FIRST SEMESTER M.Ed. DEGREE EXAMINATION, 20M.Ed. 04

Prepared by: Anupama.VM (2022-24)

1) Explain the meaning of descriptive research?

A) Descriptive research aims to accurately and systematically describe a population, situation ,and phenomenon. It can answer what, where ,when and how questions, but not why questions. In descriptive research variables were only observed and measured.

2) Explain the steps in educational research?

A) Identifying the problem, formulating objectives and hypothesis, deciding the method of research, collecting the data, analysing and interpreting the data, writing a report.

3) Explain the meaning of purposive sampling?

A) Purposive sampling is a nonprobability sampling method. It is also known as judgment, selective, or subjective sampling in which researcher relies on his on her own judgment when choosing members of population to participate in the study.

4) Meaning of variables in research?

A) A variable is a characteristic that can be measured and that can assume different values. Height, age, income, province, or country of birth, grades obtained at school and type of housing are all examples of variable.

5) Explain the use of statistics in research?

A) Statistical methods involved in carrying out a study include planning, designing, collecting data, analysing, drawing meaningful interpretation and reporting of the research findings.

6**) Explain the use of primary sources and secondary sources of review of literature?**

A) When we do research we have to gather information and evidence from a variety of sources. Primary sources gives you direct access to the subject of your research. Secondary sources provide second hand information and commentary from other researchers. Examples include journal

articles, reviews, and academic books. A secondary source describes, interprets, or synthesize primary sources. Primary sources are more credible as evidence, but good research uses both primary and secondary sources.

A primary source is anything that gives you direct evidence about the people, events, or phenomena that you are researching. Primary sources will usually be the main objects of your analysis. If you are researching the past ,you cannot directly access it yourself, so you need primary sources that were produced at the time by participants or witnesses (eg; letters, photographs, newspapers).If you are researching something current, your primary sources can either be qualitative or quantitative data that you collect yourself (eg; through interviews, survey, experiments) or sources produced by people directly involved in the topic(eg; official documents or media text).

Primary source example

History

Letters and diaries

Photographs and video footage

Official documents and records

Physical objects

Art and literature

Novels and poems

Paintings and art installation

Films and performances

Communication and social studies

Interview transcripts

Recordings of speeches

Newspapers and magazines

Social media posts

A secondary source is anything that describes, interprets, evaluates or analyses information from primary sources. Common examples include

Books , articles, documentaries that synthesize information on a topic.

Synopses and description of artistic works

Encyclopedias and textbooks that summarize information and ideas.

Reviews and essays that evaluate or interpret something.

When you cite a secondary source, it's usually not to analyse it directly. Instead you will probably test it's arguments against new evidence, or use its ideas to formulate your own.

Secondary source examples
Primary source
Novel

Secondary source
Article analysing the novel

Primary source
Painting

Secondary source
Exhibition catalog explaining the painting

Primary source
Letters and diaries written by a historical figure

Secondary source
Biography of the historical figure

7) Explain the role of dependent and independent variables in research?

A)In research, variables are any characteristic that can take on different values, such as height ,age , temperature, or test scores. Researchers often manipulate or measure independent and dependent variables in studies to test cause and effect relationships. The independent variable is the cause. Its value is independent of other variables in the study. The dependent variable is the effect. Its value depends on changes in the independent variable.

An independent variable is the variable you manipulate or vary in an experimental study to explore its effects. Its called "independent " because it's not influenced by any other variables in the study. It is also called explanatory variables, predictor variables, right hand side variables.

Types of independent variables

There are two main types of independent variables

Experimental independent variables can be directly manipulated by the researchers.

Subject variables cannot be manipulated by researchers, but they can be used to group research subjects categorically.

A dependent variable is the variable is that changes as a result of the independent variable manipulation. Its the outcome you are interested in measuring, and it depends on your independent variable. In statistics dependent variables are also called response variables, outcome variables, left-hand variables.

The dependent variable is what you record after you have manipulated the independent variable. You use this measurement data to check whether and to what extent your independent variable influences the dependent variable by conducting statistical analysis.

Based on your findings, you can estimate the degree to which your independent variable variation drives changes in your dependent variable. You can also predict how much your dependent variable will change as a result of variation in the independent variable.

8) Discuss the need of classification of quantitative data?

A) Understanding data is not like understanding complexities of life. Data-specifically quantitative data is no different. In order to analyse and interpret this data type, you will need to understand nuances, details, and other factors.

Quantitative data is information that can be quantified. Its data that can be counted or measured , and given a numerical value. Quantitative variables can tell you "how much", "how many", or "how often".

Different categories of quantitative data include

Measurements counts, and calculations. For example, the attendees of a webinar.

Sensors: a photo sensor detecting the presence of light

Quantification of qualitative data: how a categorical group of people felt about a quiz.

Projections: Expected revenue for knowing what type of quantitative data also called numerical data. You are working with help teams apply the correct type of statistical analysis.

Since quantitative data is statistically analysed and mathematical in nature research can be extremely detailed. And due to its numerical nature ,there is less potential for bias .The results obtained are objective in nature, so they are extremely accurate. There are three types of quantitative data; discrete, continuous, and interval(as compared to ratio data).

Discrete data

Discrete data is information that can only taken numerical values. It is represented using tally charts, bar graphs or piecharts.

Continuous data

Continuous data is data that can take any value and can be infinitely broken down into smaller parts. It differs from discrete data, in that it can change its value.

Interval v/s ratio data

Difference between interval and ratio data derives from their ability to fall below zero. Interval scales can represent values below zero. Ratio variables ,on the other hand, never fall below zero .Ratio scale has an absolute zero. Quantitative data can be collected using sampling, analytic tools ,questionnaires, and open source data sets ,among others.

9) Differentiate evaluation study from cross sectional data?

A) Evaluation research also referred to as program evaluation or evaluative research,is the systematic assessment of the value of resources and time committed to a project, product or specific goal.The main objective of evaluation research is to determine whether or not a process has achieved a goal or yielded the desired results.

Evaluation research is an integral part of the product development process, especially in the early design phases,and is continually utilized until the product is finalized.It is also used to monitor user experience after the product is launched in the market by gathering their feedback.

Evaluative research uses quantitative and or qualitative research methods to collect customer feedback on aspects such as usability, accessibility, desirability etc to improve the product and ensure that the solution meet the users needs.

Cross-section of a study population,in statistics,and econometrics, is a type of data collected by observing many subjects (such as individuals,firms, countries or regions) at one point or period of time.The analysis might also have regard to differences in time.

Analysis of cross sectional data usually consists of comparing the differences among selected subjects.For example,if we want to measure current obesity levels in a population,we could draw a sample of 1000 people randomly from that population (also known as cross section of that population), measure their weight and height and calculate what percentage of that sample is categorized as obese.This cross sectional sample provides us with a snapshot of that population, at that one point in time.Note that we do not know based on one cross-sectional sample,if obesity is increasing or decreasing, we can only describe the current proportion.

10)Explain the meaning of multistage sampling in research?

A)In multistage sampling or multistage cluster sampling ,you draw a sample from a population using smaller and smaller groups (units) at each stage.Its often used to collect data from a large ,geographically spread group of people in national surveys.

Multistage sampling is often considered an extended version of cluster sampling.In multistage sampling , you divide the population into clusters and select some clusters at the first stage.At each subsequent stage ,you further divide up those selected clusters into small clusters,and repeat the process until you get to the last step.At the last step ,you only select some members of each cluster of your sample.But in multistage sampling,you don't need a sampling frame that lists every member of the population.That's why this method is used for collecting data from large dispersed populations.

Examples of multistage sampling

1)In the first stage,you make a list of school districts within the state.You select 15 school districts as your primary sampling units.

2) In the second stage, you list all schools within those school districts.You select to schools from each district as your secondary sampling units.

3)In the third stage, you obtain a list of all students within those schools.You select 50 students from every school as your ultimate sampling units,and collect data from those students.

11) Explain the use of normal probability curve?

A)The bell shaped curve is called normal probability curve or simply normal curve.It is not an actual distribution of scores or any test of ability or achievement; but is instead a mathematical model.The distributions of test scores approach the theoretical normal distribution as a limit,which is rarely,ideal and perfect.It was first discovered by Abraham de moivre, a French mathematician, who obtained it while working on certain problem in the game of chances.

Measurements of many natural phenomena,and of many mental and social traits under certain conditions tend to be described symmetrically about their means in proportions which approximate those of the normal probability distribution.Phenomena which follow the normal probability curve (at least approximately) may be classified as follows:

1) Biological statistics -Proportion of male to female births for the same country or community over a period of years , proportion of different types of plants and animals in cross fertilization (Mendeliean ratio).

2)Anthropometric data-height, weight etc for large groups of the same age and sex.

3) Social and economic data-birth rate, death rate, wage and output of large number of workers.

4) Psychological measurement-Intelligence as measured by standard tests ,speed of association, perception spun, reaction time .

5) Errors of observation -measures of height, speed of movement, linear magnitude, physical and mental traits , certain errors which likely cause to deviate above or below their true values.

12) Discuss the important features of anyone computer software which is used in research?

A)NVIVO is a software program used for qualitative and mixed methods research.Specifically , it is used for the analysis of unstructured text ,audio, video and image data, including (but not limited to interviews, focus groups, surveys, social media and journal articles).It is produced by QSR international.As of July 2014, it is available for both windows and Macintosh operating systems.

1)It is used to analyze and organize unstructured text, audio, video or imaged data.

2)Playback ability for audio and video files, so that interviews can be transcribed in NVIVO.

3) Ability to capture social media data from facebook and twitter using the Ncapture browser plugin.

4) Important notes and captures from Evernote -great for field research.

5) Important citations from Endnote, Mendeley, Zotero, or other bibliographic management software great for literature reviews.

6)Perform simple text analysis queries such as text search or word frequencies for text data in English, French,German,Spanish,Portuguese, Japanese and simplified Chinese .

13) Explain the uses of measures of histogram?

A)A histogram is very similar to barcharts.This chart is a combination of a line chart and a vertical bar chart.A histogram is used to display the distribution of information or data over a continuous time period.The area of the rectangular bars is proportional to the frequency of the variable.They are used to plot the frequency of variable occurrences in continuous data.

Uses of measures of histogram

1) Histograms help in displaying a large amount of data- graphically, that is difficult to be put into tabular form.

2)It makes it easier to display data that are of various types and frequencies.

3)It is useful for the visualization of the distribution of data.

4)With the use of a histogram the median ,distribution and variations in data can be found out.

5) Histogram tells us about the skewness of data plotted.

6)It is helpful in calculating the standard deviation of data.

7)It is helpful in decision making

14) a)Explain the different types of hypotheses with the help of illustrations?

b) Explain the use of measures of central tendency and variability in educational research?

A)a)A hypothesis is an assumption that is made based on some evidence.This is the initial point of any investigation that translates the research questions into predictions.It includes components like variables, population and the relation between the variables.A research hypothesis is a hypothesis that is used to test the relationship between two or more variables.

Simple hypothesis

It shows a relationship between one dependent variable and a single independent variable.For eg- If you eat more vegetables, you will lose weight faster.Here eating vegetables is an independent variable,while losing weight is the dependent variable.

Complex hypothesis

It shows the relationship between two or more dependent variables and two or more independent variables.Eating more vegetables and fruits leads to weight loss, glowskin and reduces the risk of many heart diseases such as heart disease.

Directional hypothesis

It shows how a researcher is intellectual and committed to a particular outcome.The relationship between the variables can also predict its nature.For eg; children aged four eating proper food over a five year period are having higher IQ levels than children not having a proper meal.This shows the effect and direction of the effect.

Null hypothesis

It provides a statement which is contrary to the hypothesis.Its a negative statement, and there is no relationship between independent and dependent variables.The symbol is denoted by "Ho".

b)Measures of central tendency are a combination of two words that is 'measure' and 'central tendency '.Measure means methods and central tendency means average value of any Statistical series.

Thus we can say that central tendency means the methods of finding out the central value or average value of a statistical series of quantitative information.

There are three measures of central tendency, such as:

1)The arithmetic mean

2)The Median

3)The mode

The mean

The mean is the sum of the value of each observation in a dataset by the number of observations.The mean can be used for both continuous and discrete numeric data.

The Median

The median is the middle value in distribution when the values are arranged in ascending or descending order.The median is less affected by outliers and skewed data than the mean,and is usually the preferred measure of central tendency when the distribution is not symmetrical.

The Mode

The Mode is the most commonly occuring value in a distribution.The mode has an advantage over the median and the mean as it can be found for both numerical and categorical (non numerical)data.

15**) a)Explain the need and importance of longitudinal studies in research with the help of illustrations?**

b) Write the use of survey on research in education?

A)a) Longitudinal studies employ continuous or repeated measures to follow particular individuals over prolonged periods of time -often years or decades.They are generally observational in nature, with quantitative and or qualitative data being collected on any combination of exposures and outcomes,without any external influenced being applied.This study type is particularly useful for evaluating the relationship between risk factors and the development of diseases,and the outcomes of treatments over different lengths of time.Similarly , because data is collected for given individuals within a predefined group, appropriate statistical testing may be employed to analyze change over time for the group as a whole,or for particular individuals.

Advantages of longitudinal studies

1)The ability to identify and relate events to particular exposures, and to further define these exposures with regards to presence ,timing ,and chronicity.

2)Establishing a sequence of events.

3) Following change over time in particular individuals within the cohort.

4) Excluding recall bias in participants , by collecting data prospectively and prior knowledge of a possible subsequent event occuring.

b)Due to its anonymous nature, surveys are perfect for privacy seeking students,adults,or even teachers.These powerful tools give them the space to air out concerns while offering in-depth insights into their way of thinking characteristics even prestigious institutions like IITs,IIMs,etc

can benefit from.It also enables a healthy dialogue with the students -clear signs of a system fostering good practices.On the education front,it helps improves the quality of education and courses being provided,enhances teaching methods and processes,and helps everyone understand the learning environment at large in a school or college.

ODUCTION TO EDUCATIONAL, RESEARCH AND STATISTICS

FIRST SEMESTER M.Ed. DEGREE EXAMINATION, NOVEMBER 2019

M.Ed. 04 **Prepared by: Dhanya. j (2022-24)**

INTRODUCTION TO EDUCATIONAL, RESEARCH AND STATISTICS

1)List any four types of qualitative research.

A) Descriptive, Correlational, Causal-Comparative/Quasi-Experimental, and Experimental research

2) Distinguish between dichotomous and continuous variables.

A) Dichotomous variables are nominal variables which have only two categories or levels. For example, if we were looking at gender, we would most probably categorize somebody as either "male" or "female".

Continuous variables are also known as quantitative variables. Continuous variables can be further categorized as either interval or ratio variables.

3) Enumerate the criteria that one can use to evaluate a research problem.

A) A research problem is deemed appropriate when it is supported by the literature, and considered significant, timely, novel, specific, and researchable. Stronger research problems are more likely to succeed in publication, presentation, and application.

4. Give two reasons for stating a hypothesis before the data-gathering phase of a quantitative study.

A) ● Hypothesis helps in making an observation and experiments possible.

● It becomes the start point for the investigation.

● Hypothesis helps in verifying the observations.

● It helps in directing the inquiries in the right directions.

5. Write a directional and a non-directional hypothesis based on the research question: "What is the relationship between the rate of maturation of adolescent boys and their self-concepts?" 6. Discuss the role of review of related literature in quantitative research.

• The literature review demonstrates the underlying assumptions (propositions) behind the research questions that are central to the research proposal.

• The literature review provides a way for the novice researcher to convince the proposal the reviewers that she is knowledgeable about the related research and the "intellectual traditions" that support the proposed study.

• The literature review provides the researcher with an opportunity to identify any gaps that may exist in the body of literature and to provide a rationale for how the proposed study may contribute to the existing body of knowledge.

• The literature review helps the researcher to refine the research questions and embed them in guiding hypotheses that provide possible directions the researcher may follow.

7. Write a short note on simple random sampling.

Simple random sampling: One of the best probability sampling techniques that helps in saving time and resources, is the Simple Random Sampling method. It is a reliable method of obtaining information where every single a member of a population is chosen randomly, merely by chance. Each

an individual has the same probability of being chosen to be a part of a sample. 8. Distinguish between Type I and Type II errors with the help of an example.

In statistics, a Type I error is a false positive conclusion, while a Type II error is a false negative conclusion.

Making a statistical decision always involves uncertainties, so the risks of making these errors are unavoidable in hypothesis testing.

Example: Type I vs Type II error

You decide to get tested for COVID-19 based on mild symptoms. There are two errors that could potentially occur:

Type I error (false positive): the test result says you have coronavirus, but you actually don't.
Type II error (false negative): the test result says you don't have coronavirus, but you actually do.

9) Explain the necessity of mastering online database searching with examples.

• To get accurate and pinpointed information on the topic of specialization • One can get the exact information from the published literature either from the print or digital source.

• Authenticity and reliability of data

• For both retrospective and current information

• searching and retrieving data with various combinations and extract the exact data from the database (Boolean, free text, browse, root search etc)

• faster than manual searching with less chances of missing data .

Wide coverage of information

• mUltiple database searching and removal of duplication is possible • user friendly searching tips

• Save time, space, money, effort, person-hours etc

• Easier to assess quality information for problem solving .

10) Discuss the criteria and qualities of good research.

One expects scientific research to satisfy the following criteria:

1. The purpose of the research should be clearly defined and common concepts be used.

2. The research procedure used should be described in sufficient detail to permit another researcher to repeat the research for further advancement, keeping the continuity of what has already been attained.

3. The procedural design of the research should be carefully planned to yield results that are as objective as possible.

4. The researcher should report with complete frankness, flaws in procedural design and estimate their effects upon the findings.

5. The analysis of data should be sufficiently adequate to reveal its significance and the methods of analysis used should be appropriate. The validity and reliability of the data should be checked carefully.

6. Conclusions should be confined to those justified by the data of the research and limited to those for which the data provide an adequate basis.

7. Greater confidence in research is warranted if the researcher is experienced, has a good reputation in research and is a person of integrity.

In other words, we can state the qualities of a good research as under:

1. Good research is systematic: It means that research is structured with specified steps to be taken in a specified sequence in accordance with the well-defined set of rules. Systematic characteristic of the research does not rule out creative thinking but it certainly does reject the use of guessing and intuition in arriving at conclusions.

2. Good research is logical: This implies that research is guided by the rules of logical reasoning and the logical process of induction and deduction are of great value in carrying out research.
Induction is the process of reasoning from a part to the whole whereas deduction is the process of reasoning from some premise to a conclusion which follows from that very premise. In fact, logical reasoning makes research more meaningful in the context of decision making.

3. Good research is empirical: It implies that research is related basically to one or more aspects of a real situation and deals with concrete data that provides a basis for external validity to research results.

4. Good research is replicable: This characteristic allows research results to be verified by replicating the study and thereby building a sound basis for decisions.

11) Write a short note on APA style of referencing.

● The American Psychologicali Association is an organization created for individuals in the psychology field.

● With close to 121,000 numbers, they provide educational opportunities, finding guidance and research information for everything psychology related.

● They also have numerous high-quality databases Peer reviewed journals and books that revolve around mental health.

● The APA is also created with creating their own specific citation and reference style. ● Education, economics, business and social sciences also use APA style quite frequently.

APA style referencing

● Standardized by the American Psychological Association.

● Uses the author/date method of citation in which the author's last name and the year of publication are inserted in the text of the paper.

● Use only the surname of the author followed by a comma and the year of publication. ● References must be listed in an alphabetical order by author, and
then chronologically .

12) Distinguish between cross-sectional and longitudinal research with examples. Cross-sectional research

Cross-sectional studies is a research method which is descriptive and observational in nature, researcher's record information that is present in the population without manipulating variables. This type of study takes place at a given point in time (like a snapshot) without any delay.

The most basic design for a cross-sectional study involves just two variables. These variables may both be scores, may both be nominal categories, or there may be a mixture of nominal and score variables. For example, we could examine the relationship between gender and a diagnosis of depression. In this case, both variables would consist of two binary values – male versus female, diagnosed as depressed versus not diagnosed as depressed.

Examples

● A random sample of schools across Delhi was used to assess the prevalence of asthma among 12-14 years old.

● A study to assess the covid-19 cases in a population suffering from anxiety disorder. ● To measure the cholesterol level among daily walkers and non-walkers.

Longitudinal research

Longitudinal research is a type of correlational research that is observational and involves collecting data over an extended period, often years or even decades. This enables a fuller interpretation of data than is possible with a cross-sectional design. So, for example, it is important to study how human memory changes (or does not change) during the life cycle. Change is so characteristic of humans that researchers would be failing in their obligations if they did not study change over time.

Examples

● To measure the level of cholesterol in a population over an extended period of time. ● A study was conducted to understand the similarities or differences between identical twins that are brought up together versus identical twins that were not. Researchers, in this case, would want to observe the twins from childhood to adulthood to understand the trait habits and personality.

The difference between these studies is the timeline and variable. In a cross-sectional study, researchers observe the same constant variable, and the study is carried out only once. In a longitudinal study, a researcher observes different variables over time and collects data based on those studies.

13) Discuss the significance of formulating hypotheses. What are the different types of hypotheses in educational research ? illustrate

Types of Hypothesis

There are six forms of hypothesis and they are:

● Simple hypothesis
● Complex hypothesis

- Directional hypothesis
- Non-directional hypothesis
- Null hypothesis
- Associative and causal hypothesis

1) Simple Hypothesis :

It shows a relationship between one dependent variable and a single independent variable. For example - If you eat more vegetables, you will lose weight faster. Here, eating more vegetables is an independent variable, while losing weight is the dependent variable.

2) Complex Hypothesis

It shows the relationship between two or more dependent variables and two or more independent variables. Eating more vegetables and fruits leads to weight loss, glowing skin, reduces the risk of many diseases such as heart disease, high blood pressure and some cancers.

3) Directional Hypothesis

It shows how a researcher is intellectual and committed to a particular outcome. The relationship between the variables can also predict its nature. For example- children aged four years eating proper food over a five-year period are having higher 1Q levels than children not having a proper meal. This shows the effect and direction of effect.

4) Non directional Hypothesis

It is used when there is no theory involved. It is a statement that a relationship exists between two variables, without predicting the exact nature (direction) of the relations.

5) Null Hypothesis

It provides a statement which is contrary to the hypothesis. It's a negative statement, and there is no relationship between independent and dependent variables. The symbol is denoted by "HO".

6) Associative and Causal Hypothesis

Associative hypothesis occurs when there is a change in one variable resulting in a change in the other variable. Whereas, the causal hypothesis proposes a cause and effect interaction between two or more variables.

Formulating a Hypothesis

● In many cases, researchers might draw a hypothesis from a specific theory or build on previous research.

● For example, prior research has shown that stress can impact the immune system. ● So a researcher might hypothesize: "People with high-stress levels will be more likely to contract a common cold after being exposed to the virus than people who have low-stress levels."
● "Birds of a feather flock together" is one example of folk wisdom that a psychologist might try to investigate.
● The researcher might pose a specific hypothesis that "People tend to select romantic partners who are similar to them in interests and educational level".
14. Explain the merits and demerits of different measures of dispersion.

There are different measures of dispersion with their relative merits and demerits. But the merits and demerits common to all types of measures of dispersion are outlined as under:

Merits :

● They indicate the dispersal character of a statistical series.
● They speak of the reliability, or dependability of the average value of a series. ● They enable the statisticians to make a comparison between two or more statistical series with regard to the character of their stability or consistency.
● They facilitate in controlling the variability of a phenomenon under his purview. ● They facilitate further statistical analysis of the series through the devices like coefficient of skewness, coefficient of correlation, variance analysis etc.
● They supplement the measures of central tendency in finding out more and more information relating to the nature of a series.

Demerits :

● They are liable to misinterpretations, and wrong generalizations by a statistician of based character.

● They are liable to yield inappropriate results as there are different methods of calculating the dispersions.

● Exceptions or two, of the methods of dispersion involve complicated processes of computation.

● They may give a value of variation, which may not be practically found with the items of the series.

● They, by themselves, cannot give any idea about the symmetricity, or skewed character of a series.

Like the measures of central tendency, most of the measures of dispersion do not give a convincing idea about a series to a layman.

15) Illustrate the applications of the normal probability curve.

Normal curve has wide significance and applications in the field of measurement. Some of the main applications are given below.

1.Use as a model Normal curve represents a model distribution. Hence it may be used as a model

a. To compare various distributions with it, to say, whether the distribution is normal or not, if not in what way it diverges from normal.

b. To compare 2 or more distributions in terms of overlapping

2. To convert raw scores into comparable standard normalized scores. Z score = (where X raw score or actual score, M.5 mean, and o standard deviation. From this, the following interpretations can be made. S.0) and the standard deviati on is 1.0 (6 -1.0) the scores are a

a. If raw score = mean, Z score is 0. 0 6.0 b. If raw score> mean, Z score is positives) -38) C. raw score < mean, Z Score is negative.

3.To compute percentiles and percentile ranks.

4 To understand and apply the concept of standard errors of measurement.

5. For ability grouping .

6. To transform and combine qualitative data.

7.To determine relative difficulty of test items.

8. To determine the percentage of cases in a normal distribution within given limits.

9. To determine the limits in any normal distribution which include a given percentage of the cases.

17) Explain the concept and types of correlation in educational research.

Correlation

Correlation is used to describe how data sets are related to one another. Correlation can be seen when two sets of data are graphed on a scatter plot, which is a graph with an X and Y axis and dots representing the data points.

Types of Correlation

Correlation can be positive, negative, or no correlation. Positive correlation means that as one data set increases, the other data set increases as well. The data in Image 1 has a positive correlation because as years of education increase, so does income. Typically, positively correlated data sets are seen as a line that goes up and to the right on a scatter plot.

Negative correlation means that as one data set increases, the other decreases. Image 2 shows two sets of unknown data with a negative correlation. As the data set on the x axis increases, the data set on the y axis decreases. Typically, negatively correlated data sets are seen as a line that goes down and to the right on a scatter plot.

No correlation means that the two sets of data are not related at all. In other words, this means that one set of data does not increase or decrease with the other.

18) (a) What do you understand about the term 'Research'?

(b) Discuss the criteria, objectives and qualities of a good researcher.

(c) Describe the different methods of non-probability sampling

● Meaning of research

Research is a way of looking for new information, new understanding, and new facts. Research is often used for solving problems or increasing available knowledge. This can be done by testing theories and making observations. Research usually prefers to be systematic, organized, and objective.

One expects scientific research to satisfy the following criteria :

● The purpose of the research should be clearly defined and common concepts be used.

● The research procedure used should be described in sufficient detail to permit another researcher to repeat the research for further advancement, keeping the continuity of what has already been attained.

● The procedural design of the research should be carefully planned to yield results that are as objective as possible.

● The researcher should report with complete frankness, flaws in procedural design and estimate their effects upon the findings.

● The analysis of data should be sufficiently adequate to reveal its significance and the methods of analysis used should be appropriate. The validity and reliability of the data should be checked carefully.

Objectives

1. To find out the real facts

As we know, every type of research has its own objective but the basic aim of the research is always to find out or obtain the information from the markets and societies and their number of respondents. A researcher evaluates or finds the real or exact information for our problem-related questions.

2. To achieve the new thoughts

In this objective of the research, anybody can find new thoughts from the research. Research is the process of finding the exact information through proper observation, optimization, and experiments.

These are the scientific methods to find out or evaluate the information which is very necessary for evaluating the problem task.

3. To evaluate the information

The first aim of the research is to find out the information and then evaluate them in an appropriate or efficient manner so that they can easily design the research problem and solve them also.

A researcher evaluates the information through various scientific approaches and methods, statistical analysis and procedures, and another type of tables and graphs.

4. To test a hypothesis

In this objective of the research, the researcher does the causal relationship between the variables (it can also be said that the hypothesis testing research studies). The hypothesis testing study represents the number of actions like these terms:

(a) Making a formal statement,

(b) Selecting a significance level,

(c) Deciding the distribution use,

(d) Selecting a random sample and computing an appropriate value,

(e) Calculation of the probability,

(f) Comparing the probability.

5. To design or implement the research

After the collection of all information, the researcher prepares the structure of a research design for the company so that they can easily describe or identify the structure of a particular research theme. The research designs can be broadcasted into two forms such as experimental designs and non-experimental designs.

After the structure of the research design, the researcher implements them in a problem and find out the optimum factor to solve them

Qualities of a good research

● The research procedures used should be described in sufficient details to permit another researcher to repeat the research.
● The procedural design of the research should be carefully planned to yield results that are as objective as possible.
● Greater confidence in the research is expected if the researcher is experienced and has a good reputation in research and if he is a person of integrity.
● The purpose of the research or research problem involved should be clearly defined. ● The research should report with frankness, flaws in the procedural design and estimate their effect upon the findings
● Analysis of the data should be sufficiently adequate to reveal it's significant and the methods of analysis used should be appropriate. It should be emphasized that data analysis is one of the most difficult phases of research for the beginners as a result; the validity and reliability of the data should be carefully checked.

c) Four types of non-probability sampling

Convenience sampling:

This method is dependent on the ease of access to subjects such as surveying customers at a mall or passers-by on a busy street. It is usually termed as convenience sampling, because of the researcher's ease of carrying it out and getting in touch with the subjects. Researchers have nearly no authority to select the sample elements, and its purely done based on proximity and not representativeness.

Judgmental or purposive sampling:

Judgemental or purposive samples are formed by the discretion of the researcher. Researchers purely consider the purpose of the study, along with the understanding of the target audience.

Snowball sampling:

Snowball sampling is a sampling method that researchers apply when the subjects are difficult to trace. For example, it will be extremely challenging to survey shelterless people or illegal immigrants. In such cases, using the snowball theory, researchers can track a few categories to interview and derive results.

Quota sampling:

In Quota sampling, the selection of members in this sampling technique happens based on a pre-set standard. In this case, as a sample is formed based on specific attributes, the created sample will have the same qualities found in the total population. It is a rapid method of collecting samples.

19. (a) Describe the ways in which a researcher reaches a problem formulation. (b) illustrate with examples different types of variables in quantitative research. (c) illustrate the need and importance of graphical representation of data in educational research. ● Distinguish the subject area of interest:

The first step in the formulation of the research problem is to decide on a broad subject area on which you have thorough knowledge. Your knowledge in that particular subject area will enable you to decide about the research problem. It will also help you to carry out the overall research. The subject area can be selected with the help of your adviser if you are a student. When you are

doing independent research you can consult some senior researchers in your particular field. Distinguishing or selecting the subject area is the first step in the research problem formulation.

● Dissect the subject area into sub-areas:

The subject area of your interest will be broad and you need to dissect it into small areas. In this way you will be able to select one of your interests and convenience. You can get help from an experienced person in this regard.

● Decide about an area:

Select an area among all the sub-areas, one that interests you most. This will help you in creating the research problems among which you have to choose one. During all these steps you need to constantly look at the possibilities of further narrowing down the subject area in order to become more specific.

● Generate research questions:

Generate as many research questions as possible, from these questions you have to choose those questions that you want to answer through your research. You should take as much time as possible to generate many questions so that you have plenty of choices. Research questions should be such that they can be answered using scientific techniques and research procedures. Suppose does God exist is a question that you want to be answered but this question cannot be answered using current scientific techniques therefore such questions should be avoided.

● Decide about the objectives:

Objectives are the possible answers to the research question or the research problem that you have formulated. They should be formulated in a clear manner. Objectives make you specific, as you conduct your study around the objectives that you have decided. The objectives need to be specific in nature but you can also generate general objectives. General objectives and specific objectives both will have their own importance in the research.

b) Quantitative variables

Quantitative variables, also called numeric variables, are those variables that are measured in terms of numbers. A simple example of a quantitative variable is a person's age. The age can take on different values because a person can be 20 years old, 35 years old, and so on. Likewise, family size is a quantitative
Variable, because a family might be one, two, three members and so on.

Discrete and continuous variables

Quantitative variable are again of two types:

Discrete and continuous.
Discrete variable:a discrete variable, restricted to certain values, usually(but not necessarily)consists of whole numbers, such as the family size, number of defective items in a box. They are often the results of enumeration or counting.
eg:the number of accidents in the twelve months. The number of mobile cards sold in a store within seven days.

Continuous variable:A continuous variable is one that may take on an infinite number of intermediate values along a specified interval.
eg: the sugar level in the human body. Blood pressure reading Temperature
Height or weight of the human body.

c) The primary advantages of graphical representation of data are:

● Facilitates and improves learning: graphics make data easy to understand and eliminate language and literacy barriers
● Understanding content: visuals are more effective than text in human understanding ● Flexibility of use: graphical representation can be leveraged in nearly every field involving data

● Increases structured thinking: users can make quick, data-driven decisions at a glance with visual aids. Supports creative, personalized reports for more engaging and stimulating visual presentations.

● Improves communication: analysing graphs that highlight relevant themes is significantly faster than reading through a descriptive report line by line.

20. (a) Define Kurtosis. What are different types of curves based on Kurtosis values?

(b) What do we mean by probability sampling techniques ? Describe how you would draw a (i) Stratified random sample; and (ii) Cluster sample ?

(c) An achievement test was administered on 2500 students. The mean is 85 and standard deviation is 4.25. Find the 67th percentile and 39th percentile and then write interpretation.

a) Kurtosis refers to the
divergence in the height
of the curve, specially in
the peakness.

They are of 3 types
1. Platy kurtic – flatter
peak at the center
than a normal curve.
2. Lepto kurtic – more
peaked at the center
than a normal curve.
3. Meso kurtic – almost
resembles a normal
curve.

b) Probability sampling refers to the selection of a sample from a population, when this selection is based on the principle of randomization, that is, random selection or chance. Probability sampling is more complex, more time-consuming and usually more costly than non-probability sampling.

How Stratified Random Sampling Works

When completing analysis or research on a group of entities with similar characteristics, a researcher may find that the population size is too large to complete research on it. To save time and money, an analyst may take on a more feasible approach by selecting a small group from the population. The small group is referred to as a sample size, which is a subset of the population used to represent the entire population. A sample may be selected from a population through a number of ways, one of which is the stratified random sampling method.

Stratified random sampling involves dividing the entire population into homogeneous groups called strata (plural for stratum). Random samples are then selected from each stratum.

Cluster sampling involves dividing a large target group into multiple smaller groups or clusters for research purposes. Researchers then form a sample by randomly selecting these groups.

The random selection gives every group in that target population an equal chance to be a part of the sample group. Though only a few relevant groups are selected and the rest are eliminated.

How to Conduct Cluster Sampling in 4 Simple Steps

Here's how to conduct single-stage cluster sampling and find the correct representative sample:

Step 1: Define Your Audience

Decide on your target population and desired sample size.

Step 2: Create Clusters or Subgroups

Now divide your target population into smaller subgroups or clusters based on a specific criteria. Getting this step right is crucial as it affects the quality of your segment or cluster and how well it represents your target population.

Step 3: Randomly Select Your Clusters

Pick a cluster or group that closely resembles the audience that you're looking to research. You can pick a cluster based on a method of random selection. Make sure to keep the sample size in mind while you select one.

Step 4: Collect Data from the Sample

Finally, conduct your research and collect data from your selected clusters.

C) percentile value = $\mu + z\sigma$

where:

μ: Mean

z: z-score from z table that corresponds to percentile value

σ: Standard deviation

67 th

Z = 0.2486 (value of 0.67)

Percentile value = 85+(0.2486*4.25)

= 86.0565

39 th

Z= 0.1517(value of 0.39)

Percentile value = 85+(0.1517*4.25)

= 85.6447

21. (a) Elucidate the meaning and characteristics of standard deviation.

(b) Explain the use of derived z-score and stanine scores with examples.

(c) Compute the standard deviation for the following data and write the interpretation :

Marks	No. of students
90 - 99	10
80 - 89	17
70 - 79	23
60 - 69	40
50 - 59	45
40 - 49	25
30 - 39	20
20 -29	13
10-19	7

a) Standard deviation is the measure of dispersion of a set of data from its mean. It measures the absolute variability of a distribution; the higher the dispersion or variability, the greater is the standard deviation and greater will be the magnitude of the deviation of the value from their mean.

Characteristics

1. It cannot be negative.
2. It is only used to measure spread or dispersion around the mean of a data set.
3. It shows how much variation or dispersion exists from the average value.
4. It is sensitive to outliers. A single outlier can raise σ and, in turn, distort the picture of the spread.

5. For data with almost the same mean, the greater the spread, the greater the standard deviation.

6. Standard deviation can be used in conjunction with the mean in order to calculate data intervals when analysing normally distributed data.

b) In statistics, a z-score tells us how many standard deviations away a given value lies from a population mean.

We use the following formula to calculate a z-score for a given value:

$z = (x – \mu) / \sigma$

where:

x: Individual data value

μ: Mean of population

σ: Standard deviation of population

The following examples show how z-scores are used in real life in different scenarios.

Example 1: Exam Scores

Z-scores are often used in academic settings to analyze how well a student's score compares to the mean score on a given exam.

A stanine ("standard nine") score is a way to scale scores on a nine-point scale. It can be used to convert any test score to a single-digit score. Like z-scores and t-scores, stanines are a way to assign a number to a member of a group, relative to all members in that group. However, while z-scores and t-scores can be expressed with decimals like 1.2 or 3.25, stanines are always positive whole numbers from 0 to 9.

Stanines are also similar to normal distributions. You can think of these scores as a bell curve that has been sliced up into 9 pieces. These pieces are numbered 1 through 9, starting at the left hand section. However, where a standard normal distribution has a mean of 0 and a standard deviation of 1, stanines have a mean of 5 and a standard deviation of 2.

Example:

We are given the heights of 1000 people living in a particular town. The average height of the people is 168cm with a standard deviation of 3cm. If John has a height of 163cm, express it in the form of a stanine score.

Solution:

Step 1: We note down the mean and standard deviation given as,

$\mu = 168$ cm and $\sigma = 3$cm

Step 2: Given X = 163cm we convert it to Z score using the formula,

$Z=(X-\mu)/\sigma$

$Z = (163-168)/3 = -1.67$

Step 3: Convert the Z score to the stanine scale by looking at the above table.

Hence, Stanine score = 2 which means that John has a much smaller height compared to the average.

c)

Marks	No of students (f)	Mid point fx Deviation (x) from mean (d) d^2	fd^2
90-99	10	94.5 945 1425.06	14250.6
80-89	17	84.5 1436.5 770.06	13091.02
70-79	23	74.5 1713.5 315.06	7246.38
60-69	40	64.5 2580 60.06	2402.4
50-59	45	54.5 2452.5 5.06	227.7
40-49	25	44.5 1112.5 150.06	3751.5
30-39	20	34.5 690 495.06	9901.2
20-29	13	24.5 318.5 1040.06	13520.78
10-19	7	14.5 101.5 1785.06	12495.42
	N= 200	=11350	= 76,887

Mean = Σ fx/N

= 11350÷ 200

= 56.75

Std deviation

Std deviation = √76887÷200

= 19.6070

MED 04 – INTRODUCTION TO EDUCATIONAL RESEARCH AND STATISTICS

Prepared by: Nidhin Raj A

Batch: 2022-2024

Year of Question Paper: 2018

Part 1

1.Explain the meaning of Descriptive Research.

Ans: Descriptive research studies are designed to obtain a pertinent and precise information concerning the current status of phenomena and whenever possible to draw valid general conclusions from the facts discovered. Descriptive studies involve measurement, collection, classification, analysis, comparison and interpretation. They collect and provide three types of information.

1.What exists with respect to variables of conditions in a situation.

2.What we want by identifying standards of norms desirable.

3.How to achieve goals by exploring possible ways and means on the basis of experience of others or the opinions of experts.

II.Explain the steps in educational research

Ans: Educational research is the application of scientific method to the study of educational problems. The steps in the educational research, therefore are more or less identical, but more structured, systematic and intensive to those of scientific method.

1. Selection of the research problem
2. Formulation of hypothesis
3. The method to be used
4. Data collection
5. Analysis and Interpretation of the data
6. Reporting the result

III. Explain the meaning of purposive sampling.

Ans: In this sampling method, the researcher selects a typical groups of individuals who might represent the larger population and then collects data from this group. For example, if a researcher wants to survey the attitude towards the teaching profession of teachers teaching students from lower socio-economic stratum, he or she might survey the teachers teaching in schools catering to students from slums (more specifically, teachers teaching in Municipal Schools) with the assumption that since all teachers teaching in Municipal Schools cater to students from lower socioeconomic stratum, they are representative of all the teachers teaching students from lower socio-economic stratum.

IV. Explain the Meaning of variables in research.

Ans: All research projects are based around variables. "A variable is the characteristic or attribute of an individual, group, educational system or the environment that is of interest in a research study."

V. Explain the use of statistics in research

Ans: Statistical methods are essential for scientific research. In fact, statistical methods dominate the scientific research as they include planning, designing, collecting data, analysing, drawing meaningful interpretation and reporting of research findings. Furthermore, the results acquired from research project are meaningless raw data unless analysed with statistical tools.

Part 2

VI. Differentiate evaluation study from cross sectional study.

Ans: Cross-sectional study design is a type of observational study design. As discussed in the earlier articles, we have highlighted that in an observational study, the investigator does not alter the exposure status. The investigator measures the outcome and the exposure(s) in the population, and may study their association. n a cross-sectional study, the investigator measures the outcome and the exposures in the study participants at the same time. Unlike in case–control studies (participants selected based on the outcome status) or cohort studies (participants selected based on the exposure status), the participants in a cross-sectional study are just selected based on the inclusion and exclusion criteria set for the study. Once the participants have been selected for the study, the investigator follows the study to assess the exposure and the outcomes. After the entry into the study, the participants are measured for outcome and exposure at the same time [Figure 1]. The investigator can study the association between these variables. It is also possible that the investigator will recruit the study participants and examine the outcomes in this population. The investigator may also estimate the prevalence of the outcome in those surveyed.

Evaluation research, also known as program evaluation, refers to research purpose instead of a specific method. Evaluation research is the systematic assessment of the worth or merit of time, money, effort and resources spent in order to achieve a goal.

Evaluation research is closely related to but slightly different from more conventional social research. It uses many of the same methods used in traditional social research, but because it takes place within an organizational context, it requires team skills, interpersonal skills, management skills, political smartness, and other skills that social research does not need much. Evaluation research also requires one to keep in mind the interests of the stakeholders.

Evaluation research is a type of applied research, and so it is intended to have some real-world effect. Many methods like surveys and experiments can be used to do evaluation research. The process of evaluation research consisting of data analysis and reporting is a rigorous, systematic process that involves collecting data about organizations, processes, projects, services, and/or resources. Evaluation research enhances knowledge and decision-making, and leads to practical applications.

VII. Explain the role of dependent and independent variable.

A dependent variable is an attribute or characteristic that is dependent on or influenced by the independent variable. If one variable depends upon or is a consequence of the other variable it is termed as a dependent variable. Researchers typically investigate multiple dependent variables in a single study (eg. You hit another car, the other car hit the car in front of it and so forth), although in many studies, one of the dependent variables is typically of central interest. Dependent variables can be measured using continuous or categorical scores. Examples of dependent variables in education are achievement scores on a test, organizational climate of a junior high school, leadership skills of principals, or cost effectiveness of student affairs programmes in colleges. To locate dependent variables in a study, examine purpose statements, research questions, and hypotheses for outcomes that the researcher wishes to predict or explain. Ask yourself, "What is the outcome in this study?" An independent variable is an attribute or characteristic that influences or affects an outcome or dependent variable. Variable that is antecedent to the dependent variable is termed as an independent variable. Eg: Height depends upon age. So height is a dependent variable and age is an independent variable. So is sex, so age and sex are independent variables. There are four types of independent variables, and each serves a slightly different purpose. ▪ Measures variable ▪ Control variable ▪ Treatment variable ▪ Moderating variable An independent variable is that which you have control over and what you can choose and manipulate. In some cases it may not be able to manipulate the independent variable. It may not be able to manipulate the independent variable. It may be something that is already there and is fixed, something you would like to evaluate with respect to how it affects something else, the dependent variable like color, king and time etc.

VIII. **Discuss the need of classification of quantitative data.**

Quantitative data is the value of data in the form of counts or numbers where each data set has a unique numerical value. This data is any quantifiable information that researchers can use for mathematical calculations and statistical analysis to make real-life decisions based on these mathematical derivations. For example, there are quantities corresponding to various parameters. For instance, “How much did that laptop cost?” is a question that will collect quantitative data. Values are associated with most measuring parameters such as pounds or kilograms for weight, dollars for cost, etc. It makes measuring various parameters controllable due to the ease of mathematical derivations they come with. It is usually collected for statistical analysis using surveys, polls, or questionnaires sent across to a specific section of a population. Researches can establish the retrieved results across a population.

The most common types of quantitative data are as below:

Count equated with entities. For example, the number of people downloading a particular application from the App Store.

Measurement of physical objects: Calculating measurement of any physical thing. For example, the HR executive carefully measures the size of each cubicle assigned to the newly joined employees.

Sensory calculation: Mechanism to naturally “sense” the measured parameters to create a constant source of information. For example, a digital camera converts electromagnetic information to a string of numerical data.

Projection of data: Future data projection can be made using algorithms and other mathematical analysis tools. For example, a marketer will predict an increase in sales after launching a new product with a thorough analysis.

Quantification of qualitative entities: Identify numbers to qualitative information. For example, asking respondents of an online survey to share the likelihood of recommendation on a scale of 0-10.

IX. Explain the meaning of multi stage sampling.

Multistage sampling is defined as a sampling method that divides the population into groups (or clusters) for conducting research. It is a complex form of cluster sampling, sometimes, also

known as multistage cluster sampling. During this sampling method, significant clusters of the selected people are split into sub-groups at various stages to make it simpler for primary data collection.

There are four multistage steps to conduct multistage sampling:

Step one: Choose a sampling frame, considering the population of interest. The researcher allocates a number to every group and selects a small sample of relevant separate groups.

Step two: Select a sampling frame of relevant separate sub-groups. Do this from related, different discrete groups selected in the previous stage.

Step three: Repeat the second step if necessary.

Step four: Using some variation of probability sampling, choose the members of the sample group from the sub-groups.

Multistage sampling is applied mainly to these areas:

Applied to a multistage design where the population is too vast and researching every individual is impossible.

To gather student perceptions from students belonging to various colleges, studying different courses and located throughout the country.

To survey employees of a multinational company belonging to multiple locations in multiple countries.

Government bureaus use this method all the time to draw inferences from the population.

Multiphase sampling reduces the time taken to research an area. It also keeps a tab on the cost of the research. The information collected from the samples is used to draw inferences from the population as a whole.

Here are the top 8 benefits obtained from multistage sampling:

- It allows researchers to apply cluster or random sampling after determining the groups.
- Researchers can apply multistage sampling to make clusters and sub-clusters until the researcher reaches the desired size or type of group.

- Researchers can divide the population into groups without restrictions. It allows flexibility to the researchers to choose the sample carefully.
- It is useful while collecting primary data from a geographically dispersed population.
- Cost-effective and time-effective because this method helps cut down the population into smaller groups.
- Finding the right survey sample becomes very convenient for researchers.
- The researcher mindfully chooses the audience. It decreases the issues faced during random sampling.
- It does not need a complete list of all the members of the target population, dramatically reducing sample preparation cost.

X. Explain the use of primary sources and secondary sources of review of related research.

Ans: The first step in reviewing the related literature is identifying the material that is to be read and evaluated.

The identification can be made through the use of primary and secondary sources available in the library.

Primary Sources

In the primary sources of information, the author reports his own work directly in the form of research articles, books, monographs, dissertations or theses. Such sources provide more information about a study that can be found elsewhere.

Primary sources give the researcher a basis on which to make his own judgement of the study. Though consulting such sources is a time consuming process for a researcher, yet they provide a good source of information on the research methods used in Secondary sources

In secondary sources, the author complies and summarises the findings of the work done by the other and gives interpretation of these findings. In them, the author usually attempts to cover all of the important studies in an area reported in

encyclopedia of education, education indices, abstracts, bibliographies, bibliographical references and quotation sources. Working with secondary sources is not time-consuming because of the amount of reading required. The disadvantage of the secondary sources, however, is that the reader is depending upon someone else's judgements about the important and significant aspects of the study.

When to use primary and secondary sources

The decision concerning the use of primary or secondary sources depends largely on the nature of the research study proposed by the researcher.

If it is a study in an area in which much research has been reported, a review of the primary sources would be a logical first step. If the study is in an area in which little or no research has been conducted, a check of the secondary sources is more logical.

Sources of information, whether primary or secondary, are found in a library. The researcher must, therefore, develop the expertise to use resources without much loss of time and energy.

XI. Discuss the steps in Construction of pie diagram.

Ans: Steps of construction of pie chart for a given data:

● Find the central angle for each component using the formula given on the previous page.

● Draw a circle of any radius.

● Draw a horizontal radius

● Starting with the horizontal radius, draw radii, making central angles corresponding to the values of respective components.

● Repeat the process for all the components of the given data.

● These radii divide the whole circle into various sectors.

● Now, shade the sectors with different colors to denote various components.

● Thus, we obtain the required pie chart.

XII. Explain the uses of measures of Histogram.

Histogram in Statistics is a very important topic. Let's understand it with an example. A container shows the number of information that focuses inside a reach (a stretch). Regularly, you pick the reach that best accommodates your information. A histogram chart is a bar chart representation of data. It is a representation of a series of results into columns along the x-axis. in the same histogram, the number of numbers or more occurrences in the data for each column is represented by the y-axis. It is the simplest method that can be used to visualize the distribution of data.

Given below are the main part of the Histogram.

- **The Title:** The title is the first and the foremost thing it describes all the information which is given in the Histogram.
- **X-axis:** The intervals under which the measurement falls is shown in the X-axis intervals.
- **Y-axis:** The values that occurred within the intervals set by the X-axis, is shown in the Y-axis.
- **The Bars:** The bars are used for showing the value of the data. And for knowing the total number of times the values occurred within the interval, the height of the bar is helpful, while the interval that is covered is shown by the width of the bar. And hence, it is obvious that the Histogram which has all the bins equal must have the width same as well, across all the bars.

How Histogram Works

In statistics Histograms, for the most part, are used widely because it shows how many of a specific type of variable occurs within a certain range. That is to say, it helps in showing the data or the numbers in graphical format and hence makes it much easier for us to understand and interpret the data.

Histograms can work and serve so many different purposes, such as from the cenus, Histogram can be used for showing the range of people between eh certain age, such as how many people are there in the country between the age of 10 and 20 etc. In many various operations, Histograms are useful.

Also, if you wish to know about the Bar Graph and Histogram, because more often than not both are confused with each other, you may find this link helpful Bar graph– Definition, Types, Uses and Key Difference (vedantu.com)

Characteristics of a Histogram

- A Histogram is used to display continuous data in a categorical form.
- In a Histogram, there are no gaps between the bars, unlike a bar graph.
- The width of the bins is equal.
-
-

It is the Area, Not the Height of the Bars

In a Histogram, it is the area and not the height of the bar that indicates the frequency of occurrences for each bin. The height of the bar does not indicate how many occurrences of scores are there in each individual bin. It is always the product of the height and width of the bin that indicates the frequency of occurrences within that bin.

How to Create a Frequency Histogram Graph

To construct a Histogram graph from a continuous variable there are a few steps that we need to follow. They are given below;

- **Step 1)** Firstly, we need to split the data into class intervals which are also known as bins and frequencies.
- **Step 2)** In this step, we have to draw a Histogram graph with X-axis and Y-axis. Then write down the class intervals on the X-axis and the frequencies on the Y-axis.

- **Step 3)** Draw vertical rectangles using the X-axis and the Y-axis.

A histogram Graphical representation points organized into user-specified ranges. Similar in appearance to a bar graph, the histogram condenses a data series into an easily interpreted visual by taking many data points and grouping them into logical ranges or bins. It is frequently used to easily depict the main characteristics of the data distribution.

XIII. Explain the use of Median and quartile deviation.

Ans: Quartile deviation is a statistic that measures the deviation. It measures the deviation of the data from the average value. Here quartile deviation gives the spread of the data, which helps to understand the distribution of the data. Before understanding more about quartile deviation let us understand more about quartiles. Here we have three quartiles Q_1, Q_2, Q_3 which divide the data into three quarters. The median of the data has been referred as the second quartile Q_2. Also, the first quartile Q_1 is the median of the first half of the data, and the third quartile Q_3 is the median of the second half of the data.

Quartile deviation is the dispersion in the middle of the data. The difference between the first quartile Q_1 and the third quartile Q_3 is called the interquartile range, and half of this interquartile range is called the quartile deviation. This quartile deviation is also referred to as a semi-interquartile range.

Quartile Deviation = (Third Quartile – First Quartile) / 2

Quartile Deviation =$(Q_3 – Q_1) / 2$

Quartile deviation can be calculated for both the grouped data and the ungrouped data. Quartile deviation measures the absolute level of dispersion and is not affected by the extreme values. And the relative measure with reference to quartile deviation is known as the coefficient of quartile deviation.

Coefficient of Quartile Deviation = $(Q_3 – Q_1) / (Q_3 + Q_1)$

The quartile deviation can be calculated in two different methods, based on the type of given data. The quartile deviation is calculated differently for ungrouped data and for the grouped data. The quartile deviation is

- Arrange the available data in ascending or both the grouped and ungrouped data.
- Find the first quartile value using one of these formulas. For ungrouped data use the formula $Q_1 = (n + 1)/4$, and for ungrouped data use the formula $Q1=l1+(N/4)−cf(l2−l1)Q1=l1+(N/4)−cf(l2−l1)$. Here n is for the particular quartile, N is the total frequency, f is the frequency of the particular class, c is the cumulative frequency of the preceding class, and l_1, l_2 are the lower and upper boundaries of the class interval.

In statistics, the median is the value that splits an ordered list of data values in half. Half the values are below it and half are above—it's right in the middle of the dataset. The median is the same as the second quartile or the 50th percentile. It is one of several measures of central tendency.

Uses of Median:

1. Median is used when the exact midpoint of the distribution is needed or the 50% point is wanted.
2. When extreme scores affect the mean at that time median is the best measure of central tendency.
3. Median is used when it is required that certain scores should affect the central tendency, but all that is known about them is that they are above or below the median.
4. Median is used when the classes are open ended or it is of un equal cell size.

Merits of Median:

1. It is easy to compute and understand.
2. All the observations are not required for its computation.
3. Extreme scores does not affect the median.
4. It can be determined from open ended series.
5. It can be determined from un-equal class intervals.

XIV. Explain the method of correlation.

The **Correlation** is a statistical tool used to measure the relationship between two or more variables, i.e. the degree to which the variables are associated with each other, such that the change in one is accompanied by the change in another.

The **Scatter Diagram Method** is the simplest method to study the correlation between two variables wherein the values for each pair of a variable is plotted on a graph in the form of dots thereby obtaining as many points as the number of observations. Then by looking at the scatter of several points, the degree of correlation is ascertained.

The degree to which the variables are related to each other depends on the manner in which the points are scattered over the chart. The more the points plotted are scattered over the chart, the lesser is the degree of correlation between the variables. The more the points plotted are closer to the line, the higher is the degree of correlation. The degree of correlation is denoted by **"r".**

Karl Pearson's Coefficient of Correlation is widely used mathematical method wherein the numerical expression is used to calculate the degree and direction of the relationship between linear related variables. Pearson's method, popularly known as a **Pearsonian Coefficient of Correlation,** is the most extensively used quantitative methods in practice. The coefficient of correlation is denoted by **"r".**

The **Spearman's Rank Correlation Coefficient** is the non-parametric statistical measure used to study the strength of association between the two ranked variables. This method is applied to the ordinal set of numbers, which can be arranged in order, i.e. one after the other so that ranks can be given to each.

XV. Explain the use of biserial correlation.

Ans: The biserial correlation is a correlation between on one hand, one or more quantitative variables, and on the other hand one or more binary variables. It was introduced by Pearson (1909). The biserial correlation can be calculated with XLSTAT.

As for the Pearson correlation, the biserial correlation coefficient varies between -1 and 1. 0 corresponds to no association (the means of the quantitative variable for the two categories of the qualitative variable are identical).

Biserial correlation in XLSTAT

XLSTAT allows testing if the value of the biserial correlation r that has been obtained is different from 0 or not.

For the two-tailed test, the null H0 and alternative Ha hypotheses are as follows:

$H0 : r = 0$

$Ha : r \neq 0$

In the left one-tailed test, the following hypotheses are used:

$H0 : r = 0$

$Ha : r < 0$

In the right one-tailed test, the following hypotheses are used:

$H0 : r = 0$

$Ha : r > 0$

Two methods to compute the p-value are proposed by XLSTAT. The user can choose between a p-value computed using on a large sample approximation, and a p-value computed using Monte Carlo re samplings. The second method is recommended.

XVI. Discuss the important Features of any one computer software which is used in research.

Ans: Stata is one of statistical software packages, like SAS, SPSS, MINITAB , or BMDP. Similar tasks across this software: • Data Management • Data Analysis • Ability to use graphs to present analysis results Differences among this software: • User interface • Data format • Efficiency in managing/analysing data and presenting results • Syntax rules • Some statistical analysis may be available in one package.

The main features of Stata are: Stata is cross-platform compatible. Stata software help you transfer data between Stata and other statistical software. You can easily learn how to use Stata even if you do not know the syntax. Stata is easily extensible. Some special statistical analyses were not available in Stata, e.g. structural equation modelling or item response analysis.

Stata is a power software and very easy to use. • Use the interactive mode to learn about Stata, and the use text mode for doing research • Your ability as a researcher is the main determinant of the quality of your research.

XVII. Explain the use of Normal probability curve.

Ans: The normal distribution is by far the most used distribution in statistics. Some important reasons for the same are

1. The normal distribution appears to be a reasonable model of the behaviour of many of the random phenomena.
2. It may be convenient on mathematical grounds alone to assume a normally distributed population
3. The normal distribution can be used as a good approximation to a number of theoretical distributions like the binomial, Poisson, etc.
4. There is a very intimate connection between size of sample and the extent to which a sampling distribution approaches the normal form.
5. Biological Statistics – proportion of male to female births for the same country or community over a period of years, proportion of different types of plants and animals in cross fertilization (mendelian ratio).
6. Anthropometrical data – height, weight, etc., for large groups of the same age and sex.
7. Psychological measurements – intelligence as measured by standard tests, speed of association, perception spun, reaction time, etc.
8. Social and economic data – birth rate, death rate, wage and output of large number of workers, etc.
9. Errors of observation – measures of height, speed of movement, linear magnitudes, physical and mental traits, certain errors which likely cause to deviate above or below their true values.
10. Even if the variable is not normally distributed, it can sometimes be brought to normal form by simple transformation of variable.
11. The entire theory of small sample tests, viz, t, F, test etc. is based on the fundamental assumption that the parent population from which the samples have been drawn follow normal distribution.
12. Normal distribution is invariably applied in statistical quality control for setting control limits.

Part- 3

XVIII. Compute the mean, median, standard deviation, 67^{th} percentile and skewness for the following test score of 100 students and then write the interpretation based on test scores.

Marks : 90 80 70 60 50 40 30 20 10

No Of Students : 8 6 4 12 30 20 10 8 2

Ans: Computer the Mean, median and Standard deviation,67th percentiles coefficient of skewness for the following test scores of 100 students and then write the interpretation based on the scores

Mark : 90 80 70 60 50 40 30 20 10

No. of students: 8 6 4 12 30 20 10 8 2

MARKS(x)	**NO OF STUDENTS(f)**	**fx**	**CUMULATIVE FREQUENCY**	**x-mean**	**(x-mean)^2**	**F(x-mean)^2**	
10	2	20	2	-39.8	1584.04	3168.08	
20	8	160	10	-29.8	888.04	7104.32	
30	10	300	20	-19.8	392.04	3920.4	
40	20	800	40	-9.8	96.04	1920.8	
50	30	1500	70	.2	.04	1.2	
60	12	720	82	10.2	104.04	1248.48	
70	4	280	86	20.2	408.04	1632.16	
80	6	480	92	30.2	912.04	5472.24	
90	8	720	100	40.2	1616.04	12928.32	
	100	4980				37396	

MEAN=fx/∑f =4980/100= 49.8

MEDIAN

40+50/2 =45

STANDARD DEVIATION=√f(x-mean)^2/∑f=√37396/100=√373.96=19.33.

XIX. a) Explain the different type of Hypothesis with the help of illustration.

b) Explain the use of survey on research in education.

Ans: a) A hypothesis is an assumption that is made based on some evidence. This is the initial point of any investigation that translates the research questions into predictions. It includes

components like variables, population and the relation between the variables. A research hypothesis is a hypothesis that is used to test the relationship between two or more variables.

Simple Hypothesis

It shows a relationship between one dependent variable and a single independent variable. For example – If you eat more vegetables, you will lose weight faster. Here, eating more vegetables is an independent variable, while losing weight is the dependent variable.

Complex Hypothesis

It shows the relationship between two or more dependent variables and two or more independent variables. Eating more vegetables and fruits leads to weight loss, glowing skin, and reduces the risk of many diseases such as heart disease.

Directional Hypothesis

It shows how a researcher is intellectual and committed to a particular outcome. The relationship between the variables can also predict its nature. For example- children aged four years eating proper food over a five-year period are having higher IQ levels than children not having a proper meal. This shows the effect and direction of the effect.

Non-directional Hypothesis

It is used when there is no theory involved. It is a statement that a relationship exists between two variables, without predicting the exact nature (direction) of the relationship.

Null Hypothesis

It provides a statement which is contrary to the hypothesis. It's a negative statement, and there is no relationship between independent and dependent variables. The symbol is denoted by "H_O".

Associative and Causal Hypothesis

Associative hypothesis occurs when there is a change in one variable resulting in a change in the other variable. Whereas, the causal hypothesis proposes a cause and effect interaction between two or more variables.

b)

- Measures of central tendency tell us what is common or typical about our variable.
- Three measures of central tendency are the mode, the median and the mean.

- The mode is used almost exclusively with nominal-level data, as it is the only measure of central tendency available for such variables. The median is used with ordinal-level data or when an interval/ratio-level variable is skewed (think of the Bill Gates example). The mean can only be used with interval/ratio level data.
- Measures of variability are numbers that describe how much variation or diversity there is in a distribution.
- Four measures of variability are the range (the difference between the larges and smallest observations), the interquartile range (the difference between the 75th and 25th percentiles) the variance and the standard deviation.
- The variance and standard deviation are two closely related measures of variability for interval/ratio-level variables that increase or decrease depending on how closely the observations are clustered around the mean.

XX. a) Explain the need and importance of longitudinal and studies in research with the help of illustration.

b) Write the use of survey on research in education.

Ans) A longitudinal study is a correlation method that helps discover the relationship between variables in a specific target population. It is pretty similar to a cross_ sectional study, although in its case, the researcher observes the variables for a longer time, sometimes lasting many years.

For example, let's say you are researching social interactions among wild cats. You go ahead to recruit a set of newly-born lion cubs and study how they relate with each other as they grow. Periodically, you collect the same type of data from the group to track their development.

The advantage of this extended observation is that the researcher can witness the sequence of events leading to the changes in the traits of both the target population and the different groups. It can identify the causal factors for these changes and their long-term impact.

Advantages of a Longitudinal Study

1. Longitudinal studies help you discover variable patterns over time, leading to more precise causal relationships and research outcomes.
2. When researching developmental trends, longitudinal studies allow you to discover changes across lifespans and arrive at valid research outcomes.
3. They are highly flexible, which means the researcher can adjust the study's focus while it is ongoing.
4. Unlike other research methods, longitudinal studies collect unique, long-term data and highlight relationships that cannot be discovered in a short-term investigation.
5. You can collect additional data to study unexpected findings at any point in your systematic investigation.

Disadvantages and Limitations of a Longitudinal Study

1. It's difficult to predict the results of longitudinal studies because of the extended time frame. Also, it may take several years before the data begins to produce observable patterns or relationships that can be monitored.
2. It costs lots of money to sustain a research effort for years. You'll keep incurring costs every year compared to other forms of research that can be completed in a smaller fraction of the time.
3. Longitudinal studies require a large sample size which might be challenging to achieve. Without this, the entire investigation will have little or no impact.
4. Longitudinal studies often experience panel attrition. This happens when some members of the research sample are unable to complete the study due to several reasons like changes in contact details, refusal, incapacity, and even death.

Longitudinal Studies Examples

How does a longitudinal study work in the real world? To answer this, let's consider a few typical scenarios.

Example 1

A researcher wants to know the effects of a low-carb diet on weight loss. So, he gathers a group of obese men and kicks off the systematic investigation using his preferred longitudinal study

method. He records information like how much they weigh, the number of carbs in their diet, and the like at different points. All these data help him to arrive at valid research outcomes.

b)

The uses of Survey Research

Cost

Surveys are relatively inexpensive. Online surveys and mobile surveys, in particular, have a very small cost per respondent. Even if incentives are given to respondents, the cost per response is often far less than the cost of administering a paper survey or phone survey, and the number of potential responses can be in the thousands.

Extensive

Surveys are useful in describing the characteristics of a large population. No other research method can provide this broad capability, which ensures a more accurate sample to gather targeted results in which to draw conclusions and make important decisions.

Flexible

Surveys can be administered in many modes, including: online surveys, email surveys, social media surveys, paper surveys, mobile surveys, telephone surveys, and face-to-face interview surveys. For remote or hard-to-reach respondents, using a mixed mode of survey research may be necessary (e.g. administer both online surveys and paper surveys to collect responses and compile survey results into one data set, ready for analysis).

Dependable

The anonymity of surveys allows respondents to answer with more candid and valid answers. To get the most accurate data, you need respondents to be as open and honest as possible with their answers. Surveys conducted anonymously provide an avenue for more honest and unambiguous responses than other types of research methodologies, especially if it is clearly stated that survey answers will remain completely confidential.

FIRST SEMESTER M.Ed. DEGREE EXAMINATION, NOVEMBER 2019

M.Ed. 04 **Prepared by: Ranya.kp (2022-24)**

INTRODUCTION TO EDUCATIONAL, RESEARCH AND STATISTICS

Part A (Answer all question. Each question carries 2 marks.)

1. **Mention the two paradigms of research.**

Ans: Conceptual framework for seeing and making sense of the social world.

1.Quantitative paradigm: Quantitative research is termed the traditional, Quantitative research is termed the traditional, positivists, the experimental or empirical tradition by some of the authorities like Comte, Mill, Durkheim, Newton, etc. Quantitative approach uses 'measurement' as the most precise and universally accepted method for assigning quantitative values. Measurement is defined as the assignment of numbers to objects and events according to logically accepted rules.

2.Qualitative paradigm: Qualitative approach is used in phenomenological inquiry for the verification of proposition, which takes into consideration the totality of phenomenon and does not attempt at analysing it into quantifiable components. Qualitative research in contrast to quantitative approach is by some regarded as less 'scientific' and 'softer.' Detailed perspectives of the participants in the 'naturalistic' settings.

2.What are the type of sampling?

Ans: Sampling in market research is of two types – probability sampling and non-probability sampling.

Probability sampling: Probability sampling is a sampling technique where a researcher sets a selection of a few criteria and chooses members of a population randomly. All the members have an equal opportunity to be a part of the sample with this selection parameter.

Non-probability sampling: In non-probability sampling, the researcher chooses members for research at random. This sampling method is not a fixed or predefined selection process. This makes it difficult for all elements of a population to have equal opportunities to be included in a sample.

3.what is variable?

Ans: A variable is, as the name applies, something that varies.

Age, sex, export, income, and expenses, family size, class grades, blood pressure readings, eye colour, and vehicle type are all examples of variables because each of these properties varies or differs from one individual to another.

Variable definition in research

A variable is any property, a characteristic, a number, or a quantity that increases or decreases over time or can take on different values in different situations.

4.write any two applications of statistics?

Ans: The important areas of application of statistics are: The State. Economics. Business Management and Industry. In the finance sector, statistical data facilitate decision-making. For instance, a watch manufacturing company can use statistical tools to determine the percentage of defective watches in every lot at a macro level, it helps in understanding a country's financial state and measuring economic growth. At a micro-level, statistics helps analysts determine a company's business income, earnings, and revenue-generating capacity. Be it preparing budgets, financial forecasts, monitoring a company, or a country's performance, statistics is everywhere.

5.Define primary data?

Ans: Primary data that has been generated by the researcher himself/herself, surveys, interviews, experiments, specially designed for understanding and solving the research problem at hand. It is a real time data. Primary data very involved process. The main source of primary data are surveys, observations, experiments, questionnaires, personal interviews, etc…. it always specific to the researchers needs.

Part B (Answer any eight questions. Each question carries 5 marks.)

6.Describe the nature of action research?

Ans: Action research is situation and context-based. Action research develops reflection practices based on the interpretations made by participants. Knowledge is created through action and application. Action research can be based in problem-solving, if the solution to the problem results in the improvement of practice.

In view of these issues, a particularly important kind of investigation for teachers is action research (sometimes also teacher research), an activity referring to systematic, intentional inquiry by teachers for the purpose of improving their own practice. Action research is not to be confused with research about teaching and learning, which are investigations by professional researchers on topics of teachers, teaching, or learning.

Action research has several defining characteristics, in addition to being planned and conducted by teachers.

First, it originates in the problems and dilemmas of classroom practice, or in chronic problems with certain students, materials, or activities.

Second, its outcomes offer information focused on particular teachers and classrooms, rather than about teachers in general or students in general.

Although this feature might make action research seem less useful as a source of advice or knowledge that is truly general, supporters argue that focusing on specific learning contexts makes action research more credible or valid as a source of practical information and ideas. It is, they argue, simply more attuned to the context of real classrooms.

Third, while the audience for action research can certainly include professors and educational administrators, the audience tends to be other teachers.

Action research is therefore in an especially strong position to provide "insider" perspectives on educational problems.

7. Describe the needs and different ways of stating hypothesis.

Ans: Hypothesis is an assumption that is made on the basis of some evidence. This is the initial point of any investigation that translates the research questions in to a prediction it includes components like variables, population and the relation between the variables. A research hypothesis is a hypothesis that is used to test the relationship between two or more variables.

Following are the characteristics of hypothesis.

- The hypothesis should be clear and precise to consider it to be reliable.
- if the hypothesis is a relational hypothesis, then it should be stating in the relations between variables.
- The hypothesis must be specific and should have scope for conducting more tests.
- The way of explanation of the hypothesis must be very simple and it should also be understood that the simplicity of the hypothesis is not related to its significance.
- There are six forms of hypothesis and they are:
    - Simple hypothesis
    - Complex hypothesis
    - Directional hypothesis
    - Non-directional hypothesis
    - Null hypothesis
    - Associative and causal hypothesis
- Simple Hypothesis

  It shows a relationship between one dependent variable and a single independent variable. For example – If you eat more vegetables, you will lose weight faster. Here, eating more vegetables is an independent variable, while losing weight is the dependent variable.
- Complex Hypothesis

  It shows the relationship between two or more dependent variables and two or more independent variables. Eating more vegetables and fruits leads to weight loss, glowing skin,

reduces the risk of many diseases such as heart disease, high blood pressure and some cancers.

- Null Hypothesis

  It provides the statement which is contrary to the hypothesis. It's a negative statement, and there is no relationship between independent and dependent variables. The symbol is denoted by "HO".

- Associative and Causal Hypothesis

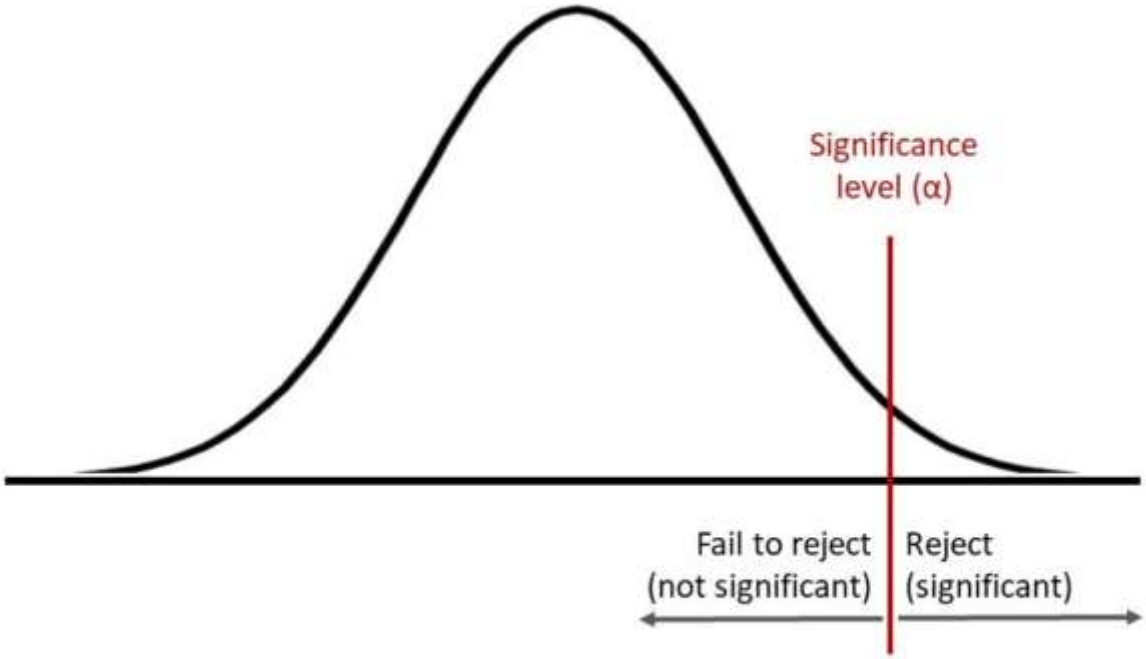
  Associative hypothesis occurs when there is a change in one variable resulting in a change in the other variable. Whereas, causal hypothesis proposes a causes and effect interaction between two or more variables.

8.Explain the term level of significance.

Ans: The level of significance is defined as the fixed probability of wrong elimination of null hypothesis when in fact, it is true. The level of significance is stated to be the probability of type I error and is present by the researcher with the outcomes of error. The level of significance is the measurement of the statistical significance. It defines whether the null hypothesis is assumed to be accepted or rejected. It is expected to identify if the result is statistically significant for the null hypothesis to be false or rejected.

Levels of significance of hypothesis;

In hypothesis testing, the level of significance is a measure of how confident you can be about rejecting the null hypothesis. This blog post will explore what hypothesis testing is and why understanding significance levels are important for your data science projects. In addition, you will also get to test your knowledge of level of significance towards the end of the blog with the help of quiz. These questions can help you test your understanding and prepare for data science / statistics interviews**.** Before we look into what level of significance is, let's quickly understand what is hypothesis testing.



9.state the characteristics of ethnographic research.

Ans: The ethnography**,** ethnographic research or village science is considered a method of social research that originated in the seventies. It has its origins in countries such as Great Britain, the United States and Australia, to improve the quality of education and solve problems at a pedagogical level.

The research has covered several areas in recent decades. Undoubtedly, the study of peoples and their behaviour is something that has captured the attention of professionals in order to better understand the issues of the sociocultural field.

Main characteristics of ethnography

According to Del Rincón, the characteristics of ethnography as a form of social research are:

Phenomenological or emico character

It consists of interpreting social phenomena from an "inside" perspective of the participants of that group of people. This allows the researcher to have a vision of what social life is like.

Through the description and interpretation, the professional can know the social phenomena seen as a member of this society. The term emic refers to the differences that exist within the same culture.

Relatively persistent stay

The professional must be involved in the group to study to get their acceptance and trust. This will allow a connection between the researcher and the members of the society to be created, a relationship that will make it possible to know the details of the grouping.

It is important for the expert to understand the culture he is studying. For this, many ethnographers decide to live the experience in the first person, since they can see events when they happen on a regular or daily basis.

It is holistic and naturalistic

Study the general reality of the facts seen from two points of view: an internal one, as if it were a member of the group; and the other external, properly the interpretation of the researcher as an outsider of said society.

Inductive character

The experience and the exploration are the tools to know a social scenario at first hand through the participant observation. From this strategy, information is obtained that generates conceptual categories.

The idea of research is to discover regularities and associations between social phenomena observed to analyzed them based on models, hypotheses and explanatory theories.

10.Explain the step of correlational research.

Ans: A correlational research design investigates relationships between variables without the researcher controlling or manipulating any of them. A correlation reflects the strength and/or direction of the relationship between two (or more) variables. The direction of a correlation can be either positive or negative.Here are five steps you can take to conduct a correlational study:

1. Make a claim or create a hypothesis. Making a claim or a hypothesis is often the first step in any study. ...
2. Choose a data collection method. ...
3. Collect your data. ...
4. Analyse the results. ...

5. Conduct additional research.
1. Make a claim or create a hypothesis. Making a claim or a hypothesis is often the first step in any study. ...

Making a claim or a hypothesis is often the first step in any study. In correlational studies, there may or may not be a clear relationship. For example, you can hypothesize that younger people are more likely to use digital coupons. Your hypothesis can clearly state an assumption about the relationship between two things. Identifying both of your variables when creating your hypothesis can help determine how you might collect your data.

2.Choose a data collection method. ...

There are a few ways you can choose to collect your data depending on the information you want to understand. Three of the most common data-collection methods are:

Surveys; One quick and flexible way to collect data is by asking questions. For example, if you want to determine the relationship between income and eating at a restaurant, you can provide questionnaires to willing participants

Observation; This is when you observe and record variable behaviour or quantitative data in the natural environment. If you want to determine the relationship between parents with children shopping at a specific store in the morning, you can tally how many people fit this variable for a few mornings to gather a larger sample size.

Secondary data; With secondary data, you might leverage existing data from polls or other studies to identify correlations. For example, you might collect international order data to see if there is a correlation between international customers and sales in a particular month.

Choosing the right method for your process can depend on what data you're collecting and what you want to do with it. For example, if one of your variables is highly subjective like preference, you might consider a survey. If you have access and time to record data using observation, that could be the most effective way to identify any correlations.

3. Collect your data.

Once you establish your method, collect the data. Consider a sample size that is big enough to determine correlation. For example, if your variables are people with brown hair shopping in the afternoon, you may want to observe these variables over many afternoons rather than just one.

Similarly, asking 10 people for their income and if they prefer picking up dinner or delivery might not provide you with a clear correlation.

4. Analyse the results. ...

Analysing data can help you identify if you have positive, negative or no correlation. Although there may be a positive or negative correlation, a third variable still might be the cause. For example, after studying the relationship between customer happiness and shopping in the morning, this could show shoppers have better moods because of the time of day, or it could be due to another factor like income. You might develop a causal hypothesis or a theory between what's causing the variables to affect one another, but it could be difficult to determine causation without research.

5.Conduct additional research

Correlational studies can create causal hypotheses, which can lead to conducting more specified research with controlled variables or additional correlational studies. For example, if a study at a coffee shop shows an increase in the number of people drinking coffee in the morning, you might see what other factors may influence this like music, staff or self-serve resources.

11.Describe the relationship between variables.

This relationship between the two variables is called a correlation. The amount of correlation, or relationship, can be explained in a numerical form called a correlation coefficient. A correlation coefficient is defined as a numerical representation of the strength and direction of the relationship. It is very important to understand relationship between variables to draw the right conclusion from a statistical analysis. The relationship between variables determines how the right conclusions are reached. Without an understanding of this, you can fall into many pitfalls that accompany statistical analysis and infer wrong results from your data.

There are several different kinds of relationships between variables. Before drawing a conclusion, you should first understand how one variable changes with the other. This means you need to establish how the variables are related - is the relationship linear or quadratic or inverse or logarithmic or something else

Suppose you measure a volume of a gas in a cylinder and measure its pressure. Now you start compressing the gas by pushing a piston all while maintaining the gas at the room temperature. The volume of gas decreases while the pressure increases. You note down different values on a

graph paper if you take enough measurements, you can see a shape of a parabola defined by x y = constant. This is because gases follow Boyle's law that says when temperature is constant, PV = constant. Here, by taking data you are relating the pressure of the gas with its volume. Similarly, many relationships are linear in nature.

12.what are the sources of hypothesis?

A hypothesis is an assumption that is made based on some evidence. This is the initial point of any investigation that translates the research questions into predictions. It includes components like variables, population and the relation between the variables. A research hypothesis is a hypothesis that is u Following are the characteristics of the hypothesis:

- The hypothesis should be clear and precise to consider it to be reliable.
- If the hypothesis is a relational hypothesis, then it should be stating the relationship between variables.
- The hypothesis must be specific and should have scope for conducting more tests.
- The way of explanation of the hypothesis must be very simple and it should also be understood that the simplicity of the hypothesis is not related to its significance.

sed to test the relationship between two or more variables.

Following are the sources of hypothesis:

- The resemblance between the phenomenon.
- Observations from past studies, present-day experiences and from the competitors.
- Scientific theories.
- General patterns that influence the thinking process of people.

There are six forms of hypothesis and they are:

- Simple hypothesis
- Complex hypothesis
- Directional hypothesis
- Non-directional hypothesis

- Null hypothesis
- Associative and casual hypothesis

13.Enumarate the advantages and dis advantages of sampling?

Sampling is a technique of selecting individual members or a subset of the population to make statistical inferences from them and estimate characteristics of the whole population. Different sampling methods are widely used by researchers in market research so that they do not need to research the entire population to collect actionable insights it is also a time-convenient and a cost-effective method and hence forms the basis of any research design. Sampling techniques can be used in a research survey software for optimum derivation.

Advantages of Sampling

- **Less expense of sampling:** If information somehow managed to be gathered for the whole population, the expense will be very high. A Sample is a little extent of a population. In this way, the expense will be lower assuming the information is gathered for a Sample of the population which is a major benefit.
- **Less tedious in Sampling:** Utilization of Sampling takes less time too. It consumes less time than the evaluation method. Arrangement, examination and so forth, take significantly less time on account of a Sample than on account of a population.
- **Extent of Sampling is high:** The specialist is worried about the speculation of information. To concentrate on an entire population to show up at speculations would be illogical. A few populations are huge to the point that their attributes couldn't be estimated. Before the estimation has been finished, the population would have changed. However, the most common way of examining makes it conceivable to show up at speculations by concentrating on the factors inside a moderately little extent of the population.
- **Exactness of information is high:** Having drawn a Sample and figured out the ideal graphic measurements, it is feasible to decide the security of the got test esteem. A Sample addresses the population from which it is drawn. It allows a serious level of precision because of a restricted area of activities. In addition, cautious execution of

fieldwork is conceivable. Eventually, the consequences of Sampling concentrates on end up being adequately exact.

- **Association of comfort:** Hierarchical issues engaged with Sampling are not many. Since the test of a little size, immense offices are not needed. Examining is in this manner affordable in regard to assets. Investigation of tests includes less space and hardware.
- **Escalated and thorough information:** In Sample studies, estimations or perceptions are made of a predetermined number. Along these lines, escalated and thorough information is gathered.
- **Reasonable in restricted assets:** The assets accessible inside an association might be restricted. Concentrating on the whole universe isn't feasible. The population can be acceptably covered through Sampling. Where restricted assets exist, utilization of Sampling is a fitting technique while directing promoting research.

Disadvantages of sampling

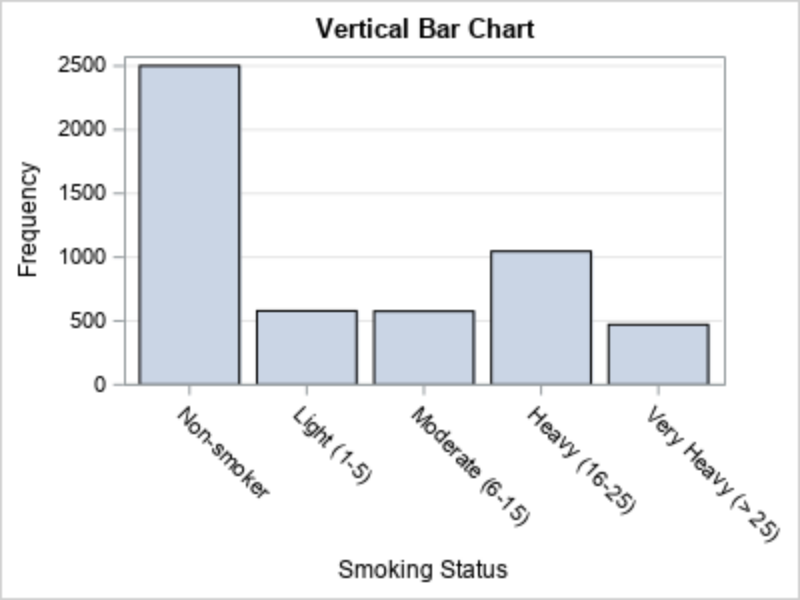
- **Chances of predisposition:** The genuine constraint of the examining technique is that it includes one-sided choice and in this manner drives us to reach incorrect determinations. Inclination emerges when the technique for choice of test utilized is broken. Relative little Samples appropriately chosen might be substantially more dependable than huge Samples ineffectively chosen.
- **Troubles in choosing a really delegated test:** Hardships in choosing a genuinely delegated test produces solid and precise outcomes just when they are illustrative of the entire gathering. Choice of a genuine agent test is troublesome when the peculiarities under review are of an intricate sort. Choosing great Samples is troublesome.
- **In satisfactory information in the subject:** Utilization of examining strategy requires sufficient subject explicit information in the Sampling method. A sampling includes factual examination and estimation of likely blunders. Whenever the analyst needs specific information in Sampling, he might submit genuine errors. Thusly, the aftereffects of the review will be misdirected.

- **Alterability of units:** Whenever the units of the population are not inhomogeneous, the Sampling procedure will be informal. In examining, however, the quantity of cases is little, it isn't simple 100% of the time to adhere to the, chosen cases. The units of the test might be broadly scattered. A portion of the instances of the test may not help out the analyst and some others might be blocked off. Due to these issues, every one of the cases may not be taken up. The chosen cases might need to be supplanted by different cases. Variability of units holds up traffic of consequences of the review.
- **Inconceivability of examining:** Inferring a delegate test is troublesome when the universe is excessively little or excessively heterogeneous. For this situation, registration study is the main other option. Additionally, in examinations requiring an exceptionally exclusive expectation of precision, the Sampling strategy might be unsatisfactory. There will be chances of mistakes regardless of whether tests are drawn most cautiously.

14.Discribe two type of bar chart with suitable examples.

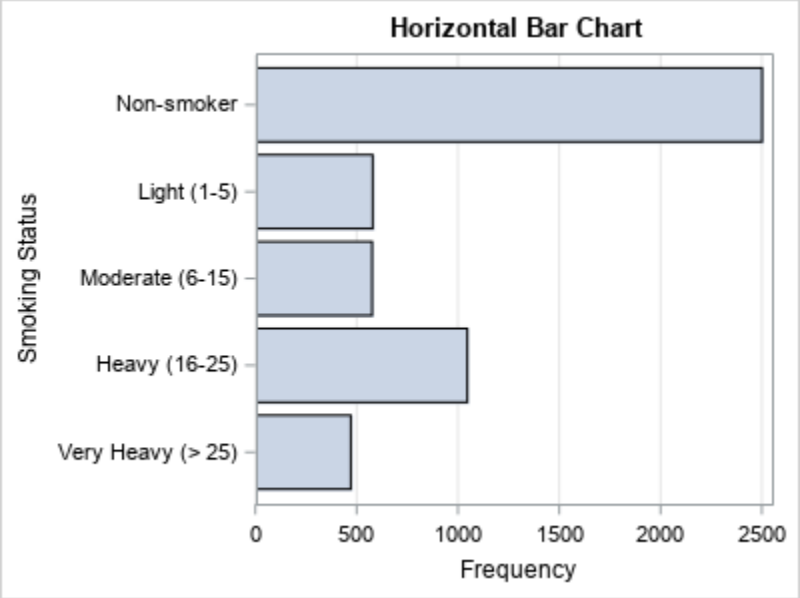
Ans: Vertical graph chart

When the grouped data are represented vertically in a graph or chart with the help of bars, where the bars denote the measure of data, such graphs are called vertical bar graphs. The data is represented along the y-axis of the graph, and the height of the bars shows the values.



Horizontal graph chart

When the grouped data are represented horizontally in a chart with the help of bars, then such graphs are called horizontal bar graphs, where the bars show the measure of data. The data is depicted here along the x-axis of the graph, and the length of the bars denote the values.



15.when do you use secondary data? Mention the sources of secondary data.

Ans: Secondary data is usually gathered from the published (printed) sources. A few major sources of published information are as follows: Published articles of local bodies, and central and state governments. Statistical synopses, census records, and other reports issued by the different departments of the government.

Secondary data refers to data that is collected by someone other than the primary user. Common sources of secondary data for social science include censuses, information collected by government departments, organizational records and data that was originally collected for other research purposes.

Sources of secondary data. Secondary data can be obtained from many sources:

- censuses and government departments like housing, social security, electoral statistics, tax records
- internet searches and libraries
- GPS and remote sensing
- km progress reports
- journals, newspapers and magazines

6.Describe the scope of statistics.

Ans: Statistics is the study of collection, organization, analysis, interpretation and presentation of data with the use of quantified models. In short, it is a mathematical tool that is used to collect and summarize data. Scope of Statistics: It presents the facts in numerical figures. Statistics is the branch of mathematics for collecting, analysing and interpreting data. Statistics can be used to predict the future, determine the probability that a specific event will happen, or help answer questions about a survey. Statistics is used in many different fields such as business, medicine, biology, psychology and social sciences.

The following are the main scope of statistics:

1. Presents facts in numerical figures
2. Presents complex facts in a simplified form
3. Studies relationship between two or more phenomena
4. Helps in the formulation of policies
5. Helps in forecasting

6. Provides techniques for testing of hypothesis
7. Provides techniques for making decisions under uncertainty

1.Presents facts in numerical figures

The first function of statistics is to present a given problem in terms of numerical figures. We know that the numerical presentation helps in having a better understanding of nature an of problem.

2.Presents complex facts in a simplified form

Generally, a problem to be investigated is represented by a large mass of numerical figures which are very difficult to understand and remember. Using various statistical methods, this large mass of data can be presented in a simplified form.

3.Studies relationship between two or more phenomena

Statistics can be used to investigate whether two or more phenomena are related. For example, the relationship between income and consumption, demand and supply, etc.

4.Helps in the formulation of policies

Statistical analysis of data is the starting point in the formulation of policies in various economic, business and government activities.
For example, using statistical techniques a firm can know the tastes and preferences of the consumers and decide to make its product accordingly.

5.Helps in forecasting

The success of planning by the Government or of a business depends to a large extent upon the accuracy of their forecasts. Statistics provides a scientific basis for making such forecasts.

6.Provides techniques for testing of hypothesis.

A hypothesis is a statement about some characteristics of a population (or universe).

7.Provides techniques for making decisions under uncertainty

Many time we face an uncertain situation where any one of the many alternatives may be adopted. A businessman might face a situation of uncertain investment opportunities in which he can lose or gain. He is may be interested in knowing whether to undertake a particular investment or not. The answer to such problem are provided by the statistical techniques of decision-making under uncertainty.

17. Write a brief note on kurtosis.

- **Ans:** A statistical measure that is used to describe distribution which measures extreme values in either tail.

- Distributions with large kurtosis exhibit tail data exceeding the tails of the normal distribution.
- Distributions with low kurtosis exhibit tail data that are generally less extreme than the tails of the normal distribution.
- For investors, high kurtosis of the return distribution implies the investor will experience occasional extreme returns (either positive or negative), more extreme than the usual + or - three standard deviations from the mean that is predicted by the normal distribution of returns.
- This phenomenon is known as kurtosis risk
- Kurtosis is a measure of the combined weight of a distribution's tails relative to the center of the distribution.
- kurtosis is a measure that describes the shape of a distribution's tails in relation to its overall shape.
- A distribution can be infinitely peaked with low kurtosis, and a distribution can be perfectly flat-topped with infinite kurtosis. Thus, kurtosis measures "tailedness," not "peakedness."

There are three categories of kurtosis

- The first category of kurtosis is a mesokurtic distribution. This distribution has a kurtosis statistic similar to that of the normal distribution, meaning the extreme value characteristic of the distribution is similar to that of a normal distribution
- The second category is a leptokurtic distribution. Any distribution that is leptokurtic displays greater kurtosis than a mesokurtic distribution. Characteristics of this distribution is one with long tails (outliers.).
- Examples of leptokurtic distributions are the T-distributions with small degrees of freedom.
- The final type of distribution is a platykurtic distribution.
- These types of distributions have short tails (paucity of outliers.) The prefix of "platy-" means "broad," and it is meant to describe a short and broad-looking peak, but this is an historical error.
- Uniform distributions are platykurtic and have broad peaks.

Part C (Answer any two questions. Each question carries 15 marks.)

18.Discuss the type of probability sampling in detail.

Ans: Probability sampling is a sampling technique in which researchers choose samples from a larger population using a method based on the theory of probability. This sampling method considers every member of the population and forms samples based on a fixed process.

Here are four types of probability sampling techniques:

1.Simple random sampling: One of the best probability sampling techniques that helps in saving time and resources, is the Simple Random Sampling method. It is a reliable method of obtaining information where every single member of a population is chosen randomly, merely by chance. Each individual has the same probability of being chosen to be a part of a sample.

2.Cluster sampling: Cluster sampling is a method where the researchers divide the entire population into sections or clusters that represent a population. Clusters are identified and included in a sample based on demographic parameters like age, sex, location, etc. This makes it very simple for a survey creator to derive effective inference from the feedback.

3.Systematic sampling: Researchers use the systematic sampling method to choose the sample members of a population at regular intervals. It requires the selection of a starting point for the sample and sample size that can be repeated at regular intervals. This type of sampling method has a predefined range, and hence this sampling technique is the least time-consuming.

4.Stratified random sampling: Stratified random sampling is a method in which the researcher divides the population into smaller groups that don't overlap but represent the entire population. While sampling, these groups can be organized and then draw a sample from each group separately.

There are multiple uses of probability sampling:

Reduce Sample Bias: The selection of the sample mainly depicts the understanding and the inference of the researcher. Probability sampling leads to higher quality data collection as the sample appropriately represents the population.

Diverse Population: When the population is vast and diverse, it is essential to have adequate representation so that the data is not skewed towards one demographic.

Types of non-probability sampling with examples;

The non-probability method is a sampling method that involves a collection of feedback based on a researcher or statistician's sample selection capabilities and not on a fixed selection process.

Four types of non-probability sampling

1.Convenience sampling: This method is dependent on the ease of access to subjects such as surveying customers at a mall or passers-by on a busy street. It is usually termed as convenience sampling, because of the researcher's ease of carrying it out and getting in touch with the subjects. Researchers have nearly no authority to select the sample elements, and it's purely done based on proximity and not representativeness.

2.Judgmental or purposive sampling: Judgemental or purposive samples are formed by the discretion of the researcher. Researchers purely consider the purpose of the study, along with the understanding of the target audience.

3.Snowball sampling: Snowball sampling is a sampling method that researchers apply when the subjects are difficult to trace. For example, it will be extremely challenging to survey shelter less people or illegal immigrants. In such cases, using the snowball theory, researchers can track a few categories to interview and derive results.

4.Quota sampling: In Quota sampling, the selection of members in this sampling technique happens based on a pre-set standard. In this case, as a sample is formed based on specific attributes, the created sample will have the same qualities found in the total population. It is a rapid method of collecting samples.

Non-probability sampling is used for the following:

Create a hypothesis: Researchers use the non-probability sampling method to create an assumption when limited to no prior information is available. This method helps with the immediate return of data and builds a base for further research.

Exploratory research: Researchers use this sampling technique widely when conducting qualitative research, pilot studies, or exploratory research.

Budget and time constraints: The non-probability method when there are budget and time constraints, and some preliminary data must be collected. Since the survey design is not rigid, it is easier to pick respondents at random and have them take the survey or questionnaire.

18. Explain the steps of case study research.

Ans: A **case study** is a detailed study of a specific subject, such as a person, group, place, event, organization, or phenomenon. Case studies are commonly used in social, educational, clinical, and business research.

A case study research design usually involves qualitative methods, but quantitative methods are sometimes also used. Case studies are good for describing, comparing, evaluating and understanding different aspects of a research problem.

A case study is an appropriate research design when you want to gain concrete, contextual, in-depth knowledge about a specific real-world subject. It allows you to explore the key characteristics, meanings, and implications of the case.

Case studies are often a good choice in a thesis or dissertation. They keep your project focused and manageable when you don't have the time or resources to do large-scale research.

You might use just one complex case study where you explore a single subject in depth, or conduct multiple case studies to compare and illuminate different aspects of your research problem.

Step 1: Select a case

Once you have developed your problem statement and research questions, you should be ready to choose the specific case that you want to focus on. A good case study should have the potential to:

- Provide new or unexpected insights into the subject
- Challenge or complicate existing assumptions and theories
- Propose practical courses of action to resolve a problem
- Open up new directions for future research

Unlike quantitative or experimental research, a strong case study does not require a random or representative sample. In fact, case studies often deliberately focus on unusual, neglected, or outlying cases which may shed new light on the research problem.

Step 2: Build a theoretical framework

While case studies focus more on concrete details than general theories, they should usually have some connection with theory in the field. This way the case study is not just an isolated description, but is integrated into existing knowledge about the topic. It might aim to:

- **Exemplify** a theory by showing how it explains the case under investigation
- **Expand** on a theory by uncovering new concepts and ideas that need to be incorporated
- **Challenge** a theory by exploring an outlier case that doesn't fit with established assumptions

To ensure that your analysis of the case has a solid academic grounding, you should conduct a literature review of sources related to the topic and develop a theoretical framework. This means identifying key concepts and theories to guide your analysis and interpretation.

Step 3: Collect your data

There are many different research methods you can use to collect data on your subject. Case studies tend to focus on qualitative data using methods such as interviews, observations, and analysis of primary and secondary sources (e.g., newspaper articles, photographs, official records). Sometimes a case study will also collect quantitative data.

Examples of a mixed methods case study

For a case study of a wind farm development in a rural area, you could collect quantitative data on employment rates and business revenue, collect qualitative data on local people's perceptions and experiences, and analyse local and national media coverage of the development.
The aim is to gain as thorough an understanding as possible of the case and its context.

Step 4: describe and analyse the case

In writing up the case study, you need to bring together all the relevant aspects to give as complete a picture as possible of the subject. How do you report your findings depends on the type of research you are doing. Some case studies are structured like a standard scientific paper or thesis, with separate sections or chapters for the methods, results and discussion.

Others are written in a more narrative style, aiming to explore the case from various angles and analyse its meanings and implications (for example, by using textual analysis or discourse analysis).

In all cases, though, make sure to give contextual details about the case, connect it back to the literature and theory, and discuss how it fits into wider patterns or debates.

20.Describe the diagrammatic and graphical representation of data and their significance.

Ans: Representation of any numerical data by using diagrams is known as diagrammatic representation. Diagrammatic data representations give a simple and easy understanding of any numerical data collected as compared with the tabular form of the data or textual form of the data. Diagrammatic representations use geometrical figures as diagrams to improve the representation of the data. Diagrammatic representations are like visual assistance to the readers.

Basics of diagrammatic presentations

Diagrammatic representation of data gives a lot of information regarding numerical data. It is a more attractive and easy way of representing any numerical data in statistics. Diagrammatic

representations are like visual assistance to the readers. Diagrammatic representations use the geometrical figures as diagrams to improve the data representation, such as cartography, pictographs, Pie charts, bar diagrams, etc

In pictographic representation of the data, we use pictures to represent the data

1. In the cartograms, we represent the geographical location of certain things, and we use maps.
2. Bar graphs are represented by rectangle bars. The height of the bars gives the value or frequency of the variable. All rectangular bars should have equal width.
3. In the pie charts, a circle is divided into parts, such that each part shows the proportion of various data.
4. In a line representation of data, we use the line to connect the various portions or parts of the plotted data on the graph.

Advantages of diagrammatic presentation

The various advantages of the diagrammatic representations are listed below:

1. The diagrammatic representations of the data are more attractive and pretty impressive compared with the tabular form of the data or textual form of the data.
2. The diagrammatic representations of the data are easy to remember as they use the geometrical figures as the diagrams.
3. The diagrammatic representation of data is easy to understand.
4. Diagrammatic data representations translate the highly complex ideas included in the given numerical data into concrete and pretty effectively in a simple, understandable manner.
5. Diagrammatic representations also help identify hidden facts or relations in the data that are not observed in the tabular form.
6. Diagrammatic representations of the data are a handy tool in the comparison of data.
7. **Types of one - dimensional diagrams**

In one-dimensional diagrammatic representations of the data, we will consider only the length of the diagram. We have different types of one-dimensional diagrams that are listed below:

- **Simple bar diagram**
- **Multiple bar diagrams**

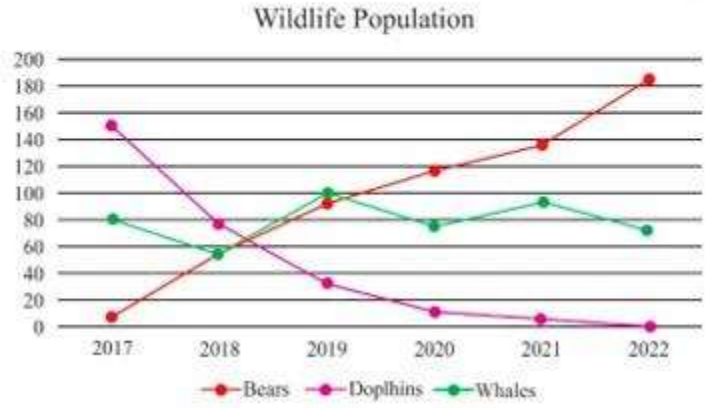
- **Subdivided bar diagrams**
- **Percentage bar diagram**
- **Deviation bar diagram**

Types of of Diagrammatic Representations

Diagrammatic representations use the geometrical figures as diagrams to improve the data representation, such as cartographs, pictographs, Pie charts, bar diagrams, etc.

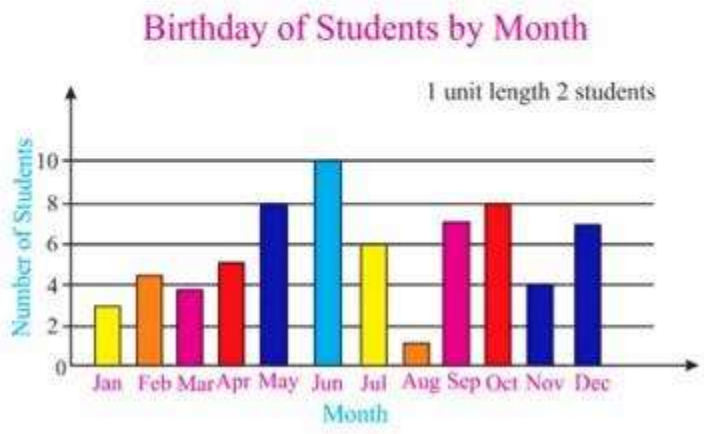
1. Line Diagrams

In the linear diagrammatic representations of the data, we will use the line that connects the points or portions of the various data in the graph by taking two variables on
horizontal



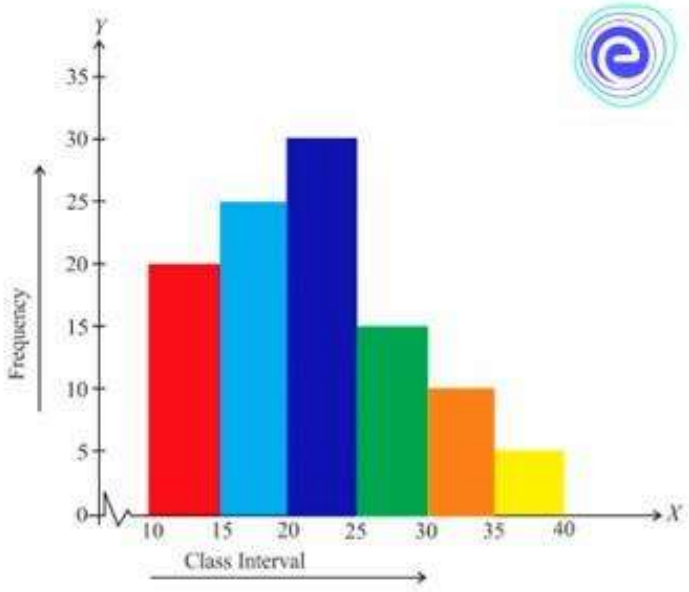
2. Bar Diagrams

In the bar diagrammatic representation of data, the data can be represented by rectangular bars. The height of the bars gives the value or frequency of the variable. All rectangular bars should have equal width. This is one of the best-used tools for the comparison of the data.



3. Histograms

Histograms are also similar to bar diagrams; they use rectangular bars to represent the data. But all the rectangular bars are kept without any gaps.

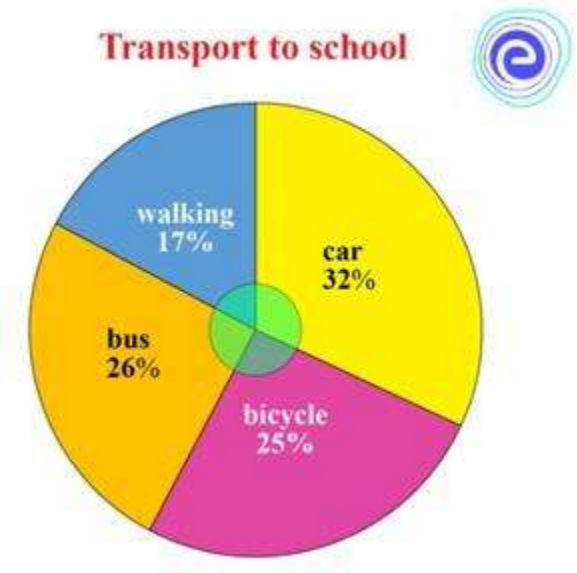


4. Pie Diagrams

Pie Diagram is a diagrammatic representation of data by using circles and spheres. In the pie diagrams, a circle is divided into parts, such that each part shows the proportion of various data.

Example:

The below pie diagram represents the different modes of transport used by the students.



5. Pictographs

The pictographic representation shows the given data graphically by using images or symbols. The symbol or image is used in the pictographic diagrams describes the frequency of the object in the given set of data. Pictographs provided the information of the given data by using symbols or images.

Example:

The pictograph diagram below shows the mode of transport used by the number of students using the image, and each image represents the value.

Mode of transport	Number of students
Bus	
Car	
Walking	
Bicycle	
Key:  Represents 3 children	

21.(a)The median is midpoint between the highest and the lowest scores in a distribution. Do you agree or disagree? why?

(b)Differentiate between point biserial and biserial correlation?

Ans: The median is the middle value of a series of values laid out in numerical order. It is the middle point of the data set; it's also known as the midpoint. Half of the values in your list will be less than the median, and half will be higher than the median.

The Difference Between Mean and Median

Now if we compare our median value of $55,000 to the mean we calculated of $52,800, you might be thinking those numbers are pretty close so what's the big difference? That's an excellent question.

If the values in the data set are not to spread out, like in our example, the mean and the median may not be very different.

company for 20 years, each at $88,000 and a controller at $120,000. Now let's see what we get when we do the math.

Finding the median

1. Arrange the data points from smallest to largest.
2. If the number of data points is odd, the median is the middle data point in the list.
3. If the number of data points is even, the median is the average of the two middle data points in the list.

Whenever you find yourself writing the words, "the average worker" this, or "the average household" that, you don't want to use the mean to describe those situations. You want a statistic that tells you something about the worker or the household in the middle. That's the **median**.
Again, this statistic is easy to determine because the median literally **is** the value in the middle. Just line up the values in your set of data, from largest to smallest. The one in the dead- center is your median.
For the WorldWideWidget Co., here are the worker's salaries:

The biserial correlation is a correlation between on one hand, one or more quantitative variables, and on the other hand one or more binary variables. It was introduced by Pearson (1909)
The biserial correlation is a correlation between on one hand, one or more quantitative variables, and on the other hand one or more binary variables. It was introduced by Pearson (1909). The biserial correlation can be calculated with XLSTAT.

As for the Pearson correlation, the biserial correlation coefficient varies between -1 and 1. 0 corresponds to no association (the means of the quantitative variable for the two categories of the qualitative variable are identical).
Biserial correlation in XLSTAT
XLSTAT allows testing if the value of the biserial correlation r that has been obtained is different from 0 or not.For the two-tailed test, the null H0 and alternative Ha hypotheses are as follows:

- H0: $r = 0$
- Ha: $r \neq 0$

In the left one-tailed test, the following hypotheses are used:

- H0: r = 0
- Ha: r < 0

In the right one-tailed test, the following hypotheses are used:

- H0: r = 0
- Ha: r > 0

Two methods to compute the p-value are proposed by XLSTAT. The user can choose between a p-value computed using on a large sample approximation, and a p-value computed using Monte Carlo re samplings. The second method is recommended.
Note: the XLSTAT_ Biserial spreadsheet function can be used to compute the biserial correlation between a quantitative variable and a binary variable.

A point-biserial correlation is used to measure the strength and direction of the association that exists between one continuous variable and one dichotomous variable. It is a special case of the Pearson's product-moment correlation, which is applied when you have two continuous variables, whereas in this case one of the variables is measured on a dichotomous scale.

For example, you could use a point-biserial correlation to determine whether there is an association between salaries, measured in US dollars, and gender (i.e., your continuous variable would be "salary" and your dichotomous variable would be "gender", which has two categories: "males" and "females"). Alternately, you could use a point-biserial correlation to determine whether there is an association between cholesterol concentration, measured in mmol/L, and smoking status (i.e., your continuous variable would be "cholesterol concentration", a marker of heart disease, and your dichotomous variable would be "smoking status", which has two categories: "smoker" and "non-smoker").

This "quick start" guide shows you how to carry out a point-biserial correlation using SPSS Statistics, as well as how to interpret and report the results from this test. However, before we introduce you to this procedure, you need to understand the different assumptions that your data

must meet in order for a point-biserial correlation to give you a valid result. We discuss these assumptions.

FIRST SEMESTER M.Ed. DEGREE EXAMINATION, DECEMBER 2021

INTRODUCTION TO EDUCATIONAL RESARCH AND STATISTICS

M.Ed. 04 **Prepared by: Soumya. V (2022-24)**

PART A

1.Mention the types of educational research

Three Types of research

1.Descripitive research

2.Correlational research

3.Experimental research

2.What is moderator variable?

A moderating variable can either be categorical or continuous, and is used exclusively in quantitative, rather than qualitative, research. Moderating variables are useful because they help explain the links between the independent and dependent variables.

3. Define sampling error

A statistical error that occurs when the analyst selects a sample that is not representative of the population being studied.

4. Give any two merits of quartile deviation

1.Easily understood

2. It is superior to range as it is not affected by extremity large or small observations.

3. It can be calculated in case of open-end frequency distribution as well.

5. What are the three types of kurtoses

1.Mesokurtosis: An excess kurtosis of 0. Normal distributions are mesokurtic.

2. Platykurtic: A negative excess kurtosis. Platy kurtic distributions are thin-tailed meaning that they have few outliers.

3.Leptokurtosis: A positive excess kurtosis.

(5*2=10)

PART B

6. Critically analyse the scope of educational research

Critical Studies in Education is a key global forum for advancing original research that addresses the implications of intensifying social, political and economic challenges and how this shape education.

Scope of educational research

1.educational psychology

2. philosophy of education

3. sociology of education

4. educational administration

5. cooperative education

6. teacher educational and teaching behaviour

7. educational technology

7. Write short notes on paradigms of educational research

Paradigm conceptual framework for seeing and making sense of the social world

A paradigm comprises four elements epistemology, ontology, methodology and axiology.

Method of paradigm is research design; methods approaches and procedures used in an investigation. How shall go about obtaining the desired data, knowledge. A model or pattern we

follow when conducting research, including both our surface-level methods as well as our deeper, often hidden, beliefs, attitudes, expectations, and values.

8 What do you understand by selection of research problem?

Selection of research problem depends on several factors such as researchers' knowledge, skills, interest, expertise, motivation& creativity with respect to the subject of inquiry. It is believed that most of the good research studies need lots of time for selection of a research problem.

9 Describe briefly the preparation of related literature report.

- Narrow your topic and select papers accordingly.
- Search for literature
- Read the selected articles thoroughly and evaluate them.
- Organise the selected papers by looking for patterns and by developing subtopics.
- Develop a thesis or purpose statement.
- Write the paper.
- Review your work.

10 Briefly explain the various quantitative research methods

Quantitative research is explaining phenomena by collecting numerical data that are analysed using mathematically based method. Quantitative methods emphasize objective measurements and the statistical, mathematical or numerical analysis of data collected through polls, questionnaires and surveys, or by manipulating pre-existing statistical data using computational techniques. Quantitative research focuses on gathering numerical data and generalizing it across group of people or to explain a particular phenomenon.

11 What is the basic purpose of statistical methods in correlation studies?

- evaluating relationships for numerical scores
- Scores in each pair are identified as X and Y
- Data can be presented in a list showing the two scores for each individual.
- Scores can be shown in a scatter plot graph

12. Differentiate between qualitative and quantitative data

	Qualitative	Qualitative
Conceptual	Concerned with understanding Human behaviour from the information's perspective.	Concerned with discovering facts about social phenomena.
	Assumes a dynamic and negotiated reality	Assumes a fixed and measurable reality.
Methodological	Data are analysed by themes from descriptions by informants	Data are analysed through numerical comparisons and statistical inference
	Data are collected through participant observation and interviews	Data are collected through measuring things
	Data are reported in the language of the informant	Data are reported through statistical analyses.

13.Explain about size of sampling.

Sample size refers to number to the number of participants or observations included in a study. This number is usually represented by n. The size of a sample influences two statistical properties.

1. The precision of our estimates
2. The power of the study to draw conclusions

14. Clarify the importance and purpose of hypothesis.

It helps to provide link to the underlying theory and specific research question. It helps in data analysis and measure the validity and reliability of the research. It provides a basis or evidence to prove the validity of the research.

Importance of Hypothesis

1 It ensure the entire research methodologies are scientific and valid

2.It help to assume the probability of research failure and progress.

3.Ihelp to provide link to the underlying theory and specific research question.

4.It help in data analysis and measure the validity and reliability of the research.

5. It provides a basis or evidence to prove the validity of the research.

15. What is called normal probability curve? Explain with examples

The normal probability curve shows the probability distribution for continuous random variable and the curve is bell- shaped. The graph of the probability density function of the normal distribution is symmetrical about the mean and it is called Normal probability curve

16. Describe the types of correlation.

There are three types correlation

1.positive and negative correlation.

2. Linear and non-linear correlation.

3. simple, multiple, and partial correlation.

PART -C

17.a. Explain the process of sampling and its methods.

b. Explain the types of hypotheses.

Sampling means selecting the group that you will actually collect data from in you research. For example, if you are researching the opinions of students in your university, you could survey a sample of 100 students. In statistics, sampling allows you to test a hypothesis about the characteristics of population.

The sample is the group of individuals who will actually participate in the research. There are two primary types of sampling methods that you can use in your research.

1.Probability sampling: involves random selection, allowing you to make strong statistical inferences about the whole group.

2. non-probability sampling: involves non- random selection based on convenience or other criteria, allowing you to easy collect data.

19.Write an essay on kinds of dispersion.

Measure of dispersion indicates the scattering of data. It explains the disparity of data from one another, delivering a precise view of their distribution. The measure of dispersion displays and gives us an idea about the variation and the central value of an individual item. In other words, dispersion is the extent to which values in a distribution differ from the average of the distribution. It gives us an idea about the extent to which individual items vary from one another, and from the central value.

The variation can be measured in different numerical measures

1.Range

2.Quatile deviation

3.Mean deviation

4.Standard deviation

Characteristics of a Good Measure of Dispersion

1.It should be easy to calculate and simple to understand

2.It should be based on all the observation of the series

3. It should be rigidly defined.

4.It should not be affected by extreme values.

5.It should not be unduly affected by sampling

20 a Explain the sources of primary data

Primary data collection is the process of collecting information directly from users.

This type of data collection is usually done through surveyes or interviews

Mainly primary data sources

- Autobiographies and memoirs.

- Diaries, personal letters, and correspondence.
- Interviews, surveys and fieldwork.
- Internet communications on email, blogs, listservs, and newsgroups.

b. from the following data calculate quartile mean and standard deviation

Class interval	F	C.I	F	C.F
60-65	2	0-5	5	5
54-59	2	6-11	6	11
48-53	6	12-17	10	21
42-47	3	18-23	6	27
36-41	8	24-29	4	31
30-35	8	30-35	8	39
24-29	4	36-41	8	47
18-23	6	42-47	3	50
12-17	10	48-53	6	56
6-11	6	54-59	2	58
0.5	5	60-65	2	60

N=60

---

21 a By Spearman's rank Difference Method calculate correlation coefficient

student	A	B	C	D	E	F	G	H	I	J
Maths	78	36	98	25	75	80	25	62	36	40
Hindi	84	54	36	60	54	92	36	62	36	98