B.COM., (IV SEMESTER)—16CCCCM8

BUSINESS TOOLS FOR DECISION MAKING

1. **What is Statistics?**

In simple ‘Statistics’ means numerical information expressed in quantitative terms.

In the beginning, it may be noted that the word ‘statistics’ is used rather curiously in two senses plural and singular. In the plural sense, it refers to a set of figures or data. In the singular sense, statistics refers to the whole body of tools that are used to collect data, organize and interpret them and, finally, to draw conclusions from the data.

1. **What are the various stages in Statistics?**



1. **What are the various methods of collecting primary Data?**

Primary data is the one, which is collected by the investigator himself for the purpose of a specific inquiry or study.

The primary data can be collected by the following five methods. 1. Direct personal interviews. 2. Indirect Oral interviews. 3. Information from correspondents. 4. Mailed questionnaire method. 5. Schedules sent through enumerators.

1. **Explain Secondary Data.**

Secondary data are those data which have been already collected and analyzed by some earlier agency for its own use; and later the same data are used by a different agency.

There is a vast amount of published information from which statistical studies may be made and fresh statistics are constantly in a state of production.

 The sources of secondary data can broadly be classified under two heads:

 1. Published sources, and

 2. Unpublished sources.

1. **What is meant by Discrete Series?**

A discrete distribution is one in which the data can only take on certain values, for example integers. Frequency of each value multiplied with the respective size All data have a frequency in this method.

AM= Σfx ÷Σf or N

1. **Define Individual Series**:

Individual Series is a statistical series in which the all the observations are listed out and all the observations have a frequency of 1

1. **Define continuous series.**

In continuous frequency distribution each individual frequency distribution is unknown an assumption in frequency distribution concentrate only in the midpoint or class interval

A continuous distribution is one in which data can take on any value within a specified range (which may be infinite) Ex.10-20,20-30,30-40,40-50 …..It has a class intervals

1. **What is class Interval?**

A class interval is a way to divide data and group certain answers together. The difference between lower limit and higher limit is called class interval

 Methods of classification of Data can be classified into Exclusive Method, Inclusive method.

1. **Define the diagrammatic presentation of statistical data.**

 A diagram is a visual form for presenting statistical data for highlighting the basic facts and relationship which are inherent in the data. The diagrammatic presentation is more understandable and it is appreciated by everyone. It attracts the attention and it is a quicker way of grasping the results saving the time. It is very much required, particularly, in presenting qualitative data.

1. **Define Bar Diagram?**

 Bar diagram is a diagram that presents data in the form of vertical rectangles or horizontal rectangles. Bar diagram is generally used to describe the development value of a research object within a certain time. The bar diagram shows descriptions with straight or horizontal bars and is the same width with separate trunks. Bar diagrams further classified into Multiple bar diagram, sub divided bar diagram and percentage bar diagram.

1. **What is Histogram?**

Histogram is the most Popular and widely used method of graphical presentation of data. Histogram is a set of vertical bars.

 A histogram is a block beam showing one kind of measurement of a process or event. This graph is very suitable for the data in the grouping. Histogram is a neighboring frequency diagram that looks like a bar chart

1. **What is mean by Pie Diagram**?

The pie chart is a diagram to describe or represent the data as a circle image. Because of its representation as a circle then the data must form a sum and each datum (data item) can be expressed as a percent (having a certain portion between 0 to 100) against the data.

1. **Define Ogive?**

When cumulative frequencies are plotted in a graph sheet We obtained frequency curve is called ogive or cumulative frequency curve it determine median, quartiles, percentiles. Class limits are in x axis cumulative frequency are in y axis. Cumulative frequency is plotted at the upper limit of the class interval, the successive points later join together to get an ogive curve, there are two methods of:

i)Less than ogive

ii)more than ogive

1. **What is tabulation?**

 Tabulation is a process of systematic and specific presentation of data. Tabulation is the process of summarizing classified or grouped data in the form of a table so that it is easily understood and an investigator is quickly able to locate the desired information. A table is a systematic arrangement of classified data in columns and rows. Thus, a statistical table makes it possible for the investigator to present a huge mass of data in a detailed and orderly form.

1. **What do you mean by classification of data?**

Classification is the process of arranging available facts into homogeneous groups and classes according to resemblances and similarities.

 “The process grouping large number of individual facts or observation on the basis of similarities among the items are called classification

1. **Measures of Central Tendency-Define**

In statistics, the three most common **measures of central tendency** are the **mean**, **median**, and **mode.**

1. **What is an Arithmetic Mean?**

It is also called Arithmetic average .AM is the most commonly used method of measure the central tendency. Arithmetic Mean is Adding the numbers in a data set and dividing by how many numbers there are.

AM=Sum of variables /No, of observation

1. **Write a short note on Mean, Median &Mode.**

**Mean** is Adding the numbers in a data set and dividing by how many numbers there are.

**Median** is the middle number in a data set when the numbers are listed in either ascending or descending order.

 **Mode** is the value that occurs the most often in a data set

1. **What is meant by ‘MODE’?**

In statistics, the three most common **measures of central tendency** are the **mean**, **median**, and **mode.**

Mode means the value that occurs most frequently in the statistical distribution. It Is otherwise called as the value which occurs the greatest number of times in a series or group of data Example: Marks obtained by 10 students in a class are 17,42,35,42,35,80,72 and 65. In this example the variable 42 occurred the maximum number of time. Hence 42 is called mode of the distribution.

**Mode** is the value that occurs the most often in a data set, It is derived from the Latin Word ‘La mode ‘which means fashion.

MODE=3MEDIAN -2 MEAN

1. **Write the methods of locate mode Graphically.**

Draw a histogram and locate points X AND Y Axis Highest rectangular will be the mode.

1. **Write down the formula for Mode?**

Individual observation: **Mode** is the value that occurs the most often in a data set.

Discrete series: Highest frequency related X called mode or Prepare grouping table.

Continuous series=L+ f1-f0/2f1-f0-f2\*i

1. **What are the various types of mode ?**

i)Unimodal – 1,2,3,3,4,5

ii)Bimodal -1,2,2,3,4,4,5

iii)Trimodal - 1,2,2,3,3,4,5,6,6.

iv)Multimodal – 1,1,2,2,2,3,3,3,3,4,4,4,5,5,5,5

1. **What is Harmonic mean?**

It is like a geometric mean. It is a measure of central tendency in solving special types of problems. Reciprocal of the arithmetic averages is called Harmonic mean.

Ex ;7 is $\frac{1}{7}$ , 9 is $\frac{1}{9}$

1. **Define the term Dispersion.**

 It measures of variation of the item. The degree in which numerical data tend to spread about an average values is called the variation or dispersion of the data.

1. **Find the Quartile Deviation and its co-efficient Q1=40, Q3=60**

Q.D=Q3-Q1/2

60-40/2=10

co-efficient of Quartile Deviation

 Q3-Q1/Q3+Q1=60-40/60+40=0.2

1. **Define Range.**

The range is the simplest measure of a dispersion. It is a rough measure of a dispersion.

Its measure depends upon the extreme items and not on all the items. It does not tell as anything about the dispersion of values in the series relative to a typical value. Thus

Range=Largest value - Smallest value

Relative measure of dispersion: R=L-S

Co-efficient of Range= L-S/L+S

1. **Find Range and its Co-efficient of Range**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 27 | 32 | 16 | 15 | 10 | 30 | 15 | 29 | 19 | 35 |

Range=L-S

35-10=25

Co-Efficient of Range=L-S/L+S

35-10/35+10=25/45=.55

1. **What is standard deviation?**

Karl Pearson introduced the concept of standard deviation in 1893.it is the most important measure of dispersion and is widely as in many statistical formulae. Standard deviation is also called Root -mean square deviation or Mean error or Mean square error deviation. The Standard deviation is denoted by the Greek letter(sigma).

1. **Define Mean Deviation**. Mean Deviation is the arithmetic mean of deviation of a series computed from any measures of central Tendency.” ---“signs are ignored. It is a absolute measures & wide application mode

MD=ΣΙ D Ι÷N

Co-efficient of MD=MD ÷Mean or Median or Mode

1. **What is Quartile Deviation?**

It is the absolute measures of dispersion. The relative measures of dispersion between 2 quartiles. Q3&Q1

QD = Q3-Q1/2 Co- Efficient of QD=Q3-Q1/Q3+Q1

1. **Give the formula for co-efficient of Variation**.

Square of standard deviation is called variance, multiplied by 100 gives the CV

Therefore, co-efficient of Variation(CV)

CV= standard deviation / mean \*100. The standard deviation must be converted into relative measures of dispersion for the purpose of comparison relative measures is known as co-efficient of Variation.

1. **What do mean by Time series?**

A n arrangement of Statistical Data in accordance with time of occurrence in a chronological order is called Time Series. Data which we get at different points of time and its arrangement.

1. **List out the uses of time series**

Time series analysis is very useful in every sphere. It is important in economic business… etc.

i)It helps in understanding past behaviour and estimating future behaviour.

ii)It helps in planning and forecasting.,

iii)comparison between one period of data and another period of data.

1. **Define the method of least square**

A straight line trend can be fitted. This method can be used to explain linear and nonlinear i.e That is straight line trend and parabolic trend.

1. **List out the components Time Series Analysis**.

Components or elements of Time Series under 4 methods.

1.Secular Trend Method

2.Seasonal Variation Method

3.Cyclical fluctuation Method

4.Irregular or Random Fluctuation Method

1. **What is meant by Moving Average?**

 To obtain the trend values with the help of moving average. It is also based on the arithmetic mean. moving average can be calculated for 3,4,5,6,7,8, or 9 yearly period.,a+b+c/3, b+c+d/3, c+d+e/3

1. **What is Seasonal Variation?**

Economic and business activities mostly depend upon seasonal variations. A seasonal variation could not be identified when the expressed annually. It can identify only when the data are expressed monthly or quarterly. Ex Cool drinks

1. **What is Skewness?**

 “Skewness or asymmetry is the attribute of a frequency distribution that extends on one side of the class with the highest frequency that on the other.” -Simpson and Kafka.

“when a series is not symmetrical it is said to be a asymmetrical or skewed.” -croton and Cowden.

1. **Write down the formula for Karl Pearson’s coefficient of Skewness?**

The absolute skewness is Mean – Mode. This measures is not suitable for making valid comparison of the skewness in two or more distribution. Hence Mean, Mode difference divided by standard deviation.

SKp = Mean – Mode / standard deviation

1. **What do you mean by ‘Correlation’?**

Correlation means relationship between two variables and its degree

According to Ya Lun Chou, “Correlation analysis attempts to determine the degree of relationship between variables.”

According to w.I.King, “Correlation means that between two series or group of data there exits some causal connection.”

According to L.R..Conon, “If two or more quantities vary in sympathy, so that moments in one tend to be accompanied by corresponding movements in the other(s), then they are said to be Correlated

1. **What do you understand by Interpolation?**

Interpolation relates to the past. It is an estimation of the dependent variable from the given independent variables. It is freely used to compute the value of median and mode in continuous series.

In population census figure 1951,1961, 1971, 1981.But we need population of 1975 census in the given data is called interpolation.

1. **What is extrapolation?**

Extrapolation is subject to greater [uncertainty](https://en.wikipedia.org/wiki/Uncertainty) and a higher risk of producing meaningless results. Extrapolation may also mean extension of a [method](https://en.wiktionary.org/wiki/method), Extrapolation may also apply to human [experience](https://en.wikipedia.org/wiki/Experience) to project, extend, or expand known experience into an area not known or previously experienced so as to arrive at a knowledge of the unknown .

1. **Write a note on positive Correlation.**

Correlation means relationship between two variables and its degree. Correlation is classified into

1. positive and negative

2.simple and multiple

3. partial and total

4. linear and non linear

 Positive and negative correlation depend upon the direction of change of variable. If two variables move in the same direction, height and weight, price and supply are the examples of positive correlation.

1. **Define Rank Correlation**

It is one of the methods of correlation. Rank correlation is only applicable to individual observations. Ranking method original value is not taken into a/c.

P=1-6Σ$D^{2}$/$N^{3}$-N When repeated ranks are given

P=1-$\frac{6[ΣD^{2}+1/12(m^{3}-m)+1/12(m^{3}-m)}{N^{3}-N}$$Type equation here.$

1. **What do you mean by Regression Analysis?**

 Regression means going back and it is mathematical measures showing the average relationship between two variables “Regression is the measures of the average relationship between two or more variable in terms of original unit of data

Here x is a random variable and y is a fixed variable. Sometimes both the variables may be random variables.

1. **What is an Index Numbers?**

An index number in [statistics](https://www.toppr.com/guides/business-economics-cs/descriptive-statistics/application-of-statistics/) is a tool that we generally use to measure the difference in relative changes from time to time. The difference can also be from place to place. It can be thought of as the [arithmetic](https://www.toppr.com/guides/maths/sequences-and-series/arithmetic-progression/) mean that we use to find or represent some values of a particular data set.

1. **How would you classify the index Numbers?**

It is a numerical value characterizing the change in complex phenomena over a period of time or space.

Classification:

1.Unweighted

a. Simple aggregate

b. Simple average price relative

2.Weighted

a. weighted aggregate

b. Weighted average price relative

1. **What is Time Reversal Test?**

Reversibility is an important property that an index number should possesses. A good index number should satisfy the time Reversal tests.

P01X P10=1

1. **What is Factor Reversal Test?**

Interchanging the price and quantities without giving inconsistent results that is two results multiply together should give true value ratio. A good index number not only satisfy the time reversal test also satisfy factor reversal test. P01XQ01=$ΣP1Q1/ΣP0Q0$

1. **What is cost of Living Index ?**

Cost of Living Index is also known as consumer price index, or price of living index. Consumer price index numbers are designed to measure the average change over time in the price paid by ultimate consumer for a specified quantity of goods and services. Consumer price indices or cost of living indices measure the change in the cost of living of workers due to the change in the retail price. A change in price level affects the cost of living of different classes of people differently.

1. **Write a short note on Family Budget Index**

 In this method, the family budgets of a large number of people are carefully studied and the aggregate expenditure of the average family for various items is estimated. These values are used as weights. The current year’s prices are converted into price relatives on the basis of the base year’s prices, and these price relatives are multiplied by the respective values of the commodities in the base year. The total of these products is divided by the sum of the weights and the resulting figure is the required index numbers.
 $P\_{ON}$ = $\frac{ΣWI}{ΣW}$

**Here, I=**$\frac{P\_{n}}{p\_{0}} $**×100 and   W=**$P\_{0}q\_{0}$