPLE OF THE THEORY OF PROBABILITY

In chapter 1, we considered some elementary methods of mathematics of counting essential for determining probabilities of events. In this chapter, we continue our study by discussing how to apply the knowledge gained in chapter 1 to determine probabilities of events and general properties of probabilities. We present some theorems and definitions that are basic to the understanding of commonly encountered problems.

Example 2.1

Suppose a fair die is rolled once. There are six possible outcomes. The sample space is

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

let A_1, A_2 and A_3, represent the following events

$$A_1 = \text{"an even number occurs"}$$

$$A_2 = \text{``an odd number occurs}$$

$$A_3 = \text{a prime number occurs''}$$

The six possible outcomes in Ω are equally likely. The probability that A occurs is denoted by P(A) and the probability that A does not occur by $P(A^C)$. if we let n_A be the number of outcomes that have attribute A, then

$$A_1 = \{2,4,6\}$$
$$A_2 = \{1,3,5\}$$
$$A_3 = \{2,3,5\}$$

Thus,

$$P(A_1) = \frac{nA_1}{n\Omega} = 3/6 = 1/2, \; p(A_2) = \frac{3}{6}, \; P(A_3) = \frac{3}{6}$$

Example 2.2

Suppose that a box contains 10 items of which are defective. Two items are selected at random without replacement. Find the probabilities that:

(i)   both items are non-defective   (ii)   only one item is defective

(iii)   both items are defective: (iv)   at least one item is defective

Let A_1, A_2, A_3, A_4, denote the events

"both items are non-defective

"only one item is defective"

"both items are defective"

"at least one item is defective"

respectively. The number of ways of selecting 2 items from 10 is

$$^{10}C_2 = 45 \, ways.$$

So there are 45 elements in the sample space.

(i)   The number of ways of selecting 2 items from the non-defective items is $^6C_2 = 15$. That is A_1 can occur in 15 ways. Thus.

$$P(A_1) = \frac{^6C_2}{^{10}C_2} = \frac{nA_1}{n\Omega} = \frac{15}{45} = \frac{1}{3}$$

(ii)   The number of ways of selecting 1 item from the 4 defective items and 1 item from the 6 non-defective items is $^4C_1 \times {}^6C_1 = 24$ so, A_2 can occur in 2 ways. Thus,

$$P(A_2) = \frac{^4C_1 \times {}^6C_1}{^{10}C_2} = \frac{24}{45} = \frac{8}{15}$$

Similarly

(iii) $P(A_3) = \frac{^4C_2}{^{10}C_2} = \frac{6}{45} = \frac{2}{15}$

(iv)   At least one defective means 1 or 2 defective items. So,

$P(A_4) = P(1 \text{ defective item}) + P(2 \text{ defective items}) = P(A_2) + P(A_3) = \frac{8}{15} + \frac{2}{15} = \frac{2}{3}.$

Example 2.3

Suppose a fair die is rolled twice. Find the probability that the sum of the numbers on the two faces is (i) even, (ii) less than 5.

The sample space Ω, consists of 36 elements.

(i)   Let A be the event "the sum of the two faces is even". Possible outcome are:

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

        (3,1) (5,1) (4,2) (6,2),…

A can occur in 18 ways. Thus,

$$P(A) = \frac{18}{36} = 1/2$$

(ii)   Let B be the event "the sum is less than 5". B occurs if the sum is 2, 3, or 4. The sum is 2 if the outcome is (1,1) the sum is 3 if the outcome is (1,2) or (2,1) and the sum is four if the outcome is (3,1), (1,3) or (2,2). Therefore B can occur in

$$1 + 2 + 3 = 6 \text{ ways}$$

Thus

$$P(B) = \frac{6}{36} = \frac{1}{6}.$$

Example 2.4

3 balls are drawn at random with replacement from a box containing 8 red and 3 white balls. Find the probability that (i) all 3 are red; (ii) 1 is red and 2 are white.

The sample space consists of 11^3 possible outcomes. Let A be "the evet all 3 are red" and B the event "1 is red and 2 white"

$$P(A) = \frac{nA}{n\Omega} = \frac{8^3}{11^3}$$

(The first red can be chosen in 8 ways the second in 8 ways and the third Red in 8 ways).

$$P(B) = \frac{nB}{n\Omega}, \quad nB = 8,3.3 + 3.8.3 + 3.3.8$$

$(8 \times 3 \times 3)$ = number of ways of picking the first to be red, second white and third white (RWW), $3 \times 8 \times 3$ is for WRW and $3 \times 3 \times 8$ is for WWR). Thus

$$P(B) = \frac{3.8.3.3}{11^3}.$$

2.1   Properties of Probability

Definition 2.1: Mutually Exclusive events

Two events A_1 and A_2 are mutually exclusive if an only if

$$A_i \cap A_j = \phi \text{ for every } i \neq j.$$

That is, two or more events are mutually exclusive if no two of them have points in common. In example 2.1

$$A_1 = \{2, 4, 6\}$$

$$A_2 = \{1, 3, 5\}$$

$$A_1 \cap A_2 = \phi,$$

Therefore A_1 and A_2 are mutually exclusive events.

Definition 2.1

If A_1 and A_2 are any two events, $A_1 \cap A_2$ is the event that occurs if and only if both A_1 and A_2 occur. In general, if $A_1, A_2, \ldots, A_n$ occur.

In example 2.1. $A_1 \cap A_2 \cap A_3$ is the event that an even prime occurs.

$$A_1 \cap A_3 = \{2\}$$

Thus

$$P(A_1 \cap A_3) = \frac{1}{6}$$

The event $A_2 \cap A_3$ is the event that an odd prime occurs $A_2 \cap A_3 = \{3,5\}$.

Thus

$$P(A_1 \cap A_2) = \frac{2}{6} = \frac{1}{3}.$$

Complementary Event

The event "A occurs" and "A does not occur" are mutually exclusive. The event "A does nt occur" is called the complement of A and is denoted by A^C.

Theorem 2.1

If A^C is the complement of an event A, then $P(A^c) = 1 - P(A)$. This theorem states that the probability that an event will not occur is equal to 1 minus the probability that it will occur.

In example 2.2 (iv); The event "no defective item" is the same as the event both items are non-defective".

From Theorem 2.1, we have

$$P(A_4) = 1 - P(A_4^c) = 1 - P(A_1) = 1 - \frac{1}{3} = \frac{2}{3}$$

$A_4^c = A_1$ since the event "no defective item" is the same as the event" both items are non-defective.

The following events are complementary events

(i)   at most 2 and greater than 2

(ii)   at least 4 and less than 4.

Definition 2.2

If A_1 and A_2 are any two events, $A_1 \cup A_2$ is the event that occurs if t least one of A_1 and A_2 occurs.

In general, if A_1, A_2,..., A_n are any events, $A_1 \cup A_2$ is the event that occurs if at least one of A_1 i = 1, 2,..., n, occurs.

In example 2.1, the event $A_1 \cup A_3$ is the event that an even or a prime number occurs.

$$A_1 \cup A_3 = \{2, 3, 4, 5, 6\}.$$

Thus

$$P(A_1 \cup A_2) = \frac{n(A_1 \cup A_2)}{n\Omega} = \frac{5}{6}$$

The event $A_1 \cup A_2$ is the event that an event or an odd number occurs.

$$A_1 \cup A_2 = \{1, 2, 3, 4, 5, 6\} = \Omega,$$

thus

$$P(A_1 \cup A_2) = P(\Omega) = \frac{6}{6} = 1.$$

Theorem 2.2

If A_1 and A_2 are any events, then

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

The proof of this theorem follows directly from the definition

$$P(A_1 \cup A_2) = \frac{n(A_1 \cup A_2)}{n\Omega}$$

$$n(A_1 \cup A_2) = nA_1 + nA_2 - n(A_1 \cap A_2)$$

dividing through by $n\Omega$ the result follows.

Corollary

(i)   If A_1 and A_2 are mutually exclusive event then

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

since A_1 and A_2 are mutually exclusive,

$$A_1 \cap A_2 = \phi$$

hence

$$P(A_1 \cap A_2) = 0$$

(ii)   If A_1 and A_2 are any two events

$$P(A_1 \cup A_2) \leq P(A_1) + P(A_2)$$

In Example 2.1 above,

$$P(A_1 \cup A_3) = P(A_1) + P(A_3) - P(A_1 \cap A_3) = \frac{1}{2} + \frac{1}{2} - \frac{1}{6} = \frac{5}{6}.$$

(iii)   If A and B are any two events defined on the sample space then

$$P(A) = P(A \cap B) + P(A \cap B^c)$$

$$P(B) = P(A \cap B) + P(A^c \cap B)$$

In general, if A_1, A_2,…A_n are any n mutually exclusive events, then

$$P(A_1 \cup A_2 \cup \ldots \cup A_n) = P(A_1) + \ldots + P(A_n).$$

Theorem 2.2 can be generalized for n > 2 events. It can easily be shown that for any 3 events A_1, A_2, A_3

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

$$- \quad P(A_1 \cap A_3) - P(A_2 \cap A_3) + \ldots + P(A_3).$$

To see this, let $B = A_1 \cup A_2$, then

$$P(A_1 \cup A_2 \cup A_3) = P(B \cup A_3) = P(B) + P(A_3) - P(B \cap A_3)$$

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

$$P(B \cap A_3) = P(A_3 \cap B) = P\{(A_3 \cap (A_1 \cup A_2)\} = P\{(A_3 \cap A_1) \cup (A_3 \cap A_2)\}$$

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

42

(by theorem 2.2). thus

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

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

Note:

$A_1 \cup A_2 \cup A_3$ is the event that at least one of them will occur.

Definition 2.3

Two or more events are said to be exhaustive if their union equals the whole sample space. In other words, A_1, A_2,…, A_n are exhaustive events if

$$P(A_1 \cup A_2 \cup \ldots \cup A_n) = 1$$

In other words it is certain that at least one of them will occur.

Note:

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

2. $P(\Omega) = 1$.

3.   For an sequence of mutually excusive events A_1, A_2,…

$$P(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i).$$

Examples:

2.5   A box contains 6 balls numbered 1 to 6. A ball was drawn from the box at random. Find the probability that the number on the ball drawn was either 1, 2, or 6.

Solution

Let A_1, A_2, A_3 denote the events that the ball drawn was 1, 2 and 6 respectively.

$A_1 \cup A_2 \cup A_3$ denote the event the number on the ball drawn was either 1, 2 or 6.

A_1, A_2 and A_3 are mutually exclusive events. Thus

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

$$= 1/6 + 1/6 + 1/6 = 12.$$

2.6   Four fair dice are tossed once, what is the probability that the sum of the numbers on the four dice is 23?

Solution

The possible outcomes to get 23 are

(6,6,6,5), (5,6,6,6), (6,5,6,6), (6,6,5,6).

The sample space Ω consists of 6^4 elements.

Therefore, the probability that the sum of the numbers on the four dice is 23 is $4/6^4$.

2.7   If $P(A_1) = 2/3$, $P(A_1 \cap A_2) = \frac{1}{4}$ and $P(A_1 \cup A_2) = 5/6$, find $P(A_2)$.

Solution

From the addition law of probability
$$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)$$
We have
$$P(A_2) = P(A_1 \cup A_2) \ P(A_1 \cap A_2) - P(A_1) = 5/6 + \frac{1}{4} - 2/3 = 5/12.$$

2.8   Suppose a fair on is tossed three times, what is the probability that at least one head occurs?

Solution

Let A_1 be the event that the first toss lands heads, A_2 the event that the second toss lands, and A_3 the event that the third toss lands heads.

$A_1 \cup A_2 \cup A_3$ is the event that at least one head occurs.

We are required to compute $P(A_1 \cup A_2 \cup A_3)$ The complement of $A_1 \cup A_2 \cup A_3$ is $A_1^c \cap A_2^c \cap A_3^c$. That is, the first does not land heads, the second does not and the third does not.

$$A_1^c \cap A_2^c \cap A_3^c = \{T, T, T\}.$$

Thus

$$P(A_1^c \cap A_2^c \cap A_3^c) = 1/8.$$

Hence

$$P(A_1 \cup A_2 \cup A_3) = 1 - 1/8 = 7/8.$$

2.9   If $A_1 \subset A_2$, then $P(A_1) \leq P(A_2)$

$$A_2 = A_1 \cup (A_2/A_1).$$

Since A_1 and A_2/A_1 are mutually exclusive, we have

$$P(A_2) = P(A_1) + P(A_2/A_1) \geq P(A_1)$$

Since $P(A_2/A_1) \geq 0$.

2.9   If $A_1, A_2, \ldots, A_n$ are n events, then

$$P\left(\bigcup_{i=1}^{n} A_1\right)^c = P\left(\bigcup_{i=1}^{n} A_1^c\right)$$

From de Morgan's law, we have

$$\left(\bigcup_{i=1}^{n} A_1\right)^c = \bigcap_{i=1}^{n} A_1^c$$

hence

$$P\left(\bigcup_{i=1}^{n} A_1\right)^c = P\left(\bigcup_{i=1}^{n} A_1^c\right)$$

2.2   Conditional Probability

The conditional probability of an event A given that an event B has occurred is denoted as $P(A|B)$. the word "given" is represented by the upright stroke. We wish to determine

the probability that an event A will occur "conditional on" the knowledge that another event B has occurred.

Suppose a fair die is rolled and it is known that an even number appeared uppermost. Let A be the event that the number was greater than 3 and B the event that the number that appeared was even. The problem is to find the conditional probability that the event A occurred given that the event B has occurred, P(A|B). since we know that the number was even, the number must be either 2,4 or 6. Therefore, the conditional sample space contains 3 elements. The event A occurs if the number showing is 4 or 6, thus

$$P(A|B) = 2/3.$$

We can therefore define conditional probability of A given B as the number of ways AnB can occur divided by number of elements in the conditional sample space. That is,

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

Where $n\Omega_B \neq 0$, where Ω_B is the condtional sample space given that B has occurred.

Divide the numerator and denominator by $n\Omega$

$$P(A|B) = \frac{\frac{n(A \cap B)}{n\Omega_B}}{\frac{nB}{n\Omega_B}} = \frac{P(A \cap B)}{P(B)}.$$

In the above example,

$$n(A \cap B) = 2$$

$$n(B) = 3$$

Definition 2.4

Let A and B be two events such that P(B) > 0. Then the conditional probability of A given B, denoted by P(A|B) is defined to be $P(A|B) = \frac{P(A \cap B)}{P(B)}$ 2.2

Examples

2.11   Suppose a box contains 4 red balls and 3 black balls. Compute the probability that

     (i)   the second ball drawn is red if the first ball drawn was red; without replacement,

     (ii)   the second ball drawn is red if the first ball drawn was black

Solution

If the first ball drawn is red, there remains 6 balls, 3 red balls and 3 black balls. The probability of the second ball being red is $\frac{3}{6}=\frac{1}{2}$. But if the first ball is black, the box is left with 4 red and 2 black so the probability of the second ball being red is then 4/6 = 2/3. Thus

$$P(2^{nd} \text{ is re/first was red}) = 3/6/ = ½$$

$$P(2^{nd} \text{ is red/ first was black}) = 4/6 = 2/3.$$

This shows that the probability of the event "the second ball drawn is red" depends on the colour of the first ball drawn.

2.12 Suppose two fair dice are rolled. If the sum of the numbers appearing is 6, what is the probability that one of the number is 2?

Solution

Let A be the event "one of the numbers is and B the sum is 6. there are five ways for the event B to occur: (3,3), (2,4), (4,2), (5,1) and (1,5) and there are two ways for the event AnB to occur: (2,4) and (4,2).

Thus,

$$P(A \cap B) = 2/36, P(B) = 5/36$$

Hence,

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

2.13 There are two children in a family. If there is at least a girl in this family, what is the conditional probability that both are girls.

Solution

The sample space is

$$\Omega = \{BB, GB, BG, GG\}$$

Let A be the event "both children are girls and B "a least a girl in the family.

B = {GB, BG, GG}, A = {GG}, A∩B = {GG}.

$$P(A \mid B) = \frac{P(AnB)}{P(B)} = \frac{1/4}{3/4} = \frac{1}{3}.$$

2.14 There are three children in a family. If there is at least one boy and at most two boys in this family. What is the conditional probability that there are exactly two boys in this family.

The sample space is

$$\Omega = \{BBB, BBG, BGG, BGB, GBB, GBG, GGB, GGG\}$$

Let B be the event "at least one boy and at most 2 boys in the family" and let A be the event "exactly two boys in the family". Then

$$B = \{BBG, BGG, BGB, GBB, GBG. GGB\}$$

$$A \cap B = \{BBG, BGB, GBB\}.$$

Therefore

$$P(A \mid B) = \frac{P(AnB)}{P(B)} = \frac{3/8}{6/8} = \frac{3}{6} = \frac{1}{2}.$$

Exercises 2.1

1. A fair die is thrown twice.

(i) If it is known that the sum of the numbers appearing was 8, what is the probability that the difference between the two numbers was 2.

(ii) If it is known that the difference the two numbers was 3, what is the probability that the sum of the two numbers was 7?

2. Two unbiased dice are thrown once. What is the probability that

(i) at least one 5 is thrown.

(ii) the sum is 10

(iii) the sum is 10 given that no 5 is thrown?

3. Suppose events A and B are sun that $P(A) = 1/5$, $P(AnB) = 1/6$.

Find (i) $P(BIA)$; (ii) $P(A^c \cup B^c)$.

4. If A and B are two events defined on the same probability space, show that: (i) $P(A) = P(A \cap B) + P(A \cap B^c) = P(B) + P(A \cap B^c) - P(A^c \cap B)$.

5. Prove that $P(A^c \cap B^c) = 1 - P(A) - P(B) + P(A \cap B)$.

6. Suppose a well balanced coin is tossed twice.

Find the conditional probability that

(i) both coins show a tail given that the first shows a head;
(ii) both are heads given that at least one of them is a head.

7. A red die and a green die are rolled once. Find the conditional probability that:

(i) the number on red die is odd, given that the sum of the two numbers showing is 9;

(ii) the sum of the two numbers is 9 given that one of the numbers is odd and the other even?

2.3   Bayes Theorem

suppose a box contains r red balls and b black balls. Two balls are drawn at random without replacement. Assume that the probability of drawing any particular ball is $\frac{1}{r+b}$.
Let A_1 be the event "the first ball drawn is red and let A_2 be the event" the second ball drawn is red. Then

$$P(A_1)=\frac{r}{r+b}, P(A_2 \mathrm{I} A_1)=\frac{r-1}{r+b-1}$$
$$P\left(A_2 \mathrm{I} A_1{}^c\right)= \frac{r}{r+b-1}$$

The probability of the event A_2 depends on A_1 and $A_1{}^c$.

That is A_2 is equivalent to A_2 n $A_1 \cap A_1{}^c$

$$A_2 = A_1 \cap A_2 \text{ or } A_1{}^c \cap A_2.$$

Therefore,

$$P(A_2) = P(A_2 \cap A_1) + P(A_2 \cap A_1{}^c) = P(A_1)\, P(A_1)\, P(A_2 \mathrm{I} A_1) + P(A_1{}^c)\, P(A_2 \mathrm{I} A_1{}^c).$$

Theorem 2.3

If the probability of an event B, depends on k mutually exclusive and exhaustive events A_1, A_2 +…+ A_k, then

$$P(B) = P(B \cap A_1) + P(B \cap A_2) +…+ P(B \cap A_k)$$

$$= \sum_{i=1}^{k} P(B \cap A_1) = \sum_{i=1}^{k} P(A_1)\, P(B \mathrm{I} A_1).$$

Example 2.15

Suppose a box contains 3 red balls, 2 black balls and 5 green balls. Two balls are drawn at random without replacement. Find the probability that the second ball drawn is red. Let A_1 be the event "the first ball drawn is red A_2 the event the first ball drawn is black and

A$_3$, the event "the first ball drawn is green. Let B be the event "the second ball drawn is red. The event B occurs if

(i)   the first ball is red and the second is red or

(ii)   the first ball is black and the second is red or

(iii)   the first ball is green and the second is red.

Thus

$$B = B \cap A_1 \text{ or } B \cap A_2 \text{ or } B \cap A_3, P(B \mid A_1) = 2/9, P(B \mid A_2) = 3/9,$$

$$P(B \mid A_3) = 3/9.$$

P(B) depends on A$_1$, A$_2$, A$_3$ which are mutually and exhaustive events. Therefore,

$$P(B) = P(B \cap A_1) + P(B \cap A_2) + P(B \cap A_3)$$

$$= P(A_1) \, P(B \mid A_1) + P(A_2) \, P(B \mid A_2) + P(A_3) \, P(B \mid A_3) = 3/10.2/9 + 2/10.3/9 + 5/10.3/9$$

since $P(A_1) = 3/10, P(A_2) = 2/10, P(A_3) = 5/10$.

Thus

$$P(B) = 1/90 \, (6 + 6 + 15) = 27/90 = 3/10.$$

Example 2.16

Suppose a factory has three machines M$_1$, M$_2$, M$_3$ which produce 60%, 30% and 10% of the total production respectively. Of their output, machine M$_1$ produces 2% defective items, machine M$_2$ produce 3% defective items while machine M$_3$ produces 4% defective items. Find the probability that a part selected at random is defective.

Solution

Let B be the event "a part selected at random is defective". A defective item could have been produced by either machine M_1, M_2 or M_3. Thus

$$B = (B \cap M_1) \cup (B \cap M_2) \cup (B \cap M_3)$$

Since $(B \cap M_1)$, $(B \cap M_2)$, $(B \cap M_3)$ are mutually exclusive events. The following information is contained din the equation.

$$P(M_1) = 60\% = 0.6, \ P(M_2) = 0.3, \ P(M_3) = 0.1$$

$$P(B|M_1) = 0.02, \ P(B|M_2) = 0.03, \ P(B|M_3) = 0.04$$

Hence

$$P(B) = P(M_1) \ P(B|M_1) + P(M_2) \ P(B|M_2) + P(B|M_3) \ P(M_3)$$

$$=(0.6 \times 0.02) + (0.3 \times 0.03) + (0.1 \times 0.04) = 0.002 + 0.009 + 0.004 = 0.025.$$

suppose you are now asked, what is the probability that a given defective part was produced by machine M_1. that is, you are to find $P(M_1|B) = P$ (a part was produced by machine M_1 given that the part was defective).

NOTE:

1. $P(M_1 \mid B) \neq P(B|M_1)$, but
2. $P(M_1 n B) = P(BnM_1)$

$$P(B) = \sum_{i=1}^{3} P(M_1 \ n \ B)$$

hence,

$$P(M_1|B) = \frac{P(M_i \ n \ B)}{\sum_{i=1}^{3} P(M_i \ n \ B)} = \frac{P(M_i) \ P(B \mid M_i)}{\sum_{i=1}^{3} P(M_i) \ P(B \mid M_i)} \qquad 2.4$$

Thus,

$$P(M_1|B) = \frac{0.6 \times 0.02}{0.025} = \frac{0.012}{0.025} = 0.48$$

$$P(M_2|B) = \frac{0.3 \times 0.03}{0.025} = \frac{0.009}{0.025} = 0.36$$

$$P(M_3|B) = \frac{0.1 \times 0.04}{0.025} = \frac{0.004}{0.025} = 0.16$$

Equation (2.4) is an example of Bayes theorem which may be stated as follows.

Bayes Theorem

If A_1, A_2,…A_k are set of mutual exclusive and exhaustive events in a sample space Ω and B is any other event in Ω such that $P(B) > 0$, then

$$P(A_i|B) = \frac{P(A_i)P(B \mid A_i)}{\sum_{i=1}^{k} P(A_i)P(B \mid A_i)}; i = 1, 2, ..., k \qquad 2.5$$

Example 2.17

Suppose a college is composed of 70% male and 30% female students. It is known that 40% of the male students and 20% of the female students smoke cigarette. Find the probability that a student observed smoking a cigarette is male?

Let M,F denote male and female respectively and S denotes smoker. The above problem contains the following information.

P(M) = P(A student selected at random is male) = 0.7

P(F) = P(A student selected at random is female) = 0.3

P(S|M) = P(a student selected at random smokes given that the selected student is male) = 0.4

P(S|F) = P(a student selected at random smokes given that the selected student is female)

= 0.2

By Bayes theorem, we have

P(M|S) = P(a student observed smoking is male)

= P(A student selected at random is male given that the selected student is a smoker)

$$= \frac{P(M \cap S)}{P(S)}$$

Where P(S) = P(a student selected at random is a smoker)

$$= P(S \cap M) + P(S \cap F) = P(M)P(S|F)$$

$$= (0.7 \times 0.4) + (0.3 \times 0.2) = 0.28 + 0.06 = 0.34$$

Thus

$$P(M|S) = \frac{P(M)P(S \mid M)}{P(S)} = \frac{0.7 \times 0.4}{0.34} = \frac{0.28}{0.34} = \frac{14}{17}.$$

Example 2.18

A table has drawers. Drawer 1 contains two red and five black biros, drawer II contains four red and three black biros and drawer III contains one red and six black biros. A drawer is chosen at random and a biro is chosen from the drawer. Find the probability that

(i)   the biro chosen is black

(ii)   the biro chosen is from drawer I if the chosen biro is black.

Solution

Let P(i) denote the probability that drawer I is selected (I – 1, 2, 3) and R and B representing red and black biros respectively.

Then

$$P(1) = 1/3, P(2) = 1/3, P(3), = 1/3$$

$$P(B|1) = 5/7, P(B|2) = 3/7, p(B|3) = 6/7$$

From equation (2.3)

(i)   P(B) = P(Bn1) + P(Bn2) + P(Bn3) = P(1)P(B|1) + P(2)P(B|2) + P(3)P(B|3)

= 1/3.5/7 + 1/3.3/7 +1/3.6/7 = 1/3 (5/7 + 3/7 + 6/7) = 2/3.

(ii)   P(1|B) = P(the biro chosen came from drawer 1 given that the chosen biro is black).

$$= \frac{P(1 \cap B)}{P(B)} = \frac{P(B \cap 1)}{P(B)} = \frac{P(1)P(BI1)}{P(B)} = \frac{\frac{1}{3} \cdot \frac{5}{7}}{\frac{2}{3}} = \frac{5}{14}$$

Definition 2.5

Let A, B and C be three events such that P(C) > 0. then the conditional probability of A $\cup$ B given C is defined by

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

Exercise 2.2

1.   Suppose that a box contains 5 balls labeled 1 to 5. two balls are drawn at random (one after the other without replacement).

What is the probability that

(i)   the sum of the numbers on the two balls selected is even?

(ii)   The number on the first ball drawn is even if it is know that the sum of the two numbers is even.

2.   In a large population, it is observed that 30 percent of the people that are black have caner and 25 percent of the people are not black have cancer. Assume that 10 percent of the population is black. What is the probability that a person selected at random and found to have cancer is not black.

3.   A vaccine produces immunity against smallpox in 98 percent of cases.

Suppose that, in a large population, 20 percent have been vaccinated. Find the probability that a person who contracts smallpox has been vaccinated, assuming that a vaccinated person without immunity has the same probability of contracting smallpox as an unvaccinated person.

4.   In a factory machines A,B,C produce respectively 20, 30 and 50 percent of the total production. Of their output 3, 4, 5 percent respectively are defective items. An item is drawn at random from the total production and is found defective. What are the probability that it was produced by machine A, B,C?

5.   In a faculty of a certain college, 60% of the students are female: 20% of the females and 50% of the male are studying mathematics. If a student data card is selected at random and the student is found to be studying mathematics, what is the probability that the selected student is a male?

6.   In JAMB examination each question has 5 possible answers, exactly one of which is correct. If a student knows the answer he selects the correct answer. Otherwise he selects one answer at random from the 5 possible answers. Suppose that the student knows the answer to 70% of the questions.

(i)   What is the probability that on a given question the student gets the correct answer?

(ii)   If the student gets the correct answer to a question, what is the probability that he knows the answer?

(iii)   what is the expected score of the student in the examination?

7.   A television set retailer finds out that 80% of this customers buy coloured T.V., and that 4 out of every 20 customers who buy coloured T.V set also buy antenna. Calculate the probability that:

(i)   a randomly selected customer buys antenna

(ii)   a randomly selected customers who buys antenna has bought a

colored T.V

(iii)   a randomly selected customer who has not bought an antenna has bought a colored T.V set.

8.   In a factory, machines A, B and C produce 20, 30 and 40 percent of the total output, respectively. Of their output 5,4 and 3 percent respectively are defective bolts. A bolt is chosen at random and found to be defective. What is the probability that the bolt came from machine (i) A (ii) B (iii) C?

9.   Four percent of an article manufactured by a company are defective. All the articles produced are regularly inspected and those found defective are rejected. It is found that one out of every eight of defective article produced is missed by the inspector, while every good article passes inspection. What is the probability that a customer buys a defective article produced by the company?

2.4   Independence

The notion of independence is a basic tool of probability theory. Consider tossing a die twice, and let A_1 be the event that the first toss gives an even number and A_2 the event that the second toss gives an even number. The event $A_1 \cap A_2$ is the event that both tosses give even numbers. The occurrence of A_1 does affect the probability of A_2 occurring.

Therefore

$$P(A_2|A_1) = P(A_2) \text{ and } P(A_1 \cap A_2) = P(A_1)P(A_2|A_1) = P(A_1)P(A_2).$$

Hence, we say that two events A_1 and A_2 are independent if

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

Example 1.19

Toss a fair die twice and let A be the event "the first toss shows 3" and B be the event the sum of the two numbers showing is 7.

$$P(A) = \frac{1}{6},$$

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

$$P(B) = \frac{3}{36} = \frac{1}{6}$$

$$A \cap B \equiv (the\ first\ is\ 3\ and\ sum\ is\ 7) = \{(3,4)\}$$

Therefore,

$$P(A \cap B) = \frac{1}{36}$$

Thus $P(A \cap B) = P(A)P(B)$ hence, A and B are independent events.

In general case of n events, we have the following definition.

Definition 2.6

The events A_1, A_2,…A_n are independent if

(i) $P(A_i \cap A_j) = P(A_i)P(A_j)$ for all $I \neq j$

(ii) $P(A_i \cap A_j \cap A_k) = P(A_i)P(A_j)P(A_k)$ for all I, j, k such that $i \neq j \neq k$
(2.7)

(n-1) $P(A_1 \cap A_2 \cap A_3 \ldots A_n) = P(A_1 P(A_2) P(A_3) \ldots P(A_n)$

That is, the events $A_1, A_2, \ldots A_n$ are said to be mutually independent if

$$P(A_{n1} A_{n2} \ldots A_{nk}) = P(A_{n1}) P(A_{n2}) \ldots P(A_{nk})$$

For every subsequence of two or more events.

$A_{n1}, A_{n2} \ldots A_{nk}$ is a subsequence of $A_1, A_2 \ldots A_n$ if the subscript integers satisfy

$1 \leq n_1 < n_2 \ldots n_k \leq n$.

If $n = 3$, we have, A_1, A_2, A_3 are independent if

(i) $P(A_1 \cap A_2) = P(A_1) P(A_2)$

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

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

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

Condition (i) is called pairwise independent. We might think that pairwise independence always implies independence. But this is not necessarily so, as illustrated by the following example.

Example 2.20

Let a pair of fair dice be rolled once. Consider the events. A_1 number appearing on the first die is even, A_2 = the number appearing on the second die is odd = {1,3, 5} and A_3 = the difference of the two numbers is even

$A_1 \equiv \{(2, 1), (2, 2), (2, 3), (2, 4), (2, 5), (2, 6), (4, 1), (4, 2), \ldots, (6, 1), (6, 2), \ldots\}$

$A_2 \equiv \{(1, 1), (1, 3), (1, 5), (2, 1), (2, 3), (2, 5), (3, 1), \ldots\}$

$A_3 \equiv \{(1, 1), (1, 3), (1, 5), (2, 2), (2, 4), (2, 6), (3, 1), (3, 3), (3, 5), (4, 2), (4,4),$

$(4, 6), (5, 1), (5, 3), (5, 5), (6, 2), (6, 4), (5, 6)\}$

$A_1 \cap A_2 \equiv \{(2, 1), (2, 3), (2, 5), (4, 1), (4, 3), (4, 5), (6, 1), (6, 3), (6, 5)\}$

$A_2 \cap A_3 \equiv \{(1, 1), (1, 3), (1, 5), (3, 1), (3, 3), (3, 5), (5, 1), (5, 3), (5, 5)$

$A_1 \cap A_2 \cap A_3 = \Phi$.

Thus

$$P(A_1) = \tfrac{1}{2}, \; P(A_2) = \tfrac{1}{2}, \; (P(A_3) = \tfrac{1}{2}$$

$$P(A_1 \cap A_2) = \frac{9}{36} = \frac{1}{4}, \; P(A_1 \cap A_2) = \frac{1}{4}, \; P(A_2 \cap A_3) = \frac{1}{4}$$

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

That is, A_1, A_2, A_3 are pairwise independent but

$$P(A_1 \cap A_2 \cap A_3) = 0 \neq P(A_1)P(A_2)P(A_3) = \frac{1}{8}$$

Since condition (ii) is not satisfied, we conclude that A_1, A_2 and A_3 are not independent.

Definition 2.7

If the events A_1, A_2,…A_n are independent then

$$P(A_1 \cap A_2 \dots A_n) = P(A_1)P(A_2)\dots P(A_n) \tag{2.8}$$

Example 2.21

A man fires 10 shots independently at a target. What is the probability that he hits the target (i) 10times; (ii) at least once

If he has probability 1/3 of hitting the target on any given shot.

(i)   Let A_i be the event "he hits the target at the ith shot" (I = 1, 2, 3,…10)

A_1 ∩ A_2 ∩ …∩ A_10 is the event he hits the target 10 times A_1, A_2,… A_{10} are independent events.

Therefore, the probability of hitting the target 10 times is

$$P(A_1 \cap A_2 \cap \dots, \cap A_{10}) = P(A_1)P(A_2)\dots p(A_{10}) = \frac{1}{3}.\frac{1}{3}\dots\frac{1}{3} = \left(\frac{1}{3}\right)^{10}$$

(ii)   P(hitting the target at least once) = 1 – P(not hitting the target at all).

P(not hitting the target at all) = $P(A_1 \cap A_2 \cap \dots \cap A_{10})$

Where, $A_i \equiv$ not hitting the target at the ith shot. A_1, A_2, …, A_{10} are independent events and $P(A_i) = 1 - \frac{1}{3} = \frac{2}{3}$.

Hence

$$P(\bar{A}_1 \cap \bar{A}_2 \cap \dots \cap \bar{A}_{10}) = P(\bar{A}_1)P(\bar{A}_2)\dots P(\bar{A}_{10}) = \frac{2}{3}.\frac{2}{3}\dots\frac{2}{3} = \left(\frac{2}{3}\right)^{10}.$$

Examples

2.22   Suppose a box contains 5 red and 3 black balls. A ball is chosen at random from the box and then a second ball is drawn at random from the remaining balls in the box. Find the probability that

  (i)   both balls are black

  (ii)   both balls are red

  (iii)   the first ball is black and the second is red

  (iv)   the second is red

  (v)   the second is black.

Solution

Let A_1 be the event "the first ball is red

$\bar{A}_1$ be the event "the first ball is black

$\bar{A}_2$ be the event "the second ball is red

$\bar{A}_2$ be the event "the second ball is black

(i)   P(both balls are black) $= P(\bar{A}_1 \cap \bar{A}_2)$

$$P(\bar{A}_1 \cap \bar{A}_2) = P(\bar{A}_1)P(\bar{A}_2) \mid \bar{A}_1)$$

$$P(\bar{A}_1) = \frac{3}{8}, P(\bar{A}_2 \mid \bar{A}_1) = \frac{2}{7}$$

Thus

$$P(\bar{A}_1 \cap \bar{A}_2) = \frac{3}{8} \times \frac{2}{7} = \frac{3}{28}$$

(ii)   P(both balls are red) $= P(A_1 \cap A_2)$

$$P(A_1 \cap A_2) = P(A_1)P(A_2) \mid A_1)$$

$$P(A_1) = \frac{5}{8}, P(A_2 \mid A_1) = \frac{4}{7}$$

Thus

$$P(A_1 \cap A_2) = \frac{5}{8} \times \frac{4}{7} = \frac{5}{14}$$

(iii) $P(A_1 \cap A_2) = $ P(first ball is black and the second ball is red) $= P(\bar{A}_1)P(\bar{A}_2) \bar{A}_1)$

$$P(\bar{A}_1) = \frac{3}{8}, P(A_2 \mid \bar{A}_1) = \frac{5}{7}$$

Thus

$$P(\bar{A}_1 \cap A_2) = \frac{3}{8} \times \frac{5}{7} = \frac{15}{56}.$$

(iv)　　$P(A_2) = P(A_2 \cap A_1) + P(A_2 \cap \bar{A}_1)$

From (ii) and (iii) we have

$$\frac{5}{14} + \frac{15}{56} = \frac{35}{56} = \frac{5}{8}$$

(v)　　$P(\bar{A}_2) = P(A_1 \cap \bar{A}_2) + P(\bar{A}_1 \cap \bar{A}_2) = \frac{5}{8} \cdot \frac{3}{7} + \frac{3}{28} = \frac{3}{8}.$

2.23　Two fair dice are rolled. Given that the dice show different numbers, what is the probability that at least one die shows a 6?

Solution

Let A be the event: the dice show different numbers.

A ≡ {(1,2), (2,1), (1, 3), (3, 1), (1, 4), (4, 1), (1,5), (5, 1), (1, 6), (6, 1),

　　　(2, 3),…(5, 6), (6, 5)}

B ≡ At least one die show a 6.

　　≡ {(1,6), (6,1), (2, 6), (6, 2), (3, 6), (6, 3), (4,6), (6, 4), (5, 6), (6, 5)}

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

$$P(B \mid A) = \frac{10}{30} = \frac{1}{3}$$

2.24　Let A and B be any two events defined on the same sample space. Suppose P(A) = 0.3 and P(A∪B) = 0.6. Find P(B) such that

(i)　　A and B are independent

(ii)　　A and B are mutually exclusive.

Solution

(i)　　If A and B are independent, then $P(A \cap B) = P(A)P(B)$

Thus

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

We have

$$0.6 = 0.3 + P(B) - 0.3\, P(B) = 0.3 + 0.7\, P(B)$$

$$P(B) = \frac{0.3}{0.7} = 0.43$$

(ii)   If A and B are mutually exclusive, then

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

Thus

$$P(A \cup B) = P(A) + P(B)$$
$$0.6 = 0.3 + P(B)$$
$$P(B) = 0.3$$

2.25   Two women A, and B share an office with a single telephone. The probability that any call will be for A is $\frac{2}{3}$. Suppose that A is out of her office during the office hours half of the time and B one third. Find the probability that for any call during the working hours

(i)   no one is in to answer the call

(ii)   A call can be answered by the person being called

(iii)   Two successive calls are for the same woman

(iv)   A caller who wants A has to try more than two times to get her.

Solution

(i)   P(A and B are not in the office) $= \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{6}$.

(ii)   P(A call can be answered by the person being called)

= P(call for A and A is in the office) + P(call for B and B is in the office)

$= \frac{2}{3} \cdot \frac{1}{2} + \frac{1}{3} \cdot \frac{2}{3} = \frac{5}{9}$

(iii)   P(for AA) + P(for BB) $= \frac{2}{3} \cdot \frac{2}{3} + \frac{1}{3} \cdot \frac{1}{3} = \frac{5}{9}$.

(iv) $P(X > 2) = 1 - P(X = 1) - P(X = 2)$.

Where X number of time for a caller who wants A has to try to get her. Thus,

$$P(X > 2) = 1 - \left(\frac{2}{3} \cdot \frac{1}{2}\right) - \left(1 - \frac{2}{3} \cdot \frac{1}{2}\right)\left(\frac{2}{3} \cdot \frac{1}{2}\right) = 1 - \frac{1}{3} - \frac{2}{9} = \frac{4}{9}.$$

Exercise 2.3

1. (a) Show that if A and B are independent events, then

   (i) A and B^c, (ii) A^c and B^c are also independent.

2. Let A and B denote two independent events such that A is a subset of B. prove that either $P(A) = 0$ or $P(B) = 1$.

3. A man fires 10 shots independently at a target. The probability of hitting the target at any shot is 1/3. Calculate the probability that

   (i) noneof the shots hits the target (ii) At least one shot hits the target

   (iii) The target is hit at least twice if it is know that it is hit at least once.

4. A box contains 6 red balls and 4 white balls. Three balls are drawn from the box one after the other without replacement. Find the probability that

   (i) the first two are white and the third red

   (ii) the first two are white and the third white.

   (iii) two are red and one is white

   (iv) the second ball drawn is red

   (v) the third ball drawn is white

5. A die is rolled 8 times. What is the probability that

   (i) exactly 2 sixes appear.

   (ii) at least 2 sixes appear.

   (iii) at most 2 sixes appear.

6. Prove that if $A_1,\ldots,A_n$ are independent events then

   (i) $P(A_1 \cup A_2 \cup \ldots \cup A_n) = 1 - [1 - P(A_1)]\{1 - P(A_2)]\ldots[1 - P(A_n)]$

   (ii) $P(A_1 \cup A_2 \cup \ldots \cup A_n) \le 1 - e^{-[P(A_1) + P(A_2) + \ldots + P(A_n)]}$

   Hint: $(1 - x_1)(1 - x_2)\ldots(1 - x_n) \ge e^{-(x_1 + x_2 + \ldots + x_n)} x_i \le 1$

7. Show that $P(A \cap B \cap C) = P(A \mid B \cap C)P(B \mid C)P(C)$.

8. A die is tossed n times. What is the probability that a 6 appears at least two times in the n tosses.

9. Suppose that A or B occurs is 0.7 while $P(A) = 0.2$, find $P(B)$.

10. A boy decides to continue tossing a fair coin until he has thrown total of three heads. Find P_n, the probability that exactly n tosses will needed.

11. Six blood samples are selected from 40 blood samples, of which four are cancerous. What is the probability that exactly two of the blood samples selected are cancerous?

12. Prove that if $A_1, A_2, \ldots A_n$ are any n events, then

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

13. Prove that

    (i) $P(A \cap B^c) = P(A) - P(A \cap B)$

    (ii) $P(A \cap B^c) = 1 - P(A) - P(B) \quad P(A \cap B)$

    (iii) $P(A) = P(A \cap B) + P(A \cap B^c)$

    (iv) $P(A \cap B) \geq P(A) + P(B) - 1.$

14. three coins have probability 0.5, 0.6 and 0.8 for heads respectively. One of then is selected at random, that is, with equal chance for each, and tossed. If the outcome is head, what is the probability that the coin with probability 0.8 for heads was selected?

15. On a certain weekend there are 4 movies. Calculate the probability that at least one of A and B will be selected by one or more of the 3 students.

16. A box contains n white balls numbered 1 to n, n black balls numbered 1 to n, and n red balls numbered 1 to n. if two balls are drawn at random without replacement, what is the probability that both balls will be of the same colour or bear the same numbers.

17. On the first round, three fair coins are flipped at random. The coins resulting in heads are flipped at random on the second round. If the second round results in exactly one head, what is the conditional probability that the first round ended inexactly two head?

UNIT THREE
DISCRETE RANDOM VARIABLES

3.1 Introduction

This chapter introduces idea of a random variable and its probability density function. A random variable is a variable whose actual numerical value is determined by chance. There are two easily identifiable types of random variables, discrete and continuous. A discrete variable is one that takes only a limited number of possible values, otherwise the variable is called continuous. This chapter is devoted to discrete random variables.

Section 3.2-3.4 are devoted to some special discrete random variables. Bernoulli, Binomial Poisson, uniform,Geometric and negative bionmial.

Example 3.1

Consider variable X, the number of heads in three tosses of a coin. There are four possible values (0, 1,2, 3,) of X.

The actual value assumed is due to chance therefore X is a random variable. The sample space for this experiment is

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

$$X = 0 \text{ if the outcome is TTT}$$

$$X = 1 \text{ if the outcomes is HTT, or THT, or TTH}$$

$$X = 2 \text{ if the outcome is HHT or HTH or THH}$$

$$X = 3 \text{ if the outcome is HHH}$$

Let p be the probability of the con landing tail. Since landing tail and landing head are exhaustive events the probability of the coin landing head is 1-p.

$$P(T \cap T \cap T) = P(T)P(T)P(T)$$

Since the outcome at each trials are independent.

$$P(TTT) = p.p.p = P^3$$

$$P(HTH) = P(T)P(H)P(H) = p (1-P) (1-P) = (1-P)^2$$
$$P(HTH) = P(H)P(H)P(H) = (1-P) (1-P) = P(1-P)^2$$
$$P(HHT) = P(H)P(H)P(T) = (1-P)(1-P)P = = (1-P)^2$$

The probability of getting two heads $= P (THH) + P(HTH) + P(HHT) = 3P (1-p)^2$

Similarly we have

$$P (o\ head) = P(TTT) = p^3$$
$$P (1\ head) = P(HTT) + P(TTH) + P(THT) = 3p^2(1-p)$$
$$P (3\ heads) = P(HHH) = (1-p)^3$$

Thus we have the following table

Heads	0	1	2	3
Probability	P^3	$3p^2 (1-p)$	$3p (1-p)^2$	$(1-p)^3$

$P (0\ head) = P (X\ is\ 0) = p^3$, $p(X=1)\ 3p^2 (1-p)$ and so on

The random variable X defined above is an example of what is called a discrete random variable. A random variable is denoted by a capital letter such as X. Y, Z… and the values that the random variable takes on is denoted by a lower case letter x, y, z…

The notation $(P(X=x)$ means the probability that the random variable X takes on the value x.

Definition 3.1

A random variable X on a sample space Ω is a function assigns to each element Ω one and only one real number $X ( ) = x$, the space ofX is the set of real number
$$\Phi = \{x:x= x ( ),\ we\ \Omega).$$

Definition 3.2

A random variable x is discrete if t can assume at most a finite or a countable infinite number of possible values.

In the above example,
$$\Omega =$$
Where $w_1 = HHH$, $w_2 = HHT…$, $w_8 = TTT$

= {0, 1, 2, 3} and X (w_1) = 3, X (w_2) = 2.X(w_3) = 2, X (w_4) = 2

$$X (w_5 = 1, X (w_6) = 1, X (w_7 = 1, X (w_8) = 0$$

That is {w:X(w) = x_1} id an event

Definition 3.3

The real valued function f defined on R by f(x) = p(X = x) is called the discrete probability density function of X.

Let X be a discrete random variable and suppose that the values it can assume are x_1, x_2...,x_n

The probability can be written as

$P(X = x_1) = f(x_1)$, $_{=P(X} = x_2) = f(x_2)$..., $P(X = x_n) = f(x_n)$

Such that

F(x) is calledprobability density function of X

For an illustration, let us consider the following examples.

Example 3.2

Suppose a pair of fair dice is tossed onc, Let X,Y,Z represent the sum, maximum and minimum respectively of the two numbers appearing find the probability density function of

   (i) X, (ii) Y,   (iii) Z.

Solution

The sample space Ω = {(1, 1), (1, 2) ,..., (6, 6)} consist of 36 elements

     (i)   P (X =2) = P {(1, 1)} = $\dfrac{1}{36}$

          P (X = 3) = P{(1, 2), (2, 1) = $\dfrac{2}{36}$

          P (X = 4) = P{(1,3), (3, 1) , (2, 2)} = $\dfrac{3}{36}$

          .

          .

   P (X = 12) = P {(6, 6)} = $\dfrac{1}{36}$

   Thus,

          F(2)= $\underline{1}$, f(3) = $\underline{2}$, f(4) = $\underline{3}$ ---

$$\frac{}{36} \qquad \frac{}{36} \qquad \frac{}{36}$$

In tabular form we have

X	2	3	4	5	6	7	8	9	10	11	12
F(x)	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

in function form

      (ii)   Y =Maximum of the two number

            Possible values of Y are 1, 2, 3, 4, 5, 6

            $g\,(1) = P(Y=1)= P\,\{(1,1)\}= \frac{1}{36}$

            $g(2) = P(Y = 2) = P\{(\,1, 2), (2, 1)\,(2, 2)= \frac{3}{36}$

            $g(3) = P\,(Y = 3)= P\,(1, 3), (3, 1), (2, 3)\,(3, 2)\,(3, 3)\} = \frac{5}{36}$

            $g(4) = P\,(Y=\$) = P\{(\,1,4)(4,1)\,2, 4)\,(4, 2)\,(3, 4)\,(4, 3)\,(4, 4) = \frac{7}{36}$

            $g(5) = P\,(Y = 5)$

              $= P\,\{(1,5), (5,1), (2, 5)\,(5,2)\,(3, 5)\,(5, 3)\,(4, 5)(5, 4)\,(5, 5) = \frac{9}{36}$

            $g(6) = P\,(Y = 6)$

              $= P\,\{(\,1,6), (6,1)\,(2,6)\,(6,2)\,(3, 6)\,(6, 3)(4, 6)\,(6, 4)\,(5, 6)\,(6,5)\,(6,6)$

        $= \frac{11}{36}$

Putting in form a table we have

Y	1	2	3	4	5	6
g(y)	$\frac{1}{36}$	$\frac{3}{36}$	$\frac{5}{36}$	$\frac{7}{36}$	$\frac{9}{36}$	$\frac{11}{36}$

This can be written in a functional form as

(iii)   Z = Minimum of the two numbers

The possible values of Z are 1, 2, 3, 4,5,6

$P(Z-1) = P\{(1,6), (6,1)(1,5)(5,1)(1,4)(4,1)$

$(1,3)(3,1)(1,2)(2,1)(1,1)! = \underline{11}$
$$36$$

$P(Z = 2) = P (2,6)(6,2)(2,5)(5,2)(2,4(4,2)$

$(2,3)(3,2) (2,2)= 9/36$

$P(Z= 4)P\{(4,6) (6,4) (4, 5) (5,4) (4,4) 5.36$

$P(Z =5) = P(5,6) (6,4) (5,5) 3/36$

$P(Z= 6) = P(6,6) 1/36$

Putting in form of a table we have

Z	1	2	3	4	5	6
h(z)	$\dfrac{11}{36}$	$\dfrac{9}{36}$	$\dfrac{7}{36}$	$\dfrac{5}{36}$	$\dfrac{3}{36}$	$\dfrac{1}{36}$

$$H(Z) = \underline{13 - 2z} \ z = 1, 2, 3, 4, 5, 6$$
$$36$$

$$= 0 \text{ for other values of x.}$$

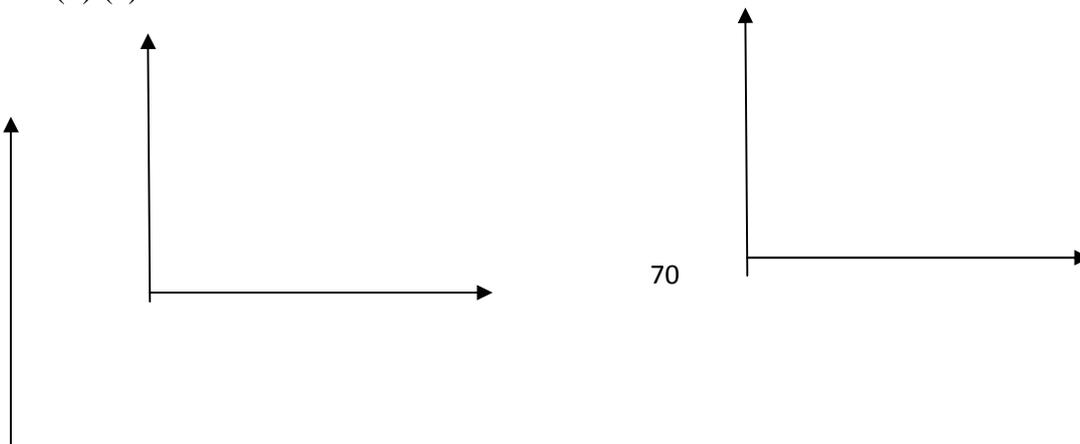The probability density function of a discrete random variable X has the following properties

(i) $O \leq f(x) \leq 1$, x E R

(ii) $\{X:f(x) = 0\}$ is a finite or countable infinite subset of R

(iii) $\sum f(x_{i)} = 1$

The above densities can be represented in terms of a diagram as illustrated in figure (a), (b) (c)

Example 3.3

Let X,Y,Z be the random viable introduced in example 3.1 above

The above three properties are satisfied, properties (i) & (ii) are immediate from definition of probabilities. To check (iii) we have

$\sum f(x_1) = 1/36 + 2/36 + 3/36 + 4/36 + 5/36 + 6/36 + 5/36 + 4/36 + 3/36 + 2/36 + 1/36 = 1$

$\sum G(y_i) = 1/36 + 3/36 + 5/36 + 7/36 + 9/36 + 11/36 = 1$

$\sum h(z_i) = 11/36 + 9/36 + 7/36 + 5/36 + 5/36 + 3/36 + 1/36 = 1$

Example 3.4

Suppose a box contains 1 balls of which 4 are red and 6 are black. A random sample of size 3 is selected. Let X denote the number of red balls selected. Find the probability density function of x if

   (i) Sampling is without replacement

   (ii) Sampling is with replacement

   (iii)   The possible values of X are 0, 1, 2, 3.

$P(X = 0) = (P$ no red ball in the three balls selected$) = \dfrac{^6C_3}{10_{c5}} = \dfrac{6 \; 5 \; 4}{9 \; 8 \; 6} = 1$

$P(X = 1) = P(1$ red and 2 black balls $) = \dfrac{^4C_1 \; x \; ^6C_2}{^{10}C_3} = \dfrac{1}{2}$

$P(X = 2) = \dfrac{^4C_2 x \; ^6C_1}{^{10}C_3} = \dfrac{3}{10}$

$P(X = 3) = ^4C_3/^{10}C_3 = 1$

71

Thus the p.d.f is

X	0	1	2	3
F(x)	$\dfrac{1}{6}$	$\dfrac{1}{2}$	$\dfrac{3}{10}$	$\dfrac{1}{30}$

As a check, we add up all the probabilities

1/6+ ½ + 3/10 + 1/30=1

(ii) Sampling with replacement

P(X = 0) =P (first ball is black, second black and the third black) = P(bbb).

The probability of black at any drawing is

6/10= 3/5

This value is constant since drawing is with replacement. Therefore

$$P(bbb) = \frac{3}{5} \cdot \frac{3}{5} \cdot \frac{3}{5} = \left(\frac{3}{5}\right)^3$$

$$P(X = 1) = P(Rbb) + P(bRb) + P(bbR) = \frac{2}{5} \cdot \frac{3}{5} \cdot \frac{3}{5} + \frac{3}{5} \cdot \frac{2}{5} \cdot \frac{3}{5} + \frac{3}{5} \cdot \frac{3}{5} \cdot \frac{2}{5}$$

$$= 3 \cdot \frac{2 \times 3 \times 3}{5 \times 5 \times 5} = 3 \cdot \left(\frac{2}{5}\right) \left(\frac{3}{5}\right)^2.$$

Similarly ,

$$P(X = 2) = \frac{2}{5} \cdot \frac{2}{5} \cdot \frac{3}{5} + \frac{2}{5} \cdot \frac{3}{5} \cdot \frac{2}{5} + \frac{3}{5} \cdot \frac{2}{5} \cdot \frac{2}{5} = 3 \times \left(\frac{2}{5}\right)^2 \left(\frac{3}{5}\right)$$

$$P(X = 3) = P(RRR) = \frac{2}{5} \cdot \frac{2}{5} \cdot \frac{2}{5} = \left(\frac{2}{5}\right)^3$$

Thus, the p.d.f of X is

X	0	1	2	3
f(x)	$\left(\dfrac{3}{5}\right)^3$	$3\left(\dfrac{2}{5}\right)\left(\dfrac{3}{5}\right)^3$	$3\left(\dfrac{2}{5}\right)^2\left(\dfrac{3}{5}\right)$	$\left(\dfrac{2}{5}\right)^3$

This can be written as

$$f(x) = \begin{cases} \left(\frac{2}{5}\right)^x \left(\frac{2}{5}\right)^{3-x} & x = 0, 1, 2, 3 \\ 0 & \text{elsewhere} \end{cases}$$

Example3.5

A box contains 6 balls labeled 1, 2, 3, 4, 5, 6 two balls are drawn at random one after the other. Let X denote the larger of the two numbers on the balls selected, obtain the probability density function of X if

(i) Sampling is without replacement   (ii) Sampling is with replacement

Solution

(i) The possible values of X are 2, 3, 4, 5, 6. The larger of the two numbers can not be 1 since sampling is without replacement and we can not get same number twice.

$$P(X = 2) = P\{(1, 2), (2, 1)\} = \frac{4}{6.5}$$

$$P(X = 3) = P\{(1, 3), (3, 1), (2, 3), (3, 2)\} = \frac{4}{6.5}$$

$$P(X = 5) = P\{(1, 5), (5, 1), (2, 5), (5, 2), (3, 5), (5, 3), (5, 4), (4, 5)\} = \frac{8}{6.5}$$

Similarly,

$$P(X = 6) = \frac{10}{6.5}$$

Thus the probability density function of X is

X	2	3	4	5	6

| f(x) | 1/15 | 2/15 | 1/5 | 4/15 | 1/3 |

In functional, form, this can be written as

$$f(x) = \begin{cases} \dfrac{2(x-1)}{6.5} = \dfrac{x-1}{15}, & x = 2, 3, 4, 5 \\ \\ 0 & \text{elsewhere} \end{cases}$$

(ii)   The possible values of X are 1, 2, 3, 4, 5, 6. The larger can be 1 in this case since (1, 1) is a possible outcome. The sample space consists 36 possible outcomes

$$P(X = 1) = P(1, 1) = \dfrac{1}{36}$$

$$P(X = 6) = P\{(1, 6), (6, 1), (2, 6), (6, 2).,,, (6, 6)\} = \dfrac{11}{36}$$

Hence the probability density function X is

X	1	2	3	4	5	6
f(x)	1/36	3/36	5/36	7/36	9/36	11/36

In functional form we have

$$f(x) = \dfrac{112x-1}{36}, x = 1, 2, 3, 4, 5, 6$$

3.6   A couple decides that they will continue to have children until either they have a boy and a girl in the family or they have four children. Assuming that boys and girls are equally likely to be born. Let X denote the number of children in the family. Find the probability density function of X.

The number of children in the family can either be 2, 3, or 4.

$$P(X = 2) = P\{(B, G), (G, B)\}$$

That is, the couple will stop having more children if the first child is a boy and the second of girl or the first is a girl and the second a boy.

Thus,

$$P(X = 2) = P(B, G) + P(G, B) = P(B)P(G) + P(G)P(B)$$

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

Similarly,

$$P(X = 3) = P\{(BBG),(GGB)\} = P(BBG) + P(GGB) = \frac{1}{8} + \frac{1}{8} = \frac{1}{4}.$$

Hence, the probability density function of X is

X	2	3	4
f(x)	½	¼	¼

Exercise 3.1

1.   A fair coin is tossed until a head or five tails occur. Let X denote the number of tosses of the coin. Compute the probability density function of X.

2.   A box contains 2 red balls and 3 blue balls. Balls are successively drawn without replacement until a blue ball is drawn. Let X denote the number of draws required. Compute the p.d.f of X

3.   The pdf of a random variable X is given by

X	1	3	4	6	8
f(x)	K	$\frac{k}{2}$	$\frac{k}{3}$	$\frac{k}{4}$	$\frac{k}{5}$

Find (i) K (ii) $P(X \geq 3)$.

4. Toss a fair coin two times. Let X be the number of heads obtained. Find the pdf of X.

5. A coin with probability p of a head is tossed until a head appears. Let X denote the number of times the coin is tossed. Find the pdf of X.

6. A fair die is tossed twice. Let X denote the product of the two numbers appearing. Find the pdf of X

7. A fair coin is tossed 3 times. Let X represent the difference between the number of heads and the number of tails obtained. Find the pdf of X.

8. The pdf of a random variable X is given by

    $f(x) = \{k\, 2^x,\ x = 1, 2, 3, \ldots N,\ \text{zero elsewhere}\}$. Find the value of K.

Definition 3.3

Probability Distribution function. Let X be a random variable with probability density function $f(x)$. The probability distribution function of X denoted $f(x)$ is defined by

$$F(x) = p(X \leq x) \text{ for x real.}$$
$$= \sum_{y \leq x} f(y)$$

Properties of the Probability distribution function

1. F is a non decreasing function, that is if $a < b$, then

    $F(a) < F(b)$.

2. $\lim F(b) = 1$.

    b-∞

    $\lim F(b) = 0$

    b-∞

3. F is right continuous. That is $F(b +) = F(b)$.

Let A be any subset of R and let $f(x)$ be the probability density function of X. ew can compute $P(X \varepsilon A)$ by noting that $\{w : X (w) \varepsilon A\}$ is an event and that

$$\{w : X(w) \varepsilon A\} = U \{ w : X(w) = x_i \}.$$

$$X_i \varepsilon A$$

Thus

$$P(X \, \varepsilon \, A) = \sum_{Xi \varepsilon A} f(X_i).$$

If A is an interval with end points a and b, say A = [a,b].

Then

$$\{(X \, \varepsilon \, A) = P(a \le X \le b) = \sum_{a}^{b} f(x_i).$$

Example 3.7

Consider the random variable X the sum of the two numbers appearing when a fair die is tossed twice of Example 3.2. The probability distribution function for X is given by

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

Suppose we wish to find the probability that X is between 4 and 9 (4 and 9 inclusive) we write it as

$$P(4 \le X \le 9) = P(X = 4) + P(X = 5) + P(X = 6) + P(X = 7) + P(X = 8) + P(X = 9)$$

$$= \frac{3}{36} + \frac{4}{36} + \frac{5}{36} + \frac{6}{36} + \frac{5}{36} + \frac{4}{36} = \frac{27}{36} = \frac{3}{4}.$$

This can also be written as $P(X \le 9) - P(X \le 3) = \frac{30}{36} - \frac{3}{36} = \frac{27}{36} = \frac{27}{36} = \frac{3}{4} = F(9) - F(3)$

3.2   SPECIAL DISCRETE RANDOM VARIABLES

3.2.1   In Section 3.1 we introduced discrete random variables and the probability density function of some random variables were determined. Some variables are so common and important that names are given to them.

In this section we shall consider in considerable detail a number of important discrete random variables.

Bernoulli Random variables

Definition 3.4

Bernoulli Trail:

A random trail or experiment is which the outcome can be classified into one of two mutually exclusive and exhaustive ways usually called success or failure is called a Bernoulli trail. The random variable associated with Bernoulli trail is called a Bernoulli random variable (X).let X=0 if outcome is a failure and X =1 if the outcome is a success. That is any variable assuming only two values is called a Bernoulli random variable.

Suppose that we toss a coin once. Let the probability of it landing head be p. p= $\frac{1}{2}$, if the coin is fair and let denote the outcome of the toss. Then there are two possible values for X, Heads or tails. These two values are mutually, exclusive and exhaustive and we may associate the two possible outcomes of the toss with values 1, 0 of the random variables X. That X = 1 when a head appears and X = 0 when a tail appears.

$$P(X = 1) = p, P(X = 0) = 1 - p.$$

The p.d.f. of X is

X	0	1
f(x)	1 – p	P

Or in functional form

$$F(x) = p^X(1-p)^{1-x}, \quad x = 0, \ 1$$
$$0 \qquad \text{elsewhere} \qquad 3.1$$

f(x) as defined above is called the Bernoulli probability density function and any variable X having (3.1) has its probability density function is called a Bernoulli random variable and is said to have the Bernoulli distribution.

3.2.2 The Binomial Random Variable

This is one of the most important random variables in statistics and the most important discrete random variable. Consider n independent repetitions of Bernoulli trails. Let X_i, I = 1, 2,…, n be Bernoulli random variables associated with the trails. The random

variables X_1, X_2,…X_n are independent Bernoulli random variables. Let us assume the probability of success is p and failure 1-p and

$$P(x_1 = 1) = p$$

Then, sum $S_n = + x_2 + \cdots + X_n$ is the number of successes in n Bemoulli trials. That is, S_n is a counting- variable counting the number of successes in n repeated trials. This random variable S_n is called the Binomial random variable. The possible values of S_n are 0, 1, 2, 3, 4,----, n.

$P(S_n = 0) = p(\text{no success})$

$= p(\text{1sttrail is a failure}) \, P(2^{nd} \text{ trail is a failure})…P(\text{nth trail is a failure}).$

$= P(X_1 = 0) \, P(X_2 = 0) \, P(X_3 = 0)…P(X_n = 0)$

$= (! - p) \, (1 - p)(1 - p)…(1 - p) = (1-p)^n.$

S_n is 1 if the sequence of outcome is

1 0 0 0 0 0 0 …0 or 0 1 0 …0 or 0.0 1 0 …0,…0 0 0 0 0 0 1

$P(S_n = 1) = P(1, 0, 0, 0,…0) + P(), 1, 0, 0,…, 0) + …+ P(0, 0,…1)$

$= P(X_1 = 1)P(X_2 = 0)…P(X_n = 0) + P(X_1 = 0)P(X_2 = 1)…P(X_n = 0) +…+$

$+ P(X_1 = 0)…P(X_n = 1)$

$= P(1 - p)…(1 - p) + (1 - 1)P(1 - p)…(1 - p)…(1 - p)+ (1 - p)(1 - p)$

$= P(1-p)^{n-1} + P(1-p)^{n-1} +…+ P(1-p)^{n-1} = np(1-p)^{n-1}$

Similarly,

$$P(S_n = 2) = {}^nC_2 \, P^2 \, (1-p)^{n-2}$$

Where ${}^nC_2$ is the number of sequences in which exactly 2 have value 1 and the others 0 e.g. (1, 1, 0, 0, 0, 0,0,…, 0), (1, 0, 1, 0,…, 0)…

In general, it can easily be seen that

$$P(S_n = k) = {}^nC_2 \, p^k \, (1-p)^{n-k}, \, k = 0, 1,…, n$$

Where ${}^nC_2$ is the number of sequences in which exactly k have value 1 and others 0.

For example when n = 4, possible sequences of outcomes are given below.

Sequence	S_n	$P(S_n)$

(0,0,0,0)	0	$(1-p)^4$
(1,0,0,0)	1	$P(1-p)^3$
(0,1,0,0)	1	$P(1-p)^3$
(0,0,1,0)	1	$P(1-p)^3$
(0,0,0,1)	1	$P(1-p)^3$
(1,1,0,0)	2	$P^2(1-p)^2$
(1,0,1,0)	2	$P^2(1-p)^2$
(1,0,0,1)	2	$P^2(1-p)^2$
(0,1,1,0)	2	$P^2(1-p)^2$
(0,0,1,1)	2	$P^2(1-p)^2$
(0,1,0,1)	2	$P^2(1-p)^2$
(1,1,1,0)	3	$P^3(1-p)$
(1,1,0,1)	3	$P^3(1-p)$
(1,0,1,1)	3	$P^3(1-p)$
(0,1,1,1)	3	$P^3(1-p)$
(1,1,1,1)	4	P^4

Thus,

$$P(S_n = 0) = (1-p)^4$$
$$P(S_n = 1) = 4p(1-p)^3$$
$$P(S_n = 2) = 6p^2(1-p)^2$$
$$P(S_n = 3) = 4p^3(1-p)$$
$$P(S_n = 4) = p^4$$

Theorem 3.1

Let S_n denote number of successes in n repeated Bernoulli trails, with probability of success p. the probability density function of S_n is given by

80

$$f(x) = P(S_n = x) = {}^nC_x \, p^x (1 - p)^{n-x} \quad x = 0, 1, \ldots, n \qquad (3.2)$$

$$0$$

Definition 3.5

A discrete random variable X denoting total number of successes in n trails is said to have the binomial distribution if

$$P(X = x) = {}^nC_x p^x (1-p)^{n-x} \; ; x = 0, 1, 2, \ldots, n$$
$$0 \le p \le 1$$

The conditions under which binomial distribution will arise are
(i)   The number of trails is fixed
(ii)   There are only two possible outcome 'success' or 'failure' at each trial.
(iii)   The trails are independent
(iv)   The probability p of success at each trail is constant
(v)   The variable is the total number of successes in n trails.

Example 3.8
A soldier fires 10 independently at a target. Find the probability that he hits the target.
(i) once   (ii) at least 9 times   (iii) at most two times.
If he has probability 0.8 of hitting the target at any given time? Let X denote the number of times he hits the target. Then X is a binomial variable with n =10 and p = 0.8
From equation (3.2), we have
$$P(X = x) = {}^{10}C_x \, (0.8)^x \, (0.2)^{10-x}$$
(i) $P(X = 1) = {}^{10}C_x \, 0.8(0.2)^9 = 8(0.2)^9$
(ii)   P(He hits the target at least 9 times) = $P(X \ge 9)$
$$= P(X = 9) + P(X = 10)$$
$$P(X = 9) = {}^{10}C_9 \, (0.8)^9 \, (0.2) = 10 \times (0.8)^9 \times 0.2 = 2(0.8)^9$$
$$P(X = 10) = {}^{10}C_{10}(0.8)^{10} = (0.8)^{10}$$

Hence,
$$P(X \ge 9) = 2(0.8)^9 + (0.8)^{10} = (0.8)^9 (2 + 0.8) = (0.8)^9 (2.8) = 0.3758$$
(iii)   P(at most twice) = $P(X \le 2) = P(X = 1) + P(X = 2)$
$$P(X = 0) = (0.2)^{10};$$
$$P(X = 1) = 8(0.2)^9;$$
$$P(X = 2) = 45 \, (0.8)^2 \, (0.2)^8$$
Thus,
$$P(X \le 2) = (0.2)^{10} + 8(0.2)^9 + 45(0.8)^2 \, (0.2)^8 = 0.00008.$$

Example 3.9
A fair die is rolled four times. Find the probability of getting 2 sixes. Let us call a six a success on a toss of a die and let X be number of sixes (successes) in 4 trails. X is binomial random variable with n = 4 and p = 1/6. Thus,

$$P(X = 2) = {}^4C_2 \left(\frac{1}{6}\right)^2 \left(\frac{5}{6}\right)^2 = \frac{25}{216}.$$

Example 3.10

Suppose that a certain type of electric bulb has a probability of 0.3 of functioning more than 800hours. Out of 50 bulbs, what is probability that less than 3 will function more than 800 hours. Let X be the number of bulbs functioning more than 800 hours. Assuming that X has a Binomial distribution,

$$P(X = x) = 50\,C_x(0.3)^x(0.7)^{50-x}$$
$$P(X < 3) = P(X = 0) + P(X = 1) + P(X = 2)$$
$$P(X = 0) = (0.7)^{50}; \quad P(X = 1) = 50.(0.3)\,(0.7)^{49}$$
$$P(X = 2) = {}^{50}C_2\,(0.3)^2(0.7)^{48}.$$

Thus,

$$P(X < 3) = (0.7)^{50} + 15(0.7)^{49} + 110.25(0.7)^{48} = 0.0000046$$

Exercise 3.2

1. A fair coin is tossed 4 times. Compute the probability that (i) exactly two heads occur (ii) at least 3 heads occur.

2. An investigation reveals that four out of every five patients are cured of malaria when treated with a new drug. If a sample of ten patients is treated by the new drug, compute the probability that (i) exactly six patients are cured, (ii) at most four patients are cured

3. A man fires 12 shots independently at a target. The probability of hitting his target is ¼.

    (i)   What is the probability of his hitting the target at least two times?
    (ii) How many times must he fire so that the probability of his hitting the target at least once is greater than 7/9?

4. Six children are born in a hospital in a given day. Calculate

    (i)   the probability that the number of boys is the same as the number of girls.
    (ii)   the probability that there are more girls tha boys
    (iii) the most likely number of boys.

5.   An experiment has 90 percent probability of success and 10 percent probability of failure. The experiment is repeated four times. Find the probability of obtaining

(i) No success   (ii) No failure   (iii) two successes and two failures.

6.   Let X be a Binomial random variable with paraments (n.p). show that

(i) $P(X = k)$ first increases monotonically and then decreases monotonically.

(ii)   The value of k that maximizes $P(X = k)$ is the largest integer less than or equal to $(n + 1)P$

(iii) $P(X = k) = \frac{n-k+1}{k} \cdot \frac{P}{1-P} \, P(X = k - 1)$.

7.   A fair coin is tossed repeatedly until three are obtained. Find P_n, the probability that exactly n tosses are needed.

8.   Show that (i) $\sum_{x=0}^{n} {}^nC_x p^x (1 - p)^{n-x} = 1$ (ii) $\sum_{x=0}^{n} x\,{}^nC_x p^x (1 - p)^{n-x} = np$
      $0<p<1$.

3.3   Geometric Random Variable

Consider a Bernoulli trail with probability p of a success on one trail. The trail is continued until a success occurs. Let X denote number of trails before the first success. For example, a student decides to continue taking JAMB examination until he passes. X in this case denote number of times he takes the examination before the first success.

The probability that the first $x - 1$ trails are failures and the xth trail is a success is given by $(1 - p)^{x=1}$ p.

To see this, the required probability is

   $P(F\ F\ F \ldots F\ S) = P(F)P(F) \ldots P(F)P(S) = (1 - p)(1 - p)(1 - p) \ldots (1 - p)P = (1 - p)^{x-1}$ p.

The probability that x trails are needed for the first success is the same as the probability that the first $x - 1$ trails are failures and the xth is a success. Thus,

$$P(X = x) = \begin{cases} P(1 - p)^{x-1}, & x = 1,2\ldots \\ 0 & \text{otherwise} \end{cases}$$

Where X is the number of trials before the first success and if we define Y as the number of failures preceding the first success, we have

$$P(Y= y) = \begin{cases} P(1-p)^y & y = 0,1,2... \\ 0 & \text{otherwise} \end{cases}$$

Definition 3.6

A discrete random variable X is called a geometric random variable if its probability density function is given by

$$F(x) = p(X= x) = \begin{cases} (1-p)^{x-1}\,p & x = ,1,2,3... \\ 0 & \text{otherwise} \end{cases}$$

$$0 < p < 1 \qquad (3.3)$$

Where X is the number of independent Bernouilli trials taken for the first success to occur (the successful trial is included in the count). To see that (f(x) is a probability density function, all that needed to be checked is that

$$\sum_{x=1}^{\infty} (1\text{-}p)^{x=1} = 1 = P[1 + (1\text{-}p) + (1\text{-}p)^2 +...]$$

From geometric series,

$$\sum_{x=1}^{\infty} (1\text{-}p)^{x=1} = \frac{1}{1-(1-p)} = \frac{1}{p}.$$

Example 3.11

A fair is tossed until a head appears (a) What is the probability that three tosses are needed? (b) What is the probability that at most three tosses are needed?

Solution

Let X denote the number of tosses until a success (a head) occurs. Since the coin is fair, $p = \frac{1}{2}$.

$$P(X = 3) = (1-p)^{3-1}p \left(\frac{1}{2}\right)^2 \frac{1}{2} = \left(\frac{1}{2}\right)^3$$

P(at most three tosses are needed) = $P(X \le 3)$

$$= P(X = 1) + P(X = 2) + P(X = 3)$$

$$= \frac{1}{2} + \frac{1}{2}\cdot\frac{1}{2} + \left(\frac{1}{2}\right)^2 \cdot\frac{1}{2} = \frac{1}{2} + \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^3 = \frac{7}{8}.$$

3.12 A fair die is rolled until a six appears.

What is the probability that (i) at most 4 rolls are needed. (ii) at least 3 rolls are needed?

Solution

(i) P(at most 4 rolls are needed)= $P(X \leq 4)$

$$= P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4)$$

X is a geometric random variable with $P = \dfrac{1}{6}$

$$P(X = x) = \left(\frac{5}{6}\right)^{x-1} \frac{1}{6}$$

$$P(X = 1) = \frac{1}{6}, \ P(X = 2) = \frac{5}{6} \cdot \frac{1}{6}, P(X = 3) = \left(\frac{5}{6}\right)^{2} \cdot \frac{1}{6}$$

$$P(X = 4) = \left(\frac{5}{6}\right)^{3} \cdot \frac{1}{6}.$$

Thus

$$P(X \leq 4) = \frac{1}{6} + \frac{5}{6} \cdot \frac{1}{6} + \left(\frac{5}{6}\right)^{2} \cdot \frac{1}{6} + \left(\frac{5}{6}\right)^{3} \cdot \frac{1}{6}$$

$$= \frac{1}{6} \sum_{x=1}^{4} \left(\frac{5}{6}\right)^{x-1} = \frac{1}{6} \cdot \frac{1 - (5/6)^{4}}{1 - 5/6} = 1 - (5/6)^{4}$$

(ii) P(at least 3 rolls are needed) $= P(X \geq 3) = P(X = 3) + P(X = 4) + \ldots$

$$P(X \geq 3) = \sum_{x=3}^{\infty} f(x) = \sum_{x=3}^{\infty} \left(\frac{5}{6}\right)^{x-1} \frac{1}{6} = \frac{1}{6} \sum_{x=3}^{\infty} \left(\frac{5}{6}\right)^{x-1} = \frac{1}{6} \cdot \frac{1 - (5/6)^{2}}{1 - 5/6} = \left(\frac{5}{6}\right)^{2} = \frac{25}{36}$$

or

$$P(X \geq 3) = 1 - \{P(X = 1) + P(X = 2)\} = 1 - \frac{11}{36} = \frac{25}{36}$$

Example 3.13

The probability that a certain test yields a "positive" reaction is 0.6. what is the probability that not more than 4 negative reactions occur before the first positive one?

Solution

85

Let X denote the number of negative reactions before the first positive one, then

$$P(X = x) = (0.4)^x(0.6); \; x = 0,1, 2,\ldots$$

Thus,

$$P(X \leq 4) = \sum_{0}^{4} P(X = x) = \sum_{x=0}^{4} (0.4)^X (0.6) = (0.6)\sum_{x=0}^{4}(0.4)^x = (1-(0.4)^5) = 0.9898$$

The geometric random variable has an interesting property which is summarized in the following theorem.

Theorem: 3.2

Suppose that X is a geometric random variable. Then for any given two positive integers s and t,

$$P(X > s + t | X > s) = P(X > t).$$

Proof:

$$P(X > s + t \mid X > s) = \frac{P(X > s + t \; and \; X > s)}{P(X > s)} = \frac{P(X > s+t)}{P(X > s)}$$

$$P(X = x) = p(1- p)^{x-1}, \; x = 1, 2, 3, \ldots$$

$$P(X > s + t) = P \sum_{X=S+t+1}^{\infty} (1-p)^{X-1} = P\frac{(1-p)^{s+t}}{1-(1-p)} = (1-p)^{s+t}$$

$$P(X > s + t) = P \sum_{S+1}^{\infty} (1-p)^{X-1} = \frac{P(1-p)^s}{1-(1-p)} = (1-p)^s$$

Thus,

$$P(X > s + t | X > s) = \frac{(1-p)^{s+t}}{(1-p)^s} = (1-p)^t$$

$$P(X > t) = \sum_{t+1}^{\infty} p(1-p)^{X-1} = \frac{p(1-p)^t}{p} = (1-p)^t.$$

Hence,

$$P(X > s + t \mid X > s) = P(X > t).$$

The above theorem states that if a success has not occurred during the first s repetitions of Bernouli trails, then the probability that it will not occur during the next t repetitions is the same as the probability that it will not occur during the first t repetitions of Bernoulli trails. Therefore the distribution is said to have "no memory".

Exercise 3.3

1. On a certain road the probability of an accident on any day is 0.05 (assuming not more than one accident can occur on any day). Assuming independence of accidents from day to day on this road, what is the probability that the first accident of the year occurs in the month of March.

2. A fair die is rolled until a six appears. Calculate the probability that he has to throw the die more than three times before he gets a six.

3. Let X be a geometric random variable with p = 0.2. Calculate the following probabilities. (i) $P(3 < x \leq 6)$ (ii) $P(2 \leq X \leq 4)$ (iii) $P(X \leq 2)$.

4. Prove that $\sum_{x=1}^{\infty} xp(1-p)^{x-1} = \frac{1}{p}$.

5. Suppose a Bernoulli trail with probability p of success is continued until rth success occurs. Let X be the number of independent trails needed in order to have r successes. Show that

$$P(X = K) = \binom{k-1}{r-1} p^{r}(1-p)^{k-r}, k = r, r+1,...$$

3.13 Pascal or Negative Binomial Random variable

A probability distribution closely related to the geometric distribution is the Pascal or negative binomial distribution.

Suppose that independent repetition of Bernoulli trail needed to have "success" occurs exactly r times.

X = x if and only if success occurs on the xth trail and success occurs exactly (r – 1) times in the previous x -1 trails. The probability of this event is determined as follows:

The probability of r – 1 successes in x – 1 trails is given by the binomial formula

$$^{X=1}C_{r-1}p^{r-1}(1-p)^{x=1-(r-1)=x-1}C_{r-1}\,p^{\,r-1}\,(1-p)^{\,x-r}$$

Therefore, the probability of this event is

$$P.^{x-1}C_{r-1}p^{r-1}(11-p)^{x-r+}$$

hence,

$$f(x) = P(X = x) = {}^{x-1}C_{r-1}\,p^{r}(1-p)^{x-r}, x = r, r+1,\ldots \qquad (3.4)$$

Similarly, if we let Y be the number of failures before the rth success we have

$$P(Y = y) = p. \, (^{y+r-1}C_y p^{r-1}(1-p)^y$$

$$F_Y(y) = {}^{y+r-1}C_y{}^r(1-p)^y, \, y = 0, 1, 2,\ldots$$

Definition 3.7

A random variable having its probability density function given by (3.4) or (3.5) is said to gave a Negative Binomial or Pascal distribution.

Example 3.14

Show that $\sum_{X=r}^{\infty} \, {}^{x-1}C_{r-1}P^r(1-P)^{x-r} = 1.$

Solution

By Binomial theorem,

$$P^{-n} = (1-(1-p)^{-n} = \left\{1+n(1-p)+\frac{n(n+1)}{2!}(1-P)^2 + \ldots\right\}$$

$$\sum_{X=r}^{\infty} \, {}^{x-1}C_{r-1}P^r(1-P)^{x-r} = P^r(1-p)^{-r} \sum_{X=r}^{\infty} \, {}^{(x-1)}C_{r-1}(1-P)^x$$

$$= p^r(1-p)^{-r}\left\{(1-p)^r + r(1-p)^{r+1} + \frac{r(r+1)}{2!}(1-p)^{r+2} + \cdots\right\}$$

$$= p^r\left\{1+r(1-p) \; + \frac{r(r+1)}{2!}(1-p)^2 + \cdots\right\} = p^r p^{-r} = 1.$$

Example 3.15

A fair die is rolled until two sixes occurs, find the probability that

(i) exactly 5 tosses are needed

(ii) at most 5 tosses are needed.

Solution

Let X be the number of tosses needed to get two sixes. X is a Pascal random variable with pdf: $f(x) = {}^{x-1}C_1 p^2(1-p)^{x-2}, \, x = 2,3,\ldots$

Where $p = \frac{1}{2}$, $r = 1$.

$$f(x) = (x-1)\left(\frac{1}{2}\right)^x$$

(i) $P(X = 5) - 4\left(\frac{1}{2}\right)^5 = 0.125$

(ii) $P(X \le 5) = \sum_{x=2}^{5} f(x) = P(X = 2) + P(X = 3) + P(X = 4) + P(X = 5)$

$$P(X = 2) = \left(\frac{1}{2}\right)^2, P(X = 3) = 2\left(\frac{1}{2}\right)^3, P(X = 4) = 3\left(\frac{1}{2}\right)^4.$$

Thus,

$$P(X \leq 5) = \frac{1}{4} + \frac{1}{4} + \frac{3}{16} + \frac{1}{8} = \frac{13}{16} = 0.8125.$$

3.3.2 The Hyper geometric Random Variable

Suppose that we have a box containing n items of which n_1 are defective and $n - n_1$ are non-defective. Suppose that we choose at random k items from the box without replacement. Let X denote the number of defective items in the k items selected. Then

$$P(X = x) = \frac{{}^{n_1}C_x \; {}^{n-n_1}C_{k-x}}{{}^{n}C_x}, x = 0, 1, 2, 3, \ldots k \qquad (3.4)$$

The reader will notice that $X = x$ if and only if we obtain x defective items from n_1 defective items in the box (${}^{n}_1C_x$ ways of doing this).

Any random variable having its probability density function as given by (3.5) is called hyper geometric random variable and is said to have hyper geometric distribution.

3.4 The Poisson Random Variable

Definition: *Rare Event*

An event is said to be rare if the probability p of observing the event is very small.

Consider n repeated Bernoulli trails, where n is very large and P very small, let X be be the number of successes in n trials. Then

$$P(X = x) = {}^{n}C_x p^x (1 - p)^{n-x}.$$

Setting $\lambda = n\,p$, we have

$$P(X = x) = {}^{n}C_x \left(\frac{\lambda}{n}\right)^x \left(1 - \frac{\lambda}{n}\right)^{n-x} = \frac{n(n-1)\ldots(n-x+1)}{x!} \cdot \left(\frac{\lambda}{n}\right)^x \left(1 - \frac{\lambda}{n}\right)^{n-x}$$

$$= \frac{n}{n} \cdot \frac{n-1}{n} \cdot \frac{(n-2)}{n} \ldots \frac{n-x+1}{n} \cdot \frac{\lambda^X}{X!} \left(1 - \frac{\lambda}{n}\right)^{n-X}$$

Now as $n \qquad \infty$.

1. $\left(1 - \frac{\lambda}{n}\right)^n \to e^{-\lambda}$

2. $\left(1 - \frac{\lambda}{n}\right)^{-x} \to 1$

3. $\frac{n-X+1}{n} \to 1$

Consequently,

$$\lim_{n \to \infty} {}^{n}C_x p^x (1 - p)^{n-x} = \frac{\lambda}{X!} e^{-\lambda}$$

(3.5) gives an approximation to the binomial distribution with $\lambda = np$, when n is large and p is small, where e = 2.71828 is the base of natural logarithms.

Definition:

A random variable X is called a Poisson random variable if its probability density function is given by

$$f(x) = \begin{cases} \dfrac{e^{-\lambda t}(\lambda t)^x}{x!} & x = 0, 1, 2, \ldots \\ & (\lambda > 0) \\ 0 & \text{otherwise} \end{cases} \qquad (3.8)$$

X represents the total number of events which have occurred up to time t. When t= 1 then f(x) corresponds to the probability density function of number of events in a unit interval. For examples, the number of calls that come into a telephone exchange in a unit time interval, the number of vehicles passing through a designated point in a unit time interval. In order to motivate our discussion, let us consider the following examples.

Example 3.16

Suppose a rare disease occurs in 2 percent of a large population. A random sample of 10,000 people are chosen at random from this population and tested for the disease. Calculate the probability that at least two people have the rate disease.

Solution

The probability p of having the disease is 0.02 and n = 10,000. Let X be the number of people having the disease. X is a binomial random variable with parameters n= 10,000 and p = 0.02. we shall apply the result (3.5) since n is large and p very small.
Hence,

$$Np = 0.02 \times 10,000 = 20,,$$

Thus,

$$P(X = x) = \frac{e^{-200}(200)^X}{X!}$$
$$P(X =) = e^{-200}$$
$$P(X = 1) = 200\, e^{-200}$$

Hence

$$P(X \geq 2) = 1 - P(X < 2) = 1 - e^{-200}(1 + 200) = 1 - 201\, e^{-200}.$$

Example 3.17

On a given road, an average of five accidents occur every month. Calculate the probability that over a year period there will be (i) no accident (ii) at most 2 accidents.

Solution

In this case, $\lambda = 5$, t = 12. From (3.8) we have

$$f(x) = \frac{e^{-\lambda t}(\lambda t)^X}{X!}$$
$$f(0) = e^{-\lambda t} = e^{-60}$$
$$f(1) = \frac{e^{-\lambda t}(\lambda t)^1}{1!} = 60e^{-60}$$
$$f(2) = \frac{e^{-\lambda t}(\lambda t)^2}{2!} = \frac{60^2}{2}\, e^{-60}$$

thus,

$$P(\text{no accident}) = f(0) = e^{-60}$$

$$P(\text{at most two accidents}) = f(0) + f(1) + f(2) = e^{-60} + 60\, e^{-60} + \frac{60^2}{2} e^{-60}$$

$$= e^{-60} (1 + 60 + 1800) = 1861 \, e^{-60}.$$

Tables for the Poisson distribution are available and brief abulation is given in the Appendix.

Exercises 3.4

1. If 3% of the items manufactured in a factory are defective. Compute the probability that in a simple of 100 items. (i) 2 items will be defective (ii) at least 2 items will be defective. Use Poisson approximation to the Binomial.

2. Use the Poisson approximation to calculate the probability that at least two sixes are obtained when six dice are rolled once.

3. The telephone switchboard of a University has an average of two incoming calls per minutes. Calculate the probability that, over a three-minute interval, there will be (i) no incoming calls (ii) exactly one incoming call. (iii) at most two incoming calls.

4. Let P_r be the negative binomial pdf with parameters r and P. prove that
    $$\frac{P_r(k)}{P_r(k-1)} = 1 + \frac{1}{k}\left[(r-1)\frac{a}{p} - k\right] k = 1, 2, , ..., \text{ where } P_r(k) = P(X = k).$$

5. Prove that $\sum_{x=0}^{n} \frac{e^{-\lambda}\lambda^x}{x!} = \frac{1}{n!}\int_{\lambda}^{\infty} e^{-y}y^n dy.$ Hint: Use integration by parts.

6. A fair die is tossed 20 times. Let a success on the ith toss corresponds to at least a five appears. What is the probability that
    (i)   exactly 10 failure prior to the first success?
    (ii) exactly 10 failures prior to the fifth success?
    (iii) exactly 8 failure and 3 successes for the first II throws.

7. if (i) $f(x) = \frac{e^{-\lambda}\lambda^x}{x!}$, show that $(x + 1)f(x +1)= \lambda \, f(x).$
    (ii) $f(x) = {}^nC_c p^x(1-p)$, show that $(1-p)(x + 1)f(x + 1) = p(n-x)f(x).$
    In each case, show that f(x) increases monotonically and then decreases monotonically as x increases. Find x that maximizes f(x).

8. Suppose that the number of accidents occurring on a highway each day is a Poisson random variable with parameter $\lambda = 2.$
    (a) What is the probability that at least two accidents occur in a day?
    (b) what is the probability that at least three accidents occur in a day given that at least one accident occurs in a day?

9. Describe the Poisson distribution, stating clearly the meanings of the symbols used. Show that the variance of the distribution is equal to its mean.

10. Show that $f(x) = {}^{x-1}C_{n-1}P^n(1-p)^{x-n}$, x, x = n, n+1,… is a pdf.
    Note $P^{-n}= (1 - (1-p))^{-n} = \left\{1 + n(1-p) + \frac{n(n+1)}{2!}(1-p)^2 + \cdots\right\}.$

11. Show that $f(x) = \frac{(1-p)^x}{X \log (1-p)}, x = 1,2, …, 0< p<1$ is a pdf. Find the probability generating function of X.

12. Solve the problem of exercise 3.2, No 7.

13. If $n \geq 8$ and in s binomial distribution the probability of 7 successes in n trails is equal to the probability of 8 successes in n trail, find the probability of success on any given trail.

14. Let X be a random variable with moment generating function $M_X(t)$. If $R(t) = In''$ $M_X(t)$ for all t. Find $R''(0)$.

15. If a fair die is rolled repeatedly, find the probability that the first "six" will appear on an odd-numbered roll.

16. $f(x) = \begin{cases} p, x = -1 \\ \\ (1-p)^2 p^x, X = 0,1,2 \end{cases}$

    (i) Show that f(x) is a pdf

UNIT FOUR

EXPECTATION OF DISCRETE RANDOM VARIABLE

In this chapter we introduce the concept of the mean value of a random variable. It is closely related to the notion of weighted average. The moment and probability generating functions of a random variable are also introduced.

4.1

Let X be a random variable having possible values x_1, x_2, …, x_n, and let the experiment on X be performed n times. For example let X be the outcome of rolling a die. There are six possible values $x_1 = 1$, $x_2 = 2$, $x_3 = 3$, $x_4 = 4$. $x_5 = 5$ and $x_6 = 6$. Suppose the die is rolled n times. The successive rolls constitute independent repetitions of the same experiment. Let X_1, X_2, . …, X_n denote the outcomes of the experiment of rolling a die n times (that is X_i denote the outcome of the ith toss). Then

$$\overline{X} = \frac{X_2 + X_2 + \cdots + X_n}{n}$$

Is the average of the numbers that appeared. Let f_i denote the number of times x_i occur. Then we have

$$\frac{1}{n}(X_1 + X_2 + \cdots + X_n) = \frac{1}{n}\sum_{i=1}^{k} x_i f_i = \sum_{i=1}^{k} x_i \cdot \frac{f_1}{n}.$$

We   know   that

$$\lim_{n \to \infty} \frac{f_i}{n} = P(X = x_i) = f(x_i). \text{ When } \frac{f_i}{n} \text{ is replaced by } f(x_i), \text{ the sum } \sum x_i f(x_i) \text{ is}$$

called the expectation of the random variable X. thus the expected value of a random variable X is the long-run theoretical average value of X.

Definition 1
Mathematical Expectation. Let X be a random variable with probability density function as follows:

Value of X, x	x_1	x_2	x_3	....	x_k
Probability	$f(x_1)$	$f(x_2)$	$f(x_3)$		$f(x_k)$

The mathematical expectation of X, denoted by E(X), is defined to be

$$E(X) = x_1f(x_2) + x_2f(x_2) + \ldots + x_kf(x_k)$$
$$E(X) = \sum_{i=1}^{k} x_i f(x_i).$$

E(X) is weighted average of possible values of x, the weight attached to the value xi is its probability $f(x_i)$. E(X) is also called the mean of X or ***the population mean.***
We may express the result of (1) in words.
To compute the expected value or mean of a random variable, multiply each possible value of the variable by its probability and add these products.

Examples
4.1   A fair coin is tossed three times. Let X denote the number of heads obtained. Find the mathematical expectation of X.

Solution
The probability density function of X is as follows.

X	0	1	2	3
f(x)	$\dfrac{1}{8}$	$\dfrac{3}{8}$	$\dfrac{3}{8}$	$\dfrac{1}{8}$

Hence,
$$E(X) = (1 \times 1/8) + (1 \times 3/8) + (2 \times 3/8) + (3 \times 1/8)$$
$$= 0 + 3/8 + 6/8 + 3/8 = 12/8 = 1.5$$

4.2   What ate the mathematical expectations of the variables X, Y and Z as defined in example 3.1.1 of chapter 3.

Solution
The probability density function of X is

X	2	3	4	5	6	7	8	9	10	11	12
f(x)	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36

$$E(X) = (2 \times 1/36) + (3 \times 2/36) + (4 \times 3/36) + (5 \times 4/36) + (6 \times 5/36) + (7 \times 6/36) +$$

$(8 \times 5/36) + (9 \times 4/36) + (10 \times 3/36) + (11 \times 2/36) + (12 \times 1/36)$
$= 7.0$

The p.d.f of Y is

Y	1	2	3	4	5	6
g(y)	1/36	3/36	5/36	7/36	9/36	11/36

$E(Y) = (1 \times 1/36) + (2 \times 3/36) + (3 \times 5/36) + (4 \times 7/36) + (5 \times 9/36) + (6 \times 11/36)$
$= 4.47$

The p.d.f of Z is

Z	1	2	3	4	5	6
h(Z)	11/36	9/36	7/36	5/36	3/36	1/36

$E(Z) = (1 \times 11/36) + (2 \times 9/36) + (3 \times 7/36) + (4 \times 5/36) + (5 \times 3/36) + (6 \times 1/36)$
$= 2.53.$

Exercises 4.1
1. Suppose a box contains 10 balls of which 4 are red and 6 are black. A random sample of size 3 is selected. Let X denote the number of red balls selected. Find (E(X) if
    (i)   Sampling is with replacement.
    (ii) Sampling is without replacement.
2. A box contains 6 balls labeled 1, 2, 3, 4, 5, 6. Two balls are drawn at random one after the other. Let X denote the larger of the two numbers on the balls selected. Compute E(X) if (i) Sampling is with replacement, (ii) Sampling is without replacement.
3. A box contains 3 balls and 2 white balls. A ball is without replacement one after the other until a white ball is drawn. Find the Expected number of draws required.

Expectation of a Binomial Random Variable
The probability density function of a binomial variable X is defined by
$$f(x) = {}^nC_x p^x (1 - p)^{n-x}; \quad x = 0, 1, 2, \dots, n.$$
the mean or expected value of X is
$$E(X) = \sum_{x=0}^{n} x \;\; {}^nC_x p^x (1-p)^{n-x} = \sum_{x=0}^{n} x . \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}$$
and since the term of x = 0 is zero,
$$E(X) = \sum_{x=1}^{n} \frac{n!}{(x-1)!(n-X)!} p^X (1-p)^{n-x} = np \sum_{x=1}^{n} \frac{(n-1)!}{(x-1)!(n-X)!} p^{X-1} (1-p)^{n-x}$$
Using the binomial theorem, we have

$$\sum_{x=1}^{n} \frac{(n-1)!}{(x-1)!\,(n-X)!} p^{X-1}(1-p)^{n-x} = \sum_{x=0}^{n-1} {}^{n-1}C_x\, p^{x-1}(1-p)^{n-x}$$

$$= (p + (1-P))^{n-1} = 1$$

since p + 1 – p = 1. Hence,

$$E(X) = np.$$

Example 4.3

A fair die is rolled 12 times, what is the expected number of six appear?

Solution

Let X be the number of sixes that appear. X is a binomial random variable with n = 12 and P = 1/6 (probability of a six). Hence

$$E(X) = np = 12 \times 1/6 = 2.$$

That is, we expected to get 2 sixes when a die is rolled 12 times.

Definition 2

Let X be a discrete random variable having probability density function f(x).

$\sum_{i=1}^{\infty} |X_i f(x_i)| < \infty$, then we say that X has finite expectation otherwise we say that X does not have finite expectation.

Examples

4.4 The Random Variables X has Probability

X	-2	-1	0	3	5
f(x)	¼	1/8	1/8	¼	¼

Compute the expected value of the following random variables.

(i)   X,   (ii) 3X,   (iii) X + 5,   (iv) X^2

Solution

(ii) $Ex = \left(-2x\frac{1}{4}\right) + \left(-1x\frac{1}{8}\right) + \left(0x\frac{1}{8}\right) + \left(3x\frac{1}{4}\right) + \left(5x\frac{1}{4}\right)$

$= -2/4 - 1/8 + 3/4 + 5/4 = 11/8 = 1\frac{3}{8}$

(iii)   The possible values of 3x are -6, -3,0, 9, 15.
         The probability density function of 3x is

3x	-6	-3	0	9	15
f(x)	¼	1/8	1/8	¼	¼

E(3X) = (-6 × ¼) + (-3 × 1/8) + (0 × 1/8) + (9 ×1/4) + (5 × ¼)

  = -6/4 – 3/8 + 9/4 + 15/4 = 33/8

Note that

  P(3X = -6) = P(X = -2_,

And so on and

  E(3X) = 33/8 = 3 × 11/8 = 3E(X)

(iii)

x + 5	3	4	5	8	10
f(x)	¼	1/8	1/8	¼	¼

  Note that P(X +5) = 3 =P(X = 3 – 5) = P(X = -2)

  E(X + 5) = (3 × ¼) + (4 × 1/8) + (5 × 1/8) + (8 × ¼) + (10 ×1/4)

  = ¾ + 4/8 + 5/8 + 8/4 + 10/4 + 51/8 = $6\frac{3}{8}$ = 5 + $1\frac{3}{8}$ = 5 + E(X).

(iv)

X^2	4	1	0	9	25
f(x)	¼	1/8	1/8	¼	¼

  E(X^2) = (4 × ¼) + (1 × 1/8) + (0 × 1/8) + (9 × ¼) + (25 × ¼)

  = 1 + 1/8 + 9/4 + 25/4 = 77/8 = $9\frac{5}{8}$

  Note that E(X^2) ≠ $[E(X)]^2$.

Definition 3

Let X be a random variable whose probability density function is given by

x	x_1	x_2	………….	x_k
f(x)	f(x_1)	f(x_2)		f(x_k)

Let ψ(X) be a function of X. then the mean or expected value of the new random ψ(X) is given by

  E[ψ(X)] = ψ(x_1) f(x_1) + ψ(x_2) f(x_2) +…+ψ(x_k) f(x_k)

That is

  $E[\psi(X)] = \sum_{i=1}^{k} \psi(x_i) f(x_i)$

Theorem 1

Let X be a random variable and let ψ(X) = a X + b where a, b are constants, then

$$E(\psi(X)) = aE(X) + b$$

Proof:
Suppose the probability density function of X is
$$\{x_i, f(X_i), I = 1, 2, \ldots, k\}$$

From the above definition we have
$$E(aX + b) = (ax_1 + b)\, f(x_1) + (ax_2 + b)\, f(x_2) + \ldots + (ax_k + b)\, f(x_k) + bf(x_1) + \ldots + bf(x_k)$$
$$= a[x_1 f(x_1) + x_2 f(x_2) + \ldots x_k f(x_k)] + b[f(x_1) + f(x_2) + \ldots +$$
$$f(x_k)] = \sum_{i=1}^{k} x_i f(x_i) + b \sum_{i=1}^{k} f(x_i) = aE(x) + b$$
since $\{E(X) = \sum x f(x),\ \sum f(x_i) = 1\}$

Properties of Excitations
1. If c is a constant and $P(X = c) = 1$, then $E(X) = c$.
2. If a and b are constants and X has a finite expectation, then aX has finite expectation and
   (i) $E(aX) = aE(X) + b$.
   (ii) $E(aX + b) = aE(X) + b$.
3. If X and Y are two random variables having finite expectations, then
   (i) $X + Y$ has random finite expectation and $E(X + Y) = E(X) + E(Y)$
   (ii) $E(X) \geq E(Y)$ if $P(X \geq Y) = 1$.
   (iii) $|EX| \leq E\,|X|$.
4. A bounded random variable has a finite expectation. That is if $P(X \leq M) = 1$, then X has a finite expectation and $|EX| \leq M$.

Definition 4
Variable. Let X be a random variable with mean $E(X) = \mu$.
The variance of X, denoted by Var(X) is defined by
$$Ver(X) = E\{X - \mu)^2\} = \sum (x_t - \mu)^2 f(x_i)$$
A general formula that is usually simple for computing the variable is given below
$$Var(X) = E\{(X - \mu)^2\} = E\{X^2 - 2X\mu + \mu^2\}$$
$$= E(X^2) - E(2\mu X) + E(\mu^2) = E(X^2) - 2\mu E(X) + \mu^2 \text{ (property 2)}$$
$$= E(X^2) - 2\mu^2 + \mu^2 = E(X^2) - \mu^2.$$
Thus, the computing definition of variance is
$$Var(X) = E(X^2) - [E(X)]^2 \qquad\qquad (7)$$
The variance of X is interpreted as a numerical measure of spread or dispersion about its mean.

Example 4.5
1. Find the variance of a random variable having the following probability density function.

x	1	2	3	4	5	6

f(x)	1/6	1/6	1/6	1/6	1/6	1/6

$E(X) = (1 \times 1/6) + (2 \times 1/6) + (3 \times 1/6) + (4 \times 1/6) + (5 \times 1/6) + (6 \times 1/6).$
$\quad = 1/6 + 2/6 + 3/6 + 4/6 + 5/6 + 6/6 = 21/6$

$E(X^2) = (1^2 \times 1/6) + (2^2 \times 1/6) + (3^2 \times 1/6) + (4^2 \times 1/6) + (5^2 \times 1/6) + (6^2 \times 1/6)$
$\quad = 1/6 + 4/6 + 9/16 + 16/6 + 25/6 + 36/6 = 91/6.$
Hence from (7) we have
$\quad\quad Var(X) = 91/6 - (21/6)^2 = 105/36 = 2\,(11/12).$

Example 4.6
Find the variance of a Bernoulli random variable with parameter p. that is X has the following p.d.f.

X	0	1
f(x)	1 – p	P

$\quad\quad E(X) - 0 \times (1\text{-}p) + (1 \times P) = p$
$\quad\quad E(X^2) = 0^2 \times (1\text{-}p) + (1^2 \times p) = p.$

Hence,
$\quad\quad Var(X) = p - p^2 = p(1 - p)$

Properties of Variance
1. If c is a constant and $P(X) = c) = 1$, then Var $(X) = 0$.
2. If a and b are contents, then
    (i) $\quad Var(aX) = a^2\,Var(X)$
    (ii) $\quad Var(aX + b) = a^2\,Var(X)$
    (iii) $\quad Var(X) \geq 0$

The proof of property (1) is very trivial and this is left as an exercise to the reader, we shall now give a proof of (ii). From (7),
$Var(aX) = E[aX^2] - [E(aX)]^2 = E\,a^2\,X^2 - [aE(X)]^2$

$\quad\quad = a^2\,E(X^2) - a^2[E(X)]^2 = a^2\{(E(X^2)]^2\} = a^2\,Var(X).$

Definition 5

Standard deviation. Let X be a random variable with mean μ.

The standard deviation is the positive square root of the variance, and is given by

$$\sqrt{Var(X)} = \sqrt{E(X^2) - (E(X))^2}$$

The variance of the binomial random variable.

To calculate the variable of X, we need $E(X)$ and $E(X^2)$.

$$Var(X) = (E(X))^2.$$

From (3) we have

$$E(X) = np.$$

$E[X(X-1)] = \sum_{x=0}^{n} x(x-1)^n({}_x p^x(1-p)^{n-x} = \sum_{k=0}^{n} x(x-1)\frac{n!}{x!(n-x)!} p^x(1-p)^{n-x}$

(w2hen x = 1 or 1 the expression is zero)

$$= \sum_{x=2}^{n} \frac{n!}{(x-2)!(n-x)!} p^x(1-p)^{n-x}$$

$$= n(n-1)\, p^2 \sum_{x=2}^{n} \frac{(n-2)!}{(x-2)!(n-x)!} p^{x-2}(1-p)^{n-x}$$

Let $y = x - 2 = n(n-1)\, p^2 \sum_{y=0}^{n-2} n - 2({}_y p^y (1-p)^{n-2-y} = n(n-1)\, p^2\, [\, P + (1-p)]^{n-1}$

Hence

$$E(X(X-1) = n(n-1)\, p^2$$
$$E(X(X-1)) = E(X^2) - E(X) = n(n-1)p^2$$
$$E(X^2) = n(n-1)\, p^2 + E(X) = n(n-1)\, p^2 + np.$$

Thus,

$$Var(X) = N(n-1)\, p^2 + np - n^2p^2 = n^2p^2 - np^2 + np - n^2p^2 = np(1-p).$$

The Variance of a Geometric Random Variable

The probability density function of X is given by

$$f(x) = p(1-p)^{x-1}, \; x = 1, 2, \ldots$$
$$E(X) = \sum_{x=1}^{\infty} x p(1-p)^{x-1} = \frac{1}{p}$$

$E(X) = P + 2p(1-p) + 3p(1-p)^2 + \ldots = P\{1 + 2(1+p) + 3(1-p)^2 + \ldots\}.$

From the Binomial theorem

$$(1-a)^{-2} = 1 + 2a + 3a^2 + \ldots \quad |a| < |.$$

Putting $a = (1-p)$, we obtain1

$$E(X) = P. (1-(1-p))^{-2} = \frac{1}{p}$$

$E[X(X-1)] = \sum_{x=2}^{\infty} x(x-1) P(1-p)^{x-1}$

$$= p[2.1(1-p) + 3.2(1-p)^2 + \ldots + r(r-1)(1-p)^{r-1} + \ldots].$$

Multiply both sides by 1-p,

$$(1-p)\, E[X(X-1)] = p(2.1(1-p)^2 + (3.2(1-p)^3 + \ldots + r(r-1)(1-p)^r + \ldots+]$$

$E(X(X-1)] - (1-p)E[X(X-1)] = p[2(1-p) + 4(1-p)^2 + 6(1-p)^3 + \ldots + 2r(1-p)^r + \ldots]$

$$= 2p(1-p)\, [\, + 2(1-p) + 3(1-p)^2 + \ldots] = 2p(1-p).\frac{1}{p^2} = \frac{2(1-p)}{p}$$

Hence

$$E[X(X)] - (1-p)\,E[X(X-1)] = \frac{2(1-p)}{p}$$

$$E[X(X-1)]\,\{1 - (1-p)\} = \frac{2(1-p)}{p}$$

which implies that

$$E[X(X-1)] = \frac{2(1-p)}{p^2}$$

$$E(X^2) = \frac{2(1-p)}{p^2} + E(X) = \frac{2(1-p)}{p^2} + \frac{1}{p}.$$

Thus

$$Var(X) = \frac{2(1-p)}{p^2} + \frac{1}{p} - \frac{1}{p^2} = \frac{2(1-p)+p-1}{p^2} = \frac{1-p}{p^2}$$

The Variance of a Poisson Random Variable

Let X be a poisson random variable with parament λ.

$$E(X) = \sum_{x=0}^{\infty} x\, e^{-\lambda}\frac{\lambda^x}{x!} = \sum_{x=1}^{\infty} e^{-\lambda}\frac{\lambda^x}{(x-1)!} = \lambda\, e^{-\lambda}\sum_{x=1}^{\infty}\frac{\lambda^{x-1}}{(x-1)!}$$

$$= \lambda e^{-\lambda}\sum_{y=0}^{\infty}\frac{\lambda^y}{y!} = \lambda\, e^{-\lambda}\, e^{\lambda} = \lambda$$

$$E[X(X-1)] = \sum_{x=0}^{\infty} x(x-1)e^{-\lambda}\frac{\lambda^x}{x!} = \sum_{x=2}^{\infty} x(x-1)\frac{e^{-\lambda}\lambda^x}{x!}$$

$$= e^{-\lambda}\lambda^2\sum_{x=2}^{\infty}\frac{\lambda^{x-2}}{(x-2)!} = e^{-\lambda}\lambda^2\sum_{y=0}^{\infty}\frac{\lambda^y}{y!} = \lambda^2$$

$$E(X^2) - E(X) = \lambda^2$$

$$E(X^2) = \lambda^2 + E(X) = \lambda^2 + \lambda$$

Hence,

$$Var(X) = \lambda^2 + \lambda - \lambda^2 = \lambda.$$

This shows that the mean and variance of a passion random variable are both equal.

Exercise 4.2

1. Compute the variance and standard deviation of the random variable X defined in exercises 4.1.1

2. A fair coin is tossed twice. Let X denote the number of heads that appear. Compute

    (i) the expected value of X

    (ii) the variance of X

(iii)   the expected value of $\frac{1}{x}$.

3.   Calculate the mean, variance and standard deviation of the random variable having the following probability density function.

X	-2	1	0	1	2
f(x)	3/10	1/5	1/10	1/5	1/5

4.   Let X be any random variance, show that $E(AX + B) = aE(X) + B$

Where a, b are contants.

5.   Let $f(x) = \frac{1-p}{2p} p^{|x|}$, $x = \pm 1, \pm 2,\ldots, 0 < p < 1$

Show that f(x) is a pdf. Find E(X).

6.   Let X be any random variable having finite expectation. Prove that
$$\left|EX\right| \le E\left|X\right|.$$

7.   Let X be a random variable such that for some constant K, $P(\left|X\right| \le K)$. Prove that X has finite expectation and $\left|EX\right| \le k$.

8.   Show that if X and Y are any two dependent random variables with finite variances, then $Var(X + Y) \ne Var(X) + Var(Y)$.

4.3   Probability Generating Function

In this section we shall introduce important mathematical concepts in probability theory. The rth moment of a random variable X is defined by $E(X^r)$, where $E(X^r)$ and $E(X - \mu)^r$ are called the rth moment of X about 0 and μ respectively.

From (2)
$$E(X^r) = \sum_{i=1}^{\infty} x_i^r f(x) \; and \; E(X - \mu)^r = \sum (x_i - \mu)^r f(x_i).$$

Note that the second moment of X about μ is the variance of X. an indirect way of calculating expectation is the use of probability generating function or moment generating function. it is a mathematical device to simplify the calculations of moments of X. its more general usefulness will not be apparent until we get to chapter 6.

Definition 6

The probability generating function of a non-negative, integer-valued random variable X is defined by

$$E(S^x) = P(S) = \sum_{x=0}^{\infty} S^x f(x); \quad -1 \leq s \leq 1; \quad -s_0 < s < s_0 \qquad (8)$$

From (6), we see that

$$P(S) = E(S^x)$$

By differentiating with respect to s we have

$$P'(s) = E(XS^{X-1})$$

$$P''(s) = E[X(X-1)S^{X-2}]$$

$$P^{(r)}(s) = E(X(X-1)(X-2)...(X-r+1)s^{X-1})$$

Putting s =1 in the above derivatives, we have

$$P'(1) = E(X)$$

$$P''(1) = E[X(X-1)] = E(X^2) - E(X)$$

$$P^{(r)}(1) = E[X(X-1)(X-2)...(X-+1r)].$$

Thus the mean and variance of X can be obtained from P(s) by the following formulas.

$$E(X) = P'(1)$$

$$E(X^2) - E(X) = P''(!)$$

$$E(X^2) = P''(1) + E(X) = P''(1) + P'(1)$$

And

$$Var(X) = E(X^2) - [E(X)]^2 = P''(1) + P'(1) - [P'(1)]^2$$

We now illustrate the use of these formulas with the following examples.

Examples 4.7

Let X be a poisson random variable wit parameter λ. Find the mean variance of X

$$f(x) = \frac{e^{-\lambda}\lambda^x}{x!}, x = 0, 1, 2, ...$$

The probability generating function is defined by

$$P(S) = E(S^X) = \sum_{x=0}^{\infty} s^x \, {}^n(_x p^X(1-p)^{n-x} = \sum_{x=0}^{\infty} \, {}^n(^x(sp)^x(1-p)^{n-x}.$$

By the Binomial theorem

$$= [sp + (1-p)]^n$$

Thus,

$$P(S) = [sp + (1 – p)]^n \qquad (12)$$

Differentiating, we have

$$P'(s) = np [sp + (1 – p)]^{n-1}$$

$$P''(s) = n(n – 1)p^2 [sp + 1 – p]^{n-2}$$

$$Var(X) = p'' (1) + - [p'(1) – [p'(1)]^2 = n(n – 1) p^2 + np – (np)^2 = np(1 – p).$$

The table below summarizes some of the results of this chapter

Random variable	Probability density function	Mean	Variance	Probability generating function	Moment generating function
Bernoulli	$P^x(1-p)^{1-x}$	P	P(1-p)	$(sp+1 – p)$	$pe^t + (1-p)$ - $\infty< t <\infty$
Binomial	$^nC_x p^x(1-p)^{n-x}$ x = 0, 1, 2, ….n	Np	Np(1-p)	$[sp +(1-p)]^n$	$[pe^t+(1-p)]^n$
Geometric	$p(-p)^{x-1}$ x = 1,2,…	$\dfrac{1}{p}$	$\dfrac{1-p}{p^2}$	$P[1-s(1-p)]^{-1}$	$\dfrac{ps}{1 - (1 - p)e^t}$
Poisson	$\dfrac{e^{-\lambda}\lambda^x}{x!}$	Λ	Λ	$e^{\lambda(s-1)}$	$e^{\lambda(e^t – 1)}$
Uniform	$\dfrac{1}{b - a + 1}$ X= a, a + 1,…b	$\dfrac{a+b}{2}$	$\dfrac{(b - a)^2}{12}$	$\dfrac{s^a - s^{b+1}}{(b - a + 1)(1 - s)}$	

The Moment Generating Function

Definition 7

Let X be a discrete random variable with pdf f(x). the moment generating function denoted by $M_x(t)$ is defined by

$$M_X(t) = E(e^{tX})$$

Where t is any real constant for which the expectation exists. $\qquad$ (13)

$$M_X(t) = \sum_{Vx} e^{tx} f(x)$$

$$R^{tX} = 1 + tX + \frac{(tX)^2}{2!} + \dots$$

$$E(e^{tX}) = E(1 + tX + \frac{(tX)^2}{2!} + \dots + \frac{t^n X^n}{n!}$$

Under fairly general conditions we shall assume that expectation of infinite sum equals the sum of the expected value but this is true in general for finite sum

$$1 + tE(X) + \frac{t^2}{2!} E(X^2) + \frac{t^n X^n}{n!} + \dots$$

That is

$$M_X(t) = 1 + tE(X) + \frac{t^2}{2!} E(X^2) + \dots + \frac{t^n}{n!} E(X^n) + \dots$$

Differentiating with respect to t, we have

$$M'_{X(t)} = E(X^2) + \dots + \frac{n(n-1)t^{n-2}}{n!} E(X^n) + \dots$$

$$M''_x(0) = E(X^{2+})$$

hence,

$$Var(X) = M_x''(0) - [M_X'(0)^2]$$

And

$$M_x(t) = E(e^{tX}), \quad M_x^{(r)}(t) = E[X^r e^{tX}] \tag{15}$$

Putting t = 0, we have

$$M^{(r)}(0) = E(X_r)$$

that is, $E(X^r)$ is the rth derivative of $M_x(t)$ evaluated at t=0. (16)

Note: It is assumed that $\sum e^{tX} f(x)$ converges for all values of t.

The domain of $M_x(t)$ is all real number t such that $\sum e^{tX} f(x)$ converges. If $M_x(t)$ is defined for some $t_0 > 0$, then it is defined in the interval $0 \le t \le t_0$.

Examples 4.9

Let X be a binomial random variable with parameters n and p. then

$$M_X(t) = E(e^{tX}) = \sum_{x=0}^{n} e^{tX} {}^nC_x P^x (1-p)^{n-x} = \sum_{x=0}^{n} {}^nC_x (Pe^t)^x (1-p)^{n-x}$$

from the binomial theorem

$$(a + b)^n = \sum_{x=0}^{n} {}^nC_x a^x b^{n-x} = [pe^t + (1-p)]^n,$$

on putting $a = pe^t$, $b = 1 - p$.

Example 4.10

Let X be a Poisson random variable with parameter λ. Then

$$M_X(t) = \sum_{x=0}^{\infty} e^{tX} \frac{e^{-\lambda}\lambda^x}{x!} = e^{-\lambda} \sum_{x=0}^{\infty} \frac{(\lambda e^t)^x}{x!} = e^{-\lambda}e^{\lambda e^t} = e^{\lambda(e^t - 1)}$$

Differentiating, we have

$$M'_X(0) = \lambda e^t e^{-\lambda(1-e^t)}$$

$$M'_X(0) = \lambda e^t e^{-\lambda(1-e^t)} \Big|_{\lambda=0} = \lambda\, E(X)$$

$$M_x''(t) = \lambda e^t\, e^{-\lambda(1-e^t)} + \lambda^2\, e^{2t}e^{-\lambda(1-e^t)}$$

$$M_x''(0) = \lambda + \lambda^2 = E(X^2)$$

Hence,

$$\text{Var}(x) = \lambda + \lambda^2 - \lambda^2 = \lambda.$$

Example 4.11

Let X be a geometric random variable with parameter p. them

$$M_x(t) = \sum_{x=1}^{\infty} e^{tX}\, P(1-p)^{x-1} = \frac{p}{1-p} \sum_{x=1}^{\infty} e^{tX}\, P(1-p)^x$$

$$= \frac{P}{1-p} \sum_{x=1}[e^t(1-p)]^x = \frac{p}{1-p}\cdot\frac{e^t(1-p)}{1-e^t(1-p)} = pe^t[1 - e^t(1-p)]^{-1}.$$

The fourth equality follows from sun to infinity of geometric series $\sum_{x=1}^{\infty} a^x$. Further applications of $M_x(t)$ will be considered later in Chapter 6.

Exercise 4.3

1. Find the probability generating function of the geometric distribution with parameter P. hence determine the mean and variance of the distribution.

2. Let X be uniformly distributed on (0, 1, 2,…n). Find the mean and variance of X.

    Hint: $\sum_1^n x = \frac{n(x+1)}{2}, \sum_1^n x^2 = \frac{n(n+1)(2n+1)}{6}$.

3. Let X be a random variable with finite variance. Prove that for any real number a, $\text{Var}(X) = E[X - a)^2] - [E(X) + a]^2$.

4. A fair die is tossed 50 times. Let X be the number oftimes six appears. Evaluate

(i) E(X)   (ii) E(X²)   (iii) VAR(X)

5.   Find the expected value and variance of the random variable X os Example 3.1.1.

6.   Show that if c is a constant, $Var(x + c) = Var(X)$.

7.   Let X be uniformly distributed on (0, 1, 2,..N). Find P(S), mean and variance of X.

8.   Let X be defined by

x	1	2	3
f(x)	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$

Find P(S) and hence the mean and variance of X.

9.   Prove that if X has variance σ^2, and mean μ, then

$\sum_{i=1}^{\infty}(x_i - a)^2 f(x_i) = \sigma^2 + (a - \mu)^2$ every real number a.

10.   Derive the MGF for the following: (a) Bernoulli, (b) Negative Binomial, (c) Binomial and use it to find the mean and variance.

11.   A random variable X has the pdf f(x) defined by

$$f(x) = \begin{cases} c(x + 2), & x = 1, 2, 3, 4 \\ 0 & \text{elsewhere} \end{cases}$$

Find (i) c, (ii) moment generating function of X, (iii) use result of (ii) tofind the mean and variance of X.

12.   Prove that for any random variable, X, $E(X^2) > [E(X)]^2$.

13.   Show that $[E(X - a)^2]$ is minimized when a = E(X).

14.   A fair die is rolled until all the 6 sides appeared at least once. Find the expected number of rolls needed.

15.   Let $f(x) = \begin{cases} p, & x = -1 \\ P^x(1-p)^2, & x = 0,1, 2,... \end{cases}$

(a) Show that f(x) is a pdf. (b) Find $M_x(t)$, the moment generating function of (b) Hence, the mean and variance of X.

MODULE THREE (UNIT 1)

CONTINUES RANDOM

5.1   INTRODUCTION

In the chapter so far, we consider discrete random variables and their probability density functions e.g. Binomial, Poisson, Geometric. These are essentially counting variables. However, there exist random variables whose set of possible uncountable, such random variables are called continuous random variables.

In this chapter, we shall study in considerable detail, a number of important continuous random variables and their characteristics. We shall consider an idealized range space for a random variable X, in which all possible real number (in some specific intervals or set of intervals) many be considered as possible outcomes. We shall essentially replace the summation in the previous chapters in the discrete probability function case by integration for the continuous probability density function case.

Example 5.1

Let a point be chosen at random from the interval (0, 2) and let X represent the point chosen. Then 0 ≤ X≤ 2. the range space of X consists of unaccountable real numbers. Any real number in the interval (0, 2) is a sample point. Let A be the event that the point chosen is 0.3. From the basic definition of probability we have

$$P(A) = \frac{number\ of\ point\ A}{number\ of\ points\ in\ \Omega} = \frac{1}{\infty} = 0.$$

Hence,

In general, if X is a continuous random variable, probability that X takes a specific value is 0. random variables denoting measurements of physical quantities such as time, weight, temperature, length are continuous random variables. The proability that X lies between two defined values is given by

$$P(a \leq X \leq b).$$

$$P(X \varepsilon A) = \int_A f(x)dx$$

For any set A of real numbers

$$P(X) \text{ is denoted by } F(X).$$

Definition

The distribution function F of a random variable X is the function

$$P(X) = P(X \leq x),\ -\infty < x < \infty,$$

If we know the probability distribution function of X, then we can calculate the probability that X fails in a specific interval. For example

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

Properties of Distribution Function

A probability distribution function is any function F satisfying the following conditions:

(i) $0 \leq F(x) \leq 1$ for all x

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

(iii) $F(-\infty) = 0$ and $F(+\infty) = 1$.

(iv) $F(x +) = F(x)$ for all values of x where $F(x+) = \underset{\delta-0}{Lim} F(x + \delta)$.

Graphs of a typical distribution function

Example   2

Show that

$$F(x) = \begin{cases} e^x & x < 0 \\ 1 & x > 0 \end{cases}$$

is a probability distribution function

Solution

$0 \le e^x \le 1$ for x ≤ 0 and e^x is a non decreasing continuous function, hence F(x) is a probability distribution function.

Probability density function P.df

For discrete random variable

$$F(x) = P(X \le x) = \sum_{-\infty}^{x} f(y)$$

Where f(y) = P(X = y).

Similarly, for continuous random variables, there exists a function f(x) such that

$$F(x) = P_x(X \le x) = \int_{-\infty}^{x} f(y)dy$$

Such a function f is called probability density function of X. to see this consider the following.

The probability that the random variable X takes on a value between x and x + Δx is

P(x < X < x + Δx) = P(X ≤ x + Δx) − P(X < x)

= F(x + Δx) − F(x).

this is the probability associated with an interval of length Δx

$$\frac{F(x+ \Delta x) - F(x)}{\Delta x}$$

is the probability per unit length in the interval. If we now let Δx tends to zero, we have

$$\lim_{\Delta x - 0} \frac{F(x+x) - F(x)}{\Delta x} = F'(x) = f(x)$$

Hence, F(x) $= \int_{-\infty}^{x} f(y)dy$ where F(-∞) = 0

F(x) represents the probability per unit distance along the x acis and

$$F(x) = \frac{d}{dx} F(x)$$

This shows that the p.d.f of X is obtained by differentiating the distribution function of X.

Properties of probability density function

A probability density function is any function satisfying the following conditions.

(i) $f(x) \geq 0$

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

(iii)   P(X < a) = $\int_{-\infty}^{a} f(x)\,dx$ = 1 - $\int_{a}^{\infty} f(x)\,dx = 1 - P(X > a)$

(iv)   if A is any interval, P(X ε A) = $\int_{-\infty} f(x)\,dx$.

Figure (2) represents a graph of the function f(x). the area enclosed by the curve f(x) between a and b is

$$\int_{a}^{b} f(x)\,dx = P(a \leq X \leq b) = F(b) - F(a).$$

Graphical Representation of a Continuous Random Variable.

Theorem

(i)   The function F is non decreasing

(ii) $\lim\limits_{x \to \infty}$ F(x) = 0 and $\lim\limits_{x \to \infty}$ F(x) = 1.

Proof:

(i)   Suppose $x_1 \leq x_2$ and let A = { X $\leq$ x_1}, B = {X $\leq$ x_2), then

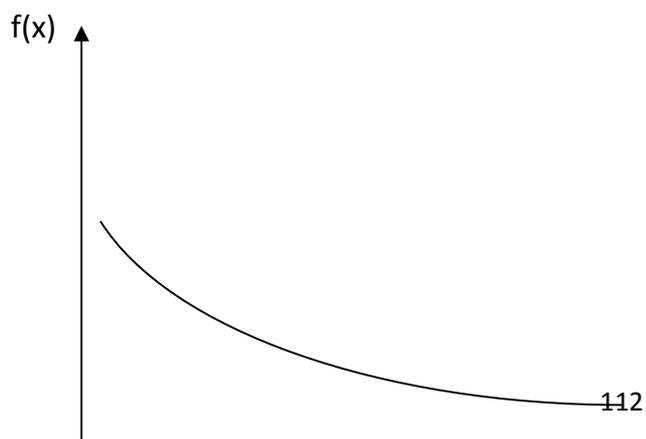$$A \subset B \text{ since } x_1 \leq x_2$$

Thus

$$P(A) \leq P(B)).$$

(ii) $\lim\limits_{x \to \infty}$ F(X) = $\lim\limits_{x \to \infty} \int_{-\infty}^{x} f(y)dy = 0$

$$\lim\limits_{x \to \infty} F(x) = \lim\limits_{x \to \infty} \int_{-\infty}^{x} f(y)dy = 1$$

Examples

Exponential Distribution. Let f(x) = $\begin{cases} e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$
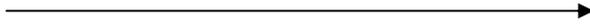
The graph of f(x) is shown in Figure 2 below

f(x)

112

Figure 3. Density function of exponential distribution.

The random variable X having its probability density function given by (1) is said to have exponential distribution with parameter λ.

5.4   Normal Distribution

Let f(x) = $\dfrac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}, -\infty < x < \infty.$

The graph of f(x) is shown in Figure 4 below.

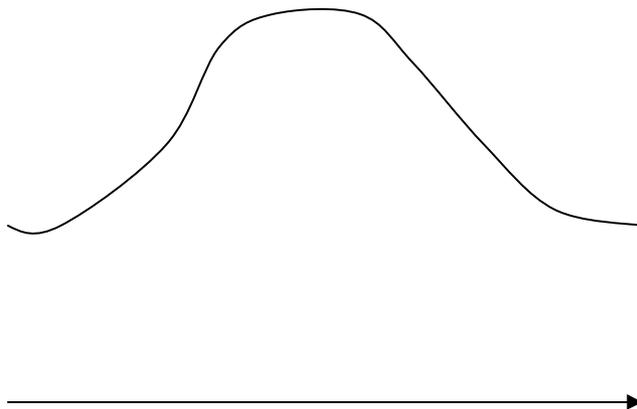Figure 4. Standard Normal Distribution

5.5   Let the random variable X have the probability density function $f(x) - 1/2e^{-1/2x}$ $0 \leq x \leq \infty$. Find (i) $P(2 \leq X \leq 7)$ (ii) $P(X \geq 10)$.

Solution

$$P(2 \leq x \leq 7) = \int_2^7 1/2e^{-1/2x}\,dx = -e^{-\frac{x}{2}}I\Big|_2^7 = e^{-1} - e^{-7/2} = 0.338$$

And

$$P(X \geq 10) = \int_{10}^{\infty} 1/2e^{-1/2x}\,dx = -e^{-1/2x}I\Big|_{10}^{\infty} = e^{-5}.$$

Suppose that X is a continuous random variable whose probability density function is given by

$$f(x) = \begin{cases} A\ e^{-2x} & 0 < x < \infty \\ \\ 0 & \text{otherwise} \end{cases}$$

(i)   what is the value of a? and (ii) find $P(X > 1)$.

Solution

(i)   From property (ii) of a probability density function, we have

$$\int_{-\infty}^{\infty} f(x)\,dx = 1$$

implying that

$$\int_0^{\infty} ae^{-2x}\,dx = 1 = a \int_0^{\infty} e^{-2x}\,dx = 1/2$$

$$==> a = 2$$

hence

$$f(x) = \begin{cases} 2e^{-2x} & 0 < x < \infty \\ 0 & \text{otherwise} \end{cases}$$

(ii) $P(X > 1) = \int_{1}^{\infty} 2e^{-2x}\, dx = e^{-2} = 0.14$

5.2   Expectation and variance of continuous random variables.

Let X be a continuous random variable with probability density function f(x). Suppose we divide the range of X into small intervals, each of length Δx (by the definition of f(x)).

The expected value of X is therefore approximately

$$\sum xf(x)\Delta x$$

The limiting value of this as Δx   0

$$\int xf(x)\, dx.$$

This leads us to the following definition.

Definition 3

If X is a continuous random variable with probability density function f(x), then

$$E(x) = \int_{-\infty}^{\infty} xf(x)\, dx$$

In general,

$$E(X^r) = \int_{-\infty}^{\infty} x^r f(x)\, dx$$

115

$$E(\Phi(X)) = \int_{-\infty}^{\infty} \Phi(x)\, f(x)\, dx.$$

The variance $\sigma^2 = E[(X - \mu)^2] = \int_{-\infty}^{\infty} (x - \mu)^2\, f(x)\, dx = E(x^2) - \mu^2$

Where $\mu = E(x)$.

Examples

5.6 Suppose X is a continuous random variable with p.d.f $f(x) = \lambda e^{-\lambda x}$, $0 < x < \infty$.

Find (i) the mean and variance of X and (ii) the probability distribution function F(x), and the probability that X lies between 2 and 5.

Solution

$$E(X) = \int_{-\infty}^{\infty} x.\lambda e^{-\lambda x} dx = \int_{0}^{\infty} x\lambda e^{-\lambda x} dx$$

By integrating by parts

$$= \int_{-\infty}^{\infty} -x \frac{d}{dx}(e^{-\lambda x}) = -xe^{-\lambda x} \, \mathrm{I} \, \Big|_{0}^{\infty} + \int_{0}^{\infty} e^{-\lambda x} dx = \frac{1}{\lambda}.$$

Similarly,

$$E(x^2) = \int_{0}^{\infty} x^2.\lambda e^{-\lambda x} dx = \int_{0}^{\infty} -x^2 \frac{d}{dx}(e^{-\lambda x}) = -x^2 e^{-\lambda x} \, \mathrm{I} \, \Big|_{0}^{\infty} + 2\int_{0}^{\infty} xe^{-\lambda x} dx$$

$$= 2\int_{o}^{0} x^{o-2x} dx = \frac{2}{\lambda} \int x\lambda \ell^{-2x} dx = 2E(x) = \frac{2}{\lambda^2}$$

hence, $\text{var(x)} = \dfrac{2}{\lambda^2} - \dfrac{1}{\lambda^2} = \dfrac{1}{\lambda^2}$

(u) $F(x) = \displaystyle\int_{-}^{xf(y)dy=} \int_{=}^{x} f(y)dy = \int_{=}^{x} \lambda\ell - \lambda y dy = -e^{-\lambda y[\frac{x}{0}=1-e^{1x}}$

$p(2 < x < 5) = \displaystyle\int_{2}^{5} \lambda e^{2x} dx = F(5) - F(2)1 - e^{-5} - 1 + e^{-\mu} = e^{-\mu} - e^{-\mu}$

5.7   Let X be a continuous random variable with probability density function

$F(x) = \begin{cases} \dfrac{1}{6} & \\ & 0<x<6 \\ & \\ \text{elesewhere} \end{cases}$

find (1) F(x) (ii), E(x), $E(x^2)$ and var (x).

solution

(1) $F(x) = \displaystyle\int_{0}^{x} \dfrac{1}{6} = p(x_{\leq}x)$

$E(x) = \displaystyle\int_{0}^{0} x\dfrac{1}{6}dx = \dfrac{1}{6-2}\dfrac{1}{2}x^2[\dfrac{6}{0} = \dfrac{1}{6} - \dfrac{1}{2} - 36 = 3$

$E(x^2) = \displaystyle\int_{0}^{6} x^2 - \dfrac{1}{6}dx = \dfrac{1}{6} - \dfrac{1}{3}x^3[_{0}^{6} = \dfrac{1}{6} - \dfrac{1}{3} - 6^3 = 12$

hence   var(X) = 12-9 =3

5.3 symmetric probability Density Function

definition 4

a probability density function f(x) is called a symmetric probability function if

$$f(-x) =f(x) \text{ for all x.}$$

for example. If $f(x) = \sqrt{2x}e\dfrac{1}{2}^{x2}$ $isasynmericfunction.$

similarly,

$$f(x) = \frac{1}{2x} - 0 < x < a$$

is a symmetric probability density function

Definition.

A probability density function f(x) is said to be symmetric about a if

$$F(a + x)=f(a - x) \text{ for all x.}$$

If x = 0. then f(x) is symmetrical about 0. for example, if

$$F(x) = \sqrt{2\pi}e\frac{-1}{2}(x-\pi)^2 =< x <=$$

Then

$$F((\mu - x) = \frac{1}{\sqrt{2x}} e \frac{1}{2} (\mu + x- \mu)2 = \frac{1}{2\pi} e \frac{-1}{2}$$

2

$$F(\mu - x) = \frac{1}{\sqrt{2x}} e \frac{1}{2} (\mu + x- \mu)2 = \frac{1}{2\pi} e \frac{-1}{2}$$

hence, f(μ +x) = f(μ -x). thus f(x) is symmetrical about μ .

theorem 1

if f(x) is a symmetric probability density function then

$$f(x) = 1-f(-x)$$

where f(x) is the distribution function of x.

proof

$$f(-x) = \int_{-}^{x} f(y)dy = \int_{-}^{-x} f(-y)dy$$

since f(y) =f(-y). Let z = -y, then

$$f(-x) = \int_{-}^{x} f(z)(-1)dz = \int_{x}^{-} f(z)dz = \int_{-}^{-} f(z)dz - \int_{-}^{x} f(z)dz = 1 - f(x)$$

and hence

$$f(x) = 1\text{-}f(x)$$

terminology

when we speak of the probability distribution of a random variable X we mean its probability density function and when we speak of the distribution function or cumulative distribution function we always mean f(x).

5.2 some important continuous random variables.

5.2.1   The Normal Distribution

one of the most important continuous random variables is the Normal random variable, X. The range of X is the real line.

Definition

The Normal Distribution. The continuous random variable X has the normal distribution with mean μ and variance^3 if μ has the probability density function.

$$F(x) = \frac{1}{\sqrt{2\pi}}\, e\, \frac{-1}{2}\frac{(x-\mu)^3}{0^2} =< x <=$$

The graph of f(x) is given in Figure (4)

It can be verified that f(x) satisfies

120

1. $\int f(x)dx = 1$

2. $\int xf(x)dx = \mu$

3. $\int (x-\mu)^2 f(x)dx = 0^2$

Properties of the Normal Distribution.

1. $\int f(x)dx = 1$ this can be verified.

Thus let z = $\dfrac{x-\mu}{\sigma}$

$$\int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}} dx = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2x}} e^{-\frac{1}{2}} dz = 1$$

$$I^2 = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-\frac{1}{2}y^2} dy = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(z^2+y^2)} dz\, dy.$$

Let us introduce polar coordinates to evaluate this double integral:

$$Z = r\cos\theta,\ y = r\sin\theta$$

$$I^2 = \frac{1}{2} \int_{0}^{2\pi} \int_{0}^{\infty} re^{-r^2} dr\, d\theta = \frac{1}{2\pi} \int_{0}^{2\pi} -e^{-\frac{r^2}{2}} \Big|_{0}^{\infty} d\theta = \frac{1}{2\pi} \int_{0}^{2\pi} d\theta = 1.$$

where dzdy = rdrdθ, Hence

$$\int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}} dx = 1.$$

2. $E(X) = \mu$, $Var(X) = \sigma^2$.

The p.d.f is completely specified when μ and σ are known.

When $\mu = 0$ and $\sigma^2 = 1$. X is said to have a standardized normal distribution. This is, the pdf of X is

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}, -\infty < x < \infty$$

The importance of the standardized normal distribution is the fact that it is tabulated.

3. $f(x)$ is symmetrical about the point μ, hence $F(-x) = 1 - F(x)$.

Most available tables are tabulated only for positive values of x since $F(-x)$ can be found from $f(x)$ and $f(x)$ are given at the back of this book. The graph of $f(x)$ is bell-shaped curve centred at its mean.

4. IF X has a normal distribution with mean μ and variance σ^2, then $Y = \dfrac{X - \mu}{o}$ has a standardized normal distribution. The proof of this is left as an exercise. Y is said to have a standard Normal distribution.

5. The moment generating function of X is defined by

$$M_x(t) = E(e^{tx}) = \int_{-\infty}^{\infty} e^{tx} \frac{1}{\sqrt[q]{2\pi}} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}} dx.$$

Let $y = \dfrac{X - \mu}{\sigma}$, then $dy = \dfrac{1}{\sigma} dx$.

$$x = \mu + y\,\sigma.$$

Thus,

$$M_x(t) = \int_{-\infty}^{\infty} e^{t(\mu + y\sigma)} \frac{1}{\sqrt[q]{2\pi}} e^{-\frac{1}{2}y^2} . \sigma\,dx\,dy.$$

$$= \frac{e^{t\mu}}{\sqrt{2x}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(y-t\sigma)^2} \, dy = e^{t\mu + \frac{1}{2}\sigma^2 t^2} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(y-t\sigma)^2} \, dy.$$

Let y – tσ = y, then

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(y-t\sigma)^2} \, dy = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}y^2} \, dy = 1.$$

Hence,

$$M_x(t) = e^{t\mu + \frac{1}{2}\sigma^2 t^2}$$

Examples

5.8   If X has a standard normal distribution i.e. mean of X, μ=0, and variance = 1. Find

    (i)   P(X ≤ 0.5)   (ii)   P(X ≥ 0.5)

    (iii)   P(X ≤ -0.3)   (iv)   P(0.3 ≤ X ≤ 0.5)

    (v)   P(-0.3 ≤ X ≤ 0.5)

Solution

(i)   P(X ≤ 0.5) = $\int_{-\infty}^{\infty} f(x)\,dx = F(0.5)$.

From table (3) F(0.5) is found by locating the first two digits (0.5) in the column headed x, the desired probability of 0.6915 is found in the row labeled 0.5 and the labeled 0.00. That is

$$P(X \le 0.5) = \int_{-\infty}^{0.5} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \, dx = 0.6915.$$

and

(ii)   P(X ≥0.5) = 1 – P(X < 0.5) = 1 – 0.6915 = 0.3085.

(iii)   P(X ≤ -0.3) = F(-0.3) = 1 – F(0.3)

From the table, F(0.3) = 0.6179.

(iv) $P(0.3 \leq X \leq 0.5) = F(0.5) - F(0.3) = 0.6915 - 0.6179 = 0.0736$

(v) $P(-0.3 \leq X \leq 0.5) = F(0.5) - F(-0.3) = 0.6915 - 0.3821 = 0.3074.$

Example 5.9

If x is a normal variable = 5 and variance $0^2 = 16$, find (i) $P(X) \leq 6)$

(ii) $P(3 < x < 7)$ (iii) $P\{X-5\}>2\}$.

<div align="center">**Solution**</div>

From property (4)

(i) $P (X <(6) = P\left(\frac{X-\mu}{4} \leq \frac{6-5}{4}\right) = P\left(Y \leq \frac{1}{4}\right) = 0.5987$

(ii) $P(3 < X < 7) = P\left(\frac{3-5}{4} < \frac{X-5}{4} < \frac{7-5}{4}\right) = (-0.5 < Y < 0.5)$

   $= \Phi(0.5) - \Phi(-0.5) = 0.6915 - 0.3085 = 0.3830$

(iii) $P\{|X-5|>2\} = P(X - 5 > 2 \text{ or } X - 5 < -2) = P(X > 7 \text{ or } X < 3\}$

   $= P(X > 7) + P(X<3) = P\left(Y > \frac{7-5}{4}\right) + \left(Y < \frac{3-5}{4}\right)$

   $= P(Y > 0.5) + P(Y < -0.5) = 0.3085 + 0.3085 = 0.6170.$

Notation:

When we write X is N (0^2), we mean that X has a Normal probability distribution with mean and variance 0^2

5.3.2 The Normal Approximation to the Binomial

Let X be a binomial random variable with parameters n and p, then

<div align="center">E(X) = np, var (X) = np (1-p)</div>

An approximation to the distribution of X is given by De Moivre Laplace limit theorem.

Theorem (De Moivre laplace Limit Thoerem)

If X denote a binomial random variable with mean up and variance np(1-p), then for any a < b,

$$\lim_{n\to\infty} P\left\{a \le \frac{X - np}{\sqrt{np(1-p)}} \le b\right\} = \Phi(b) - \Phi(a).$$

The accuracy of this approximation is quite good for values of n greater than 10 P(1-p).

This theorem says that if X is B(n,p) then large n, X is N(np, np (1-p)).

Example 5.10

Let X be a binomial random variable with parameters n = 40, p = 0.20. find the probability that X = 5.

Since X is discrete and normal distribution is for continuous random variable, we write

P (X = 5) = {(4.5 < 5.5 = $P\left(\frac{4.5-4}{1.90} < \frac{X-4}{1.90} > \frac{5.5-4}{1.90}\right)$

= p (0.26 < y < 0.79) = Φ (0.79) – Φ(0.26) = 0.18226.

Notation: Whenwe write X is N_0 (μ o), we mean that X has a normal probability distribution with mean μ variance 0^2

5.2.4   The Exponential Distribution

A continuous random variable, X is said to have an exponential distribution with parameter > 0 if its pdf is given by,

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & if\ x \ge 0 \\ 0 & x < 0 \end{cases}$$

The pdf has the graph shown in Figure 3. It can easily be shown that

$$\int_0^\infty \lambda e^{-\lambda x} = 1.$$

The exponential distribution plays an important role in point processes and in describing random phenomena. A number of real life situations can be modeled using an exponential distribution.

For example, time to failure of an equipment.

Properties of the Exponential Distribution

(1)   E(X) and standard deviation of X are the same. The expected value of X is given by

$$E(X) = \int_0^\infty x\lambda e^{-\lambda x}dx$$

On integrating by parts we have

$$\int_0^\infty x\lambda e^{-\lambda x}dx = \int_0^\infty -xd\left(e^{-\lambda x}\right) = -xe^{-\lambda x}|_0^\infty + \int_0^\infty e^{-\lambda x}dx = \frac{1}{\lambda}$$

(ii)   The variance of X is given by

$$Var(X) = E(X^2) - (E(x))^2$$

$$E(X^2) = \int_0^\infty x^2\,\lambda e^{-\lambda x}dx = \int_0^\infty -x^2 d\left(e^{-\lambda x}\right) = -x^2 e^{-\lambda x}|_0^\infty + 2\int_0^\infty xe^{-\lambda x}dx = \frac{2}{\lambda^2}$$

Hence,

$$Var(X) = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}\;.$$

Thus, the mean and standard deviation of an exponential distribution are the same.

(ii) Lack of Memory of the Exponential Distribution

A random variable X is said to be without memory =, or memorable, if

$$P(X> s + t\backslash X> t) = P(X > s) \text{ for all } s, t \geq 0$$

$$P(X > s + t)\; X > t = \underline{P\,(X > s + t \text{ and } X > t} = \underline{P(X> s + t}$$

$$P(X> t) \qquad\qquad P(X > t)$$

By the definition of memoryless, we have,

$$P(x > s) = \frac{1 - F(x + t)}{1 - F(t)}$$

That is,

$$1 - F(s) = \frac{F(s+t)}{1-F(t)}$$

$$(1-F(s))(1-F(t) = 1 -F(s+t)$$

$$F(X) = P(X \leq x) = dy = 1-e^e$$

$$1 - F(x) = ex$$

$$P(X > s) = 1-F(s) = e^x$$

$$P(X > s + t) = 1 - F(s + t) = e$$

$$(p(x>t) = 1 - F(t) = e$$

$$P(X > s) = e$$

$$e^-\mu = e$$

thus, an exponential random variable is memoryless. The only continuous random variable X assuming non negative values satisfying

$$P(X < s + t/X > t) = p(X > s) \text{ for all s, t} > 0$$

(iii)   The moment generating functions of X is given by (definition 7 of chapter 4).

5,2,5   The Gamma Distribution

Definition 6

The gamma function, denoted by I', is defined as follows:

$$I(x) =$$

By integrating by parts, we have

$$F(X) = \int_0^\infty -y^{x-1} d(e^{-y})dy = -y^{x-1}e^{-y}\Big|_0^\infty + \int_0^\infty (x-1)y^{x-2}e^{-y}dy$$

$$= (x-1) \int_0^\infty y^{x-2} e^{-y} dy = (x-1)\Gamma(x-1)$$

$\Gamma(x) = (x-1)\Gamma(x-1) = (x-1). (x-2)\, \Gamma(X-2) = (X-1)!\Gamma(1)\backslash$

$$(1) = \int_0^\infty e^{-y} dy = 1.$$

Hence

$$\Gamma(x) = (x-1)!$$

It is also easy to verity that

(i) $\quad \Gamma\left(\frac{1}{2}\right) = \int_0^\infty y^{-\frac{1}{2}} e^{-y}\, dy = \sqrt{\pi}$

(ii) $\quad \int_0^\infty \frac{y^{n-1} e^{-y} dy}{\Gamma(n)} = 1.$

Definition 7

Let X be a continuous random variable assuming only non-negative values. Then X is said to have a gamma distribution if its pdf is given by

$$f(x) = \begin{cases} \dfrac{\lambda(\lambda x)^{n-1}}{\Gamma(n)} e^{-\lambda x} & x > 0 \\ 0 & elsewhere \end{cases}$$

λ and n are called the paramenters of the distribution. f(x) is denoted by $\Gamma(n, \lambda)$.

Properties of the Gamma distribution

1. E(X) $= \frac{n}{\lambda}, Var(X) = \frac{n}{\lambda^2}$

To see this, we have

E(X) $= \int_0^\infty x. \frac{\lambda(\lambda x)^{n-1} e^{-\lambda x}}{\Gamma(n)} dx = \int_0^\infty \frac{\lambda^n X^n e^{-\lambda x}}{\Gamma(n)} dx = \frac{\lambda^n}{\Gamma(n)} \int_0^\infty X^n e^{-\lambda x} dx.$

Let y = λx, dy = λdx

$$\int_0^\infty x^n e - \lambda x\, dx = \int_0^\infty \left(\frac{y}{\lambda}\right)^n e^{-y} \frac{1}{\lambda}\, dy = \frac{1}{\lambda^{n+1}} \int_0^\infty y^n e^{-y}\, dy.$$

Hence,

$$E(X) = \frac{\lambda^n}{\Gamma(n)} \cdot \frac{\Gamma(n+1)}{\lambda^{n+1}} = \frac{n}{\lambda}.$$

Similarly,

$$E(X^2) = \int_0^\infty x^2 \cdot \frac{\lambda(\lambda x)^{n-1} e^{-\lambda x}}{(n)} dx = \frac{\lambda^n}{\Gamma(n)} \int_0^\infty x^{n+1} e^{-\lambda x}\, dx.$$

$$\int_0^\infty x^{n+1} e^{-\lambda x}\, dx = \frac{\Gamma(n+2)}{\lambda^{n+2}}.$$

Hence,

$$E(X^2) = \frac{\lambda^n}{\Gamma(n)} \cdot \frac{\Gamma(n+2)}{\lambda^{n+2}} = \frac{n(n+1)}{\lambda^2}.$$

Thus

$$Var(X) = \frac{n(n+1)}{\lambda^2} - \frac{n^2}{\lambda^2} = \frac{n}{\lambda^2}.$$

The moment generating function of x is given by

$$M_x(t) = \int_0^\infty e^{tx} \frac{\lambda^n x^{n-1} e^{-\lambda x}}{\Gamma(n)}\, dx = \frac{\lambda^2}{\Gamma(n)} \int_0^\infty x^{n-1} e^{-x(\lambda-t)} dx$$

$$= \frac{\lambda^n}{\Gamma(n)} \cdot \frac{1}{(\lambda-t)^{n-1}} \int_0^\infty [x(\lambda-t)]^{n-1} e^{-x(\lambda-t)} dx.$$

Let y = x(λ − t), dy = (λ − t) dx. Therefore

$$M_x(t) = \frac{\lambda^n}{\Gamma(n)(\lambda-t)^n} \int_0^\infty y^{n-1} e^{-y} dy = \frac{\lambda^n}{\Gamma(n)(\lambda-t)^n} \cdot \Gamma(n) = \left[\frac{\lambda}{\lambda-t}\right]^n \qquad (8)$$

2. If X is a random variable having Γ(n, λ) distribution and Y is a poisson random variable with parameter λx, then

$$P(X \le x) = P(Y \ge n)$$

That is

$$\int_0^\infty \frac{\lambda^n y^{n-1} e^{-\lambda y}}{\Gamma(n)}\, dy = \sum_n^\infty \frac{(\lambda x)^k e^{-\lambda x}}{k!}.$$

3. If n = 1, f(x) becomes

129

$$f(x) = \lambda e^{-\lambda x}.$$

Hence, the exponential distribution is a special case of the Gamma distribution.

5.2.6 The Chi-square Distribution

A special case of the gamma distribution in which $\lambda = \frac{1}{2}$ and n replaced by n/2, where n is a positive integer is called the chi-square distribution. The pdf is given by

$$f(x) = \frac{x^{\frac{n}{2}-1}\left(\frac{1}{2}\right)^{\frac{n}{2}} e^{-\frac{n}{2}x}}{\Gamma\left(\frac{n}{2}\right)} = \frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)} x^{\frac{n}{2}-1} e^{-\frac{x}{2}}, x > 0$$

A continuous random variable having pdf f(x) given by (7) is said to have a chi-square distribution with n degrees of freedom (denoted by χ_n). Because of its importance, the chi-square distribution is tabulated for various value of the parameter n) see table 4). If X has a chi-square distribution with n degrees of freedom, we have

(i)   E(X) = n;

(ii)   Var(X) = 2n;

(iii) $M_x(t) = (1 - 2t)^{-n/2}$; $t > \frac{1}{2}$.

The graph of f(x) is shown in figure below.

5.2.7 The uniform distribution

A continuous random variable X assuming all values in the interval (a, b) where both a and b are finite is said to be uniformly distributed over the interval (a, b) if its pdf is given by

$$f(x) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & elsewhere \end{cases}$$

The graph of f(x) is shown in figure below:

Properties of the Uniform Distribution

1. $E(X)$ $\frac{a+b}{2}$, $M_x(t) = \frac{e^{bt} - e^{at}}{t(b-a)}$

$e^{ta} = 1 + ta + \frac{(ta)^2}{2!} + \cdots$

$e^{ta} + 1 + tb + \frac{(tb)^2}{2!} + \cdots$

$\frac{e^{bt} - e^{at}}{t(b-a)} = \frac{1}{b-a}\left[(b-a) + \frac{(b^2 - a^2)t}{2!} + \frac{(b^3 - a^3)t^2}{3!} + \cdots\right]$

$M_X(t) = \frac{1}{b-a}\left[\frac{(b^2 - a^2)t}{2!} + \frac{(b^3 - a^3)t^2}{3!} + \cdots\right]$

$M_x'(0) = \frac{(b^2 - a^2)}{2(b-a)} = \frac{(b-a)(b+a)}{2(b-a)} = \frac{b+a}{2}$

Similarly,

$M_x'(0) = \frac{(b^3 - a^3)}{3(b-a)}$

Hence,

$\text{Var}(X) = \frac{(b^3 - a^3)}{3(b-a)} - \frac{a+b}{2} = \frac{b^2 + ab + a^2)}{3} - \frac{a+b}{2} = \frac{1}{6}[2(b^2 + ab + a^2) - 3(a+b)]$

$F(x) = P(X \le x) = \begin{cases} \frac{x-a}{b-a} & a \le x \le b \\ 1 & x > b \end{cases}$

5.2.8 The Beta Distribution

A continuous random variable X is said to have a beta distribution if its pdf is given by α, β > -1

$f(X) = \begin{cases} \frac{\alpha + \beta + 1)!}{\alpha!\,\beta!} x^\alpha (1-x)^\beta, & 0 < x < 1 \\ 0 & elsewhere \end{cases}$

f(x) becomes the pdf of a uniform distribution over (0, 1) when α = β = 0.

Properties of the Beta distribution

132

1. $E(x^r) = \int_0^1 x^r \frac{(\alpha + \beta + 1)!}{\alpha! \, \beta!} x^\alpha (1 - x)^\beta \, dx = \frac{(\alpha + \beta + 1)!}{\beta! \, \alpha!} \int_0^1 x^{r+\alpha}(1 - x)^\beta \, dx.$

From definition (Beta function)

$$\frac{\alpha! \, \beta!}{(\alpha + \beta + 1)!} = \int_0^1 x^\alpha (1 - x)^\beta \, dx$$

and

$$\int_0^1 x^{r+\alpha}(1 - \alpha)^\beta \, dx = \frac{(r + \alpha)! \, \beta!}{(r + \alpha + \beta + 1)!}$$

We have

$$E(x^r) = \frac{(\alpha + \beta + 1)!}{\beta! \, \alpha!} \times \frac{(r + \alpha)! \, \beta!}{(r + \beta + 1)!} = \frac{(\alpha + \beta + 1)!(r + \alpha)!}{\alpha!(r + \beta + 1)!}.$$

The moment generating function of X does not have a simple form.

5.2.9 The Weibull Distribution

A random variable is said to have a Cauchy distribution with parameter θ, $-\infty < \theta < \infty$, if its pdf is given by

$$f(x) = \frac{1}{\pi} \cdot \frac{1}{[1 + (x - \theta)^2]}, \quad -\infty < x < \infty$$

For an illustration of a Cauchy distribution random variable. See example 6.12

Exercises

1.   If X is N(0,1), determine the following probabilities

   (i) $P(X \geq 0.9)$
   (ii) $P(X \leq -0.7)$
   (iii) $P(0.7 \leq X \leq 0.9)$
   (iv) Find a such that $P(X < -a)$ 0.10
   (v) $P(|X| \leq 1.5)$
2.   A random value X has the pdf.

$$f(x) = \begin{cases} cx & 0 < x < 1 \\ 0 & elsewhere \end{cases}$$

Determine the value of the constant c. find F(x) and the mean and variance of X.

3. A random variable X has the pdf.

$$f(x) = \begin{cases} x & 0 \le x \le 1 \\ 2 - x & 1 \, x \le 2 \\ 0 & elsewhere \end{cases}$$

(i) Show that f(x) is actually a pdf
(ii) Find the means and variance of X.

4. Let X bed a continuous random variable having probability density function

$$f(x) = \frac{1}{2} e^{|x|}, \, -\infty < x < \infty$$

Find (i) P(-2<X<2) (ii) EW(X) (iii) Var (X)

5. Let X be a Normal random variable with mean 8 and variance 2. Find P(-2<X<6). Determine the values of a and b such that $P(X \le a) = 0.25$ and $P(X \ge b) = 0.25$

6. Suppose that X is N(5, 9). Find a constant c such that $P(X > c) = 2P(X \le c)$.

7. Let a point be chosen uniformly in the interval (a, b).). Let let X denote the distance of the point chosen from a. find the probability distribution function F(x) of X.

8. (a) Show that $\int_{-\infty}^{\infty} e^{-\frac{1}{2}x^2} \, dx = \sqrt{2\pi}$

(b) Show that f(x) = 1-|α|, -1 < x < 1 is a pdf

9. Show that $\int_{\lambda}^{\infty} \frac{y^{k-1}}{\Gamma(k)} e^{-y} \, dy = \sum_{x=0}^{k-1} \frac{\lambda^x e^{-\lambda}}{x!}$.

10. The lifetime in hours of an electric light bulb is a random variable with probability density function

$$f(x) = \theta \exp\left(-\frac{1}{250} x\right), x > 0$$

Determine the value of θ

11. A probability density function f of random variable X is defined by

$$f(x) = \begin{cases} 0, & x < 0 \\ x, & 0 < x < 1 \\ \frac{1}{2}, & 1 < x < c \\ 0, & x > c \end{cases}$$

Where c is a constant

   (i) Calculated the value of c and hence the mean and variance of X.
   (ii) Sketch the graph of f.
12.   A random variable X has pdf.

   $F(x) = k \exp \{-(x-2)^2/9\} < x <$

   (i) Find the value of k,
   (ii) What is the probability distribution of X?
13.   Suppose $(x) = \dfrac{f(x)}{1 - F(x)}$, where F(x) = f(x) x > 0

   Show that $F(x) = 1 - \exp - \{\int_0^\lambda \lambda(x)dx\}$

14.   The pdf of a random variable X is defined by

   $f(x) = \frac{1}{2} e^{|x|}$, $-\infty < x < \infty$

   Find $P(1\leq) /x/ \leq 2$

15.   (a) If X is N (o^2) find k (as a function of u and 0^2 such that P(X > k) = 2P (X $\leq$k)

16.   Let X be a random variable with pdf

   (a) $f(x) = \lambda e^{-\lambda(x-a)}$ $x \geq a$;

   (b) $f(x) = \frac{1}{2}{}^{-\lambda(x-a)}$ $-\infty < x < \infty$

   Find the M.G.F of X and hence the mean and variance of x in each case

   (a) F(x) = e x $\geq$ a;
   (b) F(x) = 1/2 e < x <
   Find the M.G.F of X and hence the mean and variance of X in each case

17.   A random variable X has mean 5 and variance 3. If the pdf of X is given by f(x) =
   $\dfrac{1}{b-a}$, if a < x < b and zero otherwise. Find a and b

18. If f(x) = in 1/x, 0 < x < 1 and 0, otherwise show that f(x) is a pdf

19. Let X be a random variable with zero expectation and probability density function

    given by f(x) = $\begin{cases} \alpha\, e^x & -\infty < x < \beta \\ 0 & elsewhere \end{cases}$

    Find α.

20. Let f(x) = ½ (k) = ½ (k-1)/ (1+/x/)k, k ≥ 1-∞ < x< ∞show that f is pdf

UNIT 2

JOINTLY DISTRIBUTED RANDOM VARIABLES

6.1   Bivariate Distribution

In our study so far, we have considered only univariate distribution, Univariate distribution depends on only one random variable but we are frequently interested in specifying two or more random variables for a member of a population and interested in probability statements. For example, the age A and the systolic blood pressure, P may be of interest and we would consider (A,P) as a single experimental outcome. We might study the height H and the weight W of a chosen person given rise to the outcome (h, w). A probability distribution which depends on two random variables. Table (1) gives the joint variables. Table (1) gives the joint probability density function of two discrete random variables X, number of white marbles, and Y, number of blue marbles in a sample of 3 chosen at random from a box containing 4 red, 3 white and 5 blue marbles.

Table 1

X/Y	0	1	2	3	
0	1/55	3/22	2/11	1/22	21/55
1	9/110	3/11	3/22	0	2755
2	3/55	3/44	0	0	27/220
3	1/220	0	0	0	1/220
	35/220	21/44	7/22	1/22	1

We denote the probability that X takes the value of x and Y the value y by f(x, y) where f(x, y) = P(X = x, Y = y).

For example, the probability that the sample contains 1 white and 2 blue marbles is denoted by f(1, 2). From the table,

137

$$f(1, 2) = 3/22$$

This is calculated thus:

$$P(X = 1, Y = 2) = \frac{{}^3C_1 \times {}^5C_2}{{}^{12}C_3} = \frac{3 \times 10}{220} = 3/22.$$

The reader should verify the entries in table (1).

Note:

(i) $\sum\sum f(x, y) = 1$.

(ii) $0 \leq f(x, y) \leq 1$ for all (x, y).

(iii) the row totals form what is called the marginal distribution of X denoted by $f_x(x)$ and the column totals is called the marginal distribution of Y denoted by $f_y(y)$.

From table (1) the probability of obtaining a particular value of X irrespective of the value of Y is given below.

Table 2

x	0	1	2	3
$f_x(x)$	21/55	27/55	27/220	1/220

Similarly for Y, we have

y	0	1	2	3
$f_Y(x)$	35/220	21/44	7/22	1/22

If we know the joint probability density function of two random variables X and Y, then we can compute the marginal probability density function $f_x(x)$ by summing over y and the marginal probability density function $f_y(y)$ by summing over x i.e.

$$f_x(x) = \sum_y f(x,y)$$

$$f_y(y) = \sum_x (f(x,y)) \tag{1}$$

Note: It may happen that X is discrete while Y is continuous, but in most applications it is either both are discrete or both are continuous.

Definition 1

Two discrete random variables X and Y are said to be independent if their joint density function is give by

$$f(x, y) = f_x(x) f_Y(y)$$

i.e.

$$P(X = x, Y = y) = P(X = x) P(Y = y) \text{ for all pairs of x, y.}$$

In the above example,

$$f_x(1) = \frac{27}{55}$$

$$f_y(2) = \frac{7}{22}$$

$$f(1, 2) = \frac{3}{22} \neq \frac{27}{55} \times \frac{7}{22}$$

hence X and Y are dependent (not independent) random variables.

6.2 Expectation of Sum and Product of Jointly Distributed Discrete Random Variables

Let f(x, y) be the joint probability density function of X and Y and $f_x(x)$, $f_y(y)$, the marginal probability densities of X and Y respectively. Then

$$E(X) = \sum x f_x(x)$$

and

$$E(Y) = \sum_y y f_y(y).$$

For example, consider the variables X and Y having the joint pdf given by the table (1). The marginal pdf is given by table (2). Then

$$E(X) = \left(0 \times \frac{21}{55}\right) + \left(1 \times \frac{27}{220}\right) + \left(2 \times \frac{27}{220}\right) + \left(3 \times \frac{1}{220}\right) = 0.38$$

$$E(Y) = \left(0 \times \frac{35}{230}\right) + \left(1 \times \frac{21}{44}\right) + \left(2 \times \frac{7}{22}\right) + \left(3 \times \frac{1}{22}\right) = 2.61$$

and

$$E(X + Y) = E(X) + E(Y) = 2.99.$$

Expectation of Product

The expectation of product of two discrete random variables is defined as

$$E(XY) = \Sigma_y \quad \Sigma_x xyf(x,y)$$

For the above example we have

$$\left(\quad (0 \times 0) \times \frac{1}{55}\right) + \left(\quad (0 \times 1) \times \frac{3}{22}\right) + \left(\quad (0 \times 2) \times \frac{2}{11}\right) + \left(\quad (0 \times 3) \times \frac{1}{22}\right)$$

$$+ \left((1 \times 0) \times \frac{9}{110}\right) + \left((1 \times 1) \times \frac{3}{11}\right) + \left((1 \times 2) \times \frac{3}{22}\right)$$

$$+ \left((2 \times 0) \times \frac{3}{55}\right) + \left((2 \times 1) \times \frac{3}{44}\right) \times \left((3 \times 0) \times \frac{1}{220}\right)$$

$$= \left(1 \times \frac{3}{11}\right) + \left(2 \times \frac{3}{22}\right) + \left(2 \times \frac{3}{44}\right) = \frac{30}{44}$$

Variance of Sum of Two Random Variables

From the definition variance $E(X^2) - [E(X)]^2$ replacing X by X + Y we have Var(X + Y)
$= E(X + Y)]^2$

$$= E(X^2 + Y^2 + 2XY) - [E(X)]^2 - [E(Y)]^2 - 2E(X)E(Y)$$
$$= E(X^2) + E(Y^2) + 2E(XY) - [E(X)]^2 - [E(Y)]^2 - 2E(X)E(Y)$$
$$= E(X^2) - [E(X)]^2 + E(Y^2) - [E(Y)]^2 + 2[E(XY) - E(X)E(Y)]$$
$$= Var(X) + Var(Y) + 2[E(XY) - E(X)E(Y)]$$
$$= Var(X) + Var(Y) + 2Cov(X, Y)]$$

Where Cov(X<Y)= E(XY) – E(X)E(Y)

Definition 2: Covariance

The covariance of two random variables X and YU denoted by Cov(X< Y) is denoted by

$$Cov(X, y) = E\{(X - E(X))((Y - E(Y)\}$$
$$= E[XY - XE(Y) - YE(X) + E(X)E(Y) = E(XY) - E(X) E(Y)$$

Example 1

Given the following joint probability density function of X and Y. Find

(i)   The marginal probability density function of X and Y

(ii)   The expectation of X and Y

(iii)   The covariance of X and Y

(iv)   The variance of X + Y.

X/Y	1	2	3
3	.1	.2	0
4	0.2	0	.1
6	.15	.15	.1

(i)   Summing along the rows we have

x	3	4	6
$f_X(x)$	.3	.3	.4

Summing along the column, we have

y	1	2	3
$f_Y(y)$	0.45	0.35	0.2

(ii) $E(X) = (3 \times 0.3) + (4 \times 0.3) + (6 \times 0.4) = 4.5$

Similarly $E(Y) = (1 \times 0.45) + (2 \times 0.35) + (3 \times 0.2) = 1.75$

(iii) $Cov(X, y) = E(XY) - E(X)E(Y)$

$E(XY) = \sum\sum xyf(x, y)$

$= ((3 \times 1) \times 0.1) + ((3 \times 2) \times 0.2) + ((4 \times 1) \times 0.2)$

$+ ((4 \times 3) \times 0.1) + ((6 \times 1) \times 0.15) + (12 \times 1.5) + ((6 \times 3) \times 0.1) = 8.0$

Hence,

$$Cov(X, Y) = 8.0 - (4.5)(1.75) = 0.125$$

(iv) $E(X^2) = 21.9$

$$Var(X) = 21.9 - 20.25 = 1.65$$

$$E(Y^2) = 3.65$$

Hence,

141

$$Var(X + Y) = Var(X) + Var(Y) + 2Cov(X, Y) = 1.65 \ + \ 0.25 \ = 2.49.$$

6.3   Independent Random Variables

X and Y are said to be independent if

$$P(X = x, Y = y) = P(X = x) \, P(Y = y)$$

For all pairs (x. y). That is,

$$f(x, y) = f_x(x)f_y(y) \text{ for all pairs } (x, y).$$

thus for any subset A and Bof R

$$P(X \ \varepsilon \ A, Y \ \varepsilon \ B) = P(X \ \varepsilon \ A) \, P(Y \ \varepsilon \ B).$$

Example 2

Two random variables X and Y have the joint probability distribution given by

X/Y	1	2
1	0.2	0.4
2	0.3	0.1

Show that X and Y are not independent.

<div align="center">

Solution

</div>

The marginal pdf of X is

x	1	2
$f_X(x)$	0.6	0.4

And that of Y is

Y	1	2
$f_y(y)$	0.5	0.5

$$P(X = 1, Y = 1) = 0.2$$

$$P(X = 1) = 0.6, P(Y = 1) = 0.5$$

$$P(X = 1) \, P(Y = 1) = 0.30 \neq P(X = 1, Y = 1) = 0.2$$

Hence X and Y are not independent.

Consider two independent random variables X and Y. then

$$E(XY) = E(X)E(Y).$$

To see this, note that

$$E(XY) = \sum_x \; \sum_y xyf(x,y) \; = \; \sum_x \; \sum_y xy f_x(x) f_y(y)$$

$$= [\sum xf_x(x)] \, [\sum yf_y(y)] = E(X)E(Y).$$

Thus if two random variables X and Y are independent, then

$$E(XY) = E(X)E(Y) \hspace{4cm} (3)$$

That is Cov(X, Y) = 0.

Note that independent and zero covariance are not the same. That is if $E(XY) = E(X)E(Y)$ it does not imply that X and Y are independent.

Definition 3: Correlation Coefficient

Let X and Y be two random variable having finite non-zero variance. One measure of degree of dependence between X and Y is the correlation coefficient denoted by p(X, Y) and defined by

$$\rho = \; \rho \, (X, Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}$$

it can be shown that $-1 \le p \le 1$.

If X and Y are independent random variables then

(i) $\rho = 0$

(ii) $Var(X + Y) = Var(X) + Var(Y)$

(iii) $Var(aX + bY) = a^2 Var(X) + b^2 Var(Y)$ where a and b are constants.

(iv) $Var(X - Y) = Var(X) + Var(Y)$.

6.4   Conditional Probability Density Function

In chapter Two, we defined the conditional probability of an event A given event B by

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

Similarly, this idea can be extended to random variables. Thus

$$P(X = x \mid Y = y) = \frac{P(X = x \ and \ Y = y)}{P(Y = y)}$$

Definition 4

If X and Y are two jointly distributed discrete random variables, trhe conditional probability density function of X given Y = y is denoted and defined by

$$f_X\mid_y(x \mid y) = P(X = x \mid Y = y) = \frac{P(X = x, Y = y)}{P(Y = y)} = \frac{f(X, y)}{f_Y(y)}.$$

It follows that

$$f_X\mid_y(x \mid y) = \frac{f(X, y)}{f_Y(y)}.$$

And

$$F(x, y) = f_X\mid_y(x \mid y)f_y(y).$$

Example 3

Let the joint probability density function of X and Y be given by the table below.

X/Y	1	2	3
3	.1	.2	0
4	.2	0	.1
6	.15	.15	.1

Compute the conditional probability density function of X given that Y = 2.

Solution

From the joint p.d.f. table we have using the formula

$$f_X\mid_y(x \mid y) = \frac{f(X, y)}{f_Y(y)}.$$

$$f_X\mid_{y=2}(3 \mid 2) = \frac{f(3, 2)}{f_Y(2)} = \frac{P(X = 3, Y = 2)}{P(Y = 2)} = \frac{0.2}{0.35} = \frac{4}{7}.$$

$$f_X\mid_{y=2}(4 \mid 2) = \frac{P(X = 4, Y = 2)}{P(Y = 2)} = 0$$

$$f_X|_{y=2}(6|2) = \frac{P(X=6,Y=2)}{P(Y=2)} = \frac{0.15}{0.35} = \frac{3}{7}.$$

Hence, the conditional pdf of X given Y = 2 is

X	3	4	6
$f_X|_{y=2}(x|2)$	$\frac{4}{7}$	0	$\frac{3}{7}$

It is left as an exercise to show that the conditional pdf of X given that Y= 1,3 are

1. $f_X|_1(x|1)$

x	3	4	6
$f_X|_1(x|1)$	$\frac{2}{9}$	$\frac{4}{9}$	$\frac{3}{9}$

2. $f_X|_3(x|3)$

X	3	4	6
$f_X|_3(x|3)$	0	$\frac{1}{2}$	$\frac{1}{2}$

Definition 5: Conditional Expectation

The conditional expectation of X given Y = y is defined by

$$E(X|y) = \sum x f_X|_y(x|y)$$

Where $f_X|_y(x|y)$ is the conditional probability density function of X given Y = y.

Definition 6: Conditional Variance

The conditional variance of X given y is defined by

$$Var(X|y) = E(X^2|y) - [E(X|y)]^2$$

Example 4

Calculate the conditional expectation and conditional variance of X given that Y = 2 where the joint pdf of X and Y is given in Example 3. From example 3, we have

$$E(X \mid Y = 2) = \left(3 \times \frac{4}{7}\right) + \left(6 \times \frac{3}{7}\right) = \frac{30}{7} = 4\frac{2}{7}$$

$$E(X^2 \mid Y = 2) = \left(9 \times \frac{4}{7}\right) + \left(36 \times \frac{3}{7}\right) = \frac{144}{7}$$

Hence condition variance of X given that Y = 2 is

$$\text{Var}(X \mid Y = 2) = E(X^2 \mid Y = 2) - [E(X \mid Y = 2)]^2 = \frac{144}{7} - \left(\frac{30}{7}\right)^2 = 2.20$$

Exercise 6.1

1. Two discrete variables X and Y have the joint probability distribution given below.

X/Y	-1	5	7
0	0.1	0.2	K
1	0.05	0	0.15
2	0.3	0.1	0

(a) Find the value of k,   (b) Find the marginal distribution of X and Y and

(c) Are X and Y independent?

2. Given the following joint probability density function of random variables X and Y.

X/Y	1	2	3
1	1/8	1/8	¼
2	3/8	0	0
3	0	1/8	0

Find

1. The marginal probability density function of X and Y and hence E(X) and E(Y).

2. Are X, Y independent?

3. Find the conditional pdf of Y given that X = x for which $E(Y \mid x)$ is defined.

3. The joint probability distribution of two random variables X and Y is given by

X/Y	-1	0	3
2	.1	0	.1
3	.2	.1	0
4	0	.3	.2

(a) Determine whether or not X and Y are independent.

(b) Calculate the correlation coefficient, P(X, Y).

(c) Find the variance of X + Y.

4. Suppose a box has 3 balls labeled 1, 2, 3. Two balls are drawn at random one after the other without replacement. Let X and Y denote the number on the first and second balls drawn respectively

(a) Find the joint probability distribution of X and Y

(b) Determine whether or not X and Y are independent.

5. Suppose the situation is as an exercise 4, except now that the two balls are selected with replacement.

(a) Find the joint probability distribution of X and Y

(b) Determine whether or not X and Y are independent.

(c) Find the Var(X + Y).

6. Let X and Y be two random variables having the joint probability density function given by the following table.

X/Y	-3	-1	0	2	5
-1	1/20	1/10	0	0	3/20
3	1/40	1/40	1/20	1/20	0
4	3/40	1/40	1/40	1/20	1/20
5	1/20	1/40	1/10	1/20	1/10

Compute the following probabilities

(i)   X is even   (ii)   Y is odd

(ii)   X + Y is even   (iv)   XY is odd

(v)   X > 0 and Y > 0   (vi)   X is even given that Y is even

(vii)   X > 0 given that Y is ≤ 0.

6.5   Multinomial Distribution

Multinomial distribution is a generalization of the binomial distribution in which each of n independent identical experiments can result in any of k possible outcomes with probabilities

$$p_1, p_2, \ldots p_k, \sum p_k = 1.$$

For k= 2 we have the binomial distribution.

Consider an experiment, such as rolling a die, that can result in only a finite number of k distinct outcomes at let X_i be the number of the n experiments that result in outcome number i.

Then

$$P(X_1 = x_1, X_2 = x_2 = \ldots X_k = x_k) = p_1^{x_1} \cdot p_2^{x_2} \ldots p_k^{x_k} \cdot c(x_1, x_2, \ldots x_k)$$

Where $c(x_1, x_2, \dots x_k)$ = number of possible sequences of outcomes for the n experiments to yield

$$X_1 = x_1, X_2 = x_2 = \dots X_n = x_n.$$

From our knowledge of mathematics of counting,

$$c(x_1, x_2, \dots x_k) = \frac{n!}{x_1! x_2! \dots x_k!}$$

can be regarded as the number of ways n objects can be partitioned into k classes such that class contains x_i objects (I = 1, 2, …k). Thus,

$$P(X_1 = x_1, X_2 = x_2, \dots X_k = x_k) = \frac{n!}{x_1! x_2! \dots x_k!} p_1^{x_1} p_2^{x_2} \dots p_k^{x_k}$$

$$\sum_1^k x_i = n, \ \sum_1^k p_i = 1.$$

Example 5

Suppose that a fair die is rolled 12 times. Find the probability that 1 and 6 appear 2 times each. 3 thrice each, 3 times, and 4 and 5 once each.

Solution

Let X_i be the number of times ith outcome appears.

$$n=12, \ x_1 = 2, \ x_2 = 3, \ x_3 = 3, \ x_4 = 1, \ x_5 = 1, \ x_6 = 2, \ p_i = \frac{1}{6}.$$

Hence,

$$P(X_1 = 2, X_2 = 3, x_3 = 3, x_4 = 1, x_5 = 1, x_6 = 2)$$

$$= \frac{12!}{2!3!3!1!2!} \left(\frac{1}{6}\right)^2 \left(\frac{1}{6}\right)^3 \left(\frac{1}{6}\right) \left(\frac{1}{6}\right) \left(\frac{1}{6}\right)^2 \left(\frac{1}{6}\right)^3 = \frac{12!}{2!3!2!3!} \left(\frac{1}{6}\right)^{12}.$$

Example 6

In a certain large population, 70% are right-handed, 20% left handed and 10% are ambidextrous. If 10 persons are chosen at random from the population, what is the probability that (i) all are right-handed? (ii) 7 are right-handed, 2 are left handed and 1 is ambidextrous.

Solution

$$P_1 = 0.7, P_2 = 0.2, P_3 = 0.1$$

Let X_1, X_2, X_3 denote number that are right-handed and ambidextrous respectively, then

$$P(X_1 = x_1, X_2 = x_2 = X_3 = x_3) = \frac{n!}{x_1! x_2! x_3!} p_1{}^{x_1} p_2{}^{x_2} p_3{}^{x_3}$$

(i) $P(X_1 = 10, X_2 = 0 = X_3 = 0) = \frac{10!}{10!}(0.7)^{10}(0.2)^0(0.1)^0 = (0.7)^{10} = 0.028$

(ii) $P(X_1 = 7, X_2 = 2 = X_3 = 1) = \frac{10!}{7!2!1!}(0.7)^7(0.2)^2(0.1)^1 = 0.0119.$

6.6   Continuous Random Variables

In this section we will consider a pair of continuous random variables X and Y and some of their properties. The results in the earlier section, which have been proved for discrete random variables are also applicable to continuous random variables. The joint probability density function of X and Y is denoted by f(x, y) such that

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)\, dx\, dy = 1$$

and

$$f(x,y) \geq \text{ for all } (x, y)\, \varepsilon\, \mathbb{R}$$

(that is, summation in discrete case is replaced by integrals) and the marginal probability density functions are given by

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y)\, dy$$

$$f_Y(y) = \int_{-\infty}^{\infty} f(x, y)\, dx$$

The joint probability distribution function F is defined by

$$F(x, y) = P(X \leq x, Y \leq y)$$

Where X, Y are defined on the same probability space.

$$F(x, y) = \int_{-\infty}^{y} \int_{-\infty}^{x} f(u, v)\, du\, dv$$

If B is in the range space of (X, Y)

$$P(X, \varepsilon <\, = B) = P(B) = \int \int_B f(x,y)dxdy$$

and

$$\frac{\partial^2}{\partial x\, \partial y} F(x,y) = f(x,y).$$

The marginal probability distribution function are defined yb

$$F_x(x) = P(X \leq x) \text{ and } F_y(y) = P(Y \leq y).$$

From the joint probability distribution function we can get the marginal distribution functions by the following relationship

$$F_X(x) = F(x, \infty) = \lim_{y \to \infty} F(X, y).$$

$$F_Y(y) = F(\infty, y) = \lim_{x \to \infty} F(x, y).$$

Let X and Y be any continuous random variables. Then X and Y are independent if and only if

$$f(x, y) = f_X(x)f_Y(y), \quad -\infty < x < \infty, \quad -\infty < y < \infty$$

The conditional probability density function of X is defined by

$$f_{x|y}(x|y) = \frac{f(x,y)}{f_y(y)}.$$

Similarly, the conditional probability function of Y given X is

$$f_{Y|x}(y|x) = \frac{f(x,y)}{f_x(x)}.$$

The conditional expectation is defined by

$$E(Y \mid X=x) = \int_{-\infty}^{\infty} y\, f_Y \mid_x \left(y \mid_x \right) dy$$

151

Example 7

Let X and Y have the joint pdf

$$f(x, y) = \begin{cases} 2 & 0 < x < y < 1 \\ 0 & \text{elsewhere} \end{cases}$$

Find
(i)   The marginal probability density functions.
(ii)   Determine whether or not X and Y are independent
(iii)   Find the conditional expectation of X given Y = y.

Solution

The marginal probability density function is given by
(i) $f_X(x) = \int f(x,y)dy, x < y < 1, 0 < x < 1$

$$f_X(x) = \int_x^1 2\,dy = 2yI\Big|_x^1 = 2(1-x)$$

and

$$f_Y(y) = \int f(x,y)dx, 0 < x < y$$
$$= \int_0^y 2dx = 2xI\Big|_0^y = 2y$$

Hence

$$f_x(x) = \begin{cases} 2(1-x) & 0 < x < 1 \\ 0 & \text{elsewhere} \end{cases}$$

and

$$f_y(y) = \begin{cases} 2y & 0 < y < 1 \\ 0 & \text{elsewhere} \end{cases}$$

(ii) $f_x(x)f_Y(y) = 2(1 - x)\,2y = 4y(1-x) \neq f(x, y)$
        So X and Y are not independent.

(iii) $E(X\,|\,y) = \int xf_{x\,|\,y}\left(x\,\Big|\,y\right)dx = \int x \frac{f(x,y)}{f(y)}dx = \int x \frac{2}{2y}dx = \int \frac{x}{y}dx.$

$$= \frac{1}{y} \int_0^y x\,dx, \quad 0 < x < y$$
$$= \frac{1}{2}\,y, 0 < y , 1.$$

Example 8
Let the random variables X and Y have the joint pdf.

$$x+y \quad 0 < x < 1, 0 < y < 1$$

$$f(x, y) =$$
$$0 \qquad \text{elsewhere}$$

Find (i) the marginal probability density functions, (ii) are X, Y independent? And (iii) find the Cov(x, Y).

Solution

(i) $f_X(x) = \int f(x, y)\, dy,\ 0 < y < 1$

$$= \int_0^1 (x + y)\, dy = xy + \tfrac{1}{2}y^2 \, I_0^1 = x + \tfrac{1}{2}$$

$$f_Y(y) = \int_0^1 (x + y)\, dx,\ 0 < x < 1$$

$$= \tfrac{1}{2}x^2 + xy \left[I_0^1 \right. = \tfrac{1}{2} + y.$$

Hence,

$$f_X(x) \ = \ \begin{cases} \tfrac{1}{2} + x & 0 < x < 1 \\ \\ 0 & \text{elsewhere} \end{cases}$$

$$f_Y(y) \ = \ \begin{cases} \tfrac{1}{2} + y & 0 < y < 1 \\ \\ 0 & \text{elsewhere} \end{cases}$$

(ii) $f(x, y) \neq f_X(x)f_Y(x))$. Thus, X and Y are not independent.

(iii) $E(X) = \int_0^1 x f_x(x)\, dx = \int_0^1 x \left(\tfrac{1}{2} + x \right) dx = \tfrac{7}{12}$

$$E(Y) = \int_0^1 y\, f_Y(y)\, dy$$
$$E(Y) = \tfrac{7}{12}.$$

$$E(XY) = \int \int xyf(x,y)\,dxdy = \int_0^1 \int_0^1 xy(x + y)\,dxdy = \int_0^1 \left[\int_0^1 (x^2 y + xy^2)\,dx \right] dy$$

$$\int_0^1 (x^2 y + xy^2)\,dx = 1/3x^3 y + \tfrac{1}{2}x^2 y^2 \, I_0^1 = \left(\tfrac{1}{3} \right) y + \tfrac{1}{2}y^2$$

$$\int_0^1 \left(\tfrac{y}{3}y + \tfrac{1}{2}y^2 \right) dy = \tfrac{y^2}{6} + \tfrac{y^3}{6} \, I_0^1$$

Hence

$$E(XY) = 1/3$$

Thus,

$$\text{Cov}(X, Y) = 1/3 - 7/12.7/12. = -1/144.$$

Example 9

Let the random variables X and Y have the joint pdf

$$f(x, y) = \begin{cases} \dfrac{1}{x} & 0 < y < x < 1 \\ \\ 0 & \text{elsewhere} \end{cases}$$

Find the marginal p.d.f of X and of Y and the conditional pdf of Y given X = x.

Solution

$$f_X(x) = \int f(x, y)dy, \quad 0 < y < x, 0 < x < 1$$

$$= \int_0^x \frac{1}{x} dy = \frac{1}{x} y \Big|_0^x = 1.$$

Therefore

$$f_X(x) = \begin{cases} 1, & 0 < x < 1 \\ \\ 0 & \text{elsewhere} \end{cases}$$

$$F_Y(y) = \int_y^1 f(x, y)dx = \int_y^1 \frac{1}{x} dx = -\log y,$$

Therefore

$$f_y(y) = \begin{cases} -\log y & 0 < y < 1 \\ \\ 0 & \text{elsewhere} \end{cases}$$

$$f_Y|_x(y \,|\, x) = \frac{f(x,y)}{f_x(x)} = \begin{cases} \dfrac{1}{x} & 0 < y < x < 1 \\ \\ 0 \end{cases}$$

Hence X and Y/x are uniformly distributed over (0, 1), (0, x) respectively.

Example 10

Let the random variables X and Y have the joint pdf

$$f(x, y) = \begin{cases} x^2 + \dfrac{xy}{3} & 0 \le x \le 1, 0 \le y \le 2 \\ \\ 0 & \text{elsewhere} \end{cases}$$

Find the $P(X + Y \le 2)$.

<div align="center">**Solution**</div>

Let $A \equiv \{x, y: x + y \le 2\}$.

$$P(A) = \int_0^1 \int_0^{2-x} f(x, y)\, dy\, dx = \int_0^1 \int_0^{2-x}\left[x^2 + \frac{xy}{3}\right] dy\, dx$$

$$= \int_0^1 \frac{4x + 8x^2 - 5x^3}{6}\, dx = 0.57.$$

Exercise 6.2

1. Suppose that X and Y have the joint p.d.f

$$f(x, y) = \begin{cases} K\, e^{-\lambda(x+y)} & x \ge 0,\ y \ge 0 \\ 0 & \text{elsewhere} \end{cases}$$

   Find (i) K and (ii) the marginal pdf of X and of Y.

2. Suppose that X and Y have the joint pdf

$$f(x, y) = \begin{cases} C(x-y) & 0 < x < 2 \\ & -x < y < x. \\ 0 & \text{elsewhere} \end{cases}$$

   (a) Find (i) c and (ii) the marginal pdf of X and of Y.
   (b) Compute the following (i) $P(x < 1)$ and (ii) $P(0 < Y < 2)$.

3. Suppose that X and Y have the joint pdf

$$f(x, y) = \begin{cases} \left[\frac{5x^2}{12} + \frac{xy}{3}\right]K & 0 \le x \le 2 \\ 0 & \text{elsewhere} \end{cases}$$

   (i) Find K using the fact that $\int\int f(x, y)\, dx\, dy = 1$.
   (ii) Are X and Y independent random variables.

4. Two random variables X and Y have the joint pdf given by

$$f(x, y) = \begin{cases} c(x - 2xy + y) & 0 \le x \le 1,\ 0 \le y \le 1 \\ 0 & \text{elsewhere} \end{cases}$$

   (a) Show that X and Y are uniformly distributing over [0, 1].
   (b) Find the joint probability distribution function of X and Y. Hence, find the marginal distribution functions of X and Y.

5. Suppose that X and Y have the joint pdf

$$f(x, y) = \begin{cases} \lambda^2 e^{-\lambda y} & 0 \le x \le y \\ 0 & \text{elsewhere} \end{cases}$$

   (i) Find the marginal probability density functions of X and of Y.

6. Suppose that X and Y have the joint pdf

$$f(x, y) = \begin{cases} c\, e^{-\frac{1}{2}(x^2 - xy + y^2)} & -\infty < x < \infty \\ 0 & -\infty < x < \infty \end{cases}$$

   (i) Determine the value of c.

(ii)   Are X and Y independent random variables?

7.   Let X and Y have the probability density function defined by

$$f(x, y) = \begin{cases} 2 & \text{if (x,y) is inside the triangle having variances at} \\ & (0, 0), (0,1), (1, 0) \\ 0 & \text{elsewhere} \end{cases}$$

Find the marginal probability density function of X and Y

8.   Let $X_1/X_1 + X_2$)   (ii) $E(X_1 | X_1 + X_2 = m)$.

9.   A fair die is successively rolled. Let X and Y denote, respectively, the number of rolls necessary to obtain a 6 and a 5. Find (a) E(X)   (b) $E(X | Y = 1)$, and (c) $E(X | Y = 5)$.

10.   Let the random variables X and Y have the joint pdf

$$F(x, y) = \frac{e^{-y}e^{-x/y}}{y}, 0 < x < \infty; \qquad 0 < y < \infty. \text{ Find } E(X | Y = y).$$

11.   If it is assumed that, in single – car accidents in a certain area, the probabilities of minor severe, and fatal injuries to the driver are 0.5, 0.4 and 0.1 respectively, find the probability that in 10 accidents there are 4 minor, 4 severe and 2 fatal injuries to the 10 drivers.

Example 2

Let X have the poisson pdf with parameter λ. Find the pdf of $Y = X^2$

$$f(x) = \frac{e^{-\lambda}\lambda^X}{x!}; X = 0, 1, 2, ...$$

Solution

Possible values of Y are: 0, 1, 4, 9, 16…

$$P(Y = y) = P(X^2 = y) = P(X = + \sqrt{y})$$

(single-valued inverse function)

$$= f(\sqrt{y}), x = \sqrt{y}$$

$$= \begin{cases} \dfrac{e^{-\lambda}\lambda^{\sqrt{y}}}{\sqrt{y}}, & y = 0, 1, 4, 9, 16 ... \\ 0 & elsewhere \end{cases}$$

Since X is a positive random variable), y = 0, 1, 4, 9, 16,…

Example 3

Suppose X is a random variable having probability density function f(x) given by

X	-2	-1	0	1	2	3
f(x)	.15	.25	0.15	.2	.15	.1

Find the p.d.f. of $Y = X^2$

Solution

Possible values of Y are 0, 1, 4, 9.

This is not a one-to-one function

$$P(Y = 0) = P(X = 0) = 0.15$$
$$P(Y = 1) = P(X = -1 \text{ or } 1) = P(X = -1) + P(X = 1) = 0.25 + 0.2 = 0.45$$
$$P(Y = 4) = P(X = -2 \text{ or } 2) = P(X = -2) + P(X = 2) = 0.15 + 0.15 = 0.3$$
$$P(Y = 9) = P(X = 3) = 0.1$$

Hence, the pdf of Y is given by

y	0	1	4	9
h(y)	0.15	0.45	0.3	0.1

Example 4

Let X has the pdf given by

$$f(x) = \begin{cases} \left(\frac{1}{2}\right)^x, & x = 1, 2 \\ 0 & elsewhere \end{cases}$$

Find the pdf of Y defined by

$$Y = \begin{cases} 1 \text{ if X is even} \\ 1 \text{ if X is odd} \end{cases}$$

Solution

$$P(Y = 1) = P(X = 2, 4, 6, 8, \ldots$$
$$= P(X = 2) + P(X = 4) + P(X = 6) + \ldots$$
$$= \left(\tfrac{1}{2}\right)^2 + \left(\tfrac{1}{2}\right)^4 + \left(\tfrac{1}{2}\right)^6 + \ldots = \tfrac{1}{3}$$

Similarly,
$$P(Y = -1) = \tfrac{2}{3}.$$

Hence, the pdf of Y is given by

Y	1	1
g(y)	$\dfrac{1}{3}$	$\dfrac{2}{3}$

Continuous Case

Let X be a continuous random variable having probability density function f(x). we wish to find the probability density of a variable $Y = g(X)$. We shall discuss methods for finding the pdf of Y.

Let $F_Y(y)$ be the probability distribution function of Y. Then, expressing $F_Y(y)$ in terms of X, we obtain
$$F_Y(y) = P(Y \leq y) = P(g(X)) \leq y)$$

$$= \begin{cases} P\,(X \leq g^{-1}(y)) & \text{if Y is a strictly increasing function of X} \\ P\,(X \geq g^{-1}(y)) & \text{if Y is a strictly increasing function of X} \end{cases}$$

Where $X = g^{-1}(y)$ is the inverse of the transformation $Y = g(X)$ and $g(X)$ is a one-to-one function.

Hence,
$$F_Y(y) = F_X(g^{-1}(y)).$$

Thus, the pdf of Y is given by
$$F_Y(y) = F_Y(y) = F_X(g^{-1}(y)).$$

Example 5

Let X be uniformly distributed over (0, 1). Find the pdf of $Y = \dfrac{-1}{\lambda}\log_e(1 - x), \ x > 0.$

Solution

$$F_Y(y) = P(Y \leq y) = P\,\tfrac{1}{\lambda}\log_e(1 - x) \ \leq y) = P(-\log(1 - x) \leq \lambda y)$$
$$= P(\log(1 - X) \geq -\lambda y) = P(1 - X \geq e^{-\lambda y}) = P(X \leq 1 - e^{-\lambda y})$$

The pdf of X is given by

$$f(x) = \begin{cases} 1 & 0 < x < 1 \\ 0 & \text{elsewhere} \end{cases}$$

and

158

$$F_X(u) = \int_0^u f(x)dx = \int_0^u 1\, dx = u$$

Hence

$$F_Y(y) = (1 - e^{-\lambda y})$$

Thus, the pdf of Y is given by

$$f(x) = \begin{cases} \lambda e^{-\lambda y} & y \geq 0 \\ 0 & \text{elsewhere} \end{cases}$$

Example 6

Let X has the pdf, $f(x) = e^{-x}$, $x > 0$. Find the pdf of $Y = 2X + 1$.

Solution

$$P(Y \leq y) = P(2X + 1 \leq y) = P\left(X \leq \frac{y-1}{2}\right)$$

$$P(X \leq x) = \int_0^x e^{-x}dx = 1 - e^{-x}$$

$$F_X\left(\frac{y-1}{2}\right) = P\left(X \leq \frac{y-1}{2}\right) = 1 - e^{-(y-1)/2}.$$

On differentiating, we obtain

$$F_Y(y) = F'_y(y) = \frac{1}{2}e^{-\frac{1}{2}(y-1)}.$$

Hence,

$$f_Y(y) = \begin{cases} \frac{1}{2}e^{-\frac{1}{2}(y-1)}, & y \geq 1 \\ 0 & \text{elsewhere} \end{cases}$$

Example 7

let X has the density function given by

$$f_x(x) = \begin{cases} \lambda x^{\lambda-1} & 0 \leq x \leq 1 \\ 0 & \text{elsewhere} \end{cases}$$

Show that $Y = -\log_e X$ is an exponential random variable.

Solution

$$P(Y \leq y) = P(-\log_e X \leq y) = P(X \geq e^{-y}) = 1 - P(X \leq e^{-y})$$

$$P(X \leq u) = \int_0^u \lambda x^{\lambda-1}dx = x^\lambda\Big|_0^u = u^\lambda$$

thus,

$$P(X < e^{-y}) = e^{-\lambda y}$$

$$P(Y \leq y) = 1 - e^{-\lambda y}; \quad f_Y(y) = \lambda e^{-\lambda y}, \quad y \geq 0$$

Hence,

$$F_Y(y) = P(Y \leq y) = 1 - e^{-\lambda y},$$

$$F_y(y) = F'_Y(y) = \begin{cases} \lambda e^{-\lambda y} & , \quad y \geq 0 \\ 0 & \text{elsewhere} \end{cases}$$

This shows that Y is an exponential random variable.

Example 8

Let X be a random variable having pdf

$$f_x(x) = \begin{cases} 2x & 0 \le x \le 1 \\ 0 & \text{elsewhere} \end{cases}$$

Find the pdf of $Y = 4x^2$.

Solution

$$F_Y(y) = P(Y \le y) = P(4X^2 \le y) = P\left(X^2 \le \frac{y}{4}\right)$$

$$= P\left(X^2 \le \frac{y}{4}\right) = P\left(\frac{-\sqrt{y}}{2} \le X \le \frac{+\sqrt{y}}{2}\right) = F_X\left(\frac{+\sqrt{y}}{2}\right) - F_X\left(\frac{-\sqrt{y}}{2}\right)$$

$$F_X(u) = \int_0^u f(x)dx = \int_0^u 2x\, dx = x^2\Big|_0^u = u^2$$

Thus,

$$F_Y(y) = F_X\left(\frac{\sqrt{y}}{2}\right)^2 ; F_X\left(\frac{-\sqrt{y}}{2}\right) = 0$$

since X cannot be less than zero. Hence,

$$F_Y(y) = \left(\frac{\sqrt{y}}{2}\right)^2 = \frac{y}{4}.$$

Thus, the pdf of Y is

$$F'_Y(y) = f_Y(y) = \begin{cases} \frac{1}{4} & 0 < Y < 4 \\ 0 & \text{elsewhere} \end{cases}$$

that is, Y is uniformly distributed over (0, 4).

Example 9
Let X be a continuous random variable having probability density function f(x). derive an expression for the probability density function of the random variable $Y = X^2$.

Solution

$$F_Y(y) = P(Y \le y) = P(X^2 \le y) = (-\sqrt{y} \le X \le +\sqrt{y}) = F_X(+\sqrt{y}) - F_X(-\sqrt{y}).$$

Differentiation gives

$$F'_X(\sqrt{y}) = \frac{1}{2\sqrt{y}} f_X(\sqrt{y}), F'_X(-\sqrt{y}) = \frac{-1}{2\sqrt{y}} f_X(-\sqrt{y}).$$

Hence

$$F'_Y(y) = f_Y(y) = \begin{cases} \frac{1}{2\sqrt{y}} & (f_X(\sqrt{y}) + f_X(-\sqrt{y}); \quad y > 0 \\ 0 & \text{elsewhere} \end{cases}$$

Example 10
Lex X be a random variable having the pdf.

$$f(x) = \begin{cases} \frac{1}{2}(x + 1); & -1 \le x < 1 \\ \end{cases}$$

<center>0   elsewhere</center>

Find the pdf of $Y = X^2$.

<center>**Solution**</center>

From example 10above, we have

$$f_Y(y) = \frac{1}{2\sqrt{y}}(f_X(\sqrt{y}) + f_X(-\sqrt{y})$$

$$f_X(\sqrt{y}) = \frac{1}{2}(\sqrt{y} + 1)$$

hence,

$$f_X(y) = \frac{1}{2\sqrt{y}}\left\{\frac{1}{2}(\sqrt{y} + 1) + \frac{1}{2}(-\sqrt{y} + 1)\right\} = \frac{1}{2\sqrt{y}}(1) = \begin{cases} \frac{1}{2}y^{\frac{1}{2}} & 0 < y < 1 \\ 0 & elsewhere \end{cases}$$

Note: The transformation $Y = X^2$ mapped $x: -1 \leq x < 1$ onto the set $y: 0 < y < 1$.

Example 11
Let X be a continuous random variable having pdf f(x). Derive the formula for the pdf of $Y = |X|$

$$F_Y(y) = P(Y \leq y) = P(|X| \leq y) = P(-y < X < y) = F_X(y) - F_X(-y)$$

By differentiation we see that

$$F_Y(y) = \begin{cases} (f_x(y) + (f_x(-y) & y > 0 \\ 0 & y \leq 0 \end{cases}$$

Theorem 1
Let $Y = g(X)$ be a differentiable strictly increasing or strictly decreasing function on an interval D and let D^* denote the range of g. let X e a continuous random variable with pdf f(x) such that f(x) = 0 for x □ D. Then the pdf of Y is given by

$$f_Y(y) = f_X(g^{-1}(y)) \left|\frac{dg^{-1}(y)}{dy}\right|; y \in D^*$$

where $x = g^{-1}(y)$ is the inverse function of g. equivalently

$$f_Y(y) = f(x) \left|\frac{dx}{dy}\right|$$

where x is expressed in terms of y, y □ D^* and $x = g^{-1}(y)$.

Proof:
Case 1: suppose that g(X) is strictly increasing
$$F_Y(y) = P(Y \leq y) = P(g(X) \leq y) = P(X \leq g^{-1}(y)) = F_X(g^{-1}(y))$$
Thus by differentiation, we have
$$F'_Y(y) = f_X(g^{-1}(y)) \frac{d}{dy}(g^{-1}(y)).$$
Since g(X) is strictly increasing so also is $g^{-1}(y)$.
Hence
$$\frac{d}{dy}[g^{-1}(y))] = \left|\frac{d}{dy}[g^{-1}(y)\right|.$$
Since $\frac{d}{dy}[g^{-1}(y)) > 0$. Thus

<center>161</center>

$$F_Y(y) = f_X(g^{-1}(y)) \left| \frac{d}{dy} [g^{-1}(y)] \right| = f_X(x) \left| \frac{dx}{dy} \right| ; x = g^{-1}(y).$$

Case II: When g(X) is a strictly decreasing function of X

$$F_Y(y) = P(Y \le y) = P(g(X) \le y) = P(X \ge g^{-1}(y))$$
$$= 1 - P(X \le g^{-1}(y)) = 1 - F_X(g^{-1})y)).$$

Thus,
$$F'_Y(y) = f_Y(y) = -f_X(g^{-1})y)).$$

Since g(X) is strictly decreasing so also is $g^{-1}(y)$.

Therefore, $\frac{d}{dy} g^{-1}(y)$ is negative and $\frac{d}{dy} g^{-1}(y) = \left| \frac{d}{dy} g - 1(y) \right|$.

Thus,
$$f_y(y) = f_X(g^{-1}(y)) \left| \frac{d}{dy} g - 1(y) \right| = f_X(x) \left| \frac{dx}{dy} \right| ; x = g^{-1}(y)$$

Example 12

Let X be a random variable having the pdf

$$f(x) = \begin{cases} 1 & 0 < x < 1 \\ 0 & \text{elsewhere} \end{cases}$$

Find the pdf of $Y = -2 \log_e x$.

Solution

Y is a strictly decreasing differentiable function of X. hence by Theorem 1

$$f_Y(y) = f_X(x) \left| \frac{dx}{dy} \right| ; x = e^{-y/2}$$
$$Y = -2\log_e x,$$

then
$$\frac{dx}{dy} = -\frac{1}{2}e^{-y/2}, \left| \frac{dx}{dy} \right| = \frac{1}{2}e^{-y/2}$$

Hence
$$f_Y(y) = 1, \frac{1}{2}e^{-y/2}, y \ge 0.$$

Thus
$$f_Y(y) = \begin{cases} \frac{1}{2} e^{-y/2} & y > 0 \\ 0 & \text{elsewhere} \end{cases}$$

Example 13

Let X be a random variable with pdf

$$f_X(x) = \begin{cases} 2x & 0 \le x \le 1 \\ 0 & \text{elsewhere} \end{cases}$$

Find the pdf of (i) $Y = 2X - 5$ (ii) $Y = X^2$.

Solution

(i)   Y is a strictly increasing function of X. thus,

$$f_Y(y) = f_X(x)\frac{dx}{dy}, x = \frac{y+5}{2}$$

$$\frac{dx}{dy} = \frac{1}{2}$$

Thus,

$$f_Y(y) = 2x, \frac{1}{2} = x = \frac{y+5}{2}$$

Hence

$$f_Y(y) = \begin{cases} \frac{y+5}{2} & -5 \le y \le -3 \\ \\ 0 & \text{elsewhere} \end{cases}$$

(ii)   Y is strictly increasing on (0, 1) and $0 \le Y \le 1$.

$$X = +\sqrt{y} \text{ and } \frac{dx}{dy} = \frac{1}{2}y^{-\frac{1}{2}}$$

thus,

$$f_Y(y) = f_X(x)\frac{dx}{dy}$$

$$f_Y(y) = 2x \left|\frac{1}{2}y^{-\frac{1}{2}}\right| = 2.y^{-\frac{1}{2}}.\frac{1}{2}y^{-\frac{1}{2}}, x = \sqrt{y} = 1.$$

Hence,

$$f_Y(y) = \begin{cases} 1 & 0 \le y \le 1 \\ \\ 0 & \text{elsewhere} \end{cases}$$

Example 14:*Chi-square Distribution*
Let X be a random variable having the $N(\mu, \sigma^2)$. Find the pdf of (i) $Y = \frac{X-\mu}{\sigma}$ (ii) $Z = Y^2$.

Solution

(i) $Y = \frac{X-\mu}{\sigma}, X = \mu + Y\sigma, \frac{dx}{dy} = \sigma$

Y is a strictly increasing function of X, hence

$$f_Y(y) = f_X(x)\left|\frac{dx}{dy}\right|; x \mu + y\sigma$$

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}}.$$

Substituting for x, we see that

$$f_X(\mu + y\sigma) = \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}y^2}$$

Thus

$$f_Y(y) = \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}y^2}.\sigma = \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}y^2}, -\infty < y < \infty.$$

That is Y is a standard normal random variable (N(0, 1)).

(ii) $Z = Y^2$

$$f_Z(z) = f_Y(y) \left|\frac{dy}{dz}\right|_{-\infty}^{0} + f_Y(y) \left|\frac{dy}{dz}\right|_{-\infty}^{0}$$

$$y = z^{\frac{1}{2}}, \frac{dy}{dz} = \frac{1}{2} z^{\frac{1}{2}}$$

Hence

$$f_z(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z} \cdot \frac{1}{2} Z^{\frac{1}{2}} + \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z} \cdot \frac{1}{2} Z^{\frac{1}{2}}$$

$$= \frac{z^{-\frac{1}{2}}}{\sqrt{2\pi}} e^{-\frac{1}{2}z} = \begin{cases} \dfrac{\left(\frac{1}{2}\right)^{\frac{1}{2}} z^{\frac{1}{2}} e^{-\frac{1}{2}z}}{\sqrt{\pi}} & 0 < z < \infty \\ 0 & elsewhere \end{cases}$$

Definition 3

A continuous random variable X that has the pdf

$$f(x) = \begin{cases} \dfrac{\left(\frac{1}{2}\right)^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2}\right)} x^{\frac{n}{2}-1} e^{-\frac{x}{2}}, & 0 < x < \infty \\ 0 & elsewhere \end{cases}$$

is said to have a chi square distribution with n degrees of freedom and is written $X^2_{(n)}$.

Let X be a random variable having the $N(\mu, \sigma^2)$ distribution.

Then the pdf of $Y = \left(\frac{X-\mu}{\sigma}\right)^2$ is $X^2_{(1)}$.

Proof: See Example 14.

The pdf of Y = g(X) when g is not one to one.

Theorem 2

Let X be a continuous random variable having pdf f(x) such that f(x) > 0, x ∈ D and f(x) = 0, x ∉ D. Suppose D can be partitioned into two sets D_1 and D_2 such that the transformation Y = g(x) which is a monotone differentiable function in D and is a strictly increasing function on D_1 and strictly decreasing differentiable function of X on D_2, then the pdf of Y is given by

$$f_Y(y) = f_X(x) \left|\frac{dx}{dy}\right|_{onD_1} + f_X(x) \frac{dx}{dy}\bigg|_{onD_2}$$

Where x = g^{-1}(y), D = $D_1 \cup D_2$

Consider the following examples.

Example 15

Let X be a continuous random variable having pdf f(x) given by

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \quad -\infty < x < \infty$$

Find the pdf of (i) $Y = |X|$ (ii) $Y = X^2$.
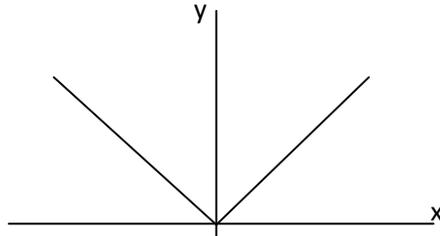
(i)   The graph of Y is given below.

Figure 1

Y is strictly increasing on $(0, \infty) = D_1$ and strictly decreasing on $(-\infty, 0) = D_2$

On D_1, $y = x$, $\dfrac{dx}{dy} = 1$.

$$f_X(y)\left|\dfrac{dx}{dy}\right| = \dfrac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}y^2}$$

On D_2, $y = x$, $\dfrac{dx}{dy} = 1$.

$$f_X(-y)\left|\dfrac{dx}{dy}\right| = \dfrac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}y^2}.$$

Thus,

$$f_Y(y) = f_X(y)\left|\dfrac{dx}{dy}\right|_{D_1} + f_X(-y)\left|\dfrac{dx}{dy}\right|_{D_2} = \dfrac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}y^2} + \dfrac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}y^2},$$

Hence

(ii)   The function $Y = X^2$ is a strictly decreasing differentiable function of X on $(-\infty, 0)$ and is a strictly increasing function of X on $(0, \infty)$. Thus $D = (-\infty, \infty)$ can be partitioned in to $D_1 = (-\infty, 0)$ and $D_2 = (0, \infty)$ so that $D = D_1 \cup D_2$.

Thus, by Theorem 2, we have

$$f_Y(y) = f_X(x)\left|\dfrac{dx}{dy}\right|_{onD_1} + f_X(x)\left|\dfrac{dx}{dy}\right|_{onD_2}$$

166

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$$

$$x = \pm\sqrt{y}$$

on D_1, $x = -\sqrt{y}$ and $\left|\frac{dx}{dy}\right| = \frac{1}{2}y^{-\frac{1}{2}}$

$$f_X(-\sqrt{y})\left|\frac{dx}{dy}\right| = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y} \cdot \frac{1}{2} y^{-\frac{1}{2}}$$

On D_2, $x = +\sqrt{y}$ and

$$f_X(\sqrt{y})\left|\frac{dx}{dy}\right| = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y} \cdot \frac{1}{2} y^{-\frac{1}{2}}$$

Hence

$$f_Y(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y} \cdot \frac{1}{2} y^{\frac{1}{2}} + \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y} \cdot \frac{1}{2} y^{\frac{1}{2}} = \begin{cases} \dfrac{y^{\frac{1}{2}} e^{-\frac{1}{2}y}}{\sqrt{2\pi}} & y \geq 0 \\ 0 & elsew\square ere \end{cases}$$

Example 16

Let X be a uniform random variable over the interval $\left(\frac{-\pi}{2}, \frac{\pi}{2}\right)$. Find the probability density function of $Y = \tan X$.

Solution

$$f_x(X) = \begin{cases} \dfrac{1}{\pi} \dfrac{-\pi}{2} < x < \dfrac{\pi}{2} \\ \\ 0 \qquad elsewhere \end{cases}$$

$Y = \tan X$ is a strictly increasing function of X in the interval $\left(\frac{-\pi}{2}, \frac{\pi}{2}\right)$ *Hence,*

$$f_Y(y) = f(x)\left|\frac{dx}{dy}\right|; x = \tan^{-1}y$$

$$\frac{1}{\pi} \cdot \frac{1}{1+y^2}$$

Thus, the pdf of Y is given by

$$f_Y(y) = \frac{1}{\pi(1 + y^2)} \qquad -\infty < y < \infty$$

This pdf is called the Cauch pdf.

Exercise 7.1

1.   Let X be a uniform random variable over the interval (-1, 1).

   (i)   Let $Y = 2X - 1$. Find the pdf of Y.

   (ii)   Find the pdf of (a) $Z = |X|$ (b) $W = \sin\left(\frac{\pi}{2}\right)X$

2.   Let X has the pdf given by

$$f(x) = \begin{cases} \frac{2^x}{2^{N+1}-2}, & x = 1, 2, \dots N(even) \\ \\ 0 & \text{elsewhere} \end{cases}$$

Find the pdf of Y defined below.

$$Y = \begin{cases} 2 \text{ if X is even} \\ \\ 5 \text{ is X is odd} \end{cases}$$

Hence, find the mean and variance of Y.

3.   Let X has pdf

$$f(x) = \begin{cases} 2xe^{-x^2}, & 0 < x < \infty \\ \\ 0 & \text{elsewhere} \end{cases}$$

Find the pdf of (i) $Y = X^2$ (ii) $Y = |X|$.

4.   Suppose that X is a continuous random variable defined over (0, 1) such that $P(X \leq 0.45) = 0$. Let $Y = 1 - X$. Find the value of K such that $P(Y \leq k) = 0.40$.

5.   Suppose that X has the pdf $f(x) = e^{-x}$, $x > 0$. Find the pdf of the following random variables, (a) $Y = X^3$ (b) $z = \frac{1}{(X+1)^2}$.

6.   Suppose that X is a geometric random variable with parameter p. Find the pdf of

(a) $Y = X^2$ (b) $Z = X + r$, where r is a constant.

7.   Find the pdf of $Y = e^x$ if X has

      (i)   Normal distribution with mean μ and variance σ^2

      (ii)   Exponential distribution with mean λ.

8.   Let X be uniformly distributed over (0, 1).

      Find the pdf of $Y = X^n$.

9.   Let X be a continuous random variable having probability distribution function F. show that $Y = F(X)$ is a uniform random variable over (0, 1).

10.   Let X be uniformly distributed over (0, 1). Find the pdf of

      (a) $Y = e^X$ (b) $Z = X^2 + 1$

      (c) $W = \dfrac{1}{X + 1}$ (d) $Y = X^{1/\beta}$, where $\beta \neq 0$.

11.   Let X be exponentially distributed with parameter λ. Find the pdf of $Y = \log_e X$.

12.   Let X be have gamma density with parameters (α, λ). Find the pdf of

      (i) $Y = \sqrt{X}$ (ii) $Z = cX$, where $C > 0$.

13.   An angle θ is chosen at random from $\left(\dfrac{-\pi}{2}, \dfrac{\pi}{2}\right)$. Let $X = \sin \theta$. Find the pdf of X.

14.   Let X be exponentially distributed with parameter λ.

      Let $Y = n$ if $n \leq X < n + 1$ where n is a nonnegative integer. Find the pdf of Y.

15.   Let X and Y be independent random variable each having uniform pdf on (0, 1,…N).

      Find the pdf of (i) min (x, Y) (ii) Max (X, Y).

7.2   Functions of Two Random Variables

The method discussed in section 7.1 of finding the pdf of a function of one random variable of the continuous type will now be extended to functions of two random variables

Let X and Y be continuous random variables and $Z = X + Y$.

The problem of finding the pdf of Z is somewhat more involved. The following steps are useful.

1.   Introduce a second random variable $W = K(X, Y)$.

2.   Obtain the joint pdf of Z and W, say d(z, w).

3.   Obtain the desired pdf of Z by simply integrating d(z, w) with respect to w. that is

$$\int d(z, w)\, dw.$$

The following will be assumed in obtaining the jointing pdf of Z and w.

(i)   The functions h(X, Y) and k(X, Y) satisfy the following conditions.

    (a)   The equations

$$Z = h(x, y) \text{ and } w = k(x, y)$$

    may be uniquely solved for x and y in termsof z and w, ssay $x = g_1(z, w)$,

$$y = g_2(x, w)$$

    (b)   The partial derivatives $\frac{\partial x}{\partial z}$, $\frac{\partial x}{\partial w}$, $\frac{\partial y}{\partial z}$, $\frac{\partial y}{\partial w}$ exist and are continuous

From the result of section 7.1 it can be proved using results of advanced calculus that under the assumptions (i) and (ii), the joint pdf of Z and W is given by d(z, w) = f(x, y) |J| where x = g₁(z, w), y = g₂(z, w) and

$$J = \begin{vmatrix} \dfrac{\partial x}{\partial z} & \dfrac{\partial x}{\partial w} \\ \dfrac{\partial y}{\partial z} & \dfrac{\partial y}{\partial w} \end{vmatrix}.$$

The determinant J, is called the Jacobian of the transformation (x, y)-→ (z, w). some of the important random variables we shall be interested in are

$$X + Y,\ XY,\ X/Y,\ \min.\ (X, Y),\ \max.\ (X, Y), \text{ etc.}$$

Distribution of Sum

Let Z = X + Y and W = X, the Jacobian,

$$J = \begin{vmatrix} 0 & 1 \\ 1 & -1 \end{vmatrix} = -1$$

Thus d(z, w) = f(x, y) |J| = f(w, (z − w)),

$$F_Z(z) = \int_{-\infty}^{\infty}(w, z - w)\, dw = \int_{-\infty}^{\infty} f(x, z - x)\, dx$$

If X and Y are nonnegative independent random variables, then

$$F_Z(z) = \int_{0}^{z} f_X(x) f_Y(z - x)\, dx$$

Example 17

Let X and Y have the joint pdf $f(x, y) = \begin{cases} 1 & 0 < x < 1 \qquad 0 < y < 1 \\ 0 & \text{elsewhere} \end{cases}$

Find the pdf of $Z = X + Y$.

Solution

Let $W = X$; $Z = X + Y$, then $x = w$; $y = z - w$ and

$$J = \begin{vmatrix} 0 & 1 \\ 1 & -1 \end{vmatrix} = -1$$

Hence, $d(z, w) = 1.1 = 1$.

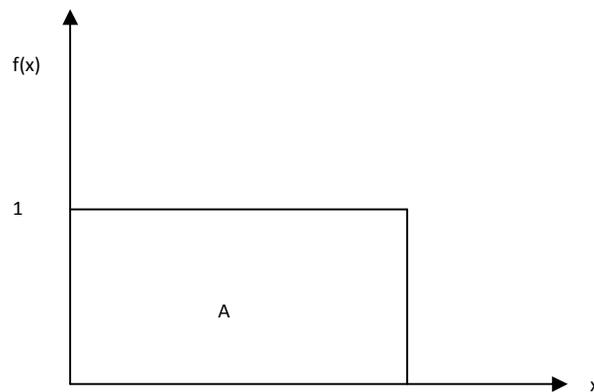Let A be the set $A = \{(x, y), 0 \leq x \leq 1, 0 \leq 1\}$. A is shown in figure below.

Figure 2

Under the transformation, $Z = X + Y$, $W = X$, the boundaries of A are $x = 0$, $x = 1$, $y = 0$ and $y = 1$. The set B in the z, w –plane which is the mapping of A under the one-to-one transformation is determined as follows:

$x = 0$ implies that $w = 0$; $y = 1$ implies that $w = 1$

$y = 0$ implies that $z = w$; $y = 1$ implies that $z = w + 1$.

Thus, the boundaries of B are the lines w = 0, w = 1, z = w, z = w + 1. This is show in the figure below.
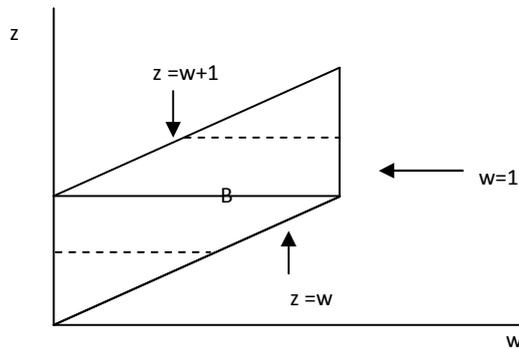
Figure 3

Hence

$$d(z, w) = 1$$

thus,   (z, w) ∈ B

$$d(z) = \int_0^z 1 \; dw \; = z \qquad 0 < z < 1$$

$$\int_{z-1} 1 \; dw = 1 - z \qquad 1 < z < 2$$

$$d(z)= \begin{cases} Z & 0 < z \le 1 \\ 2\text{-}z & 1 < z < 2 \\ 0 & \text{elsewhere} \end{cases}$$

Example 18

Suppose a point (x, y)is selected at random in a circle centre (0, 0) and radius 1 and that this point is uniformly distributed over the circle. Find the pdf of

$R = \sqrt{(X^2 + Y^2)}$, the distance from the origin.

The joint pdf of X and Y is

$$d(z)= \begin{cases} \dfrac{1}{\pi} & \text{if (x, y) lies inside or on the circle} \\ \\ 0 & \text{elsewhere} \end{cases}$$

172

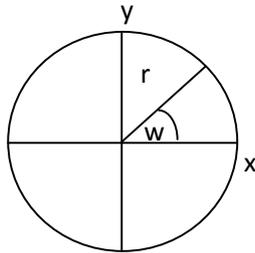Let W = $\tan^{-1}$(Y/X) then tan W = Y/X, x = r cos w, y = r sin w.
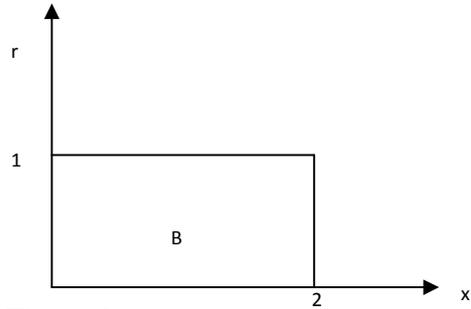
Figure 4

The Jacobian is

$$J = \begin{vmatrix} \dfrac{\partial x}{\partial r} & \dfrac{\partial x}{\partial w} \\ \dfrac{\partial y}{\partial x} & \dfrac{\partial y}{\partial w} \end{vmatrix} \begin{vmatrix} \cos w & -r\sin w \\ \sin w & r\cos w \end{vmatrix} = r(cos^2w + sin^2w) = r.$$

hence

$$d(r, w) = f(x, y) \, |J| = \begin{cases} \dfrac{1}{\pi} r & (r, w) \in B \\ 0 & \text{elsewhere} \end{cases}$$

Thus,

$$d(r) = \int d(r, w) \, dw = r \int_0^{2\pi} \frac{1}{\pi} dw = \begin{cases} 2r & 0 \le r \le 1 \\ 0 & elsewhere \end{cases}$$

Example 19

$$\text{If } f(x, y) = \begin{cases} 3x & 0 < y < x \\ & 0 < x < 1. \text{ Find the pdf of } Z = X - Y. \\ 0 & \text{elsewhere} \end{cases}$$

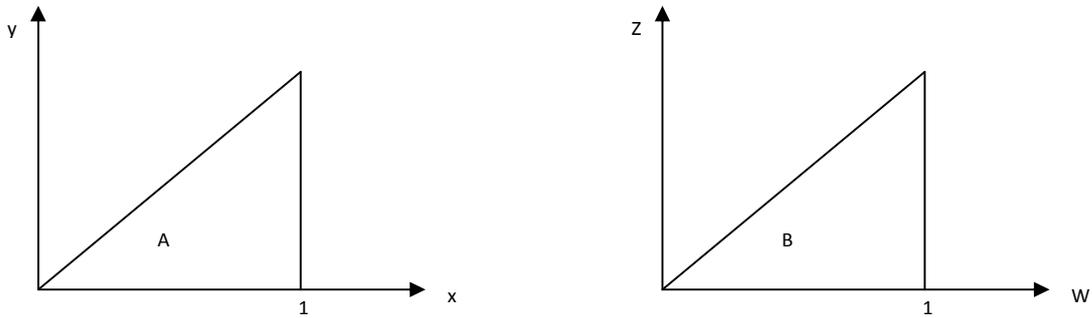Let W = X, A = {(x, y), 0< y < x, 0 < x < 1} and B ={(z, w) : z=x − y, w = x}

173

Figure 5

The Jacobian is

$$J = \begin{vmatrix} 0 & 1 \\ -1 & 1 \end{vmatrix}$$

hence

$$d(z, w) = \begin{cases} 3w & 0 < z < w \\ & 0 < w < 1 \\ 0 & \text{elsewhere} \end{cases}$$

$$d(z) = = = \int_z^1 3\,w\,dw$$

$$d(z) = \begin{cases} \dfrac{3}{2}(1 - z^2) & 0 < z < 1 \\ 0 & \text{elsewhere} \end{cases}$$

Distribution of Product

Suppose X and Y are continuous random variables having the joint pdf (x, y). show that the pdf of Z = XY is given by

$$\int \left|\frac{1}{x}\right| f\left(x, \frac{z}{x}\right) dx$$

Solution

Let W = X, thus, x = w, $y = \dfrac{z}{w}$. The Jacobian

174

$$J = \begin{vmatrix} 1 & 0 \\ \frac{-z}{w^2} & \frac{1}{w} \end{vmatrix} = \frac{1}{w}$$

Hence,

$$d(z,w) = f\left(w, \frac{z}{w}\right)|J|$$

Thus,

$$d(z) = \int \left|\frac{1}{w}\right| f\left(w, \frac{z}{w}\right) dw \qquad (2)$$

Example 20

Let X and Y be independent random variables with the following pdfs.

$$f(x) = \begin{cases} 2x & 0 \le x \le 1 \\ 0 & elsewhere \end{cases}, \quad g(y) = \begin{cases} 4y^3/15 & 1 \le y \le 2 \\ 0 & elsewhere \end{cases}$$

Find the pdf of Z = XY.

Solution

The joint pdf of X and Y is f(x. y) = f(x) g(y).

Thus, from 2.

$$d(z) = \int \left|\frac{1}{w}\right| g(w) g\left(\frac{z}{w}\right) dw, \text{ where } W = X$$

$$d(z) = \int \left|\frac{1}{w}\right| \cdot 2w \frac{4}{15} \left(\frac{z}{w}\right)^3 dw = \frac{8}{15} z^3 \int \frac{1}{w^3} dw.$$

Let A = {(x. y): 0< x < 1, 1 ≤ y ≤ 2}. The transformation, Z = XY, W = X gives the set B determined as follows:

When x = 0, w = 0; when x = 1, w = 1,

When y = 1, Z = w, when y = 2, z = 2w.

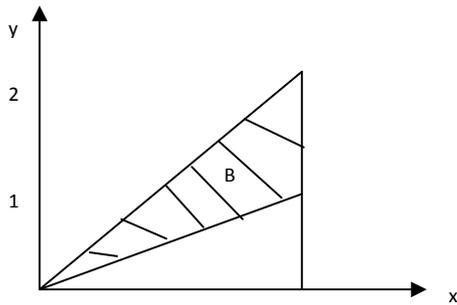The lines w = 0; w = 1; x = w, z = 2w are the boundaries of B as shown in the figure below.

Figure 6

Thus

$$d(z) = \frac{8z^3}{15} \int_{z/2}^{z} \frac{1}{w^3} dw = \frac{4}{5} z, \quad 0 < z < 1$$

$$= \frac{8z^3}{15} \int_{z/2}^{1} \frac{1}{w^3} dw = \frac{4z}{15} (4 - z^2), \quad 1 < z < 2$$

$$d(z) = \begin{cases} \frac{4z}{5} & 0 < z < 1 \\ \frac{4z}{15}(4 - z^2) & 1 < z < 2 \end{cases}$$

Distribution of Quotients

Suppose X and Y are continuous random variables having the joint pdf f(x, y). show that the pdf of Z = Y/X is given by

$$f_z(z) = \int_{-\infty}^{\infty} |x| f(x, xz) dx - \infty < z < \infty$$

Let W = X, then x = w, y = zw, the Jacobian is

$$J = \begin{vmatrix} 1 & 0 \\ z & w \end{vmatrix} = w.$$

Hence

$$d(z, w) = f(w, zw)|J| = f(w, zw)|w|$$

Thus

$$f_z(z) = \int |w| f(w, zw) dw. \tag{3}$$

Example 21

Let X and Y be independent random variables having the respective gamma densities $\Gamma(\alpha_1, \lambda)$ and $\Gamma(\alpha_2, \lambda)$. Find the probability density function of $Z = Y/X$.

$$f_X(x) = \frac{\lambda^{\alpha_1} x^{\alpha_1-1} e^{-\lambda x}}{\Gamma(\alpha_1)} \qquad\qquad x > 0$$

$$f_Y(y) = \frac{\lambda^{\alpha_2} x^{\alpha_2-1} e^{-\lambda x}}{\Gamma(\alpha_2)} \qquad\qquad y > 0$$

The joint pdf of X and Y is given by

$$f(x,\ y) = f_X(x) f_Y(y) = \frac{\lambda^{\alpha_1+\alpha_2} x^{\alpha_1-1} e^{-\lambda(x+y)}}{\Gamma(\alpha_1)\Gamma(\alpha_2)}.$$

From (3) we have

$$f_Z(z) = \int w f(x, zw) dw = \int w \frac{\lambda^{\alpha_1+\alpha_2} w^{\alpha_1-1} (zw)^{\alpha_2-1}}{\Gamma(\alpha_1)\,\Gamma(\alpha_2)} e^{\lambda(w+wz)} dw$$

$$= \frac{\lambda^{\alpha_1+\alpha_2} z^{\alpha_2-1}}{\Gamma(\alpha_1)\,\Gamma(\alpha_2)} \int_0^\infty x^{\alpha_1+\alpha_2-1} e^{-\lambda x(1+z)}\, dw$$

But

$$\int_0^\infty w^{\alpha_1+\alpha_2-1} e^{-\lambda w(1+z)} dw = \frac{\Gamma(\alpha_1+\alpha_2)}{(\lambda(1+z))^{\alpha_1+\alpha_2}}$$

Hence

$$f_Z(z) = \frac{\Gamma(\alpha_1+\alpha_2)}{\Gamma(\alpha_1)\,\Gamma(\alpha_2)} \frac{z^{\alpha_2-1}}{(1+z)^{\alpha_1+\alpha_2}}, \qquad\qquad z > 0$$

Example 22

Let X and Y be independent exponential random variables with parameter λ. Find the joint pdf of $Z = X + Y$ and $W = \dfrac{X}{X + Y}$.

Solution

The joint pdf of X and Y is given by
$$f_{X,Y}(x, y) = \lambda^2 e^{-\lambda(x+y)} \quad x > 0, \quad y > 0$$

The transformation gives
$$x = wz,\ y = z(1 - w)$$

The Jacobian of the transformation is
$$|J| = z$$

The joint pdf of Z and W is given by
$$f_{Z,W}(z, w) = \begin{cases} z\lambda^2 e^{-\lambda z}, & z > 0, 0 < w < 1 \\ 0 & , \qquad elsewhere \end{cases}$$

Example 23

Let X and Y have the joint pdf $f_{X,Y}(x,y) = \frac{1}{x^2 y^2}$ $x \geq 1, y \geq 1$, find the joint pdf of $Z = XY$ and $W = \frac{X}{Y}$.

Solution

From the transformation $z = wy$, $w = x/y$, we have

$$J = \begin{vmatrix} y & x \\ \frac{1}{y} & \frac{-x}{y^2} \end{vmatrix} = \frac{-2x}{y}$$

$$|J| = 2w.$$

Thus

$$f_{Z,w}(z,w) = f_{Z,w}(x,y)|J|^{-1} = \frac{1}{z^2} \cdot \frac{1}{2w}$$

$$= \frac{1}{2wz^2} \ z \geq 1, \ 0 < w < z$$

$$= 0, \text{ otherwise.}$$

Example 24

Let X_1 and X_2 be random variable with joint pdf $f_{x_1, x_2}(x_1, x_2)$. Let $Y_1 = X_1 + X_2$, $Y_2 = X_1 - X_2$. Find the joint pdf of Y_1 and Y_2 in terms of $f_{x_1, x_2}(x_1, x_2)$.

Solution

$$f_{Y_1, Y_2}(y_1, y_2) = f_{x_1, x_2}(x_1, x_2)|J|$$

$$x_1 = \frac{1}{2}(y_1 + y_2), x_2 = \frac{1}{2}(y_1 - y_2)$$

$$J = \begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} \\ \frac{\partial y_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} \end{vmatrix} = \begin{vmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{vmatrix} = -\frac{1}{2}.$$

Hence

$$f_{Y_1, Y_2}(y_1, y_2) = \frac{1}{2} f_{x_1, x_2}\left(\frac{y_1 + y_2}{2}, \frac{y_1 - y_2}{2}\right)$$

Example 25

Let X and Y be independent exponential random variables with parameters λ_1 and λ_2 respectively.

(i)   find the joint pdf of $U = X + Y$ and $V = \frac{X}{X + Y}$

(ii)   show that U and V are independent if $\lambda_1 = \lambda_2$.

Solution

The joint pdf of X and Y is given by

$$f_{X,Y}(x, y) = \lambda_1 \lambda_2 e^{-(\lambda_1 x + \lambda_2 y)}$$

$$x = uv \text{ and } y = u - uv.$$

The joint pdf of u and v is given by

$$f_{u,v}(u,v) = f_{X,Y}(x,y) |J|.$$

Substituting for x and y in terms of u and v

$$f_{u,v}(u,v) = \lambda_1 \lambda_2 e^{-(\lambda_1 uv + \lambda_2 u(1-v))}|J| = u\lambda_1 \lambda_2 e^{-uv(\lambda_1 - \lambda_2) - u\lambda_2}$$

(ii)   Putting $\lambda_1 = \lambda_2 = \lambda$. We obtain

$$f_{u,v}(u,v) = u\lambda^2 e^{-\lambda u}, u \geq 0, 0 < v < 1$$

$$f_u(u) = \int_0^1 u\lambda^2 e^{-\lambda u} dv = u\lambda^2 e^{\lambda u}, u \geq 0$$
$$f_v(v) = \int_0^\infty u\lambda^2 e^{-\lambda u} du = 1 \quad 0 < v < 1.$$

Example 26

The joint pdf of X and Y is given by
$$f(x,y) = \begin{cases} e^{-(x+y)} & X > 0, \ y > 0 \\ 0 & otherwise \end{cases}$$
find (i) P(X > 2, Y > 3), (ii) P(Y < X) and (iii) the pdf of Z = X/Y.

Solution

Let A = {x, y)! x > 2, y > 3), we have

$$P(X > 2, Y > 3) = \iint_A f(x,y)dx\,dy = \int_3^\infty \int_2^\infty f(x,y)dx\,dy$$
$$= \int_3^\infty \left[\int_2^\infty e^{-(x+y)}dx\right]dy = \int_3^\infty e^{-y}\left[\int_2^\infty e^{-x}dx\right]dy = e^{-2}\int_3^\infty e^{-y}\,dy = e^{-5}$$

(ii)   Let A {(x, y); y < x}

$$P(Y < X) = \iint_A f(x,y)dxdy = \int_0^\infty \left[\int_y^\infty e^{-(x+y)}dx\right]dy = \int_0^\infty e^{-2y}\,dy = \frac{1}{2}.$$

Example27

Let f(x, y) = $\begin{cases} \frac{1}{\pi r^2} x^2 & + \quad y^2 \ \leq \ r^2 \\ 0 & elsewhere \end{cases}$

find (i)$f_X(x)$ and (ii) $f_Z(z)$ where $Z = \sqrt{X^2 + Y^2}$.

Solution

(i) $f_x(x) = \int_{-\infty}^\infty f(x,y)dy = \int \frac{1}{\pi r^2}dy = \frac{y}{\pi r^2}\bigg|_{\sqrt{r^2-x^2}}^{\sqrt{r^2+x^2}} = \frac{2}{\pi x^2}\sqrt{r^2 - x^2}$

(ii) $f_Z(z) = P\{X^2 + Y^2 \leq z^2) = \iint_{x^2+y^2\leq z^2} f(x,y)dxdy = \iint_{x^2+y^2\leq z^2} \frac{1}{\pi r} dxdy$

      $= \frac{\pi z^2}{\pi r^2} = \frac{z^2}{r^2}$ (since $\iint_{x^2+y^2\leq z^2} dxdy$ = area of a circle radius z.)

(iii) $F_z = P(z \leq z) = P\left(\frac{X}{Y} \leq x\right) = \iint_{\frac{X}{Y}\leq Z} f(x,y)dx\,dy$

      $= \int_0^\infty \int_0^{yz} e^{-(x+y)}dxdy = 1 - \frac{1}{z+1}$

Differentiating we obtain
$$f_z(z) = \frac{1}{(z+1)^2}, 0 < z < \infty.$$

Example 28

Distribution of sum. Let X and Y have the joint pdf f(x, y) and let Z = X + Y then
$$F_Z(z) = P(X + Y \leq z) = \iint_{x+y<z} f(x,y)dx\,dy = \int_{-\infty}^\infty \int_{-\infty}^{z-y} f(x,y)dx\,dy.$$
Changing variable, x = v − y
$$F_z(z) = \int_{-\infty}^\infty \int_{-\infty}^{z-y} f(v-y,y)dv\,dy = \int_{-\infty}^z \left[\int_{-\infty}^z f(v-y,y)dy\right]dv.$$
Thus
$$F_Z(z) = \int_{-\infty}^\infty f(z-y,y)dy.$$

Example 29: *Student's t-Distribution*

Let X be a standard normal distribution and Y a chi-square random variable with v degrees of freedom. Find the pdf of $Z = \frac{X\sqrt{v}}{\sqrt{Y}}$ where X and Y are independent.

Solution

Let $Z = \frac{X\sqrt{v}}{\sqrt{Y}}$ and $W = Y$ then $y = w$ and $x = z\frac{\sqrt{w}}{\sqrt{v}}$.

Then Jacobian of the transformation is

$$J = \begin{vmatrix} 1 & 0 \\ -\frac{1}{2}z\frac{\sqrt{w}}{v^{-\frac{3}{2}}} & \frac{\sqrt{w}}{\sqrt{v}} \end{vmatrix} = \frac{\sqrt{y}}{\sqrt{w}}.$$

Thus

$$d_{Z,w}(z,w) = f_x(x)f_Y(y)\left|\frac{\sqrt{w}}{v}\right| = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}\frac{1}{\Gamma\left(\frac{V}{2}\right)2^{v/2}}y^{\frac{v}{2}-1}e^{-\frac{y}{2}}\frac{\sqrt{W}}{V}$$

substituting for x and y, we have

$$d_{Z,w}(z,w) = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}\frac{z^2w}{v}} \cdot \frac{1}{\Gamma\left(\frac{V}{2}\right)2^{\frac{n}{2}}}w^{\frac{v}{2}-1}e^{-\frac{w}{2}}\cdot\frac{\sqrt{W}}{\sqrt{V}} = \frac{w^{\frac{v-1}{2}}}{\sqrt{2\pi}\Gamma\left(\frac{V}{2}\right)2^{\frac{v}{2}}}e^{-\frac{1}{2}\frac{z^2w}{v}-\frac{w}{2}}.$$

The marginal pdf of z is given by

$$d_z(z) = \int \frac{w^{\frac{v-1}{2}}}{\sqrt{2\pi}\Gamma\left(\frac{V}{2}\right)2^{\frac{n}{2}}}e^{-\frac{1}{2}\frac{z^2w}{v}-\frac{w}{2}}\cdot\frac{\sqrt{W}}{\sqrt{V}}\,dw = \frac{1}{\sqrt{2\pi}\Gamma\left(\frac{V}{2}\right)2^{\frac{v}{2}}}\int_0^\infty w^{\frac{v-1}{2}}e^{-\frac{w}{2}\left(1+\frac{z^2}{v}\right)}\frac{\sqrt{W}}{\sqrt{V}}\,dw$$

Let $u = \frac{w}{2}\left(1+\frac{z^2}{v}\right)$; $\frac{du}{dw} = \frac{1}{2}\left(1+\frac{z^2}{v}\right)$, $w = \frac{2u}{1+\frac{z^2}{V}}$

$$d_Z(z) = \frac{2^{\frac{v-1}{2}}}{\sqrt{2\pi v}\Gamma\left(\frac{V}{2}\right)2^{\frac{v}{2}-1}}\int_0^\infty \left[\frac{2u}{1+\frac{z^2}{V}}\right]e^{-u}\frac{du}{\frac{1}{2}\left(1+\frac{z^2}{V}\right)}$$

$$\frac{1}{\sqrt{2\pi}\Gamma\left(\frac{V}{2}\right)2^{\frac{v}{2}}\left(1+\frac{z^2}{V}\right)^{\frac{v-1}{2}}}\int u^{\frac{v-1}{2}-1}e^{-u}\,du.$$

Note

$$\int_0^\infty u^{\frac{v}{2}-1}e^{-u}\,du = \Gamma\left(\frac{v+1}{2}\right).$$

Hence

$$d_Z(z) = \frac{\Gamma\left(\frac{v+1}{2}\right)}{\sqrt{2\pi v}\Gamma\left(\frac{V}{2}\right)}\cdot\frac{1}{\left(1+\frac{z^2}{V}\right)^{\frac{v+1}{2}}}, \quad -\infty < z < \infty.$$

The above distribution, $d_Z(z)$ is called a t-distribution.

Note
(i) That a t-distribution is completely determined by the parameter v, the number of degrees of freedom of the random variable that has the chi-square distribution.
(ii) The t-distribution is a bell-shaped distribution, symmetric about zero.
(iii) The t-distribution is flatter than the Normal distribution. As the sample size becomes large, t-distribution approaches the standard normal distribution ($n > 30$ is sufficient).
(iv) The values of the t-distribution for various values of degrees of freedom have been tabulated.

Example 30
Exercise 7.2

1. Let X_1 and X_2 be independent normal random variables $N(\mu_1, \sigma^2_1)$, $N(\mu_2, \sigma^2_2)$ respectively. Find the pdf of $Y = X_1 + X_2$.
2. Let X and Y be independent random variables having the respective pdfs.
    $f(x) = e^{-x}$, $x \geq 0$, $g(y) = 2e^{-y}$ $y \geq 0$. Find the pdf of $Z = \dfrac{X}{Y}$
3. Let X and Y have the joint pdf $f(x, y) = \lambda^2 e^{-\lambda(x+y)}$, $x > 0$, $y > 0$. Find the pdf of $Z = X + Y$.
4. Let X and Y be independent random variables each having the normal probability density function $N(0, \sigma^2)$. Find the probability density function of
    (i) $\dfrac{Y}{X}$ (ii) $\dfrac{Y^2}{X^2}$ (iii) $\dfrac{Y}{|X|}$ (iv) $\dfrac{|Y|}{|X|}$
5. Suppose X and Y are independent normal random variables with mean 0 and variance σ^2. Find the pdf of (i) $X + Y$ (ii) $X^2 + Y^2$.
6. If X and Y are independent exponential random variables with parameter λ. Find the joint pdf of $Z = X + Y$ and $W = \dfrac{X}{X+Y}$. Hence find the pdf of Z and of W.
7. Let X and Y have uniform distribution over the interval (a, b). Find the pdf of $Z = |Y - X|$.
8. Let X be uniformly distributed over (0, 1) and let Y be uniformly distributed over (0, x). Find
    (i)   The conditional pdf of Y given X = x.
    (ii)   The joint pdf of X and Y
    (iii)   The marginal pdf of Y.
9. Let $f(x, y) = c(y - x)^2$, $0 \leq x < y \leq 1$
    $= 0$,   elsewhere
    Find the pdf of $Z = X + Y$.

7.3   Sum of Independent Identically Distributed RandomVariables
The Moment Generating Function Technique
Let X_1, X_2,....X_n be independent identically distributed random variables. In this section we will find the pdf of sum of the X's. The moment generating function technique will be used to establish the pdf of sum Bernoulli, Poisson, exponential, Normal, Gamma random

variables. All the probability distributions above except Bernoulli and exponential has the following remarkable and very useful property called Reproductive properties.

Definition: Reproductive Property
A probability distribution F is said to have reproductive property if when two or more independent random variables having the probability distribution F are added, the resulting random variable has the distribution F. we shall establish this result using moment generating function defined in Chapter 4, section 4.3.

Moment Generating Function
The moment generating function $M_X(t)$ of a random variable X is defined by
$$M_X(t) = E(e^{tX})$$
The domain of $M_x(t)$ is all numbers t such that e^{tX} has finite expectation.

Theorem 3
Suppose that a random variable X has mgf $M_x(t)$. Let $Y = aX + b$. Then, the mgf of the random variable Y is given by $M_Y(t) = e^{bt} M_x(at)$.
Proof:
$$M_Y(t) = E(e^{ty}) = E(e^{(aX + b)t}) = E(e^{atX + bt}) = E(e^{atX} e^{bt}) = e^{bt} E(e^{atX}) = e^{bt} M_X(at).$$

Theorem 4
Let X and Y be two random variables with mgf's $M_x(t)$ and $M_Y(t)$, respectively. If $M_x(t) = M_Y(t)$ for all values of t, then X and Y have the probability distribution. That is, the mgf **uniquely** determines the probability distribution of the random variable. The proof of this theorem is outside the scope of this work, the reader is referred to volume II of this book.

This theorem will now be used to show that a linear function of a random variable also has a normal distribution.

Example 31
Suppose that X has distribution $N(\mu, \sigma^2)$.
Let $Y = aX + b$. show that Y also has a Normal distribution with mean $a\mu + b$ and variance $a^2 \sigma^2$. The mgf of Y is
$$M_Y(t) = E(e^{ty}) = e^{bt}(M_X(at))$$
from theorem I.
$$M_X(t) = E(e^{tx})$$
$$\int_{-\infty}^{\infty} e^{tx} = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1/2(x-\mu)^2}{\sigma^2}} \, dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{tx} e^{-\frac{1/2(x-\mu)^2}{\sigma^2}} \, dx$$
Let $s = \frac{x-\mu}{\sigma}$, then $M_X(t)$ becomes
$$M_X(t) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{t(\sigma s + \mu)} e^{-1/2 s^2} \sigma ds = \frac{e^{t\mu}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-1/2 s^2 + t\sigma s} \, ds$$

$$\frac{e^{t\mu + \sigma^2 \frac{t^2}{2}}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-1/2 S^2 + t\sigma s} \, ds.$$

Making the change of variable u = s − σt, we see that

$$M_x(t) = e^{t\mu + \frac{1}{2}\sigma^2 t^2} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-u^2} \, du = t\mu + e^{tu + \frac{\sigma^2 t^2}{2}}$$

Substituting at for t, we have

$$M_X(at) = e^{at\mu + \sigma^2 a^2 t^2/2}$$

Hence,

$$M_Y(t) = e^{bt} \left(e^{at\mu} + \frac{\sigma^2 a^2 t^2}{2} \right) = e^{t(b+a\mu) + \frac{\sigma^2 a^2 t^2}{2}} = e^{t\gamma + \sigma^2 t^2 a^2/2}$$

Where $\gamma = b + \mu a = E(Y)$, Vara = E(Y), Var(Y) = $a^2\sigma^2$.
This is the mgf of a normal variable with mean $\gamma = a\mu + b$ and variance $a^2\sigma^2$. Thus, by theorem 2, Y is a normal random variable with mean $a\mu + b$ and variance $a^2\sigma^2$.

Example 32
Let X be N(μ, σ^2). Show that $Y = \frac{X-\mu}{\sigma}$ is N(0, 1).

Solution

$$Y = \frac{X-\mu}{\sigma} = \frac{X}{\sigma} - \frac{\mu}{\sigma}.$$

Let $a = \frac{1}{\sigma}$ and $b = -\frac{\mu}{\sigma}$, then Y = aX + b.
From Example 312, Y is a normal random variable with mean $a\mu + b$ and variance $a^2\sigma^2$.
Substituting for a and b, we see that

(i) $a\mu + b = \frac{\mu}{\sigma} - \frac{\mu}{\sigma} = 0$

(ii) $a^2\sigma^2 = \frac{\sigma^2}{\sigma^2} = 1$

Thus Y is standard normal (N(0, 1)) random variable.

Theorem 5
Let X and Y be independent random variables and Z = X + Y.
Let $M_X(t)$, $M_Y(t)$ and $M_Z(t)$ be the mgfs of the random variables X, Y and Z respectively.
Then

$$M_Z(t) = M_X(t) \, M_Y(t) \qquad\qquad (4)$$

Proof:
By definition,

$$M_Z(t) = E(e^{Zt}) = E(e^{(X + Y)t}) = E(e^{Xt} + e^{Yt}) = E^{(e+Xt)} E(e^{Yt}) = M_X(t) \, M_Y(t).$$

In general, if X_1, X_2,…, X_n are independent random variables.
The mfg of Z = $X_1 + X_2 + …+ X_n$ is

$$M_Z(t) = M_{X_1}(t) M_{X_2}(t) \dots M_{X_n}(t)$$

183

If $X_1, X_2, \ldots, X_n$ are independent identically distributed with mgf $M_X(t)$, then
$$M_Z(t) = [M_X(t)]^n.$$

Example 33: Sum of Binomial Random Variables
Let X and Y be independent binomial random variables with parameters (n_1, p), (n_2, p) respectively. Find the pdf of Z = X + Y.
$$M_X(t) = E(e^{tX}) = \sum_{x=0}^{n_1} e^{tX} \binom{n_1}{x} p^X (1-p)^{n_1-x} = [pe^t + (1-p)]^{n_2}.$$
The mgf of Z is given by (from Theorem 5).
$$M_Z(t) = M_X(t)\, M_Y(t) = [pe^t + (1-p)]^{n_1+n_2}$$
but this is the mgf of a binomially distributed random variable with parameters $(n_1 + n_2, P)$.
thus, by theorem 2, Z is a binomial random variable and its pdf is given y
$$f_Z(z) = \begin{cases} {}^{n_1+n_2}C_z p^Z (1-p)^{n_1+n_2-z} & z = 0, 1, 2, \ldots n_1 + n_2 \\ 0 & elsewhere \end{cases}$$
In general, if $X_1, X_2, \ldots X_K$ are independent binomial random variables with parameters (n_1, p), (n_2, p), $\ldots (n_k, p)$ respectively, then the moment generating function of $Z = X_1$, $X_2, \ldots, X_k$ is given by
$$M_Z(t) = [pe^t + (1-p)]^{n_1+n_2+\cdots+n_k}.$$
Thus Z is a binomial variable with parameters (n, p) where $n = n_1 + n_2 \ldots n_k$.

Example 34: *Sum of normal random variables*
Suppose that X and Y are independent random variables with distribution $N(\mu_1, \sigma^2_1)$ and $N(\mu_2, \sigma^2_2)$ respectively. Show that Z = X + Y is a normal random variable with mean $\mu_1 + \mu_2$ and variance $\sigma^2_1 + \sigma^2_2$
$$M_X(t) = e^{\mu_1 t + 1/2\sigma^2_1 t^2}$$
$$M_Y(t) = e^{\mu_2 t + 1/2\sigma^2_2 t^2}$$

Hence
$$M_Z(t) = M_X(t)M_X(t) = e^{\mu_1 t + 1/2\sigma^2_1 t^2} e^{\mu_2 t + \frac{1}{2}\sigma^2_2 t^2}$$
$$= e^{t(\mu_1 + \mu_2) + 1/2t^2(\sigma^2_1 + \sigma^2_2)} = e^{t\mu + 1/2t^2\sigma^2}$$
where $\mu = \mu_1 + \mu_2$, $\sigma^2 = \sigma^2_1 + \sigma^2_2$. Since $e^{t\mu + 1/2t^2\sigma^2}$ is the mgf of $N(\mu, \sigma^2)$, hence by Theorem 4, z is a normal random variable with mean $\mu = \mu_1 + \mu_2$ and variance $\sigma^2 = \sigma^2_1 + \sigma^2_2$.
In general, if $X_1, X_2, \ldots, X_n$ are independent $N(\mu_1, \sigma^2_1)$ then
$$Z = X_1 + X_2 + \ldots + X_k$$
is $N(\mu, \sigma^2)$ where $\mu = \mu_1 + \mu_2 + \ldots + \mu_n$, $\sigma^2 = \sigma^2_1 + \sigma^2_2 + \ldots + \sigma^2_k$

Example 35: *Sum of Poisson Random Variables.*
Suppose that X and Y are independent poisson random variable with mean λ_1 and λ_2 respectively. Show that Z = X + Y is a Poisson random with mean $\lambda_1 + \lambda_2$
$$M_X(t) = E(e^{tX}) = \sum_{x=0}^{\infty} e^{tX} \frac{e^{-\lambda_1}\lambda^x_1}{x!} = e^{-\lambda_1(1-e^t)}.$$
Similarly,
$$M_Y(t) = e^{-\lambda_1(1-e^t)}.$$

Hence,
$$M_Z(t) = e^{-\lambda_1(1-e^t)}e^{-\lambda_2(1-e^t)} = e^{-(1-e^t)(\lambda_1+\lambda_2)} = e^{-(1-e^t)\lambda}, \lambda = \lambda_1 + \lambda_2$$

The mgf of Poisson distribution with parameter λ is $e^{-\lambda(1-e^t)}$, thus Z is a poisson random variable with mean (parameter) $\lambda = \lambda_1 + \lambda_2$. In general, if $X_1, X_2,\ldots,X_n$ are independent $P(\lambda_1)$, then
$$Z = X_1 + X_2 +\ldots + X_n$$
is a Poisson random variable with parameter $\lambda = \lambda_1 + \lambda_2 +\ldots+ \lambda_n$.

Example 36
Let X and Y be independent gamma random variable with parameters (n_1, λ) and (n_2, λ) respectively. Show that $Z = X + Y$ is a gamma random variable withparameter $(n_1 + n_2, \lambda)$.

Solution
$$M_X(t) = E(e^{tX}) = \int_{-\infty}^{\infty} e^{tx} \frac{\lambda^{n_1}x_1^{n-1}e^{-\lambda x}}{\Gamma(n_1)}\, dx = \int_0^{\infty} \lambda^{n_1} \frac{(x)^{n_1-1}e^{-x(\lambda-t)}}{\Gamma(n_1)}\, dx$$

$$= \frac{\lambda^{n_1}}{\Gamma(n_1)}\int_0^{\infty} x^{n-1}e^{-x(\lambda-t)}dx = \frac{\lambda^{n_1}}{(\lambda-t)^{n_1-1}\Gamma(n_1)}\int_0^{\infty} u^{n_1-1}e^{-u}\, du$$

But $\int_0^{\infty} u^{n_1-1}e^{-u}\, du = \Gamma(n)$.

Hence,
$$M_X(t) = \left[\frac{\lambda}{\lambda-t}\right]^{n_1}.$$

Similarly $\quad M_Y(t) = \left[\frac{\lambda}{\lambda-t}\right]^{n_2}.$

Thus,
$$M_Z(t) = \left[\frac{\lambda}{\lambda-t}\right]^{n_1} \cdot \left[\frac{\lambda}{\lambda-t}\right]^{n_2} = \left[\frac{\lambda}{\lambda-t}\right]^{n_1+n_2} = \left[\frac{\lambda}{\lambda-t}\right]^{n} . n = n_1 + n_2.$$

This show that Z is a gamma random variable with parameters $(n_1 + n_2, \lambda)$.
In general, if $X_1, X_2 ,\ldots + X_k$ are independent gamma random variables with parameters (n_1, λ), $I = 1,2,\ldots k$. then $Z = X_1 + X_2 +\ldots + X_k$ has a gamma distribution with parameters (n, λ) where $n = n_1 + n_2 +\ldots+ n_k$.

Example 37
Suppose that X_i has $\chi^2 (n_1)$ (chi-square distribution with n_i degrees of freedom) $I = 1, 2,\ldots k$ where the X_i's are independent random variables.

Let $Z = X_1 + X_2 +\ldots + X_k$. Show that Z has distribution χ^2_n where $n = n_1 + n_2 +\ldots+ n_k$. Recall (5.9) that chi-square distribution is a special case of the gamma distribution in which $\lambda = 1/2$ and n replaced by n/2. Since the mgf of gamma (n, λ) is given by
$$\left[\frac{\lambda}{\lambda-t}\right]^{n} ;$$
substituting for $\lambda = 1/2$ and replacing n by n/2 we have
$$M_X(t) = \left[\frac{1/2}{1/2-t}\right]^{n/2} = \left[\frac{1}{1-2t}\right]^{n/2}.$$
The mgf of X_i is therefore given by

185

$$M_X(t) = \left[\frac{1}{1-2t}\right]^{n/2} I = 1, 2,\ldots$$

Hence, the mgf of Z is

$$M_Z(t) = M_{X_1}(t)M_{X_2}(t)\ldots M_{X_k}(t) = \left[\frac{1}{1-2t}\right]^{\frac{n_1+n_2+\cdots+n_k}{2}} = \left[\frac{1}{1-2t}\right]^{\frac{n}{2}}.$$

where $n = n_1 + n_2 + \ldots + n_k$ which is the mgf of χ^2_n distribution.

Example 38

Let $X_1, X_2,\ldots X_n$ be independent identically normal variables with mean μ and variance σ^2. Show that (i) $E(\bar{X}) = \mu$, where $\bar{X} = \frac{1}{n}(X_1 + X_2 + \ldots + X_n)$ (ii) $Var(\bar{X}) = \frac{\sigma^2}{n}$ and

(iii) $Y = \frac{\sum_1^n (X_1 - \mu)^2}{\sigma^2}$ has a $\chi^2(n)$.

(i) $E(\bar{X}) = E\left(\frac{1}{n}\sum_1^n X_i\right) = \frac{1}{n}\sum_1^n E(X_i) = \frac{1}{n}.n\mu = \mu$

(ii) $Var(\bar{X}) = Var\left(\frac{1}{n}\sum_1^n X_i\right) = \frac{1}{n^2}\sum Var(X_i) = \frac{1}{n^2}n\sigma^2 = \frac{\sigma^2}{n}$

(iii) $Y = \frac{\sum_1^n (X_1-\mu)^2}{\sigma^2} = \frac{(X_1-\mu)^2}{\sigma^2} + \frac{(X_2-\mu)^2}{\sigma^2} + \cdots + \frac{(X_n-\mu)^2}{\sigma^2}$

$= \left(\frac{X_1-\mu}{\sigma}\right)^2 + \left(\frac{X_2-\mu}{\sigma}\right)^2 + \cdots + \left(\frac{X_n-\mu}{\sigma}\right)^2 = Y_1^2 + Y_2^2 + \ldots + Y_n^2.$

From Example 14, we know that Y_1^2 has $\chi^2(1)$ and from Example 30, since Y_1^2 are independent $\chi^2(1)$ we have that Y is $\chi^2(n)$.

Example 39

Let $X_1, X_2,\ldots X_n$ be independent identical normal random variables with mean μ and variance σ^2

Show that $S^2 = \frac{\sum_{i=1}^n (X_1-\bar{X})^2}{\sigma^2}$ has a $\chi^2(n-1)$.

We shall proceed as follows:

$$X_1 - \mu = X_i - \bar{X} + \bar{X} - \mu$$
$$(X_1 - \mu)^2 = (X_1 - \bar{X})^2 + (\bar{X} - \mu)^2 + 2(X_i - \bar{X})(\bar{X} - \mu)$$

Summing, we have

$$\sum_{i=1}^n (X_i - \mu)^2 = \sum_{i=1}^n (X_i - \bar{X})^2 + \sum_{i=1}^n (\bar{X} - \mu)^2 + 2\sum_{i=1}^n (X_i - \bar{X})(\bar{X} - \mu)$$

$$\sum_{i=1}^n (X_i - \bar{X})(\bar{X} - \mu) = (\bar{X} - \mu)\sum_{i=1}^n (X_i - \bar{X})\backslash$$

since $\sum(X_i - \bar{X}) = 0$, we see that

$$\sum_{i=1}^n (X_i - \mu)^2 = \sum_{i=1}^n (X_i - \bar{X})^2 + \sum_{i=1}^n (\bar{X} - \mu)^2$$

dividing through by σ^2, we have

$$\frac{\sum_{i=1}^n (X_i - \mu)^2}{\sigma^2} = \frac{\sum(X_1-\bar{X})}{\sigma^2} + \frac{\sum_{i=1}^n (\bar{X} - \mu)^2}{\sigma^2}$$

From Example 37, we have $\sum_{i=1}^{n} \frac{(X_i - \mu)^2}{\sigma^2}$ is $\chi^2(n)$ and (it is left as an exercise to show that $\frac{\Sigma(X_i - \mu)^2}{\sigma^2}$ is $\chi^2(1)$.

Let $Z = \frac{\Sigma(X_i - \mu)^2}{\sigma^2}$ and $Y = \frac{\Sigma(X_i - \mu)^2}{\sigma^2}$

$$S^2 = \frac{\Sigma(X_i - \bar{X})^2}{\sigma^2}$$

Since $\bar{X}$ and S^2 are stochastically independent, we have

$$E(e^{tZ}) = E(e^{t(Y + S2)} = E(e^{tY})E(e^{ts^2})$$

but $M_Z(t) = E(e^{tZ}) = \left[\frac{1}{1-2t}\right]^{n/2}$, since Z is $\chi(n)$,

$$M_Y(t) = E(e^{tY}) = \left[\frac{1}{1-2t}\right]^{t/2} \text{ (Y is } Z^2_{(1)})$$

hence,

$$E(e^{tS})^2 = \left[\frac{1}{1-2t}\right]^{\frac{n-1}{2}}$$

and this is the mgf of a chi-square distribution with $n - 1$ degrees of freedom. This shows that S^2 has a chi-square distribution with $n - 1$ degrees of freedom.

Exercise 7.3

1. Let $X_1, X_2,\ldots X_n$ be independent identically distributed random variables, each of which is $N(\mu, \sigma^2)$. Show that $\bar{X} = \frac{X_1 + X_2 + \cdots + X_n}{n}$ is $N\left(\mu, \frac{\sigma^2}{n}\right)$.

2. Let $X_1, X_2,\ldots X_n$ be independent identically distributed exponential random variables with parameters λ. Show that $Z = X_1 + X_2 + \ldots + X_n$ has a gamma distribution with parameters n and λ.

3. Let $X_1, X_2,\ldots X_n$ be independent $N(\mu_1, \sigma^2_1)$. Show that $Y = a_1 X_1 + a_2X_2 + \ldots + a_nX_n$ is a normal random variable with mean $\mu = a_1\mu_1 + a_2\mu_2 + \ldots + a_n\mu_n$ and variance $\sigma^2 = a^2_1\sigma^2_1 + a_2\sigma^2_2 + \ldots + a^2_n \sigma_n^2$.

4. Let $S^2 = \frac{\Sigma(X_i - \bar{X})^2}{n-1}$ where $X_1, X_2,\ldots X_n$ are independent identically distributed normal random variable with mean μ and variance σ^2. Show that $\frac{(n-1)S^2}{\sigma^2}$ has a chi-square distribution with $n - 1$ degrees of freedom.

5. If X has $\chi^2(n)$, show that $E(X) = n$ and $Var(X) = 2n$.

6. If $X - \chi^2_{k1}$ and $Y - \chi^2_{k1}$ where X and Y are independent random varaible. Prove that $X + Y - \chi^2_{k1+k2}$.

7. (a) If $X - N(0, 1)$ prove that (i) $Y = X^2$ has $\chi^2(1)$

   (b) If $X - N(\mu, \sigma^2)$ prove that $\left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}}\right)^2 - \chi^2(1)$

8. Let $S^2 = \frac{\Sigma(X_i - \bar{X})^2}{n-1}$ where X_i are independent identically distributed random variables $N(\mu, \sigma^2)$, prove that (i) $E(S^2) = \sigma^2$ (ii) $Var(S^2) = \frac{2\sigma^4}{n-1}$.

   (iii) Show that $\bar{X}$ and S^2 are stochastically independent.

9. Suppose X is uniformly distributed over (0, 1). Find the moment generating function of (i) $Y = -2 \ln X$ and (ii) $Z = -2 \sum_{i=1}^{n} \ln X_i$, where the X_i's are independent uniformly distributed over (0, 1) random variables.

10. Let $X_1, X_2,..$ be independent identically distributed geometric random variable with parameter p. Find the moment generating function of $Y = X_1 + X_2 + ... + X_n$.

11. Let X and Y be independent random variables with the respective pdf N(0, 1) and $\chi^2{}_n$. Find the pdf of $Z = \dfrac{X}{\sqrt{\dfrac{Y}{n}}}$.

12. Let X and Y be independent chi-square random varaibles with parameters k_1 and k_2. Find the pdf if $Z = \dfrac{X/k_1}{Y/k_2}$.

13. Let X_1, X_2 be independent random variables each having an exponential pdf with parameter λ. $Y_1 = X_1 +, ... + X_i; 1 \le i \le n$. Find the joint pdf of $Y_1, Y_2,...Y_n$.

14. Let X_1, X_2 be independent Poisson variables with mean λ. Show that the conditional pdf of X_1 given t has the binomial pdf $b\left(\dfrac{1}{2}, t\right)$, where $t = X_1 + X_2$

7.4 Order Statistics
7.4.1 Independent Identically Distributed Random Variables
Let $X_1, X_2,..., X_n$ be independent continuous random variables, each having distribution function F(x) and probability density function f(x). define

$$X_{(1)} = \text{smallest of } X_1, X_2,... X_n$$
$$X_{(2)} = \text{second smallest of } X_1, X_2,... X_n$$
.
.
.
$$X_{(n)} = \text{largest of } X_1, X_2,... X_n$$

That is, $X_{(1)}, X_{(2)},... X_{(n)}$ are random variable obtained by arranging $X_1, X_2,... X_n$ in an increasing order.

$$X_{(1)} = \min (X_1, X_2,... X_n)$$
$$X_{(n)} = \max (X_1, X_2,... X_n).$$

The random variables $X_{(i)}$ is called the ith order statistics and the ordered values $x_{(1)} \le x_{(2)} \le ... \le x_{(n)}$ are known as the order statistics corresponding to the random variables $X_1, X_2,...X_n$.

For Example, let x_1, x_2, x_3 be a random sample of size 3 from a random variables X having probability density function f(x) $-\infty < x < \infty$.
The possible values of $X_{(1)}, X_{(2)}$ and $X_{(3)}$ are as follows:

$$x_{(1)} = x_1, \quad x_{(2)} = x_2, \quad x_{(3)} = x_3 \text{ if } x_1 < x_2 < x_3$$
$$x_{(1)} = x_2, \quad x_{(2)} = x_1, \quad x_{(3)} = x_3 \text{ if } x_2 < x_1 < x_3$$
$$x_{(1)} = x_3, \quad x_{(2)} = x_2, \quad x_{(3)} = x_1 \text{ if } x_3 < x_2 < x_1$$
.
.
$$x_{(1)} = x_3, \quad x_{(2)} = x_1, \quad x_{(3)} = x_2 \text{ if } x_3 < x_1 < x_2$$

188

There are 3! = 6 permutations of (x_1, x_2, x_3) and hence 6 possible ordered statistics. We shall now find the joint probability density function of $X_{(1)}, X_{(2)}, X_{(3)}$. Since for any permutation, say $(x_2 < x_1 < x_3)$

$$P\left(x_2 - \frac{\in}{2} < X_{(1)} < x_2 + \frac{\in}{2}, x_1 + \frac{\in}{2} < X_{(2)} < x_1 + \frac{\in}{2}, x_{(3)} < x_3 + \frac{\in}{2}\right) = \in f(x_2)$$

$$\in f(x_1) \in f(x_3) = \in^3 f(x_1)f(x_2)f(x_3),$$

as illustrated in the Figure (1) below
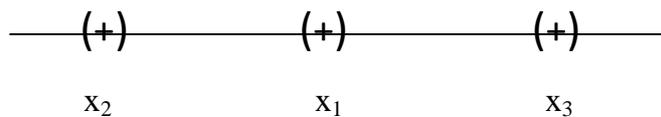

Figure 1

Since there are 3! Possible permutations, the region D is partitioned into six dejoint sets $D_1, D_{2=,...} D_6$. We have $X_{(1)} X_{(2)} X_{(3)}$.

Thus, for $x_1 < x_2 < x_3$

$$P\left(x_1 - \frac{\in}{2} < X_{(1)} < x_1 + \frac{\in}{2} < x_{(2)} < x_2 - \frac{\in}{2}, x_3 - \frac{\in}{2} < X_{(3)} < x_2 - \frac{\in}{2}\right)$$

$$= 3! \ \in^3 f(x_1)(f(x_2))f(x_3)$$

Dividing by $\square^3$ and letting $\varepsilon \rightarrow 0$, we have

$$f(x_{(1)}, x_{(2)}, x_{(3)}) = n! \ f(x_1 \ f(x_2) \ f(x_3); \ x_1 < x_2 < x_3$$

in general, we can extend this idea to a sample of size n and we have for $x_1 \ x_2 \ ... \ x_n$

$$f_{x(1), X_{(2)}, \ ... \ X_{(n)}} (x_1, x_2, ... \ x_n) = n! \ f(x_1)f(x_2)...f(x_n);$$

$$x_1 < x_2 < ... < x_n.$$

the marginal density function of $X_{(i)}$ the ith order statistic can be obtained by integrating the probability density function (5)

$$f_{x_1}(x) = \int ... \int f_{x_{(1)}, x_{(2)}...x_{(n)}}(x_1 x_2, ... x_n) \ dx_2 \ dx_3 \ ... dx_n$$

$$f_{x_{(1)}}(x_i) = \int ... \int f_{x_{(1)}...,x_{(n)}}(x_1, x_2, ... x_n) \ dx_1 \ dx_2 \ ... \ dx_{i-1} dx_{i+1} \ ... dx_n.$$

Example 33

Let x_1, x_2, x_3 be a random sample of a random varaible X having the p.d.f.

$$f(x) = \begin{cases} 1 & 0 < x < 1 \\ 0 & elsewhere \end{cases}$$

Find

(i)   the joint p.d.f. of X_1, X_2 and X_3
(ii) the joint p.d.f of $X_{(1)}, X_{(2)}$ and $X_{(3)}$ the ordered statistics
(iii) the marginal p.d.f of $X_{(1)}$.

Solution

$$f_{x_1, x_2, x_3}(x_1, x_2, x_3) = f_{x_1}(x_1)f_{x_2}(x_2)f_{x_3}(x_3) - 1; \ 0 < x_1, x_2, x_3 < 1$$

$$f_{x_{(1)}, x_{(2)}, x_{(3)}}(x_{(1)}, x_{(2)}, x_{(3)}) = 3! \ f_{x_1}(x_1)f_{x_2}(x_2)f_{x_3}(x_3); \ 0 < x_1, x_2, x_3 < 1$$

$$f_{x_1}(x_1) = \int_{x_1}^{1} \left[\int_{x_2}^{1} 3! \, dx_{(3)} \right] dx_{(2)}$$

$$= 3! \int_{x_1}^{1} (1 - x_2) dx_2 = 3! \left[\frac{1}{2} - x_1 - \frac{1}{2} x_1^2 \right] = \frac{3!}{2} [1 - 2x_1 - x_1^2]$$

$$= 3(1 - x_1)^2$$

Thus,

$$f_{x_{(1)}}(x) = \begin{cases} 3(1 - x)^2 & 0 < x < 1 \\ 0 & elsewhere \end{cases}.$$

Another method of finding the joint marginal p.d.f of $X_{(1)}$, $X_{(2)}$,... $X_{(n)}$ is by direct reasoning as follows:

(i)   For $X_{(i)}$ to equal x, $I - 1$ of $(x_1, x_2....x_n)$ should be less than x and n-1 should be greater than x and 1 of them equals x.

P(I – 1 of the X_i's are less than x) = $[F(x)]^{i-1}$
P(n – 1 of the X_i's are greater than x) = $[1 - F(x)]^{n-i}$
Thus for any given set of partition of $(X_1, X_2...X_n)$ into 3 as stated above, the required probability is from Chapter 1 section 14 number of ways of partitioning a set of n object into the 3 group above is $\frac{n!}{(i-1)!1!(n-i)!}$. *given by*

$$f_{x_1}(x) = \frac{n!}{(i-1)!1!(n-i)!} [F(x)]^{i=1}[1 - F(x)]^{n=1} f(x). \quad (6)$$

(ii)   The joint p.d.f. of $X_{(i)}$ and $X_{(j)} I < j$

(i – 1) values   (j – i – 1)values   (n – j) values

X_i X_j

Figure 3

The probability that exactly i of the X_i's lie in $(-\infty, x)$ and (n-1) lie in (x, ∞) is
$$^nC_i \, [F(x)]^i [1 - F(x_i)]^{n-1}$$
since F(x) = P(X ≤ x) = P(any of the $x_1, ..., X_n$ falls in $(-\infty, x)$.
the event $\{X_{(i)} \le x\}$ occurs if and only i or more of the X_i's lie in $(-\infty, x)$. Thus
$$F_{x_{(1)}}(x) = P(X_{(i)} \le x) = \sum_{j=1}^{n} \, ^nC_j \, [F(x)]^j \, [1 - F(x)]^{n-j} \qquad -\infty < x < \infty \qquad (8)$$
If i = 1, we have
$$F_{x_{(1)}}(x) = \sum_{j=1}^{n} \, ^nC_j \, [F(x)]^j \, [1 - F(x)]^{n-j} \qquad\qquad (9)$$
By Binomial theorem $\sum_{j=1}^{n} \, ^nC_j \, [F(x)]^j \, [1 - F(x)]^{n-j} = 1$
Thus
$$F_{x_{(1)}}(x) = 1 - \, ^nC_j \, [F(x)]^0 \, [1 - F(x)]^{n-0} = 1 - [1 - F(x)]^n \qquad -\infty < x < \infty \qquad (10)$$

Similarly, when we put i = n in (8)

$$F_{x_{(1)}}(x) = [F(x)]^n \qquad -\infty < x < \infty \qquad (11)$$

By differentiating, we have

$$f_{x_{(1)}}(x) = F_{x_1}(x) = n\,[1 - F(x)]^{n-1}\,f(x) \qquad -\infty < x < \infty$$

$$F_{x_{(n)}}(x) = F_{x_{(n)}}(x) = n[F(x)]^{n-1}f(x) \qquad -\infty < x < \infty$$

2.   The Joint Distribution Function of $X_{(1)}$, $X_{(n)}$.

To find the joint distribution function of X_i and X_n, we proceed as follows.

$$\{(x_1, x_n) : X_{(1)} \le x_1, x_{(n)} \le x_{(n)}\} = \{(x_1, x_n) : X_{(n)} \le x_{(n)}\}\backslash\{x, y)\}:$$
$$X_{(1)} > x_1, x_{(n)} \le x_n\}.$$

See figure 5 below.
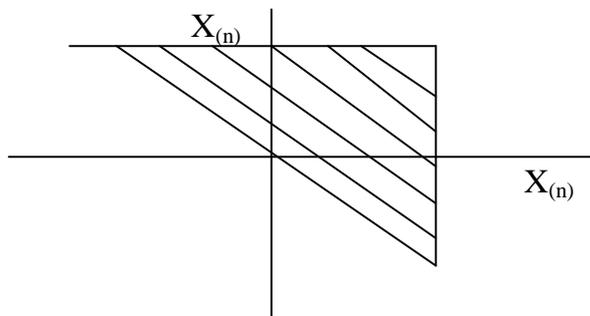
Figure 5

We know that $P(X_{(n)} \le x_n) = [F(x_n)]^n$, therefore

$P(X_{(1)} > x_1, X_{(n)}) = P(x_1 < X_1 \le x_n, x_1 < X_2 \le x_n \dots x_1 < X_n \le x_n)$
$$= (F(x_n) - F(x_1))^n$$

Thus

$$f_{x_{(1)}x_{(n)}}(x_1, x_n) = P(X_{(1)} \le x_1, X_{(n)} \le x_n) = [F(x_n)]^n - \{F(x_n) - F(x_1)\}^n \qquad (12)$$

from which the density function can be derived by differentiation:

$$f_{x_{(1)},x_{(2)}}(x_1, x_n) = \frac{\partial^2}{\partial x_1 \partial x_n} F_{x_{(1)}x_{(n)}}(x_1, x_n)$$

$$= n(n-1)\,f(x_1)f(x_n)\,[F(x_n) - F(x_1)]^{n-2}, \; x_1 \le x_n \qquad (13)$$

Distribution Function of the Range of a Random Sample

The random variable R, defined by $R = X_{(n)} - X_{(1)}$ is called the range of the observed random variables.

$$P(R \le r) = P(X_{(n)} - X_{(1)} \le r) = \iint_{x_n - x_1 \le r} f_{x_{(1)},x_{(n)}}(x_1, x_n)\,dx_1 dx_2$$

$$= \int_{-\infty}^{\infty} \int_{x_1}^{x_1 + r} n(n-1)\,[F(x_n) - f(x_1)]^{n-2}\,f(x_1)\,f(x_n)\,dx_n dx_1.$$

Making the change of variable $y = F(X_n) - F(x_1)$, $dy = f(x_n)\,dx_n$

$$\int_{x_1}^{x_1+r} [F(x_n) - F(x_1)]^{n-2} f(x_n) dx_n = \int^{F(x_1+r)-F(x_1)} y^{n-2} dy = \frac{1}{n-1} [F(x_1 + r)$$
$$- F(x_1)]^{n-1}$$

Thus
$$F_R(r) = P(R \le r) = n\int_{-\infty}^{\infty}[F(x_1 + r) - F(x_1)]^{n-1}f(x_1)dx_1 \qquad (14)$$
$$F_R(r) = n(n-1)\int_{-\infty}^{\infty}F(x_1)f(x_1 + r)[F(x_1 + r) - F(x_1)]^{n-2}dx_1$$

Example 40
Let $X_1, X_2,\ldots X_n$ be a random sample from uniform distribution on (0, 1). Find
(i)   The distribution function of $R = X_{(n)} - X_{(1)}$
(ii)   The probability density function of R.

Solution
(i) $P(R \le r) = \int_0^x [F(x_1 + r) - F(x_1)]^{n-1}f(x_1)dx_1$
        Where $F(x) = \int_0^x 1 \, dx = x$;
$$F(x_1 + r) = 1 \text{ if } x_1 + r > 1$$
$$= x_1 + r \text{ if } x_1 + r < 1.$$
Thus
$$P(R \le r) = n\int_0^{1-r} r^{n-1}dx_1 + n\int_{1-r}^1 (1 - x_1)^{n-1}dx_1$$
$$F_R(r) = n(1-r)r^{n-1} + r^n$$
(ii) $f_R(r) = F'_R(r) = \begin{cases} n(n-1)(1-r)r^{n-2} & 0 \le r \le 1 \\ 0 & otherwise \end{cases}$.

Another method of finding the probability density function of R is by using the formula (1) for sum of random variables.
$$f_R(r) = \int_{-\infty}^{\infty} f_{X_{(1)},X_{(n)}}(x_1, r + x_1)dx_1$$
$$= \begin{cases} n(n-1)\int_{-\infty}^{\infty} f(x)f(r + x)[F(r + x) - F(x)]^{n-2}dx & r > 0 \\ 0 & r < 0 \end{cases}$$

Example 41
Let $X_1, X_2, \ldots X_n$ be independent random variables having the uniform distribution on (a, b). Find the pdf of (i) $X_{(1)} = \min(X_1, X_2,\ldots X_n)$ (ii) $X_{(n)} = \max(X_1, X_2,\ldots X_n)$.

Solution
The pdf of X is
$$f_x(x) = \frac{1}{b-a}, a < x < b.$$
The pdf of $X_{(i)}$ is given by
$$f_{X_{(1)}}(x) = \frac{n!}{(i-1)!(n-1)!} [F(x)]^{i-1}[1 - F(x)]^{n-1}f(x)f(x)$$
where
$$F(X) = \int_a^x \frac{1}{b-a} \, dx = \frac{x-a}{b-a}$$
$$f_{X_{(1)}}(x) = n\left[1 - \frac{x-a}{b-a}\right]^{n-1} \frac{1}{b-a} = \frac{n(b-x)^{n-1}}{(b-a)^n}, a < x < b$$
(ii) $f_{X_{(1)}}(x) = n\left[\frac{x-a}{b-a}\right]^{n-1} \frac{1}{b-a} = \begin{cases} \frac{n(b-x)^{n-1}}{(b-a)^n}, a < x < b \\ 0 \qquad elsewhere \end{cases}$

Exercise 7.4

1. Let $X_1, X_2,\ldots X_n$ be independent and each is uniformly distributed over (a, b). find the joint p.d.f of $X_{(1)}, X_{(2)},\ldots X_{(n)}$.

2. Let $X_1, X_2,\ldots X_n$ be independent random variables each having an exponential density with parameter λ. Find the pdf of (i) $X_{(1)} = \min(X_1, X_2,\ldots X_n)$
   (ii) $X_n = \max(X_1, X_2,\ldots X_n)$.

3. Let $X_1, X_2,\ldots X_n$ (n odd) be independent random varaibles each uniformly distributed over $(0, 1)$. Find the p.d.f of the median of the sample.
   Note: When n is odd, the median is $X_{\left(\frac{n+1}{2}\right)}$.

4. Show that if n people are distributed at random along road Y miles long, then the probability that no 2 people are less than a distance k miles apart is
   $$\left[1 - \frac{(n-1)k}{y}\right]^n, k < \frac{Y}{n-1}$$

5. Let $X_1, X_2,\ldots X_{2n-1}$ be a random sample from uniform distribution on $(0, 1)$. Show that the median of the sample has a beta distribution with parameter $(n + 1, n + 1)$.

LIMIT THEOREMS

The Law of Large Numbers and the Central Limit Theorem
The most important theoretical results in probability are limit theorems. In this chapter we prove often useful tools in probability, the Chebychev's inequality. This inequality is then used to deduce the law of large numbers for independent and identically distributed random variables. In section 8.5 the central limit theorem which is a fundamental result in the theory and applications f probability theory is given.

8.1   Chebychev's Inequality
Chebychev;s inequality gives an upper bound in terms of variance of a random variable X for probability that X deviates from its mean by more than k units.

Theorem 8.1:*Chebychev's Inequality*
Let X be a random variable (discrete or continuous) with mean μ and variance σ^2. Then for any positive number k we have

$$P\{ \left| X\text{-}\mu \right| > k\} \le \frac{\sigma^2}{k^2}$$

Or equivalently

$$P\{ \left| X\text{-}\mu \right| \ge k\,\sigma\}\} \le \frac{1}{k^2}.$$

Proof:
Let X be a non-negative random variable such that $E(X) = \mu <\infty$.
Define another random variable Y as follows:

$$Y = \begin{cases} 0 & \text{if } X < k \\ k & \text{if } X > k \end{cases}$$

This new variable Y is a discrete variable having two values 0,k. The probability density function of Y can be written thus:

y	0	k
P(Y = y)	P(X < k)	P(X > k)

Hence,

$$E(Y) = OP(X < k) + kP(X > k)$$

That is

$$E(Y) = kP(X \ge k).$$

Since the variable $X \geq Y$ for all possible values, we have
$$E(X) \geq E(Y) = kP(X \geq k),$$
Thus,
$$P(X \geq k) \leq \frac{E(X)}{k}. \qquad (2)$$

Equation (2) is called the Markov inequality and can be generalized thus: For any $j \geq 0$, $k > 0$
$$P\{ |W| \geq k \} \leq \frac{E|W|^j}{k^j} \qquad (3)$$
The proof of Chebychev's inequality is an immediate consequences of (3) by putting $j = 2$ and $W = X - \mu$, we obtain
$$P\{ |X - \mu| \geq k \} \leq \frac{E(X-\mu)^2}{k^2} = \frac{\sigma^2}{k^2}$$
Since $(X - \mu)^2 \geq k^2 <==> |X-\mu| > k$, we have
$$P\{ |X-\mu| > k \} = P\{(X - \mu)^2 \geq k^2\} \leq \frac{\sigma^2}{k^2}.$$
Thus
$$P\{ |X-\mu| \leq k \} \leq \frac{\sigma^2}{k^2}.$$

Example 8.1
Let X be a random variable having Poisson distribution with mean λ and variance σ^2. Use Chebychev's inequality to show that (i) $P\{P |X-\lambda| \geq 1\} \leq \lambda$ (ii) $P\left(X > \frac{3\lambda}{2} \right) \leq \frac{4}{\lambda}$.

Solution .
Using Chebychev's inequality with $\mu = \lambda$, and $\sigma^2 = \lambda$ we have
(i) $P\{ |X-\mu| \geq k \} \leq \frac{\lambda}{k^2}.$
     Putting $k = 1$ we obtain
     $$P\{ |X-\mu| \geq k \} \leq \lambda.$$
(ii) $Let \frac{\sigma^2}{k^2} = \frac{4}{\lambda} ===> \sigma^2 = \frac{4k^2}{\lambda}$
     On substituting for σ^2 we have
     $$\lambda^2 = 4k^2, k^2 = \left(\frac{\lambda}{2}\right)^2, k = \frac{\lambda}{2}.$$

Thus
$$P\{ |X-\mu| \geq \frac{\lambda}{2} \} \leq \frac{4}{\lambda}.$$
Since $|X-\lambda| \geq \frac{\lambda}{2}$ if and only if $X-\lambda < \frac{-\lambda}{2}$ or $X-\lambda > \lambda/2$, we see that
$$P\{ |X-\lambda| \geq \frac{\lambda}{2} \} = P\{X-\lambda < -\frac{\lambda}{2}) + P\left(X - \lambda > \frac{\lambda}{2}\right)$$
$$= P((X < \frac{\lambda}{2}) + P\left(X > \frac{3\lambda}{2}\right) \leq \frac{4}{\lambda}$$
Since $P\left(X < \frac{\lambda}{2}\right) \geq, we\ have$
$$P\left(X > \frac{3\lambda}{2}\right) \leq \frac{4}{\lambda}.$$

8.2 The Law of Large Numbers of Bernoulli Trails

Let $X_1, X_2, \ldots, X_n$ be n independent and identically distributed Bernoulli random variables and let $X = X_1 + X_2 + \ldots + X_n$ be Binomial random variables (number of successes), with parameters n and p. The mean and variable of X are np and npq respectively.

The mean μ grows as n increases but the standard deviation grows only as $\sqrt{n}$, Using Chebychev's inequality

$$P\{ \left| X - np \right| > \varepsilon \} \le \frac{npq}{\varepsilon^2} \le \frac{n}{\varepsilon^2}.$$

Theorem 8.2

Let X be the number of successes in n independent Bernoulli trails with probability p of success. For any $\varepsilon > 0$.

$$P\left\{ I\frac{X}{n} - PI > \varepsilon \right\} \le \frac{pq}{n\epsilon^2} \le \frac{1}{4\epsilon^2 n}$$

and

$$\lim_{n \to \infty} P\left\{ I\frac{X}{n} - PI > \varepsilon \right\} = 0$$

Proof:

Applying Chebychev's inequality

$$E\left(\frac{X}{n}\right) = P, Var\left(\frac{X}{n}\right) = \frac{1}{n^2} \cdot npq = \frac{pq}{n}$$

$$\frac{Var\left(\frac{X}{n}\right)}{\epsilon^2} = \frac{pq}{n\epsilon^2} -> 0 \text{ as } n-> \infty$$

$$pq = (1-p)p \le \frac{1}{4} \text{ for all } p, 0 \le p < 1$$

Hence

$$P\left\{ \left|\frac{X}{n} - p\right| \ge \epsilon \right\} -> 0 \text{ as } n-> \infty$$

This means that for large number we can be almost certain that $\frac{X}{n}$ will be very close to probability p of success. This shows that the relative frequency of success in independent Bernoulli trails converges (in a probabilistic sense) to the theoretical probability p of success at each trail. This important result can be stated more precisely as the law of large numbers.

The Law of Large Numbers
8.3 The Weak Law of Large Numbers:

Let $X_1, X_2, \ldots, X_n$ be independent and identically distribution random variables with all $E(X_i^2) < \infty$ and let $S_n - X_1 + X_2 + \ldots + X_h$.

We know that

$$E(X_1) = E(X_2) = \ldots E(X_n), E(S_n) = nE(X_1) \text{ and } Var(S_n) = nVarX_1$$

$$E\left(\frac{S_n}{n}\right) = \frac{1}{n} E(S_n) = \frac{1}{n} \cdot n E(X_1) = E(X_1)$$

$$Var\left(\frac{S_n}{n}\right) = \frac{1}{n^2} Var(S_n) = \frac{nVarX_1}{n^2} = \frac{1}{n} Var(X_1)$$

This shows that expectation of $\frac{S}{n}$ is equal to the expectation of X_i and the standard deviation of $\frac{S_n}{n}$ is

$$\sqrt{\left\{\frac{VarX}{n}\right\}} = \frac{StdX_1}{\sqrt{n}}$$

Which tends to zero as n tends to infinity. Thus, the distribution of S_n/n becomes more and more concentrated near $E(X_1)$.

We state the more general result in the following theorem.

Theorem 8.3: A (WLLN) Weak Law of Large Numbers for Independent Random Variables. Let X_1, X_2...X_n be independent and identically distributed random variables with finite mean μ and σ^2. Let $S_n = X_1 - X_2 + \ldots + X_n$. then for $\delta > 0$.

$$\lim_{n \to \infty} P\left(\left|\frac{S_n}{n} - \mu\right| > \delta\right) = 0 \qquad (4)$$

Proof:

Applying Chebychev's inequality to $\frac{S_n}{n}$, we have

$$\mu = E\left(\frac{S_n}{n}\right). Var\left(\frac{S_n}{n}\right) = \frac{1}{n^2} Var(S_n)$$

$$P\left\{\left|\frac{1}{k}S_n - E(X)\right| > \delta\right\} \le \frac{2\sigma^2 X}{\sigma^2 n}$$

For at least one k satisfying $n \le k \le m$}.

Corollary 1:

Let f(x) be a continuous real function on [0, 1]. Then as n -> ∞ uniformly with respect to p ε [0, 1]

$$E\left\{f\left(\frac{S_n}{n}\right)\right\} \to f(p).$$

Corollary II: Weierstress Approximation Theorem.

Let f(X) be a continuous real function on [0, 1] defined on the real interval $0 \le p \le 1$ and let

$$P_n(p) = \sum_{k=0}^{n} f\left(\frac{k}{n}\right) {}^nC_k p^k (1-p)^{n-k}.$$

Then $P_n(p)$ -> ∞ f(p) as n -> ∞ and the convergence is uniform in p. $P_n(x)$ is called the Bernstein Polynomials.

Theorem 8.5

A WLLN for independent but not necessarily identical random variables.

Let X_1, X_2,...X_n be independent random variables with all

$E((X^2) < \infty$ and let $S_n = X_1 + X_2 + \ldots + X_n$ and $\mu = E\left(\frac{S_n}{n}\right)$, then

$$P\left\{\left|\frac{S_n}{n} - \mu_n\right| \ge \epsilon\right\} \le \frac{1}{\epsilon^2 n^2}\{Var(X_1) + Var(X_2) + \cdots + Var(X_n)\}. \varepsilon > 0.$$

When $VarX_i \le M$ for all I, that is when the variances are all bounded bya constant which does not depend on n, then we have

$$P\left\{\left|\frac{S_n}{n} - \mu_n\right| \ge \epsilon\right\} \le \frac{M}{\epsilon^2 n} -> 0 \ as \ n-> \infty.$$

8.5 The Central Limit Theorem

The central limit theorem is a fundamental result in the theory and application of

probability. The law of large numbers asserts that $\frac{S_n}{n}$ in a series of n independent trails

with constant probability p of success in one trail tends (in a certain probabilistic sense) to p as n increases. But this assertion does not tell us anything about the distribution of $\frac{S_n}{n}$ as n becomes large. The answer to this question is given by the so called central limit theorem. The theorem asserts that under quite general conditions the sum of independent variables has the Normal distribution in the limit.

Theorem 8.6: Central limit theorem

Let X_1, X_2, …, X_n be a sequence of independent random variables with $E(X_i) = \mu$ and $Var(X_i) = \sigma_i^2$ I = 1, 2,…Let $S_n = X_1 + X_2 +…+ X_n$. Then under general conditions,

$$Z_n = \frac{S_n - E(S_n)}{\sqrt{Var\ S_n}} = \frac{S_n - \sum_{i=1}^{n} \mu_i}{\sqrt{\sum_{i=1}^{n} \sigma_i^2}}$$

Has approximately the standard Normal distribution.

The Identically Distribution Case
Theorem 8.7
Let X_1, X_2, …, X_n be independent identically distributed random variables. Then

(i) $\quad Z_n = \frac{S_n - n\mu}{\sqrt{n\sigma^2}}$ (ii) $\quad Z_n = \frac{\frac{S_n}{n} - \mu}{\sqrt{\sigma^2/n}}$

tend to the standard normal distribution as n -> ∞.

What is remarkable about this theorem is that all we need to know about X is its mean and variance and there is no need to specify the nature of its distribution. Also X can be discrete, continuous or both. The outline of the proof is given below,

Proof:

Let $M_X(t) = E(e^{tX})$, $S_n(t) = [E(e^{tX})]^n$

Let $Z_n = \dfrac{sn - n\mu}{\sqrt{n\sigma^2}} = \dfrac{sn}{\sqrt{n}\sigma} - \sqrt{n}.\dfrac{\mu}{\sigma}$.

The m.g.f of Z_n is

$$M_{Z_n}(t) = e^{(\sqrt{n\mu/\sigma})t}\left[M_X\left(\frac{t}{\sigma\sqrt{n}}\right)\right]^n$$

Taking log we have

$$In\, M_{Z_n}(t) = -\sqrt{n}\frac{\mu}{\sigma}t - n\, In\, M_X\left(\frac{1}{\sigma\sqrt{n}}\right).$$

Using Maclaurin series

$$M(t) = 1 + M'(0)t + M''\frac{(0)t^2}{2!} + \cdots$$

From above we have

$$In\, M_{Z_n}(t) = -\sqrt{n}\frac{\mu t}{\sigma}t + n\, In\left[1 + \frac{\mu t}{\sigma\sqrt{n}} + \frac{(\mu^2 - \sigma^2)t^2}{2n\sigma^2} + A\right] \qquad (7)$$

where A is the remainder term from Maclaurin series. Similarly, when the Maclaurin series

$$In\,(1 + x) = x - \frac{X^2}{2} + \frac{X^3}{3} + \cdots$$

is used to expand $In\left(1 + \frac{\mu t}{\sqrt{n}\sigma} + \frac{(\mu^2 + \sigma^2)}{n\,\sigma^2}t + A\right). setting\; x = \frac{\mu t}{\sqrt{n}\sigma} + \frac{(\mu^2 + \sigma^2)t^2}{2n\sigma^2} + A,$

we find that

$In\left(1 + \frac{\mu t}{\sqrt{n}\sigma} + \frac{(\mu^2 + \sigma^2)}{n\,\sigma^2}t + A\right)$

$= \frac{\mu t}{\sqrt{n}\sigma} + \frac{(\mu^2 + \sigma^2)}{n\,\sigma^2}t + A - \frac{1}{2}\left[\frac{\mu t}{\sqrt{n}\sigma} + \frac{(\mu^2 + \sigma^2)}{n\,\sigma^2}t + A\right]^2 + \cdots$

It is left as an exercise to show that for large n.

$$\left|\frac{\mu t}{\sqrt{n}\sigma} + \frac{(\mu^2 + \sigma^2)}{2n\sigma^2} + A\right| < 1.$$

On substituting in (7) we obtain

$$In\ M_{Z_n}(t) = -\sqrt{n}\frac{\mu t}{\sigma}$$

$$+ n\left\{\frac{\mu t}{\sigma\sqrt{n}} + \frac{(\mu^2 - \mu^2)}{2n\sigma^2} + A - \frac{1}{2}\left[\frac{\mu t}{\sigma\sqrt{n}} + \frac{(\mu^2 + \sigma^2)t^2}{2n\sigma^2} + A\right]^2 + \ldots\right\}$$

It can easily be shown that as n -> ∞

$$In\ M_{Z_n}(t) = \frac{t^2}{2}.$$

Hence,

$$\lim_{n\to\infty} M_{Z_n}(t) = e^{t^2/2},$$

the moment generating function of a random variable having the standard normal distribution.

Theorem

Let X_n- n > 1, and X be random variables such that $\lim_{n\to\infty} M_{X_\infty}(t) = M_X(t), -\infty < t <$ ∞. Then Lim *FXnx= FXxat all points where FXx* is continuous.

Example 8.2

A boy throws a fair die 100 times. What is the probability that his mean score will exceed 3.

Let X represent the outcome of each toss.

Possible values of X are 1, 2, 3, 4, 5, 6.

E(X) = 3.5. Var(X) = 2.92.

That is μ = 3.5. σ^2 = 2.92

Let X be the mean score. By the central limit theorem X is normally distributed with mean 3.5 and Varσ^2/n.

$$\frac{\sigma^2}{n} = \frac{2.92}{100}.$$

Hence,

$$Z = \frac{\bar{X}- 3.5}{\sqrt{0.029}}$$

has a standard normal distribution.

$$P(\bar{X} > 3) = P\left(Z > \frac{-0.5}{0.170}\right) = P(Z > -2.94) = 0.9984.$$

Example 8.3

Marks in an I.Q examination are normally distributed with mean 55 and standard deviation 10. What is the probability that

(i)   the mean mark of a group of 10 students will be above 50

(ii)   the mean mark of a group of 20 students will be between 40 and 50.

(iii)   the sum of the marks of the 10 students will be less than 500.

Solution

Let X_1, X_2,… X_{10} be the mark scored by each student respectively.

$$E(X_1) = E(X_2) = … = E(X_{10}) = 55$$

$$\sigma_{X_1} = \sigma_{X_2} = \sigma_{X_n} = 10.$$

(i) $\bar{X}$ = mean, be the central limit theory (assuming n = 10 is large enough for this distribution) $\frac{\bar{X} - 55}{\sigma/\sqrt{n}}$ has a standard Normal distribution.

$$\frac{\bar{X} - 55}{10/\sqrt{10}} = \frac{\bar{X} - 55}{\sqrt{10}}$$

$$P(\bar{X} > 50) = P\left(\frac{\bar{X} - 55}{\sqrt{10}} > \frac{50 - 55}{\sqrt{10}}\right) = 0.9306.$$

(ii)   The mean mark of a group of 20 students between will be between 40 and 50. In this case, n = 20.

$$P(40 < \bar{X} < 50) = P(\bar{X} < 50) - P(\bar{X} \le 40)$$

$$P(\bar{X} \le 40) = P\left(z \le \frac{40 - 55}{\sqrt{5}}\right) = P(Z \le -3\sqrt{5}) = 0.00$$

$$P(\bar{X} < 50) = P\left(z < \frac{50 - 55}{\sqrt{5}}\right) = P(Z < -\sqrt{5}) = 0.01$$

Hence

$$P(40 < \bar{X} < 50) = 0.01 - 0.00 = 0.01$$

(iii) $S_{10} = X_1 + X_2 + … + X_{10}$

$\frac{S_{10} - n\mu}{\sqrt{n\sigma^2}}$ has a standard Normal

$$\frac{S_{10} - 550}{\sqrt{1000}} = \frac{500 - 550}{10\sqrt{10}} = \frac{-50}{10\sqrt{10}} = -1.53$$

Hence,

$$P(S_n < 500) = 0.057.$$

8.6   Normal Approximately to the Binomial Distribution

Let X_1, X_2, …, X_n be n independent identically distributed Bernoulli trails where

$$X_i = \begin{cases} 1 & \text{with probability P} \\ 0 & \text{with probability } 1 - P \end{cases}$$

$E(X_i) = p$ and $Var(X_i) = p(1 - p)$. From theorem 2, the distribution

$\frac{S_n - E(S_n)}{\sqrt{n\sigma^2}}$ --> standard Normal.

where $S_n = X_1 + X_2 + … + X_n$

$$E(S_n) = np$$

$$Var(S_n) = np(1 - p), \sigma^2 = p(1 - p).$$

Note that S_n has binomial distribution with mean np and variance np(1-P).

$$P(S_n \leq y) = \sum_{k=0}^{y} {}^{n}C_k P^k (1 - p)^{n-k}$$

For large n

$$P(S_n \leq y) = P\left\{ Z_n \leq \frac{y - np}{\sqrt{np(1 - p)}} \right\}$$

$$\varphi = \left(\frac{y - np}{\sqrt{np(1 - p)}} \right)$$

where Z_n is standard Normal random variable. Thus, from above we have

$$F_{S_n}(y) = P(S_n \leq y) = \sum_{k=0}^{y} {}^{n}C_k p^k (1 - p)^{n-k} = \varphi\left(\frac{y - np}{np(1 - p)} \right)$$

Example 8.3

It is claimed that 60% of the voters I a given ward are going to vote for party A. assuming that all voters will vote, and that there are 100 voters. What is the probability that Party A receives at least 50 votes.

Let p = 0.6. prob. Of voting for A. n = 100. Assuming independent voting

$P(S_n > 50) = \sum_{k=50}^{100} 100C_k p^k (1 - p)^{100-k} = \sum_{k=50}^{100} 100C_k (0.6)^k (0.4)^{100-k}$

Using normal approximation (6) above, we have

$P(S_n > 50) = P\left(Z \geq \frac{50 - 100 \times 0.6}{\sqrt{100(0.6)(0.4)}} \right) = P\left(Z \geq \frac{50 - 60}{\sqrt{24}} \right) = 0.9793$

8.7 The Normal Approximation to the Poisson Distribution

If a random variable X has a poisson distribution whose mean is λ, then for large λ the standardized random variable.

$$Z = \frac{X - \lambda}{\sqrt{\lambda}}$$

has a standard Normal Distribution. you will recall that $E(X) = \lambda$, $Var(X) = \lambda$,

Continuity correction may be introduced.

It has been shown that the approximation is good for $\lambda > 5$.

Example 8.4

A system suffers random breakdown at a constant rate of 10 per month. Find the probability that there will be at least 8 breakdowns in any month.

Let X be the number of time the system breaks down in any month. Then X has a poisson distribution with $\lambda = 10$. Thus

$$P(X > 8) = \sum_{k=8}^{\infty} \frac{e^{10}(10)^k}{k!}$$

Using Normal approximation, we have

$$\frac{X - 10}{\sqrt{10}} \text{ has N(0.1)}$$

$$P(X \geq 8) = P\left(Z > \frac{8 - 10}{\sqrt{10}}\right) = P\left(Z > \frac{-2}{\sqrt{10}}\right) = 0.7357$$

Exercise 8

1. It is claimed that 30% of the voters in a given local government area are going to vote for party A in a local government election. Assuming that all voters will vote, there are 4000 voters and the claim is based on a proper (unbiased) sampling method. What is the probability that Party A will receive more than 1,500 votes.

2. (a) Suppose that a system consists of components each of which has a probability of 0.05 of failing during a specific time. The system functions properly when at least 150 components function. Assuming these components function independently of one another, what is the probability that the system functions properly during a specific time.

    (b) Suppose that the above system is made of n components each having probability of 0.90 of not failing during a given time. The system will

function if at least 80 percent of the components function properly. Determine n so that the probability that the system functions properly during a specific time is 0.96.

3. Let X be a non-negative interger-valued random variable whose probability generating function $P_X(s) = E(s^X)$ is finite all s.

Use Chebychev's inequality to verify the following inequalities:

(a) $P(X \leq k) \leq \frac{P_X(s)}{s^k}.0 \leq s \leq 1, k$ any positive inequalities:

(b) $P(X > k) \leq \frac{P_X(s)}{s^k}.s > 1$

4. Prove the Chebychev's inequality.

Let X be a random variable (discrete or continuous) with $E(X) = \mu$ and $Var(X) = \sigma^2$. Then for any positive constant number k we have

$$P\{|X - \mu| \geq k\sigma\} \leq \frac{1}{k^2}.$$

5. Prove that (i) $1 - x \leq e^{-x}$ for real x, (ii) $\log x \leq x - 1$ for x > 0, and

(iii) $\prod_{i=1}^{N} X_i \leq e^{-\sum_{i=1}^{N} X_i}$ if $X_i \leq 1, N = 1,2,\ldots$