AMC Group of Educational Institutions, Manissery

Department of Commerce

e- resources for students

M.Com First Semester

(Question Bank with Answers)

Quantitative Techniques for Business Decision

Contents

Short Questions

1. What do you mean by non - parametric tests?

Ans :- A non parametric test is a test which is not concerned with testing of parameters. Non parametric tests do not depend on the particular form of the distribution of the population. That is, they do not make any assumption regarding the form of the population. A non parametric test is used only if the researcher concludes that a parametric test is not applicable.

2. What is meant by analysis of variance?

Ans :- Analysis of variance may be defined as a technique which analysis the variances of two or more comparable series (or sample) for determining the significance of differences in their Arithmetic means, and for determining whether different samples under study are drawn from same population or not, with the help of the statistical technique called F - test.

3. What are sign tests?

Ans :- t - test is generally used when sample is small and there is an assumption that the population is normal. Therefore, when sample is small but it is not possible to make assumption about the nature of population distribution, t - test cannot be applied. In such a case sign test is used. In sign test, to find the value of test static, we use the proportion of signs "+ve or -ve signs) that is why, the test is known as sign test.

4. What is meant by CODING with reference to Analysis of variances?

Ans :- Coding method refers to the addition, multiplication, subtraction or division of data by a constant or common factor. In the computation of analysis of variance, the final quantity tested is a ratio, hence dimensionless. Thus, the original values can be coded to simplify calculations without the need for any subsequent adjustment of the results. Coding can be applied for both one way and two way analysis

5. What are the characteristics of Analysis of Variance?

Ans :- It makes statistical analysis of variances of 2 or more samples.

It tests whether the difference in the means of different samples is due to chance or due to any significant cause.

It uses the statistical test called F - test by finding the appropriate variance ratio.

6. What is the null hypothesis in the analysis of variance?

Ans :- Null hypothesis of the test is "population means are equal" or " all the samples belong to the population having same variance".

7. What do you mean by Rank Sum Tests?

Ans :- In rank sum test we are replacing the values by ranks .All values are taken together and they are assigned ranks. Rank tests applied to test whether the population are identical. Two important rank sum tests are Wilcoxon - Mann - Whitney test (or U test) and The Kruskal - Wallis test (or H - test).

8. What is One sample sign test ?

Ans :- One sample sign test is a very simple non-parametric test applicable, when (1) sample is taken from a continuous population, (2) P(sample value<mean)=1/2 and P (sample value>mean)=1/2.

9. What are the different types of variances in 2 way classification?

Ans :- Variance between samples due to column variable.

Variance between samples due to row variable.

Variance within samples.

Variance about the samples .

10. What are the basic steps involved in Non-Parametric test?

Ans :- 1) Make the assumption necessary for the validity of the test procedure

2) Collect the sample data.

3) Lay down the null and alternative hypothesis.

4) Decide the test procedure and alternative hypothesis.

5) Decide the degree of freedom and level of significance.

6) Obtain the table value.

7) Arrive at a conclusion.

11) Define the term Correlation?

Ans) Correlation is defined as " the tendency of two or more groups or series of items to vary together directly or inversely ". According to A. M. Tuttle " Correlation is an analysis of the association between two or more variables ". The word correlation usually implies cause and effect relationship, that is, mutual interdependence. The different type of correlation are, Positive and Negative Correlation, Linear and Non- linear Correlation, Simple, Multiple and Partial Correlation.

12) What are the method of studying correlation?

Ans) a) Scatter Diagram

This is a graphical method of studying correlation between two variables. One of the variables is shown on the X - axis and the other on the Y- axis. Each pair values is plotted on the graph by means of a dot mark. The scatter diagram is a visual aid to show the presence or absence of correlation between two variables.

b) Correlation Graph

Under this method, separate curves are drawn for the X variable and Y variable on the same graph paper. The values of the variables are taken as ordinates of the points plotted.

c) Coefficient of Correlation

Coefficient of Correlation ia an algebraic method of measuring correlation. Under this method we measure correlation by finding a value known as the coefficient of Correlation using an appropriate formula. Correlation coefficient is a numerical value. It shows the degree or the extent of Correlation between two variables.

13) What do you mean by karl Pearson's method?

Ans) It is a method of computing Coefficient of Correlation. It is a mathematical method. The karl Pearson's method popularly known as pearsonian coefficient of Correlation is most widely used in practice. Karl Pearson the great biologist and statistician has given a formula for calculation of coefficient of Correlation. The pearsonian coefficient of

Correlation is denoted by the symbol 'r'. This is also known as product moment Correlation coefficient.

14) What are the properties of Correlation coefficient?

Ans) 1. Correlation coefficient has a well defined formula.

2. Correlation coefficient is a pure number and is independent of the units of measurement.

3. It lies between -1and +1.

4. Correlation coefficient does not change with reference to change of origin or change of scale.

5. Coefficient of Correlation between x and y is same as that between y and x.

15) Define Probable Error?

Ans) Probable Error of the coefficient of Correlation is a statistical measure which measures reliability and dependability of the value of coefficient of Correlation. If probable error is added to or subtracted from the coefficient of Correlation it would give two such limits within which we can reasonably expect the value of coefficient of Correlation to vary. Usually the coefficient of Correlation is calculated from samples. For different samples drawn from the same population the coefficient of Correlation may vary. But the numerical value of such variations is expected to be less than the probable error.

16) Define Regression Analysis?

Ans) Regression Analysis in the general sense means the estimation or the prediction of the unknown value of one variable from the known value of the other variable. It is statistical device used to study the relationship between two or more variables that are related. In the words of M. M. Blair " Regression Analysis is a mathematical measure of the average relationship between two or more variables in terms of the original units of the data ".

17) What is Line of best fit.?

Ans) It is also known as Regression line. When the given bivariate data are plotted on a graph, we get the scatter diagram. If the points of the scatter diagram concentrate around a straight line, that line is called the line of best fit. That is the line of best fit is that line which is closer to the points of the scatter diagram. This line is also known as Regression line. So regression line is a graphic technique to show the functional relationship

between the dependent and the independent variables. It shows average relationship between the variables.

18) What do you mean by Freehand Curve method?

Ans) A Freehand Curve method is an easy method for obtaining a regression line. According to this method original data are plotted on a graph paper. Usually original data when plotted on a graph gives a wave like curve but it depicts a general tendency of the data. Independent variable is taken along the horizontal axis and dependent variable along the vertical axis. We draw a smooth Freehand line in such a way that it clearly indicate the tendency of the original data. This line is fitted merely by inspection.

19) Define Method of least squares (curve fitting).?

Ans) Method of least squares is a method of drawing regression line by applying principle of least squares. The principle of least squares is that principle which States that the line of best fit should be drawn in such a manner that the sum of the square of different between the known values of the dependent variable and the corresponding values of it obtained from the line of best fit should be the least.

20) What are the limitations of regression?

Ans) The regression equation are based on the given values only. It is presumed that the trend of the given values will continue in future also. But this need not happen and therefore the regression equations need not show the correct relation for the future also.

* we are only considering LINEAR relationship.

* r and least square regression are not resistant to outliers.

* Extrapolation is dangerous.

* There may be variable other than x which are not studied, yet do influence the response variables.

21)What is SPSS ?

Ans:SPSS is used by market researchers, health researchers, survey companies, government entities, education researchers, marketing organizations, data miners, and many more for the processing and analyzing of survey data.SPSS is used by market researchers, health researchers, survey companies, government entities, education researchers, marketing organizations, data miners, and many more for the processing and analyzing of survey data.

22)What are the other statistical packages useful for analysis ?

Ans:SPSS :-offers the ability to easily compile descriptive statistics, parametric and non-parametric analyses.

R(R Foundation for Statistical Computing):-It does however come with an active community engaged in building and improving R and the associated plugins, which ensures that help is never too far away.

Graphpad prism:-GraphPad Prism is premium software primarily used within statistics related to biology, but offers a range of capabilities that can be used across various fields.

SAS(Statistical Analysis Software):-SAS is a statistical analysis platform that offers options to use either the GUI, or to create scripts for more advanced analyses.

Microsoft Excel:-While not a cutting-edge solution for statistical analysis, MS Excel does offer a wide variety of tools for data visualization and simple statistics. It's simple to generate summary metrics and customizable graphics and figures, making it a usable tool for many who want to see the basics of their data.

23)What are the major benefits of SPSS compared with other packages ?

Ans:The advantages of using SPSS as a software package compared to other are: SPSS is a comprehensive statistical software. Many complex statistical tests are available as a built in feature. Interpretation of results is relatively easy.

The very first version of SPSS was released in 1968 and was developed by Norman Nie, Dale Bent and Hadlai Hull. In 1975 the SPSS Inc. was originally founded. The earlier versions of the software were designed for processes such as batch processing. In 1983 the SPSS version 10 called SPSS-X was released which could contain multiple records. Since then many other versions...

ADVANTAGES:

The advantages of using SPSS as a software package compared to other are:

- SPSS is a comprehensive statistical software.

- Many complex statistical tests are available as a built in feature.

- Interpretation of results is relatively easy.

- Easily and quickly displays data tables.

- Can be expanded.

24) Examples of Statistical Functions ?

Ans:Statistical functions available in excel that are particularly relevant to the analysis techniques:

1.AVERAGE:-Returns the arithmetic mean of the given number .

2.MAX:-Returns the maximum value of the given numbers.

3.MIN:-Returns the minimum value of the given numbers .

4.MEDIAN:-Returns the median of the given numbers

25) What is the Use of Excel for qualitative methods?

Ans:You can use Excel to code and track themes in your qualitative data. You can create new themes and sub-themes in additional columns and clearly see What themes each participant's responses relate to. Create a new Excel worksheet tab for each question so your data stays organized and manageable.An approach to decision-making in evaluation that involves identifying the primary intended users and uses of an evaluation and then making all decisions in terms of the evaluation design and plan with reference to these.The aim of this short guide is to provide an introduction to using Microsoft Excel for quantitative data analysis within the context of a business and Management Research project.it covers some of the key features of Excel that are particularly useful when doing a research project .even if using bespoke statistical software,Excel can be helpful when preparing data for analysis in those packages.many basic analysis project involving primarly data exploration,descriptive statistics and inferential statistics can be successfully completed using standard Excel.

26)What is Variable View?

Ans:Variable View contains descriptions of the attributes of each variable in the data file. In Variable View: Rows are variables. Columns are variable attributes.In the variable View that we can see and edit the information that defines each variables sometimes called metadata in data set each columns of the data view is described by a row of the vaiable view.The first attributes of each variable is its name . The variable name is how the data column is identified in the programming language, and in order for the programming language to work gragracefully variable names have to abide by certain

restrictions :names must begin with a letter,and may be made up of the characters , numerals,non-punctuation Characters,and the period . Capitalization is ignored . Variable names may be up to 64 chctcharactersl.variable name may be added or changed simply by typing them in.However , just to make things confusing , SPSS allows us to select among several different standard formats for displaying numeric data and calls it type.we can set the variable type by clicking in the column,thon the gray button that appears and working in a dialog box .

27) What is Annova and Tests?

Ans:ANOVA tables are a core concept in statistics,and they are produced by several different commands in SPSS,including oneway,glm,and unianova.The unianova command is perhaps the easiest to use overall,because it allows to use string variables as factors . If we are doing a one way ANOVA and our factor is coded in numeric form,then one way is even easier to use.

28) What is Pivot tables?

Ans:A pivot table is a table of statistics that summarizes the data of a more extensive table. This summary might include sums, averages, or other statistics, which the pivot table groups together in a meaningful way. Pivot tables are a technique in data processing.Pivot tables provide a way of generating summaries of your data and organizing data in ways that are more useful for lparticulartasks.they are extremely useful for creating contingency tables,cross-tabulation and tables of means or other summary statistics .

29)What is the disadvantages of SPSS?

Ans:SPSS is not free and is fairly pricey. For those requiring advanced procedures not on SPSS, programmable coding, like in R Statistics, may be more limited and not as robust and use of R or other programs might be needed.

*SPSS can be expensive to purchase for students.

*Usually involves added training to completely exploit all the available features.

*The graph features are not as simple as of Microsoft Excel.

*Lack of many regressionanalysis techniques .

*Limited functionality for complex analysis .

*SPSS does not have very strong data management tools

*It can be slow at times if not using the latest hardware .

 * It can be expensive .

*No intuitive to use.

* Graphic features not as simple as Excel .

*Difficulties in editing graphs.

*Difficult to edit adopted .

*Need to have some some of the newer grgraphs such as dot plots,comparative bar,chart etc.

30)Compare SAS STATA & SPSS?

 Each package offers its own unique strengths and weaknesses. As a whole, SAS, Stata and SPSS form a set of tools that can be used for a wide variety of statistical analysis. With Stat/Transfer it is very easy to convert data files from one package to another in just a matter of seconds or minutes. Therefore, there can be quite an advantage to switching from one analysis package to another depending on the nature of your problem. For example, if you were performing analysis using mixed models you might choose SAS, but if you were doing logistic regression you might choose Stata, and if you were doing analysis of variance you might choose SPSS. If you are frequently performing statistical analysis, we would strongly urge you to consider making each one of these packages part of your toolkit for data analysis.

<p align="center">**Short Essay Questions**</p>

1.Discuss different type of significance test?

      There are numerous type of significance tests, according to situations and criteria of testing. Test may be;

     1.parametric and non parametric: on the basis of focus of the test, they will be classified as parametric and non parametric tests. On certain tests, assumption about population distribution can be made. For example in large sample test or z test we assume that samples are drawn from population following normal distribution. Such tests which are based

assumptions about population are called pata parametric tests. All mean tests, proportion tests, variance test are non parametric test.

But in certain situations it is not possible to make an assumption about population, distribution, from which samples are drawn. Besides they do not focus on parameters like mean, proportion or variance. Such tests are called non parametric test. Non parametric test include chi square test, rank test, sign test etc.

2.Sample and large sample tests: According to the number of items included in a sample, tests can divided as small sample tests and large sample tests. If a test includes a sample of size less than 30,it is small sample test. If the size is 30 or more, it is large sample test.

Small sample tests follows t students distribution large samples test follow normal distribution. Mean test may be conducted as large or small sample tess. But proportion are conducted large sample test only.

3.One tailed or two tailed test: One the basis of location of rejection region, test may be one tailed or two tailed. When a test examine the significance of difference of either more than a specific value or less than a specific value, rejection appears only on one side of the curve. It is called 1 tailed test.

When test examine both more than or lower than a specific value at the same time, rejection region appears on both sides of the curve, and such test is called two tailed test. Most of the tests are two tailed teststs.

2.What is F test? Mention uses of F tests?

F-test is used to determine whether the two independent estimates of population variance significantly different between themselves or to establish whether both variables have come from the same universe. For this purpose the ratio of variance given by two samples are used. This ratio is called F ratio. Ftest is named in the honour of sri. R. A. Fisher.

The assumptions of F test are;

1. The population for each sample is normally distributed with the same or identical mean or variance.

2.All observations are randomly selected.

3.F distribution is always formed by the ratio squared values. Therefore it can never be a negative number.

Uses of F distribution are;

- F test is used to test the hypothesis.
- It is used to test the linearity of regression.
- It is used to test the equality of means of three or more population.
- It is used for testing the significance of an observed sample multiple correlation.
- It can be used to test the equality of variances of two population when samples are drawn.

3. Distinguish between one way and two way analysis of variance?

1. One way Annova only involves one factor or independent variable, with as there are two independent variables in two way Annova.

In one way Annova one factor or independent variable Analysised as three or more Categorical groups. A two way Annova compares multiple group of two factors.

2. One way Annova needs to satisfy only two principles of design of experiment ie, replication and randomisation. Two way Annova which meets all three principles of design of experiments which are replication, randomisation, and local control.

3. Analysis of variance Table one way Annova

Source of variation	Sum of squares	DF	MS	Variance ratio of F
Between the samples With the samples	SSC SSE	C-1 n-c	MSC=SSC/(C-1) MSE=SSE/n-c	MSC/MSS
Total	SST	n-1		

Analysis of variance table Two way Annova

Source of variation	Sum of squares	D. F	Mean squares	F ratio
Between samples (Coloumn) Between rows Residual	SSC SSR SSE	(c-1) (r-1) (c-1) ×(r-1)	MSC=SSC/(C-1) MSR=SSR/(r-1) MSE=SSE/(r-1) (c-1)	MSC/MSE MSR/MSE
Total	SST	n-1		

4.What are the merits of non parametric test?

Non parametric tests are used when assumptions assumed by the parametric tests are not met or are questionable. All test involving ranked data are non parametric. In non parametric tests no assumptions are made about the parameters population or population from which we draw our sampl. These tests are distribution free.

The important assumptions of non parametric test are;

    1. Sample observation are independent.

   2.The variables are continuous.

   3. Sample drawn is a random.

   4.Observations are measured on ordinal scale.

The merits of non parametric test are;

- The main advantage of distribution free tests are that they proud more power than traditional tests when the samples are from highly skewed distributions.
- **Assumption free :** They do not require a assumption that a population is distributed in the shape of probability distribution.
- **Simple:** They are simple to understand, explain and solve. No mean, variance, SD to be used.

- Measurement: They requires only basic level of measurement, such as ordinal and normal. Rank and signs form basis of non parametric tests.
- **Flexible:** Non parametric test allow considerable deviation from normal distribution, it need to follow the rigid of normal curve and central limit of theorem.
- **Natural option:** It is the only choice when hypothesis is to be tested, but no Parameetres is available.
- **Realistic:** It's more realistic, because they considered observation themselves, rather than deviation like mean and standard deviation.

5.What is one sample sign test?

All tests concerning means that are based on the assumptions that samples are taken from a population having the shape of a normal distribution. When this assumption is not possible, or any statistic is not available, vital questions still arise;

Is there is significant difference between actual observations and theoretically expected observations

It is responsible to believe that samples have been taken from a probabilistic sampling distribution.

Is it reasonable to accept that the sample as a random sample from a population etc. One sample sign test is applicable, when sample is taken from a population which is continuous, in this case, the probability that the sample value is less than mean, and the probability that sample value is grater than mean Are both ie, half and half. Here the sign test is used to test hypothesis as a population median. The medin divide the distribution in to two equal halves. Now we may examine wheather two halves are equal or not. If these halves are equal or approximately equal, the distribution is and the same prediction is possible. This is the rationale behind one sample sign test.

In this test, each sample value greater than given median value is replaced by plus sign and each sample value less than the given value is replaced by negative sign. Then we test the hypothesis that these plus signs and minussigns ate equal or approximately equal or there is no significant difference between them.

6. What are the different type of Rank test?

In the rank sum tests we replace the values with ranks. All values are taken together and ranks are assigned. The major types of rank tests are Wilcoxon Mann Whitney Utest and kruskal wallis test(H test).

Kruskal wallis test (H) test:

The test is a non parametric test used to compare three or more samples. The H test is a non parametric test, and is used when the assumptions of ANOVA are not met. It is used to test the null hypothesis that all population have identical distribution functions against the alternative hypothesis that at least two of the samples differ only with respect to location, if at all. It is the analogue to the F test used in analysis of variance. While analysis of variance tests depend on the assumption that all populations under comparison are normally distributed, the H test places no such restriction on the comparison. It is a logical extension of the wilcoxon Mann Whitney test.

Wilcoxon - Mann-Whitney U test :

This test examines whether two samples groups come from same population, or if there is any significant variations between two variables. To perform this test we first of all the data jointly, considering them as belonging to a single sample, in either increasing or decreasing order. We start from low to high. If tie occurs, ranks will be distributed between them applying average principle.

7.What are the uses of chi- Square distribution?

Chi-squre test is a stastistical test used to test the significant of difference between observed frequencies and the corresponding theoretical frequency of a distribution, without any assumptions about the nature of distribution about the population. This is most popular widely used non parametric test. It was developed by pro.Karl Pearson.

The uses of chi square test are;

- **Used to test goodness of fit:** As a test for goodness of fit, chi squre test used to far the theoretical frequencies fit to the observed frequencies.
- **Used to test independence:** As a test of Independence, chi square test used to test the attributes of a sample are associated or not.
- **Used to test homogeneity:** As a test of homogeneity, chi square test used to test different samples are homogeneous as far as a particular attribute is concerned.
- **Used to test population variance:** Here, chi square test used for testing the given population variance when the sample is small. In other

words it is used to test whether there is any significant difference between sample variance and population variance.

8. Explain procedure for testing hypothesis

The steps followed in a test of hypothesis are

★ Set up a null hypothesis and an alternative hypothesis appropriate to the test to be conducted
★ Decide the appropriate test criteria (such as Z-test, t-test,t-test etc)
★ Specify level of significance desired. Usually level of significance specified is 5%(or1%) determine suitable level of significance. Hindi absence of any specific instruction, it should be normally 5%
★ Calculate the value of the test statistics using appropriate formula
★ Then obtain the table value of the test statistics corresponding to the level of significance and the degree of freedom
★ Make decision about accepting or rejecting the null hypothesis. The decision depends on whether the computed value of the test statistics falls in the acceptance region or rejection region. When d computed value of the test statistics is numerically less than the table value of the test statistics it falls in the acceptance region. Otherwise in the rejection region

9. Explain standard error. What are its uses?

The standard error (SE) of a statistic is the approximate standard deviation of a statistical sample population. The standard error is a statistical term that measures the accuracy with which a sample distribution represents a population by using standard deviation. In statistics, a sample mean deviates from the actual mean of a population; this deviation is the standard error of the mean.

Uses of standard error

1. S. E is used for testing a given hypothesis

2.S.E give an idea about the reliability of a Sample. The reciprocal of S.E is a measure of reliability of the sample

3. S.E can be used to determine the confidence limits for population values like mean, proportion and standard deviation

10. Difference between standard deviation and standard error?

- Standard deviation is a measure of dispersion of statistical data (or a probability distribution). Standard error is a measure of dispersion of a sampling distribution
- Standard deviation measures the variability or consistency of statistical series. Standard error determines precisious or reliability of an estimated value
- Standard deviation is calculated in relation to the mean of a series. Standard error can be calculated in relation to Mean, Median, S.D, correlation coefficient etc.
- Standard Deviation can be used to find out S.E standard error can be used for estimation and testing of hypothesis

11. What is z test and uses of z test?

A z-test is a statistical test used to determine whether two population means are different when the variances are known and the sample size is large. The test statistic is assumed to have a normal distribution, and nuisance parameters such as standard deviation should be known in order for an accurate z-test to be performed.

Users of Z-test

- To test the given population mean when the sample is large or when the population SD is known
- To test the equality of two sample means when the sample are large or when the population SD is known
- To test the population proportion
- To test the equality of two sample proportions
- To test the population SD when the sample is large
- To test the equality of two sample standard deviations when the sample are are large or when population standard deviations are known

Assumption in Z-test

1. Sampling distribution of test statistics is normal

2. Sample statistics are close to the population parameter and therefore for finding standard error, sample statistics are used in places worker population parameters are to be used

5.What is chi square test and the characteristics of chi square test?

Ans:A chi-squared test, also written as x^2 test, is a statistical hypothesis test that is valid to perform when the test statistic is chi-squared distributed

under the null hypothesis, specifically Pearson's chi-squared test and variants thereof. Pearson's chi-squared test is used to determine whether there is a statistically significant difference between the expected frequencies and the observed frequencies in one or more categories of a contingency table. In the standard applications of this test, the observations are classified into mutually exclusive classes. If the null hypothesis that there are no differences between the classes in the population is true, the test statistic computed from the observations follows a $x2$ frequency distribution.

Characteristics of chi square test

- It is a non- parametric test. Assumptions about the form of the distribution or its parameters is not required
- It is a distribution free test, which can be used in any e type of distribution of population
- It is easy to calculate chi square test statistics
- It analyses the difference between a set of observed frequencies and a set off corresponding expected frequencies

12. What are type 1 and type 2 errors?

 in any test of hypothesis the decision is to accept or to reject a null hypothesis. The decision is is based on the information supplied buy the sample data.the four possibilities of the decisions are:

1. Accepting a null hypothesis when it is true

2. Rejecting a null hypothesis when it is false

3. Rejecting a null hypothesis when it is true

4. Accepting a null hypothesis when it is a false

Obviously (1) and (2) are correct while (3) and (4) are errors. The last two cases are respectively known as type 1 and type 2 errors.

In the absence of the proper and adequate sample information it is likely that errors may occur. Type 1 error is committed by rejecting null hypothesis even when that hypothesis is true. Similarly type 2 error is committed by accepting a hypothesis when it is files.

13. What is statistics and parameter?

 In statistics, as opposed to its general use in mathematics, a parameter is any measured quantity of a statistical population that summarises or

describes an aspect of the population, such as a mean or a standard deviation. If a population exactly follows a known and defined distribution, for example the normal distribution, then a small set of parameters can be measured which completely describes the population, and can be considered to define a probability distribution for the purposes of extracting samples from this population.

parameter is to a population as a statistic is to a sample; that is to say, a parameter describes the true value calculated from the full population, whereas a statistic is an estimated measurement of the parameter based on a subsample. Thus a "statistical parameter" can be more specifically referred to as a population parameter

14 . Difine the term correlation and different kinds of correlation?

Correlation

Two variables are said to be correlation if the change in one variable results in a corresponding change in other variable. That is, when two variables move together, we say they are correlated. Hence price and supply are correlated.

Definition

According to A.M Tuttle " correlation is an analysis of the association between two or more variables".

Different kinds of correlation

The different types of correlation are :

1. Positive and negative correlation
2. Linear and non - linear correlation
3. Simple, multiple and partial correlation

1 . Positive and negative correlation :- Correlation can be either positive or negative . When the value of two variables move in the same direction, correlation is said to be positive ie an increase in the one variable alsi if a decrease in the value of one variable, results into a decrease in the value of the other variable also correlation is said to be positive. If on the other hand, the value of two variables move in opposite directions so that an increase in the value of one variables , results into a decrease in the value of one variable results into an increase in the value of the other variable the correlation is said to be negative.

2 . Linear and non- linear correlation :- correlation may be linear or non linear. When the amount of change in one variable leads to a constant ratio of change in the other variable , correlation is said to be linear.

3 . Simple , partial and multiple correlation :- in the study of relationship between variables, if there are only two variables, the correlation is said to be simple.eg . the correlation between price and demand is simple.

When one variable is related to a number of others, correlation is not simple. Then the study of relationship between the variables is done with the help of partial or relationship between the variables is done with the help of partial or multiple correlation.

In the study of multiple correlation we measure the degree of association between one variable on one side and all the other variables to gether on the other side. Thus the relationship between yield with both rainfall and temperature together is multiple correlation. In partial correlation we study the relationship of one variable with one of the other variables presuming that the other variables remain constant.

15 . What is multiple correlation? And What are the properties of multiple correlation?

When there are many variables and we want to study relation of one variable with all the other variables taken together, the correlation obtained is colled multiple correlation.

Properties of Multiple Correlation Coefficient

- Multiple correlation coefficient measures the closeness of the association between the observed values and the expected values of a variable obtained from the multiple linear regression.
- Multiple correlation coefficient between observed values and expected values, when the expected values are calculated from a linear relation of the variables determined by the method of least squares, is always grater than that when expected value are calculated from any other linear combination of the variables.
- $0 < R_{1.23} < 1$
- If $R_{1.23} = 1$, the association is perfect and all the regression residual are zero. In this case , multiple linear regression equation of X_1 onX_2 and X_3 may be said to be perfect prediction formula.
- If $R_{1.23} = 0$, then all total and partial correlation involving X_1 are zero. So X_1 is completely uncorrelated with all the other variables.

16 . What are multiple regression? Write down the multiple regression equation of X_1

onX$_2$ and X$_3$.

Multiple regression

It is there are three or more variables and we try to find out the effect of two or more independent variables on one dependent variable. Let X$_1$,X$_2$ and X$_3$ be three variables.Let X$_1$ for a dependent variable X$_2$ and X$_3$ be the two independent variables. Then we study, in multiple regression analysis, the relative movement of X$_1$ for a unit movement in X$_2$ and X$_3$.

A multiple regression equation is an equation for estimating a depending variables say X$_1$, from the independent variables X$_2$ and X$_3$ and it is called a regression equation of X$_1$ on X$_2$ and X$_3$.

Since there are more than two variables,we can use the term plane of regression.

The equation of plane of regression of X1 onX2 and X3 is

(X$_1$-X bar $_1$)= b$_{12.3}$(X$_2$-Xbar $_2$)+b$_{13.2}$(X$_3$-Xbar $_3$) where Xbar $_1$,Xbar $_2$, X bar means of X$_1$,X$_2$ and X$_3$.

Similarly equation of plane of regression of X$_2$ on X$_1$ and X$_3$ is X$_2$-X BAR$_2$ =b$_{21.3}$(X$_1$-XBAR $_1$)+b$_{23.1}$(X$_3$-XBAR $_3$) and equations of plane of regression of X$_3$ on X$_1$ and X$_2$ is X$_3$-XBAR $_3$= b$_{31.2}$ (X$_1$-Xbar $_1$)+b$_{32.1}$(X$_2$-XBAR $_2$).

17 . What is scatter Diagram? What are the merits and demerits?

Scatter Diagram

This is graphical method of studying correlation between two variables. One of the variables is shown on the X- axis and the other on the Y- axis. Each pair of values is plotted on the graph by means of a dot mark. After ,all the items are plotted we get as many dots on the graph paper as the number of points.if these points show some trend either upward or downward, the two variables are said to be correlated. If the plotted points do not show any trend , the two variables are not correlated. If on the other hand the tendency is reverse so that the points show a downward trend from the left top to the right bottom, correlation is negative.

The scatter Diagram is a visual aid to show the presence or absence of correlation between two variables. A line of best fit can be drawn using the method of leat square. This line will be as close to the points as possible,If the points are falling very close to this line, there is very high degree of correlation. If they lie Very much away from this line it shows that the correlation is not much.

Merits of scatter Diagram

- It is easy to plot the points.
- It is simple to understand.
- Abnormal values in the data can be easily detected.
- The value of dependent variable for a given value of independent variable can be detected.
- The extreme values do not affect it.

Demerits of scatter Diagram

- The degree of correlation cannot be easily estimated.
- Algebraic treatment of is not possible.
- When the number of pairs of observations is either very big or very samll, the method is not easy.

18 . Difine Degree of Correlation , and are the classifications?

 Degree of Correlation

The degree or the intensity of the relationship between two variables can be ascertained by finding the value of coefficient of correlation. The degree of correlation can be classified into three :

- Perfect correlation
- No correlation
- Limited degree of correlation

a . Perfect correlation

When the change in the two variables is such that with an increase in the value of one, the value of the other increase in a fixed proportion, correlation is said to be perfect. Perfect correlation may be positive or negative. Coefficient of correlation is +1 for perfect positive correlation and it is -1 for perfect negative correlation.

b . No correlation

If change in the value of one variable are not associated with change in the value of the other variables, there will be no correlation. When there is no correlation the coefficient of correlation is zero.

c . Limited degree of correlation

In between perfect correlation and no correlation there may be limited degree of correlation. Limited degree of correlation may be also be positive or

negative. Limited degree of correlation may be termed as high, moderate or low. For limited degree of correlation the coefficient of correlation lies between 0 and 1 numerically.

19 . Define correlation.What arethe uses and importance of correlation in Economics and Business activities?

Correlation

Correlation is defined as "the tendency of two or more groups or series of items to items to vary together directly or indirectly".

Uses of correlation

- It helps to study the association between two variables.
- Correlation measures degree of relation between two variables. Karal Pearson's coefficient of correlation provides a formula for finding the degree of relation between two variables.
- From the correlation coefficient, we can develop a measure called probable error. Probable error indicates whether the correlation is significant or not.
- Correlation analysis help to estimate the future values. For ex- ample, from the correlation coefficient between income and investment one can predict the possible quantum of investment for a particular amount of income.

Importance of correlation in economics and business activities

Correlation analysis is useful in understanding economic behaviour. It helps in finding out interrelated variables. For example, we can find out the factors which will stabilise the disturbance economic situation.

Economic theory and Business show relationship between two variables.eg. price and demand, advertisement expenditure and sales etc.are interrelated. Correlation analysis helps to study such relation. The relationship between two variables can be summed up by a single number called correlation coefficient.

Correlation study is useful in estimating the likely change in a variable with a particular amount of change in another variable.

Prediction of values of variables based on correlation analysis will be more reliable and near to reality. Therefore in Economics forecasting and Business forecasting, the correlation analysis is very useful.

20 . Explain Rank correlation. What are the merits and demerits of rank correlation?

(Ans). <u>Rank correlation</u>

Karal Pearson's coefficient is applicable when variables are measured in quantitative from. But in many cases measurement is not possible because they are in quantitative from.

There fore rank correlation is the correlation is the Correlation obtained from ranks, instead of their quantitative measurements.

Thus ,when the value of two variables are expressed in ranks and thereform correlation is obtained, that correlation is known as rank correlation. Spearman has devised a formula known as spearman's rank correlation coefficient to find the correlation coefficient to find the correlation coefficient from the ranks. According to spearman's method, the formula for :

Rank correlation coefficient $=1-6\pounds D^2/n(n^2-1)$

Merits and demerits of rank correlation

<u>Merits</u>

- It is easy to calculate.
- It is simple to understand.
- It can be applied to both quantitative and qualitative data.
  <u>Demerits</u>
  - Rank correlation coefficient is only approximate measure as the actual values are not used.
  - It is not convenient when 'n ' is large.
  - Further algebraic treatment is not possible.

Essay Questions

1)   Define the term Correlation.Expalin the various methods of studying correlation.

Ans   According to A.M.Tuttle "Correlation is an analysis of the association between two or more variables".

Bowley defines correlation as,"when two quantities are so related that the fluctuations in one are sympathy with the fluctuations of the other, that an increase or decrease of the one is found in connecconnection with an increase or decrease of the other and greater the magnitude of change in

one,greater is the magnitude of change in the other, the quantities are held to be correlated ".

The word correlation usually implies cause and effect relationship,that is mutual interdependence.

There are different kinds of correlation:-

a) positive correlation and negative correlation

b)Linear and nonlinear correlation

c)Simple,multiple and partial correlation

Methods for studying correlation or Measures of correlation

Correlation between two variables can be measured by both graphic and algebraic methods.

GRAPHIC METHODS:

a) Scatter diagram

b)Correlation graph

ALGEBRAIC METHOD

a) coefficient of correlation

I. SCATTER DIAGRAM

- This is a graphical method of studying correlation between two variables. One of the variable is shown on Xaxis and other on Y axis. Each value is plotted on the graph by means of a dot mark.
- If these points show either upward or downward, the two variables are said to be correlated.
- If the plotted points do not show any trend,the two variables are not correlated.

MERITS

- It is easy to plot the points
- It is simple to understand
- Abnormal values in the data can be easily detected
- The value of dependent variable for a given value of independent variable can be detected
- The extreme value donot affect it.

DEMERITS

- The degree of correlation cannot be easily estimated
- Algebraic treatment is not possible

II . CORRELATION GRAPH (GRAPHIC METHOD)

Here separate curves are drawn for the X variable and Y variable on the same graph paper.

The values of the variable are taken as coordinates of the points plotted. From the direction and closeness of the two curves we can infer whether the variables are related.

If not curve move in same direction,correlation is positive

If the curves move in opposite direction ,correlation is negative.

III. COEFFICIENT OF CORRELATION

Under this method we measure correlation by finding a value known as coefficient of correlation using an appropriate formula. Correlation coefficient is a numerical value.

It can be computed by applying the methods given below:

- Karl Pearson's method
- Spearman 's method
- Concurrent deviation method.

2) What is Regression analysis and its limitations?

Ans.   Regression analysis is a mathematical measure of the average relationship between two or more variables in terms of the original units of the data

- Simple regression: when there are only two variables the regression equation obtained is called simple regression equation.
- Multiple regression:Here there are more than two variables and we try to find out the effect of two or more independent variables on one dependent variable.

METHODS OF DRAWING REGRESSION LINES

a) free hand curve

We draw a smooth freehand line in such a way it clearly indicates the tendency of the original data.

b) The method of least square

Method of least square is a method of drawing regression line by applying principle of least square.

Regression equation are the equations of regression lines.There are two regression lines .They are

- Regression equation of Y on x
- Regression equation of x on y

The properties of regression lines are

- The two lines intersect
- When r=1, two lines coincide
- When r=0,the two lines are mutually perpendicular

Properties of regression coefficient are the sign of both regression coefficient will be same.that is both will be positive or both will be negative.

Product of regression coefficient is the square of correlation coefficient. When there is perfect correlation byx and bxy are reciprocal of each other.

Both the regression coefficient will not be greater than 1.

LIMITATIONS

The regression equation are based on the given value only.

It is presumed that the trend of the given values will continue in future also.

But this need not happen and therefore the regression equations need not show the correct relation for the future also.

3) Explain the term coefficient of correlation. What are the properties of correlation coefficient?

Ans   Coefficient of correlation is an algebraic method of measuring correlation. Under this method, we measure correlation by finding a value known as coefficient of correlation using appropriate formula.

Correlation coefficient is a numerical value.It shows the degree or the extent of correlation between two variables.

- Coefficient of correlation is a pure number lying between -1 and +1.
- when the correlation is negative, it lies between -1 and 0.
- when the correlation is positive, it lies between 0.
- when the correlation of coefficient is zero,it indicates that there is no correlation between the variables.

- When the correlation coefficient is 1,there is perfect correlation.
- Between no correlation and perfect correlation there are varying degrees of correlation.

Coefficient of correlation can be computed by applying the methods given below:
1. Karl Pearson's method
2. spearmans method
3. concurrent deviation method

KARL PEARSON 'S COEFFICIENT OF CORRELATION

It is popularly known as pearsonian coefficient of correlation Karl Pearson, the great biologist and statistician has given a formula for calculation of coefficient of correlation. The Pearsonian coefficient of correlation is denoted by the symbol 'r'.

Degree of correlation includes

- Perfect correlation
- No correlation
- Limited degree of correlation

RANK CORRELATION

The rank correlation is the correlation obtained from ranks,instead of their quantitative measurement.Spearman has devised a formula known as Spearmans rank correlation coefficient to find the correlation coefficient from the ranks.

CONCURRENT DEVIATION METHOD

This is a very simple method of measuring correlation. Under this method the directions of deviation are only taken.The magnitude of the values are ignored.

Under this method the nature of correlation is known from the direction of change in the values of variables.

If the deviation of two variables are concurrent then they move in the same direction, otherwise in the opposite direction.

PROPERTIES OF CORRELATION COEFFICIENT

- Correlation coefficient ha a well defined formula.
- Correlation coefficient is a pure numbnumber and is independent of the units of measurement.
- It lies between -1 and +1.

- Correlation coefficient does not change with reference to change of origin or change of scale.
- Coefficient of correlation between x and y is same as that between y and x.

4. Describe the various Methods of classifying of Quantitative Techniques?

Ans) <u>Classification of Quantitative Techniques</u>

There are different methods which can be broadly classified into three categories,

A. Mathematical Quantitative Techniques
B. Statistical Quantitative Techniques
C. Programming Quantitative Techniques

A. Mathematical Quantitative Techniques

A technique is which quantitative data are used along with the principles of mathematics is known as Mathematical Quantitative Techniques.

1) permutations and Combinations

2) Set Theory

3) Matrix Algebra

4) Determinants

5) Differential Calculus and Integral Calculus

6) Integration

7) Differential Equation

B. Statistical Quantitative Techniques

Statistical techniques are those techniques which are used in conducting the statistical inquiry concerning a certain phenomenon. Important methods are :

1) Collection of Data

2) Classification and Tabulation

3) Measures of Central tendency, Dispersion, Skewness and Kurtosis

4) Correlation and Regression analysis

5) Interpolation and Extrapolation

6) Index Numbers

7) Time Series Analysis

8) Statistical Quality Control

9) Ratio Analysis

10) Analysis of Variance

11) Theory of Attributes

C. Programming Quantitative Techniques

These techniques are model building techniques used by decision makers in modern times. Important Programming methods are :

1) Linear Programming

2) Queuing Theory

3) Inventory Theory

4) Theory of Games

5) Decision Theory

6) Network Programming

7) Simulation

8) Search Theory

9) Non Linear programmin

10) Dynamic Programming

11) Integer Programming

12) Parametric Programming

13) Branch and Bound Technique

14) Quadratic Programming

5. What is Normal distribution and What are the merits and demerits of normal distribution?

Ans) <u>Normal Distribution</u>

It is a probability distribution that is symmetric about the mean, showing that data near to the mean are more frequent in occurance than data far from the mean. In graph form, normal distribution will appear as a bell curve.

Advantages of Normal distribution

- Normal distribution is the mostly used distribution in Inferential statistics.
- Most of error of measurements and a large variety of physical observations have approximately Normal distributions.
- The measurements of linear dimensions of large number of articles produced may show individual variations. These follow Normal distributions.
- The standard normal distribution table shows exhaustively areas for the different intervals of the values of the variable.
- The Normal distribution has a number of mathematical properties.
- Most of the distributions in nature are either normal or that can be approximated to normal.

Disadvantages of Normal Distribution

- The variables which are not continuous cannot be normally distributed. Therefore many distributions in Economics like distribution of number of children per family, cannot be studied under Normal Distribution.
- The Normal distribution cannot be applied to situations where the distribution is highly skewed.

3. What is point estimation and What are the methods used for point estimation?

Ans) <u>Point Estimation</u>

Any strategic suggested as an estimate of an unknown parameter is called a point estimate of that parameter. So under point estimation, we determine a value which may be taken as an estimate of the population parameter.

<u>Methods used for Point estimation</u>

Commonly used methods for obtaining the point estimation are ;

- Method of Maximum likelihood
- Method of moments
- Method of minimum square
- Method of least squares
- Method of minimum chi-square
- Method of inverse probability

1. Method of Maximum likelihood

This is the most commonly used method of estimating the population parameter. Maximum likelihood estimation is based on the assumption that different populations generate different samples and that any given sample is more likely to have come from the given population than others.

2. Method of moments

This is a simple method. Under this method, we find moments from the sample and equate them with the corresponding moments of the population.

3. Method of minimum variance

It is a unbiased estimator that has lower variance than any other unbiased estimator for all possible values of the parameter.

4. Method of least squares

It is standard approach in regression analysis to approximate the solution of overdetermined systems by minimizing the sum of the squares of the residuals made in the results of every single equation.

5. Method of minimum chi-square

It is a method of estimation of unobserved quantities based on observed data among the consequences of its use is that the test statistic actually does have approximately a chi-square distribution when the sample size is large.

6. Method of inverse probability

It is the posterior distribution which depends both on the likelihood function and a prior distribution.