Unit 3- RESEARCH DESIGN AND ANALYSIS

SYLLABUS:

Meaning of research design - Need of research design - Different research designs - Basic principles of experimental design - Developing a research plan - Design of experimental set-up - Use of standards and codes. Overview of Multivariate analysis, Hypotheses testing and Measures of Association. Presenting Insights and findings using written reports and oral presentation. 9 Hours

3.1 MEANING OF RESEARCH DESIGN

(Q: Define research design. Explain the components and important features of research design.

Q: Discuss the various types of research designs and their applications.

Q: Highlight the differences between exploratory and descriptive/ diagnostic type of research designs)

- Research design refers to '**designing the research project'**.
- **Research design** stands for advance planning of the methods to be adopted for collecting the relevant data, techniques for data analysis, keeping in view the objective of the research and the availability of staff, time and money.

3.1.1 The parts of research design:

(a) **Sampling design**: It deals with the method of selecting items to be observed for the given study;

(b) **Observational design**: It relates to the conditions under which the observations are to be made;

(c) **Statistical design**: It is concerned with the question of how many items are to be observed and how the information and data gathered are to be analysed.

(d) **Operational design**: It deals with the techniques of carrying out the procedures specified in the sampling, statistical and observational designs.

3.1.2 The important features of a research design:

(a) A clear statement of the research problem;

(b) Procedures and techniques to be used for gathering information;

(c) The population to be studied;

(d) Methods to be used in processing and analysing data.

(e) It includes the constraints time and cost budgets.

3.2 NEED (Importance) FOR RESEARCH DESIGN

- Facilitates the smooth sailing of the various research operations, making research as efficient as possible yielding maximal information with minimal expenditure.
- Ex: Better, economical and attractive construction of a house, needs a blueprint well thought out and prepared by an expert architect. Similarly, for the data collection and analysis, research project needs a research design.

3.3 DIFFERENT RESEARCH DESIGNS

Different research designs are:

(1) Research design in case of **exploratory research** studies

(2) Research design in case of **descriptive and diagnostic research** studies

(3) Research design in case of **hypothesis-testing research** studies.

1. Research design in case of exploratory research (formulative research) studies:

- The main purpose is to formulate a problem for investigation or to develop the working hypotheses from an operational point of view.
- The major emphasis is on the discovery of ideas and insights; the design must be **flexible**.

2. Research design in case of descriptive and diagnostic research studies:

- Descriptive research studies are concerned with describing the characteristics of a particular Individual/ group, whereas diagnostic research studies determine the frequency with which something occurs / its association with something else.
- This survey design **must be rigid** and **not flexible**. It includes:

(a) **Formulating the objective of the study with precision** – deals with what the study is about and why is it being made.

(b) **Designing the methods of data collection** –deals with what techniques of gathering data will be adopted. Observation, questionnaires, interviewing, examination of records, etc. are the methods available. Researcher should safeguard against bias and unreliability.

(c) **Selecting the sample** –deals with how much material will be needed. Researcher takes out sample(s) and then infers about the population on the basis of the sample analysis.

(d) **Collecting the data** –deals with where can the required data be found and with what time period should the data be related. As data then should be examined for completeness, comprehensibility, consistency and reliability.

(e) **Processing and analyzing the data**: This includes coding the interview replies, observations, tabulating the data; and performing several statistical computations. Coding should avoid errors, tabulation should be accurate, statistical operations and tests of significance should be appropriate to draw proper conclusions.

(f) **Reporting the findings**: To communicate the findings to others efficiently, the layout of the report needs to be well planned to present in simple and effective style.

The difference between two types of research designs are summarized in Table 3.1:

Table 3.1

	Type of study	
Research Design	*Exploratory of Formulative*	*Descriptive/Diagnostic*
Overall design	Flexible design (design must provide opportunity for considering different aspects of the problem)	Rigid design (design must make enough provision for protection against bias and must maximise reliability)
(i) Sampling design	Non-probability sampling design (purposive or judgement sampling)	Probability sampling design (random sampling)
(ii) Statistical design	No pre-planned design for analysis	Pre-planned design for analysis
(iii) Observational design	Unstructured instruments for collection of data	Structured or well thought out instruments for collection of data
(iv) Operational design	No fixed decisions about the operational procedures	Advanced decisions about operational procedures.

3. Research design in case of hypothesis-testing research (experimental studies) **studies:**

- In this, the researcher tests the hypotheses of causal relationships between variables. Such studies require procedures to reduce bias, increase reliability, and inferences about causality.

3.4 BASIC PRINCIPLES OF EXPERIMENTAL DESIGNS

(Q: What is meant by experimental design? What are its benefits ?
Q: Differentiate between:
i. factors vs. levels ii. dependent vs. independent variables
iii. full factorial designs vs. fractional factorial designs
Q: Enumerate the principles of experimental design.
Q: Classify the Formal experimental designs and illustrate each with example
Q: Clarify the use of orthogonal arrays and response surface designs)

3.4.1 Experimental design: It provides a powerful framework for conducting research that can yield valid, reliable, and insightful results. It outlines how data is collected, what variables are manipulated, how results are interpreted and allows researchers to investigate cause-and-effect relationships between variables.

3.4.2 Benefits of experimental design:

- **Resource efficiency**: More information per experiment will be obtained when compared with unplanned experiments.
- **Control over variables**: Controls external variables (confounding factors) that might otherwise influence the dependent variable
- **Establishing causality**: Helps to draw conclusions about cause-and-effect relationships between variables.
- **Replicability**: Results can be tested in different contexts, strengthening reliability of the findings.
- **Random assignment**: Reduces bias, ensures random allocation.
- **Precision**: Facilitates accurate measurement of effects of an independent variable on a dependent variable

3.4.3 Important terms used:

- **Dependent and Independent variables:** Independent variables are the factors manipulated or controlled by the experimenter to observe their effect on outcomes. **Dependent variables** (response variable) are the outcomes measured in an experiment, and they depend on changes in the independent variables.
  Ex: In a study on medication effects, the **dosage of a drug** (e.g., 10 mg, 20 mg, 30 mg) is the **independent variable** (response variable, outcome) that is controlled, while **blood pressure levels** are the **dependent variables (factors)** measured to observe the drug's effect on lowering blood pressure.
- **Factors and Levels:**

In experimental design, factors are the variables of interest, and levels are the specific values those variables take in the experiment:

Factor: A factor is an independent variable that is expected to influence the outcome of an experiment. For example, in a chemical reaction study, factors might include temperature, pressure, and concentration.

Level: A level is a specific value or setting of a factor. Each factor can have multiple levels. For example, if "temperature" is a factor, its levels might be 100°C, 150°C, and 200°C. Levels represent the different conditions under which the factor is tested.

Treatment (condition): A "treatment" refers to a specific combination of factors and their corresponding levels that are applied to the experimental group (group of units that receive the treatment) to observe its effect on the outcome being studied.

3.4.4 Principles of experimental designs

Three **principles of experimental design,** as per Professor Fisher:

(1) The Principle of Replication: As per this, the experiment should be repeated more than once and each treatment is applied in many experimental units to improve the statistical accuracy. **Ex:** Testing two rice varieties: **without replication**: Divide the field into two parts and grow one variety in each, and compare their yields. **With replication:** First divide the field into several parts, grow one variety in half and the other variety in the rest. Then collect the yield and compare the results.

(2) The Principle of Randomization: It helps to protect experiments, from the influence of external factors by randomly assigning treatments. **Ex:** if one rice variety is planted in one half of the field and another variety in the other half, any external factors, like soil fertility differences could affect the results. To avoid this, randomly assign the rice varieties to different parts of the field.

(3) **Principle of Local Control**: Through this, one can eliminate the variability due to external factors reducing the experimental error. **Ex:** researchers should design experiments so that a two-way analysis of variance can be used. This method separates the total variability into three parts: treatments (e.g., rice varieties), extraneous factors (e.g., soil fertility), and experimental error. To apply this, the field is divided into homogeneous sections, called blocks. Each block is then split into parts equal to the number of treatments. Treatments are randomly assigned to each part of a block. "Blocking" (Explained in next section) ensures that the influence of external factors is controlled, allowing their impact on the results to be measured accurately.

3.4.5 Important Experimental Designs

Experimental designs are broadly classified into two categories:

i) Informal experimental designs: Includes less sophisticated analysis**.**

Types:

- **a) Before-and-After Without Control Design:** In this design, a single test group is selected and a dependent variable (DV) is measured at two points in time: once before the treatment and once after the treatment.
- **b) After-only with control design**: In this design, two groups (test group and control group) are selected and treatment is given to test group only. The impact of treatment is assessed by subtracting the value of the DV in the control group from the value in the test group.
- **c) Before-and-After With Control Design:** In this design, two groups (test group and control group) are selected and DV is measured in both the groups before the treatment. The treatment is given to test group only and the DV is measured in both the groups again.

ii) Formal experimental designs: Offer better control and uses statistically designed experiments. They are classified as:

1. **Blocking Designs**: A "block" is a group of experimental units that share similar background characteristics. For example, in an agricultural experiment, blocks could be different plots of land with similar soil conditions or exposure to sunlight. The aim of blocking is to **reduce the variability** in results caused by these background variables, which are not of primary interest but might influence the outcomes.
2. **Factorial Designs:** These are the experiments in which the levels of each factor are combined with the levels of all other factors to conduct experiment simultaneously. For example, if there are

three factors with each factor having four levels, then there will be 64 combinations (4^3) of experiments to be conducted. In factorial designs we have:

a) **Full factorial designs:** It is the one in which we control several factors and investigate their effects at each of two or more levels. The experimental design consists of making an observation at each of all possible combinations that can be formed for the different levels of factors. Each different combination is called a "treatment combination".

b) **Fractional factorial design**: This method is used in situations where testing every combination of factors (as done in a full factorial design) would be too time-consuming or expensive. This reduces the number of trials needed to study multiple factors, especially when each factor has several levels and allows researchers to study the main effects and some interactions with fewer experiments. Ex: if n = 5 factors, with 2 levels, the number of combinations will be, $2^5 = 32$ trials. In a fractional factorial design, instead of testing all 32 combinations, you would only test a subset of those combinations—just 8 trials, you could still get useful insights into how the factors affect DV, while saving time and resources.

c) Use of orthogonal arrays: This experimental design is a systematic and efficient way to investigate multiple factors by structuring the experiment so that each factor's impact is independently and accurately measured, with reduced number of trials. An orthogonal array is a matrix where each row represents a trial (or experimental run), and each column represents a factor (an independent variable). Each combination of factor levels appears in the array in a balanced way, ensuring all levels are represented equally across trials. **Taguchi Method**: The Taguchi method uses orthogonal arrays to design robust experiments that maximize performance while minimizing variability and resource requirements.

Ex. of orthogonal arrays: Suppose a researcher wants to study three factors (A, B, and C) that each have three levels (e.g., low, medium, high). A full factorial design would require $3^3=27$ experiments. Using a suitable orthogonal array, such as the L9 array, the researcher can reduce the number of required experiments to 9 while ensuring a balanced representation of each factor and level.

3. Response surface designs: They are a set of statistical techniques used in experiments to model and optimize responses influenced by several continuous variables. By fitting a quadratic surface to the data, they allow researchers to identify the optimal conditions for the factors being studied, making them useful for fine-tuning and improving processes.

3.6 DEVELOPING A RESEARCH PLAN

(Q: Define a research plan. What is its significance?)

'Research Plan': After identifying and defining the problem researcher must arrange his ideas in order and write them in the form of an experimental plan, known as 'Research Plan'.

3.6.1 Importance of Research Plan:

(a) It helps to organize the ideas and identify the inadequacies.
(b) It suggests what must be done further and to begin with which materials have to be collected.
(c) It is a document that can be given to others for comment.

3.6.2 Research plan must contain the following items:

1. A clearly stated research objective, outlining what the researcher is expected to do.
2. An explicitly stated research problem stating what information is needed to solve the problem.
3. Each major concept (to be measured) defined in operational terms.
4. An overall description of the approach to be adopted.
5. The details of the techniques to be adopted. For instance, if interview method is to be used, detailed interview procedure, if tests are to be given, the conditions under which they are to be administered and the nature of instruments to be used.

3.7 DESIGN OF EXPERIMENTAL SET-UP

(Q: Identify and explain the phases in the process of designing an experimental setup.)

The experimental set-up depends on objective of the research.

- Process is divided into three main phases: conceptual, substantive and detail design.
- During the **conceptual phase** one should avoid thinking about physical hardware and should concentrate on functional requirements. The approaches are: (i) literature search (ii) known technical systems (iii) analogies (iv) previous experimental studies (v) design catalogues (vi) discussion with colleagues and experts etc.
- In the **substantive or embodiment phase**, the researcher should concentrate on how to convert the conceptual set-ups into physical or simulated structures. Here, the main emphasis is identification and assessment of equipment, instruments, software available, sources of supply, their method of functioning, efficacy, precision, cost etc.
- In **detailed design phase**, each of the alternative experimental set-up has to be thoroughly evaluated considering the study objective, expected outcome, feasibility, cost involved and time frame required etc. and most feasible one should be selected for adoption.

3.8 USE OF STANDARDS AND CODES

(Q: Give an overview of codes and standards used by the researcher in the design of the experimental set-up.)

- The **codes and standards** are an important source of documents to help the researcher in the design of the experimental set-up. If these standardized procedures are adopted, the credibility of the approach gets enhanced.
- Standards specify the number of specimens to be taken, the statistical tools to be used and cover instrument accuracy, test apparatus, testing methods and environment.
- Many countries have their national standards organizations, which ensure coordination and preparation of standards and codes and publish them for the benefit of industry and public. Ex: American National Standards Institute (ANSI) of USA, British Standards (BS) of UK, DIN standards of Germany, Bureau of Indian Standards (BIS) of India. The International Standards Organization (ISO) examines the standards prepared by member countries and evolves common standards.
- The professional bodies are also involved in the preparation of specification, standards and codes on various aspects. Ex: the American Society for Testing Materials (ASTM), Society of Automotive Engineers (SAE), American Concrete Institute (ACI), Institute of Electrical and Electronic Engineers (IEEE), Association of Computing Machinery (ACM) and Indian Road Congress (IRC) etc.

3.9 OVERVIEW OF MULTIVARIATE ANALYSIS

(Q: Throw light on the important techniques used in Multivariate Analysis)

3.9.1 Definition

Multivariate analysis is a powerful statistical tool that enables researchers to investigate and understand complex data sets by considering multiple variables simultaneously. Unlike univariate analysis, which examines one variable at a time, multivariate analysis explores the relationships and interactions among several variables.

3.9.2 Purpose

The main goals of multivariate analysis include:

- Understanding complex relationships among variables.
- Identifying patterns and trends in data.
- Reducing data dimensionality while retaining essential information.
- Making predictions based on multiple inputs.

3.9.3 Common Techniques

1. **Multiple Regression Analysis**: Examines the relationship between one dependent variable and multiple independent variables to predict outcomes. Ex: researchers might investigate how various factors, such as age, income, education level, and lifestyle choices, impact individuals' health outcomes (the dependent variable).
2. **Factor Analysis:** Identifies underlying factors that explain the correlations among multiple observed variables, useful for data reduction. Ex: Researchers might collect survey data on responses to questions about consumer preferences for various smartphone features such as battery life, camera quality, screen size, brand reputation, and price.
   Through factor analysis, the researchers could identify:
   Performance: includes battery life and processing speed.
   Camera Quality: encompasses camera resolution and features.
   Value: combines brand reputation and price sensitivity.
3. **Principal Component Analysis (PCA):** Transforms correlated variables into a set of uncorrelated variables (principal components) to simplify data analysis. These variables represent the directions of maximum variance in the data.
   Ex: In analyzing air quality data, researchers collect data on various air pollutants, including levels of carbon monoxide, nitrogen dioxide, sulfur dioxide, ozone etc. across multiple cities. Using PCA, the researchers can reduce the complexity of this data as follows:
   Component 1: Represents overall pollution levels (combining effects of all pollutants).
   Component 2: Distinguishes between traffic-related pollutants (e.g., carbon monoxide and nitrogen dioxide) and industrial pollutants (e.g., sulfur dioxide).
4. **Cluster Analysis:** Groups similar observations or variables into clusters, helping to identify natural groupings within data.
   Ex: A retail company collects data on various customer attributes, such as age, income, shopping frequency, and preferences for product categories (e.g., electronics, clothing, groceries).
   Using cluster analysis, the company can identify:
   Cluster1 Younger customers with high incomes frequently buy electronics.
   Cluster2: Middle-aged customers with moderate incomes look for discounts.
   Cluster3: Young adults interested in clothing and accessories, shop regularly for latest trends.
5. **MANOVA (Multivariate Analysis of Variance):** Tests differences in multiple dependent variables across groups based on independent variables.
   Ex: A study is investigating the effects of three teaching methods (traditional, online, and hybrid) on student performance across multiple subjects, such as math, science, and language arts.
   The independent variable would be the teaching method, while the dependent variables would be the test scores in math, science, and language arts. Using MANOVA, the researchers can determine if there are any statistically significant differences in student performance across the three teaching methods for the combined scores in all subjects.

3.9.4 Applications

- Market Research: Understanding consumer preferences and behavior.
- Healthcare: Analyzing patient data to identify risk factors for diseases.

- Social Sciences: Studying relationships between socioeconomic factors and outcomes.
- Finance: Portfolio management and risk assessment.

3.7 HYPOTHESIS TESTING OVERVIEW

(Q: What is a hypothesis? Describe the characteristics of a good hypothesis and explain the process of hypothesis testing.

Q: Differentiate between: i. Null hypothesis and alternative hypothesis ii. Type I and Type II errors iii. Two-tailed and One-tailed tests

Q: Explain: The level of significance and decision rule for testing of hypothesis)

Hypothesis testing is a statistical technique used to draw conclusions about an entire population based on a representative sample. It involves formulating a hypothesis, collecting sample data, and evaluating the evidence to determine whether the null hypothesis (expected result) or the alternative hypothesis (the proposed change) is more likely.

3.7.1 What is hypothesis? Research hypothesis is a predictive statement, capable of being tested by scientific methods. They state the variables of concern, the relationships among them, and the target group being studied.
Ex: 1. “Students who receive counselling will show a greater increase in creativity than students not receiving counselling”
2. “The automobile A is performing as well as automobile B.”

3.7.2 Importance of Hypothesis Testing:

- It provides a framework for making informed decisions based on data.
- It helps to avoid Type I and Type II errors.
- It enables researchers to draw conclusions about populations based on sample data.

By understanding hypothesis testing, researchers can effectively evaluate evidence, make data-driven decisions, and communicate their findings to stakeholders.

3.7.3 Characteristics of hypothesis:

(i) Hypothesis should be clear, precise and specific.
(ii) Hypothesis should be capable of being tested.
(iii) Hypothesis should state relationship between variables, if it is a relational hypothesis.
(iv) Hypothesis should be stated in most simple terms so that it is easily understandable.
(v) Hypothesis should be consistent with most known facts, reasonable, credible, supported by existing evidence and most likely prediction.
(vi) Hypothesis should be agreeable to testing within a reasonable time.
(vii) Hypothesis must explain the facts that gave rise to the need for explanation.

3.7.4 Basic concepts related to testing of hypotheses

(a) Null hypothesis and alternative hypothesis:

Null hypothesis refers to the default assumption that there is no significant difference or relationship between variables.

Ex1: If method A is to be compared with method B for superiority, then null hypothesis assumes that both methods are equally effective, and is denoted by H_0. In contrast, alternative hypothesis assumes that method A is superior or method B is inferior, and is denoted by Ha.

Ex2: If it is needed to test the hypothesis that the population mean (μ) is equal to the hypothesized mean (μ_{H0}) = 100.

Then **null hypothesis:** 'The population mean is equal to the hypothesized mean 100' and

$H_0 : \mu = \mu_{H0} = 100$.

If sample results do not support this null hypothesis, researcher should conclude that something else is true. Rejecting the null hypothesis is known as accepting alternative hypothesis. If we accept H_0, then we are rejecting H_a and if we reject H_0, then we are accepting H_a.

For $H_0 : \mu = \mu_{H0} = 100$, there are three possible alternative hypotheses:

Alternative hypothesis	*To be read as follows*
$H_a : \mu \neq \mu_{H_0}$	(The alternative hypothesis is that the population mean is not equal to 100 i.e., it may be more or less than 100)
$H_a : \mu > \mu_{H_0}$	(The alternative hypothesis is that the population mean is greater than 100)
$H_a : \mu < \mu_{H_0}$	(The alternative hypothesis is that the population mean is less than 100)

The null hypothesis and the alternative hypothesis are chosen before the sample is drawn. Important points related to hypothesis testing:

(a) The **null hypothesis** represents the default assumption that is being tested and is typically assumed to be true until evidence suggests otherwise. It is the hypothesis that researchers aim to **reject**.

(b) The **alternative hypothesis** is the hypothesis that the researcher seeks to **prove** or support. It encompasses all other possible outcomes that are considered if the null hypothesis is rejected based on the evidence.

(c) If the rejection of a certain hypothesis when it is actually true involves great risk.

(d) Null hypothesis should always be specific hypothesis i.e., it should not state approximately a certain value.

(b) **The level of significance:** Thus the significance level is the maximum value of the probability of rejecting H_0 when it is true and is determined in advance before testing the hypothesis with great care, thought and reason. The actors affecting the level of significance: (a) the magnitude of the difference between sample means; (b) the size of the samples; (c) the variability of measurements within samples; and (d) whether the hypothesis is directional or non-directional. **Ex:** 5 % level of significance means that researcher is willing to take maximum 5 per cent risk of rejecting the null hypothesis when it (H_0) it is true.

(c) **Decision rule for testing of hypothesis:** Given a hypothesis H_0 and an alternative hypothesis H_a, researcher makes decision rule according to which he accepts H_0 (i.e., reject H_a) or reject H_0 (i.e., accept H_a). **Ex:** if H_0 is that a certain lot is good (=there are very few defective items in it) against H_a, that the lot is not good (=there are too many defective items in it), then the researcher must decide the number of items to be tested (say, test 10 items in the lot) and the criterion for accepting or rejecting the hypothesis (say, if there is only 1 defective item per 10 items, then accept H_0 otherwise reject H_0.

(d) **Type I and Type II errors:** While testing of hypotheses, the researcher can make two types of errors. **Type I error (Alpha error):** rejecting H_0 when H_0 is true. It means rejection of hypothesis which should have been accepted and denoted by α (alpha).

Type II error (Beta error): Accepting H_0 when H_0 is not true. In other words, accepting the hypothesis which should have been rejected and is denoted by β (beta).
For more clarity consider the table below:

	Decision	
	Accept H_0	Reject H_0
H_0 (true)	Correct decision	Type I error (α error)
H_0 (false)	Type II error (β error)	Correct decision

The probability of Type I error is determined in advance and is known as the level of significance of testing the hypothesis. If type I error is fixed at 5 per cent, it means there are about 5 chances in 100 that we will reject H_0 when H_0 is true. We can control Type I error just by fixing it at a lower level. But with a fixed sample size, n, with this, the probability of committing Type II error increases. Hence, in the testing of hypothesis, one must make all possible effort to strike an adequate balance between Type I and Type II errors.

(e) Two-tailed and One-tailed tests (see fig. 3.1):
Two-tailed test rejects the null hypothesis if, the sample mean is significantly higher or lower than the hypothesised value of the mean of the population. Such a test is appropriate when the null hypothesis is some specified value and the alternative hypothesis is a value not equal to the specified value of the null hypothesis. Symbolically, the two tailed test is appropriate for $H_0 : \mu = \mu_{H0}$ and $H_a : \mu \neq \mu_{H0}$ $(= \mu > \mu_{H0}$ or $\mu < \mu_{H0})$

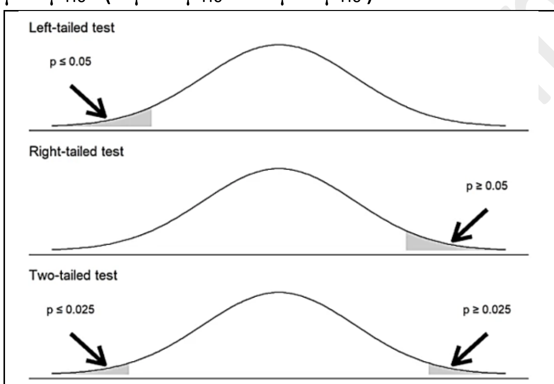

Fig. 3.1 Two-tailed and One-tailed tests

A one-tailed test would be used when we are to test, whether the population mean is either lower than or higher than some hypothesised value. Ex: if $H_0 : \mu = \mu_{H0}$ and $H_a : \mu < \mu_{H0}$, then we are interested in left-tailed test (wherein there is one rejection region only on the left tail)
In case $H_0 : \mu = \mu_{H0}$ and $H_a : \mu > \mu_{H0}$, we are then interested in what is known as right tailed test.
Note: Accepting H_0 on the basis of sample information does not constitute the proof that H_0 is true. We only mean that there is no statistical evidence to reject it, but we are certainly not saying that H_0 is true (although we behave as if H_0 is true).

3.7.5 Procedure for Hypothesis Testing:

Hypothesis testing means to tell on the basis of the data the researcher has collected whether or not the hypothesis seems to be valid. **In hypothesis testing the main question is:** whether to accept the null hypothesis or not to accept the null hypothesis.
(i) **Making a formal statement:** This step consists of making a formal statement of the null hypothesis (H_0), alternative hypothesis (H_a), one-tailed test or a two-tailed test. If H_a is of the type greater than or of

lesser than, we use a one-tailed test, but when Ha is of the type "whether greater or smaller" then we use a two-tailed test.

Ex: 1. A researcher of the Civil Engineering Department wants to test the load bearing capacity of an old bridge which must be more than 10 tons; The hypotheses as under:

Null hypothesis H_0: $\mu = 10$ tons

Alternative Hypothesis Ha: $\mu > 10$ tons

2. The average score in an aptitude test administered at the national level is 80. To evaluate a state's education system, the average score of 100 of the state's students selected on random basis was 75. The state wants to know if there is a significant difference between the local scores and the national scores. Then the hypotheses:

Null hypothesis H_0: $\mu = 80$

Alternative Hypothesis H_a: $\mu \neq 80$

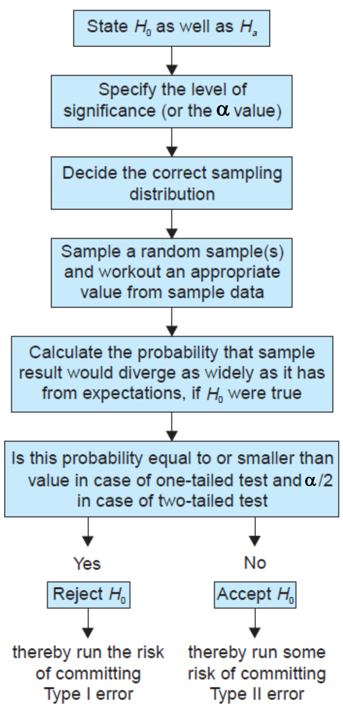

Fig. 3.2 Flow Diagram for Hypothesis Testing

(ii) Selecting a significance level: The hypotheses are tested on a pre-determined level of significance either 5% level or 1% level and should be specified based on the purpose and nature of enquiry.

(iii) Deciding the distribution to use: It includes determining the appropriate sampling distribution based on the rules for selection.

(iv) Selecting a random sample and computing an appropriate value: This refers to select a random sample and compute an appropriate value from the sample data concerning the test statistic utilizing the relevant distribution.

(v) Calculation of the probability: Then calculate the probability that the sample result would diverge as widely as it has from expectations, if the null hypothesis were true.

(vi) Comparing the probability: It refers to comparing the probability thus calculated with the specified value for α, the significance level. If the

calculated probability =< α value in case of one-tailed test

and =< α /2 in case of two-tailed test, then reject the H_0 (i.e., accept the H_a).

but if the calculated probability is greater, then accept the null hypothesis.

The hypothesis testing is depicted in Fig. 3.2

3.8 MEASURES OF ASSOCIATION

(Q: Explain the concept of measures of association, highlighting its importance and types.)

3.8.1 Definition:

Measures of association are statistical tools used to quantify the strength and direction of relationships between two or more variables. They help researchers understand how variables interact and can indicate whether changes in one variable are associated with changes in another. By understanding these associations, researchers can draw meaningful conclusions and inform practical applications across various disciplines.

3.8.2 Importance:

Understanding measures of association allows researchers to:

- Identify patterns and relationships in data.
- Inform decision-making and policy development.
- Design interventions and strategies based on identified associations.

3.8.3 Types of Measures:

1. **Correlation Coefficients**:
    - **Pearson's Correlation Coefficient (r)**: Measures the linear relationship between two continuous variables, ranging from -1 to +1. A value close to +1 indicates a strong positive relationship, while a value close to -1 indicates a strong negative relationship.
    - **Spearman's Rank Correlation Coefficient**: Used for ordinal data or non-normally distributed continuous data, measuring the strength and direction of a monotonic relationship.
2. **Contingency Coefficients**:
    - **Chi-Square Test of Independence**: Assesses the association between two categorical variables in a contingency table(A contingency table is a matrix that displays the frequency distribution of the variables; how often each combination of categories occurs). It indicates whether the observed frequencies differ from expected frequencies under the assumption of independence.
    - **Cramér's V**: A measure derived from the Chi-Square statistic, providing a value between 0 and 1 to indicate the strength of association between categorical variables.

3.8.4 Applications

Measures of association are widely used in various fields, including:

- **Market Research**: To evaluate the association between consumer preferences and purchasing behavior.
- **Social Sciences**: To analyze relationships between demographic variables and social behaviors.
- **Medicine**: To assess the relationship between risk factors and health outcomes.

3.9 PRESENTING INSIGHTS AND FINDINGS

(Q: Discuss the importance of both written and oral presentations in research methodology.
Q:

3.9.1 Written Reports

- A quality presentation of research findings can have an excessive effect on a reader's perceptions of a study's quality. Hence a researcher to make a special effort to communicate skillfully and clearly.
- Research reports contain findings, analysis, interpretations, conclusions, and sometimes recommendations.
- The writer of research reports should be guided by four questions:
    - ✓ What is the purpose of this report?
    - ✓ Who will read it?
    - ✓ What are the circumstances and limitations under which it is written?
    - ✓ How will the report be used?

Reports should be clearly organized, physically inviting, and easy to read. Most statistics should be placed in tables, charts, or graphs based on the specific data and presentation purpose.

3.9.2 Types of Reports

(Q: Explain the different types of report)

They may be **short**, informal format typical of memoranda and letters, or they may be **longer** and more complex.

i. Short Reports: Short reports are appropriate when the problem is well defined, is of limited scope, and has a simple and straightforward methodology. The purpose is to distribute information quickly in an easy-to-use format. Informational, progress, inexpensive research projects and interim reports are of this kind. Ex: a report of cost-of-living changes for upcoming labor negotiations. Short reports are about five pages. It should have a brief statement about the authorization for the study, the problem examined, and its important details, the conclusions and recommendations and the findings that support them. The report should be direct, make ample use of graphics to show trends.

ii. Long Reports: Long reports are of two types, the technical (or base) report and the management report

A technical report is written for an audience of researchers. Completeness is a goal. It should include full documentation and detail. This includes sources of data, research procedures, sampling design, data gathering instruments, index construction, and data analysis methods.

Management report is written for the nontechnical client and the manager and should make liberal use of visual displays such as pictures and graphs. The reader has little time to absorb details and needs an exposure to the most critical findings; thus the report's sections are in an inverted order. After the prefatory and introductory sections, the conclusions with recommendations are presented to grasp quickly. Individual findings are presented next, supporting the conclusions already made. The appendices present any required methodological details.

3.9.3 Research Report Components

(Q: Discuss the components of a research report that assist the reader in understanding the structure and context of the report)

1. Prefatory Items: These assist the reader in using the research report.

2. Letter of Transmittal: It should refer to the authorization for the project, instructions, limitations, state the purpose and the scope of the study.

3. Title Page: It should include four items: the title of the report, the date, and for whom and by whom it was prepared. The title should be brief but include the following three elements: (1) the variables included in the study, (2) the type of relationship among the variables, and (3) the population to which the results may be applied. **Ex:**

Descriptive study: The Five-Year Demand Outlook for Consumer Packaged Goods in India.

Correlation study: The Relationship between Relative National Inflation Rates and Household Purchases of Brand X in International Markets.

Causal study: The Effect of Various Motivation Methods on Retail Sales Associates' Attitudes and Performance.

4. Authorization Letter: When the report is sent to a public organization, this letter is included which not only shows who sponsored the research but also defines the original request.

5. Executive Summary: It is of 2 pages and serve two purposes. It may be a report covering all the aspects in the body of the report in abbreviated form, or it may be a concise summary of the major findings and conclusions, including recommendations.

6. Table of Contents: A report of several sections should have a table of contents, list of tables, charts, or other exhibits.

7. Introduction: It describes the parts of the project: the problem statement, research objectives, and background material.

8. Problem Statement: It contains the need for the research project and represented by a management question.

9. Research Objectives: These objectives are research questions and associated investigative questions. They address the purpose of the project. In correlational or causal studies, the hypothesis statements are included.

10. Background: Background material may be of two types. It may be the preliminary results of exploration from an experience survey, focus group, or another source. It could be secondary data from the literature review. The background includes definitions, qualifications, and assumptions. It gives the reader the information needed to understand the remainder of the research report.

11. Methodology: For technical report, the methodology is an important section.

11.1 Sampling Design: This includes the target population being studied, uniqueness of the chosen parameters and the sampling methods used.

11.2 Research Design: The coverage of the design must be adapted to the purpose. In an experimental study, the materials, tests, equipment, control conditions, and other devices should be described. Also cover the rationale for using one design instead of competing alternatives.

12. Data Collection: This part describes the specifics of gathering the data. How many people were involved? When were the data collected? etc. The use of standardized procedures and protocols, the administration of tests, manipulation of the variables etc. used in the experiment should be explained.

13. Data Analysis: This section summarizes the methods used to analyze the data and describes data handling, preliminary analysis, statistical tests, computer programs, and other technical information. The rationale for the

choice of analysis, approaches, assumptions and appropriateness of use should be presented.

14. Limitations: Acknowledging limitations helps readers assess the study's validity and understand the context of the findings.

15. Findings: The objective is to explain the results9facts) rather than draw interpretations or conclusions. It is useful to present findings in numbered paragraphs with the quantitative data supporting the findings presented in a small table or chart on the same page.

16. Conclusions:

16.1 Summary and Conclusions: The summary is a brief statement of the essential findings. Findings state facts; conclusions represent inferences drawn from the findings.

16.2 Recommendations: In this section, researchers offer ideas for corrective actions and suggestions for further research initiatives that broaden or test the understandings of a subject area.

17. Appendices: They include complex tables, statistical tests, supporting documents, copies of forms and questionnaires, detailed descriptions of the methodology, instructions to field workers, and other evidence important for later support.

18. Bibliography: It documents the sources used by the writer and requires to be adhered to specific citation styles and formats, which can vary based on the instructor, program, or institution. Recommended style manuals include the APA Publication Manual, Kate L. Turabian's manual, and the MLA Handbook.

3.10 ORAL PRESENTATION IN RESEARCH METHODOLOGY

(Q: Explain the significance of oral presentations in research methodology. Discuss the key elements involved in preparing an effective oral presentation, including the structure, audience engagement, and delivery techniques that contribute to a successful presentation.)

- It involves the effective communication of research findings to an audience, in a structured and concise format. It is a vital skill for researchers, as it allows them to share their work, engage with other scholars, and receive feedback.
- A successful oral presentation not only communicates the research effectively but also builds the presenter's credibility and helps establish professional networks in the academic community.
- The presentation includes an introduction to the research problem, review related literature and research gap, objectives, methodology, results, and conclusions.
- Preparing for an oral presentation requires organizing key points logically, selecting visuals such as slides to enhance understanding, and practicing clear and confident delivery.
- Audience engagement is also crucial, often achieved through interactive elements, such as Q&A sessions. Good time management during the presentation is essential to ensure that all main points are covered within the allotted time.
- Attention to the tone, body language voice modulation and avoiding fillers (“ah,” “um,” “you know,” “like,”) can make the presentation more effective. Anticipating potential questions from the audience allows the presenter to prepare informed responses.

References:

1. Ganesan R, Research Methodology for Engineers, MJP Publishers, Chennai. 2011
2. Cooper, Donald R. and Schindler, Pamela S., Business Research Methods, Tata McGraw-hill Publishing Company Limited, New Delhi, India. 2012
3. Internet resources

(Note: Questions are given for illustrative purpose only)