PROBABILITY AND DISTRIBUTIONS (DBSTT11/DSTT11) (BSC STASTICS - I)

ACHARYA NAGARJUNA UNIVERSITY

CENTRE FOR DISTANCE EDUCATION

NAGARJUNA NAGAR,

GUNTUR

ANDHRA PRADESH

Lesson - 1

ANALYSIS OF QUANTITATIVE DATA

Object:

After studing the lesson the students are expected to have clear comprehension of the theory and practical utility about the concepts of measures of central tendency-mean, median and mode; Measures of Dispersion, relative dispersion and their area of applications.

Structure of the lesson:

This consists of .... sections as detailed below:

1.1 Measure of Central Tendency - Introduction:

According to professor Boweley averages are "Statistical Constants which enable us to understand in a single effort the significance of the whole." They give us an idea about the concentration of the values in the central part of the distribution. So they are called measures of central tendency.

1.1.1 Characteristics of a good average: According to G.V. Yule, the properties of a good average are as follows:

1. It should be well defined.
2. It should be easy to calculate.
3. It should be capable of further algebraic treatment.
4. It should be based on all the observations.
5. It should not be affected by extreme observations.
6. It should not be affected by fluctuations in sampling.

1.1.2 Arthmetic Mean: The arithmetic mean or simple mean of set of observations is defined as teir sum, divided by the number of observations.

$$\text{i.e.,}\quad \text{mean} = \frac{\text{sum of observation}}{\text{number of observations}}$$

Let $x_1, x_2, \cdots\cdots\cdots, x_n$ be the n observations then their mean is denoted as $\bar{x}$ and it is given by

$$\text{mean}\quad \bar{x} = \frac{x_1 + x_2 + \cdots\cdots + x_n}{n}$$

$$= \frac{1}{n}\sum_{i=1}^{n} x_i$$

In the case of grouped frequency distribution, if $x_1, x_2, \cdots\cdots, x_n$ be the mid values of class interval with corresponding frequencies $f_1, f_2, \cdots\cdots, f_n$ respectively, then their mean $\bar{x}$ is given by

$$\text{mean}\quad \bar{x} = \frac{f_1x_1 + f_2x_2 + \cdots\cdots + f_nx_n}{f_1 + f_2 + \cdots\cdots + f_n}$$

$$= \frac{\sum_{i=1}^{n} f_ix_i}{\sum f_i}$$

$$= \frac{1}{N}\sum_{i=1}^{n} f_ix_i$$ where N in total frequency is $N = \sum f_i$

Calculation of Mean by Change of origin and Scale: If x values and corresponding frequencies (f) are large then the calculation of mean takes large time. This can be reduced by taking th deviations of the given values from any arbitrary point "A" as below:

Let $d_i = x_i - A$

then $f_i - d_i = f_i x_i - f_i \cdot A$

$$\Rightarrow \sum f_i d_i = \sum f_i x_i - A \sum f_i$$

$$\Rightarrow \frac{1}{N} \sum f_i d_i = \frac{1}{N} \sum f_i x_i - A \frac{1}{N} \sum f_i$$

$$\Rightarrow \frac{1}{N} \sum f_i d_i = \bar{x} - A \qquad \because \sum f_i = N$$

$$\therefore \bar{x} = A + \frac{1}{N} \sum f_i d_i$$

This formula is much convenient to apply. It is also called short-cut method.

In the case of frequency distribution having equal class interval say having width "h", it is convenient to use change of origin (A) and scale (h).

If $U_i = \frac{x_i - A}{h}$ then

$$x_i = A + hU_i$$

$$\Rightarrow \sum f_i x_i = A \sum f_i + h \sum f_i U_i$$

$$\Rightarrow \frac{1}{N} \sum f_i x_i = A \frac{1}{N} \sum f_i + h \cdot \frac{1}{N} \sum f_i U_i$$

$$\Rightarrow \bar{x} = A + h \, \bar{U} \qquad \because \sum f_i = N$$

Hence mean is effected by both change of origin and scale. This formula is also called step deviation method.

Properties of mean:

1. Algebric sum of the deviations of a set of values from their mean is zero. i.e. $\sum \left(x_i - \bar{x} \right) = 0$

2. Sum of the squared deviations of set of values is minimum when deviations are taken from mean of the observation.

3. Let $\bar{x}_1$ and $\bar{x}_2$ be the man of n_1 and n_2 observations. The combined mean of $(n_1 + n_2)$ observations is given by

$$\bar{x} = \frac{n_1\bar{x}_1 + n_2\bar{x}_2}{n_1 + n_2}$$

4. Let W_i be the weights attached to the items x_i ; $i = 1, 2, ..., n$ then the weighted mean is given by $\bar{x}_w = \frac{\sum w_i x_i}{\sum w_i}$.

Merits and Demerits of Man:

Merits of Mean:
1. It is rigidly defined
2. It is easy to understand and easy to calculate
3. It is based on all the observations
4. It is sutable for algebric treatment
5. Mean is an ideal average.

Demerits of Mean:
1. It can not b determined by inspection or graphical method.
2. It is not sutable for qualitative data.
3. It can not be obtained if one or more observations is missed.
4. It is affected very much by extreme values.
5. It can not be suitable if te extreme class is open.

1.1.3 Median: The median is defined as the middle most or the central value of the variate, when the observations are arranged in ascending or decending order of their magnitudes.

In the case of ungrouped data, having size n.

if n is odd then median is middle most observation.

i.e. $\text{median} = \left(\frac{n+1}{2}\right)^{th}$ observation.

If n is even the median is mean of two middle terms.

i.e. $\text{median} = \frac{\left(\frac{n}{2}\right)^{th} + \left(\frac{n}{2}+1\right)^{th}}{2}$ observation.

For the grouped data the median is defined by

$$\text{median} = \ell + \frac{\left(\frac{N}{2} - m\right) \times h}{f}$$

Where ℓ is the lower limit of the median class.

f is frequency of median class

m is cumutative frequency of the class preceeding the median class

N is total frequency

h is width of class interval of median class

Note: To decide the median class calculate $\frac{N}{2}$ value and see the cumulative frequency which is more than or equal to $\frac{N}{2}$, the corresponding class is called Median Class.

Merits and Demerits of Median:

Merits:

1. It is rigidly defined
2. It is easy to understand and easy to calculate
3. It is not effected by extreme values.
4. It can be used to calculate for distributions with open end classes.

Demerits:

1. It is not based on all the observations.
2. It is not sutable for mathematical treatment.
3. As compared with mean, it is affected much by fluctuations of sampling.
4. In case of even number of observations median can not be determined exactly.

Note: To apply median formula, the frequency distribution must be continuous frequency distribution.

1.1.4 Mode: Mode is the value which occurs most frequently in a set of observations.

In the case of grouped continuous frequency distribution mode is given by the formula.

$$\text{Mode} = \ell + \frac{(f_1 - f_0)h}{2f_1 - (f_0 + f_2)}$$

Where ℓ is the lower limit of model class

f_1 is the frequency of the modal class

f_0 is the frequency of class preceding model class

f_2 is the requency of class succeeding the model class.

h is width of model class.

If the distribution is moderately asymmetrical, the mean, median and mode satisfy the following empirical relationship.

$$\text{mean} - \text{mode} = 3(\text{mean} - \text{median})$$

$$\Rightarrow \text{mode} = 3\ \text{median} - 2\ \text{mean}$$

Merits and Demerits of Mode:

Merits:
1. Mode can obtain some times by inspection.
2. Mode is readily comprehensible and easy to calculate.
3. Mode is not affected by extreme values.
4. Mode can be conveniently located even if t frequency distribution has unequal class intervals, but same width for model class, preceding and succeeding classes.

Demerits:
1. Mode is ill-defined. It is not always possible to find a clearly defined mode.
2. It is not based upon all the observations.
3. It is not suitable of further mathematical treatment.
4. Mode is effected by fluctuations of sampling.

1.1.5 Workedout Examples:

Example 1: Find the arithematic mean of the numbers 80, 30, 50, 120, 100.

Solution: $\text{Mean } (\bar{x}) = \frac{80+30+50+120+100}{5} = \frac{380}{5} = 76$

Example 2: Find the mean of the weekly earnings from the following:

Wekly earnings in Rs (X):	10	12	14	16	18	20	22
Number of employees (f):	3	6	10	15	24	42	75

Solution:

x	f	fx
10	3	30
12	6	72
14	10	140
16	15	240
18	24	432
20	42	840
22	75	1050
	175	3404

$$\therefore \text{ Mean } (\bar{x}) = \frac{1}{N} \sum f_i x_i$$

$$= \frac{3404}{175} = 19.45$$

Step Deviation Method:

x	f	$U = \frac{x-16}{2}$	f U
10	3	- 3	- 9
12	6	- 2	- 12
14	10	- 1	- 10
16	15	0	0
18	24	1	24
20	42	2	84
22	75	3	225
	175		302

Here A = 16, h = 2

$$\therefore \text{ Mean } \bar{x} = A + h\,\bar{U}$$

$$= 16 + 2\left(\frac{302}{175}\right)$$

$$= 16 + 3.45 = 19.45$$

Example 3: A distribution of 3 components with frequencies 45, 40, 65 having their means 2, 2.5 and 2 respectively. Findthe combined mean.

Solution: Given that $n_1 = 45, \quad n_2 = 40, \ n_3 = 65$

$$\bar{x}_1 = 2, \quad \bar{x}_2 = 2 \cdot 5, \ \bar{x}_3 = 2$$

$$\therefore \text{ Combined Mean } \bar{x} = \frac{n_1\bar{x}_1 + n_2\bar{x}_2 + n_3\bar{x}_3}{n_1 + n_2 + n_3}$$

$$= \frac{45 \times 2 + 40 \times 2 \cdot 5 + 65 \times 2}{45 + 40 + 65}$$

$$= \frac{320}{150} = 2 \cdot 13$$

Example 4: Calculate the man for the following frequency distribution.

Class interval:	0 - 8	8 - 16	16 - 24	24 - 32	32 - 40	40 - 48
Frequency:	8	7	16	24	15	7

Solution:

C.I.	**Frequecny f**	**Mid Value x**	$U = \frac{x-28}{8}$	f U
0 - 8	8	4	- 3	- 24
8 - 16	7	12	- 2	- 14
16 - 24	16	20	- 1	- 16
24 - 32	24	28	0	0
32 - 40	15	36	1	15
40 - 48	7	44	2	14
	77			-25

Hence $A = 28, \ h = 8$

$$\therefore \text{ Mean } \bar{x} = A + h\,\bar{U}$$

$$= 28 + \left(\frac{-25}{77}\right) = 25 \cdot 404$$

Example 5: Show that the weighted mean of first n natural numbers whose weigths are equal to the corresponding numbers is equal to $\frac{1}{3}(2n+1)$.

Solution:

$$\text{Mean} = \frac{1\times1 + 2\times2 + 3\times3 + \cdots\cdots\cdots + n\times n}{1+2+3+\cdots\cdots\cdots+n}$$

$$= \frac{1^2 + 2^2 + 3^2 + \cdots\cdots\cdots + n^2}{1+2+3+\cdots\cdots+n}$$

$$= \frac{n(n+1)(2n+1)}{6\left[\frac{n(n+1)}{2}\right]}$$

$$= \frac{2(2n+1)}{6} = \frac{1}{3}(2n+1)$$

Example 6: Pre average salary of male employees in a firm was Rs. 520 and that of female was Rs. 420. The mean salary of all the employees was Rs. 500. Find the percentage of male and female employees.

Solution: Let n_1 and n_2 denote respectively the number of male and female employees.

$\bar{x}_1$ and $\bar{x}_2$ their averages respectively.

Let $\bar{x}$ be the average of salary of all the employees.

$\therefore$ given that $\bar{x}_1 = 520$; $\bar{x}_2 = 420$ and $\bar{x} = 500$.

We know that $\bar{x} = \frac{n_1\bar{x}_1 + n_2\bar{x}_2}{n_1 + n_2}$

$$\Rightarrow 500 = \frac{n_1\, 520 + n_2\, 420}{n_1 + n_2}$$

$$\Rightarrow n_1\, 500 + n_2\, 500 = n_1\, 520 + n_2\, 420$$

$$\Rightarrow n_2\,(500 - 420) = n_1\,(520 - 500)$$

$$\Rightarrow 80\, n_2 = 20\, n_1$$

$$\Rightarrow \frac{n_1}{n_2} = \frac{4}{1} \Rightarrow n_1 : n_2 = 4:1$$

Hence the percentage of male employees $= \frac{4}{4+1} \times 100 = 80\ \%$

$\therefore$ Percentage of female employees $= \frac{1}{5} \times 100 = 20\ \%$

Example 7: From the following data find the value of median

35, 49, 225, 50, 30, 65, 40, 55, 52, 76, 48, 325, 47, 32.

Solution: Arrange the data in order the data becomes,

30, 32, 35, 40, 47, 48, 49, 50, 52, 55, 60, 65, 76, 225, 325.

Here $n = 15$

$\therefore$ Median $= \left(\frac{n+1}{2}\right)^{th}$ observation

$= \left(\frac{16}{2}\right)^{th}$ observation

$= 8^{th}$ observation

$= 50$

$\therefore$ Median value is 50.

Example 8: From the following data find the value of median.

Income (in Rs.):	1000	1500	800	2000	2500	1800
No. of Persons:	24	26	16	20	6	30

Solution:

Income (in Rs.):	800	1000	1500	1800	2000	2500
No. of Persons(f):	16	24	26	30	20	6
Cumulative Frequency (c.f.):	16	40	66	96	116	122

$N = 122$

$$\therefore \frac{N+1}{2} = \frac{122+1}{2} = 61.5$$

$$\therefore \text{ Median } = \left(\frac{N+1}{2}\right)^{th} \text{ item } = (61.5)^{th} \text{ item } = 1500$$

Example 9: Calculate median from the following data:

Wage (in Rs.):	0 - 99	100 - 199	200 - 299	300 - 399	400 - 499	500 - 599
No of Persons:	10	18	25	12	8	3

Solution: This is grouped discontinuous data. To convert this into continuous form average of first class upper limit and second class lower limit.

$$\text{i.e. } \frac{99+100}{2} = 99.5$$

i.e. the first class upper limit is 99.5 and second class lower limit is 99.5

$\therefore$ The continuous form is 0 - 99.5, 99.5 - 199.5, ...

Wage (in Rs.)	No. of Persons (f)	Cumulative Frequency (cf)
0 - 99.5	10	10
99.5 - 99.5	18	28
199.5 - 299.5	25	53
299.5 - 399.5	12	65
399.5 - 499.5	8	73
499.5 - 599.5	3	76

N = 76

$$\therefore \frac{N}{2} = \frac{76}{2} = 38$$

i.e. Median class is 199.5 - 299.5

$$\therefore \ell = 199.5,\ f = 25,\ m = 28,\ h = 100$$

$$\therefore \text{Median} = \ell + \frac{\left(\frac{N}{2} - m\right)h}{f}$$

$$= 199.5 + \frac{(38-28)100}{25} = 239.5$$

Example 10: Calculate median value from the following data:

C.I.:	5 - 15	15 - 25	25 - 35	35 - 45	45 - 55	55 - 65
f :	20	5	15	5	35	20

Solution:

C.I.:	5 - 15	15 - 25	25 - 35	35 - 45	45 - 55	55 - 65
f :	20	5	15	5	35	20
C.f.:	20	25	40	45	80	100

Here N = 100

$$\therefore \frac{N}{2} = \frac{100}{2} = 50$$

i.e. Median class is 45 - 55

$$\therefore \ \ell = 45, \ f = 35, \ m = 45, \ h = 10.$$

$$\therefore \text{ Median } = \ell + \frac{\left(\frac{N}{2} - m\right)h}{f}$$

$$= 45 + \frac{(50 - 45)\,10}{35}$$

$$= 45 + \frac{50}{35}$$

$$= 45 + 1 \cdot 43 = 46 \cdot 43$$

Example 11: Calculate mode of the following data

C.I.:	10 - 19	20 - 29	30 - 39	40 - 49	50 - 59	60 - 69	70 - 79	80 - 89	90 - 99
f :	7	10	12	18	10	6	3	2	1

Solution: The give a grouped discontinuaes data can be converted into continuous form is as follows:

C.I.:	9.5 - 19.5	19.5 - 29.5	29.5 - 39.5	39.5 - 49.5	49.5 - 59.5
f :	7	10	12	18	10
C.I.:	59.5 - 69.5	69.5 - 79.5	79.5 - 89.5	89.5 - 99.5	
f :	6	3	2	1	

Highest frequency is 18 i.e. Mode Class is 39.5 - 49.5

$$\therefore \ \ell = 39.5, \ f_1 = 18, \ f_0 = 12, f_2 \ \ 10, \ h = 10 .$$

$$\text{Mode} = \ell + \frac{(f_1 - f_0)\, h}{2f_1 - (f_0 + f_2)}$$

$$= 39 \cdot 5 + \frac{(18 - 12)\, 10}{(2 \times 18 - 12 - 10)} = 43 \cdot 3$$

Example 12: Calculate the mode of the following distribution.

C.I.:	0 - 10	10 - 20	20 - 30	30 - 40	40 - 50
f :	2	6	11	20	40
C.I.:	50 - 60	60 - 70	70 - 80	80 - 90	90 - 100
f :	75	45	25	18	8

Solution: Highest frequency is 75 and it is in the class 50 - 60.

$\therefore\ \ell = 50,\ f_1 = 75,\ f_0 = 40,\ f_2 = 45,\ h = 10$.

$$\text{Mode} = \ell + \frac{(f_1 - f_0)\, h}{2f_1 - (f_0 + f_2)}$$

$$= 50 + \frac{(75 - 40)\, 10}{2 \times 75 - (40 + 45)}$$

$$= 50 + \frac{350}{65} = 55.4 \ (\text{app})$$

1.2 Measures of Dispersion:

1.2.1 Introduction: Averages give us an idea of the concentration of observations about the central part of the distribution. If we know the average alone, we cannot form complete idea about the distribution. To illustrate this consider the following:

Series A:	15	20	25	30	35
Series B:	5	18	25	40	65

These two sets of observations have the same median. However, it may be noticed that observations in series - A are less deviations from average while series - B are large deviations from average (Median). Thus we can say that variability of series - B is more than that of series - A. Hence average alone we cannot study completely the characteristics of data and hence the necessity of measure of dispersion or variation. The various measures of dispersions are (i) Range, (ii) Mean Deviation, (iii) Quartile Deviation, (iv) Standard Deviation.

According to G.V. Yule, a good measure of Dispersion should have the following:

i) It should be rigidly defined.

ii) It should be easy to calculate and easy to understand.

iii) It should be based on all the observations.

iv) It should be readily comprehensible.

v) It should be capable of further algebric treatment.

vi) It should be affected as little as possible fluctuations of sampling.

1.2.3 Range: Range is the difference between the hightest and lowest values in the data. It is very useful in statistical quality control.

If x_ℓ is largest observation and x_s is smallest observation then $\text{Range} = x_\ell - x_s$

This is an absolute measure and it is not suitable to compare two or more data with different units of measurement. To compare two or more situation we use relative measures of dispersion called as coefficient measures. These are pure numbers and independent of units of measurement.

The relative measure of range is called coefficient of range and it is given by

$$\text{Coefficient of Range } = \frac{x_\ell - x_s}{x_\ell + x_s}$$

1.2.4 Quartile Deviation (Q.D.): It is defined as half the difference between the lower and upper quartiles. It is also called as semi-interquartile range. If Q_1 is first quartile and Q_3 is the third quartile the quartile deviation (Q.D.) is given by

$$\text{Q.D.} = \frac{Q_3 - Q_1}{2}$$

$$\text{Coefficient of Q.D.} = \frac{Q_3 - Q_1}{Q_3 + Q_1}$$

1.2.5 Mean Deviation (M.D.): It is sum absolute deviations taken from average divided by number of observations. If $x_1, x_2, \cdots\cdots, x_n$ be te mid values of class intervals with corresponding frequencies $f_1, f_2, \cdots\cdots, f_n$ respectively then mean deviation is given by

$$\text{Mean Deviation} = \frac{1}{N} \sum f_i \left| x_i - A \right|$$

Where A is any average.

If A is mean then M.D. about mean is given by

$$\text{M.D. about mean} = \frac{1}{N} \sum f_i \left| x_i - \bar{x} \right|$$

$$\text{Coefficient of M.D.} = \frac{\text{M.D. Value}}{\text{The Average used in it}}$$

$$\text{i.e. coefficient of M.D. about mean} = \frac{\text{M.D. about mean}}{\text{mean}}$$

1.2.6 Standard Deviation (S.D.): Standard Deviation is the positive square root of the arthemetic mean of squares of deviation from mean.

It is denoted by σ and it is given by

$$\text{S.D.} (\sigma) = \sqrt{\frac{1}{N} \sum f_i \left(x_i - \bar{x} \right)^2}$$

The square of S.D. is known as variance and it is denoted as σ^2.

$$\therefore \text{ Variance } \sigma^2 = \frac{1}{N} \sum f_i \left(x_i - \bar{x} \right)^2$$

$$\text{Coefficient of variation (C.V.)} = \frac{\text{S.D.}}{\text{mean}} \times 100$$

$$= \frac{\sigma}{\bar{x}} \times 100$$

Among the relative measures the coefficient of variation is the most important and is used in almost all cases.

The simplified form of caliculation of variance is variance $\sigma^2 = \frac{1}{N} \sum f_i x_i^2 - \left(\bar{x} \right)^2$

The shortcut method of calculation of variance is given by

$$\text{if } U = \frac{x - a}{h}$$

$$\Rightarrow x = a + h \, \overline{U}$$

$$\Rightarrow \bar{x} = a + h \, \overline{U}$$

$$\therefore\ x_i - \bar{x} = h\left(U_i - \bar{U}\right)$$

$$\Rightarrow \left(x_i - \bar{x}\right)^2 = h^2\left(U_i - \bar{U}\right)^2$$

$$\sigma_x^2 = \frac{1}{N}\sum f_i\left(x_i - \bar{x}\right)^2$$

$$= \frac{1}{N}\sum f_i\ h^2\left(U_i - \bar{U}\right)^2$$

$$= h^2\ \frac{1}{N}\sum f_i\left(U_i - \bar{U}\right)^2$$

$$= h^2\ \sigma_U^2$$

$\therefore$ VAriance is effected by change of origin but is effected by scale.

1.2.7 Workedout Examples:

Example 1: Find the range and coefficient of range for the observations.

10, 8, 5, 10, 9, 14, 7, 4, 20.

Solution: Range $= 20 - 4 = 16$

Coefficient of range $= \dfrac{x_\ell - x_s}{x_\ell + x_s} = \dfrac{20-4}{20+4} = \dfrac{16}{24} = \dfrac{4}{6} = \dfrac{2}{3}$

Example 2: The following are the marks of 80 students of a class. Find the range and coefficiant of range.

Marks:	0 - 10	10 - 20	20 - 30	30 - 40	40 - 50	50 - 60	60 - 70	70 - 80
No. of Students:	4	12	20	18	15	8	2	1

Solution: Range $= x_\ell - x_s = 80 - 0 = 80$

Coefficient of Range $= \dfrac{x_\ell - x_s}{x_\ell + x_s} = \dfrac{80-0}{80+1} = 1$

Example 3: Calculate the quartile deviation and its coefficient from the following data.

Marks: 39, 40, 40, 41, 41, 42, 42, 43, 43, 44, 44, 45.

Solution: Arrange the data in order we get

39, 40, 40, 41, 41, 42, 42, 43, 43, 44, 44, 45

$$Q_1 = \text{size of } \left(\frac{n+1}{4}\right)^{th} \text{ item } = \left(\frac{12+1}{4}\right)^{th} = 3\cdot 25^{th} \text{ item}$$

$$= 3^{rd} \text{ item} + 0\cdot 25\left(4^{th} - 3^{rd}\right)$$

$$= 40 + 0\cdot 25\,(41-40)$$

$$= 40 + 0\cdot 25 = 40\cdot 25$$

$$Q_3 = \text{size of } \frac{3(n+1)^{th}}{4} \text{ item } = \frac{3(12+1)}{4} = 9\cdot 75^{th} \text{ item}$$

$$= 9^{th} \text{ item} + 0\cdot 75\left(10^{th} - 9^{th}\right)$$

$$= 43 + 0\cdot 75\,(44-43)$$

$$= 43 + 0\cdot 75 = 43.75$$

$$\therefore \text{ Q.D.} = \frac{Q_3 - Q_1}{2}$$

$$= \frac{43\cdot 75 - 40\cdot 25}{2} = \frac{3\cdot 5}{2} = 1\cdot 75$$

Coefficient of Q.D. $= \dfrac{Q_3 - Q_1}{Q_3 + Q_1} = \dfrac{43\cdot 75 - 40\cdot 25}{43\cdot 75 + 40\cdot 25}$

$$= \frac{3\cdot 5}{84} = 0\cdot 042$$

Example 4: Compute coefficient of Quartile Deviation from the following data.

Sales in Rs. lakhs:	4 - 8	8 - 12	12 - 16	16 - 20	
No. of Companies:	6	10	18	30	
Sales in Rs. lakhs:	20 - 24	24 - 28	28 - 32	32 - 36	36 - 40
No. of Companies:	15	12	10	6	2

Solution:

Sales in Rs. Lakhs (C.I.)	**No. of Companies (f)**	**c.f.**
4 - 8	6	6
8 - 12	10	16
12 - 16	18	34
16 - 20	30	64
20 - 24	15	79
24 - 28	12	91
28 - 32	10	101
32 - 36	6	107
36 - 40	2	109
	N = 109	

$$\frac{N}{4} = \frac{109}{4} = 27 \cdot 25^{th} \text{ item}$$

i.e. Q_1 is in the class 12 - 16

$$\ell_1 = 12, \; f = 18, \; m = 16, \; h = 4$$

$$Q_1 = \ell + \frac{\left(\frac{N}{4} - m\right)h}{f}$$

$$= 12 + \frac{(27 \cdot 25 - 16)\, 4}{18}$$

$$= 12 \;\; \frac{11 \cdot 25}{18} \times 4$$

$$= 12 + 2 \cdot 5 \; = 14 \cdot 5$$

$$\frac{3N}{4} = 3(27 \cdot 25) = 81 \cdot 75$$

i.e. Q_3 is in the class 24 - 28.

$\ell_3 = 24,\ f = 12,\ m = 79,\ h = 4$

$$Q_3 = \ell_3 + \frac{\left(\frac{3N}{4} - m_3\right)h}{f_3} = 24 + \frac{(81\cdot 75 - 79)\,4}{12} = 24 + \frac{11}{12} = 24\cdot 917$$

$$Q\cdot D\cdot = \frac{Q_3 - Q_1}{2}$$

$$= \frac{24\cdot 917 - 14\cdot 5}{2}$$

$$= \frac{10\cdot 417}{2} = 5\cdot 2085$$

Coefficient of $Q\cdot D\cdot = \frac{Q_3 - Q_1}{Q_3 + Q_1} = \frac{24\cdot 917 - 14\cdot 5}{24\cdot 917 + 14\cdot 5}$

$$= \frac{10\cdot 417}{39\cdot 417} = 0\cdot 264$$

Example 5: Calculate mean deviation from mean for the following data 5, 10, 15, 20, 25, 30, 35.

Solution: $\text{Mean} = \frac{1}{n}\sum x_i = \frac{5+10+15+20+25+30+35}{7} = \frac{140}{7} = 20$

x	$x - \bar{x}$	$\lvert x - \bar{x} \rvert$
5	- 15	15
10	- 10	10
15	- 5	5
20	0	0
25	5	5
30	10	10
35	15	15
		60

$\therefore$ Mean Deviation about mean $= \frac{1}{N} \Sigma \left| x_i - \bar{x} \right|$

$$= \frac{60}{7} = 8 \cdot 57$$

Example 6: Calculate mean deviation from mean to the following data:

x :	2	4	6	8	10
f :	1	4	6	4	1

Solution:

x	f	f x	$\left\lvert x_i - \bar{x} \right\rvert$	$f_i \left\lvert x_i - \bar{x} \right\rvert$
2	1	2	4	4
4	4	16	2	8
6	6	36	0	0
8	4	32	2	8
10	1	10	4	4
	16	96		24

mean $\bar{x} = \frac{1}{N} \sum f_i x_i = \frac{96}{16} = 5$

Mean Deviation $= \frac{1}{N} \sum f_i \left| x_i - \bar{x} \right|$

$$= \frac{24}{16} = 1 \cdot 5$$

Example 7: From the following data calculate mean deviation and its coefficient.

C.I. :	0 - 5	5 - 10	10 - 15	15 - 20	20 - 25	25 - 30
f :	5	8	10	7	6	4

Solution:

C.I.	f	mid x	c.f.	$\lvert x_i - med \rvert$	$f_i \lvert x_i - md \rvert$
0 - 5	5	2.5	5	11	55
5 - 10	8	7.5	13	6	48
10 - 15	10	12.5	23	1	10
15 - 20	7	17.5	30	4	28
20 - 25	6	22.5	36	9	54
25 - 30	4	27.5	40	14	56
	40				251

Median $= \ell + \frac{\left(\frac{N}{2} - m\right) h}{f}$

$$= 10 + \frac{(20 - 13)\, 5}{10} = 13 \cdot 5$$

M.D. about median $= \frac{1}{N} \sum f_i \lvert x_i - md \rvert$

$$= \frac{251}{40} = 6 \cdot 275$$

Coefficient of M.D. $= \frac{\text{M.D. Value}}{\text{Median}}$

$$= \frac{6 \cdot 275}{13 \cdot 5} = 0 \cdot 465$$

Example 8: Ten measurements were made with the following results:

Length in Cms : 77, 73, 75, 70, 72, 76, 75, 72, 74, 76.

Find the standard deviation.

Solution:

x	$d = x - 75$	d^2
77	2	4
73	- 2	4
75	0	0
70	- 5	25
72	- 3	9
76	1	1
75	0	0
74	- 1	1
76	1	1
	- 10	54

$$\therefore \ S \cdot D \cdot \ \sigma = \sqrt{\frac{1}{M}\sum d_i^2 - \left(\frac{1}{M}\sum d_i\right)^2}$$

$$= \sqrt{\frac{54}{10} - \left(\frac{-10}{10}\right)^2}$$

$$= \sqrt{5 \cdot 4 - 1}$$

$$= \sqrt{4 \cdot 4} \ = 2 \cdot 09$$

Example 9: Calculate mean and standard deviation to the following data also C.V.

Marks :	10	20	30	40	50	60
No. of Students :	8	12	20	10	7	3

Solution:

x	f	$U = \frac{x-30}{10}$	f U	$f U^2$
10	8	- 2	- 16	32
20	12	- 1	- 12	12
30	20	0	0	0
40	10	1	10	10
50	7	2	14	28
60	3	3	9	27
	60		5	109

$$\overline{U} = \frac{1}{N} \sum f_i U_i$$

$$= \frac{5}{60} = 0 \cdot 083$$

$$\sigma_U^L = \frac{1}{N} \sum f U^2 - \left(\overline{U}\right)^2$$

$$= \frac{109}{60} - \left(\frac{5}{60}\right)^2$$

$$= 1 \cdot 817 - 0 \cdot 0069$$

$$= 1 \cdot 81$$

$\therefore$ Mean $\overline{x} = a + h\,\overline{U}$

$$= 30 + 10\,(0 \cdot 083) = 30 \cdot 83$$

Variance $\sigma^2 = h^2 \;\sigma_U^2$

$$= (10)^2 \;(1 \cdot 81)$$

$$= 100 \times 1 \cdot 81 = 181$$

$\therefore$ S·D· $\sigma = \sqrt{181}$

$= 13 \cdot 5$

Coefficient of variation C.V. $= \frac{\sigma}{\bar{x}} \times 100$

$= \frac{13 \cdot 5}{30 \cdot 83} \times 100$

$= 43 \cdot 7885$

Example 10: Calculate mean and standard deviation for the data given below.

Marks:	0 - 10	10 - 20	20 - 30	30 - 40	40 - 50
No. of students:	7	12	24	10	7

Solution:

C.I.	f	Mid x	$U = \frac{x-25}{10}$	f U	f U^2
0 - 10	7	5	- 2	- 14	28
10 - 20	12	15	- 1	- 12	12
20 - 30	24	25	0	0	0
30 - 40	10	35	1	10	10
40 - 50	7	45	2	14	28
	60			- 2	78

$\overline{U} = \frac{1}{N} \sum f\,U$

$= \frac{-2}{60}$

A = 25, h = 10

$\therefore$ Mean $\bar{x} = a + h\,\overline{U}$

$= 25 + 1\not{0}\left(\frac{-2}{6\not{0}}\right)$

$= 25 - 0 \cdot 33 = 24 \cdot 67$

$$S \cdot D \cdot \quad \sigma = h\sqrt{\frac{1}{N}\Sigma f\,U^2 - (\overline{U})^2}$$

$$= 10\sqrt{\frac{78}{60} - \left(\frac{-2}{60}\right)^2}$$

$$= 10\sqrt{1 \cdot 3 - 0 \cdot 001}$$

$$= 10 \times 1 \cdot 1397 = 11 \cdot 397$$

1.3 Exercise:

1. Define various measures of central tendencies.
2. Define various measures of dispersions.
3. Calculate median from the following data :

   70, 60, 75, 90, 65, 80, 42, 65, 72.
4. The mean of marks in statistics of 100 students of a class was 72. The mean of marks of boys was 75 while their number was 70. Find out the mean marks of girls in the class.
5. Calculate the mean marks from the following data by direct method and step deviation method.

Marks :	0 - 10	10 - 20	20 - 30	30 - 40	40 - 50	50 - 60	60 -70
No. of students :	5	12	15	25	8	3	2

6. Compute median from the following data:

C.I. :	10 - 19	20 - 29	30 - 39	40 - 49	50 - 59	60 - 69
f :	12	19	31	27	16	8

7. Calculate mode value from the following data:

Age in years:	55 - 60	50 - 55	45 - 50	40 - 45	35 - 40	30 - 35	25 - 30	20 - 25
No. of persons:	8	12	30	40	20	8	7	2

8. The following are the wages of 8 workers of a factory. Findthe range and coefficient of range.

   Wages in Rs.: 1400, 1450, 1520, 1380, 1485, 1495, 1575, 1440.
9. Calculate the appropriate measure of dispersion from the following data:

Wages in Rs. :	Below 35	35 - 37	38 - 40	41 - 43	Over 43
No. of wage earners:	14	60	95	24	7

10. Calculate mean and standard deviation to the following data:

    600, 620, 640, 620, 680, 670, 680, 640, 700, 650.

11. TWo cricketers scored the following runs in the several innings. Find who is a better run - getter and who is more consistent players.

    Player A : 42, 17, 83, 59, 72, 76, 64, 45, 40, 32.

    Player B : 28, 70, 31, 0, 59, 108, 82, 14, 3, 95.

12. Coefficient of variation of two series are 58 % and 69 % their standard deviations are 21.2 and 15.6 respectively. What are their means.

13. Define standard deviation and C.V. line the step deviation formula for S.D.

14. Distinquish between mean and median.

15. Distinquish between mean and mode.

1.4 Answers:

3. Median = 70

4. Number of girls are 65

5. Mean = 30.14

6. Median = 37.73

7. Mode = 43.33

8. Range = Rs. 195

   Coefficient of Range = 0.066

9. Q.D. = 1.8

10. Mean = 650

    S.D. = 30.33

11. C.V. for A = 37.92

    C.V. for B = 75.6

    $\therefore$ A is more consistent

12. Means are 36.55 and 22.6

Lesson – 2

MOMENTS

Syllabus :

Importance of moments, Central and non central momets shappard's corrections for moments for grouped data. Skewness and Kurtosis - their measures including those based on quartile and moments with real life examples.

Structure of the Lesson :

Object :

After studying the lesson the students are expected to have clear comprehension of the theory and practical utility about the concepts of non-central moments, central moments, skewness, kurtosis with real life examples.

2.1 Introduction :

To form an idea about the nature of the distribution averages and dispersions are not enough to give clear idea. To study the pattern of distribution there are other comparable characteristics also known as symmetry and peakedness of which the former is known as skewness and the latter as kurtsis. To study these measures first we have to study the idea of moments.

2.2 Definition :

The moments are of three types, they are

(i) Raw moments, (ii) Moments ablut any point, (iii) Central moments.

(i) Raw Moments : If $x_1, x_2, ..., x_n$ are mid values of class interval with corresponding frequencies $f_1, f_2, ..., f_n$ respectively then r^{th} moment about orgin or r^{th} raw moment is denoted as μ_r^1 and it is given by

$$\mu_r^1 = \frac{1}{N}\sum_{i=1}^{n} f_i x_i^r \quad ; \; r = 0, 1, 2, ...$$

inarticular if r = 1, $\mu_1^1 = \frac{1}{N}\sum_{i=1}^{n} f_i X_i = \bar{x}$

Therefore first moment about orgin is mean.

(ii) Non-Central Moments (or) Moments about any Point : If $x_1, x_2, ..., x_n$ be the mid values of class interval with corresponding frequencies $f_1, f_2, ...f_n$ respectively then the r^{th} moment about a point A is denoted as μ_r^1 and it is defined as

$$\mu_r^1 = \frac{1}{N}\sum_{i=1}^{n} f_i (x_i - A)^r$$

$$\mu_r^1 = \frac{1}{N}\sum_{i=1}^{n} f_i (x_i - A)^r$$

where $xN = \sum_{i=1}^{n} f_i$

It is also called as r^{th} non-central moment.

If $d_i = x_i - A$ then $\mu_r^1 = \frac{1}{N}\sum f_i d_i^r$

In particular $\mu_0^1 = \frac{1}{N}\sum f_i (x_i - A)^0 = \frac{1}{N}\sum f_i = 1$

$$\mu_1^1 = \frac{1}{N}\sum f_i (x_i - A)$$

$$= \frac{1}{N}\sum f_i x_i - A\frac{1}{N}\sum f_i$$

$$= \bar{x} - A \qquad \because \sum f_i = N$$

$\therefore$ Mean $\bar{x} = A + \mu_1^1$

(iii) Cental Moments : If $x_1, x_2, \ldots\ldots\ldots, x_n$ be the mid values of class interval with corresponding frequencies $f_1, f_2, \ldots\ldots f_n$ respectively then r^{th} moment about mean $\bar{x}$ is denoted as μ_r and it is given by

$$\mu_r = \frac{1}{N}\sum t_i\left(x_i - \bar{x}\right)^r \qquad r = 0,1,2,\ldots\ldots\ldots$$

$$r = \frac{1}{N}\sum t_i Z_i^r \qquad \text{where } z_i = x_i - \bar{x}$$

It is also called as r^{th} central moment

Imparticular $\mu_o = \frac{1}{N}\sum f_i\left(x_i - \bar{x}\right)^0 = \frac{1}{N}\sum t_i = 1$

$$\mu_1 = \frac{1}{N}\sum f_i\left(x_i - \bar{x}\right)$$

$$= \frac{1}{N}\sum f_i x_i - \bar{x}\cdot\frac{1}{N}\sum t_i$$

$$= \bar{x} - \bar{x} \qquad \because \sum t_i = N$$

$= 0$

$$\mu_2 = \frac{1}{N}\sum f_i\left(x_i - \bar{x}\right)^2 = \text{Variance}$$

2.3 Expression of Central Moments in Terms of Non-Central Moments:

We have, by definition of central moments,

$$\mu_r = \frac{1}{N}\sum t_i\left(x_i - \bar{x}\right)^r$$

$$= \frac{1}{N}\sum f_i\left(x_i - A + A - \bar{x}\right)^r$$ where "A" is a constant

$$= \frac{1}{N}\sum f_i\left[(x_i - A) - (\bar{x} - A)\right]^r$$

$$= \frac{1}{N}\sum f_i\left(d_i - \mu_1^1\right)^r \qquad \because d_i = x_i - A,\quad \mu_1^1 = \bar{x} - A$$

$$\mu_r = \frac{1}{N}\sum f_i\left(d_i - \mu_1^1\right)^r$$

$$= \frac{1}{N}\sum f_i\left[d_i^r - rc_1 d_i^{r-1}\mu_1^1 + rc_2 d_i^{r-2}\left(\mu_1^1\right)^2 - rc_3 d_i^{r-3}\left(\mu_1^1\right)^3 + \ldots\ldots\ldots\ldots + (-1)^r\left(\mu_1^1\right)^r\right]$$

$$= \frac{1}{N}\sum f_i d_i^r - rc_1\left(\frac{1}{N}\sum f_i d_i^{r-1}\right)\mu_1^1 + \ldots\ldots\ldots\ldots + (-1)^r\left(\mu_1^1\right)^r$$

$$= \mu_r^1 - rc_1\mu_{r-1}^1\mu_1^1 + rc_2\mu_{r-2}^1\left(\mu_1^1\right)^2 - rc_3\mu_{r-3}^1\left(\mu_1^1\right)^3 + \ldots\ldots\ldots + (-1)^r\left(\mu_1^1\right)^r$$

Hence the expression of central moments interms non-central moments is

$$= \mu_r^1 - rc_1\mu_{r-1}^1\mu_1^1 + rc_2\mu_{r-2}^1\left(\mu_1^1\right)^2 - rc_3\mu_{r-3}^1\left(\mu_1^1\right)^3 + \ldots\ldots\ldots + (-1)^r\left(\mu_1^1\right)^r$$

2.4 Expression of Non - Central Moments in Terms of Central Moments:

By the definition of non-central moments, we have

$$\mu_r^1 = \frac{1}{N}\sum f_i\left(x_i - A\right)^r$$

$$= \frac{1}{N}\sum f_i\left[x_i - \bar{x} + \bar{x} - A\right]^r$$

$$= \frac{1}{N}\sum f_i\left[z_i + \mu_1^1\right]^r;$$ where $z_i = x_i - \bar{x}$

$$\mu_1^1 = \bar{x} - A$$

$$= \frac{1}{N}\sum f_i\left[z_i^r + rc_1 z_i^{r-1}\mu_1^1 + rc_2 z_i^{r-2}\left(\mu_1^1\right)^2 + \ldots\ldots\ldots + \left(\mu_1^1\right)^r\right]$$

$$= \frac{1}{N}\sum f_i z_i^r + rc_1\left(\frac{1}{N}\sum f_i z_i^{r-1}\right)\mu_1^1 + rc_2\left(\frac{1}{N}\sum f_i z_i^{r-2}\right)\left(\mu_1^1\right)^2 + \ldots\ldots\ldots\ldots + \left(\mu_1^1\right)^r$$

$$= \mu_r + rc_1\mu_{r-1}\left(\mu_1^1\right) + rc_2\mu_{r-2}\left(\mu_1^1\right)^2 + \ldots\ldots\ldots + \left(\mu_1^1\right)^r \quad \ldots\ldots\ldots\ldots\ldots\ldots(2.4.1)$$

Summary :

Subsituting r = 0,1,2,3,4 in (2.3.1) and (2.4.1) we get

$\mu_0^1 = 1$	$\mu_0 = 1$
$\mu_1^1 = \bar{x} - A$	$\mu_1 = 0$
$\mu_2^1 = \mu_2 + \left(\mu_1^1\right)^2$	$\mu_2 = \mu_2^1 - \left(\mu_1^1\right)^2$
$\mu_3^1 = \mu_3 + 3\mu_2\mu_1^1 + \left(\mu_1^1\right)^3$	$\mu_3 = \mu_3^1 - 3\mu_2^1\mu_1^1 + 2\left(\mu_i\right)^3$
$\mu_4^1 = \mu_4 + 4\mu_3\mu_1^1 + 6\mu_2\left(\mu_1^1\right)^2 + \left(\mu_1^1\right)^4$	$\mu_4 = \mu_4^1 - 4\mu_3^1\mu_1^1 + 6\mu_2^1\left(\mu_1^1\right)^2 - 3\left(\mu_1^1\right)^4$

2.5 Effect of Change of Origin and Scale on Moments:

Theorem (2.5.1) : Cental moments are independent of change of origin but not scale.

Proof : Let $U_i = \frac{x_i - A}{h}$ so that

$$x_i = A + hU_i$$

$$\bar{x} = A + h\bar{U}$$

$$\therefore\ x_i - \bar{x} = h\left(U_i - \bar{U}\right)$$

Thus $\mu_r = \frac{1}{N}\sum f_i\left(x_i - \bar{x}\right)^r$

$$= \frac{1}{N}\sum f_i\left[h\left(U_i - \bar{U}\right)\right]^r$$

$$= h^r \frac{1}{N} \sum f_i \left(U_i - \overline{U}\right)^r$$

$$= h^r \cdot \mu_r \text{ of } U$$

$\therefore r^{th}$ central moment of X - distribution = h^r times r^{th} central moment of U - distibution.

Theorem (2.5.2) : non - Central moments are independent of change of origin but not scale.

Proof : Let $U_i = \frac{x_i - A}{h}$ so that

$$x_i = A + hU_i$$

$$\Rightarrow x_i - A = hU_i$$

Thus $\mu_r^1 = \frac{1}{N} \sum f_i \left(x_i - A\right)^r$

$$= \frac{1}{N} \sum f_i \left(hU_i\right)^r$$

$$= h^r \frac{1}{N} \sum f_i U_i^r$$

$$= h^r \cdot \mu_r^1 \text{ of U about origin.}$$

Thus r^{th} non - central moment of distribution of $x = h^r$ times r^{th} moment about origin of U.

Hence Non - Central moments are independent of change of origin but not scale.

2.4 Shappard's Corrections for moments :

In case of frequency distibution, we assume that the frequencies are concentrated at mid - points of class intervals. If the distribution is symmetrical or slightly symmetrical and the class intervals are not greater than $\left(\frac{1}{20}\right)^{th}$ of the range of the distribution then the assumption is true.

This presumption is likely to gives rise to some error in the values of the moments and is called the "grouping error". W.F. shappard corrected the effect due to grouping in the mid point of the intervals by the following formulas known as Shappard's corrections. He proved that the correction is made for, if (i) the frequency distribution is continuous (ii) the frequency tails off to zero in both directions. The corrected moments are

$$\mu_2\,(\text{Corrected}) = \mu_2 - \frac{h^2}{12}$$

$$\mu_3\,(\text{Corrected}) = \mu_3$$

$$\mu_4\,(\text{Corrected}) = \mu_4 - \frac{1}{2}h^2\mu_2 + \frac{7}{240}h^4$$

Where h is the width of the class interval.

Note : Correction is not necessary for odd moments, because in these cases the algebric signs of the deviations +, - remains as they are. Hence, the error is neutralized because of its compensatory nature.

Pearson's β and γ Coefficients :

Karl Pearson defined the following four coefficients, based upon the first four central moments they are

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} \qquad \gamma_1 = \pm\sqrt{\beta_1}$$

$$\beta_2 = \frac{\mu_4}{\mu_2^2} \qquad \gamma_2 = \beta_2 - 3$$

The sign of γ_1 will be that of μ_3.

β_1 and γ_1 are the measures of skewness.

β_2 and γ_2 are measures of kurtsis.

2.7 Measures of Skewness :

2.7.1 Symmetrical distribution : A distribution is said to be symmetrical when (i) the frequencies are symmetrically distributed about the mean. (ii) For symmetrical distribution the mean, mode and median coincide. (iii) Median lies half way between the two quantites. i.e. Q_3 - median = median - Q_1

Ex: The following distribution is symmetical about its mean 5

x	1	2	3	4	5	6	7	8	9
f	3	4	6	9	10	9	6	4	3

2.7.2 Skewness :

Skewness means "lack of symmetry". We study skewness to have an idea about the shape of the curve which we can draw with the help of the given data. The distribution is said to be skewed if

(i) Mean, Median and Mode fall at different points.

i.e. Mean $\neq$ Median $\neq$ Mode.

(ii) Quantites are not equidistant from median.

i.e. Q_3 - Median $\neq$ Medium - Q_1

(iii) The curve drawn with the help of the given data is not symmetrical but bent more to one side than to the other and more points lie on that side.

Skewness is said to be positive, if the curve of the distribution has more longer tail on right side. the skewness in said to be negative, if the curve of the distribution has more longer tail on left side. The following are the figures.

2.7.3 Coefficient of skewness :

The various measures of coefficient of skewness are

(i) Karl Pearson's coefficient of skewness

(ii) Bowley's coefficient of skewness

(iii) Coefficient of skewness based on moments.

(i) Karl Pearson's Coefficient of Skewness :

The Karl Pearson's coefficient of skewness is given by the formula.

Coefficient of skewness = sk = $\frac{\text{Mean} - \text{Mode}}{\text{Skandard Deviation}}$

$$= \frac{m - m_o}{\sigma}$$

Some times mode is diffecult to obtain but median is always easy to locate. If the mode is ill - defined then we may use the relation.

Mean - Mode = 3(mean - median)

In this case, Karl Pearson's coefficient of skewness become.

Coefficient of skewness = $\frac{3(\text{Mean} - \text{Median})}{\text{S tan dard Deviation}}$

The limits of eocfficient of skewness based on mode are ± 1 and based on median are ± 3.

Skewness is positive of Mean > Mode or Mean > Median.

Skewness is negative if mean < mode or mean < median.

(ii) Bowley's coefficient of skewness :

Bowley's coefficient of skewness is based on quartiles and it is given by the formula.

$$\text{coefficient of skewness} = \frac{(Q_3 - \text{Median}) - (\text{Median} - Q_1)}{(Q_3 - \text{Median}) + (\text{Median} - Q_1)}$$

$$= \frac{Q_3 + Q_1 - 2\,\text{Median}}{Q_3 - Q_1}$$

It is also known as Quantile coefficent of skewness. Where Q_3 is third quartile and Q_1 is first quartile. It lies between -1 and +1.

If coefficient of skewness > 0 then the distribution is positively skewed. If coefficent of skewness < 0 then the distribution is negatively skewed.

(iii) Coefficient of Skewness based on Moments :

The coefficient of skewness based on moments is given by

$$\text{Coefficinet of skewness} = \frac{\sqrt{\beta_1}(\beta_2 + 3)}{2(5\beta_2 - 6\beta_1 - 9)}$$

Coefficient of skewness is zero if either $\beta_1 = 0$ or $\beta_2 = -3$. But $\beta_2 \neq -3$ since $\beta_2 = \mu_4 / \mu_2^2$ can not be negative. Hence coefficent of skewness = 0 if and only at $\beta_1 = 0$.

Thus for symmetrical distribution $\beta_1 = 0$

To study the distribution is skewed or not it is convenient to study by using γ_1 where $\gamma_1 = \pm\sqrt{\beta_1} = \frac{\mu_3}{\mu_2^{3/2}}$.

If $\gamma_1 > 0$ then the distribution is positively skewed.

If $\gamma_1 < 0$ then the distribution is negatively skewed.

If $\gamma_1 = 0$ then the distributionis symmetrical.

Theorem (2.7.3.1) : Show that Bowley's coefficient of skewness lies between -1 and +1.

Proof : We know that for any two real positive numbers a and b (a, b > 0).

$$|a-b| \le |a+b|$$

$$\Rightarrow \left|\frac{a-b}{a+b}\right| \le 1$$

We know that Q_3 - Median and Median - Q_1 are both non - negative.

Thus if a = Q_3 - Median, b = Median - Q_1 then we get

$$\left|\frac{(Q_3-\text{Median})-(\text{Median}-Q_1)}{(Q_3-\text{Median})+(\text{Median}-Q_1)}\right| \le 1$$

$$\Rightarrow |\text{Bowley's Coefficient of sk}| \le 1$$

$$\Rightarrow -1 \le \text{Bowley's Coefficient of sk} \le 1$$

Hence Bowley's coefficient of skewness is always lies -1 and +1.

2.8 Kurtosis:

The skewness was mainly concerned with the identification of the right and left tails of distribution. In addition to this measure karl pearson gave another measure called "Convexity of a Curve" or Kurtosis. Kurtosis enables us to have an idea about the flatness or peakedness of the Curve. Kurtosis is measured by coefficent β_2 or γ_2 and given by

$\beta_2 = \frac{\mu_4}{\mu_2^2}$ and $\gamma_2 = \beta_2 - 3$.

The curve which is neither flat nor peaked is called the normal curve or mesokurtic curve and for that curve $\beta_2 = 3$ or $\gamma_2 = 0$.

The curve which is flatter than the normal curve is known as platykurtic and for such a curve $\beta_2 < 3$ or $\gamma_2 < 0$.

The curve which is more peaked than the normal curve is called leptokurtic and for such a curve $\beta_2 > 3$ or $\gamma_2 > 0$.

2.9 Merits and Demerits of Moments:

Merits : moments are used to study the general nature of the frequency distribution. We compare one distribution with another distribution with respect to those characteristics. In fact, mean, variance, skewness and kurtosis etc are nothing but moments. Thus for analysing a statistical data moments are also used in describing the shape and location of a frequency distribution. These are useful in fitting of distributions.

Demerits :

1. The higher themoments the larger error would be subjected. Hence the use of higher order moments are avoided in practice.
2. In studying a distribution through moments, we compare it .with normal distribution. So this method of comparision becomes less effective when a distribution is very much away from the normal conditions.
3. In symmetrical distributions all the odd order moments are zero. Thus, our scope of knowledge is reduced to half.
4. In some theoritical distributions some of moments do not exist.

   Ex: Mean does not exist in Cavchy distribution.

Theorem 2.8.1 : Show that for discrete distribution $\beta_2 > 1$.

Proof : If $x_1, x_2, ..., x_n$ be the mid values of class interval with conresponding frequencing $f_1, f_2, ..., f_n$ respectively.

by definition $\beta_2 = \frac{\mu_4}{\mu_2^2}$

where $\mu_2 = \frac{1}{N} \Sigma f_i \left(x_i - \bar{x}\right)^2$

$$\mu_4 = \frac{1}{N} \Sigma f_i \left(x_i - \bar{x}\right)^4$$

We have to prove that $\beta_2 \geq 1$

i.e. $\frac{\mu_4}{\mu_2^2} \geq 1$

i.e. $\mu_4 \geq \mu_2^2$

i.e. $\left[\frac{1}{N}\sum f_i\left(x_i-\bar{x}\right)^4\right] \geq \left[\frac{1}{N}\sum f_i\left(x_i-\bar{x}\right)^2\right]^2$

if $\left(x_i-\bar{x}\right)^2 = Zi$ then we have to prove

that $\frac{1}{N}\sum f_i z_i^2 \geq \left(\frac{1}{N}\sum f_i z_i\right)^2$

i.e. $\frac{1}{N}\sum f_i z_i^2 - \left(\frac{1}{N}\sum f_i z_i\right)^2 \geq 0$

i.e. $\frac{1}{N}\sum f_i\left(z_i-\bar{z}\right)^2 \geq 0$

Which is true since square of the quantity is non - negative. Hence $\beta_2 \geq 1$

Theorem 2.8.2 : Show that $\beta_2 \geq \beta_1$

Proof : If $x_1, x_2, ..., x_n$ be the mid values of class interval with corresponding frequencives $f_1, f_2, ..., f_n$ respectively.

By definition $\mu_2 = \frac{1}{N}\sum f_i\left(x_i-\bar{x}\right)^2$ $\mu_3 = \frac{1}{N}\sum f_i\left(x_\propto-\bar{x}\right)^3$ $\mu_4 = \frac{1}{N}\sum f_i\left(x_i-\bar{x}\right)^4$

Let $y = x_i - \bar{x}$ then $\mu_2 = \frac{1}{N}\sum f_i y_i^2$ $\mu_3 = \frac{1}{N}\sum f_i y_i^3$ $\mu_4 = \frac{1}{N}\sum f_i y_i^4$

We know that square of the quantity is non negative

i.e. $\frac{1}{N}\sum f_i\left[y_i^2+ty_i\right]^2 \geq 0$

$\Rightarrow \frac{1}{N}\sum f_i y_i^4 + \frac{1}{N}t^2\sum y_i^2 + 2t\frac{1}{N}\sum f_i y_i^3 \geq 0$

$\Rightarrow \mu_4 + t^2\mu_2 + 2t\mu_3 \geq 0$

$\therefore \left(2\mu_3\right)^2 - 4\cdot\mu_2\cdot\mu_4 \leq 0$ $\therefore at^2 + bt + c \geq 0 \text{ then } b^2 - 4ac \leq 0$

$$\Rightarrow 4\mu_3^2 - 4\mu_2\ \mu_4 \le 0$$

$$\Rightarrow \mu_3^2 \le \mu_2\,\mu_4$$

$$\Rightarrow \mu_2\mu_4 \ \ge \mu_3^2$$

$$\Rightarrow \frac{\mu_2\mu_4}{\mu_2^3} \ge \frac{\mu_3^2}{\mu_2^3}$$

$$\Rightarrow \frac{\mu_4}{\mu_2^2} \ \ge \frac{\mu_3^2}{\mu_2^3}$$

$$\Rightarrow \beta_2 \ge \beta_1$$

2.10 Examples:

Ex 1 : The first three moments of a distribution about value 2 are 1,16 and -40. Show that the variable mean = 3, variance 15 and $\mu_3 = -86$. Also show that the first three moments about zero are 3, 24 and 76.

Sol : Given the momnets about the point A = 2. They are

$$\mu_1^1 = 1,\ \mu_2^1 = 16,\quad \mu_3^1 = -40$$

$\therefore$ Mean $\bar{x} = A + \mu_1^1 = 2 + 1 = 3$

Variance $\mu_2 = \mu_2^1 - \left(\mu_1^1\right)^2$

= 16 - 1 = 15

$$\mu_3 = \mu_3^1 - 3\mu_2^1\left(\mu_1^1\right) + 2\left(\mu_1^1\right)^3$$

$$= -40 - 3(16)(1) + 2(1)^3 = -86$$

The first moment about zero (origin) $\mu_1^1 = \bar{x} = 3$

$$\mu_2^1 = \mu_2 + \left(\mu_1^1\right)^2 = 15 + (3)^2 = 24$$

$$\mu_3^1 = \mu_3 + 3\mu_2\left(\mu_1^1\right) + \left(\mu_1^1\right)^3$$

$$= -86 + 3(15)(3) + (3)^3 = 76$$

Ex 2 : The first four moments of a distribution about the value 5 of the variable are 2, 20, 40 and 50. Find mean, variance, μ_3 and μ_4 values.

Sol : Given that A = 5, $\mu_1^1 = 2$, $\mu_2^1 = 20$, $\mu_3^1 = 40$, $\mu_4^1 = 50$

We know that $\bar{x} = A + \left(\mu_1^1\right)$

$$= 5 + 2 = 7$$

$\therefore$ Mean = 7.

Moments about the mean are

$$\mu_2 = \mu_2^1 - \left(\mu_1^1\right)^2$$

$$= 20 - (2)^2 = 16$$

$$\mu_3 = \mu_3^1 - 3\mu_2^1\mu_1^1 + 2\left(\mu_1^1\right)^3$$

$$= 40 - 3(20)(2) + 2(2)^3$$

= - 64

$$\mu_4 = \mu_4^1 - 4\mu_3^1\mu_1^1 + 6\mu_2^1\left(\mu_1^1\right)^{21} - 3\left(\mu_1^1\right)^4$$

$$= 50 - 4(40)(2) + 6(20)(2)^2 - 3(2)^4$$

= 162.

Ex 3 : For a distribution the mean 10, variance is 16, $v_1 = 1$ and β_2 is 4. Find the first four moments about origin.

Sol : Given that $\bar{x} = 10$, $\sigma^2 = \mu_2 = 16$

$$v_1 = \sqrt{\beta_1} = 1$$

$\Rightarrow \beta_1 = 1$

$\Rightarrow \frac{\mu_3^2}{\mu_2^3} = 1$

$\Rightarrow \mu_3^2 = \mu_2^3 = (16)^3 = 4096$

$\therefore \mu_3 = \sqrt{4096} = 64$

Given that $\beta_2 = 4$

$\Rightarrow \frac{\mu_4}{\mu_2^2} = 4$

$\Rightarrow \mu_4 = 4\mu_2^2$

$= 4(16)^2$

= 1024

$\therefore \ \overline{x} = 10, \ \mu_2 = 16, \mu_3 = 64, \mu_4 = 1024$

Moments about orgin are

$\mu_1^1 = \overline{x} = 10$

$\mu_2^1 = \mu_2 + \left(\mu_1^1\right)^2 = 16 + (10)^2 = 116$

$\mu_3^1 = \mu_3 + 3\mu_2\mu_1^1 + \left(\mu_1^1\right)^3$

$= 64 + 3(16)(10) + (10)^3$

= 1544

$\mu_4^1 = \mu_4 + 4\mu_3\mu_1^1 + 6\mu_2\left(\mu_1^1\right)^2 + \left(\mu_1^1\right)^4$

$= 1024 + 4(64)10 + 6(16)(10)^2 + (10)^4$

= 23784

Ex 4 : The first four moments of a distribution about the point 7 are 2, 8, 11 and 15 respectively. Obtain β_1 and β_2 coefficients and comment on the nature of the distribution.

Sol : Given that

A = 7, $\mu_1^1 = 2,\ \mu_2^1 = 8,\ \mu_3^1 = 11,\ \mu_4^1 = 15$

$$\therefore\ \mu_2 = \mu_2^1 - \left(\mu_1^1\right)^2 = 8 - (2)^2 = 4$$

$$\mu_3 = \mu_3^1 - 3\mu_2^1\mu_1^1 + 2\left(\mu_1^1\right)^3$$

$$= 11 - 3(8)^2 + 2(2)^3$$

= 21

$$\mu_4 = \mu_4^1 - 4\mu_3^1\mu_1^1 + 6\mu_2^1\left(\mu_1^1\right)^2 - 3\left(\mu_1^1\right)^4$$

$$= 15 - 4(11)_2 + 6(8)(2)^2 - 3(2)^4$$

= 15 - 88 + 192 - 48

= 71

$$\therefore\ \beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{(21)^2}{64} = 6.89$$

$$\therefore\ \gamma_1 = \frac{\mu_3}{\mu_2^2} = \frac{21}{4^{3/2}} = \frac{21}{\sqrt{64}} = \frac{21}{8} = 2.625$$

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{71}{(4)^2} = \frac{71}{16} = 4.4375$$

Comment : Since $\gamma_1 > 0$, the distribution is postively skewed and since $\beta_2 > 3$, the curve of the distribution is leptokurtic.

Ex 5 : If the first four moments of a distribution about the value 5 are equal to -4, 22, -117 and 560. Determine the corresponding moments (i) about mean, (ii) about origin.

Sol : Given that A = 5, $\mu_1^1 = -4$, $\mu_2^1 = 22$, $\mu_3^1 = -117$, $\mu_4^1 = 560$

Mean $\left(\overline{x}\right) = A + \mu_1^1$ = 5 - 4 = 1

(i) Moment about mean

$$\mu_2 = \mu_2^1 - \left(\mu_1^1\right)^2 = 22 - 16 = 6$$

$$\mu_3 = \mu_3^1 - 3\mu_2^1\,\mu_1^1 + 2\left(\mu_1^1\right)^3$$

$$= -117 - 3(22)(-4) + 2(-4)^3$$

$$= -117 + 264 - 128 = 19$$

$$\mu_4 = \mu_4^1 - 4\mu_3^1\mu_1^1 + 6\mu_2^1\left(\mu_1^1\right)^2 - 3\left(\mu_1^1\right)^4$$

$$= 560 - 4(-117)(-4) + 6(22)(-4)^2 - 3(-4)^4$$

$$= 560 - 936 + 212 - 768$$

= 968

(ii) Moments about origin

We know that first moment about origin is mean

i.e. $\mu_1^1 = \bar{x} = 1$

$$\mu_2^1 = \mu_2 + \left(\mu_1^1\right)^2 = 6 + 1 = 7$$

$$\mu_3^1 = \mu_3 + 3\mu_2\mu_1^1 + \left(\mu_1^1\right)^3$$

$$= 19 + 3(6)(1) + (1)^3$$

= 38

$$\mu_4^1 = \mu_4 + 4\mu_3\mu_1^1 + 6\mu_2\left(\mu_1^1\right)^2 + \left(\mu_1^1\right)^4$$

$$= 968 + 4(19)(1) + 6(6)(1)^2 + (1)^4$$

$$= 968 + 76 + 36 + 1$$

= 1071

Ex 6 : In a frequency distibution coefficient of skewness based upon the quartiles is 0.6. If the sum of the upper and lower quartiles are 100 and median is 38. Find the values of upper and lower quantites.

Sol : Coefficient of SK $= \frac{Q_3 + Q_1 - 2 \text{ Median}}{Q_3 - Q_1}$

Given that $Q_3 + Q_1 = 100, \text{Median} = 38$ and coefficient of skawness = 0.6

$$\therefore \quad 0.6 = \frac{100 - 2 \times 38}{Q_3 - Q_1}$$

$$\Rightarrow Q_3 - Q_1 = \frac{100 - 76}{0.6} = 40$$

$$\therefore Q_3 + Q_1 = 100$$

$$Q_3 - Q_1 = 40 \qquad \Rightarrow 2Q_3 = 140$$

$$\Rightarrow Q_3 = 70$$

$$Q_1 = 100 - Q_3 = 100 - 70 = 30$$

$$\therefore Q_1 = 30 \quad \text{and} \quad Q_3 = 70$$

Ex 7 : A frequency distribution gives the following results. C.V. = 5, Karl Pearsons SK = 0.5 and $\sigma = 2$. Find mean and mode of the distribution.

Sol : Given that

C.V. = 5

$$\Rightarrow \frac{\sigma}{\bar{x}} \times 100 = 5$$

$$\Rightarrow \frac{2}{\bar{x}} \times 100 = 5$$

$$\Rightarrow \bar{x} = \frac{2 \times 100}{5} = 40$$

Karl Pearson's Coefficient of sk = 0.5

$$\Rightarrow \frac{\text{Mean} - \text{Mode}}{\text{S.D}} = 0.5$$

$$\Rightarrow \frac{40 - \text{Mode}}{2} = 0.5$$

$$\Rightarrow 40 - \text{Mode} = 2 \times 0.5 = 1$$

$\therefore$ Mode = 40 -1 = 39

$\therefore$ Mean = 40, Mode = 39.

EXERCISE:

1. Define central and non-central moments. Define central moments interms of non-central moments.
2. Define skewness and explain the various measures of coefficient of skewness.
3. Define kurtosis and give its various measures and curves.
4. Show that for any frequency distribution coefficient of Kurtsis is grater than unity.
5. Show that Bowley's coefficient of skewness lies between -1 and +1.
6. What are Shappard's corrections ? Explain them.
7. Show that moments are not effected by origin but effected by scale.
8. The first four moments of a distribution about the value 4 of the variable are -1.5, 17, -30 and 108. Find the moments about mean, β_1 and β_2. Also find the moments about origin.
9. The first four moments of a distribution about 4 are 1, 4, 10 and 45. Obtain the moments about mean, β_1 and β_2. Comment on their values.
10. For a distribution the mean is 10, variance is 16, γ_1 is + 1 and β_2 is 4. Obtain the first four raw moments.
11. The first four central moments of a distribution are 0, 2.5, 0.7 and 18.75. Compute coefficient of skewness and kurtosis and comment.
12. The first three moments about the origin are

$$\mu_1^1 = \frac{n+1}{2}, \quad \mu_2^1 = \frac{(n+1)(2n+1)}{6}, \quad \mu_3^1 = \frac{n(n+1)^2}{4}$$

obtain the variance and β_1 coefficient.

13. Prove that $\mu_r = \mu_r^1 - rc_1\mu_{r-1}^1\left(\mu_1^1\right) + ... + (-1)^r\left(\mu_1^1\right)^r$.

14. Obtain Karl Pearson's coefficient of skewness for the following data.

Class :	5 - 15	15 - 25	25 - 35	35 - 45	45 - 55
Frequency :	14	22	36	18	10

Answers :

8. mean = 2.5, μ_2 = 14.75, μ_3 = 39.75, μ_4 = 142.3125

μ_1^1 = 2.5, $\mu_2^1 = 21$, μ_3^1 = 166, μ_4^1 = 1132

9. mean = 5, μ_2 = 3, μ_3 = 0, μ_4 = 26

μ_1^1 = 5, μ_2^1 = 28, μ_3^1 = 170, μ_4^1 = 1101

$\beta_1 = 0, \ \beta_2 = 2.889$

The distribution is symmetric and platykurtic.

10. $\mu_1^1 = 10, \ \mu_2^1 = 116, \ \mu_3^1 = 1544, \ \mu_4^1 = 23184$

11. $\beta_1 = 0.0314, \ \beta_2 = 3$

12. Variance $\mu_2 = \frac{n^2 - 1}{12}, \ \beta_1 = 0$.

13. mean = 216, mode = 22, σ = 7.137

Karl Pearson coefficient of skewness = -0.056

14. $\mu_2 = 134.56 \qquad \mu_3 = 126.144 \qquad \mu_4 = 41853.26$

$\beta_1 = 0.006531$ the distribution is positively skewed

$\beta_2 = 2.31151$ the distribution is Leptokurtic Curve.

Lesson – 3

ANALYSIS OF CATEGORICAL DATA

Syllabus:

Consistency of categorical data, independene and association of attributes, various meaures of association for two way data with real life examples.

Objectives:

After studying the lesson the students are expected to have clear comprehension of the theory and practical utility about the concepts of consistency of categorical data, independence and association of attributes, various measures of association for two way data with real life examples.

Structure of The Lesson:

This lesson consists of sections as detailed below:

3.1 Introduction:

The statistical data my be classified as two catagiries, they are quantitative data and qualitative data. Quantitative data means the data which is measured in terms of numbers.

Ex: Weight in Kgs of a man, height of a person in inches etc...

The data obtained in this way are known as statistics of variable. The other type of statistical data is called statistics of attributes. Literally an attribute means a quality or characteristic. Theory of attributes deals with qualitative characteristics which are not quantitatively measurable. Examples of such situations arise when one deals with characters like smoking, health, honesty etc. In this case, it is not possible to measure the extent of smoking or honesty but one can covert the number of persons who possess a particular quality and who do not possess it.

3.2 Notations:

To understand the theory of attributes, it is necessary to introduce some notations for the classes formed and the number of observations assigned to each. The capital letters A,B,C,......... will be used to denote the presence or posses the several attributes. The greek letters $\alpha, \beta, \gamma, \cdots\cdots\cdots$ are generally used to represent the absence of the attributes A,B,C, ... respectively. For example if A represents the blindness then α represent not blindness, if B represents smoker then β represents not smoker etc.

If A denotes the attribute of being blind, B represents the attribute of smoking then the combination of attributes will be represented by grouping together the letters that indicate the attributes concerned.

Thus AB stands for blind and smoker

$A\beta$ stands for blind and non-smoker

αB stands for not blind and smoker

$\alpha\beta$ stands for not blind and non-smoker

If a third attribute C be included to represent, say male, then ABC will stand for male blind smokers. Similarly $AB\gamma,\ A\beta C,\ A\beta\gamma$ etc$\cdots\cdots$

3.3 Class and Class - Frequencies:

Different attributes and their combinations are called different classes and the number of observations assigned to them are called class frequencies. The class frequencies are denoted by putting the letter or letters within brackets.

If class of the certain attribute can be denoted by letter A then the class frequency can be denoted as (A).

i.e. (A) means the number of objects belonging to class A.

(AB) means number of objects passessingthe attributes A and B.

Similarly $(A\beta), (\alpha\beta), (A\beta\gamma), (ABC)$ etc are the number of objects possessing the attributes $A\beta,\ \alpha\beta,\ AB\gamma,\ ABC$ respectively.

A class frequency of capital letters or positive attributes only is called positive class frequency and greek letters or nagative attributes are called the negative class frequencies.

Ex: (AB) is a positive class frequency

$(\alpha\beta)$ is a negative class frequency

For conveniance total frequency "N" is taken as a positive class frequency.

3.4 Order of Class Frequencies:

A class containing " r " attributes then it is taken as of r^{th} order class and a class frequecny containing " r " attributes init is taken as of r^{th} order class frequecny.

Ex: AB, BC, CA are the second order classes and their frequencies $(AB), (BC), (CA)$ are the second order class frequencies.

Similarly $(A), (AB), (ABC)$ are respectively first, second, third order class frequencies.

Note: Conventionally, total frequency N is taken as the class frequency of order zero.

Class frequency in terms higher order class frequencies:

All the class frequencies of various orders are not indipendent of each other and any class frequency can always be expressed in terms of class frequencies of higher order.

Thus $N = (A) + (\alpha) = (B) + (\beta) = (C) + (\gamma)$.

The numbers of A's is equal to the number of A's which are B's added to the number of A's which are not B's.

i.e. $(A) = (AB) + (A\beta)$

$(\alpha) = (\alpha B)(\alpha\beta)$

Similarly $(AB) = (ABC) + (AB\gamma)$

$(A\beta) = (A\beta C) + (A\beta\gamma)$

$(\alpha\beta) = (\alpha BC) + (\alpha B\gamma)$

$(\alpha\beta) = (\alpha\beta C) + (\alpha\beta\gamma)$ and so on

$\therefore (A) = (AB) + (A\beta)$

$$= (ABC) + (AB\gamma) + (A\beta C) + (A\beta\gamma)$$

$$(\alpha) = (\alpha B) + (\alpha\beta)$$

$$= (\alpha BC) + (\alpha B\gamma) + (\alpha\beta C) + (\alpha\beta\gamma)$$

Also $N = (A) + (\alpha)$

$$= (AB) + (A\beta) + (\alpha B) + (\alpha\beta)$$

$$= (ABC) + (AB\gamma) + (A\beta C) + (A\beta\gamma) + (\alpha BC) + (\alpha B\gamma) + (\alpha\beta C) + (\alpha\beta\gamma)$$

3.5 Ultimate Class Frequencies:

Any class frequency is expressable in terms of class frequencies of higher order are called the ultimate class frequencies. Thus in case of n attributes the ultimate class frequencies will be the frequency of n^{th} order.

Note:

(1) In case of n attributes, the ultimate class frequencies each contain n symbols and since each symbol may be written in two ways there are positive part and negative part, for example A or α, B or β etc. Hence the total of ultimate class frequencies of n attributes is 2^n.

If A, B, C are three attributes then the total number of ultimate class frequencies are $2^3 = 8$.

(2) Any class frequency can express as the sum of some of the 2^n ultimate class frequencies.

(3) The total of number of ultimate class frequencies specify the data completely.

Classification:

The objects or individuals possessing or not possessing a particular attribute for two distinct classes. They are (i) Dichotomy, (ii) Manifold Classifications.

Dichotomy: The process of dividing the collection of individuals into two classes according to the presence or absence of an attribute is called Dichotomy.

Ex: Total population can be classified into two classes as blind and not blind people.

Manifold Classification: If the number of sub classes are more than two then the grouping is called manifold classification.

Ex: Grouping of university students into three classes as fail, second class, first class is a manifold classification.

Theorem 1: Prove that the total number of class frequencies with 3 attributes is 27.

Proof: Let A, B, C are three attributes. The following are the various orders and their class frequences.

Order	Frequencies	Number of frequencies
0	N	1
1	$(A), (B), (C), (\alpha), (\beta), (\gamma)$	6
2	$(AB), (AC), (BC)$ $(A\beta), (A\gamma), (B\gamma)$ $(\alpha B), (\alpha C), (\beta C)$ $(\alpha\beta), (\alpha\gamma), (\beta\gamma)$	12
3	$(ABC), (AB\gamma), (A\beta C), (A\beta\gamma)$ $(\alpha BC), (\alpha B\gamma), (\alpha\beta C), (\alpha\beta\gamma)$	8
		27

Hence the total number of class frequencies with 3 attributes is 27.

Theorem 2: Prove that total number of class frequencies with n attributes are 3^n.

Proof: Suppose the " n " attributes are $A, B, C, \cdots\cdots\cdots\cdots, M$

of order zero number of class frequencies = 1

of order 1 number of class frequences = $2 \times nC_1$

[Since out of n attributes one can be selected in nC_1 ways and for each of the nC_1 first order classes there would be two class frequencies Ex: (A) and (α) for A, (B) and (β) for B etc]

of order 2 number of class frequencies $= nC_2 \ 2^2$

[Since out of n attributes 2 can select nC_2 ways and for each of nC_2 combinations of two attributes we can form 4 class frequencies Ex: with two attributes A and B we form (AB), $(A\beta)$, (αB), $(\alpha\beta)$ are 2^2 class frequencies]

Similarly of order 3 number of class frequencies $= nC_3\ 2^3$

and so on

of order n the number of class frequencies $= nC_n\ 2^n$

Hence the total number of class frequencies with " n "

Attributes $= 1 + nc_1 2 + nc_2 2^2 + nc_3 2^3 + \cdots\cdots\cdots + nc_n 2^n$

$= (1+2)^n$

$= 3^n$

Theorem 3: Prove that the total number of positive class frequencies with n attributes in 2^n .

Proof: Suppose there are n attributes $A, B, C, \cdots\cdots\cdots, M$

N is the only positive class frequency of order zero.

$\therefore$ number of positive class frequencies of order zero = 1

number of positive class frequencies of order one = nc_1

[since one attribute will give only one positive class freuency, Ex. for A only (A)].

number of positive class frequencies of order $2 = nc_2$

[Out of two attributes 2 can select in nc_2 ways and combination of any two attributes will give only one positive class frequency for example (AB) for A and B].

Similary number of positive class frequencies of order 3 = $nc_3 \cdot 1$ and so on.

number of positive class frequencies of order $n = nc_n \cdot 1$

$\therefore$ The total number of positive class frequencies of with n attributes

$= 1 + nc_1 + nc_2 + \cdots\cdots + nc_n$

$= (1+1)^n$

$= 2^n$

Note: The total number of negative class frequencies with n attributes are $2^n - 1$.

Class Symbols as Operators:

Usually the some of attributes will be given and remaining must be found by using them. It is not easy to remember always. For finding the unknown attributes without errors the easy method is as follows:

Let us write symbolically

$$AN = (A)$$

$$\alpha N = (\alpha)$$

Adding we get $AN + \alpha N = (A) + (\alpha)$

$$\Rightarrow (A + \alpha) N = N$$

$$\Rightarrow A + \alpha = 1$$

$$\therefore \alpha = 1 - A$$

(or) $A = 1 - \alpha$

Thus in symbolic expression we can replace A by $1 - \alpha$ and α by $(1 - A)$. Similarly B can be replaced by $(1-\beta)$ and β can be replaced by $(1-B)$ etc.

Dichotomising (B) according to A,

Let us write

$$A(B) = (AB)$$

Similarly $B(A) = (BA)$

$$\therefore A(B) = B(A) = (AB) = AB \cdot N$$

These are operators only, they are not numbers. By using these operators we can express any formula easily.

Ex: $(\alpha\beta) = \alpha\beta N$

$$= (1-A)(1-B)N$$

$$= N - AN - BN + ABN$$

$$= N - (A) - (B) + (AB)$$

$$(AB\gamma) = AB\gamma N$$

$$= AB(1-C)N$$

$$= ABN - ABCN$$

$$= (AB) - (ABC)$$

$$(\alpha\beta\gamma) = \alpha\beta\gamma N$$

$$= (1-A)(1-B)(1-C)N$$

$$= N - AN - BN - CN + ABN + ACN + BCN - ABCN$$

$$= N - (A) - (B) - (C) + (AB) + (AC) + (BC) - (ABC)$$

$$(A\beta\gamma) = A\beta\gamma N$$

$$= A(1-B)(1-C)N$$

$$= AN - ABN - ACN + ABCN$$

$$= (A) - (AB) - (AC) + (ABC)$$

and so on.

Theorem: Prove that for n attributes $A_1, A_2, \cdots\cdots, A_n$ as

$$(A_1 A_2 \cdots\cdots A_n) \geq (A_1) + (A_2) + \cdots\cdots + (A_n) - (n-1)N$$

where N is the total number of observations.

Proof: This can be proved by mathematical induction.

We have

$$(\alpha_1 \alpha_2) = \alpha_1 \alpha_2 N$$

$$= (1-A_1)(1-A_2)N$$

$$= N - A_1N - A_2N + A_1A_2N$$

$$= N - (A_1) - (A_2) + (A_1A_2)$$

Since the class frequency is always non - negative.

$\therefore (\alpha_1 \alpha_2) \geq 0$

$\Rightarrow N - (A_1) - (A_2) + (A_1 A_2) \geq 0$

$\Rightarrow (A_1 A_2) \geq (A_1) + (A_2) - N \cdots\cdots\cdots\cdots (1)$

Hence the theorem is true for $n = 2$

By substituting A_2A_3 in place of A_2 in (1) we have

$(A_1 A_2 A_3) \geq (A_1) + (A_2 A_3) - N$

$\geq (A_1) + (A_2) + (A_3) - N - N$

$\therefore (A_1 A_2 A_3) \geq (A_1) + (A_2) + (A_3) - 2N$

Hence the theorem is true for $n = 3$.

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

that is $(A_1 A_2 \cdots\cdots\cdots\cdots A_r) \geq (A_1) + (A_2) + \cdots\cdots\cdots\cdots + (A_r) - (r-1) N \cdots\cdots (2)$

Replacing the attribute A_r by $A_r A_{r+1}$ in (2) we get

$$(A_1 A_2 \cdots\cdots A_r A_{r+1}) \geq (A_1) + (A_2) + \cdots\cdots\cdots\cdots + (A_r A_{r+1}) - (r-1)N$$

$$\geq (A_1) + (A_2) + \cdots\cdots\cdots\cdots + \{(A_r) + (A_{r+1}) - N\} - (r-1)N$$

$$\geq (A_1) + (A_2) + \cdots\cdots\cdots + (A_r) + (A_{r+1}) - rN$$

Thus the theorem is true for $n = r + 1$.

Hence by induction, the theorem istrue for all positive integer values of n.

3.6 Consistencey:

Definition:

Consistency of a set of class frequencies may be defined as the property that none of them is negative, otherwise the data for class frequencies are said to be inconsistent.

Sine any class frequency can be expressed as the sum of some ultimate class frequencies it is necesarily non-negative if all the ultimate class frequencies are non-negative.

Result:

The necessary and sufficient condition for the consistency of a set of independent class frequencies is that no ultimate class frequency is negative.

Conclusion:

To determine whether the given frequencies are consistent or inconsistent, the given frequencies are to be expressed in ultimate class frequencies. If any of them is negative then the data are inconsistent otherwise the data are consistent.

3.6.1 Consistency Conditions: Consistency conditions for a single attribute A are

$$(A) \geq 0$$

$$(\alpha) \geq 0 \Rightarrow (A) \leq N \qquad (3.6.1.1)$$

Consistency Conditions for two attributes A and B are

$$(AB) \geq 0$$

$$(A\beta) \geq 0 \Rightarrow (AB) \leq (A)$$

$$(\alpha B) \geq 0 \Rightarrow (AB) \leq (B)$$

$$(\alpha\beta) \geq 0 \Rightarrow (AB) \geq (A) + (B) - N$$

Consistency Conditions for three attributes A, B and C are

(i) $(ABC) \geq 0$

(ii) $(AB\gamma) \geq 0 \Rightarrow (ABC) \leq (AB)$

(iii) $(A\beta C) \geq 0 \Rightarrow (ABC) \leq (AC)$

(iv) $(\alpha BC) \geq 0 \Rightarrow (ABC) \leq (BC)$

(v) $(A\beta\gamma) \geq 0 \Rightarrow (ABC) \geq (AB) + (AC) - (A)$

(vi) $(\alpha B\gamma) \geq 0 \Rightarrow (ABC) \geq (AB) + (BC) - (B)$

(vii) $(\alpha\beta C) \geq 0 \Rightarrow (ABC) \geq (AC) + (BC) - (C)$

(viii) $(\alpha\beta\gamma) \geq 0 \Rightarrow (ABC) \leq (AB) + (BC) + (AC) - (A) - (B) - (C) + N$

From (i) and (viii) we get

$$(AB)+(BC)+(AC) \geq (A)+(B)+(C)-N$$

From (ii) and (vii) we get

$$(AC)+(BC)-(AB) \leq (C)$$

From (iii) and (vi) we get

$$(AB)+(BC)-(AC) \leq (B)$$

From (iv) and (v) we get

$$(AB)+(AC)-(BC) \leq (A)$$

3.7 Independence of Attributes:

Two attributes A and B are said to be independent if there exists no relationship of any kind between them.

If A and B are independent we would expect (i) The same proportion of A's amongst B's as amongst β's (ii) The proportion of B's amongst A's same as that amongst the α's.

Criterion of Independence:

1. If A and B are independent then from (i)

$$\frac{(AB)}{(B)} = \frac{(AB)}{(\beta)} \Rightarrow \frac{(\alpha B)}{(B)} = \frac{(\alpha\beta)}{(\beta)} \cdots\cdots\cdots\cdots\cdots (3.6.2.1)$$

Similarly from (ii) we have

$$\frac{(AB)}{(A)} = \frac{(\alpha B)}{(\alpha)} \Rightarrow \frac{(A\beta)}{(A)} = \frac{(\alpha\beta)}{(\alpha)} \cdots\cdots\cdots\cdots\cdots (3.6.2.2)$$

from equation (3.6.2.1)

$$\frac{(AB)}{(B)} = \frac{(AB)}{(\beta)} = \frac{(AB)+(A\beta)}{(B)+(\beta)} = \frac{(A)}{N}$$

It becomes easier to grasp the nature of the above relations if the frequencies are supposed to be grouped into a table with two rows and two columns as follows:

	A	α	total
B	(AB)	(αB)	(B)
β	$(A\beta)$	$(\alpha\beta)$	(β)
total	(A)	(α)	N

From table we express as

(i) $$\frac{(AB)}{(A)} = \frac{(\alpha B)}{(\alpha)} = \frac{(B)}{N}$$

(ii) $$\frac{(A\beta)}{(A)} = \frac{(\alpha\beta)}{(\alpha)} = \frac{(\beta)}{N}$$

(iii) $$\frac{(AB)}{(B)} = \frac{(A\beta)}{(\beta)} = \frac{(A)}{N}$$

(iv) $$\frac{(\alpha B)}{(B)} = \frac{(\alpha\beta)}{(\beta)} = \frac{(\alpha)}{N}$$

(v) $$\frac{(AB)}{(A\beta)} = \frac{(\alpha B)}{(\alpha\beta)} = \frac{(B)}{(\beta)}$$

(vi) $$\frac{(AB)}{(\alpha B)} = \frac{(A\beta)}{(\alpha\beta)} = \frac{(A)}{(\alpha)}$$

2. The criterion of independence may be obtained in terms of class frequencies of first order and it is given by

$$(AB) = \frac{(A)(B)}{N}$$

$$\Rightarrow \frac{(AB)}{N} = \frac{(A)}{N} \cdot \frac{(B)}{N}$$

which leads to the following fundamental rule.

If the attributes A and B are independent, the proportion of AB's in the population is equal to te product of proportions of A's and B's in the population.

3. We may obtain the third criterion of independene in terms of second order class frequencies as follows:

$$(\mathrm{AB})(\alpha\beta) = (\mathrm{A}\beta)(\alpha\mathrm{B})$$

3.8 Asociation of Attributes:

Two attributes A and B are said to be associated if they are not independent but are related in some way or the other.

The association are of two types they are (i) positive association, (ii) Negative association.

Positive Association:

Two attributes A and B are said to be positively associated or simply associated if

$$(\mathrm{AB}) > \frac{(\mathrm{A})(\mathrm{B})}{\mathrm{N}}.$$

Negative Association:

Two attributes A and B are said to be Negatively Associated or disassociated if

$$(\mathrm{AB}) < \frac{(\mathrm{A})\ (\mathrm{B})}{\mathrm{N}}$$

Complete Association:

If all A's are B's or all B's are A's then A and B are said to be completely associated. So A and B are completely associated if either $(\mathrm{AB}) = (\mathrm{A})$ or $(\mathrm{AB}) = (\mathrm{B})$ which even is less.

Complete Disassociation:

If none of the A's are B's that is $(\mathrm{AB}) = 0$ or more of the α's are β's that is $(\alpha\beta) = 0$ then A and B are said to be completely negatively associated.

Thus when A and B are completely disassociated if $(\mathrm{AB}) = 0$ or $(\mathrm{AB}) = (\mathrm{A}) + (\mathrm{B}) - \mathrm{N}$, whichever is more.

3.8.1 Coefficient of Association:

To measure the intensity of association, we use the two measures of coefficients, they are

1. Yule's coefficient of Association
2. Coefficient of colligation

1. Yule's Coefficient of Association: For measuring the intensity of association between two attributes A and B G. Vdny Yule gave the coefficient of association Q defined as follows:

$$Q = \frac{(AB)(\alpha\beta) - (A\beta)(\alpha B)}{(AB)(\alpha\beta) + (A\beta)(\alpha B)}$$

Properties:

i) When A and B are independent then Q = 0

ii) If A and B are completely associated Q then Q = 1

iii) If A and B are completely disassociated then Q = -1

iv) The range of Q is -1 to +1

v) Yule's coefficient of association will not change if the terms containing any of the attributes are multiplied by the same constant.

3.8.2 Coefficient of Colligation: The coefficient of colligation between A and B are given by

$$Y = \frac{\sqrt{(AB)(\alpha\beta)} - \sqrt{(A\beta)(\alpha B)}}{\sqrt{(AB)(\alpha\beta)} + \sqrt{(A\beta)(\alpha B)}}$$

Propertions:

1. If A and B are independent then $Y = 0$
2. If A and B are completely associated then $Y = +1$
3. If A and B are completely dissassociated then $Y = -1$
4. The range of colligation in -1 to $+1$
5. The relation between Q and Y is $Q = \frac{2Y}{1+Y^2}$
6. The value of Y will not change if all the terms containing any of the attributes are multiplied by the same constant.

Theorem 3.8.2: Prove that in the usual notations $Q = \frac{2Y}{1+Y^2}$

Proof: The coefficient of colligation Y is given by

$$Y = \frac{1 - \sqrt{\frac{(A\beta)(\alpha B)}{(AB)(\alpha\beta)}}}{1 + \sqrt{\frac{(A\beta)(\alpha B)}{(AB)(\alpha\beta)}}}$$

Let $K = \frac{(A\beta)(\alpha B)}{(AB)(\alpha\beta)}$ then we have

$$Y = \frac{1 - \sqrt{K}}{1 + \sqrt{K}}$$

$$Y^2 = \frac{1 + K - 2\sqrt{K}}{1 + K + 2\sqrt{K}}$$

i.e. $1 + Y^2 = 1 + \frac{1 + K - 2\sqrt{K}}{1 + K + 2\sqrt{K}}$

$$= \frac{1 + K + 2\sqrt{K} + 1 + K - 2\sqrt{K}}{1 + K + 2\sqrt{K}}$$

$$= \frac{2(1 + K)}{1 + K + 2\sqrt{K}}$$

$$= \frac{2(1 + K)}{\left(1 + \sqrt{K}\right)^2}$$

$$\therefore \frac{2Y}{1 + Y^2} = \frac{2\left(1 - \sqrt{K}\right)}{1 + \sqrt{K}} \Bigg/ \frac{2(1 + K)}{\left(1 + \sqrt{K}\right)^2}$$

$$= \frac{\left(1 - \sqrt{K}\right)\left(1 + \sqrt{K}\right)}{1 + K}$$

$$= \frac{1 - K}{1 + K}$$

$$= \frac{1 - \frac{(A\beta)(\alpha B)}{(AB)(\alpha\beta)}}{1 + \frac{(A\beta)(\alpha B)}{(AB)(\alpha\beta)}}$$

$$= Q$$

$$\therefore\ Q = \frac{2Y}{1+Y^2}$$

3.9 Examples:

Example 1: Find the remainiing class frequencies, given the following data for two attributes A and B.

$$(AB) = 250,\ (A\beta) = 120,\ (\alpha B) = 200,\ (\alpha\beta) = 70$$

Solution: Here $N = (AB) + (A\beta) + (\alpha B) + (\alpha\beta)$

$$= 250 + 120 + 200 + 70$$

$$= 640$$

$$(A) = (AB) + (A\beta)$$

$$= 250 + 120 = 370$$

$$(B) = (AB) + (\alpha\beta)$$

$$= 250 + 200 = 450$$

$$(\alpha) = N - (A) = 640 - 370 = 270$$

$$(\beta) = N - (B) = 640 - 450 = 190$$

Example 2: Find the remaining class frequencies, given the following data.

$$(A) = 50,\ (B) = 40,\ (AB) = 30,\ N = 100$$

Solution: $(A\beta) = (A) - (AB)$

$$= 50 - 30 = 20$$

$(\alpha B) = (B) - (AB) = 40 - 30 = 10$

$(\beta) = N - (B) = 100 - 40 = 60$

$(\alpha) = N - (A) = 100 - 50 = 50$

$(\alpha\beta) = N - (A) - (B) + (AB) = 100 - 50 - 40 + 30 = 40$

Example 3: Find the remaining class frequencies, given the following data.

$$(ABC) = 57, (\alpha BC) = 78, (AB\gamma) = 281, (\alpha B\gamma) = 620,$$

$$(A\beta C) = 86, (\alpha\beta C) = 65, (A\beta\gamma) = 453, (\alpha\beta\gamma) = 8310.$$

Solution: For three attributes there are $3^3 = 27$, class frequencies given 8 class frequencies, so we have to find the remaining 19 class frequencies.

Order 2: $(AB) = (ABC) + (AB\gamma)$

$= 57 + 281$

$= 338$

$(AC) = (ABC) + (A\beta C) = 57 + 86 = 143$

$(BC) = (ABC) + (\alpha BC) = 57 + 78 = 135$

$(\alpha B) = (\alpha BC) + (\alpha B\gamma) = 78 + 670 = 748$

$(\alpha C) = (\alpha BC) + (\alpha\beta C) = 78 + 65 = 143$

$(\alpha\beta) = (\alpha\beta C) + (\alpha\beta\gamma) = 65 + 8310 = 8375$

$(\alpha\gamma) = (\alpha B\gamma) + (\alpha\beta\gamma) = 670 + 8310 = 8980$

$(A\beta) = (A\beta C) + (A\beta\gamma) = 86 + 453 = 539$

$(A\gamma) = (AB\gamma) + (A\beta\gamma) = 281 + 453 = 734$

$(B\gamma) = (AB\gamma) + (\alpha B\gamma) = 281 + 670 = 951$

$(\beta C) = (A\beta C) + (\alpha\beta C) = 86 + 65 = 151$

$(\beta\gamma) = (A\beta\gamma) + (\alpha\beta\gamma) = 453 + 8310 = 8763$

Order 1: $(A) = (AB) + (A\beta) = 338 + 539 = 877$

$(B) = (AB) + (\alpha B) = 338 + 748 = 1086$

$(C) = (BC) + (\beta C) = 135 + 151 = 286$

$(\alpha) = (\alpha B) + (\alpha\beta) = 748 + 8375 = 9123$

$(\beta) = (A\beta) + (\alpha\beta) = 539 + 8375 = 8914$

$(\gamma) = (A\gamma) + (\alpha\gamma) = 734 + 8980 = 9714$

Order 0: $N = (A) + (\alpha)$

$= 877 + 9123$

$= 10000$

Example 4: Given the following positive class frequencies, findthe ultimate class frequencies.

$N = 12000, (A) = 977, (B) = 1185, (C) = 596,$

$(AB) = 153, (AC) = 284, (BC) = 250, (ABC) = 127.$

Solution: The ultimate class frequencies are

$(ABC) = 127$ (given)

$(\alpha BC) = (BC) - (ABC)$

$= 250 - 127 = 123$

$(A\beta C) = (AC) - (ABC)$

$= 284 - 127 = 157$

$(AB\gamma) = (AB) - (ABC)$

$= 453 - 127 = 326$

$(\alpha B\gamma) = (B) - (BC) - (AB) + (ABC)$

$= 596 - 50 - 284 + 127 = 189$

$$(A\,\beta\,\gamma) = (A) - (AB) - (AC) + (ABC)$$

$$= 977 - 453 - 284 + 127 = 367$$

$$(\alpha\,\beta\,\gamma) = N - (A) - (B) - (C) + (AB) + (BC) + (AC) - (ABC)$$

$$= 12000 - 977 - 1185 - 596 + 453 + 284 + 250 - 127$$

$$= 10102$$

Example 5: If 598 men in a locality exposed to cholera 147 in all were attacked. 137 were inoculated and of these only 14 were atttacked. Find the number of persons not inoculated not attacked, inoculated not attacked and not inoculated attacked.

Solution: Let A be the attribute of attacked.

Let B be the attribute of inoculated.

Given that $N = 598, (A) = 147, (B) = 137, (AB) = 14,$

So we have to obtain the values of $(\alpha\beta), (\alpha B)$ and $(A\beta)$.

$$(\alpha\,\beta) = N - (A) - (B) + (AB)$$

$$= 598 - 147 - 137 + 14 = 328$$

$$(\alpha\,B) = (B) - (AB)$$

$$= 137 - 14 = 123$$

$$(A\,\beta) = (A) - (AB)$$

$$= 147 - 14 = 133$$

Example 6: Given that $(A) = (B) = (C) = \frac{N}{2}$ and 80% of A's are B's, 75% of A's are C's. Find the limits to the percentage of B's that are C's.

Solution: Let $(A) = (B) = (C) = \frac{N}{2} = 100$

then $(AB) = 80,\ (AC) = 75$.

We have to find the limits of (BC).

By using the conditions for consistency, we have

$$(AB)+(BC)+(AC) \geq (A)+(B)+(C)-N$$

$$\Rightarrow (BC) \geq (A)+(B)+(C)-N-(AB)-(AC)$$

$$\geq 100+100+100-200-80-75$$

$\Rightarrow (BC) \geq -55 \cdots\cdots\cdots\cdots(1)$ which is wrong

$$(AC)+(BC)-(AB) \leq (C)$$

$$\Rightarrow (BC) \leq (C)+(AB)-(AC)$$

$$\leq 100+80-75$$

$\Rightarrow (BC) \leq 105 \cdots\cdots\cdots\cdots \geq (2)$ which is wrong

$(AB)+(BC)-(AC) \leq (B)$ $\qquad \because (BC) \leq (B)$

$$\Rightarrow (BC) \leq (B)+(AC)-(AB)$$

$$\leq 100+75-80$$

$$\therefore (BC) \leq 95 \cdots\cdots\cdots\cdots(3)$$

$$(AB)+(AC)-(BC) \leq (A)$$

$$\Rightarrow (AB)+(AC)-(A) \leq (BC)$$

$$\Rightarrow 80+75-100 \leq (BC)$$

$$\Rightarrow 55 \leq (BC) \cdots\cdots\cdots(4)$$

From (1), (2), (3), (4) we get $55 \leq (BC) \leq 95$

Hence (BC) lies between 55% and 95%.

Example 7: Given that $(A)=(B)=(C)=\frac{N}{2}=50$ and $(AB)=30$, $(AC)=25$. Find the limits within which (BC) will lie.

Solution: By using conditions consistency we have

$(AB)+(BC)+(AC)\geq(A)+(B)+(C)-N$

$\Rightarrow(BC)\geq(A)+(B)+(C)-N-(AB)-(AC)$

$\geq 50+50+50-100-30-25$

$\Rightarrow(BC)\geq -5\cdots\cdots\cdots\cdots(1)$ which is wrong

$(AB)+(AC)-(BC)\leq(A)$

$\Rightarrow (AB)+(AC)-(A)\leq(BC)$

$\Rightarrow 30+25-50\leq(BC)$

$\Rightarrow 5\leq(BC)\cdots\cdots\cdots\cdots(2)$

$(AB)+(BC)-(AC)\leq(B)$

$\Rightarrow (BC)\leq(B)+(AC)-(AB)$

$\Rightarrow (BC)\leq 50+25-30$

$\Rightarrow (BC)\leq 45\cdots\cdots\cdots\cdots(3)$

$(AC)+(BC)-(AB)\leq(C)$

$\Rightarrow (BC)\leq(C)+(AB)-(AC)$

$\leq 50+30-25$

$\Rightarrow (BC)\leq 55\cdots\cdots\cdots\cdots(4)$ which is wrong since $(BC)\not<(B)$

$\therefore\ 5\leq(BC)\leq 45$

Example 8: If A and B are two independent attributes and $N=1024,\ (A)=144, (B)=384$ then find $(AB), (A\beta),\ (\alpha B)$ and $(\alpha\,\beta)$ values.

Solution: $(\alpha)=N-(A)$

$=1024-144=880$

$(\beta) = N - (B)$

$= 1024 - 384 = 640$

Since A and B are independent attributes

then $(AB) = \frac{(A)(B)}{N} = \frac{144 \times 384}{1024} = 54$

$(A\,\beta) = \frac{(A)(\beta)}{N} = \frac{144 \times 640}{1024} = 95$

$(\alpha\,B) = \frac{(\alpha)(B)}{N} = \frac{880 \times 384}{1024} = 330$

$(\alpha\,\beta) = \frac{(\alpha)(\beta)}{N} = \frac{880 \times 640}{1024} = 550$

Example 9: Can vaccination be regarded as a preventive measure for small pox from the data given below? Of 1482 persons in a locality exposed to small Pox 368 in all were attacked. of 1482 persons, 343 had been vaccinated and of these only 35 were attacked.

Solution: Let A denote the attribute of vaccination and B denote attack. Then the given data is $N = 1482,\ (A) = 368, (B) = 343,\ (AB) = 35$.

$(\alpha\,\beta) = N - (A) - (B) + (AB)$

$= 1482 - 368 - 343 + 35 = 806$

$(A\,\beta) = (A) - (AB) = 368 - 35 = 333$

$(\alpha B) = (B) - (AB) = 343 - 35 = 308$

$\therefore$ Yules coefficient of association Q is given by

$$Q = \frac{(A\,B)(\alpha\,\beta) - (A\,\beta)(\alpha\,B)}{(A\,B)(\alpha\,\beta) + (A\,\beta)(\alpha\,B)}$$

$$= \frac{35 \times 806 - 333 \times 308}{35 \times 806 + 333 \times 308}$$

$$= -0 \cdot 5$$

$\therefore$ There is a negative association between A and B.

i.e. there is positive association between not attacked and vaccinated.

Hence vaccination can be regarded as a preventive measure for small pox.

Example 10: Find the association between proficiency in English and in Hindi among candidates at a certain test if 245 of them passed in Hindi, 285 failed in Hindi, 190 failed in Hindi but passed in English and 147 passed in both.

Solution: Let A denotes attribute of passed in English

B denotes attribute of passed in Hindi

given that $(B) = 245$

$(\beta) = 285$

$(A\ \beta) = 190$

$(A\ B) = 147$

$(A\ \beta) = 190$

$\Rightarrow (A) - (A B) = 190$

$\Rightarrow (A) = 190 + (AB)$

$= 190 + 147 = 337$

$N = (B) + (\beta) = 245 + 285 = 530$

$(\alpha\ B) = (B) - (AB) = 245 - 147 = 98$

$(\alpha\ \beta) = N - (A) - (B) + (A B)$

$= 530 - 337 - 245 + 147$

$= 95$

$\therefore$ Yules coefficient of association is given by

$$Q = \frac{(AB)(\alpha\beta) - (A\ \beta)(\alpha\ B)}{(AB)(\alpha\ \beta) + (A\beta)(\alpha\ B)}$$

$$= \frac{147 \times 95 - 190 \times 98}{147 \times 95 + 190 \times 98}$$

$$= -\,0 \cdot 142857$$

Hence association between English and Hindi is $-\,0 \cdot 142857$.

Example 11: The male population of a state is 250 lakhs. The number of literate males is 20 lakhs and total number of male criminals is 26 thousands. The number of literates male criminals is 2 thousands. Do you find any association between literacy and criminality.

Solution: Let attribute A denote literate males and B denotes male criminals.

Thus given that $N = 2,50,00000$

$$(A) = 20,00000$$

$$(B) = 26000$$

$$(AB) = 2000$$

$$(AB) = (A) - (AB) = 2000000 - 2000 = 1998000$$

$$(\alpha B) = (B) - (AB) = 26000 - 2000 = 24000$$

$$(\alpha\beta) = N - (A) - (B) + (AB)$$

$$= 25000000 - 2000000 - 26000 + 2000$$

$$= 2,29,76,000$$

$$Q = \frac{(AB)(\alpha\beta) - (A\beta)(\alpha B)}{(AB)(\alpha\beta) + (A\beta)(\alpha B)}$$

$$= \frac{2000 \times 22976000 - 1998000 \times 24000}{2000 \times 22976000 + 1998000 \times 24000}$$

$$= \frac{45952 - 47952}{45952 + 47952}$$

$$= -\,0 \cdot 0213$$

Hence literacy and criminality are negatively associated.

3.10 Exercises:

1. Define class and class frequency.

2. Explain the following (i) order of a class, (ii) Ultimate class.

3. What do you understand by consistency of given data.

4. What do you mean by independence of attributes? Give a criteria of independence for attributes A and B.

5. What are te various methods of finding whether two attributes are associated, dissociated or independent ?

6. What is association of attributes ? Give the measure of coefficient of association.

7. Define Yule's coefficient of association (Q) and colligation (Y). Prove that $Q = 2Y/\left(1+Y^2\right)$.

8. Given the following ultimate class frequencies, find the frequencies of positive class,

$$(ABC)=149, (AB\gamma)=738, (A\beta C)=225, (A\beta\gamma)=1196,$$

$$(\alpha BC)=204, (\alpha B\gamma)=1762, (\alpha\beta C)=171, (\alpha\beta\gamma)=21842.$$

9. Show that for n attributes $A_1, A_2, \cdots\cdots\cdots, A_n$.

$$(A_1, A_2, \cdots\cdots\cdots, A_n) \geq (A_1)+(A_2)+\cdots\cdots\cdots+(A_n)-(n-1)N$$

where N is the total number of observations.

10. Among the adult population of a certain town 50% are male, 60% are wage - earners and 50% are 45 years of age or over 10% of the males are not wage earners and 40% of the males are under 45 make the best possible inference about the limits within which the percentage of persons of 45 years or over are wage - earners.

11. If $1000 = N = 1\frac{2}{3}(A) = 2(B) = 2\frac{1}{2}(C) = 5(AB)$ and $(AC)=(BC)$, what should be the minimum value of (BC)?

12. Investigate the association between darkness of eye colour in father and son from the following data:

    Fathers with dark eyes and sons with dark eyes : 50

    Fathers with dark eyes and sons with not dark eyes : 79

    Father with not dark eyes and sons with dark eyes : 89

    Father with not dark eyes and sons with not dark eyes : 782

3.11 Answers:

8. $(A) = 2308, (B) = 2853, (C) = 749,$

   $(AB) = 887, (AC) = 374, (BC) = 353, N = 26287.$

10. $25 \le (BC) \le 45$

11. 150

12. $Q = 0 \cdot 6951$

    $Y = 0 \cdot 4052$

3.12 Summary:

In statistical data, the concept of non numerical data called qualitative characteristic is defined. construction of frequency distributions of such a data is given. Notion of consistency, independence, association are introduced. These concepts are applied in analysis of qualitative data.

3.13 Technical Terms:

Altribute

Independence

Association

Contigency Table

Consistency of Data

Class Frequency

Lesson – 4

PROBABILITY

Syllabus:

Definition of probability, classical and relative frequency approach to probability, merits and demerits of these approaches, random experiment sample point and sample space, definition of an event, operation of events, properties of probability based on axioamatic approach, addition theorem for 'n' events.

Objective:

This lesson is prepared in such a way that after studying the material the student is expected to have a thorough comprehension of the concept "probability" - the breath of any statistical investigation and analysis. The student would be equipped with theoretical as well as practical aspects of probability of an event or combination of events.

Structure of the lesson:

4.1 Introduction

4.2 Basic pre-requisites

4.3 Relative frequency approach

4.4 Classical Definition

4.5 Axiomatic Approach

4.6 Addition Theorem

4.7 Examples

4.8 Exercises

4.9 Summary

4.10 Technical Terms

4.1 Introduction:

Quite often we come across statements that are not always true and not always false. The weather forecast in news bulletins. The announcements about arrivals and departures of trains in a railway station, the results of pre-poll surveys in general elections etc. are some situations. In all these examples we see an element of uncertainity associated with them that would prevent us from taking an appropriate decision. Therefore if there is a method of expressing uncertainity in numerical quantity, depending on the magnitude of the numerical quantity one can decide whether are not to go ahead with a decision.

For example, let us take the news presentation in weather fore cast in TV/ Radio. It generally says that in a particular place heavy rains are likely to be experienced, as per the satelite pictures. The word 'likely' in the news presentation makes an individual to become alert and take an umbrella/ rain coat while going outor post-pone his/her out door works and so on. That is, the importance is for rain rather than no rain. The above narration indicates that the weather forecast people have some method that quantities the uncertain incident rain or no rain to say in the news. Effectively this is the phenomenon called probability. Hence a descriptive explanation for the concept probability can be given as 'quantification of uncertainty'. In this lesson we discuss at length the notion of probability, the various developments in its definition, some standared theorem results along with spectific applications.

4.2 Basic Prerequisites :

In this section we present some concepts required to expalin probability.

4.2.1 Definition : Random Experiment : An experiment whose result is not known with certainity, unless the experiment is performed completely.

Example : 1. Applying for an admission in a college

2. Trying to catch a bus in a bus station
3. Winning or losing a match
4. Hitting a target in a shooting test and of course - tossing a coin throwing a dice. etc.

In all these examples some action is performed with an intended result. But the expected result may or may not happen. Infact we experience many random experiments in daily observations.

4.2.2 Definition : Sample Space : In a random experiment, though we can not say exactly - guess out come of the action, with some enlightened vision we can say the various possible results, for the experiment. The set of all possible out comes of a random experiment without any omission is called sample space. In the examples of the definition 4.2.1 the following sets are sample spaces respectively.

{ admission, no admission}

{Catching the Bus, Missing the Bus}

{Win, lose, dran}

{hitting the target, missing the target}

{head, tail}

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

Sample space is similar to the universal set in set theory and is denoted by Ω. The elements of Ω are called sample points are also called simple events. Combination of simple events is called an event. That is sub sets of Ω are called events. For example in a die throwing example the signleton sets {1}, {2}, {3}, {4}, {5}, {6} are called sample points. The subset {1,3,5} is an event - denoting getting an odd number and the set {3, 6} denotes the event getting a multiple of 3.

Hence we can think of a parallel between set theory and events in sample space. If A and B are any two subsets of Ω.

$A^1 = \Omega - A$ is called compumentary event to A

$A \cup B$ = Occurence of either of the events

$A \cap B$ = Occurence of both the events A, B

In particular if $A \cap B = \phi$ - the null set then A, B are specifically called muthally exclusive events.

4.3 Relative Frequency Approach:

Let us consider the repelitions of a random experment say 'n' times. Suppose the event 'E' in which we are interested appears say r times. Then $\frac{r}{n}$ is called relative frequency of the event E, which can be regarded as sequence of real fractions based on 'n' limit of $\frac{r}{n}$ as n approaches to infinity (i.e. limiting value of relative frequency as the number of repetitions becomes larger and larger) iscalled probability of the event 'E'.

The relative frequency approach to the definition of probability is basically a limit of a sequence. Hence unless the sequence is convergent, we cannot get the probability. Even if it is convergent, one may have to do a number of repetitions of the experiment, which may be a costly affair. Therefore this remains more of a theoretical proposition than a practically adaptable definition.

4.4 Classical Definition:

Let the sample space Ω of a random experiment contain 'n' simple events all of which are mutually exclusive, exhaustive and equally likely. Let m (m < n) of these are in favour of the occurence of an event E. Then the ratio $\frac{m}{n}$ is called the probability of the event E.

Example 4.4.1 : In a die throwing the sample space $\Omega = \{1,2,3,4,5,6\}$. Suppose we are intenested in getting a prime number. Then the set {2,3,5} is the interested event say E. Here Ω contains 6 elements and E contains 3 elements i.e. m = 3, n = 6. According to classical definition the probability of the event E is $P(E) = \frac{3}{6}$.

In this definition, if the number points in thesample space is not finite, if the elements of Ω are not equally likely we can not use classical definition. More over, the phrase equally likely events is not clear. It hints that the elements of Ω should have equal chance of happening which in turn means that they should have the same probability of occurence. That is, the notion of probability is imposed on the events to define the concept probability in classical approach. Hence this approach is not totally admissible. Over coming all the demerits of relative frequency approach and classical approach, probability is defined in axiomatic approach by A.N. Kolmogrov in the early part of 20th century. We explain this approach in section 4.5.

4.5 Axiomatic Definition of Probability:

Let the sample space Ω of a random experiment be considered. Let to be the collection of all subsets of Ω. Let P be a function to [0, 1] of the real line. such a trans formation P is called a set function. The set function P is called a probability set function or simply probability if it obeys the following rules - (also called axioms).

(i) $0 \le P(A) \le 1 \forall A \in C$

(ii) $P(\Omega) = 1, \; P(\phi) = 0$

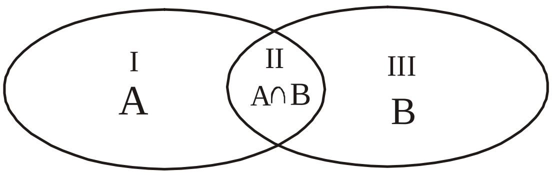
(iii) $P\left(\bigcup_{i=1}^{n} Ai\right) = \sum_{i=1}^{n} P(Ai)$ where $Ai \cap Aj = \phi$ for all $i \neq j$

4.6 Addition Theorem:

For any two events A and B

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

Proof : Let the events A and B be represented as the sets show in the figure.



The regions I, II, III in the figure are mutually disjoint.

Also $I \cup II \cup III = A \cup B$

$I \cup II = A$

$II \cup III = B$

$\therefore$ by the third rule in the axiomatic definition of the probability we get the following identities.

$$P(A \cup B) = P(I) + P(II) + P(III) \cdots\cdots\cdots\cdots (1)$$

$$P(A) = P(I) + P(II) \cdots\cdots\cdots (2)$$

$$P(B) \ = P(II) + P(III) \cdots\cdots\cdots\cdots (3)$$

Subtracting the sum of equations (2) and (3) from (1) we set

$$P(A \cup B) - P(A) - P(B) = -P(II) \cdots\cdots\cdots\cdots (4)$$

But region II is $A \cap B$

$\therefore$ equation (4) becomes

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

4.7 Examples:

1. Prove that the probability of obtaining a total of 7 in a single thron with two dice is 1/9.

Sol. When two dice are thrown once we get a sample space contains 36 pairs points givenly

$$\{(i, j)/i = 1, 2, 3, 4, 5, 6; j = 1, 2, 3, 4, 5, 6\}$$

The total of the two digits that appear is i + j. When we want i + j = 9, the pairs (3,6), (4,5), (5,4), (6,3) are favour of giving a total 9.

$\therefore$ four points in the sample space of 36 points are in favour of getting a total '9'.

Hence by classical definition the probability = $\frac{4}{36} = \frac{1}{9}$.

2. Show that in a single throw with two dice, the chance of throwing more than 7 is equal to that of throwing lessthan 7.

Sol. Of the (36) poissible pairs the following pairs give the totals written against the repective groups of pairs.

(1,1) : 2

(1,2) ; (2,1) : 3

(1,3), (2,2), (3,1) : 4

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

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

(2,6), (3,5), (4,4), (5,3), (6,2) : 8

(3,6), (4,5), (5,4), (6,3) : 9

(4,6), (5,5), (6,4) : 10

(5,6), (6,5) : 11

(6,6) : 12

In the above the number of pairs that give a total of 2 or 3 or 4 or 5 or 6 i.e. a total less than 7 is 15. Hence probability of getting a total less than 7 is $\frac{15}{36}$ similarly the total number of pairs to set a total of 8 or 9, or 10 or 11 or 12 i.e. a total of more than 7 is 15.

$\therefore$ The probability of setting a total of more than 7 is also $\frac{15}{36}$, because these two probabilities are equal the result follows.

3. In a single throw with two dice what is the number such that it is a total with minimum probability.

Ans: We know that when two dice are thrown once weget the following totals with corresponding pairs of results.

Total	**Pairs of Observations**
2	(1,1) = 1 pair of out 36 pairs
3	(1,2), (2,1) = 2 Pairs out of 36 pairs
4	(1,3), (2,2), (3,1) = 3 Pairs out of 36 pairs
5	(1,4), (2,3), (3,2), (4,1) = 4 Pairs out of 36 pairs
6	(1,5), (2,4), (3,3), (4,2), (5,1) = 5 Pairs out of 36 pairs
7	(1,6), (2,5), (3,4), (4,3), (5,2) = 5 Pairs out of 36 pairs
8	(2,6), (3,5), (4,4), (5,3), (6,2) = 5 Pairs out of 36 pairs
9	(3,6), (4,5),(5,4), (6,3) = 4 Pairs out of 36 pairs
10	(4,6), (5,5), (6,4) = 3 Pairs out of 36 pairs
11	(5,6), (6,5) = 2 Pairs out of 36 pairs
12	(6,6) = 1 Pairs out of 36 pairs

The probability of setting a total 2 is $1/36$.

Also the probabilites getting a total 6 is $1/36$

Which isthe minimum probability

4. Two different digits are chosen at random from the set 1,2,3,........,8. Show that the probability that the sum of the digits will be equal to 5 is the same asthe probability that their sum will exceed 13. Also show that the chance of both digits exceeding 5 is $3/28$.

Ans : If both the digits were to add up to 5, they should be one of the following pairs (1,4), (2,3). Two different digits our of 8 can be taken in ${}^8C_2$ ways.

$\therefore$ probability of getting 5 as the total = $\frac{2}{{}^8C_2} = \frac{2\times1\times2}{8\times7}$

$$= \frac{1}{14}$$

In order to get a total of more than 13, the required pairs are (8,6), (8,7).

Hence the required probability is again $\frac{2}{{}^8C_2} = \frac{1}{14}$

Hence the two probabilitis are the same

The chance of both the digits exceeding 5 is as follows. The pairs should be from 6,7,8 which can be in ${}^3C_2$ ways.

Hence the required probability $= \frac{{}^3C_2}{{}^8C_2} = \frac{3\times2}{8\times7} = \frac{3}{28}$

5. Four persons are chosen at random from a group containing 3 men, 2 women and 4 children, show that the chance that exactly two of them will be children is $\frac{10}{21}$

Ans : Total number of persons in the group = 9

4 persons out of these 9 can be drawn in ${}^9C_4$ ways.

Among these 2 should come from the 4 children and 2 should come from out of 5 persons of a mixture of men and women.

These two can be done ${}^4C_2 \times {}^5C_2$ ways.

Hence the required probability by applying classical definition = $\frac{{}^4C_2 \times {}^5C_2}{{}^9C_4}$

$$= \left(\frac{4\times3}{1\times2}\times\frac{5\times4}{1\times2}\right) \div \frac{9\times8\times7\times6}{1\times2\times3\times4}$$

$$= \frac{6\times10}{9\times7\times2} = \frac{60}{126} = \frac{10}{21}$$

6. Four tickets marked 00, 01, 10, 11 respectively are placed in a bag. A ticket is drawn at random five times being replaced each time. Find the probability that the sum of the numbers or the tickets drawn is 23.

Ans : The problem is equivalent to finding how many times each of the tickets 00, 01, 10, 11 out of 5 draws are to be obtained to get a total of 23 for the numbers 00, 01, 10, 11.

It can be earily seen that

00 : once, 01 : twice, 10 : once, 11 : once

or

00 : twice, 01 : once, 10 : none , 11 : twice

or

00 : none, 0.1 : twice, 10 : twice, 11 : none

Only will given a total of 23 in 5 draws.

The probabilities according to these are

$$\left(\frac{1}{5}\times\frac{2}{5}\times\frac{1}{5}\times\frac{1}{5}\right)+\left(\frac{2}{5}\times\frac{1}{5}\times\frac{2}{5}\right)+\left(\frac{3}{5}\times\frac{2}{5}\right)$$

$$=\frac{2}{25}\left[\frac{1}{25}+\frac{2}{5}+3\right]=\frac{2}{25}\times\frac{(1+10+75)}{25}$$

$$=\frac{2\times 86}{25\times 25}=\frac{172}{625}$$

7. If A, B and C are three events express the following in appropriate symbols.

(a) Simuntaneous occurence of A, B and C

(b) Occurence of at least one of them

(c) A, B and C are mutualy exclisive events

(d) Every point of A is contained in B

(e) The event 'B' but not A occurs

Ans : (a) $A\cap B\cap C$

(b) $A\cup B\cup C$

(c) $A\cap B\cap C=\phi$

(d) $A\subset B$

(e) $A^1B=B-A$

8. A sample space S contains four points x_1, x_2, x_3 and x_4 and the value of a set function P(A) are known for the following sets

$$A_1 = (x_1, x_2),\ A_2 = (x_3, x_4)$$

$$A_3 = (x_1, x_2, x_3),\ A_4 = (x_2, x_3, x_4)$$

$$P(A_1) = \frac{4}{10},\ P(A_2) = \frac{6}{10},\ P(A_3) = \frac{4}{10},\ P(A_4) = \frac{7}{10}$$

(i) Find the total number of sets including the null set

(ii) Although the set contains no sample point has the probability, bring out an example to show that the converse is not true.

Ans : (i) We have $\Omega = \{x_1, x_2, x_3, x_4\}$

Ω contains 4 ponts. The sub sets along with the number of points in each set is as follows.

null set = $\{\phi\}$ = non of points / sets one

single ton sets = $\{x_1\}, \{x_2\}, \{x_3\}, \{x_4\} = 4$

number of sets having two elements = ${}^4C_2 = 6$

number of sets with 3 elements = ${}^4C_3 = 4$

number of sets having all the fourth = 1

Total number of sets = 1 + 4 + 6 + 4 + 1 = 16

(ii) Consider the set of points that contain elements of A_3 but not of A_1

i.e. $A_3 - A_1 = \{x_3\}$

i.e. $A_3 - A_1$ is not a null set

since $A_1 \subset A_3$ we know that $P(A_3 - A_1) = P(A_3) - P(A_1)$

$$= \frac{4}{10} - \frac{4}{10} = 0$$

i.e. from the given information $P(A_3 - A_1) = 0$

But $A_3 - A_1$ is not a null set. Hence this is an example to show that zero probability doesnot imply that the set whole considaration not necessarily a null set.

9. If A and B are two muthally exclusive events and $P(A)=\frac{1}{2}, P(B)=\frac{1}{3}$ find $P(A \cup B),\ P(A \cap B)$.

Ans : Since A and B are mutually exclusive events.

$A \cap B = \phi$

$P(A \cap B) = 0$

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

$\therefore \quad = \frac{1}{2} + \frac{1}{3} = \frac{5}{6}$

$\therefore \quad P(A \cup B) = \frac{5}{6}$

$P(A \cap B) = 0$

10. A man forget the last digit of a telephone number and dails the last digit at random. What is the probability of not calling more than 3 wrong numbers.

Ans : The last digit would be one omons the 10 digits 0,1,2,3,4,5,6,7,8,9.

Probability of choosing a correct digit $= \frac{1}{10}$

Hence probability of choosing a wrong digit $= 1 - \frac{1}{10}$

$= \frac{9}{10}$

The probability of not making more than three wrong calls is obetained as follows. In a series of 10 trails, either the first or the second or the third call should be a correct call. It is explained as follows :

(i) First call may be correct call

(or)

(ii) First wrong and second correct call

(or)

(iii) First and second one wrong calls and the third is correct calls

Probability according to (i) $=\frac{1}{10}=\frac{1}{10}$

Probability according to (ii) $=\frac{9}{10}\times\frac{1}{10}=\frac{9}{100}$

Probability according to (iii) $=\frac{9}{10}\times\frac{9}{10}\times\frac{1}{10}=\frac{81}{1000}$

Hence required probability $=\frac{1}{10}+\frac{9}{100}+\frac{81}{1000}$

$$=\frac{271}{1000}$$

11. Cet A and B be two possible out comes of an expairment and suppose that P(A) = 0.4, $P(A \cup B)=0.7$ and $P(B)=p$.

For what choice of p are A and B muthally exclusive.

Ans : If A and B were to be mutually exclusive then $P(A \cup B)=P(A)+P(B)$

i.e. 0.7 = 0.4 + p

$\therefore$ p = 0.7 - 0.4 = 0.3

12. Suppose A and B are any two events and $P(A)=p_1$, $P(B)=p_2$, $P(A \cap B)=p_3$ prove the following identities.

(i) $P(\overline{A} \cup \overline{B})=1-p_3$

(ii) $P(\overline{A} \cap \overline{B})=1-p_1-p_2+p_3$

(iii) $P(A \cap \overline{B})=p_1-p_3$

(iv) $P(\overline{A} \cap B)=p_2-p_3$

(v) $P(\overline{A \cap B})=1-p_3$

(vi) $P(\overline{A} \cup B)=(1-p_1+p_3)$

(vii) $P(\overline{A \cup B})=1-p_1-p_2+p_3$

(viii) $P(\overline{A} \cap (A \cup B))=p_2-p_3$

(ix) $P(A \cup (\overline{A} \cap B))=p_1+p_2-p_3$

Ans : From the results of set then we know that

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

$\therefore P(\overline{A}\cup\overline{B})=P(\overline{A\cap B})=P(A\cap B)=1-p_3$

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

i.e. $P(\overline{A}\cap\overline{B})=1-p_1-p_2+p_3 \quad \overline{1-P(A\cup B)}$

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

(iii) $A\cap\overline{B}$ can be interpreted as follows :

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

Also $A\cap\overline{B}$ and $A\cap B$ are muthally exclusive events.

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

i.e. $p_1=P(A\cap\overline{B})+p_3$

i.e. $P(A\cap\overline{B})=p_1-p_3$

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

Also $\overline{A}\cap B$, $A\cap B$ are mutually exclusive

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

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

$P(\overline{A}\cap B)+p_3=p_2$

$\therefore P(\overline{A}\cap B)=p_2-p_3$

(v) $P(\overline{A\cap B})=1-P(A\cap B)=1-p_3$

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

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

$=1-(p_1-p_3)$ from (iii)

$=1-p_1+p_3$

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

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

$= 1 - [p_1 + p_2 + p_3]$

$= 1 - p_1 - p_2 + p_3$

(viii) Consider $\overline{A} \cap (A \cup B)$

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

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

from (iv) we know that $P(\overline{A} \cap B) = p_2 - p_3$

$\therefore P[\overline{A} \cap (A \cup B] = p_2 - p_3$

(ix) Consider $A \cup (A \cap B)$

Applying addition law of probabilities to the sets A and $\overline{A} \cap B$ we get

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

$= P(A) + P(\overline{A} \cap B)$ since A, $\overline{A} \cap B$ are dispoint set s

$= p_1 + (p_2 - p_3) = p_1 + p_2 - p_3$

(because we know from (iv) that $P(\overline{A} \cap B) = p_2 - p_3$

13. Two six fased unbaised dice are thrown. Find the probability that the sum of the numbes shown is 7 or their product is 12.

Ans : The sample space of int comes when two unbioned dice are thrown once can be represented in the following matrix form :

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

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

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

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

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

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

The set of pairs that give a total 7 is

A = {(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)} : 6 events

The set of pairs that give a produt 12 is

B = {(2,6), (3,4) (6,2), (4,3)} : 4 Elements

$A \cap B = \{(3,4), (4,3)\}$: 2 elements

Total points in the sample space = 36 elements.

We konw that $P(A) = \frac{6}{36}$

$P(B) = \frac{4}{36}$

$P(A \cap B) = \frac{2}{36}$

Probability of setting a total 7 or a product 12 for the numbers on the two dice.

$$= P(A \cup B) = P(A) + P(B) - P(A \cap B) = \frac{6}{36} + \frac{4}{36} - \frac{2}{36} = \frac{8}{36} = \frac{2}{9}$$

14. Defects are Clasified as A, B or C and the following probabilities have been determined from available production data.

$$P(A) = \cdot 20,\ P(B) = 0.16,\ P(C) = 0.14,\ P(A \cap B) = 0.08$$

$$P(B \cap C) = 0.04,\ P(A \cap C) = 0.05,\ P(A \cap B \cap C) = 0.02$$

What is the probability that a randomly selected lot exhibits at least one one type of defect? What is the probability that it exhibits A and B defects but is free from type C defect.

Ans : Probability that the selected product to exhibit at least one defect

$$P(A \text{ or } B \text{ or } C) = P(A \cup B \cup C)$$

(i) We know that $P(A \cup B \cup C) = P(A) + P(B) + P(C)$

$-P(AiB) \cdot P(B \cap C) - P(A \cap C)$

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

= 0.20 + 0.16 + 0.14 - 0.08 - 0.04 - 0.005 + 0.02

= 0.35

(ii) Probability that the product exhibits both A and B defects but is free from type C defect

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

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

Also the events $A \cap B \cap C$ and $A \cap B \cap \overline{C}$ are muthally exclusive. Hence

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

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

$$0.02 + P(A \cap B \cap \overline{C}) = 0.08$$

$$\therefore P(A \cap B \cap \overline{C}) = 0.08 - 0.02 = 0.06$$

22. Out of 110 students interviewed at a job fair, 22 were taking a finance couse, 20 were taking an accounting course. One of these 110 students is selected at random. What is the probability that the student is

(a) not taking an accounting course

(b) Taking both a finance and an accounting course

(c) neither taking a finance course nor an accounting course

Let A, F respectively denote the events that the student takes an accounting, finance course respectively. Given that

$$P(A) = \frac{20}{110},\ P(F) = \frac{22}{110},\ P(A \cup F) = \frac{30}{110}$$

(a) We are asked to find $P(\overline{A}) = 1 - p(A) = 1 - \frac{20}{110} = \frac{110 - 20}{110}$

$$= \frac{90}{110} = \frac{9}{11}$$

(b) We are asked to find $P(A \cap F)$

We know that $P(A \cup F) = P(A) + P(F) - P(A \cap F)$

i.e. $$\frac{30}{110} = \frac{20}{110} + \frac{22}{110} - P(A \cap F)$$

(with $\frac{22}{110}$ cancelled to $\frac{11}{55}$)

i.e. $$P(A \cap F) = \frac{2}{11} + \frac{11}{55} - \frac{3}{11}$$

$$= \frac{11}{55} - \frac{1}{11} = \frac{11}{55} - \frac{5}{55} = \frac{6}{55}$$

(c) We are asked to find $P(\overline{A} \cap \overline{F})$

We know that $\overline{A} \cap \overline{F} = \overline{(A \cup F)}$

$$\therefore P(\overline{A} \cap \overline{F}) = P(\overline{A \cup F}) = 1 - P(A \cup F)$$

$$= 1 - \left\{\frac{30}{110}\right\} = \frac{80}{110} = \frac{8}{11}$$

4.8 Exercises:

1. Two cards are drawn at randim from a well shuffled pack of cards show that the probability of drawing two aces is $\frac{1}{221}$.

2. Among the digits 1,2,3,4,5 at first one is closen and then a second selection is made among we remaining four digits. Assuming that all 20 possible out comes have equal probabilite find the probability that an odd digit will be selected (i) the first time, (ii) the second time, (iii) both times.

   (Ans: (i) $\frac{3}{5}$, (ii) $\frac{3}{5}$, (iii) $\frac{3}{10}$)

3. A committee of 4 people is to be appointed from 3 offians of the production department, 4 officers of the purchase department, two officers of the sales department and 1 charted accountant. Find the probability of formming in the following manner.

   (i) There must be one from each catesam

   (ii) It should have at least one from the purchase department

(iii) The chartered accountant must be in the committee

(Ans : (i) $\frac{8}{70}$, (ii) $\frac{13}{14}$, (iii) $\frac{2}{5}$)

4. Each coefficient of the equation $ax^2 + bc + c = 0$ is dethmined by throwing an ordinary die. Find the probability that the equation will have real roots.

(Ans : $\frac{43}{216}$)

5. Out of (2n + 1) tickets conseclrtinady numbered three are drawn at random. Find the chance that the numbers are in A.P.

(Ans : $3n/(4n^2 - 1)$)

6. A, B and C are three ordinary events. Find expression for the events noted below in the context of A, B and C (i) only A occurs, (ii) Both A and B but not C occur, (iii) All three events occur, (iv) At one occurs, (v) At least two occur, (vi) one and no more occurs, (vii) Two and no more occur, (viii) none occurs.

(Ans : (i) $A \cap \overline{B} \cap \overline{C}$, (ii) $A \cap B \cap \overline{C}$, (iii) $A \cap B \cap C$, (iv) $A \cup B \cup C$

(v) $(A \cap B \cap \overline{C}) \cup (A \cap \overline{B} \cap C) \cup (\overline{A} \cap B \cap C) \cup (A \cap B \cap C)$

(vi) $(A \cap \overline{B} \cap \overline{C}) \cup (\overline{A} \cap B \cap \overline{C}) \cup (\overline{A} \cap \overline{B} \cap C)$

(vii) $(A \cap B \cap \overline{C}) \cup (\overline{A} \cap B \cap C) \cup (A \cap \overline{B} \cap C)$

(viii) $\overline{A} \cap \overline{B} \cap \overline{C}$ or $\overline{A \cup B \cup C}$

7. If two dice are thrown what is the probability that sumes (a) greater than 8, (b) neither 7 nor 11.

(Ans : (a) $\frac{5}{18}$, (b) $\frac{7}{9}$)

8. A box contains 6 red, 4 white and 5 black balls. A person draws 4 balls from the box at random. Find the probability that among the balls drawn there is at least one ball of each colour.

(Ans : 0.5275)

9. IF $A \cap B = \phi$ then show that $P(A) \leq P(B)$

10. If A and B are two events such that

$P(A) = \frac{3}{4}, P(B) = \frac{5}{8}$ show that

(a) $P(A \cup B) \geq \frac{3}{4}$, (b) $\frac{3}{8} \leq P(A \cap B) \leq 5/8$

11. A special dice is prepared such that the probabilities of throwing 1, 2, 3, 4, 5, 6 are respectively

$\frac{1-k}{6}, \frac{1+2k}{6}, \frac{1-k}{6}, \frac{1+k}{6}, \frac{1-2k}{6}, \frac{1+k}{6}$ respectively.

If two such dice are thrown find the probability of getting a sum equal to 9.

(Ans : $(1+k)(2-3k)/18$)

12. In these of probabilities $P(A) - p_1. P(B) = p_2, P(A \cap B) = p_3$

Express $P(A \cup B)$ $P(\overline{A} \cap B)$ under the condition that A and B are mutually exclusive.

13. Let A and B be the possible and comes of an experiments and suppose P(A) = 0.4, $P(A \cup B) = 0.7$, P(B) = p. For what choice of o are A and B be muthally exclusive ?

(Ans : p = 0.3)

4.9 Summary:

The concept 'probability' is defined in various ways starting from the classical definition to the most moden way the axiomatic approach. Some general laws of probability upto the concept of additive law for two and more than two events are estabilished besides showing the applications of these laws in a number of examples. Some exercices in the answers one also provided for the students to try on their own.

4.10 Technical Terms:

Relative Frequency, Random Experment, Sample Space, Simple Event, Compound Event, Axious Operations on sets.

Lesson - 5

CONDITIONAL PROBABILITY

Syllabus:

Conditional Probability, multiplication rule of probability for n events, Broole's inequality, independence of events, Baye's Theorem and its applications (with examples in real life).

Objective:

After studying this lesson the student is expected to have a clear notion of probabilities of dependent events. Its application in making decision about conditional events and the principle be find BAye's Theorem in assessing the performance of devices with builtin structural probability.

Structure of the lesson:

5.1 Introduction:

In the theory of probability if we consider the probability of occurence of more than one event in succession some times the sequence of order in which the events occur makes a difference and some times it will not make any difference. For example from a box containging '9' cards of identical size marked with the digits 1, 2, 3, 4, 5, 6, 7, 8, 9, let us draw two cards one after the other. This is suggested in two ways.

(i) The card drawn in the 1st draw is placed back into the box before the second draw.

(ii) The card drawn in the first draw is not placed back into the box before the second draw.

According to the first scheme the probability of drawing '9' in the second draw will be the same what ever may be the result of the first draw, where as according to the second scheme, probability of drawing 9 in the second draw depends on the result of the first draw. The second scheme gives rise to the notion of conditional probabilities. In this lesson we discuss the need for conditional probability, its definition, independent events, applications, the Bayer's Theorem its importance in evaluating probabilities.

5.2 Conditional Probability:

As defined in lesson - 4, let us consider a probability space. (Ω, C, p). Let A and B be any two subsets of Ω. Then the conditional probability of occurrence of A after the occurence of B is demoted by $P\left(A/B\right)$. (to be read as probability of A given B). It is defined as

$$P\left(A/B\right) = \frac{P\left(A \cap B\right)}{P\left(B\right)}.$$

For the above definition to be valid, $P(B) \neq 0$. Similarly the conditional probability of occurence of B after the occurence of A is denoted by $P\left(B/A\right)$ and is defined as

$$P\left(B/A\right) = \frac{P\left(B \cap A\right)}{P\left(A\right)} \qquad \text{where } P(A) \neq 0$$

Since $A \cap B$ is same as $B \cap A$ we can write that

$$P\left(A/B\right) = \frac{P\left(A \cap B\right)}{P\left(B\right)}, \; P(B) \neq 0$$

$$P\left(B/A\right) = \frac{P\left(A \cap B\right)}{P\left(A\right)}, \; P(A) \neq 0.$$

Example:

A bag contains 10 gold and 8 silver coins. Two successive draws of one coin in each draw are made such that the coin drawn in the first draw is not replaced before the second draw is made. Find the probability that both the draws give gold coins.

Let A, B be the events of drawing a gold coin in the first draw and second draw respectively we are to find $P\left(A \cap B\right) = P(A) \cdot P\left(B/A\right)$.

We know that, $P(A) = \frac{10}{18}$.

$P(B/A) =$ probability of drawinga gold coin in the second draw given that a gold coin is drawn in the first draw. Since the coin drawn is not replaced we will have a total of 17 coins of which 9 would be gold and hence the probability of drawing a gold coin in the second draw given that a gold coin is drawn in the first draw $= \frac{9}{17}$.

i.e. $$P(B/A) = \frac{9}{17}$$

$$P(A \cap B) = P(A) \cdot P(B/A) = \frac{10}{18} \cdot \frac{9}{17} = \frac{5}{17}$$

5.3 Multiplication Rule:

Let A, be any two events such that $P(A) \neq 0$, $P(B) \neq 0$. Then by definition we know that

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

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

By cross multiplication we get

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

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

These two equations are called multiplication rule of probability for two events. We can establish multiplication rule of probability for n events, as a theorem.

5.3.1 Theorem: If $A_1, A_2, \cdots\cdots, A_n$ are events. Then

$$P(A_1 \cap A_2 \cap A_3 \cap \cdots \cap A_n) = P(A_1) \cdot P(A_3/A_1 \cap A_2) \cdots P(A_n/A_1 \cap A_2 \cap \cdots \cap A_{n-1})$$

Proof: we prove the result by the principle of mathematical induction. It is obvious that the minimum number of events for the definition of conditional probability is 2. Therefore in our theorem $n \geq 2$. For two events the statement of the theorem is

$P(A_1 \cap A_2) = P(A_1) \cdot P(A_2/A_1)$ and this follows from the definition and cross multiplication of conditional probability.

Let $n = 3$. Then L.H.S. is

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

But $P(A_1 \cap B) = P(A_1)\, P(B/A_1)$ since the statement is time for two events.

i.e. $$P(A_1 \cap A_2 \cap A_3) = P(A_1) \cdot \left[P\left\{(A_2 \cap A_3)/A_1\right\}\right]$$

$$= P(A_1) \cdot \left[\frac{P(A_2 \cap A_3 \cap A_1)}{P(A_1)}\right]$$

$$= P(A_1) \cdot \left[\frac{P(A_2 \cap A_3 \cap A_1)}{P(A_1 \cap A_2)} \cdot \frac{P(A_1 \cap A_2)}{P(A_1)}\right]$$

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

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

Hence the result is proved for 3 events.

In a similar maner suppose the result istrue for $n = k$.

i.e.

$$P(A_1 \cap A_2 \cap A_3 \cdots \cap A_k) = P(A_1) \cdot P(A_2/A_1) \cdot P(A_3/A_1 \cap A_2) \cdots P(A_k/A_1 \cap A_2 \cap \cdots A_{k-1})$$

we shall prove it for $n = k+1$

consider

$$P(A_1 \cap A_2 \cap A_3 \cap \cdots\cdots \cap A_{k+1})$$

$$= P(B \cap A_{k+1}) \text{ where } (B = A_1 \cap A_2 \cap \cdots\cdots \cap A_k)$$

i.e. $$P(B \cap A_{k+1}) = P(B) \cdot P(A_{k+1}/B)$$

$$= P(A_1 \cap A_2 \cap A_3 \cap \cdots\cdots \cap A_k) \cdot P(A_{k+1}/A_1 \cap A_2 \cap \cdots\cdots A_k)$$

Since we have assumed that the result is true for $n = k$ the above becomes

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

$$\cdot P(A_k/A_1 \cap A_2 \cap \cdots\cdots \cap A_{k-1}) \cdot P(A_{k+1}/A_1 \cap A_2 \cdots\cdots \cap A_k)$$

Therefore the result if true for any natural number 'n' by the principle of mathematical induction.

5.4 Boole's Inequality:

It $A_1, A_2, \cdots\cdots\cdots, A_n$ are n events then

$$\text{(i)} \qquad P\left(\bigcap_{i\equiv 1}^{n} A_i\right) \geq \sum_{i=1}^{n} P(A_i) - (n-1)$$

$$\text{(ii)} \qquad P\left(\bigcup_{i\equiv 1}^{n} A_i\right) \leq \sum_{i=1}^{n} P(A_i)$$

Proof: We shall prove both the inequalities by the principle of mathematical induction.

(i) suppose $n = 2$, then by addition theorem for two events, we know that

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

Since L.H.S. is a probability it is ≤ 1

$\therefore$ R.H.S. is also ≤ 1 i.e.

$$P(A_1) + P(A_2) - P(A_1 \cap A_2) \leq 1$$

$\therefore$ $P(A_1 \cap A_2) \geq P(A_1) + P(A_2) - 1 = P(A_1) + P(A_2) - (2-1)$

Hene we have $P(A_1 \cap A_2) \geq P(A_1) + P(A_2) - (2-1)$

i.e. the result is true for $n = 2$

Suppose the result is true for $n = k$

i.e. $P(A_1 \cap A_2 \cap A_3 \cap \cdots\cdots \cap A_k) = P(A_1) + P(A_2) + \cdots\cdots + P(A_k) - (k-1)$

Let us shown that the result is true for $n = k+1$. The L.H.S. of the 'to be shown' experssion is

$P(A_1 \cap A_2 \cap \cdots\cdots \cap A_k \cap A_{k+1}) = P(B \cap A_{k+1})$ where

$$B = A_1 \cap A_2 \cap A_3 \cap \cdots\cdots \cap A_k$$

Applying the result fortwo events $B, \; A_{k+1}$ we get

$$P(B \cap A_{k+1}) \geq P(B) + P(A_{k+1}) - 1$$

i.e. $P(A_1 \cap A_2 \cap \cdots\cdots \cap A_{k+1}) \geq P(A_1 \cap A_2 \cap \cdots \cap A_k) + P(A_{k+1}) - 1$

In the R.H.S. of the above $P(A_1 \cap A_2 \cap \cdots\cdots \cap A_k)$ is greater than or equal to $P(A_1) + P(A_2) + \cdots\cdots + P(A_k) - (k-1)$ since we assumed that the result is true for k events.

Hence the above inequality be comes

$$P(A_1 \cap A_2 \cap \cdots\cdots \cap A_{k+1}) \geq P(A_1) + P(A_2) + \cdots\cdots + P(A_k) - (k-1) + P(A_{k+1}) - 1$$

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

$$P(A_1 \cap A_2 \cap \cdots\cdots \cap A_{k+1}) = \sum_{i=1}^{k+1} P(A_i) - k$$

Hence by the principle of mathematical induction theresult is for any natural number.

5.5 Independence of Events:

5.5.1 Definition: Two events A, B are said to be statistically independent if the probability of their joint occurence is same as the product of the probabilities of their individul occurances.

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

5.5.2 Definition: In the case of three events A_1, A_2, A_3 the concepts of independence is of two types.

Pairwise independence and mutual independence.

If the events are independent taken two at time a we say that they are pairwise independent.

In the case of three events this means

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

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

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

In addition to this if the events are independent taken all at a time (in the case of three events).

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

If all the above fourconditions are true we say that a set of three events A_1, A_2, A_3 are mutually independent.

In general if we have n events say $A_1, A_2, \cdots\cdots, A_n$ we say that hese are pair wise independent if

$$P(A_i \cap A_j) = P(A_i) P(A_j) \text{ for all } i \neq j \qquad i, j = 1, 2, 3, \cdots\cdots, n$$

These conditions are n_{C_2} in number.

These n events shall be mutually independent if in addition to the above n_{C_2} conditions the following are also true.

$$P(A_i \cap A_j \cap A_k) = P(A_i)\, P(A_j)\, P(A_k) \quad i \neq j \neq k$$

$$P(A_i \cap A_j \cap A_k \cap A_\ell) = P(A_i)\, P(A_j)\, P(A_k) P(A_\ell) \quad i \neq j \neq k \neq \ell$$

$$\cdots\cdots\cdots\cdots$$

$$P(A_1 \cap A_2 \cap \cdots\cdots \cap A_n) = P(A_1) P(A_2) \cdots\cdots P(A_n)$$

These sets of conditions are respectively $n_{C_2}, n_{C_3}, \cdots\cdots, n_{C_n}$ in number.

Hence total number of conditions required for the mutual independence of n events is $n_{C_2} + n_{C_3} + \cdots\cdots + n_{C_n} = (1+1)^n - n_{C_1} - n_{C_0} = 2^n - n - 1$.

On the other hand number of conditions required for pairwise independence of n events is only $^{n}C_2$. This is true for $n \geq 2$.

It can be seen that mutually independent events are always passwise independent while the converse is not true as can be explained by the following example.

5.5.3 Example: Consider box containing 4 cards marked wit the digits 100, 010, 001, 111. Let A, B, C - be the events representing drawing a card at random with 1 in hundredth place, 1 in 10th place, 1 in 1st place respectively. Then it can be seen that

$$P(A) = \frac{2}{4} = 1/2$$

$$P(B) = 1/2$$

$$P(C) = 1/2$$

$$P(A \cap B) = \frac{1}{4} = P(A) \cdot P(B) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$$

$$P(B \cap C) = \frac{1}{4} = P(B) \cdot P(C) \quad \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$$

$$P(A \cap B \cap C) = \frac{1}{4} \neq \frac{1}{8} = P(A) \cdot P(B) \cdot P(C)$$

Hence the events A, B, C are pairwise independent but not mutually independent.

5.5.4 Example: If A nad B are independent then A and $\overline{B}$, $\overline{A}$ and $\overline{B}$ are also independent where $\overline{A}$, $\overline{B}$ are complementary events for A, B respectively.

Solution: Given $P(A \cap B) = P(A) \cdot P(B)$

To show that $P(A \cap \overline{B}) = P(A) \cdot P(\overline{B})$

We know that $P(\overline{B}) = 1 - P(B)$

Multiplying with $P(A)$ we get

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

$$= P(A) - P(A \cap B) \cdots\cdots (1) \quad \because A, B \text{ are independent}$$

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

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

$$= P\left(A \cap \overline{B}\right)$$ using propertis of sets

In a similar way we can prove that

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

5.5.5 Example: Given that $P\left(A_1 \cup A_2\right) = \frac{5}{6}$, $P\left(A_1 \cap A_2\right) = \frac{1}{3}$, $P\left(\overline{A}_2\right) = \frac{1}{2}$. Find $P\left(A_1\right)$, $P\left(A_2\right)$. Hence show that A_1, A_2 are independent.

Solution:

$$P\left(\overline{A}_2\right) = \frac{1}{2} \Rightarrow P\left(A_2\right) = 1/2$$

we know that

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

$$\frac{5}{6} = P\left(A_1\right) + \frac{1}{2} - \frac{1}{3}$$

$$\Rightarrow P\left(A_1\right) = \frac{5}{6} - \frac{1}{2} + \frac{1}{3} = \frac{4}{6} = \frac{2}{3}$$

$$P\left(A_1\right) \cdot P\left(A_2\right) = \frac{2}{3} \cdot \frac{1}{2} = \frac{2}{6} = \frac{1}{3}$$

Given that

$$P\left(A_1 \cap A_2\right) = \frac{1}{3}$$

Hence

$$P\left(A_1 \cap A_2\right) = P\left(A_1\right) \cdot P\left(A_2\right)$$

$\therefore$ A_1, A_2 are independent.

5.6 Baye's Theorem:

Let $E_1, E_2, \cdots\cdots\cdots, E_n$ be mutually disjoint exhaustive events with $P(E_i) \neq 0$. If A is any event that can occur with any of $E_1, E_2, \cdots\cdots\cdots, E_n$ then

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

The L.H.S. is called posterior or inverse probability. In the numerator of R.H.S. namely $P(E_i)$ is prior probability.

Proof:

Given that $\bigcup_{i=1}^{n} E_i = \Omega$ - the sample space and $E_i \cap E_j = \phi$ for $i \neq j$. $i, j = 1, 2, \cdots\cdots, n$

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

$$P(A) = \sum_{i=1}^{n} P(A \cap E_i) = \sum_{i=1}^{n} P(E_i) \cdot P(A/E_i)$$

Consider $P(A \cap E_i) = P(A)\, P(E_i/A)$

$$\therefore P(E_i/A) = \frac{P(A \cap E_i)}{P(A)} = \frac{P(E_i) \cdot P(A/E_i)}{\sum_{i=1}^{n} P(E_i)\, P(A/E_i)}$$

i.e., $$P(E_i/A) = \frac{P(E_i) \cdot P(A/E_i)}{\sum_{i=1}^{n} P(E_i) \cdot P(A/E_i)}$$

Hence the theorem is proved.

5.6.1 Example: In a factory machines A and B are producing springs of the same type of this production machine A and B produce 5% and 10% defective springs respectively. Machines A and B produce 40 % and 60% of the total output of the factory. One spring is selected at randim and it is found to be defective. What is the probability that this defective spring is produced by machine A.

Solution:

Given $P(A) = \frac{40}{100} = 0 \cdot 4$

$P(B) = \frac{60}{100} = 0 \cdot 6$

$P(D/A) = \frac{5}{100} = 0 \cdot 05$ D: stands for defective

$P(D/B) = \frac{10}{100} = 0 \cdot 1$

we have to find $P(A/D)$

$$P(A/D) = \frac{P(D/A) \cdot P(A)}{P(D/A) \cdot P(A) + P(D/B) \cdot P(B)}$$

$$= \frac{0 \cdot 05 \times 0 \cdot 4}{(0 \cdot 05 \times 0 \cdot 4) + (0 \cdot 1) 0 \cdot 6} = \frac{.020}{.020 + .06}$$

$$= \frac{0 \cdot 020}{0 \cdot 080} = \frac{2}{8} = \frac{1}{4}.$$

5.7 Examples:

5.7.1 Example: Let $S = \left\{1, \frac{1}{2}, \left(\frac{1}{2}\right)^2, \cdots\cdots\cdots, \left(\frac{1}{2}\right)^n\right\}$ be a classical event space. A, B be events given by $A = \left\{1, \frac{1}{2}\right\}$; $B = \left\{\left(\frac{1}{2}\right)^k \middle/ K \text{ is an even positive integer}\right\}$ find $P(\overline{A} \cap \overline{B})$

Solution: We know that $\overline{A} \cap \overline{B} = \overline{(A \cup B)}$

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

$$= 1 - \left\{P(A) + P(B) - P(A \cap B)\right\} \cdots\cdot(1)$$

Given $A = \left\{1, \frac{1}{2}\right\}, B = \left\{\left(\frac{1}{2}\right)^0, \left(\frac{1}{2}\right)^2, \left(\frac{1}{2}\right)^4 \cdots\cdot\right\}$

$$A \cap B = \{1\}$$

$P(A \cap B) = \frac{1}{n+1}$ since S contains $n+1$ elements.

$P(A) = \frac{2}{n+1}$ $\quad P(B) = \frac{n+2}{2(n+1)}$ if n is even

$P(B) = \frac{1}{2}$ if n is odd

Suppose n is even using (1)

$$P\left(\overline{A} \cap \overline{B}\right) = 1 - \left\{\frac{2}{n+1} + \frac{n+2}{2(n+1)} - \frac{1}{n+1}\right\}$$

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

$$P\left(\overline{A} \cap \overline{B}\right) = \frac{2n + 2 - n - 4}{2(n+1)} = \frac{n-2}{2(n+1)}$$

Suppose n is even then

$$P\left(\overline{A} \cap \overline{B}\right) = 1 - \left\{\frac{2}{n+1} + \frac{1}{2} - \frac{1}{n+1}\right\}$$

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

$$= 1 - \frac{n+3}{2(n+1)} = \frac{2n + 2 - n - 3}{2(n+1)} = \frac{n-1}{2(n+1)}$$

5.7.8 Example: If $P(B/\overline{A}) = P(B/A)$, then A and B are independent

Proof: $P(B/\overline{A}) = P(B/A)$

i.e. $$\frac{P(B\cap\overline{A})}{p(\overline{A})} = \frac{P(B\cap A)}{P(A)}$$

i.e. $P(A)\cdot P(B\cap\overline{A}) = P(A\cap B)\, P(\overline{A})$

$= P(A)\cdot \left\{P(B)+P(\overline{A}) - P(\overline{A}\cup B)\right\} = P(A\cap B)\, P(\overline{A})$

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

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

i.e. $P(A)\left\{P(\overline{A}\cup B) - P(B)\right\} = P(\overline{A})\cdot\left\{P(A) - P(A\cap B)\right\}$

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

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

$P(A)\left\{1-P(A\cap B)\right\} = P(B) - P(A)\cdot P(B)$

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

Hence it is proved that A and B are independent.

5.7.9 Example: The chances of X, Y, Z becoming managers of a company are $4:2:3$. The probabilities that bonus sheme will be introduced if X, Y, Z become managers are respectively are $0\cdot 3, 0\cdot 5, 0\cdot 8$. If the bonus scheme has been introduced what is the probability that X is appointed as the manager.

Solution: Given $P(X) = \frac{4}{9}$, $P(Y) = \frac{2}{9}$, $P(Z) = \frac{3}{9}$

Also $P(B/X) = 0\cdot 3$, $P(B/Y) = 0\cdot 5$, $P(B/Z) = 0\cdot 8$.

where B stands for bonus. We have to find $f(X/B)$.

Baye's Theorem gives as

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

$$= \frac{0\cdot 3 \times \frac{4}{9}}{0\cdot 3 \times \frac{4}{9} + 0\cdot 5 \times \frac{2}{9} + 0\cdot 8 \times \frac{3}{9}}$$

$$= \frac{1\cdot 2}{1\cdot 2 + 1\cdot 0 + 2\cdot 4} = \frac{1\cdot 2}{4\cdot 6} = \frac{0\cdot 6}{2\cdot 3} = 0\cdot 218$$

5.7 Exercises:

1. Show that in a single throw with two dice the chance of throwing more than seven is equal to that of throwing less than seven.

2. Two different digits are chosen at random from the set 1, 2, 3, ... 8. Show that the probability that the sum of the digits will be equal to 5 is the same as the probability that their sum will exceed 13 each being $\frac{1}{14}$.

3. A coin is tossed until there are either two consecutive heads or two consecutive tails or the number of tosses become 5. Describe the sample space.

4. Let A and B be two events, neither of which has probability zero. Then show that if A and B are disjoint then A and B are independent.

5. If A and B are two mutually exclusive events then show that $P\left(A/\overline{B}\right) = \frac{P(A)}{1-P\left(\overline{B}\right)}$

6. A man takes advice regarding one of two possible courses of action from three advisers, who arrived at their recommendations independently. He follows the recommendation of the majority the probability that the individual advisers are wrong are 0.1, 0.5 and 0.05 respectively. What is the probability that the man takes incorrect advice.

7. A person takes four tests in succession. There probability of this passing the first test is p, that of his passing each succeeding test is p or $p/2$ according as he passes or fails the preceding one. He qualifies provided he passes atleast three tests. What is the chance of his qualifying.

8. $P(A) = 0\cdot7$, $P(\overline{B}) = 0\cdot5$, $P(\overline{A} \cup \overline{B}) = 0\cdot6$ find $P(A \cup B)$.

9. A certain drug manufactured by a comapny is tested chemically for its toxic nature. Let the event the drug is toxic be denoted by E. The event the chemical test reveals that the drug is toxic be denoted by F. Let $P(E) = \theta$, $P(F/E) = P(\overline{F}/\overline{E}) = 1-\theta$. Then show that probability of the drug not being toxic when the chemical test reveals that it is toxic is $1/2$.

5.9 Summary:

In this lesson an attempt is made to explain the concepts of conditional probability and related aspects along with examples. The most important aspect isthe Baye's Theorem and inverse probabilities. A number of examples are worked out and a good number of exercises are also given.

5.10 Technical Terms:

Dependent events

Independent events

Product law

Pairwise independece

Mutual independence

Prior probability

Posterior / Inverse probability

Mathematical induction

Restricted sample space

Relative probability

Lesson - 6

RANDOM VARIABLES

Syllabus:

Notion of a random variable, distribution function and its properties, discrete random variable, probability mass function, Cnlinum random variable, probability density function, transformation of one dimensional random variable (simple 1 -1 functions only)

Objective of The Lesson:

After studying this lesson the student is expect to have a clear comprehension of the theory and the practical utility about the concepts of Random variable, distribution function - its properties, Discrete random variable, probability mass function, continuous random variable, probability denisty function and transformation of one dimensional random variable.

Structure of The Lesson:

6.1 Notion of Random Variable:

When a statistical experiment is conducted, all the possible outcomes of it will generate a set, called the sample space, denoted by S. We are often much interested in its numerical description rather than its specific outcomes. For example when we toss a coin three times, we get the following outcomes.

TTT, HTT, THT, TTH, HHT, HTH, THH, HHH

which generates a sample space S. Suppose we are interested in number of heads then a numerical value 0, 1, 2, or 3 will be assigned to each sample points. The numerical values 0, 1, 2, 3 are random observations which may be assumed by some random variable, denoted by X. In the example, the random variable X represents the number of heads The value sof different sample points of the sample space S are denoted by W. Also from the example i.e., three tosses of a fair coin the sample points in S are

$$W_1 = (T\ T\ T),\ W_2 = (H\ T\ T),\ W_3 = (T\ H\ T),\ W_4 = (T\ T\ H),\ W_5 = (HHT), W_6 = (H\ T\ H),$$
$$W_7 = (THH),\ \ W_8 = (HHH)$$

Since we define a function X, which denotes the number of heads, we note that

$$X(W_1) = 0,\ X(W_2) = X(W_3) = X(W_4) = 1,\ X(W_5) = X(W_6) = X(W_7) = 2, X(W_8) = 3$$

The inverse image

$X^{-1}(W_i)$ is the event $\{0 \le X < 1\}$ for $i = 1$

$X^{-1}(W_i)$ is the event $\{1 \le X < 2\}$ for $i = 2,3,4$

$X^{-1}(W_i)$ is the event $\{2 \le X < 3\}$ for $i = 5, 6, 7$

$X^{-1}(W_i)$ is the event $\{X = 3\}$ for $i = 8$

Hence X, which denotes the number of heads in three tosses, is called a random variable.

Hence a real valued function defined on a sample space S associated with a given random experiment and taking values in $R(-\infty, \infty)$ is called a random variable.

6.2 DISTRIBUTION FUNCTION AND ITS PROPERTIES

Let X be a random variable then the function $F(x)$ defined for all real x,

$$F(x) = P(X \le x) = P\{w : X(w) \le x\}, -\infty < x < \infty$$

is called the distribution function $(d.f)$ of X.

Properties of Distribution function :

Property 1 : If $F(x)$ is the distribution function of random variable X, and if $x < y$ then (a) $0 \le F(x) \le 1 \forall\ x \in R,\ F$ is bounded (b) $F(x) \le F(y)$, F is monotonically non-decreasing.

Proof :

(a) Since probability is a non-negative quantity and lies between 0 and 1, i.e., $0 \le P \le 1$ therefore we can write $0 \le P(X \le x) \le 1$

$$\Rightarrow 0 \le F(x) \le 1 \qquad (\because P(X \le x) = F(x))$$

(b) As $F(x)$ is a monotonically non-decreasing function of x and x, y be any value in R, such that $x < y$. Since $(-\infty, x]$ is a subset of $(-\infty, y]$ we can write

$$(x, y] = (-\infty, y] - (-\infty, x]$$

$$\therefore P(x, y] = P(-\infty, y] - P(-\infty, x]$$

$$= P(X \le y) - P(X \le x)$$

$$P(x, y] = F(y) - F(x) \cdot \quad \cdot \quad \cdot (1)$$

Also since $P(x, y] \ge 0$ we have

$$F(y) - F(x) \ge 0 \qquad (\because \text{ from (1)})$$

$$\Rightarrow F(x) \le F(y).$$

Property 3: If $F(x)$ is distribution function of random variable X, then

(a). $\lim\limits_{x \to -\infty} F(x) = F(-\infty) = 0$, (b). $\lim\limits_{x \to -\infty} F(x) = F(+\infty) = 1$.

Proof: (a) Let us define that the sequence of events $A_n = \{x \le n\}$. Here the sequence $\{A_n\}$ is a decreasing sequence of events with

$$\lim_{n \to \infty} A_n = \phi \cdot \quad \cdot \quad \cdot (1)$$

Therefore by the continuity axiom on probability we have

$$\lim_{n \to \infty} P(A_n) = P\left(\lim_{n \to \infty} A_n\right) = P(\phi) = 0 \quad (\because \text{ from (1)}) \text{ -------- (2)}$$

But $P(A_n) = P(X \le -n) = F(-n)$ (by defition of d.f.)

$$\therefore \lim_{n\to\infty} P(A_n) = \lim_{n\to\infty} F(-n) = 0$$ ($\because$ from (2))

i.e., $F(-\infty) = 0$

(b) Similarly define that the sequences of events $A_n = \{X \le n\}$, the sequence $\{A_n\}$ is a increasing sequence of events with

$$\lim_{n\to\infty} A_n = S \quad \cdot \quad \cdot \quad \cdot \; (1)$$

Hence by the continuity axiom on probability we have

$$\lim_{n\to\infty} P(A_n) = P\left(\lim_{n\to\infty} A_n\right) = P(S) = 1$$ ($\because$ from (1)) $\cdot \quad \cdot \quad \cdot \; (2)$

$\therefore P(A_n) = P(X \le n) = F(n)$ (by defn. of d.f.)

$$\Rightarrow \lim_{n\to\infty} P(A_n) = \lim_{n\to\infty} F(n) = 1$$ ($\because$ from (2))

i.e., $F(+\infty) = 0$

Property 4 : If $F(x)$ is the distribution function of the random variable X and if $x < y$ then $P(x < X \le y) = F(y) - F(x)$.

Proof : Since the events $x < X \le y$ and $X \le x$ are disjoint and their union is the event $X \le y$.

Hence by addition theorem of probability

$$P(x < X \le y) + P(X \le x) = P(X \le y)$$

$$\Rightarrow P(x < X \le y) = P(X \le y) - P(X \le x) \quad (\because P(X \le x) = F(x) \;\&\; P(X \le y) = F(y))$$

We have $P(x < X \le y) = F(y) - F(x)$

Property 5 : $F(x)$ is continuous from the right i.e., $F(x+0) = F(x)$ for each x.

Proof : Let $A_n = \left(X \le x + \frac{1}{n}\right)$ be a sequence of events and for a fixed value of x, sequence of events and for a fixed value of x, sequence $\{A_n\}$ is a decreasing sequence of events with

$$\lim_{n\to\infty} P(A_n) = P(X \le x) = F(x)$$

or $\lim_{n\to\infty} P\left(X \le x + \frac{1}{n}\right) = F(x)$ $\left(\because A_n = \left(X \le x + \frac{1}{n}\right)\right)$

$\lim_{n\to\infty} F\left(x + \frac{1}{n}\right) = F(x)$

or $F\left(x + \frac{1}{\infty}\right) = F(x)$

or $F(x+0) = F(x)$

6.3 Discrete Random Variable:

If the sample space S contains a finite number of points or countably infinite number of points, it is called a discrete sample space. A random variable X defined over a discrete sample space is called a discrete random variable.

For example if we collect data about number of persons in families of certain town, then it is certain that number of persons in each family would be in whole numbers. Therefore there would be no family with 2.5 or 2.67 or 1.97 persons. The variable i.e., the number of persons in a family, in this case is a discrete random variable - some max examples of discrete random variable are given below.

Example 1 : The number of heads in tossing of a coin.

Example 2 : The number of points on the dice when it is rolled.

Example 3 : The number of insects survived when an insecticide is sprayed.

Example 4 : The number of accidents occured in a year.

Example 5 : The number of defective items in a sample of size 'n'.

6.4 Probability Mass Function:

If 'X' is a discrete random variable defined on the sample space S which takes the values $x_1, x_2, \cdots$ with each possible outcome x_i, then a number is associated that is $p_i = p(x = x_i) = p(x_i)$, called the probability of x_i, the numbers $p(x_i), i = 1, 2, \cdots$ must satisfy the following conditions.

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

the function p is called the probability mass function of the random variable X and the set $\{x_i, p(x_i)\}$ is called the probability distribution of the random variable X.

6.5 Continuous Random Variable:

If sample space S contains an infinite number of points or continuity of points on a line segment or with more than one interval of points is called continuous sample space. A random variable defined over the continuous sample space is called a continuous random variable.

For example 1 : The weight of middle aged people in India lying between 40 Kg and 150 Kg is a continuous variable.

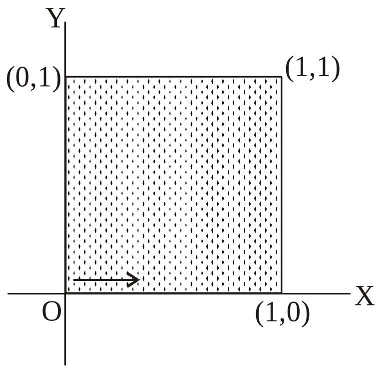
i.e., $X(x)=\{x:40 \le x \le 150\}$

2 : The maximum breaking strength 250 Kg of a wire is a continous variable.

i.e., $X(x)=\{x:0 \le x \le 250\}$

6.6 Probability Density Function:

Since in dealing with the distriction of a continuous random variable will be necessary to express the probabilities in the form of intervals. So with the help of an example continuous probability distribution is explained as follows.



Let us consider a squared target of one unit dimensions and a riffled is aimed at it which is triggered several times and after some fixings it will appear as shown in figure 6.6. Whenever a bullet is fixed it is equally likely to strike any where on the squared target. Let us consider the left side of the target as Y - axis and the bottom as X - axis, obviously the variate X would defined horizontal distance of a hit from the vertical axis ($X = 0$) and as such it may have any value between 0 and 1 and will be called as continuous random variable.

From Geometric probability it can be seen that the chance of a hit into any internal is equal to the horizontal length of that interval divided by the total length of the board which will be equal to one. For instance the probability that the hit strikes between 0.3 and 0.8, horizontal distance $\frac{0.5}{1.0} = 0.5$. Since for the continuous distribution it is not possible to have a finite probability associated with single point as in the discrete distribution.

6.7 Transofrmation of One Dimenssional Random Variable:

Let X be a random variable defined on the event sace S and let $g(\ )$ be a function such that $Y = g(X)$ is also a.r.v. defined on S. This can be shown by following theorem.

Theorem : Let X be a continuous r.v. with p.d.f. $f_X(x)$. Let $y = g(x)$ be strictly monotonic increasing (or decreasing) function of x. Assume that $g(x)$ is differentiable and hence continuous for all x. Then the p.d.f. of the r.v. y is given by

$$h_y(y) = f_X(x)\left|\frac{dx}{dy}\right|$$

where x is expressed in terms of y.

Proof :

Case (i): $y = g(x)$ is strictly increasing function of x

i.e., $\frac{dy}{dx} > 0$. The d.f. of y is given by

$$H_Y(y) = P(Y \le y) = P[g(X) \le y] = P\left[X \le g^{-1}(y)\right]$$

the inverse exists and is unique, since $g(\cdot)$ is strictly increasing.

$$\therefore H_Y(y) = F_X\left[g^{-1}(y)\right],$$ where F is the d.f. of X.

$$= F_X(x) \quad \left[\because y = g(x) \Rightarrow g^{-1}(y) = x\right]$$

Differentiating w.r.t. 'y', we get

$$h_Y(y) = \frac{d}{dy}\left[F_X(x)\right] = \frac{d}{dx}\left[F_X(x)\right]\frac{dx}{dy}$$

$$= f_X(x)\frac{dx}{dy} \quad \cdot \quad \cdot \quad \cdot (1)$$

Case (ii): $y = g(x)$ is strictly monotonic decreasing

$$H_Y(y) = P(Y \le y) = P[g(X) \le y]$$

$$= P\left[X \ge g^{-1}(y)\right]$$

$$= 1 - P\left[X \le g^{-1}(y)\right]$$

$$= 1 - F_X\left[g^{-1}(y)\right]$$

$$= 1 - F_X(x)$$

where $x = g^{-1}(y)$, the inverse exists and is unique. Differentiating w.r.t. 'y', we get

$$h_Y(y) = \frac{d}{dx}\left[1 - F_x(x)\right]\frac{dx}{dy} = -f_X(x)\cdot\frac{dx}{dy} = f_X(x) - \frac{dx}{dy} \quad \cdot \quad \cdot \quad \cdot (2)$$

the algebraic sign (-ve) obtained in (2) is since y is a decreasing function of $x \Rightarrow x$ is a decreasing function of $y \Rightarrow \frac{dx}{dy} < 0$.

By combining equation (1) & (2) gives

$$h_Y(y) = f_X(x)\left|\frac{dx}{dy}\right|$$

6.8 Workedout Examples:

Example 1 : A random variable X has the following probability distribution.

x :	0	1	2	3	4	5	6	7	8
$P(x)$:	K	3K	5K	7K	9K	11K	13K	15K	17K

a. determine the value of K.

b. find the distribution function.

c. find the smallest value of x for which $P(X \le x) > 0.5$

Solution :

(a) $\Sigma P(x_i) = 1 \Rightarrow K(1+3+5+7+9+11+13+15+17) = 1$

$$\Rightarrow 81\,K = 1 \Rightarrow K = \frac{1}{81}$$

(b) $P(X<4)=P\{(X=0)\cup(X=1)\cup(X=2)\cup(X=3)\}=p_0+p_1+p_2+p_3$

$$=K+3K+5K+7K=16K=\frac{16}{81}$$

$$P(X\geq 5)=P\{(X=5)\cup(X=6)\cup(X=7)\cup(X=8)\}$$

$$=p_5+p_6+p_7+p_8=11K+13K+15K+17K=56K=\frac{56}{81}$$

$$P(0<X<4)=P\{(X=1)\cup(X=2)\cup(X=3)\}=p_1+p_2+p_3=3K+5K+7K=15K=\frac{15}{81}$$

(c) the distribution is $F(t)=P(x\leq t)$

$F(0)=P(X\leq 0)=K$

$F(1)=P(X\leq 1)=4K$

$F(2)=P(X\leq 2)=9K$

$F(3)=P(X\leq 3)=16K$

$F(4)=P(X\leq 4)=25K$

$F(5)=P(X\leq 5)=36K$

$F(6)=P(X\leq 6)=49K$

$F(7)=P(X\leq 7)=64K$

$F(8)=P(X\leq 8)=81K$

We observe that $p(X<4)=p(X\leq 3)=F(3)=16K,$ etc...

$$P(X\geq 5)=1-P(X<5)=1-P(X\leq 4)=1-25K=\frac{86}{81}$$

$$P(0<X<4)=F(4)-F(0)-P_4=(25-1-9)K=15K=\frac{15}{81}$$

(d) $F(x) > \frac{1}{2}; F(5) = \frac{36}{81} = 0.44; F(6) = \frac{49}{81} = 0.61$

then the smallest value of x for which $F(x) > \frac{1}{2}$ is $x = 6$.

Example 2 : Consider the distribution function

$$F(t) = 0, \qquad t < 0$$
$$= t, \qquad 0 \le t < 1$$
$$= 1, \qquad t \ge 1$$

Find the density function and compute

(a) $P\left[\frac{1}{4} < t < \frac{3}{4}\right]$ (b) $\left[-1 < t < \frac{1}{2}\right]$

Solution : The derivative of F(t) at t = x is given by

$$f(x) = \left[\frac{d}{dt}F(t)\right] \quad \text{at} \quad t = x$$

f(x) = 0, $x < 0$

= 1, $0 \le x < 1$

= 0, $x \ge 1$

Hence we define the density function

f(x) = 1 $0 < x < 1$

= 0 otherwise

As the derivative $\frac{d}{dt}F(t)$ does not exist at t = 0 and t = 1, so we take the range $0 \le t < 1$ as $0 < t < 1$

(a) $$0 < P\left[\frac{1}{4} < x < \frac{3}{4}\right] = \int_{1/4}^{3/4} 1 \cdot dx = [x]_{1/4}^{3/4} = \frac{3}{4} - \frac{1}{4} = 1/2$$

(b) $P\left[-1 < x < \frac{1}{2}\right] = \int_{-1}^{1/2} f(x)dx = \int_{-1}^{0} f(x)dx + \int_{0}^{1/2} f(x)dx = \int_{-1}^{0} 0 \cdot dx + \int_{0}^{1/2} 1 \cdot dx$

$$= [x]_0^{1/2} = \frac{1}{2} - 0 = 1/2$$

Example 3 : Verify that F(t) = 0 $\quad t < 0$

$= t^2 \quad 0 \le t < 1/2$

$= 1 - 3(1 - t^2) \quad 1/2 \le t < 1$

$= 1 \quad t \ge 1$

is a distribution function and derive the density funtion of X.

Solution : 1. It satisfies that $0 \le F(t) \le 1 \ \forall t$

2. It satisfies that $F(-\infty) = 0$ and $F(\infty) = 1$

3. F(t) is a non-decreasing funtion and it is right continuous for all t

Hence it is a distribution function

the derivative of F(t) at t = x

$$f(x) = \left[\frac{d}{dt} F(t)\right] \text{ at } t = x$$

f(x) = 0, $\quad x < 0$

= 2x, $\quad 0 \le x < 1/2$

= 6x, $\quad 1/2 \le x < 1$

= 0 $\quad x \ge 1$

the density function is f(x) = 2x, $\quad 0 \le x < 1/2$

= 6x, $\quad 1/2 \le x < 1$

= 0, otherwise

Example 4: Check whether the function given by

$f(x)=\frac{x+2}{25}$ for x = 1,2,3,4,5 is a p.d.t of desuete random variable.

Solution : The given function is $f(x)=\frac{x+2}{25} \cdot \quad \cdot \quad \cdot (1)$

Now substituting the different values of x we get from (1) x = 1,2,3,4,5

$f(1)=3/25$, $f(4)=6/25$

$f(2)=4/25$, and $f(5)=7/25$

$f(3)=5/25$

Since these values are all non negative.

f(1) + f(2) + f(3) + f(4) + f(5) = $\frac{3}{25}+\frac{4}{25}+\frac{5}{25}+\frac{6}{25}+\frac{7}{25}=1$

$\because f(x)\geq 0$ and $\sum f(n)=1$ conditions are satisfied thus, the given function is $\frac{x}{a}$ p.d.f. of a random variable hairing the range {1,2,3,4,5}.

Example 5 : Find the distribution function of the total number of heads obtained in four tosses of a balanced coin.

Solution : Given $f(0)=1/16$, $f(1)=4/16$, $f(2)=6/16$, $f(3)=4/16$, $f(4)=1/16$

Which follows

$$F(0)=f(0)=1/16$$

$$F(1)=f(0)+f(1)=1/16+\frac{4}{16}=5/16$$

$$F(2)=f(0)+f(1)+f(2)=\frac{1}{16}+\frac{4}{16}+\frac{6}{16}=\frac{11}{16}$$

$$F(3)=f(0)+f(1)+f(2)+f(3)=\frac{1}{16}+\frac{4}{16}+\frac{6}{16}+\frac{4}{16}=\frac{15}{16}$$

$$F(4)=f(0)+f(1)+f(2)+f(3)+f(4)=\frac{1}{16}+\frac{4}{16}+\frac{6}{16}+\frac{4}{16}+\frac{1}{16}=1$$

Hence the distribution function is given by

$$F(x)=\begin{cases} 0 & \text{for} \quad x<0 \\ \frac{1}{16} & \text{for} \quad 0\le x<1 \\ \frac{5}{16} & \text{for} \quad 1\le x<2 \\ \frac{11}{16} & \text{for} \quad 2\le x<3 \\ \frac{15}{16} & \text{for} \quad 3\le x<4 \\ 1 & \text{for} \quad x\ge 4 \end{cases}$$

This distribution function is defined not only for the values taken on by the given random variable, but for all real numbers is observed. For example we can write $F(1\cdot 7)=\frac{5}{16}$ and $F(100)=1$ although the probabilities of getting "atmost 1.7 heads" or "atmost 100 heads" in four tosser of a balanced coin may not be of any real significance.

Example 6 : If X has probability density

$$f(x)=\begin{cases} k\cdot e^{-3x}, & \text{for } x>0 \\ 0 & \text{otherwise} \end{cases}$$

Since the given function is continuous we have

$$\int_{-\infty}^{\infty} f(x)dx=\int_{0}^{\infty} k.e^{-3x}dx=k.\int_{0}^{\infty} e^{-3x}dx=k\cdot\frac{e^{-3x}\Big|_0^{\infty}}{3}$$

$$=\frac{-k}{3}\left(e^{-\infty}-e^{-0}\right)=\frac{k}{3}=1\Rightarrow k=3$$

and

$$P(0.5\le x\le 1)=\int_{0.5}^{1} 3e^{-3x}dx=-e^{-3x}\Big|_{0.5}^{1}=-e^{-3}+e^{-1.5}=0.173$$

For x > 0,

$$F(x)=\int_{-\infty}^{x} f(t)dt=\int_{0}^{x} 3e^{-3t}dt=-e^{-3t}\Big|_0^x=1-e^{-3x}$$

and since F(x) = 0 for $x \le 0$ we can write

$$F(x) = \begin{cases} 0 & \text{for} \quad x \le 0 \\ 1 - e^{-3x} & \text{for} \quad x > 0 \end{cases}$$

$$P(0.5 \le x \le 1) = F(1) - F(0.5)$$

$$= \left(1 - e^{-3}\right) - \left(1 - e^{-1.5}\right)$$

= 0.173

Example 7 : Let f(x) = 2x, 0 < x < 1 and 0, otherwise, be the p.d.f. of x. Find the distribution function of p.d.f. of $y = \sqrt{x}$.

Solution : $P(Y \le y) = P\left(\sqrt{x} \le y\right) = P\left(x \le y^2\right)$

$$\therefore F_x(y) = \int_0^{y^2} f(x)dx$$

$$= \int_0^{y^2} 2xdx = 2 \cdot \int_0^{y^2} xdx = \left.\frac{2 \cdot x^2}{2}\right|_0^{y^2} = y^4, 0 \le y \le 1$$

This gives the distribution function of y. the p.d.f. of y is

$$f_y(y) = F_y^1(Y) = 4y^3, \; 0 \le y \le 1$$

Example 8 : A variate x has p.d.f. $f(x) = 1/x^2, \; x \ge 1;$

f(x) = 0, x < 1 find the p.d.f. of e^{-x}.

Solution : Let $y = e^{-x}$ and use the transformation $y = e^{-x}$

$$\Rightarrow x = -\log y$$

this gives $\left|\frac{dx}{dy}\right| = \frac{1}{y}; \; x \ge 1 \Rightarrow y \le \left(1/e\right)$

Further y > 0

$f_x(x) = \left|\frac{dy}{dx}\right| f_y(y)$ we get

$f_y(y) = [y \ \log\ y]^{-2} \quad 0 < y \le e^{-1},\ f_y(y) = 0$ otherwise

Example 9 : Variate X has the p.d.f. $f(x) = \frac{1}{\pi(1+x^2)},\ -\infty < x < \infty$

Find the p.d.f. of $y = \tan^{-1} x$

Solution : $Y = \tan^{-1} x \Rightarrow x \ \tan\ y$ so that $\frac{dx}{dy} = \sec^2 y > 0$

Thus the function is encreasing. Also

$1 + x^2 = 1 + \tan^2 y = \sec^2 y$

hence $f_y(Y) = \left[f_x(x)\left|\frac{dx}{dy}\right|\right]_y = \frac{\sec^2 y}{\pi \sec^2 y} = \frac{1}{\pi},\ \frac{-\pi}{2} < y < \frac{\pi}{2}$

Example 10 : If $Y = |X|$. Show that

$$F_y(Y) = \begin{cases} F_x(y) \quad -F_x(-y) + P(x = -y), y > 0 \\ 0 \qquad\qquad y \le 0 \end{cases}$$

Solution : $F_y(Y) = P(Y \le y)$

$= P(1 \times 1 \le y)$

$= P(-y \le x \le y)$

$= P\{(-y < x \le y) \cup (x = -y)\}$

$= P\{-y < x \le y\} + P(x = -y)$

$= F_x(y) - F_x(-y) + P(x = -y)$

($\because$ By using distribution properties)

6.9 Exercise:

1. An experiment consists ofthree independent to uses of a fair win. Let x = the number of heads, y = the number of head runs, z = the length of head runs. A head run being defined as consewtive occurrence together in three tosses of the win. Find the probability function of (i) X, (ii) Y, (iii) Z, (iv) X + Y and (v) XY and construct probability tables and draw their probability charts.

2. A random variable X has the following probability function

x	0	1	2	3	4	5	6	7
p(x)	0	k	2k	2k	3k	k^2	$2k^2$	$7k^2 + k$

(i) Find k,

(ii) Evaluate P(x < 6), $P(x \geq 6)$ and P(0 < x < 5)

(iii) If $P(x \leq k) > \frac{1}{2}$, find the minimum value of k and

(iv) Determine the distribution function of X.

(Ans: (i) $k = 1/10$, (ii) $\frac{81}{100}, \frac{19}{100}, 4/5$, (iii) k = 4)

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

(i) check that above is p.d.f.

(ii) Determine a number of such that $P(x < b) = P(x > b)$

(Ans: (i) p.d.f., (ii) b=1/2)

4. A continuous random variable X has a p.d.f. $f(x) = 3x^2$, $0 \leq x \leq 1$

Find a and b such that (i) $P(x \leq a) = P(x > a)$, (ii) $P(x > b) = 0.05$

(Ans: (i) $a = \left(\frac{1}{2}\right)^{1/3}$, (ii) $b = \left(\frac{19}{20}\right)^{1/3}$)

5. Let x be a continuous random variate with p.d.f.

f(x) = ax, $0 \leq x \leq 1$

= a, $1 \leq x \leq 2$

= -ax + 3a, $2 \le x \le 3$

= 0 otherwise

(i) Determine the constant a

(ii) Compute $P(x \le 1.5)$

(Ans: (i) a = 1/2, (ii) a = 1/2)

6. The milege C in thousands of miles which car owners get with a certain kind of type is a random variable having probability density function

$$f(x) = \frac{1}{20} e^{-x/20}, \text{ for } x > 0$$

$= 0$ for $x \le 0$.

Find the probability that one of these tyres will last

(i) at most 10,000 miles

(ii) any where form 16,000 to 24,000 miles

(iii) at least 30,000 miles.

(Ans: (i) 0.3935, (ii) 0.1481, (iii) 0.2231)

7. Verify that the following function is a distrilention function

$$F(x) = \begin{cases} 0 & x < -a \\ \frac{1}{2}\left(\frac{x}{a} + 1\right) & -a \le x \le a \\ 1 & x > a \end{cases}$$

8. A petrol pump is sopplied with petrol once a day. If its daily volume X of sales in thousands of litres is distrilented by

$$f(x) = 5(1-x)^4 \ ; 0 \le x \le 1$$

What must be the capacity of its tank in order that the probability that its supply will be exhansted in a given day shall be 0.01 ?

(Ans : a = 0.6019, 601.9 litres)

9. If the comulative distribution function of X if F(x), Find the c.d.f. of

(i) Y = x+a, (ii) Y = x - b, (iii) y = ax, (iv) $y = x^3$, (v) $y = x^2$

What are corresponding probability density functions ?

10. Let $f(x)=\begin{cases} \frac{1}{2} & -1<x<1 \\ 0 & \text{otherwise} \end{cases}$

be the p.d.f. of the r.u.x. Find distribution function and the p.d.f. of $Y=x^2$?

6.10 Summary :

The conceproja random variable, its associated distribution function are defined and a number of example, are given prigunction of random variables into discrate and continution types - the associated mass functions and density functions are presulted.

6.11 Technical Terms :

Random, Variables, Distribution function, Probabilites mass function, Probability density function.

Lesson – 7

MATHEMATICAL EXPECTATION

Objective of The Lesson:

After studying this lesson the student is expected to have a clear comprehension of the theory and the practical utility about the concepts of mathematical expectation, moments, M.G.F., C.G.F., P.G.F. definitions, properties and their applications.

Structure of The Lesson:

7.1 Introduction:

When real variable is associated with a probability distribution, no value of the variable is a certainty and accordingly one is not since of what value is assumed us the real variable. In situations like temperature in atmosphere, rainfall in monsoon, profits / losses in business, sensex in share market etc. We can talk of only average value rater than exact value. If the probability distribution of the underlying variable is known / specified the average value can be calculated by a concept called Mathematical Expectation. This lesson is devoted to introduce this concept in a theoretical way and present its practical utility along with its other related concepts in the following sections.

7.2 Mathematical Expectation of Random Variable:

7.2.1 Definition: Let X be a random variable defined on a probability space. Suppose if X be a discrete and Let $\{x_i\}$ be the countable set of its possible values, such that $p\left(x = x_i\right) = p_i$

the mathematical expectation of X, (or the mean of X) written as $E(X)$, is a real number defined by

$$E(X) = \sum p_i x_i, \ \forall i \qquad \cdots\cdots\cdots\cdots\cdots(1)$$

Provided the series $\sum p_i |x_i|$ is convergent.

$$\int_{-\infty}^{\infty} f(x)|x| \, dx < \infty$$

i.e., convergent, the expectation of X, written as $E(x)$, is the real number defined by

$$E(X) = \int_{-\infty}^{\infty} x f(x) \, dx \qquad \cdots\cdots\cdots\cdots\cdots\cdots(2)$$

If the series (1), or integral (2) is conditionally convergent, $E(X)$ does not exist. Therefore the series (1) or integral (2) must be absolute convergent and hence $E(X)$ exist. Thus $E(X)$ exists iff $E(|X|)$ exists.

Illustration 1:

For example if the random variable X takes the values $0!1!\ 2!\cdots\cdots\cdots\cdots$ with probability law

$$P(X = x!) = \frac{e^{-1}}{x!}, \quad x = 0,1,2\cdots\cdots\cdots$$

then $\sum_{x=0}^{\infty} x!\, P(X = x!) = e^{-1} \sum_{x=0}^{\infty} 1$

Which is a divergnt series. In this case $E(X)$ does not exist.

For example if the random variable X which takes the values

$$x_i = (-1)^{i+1}(i+1) \ ; \ i = 1,2,3,\cdots\cdots\cdots\cdots$$

with the probability law $p_i = P(X = x_i) = \frac{1}{i(i+1)}$; $i = 1,2,3,\cdots\cdots\cdots$

Here $\sum_{i=1}^{n} x_i \; p(X = x_i) = \sum_{i=1}^{n} (-1)^{i+1} \left(\frac{1}{i}\right) = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots\cdots$

The series on R.M.S. is conditionally convergent since the terms alternate in sign, are monotonically decreasing and converge to zero when we use leibnitz test for alternating series. By conditional convergence although $\sum_{i=1}^{\infty} p_i x_i$ converges, $\sum_{i=1}^{\infty} |p_i \, x_i|$ does not converge. So in the above example $E(X)$ does not exist even though $\sum_{i=1}^{\infty} p_i x_i$ is finite i.e., $\log_e^2$.

Illustration 2:

Let us consider the r.u.X. which takes te values

$$x_t = \frac{(-1)^t \cdot 2^t}{t} \; ; \; t = 1, 2, 3, \cdots\cdots\cdots$$

with probabilities $p_t = 2^{-t}$

Here also we get $\sum_{t=1}^{\infty} x_t p_t = \sum_{t=1}^{\infty} \frac{(-1)^t}{t}$

$$= -\left[1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots\cdots\right] = -\log_e^2$$

and $\sum_{t=1}^{\infty} |x_t| \, p_t = \sum_{t=1}^{\infty} \frac{1}{t}$

Which is a divergent series. Hence in this case also expectation does not exist.

Illustration 3:

Let us consider a continuous r.v. 'X' and p.d.f. is

$$f(x) = \frac{1}{\pi} \cdot \frac{1}{\left(1 + x^2\right)} \; ; -\infty < x < \infty$$

which isthe p.d.f. of standard cauchy distribution

$$\int_{-\infty}^{\infty} |x|\, f(x)\, dx = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{|x|}{1+x^2} dx = \frac{2}{\pi} \int_{0}^{\infty} \frac{x}{1+x^2}\, dx$$

($\because$ Integrand is an even function of X)

$$= \frac{1}{\pi} \left| \log\left(1+x^2\right) \right|_0^{\infty} \to \infty$$

Since this integral does not converge to a finite limit. Hene $E(X)$ does not exist.

7.2.2 Properties:

Expectation of a Function of a Random Variable:

Consider a r.v. X with p.d.f. (p.m.f.) $f(x)$ and distribution $F(x)$. If $g(x)$ is a function such that $g(x)$ is a r.v. and $E(g(x))$ exists, then

$$E[g(x)] = \int_{-\infty}^{\infty} g(x) \cdot dF(x) = \int_{-\infty}^{\infty} g(x) \cdot f(x) dx \cdots\cdots\cdots\cdots (1)$$

(for continuous r.v.)

$$= \sum_{x} g(x) \cdot f(x) \cdots\cdots\cdots\cdots\cdots (2)$$

(for discrete r.v.)

Case 1: If we take $g(x) = x^2$, r being a positive integer, in (1) we get

$$E(X^r) = \int_{-\infty}^{\infty} x^2 \cdot f(x)\, dx \cdots\cdots\cdots\cdots (3)$$

which isdefined as μ_r^1, the r^{th} moment (about origin) of the probability distrilention.

thus $\mu_r^1 \left(\text{about origin}\right) = E(X^r)$

$\mu_1^1 \left(\text{about origin}\right) = E(X)$

$$\mu_2^1 \text{ (about origin)} = E(X^2)$$

Hence Mean $= \bar{x} = \mu_1^1 \text{ (about origin)} = E(X) \cdots\cdots\cdots (4)$

and $\mu_2 = \mu_2^1 - \mu_1^{1^2} = E(X^2) - [E(X)]^2 \cdots\cdots\cdots\cdots (5)$

Case 2: If $g(x) = [X - E(X)]^2 = (X - \bar{X})^2$

then from equation (1) we get

$$E(X - E(X))^r = \int_{-\infty}^{\infty} [x - E(X)]^2 - f(x)\,dx = \int_{-\infty}^{\infty} (x - \bar{x})^2 f(x)\,dx \cdots\cdots\cdots (6)$$

which is μ_r, the r^{th} moment about mean.

In particular if $r = 2$, we get

$$\mu_2 = E[X - E(X)]^2 = \int_{-\infty}^{\infty} (x - \bar{x})^2 \cdot f(x)dx \cdots\cdots\cdots\cdots (7)$$

Equations (5) and (7) gives the variance of the probability distribution of a r.v. X in terms of expectation.

Case 3: Taking $g(x) = \text{cons tan t} = C(\text{say})$ in (1) we get

$$E(C) = \int_{-\infty}^{\infty} C \cdot f(x)\,dx = C \int_{-\infty}^{\infty} f(x)\,dx = C \cdots\cdots\cdots\cdots (8)$$

$$\therefore\ E(C) = C \cdots\cdots\cdots\cdots \geq (9)$$

Addition theorem of Expectation:

Statement: If X and Y are random variables then

$$E(X + Y) = E(X) + E(Y) \qquad \cdots\cdots\cdots\cdots\cdots\cdots (1)$$

Provided all the exectations exists.

Proof: Let X and Y be continuous r.v.'s with joint p.d.f. $f_{x,y}(x, y)$ and marginal p.d.f's $f_x(x)$ and $f_y(y)$ respectively. Then by definition

$$E(X) = \int_{-\infty}^{\infty} x \cdot f_x(x) dx \qquad \cdots\cdots\cdots\cdots\cdots(2)$$

$$E(Y) = \int_{-\infty}^{\infty} y \cdot f_y(y)\, dy \qquad \cdots\cdots\cdots\cdots\cdots(3)$$

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

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x \cdot f_{x,y}(x, y)\, dx\; dy + \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y \cdot f_{xy}(x, y)\, dx\, dy$$

$$= \int_{-\infty}^{\infty} x \cdot \left[\int_{-\infty}^{\infty} f_{xy}(x, y)\, dy \right] dx + \int_{-\infty}^{\infty} y \left[\int_{-\infty}^{\infty} f_{xy}(x, y)\, dx \right] dy$$

$$= \int_{-\infty}^{\infty} x\, f_x(x)\, dx + \int_{-\infty}^{\infty} y\, f_y(y)\, dy$$

$$= E(X) + E(Y) \qquad \text{[From equations (2) \& (3)]}$$

The above result can be extended to n variables also by mathematical induction.

Multiplication theorem of Expectation:

Statement: If X and Y are independent random variables, then $E(X\,Y) = E(X) \cdot E(Y)$.

Proof: By definition of Mathematical expectation if X, Y are continuous r.v.'s then

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

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy \cdot f_x(x) \cdot f_y(y)\, dy\, dx \qquad (\because \text{ x and y are independent})$$

$$= \int_{-\infty}^{\infty} x \cdot f_x(x)\, dx \int_{-\infty}^{\infty} y \cdot f_y(y)\, dy$$

$$= E(X) \cdot E(Y) \qquad \left[\because \int_{-\infty}^{\infty} x \cdot f_x(x)\, dx = E(X), \int_{-\infty}^{\infty} y\, f_y(y)\, dy = E(Y)\right]$$

The above result can be extended to n variables also by mathematical induction.

Theorem:

If X is a random variable and 'a' is constant then

(i) $E(a\, \psi(X)) = a\, E[\psi(x)]$ (ii) $E(\psi(X)+a) = E[\psi(X)] + a$

where $\phi(x)$, a function of X, is a r.v. and all the expectations exist.

Proof:

(i) $E[a\psi(X)] = \int_{-\infty}^{\infty} a \cdot \psi(X) \cdot f(x)\, dx = a \int_{-\infty}^{\infty} \psi(X)\, f(x)\, dx = a \cdot E(\psi(X))$

(ii) $E[\psi(x) + a] = \int_{-\infty}^{\infty} [\psi(X)+a]\, f(x)\, dx$

$$= \int_{-\infty}^{\infty} \psi(X)\, f(x)\, dx + a \cdot \int_{-\infty}^{\infty} f(x)\, dx$$

$$= E[\psi(X)] + a \qquad \left[\because \int_{-\infty}^{\infty} f(x)dx = 1\right]$$

Corollary: (i) If $\psi(X) = x$ then $E[aX] = A\ E[X]$ and $E[X+a] = E(X) + a$

(ii) If $\psi(X) = 1$, then $E[a] = a$

Theorem: If X is a random variable and a and b are constants, then $E(ax+b) = aE(X)+b$. Provided all the expectations exists.

Proof: By definition of mathematical expectation we have

$$E(aX+b) = \int_{-\infty}^{\infty} (ax+b)\ f(x)\ dx$$

$$= a \cdot \int_{-\infty}^{\infty} x \cdot f(x)\ dx + b \cdot \int_{-\infty}^{\infty} f(x)\ dx$$

$$= a \cdot E(X) + b$$

Corollary: 1. If $b = 0$, then we get $E(aX) = a \cdot E(X)$

2. Taking $a = 1,\ b = -\overline{X} = -E(X)$, we get $E(X - \overline{X}) = 0$

Theorem: If $X \geq 0$ then $E(X) \geq 0$.

Proof: If X is a continuous r.v. s.t. $x \geq 0$ then

$$E(X) = \int_{-\infty}^{\infty} x\ f(x)\ dx = \int_{0}^{\infty} x \cdot f(x)\ dx > 0 \qquad [\because \text{ if } X \geq 0, P(x) = 0 \text{ for } x < 0]$$

$$\therefore\ E(X) \geq 0$$

Provided the expectation exists.

Theorem: Let X and Y betwo random variables such that $Y \leq X$ then

$$E(Y) \leq E(X)$$

Provided the expectation exists.

Proof: Since $Y \leq X$, we have the r.v.

$$Y - X \leq 0 \Rightarrow X - Y \geq 0$$

Hence $E(X-Y)\geq 0 \Rightarrow E(X)-E(Y)\geq 0$

$\Rightarrow E(X)\geq E(Y) \Rightarrow E(Y)\leq E(X)$

Theorem: $|E(X)|\leq E|(X)|$ provided the expectation exist.

Proof: Since $X\leq |X|$ we have by theorem $\left(\text{i.e. } E(Y)\leq E(X)\right)$

$E(X)\leq E|X| \quad \cdots\cdots\cdots\cdots\cdots\cdots(1)$

Again since $-X\leq |X|$ again by theorem $\left(\text{i.e., } E(Y)\leq E(X)\right)$

$E(-X)\leq E|X|$

$\Rightarrow E(X)\leq E|X| \quad \cdots\cdots\cdots\cdots\cdots\cdots(2)$

From (1) & (2) we get $|E(X)|\leq E|X|$

Theorem: If X is a random variable, then $V(aX+b)=a^2\cdot V(X)$ where a and b are constants.

Proof: Let $Y=aX+b\cdots\cdots\cdots\cdots\cdots\cdots(1)$

Taking expectation on both sides of equation (1) we get

$E(Y)=E(aX+b)$

$\Rightarrow E(Y)=aE(X)+b$

$\therefore\ Y-E(Y)=a\left\{X-E(X)\right\}$

Squaring and taking expectation of both sides, we get

$E\left\{Y-E(Y)\right\}^2=a^2\ E\left\{X-E(X)\right\}^2$

$\Rightarrow V(Y)=a^2\cdot V(X)\Rightarrow V(aX+b)=a^2V(X)\cdots\cdots\cdots(2) \qquad (\because \text{from equation (1)})$

If $b = 0$, then $V(aX) = a^2 \cdot V(X)$

$\Rightarrow$ Variance is not independent of change of scale.

If $a = 0$, then $V(b) = 0$

$\Rightarrow$ Variance of a constant is zero

If $a = 1$ then $V(X + b) = V(X) \Rightarrow$ variance is independent of change of origins.

7.3 Moments and Central Moments:

As explained in section 7.2 we can set the expectation any function $g(x)$ of a random variable X with the help of the probability model of that random variable. Specifically if we take $g(x) = (X - a)^k$ where 'a' is any constant and k is a natural number i.e.

$$E(g(x)) = E(X - a)^k$$

then we call this k-th moment of X about 'a' and denote is by μ_k^1. In particular if $a = 0$ then it is called k-th raw moment. If 'a' is taken as $E(X)$ itself then it is called k-th central moment and is denoted by μ_k.

In all we introduce the following in terms of mathematical expectation.

k-th moment bout an arbitrary constant:

$$\mu_k^1 = E(X - a)^k$$

k-th raw moment $m_k = E(X^k)$

k-th central moment $\mu_k = E(X - E(X))^k$

It can be seen that first raw moment is mean of the random variable, first central moment is zero always, second central moment is variance of the random variable.

7.4 Moment Generating Function:

The moment generating funtion (m.g.f.) of a random variable X (about origin) haing the probability density function $f(x)$ is given by

$$M_X(t) = E\left(e^{tX}\right) = \int e^{tX} \cdot f(x)\, dx.$$ (for continuous probability distribution)

$$= \sum_x e^{tx} f(x)$$ (for descrete probability distribution) $\cdot \quad \cdot \quad \cdot (1)$

Here the integration or summation being extended to the entire range of x, t being a real number and it is being assumed that the R.H.S. of (1) is absolvetely convergent for some positive number h such that $-h < t < h$. Thus

$$M_X(t) = E\left(e^{tX}\right) = E\left[1 + tX + \frac{t^2X^2}{2} + \cdots\cdots + \frac{t^rX^2}{r!} + \cdots\cdots\right]$$

$$= 1 + t\,E(X) + \frac{t^2}{2!} E\left(X^2\right) + \cdots\cdots + \frac{t^2}{r\,!} E\left(X^2\right) + \cdots\cdots$$

$$= 1 + t \cdot \mu_1^1 + \frac{t^2}{2!} \mu_2^1 + \cdots\cdots\cdots + \frac{t^2}{r\,!} \mu_r^1 + \cdots\cdots\cdots\cdots (2)$$

where $\mu_r^1 = E\left(X^r\right) = \int x^2 \cdot f(x)\, dx$, for continuous distribution.

$= \sum_x x^2 \cdot p(x)$, for discrete distribution.

is the r^{th} moment of X about origin. Thus the coefficient of $\frac{t^2}{r\,!}$ in $M_X(t)$ gives μ_r^1. Since $M_X(t)$ generates moments it is known as moment generating function.

Differentiating w.r.t. 't' and then putting $t = 0$ we get

$$\left|\frac{d^r}{dt^r}\{M_X(t)\}\right|_{t=0} = \left|\frac{\mu_r^1}{r!} \cdot r! + \mu_{r+1}^1 \cdot t + \mu_{r+2}^1 \cdot \frac{t^2}{2!} + \cdots\cdots\right|_{t=0}$$

$$\Rightarrow \mu_r^1 = \left|\frac{d^r}{dt^r} \cdot \{M_X(t)\}\right|_{t=0} \quad \cdots\cdots\cdots\cdots (3)$$

In general, the moment generating function of X about the point $X = a$ is defined as

$$M_X(t) \quad (\text{about } X = a) = E\left[e^{t(X-a)}\right]$$

$$= E\left[1 + t(X-a) + \frac{t^2}{2!}(X-a)^2 + \cdots\cdots + \frac{t^2}{r!}(X-a)^2 + \cdots\cdots\right]$$

$$= \left[1 + t\, E(X-a) + \frac{t^2}{2\,!} E(X-a)^2 + \cdots\cdots + \frac{t^2}{r!}(X-a)^2 + \cdots\cdots\right]$$

$$= 1 + t\, \mu_1^1 + \frac{t^2}{2!}\mu_2^1 + \cdots\cdots + \frac{t^2}{r!}\mu_2^1 + \cdots\cdots\cdots(4)$$

Properties of M.G.F.:

1. $M_{CX}(t) = M_X(Ct),\ C$ beinga constant $\cdots\cdots\cdots\cdots(1)$

By definition of M.C.F. we have

i.e., $M_X(t) = E\left(e^{tX}\right)$

$\therefore$ From L.H.S. of (1) is $M_{CX}(t) = E\left(e^{tCX}\right)$

From R.H.S. of (1) is $M_X(Ct) = E\left(e^{CtX}\right) = L \cdot H > S.$

Hene $M_{CX}(t) = M_X(Ct)$

2. The moment generating funtion of the sum of a number of independent random variables is equal to the product of their respective moment gnerating functions.

Symbolically, if $X_1, X_2, \cdots\cdots, X_n$ are independent random variables, then the moment generating function of their sum $X_1 + X_2 + \cdots\cdots + X_n$ is given by

$$M_{X_1+X_2+\cdots\cdots+X_n}(t) = M_{X_1}(t) \cdot M_{X_2}(t) \cdots\cdots M_{X_n}(t) \cdots\cdots\cdots(1)$$

By definition of M.G.F. we have

$$M_{X_1+X_2+\cdots\cdots+X_n}(t) = E\left[e^{t(X_1+X_2+\cdots\cdots+X_n)}\right]$$

$$= E\left[e^{tX_1} \cdot e^{tX_2} \cdots\cdots e^{tX_n}\right]$$

$$= E\left(e^{tX_1}\right).\left(e^{tX_2}\right)\ldots\ldots\ldots\ldots\ldots\left(e^{tX_n}\right)$$

$$\left(\because X_1, X_2, \cdots\cdots, X_n \text{ are independent}\right)$$

$$= M_{X_1}(t) \cdot M_{X_2}(t) \cdots\cdots M_{X_n}(t)$$

3. Effect of change of origin and scale on M.G.F.

By definition of M.G.F. we have

$$M_X(t) = E\left(e^{tX}\right) \ldots\ldots\ldots\ldots\ldots\ldots (1)$$

Let us transform X to the new variable U by changing both theorigin and scale in X as

$U = \dfrac{X-a}{h}$, a, h are constants.

M.G.F. of U (about origin) is given by

$$M_U(t) = E\left(e^{tU}\right) = E\left[\exp\left\{\frac{t(x-a)}{h}\right\}\right]$$

$$= E\left[e^{tX/h} \cdot e^{-at/h}\right]$$

$$= e^{-at/h} \cdot E\left(e^{tX/h}\right)$$

$$= = e^{-at/h} \cdot M_X(t/h) \qquad \text{[from equation (1)]} \cdots\cdots\cdots\cdots (2)$$

where $M_X(t)$ is the m.g.f. of X about origin

Also if $a = E(X) = \mu$ (say) and $h = \sigma_X = \sigma$ (say), then

$$U = \frac{X - E(X)}{\sigma_X} = \frac{X-\mu}{\sigma} = Z \quad \text{(say)}$$

is known as a standard variate. Thus the m.g.f. of a standard variate Z is given by

$$M_Z(t) = e^{-\mu t/\sigma} \cdot M_X(t/\sigma) \cdots\cdots\cdots\cdots (3)$$

Remark: $E(Z) = E\left(\dfrac{X-\mu}{\sigma}\right) = \dfrac{1}{\sigma} E(X-\mu)$

$$= \frac{1}{\sigma}\left\{E(X) - \mu\right\} = \frac{1}{\sigma}(\mu - \mu) = 0$$

and $$V(Z) = V\left(\frac{X-\mu}{\sigma}\right) = \frac{1}{\sigma^2} V(X-\mu)$$

$$= \frac{1}{\sigma^2} V(X)$$

$$= \frac{1}{\sigma^2}\ \sigma^2$$

$$= 1$$

$E(Z) = 0$ and $V(Z) = 1$ i.e., the mean and vareaince of a standard variate are 0 and 1 respectively.

Limitations of Moment Generating Function:

Moment Generating Function is restricted its use in statistics. Here the deficiencies of m.g.f's with illustrations are explained.

1. A random variable X may have no moments although its m.g.f. exists for example let us consider a discrete random variable with probability function.

$$\left.\begin{aligned} f(x) &= \frac{1}{x(x+1)} \ ; \ x = 1, 2, \cdots\cdots\cdots \\ &= 0, \qquad \text{otherwise} \end{aligned}\right\}$$

Here $$E(X) = \sum_{x=1}^{\infty} x\, f(x) = \sum_{x=1}^{\infty} \frac{1}{(x+1)}$$

$$= \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots\cdots\cdots\cdots\cdots$$

$$= \sum_{x=1}^{\infty} \frac{1}{x}$$

Since $\sum\limits_{x=1}^{\infty} \frac{1}{x}$ is a divergent series $E(X)$ does not exist and consequently no moment of X exists. However, m.g.f. of X is given by

$$M_X(t) = \sum_{x=1}^{\infty} e^{tX} \cdot f(x) = \sum_{x=1}^{\infty} \frac{e^{tx}}{x(x+1)}$$

$$= \sum_{x=1}^{\infty} \frac{Z^x}{x(x+1)}, \quad \left(Z = e^t\right) \cdots\cdots\cdots\cdots(1)$$

$$= \frac{Z}{1\cdot 2} + \frac{Z^2}{2\cdot 3} + \frac{Z^3}{3\cdot 4} + \frac{Z^4}{4\cdot 5} + \cdots\cdots\cdots$$

$$= Z\left[1 - \frac{1}{2}\right] + Z^2\left[\frac{1}{2} - \frac{1}{3}\right] + Z^3\left[\frac{1}{3} - \frac{1}{4}\right] + Z^4\left[\frac{1}{4} - \frac{1}{5}\right] + \cdots\cdots\cdots$$

$$= \left[Z + \frac{Z^2}{2} + \frac{Z^3}{3} + \frac{Z^4}{4} + \cdots\cdots\right] - \frac{Z}{2} - \frac{Z^2}{3} - \frac{Z^3}{4} - \frac{Z^4}{5} \cdots\cdots\cdots$$

$$= -\log(1-Z) - \frac{1}{2}\left[\frac{Z^2}{2} + \frac{Z^3}{3} + \frac{Z^4}{4} + \cdots\cdots\right]$$

$$= -\log(1-Z) + 1 + \frac{1}{2}\log(1-Z), \ |Z| < 1$$

$$= 1 + \left(\frac{1}{Z} - 1\right)\log(1-Z), \ |Z| < 1$$

$$= 1 + \left(e^{-t} - 1\right)\log\left(1 - e^t\right), \ t < 0$$

$$[\text{from } (1)] \quad \left[\because |Z| < 1 \Rightarrow \left|e^t\right| < 1 \Rightarrow t < 0\right]$$

$M_X(t) = 1$, for $t = 0$

while for $t > 0$, $M_X(t)$ does not exist.

2. A random variable X can have m.g.f. and some (or all) moments, yet the m.g.f. does not generate the moments.

For example consider a discrete r.v. with probability function

$$P\left(X = 2^x\right) = \frac{e^{-1}}{x!} \; ; \; x = 0,1,2,\cdots\cdots\cdots$$

Here $E\left(X^r\right) = \sum_{x=0}^{\infty} \left(2^x\right)^r \cdot P\left(x = 2^x\right) = e^{-1} \sum_{x=0}^{\infty} \frac{\left(2^r\right)^x}{x!}$

$$= e^{-1} \cdot \exp\left(2^r\right) = \exp\left(2^r - 1\right)$$

Hence all the moments of X exist.

The m.g.f. of X, if it exusts is given by

$$M_X(t) = \sum_{x=0}^{\infty} \exp\left(t \cdot 2^x\right)\left(\frac{e^{-1}}{x!}\right)$$

$$= e^{-1} \cdot \sum_{x=0}^{\infty} \exp\left(t \cdot 2^x\right) \cdot \frac{1}{x!}$$

By D'Alemborts ratio test, the series on the R.H.S. converges for $t \le 0$ and diverges for $t > 0$. Hence $M_X(t)$ cannot be differentiated at $t = 0$ and has no Mactaurin's expansion and lonsequently it does not generate moments.

3. A r.v. X can have all or some moments, but m.g.f. does not exist except perhaps at one point.

For example Let X be a r.v. with probability function.

$$P\left(X = \pm\, 2^x\right) = \frac{e^{-1}}{2x!} \; ; \; x = 0,1,2,\cdots\cdots\cdots$$

$= 0$, otherwise.

Since the distribution is symmetric about the line $X = 0$, all moments of odd order about origin vanish.

i.e., $E\left(X^{2r+1}\right) = 0 \Rightarrow \mu_{2r+1} = 0$

$$E\left(X^{2r}\right) = \sum_{x=0}^{\infty} \left(\pm 2^x\right)^{2r} \cdot \frac{1}{2e \cdot x!} = \frac{1}{e} \sum_{x=0}^{\infty} \frac{\left(2^{2r}\right)^x}{x!}$$

$$= \frac{1}{e} \cdot \exp\left(2^{2r}\right) = \exp\left(2^{2r} - 1\right)$$

Thus all the moments of X exist. The m.g.f. of X, if it exists is given by

$$M_X(t) = \sum_{x=0}^{\infty} \left\{\left[e^{t \cdot 2^x} + e^{-t \cdot 2^x}\right] \frac{1}{ex!}\right\}$$

$$= e^{-1} \cdot \sum_{x=0}^{\infty} \left[\frac{\cos\left(t \cdot 2^x\right)}{x!}\right]$$

wich converges only for $t = 0$.

In case of continuous probability distribution consider Poneto distribution with p.d.f.

$$P(x) = \frac{\theta \cdot a^\theta}{x^{\theta+1}} \; ; \; x \geq a \; ; \; \theta > 1$$

$$E\left(X^r\right) = \theta \cdot a^\theta \int_a^\infty x^{r-\theta-1}\, dx = \theta \cdot a^\theta \cdot \left|\frac{x^{r-\theta}}{r-\theta}\right|_a^\infty$$

which is finite iff $r - \theta < 0 \Rightarrow \theta > r$ and then

$$E\left(X^r\right) = \theta \cdot a^\theta \left[0 - \frac{a^2 - \theta}{r - \theta}\right] = \frac{\theta \cdot a^2}{\theta - r} \; ; \; \theta > r$$

However, the m.g.f. is given by

$$M_X(t) = \theta \cdot a^\theta \cdot \int_a^\infty \frac{e^{tx}}{x^{\theta+1}}\, dx$$

The integral is not convergent since e^{tx} dominates $x^{\theta+1}$ and $\left(e^{tx}/x^{\theta+1}\right) \to \infty$ as $x \to \infty$, hece $M_X(t)$ does not exist.

7.5 Cumulant Generating Function:

Cumulants generating function $K(t)$ is defined as

$$K_X(t) = \log_e M_X(t) \quad \cdots\cdots\cdots\cdots\cdots(1)$$

The R.H.S. of above equation (1) can be expanded as a convergent series in powers of t. Thus

$$K_X(t) = K_1 t + K_2 \frac{t^2}{2!} + \cdots\cdots\cdots + K_r \frac{t^r}{r!} + \cdots\cdots\cdots\cdots = \log M_X(t)$$

$$= \log\left[1 + \mu_1^1 t + \mu_2^1 \frac{t^2}{2!} + \mu_3^1 \frac{t^3}{3!} + \cdots\cdots\cdots + \mu_r^1 \frac{t^2}{r!} + \cdots\cdots\right] \cdots\cdots\cdots\cdots\cdots(2)$$

Where $K_r =$ coefficient of $\frac{t^r}{r\,!}$ in $K_X(t)$ is called the r^{th} cumulant.

Hene

$$K_1 t + K_2 \frac{t^2}{2!} + K_3 \frac{t^3}{3!} + K_4 \frac{t^4}{4!} + \cdots\cdots\cdots\cdots = \left[\left[\mu_1^1 t + \mu_2^1 \frac{t^2}{2!} + \mu_3^1 \frac{t^3}{3!} + \mu_4^1 \frac{t^4}{4!} + \cdots\cdots\right]\right.$$

$$-\frac{1}{2}\left[\mu_1^1 t + \mu_2^1 \frac{t^2}{2!} + \mu_3^1 \frac{t^3}{3!} + \cdots\cdots\cdots\right]^2 + \frac{1}{3}\left[\mu_1^1 t + \mu_2^1 \frac{t^2}{2\,!} + \cdots\cdots\right]^3$$

$$\left.-\frac{1}{4}\left[\mu_1^1 t + \mu_2^1 \frac{t^2}{2\,!} + \cdots\cdots\cdots\cdots\right]^4 + \cdots\cdots\cdots\cdots\right]$$

Comparing the coefficients of like powers of 't' on both sides, we get the relationship between the moments and cumulants. Hence we have

$$K_1 = \mu_1^1 = \text{Mean}, \quad \frac{K_2}{2!} = \frac{\mu_2^1}{2!} - \frac{\mu_1^{1^2}}{2!} \Rightarrow K_2 = \mu_2^1 - \mu_1^{1^2} = \mu_2$$

$$\frac{K_3}{3!} = \frac{\mu_3^1}{3!} - \frac{1}{2}\,\frac{2\,\mu_1^1\,\mu_2^1}{2!} + \frac{\mu_1^{1^3}}{3} \Rightarrow K_3 = \mu_3^1 - 3\,\mu_2^1\mu_1^1 + 2\,\mu_1^{1^3} = \mu_3$$

$$\frac{K_4}{4\,!} = \frac{\mu_4^1}{4} - \frac{1}{2}\left[\frac{\mu_1^{1^2}}{4} + \frac{2\mu_1^1\,\mu_3^1}{3\,!}\right] + \frac{1}{3}\,\frac{3\mu_1^{1^2}\,\mu_2^1}{2} - \mu_1^{1^4} \Rightarrow K_4 = \mu_4^1 - 3\mu_2^{1^2} - 4\mu_1^1\,\mu_3^1 + 12\mu_1^{1^2}\mu_2^1 - 6\mu_1^{1^4}$$

$$\Rightarrow K_4 = \left(\mu_4^1 - 4\mu_3^1\,\mu_1^1 + 6\mu_2^1\,\mu_1^{1^2} - 3\mu_1^{1^4}\right) - 3\left(\mu_2^{1^2} - 2\mu_2^1\,\mu_1^{1^2} + \mu_1^{1^4}\right)$$

$$= \mu_4 - 3\left(\mu_2^1 - \mu_1^{1^2}\right)^2 = \mu_4 - 3\mu_2^2 = \mu_4 - 3K_2^2 \qquad (\because\ \mu_2 = K_2)$$

$$\left.\begin{aligned} \Rightarrow \mu_4 &= K_4 + 3K_2^2 \\ \text{Hence Mean} &= K \\ \mu_2 = K_2 = \text{Variance},\ \mu_3 = K_3,\ \mu_4 &= K_4 + 3K_2^2 \end{aligned}\right\} \cdots\cdots\cdots\cdots\cdots(3)$$

If we differentiate both sides of (2) w.r.t. 't' 'r' times and then put $t = 0$, we get

$$K_r = \left[\frac{d^2}{d\,t^2}\,K_X(t)\right]_{t=0} \qquad \cdots\cdots\cdots\cdots\cdots(4)$$

Additive Property of Cumulants:

The r^{th} cumulant of the sum of independent random variables is equal to the sum of the r^{th} cumulants of the individual variables.

i.e., $K_r\left(X_1 + X_2 + \cdots\cdots + X_n\right) = K_r(X_1) + K_r(X_2) + \cdots\cdots + K_r(X_n) \cdots\cdots\cdots(1)$

Where X_i ; $i = 1, 2, \cdots\cdots\cdots, n$ are independent random variables.

Proof: Since X_i 's are independent we have

$$M_{X_1+X_2+\cdots\cdots+X_n}(t) = M_{X_1}(t)\cdot M_{X_2}(t)\cdots\cdots M_{X_n}(t)$$

Taking logarithm on both sides we get

$$\log\ M_{X_1+X_2+\cdots\cdots+X_n}(t) = \log\left[M_{X_1}(t)\cdot M_{X_2}(t)\cdots\cdots M_{X_n}(t)\right]$$

$$\log\ M_{X_1+X_2+\cdots\cdots+X_n}(t) = \log\ M_{X_1}(t) + \log\ M_{X_2}(t) + \cdots\cdots + \log\ M_{X_n}(t)$$

$$K_{X_1+X_2+\cdots\cdots+X_n}(t) = K_{X_1}(t) + K_{X_2}(t) + \cdots\cdots + K_{X_n}(t) \cdots\cdots\cdots (2)$$

Differentiating both sides w.r.t. 't' 'r' times and putting $t = 0$, we get

$$\left[\frac{d^2}{st^2} K_{X_1+X_2+\cdots\cdots+X_n}(t)\right]_{t=0} = \left[\frac{d^2}{dt^2} K_{X_1}(t)\right]_{t=0} + \left[\frac{d^r}{dt^r} K_{X_2}(t)\right]_{t=0} + \cdots\cdots + \left[\frac{d^2}{dt^2} K_{X_n}(t)\right]_{t=0}$$

$$\Rightarrow K_r(X_1 + X_2 + \cdots\cdots + X_n) = K_r(X_1) + K_2(X_2) + \cdots\cdots + K_r(X_n)$$

Effect of Change of origin and scale on cumulants:

If we take $U = \frac{X-a}{h}$, then

$$M_U(t) = \exp(-at/h)\ M_X\ (t/h)$$

$$\therefore\ K_U\ (t) = \frac{-at}{h} + K_X\ (t/h)$$

$$\Rightarrow K_1^1 + K_1^1 t + K_2^1 \frac{t^2}{2!} + \cdots\cdots + K_2^1 \frac{t^2}{r!} + \cdots\cdots = \frac{-at}{h} + K_1(t/h) + K_2 \frac{(t/h)^2}{2!} + \cdots\cdots + K_r \frac{(t/h)^2}{r!} + \cdots\cdots$$

Where K_r^1 and K_r are the r^{th} cumulants of U and X respectively.

Comparing the coefficients, we get

$$K_1^1 = \frac{K_1 - a}{h} \quad \text{and} \quad K_1^1 = \frac{K_r}{h^r}\ ;\ \ r = 2, 3, \cdots\cdots\cdots$$

Hence we see that except the first cumulant, all cumulants are independent of change of origin. But the cumulants are not invariant of change of scale as the r^{th} cumulant of U is $\left(1/h^2\right)$ times the r^{th} cumulant of the distribution of X.

7.6 Probability Generating Function:

If $a_0, a_1, \cdots\cdots\cdots$ is a sequence of real numbers and if

$$G(s) = a_0 + a_1 S + a_2 S^2 + \cdots\cdots\cdots = \sum_{i=0}^{\infty} a_i S^i \qquad \cdots\cdots\cdots(1)$$

Converges in some interval $-S_0 < S < S_0$, when the sequence is infinite then the function $G(s)$ is known as the generating function of the sequence $\{a_i\}$.

The variable S has no significance of its own and is introduced to identify a_i as the coefficient of S_i in the expansion $G(S)$. If thesequence $\{a_i\}$ is bounded, then the comparison with the geometric series shows that $A(S)$ converges at least for $|S| < 1$.

The case when a_i is the probability that an integral valued discrete ariable X takes the value i,

i.e., $a_i = p_i = P(X = i)$; $i = 0, 1, 2, \cdots\cdots\cdots$ with $\sum p_i = 1$, then the probability generating function (p.g.f.) of r.v. X is defined as

$$G(S) = E(S^X) = \sum_{x=0}^{\infty} S^x \cdot P_x \qquad \cdots\cdots\cdots(2)$$

Effect of linear transformation on P.G.F.:

$$G(s : a + bX) = s^a \; G(s^a : X)$$

Proof: $G(s : a + bX) = E(s^{a+bX})$ (By definition)

$$= E(s^a \cdot s^{bX}) = s^a \cdot E\left[(s^b)^X\right] = s^a \cdot G \cdot (s^b : X)$$

Additive Property:

If X, Y are independent variates, then for constants a, b

$$G(s : aX + bY) = G(s^a : X) \cdot G(s^b : Y)$$

Proof: $G(s : aX + bY) = E(s^{aX+bY})$ (By definition)

$$= E(s^{aX} \cdot s^{bY})$$

$= E[s^{aX}] \cdot E(s^{bY})$ ($\because$ X and Y are independent)

$= G(s^a : X) \cdot G(s^b : Y)$

In general $G(s : X+Y) = G(s : X) \cdot G(s : Y)$

Relation between the P.G.F. and M.G.F.:

$M(t : X) = E(e^{tX}) \cdots\cdots\cdots\cdots (1)$

$G(s : X) = E(s^x) \cdots\cdots\cdots\cdots\cdots (2)$

Thus (1) is obtainable from (2) by changing s to e^t and conversely to obtain (2) from (1) changing e^t to s.

$\therefore M(t : X) = G(e^t : X)$;

$G(s : X) = M(\log s : X)$

For example (1) if $M(t : X) = (q + pe^t)^n$

then $G(s : X) = (q + ps)^n$

(2) if $G(s : X) = e^{m(s-1)}$

then $M(t : X) = e^{m(e^t - 1)}$

7.7 Characteristic Function:

For some distributions m.g.f. does not exist, since the series $\sum_x e^{tX} \cdot P(x)$ or the integral $\int e^{tX} f(x)\, dx$ does not converge for real values of the auxiliary parameter t. As such the m.g.f. of such discrete or continuous distribution fails to exist. $P(x) = \frac{K}{x^2}, \; x = 0, 1, 2, \cdots\cdots\cdots, n$ (distribution) or $f(x) = \frac{K}{(1+x^2)^n}, -\infty \le x \le \infty$ (continuous distribution) can be cited as example of such cases out of many. In this type of situations there is a more useful generating function, known as the characteristic function and is defined as under.

$\phi_X(t) = E(e^{itX}) = \sum_x e^{itX} f(x)$ (for descrete distribution $i = \sqrt{-1}$)

$$= \int_{-\infty}^{\infty} e^{itX} f(x)\, dx \qquad \text{for continuous distribution}$$

Also $\phi_X(t) = \int_{-\infty}^{\infty} e^{itX}\, d\, F(x)$ $\qquad \left[\because\ f(x) = F^1(x)\right]$

Here $\left|\phi_X(t)\right| = \left|\int e^{tX} \cdot f(x) dx\right| \leq \int \left|e^{itx}\right| f(x)\, dx$

i.e., $\leq \int \left|\cos tx + i \sin tx\right| f(x)\, dx$

$$\leq \int \left(\cos^2 tx + \sin^2 tx\right) f(x)\, dx$$

$$\leq \int f(x)\, dx$$

$$\leq 1 \qquad \left(\because \cos^2 tx + \sin^2 tx = 1\right)$$

$\because \left|\phi_X(t)\right| \leq 1$ so the characteristic funtion always exists though $M_X(t)$ may not exist. Thisis an advantage with the characteristic function over the m.g.f.

Properties of Characteristic Function:

1. $\phi_X(t) = 1 + it\, \mu_1^1 - \frac{t^2}{2} \mu_2^1 - \frac{it^3}{3!} \mu_3^1 + \frac{t^4}{4!} \mu_4^1 + \cdots\cdots\cdots$

where $\mu_r^1 = (-i)^{+r} \left[\frac{\partial^r}{\partial t^2} \phi_X(t)\right]_{t=0}$

Since by definition of $\phi_X(t)$ we have

$$\phi_X(t) = \int_{-\infty}^{\infty} e^{itX} f(x)\, dx \cdots\cdots\cdots\cdots\cdots (1)$$

Differentiating r times w.r.t. 't' and putting $t = 0$ we get from equation (1)

$$\left[\frac{\partial^r \phi_X(t)}{\partial t^2}\right]_{t=0} = \left[\int_{-\infty}^{\infty} (ix)^2 e^{itx} \cdot f(x)\,dx\right]_{t=0}$$

$$= (i)^r \int_{-\infty}^{\infty} x^r f(x)\,dx$$

$$= (i)^r \cdot E(X^r) = i^r \cdot \mu_r^1$$

$$\therefore \mu_1^1 = \left(\frac{1}{i}\right)^r \left[\frac{\partial^2}{\partial t^2} \phi_X(t)\right]_{t=0}$$

$$= (-i)^r \left[\frac{\partial^r}{\partial t^r} \cdot \phi_X(t)\right]_{t=0} \qquad \left(\because \frac{1}{i} = \frac{i}{i^2} = -i\right)$$

Thus, the coefficient of $\frac{(it)^r}{r!}$ in the expansion of $\phi_X(t)$ gives the r^{th} moment about zero $= \mu_r^1$. So like m.g.f. about zero $\phi_X(t)$ also generates the raw or crude moments.

2. $\phi_X(t)$ is defined or exists always with limite modulus This is the prominent advantage of the characteristic function over the m.g.f., as the later does not exists always.

   For example cauchy distribution

3. $\phi_X(t)$ is a uniformly continuous function of t.

4. $\phi_X(0) = \int_x f(x)\,dx = 1$, we have $\phi_X(t) = E(e^{itx}) \Rightarrow \phi_X(0) = E(e^0) = E(1) = 1$

5. $\phi_X(t)$ and $\phi_X(-t)$ are conjugate quantities and $\phi_X(t) = \phi_X(-t)$

$$\phi_X(t) = E(e^{itX}) = E[\cos tX + i \sin tX]$$

$$\overline{\phi_X(t)} = E[\cos t X - i \sin t X]$$

$$= E[\cos(-t) \times + i \sin(-t) X]$$

$$= E(e^{-itX}) = \phi_X(-t)$$

6. The characteristic function of the sum of two independent random variates is equal to the product of their individual characteristic functions.

Let $\phi_{X_1}(t)$ and $\phi_{X_2}(t)$ be the c.f.'s of two independent random variates X_1 and X_2 and let $\phi_{X_1+X_2}(t)$ be te c.f. of X_1+X_2, then

$$\phi_{X_1+X_2}(t) = E\left\{e^{it(X_1+X_2)}\right\} = E\left\{e^{itX_1} \cdot e^{itX_2}\right\} = E\left(e^{itX_1}\right) \cdot E\left(e^{itX_2}\right) = \phi_{X_1}(t) \cdot \phi_{X_2} t$$

which also can be extended to any number of variates X_i, $i = 1, 2, \cdots\cdots\cdots, n$ as $\phi_{X_1+X_2+\cdots\cdots+X_n}(t) = \phi_{X_1}(t) \cdot \phi_{X_2}(t) \cdots\cdots\cdots \phi_{X_n}(t)$.

7. Effect of changing the origin and scale.

We have $\phi_X(t) = E\left(e^{itX}\right)$

Changing the origin and scale by introducing a new variate U such that

$U = \frac{X-a}{h}$ i.e., $X = a + hU$, we have

$$\phi_X(t) = \phi_{a+hU}(t) = E\left[e^{it(a+hU)}\right] = e^{ait} \cdot E\left(e^{ithU}\right)$$

$$= e^{ait}\ E\left[e^{i(th)U}\right] = e^{iat} \cdot \phi_U(th)$$

or $\phi_U(th) = e^{-iat}\ \phi_X(t)$

Replacing ' t ' by t/h, this results

$$\phi_U(t) = e^{-iat/h} \cdot \phi_X(t/h) \text{ or } \phi_{\frac{(x-a)}{n}}(t) = e^{iat/h} \cdot \phi_X(t/h)$$

If $a = m, h = \sigma$ then

$$\phi \frac{(x-M)^{(t)}}{\sigma} = e^{-int/\sigma} \cdot \phi_X(t/\sigma)$$

8. If $F(x)$ and $\phi(t)$ be respectively, the distribution and characteristic functions, then density of the function is given by

$$f(x) = F^1(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \, \phi_X(t) \, dt$$

Provided $\phi(t)$ is integrable.

9. Uniqueness theorem of characteristic functions.

   The characteristic function uniquely determines the distribution function and conversely a particular distribution has unique characteristic function.

7.8 Tchebyshev's Inequality:

It is well known that the standard deviation is a measure of dispersion. Thus if the variance is small, large deviations from the mean are improbable. Hene this role of standard deviation is made quite obvious by a well - known Tchebyshev's inequality.

In equality: If X be random variable taking only non negative values, possess a finite mean μ and variance $\sigma^2 = \text{Var}(X)$, then for any $t > 0$.

$$P\left(|X - M| \geq t\right) \leq \frac{\sigma^2}{t^2}$$

Proof: We shall prove this inequality for a continuous variable we have for any $k > 0$.

$$E(X) = \mu = \int_0^{\infty} x \cdot f(x) \, dx$$

$$= \int_0^{K} x \cdot f(x) \, dx + \int_K^{\infty} x \, f(x) \, dx$$

$$\geq \int_K^{\infty} x \cdot f(x) \, dx \geq K \int_0^{\infty} f(x) \, dx = K \cdot P(X \geq K)$$

$$\therefore P(X \geq K) \leq \frac{E(X)}{K}$$

Now let us consider the random variable $Y = (X - \mu)^2$

then $E(Y) = E(X - \mu)^2 = \sigma^2$

Applying the result above, we get

$$P\left(Y \geq t^2\right) \leq \frac{\sigma^2}{t^2} \qquad \text{where} \quad K = t^2$$

or $\quad P\left((X-\mu)^2 \geq t^2\right) \leq \frac{\sigma^2}{t^2}$

i.e., $\quad P\left(|X-\mu| \geq t\right) \leq \frac{\sigma^2}{t^2}$

or $\quad P\left(|X-\mu| \leq t\right) \geq 1-\frac{\sigma^2}{t^2}$

Known as Tchebeyshev's inequality.

7.9 Worked Examples:

Example 1: A boy throws a coin four times an guesses ach time whether the head or the tail has been thrown. He was not allowed to see the results. He is to recieve 6 rupees for 2 heads, 1 rupee for 3 heads and 50 paise for 4 heads. Find his expectation.

Solution: The probability of getting x heads in 4 throws

$$= {}^4C_x \left(\frac{1}{2}\right)^x \cdot \left(\frac{1}{2}\right)^{4-x} ; x = 0,1,2,3,4. \qquad \left(\because\ p = q = \tfrac{1}{2}\right)$$

$\Rightarrow$ the probability of getting 2 heads in 4 throws is

$$= {}^4C_2 \left(\frac{1}{2}\right)^2 \left(\frac{1}{2}\right)^{4-2} = \frac{3}{8}$$

the probability of getting 3 heads in 4 throws is

$$= {}^4C_3 \left(\frac{1}{2}\right)^3 \cdot \left(\frac{1}{2}\right)^{4-3} = \frac{1}{4}$$

the probability of getting 4 heads in 4 throws is

$$= {}^4C_4 \left(\frac{1}{2}\right)^4 \cdot \left(\frac{1}{2}\right)^{4-4} = \frac{1}{16}$$

Theexpectation is, the boy is to recieve 6 Rs for 2 heads 1 Rs for 3 heads and 0.50 Rs for 4 heads.

$\therefore$ The expectation $= \frac{3}{8}$ of 6 Rs + $\frac{1}{4}$ of 1 Rs + $\frac{1}{16}$ of 0.50 Rs.

$$= \frac{81}{32} \text{ Rs.}$$

Example 2: Find $E(X)$ if a random variable X takes the values $x_K = \frac{(-1)^K \, 2^K}{K} \; (K = 1, 2, \cdots\cdots\cdots)$

with probabilities $p_K = \frac{1}{2^K}$

Solution: Since X is a random variable and takes te values.

$$x_K = \frac{(-1)^K 2^K}{K} \qquad \text{(given)} \; (K = 1, 2, \cdots\cdots\cdots)$$

$$x_1 = \frac{(-1)\,2}{1} = -2, \; x_2 = \frac{(-1)^2 \; 2^2}{2} = \frac{4}{2} = 2$$

$x_3 = \frac{(-1)^3 \; 2^3}{3} = \frac{-8}{3}, \; \cdots\cdots\cdots\cdots$ and probabilities $p_1 = \frac{1}{2}, \;\; p_2 = \frac{1}{2^2}, \, p_3 = \frac{1}{2^3}, \cdots\cdots\cdots$

we have $E(X) = p_1x_1 + p_2x_2 + p_3x_3 + \cdots\cdots\cdots$

$$= \frac{1}{2}(-2) + \frac{1}{2^2}(2) + \frac{1}{2^3}\left(\frac{-8}{3}\right) + \cdots\cdots\cdots\cdots$$

$$= -1 + \frac{1}{2} - \frac{1}{3} + \cdots\cdots\cdots\cdots\cdots$$

$$\backslash = -\left[1 - \frac{1}{2} + \frac{1}{3} \cdots\cdots\right] = -\log\,(1+1) = -\log 2 = -\log\,\frac{1}{2}$$

Example 3: If n dice are tossed and X denotes sum of the numbers on them, then find $E(X)$.

Solution: Denoting the number of i^{th} dice by x_i we have the sum of the numbers on n dice

$= X = x_1 + x_2 + \cdots\cdots\cdots + x_n$

$$\therefore\ E(X) = E(x_1 + x_2 + \cdots\cdots + x_n) = E(x_1) + E(x_2) + \cdots\cdots\cdot + E(x_n)$$

But for the i^{th} dice, the variate x_i can take the values as $1,2,3,4,5,6$ each with probability $1/6$ thus

$$E(x_i) = \frac{1}{6}\cdot 1 + \frac{1}{6}\cdot 2 + \frac{1}{6}\cdot 3 + \frac{1}{6}\cdot 4 + \frac{1}{6}\cdot 5 + \frac{1}{6}\cdot 6$$

$$= \frac{21}{6} = \frac{7}{2}$$

Hence $E(X) = \frac{7}{2} + \frac{7}{2} + \cdots\cdots\cdots n$ times $= \frac{7n}{2}$

Example 4: A box contains ' a ' white and ' b ' black balls, ' c ' balls are drawn. Find the expectation of the number of white balls drawn.

Solution: Let 'X' we the number of whites among the a balls drawn then dejining a variate x_i such that

$$\left.\begin{aligned} x_i &= 1 \text{ if } i^{th} \text{ ball drawn is white} \\ &= 0 \text{ y } i^{th} \text{ ball drawn is black} \end{aligned}\right\} i = 1,2,3,\cdots\cdots\cdots c$$

then $X = x_1 + x_2 + \cdots\cdots\cdots + x_c$ and $E(X) = E(x_1 + x_2 + \cdots\cdots\cdots + x_c)$

i.e., $E(X) = E(x_1) + E(x_2) + \cdots\cdots\cdots\cdot + E(x_c)$

$\because\ x_i$ takes values 1 and 0 with probabilities $a/a+b$ and $b/a+b$.

$$\therefore\ E(x_i) = 1\cdot\frac{a}{a+b} + 0\cdot\frac{b}{a+b} = \frac{a}{a+b}$$

Hene $E(X) = \frac{a}{a+b} + \frac{a}{a+b} + \cdots\cdots\cdots\cdots c$ times $= \frac{ac}{a+b}$

Example 5: Find the m.g.f. of mean of ' n ' independent observations of the variate X in terms of the m.g.f. of X.

Solution: Suppose the variate X assumes values $X_1 + X_2 + \cdots\cdots + X_n$. Distribution of each $X_i \ (i = 1, 2, \cdots\cdots\cdots, n)$ would be the same as that of the variate itself. Hence m.g.f. of each X_i would be equal to $M_X(t)$.

$$\therefore \text{ m.g.f. of } \frac{X_1}{n} = M_X\left(\frac{t}{n}\right) ; \text{ for } i = 1, 2, \cdots\cdots\cdots, n$$

also $\text{m.g.f. of } \dfrac{X_1 + X_2 + \cdots\cdots + X_n}{n} = M_X\left(\dfrac{t}{n}\right) \cdot M_X\left(t/n\right) \cdots\cdots\cdots\cdots n \text{ times}$

$$\text{m.g.f. of } \overline{X} = \left[M_X\left(t/n\right)\right]^n$$

Example 6: Find the m.g.f. of a random variable whose moments are $\mu_r^1 = (r+1)!\, 2^r$

Solution: Sine $M_X(t) = 1 + t\,\mu_1^1 + \dfrac{t^2}{2!}\mu_2^1 + \cdots\cdots\cdots + \dfrac{t^2}{r!}\mu_r^1 + \cdots\cdots\cdots$

$$= \sum_{r=0}^{\infty} \frac{t^2}{r!}\mu_r^1$$

$$= = \sum_{r=0}^{\infty} \frac{t^2}{r!}(r+1)!\, 2^r = \sum_{r=0}^{\infty} (r+1)(2t)^r$$

$$= 1 + 2(2t) + 3(2t)^2 + 4(2t)^3 + \cdots\cdots\cdots\cdots$$

$$= (1 - 2t)^{-2} = M_X(t)$$

Example 7: Find $M_X(t)$ when the c.d.f. of X is $F_X(x) = 0,\ x < 0$, $F_X(x) = 1 - \dfrac{1}{2}e^{-x},\ x \geq 0$.

Solution: Here $F(x) = 0$, when $x < 0, F(0) = 1/2$;

hene $P(X = 0) = \dfrac{1}{2}$ for $x > 0,\ f(x) = F^1(x) = \dfrac{1}{2}e^{-x}$, Hence

$$M_X(t) = E\left(e^{tX}\right) = \frac{1}{2} e^{ot} + \int_0^{\infty} e^{tx} \frac{1}{2} e^{-x} \, dx$$

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

$$= \frac{2-t}{2(1-t)}, \; t < !$$

Example 8: Show that $\phi_X(t)$ is a uniformly continuous function of t.

Solution: Since $\phi_X(t) = \int_{-\infty}^{\infty} e^{itx} f(x) \, dx \Rightarrow \phi_X(t+h) = \int_{-\infty}^{\infty} e^{i(t+h)x} f(x) \, dx$

$$\left|\phi_X(t+h) - \phi_X(t)\right| = \left| \int_{-\infty}^{\infty} \left\{e^{i(t+h)x} - e^{itx}\right\} \right| dF(x)$$

$$\leq \int_{-\infty}^{\infty} \left|e^{itx}\right| \left|e^{ihx} - 1\right| \, dF(x)$$

$$\leq \int_{-\infty}^{\infty} \left|e^{ihx} - 1\right| \, dF(x) \text{ as } \left|e^{itx}\right| = 1$$

$$\leq \int_{-\infty}^{\infty} \left\{\left|e^{ihx}\right| + 1\right\} dF(x) \quad \text{as} \quad \left|e^{ihx} - 1\right| \leq \left|e^{ihx}\right| + 1$$

$$\leq 2 \int_{-\infty}^{\infty} dF(x) \text{ as } \left|e^{ihx}\right| = 1$$

$$\leq 2 \quad \left(\because \int_{-\infty}^{\infty} dF(x) = 1 \text{ for p.d.f}\right)$$

i.e., $\phi_X(t)$ is bounded.

Also $\lim\limits_{h \to 0} \left|\phi_X(t+h) - \phi_X(t)\right| \le \int\limits_{-\infty}^{\infty} \lim\limits_{h \to 0} \left|e^{ihx} - 1\right| d\,F(x) = 0$

$$\Rightarrow \lim_{h \to 0} \phi_X(t+h) = \phi_X(t)\ \forall\ t$$

Hence $\phi_X(t)$ is a uniformly continuous function of t.

Example 9: Find the characteristic function of the random variate X which assums, values $x_1 = -1$ and $x_2 = 1$ with probabilities $\frac{1}{2}, \frac{1}{2}$ each.

Solution: Let x_1, x_2 be the values assumes by random variate X with probabilities p_1, p_2 then by definition of characteristic function we have

$$\phi_X(t) = E\left(e^{itx}\right) = p_1\, e^{itx_1} + p_2\, e^{itx_2}$$

$$= \frac{1}{2} e^{-it} + \frac{1}{2} e^{it} = \cos t$$

$$\left(\because x_1 = -1,\ x_2 = 1,\ p_1 = 1/2,\ p_2 = 1/2 \text{ given}\right)$$

Example 10: A random variable x has the density function e^{-x} for $x \ge 0$ show that Tchelochev's inequality gives $P\left\{|X - 1| > 2\right\} < \frac{1}{4}$ and show that the actual probability is e^{-3}.

Solution: Here $t = 2$,

Mean is given by

$$E(X) = \mu_1^1 = \int_0^\infty x \cdot e^{-x}\, dx = \int_0^\infty x^{2-1} e^{-x} dx = \Gamma(2) = 1$$

$$\mu_2^1 = \int_0^\infty x^2\, e^{-x}\, dx = \int_0^\infty x^{3-1}\, e^{-x}\, dx = \Gamma(3) = 2$$

$$\therefore\ \sigma^2 = \mu_2^1 - \mu_1^{1^2} = 2 - 1 = 1$$

Substituting the values of t, mean and variance on Tchebychev's inequality we get

$$P\{1 \times -11 \geq 2\} \leq 1/4$$

Since $x \geq 0, |x-1| > 2 \Rightarrow x + > 2\ x > 3$

Hence the actual probability is given by

$$p = \int_3^\infty e^{-x}\, dx = \left[-e^{-x}\right]_3^\infty = e^{-3} - e^{-\infty} = e^{-3} - 0 = e^{-3}$$

Example 11: A variate X has mean 50 and variance 100. Find the following (i) $P(X \geq 65) + P(X \leq 35)$, (ii) $P\{|X-50| < 20\}$, (iii) $P(30 < X < 70)$, (iv) Values of t that make $P(|X-50| \geq t) \leq 0 \cdot 01$.

Solution: If $Z = (x-M)/\sigma$ is standard r.v. then Techebychev's in equality is

$$P\{|Z| \geq K\} \leq 1/K^2, \quad \text{or} \quad P\{|Z| < K\} > 1 - \left(1/K^2\right) \cdots\cdots\cdots\cdots\cdots (1)$$

Here $Z = \dfrac{(X-50)}{10}$

(i) $P\left(Z > 3/2\right) + P\left(Z \leq -3/2\right) = P\left(|Z| \geq 3/2\right) \leq 4/9$

(ii) $P\left(|Z| < 2\right) > 1 - 1/4 = 3/4$

(iii) $P(-2 < Z < 2) = P(|Z| < 2) = 3/4$

(iv) $P\left(|Z| \geq t/10\right) \geq 100/t^2 \Rightarrow \left(\dfrac{100}{t^2}\right) \leq 0 \cdot 01 \Rightarrow t^2 \geq (100)^2$, i.e., $t \geq 100$

Example 12: A sample of size n is drawn from a population whose mean is 5 and S.D.I. prove that

$$P\left\{|\overline{X} - 5| < 0 \cdot 001\right\} \geq 1 - \left(\frac{10^6}{n}\right)$$

Solution: Here $E(\overline{X}) = \mu = 5,\ Var(\overline{X}) = \frac{\sigma^2}{n} = \frac{1}{n},\ K = 10^{-3}$

By using Tchebychev's inequality we have

$$P\{|X-M| < K\} \geq 1-\left(\frac{\sigma^2}{K^2}\right)$$

$$\Rightarrow P\{|X-5| < 10^{-3}\} \geq 1-\frac{10^6}{n}$$

$$\Rightarrow P\{|X-5| < 0 \cdot 001\} \geq 1-\frac{10^6}{n}$$

Example 13: A discrete variate X is specified by $f(-a) = f(a) = 1/8,\ \ f(0) = 3/4$ compute $P(|X| \geq 2\sigma)$ and compare it with Techychev's inequality bound.

Solution: Here $E(X) = \frac{-a}{8} + \frac{a}{8} + \frac{0 \cdot 3}{4} = 0$

$$E(X^2) = \frac{a^2}{8} + \frac{a^2}{8} = \frac{1}{4}a^2$$

$$\mu = 0,\ \sigma^2 = V(X) = E(X^2) - \{E(X)\}^2 = \frac{a^2}{4} - 0 = a^2/4$$

or $\sigma = \sqrt{a^2/4} = a/2$

Now by Techebychev's inequality we have

$$P\{|X| \geq 2\sigma\} = P\{|X-M| \geq a\} \leq 1/4$$

Actually $p = P\{|X| \geq 2\sigma\} = 1-\{|X| < a\} = 1-P\{-a < X < a\}$

$$= 1 - P(X = 0) = 1 - 3/4 = 1/4$$

Hene Techychev's upper bound coincides with actual values and so the upper bound is atttained.

7.10 Exercise:

1. Let X be a random variable with the following probability distribution.

x	:	-3	6	9
$P(X=x)$	:	$1/6$	$1/2$	$1/3$

Find $E(X)$ and $E(X^2)$ and using the laws of expectation evaluate $E(2X+1)^2$.

$$\left[\text{Ans}: \frac{11}{2}, \frac{93}{2}, 209\right]$$

2. Find the expectation fo the number on a die when thrown and also find the expected values of the sum of numbers of points when two unbiased dice are throuwn.

$$\left[\text{Ans}: E(X)=\frac{1}{2}, E(X)=7\right]$$

3. Let X be r.v. with man μ and variance σ^2. Show that $E(X-b)^2$, as a function of b, is minimised when $b=\mu$.

4. If t is any positive real number, show that the function defind by

$$P(x)=e^{-t}\left(1-e^{-t}\right)^{x-1}$$

can represent a probability function of a random variable X assuming the values $1,2,3,\cdots\cdots\cdots$ find the $E(X)$ and $Var(X)$ of the distribution.

$$\left[\text{Ans}: E(X)=e^{t}, V(X)=e^{t}\left(e^{t}-1\right)\right]$$

5. Let the random variable X assume the value ' r ' with the probability law

$$P(X=r)=q^{r-1}\cdot p \; ; \; r=1,2,3,\cdots\cdots\cdots$$

Find the m.g.f. of X and hence its mean and variance.

$$\left[\text{Ans}: M_X(t)=\frac{pet}{1-qet}, \mu_1^1=\frac{1}{p}, \mu_2 = q/p^2\right]$$

6. The probability density funtion of the random variable X follows the following probability law.

$$P(x) = \frac{1}{2\theta} \exp\left(\frac{-|x-\theta|}{\theta}\right), -\infty < x < \infty$$

Find the m.g.f. of X. Hence or otherwise find $E(X)$ and $V(X)$.

$$\left[\text{Ans}: M_X(t) = 1 + \theta t + \frac{3\theta^2 t^2}{2!} + \cdots, \mu_1^1 = \theta, \mu_2 = 2\theta^2\right]$$

7. For a distribution, the cymmulants are given by

$$K_r = n\left[(r-1)!\right], n > 0$$

Find the characteristic function.

8. Find the density function $f(x)$ corresponding to the characteristic function defined as follows:

$$\phi(t) = \begin{cases} 1-|t|, & |t| \le 1 \\ 0, & |t| > 1 \end{cases}$$

9. Let X be a random variable with generatingfunction $P(S)$. Find the generating funtion of (a) $X+1$, (b) $2X$.

$$\left[\text{Ans}: (a)\, S \cdot P(S),\ (b)\, P(S)^2\right]$$

10. Find the generating function of (a) $P(X \le n)$, (b) $P(X < n)$ and (c) $P(X = 2n)$.

11. If $P(S)$ is the probability generating function for X, find the generating function for $(X-a)/b$. $\left[\text{Ans}: S^{-a/b} \cdot P\left(S^{1/b}\right)\right]$

12. For geometric distribution $P(x) = 2^{-x}; x = 1, 2, 3, \cdots\cdots$ prove that Theloychev's inequality gives

$$P\{|X-2| \le 2\} > 1/2$$

while the actual probability is $15/16$.

13. Does there exist a variate X for which $P\{\mu_x - 2\sigma \le X \le \mu_x + 2\sigma\} = 0\cdot 6$

14. If X is the number scored in a throw of a fair die, show that the Tchetycev's inequality gives $P\{|X-\mu| > \cdot 5\} < 0\cdot 47$ where μ is the mean of X, while the actual probability is zero.

15. Two unbiased dice are thrown. If X is the sum of the numbers showing up, prove that $P\{|X-7| > 3\} \le 35/54$. Compare this with the actual probability.

16. A symmetric die is throiwn 600 times. Find the where bound for the probability of getting 80 to 120 sixes.

$\left[\text{Ans}: P\{80 \le S \le 120\} \ge 19/24\right]$

17. Use Tchebychev's inequality to determine how many times a fair win must be bassed in order that the probability will be at least 0.90 that the ratio of the observed number of heads to the number of tosses will lie between 0.4 and 0.6

$[\text{Ans}: n = 250]$

18. Thirteen eards are drawn simultaneously from a pack of 52. If aces count 1, face cards 10 and other according to their denominations, find the expectation of the total score on the 13 cards. $[\text{Ans}: E(X) = 85]$

Lesson – 8

WEAK LAW OF LARGE NUMBERS & CENTRAL LIMIT THEOREM

Object of Lesson:

After studying this lesson the student is expected to have a clear comprehension of the theory and the practical utility about the concepts of weak law of large numbers, central limit theorem and their applications.

Structure of The Lesson:

This consists of sections as detailed below :

8.1 Introduction

8.2 Elements of Weak Law of Large Numbers

8.3 Elements of Central Limit Theorem

8.4 Workedout Examples

8.5 Exercise

8.1 Introduction:

In repeated experimentation, the observations of the experiment start stabilizing at some value or around some value thus giving regularities of the experiment. This phenomenon can be explained by law called Law of Large Numbers.

8.2 Elements of Weak Law of Large Numbers:

Let 'X' be random variable with density $f(X)$ and let its expected value be $E(X) = \mu$. Then $E(X)$ is an average of an infinite number of values. Here the problem is that using a finite number of values of x say n, can a reliable conclusion be made about $E(X)$, the average of an infinite number of values of X. For this answer is given by the weak law of large numbers which states that, if a random sample of size n or larger is taken from a population with density $f(x)$, the probability that the sample mean $\overline{X}$ will deviate from $E(X) = \mu$, the population mean, by any arbitary small quantity can be made as near to one as desired, or in other words, for an $\in > 0$ and $0 < \delta < 1$, there exists an integer 'n' such that for $m \geq n$.

$$P\left\{\left|\overline{X}_m - \mu\right| < E\right\} > 1 - \delta \quad \text{as } n \to \infty$$

Such types of weak law of large numbers are of different variaties proposed by different persons by changing the hypothetical conditions. The statements and proofs of all such results are beyond the scope of this study material. However if we want to know the applications, a few of them are -

1. The converging limit of the arithmatic means of a series of observations such as the average temperature of a resion the average rainfall in a monsoon the average height to which an individual can grow etc. In all these cases the mean of the first 'n' observations will be calculated and the limiting behaviour of all these means as 'n' becomes larger can be assessed through probability distributions and expectations.

8.3 Elements of Central Limit Theorem:

In a collection of indefinitely many observations for a randim variables. The arithmatic mean is calculated the distribution of the arithmatic mean may not be known exactly in an analytic form. But as the size of the sample increases the graph of the distribution function for the arithmatic mean is likely to follow a definite shape and stabilizes to that shpae ultimately. In statistical science we always consider mean as the best average and standard deviation as the best dispersion method given these two measures a complete discription of the distribution spread over the entire real line is given by only normal distribution. Hence if the limiting distribution of the arithmatic mean of the sample whatever may be its parent is modelled by a normal distribution we say that central limit property holds good for that data.

In literature central limit theorem proved for all the cases of i.i.d. random variables. This result can be used by various practitioners in this hypotesis testing, analysis of variance interval estimation etc., whenever use of normal distribution rises the only requiremet is the data should be sufficiently large.

8.4 Workedout Examples:

Example 1: If the variable X_k assumes the value $2^{r-2\log r}$ with probability 2^{-r} ; $r = 1, 2, \cdots\cdots$ examine whether WLLN holds in this case.

Solutionl: Putting $k = 1, 2, 3, \cdots\cdots$ the values of the identical variables $X_1, X_2, X_3, \cdots\cdots\cdots$ are respectively $2^{1-2\log^1}, 2^{2-2\log^2}, 2^{3-2\log^3}, \ldots\ldots\ldots\ldots$ with probabilities $\frac{1}{2}, \frac{1}{2^2}, \frac{1}{2^3}, \cdots\cdots\cdots$; so that

$$E(X_r) = E(X_1 + X_2 + X_3 + \cdots\cdots\cdots) = E(X_1) + E(X_2) + E(X_3) + \cdots\cdots\cdots$$

$$= 2^{1-2\log^1} \cdot \frac{1}{2} + 2^{2-2\log^2} \cdot \frac{1}{2^2} + 2^{3-2\log^3} \cdot \frac{1}{2^3} + \cdots\cdots\cdots$$

$$= \sum_{r=1}^{n} 2^{r-2\log^r} \cdot 2^{-r}$$

$$= \sum_{r=1}^{\infty} \frac{1}{2^{2\log r}} = \sum_{r=1}^{\infty} \frac{1}{e^{\log 2^{2\log^{r}}}} \qquad \left(\because\ k = e^{\log\ k}\right)$$

$$= \sum_{r=1}^{\infty} \frac{1}{e^{\log^{r}\cdot\log^{4}}} = \sum_{r=1}^{\infty} \frac{1}{e^{\log^{4}\cdot\log^{r}}} = \sum_{r=1}^{\infty} \frac{1}{e^{\log^{r}\ \log^{4}}}$$

$$= \sum_{r=1}^{\infty} \frac{1}{r^{\log^{4}}}$$

$$= 1 + \frac{1}{2^{\log^{4}}} + \frac{1}{3^{\log^{4}}} + \cdots\cdots\cdots$$

where the R.H.S. is a convergent series since $\log_e^4 > 1$ (In the test $\sum 1/n^p,\ p > 1$), there by showing that the mathematical expectation of variates $X_r,\ r = 1, 2, \cdots\cdots$ exists and hence the weak law of large numbers holds in this case (Due to Khintchine theorem).

Example 2: If $\{X_i\}$ be mutually independent and identically distributed random variables with mean μ and finite variance and if $S_n = X_1 + X_2 + \cdots\cdots\cdots + X_n$, then prove that the law of large numbers does not hold for the sequence $\langle S_n \rangle$.

Solution: Since $S_n = X_1 + X_2 + \cdots\cdots\cdots + X_n$

$\therefore \quad S_1 = X_1\ ;\ S_2 = X_1 + X_2\ ;\ S_3 = X_1 + X_2 + X_3\ ;\ \cdots\cdots\cdots\cdots$ etc.

Thus the variates are $S_1, S_2, \cdots\cdots\cdots, S_n$

$$B_n = \text{Var}\left(S_1 + S_2 + \cdots\cdots\cdots\cdots + S_n\right)$$

$$= \text{Var}\left(X_1 + (X_1 + X_2) + \cdots\cdots\cdots\cdots + (X_1 + X_2 + \cdots\cdots\cdots + X_n)\right)$$

$$= \text{Var}\left(n X_1 + (n-1) X_2 + \cdots\cdots\cdots + 2X_{n-1} + X_n\right)$$

$$= n^2\ \text{Var}\left(X_1\right) + (n-1)^2\ \text{Var}\left(X_2\right) + \cdots\cdots\cdots + 2^2\ \text{Var}\left(X_{n-1}\right) + 1^2\ \text{Var}\left(X_n\right)$$

Also, variables being identically distributed,

$$\text{Var}\left(X_1\right) = \text{Var}\left(X_2\right) = \cdots\cdots\cdots = \text{Var}\left(X_n\right) = \sigma^2$$

$$\therefore\ B_n = \left[n^2 + (n-1)^2 + \cdots\cdots\cdots + 2^2 + 1^2 \right] \sigma^2$$

$$= \frac{n(n+1)(2n+1)}{6} \sigma^2$$

or $$\frac{Bn}{n^2} = \frac{(n+1)(2n+1)\sigma^2}{6n}$$

giving $$\lim_{n \to \infty} \frac{Bn}{n^2} = \infty \neq 0$$

Hence the law of large numbers does not hold for the sequence $\{S_n\}$.

Example 3: Show that the following sequence does not obey WLLN:

$$P\left\{ X_k = \pm (2k-1)^{1/2} \right\} = 1/2$$

Solution: We have

$$E(X_k) = \frac{1}{2}(2k-1)^{1/2} - 1/2\,(2k-1)^{1/2} = 0$$

$$Var(X_k) = E(X_k)^2 - \left[E(X_k)\right]^2$$

$$= E(X_k)^2 - 0 \qquad (\because E(X_k) = 0)$$

$$= E(X_k)^2 = \frac{1}{2}(2k-1) + \frac{1}{2}(2k-1) = 2k-1$$

$$B_n = \sum_{k=1}^{n} V(X_k) = \sum_{k=1}^{n} (2k-1) = n^2$$

$\because \lim_{n \to \infty} \frac{B_n}{n^2} = 1 \neq 0$ it follows that for $\{X_n\}$ the WLLN does not hold.

Example 4: Avariate X_k has the distribution

$$P(X_k = 0) = 1 - \left(\frac{2}{3^{2k+2}} \right), \; P\left(X_k = 3^k\right) = p\left(X_k = -3^k\right) = 3^{-(2k+2)}$$

Does the WLLN hold for the sequence $\{X_k\}$?

Solution: Here $E(X_k) = 3^k \cdot 3^{-(2k+2)} - 3^k \cdot 3^{-(2k+2)} = 0$

$$E\left(X_k^2\right) = 3^{2k} \cdot 3^{-2k-2} + 3^{2k} \cdot 3^{-2k-2} = \frac{2}{9}$$

$$\therefore \quad Var(X_k) = E\left(X_k^2\right) - \left[E(X_k)\right]^2$$

$$Var\,(X_k) = E\left(X_k^2\right) = \frac{2}{9}$$

$$B_n = \sum Var\,(X_i) = \left(\frac{2}{9n}\right) \qquad i = 1, 2, \cdots\cdots\cdots\cdots, n$$

As $\frac{B_n}{n^2} = \left(\frac{2}{9n}\right) \to 0$ as $n \to \infty$ it follows that WLLN holds for the sequence $\{X_n\}$.

Example 5: Let $\{X_n\}$ be a sequence of mutually independent variates such that

$$P\{X_n = \pm 1\} = \frac{1}{2}\left(1 - 2^{-n}\right), \; P\left\{X_n = \pm 2^{-n}\right\} = 2^{-n-1}$$

Does the WLLN hold for this sequence ?

Solution: Here $E(X_n) = \frac{1}{2}\left(1 - 2^{-n}\right) - \frac{1}{2}\left(1 - 2^{-n}\right) + 2^{-n} \cdot 2^{-n-1} - 2^{-n}\left(2^{-n-1}\right) = 0$

$$E\left(X_n^2\right) = \frac{1}{2}\left(1 - 2^{-n}\right) + \frac{1}{2}\left(1 - 2^{-n}\right) + 2^{-2n} \cdot 2^{-n-1} + 2^{-2n} \cdot 2^{-n-1}$$

$$= 1 - 2^{-n} + 2^{-3n} = Var(X_n)$$

$$B_n = \sum_{r=1}^{n} \sigma_r^2 = \sum_{r=1}^{n} \left(1 - 2^{-2} + 2^{-3r}\right) = n - \left(1 - 2^{-n}\right) + \frac{1}{7}\left(1 - 8^{-n}\right)$$

$$(r = 1, 2, \cdots\cdots\cdots\cdots, n)$$

$$\frac{B_n}{n^2} = \frac{1}{n} + \frac{1}{2^n \cdot n^2} - \frac{6/7}{n^2} - \frac{1}{7}\frac{1}{8^n} \cdot \frac{1}{n^2} \to 0 \quad \text{as } n \to \infty$$

It follows that the WLLN holds for the given sequence.

Example 6: Prove that the WLLN is applicable to the arithmetic mean of a sequence of independent variates X_k specified by

$$P\left\{ X_k = \pm\sqrt{\log k} \right\} = 1/2$$

Solution: Here $E(X_k) = \frac{1}{2}(\log K)^{1/2} - \frac{1}{2}(\log k)^{1/2} = 0$

Var $(X_k) = E\left(X_k^2\right) - \left[E(X_k)\right]^2 = E\left(X_k^2\right)$

$$= \frac{1}{2}\log k + \frac{1}{2}\log k = \log k$$

$\therefore B_n = \sum \sigma_1^2 = \log 1 + \log 2 + \cdots\cdots\cdots + \log n = \log n!$

Using Stirlings approximation to n!,

$$\frac{Bn}{n^2} = \frac{\log\left(e^{-n} \; n^{n+1/2} \cdot \sqrt{2\pi}\right)}{n^2} = \frac{1}{n^2}\left\{\left(n + \frac{1}{2}\right)\log n + \log\sqrt{2\pi}\right\}$$

$$= \left(1 + \frac{1}{2n}\right)\log(n)^{1/n} + \frac{\left(\log\sqrt{2\pi} - n\right)}{n^2}$$

Now as $n \to \infty$, $n^{1/n} \to 1$, so that each term $\to 0$.

Hence $\lim_{n \to \infty} \frac{Bn}{n^2} = 0$, as $n \to \infty$.

This proves the apllicability of the weak law of large numbers.

Example 7: If $X_1, \cdots\cdots\cdots\cdots, X_n$ are i.i.d. variaties, with p.m.t. $P(X_i = \pm 1) = 1/2$ show that central limit theorem holds for this sequence.

Solution: Here $E(X) = 1\left(\frac{1}{2}\right) + (-1)\left(\frac{1}{2}\right) = 0$

$$\text{Var}(X) = E\left(X^2\right) = 1$$

Let $\frac{1}{n} = \frac{\left(X_1 + X_2 + \cdots\cdots + X_n\right)}{\sqrt{n}}$

$$\therefore\ E\left(Y_n\right) = 0,\ \text{Var}\left(Y_n\right) = n^{-1}\ \text{Var}\left(S_n\right) = 1 \qquad \left[\because\ \text{Var}\left(S_n\right) = n\right]$$

$$\phi\ (t : X) = E\left(e^{it\,X}\right)$$

$$= \frac{1}{2}e^{it} + \frac{1}{2}e^{-it} = \cos t$$

$$\therefore\quad \phi\left(t : Y_n\right) = \phi\left(\frac{t}{\sqrt{n}} : \Sigma X_i\right) = \left[\phi_X\left(t/\sqrt{n}\right)\right]^n \qquad \left[\because\ X_i\ \text{are i.i.d.}\right]$$

$$= \left[\cos\left(t/\sqrt{n}\right)\right]^n = \left[1 - \left(\frac{t^2}{2n}\right) + O\left(\frac{1}{n}\right)\right]^n$$

as $n \to \infty$

$$\ell\text{im}\ \phi\ \left(t : Y_n\right) = \ell\text{im}\ \left[1 - \left(\frac{t^2}{2n}\right) + O\left(\frac{1}{n}\right)\right]^n = e^{-1/2^{+2}}$$

Which is the characteristic function of $N(0,\ 1)$. It follows that $Y_n \to N(0, 1)$ as $n \to \infty$, thus showing that L.L.T. holds for the sequence $\{X_n\}$.

8.5 Exercise:

1. Examine whether the weak law of large numbers holds for the sequence $\{X_k\}$ of independent random variables defined as follows:

$$P\left[X_k = \pm\, 2^k\right] = 2^{-(2k+1)}$$

$$P\left[X_k = 0\right] = 1 - 2^{-2k}$$

2. If X_n takes the values 1 and 0 with corresponding probabilities p_n and $1-p_n$, examine whether the weak law of large numbers can be applied to the sequence $\{X_n\}$ where the variables X_n, $n=1,2,\cdots\cdots\cdots$ are independent.

3. $\{X_i\}$, $i=1,2,\cdots\cdots\cdots$ is a sequence of independent random variables with expected value of X_i equal to m_i and variance of X_i is σ_i^2. If $\frac{1}{n^2}\sum_{i=1}^{n}\sigma_i^2$ tends to zero as n tends to infinity, show that the weak law of large numbers holds good to the sequence.

4. Examine whether the weak law of large numbers holds good for the sequence X_n of independent random variables, where

$$P\left\{X_n=\frac{1}{\sqrt{n}}\right\}=\frac{2}{3},\ P\left\{X_n=\frac{-1}{\sqrt{n}}\right\}=\frac{1}{3}.$$

5. If $X_1+X_2+\cdots\cdots\cdots+X_n$ be r.v.'s with means $\mu_1,\mu_2,\cdots\cdots\cdots\cdots,\mu_n$ and standard deviations $\sigma_1,\sigma_2,\cdots\cdots\cdots\cdots,\sigma_n$ and if $\sigma_n\to 0$ as $n\to\infty$, show that $\overline{X}_n-\overline{\mu}_n$ convergs to zero stochastically.

6. Show that if m is the number of successes in n independent trials, the probability of success at the i^{th} trial being p_i then m/n converges in probability to $(p_1+p_2+\cdots\cdots+p_n)/n$.

Lesson – 9

UNIFORM AND BINOMIAL DISTRIBUTIONS

Objectives:

After studying the lesson the student is expected to have clear comprehension of the theory and practical utility about the concepts of discrete Uniform Distribution and Biononial Distribution.

Structure of The Lesson:

This lesson consists of 3 sections as detailed below:

9.1 Discrete Uniform Distribution:

Definition: A random variable X is said to have a discrete uniform distribution over the range $[1, n]$ if its p.m.f. is given as follows:

$$P(X = x) = \frac{1}{n} \; ; \; x = 1, 2, \cdots\cdots, n \qquad (9.1.1)$$

$$= 0 \text{ otherwise}$$

Here n is known as the parameter of the distribution and it takes the set of all positive integers. Equation (9.1.1) is also called a discrete rectangular distribution.

9.1.1 Moments:

$$\text{Mean } = E(X) = \sum x \, P(x)$$

$$= \frac{1}{n} \sum_{x=1}^{n} x$$

$$= \frac{n(n+1)}{2n}$$

$$= \frac{n+1}{2}$$

$$E(X^2) = \sum x^2 \, P(x)$$

$$= \frac{1}{n} \sum_{x=1}^{n} x^2$$

$$= \frac{1}{n} \left[\frac{n(n+1)(2n+1)}{6} \right]$$

$$= \frac{(n+1)(2n+1)}{6}$$

$$\text{Variance } = V(X) = E(X^2) - [E(X)]^2$$

$$= \frac{(n+1)(2n+1)}{6} - \frac{(n+1)^2}{4}$$

$$= \frac{n^2 - 1}{12}$$

9.1.2 Moment Generating Function:

The M.G.F. of uniform distribution is given by

$$M_X(t) = E\left(e^{tX}\right) = \sum_{x=\phi}^{n} e^{tx} \cdot P(x)$$

$$= \frac{1}{n} \sum_{n=\phi}^{n} e^{tx}$$

$$= \frac{e^t\left(1-e^{nt}\right)}{n\left(1-e^t\right)}$$

9.1.3 Characteristic Function:

The c.f. of uniform distribution is given by

$$d_X(t) = E\left(e^{itX}\right)$$

$$= \sum_{x=1}^{n} e^{itX} P(x)$$

$$= \frac{1}{n} \sum_{x=1}^{n} e^{itx}$$

$$= \frac{e^{it}\left(1-e^{n\,it}\right)}{n\left(1-e^{it}\right)}$$

9.2 Bernoull's Distribution:

Definition: A random variable X is said to have a Bernoulli distribution with parameter p if its p.m.f. is given by

$$P(X=x) = P^x\, q^{1-x} \quad ; \quad x = 0, 1, \cdots\cdots\cdots$$

$$= 0 \quad \text{otherwise} \qquad\qquad \text{where } q = 1-p$$

The parameter p always lies btween 0 and 1.

A random experiment whose outcomes are of two types they are success (S) and failure (F) occurring with probabilities of p and q respectively is called a Bernoulli experiment. In this experiment r.v. X takes the values 1 and 0 respectively with occurance of S and F.

9.2.1 Mean and Variance of Bernoulli Distribution:

$$\text{Mean} = E(X) = \sum x \cdot P(x)$$

$$= 0 \cdot q + 1 \cdot p = p$$

$$E(X^2) = \sum x^2 \, P(x)$$

$$= 0^2 \cdot q + 1^2 \cdot p$$

$$= p$$

$$\therefore \text{ Variance } = V(X) = E(X^2) - [E(X)]^2$$

$$= p - p^2$$

$$= P(1-P)$$

$$= Pq$$

$\therefore$ Mean = p and Variance = Pq

9.2.2 Moment Generating Function:

If X is a Bernoulli variate with parameter p then its M.G.F. is given by

$$M_X(t) = E(e^{tX})$$

$$= \sum e^{tX} \, P(x)$$

$$= e^{t(0)} \, P(X=0) + e^{t(1)} \, P(X=1)$$

$$= 1 \cdot q + e^t \, p$$

$$= q + Pe^t$$

9.2.3 Characteristic Function:

The c.f. of Bernoulli Distribution is given by

$$d_X(t) = E[e^{itX}]$$

$$= \sum_{x=0}^{1} e^{itx} \cdot P(x)$$

$$= e^{it(0)}\,P(X=0) + e^{it(1)}\cdot P(X=1)$$

$$= 1\cdot q + e^{it}\cdot p$$

$$= q + Pe^{it}$$

9.3 Binomial Distribution:

Introduction:

Binomial distribution was discovered by James Bernoulli in the year 1700 and was first published in 1713. Let a random experiment be performed repeatedly and let the occurrence of an event in a trial be called a success (S) andits non-occurrence is called failure (F). Let a set of n independent Bernoulli trials in which the probability of success (p) in any trial is constant for each trial and q is the probability of failure in any trial. Let the random variable X be the number of successes in "n" Bernoulli trails. The possible values of X are $0,1,2,\cdots\cdots,n$. This probability model is the most widely used model and it is appropriate in the following experimental situations.

1. The result of each trial can be classified into one of the two mutually exclusive outcomes say success and failure.
2. Probability of success (p) remais constant for each trail.
3. The outcomes of all trials are independent of each other.

Then such experimental situation is called Binomial.

If X denotes the exact number of sucess in "n" trials then X takes the values $0,1,2,\cdots\cdots\cdots,n$. If they are exactly x successes then the remaining $(n-x)$ are failures. Since, the probability of success is p and that of a failure is q. The probability of x successes and consequently $(n-x)$ failures in n independent trials in a specified order is given by the compound probability theorem by the expression $p^x\, q^{n-x}$. But x successes in n trials can occur in $\binom{n}{x}$ ways. Hence the requised probability is $\binom{n}{x}\, p^x\, q^{n-x}$.

The probability distribution of the number of success so obtained is called the Binomial Probability distribution, the reason is obvious that the probabilities of $0,1,2,\cdots\cdots,n$ successes are the successive terms of the binomial expansion of $(q+p)^n$.

Definition: A random variable X is said to follow binomial distribution if it assumes only non - negative integer values and its p.m.f. is given by

$$P(X=x) = P(x) = \begin{cases} \binom{n}{x} p^x q^{n-x} \ ; & x = 0,1,2,\cdots\cdots,n \\ 0 \ \ \text{otherwise} \ ; & q = 1-p \end{cases} \qquad (9.3.1)$$

The two constants n and p are known as the parameters of the distribution. Any random variable which follows binomial distribution is known as binomial variate.

Note: 1. If r.v. X follows binomial distribution with parameters n and p then it is denoted as $X \sim B(n, p)$.

2. Total probability $= \sum_{x=0}^{n} p(x) = \sum_{x=0}^{n} \binom{n}{x} p^x q^{n-x} = (q+p)^n = 1$

9.3.1 Moments of Binomial Distribution: The first four moments about origin of binomial distribution are obtained as follows:

$$\mu_1^1 = E(X) = \sum_{x=0}^{n} x \cdot p(x)$$

$$= \sum_{x=0}^{n} x \cdot \binom{n}{x} p^x q^{n-x}$$

$$= np \sum_{x=1}^{n} \binom{n-1}{x-1} p^{x-1} q^{n-x}$$

$$= np\ (q+p)^{n-1} \qquad \because \binom{n}{x} = \frac{n}{x}\binom{n-1}{x-1}$$

$$= n\ p \qquad = \left(\frac{n}{x}\right)\binom{n-1}{x-1}\binom{n-2}{x-2}\dots\dots\dots$$

$\therefore$ Mean of B.D. is $n\ p$

$$\mu_2^1 = E\left(X^2\right)$$

$$= E\left[X\ (X-1) + X\right]$$

$$= E\left[X(X-1)\right] + \sqrt{2}(X)$$

$$= \sum_{x=0}^{n} x(x-1)\ P(x) + \sum_{x=0}^{n} x\ P(x)$$

$$= \sum_{x=2}^{n} x(x-1)\left(\frac{n}{x}\right)\left(\frac{n-1}{x-1}\right)\binom{n-2}{x-2} p^x q^{n-x} + n\,p$$

$$= n(n-1)\,p^2 \sum_{x=2}^{n} \binom{n-2}{x-2} p^{x-2}\, q^{n-x} + n\,p$$

$$= n(n-1)p^2 \sum_{x=2}^{n} \binom{n-2}{x-2} p^{x-2}\, q^{n-x} + n\,p$$

$$= n(n-1)p^2\,(q+p)^{n-2} + n\,p$$

$$= n(n-1)\,p^2 + n\,p$$

$$\mu_3^1 = E\left(X^3\right)$$

$$= E\left[X(x-1)\,(X-2) + 3X\,(X-1) + X\right]$$

$$= E\left[X\,(X-1)(X-2)\right] + 3E\left[X(X-1)\right] + E(X)$$

$$= \sum_{x=0}^{n} x(x-1)\,(x-2)\binom{n}{x} p^x q^{n-x} + 3\sum_{x=0}^{n} x(x-1)\binom{n}{x} p^x q^{n-x} + n\,p$$

$$= n(n-1)(n-2)p^3 \sum_{x=3}^{n} \binom{n-3}{x-3} p^{x-3} q^{n-x} + 3n(n-1)p^2 \sum_{x=2}^{n} \binom{n-2}{x-2} p^{x-2} q^{n-x} + n\,p$$

$$= n(n-1)\,(n-2)\,p^3\,(q+p)^{n-3} + 3n(n-1)p^2\,(q+p)^{n-2} + n\,p$$

$$\therefore \mu_3^1 = n(n-1)(n-2)p^3 + 3n\,(n-1)p^2 + n\,p$$

Similarly $x^4 = x(x-1)(x-2)(x-3) + 6x(x-1)(x-2) + 7x(x-1) + x$

$$\therefore \mu_4^1 = E\left(X^4\right)$$

$$= E\left[X(X-1)(X-2)(X-3) + 6X(X-1)(X-2) + 7X(X-1) + X\right]$$

$$= E[X(X-1)(X-2)(X-3)] + [X(X-1)(X-2)] + 7E[X(X-1)] + E(X)$$

$$= n(n-1)(n-2)(n-3)p^4 + 6n(n-1)(n-2)p^3 + 7n(n-1)p^2 + n\,p$$

[After simplification]

Central Moments:

The first four central moments are given by

$$\mu_1 = 0$$

$$\mu_2 = \mu_2^1 - \left(\mu_1^1\right)^2$$

$$= n(n-1)p^2 + np - n^2p^2$$

$$= n^2p^2 - np^2 + np - n^2p^2$$

$$= np(1-p)$$

$$\therefore \quad \text{variance} \quad \mu_2 = n\,p\,q$$

$$\mu_3 = \mu_3^1 - 3\mu_2^1\,\mu_1^1 + 2\left(\mu_1^1\right)^3$$

$$= n(n-1)(n-2)p^3 + 3n(n-1)p^2 + np - 3\left\{n(n-1)p^2 + n\,p\right\} n\,p + 2(n\,p)^3$$

$$= n\,p\left[3n\,p(1-p) + 2p^2 - 3p + 1 - 3n\,p\,q\right]$$

$$= n\,p\left[3n\,p\,(1-p) + 2p^2 - 3p + 1 - 3n\,p\,q\right]$$

$$= np\left(2p^2 - 3p + 1\right)$$

$$= n\,p\left(2p^2 - 2p + q\right)$$

$$= n\,p\,(1-p)\,(1-2p)$$

$$= n\,p\,q(1-2p)$$

$$= n\,p\,q\,(q+p-2p)$$

$$= n\,p\,q(q - p)$$

$$\therefore \mu_3 = npq(q-p)$$

$$\mu_4 = \mu_4^1 - 4\,\mu_3^1\,\mu_1^1 + 6\mu_2^1\left(\mu_1^1\right)^2 - 3\left(\mu_1^1\right)^4$$

$$= n(n-1)(n-2)(n-3)p^4 + 6n(n-1)(n-2)p^3 + 7n(n-1)p^2 + np$$

$$-4\left[n(n-1)(n-2)p^3 + 3n(n-1)p^2 + n\,p\right]n\,p + 6\left[n(n-1)p^2 + n\,p\right](n\,p)^2 - 3(n\,p)^4$$

$$= npq\left[1 + 3(n-2)p\,q\right]$$

Coefficient of Skewness $\beta_1 = \dfrac{\mu_3^2}{\mu_2^3} = \dfrac{n^2p^2q^2(q-p)^2}{n^3p^3q^3} = \dfrac{(q-p)^2}{npq}$

Coefficient of Kurtosis $\beta_2 = \dfrac{\mu_4}{\mu_2^2} = \dfrac{npq\left[1+3(n-2)\,p\,q\right]}{n^2p^2q^2} = \dfrac{1+3(n-2)p\,q}{npq}$

$$= 3 + \frac{(1-6p\,q)}{n\,p\,q}$$

Measure of Skewness $\gamma_1 = \sqrt{\beta_1} = \dfrac{q-p}{\sqrt{npq}} = \dfrac{1-2p}{\sqrt{npq}}$

Measure of Kurtosis $\gamma_2 = \beta_2 - 3$

$$= \frac{1-6pq}{npq}$$

9.3.2 Moment Generating Function:

If $X \sim B(n, p)$ then its p.m.f. is given by

$$p(x) = nc_x p^x q^{n-x} \quad ; \quad x = 0,1,2,\cdots\cdots,n\,.$$

The M.G.F. of X is given by

$$M_X(t) = E\left(e^{tX}\right)$$

$$= \sum_{x=0}^{n} e^{tx} p(x)$$

$$= \sum_{x=0}^{n} e^{tx} \cdot \binom{n}{x} p^x q^{n-x}$$

$$= \sum_{x=0}^{n} \binom{n}{x} \left(pe^t\right)^x q^{n-x}$$

$$= \left(q + pe^t\right)^n$$

$$\therefore M_X(t) = \left(q + pe^t\right)^n$$

Calculation of Mean and Variance using M.G.F.

The M.G.F. of Binomial distribution is given by

$$M_X(t) = \left(q + pe^t\right)^n$$

$$\mu_1^1 = \left[\frac{d}{dt} M_X(t)\right]_{t=0}$$

$$= \left[\frac{d}{dt}\left(q + pe^t\right)^n\right]_{t=0}$$

$$= n\left(q + p\right)^{n-1} p$$

$$= n\,p \qquad \because p + q = 1$$

$$\therefore \text{ Mean } \mu_1^1 = n\,p$$

$$\mu_2^1 = \left[\frac{d^2 M_X(t)}{dt^2}\right]_{t=0}$$

$$= \left[\frac{d}{dt}\, n\left(q + pe^t\right)^{n-1} pe^t\right]_{t=0}$$

$$= n\,p\left[(n-1)\left(q+pe^t\right)^{n-2} p.e^t.e^t + \left(q+pe^t\right)^{n-1} e^t\right]_{t=0}$$

$$= n\,p\left[(n-1)\,p+1\right]$$

$$= n\,(n-1)\,p^2 + n\,p$$

$$\therefore \text{Variance} \quad \mu_2 = \mu_2^1 - \left(\mu_1^1\right)^2$$

$$= n(n-1)p^2 + n\,p - n^2p^2$$

$$= n^2p^2 - np^2 + np - n^2p^2$$

$$= n\,p\,(1-p)$$

$$= n\,p\,q$$

$\therefore$ Mean of B.D. is $n\,p$

Variance of B.D. is $n\,p\,q$

9.3.3 Cumulant Generating Function:

We know that $M_X(t) = \left(q+pe^t\right)^n$

By defination C.G.F. of r.v. X is given by

$$K_X(t) = \ell og\, M_X(t)$$

$$= \ell og\left[\left(q+pe^t\right)^n\right]$$

$$= n \log\left(q+pe^t\right)$$

$$= n\,\ell og\left[q+p\left(1+t+\frac{t^2}{2!}+\frac{t^3}{3!}+\cdots\cdots\cdots\right)\right]$$

$$= n\,\ell og\left[1+p\left(t+\frac{t^2}{2!}+\frac{t^3}{3!}+\cdots\cdots\right)\right]$$

$$= n\left[p\left(t+\frac{t^2}{2!}+\frac{t^3}{3!}+\frac{t^4}{4!}+\cdots\cdots\right)-\frac{p^2}{2}\left(t+\frac{t^2}{2!}+\frac{t^3}{3!}+\cdots\cdots\right)^2\right.$$

$$\left.+\frac{p^3}{3}\left(t+\frac{t^2}{2!}+\frac{t^3}{3!}+\cdots\cdots\right)^3 \frac{-p^4}{4}\left(t+\frac{t^2}{2!}+\frac{t^3}{3!}+\cdots\right)^4+\cdots\cdots\right]$$

$\therefore$ Mean $\mu_1^1 = k_1 =$ The coefficient of $\frac{t}{1!}$ in $K_X(t) = np$

Variance $\mu_2 = k_2 =$ The coefficient of $\frac{t^2}{2!}$ in $K_X(t) = n\left(p-p^2\right)$

$$= n\,p\,(1-p)$$

$$= npq$$

$K_3 =$ The coefficient of $\frac{t^3}{3!}$ in $K_X(t) = n\left[p-3p^2+2p^3\right]$

$$= np\left(1-3p+2p^2\right)$$

$$= n\,p\,(1-p)(1-2p)$$

$$= n\,p\,q\,(1-2p)$$

$$\therefore \mu_3 = K_3 = n\,pq\,(1-2p) = n\,p\,q(q-p)$$

$K_4 =$ The coefficient of $\frac{t^4}{4!}$ in $K_X(t) = n\left[p-7p^2+12p^3-6p^4\right]$

$$= np\left[1-7p+12p^2-6p^3\right]$$

$$= np(1-p)\left(1-6p+6p^2\right)$$

$$= npq\left[1-6p(1-p)\right] = npq(1-6pq)$$

$$\therefore \mu_4 = K_4 + 3K_2^2$$

$$= npq(1-6pq) + 3n^2p^2q^2$$

$$= npq(1-6pq+3npq)$$

$$= npq\left[1+3pq(n-2)\right]$$

9.3.4 Characteristic Function:

By definition $\quad \phi_X(t) = E\left[e^{itX}\right]$

$$= \sum_{x=0}^{n} e^{itX} p(x)$$

$$= \sum_{x=0}^{n} e^{itX} \binom{n}{x} p^x \, q^{n-x}$$

$$= \sum_{x=0}^{n} \binom{n}{x} \left(pe^{it}\right)^x \cdot q^{n-x}$$

$$= \left(q + pe^{it}\right)^n$$

$\therefore$ C.F. of B.D. is given by

$$\phi_X(t) = \left(q + pe^{it}\right)^n$$

9.3.5 Probability Generating Function:

By definition P.G.F. of r.v. X is given by

$$p_X(S) = E\left[S^X\right]$$

$$= \sum_{x=0}^{n} S^X p(x)$$

$$= \sum_{x=0}^{n} S^x \binom{n}{x} p^x \, q^{n-x}$$

$$= \sum_{x=0}^{n} \binom{n}{x} (PS)^x q^{n-x}$$

$$= (q + PS)^n$$

$\therefore$ P.G.F. of B.D. is $P(S) = (q + PS)^n$

9.3.6 Recurrence Relation for The Moments of Binomial Distribution:

By defination the r^{th} moment about mean is given by

$$\mu_r = E[X - E(X)]^r$$

$$= \sum_{x=0}^{n} (x - nP)^r P(x)$$

$$= \sum_{x=0}^{n} (x - nP)^r \binom{n}{x} P^x q^{n-x} \qquad q = 1 - P$$

Differentiating w.r.t. P we get

$$\frac{d\mu_r}{dP} = \sum_{x=0}^{n} \binom{n}{x} \left[\begin{array}{l} r(x - nP)^{r-1}(-n)P^x(1-P)^{n-x} + (x - nP)^r \\ \left\{P^x(n-x)(1-P)^{n-x-1}(-1) + (1-p)^{n-x} xP^{x-1}\right\} \end{array} \right]$$

$$= \sum_{x=0}^{n} (-nr)(x - nP)^{r-1} \binom{n}{x} P^x (1-P)^{n-x} - \sum_{x=0}^{n} (x - nP)^r \binom{n}{x} P^x (1-P)^{n-x-1}(n-x)$$

$$+ \sum_{x=0}^{n} \frac{x}{P} \binom{n}{x} P^x (1-P)^{n-x} (x - nP)^r$$

$$= nr \sum_{x=0}^{n} (x - nP)^{r-1} P(x) - \sum_{x=0}^{n} (x - nP)^r \binom{n}{x} P^x \frac{(1-P)^{n-x}}{(1-P)} \cdot (n-x)$$

$$+ \sum_{x=0}^{n} \left(\frac{x}{P}\right) (x - nP)^r \binom{n}{x} P^x (1-P)^{n-x}$$

$$= -nr\sum_{x=0}^{n}(x-nP)^{r-1}P(x) - \sum_{x=0}^{n}(x-nP)^{r}\cdot\frac{(n-x)}{q}P(x) + \sum_{x=0}^{n}\frac{x}{P}(x-nP)^{r}P(x)$$

$$= -nr\sum_{x=0}^{n}(x-nP)^{r-1}P(x) - \sum_{x=0}^{n}(x-nP)^{r}P(x)\left[\frac{n-x}{q}-\frac{x}{P}\right]$$

$$= -nr\,\mu_{r-1} - \sum_{x=0}^{n}(x-nP)^{r}P(x)\left[\frac{nP-xP-xq}{Pq}\right]$$

$$\therefore \frac{d\mu_r}{dP} = -nr\,\mu_{r-1} - \sum_{x=0}^{n}(x-nP)^{r}P(x)\left[\frac{nP-x(p+q)}{Pq}\right]$$

$$= -nr\,\mu_{r-1} - \sum_{x=0}^{n}(x-nP)^{r}P(x)\left[\frac{nP-x}{Pq}\right]$$

$$= -nr\,\mu_{r-1} + \sum_{x=0}^{n}(x-nP)^{r}P(x)\cdot\frac{(x-nP)}{Pq}$$

$$= -nr\,\mu_{r-1} + \frac{1}{Pq}\sum_{x=0}^{n}(x-nP)^{r+1}P(x)$$

$$= -nr\,\mu_{r-1} + \frac{1}{Pq}\mu_{r+1}$$

$$\therefore \frac{1}{Pq}\mu_{r+1} = \frac{d\mu_r}{dP} + nr\mu_{r-1}$$

$$\Rightarrow \mu_{r+1} = Pq\left[\frac{d\mu_r}{dP} + nr\mu_{r-1}\right];\quad r = 1, 2, \cdots\cdots\cdots$$

we know that $\mu_0 = 1, \mu_1 = 0$

If $r = 1$ then $\mu_2 = Pq\left[\frac{d\mu_1}{dp} + n(1)\,\mu_0\right]$

$$= Pq\left[0+n\right]$$

$$= nPq$$

$$\therefore \mu_2 = nPq$$

If $r = 2$ then $\mu_3 = Pq\left[\frac{d\mu_2}{dP} + 2n\mu_1\right]$

$$= Pq\left[\frac{d}{dP}(nPq) + 2n(0)\right]$$

$$= nPq\left[P(-1)+q(1)\right] \qquad \because q = 1-P$$

$$= nPq(q-P)$$

$$\therefore \mu_3 = nPq \ (q-P)$$

If $r = 3$, then $\mu_4 = Pq\left[\frac{d\mu_3}{dP} + 3(n)\mu_2\right]$

$$= Pq\left[\frac{d}{dP}\{nPq(q-P)\} + 3n(nPq)\right]$$

$$= nPq\left[\frac{d}{dP}[P(1-P)(1-2P)] + 3nPq\right]$$

$$= nPq\left[\frac{d}{dP}\left(P-P^2\right)(1-2P) + 3nPq\right]$$

$$= nPq\left[1-6Pq+3nPq\right]$$

$$= nPq\left[1-6Pq+3nPq\right]$$

$$= nPq\left[1-6Pq+3nPq\right]$$

$$= nPq\left[1+3Pq(n-2)\right]$$

9.3.7 Recurrence Relation For The Probabilities of Binomial Distribution:

By definition P.M.F. of Binomial Distribution is $P(x) = \binom{n}{x} P^x \cdot q^{n-x} \quad ; \; x = 0, 1, 2, \cdots\cdots, n$

We have $\dfrac{P(x+1)}{P(x)} = \dfrac{\binom{n}{x+1} P^{x+1} q^{n-x-1}}{\binom{n}{x} p^x q^{n-x}}$

$$= \frac{n!}{(x+1)!\,(n-x-1)!} \cdot \frac{x!(n-x)!}{n!} \cdot \frac{P}{q}$$

$$= \left(\frac{n-x}{x+1}\right) \frac{P}{q}$$

$\therefore$ The required recurrence formula for the probabilities of Binomial Distribution is

$$P(x+1) = \left(\frac{n-x}{x+1}\right) \frac{P}{q} P(x) \quad ; \; x = 0, 1, 2, \cdots\cdots, n-1$$

This formula is very convenient to obtain probabilities of Binomial Distribution for the given data. The only probability we need to calculate is $P(0)$ which is given by $P(0) = q^n$.

Where q is estimated from the given data by equating the mean $\bar{x}$ of the distribution to nP, the mean of B.D.

Thus $\hat{P} = \bar{x}/n$

$\hat{q} = 1 - \hat{P}$

The remaining probabilities $P(1), P(2), \cdots\cdots\cdots\cdots, P(n)$ can be obtained using recurrence formula on substitution of $x = 0, 1, 2, \cdots\cdots, n-1$ respectively.

9.3.8 Additive Property (or) Reproductive Property:

Statement: If X_1 and X_2 are two independent random variables with parameters (n_1, P) and (n_2, P) respectively then $X_1 + X_2$ also follows Binomial Distribution with parameters $(n_1 + n_2, P)$.

Proof: Given that $X_1 \sim B(n_1, P)$ then its M.G.F. is given by $M_{X_1}(t) = (q + Pe^t)^{n_1}$

Also $X_2 \sim B(n_2, P)$ then its M.G.F. is given by $M_{X_2}(t) = (q + Pe^t)^{n_2}$

If X_1 and X_2 are independent then from the properties of M.G.F. we have

$$M_{X_1+X_2}(t) = M_{X_1}(t) \cdot M_{X_2}(t)$$

$$= (q + Pe^t)^{n_1} (q + Pe^t)^{n_2}$$

$$= (q + Pe^t)^{n_1+n_2}$$

Which is the M.G.F. of a Binomial Variate with parameters $(n_1 + n_2, P)$. Hence by uniqueness theorem of M.G.F.'s $X_1 + X_2 \sim B(n_1 + n_2, P)$.

Remark:

If X_1 and X_2 are two independent random variables with parameters (n_1, P_1) and (n_2, P_2) respectively then $X_1 + X_2$ does not follow Binomial Distribution.

Since $X_1 \sim B(n_1, P_1)$ and $X_2 \sim B(n_2, P_2)$ then M.L.F.'s are $M_{X_1}(t) = (q_1 + P_1e^t)^{n_1}$ and $M_{X_2}(t) = (q_2 + P_2e^t)^{n_2}$.

Using the property of M.G.F. we have

$$M_{X_1+X_2}(t) = M_{X_1}(t) \cdot M_{X_2}(t)$$

$$= (q_1 + P_1e^t)^{n_1} \cdot (q_2 + P_2e^t)^{n_2}$$

Which cannot be expressed in the form $(q + Pe^t)^n$

Hence $X_1 + X_2$ is not a Binomial Variate.

Hence in general the sum of two independent binomial variates is not a binomial variate.

9.3.9 Mode of The Binomial Distribution:

We have $\frac{P(x)}{P(x-1)} = \frac{\binom{n}{x} P \quad q^{n-x}}{\binom{n}{x-1} P^{x-1} q^{n-x+1}}$

$$= \frac{n!}{x!(n-x)!} \cdot \frac{(x-1)!(n-x+1)!}{n!} \cdot \frac{P}{q}$$

$$= \frac{(n-x+1)}{x} \cdot \frac{P}{q}$$

$$= \frac{xq + (n-x+1)P - xq}{xq}$$

$$= 1 + \frac{(n+1)P - x(P+q)}{xq}$$

$$= 1 + \frac{(n=1)\ P - x}{x\ q} \qquad (9.3.9.1)$$

Mode is the value of x for which $P(x)$ is maximum.

Case (i): If $(n+1)P$ is not an integer.

Let $(n+1)P = m + f$, where m is an integral part and f is fraction part of $(n+1)P$ where $0 < f < 1$.

Substituting in (9.3.9.1) we get

$$\frac{P(x)}{P(x-1)} = 1 + \frac{(m+f)-x}{xq} \quad ; \quad x = 1, 2, \cdots\cdots\cdots, n \qquad (9.3.9.2)$$

from (9.3.9.2) it is obvious that

$$\frac{P(x)}{P(x-1)} > 1 \text{ for } x = 1, 2, \cdots\cdots, m \text{ and } \frac{P(x)}{P(x-1)} < 1 \text{ for } x = n+1, n+2, \cdots\cdots, n$$

$$\Rightarrow \frac{P(1)}{P(0)} > 1, \frac{P(2)}{P(1)} > 1, \cdots\cdots\cdots, \frac{P(m)}{P(m-1)} > 1, \frac{P(m+1)}{P(m)} < 1, \frac{P(m+2)}{P(m+1)} < 1, \cdots, \frac{P(n)}{P(n-1)} < 1$$

$$\Rightarrow P(0) < P(1) < P(2) < \cdot\cdot\cdot < P(m-1) < P(m) > P(m+1) > P(m+2) > \cdot\cdot\cdot > P(n)$$

$\Rightarrow P(x)$ is maximum at $x = m$.

$\therefore$ Mode = m = integral part of $(n+1)P$

Case (ii): If $(n+1)P$ is an integer

Let $(n+1)P = m$

Substituting in (9.3.9.1) we get

$$\frac{P(x)}{P(x-1)} > 1 \text{ for } x = 1, 2, \cdot\cdot\cdot (m-1)$$

$$\frac{P(x)}{P(x-1)} = 1 \text{ for } x = m$$

$$\frac{P(x)}{P(x-1)} < 1 \text{ for } x = m+1, m+2, \cdot\cdot\cdot, n$$

Now proceeding as in case (i), we have

$$P(0) < P(1) < P(2) < \cdot\cdot\cdot < P(m-1) = P(m) > P(m+1) > \cdot\cdot\cdot > P(n)$$

Thus maximum probability is at $P(m-1)$ & $P(m)$

Therefore m and $m-1$ are two modes.

9.3.10 Worked Examples:

Example 1: The mean and variance of Binomial Variate X with parameters n and P are 16 and 8. Find (i) $P(X=0)$, (ii) $P(X=1)$, (iii) $P(X \geq 2)$

Solution: If $X \sim B(n, P)$ then its mean $= nP$ and Variance $= nPq$

Given that $nP = 16$

$$nPq = 8$$

$$\therefore \frac{nPq}{nP} = \frac{8}{16}$$

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

$$\therefore P = 1 - q$$

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

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

Also $nP = 16$

$$\Rightarrow n = \frac{16}{\left(1/2\right)} = 32$$

$\therefore$ The parameters of B.D. are $n = 32,\ P = \frac{1}{2}$

It P.M.F. is given by $P(X = x) = 32\,C_x \left(\frac{1}{2}\right)^x \left(\frac{1}{2}\right)^{32-x}$; $x = 0, 1, \cdots\cdots, 32$

(i) $P(X=0) = 32\,C_0 \left(\frac{1}{2}\right)^0 \left(\frac{1}{2}\right)^{32-0} = \left(\frac{1}{2}\right)^{32} = \frac{1}{2^{32}}$

(ii) $P(X=1) = 32C, \left(\frac{1}{2}\right)^1 \left(\frac{1}{2}\right)^{32-1} = 32\left(\frac{1}{2}\right)^{32} = \frac{32}{2^{32}}$

(iii) $P(X \geq 2) = 1 - P(X < 2)$

$$= 1 - \left[P(X=0) + P(X=1)\right]$$

$$= 1 - \left[\left(\frac{1}{2}\right)^{32} + 32\left(\frac{1}{2}\right)^{32}\right]$$

$$= 1 - \frac{33}{2^{32}}$$

Example 2: In a certain town 50% of the population is literates and assume that 100 investigators take a sample of 10 individuals each to see would you expect to report that three people or less are literates in the sample ?

Solution: The probability of literates $= p = 50\% = \frac{50}{100} = \frac{1}{2}$

$$\therefore q = \frac{1}{2}$$

Given that $n = 10,\ N = 100$.

Let X denotes no of literates

$\therefore$ Probability that three people or less are literates out of 10 samples $= P(X \le 3)$

$$= P(X=0) + P(X=1) + P(X=2) + P(X=3)$$

$$= {}^{10}C_0\left(\frac{1}{2}\right)^0\left(\frac{1}{2}\right)^{10-0} + {}^{10}C_1\left(\frac{1}{2}\right)^1\left(\tfrac{1}{2}\right)^{10-1} + {}^{10}C_2\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^{10-2} + {}^{10}C_3\left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right)^{10-3}$$

$$= \frac{1}{2^{10}}\left[1 + {}^{10}C_1 + {}^{10}C_2 + {}^{10}C_3\right]$$

$$= \frac{1}{2^{10}}\left[1 + 10 + 45 + 120\right]$$

$$= \frac{176}{1024}$$

$\therefore$ No. of investigators expect to report that there or less people are literates

$$= 100\ P(X \le 3)$$

$$= 100 \times \frac{176}{1024}$$

$$= 17 \cdot 8$$

$\therefore$ 17 are the investigators.

Example 3: In 256 sets of twelve tosses of a fair coin in how many cases may one expect eight heads and four tails.

Solution: $P =$ Probability of getting head $= \frac{1}{2}$

$$\therefore \; q = \frac{1}{2}$$

given that $n = 12, \; N = 256$.

Let X denotes no of heads.

$$\therefore P(x) = {}^{n}C_{x} \; P^{x} \; q^{n-x} \;\; ; x = 0, 1, 2, \cdot \cdot \cdot, n$$

The probability of getting eight heads $= P(X = 8)$

$$= {}^{12}C_{8} \left(\frac{1}{2}\right)^{8} \left(\frac{1}{2}\right)^{12-8}$$

$$= \frac{495}{2^{12}}$$

$$= \frac{495}{4096}$$

Out of 256 sets, the no of cases of getting 8 heads

$$= 256 \times \frac{495}{4096}$$

$$= 31$$

$\therefore$ Expected no. of case is 31.

Example 4: A perfect cubic die is throw a large number of times in sets of 8. The occurence of a 5 or 6 is called a success. In what proportion of the sets would you expect 3 successes?

Solution: The probability of getting a 5 or 6 with a die $= P = \frac{2}{6} = \frac{1}{3}$

$$\therefore q = 1 - \frac{1}{3} = \frac{2}{3}$$

given that $n = 8$

If X denotes no. of success then $X \sim B\left(8, \frac{1}{3}\right)$

$$\therefore P(x) = {}^{8}C_{x} \left(\frac{1}{3}\right)^{x} \left(\frac{2}{3}\right)^{8-3} \quad ; \; x = 0,1,2, \cdots ,8$$

The probability of 3 successes in one set of $8 = P(X = 3)$

$$= {}^{8}C_{3} \left(\frac{1}{3}\right)^{3} \left(\frac{2}{3}\right)^{5}$$

$$= 56 \times \frac{1}{27} \times \frac{32}{243}$$

$$= \frac{1792}{6561}$$

$$= 0 \cdot 2731$$

$\therefore$ The proportion of the sets giving 3 successes $= 100 \times 0 \cdot 2731$

$= 27 \cdot 31$

≈ 27

Example 5: Obtain the M.G.F. of Binomial distribution with $n = 7$ and $P = 0 \cdot 6$. Find the first three central moments.

Solution: If $X \sim B(n, P)$ then its M.G.F. is given by $M_X(t) = (q + Pe^t)^n$

given that $n = 7, P = 0 \cdot 6, q = 1 - P = 0 \cdot 4$

$\therefore$ M.G.F. of given B.D. is $M_X(t) = (0 \cdot 4 + 0 \cdot 6 e^t)^7$

Mean $= n P = 7(0 \cdot 6) = 4 \cdot 2$

Variance $= \mu_2 = nPq = 7(0 \cdot 6)(0 \cdot 4) = 1 \cdot 68$

$$\mu_3 = nPq(q - P)$$

$$= 7(0 \cdot 6)(0 \cdot 4)(0 \cdot 4 - 0 \cdot 6)$$

$$= -0 \cdot 336$$

Example 6: With the usual notation, find P for a Binomial r.v. X if $n = 6$ and if $9P(X=4) = P(X=2)$.

Solution: Given that $n = 6$

and $9P(X = 4) = P(X = 2)$

$$\Rightarrow 9 \times 6C_4 P^4 q^{6-4} = 6C_2 P^2 q^{6-2}$$

$$\Rightarrow 9 \times 15 P^4 q^2 = 15 P^2 q^4$$

$$\Rightarrow 9 = \frac{P^2 q^4}{P^4 q^2} = \frac{q^2}{P^2}$$

$$\Rightarrow \frac{q}{P} = 3$$

$$\Rightarrow q = 3P$$

$$\Rightarrow 1 - P = 3P$$

$$\Rightarrow 4P = 1$$

$$\Rightarrow P = \frac{1}{4}$$

$\therefore\ q = 1 - P$

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

$= \frac{3}{4}$

$\therefore$ parameters of B.D. are $n = 6$ and $P = \frac{1}{4}$

Example 7: If X is a r.v. following B.D. with mean 2.4 and variance 1.44. Find $P(X \geq 5)$ and $(1 < X \leq 4)$.

Solution: Given that $n\,P < 2 \cdot 4$

$n\,Pq = 1.44$

$\therefore\ q = \frac{n\,P\,q}{n\,P} = \frac{1 \cdot 44}{2 \cdot 4} = \frac{3}{5}$

$\therefore\ P = 1 - q = 1 - \frac{3}{5} = \frac{2}{5}$

$\therefore\ n = \frac{2 \cdot 4}{P} = \frac{2 \cdot 4}{\left(2/5\right)} = 6$

The parameters of B.D. are $n = 6,\ P = 2/5$.

Its P.M.F. is given by $P(x) = 6C_x \left(\frac{2}{5}\right)^x \left(\frac{3}{5}\right)^{6-x}$; $x = 0, 1, \cdots\cdots\cdots, 6$

Now $P(X \geq 5) = P(X = 5) + P(X = 6)$

$$= 6C_5 \left(\frac{2}{5}\right)^5 \left(\frac{3}{5}\right)^1 + 6C_6 \left(\frac{2}{5}\right)^6$$

$$= \frac{6 \times 3 \times 2^5}{5^6} + \frac{2^6}{5^6}$$

$$= \frac{2^5 [18 + 2]}{5^6} = \frac{128}{3125}$$

$$P(1 < X \le 4) = P(X = 2) + P(X = 3) + P(X = 4)$$

$$= 6C_2 \left(\frac{2}{5}\right)^2 \left(\frac{3}{5}\right)^4 + 6C_3 \left(\frac{2}{5}\right)^3 \left(\frac{3}{5}\right)^3 + 6C_4 \left(\frac{2}{5}\right)^4 \left(\frac{3}{5}\right)^2$$

$$= \frac{1}{5^6} \left[15 \times 4 \times 3^4 + 20 \times 2^3 \times 3^3 + 15 \times 2^4 \times 3^2 \right]$$

$$=$$

Example 8: For a binomial distribution, mean = 20, S.D. = 4. Calculate mode value.

Solution: Given that mean $nP = 20$

$$\text{S.D. } \sqrt{nPq} = 4$$

$$\Rightarrow n\,P\,q = 4^2 = 16$$

$$\therefore q = \frac{nPq}{nP} = \frac{16}{20} = \frac{4}{5} = 0 \cdot 8$$

$$P = 1 - q = 1 - 0 \cdot 8 = 0 \cdot 2$$

$$\therefore n = \frac{20}{P} = \frac{20}{0 \cdot 2} = 100$$

$\therefore$ Mode = Integral of $(n+1)\,P = (100+1)\,0 \cdot 2 = 20 \cdot 2$

$\therefore$ Mode = 20

9.4 Exercise:

1. Obtain the moment generating function of B.D. and hence find its mean and variance.
2. Show that relation between moments about origin is $\mu^1_{r+1} = P\left[n\,\mu^1_r + q\,\frac{d\,\mu^1_r}{d\,P}\right]$

3. Obtain the probability generating function of a Binomial Distribution. Hence or otherwise obtain the mean and variance of the distribution.
4. State and Prove the reproductive property of B.D.
5. Define Binomial Distribution and derive its mean and variance.
6. Obtain the characteristic function of B.D. hence or otherwise find its mean and variance.
7. Derive the mode of the Binomial Distribution.
8. Derive the cumulant generating function of B.D. and hence find the first four central moments.
9. Determine the Binomial Distribution for which the mean is 4 and variance 3 and find its mode.
10. In a shooting competition, the probability of a man hitting a target is $\frac{1}{5}$. If he shoots 5 times, what is the probability of hitting the target at least twice.
11. In a Binomial Distribution consisting of 5 independent trials, probabilities of 1 and 2 successes are 0.4096 and 0.2048 respectively. Find the parameter P of the distribution.
12. Ten coins are thrown simultaneously. Find the probability of getting at least seven heads.
13. A and B play a game in which their chances of winning are in the ratio 3 : 2. Find A's chance of winning at least three games out of the five games played.
14. Define discrete uniform distribution and hence find its mean and variance.
15. Define Bernoulli Distribution. Find its M.G.F. and hence find its mean and variance.

9.5 Answers:

9. $n = 16,\ P = \frac{1}{4},\quad \text{Mode} = 4$
10. $P\left(X \geq 2\right) = 0 \cdot 2634$
11. $P = 0 \cdot 2$
12. $P\left(X \geq 7\right) = \frac{176}{1024}$
13. $P\left(X \geq 3\right) = 0 \cdot 68$

Lesson – 10

POISSON DISTRIBUTION

Objective:

After studying this lesson the student is expected to have clear comprehencsion of thetheory and practical utility about the concepts of poisson distribution and its properties.

Structure of The Lesson:

This lesson consists of sections as detailed below:

10.1 Introduction:

Poisson Distribution was discovered by the French mathematician and Physicist Simeon Denis Poisson in 1837. Some times we come across a rare event which occurs once in number of trials. For example, consider the event of a recieving telephone calls at a particular telephone exchange in some specified time. If we consider a trial as a number of calls on particular time and the outcome of the trial as to recieve a call or not to recieve a call, then clearly, "n" represents the number of calls during a particular time period is very large and it is difficult to find it exactly. Also, the probability "p" of recieving a call is very small. However, the mean number of calls in the time period is $np = \lambda$ (say) is finite constant. In these situations if X denotes umber of calls then the probability function of randim variable in the given time period can be given by $p(X=x) = e^{-\lambda} \cdot \lambda^x / x!$.

Where e is a constant with approximate value 2.7183. Thisdistribution of X is called Poisson Distribution and the variable X is called a poisson variate.

10.2 Defination:

A random variable X is said to follow a Poisson Distribution if it assumes only non - negative integer values and its probability mass function (p.m.f.) is given by

$$p(X=x) = \begin{cases} \dfrac{e^{-\lambda}\lambda^x}{x!} \ ; & x = 0,1,2,\cdots\cdots \\ & \lambda > 0 \\ 0 \ ; & \text{otherwise} \end{cases}$$

Here λ is called as the paramter of the distribution.

Remarks: 1. If X is a Poisson Variate with parameter λ then it is denoted as $X \sim p(\lambda)$

2. The total probability $= \sum_{x=0}^{\infty} (x) = \sum_{x=0}^{\infty} \dfrac{e^{\lambda} \cdot \lambda^x}{x!}$

$$= e^{-\lambda} \cdot \sum_{x=0}^{\infty} \frac{\lambda^x}{x!}$$

$$= e^{-\lambda} \cdot e^{\lambda}$$

$$= 1$$

3. Distribution function of Poisson Distribution is given by

$$F(x) = p(X \le x) = \sum_{r=0}^{x} p(r)$$

$$= \sum_{r=0}^{x} \frac{e^{-\lambda} \cdot \lambda^r}{r!}$$

$$= \sum_{r=0}^{x} \frac{e^{-\lambda} \cdot \lambda^r}{r!} \quad ; \quad x = 0,1,2,\cdots\cdots\cdots$$

10.3 Uses or Live Examples of Poisson Distribution:

The poisson distribution may be useful in the following some instances.

1. Number of sucides reported in a particular city.
2. Number of deaths from a disease such as heart attack or due to snake bite.
3. Number of faulty blads in as acket of large number of blades.
4. Number of air accidents in some unit of time.
5. Number of printing mistakes at each page of the book.
6. Number of telephone calls received at a particular telephone exchage in some unit of time.
7. The number of defective material in a packing manufactured by a company.

10.4 Poisson Distribution is a Limiting Case of Binomial Distribution:

The poisson distribution is a limiting case of Binomial Distribution under the following conditions.

(i) The number of trials "n" is large, i.e. $n \to \infty$

(ii) The constant probability of success "p" for each trial is very small, i.e. $p \to 0$

(iii) $n\,p$ is finite, say $\lambda = n\,p$

$$\therefore\ p = \frac{\lambda}{n} \quad \text{and} \quad q = 1 - \frac{\lambda}{n}$$

By definition the p.m.f. of Binomial Distribution is

$$p(x) = \binom{n}{x} p^x\, q^{n-x} \qquad x = 0, 1, 2, \cdots\cdots, n$$

$$= \frac{n(n-1)(n-2)\cdots\cdots[n-x+1]}{x\,!}\left(\frac{\lambda}{n}\right)^x \left(1-\frac{\lambda}{n}\right)^{n-x}$$

$$= \frac{\lambda^x}{x\,!}\ \frac{n(n-1)(n-2)\cdots\cdots\cdots[n-(x-1)]}{n^x} \cdot \frac{\left(1-\frac{\lambda}{n}\right)^{n-x}}{\left(1-\frac{\lambda}{n}\right)^x}$$

$$= \frac{\lambda^x}{x\,!}\left(\frac{n}{n}\right)\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\cdots\left[1-\frac{(x-1)}{n}\right]\frac{\left(1-\frac{\lambda}{n}\right)^n}{\left(1-\frac{\lambda}{n}\right)^x}$$

$$\therefore \lim_{n\to\infty} p(x) = \frac{\lambda^x}{x!} \lim_{x\to\infty}\left\{\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\cdots\cdots\left[1-\frac{(x-1)}{n}\right]\right\} \frac{\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^n}{\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^x}$$

$$= \frac{\lambda^x}{x!}\,\frac{e^{-\lambda}}{1}$$

$$= \frac{e^{-\lambda}\cdot\lambda^x}{x!}\ ;\ x = 0,1,2,\cdots\cdots\cdots \qquad \because \lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^n = e^{-\lambda}$$

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

Hence poisson distribution is a limiting case of B.D.

10.5 Moments of Poisson Distribution:

The p.m.f. of Poisson Distribution with parameter λ is given by

$$p(x) = \frac{e^{-\lambda}\cdot\lambda^x}{x!} \quad ; \quad x = 0,1,2,\cdots\cdots\cdots,\infty$$

Moments about origin:

$$\therefore \mu_1^1 = E(X) = \sum_{x=0}^{\infty} x\cdot p(x)$$

$$= \sum_{x=0}^{\infty} x\cdot\frac{e^{-\lambda}\cdot\lambda^x}{x!}$$

$$= \lambda\, e^{-\lambda} \sum_{x=1}^{\infty}\frac{\lambda^{x-1}}{(x-1)!}$$

$$= \lambda\, e^{-\lambda}\cdot e^{\lambda}$$

$$= \lambda$$

$\therefore$ Mean of The Poisson distribution is λ.

$$\mu_2^1 = E\left(X^2\right)$$

$$= E[X(X-1)(X-2)+3X(X-1)+X]$$

$$= \sum_{x=0}^{\infty} x(x-1)(x-2)\frac{e^{-\lambda}\lambda^x}{x!} + 3\sum_{x=0}^{\infty} x(x-1)\frac{e^{-\lambda}\cdot\lambda^x}{x!} + \lambda$$

$$= e^{-\lambda}\cdot\lambda^3 \sum_{x=3}^{\infty} x(x-1)(x-2)\cdot\frac{\lambda^{x-3}}{x(x-1)(x-2)(x-3)!} + 3e^{-\lambda}\cdot\lambda^2 \sum_{x=2}^{\infty} x(x-1)\frac{\lambda^{x-2}}{x(x-1)(x-2)!} + \lambda$$

$$= e^{-\lambda}\lambda^3 \sum_{x=3}^{\infty} \frac{\lambda^{x-3}}{(x-3)!} + 3\lambda^2 e^{-\lambda}\sum \frac{\lambda^{x-2}}{(x-2)!} + \lambda$$

$$= e^{-\lambda}\,\lambda^3\, e^{\lambda} + 3\lambda^2\, e^{-\lambda}\, e^{\lambda} + \lambda$$

$$= \lambda^3 + 3\lambda^2 + \lambda$$

$$\therefore\ \mu_3^1 = \lambda^3 + 3\lambda^2 + \lambda$$

$$\mu_4^1 = E\left(X^4\right)$$

$$= E[X(X-1)(X-2)(X-3)+6X(X-1)(X-2)+7X(X-1)+X]$$

$$\mu_4^1 = \lambda^4 + 6\lambda^3 + 7\lambda^2 + \lambda$$

Central Moments:

$$\mu_2 = \mu_2^1 - \left(\mu_1^1\right)^2$$

$$= \lambda^2 + \lambda - \lambda^2$$

$$= \lambda$$

$\therefore$ Variance of P.D. is λ

Thus Poisson Distribution is a discrete distribution in which mean and variance are equal.

$$\mu_3 = \mu_3^1 - 3\mu_2^1\mu_1^1 + 2\left(\mu_1^1\right)^3$$

$$= \lambda^3 + 3\lambda^2 + \lambda - 3\left(\lambda^2 - \lambda\right)\lambda + 2\lambda^3$$

$$= \lambda$$

$$\mu_4 = \mu_4^1 - 4\mu_3^1\,\mu_1^1 + 6\mu_2^1\left(\mu_1^1\right)^2 - 3\left(\mu_1^1\right)^4$$

$$= \lambda^4 + 6\lambda^3 + 7\lambda^2 + \lambda - 4\left(\lambda^3 + 3\lambda^2 + \lambda\right)\lambda + 6\left(\lambda^2 + \lambda\right)\lambda^2 - 3\lambda^4$$

$$= 3\lambda^2 + \lambda$$

Coefficient of skewness $\beta_1 = \dfrac{\mu_3^2}{\mu_2^3} = \dfrac{\lambda^2}{\lambda^3} = \dfrac{1}{\lambda}$

Coefficinet of kurtosis $\beta_2 = \dfrac{\mu_4}{\mu_2^2} = \dfrac{3\lambda^2 + \lambda}{\lambda^2} = 3 + \dfrac{1}{\lambda}$

Measure of skewness $V_1 = \sqrt{\beta_1} = \dfrac{1}{\sqrt{\lambda}}$

Measure of kurtosis $V_2 = \beta_2 - 3$

$$= 3 + \frac{1}{\lambda} - 3$$

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

Note: As $\lambda \to \infty$, $\beta_1 \to 0$ and $\beta_2 \to 3$. Hence for large λ poisson distribution tends to normal distribution.

10.6 Recurrence Relation for Moments of The Poisson Distribution:

The p.m.f. of Poisson Distribution with parameter λ is given by

$$p(x) = \frac{e^{-\lambda}\,\lambda^{x}}{x\,!} \quad ; x = 0,1,2,\cdots\cdots$$

$$E(X) = T$$

By definition $\mu_r = E[X - E(X)]^r$

$$= \sum_{x=0}^{\infty} (x-\lambda)^r \; p(x)$$

$$= \sum_{x=0}^{\infty} (x-\lambda)^r \cdot \frac{e^{-\lambda}\lambda^{x}}{x\,!}$$

Defferentiate w.r.t. λ, we get

$$\frac{d\mu_r}{d_\lambda} = \sum_{x=0}^{\infty} r\,(x-\lambda)^{r-1}\,(-1) + \sum_{x=0}^{\infty} \frac{(x-\lambda)^r}{x\,!}\left(\lambda^{x-\lambda} e(-1) + e^{-\lambda}\cdot x\cdot\lambda^{x-1}\right)$$

$$= -r\sum_{x=0}^{\infty}(x-\lambda)^{r-1}\frac{e^{-\lambda}\lambda^{x}}{x!} + \sum_{x=0}^{\infty}\frac{(x-\lambda)^r}{x!}\left(x e^{-\lambda}\lambda^{x-1} - \lambda^{x} e^{-\lambda}\right)$$

$$= -r\sum_{x=0}^{\infty}(x-\lambda)^{r-1}p(x) + \sum_{x=0}^{\infty}(x-\lambda)^r\frac{e^{-\lambda}\lambda^{x}}{x\,!}\left(\frac{x}{\lambda}-1\right)$$

$$= -r\sum_{x=0}^{\infty}(x-\lambda)^{r-1}p(x) + \sum_{x=0}^{\infty}(x-\lambda)^r\,p(x)\cdot\frac{(x-\lambda)}{\lambda}$$

$$= -r\sum_{x=0}^{\infty}(x-\lambda)^{r-1}p(x) + \frac{1}{\lambda}\sum_{x=0}^{\infty}(x-\lambda)^{r+1}p(x)$$

$$= -r\,\mu_{r-1} + \frac{1}{\lambda}\,\mu_{r+1}$$

$$\therefore \frac{d\mu_r}{d\lambda} = -r\mu_{r-1} + \frac{1}{\lambda}\mu_{r+1}$$

$$\Rightarrow \frac{1}{\lambda}\mu_{r+1} = \frac{d\mu_r}{d\lambda} + r\,\mu_{r-1}$$

$$\Rightarrow \mu_{r+1} = \lambda\left[\frac{d\mu_r}{d\lambda} + r\,\mu_{r-1}\right]$$

Hence the recurrence relation for the moments is $\mu_{r+1} = \lambda\left[\frac{d\mu_r}{d\lambda} + r\mu_{r-1}\right]$

Putting $r = 1$, we get $\mu_2 = \lambda\left[\frac{d\mu_1}{d\lambda} + \mu_0\right]$

$$= \lambda\,[0+1] \qquad \because \mu_1 = 0$$

$$= \lambda \qquad \mu_0 = 1$$

For $r = 2$, we get $\mu_3 = \lambda\left[\frac{d\mu_2}{d\lambda} + 2\mu_1\right]$

$$= \lambda\left[\frac{d}{d\lambda}(\lambda) + 0\right]$$

$$= \lambda\,[1+0]$$

$$= \lambda$$

for $r = 3$, we get $\mu_4 = \lambda\left[\frac{d\mu_3}{d\lambda} + 3\mu_2\right]$

$$= \lambda\left[\frac{d}{d\lambda}(\lambda) + 3\lambda\right]$$

$$= \lambda\,[1+3\lambda]$$

$$= \lambda + 3\lambda^2$$

10.7 Moment Generating Function of Poisson Distribution:

If X follows P.D. with parameter λ then its p.m.f. is given by

$$p(x) = \frac{e^{-\lambda}\lambda^{x}}{x\,!} \quad ; \quad x = 0,1,2,\cdots\cdots$$

The M.G.F. of X is given by

$$M_X(t) = E\left[e^{tX}\right]$$

$$= \sum_{x=0}^{\infty} e^{tx} p(x)$$

$$= \sum_{x=0}^{\infty} e^{tx} \cdot \frac{e^{-\lambda}\lambda^{x}}{x\,!}$$

$$= \sum_{x=0}^{n} e^{-\lambda} \cdot \frac{\left(\lambda e^{t}\right)^{x}}{x\,!}$$

$$= e^{-\lambda} \sum_{x=0}^{\infty} \frac{\left(\lambda e^{t}\right)^{x}}{x\,!}$$

$$= e^{-\lambda} \cdot e^{\lambda e^{t}}$$

$$= e^{\lambda\left(e^{t}-1\right)}$$

Hence M.G.F. of P.D. is $M_X(t) = e^{\lambda\left(e^{t}-1\right)}$

Mean and variance using M.G.F.

$$\mu_1^1 = \left[\frac{d}{dt} M_X(t)\right]_{t=0}$$

$$= \left[\lambda e^{t} e^{\lambda\left(e^{t}-1\right)}\right]_{t=0}$$

$$= \lambda e^{0} e^{0} \qquad\qquad e^{0} = 1$$

$$= \lambda$$

$\therefore$ Mean $\mu_1^1 = \lambda$

$$\mu_2^1 = \left[\frac{d^2 M_X(t)}{dt^2}\right]_{t=0}$$

$$= \left[\frac{d}{dt} \lambda e^{t} \cdot e^{\lambda\left(e^{t}-1\right)}\right]_{t=0}$$

$$= \lambda\left[\lambda e^{0} \cdot e^{0} e^{0} + e^{0} \cdot e^{0}\right]$$

$$= \lambda[\lambda + 1]$$

$$= \lambda^2 + \lambda$$

$$\therefore \mu_2^1 = \lambda^2 + \lambda$$

$\therefore$ Variance $\mu_2 = \mu_2^1 - \left(\mu_1^1\right)^2$

$$= \lambda^2 + \lambda - \lambda^2$$

$$= \lambda$$

10.8 Characteristic Function of The Poisson Distribution:

By definition $d_X(t) = E\left[e^{itX}\right]$

$$= \sum_{x=0}^{\infty} e^{itx} p(x)$$

$$= \sum_{x=0}^{\infty} e^{itx} \cdot \frac{e^{-\lambda}\lambda^{x}}{x!}$$

$$= e^{-\lambda} \sum_{x=0}^{\infty} \frac{\left(\lambda e^{it}\right)^{x}}{x!}$$

$$= e^{-\lambda} \cdot e^{\lambda e^{it}}$$

$$= e^{\lambda\left(e^{it}-1\right)}$$

$\therefore$ c. f. of P.D. is $d_X(t) = e^{\lambda\left(e^{it}-1\right)}$

Mean and Variance using c.f.

$$\mu_1^1 = (-i)\left[\frac{\partial d_X(t)}{\partial t}\right]_{t=0}$$

$$= (-i)\left[\lambda i\, e^{it} \cdot e^{\lambda\left(e^{it}-1\right)}\right]_{t=0}$$

$$= (-i)\left[\lambda\, i\, e^{0}\, e^{0}\right]$$

$$= -i^2\, \lambda$$

$$= \lambda \qquad \because i^2 = -1$$

$\therefore$ Mean $\mu_1^1 = \lambda$

$$\mu_2^1 = (-i)^2\left[\frac{\partial^2 d_X(t)}{\partial t^2}\right]_{t=0}$$

$$= (-i)^2\left[\frac{\partial}{\partial t}\lambda\, i e^{it}\, e^{\lambda\left(e^{it}-1\right)}\right]_{t=0}$$

$$= (-i)^2 \, i \, \lambda \left[e^{it} \cdot \lambda i \, e^0 e^0 + e^0 \, i \cdot e^0 \right]$$

$$= (-i)^2 \, i \, \lambda \, [\lambda \, i + i]$$

$$= (-i)^2 \, i \, \lambda [\lambda \, i + i]$$

$$= (-i)^2 \, i^2 \, \lambda \, (\lambda + 1)$$

$$= i^4 \, \lambda \, (\lambda + 1)$$

$$= \lambda^2 + \lambda \qquad \because \; i^2 = -1$$

$$\therefore \; \mu_2^1 = \lambda^2 + \lambda$$

$$\therefore \quad \text{variance} \quad \mu_2 = \mu_2^1 - \left(\mu_1^1\right)^2$$

$$= \lambda^2 + \lambda - \lambda^2$$

$$= \lambda$$

10.9 Cumulants of The Poisson Distribution:

By definition $K_X(t) \; = \; \ell og \, M_X(t)$

$$= \ell og \left[e^{\lambda \left(e^t - 1\right)} \right]$$

$$= \lambda \left(e^t - 1\right)$$

$$= \lambda \left[1 + t + \frac{t^2}{2!} + \frac{t^3}{3!} + \cdots\cdots\cdots - 1 \right]$$

$$= \lambda \left[\frac{t}{1!} + \frac{t^2}{2!} + \frac{t^3}{3!} + \cdots\cdots\cdots + \frac{t^r}{r!} + \cdots\cdots\cdots \right]$$

$\therefore$ The r^{th} cumulant k_r = The Coefficient of $\frac{t^r}{r!}$ in $K_X(t) = \lambda$

Hence all the cumulants of P.D. are equal to λ

$$\therefore \text{ Mean } = \mu_1^1 = k_1 = \lambda$$

$$\text{Variance } \mu_2 = k_2 = \lambda$$

$$\mu_3 = k_3 = \lambda$$

$$\mu_4 = k_4 + 3k_2^2$$

$$= \lambda + 3\lambda^2$$

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{\lambda^2}{\lambda^3} = \frac{1}{\lambda}$$

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{\lambda + 3\lambda^2}{\lambda^2} = 3 + \frac{1}{\lambda}$$

10.10 Probability Generating Function of Poisson Distribution:

By definition P.G.F. of P.D. is given by

$$p(S) = E\left(S^X\right)$$

$$= \sum_{x=0}^{\infty} S^x \cdot \frac{e^{-\lambda} \lambda^x}{x!}$$

$$= e^{-\lambda} \cdot \sum_{x=0}^{\infty} \frac{(\lambda S)^x}{x!}$$

$$= e^{-\lambda} \cdot e^{\lambda S}$$

$$= e^{\lambda(S-1)}$$

10.11 Additive or Reproductive Property of Poisson Distribution:

Statement: Prove that sum of independent poisson variates is also a Poisson variate.

Proof: Let $X_1, X_2, \cdots\cdots, X_n$ are independent poisson variates with parameters $\lambda_1, \lambda_2, \cdots\cdots, \lambda_n$ respectively.

i.e. $X_i \sim P(\lambda_i)$; $i = 1, 2, \cdots\cdots, n$

then its M.G.F. is given by $M_{X_i}(t) = e^{\lambda_i (e^t - 1)}$; $i = 1, 2, \cdots\cdots, n$

$\therefore$ The M.G.F. of sum of poisson variates are given by

$$M_{X_1, X_2, \cdots\cdots, X_n}(t) = M_{X_1}(t) \cdot M_{X_2}(t) \cdots\cdots M_{X_n}(t) \quad \because X_i\text{'s are ind...}$$

$$= e^{\lambda_1 (e^t - 1)} \cdot e^{\lambda_2 (e^t - 1)} \cdots\cdots e^{\lambda_n (e^t - 1)}$$

$$= e^{(\lambda_1 + \lambda_2 + \cdots\cdots + \lambda_n)(e^t - 1)}$$

Which is the M.G.F. of a Poisson Variate with parameters $\lambda_1 + \lambda_2 + \cdots\cdots + \lambda_n$.

Hence by uniqueness theorem of M.G.F.'s $\sum_{x=1}^{n} X_i$ is also a Poisson Variate with parameter $\sum_{i=1}^{n} \lambda_i$.

Hence sum of the independent Poisson Variates is also a Poisson Variate.

Note: The difference of two independent Poisson Variates is not a Poisson Variate.

10.12 Mode of The Poisson Distribution:

If $X \sim p(\lambda)$ then its p.m.f. is given by

$$p(x) = \frac{e^{-\lambda} \lambda^x}{x!} \quad ; \quad x = 0, 1, 2, \cdots\cdots$$

Consider $\dfrac{p(x)}{p(x-1)} = \dfrac{\dfrac{e^{-\lambda}\lambda^x}{x!}}{\dfrac{e^{-\lambda}\lambda^{x-1}}{(x-1)!}} = \dfrac{e^{-\lambda}\lambda^x}{x(x-1)!} \times \dfrac{(x-1)!}{e^{-\lambda}\lambda^{x-1}}$

$$\frac{p(x)}{p(x-1)} = \frac{\lambda}{x}$$

Case (i): If λ is not an integer, then $\lambda = m + t$ where m is an integral part and $t\,(0 < t < 1)$ is a fractional part of the λ.

$$\frac{p(x)}{p(x-1)} = \frac{m+t}{x} \quad ; \quad x = 0,1,2,\cdots\cdots, m,\ m+1,\cdots\cdots$$

It is clear that $\dfrac{p(1)}{p(0)} > 1,\ \dfrac{p(2)}{p(1)} > 1, \cdots\cdots, \dfrac{p(m)}{p(m-1)} > 1,\ \dfrac{p(m+1)}{p(m)} < 1, \cdots\cdots$

$$\therefore\ p(0) < p(1) < p(2) < \cdots\cdots < p(m-1) < p(m) > p(m+1) > p(m+2) > \cdots\cdots$$

Which shows that $p(m)$ is the maximum value.

$\therefore$ m is the mode and which is an integral part of λ.

Case (ii): If λ is an integer say $\lambda = m$ then

$$\frac{p(x)}{p(x-1)} = \frac{m}{x} \quad ; \quad x = 0,1,2,\cdots\cdots, m,\ m+1,\cdots\cdots$$

$$\therefore\ \frac{p(1)}{p(0)} > 1,\ \frac{p(2)}{p(1)} > 1, \cdots\cdots, \frac{p(m-1)}{p(m-2)} > 1,\ \frac{p(m)}{p(m-1)} = 1,\ \frac{p(m+1)}{p(m)} < 1, \cdots\cdots$$

$$\Rightarrow p(0) < p(1) < p(2) < \cdots\cdots < p(m-1) = p(m) > p(m+1) > \cdots\cdots$$

In this case we have two maximum values they are $p(m-1)$ and $p(m)$.

$\therefore$ Mode values are m and $m-1$ when λ is integer.

10.13 Recurrence Relation For The Probabilities of P.D.:

If $X \sim p(\lambda)$ then its p.m.f. is given by

$$p(x) = \frac{e^{-\lambda}\,\lambda^{x}}{x\,!} \quad ; \quad x = 0,1,2,\cdots\cdots$$

$$\Rightarrow p(x+1) = \frac{e^{-\lambda}\,\lambda^{x+1}}{(x+1)\,!} \quad ; \quad x = 0,1,2,\cdots\cdots$$

$$\therefore \frac{p(x+1)}{p(x)} = \frac{e^{-\lambda}\,\lambda^{x+1}}{(x+1)\,!} \times \frac{x\,!}{e^{-\lambda}\,\lambda^{x}}$$

$$= \frac{\lambda}{x+1}$$

$$\Rightarrow p(x+1) = \left(\frac{\lambda}{x+1}\right) p(x)$$

Hence recurrence relation between probability of P.D. is

$$p(x+1) = \left(\frac{\lambda}{x+1}\right) p(x)$$

This formula is very convenient to calculate the probability of P.D. The value of P_0 given by $p(0) = e^{-\lambda}$. Where λ is the mean of the given frequency distribution. The other probabilities can be obtained by using above recurrence relation.

10.14 Worked Examples:

Example 1: If the probabilities of a Poisson Variate taking the values 3 and 4 are equal, calculate the probabilities of the variable taking the values 0 and 2.

Solution: If $X \sim p(X)$ the its p.m.f. is given by

$$p(X = x) = \frac{e^{-\lambda} \cdot \lambda^{x}}{x!} \quad ; \quad x = 0,1,2,\cdots\cdots$$

Given that $p(x=3) = p(x=4)$

$$\Rightarrow \frac{e^{-\lambda}\lambda^3}{3!} = \frac{e^{-\lambda}\cdot\lambda^4}{4\,!}$$

$$\Rightarrow \frac{\lambda^3}{6} = \frac{\lambda^4}{24}$$

$$\Rightarrow \lambda = 4$$

$$\therefore\ p(X=0) = \frac{e^{-4}\cdot 4^0}{0\,!} = e^{-4} = 0\cdot 0183$$

$$p(x=2) = \frac{e^{-4}\cdot 4^2}{2!} = e^{-4}\times 8 = 0\cdot 146$$

Example 2: The probability of getting no misprint in a page of a book is e^{-4}. Determine the probability that a page of a book contains more than 2 misprints.

Solution: Given that $p(x=0) = e^{-4}$

$$\Rightarrow e^{-\lambda} = e^{-4}$$

$$\Rightarrow \lambda = 4$$

$$\therefore p(X>2) = 1 - p(X \le 2)$$

$$= 1 - [p(X=0) + p(X=1) + p(X=2)]$$

$$= 1 - \left[\frac{e^{-4}\cdot 4^0}{0\,!} + \frac{e^{-4}\times 4^1}{1\,!} + \frac{e^{-4}\times 4^2}{2\,!}\right]$$

$$= 1 - e^{-4}\left[1 + 4 + \frac{16}{2}\right]$$

$$= 1 - e^{-4}\times 13$$

$$= 1 - 13\times 0\cdot 0183$$

$$= 0\cdot 762$$

Example 3: A telephone switch board receives 20 calls on an average during an hour. Find the probability that during a period of 5 minutes (i) no calls is received, (ii) exactly 3 calls are received, (iii) more than 5 calls are received.

Solution: We assume that the number of incoming cells during any time period follows Poisson Process. 20 calls per hour is equivalent to 0.33 calls perminute, which is the mean rate of occurrance. Hence number of cells in 5 minute period follows a Poisson Dostribution with parameter $\lambda = 1 \cdot 65$

$$\because \lambda\ t = 0 \cdot 33 \times 5 = 1 \cdot 65$$

(i) p (no calls in a 5 minute period) $= e^{-1\cdot 65} = 0 \cdot 192$

(ii) p (3 calls in a 5 minute period) $= \dfrac{e^{-1\cdot 65}\ (1 \cdot 65)^3}{3\,!} = 0 \cdot 144$

(iii) p (more than 5 calls in a 5 minute period) $= \sum_{x=6}^{\infty} \dfrac{e^{-1\cdot 65}\ (1 \cdot 65)^x}{x\,!} = 0 \cdot 007$

Example 4: Assuming that one in 80 births is a case of twins, calculate the probability of 2 or more births of twins on a day when 30 births occur using (i) Binomial Distribution, (ii) Poisson Approximation.

Solution: (i) Assuming X to be a Binomial Variate.

$p =$ probability of twin births $= \dfrac{1}{80} = 0 \cdot 0125$

$q = 1 - p$

$= 0 \cdot 9875$

given $n = 30$

$$\therefore p(X = x) = 30\ C_x\ (0 \cdot 0125)^x\ (0 \cdot 9875)^{30-x}$$

$\therefore$ Probability of 2 or more births of twins on a day is

$$p(X \geq 2) = 1 - p(X < 2)$$

$$= 1 - \left[p(X = 0) + p(X = 1)\right]$$

$$= 1 - \left[30C_0\,(0\cdot 0125)^0\,(0\cdot 9875)^{30} + 30C_1\,(0\cdot 0125)^1\,(0\cdot 9875)^{29}\right]$$

$$\Rightarrow \frac{e^{-\lambda}\,\lambda^2}{2} = \frac{2}{3}\,e^{-\lambda}\cdot\lambda$$

$$\Rightarrow 3\,\lambda^2 = 4\,\lambda$$

$$\Rightarrow \lambda = \frac{4}{3}$$

$$\therefore\ p(X=0) = \frac{e^{-\lambda}\cdot\lambda^0}{0\,!} = e^{-\lambda} = e^{-4/3}$$

$$p(X=3) = \frac{e^{-\lambda}\cdot\lambda^3}{3\,!} = \frac{e^{-4/3}\left(\frac{4}{3}\right)^3}{6} = \frac{64}{27\times 6}\,e^{-4/3}$$

$$= \frac{32}{81}\,e^{-4/3}$$

10.15 Exercise:

1. Derive Poisson distribution as a limiting form of a Binomial Distribution.
2. State and prove reproductive property of Poisson Distribution.
3. Derive the recurrence relation between the moments of Poisson Distribution and hence obtain the Skewness and Kurtosis.
4. If X is a Poisson variate with parameter λ and r is a non - negative integer then prove that

$$\mu^1_{r+1} = \lambda\left(\mu^1_r + \frac{d\mu^1_r}{d\lambda}\right) \qquad \text{where } \mu^1_r = E\left(X^r\right)$$

5. Show that all the Cumulants are equal to the parameter λ.
6. Derive M.G.F. of Poisson Distribution and hence find its mean and variance.
7. Derive P.G.F. of Poisson Distribution and hence find its mean and variance.
8. Find the cumulant generating function of Poisson Distribution. Using cumulants find the first form Central Moments.

9. Derive the characteristic function fo poisson distribution and find the mean and variance from it.

10. Derive mode of Poisson Distribution.

11. Assuming that the probability that a bomb dropped from an aeroplane will hit a largest is $\frac{1}{5}$. If 6 bombs are dropped, find the probability that

(i) Exactly two will hit the target.

(ii) Atleast two will hit the target.

12. If X and Y are independent Poisson Variates having means 1 and 3 respectively, find the mean and variance of $3X+Y$.

13. Show that for a Poisson Distribution $\beta_1^{1/2}(\beta_2-3)\mu_1^1\sigma=1$.

14. In a book of 520 pages, 390 typo - graphical errors occur. Assuming Poisson law for the number of errors per page, find the probability that a random sample of 5 pages will contain no error.

15. If X is a Poisson variate such that

$$p(X=2)=9p(X=4)+90\ p(X=6)$$

Find X value, mean, variance and coefficient of skewness.

10.16 Answers:

11. (i) $0\cdot 24576$ (ii) $0\cdot 34464$

12. 6, 12.

14. $[p(X=0)]^5=e^{-3\cdot 75}$

15. $\lambda=1,\ \text{mean}=1,\ \text{variance}=1,\ p_1=1$

Lesson – 11

NEGATIVE BINOMIAL DISTRIBUTION

Objective:

After studying the material of this lesson, the student supposed to have a clear concept about the negative binomial distribution, its dualities with the ordinary binomial distribution, the situations where if works well.

Structure of The Lesson:

11.1 Introduction:

The important characteristic of the Binomial Distribution is mean value is always greater than the variance. The Negative Binomial Distribution obtained by the same process that gives rise to Binomial Distribution, but its mean is always less than the variance. In Binomial Distribution, no of successes "X" varies from 0 to n where n is number of trials and is fixed, where as in Negative Binomial Distribution x is fixed and n is allowed to vary. Some of such sitvations are (i) Death of insects, (ii) Number of insects bites, (iii) Bacterial clustering etc are leads to the Negative Binomial Distribution.

The random experiment with the following properties lead to Negative Binomial Distribution.

1. The result of each trial can be classified into one of the two mutvally exclusive out comes say success and failure.

2. Probability of success "p" remains constant for each trial.

3. The outcomes of all trials are independent of each other.

4. The series of thetrials is performed until a fixed number of success in achieved.

1.2 Defination:

A random variable X is said to follow a negative Binomial Distribution with parameters r and p if its p.m.f. is given by

$$P(x) = \begin{cases} \binom{x+r-1}{r-1} p^r q^x \; ; & x = 0,1,2,\cdots\cdots \\ 0 & \text{otherwise} \end{cases}$$

Simplified form of Negative Binomial Distribution

$$\binom{x+r-1}{r-1} = \binom{x+r-1}{x}$$

$$= \frac{(x+r-1)(x+r-2)\cdots\cdots\cdots(r+1)r}{x\,!}$$

$$= \frac{(-1)^x(-r)(-r-1)\cdots\cdots\cdots(-r-x+2)(-r-x+1)}{x!}$$

$$= (-1)^x \binom{-r}{x}$$

$$\therefore \; P(x) = \binom{-r}{x} p^r (-q)^x \quad ; \quad x = 0,1,2,\cdots\cdots$$

which is the $(x+1)^{th}$ term in the expansion of $p^r(1-q)^{-r}$, a binomial expansion with negative index. Hence the distribution is known as negative binomial distribution.

The relation which establishes a similarity to Binomial Distribution:

If $P = \frac{1}{Q}$ and $q = \frac{p}{Q}$ so that $Q - P = 1$ $\quad \therefore \; p + q = 1$

$\therefore$ The p.m.f. of NBN becomes

$$p(x) = \binom{-r}{x} Q^{-r} \left(\frac{-p}{Q}\right)^x \quad ; \quad x = 0,1,2,\cdots\cdots$$

This is the general term in the Negative Binomial expansion $(Q-P)^{-r}$.

That is the relation with B.D. is $p=-P$

$$n=-r$$

11.3 Moments of Negative Binomial Distribution:

$$\mu_1^1 = E(X) = \sum_{x=0}^{\infty} x\, p(x)$$

$$= \sum_{x=1}^{\infty} x \cdot \binom{x+r-1}{x} p^r q^x$$

$$= \sum_{x=1}^{\infty} x \frac{(x+r-1)(x+r-2)\cdots\cdots(r+1)r}{x\,(x-1)!} p^r q^x$$

$$= r\, p^r q \sum_{x=0}^{\infty} \binom{x+r-1}{x-1} q^{x-1}$$

$$= r p^r q\, (1-q)^{-(r+1)}$$

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

$$= \frac{rq}{p}$$

$$\therefore \text{ Mean } = \mu_1^1 = rq/p$$

$$\mu_2^1 = E\left[X^2\right]$$

$$= E\left[X(X-1)+X\right]$$

$$= E\left[X(X-1)\right] + E(X) \qquad (11.3.1)$$

Now $E[X(X-1)] = \sum_{x=0}^{\infty} x(x-1)\, p(x)$

$$= \sum_{x=2}^{\infty} x(x-1)\binom{x+r-1}{x} p^r q^x$$

$$= p^r\, r(r+1)\, q^2 \sum_{x=2}^{\infty} \frac{(x+r-1)(x+r-2)\cdots\cdots(r+2)q^{x-2}}{(x-2)!}$$

$$= r(r+1)\, p^r q^2 \sum_{x=2}^{\infty} \binom{x+r-1}{x-2} q^{x-2}$$

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

$$= \frac{r(r+1)\, p^r q^2}{(1-q)^{r+2}}$$

$$= \frac{r(r+1)\, p^r q^2}{p^{r+2}} = \frac{r(r+1) p^r q^2}{p^r p^2} = \frac{r(r+1)\, q^2}{p^2} \qquad (11.3.2)$$

substituting (11.3.2) in (11.3.1) we get

$$\mu_2^1 = r(r+1)\frac{q^2}{p^2} + {}^{rq}\!/_{p} \qquad \because E(X) = {}^{rq}\!/_{p}$$

$$\therefore \mu_2 = \mu_2^1 - \left(\mu_1^1\right)^2$$

$$= r(r+1)\frac{q^2}{p^2} + \frac{rq}{p} - \frac{r^2q^2}{p^2}$$

$$= \frac{r^2q^2}{p^2} + \frac{rq^2}{p^2} + \frac{rq}{p} - \frac{r^2q^2}{p^2}$$

$$= \frac{rq}{p}\left(\frac{q}{p}+1\right)$$

$$= \frac{rq}{p}\left(\frac{q+p}{p}\right)$$

$$= \frac{rq}{p}\left(\frac{1}{p}\right)$$

$$= \frac{r\,q}{p^2} \qquad\qquad \therefore\ q+p=1$$

$\therefore$ variance $\mu_2 = rq/p^2$

Note: Mean < variance, which is a special feature of the NBD.

11.4 Moment Generating Function of Negative Binomial Distribution:

The p.m.f. of Negative Binomial Distribution is given by

$$p(x) = \binom{x+r-1}{x} p^r q^x \qquad ; \qquad x = 0,1,2,\cdots\cdots$$

The M.G.F. of NBD is given by

$$M_X(t) = E\left[e^{tX}\right]$$

$$= \sum_{x=0}^{\infty} e^{tX}\, p(x)$$

$$= \sum_{x=0}^{\infty} e^{tX} \binom{x+r-1}{x} p^r q^x$$

$$= p^r \sum_{x=0}^{\infty} \binom{x+r-1}{x} \left(q\,e^t\right)^x$$

$$= p^r \left(1-qe^t\right)^{-r} \qquad \because \sum_{x=0}^{\infty} \binom{x+r-1}{x} q^x = (1-q)^{-r}$$

$$\therefore M_X(t) = p^r \left(1-qe^t\right)^{-r}$$

Mean and variance using M.G.F.

$$\mu_1^1 = \left[\frac{d M_X(t)}{dt}\right]_{t=0}$$

$$= p^r \left[\frac{d}{dt}\left(1-qe^t\right)^{-r}\right]_{t=0}$$

$$= p^r \left[-r\left(1-qe^t\right)^{-r-1}\left(-qe^t\right)\right]_{t=0}$$

$$= p^r \left[rq\,(1-q)^{-(r+1)}\right]$$

$$= p^r \; \frac{rq}{p^{r+1}}$$

$$= r\,\frac{q}{p}$$

$$\therefore \text{ Mean } = \mu_1^1 = {rq}/{p}$$

$$\mu_2^1 = \left[\frac{d^2 M_X(t)}{dt^2}\right]_{t=0}$$

$$= \left\{\frac{d}{dt}\left[p^r\; r\,q\,e^t\left(1-qe^t\right)^{-r-1}\right]\right\}_{t=0}$$

$$= r\,q\,p^r \left[\left(1-qe^t\right)^{-(r+1)} e^t + \left\{-(r+1)\left(1-qe^t\right)^{-(r+2)}\left(-qe^t\right)e^t\right\}\right]_{t=0}$$

$$= r q p^r \left[(1-q)^{-(r+1)} + \left\{ (r+1) q (1-q)^{-(r+2)} \right\} \right]$$

$$= r q p^r \left[\frac{1}{p^{r+1}} + \frac{(r+1) q}{p^{r+2}} \right]$$

$$= \frac{r q}{p} + r(r+1) \frac{q^2}{p^2}$$

$$\therefore \quad \text{Variance } \mu_2 = \mu_2^1 - \left(\mu_1^1\right)^2$$

$$= \frac{rq}{p} + r(r+1) \frac{q^2}{p^2} - \frac{r^2 q^2}{p^2}$$

$$= rq/p^2$$

11.5 Characteristic Function of Ngative Binomial Distribution:

If X follows Negative Binomial Distribution with parameters p and r then its p.m.f. is given by

$$p(x) = \binom{x+r-1}{x} p^r q^x \qquad ; \qquad x = 0,1,2,\cdots\cdots$$

The c.f. of X is given by

$$d_X(t) = E\left[e^{itX}\right]$$

$$= \sum_{x=0}^{\infty} e^{itX} p(x)$$

$$= \sum_{x=0}^{\infty} e^{itX} \binom{x+r-1}{x} p^r q^x$$

$$= p^r \sum_{x=0}^{\infty} \binom{x+r-1}{x} \left(q e^{it}\right)^x$$

$$= p^r \left(1 - q e^{it}\right)^{-r}$$

$$\therefore \phi_X(t) = \frac{p^r}{\left(1 - q e^{it}\right)^r}$$

11.6 Probability Generating Function:

If X follows NBD with parameter r and p then its p.m.f. is given by

$$p(x) = \binom{x+r-1}{x} p^r q^x \quad ; \quad x = 0, 1, 2, \cdots\cdots$$

The P.G.F. of X is given by

$$P(S) = E\left(S^x\right)$$

$$= \sum_{x=0}^{\infty} S^x p(x)$$

$$= \sum_{x=0}^{\infty} S^x \binom{x+r-1}{x} p^r q^x$$

$$= \sum_{x=0}^{\infty} \binom{x+r-1}{x} p^r (qS)^x$$

$$= p^r \sum \binom{x+r-1}{x} (qS)^x$$

$$\therefore P(S) = p^r (1 - qS)^{-r}$$

Mean and variance using P.G.F.

$$\text{mean} = \mu_1^1 = E(X) = \left[\frac{d}{ds} p(s)\right]_{s=1}$$

$$= \left[\frac{d}{ds} p^r (1 - qs)^{-r}\right]_{s=1}$$

$$= \left[\frac{d}{ds} p^r (1-qs)^{-r} \right]_{s=1}$$

$$= p^r \left[r\,q\,(1-q)^{-(r+1)} \right]$$

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

$$= \frac{rqp^r}{p^{r+1}}$$

$$= \frac{r\,q}{p}$$

$$\therefore \text{ Man } \mu_1^1 = \frac{rq}{p}$$

We know that $\mu_2^1 = \left[p^{11}(s) + p^1(s) \right]_{s=1}$

$$p^{11}(s) = \frac{d}{ds}\left[p^1(s) \right]$$

$$= \frac{d}{ds}\left[rp^r q\,(1-qs)^{-(r+1)} \right]$$

$$= p^r rq \left[-(r+1)(1-qs)^{-(r+2)} (-q) \right]_{s=1}$$

$$= p^r\ rq \left[-(r+1)(1-q)^{-(r+2)} (-q) \right]$$

$$= \frac{p^r\ r\,(r+1)q^2}{(1-q)^{r+2}} = \frac{p^r\,r(r+1)q^2}{p^{r+2}}$$

$$= r(r+1)\frac{q^2}{p^2}$$

$$\therefore \quad \mu_2^1 = r(r+1)\frac{q^2}{p^2} + \frac{rq}{p} \qquad \because \left[p^1(s)\right]_{s=1} = {}^{rq}\!/_{p}$$

$$\therefore \text{ Variance } \mu_2 = \mu_2^1 - \left(\mu_1^1\right)^2$$

$$= r(r+1)\frac{q^2}{p^2} + \frac{rq}{p} - \frac{r^2q^2}{p^2}$$

$$= \frac{r^2q^2}{p^2} + \frac{rq^2}{p^2} + \frac{rq}{p} - \frac{r^2q^2}{p^2}$$

$$= \frac{rq}{p}\left(\frac{q}{p}+1\right)$$

$$= \frac{r\,q}{p}\left(\frac{q+p}{p}\right)$$

$$= \frac{rq}{p^2} \qquad \therefore q + p = 1$$

Hence mean $= {}^{rq}\!/_{p}$ and variance $= {}^{rq}\!/_{p^2}$

11.7 Poisson Distribution As A Limiting Case of The Neagative Binomial Distribution:

Negative Binomial Distribution tends to Poisson Distribution under the following conditions.

1. $P \to 0$
2. $r \to \infty$
3. $r\,P = \lambda$, finite

The p.m.f. of NBD is given by

$$P(x)=\binom{x+r-1}{x}p^r\,q^x$$

$$=\binom{x+r-1}{x}Q^{-r}\left(\frac{P}{Q}\right)^x$$

$$=\frac{(x+r-1)\,(x+r-2)\cdots\cdots(r+1)^r\,(1+p)^{-r}\left(\frac{p}{1+p}\right)^x}{x\,!}$$

$$=\frac{\left[1+\left(\frac{x-1}{r}\right)\right]\left[1+\left(\frac{x-2}{r}\right)\right]\cdots\cdots\cdots\cdots\left[1+\frac{1}{r}\right]}{x\,!}r^x\,p^x\cdot(1+p)^{-r}\left(\frac{1}{1+p}\right)^x$$

$$=\frac{\left[1+\left(\frac{x-1}{r}\right)\right]\left[1+\left(\frac{x-2}{r}\right)\right]\cdots\cdots\cdots\cdots\left[1+\frac{1}{r}\right]}{x!}(r\,p)^x\,(1+p)^{-r}\left(\frac{1}{1+p}\right)^x$$

Now proceeding to the limits, we get

$$\lim_{r\to\infty}P(x)=\frac{1}{x!}\lim_{r\to\infty}\left\{\left[1+\left(\frac{x-1}{r}\right)\right]\left[1+\left(\frac{x-2}{r}\right)\right]\cdots\cdots\left[1+\frac{1}{r}\right]\right\}\lambda^x\cdot\lim_{r\to\infty}\left(1+\frac{\lambda}{r}\right)^{-r}\lim_{r\to\infty}\left(1+\frac{\lambda}{r}\right)^{-x}$$

$$=\frac{1}{x!}\lambda^x\cdot e^{-\lambda}\cdot 1$$

$$=\frac{e^{-\lambda}\,\lambda^x}{x!}$$

Which is the p.m.f. of the Poisson Distribution with parameter λ.

11.8 Recurrence Relation Betwen Central Moments of NBD:

Theorem: Show that the recurrence relation between central moments of NBD is

$$\mu_{k+1} = q\left[\frac{kr}{p^2}\mu_{k-1} - \frac{d\mu_k}{dp}\right]$$

Proof: By definition, p.m.f. of NBD is given by

$$P(x) = \binom{x+r-1}{x} p^r q^x$$

$$\therefore \mu_k = \sum_{x=0}^{\infty}\left(x-\frac{rq}{p}\right)^k p(x)$$

$$= \sum_{x=0}^{\infty}\left(x-\frac{rq}{p}\right)^k \binom{x+r-1}{x} p^r q^x$$

Differentiating w.r.t. p we get

$$\frac{d\mu_k}{dp} = \sum_{x=0}^{\infty}\binom{x+r-1}{x}\left[\left(x-\frac{rq}{p}\right)^k \left\{p^r\, x\, q^{x-1}(1) + q^x\, r p^{r-1}\right\} + p^r q^x k\left(x-\frac{rq}{p}\right)^{k-1}\frac{r}{p^2}\right]$$

$$= \sum_{x=0}^{\infty}\binom{x+r-1}{x}\left(x-\frac{rq}{p}\right)^k p^r q^x \left\{\frac{r}{p}-\frac{x}{q}\right\} + \sum_{x=0}^{\infty}\binom{x+r-1}{x}\left(x-\frac{rq}{p}\right)^{k-1} p^r q^x \frac{kr}{p^2}$$

$$= -\frac{1}{q}\sum_{x=0}^{\infty}\binom{x+r-1}{x} p^r q^x \left(x-\frac{rq}{p}\right)^k\left(x-\frac{rq}{p}\right) + \frac{rk}{p^2}\sum_{x=0}^{\infty}\left(x-\frac{rq}{p}\right)^{k-1}\binom{x+r-1}{x} p^r q^x$$

$$= -\frac{1}{q}\sum_{x=0}^{\infty}\left(x-\frac{rq}{p}\right)^{k+1} p(x) + \frac{rk}{p^2}\sum_{x=0}^{\infty}\left(x-\frac{rq}{p}\right)^{k-1} p(x)$$

$$= -\frac{1}{q}\mu_{k+1} + \frac{rk}{p^2}\mu_{k-1}$$

$$\Rightarrow \frac{d\mu_k}{dp} = -\frac{1}{q}\mu_{k+1} + \frac{rk}{p^2}\mu_{k-1}$$

$$\Rightarrow \frac{1}{q}\mu_{k+1} = \frac{rk}{p^2}\mu_{k-1} - \frac{d\mu_k}{dp}$$

$$\Rightarrow \mu_{k+1} = q\left[\frac{rk}{p^2}\mu_{k-1} - \frac{d\mu_k}{dp}\right]$$

In particular if $k = 1$ then

$$\mu_2 = q\left[\frac{n}{p^2}\mu_0 - \frac{d}{dp}\mu_1\right]$$

$$= q\left[\frac{r}{p}\right] \qquad \mu_0 = 1,\ \mu_1 = 0$$

$$= {}^{rq}\!\!/_{p}$$

If $k = 2$ we get

$$\mu_3 = q\left[\frac{2r}{p^2}\mu_1 - \frac{d}{dp}\mu_2\right]$$

$$= q\left[-\frac{d}{dp}\left(\frac{rq}{p}\right)\right]$$

$$= -rq\left(\frac{-2}{p^3} + \frac{1}{p^2}\right)$$

$$= \frac{2rq}{p^3} - \frac{\lambda qp}{p^3}$$

$$= \frac{rq}{p^3}(1+1-p)$$

$$\mu_3 = \frac{rq}{p^3}(1+q)$$

Similarly if $k=3$, then $\mu_4 = \frac{rq}{p^4}\left[p^2 + 3q(r+2)\right]$

11.9 Recurrence Relation For The Probabilities of NBD:

By definition $P(x) = \binom{x+r-1}{r-1} p^r q^x$; $x = 0,1,2,\cdots\cdots$

$$\Rightarrow p(x+1) = \binom{x+r}{r-1} p^r q^{x+1}$$

$$\therefore \frac{p(x+1)}{p(x)} = \frac{\binom{x+r}{r-1} p^r q^{x+1}}{\binom{x+r-1}{r-1} p^r q^x}$$

$$= \frac{(x+r)!\,(r-1)!\,x!}{(r-1)!\,(x+1)!\,(x+r-1)!}\, q = \left(\frac{x+r}{x+1}\right) q$$

$$\therefore p(x+1) = \left(\frac{x+r}{x+1}\right) q\ p(x)$$

Hence recurrence relation between probabilities of Negative Binomial Distribution is

$$p(x+1) = \left(\frac{x+r}{x+1}\right) q\ p(x)$$

11.10 Workedout Examples:

Example 1: What is the probability that we need 5 trials to get the two successes, if the probability of success is $\frac{1}{4}$.

Solution: Given that $p = \frac{1}{4}$

$$q = \frac{3}{4}$$

$$r = 2$$

$$\therefore\ p(r = 2) = \binom{3+2-1}{2-1} \left(\frac{1}{4}\right)^2 \left(\frac{3}{4}\right)^3$$

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

$$= \frac{4 \times 27}{45}$$

$$= \frac{27}{256}$$

$$= 0 \cdot 1054$$

Example 2: If a boy is throwing stones at a target, what is the probability that his 10^{th} throw is his 5^{th} hit, if the probability of hitting the target at any trial is 0.5 ?

Solution: Given that $p = 0 \cdot 5$

$$q = 0 \cdot 5$$

$$r = 5 \qquad\qquad x + r = 10$$

$$\therefore\ p(x = 5) = \binom{5+5-1}{5-1} (0 \cdot 5)^5\ (0 \cdot 5)^5$$

$$= \binom{9}{4} (0 \cdot 5)^{10}$$

$$= 126 \times (0 \cdot 5)^{10}$$

$$= 0 \cdot 12305$$

Example 3: If the probability is 0.40 that a child exposed to a certain contagious diseas will catch it, what is the probability that the tenth child exposed to the disease will be third to catch it.

Solution: Given that $p = 0 \cdot 4$

$q = 0 \cdot 6$

$r = 3$ $\qquad$ $x + r = 10$

$$\therefore \; p(r = 3) = \binom{7+3-1}{3-1} (0 \cdot 4)^3 (0 \cdot 6)^7$$

$$= \binom{9}{2} (0 \cdot 4)^3 (0 \cdot 6)^7$$

$$= 0 \cdot 0645$$

Example 4: If the probability of getting a head is $\frac{1}{2}$, find the probability that a fourth toss is the getting of head first time.

Solution: Given that $p = \frac{1}{2}$

$q = \frac{1}{2}$

$r = 1$ $\qquad$ $x + r = 4$

$$p(r = 1) = \binom{3+1-1}{1-1} \left(\frac{1}{2}\right)^1 \left(\frac{1}{2}\right)^3$$

$$= \binom{3}{0} \left(\frac{1}{2}\right)^4$$

$$= \frac{1}{16}$$

$$= 0 \cdot 0625$$

Example 5: If the probability is 0.75 that a person will believe a rumour about the corruption of certain politicians, find the probabilities that

(i) The eighth person to hear the rumour will be the fifth to believe it.

(ii) The fifteenth person to hear the rumour will be the tenth to believe it.

Solution: (i) Given that $p = 0 \cdot 75$

$$q = 0 \cdot 25$$

$$r = 5$$

$$\therefore \; p(r-5) = \binom{5+3-1}{5-1} (0 \cdot 75)^5 \, (0 \cdot 25)^3$$

$$= \binom{7}{4} (0 \cdot 75)^5 \, (0 \cdot 25)^3$$

$$= 0 \cdot 1298$$

(ii) $$p(r=10) = \binom{14}{9} (0 \cdot 75)^{10} \, (0 \cdot 25)^5$$

$$= 0 \cdot 1101$$

Example 6: Obtain the characteristic function of the Negative Binomial Distribution given in the form $p(x) = \binom{-\lambda}{x} \left(\frac{1}{\alpha+1}\right)^{\lambda} \left(\frac{-\alpha}{1+\alpha}\right)^{x}$; $x = 0,1,2,\cdots\cdots$ and hence evaluate its first two moments.

Solution: The c.f. of X is given by

$$\phi_X(t) = E\left[e^{itX}\right]$$

$$= \sum_{x=0}^{\infty} \binom{-\lambda}{x} \left(\frac{1}{1+\alpha}\right)^{\lambda} \left(\frac{-\alpha e^{it}}{1+\alpha}\right)^{x}$$

$$= \left[(1+\alpha) - \alpha e^{it}\right]^{-\lambda}$$

$$= \left(1+\alpha-\alpha e^{it}\right)^{-\lambda}$$

$$= \left[\,1-\alpha\left(e^{it}-1\right)\right]^{-\lambda}$$

$$= 1 + \lambda\, \alpha \left(e^{it} - 1\right) + \lambda \frac{(\lambda+1)}{2} \lambda^2 \left(e^{it} - 1\right) + \cdots\cdots\cdots$$

$$= 1 + \lambda < \left[1 + \frac{it}{1!} + \frac{(it)^2}{2!} + \cdots\cdots\cdots - 1\right] + \frac{\lambda(\lambda+1)}{2}\alpha^2 \left[1 + \frac{it}{1!} + \frac{(it)^2}{2!} + \cdots\cdots\cdots - 1\right]^2 + \cdots\cdots$$

$$\therefore\ d_X(t) = 1 + \lambda\alpha \left[it + \frac{(it)^2}{2!} + \frac{(it)^3}{3!} + \cdots\cdots\right] + \frac{\lambda(\lambda+1)}{2!}\alpha^2 \left[it + \frac{(it)^2}{2!} + \cdots\cdots\right]$$

$\therefore\ \mu_1^1 =$ The coefficient of $\frac{(it)}{1!}$ in $d_X(t) = \alpha\lambda$

$\mu_2^1 =$ The coefficeint of $\frac{(it)^2}{2!}$ in $d_X(t) = \lambda\,\alpha + \alpha^2\,\lambda\,(\lambda+1)$

$$= \lambda\alpha + \lambda^2\alpha^2 + \lambda\alpha^2$$

$$\therefore\ \mu_2 = \mu_2^1 - \left(\mu_1^1\right)^2$$

$$= \lambda\alpha + \lambda^2\alpha^2 + \lambda\alpha^2 - \lambda^2\alpha^2$$

$$= \lambda\alpha + \lambda\alpha^2$$

$$= \lambda\,\alpha\,(1+\alpha)$$

$\therefore$ Mean $= \alpha\lambda$, Variance $= \lambda\alpha(1+\alpha)$.

11.11 Exercise:

1. Define Negative Binomial Distribution and find its mean and variance.
2. Obtain M.G.F. of Negative Binomial Distribution and hence find its mean and variance.
3. Obtain c.f. of Negative Binomial Distribution and hence find its mean and variance.
4. Obtain P.G.F. of Negative Binomial Distribution and hence find its mean and variance.

5. Obtain cumulant generating function of Negative Binomial Distribution and hence find its mean and variance.

6. Show that Poisson Distribution is a limiting case of Negative Binomial Distribution.

7. If X is Negative Binomial Variate with p.m.f.

$$p(x) = \binom{x+r-1}{x} p^r q^x \quad ; \qquad x = 0,1,2,\cdots\cdots$$

then show that the recurrence relation between central moments is $\mu_{k+1} = q\left[\frac{d\mu_k}{dq} + \frac{rk}{p^2}\mu_{k-1}\right]$.

8. Deduce the moments of Negative Binomial Distribution from those of Binomial Distribution.

9. An item is produced in large numbers. The machine is known to produce 5% defectives. A quality control inspector is examining the items by taking them at random. What is the probability that at least 4 items are to be examined in order to get 2 defectives?

10. Find the probability that a person tossing 3 coins will get either all heads or all tails, for the second time on the fifth toss.

11.12 Answers:

9. $$p(x \geq 4) = \sum_{x=4}^{\infty} \binom{x-r}{2-1} (0\cdot05)^2 (0\cdot95)^{x-2} = 0.9928$$

10. $$p(X = 2) = \binom{4}{1}\left(\frac{1}{4}\right)^2 \left(\frac{3}{4}\right)^3 = 0\cdot103$$

Lesson - 12

GEOMETRIC DISTRIBUTION

Objective:

After study this lesson the students are expected to have clear comprehension of the theory and practical utility about the concepts of mean, variance, moments, generating functions and properties of Geometric Distribution.

Structure of The Lesson:

12.1 Introduction:

Suppose we have a series of independent trials or repetitions and on each trial or repetition the probability of success "p" remains fixed. In Geometric Distribution X can be defind as the number of failures before the first success. It can also be defined as number of trials required for getting the first success. The probability that there are X failures prceding the first success is given by $q^x p$. Some of the situations where the Geometric Distribution are as follows:

1. Suppose a man is hitting the torget. Number of failures before hitting the target.
2. Number of balls required for a cricket bowler to make the batsman out on the assumption of probability that the bats man will be out for any ball is same.

12.2 Defination:

A random variable X is said to have a Geometric Distribution if it assumes only non - negative values and its p.m.f. is given by

$$p(x) = \begin{cases} q^x p & ; \quad x = 0,1,2,\cdots\cdots \\ 0 & \text{otherwise} ; \quad p+q=1 \end{cases}$$

Note: (i) Since the various probabilities for $x = 0,1,2,\cdots\cdots$ are the various terms of Geometric progression, hene the distribution named as Geometric Distribution.

(ii) Total Probability $= \sum_{x=0}^{\infty} p(x)$

$$= \sum_{x=0}^{\infty} q^x p$$

$$= p\left(1+q+q^2+\cdots\cdots\right)$$

$$= p(1-q)^{-1}$$

$$= \frac{p}{1-q}$$

$$= \frac{p}{q}$$

$$=1$$

Hence total probability is one.

12.3 Derivation of Geometric Distribution:

If X denotes the number of failures before the first success then X can take any one of the values $0,1,2,\cdots\cdots$ if the probability of success is p then probability of failure will be $q=1-p$.

Now $p(X=0)$ = Probability of zero failures

= Probability of success in the first trial

$= p$

$p(X=1) = p(\text{one failure})$

$= p$ {failure in the first trial and success in the next trial}

$= q\,p$

$p(x=2) = p$ [two failures preceeding first success]

$= q^2 p$

Proceding in the similar manner, we get

$p(X-x) = q^x p \;;\; x = 0,1,2,\cdots\cdots$

It is a Geometric Progression with common ratio q - hene the distribution is Geometric Distribution.

12.4 Moments of Geometric Distribution:

The p.m.f. of Geometric Distribution is given by

$p(x) = q^x p \quad ; \; x = 0,1,2,\cdots\cdots$

By defination $\mu_1^1 = E(X) = \sum_{x=0}^{\infty} x \; p(x)$

$$= \sum_{x=0}^{\infty} x \cdot q^x \cdot p$$

$$= p\,q \sum_{x=1}^{\infty} x\, q^{x-1}$$

$$= p\,q\left[1 + 2q + 3q^2 + \cdots\cdots\right]$$

$$= p\,q\,(1-q)^{-2}$$

$$= \frac{pq}{(1-q)^2}$$

$$= \frac{pq}{p^2}$$

$$= \frac{q}{p}$$

$$\therefore \quad \text{Mean} = \mu_1^1 = q/p$$

$$\mu_2^1 = E\left(X^2\right)$$

$$= E\left[X(X-1)+X\right]$$

$$= E\left[X(X-1)\right] + E(X)$$

Now $E\left[X(X-1)\right] = \sum_{x=2}^{\infty} x(x-1)p(x)$

$$= \sum_{x=2}^{\infty} x(x-1)q^x p$$

$$= pq^2 \sum_{x=2}^{\infty} x(x-1)q^{x-2}$$

$$= pq^2 \left[2 + 3 \text{ x } 2q + 4x3q^2 + 5x4q^3 + \cdots\cdots\right]$$

$$= 2pq^2 \left[1 + 3q + 6q^2 + 10q^3 + \cdots\cdots\right]$$

$$= 2pq^2 (1-q)^{-3}$$

$$= \frac{2pq^2}{(1-q)^3}$$

$$= \frac{2pq^2}{p^3}$$

$$= \frac{2q^2}{p^2}$$

$$\therefore \quad \mu_2^1 = \frac{2q^2}{p^2} + q/p \qquad \because E(X) = q/p$$

$$\mu_2 = \mu_2^1 - \left(\mu_1^1\right)^2$$

$$= \frac{2q^2}{p^2} + q/p - \frac{q^2}{p^2}$$

$$= \frac{q^2}{p^2} + q/p$$

$$= \frac{q}{p}\left(\frac{q}{p} + 1\right)$$

$$= \frac{q}{p}\left(\frac{q+p}{p}\right)$$

$$= \frac{q}{p^2} \qquad \because q + p = 1$$

Hence variance $\mu_2 = q/p^2$

$$\mu_3^1 = E\left(X^3\right)$$

$$= E\left[X(X-1)(X-2) + 3X(X-1) + X\right]$$

$$= E\left[X(X-1)(X-2)\right] + 3E\left[X(X-1)\right] + E(X) \qquad (12.4.1)$$

Now $E\left[X(X-1)(X-2)\right] = \sum_{x=0}^{\infty} x(x-1)(x-2)\, q^x\, p$

$$= 3!\,pq^3 \sum_{x=3}^{\infty} x \frac{(x-1)(x-2)}{3!} q^{x-3}$$

$$= 6pq^3 \left(1+4q+10q^2 + 20q^3 + \cdots\cdots\right)$$

$$= 6p\,q^3\,(1-q)^{-4}$$

$$= \frac{6pq^3}{(1-q)^4}$$

$$= \frac{6pq^3}{p^4}$$

$$= \frac{6q^3}{p^3} \qquad (12.4.2)$$

We know that $E[X(X-1)] = 2q^2/p^2$ and $E(X) = q/p$ (12.4.3)

Substituting (12.4.2) and (12.4.3) in (12.4.1) we get

$$\mu_3^1 = \frac{6q^3}{p^3} + \frac{3X\,2q^2}{p^2} + q/p$$

$$= \frac{q}{p}\left(\frac{6q^2}{p^2} + \frac{6q}{p} + 1\right)$$

Similarly $\mu_4^1 = E\left(X^4\right)$

$$= E\left[X(X-1)(X-2)(X-3)\ \ (X-1)(X-2) + 7X(X-1) + X\right]$$

$$= \frac{24q^4}{p^4} + \frac{36q^3}{p^3} + \frac{14q^2}{p^2} + q/p \qquad \text{on simplification}$$

12.5 Moment Generating Function:

The p.m.f. of G.D. is given by

$$p(x) = q^x p \quad ; \quad x = 0, 1, 2, \cdots\cdots$$

The M.G.F. of G.D. is given by

$$M_X(t) = E\left(e^{tX}\right)$$

$$= \sum_{x=0}^{\infty} e^{tx}\, p(x)$$

$$= \sum_{x=0}^{\infty} e^{tx} q^x p$$

$$= p \sum_{x=0}^{\infty} \left(q e^t\right)^x$$

$$= p\left(1 - q e^t\right)^{-1}$$

$$= \frac{p}{\left(1 - q e^t\right)}$$

$$\therefore \quad M_X(t) = \frac{p}{\left(1 - q e^t\right)}$$

Mean and variance using M.G.F.

Mean $\mu_1^1 = \left[\dfrac{d M_X(t)}{dt}\right]_{t=0}$

$$= \left[\frac{d}{dt}\, p\left(1 - q e^t\right)^{-1}\right]_{t=0}$$

$$= p\left[(-1)\left(1 - q e^t\right)^{-2}\left(-q e^t\right)\right]_{t=0}$$

$$= p\,q\,(1-q)^{-2}$$

$$= \frac{pq}{p^2}$$

$$= q/p \qquad\qquad \because 1-q=p$$

$$\therefore \text{ Mean } \mu_1^1 = q/p$$

$$\mu_2^1 = \left[\frac{d^2 M_X(t)}{dt^2}\right]_{t=0}$$

$$= \left[\frac{d}{dt}\, pq\left(1-qe^t\right)^{-2} e^t\right]_{t=0}$$

$$= pq\left[\left(1-qe^t\right)^{-2} e^t - 2\left(1-qe^t\right)^{-3}\left(-qe^t\right) e^t\right]_{t=0}$$

$$= p\,q\left[\frac{1}{p^2} + \frac{2q}{p^3}\right]$$

$$= \frac{q}{p} + \frac{2q^2}{p^2}$$

$$\therefore \text{ variance } \quad \mu_2 = \mu_2^1 - \left(\mu_1^1\right)^2$$

$$= \frac{q}{p} + \frac{2q^2}{p^2} - \frac{q^2}{p^2}$$

$$= \frac{q}{p} + \frac{q^2}{p^2}$$

$$= \frac{q}{p}\left(1+\frac{q}{p}\right)$$

$$= \frac{q}{p}\left(\frac{p+q}{p}\right)$$

$$= \frac{q}{p^2} \qquad \because\ p+q=1$$

$$\therefore\ \text{variance} = q/p^2$$

12.6 Characteristic Function:

If X follows G.D. then its c.f. is given by

$$d_X(t) = E\left[e^{itX}\right]$$

$$= \sum_{x=0}^{\infty} e^{itX}\ p(x)$$

$$= \sum_{x=0}^{\infty} e^{itX}\ q^x\ p$$

$$= p\sum_{x=0}^{\infty}\left(q\,e^{it}\right)^x$$

$$= p\left(1-q\,e^{it}\right)^{-1}$$

$$\therefore\ d_X(t) = \frac{p}{\left(1-q e^{it}\right)}$$

Mean and Variance using C.F.

We know that $\mu_1^1 = (-i)\left[\frac{\partial\, d_X(t)}{\partial t}\right]_{t=0}$

$$= (-i)\left[\frac{\partial}{\partial t}\, p\left(1-q\,e^{it}\right)^{-1}\right]_{t=0}$$

$$= (-i)\, p\left[(-1)\left(1-q\,e^{it}\right)^{-2}\left(-iqe^{it}\right)\right]_{t=0}$$

$$= (-i)\, pq\, i\left[\left(1-qe^{it}\right)^{-2} e^{it}\right]_{t=0}$$

$$= -i^2\, p\, q\, (1-q)^{-2} \qquad \because i^2 = -1$$

$$= \frac{(1)\, pq}{p^2}$$

$$= q/p$$

$$\therefore \text{ Mean } \mu_1^1 = q/p$$

$$\mu_2^1 = (-i)^2\left[\frac{\partial^2 d_X(t)}{\partial t^2}\right]_{t=0}$$

$$= (-i)^2\left[\frac{\partial}{\partial t}\, i\, pq\, e^{it}\left(1-qe^{it}\right)^{-2}\right]_{t=0}$$

$$= (-i)^2\, ipq\left[ie^{it}(1-q)^{-2} + e^{it}\left(1-qe^{it}\right)^{-3}(-2)\left(-qe^{it}\right)i\right]_{t=0}$$

$$= (-i)^2\, ipq\left[i(1-q)^{-2} + 2qi(1-q)^{-3}\right]$$

$$= i^4\, pq\left[\frac{1}{p^2} + \frac{2q}{p^3}\right]$$

$$= (1)\; pq\left[\frac{1}{p^2} + \frac{2q}{p^3}\right]$$

$$= \frac{pq}{p^2} + \frac{2\,p\,q^2}{p^3}$$

$$= \frac{q}{p} + \frac{2q^2}{p^2}$$

$$\therefore \quad \mu_2 = \mu_2^1 - \left(\mu_1^1\right)^2$$

$$= \frac{q}{p} + \frac{2q^2}{p^2} - \frac{q^2}{p^2}$$

$$= q/p^2$$

$$\therefore \text{ variance } \left(\mu_2\right) = q/p^2$$

12.7 Probability Generating Function:

If X follows G.D. then its P.G.F. is given by

$$P(S) = E\left(S^X\right)$$

$$= \sum_{x=0}^{\infty} S^x \, P(x)$$

$$= \sum_{x=0}^{\infty} S^x \, q^x \, P$$

$$= P \sum_{x=0}^{\infty} (q\,S)^x$$

$$= P\,(1 - qS)^{-1}$$

$$\therefore \; P(S) = \frac{P}{(1 - q\,S)}$$

Mean and Variance using P.G.F.

$$\mu_1^1 = E(X) = \left[\frac{d}{ds} p(s)\right]_{s=1}$$

$$= \left[\frac{d}{ds} p\,(1-qs)^{-1}\right]_{s=1}$$

$$= p\left[-1\,(1-qs)^{-2}(-qs)\right]_{s=1}$$

$$= p\,q\,(1-q)^{-2}$$

$$= \frac{p\,q}{p^2}$$

$$= q/p$$

$$\therefore \text{ Mean } \quad \mu_1^1 = q/p$$

$$\mu_2^1 = \left[\frac{d^2\,p(s)}{s\,s^2}\right]_{s=1}$$

$$= pq\left[\frac{d}{ds}\left\{s(1-qs)^{-2}\right\}\right]_{s=1}$$

$$= pq\left[(1-qs)^{-2}1 - 2(1-qs)^{-3}(-qs)\,s\right]_{s=1}$$

$$= pq\left[(1-q)^{-2} + 2q(1-q)^{-3}\right]$$

$$= p\,q\left[\frac{1}{p^2} + \frac{2q}{p^3}\right]$$

$$= \frac{pq}{p^2} + \frac{2q^2\,p}{p^3}$$

$$= \frac{q}{p} + \frac{2q^2}{p^2}$$

$$\therefore \quad \mu_2 = \mu_2^1 - \left(\mu_1^1\right)^2$$

$$= \frac{q}{p} + \frac{2q^2}{p^2} - \frac{q^2}{p^2}$$

$$= q/p^2$$

$$\therefore \quad \text{variance} \quad V(X) = \mu_2 = q/p^2$$

12.8 Recurrence Relation For The Moments of Geometric Distribution:

The p.m.f. of G.D. is given by

$$p(x) = q^x \, p \qquad ; \; x = 0, 1, 2, \cdots\cdots$$

and we know that $E(X) = q/p$

The r^{th} moment about mean is given by

$$\mu_r = E\left[\{X - E(X)\}^r\right]$$

$$= \sum_{x=0}^{\infty} \left(x - q/p\right)^r p(x)$$

$$= \sum_{x=0}^{\infty} \left(x - q/p\right)^r q^x \, p$$

Differentiating w.r.t. q we get

$$\frac{d\mu_r}{dq} = \sum_{x=0}^{\infty} \left[r\left(x - q/p\right)^{r-1} \frac{d}{dq}\left(x - q/p\right) p q^x + \left(x - q/p\right)^r \left\{ p x q^{x-1} + q^x \frac{dp}{dq} \right\} \right]$$

$$= \sum_{x=0}^{\infty}\left[r\left(x - q/p\right)^{r-1}\left(\frac{-1}{p^2}\right) p\,q^x + \left(x - \frac{q}{p}\right)^r \left(x p q^{x-1} - q^x\right)\right]$$

$$= \frac{-r}{p^2}\sum_{x=0}^{\infty}\left(x - q/p\right)^{r-1} p(x) + \sum_{x=0}^{\infty}\left(x - q/p\right)^r p q^x \left(\frac{x}{q} - \frac{1}{p}\right)$$

$$= \frac{-r}{p^2}\sum_{x=0}^{\infty}\left(x - q/p\right)^{r-1} p(x) + \frac{1}{q}\sum_{x=0}^{\infty}\left(x - q/p\right)^r p q^x \left(x - q/p\right)$$

$$= \frac{-r}{p^2}\mu_{r-1} + \frac{1}{q}\sum_{x=0}^{\infty}\left(x - q/p\right)^{r+1} p(x)$$

$$= \frac{-r}{p^2}\mu_{r-1} + \frac{1}{q}\mu_{r+1}$$

$$\Rightarrow \frac{1}{q}\mu_{r+1} = \frac{d\,\mu_r}{d\,q} + \frac{r}{p^2}\mu_{r+1}$$

$$\Rightarrow \mu_{r+1} = q\left[\frac{d\mu_r}{dq} + \frac{r}{p^2}\mu_{r-1}\right]$$

Hence the recurrence relation between moments is $\mu_{r+1} = q\left[\frac{d\mu_r}{dq} + \frac{r}{p^2}\mu_{r-1}\right]$

If $r = 1$, then $\mu_2 = q\left[\frac{d\,\mu_1}{dq} + \frac{r}{p^2}\mu_0\right] = q/p^2$ $\quad \because \mu_0 = 1,\ \mu_1 = 0$.

If $r = 2$, then $\mu_3 = q\left[\frac{d\mu_2}{dq} + \frac{2}{p^2}\mu_1\right] = q\left[\frac{d}{dq}\left(q/p^2\right)\right] = q(1+q)/p^3$

If $r = 3$, then $\mu_4 = q\left[\frac{d\mu_3}{d\,q} + \frac{3}{p^2}\mu_2\right]$

$$= q\left[\frac{d}{dq}\left\{\frac{q(1+q)}{p^3}\right\} + \frac{3}{p^2}\left(q/p^2\right)\right]$$

$$= \frac{q}{p^4}\left[p^2 + 3q(3)\right]$$

$$= \frac{q}{p^4}\left(p^2 + pq\right)$$

$$\therefore \mu_4 = q\left(p^2 + pq\right)/p^4$$

12.9 Lack of Memory Property of Geometric Distribution:

The special property of Geometric Distribution is lack of memory. It means that the probability of additional number of failures before the first success is equal to "t", given that the number of failures preceding the first success is greater than or equal to "k" is the same as the unconditional probability of the number of failures before the first success is "t".

By definition $p(X = t) = pq^t$; $t = 0,1,2,\cdots\cdots$

We have to show that $p\left[X = k+t / X \geq k\right] = p(X = t)$

Consider,

$$p\left[X = k+t / X \geq k\right] = \frac{p[X = k + t \cap X \geq k]}{p(X \geq k)}$$

$$= \frac{p[X = k + t]}{p(X \geq k)} \qquad (12.9.1)$$

Now $p(X \geq k) = \sum_{x=k}^{\infty} p\,q^x$

$$= p\left[q^k + q^{k+1} + q^{k+r} + \cdots\cdots\cdots\right]$$

$$= p\,q^k\left(1 + q + q^2 + \cdots\cdots\right)$$

$$= p\,q^k\,(1-q)^{-1}$$

$$= \frac{p\,q^k}{p} = q^k \qquad (12.9.2)$$

Substituting (12.9.2) in (12.9.1) we get

$$p\left(X = k+t \big/ X \geq k\right) = \frac{p\,q^{k+t}}{q^k}$$

$$= p\,q^t$$

$$= p(X = t)$$

Hence $p\left[X = k + t \big/ X \geq k\right] = P(X = t)$

Hence G.D. lacks memory.

12.10 Recurrence Relation For The Probabilities of G.D.:

The p.m.f. of G.D. is given by

$$p(x) = q^x\,p \quad ; \quad x = 0,1,2,\cdots\cdots\cdots$$

$$\Rightarrow p(x+1) = q^{x+1}\,p$$

$$\therefore \frac{p(x+1)}{p(x)} = \frac{q^{x+1}p}{q^x\,p} = q$$

$$\therefore p(x+1) = q\,p(x) \quad ; \quad x = 0,1,2,\cdots\cdots\cdots$$

For apply this recurrence relation we need $P(0) = p$

The other probabilities $P(1), P(2), \cdots\cdots\cdots$ can be easily obtained by using recurrence relation.

12.11 Additive Property of G.D.:

If X_1 and X_2 are two independent Geometric variables with parameter p.

i.e. $X_1 \sim G \cdot D \cdot (p)$ and $X_2 \sim G \cdot D \cdot (p)$

The M.G.F. of X_1 & X_2 are given by

$$M_{X_1}(t) = p\left(1 - q e^t\right)^{-1} \qquad M_{X_2}(t) = p\left(1 - q e^t\right)^{-1}$$

$$\therefore \quad M_{X_1+X_2}(t) = M_{X_1}(t) \cdot M_{X_2}(t)$$

$$M_{X_1+X_2}(t) = p^2 \left(1 - q\, e^t\right)^{-2}$$

This is M.G.F. of Negative Binomial Distribution with $r = 2$.

Hence by uniqueness theorem of M.G.F.'s $X_1 + X_2$ follows Negative Binomial Distribution.

12.12 Workedout Examples:

Example 1: If the probability that a target is destroyed on any shot is 0.5. What is the probability that is would be destroyed on 6^{th} attempt?

Solution: Given that $p = 0 \cdot 5$

$$q = 1 - p = 0 \cdot 5$$

The probability that target would be destroyed on 6^{th} attempt

$$= P(X = 5) = q^5 P$$

$$= (0 \cdot 5)^5 \, (0 \cdot 5)$$

$$= (0 \cdot 5)^6$$

Example 2: An unbiased dice is tossed until the occurrence of a six. Find probability that the number of trials required in all is more than 6.

Solution: Let x be the number of trials required in all

$\therefore P(X = x) =$ P [$(x-1)$ trials does not give six and x^{th} trial gives 6]

$$= \left(\frac{5}{6}\right)^{x-1}\left(\frac{1}{6}\right) \quad ; \quad x = 1, 2, \cdots\cdots$$

$$\therefore\ P(X > 6) = P(X \geq 7) = \sum_{x=7}^{\infty} \left(\frac{5}{6}\right)^{x-1}\left(\frac{1}{6}\right)$$

$$= \frac{1}{6}\left[\left(\frac{5}{6}\right)^{6} + \left(\frac{5}{6}\right)^{7} + \cdots\cdots\cdots\right]$$

$$= \left(\frac{1}{6}\right)\left(\frac{5}{6}\right)^{6}\left[1 + \frac{5}{6} + \left(\frac{5}{6}\right)^{2} + \cdots\cdots\cdots\right]$$

$$= \left(\frac{1}{6}\right)\left(\frac{5}{6}\right)^{6}\left[1 - \frac{5}{6}\right]^{-1}$$

$$= \frac{\left(\frac{1}{6}\right)\left(\frac{5}{6}\right)^{6}}{\left(\frac{1}{6}\right)}$$

$$\therefore\ P(X > 6) = \left(\frac{5}{6}\right)^{6}$$

Example 3: Assume that a student is waiting for school bus. The probability that exactly X buses will pass the student before the school bus arrives is given by

$$P(X = x) = \frac{1}{3}\left(\frac{2}{3}\right)^{x} \quad ; \quad x = 0, 1, 2, \cdots\cdots$$

(i) What is the probability that five buses will pass the student before the school bus arrives ?

(ii) What is the probability tht additional five buses will pass before the school bus arrives given that more than three have already passed?

Solution: Given that $P(X = x) = \frac{1}{3}\left(\frac{2}{3}\right)^x \quad ; \quad x = 0, 1, 2, \cdots\cdots$

(i) $\therefore \quad P(X = 5) = \frac{1}{3}\left(\frac{2}{3}\right)^5 = \frac{32}{729}$

(ii) $P\left(x \geq 8 / x \geq 3\right) = \frac{P(X \geq 8)}{P(X \geq 3)} = \frac{q^8}{q^3} = q^5 = \left(\frac{2}{3}\right)^5$

$$\therefore \quad P\left(X = 8 / x \geq 3\right) = P\left(X \geq 8 / X \geq 3\right) - P\left(X \geq 9 / X \geq 3\right)$$

$$= q^5 - q^6$$

$$= q^5 (1 - q)$$

$$= q^5 \, p$$

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

$$= \frac{32}{729}$$

12.13 Exercise:

1. Define Geometric Distribution and state where the distribution is useful.
2. Find the M.G.F. of Geometric Distribution and hence find its mean and variance.
3. Find the characteristic function of Geometric Distribution and hence find its mean and variance.
4. Find the probability generating function of Geometric Distribution and hence find its mean and variance.
5. Explain memory less property of Geometric Distribution.
6. Obtain the recurrence relation between moments of Geometric Distribution.

7. If the probability is 0.75 that an applicant for a driver's licence will pass the road test on ny given try, what is the probability that an applicant will finally pass the test on the fourth try.

8. If the probability that a target is destroyed on any one shot is 0.5. What is te probability that it would be destroyed on 8^{th} attempt?

9. A die is cost until 6 appeons. What is the probability that it must be cast more than five times.

12.14 Answers:

7. 0.1055

8. $(0 \cdot 5)^8$

9. $\left(\frac{5}{6}\right)^5$

Lesson - 13

HYPERGEOMETRIC DISTRIBUTION

Objective:

After studying the lesson the students are expected to have clear comprehension of the theory and practicl utility about the concepts of mean, variance, limiting case, recurrence relation of Hypergeometric Distribution.

Structure of The Lesson:

This lesson consists of the following sections as detailed below:

13.1 Introduction:

Hypergeometric Distribution is used when sampling conducted without replacement from a finite population. In this distribution, the rpobability of an outcome in any trial is not same as in any other trial. Consider a box with N balls, M of which are white and N - m are red. Suppose that we draw a sample of n balls at random by without replacement from the box. then the probability of getting x white balls out of n balls $(x < n)$ is $\binom{M}{x}\binom{N-M}{n-x}\Big/\binom{N}{n}$.

Since x white balls can be drawn from "m" white balls in $\binom{M}{x}$ ways and out of the remaining $N-m$ red balls, $(n-x)$ balls can be choosen in $\binom{N-M}{n-x}$ ways.

i.e. favourable number of cases $=\binom{M}{x}\binom{N-M}{n-x}$

total possible number of cases $=\binom{N}{n}$

$\therefore$ probability of getting x white balls $= P(X=x) = \dfrac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$

13.2 Defination:

A discrete random variable X is said to follow the Hypergeometric Distribution with parameter N, m and n if it assumes only non - negative values and its probability mass function is given by

$$P(X=x) = \begin{cases} \dfrac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}} \; ; & x = 0,1,2,\cdots\cdots,\min(n, M) \\ 0 & \text{otherwise} \end{cases}$$

Where N is a positive integer, m is a positive integer not exceeding N and n is a positive integer that is at most N.

Note: Total probability $= \sum_{x=0}^{n} P(x)$

$$= \sum_{x=0}^{n} \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$$

$$= \frac{\binom{N}{n}}{\binom{N}{n}}$$

$$= 1$$

Here N, m, n are parameters of Hypergeometric Distribution.

13.3 Mean and Variance of Hypergeometric Distribution:

The p.m.f. of Hypergeometric Distribution (H.G.D.) is given by

$$P(x) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}} \quad ; \quad x = 0, 1, 2, \cdots\cdots, n$$

By defination $E(X) = \sum_{x=0}^{n} x \cdot \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$

$$= \frac{1}{\binom{N}{n}} \sum_{x=0}^{n} x \frac{M}{x} \binom{M-1}{x-1}\binom{N-M}{n-x}$$

$$= \frac{M}{\binom{N}{n}} \sum_{x=1}^{n} \binom{M-1}{x-1}\binom{N-M}{n-x}$$

$$= \frac{M}{\binom{N}{n}} \binom{N-1}{n-1} \qquad \because \sum_{x=1}^{n} \binom{M-1}{x-1}\binom{N-M}{n-x} \binom{N-1}{n-1}$$

$$= \frac{M}{\frac{N}{n}\binom{N-1}{n-1}} \cdot \binom{N-1}{n-1}$$

$$\therefore \quad E(X) = \frac{nM}{N}$$

Hence mean of HGD is $\frac{nM}{N}$

$$E(X^2) = E[X(X-1) + X]$$

$$= E[X(X-1)] + E(X) \qquad (13.3.1)$$

Consider $E[X(X-1)] = \sum_{x=0}^{n} x(x-1)\,P(x)$

$$= \sum_{x=2}^{n} x(x-1)\,\frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$$

$$= \frac{1}{\binom{N}{n}} \sum_{x=2}^{n} x(x-1)\,\frac{M}{x}\cdot\frac{M-1}{x-1}\cdot\binom{M-2}{x-2}\binom{N-M}{n-x}$$

$$= \frac{M(M-1)}{\binom{N}{n}}\binom{N-2}{n-2}$$

$$= \frac{M(M-1)}{\frac{N}{n}\cdot\frac{N-1}{n-1}\binom{N-2}{n-2}}\binom{N-2}{n-2}$$

$$= \frac{n(n-1)\,M(M-1)}{N(N-1)} \qquad (13.3.2)$$

Substituting (13.3.2) in (13.3.1) we get

$$E\left(X^2\right) = \frac{n(n-1)\,M(M-1)}{N(N-1)} + \frac{nM}{N} \qquad \because\ E(X) = \frac{nM}{N}$$

$$\therefore\ V(X) = E\left(X^2\right) - [E(X)]^2$$

$$= \frac{n(n-1)\,M(M-1)}{N\,(N-1)} + \frac{nM}{N} - \frac{n^2M^2}{N^2}$$

$$= \frac{nM}{N}\left[\frac{(n-1)(M-1)}{N-1} + 1 - \frac{nM}{N}\right]$$

$$= \frac{nM}{N}\left[\frac{N(n-1)(M-1)+N(N-1)-(N-1)nM}{N(N-1)}\right]$$

$$= \frac{nM}{N^2(N-1)}\left[NMn - NM - Nn + N - NMn + Mn + N^2 - N\right]$$

$$= \frac{nM}{N^2(N-1)}\left[N(N-M) - n(N-M)\right]$$

$$= \frac{(N-n)(N-M)\,nM}{N^2(N-1)}$$

$\therefore$ variance $V(X) = \dfrac{nM(N-n)(N-M)}{N^2(N-1)}$

13.4 Hypergeometric Distribution Tends To Binomial Distribution:

Hypergeometric Distribution tends to Binomial Distribution, under the following conditions.

(i) $N \to \infty$

(ii) $\dfrac{M}{N} \to p$

The p.m.f. of H.G.D. is given by

$$P(X=x) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$$

$$= \frac{M!}{(M-x)!(x!)} \cdot \frac{(N-M)!}{(N-M-n+x)!(n-x)!} \cdot \frac{(N-n)!\,n!}{N!}$$

$$= \frac{M(M-1)\cdots\cdots(M-x+1)(M-x)!}{x!(M-x!)} \frac{(N-M)(N-M-1)\cdots\cdots(N-M-n+x+1)(N-M-n+x)!}{(n-x)!(N-M-n+x)!}$$

$$\times \frac{n!(N-n)!}{N(N-1)(N-2)\cdots\cdots(N-n+1)(N-n)!}$$

$$= \frac{n!}{x!(n-x)!} \; \frac{M(M-1)\cdots(M-+1)(N-M)(N-M-1)\cdots\cdots(N-M-n+x+1)}{N(N-1)\cdots\cdots(N-n+1)}$$

$$= \frac{\binom{n}{x}\left(\frac{M}{n}\right)\left(\frac{M-1}{N}\right)\cdots\cdots\left(\frac{M-x+1}{N}\right)\times\left(\frac{N-M}{N}\right)\left(\frac{N-M-1}{N}\right)\cdots\cdots\left(\frac{N-M-n+x+1}{N}\right)}{\left(\frac{N}{N}\right)\left(\frac{N-1}{N}\right)\cdots\cdots\cdots\left(\frac{N-n+1}{N}\right)}$$

Proceeding to the limit as $N \to \infty$ and putting $\frac{M}{N} = p$ we get

$$\lim_{N\to\infty} P(x) = \binom{n}{x} \underbrace{p\cdot q\cdots\cdots p}_{x \text{ times}} \cdot \underbrace{(1-p)(1-p)\cdots\cdots(1-p)}_{(n-x) \text{ times}}$$

$$= \binom{n}{x} p^x q^{n-x} \qquad \because \; q = 1-p$$

Hence HGD tends Binomial Distribution.

13.5 Recurrence Relation Between Probabilities of HGD:

By definition $$p(x) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}} \quad ; \quad x = 0,1,2,\cdots\cdots,n$$

$$\therefore \quad p(x+1) = \frac{\binom{M}{x+1}\binom{N-M}{n-x-1}}{\binom{N}{n}}$$

$$\therefore \quad \frac{p(x+1)}{p(x)} = \frac{\binom{M}{x+1}\binom{N-M}{n-x-1}}{\binom{N}{n}} \cdot \frac{\binom{N}{n}}{\binom{M}{x}\binom{N-M}{n-x}}$$

$$= \frac{M!}{(x+1)!\,(M-x-1)!} \cdot \frac{(N-M)!}{(n-x-1)!\,(N-M-n+x+1)!}$$

$$\times \frac{x!(M-x)!\,(n-x)!\,(N-M-n+x)!}{M!\,(N-M)!}$$

$$= \frac{x!\,(M-x)(M-x-1)!\,(N-M-n+x)!\,(n-x)(n-x-1)!}{(x+1)x!\,(M-x-1)!\,(n-x-1)!\,(N-M-n+x+1)(N-M-n+x)!}$$

$$= \frac{(M-x)\,(n-x)}{(x+1)\,(N-M-n+x+1)}$$

$$\therefore \; p(x+1) = \frac{(M-x)\,(n-x)}{(x+1)\,(N-M-n+x+1)} \cdot p(x) \;\; ; \;\; x = 0,1,2,\cdots\cdots$$

which is the required recurrence relation.

13.6 Workedout Eamples:

Example 1: If p.m.f. of HGD is $p(x) = \frac{{}^{M}C_{x}\; {}^{n}C_{r-x}}{{}^{(M+n)}C_{r}}$; $x = 0,1,2,\cdots\cdots,r$ then find its mean and variane.

Solution: $\text{Mean} = E(X) = \sum_{x=0}^{r} x \cdot p(x)$

$$= \sum_{x=0}^{r} x \; \frac{\binom{M}{x}\binom{n}{r-x}}{\binom{M+n}{r}}$$

$$= \frac{M}{\binom{M+n}{r}} \sum_{x=0}^{r} \binom{M-1}{x-1}\binom{n}{r-x}$$

$$= \frac{M}{\binom{M+n}{r}} \cdot \binom{M+n-1}{r-1} \qquad \because \sum_{x=0}^{r} \binom{M-1}{r-1}\binom{n}{r-x} = \binom{M+n-1}{r-1}$$

$$= \frac{M\binom{M+n-1}{r-1}}{\left(\frac{M+n}{r}\right)\binom{M+n-1}{r-1}}$$

$$\therefore \quad \text{Mean} \;=\; \frac{Mr}{M+n}$$

$$E\left(X^2\right) = E\left[X(X-1)+X\right]$$

$$= E\left[X(X-1)\right] + E(X) \cdot \quad \cdot \quad \cdot \quad (1)$$

Now

$$E\left[X(X-1)\right] = \sum_{x=2}^{r} x(x-1)\frac{\binom{M}{x}\binom{n}{r-x}}{\binom{M+n}{r}}$$

$$= \sum_{x=2}^{r} x(x-1)\,\frac{M(M-1)}{x(x-1)} \cdot \frac{\binom{M-2}{x-2}\binom{n}{r-x}}{\binom{M+n}{r}}$$

$$= \frac{M(M-1)}{\binom{M+n}{r}} \sum_{x=2}^{r} \binom{M-2}{x-2}\binom{n}{r-x}$$

$$= \frac{M(M-1)}{\binom{M+n}{r}} \cdot \binom{M+n-2}{r-2}$$

$$= \frac{M(M-1)\binom{M+n-2}{r-2}}{\left(\frac{M+n}{r}\right)\left(\frac{M+n-1}{r-1}\right)\binom{M+n-2}{r-2}}$$

$$= \frac{r(r-1)\,M(M-1)}{(M+n)\,(M+n-1)} \quad \cdot \cdot \cdot (2)$$

Substituting (2) in (1) we get

$$E\left(X^2\right) = \frac{r(r-1)(M)(M-1)}{(M+n)(M+n-1)} + \frac{Mr}{M+n} \qquad \because\ E(X) = \frac{Mr}{M+n}$$

$$\therefore \quad \text{variance } V(X) = E\left(X^2\right) - [E(X)]^2$$

$$= \frac{M(M-1)\,r(r-1)}{(M+n)\,(M+n-1)} + \frac{Mr}{M+n} - \frac{M^2 r^2}{(M+n)^2}$$

$$= \frac{Mr}{(M+n)}\left[\frac{(M+n)(M-1)(r-1)+(M+n)(M+n-1)-Mr(M+n-1)}{(M+n)\ (M+n-1)}\right]$$

$$= \frac{Mr}{(M+n)^2(M+n-1)}\left[\begin{array}{r} M^2r+Mnr-Mr-nr-M^2-Mn+n+n+M^2+ \\ 2Mn-n-n+n^2-M^2r-Mnr+Mr \end{array}\right]$$

$$= \frac{Mr}{(M+n)^2\ (M+n-1)}\left[Mn-nr+n^2\right]$$

$$\therefore \quad \text{Variance} = \frac{Mnr\,(M+n-r)}{(M+n)^2(M+n-1)}$$

Example 2: As part of an air pollution survey, an inspector decides to examine the exhavst of 6 of a company's 24 trucks. If 4 of the company's trucks emit excessive amounts of pollutants, what is the probability that none of them will be included in the inspector's sample?

Solution: $n = 6,\ N = 24,\ M = 4$.

$$\therefore \quad p(X=0) = \frac{\binom{4}{0}\binom{20}{6}}{\binom{24}{6}} = 0 \cdot 2880$$

Example 3: Among the 16 applicants for a job, ten have college degrees. If three of the applicants are randamly choosen for interviews, what are the probabilities that

(i) none has a college degree

(ii) one has a college degree

(iii) two have college degree

(ig) All three have college degree

Solution: Given that $N = 16, \ M = 10, \ n = 3$.

Using HGD

(i) $$P(X=0) = \frac{\binom{10}{0}\binom{6}{3}}{\binom{16}{3}} = \frac{20}{560} = \frac{1}{28}$$

(ii) $$P(X=1) = \frac{\binom{10}{1}\binom{6}{2}}{\binom{16}{3}} = \frac{10 \times 15}{560} = \frac{15}{56}$$

(iii) $$P(X=2) = \frac{\binom{10}{2}\binom{6}{1}}{\binom{16}{3}} = \frac{45 \times 6}{560} = \frac{27}{56}$$

(iv) $$P(X=3) = \frac{\binom{10}{3}\binom{6}{0}}{\binom{16}{3}} = \frac{120 \times 1}{560} = \frac{3}{14}$$

Example 4: What is the probability that an IRS avditor will catch only two incometax returns with illegitimate deductions, if she randomly select five returns from among 15 returns of which 9 contain illegitmate deductions?

Solution: Given that $N = 15, \ M = 9, \ n = 5, \ X = 2$.

Substituting $X = 2$ in HGD we get

$$P(X=2) = \frac{\binom{9}{2}\binom{6}{3}}{\binom{15}{5}} = \frac{36 \times 20}{3003} = \frac{240}{1001}$$

Example 5: A box consists of 15 red balls and 5 black balls. 5 balls are drawn at random what is the probability that drawn balls consists of 2 black balls.

Solution: Using HGD we can get probability

Given that $N = 20, n = 5, M = 5, X = 2$.

$$\therefore \quad P(X=2) = \frac{\binom{5}{2}\binom{15}{3}}{\binom{20}{5}}$$

$$= \frac{10 \times 455}{15504}$$

$$= 0 \cdot 29$$

Example 6: A bag contains 4 white balls and 3 green balls. 3 balls are drawn. What is the probability in that 2 are white.

Solution: $N = 4+3 = 7, \quad M = 4, \quad n = 3, \quad X = 2$.

$$\therefore \quad P(X=2) = \frac{\binom{4}{2}\binom{3}{1}}{\binom{7}{3}} = \frac{18}{35}$$

13.7 Exercise:

1. Define Hypergeometric Distribution.
2. Obtain the mean and variane of Hypergeometric Distribution.
3. Obtain the recurrence relation between probabilities of Hypergeometric Distribution.
4. Show that Binomial Distribution as a limiting case of Hypergeometric Distribution.
5. In a group of 10 people there are 5 drinkers. Find te probability distribution of the number of drinkers X, in a random sample of 6 people selected. Hence find its mean and variance.
6. A taxi cab company has 12 ambassadors and 8 fiats. If 5 of these cabs are in the shop for repairs and Ambassador is as likely to be in for repairs as a fiat. What is the probability

   (i) 3 of them are Ambassadors and 2 are Fiats.

   (ii) At least 3 of them are Ambassadors.

7. A basket consists of 5 Apples and 4 Mangoes. If 4 fruits are randomly drawn, find the probability that

   (i) no mangoes were selected

   (ii) exactly two mangoes were selected.

8. From a group of 10 boys and 6 girls, a committee of 5 students is to be formed at random

   (i) Find the probability that the committee consists of 2 boys and 3 girls.

   (ii) How many girls are expected in the committee.

9. A box contains 7 black and 4 white balls. If 5 balls are drawn, find the probability that they consists of 2 white balls.

10. A Jury of 12 members is drawn at random from a voters list of 1000 persons out of which 700 are non-graduates and 300 are graduates. What is the probability that the Jury will consists of all graduates?

    (i) Compute the probability by Hypergeometric Distribution.

    (ii) Compute the probability by Binomial Distribution.

13.8 Answers:

5. mean = 2.9991, variance = 0.6.

6. (i) 0.3973

   (ii) 0.7038

7. (i) 0.03968

   (ii) 0.47619

8. (i) 0.20604

   (ii) 2

9. $\frac{5}{11}$

10. (i) 0.000000454

    (ii) 0.00000053

Lesson - 14

NORMAL DISTRIBUTION

Object of The Lesson:

After studying this lesson the student is expected to have a clear comprehension of the theory and the practical utility about the concepts of definition, properties, importance in statistics of Normal Distribution and as limiting case of Binomial and Poisson Distribution.

Structure of The Lesson:

This lesson consists of the following sections as detailed below:

14.1 Introduction:

The Normal Distribution was introduced in 1733 by Mathematician De-Moivre, who obtained this continuous distribution as a limiting case of the binomial distribution and applied it to problems arising in the game of chance. Later Laplace and Gauss derived it independently of each other as the distribution of erross in physical measurements. Thus the normal distribution has got wide applications in the theory of statistics.

14.2 Definition:

A random variable X is said to have a Normal Distribution with parameters μ, called mean and σ^2 the variance if its density function is given by the probability law

$$f(x\,;\,\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} \cdot \exp\left[\frac{-1}{2}\left\{\frac{x-\mu}{\sigma}\right\}^2\right]$$

or $$f(x\,;\,\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{\frac{-(x-\mu)^2}{2\sigma^2}} \qquad -\infty < x < \infty,\ -\infty < \mu < \infty,\ \sigma > 0$$

Therefore, A random variable X with mean μ and variance σ^2 following the Normal Distribution is expressed as $X \sim N(\mu, \sigma^2)$. If $X \sim N(\mu, \sigma^2)$, then $Z = \frac{X-\mu}{\sigma}$ is a standard normal variate with mean 0 i.e., $E(Z) = 0$ and variance 1 i.e., $V(Z) = 1$ and is denoted by $Z \sim N(0, 1)$.

Hence the probability density function of standard normal variate Z is given by

$$\phi(Z) = \frac{1}{\sqrt{2\pi}} \cdot e^{-Z^2/2}\ , -\infty < Z < \infty$$

and the corresponding distribution function, denoted by $\Phi(Z)$ is given by

$$\Phi(Z) = P(Z \le Z) = \int_{-\infty}^{Z} \Phi(u)\,du = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{Z} e^{-u^2/2}\,du$$

Here below two important results on the distribution function $\Phi(Z)$ of standard normal variate.

Result 1: To show that $\Phi(-Z) = 1 - \Phi(Z)$

Proof: $\Phi(-Z) = P(Z \le -Z) = P(Z \ge Z)$ (By symmetry)

$$= 1 - P(Z \le Z)$$

$$= 1 - \Phi(Z)$$

Result 2: To show that $P(a \le X \le b) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$

where $X \sim N(\mu, \sigma^2)$

Proof: $P(a \le X \le b) = P\left\{\frac{a-\mu}{\sigma} \le Z \le \frac{b-\mu}{\sigma}\right\}$ $\left(\because Z = \frac{X-\mu}{\sigma}\right)$

$$= P\left[Z \le \frac{b-\mu}{\sigma}\right] - P\left[Z \le \frac{a-\mu}{\sigma}\right]$$

$$= \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$$

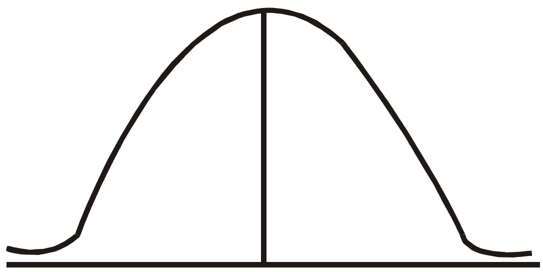
14.3 Properties of Normal Distribution:

The normal probability curve with mean μ and standard deviation σ is given by the equation

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2 / 2\sigma^2}, \quad -\infty < x < \infty$$

and has the following properties

1. The curve is bell shaped and symmetrical about the line $x = \mu$.



$x = \mu$

2. Mean, median and mode of the distribution coincide.

Mean of Normal Distribution:

$$E(X) = \int_{-\infty}^{\infty} x \cdot \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}} \, dx$$

put $\frac{x-\mu}{\sigma} = Z \Rightarrow x = \mu + \sigma Z, \qquad dx = \sigma \, dZ$

$x \quad \infty, \; Z = \infty, \; x = -\infty, Z = -\infty$.

$$E(Z) = \int_{-\infty}^{\infty} (\mu + \sigma Z) \frac{1}{\not\sigma \sqrt{2\pi}} \cdot e^{-Z^2/2} \not\sigma \, dZ$$

$$= \int_{-\infty}^{\infty} \frac{\mu}{\sqrt{2\pi}} \cdot e^{-Z^2/2} \, dZ + \int_{-\infty}^{\infty} \frac{\sigma}{\sqrt{2\pi}} Z \cdot e^{-Z^2/2} \, dZ$$

$$= \frac{\mu}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-Z^2/2} \, dZ + \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{\infty} Z \cdot e^{-Z^2/2} \, dZ$$

$$= \mu \cdot 1 + 0$$

$$= \mu$$

[$\because \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-Z^2/2} \, dZ = 1$ and the second integral $\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} Z \cdot e^{-Z^2/2} \, dZ = 0$ is an odd function].

Median of Normal Distribution:

If M is the median of the Normal Distribution we have

$$\int_{-\infty}^{M} f(x) \, dx = \frac{1}{2} \quad \cdot \quad \cdot \quad \cdot (1)$$

But $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}} \quad , \quad -\infty < x < \infty$

Substituting the value of $f(x)$ in equation (1) we get

$$\int_{-\infty}^{M} \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}} dx = 1/2$$

$$\Rightarrow \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\mu} e^{\frac{-(x-\mu)^2}{2\sigma^2}} dx + \frac{1}{\sigma\sqrt{2\pi}} \int_{\mu}^{M} e^{\frac{-(x-\mu)^2}{2\sigma^2}} dx = \frac{1}{2} \quad \cdot \quad \cdot \quad \cdot (2)$$

But $\frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\mu} e^{\frac{-(x-\mu)^2}{2\sigma^2}} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{0} e^{-z^2/2} dZ = 1/2$

$\therefore$ from (2) we get

$$\frac{1}{2} + \frac{1}{\sigma\sqrt{2\pi}} \int_{\mu}^{M} e^{\frac{-(x-\mu)^2}{2\sigma^2}} dx = 1/2$$

$$\Rightarrow \frac{1}{\sigma\sqrt{2\pi}} \int_{\mu}^{M} e^{\frac{-(x-\mu)^2}{2\sigma^2}} dx = 0$$

$$\Rightarrow \quad \mu = M$$

Hence for the Normal Distribution Mean = Median.

Mode of Normal Distribution:

Mode is the value of x for which $f(x)$ is maximum that is mode is the solution of $f'(x) = 0$ and $f''(x) < 0$. For Normal Distribution with mean μ and standard deviation σ $f(x)$ is given by

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}} \quad , \quad -\infty < x < \infty$$

Taking logarithms on both sides of above equation becomes

$$\ell og\ f(x) = \ell og_e \left(\frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}} \right)$$

$$= \ell og_e \left(\frac{1}{\sigma\sqrt{2\pi}} \right) + \ell og_e\ e^{\frac{-(x-\mu)^2}{2\sigma^2}}$$

$$= K - \frac{(x-\mu)^2}{2\sigma^2}\ \ell og_e^e \qquad \left(\because \ell og_e^e = 1,\ K = \ell og_e \frac{1}{\sigma\sqrt{2\pi}} \right)$$

$$\ell og\ f(x) = K - \frac{(x-\mu)^2}{2\sigma^2}$$ K is constant.

Differentiating w.r.t. 'x' above equation becomes

$$\frac{1}{f(x)} \cdot f'(x) = \frac{-1}{\sigma^2}(x-\mu) \Rightarrow f'(x) = \frac{-1}{\sigma^2}(x-\mu)\, f(x)$$

Also $f''(x) = \frac{-1}{\sigma^2} \left[1 \cdot f(x) + (x-\mu) f'(x) \right]$

$$= \frac{-1}{\sigma^2} \left[f(x) + (x-\mu)^2 \left(\frac{-f(x)}{\sigma^2} \right) \right] \qquad \left(\because\ f'(x) = \frac{-1}{\sigma^2}(x-\mu) f(x) \right)$$

$$= \frac{-1}{\sigma^2} \left[f(x) - \frac{(x-\mu)^2}{\sigma^2} f(x) \right]$$

$$= \frac{-f(x)}{\sigma^2} \left[1 - \frac{(x-\mu)^2}{\sigma^2} \right]$$

As $f'(x) = 0 \Rightarrow \frac{-1}{\sigma^2}(x-\mu)\, f(x) = 0$

$\Rightarrow (x - \mu) = 0$

$\Rightarrow \ x = \mu$

At the point $x = \mu$ we have from equation

$$f''(x) = \frac{-1}{\sigma^2}[f(x)]_{x=\mu} = \frac{-1}{\sigma^2} \cdot \frac{1}{\sigma\sqrt{2\pi}} < 0$$

Hence $x = \mu$ is the mode of the Normal Distribution. Therefore mean , median and mode coincide at $x = \mu$.

3. Mean deviation about mean is $\sqrt{\frac{2}{\pi}} \cdot \sigma \cong \frac{4}{5}\sigma$ (approximately)

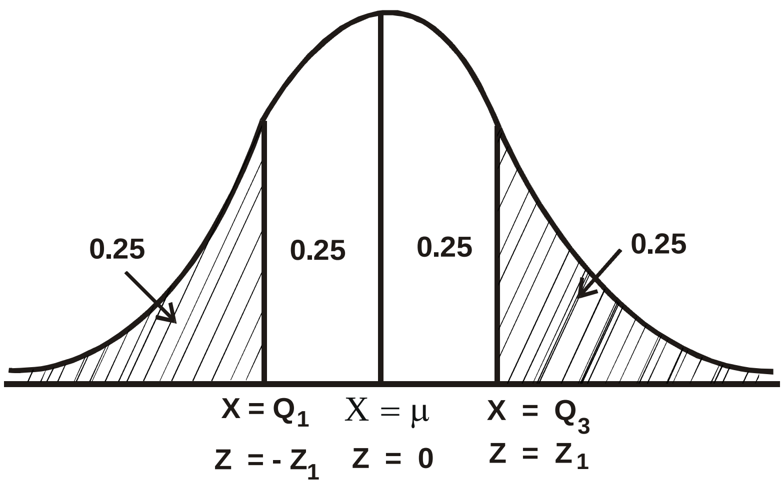
Quartile deviation $(Q \cdot D \cdot) = \frac{Q_3 - Q_1}{2} \cong \frac{2}{3}\sigma$ (approximately)

or $Q \cdot D \cdot : M \cdot D \cdot : S \cdot D \cdot :: \frac{2}{3}\sigma : \frac{4}{5}\sigma : \sigma :: \frac{2}{3} : \frac{4}{5} : 1 \Rightarrow Q \cdot D \cdot : M \cdot D \cdot : S \cdot D \cdot :: 10 : 12 : 15$

Let X be a $N(\mu, \sigma^2)$. If Q_1 and Q_3 are the first and third quartiles respectively then by definition.

$$P(X < Q_1) = 0 \cdot 25 \quad \text{and} \quad P(X > Q_3) = 0 \cdot 25$$

The points Q_1 and Q_3 are located as shown in the figure given below.



When $X = Q_3,\ Z = \frac{Q_3 - \mu}{\sigma} = Z_1\ (\text{say}) \quad \cdot \quad \cdot \quad \cdot\ (1)$

and when $X = Q_1,\ Z = \frac{Q_1 - \mu}{\sigma} = -Z_1$ (as from the figure) $\cdot \quad \cdot \quad \cdot\ (2)$

Substracting (1) & (2) we get

$$\frac{Q_3 - Q_1}{\sigma} = 2Z_1$$

The quartile deviation is given by $Q \cdot D \cdot = \frac{Q_3 - Q_1}{2} = \sigma Z_1$

From the figure

$P(0 < Z < Z_1) = 0 \cdot 25 \Rightarrow Z_1 = 0 \cdot 67$ (approximately) (From Normal Tables)

$$\therefore\ Q \cdot D \cdot = \sigma Z_1 = 0 \cdot 67\ \sigma \cong \frac{2}{3}\sigma$$

For Normal Distribution mean deviation about mean is given by

$$M \cdot D \cdot = \sqrt{\frac{2}{\pi}}\ \sigma \cong \frac{4}{5}\sigma$$

Also for Normal Distribution standard deviation $S \cdot D \cdot = \sigma$

Hence $Q \cdot D \cdot : M \cdot D \cdot : S \cdot D \cdot :: \frac{2}{3}\sigma : \frac{4}{5}\sigma : \sigma :: \frac{2}{3} : \frac{4}{5} : 1 :: 10 : 12 : 15$

4. As x increases numerically, $f(x)$ decreases rapidly, the maximum probability occurring at the point $x = \mu$ and given by

$$[P(x)]_{max} = \frac{1}{\sigma\sqrt{2\pi}}$$

5. Area under the normal curve is unity

As area $A = \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}}\ e^{\frac{-(x-\mu)^2}{2\sigma^2}}\ dx \quad \cdot \quad \cdot \quad \cdot\ (1)$

Put $Z = \dfrac{x-\mu}{\sigma} \Rightarrow \quad x = \infty, \quad Z = \infty$

$x = -\infty, \quad Z = -\infty$

$\sigma\, dZ = dx$

Substituting in equation (1) we get

$$A = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\, dZ = 2\int_{0}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-Z^2/2}\, dZ \quad \text{(from properties of Integral)}$$

Fake $\dfrac{z^2}{2} = t \Rightarrow z = \sqrt{2t},\ dz = \sqrt{2}\cdot \dfrac{1}{2} t^{-1/2}\, dt$

$z = 0, \quad t = 0$

$z = \infty, \quad t = \infty$

$$\therefore\ A = \cancel{2}\cdot \int_{0}^{\infty} \frac{1}{\sqrt{2\pi}} \sqrt{\cancel{2}}\cdot \frac{1}{\cancel{2}} t^{-1/2}\cdot e^{-t}\, dt$$

$$= \frac{1}{\sqrt{\pi}} \int_{0}^{\infty} t^{-1/2}\cdot e^{-t}\, dt = \frac{1}{\sqrt{\pi}} \int_{0}^{\infty} t^{\frac{1}{2}-1}\cdot e^{-t}\, dt$$

$$A = \frac{1}{\sqrt{\pi}}\cdot \Gamma\left(\tfrac{1}{2}\right) \qquad \left(\because \int_{0}^{\infty} t^{1/2-1} e^{-t}\, dt = \Gamma\left(\tfrac{1}{2}\right)\right)$$

$$= \frac{1}{\cancel{\sqrt{\pi}}}\cdot \cancel{\sqrt{\pi}} \qquad \left(\because \Gamma\left(\tfrac{1}{2}\right) = \sqrt{\pi}\right)$$

$A = 1$

$\therefore$ Area under the normal curve is unity.

6. Since $f(x)$ being the probability, which will never be negative therefore no portion of the curve lies below the x - axis.

7. x - axis is an asymptok to the curve.

8. The distribution has points of inflexion at $x = \mu \pm \sigma$

At the points of inflexion of normal curve, we should have

$f''(x) = 0$ and $f'''(x) \neq 0$

For normal curve we get

$$f''(x) = \frac{-f(x)}{\sigma^2}\left[1 - \frac{(x-\mu)^2}{\sigma^2}\right]$$

$$\therefore\ f''(x) = 0 \Rightarrow \frac{-f(x)}{\sigma^2}\left[1 - \frac{(x-\mu)^2}{\sigma^2}\right] = 0$$

$$\Rightarrow 1 - \frac{(x-\mu)^2}{\sigma^2} = 0$$

$$\Rightarrow x = \mu \pm \sigma.$$

Thus the points of inflexion of the normal curve are given by $x = \mu \pm \sigma$ and $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-1/2}$. That is they are equidistant (at a distance σ) from the mean.

9. Mean Deviation from the mean for Normal Distribution:

Mean Deviation about mean $= \int_{-\infty}^{\infty} |x-\mu|\, f(x)\, dx$

$$= \int_{-\infty}^{\infty} |x-\mu| \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}\, dx \qquad \left(\because\ f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}\right.$$ for Normal Distribution)

Put $z = \frac{x-\mu}{\sigma} \Rightarrow \sigma\, dZ = dx$, also $x-\mu = \sigma z$

$x = \infty, \quad z = \infty$

$x = -\infty, \quad z = -\infty$

$$= \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{\infty} |z|\, e^{-z^2/2}\, dz$$

$$= \frac{2\sigma}{\sqrt{2\pi}} \int_{0}^{\infty} |z|\, e^{-z^2/2}\, dz$$ (By property of Integrales)

Since in the above equation the integrand $|z|\, e^{-z^2/2}$ is an even function of z. Also in $[0, \infty]$, $|z| = z$ we have

$$M \cdot D \cdot (\text{about mean}) = \sqrt{\frac{2}{\pi}} \cdot \sigma \int_{0}^{\infty} z \cdot e^{-z^2/2}\, dz$$

$$= \sqrt{\frac{2}{\pi}} \cdot \sigma \int_{0}^{\infty} e^{-t}\, dt \qquad \left(\text{putting } \frac{z^2}{2} = t\right)$$

$$= \sqrt{\frac{2}{\pi}} \cdot \sigma \left[\frac{e^{-t}}{-1}\right]_{0}^{\infty}$$

$$= \sqrt{\frac{2}{\pi}} \cdot \sigma \cdot \left(e^{-0} - e^{-\infty}\right)$$

$$= \sqrt{\frac{2}{\pi}} \cdot \sigma$$

$$= \frac{4}{5}\sigma$$ (approximately)

10. Area property of Normal Distribution:

If $X \sim N(\mu, \sigma^2)$ then the probability that random value of X will lew between $X = \mu$ and $X = x_1$ is given by

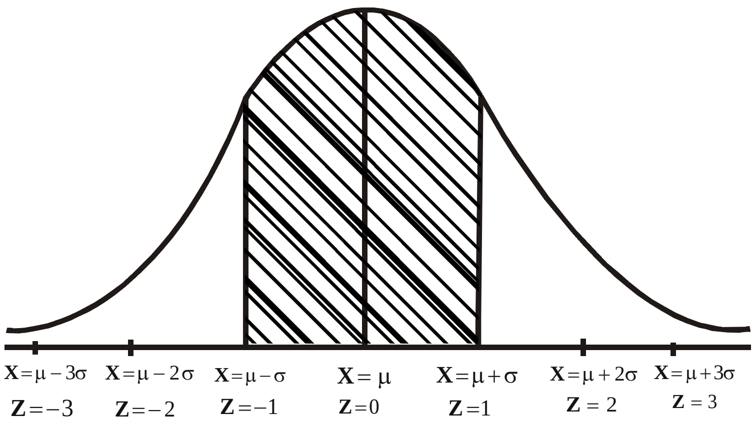
$$P(\mu < X < x_1) = \int_{\mu}^{x_1} f(x)\,dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{\mu}^{x_1} e^{\frac{-(x-\mu)^2}{2\sigma^2}}\,dx$$

Put $\frac{x-\mu}{\sigma} = z$, i.e., $X - \mu = \sigma z$, $dx = \sigma dz$

When $X = \mu,\ Z = 0$; $X = x_1,\ z = \frac{x_1 - \mu}{\sigma} = z_1$ (say)

$$\therefore\quad P(\mu < X < x_1) = P(0 < Z < z_1) = \frac{1}{\sqrt{2\pi}} \int_0^{z_1} e^{-z^2/2}\,dz = \int_0^{z_1} \phi(z)\,dz$$

Where $\phi(z) = \frac{1}{\sqrt{2\pi}} \cdot e^{-z^2/2}$ is the probability density function of standard normal variate. The deinite integral $\int_0^{z_1} \phi(z)\,dz$ is known as normal probability integral and it gives the area under standard normal curve between the ordinates at $z = 0$ and $Z = z_1$. These areas have been tabulated for different values of z_1, at internals of $0 \cdot 01$.



Probability that a random value of X lies in the internal $(\mu - \sigma, \mu + \sigma)$ is given by

$$P(\mu-\sigma < X < \mu + \sigma) = \int_{\mu-\sigma}^{\mu+\sigma} f(x)\, dx$$

$$\Rightarrow P\left(-1 < z = \frac{x-\mu}{\sigma} < +1\right) = \int_{-1}^{1} \phi(z)\, dz$$

$$\Rightarrow P(-1 < z < 1) = z \cdot \int_{0}^{1} \phi(z)\, dz \qquad \text{(By symmetry)}$$

$$= 2 \cdot [\Phi(1) - \Phi(0)]$$

$$= 2 \times 0 \cdot 3413 \qquad (\because \phi(1) = 0 \cdot 3413 \text{ are taken from tables)}$$

$$= 0 \cdot 6826$$

Similarly

$$P(\mu-2\sigma < x < \mu+2\sigma) = P(-2 < z < 2) = \int_{-2}^{2} \phi(z)\, dZ$$

$$= 2 \cdot \int_{0}^{2} \phi(z)\, dZ$$

$$= 2 \cdot [\Phi(2) - \Phi(0)] = 2 \times 0 \cdot 4772 = 0 \cdot 9544 \qquad (\because \Phi(2) = 0 \cdot 4772,\ \phi(0) = 0)$$

$$P(\mu-3\sigma < X < \mu + 3\sigma) = P(-3 < z < 3) = \int_{-3}^{3} \phi(z)\, dZ$$

$$= 2 \cdot \int_{0}^{3} \phi(z)\, dZ$$

$$= 2 \cdot [\Phi(3) - \Phi(0)]$$

$$= 2 \times 0 \cdot 49865 \qquad \begin{pmatrix} \because & \Phi(3) = \dot{0}49865, \\ & \Phi(0) = 0 \end{pmatrix}$$

$$= 0 \cdot 9973$$

Therefore the probability that a normal variate X lies outside the range $\mu \pm 3\sigma$ is given by

$$P\left(|X - \mu| > 3\sigma \right) = P\left(|Z| > 3 \right) = 1 - P\left(-3 \le z \le 3 \right) = 1 - 0 \cdot 9973$$

$$= 0 \cdot 0027$$

Thus the probability of normal variate lie with in the range $\mu \pm 3\sigma$, though theoritically, it may range from $-\infty$ to ∞.

11. Linear combination of independent normal variates is also a normal variate.

Let X_i, $(i = 1, 2, \cdots\cdots, n)$ be n independent normal variates with mean μ_i and variance σ_i^2 respectively. Then

$$M_{X_i}(t) = \exp\left\{ \mu_i \; t + t^2 \; \sigma_i^2 \big/ 2 \right\} \cdot \quad \cdot \quad \cdot (1)$$

The m.g.f. of their linear combimation $\sum_{i=1}^{n} a_i X_i$, where $a_1, a_2, \cdot \cdot \cdot, a_n$ are constatns, is given by

$$M_{\sum_i a_i X_i}(t) = M_{a_1 X_1 + a_2 X_2 + \cdot \cdot \cdot + a_n X_n}(t)$$

$$= M_{a_1 X_1}(t) \cdot M_{a_2 X_2}(t) \cdot \cdot \cdot M_{a_n X_n}(t)$$

($\because$ $X_i^{'}s$ are independent)

$$= M_{X_1}(a_1 t) \cdot M_{X_2}(a_2 t) \cdot \cdot \cdot M_{X_n}(a_n t) \cdot \quad \cdot \quad \cdot (2)$$

$$\left(\because \; M_{CX}(t) = M_X(Ct) \right)$$

From equation (1) we have

$$M_{X_i}(a_i\, t) = \exp\left\{\mu_i a_i t + {t^2\, a_i^2\, \sigma_i^2}/{2}\right\}$$

Equation (2) gives

$$M_{\sum_i a_i X_i}(t) = \left[\exp\left(\mu_1 a_1 t + {t^2 a_1^2 \sigma_1^2}/{2}\right) \cdot \exp\left(\mu_2 a_2 t + {t^2 a_2^2 \sigma_2^2}/{2}\right) \cdot \cdot \cdot \exp\left(\mu_n a_n t + {t^2 a_n^2 \sigma_n^2}/{2}\right)\right]$$

$$= \exp\left[\left(\sum_{i=1}^{n} a_i \mu_i\right) t + t^2 \left(\sum_{i=1}^{n} a_i^2\, \sigma_i^2\right)\Big/2\right] \quad \cdot \cdot \cdot (3)$$

Equation (3) is the m.g.f. of a normal variate with mean $\sum_{i=1}^{n} a_i \mu_i$ and variance $\sum_{i=1}^{n} a_i^2\, \sigma_i^2$. Hence by uniqueness theorem of m.g.f.

$$\sum_{i=1}^{n} a_i\, X_i \sim N\left[\sum_{i=1}^{n} a_i^2\, \sigma_i^2\right]$$

Hence linear combination of independent normal variates is also a normal variate.

14.4 Moment Generating Function of Normal Distribution:

The m.g.f. (about origin) is given by

$$M_X(t) = \int_{-\infty}^{\infty} e^{tx} \cdot f(x)\, dx$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{tx} \cdot \exp\left\{-(x-\mu)^2 \big/ 2\sigma^2\right\} dx$$

$$\left(\because f(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot \exp\left\{ -(x-\mu)^2 \Big/ 2\sigma^2 \right\}\right)$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\{t(\mu + \sigma z)\} \exp\left(-z^2 \big/ 2\right) dz$$

$$\left(\because \frac{x-\mu}{\sigma} = z \Rightarrow x = \mu + \sigma z \ \& \ dx = \sigma dz\right)$$

$$= \frac{e^{\mu t}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left\{\frac{-1}{2}\left(z^2 - 2t\sigma z\right)\right\} dz$$

$$= e^{\mu t} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left\{\frac{-1}{2}\left(z^2 + \sigma^2 t^2 - 2t\sigma z - \sigma^2 t^2\right)\right\} dz$$

$$= e^{\mu t} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left\{\frac{-1}{2}(z - \sigma t)^2 - \sigma^2 t^2\right\} dz$$

$$= e^{\mu t} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left(+t^2 \sigma^2 \big/ 2\right) \cdot \exp\left(\frac{-(z-\sigma t)^2}{2}\right) dz$$

$$= e^{\mu t + \frac{t^2\sigma^2}{2}} \cdot \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left(-\mu^2 \big/ 2\right) du \qquad (\because z - \sigma t = u, dz = du)$$

$$M_X(t) = e^{\mu t + \frac{t^2\sigma^2}{2}} \cdot 1 = e^{\mu t + t^2 \sigma^2 / 2}$$

($\because$ Area under normal curve is unity)

Note: M.G.F. of standard Normal Variate. If $X \sim N\left(\mu, \sigma^2\right)$ then standard normal variate is given by $Z = \frac{(X-\mu)}{\sigma}$.

Now $M_z(t) = e^{-\mu t/\sigma} \cdot M_X\left(t/\sigma\right) = e^{-\mu t/\sigma} \cdot e^{\frac{\mu t}{\sigma} + \frac{t^2}{\sigma^2} \cdot \frac{\sigma^2}{2}} = e^{t^2/2}$

14.5 Cumulant Generating Function:

The c.g.f. of Normal Distribution is given by

$$K_X(t) = \ell og_e M_X(t) = \ell og_e\left(e^{\mu t + t^2\sigma^2/2}\right) = \left(\mu t + \frac{t^2\sigma^2}{2}\right) \cdot \ell og_e^e = \mu t + \frac{t^2\sigma^2}{2}$$

$$\left(\because\ \ell og_e^e = 1\right)$$

$\therefore$ Mean = K_1 = Coefficient of t in $K_X(t) = \mu$

Variance = K_2 = Coefficient of $\frac{t^2}{2!}$ in $K_X(t) = \sigma^2$

and K_r = coefficient of $\frac{t^2}{r!}$ in $K_X(t) = 0$; for $r = 3, 4, \cdots\cdots$

$\mu_3 = K_3 = 0$ and $\mu_4 = K_4 + 3K_2^2 = 3 \cdot \sigma^4$

Hene $\beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{0}{\left(\sigma^2\right)^3} = 0$ and $\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{3\sigma^4}{\left(\sigma^2\right)^2} = 3$

14.6 Characteristic Function of Normal Distribution:

From the definition Characteristic Function we have

$$\phi_X(t) = E\left(e^{itX}\right) = \int_{-\infty}^{\infty} e^{itX} f(x)\, dx$$

$$= \int_{-\infty}^{\infty} e^{itX} \cdot \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}} dx$$

put $\frac{x-\mu}{\sigma} = z \Rightarrow x = \mu + \sigma z$; $dx = \sigma dz$

$$\therefore \phi_z(t) = \int_{-\infty}^{\infty} \frac{e^{it(\mu+\sigma z)}}{\sqrt{2\pi}} \cdot e^{-z^2/2} dz$$

$$= \frac{e^{it\mu}}{\sqrt{2\pi}} \cdot \int_{-\infty}^{\infty} e^{\frac{-z^2}{2}+it\sigma z} dz$$

$$= \frac{e^{it\mu}}{\sqrt{2\pi}} \cdot \int_{-\infty}^{\infty} e^{\frac{-1}{2}\left(z^2-2it\sigma z+(it\sigma)^2-(it\sigma^2)\right)} dz$$

$$= \frac{e^{it\mu} \cdot e^{\frac{-1}{2}t^2\sigma^2}}{\sqrt{2\pi}} \cdot \int_{-\infty}^{\infty} e^{\frac{-1}{2}(z-it\sigma)^2} dz$$

put $z - it\sigma = K,\ dz = dK$

$$\therefore \phi_z(t) = \frac{e^{it\mu-\frac{1}{2}t^2\sigma^2}}{\sqrt{2\pi}} \cdot \int_{-\infty}^{\infty} e^{-k^2/2} dk$$

$$= \frac{2}{\sqrt{2\pi}} \cdot e^{it\mu-\frac{1}{2}t^2\sigma^2} \cdot \int_{0}^{\infty} e^{-k^2/2} dk$$

put $\frac{k^2}{2} = h$ i.e., $dk = \frac{dh}{\sqrt{2h}}$

$$\phi_z(t) = \frac{1}{\sqrt{\pi}} \cdot e^{it\mu - \frac{1}{2}t^2\sigma^2} \cdot \int_0^\infty h^{-1/2} \cdot e^{-h}\, dh$$

$$= \frac{1}{\sqrt{\pi}} \cdot e^{it\mu - \frac{1}{2}t^2\sigma^2} \cdot \Gamma\left(1/2\right) = \frac{1}{\sqrt{\pi}} \cdot e^{it\mu - \frac{1}{2}t^2\sigma^2} \cdot \sqrt{\pi}$$

$$= e^{it\mu - \frac{1}{2}t^2\sigma^2} \qquad \left(\because \Gamma\left(1/2\right) = \sqrt{\pi}\right)$$

14.7 Moments of Normal Distribution:

Odd order moments about mean are given by

$$\mu_{2n+1} = \int_{-\infty}^{\infty} (x-\mu)^{2n+1} f(x)\, dx$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} (x-\mu)^{2n+1} \exp\left\{-(x-\mu)^2 / 2\sigma^2\right\} dx$$

$$\therefore \quad \mu_{2n+1} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (\sigma z)^{2n+1} \cdot \exp\left(-z^2/2\right) dz$$

$$\left(\because \frac{x-\mu}{\sigma} = z \Rightarrow x - \mu = \sigma z,\ dx = \sigma dz\right)$$

$$= \frac{\sigma^{2n+1}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z^{2n+1} \exp\left(-z^2/2\right) dz = 0 \quad \cdot \quad \cdot \quad \cdot (1)$$

Since the integrand $z^{2n+1} \cdot \exp\left(-z^2/2\right)$ is an odd function of z.

Even order moments about mean are given by

$$\mu_{2n} = \int_{-\infty}^{\infty} (x-\mu)^{2n} \cdot f(x)\,dx$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} (x-\mu)^{2n} \cdot \exp\left\{ -(x-\mu)^2 \big/ 2\sigma^2 \right\} dx$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (\sigma z)^{2n} \cdot \exp\left(-z^2 \big/ 2 \right) dz$$

$$= \frac{\sigma^{2n}}{\sqrt{2\pi}} \cdot \int_{-\infty}^{\infty} z^{2n} \exp\left(-z^2 \big/ 2 \right) dz$$

$$= \frac{\sigma^{2n}}{\sqrt{2\pi}} \cdot \int_{-\infty}^{\infty} z^{2n} \cdot \exp\left(-z^2 \big/ 2 \right) dz$$

($\because$ Integrand is an even function of z)

$$\therefore \mu_{2n} = \frac{2 \cdot \sigma^{2n}}{\sqrt{2\pi}} \int_0^{\infty} (2t)^n \cdot e^{-t} \cdot \frac{dt}{\sqrt{2t}} \qquad \left(\because \frac{z^2}{2} = t \Rightarrow z = \sqrt{2t}\,;\, dz = \frac{1}{\sqrt{2t}}\,dt \right)$$

$$= \frac{2^n \cdot \sigma^{2n}}{\sqrt{\pi}} \cdot \int_0^{\infty} e^{-t} \cdot t^{\left(n+1/2\right)-1}\,dt$$

$$\Rightarrow \mu_{2n} = \frac{2^n \cdot \sigma^{2n}}{\sqrt{\pi}} \cdot \Gamma\left(n + 1/2\right) \qquad \left(\because \int_0^{\infty} e^{-t} \cdot t^{\left(n+1/2\right)-1}\,dt = \left(n + 1/2\right) \right)$$

Changing $n \ \ell og \ (n-1)$ in the above expression we get

$$\mu_{2n-2} = \frac{2^{n-1} \cdot \sigma^{2n-2}}{\sqrt{\pi}} \cdot \Gamma\left(n - 1/2\right)$$

$$\therefore \quad \frac{\mu_{2n}}{\mu_{2n-2}} = \frac{2^n \cdot \sigma^{2n}}{\sqrt{\pi}} \frac{\Gamma\left(n+1/2\right)}{\Gamma\left(n-1/2\right)} \frac{\sqrt{\pi}}{2^{n-1} \cdot \sigma^{2n-2}}$$

$$= \quad 2 \cdot \sigma^2 \cdot \frac{\Gamma\left(n+1/2\right)}{\Gamma\left(n-1/2\right)}$$

$$= 2\sigma^2 \cdot \frac{\left(n+\frac{1}{2}-1\right) \Gamma\left(n+\frac{1}{2}-1\right)}{\Gamma\left(n-\frac{1}{2}\right)} = \frac{2\sigma^2 \left(n-\frac{1}{2}\right) \Gamma\left(n-\frac{1}{2}\right)}{\Gamma\left(n-\frac{1}{2}\right)}$$

$$= 2\,\sigma^2 \left(n - \frac{1}{2}\right) \qquad \left[\because \Gamma(n) = (n-1)\,\Gamma(n-1)\right]$$

$$\Rightarrow \mu_{2n} = \sigma^2\,(2n-1)\,\mu_{2n-2} \quad \cdot \quad \cdot \quad \cdot (2)$$

Which gives the recurrence relation for the moments of Normal Distribution. From equation (2) we get

$$\mu_{2n} = \left[(2n-1)\sigma^2\right]\left[(2n-3)\sigma^2\right]\mu_{2n-4}$$

$$= \left[(2n-1)\,\sigma^2\right]\left[(2n-3)\,\sigma^2\right]\left[(2n-5)\,\sigma^2\right]\mu_{2n-6}$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$= \left[(2n-1)\sigma^2\right]\left[(2n-3)\sigma^2\right]\left[(2n-5)\sigma^2\right] \cdot \quad \cdot \quad \cdot \left(3\sigma^2\right)\left(1\sigma^2\right)\mu_0$$

$$= 1 \cdot 3 \cdot 5 \cdot \quad \cdot \quad \cdot (2n-1)\,\sigma^{2n} \quad \cdot \quad \cdot \quad \cdot (3)$$

From equations (1) & (2) we conclude that for the Normal Distribution all odd order moments about mean vanish and the even order moments are given by equation (3).

14.8 Normal Distribution as Limiting Case of Binomial Distribution:

If $X \sim B(n, p)$, then the m.g.f. is $M_X(t) = \left(q + pe^t\right)^n$

The m.g.f. of standard binomial variate is

$$Z = \frac{X - np}{\sqrt{npq}} = \frac{X - \mu}{\sigma}$$

If $\mu = np$ and $\sigma^2 = npq$, is given by

$$M_Z(t) = e^{-\mu t/\sigma} \cdot M_X\left(t/\sigma\right)$$

$$M_Z(t) = e^{-npt/\sqrt{npq}} \cdot \left(q + pe^{t/\sqrt{npq}}\right)^n$$

$$= \left[e^{-pt/\sqrt{npq}} \left(q + pe^{t/\sqrt{npq}}\right)\right]^n$$

$$= \left[q \cdot e^{-pt/\sqrt{npq}} + pe^{t/\sqrt{npq}} \cdot e^{-pt/\sqrt{npq}}\right]^n$$

$$= \left[qe^{-pt/\sqrt{npq}} + pe^{\frac{t-(1-p)}{\sqrt{npq}}}\right]^n$$

$$= \left[q \cdot e^{-pt/\sqrt{npq}} + p \cdot e^{tq/\sqrt{npq}}\right]^n \qquad (\because 1 - p = q)$$

$$= \left[q\left\{1 - \frac{pt}{\sqrt{npq}} + \frac{p^2t^2}{2npq} + 0^1\left(n^{-3/2}\right)\right\} + p\left\{1 + \frac{qt}{\sqrt{npq}} + \frac{q^2t^2}{2npq} + 0^{11}\left(n^{-3/2}\right)\right\}\right]^n$$

$$\left(\because \; e^{-x} = 1 - \frac{x}{1!} + \frac{x^2}{2!} - \frac{x^3}{3!} + \cdot \quad \cdot \quad \cdot ;\; e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdot \quad \cdot \quad \cdot\right)$$

Where $0^1\left(n^{-3/2}\right)$ and $0^{11}\left(n^{-3/2}\right)$ involve terms containing $n^{3/2}$ and higher powers of n in the denominator.

$$\therefore M_Z(t) = \left[(q+p) + \frac{t^2 \, pq}{2n \, pq}(p+q) + 0\left(n^{-3/2}\right)\right]^n$$

$$= \left[1 + \frac{t^2 \, pq}{2n \, pq} + 0\left(n^{-3/2}\right)\right]^n \qquad (\because \; p+q = 1)$$

$$= \left[1 + \frac{t^2}{2n} + 0\left(n^{-3/2}\right)\right]^n$$

Where $0\left(n^{-3/2}\right)$ involves terms with $n^{3/2}$ and higher powers of n in the denominator.

$$\therefore \log M_Z(t) = n \cdot \log\left[1 + \frac{t^2}{2n} + 0\left(n^{-3/2}\right)\right] \qquad \left(\because \log m^n = n \log m\right)$$

$$= n\left[\left\{\frac{t^2}{2n} + 0\left(n^{-3/2}\right)\right\} - \frac{1}{2}\left\{\frac{t^2}{2n} + 0\left(n^{-3/2}\right)\right\}^2 + \cdot \quad \cdot \quad \cdot\right]$$

$$\left[\because \log(1+x) = x - \frac{x^2}{x} + \frac{x^3}{3} - \frac{x^4}{4} + \cdot \quad \cdot \quad \cdot\right]$$

$$= \frac{t^2}{2} + 0^{111}\left(n^{-1/2}\right)$$

Where $0^{111}\left(n^{-1/2}\right)$ involve terms with $n^{1/2}$ and higher powers of n in the denominator. As $n \to \infty$ we get

$$\lim_{n \to \infty} \log M_Z(t) = \frac{t^2}{2}$$

$$\Rightarrow \lim_{n \to \infty} M_Z(t) = \exp\left(t^2/2\right) \quad \cdot \quad \cdot \quad \cdot (1)$$

The above expression (1) is the m.g.f. of standard normal variate. Hence by uniqueness theorem of moment generating functions standard binomial variate tends to standard normal variate as $n \to \infty$. In otherwords, binomial distribution tends to normal distribution as $n \to \infty$.

14.9 Normal Distribution as Limiting Case of Poisson Distribution:

$$\text{Let} \quad Z = \frac{X-\lambda}{\sqrt{\lambda}} = \frac{X-\mu}{\sigma}$$

Where mean $\mu = \lambda$ and $\sigma^2 = \lambda \Rightarrow \sigma = \sqrt{\lambda}$ for poisson distribution.

i.e., $E(Z) = \lambda,\ \ Var(Z) = \lambda$.

$$\therefore\ E(Z) = E\left(\frac{X-\lambda}{\sqrt{\lambda}}\right) = \frac{1}{\sqrt{\lambda}} \cdot E(X-\lambda) = E(X) - \lambda = \lambda - \lambda = 0$$

$$V(Z) = E\left(Z^2\right) - [E(Z)]^2 = E\left[\frac{X-T}{\sqrt{\lambda}}\right]^2 - \left\{E\left[\frac{X-\lambda}{\sqrt{\lambda}}\right]\right\}^2$$

$$= E\left[\frac{X-\lambda}{\sqrt{\lambda}}\right]^2 \qquad \left(\because\ E\left(\frac{X-\lambda}{\sqrt{\lambda}}\right) = 0\right)$$

$$= \frac{1}{\lambda}\ E(X-\lambda)^2 = \frac{1}{\lambda}\mu_2 = \frac{1}{\chi}\chi = 1 \qquad \left(\because\ \mu_2 = V(Z) = \lambda\right)$$

M.G.F. of $Z = M_Z(t) = E\left(e^{tZ}\right) = E\left[e^{t(X-\lambda)/\sqrt{\lambda}}\right]$

$$= E\left[e^{tX/\sqrt{\lambda}} \cdot e^{-t\lambda/\sqrt{\lambda}}\right]$$

$$= E\left[e^{tX/\sqrt{\lambda}} \cdot e^{-t\sqrt{\lambda}}\right]$$

$$= e^{-t\sqrt{\lambda}} \cdot E\left(e^{tX/\sqrt{\lambda}}\right)$$

$$= e^{-t\sqrt{\lambda}} \sum_{x=0}^{\infty} \frac{e^{-\lambda}\lambda^x}{x!} \cdot e^{tx/\sqrt{m}} \qquad \left(\because\ E(X) = \sum_{x=0}^{\infty} x \cdot P(x,\lambda)\right)$$

$$= e^{-t\sqrt{\lambda}} \cdot e^{-\lambda} \sum_{x=0}^{\infty} \frac{\left(\lambda\ e^{t/\sqrt{m}}\right)^x}{x!}$$

$$= \quad e^{-t\sqrt{\lambda}-\lambda} \cdot \left[1+\frac{\lambda e^{t/\sqrt{\lambda}}}{1!}+\frac{\left(\lambda \cdot e^{t/\sqrt{\lambda}}\right)^2}{2!}+\cdot \quad \cdot \quad \cdot\right]$$

$$= e^{-\lambda-t\sqrt{\lambda}} \cdot \exp\left(\lambda \cdot e^{t/\sqrt{\lambda}}\right)$$

$$= \exp\left[-\lambda - t\sqrt{\lambda} + \lambda e^{t/\sqrt{\lambda}}\right]$$

$$= \exp\left[-\lambda - t\sqrt{\lambda} + \lambda\left(1+\frac{t}{\sqrt{\lambda}}+\frac{t^2}{2!\,\lambda}+\frac{t^3}{3!\,\lambda^{3/2}}+\cdot \quad \cdot \quad \cdot\right)\right]$$

$$M_Z(t) = \exp\left[\frac{1}{2}t^2+\frac{1}{3!}\frac{t^2}{\sqrt{\lambda}}+\cdot \quad \cdot \quad \cdot\right]$$

As $\lambda \to \infty$, we get

$$M_Z(t) = \exp\left(t^2/2\right) \cdot \quad \cdot \quad \cdot (1)$$

Above expression (1) is the m.g.f. of standard normal variate. Therefore by uniqueness theorem of m.g.f.'s standard poisson variate tends to standard normal variate as $\lambda \to \infty$. Hence Poisson Distribution tends Normal Distribution for large values of parameter.

14.10 Importance of Normal Distribution in Statistics:

This was an important distribution which was initially discovered for studying the random errors of measurements, that isduring the calculations of orbits of celestial bodies. It happened because of a remarkable coincidence that normal distribution follows all the basic principles of errors. It is mainly for this quality that the distribution has a wide range of application these days in the theory of statistics. To count a few are industrial quality control, testing of significance, sampling distributions of various statistics, graduation of non-normal curves etc. Length of the leaves from a particular point of time, weights of trees of the same variety, weights taken for a group of students taken of the same age, intelligence, proportion of male to female births for some particular geographical region over a period of years and many other examples from variousfields can be given which are studied through normal distribution. Some facts of Normal Distribution we given below:

1. Normal Distribution approximates the p.d.f's of most of the commonly occurring distributions such as Binomial, Poisson, Hyper - Geometric ... etc.

2. Many of the sampling distributions such as students t, Snedecor's F, Pearson's X^2 etc. are asynptotically normal. Also most of sampling distributions tend to normality as $n \to \infty$.

3. Sometimes a non - normal variate begins to exhibit normality properties under suitable transformations.

4. $P\{|Z| \geq 1 \cdot 96\} = 0 \cdot 05$ and $P\{|Z| \geq 3\} = 0 \cdot 0027$, if Z is $N(0, 1)$. These properties of $N(0, 1)$ form the basis of "Large Sample Theory".

5. For large number (> 30) of variate observations the sample can always be treated as normal, even though the parent population is non - normal (central limit theorem).

6. In tests of significance the posent population is assumed to be normal.

7. Normal distribution finds large applications in statistical quality control in industries and graduation of non - normal curves.

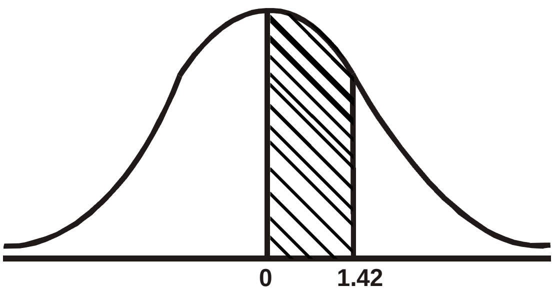
14.11 Workedout Examples:

Example 1: Let X be a random variable with the Standard Normal Distribution ϕ.

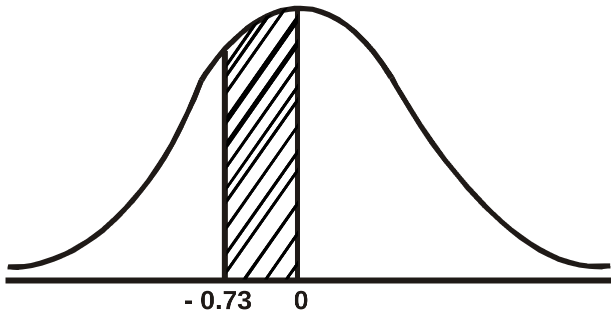
Find (i) $P(0 \leq X \leq 1 \cdot 42)$ (ii) $P(-0 \cdot 73 \leq X \leq 0)$ (iii) $P(X \geq 1 \cdot 13)$

(iv) Determine the value of t if $P(X \leq t) = 0 \cdot 7967$

Solution: (i) Here $P(0 \leq X \leq 1 \cdot 42)$ is equal to the area under the standard normal curve between 0 and 1.42. Thus in Table of areas under standard normal curve, page ( ) look down the first column until 1.4 reached, and then continue right to column 2. The entry is 0.4222. Hence $P(0 \leq X \leq 1 \cdot 42) = 0 \cdot 4222$.



(ii) $P(-0 \cdot 73 \leq X \leq 0) = P(0 \leq X \leq 0 \cdot 73)$ (By symmetry) is equal to the area under the standard normal curve between 0 and 0.73. Thus in Table of areas under standard normal curve, page look down the first column until 0.73 reached and then continue right to column 2. The entry is 0.2673, hence $P(0 \le X \le 0 \cdot 73) = 0 \cdot 2673$.

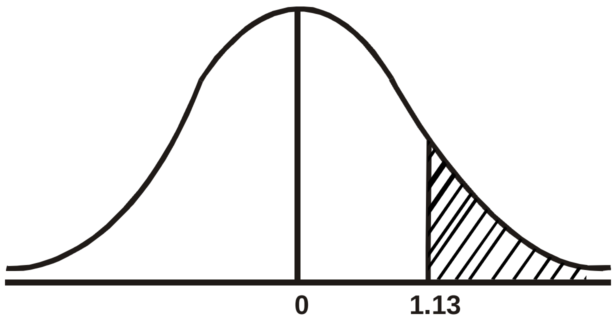


(iii) $P(X \ge 1 \cdot 13)$ which can be written as

$$P(X \ge 1 \cdot 13) = P(X \ge 0) - P(0 \le X \le 1 \cdot 13)$$

$$= 0 \cdot 5000 - 0 \cdot 3708 = 0 \cdot 1292$$

From table of area under the standard normal curve gives the value of 1.13 as 0.3708 & $P(X \ge 0) = 0 \cdot 5000$.



(iv) $P(X \le t) = 0 \cdot 7967$

Here t must be positive since the probability is greater than $\frac{1}{2}$. Therefore

we write $P(0 \le X \le t) = P(X \le t) - \frac{1}{2}$

$$= 0 \cdot 7967 - 0 \cdot 5000 \qquad (\because P(X \le t) = 0 \cdot 7967)$$

$$P(0 \le X \le t) = 0 \cdot 2967$$

Now observing the value 0.2967 in Table of areas under standard normal curve which lies at $t = 0 \cdot 83$.

Therefore we obtain the value of t as 0.83

Example 2: In a Normal Distribution 31 percent of the items are under 45 and 8 percent are over 64. Find the mean and standard deviation of the distribution. Given $\phi(0\cdot5)=0\cdot19$, $\phi(1\cdot405)=0\cdot42$.

Solution:

Example 3: If $f(x)=e^{-x}, 0\le x\le\infty$ then find (i) $P(1<x<2)$, (ii) $F(x),\ x\ge 0$. Show tht it is a p.d.f. and obtain its mean μ_2, μ_3 and μ_4.

Solution: Given $f(x)=e^{-x},\ 0\le x\le\infty$

(i) $P(1<x<2)=\int_1^2 f(x)\,dx=\int_1^2 e^{-x}\,dx=\left[-e^{-x}\right]_1^2=e^{-1}-e^{-2}=\frac{1}{e}-\frac{1}{e^2}=\frac{e-1}{e^2}$

(ii) $F(x)=\int_0^x f(x)\,dx=\int_0^x e^{-x}\,dx=\left[-e^{-x}\right]_0^x=1-e^{-x}$

(iii) Since $f(x)\ge 0\ \forall\ x$ in $[0,\infty]$ and $\int_0^\infty f(x)\,dx=\int_0^\infty e^{-x}\,dx=1$

$\therefore\ f(x)$ is a p.d.f.

Now mean $=\mu_1^1=\int_0^\infty x^2\cdot e^{-x}\,dx=\int_0^\infty x^{2-1}\cdot e^{-x}\,dx=\Gamma(2)=1$

$$\mu_2^1=\int_0^\infty x^2\cdot e^{-x}\,dx=\Gamma(3)=2$$

$$\mu_1^1\ \&\ \mu_3^1=\Gamma(4)=6,\ \mu_4^1=\Gamma(5)=24$$

$$\therefore\ \mu_2=\mu_2^1-{\mu_1^1}^2=2-1=1$$

$$\mu_3=\mu_3^1-3\mu_2^1\,\mu_1^1+2{\mu_1^1}^3=6-3(2)(1)+2(1)^3=2$$

and $\mu_4=\mu_4^1-4\mu_3^1\,\mu_1^1+6\mu_2^1\,{\mu_1^1}^2-3{\mu_1^1}^4=24-4(6)(1)+6(2)(1)^2-3(1)^4=9$

Example 4: The water consumption of a city, in excess of 20,000 gallons, is exponentially distributed with mean $\sigma = 20{,}000$. The city's water works has a daily stock of 4,00,000 gallons. What is the probability that the stock is in sufficient for atleast two of the three days selected at random?

Solution: If Y is the total consumption in a day, then $X = Y - 20{,}000$ has an exponential distribution with mean $20{,}000$ i.e., with the probbaility density function

$$f(x) = \frac{1}{20{,}000} e^{-x/20{,}000}, \text{ for } 0 \le X \le \infty.$$

$\because$ The stock will be proved insufficient, if the demand exceeds 40,000 gallons, i.e., $X \ge 40{,}000 - 20{,}000$, i.e., $\ge 20{,}000$.

$\therefore$ The probability that the stock remains insufficient on any particular day

$$= p\{X \ge 20{,}000\} = \int_{20{,}000}^{\infty} \frac{1}{20{,}000} \cdot e^{-x/20{,}000}\, dx = e^{-10}$$

$\therefore$ The probability that the stock is insufficient for atleast two or three days selected at random = probability that it is insufficient for all the three days + probability that it is insufficient for two or the three days.

$$= \left(e^{-10}\right)^3 + 3C_2 \left(e^{-10}\right)^2 \left(1 - e^{-10}\right)$$

$$= e^{-20}\left\{e^{-10} + 3\left(1 - e^{-10}\right)\right\} = e^{-20}\left\{3 - 2e^{-10}\right\}$$

Example 5: Suppose that X has an exponentil distribution with parameter λ. What is the probbaility that X exceeds its expected value ?

Solution: Since X has an exponential distribution with parameter λ

we have

$$f(x) = \lambda \cdot e^{-\lambda x}, \; x > 0$$

then $E(X) = \int_0^{\infty} x \cdot f(x)\, dx = \lambda \int_0^{\infty} x \cdot e^{-\lambda x}\, dx = \frac{1}{\lambda}$

and $p\left(X > \frac{1}{\lambda}\right) = 1 - p\left(X \le \frac{1}{\lambda}\right) = e^{-\lambda \cdot \frac{1}{\lambda}} = e^{-1}$

Example 6: Cars arrive at a petrol bunk randomly every 2 minutes on the average. Determine the probability that the arrival time of cars does not exceed 1 minute.

Solution: Here we have to find $P(X \le 1)$. Since here X is the arrival time of cars follows an exponential distribution with parameter λ.

$$\therefore f(x) = \lambda \cdot e^{-\lambda x}, \; \lambda > 0, x > 0$$

$$= 0, \quad \text{otherwise}$$

the c.d.f. of exponential distribution is $F(x) = 1 - e^{-\lambda x}$

$$= p(X \le x)$$

Since the rate of arrival is $\lambda = \frac{1}{2}$ arrival per minute.

then $P(X \le 1) = 1 - e^{\frac{-1}{2}(1)} = 1 - e^{\frac{-1}{2}} = 0 \cdot 39$

Example 7: If X is normally distributed with $\mu = 0$ nd variance 1 what is the expectation and variance of (i) (X^2), (ii) e^{ax}?

Solution: Since $X \sim N(0, 1)$ we have the p.d.f.

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-x^2}{2}\right) dx \cdot \quad \cdot \quad \cdot (1)$$

(i) Mean value of X^2 is given by

$$E(X^2) = \mu_2^1 = \int_{-\infty}^{\infty} x^2 \cdot f(x) dx$$

$$= \int_{-\infty}^{\infty} x^2 \cdot \frac{1}{\sqrt{2\pi}} \cdot \exp\left(\frac{-x^2}{2}\right) dx \qquad (\because \text{ form } (1))$$

$$= \frac{2}{\sqrt{2\pi}} \int_0^{\infty} x^2 \cdot \exp\left(\frac{-x^2}{2}\right) dx$$ (By properties of Integrals)

Putting $\frac{x^2}{2} = t$

$x^2 = 2t \Rightarrow x = \sqrt{2t}$ $\quad$ $x = 0, t = 0$

$dx = \frac{1}{2\sqrt{2t}} \cdot dt$ $\quad$ $x = \infty, t = \infty$

$$= \frac{2}{\sqrt{2\pi}} \cdot \int_0^{\infty} 2t \cdot \exp(-t) \frac{1}{2} \frac{1}{\sqrt{2}} \frac{1}{\sqrt{t}}^{2} \cdot dt$$

$$= \frac{2}{\sqrt{\pi}} \int_0^{\infty} t^{1-1/2} e^{-t} dt$$

$$= \frac{2}{\sqrt{\pi}} \int_0^{\infty} t^{1/2} \cdot e^{-t} dt$$

$$= \frac{2}{\sqrt{\pi}} \int_0^{\infty} t^{\frac{3}{2}-1} \cdot e^{-t} dt$$

$$E(X^2) = \frac{2}{\sqrt{\pi}} \cdot \Gamma(3/2) = \frac{2}{\sqrt{\pi}} \cdot \frac{1}{2} \cdot \Gamma\left(\frac{1}{2}\right) = \frac{1}{\sqrt{\pi}} \cdot \sqrt{\pi} = 1$$

$\therefore E(X^2) = \mu_1^1 = 1$ $\quad$ i.e., Mean of X^2

Also μ_2^1 of X^2 is given by

$$E\left((X^2)^2\right) = E(X^4)$$

$$= \int_{-\infty}^{\infty} x^4 \cdot f(x)\, dx$$

$$= \int_{-\infty}^{\infty} x^4 \cdot \frac{1}{\sqrt{2\pi}} \cdot \exp\left(\frac{-x^2}{2}\right) dx \qquad (\because \text{ from } (1))$$

$$= \frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^{\infty} x^4 \cdot \exp\left(-x^2/x\right) dx = \frac{2}{\sqrt{2\pi}} \int_0^{\infty} x^4 \cdot \exp\left(\frac{-x^2}{2}\right) dx$$

$$= \frac{2}{\sqrt{2\pi}} \cdot \int_0^{\infty} \left(\sqrt{2}\, t\right)^4 \cdot e^{-t} \cdot \sqrt{2}\frac{1}{2}\, t^{-1/2}\, dt$$

$$= \frac{4}{\sqrt{\pi}} \cdot \int_0^{\infty} t^{3/2} \cdot e^{-t}\, dt \qquad \left(\begin{array}{l} \because \frac{x^2}{2} = t = 1\ x = \sqrt{2t} \\ dx = \frac{1}{2\sqrt{t}}\sqrt{2}\ \ dt \end{array}\right)$$

$$= \frac{4}{\sqrt{\pi}} \cdot \int_0^{\infty} t^{3/2 - 1} \cdot e^{-t}\, dt$$

$$= \frac{4}{\sqrt{\pi}} \int_0^{\infty} t^{5/2 - 1} \cdot e^{t}\, dt \qquad \begin{array}{l}(x = 0,\ \ t = 0, \\ \ x = \infty,\ t = \infty)\end{array}$$

$$= \frac{4}{\sqrt{\pi}} \cdot 5\left(5/2\right) = \frac{4}{\sqrt{\pi}} \cdot \frac{3}{2} \cdot \Gamma\left(3/2\right) = \frac{4}{\sqrt{\pi}} \cdot \frac{3}{2} \cdot \frac{1}{2} \cdot \Gamma\left(1/2\right)$$

$$= \frac{3}{\sqrt{\pi}} \cdot \sqrt{\pi} = 3$$

$\left(\because \quad \Gamma\left(\tfrac{1}{2}\right) = \sqrt{\pi}\right)$

$$\therefore \text{ variance of } \left(X^2\right) = \mu_2 = \mu_2^1 - {\mu_1^1}^2 = E\left(X^4\right) - \left[E\left(X^2\right)\right]^2$$

$$= 3 - 1 = 2.$$

(ii) Mean value of e^{ax}

$$\mu_1^1 = E\left(e^{ax}\right) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \cdot \exp\left(\frac{-x^2}{2}\right) \cdot e^{ax}\, dx$$

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \cdot \exp\left\{\frac{-1}{2}\left(x^2 - 2ax + a^2 - a^2\right)\right\} dx$$

$$= \int_{-\infty}^{\infty} e^{a^2/2} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(x - a^2\right)}\, dx$$

$$= e^{a^2/2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}}\, e^{-\frac{1}{2}(x-a)^2}\, dx$$

$$\mu_1^1 = e^{a^2/2}$$

μ_2^1 of $e^{ax} = E\left(e^{2ax}\right)$

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}}\, e^{\frac{-x^2}{2}} \cdot e^{2ax}\, dx$$

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left[\left\{x^2 - 4ax + 2a^2\right\} - 2a^2\right]}$$

$$= e^{2a^2} \cdot \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left\{\frac{-1}{2}(x-2a)^2\right\} dx$$

$$= e^{2a^2}$$

$\therefore$ variance of e^{ax} is

$$V\left(e^{ax}\right) = \mu_2 = \mu_2^1 - \mu_1^{1^2}$$

$$= e^{2a^2} - e^{a^2}$$

Example 8: If X is a normal variate with mean 1 and variance 4, Y is another normal variate independent of X with mean 2 and variance 3. What isthe distribution of $X + 2Y$?

Solution: We are given $X \sim N(1, 4)$

i.e., given $E(X) = 1,\ V(X) = 4$.

Again $Y \sim N(2, 3)$

i.e., given $E(Y) = 2,\ V(Y) = 3$.

Here X, Y are independent.

$\therefore\ E(X + 2Y) = E(X) + 2 \cdot E(Y) = 1 + 2 \cdot 2 = 1 + 4 = 5$

$V(X + 2Y) = V(X) + V(2Y) = V(X) + 4 \cdot V(Y) = 4 + 4 \cdot 3$

$= 4 + 12 = 16$

$\therefore\ x + 2Y \sim N(5, 16)$

Example 9: If two normal populations have the same total frequency but the standard deviation of one is K times that of the other, show that the maximum frequency of the first is $1/k$ that of the other.

Solution: If σ and $K\sigma$ are standard deviations of the two populations and Let N be the total frequency. Then we have

$$Y = \frac{N}{\sigma\sqrt{2\pi}} \cdot \exp\left[\frac{-(x-\mu)^2}{2\sigma^2}\right] \quad \cdot \quad \cdot \quad \cdot (1)$$

& $$Y = \frac{N}{K\sigma\sqrt{2\pi}} \cdot \exp\left[\frac{-(x-M)^2}{2K^2\sigma^2}\right] \quad \cdot \quad \cdot \quad \cdot (2)$$

But for the normal distribution the maximum frequency is obtained against mode which is m here in both the cases.

$\therefore$ In case of eqution (1) we have

$$Y = \frac{N}{\sigma\sqrt{2\pi}} \exp\left[\frac{-(x-m)^2}{2\sigma^2}\right] \quad \cdot \quad \cdot \quad \cdot (3)$$

& $$Y = \frac{N}{K\sigma\sqrt{2\pi}} \exp\left[\frac{-(x-m)^2}{2K^2\sigma^2}\right] \quad \cdot \quad \cdot \quad \cdot (4)$$

$\therefore$ In case of equation (3), the maximum frequency is $= \frac{N}{\sigma\sqrt{2\pi}} = Y_0$ (say)

In case of equation (4) the maximum frequency is $= \frac{N}{K\sigma\sqrt{2\pi}} = \frac{1}{K} Y_0$

Hence the result

Example 10: If X and Y are independent normal variates with means 3, 4 and variances 4, 9 respectively, find the value of λ and show that

$$P(3x + 2y \le \lambda) = P(5X - 3Y \ge 3).$$

Solution: Let $U = 3X + 2Y$ and $V = 5X - 3Y$

Since X and Y are independent normal variates, Hence U and V are also normally distributed with $X \sim N(3, 4), \; Y \sim N(4, 9)$.

$$\therefore \; E(U) = 3\times 3 + 2\times 4 = 17, \; E(V) = 5\times 3 - 3\times 4 = 15 - 12 = 3$$

$$Var(U) = 3^2 \; V(X) + 2^2 \; V(Y) = 9\cdot 4 + 4\cdot 9 = 36 + 36 = 72$$

$$Var(V) = 5^2 \; V(X) + 3^2 \; V(Y) = 25\cdot 4 + 9\cdot 9 = 100 + 81 = 181$$

$$\therefore \; P(3X + 2Y \le \lambda) = P(U \le \lambda) = P\left[\frac{U-17}{\sqrt{72}} \le \frac{\lambda - 17}{\sqrt{72}}\right]$$

$$= P\left[Z \le \frac{\lambda - 17}{\sqrt{72}}\right]$$

Where $Z \sim N(0, 1)$

$$P(5X - 3Y \ge 3\lambda) = P(V \ge 3\lambda) = P\left[\frac{V-3}{\sqrt{181}} \ge \frac{3\lambda - 3}{\sqrt{181}}\right]$$

$$= P\left[Z > \frac{3\lambda - 3}{\sqrt{181}}\right]$$

From the given condition

$$P\left(Z \le \frac{\lambda - 17}{\sqrt{72}}\right) = P\left(Z \ge \frac{3\lambda - 3}{\sqrt{181}}\right) \quad \cdot \quad \cdot \quad \cdot (1)$$

Since $P(Z \le a) = P(Z \le b)$ then $a = -b$

$$\therefore \; \frac{\lambda - 17}{\sqrt{72}} = \frac{3\lambda - 3}{\sqrt{181}}$$

$$\Rightarrow 13\cdot 45\lambda - 228\cdot 71 = 25\cdot 456\lambda - 25\cdot 456$$

$$\Rightarrow \lambda = \frac{-203\cdot 254}{12\cdot 006} = -16\cdot 929$$

Using the value λ in (1) we have

$$P\left(Z \le \frac{-33\cdot 929}{8\cdot 485}\right) = P(Z \le -3\cdot 9987)$$

$$= P(Z \le +3\cdot 9987) \qquad (\because \text{ By symetry})$$

$$= \phi(3\cdot 9987) = 0\cdot 500 \quad \cdots (2)$$

$$(\because \text{ from Tables})$$

Thly $$P\left(Z \ge \frac{-53\cdot 786}{13\cdot 4536}\right) = P(Z \ge -3\cdot 9979)$$

$$= 1 - P(Z \le -3\cdot 9979)$$

$$= 1 - P(Z \le 3\cdot 9979)$$

$$= 1 - \phi(3\cdot 9979)$$

$$= 1 - 0\cdot 5$$

$$= 0\cdot 5 \quad \cdots (3)$$

from (2) & (3) we get

$$P(3X + 2Y \le \lambda) = P(5X - 3Y \ge 3\lambda)$$

Example 11: The linear measurements of the items of a product are approximately normally distributed with a mean of 20 cms and standard deviation of 4 cms. Items which measure between 18 cms and 23 cms are sold at 50p each and the other items at 30p, each. Findthe total amount collected if in all 10,000 items are sold. How many items are of measurements 26 cms or more?

Solution: Let X be the linear measures of the items of a product which is normally distributed with mean $\mu = 20$ cms and $S\cdot D\cdot \sigma = 4$ cms.

So when $X = 18$ cms the corresponding standard normal variate Z is given by

$$Z = \frac{X-\mu}{\sigma} = \frac{18-20}{4} = -0\cdot 5$$

Again for $X = 23$ cms the standard normal variate Z is given by

$$Z = \frac{X-\mu}{\sigma} = \frac{23-20}{4} = 0.75$$

$$\therefore \; P\{18 \le X \le 23\} = P\left\{\frac{18-20}{4} \le \frac{X-\mu}{\sigma} \le \frac{23-20}{4}\right\}$$

$$= P\{-0\cdot 5 \le Z \le 0\cdot 75\}$$

$$= P\{0 \le Z \le 0\cdot 5\} + P\{0 \le Z \le 0\cdot 75\}$$

$$= 0\cdot 1915 + 0\cdot 2734$$

$$= 0\cdot 4649$$

The area under the curve between $Z = -0\cdot 5 + 00\cdot 5 = 0\cdot 3830$. Since $\phi(Z) = Z\phi(Z)$, i.e., $\phi(-0\cdot 5) = 2\phi(0\cdot 5)$ and the area between $Z = -\,0\cdot 75 + 00\cdot 75$ is $0\cdot 5468$.

$\therefore$ The number of items which measure between 18 cms and 23 cms

$$= 10,000 \times P\{18 \le X \le 23\} = 10,000 \times 0\cdot 4649$$

$$= 4649$$

The other items $= 10,000 - 4649 = 5351$

Out of 10,000 items sold the total amount collected is

$$= 4649 \times 50 + 5351 \times 30 = 232450 + 160530$$

$$= \text{Rs. } 3929.80$$

Since the items measured between 18 cms and 23 cms are sold at 50 paise and the other items are sold at 30 paise each.

The number of items having measurement 26 cms or more

$$= P(X > 26) = P\left\{\frac{X-\mu}{\sigma} > \frac{26-20}{4}\right\} = P\{Z > 1\cdot 5\} = P\{1\cdot 5 < Z < \infty\}$$

$$= \phi(\infty) - \phi(1\cdot 5) = 1 - 0\cdot 4332 = 0\cdot 5668$$

Since the probability greater than $0 \cdot 5$ we have

$$P(X > 26) = 0 \cdot 5668 - 0 \cdot 5 = 0 \cdot 0668$$

$\therefore$ The number of items $= 0 \cdot 0668 \times 10{,}000 = 668$

Example 12: In a normal distribution 31% of the items are under 45 and 8% are over 64. Find the mean and S.D. of the distribution.

Solution: If μ be the mean and σ be the standard deviation of the distribution and given that 31% of items are under 45.

i.e., $P(X < 45) = 0 \cdot 31$

$$\Rightarrow P(X > 45) = 1 - 0 \cdot 31 = 0 \cdot 69$$

But probability greater than 0.5 therefore

$$P(X > 45) = 0 \cdot 69 - 0 \cdot 5 = 0 \cdot 19$$

Also 8% are over 64.

Example 13: Assume the mean height of soldurs to be 68.22 inches with variance of 10.8 square inches, how many soldurs in a reginent of 1000 would you expect to be of 6 feet tall? Given that the area under the standard normal curve between $x = 0$ and $x = 0 \cdot 35$ is $0 \cdot 1368$ and between $x = 0$ and $x = 1 \cdot 15$ is $0 \cdot 3746$.

Solution: Let X denote the height of an individual then by hypothesis $X \sim N(68 \cdot 22, 10 \cdot 8)$. If p denotes the probability that an individual is over 6 feet (i.e., 72 inches) then

$$p = P(X \geq 72) = P\left\{\frac{X - 68 \cdot 22}{\sqrt{10 \cdot 8}} \geq \frac{72 - 68 \cdot 22}{3 \cdot 28}\right\}$$

$$p = P(Z \geq 1 \cdot 15)$$

or $p = 1 - P(Z < 1 \cdot 15)$

$$= 1 - 0 \cdot 3749 - 0 \cdot 5$$

$$p = 0 \cdot 1254$$

From this we can say that the number of soldurs with height over 6 feet is

$$N = 1000\,p = 1000 \times 0 \cdot 1254 = 125 \cdot 4 = 125$$

Example 14: If X and Y are independent normal variates with the same mean μ (known to be less than 2) and standard deviatiors σ_1 and σ_2 respectuvely. Such that

$$P(4X - 3Y \le 6) + P(5X + 12Y \ge 30) = 1 \quad \text{and}$$

$$P(4X + 3Y \le 12) + P(5X - 12Y \le -20) = 1$$

determine the common mean and the ratio of the variances.

Solution: Let us suppose that

$$V(X) = \sigma_1^2 \text{ and } V(Y) = \sigma_2^2 \ ; \qquad E(X) = E(Y) = \mu$$

$$\therefore\ 4X - 3Y \sim N\left(\mu, 16\sigma_1^2 + 9\sigma_2^2\right)$$

$$5X + 12\,Y \sim N\left(17\mu,\ 25\sigma_1^2 + 144\sigma_2^2\right)$$

$$4X + 3Y \sim N\left[7\mu, 16\sigma_1^2 + 9\sigma_2^2\right]$$

$$5X - 12Y \sim N\left(-7\mu,\ 25\sigma_1^2 + 144\sigma_2^2\right)$$

Let us denote $16\sigma_1^2 + 9\sigma_2^2$ by α^2 and $25\,\sigma_1^2 + 144\,\sigma_2^2$ by β^2.

$$\therefore\ P(4X - 3Y \le 6) + P(5X + 12Y \ge 30)$$

$$\Rightarrow P\left(Z \le \frac{6-\mu}{\alpha}\right) + P\left(Z \ge \frac{30-17\mu}{\beta}\right) = 1$$

and $$P(4X + 3Y \le 12) + P(5X - 12 \le -20)$$

$$\Rightarrow P\left(Z \le \frac{12 - 7\mu}{\alpha}\right) + P\left(Z \le \frac{-20 + 7\mu}{\beta}\right) = 1$$

$$\therefore\ P\left(Z \le \frac{6-\mu}{\alpha}\right) = 1 - P\left(Z \ge \frac{30-17\mu}{\beta}\right) \quad \text{and}$$

$$P\left(Z \le \frac{12-7\mu}{\alpha}\right) = 1 - P\left(Z \le \frac{-23+7\mu}{\beta}\right)$$

Also $$P\left(Z \le \frac{6-\mu}{\beta}\right) = P\left(Z \le \frac{30-17\mu}{\beta}\right) \ \& \ P\left(Z \le \frac{12-7\mu}{\alpha}\right) = P\left(2 \ge \frac{-20+7\mu}{\beta}\right)$$

Hence $$\frac{6-\mu}{\alpha} = \frac{30-17\mu}{\beta} \quad \cdot \quad \cdot \quad \cdot (1)$$

$$\frac{12-7\mu}{\alpha} = \frac{-20+7\mu}{\beta} = \frac{20-7\mu}{\beta} \quad \cdot \quad \cdot \quad \cdot (2)$$

Solving (1) & (2)

$$\frac{\alpha}{\beta} = \frac{6-\mu}{30-17\mu} = \frac{12-7\mu}{20-7\mu}$$

$$\Rightarrow 112\,\mu^2 - 352\mu + 240 = 0$$

i.e., $(\mu-1)(112\mu - 240) = 0$

$$\therefore\ \mu = 1 \text{ and } \mu = \frac{240}{112} > 2$$

But we are given the $\mu < 2$. So the permissible value of $\mu = 1$

$$\therefore\ \frac{\alpha}{\beta} = \frac{6-\mu}{30-17\mu}\Big/ \mu = 1 \ = \frac{6-1}{30-17} = \frac{5}{13}$$

$$\frac{\alpha^2}{\beta^2} = \frac{25}{169} = 0\cdot 15$$

14.12 Exponential Distribution - Exercise:

1. If $X_1, X_2, \cdots\cdots, X_n$ are independent r.v.'s X_i haivng an exponential distribution with parameter θ_i, $i = 1, 2, \cdots\cdots, n$. Then prove that $Z = \min(X_1, X_2, \cdots\cdots, X_n)$ has an exponential distribution with parameter $\sum_{i=1}^{n} \theta_i$.

2. X and Y are independent with a common p.d.f.

$$f(x) = \begin{cases} e^{-x}, & x \ge 0 \\ 0, & x < 0 \end{cases}$$

Find a p.d.f. for $X - Y$. (Ans: $g(u) = \frac{1}{2} e^{-|u|}, -\infty < u < \infty$)

3. Let X and Y having common p.d.f. $\alpha \cdot e^{-\alpha x}, 0 < x < \infty, \alpha > 0$. Find the p.d.f. of

(i) X^3 (ii) $3 + 2X$ (iii) $X - Y$ and (iv) $|X - Y|$

Ans: (i) $\frac{\alpha}{3} x^{-2/3}$ **exp** $\left(-\alpha x^{1/3}\right)$ (ii) $\frac{\alpha}{2} e^{\frac{-\alpha(x-3)}{2}}$, $x > 3$

(iii) $\frac{\alpha}{2} e^{-\alpha|x|}$, $\forall x$ and (iv) $\alpha e^{-\alpha x} \le x > 0$

4. If X has exponential distribution with mean 2. Find $P(X < 1/X < 2)$.

5. Supose that during rainy season on a tropical island the length of the shower has an exponential distribution, with parameter $\lambda \;\; 2$, time being measured in minutes. What is the probability that a shower will last more than three minutes ? If a shower has already lasted for 2 minutes, what is the probability that it will last for atleast one more minute?

6. A continuous random variable X has the probability density function $f(x)$ given by

$$f(x) = A \cdot e^{-x/5}, \; x > 0$$

$$= 0, \quad \text{otherwise}$$

Find the value of A and show that for any two positive numbers s and t, $P\left[{X > S + t}/{X > S}\right] = P[X > t]$.

7. If $X \sim \text{Expo}(\lambda)$, find the value x_a such that $P\{X > x_a\} / P\{X \le x_a\} = a$.

8. Show that $Y = -\left(1/\lambda\right) \log F(X)$ is $\text{Expo}(\lambda)$.

9. Suppose the life of automobile batteries is exponentially distributed with parameter $\lambda = 0 \cdot 001$ days.

(a) What is the probability that a battery will last more than 1200 days?

(b) What is the probability that a battery will last more than 1200 days given that it has already served 1000 days?

[Ans: (a) 0.301, (b) 0.819]

10. The life time X in hours of a TV tube of certain type obeys an exponential distribution with $\lambda = 0 \cdot 001$ hrs. Find

(a) $P[X > 1000]$

(b) $P[700 \le X \le 1000]$ [Ans: (a) 0.368, (b) 0.129]

Lesson - 15

RECTANGULAR DISTRIBUTION

Objectives:

After studying this lesson the student is expected to have a clear comprehension of thetheory and the practical utility about the concepts of definition and properties such as m.g.f., c.g.f. characteristic function, moments upto fourth order of Rectangular Distribution.

Structure of The Lesson:

This lesson consists of the following sections as detailed below:

15.1 Introduction

15.2 Definition

15.3 Properties

15.4 Applications in Real Life

15.5 Workedout Examples

15.6 Exercise

15.1 Introduction:

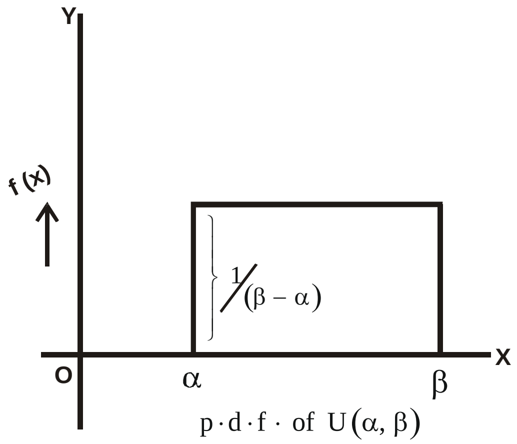
The simplest of all continuous random variables is the one in which the probability of its values is constant every where over an interval of the real line called uniform or rectangular distribution. We study this in the present lesson.

15.2 Definition:

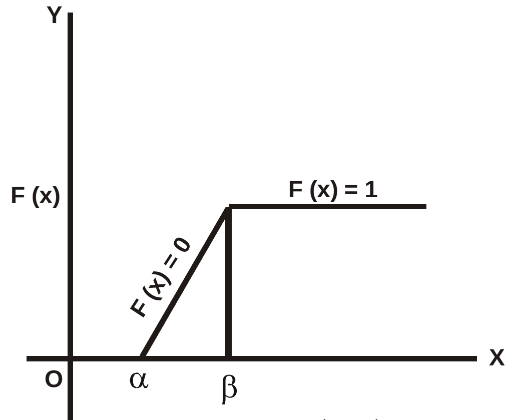
The random variable 'X' with a p.d.f. given by

$$\left.\begin{aligned} f(x) &= \frac{1}{\beta-\alpha}, \quad \alpha \le X \le \beta,\ \beta > \alpha \\ &= 0, \qquad \text{otherwise} \end{aligned}\right\} \cdots\cdots\cdots\cdots (1)$$

is said to be a uniform (or) rectangul;ar variable on the interval (α, β) and $f(x)$ given by (1) is called uniform or rectangular density function. If X has the density function (1), it is expressed as $X \sim U(\alpha, \beta)$. α and β are called parameters of the distribution the distribution is said to be rectangular because it takes the rectangular space and as the p.d.f. remains uniform for all the variate values within the range it is called uniform distribution.



p·d·f· of U(α, β)



c·d·f at U(α, β)

Distribution Function:

$F(x)$ is given by $F(x)=\begin{cases} 0 & , \text{if} \quad -\infty < x < \alpha \\ \dfrac{x-\alpha}{\beta-\alpha} & , \text{if} \quad \alpha \le x \le \beta \\ 0 & , \text{if} \quad \beta < x < \infty \end{cases}$

Since $F^1(x)=\dfrac{dF(x)}{dx}=\dfrac{1}{\beta-\alpha}=f(x)\neq 0$ in $\alpha \le x \le \beta$

$f(x)=0$, for $X>\beta$ and for $X<\alpha$, that is the density function is discontinuous at $x=\alpha$ & $x=\beta$.

15.3 Properties:

1. Moments:

If μ_r^1 and μ_r denote r^{th} moment about origin and central moments, then

$$\mu_r^1 = E(X^r) = \int_\alpha^\beta \frac{x^2}{\beta-\alpha}\,dx = \frac{1}{\beta-\alpha}\cdot\int_\alpha^\beta x^2\cdot dx = \frac{1}{\beta-\alpha}\,\frac{x^{r+1}}{r+1}\int_\alpha^\beta$$

$$\mu_r^1 = \frac{1}{(\beta-\alpha)}\left[\frac{\beta^{r+1}-\alpha^{r+1}}{r+1}\right] r = 1,2,3,\cdot \quad \cdot \quad \cdot (1)$$

For $r = 1, 2, 3, \cdot \quad \cdot \quad \cdot$ in equation (1) we get

$$\mu_1^1 = \text{mean} = E(X) = \frac{1}{(\beta - \alpha)} \; \frac{\beta^2 - \alpha^2}{2} = \frac{\cancel{(\beta - \alpha)} \; (\beta + \alpha)}{\cancel{(\beta - \alpha)} \; 2}$$

$$\mu_1^1 = \frac{\beta + \alpha}{2}$$

$$\mu_2^1 = E\left(X^2\right) = \frac{1}{(\beta - \alpha)} \; \frac{\left(\beta^3 - \alpha^3\right)}{3} = \frac{\cancel{(\beta - \alpha)} \left(\beta^2 + \alpha^2 + \beta\alpha\right)}{\cancel{(\beta - \alpha)} \quad 3}$$

$$\mu_2^1 = \frac{\beta^2 + \alpha^2 + \beta\,\alpha}{3}$$

$$\therefore \quad \text{Variance} \quad V(X) = E\left(X^2\right) - [E(X)]^2 = \mu_2^1 - \left(\mu_1^1\right)^2$$

$$V(X) = \frac{\beta^2 + \alpha^2 + \beta\alpha}{3} - \frac{(\beta - \alpha)^2}{4}$$

$$= \frac{4\beta^2 + 4\alpha^2 + 4\beta\alpha - 3\beta^2 - 3\alpha^2 - 64^3}{12}$$

$$= \frac{\beta^2 + \alpha^2 - 2\beta\alpha}{12}$$

$$= \frac{(\beta - \alpha)^2}{12}$$

$$\mu_3^1 = E\left(X^3\right) = \frac{1}{(\beta - \alpha)} \; \frac{\left[\beta^4 - \alpha^4\right]}{4} = \frac{\cancel{(\beta - \alpha)} \; (\beta + \alpha)\left(\beta^2 + \alpha^2\right)}{\cancel{(\beta - \alpha)} \quad 4}$$

$$= \frac{\beta^3 + \alpha^3 + \beta\alpha^2 + \alpha\beta^2}{4}$$

For finding μ_3 we have

$$\mu_3 = \mu_3^1 - 3\mu_2^1 \mu_1^1 + 2\mu_1^{1^3}$$

Now substituting thevalues of $\mu_3^1, \mu_2^1, \mu_1^1$ and after simplification we get

$$\mu_3 = \frac{\beta^3+\alpha^3+\beta\alpha^2+\alpha\beta^2}{4} - \frac{\not{3}\left(\beta^2+\alpha^2+\beta\alpha\right)}{\not{3}} \frac{(\beta+\alpha)}{2} + \frac{\not{2}(\beta+\alpha)^3}{\not{8}\ 4}$$

$$= \frac{\beta^3+\alpha^3+\beta\alpha^2+\alpha\beta^2}{4} + \frac{\beta^3+3\beta^2\alpha+3\beta\alpha^2+\alpha^3}{4} - \frac{\left[\beta^3+\alpha^3+2\beta^2\alpha+2\beta\alpha^2\right]}{2}$$

$$= \frac{2\beta^3+2\alpha^3+4\beta^2\alpha+4\beta\alpha^2}{4} - \frac{\left[\beta^3+\alpha^3+2\beta^2\alpha+2\beta\alpha^2\right]}{2}$$

$$= \frac{4\beta^3 + 4\alpha^3 + 8\beta^2\alpha + 8\beta\alpha^2 - 4\beta^3 - 4\alpha^3 - 8\beta^2\alpha - 8\beta\alpha^2}{8}$$

$$= 0$$

$$\mu_4^1 = E\left(X^4\right) = \frac{1}{(\beta-\alpha)} \frac{\left(\beta^5-\alpha^5\right)}{5}$$

$$\mu_4 = \mu_4^1 - 4\mu_3^1\mu_1^1 + 6\mu_2^1\mu_1^{1^2} - 3\mu_1^{1^4}$$

Now substituting the values of $\mu_4^1, \mu_3^1, \mu_2^1, \mu_1^1$ and after simplification we get

$$\mu_4 = \frac{(\beta-\alpha)^4}{80}$$

2. Pearson's Coefficient Skewness:

For uniform or rectangular distribution the skewness is

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = 0 \qquad \left(\because \mu_3 = 0,\ \mu_2 = \frac{\left(\beta-\alpha^2\right)}{12}\right)$$

or $\gamma_1 = \sqrt{\beta_1} = 0$

Since $\gamma_1 = 0$, the distribution $U(\alpha, \beta)$ is symmetric.

3. Pearson's Coefficient Kurtosis:

For uniform or rectangular distribution the Kurtosis is

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = 9/5 \qquad \left(\because \mu_4 \frac{(\beta-\alpha)^4}{80}, \ \mu_2 = \frac{(\beta-\alpha)^2}{12}\right)$$

$$\gamma_2 = \beta_2 - 3 = \frac{9}{5} - 3 = -6/5$$

Since $\gamma_2 < 0$ the distribution $U(\alpha, \beta)$ is platycurtic.

4. Moment Generating Function:

Accoridng to the definition of m.g.f. we have

$$M_X(t) = \int e^{tX} \cdot f(x)\,dx$$

If $X \sim U(\alpha, \beta)$ we have

$$M_X(t) = \int_\alpha^\beta e^{tX} \cdot \frac{1}{\beta-\alpha}\,dx$$

$$= \frac{1}{(\beta-\alpha)} \frac{e^{tX}}{t} \cdot \Big|_\alpha^\beta$$

$$= \frac{1}{t(\beta-\alpha)}\left(e^{t\beta} - e^{t\alpha}\right) = \frac{e^{t\beta} - e^{t\alpha}}{t(\beta-\alpha)}$$

5. Characteristic Function:

According to definition of characteristic function we have

$$\phi_X(t) = \int e^{itX} \cdot f(x)\,dx$$

If $X \sim U(\alpha, \beta)$ we have

$$\phi_X(t) = \int_\alpha^\beta e^{itX} \cdot \frac{1}{\beta-\alpha} dx \qquad \left(\because f(x) = \frac{1}{\beta-\alpha}\right)$$

$$= \frac{1}{(\beta-\alpha)} \frac{e^{i\beta t} - e^{i\alpha t}}{it}$$

$$\phi_X(t) = \frac{e^{i\beta t} - e^{i\alpha t}}{it\,(\beta-\alpha)}$$

6. Mean Deviation About Mean:

Mean deviation about mean for rectangular distribution is given by

$$M \cdot D \cdot = E(1X-M) = \int_\alpha^\beta \left| x - \frac{(\beta+\alpha)}{2} \right| \cdot \frac{1}{(\beta-\alpha)} dx$$

$$= \frac{1}{(\beta-\alpha)} \int_{-C}^{C} |Z|\, dZ \qquad \left(\because Z = x - \frac{(\beta+\alpha)}{2},\right.$$

$$x = \alpha,\ C = \frac{-(\beta-\alpha)}{2},$$

$$\left. x = \beta,\ C = \frac{(\beta-\alpha)}{2}.\right)$$

$$= \frac{2}{(\beta-\alpha)} \int_0^C Z\, dZ = \frac{\cancel{2}}{(\beta-\alpha)} \left.\frac{Z^2}{\cancel{2}}\right|_0^C = \frac{C^2 - 0}{(\beta-\alpha)} = \frac{C^2}{(\beta-\alpha)}$$

$$\because C = \frac{(\beta-\alpha)}{2}$$

We have $M \cdot D \cdot = \frac{C^2}{(\beta-\alpha)} = \frac{(\beta-\alpha)^{\cancel{2}}}{4} \cdot \frac{1}{\cancel{(\beta-\alpha)}} = \frac{(\beta-\alpha)}{4}$

7. Median:

If m is the median then

$$\int_{\alpha}^{m} f(x)\,dx = \frac{1}{2}$$

$$\Rightarrow \int_{\alpha}^{m} \frac{1}{\beta-\alpha}\,dx = 1/2 \qquad \left(\because\ f(x) = \frac{1}{\beta-\alpha}\right)$$

$$\Rightarrow \frac{1}{\beta-\alpha}\int_{\alpha}^{m} dx = \frac{1}{2}$$

$$\Rightarrow \frac{1}{\beta-\alpha}\, x \Big|_{\alpha}^{m} = \frac{1}{2}$$

$$\Rightarrow \frac{1}{\beta-\alpha}(m-\alpha) = \frac{1}{2}$$

$$\Rightarrow m = \frac{1}{2}(\beta+\alpha)$$

8. Mode:

Since in the rectangular distribution $(b-a)\times(b-a)^{-1}$ each point in the internal (a, b) has the maximum probability it follows that each point of (a, b) is a mode.

9. If X and Y are independently and identically distributed (i. i. d.) rectangular or uniform $U(0,1)$, the distribution of the variates $(x+y)$, $(x-y)$, xy and x/y are as follows:

$$f(x+y) = \begin{cases} x+y, & 0 \le x+y \le 1 \\ 2-(x+y), & 1 \le x+y \le 2 \end{cases}$$

$$f(x-y) = \begin{cases} x-y+1, & -1 \le x- \le 0 \\ 1-(x-y), & 0 \le x-y \le 1 \end{cases}$$

$$f(xy) = -\ell og\,(xy),\ 0 < xy < 1$$

$$f\left(\tfrac{x}{y}\right) = \begin{cases} \tfrac{1}{2}, & 0 < \tfrac{x}{y} < 1 \\ \tfrac{y^2}{2x^2}, & 1 < \tfrac{x}{y} < \infty \end{cases}$$

15.5 Worked Examples:

Example 1: Obtain the m.g.f. for the rectangular distribution

$$f(x) = 1$$

Hence obtain mean, variance, mean deviation.

Solution: The p.d.f. of the rectangular distribution is

$$f(x) = 1 \quad \cdot \quad \cdot \quad \cdot (1)$$

By the definition of m.g.f., we have

$$M_X(t) = E\left(e^{tX}\right) = \int e^t x \cdot f(x)\, dx$$

If $X \sim U(0, 1)$ we have

$$M_X(t) = \int_0^1 e^{tX}\, dx = \frac{e^{tX}}{t} \int_0^1 = \frac{e^t - e^0}{t} = \frac{e^t - 1}{t} \quad \cdot \quad \cdot \quad \cdot (2)$$

for mean we have

$$\mu_1^1 = E(X) = \int_0^1 x\, dx = \frac{x^2}{x} \int_0^1 = \frac{1^2 - 0}{2} = \tfrac{1}{2}$$

& $$\mu_2^1 = E\left(X^2\right) = \int_0^1 x^2\, dx = \frac{x^3}{3} \int_0^1 = \frac{1^3 - 0}{3} = \frac{1}{3}$$

Hence variance μ_2 is given by

$$\mu_2 = \mu_2^1 - \left(\mu_1^1\right)^2 = \frac{1}{3} - \left(\frac{1}{2}\right)^2 = \frac{1}{3} - \frac{1}{4} = \frac{1}{12}$$

Mean deviation about the mean is

$$E\left(\left|X-M\right|\right) = \int_0^1 \left|x-M\right| \cdot f(x)\,dx = \int_0^1 \left|x-\frac{1}{2}\right| dx = \int_0^{1/2}\left(\frac{1}{2}-x\right)dx + \int_{1/2}^1\left(x-\frac{1}{2}\right)dx$$

$$= \frac{1}{2}x\Big|_0^{1/2} - \frac{x^2}{2}\Big|_0^{1/2} + \frac{x^2}{2}\Big|_{1/2}^1 - \frac{1}{2}\cdot x\Big|_{1/2}^1$$

$$= \frac{1}{4} - \frac{1}{8} + \cancel{\frac{1}{2}} - \frac{1}{8} - \cancel{\frac{1}{2}} + \frac{1}{4}$$

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

Example 2: A circular is marked from 1 to 10. It has a balanced pointer pivoted at the centre such that when it is whirled and allowed to stop, it is equally likely to stop any where. What is the probbility that in 2 out of 3 trials the pointer stops in between 2 and 3?

Solution: Let X be the distance of the stopping point measured from the zero point, implies that X has rectangular distribution with the p.d.f.

$$\left.\begin{aligned} f(x) &= \frac{1}{10}, \quad 0 \le x \le 10 \\ &= 0, \quad \text{otherwise} \end{aligned}\right\} \quad \cdot \quad \cdot \quad \cdot (1)$$

$\therefore$ The probability that the pointer stops in between 2 and 3 at any single trial is

$$\int_2^3 f(x)\,dx = \int_2^3 \frac{1}{10}dx \qquad (\because \text{ from } (1))$$

$$= \frac{1}{10} \int_2^3 1 dx$$

$$= \frac{1}{10} x \int_2^3 = \frac{1}{10}(3-2) = \frac{1}{10} = p \quad (say)$$

If the following of the pointer between 2 & 3 gives probability of success as 'p', then it will be same for all the trials.

$\therefore$ The probability of 2 successes out of 3 trials

$$= {}^3C_2 \cdot \left(\frac{1}{10}\right)^2 \left(1 - \frac{1}{10}\right)^{3-2}$$

$$= 3 \times \frac{1}{10} \times \frac{1}{10} \times \frac{9}{10}$$

$$= \frac{27}{1000}$$

$$= 0 \cdot 027$$

Example 3: If X is $U(1, 2)$ find Z such that $P(X > Z + M_x) = \frac{1}{4}$

Solution: Since $X \sim U(1, 2)$ its p.d.f. is given by

$$\left.\begin{aligned} f(x) &= 1, \quad 1 < x < 2 \\ &= 0, \quad \text{otherwise} \end{aligned}\right\} \quad \cdot \quad \cdot \quad \cdot (1)$$

$$\mu_X = E(X) = \int_1^2 x\, dx = \frac{x^2}{2} \int_1^2 = \frac{2^2 - 1}{2} = \frac{4-1}{2} = \frac{3}{2}$$

Also given that $P(X > z + M_X) = \frac{1}{4}$

$$\Rightarrow p\left(X > z + \frac{3}{2}\right) = \int_{z+\frac{3}{2}}^{2} dx = \left(\frac{1}{2} - Z\right) = \frac{1}{4}$$

$$\therefore \frac{1}{2} - Z = \frac{1}{4} \quad \text{or} \quad Z = \frac{1}{2} - \frac{1}{4} = \frac{1}{2}$$

Example 4: If X is $U(-b, b)$, determine b if $P(1 \times 1 > 2) = 3/4$.

Solution: Since $X \sim U(-b,\ b)$ its p.d.f. is

$$f(x) = 1/2b, \quad -b < x < b$$

$$= 0, \quad \text{otherwise}$$

So $\frac{3}{4} = P(|X| > 2) = 1 - P(|X| \le 2) = 1 - P(-2 \le X \le 2)$

$$= 1 - \int_{-2}^{2} \frac{dx}{2b} = 1 - \frac{1}{2b}[x]_{-2}^{2} = 1 - \frac{1}{2b}[\ 2 - (-2)] = 1 - \frac{\not{4}\,2}{\not{2}\,b}$$

$$\frac{3}{4} = 1 - \frac{2}{b} \Rightarrow \frac{2}{b} = 1 - \frac{3}{4} = 1/4 \Rightarrow \frac{2}{b} = \frac{1}{4} \text{ or } b = 8$$

Example 5: Let X an Y be i.i.d. $U(a, b)$ variates and $K \in (a, b)$. Find the number K such that the probability that at least one of X and Y exceeds K is p.

Solution: Let $A = \{X > K\},\ B = \{Y > K\}$; then

$$P(A) = P(B) = \int_{K}^{b} \frac{dx}{b-a} = \frac{b-K}{b-a} \quad ; \quad P(\overline{A}) = \frac{K-a}{b-a} \cdot \quad \cdot \quad \cdot (1)$$

$$p = p(A \cup B) = 1 - p(\overline{A}) \cdot p(\overline{B}) = 1 - \left[(K-a)/(b-a)\right]^2$$

$$\therefore\ (K-a)/(b-a) = \sqrt{1-p} \Rightarrow K = a + (b-a)\sqrt{1-p}\,.$$

Example 6: A variate X has the uniform distribution with the density function given by

$$f(x) = \frac{1}{100}, \quad 0 < x < 100$$

$$= 0, \quad \text{otherwise}$$

compute $p(X > 60)$ and $P(20 \le x \le 40)$

Solution: Variate X has the uniform distribution with the density function given by

$$\left.\begin{aligned} f(x) &= \frac{1}{100}, \quad 0 < \ < 100 \\ &= 0, \qquad \text{otherwise} \end{aligned}\right\} \quad . \quad . \quad \cdot (1)$$

$$P(X>60) = \int_{60}^{100} f(x)\,dx = \int_{60}^{100} \frac{1}{100}\,dx = \frac{1}{100}[x]_{60}^{100} = \frac{1}{100}[100-60]$$

$$= \frac{40}{100} = 0 \cdot 4$$

$$P(20 \le x \le 40) = \int_{20}^{40} f(x)\,dx = \int_{20}^{40} \frac{1}{100}\,dx = \frac{1}{100}[x]_{20}^{40} = \frac{1}{100}[40-20]$$

$$= \frac{20}{100} = 0 \cdot 20$$

Example 7: The variates a nad b are independently and uniformly distributed in the intervals $[0, 6]$ and $[0, 9]$ respectively. Find the probability that $x^2 - ax + b = 0$ has two real roots.

Solution: Solving the given quadratic equation we get

$$x = \frac{1}{2}\left[\ a \pm \sqrt{a^2 - 4b}\right]$$

The roots are imaginary if $b > a^2/4$

$$\therefore\ P\left(b > \frac{a^2}{4}\right) = \int_0^6 \int_{a^2/4}^{9} \frac{1}{\sqrt{4}}\,da\,db \qquad \left[\because\ f(a, b) = f(a) \cdot f(b) = \frac{1}{6} \cdot \frac{1}{9}\right]$$

$$= 1 - \left(1/3\right) = 2/3$$

$$\therefore\ P\left(b \le\ a^2/4\right) = P\left(\text{Roots are real}\right) = 1/3$$

15.6 Exercise:

1. If X is a random variable with a continuous distributuion function F, then $F(X)$ has a uniform distribution on $[0, 1]$.

2. Show that for the rectangular distribution

$$f(x) = \frac{1}{2a}, -a < x < a$$

The m.g.f. about origin is $\frac{1}{at}$ (sinhaf). Also show that moments of even order are given by $\mu_{2n} = \frac{a^{2n}}{(2n+1)}$.

3. If X is uniformly distributed with mean 1 and variance $4/3$ find $P(X < 0)$.

$\left(\text{Ans}: 1/4\right)$

4. If X has a uniform distribution in $[0, 1]$ find the distribution (p.d.f.) of $-2 \ell og X$. Identify the distribution also.

5. Subway trains on a certain line run every $1/2$ hour between mid - night and six in the morning. What is the probability that a man entering the station at a random time during this period will have to wait at least twenty minutes ?

6. For the rectangular distribution

$$dF = K dx, 1 \le x \le 2$$

Show that Arithmetic Mean > Geometric Mean > Harmonic Mean.

7. Assume a string, 1 meter long, is to be cut in two at a random point along its length. Let X be the point where cut occurs and let its p.d.f. be

$$f(x) = 1, 0 < x < 1$$

what is the probability that the longer piece is at least twice the length of the shorter?

8. If X is $U(-3, 3)$ find K such that $P(X < K) = 1/3$. Also compute $P(X(2))$; $P\{|X-2| < 2\}, P(X=2)$ and $P\{|X| < 2\}$.

9. Suppose that X is uniformly distributed over $(-\alpha, \alpha)$ where $\alpha > 0$ determine α so that $P(|X| < 1) = P(|X| > 1)$. $P(X > 1) = 1/3$, $P\left(X < 1/2\right) = 0 \cdot 3$.

$\left(\text{Ans}: \alpha = 2, \alpha = 3, \alpha = 5/6\right)$

Lesson - 16

EXPONENTIAL DISTRIBUTION

Object of The Lesson:

After studying this lesson the student is expected to have a clear comprehension of the theory and the practical utility about the concepts, properties of exponential distribution.

Structure of The Lesson:

This lesson consists of the following sections as detailed below:

16.1 Introduction

16.2 Definition

16.3 Properties

16.4 Workedout Examples

16.5 Applications in Real Life

16.6 Exercise

16.1 Introduction:

In the theory of continuous distributions we commonly say normal distribution is an example of model for any data. Because its range is spread over $\left(-\infty, +\infty\right)$, for data that is generally positive valued, application of normal distribution is not suitable sometimes. Then the simple model that can be useful is exponential distribution. The random variable of exponential distribution is positive valued. Like normal distribution it has many smooth properties. It can be used on a good model for life time of a number of industrial products. In this lesson we study a number of theoretical aspects of exponential distribution.

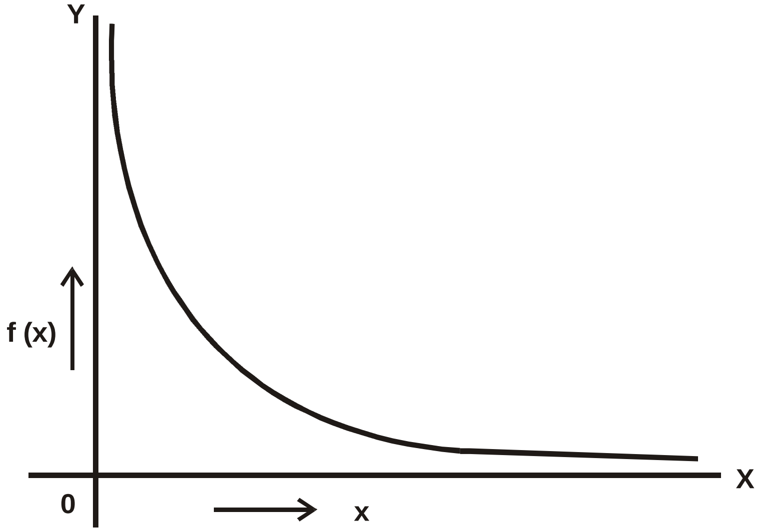
16.2 Definition:

The continuous random variable X which is distributed according to the probability law

$$f(x) = \lambda \cdot e^{-\lambda x}, \quad 0 < x < \infty, \ \lambda > 0 \quad \cdot \quad \cdot \quad \cdot (1)$$

$$= 0, \ \text{otherwise}$$

is called the exponential variate with parameter λ, $f(x)$ is called the probability density function of exponential distribution. Any variate having the p.d.f. (1) is expressed as $X \sim \text{Expo}(\lambda)$.

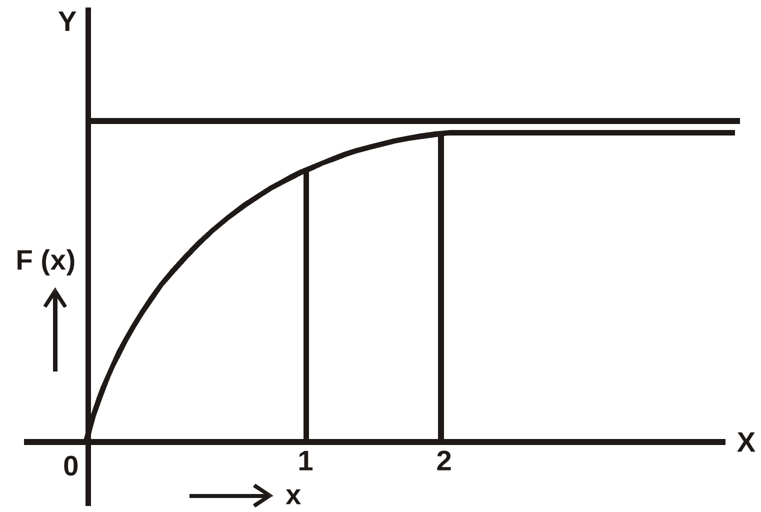


Exponenutial Curve

The cumulative distribution function $F(x)$ is given by

$$F(x) = \int_0^x \lambda\, e^{-\lambda t}\, dt = 1 - e^{-\lambda x},\ x > 0$$

$$\therefore\ F(x) = \begin{cases} 1 - e^{-\lambda x}, & x > 0 \\ 0 \leq & \text{otherwise} \end{cases}$$



Exponential Distribution Function

16.3 Properties:

1. Mean of exponential distribution is $1/\lambda$.

We have $\mu_1 = E(x) = \int_0^\infty x \cdot f(x)\, dx$

$$= \int_0^\infty x \cdot \lambda \cdot e^{-\lambda x}\, dx \qquad \left(\because\ f(x) = \lambda \cdot e^{-\lambda x}\right)$$

$$= \lambda \cdot \int_0^\infty x \cdot e^{-\lambda x}\, dx$$

$$= \lambda \left[x \cdot \int_0^\infty e^{-\lambda x}\, dx - \int_0^\infty \left[\left(\frac{d}{dx} x \right) \cdot \int e^{-\lambda x}\, dx \right] \right] dx$$

$$= \lambda \cdot \left[\frac{x \cdot e^{-\lambda x}}{-\lambda} \Big|_0^\infty + \int_0^\infty \frac{e^{-\lambda x}}{\lambda}\, dx \right]$$

$$= 0 + \cancel{\lambda} \cdot \int_0^\infty \frac{e^{-\lambda x}}{\cancel{\lambda}}\, dx = \frac{-e^{-\lambda x}}{\lambda} \Big|_0^\infty = \frac{-1}{\lambda}\left(e^{-\infty} - e^{0}\right)$$

$$\mu_1 = E(x) = 1/\lambda.$$

2. Variance of Exponential Distribution is $1/\lambda^2$

We know $V(x) = E\left(x^2\right) - [E(x)]^2$

or $\mu_2 = \mu_2^1 - \left(\mu_1^1\right)^2$

But $E(x) = 1/\lambda \quad \cdot \quad \cdot \quad \cdot (1)$ & $E\left(x^2\right)$ is given by

$$E(x^2) = \mu_2^1 = \int_0^\infty x^2 \cdot f(x)\,dx = \int_0^\infty x^2 \cdot \lambda \cdot e^{-\lambda x}\,dx$$

$$= \lambda \cdot \int_0^\infty x^2 \cdot e^{-\lambda x}\,dx = \lambda\left[x^2 \cdot \int_0^\infty e^{-\lambda x}\,dx - \int_0^\infty \left[\left(\frac{d}{dx}x^2\right) \cdot \int_0^\infty e^{-\lambda}\,dx\right]\right]dx$$

$$= \chi \cdot \left[\frac{x^2 \cdot e^{-\lambda x}}{-\chi}\int_0^\infty - \chi\int_0^\infty \frac{2x \cdot e^{-\lambda x}}{-\chi}\,dx = -\left(x^2 \cdot e^{-\lambda x}\right)_0^\infty + 2 \cdot \int_0^\infty x \cdot e^{-\lambda x}\,dx\right]$$

$$= 0 + 2 \cdot \left[x \cdot \int_0^\infty e^{-\lambda x}\,dx - \int_0^\infty \left(\left(\frac{d}{dx}x\right) \cdot \int e^{-\lambda x}\,dx\right)dx\right]$$

$$= 2\,\frac{x \cdot e^{-\lambda x}}{-\lambda}\int_0^\infty + 2\int_0^\infty \frac{e^{-\lambda x}}{\lambda}\,dx = 0 + 2 \cdot \frac{e^{-\lambda x}}{-\lambda^2}\int_0^\infty$$

$$E(x^2) = \frac{-2}{\lambda^2}\left(e^{-\infty} - e^0\right) = 2/\lambda^2\,[-1] = 2/\lambda^2 \quad \cdot \quad \cdot \quad \cdot (2)$$

$$\therefore \quad \mu_2^1 \quad = \quad E(x^2) = 2/\lambda^2$$

Hence $V(x) = E(x^2) - [E(x)]^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = 1/\lambda^2$ (from (1) & (2))

$$\therefore \mu_2 = V(x) = 1/\lambda^2$$

3. Moments of all order exist.

The r^{th} moment about origin is given by

$$\mu_r^1 = E\left(X^r\right) = \int_0^\infty \lambda \cdot e^{-\lambda x} \cdot x^r \, dx$$

$$= \lambda \cdot \int_0^\infty e^{-\lambda x} \cdot x^{(r+1)-1} \, dx$$

$$= \lambda \cdot \frac{\Gamma(r+1)}{\lambda^{r+1}} \qquad \left(\because \int_0^\infty e^{-ky} \, y^{n-1} dy = \frac{\Gamma(n)}{k^n}\right)$$

$$= \frac{\Gamma(r+1)}{\lambda^r}$$

$$\mu_r^1 = \frac{r!}{\lambda^r} \qquad \left(\because \Gamma(n) = (n-1)!\right)$$

thus for $r = 1, 2, 3, 4$ are

$$\mu_1^1 = \frac{1}{\lambda}, \quad \mu_2^1 = 2/\lambda^2, \ \mu_2 = \mu_2^1 - {\mu_1^1}^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = 1/\lambda^2$$

$$\mu_3^1 = \frac{6}{\lambda^3}, \quad \mu_3 = \mu_3^1 - 3\mu_2^1 \mu_1^1 + 2{\mu_1^1}^3 = \frac{6}{\lambda^3} - 3 \cdot \frac{2}{\lambda^2} \cdot \frac{1}{\lambda} + 2 \cdot \frac{1}{\lambda^3}$$

$$\mu_3 = \frac{6}{\lambda^3} - \frac{6}{\lambda^3} + 2/\lambda^3 = 2/\lambda^3$$

$$\mu_4^1 = \frac{24}{\lambda^4}, \quad \mu_4 = \mu_4^1 - 4\,\mu_3^1\,\mu_1^1 + 6\mu_2^1\,{\mu_1^1}^2 - 3{\mu_1^1}^4$$

$$= \frac{24}{\lambda^4} - 4 \cdot \frac{6}{\lambda^3} \cdot \frac{1}{\lambda} + 6 \cdot \frac{2}{\lambda^2} \cdot \frac{1}{\lambda^2} - 3 \cdot \frac{1}{\lambda^4}$$

$$= \frac{24}{\lambda^4} - \frac{24}{\lambda^4} + \frac{12}{\lambda^4} - 3/\lambda^4 = 9/\lambda^4$$

5. Characteristic Function:

The characteristic function of a random variable X having the probability function $f(x)$ is given by

$$\phi_X(t) = \int e^{itx} \cdot f(x)\, dx$$

$\because\ X \sim \text{Expo}(\lambda)$ we have

$$\phi_X(t) = \int_0^\infty e^{itx} \cdot \lambda \cdot e^{-\lambda x}\, dx$$

$$= \lambda \cdot \int_0^\infty e^{-(\lambda - it)x}\, dx$$

$$= \lambda \cdot \frac{\left. e^{-(\lambda - it)x} \right|_0^\infty}{-(\lambda - it)}$$

$$= \frac{-\lambda}{\lambda - it}\left(e^{-\infty} - e^0\right) \qquad \left(\because\ e^{-\infty} = 0,\ e^0 = 1\right)$$

$$\phi_X(t) = \frac{\lambda}{(\lambda - it)} = \left(1 - {it}/{\lambda}\right)^{-1}$$

6. Cumulant Generating Function:

The cumulatn generating function of a random variable 'X' is given by

$$K_X(t) = \log M_X(t)$$

If $X \sim \text{Expo}(\lambda)$

$$K_X(t) = \log\left(1 - {t}/{\lambda}\right)^{-1} \qquad \left(\because M_X(t) = \left(1 - {t}/{\lambda}\right)^{-1}\right)$$

$$= -1 \log\left(1 - {t}/{\lambda}\right) = -1\left[\frac{-t}{\lambda} - \frac{t^2}{\lambda^2} \cdot \frac{1}{2} - \frac{t^3}{\lambda^3} \cdot \frac{1}{3} - \frac{t^4}{\lambda^4} \cdot \frac{1}{4} \cdot \quad \cdot \quad \cdot\right]$$

$$= t/\lambda + \frac{t^2}{2} \cdot \frac{1}{\lambda^2} + \frac{t^3}{3} \cdot \frac{1}{\lambda^3} + \frac{t^4}{4} \cdot \frac{1}{\lambda^4} + \cdot \quad \cdot \quad \cdot$$

$$\therefore \quad \mu_1 = 1/\lambda,\ \mu_2 = 1/\lambda^2,\ \mu_3 = 2/\lambda^3,\ \mu_4 = 9/\lambda^4$$

If $\lambda > 1$, mean is greater than variance

$\lambda < 1$, mean is less than variance

$\lambda = 1$, mean is equal to variance

This is an important feature of this distribution.

7. Moment Generating Function:

The moment generating function (.g.f.) of a random variable X (about origin) having the probability function $f(x)$ is given by

$$M_X(t) = E\left(e^{tx}\right) = \int e^{tx} \cdot f(x)\, dx$$

$\because \quad X \sim \text{Expo}(\lambda)$ we have

$$M_X(t) = \lambda \int_0^\infty \cdot e^{tx} \cdot e^{-\lambda x}\, dx$$

$$= \lambda \cdot \int_0^\infty e^{-(\lambda - t)x}\, dx$$

$$= \lambda \cdot \frac{e^{-(\lambda-t)x}}{-(\lambda - t)} \int_0^\infty$$

$$= \frac{\lambda}{-(\lambda - t)}\left(e^{-\infty} - e^0\right)$$

$$= \frac{\lambda}{(\lambda - t)} = \left(1 - t/\lambda\right)^{-1}$$

Using the binomial expansion we get

$$M_X(t) = \sum_{r=0}^{\infty}\left(\frac{t}{\lambda}\right)^r$$

$$\mu_r^1 = E\left(X^2\right) = \text{ coefficient of } \frac{t^r}{r!} \text{ in } M_X(t)$$

$$= \frac{r!}{\lambda^r} \;;\; r = 1, 2, \cdots\cdots$$

$$K_X(t) = -\left[-\frac{t}{\lambda} - \left(\frac{t}{\lambda}\right)^2 \frac{1}{2} - \left(\frac{t}{\lambda}\right)^3 \cdot \frac{1}{3} - \left(\frac{t}{\lambda}\right)^4 \cdot \frac{1}{4} \cdots\cdots\cdots \left(\frac{t}{\lambda}\right)^r \cdot \frac{1}{r}\right]$$

$$\left(\because \ell og\,(1-x) = -x - \frac{x^2}{2} - \frac{x^3}{3} - \frac{x^4}{4} \cdots\cdots\cdots \frac{(-1)^r x^r}{r} \cdots\cdots\right)$$

$$K_X(t) = \frac{t}{\lambda} + \left(\frac{t}{\lambda}\right)^2 \frac{1}{2} + \left(\frac{t}{\lambda}\right)^3 \cdot \frac{1}{3} + \cdots\cdots\cdots \left(\frac{t}{\lambda}\right)^r \cdot \frac{1}{r} + \cdots\cdots$$

$$= \sum_{r=1}^{\infty}\left(\frac{t}{\lambda}\right)^r \cdot \frac{1}{r} = \sum_{r=1}^{\infty} \frac{(r-1)!}{\lambda^r} \cdot \frac{t^r}{r!}$$

$\therefore$ coefficient of $\frac{t^r}{r!}$ in the above expression is $K_r(t) = \frac{(r-1)!}{\lambda^r}$

Thus $K_1(t) = \frac{1}{\lambda}$, $K_2(t) = \frac{1}{\lambda^2}$, $K_3(t) = \frac{2}{\lambda^3}$, $K_4(t) = \frac{6}{\lambda^4} \cdots\cdots$

8. Therelationship between central moments and cumulants are

$$\mu_1 = K_1(t),\ \mu_2 = K_2(t),\ \mu_3 = K_3(t),\ \mu_4 = K_4(t) + 3 \cdot K_2^2(t)$$

9 (a). Pearson's Measure of Skewness:

Since $\mu_1 = K_1(t) = 1/\lambda$, $\mu_2 = K_2(t) = 1/\lambda^2$,

$$\mu_3 = K_3(t) = 2/\lambda^3, \quad \mu_4 = K_4(t) + 3K_2^2(t) = 9/\lambda^4$$

We have

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{\left(2/\lambda^3\right)^2}{\left(1/\lambda^2\right)^3} = \frac{4}{\lambda^6} \times \frac{\lambda^6}{1} = 4$$

or $\gamma_1 = \sqrt{\beta_1} = \sqrt{4} = 2$

(b). Pearson's Measure of Kurtosis:

We have $\beta_2 = \dfrac{\mu_4}{\mu_2^2} = \dfrac{9/\lambda^4}{\left(1/\lambda^2\right)^2} = \dfrac{p}{\lambda^4} \times \dfrac{\lambda^4}{1} = 9$

$$\left(\because \mu_4 = 9/\lambda^4, \ \mu_2 = 1/\lambda^2\right)$$

or $\gamma_2 = \beta_2 - 3 = 9 - 3 = 6$

The values of β_1 or γ_1 and β_2 or γ_2 clearly staks that exponential distribution is positively skewed and is teptokurtic.

10. Median of exponential distribution:

If m is the median of exponential distribution then

$$\frac{1}{2} = \int_0^m \lambda \cdot e^{-\lambda x} \, dx = 1 - e^{-\lambda m}$$

$$\Rightarrow e^{-\lambda m} = 1/2$$

$$\Rightarrow m = \lambda^{-1} \cdot \ell og_e^2 .$$

11. Relation with uniform distribution

If $X \sim \text{Expo}(\lambda)$, then $Y = e^{-\lambda x}$ is $U(0,1)$

i.e., consider the robbaility differential of exponential variate

$$dF_1(x) = \lambda \cdot e^{-\lambda x}\, dx \;,\quad 0 \le x < \infty \quad \cdot \quad \cdot \quad \cdot (1)$$

Put $y = e^{-\lambda x}, \quad dy = -\lambda \cdot e^{-\lambda x} \cdot dx$

if $x = 0 \;;\; y = 1$

$x = \infty \;;\; y = 0$

Hence changing the sign $dF_2(y) = 1dy, \;\; 0 \le y \le 1$

i.e., $Y \sim U(0, 1)$

12. Memory - less Property:

The p.d.f. of the exponential distribution with parameter λ is

$$f(x) = \lambda \cdot e^{-\lambda x} \cdot \lambda > 0, 0 < x < \infty$$

We have

$$P(Y \le x \cap X \ge a) = P(X - a \le x \cap X \ge a) \qquad (\because \; Y = X - a)$$

$$= P(X \le a + x \cap X \ge a)$$

$$= P(a \le X \le a + x)$$

$$= \lambda \int_a^{a+x} e^{-\lambda x}\, dx$$

$$= \chi \cdot \frac{e^{-\lambda x}}{-\chi} \cdot \int_a^{a+x}$$

$$= -\left(e^{-\lambda(a+x)} - e^{-\lambda a}\right)$$

$$= e^{-a\lambda}\left(1 - e^{-\lambda x}\right)$$

and $P(X \geq a) = \lambda \cdot \int_a^\infty \cdot e^{-\lambda x}\, dx$

$$= \chi \cdot \frac{e^{-\lambda x}}{-\chi}\Big|_a^\infty = -\left(e^{-\infty} - e^{-a\lambda}\right) = e^{-a\lambda}$$

$$\therefore\ p\left(\frac{Y \leq x}{X \geq a}\right) = \frac{P(Y \leq x \cap X \geq a)}{P(X \geq a)} = \frac{e^{-a\lambda}\left(1 - e^{-\lambda x}\right)}{e^{-a\lambda}}$$

$$= \left(1 - e^{-\lambda x}\right) \quad \cdot \quad \cdot \quad \cdot (1)$$

Also $P(X \leq x) = \lambda \cdot \int_0^x e^{-\lambda x}\, dx = 1 - e^{-\lambda x} \quad \cdot \quad \cdot \quad \cdot (2)$

$\therefore$ from (1) & (2) we get

$$P\left(\frac{Y \leq x}{X \geq a}\right) = P(X \leq x)$$

i.e., exponential distribution lacks memory.

16.4 Workedout Examples:

Example 1: Show that for the exponential distribution

$dP(x) = y_0 \cdot e^{-x/\sigma}\, dx,\ \ 0 \leq x < \infty\ ;\ \sigma > 0\ ,\ \ y_0$ is a constant.

mean and variance are equal. Also obtain the interquantile range of the distribution.

Solution: In order to change the given distribution into probability density function, we must have

$$\int_0^\infty y_0 \cdot e^{-x/\sigma}\, dx = 1$$ ($\because$ the total area under curve is unity)

$$\Rightarrow\ y_0 \cdot \int_0^\infty e^{-x/\sigma}\, dx = 1$$

$$\Rightarrow y_0 \cdot \sigma \cdot \left[-e^{-x/\sigma}\right]_0^{\infty} = 1 \Rightarrow y_0 \cdot \sigma\left[-\left(e^{-\infty} - e^{0}\right)\right] = 1$$

$$\Rightarrow y_0 \cdot \sigma = 1 \Rightarrow y_0 = 1/\sigma$$

Hence the p.d.f. $f(x) = \frac{1}{\sigma} \cdot e^{-x/\sigma}, \ 0 \le x \le \infty, \sigma > 0$

Mean $= E(X) = \mu_1^1 = \int_0^{\infty} x \cdot f(x)\, dx = \int_0^{\infty} x \cdot \frac{1}{\sigma} \cdot e^{-x/\sigma}\, dx$

$$= \frac{1}{\sigma}\left[x \cdot \frac{e^{-x/\sigma}}{-1/\sigma}\right]_0^{\infty} - \frac{1}{\sigma}\int_0^{\infty} \cdot \frac{e^{-x/\sigma}}{-1/\sigma}\, dx$$

($\because$ By integrating by parts)

$$= 0 + \int_0^{\infty} e^{-x/\sigma}\, dx$$

$$= \left(\frac{e^{-x/\sigma}}{-1/\sigma}\right)_0^{\infty} = -\sigma\left(e^{-\infty} - e^{0}\right) = \sigma \qquad \left(\because e^{-\infty} = 0, e^{0} = 1\right)$$

$$\therefore \quad \mu_1^1 = \sigma \quad \cdot \quad \cdot \quad \cdot (1)$$

Similarly $\mu_2^1 = \int_0^{\infty} x^2 \cdot f(x)\, dx = \int_0^{\infty} x^2 \cdot \frac{1}{\sigma} \cdot e^{-x/\sigma}\, dx$

$$= \frac{1}{\sigma}\int_0^{\infty} x^2 \cdot e^{-x/\sigma}\, dx$$

$$= \frac{1}{\sigma}\left[x^2 \cdot \frac{e^{-x/\sigma}}{-1/\sigma}\right]_0^\infty - \frac{1}{\sigma}\int_0^\infty 2x \cdot \frac{e^{-x/\sigma}}{-1/\sigma} \int_0^\infty dx$$

$$= 0 + 2\int_0^\infty x \cdot e^{-x/\sigma}\, dx$$

$$= 2 \cdot x \cdot \frac{e^{-x/\sigma}}{-1/\sigma}\int_0^\infty - 2\cdot \int_0^\infty \cdot \frac{e^{-x/\sigma}}{-1/\sigma}\, dx$$

$$= 0 + 2\sigma \cdot \int_0^\infty e^{-x/\sigma}\, dx$$

$$\mu_2^1 = 2\,\sigma \cdot \frac{e^{-x/\sigma}}{-1/\sigma}\int_0^\infty = 2\sigma^2\left[-\left(e^{-\infty} - e^0\right)\right] = 2\sigma^2 \cdot \quad \cdot \quad \cdot (2)$$

$$\therefore\ V(x) = \mu_2 = \mu_2^1 - \mu_1^{1^2} = 2\sigma^2 - \sigma^2 \qquad (\because \text{ from } (1) \,\&\, (2))$$

Hence $S \cdot D \cdot = \sqrt{V(x)} = \sqrt{\sigma^2} = \sigma \cdot \quad \cdot \quad \cdot (3)$

Therefore the mean and S.D. are equal to σ.

Interquartile Range

If Q_1, Q_3 be the first and the third qurtiles then

$$\int_0^{Q_1} \frac{1}{\sigma} \cdot e^{-x/\sigma}\, dx = \frac{1}{4} \Rightarrow \frac{1}{\sigma} \cdot \frac{e^{-x/\sigma}}{-1/\sigma}\int_0^{Q_1} = \frac{1}{4}$$

$$\Rightarrow -\left(e^{-Q_1/\sigma} - e^0\right) = 1/4$$

$$\Rightarrow 1 - e^{-Q_1/\sigma} = 1/4 \Rightarrow e^{-Q_1/\sigma} = 1 - 1/4 = 3/4$$

$$\Rightarrow \frac{-Q_1}{\sigma} = \ell og\ 3/4 \Rightarrow \frac{Q_1}{\sigma} = \ell og\ 4/3 \quad \text{or} \quad Q_1 = \sigma\, \ell og\ 4/3$$

Also $\displaystyle\int_0^{Q_3} \frac{1}{\sigma} \cdot e^{-x/\sigma} \cdot dx = 3/4 \Rightarrow \frac{1}{\sigma} \cdot \frac{e^{-x/\sigma}}{-1/\sigma} \Big]_0^{Q_3} = \frac{3}{4}$

$$\Rightarrow -\left(e^{-Q_3/\sigma} - e^0\right) = 3/4 \Rightarrow 1 - e^{-Q_3/\sigma} = 3/4$$

$$\Rightarrow e^{-Q_3/\sigma} = 1/4 \quad \Rightarrow \frac{-Q_3}{\sigma} = \ell og\ 1/4 \quad \text{or} \quad Q_3 = \sigma\, \ell og\ 4$$

$\therefore$ Interquartile range is $Q_3 - Q_1 = \sigma \cdot \ell og\ 4 - \sigma \cdot \ell og\ 4/3$

$$= \sigma\left(\ell og\ 4 - \ell og\ 4 + \ell og\ 3\right) = \sigma \cdot \ell og_e^3$$

Example 2: The water consumption of a city, in excess of 20,000 gallons, is exponentially distributed with mean 20,000. The city's water works has a daily stock of 4,00,000 gallons. What is the probability that the stock is insufficient for atleast two of the three days selected at random ?

Solution: If Y is the total consumption in a day, then $X = Y - 20,000$ has an exponential distribution with mean $20,000$ with the probability density function

$$f(x) = \frac{1}{20,000}\, e^{-x/20,000}, \qquad \text{for} \quad 0 \le X \le \infty$$

$\because$ If the demand exceeds 40,000 gallons then the stock will be proved insufficient.

i.e., $X \ge 40,000 - 20,000$

$$\Rightarrow \quad X \ge 20,000$$

The probability that the stock remains insufficient on any particular day is given by

$$P(X \geq 20,000) = \int_{20,000}^{\infty} \frac{1}{20,000} \cdot e^{-x/20,000}\, dx = e^{-10}$$

The probability that the stock is insufficient for atleast two of three days selected at random is equal to sum of probability that it is insufficient for all the three days and p[robability that it is insufficient for two of the three days.

$$= \left(e^{-10}\right)^3 + 3C_2 \left(e^{-10}\right)^2 \left(1 - e^{-10}\right)$$

$$= e^{-30} + 3 \cdot e^{-20}\left(1 - e^{-10}\right)$$

$$= e^{-20} \left(e^{10} + 3\left(1 - e^{-10}\right)\right) = e^{-20} \left(e^{-10} + 3 - 3e^{-10}\right)$$

$$= e^{-20} \left(3 - 2e^{-10}\right)$$

Lesson – 17

GAMMA DISTRIBUTION

Objecti of The Lesson:

After studying this lesson the student expect to have a clear comprehension of the theory and the practical utility about the definition, properties and applications of Gamma Distribution.

Structure of The Lesson:

This lesson consists of the following sections as detailed below:

17.1 Introduction

17.2 Definition

17.3 Properties

17.4 Workedout Examples

17.5 Exercise

17.1 Introduction:

We studied exponential distribution lesson 16. One of the properties of exponential distribution is that sum of the i dependent and identically distribution exponential variates follows a distribution, whose probability density function is given by a mathematical expression involving a gamma function. Even directly also we can define the gamma distribution from the improper integrals of mathematics. In this lesson we consider such a distribution and discuss its properties and applications.

17.2 Definition:

The continuous variate X which is distributed according the probability law

$$f(x) = \frac{e^{-x} \cdot x^{n-1}}{\Gamma(n)} \quad ; \; n > 0, 0 < x < \infty \cdot \quad \cdot \quad \cdot (1)$$

$$= 0 , \quad \text{otherwise}$$

is known as a Gamma variate with parameter n, $f(x)$ is called the probability density function of Gamma Distribution. Any variate having the p.d.f. (1) is expressed as $X \sim \Gamma(n)$. The function $f(x)$ defined above represents a probability function since

$$\int_0^\infty f(x)\,dx = \frac{1}{\Gamma(n)} \int_0^\infty e^{-x} \cdot x^{n-1}\,dx = \frac{1}{\Gamma(n)} \cdot \Gamma(n) = 1$$

Also a continuous random variable X having the following p.d.f. is said to have a Gamma Distribution with two parameters α and n is

$$f(x) = \frac{\alpha^n}{\Gamma(n)} \cdot e^{-\alpha x} \cdot x^{n-1} \;\; ; \;\; \alpha > 0,\, n > 0 \;\; ; \;\; 0 < x < \infty \cdot \quad \cdot \quad \cdot (2)$$

$$= 0, \quad \text{otherwise}$$

If X has this density we express it as $X \sim \Gamma(\alpha, n)$. Taking $a = 1$ in (2) we get (1). Hence we may denote it as $X \sim \Gamma(n) = (1,\; n)$.

The cumulative distribution function of gamma variate x is

$$F(x) = \int_0^x (u)\,du = \frac{1}{\Gamma(n)} \cdot \int_0^x e^{-u} \cdot u^{n-1}\,du \;\; , \quad x > 0$$

$$= 0, \quad \text{otherwise}$$

17.3 Properties - Moments:

Moments about origin is given by $\mu_r^1 = E(X^r) = \int_{-\infty}^{\infty} X^2 \cdot f(x)\,dx \cdot \quad \cdot \quad \cdot (1)$

It $X \sim \Gamma(n)$ then from equation (1)

$$\mu_r^1 = E(X^r) = \int_0^\infty x^r \cdot \frac{1}{\Gamma(n)} \cdot e^{-x} \cdot x^{n-1}\,dx$$

$$= \int_0^\infty \frac{1}{\Gamma(n)}\, e^{-x} \cdot x^{n+r-1}\,dx$$

$$= \frac{1}{\Gamma(n)} \int_0^\infty e^{-x} \cdot x^{n+r-1}\,dx$$

$$\mu_r^1 = \frac{1}{\Gamma(n)}\Gamma(n+r) \qquad \left(\because \Gamma(n) = \int_0^\infty e^{-x}\cdot x^{n-1}\,dx\right) \cdot \quad \cdot \quad \cdot (2)$$

for $r = 1, 2, 3, \cdot \quad \cdot \quad \cdot$ th result (2) gives

$$r = 1,\ \mu_1^1 = \frac{\Gamma(n+1)}{\Gamma(n)} = \frac{(n+1-1)!}{(n-1)!} \qquad (\because \Gamma(n) = (n-1)!)$$

$$= \frac{n!}{(n-1)!} = n$$

$$r = 2,\ \mu_2^1 = \frac{\Gamma(n+2)}{\Gamma(n)} = \frac{(n+2-1)!}{(n-1)!} = \frac{(n+1)!}{(n-1)!} = n(n+1)$$

$$\therefore\ \mu_2 = \mu_2^1 - \mu_1^{1^2} = n(n+1) - n^2 = \cancel{n^2} + n - \cancel{n^2} = n$$

$$r = 3,\ \mu_3^1 = \frac{\Gamma(n+3)}{\Gamma(n)} = \frac{(n+3-1)!}{(n-1)!} = \frac{(n+2)!}{(n-1)!} = n(n+1)(n+2)$$

$$\therefore\ \mu_3 = \mu_3^1 - 3\mu_2^1\,\mu_1^1 + 2\mu_1^{1^3} = n(n+1)(n+2) - 3n(n+1)\,n + 2\cdot n^3$$

$$\mu_4^1 = \frac{\Gamma(n+4)}{\Gamma(n)} = \frac{(n+4-1)!}{(n-1)!} = \frac{(n+3)!}{(n-1)!} = n(n+1)(n+2)(n+3)$$

$$\mu_4 = \mu_4^1 - 4\,\mu_3^1\,\mu_1^1 + 6\,\mu_2^1\,\mu_1^{1^2} - 3\,\mu_1^{1^4}$$

$$= n(n+1)(n+2)(n+3) - 4n(n+1)(n+2)n + 6\cdot n(n+1)\cdot n^2 - 3n^4$$

$$= 3n^2 + 6n$$

$$\because\ \beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{(2n)^2}{n^3} = \frac{4}{n}, \qquad \beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{3n^2+6n}{n^2} = 3 + 6/n$$

Harnonic Mean:

For Harnonic Mean H we evaluate

$$E\left(\frac{1}{H}\right) = E\left(\frac{1}{X}\right) = \int_0^\infty \frac{1}{X} \cdot f(x)dx = \frac{1}{\Gamma(n)} \cdot \int_0^\infty \frac{1}{x} \cdot x^{n-1} \cdot e^{-x}\, dx$$

$$= \frac{1}{\Gamma(n)} \int_0^\infty x^{n-2} \cdot e^{-x}\, dx = \frac{1}{\Gamma(n)} \int_0^\infty x^{(n-1)-1} \cdot e^{-x} dx$$

$$E\left(\frac{1}{H}\right) = \frac{\Gamma(n-1)}{\Gamma(n)} = \frac{(n-2)!}{(n-1)!} = \frac{1}{n-1}$$

$$\therefore\ H = n-1$$

Mode:

If $X \sim \Gamma(n)$ then $f(x) = \frac{1}{\Gamma(n)} \cdot e^{-x} \cdot x^{n-1}$, $n > 0,\ 0 < x < \infty$

$\therefore$ Mode M_0 is obtained by taking $f^1(x) = 0$

$$\Rightarrow f^1(x) = \frac{d}{dx} f(x) = \frac{d}{dx}\left\{\frac{1}{\Gamma(n)} e^{-x} \cdot x^{n-1}\right\} = \frac{1}{\Gamma(n)}\left\{-e^{-x} x^{n-1} + e^{-x} \cdot x^{n-2}(n-1)\right\}$$

$$= \frac{e^{-x} \cdot x^{n-2}}{\Gamma(n)} \left\{-x + (n-1)\right\}$$

$$\therefore\ f^1(x) = 0 \Rightarrow \frac{e^{-x} \cdot x^{n-2}}{\Gamma(n)} \left\{-x + (n-1)\right\} = 0$$

$$\Rightarrow -x + (n-1) = 0 \quad \text{or} \quad x = n-1.$$

Giving the possible value of mode $= n-1$. Also for $x = n-1$, $f^1(x)$ is negative, which confirms that $x = n-1$ is the modal value of the distribution.

If $X_1, X_2, \cdots\cdots, X_n$ are i.i.d. Expo (λ) variates, then $S_n = X_1, X_2, \cdots\cdots, X_n$ is a gamma $(\lambda,\ n)$ variate.

Proof: If $X \sim \text{Expo}(\lambda)$, then its m.g.f. is given by

$$M_X(t) = \left[1 - t/\lambda\right]^{-1} \quad \cdot \quad \cdot \quad \cdot (1)$$

Now $M_{S_n}(t) = M_{X_1, X_2, \ldots\ldots, X_n}(t) = M_{X_1}(t) \cdot M_{X_2}(t) \cdot \quad \cdot \quad \cdot M_{X_n}(t)$

$$= \underbrace{\left(1 - t/\lambda\right)^{-1} \cdot \left(1 - t/\lambda\right)^{-1} \cdot \quad \cdot \quad \cdot \left(1 - t/\lambda\right)^{-1}}_{\text{n times}}$$

$$= \left(1 - t/\lambda\right)^{-n} \quad \left(\because X_i^{'}s \text{ are } i \cdot i \cdot d \cdot \text{Expo}(\lambda)\right)$$

The shows that $S_n \sim \text{gamma}(\lambda, n)$ variate.

Limiting form of Gamma Distribution as $n \to \infty$:

If $X \sim \Gamma(n)$, then $E(X) = n = \mu \text{ (say)}$ and $Var(X) = n = \sigma^2$ (say).

Then standard gamma variate is given by

$$Z = \frac{X - \mu}{\sigma} = \frac{X - n}{\sqrt{n}}$$

But $M_Z(t) = e^{-\mu t/\sigma} \cdot M_X\left(t/\sigma\right) = e^{-\mu t/\sigma}\left(1 - t/\sigma\right)^{-n}$

$$= e^{-nt/\sqrt{n}} \cdot \left(1 - t/\sqrt{n}\right)^{-n}$$

$$= e^{-\sqrt{n} \cdot t} \cdot \left(1 - t/\sqrt{n}\right)^{-n}$$

$$\Rightarrow K_z(t) = \log M_z(t) = \log\left(e^{-\sqrt{n} \cdot t} \cdot \left(1 - t/\sqrt{n}\right)^{-n}\right)$$

$$= -\sqrt{n}\, t - n \cdot \log\left(1 - t/\sqrt{n}\right)$$

$$= -\sqrt{n}\, t - n\left[-\frac{t}{\sqrt{n}} - \frac{t^2}{n}\cdot\frac{1}{2} - \frac{t^3}{n^{3/2}}\cdot\frac{1}{3} \cdot \quad \cdot \quad \cdot\right]$$

$$= -\sqrt{n}\, t + \sqrt{n}\, t + t^2/2 + 0\left(n^{-3/2}\right)$$

Where $0\left(n^{-1/2}\right)$ are terms containing $\frac{1}{2}$ and higher powers of n in the denaminator.

$$\therefore \lim_{n\to\infty} K_z(t) = \frac{t^2}{2} \Rightarrow \lim_{n\to 8} M_z(t) = e^{t^2/2}$$

Which is the m.g.f. of a standard normal variate therefore Gamma Distribution tends to Normal Distribution for large values of parameter n.

Moment Generating Function:

According to definition of Moment Generating Function we have

$$M_X(t) = E\left(e^{tX}\right) = \int e^{tX}\, f(x)\, dx$$

If $X \sim \Gamma(n)$ then we have

$$M_X(t) = \int_0^\infty e^{tX} \cdot \frac{e^{-x}\cdot x^{n-1}}{\Gamma(n)}\, dx$$

$$= \frac{1}{\Gamma(n)} \cdot \int_0^\infty e^{-(1-t)x} \cdot x^{n-1}\, dx$$

$$= \frac{1}{\Gamma(n)} \cdot \frac{\Gamma(n)}{(1-t)^n}\ ,\ |t| < 1 \qquad \left(\because \int_0^\infty e^{-ky}\cdot y^{n-1}\, dy = \frac{\Gamma(n)}{k^n}\right)$$

$$= \frac{1}{(1-t)^n} = (1-t)^{-n} = 1 + nt + \frac{n(n+1)}{2!}\cdot t^2 + \cdot \quad \cdot \quad \cdot + \frac{n(n+1)\cdot \quad \cdot \quad \cdot (n+r-1)t^2}{r!}$$

Which gives $\mu_r^1 =$ the coefficience of $\frac{t^2}{r!} = n(n+1) \cdot \quad \cdot \quad \cdot (n+r-1)$

$$\therefore \quad M_X(t) = (1-t)^{-n} \quad |t| < 1 \qquad\qquad (r = 1, 2, 3, 4, \cdot \quad \cdot \quad \cdot)$$

Mean $\mu_1^1 = n, \mu_2^1 = n(n+1), \mu_3^1 = n(n+1)(n+2), \mu_4^1 = n(n+1)(n+2)(n+3), \mu_2 = n$.

Cumulant Generating Function of Gamma Distribution:

By definition of cumulant generating function

We have

$$K_X(t) = \ell og\ M_X(t) \cdot \quad \cdot \quad \cdot (1)$$

If $X \sim \Gamma(n)$, then we have $M_X(t) = (1-t)^{-n}, \ |t| < 1$

$\therefore$ substituting the value of $M_X(t)$ in (1) we get

$$K_X(t) = \ell og\ (1-t)^{-n}$$

$$= -n\ \ell og\ (1-t)$$

$$= -n \cdot \left[-t - \frac{t^2}{2} - \frac{t^3}{3} - \frac{t^4}{4} \cdot \quad \cdot \quad \cdot \right]$$

$$\left(\because \ \ell og\ (1-x) = -x - \frac{x^2}{2} - \frac{x^3}{3} - \frac{x^4}{4} \cdot \quad \cdot \quad \cdot \right)$$

$$K_X(t) = n\left[t + \frac{t^2}{2} + \frac{t^3}{3} + \frac{t^4}{4} + \cdot \quad \cdot \quad \cdot \right] \cdot \quad \cdot \quad \cdot (2)$$

$$K_X(t) = n \cdot \sum_{r=1}^{\infty} t^r / r$$

$$\therefore \quad K_X(t) = n \cdot \sum_{r=1}^{\infty} \frac{t^r}{r!}(r-1)! \cdot \quad \cdot \quad \cdot (3)$$

$\therefore$ Mean = K_1 is obtained by taking coefficient of t in $K_X(t)$

i.e., $K_1 = n$

$\mu_2 = K_2$ is obtained by taking coefficient of $\frac{t^2}{2!}$ in $K_X(t)$

i.e., $K_2 = n$

$\mu_3 = K_3$ is obtained by taking coefficient of $\frac{t^3}{3!}$ in $K_X(t)$

i.e., $K_3 = 2n$

is obtained by tking coefficient of $\frac{t^4}{4!}$ in $K_X(t)$

i.e., $K_4 = 6n$

$\therefore \quad \mu_4 + K_4 + 3K_2^2 = 6n + 3n^2$

Like Poisson Distribution, the mean and variance of the Gamma Distribution are equal.

Pearson's Measure of Skewness:

Since $\mu_1 = K_1(t) = n, \ \ \mu_2 = K_2(t) = n, \ \ \mu_3 = K_3(t) = 2n, \ \ \mu_4 = K_4 + 3K_2^2 = 6n + 3n^2$

We have $\beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{(2n)^2}{n^3} = \frac{4n^2}{n^3} = 4/n$

or $\gamma_1 = \sqrt{\beta_1} = \sqrt{4/n} = \frac{2}{\sqrt{n}} > 0$

Pearson's Measure of Kurtosis:

We have $\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{6n + 3n^2}{n^2} = \frac{n(3n+6)}{n^2} = \left(3 + \frac{6}{n}\right)$

or $\gamma_2 = \beta_2 - 3 = 3 + \frac{6}{n} - 3 = \frac{6}{n} > 0$

Since $\gamma_1 > 0$, the Gamma Distribution is positively skewed and $\gamma_2 > 0$ implies that it is leptokurtic. Obviously, the point (β_1, β_2) lies on the straight line $2\beta_2 - 3\beta_1 - 6 = 0$.

Additive Property of Gamma Distribution:

The sum of independent gamma variate is also a gamma variate. More specifically, if $X_1, X_2, \cdots\cdots, X_K$ are independent gamma variates with parameters $n_1, n_2, \cdots\cdots\cdots, n_k$ respectively then $X_1 + X_2 + \cdots\cdots + X_k$ is also a gamma variate with parameter $n_1 + n_2 + \cdots\cdots\cdots + n_k$.

Proof: Since X_i is gamma variate with parameters n_i.

i.e., $X_i \sim \Gamma(n_i)$

Then according to m.g.f. of Gamma Distribution we have

$$M_{X_i}(t) = (1-t)^{-n_i}$$

The m.g.f. of the sum $X_1 + X_2 + \cdots\cdots + X_k$ is given by

$$M_{X_1+X_2+\cdots\cdots+X_k}(t) = M_{X_1}(t) \cdot M_{X_2}(t) \cdots\cdots\cdots\cdots M_{X_k}(t)$$

$$(\because \; X_1, X_2, \cdots\cdots, X_k \text{ are independent})$$

$$= (1-t)^{-n_1} \cdot (1-t)^{-n_2} \cdots\cdots\cdots (1-t)^{-n_k}$$

$$= (1-t)^{-(n_1+n_2+\cdots\cdots+n_k)}$$

Which is again the m.g.f. of a gamma variate with parameter $(n_1 + n_2 + \cdots\cdots\cdots + n_k)$. Hence by the uniqueness theorem of m.g.f. the sum of independent gamma variates is also a gamma variate.

17.4 Workedout Examples:

Example 1: Show that the Harnonic mean of the variate which ranges from 0 to ∞ with p.d.f. $f(x) = \frac{x^n \cdot e^{-x}}{n!}$ is n. Where n is positive.

Solution: Given $f(x) = \frac{x^n \cdot e^{-x}}{n!}$, $0 < x < \infty \quad \cdot \quad \cdot \quad \cdot (1)$

Then we have Harmonic mean of the variate as

$$\frac{1}{H} = \int_0^\infty \frac{1}{x} f(x)\, dx$$

$$\frac{1}{H} = \int_0^\infty \frac{1}{x} \cdot \frac{x^n \cdot e^{-x}}{n!}\, dx \qquad (\because \text{ from } (1))$$

$$= \frac{1}{n!} \int_0^\infty x^{n-1} e^{-x}\, dx$$

$$= \frac{1}{n!} \Gamma(n) \qquad (\because \text{ By definition of p.d.f. of gamma function})$$

$$= \frac{(n-1)!}{n!} \qquad (\because \Gamma(n) = (n-1)!)$$

$$\frac{1}{\mu} = \frac{1}{n} \Rightarrow H = n$$

$\therefore$ The Harnonic mean of the variate which ranges from 0 to ∞ with p.d.f. $\frac{x^n \cdot e^{-x}}{n!}$ is n.

Example 2: Prove that the sum of two independent gamma variates with parameters (n_1, σ_1) and (n_2, σ_2) is a gamma variate with parameter $n_1 + n_2$ provided $\sigma_1 = \sigma_2$.

Solution: Let X_1 and X_2 betwo independent gamma variates with parameters (n_1, σ_1) and (n_2, σ_2) respectively. The moment generating function of their sum $X_1 + X_2$ is given by

$$M_{X_1+X_2}(t) = E\left[e^{t(X_1+X_2)}\right] = E\left[e^{tX_1} \cdot e^{tX_2}\right]$$

$$= E\left[e^{tX_1}\right] \cdot E\left[e^{tX_2}\right]$$

$$= M_{X_1}(t) \cdot M_{X_2}(t)$$

Since $M_X(t)$ for gamma variate $= (1 - \sigma t)^{-n}$

Hence $M_{X_1+X_2}(t) = (1 - \sigma_1 t)^{-n_1} \cdot (1 - \sigma_2 t)^{-n_2}$

If $\sigma_1 = \sigma_2 = \sigma$, then

$$M_{X_1+X_2}(t) = (1-\sigma t)^{-n_1} \cdot (1-\sigma t)^{-n_2}$$

$$= (1-\sigma t)^{-(n_1+n_2)}$$

Which is the moment generating function of gamma variate with parameter (n_1+n_2). Provided $\sigma_1 = \sigma_2 = \sigma$.

Example 3: Consumer demand for milk in a certain locality, per month is known to be a gamma variate. If the average demand is 'a' litres and the most likely demand is 'b' litres $(b < a)$ what isthe variand of the demand?

Solution: If $X \sim \Gamma(n)$ then we have

$$E(X) = n, \ \text{Mode} = n-1$$

Since the average demand is 'a' litres i.e., $a = n \quad \cdot \quad \cdot \quad \cdot (1)$

The most likely demand is 'b' litres i.e., $b = n-1 \quad \cdot \quad \cdot \quad \cdot (2)$

Thus a - b gives, $a - b = \not{n} - \not{n} + 1 = 1 \quad \cdot \quad \cdot \quad \cdot (3)$

But the variance of X is given by $V(X) = n$

$V(X)$ can be written as $V(X) = n \cdot 1 = a(a-b)$

$(\because \text{ from } (1) \;\&\; (3) \text{ i.e., } a = n,\ (a-b) = 1)$

Example 4: A random sample of size n is taken from population. If $\overline{X}$ is the sample, show that $n\lambda\overline{X}$ is $\Gamma(1, n)$ and that $\leq E$ of $\overline{X}$ is $1/\lambda\sqrt{n}$.

Solution: If $X \sim \text{Expo}(\lambda)$ then m.g.f. is given by

$$M_X(t) = \left(1 - t/\lambda\right)^{-1}$$

$$\therefore M_{n\lambda\overline{X}}(t) = M_{X_1+X_2+\cdots\cdots+X_n}(\lambda t) = \left[M_{X_i}(\lambda t)\right]^n = (1-t)^{-n} \cdot \quad \cdot \quad \cdot (1)$$

$(\because X_i \text{ are i.i.d.})$

This shows that $n\lambda\overline{X} \sim \Gamma(1, n)$

Let $Y = n\lambda\overline{X}$ and expanding equation (1) we get

$$M_Y(t) = 1 + nt + n(n+1)\frac{t^2}{2} + \cdots\cdots$$

$$E(Y) = n \Rightarrow E(n\lambda\overline{X}) = n \quad \text{or} \quad E(\overline{X}) = \frac{n}{n\lambda} = 1/\lambda$$

$$E(Y^2) = n(n+1) \quad \text{i.e., coefficient of } t^2/2$$

$$\Rightarrow E\left(n^2\lambda^2\overline{X^2}\right) = n(n+1) \quad \text{or} \quad E\left(\overline{X^2}\right) = \frac{n(n+1)}{n^2\lambda^2} = \frac{(n+1)}{n\lambda^2}$$

$$\therefore \text{Var}(\overline{X}) = E\left(\overline{X^2}\right) - \left[E(\overline{X})\right]^2$$

$$\text{Var}(\overline{X}) = \frac{(n+1)}{n \cdot \lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}\left[\frac{n+1-n}{n}\right] = \frac{1}{n\lambda^2}$$

$$\therefore S \cdot E \cdot (\overline{X}) = \sqrt{\text{Var}(\overline{X})} = \sqrt{\frac{1}{n\lambda^2}} = \frac{1}{\lambda\sqrt{n}}$$

17.5 Exercise:

1. Prove that $\int_0^x x \cdot e^{-x^2}\, dx = x e^{-x^2}\left[1+\frac{1}{3}\cdot\left(2x^2\right)+\frac{1}{3\cdot 5}\left(2x^2\right)^2+\cdot\quad\cdot\quad\cdot\right]$

2. Show that the function $f(x)=\frac{e^{-x}\cdot x^{n-1}}{\Gamma(n)},\ n>0,\ 0<x<\infty$ and '0' otherwise is a p.d.f.

3. Find the M.G.F. of a $\Gamma(n)$ variate, i.e., a gamma variate with parameter n, w.r.t. oxigin.

4. Show that the sum of two gamma variates, with parameters ℓ and m is a gamma variate with parameter $\ell + m$.

5. Show that for the distribution $f(x)=\frac{1}{\Gamma(n)}e^{-x}\cdot x^{n-1},\quad 0\le x\le\infty,\quad n>0$ mean and variance are equal.

6. If X be a normal variate with mean μ and S.D. σ, then $\frac{(x-\mu)^2}{2\sigma^2}$ is a gamma variate with parameter $1/2$.

7. Show that the mean value of positive square root of a $\Gamma(n)$ variate is $\Gamma\left(n+1/2\right)/\Gamma(n)$.

   Hence prove that the mean deviation of a normal variate from its mean is $\sigma\sqrt{2/\pi}$, σ being the S.D. of the distribution.

8. If a random variable X has a gamma distribution with $n=2$ and $\sigma=1$, find $P\left[1\cdot 8<X<2\cdot 4\right]$. $\left[\text{Ans}: 2\cdot 8e^{-1\cdot 8}-3\cdot 4e^{-2\cdot 4}, 1545\right]$

9. In a certain city the daily consumption fo water (in millions of gallons) follows approximately a Gamma Distribution with $n=2$ & $\sigma=3$. If the daily capacity of this city is 9 million gallons of water what is the probability that on any given day the water supply is inadequate? $\left[\text{Ans}:\ 4e^{-3}=0\cdot 1992\right]$

Lesson - 18

CAUCHY DISTRIBUTION

Object of The Lesson:

After studying this lesson the student is expected to have a clear comprehension of the theory and the practical utility about the concepts of definition, properties and applications of Cauchy Distribution.

Structure of The Lesson:

This lesson consists of the following sections as detailed below:

18.1 Introduction:

In mathematics it is well known that, to prove any result a logical construction is essential. On the other hand to disprove a statement a counter example is enough. The same is true in statistics also. One such distribution to disprove a number of theoretical propositions is the Cauchy Distribution. It is infact the probability of a random angle with uniform probability in $\left(-\pi/2, \pi/2\right)$ described by a distribution. After being converted into Cartesian coordinates it becomes the Cauchy Distribution. In this lesson we explain the development and properties of a Cauchy Distribution.

18.2 Definition:

Probbability Denisity Function:

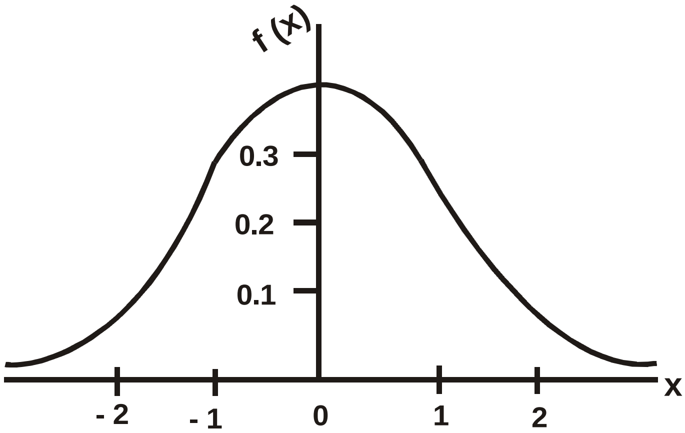
A random variable X having the p.d.f.

$$f(x) = \frac{\lambda}{\pi}\left[\lambda^2 + (x-\mu)^2\right], \quad -\infty < x < \infty, \quad \lambda > 0 \quad \cdot \quad \cdot \quad \cdot (1)$$

is said to be Cauchy Distribution with location parameter μ and scale parameter $\lambda > 0$. A variate with density (1) is expressed as $X \sim C(\mu, \lambda)$. The case $\mu = 0, \ \lambda = 1$ equation (1) gives

$$f(x) = \frac{1}{\pi} \cdot \frac{1}{(1+x^2)}, \quad -\infty < x < \infty \cdot \quad \cdot \quad \cdot (2)$$

and if a variate with density (2), is expressed as $X \sim C(0, 1)$



Cauchy Distribution

Cumulative Distribution Function:

$$F(x) = \int_{-\infty}^{x} \frac{\lambda}{\pi} \frac{dt}{\lambda^2 + (t-\mu)^2}$$

$$\therefore P(X \le x) = \frac{1}{2} + \pi^{-1} \tan^{-1}\left[\frac{(x-\mu)}{\lambda}\right]$$

The shape of Cauchy Distribution generally is similar to normal curve, but it decreases more slowly for large values of x, more specifically using c.d.f. we observe that

$$P\{|X-\mu| < \lambda\} = F(\mu+\lambda) - F(\mu-\lambda) = \left(\frac{1}{2}+\pi^{-1}\tan^{-1}(1)\right) - \left(\frac{1}{2}+\pi^{-1}\tan^{-1}(-1)\right) = \frac{1}{2} = 0\cdot 5$$

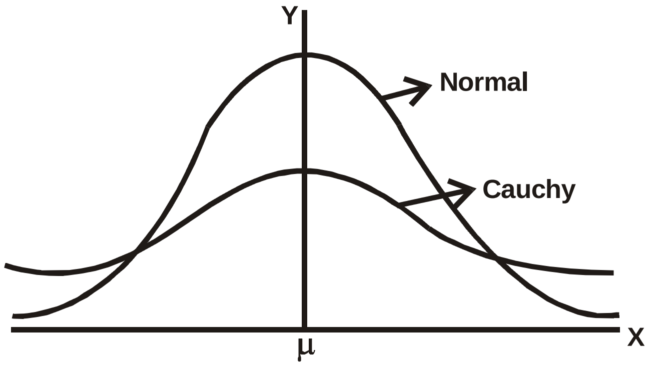
$$P\{|X-\mu| < 2\lambda\} = F(\mu+2\lambda) - F(\mu-2\lambda) = 2\pi^{-1}\tan^{-1}(2) = 0\cdot 706$$

$$P\{|X-\mu| < 3\lambda\} = F(\mu+3\lambda) - F(\mu-3\lambda) = 2\pi^{-1}\tan^{-1}(3) = 0\cdot 795$$

For $N(\mu, \sigma^2)$ the corresponding values are

$$P\{|X-\mu| < \sigma\} = 0\cdot 6826,\ P\{|X-\mu| < 2\sigma\} = 0\cdot 9544,\ P\{|X-\mu| < 3\sigma\} = 0\cdot 9973$$

The higher values for $N(\mu, \sigma^2)$ indicate a larger concentration of area close to μ than for Cauchy μ with implications their rate of fall is slow - the following figure indicates this situation.



18.3 Properties:

Moments:

$$E(y) = \int_{-\infty}^{\infty} y \cdot f(y)\, dy = \frac{\lambda}{\pi} \int_{-\infty}^{\infty} \frac{y}{\lambda^2 + (y-\mu)^2}\, dy$$

$$= \frac{\lambda}{\pi} \int_{-\infty}^{\infty} \frac{(y-\mu) + \mu}{\lambda^2 + (y-\mu)^2}\, dy$$

$$= \mu \cdot \frac{\lambda}{\pi} \int_{-\infty}^{\infty} \frac{dy}{\lambda^2 + (y-\mu)^2} + \frac{\lambda}{\pi} \int_{-\infty}^{\infty} \frac{(y-\mu)}{\lambda^2 + (y-\mu)^2}\, dy$$

$$= \mu \cdot 1 + \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{z}{\lambda^2 + z^2} dz$$

Although the integral $\int_{-\infty}^{\infty} \frac{z}{\lambda^2 + z^2} dz$ is not completely convergent.

i.e., $\lim_{\substack{n \to \infty \\ n^1 \to \infty}} \int_{-n}^{n^1} \frac{z}{\lambda^2 + z^2} dz$ does not exist its principal value,

$\lim_{n \to \infty} \int_{-n}^{n} \frac{z}{\lambda^2 + z^2} dz$ exists and is equal to zero. Thus, in general mean of Cauchy Distribution does not exist. But, if we assume that the mean of Cauchy Distribution exists (by taking the principal value), then it is located at $x = \mu$. Also the probability curve is symmetrical about the point $x = \mu$. Hence for this distribution, the mean, median and mode coincide at the point $x = \mu$.

$$\mu_2 = E(y-\mu)^2 = \int_{-\infty}^{\infty} (y-\mu)^2 \, f(y) dy = \frac{\lambda}{\pi} \int_{-\infty}^{\infty} \frac{(y-\mu)^2}{\lambda^2 + (y-\mu)^2} dy$$

Which does not exist since the integral is not convergent thus for the Cauchy's Distribution $\mu_r \; (r \geq 2)$ do not exist.

Moment Generating Function:

The m.g.f. of Cauchy Distribution does not exist for $t \neq 0$. According to definition of m.g.f.

$$M_X(t) = \frac{\lambda}{\pi}\int_{-\infty}^{\infty}\frac{e^{tx}}{\lambda^2+x^2}\,dx \;\geq\; \frac{\lambda}{\pi}\int_{0}^{\infty}\frac{e^{tx}}{\lambda^2+x^2}\,dx \quad \cdot \quad \cdot \quad \cdot (1)$$

(By taking $\mu = 0$ in $f(x)$)

Since $e^{tx} = 1 + tx + 0(t)$, where $0(t) > 0$

for $b > 0$, the above equation (1) gives $M_X(t) \geq \int_{0}^{\infty}\frac{\lambda(1+tx)}{\pi(\lambda^2+x^2)}\,dx \to \infty$

From this it follows that no moments of Cauchy variate exist.

Characteristic Function:

If X is Laplace $(\mu,\ \lambda)$, its p.d.f. and characteristic function are

$$f(x) = \frac{1}{2}\lambda\, e^{-\lambda|x-\mu|},\ -\infty < x < \infty \quad \cdot \quad \cdot \quad \cdot (1)$$

$$\phi(t) = e^{\mu t i}\cdot\frac{\lambda^2}{(\lambda^2+t^2)} \quad \cdot \quad \cdot \quad \cdot (2)$$

By using the Inversion formula

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\phi(t)\,e^{itx}\cdot dt$$

$$\therefore \quad \frac{1}{2}\lambda\cdot e^{-\lambda|x-\mu|} = \frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{e^{\mu t i}\cdot\lambda^2\cdot e^{-itx}}{\lambda^2+t^2}\,dt \quad (\because \text{ from equations (1) \& (2)})$$

or $$e^{-\lambda|z|} = \frac{\lambda}{\pi}\int_{-\infty}^{\infty} \frac{e^{itz}}{\lambda^2+t^2}\,dt \qquad (\because\ z=\mu-x)$$

Now transforming variable t to $y-\mu$ & taking $z=\theta$ we get

$$e^{-\lambda|\theta|} = \frac{\lambda}{\pi}\int_{-\infty}^{\infty} \frac{e^{i\theta(y-\mu)}}{\lambda^2+(y-\mu)^2}\,dy$$

$$= e^{-\mu i\theta}\int_{-\infty}^{\infty} \frac{\lambda e^{i\theta y}\,dy}{\pi\left[\lambda^2+(y-\mu)^2\right]}$$

Againing rewritting θ to t we have

$$e^{\mu ti-\lambda|t|} = \int_{-\infty}^{\infty} \frac{\lambda e^{ity}\,dy}{\pi\left[\lambda^2+(y-\mu)^2\right]} = \int_{-\infty}^{\infty} e^{ity}\,g(y)\,dy$$

where $g(y) = \frac{\lambda}{\pi}\left[\lambda^2+(y-\mu)^2\right], \quad -\infty < y < \infty$

Which follows that $g(y)$ is $C(\mu,\lambda)$ density and hence $\phi(\,t:y) = e^{\mu it-\lambda|t|}$

Mode and Points of Inflexion:

Let $y = \frac{\lambda}{\pi}\left(\frac{1}{\lambda^2+z^2}\right),\ z = x-\mu$

Taking logarithm on bothe sides and differentiating twice we get w.r.t. 'z'

$$\frac{1}{y} y^1 = \frac{-2z}{\left(\lambda^2+z^2\right)} \quad \cdot \quad \cdot \quad \cdot (1) \qquad \left(\because y^1 = \frac{dy}{dz}\right)$$

$$\frac{1}{y} y^{11} - \left(\frac{y^1}{y}\right)^2 = \frac{-2\left(\lambda^2-z^2\right)}{\left(\lambda^2+z^2\right)^2} \quad \cdot \quad \cdot \quad \cdot (2) \qquad \left(\because y^{11} = \frac{d^2y}{dz^2}\right)$$

$y^1 = 0 \Rightarrow z = 0$ and $y^{11} < 0$ at $z = 0$

Thus $z = 0$, i.e., $x = \mu$ gives the modal value, which means $C(\mu, \lambda)$ has unique maximum at $x = \mu$.

Now taking $y^{11} = 0$ and eliminating y^1 betwen (1) and (2) which gives $3z^2 = \lambda^2 \Rightarrow \; = \mu \pm \left(\lambda/\sqrt{3}\right) = \mu \pm 0{\cdot}577\lambda$. These are the two points of inflexion of $C(\mu, \lambda)$ and $\mu \pm \sigma$ is the two points of inflexion of $N\left(\mu, \sigma^2\right)$ are compared.

Median:

The c.d.f. of $C(\mu, \lambda)$ is given by

$$F(x) = \frac{1}{2} + \pi^{-1} \tan^{-1}\left(\frac{(x-\mu)}{\lambda}\right)$$

If m is the median of $C(\mu, \lambda)$ then

$$F(m) = \frac{1}{2} \Rightarrow \frac{1}{2} + \pi^{-1} \tan^{-1}\left[\frac{m-\mu}{\lambda}\right] = \frac{1}{2}$$

$$\Rightarrow \pi^{-1} \tan^{-1}\left[\frac{m-\mu}{\lambda}\right] = 0$$

$$\Rightarrow \tan^{-1}\left[\frac{m-\mu}{\lambda}\right] = 0$$

$$\Rightarrow \frac{m-\mu}{\lambda} = 0 \qquad \left(\because \tan 0^0 = 0\right)$$

$$\Rightarrow \; m - \mu = 0$$

$$\Rightarrow \; m = \mu$$

Reproductive Property:

Statement: If $X_1 \sim C(\mu_1, \lambda_1)$ and $X_2 \sim (\mu_2, \lambda_2)$ are independent, then

$$(X_1 + X_2) \sim C(\mu_1 + \mu_2, \; \lambda_1 + \lambda_2).$$

Proof: From the definition of characteristic function $\phi_X(t) = E(e^{itX})$

If $\quad X_1 \sim (\mu_1, \lambda_1) \Rightarrow \phi_{X_1}(t) = e^{it\mu_1 - \lambda_1|t|}$

$\quad X_2 \sim C(\mu_2, \lambda_2) \Rightarrow \phi_{X_2}(t) = e^{it\mu_2 - \lambda_2|t|}$

Since X_1, X_2 are independent as stated we have

$$\phi_{X_1+X_2}(t) = \phi_{X_1}(t) \cdot \phi_{X_2}(t)$$

$$= e^{it\mu_1 - \lambda_1|t|} \cdot e^{it\mu_2 - \lambda_2|t|}$$

$$= e^{it\mu - \lambda|t|} \qquad (\because \mu = \mu_1 + \mu_2, \lambda = \lambda_1 + \lambda 2)$$

By uniqueness theorem, if follows that $X_1 + X_2 \sim C(\mu_1 + \mu_2, \lambda_1 + \lambda_2)$

Distribution of The Mean of Cauchy Variates:

Statement: If $X_1, X_2, \cdots\cdots, X_n$ are i.i.d. $C(\mu, \lambda)$ variates then

$$\overline{X} = \frac{(X_1 + X_2 + \cdots\cdots\cdots\cdots + X_n)}{n} \sim C(\mu, \lambda)$$

Proof: If $X \sim C(\mu, \lambda)$ then we have

$$\phi_X(t) = e^{i\mu t - \lambda(t)} \quad . \quad . \quad . (1)$$

$$\phi_{\frac{(X_1+X_2+\cdots+X_n)}{n}}(t) = \phi_{X_1+X_2+\cdots+X_n}\left(t/n\right)$$

$$= \phi_{X_1}\left(t/n\right) \cdot \phi_{X_2}\left(t/n\right) \cdots\cdots \phi_{X_n}\left(t/n\right)$$

$$= \left[\phi_{X_i}\left(t/n\right)\right]^n \qquad \left(\because X_i^{'}s \text{ are } i \cdot i \cdot d \cdot\right)$$

$$= \left[e^{\frac{i\mu t}{n} - \frac{\lambda|t|}{n}}\right]^n \qquad \text{[by equation (1)]}$$

$$= e^{i\mu t - \lambda|t|}$$

This shows that $\overline{X} \sim C(\mu, \lambda)$.