Practical Manual

Elementary Statistics & Computer Applications

ABB 158 3(2+1)
For B.Sc. Horticulture I Semester

Dr. Shailendra Kumar
Dr. Amit Kumar Jain

2020

College of Agriculture
Rani Lakshmi Bai Central Agricultural University, Jhansi-284003

Syllabus:

Construction of frequency distribution table and its graphical representation, histogram, frequency polygon, frequency curve, bar chart, simple, multiple, component and percentage bar charts, pie chart, mean, mode for row and grouped data, percentiles, quartiles, and median for row data grouped data, coefficient of variation, 't' test for independent will equal and unequal variances, paired 't' test, chi-square test for contingency tables and theoretical ratios, correlation and linear regression. Studies on computer components- Basic language, visual basic, programming techniques, MS Office, Excel, power point

Name of Students…………………………………………………………………………………

Roll No………………………………………………….Batch……………………………….

Session………………………………………………………………Semester…………..

Course Name: …………………………………………………………………………………

Course No:……………………………………………………………………………….

Credit……………………………………………………………………………………….

No. of Copies:

Price: Rs.

Year: 2020

CERTIFICATE

This is to certified that Shri./Km. …………………………………………………… ID No ………………..

has completed the practical of course …………………………..……………………………………..Course No…………………….. as per the syllabus of B.Sc. (Hons.) Horticulture……………………………………… ....Semester in the year……………………………..in the respective lab of College.

Date: Course Teacher

Contents

Practical No. 1

Objective: To construct frequency distribution table

Problem: The weight of 40 apples are following 85, 86, 89, 76, 76, 75, 111, 123, 56, 68, 67, 69, 91, 93, 99, 122, 112, 144, 123, 142, 151, 165, 88, 86, 92, 89, 90, 190, 145, 167, 173, 113, 118, 115, 119, 123, 126, 127, 88, 78. Construct a frequency table.

Solution:

Frequency distribution: Itis used to condense the large amount of data in tabulation form and to provide information of our interest in systematic way.

For construction of the frequency distribution table, we follow following steps`

Step-I: calculate the range= maximum value –minimum value. = ...

Step-II: determine an approximate value of

k = 1+3.322 $\log_{10}N$, where N is the total frequency

or k = $\frac{\log(\text{no.of observations})}{\log(2)}$ =...

Round up the value of k to the next integer which integer is take as a number of classes

Step-III: Width of class interval = $\frac{\text{range}}{\text{number of classes}}$ =..

Step-IV: Construct a table

Class Interval	Tally Marks	Frequency
………………..…..	……………………………..…………………	……….…..……
………………..…..	……………………………..…………………	……….…..……
………………..…..	……………………………..…………………	……….…..……
………………..…..	……………………………..…………………	……….…..……
………………..…..	……………………………..…………………	……….…..……
………………..…..	……………………………..…………………	……….…..……
………………..…..	……………………………..…………………	……….…..……
………………..…..	……………………………..…………………	……….…..……
………………..…..	……………………………..…………………	……….…..……
………………..…..	……………………………..…………………	……….…..……

.

Practical No. 2

Objective: To draw Histogram, Polygon, Frequency Curve and Ogive.

Problem: The weight of 40 apples are following 85, 86, 89, 76, 76, 75, 111, 123, 56, 68, 67, 69, 91, 93, 99, 122, 112, 144, 123, 142, 151, 165, 88, 86, 92, 89, 90, 190, 145, 167, 173, 113, 118, 115, 119, 123, 126, 127, 88, 78. Construct a frequency table and draw histogram, polygon, frequency curve, and ogive of the given data.

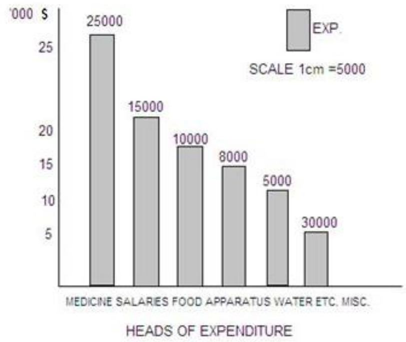
....
....
....
....
....
....
....
....
....
....
....
....
....
....
....
....
....
....
....
....
....
....
....
....
....
....
....
....
....
....
....
....
....
....

Practical No. 3

Objective: To draw Simple, Component, Multiple and Percentage Bar Diagram and Pie Chart.

Simple Bar diagram

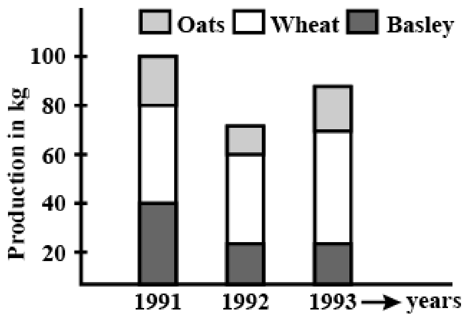
Problem: The cropping pattern in U.P. in the year 2012-13 was as follows. Draw simple bar diagram for this data is given below.



Crops	Area In 1,000 hectares
Cereals	6950
Oilseeds	3165
Pulses	1660
Others	2250

Component Bar diagram

Problem: Draw a Component bar diagram for the following data:

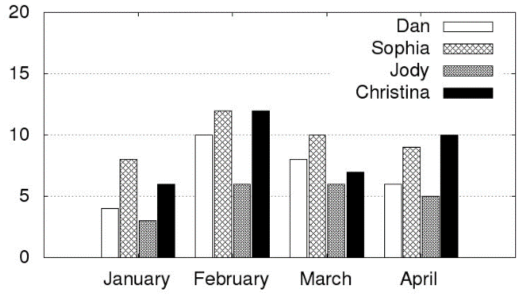


Year	Sales (Rs.)	Gross Profit (Rs.)	Net Profit (Rs.)
2015	110	40	20
2016	125	55	25
2017	130	60	30
2018	145	65	35
2019	160	70	40

Multiple Bar diagram

Problem: Draw a multiple bar diagram for the following data which represented agricultural production for the period from 2015-2019

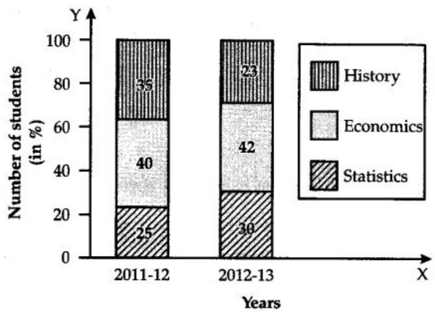
Year	Food grains (tones)	Vegetable(tones)	others (tones)
2015	120	45	15
2016	125	50	25
2017	135	60	30
2018	140	65	35
2019	160	70	40



Percentage Bar diagram

Problem: Draw a Percentage bar diagram for the following data:

Year	Sales (Rs.)	Gross Profit (Rs.)	Net Profit (Rs.)
2015	110	40	20
2016	125	55	25
2017	130	60	30
2018	145	65	35
2019	160	70	40



Percentage = $\frac{Actual\ value}{Total\ of\ the\ actual\ value} \times 100$

……

……

……

……

……

……

……

……

……

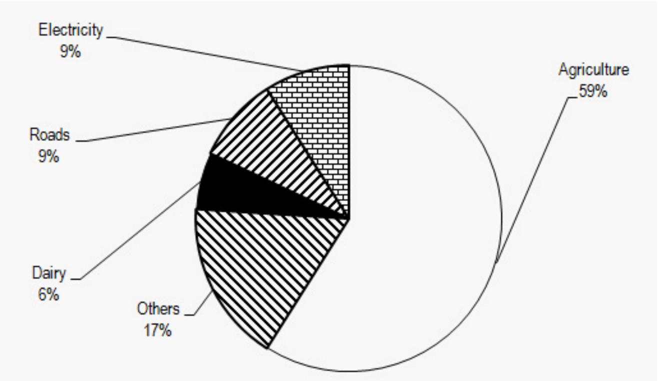
……

……

Pie chart

Problem: Given the cultivable land area in five Bundelkhand region districts of UP. Construct a pie diagram for the following data.

Districts	Cultivable area(in hectares)
Jhansi	425
Lalitpur	300
Banda	350
Hamirpur	250
Jalaun	500



Angle = $\frac{Actual\ value}{Total\ of\ the\ actual\ value} \times 360^0$

……

……

……

……

……

……

……

……

……

……

……

Practical No. 4

Objective: To calculate Average (Mean, Median and Mode) of ungrouped and grouped data.

Measure of Central Tendency: The purpose of measure central tendency is to find central value of the gathering data or frequency distribution. There are five method of Measure of central tendency.

Arithmetic Mean: It is useful when data is quantitative (Interval scale or Ratio scale). Arithmetic mean is more fluctuate by extreme values so if data set have very larger or very smaller value then we avoid to calculate arithmetic mean. In this case, we prefer to calculate median. If the distribution is in arithmetic progression then arithmetic mean gives best measure of central tendency. There are following method to calculation of arithmetic mean in different format of data

(a) For Individual Series (or Ungrouped data):

Problem: The following data is related to crop yield in quintals of last 7 years of a farmer. To find the average crop yield?

112, 124, 104, 140, 136, 132, 144.

Solution: Step –I: count number of observation (or values) = n =

Step-II: to calculate the total of all of observation (or values) = $\sum_{i=1}^{n} x_i$

= ....

Step-III: Finally we calculate the arithmetic mean of the following formula,

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

(b) For Discrete Series (Grouped data or discrete frequency distribution):

Problem: The following data is related to yearly income of 40 farmers of a village which are selected randomly. To find the average (Arithmetic Mean) income of farmers of the given village?

Income in Lakh (x_i):	2	3	4	5	6	7	8
No. of farmer (f_i):	1	3	6	12	8	7	3

Solution:

Step-I to construct the following Table:

Variable (x_i)	frequency(f_i)	$f_i \times x_i$
Total	$\sum_{i=1}^{n} f_i$=	$\sum_{i=1}^{n} f_i x_i$=

Step-II: To calculate the arithmetic mean, $\overline{x} = \frac{\sum_{i=1}^{n} f_i x_i}{N}$ =

(c) For Continuous Series (Grouped data in class interval):

Problem: The following data is related to milk production in liter of 240 farmers of a village which are selected randomly. Calculate average milk production of the given village?

Milk in liter(x_i)	0 - 4	4 - 8	8 - 12	12-16	16 - 20

No. of farmer (f_i):	30	90	60	40	20

Solution:

Step-I: to construct the following table:

Class Interval ($x_i - x_{i+1}$)	**frequency** **(f_i)**	**Mid Value of Class Interval** $m_i = \frac{x_i + x_{i+1}}{2}$	$f_i \times \boldsymbol{m_i}$
Total	$N = \sum_{i=1}^{n} f_i =$		$\sum_{i=1}^{n} f_i m_i =$

Step-II: To calculate the arithmetic of the following formula, $\overline{x} = \frac{\sum_{i=1}^{n} f_i x_i}{N}$=..

………..…………

……...

Median: It is most useful when data is qualitative (Ordinal). Median is also preferable in quantitative data if the data set have very larger or very smaller value and some values are missing but their rank is known. Median gives the middle value of the arranged data. It is also known as positional average. There are following method to calculation of Median in different format of data.

(a) For Individual Series (or ungrouped data):

Problem: The following data is related to plant height in cm. To find the median height of the plants?

12, 14, 18, 15, 21, 22, 24, 25, 3

Solution:

Step-I: arrange the data in ascending or descending order

....

Step –II: count number of observation (or values) = n =................

Step –III: Since n is odd number so we find median by the following formula:

Median = $(\frac{n+1}{2})^{th}\ item$ =.....................

....

Problem: The following data is related to marks of 12 students getting from 30 marks Mid-term examination. To find the median marks of the students?

19, 20, 22, 12, 14, 18, 15, 21, 22, 24, 25, 2

Solution:

Step-I: arrange the data in ascending or descending order

....

Step –II: count number of observation (or values) = n =................

Step –III: Since n is even number so we find median by the following formula:

Median = $\frac{(\frac{n}{2})^{th}\ item + (\frac{n}{2}+1)^{th}\ item}{2}$ =....................................

....

(b) For Discrete Series (Grouped data or discrete frequency distribution):

Problem: The following data is related to yearly income of 50 farmers of a village which are selected randomly. To find the median income of farmers of the given village?

Income in Lakh(x_i):	2	3	4	5	6	7	8
No. of farmer (f_i):	1	3	6	12	18	7	3

Solution:

Step-I: To construct the following table

Variable (x_i)	**Frequency (f_i)**	**cumulative frequency**
2	1	1
3	3	1+3=4
4	6	4+6=10
5	12	
6	18	
7	7	
8	3	
Total	N=$\sum_{i=1}^{n} f_i$=	

Step-II: To calculate $\frac{N}{2}$=...............

Step-III: To see the value just greater than $\frac{N}{2}$ in cumulative frequency column =..........

Step-IV: In first column, see the value of x correspond to the value getting in step-III,
Median =....................

(c) For Continuous Series (Grouped data in class interval):

Problem: The following data is related to height of tree in feet of 250 trees of a forest which are selected randomly. To calculate median height of tree of the given forest?

Height of tree (x_i)	35-40	40-45	45-50	50-55	55-60
No. of tree (f_i):	30	70	90	40	20

Solution:

Step-I to construct the following table:

Class Interval	**frequency(f_i)**	**cumulative frequency**
Total	N=$\sum_{i=1}^{n} f_i$=	

Step-II: To calculate $\frac{N}{2}$=.........................

Step-III: To see the value just greater than $\frac{N}{2}$ in cumulative frequency column =....................

Step-IV: corresponding class of step-III is called middle class

Step-V: Median, M_d = L + [h ($\frac{\frac{N}{2}-C}{f}$)] Where L is lower limit of middle class, h is magnitude value of class interval, f is frequency of middle class and C is cumulative frequency of preceding middle class.

Median, M_d=...
....

Mode: It is the good measure of central tendency for nominal data or categorical data. Mode is value of variable which have maximum frequency in the distribution. It is use to find average sizes. Mode gives good measures of central tendency in the skew distribution.

(a) For Individual Series (or Ungrouped data):

Problem: Suppose that 15 student's shoes size number are following:

8, 7, 6, 7, 10, 7, 8, 9, 9, 7, 7, 7, 8, 7, 7

Find average (mode) size of shoes?

Solution:

Mode =........................

(b) For Discrete Series (Grouped data or discrete frequency distribution):

Problem: Following table gives the category of 250 trees in a forest.

Tree's Name	A	B	C	D	E
No. of trees	12	150	50	20	18

Find the measure of central tendency (mode) category of trees.

Solution:

Mode =........................

(c) For Continuous Series (Grouped data in class interval):

Problem: The following data is related to milk production in liter of 225 farmers of a village which are selected randomly. To calculate average (mode) milk production of the given village?

Milk in liter(x_i)	0 - 4	4 - 8	8 - 12	12-16	16 - 20	20-24	24-28
No. of farmer (f_i):	30	90	60	12	8	4	1

Solution:

Step-I: See the maximum frequency. Correspond class to maximum frequency is called model class.,f_1=.............

Step-II: Mode = L + [$\frac{h(f_1-f_0)}{2f_1-f_0-f_2}$] where f_1 is frequency of model class, f_0 is frequency of preceding model class, f_2 is frequency of succeeding model class, L is lower limit of model class and h is magnitude value of class interval.

Mode

=...

....

....

....

....

....

....

Practical No. 5

Objective: To calculate Average (Geometric Mean and Harmonic Mean) of ungrouped and grouped data.

Geometric Mean is very suitable average for geometric progression data (or relative change and index data). It is calculate as the n^{th} root of the product of n numbers.

Geometric mean, G= $[x_1 \times x_2 \dots x_n]^{\frac{1}{n}}$

(a) For Individual Series (or ungrouped data):

Problem: The following data are related to Cow population in a village every fifth year.

No. of fifth year	First	Second	Third	Fourth	Fifth	Six
Population	80	120	185	300	448	680

Find the average (geometric mean) population of cow?

Solution:

Step-I

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

Step-II:

Geometric mean, G= Antilog $[\frac{\sum_{i=1}^{n} log x_i}{n}]$ =...

....

....

(b) For Discrete Series (Grouped data or discrete frequency distribution

Problem: Find the Geometric mean of the given data:

x_i	2	4	8	16	32
f_i	4	5	7	6	3

Solution:

Step-I:

x_i	f_i	log x_i	f_i log x_i
	N=$\sum_{i=1}^{n} f_i$=		$\sum_{i=1}^{n} f_i log x_i$=

Step-II:

Geometric mean, G= Antilog $[\frac{\sum_{i=1}^{n} f_i log x_i}{N}]$ =...

....

....

(c) For Continuous Series (Grouped data in class interval):

Problem: Find the Geometric mean of the given data:

x_i	0 - 4	4 – 8	8 - 12	12-16	16 - 20	20-24	24-28
f_i	5	7	10	15	11	8	4

Solution:

Step-I: To construct table:

Class Interval ($x_i - x_{i+1}$)	frequency (f_i)	Mid-Value $m_i=\frac{x_i+x_{i+1}}{2}$	log m_i	f_i log m_i
	$N=\sum_{i=1}^{n} f_i$=			$\sum_{i=1}^{n} f_i log m_i$=

Step-II: Geometric mean, G= Antilog $[\frac{\sum_{i=1}^{n} f_i log m_i}{N}]$ =..

....

....

Harmonic Mean: It is used to determine average speed. It is a reciprocal of arithmetic mean of reciprocals of the data.

Problem: Find Harmonic mean of the given data

1/3, 1/4, 1/5, 1/6, 1/7

Solution: Step-I:

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

Step-II:

Harmonic mean, H = $\frac{n}{\sum_{i=1}^{n} \frac{1}{x_i}}$ =..

....

....

(b) For Discrete Series (Grouped data or discrete frequency distribution):

Problem: A person went from City-A to City-B by different transport mode which speed and cover distance are given below:

transport mode	by foot	Taxi	train	Airplane	taxi
Speeds in km/h(x_i)	5	30	70	800	40
distances in km (f_i)	1	10	430	1200	50

Find the average speed (Harmonic Mean) of the complete journey.

Solution:

Step-I:

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

Step-II: Harmonic mean, H $=\frac{\sum_{i=1}^{n}\frac{f_i}{x_i}}{N}$ =...

....

....

(c) For Continuous Series (Grouped data in class interval):

Problem: Find the Harmonic mean of the given data:

x_i	10-20	20-30	30-40	40-50	50-60	60-70	70-80
f_i	2	3	5	8	6	4	2

Solution:

Step-I:

Class Interval ($x_i - x_{i+1}$)	frequency (f_i)	Mid-Value $m_i=\frac{x_i+x_{i+1}}{2}$	$\frac{1}{m_i}$	$\frac{f_i}{m_i}$
	$N=\sum_{i=1}^{n} f_i=$			$\sum_{i=1}^{n} \frac{f_i}{x_i}$ =

Step-II: Harmonic mean, H $=\frac{\sum_{i=1}^{n}\frac{f_i}{m_i}}{N}$

=..

....

....

Practical No. 6

Objective: To calculate Percentiles and Quartiles of ungrouped and grouped data.

Percentile: The percentile values divide the distribution into 100 parts each containing 1 percent of the cases. The kth percentile is that value below which k percent of values in the distribution fall.

(a) For Individual Series (or ungrouped data):

Problem: The following data is related to plant height in cm. Compute the 20th percentile height of the plants?

20, 22, 12, 14, 18, 15, 21, 22, 24, 25, 3, 10, 26, 28, 32, 33, 39

Solution:

Step-I: arrange the data in ascending order

....

Step –II: count number of observation (or values), n =...............

Step –III: kth percentile = $[k(\frac{n+1}{100})]^{th}\ item$

20th percentile = $[20(\frac{n+1}{100})]^{th}\ item$ =...

....

....

(b) For Discrete Series (Grouped data or discrete frequency distribution):

Problem: The following data is related to yearly income of 80 farmers of a village which are selected randomly. Compute the 30th percentile income of farmers of the given village?

Income in Lakh(x_i):	2	3	4	5	6	7	8	9	10
No. of farmer (f_i):	1	3	6	12	18	17	13	8	2

Solution: Step-I: To construct the following table

:

Variable (x_i)	Frequency (f_i)	Cumulative Frequency
Total	N=$\sum_{i=1}^{n} f_i$=	

Step-II: To calculate $\frac{k.N}{100}$ =..............

Step-III: To see the value just greater than $\frac{k.N}{100}$ in cumulative frequency column =.........

Step-IV: In first column, see the value of x correspond to the value getting in step-III,

kth percentile =

30th percentile =

(c) For Continuous Series (Grouped data in class interval):

Problem: The following data is related to height of tree in feet of 270 trees of a forest which are selected randomly. To calculate 30th percentile height of tree of the given forest?

Height of tree (x_i)	35-40	40-45	45-50	50-55	55-60	60-65	65-70
No. of tree (f_i):	30	70	90	40	20	12	8

Solution: Step-I to construct the following table:

Class Interval	**frequency(f_i)**	**cumulative frequency**
Total	N=$\sum_{i=1}^{n} f_i$=	

Step-II: To calculate $\frac{k.N}{100}$ =...............

Step-III: To see the value just greater than $\frac{k.N}{100}$ in cumulative frequency column =..........

Step-IV: corresponding class of step-III is called k^{th} percentile class

Step-V: k^{th} percentile = L + [h $(\frac{\frac{k.N}{100} - C}{f})$] Where L is lower limit of kth percentile class, h is magnitude value of class interval, f is frequency of k^{th} percentile class and C is cumulative frequency of preceding k^{th} percentile class.

30th percentile =L + [h $(\frac{\frac{30.N}{100} - C}{f})$] =...

....

....

....

....

Quartiles: The quartiles divide the distribution into 4equal parts so there are three quartiles say Lower Quartile (Q_1), Middle Quartile (Q_2) and Upper Quartile (Q_3).

(a) For Individual Series (or ungrouped data):

Problem: The following data is related to plant height in cm. Compute the First Quartile (Q_1) height of the plants?

20, 22, 12, 14, 18, 15, 21, 22, 24, 25, 3, 10, 26, 28, 32, 33, 39

Solution:

Step-I: arrange the data in ascending order

....

Step –II: count number of observation (or values), n =................

Step –III: First Quartile (Q1) = $[(\frac{n+1}{4})]^{th}\ item$ =..

....

....

(b) For Discrete Series (Grouped data or discrete frequency distribution):

Problem: The following data is related to yearly income of 80 farmers of a village. Compute the Third Quartile (Q_3) income of farmers of the given village?

Income in Lakh (x_i):	2	3	4	5	6	7	8	9	10
No. of farmer (f_i):	1	3	6	12	18	17	13	8	2

Solution: Step-I: To construct the following table

Variable (x_i)	Frequency (f_i)	Cumulative Frequency
Total	N=$\sum_{i=1}^{n} f_i$=	

:

Step-II: To calculate $\frac{3N}{4}$ =..............

Step-III: To see the value just greater than $\frac{3N}{4}$ in cumulative frequency column =..........

Step-IV: In first column, see the value of x correspond to the value getting in step-III,
Third Quartile (Q_3) =

(c) For Continuous Series (Grouped data in class interval):

Problem: The following data is related to height of tree in feet of 270 trees of a forest. Calculate Third Quartile (Q_3) height of tree of the given forest?

Height of tree (x_i)	35-40	40-45	45-50	50-55	55-60	60-65	65-70
No. of tree (f_i):	30	70	90	40	20	12	8

Solution: Step-I: To construct the following table

Variable (x_i)	Frequency (f_i)	Cumulative Frequency
Total	N=$\sum_{i=1}^{n} f_i$=	

:

Step-II: To calculate $\frac{3N}{4}$ =..............

Step-III: To see the value just greater than $\frac{3N}{4}$ in cumulative frequency column =..........

Step-IV: Corresponding class is called Third Quartile class

Step-V: Third Quartile (Q_3) = L +$\left[\frac{\frac{3N}{4}-C}{f}\right]\times h$ =...................................

Practical No. 7

Objective: To Standard Deviation and Coefficient of Variance (CV).

Standard Deviation: It is the positive square root of the arithmetic mean of the square values of the deviation from arithmetic mean. The notation of standard deviation is σ.

Coefficient of Variation: It is the standard deviation in terms of percentage of arithmetic mean. It is a relative (or unit less) measure of dispersion.

Coefficient of Variance (CV) = $\frac{\sigma}{\overline{x}} \times 100$ where $\overline{x}$ is arithmetic mean.

Variance = (Standard Deviation)^2 = σ^2

Coefficient of Standard Deviation = $\frac{\sigma}{\overline{x}}$

(a) For Individual Series (or ungrouped data):

Problem: The yield of wheat of 10 plots are given below. Find the standard deviation and Coefficient of Variance (CV) in yield data.

22.0, 22.5, 22.9, 23.0, 23.2, 23.1, 23.4, 23.0, 24.1, 23.6 in quintal.

Solution:

Step-I: To calculate Arithmetic Mean, $\overline{x}$ = $\frac{\sum_{i=1}^{n} x_i}{n}$ =...

Step-II: Construct a table:

x_i	$x_i - \overline{x}$	$(xi - \overline{x})^2$
$\sum_{i=1}^{n} x_i$=		$\sum_{i=1}^{n}(xi - \overline{x})^2$=

Step-III: Standard Deviation, σ = $\frac{\sum_{i=1}^{n}(xi - \overline{x})^2}{n}$=...

....

Step-IV: Coefficient of Variance (CV) = $\frac{\sigma}{\overline{x}} \times 100$ =..

(b) For Discrete Series (Grouped data or discrete frequency distribution):

Problem: The following data is related to income of farmer. Compute the standard deviation and CV of the data.

Income ('00000') Rs.	8	9	10	11	12	13	14	15
No. of Farmers	2	5	16	20	10	10	5	2

Solution:

Step-I: To calculate Arithmetic Mean, $\overline{x}=\frac{\sum_{i=1}^{n} f_i x_i}{N}$ =...

Step-II: Construct a table:

x_i	f_i	$f_i x_i$	$x_i - \overline{x}$	$(xi - \overline{x})^2$	$f_i(xi - \overline{x})^2$
		$\sum_{i=1}^{n} f_i x_i$=			$\sum_{i=1}^{n} f_i (xi - \overline{x})^2$ =

Step-III: Standard Deviation, σ = $\frac{\sum_{i=1}^{n} f_i (xi - \overline{x})^2}{N}$=...

....

Step-IV: Coefficient of Variance (CV) = $\frac{\sigma}{\overline{x}} \times 100$ =..

(c) For Continuous Series (Grouped data in class interval):

Problem: Find the standard deviation, coefficient of standard deviation, variance and CV of the given data.

Class Interval	10-20	20-30	30-40	40-50	50-60
frequency	8	15	45	20	12

Solution:

Step-I: To calculate Arithmetic Mean, $\overline{x}$ = $\frac{\sum_{i=1}^{n} f_i m_i}{N}$ =...

Step-II: Construct a table:

Class Interval	f_i	Mid-Value (m_i)	$f_i m_i$	$m_i - \overline{x}$	$(m_i - \overline{x})^2$	$f_i(m_i - \overline{x})^2$
			$\sum_{i=1}^{n} f_i m_i$=			$\sum_{i=1}^{n} f_i (m_i - \overline{x})^2$=

Step-III: Standard Deviation, σ = $\frac{\sum_{i=1}^{n} f_i (m_i - \overline{x})^2}{N}$=...

....

Coefficient of Standard Deviation = $\frac{\sigma}{\overline{x}}$ =..

Variance = (Standard Deviation)^2 = σ^2 =..

Coefficient of Variance (CV) = $\frac{\sigma}{\overline{x}} \times 100$ ....

Practical No.-8

Objective: To test significant difference between means in case of single sample, two samples (independent) and two samples (paired) using t-test (small sample test)

t - Test for single Mean: Used when we want to test significant difference between sample mean and population mean or testing a sample come from a specific population which mean is specific.

Problem: A sample of 10 trees which height are 10.5, 10.4, 10.8, 11.3, 12.5, 12.7, 11.5, 11.8, 12.1 and 11.5 meter. Test whether the sample comes from a forest which trees mean height is 11.

Solution:

Step-I (a): Null Hypothesis H_0: Here we set up the null hypothesis that there is no significant difference between sample mean and population mean.

(b): Alternative hypothesis H_1: we set up the alternative hypothesis that there is significant difference between sample mean and population mean.

Step-II: to calculate test statistic t

x_i	$x_i - \overline{x}$	$(x_i - \overline{x})^2$
$\sum_{i=1}^{n} x_i$ =		$\sum_{i=1}^{n}(x_i - \overline{x})^2$ =

Sample mean, $\overline{x} = \frac{\sum_{i=1}^{n} x_i}{n}$ =..

$S = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \overline{x})^2}{n-1}}$ =..

$t = \frac{\overline{x} - \mu}{S/\sqrt{n}}$ =..

....

Step-III: Conclusion: If calculated value of |t| is .... than tabulated value of t at 5 % level of significance then null hypothesis is..

....

....

t - Test for difference Mean: It is used when we want to test significant difference between two sample means or testing two samples comes from a same population.

Problem: A sample of 10 trees from forest A which height are 11.5, 11.4, 11.8, 12.3, 12.5, 13.3, 12.5, 12.8, 13.1 and 11.5 meter and second sample of 8 trees from forest B which height are 12.8, 12.3, 12.5, 13.8, 12.5, 12.8, 13.1 and 14.1 meter. Test whether forest A and B have same average height trees or not?

Solution:

Step-I:(a) Null Hypothesis H_0: Here we set up the null hypothesis that there is no significant difference between two sample means.

(b) Alternative hypothesis H_1: we set up the alternative hypothesis that there is significant difference between two sample means.

Step-II: to calculate test statistic t

x_i	$x_i - \overline{x}$	$(x_i - \overline{x})^2$
$\sum_{i=1}^{n_1} x_i$=		$\sum_{i=1}^{n_1}(x_i - \overline{x})^2$=

Sample mean, $\overline{x} = \frac{\sum_{i=1}^{n_1} x_i}{n_1}$ =...

y_i	$y_i - \overline{y}$	$(y_i - \overline{y})^2$
$\sum_{i=1}^{n_2} y_i$=		$\sum_{i=1}^{n_2}(y_i - \overline{y})^2$=

Sample mean, $\overline{y} = \frac{\sum_{i=1}^{n_2} y_i}{n_2}$ =...

$S = \sqrt{\frac{\sum_{i=1}^{n_1}(x_i - \overline{x})^2 + \sum_{i=1}^{n_2}(y_i - \overline{y})^2}{n_1 + n_2 - 2}}$ =...

....

$t = \frac{\overline{x} - \overline{y}}{S\sqrt{[\frac{1}{n_1} + \frac{1}{n_2}]}}$ =...

....

Step-III: Conclusion: If calculated value of |t| is ....than tabulated value of t at 5% level of significance then null hypothesis is ....

....

....

Paired t-test: It is used for testing significance difference between two sample means in paired data.

Problem: The following given yield data is related to before and after applying a soil treatment

Before treatment	7.9	8.5	7.3	9.7	10.3	10.2	11.1	8.9	8.5
After treatment	9.6	10.1	8.9	10.8	12.1	11.5	11.0	10.1	8.9

Test whether there is any significant effect of the soil treatment on the yield or not?

Solution:

Step-I:(a) Null Hypothesis H_0: Here we set up the null hypothesis that there is no significant difference between two sample means.

(b) Alternative hypothesis H_1: we set up the alternative hypothesis that there is significant difference between two sample means.

Step-II: to calculate test statistic t

x_i	y_i	$d_i = y_i - x_i$	$d_i - \overline{d}$	$(d_i - \overline{d})^2$
		$\sum_{i=1}^{n} d_i$=		$\sum_{i=1}^{n}(d_i - \overline{d})^2$=

Sample mean, $\overline{d} = \frac{\sum_{i=1}^{n} d_i}{n}$ =...

S = $\sqrt{\frac{\sum_{i=1}^{n}(d_i - \overline{d})^2}{n-1}}$ =...

t = $\frac{\overline{d}}{S/\sqrt{n}}$ =...

....

Step-III: Conclusion: If calculated value of |t| is....................than tabulated value of t at 5% level of significance then null hypothesis is...

....

....

Practical No.-9

Objective: To test goodness fit of the distribution and association between two attributes using chi Square test.

Chi Square test for goodness of fit: It is used for testing significant difference between Observed (or Experimental) and Expected (or Theoretical) frequencies.

Problem: The no. of deaths due to covid-19 over the days of week is following:

Day's	Sun	Mon	Tue	Wed	Fri	Sat	Sun
No. of death	14	11	13	10	12	9	15

Test whether death due to covid-19 is distributed uniformly over the days of week or not?

Solution:

Step-I:(a) Null Hypothesis H_0: Here we set up the null hypothesis that there is no significant difference between observed frequencies and expected frequencies.

(b) Alternative hypothesis H_1: we set up the alternative hypothesis that there is significant difference between observed frequencies and expected frequencies.

Step-II: Calculate test statistic, χ^2

O_i	E_i	$O_i - E_i$	$(O_i - E_i)^2$	$\frac{(O_i - E_i)^2}{E_i}$
				$\chi^2 = \sum_{i=1}^{n} \frac{(O_i-E_i)^2}{E_i}$=

Step-III: Conclusion: If calculated value of $\chi^2$is .... than tabulated value of χ^2at α % level of significance then null hypothesis is...

....

....

Chi Square test for Independence: Used for testing the significance of independence of two attributes.

Problem: The following data is related to eye colour father and their son.

		Father's eye colour		
		Blue	Black	Total
Son's eye colour	*Blue*	70	30	100
	Black	20	8o	100
	Total	90	110	200

Test whether there is any association between father's eye colour and son's eye colour or not?

Solution:

Step-I: (a) Null Hypothesis H_0: Here we set up the null hypothesis that there is no association between two attributes.

(b) Alternative hypothesis H_1: we set up the alternative hypothesis that there is any association between two attributes.

Step-II: calculate expected frequencies:

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

Expected frequencies,

E$(\alpha\beta)$ = $\frac{(\beta)(\alpha)}{N}$=...

E$(A\beta)$ = $\frac{(\beta)(A)}{N}$=..

E(αB) = $\frac{(B)(\alpha)}{N}$=..

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

Step-III: to calculate test statistic, χ^2

O_i	E_i	$O_i - E_i$	$(O_i - E_i)^2$	$\frac{(O_i - E_i)^2}{E_i}$
				$\chi^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$=

Step-IV: Conclusion: If calculated value of $\chi^2$is ....than tabulated value of χ^2_αat α % level of significance then null hypothesis is...

....

Practical No. 10

Objective: To calculate correlation co-efficient between two variables

Karl Pearson Correlation Coefficient: It measures degree of relationship between two variables and gives numerical value of the correlation which is called coefficient of correlation. The range of correlation coefficient is -1 to 1. Karl Pearson correlation coefficient is only apply in ratio and interval scale.

Problem: Find the correlation coefficient between height and weight of yield of the plants. Data are given below:

Height in cm	6	7	8	9	10
weight in gm	20	23	24	26	26

Solution:

Step-I: To construct table:

X	y	x²	y²	xy
$\sum_{i=1}^{n} x_i$=	$\sum_{i=1}^{n} y_i$=	$\sum_{i=1}^{n} x_i^2$=	$\sum_{i=1}^{n} y_i^2$=	$\sum_{i=1}^{n} x_i y_i$=

Step-II: (a) n= number of paired observations =

(b) $\overline{x} = \frac{\sum_{i=1}^{n} x_i}{n} =$

(c) $\overline{y} = \frac{\sum_{i=1}^{n} y_i}{n} =$

Step-III: (a) $\sigma_x = \sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^2 - \overline{x}^2} =$

(b) $\sigma_y = \sqrt{\frac{1}{n}\sum_{i=1}^{n} y_i^2 - \overline{y}^2} =$

(c) Cov(x, y) = $\frac{1}{n}\sum_{i=1}^{n} x_i y_i - \overline{x} \times \overline{y} =$

Step-IV: Karl Pearson correlation coefficient, $r_{xy} = \frac{Cov(x,y)}{\sigma_x \sigma_y} =$

Spearman's Correlation Coefficient: It is also known as rank correlation coefficient. It is useful to measure correlation coefficient when data is ordinal.

Problem: Find the rank correlation coefficient between height and biomass of the plants. Data are given below:

Rank in Height	1	2	3	4	5	6	7	8
Rank in biomass	1	2	3	5	6	4	7	8

Solution:

Step-I: Count number of paired observation, n =

Step-II: To construct table

R_x	R_y	$d_i = R_y - R_x$	d_i^2
			$\sum_{i=1}^{n} d_i^2$=

Step-III: Rank correlation coefficient, $\rho = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n(n^2-1)}$ =

Practical No. 11

Objective: To fit linear regression equation on the given data.

Linear Regression: It gives linear relationship between dependent and independent variables.

Problem: Fit the regression equation of y (yield in kg) on x (number of root fibers) of turmeric crop from the following data.

No. of roots	8	7	5	10	11	9	12	14	13
Yield (in kg)	1.2	1.1	0.7	1.3	1.3	1.0	1.4	1.3	1.4

Solution:

Step-I: To construct table:

X	y	x^2	y^2	xy
$\sum_{i=1}^{n} x_i$=	$\sum_{i=1}^{n} y_i$=	$\sum_{i=1}^{n} x_i^2$=	$\sum_{i=1}^{n} y_i^2$=	$\sum_{i=1}^{n} x_i y_i$=

Step-II: (a) n= number of paired observations =

(b) $\overline{x} = \frac{\sum_{i=1}^{n} x_i}{n}$ =

(c) $\overline{y} = \frac{\sum_{i=1}^{n} y_i}{n}$ =

Step-III: (a) $\sigma_x = \sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^2 - \overline{x}^2}$ =

(b) $\sigma_y = \sqrt{\frac{1}{n}\sum_{i=1}^{n} y_i^2 - \overline{y}^2}$ =

(c) Cov(x, y) = $\frac{1}{n}\sum_{i=1}^{n} x_i y_i - \overline{x} \times \overline{y}$ =

Step-IV: Karl Pearson correlation coefficient, $r_{xy} = \frac{Cov(x,y)}{\sigma_x \sigma_y}$=

Step-V: regression coefficient, $b_{yx} = \frac{\sigma_y}{\sigma_x} \times r_{xy}$=

Step-VI: regression equation, $y - \overline{y} = b_{yx} (x - \overline{x})$

....

Practical No. 12

Objective: To learn about DOS Commands

DOS (Disk Operating System)…

....

....

....

....

....

....

....

Command	Description	Type
ansi.sys		
append		
arp		
assign		
assoc		
at		
atmadm		
attrib		
batch		
bcdedit		
break		
cacls		
call		
cd		
chcp		
chdir		
chkdsk		
chkntfs		
choice		
clip		
cls		

cmd		
color		
command		
comp		
compact		
control		
convert		
copy		
ctty		
date		
debug		
defrag		
del		
delete		
deltree		
dir		
diskcomp		
diskcopy		
doskey		
dosshell		
driverquery		
drivparm		
echo		
edit		
edlin		
emm386		
endlocal		
erase		
exit		
expand		
extract		
fasthelp		
fc		
fdisk		
find		

findstr		
for		
format		
ftp		
fType		
goto		
graftabl		
help		
if		
ifshlp.sys		
ipconfig		
keyb		
label		
lh		
listsvc		
loadfix		
loadhigh		
lock		
logoff		
md		
mem		
mkdir		
mklink		
mode		
more		
move		
msav		
msd		
mscdex		
nbtstat		
net		
netsh		
netstat		
nlsfunc		
nslookup		

path		
pathping		
pause		
ping		
popd		
power		
print		
prompt		
pushd		
qbasic		
rd		
ren		
rename		
rmdir		
robocopy		
route		
runas		
sc		
scandisk		
scanreg		
set		
setlocal		
setver		
share		
shift		
shutdown		
smartdrv		
sort		
start		
subst		
switches		
sys		
telnet		
time		
title		

tracert		
tree		
Type		
undelete		
unformat		
unlock		
ver		
verify		
vol		
xcopy		

Practical No. 13

Objective: To learn Microsoft office Word

Microsoft office package: …………………………………………………………………………………….

……

……

……

……

……

Word Processing:………………………………………………………………………………………..

……

……

……

……

……

Word-wrap: ………………………………………………………………………………………………...

……

Justification: …………………………………………………………………………………………....

Adjustment: ……………………………………………………………………………………………….

Alignment: ……………………………………………………………………………………………...

Decimal Alignment: ………………………………………………………………………………..…..

Indents: ………………………………………………………………………………………………..

Insertion: ……………………………………………………………………………………...…..

Overstriking: ……………………………………………………………………………….....……

Deletion: ………………………………………………………………………………...……….

Search and Replace: ……………………………………………………………………………..…..

Copying or Cutting: …………………………………………………………………..……….

Boilerplate: ……………………………………………………………………………..…………..

……

Pagination: …………………………………………………………………………………………...

Page Numbering: ……………………………………………………………………………………….

Headers and Footers: …………………………………………………………………………………….

……

Footnoting: ……………………………………………………………………………………….…….

……

Table of Contents and Index Generators:.………………………………………………………...…..

Form Letter Merging: …………………………………………………………………………….…….

………..………………

Automatic Spelling Checker and Corrector. ………………………………………………………………………………

………..………………

………..………………

Exercise on Computer:

1. **How to open the MS word File**
2. **How to save the MS word File**
3. **How to change design and fonts size**
4. **How to Secure Word Document**
5. **Create the Table in the word.**
6. **How to insert the pics into word document**

Practical No. 14

Objective: To learn working with MS Excel

MS Excel: ……

……

……

……

……

……

……

Features:

Hyperlink. ……...

……

Clip art: ……..

……

Charts: ……..

……

……

Tables: ……

……

……

Macros: ……..

……

Database: ……..

……

Sorting and filtering. ………………………………………………………………………………………………………..

……

Data validations. …….

Grouping. ……

Page layout. ………

……

Exercise on Computer:

1. **How to insert image, in excel sheet**
2. **How to make a graph using Excel**
3. **How to create the charts in the excel sheet.**

Practical No. 15

Objective: To learn working with MS Excel and Formula Bar

MS excel: ……………………………………………………………………………………………………..

………

………

………

………

Exercise on Computer

1. Insert the formula in the excel sheet.

X	2	3	4	7	8
Y	8	5	6	71	12

Find the value of z; Z= (x+y), Z= (x-y), Z= (x*y)

………

………

………

………

………

………

………

………

………

………

………

………

………

………

………

………

………

………

………

………

………

………

………

Practical No. 16

Objective: To perform statistical analysis through MSExcel

Statistical analysis: ………………………………………………………………………………………………

………

………

………

Exercise on Computer

Find the correlation of given data

X	2	3	4	7	8
Y	8	5	6	71	12

………

………

………

………

………

………

………

………

………

………

………

………

………

………

………

………

………

………

………

………

………

………

………

………

………

Practical No. 17

Objective: To learn working with MS PowerPoint

Microsoft power point presentation: ……………………………………………………………………………

………

………

………

………

Design: ……

………

………

………

Animation: ………

………

………

………

………

………

Presentation: ……

………

………

………

Integration: ………

………

………

………

Different menu and Submenu of Power point: ……………………………………………………………………

………

………

………

………

………

………

………

………

………

………

Practical No. 18

Objective: To make presentation with MS PowerPoint

Exercise on Computer:

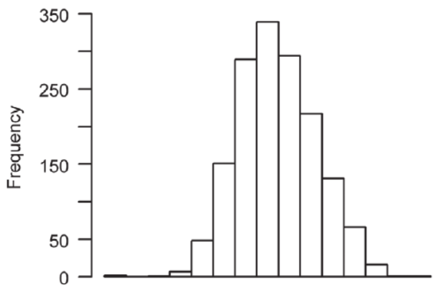
1. **Create First Slide using MS Power point**
2. **How to insert animation on slide**
3. **Create the PowerPoint presentation on introduction to computer.**
4. **What is the difference between ppt and pptx.**
5. **How to insert effect on ppt.**

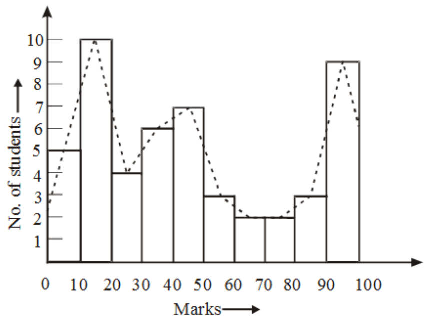
HISTOGRAM, POLYGON, FREQUENCY CURVE AND OGIVE

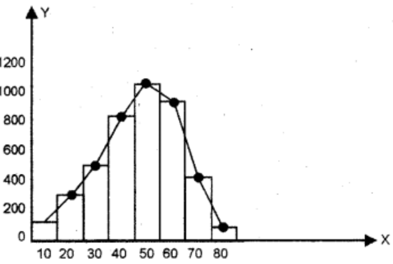
Histogram: Histogram is a part of graphical representation of frequency distribution. In which we plot class interval along with x-axis and frequency along with y axis with suitable scale and draw rectangle on x axis with proportional height of frequency.

Frequency Polygon: First draw histogram and then take mid-point on the top of each rectangle of histogram. After that we join mid-point by draw line with help of scale.

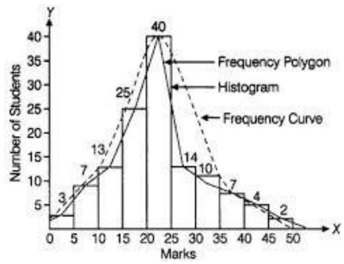
Frequency Curve: The procedure for drawing a frequency curve is same as for frequency polygon, but the points are joined by smooth or free hand curve.







Frequency Curve



Ogives: In which we plot class interval along with x-axis and cumulative frequency along with y axis with suitable scale. Ogives are also known as cumulative frequency curves and there are two kinds of cumulative frequency curves

Less Than Cumulative Frequency Curve: here the cumulative frequencies are plotted against the upper boundary of respective class interval.

More than cumulative frequency Curve: here the cumulative frequencies are plotted against the lower boundaries of respective class intervals.